REPURPOSING CLINIC-TESTED DRUGS TO TREAT RARE NEUROGENETIC DISEASES BY TRANSCRIPTIONAL MODULATION

Jeremiah Hadwen

Thesis submitted to the Faculty of Graduate and Postdoctoral Studies

in partial fulfillment of the requirements for the

Ph.D. degree in Cellular and Molecular Medicine

Department of Cellular and Molecular Medicine

Faculty of Medicine

University of Ottawa

April 27, 2018

AUTHORIZATIONS

- Table 1 is an adaptation of Table 1 in "Mining the transcriptome for rare disease therapies: a comparison of the efficiencies of two data mining approaches and a targeted cell-based drug screen" by Mears AJ et al., 2017. The table has been modified with some information removed and some information added and is included in this dissertation according to permission from the Creative Commons Attribution 4.0 International License which can be found at: https://creativecommons.org/licenses/by/4.0/legalcode

CONTRIBUTIONS

- Jeremiah Hadwen[1,2] was involved with experiment design, experiment implementation and writing and editing the present dissertation.
- Dr. Alex MacKenzie[1,2] was involved with experiment design and editing the dissertation.
- Dr. Sarah Schock[1,2] helped with cerebrocortical dissections, DRG culturing, experiment design and dissertation editing.
- Dr. Alan Mears[2] helped with experiment design and qRT-PCR.
- Dr. Robert Yang[3] was involved in RNAseq differential expression analysis.
- Liying Zhang[3] helped to prepare Figure 2.5.
- Dr. Roxanne Lariviere[4] performed the western blots for Figure 3.7B.
- Dr. Jason Vanstone[2] helped with cerebrocortical dissections.
- Kate Daniels[2] helped with cerebrocortical dissections, cDNA library preparation, RNA extractions.
- Philippe Charron[1,2] helped to create the Neuron Screen database (Figure 2.10)

Contributors' Affiliations (at the time of contributing to the work)

1. Department of Cellular and Molecular Medicine, University of Ottawa, Ottawa, ON, Canada.
2. Children's Hospital of Easter Ontario (CHEO) Research Institute, University of Ottawa, Ottawa, ON, Canada.
3. Computational Sciences Center of Emphasis, Pfizer Worldwide Research and Development, Cambridge, MA, USA.
4. Montreal Neurological Institute and Hospital, McGill University, Montreal, QC, Canada.

ACKNOWLEDGEMENTS

First and foremost, I want to thank my amazing wife Anna for all her support and sacrifice that has enabled me to continue working on this PhD over the past 3 years. Countless times she has encouraged me and has advised me and none of this would be possible without her help. She is also a fantastic mother to our two beautiful daughters Elaina and Esther.

I also want to thank my parents, Matthew and Simone, without whose love and guidance throughout my life there is no way I would be where I am today.

I am deeply indebted to my supervisor Dr. Alex Mackenzie who has been a tremendous role model for me as a clinician scientist and has given me the support and the inspiration to work to help children with rare diseases.

I would also like to specifically thank Drs. Sarah Schock and Alan Mears for their constant support and always having their doors open for me to come ask a question or get their advice.

I am also grateful to my TAC members (Drs. Bill Staines, Ilya Ioschikes, and David Dyment) for their valuable input and guidance.

I also want to thank all my colleagues at the CHEO-RI, especially Izabella Pena, Kevin Mongeon, Nafisa Tasnim, Kristin Kernohan for all their help and support.

I want to acknowledge all the CHEO-RI administrative and support staff especially Sandy Boehmer and Lynn Kyte.

I also want to thank our collaborators at the Pfizer Computational Sciences Center of Emphasis especially Simon Xi and Robert Yang for their tremendous support in RNAseq analysis.

Finally, I would like to thank the many family members and my church family who have been a tremendous support to my wife and I during the last three years and whose help caring for our two daughters and providing meals enabled me to continue working on my thesis.

ABSTRACT

Rare diseases caused by single-gene mutations affect almost one million Canadians. According to the Online Mendelian Inheritance in Man database, ~4,500 rare monogenic diseases have a known cause; but less than 5% of patients have access to disease-modifying drugs. The dearth of accessible drugs for patients suffering from rare genetic diseases is partly due to the astronomical costs of traditional drug development which, when combined with the small target population, make rare disease therapeutics unattractive ventures for the pharmaceutical establishment. The paucity of cost-effective treatments for rare diseases has resulted in the promotion of clinic-ready drug repurposing as a tenable strategy for rare disease therapeutics. To identify repurposed candidates for rare neurogenetic diseases, I conducted a transcriptome-wide drug screen in mouse primary cerebrocortical cultures. RNA sequencing was used to develop a database of transcriptome-wide differential expression for 218 clinic-tested drugs. The "Neuron Screen" database was queried to identify drugs that upregulate ~60 rare neurogenetic disease genes (type I hits). Gene set enrichment pathway analysis by Ingenuity Pathway Analysis (IPA) was used to identify network associated drug-gene interactions (type II hits). Both types of drug-gene hits were further assessed *in vitro* and *in vivo* by qRT-PCR and western blot analysis. This analysis showed that the IPA-based network-associated approach reduces the false positive rate when identifying differentially expressed genes in transcriptome-wide data-sets. The analysis also identified two drug-gene interactions with genes that cause rare neurogenetic disease, thyroid hormone-*Pmp22* and dexamethasone-*Mfsd2a*, that merit further investigation. This work proves the utility of the Neuron Screen database to connect rare disease genes with transcript-modulating drugs and provides a

starting point to understand the transcriptional effects of pharmacologic agents on the mammalian brain.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Chapter 2

Chapter 3

Chapter 4

LIST OF ABBREVIATIONS

OMIM	Online Mendelian Inheritance in Man
3H	Triple hybrid
ADA-SCID	Severe combined immunodeficiency disease caused by adenosine deaminase deficiency
ALDH1L1	Aldehyde dehydrogenase 1 family, member L1
ARSACS	Autosomal recessive spastic ataxia of Charlevoix-Saguenay
ASO	Antisense oligonucleotide
BBB	Blood-brain-barrier
C4R	Canadian Care for Rare consortium
CMap	Connectivity Map
C-mapping	Connectivity mapping
CMT1A	Charcot-Marie-Tooth 1A
CNS	Central nervous system
CRE	Causal Reasoning Engine
CSCoE	Computational Sciences Center of Emphasis
CSF	Cerebral spinal fluid
DEX	Dexamethasone
DMD	Duchenne muscular dystrophy

E16.5	Embryonic day 16.5
ER	Endoplasmic reticulum
FDA	Federal Drug Association
FDR	False discovery rate
GABA	Gamma-aminobutyric acid
GFAP	glial fibrillary acidic protein
GOF	Gain-of-function
HNPP	Hereditary neuropathy with liability to pressure palsies
HTS	High-throughput drug screening
HU	Hydroxyurea
iNs	iPSC derived neurons
IPA	Ingenuity Pathway Analysis
iPSCs	Induced pluripotent stem cells
LCLs	Lymphoblastoid cells
LDH	Lactate dehydrogenase
LPLD	Familial lipoprotein lipase deficiency
MAP2	Microtubule associated protein 2
MEA	Multiple electrode arrays
mRNA	Messenger RNA
NeuN	Neuronal nuclei

NGS	Next generation sequencing
p-adj	Adjusted p-value
PMP22	Peripheral myelin protein 22
qRT-PCR	Quantitative reverse transcriptase polymerase chain reaction
RNAseq	RNA sequencing
SCD	Sickle cell disease
SMA	Spinal muscular atrophy
SMN2	Survival of motor neuron 2
T3	Liothyronine
T4	Levothyroxine
UPR	Unfolded protein response
vGAT	Vesicular GABA transporter
vGLUT	Vesicular glutamate transporter
VLCFA	Very-long chain fatty acids
X-ALD	X-linked adrenoleukodystrophy

INTRODUCTION

Transcriptional drug repurposing to treat rare neurogenetic diseases

Rare monogenic diseases

Rare diseases caused by single-gene mutations are often chronic, debilitating, and life limiting and thus major contributors to human disability and illness. The majority of rare monogenic diseases have a childhood onset and patients suffering from them require a disproportionately larger number of longer hospitalizations than do other patients (Yoon et al. 1997; Dye et al. 2011). Although by definition affecting fewer than 1 in 2,000 people (Europe) or fewer than 200,000 individuals in the United States (Beaulieu et al. 2012), rare genetic diseases are collectively common and pose a significant personal and public financial burden to our society (Carter 1977; Baird et al. 1988). Queries of the human gene database Online Mendelian Inheritance in Man (OMIM) (McKusick 2007) and the rare disease portal Orphanet (Ayme et al. 1998) identify between 6,000 and 7,000 rare monogenetic diseases (Boycott et al. 2013); more than half have a known aetiology. The number of genes causally linked to rare diseases is increasing rapidly, largely a result of the application of next generation sequencing (NGS) technologies. The identification of novel rare disease genes has also been accelerated by recently established national and international collaborations between clinicians and scientists that aim to provide genetic diagnoses for all rare genetic diseases (Philippakis et al. 2015; Boycott et al. 2017). One such initiative is the pan-Canadian Care for Rare (C4R) consortium, a collaboration between researchers, scientists, informaticians, clinicians, and patients that has published over 30 manuscripts on new characterizations of rare genetic diseases within the past three years (T. Hartley et al. 2017). Although the identification, and consequently diagnosis, of rare genetic

diseases is progressing rapidly, the development of effective treatments for patients diagnosed with rare diseases is increasing at a far slower pace. Boycott *et al.* (2013) estimate that at the current drug discovery rate, over the next 20 years only 75 new disease-modifying drugs will be approved for rare monogenic diseases. Presently, most treatments for rare genetic diseases moderate symptoms and improve quality of life rather than being curative or even significantly disease altering; surgical treatments are often used to improve the quality of life for neuromuscular disease patients (Halawi et al. 2015), enzyme replacement strategies are effective for some lysosomal storage diseases (Desnick & Schuchman 2002; Ratko et al. 2013), and diet restriction/supplementation is used to treat inborn errors of metabolism such as phenylketonuria (Camp et al. 2012; Yuzyuk et al. 2016).

Gene-level treatment strategies

The most intuitive strategy for curing a genetic disease is to correct the disease-causing mutation at the nucleic acid level (DNA or RNA), thus re-establishing physiological gene function. At present, the most promising avenues of direct gene correction for rare genetic diseases are virus-mediated gene delivery (gene-therapy) and antisense oligonucleotide (ASO) therapy. Gene-therapy is a technique that was first developed in the early 1970s and consists of viral-mediated delivery of a gene or partial gene to replace the defective target gene (Giacca & Zacchigna 2012). Numerous clinical trials have shown promising results, yet almost 30 years after the first trial of gene therapy in a human (Wirth et al. 2013), the one instance of approval in the western

world is a cautionary tale. The gene therapy Glybera was approved (in Europe) almost four years ago for the treatment of the rare monogenic disease familial lipoprotein lipase deficiency (LPLD) (Watanabe et al. 2015). However, the cost of Glybera was so prohibitive ($1.2 million dollars) and the prevalence of LPLD so low (1:100,000), that only one patient was treated (Melchiorri et al. 2013). The treatment was a success but the maker of Glybera (Uniqure) has not sought to extend drug approval for financial reasons. One of the most promising ongoing clinical trials of gene therapy is for the rare monogenic severe combined immunodeficiency disease caused by adenosine deaminase deficiency (ADA-SCID). Over the past few decades, several clinical trials have shown that ADA-SCID patients can be cured by gene-therapy (Ferrua et al. 2010). The most recent clinical trial showed a ~90% effect rate and no adverse events 4 years post-treatment (Shaw et al. 2017). However haematological disease like ADA-SCID have the advantage of utilizing *ex vivo* gene delivery to a patient's lymphocytes followed by autologous transplantation (Tani 2016). For most rare genetic diseases, gene therapy would require systemic administration or multiple administrations to target tissues, posing significant difficulty for treating multi-system diseases and diseases involving the central nervous system (Maguire et al. 2014).

ASO therapy consists of using short DNA molecules that for example can bind to mutated genomic DNA correcting mis-splicing, thus increasing functional transcript formation (Evers et al. 2015). The functional transcript can lead to a functional protein and thus correct the disease-causing protein dosage problem (responsible for the pathophysiology of most monogenic disease). The ASO strategy recently broke into the

clinical realm with the 2016 Federal Drug Association (FDA) approval of ASO drugs for the treatment of two progressive and fatal rare genetic neuromuscular diseases (Mullard 2017); Eteplirsen was approved to treat Duchenne muscular dystrophy (DMD) (Young & Pyle 2016) and Nusinersen to treat spinal muscular atrophy (SMA) (Corey 2017; Hoy 2017). Although Nusinersen, which has also received Health Canada approval, can be used for most SMA patients, Eteplirsen only applies to 13% of DMD patients (Dowling 2016). Also, many insurance providers (public and private) refuse to pay the high price-tag for these treatments ($300,000 USD/year for Eteplirsen; $750,000 USD/year for Nusinersen), impeding access for many patients. In sum, very few genetic treatments are approved for rare diseases, and even fewer are available to patients from an economic point of view.

Drug repurposing

Repurposing clinically-tested small molecules and biologics has received increasing recognition during the past decade (Keiser et al. 2009; Campillos et al. 2008). The interest in finding new clinical applications for drugs is fueled by the unmet medical need of most rare diseases (as well as neglected infectious diseases), the increasingly inefficient new therapeutic drug approval process and the availability of a growing library of clinic-tested drugs. Over the past 50 years, the ratio of cost of development to number of new drug approvals has consistently increased (Scannell et al. 2012), a trend that even the dramatic breakthroughs in DNA sequencing and protein structure resolution technologies have not moderated. According to recent studies, each approved

new drug costs between USD$1.9 billion and $3.2 billion and takes 12-15 years to develop (DiMasi et al. 2016; Nosengo 2016). Most of the cost associated to new drug development is not related to the drug that is approved, but to the research and development costs related to drugs that are not approved. For every drug approved for human use, nine drugs fail at some stage in the development pipeline. Attrition occurs most frequently after safety testing (Phase I), during Phase II and Phase III with 66% and 30% attrition rates, respectively (Paul et al. 2010). Consequently, together with clinic-approved drugs, there exist ~10,000 drugs that have been trialed in humans, and ~5,700 of these compounds can be readily purchased for experimental purposes (Corsello et al. 2017). Many of these drugs that have passed safety testing in humans, but ultimately failed to elicit the intended clinical effect, are prime targets for drug repurposing. Compared to new drug development, the cost to repurpose a clinic-tested drug is roughly one tenth, and takes about half the time (Nosengo 2016). The reduced cost is a big reason for the interest in drug repurposing, especially for rare diseases that have not received the type of funding accorded to more common human ailments (diabetes, prevalent cancers, etc.). This is largely because the rarer a disease, the fewer number of eventual users of a new therapy, and thus a lower amount of revenue and return on investment for the drug maker (Boycott et al. 2013). Economic incentives from governments and rare disease patient groups have prompted an increased interest in developing rare disease therapeutics (Dharssi et al. 2017), but most of the drugs that have been developed and approved for rare diseases are prohibitively expensive. Thus, because repurposing safety-tested and clinic-approved drugs significantly reduces the

cost and time of drug development, it is an attractive way to identify treatments for orphan diseases – i.e. diseases that have been "left behind" by the drug-development industry, as well as to develop lower cost treatments for diseases that have a prohibitively expensive treatment.

Classic examples of repurposed drugs include sildenafil (erectile dysfunction – originally for angina), thalidomide (multiple myeloma – originally for hyperemesis gravidarum), and avastin (metastatic breast cancer – originally for colon and lung cancer) (Ashburn & Thor 2004; Chong & Sullivan Jr 2007). The repurposing of such drugs has benefited millions of patients and has led to billions of dollars in sales for the pharmaceutical industry (Swamidass 2011; Dudley et al. 2011). These early instances of drug repurposing occurred serendipitously because of unanticipated side-effects. As drug repurposing research has gained momentum, more structured and efficient approaches have been developed to identify repurposing candidates supported by an expanding list of public, private, and not-for-profit organizations. These efforts are bearing fruit as evidenced by the fact that ~30% of drug approvals in recent years have been repurposed compounds (Jin & Wong 2014). The majority of interconnected and overlapping approaches to drug repurposing (reviewed by Jin & Wong 2014) rely on the traditional drug-development principle of targeting proteins/pathways (enzymes, receptors, ion channels, transport proteins etc.) (Imming et al. 2006). However, it is estimated that a full ~80% of proteins are not candidates for modulation with a small molecule or biologic drug, and thus many conditions may not be treatable if only proteins are considered as druggable targets (Plenge 2016). Furthermore, the protein

structure of many rare disease genes has not been resolved or studied sufficiently to be included in a typical high-throughput drug assay.

Transcriptional drug repurposing

A relatively recent strategy for drug repurposing relies on the theory that some pharmaceuticals have effects at the transcript level that are therapeutically beneficial to some diseases and may be repurposed by studying their effects at the mRNA level. Transcriptional drug repurposing is a strategy of drug development that has become increasing popular and feasible due to the development of transcriptome-wide interrogation tools and bioinformatic approaches to analyze and compare transcriptomic data. The most widely used approach to identify transcriptional drug repurposing candidates involves the use of the Connectivity Map (CMap) – i.e. a database of transcriptome-wide drug effects obtained by treating 3 cancer cell lines with ~1300 drugs. Lamb *et al* (2006) proposed the connectivity mapping (c-mapping) method of using the CMap data to identifying candidates for drug repurposing (Lamb 2007). Based on the assumption that each disease state (e.g. monogenic disease, cancerous disease) causes certain transcriptomic changes (i.e. disease transcriptomic signature), the goal of c-mapping is to identify a drug that elicits a signature that is the inverse or ortholog of the disease signature. The underlying (and admittedly somewhat speculative) hypothesis of c-mapping is that a drug that creates a signature which is the inverse of or orthogonal to a disease signature can be used to moderate the faulty gene signature of the disease and thus bring therapeutic benefit. The c-mapping method has been used to

identify potential drug treatments against diseases with signatures comprised of as few as 23 (Yuen et al. 2012) and as many as 320 genes (Delahaye-Duriez et al. 2016). The c-mapping method can also be used as a novel means of classifying drugs by signature, possibly identifying alternative drugs for a disease that already has a functional drug (Lamb et al. 2006). In the more than 10 years since publication of the CMap, hundreds of labs have used the CMap approach or variations of the CMap approach to identify candidate therapies for a variety of diseases, mostly for various types of cancer (Sirota et al. 2011; McArt & Zhang 2011; Shigemizu et al. 2012). Connectivity mapping has also led to drug candidates for some non-cancerous diseases including the two anthelmintic compounds (albendazole and oxamniquine) identified for Gaucher disease (rare monogenic disease) (Yuen et al. 2012) and the nutraceutical ursolic acid for the treatment of skeletal muscle atrophy (Kunkel et al. 2011). Several clinical trials have been initiated for drugs identified by the connectivity mapping technique, but to date none have been approved for a new disease. In addition to the modest clinical success, the c-mapping approach has limitations, especially for the treatment of rare genetic diseases. The use of the CMap for drug candidate selection depends to a degree upon the availability of transcriptomic data for a given disease to enable the identification of a disease-related signature. Disease signatures are known for many common diseases, but for most rare diseases a druggable disease signature is not established. Therefore, c-mapping remains mostly applicable to common diseases.

Another method for transcriptional drug repurposing is single transcript-modulating or pharmacological gene therapy – a single-gene approach that is

particularly suitable to treat monogenic diseases. The transcript-modulating approach rests on the theory that rare monogenic disease can be viewed as a disorder of protein level/activity that might be fixed by pharmacological transcript modulation. Given a central dogma of biology is that DNA codes for mRNA that is translated into protein, it is anticipated that transcripts for most genes, when present in greater abundance, will lead to the translation and production of greater amounts of protein. Accordingly, transcriptional modulation of a monogenic disease-related gene to restore physiological levels of mRNA and protein can thus be curative (Beaulieu et al. 2012; Mears et al. 2017), a principle previously demonstrated by our lab in SMA disease-model mice (Farooq et al. 2013; Farooq et al. 2011). The transcript-modulating approach does not require a disease-related transcriptomic signature or complex bioinformatic work but depends on the correct categorization of a gene mutation. If a disease is caused by haploinsufficiency, expression of the wild type allele can be up-regulated to offset the gene dosage deficiency. If a mutation is hypomorphic, upregulation of the gene can lead to greater levels of mutated protein with increased residual activity providing the protein's sub-cellular localization has not been impacted. If a mutation is a duplication or leads to a toxic gain-of-function (GOF), downregulation of the gene can reduce the levels of the excess/toxic transcript or protein and thus moderate the mutation's deleterious consequences. The *PMP22* gene that codes for peripheral myelin protein 22 (PMP22) is an example of how both up and down regulation of a given gene can result in disease. Heterozygous loss of function mutation in *PMP22* causes a lack of PMP22 protein resulting in the rare hereditary neuropathy with liability to pressure palsies

(HNPP) (Chance et al. 1993). Conversely, duplication of *PMP22* results in an over-abundance of PMP22 protein (GOF) leading to the rare genetic disease Charcot-Marie-Tooth 1A (CMT1A) (Lupski et al. 1991; Raeymaekers et al. 1991; Wise et al. 1993). Since both HNPP and CMT1A have related pathologies, (peripheral nerve dysmyelination with reduction in nerve conduction velocity) (van Paassen et al. 2014), it appears that the PMP22 dosage must fall within a narrow window to enable proper formation and function of myelin in the peripheral nervous system (PNS). While patients with HNPP may benefit from upregulation of *PMP22*, patients with CMT1A may benefit from downregulation of *PMP22*. Several repurposing efforts aimed at reducing *PMP22* expression have been made for treating CMT1A. The most successful to date is the drug PXT3003, a low-dose formulation of the three clinically approved drugs (RS)-baclofen, naltrexone hydrochloride and D-sorbitol (Chumakov et al. 2014). PXT3003 was approved for phase III efficacy testing at the end of 2016 after a successful phase II clinical trial (Attarian et al. 2014; Attarian et al. 2016).

For some disease genes, a paralogous gene exists in the genome with similar sequence and function. These genes are sometimes referred to as "rescuing paralogs" because upregulation at sufficient levels in the appropriate tissues may compensate for the loss of function caused by mutation of the disease gene; SMN2 for SMA and HbF for SCD being two well established examples of this (Beaulieu et al. 2012). Another example of a disease that may be treated by upregulation of a rescuing gene paralog is the rare peroxisomal disease X-linked adrenoleukodystrophy (X-ALD). Mutations in the *ABCD1* gene leads to accumulation of very-long chain fatty acids (VLCFA) and to a range of

neurological symptoms. The *ABCD2* gene was identified as a rescuing paralog of *ABCD1* because the protein products of both genes are molecularly similar and over-expression of *ABCD2* can rescue the X-ALD phenotype *in vitro* and *in vivo* (Kemp et al. 1998). Upregulation of *ABCD2* by thyromimetic compounds has been postulated as an effective treatment for X-ALD in humans (M. D. Hartley et al. 2017).

Along with characterizing new rare genetic diseases, the C4R consortium aims to identify repurposing candidates for rare monogenic diseases. To this end, C4R prioritized 76 rare monogenic diseases of known aetiology as potential targets for drug repurposing by transcript-targeting therapy (Mears et al. 2017). These diseases were identified because they are clinically tractable and have the potential to be treated by transcriptional modulation of an individual gene (i.e. each disease is hypomorphic, haploinsufficient, toxic GOF or has a known paralogous gene). Clinical tractability, the relative feasibility of treating a known disease, was determined by C4R clinical geneticists and was based primarily on age of disease onset, rate of disease progression, and likelihood that the phenotype can be reversed (Mears et al. 2017).

Small molecule therapies that target an individual gene's expression is a relatively novel approach to treating rare diseases and the method used to identify candidates for gene expression modulation relies to a large degree on knowledge of the gene in question. If mechanisms for transcriptional regulation of a gene of interest have been characterized, then a non-screening rational drug discovery approach can be used and molecules that modulate pathways related to gene expression can be identified and

tested as potential therapies. For example, knowledge that the STAT-5 transcription factor enhances promoter activity of the SMA rescuing paralog (survival of motor neuron 2 (SMN2)) (Ting et al. 2007), led to the identification of the STAT-5 inducing hormone prolactin as a potential treatment for SMA (Farooq et al. 2011). Another example of knowledge-driven identification of a transcriptional repurposing candidate to treat rare monogenic disease is the pre-clinical validation of thyromimetic compounds for X-ALD (Genin et al. 2009). Characterization of the *ABCD2* promoter region enabled the identification of thyroid hormone response elements that were shown to mediate upregulation of *ABCD2*.

Unfortunately, most rare and newly characterized rare diseases do not benefit from the same level of knowledge as comparatively common rare diseases (e.g. SMA, X-ALD). For many rare diseases, only a handful of patients have been identified, and little is known other than the identity of the gene mutation and resulting impact on functional protein. Previous work done by our lab outlined how transcriptional data from the CMap could be used to identify individual drug-gene interactions that validate *in vitro* and *in vivo* as potential treatments for myotonic dystrophy type I (Witherspoon et al. 2015). The validation process was based on the hypothesis that if a drug had a certain transcriptional effect in cancer cells (CMap), it would have a similar transcriptional effect on non-cancerous somatic cells of various kinds. This transcriptional effect could translate to a rescue of gene function and the drug would then be repurposed to treat the associated disease. This kind of individual drug-gene interaction approach to repositioning drugs differs from the well-recognized orthogonal drug signature

approach developed by Lamb et al. (2006). Our lab recently compared three different approaches that could be used to identify transcript-targeting drugs for any rare monogenic disease regardless of the level of knowledge about the disease. Mears et al. (2017) studied the *in vitro* transcriptional validation rate of drug-gene hits from the 76 C4R prioritized disease genes. Three candidate identification methods were compared: 1) drug-gene hits identified by mining the transcriptomic data that forms the CMap; 2) drug-gene hits identified by the Causal Reasoning Engine (CRE) – i.e. Pfizer-developed literature mining tool (Chindelevitch et al. 2012); 3) drug-gene hits identified through a targeted screen of ~300 FDA approved drugs in primary human fibroblasts. The rate of hits identified from the CMap that validated (i.e. demonstrated reproducible modulation of mRNA) in primary human fibroblasts was ~5%, while the rate of validation of hits from the fibroblast screen was ~9% (Mears et al. 2017).

Differential expression analysis

Differential expression analysis is used to compare the transcript abundance for a given gene between tissue types, between control and disease tissue, or between treated and untreated conditions. The gold-standard of differential expression is to study individual genes by designing gene-unique primers and using quantitative reverse transcriptase polymerase chain reaction (qRT-PCR) or TaqMan PCR to amplify the gene and compare its abundance between samples. Because PCR technology is highly specific and necessitates a separate reaction for each gene of interest, it is not feasible for transcriptome-wide analyses. The development of DNA microarray technology

revolutionized the field of genomics in the early 1990s, and is a mainstay tool for drug discovery, enabling simultaneous differential expression analysis of thousands of genes on a single chip (Marshall 1999). The development of next generation sequencing (NGS) further revolutionized the field of genomics by vastly increasing the speed and reducing the cost of nucleic acid sequencing (Heather & Chain 2016). NGS uses a shotgun approach to randomly sequence short fragments of nucleotides that are representative of the much larger sequence of interest (e.g. entire transcriptome). One major application of NGS, RNA sequencing (RNAseq), has revolutionized the field of transcriptomics: the ability to study the complete set of transcripts from a given cellular or tissue source. Although early transcriptome-wide drug screens used transcript hybridization technologies like microarray, RNAseq is fast becoming the preferred method for transcriptome-wide studies (Wang et al. 2009; McGettigan 2013). In comparison to microarrays, RNAseq is able to identify more low-abundant transcripts and, due to its greater dynamic range, can identify more differentially expressed genes in response to drug treatment (Zhao et al. 2014). RNAseq also requires fewer replicates because it identifies differential expression more accurately than microarray technology (Marioni et al. 2008). The major draw-back of RNAseq is the cost associated to the reagents and the sequencing. RNAseq also suffers from the same problems associated with multiple hypothesis testing, thus statistically reliable methods to reduce the false discovery rate (FDR) are necessary (Noble 2009). Even with statistical methods to reduce the FDR, gene expression that is identified by RNAseq must be validated at the individual gene level.

Thus qRT-PCR retains its usefulness as it is the gold-standard for validating gene expression results from transcriptomic screens.

Rare neurogenetic diseases

The optimal choice of a culture system for a transcriptome-screening experiment is a balance between feasibility and physiological relevance. A correlation exists between the scope of an experiment and the feasibility of an experiment; the greater the scope, the more feasible the experiment must be and thus a model with less physiological relevance must often be used. Traditional high-throughput drug screening (HTS) used non-cellular protein based approaches to screen hundreds of thousands of compounds (Hertzberg & Pope 2000; Macarron 2006; Macarron & Hertzberg 2011). To increase physiological relevance, HTS has been adapted for cell-based assays using immortalized cell lines that are ideal for high-throughput experiments because of their ability to generate large quantities of cells in a timely and cost-effective manner. Yet immortalized cells have also been shown to in general be poorly predictive of the effects of drugs in living organisms (Eglen et al. 2008; Xia & Wong 2012). Primary human cell cultures are more physiologically relevant than immortalized cells but require biopsies from a patient or volunteer to harvest the cells. However, biopsies are not possible for many organs including the brain, thus a physiologically relevant alternative is needed for identifying drug candidates for rare neurological genetic (neurogenetic) diseases. The Genetic and Rare Diseases Information Center of the United States National Center for Advancing Translational Sciences identifies ~1200 rare genetic diseases as nervous

system diseases (NIH n.d.). This means that about one third of the genetically characterized rare monogenic diseases have a neurological component. Neurological diseases also make up over two-thirds of the C4R prioritized disease list. Query of the FDA's clinical trials database (https://clinicaltrials.gov) revealed that few C4R prioritized diseases have associated clinical trials, and many of the clinical trials that are listed are either for enzyme replacement therapy or dietary treatments (**Table 1**).

Table 1. Clinically tractable neurogenetic C4R diseases. Each disease has a defined genetic aetiology, is monogenetic, and has a neurologic phenotype. Diseases are classified by inheritance pattern, genotype/phenotype relationship and the potential rescue mechanism. Only diseases that might be treated by gene upregulation are included. Number of clinical trials were obtained from ClinicalTrials.gov based on "gene symbol" + "disease name" searches. Clinical trials of devices and surgeries were not included in the number of clinical trials.

Gene	Disorder	OMIM#	Inheritance	Gene to upregulate	# of clinical trials
		Hypomorphic			
AFG3L2	Spastic ataxia type 5	614487	AR	AFG3L2	0
ALDH18A1	Autosomal recessive cutis laxa type IIIA	219150	AR	ALDH18A1	0
AMACR	Apha-methylacyl-CoA racemase deficiency	614307	AR	AMACR	1
ARSA	Metachromatic leukodystrophy	250100	AR	ARSA	22
ASAH1	Farber disease	228000	AR	ASAH1	0
	Spinal muscular atrophy with progressive myoclonic epilepsy	159950	AR	ASAH1	0
ASL	Argininosuccinic aciduria	207900	AR	ASL	0
ASPA	Canavan disease	271900	AR	ASPA	4
ATP7B	Wilson disease	277900	AR	ATP7B	10
BCKDHA	Maple Syrup urine disease	248600	AR	BCKDHA	1
BCKDHB	Maple Syrup urine disease	248600	AR	BCKDHB	1
CLN3	Ceroid Lipofuscinosis type 3	204200	AR	CLN3	13
CTSA	Galactosialidosis	256540	AR	CTSA	0
EIF2B5	Central hypomyelination and vanishing white matter disease	603896	AR	EIF2B5	0
FH	Fumarase deficiency	606812	AR	FH	0
GALC	Krabbe disease	245200	AR	GALC	11
GLB1	GM1-Gangliosidoses Type 3	230650	AR	GLB1	0
GUSB	Mucopolysaccharidoses VII	253220	AR	GUSB	2
HARS	Usher syndrome type IIIB	614504	AR	HARS	0
HEXA	Tay Sachs disease	272800	AR	HEXA	12
HEXB	Sandhoff disease	268800	AR	HEXB	0
HSD17B4	D-Bifunctional protein deficiency	261515	AR	HSD17B4	0
IDS	Mucopolysaccharidosis II	309900	AR	IDS	22
IDUA	Mucopolysaccharidosis I	607014	AR	IDUA	13
MAN2B1	Alpha-Mannosidosis type I	248500	AR	MAN2B1	0
MUT	Methylmalonic aciduria, mut type	251000	AR	MUT	0
NEU1	Sialidosis type 1	256550	AR	NEU1	0
OTC	Ornithine transcarbamylase deficiency	300461	XLR	OTC	9

PHYH	Adult Refsum disease	266500	AR	PHYH	0
PLP1	Pelizaeus-Merzbacher disease	312080	AR	PLP1	3
PMM2	Congenital disorder of glycosylation type 1C	212065	AR	PMM2	0
POLR3A	Hypomyelinating leukodystrophy 7	607694	AR	POLR3A	0
PPT1	Ceroid lipofuscinosis type1	600722	AR	PPT1	1
RDH12	Leber congenital amaurosis 13	612712	AR	RDH12	0
SACS	ARSACS	270550	AR	SACS	0
SCARB2	Action myoclonus renal failure syndrome	254900	AR	SCARB2	0
SGSH	Mucopolysaccharidoses IIIA	252900	AR	SGSH	5
SLC16A2	Allan-Herndon-Dudley syndrome	300523	XLD	SLC16A2	1
SLC6A8	Creatine transporter deficiency	300352	XLR	SLC6A8	0
SUMF1	Multiple sulfatase deficiency	272200	AR	SUMF1	0
TYMP	Mitochondrial DNA depletion syndrome type 1	603041	AR	TYMP	0
		Haploinsufficient			
AFG3L2	Spinocerebellar ataxia type 28	610246	AD	AFG3L2	0
ATP1A2	Familial hemiplegic migraine type 2	602481	AD	ATP1A2	0
CSF1R	Hereditary diffuse leukoencephalopathy with spheroids	221820	AD	CSF1R	0
FGF14	Spinocerebellar ataxia type 27	609307	AD	FGF14	0
GRN	Ubiquitin-positive frontotemporal dementia	607485	AD	GRN	1
ITPR1	Spinocerebellar ataxia type 15	606658	AD	ITPR1	0
MAPT	Wilhelmsen-Lynch disease	600274	AD	MAPT	0
MPZ	Charcot-Marie-Tooth disease type 1B	118200	AD	MPZ	0
NKX2-1	Hereditary benign chorea	118700	AD	NKX2-1	0
OPA1	Optic atrophy type 1	605290	AD	OPA1	0
PMP22	Hereditary neuropathy with liability to pressure palsies	162500	AD	PMP22	0
PRPF31	Retinitis pigmentosa type 11	600138	AD	PRPF31	0
SCN1A	Dravet syndrome	607208	AD	SCN1A	0
SLC2A1	GLUT1 deficiency syndrome 2	612126	AD	SLC2A1	12
SPAST	Hereditary spastic paraparesis type 4	182601	AD	SPAST	0
		Null			
ABCD1	X-linked adrenoleukodystrophy	300100	XLR	ABCD2 (paralog)	23
DDHD2	Spastic paraplegia 54	615033	AR	DDHD1 (paralog)	0

Abbreviations: AD (autosomal dominant), AR (autosomal recessive), XLR (X-linked recessive), XLD (X-linked dominant).

Neuronal tissue culture models

The utility of a neuronal culture system to identify treatments for neurogenetic central nervous system (CNS)-related diseases depends on the degree to which the culture approaches the physiological characteristics of the human brain. Two important considerations for neuronal cultures are the cellular heterogeneity and the level of maturity of the cultures. Although mammalian neuronal complexity is no doubt far greater than this, historically neurons in the cerebral cortex have been classified in six layers, each with unique functional and morphological characteristics (Gupta et al. 2002). The neocortex consists of ~80% projection neurons and ~20% interneurons (Hendry et al. 1987; Meinecke & Peters 1987). The projection neurons are pyramidal glutamate releasing excitatory neurons while the interneurons are a variety of short-range gamma-aminobutyric acid (GABA) releasing inhibitory neurons. The variety of GABAergic interneurons are essential to neuron circuit formation, regulation of synaptic excitability, and cortical projection neuron plasticity (Wonders & Anderson 2006). In the mouse, cortical neurogenesis begins at embryonic day 10.5 and is mostly complete by embryonic day 16.5 (E16.5) (Greig et al. 2013) The development of astrocytes, the most common glial cell in the brain and in the neocortex, begins between E16 and E18 and peaks in the early post-natal period (Kaur et al. 2014; Miller & Gauthier 2007). In neuronal culture, the presence of some astrocytes is crucial for healthy neuron development and normal network formation (Nedergaard 1994; Araque & Navarrete 2010). However, an overabundance of astrocytes in a neuronal culture could mask the transcriptional response of neurons to drug treatments. For the purposes of a drug

screen, it is also essential to treat neuronal cultures that have a transcriptome profile that is stable and mature. This is because transcriptional noise from maturing neurons could mask drug-induced transcriptional changes of neuronal genes. This results in the following dilemma: because it is difficult to culture neurons long-term *in vitro*, the majority of primary neuron experiments are performed on neurons that have only been grown *in vitro* for 7-14 days (Schock et al. 2012) whereas several reports have shown that cortical neurons isolated from embryonic mice or rats are not transcriptionally mature until at least 21 days *in vitro* (DIV21) (Schock et al. 2012; Lesuisse & Martin 2002). Vesicular glutamate transporter (vGLUT) and vesicular GABA transporter (vGAT) are required for packaging of their respective neurotransmitters (glutamate and GABA) into presynaptic vesicles that are subsequently released into the synaptic cleft and thus are taken as hallmarks of cortical neuron synaptic maturity (Bellocchio et al. 2000; Wojcik et al. 2006). Schock *et al.* (2012) studied the expression of both transporters in embryonic rat cortical neuron cultures and showed that neither is expressed at significant levels before DIV21.

The three *in vitro* systems most commonly used as surrogates of the human brain for drug screening are (1) human neuron-like immortalized cells, (2) human induced pluripotent stem cell (iPSC) derived neurons (iNs), and (3) rodent primary neuronal cultures (mouse or rat).

1. Due to the lack of access to healthy cells from the human central nervous system, most neurological diseases have been studied in neuron-like cancer cells, most

commonly SH-SY5Y (human neuroblastoma cells) and NTera (human neuronal committed teratocarcinoma cells). Immortalized cells like SH-SY5Y and NTera are very useful for drug screening because they can divide indefinitely, are easy to maintain, and are very homogeneous with reduced batch to batch variability. The SH-SY5Y and NTera cells can also be pharmacologically differentiated into cells that are morphologically similar to neurons and express neuronal markers (Agholme et al. 2010; Gordon et al. 2013). However, because of culturing over many generations, these immortalized cells nonetheless are prone to genetic and epigenetic drift accumulating many mutations and genomic reorganization resulting in a loss of physiological relevance (Krishna et al. 2014; Nestor et al. 2015).

2. Recent advances in stem cell technology have led to a revolution in the field of neuronal cultures. Induced pluripotent stem cells (iPSCs) can be generated from human somatic cells and then the iPSCs can be differentiated to neuronal cells (iNs). There now exist protocols to generate iNs of many different lineages from iPSCs or even directly from patient peripheral cells (Barral & Kurian 2016). Critically, iNs allow researchers to study the effect of gene mutations and disease on patient derived cells in the neuronal context (Barral & Kurian 2016). Induced neurons are increasingly being used to screen drugs for many neurological diseases (Xu & Zhong 2013; D'Aiuto et al. 2014). However, a significant drawback to the use of iPSC derived neurons is reduced physiological relevance due to the persistence of epigenetic marks (from the somatic cells) that can alter

the landscape of transcriptional expression of iNs (Vaskova et al. 2013). Also, human iNs appear to need to be co-cultured with induced astrocytes and then allowed to mature for ~30 weeks before they reach electrophysiological maturity similar to the human brain (Odawara et al. 2016).

3. Rodent primary cell cultures are physiologically relevant and can contain a good complement of neuronal and supporting cells. Although a major drawback of primary cultures is the non-human sourcing of cells, there is nonetheless a high degree of similarity between the mouse and human genomes with similar gene expression (Consortium et al. 2002). Many different types of primary neuronal cultures from mice and rats have been described. Rodent primary dissociated cultures are commonly generated from cerebral cortex and striatum to model the CNS, and dorsal root ganglia to model the PNS (Brewer et al. 1993; Lesuisse & Martin 2002; Chumakov et al. 2014). These cultures have been shown to achieve physiological and morphological maturity *in vitro* within 21-35 days. Although primary cultures are difficult to generate en masse to produce large enough quantities for drug screening, several recent reports have conducted transcriptome-wide studies on primary neuronal cultures. One such study used mouse cortical cultures at DIV9 with RNAseq to characterize the role of thyroid hormone on neocortical development(Gil-Ibanez et al. 2017). Another report studied the effect of rotenone on primary cortical neurons performed by micro-array (Yap et al. 2013). Primary neuronal cultures have also been proposed as a

good model for high-content screening to identify candidates for Parkinson's disease (Daub et al. 2009)(Sharma et al. 2012).

Regardless of the cell culture model used for drug screening, there is no feasible way to model the effect of the blood-brain-barrier (BBB) on drug delivery to the brain. The BBB is a diffusion barrier composed of specialized endothelial cells that tightly regulates transfer of hydrophilic molecules from circulating blood to the brain (Ballabh et al. 2004). Although the BBB is crucial to normal physiological function of the brain, it is only permeable to small, lipophilic molecules and thus poses problems for drug and biologics delivery to the brain (Nau et al. 2010). Thus, one of the major challenges for neurogenetic disease drug development is ensuring that a drug can penetrate the BBB. The ratio of systemic blood drug levels to cerebral spinal fluid (CSF) levels is a commonly used BBB penetrance metric; when viewed from this perspective, some clinic-tested drugs are seen to be BBB penetrant and thus are suitable candidates to test as neurogenetic disease therapeutics.

Primary cortical culture drug screening for rare neurogenetic disease therapeutics

Because drug screening has traditionally been performed in non-cellular assays or immortalized cell line models, there is a lack of transcriptional data measuring the effect of clinic-tested drugs in the context of the CNS. Thus, at the present, it is difficult to identify physiologically relevant drug-gene hits with therapeutic potential for rare neurogenetic diseases. I therefore performed a drug screen of 218 clinic-tested BBB penetrant drugs in mouse primary cerebrocortical cultures – hereafter referred to as the "Neuron Screen". Cerebrocortical cultures collected from E16.5 outbred mice were treated at DIV21 with the clinic-tested drug library. Transcriptomic-profiles of drug and vehicle treated cerebrocortical cultures were established by RNAseq and bioinformatic analysis of the sequencing was performed in collaboration with Pfizer's Computational Sciences Center of Emphasis (CSCoE).

In the past, our lab has interrogated the CMap to identify candidates to treat rare genetic diseases by transcript-modulating therapy. In similar fashion, the differential expression data from the Neuron Screen was queried for drugs causing modulation of the ~60 C4R neurogenetic disease-related genes. Because the majority of neurogenetic diseases are caused by a lack of gene expression (hypomorphic, haploinsufficient), I chose to focus on identifying drugs that upregulate disease-related genes; in this manner, a list of drug-gene hits that involved a C4R neurogenetic disease gene was assembled (type I hits). Because a high false discovery rate plagues individual drug-gene hits that are identified from transcriptomic data (Mears et al. 2017), I hypothesized that

the validation rate of drug-gene hits that are part of a common network would be higher than for drug-gene hits that are individually identified. A network is defined as a set of structurally and/or functionally related genes that are modulated by the same drug treatment (Abatangelo et al. 2009). To facilitate the network approach to identifying drug-gene hits, the commercially available bioinformatic software Ingenuity Pathway Analysis (IPA) was used to identify network-associated drug-gene hits (type II hits). Both type I and type II hits were further investigated *in vitro* and *in vivo* by means of qRT-PCR and western blotting; a higher percentage of type II hits were validated compared with type I hits while one drug-gene hit of both types showed promising results *in vivo*.

Goals of the Neuron Screen project

1. Create a transcriptome-wide differential expression database of clinic-tested, BBB penetrant drug activity in cerebrocortical cultures relevant to the human brain.
2. Identify drug repurposing candidates to treat at least one of the ~60 C4R-prioritized rare neurogenetic diseases, based on the BBB penetrant drug screen.
3. Investigate the reproducibility of drug-gene hits from a novel gene-network approach to studying the Neuron Screen.

MATERIALS & METHODS

Animals and treatments

All protocols were approved by the Animal Care and Veterinary Services (ACVS) and Ethics board of the University of Ottawa. For *in vivo* drug studies, male C57BL6 mice (8 weeks old) and male Sprague-Dawley rats (7 weeks old) were obtained from Charles River. Mice and rats were treated as per the details found in **Table 3**. At the end of each treatment experiment, ~4 hours after the last dose, animals were anesthetized by isoflurane and euthanized by cervical dislocation (mice) or decapitation (rats). Bilateral cortex and cerebellum were collected from each animal and flash-frozen in liquid nitrogen. For the levothyroxine-treated rats, bilateral sciatic nerve was also collected.

Tissue culture

Mouse 3H cortical culture

Female FVB/N mice were crossed with male C3B6F1 (Jackson Labs) to create triple hybrid (3H) embryos. At gestational age 16.5 days, dams were anaesthetized by isoflurane and euthanized by cervical dislocation. Embryonic day 16.5 (E16.5) 3H embryos were placed in ice cold sterile HBSS and decapitated. Cortex was dissected and leptomeninges were removed. Cortices collected on each dissection day were pooled and cells were dissociated mechanically by trituration. Tissue that was not dissociated was allowed to settle for 1 minute and the supernatant containing the neurons was transferred to a new tube. Dissociated cortices were then centrifuged at 200g for 5 minutes and single-cell filtered through a 0.4μm filter. Cortical cultures were plated on

6-well and 96-well poly-D-Lysine (PDL, 50μg/mL) coated polystyrene plates and maintained in Neurobasal medium containing B27 supplement (2%), GlutaMAX (2%), 100 I.U./mL penicillin, and 100 μg/mL streptomycin. Plating densities were one million and 60,000 cells per well for the 6-well and 96-well plates respectively. All cells were cultured at 37°C with 5% CO2. For the Neuron Screen and follow-up validation experiments, cortical cultures were grown for 21 days *in vitro* (DIV21). One week after plating a 75% media change was performed, with subsequent 75% media changes every 2-3 days until DIV21.

Dorsal root ganglion (DRG) culture

Cryopreserved rat DRGs were a generous gift from QBM Cell Science. DRGs were grown on 12-well (60,000 cells/well) and 96-well (5,000 cells/well) plates coated overnight with PDL (50ug/ml) and subsequently with laminin (20μg/mL) for 1 hour at 37°C. DRG cultures were seeded in Neurobasal medium containing B27 supplement (2%), GlutaMAX (2%), 100 I.U./mL penicillin, 100 μg/mL streptomycin, and 50ng/mL nerve growth factor (NGF). A 50% media change was performed after 4 days, and subsequent 75% media changes were performed every 2-3 days. After 7 days *in vitro*, media was supplemented with 50μg/mL ascorbic acid (AA) to induce myelination and treated with DMSO (0.1%) or triiodothyronine (T3) (10nM, 31.6nM, 100nM). Cultures were maintained for 21 or 28 days *in vitro* after addition of AA and T3.

Human glioblastoma cell culture

Human U87 glioblastoma cells (ATCC) were cultured in DMEM high-glucose supplemented with 10% fetal calf serum (FCS) and 2mM L-glutamine. For the initial validation of FOXM1 targets, U87 cells were plated in 10 cm dishes at $1.2x10^6$ cells per dish in a volume of 10mL. After an over-night settling period, U87 cells were treated for 0, 4 or 8 hours with 0.25mM hydroxyurea (HU). For the dose-response curve, cells were plated as before and treated for 8 hours with 0, 0.0125, 0.025, 0.05, 0.125, 0.25, 0.5, 1, or 2mM HU. For the time-course experiment, cells were treated with 0.5, 2mM HU or 0.2% DMSO for 4 and 8 hours. After treatment end-points, cells were washed with 1x phosphate buffered saline (PBS), trypsinized and pelleted by centrifugation at 300g for 5min. Pellets were resuspended in 10mL of 1xPBS and divided ~30% for RNA extraction and ~70% for protein extraction. Cells were centrifuged again and pellets for RNA extraction were stored -80C. Pellets for protein extraction were washed two times in 1xPBS, pelleted, and frozen at -80C.

Human lymphoblastoid cell line culture

Human GM16119 lymphoblastoid cells (obtained from Coriell Institute) were cultured in RPMI supplemented with 10% FCS, 2mM L-glutamine, 100 I.U./mL penicillin, and 100µg/mL streptomycin. The GM16119 cells are B-lymphocytes from the peripheral vein of a healthy 19-year old male and immortalized by Epstein-Barr virus transformation. Cells were maintained and treated in 25mL flasks with a volume of 10mL. A few hours prior to treatments, cells were passaged into new flasks at a density

of $1x10^6$ cells/mL. For the time-course experiment, cells were treated with 0.5mM HU for 0, 1, 2, 4, 8, 12, 24 hours. After treatment, cells were divided into separate 15mL falcon tubes for RNA and protein extraction (~30% for RNA and ~70% for protein). Cell pellets for RNA were frozen at -80C immediately. Cell pellets for protein were washed two times in 10mL 1xPBS and then frozen at -80C.

Immunofluorescence

Immunofluorescence (IF) staining was performed on cortical cultures (DIV21) in 96-well polystyrene plates. Cells were fixed in 4% formaldehyde containing 7% (v/v) picric acid for 20 minutes at room temperature (Schock et al. 2012). Neurons were washed three times in 10mM PBS and subsequently stained for antigens of interest. Primary antibodies were diluted in 10mM PBS containing 0.3% triton X-100 (PBST). All primary antibodies except NeuN were incubated for 2 hours with vibration at room temperature or overnight at 4°C without vibration (Jacobsen & Staines 2004). NeuN was incubated with vibration at 4°C for ≥24 hours. Secondary Alexa Fluor antibodies (Abcam) were used at 1:400 in PBST and incubated for 35 min at 37°C. Secondary antibody for NeuN was incubated for 3 hours with vibration at room temperature. Imaging was performed using a Zeiss Axioskop II microscope, and whole-well imaging of DRG cultures was performed with an EVOS FL Auto 2 imaging system.

The concentrations used and sources for IF primary antibodies were as follows: mouse NeuN (1:10, gift from QBM Cell Science), mouse anti-GFAP (1:200, Chemicon), rabbit anti-MAP2 (1:2,000, Abcam), mouse anti-ALDH1L1 (1:200, NeuroMab), rabbit

anti-vGAT (1:500, Alpha diagnostics), guinea pig anti-vGLUT-1 (1:500, Millipore), mouse anti-Beta-3 Tubulin (Tuj-1) (1:10), rabbit anti-PMP22 (1:50, Novus Biologicals).

TUNEL staining

TUNEL staining was performed on 3H cortical cultures at DIV21 in 96-well format. Cells were prepared for IF staining as described above. After performing NeuN staining, cultures were stained using the *In Situ* Cell Death Detection Kit, Fluorescein (Roche). The procedure for performing TUNEL staining was followed according to the manufacturer's protocol for labeling adherent cells. Briefly, cells were incubated in the TUNEL reaction mixture in a humidified atmosphere for 60 min at 37°C in the dark. Cells were rinsed 3x in PBS and images were captured on a confocal microscope at 20x.

Neuronal network recordings

Network extracellular electrophysiological recordings of cortical cultures were performed with a microelectrode array (MEA) system from MultiChannel Systems. Triple hybrid mouse cortical cultures were grown on PDL (50μg/ml) and laminin (20 μg/ml) coated glass MEAs and recordings were obtained as described by Schock et al. 2012. Recordings were taken at DIV7, DIV15, and DIV21 and performed for a total of 300 seconds on each day from all 60 electrodes on the MEA. Analysis of burst frequency and synchrony was done by using the software SpAnNer (as outlined by Otto *et al*., 2003; Illes *et al*., 2009).

3H Cortical culture drug treatments

For the drug screen, 3H cortical cultures (DIV21) in 6-well plates were treated for 8 hours with vehicle (0.2% DMSO) or one of the 218 BBB-penetrant drugs used for the screen. 207 drugs were selected from the Screen-well v2 FDA approved drug library (Enzo Life Sciences) and 11 Pfizer shelf compounds were included in the screen. One to two vehicle treatments were performed for each neuron isolation day. Drug concentrations for the 207 clinic-tested drugs (Table S1) were determined based on literature review and in consultation with CHEO pharmacists. The 11 Pfizer drugs were used at a concentration of 1uM (per Pfizer's request). DIV21 cortical cultures were also used for investigating the drug-gene hits. For initial transcriptional investigation, 3H cultures were treated the same as for the screen, with the same dose and time (8 hours). For the half-log dose-curve, 3H cultures were treated with half-log dose-curves of the drugs outlined in **Appendix VII** for 8 hours. For the LDH assay, the same half-log dose-curves were used but cultures were treated for 8 and 14 hours.

cDNA library preparation and RNAseq

RNA was isolated from cortical cultures using the NucleoSpin RNA kit (Macherey-Nagel). Purity and concentration of total RNA was tested by Nanodrop analysis while the RNA integrity number (RIN) of each total RNA sample was determined using the Agilent RNA 6000 nano kit and the Agilent 2100 Bioanalyzer. Complementary DNA (cDNA) libraries were created using the KAPA Biosystems Stranded mRNAseq kit from high-quality RNA samples (RIN≥8). Chemical

fragmentation was performed per kit specifications to generate an average fragment size 200-300bps (confirmed by Agilent DNA 1000 kit). NEXTflex™ RNAseq barcode adapters were ligated onto fragmented RNA samples and adapter ligated libraries were amplified by polymerase chain reaction (PCR) for 10 cycles. Completed cDNA libraries were sequenced by Illumina HiSeq 2500 at the McGill Innovation Centre. Single-end 50bp sequencing was performed. Four cDNA libraries containing unique barcodes were multiplexed in each flow-cell lane to achieve a minimum of 30 million reads per library. The KAPA qPCR-based quantification kit was used to determine cDNA library concentration prior to multiplexing (performed by McGill Innovation Centre).

Differential expression analysis

Fastq files from Illumina sequencing were aligned to the mm10 reference genome using Star (version 2.4.0). All fastq files were validated to be single-stranded and reverse-stranded. Read counts were summarized to gene-level by the feature count module of Subread package (version 1.4.6). All samples passed a quality control performed using RSeQC (version 2.6.1). Genes that had raw counts = 0 and/or cpm < 1 across all 227 samples were filtered out of the dataset. Normalization was performed using limma/EdgeR package (version 3.10.5). Two levels of normalization were performed: 1) against library size; 2) against mean-variance relationship across count levels. The trimmed mean of M-values (TMM) method was used to calculate a weighted trimmed mean across all samples (Robinson & Oshlack 2010). Mean-variance relationship was estimated by the voom module, and appropriate observational-level

weights calculated with the input as log2CPM normalized multiplied by the normalization factors. Principle component analysis (PCA) was calculated to check for systemic bias across samples. The top 4 principle components were removed and the resultant residuals were used for subsequent analysis. Robust z-scores across values of all samples for each gene were calculated. P-values were calculated using 2-tailed test and multiple comparisons adjusted by the Benjamini Hochberg method to control for the false discovery rate (FDR). Hits were called if they showed an adjusted p-value (p-adj.) <0.05.

Ingenuity pathway analysis

Comparison analysis

The transcriptome-wide differential expression datasets from levothyroxine, liothyronine, betamethasone, and dexamethasone were entered individually into the Ingenuity Pathway Analysis (IPA) software (Qiagen Bioinformatics). The two metrics of differential expression that were entered were "z-score" and "p-adj" and data entered were limited to an absolute z-score greater >3. Each dataset was then individually analyzed by IPA. After analysis, the IPA comparison tool was used to determine the similarities between the transcriptional perturbations caused by levothyroxine and liothyronine and by the perturbations caused by dexamethasone and betamethasone. Comparison between drugs in each class was done by using the upstream regulator analysis (URA) heatmap function.

Upstream regulator analysis (URA)

The transcriptome-wide differential expression datasets of the 50 most transcriptionally active drugs were entered into the IPA software. Data-input was performed as described above in the "Comparison analysis" section. Each dataset was then individually analyzed by IPA and the URA function was used to identify upstream regulator pathways that were "activated" (z-score>2) or "inhibited" (z-score<-2). The upstream regulator networks thus identified were studied at the level of each gene-target of the upstream regulator. Only genes with a significance level of p-adj<0.05 were conserved as part of the pathway. A network was deemed "robust" if it contained ≥5 statistically significant gene-targets that were modulated in the same direction as the IPA prediction.

Quantitative RT-PCR

Quantitative RT-PCR was used to investigate the validity of hits from the Neuron Screen in cortical cultures and in DRGs. Gene-specific primers were designed for each gene of interest and *Gapdh* & *Hprt1* were used as housekeeping genes. The list of primers used for all qRT-PCR assays are shown in **Appendices II-IV**. All primers span ≥1 intron and were tested by agarose gel electrophoresis to confirm that they amplify one DNA species of the correct length. RNA was extracted after treatments and converted to complementary DNA (cDNA) using the Bio-Rad iScript advanced RT kit on a T100 Thermal Cycler using the following conditions: 25°C for 5 min, 42°C for 1 hour, 95°C for 5 min. Quantitative PCR (qPCR) was performed using the iQ SYBR Green

mastermix (Bio-Rad) on a CFX96 Touch Real-Time PCR Detection System. The conditions for qPCR amplification were as follows: 95°C for 3 min, followed by 39 cycles of 95°C for 10 sec, 57°C for 30 sec, and 72°C for 30 sec. A melting curve was performed from 55°C to 95°C at increments of 1°C/plate read at the end to confirm amplification specificity. The Bio-Rad CFX software was used to calculate geometric mean normalization of gene targets to both housekeeping genes (except when only one housekeeping gene is mentioned in a figure). For initial investigation of the 32 drug-gene interactions, hits were rejected if ≥2 biological replicates showed no difference from the control treated cells and when $p \geq 0.05$. Hits were accepted (validated) with ≥5 biological replicates and a statistically significant ($p < 0.05$) upregulation from the control.

Western Blotting

Cortical cultures were treated with aminophylline (90μM) for 24 hours and Nilotinib (10μM) for 12, 24, 48 hours. For the aminophyline-treated cells, cell lysis was performed using RIPA lysis buffer (150 mM NaCl, 1.0% NP40, 1mM EDTA, 0.5% deoxycholic acid, 0.05% SDS, 50 mM Tris HCl, pH 7.4) with freshly added Protease Inhibitor Cocktail (Sigma-Aldrich) at 1:100 dilution. Cell lysates were incubated on ice for 30 min and mixed by pipetting at 10 min intervals. Lysates were then transferred to polystyrene tubes and sonicated for 8 cycles of 15 sec "on" followed by 60 sec "off" in a Biorupter Pico with water cooler set at 4°C (Diagenode). Protein samples were subsequently centrifuged at 16,000g for 15 min at 4°C and then quantified by Bradford reagent (Bio-Rad). Protein sample (20μg) were prepared in Laemmli Sample Buffer (Bio-

Rad) (+5% of 2-mercaptoethanol), denatured for 5 min at 95°C, and run on 1.5mm 12% MiniPROTEAN TGX Stain-Free gels (Bio-Rad) at 200V for ~45 min. The gels were then UV activated for 45 sec and then the Trans-Blot Turbo blotting system (Bio-Rad) was used to perform semi-dry transfer onto low-fluorescence PVDF membranes at 1.3A for 10 min. After transfer, total protein normalization (TPN) images of the blots were obtained on the ChemiDoc MP Imaging System (Bio-Rad) using the Stain-free blot application. Membranes were then blocked in 5% milk in 1xTBS + 0.1% Tween 20 (TBST) for 1 hour and then incubated overnight at 4°C with the ALDH18A1 primary antibody (Sigma-Aldrich, HPA008333) at 1:1000 in 5% milk-TBST. After primary incubation, membranes were washed 3x for 10 min in TBST at room temperature. Blots were incubated in rabbit secondary horseradish peroxidase (HRP)-linked antibody (Cell Signaling) at 1:5000 in 5% milk-TBST for 1 hour at room temperature followed by 3x 10 min washes in TBST. Blots were then incubated in ECL-Clarity substrate (Bio-Rad) for 5 min and the chemiluminescent signal was detected on the ChemiDoc MP Imaging System (Bio-Rad). For aminophylline, protein levels were normalized to the total protein normalization (TPN). Western blots on nilotinib-treated cells were performed by Roxanne Lariviere (Montreal Neurological Institute).

For western blotting of protein from U87 and GM16119 cells, the same protocol was used as for aminophylline-treated 3H cultures, with a few modifications. Protein samples (40µg) were run on 1.5mm 10% MiniPROTEAN® TGX Stain-Free™ gels and transferred onto low-fluorescence PVDF membranes. Membranes were incubated overnight at 4°C with the following primary antibodies: FOXM1 (Cell Signaling, 5436) at

1:1000 in 5% bovine serum albumin (BSA)-TBST, CCNB1 (Cell Signaling, 12231) at 1:2000 in 5% BSA-TBST, PLK1 (Abcam, ab17056) at 1:2000 in 5% milk-TBST.

Lactate dehydrogenase (LDH) assay

The CytoTox 96 Non-Radio Cytotoxicity Assay (Promega) was used to determine the cellular toxicity of five Neuron Screen drugs on the 3H cortical culture system (96-well format cultures). Cultures at DIV21 were treated with a half-log dose curve of fenofibrate, nilotinib, aminophylline, diflunisal, and levothyroxine for 8 or 14 hours. The LDH assay was performed as per the kit specifications. Absorbance at 490nm was used to analyze the amount of LDH present in the culture media, on the SpectraMax 340 Microplate Reader (Molecular Devices). Media from living treated and non-treated cells was normalized to media from cells after complete cell death, which was achieved by freeze-thawing the cells.

Chromatin immunoprecipitation (ChIP)

Chromatin collection and preparation

The ChIP protocol was performed using the EZ-Magna ChIP kit. Human GM16119 cells were cultured as described above. Two hours prior to treatment, cells were passaged into fresh 25mL flasks at $1.2x10^7$ cells per flask in 10mL of media. Cells were treated with 0.5mM HU for a time-course of 0, 1, 2, 4, 8 hours. At the end of the treatment, cells were pelleted by centrifugation (300g for 5min) in 15mL flasks. Media was carefully aspirated and cell pellets were resuspended in 1% formaldehyde in 10mL

of RPMI culture media (no FCS). Cross-linking in formaldehyde was performed for 10min and formaldehyde was immediately quenched by addition of 1mL of 1.25M glycine. After a 5min incubation at room temperature, cells were centrifuged at 500g for 5min at 4°C. Media was aspirated and cells were washed 1x in 10mL ice cold PBS and 2x in 1mL ice cold PBS with added protease inhibitor cocktail (1:100 dilution). Cell pellets were then lysed in 500µL of lysis buffer (1% SDS, 10mM EDTA, 50mM Tris pH 8.1) for 15min with regular pipetting to mix. Chromatin was sheared by sonication on ice-water using a Vibra-Cell Ultrasonic Processor (Sonics) at an amplitude of 25%, for 8 cycles of 15sec ON/60sec OFF. Chromatin samples were then centrifuged at 12,000g for 10min at 4°C.

Immunoprecipitation

The immunoprecipitation was performed by combining and rotating over-night at 4°C the following: 44µg of chromatin, 2µg FOXM1 antibody (Diagenode, C15410232) or 5µg IgG (negative control) or 5µg Histone H3 (positive control), 20µL protein-A magnetic beads, 450µL IP dilution buffer (0.01% SDS, 1.1% Triton X-100, 1.2mM EDTA, 16.7mM Tris pH 8, 167mM NaCl). The chromatin-bead complexes were then washed on a rotating platform at 4°C for 5min with each of the following buffers: low-salt (0.1% SDS, 1% Triton X-100, 2mM EDTA, 20mM Tris pH 8.1, 150mM NaCl), high-salt (0.1% SDS, 1% Triton X-100, 2mM EDTA, 20mM Tris pH 8.1, 500mM NaCl), LiCl (0.25M LiCl, 1% IGEPAL, 1% deoxycholate Na, 1mM EDTA, 10mM Tris pH 8.1), TE (10mM Tris pH 8.1, 1mM EDTA). Decrosslinking was performed at 37°C in elution buffer (30min total)

followed by ~6hrs in a 65°C water bath. RNA contamination was removed by treatment with RNase for 30 min incubation at 37°C.

DNA purification and qPCR

Immunoprecipitated DNA was then purified by using the QIAquick PCR purification kit. Primers for the human *PLK1* gene were designed to amplify DNA in the following segments: -1kb to -500kb, -500kb to TSS (core promoter), TSS to +500kb, +500bp to +1kb (sequences available in **Appendix V**). qPCR was performed on the purified DNA with the following conditions: 95°C for 3 min followed by 39 cycles of 95°C for 10 sec, 63.5°C for 30 sec, 72°C for 30 sec. Control *Gapdh* and *Hprt1* primers were provided as part of the EZ-Magna ChIP kit.

RESULTS

CHAPTER 2

RNAseq clinic-ready drug screen in mouse primary cerebrocortical cultures

The outbred triple hybrid (3H) cortical cultures

Before performing the drug screen on mouse cerebrocortical cultures, it was necessary to validate the prospective culture as a physiologically relevant system, suitable for use in modeling the transcriptional effects of clinic-ready drugs on the mammalian brain. Cultures were derived from outbred embryos to model diverse human genetics more closely than inbred mouse models. The outbred strain used to create the cerebrocortical cultures was configured by studying the Jackson Labs mouse phylogenetic tree (Petkov et al. 2004), and was the filial 1 offspring of a cross between male hybrid B6C3F1/J mice and female FVB/NJ mice. I called this in-house outbred mouse model the triple hybrid (3H) mouse. Cortices were collected from 3H embryos at gestational day 16.5 (E16.5) to maximize the inclusion of neuronal subclasses and minimize non-neuronal mitotic cell contamination of the 3H cerebrocortical cultures. Based on the findings of several reports that studied the *in vitro* maturation of rodent cortical cultures, dissociated 3H cortices were cultured until day *in vitro* 21 (DIV21), and then characterized by immunofluorescence and electrophysiological assays.

3H cultures contain astrocytes and neurons

I first assessed the types and proportion of cells comprising the 3H cultures at DIV21. To this effect, 3H cultures were stained with antibodies against glial fibrillary acidic protein (GFAP) and neuronal nuclei (NeuN) to identify astrocytes and neurons respectively (**Fig. 2.1A**). Average cell counts of GFAP positive cells and NeuN positive cells (per 10x field) led to a ratio of ~1:5 astrocytes to neurons. Conversely, IF staining for the neuronal marker microtubule associated protein 2 (MAP2) and the astrocytic marker aldehyde dehydrogenase 1 family, member L1 (ALDH1L1) suggest equal numbers of astrocytes and neurons (**Fig. 2.1B**). However due to the lower intensity of staining and high background, ALDH1L1-stained cell counts could not be performed accurately. Surprisingly, the chromosomal counterstain 4',6-diamidino-2-phenylindole (DAPI) employed to identify the nuclei of neurons and astrocytes in 3H culture also revealed high-intensity punctate structures that were not associated with either MAP2 or ALDH1L1 stained cells (yellow arrows, **Fig. 2.1B**). Given the presence of DAPI staining that did not appear to be neuronal or astrocytic nuclei, I assumed the presence of other contaminating cells (e.g. microglial, endothelial). Thus, 3H cultures were stained with antibodies against the microglial specific marker IBA1 and the endothelial marker VE-cadherin; but IF staining for both cellular markers was negative (images not shown).

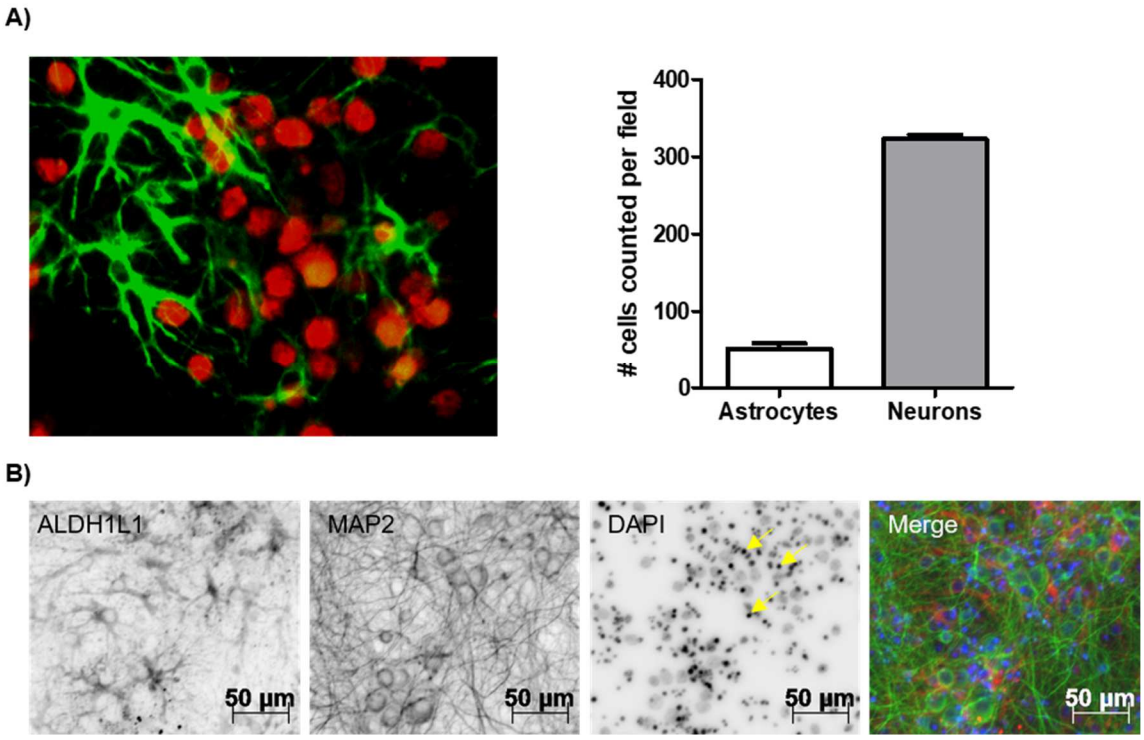

Figure 2.1. Establishing the ratio of astrocytes to neurons in 3H cultures. Cortical cultures were prepared from E16.5 3H embryos and cultured for 21 days. Cultures were fixed at DIV21 and processed for IF staining. All images were acquired on a Zeiss Axioskop II system at 10x magnification. **A)** Cultures stained with the neuron-specific stain NeuN (red) and the astrocyte-specific stain GFAP (green). The number of NeuN and GFAP positive cells were counted per microscope frame. The graph represents average counts of NeuN (neurons) and GFAP (astrocytes) positive cells (n=3). **B)** Cultures stained with MAP2 (green) for neurons, Aldh1l1 (red) for astrocytes, DAPI (blue) for nuclei. Yellow arrows indicate punctate DAPI staining.

The 3H neurons are healthy and show signs of developmentally normal pruning

Because the punctate DAPI-stained structures did not colocalize with neuron or astrocyte-specific staining and more importantly, had a much smaller diameter than the neuronal and astrocytic nuclei, I hypothesized that these might be hypercondensed chromatin. Because chromatin condensation is central to the process of apoptosis (Aras et al. 2008; Kumar et al. 2013), terminal deoxynucleotidyl transferase dUTP nick-end labeling (TUNEL) assay was employed to determine if the punctate DAPI-staining was related to apoptosis (Kyrylkova et al. 2012). To assess the 3H cultures for the presence of apoptotic nuclei and confirm the presence of healthy neuronal nuclei, cortical cultures were stained with NeuN and DAPI and then subsequently stained by a fluorescent commercially available TUNEL assay. TUNEL-stained chromatin appeared to co-localize predominantly with the punctate DAPI staining (white arrows) and not with neuronal nuclei (NeuN) staining (yellow arrows) (**Fig. 2.2)**. At DIV21, the NeuN stained neurons thus appear to be surrounded by the apoptotic remains of cells that probably died during the plating procedure and/or by physiological cellular attrition known to occur during network maturation.

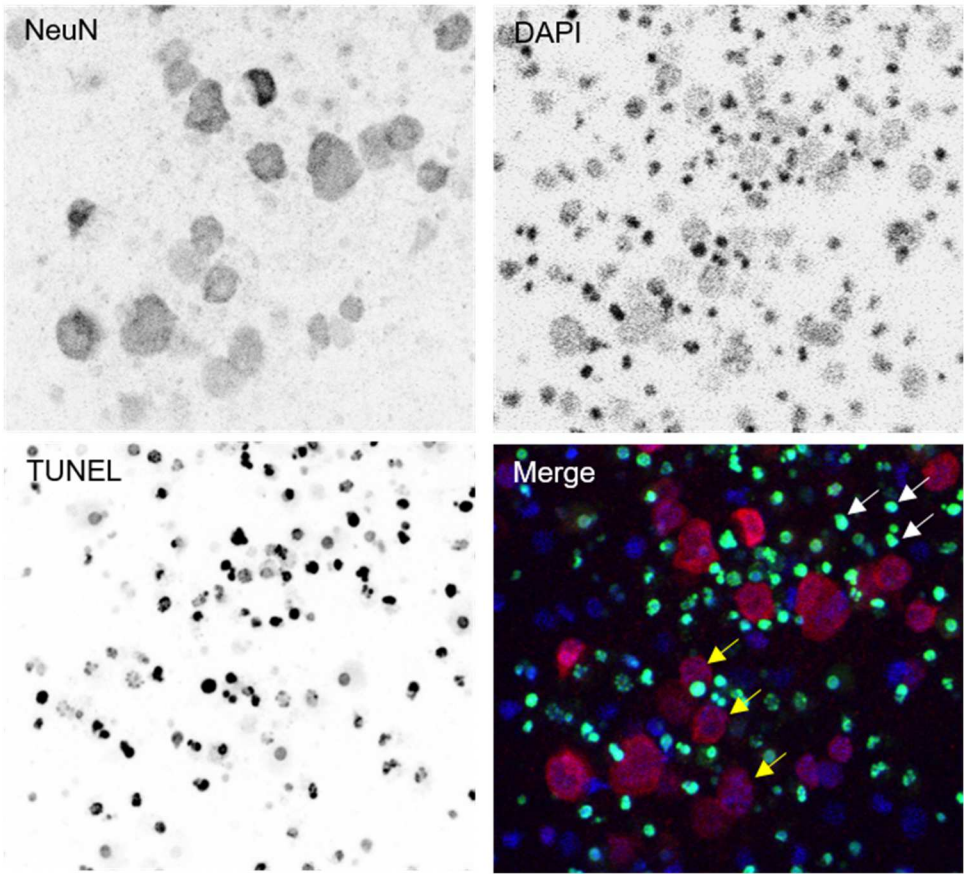

Figure 2.2. TUNEL staining of 3H cortical cultures. Cortical cultures were prepared from E16.5 3H embryos and cultured for 21 days. Cultures were fixed at DIV21. TUNEL staining was performed and cultures were subsequently stained with NeuN and DAPI. Images were taken on a confocal microscope at 20x magnification. On the merged image, white arrows show colocalized TUNEL and punctate DAPI staining, yellow arrows indicate colocalized NeuN and DAPI stained nuclei.

The 3H cortical cultures contain a good complement of GABAergic interneurons

The neuron specific markers that were used to determine the astrocyte to neuron ratio confirm the total neuron counts in the culture, but do not provide neuron sub-type information. GABAergic interneurons – i.e. interneurons that express gamma-aminobutyric acid (GABA) play a crucial role in modulating the activity of the glutamatergic projection neurons that constitute the majority of neurons in the mammalian cortex (Wonders & Anderson 2006). In mice, cortical migration of GABAergic interneurons from the medial and caudal ganglionic eminences is mostly complete by embryonic day 16.5 (Guo & Anton 2014; Miyoshi & Fishell 2011). Immunofluorescence staining of markers for three subtypes of GABAergic interneurons was performed on the 3H cultures; neuronal nitric oxide synthase (nNOS), parvalbumin (PV), and calretinin (CR) positive neurons are all present in the 3H culture model at DIV21 (**Fig. 2.3**).

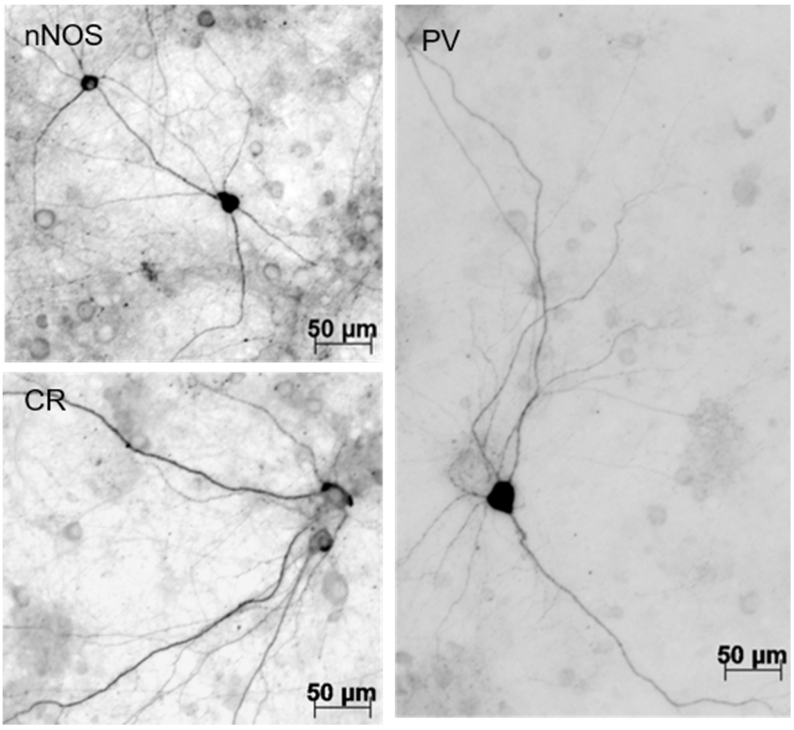

Figure 2.3. Identification of GABAergic interneurons in 3H cultures. Cortical cultures were prepared from E16.5 3H embryos and cultured for 21 days. Cultures were fixed at DIV21 and processed for IF staining. All images were acquired on a Zeiss Axioskop II system. Cortical cultures were stained for nNOS, calretinin (CR), or parvalbumin (PV). Representative neurons are shown for each class of interneuron found in the cultures.

3H cortical cultures form electrically active networks at DIV21

The maturity of the 3H cultures was assessed by immunological and electrophysiological methods. Immunofluorescence staining for vGLUT and vGAT was performed revealing punctate staining for both synaptic vesicle markers at DIV21 suggesting a good level of 3H culture synaptic maturity (**Fig. 2.4A**). To assess whether this apparently mature network of neurons was functional, the activity of 3H cultures grown on microelectrode arrays (MEAs) were recorded at DIV7, DIV15 and DIV21 (Schock et al. 2012). MEAs are tissue culture chambers embedded with metallic electrodes that are used to detect extracellular voltage changes emanating from electrically active cells such as neurons (Obien et al. 2014). The MEAs used to study the network activity of the 3H cultures consist of 60 titanium nitride electrodes. Recordings from each electrode were performed for 300 seconds to determine the presence of spontaneous neuronal depolarizations (spikes) at each electrode and the cumulative frequency and synchrony of spikes at all 60 electrodes. In the 3H cultures, spontaneous neuronal activity was observed at DIV7 and spike frequency and synchronicity increased as the cultures matured culminating in highly synchronous burst firing at DIV21, characteristic of mature cortical networks *in vitro* (**Fig. 2.4B**) (Otto et al. 2003; Illes et al. 2009).

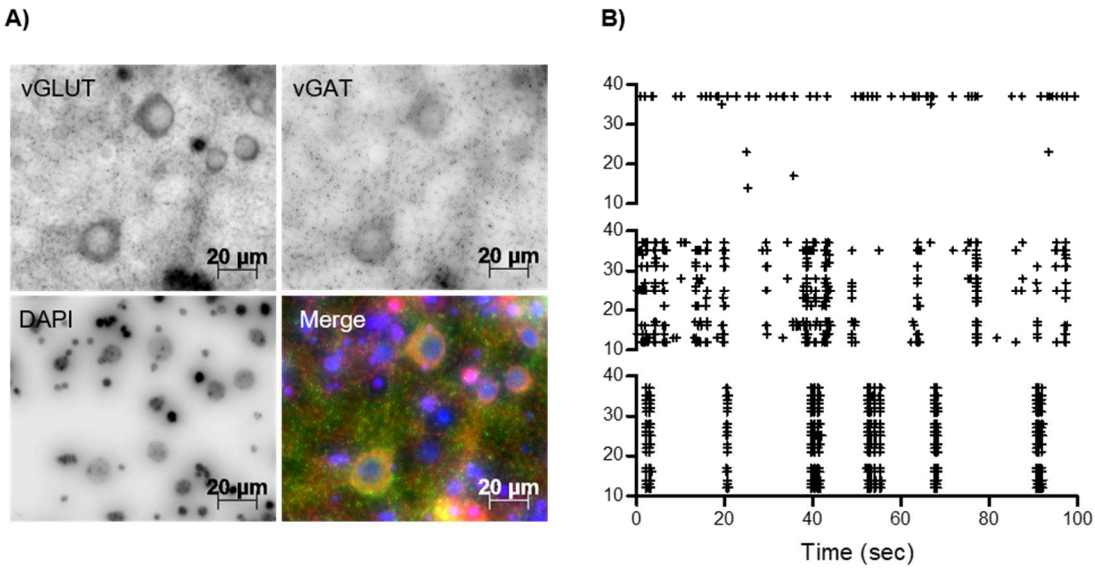

Figure 2.4. Immunological and electrophysiological characterization of 3H culture maturity at DIV21. Cortical cultures were prepared from E16.5 3H embryos and cultured for 21 days. Cultures were fixed at DIV21 and processed for IF staining. All images were acquired on a Zeiss Axioskop II system. **A)** 20x magnification of cortical cultures shows punctate vGLUT (red), vGAT (green), and nuclear DAPI (blue) staining. **B)** Cortical cultures were grown on glass microelectrode array (MEA) dishes and electrical activity was recorded once a week. Raster plots of electrical recordings are shown at DIV7 (upper), DIV15 (middle), and DIV21 (lower panel). Each mark represents burst activity for the corresponding electrode and time-point. Each panel shows burst activity from one third of the MEA dish.

3H cultures yield sufficient RNA for sequencing and PCR

In addition to confirming the physiological relevance of the 3H cortical cultures, there were logistical issues to address in using primary neuronal cultures for a drug screen; because neurons are post-mitotic, new animals must be sacrificed for each new batch of 3H cortical cultures. Furthermore, for the purposes of a transcriptome-wide screen, enough cells had to be cultured to provide adequate quantities of total RNA for RNAseq and qRT-PCR validation. To achieve the necessary yield of RNA, 3H cultures were grown on 6-well plates for the Neuron Screen and the follow-up validation experiments. Several plating densities of 3H cultures were tested; the optimal survival of neurons was determined to occur at $1x10^6$ cells per well. The total RNA yield from DIV21 cortical cultures in the 6-well format was 2-4μg, sufficient RNA for sequencing and validation experiments.

Neuron Screen implementation

The Neuron Screen was performed by treating 3H cultures at DIV21 with a subset of 207 drugs from a commercially available FDA approved drug library and 11 Pfizer shelf-compounds. Drugs were chosen for the Neuron Screen based on their putative ability to cross into the brain and based on their long-term side-effect profile; thus, most antipsychotics, chemotherapeutics, and epigenetic-modifying drugs (e.g. HDAC inhibitors) were excluded. An important consideration of drug screening is what dosage to use for each drug. Classically, pharmaceutical high-throughput cell-based drug screens have been performed at the standardized dose of 10μM (Lamb et al. 2006). However, since the 10μM dose is not physiologically relevant for all drugs, doses for the Neuron Screen were tailored according to the therapeutic serum concentration observed in humans.

To conduct the Neuron Screen, 3H cultures sufficient for 218 individual drug-treated and 9 control-treated conditions had to be established. An assembly-line to reduce the time from animal sacrifice to the plating of cells from 3H embryo cortices was therefore established; three people dissected out the brains and cortices of the embryos, while a fourth performed trituration, and cell plating. The cortices of up to three dams could thus be extracted, pooled, triturated and cultured resulting in ~1.5 hours elapsing from sacrifice to incubator. Mouse E16.5 cortices were dissociated and cultured on seven different days to achieve the number of cultures necessary to conduct the screen. The primary cortical Neuron Screen was performed by treating 3H cultures at DIV21 with

207 clinically approved drugs and 11 Pfizer shelf-compounds for 8 hours. The drugs included in the screen cover a wide range of therapeutic classes (>80 classes represented) (**Fig. 2.5**). The complete list of Neuron Screen drugs and associated concentrations can be found in **Appendix I.**

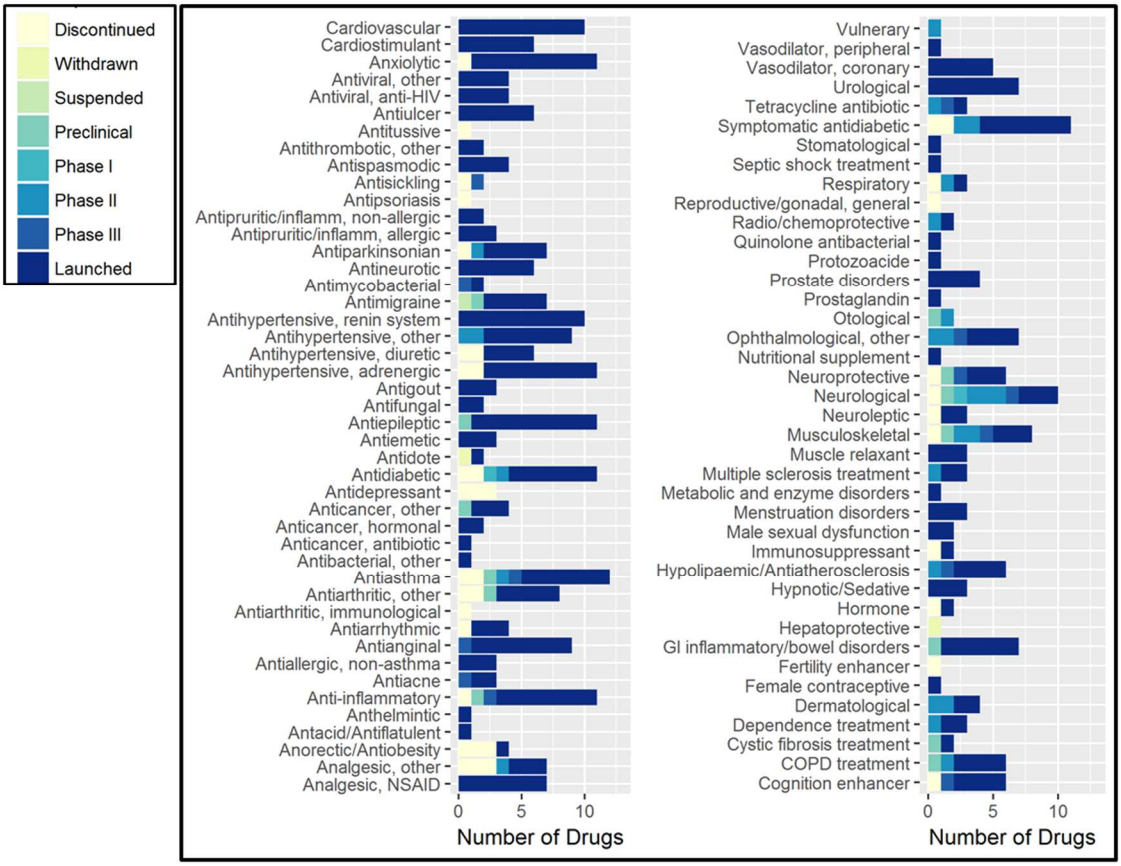

Figure 2.5. Therapeutic classification of drugs used for the Neuron Screen. Drugs were grouped by Pfizer's proprietary therapeutic class system. A wide range of classes is represented by the 219 drugs used for the screen. Each drug is further categorized by its current clinical classification. The drugs' clinical classification is represented by a different colour for each classification (see color-coded legend).

RNA sequencing of 3H cultures reveals a majority of protein coding genes

Following 8-hour drug treatments, total RNA was collected by column purification and integrity was assessed by microfluidic electrophoresis; all RNA samples had a RNA integrity number (RIN) greater than 8. Because rare disease genes are protein-coding, messenger RNA sequencing was performed. Since protein-coding mRNA genes constitute only 1-5% of the total RNA yield from mammalian cells, the exclusion of the far more abundant non-coding genes (e.g. ribosomal RNA, transfer RNA) is essential to obtain sufficient sequencing depth and maximize the coverage of protein-coding genes. Poly-A selection was used to obtain mRNA-enriched fraction of neuronal total RNA; this RNA fraction was used to construct a cDNA library corresponding to the 218 drug-treated and 9 control-treated 3H cultures. Sequencing was performed on the Illumina HiSeq 2500 platform, and differential expression analysis was completed by means of a proprietary differential expression pipeline (Pfizer). Over 96% of genes expressed in the cultures were identified as protein-coding. Long non-coding genes, a subset of which are also poly-adenylated, constitute ~3% while <1% of the genes identified in the Neuron Screen RNAseq data were from small non-coding RNA species (**Table 2**).

Table 2. Classification and numeration of Neuron Screen genes by biotype. All genes expressed in the mouse cortical neuron transcriptome were matched to their respective biotype group and Ensembl biotype. Each expressed gene had a cpm>1 for all 227 unique cDNA libraries included in the screen. Gene counts per Ensembl biotype and percent total expressed genes per biotype group are shown.

Biotype group	Ensembl biotype	# genes	% of total
Protein coding	protein_coding*	13587	96.58
	IG_C_gene*	2	
Long noncoding	lincRNA*	350	2.75
	processed_transcript	29	
	Antisense	6	
	bidirectional_promoter_lncrna	2	
Short noncoding	miRNA	20	0.38
	snoRNA	12	
	Mt_tRNA	8	
	rRNA	4	
	scaRNA	3	
	Mt_rRNA	2	
	Ribozyme	2	
	snRNA	2	
Pseudogene	Pseudogene	27	0.27
	polymorphic_pseudogene	7	
	unitary_pseudogene	2	
	processed_pseudogene	1	
	transcribed_processed_pseudogene	1	
Unspecified	-	3	0.02

Most clinic-ready drugs have very little effect on the neuronal transcriptome

Bioinformatic analysis of the RNAseq data was performed by means of a proprietary algorithm pipeline (Pfizer CSCoE) used to analyze the effect of each Neuron Screen drug on the protein-coding transcriptome. Differential expression analysis between samples was performed by normalizing the drug-induced expression of each gene (counts per million (cpm)) to its mean expression (cpm) in all 227 experimental (treatment and control) conditions. This novel technique for differential expression analysis was based on the hypothesis that most drugs have minimal transcriptional effect on the transcriptome; thus, the average response of most genes to most drugs will be equivalent to the untreated condition. As the Neuron Screen was performed as a single replicate, this normalization technique served to increase the statistical rigour of the differential expression results. Furthermore, to control for the effects of multiple hypothesis testing, Benjamini-Hochberg (BH) method was used to minimize the false discovery rate; a BH adjusted p-value (p-adj) was calculated for each drug-gene interaction with a statistical cut-off of $p\text{-adj}<0.05$.

To determine the interaction between the clinic-ready drugs and the neuronal transcriptome, the number of drugs that modulate a given gene (i.e. how many drugs affect "Gene A") (**Fig. 2.6**), and the number of protein-coding genes modulated by each drug (i.e. how many genes are affected by "Drug X") were counted (**Fig. 2.7**). Most genes were modulated by ≤1 drug, and the median number of genes modulated by the drug treatment group was equal to the median number of genes modulated by the control

treatment group; yet, the average number of genes modulated by the drug treatment group was greater than the average number of genes modulated by the control treatment group (**Fig. 2.7**). The six most transcriptionally active drugs were: fenofibrate (60), nilotinib (133), ciprofibrate (251), aminophylline (8), diflunisal (203), luteolin (194) (**Fig. 2.7**). Since most drugs in the Neuron Screen did not show an effect on the neuronal transcriptome above baseline, I reasoned that a small number of drugs were responsible for most of the gene modulation. Graphical representation of the number of gene modulations by the top 10% of drugs shows that most of the drugs modulate between 500 and 1000 genes, and the top 5 drugs modulate >1000 genes each (**Fig. 2.8**).

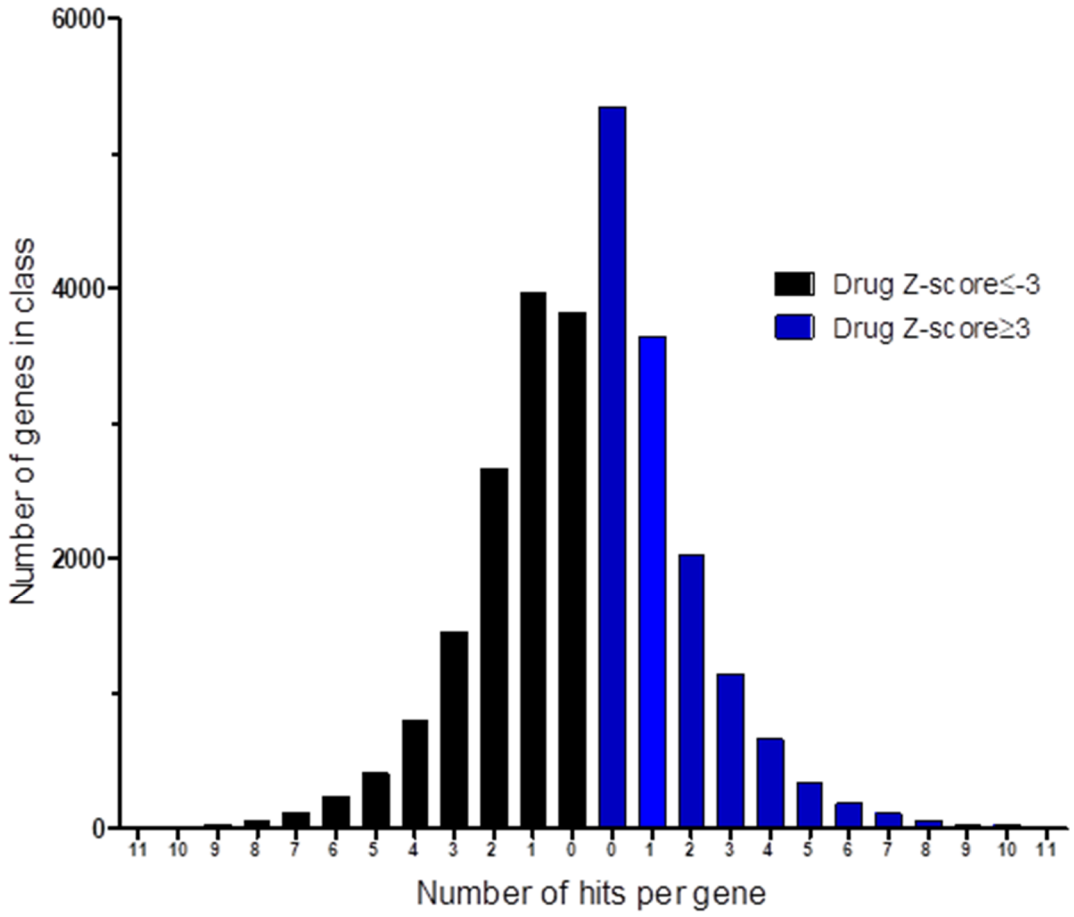

Figure 2.6. Quantification of transcriptome wide gene sensitivity in the Neuron Screen. The graph shows the total number of hits observed per protein coding gene in the neuronal transcriptome. The number of Neuron Screen drugs that modulate each protein coding gene (-3 > Z-score > 3) expressed in the neuronal transcriptome is shown on the horizontal axis. The number of genes in each classification (e.g. 0, 1, 2, 3 drug hits) is shown on the vertical axis. Negative Z-scores are shown in black, and positive Z-scores are shown in blue.

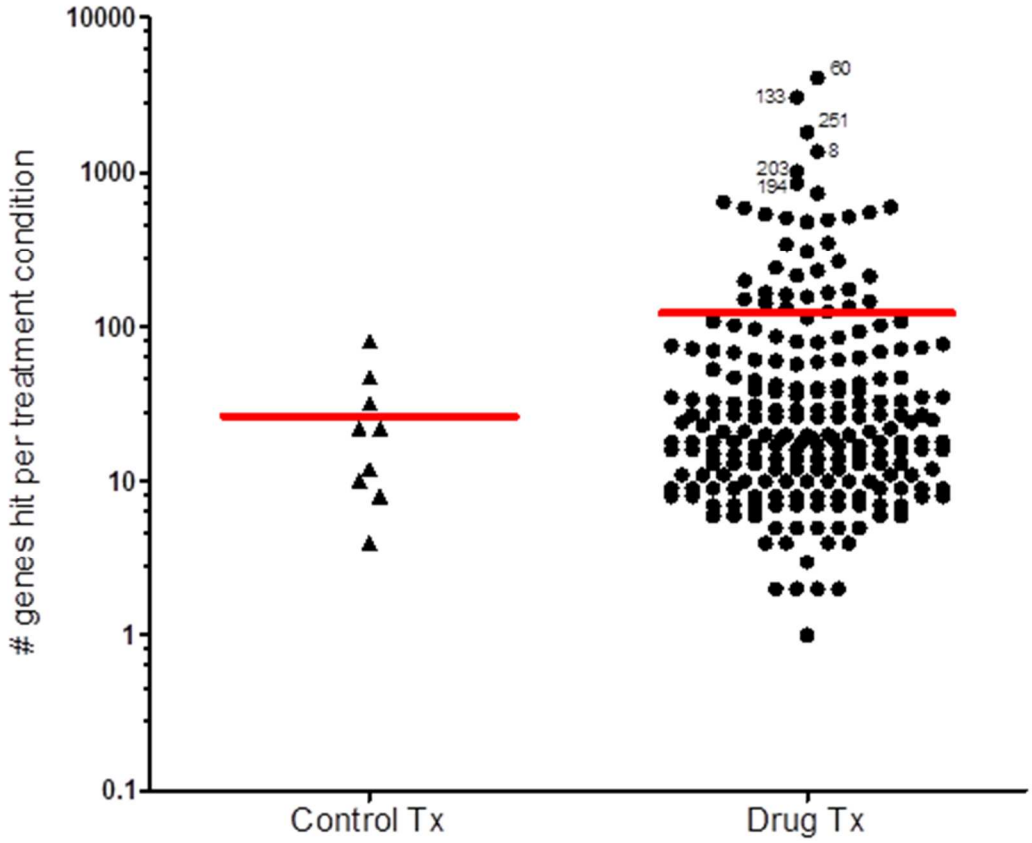

Figure 2.7. Quantification of transcriptome wide drug activity in primary mouse cortical neurons. Graph representing the total gene perturbations (p-adj<0.05) per sample from the Neuron Screen. Each triangle represents a control treated library and each circle represents a drug treated library. Solid lines represent the mean transcriptional modulation for control and drug treatment groups. The six most transcriptionally active drugs are indicated by their Neuron Screen drug number (see **Appendix 1**).

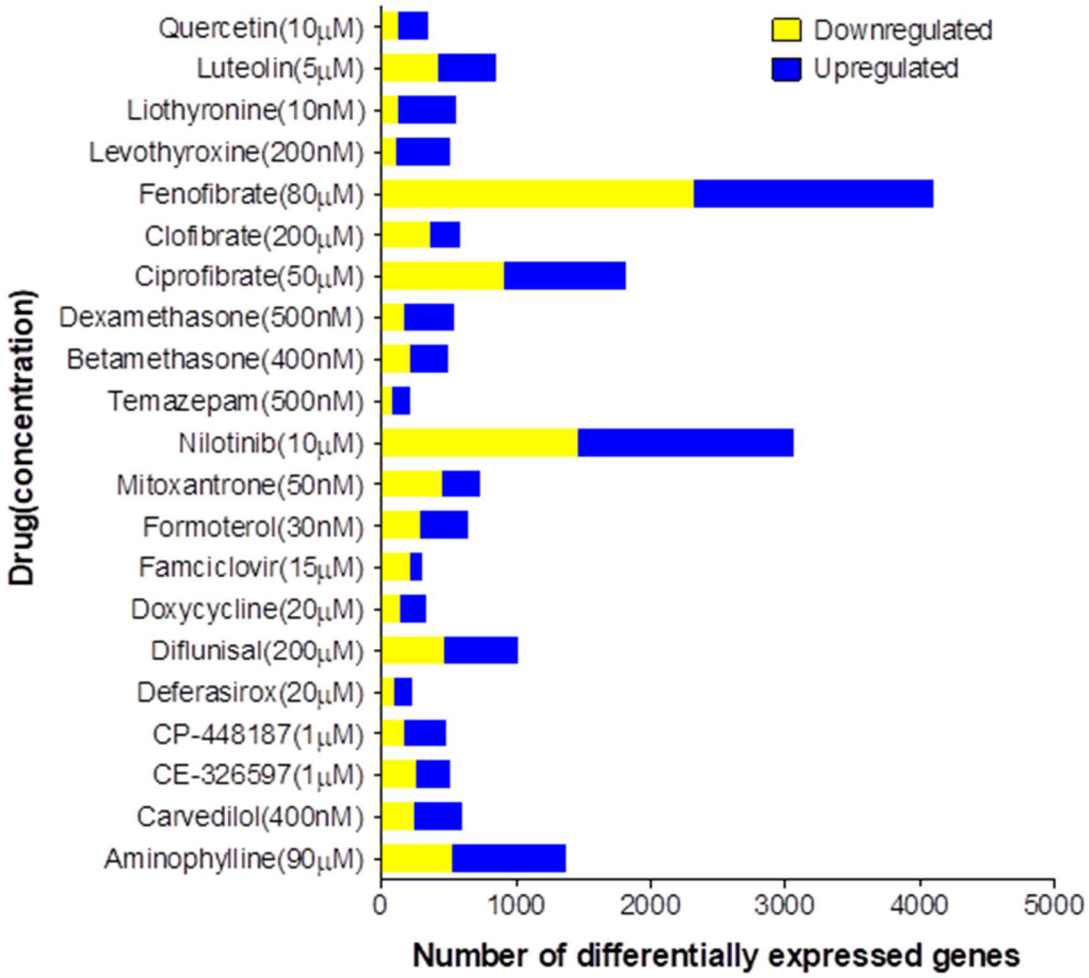

Figure 2.8. Total number of genes modulated (up/down) by the 10% most transcriptionally active drugs. Total counts of hits (p-adj<0.05) for the top 21 most transcriptionally active drugs from the Neuron Screen. Counts of downregulated genes are depicted by yellow bars and upregulated genes are depicted by blue bars. The drug concentration used for the screen is shown in parentheses after the drug name.

Thyroid hormone and corticosteroid drugs confirm reliability of Neuron Screen data

To explore the reliability of the Neuron Screen DE data, the Ingenuity Pathway Analysis (IPA) knowledge-base (comprised of hand-curated gene interactions from peer-reviewed medical literature) was used to compare the expression profiles of the thyroid hormone analogs levothyroxine (T4) and liothyronine (T3) (**Fig. 2.9A**). A similar comparison was conducted on the synthetic corticosteroids dexamethasone (DEX) and betamethasone (BMZ) (**Fig. 2.9B**). The "Comparison" tool of IPA, used to match the top putative upstream regulators (URA) in both drug classes, showed very similar profiles for both drug classes, with the most statistically significant shared upstream regulator being the drug used for the screen (thyroid hormone or DEX).

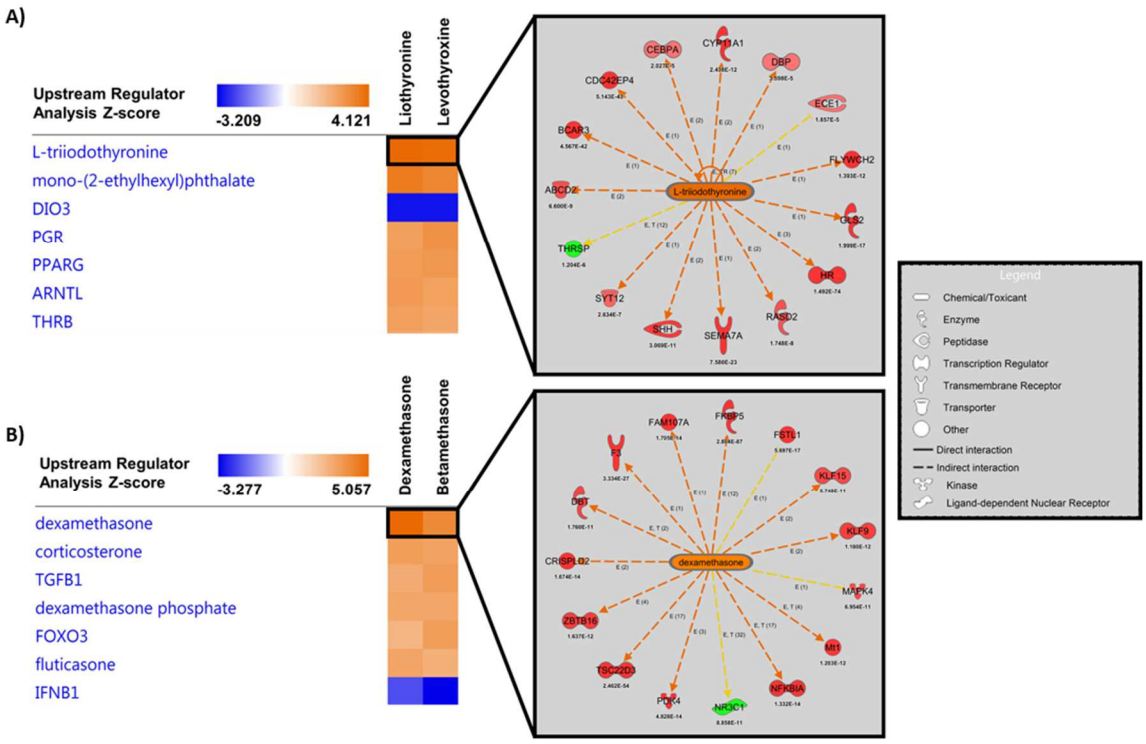

Figure 2.9. Comparison of neuronal transcriptomic drug-class effects. Transcriptome-wide data from the Neuron Screen was input into IPA for T3, T4, DEX, and BMZ. The "Analysis Comparison" tool enabled the creation of heat maps depicting the top upstream regulators for each class of molecules. **A)** Comparison between the two thyroid hormone analogs reveals that the top shared upstream regulator is L-triiodothyronine. **B)** Comparison between the two corticosteroids reveals that the top shared upstream regulator is dexamethasone.

Primary mouse cortical culture RNAseq drug screen database

I believe the neuronal RNAseq database shall serve as a valuable database not only for other neurogenetic disease researchers but in other realms of neuroscience. Therefore, the Shiny package from RStudio was used to create a user-friendly searchable database of the Neuron Screen. A screen-shot of the "Neuron Screen database" shows how the database can be searched by "gene symbol", "Ensembl ID", or "drug name", and the genes/drugs associated to the gene/drug of interest appears in a chart with corresponding z-score and p-value (**Fig. 2.10**). The database will be made publicly available as part of the publication of the results from the Neuron Screen project.

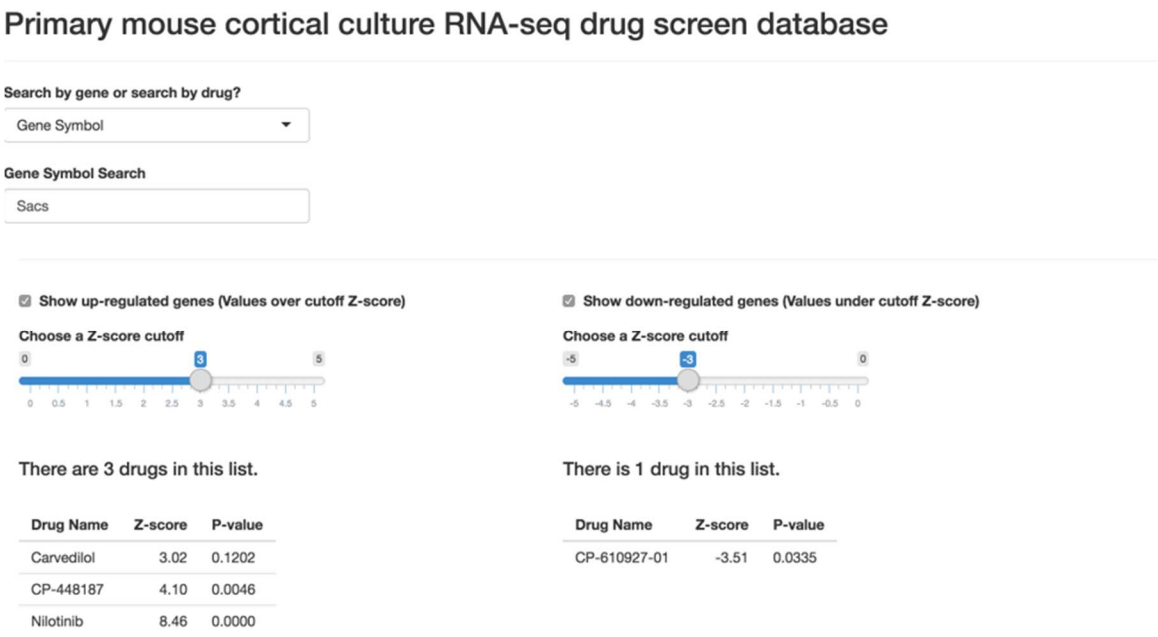

Figure 2.10. Screen-shot of the Neuron Screen database. The database was created using the Rstudio Shiny app. The transcriptome-wide differential expression results for each drug treatment from the Neuron Screen were included. The searchable interface of the screen enables a user to search by gene or by drug of their choice, to select the statistical cut-off of hits that will be shown in the table at the bottom of the screen. The table on the left shows up-regulated genes and the table on the right shows down-regulated genes.

CHAPTER 3

Mining the Neuron Screen transcriptome for neurogenetic drug-gene hits identifies thyroid hormone analogs as activators of the peripheral myelin protein 22 gene: potential implications for hereditary neuropathy with liability to pressure palsies

Repurposing candidate identification – gene-directed approach

In the past, our lab has interrogated the CMap to identify candidates to treat rare genetic diseases by transcript-modulating therapy. In similar fashion, the differential expression data from the Neuron Screen was queried for drugs causing modulation of the ~60 C4R neurogenetic disease-related genes. Because the majority of neurogenetic disease are caused by a lack gene expression (hypomorphic, haploinsufficient), I chose to focus on identifying drugs that upregulate disease-related genes; in this manner, a list of drug-gene "hits" was assembled. The drug-gene hits that involved a C4R neurogenetic disease gene were labeled "type I" hits. These hits were further studied to determine their validity by means of repeating the drug treatments from the screen in 3H cortical cultures and using qRT-PCR to confirm or reject the differential expression identified from the screen. Hits that validated in the primary cortical model were further studied at the protein level in a relevant primary neuronal cell model. The type I hits were further studied in rat or human cancer cell lines. Finally, select drug-gene hits were studied by qRT-PCR in the brain of wild-type mice and rats treated with the target drugs. A summary of the proposed type I hit validation protocol is presented in **Figure 3.1**, with accompanying number of hits studied at each validation stage.

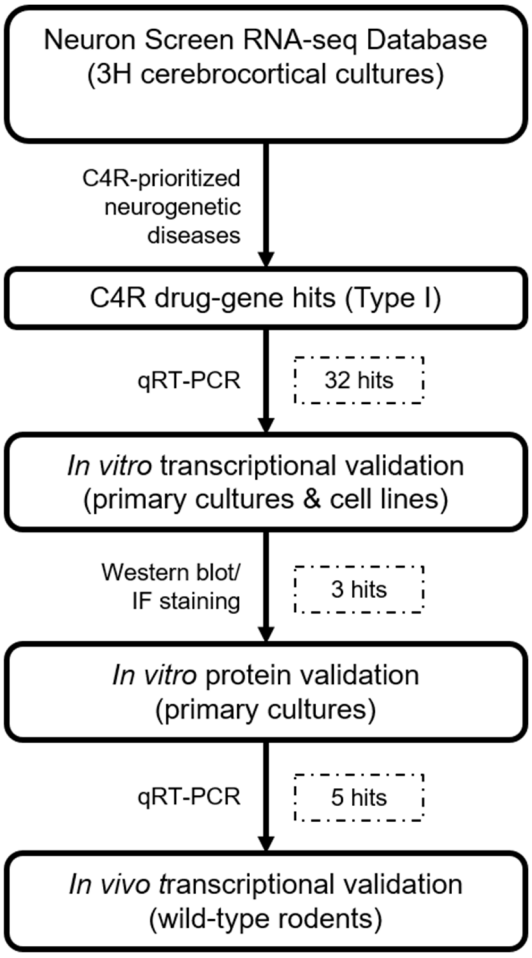

Figure 3.1. Diagram of validation strategy for Neuron Screen Type I drug-gene hits.

Dashed boxes identify the number of hits that were studied at each stage of validation.

Twenty percent of "type I" drug-gene hits identified in the Neuron Screen validate at the transcriptional level

To identify transcript-modulating repurposing candidates for rare neurogenetic diseases, the Neuron Screen database was queried for drugs that upregulate the ~60 rare neurogenetic disease genes prioritized by C4R. Such drug-gene hits were called "type I" hits and entered the validation pipeline if they were statistically significant (p-adj<0.05). The 32 type I hits (i.e. Neuron Screen drugs modulating C4R neurogenetic disease genes) are represented by a heatmap (**Fig. 3.2**); hits that did and did not validate are shown by a positive or negative sign, respectively. The type I hits were first investigated at the transcriptional level (using qRT-PCR) in 3H cortical cultures. The list of primers used for all qRT-PCR assays in mouse cells and tissue is available in **Appendix II**. The same drug dose and time (8 hours) that was used for the screen was used for initial validation. Of the 32 hits, 6 showed a statistically significant upregulation compared to control (p<0.05) and one hit showed a trend toward upregulation. Three of the hits that were validated at the transcriptional level, aminophylline-*Aldh18a1*, nilotinib-*Sacs*, diflunisal-*Slc6a8*, involved drugs that were among the top 5 transcriptionally active drugs (**Fig. 3.3A-C**). The other four hits that showed upregulation appear to be class effect hits, with both thyroid hormone analogs upregulating the gene *Pmp22* (**Fig. 3.3D**) and both corticosteroid drugs upregulating the gene *Hsd17b4* (**Fig. 3.3E**).

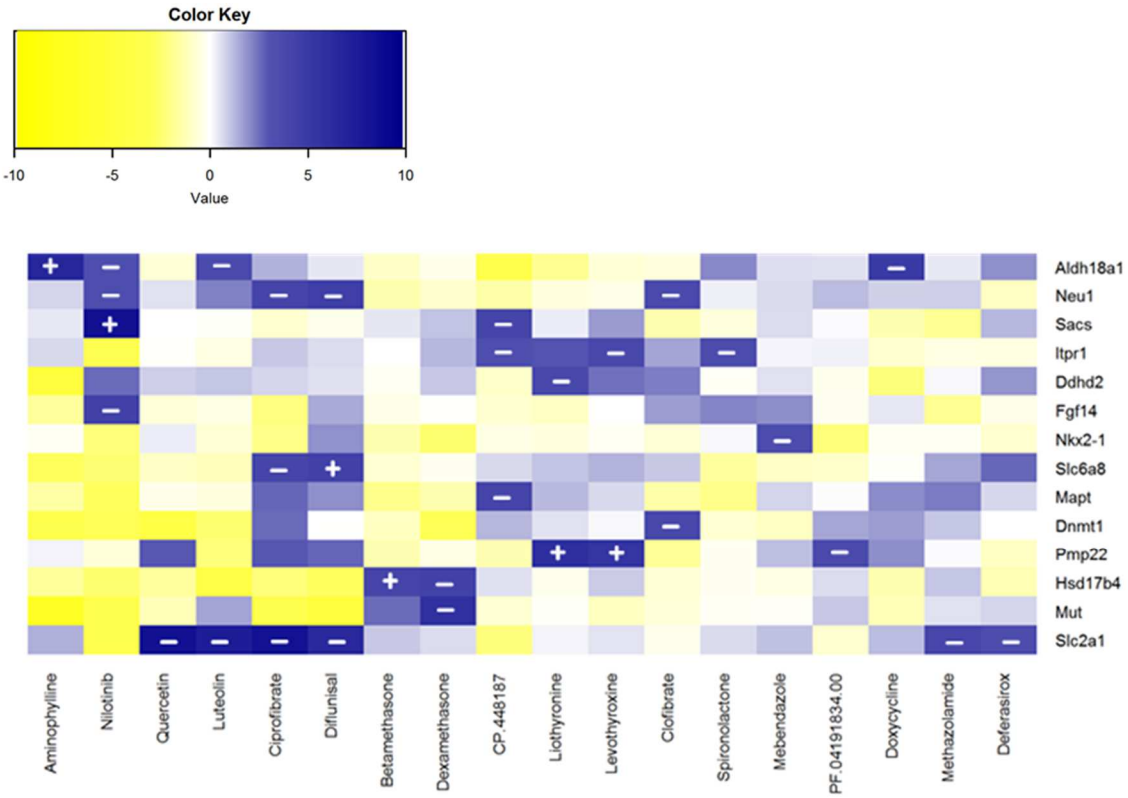

Figure 3.2: Quantitative RT-PCR investigates the C4R neurogenetic type I drug-gene hits in mouse 3H cortical cultures. Transcriptional validation was performed in 3H mouse cortical cultures at DIV21 using qRT-PCR. All treatments were performed for 8 hours and were performed at the same dose as the Neuron Screen. The heatmap representing the 32 drug-gene hits that were included for validation. Hits with a (+) were statistically significant ($n \geq 5$, $p < 0.05$) and hits with a (-) were not significant ($n \geq 2$, $p \geq 0.05$). Statistical analysis was performed using student's paired two-tailed t-test ($^{*}p < 0.05$, $^{\#}p < 0.01$).

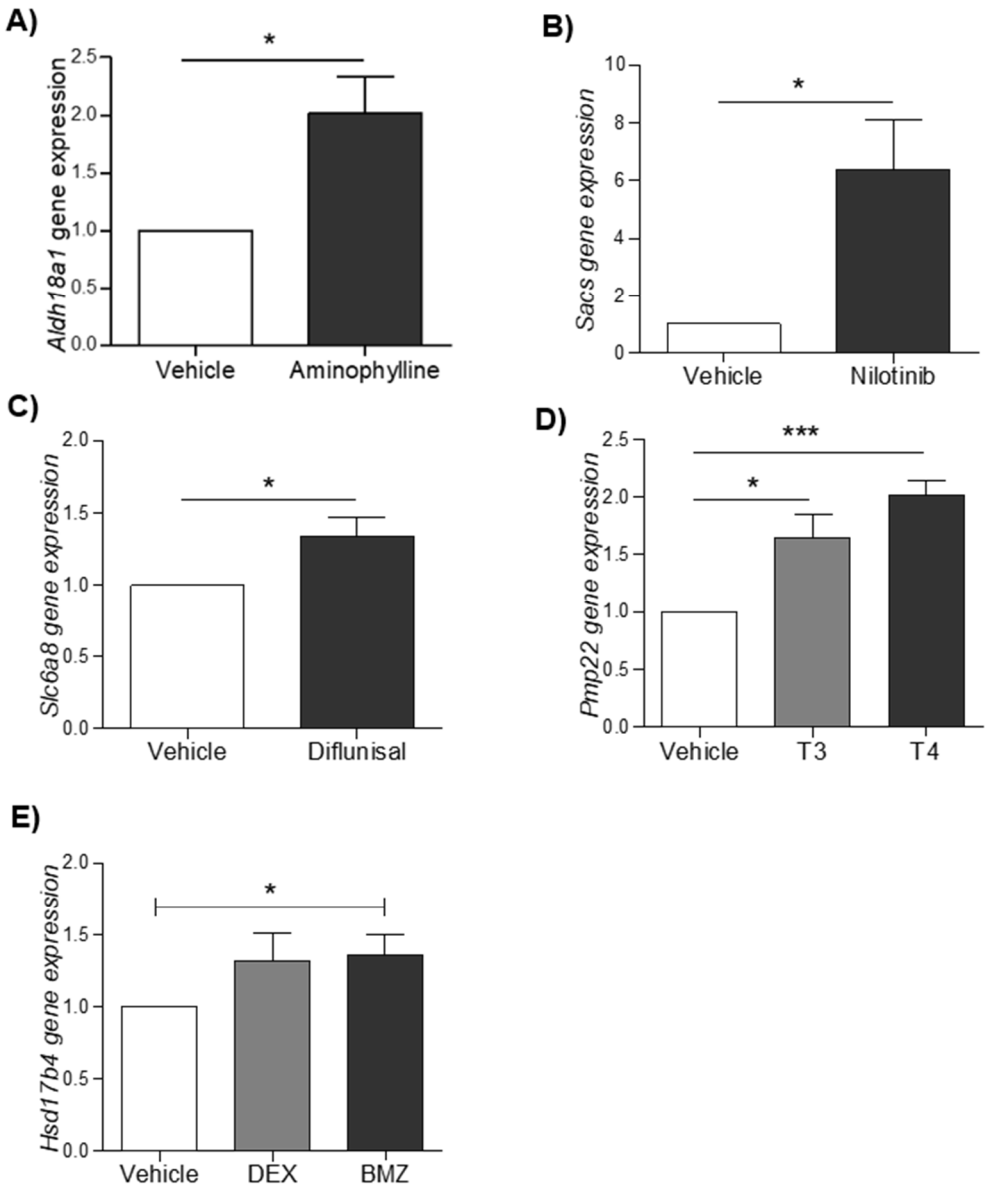

Figure 3.3: Graphical representation of the six C4R neurogenetic type I drug-gene hits that validated in mouse 3H cortical cultures. ***A)*** Aminophylline (90μM) upregulates *Aldh18a1* (n=7, *p=0.018). ***B)*** Nilotinib (10μM) upregulates *Sacs* compared to DMSO treated control (n=8, *p=0.017). ***C)*** Diflunisal (200μM) upregulates *Slc6a8* (n=7, *p=0.04). ***D)*** T3 (0.01μM) and T4 (0.2μM) upregulate *Pmp22* (n=5, *p=0.029) and (n=7, ***p=0.001) respectively. ***E)*** BMZ (0.4μM) and DEX (0.5μM) upregulate *Hsd17b4* (n=7, *p=0.049) and (n=6, n.s.) respectively.

Transcriptionally validated drug-gene hits show three different patterns of dose-curve response

To determine if the validated drug-gene interactions were indicative of true biological effect, half-log drug dose curves were performed in 3H cortical cultures for aminophylline, nilotinib, T4, and diflunisal, all drugs which mediate transcriptionally validated C4R hits. For aminophylline, nilotinib and T4, 3H cultures were treated with the validated Neuron Screen dose as well as three lower doses and one higher dose. Four lower doses were used for diflunisal, given the screen dose (200 μM) was already very high. The aminophylline-*Aldh18a1* hit showed a dose-dependent increase in gene expression (**Fig. 3.4A**). The 10μM dose of Nilotinib led to a statistically significant increase in *Sacs*, but there was not a dose-dependent increase (**Fig. 3.4B**). The 31.6μM dose of nilotinib led to downregulation of the normalizing genes, and overt cellular toxicity (determined by bright-field microscopy). Although not statistically significant, the five doses of T4 upregulated *Pmp22* in a seemingly dose-independent manner (**Fig. 3.4C**) while only the highest dose (200μM) of diflunisal upregulated *Slc6a8* (**Fig. 3.4D**).

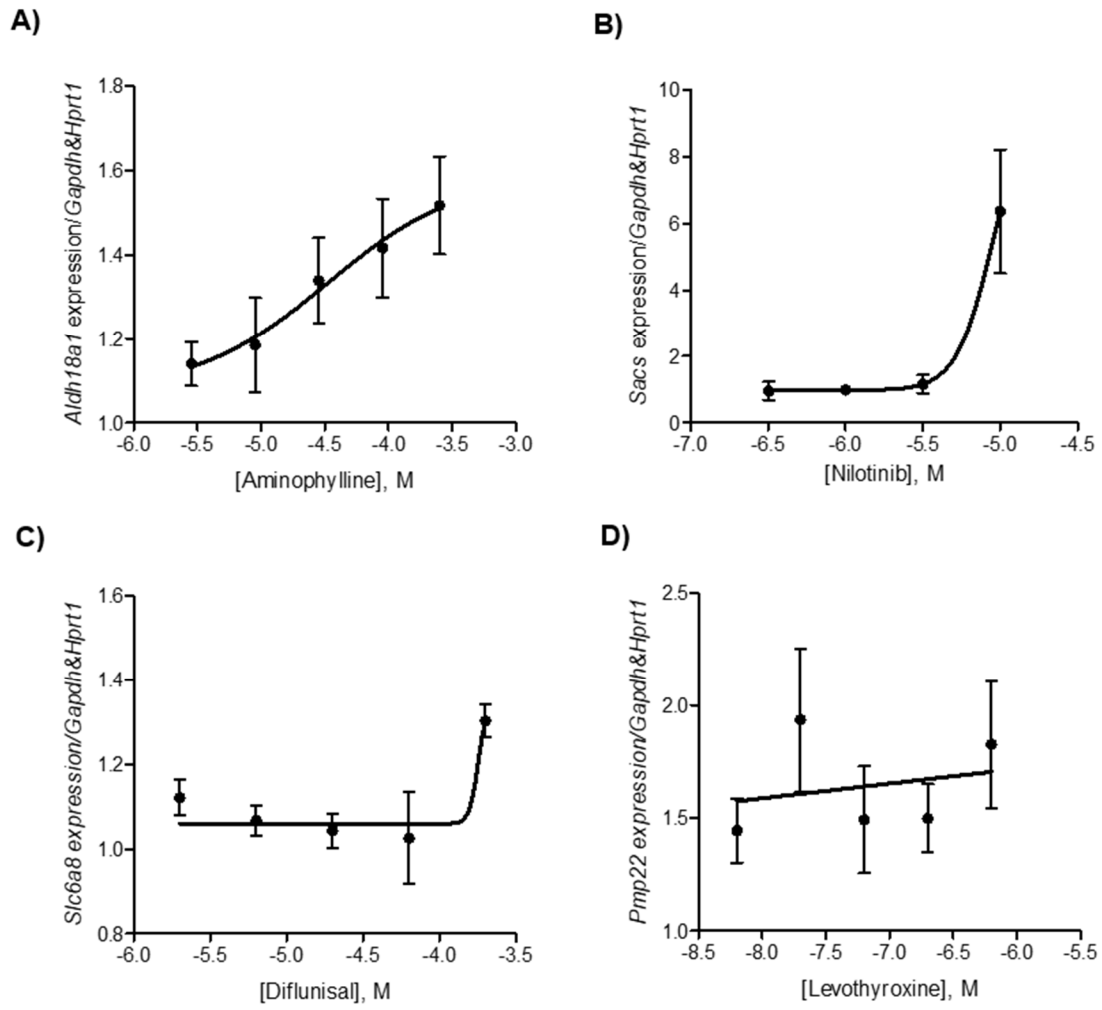

Figure 3.4. Dose-curve study of four validated C4R drug-gene hits in mouse cortical cultures. Primary mouse cortical neurons DIV16-DIV21 were treated with a half-log distribution of each drug for 8 hours. qRT-PCR was used to conduct differential expression of target genes. Each data point represents mean differential expression (± SEM) plotted against the $\log_{10}$([drug]). Statistical significance was computed by one-way ANOVA. Data fit to a non-linear regression curve for ***A)*** Aminophylline-*Aldh18a1,* (n=4) **B)** Nilotinib-*Sacs* (n=3, $p<0.01$). **C)** Diflunisal-*Slc6a8* (n=3). Data fit to a linear regression curve for **D)** T4-*Pmp22* (n=3).

High-dose fenofibrate and nilotinib are cytotoxic in 3H cortical cultures

Because the highest dose of nilotinib (31.6μM) was toxic to cortical cultures, a lactate dehydrogenase (LDH) assay was used to screen for cytotoxicity in 3H cultures at all doses used for the dose-curve experiment. The drug shown to have the broadest transcriptional impact in the screen, fenofibrate, unfortunately also conferred morphologically visible toxicity in 3H cortical cultures; for this reason, although hits from fenofibrate were not included in the validation pipeline, the drug was included in the LDH assay as a positive control. The LDH assay was performed on 3H cultures treated with a dose-curve of fenofibrate, nilotinib, aminophylline, diflunisal, and T4. The LDH assay, performed 8 and 14 hours after drug treatment, showed a dose-dependent cytotoxic effect of fenofibrate with a statistically significant increase in cytotoxicity at the 80μM dose (**Fig. 3.5A**). For the four drugs associated to type I drug-gene hits (nilotinib, aminophylline, diflunisal, T4), only the highest dose of nilotinib, half-log higher than the Neuron Screen dose, led to a statistically significant increase in cytotoxicity after 8 and 14 hours of treatment. (**Fig. 3.5B, Fig. S3.5**).

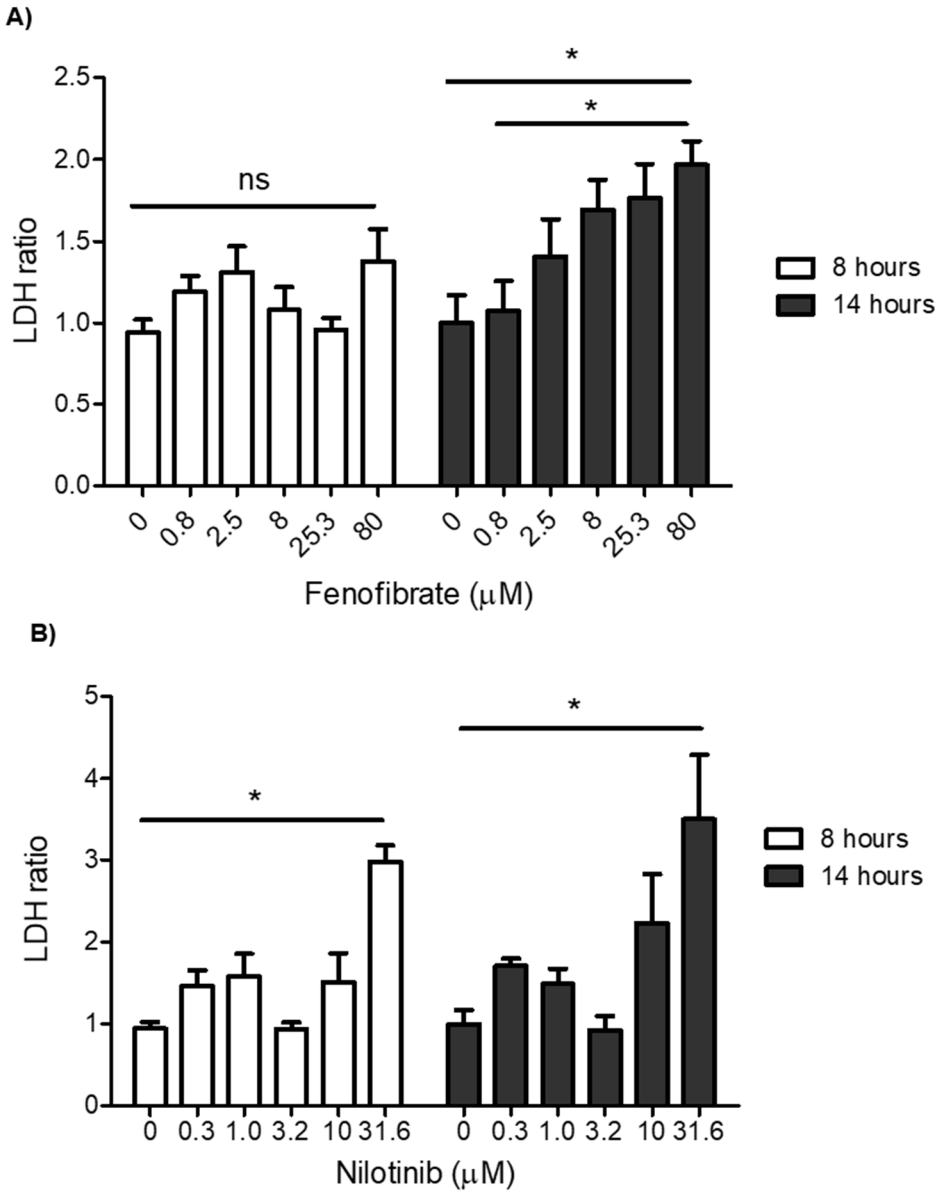

Figure 3.5. Cytotoxicity of the top two transcriptionally active Neuron Screen drugs in 3H cortical cultures. Primary mouse cortical neurons at DIV21 were treated with a half-log distribution of each drug for 8 or 14 hours. Cytotoxicity was subsequently assessed by LDH assay. Statistics were performed using one-way ANOVA and Tukey's post-hoc test. ***A)*** Fenofibrate (n=4, *p=0.009). **B)** Nilotinib (n=4, *p=0.005).

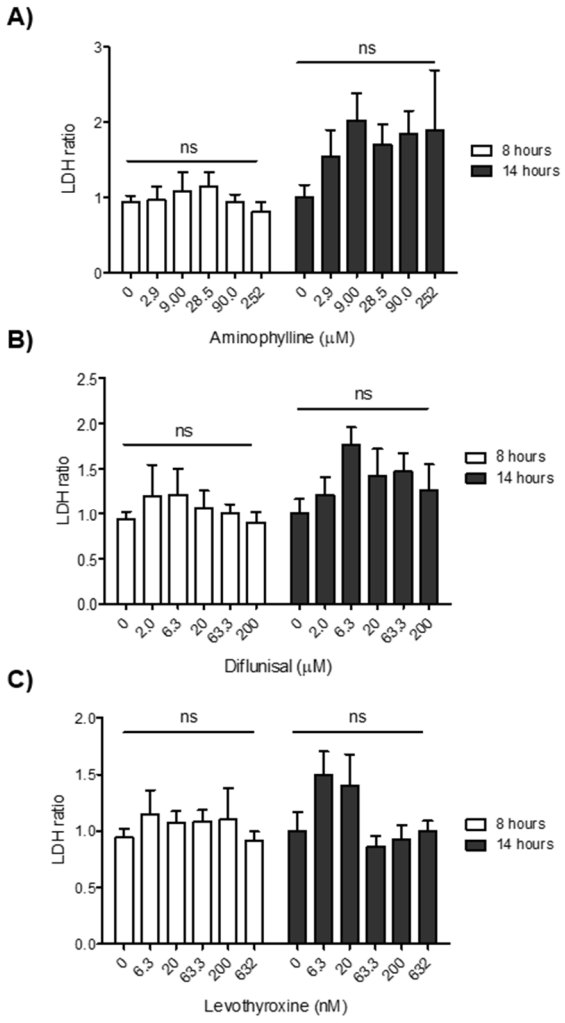

Figure S3.5. Lactate dehydrogenase assay (LDH) for aminophylline, diflunisal, levothyroxine. Primary mouse cortical neurons at DIV21 were treated with a half-log distribution of each drug for 8 or 14 hours. Cytotoxicity was subsequently assessed by LDH assay. Statistics were performed using one-way ANOVA and Tukey's post-hoc test. ***A)*** Aminophylline dose-curve. ***B)*** Diflunisal dose-curve. ***C)*** T4 dose-curve. ns = not statistically significant ($p>0.05$).

Drug-gene hits fail to validate in human immortalized neuron-like cells

In humans, mutation in *ALDH18A1* and mutation in *SACS* lead to the hypomorphic conditions autosomal recessive cutis laxa type IIIA and autosomal recessive spastic ataxia of Charlevoix-Saguenay respectively (ARSACS). Since both conditions have CNS symptoms, human SH-SY5Y cells were used to investigate the two hits related to the disease-causing genes that validated in cortical cultures. An aminophylline time-course and dose-curve was performed in differentiated and undifferentiated SH-SY5Y cells while a nilotinib dose-curve was performed in undifferentiated SH-SY5Y cells and in ARSACS patient fibroblasts. Quantitative RT-PCR was used to determine differential expression of *ALDH18A1* and *SACS* in response to the dose curves. The list of primers used for all qRT-PCR assays in human cells and tissue is available in **Appendix III**. There was no observed change of gene expression in response to nilotinib or aminophylline in the human cell lines. Also, nilotinib did not induce *Sacs* gene expression in the human primary fibroblasts (data not shown).

Thyroid hormone analog T3 upregulates *Pmp22* in rat dorsal root ganglion (DRG) cultures

Because peripheral nervous system (PNS) myelin is the primary site of pathology for the PMP22 haploinsufficient focal sensory motor neuropathy HNPP, dose-curves of T4 and T3 were performed in rat schwannoma RT4 cells and a T3 dose curve was performed in primary rat dorsal root ganglion (DRG) cultures. The list of primers used for all qRT-PCR assays in rat cells and tissue is available in **Appendix IV**. The RT4 cells did not show any *Pmp22* transcriptional response to T3 or T4 treatment after 8 and 24 hours of treatment. Rat DRG cultures were grown for 28 days *in vitro* to suitable maturity; 7 days without treatment and 21 days with either 0, 10, 31.6, or 100nM T3 treatments performed 3-4 times per week in tandem with media changes. Staining of DIV28 DRG cultures for the neuronal β-III tubulin (TuJ-1) protein showed highly interconnected sensory neurons (**Fig. 3.6A**). Quantitative RT-PCR results showed upregulation of *Pmp22* mRNA in DIV28 rat DRG cultures at the three doses of T3, achieving statistically significant induction at 100nM (**Fig. 3.6B**).

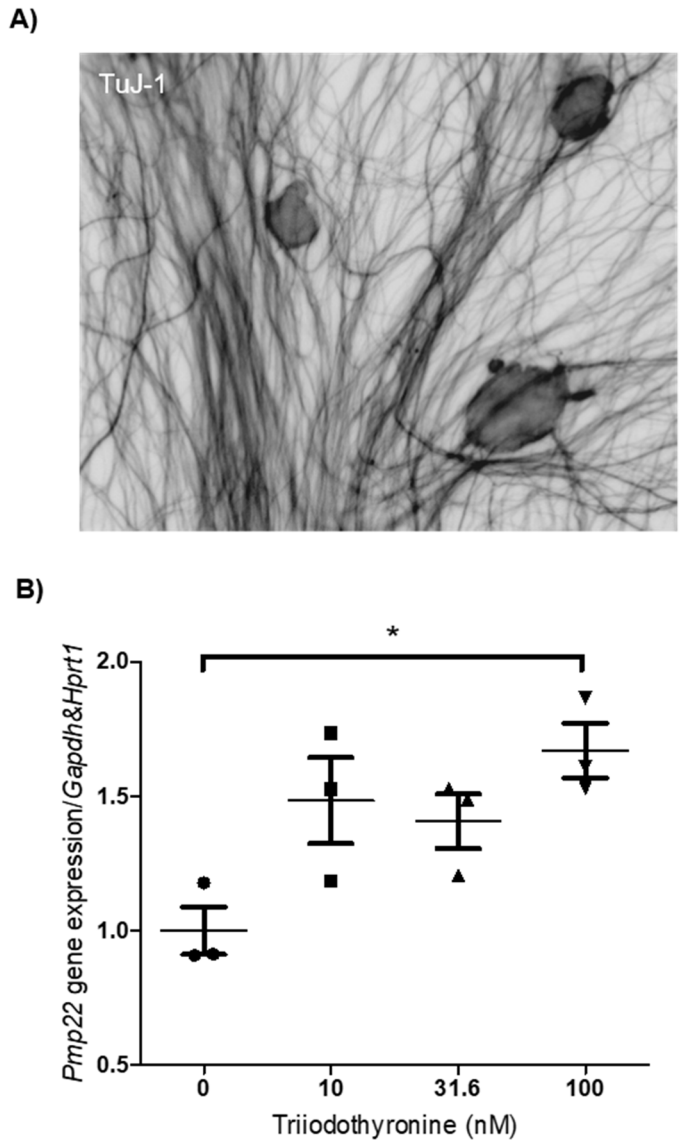

Figure 3.6. Effect of T3 on *Pmp22* transcript levels in rat dorsal root ganglion (DRG) cultures. Rat DRG cultures were grown *in vitro* for 7 days and then for an additional 21 days in the presence of vehicle (0.1% DMSO) or 10, 31.6, 100nM of T3. **A)** IF staining of DRG cultures at DIV28 for β-III tubulin (TuJ-1) (10x magnification). **B)** qRT-PCR shows the transcriptional effect of 21-day T3 dose-curve on *Pmp22* transcript levels (n=3). Statistical significance was measured by one-way ANOVA (non-parametric) with Tukey post-hoc analysis (*$p<0.05$).

The aminophylline-*Aldh18a1* and nilotinib-*Sacs* hits fail to validate at the protein level

The impact of the aminophylline-*Aldh18a1*, nilotinib-*Sacs*, and T4/T3-*Pmp22* pairings on protein level in primary neuronal cultures was next determined, exploring whether transcript upregulation translated to protein upregulation. Total soluble protein was collected from mouse 3H cultures treated with aminophylline or nilotinib, and western blotting was conducted with ALDH18A1 and Sacsin specific antibodies; unfortunately, no alteration in ALDH18A1 and Sacsin protein levels in response to aminophylline and nilotinib was observed (**Fig. 3.7**).

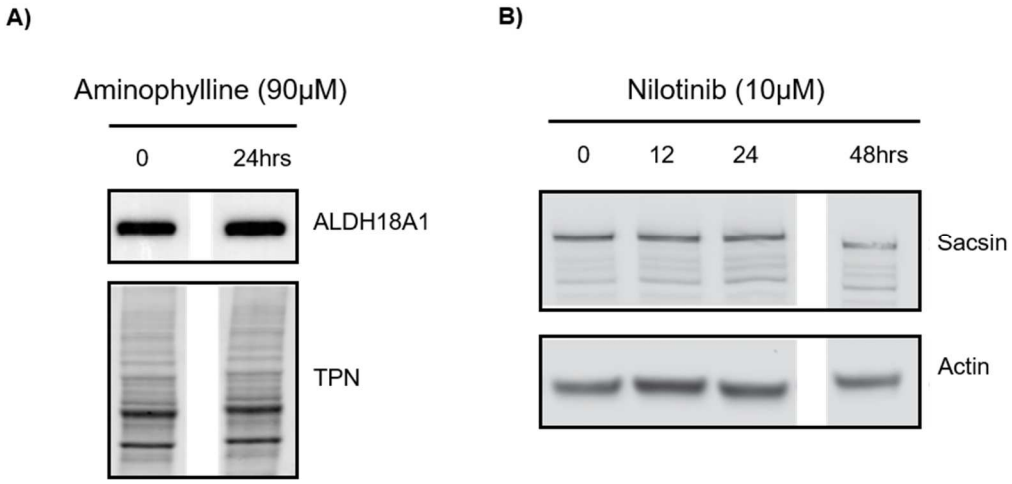

Figure 3.7. Protein expression of transcriptionally validated type I drug-gene hits in 3H cortical cultures. A) Cultures were treated with 90μM aminophylline for 24 hours. Representative western blot (non-contiguous bands from the same membrane) is shown for ALDH18A1 and total protein stain (TPN) (n=2). **B)** Cultures were treated with 10μM Nilotinib for 12, 24, 48 hours. Representative western blot (48hr treatment is non-contiguous but exposed on the same membrane) is shown for Sacsin and the house-keeping protein β-actin (n=2). Note: Sacsin blot was performed by Roxanne Lariviere (MNI, Montreal).

Liothyronine treatment increases expression of PMP22 protein in DRG cultures

Because PMP22 protein is not detectable in 3H cultures, the impact of T4/T3 on rat DRG Pmp22 protein levels was studied. Rat DRG co-cultures of peripheral sensory neurons and Schwann cells can be grown *in vitro* to express PMP22 protein in compact myelin (Notterpek et al. 1999; Nobbio et al. 2006; Callizot et al. 2011). Rat DRGs were cultured in the same way as for the transcriptional study of PMP22 (**Fig. 3.6**) and treated with the same dose curve of T3 for 21 days. Expression of PMP22 protein in response to T3 treatment was studied by western blot and IF staining. SDS-PAGE western blot analysis of soluble and insoluble protein fractions from DIV28 DRG cultures failed to reveal PMP22 immunoreactivity. Immunofluorescence staining was also too faint to detect Pmp22 expressed in myelin. However, IF staining for PMP22 in DIV28 DRG cultures showed the presence of cells with perinuclear staining pattern of PMP22 typical of non-myelinating Schwann cells (**Fig. 3.8A**) (Notterpek et al. 1999). DRGs treated with T3 showed more intensely stained PMP22-positive cells compared to vehicle treated cultures. However, signal intensity of PMP22 could not be measured due to high background levels from debris. Blinded counts of PMP22 stained cells and normalized against total cell counts (DAPI) showed a trend toward upregulation in the T3-treated cultures compared to control (**Fig. 3.8B**).

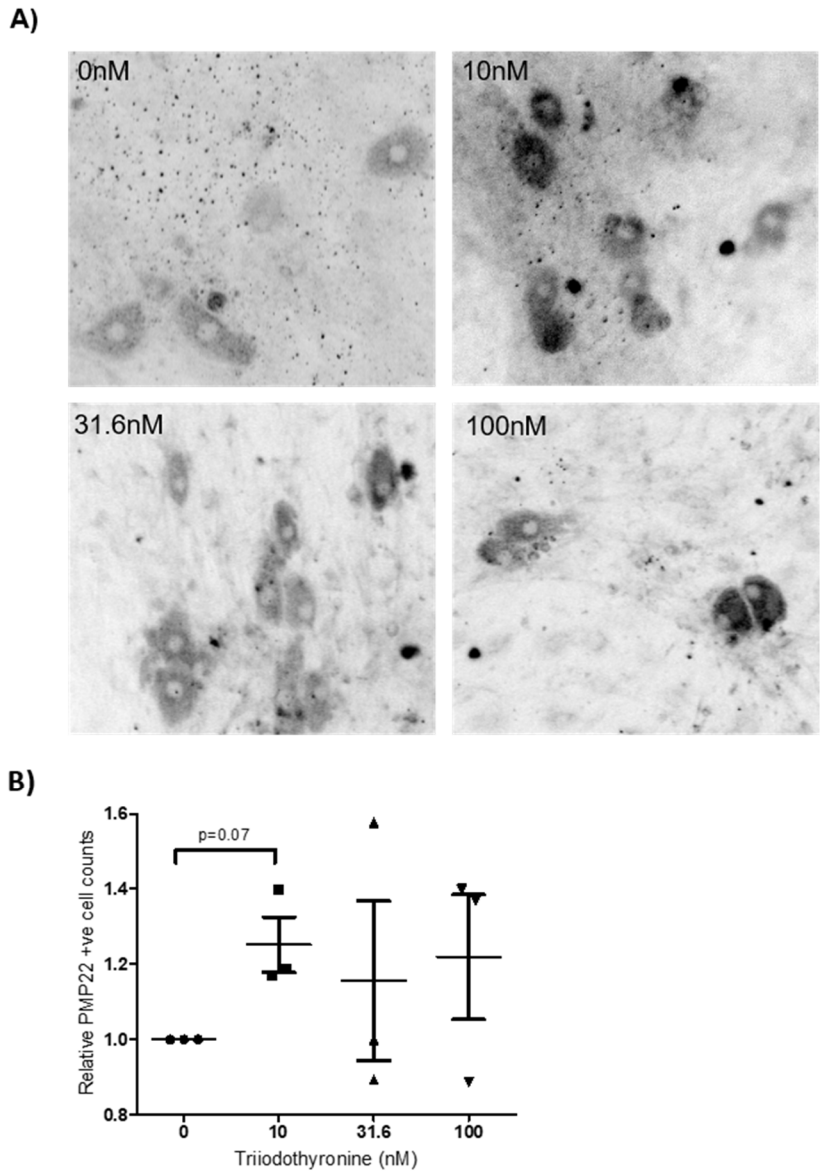

Figure 3.8. Validation of thyroid hormone effect on PMP22 protein in rat dorsal root ganglion (DRG) cultures. Rat DRG cultures were grown *in vitro* for 7 days and then for an additional 21 days in the presence of vehicle (DMSO) or 10, 31.6, 100nM of T3. **A)** Immunofluorescence staining of DIV28 DRG cultures for Pmp22 shows increased staining at 31.6nM T3 compared to DMSO treated (10x magnification). **B)** Whole-well cell counts (96-well format) of Pmp22 positive cells using ImageJ cell counter tool for quantification (n=3).

***In vivo* pilot studies of type I drug-gene hits link thyroid hormone activity and *Pmp22* upregulation in rat cortex.**

The drug-gene hits that were transcriptionally validated in cell cultures were next tested in the *in vivo* setting. Since transgenic models for the rare neurogenetic diseases in question are not commercially available, an *in vivo* pilot study of each type I hit was performed in wild-type mice (C57BL6) or wild-type rats (Sprague-Dawley). The goal of the pilot study was to triage type I hits to reduce the rate of failure in the costlier pre-clinical studies in transgenic animals. The following drug-gene hits were investigated *in vivo* using adult male C57BL6 mice: nilotinib/*Sacs*, aminophylline/*Aldh18a1*, DEX/*Hsd17b4*, T4/*Pmp22*, and diflunisal/*Slc6a8*. The T4/*Pmp22* hit was also investigated in adult Sprague Dawley (SD) rats. Mice and rats were treated by intraperitoneal (i.p.) injection or oral gavage (p.o.) for various periods of time with drug doses equivalent to the highest tolerated dose in humans (**Table 3**). For each drug-gene interaction, bilateral cortex and cerebellum were collected ~4 hours after the last dose was administered. For the T4/*Pmp22* hit, bilateral sciatic nerve was also collected.

Quantitative RT-PCR was used to analyze gene expression in each tissue relevant to that gene. Treatment of SD rats with daily intraperitoneal injections of 300μg/kg for 8 days led to upregulation of *Pmp22* in cortex of one rat. After addition biological replicates (three vehicle and six T4 treated rats) there was a ~25% statistically significant upregulation of *Pmp22* in T4 treated rat cortex (**Fig. 3.9A**). Furthermore, three well-characterized thyroid hormone-responsive genes – i.e. Hairless (*Hr*), Sonic hedgehog

(*Shh*), and Semaphorin 7a (*Sema7a*), were also upregulated (>2-fold) in the cortex of T4 treated rats confirming effect of thyroid hormone on the CNS (F**ig. 3.9B**). None of the other type I hits from this *in vivo* pilot study showed statistically significant change (**Table 3**). In contrast, *Pmp22* and the thyroid hormone-responsive genes were not upregulated in any of the mice treated with 10μg/kg. Cortex of mice treated with DEX were also positive for the DEX-responsive gene *Fkbp5* but expression of the type I hit *Hsd17b4* did not appear to be induced by DEX (**Fig. 3.9C-D**).

Table 3. Summary of *in vivo* experiments with transcriptionally validated drug-gene C4R hits. Adult mice (C57BL6) or rats (Sprague-Dawley) were treated for various time-points with the indicated drugs. Animals were treated once daily. Mice were sacrificed by isoflurane inhalation and cervical dislocation ~4hours after administration of the most recent treatment dose. All tissues were collected in bilateral fashion. Transcript response results of each target gene was determined by qRT-PCR performed on total RNA extracted from each type of tissue.

Drug	Dose (qd)	Administration Route	Animals/n	Tx length (days)	Tissues	Gene	Response (qRT-PCR)
Levothyroxine	10µg/kg	i.p.	C57BL6/3	8	Ctx, Ce	Pmp22	ns
	300µg/kg	i.p.	SD rats/6,9*	8	Ctx, Ce, ScN	Pmp22	↑
Aminophylline	10mg/kg	i.p.	C57BL6/4	8	Ctx, Ce	Aldh18a1	ns
Dexamethasone	1mg/kg	p.o.	C57BL6/3	5	Ctx, Ce	Hsd17b4	ns
Diflunisal	50mg/kg	p.o.	C57BL6/4	15	Ctx, Ce	Neu1, Slc6a8	ns
Nilotinib	25mg/kg	p.o.	C57BL6/4	8	Ctx, Ce	Sacs	ns

Abbreviations: qd (once a day), n (number of mice per treatment), Tx (treatment), SD (Sprague-Dawley), Ctx (cortex), Ce (cerebellum), ScN (sciatic nerve), ns (not significant change)

*one of the T4 treated rats died 3 days after start of treatment and was excluded from the trials.

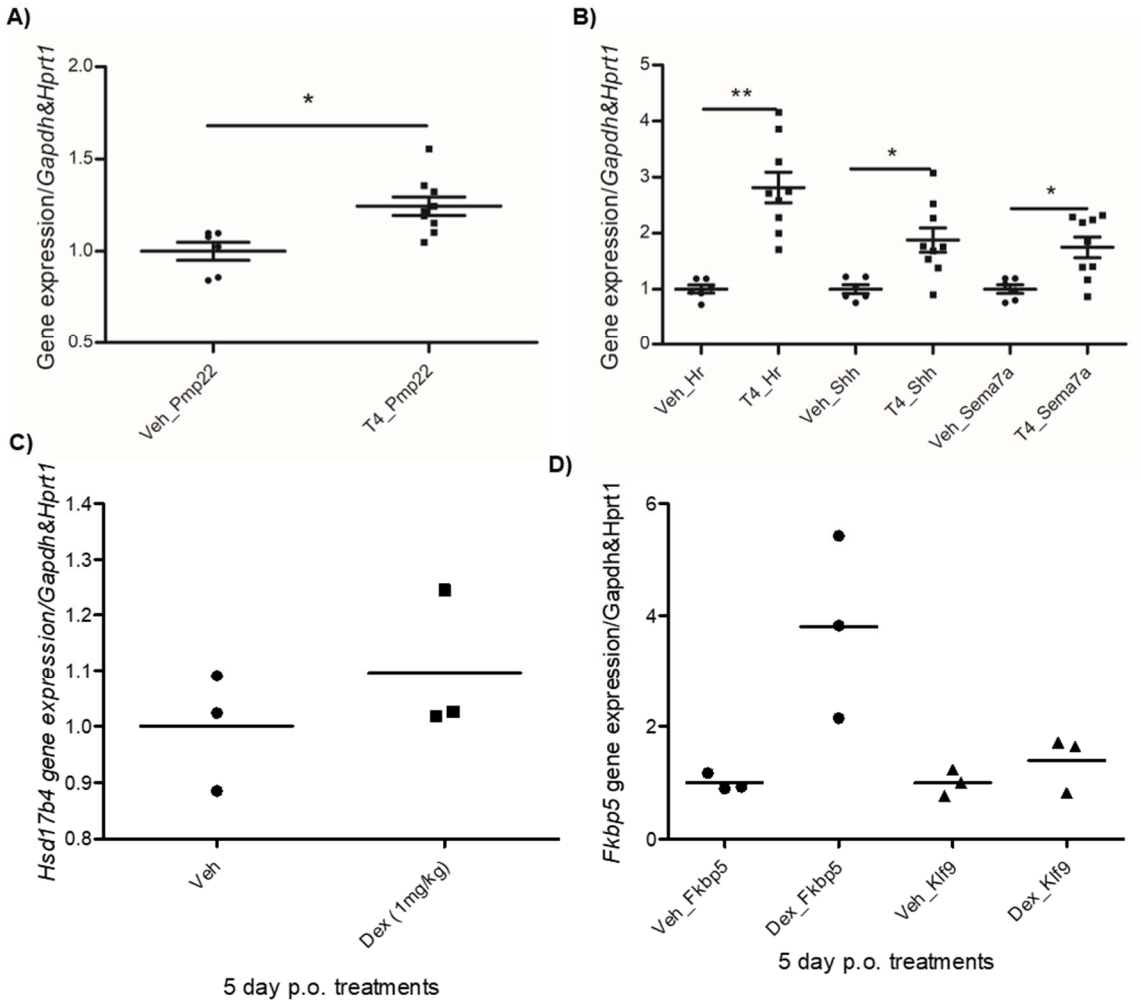

Figure 3.9. Effect of T4 on *Pmp22* and DEX on *Hsd17b4* in cortex of wild type rodents. **A,B)** Male Sprague-Dawley (SD) rats were treated with daily i.p. injections of vehicle (5% DMSO in PBS) (n=6) or T4 (300μg/kg) (n=9) for eight days. qRT-PCR was used to determine the differential expression of *Pmp22*, *Hr*, *Shh*, and *Sema7a* in response to T4 treatment in cortex. Statistical analysis performed by unpaired t-Test (n=6-8) *p<0.05, *p<0.01. **C,D)** Male C57BL6 mice were treated with daily p.o. doses of vehicle or DEX (1mg/kg) for five days. qRT-PCR was used to determine the differential expression of *Hsd17b4*, *Fkbp5*, and *Klf9*, in response to DEX treatment in cortex.

CHAPTER 4

Upstream regulator analysis of Neuron Screen data reveals novel approach to identifying drug-gene interactions with potential implications for the treatment of sickle cell disease and a rare form of microcephaly

Repurposing candidate identification – network-directed approach

A high false discovery rate plagued individual drug-gene hits that were identified from transcriptomic data (Mears et al. 2017). To address this issue, ingenuity pathway analysis (IPA) was used to develop a network-directed approach to identifying drug-gene hits from the Neuron Screen. IPA uses causal reasoning algorithms to determine the upstream biological causes and probable downstream effects of observed transcriptional changes (Krämer et al. 2014). Thus, IPA is similar to gene-set enrichment analysis platforms like DAVID (Huang et al. 2007), with the added benefit of a directional study of each pathway. Almost 5 million manually curated findings from peer-reviewed sources form the Ingenuity Knowledge Base enable the causal reasoning algorithms used to establish interactions between genes, disease states, drugs, etc. (Krämer et al. 2014). When transcriptomic data is uploaded to the IPA software, IPA compares the experimental differential expression data (e.g. Neuron Screen) to the IPA knowledge base. The IPA upstream regulator analysis (URA) tool compares experimental data with IPA knowledge-base data to develop hypotheses related to the activation or inhibition of gene networks in the experimental data. Based on the activation or inhibition of gene networks, URA identifies a common upstream regulator (e.g. transcription factor, kinase, drug) that could be responsible for modulation of a given gene network. The URA tool computes the statistical enrichment of genes involved in a network and if they are modulated in a direction that is consistent with the overall trend of the network. If most genes in a network are upregulated it is labeled "activated"; conversely if most genes are downregulated, it is labeled "inhibited".

Statistical significance of network modulation is defined by $|z|>2$ (i.e. Z-score > 2 for an activated network; Z-score < -2 for an inhibited network.

To assess the efficiency of URA networks in identifying drug-gene hits, transcriptome profiles of the 50 Neuron Screen drugs that modulated the greatest number of transcripts were analyzed by IPA and drug "activated" and "inhibited" networks were identified by URA. Statistically significant URA networks ($|z|>2$) were defined as "robust" if they contained a minimum of five downstream genes that were differentially expressed in the Neuron Screen (p-adj<0.05). Robust networks were chosen for validation if the drug that caused the network modulation was used to treat a rare genetic disease, or if a robust network included at least one gene associated to a rare genetic disease. Downstream genes that were part of robust URA networks that were modulated by a drug of interest were called "type II" drug-gene hits. Type II drug-gene hits were validated transcriptionally in human immortalized cells by qRT-PCR and western blot analysis. Type II hits were also validated in cortex of wild-type mice. A summary of the proposed type II hit validation protocol is presented in **Figure 4.1**, with accompanying number of hits studied at each validation stage.

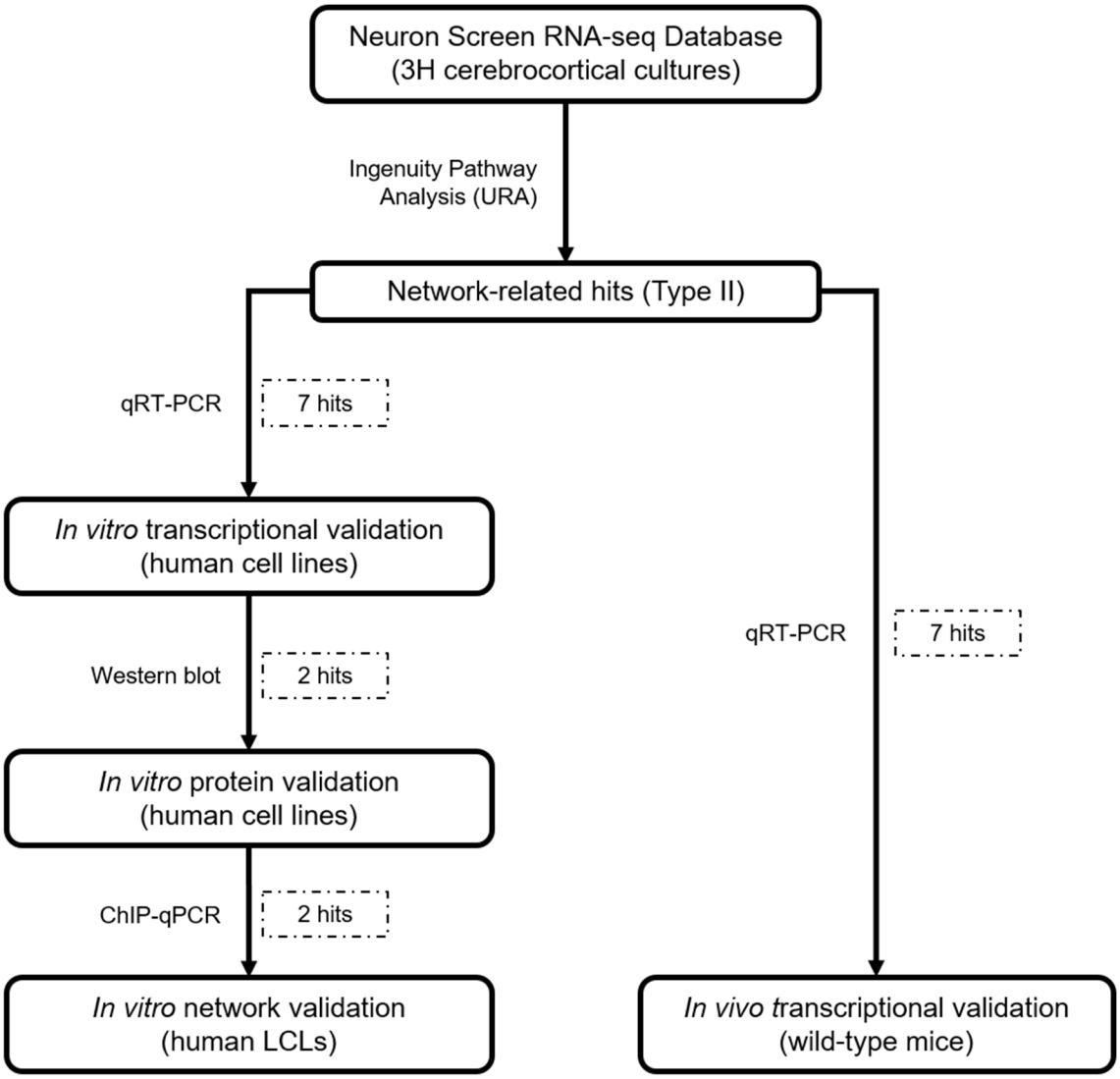

Figure 4.1. Diagram of validation strategy for Neuron Screen Type II drug-gene hits.

Dashed boxes identify the number of hits that were studied at each stage of validation.

Hydroxyurea inhibits the FOXM1 network while dexamethasone activates the PPARD network

To determine the validation rate for type II drug-gene hits, two robust URA networks were identified based on the statistical criteria previously outlined: the forkhead box protein M1 (FOXM1) network was downregulated (inhibited) by hydroxyurea (HU) and the peroxisome proliferator activated receptor delta (PPARD) network was upregulated (activated) by DEX in the Neuron Screen differential expression data. Both networks had a $|z|$-score > 2 and both networks had more than five downstream genes with a p-adj<0.05. Statistical metrics for the FOXM1 and PPARD networks are shown along with several other robust networks that could be of interest for future validation (**Table 4**). The T4-DIO3 network is shown as an example of an upstream regulator (iodothyronine deiodinase 3) with downstream genes that are well-documented targets of thyroid hormone; thus, it is a positive control for identification of robust URA networks.

Table 4. Summary of six robust upstream regulator analyses identified in the Neuron Screen data. The first row shows the well-described thyroid hormone effect on the DIO3 pathway. The last five rows show examples of novel and untested drug-upstream regulator interactions that have similar statistical metrics to the T4-DIO3 network.

Drug name	Upstream regulator	Expression Fold Change	Activation state	Activation z-score	P-value of overlap	# of genes
Levothyroxine	DIO3*	4.385	Inhibited	-2.975	1.79E-07	8
Hydroxyurea	FOXM1	-3.579	Inhibited	-3.245	1.63E-10	8
Dexamethasone	PPARD	2.522	Activated	3.126	4.58E-02	10
Dexamethasone	STAT4	0	Activated	2.933	7.36E-04	12
Vigabatrin	MKNK1	1.187	Inhibited	-3	1.40E-04	5
Pregabalin	PGR	1.013	Activated	3.376	7.77E-09	13

*Validated in literature

Hydroxyurea downregulates FOXM1 downstream genes in human glioblastoma cells

The FOXM1 network, downregulated by HU treatment, was used to determine the rate of *in vitro* validation of type II hits – i.e. drug-gene hits that are identified as part of a robust URA network. This FOXM1 transcription factor is a master regulator of genes involved in every stage of the cell cycle and plays a crucial role in activation of genes that detect and repair DNA damage (Wierstra 2013). Although the effect of HU in causing DNA damage by depletion of dNTPs has been extensively studied, its role in the transcriptional modulation of the FOXM1 pathways has not been well described. The IPA analysis showed that eight FOXM1-target genes involved in G2 (growth phase 2) and M (mitosis phase) of cell cycle progression are downregulated in response to HU treatment in mouse cortical cultures; blue arrows indicate literature findings of downregulation while green-colored genes indicated downregulated genes in response to HU treatment in the Neuron Screen (p-adj<0.05) (**Fig. 4.2A**). To validate the effect of HU on the FOXM1-associated type II hits, human U87 glioblastoma cells were treated for 0, 4 and 8 hours with 250μM HU. Quantitative RT-PCR revealed HU-mediated decrease (minimum significance of $p<0.05$) of seven of the eight FOXM1-network genes at 4 and/or 8 hours (*TOP2A* was not included for validation) (**Fig. 4.2C-E, Fig. S4.2**). Surprisingly, the level of upstream regulator gene *FOXM1* was not significantly affected by 4 or 8 hours of HU treatment (**Fig. 4.2B**).

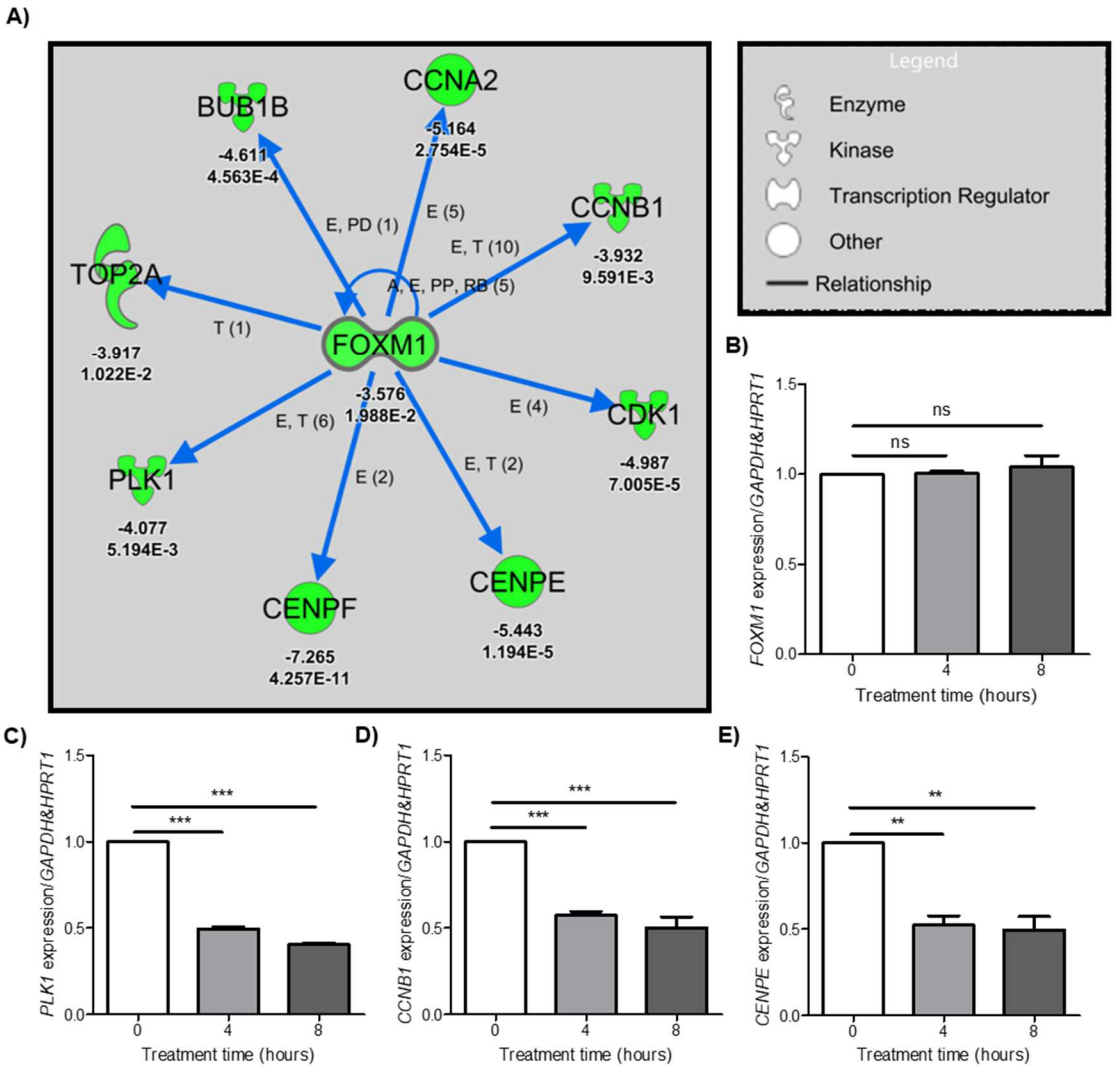

Figure 4.2. Validation of the effect of HU on the FOXM1 URA network. ***A)*** The FOXM1 network identified by URA analysis is inhibited in HU-treated cortical cultures (Neuron Screen data). Blue arrows signify inhibition and green symbols signify downregulated genes (Z-score and p-adj appear directly below each symbol). U87 glioblastoma cells were treated for 0, 4, and 8 hours and qRT-PCR was employed to determine gene expression of **A)** *FOXM1* **(B)** and four of its transcriptional targets identified in the URA. **(C)** Polo-like kinase 1 (*PLK1*). **(D)** Cyclin B1 (*CCNB1*). **(E)** Centromere protein E (*CENPE*). Statistical significance was measured by one-way ANOVA (non-parametric) with Tukey post-hoc analysis (**p<0.01, ***p<0.001, ns=not significant).

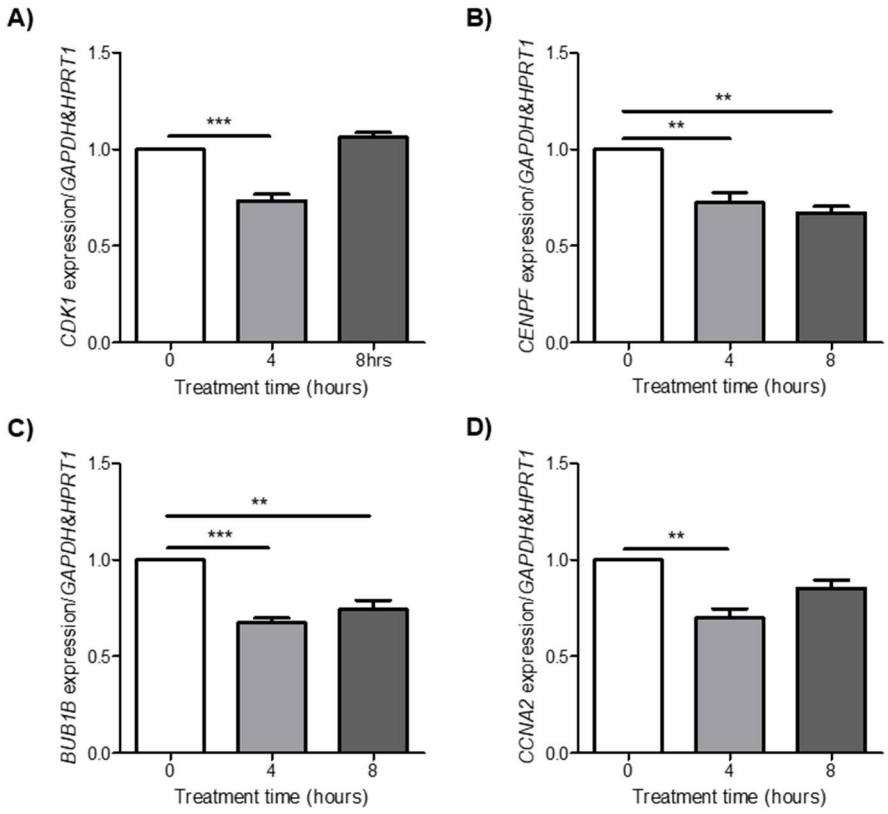

Figure S4.2. Validation of the effect of HU on four FOXM1 downstream genes. U87 glioblastoma cells were treated for 0, 4, and 8 hours and qRT-PCR was employed to determine gene expression of **A)** Cyclin-dependent kinase 1 (*CDK1*) **(B)** Centromere protein F (*CENPF*) **(C)** BUB1 mitotic checkpoint serine/threonine kinase (*BUB1*) **(D)** Cyclin A2 (*CCNA2*). Statistical significance was measured by one-way ANOVA (non-parametric) with Tukey post-hoc analysis (**p<0.01, ***p<0.001, ns=not significant).

Hydroxyurea induces a dose-dependent downregulation of *PLK1* and *CCNB1*

A dose finding study of HU impact on FOXM1-target gene mRNA and protein in U87 cells was undertaken, screening for a dose-dependent relationship and to identify the concentration for optimal gene and protein downregulation while maintaining cellular viability. U87 cells treated for eight hours with HU were divided into cell pellets for RNA and protein extraction; qRT-PCR of RNA revealed a dose-dependent downregulation of *PLK1* and *CCNB1* that followed a sigmoidal pattern of expression compared to drug concentration. However, *FOXM1* was not significantly downregulated and did not show a sigmoidal pattern of drug-induced expression (**Fig. 4.3A**). Western blot analysis showed a similar trend with lower CCNB1 and PLK1 protein levels observed after 8 hours of 0.5 and 2mM HU (**Fig. 4.3B**). Conversely, FOXM1 protein is upregulated beginning at 8 hours of HU treatment (**Fig. 4.3B**). These results indicate that the type II HU hits validate at the protein level as well as at the transcript level.

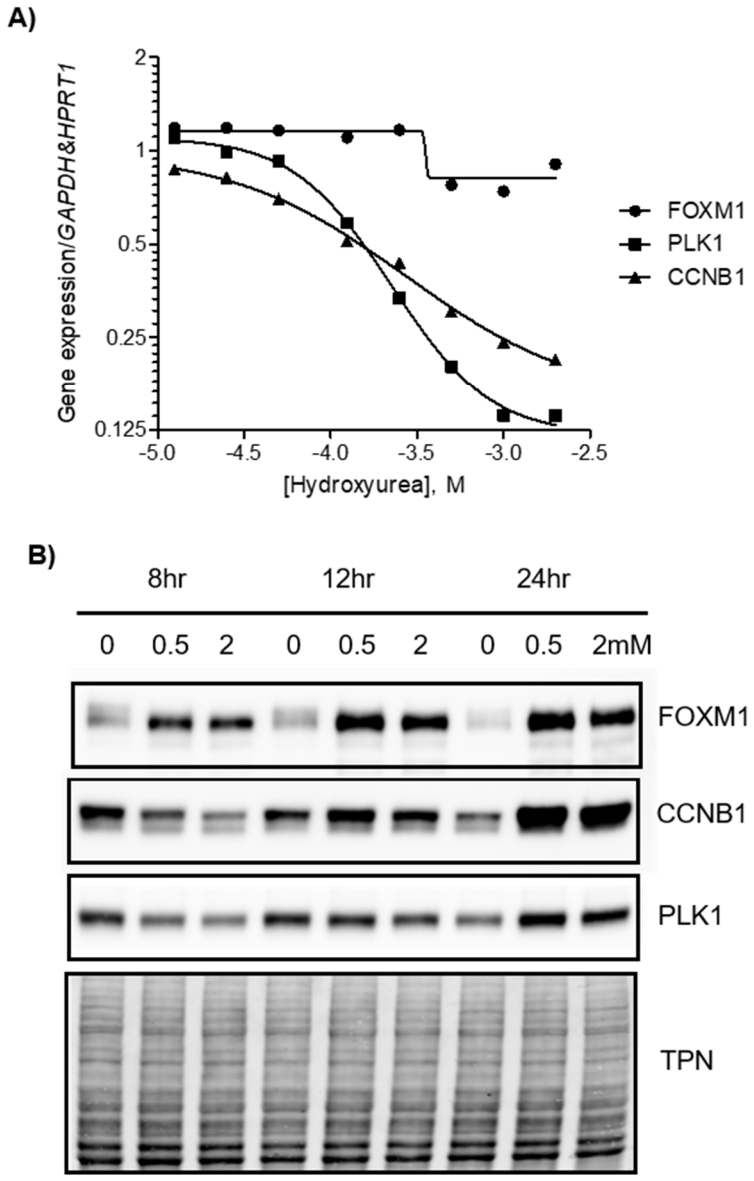

Figure 4.3. Dose-curve response experiment for HU in U87 cells. ***A)*** Dose-curve of HU effect on *FOXM1*, *CCNB1* and *PLK1*. Each data point represents differential expression (n=1) established by qRT-PCR and plotted against the $\log_{10}$([drug]). Data for CCNB1 and PLK1 fit to non-linear regression. **B)** Western blot analysis of total protein extracted from U87 cells treated with 0, 0.5, and 2mM HU for 8, 12, and 24 hours (n=1). FOXM1, CCNB1, and PLK1 proteins were visualized by chemiluminescence and normalized against total protein (TPN).

Hydroxyurea downregulates key G2/M checkpoint genes in blood-derived cells

To explore the impact of HU on the FOXM1-network genes, a blood-derived cell line, human immortalized lymphoblast cells (GM16119), were treated with 0.5mM HU over 24 hours. Overall a picture of increasing FOXM1 protein (but not mRNA) and decreasing FOXM1 target gene mRNA and protein was observed. FOXM1, PLK1, and CCNB1 transcript and protein levels were analyzed by qRT-PCR and western blot respectively. The greatest downregulation of *PLK1* and *CCNB1* transcripts (8-fold and 4-fold respectively) occurred after 4-8 hours of HU treatment (**Fig. 4.4A**). Conversely, the *FOXM1* transcript was moderately upregulated after 24 hours of HU treatment (**Fig. 4.4A**), corroborating the results observed in U87 cells. PLK1 and CCNB1 protein levels were halved at 4-8 hours of HU treatment but rebounded markedly by 24 hours (**Fig. 4.4B-C**). FOXM1 protein levels quadrupled after 8 hours of HU treatment increasing by ~8-fold at 24 hours (**Fig. 4.4B-C**).

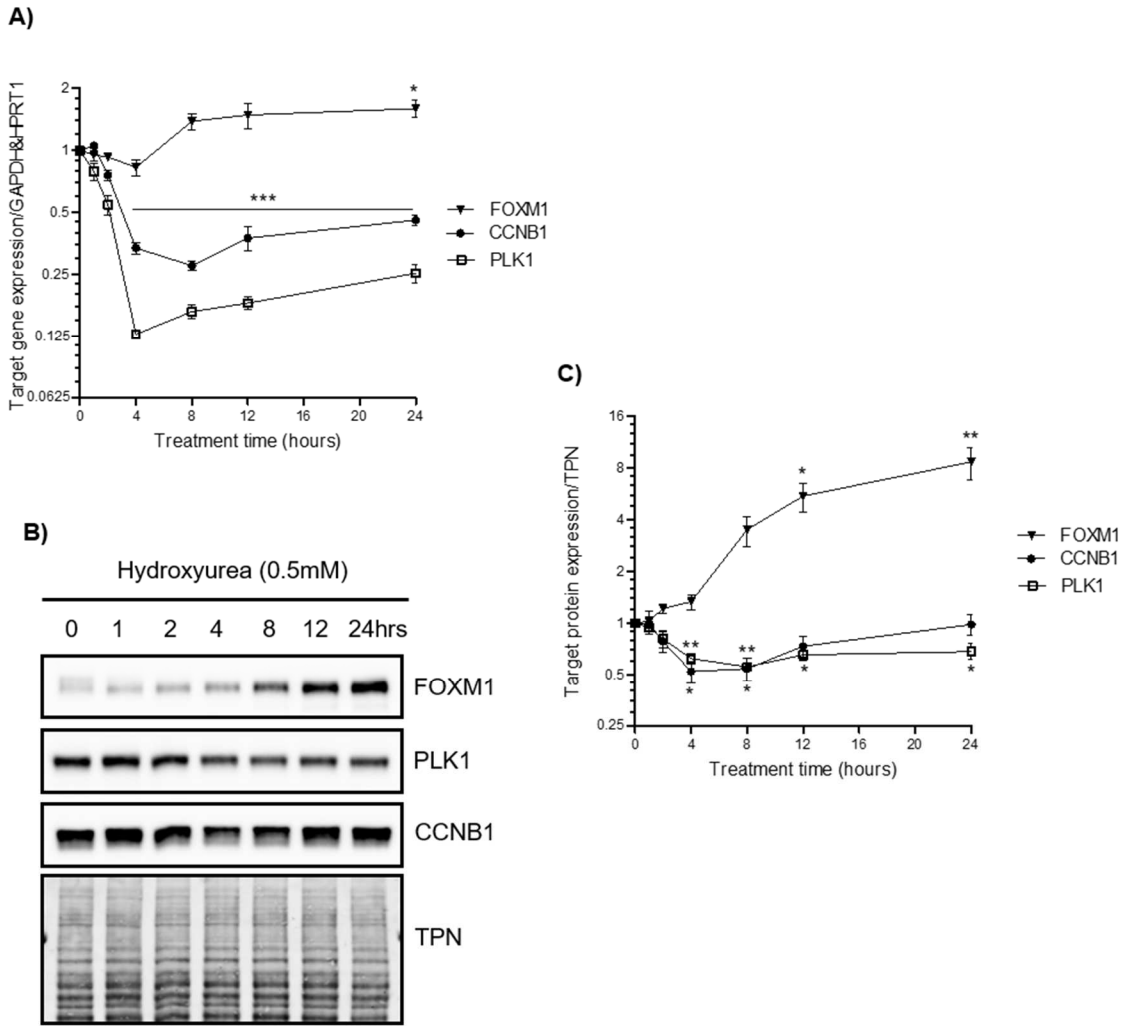

Figure 4.4. Validation of HU induced FOXM1 network GM16119 cells. Time-course of HU treatment on FOXM1 and its two target genes *CCNB1* and *PLK1*. U87 cells were treated with 0.5μM HU for 1, 2, 4, 8, 12, 24 hours (n=3). **A)** qRT-PCR results of the time-course experiment (y-axis scale log2). All time points are statistically significant except for those marked not significant (ns). **B)** Representative western blots for FOXM1, PLK1, and CCNB1 with total protein stain (TPN) used as a loading control. **C)** Pixel intensity for each protein was divided by the TPN and normalized to 0-hour time-point. One-way ANOVA with post-hoc Tukey test was used to measure statistical significance (*$p<0.05$, ***$p<0.001$).

Hydroxyurea abrogates FOXM1-mediated expression of *PLK1* in lymphoblastoid cells

Chromatin immunoprecipitation coupled with qRT-PCR (ChIP-qPCR) was used to further explore the role of FOXM1 in the HU-mediated downregulation of PLK1 in GM16119 cells. The *PLK1* gene was chosen because of a recent report linking the mitotic modulation of polo-like kinase canonical pathway to reduced fetal hemoglobin gene expression (Li et al. 2012). To confirm that HU-mediated *PLK1* downregulation is linked to a loss of FOXM1-mediated transcriptional activation, primers were designed to span the promoter region of the *PLK1* gene (**Fig. 4.5A**). GM16119 cells were treated with 0.4mM HU for 0, 1, 2, 4, and 8 hours; formaldehyde was then used to cross-link chromatin, and ChIP validated FOXM1 primary antibody was used to isolate FOXM1 associated DNA. Quantitative RT-PCR was used to measure the differential expression of *PLK1* in purified DNA from HU treated and control-treated cells. FOXM1 binding to the core-promoter of *PLK1* was confirmed by qRT-PCR of ChIP DNA (**Fig. 4.5B**). HU reduced FOXM1 occupancy of the *PLK1* core-promoter after 2, 4, and 8 hours of treatment, demonstrating that although FOXM1 is upregulated, its presence at the *PLK1* promoter is reduced in response to HU treatment (**Fig. 4.5C**).

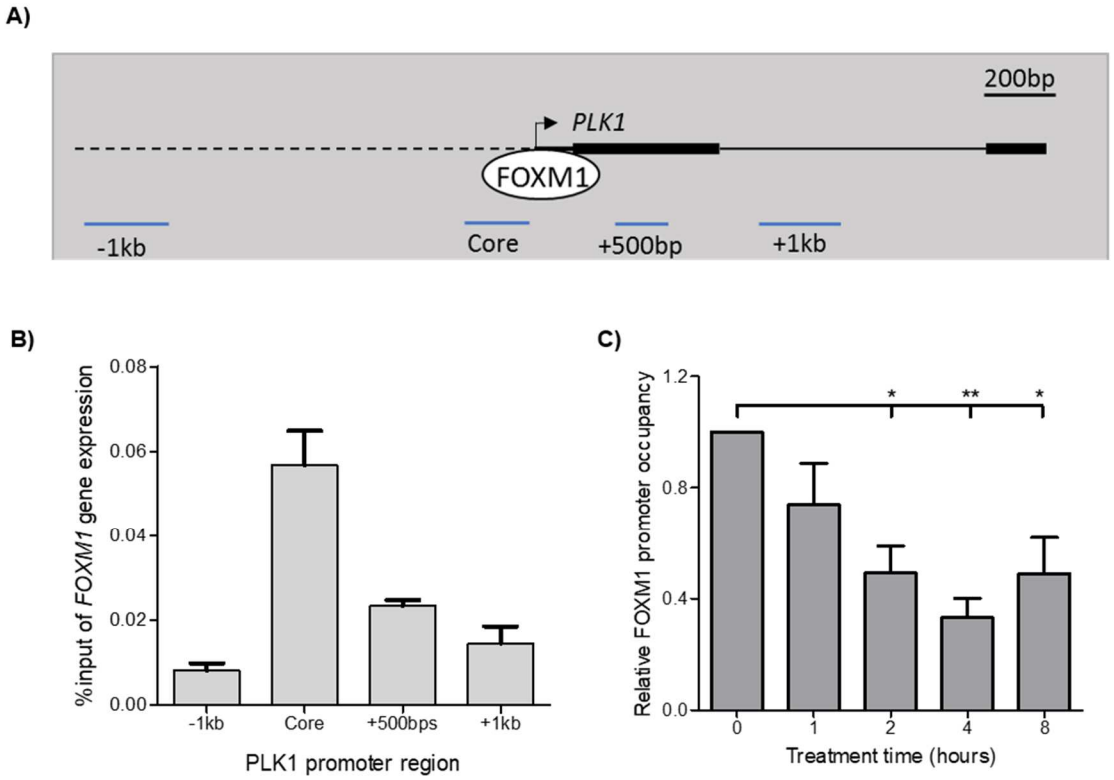

Figure 4.5. FOXM1 occupancy of the PLK1 promoter in the presence of HU treatment in GM16119 cells. GM16119 lymphoblastoid cell were treated for 0-8 hours with 0.5mM HU. Chromatin was then collected for chromatin immunoprecipitation (ChIP) by FOXM1. qRT-PCR was used to determine expression of *PLK1* in ChIP purified DNA. ***A)*** Diagram of qRT-PCR primer localization around the *PLK1* promoter sequence. Blue lines represent the amplicon for each primer pair. **B)** Confirmation of FOXM1 binding to the *PLK1* core promoter sequence. **C)** ChIP-qPCR results of the HU time course using the *PLK1* core-promoter primer set. One-way ANOVA with post-hoc Tukey test was used to measure statistical significance (*$p<0.05$, **$p<0.01$).

Dexamethasone upregulates downstream genes of the PPARD network *in vivo*

Given the significant attrition rate of type I drug-gene hit validation during the transition from the *in vitro* to *in vivo* setting, it became important to determine the rate of validation of type II hits in the whole-animal context. In addition, the validation of FOXM1-network type II hits was performed in an "inhibited" URA network, leaving open the question of whether type II hits from "activated" URA networks would also be validated. The PPARD network, upregulated by DEX, was therefore used to determine the rate of type II hit *in vivo* validation. Because DEX activity was confirmed in the cortex of mice treated with 1mg/kg for 5 days (upregulation of *Fkbp5*), a search was performed by URA in the Neuron Screen DEX treated differential expression data for robust activated networks. Two robust networks were identified; the STAT4 and PPARD upstream regulator networks were both activated with respective activation z-scores of 2.9 and 3.1 (**Table 4**). The PPARD network was chosen for validation because it is composed of 10 upregulated genes, nine of which are confirmed targets of PPARD (orange arrows), and two that are linked to rare neurogenetic diseases (*Mfsd2a*, *Mertk*) (**Fig. 4.6A**). The PPARD network was investigated *in vivo* by qRT-PCR analysis of cortex from 1mg/kg DEX-treated mice (the same samples that showed upregulation of *Fkbp5* and *Klf9*) (**Fig. 3.9D**). Seven of the ten genes in the PPARD network were studied by qRT-PCR, three of which (*Bcl2l1*, *Ilk*, *Mfsd2a*) showed statistically significant upregulation in DEX-treated mice (**Fig. 4.6B-D**), while *Pdk4* and *Kyat3* showed a trend toward upregulation in DEX treated mice (**Fig. 4.6E, Fig. S4.6A**). Only two of the seven

genes tested showed no increase in response to DEX treatment in mouse cortex (**Fig. S4.6B-C**).

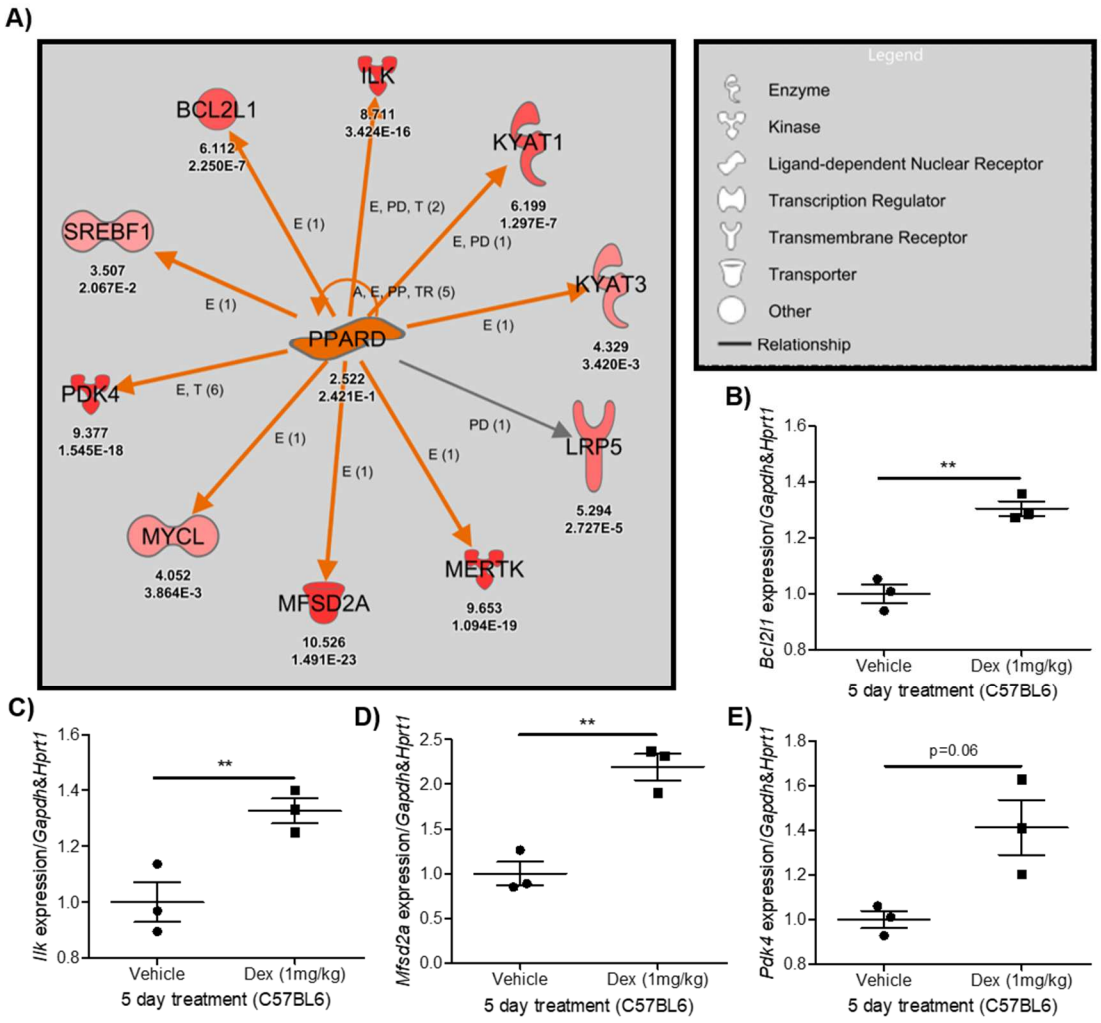

Figure 4.6. Investigation of the effect of dexamethasone on the PPARD URA network *in vivo*. *A)* The PPARD network identified by URA is activated in DEX-treated 3H cultures. Orange arrows signify activation and red symbols signify upregulated genes (Z-score and p-adj appear directly below each symbol). (***B-E***) Male C57BL6 mice were treated orally for 5 days with vehicle or DEX (1mg/kg). Gene expression of PPARD targets was determined by qRT-PCR **(B)** Bcl-2-like protein 1 **(C)** Integrin linked kinase. **(D)** Major facilitator superfamily domain containing 2A. **(E)** Pyruvate dehydrogenase kinase 4. Statistical significance measured by paired t-test (n=3, **p<0.01).

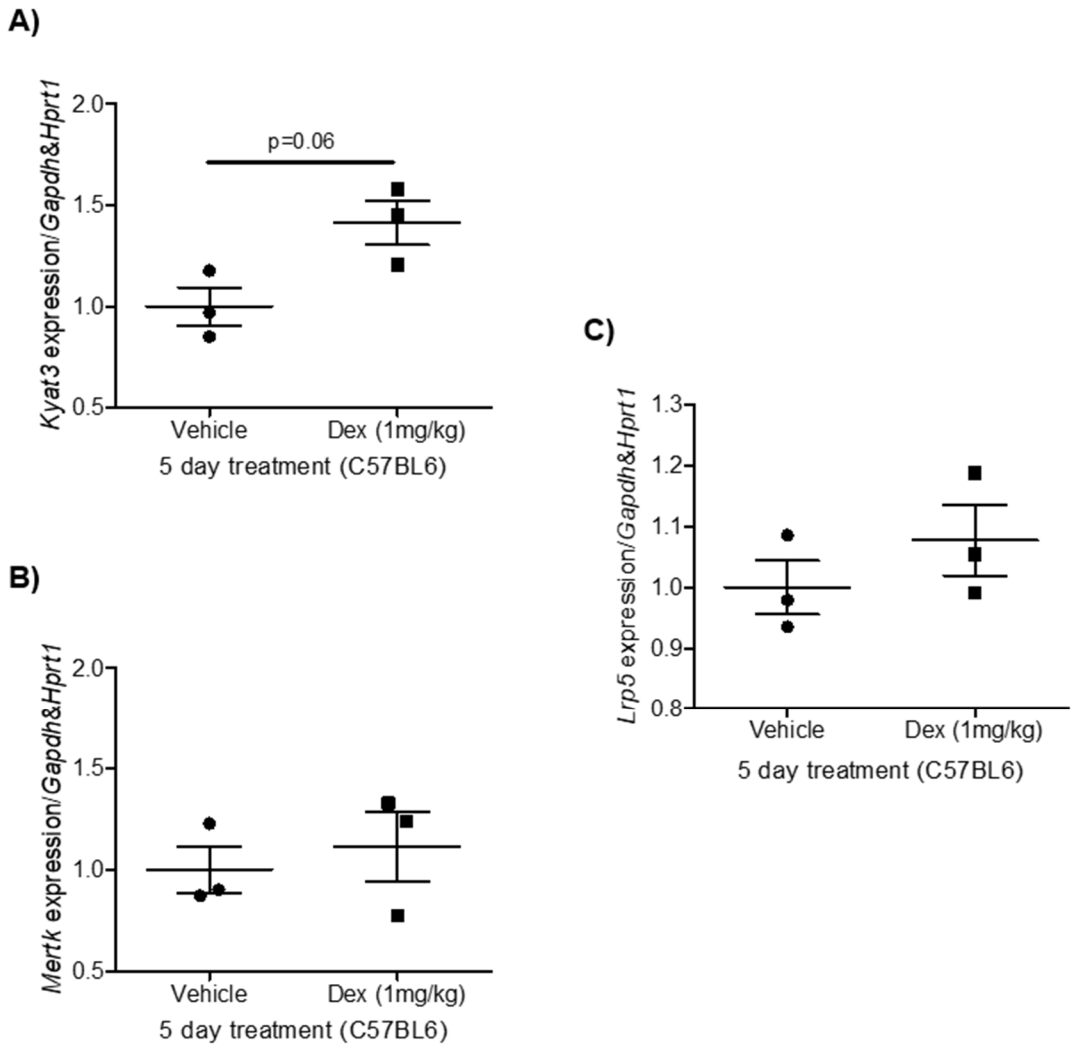

Figure S4.6. Validation of the effect of dexamethasone on three PPARD targets *in vivo*. *A)* Male C57BL6 mice were treated orally for 5 days with vehicle or DEX (1mg/kg). Gene expression of PPARD targets was determined by qRT-PCR **(A)** Kynurenine Aminotransferase 3 (*Kyat3*) **(B)** C-mer proto-oncogene tyrosine kinase (*Mertk*) **(C)** LDL receptor related protein 5 (*Lrp5*).

DISCUSSION

Overview of salient results

With the goal of identifying transcriptional modulators of therapeutic potential to rare neurogenetic diseases, primary neurons and next generation sequencing were used to develop the first transcriptome-wide clinic-ready drug screen in neuronal cells – i.e. Neuron Screen. I first explored whether primary cerebrocortical cultures from mouse could be used as a physiologically relevant model of the mammalian brain for a transcriptome-wide drug-screen. This analysis found in Chapter 2 indicates that the 3H cerebrocortical cultures have many characteristics of the mammalian cortex they are intended to model including neuronal diversity and synaptic maturity. The RNAseq based differential expression analysis of the 219 clinic-ready drugs revealed ~14,000 expressed genes, 95% of which were protein-coding. The reliability of the differential expression data from the Neuron Screen was supported by the marked similarity of profiles elicited by two pairs of chemically related drugs: two thyroid analogs and two synthetic corticosteroids. I next examined whether the differential expression contained in the database could be used to reliably predict drug-gene hits relevant to rare neurogenetic diseases. I initially studied type I hits – i.e. drugs that upregulated genes related to C4R-prioritized neurogenetic diseases (hypomorphic, haploinsufficient, presence of a rescuing paralog). Although 6 of 32 type I hits showed confirmatory mRNA upregulation in cerebrocortical cultures, unfortunately, even the most robustly induced mRNA did not translate to *in vitro* protein induction, and only one drug-gene hit showed promising transcriptional results at the *in vivo* level. Given the significant type I drug-gene interaction attrition observed in both the mRNA to protein and *in vitro*

to *in vivo* transitions, using the IPA software, I next explored whether type II hits, network-associated drug-gene interactions validated at a greater frequency. This turned out to be so, with 7 of 7 type II hits showing confirmatory mRNA modulation in cell culture and, for two of these, a corresponding modulation of protein. Furthermore, 6 out of 7 type II hits showed confirmation of mRNA modulation *in vivo* supporting the use of URA networks in identifying reliable drug-gene hits from a differential expression database like the Neuron Screen. Ultimately, the primary goal of the Neuron Screen, to identify and transcriptionally validate rare neurogenetic disease-related drug-gene interactions, was achieved albeit with a lower frequency than anticipated. Two drug-gene hits, a type I and a type II hit, showed promising results *in vivo* experiments. The results shown in Chapter 3 provide strong evidence of a relationship between thyroid hormone and the HNPP related gene *Pmp22* while those in Chapter 4 provide *in vivo* evidence of a transcriptional connection between the steroid DEX and the rare disease-related gene *Mfsd2a*. Finally, results from the Chapter 4 study of the FOXM1 network provide evidence for a novel mechanism associated to the HU-related upregulation of fetal hemoglobin.

Relevance of primary neuronal culture for drug screening

The selection of the optimal *in vitro* model for a transcriptomic drug screen, as with most high throughput screens, must strike a balance between physiological relevance and feasibility. Clearly the most physiologically relevant system for the study of rare neurogenetic diseases is the human brain itself with *in vitro* cultures of human primary embryonic neuronal and glial cells serving as a suitable proxy. However, given the legal and ethical issues associated with research using human embryonic tissue, I asked instead whether primary mouse cerebrocortical culture would be a suitable substitute in which to conduct transcriptome-wide drug screening. In this regard, immunohistochemical analyses conducted at DIV21 demonstrated neuronal and astrocytic but not non-neuronal lineage (e.g. endothelial) specific markers supporting the physiological relevance and purity of 3H cortical cultures. Two different methods were then used to estimate the ratio of astrocytes to neurons in the 3H cultures. First, cell counts of NeuN and GFAP immunolabelled populations revealed a glial-to-neuron ratio (GNR) of ~0.2 in 3H cultures at DIV21, consistent with a 0.2 GNR determined by counting NeuN and GFAP labelled cells in E17.5 mouse cerebrocortical cultures (Gil-Ibanez et al. 2017). Despite this consistency with published results, the accuracy of the traditional approach of quantifying astrocytes by anti-GFAP antibody staining has recently been called into question due both to the preferential labelling of brain white-matter astrocytes (compared with grey-matter) and increased GFAP expression in response to chemical or physical damage of neuronal tissue (Cahoy et al. 2008; Sun & Jakobs 2012). A transcriptome-wide screen for cell-type specific markers of astrocytes,

neurons, and oligodendrocytes by Cahoy et al. (2008) identified ALDH1L1 as a more faithful astrocyte marker in the CNS (and specifically the cortex) than GFAP, based both on its broader expression in the brain and identification of a greater number of astrocytes than GFAP (Cahoy et al. 2008; Yang et al. 2011). In keeping with this view, I found equal number of DIV21 3H culture cells stained for the neuronal marker MAP2 and ALDH1L1. The observed 1:1 MAP2:ALDH1L1 cell ratio is consistent with a study finding the same proportion of NeuN stained to NeuN negative cells in rat cortical cultures (Schock et al. 2012). Results from the 3H cultures are also in line with a recent study that indicated that out of four astrocyte-specific markers (including ALDH1L1) GFAP identified the least number of astrocytes (Tong et al. 2014). Therefore, counting GFAP-immunolabelled cells appears to under-estimate the number of astrocytes in the 3H cultures. However, despite the increasing evidence of astrocyte underestimation by GFAP-labeling and the more accurate estimation with anti-ALDH1L1 antibodies, the use anti-GFAP antibodies to identify astrocytes appears to be entrenched. One reason for this unwillingness to adopt ALDH1L1 as an astrocyte-specific marker may be the fainter and higher background staining observed with ALDH1L1 antibodies compared to the clear staining attained with GFAP antibodies (Cahoy et al. 2008; Tong et al. 2014) In my analysis of the 3H cultures, counts of ALDH1L1-positive cells were less straightforward than counts of GFAP-positive cells. The ALDH1L1 based GNR of the 3H cultures, although higher than initially expected, is in agreement with two recent reports of non-neuronal to neuronal cell ratios (nNNR) of human cortical grey matter ranging from 1.48 to 1.64 (Azevedo et al. 2009; Andrade-Moraes et al. 2013). Finally, given that ~30% of all

non-neuronal cells in the human brain are endothelial, the GNR drops to 1.04-1.15 (von Bartheld et al. 2016). A similar study found that the nNNR of the mouse cerebral cortex is ~3, although this study included the glial-rich white matter as part of the cortex, and did not take into consideration the number of other glial cells and endothelial cells that make up the non-neuronal fraction of cells in the mouse cortex (Herculano-Houzel et al. 2006; Tasic et al. 2016). The GNR takes on added importance given the astrocyte's crucial roles in synapse development, maturation, and maintenance (Herculano-Houzel 2014; Araque & Navarrete 2010) . The presence of astrocytes in "neuronal" cultures, especially cultures grown to achieve relative maturity is widely understood as a necessity for proper *in vitro* network formation and function (Banker 1980; Pfrieger & Barres 1997; Ullian et al. 2001). Recent reports have stressed the necessity of astrocyte co-culture to promote longevity and network functioning in human iPSC derived neurons and astrocytes are being promoted as essential for creating physiologically relevant human induced neuronal cultures if they are to be used as reliable tools for drug screening (Lam et al. 2017; Odawara et al. 2016). In keeping with this view, when mitotic inhibitors were used on the 3H cultures in an attempt to reduce the GNR, neurons did not survive (data not shown). In conclusion, the 3H cortical culture GNR of ~1 suggested by ALDH1L1 and MAP2 staining is comparable to *in vivo* estimates in both mouse and human and lends support to the physiological relevance of the 3H cerebrocortical model.

While the presence of astrocytes in cortical cultures is required for neuronal health and to support network formation and maintenance, the appropriate complement

of neuronal subtypes in the culture is essential for neuronal network function. Neurons that comprise the mammalian cerebral cortex are broadly defined as glutamatergic (projection neurons) or GABAergic (interneurons). The glutamatergic excitatory neurons account for ~80% of neurons and establish short and long-range connections within the brain and spinal cord while the remaining ~20% of neurons are a heterogeneous mix of short-range GABAergic inhibitory neurons (Hendry et al. 1987; Meinecke & Peters 1987). The presence of both classes of neurons is essential to maintain the appropriate balance between excitation and inhibition in neuronal networks. Excitatory and inhibitory neurons form highly interconnected feed-back and feed-forward loops resulting in the rhythmic synchronous membrane oscillations that are a hallmark of mature cortical networks (Isaacson & Scanziani 2011; Le Magueresse & Monyer 2013). Although much work remains to be done in deciphering the specific roles of the different classes of interneurons in the cortex, it is well documented that there exist regional differences in the cortical layers in GABAergic interneuron subpopulations (Molyneaux et al. 2007; Kubota et al. 2011). Since GABAergic interneurons migrate into the cortex at a later stage than glutamatergic neurons, it is important, when making the case for the suitability of an embryonic cortical culture in modelling the mammalian cortical system, to confirm their presence (Kelsom & Lu 2013; Guo & Anton 2014). Despite the necessity of interneurons for cortical network function, most studies employing primary cortical cultures do not look for the presence of interneurons (Sharma et al. 2012). I employed immunofluorescence staining of 3H cultures to confirm the presence of three subtypes of GABAergic interneurons that express neuronal nitric oxide synthase, calretinin, and

parvalbumin. The presence of these interneurons provides important support for the physiological relevance of the 3H cultures to model *in vivo* cortical networks.

In addition to its cellular make-up, it was also necessary to confirm the developmental and functional maturity of the 3H cultures especially given that mature neurons have different transcriptomic profiles than immature neurons and that transcriptional noise from maturing neurons could mask drug-induced transcriptional changes (63,120). I therefore next conducted an immunohistochemical and electrophysiologic assessment of culture system maturity, defined at the cellular level by immunostaining for the synaptic markers vGAT and vGLUT, and at a network level, by multi-electrode array (MEA) assay. 3H cultures showed the characteristic punctate staining of vGAT and vGLUT observed in mature cortical synapses (Schock et al. 2012). Consistent with the literature, 3H cultures on MEAs showed spontaneous burst activity by seven days *in vitro* (Otto et al. 2003). Importantly, MEA recordings at 7, 15 and 21 days *in vitro* revealed visible increases in firing rate and synchrony of bursts culminating at DIV21 with the highly synchronous firing characteristic of full cortical network maturity (Luhmann et al. 2016). In addition, at DIV21, every electrode is firing, indicating high interconnectedness of the cortical cultures at this time-point. The 3H cortical cultures thus appear to have reached maturity at the synaptic and network levels by DIV21.

Primary neuronal cultures are more difficult than immortalized cultures to use for drug screens, with greater susceptibility to environmental changes, requiring both specific growth media formulations as well as plates coated with a synthetic or

biological polymer, and are post mitotic thus their populations cannot be expanded. Notwithstanding these challenges, I confirmed that the 3H cortical system can, in a consistent and controlled fashion, be cultured on the scale required for transcriptome-wide drug screening. Moreover, our assembly-line approach enabled the rapid collection and culture of a sufficient number of embryonic cortices needed to conduct the screen of 218 drugs and 9 control treatments. Therefore, I have shown that mouse embryonic cortical cultures are a feasible model for transcriptomic screens such as our Neuron Screen.

Neuronal RNAseq transcriptome-wide screen

One of the benefits of using RNAseq for the Neuron Screen is that the impact of the drug treatments could be studied on every gene expressed in the 3H cultures. The Neuron Screen data, although primarily intended for the identification of candidate drugs for the ~60 C4R neurogenetic diseases, may also find utility for any of the 1000s of rare neurogenetic diseases as well as further our understanding of the impact of these clinically used agents in the CNS in general. A rigorous estimation of the number of mRNAs interrogated in our screen was thus an important consideration, ensuring that the Neuron Screen database would be as comprehensive as possible in capturing the neuronal transcriptome response to FDA approved drugs. RNA sequencing of the 227 individually treated 3H cultures led to the detection of 14,070 genes that were expressed in all drug and control treated cultures, of which 13,587 are protein coding genes. According to the human brain proteome, 14,548 protein coding genes are expressed in

the brain. However, whole brain proteomic studies include all regions of the brain, each with different gene expression profiles, whereas the 3H cultures exclusively represent the transcripts of the cortex (Sun et al. 2012). A recent RNAseq study of a mouse cerebrocortical culture system similar to the 3H culture system used single-end 75bp (in contrast to the Neuron Screen's 50bp) reads and a similar sequencer to identify expression of 14,801 genes (Gil-Ibanez et al. 2017). The ~6% fewer genes identified in the 3H cultures compared with the findings of Gil-Ibanez et al. (2017) could be due to variation in efficiency of mRNA capture; the more efficient the mRNA capture, the fewer non-coding genes are identified by sequencing. Thus, the gene number can be inflated by the identification of more non-coding genes; unfortunately, the number of protein-coding genes expressed in their DIV9 cerebrocortical cultures was not reported. In addition, given that 30% more genes are expressed in cultured neurons at DIV9 compared to neurons from mature mouse brain (Lovatt et al. 2014), the twelve day difference between the 3H DIV21 cultures and their DIV9 system may also be a factor. Finally, it has been estimated that 12,137 genes are expressed by single neurons in brain tissue (Lovatt et al. 2014); given 3H cultures contain a mixture of neurons and astrocytes, a greater number of genes expressed in the 3H cortical cultures might be anticipated. Regardless, the number of genes identified in the Neuron Screen indicates both a good coverage of the protein-coding transcriptome and that the RNAseq protocol and the bioinformatic pipeline used to analyze the data were appropriate.

The expense of RNAseq analysis combined with the labour/resource intensive nature of 3H culture preparation limited the Neuron Screen to single drug-treatments

with no opportunity to perform replicates. In an effort to add statistical rigor to differential expression profiles, all 227 drug and control treated conditions were treated as biological replicates based on the hypothesis that the majority of drugs have no effect the majority of genes. In general this was found to be the case with a very low transcriptional response to drug treatments as seen in the following observations (i) ~80% of genes were significantly modulated by less than 1% of the 218 drugs used for the screen (ii) only 0.2% of genes were significantly modulated by more than 5% of drugs and (iii) >50% of drugs had no greater effect on the neuronal transcriptome than the average for vehicle treatments. In addition, the top five transcriptome modulating drugs (fenofibrate, nilotinib, ciprofibrate, aminophylline, and diflunisal) collectively accounted for almost half of the differentially expressed genes identified in the screen.

Most transcriptome-wide screens have used drugs that have a well-defined gene expression signature such as histone deacetylase inhibitors (HDACi) for positive-controls (Lamb et al. 2006). However, given our selection criteria for clinically approved low toxicity BBB penetrant drugs excluded such agents and given two thyroid hormone analogs (levothyroxine, liothyronine) and two corticosteroids (dexamethasone, betamethasone) were included in the Neuron Screen drugs, these two drug classes, with well described transcriptomic impact, were used as positive controls instead. Thus, observation of similar neuronal transcriptomic profiles elicited by the two thyroid hormone analogs as well as the two corticosteroids served to validate the reproducibility of the Neuron Screen results. The IPA based comparison between thyroid hormone analogs' transcriptional effect observed in the Neuron Screen and literature on thyroid

hormone transcriptional effects was performed by upstream regulator analysis and showed that the highest z-score of activation for the upstream regulator was "triiodothyronine" – i.e. thyroid hormone. Similarly, for the two corticosteroids, the two shared URA networks with the highest z-scores were confirmed to be "dexamethasone" and "corticosterone" – i.e. endogenous glucocorticoid. Additionally, for both thyroid hormone and corticosteroid drug classes, the top seven common URA networks have all been previously implicated in transcriptional response to that drug class while the majority of downstream genes identified by the top URA networks have also been linked to thyroid hormone or DEX treatment. Interestingly, two other drug classes, the flavonoid class (quercetin, luteolin) and the fibrate class (fenofibrate, ciprofibrate, clofibrate) did not show the same degree of similarity in intra-class URA network profiles. This may be because fenofibrate, ciprofibrate, and luteolin all showed induction of important endoplasmic reticulum (ER) stress genes (*Eif2ak3, Ppp1r15a*, *Ddit3*, *Hspa5*), not genes historically known to be flavonoid or fibrate targets and possibly a result of drug-induced cellular stress signaling. In fact, the six drugs impacting the greatest number of genes also activated four ER-stress genes (**Appendix VI**) in the Neuron Screen data, genes that are important to ER-stress and the unfolded protein response (UPR) pathways. Number one on the list, the PPARalpha agonist fenofibrate showing in general the most pronounced effect on these pathways.

Nonetheless, the high similarity of the transcriptomic profiles of the thyroid hormone analog drugs along with high correlation to the IPA-based literature analysis, and the same observation with the corticosteroid drugs point to the reproducibility and

validity of the majority of the Neuron Screen data. However, caution should be used when interpreting results for the drugs that led to activation of ER stress and UPR gene expression as individual transcriptional changes may not be indicative of drug-induced changes but rather related to dose-induced toxicity. This is certainly the case with fenofibrate that led to a dose-dependent statistically significant increase in cell death measured by LDH assay. Due to this fact, none of the drug-gene hits identified by fenofibrate were included for validation. The four highest transcriptome modulating drugs in the Neuron Screen, after fenofibrate, are nilotinib, ciprofibrate, aminophylline, and diflunisal. All four of these drugs also show enrichment of ER stress gene activation and UPR (IPA analysis) but not to the same extent as fenofibrate. Nilotinib, aminophylline, and diflunisal did not however lead to significant increase in cell death by LDH assay and therefore hits from these drugs were not excluded from validation (ciprofibrate was not tested by LDH assay since no ciprofibrate drug-gene hits were validated). The activation of ER stress pathways by several of the Neuron Screen drugs points to one of the caveats of using therapeutic serum levels to guide drug concentrations used for the screen. Although the drugs used for the screen are theoretically BBB penetrant (except for fenofibrate), many such as nilotinib and aminophylline have a low CSF to serum ratio. Thus, the use of human serum levels in tissue cultures that do not have a BBB leads to the drug being present in the 3H cultures at suprapharmacologic levels compared to brain.

In vitro transcriptional validation of Neuron Screen hits

Two methods were used to identify drug-gene hits (defined here as an increase in target mRNA) for validation: the individual drug-gene hits (type I hits) related to the ~60 C4R-prioritized rare neurogenetic diseases, and the network-related drug-gene hits (type II hits) identified by IPA analysis of drug-modulated transcriptome-wide expression patterns. Roughly 20% of type I hits were validated by qRT-PCR in 3H cortical cultures at the drug concentrations originally used for the screen. The validation of type I hits from the Neuron Screen is roughly double the validation rate of hits from the fibroblast screen and roughly four times the validation rate of hits identified from CMap data by Mears et al (2017). The hit validation rate is an important quality indicator for our RNAseq data analysis, demonstrating what proportion of the drug-gene interactions identified are real and that the statistical cut-off for hit-calling was appropriately set, minimizing both type I (false positive) and type II (false negative) errors. To identify the 32 type I drug-gene hits that were investigated by qRT-PCR, the statistical boundary was set at the commonly used p-adj<0.05. However, the only drug-gene hits that validated at the transcript level in 3H cortical cultures had a p-adj<0.01. Furthermore, if only hits with a p-adj<0.01 are taken into consideration, the validation rate of type I hits was 50% higher, with no decrease in statistical power. Thus, to minimize false-negatives and maximize the efficiency of type I hit validation, only drug-gene hits from the Neuron Screen with a p-adj<0.01 should be seen as significant and further assessed for validity. The superior validation rate observed for hits detected by the Neuron Screen compared to the fibroblast screen performed by Mears et al. (2017),

may be a result of the different statistical cut-offs yielding increased statistical stringency.

An integral part of the early phases of traditional drug discovery is the use of a drug-dose curve to identify and characterize the dose-dependent effect on a therapeutic target molecule (i.e. enzyme, ion channel, etc.). The typical drug-dose curve is sigmoidal in shape and enables the calculation of the EC50 value (half-maximal concentration) that provides evidence of a drug's potency (Muller & Milton 2012). The dose-curve is also commonly used to confirm if there is a biological relationship between a drug and the intended biological target; a transcriptional effect that follows a sigmoidal drug-dose-induced response is an indicator of a physiological relationship whereas an all or nothing effect (effect only observed at high dose) is likely due to drug toxicity (and not a biological effect) (Hughes et al. 2011; Ji et al. 2009). Four of the six type I hits which validated in 3H cortical neurons were tested for a dose-dependent mRNA modulation by means of a half-log drug-dose curve performed in cerebrocortical cultures. One hit, aminophylline-*Aldh18a1*, showed a dose-dependent increase of expression with a sigmoidal shaped curve typical of a drug-induced physiological response. A roughly 50% effect size on *Aldh18a1* expression was observed from low (3μM) to high (250μM) aminophylline dose with an EC50 calculated at 33μM. Two others, nilotinib-*Sacs* and diflunisal-*Slc6a8*, showed a single dose response at the original dose used for the screen. The screening concentration of 10μM nilotinib caused significant *Sacs* upregulation while neither lower nor higher (31.6μM) did; the latter dose caused downregulation of the housekeeping gene *Gapdh* and measurable cellular toxicity according to the LDH

assay. The 10μM dose of nilotinib also caused upregulation of ER stress-related genes *Ppp1r15a*, *Ddit3*, and *Hspa5*. For the diflunisal dose-curve, the highest dose used was the Neuron Screen dose of 200μM, and only this dose upregulated *Slc6a8*. Given that only high dose led to upregulation of the target genes, it is possible that the nilotinib-*Sacs* and diflunisal-*Slc6a8* hits are the result of the activation of endoplasmic reticulum (ER) stress and associated unfolded protein response (UPR) rather than other specific pathways. Interestingly, measurement of three well-described ER-stress response genes (*Ppp1r15a*, *Ddit3*, *Hspa5*) by RNAseq (confirmed by qRT-PCR) showed that the *Sacs* upregulation by nilotinib is strongly correlated with ER-stress gene expression but not with the *Slc6a8*-inducing dose of diflunisal (**Appendix VI**). Also, ER-stress genes do correlate with an increase in *Aldh18a1* in response to aminophylline, but this does not appear to negate the dose-dependent increase in *Aldh18a1*. The aminophylline-*Aldh18a1* hit appears to be an example of how true biological interactions can still be identified even if there is activation of ER-stress and UPR genes in the transcriptomic data. The fourth type I hit studied for a dose-dependent response T4-*Pmp22* showed a third pattern; rather than dose-dependent or single-dose response; *Pmp22* was upregulated by all five doses of T4 treatment used for the dose-curve with no statistical difference between the effect of the five doses. This pattern of dose response may be indicative of a plateau effect, where the drug at a certain minimum dose saturates the target and thus there is no further increase in gene expression with drug-dose increases. Corroborating the hypothesis of a plateau effect of thyroid hormone on *Pmp22* expression, the dose curve of T3 performed in rat

DRG cultures showed a similar plateau of *Pmp22* transcript upregulation from 10nM to 100nM.

Gene Expression Correlation between Primary Neuronal Cultures and Immortalized Cells

None of the three most promising (i.e. largest effect size) type I hits validated in immortalized human and rat cell lines. The aminophylline-*Aldh18a1* and nilotinib-*Sacs* hits were studied in differentiated and undifferentiated SH-SY5Y cells. The T4-*Pmp22* hit was studied in RT4 rat schwannoma cells. The RT4 rat schwannoma cell line was used to test the thyroid hormone analogs-*Pmp22* hits because the disease related to downregulation of *PMP22* (HNPP) is a primary disease of Schwann cells. Given the failure of T4 to increase *Pmp22* gene expression, a dose-curve of T3 was also performed. This was because Schwann cells express type II deiodinase (DIO2) which convert thyroxine into the active thyroid hormone triiodothyronine (T3), only in response to chemical or mechanical injury (Li et al. 2001; Courtin et al. 2005). However, T3 treatment also failed to upregulate *Pmp22* in RT4 cells. Diseases associated with *Aldh18a1* and *Sacs* affect the CNS, therefore the commonly used human SH-SY5Y neuroblastoma cell line, that exhibits neuronal markers and can be differentiated into neuron-like networks, was used to investigate these hits (Agholme et al. 2010).

There are two factors potentially involved in the poor correlation of drug-induced gene expression between the mouse primary cortical cultures and the cell lines (human SH-SY5Y cells and rat RT4 cells); (i) inter-species differences between mouse

and human/rat and (ii) genetic and biochemical differences that exist between primary cultures and immortalized cell cultures. Comparative transcriptome-wide gene expression studies have revealed much greater correlation of gene expression between species than between tissue types (Chan et al. 2009; Zheng-Bradley et al. 2010). Comparison of gene expression correlation has shown that expression correlation between human liver and a human hepatic cell line (HepG2) is greater than between human liver and rat primary hepatocytes (Sutherland et al. 2016). Yet, correlation of base-line gene expression does not necessarily correspond to correlation of drug-induced gene modulation. For example, immortalized human umbilical endothelial cells have similar gene expression to primary human umbilical endothelial cells but treatment with atorvastatin modulates four times more genes in the immortalized cells than the primary cells (Boerma et al. 2006). In general, there is a limited literature comparing drug-induced gene modulation between species or between different *in vitro* tissue models (i.e. immortalized cells, primary cells) (Iskar et al. 2013). For the aminophylline-*Aldh18a1* and nilotinib-*Sacs* interactions, because the SH-SY5Y cells originate from a different species than the 3H cultures, it is impossible to determine conclusively if the lack of validation in SH-SY5Y cells is due to inter-species or primary to immortalized cell differences in drug-induced gene modulation or both. Yet for aminophylline-*Aldh18a1* hit, the evidence may lean toward a lack of concordance between immortalized and primary cells since alignment of the human and mouse *Aldh18a1* gene show high sequence similarity at all exons and similar transcription factor binding elements on the gene promoter between both species (Anon 2012). Because the T3-*Pmp22* hit was

validated in rat primary DRG cultures but did not validate in rat immortalized cells (RT4) of similar tissue origin (PNS), this seems to be an example of lack of concordance between primary and immortalized cells. Although these results do not enable a pronouncement on the lack of relevance of the SH-SY5Y and RT4 cells to model the drug-induced modulation of genes in the neuronal context, the results corroborate the fact that gene expression similarity between cell types does not necessarily equate similarity of gene expression response to drug treatment (Boerma et al. 2006). Clearly however, three type I drug-gene hits are not representative of the entire transcriptome and further work at a broad transcriptional level will be needed to determine the degree of correlation between cell models used to study the nervous system and primary neuronal cultures.

In contrast to validation of type I hits, 100% of type II hits that were analyzed at a single dose validated at the transcriptional level. The seven transcriptional targets of FOXM1 that were analyzed by qRT-PCR in human U87 cells showed a statistically significant reduction in gene expression after 4 and/or 8 hours of treatment with HU. Furthermore, the two hits that underwent a HU dose finding (PLK1 and CCNB1) showed a dose-dependent response with a sigmoidal shaped curve, indicating a likely biological response of these genes to HU treatment. The superior rate of type II hit validation when compared with type I transcriptional validation may be a result of several factors. First, although the same cut-off of p-adj<0.05 was used for type II hits, the average |z|-score was higher for the validated type II hits than for type I hits. The lowest Z-score for type II hits was |3.9| and was the only type II hit with a Z-score lower

than |4|. Conversely, twelve of the thirty-two type I hits had a Z-score<4. Statistically speaking, type I hits were at a disadvantage compared to type II hits. However, without the hits with Z-score<4, type I hits validated at 30% and some of the type I hits that did not validate had Z-scores>6. Second, type II hits were validated in human immortalized cells whereas type I hits did not validate in immortalized cells. The reason why type II hits identified in primary neurons showed strong validation in cancer cells is possibly ontological. The seven type II hits that validated in U87 and GM16119 cells are all involved in cell cycle progression through the G2-M checkpoint, whereas type I hits are a membrane protein gene (*Pmp22*), putative chaperone protein gene (*Sacs*), and metabolic enzyme gene (*Aldh18a1*). These latter genes may occupy specific roles in primary neurons compared to immortalized cells and thus may not be responsive to the same drug treatments in immortalized cells. Conversely, the transcriptional regulation of cell cycle-related genes is likely much more ubiquitous. A recent meta-analysis of transcriptional data that compared gene expression between human immortalized and primary cells (of different tissue origins) demonstrated high correlation of gene expression for G2-M phase genes including *BUB1*, *CDK1*, *CCNA2*, *CENPE*, *CENPF*, *CCNB1* (Giotti et al. 2017). Third, the method used for the identification of type I and type II hits may be the most important contributor to difference in validation rates. Since the type II hits investigated *in vitro* are all targets of the same upstream regulator (FOXM1), there is much greater statistical significance attributed to each individual type II hit because they exist collectively as indicators of the effect of HU on FOXM1-induced cell-cycle transcripts. Collectively, these results indicate that if drug-gene hits are

identified as part of a robust transcriptional network, like the FOXM1 network, there is high likelihood that these hits will represent physiological relationships. Such robust drug-gene hits may be more likely to transcend the immortalized-to-primary cell and mouse-to-human bottlenecks that plague drug discovery research.

In vitro protein validation of Neuron Screen hits

The success of my transcriptional modulation model for drug repurposing is contingent upon protein levels ultimately correlating with observed transcript levels. This is because, except for diseases caused by RNA toxicity (ex: myotonic dystrophy), the gene-dosage problem is really a protein-dosage problem; reduced levels of functional protein are the best explanation for the diseases caused by haploinsufficiency of *Pmp22* and hypomorphism of *Sacs* and *Aldh18a1*. Although robust transcriptional upregulation was shown in primary neuron cultures for T4/T3-*Pmp22*, nilotinib-*Sacs*, and aminophylline-*Aldh18a1*, these type I hits did not lead to statistically significant increase in protein expression. Protein expression was assayed by western blotting in 3H cortical cell cultures for the nilotinib and aminophylline related hits, while the thyroid hormone-Pmp22 hit was assayed in DRG cell cultures. This was done since Pmp22 protein is poorly expressed in most tissues apart from myelin of the PNS where it forms an essential and comparatively abundant component of compact myelin (Garbay et al. 2000; Lau et al. 2008). Pmp22 protein was not detectable by western blotting in DRGs. However, IF staining for Pmp22 in DIV28 DRG cultures showed perinuclear staining patterns in Pmp22 positive cells. Blinded cell counts of Pmp22 positive cells were challenging due to high background but there was a trend of increased Pmp22 stained

cells in T3 treated cultures compared to control. However, since thyroid hormone receptors are expressed on developing Schwann cells and Schwann cells responding to axonal damage, the increased number of Pmp22 positive cells could just indicate an increase in Schwann cells related to thyroid hormone signaling (Barakat-Walter 1999; Mercier et al. 2001).

There are several possible reasons for lack of transcript to protein correlation observed for the type I drug-gene hits. First, although traditional biological dogma states that changes in an mRNA's level should lead to changes in the encoded protein, only ~20% of genes have a high correlation between mRNA and protein abundance with a weak mRNA to protein correlation reported by many authors (Bauernfeind et al. 2015; Martinez-Nunez & Sanford 2016; Ramakrishnan et al. 2009). The mean genome-wide correlation of transcript abundance to protein abundance has recently been estimated at a comparatively paltry r=0.27 in mouse and r=0.29 in human cells (Ghazalpour et al. 2011; Ramakrishnan et al. 2009). The lack of validation of type I drug-gene hits at the protein level thus appears to reflect a significant transcript-protein concordance problem when thinking about developing rare disease therapies by analyzing transcriptional modulation. Given the lack of validation at the protein level, it is possible that Aldh18a1 and Sacs are among the ~80% of genes that do not show a high correlation between transcript and protein levels (notwithstanding the aminophylline-*Aldh18a1* clear dose-dependent increase in expression). Thus, on average, one would have to study five drug-gene hits for every hit that validates at both transcript and protein levels. One means of addressing the poor mRNA-protein level correlation is to bypass mRNA

measurement, proceeding directly to a proteomic screen of FDA approved BBB penetrant drugs. In the past ten years, significant advances in mass spectrometry has allowed analysis of the entire proteome in a cell or tissue (Altelaar et al. 2013). As the cost of "next generation" proteomic analysis decreases, whole-proteome mass spectrometry screens could complement transcriptome drug screens conducted for rare genetic disease therapies.

HU inhibition of PLK1; a role in HbF induction

Both type II hits investigated at the *in vitro* protein level showed a positive correlation with mRNA; PLK1 and CCNB1 protein were downregulated in U87 and lymphoblastoid cells (LCLs) treated for 8 hrs with HU. Thus, PLK1 and CCNB1 transcript and protein expression appear to correlate with each other, in contrast to the lack of correlation seen for the three type I hits outlined above. Interestingly PLK1 and CCNB1 protein at 12 and 24 hours rebound to normal levels possibly due to a cellular homeostatic response. In keeping with this and in contrast to the observed inhibition of the FOXM1 regulatory network, the FOXM1 protein itself showed increases after 8 hours of HU achieving statistical significance in LCLs by 12 and 24 hours. Therefore, a clear discrepancy exists between the upregulation of FOXM1, a transcription factor that activates expression of PLK1 and CCNB1, and the down-regulation of PLK1 and CCNB1. Interestingly, a rebound in both PLK1 and CCNB1 transcript and protein levels occurs several hours after the increase in FOXM1 protein levels. The FOXM1 transcription factor is an important regulator of genes involved in every stage of the cell

cycle as well as playing a crucial role in activation of genes that detect and repair DNA damage (Wierstra 2013). In response to DNA damaging agents, such as HU, FOXM1 protein levels increase due to protein stabilization (Tan et al. 2007). In the LCLs then, it appears that the HU-mediated DNA-damage-induced increase of FOXM1 protein leads to greater transcriptional activation of *PLK1* and *CCNB1* (after 12 hours) and a concomitant rebound of gene and protein expression. The ChIP assay for FOXM1, confirmed that although FOXM1 protein is upregulated by HU treatment, there is reduced occupancy of the PLK1 promoter by FOXM1 8 hours after drug treatment. The reduction in PLK1 promoter occupancy by FOXM1 likely contributes to the PLK1 downregulation observed after treatment with HU. Although the ChIP assay establishes a direct link between the effect of HU on FOXM1 and the expression of the PLK1 gene in blood-derived LCLs, more work needs to be done to determine why, although upregulated, FOXM1 binding to the PLK1 promoter is decreased.

Although HU is a well-characterized DNA damage inducing agent, no direct connection has previously been made between HU and FOXM1 or *PLK1* gene expression (Petermann et al. 2010; Hanada et al. 2007). My novel delineation of the impact of HU on *PLK1* in LCLs may lead to increased understanding of the drug's effect in treating sickle cell disease (SCD). The fetal isoform of hemoglobin has been recognized as one of the first disease rescuing paralogs; those with both persistence of fetal hemoglobin and SCD have milder disease making the drug induction of fetal hemoglobin a therapeutic target (Atweh & Fathallah 2010). Hydroxyurea, currently the only clinically approved treatment by the FDA for sickle cell disease (SCD) increases the production of fetal

hemoglobin (HbF) in reticulocytes (Platt et al. 1984; Charache et al. 1992). Although various mechanisms have been proposed to explain the HU-mediated upregulation of HbF, a clear understanding of the process remains elusive (Pule et al. 2015). One theory is that the cytotoxic impact of HU on the mitotically active and rapidly dividing late erythroid precursors leads to the recruitment of early erythroid precursors that harbour larger amounts of HbF (Mabaera et al. 2008). Recently, a negative correlation between the "mitotic roles of polo-like kinases" (PLK) canonical pathway and the SCD rescuing paralog gene fetal hemoglobin (HbF) has been reported; Li et al (2012) showed a significant association between PLK pathway upregulation and reduction of HbF as well an increase of HbA (adult hemoglobin) (Li et al. 2012). If upregulation of the PLK pathway is associated with the reduction in HbF, it may be that the converse, downregulation of PLK1, leads to upregulation of HbF. Such a relationship could partially underpin the mechanism of HU-induced HbF induction with respect to the cytotoxicity model. PLK1 knockdown in non-cancerous cells has been shown to cause minor reductions in cell cycle progression and some cytotoxicity (Liu et al. 2006; Raab et al. 2011). Low level cytotoxicity induced by HU-mediate PLK1 downregulation may be enough to enhance the recruitment of early erythroid precursors increasing the number of circulating HbF positive red blood cells. Regardless of the specific mechanism, my novel characterization of HU as an inhibitor of the FOXM1 network, and more specifically PLK1, is another avenue to explore in the study of HU-mediated upregulation of HbF. To further elucidate the interaction between PLK1 and the effect of HU in SCD, it will be necessary to confirm the effect of PLK1 inhibition on HbF levels in

an erythrocyte cell model. Since PLK1 upregulation and increased activity is observed in a number of cancers, several PLK1 inhibiting compounds have been developed and one (Volasertib) was granted orphan drug status by the FDA for the treatment of acute myeloid leukemia (2014)(Lee et al. 2015; Van den Bossche et al. 2016; Bose & Grant 2014). Thus, confirmation of the role of PLK1 inhibition in upregulating HbF might lead to the investigation of PLK1 inhibitors as a novel and repurposed treatment strategy for SCD.

Correlation of *in vitro* to *in vivo* transcriptional modulation

In vitro model systems designed to rapidly quantify small molecule impact on target cells or tissue have dramatically increased the throughput and thus efficiency of drug screening and lead compound identification when compared to whole-animal drug screening. However, *in vitro* systems do not fully replace the testing of pharmaceutical effects in whole animals before a drug can be approved for humans (Davila et al. 1998). In addition, HTS has not been as effective as predicted and some blame the use of HTS technology for the fall off in new drug discovery (Macarron et al. 2011; Scannell et al. 2012). In this regard, the gene expression of five hits that validated in 3H cortical cultures were tested in cortex and cerebellum of mice or rats treated with the respective drug candidates. The attrition of type I drug-gene hits in transitioning from the *in vitro* to the *in vivo* setting was high; none of the five *in vitro* validated drug-gene hits were validated. However, the sample size for each drug treatment was smaller than necessary to establish statistical significance in the *in vivo* context and one drug-gene hit, the T3/T4 mediated induction of Pmp22 mRNA, deserves further study. Factors underlying the

failure of drug-gene hit validation may include poor BBB penetrance, high heterogeneity of *in vivo* results that masks small changes in gene expression, and mammalian homeostatic mechanisms that work to moderate marked pharmacologic increase or decrease of transcripts *in vivo*.

Dosing considerations

To optimize the clinical relevance of the animal studies, the highest drug dose tolerated without obvious side-effects reported found in the literature for mice or rats were used. Although BBB penetrant drugs was the focus of the Neuron Screen, incomplete knowledge of pharmacokinetics and pharmacodynamics for the drugs led to the use of some that had poor BBB penetrance. In these cases, a toxic even lethal dose would be necessary to achieve sufficient CNS levels. One example is diflunisal; unlike other NSAIDS (nonsteroidal anti-inflammatory drugs), it is poorly BBB penetrant, with an estimated blood/CSF ratio of 100:1 (FDA); thus, it is unlikely that the 50mg/kg oral dose resulted in brain levels of 200μM needed to upregulate the *Slc6a8* gene *in vitro*. Although aminophylline and nilotinib are more BBB penetrant than diflunisal, they too probably did not reach intra-cerebral concentrations close to those that induced the target genes *in vitro*. In the case of nilotinib, the only dose that upregulated the *Sacs* gene *in vitro* also resulted in ER-stress related gene activation in 3H cultures. Because the blood/CSF ratio for nilotinib is a minimum of 10:1, it is likely that a lethal dose of nilotinib would have to be used to obtain the 10μM intracerebral dose in mice. Extrapolating recent clinical trial results that report 10nM concentration of nilotinib in

the CSF of patients treated with 5mg/kg, a whopping 5kg/kg of nilotinib would be needed to achieve the 10 μM dose (Pagan et al. 2016).

Technological advances in mass spectrometry have enabled improved identification of trace amounts of small molecule drugs such as diflunisal and nilotinib in the brain (Swales et al. 2015). Such methods could be used to determine if these drugs did penetrate the brain of mice used to validate type I hits. However, given that the dose needed to achieve drug levels relevant to the *in vitro* doses from the Neuron Screen would be highly toxic, it is unlikely that mass spectrometry would reveal sufficient brain levels of these drugs.

Transcriptional effect of corticosteroids on *Hsd17b4*

Both BMZ and DEX were type I drug hits for the gene *Hsd17b4*. Although only BMZ led to a statistically significant upregulation of *Hsd17b4* (35%), DEX also upregulated *Hsd17b4* by about 30% and thus was chosen for *in vivo* experiments given the extensive literature dealing with this corticosteroid. I believed that since both steroids upregulated *Hsd17b4*, albeit modestly, and DEX freely crosses the mammalian BBB (Balis et al. 1987), that it was likely that the gene was a valid target of the corticosteroids. Dexamethasone upregulates target genes by activating glucocorticoid receptor binding to glucocorticoid response elements (GRE); a consensus GRE is a 15-nucleotide sequence composed of palindrome hexamers separated by three nucleotides (John et al. 2011; Hudson et al. 2013). Corroborating the hypothesis of a glucocorticoid class-mediated upregulation of *Hsd17b4*, I identified a consensus GRE

(GGTACA(TAG)TGTTCC) in the promoter sequence of the mouse *Hsd17b4* gene. However, none of the three mice treated with 1mg/kg of DEX showed *Hsd17b4* induction in the cortex. Interestingly, *Fkbp5,* a well-characterized biomarker of DEX activity was upregulated in all three mice brains while another target gene *Klf9* was up in two of three mice suggesting at least some DEX activity in the brain (Vermeer et al. 2003; Frahm et al. 2016). It is worth noting that *Fkbp5* is highly responsive to glucocorticoids (Zannas & Binder 2014); thus, it is possible that *Hsd17b4* is simply less sensitive than *Fkbp5* and *Klf9*, and would respond to a higher steroid dose. It is also possible that, unlike the two induced genes, *Hsd17b4* is only induced in specific regions of the brain or is exclusively neuronal thus *Hsd17b4* induction is diluted by various cell types present in brain (e.g. oligodendrocytes, endothelial and microglial cells) but not 3H cortical cultures. Nonetheless whether poor sensitivity or localized expression or both underlies the lack of induction, it seems unlikely, given the lack of *in vivo* validation, that DEX could be further developed to treat the *Hsd17b4* hypomorphic condition D-bifunctional protein deficiency (Moller et al. 2001).

T4/T3 mediated upregulation of Pmp22

The T4/T3 mediated upregulation of *Pmp22* is the most promising type I hit. As previously discussed, T3 and T4 led to statistically significant transcriptional upregulation of *Pmp22* in mouse cerebrocortical cultures and T3 upregulated *Pmp22* in rat DRG cultures. Additionally, a positive trend was observed for the number of PMP22-positive cells in DRG cultures in response to thyroid hormone. With regards *in vivo*

testing, the initial dose of 10μg/kg T4 had no effect on *Pmp22* in mouse cortex or sciatic nerve, indeed none of the thyroid hormone-responsive genes tested were upregulated in mouse cortex. I next initiated 300μg/kg treatment on 3 rats. Unfortunately, one rat died suddenly after 6 days of treatment; given no necropsy was performed, it was not possible to know the cause of death. The dose was far below the oral LD50 of 20mg/kg, thus the most likely cause was acute peritonitis linked to the daily intraperitoneal injections. In the remaining two SD rats treated with 300μg/kg of T4, a correlation between *Pmp22* transcript and thyroid hormone responsive genes was observed. *Pmp22* was upregulated by ~35% while the thyroid hormone positive transcriptional markers *Hr*, *Sema7a*, and *Shh* were more than doubled in one rat; although the second didn't show *Pmp22* upregulation, thyroid-responsive genes were not upregulated to the same extent either, with only a 2-fold upregulation in *Hr* observed. Thus, the pilot study indicated that 300μg/kg T4 might upregulate *Pmp22*. Addition of biological replicates with 3 more vehicle treated rats and 7 more T4 treated rats led to a ~25% statistically significant upregulation of *Pmp22* in cortex of T4-treated rats.

Given HNPP chiefly impacts the peripheral nervous system, one caveat is that for thyroid hormone to be effective it would have to induce PNS *Pmp22.* In this regard, preliminary results indicate the, in contrast to the cortex, *Pmp22* remains unchanged in sciatic nerve of 300μg/kg T4 treated rats. This difference could be the result of a tissue specific conversion of T4 to T3 (active thyroid hormone) present in the cortex but not the PNS. Also, since PNS *Pmp22* transcript levels are roughly four orders of magnitude higher than in CNS, it may be that the addition of supraphysiologic thyroid hormone

has no significant effect on the already maximal expression of *Pmp22* in sciatic nerve. In this regard, the *in vitro* dose-curve response of *Pmp22* shows that above a threshold dose of thyroid hormone, *Pmp22* expression is relatively constant. This interpretation aligns with a previous study of the impact of thyroid hormone on the X-ALD rescuing paralog ABCD2 albeit on cortical not PNS tissue; rats treated with 1mg/kg of thyroid hormone expressed an increase in liver *Pmp22* but no significant increase in brain *Pmp22* (liver has even lower *Pmp22* expression than brain). As well, hepatic ABCD2 is upregulated in hyperthyroid and downregulated in hypothyroid rats (Fourcade et al. 2003). A similar analysis of *Pmp22* expression in hypothyroid rats might be useful in clarifying the role of thyroid hormone in its expression. Although thyroid hormone has not been directly related to Pmp22 expression, there are a few reports that indicate a potential link including a clinical case-study describing an initial worsening of HNPP symptoms with hypothyroidism followed by a return to baseline with thyroid hormone analog treatment in one individual (Kaneko et al. 2013). Another interesting link between Pmp22 and thyroid hormone is the Pharnext patent (2008(Patent #EP20080854467)) proposing the use of the thyroid peroxidase inhibiting drug methimazole to treat Charcot-Marie-Tooth disease type 1A. CMT1A in direct contrast to HNPP is caused by increased not decreased *PMP22* gene dosage. The patent was based on the thyroid hormone activation of EGR2, a gene closely linked to promyelination genes (including *Pmp22*), expression in Schwann cells (Mercier et al. 2001). Although the use of methimazole to treat CMT1A was not pursued (likely due to side-effect profile of

methimazole) the proposal of thyroid hormone down-regulation to treat CMT1A is noteworthy.

Analysis of the nucleotide sequence upstream of the *Pmp22* transcription start site revealed the presence of several putative thyroid hormone response elements. Future work exploring the effect of thyroid hormone on PMP22 could also be performed in transgenic *Pmp22* heterozygous mice, shown to have disrupted PNS myelination (there is no published rat model of HNPP) (Adlkofer et al. 1997). Hypothyroidism could also be induced in HNPP-like mice to determine if the phenotype worsens, and if so whether it improves upon return to a euthyroid state.

However, regardless of the outcome in the study of HNPP-like transgenic mice, it is likely that thyroid hormone treatment would not be feasible for HNPP patients given supraphysiologic thyroid hormone causes a range of serious side-effects likely more detrimental than actual HNPP symptoms. One avenue that could be explored is the use of thyromimetic compounds as has been proposed for X-ALD (Genin et al. 2009; M. D. Hartley et al. 2017). Nonetheless, my results could be used to inform the care of HNPP and CMT1A patients; physicians should be particularly vigilant about thyroid hormone imbalances in these patients as a change in thyroid hormone could alter the levels of PMP22 and cause a worsening of symptoms. Finally, our observation that T4 upregulates *Pmp22* in the cortex of a treated rat serves as a validation of our use of transcriptional drug-gene hits from the Neuron Screen, supporting cerebrocortical

culture as a suitable model to potentially repurpose drugs for rare neurogenetic diseases.

In contrast to the *in vivo* validation rate of type I hits, type II hits, upregulated by DEX as downstream targets of the PPARD network, showed a ~40% validation rate. In cortex of mice treated with DEX, 3 out of the 7 genes downstream of PPARD showed statistically significant upregulation and another 2 genes showed a trend toward upregulation. The gene that showed the largest increase was *Mfsd2a* (major facilitator super family domain containing 2a). The *MFSD2A* gene encodes a lysophosphatidylcholine transporter that is primarily expressed in endothelium of the BBB and the major transporter of docosahexaenoic acid (DHA), an omega-3 fatty acid essential to early brain development (Nguyen et al. 2014). Bi-allelic inactivating mutation of the *MFSD2A* gene is linked to the recently identified autosomal recessive primary microcephaly 15 (MCPH15), a rare neurogenetic disease. The disease, identified in two consanguineous families, is characterized by microcephaly, severe developmental delay and intellectual disability, hypotonia, hyperreflexia, spastic quadriparesis, seizures and infant mortality (Guemez-Gamboa et al. 2015). Several other consanguineous families with *MFSD2A* missense point mutations and a milder phenotype have also been identified (Alakbarzade et al. 2015). Thus, residual function of the *MFSD2A*-related protein lysophosphatidylcholine symporter 1 (NLS1) might benefit from *MFSD2A* upregulation. Given the clinical presentation of this disease, microcephaly developing prenatally, post-natal treatment would likely be too little too late even for the less severe syndrome, as microcephaly is already present when the

child is born. Thus, treatment for the disease would likely have to occur *in utero*. Antenatal DEX is commonly used to treat expecting mothers at risk of pre-term labour, reducing the risk of respiratory distress and intraventricular haemorrhage in premature infants (Crowley 1995; Peffer et al. 2015). However there is some evidence that antenatal DEX causes developmental abnormalities to an infant's hypothalamus (Karemaker et al. 2008; Wyrwoll & Holmes 2012) although fetal toxicity may be partially reversed after birth by use of a specialized diet (Mark et al. 2014). Further work is needed to determine if DEX treatment can also lead to upregulation of NLS1 in brain endothelium, and rescue of brain DHA levels. Independent research teams have recently developed *Mfsd2a* knockout mice that have significantly reduced brain size and reduced levels of DHA in the brain (Berger et al. 2012; Moritake et al. 2017). Mice heterozygous for *Mfsd2a* could be treated with DEX and a radiolabeled DHA assay used to determine the impact of DEX in recovery of brain DHA levels (Chen et al. 2015). In addition to the potential therapeutic utility of DEX for the treatment of MCPH15, the upregulation of *Mfsd2a* by DEX in mouse brain highlights the utility of the URA function of IPA to identify network-related drug-gene hits that have strong potential of validation *in vivo*.

General comparison of methods to identify drug-gene hits

The *in vitro* to *in vivo* validation discrepancy observed in the validation of Neuron Screen hits is a well-known phenomenon and is responsible for a large amount of the pre-clinical attrition rate of drug development. The fact that type II hits had a much greater *in vivo* validation rate than type I hits aligned with *in vitro* transcriptional

validation rates for type I and type II hits. Thus, it seems that the primary problem with the individual drug-gene method of identifying hits is that these hits are less likely to be physiologically relevant because they have not been identified in concert with other hits that share a similar pathway for transcript modulation. Of the 32 type I hits tested only one, T3/T4-*Pmp22* has been identified as a candidate for further research.

URA analysis, in addition to providing increased drug-gene hit validation, can also reveal novel transcriptional drug effects with potential to uncover new mechanisms of action of current disease treatments. The discovery of a relationship between the SCD drug HU and PLK1 in human lymphoblastoid cells is one such example. In terms of developing treatments for diseases according to the gene-dosage model, the network-based approach for drug-gene hit validation that I have described requires a minimum of five downstream genes. Because roughly a fifth of human protein coding genes have been linked to a rare genetic disease, statistically, at least one rare disease gene should be included in each robust "activated" or "inhibited" network. For example, the HU inhibited FOXM1 network contained three genes (*BUB1, CENPE, CENPF*) that are linked to rare genetic diseases of which *CENPE* is connected to a neurogenetic disease (rare form of microcephaly) (OMIM, 2017). The statistics hold true for the DEX activated PPARD network that had three downstream genes associated with rare genetic diseases (*MERTK, MFSD2A, LRP5*), two that are linked to neurogenetic diseases (*MERTK, MFSD2A*). However, simply the presence of a rare disease gene in a URA network does not mean that a potentially therapeutic drug-gene hit has been identified; the gene must be modulated in the right direction to afford rescue of protein dosage. In the case of the

FOXM1-associated rare disease genes, because mutations in these genes all lead to loss or complete lack of protein, treatment with HU to downregulate would likely only amplify the gene dosage problems. Thus drug-gene hits from the FOXM1 network are not potential therapeutic options. Another situation that can occur with network-related drug-gene hits is identification of a drug-gene hit for a non-clinically tractable disease. This was the case for the DEX-induced PPARD network that included the rare microcephaly related gene *Mfsd2a*. Although in theory, upregulation of *MFSD2A* could lead to benefit for patients suffering from the less-severe form of MCPH15, such a treatment strategy remains very unlikely as previously mentioned. This example highlights the caveat that even if a rare disease gene is identified in a URA network, the disease characteristics may not be suitable for post-natal pharmacologic gene-therapy. With the URA approach, the researcher is not able to prioritize diseases to study based on their clinical tractability or likelihood of successful therapeutic intervention by transcriptional up/downregulation of the gene.

Currently, transcriptional modulation approaches to drug repurposing for genetic diseases mostly screen for drug-induced transcriptional signatures orthogonal to the signature of the disease of interest. Disease signatures used for these purposes can vary in number of genes with gene signatures of 23(liver) & 49(spleen) identified for Gaucher disease, while a signature called the M30 network comprised of 320 downregulated genes has been identified for epilepsies (Yuen et al. 2012)(Delahaye-Duriez et al. 2016). The drawback of the orthogonal gene signature method for drug candidate identification is that it may not bring insight into the specific metabolic and

biochemical pathways that the drug candidate is acting on to achieve the therapeutic purpose. Furthermore, the signature-based model does not directly target the disease-causing gene, but instead attempts to correct the broad downstream transcriptional impact of the gene mutation. Thus, there is no guarantee that using the orthogonal signature strategy will lead to a treatment that is curative as opposed to symptomatic. In contrast, the URA network approach identifies drug-gene hits related to individual genes, potentially enabling development of a drug with direct gene-level therapeutic effect. Also, URA networks from the Neuron Screen can be used to describe neuronal transcriptional response to drugs initially tested and approved for systemic diseases. For example, nilotinib, currently approved as a systemic treatment for Philadelphia chromosome positive chronic myeloid leukemia, is being investigated as a potential treatment for Alzheimer's disease (Lonskaya et al. 2014; Lonskaya et al. 2015), and Parkinson's disease (Karuppagounder et al. 2014; Hebron et al. 2013; Pagan et al. 2016). Because the effects of nilotinib on the brain have not been widely studied, identification of URA networks that are modulated by nilotinib in the Neuron Screen could help uncover novel drug effects specific to the CNS and thus predict unintended effects of nilotinib in the brain. Yet care would have to be employed to avoid simply identifying ER-stress induced networks that tend to be upregulated by transcriptionally active drugs such as nilotinib.

Recommendations for identifying drug-gene hits from transcriptomic data

Comparison of the validation rate of type I and type II hits from the Neuron Screen has provided insight into the most effective means to interrogate transcriptome-wide differential expression data which can be summarized in the following three recommendations:

1. If the goal of a transcriptome-wide experiment is discovery of novel differentially expressed transcripts between two experimental conditions (e.g. disease vs. healthy tissues, drug treated vs. control), then the network-directed and not single gene induction method is recommended. Most of the drug-treated transcriptomes analyzed in IPA by URA had at least one network robustly modulated with the statistical metrics similar to those observed for the FOXM1 and PPARD networks. Thus, if one has little prior knowledge of the transcriptional profile of a condition or drug treatment, the network-directed approach can be the starting point for discovery of biological effects related to the condition or drug treatment of interest.
2. If the goal of a research project is to find drugs that affect the rare genetic disease causative genes with no disease-specific interest, then the network-directed approach is also recommended. Because a fifth of the human protein-coding genes have been linked to a rare genetic disease, one or even more disease linked gene(s) is likely to be present in each modulated network, since the minimum network size included five gene targets of the upstream regulator. Furthermore,

because most rare genetic disease associated genes have not been well characterized, any discovery of associated transcriptional regulation or drugs may prove beneficial in guiding future research and possibly, even leading to new treatments for a rare disease.

3. If the goal of a research project is to identify drugs that cause up/downregulation of a specific gene of interest associated with a given disorder, the network-directed approach is not recommended. Although the gene-directed approach would intuitively be the best option in this situation, this approach requires a multiplicity of hits to identify any valid drug-gene interaction. One option in this situation would be to use a much larger number of drugs such as the Pfizer chemogenomics drug library of 3000 compounds (3000 compounds covering 1000 targets) recently shared with our lab. Then the URA network approach could be used to identify which compound modulates the gene of interest in a more robust context. Another approach to identify drug targets for a specific genes would be a more traditional bioinformatic approach such as the one used to identify the CMT1A drug PXT3003 that downregulates *Pmp22* (Chumakov et al. 2014). Of course, the approaches are not mutually exclusive and depending on the screen design the opportunity to use both such as I have done may be the best option.

Concluding Remarks

The International Rare Diseases Research Consortium (IRDiRC) was established in 2011 with a goal to provide diagnoses for most rare diseases and develop 200 new or repurposed therapies for rare diseases by the year 2020 (NIH; European Commission n.d.). There exists a largely unmet need for accessible, disease-altering treatments for rare genetic diseases. The cost of new drug development is a deterrent for most companies to fund rare disease drug development, while many drugs that are developed for rare diseases are marketed at prohibitive prices. Repurposing clinic-ready drugs can reduce the cost and time needed to develop a drug for the clinic. The drug-induced transcriptional modulation – i.e. pharmacological gene therapy method for treating rare diseases is an attractive approach to treat rare genetic diseases at the source of the problem, yet only one study has successfully repurposed a drug to treat a monogenic disease by gene modulation. Drug candidate identification remains a similar challenge for transcriptional repurposing as for new drug discovery. Increasing the initial validity of drug-gene interactions at the *in vitro* stage of development can make drug discovery more cost-effective because fewer drugs will fail after the much more expensive pre-clinical and clinical trials. Comparison of the drug-directed and network-directed approaches for candidate identification has highlighted the benefits and caveats of each technique. The network-directed approach using the URA tool of IPA is a novel technique for candidate identification that leads to high rates of validation *in vitro* and *in vivo*. I have shown that this technique can be used to identify novel mechanisms of drug transcriptional effects and can be used to identify novel drug-gene hits related to rare

genetic disease genes. Utility of this technique should be further explored for the study of transcriptomic data from other experiments and other platforms (e.g. microarray).

Roughly two thirds of rare monogenic diseases have a neurological phenotype. The lack of treatments for rare neurogenetic diseases is even greater than for some other rare diseases that have benefited from various enzyme replacement and hematopoietic stem cell transplants. I have developed the first transcriptome-wide FDA-approved drug screen and created a publicly available searchable database to enable other researchers to use the data from the Neuron Screen. Although initially intended for the study of rare neurogenetic disease genes, the data contained in the Neuron Screen Database may also be used to study the effect of FDA approved drugs on genes related to more common neurological diseases. Additionally, the Neuron Screen Database could be used to identify potential neurological side-effects of the FDA-approved drugs included in the screen as well as those simply studying drug impact on the central nervous system.

In conclusion, I have identified a transcriptional relationship between thyroid hormone and the *Pmp22* gene and between DEX and the *Mfsd2a* gene, both interactions have not been reported and may lead to better understanding of the transcriptional environment of these rare neurogenetic disease genes. Also, the HU mediated downregulation of PLK1 may prove useful for improved understanding of the mechanism of HU in the context of treating SCD patients. Finally, it is my hope that the results and analysis presented in this thesis will assist researchers in the study of

transcriptomic data, and lead to the discovery of novel treatments for the millions of patients afflicted with life-altering and life-shortening rare neurogenetic diseases.

REFERENCES

Abatangelo, L. et al., 2009. Comparative study of gene set enrichment methods. *BMC bioinformatics*, 10, p.275.

Adlkofer, K. et al., 1997. Heterozygous peripheral myelin protein 22-deficient mice are affected by a progressive demyelinating tomaculous neuropathy. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, 17(12), pp.4662–4671.

Agholme, L. et al., 2010. An in vitro model for neuroscience: Differentiation of SH-SY5Y cells into cells with morphological and biochemical characteristics of mature neurons. *Journal of Alzheimer's Disease*, 20(4), pp.1069–1082.

Alakbarzade, V. et al., 2015. A partially inactivating mutation in the sodium-dependent lysophosphatidylcholine transporter MFSD2A causes a non-lethal microcephaly syndrome. *Nature genetics*, 47(7), pp.814–817.

Altelaar, A.F.M., Munoz, J. & Heck, A.J.R., 2013. Next-generation proteomics: towards an integrative view of proteome dynamics. *Nature Reviews Genetics*, 14(1), pp.35–48. Available at: http://dx.doi.org/10.1038/nrg3356.

Andrade-Moraes, C.H. et al., 2013. Cell number changes in Alzheimer's disease relate to dementia, not to plaques and tangles. *Brain : a journal of neurology*, 136(Pt 12), pp.3738–3752.

Anon, 2012. EpiTect ChIP qPCR Primers. Available at: http://www.sabiosciences.com/chipqpcrsearch.php?app=TFBS&qs=1500777363 [Accessed January 1, 2017].

Araque, A. & Navarrete, M., 2010. Glial cells in neuronal network function. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*, 365(1551), pp.2375–2381.

Aras, M.A., Hartnett, K.A. & Aizenman, E., 2008. Assessment of cell viability in primary neuronal cultures. *Current protocols in neuroscience*, Chapter 7, p.Unit 7.18.

Ashburn, T.T. & Thor, K.B., 2004. Drug repositioning: identifying and developing new uses for existing drugs. *Nature reviews.Drug discovery*, 3(8), pp.673–683.

Attarian, S. et al., 2014. An exploratory randomised double-blind and placebo-controlled phase 2 study of a combination of baclofen, naltrexone and sorbitol (PXT3003) in patients with Charcot-Marie-Tooth disease type 1A. *Orphanet journal of rare diseases*, 9, p.199.

Attarian, S. et al., 2016. Erratum to: An exploratory randomised double-blind and placebo-controlled phase 2 study of a combination of baclofen, naltrexone and sorbitol (PXT3003) in patients with Charcot-Marie-Tooth disease type 1A. *Orphanet*

journal of rare diseases, 11(1), p.92.

Atweh, G. & Fathallah, H., 2010. Pharmacologic induction of fetal hemoglobin production. *Hematology/oncology clinics of North America*, 24(6), pp.1131–1144.

Ayme, S. et al., 1998. Information on rare diseases: the Orphanet project. *La Revue de medecine interne fondee par la Societe nationale francaise de medecine interne*, 19 Suppl 3, p.376S–377S.

Azevedo, F.A.C. et al., 2009. Equal numbers of neuronal and nonneuronal cells make the human brain an isometrically scaled-up primate brain. *The Journal of comparative neurology*, 513(5), pp.532–541.

Baird, P.A. et al., 1988. Genetic disorders in children and young adults: a population study. *American journal of human genetics*, 42(5), pp.677–693.

Balis, F.M. et al., 1987. Differences in cerebrospinal fluid penetration of corticosteroids: possible relationship to the prevention of meningeal leukemia. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology*, 5(2), pp.202–207.

Ballabh, P., Braun, A. & Nedergaard, M., 2004. The blood-brain barrier: an overview: structure, regulation, and clinical implications. *Neurobiology of disease*, 16(1), pp.1–13.

Banker, G.A., 1980. Trophic interactions between astroglial cells and hippocampal neurons in culture. *Science (New York, N.Y.)*, 209(4458), pp.809–810.

Barakat-Walter, I., 1999. Role of thyroid hormones and their receptors in peripheral nerve regeneration. *Journal of neurobiology*, 40(4), pp.541–559.

Barral, S. & Kurian, M.A., 2016. Utility of Induced Pluripotent Stem Cells for the Study and Treatment of Genetic Diseases: Focus on Childhood Neurological Disorders. *Frontiers in molecular neuroscience*, 9, p.78.

von Bartheld, C.S., Bahney, J. & Herculano-Houzel, S., 2016. The search for true numbers of neurons and glial cells in the human brain: A review of 150 years of cell counting. *The Journal of comparative neurology*, 524(18), pp.3865–3895.

Bauernfeind, A.L. et al., 2015. Evolutionary Divergence of Gene and Protein Expression in the Brains of Humans and Chimpanzees. *Genome biology and evolution*, 7(8), pp.2276–2288.

Beaulieu, C.L. et al., 2012. A generalizable pre-clinical research approach for orphan disease therapy. *Orphanet journal of rare diseases*, 7, p.39.

Bellocchio, E.E. et al., 2000. Uptake of glutamate into synaptic vesicles by an inorganic phosphate transporter. *Science (New York, N.Y.)*, 289(5481), pp.957–960.

Berger, J.H., Charron, M.J. & Silver, D.L., 2012. Major facilitator superfamily domain-containing protein 2a (MFSD2A) has roles in body growth, motor function, and

lipid metabolism. *PloS one*, 7(11), p.e50629.

Boerma, M. et al., 2006. Comparative expression profiling in primary and immortalized endothelial cells: changes in gene expression in response to hydroxy methylglutaryl-coenzyme A reductase inhibition. *Blood coagulation & fibrinolysis : an international journal in haemostasis and thrombosis*, 17(3), pp.173–180.

Bose, P. & Grant, S., 2014. Orphan drug designation for pracinostat, volasertib and alvocidib in AML. *Leukemia research*, 38(8), pp.862–865.

Van den Bossche, J. et al., 2016. Spotlight on Volasertib: Preclinical and Clinical Evaluation of a Promising Plk1 Inhibitor. *Medicinal research reviews*, 36(4), pp.749–786.

Boycott, K.M. et al., 2017. International Cooperation to Enable the Diagnosis of All Rare Genetic Diseases. *American journal of human genetics*, 100(5), pp.695–705.

Boycott, K.M. et al., 2013. Rare-disease genetics in the era of next-generation sequencing: discovery to translation. *Nature reviews.Genetics*, 14(10), pp.681–691.

Brewer, G.J. et al., 1993. Optimized survival of hippocampal neurons in B27-supplemented Neurobasal, a new serum-free medium combination. *Journal of neuroscience research*, 35(5), pp.567–576.

Cahoy, J.D. et al., 2008. A transcriptome database for astrocytes, neurons, and oligodendrocytes: a new resource for understanding brain development and function. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, 28(1), pp.264–278.

Callizot, N. et al., 2011. A new long term in vitro model of myelination. *Experimental cell research*, 317(16), pp.2374–2383.

Camp, K.M., Lloyd-Puryear, M.A. & Huntington, K.L., 2012. Nutritional treatment for inborn errors of metabolism: indications, regulations, and availability of medical foods and dietary supplements using phenylketonuria as an example. *Molecular genetics and metabolism*, 107(1–2), pp.3–9.

Campillos, M. et al., 2008. Drug target identification using side-effect similarity. *Science (New York, N.Y.)*, 321(5886), pp.263–266.

Carter, C.O., 1977. Monogenic disorders. *Journal of medical genetics*, 14(5), pp.316–320.

Chan, E.T. et al., 2009. Conservation of core gene expression in vertebrate tissues. *Journal of biology*, 8(3), p.33.

Chance, P.F. et al., 1993. DNA deletion associated with hereditary neuropathy with liability to pressure palsies. *Cell*, 72(1), pp.143–151.

Charache, S. et al., 1992. Hydroxyurea: effects on hemoglobin F production in patients with sickle cell anemia. *Blood*, 79(10), pp.2555–2565.

Chen, C.T. et al., 2015. Plasma non-esterified docosahexaenoic acid is the major pool supplying the brain. *Scientific reports*, 5, p.15791.

Chindelevitch, L. et al., 2012. Causal reasoning on biological networks: interpreting transcriptional changes. *Bioinformatics (Oxford, England)*, 28(8), pp.1114–1121.

Chong, C.R. & Sullivan Jr, D.J., 2007. New uses for old drugs. *Nature*, 448(7154), pp.645–646.

Chumakov, I. et al., 2014. Polytherapy with a combination of three repurposed drugs (PXT3003) down-regulates Pmp22 over-expression and improves myelination, axonal and functional parameters in models of CMT1A neuropathy. *Orphanet journal of rare diseases*, 9, p.201.

Consortium, M.G.S. et al., 2002. Initial sequencing and comparative analysis of the mouse genome. *Nature*, 420(6915), pp.520–562.

Corey, D.R., 2017. Nusinersen, an antisense oligonucleotide drug for spinal muscular atrophy. *Nature neuroscience*, 20(4), pp.497–499.

Corsello, S.M. et al., 2017. The Drug Repurposing Hub: a next-generation drug library and information resource. *Nature medicine*, 23(4), pp.405–408.

Courtin, F. et al., 2005. Thyroid hormone deiodinases in the central and peripheral nervous system. *Thyroid : official journal of the American Thyroid Association*, 15(8), pp.931–942.

Crowley, P.A., 1995. Antenatal corticosteroid therapy: a meta-analysis of the randomized trials, 1972 to 1994. *American journal of obstetrics and gynecology*, 173(1), pp.322–335.

D'Aiuto, L. et al., 2014. Large-scale generation of human iPSC-derived neural stem cells/early neural progenitor cells and their neuronal differentiation. *Organogenesis*, 10(4), pp.365–377.

Daub, A., Sharma, P. & Finkbeiner, S., 2009. High-content screening of primary neurons: ready for prime time. *Current opinion in neurobiology*, 19(5), pp.537–543.

Davila, J.C. et al., 1998. Predictive value of in vitro model systems in toxicology. *Annual review of pharmacology and toxicology*, 38, pp.63–96.

Delahaye-Duriez, A. et al., 2016. Rare and common epilepsies converge on a shared gene regulatory network providing opportunities for novel antiepileptic drug discovery. *Genome biology*, 17(1), p.245.

Desnick, R.J. & Schuchman, E.H., 2002. Enzyme replacement and enhancement therapies: lessons from lysosomal disorders. *Nature reviews. Genetics*, 3(12), pp.954–966.

Dharssi, S. et al., 2017. Review of 11 national policies for rare diseases in the context of

key patient needs. *Orphanet journal of rare diseases*, 12(1), p.63.

DiMasi, J.A., Grabowski, H.G. & Hansen, R.W., 2016. Innovation in the pharmaceutical industry: New estimates of R&D costs. *Journal of health economics*, 47, pp.20–33.

Dowling, J.J., 2016. Eteplirsen therapy for Duchenne muscular dystrophy: skipping to the front of the line. *Nature reviews. Neurology*, 12(12), pp.675–676.

Dudley, J.T., Deshpande, T. & Butte, A.J., 2011. Exploiting drug-disease relationships for computational drug repositioning. *Briefings in bioinformatics*, 12(4), pp.303–311.

Dye, D.E. et al., 2011. The impact of single gene and chromosomal disorders on hospital admissions in an adult population. *Journal of community genetics*, 2(2), pp.81–90.

Eglen, R.M., Gilchrist, A. & Reisine, T., 2008. The use of immortalized cell lines in GPCR screening: the good, bad and ugly. *Combinatorial chemistry & high throughput screening*, 11(7), pp.560–565.

Evers, M.M., Toonen, L.J.A. & van Roon-Mom, W.M.C., 2015. Antisense oligonucleotides in therapy for neurodegenerative disorders. *Advanced drug delivery reviews*, 87, pp.90–103.

Farooq, F. et al., 2013. Celecoxib increases SMN and survival in a severe spinal muscular atrophy mouse model via p38 pathway activation. *Human Molecular Genetics*, 22(17).

Farooq, F. et al., 2011. Prolactin increases SMN expression and survival in a mouse model of severe spinal muscular atrophy via the STAT5 pathway. *Journal of Clinical Investigation*, 121(8).

Ferrua, F., Brigida, I. & Aiuti, A., 2010. Update on gene therapy for adenosine deaminase-deficient severe combined immunodeficiency. *Current opinion in allergy and clinical immunology*, 10(6), pp.551–556.

Fourcade, S. et al., 2003. Thyroid hormone induction of the adrenoleukodystrophy-related gene (ABCD2). *Molecular pharmacology*, 63(6), pp.1296–1303.

Frahm, K.A. et al., 2016. Research Resource: The Dexamethasone Transcriptome in Hypothalamic Embryonic Neural Stem Cells. *Molecular endocrinology (Baltimore, Md.)*, 30(1), pp.144–154.

Garbay, B. et al., 2000. Myelin synthesis in the peripheral nervous system. *Progress in neurobiology*, 61(3), pp.267–304.

Genin, E.C. et al., 2009. Induction of the adrenoleukodystrophy-related gene (ABCD2) by thyromimetics. *The Journal of steroid biochemistry and molecular biology*, 116(1–2), pp.37–43.

Ghazalpour, A. et al., 2011. Comparative analysis of proteome and transcriptome variation in mouse. *PLoS genetics*, 7(6), p.e1001393.

Giacca, M. & Zacchigna, S., 2012. Virus-mediated gene delivery for human gene therapy. *Journal of controlled release : official journal of the Controlled Release Society*, 161(2), pp.377–388.

Gil-Ibanez, P. et al., 2017. Global Transcriptome Analysis of Primary Cerebrocortical Cells: Identification of Genes Regulated by Triiodothyronine in Specific Cell Types. *Cerebral cortex (New York, N.Y. : 1991)*, 27(1), pp.706–717.

Giotti, B., Joshi, A. & Freeman, T.C., 2017. Meta-analysis reveals conserved cell cycle transcriptional network across multiple human cell types. *BMC genomics*, 18(1), p.30.

Gordon, J., Amini, S. & White, M.K., 2013. General overview of neuronal cell culture. *Methods in molecular biology (Clifton, N.J.)*, 1078, pp.1–8.

Greig, L.C. et al., 2013. Molecular logic of neocortical projection neuron specification, development and diversity. *Nature reviews.Neuroscience*, 14(11), pp.755–769.

Guemez-Gamboa, A. et al., 2015. Inactivating mutations in MFSD2A, required for omega-3 fatty acid transport in brain, cause a lethal microcephaly syndrome. *Nature genetics*, 47(7), pp.809–813.

Guo, J. & Anton, E.S., 2014. Decision making during interneuron migration in the developing cerebral cortex. *Trends in cell biology*, 24(6), pp.342–351.

Gupta, A., Tsai, L.H. & Wynshaw-Boris, A., 2002. Life is a journey: a genetic look at neocortical development. *Nature reviews.Genetics*, 3(5), pp.342–355.

Halawi, M.J., Lark, R.K. & Fitch, R.D., 2015. Neuromuscular Scoliosis: Current Concepts. *Orthopedics*, 38(6), pp.e452-6.

Hanada, K. et al., 2007. The structure-specific endonuclease Mus81 contributes to replication restart by generating double-strand DNA breaks. *Nature structural & molecular biology*, 14(11), pp.1096–104. Available at: http://www.ncbi.nlm.nih.gov/pubmed/17934473.

Hartley, M.D. et al., 2017. A Thyroid Hormone-Based Strategy for Correcting the Biochemical Abnormality in X-Linked Adrenoleukodystrophy. *Endocrinology*, 158(5), pp.1328–1338.

Hartley, T. et al., 2017. Whole-exome sequencing is a valuable diagnostic tool for inherited peripheral neuropathies: outcomes from a cohort of 50 families. *Clinical genetics*.

Heather, J.M. & Chain, B., 2016. The sequence of sequencers: The history of sequencing DNA. *Genomics*, 107(1), pp.1–8.

Hebron, M.L., Lonskaya, I. & Moussa, C.E.-H., 2013. Nilotinib reverses loss of dopamine neurons and improves motor behavior via autophagic degradation of alpha-

synuclein in Parkinson's disease models. *Human molecular genetics*, 22(16), pp.3315–3328.

Hendry, S.H. et al., 1987. Numbers and proportions of GABA-immunoreactive neurons in different areas of monkey cerebral cortex. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, 7(5), pp.1503–1519.

Herculano-Houzel, S., 2014. The glia/neuron ratio: how it varies uniformly across brain structures and species and what that means for brain physiology and evolution. *Glia*, 62(9), pp.1377–1391.

Herculano-Houzel, S., Mota, B. & Lent, R., 2006. Cellular scaling rules for rodent brains. *Proceedings of the National Academy of Sciences of the United States of America*, 103(32), pp.12138–12143.

Hertzberg, R.P. & Pope, A.J., 2000. High-throughput screening: new technology for the 21st century. *Current opinion in chemical biology*, 4(4), pp.445–451.

Hoy, S.M., 2017. Nusinersen: First Global Approval. *Drugs*, 77(4), pp.473–479.

Huang, D.W. et al., 2007. DAVID Bioinformatics Resources: expanded annotation database and novel algorithms to better extract biology from large gene lists. *Nucleic acids research*, 35(Web Server issue), pp.W169-75.

Hudson, W.H., Youn, C. & Ortlund, E.A., 2013. The structural basis of direct glucocorticoid-mediated transrepression. *Nature structural & molecular biology*, 20(1), pp.53–58.

Hughes, J.P. et al., 2011. Principles of early drug discovery. *British journal of pharmacology*, 162(6), pp.1239–1249.

Illes, S. et al., 2009. Niche-dependent development of functional neuronal networks from embryonic stem cell-derived neural populations. *BMC neuroscience*, 10, p.93.

Imming, P., Sinning, C. & Meyer, A., 2006. Drugs, their targets and the nature and number of drug targets. *Nature reviews. Drug discovery*, 5(10), pp.821–834.

Isaacson, J.S. & Scanziani, M., 2011. How inhibition shapes cortical activity. *Neuron*, 72(2), pp.231–243.

Iskar, M. et al., 2013. Characterization of drug-induced transcriptional modules: towards drug repositioning and functional understanding. *Molecular systems biology*, 9, p.662.

Jacobsen, K.X. & Staines, W.A., 2004. Vibration enhancement of slide-mounted immunofluorescence staining. *Journal of neuroscience methods*, 137(1), pp.71–77.

Ji, R.-R. et al., 2009. Transcriptional profiling of the dose response: a more powerful approach for characterizing drug activities. *PLoS computational biology*, 5(9), p.e1000512.

Jin, G. & Wong, S.T.C., 2014. Toward better drug repositioning: prioritizing and integrating existing methods into efficient pipelines. *Drug discovery today*, 19(5), pp.637–644.

John, S. et al., 2011. Chromatin accessibility pre-determines glucocorticoid receptor binding patterns. *Nature genetics*, 43(3), pp.264–268.

Kaneko, K. et al., 2013. Hereditary neuropathy with liability to pressure palsy emerging after hypothyroidism. *Neurology and Clinical Neuroscience*, 1(4), pp.160–161.

Karemaker, R. et al., 2008. Neonatal dexamethasone treatment for chronic lung disease of prematurity alters the hypothalamus-pituitary-adrenal axis and immune system activity at school age. *Pediatrics*, 121(4), pp.e870-8.

Karuppagounder, S.S. et al., 2014. The c-Abl inhibitor, nilotinib, protects dopaminergic neurons in a preclinical animal model of Parkinson's disease. *Scientific reports*, 4, p.4874.

Kaur, P. et al., 2014. Expression profiling of RNA transcripts during neuronal maturation and ischemic injury. *PloS one*, 9(7), p.e103525.

Keiser, M.J. et al., 2009. Predicting new molecular targets for known drugs. *Nature*, 462(7270), pp.175–181.

Kelsom, C. & Lu, W., 2013. Development and specification of GABAergic cortical interneurons. *Cell & bioscience*, 3(1), p.19.

Kemp, S. et al., 1998. Gene redundancy and pharmacological gene therapy: implications for X-linked adrenoleukodystrophy. *Nature medicine*, 4(11), pp.1261–1268.

Krämer, A. et al., 2014. Causal analysis approaches in ingenuity pathway analysis. *Bioinformatics*, 30(4), pp.523–530.

Krishna, A. et al., 2014. Systems genomics evaluation of the SH-SY5Y neuroblastoma cell line as a model for Parkinson's disease. *BMC genomics*, 15, p.1154.

Kubota, Y. et al., 2011. Selective coexpression of multiple chemical markers defines discrete populations of neocortical GABAergic neurons. *Cerebral cortex (New York, N.Y. : 1991)*, 21(8), pp.1803–1817.

Kumar, S. et al., 2013. A benzophenanthridine alkaloid, chelerythrine induces apoptosis in vitro in a Dalton's lymphoma. *Journal of cancer research and therapeutics*, 9(4), pp.693–700.

Kunkel, S.D. et al., 2011. mRNA expression signatures of human skeletal muscle atrophy identify a natural compound that increases muscle mass. *Cell metabolism*, 13(6), pp.627–638.

Kyrylkova, K. et al., 2012. Detection of apoptosis by TUNEL assay. *Methods in molecular biology (Clifton, N.J.)*, 887, pp.41–47.

Lam, R.S. et al., 2017. Functional Maturation of Human Stem Cell-Derived Neurons in Long-Term Cultures. *PloS one*, 12(1), p.e0169506.

Lamb, J., 2007. The Connectivity Map: a new tool for biomedical research. *Nature reviews. Cancer*, 7(1), pp.54–60.

Lamb, J. et al., 2006. The Connectivity Map: using gene-expression signatures to connect small molecules, genes, and disease. *Science (New York, N.Y.)*, 313(5795), pp.1929–1935.

Lau, P. et al., 2008. Identification of dynamically regulated microRNA and mRNA networks in developing oligodendrocytes. *The Journal of neuroscience : the official journal of the Society for Neuroscience*, 28(45), pp.11720–11730.

Lee, K.S. et al., 2015. Recent Advances and New Strategies in Targeting Plk1 for Anticancer Therapy. *Trends in pharmacological sciences*, 36(12), pp.858–877.

Lesuisse, C. & Martin, L.J., 2002. Long-term culture of mouse cortical neurons as a model for neuronal development, aging, and death. *Journal of neurobiology*, 51(1), pp.9–23.

Li, B. et al., 2012. Characterization of the transcriptome profiles related to globin gene switching during in vitro erythroid maturation. *BMC genomics*, 13, p.153.

Li, W.W. et al., 2001. Type 2 deiodinase in the peripheral nervous system: induction in the sciatic nerve after injury. *Neuroscience*, 107(3), pp.507–518.

Liu, X., Lei, M. & Erikson, R.L., 2006. Normal cells, but not cancer cells, survive severe Plk1 depletion. *Molecular and cellular biology*, 26(6), pp.2093–2108.

Lonskaya, I. et al., 2014. Nilotinib-induced autophagic changes increase endogenous parkin level and ubiquitination, leading to amyloid clearance. *Journal of molecular medicine (Berlin, Germany)*, 92(4), pp.373–386.

Lonskaya, I. et al., 2015. Nilotinib and bosutinib modulate pre-plaque alterations of blood immune markers and neuro-inflammation in Alzheimer's disease models. *Neuroscience*, 304, pp.316–327.

Lovatt, D. et al., 2014. Transcriptome in vivo analysis (TIVA) of spatially defined single cells in live tissue. *Nature methods*, 11(2), pp.190–196.

Luhmann, H.J. et al., 2016. Spontaneous Neuronal Activity in Developing Neocortical Networks: From Single Cells to Large-Scale Interactions. *Frontiers in neural circuits*, 10, p.40.

Lupski, J.R. et al., 1991. DNA duplication associated with Charcot-Marie-Tooth disease type 1A. *Cell*, 66(2), pp.219–232.

Mabaera, R. et al., 2008. A cell stress signaling model of fetal hemoglobin induction: what doesn't kill red blood cells may make them stronger. *Experimental hematology*, 36(9), pp.1057–1072.

Macarron, R., 2006. Critical review of the role of HTS in drug discovery. *Drug discovery today*, 11(7–8), pp.277–279.

Macarron, R. et al., 2011. Impact of high-throughput screening in biomedical research. *Nature reviews. Drug discovery*, 10(3), pp.188–195.

Macarron, R. & Hertzberg, R.P., 2011. Design and implementation of high throughput screening assays. *Molecular biotechnology*, 47(3), pp.270–285.

Le Magueresse, C. & Monyer, H., 2013. GABAergic interneurons shape the functional maturation of the cortex. *Neuron*, 77(3), pp.388–405.

Maguire, C.A. et al., 2014. Gene therapy for the nervous system: challenges and new strategies. *Neurotherapeutics : the journal of the American Society for Experimental NeuroTherapeutics*, 11(4), pp.817–839.

Marioni, J.C. et al., 2008. RNA-seq: an assessment of technical reproducibility and comparison with gene expression arrays. *Genome research*, 18(9), pp.1509–1517.

Mark, P.J. et al., 2014. Rescue of glucocorticoid-programmed adipocyte inflammation by omega-3 fatty acid supplementation in the rat. *Reproductive biology and endocrinology : RB&E*, 12, p.39.

Marshall, E., 1999. Do-it-yourself gene watching. *Science (New York, N.Y.)*, 286(5439), pp.444–447.

Martinez-Nunez, R.T. & Sanford, J.R., 2016. Studying Isoform-Specific mRNA Recruitment to Polyribosomes with Frac-seq. *Methods in molecular biology (Clifton, N.J.)*, 1358, pp.99–108.

McArt, D.G. & Zhang, S.D., 2011. Identification of candidate small-molecule therapeutics to cancer by gene-signature perturbation in connectivity mapping. *PloS one*, 6(1), p.e16382.

McGettigan, P.A., 2013. Transcriptomics in the RNA-seq era. *Current opinion in chemical biology*, 17(1), pp.4–11.

McKusick, V.A., 2007. Mendelian Inheritance in Man and its online version, OMIM. *American Journal of Human Genetics*, 80(4), pp.588–604.

Mears, A.J. et al., 2017. Mining the transcriptome for rare disease therapies: a comparison of the efficiencies of two data mining approaches and a targeted cell-based drug screen. *npj Genomic Medicine*, 2(1), p.14.

Meinecke, D.L. & Peters, A., 1987. GABA immunoreactive neurons in rat visual cortex. *The Journal of comparative neurology*, 261(3), pp.388–404.

Melchiorri, D. et al., 2013. Regulatory evaluation of Glybera in Europe - two committees, one mission. *Nature reviews. Drug discovery*, 12(9), p.719.

Mercier, G., Turque, N. & Schumacher, M., 2001. Rapid effects of triiodothyronine on immediate-early gene expression in Schwann cells. *Glia*, 35(2), pp.81–89.

Miller, F.D. & Gauthier, A.S., 2007. Timing is everything: making neurons versus glia in the developing cortex. *Neuron*, 54(3), pp.357–369.

Miyoshi, G. & Fishell, G., 2011. GABAergic interneuron lineages selectively sort into specific cortical layers during early postnatal development. *Cerebral cortex (New York, N.Y. : 1991)*, 21(4), pp.845–852.

Moller, G. et al., 2001. Molecular basis of D-bifunctional protein deficiency. *Molecular and cellular endocrinology*, 171(1–2), pp.61–70.

Molyneaux, B.J. et al., 2007. Neuronal subtype specification in the cerebral cortex. *Nature reviews. Neuroscience*, 8(6), pp.427–437.

Moritake, H. et al., 2017. A mouse model reveals that Mfsd2a is critical for unfolded protein response upon exposure to tunicamycin. *Human cell*, 30(2), pp.88–97.

Mullard, A., 2017. 2016 FDA drug approvals. *Nature reviews. Drug discovery*, 16(2), pp.73–76.

Muller, P.Y. & Milton, M.N., 2012. The determination and interpretation of the therapeutic index in drug development. *Nature reviews. Drug discovery*, 11(10), pp.751–761.

Nau, R., Sorgel, F. & Eiffert, H., 2010. Penetration of drugs through the blood-cerebrospinal fluid/blood-brain barrier for treatment of central nervous system infections. *Clinical microbiology reviews*, 23(4), pp.858–883.

Nedergaard, M., 1994. Direct signaling from astrocytes to neurons in cultures of mammalian brain cells. *Science (New York, N.Y.)*, 263(5154), pp.1768–1771.

Nestor, C.E. et al., 2015. Rapid reprogramming of epigenetic and transcriptional profiles in mammalian culture systems. *Genome biology*, 16, p.11.

Nguyen, L.N. et al., 2014. Mfsd2a is a transporter for the essential omega-3 fatty acid docosahexaenoic acid. *Nature*, 509(7501), pp.503–506.

NIH, Nervous System Diseases. Available at: https://rarediseases.info.nih.gov/diseases/diseases-by-category/17/nervous-system-diseases [Accessed June 12, 2017].

NIH; European Commission, International Rare Diseases Research Consortium. *2017*. Available at: http://www.irdirc.org/about-us/goals/ [Accessed January 1, 2017].

Nobbio, L. et al., 2006. Axonal damage and demyelination in long-term dorsal root ganglia cultures from a rat model of Charcot-Marie-Tooth type 1A disease. *The European journal of neuroscience*, 23(6), pp.1445–1452.

Noble, W.S., 2009. How does multiple testing correction work? *Nature biotechnology*, 27(12), pp.1135–1137.

Nosengo, N., 2016. Can you teach old drugs new tricks? *Nature*, 534(7607), pp.314–316.

Notterpek, L., Snipes, G.J. & Shooter, E.M., 1999. Temporal expression pattern of peripheral myelin protein 22 during in vivo and in vitro myelination. *Glia*, 25(4), pp.358–369.

Obien, M.E.J. et al., 2014. Revealing neuronal function through microelectrode array recordings. *Frontiers in neuroscience*, 8, p.423.

Odawara, A. et al., 2016. Physiological maturation and drug responses of human induced pluripotent stem cell-derived cortical neuronal networks in long-term culture. *Scientific reports*, 6, p.26181.

Otto, F. et al., 2003. Cryopreserved rat cortical cells develop functional neuronal networks on microelectrode arrays. *Journal of neuroscience methods*, 128(1–2), pp.173–181.

van Paassen, B.W. et al., 2014. PMP22 related neuropathies: Charcot-Marie-Tooth disease type 1A and Hereditary Neuropathy with liability to Pressure Palsies. *Orphanet journal of rare diseases*, 9, p.38.

Pagan, F. et al., 2016. Nilotinib Effects in Parkinson's disease and Dementia with Lewy bodies. *Journal of Parkinson's disease*, 6(3), pp.503–517.

Paul, S.M. et al., 2010. How to improve R&D productivity: the pharmaceutical industry's grand challenge. *Nature reviews. Drug discovery*, 9(3), pp.203–214.

Peffer, M.E. et al., 2015. Minireview: the impact of antenatal therapeutic synthetic glucocorticoids on the developing fetal brain. *Molecular endocrinology (Baltimore, Md.)*, 29(5), pp.658–666.

Petermann, E. et al., 2010. Hydroxyurea-Stalled Replication Forks Become Progressively Inactivated and Require Two Different RAD51-Mediated Pathways for Restart and Repair. *Molecular Cell*, 37(4), pp.492–502.

Petkov, P.M. et al., 2004. An efficient SNP system for mouse genome scanning and elucidating strain relationships. *Genome research*, 14(9), pp.1806–1811.

Pfrieger, F.W. & Barres, B.A., 1997. Synaptic efficacy enhanced by glial cells in vitro. *Science (New York, N.Y.)*, 277(5332), pp.1684–1687.

Philippakis, A.A. et al., 2015. The Matchmaker Exchange: A Platform for Rare Disease Gene Discovery. *Human mutation*, 36(10), pp.915–921.

Platt, O.S. et al., 1984. Hydroxyurea enhances fetal hemoglobin production in sickle cell anemia. *The Journal of clinical investigation*, 74(2), pp.652–656.

Plenge, R.M., 2016. Disciplined approach to drug discovery and early development. *Science translational medicine*, 8(349), p.349ps15.

Pule, G.D. et al., 2015. A systematic review of known mechanisms of hydroxyurea-induced fetal hemoglobin for treatment of sickle cell disease. *Expert review of hematology*, 8(5), pp.669–679.

Raab, M. et al., 2011. Toxicity modelling of Plk1-targeted therapies in genetically engineered mice and cultured primary mammalian cells. *Nature communications*, 2, p.395.

Raeymaekers, P. et al., 1991. Duplication in chromosome 17p11.2 in Charcot-Marie-Tooth neuropathy type 1a (CMT 1a). The HMSN Collaborative Research Group. *Neuromuscular disorders : NMD*, 1(2), pp.93–97.

Ramakrishnan, S.R. et al., 2009. Integrating shotgun proteomics and mRNA expression data to improve protein identification. *Bioinformatics (Oxford, England)*, 25(11), pp.1397–1403.

Ratko, T.A. et al., 2013. *Enzyme-Replacement Therapies for Lysosomal Storage Diseases*, Rockville (MD).

Scannell, J.W. et al., 2012. Diagnosing the decline in pharmaceutical R&D efficiency. *Nature reviews. Drug discovery*, 11(3), pp.191–200.

Schock, S.C. et al., 2012. Development of dissociated cryopreserved rat cortical neurons in vitro. *Journal of neuroscience methods*, 205(2), pp.324–333.

Sharma, P. et al., 2012. High-throughput screening in primary neurons. *Methods in enzymology*, 506, pp.331–360.

Shaw, K.L. et al., 2017. Clinical efficacy of gene-modified stem cells in adenosine deaminase-deficient immunodeficiency. *The Journal of clinical investigation*, 127(5), pp.1689–1699.

Shigemizu, D. et al., 2012. Using functional signatures to identify repositioned drugs for breast, myelogenous leukemia and prostate cancer. *PLoS computational biology*, 8(2), p.e1002347.

Sirota, M. et al., 2011. Discovery and preclinical validation of drug indications using compendia of public gene expression data. *Science translational medicine*, 3(96), p.96ra77.

Sun, D. & Jakobs, T.C., 2012. Structural remodeling of astrocytes in the injured CNS. *The Neuroscientist : a review journal bringing neurobiology, neurology and psychiatry*, 18(6), pp.567–588.

Sun, W. et al., 2012. Transcriptome atlases of mouse brain reveals differential expression across brain regions and genetic backgrounds. *G3 (Bethesda, Md.)*, 2(2), pp.203–211.

Sutherland, J.J. et al., 2016. Assessing Concordance of Drug-Induced Transcriptional Response in Rodent Liver and Cultured Hepatocytes. *PLoS computational biology*, 12(3), p.e1004847.

Swales, J.G. et al., 2015. Mapping drug distribution in brain tissue using liquid extraction surface analysis mass spectrometry imaging. *Analytical chemistry*, 87(19), pp.10146–10152.

Swamidass, S.J., 2011. Mining small-molecule screens to repurpose drugs. *Briefings in bioinformatics*, 12(4), pp.327–335.

Tan, Y., Raychaudhuri, P. & Costa, R.H., 2007. Chk2 mediates stabilization of the FoxM1 transcription factor to stimulate expression of DNA repair genes. *Molecular and cellular biology*, 27(3), pp.1007–1016.

Tani, K., 2016. Current status of ex vivo gene therapy for hematological disorders: a review of clinical trials in Japan around the world. *International journal of hematology*, 104(1), pp.42–72.

Tasic, B. et al., 2016. Adult mouse cortical cell taxonomy revealed by single cell transcriptomics. *Nature neuroscience*, 19(2), pp.335–346.

Ting, C.-H. et al., 2007. Stat5 constitutive activation rescues defects in spinal muscular atrophy. *Human molecular genetics*, 16(5), pp.499–514.

Tong, X. et al., 2014. Astrocyte Kir4.1 ion channel deficits contribute to neuronal dysfunction in Huntington's disease model mice. *Nature neuroscience*, 17(5), pp.694–703.

Ullian, E.M. et al., 2001. Control of synapse number by glia. *Science (New York, N.Y.)*, 291(5504), pp.657–661.

Vaskova, E.A. et al., 2013. "Epigenetic memory" phenomenon in induced pluripotent stem cells. *Acta naturae*, 5(4), pp.15–21.

Vermeer, H. et al., 2003. Glucocorticoid-induced increase in lymphocytic FKBP51 messenger ribonucleic acid expression: a potential marker for glucocorticoid sensitivity, potency, and bioavailability. *The Journal of clinical endocrinology and metabolism*, 88(1), pp.277–284.

Wang, Z., Gerstein, M. & Snyder, M., 2009. RNA-Seq: a revolutionary tool for transcriptomics. *Nature reviews.Genetics*, 10(1), pp.57–63.

Watanabe, N. et al., 2015. Re-examination of regulatory opinions in Europe: possible contribution for the approval of the first gene therapy product Glybera. *Molecular therapy. Methods & clinical development*, 2, p.14066.

Wierstra, I., 2013. *The transcription factor FOXM1 (Forkhead box M1): Proliferation-specific expression, transcription factor function, target genes, mouse models, and normal biological*

roles 1st ed., Copyright © 2013 Elsevier Inc. All rights reserved. Available at: http://dx.doi.org/10.1016/B978-0-12-407173-5.00004-2.

Wirth, T., Parker, N. & Yla-Herttuala, S., 2013. History of gene therapy. *Gene*, 525(2), pp.162–169.

Wise, C.A. et al., 1993. Molecular analyses of unrelated Charcot-Marie-Tooth (CMT) disease patients suggest a high frequency of the CMTIA duplication. *American journal of human genetics*, 53(4), pp.853–863.

Witherspoon, L. et al., 2015. Sodium Channel Inhibitors Reduce DMPK mRNA and Protein. *Clinical and Translational Science*, 8(4).

Wojcik, S.M. et al., 2006. A shared vesicular carrier allows synaptic corelease of GABA and glycine. *Neuron*, 50(4), pp.575–587.

Wonders, C.P. & Anderson, S.A., 2006. The origin and specification of cortical interneurons. *Nature reviews.Neuroscience*, 7(9), pp.687–696.

Wyrwoll, C.S. & Holmes, M.C., 2012. Prenatal excess glucocorticoid exposure and adult affective disorders: a role for serotonergic and catecholamine pathways. *Neuroendocrinology*, 95(1), pp.47–55.

Xia, X. & Wong, S.T., 2012. Concise review: a high-content screening approach to stem cell research and drug discovery. *Stem cells (Dayton, Ohio)*, 30(9), pp.1800–1807.

Xu, X. & Zhong, Z., 2013. Disease modeling and drug screening for neurological diseases using human induced pluripotent stem cells. *Acta pharmacologica Sinica*, 34(6), pp.755–764.

Yang, Y. et al., 2011. Molecular comparison of GLT1+ and ALDH1L1+ astrocytes in vivo in astroglial reporter mice. *Glia*, 59(2), pp.200–207.

Yap, Y.W. et al., 2013. Gene expression profiling of rotenone-mediated cortical neuronal death: evidence for inhibition of ubiquitin-proteasome system and autophagy-lysosomal pathway, and dysfunction of mitochondrial and calcium signaling. *Neurochemistry international*, 62(5), pp.653–663.

Yoon, P.W. et al., 1997. Contribution of birth defects and genetic diseases to pediatric hospitalizations. A population-based study. *Archives of Pediatrics & Adolescent Medicine*, 151(11), pp.1096–1103.

Young, C.S. & Pyle, A.D., 2016. Exon Skipping Therapy. *Cell*, 167(5), p.1144.

Yuen, T. et al., 2012. Disease-drug pairs revealed by computational genomic connectivity mapping on GBA1 deficient, Gaucher disease mice. *Biochemical and biophysical research communications*, 422(4), pp.573–577.

Yuzyuk, T. et al., 2016. Effect of dietary lysine restriction and arginine supplementation in two patients with pyridoxine-dependent epilepsy. *Molecular genetics and*

metabolism, 118(3), pp.167–172.

Zannas, A.S. & Binder, E.B., 2014. Gene-environment interactions at the FKBP5 locus: sensitive periods, mechanisms and pleiotropism. *Genes, brain, and behavior*, 13(1), pp.25–37.

Zhao, S. et al., 2014. Comparison of RNA-Seq and microarray in transcriptome profiling of activated T cells. *PloS one*, 9(1), p.e78644.

Zheng-Bradley, X. et al., 2010. Large scale comparison of global gene expression patterns in human and mouse. *Genome biology*, 11(12), p.R124.

APPENDICES

Appendix I: Spreadsheet of drugs used for the Neuron Screen. These are the drug names that can be used to search the Neuron Screen Database. The exact drug name must be used for a successful search.

Neuron Screen #	Drug name	Mol. Wt.	Test dose [final] µM
1	Rosiglitazone	357.4	8.00E-01
2	Betaxolol HCl	343.9	1.00E-01
3	Pindolol	248.3	3.00E-01
4	Naltrexone HCl	377.9	1.00E-01
5	Zolmitriptan	287.4	3.00E-02
6	Memantine HCl	215.8	7.00E-01
7	Riluzole HCl	270.7	1.80E+00
8	Aminophylline	210.2	9.00E+01
9	Imipramine HCl	316.9	9.00E-01
10	Famotidine	337.5	5.00E-01
11	Clemastine Fumarate	460.0	4.00E-03
12	Lovastatin	404.5	2.00E-02
13	Sildenafil Citrate	666.7	7.50E-01
14	Escitalopram	324.4	2.00E-01
15	Guanfacine HCl	282.6	4.00E-01
16	Tizanidine HCl	290.2	1.00E-01
17	Carvedilol	406.5	4.00E-01
18	Metformin HCl	165.6	5.00E+00
19	Galantamine HBr	368.3	2.00E-01
20	Diltiazem HCl	451.0	3.00E-01
21	Nifedipine	346.3	4.00E-01
22	Nimodipine	418.5	1.00E-01
23	Verapamil HCl	491.1	5.00E-01
24	Felodipine	384.3	3.00E-02
25	Dexamethasone	392.5	5.00E-01
26	Bumetanide	364.4	5.00E-01
27	Captopril	217.3	2.00E+00

28	Tranylcypromine Hemisulfate	182.2	1.00E+00
29	Moxifloxacin HCl	437.9	1.00E+01
30	Ketoprofen	254.3	1.00E+01
31	Meloxicam	351.4	5.00E+00
32	Simvastatin	418.6	1.00E-02
33	Rifampin (Rifampicin)	822.9	1.00E+01
34	Cetirizine 2HCl	461.8	5.00E-01
35	Pioglitazone HCl	392.9	1.00E+00
36	Rivastigmine Tartrate	400.4	2.50E-02
37	Sulindac	356.4	1.00E+01
38	Zafirlukast	575.7	1.00E-01
39	Diazoxide	230.7	8.00E+01
40	Glyburide	494.0	4.00E-01
41	Celecoxib	381.4	2.00E+00
42	Letrozole	285.3	1.50E-01
43	Anastrozole	293.4	1.50E-01
44	Dolasetron	324.4	1.50E+00
45	Fosinopril Na	585.6	8.00E-01
46	Granisetron HCl	348.9	5.00E-02
47	Oseltamivir Phosphate	410.4	2.00E+00
48	Pramipexole Dihydrochloride Monohydrate	302.3	2.00E-02
49	Acyclovir	225.2	5.00E+00
50	Allopurinol	136.1	1.00E+02
51	Betamethasone	392.5	4.00E-01
52	Buspirone HCl	422.0	9.00E-03
53	Thalidomide	258.2	5.00E+00
54	Citalopram HBr	405.3	3.00E-01
55	Dextromethorphan	271.4	1.00E-01
56	Diclofenac	318.3	8.00E+00
57	Doxycycline Monohydrate	462.4	2.00E+01
58	Enalapril	376.5	1.00E-01
59	Famciclovir	321.3	1.50E+01
60	Fenofibrate	360.8	8.00E+01
61	Finasteride	372.5	3.00E-02
62	Gemfibrozil	250.3	9.00E+01
63	Glimepiride	490.6	1.00E+00
64	Indapamide	365.8	5.00E-01

65	Lisinopril dihydrate	441.5	2.00E-01
66	Loratadine	382.9	1.00E-01
67	Losartan Potassium	461.0	2.50E+00
68	Mebendazole	295.3	3.00E-01
69	Medroxyprogesterone Acetate	386.5	8.00E-03
70	Metoprolol Tartrate	684.8	3.00E-01
71	Mitoxantrone HCl	517.4	5.00E-02
72	Nabumetone	228.3	1.00E+00
73	Omeprazole	345.4	1.00E+01
74	Oxcarbazepine	252.3	1.00E+02
75	Pantoprazole	383.4	1.00E+01
76	Paroxetine HCl	365.8	1.00E-01
77	Ramipril	416.5	1.00E-01
78	Propranolol HCl	295.8	1.00E+00
79	Spironolactone	416.6	5.00E-01
80	Tamsulosin HCl	445.0	4.00E-02
81	Telmisartan	514.6	2.00E+00
82	Terazosin HCl	423.9	2.00E-01
83	Ketotifen Fumarate	425.5	8.00E+00
84	Naloxone HCl	363.9	1.00E-01
85	Fluoxetine HCl	345.8	1.00E+00
86	Ondansetron	293.4	1.00E+00
87	Misoprostol	382.5	1.00E-03
88	Aprepitant	534.4	2.00E+00
89	Cyproheptadine HCl Sesquihydrate	350.9	1.00E-01
90	Acarbose	645.6	1.50E-01
91	Acebutolol HCl	372.9	5.00E+00
92	Acetylcysteine	163.2	1.50E+01
93	Almotriptan	335.5	2.00E-01
94	Ambrisentan	378.4	1.50E+00
95	Benztropine Mesylate	403.5	4.00E-01
96	Biperiden HCl	347.9	3.00E-01
97	Bisoprolol Fumarate	767.0	1.00E-01
98	Brompheniramine Maleate	435.3	3.00E-02
99	Bupropion	239.7	4.00E-01
100	Clonazepam	315.7	3.00E-01
101	Cortisone Acetate	402.5	6.00E-01

102	Darifenacin HBr	426.6	4.00E-03
103	Dicyclomine HCl	309.5	3.00E-01
104	Dimenhydrinate	470.0	2.00E-01
105	Doxepin HCl	315.8	5.00E-01
106	Duloxetine HCl	333.9	4.00E-01
107	Everolimus	958.2	8.00E-03
108	Febuxostat	316.4	1.00E+01
109	Fexofenadine HCl	538.1	5.00E-01
110	Fingolimod	343.9	1.50E+00
111	Fluvoxamine Maleate	434.4	5.00E-01
112	Formoterol	344.4	3.00E-02
113	Hydroxyzine 2HCl	447.8	2.00E-01
114	Irbesartan	428.5	6.00E+00
115	Isosorbide Dinitrate	236.1	2.00E-01
116	Isotretinoin	300.4	5.00E-01
117	Labetalol HCl	364.9	5.00E-01
118	Lamivudine	229.3	8.00E+00
119	Lansoprazole	369.4	2.00E+00
120	Leucovorin Calcium Pentahydrate	601.6	5.00E-01
121	Levothyroxine	798.9	2.00E-01
122	Liothyronine	673.0	1.00E-02
123	Lorazepam	321.2	8.00E-01
124	Loxapine Succinate	445.9	2.00E-01
125	Maraviroc	513.7	1.00E-01
126	Mexiletine HCl	215.7	8.00E+00
127	Mirtazapine	265.4	3.00E-01
128	Modafinil	273.4	5.00E+00
129	Nadolol	309.4	8.00E-01
130	Naratriptan HCl	371.9	1.00E-01
131	Nebivolol HCl	441.9	5.00E-02
132	Nevirapine	266.9	1.00E+01
133	Nilotinib	529.5	1.00E+01
134	Nitrofurantoin	238.2	1.00E+01
135	Nortriptyline HCl	299.8	7.00E-01
136	Orlistat	495.7	1.80E-02
137	Oxaprozin	293.3	1.40E+00
138	Oxazepam	286.7	5.00E+00

139	Oxybutynin Chloride	394.0	5.00E-02
140	Palonosetron HCl	332.9	1.60E-02
141	Perindopril Erbumine	441.6	3.00E-01
142	Phenelzine Sulfate	234.3	1.50E-01
143	Posaconazole	700.8	1.00E+00
144	Prasugrel	373.4	1.50E+00
145	Pregabalin	159.2	2.00E+01
146	Protriptyline HCl	299.8	1.00E+00
147	Pyridostigmine Bromide	261.1	5.00E-01
148	Quinidine HCl dihydrate	324.4	1.50E+01
149	Rabeprazole Na	381.4	1.50E+00
150	Raltegravir	444.4	2.00E-01
151	Rasagiline Mesylate	267.3	3.00E-01
152	Repaglinide	452.6	1.00E-02
153	Rifabutin	847.0	2.00E-01
154	Rizatriptan Benzoate	391.5	2.00E-01
155	Ropinirole HCl	296.8	2.00E-02
156	Selegiline HCl	223.7	4.00E-01
157	Sertraline HCl	342.7	7.00E-01
158	Sitagliptin Phosphate	505.3	3.00E+00
159	Tacrolimus	804.0	2.00E-02
160	Tadalafil	389.4	1.00E+00
161	Temazepam	300.7	5.00E-01
162	Tetrabenazine	371.4	4.00E-02
163	Tolterodine Tartrate	475.6	5.00E-02
164	Trandolapril	430.5	2.00E-02
165	Trazodone HCl	408.3	2.00E+00
166	Triamterene	253.3	4.00E-01
167	Triazolam	343.2	1.00E-01
168	Trihexyphenidyl HCl	337.9	6.00E-01
169	Trimipramine Maleate	410.5	6.00E-01
170	Ursodiol	392.6	1.00E+01
171	Varenicline Tartrate	361.4	3.00E-02
172	Voriconazole	349.3	1.50E+01
173	Ziprasidone	412.9	5.00E-01
174	Chlorpropamide	276.7	4.00E+02
175	Etodolac	287.4	6.00E+01

176	Hydroxyurea	76.1	1.25E+02
177	Lacosamide	250.3	2.00E+01
178	Levocarnitine	161.2	8.00E+01
179	Methazolamide	236.3	1.50E+02
180	Methocarbamol	241.2	1.60E+02
181	Primidone	218.3	5.00E+01
182	Probenecid	285.4	2.00E+02
183	Pyrazinamide	123.1	6.00E+02
184	Rufinamide	238.2	7.00E+01
185	Sulfamethoxazole	253.3	2.00E+02
186	Theophylline	180.2	8.00E+01
187	Topiramate	339.4	2.50E+01
188	Valganciclovir HCl	390.8	2.00E+01
189	Vigabatrin	129.2	5.00E+01
190	Genistein	270.2	1.00E+00
191	Kaempferol	286.2	1.00E-01
192	Benzbromarone	424.1	2.00E+01
193	Quercetin	302.2	1.00E+01
194	Luteolin	286.2	5.00E+00
195	Apigenin	270.2	5.00E+00
196	Deferasirox	373.4	2.00E+01
197	Daidzein	254.2	5.00E-01
198	Acetazolamide	222.3	8.00E+01
199	Mefenamic Acid	241.3	4.00E+01
200	Ethosuximide	141.2	7.00E+02
201	Methsuximide	203.2	2.00E+01
202	Oxtriphylline	283.3	2.00E+01
203	Diflunisal	250.2	2.00E+02
204	Levetiracetam	170.2	2.00E+02
249	Bezafibrate	361.8	2.00E+01
250	Clofibrate	242.7	2.00E+02
251	Ciprofibrate	289.2	5.00E+01
252	CP-448187	UnK	1.00E+00
253	PF-04191834-00	UnK	1.00E+00
254	PF-05416266-00	UnK	1.00E+00
255	SD-7300	UnK	1.00E+00
256	WAY-202041	UnK	1.00E+00

257	PF-04995274-00	UnK	1.00E+00
258	CP-945598	UnK	1.00E+00
259	CP-610927-01	UnK	1.00E+00
260	CE-210666	UnK	1.00E+00
261	CE-326597	UnK	1.00E+00
262	PF-05019702-00	UnK	1.00E+00

Unk: Unknown value

Appendix II: List of mouse primers used for qRT-PCR.

Gene Symbol	Forward Primer	Reverse Primer	Designer
Gapdh	CGTCCCGTAGACAAAATGGT	CTCCTGGAAGATGGTGATGG	A.M.
Hprt1	GCAAACTTTGCATTCCCTGGTT	CAAGGGCATATCCAACAACA	A.M.
Sema7a	GGGCCATCAGCAACTCAAGA	CAGTGGAACAGGGAAGGACG	J.H.
Shh	AGGGGGTTTGGAAAGAGGCG	ACTCCAGGCCACTGGTTCAT	J.H.
Klf9	GGAAACACGCCTCCGAAAAG	AACGGAACTGCTTTTCCCCA	A.M.
Hr	AAGTTTGACATTCGGGGGCA	GAGGCTAGCAGCTCACAGAG	A.M.
Aldh18a1	GGACAGAGTGGACTGATGGC	GATCACTGTTGGGCTCGGC	A.M.
Ddhd2	GTGAGACGGTGTACGTGGTT	CGACCTTGCTCAGTGGATGT	A.M.
Dnmt1	GGAGAAGCAAGTCGGACAGT	CTTCCGACTCTTCCTTGGGT	A.M.
Fgf14	CTGGCAGAGCCTGGTTTTTG	GCTTTCGGGACTGTTTCACC	A.M.
Hsd17b4	GGTTCCTGGGCTGTCATTCA	GTCCGTTTTCCACCAAAGCC	A.M.
Neu1	TTTGGAGTAAGGACGACGGC	CAGAAGACCCCATCTCGCTC	A.M.
Man2b1	GACCATGGGCTCAGACTTCC	AGCCAGTCCAGAACATGTGG	A.M.
Itpr1	TCTTCCATGCCGAGCAAGAG	CTGATCAGGGTCCACTGAGG	A.M.
Mapt	CGGAGACCTCCGATGCTAAG	GTCTCCGATGCCTGCTTCTT	A.M.
Mut	ACGGGGACCATATCCTACCAT	ACGAACTCTGGGGTTGTCTG	A.M.
Nkx2-1	CTTACCAGGACACCATGCGG	GCCATGTTCTTGCTCACGTC	A.M.
Pmp22	GGGATCCTGTTCCTGCACAT	CAAGGCGGATGTGGTACAGT	A.M.
Sacs	TTACAGAGCGGGGCTTTTGT	CGCTCCAGAGTGGTAAGAAGG	A.M.
Slc2a1	GAACGAGGCCCCTGAAGAAA	CCTGTTCACCCATCTTCCCC	A.M.
Slc6a8	AGACTTGGACACGCCAGATG	TTCCTCCAACCAGGGCAATC	A.M.
ILK	TGGACAACACAGAGAACGACC	GGGGTATCATCCCCACGATTC	J.H.
BCL2L1	CCTTGGATCCAGGAGAACGG	TCAGGAACCAGCGGTTGAAG	J.H.
PDK4	CGTACTCCACTGCTCCAACA	ACACCAGTCATCAGCTTCGG	J.H.
MFSD2A	CCTTCACTGACCCTCTGGTG	GAAGCCGTGTGAACTTTCCG	J.H.
MERTK	ACGTTGGTGGATACGTGCAT	CTCTTCCCACTTCTCGGCAG	J.H.
LRP5_1	GCCTTCATGGATGGGACCAA	GCCCGTTCAATGCTATGCAG	J.H.
KYAT3	GTCCTCGGACTCTGCACTTC	AGGATCCGCAGCCAACTTAG	J.H.
Shh	AGGGGGTTTGGAAAGAGGCG	ACTCCAGGCCACTGGTTCAT	J.H.
Sema7a	GGGCCATCAGCAACTCAAGA	CAGTGGAACAGGGAAGGACG	J.H.
Hspa5	CGGCTTCCGATAATCAGCCA	TGTCAGGCGGTTTTGGTCAT	J.H.
Ddit3	GCAGCGACAGAGCCAGAATAA	ACCAGGTTCTGCTTTCAGGT	J.H.
Dnajc3	CCCGACCTCACCAAAGTGAT	AGTGTAATCGGCACCGTCAA	J.H.
Ppp1r15a	CAGCCTGTGAAACATTGCGT	ACTGTTTTTGCCAGCCATGC	J.H.

Appendix III: List of rat primers used for qRT-PCR.

Gene symbol	Forward Primer	Reverse Primer	Designer
Gapdh	GTGCCAGCCTCGTCTCATAG	GAGAAGGCAGCCCTGGTAAC	J.H.
Hprt1	CCAGTCAACGGGGGACATAA	TCCAACAAAGTCTGGCCTGT	J.H.
Pmp22	TGTACCACATCCGCCTTGG	GAGCTGGCAGAAGAACAGGAAC	Chumakov et. al., 2014
Mpz	ACCACTCAGTTCCTTGTCCC	TCCCTGTCCGTGTAAACCAC	J.H.
Sema7a	ATGGCAAGATCCCTCGCTTT	TAACTGGGCCACTCGGGATA	J.H.
Shh	CGGCCGATATGAAGGGAAGA	CCACTGGTTCATCACGGAGA	J.H.
Hr	CCCCTGACCATCGTATGCTT	CAGGGTTCCCTGCTTGTACC	J.H.

Abbreviations: J.H. (Jeremiah Hadwen), A.M. (Alan Mears)

Appendix IV: List of human primers used for qRT-PCR.

Gene Symbol	Forward Primer	Reverse Primer	Designer
GAPDH	GTTCGACAGTCAGCCGCATC	AGTTAAAAGCAGCCCTGGTGA	A.M.
HPRT1	TGACACTGGCAAACCAATGCA	GGTCCTTTTCACCAGCAAGCT	A.M.
ALDH18A1	TGACCTGCAGGGGGTAAATG	CAACAGAAGTGCCACCTTGC	A.M.
SACS	TGATGAAACTCAATACGGAACA	CCCGATTTGGTCACCACTAA	J.H.
BUB1B	CTTCTGGGATGGGTCCTTCTG	GCTCTGAGGCAGCAATCTGT	J.H.
FOXM1	AGCGGCCACCCTACTCTTA	CCCTGGGTCCAGTGGCTTAAA	J.H.
CCNA2	CGTGAAGATGCCCTGGCTTT	AACCAGTCCACGAGGATAGC	J.H.
CENPF	CTGCGGGCAGTTTGAATTAG	CTCTTGTAGGCAGCCCTTCT	J.H.
CENPE	TGAACTCACTTCGTGCTGACT	ACTTCTGCATGCTTAACTAAATTCT	J.H.
PLK1	AGTGTCAATGCCTCCAAGCC	AGAGGATGAGGCGTGTTGAG	J.H.
CCNB1	CCTCTCCAAGCCCAATGGAA	ACTTCCCGACCCAGTAGGTA	J.H.
CDK1	CGCGGAATAATAAGCCGGGA	AGGAACCCCTTCCTCTTCACT	J.H.

Abbreviations: J.H. (Jeremiah Hadwen), A.M. (Alan Mears)

Appendix V: Primer sequences for amplification of ChIP products.

Primary name	Forward Primer	Reverse Primer	PrimerDesigner
-1kb_PLK1	ACATGGGGTAGAAAGGACGG	TGCTTTCAGTTTGCCACAGT	J.H.
TSS_PLK1	GAATTCCTCCTCTCTCGGGG	TTAAAATCCAAACCCGCCCG	J.H.
+500bp_PLK1	AGATCCCGGAGGTCCTAGTG	GCAGCAGAGACTTAGGCACA	J.H.
+1kb_PLK1	CTAGAGAAGGGTGCTGGGAG	GGCATCACTAACCTGGGACT	J.H.

Appendix VI: Upregulation of ER-stress genes in the Neuron Screen. Well-recognized ER-stress genes are shown with the z-score from the Neuron Screen indicating their expression in response to the 6 most transcriptionally active drugs from the neuron scree: fenofibrate, nilotinib, aminophylline, diflunisal, luteolin, and ciprofibrate.

	EIF2AK3	HSPA5	PPP1R15A	DDIT3
Fenofibrate	15.47	11.49	23.35	18.59
Nilotinib	16.94	5.76	10.05	9.11
Aminophylline	-0.83	-2.42	7.13	9.11
Diflunisal	3.24	-0.56	4.34	6.08
Luteolin	6.77	12.14	6.40	4.57
Ciprofibrate	6.13	12.23	8.20	9.41

Appendix VII: Drug doses used for the drug-dose curves.

Drug	Dose 1 [M]	Dose 2 [M]	Dose 3 [M]	Dose 4 [M]	Dose 5 [M]
Fenofibrate	8.00E-05	2.53E-05	8.00E-06	2.53E-06	8.00E-07
Nilotinib	3.16E-05	1.00E-05	3.16E-06	1.00E-06	3.16E-07
Aminophylline	2.52E-04	9.00E-05	2.85E-05	9.00E-06	2.85E-06
Diflunisal	2.00E-04	6.33E-05	2.00E-05	6.33E-06	2.00E-06
T4	6.32E-07	2.00E-07	6.33E-08	2.00E-08	6.33E-09