**Inaugural dissertation**

**for**

**obtaining the doctoral degree**

**of the**

**Combined Faculty of Mathematics, Engineering and Natural Sciences**

**of the**

**Ruprecht - Karls - University**

**Heidelberg**

Presented by

**M.Sc. Nikhil Oommen George**

Born in Dharwad, Karnataka India

Oral examination: 20.12.2022

# Dysregulation of alternative polyadenylation in autism spectrum disorders and its potential use in ASD diagnosis

**Referees**

Prof. Dr. Ana Martin-Villalba

Prof. Dr. Henrik Kaessmann

# Table of Contents

## Abstract

Neurogenesis is the process of generating neurons that can functionally integrate into existing neuronal circuits. In adult mammals, neurogenesis persists in specific niches such as the subventricular zone (SVZ) and the dentate gyrus (DG) which are populated by neural stem/precursor cells that produce neurons during homeostasis and injury. Processes such as alternative polyadenylation seem to play a role in NSC lineage transitions. This study also reports evidence of APA changes along the NSC lineage. Using an *in vitro* NSC lineage model, I was able to show the impact of APA upon translation. Additionally, I identified APLP1 as a potential APA regulator via the APLP1/CPEB4 axis. I observed the dysregulation of APA upon APLP1 knockout. Further, I was able to identify alterations in the composition of cells within the NSC lineage as a consequence of the APLP1 knockout. Dysregulation of neurogenesis can result in many different neurodevelopmental disorders (NDDs) such as autism spectrum disorder (ASD). ASD is characterized by deficits in social communication and the prevalence of restricted and repetitive behaviours. I showed that APLP1 knockout mice presented an ASD-like phenotype and hypothesized that APLP1/CPEB4 mediated APA may be responsible for such a phenotype by dysregulating neurogenesis. Further, owing to the 3'UTR lengthening trend observed in ASD brains, I explored the use of an APA phenotype for ASD diagnosis. The current diagnosis of ASD relies primarily on clinical symptoms that assess behavioural deficits and to some extent morphological abnormalities associated with NDDs. Early diagnosis is essential to allow for necessary intervention at the right time. Molecular diagnosis of ASD is limited to large genetic screens that can identify individual mutations that link to syndromic ASD but fail to diagnose idiopathic ASD. Therefore, I investigated the changes in the APA landscape between ASD patients and controls using whole blood and found great potential in such an application.

## Zusammenfassung

Neurogenese ist der Prozess der Bildung von Neuronen, die sich funktionell in bestehende neuronale Schaltkreise integrieren können. Bei erwachsenen Säugetieren findet die Neurogenese in bestimmten Nischen wie der subventrikulären Zone (SVZ) und dem Gyrus dentatus (DG) statt, die von neuralen Stamm-/Vorläuferzellen bevölkert werden, die bei Homöostase und Verletzungen Neuronen produzieren. Prozesse wie die alternative Polyadenylierung scheinen eine Rolle bei NSV-Linienübergängen zu spielen. In dieser Studie werden auch Hinweise auf APA-Änderungen entlang der NSC-Linie gefunden. Mit Hilfe eines in vitro NSC-Stammbaummodells konnte ich die Auswirkungen der APA auf die Translation nachweisen. Außerdem habe ich APLP1 als potenziellen APA-Regulator über die APLP1/CPEB4-Achse identifiziert. Ich beobachtete die Dysregulation von APA nach Ausschaltung von APLP1. Außerdem konnte ich als Folge des APLP1-Knockouts Veränderungen in der Zusammensetzung der Zellen innerhalb der NSC-Linie feststellen. Eine Dysregulation der Neurogenese kann zu vielen verschiedenen neurologischen Entwicklungsstörungen wie Autismus-Spektrum-Störungen (ASD) führen. ASD ist gekennzeichnet durch Defizite in der sozialen Kommunikation und das Vorherrschen von eingeschränkten und repetitiven Verhaltensweisen. Ich konnte zeigen, dass APLP1-Knockout-Mäuse einen ASD-ähnlichen Phänotyp aufweisen, und stellte die Hypothese auf, dass die APLP1/CPEB4-vermittelte APA für diesen Phänotyp verantwortlich sein könnte, indem sie die Neurogenese dysreguliert. Aufgrund der in ASD-Gehirnen beobachteten Tendenz zur Verlängerung der 3'UTR untersuchte ich außerdem die Verwendung eines APA-Phänotyps für die ASD-Diagnose. Die derzeitige Diagnose von ASD stützt sich in erster Linie auf klinische Symptome, die Verhaltensdefizite und bis zu einem gewissen Grad morphologische Anomalien im Zusammenhang mit NDDs bewerten. Eine frühzeitige Diagnose ist von entscheidender Bedeutung, um die notwendigen Maßnahmen zum richtigen Zeitpunkt ergreifen zu können. Die molekulare Diagnose von ASD beschränkt sich auf große genetische Untersuchungen, mit denen einzelne Mutationen identifiziert werden können, die mit syndromaler ASD in Verbindung stehen, die aber keine idiopathische ASD diagnostizieren können. Daher untersuchte ich die Veränderungen in der APA-Landschaft zwischen ASD-Patienten und Kontrollpersonen anhand von Vollblut und stellte fest, dass eine solche Anwendung großes Potenzial hat.

# 1. Introduction

## 1.1. Neurogenesis in the adult mammalian brain

Neurogenesis involves the production of functional neurons in the brain that arise from specialized stem/precursor cells. While most tissues in adults have a resident pool of stem cells that are essential for homeostasis or repair, for long the brain was believed to be an exception. Traditionally neurogenesis was believed to have been restricted to only embryonic and early post-natal developmental stages. However, the evidence of mammalian adult neurogenesis has been demonstrated across various studies over the years (Altman & Das, 1965; Eriksson et al., 1998; Ming & Song, 2005; Reynolds & Weiss, 1992; Richards et al., 1992). In the Adult murine brain, there are two primary niches that house neural stem cells (NSCs), namely, the ventricular-subventricular zone (V-SVZ) of the lateral ventricle and the sub granular zone (SGZ) in the dentate gyrus of the hippocampus (Ming & Song, 2005). The larger of the two is the SVZ, in which neural stem cells (NSCs) are interspersed between ependymal cells lining the wall of the lateral ventricle.

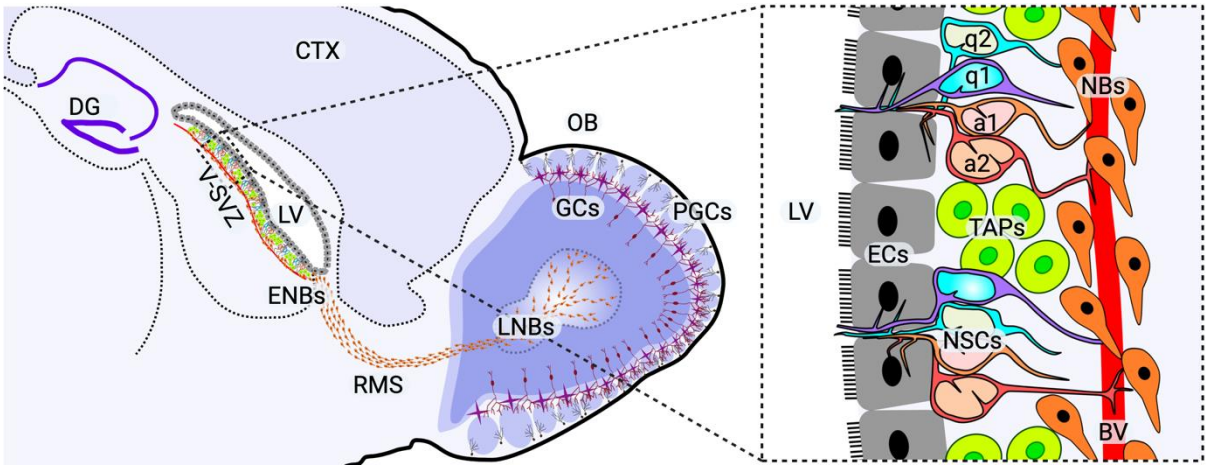


**Figure 1.1:** ***Neurogenic niches in the adult mouse brain.*** *The Dentate Gyrus (DG) of the hippocampus and the Ventricular-Sub Ventricular Zone (V-SVZ) along the lateral wall of the lateral ventricle (LV). The V-SVZ houses neural stem cells (NSCs) that are interspersed between ependymal cells (ECs). NSCs exists in different activation states; Quiescent (dormant q1 & primed q2) & Active (non-proliferating a1 & proliferating a2) in contact with the blood vessels (BV). Transitioning through transit amplifying progenitors (TAPs) they differentiate into early neuroblasts (ENBs) that migrate through the rostral migratory stream (RMS) to the olfactory bulb (OB), further differentiating into late neuroblast (LNBs) and finally mature into granule cells (GCs) and periglomerular cells (PGCs) that integrate into the olfaction circuitry. Right panel: Overview of neurogenic niches and the migration, maturation, and integration into the olfaction circuitry. Left panel: Schematic representation of the V-SVZ niche.*

The NSCs exist in a continuum of activation states namely, dormant (q1), primed (q2), active (a1) & proliferating (a2). Upon signals from the niche, activated NSCs either self-renew to

replenish the NSC pool or differentiate into neuroblasts (NBs) via an intermediate stage called the transit-amplifying progenitors (TAPs). These differentiated neuroblasts (NBs) then migrate through the rostral migratory stream (RMS) into the olfactory bulb (OB) where they mature into functional neurons (**Figure 1.1)**. These neurons finally integrate into the complex olfactory network that is essential for fine-tuning odour discrimination throughout the lifetime of the animal. Although, adult neurogenesis in the human brain has been debated (Kempermann et al., 2018), various studies show evidence of neurogenesis (Eriksson et al 1998, Spalding et al 2013, Ernst et al 2014, Boldrini et al., 2018). One study in particular, has shown new-born neurons in the striatum with the V-SVZ NSCs as the likely source (Ernst et al., 2014).

### 1.2. The Mammalian olfactory system

The murine olfactory system has two primary components; the main olfactory system responsible for the sense of smell and the vomeronasal system that is important for communication through pheromones. The olfactory placode (OP) together with the cranial neural crest (NC) cells are responsible for the dual origin of the olfactory organ (Katoh et al., 2011). During development, the invagination of the OP is responsible for the formation of the olfactory epithelium (OE) which further forms the turbinates of the main OE and the vomeronasal organ (VNO). The OP produces the olfactory sensory neurons (OSNs) and its supporting non-neuronal cells while the migratory NC cells provide for the structural aspects of the nose (Huilgol & Tole, 2016; Lyons-Warren et al., 2021). The apical surface of the nasal cavity is populated by the extensions of the OSNs as they mature to facilitate odour detection. Their OE-wrapped axons are projected via the olfactory nerve to the Olfactory Bulb (OB). The VNO produces pheromone receptor expressing cells whose projections reach the accessory olfactory bulb (AOB) (Huilgol & Tole, 2016; Lyons-Warren et al., 2021).

### 1.3. Olfactory Processing

The mammalian olfactory bulb is involved in processing the sense of smell and receives input from the olfactory sensory neurons (OSNs) lining the olfactory epithelium (OE) (Mombaerts, 2001). Specialized olfactory receptors (OR) expressed on the OSNs are capable of detecting odorants with each OSN expressing only a single OR from a repertoire of OR genes (>1000) (Barnes et al., 2020; Buck & Axel, 1991). Specialized structures in the OB called glomeruli receive input from the OE. All OSNs that converge into a single glomerulus in the OB express the same OR type (Mombaerts, 1999, 2001; Treloar et al., 2002; Zou et al., 2004).

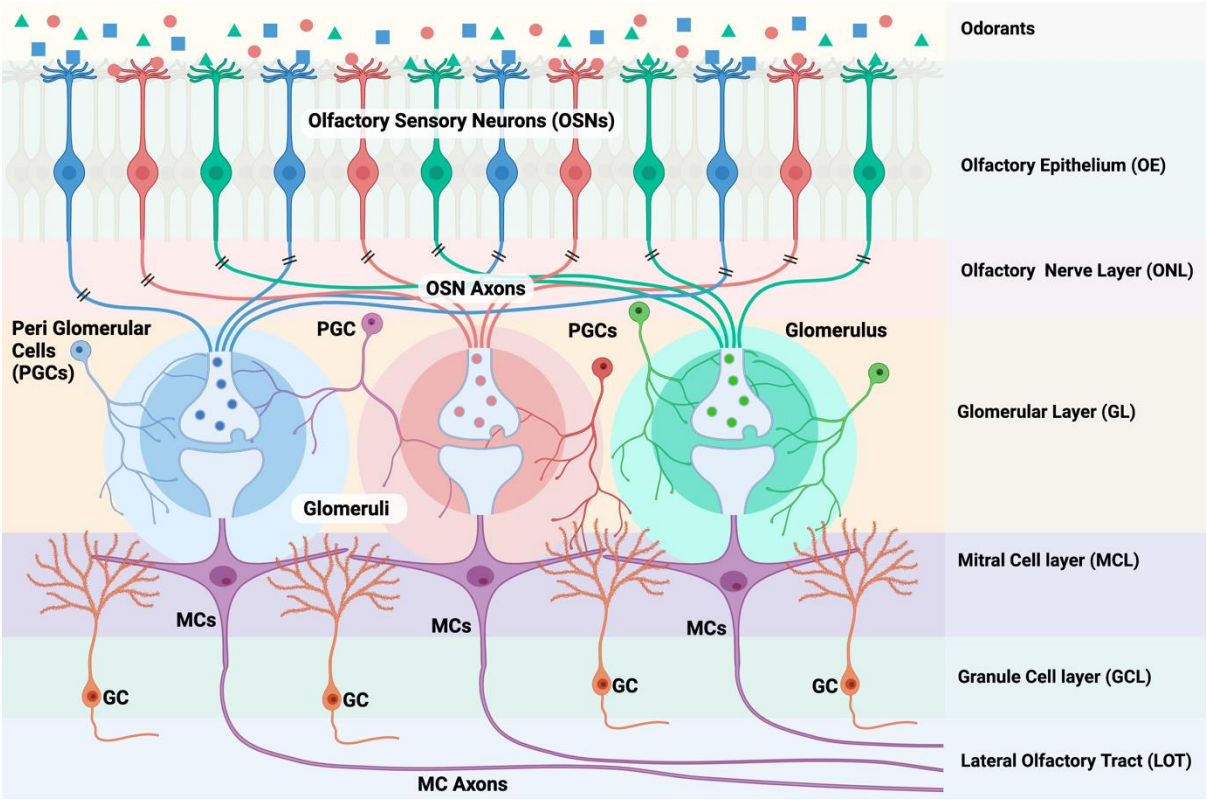


**Figure 1.2:** ***Schematic representation of the olfaction circuitry.*** *Different colours of the Olfactory sensory neuron (OSNs) represent the expression of one olfactory receptor (OR). Different colours of the glomeruli in the glomerular layer (GL) represent the corresponding input from one type of OSNs. OE, Olfactory Epithelium; PGCs, periglomerular cells; MCs, mitral cells; MCL, mitral cell layer; GCs, granule cells; GCL, granule cell layer; LOT, lateral olfactory tract; OE, olfactory epithelium. (Created with BioRender.com)*

Based on the patterns of glomerular activation a spatial map that reflects odour identity is represented in the OB (Bozza et al., 2004; Chong et al., 2020; Smear et al., 2013; Soucy et al., 2009). Glomerular activation is further fine-tuned via inhibitory and excitatory inputs from local interneurons (periglomerular cells; PGCs) in the glomerular layer (GL). These inputs are further processed by various olfactory bulb projection neurons like mitral & tufted cells (M/TCs) and their reciprocal lateral interaction with GABAergic granule cells (GCs) (Harvey & Heinbockel, 2018; Shao et al., 2019). This highly processed olfactory input is then sent to the olfactory cortex via the M/TC axons through the lateral olfactory tract (LOT) for further processing that finally translates to the sense of smell (Klingler, 2017).

### 1.4. Neural Circuitry

Developmental studies have identified key stages in circuit formation namely, axon elongation, recognition of target regions/neurons and formation of synapses (Seng et al., 2022). The elongating axon navigates its way through the brain to reach the intended target region

mediated by various molecular cues that allows for precise pathfinding (Bashaw & Klein, 2010; Seng et al., 2022). These signalling molecules comprise of four predominant families namely, slit, ephrin, netrin & semaphorins that are produced in the adjoining tissue while their receptors are expressed on the elongating axon (Seng et al., 2022). Based on the effect they have on the growth cone these cues can be characterized as either attractants or repellents and they work in tandem to guide the elongating axon (Poliak et al., 2015). Upon reaching their final projection region, recognition of appropriate neurons amid a cluster of different cells is essential to forge precise neuronal interactions within the existing circuitry. A combination of synaptic cell surface receptors and adhesion molecules such as Cadherins, Immunoglobulins & leucine-rich-repeats (LRRs), presented by both the elongating axon and surrounding neurons assist with precise target identification (Sanes & Zipursky, 2020). While the formation of neuronal circuits is predominantly a feature of the developing embryonic brain, it persists postnatally too. In rodents, circuit formation in the maturing brain is observed up to about 2 months after birth (Seng et al., 2022). In the adult brain this is studied primarily in the context of neurological disorders (such as ASD) & injury (McFadden & Minshew, 2013; Rodemer et al., 2020). Regeneration and sprouting are two mechanisms by which circuit formation takes place in the adult brain. While regeneration occurs upon injury of affected neuron, sprouting implies axonal growth from uninjured neurons, a characteristic of neurological disorders (Tuszynski & Steward, 2012). Sprouting is observed in multiple neuronal cell types and the resulting axons can either be routed locally or systemically (long range projecting neurons) across the brain. It is capable of producing both typical as well as atypical axonal projections that either concur or deviate from conventional developmental projection profiles respectively. They are also capable of producing functional synapses that may form quite rapidly (reviewed in (Seng et al., 2022)). Adult born neurons such as the ones produced by the subventricular zone (SVZ) & the dentate gyrus of the hippocampus eventually need to integrate into the existing neuronal circuitry. In the SVZ the differentiated neuroblasts that migrate through the RMS finally differentiate into two types of interneurons: periglomerular cells (PGCs) and granule cells (GCs) that eventually integrate into the olfaction circuitry (Doetsch et al., 1999; Lois & Alvarez-Buylla, 1994; Winner et al., 2002). Therefore, these differentiated neurons have the ability to re-wire healthy adult circuits through the formation of input/output synapses, axons & dendrites. In the OB, PGCs and GCs form functional synapses with the mitral and the tufted cells; the projection neurons of the OB (Gheusi et al., 2012). These adult-born neurons upon integration aid in the fine-tuning of various neuronal circuitry that affect behaviour. For example, in the OB these integrated interneurons contribute to odour discrimination,

associative learning & memory (Gheusi et al., 2012; Grelat et al., 2018; Li et al., 2018; Malvaut & Saghatelyan, 2016).

### 1.5. Functional consequences of SVZ Neurogenesis

The ability to smell is quite important in mammals already at birth. For example, the initiation of suckling in new-borns requires the recognition of maternal odour signatures (Bartocci et al., 2000; Kaitz et al., 1987; Kendrick et al., 1997; Vaglio, 2009; Winberg & Porter, 1998). A functional OB circuitry, albeit an elementary one, is therefore required even at the early postnatal stages (Logan et al., 2012). Further postnatal developmental changes in the OB circuitry continues to persist throughout adult life that allows for fine-tuning olfaction (Cecchi et al., 2001; Gheusi et al., 2000; Jurkowski et al., 2020). The ability of mammals in responding to novel environmental stimuli is augmented via continuous regeneration and replacement of granule cells that help reinforce odorant-environment association (Alvarez-Buylla & García-Verdugo, 2002; D. A. Lim & Alvarez-Buylla, 2016). Further, studies in mice have shown that alterations of SVZ neurogenesis via hormone treatment have an effect on sexual function (Lau et al., 2011). In particular, hormones such as estrogen may affect survivability of newly generated neurons from the SVZ and potentially regulating sexual behaviour (Brock et al., 2010; Burek et al., 1995; Heberden, 2017; Hidalgo et al., 1995; Ponti et al., 2018).

### 1.6. What are neurodevelopmental disorders?

Neurodevelopmental disorders (NDDs) are chronic heterogenous disorders that disrupt the functioning of the central nervous system (CNS) during development, primarily affecting cognitive abilities, motor skills, communication and/or behaviour (Ismail & Shapiro, 2019; Mullin et al., 2013; Sahin et al., 2018). These disorders can be either genetic or acquired during childhood before the individual attains the age of 22. Based on specific functional or behavioural impairments, NDDs are further classified into intellectual disability, autism spectrum disorder (ASD), communication disorders, attention deficit hyperactivity disorder (ADHD) and neurodevelopmental motor disorders (American Psychiatric Association, 2013). These different NDDs are known to co-occur with each other quite frequently. As an example, children diagnosed with ASD are often also diagnosed with intellectual disability (Mefford et al., 2012; Wilkins & Matson, 2009). The co-occurrence of NDDs with different paediatric mental health disorders is also common (Granlund et al., 2021). The current criteria for diagnosis (DSM-5) outlines groups of symptoms and behaviour patterns that categorize various NDDs (American Psychiatric Association, 2013). Some of the factors that contribute to the

cause of NDDs include genetic predisposition, metabolic defects, trauma, toxic exposure, nutrition among others.

**Table 1.1:** Major groups of neurodevelopmental disorders from the DSM-5 classification

| Disorders | Major symptoms |
|---|---|
| **Intellectual disabilities (ID)** | Cognitive impairments in functions related to communication (language, reading, writing, knowledge and interpretation), self-help (hobbies, personal care, financial management etc.) and social skills (communication, judgment, compassion, empathy, harmony) |
| **Communication disorders** | Difficulties associated with social communication either in language speech or phonetic fluency. |
| **Autism spectrum disorder (ASD)** | Difficulties in social communication in addition to restricted interests and repetitive behaviours |
| **Attention deficit hyperactivity disorder (ADHD)** | Inability to focus and be attentive during a task while having bursts of hyperactivity/ impulsivity |
| **Specific learning disorder (SLD)** | Learning difficulties in skills associated with spelling, writing or reading. |
| **Motor disorders** | Difficulties in performing tasks involving coordinated motor skills or behaviours associated with repetitive motion. |
| **Tic disorders** | Involuntary rapid, sudden, repeating and arhythmic motor movements or vocalizations (for ex. Tourette's syndrome) |

## 1.7. Autism spectrum disorders

Autism spectrum disorders (ASD) are a group of neurodevelopmental disorders primarily characterized by behavioural deficits in social interaction and communication as well as manifestations of repetitive and restrictive behaviour (Table 1.1; Alonso-Gonzalez et al., 2019).

**Table 1.2:** DSM-5 severity levels for autism spectrum disorders (ASD)

| Severity Level for ASD | Social Communication | Restricted Interests & Repetitive Behaviours (RRBs) |
|---|---|---|
| **Grade 1** 'Individuals needing some support | Deficits in social communication become pronounced impairments without some level of support. Patients show difficulty in initiating and overall decreased interest in social interactions. | Ritualistic and repetitive behaviour patterns interfere with their ability function in one or more contexts. Rigidness to change restricts any attempts to stop repetitive behaviour or divert from restricted interests. |
| **Grade 2** Individuals needing substantial support | Characterized by clear and observable deficits in social communication skills (both verbal and nonverbal); Despite the support provided, social impairments are quite apparent. Their ability to initiate social interactions is limited and their response to social overtures from others is often reduced or even abnormal. | The manifestation of repetitive behaviours and restricted interests are frequent enough that it is quite noticeable to the casual observer. These behaviours restrict and interfere with the individuals functioning in many different contexts. When interrupted during their RRBs the individuals are often distressed or frustrated and it is harder to distract them from fixated interests. |
| **Grade 3** Individuals needing very substantial support' | Deficits in verbal and nonverbal communication observed in individuals are severe which in turn results in severe impairments in their functioning. They display very limited initiation of and/or response to social interactions. | The repetitive behaviours preoccupation and restricted interests in such individuals severely interfere with their overall functioning. They display significant distress when their routines or rituals are interrupted and makes it extremely difficult to redirect their attention from their fixated interests. |

ASD is one of the most common neurodevelopmental disorders that is diagnosed in children with a higher proportion of affected males in comparison to females (American Psychiatric Association, 2013; Doernberg & Hollander, 2016; Scandurra et al., 2019). ASD is a multifactorial disorder with a variety of factors such as environment, genetics and physiology that contribute to the disease. Understanding its aetiology is therefore particularly challenging

with patients manifesting a broad spectrum of symptoms. ASD is reportedly a comorbidity of other psychiatric or neurological disorders such as ADHD, anxiety, depression, epilepsy etc. Studies have shown that altered brain development during early stages and neural reorganization potentially results in ASD (Bauman & Kemper, 2005; O'Reilly et al., 2017). The functional consequences of such alterations are often seen in the form of impeded learning abilities, difficulties in working independently, difficulty in adapting to change and novelty (American Psychiatric Association, 2013).

### 1.8. The ASD phenotype

Impairments in social interaction and communication include both verbal and non-verbal communication. Verbal/language deficits range from pedantic and or echoed speech, poor speech comprehension, to language delays that result in a complete lack of speech. The manifestation of these deficits depends on age, language ability and intellectual level. Impairment in non-verbal communication include atypical, reduced, or non-existent usage of eye-contact (compared to culturally accepted), functional gesturing, intonations of speech, facial expression etc. The ability to engage with people and communicate feelings and thoughts (social-emotional reciprocity) is documented to be impeded in children with ASD. In adults, the ability to understand social cues, process and respond to them is often observed to be diminished. Repetitive or stereotyped behaviours include repetitive object use (lining up toys or stacking), repetitive speech or simple motor stereotypies like hand flapping etc. Restrictive behaviour commonly results from rigidity to change and therefore a strong adherence to prior and familiar routines, interests, or activities. These symptoms persist since early childhood and impede everyday functioning (American Psychiatric Association, 2013). In addition to the behaviour phenotype, other clinical features have been associated with ASD. One of the characteristic features of ASD or ASD-like behaviour is changes in head circumference (Microcephaly or Macrocephaly) (Betancur, 2011; Cohen, 2003; Lenroot & Yeung, 2013; Raznahan et al., 2013). Studies have shown the association of an increased head circumference (Macrocephaly) with various genetic syndromes that also present ASD-like behaviour (Cohen, 2003; Raznahan et al., 2013). On the other hand, genetic studies have shown that various syndromes associated with microcephaly manifest ASD-like behaviour (Betancur, 2011). In addition to head circumference other morphological features have been associated with ASD (H. Ozgen et al., 2010, 2013; Heval Ozgen et al., 2011). These features are broadly classified into major and minor abnormalities (Merks et al., 2008; H. Ozgen et al., 2013). Major abnormalities can be malformations due to defects in development or other deformation while

minor abnormalities are as small morphological variations that do not present any serious medical problems or even cosmetic issues to the affected individual but can be used as a diagnostic feature (Aase, 1990; Miles et al., 2005; Miles & Hillman, 2000; H. Ozgen et al., 2010, 2013).

### 1.9. Dysregulation of Neurogenesis in ASD

Deficits in neurogenesis related to ASD have been previously reported (Gilbert & Man, 2017; Parikshak et al., 2013). Functional genomic analysis studies investigating ASD-risk associated genes have reported disruption of processes such as neurogenesis along with synapse formation in early stages of developmental in ASD (Parikshak et al., 2013). Genes such as *WDFY3*, LIS1, NDE1, TBR1, FMR1, CHD8 etc, that are involved in controlling neurogenesis and development have been shown to be associated with ASD (Gilbert & Man, 2017; Parikshak et al., 2013; Willsey et al., 2013). These findings do suggest that neurogenesis seems to be dysregulated during development in ASD and could play an important role in ASD pathogenesis. Such a dysregulation of neurogenesis could lead to the production of defective neurons that could interfere in the normal signalling of the brain. Studies have reported imbalances in the excitatory/inhibitory (E/I) ratio to be a common feature of ASD that then affects various neuronal processes involved in learning and memory, cognition, motor regulation and sensory processes such as olfaction (Uzunova et al., 2016).

### 1.10. Olfactory deficits in ASD

Many neurological disorders such as Parkinson's, Alzheimer's disease Epilepsy & Autism Spectrum Disorder have been associated with various olfactory deficits (Doty, 2012; Jung et al., 2019; Khurshid et al., 2019; Okumura et al., 2020; Tonacci et al., 2017). In many neurodevelopmental disorders in particular a critical characteristic has been shown to be olfactory dysfunction (Hornix et al., 2019; Lyons-Warren et al., 2021). For example, individuals with ASD have been reported to show atypical sensory reactivities especially deficits in olfactory processing (Okumura et al., 2020; Tonacci et al., 2017). DSM-5 has included defects in sensory sensitivity such as hyposmia and hyperosmia as a diagnostic criterion for ASD (American Psychiatric Association, 2013; Crow et al., 2020).

## 1.11. The pathogenesis of ASD

The prevalence of ASD internationally as reported by WHO is around 0.76% (Baxter et al., 2015). However, it is estimated that the actual prevalence of ASD is much higher, as this estimate takes into account only 16% of the global population of children (Baxter et al., 2015). As of 2016, the CDC has reported a 2.5% prevalence of ASD in the United States (US) (Kogan et al., 2018). According to some estimates the prevalence of ASD has doubled from 2002 to 2012 (Baio et al., 2018). Additionally, males seem to be affected more than females in a 3:1 ratio, however the symptoms may be more camouflaged in females (Corbett et al., 2021; Loomes et al., 2017). This suggest that there is a growing burden of ASD and understanding the disease pathogenesis is key to addressing this issue. The pathogenesis of ASD has not been fully understood, despite increased research interest in the field. Broadly, ASD can be classified into either secondary or idiopathic (Casanova et al., 2020). Cases where the specific cause leading to ASD has been identified are termed as secondary ASD. These include conditions such as Fragile X syndrome (FXS), down syndrome, tuberous sclerosis, and a few congenital infections (Casanova et al., 2020; Moss et al., 2013; Wheeler et al., 2015). Genetic abnormalities could be caused by either single gene mutations (5%) like FMR1, SHANK3 or MECP2, copy number variations (CNVs; 10%) produced by translocations, duplications or deletions, or by accumulation of poly-genic risk factors contributing to different aspects of the overall ASD phenotype (Varghese et al., 2017). Of the known genetic causes attributed to ASD, *de novo* mutations causing ASD makeup 20-30% (De Rubeis & Buxbaum, 2015; Pinto et al., 2014). Various collaborative studies in the recent past have uncovered numerous novel risk genes for ASD through whole-exome sequencing (Iossifov et al., 2014; O' Roak et al., 2012; Sanders et al., 2012; Satterstrom et al., 2020). On the other hand, cases where the causes are unknown are grouped as idiopathic ASD. These are the most predominant type of ASD with about 85%, while the lees prevalent secondary ASD constitute about 15% of the cases (Casanova et al., 2020). Environmental factors such as prenatal toxin exposure and maternal immune activation have been suggested as risk factors for idiopathic ASD. Being able to differentiate between secondary and idiopathic ASD provides a better understanding of the appropriate treatment protocols (both pharmacological and behavioural) to be administered (Devlin & Scherer, 2012).

## 1.12. Diagnosis of ASD

As there are no reliable biomarkers for diagnosing ASD currently, it relies on behavioural tests conducted by clinicians. They assess symptoms arising typically between 12 and 24 months of age and persisting throughout the affected individual's lifetime. Broadly the ASD spectrum identifies two major behavioural domains; a) defects in social communication and b) restricted, repetitive or unusual sensory motor behaviours. Despite ASD being highly heterogeneous, these behaviours are a pretty common feature across culture, ethnicity and/or socioeconomic groups (Khan et al., 2012). Different standardized diagnostic tools such as Screening Tool for Autism in Toddlers and Young Children (STAT), Autism Diagnostic Observation Schedule (ADOS) allow clinicians to be able monitor and categorize specific behaviour patterns in children suspected of ASD (Lord et al., 2018). The clinicians are then able to rate the observed symptoms on a variety of scales such as Social Responsiveness Scale (SRS), Childhood Autism Rating Scale (CARS), Social Communication Questionnaire (SCQ) and even some adaptive scales to assess functional behaviour for everyday living (Kenworthy et al., 2010; Lord et al., 2018; Weitlauf et al., 2014). Further, patient patients related to repetitive behaviour and level to expression through language, cognitive and motor abilities is always taken into account for diagnosing ASD. Syndromic ASD on the other hand can be tested for genetic markers. While it may not be able to accurately diagnose for ASD it can screen for known genetic causes implicated in ASD. Such form of testing is more advantageous if there is a family history of syndromic ASD. As is the feature of NDDs, ASD is usually accompanied by other disorders like ID, FXS, epilepsy, ADHD etc and are often tested for such diseases.

## 1.13. Relevance of alternative polyadenylation in Neurogenesis

Advances in sequencing technologies provided researchers more power to resolve cellular heterogeneity at single cell resolution. As previously stated NSCs exist in a continuum and transitioning between these states is regulated by both extrinsic (niche factors) as well as intrinsic factors (such as transcriptional/translational programs) (*Kalamakis et al., 2019; reviewed by Bond et al., 2015*). Distinct molecular signatures have been associated with each state. For example, a metabolic switch from a glycolysis to oxidative phosphorylation is observed when cells transition from quiescent to a more differentiated state (qNSCs to NBs) while, activation of NSCs is characterized by an upregulation in ribosome biogenesis and cell cycle transcriptional programs (Baser et al., 2019; Kalamakis et al., 2019; Llorens-Bobadilla et al., 2015). At the level of translation, a dynamic regulation of global protein synthesis is

reported along the NSC differentiation lineage (Baser et al., 2019). The onset of differentiation is accompanied by an uncoupling of transcription and translation that involves repression of a subset of mRNAs that include stem cell related genes. Transcripts of these genes contain a pyrimidine rich motif in the 5'UTR region that is regulated by mTORC1 activity (Baser et al., 2019). These data highlight the importance of post-transcriptional regulation in cell state transitions. The 3'UTR has regulatory elements such as binding sites for miRNAs and RNA-binding proteins (**Figure 1.3**). Alternative polyadenylation (APA) allows for recognition of different poly A signals (PAS) within the 3'UTR of a transcript and thereby facilitate the inclusion or exclusion of regulatory elements that affect mRNA translation, stability, and localization (Di Giammartino et al., 2011; Matoulkova et al., 2012; Miura et al., 2014). About two-thirds of all human genes are affected by APA (Derti et al., 2012). Increasing morphological and organismal complexity has been shown to correlate with 3′UTR lengthening (C. Y. Chen et al., 2012). Neuronal cells require a complex regulatory network. The longer the 3'UTRs the allow for a more controlled translational regulation (Pereira-Castro & Moreira, 2021) . Studies have shown a global lengthening of 3'UTRs during differentiation while dedifferentiation and cancer initiation have been associated with 3'UTR shortening (Ji&Tian, 2009; Ji et al, 2009; Mayr & Bartel, 2009).

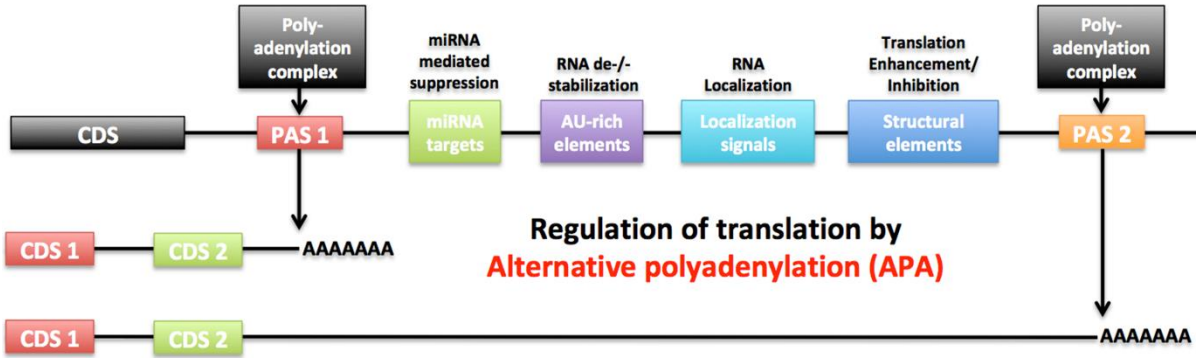


**Figure 1.3:** ***Overview of the different elements in 3'UTRs*** *that can facilitate different regulatory processes that affect translation, stability, and localization of the mRNA molecule.*

## 1.14. Relevance of APA in neurons

Neurons differ from other cell types in that their cellular structures are highly compartmentalized. This confers functional specificity within each individual compartment via structures such as the intricate branching observed in axons & dendrites and synaptic transmission. The latter in particular requires an efficient mechanism that can locally process information. mRNA localization at synapses and subsequent local translation are feature of synaptic compartmentalization that give neurons the ability to swiftly respond to incoming

information. This forms the basis of various critical process in the brain such as learning, memory, axon guidance formation of dendrites and synapses etc. In neurons, thousands of mRNA transcripts are locally compartmentalized, and this can be facilitated via sequences in their 3'UTRs. Processes such as APA can regulate the translation efficiency and stability of these transcripts. Therefore, understanding of APA within these localized compartments can provide insight into local vs global translation in neurons. Studies have shown that APA is regulated differently during the different developmental stages and in different 3'UTR isoform are expressed in a tissue specific manner within the same organism (Tushev et al., 2018). Specifically, the 3'UTR isoforms expressed in the brain are larger than those expressed in other tissues (Miura et al., 2013).

### 1.15. Alternative polyadenylation machinery

The APA machinery is a complex of many different proteins that are essential in the processing of mRNA. There are 4 main components to the APA machinery namely, cleavage and polyadenylation factor (CSPF), cleavage stimulating factor (CSTF), cleavage factor I (CFI) and CFII (A. J. Gruber & Zavolan, 2019). In addition to this these proteins others such as poly(A) polymerase (PAP), symplekin etc also play a role in APA (Figure 1.4).

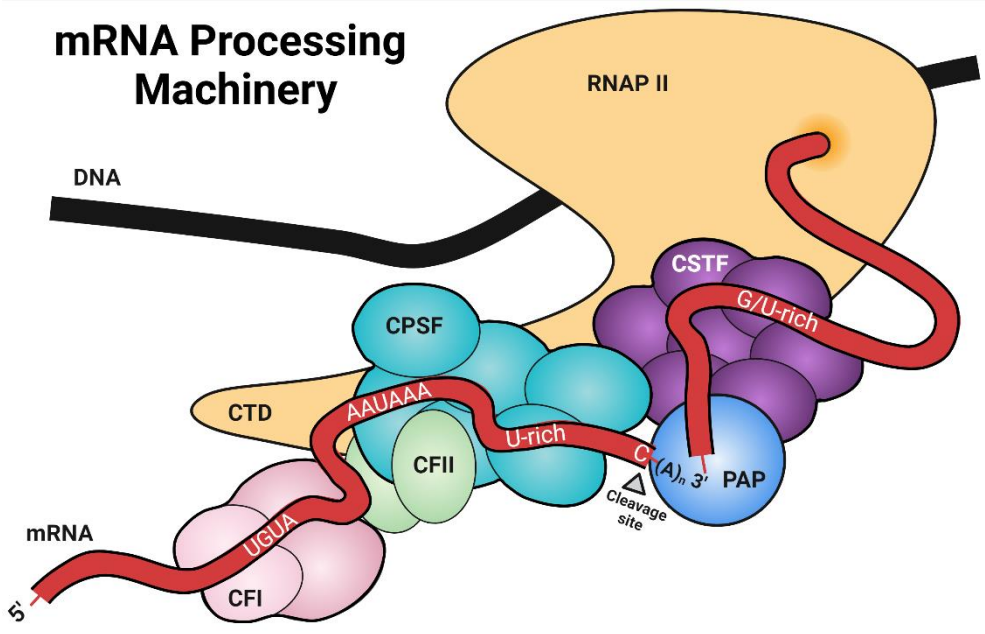


**Figure 1.4:** ***Components of the mRNA processing machinery****. This complex contains 4 major components: CPSF, CSTF, CFI & CFII. Additionally, it contains the RNA polymerase II (RNAP II) and the poly(A) polymerase (PAP). These proteins recognize various regulatory sequences on the mRNA such as the poly(A) signal (PAS; AAUAAA), the UGUA motif bound by CFIm25 (NUDT21), U-rich motif and the G/U-rich motifs. (Created with BioRender.com)*

The recognition site for the cleavage and polyadenylation also call poly(A) signal (PAS) is the canonical hexamer AAUAAA located around 20 bp upstream from the cleavage site, which is conserved across metazoans (Proudfoot & Brownlee, 1976). The recognition of the PAS occurs co-transcriptionally via the recruitment of the CPSF sub-complex (CPSF 1-4) (Dantonel et al., 1997; Shi et al., 2009). Apart from the canonical PAS other variants resembling it in sequence and function have been identified (Beaudoing et al., 2000; A. J. Gruber et al., 2016). Upon recognition of the correct PAS, CPSF3 performs an endo-nucleolytic cleavage about 20 bps downstream of PAS (Mandel et al., 2006; Wang et al., 2018). Once the cleavage is done the poly(A) polymerase now proceeds to add the poly(A) tail. The binding of the poly(A) tail by the poly(A) binding protein 1 (PABPN1) regulates the poly(A) tail length by disrupting the interaction between PAP and CPSF (Eckmann et al., 2011). In addition to the PAS there are other regulatory elements that are present in its vicinity to facilitate PAS selection. A U-rich region present upstream of the cleavage site and downstream of PAS facilitates the binding of CPSF sub complex subunit FIP1L1 (Kaufmann et al., 2004). Additionally, a G/U-rich region that is also present downstream of PAS and the cleavage site facilitates the binding of the CSTF subcomplex (CSTF1-3) influences location of the cleavage site (MacDonald et al., 1994). Finally, a UGA motif that is located about 40 bp upstream of the site of cleavage facilitated the binding of NUDT21 (also called CPSF5/CFIm 25) (Yang et al., 2010). NUDT21 is part of the tetramer protein complex CFIm that contains 2 smaller NUDT21 subunits and 2 larger subunits of CPSF6 (also called CFIm 68) and may or may not contain CPSF7 (CFIm 59) (Yang et al., 2011). The CFII complex is the least characterized among the complexes that make up the APA machinery. CFII consists of two polyadenylation factors CLP1, that is essential for the cleavage activity of CFII and PCF11, that has RNA binding affinity to G-rich regions (Schäfer et al., 2018). All these components of the polyadenylation machinery work in tandem to process the mRNA and regulate APA.

## 2. Materials & Methods

### 2.1.Experimental animals

All animal experiments were carried out in compliance with the regulations and guidelines prescribed by the Regierungspräsidium Karlsruhe, Germany. The animals were housed in a satellite facility at the German Cancer Research Centre (DKFZ) that maintained a 12-h/12-h light/dark cycle along with controlled temperature and humidity. The animals had constant access to food and water and the cages housing them were prepared with standard bedding along with enrichment. They were housed either alone or in groups (max =3 adults). The wild type mice used in this study were C57BL/6N male mice unless otherwise specified. Additionally, for knockout experiments APLP1 -/- mice (Heber et al., 2000) and the CPEB4 exon4 KO (TgCPEB4Δ4) mice (Parras et al., 2018b) were used. For behavioural experiments the WT littermates of APLP1-/- mice were used as controls.

### 2.2.SVZ dissection & tissue dissociation

The protocol to micro-dissect the lateral wall of the v-SVZ has been performed as previously described by Mirzadeh et al, with slight modifications (Mirzadeh et al., 2010). In brief, the mouse brain was extracted after opening the skull and transferred to a petri dish containing dissection solution. Excess tissue such as the olfactory bulb and the cerebellum (till lambda) was removed after which the two hemispheres were separated. A caudal (posterior) incision was made to separate the hippocampus from the cortex and the lateral wall of the ventricle was revealed by pulling the hippocampus rostrally. The medial wall along with the hippocampus was incised out at the point of intersection with the lateral wall. Now using an angled microdissection scissors and about 200 µm beneath the ventricular surface the SVZ is carefully cut out. The micro-dissected SVZs were then processed to prepare a single cell suspension using either the MACS Neural Tissue Dissociation Kit (T; Trypsin) or the (P; Papain) kit (Table 2.1).

**Table 2.1.** Composition of enzyme mix 1 & 2 for NTD kits with papain (P) and trypsin (T)

| | Enzyme Mix 1 | | Enzyme Mix 2 | |
|---|---|---|---|---|
| **MACS Kit** | **Buffer-X** | **Enzyme** | **Buffer-Y** | **Enzyme A** |
| NTDK (T) | 1750 µl | 200 µl (Trypsin) | 20 µl | 10 µl |
| NTDK (P) | 1900 µl | 50 µl (Papain) | 20 µl | 10 µl |

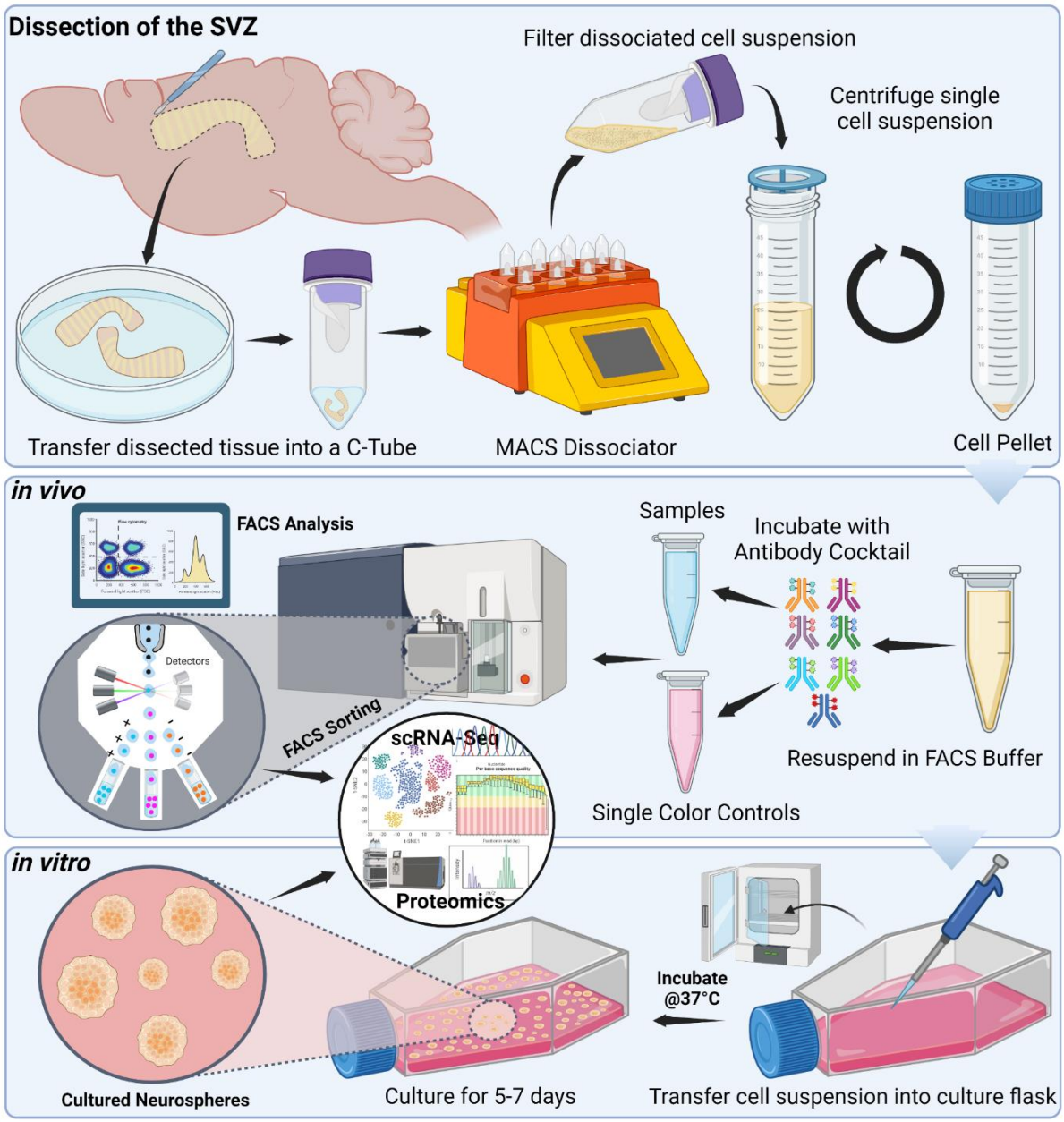


**Figure 2.1:** ***Dissection of SVZ: Schematic overview of dissection of the sub ventricular zone and downstream processing,*** *(Top) Microdissection of the SVZ and preparation of single cell suspension using the gentleMACS system. (Middle) Preparation of cells for in vivo applications; Flow cytometry Analysis and Sorting. (Bottom) Preparation of cells for in vitro application; NSC neurosphere cultures. (Created with BioRender.com)*

For experiments where replicate identity was to be preserved, the SVZ dissected from a single mouse were processed individually in separate C-Tubes. The tissues were transferred to a gentleMACS C-Tube that contained pre-warmed (37°C) enzyme mix 1 (Table). To this 30 µl of enzyme mix 2 (Table 2.1) was added. The C-Tubes were sealed by tightening the cap and then placed inverted onto the MACS Dissociator sleeves along with the temperature regulator

jacket. The program 37C_NTDK_1 was run for 22 minutes while being incubated at 37°C. The baffles attached to the cap facilitated mechanical dissociation while the enzyme mixes allowed for enzymatic dissociation of the tissue. The resulting suspension was quickly spun down and the enzymatic reaction was stopped by adding 15 ml FACS buffer (10% FCS in PBS) to it. This suspension was then passed through a 70 µm filter resulting in a single cell suspension which was spun down to pellet the cells. This pellet was subsequently processed according to the necessary downstream application (*in vivo / in vitro* applications)

### 2.3.Handling of *in vivo* animal experiments

#### 2.3.1. FACS experiments

For FACS experiments the Trypsin Kit was used as the GLAST epitope, the lineage marker, is sensitive to papain. For bulk sorting experiments the SVZs from up to 4 mice were processed together. For all FACS experiments a separate mouse was used for controls such as unstained, single colour (SC) and fluorescent minus one (FMO). The cell pellets (as described in section 3.2) from each C-tube were resuspended in 100 µl/mouse of Blocking Buffer (5% FcR Blocking reagent in FACS Buffer) and transferred to 1,5 ml Eppendorf tubes. For each of the controls, 20µl of cell suspension was used from the control mouse. An antibody staining was performed for 20 minutes with the relevant antibodies used in the following dilutions: CD45-APC-Cy7 (1:200), Ter119-APC-Cy7 (1:100), O4-APC-Vio770 (1:100), GLAST-PE (1:50), EFGR-A488 (1:100), CD133-APC (1:75), PSA-NCAM, PE-Cy7 (1:50) and Sytox-Blue (1:500). After staining the samples for 20 mins, three washes with FACS Buffer were performed and the resulting pellets were and resuspended in FACS Buffer (100 µl/mouse or control).

For scRNAseq experiments samples were sorted for the lineage marker GLAST, after depleting for Oligos (O4), Erythrocytes (Ter119), Microglia (CD45) & Dead Cells (Sytox-Blue) (Figure 3.2A). For FACS analysis of the SVZ niche, the depletion step was done as before (O4-Ter119-CD54-SytoxBlue-), and the different cell types of the niche were characterized as follows; GLAST+Prom1+ (NSCs), GLAST- Prom1- PSANCAM+ (NBs), GLAST- Prom1- EGFR+ (TAPs) (Figure 3.2B). FACS sorting was performed on FACS Aria with a 100 µm nozzle at 20 psi, while for FACS analysis FACS Canto was used. For all analysis FACS Diva software was used for analysis of the data.

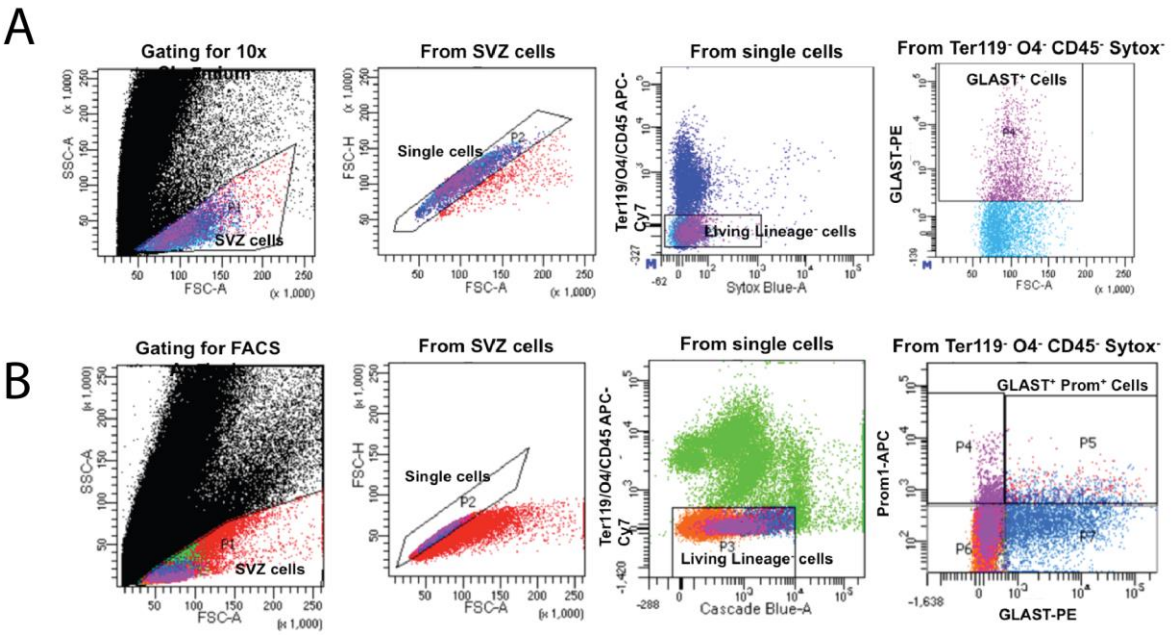


**Figure 2.2:** ***FACS gating strategies for sorting and analysis of the SVZ niche.*** *A. Sorting strategy for the enrichment of the NSC lineage (GLAST$^{+}$) for scRNASeq.* ***B.*** *FACS analysis of the SVZ niche cells to determine NSC numbers.*

### 2.3.2. Preparing cells for 10x chromium based scRNA-Seq:

For in vivo preparations (Aplp1-/- vs. WT), single cell suspensions were prepared as described above from SVZ (Section 2.2).

#### *2.3.2.1.Hash-tagging of biological replicates for sequencing experiments*

To preserve the biological replicate information, SVZs from each mouse was processed separately in an individual C-Tube and the dissociation steps were carried out as described before. The cells were stained as before for the scRNASeq experiments with the addition of unique hashtag antibody (BioLegend) to each individual mouse (1µg/100µl/mouse). The hashtag antibodies were specific against both CD45 and MHC class 1 ensuring that a broad range of cell types were targeted. Even in cell types with low expression levels of these two epitopes, whose range of expression targets almost all cell types. After the antibody staining and washing steps the replicates were combined into a single sample for downstream applications. For FACS sorting experiments the gating strategy in Figure 3.2 was employed to enrich for the neural stem cell lineage (CD45- O4- Ter119- GLAST+) cells.

#### *2.3.2.2.Sequencing on the 10X platform*

These sorted cells/single cell suspensions were then processed according to manufacturer's instruction for the 3' scRNASeq from 10x Genomics (v3.1). In Brief, the sorted cells were

loaded on to the 10x chromium chip to generate the GEMs and were further processed to generate 10x libraries. These libraries were sequenced on the NovaSeq 6000 PE 150 (for the 3'UTR analysis).

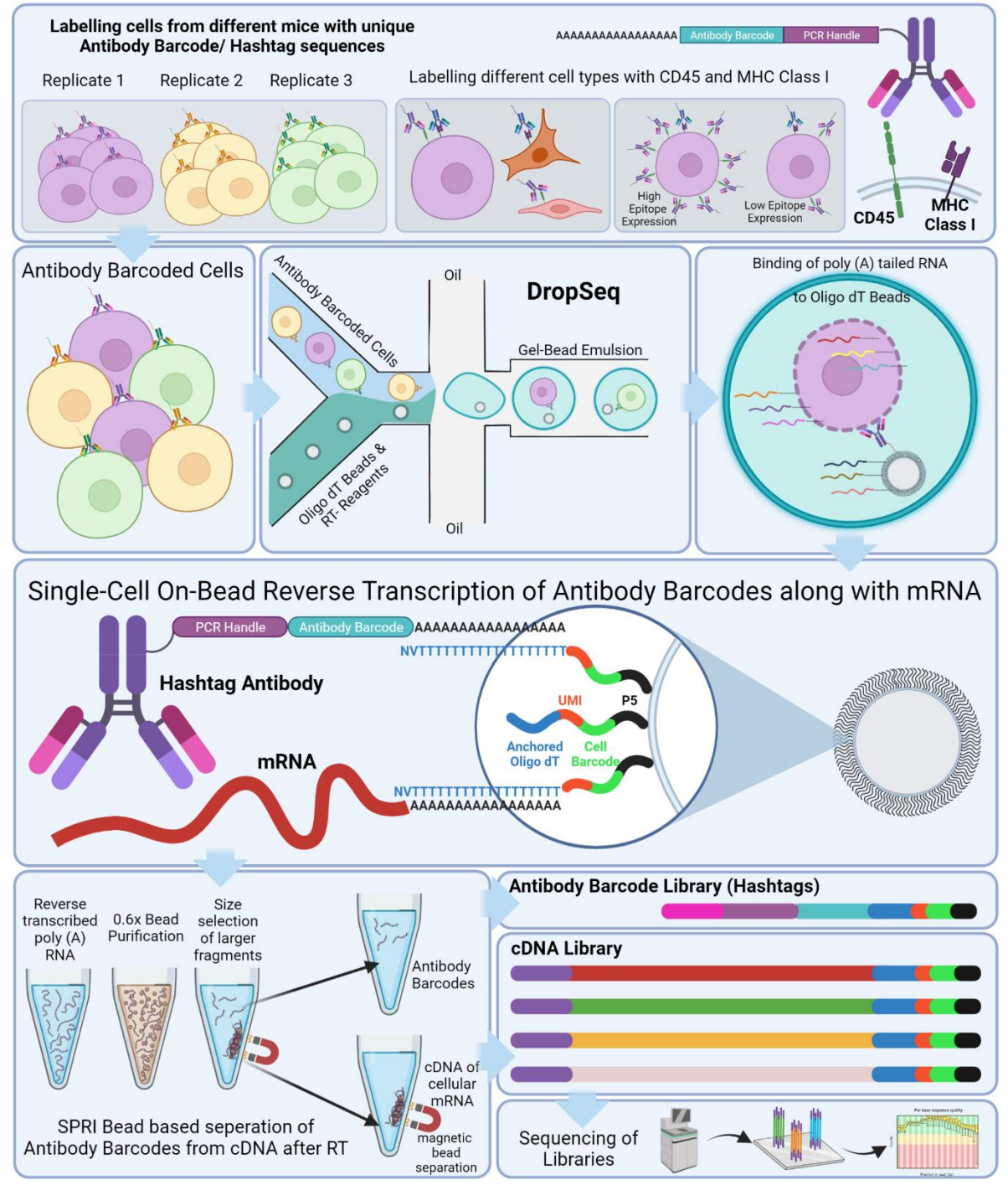


**Figure 2.3.** ***Schematic workflow for hash-tagging individual biological samples to preserve replicate identity.*** *In brief, samples are stained with antibodies targeted against CD45 and MHC Class I. These antibodies contain unique poly (A) tailed barcodes (hashtags). In combination with a DropSeq method hash-tagged cells are processed for scRNASeq where the hashtags are reverse transcribed along with mRNA. Using SPRI bead-based size selection, the shorter hashtags are separated from the cDNA. Both are them individually processed to produce libraries which are then sequenced. (Created with BioRender.com)*

## 2.4.Handling of *in vitro* animal experiments

### 2.4.1. Culturing of NSC neurospheres

For *in vitro* experiments the SVZ was dissociated using a much hasher papain kit which resulted in a better tissue dissociation. Although papain degraded the GLAST epitope, this was replenished by the NSCs themselves upon culturing. The single cell suspension from the processed SVZ was prepared as described in Section 3.2. and resulting pellet was resuspended in 7 ml of Neurobasal Medium (NBM). The medium ensured the survival and proliferation of primarily active NSCs while other cell types slowly died out. The entire suspension was seeded into a T25 flask for culturing and was labelled "Passage 0" (P0). During this time NSCs began to aggregate and proliferate as free-floating spherical clusters called neurospheres (Figure 3.1). After about 5-7 days of culturing the now fully formed neurospheres were collected and pelleted down. The neurospheres were then dissociated using 200 µl of Accutase to produce a single cell suspension. This was transferred to a T75 flask for expansion and labelled "Passage 1" (P1) and cultured till confluent. If the cells then needed to be expanded, they were transferred to a T150 flask, else they were passaged (1:10) twice a week. NSCs cultures beyond "Passage 10" (P10) were avoided as they eventually began to differentiate with increasing passages. To preserve mouse identity, the SVZs from individual mice were processed separately and further put into separate flask during culture and subsequent passaging.

**Table 2.2 :** *Composition of Neurobasal medium for NSC culture*

| Component | Volume |
|---|---|
| Neurobasal A Medium (Gibco) | 500ml |
| B27 Supplement (50x) | 10ml |
| L-Glutamine (200mM) | 5ml |
| Pen-Strep (Pen: 10,000units/ml; Strep: 10,000 µg/ml) | 5ml |
| Heparin, 2 µg/ml | 500 µl |
| **When growth factors were used**: | |
| bFGF (0.5 µg/µl in PBS) | 20 µl |
| EGF (0.5 µg/µl in PBS) | 20 µl |

### 2.4.2. BMP4 induced quiescence in neural stem cells

Neural stem cells in culture are primarily active proliferating NSCs. In order to induce quiescence *in vitro* NSCs were treated with Bone Morphogenic Protein-4 (BMP4) as described in previous studies (Martynoga et al., 2013; Mira et al., 2010). To this end, NSCs were cultured as adherent monolayers on pre-coated culture dishes or LabTek chamber slides. The coating was done with Poly-D-Lysine (PDL; 1 µg/ml) overnight at 37°C with a subsequent Laminin

coating (10µg/ml) for 2 hours at 37°C. NSCs were seeded at a density of 10,000 cells/cm$^2$ in Neurobasal Medium and allowed to attach to the coated surface overnight. The next day the culture medium was switched to a modified version of the Neurobasal Medium (NBM) in which EGF was replaced by BMP4 (50 ng/ml, R&D Systems). Before adding the new medium, the adherent cells were carefully washed twice with PBS to get rid of any residual EGF from the wells that might hinder the induction of quiescence. The BMP4 treatment was performed for 3 days with fresh medium exchanged every 2 days. Quiescence was confirmed based on morphology, proliferation, and marker expression. Over the course of the BMP4 treatment the cells ceased to proliferate and demonstrated a "Fried Egg" morphology as opposed to a spindle shape that active NSCs in culture have.

### 2.4.3. Immunofluorescence

To determine the induction of quiescence in vitro, both active and BMP4 treated NSCs were probed for the expression of Ki67 as follows. After the BMP4 treatment (described in Section 3.4.2.) these adherent cell cultures were carefully washed twice with PBS to get rid of any residual media. The cells were fixed with 4% PFA for 20 minutes under a fume hood. After fixation the PFA was discarded, and the cells were washed 3 times with PBS. The cells were stored at 4°C in PBS if antibody staining was not immediately performed. Prior to antibody staining, a 45-minute blocking step was performed to get rid of non-specific/background antibody (blocking buffer; 3% Horse Serum + 0.3% TritonX in PBS). Cells were incubated overnight at 4°C with the primary antibody (anti-MKi67; rabbit, Novus, 1:200) diluted in blocking buffer. The cells were then washed with PBS 4 times for 5 minutes each. The cells were the stained with a fluorescently labelled anti-rabbit secondary antibody (Alexa 555; 1:1000) for 1 hour at room temperature in the dark. Appropriate secondary antibody controls were used to determine background binding. Cells were washed again with PBS 4 times for 5 minutes each. For the final step the any residual PBS was removed from the wells and the plastic chambers that separated the wells were detached according to manufacturer's instruction. The coverslip was mounted onto the slide by adding Fluoromount with DAPI over the wells to ensure nuclei staining. The slides were allowed to dry in a dark place for at least 30 mins and stored at 4°C until microscopy.

### 2.4.4. Microscopy & image analysis

The microscopy images were captured using a Leica TCS SP5 microscope. For the validation of the BMP4 experiment, a UV diode (405 nm) was used to probe for the DAPI signal (nucleus)

and a helium-neon 561nm laser for the AlexaFluor$^{555}$ (MKi67) signal. A 40x oil immersion objective was used for imaging. Appropriate antibody controls were used to determine the threshold of background fluorescence signal for each channel to setup laser intensities for sample imaging. For each field of view a 6-12 z-stack (1 µm/stack) was captured at a resolution of 1024x1024 and an imaging frequency of 250 Hz. For each treatment type or biological replicate 2-3 images were captured. The images were the analysed using the ImageJ software. First a maximum projection image was created that combined all the z-stacks into a single image. Using the cell counter plugin, the percentage of proliferating cells ($DAPI^{+}Ki67^{+}$) were determined for each well. The standard deviation was calculated for the replicates and a standard t-test was performed to determine differences between the treatment types.

### 2.4.5. scRNA-Seq of *in vitro* NSCs

NSCs were cultured as described in Section 3.4.2. for the BMP4 experiment mimicking active and quiescent NSCs *in vitro*. Adherent NSCs were detached using accutase that resulted in a single cell suspension upon mild pipetting. Cell densities were estimated, and three biological replicate/treatment type were equally pooled together to get a final density of 200 cells/µl. 20,000 cells per treatment type were loaded on to the 10x Chromium chip of the single cell 3' v2 kit and the RT, cDNA and libraries were prepared according to manufacturer's instructions. The samples (BMP4 & EGF Treated) were indexed and run-in separate lanes on the HiSeq4000 with 100 bp Paired End Reads.

## 2.5.Behavioural experiments in mice

A battery of behavioural tests was conducted to investigate the APLP1-/- mice for autistic traits. Gold standard tests to identify defects in speech and social behaviour such as Ultrasonic vocalization, Sociability test and the three chamber tests were used. Apart from these, other tests such as Elevated plus maze and open field test were used to rule out anxiety in these animals.

### 2.5.1. Ultrasonic Vocalization (USV)

Previously described protocol for recording USVs was adapted. (33). For this experiment infant mice were evaluated for the USVs at two different developmental points; namely postnatal day (PD) 5 and 8. Firstly, the pups were isolated from their mother to illicit a vocalization response that was recorded. An empty glass container (6 x 9.5 x 7.5 cm) was filled with bedding material before placing the pups during the course of the USV recordings. This glass container was then

placed inside a white acryl open field box (40 × 40 × 40 cm). A Plexiglas lid was used to cover the roof a to dampen ambient noise. A microphone to record the USVs was placed through the roof at about 30 cm from above. Over a period of 5 mins the USVs were recorded after which the pups were returned back to their cage. All experiments were conducted at the same time of the day and in the same order of pups for both P5 and P8 to eliminate any bias that could arise otherwise. A total of 15 WTr6 (11 male and 4 females) and 17 APLP1-/- (11 males and 6 females) pups were used in this study.

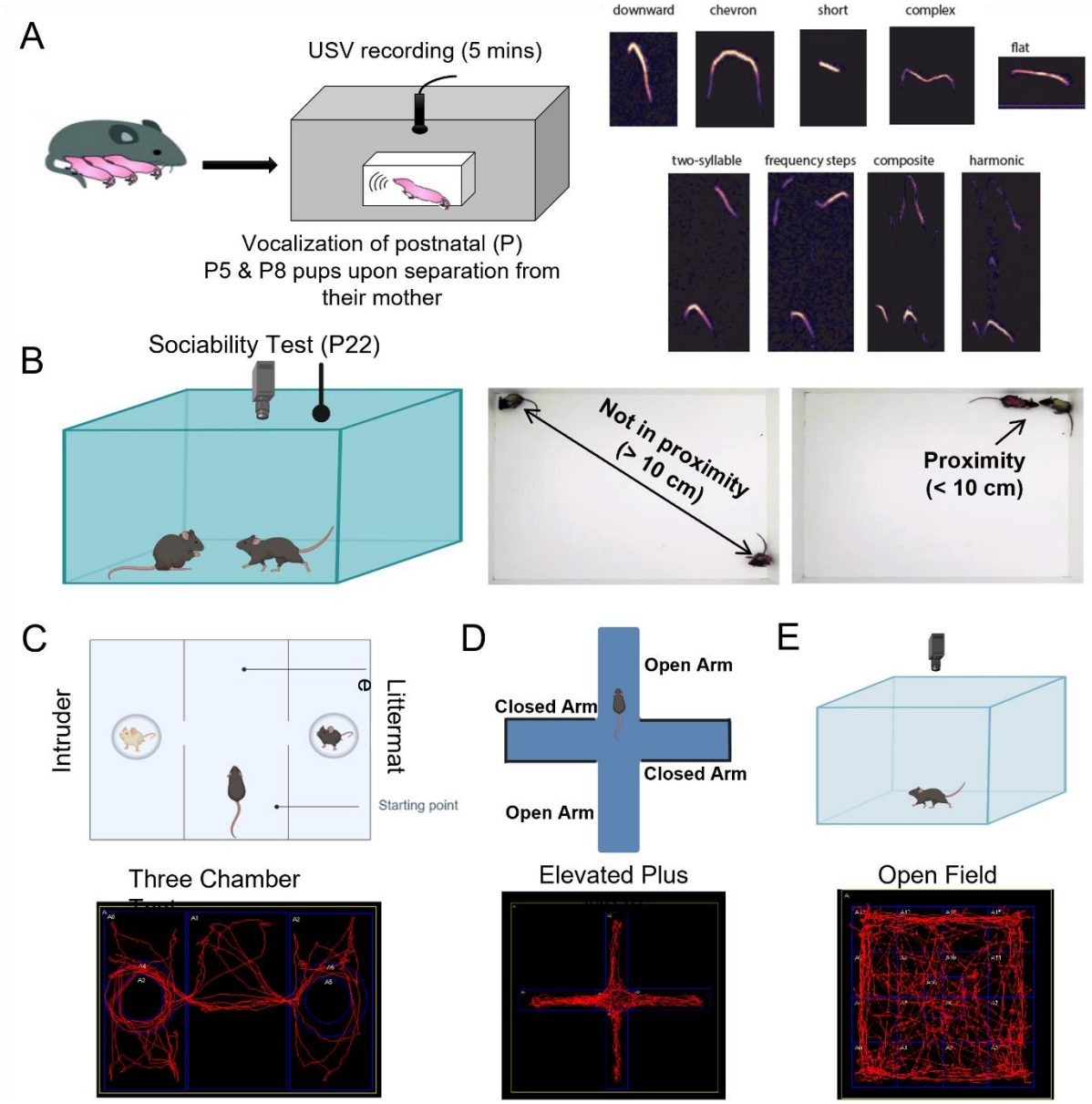


**Figure 2.4***: Overview of all the behavioural experiments performed in this study*

### 2.5.2. USV Recording and Analysis

The equipments used for recording in this study were from Avisoft Bioacoustics (Berlin, Germany). The ultrasonic condenser microphones (CM16/CMPQ) were placed above the

testing area floor (30 cm). The Ultrasound Gate 416 Hz USB audio device recorded, amplified, and digitized the USVs with a 16-bit resolution at 250 kHz. The analysis and classification of the USVs were carried out using the DeepSqueak (Coffey et al., 2019) (MATLAB package). A supervised classification of call types using DeepSqueak was performed.

### 2.5.3. Social Interaction test in adolescent mice

In adolescent mice (PD 21) the social interaction between littermates was performed using a previously described protocol (Silverman et al., 2010). The interactions were recorded between mice of the same sex and withing the same litter. From each litter and further for each gender a mouse was isolated for 24 h from their littermates. Each test composed of two trials. First, the habituation trial involved each test mouse to become habituated to the test arena (open field box; 40 × 40 × 40 cm) for 2 min. Second, in the sociability trial, each of the test mice were reunited with isolated littermate and their interactions were tracked (EthoVision XT video tracking software) for a test period of 5 min along with the recording of USVs in both trails. Interaction was defined on the basis of proximity of the two mice (distance less than 10 cm). Further, the cumulative duration of such interaction was calculated for each pair of mice. A total of 23 WTr6 (16 male and 7 females) and 15 APLP1-/- (10 males and 5 females) were used for the SI study.

### 2.5.4. Three-Chamber Sociability and Social Novelty Test

Postnatal day 40 mice were used to study both sociability and social novelty in Aplp1-/- and WTr6 young mice as described previously (Kazdoba et al., 2016; Moy et al., 2004; Sungur et al., 2016). The experiments were conducted between mice of the same gender and litter. The experimental setup involved a clear Plexiglas rectangular box that has three partitions/chambers with small openings connecting them. In the centre of both the left and right chambers a caged cylindrical enclosure was placed. A total of 3 trials were conducted for each test mouse lasting 5 min an interval of 15 mins between them. The test mice were always introduced into the central chamber for all trials. The test mice were then placed back in their respective cages in between trials to avoid getting them stressed. The habituation trail allowed test mice to explore the arena before the actual trials. For the sociability assay the interaction duration of the test mice with a previously isolated (24hrs prior) littermate (right side) was tracked in comparison to an empty cage on the other end (left). Finally for the social novelty assay the interaction of the test mice was assessed in the presence of an intruder (left) and littermate (right). The software from Sygnis Tracker allowed the tracking of the movements

of test mice within the arena. Based on the time spent in the vicinity of an empty cup, littermate, or intruder a sociability index and a social novelty index was estimated. A total of 24 WTr6 (15 male and 9 females) and 16 APLP1-/- (10 males and 6 females) were used for this study.

### 2.5.5. Open field test

The baseline activity was measured by placing each mouse individually in the centre of a 40 × 40 cm2 white box with 40 cm high walls for 10 min. The light intensity was 290 lx in the centre of the arena. The mouse activity was digitally recorded using a video camera placed 1 m above the centre of the arena. The automatic detection of the mouse path was analysed with the SYGNIS tracker software (SYGNIS). Besides the analysis of the general locomotion, the latency, duration, and the number of visits by the mouse to the inner arena (10 × 10 cm2) away from the wall were calculated for measuring the anxiety level. The open field test was performed at P22.

### 2.5.6. Elevated plus maze test

The elevated plus maze test assesses the anxiety-like behaviour by measuring the conflict between the natural tendency of mice to explore a novel environment and the aversive properties of an open arena118,119. The maze is a crossed-shaped platform (grey opaque plastic material) with equally sized arms (8 × 30 cm2) and a central intersection (8 × 8 cm2), allowing mice to move freely into each zone of the maze. Two of the arms (opposing each other) were flanked by 17 cm high opaque walls (closed arms); the remaining two arms are without walls (open arms). The maze is elevated 70 cm above the floor. The central intersection, open, and closed arms were illuminated by 230, 230 and 160 lx, respectively. Each mouse was placed in the central intersection facing one of the closed arms and allowed to explore the maze freely for 10 min. The behaviour was monitored with a digital camera and tracked with the SYGNIS tracker software (SYGNIS). The ratios of the sum of visits and duration in the two open arms compared to the closed arms were calculated. The elevated plus maze test was performed in 2-month-old adult mice.

## 2.6.Diagnosis of ASD using whole blood

In order to investigate ASD from whole blood, APA differences between individual blood cell populations were analysed by first sorting these individual populations based on marker expression (Figure 3.23).

### 2.6.1. Immunophenotyping of human blood cells

Blood was collected in BD vacutainers from 3 healthy donors (anonymized as A, B & C) and processed for FACS. The remaining blood was processed for whole blood RNA. For FACS sorting, 100 µl of blood was taken for each control/sample (unstained, single colour controls (SCC), fluorescent minus one (FMO) controls and complete staining) was taken in a 2 ml Eppendorf tube. Then, 5 µl of each appropriate antibody was added as follows (Table 3.3).

**Table 2.3** Antibody staining panel for FACS sorting of blood cell types

| Samples (100 µl) Total tubes = 14 | Unstained (100 µl) | | | Single Color Controls (100 µl) | | | | FMOs (100µl) | | | | Sample (100 µl) | | | Total Needed |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Uns A | Uns B | Uns C | CD4-FITC | CD3-PE | CD8-PE-Cy7 | CD19-APC | CD4-FITC | CD3 - PE | CD8 - PE-Cy7 | CD19 - APC | Full A | Full B | Full C | |
| CD4-FITC | - | - | - | 5 µl | - | - | - | - | 5 µl | 5 µl | 5 µl | 5 µl | 5 µl | 5 µl | 35 µl |
| CD3-PE | - | - | - | - | 5 µl | - | - | 5 µl | - | 5 µl | 5 µl | 5 µl | 5 µl | 5 µl | 35 µl |
| CD8-PE-Cy7 | - | - | - | - | - | 5 µl | - | 5 µl | 5 µl | - | 5 µl | 5 µl | 5 µl | 5 µl | 35 µl |
| CD19-APC | - | - | - | - | - | - | 5 µl | 5 µl | 5 µl | 5 µl | - | 5 µl | 5 µl | 5 µl | 35 µl |

The tubes were vortexed to mix the antibodies and incubated for 20 mins in the bark at room temperature (RT). 2 ml of 1x FACS-Lysis buffer was then added to each tube. The samples were then vortexed to lyse the erythrocyte cells. The samples were then centrifuged for 5 mins at 500 g RT and the pellet was saved after discarding the supernatant. This pellet was resuspended in 2 ml FACS buffer and centrifuged for 5 mins at 500 g RT. The resultant pellet was washed 1-2 times with FACS buffer to get a white pellet and resuspended in 400 µl FACS buffer and passed through a 40 µm filter to have a single cell suspension for FACS.

### 2.6.2. FACS sorting of different blood cell populations

The resultant single cell suspension was run on the FACS sorter (FACS Aria) to sort individual blood cell populations. Using the forward and side scatter (FSC & SSC) I distinguished three main blood populations: namely, Neutrophils, Monocytes and Lymphocytes. Each population was sorted into a 1.5 ml Eppendorf tube (containing 100 µl Lysis Buffer) directly from the unstained sample itself. Further, based on marker expression, cells were identified as either T cells ($CD3^+CD19^-$) or B cells ($CD3^-CD19^+$). The T cells were characterized as $CD8^+$ Cytotoxic T cells or $CD4^+$ Helper T cells. Each of these populations were then sorted into a 1.5 ml Eppendorf tube containing lysis buffer.

### 2.6.3. RNA Extraction from cell pellets (blood cell types)

The RNA extraction was performed according to manufacturer's instructions (*PicoPure® RNA Isolation Kit, KIT0204*). In brief, after sorting individual cell populations into lysis buffer, the samples were incubated at 42°C for 30 mins and kept on ice for 30 mins (or stored at -80°C until they further processing). Equal amounts of 70% EtOH (100 µl) was added to the lysate and mixed well by pipetting. The column was conditioned for 5 mins using 250 µl of conditioning buffer and then spun down (16,000 g for 1 min) to discard the flow through. The lysate (+EtOH) was then added to the column and centrifuged at 100 g for 2 min and immediately at 16,000 g for 30 s. The column was washed using 100 µl of wash buffer 1 (WB1) and centrifuges at 8000 g for 1 min. An on-column DNase treatment (5 µl DNase + 35 µl RDD) was performed for 15 mins at RT. Post treatment the column was washed with 40 µl WB1 followed by 2 washes with WB2 (8000 g for 1 min followed by 16,000 g for 3 min to dry). The column was transferred into 0,5 ml Eppendorf tubes and the resulting RNA was eluted in 11 µl of water after incubating for 1 min at RT (1000 g, 1 min & 16,000 g, 1 min). RNA concentration was measured using qubit and RNA quality was assessed using bioanalyzer (samples with RIN > 6.5 used for study).

### 2.6.4. RNA extraction from whole blood (Fresh/Frozen)

Whole blood was collected in EDTA tubes and depending on the need for storage of samples the blood was mixed with 3 volumes of RNALater to prevent RNA degradation. RNA extraction from whole blood was performed according to manufacturer's instruction (*RiboPure™ RNA Purification Kit, Thermo, AM1928*). Either 500µl of fresh blood or 2 ml of blood in RNA Later was used in this study. Samples with RNALater, were additionally centrifuged for 1 min at 16,000 g RT to discard the supernatant. To the pellet/fresh blood 800 µl of Lysis Solution and 50 µl of sodium acetate solution were added and the blood cells were lysed by vigorously vortexing the samples until a homogeneous mixture was observed. To this mixture 500 µl of acid-phenol:chloroform mixture was added and the sample was vortexed. The samples were incubated at RT for 5 mins and then centrifuged for 1 min at 16,000 g RT to separate the aqueous and the organic phases. The upper aqueous phase containing the RNA was then transferred to a new 2 ml Eppendorf tube while the lower organic phase was discarded. To the aqueous phase 600 µl of 100% EtOH was added, and the samples were vortexed. This mixture was passed through a filter cartridge to facilitate the binding of RNA to the column and later centrifuged (16,000 g, 10 s, RT) to discard the flow through. The column was washed with 3 times with wash solutions 1, 2/3 (16,000 g, 10 s, RT) and transferred to a

new collection tube. Finally, the RNA was eluted in 100 µl of preheated (75°C) elution solution after an incubation of 20 s by centrifuging the column at max speed for 30 s. Additionally, DNase treatment was performed to removing contaminating DNA.

### 2.6.5. cDNA preparation for Bulk 3'UTR Libraries

The cDNA prep and the subsequent library prep was performed according to a previously describe protocol (Hennig et al., 2018). 80 ng starting RNA from each sample was taken for each cDNA reaction. For samples where the RNA concentration was low, and the calculated RNA volume was more than 2 µl a 0.8x SPRI bead purification was performed to bind RNA to the beads and then perform an on-bead RT. The RNA was incubated with dNTPs, PEG (50%) and oligo dT primers for 10 mins at 70°C before proceeding to the reverse transcription step. The oligo dT Primers used in this application has a 3' end identifier sequence (CCAA) that allows to specifically enrich for the 3' ends of the transcripts in the library prep step. It also contains an 8bp UMI to identify each unique molecule of mRNA as well as a unique 6 bp sample barcode that allows multiplexing samples for sequencing. The Pre-RT mix reaction is as follows:

**Table 2.4:** Pre-RT mix for cDNA preparation **with** on-bead RT

| Component | Stock to use | Working stock | Dilution from Working | Volume (10µl rxn) |
|---|---|---|---|---|
| dNTP mix | 10 mM each | 1 mM | 1:10 | 1 |
| PEG (50%) | 50% | 5% | 1:10 | 1 |
| Water | - | - | - | 2 |
| Master Mix Total | | | | 4 µl/rxn |
| OligodT (unique) 100µM (1:10 diluted) | 10µM | 1% | 1:10 | 1 µl/rxn |
| Sample | | | | on RNA beads |
| Total | | | | 5 µl/rxn |

**Table 2.5:** Pre-RT mix for cDNA preparation **without** on-bead RT

| Component | Stock to use | Working stock | Dilution from Working | Volume (10µl rxn) |
|---|---|---|---|---|
| dNTP mix | 10 mM each | 1 mM | 1:10 | 1 |
| PEG (50%) | 50% | 5% | 1:10 | 1 |
| Sample (80 ng) | - | - | - | 2 |
| Master Mix Total | | | | 4 µl/rxn |
| OligodT (unique) 100µM (1:10 diluted) | 10µM | 1% | 1:10 | 1 µl/rxn |
| Total | | | | 5 µl/rxn |

After the 10 minutes of incubation the RT master mix is added to the pre-RT mix to synthesise the cDNA. The RT-mix was as follows:

**Table 2.6:** RT mix for cDNA preparation of all samples

| Component | Stock to use | Working stock | Dilution from Working | Volume (10µl rxn) |
|---|---|---|---|---|
| Tris-HCl pH 8.3 | 1M | 25 mM | 1:40 | 0,25 |
| Nacl | 1M | 30 mM | 1:33,33 | 0,30 |
| GTP | 100 mM | 1 mM | 1:100 | 0,10 |
| MgCl2 | 100 mM | 2,5 mM | 1:40 | 0,25 |
| DTT | 100 mM | 8 mM | 1:12,5 | 0,80 |
| RRI | 40 U/µl | 0,5 U/µl | 1:80 | 0,125 |
| TSO (for 3'UTR Seq) | 100 µM | 2 µM | 1:50 | 0,2 |
| Maxima H-minus RT enzyme | 200 U/µl | 2 U/µl | 1:100 | 0,1 |
| Water | - | - | - | 2,88 |
| RT Master Mix Total | | | | 5 µl/rxn |
| Sample + Pre-RT Mix | | | | 5 µl |
| Total | | | | 10 µl/rxn |

Once the 10 µl RT reaction is set up the following program was run on the thermocycler:

**Table 2.7:** Thermocycler cycling conditions for reverse transcription

| Step | Temp | Time |
|---|---|---|
| 1 | 42°C | 90 min |
| 2 | 50°C | 2 min |
| 3 | 42°C | 2 min |
| 4-12 | GO TO 2 | 9x |
| 13 | 85°C | 5 min |
| 14 | 4°C | Hold |

After the RT reaction the cDNA was amplified to have enough material for the tagmentation protocol. The following reaction composition was used:

**Table 2.8:** cDNA amplification PCR reaction mix

| Component | Stock to use | Working stock | Dilution from Working | Volume (25 µl rxn) |
|---|---|---|---|---|
| Kapa HiFi HotStart buffer | 5x | 1x | 1:5 | 5,00 |
| dNTP | 10 mM each | 0,3 mM | 1:33,33 | 0,75 |
| MgCl2 | 100 mM | 0,5 mM | 1:200 | 0,13 |
| IS Primers (Dilute 100 µM 1:10) | 10 µM | 0.1 µM | 1:100 | 0,25 |
| HotStart Polymerase | 1 U/µl | 0,02 U/µl | 1:50 | 0,50 |
| Water | - | - | - | 8,38 |
| Master Mix Total | | | | 15 µl/rxn |
| cDNA after RT | | | | 10 µl |
| Total | | | | 25 µl/rxn |

After setting the PCR reaction the following PCR program was set up.

**Table 2.9:** cDNA Amplification PCR program.

| Step | Temperature | Time | Cycles |
|---|---|---|---|
| Initial denaturation | 98°C | 3 min | 1x |
| Denaturation | 98°C | 20 s | **14 cycles** (Go to step 2 --> 13x) |
| Annealing | 65°C | 30 s | |
| Elongation | 72°C | 4 min | |
| Final elongation | 72°C | 5 min | 1x |
| Hold | 4°C | ∞ | |

Post PCR amplification, from fresh whole blood RNA samples, two rounds of a 0.8X SPRI bead purification was performed after the RT reaction to avoid primer carry over into the cDNA. Finally, the resulting cDNA was eluted in 11 µl of nuclease-free water. 1 µl of cDNA was used to quantify its concentration using qubit high sensitivity DNA kit and another 1 µl was used to run on the bioanalyzer to look at cDNA profiles.

### 2.6.6. RBC transcript depletion

The main issue for patient samples was that they were frozen for transportation. This caused inefficient RBC lysis as this step relied on the cell membrane integrity of RBCs that therefore resulted in RBC transcript contamination (HBB, HBA 1/2) of PBMC cDNA. In order to deplete these contaminating transcripts, I tested 2 strategies. First, I used the SPRI bead-based size exclusion of the RBC transcripts. As the size of the RBC transcripts were around 650-700 bp a 0.5x SPRI bead purification resulted in a substantial depletion of these transcripts. Two rounds of 0.5X SPRI bead purification was performed after RT in addition to the purification steps after cDNA amplification to ensure substantial reduction of RBC transcript

contamination. Despite such a depletion, the resulting 3'UTR libraries sequenced still had a large number of reads used up by the RBC transcripts. In addition to this, the size-based transcript selection arbitrarily excluded all transcripts below a certain length This implied that there was a bias to longer transcripts while shorter transcripts and even shorter 3'UTRs from the same gene could be excluded from the 3'UTR libraries. I therefore adapted a second strategy which took advantage of locked nucleic acid (LNA) primers that bind to the RBC transcripts only and prevented its specific amplification (Everaert et al., 2022). HBB_LNA (+G+C+T+C+G+G+T+G+C+C+T+T+T+A+G+T+G+A+T+G) & HBA1/2_LNA (+A+C+C +A+A+G +A+C+C +T+A+C +T+T+C +C+C+G +C+A) Thus, I could enrich the PBMC transcripts without adding any size based bias to the 3'UTR libraries and subsequent APA analysis.

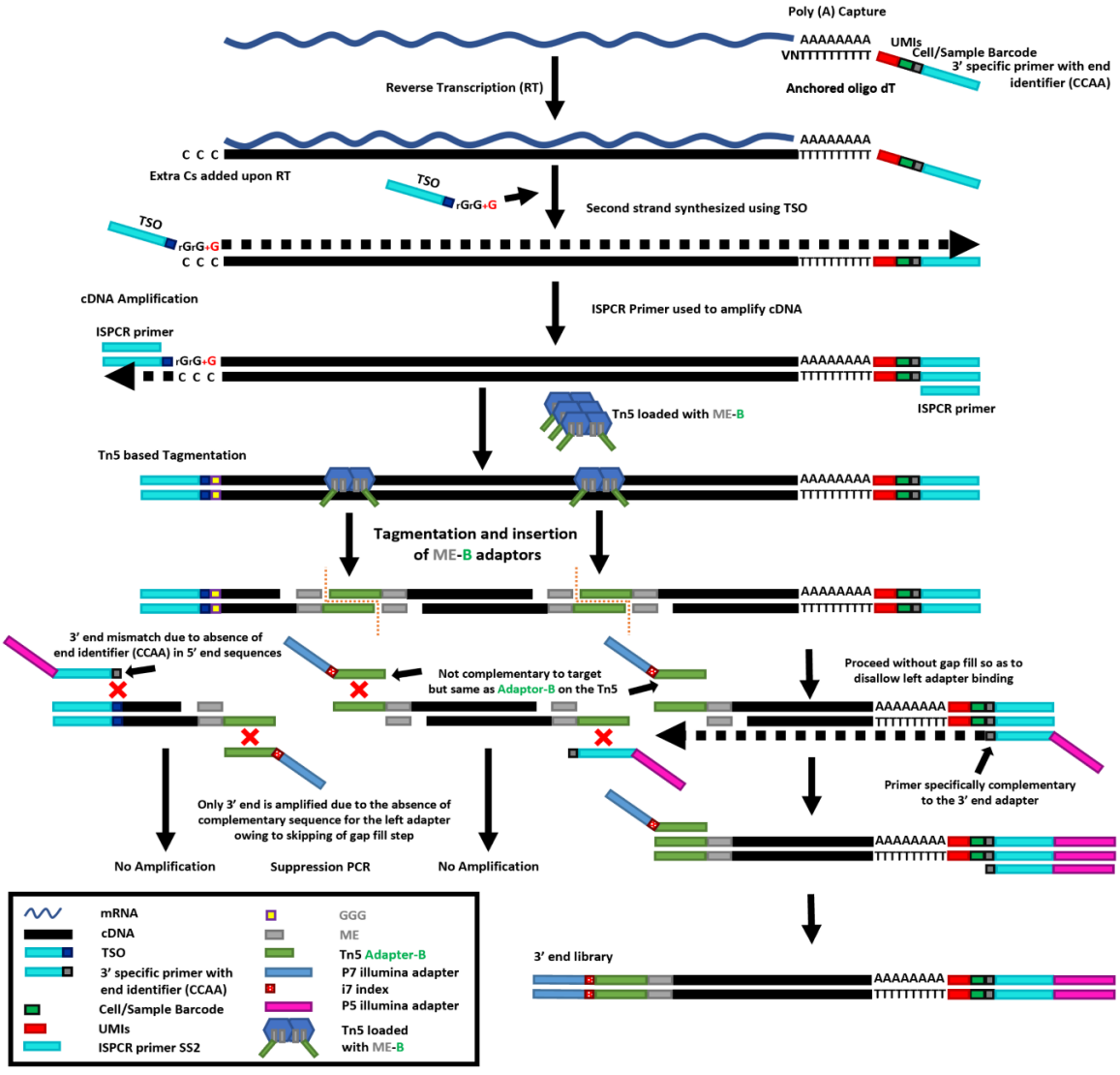


**Figure 2.5***: **Overview of the protocol for the preparation of 3'UTR library** for sequencing the APA landscape from whole blood*

### 2.6.7. Modified tagmentation approach for 3'UTR library preparation

The previously described tagmentation protocol by Hennig et al was used to enrich for the 3' ends of mRNA transcripts (Hennig et al., 2018). In brief, the transposase (Tn5; 0.85 mg/ml) was loaded with the adapter Tn5ME-B/Rev (12.5 µM) (1:10; Tn5:adapter) by incubating the mix at 23°C for 60 mins. This loaded Tn5 was then added to the cDNA and the tagmentation was performed for 5 mins at 55°C (5 µl rxn vol). The following tagmentation reaction was used.

**Table 2.10:** Tagmentation reaction mix for 3'UTR library prep

| Component | Final Conc | Amount (µl) |
|---|---|---|
| cDNA | 1 ng | 1,72 |
| 4x Tagmentation Buffer | 1x | 1,25 |
| DMF (100%) | 25% | 1,25 |
| Loaded Tn5 12.5 µM no Spacer | 120 ng/µl | 0,78 |
| Total | | 5 |

Immediately after the tagmentation, 1.25 µl of 0.2 % SDS was added to the reaction and incubated at RT for 5 mins to strip off the transposase. The normal gap filling step after tagmentation was avoided in order to enrich specifically for the 3' ends of the mRNA transcripts (Figure 2.5). In order to amplify the tagmented product the following reaction was set up (Table 2.11) with using the PCR program described in Table 2.12.

**Table 2.11:** Library PCR reaction mix post tagmentation of PBMC cDNA

| Component | Final Conc | Amount (µl) |
|---|---|---|
| Kapa HiFi HotStart buffer (5x) | 1x | 3,25 |
| DMSO (take from cell culture) | 5% | 0,81 |
| PE1.SmartSeq2 (12,5 µM) | Dilute the 100µM primer 1:8 before use | 1,5 |
| dNTPs (10 mM each) | 0,3 mM | 0,49 |
| Kapa HiFi Polymerase (1U/µl) | 0,02 U/µl | 0,33 |
| Nuclease Free Water | | 2,12 |
| Total master mix | | 8,5 |
| i7 index (12.5 µM add individually) | | 1,5 µl per rxn |
| Tagmented Sample | | 6,25 |
| **Total** | | 16.25 |

**Table 2.12:** Library PCR program to amplify final 3'UTR libraries

| Step | Temperature | Time | Cycles |
|---|---|---|---|
| Initial denaturation | 95°C | 3 min | 1x |
| Denaturation | 95°C | 10 s | 17 total cycles (Go to step 2 --> 16x) |
| Annealing | 55°C | 30 s | |
| Elongation | 72°C | 30 s | |
| Final elongation | 72°C | 5 min | 1x |
| Hold | 4°C | ∞ | |

Finally, two rounds of 0.7X bead purification was performed and the final 3'UTR libraries were eluted into 11µl of nuclease-free water. The library concentrations were determined using qubit DNA high sensitivity kit (1 µl) and then the average library peak size was determined by running 1 µl of library on the bioanalyzer. The equimolar volumes of samples were pooled for sequencing using a combination of unique dual indexed libraries (i7 + sample barcode). The final libraries were run on the NextSeq2000 sequencing platform using a P3 200 cycles kit with paired end sequencing (130+8+0+100; R1+i7+i5+R2). The amount of the library loaded on to the sequencer was 1100 pM (with 14% PhiX spike-in) and using a custom polyA-seq primer (Hennig et al., 2018) for sequencing read1. The read2 was long enough (100bp) to go through the poly(A) tail into the ends of the 3'UTR thus allowing to identify true 3'UTR ends.

**Table 2.13:** Primers used for 3'UTR bulk sequencing library preparation protocol

| Primer | Sequence |
|---|---|
| 3UTRseq-oligo-dT | AAGCAGTGGTATCAACGCAGAGTAC**CCAAATCGTG**NNNNNNNVTTTTT TTTTTTTTTTTTTTVN |
| 3UTRseq-TSO | AAGCAGTGGTATCAACGCAGAGTACATrGrG+G |
| PE1.Smart-seq2 | AATGATACGGCGACCACCGAGATCTACACAAGCAGTGGTATCAACGCA GAGTACCCAA |
| polyA-seq | AAGCAGTGGTATCAACGCAGAGTACCCAA |
| ISPCR Primer | AAGCAGTGGTATCAACGCAGAGT |

## 2.7.Bioinformatics analysis

The bioinformatics analysis for this study have been performed by Dr. Manuel Göpferich and Dr. André Lopes Martins Macedo from the department of molecular neurobiology, DKFZ. Most of the bioinformatics pipelines used to analyse the data generated as a part of this thesis have been described in greater detail in the previously published PhD thesis (Göpferich, 2020) as well as our bioRχiv paper (Göpferich et al., 2020) and is not the focus of this study. Additionally, analysis that improved upon previous methods and analysis of additional data not included previously were solely performed by Dr. André Lopes Martins Macedo for the purpose of this thesis. All credit for the bioinformatic analysis goes to the above-mentioned people. Only those aspects of the analysis that deviates from the originally published methods (Göpferich, 2020; Göpferich et al., 2020) and discussed here for more clarity.

The alignments for all sequencing data (mouse *in vivo & in vitro*, human blood ASD vs Control) in this study were performed using STAR (Dobin et al., 2013) instead of bowtie.

Read1 was used to precisely map the 3' sequenced reads onto genomic coordinates and the 3'UTR regions were identified using Ensembl annotations. 3'UTR peak calling was performed using kernel density estimation (KDE) (McCarthy & O'Callaghan, 2014). While this method has been used to solve for similar problems of peak calling in data such as chromatic accessibility, this study is the first to use it for 3'UTR peak calling. For the mouse *in vivo* data, DBScan (Ester et al., 1996) was used to isolate the NSC lineage and remove other non-NSC lineage components such as oligo, OPCs, etc. The transcriptome information was used to define different cell types within the NSC lineage as previously described (Göpferich et al., 2020; Kalamakis et al., 2019; Llorens-Bobadilla et al., 2015). A diffusion pseudotime (zero branching) was produced to assign a temporal order for each cell within the NSC lineage transition (Haghverdi et al., 2016).

In order to assess the APA changes along the NSC lineage, 4 additional replicates were added to the already 2 replicates analysed previously (Göpferich, 2020; Göpferich et al., 2020). Two multinomial regressions were performed for each gene per replicate; one with the pseudotime effect taken into consideration and another where the pseudotime information was not important. The significance in the deviation between the two MNRs was calculated using a chi-squared test and finally a one sample t-test was used to determine APA genes with significant changes along the NSC lineage combining all replicates. A correlation of 3'UTR length vs pseudotime provides the information of 3'UTR lengthening (positive correlation) or shortening (negative correlation) along the NSC lineage. The splines for the APA changes per gene were produced as previously described (Göpferich, 2020). GSEA for biological process and disease ontologies were performed now for the 6 replicates as before (Göpferich, 2020; Subramanian et al., 2005).

In order to correct for the pseudotime effect on APA changes between APLP1-/- and WT mice the cells on the pseudotime were split into bins and similar number of cells per bin was taken between each genotype. For each gene, a total of 50 iterations of randomization to acquire similar cell densities/bin was performed. On each of these 50 randomized pseudotimes (normalized by density/bin), two MNRs/pseudotime were performed for each gene per replicate; one with the genotype effect taken into consideration and another where the genotype information was not important. The significance in the deviation between the two MNRs was calculated using a chi-squared test (yielding 50 p-values) and finally a one sample t-test was used to determine APA genes with significant changes along the NSC lineage combining all replicates. The significance for each gene was defined as the median significance of the 50

iterations. Variations in PAS usage were determined via EMD as either 3'UTR lengthening (positive correlation) or shortening (negative correlation) over the NSC lineage. The significant genes were highlighted based on their MNR median p-values from the chi-squared test.

The linear regression model (LM) used for the re-analysis of the Velmeshev et al snRNASeq data to investigate APA changes in ASD vs control brains only accounted for the effect of sex and brain regions but did not account for other variables such as RNA quality (RIN), postmortem intervals (PMI), sequencing batch etc; lm(len ~ Sex + Region + Diagnosis). The newer LM has taken into account all these other variables into the model and therefore more accurately looks at APA changes primarily as an effect of diagnosis without any compounding factors. Fit1 includes the effect of diagnosis as well as the other variables while fit0 excludes diagnosis. The model is represented as follows:

fit1 <- lm(len ~ Sex + Brain region + RIN + PMI + Capbatch + Seqbatch + Diagnosis)

fit0 <- lm(len ~ Sex + Brain region + RIN + PMI + Capbatch + Seqbatch)

Additionally, in the previous analysis significant genes were highlighted on the basis of their nominal p-values while, in the current study the significant genes were identified after correcting for multiple testing (Benjamini & Hochberg, 1995).

In order to check for deviations in APA changes of individual blood cell populations compared to whole blood, a Pearson correlation as performed on average 3'UTR sizes per gene for each comparison. For the ASD blood APA data, a Dirichlet-MNR (Chen & Li, 2013; Mosimann, 1962) was performed to correct for overdispersion. An EMD was performed to visualize the APA changes in whole blood between ASD patients and healthy controls. Positive changes implied 3'UTR lengthening in ASD while negative changes implied shortening in the same. The significant genes (coloured) were highlighted based on Dirichlet-MNR. From the list of significant APA genes determined using the Dirichlet-MNR the top 10 candidates were selected on upon meeting the following criteria: a permutation p-value < 0.00001, transcripts count > 1000 UMIs, total number of PAS per gene/3'UTR < 10 & sorted by Jenson-Shannon (JS) distance.

# 3. Results

The results shown in thesis are a combination of both wet lab and bioinformatic analysis that have been performed to investigate the role of APA in neurogenesis and its dysregulation in neurodevelopmental disorders. Additionally, I tried to investigate the use of such an APA dysregulation signature as a tool for the diagnosis of ASD. **All my individual contributions to this research have been indicated in first person throughout this section.** Different data generated and the respective contributions for this study have been listed as follows:

1. NSC lineage progression dataset_1 (Kalamakis et al., 2019)
2. NSC lineage progression dataset_2 (unpublished; Dr. Santiago Cerrizuela)
3. NSC lineage *in vitro* system establishment (**Nikhil Oommen George**)
4. NSC lineage *in vitro* scRNASeq (**Nikhil Oommen George**)
5. NSC lineage *in vitro* proteomics (**Nikhil Oommen George**, Dr. Daria Fijalkowska)
6. CPEB4-RIP assay (Rosa Pascual, Dr. Raul Mendez)
7. APLP1-/- vs WT NSC Lineage scRNASeq (**Nikhil Oommen George**)
8. APLP1-/- vs WT NSC Lineage FACS & IF (**Nikhil Oommen George**)
9. Co-IP of APLP1 with CPEB4 (Ana Domingo Muelas, Alex Bizyn, Dr. Isabel Fariñas)
10. CPEB4 & APLP1 ALFA-Tag CRISPR KI & LAMA ((Marlene Aßfalg, 2021), Tobias Hub, Matthias Kaltenböck under the supervision of **Nikhil Oommen George)**
11. APLP1-/- vs WT Behaviour Studies (**Nikhil Oommen George**)
12. ASD Brain dataset (Velmeshev et al., 2019)
13. Development of 3'UTR bulk Sequencing (**Nikhil Oommen George**)
14. Whole blood vs Cell types_3'UTR bulk Seq (**Nikhil Oommen George**)
15. ASD vs Healthy Controls_3'UTR bulk Seq (**Nikhil Oommen George**)

In addition to the generation of the data, bioinformatics analysis presented in this thesis have been done the following individuals. Dr. Manuel Göpferich: Developmental of the 3'UTR analysis pipeline, Splines fits for APA changes over pseudotime. Analysis of *in vitro* NSC scRNASeq and proteomics (translation index, TI). Analysis of the CPE motif CPEB4-RIP assay. Analysis of APLP1-/- vs WT APA changes. Analysis of Velmeshev et al ASD brain data for APA changes. GSEA. Dr. André Lopes Martins Macedo: Modification of the 3'UTR pipeline and reanalysis of the APA changes in the NSC Lineage (4 additional replicates). GSEA. Effect of pseudotime on APA changes in APLP1-/- vs WT mice. Re-analysis of Velmeshev et al ASD brain data for APA changes with multiple correction. Development of pipeline for analysis 3'UTR bulk Seq. Analysis of APA in Whole blood vs cell type. Analysis of APA changes in ASD vs Controls from whole blood. Development of the ASD classifier.

Dr. Manuel Göpferich and I had begun this project as a tandem to investigate the role of APA in the NSC lineage and its dysregulation in neurodevelopmental disorders such as autism.

While I focused on the experimental part of the project, that also included method development whenever necessary, Dr. Göpferich focused on development of pipelines for the bioinformatic analysis. Additionally, I also worked closely with him providing the biologist's perspective in interpreting the results of the analysis. In addition to Dr. Gopferich's data this thesis improves on the results by adding 4 more replicates for the NSC lineage scRNAseq. Further, I established and characterized the *in vitro* system to study active and quiescent NSCs in culture that allowed to perform experiments that were restricted due to *in vivo* cell numbers. In that regard, I performed the scRNASeq on the *in vitro* NSCs to investigate APA in this system and additionally I generated the proteomics samples that helped investigate the impact on translation. I identified APLP1 as a potential APA regulator in the NSC lineage and worked on developing tools to study the APLP1/CPEB4 axis (ALFA-Tag KIs, CPEB4-ALFA-Tag IP, LAMA for CPEB4 mis-localization). Further, I generated the scRNASeq data of APLP1-/- vs WT mice in order to investigate the dysregulation of the APLP1/CPEB4 axis. I performed the functional characterization of the NSC lineage in APLP1-/- mice vs WT that identified alterations in the lineage cell distributions. Further, I performed an extensive behavioural characterization of the APLP1-/- mice compared to their WT counterparts. Owing to the ASD brain phenotype of 3'UTR lengthening, I investigated the APA landscape in blood and its potential use for ASD diagnosis. Firstly, in healthy individuals to study cell type APA differences from whole blood. Secondly in ASD individuals compared to healthy controls. To facilitate this, I established the 3'UTR bulk library protocol for whole blood and the performed the sequencing and generated 3'UTR bulk data for 116 ASD & control samples. I performed optimizations on the protocol that helped reduce RBC transcript contamination and thereby improve the data quality of the samples used to build the ASD classifier. Dr. André Lopes Martins Macedo replaced Dr. Gopferich in our tandem while we shifted gears in our attempt to use APA dysregulation for ASD diagnosis. Some of the figures shown in this thesis were previously presented by Dr. Göpferich in his PhD thesis, that focused more on the bioinformatics pipeline used for APA analysis using single cell transcriptomics (Göpferich, 2020). These figures are solely shown here to present a clear narrative and in no way is meant to infringe upon his work (Göpferich, 2020). Figure credits due unto Dr. Göpferich wherever applicable and will be mentioned in the respective figure legends.

### 3.1.Alternative Polyadenylation in the NSC lineage

To investigate any evidence of alternative polyadenylation (APA) in the SVZ NSC lineage, single cell transcriptomics (scRNASeq) on cells spanning the entire NSC lineage from a total of 6 adult mice was performed. The cells were FACS sorted (GLAST+Ter119-CD45-O4-) to exclude non-NSC lineage cells such as endothelial (Ter119), Oligodendrocytes (O4) and Microglia (CD45) while being enriched for the NSC lineage marker GLAST, before being processed for scRNASeq using 10X Chromium (Figure 3.1). Individual cell & transcript identity were preserved using the cell barcode and unique molecular identifiers (UMIs) respectively (Figure 3.1B).

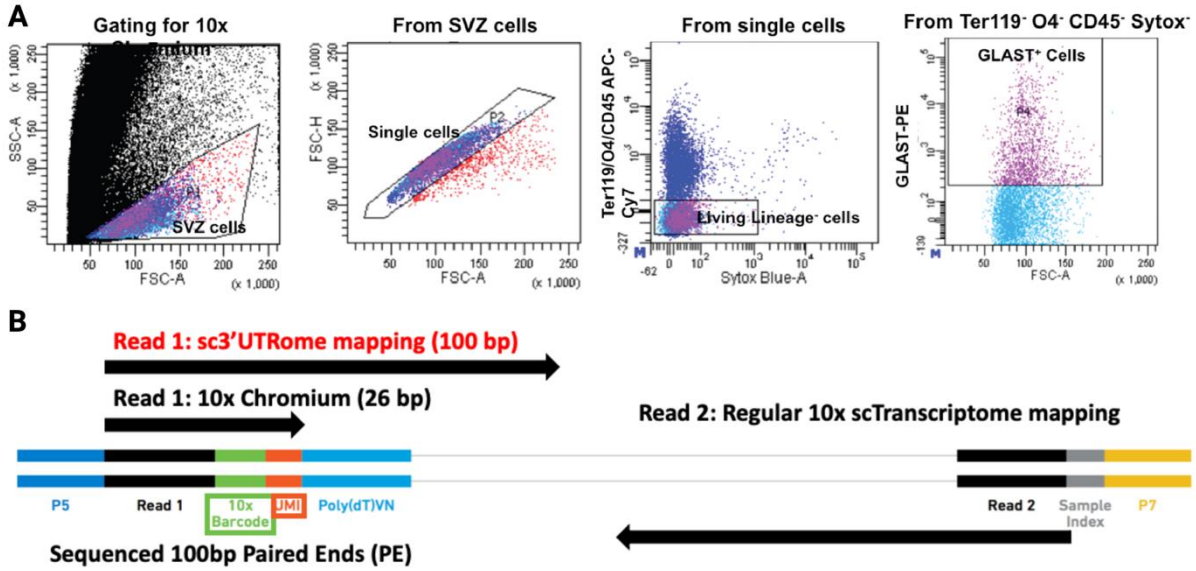


**Figure 3.1:** ***FACS sorting of the NSC lineage and scRNASeq sequencing strategy*** *A: FACS gating strategy to sort for NSC lineage cells (GLAST+Ter119-CD45-O4-) for scRNAseq using 10x chromium. B: Modified sequencing strategy to study APA. A 100 bp PE sequencing was performed instead of the conventional 98 (R2) + 26 (R1) used to sequence 10x Chromium libraries. This modification allowed for the identification of the very ends of the 3'UTRs using an extended read 1 that read through the Cell barcode, UMI & the poly(A) tail right into the 3'UTR*

The identity of single cells along the NSC lineage trajectory was determined based on gene expression and characterized into different activation states (broadly; qNSCs, aNSCs, TAPs & NBs) as previously described (Figure. 3.2A) (Kalamakis et al., 2019; Llorens-Bobadilla et al., 2015). These cells were then assigned a pseudotime value that corresponded to their position along a differentiation trajectory, starting from qNSCs (0) towards more differentiated NBs (1) (Figure 3.2B). Cell densities along the pseudotime reflected two major events/checkpoints/bottlenecks along the NSC trajectory namely, activation (qNSCs to aNSCs) and differentiation (aNSCs to NBs) (Figure 3.2C). Differential gene expression between

various NSC activation states has been extensively reported in previous studies and will not be explored as a part of this thesis (Kalamakis et al., 2019; Llorens-Bobadilla et al., 2015).

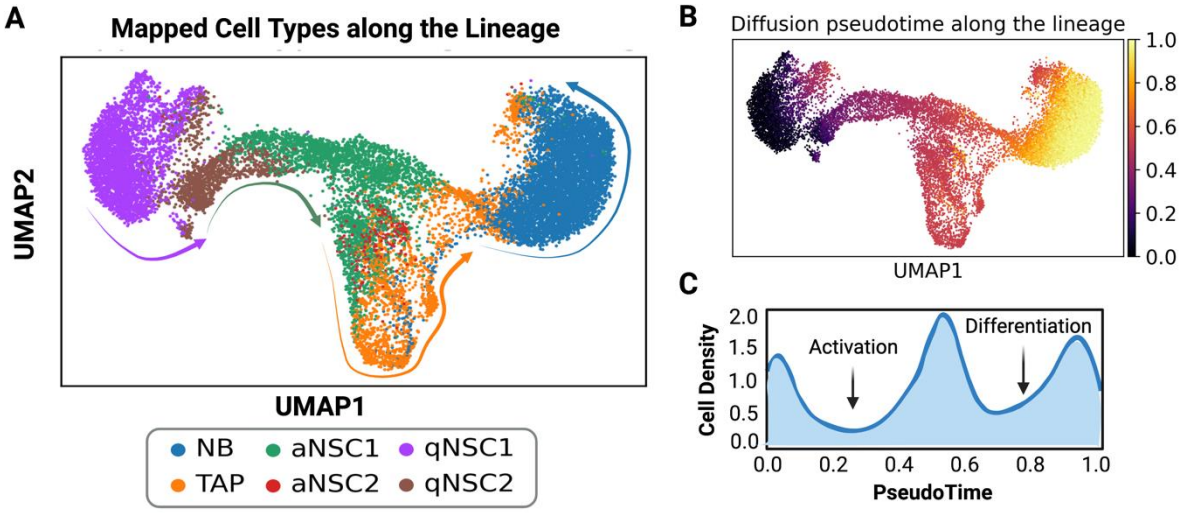


**Figure 3.2:** ***Pseudotime ordering of the NSC lineage****. A: Visualization of different cell types of the NSC lineage after dimensionality reduction using UMAP. The arrows indicate the direction of the NSC lineage progression from quiescent NSCs (qNSC1, qNSC2), active NSCs (aNSC1, aNSC2), transit amplifying progenitors (TAPs) to neuroblasts (NBs). B: The NSC lineage UMAP coloured according to assigned pseudotime. Black/Purple (Quiescent), Red/Brown/Orange (aNSCs/TAPs), Yellow (NBs). C: Plot showing cell densities of different cell populations along the pseudotime indicating 2 major checkpoints namely activation (pTime~0.3) & differentiation (pTime~0).*

The modified sequencing strategy (extended Read1; Figure 3.1B) allowed for the use of scRNAseq data to also investigate APA in the NSC lineage. Using the extended reads, 3' tail peaks were generated that identified more precisely the ends of 3'UTRs for every molecule of every gene sequenced at single cell resolution along the NSC lineage (Figure. 3.3A). For example, a look at gene Pea15a identifies 2 predominant PAS usage (~1100 & ~1900 bp) (Figure 3.3A). Many other genes showed such differential PAS usage. For all genes these 3'tail peaks were found to be in close proximity (typically 20 bp downstream) of a PAS (AAUAAA) implying that the 3'tail peaks did indeed correspond to the very ends of a 3'UTR (Figure 3.3B). Further, looking at the PAS usage dynamic changes along the lineage were observed. For example, in the same gene Pea15a a shift to a proximal PAS usage was observed upon lineage progression (qNSCs to NBs) (Figure 3.3A). To investigate this on a global scale a correlation coefficient for each gene was computed comparing average 3'UTR lengths for each cell vs the pseudotime. In the transitioning from quiescent NSCs to differentiated NBs, shortening or lengthening of 3'UTRs were indicated by a negative or a positive correlation coefficient, respectively. Genes with significant changes in PAS usage were identified using a multinomial regression (Figure 3.3C). To better visualize these more dynamic changes, fractional PAS

usage over the lineage binned along the pseudotime (Figure 3.3D) were smoothened using multinomial spline regression (Figure 3.3E). While the linear regression shows the overall trend of the distal PAS for Pea15a, shortening, (Figure 3.3A; right) the spline curves show a more dynamic change along the lineage (Figure 3.3D-E) The bioinformatics pipeline previously developed for investigating APA changes in the NSC lineage has been adapted for this study (Göpferich, 2020).

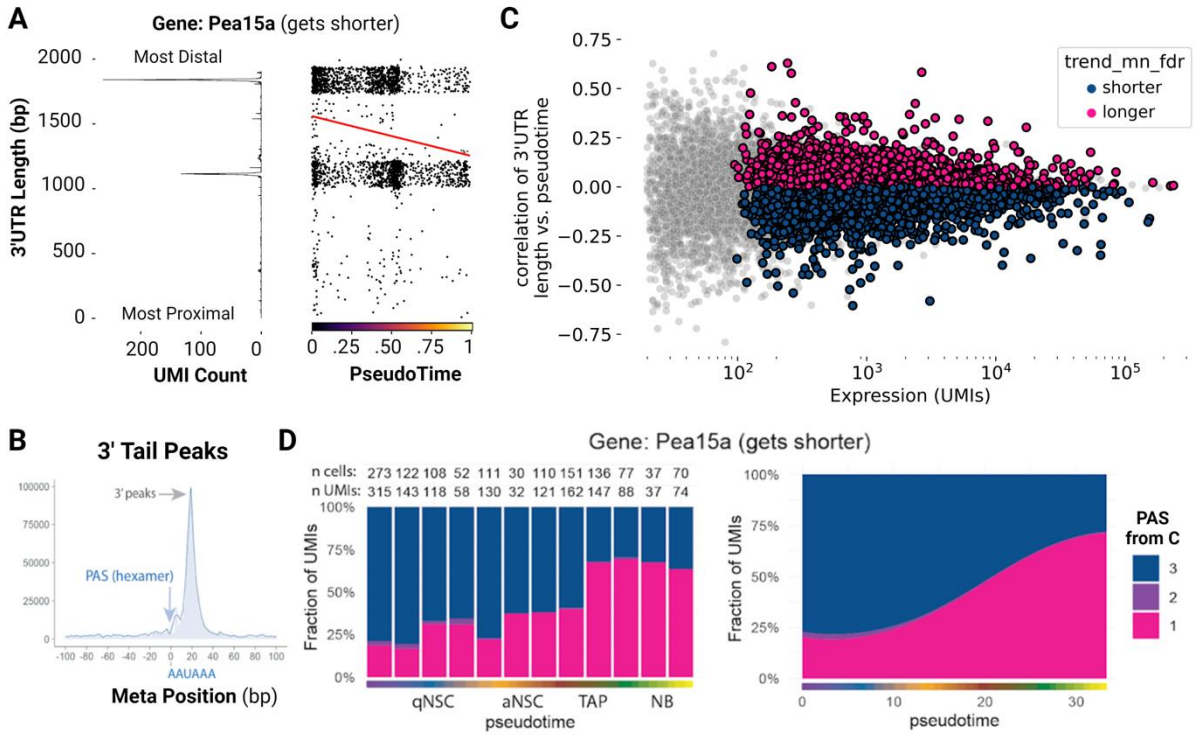


**Figure 3.3:** ***Alternative polyadenylation in the NSC lineage.*** *A: Distribution of different 3' tail peaks showing an example gene Pea15a (phosphoprotein enriched in astrocytes 15A) across the NSC lineage. The 3'UTR for Pea15a becomes shorter upon differentiation. Left panel: Mapping positions of the 3' tail peaks (y-axis) and total UMI counts aggregated over all cells, Right panel: Average mapped position of the 3'UTR for each cell (y-axis; to avoid overplotting random jitter was added to y-coordinate) along the pseudotime (x-axis, same scaling as colour scale in Fig. 3.2B), the red line indicates linear regression along the pseudotime. B: Plot showing the correlation of mean 3'UTR length per cell vs. pseudotime (y-axis) for each gene with alternative polyadenylation (APA) against their summed expression (x-axis), 3'UTR shortening, or lengthening trends estimated using multinomial regression; coloured as blue and magenta points respectively. C: Density plot showing the relative position of 3' mapping positions to the Poly A signal (PAS; AAUAAA) summed over all expressed genes. D: Fraction of UMIs for each PAS usage from C per pseudotime bin. Shown for Pea15a (blue = distal, long 3'UTR, magenta = proximal, short 3'UTR), E: Multinomial regression splines after smoothening the curves from individual bins (data from left panel). Figure B&D from* (Göpferich, 2020).

A gene set enrichment analysis (GSEA) on the genes with significant APA changes along the NSC lineage revealed enrichment for GO terms such as regulation of mRNA stability. Other

GO terms were related to synaptic transmission, dendritic spine development, organization of dendritic spines, synapse & neuronal projections to name a few (Figure 3.4A). Deficits in these processes are characteristic of neurodevelopmental disorders (Lo & Lai, 2020; Nakai et al., 2018; Pan et al., 2019; Zieger & Choquet, 2021). These genes were also enriched for categories related to neurodevelopmental disorders especially autism spectrum disorder (ASD) using a disease related database (Disease ontology; DO) (Figure 3.4B). It was hypothesized that APA could play a role in maintaining the NSC lineage progression during homeostasis, while its dysregulation could result in Neurodevelopmental disorders such as ASD.

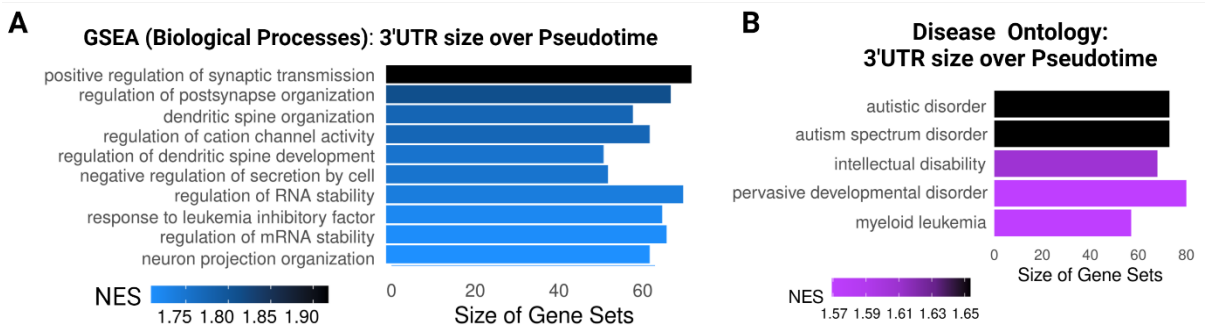


**Figure 3.4:** ***GSEA of APA genes along the NSC lineage*** *A, Gene set enrichment analysis (GSEA; Biological Processes) of APA genes ranked by the strength of 3'UTR changes along the lineage, B genes from Fig 3.3C translated into human gene orthologs, a high normalized enrichment score (NES) indicates association of disease and 3'UTR changes.*

### 3.1.1. Impact of APA on Translation

Given that APA changes along the NSC lineage and the affected genes are involved in processes such as RNA stability, RNA processing etc, I tried to investigate the impact of these changes in regulating translation. Further, I hypothesized that APA could affect the translation efficiency of various transcripts and thereby fine tune the cell's proteome which ultimately could define and alter cell states.

### 3.1.2. A system that mimics active and quiescent NSCs in vitro

To assess the impact of APA on translation, I established an *in vitro* system that mimicked the active and quiescent states of NSCs (Figure 3.3 A). I treated adherent NSC cultures with or without BMP4 for 3 days to produce quiescent and active NSCs respectively. After 3 days of BMP4 treatment, I found that the cells showed a reduction in their proliferative capacity while the untreated active NSCs proliferated much more (Figure 3.5A-C). I observed that the BMP4 treated *in vitro* quiescent NSCs (qNSCs) displayed a "fried egg" morphology upon BMP4 treatment while the active NSCs grew as spindle shaped cells (Figure 3.5A). Further, using immunofluorescence, I was able to characterize that the *in vitro* qNSCs (BMP4 treated) had a

significant reduction in the expression of MKi67, a proliferation marker, compared to active (EGF-treated) NSCs (Figure 3.5 B-C).

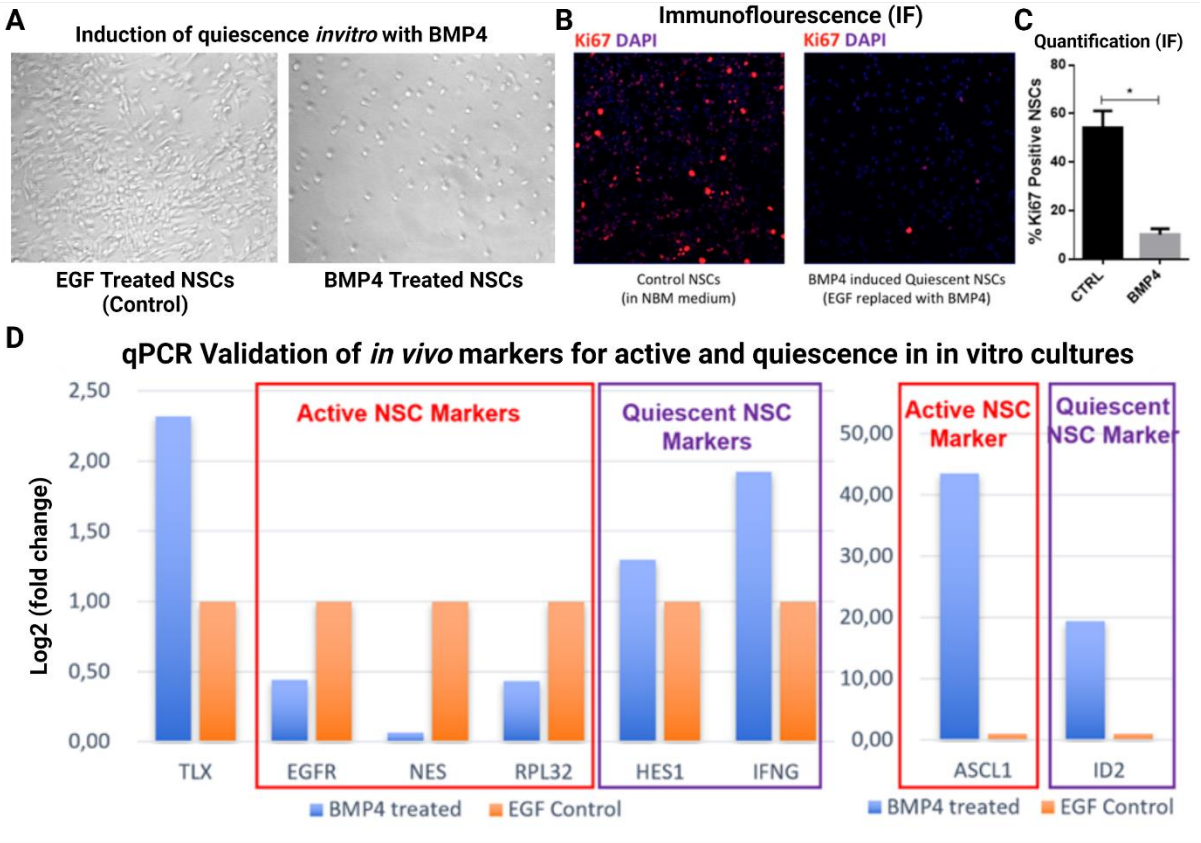


**Figure 3.5:** ***The in vitro system to mimic active and quiescent NSCs.*** *A: Brightfield images (10x objective) of NSCs treated for 3 days with EGF (left) or BMP4 (right). B: Immunofluorescence staining of EGF treated (control; left) or BMP4 treated (right) with MKi7 (red; proliferation marker) and DAPI (blue; nuclei marker). C: Quantification of proliferation (%MKi67$^{+}$ Cells) in active (EGF) and quiescent (BMP4) NSCs. D: Validation of different active and quiescent in vivo marker expression by qPCR in EGFA/BMP4-treated cultured NSCs. Normalized expression relative EGF treated NSCs shown as log2(fold change).*

In order to look at *in vivo* marker expression for active and quiescent NSCs, I initially performed qPCR validations on *in vitro* NSCs, and found coherent differences for almost all genes investigated. I found that active markers such as EGFR, NES and RPL32 showed a reduced expression in BMP4-treated NSCs, while HES1, IFNG and ID2 showed increased expression consistent with qNSCs (Figure 3.5D). I further found that the expression of ASCL1 was significantly higher in BMP4-treated quiescent NSCs despite being an active marker. I hypothesized that this was likely due to the interference of ASCL1 expression via the BMP signalling pathway (Blomfield et al., 2019).

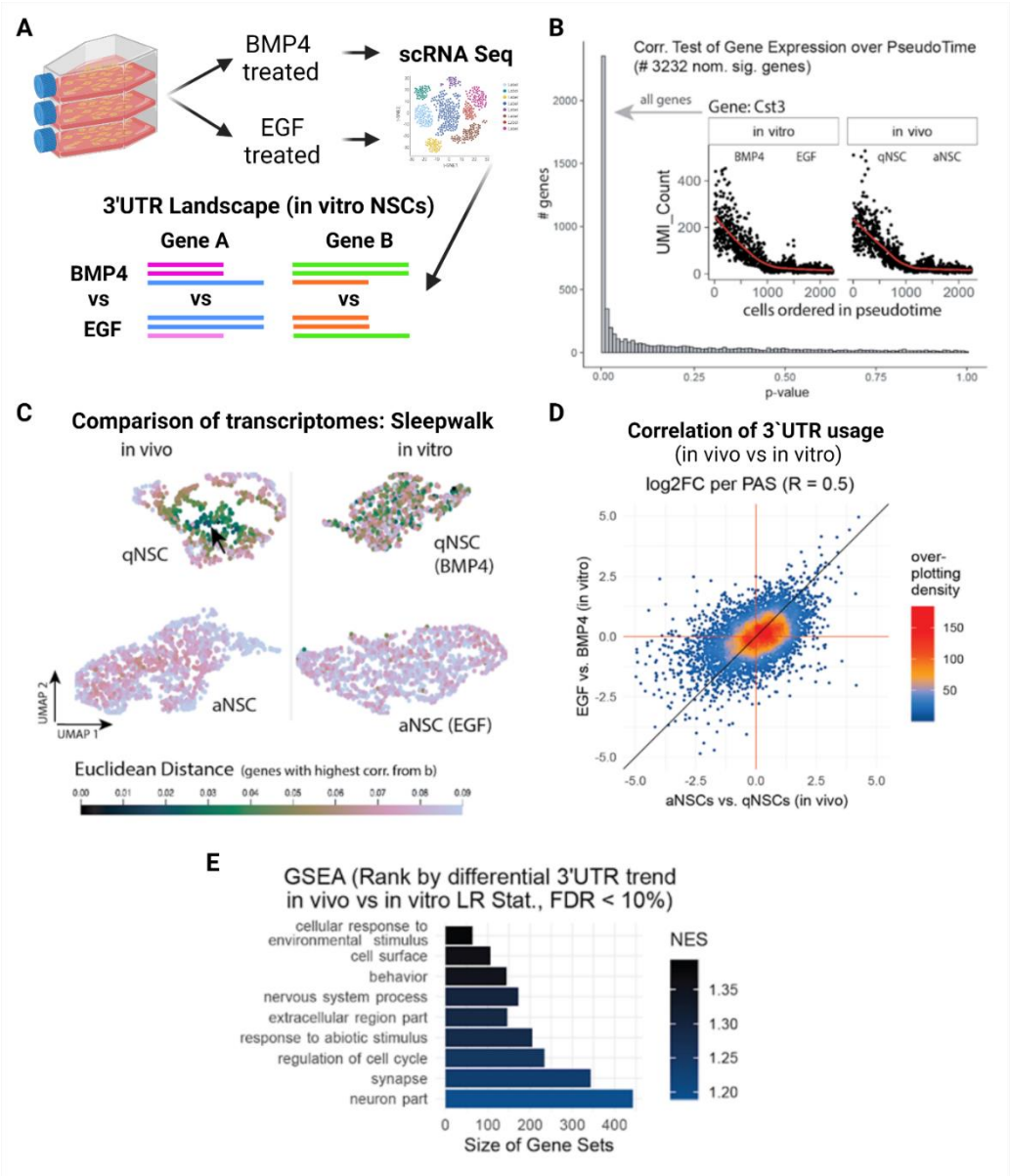


**Figure 3.6:** ***Analysis of the scRNASeq data from in vitro NSCs.*** *A: Schematic overview to generate scRNA transcriptome and APAome from in vitro cultured quiescent (BMP4 treated) and active (EGF treated) NSCs (Created with BioRender.com). B, Correlation test comparing gene expression between in vitro and in vivo NSC lineage counterparts (aNSCs/EGF-treated & qNSCs/BMP4-treated). Outer panel: histogram of p-values from the correlation test. Inner panel: Example of a gene (Cystatin C, Cst3) showing similar expression patterns between in vitro & in vivo NSC counterparts. Each dot represents single cells, and the red line shows the expression trend over the lineage. C, Comparison of similarities in gene expression between in vivo (left) & in vitro (right) cell types based of Euclidean distances (ED) using the sleepwalk tool. Each dot represents a single cell in the UMAP, and the colors indicates the ED between the highlighted cell (in vivo qNSC; cursor position) to every other cell in UMAP. Black/Green (smaller ED) Orange/Grey (larger ED). D: Correlation of log-fold changes (LFC) of 3'UTR usage between in vitro (y-axis) and in vivo (x-axis), each dot represents a single PAS (averaged over all cells). Positive LFCs imply an increased PAS usage in active (aNSCs & EGF treated), while negative LFCs imply increased PAS usage in quiescent (qNSCs & BMP4 treated), E, Gene set enrichment analysis (GSEA) for Gene Ontology. Genes were ranked, by their log likelihood ration (LR) statistics that tested differential 3'UTR usage between in vivo & in vitro*

*NSCs. Enriched gene categories are represented by the normalized enrichment score (NES). A high NES represent highly differential gene categories between in vivo & in vitro NSCs. Figure B, C & D modified and taken from* (Göpferich, 2020)

Next, I performed the sequencing *in vitro* NSCs (scRNAseq; 10x Chromium) in order to perform a more comprehensive analysis of their transcriptome compared to that of their *in vivo* counterparts. A correlation test of gene expression between *in vivo* & *in vitro* NSCs revealed over 3000 that significantly correlated (Figure 3.6B; outer panel). An example gene Cst3 shows similar gene expression along the pseudotime of both *in vitro* & *in vivo* NSCs (Figure 3.6B; inner panel). This similarity is better visualized using the sleepwalk tool that compares Euclidean distances (ED) between cells on a UMAP. Quiescent *in vivo* & *in vitro* NSCs have closer ED compared to their active counterparts (Figure 3.6C). Further, an evaluation was performed on how comparable the two systems are in their 3'UTR usage. The correlation between both data sets was 0.5 Pearson corr. which suggested that a substantial number of APA changes observed *in vivo* were also reproduced *in vitro* (Figure 3.6D). A GSEA revealed that the APA changes that did not correlate between the two were mostly associated with gene categories such as response external stimuli, cell surface etc (Figure 3.6E). Overall, the cells of the *in vitro* system do appear to mimic their *in vivo* counterparts quite reasonably and substantiates its use for sample material sensitive experiments.

### 3.1.3. Shorter 3'UTRs have a higher translation index

To study the impact of APA on translation I generated the proteomes for *in vitro* NSCs (BMP4 vs EGF treated) quantified by mass spectrometry (Figure 3.7A). A total of 3 replicates for each treatment type were used and they showed a high correlation of their proteomes with each other (Figure 3.7B). I observed that 557 proteins were upregulated in BMP4-treated NSCs while 417 proteins were downregulated in the same compared to EGF treated NSCs (Figure. 3.7C). A comparison of the protein expression to gene expression of *in vitro* NSCs revealed a substantial correlation (0.7, Pearson corr.) between the two suggesting the most proteins mirrored their transcriptome (Figure 3.8A). However, for some of the genes this was not the case.

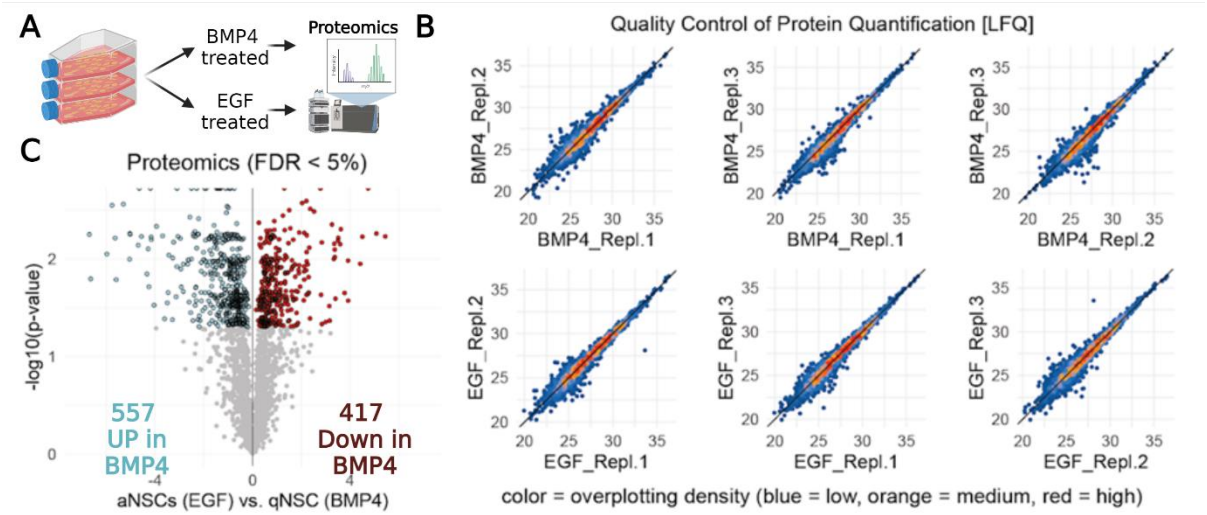


**Figure 3.7:** ***Analysis of the proteomics data from in vitro NSCs*** *A: Schematic overview to generate the proteome (mass spectrometry) of in vitro cultured quiescent (BMP4 treated) and active (EGF treated) NSCs (Created with BioRender.com). B: Quality controls plots for samples after mass spectrometry comparing replicate variability. Each dot represents LFQ values of one protein between the replicates plotted against each other. Deviation from midline represents variability between the replicates. Upper-panel: BMP4-treated NSCs (quiescent). Lower-panel: EGF-treated NSCs (active). C, A volcano plot showing differentially expressed proteins between in vitro BMP4-treated NSCs (quiescent) & EGF-treated NSCs (active). Significant proteins estimated using two-sample T-test with permutation-based FDR < 0.05 using Perseus tool. Proteins in red (sig. upregulated in BMP4-treated) and blue (sig. downregulated in BMP4-treated). Figure B&C modified and taken from* (Göpferich, 2020)

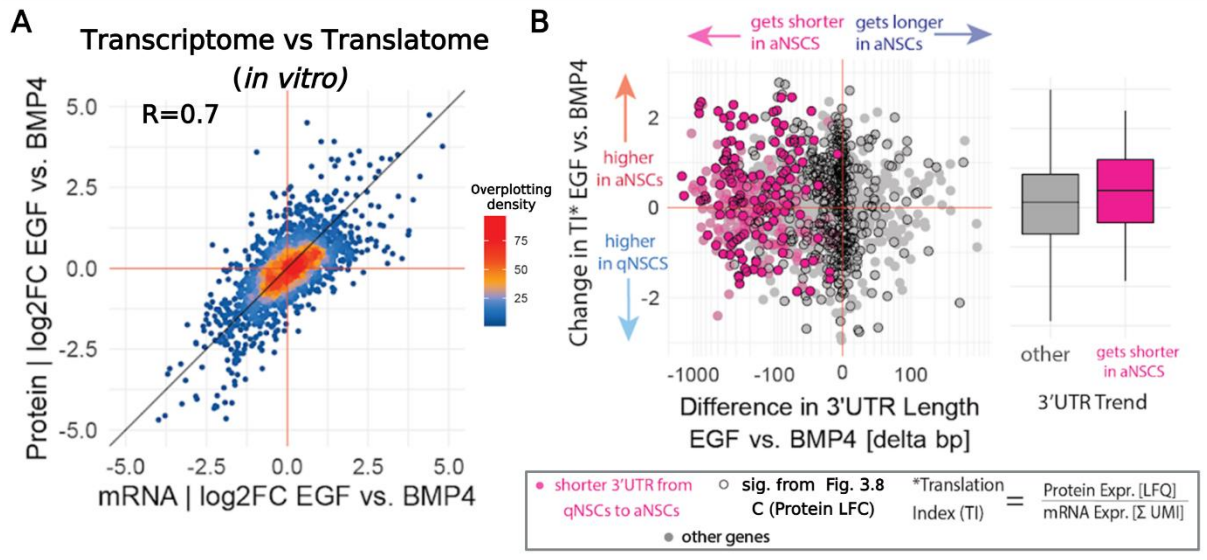


**Figure 3.8:** ***Effect of APA on translation*** *A: Comparing transcriptome to translatome of in vitro NSCs (EGF vs BMP4 treated). Log fold changes (LFC) in mRNA expression (x-axis) plotted against LFC in protein expression for each gene between EGF (active) & BMP4 (quiescent) treated NSCs. Positive LFC values indicate higher expression in EGF treated cells while negative LFC values imply higher expression in BMP4 treated NSCs. B, Plot showing the impact of 3'UTR length on translation efficiency. Left panel: plot showing changes in translation index (TI; protein/mRNA) (y-axis) computed per gene plotted against differences in 3'UTR length between EGF & BMP4 treated NSCs (x-axis). Changes in the y-axis imply an*

*increased TI in active NSCs (EGF) if positive and in quiescent (BMP4) if negative. Changes on the x-axis imply shorter 3'UTRs in active NSCs (EGF) if negative and in quiescent (BMP4) if positive. Right panel: Comparison of TI between genes with shorter 3'UTRs in aNSCs (EGF) to all other genes. Figure A&B modified and taken from* (Göpferich, 2020)

To investigate such imbalances between the proteome and transcriptome a translational index was computed by dividing protein expression with transcript levels (TI = Protein(LFQ) / mRNA(UMI). A tendency of higher protein output was correlative of shorter 3'UTRs in aNSCs (EGF treated). Owing to their restricted length, shorter 3'UTRs may not always incorporate some of the regulatory elements (repressive) that longer counterparts could (A. R. Gruber et al., 2014; Mayr & Bartel, 2009; Sandberg et al., 2008). It is therefore likely that during the transition from qNSCs to aNSCs shorter 3'UTRs could potentially restrict or even fine-tune protein output for essential genes.

### 3.2.Role of CPE/CPEB4 in regulating APA

Knowing that APA has an impact on translation the likely motifs through which APA regulation could take place were investigated. A *de novo* motif analysis of 3'UTR regions with APA identified cytoplasmic polyadenylation element (CPE) like motifs in both *in vitro* and *in vivo* NSCs (Figure 3.9A) (Charlesworth et al., 2004; Ivshina et al., 2014; Piqué et al., 2008).

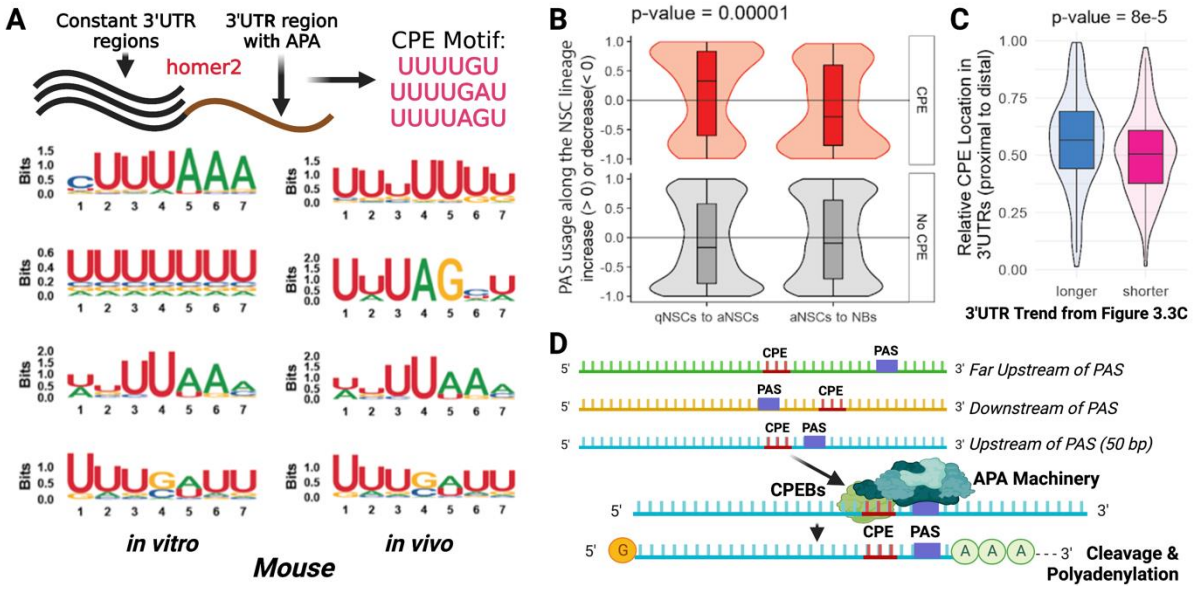


**Figure 3.9:** ***CPE mediated regulation APA in the NSC lineage*** *A: De-novo motif analysis (homer2) to identify enriched motifs in 3'UTR regions with APA over regions without. Data shown APA variable regions in mice (in vitro & in vivo). B: Comparative PAS usage with (red) or without (grey) an upstream CPE motif upon activation (qNSC to aNSCs; left) and upon differentiation (aNSCs to NBs; right). Pearson's correlation coefficients estimated by*

*multinomial spline regression that indicate PAS usage (<0 decrease, >0 increase). Interaction effects, C: Relative CPE location ANOVA. Figure A & B modified from* (Göpferich, 2020).

CPE motifs upstream of PAS (around 50 bp) appear to regulate its selection. Such selection is observed to increase during activation (qNSCs to aNSCs) and decrease upon differentiation (aNSCs to NBs) (Figure 3.9B). More specifically, in 3'UTRs getting longer along the lineage the CPE-motif was more represented in the proximity of distal PAS (Figure 3.9C). These results suggests that CPE may influence PAS selection. Further, CPE containing transcripts in neurons have been shown to localize in dendrites (Fernandez-Moya et al., 2014; Groisman et al., 2006). CPE binding proteins (CPEBs) have been shown to regulate dendritic mRNA transport and well as local translation of transcripts containing CPE motifs (Huang et al., 2003; Richter, 2007; Weill et al., 2012). Both CPEB1 & CPEB4 have been shown to regulate the length of Poly(A) tails which ultimately impacts final protein abundance (Bava et al., 2013). Alternatively, CPEB1 has also been shown to play a role in APA in the nucleus by recruiting the polyadenylation machinery to weaker PAS of CPE containing transcripts (Bava et al., 2013). In both *in vitro* & *in vivo* NSCs, CPEB4 showed higher mRNA expression compared to all other CPEBs with the highest expression in qNSCs (Figure 3.10A).

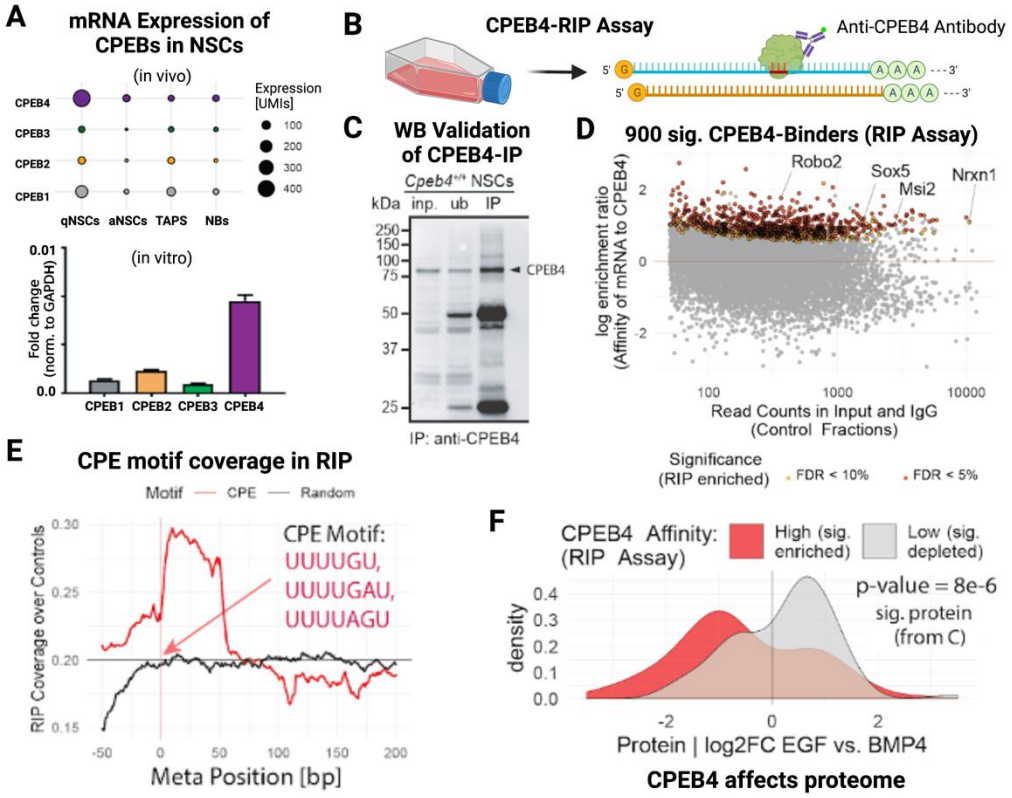


**Figure 3.10:** ***CPEB4-CPE mediated regulation APA in the NSC lineage.*** *A: Expression (mRNA) of various cytoplasmic element binding (CPEB) proteins in both in vivo & in vitro active NSCs. Gene expression (UMIs) was used for in vivo NSCs while qPCR validation was performed for in vitro NSCs (Fold change normalized to GAPDH). B: Schematic*

*representation of CPEB4-RNA immunoprecipitation (CPEB4-RIP) assay. Active NSC cultures lysates were pulled down for RNA using a CPEB4 antibody (controls: input, IgG & CPEB4 -/- NSCs) & sequenced. C: Validation of the CPEB4 RIP assay using western blot shows enriched CPEB4 in the IP fraction compared to input (controls described in B). D: Enrichment of the CPEB4 binding transcripts in the IP fraction over input & IgG fractions using DEseq2. Significant genes are coloured yellow (FDR < 10%), red (FDR < 5%). n=2 (per fraction). E: CPE motif coverage from the CPEB4 RIP assay. The RIP read coverage is show around the meta position of the CPE motif F: Comparison of protein expression of CPEB4 binders (red) to non-binders (grey) in the in vitro proteome (EGF vs. BMP4 treated). A two-sided wilcoxson rank sum test was used to determine significance. Figure A (upper), D, E & F modified and taken from* (Göpferich, 2020)

The role of CPEB4 in APA has not been explored until now. The binding of CPEB4 to CPE-like motifs in NSCs was investigated using an RNA Immunoprecipitation (RIP) against CPEB4 in *in vitro* NSCs (Figure 3.10B-C). Over 900 CPEB4-bound transcripts were observed that were enriched in the IP fraction over input (with IgG & CPEB4-/- as IP controls) (Figure 3.10D). The CPE-motifs detected in the CPEB4-bound transcripts were highly comparable to those identified by the *de novo* motif analysis. The increased RIP coverage density over controls for the CPE-motif compared to random motifs from the RIP data reaffirms the fact that CPEB4 mediated binding of the transcripts happens through CPE (Figure 3.10E). In quiescent NSCs (higher CPEB4 expression) an increased protein abundance of CPEB4 binders compared to non-binders was observed (Figure 3.10F). These results suggest the role of CPEB4/CPE in regulating protein output and/or APA.

### 3.3.APLP1 a potential regulator of the CPE/CPEB4 APA

While the regulation of protein abundance (translation) and mRNA transport (local translation) are both cytoplasmic processes, APA traditionally happens in the nucleus. This suggests a dual role of CPEB4 that requires a tight regulation of it functions in different the subcellular regions. From the bioinformatics analysis looking at potential gene candidates for APA regulation, I identified amyloid beta precursor like protein 1 (APLP1), a synaptic adhesion molecule to be an interesting candidate that could potentially regulate such balance (Figure 3.11A).When I looked at the expression of APLP1 compared to other APA regulators, I observed that it was higher in the quiescent compartment and therefore, I hypothesized that APLP1 could play a role in maintaining quiescence. Additionally, CPEB4 followed similar expression and may work in tandem to regulate APA.

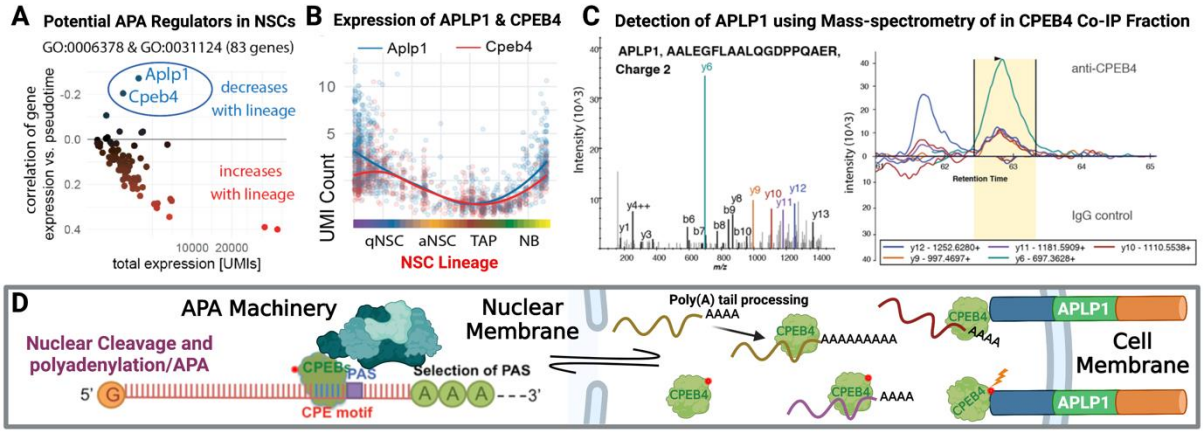


**Figure 3.11:** ***The APLP1-CPEB4 axis in regulating APA***. *A: Correlation of gene expression of potential APA regulators in the NSC lineage along the pseudotime. (GO categories: GO:0006378 and GO:0031124), based on the transcriptome data from figure 3.3. The x-axis shows the total gene expression summed over all cells (x-axis) and the y-axis shows correlation of expression vs. pseudotime. B: Correlation of expression for a subset of genes associated with APA or poly(A) tail length (GO categories: GO:0006378 and GO:0031124) along the NSC lineage to identify potential co-regulators or upstream regulators of APA. Positive correlation on the y-axis implies increased expression of said potential regulator along the NSC lineage and vice versa. Figure A & B modified and taken from* (Göpferich, 2020). *C: The identification of APLP1* protein (peptide AALEGFLAALQGDPPQAER) confirmed *via mass spectrometry in the anti-CPEB4 samples (right-up). APLP1 protein is not found in the control sample (right-down). D: Schematic representation of the hypothetical regulation of APA by the APLP1-CPEB4 axis.*

In neurons, APLP1 interacts with CPEB1 to enhance CPE-mediated mRNA translation (Cao et al., 2005). Intriguingly, APLP1 and CPEB4 have similar mRNA expression trajectories along the NSC lineage (Figure 3.11B). To investigate if the APLP1 & CPEB4 physically interacted a Co-IP of CPEB4 and its interacting proteins was performed by our collaborators (referenced, Page 43) in BMP4-treated NSCs and resulting IP fractions were analysed using mass spectrometry. Interestingly, APLP1 peptides were significantly enriched in the CPEB4-IP fractions compared to controls (Figure 3.11 C). Based on these results, I along with Dr. Goepferich hypothesized that the interaction of APLP1 and CPEB4 could potentially regulate the switch between CPEB4-mediated poly(A) tail processing & APA (Figure 3.11D).

### 3.3.1. APLP1-/- mice show an altered APA phenotype

To test this hypothesis, I wanted to look at the APA landscape upon the knockout of APLP1. In order to do this, I performed scRNAseq from the SVZ NSC lineage isolated from APLP1-/- mice and their respective WT controls (Figure 3.12A). I additionally modified the sequencing strategy that allowed for the use of hashtag antibodies to preserve replicate identity for further statistics (n=3). Using earth mover distances, a comparison of APA over the NSC lineage between the genotypes identified over 900 genes showing differential APA (Figure 3.12B).

The usage of PAS for transcripts containing CPE (50bp upstream of PAS) was observed to be dysregulated in APLP1-/- mice (Figure 3.12C). In the transition from qNSCs to aNSCs the observed preference for CPE containing transcripts in WT was lost upon APLP1 knockout (Figure 3.12C, upper). However, in the transition between aNSCs to NBs the reduced preference for CPE transcripts in WT was even lower in APLP1-/- mice suggesting that APLP1 might be essential in the maintenance of such transcripts (Figure 3.12C, lower). Comparison of the transcripts with APA changes between the genotypes to previously known CPEB binding transcripts in neurons revealed them to be enriched for CPEB4 binders and not CPEB1 (Figure 3.12D). These results implicate APLP1 as regulator of APA in a CPE/CPEB4 mediated manner. Further, a GSEA on these differential APA genes revealed that these belonged to categories of genes localized in axon, dendrites and even neuronal synapses (Figure 3.12E, upper). Such changes may impact mRNA localization & local translation and thereby affect signalling in neuronal networks (Lu et al., 2021). APLP1-/- mice have been previously reported to have imbalances of excitatory/inhibitory networks in neurons which is also observed in ASD (Schilling et al., 2017). A deeper look at these genes using a disease ontology database (DO) that revealed an enrichment for neurodevelopmental disorders especially autism spectrum disorder (ASD) (Figure 3.12E, lower).

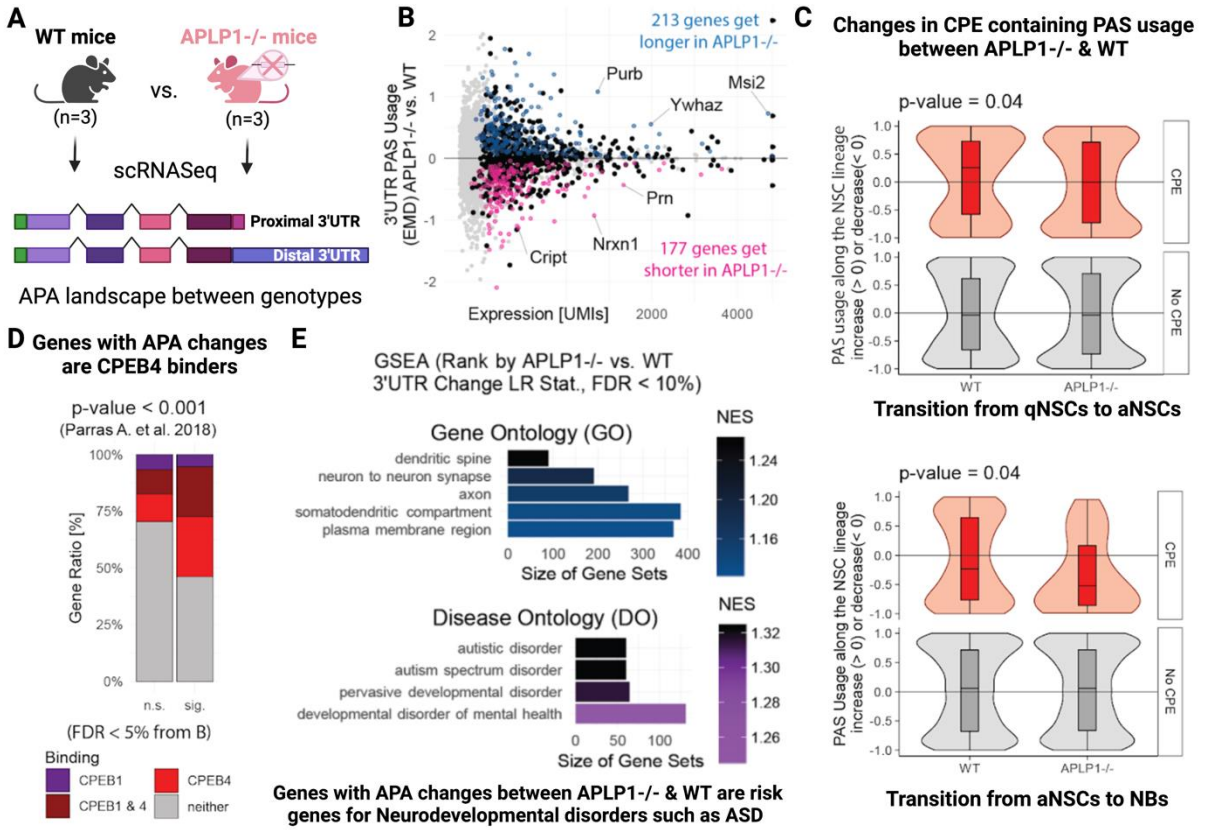

**Figure 3.12:** *APA dysregulation upon APLP1 knockout. A: Overview of scRNAseq setup to investigate APA in APLP1-/- mice compared to WT. Hashtag oligos used in the antibody staining step allowed to separate the different replicates for statistical analysis of APA changes between genotypes. B: APA changes between APLP1-/- & WT mice along the NSCs lineage using earth mover's distance (EMD; y-axis). Positive and negative EMD changes indicate overall lengthening or shortening of 3'UTRs respectively in APLP1-/- compared to controls over the lineage. This was plotted against total expression summed over single cells (x-axis). C: Comparative PAS usage with (red) or without (grey) an upstream CPE motif upon activation (qNSC to aNSCs; upper) and upon differentiation (aNSCs to NBs; lower) in WT 6 APLP-/- mice. Pearson's correlation coefficients estimated by multinomial spline regression that indicate PAS usage (<0 decrease, >0 increase). Interaction effects, ANOVA. D: Intersection of sig. genes from B with CPEB1 and CPEB4 substrates* (Parras et al., 2018b)*, colour code: gene bound by CPEB1, by CPEB4, by both or by neither, genes marked as significant (sig.) or non-significant (n.s.) from B (APLP1-/- vs. wildtype 3'UTR changes), Chi-square test.* ***E.*** *Gene set enrichment analysis: Upper panel: GSEA using Gene Ontology of genes with APA changes in APLP1-/- vs. WT mice. Lower panel: GSEA using the disease ontology database. Genes ranked by likelihood ratio (LR) statistic of APA changes. X-axis indicates the size of gene sets, and the colour indicates the normalized enrichment score (NES). Figure B, C (upper), D & E modified and taken from* (Göpferich, 2020).

### 3.3.2. Functional characterization of the APLP1-/- mice showed alterations in NSC lineage

Apart from APA changes, I was interested in investigating the functional changes that affected APLP1-/- mice. I observed changes in the composition of the SVZ NSC lineage in APLP1-/- mice compared to their WT controls. To this end, I performed a FACS analysis of the SVZ NSC lineage and I observed that APLP1-/- have lower NSC numbers compared to their WT counterparts (Figure 3.13A). Specifically from the GLAST enriched NSC lineage that I sorted for the scRNAseq, I noticed that APLP1-/- mice showed significantly reduced qNSC numbers (Figure 3.13B). Further, I observed a tendency for higher number of early NBs in APLP1-/- compared to their WT counterparts. In order to look at the number of neuroblasts that end up in the olfactory bulb, I performed a BrdU pulse chase experiment (14 days treatment + 14 days chase) in both APLP1-/- and WT mice. I observed a trend for lower number of BrdU+DcX+ late NBs in APLP1-/- compared to WT mice (Figure 3.13C&D). Based on these results, I hypothesized that the NSC pool may be depleted faster in APLP1-/- mice compared to their WT counterparts while having defects in NB migration & integration in APLP1-/- mice.

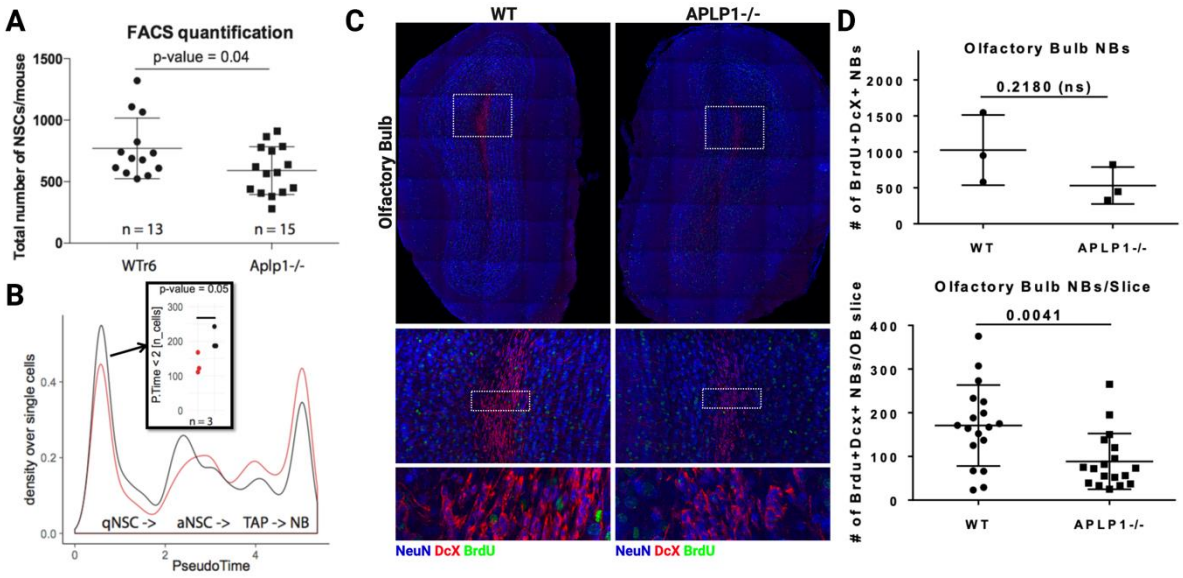


**Figure 3.13:** ***Overview of the NSC lineage cells in APLP1-/- vs WT mice.*** *A: FACS quantification of NSCs (CD45-Ter119-O4-GLAST+Prom1+) in WT & APLP1-/- mice. B: Estimation of number of qNSCs (pseudotime < 2) from scRNAseq data from WT & APLP1-/- mice (n=3). Individual mice were hash-tagged to preserve replicate identity. C: Immunofluorescence staining of the olfactory bulb of APLP1-/- and WT mice after 14-day BrdU treatment followed by a 14 day chase. Number of BrdU+DcX+ neuroblasts were counted for 6 slices per mouse (n=3). D: Quantification from C. Upper panel: Quantification of total number of BrdU+DcX+ neuroblasts in all 6 slices per mouse between APLP1-/- & Wt mice. Lower panel: Quantification of number of BrdU+DcX+ neuroblasts/slice in APLP1-/- & WT mice.*

### 3.3.3. Effect of pseudotime on APA differences between APLP1-/- vs WT mice

Owing to the changes in the composition of the NSC lineage between APLP1-/- and WT mice, a concern that needed to be addressed was, if the observed differences in APA was the result of altered cell type compositions or the genotype itself (Figure 3.14A). Dr. Macedo corrected for the pseudotime effect and still identified 718 genes with differential APA between the genotypes (Figure 3.14 B&C). This confirmed that altered APA was indeed a phenotype of APLP1 deficiency in mice. A GSEA on genes with altered APA showed that the gene categories observed before pseudotime correction remain consistently associated with neuronal function (dendritic spine, synapse etc.). GSEA using a disease ontology database showed an enrichment for disorders such as schizophrenia, cognitive disorders, and other neurodevelopmental disorders (Figure 3.14D). However, ASD was not a significant hit (FDR < 10%). An overview of the results related to CPE/CPEB4/APLP1 are summarized here below (Figure 3.15).

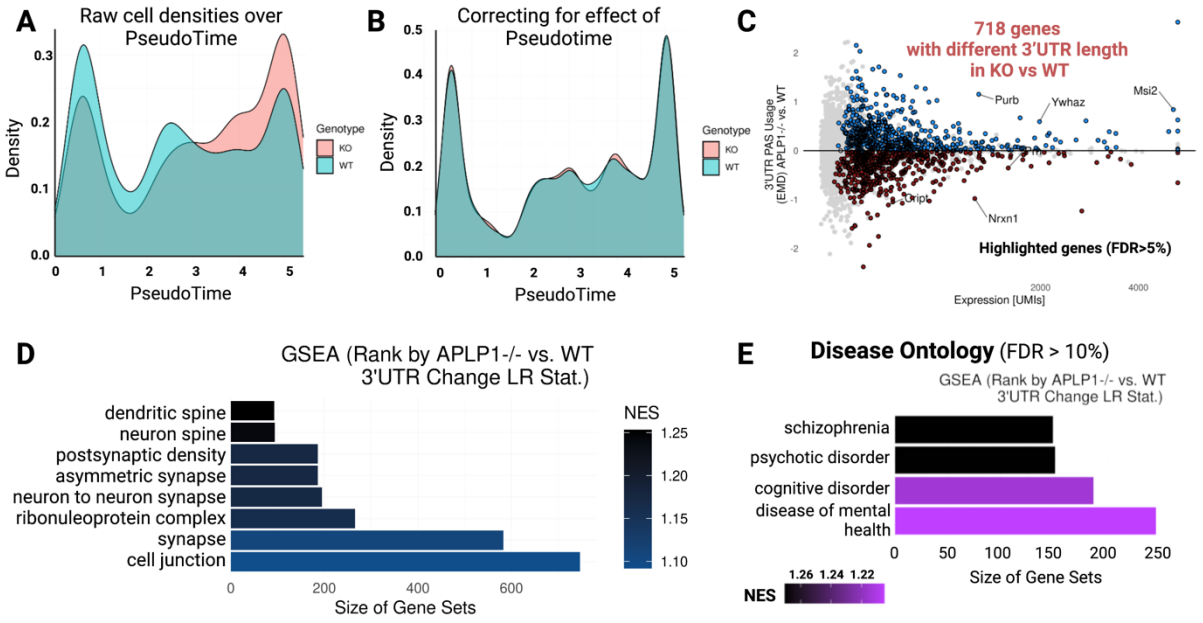


**Figure 3.14:** ***Effect skewed cell populations on APA differences between APLP1-/- vs WT mice****. A: Raw cell densities over pseudotime between APLP1-/- (pink) & WT (green). B: Normalized cell densities to remove the effect of pseudotime on downstream APA analysis. C: APA changes between APLP1-/- & WT mice along the NSCs lineage using earth mover's distance (EMD; y-axis) after correcting for the effect of pseudotime. Cells used for EMD APA analysis from B. Positive and negative EMD changes indicate overall lengthening or shortening of 3'UTRs respectively in APLP1-/- compared to controls over the lineage. This was plotted against total expression summed over single cells (x-axis). Genes with an FDR > 5% were highlighted in blue (long) brown (short) in APLP1-/- compared to WT. D: Gene set enrichment analysis using Gene Ontology of genes with APA changes in APLP1-/- vs. WT mice after correcting for pseudotime effect.* ***E.*** *GSEA using the disease ontology database of genes with APA changes in APLP1-/- vs. WT mice after correcting for pseudotime effect. For D & E genes were ranked by likelihood ratio (LR) statistic of APA changes. X-axis indicates the size of gene sets and the colour indicates the normalized enrichment score (NES). (FDR < 10%).*

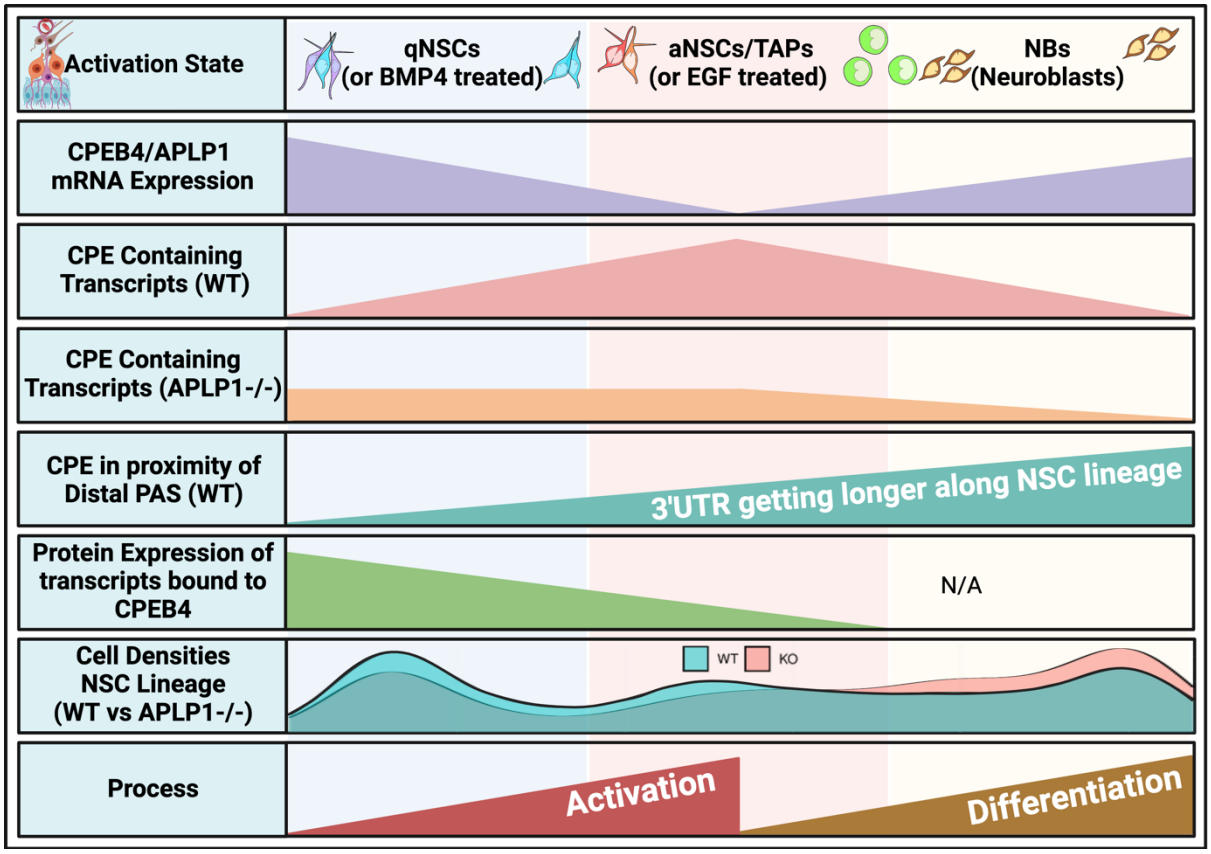


**Figure 3.15:** ***Overview of the trends over the NSC lineage for various CPE and CPEB4 results.*** *i) CPE containing transcripts ii) CPE in distal PAS in 3'UTR lengthening iii) CPEB4 mRNA expression iv) CPEB4-binders: protein expression v) Biological Process*

### 3.3.4. Exploring the molecular aspects of APLP1/CPEB4 axis

While this study identifies APLP1 and CPEB4 as potential APA regulators, it is limited in exploring the molecular mechanisms that would entail such a regulation. As presented in the results using mass spectrometry (Figure 3.11B) APLP1 was successfully pulled down in the CPEB4-IP fraction using an in house CPEB4 antibody. This suggested that the two proteins likely interacted physically to regulate APA. However, the molecular mechanism behind such an interaction is not known. APLP1 has been previously shown to interact with CPEB1 via its intracellular domain and promote polyadenylation and translation at the membrane via an aurora kinase A phosphorylation of CPEB1 (Cao et al., 2005). While CPEB4 lacks the Aurora kinase A phosphorylation site, it has additionally phosphorylation site in the neuron specific microexon, the B-region, that provides a number of other phosphorylation sites. This it is likely that the interaction of APLP1 and CPEB4 may still be facilitated through such regions. A main challenge while working with both these proteins was the availability of good antibodies that were specific. In order to circumnavigate this problem, I wanted to introduce tags (against which antibodies have previously been validated) into both proteins, APLP1 and CPEB4.

Therefore, under my supervision, ALFA-Tag CRISPR knock in clones for both APLP1 and CPEB4 was generated (Figure 3.16) (Aßfalg, 2021).

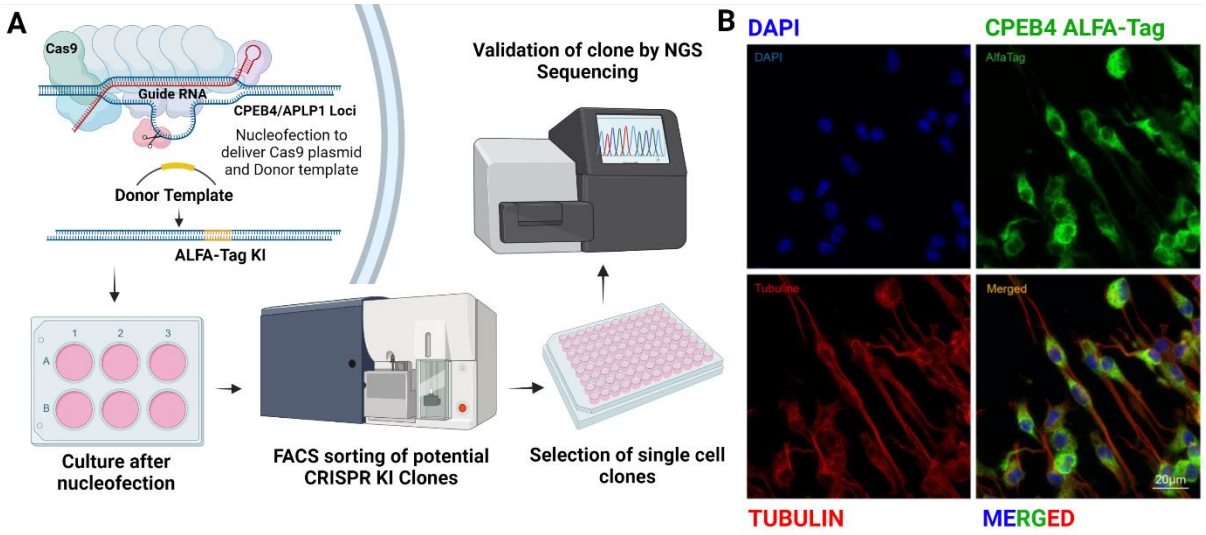


**Figure 3.16:** ***Generation of CRISPR/Cas9 based knockin of ALFA-Tag in CPEB4*** *A: Overview of the CRISPR KI strategy to generate CPEB4 and APLP1 ALFA-Tag clones. B: Validation of ALFA-Tag using immunofluorescence in a CPEB4 KI clone. A tertiary staining (Tyramide Staining: 10 mins) was performed to amplify the signal of ALFATagged CPEB4 in quiescent NSCs.*

A total of 3 homozygous and 3 heterozygous ALFA Tag knock ins for CPEB4 were generated while only a single APLP1 heterozygous KI clone was generated for APLP1. Chromatin accessibility in the vicinity of the editing site for both APLP1 and CPEB4 was quite poor, and this may be the reason for the low CRISPR efficiency in generating the KI clones. In generating the CRISPR KIs a number of optimization measures were taken into account and has been described here (Aßfalg, 2021). An additional challenge that was faced was the generally low expression of CPEB4 and APLP1 in NSCs. Despite the higher expression of CPEB4 in quiescent NSCs I needed to amplify the signal using a tertiary staining method (Tyramide staining) to visualize the tagged protein. In addition to this I plan to explore the interaction partners of CPEB4 by analysing the IP fraction of CPEB4-ALFA-Tag NSCs using mass spectrometry. I have generated the IP fractions for the same and the mass spectrometry results are awaited.

Further, having the ALFA Tagged clones for both CPEB4 and APLP1 I was interested in investigating the consequence of mis-localization of these proteins. I chose to take advantage of the LAMA system (Farrants et al., 2020). In brief, this method uses a nanobody specific to our tag of interest (ALFA Tag) coupled to a bacterial reductase complex. This reductase complex is susceptible to inhibitors like trimethoprim (TMP) that alters the binding of the nanobody to the tag. In addition, by modifying the localization signal on the LAMA one can

mis-localize the protein of interest (Figure 3.17). I have currently generated a lenti-virally transduced CPEB4-ALFA-Tag NSC line that expresses the nanobody for ALFA-Tag with an NTOM20 mitochondrial localization signal and a GFP tag for visualization. Upon LAMA expression the nanobody binds to the ALFA-Tagged protein and is sequestered in the mitochondria. This should potentially disrupt any APLP1-CPEB4 interaction at any other sub-cellular localization. I plan to explore the result of CPEB4 mis-localization and TMP rescue using the LAMA system. I currently have generated 3'UTR bulk sequencing for the CPEB4 ALFA-Tagged (LAMA/TMP rescue). These results are currently being processed but are beyond the scope of this thesis

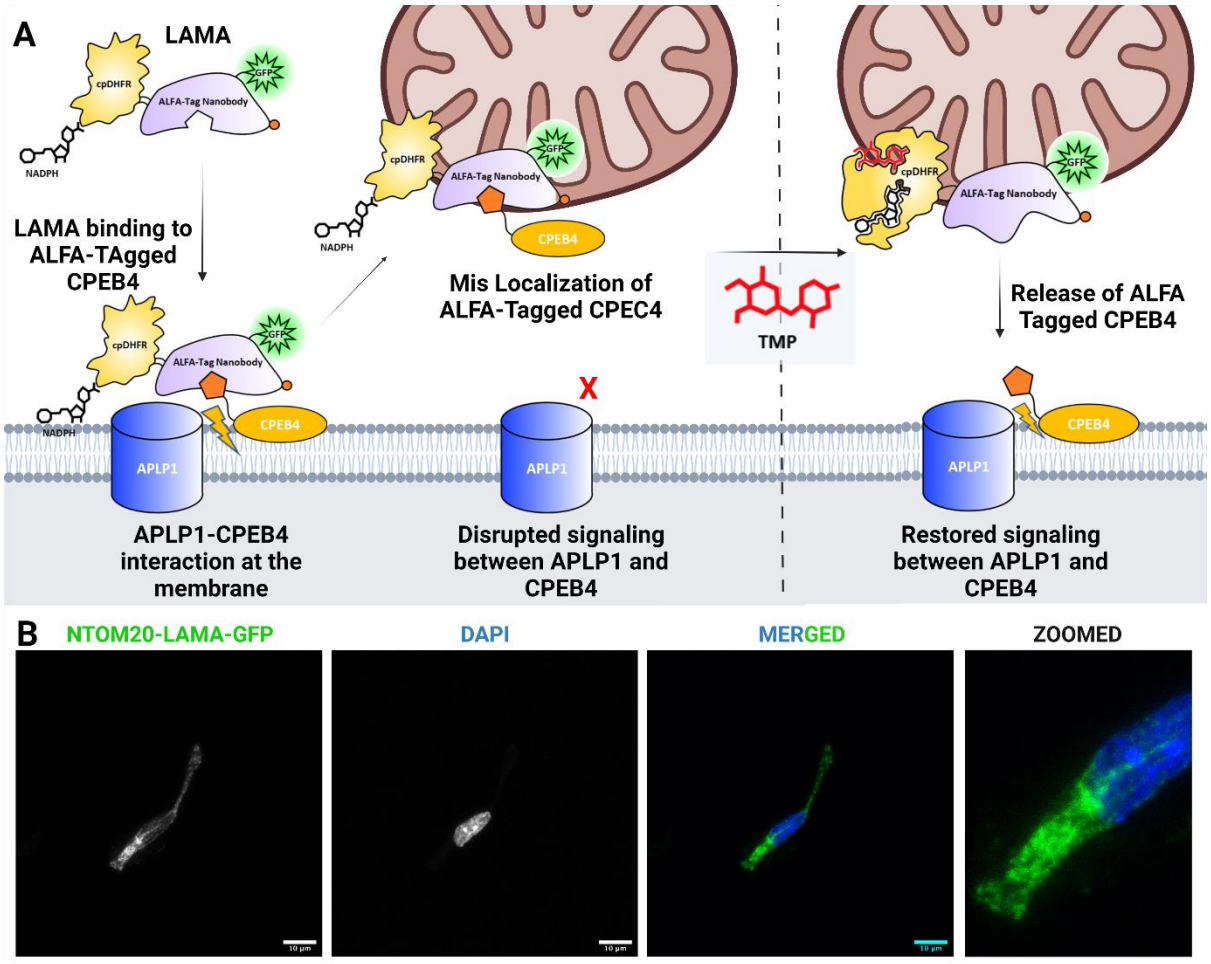


**Figure 3.17:** ***Overview of the LAMA system for mislocalization of CPEB4-ALFA-Tag*** *A. Schematic overview of the method. B: Immunofluorescence images of NTOM20-LAMA-GFP expressing CPEB4 ALFA-Tag NSCs. The localization of the GFP signal is characteristic of mitochondrial sequestering.*

### 3.4.Characterization of the behavioural phenotype of APLP1 -/- mice

Based on the GSEA results, the genes with APA changes upon APLP1 knockout were shown to be ASD-risk genes. Therefore, I hypothesized that the knockout of APLP1 may produce an ASD-like phenotype. To this end, I performed a battery of behavioural tests which are considered gold standard such as USVs, Social Interaction test & the 3-chamber test. I additionally performed tests for anxiety such as the elevated plus maze (EPM) and the open field (OF) test.

#### 3.4.1. APLP1-/- mice show altered USV patterns compared to WT mice

Young pups communicate using USVs during their early postnatal phase. To assess deficits in communication, I performed an Ultrasonic vocalization (USV) test in pups at two age points during their early stages of postnatal development (P5 and P8) (Figure 3.18A). This served as a proxy of communicative deficits observed in young children with ASD. After the USVs were recorded, I analysed them using DeepSqueak to identify real calls from background noise (Figure 3.16B). I then estimated the total frequency of USV calls per mouse in WT and APLP1-/- pups which provided a quantitative measure of communication deficits.

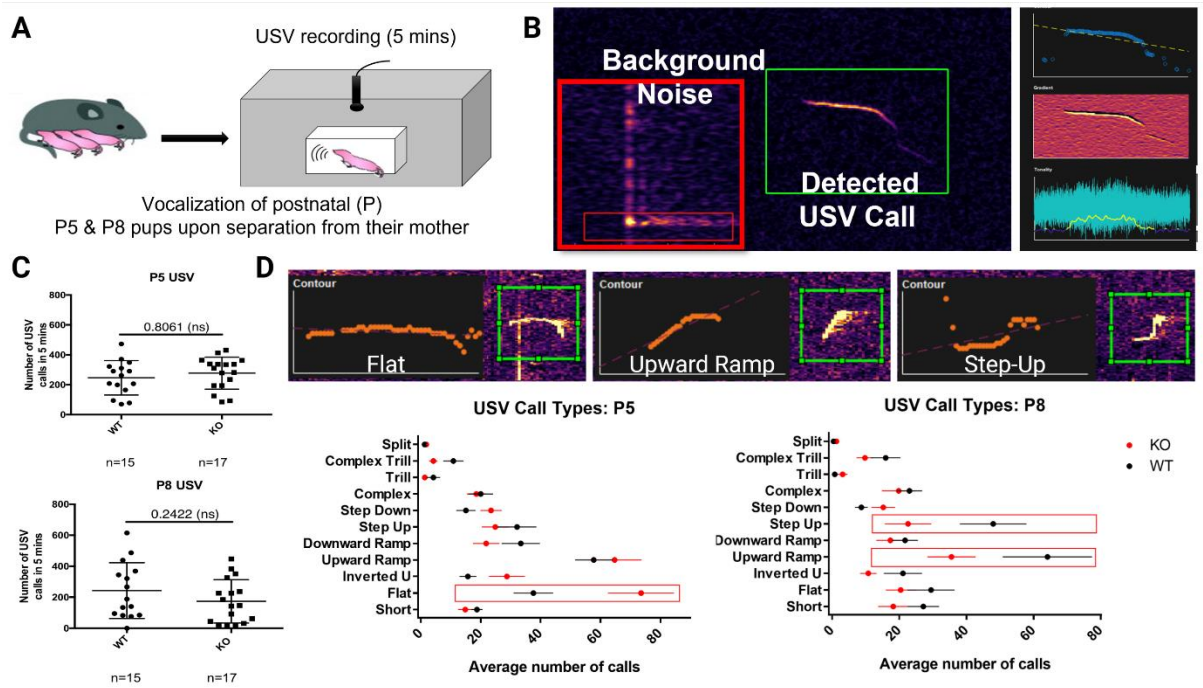


**Figure 3.18:** ***Ultrasonic vocalization experiment in APLP1-/- vs WT mice*** *A: Overview of the experimental setup ultrasonic vocalization (USV) test. B: Analysis of USVs recordings to detect and screen calls using DeepSqueak. Green (accepted calls), Red (Rejected calls) C: Plots showing the number of UVSs detected at P5 (left panel) & P8 (right panel) in both APLP1-/- (KO) & WT mice. D: USV call classification. upper: Examples of a few significant call types detected. lower: Distribution of different call types from the screened USVs (in c) between*

*APLP1-/- (KO) & WT mice at P5 (bottom left) and P8 (bottom right). APLP1-/- (KO) in red and WT in black. The dots represent the mean (plotted with SEM). The boxes (red) represent significant differences (p-value < 0.05)*

Both at P5 & P8, I observed no significant difference in call numbers between WT and APLP1-/- mice (Figure 3.18C). However, at P8 I observed a slight trend towards reduced call numbers in the APLP1-/- compared to WT mice (Figure 3.18C, lower). Further, alterations in use of specific USV call patterns (complexity in communication) in mice pups have been shown to mimic defects in social communication which is characteristic of ASD (Caruso et al., 2020; Scattoni et al., 2011). When I analyzed USV call types, I found that at P5, APLP1-/- mice had an increased number of flat (simple) calls compared to their wildtype counterparts (Figure 3.18D, lower left). Further, I noticed that the number of Upward-Ramp & Step-Up calls (relatively complex calls) showed significantly reduced numbers compared to WT at P8 (Figure 3.18D, lower right). The vocalization frequency and patterns serve as a proxy for communication and speech complexity (altered in ASD patients). Previous studies have reported a higher frequency of flat calls in a murine a model (Fmr1-KO) for fragile X syndrome which has been notably associated with ASD (Roy et al., 2012). In summary, I was able to establish that APLP1-/- mice showed alterations in the communications which presented itself in the form of reduced USV complexity compared to WT mice (Figure 3.14D).

### 3.4.2. APLP1-/- shows altered sociability compared to WT mice in Social Interaction Test (SI)

Another characteristic of ASD is that affected individuals have difficulties interacting in social situations. The social interaction test serves as a proxy to identify abnormalities in interaction with littermates in adolescent mice (P21; post weaning). To do this experiment I reunited each test pup (P21) with a previously isolated littermate of same sex & genotype. I then recorded the interactions between the two mice and also tracked their movements within the confines of an observation box using Ethovision Software (Figure 3.19 A&B). I used two parameters, latency to first contact (LFC) and percentage of cumulative contact duration (%CD) as a read out for interaction. I observed no significant difference in LFC for both males and females, between APLP1-/- and WT mice (Figure 3.19C). However, I found that the average latency was much lower in APLP1-/- (24,6 s) compared to WT (55,2 s) male mice (Figure 3.19C, left). Consistent with this trend, I observed a significantly higher %CD in APLP1-/- male mice compared to their WT counterparts (Figure 3.19D, left). However, in female mice, I observed no significant difference between the genotypes (Figure 3.19D, right). Altogether, I was able

to establish that male APLP1-/- mice showed significantly altered sociability compared to their WT counterparts while this phenotype was absent or at least less pronounced in females. This is a common feature seen in ASD with a large proportion of affected males compared to females (4:1) (Kirkovski et al., 2013).

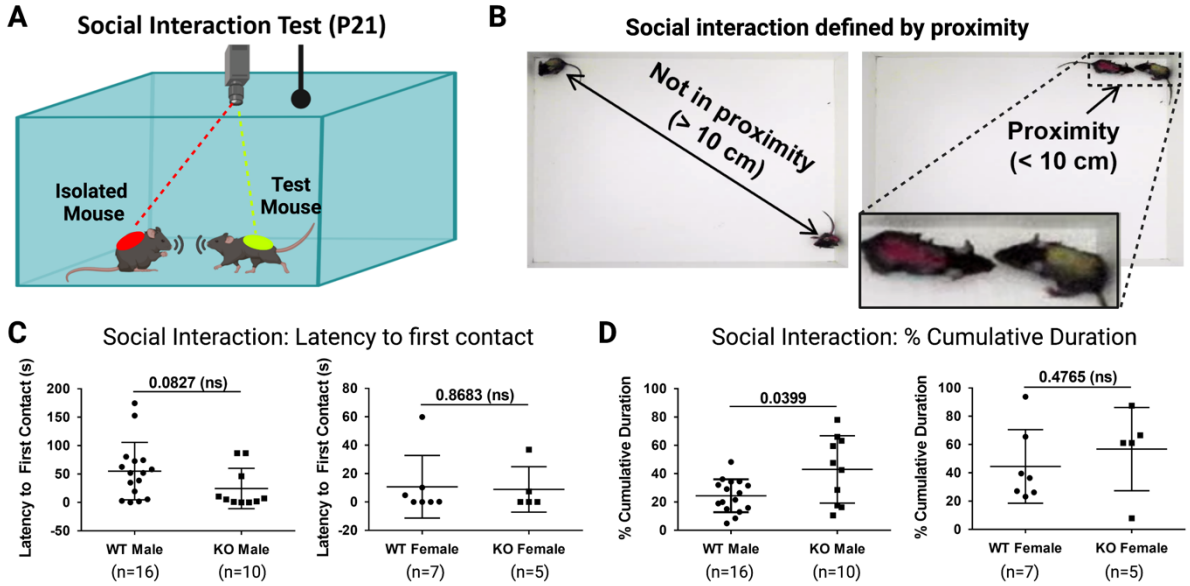


**Figure 3.19:** ***Social interaction test in APLP1-/- vs WT mice*** *A: Schematic overview of the Social Interaction Test (P21). B: Recordings of test (yellow) & isolated (red) mice showing them either in proximity (left) or not in proximity (right). Proximity was considered if distance between the centre of both tracking markers (red & yellow) was <10 cm. C: Plots showing social interaction as defined by latency to first contact (seconds) in males (left) and females (right). D: Plots showing social interaction as defined by percentage cumulative duration of proximity in males (left) and females (right).*

### 3.4.3. Male APLP1-/- mice showed decreased social novelty compared to WT mice

Further, I evaluated sociability in older mice (P40), using the three-chamber test (Figure 3.20A). I tracked the interaction time of each test mouse either in the presence or absence of a littermate and/or an intruder using SYGNIS tracker software (Figure 3.20B). Using this information, I further computed two indices namely, sociability index and novelty index, that quantified the interaction of the test mice with a littermate or an intruder respectively (Figure 3.20 C&D). I observed no significant difference in the sociability index between the two genotypes in both males and females (Figure 3.20C). I further found that these differences in sociability of APLP1-/- male mice at P21 compared to WT (Figure 3.19D, left) appeared to have diminished by P40 (Figure 3.20C; left). Additionally, I found that upon the introduction of an intruder mouse (novel social scenario), APLP1 -/- male mice showed a reduced social

novelty index compared to WT (Figure 3.20D, left). I noticed no such differences in female mice (Figure 3.20D, right). I observed that the APLP1-/- male mice were less likely to interact with an intruder as opposed to WT mice that appeared to be more socially curious. Based on these results I was able to demonstrate a dysfunction in sociability in APLP1-/- male mice, which is a core feature of ASD.

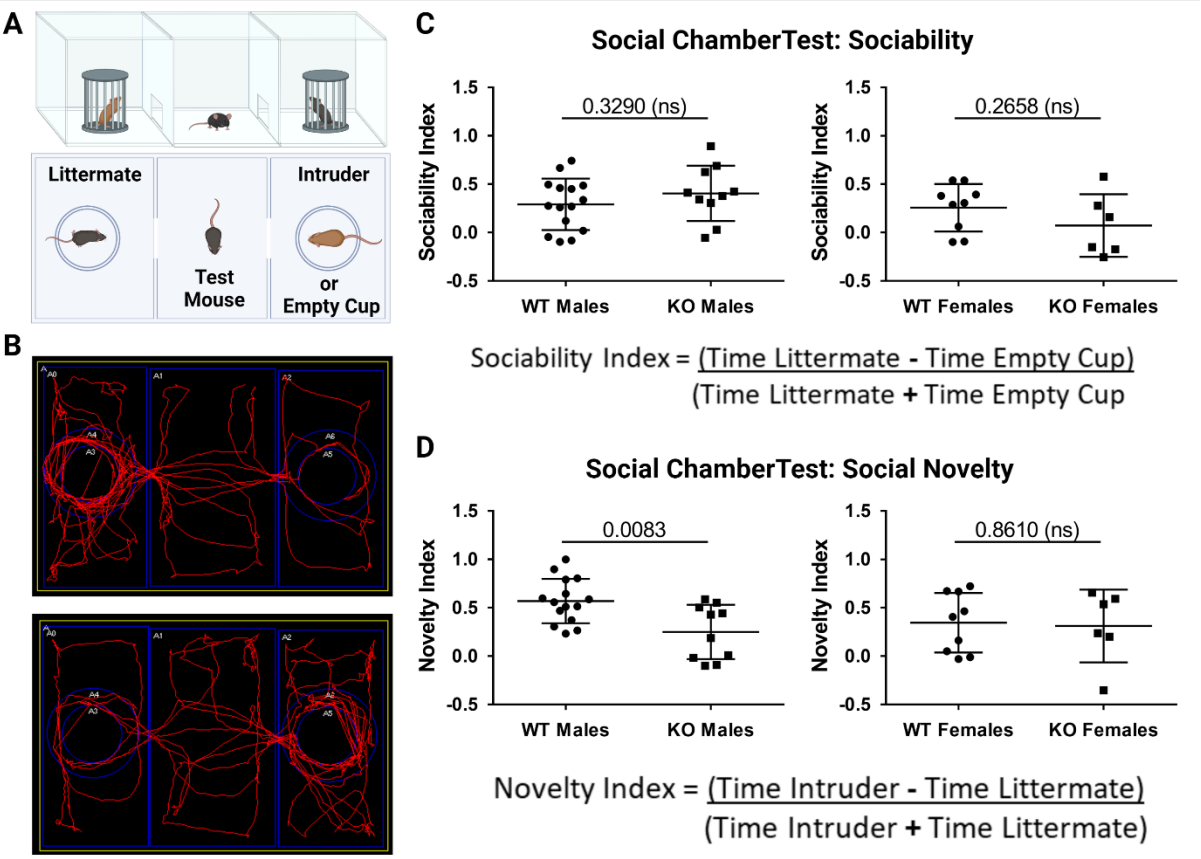


**Figure 3.20:** ***Social chamber test for sociability and social novelty*** *A: Schematic Overview of the three-chamber test for sociability (Created with BioRender.com). B: Sample tracking data (Sygnis Tracker) of mice recorded over a period of 5 mins showing preference for either the littermate (top) or the intruder (bottom). C: Plots showing sociability index calculated for Aplp1-/- & WT mice in males (left) and females (right) with the formula used to calculate the index (bottom). D: Plots showing novelty index calculated for Aplp1-/- & WT mice in males (left) and females (right) with the formula used to calculate the index (bottom).*

### 3.4.4. Aplp1-/- mice present an anxiety phenotype in males

Although, it is not a core symptom of ASD, comorbid anxiety appears in about 40% of ASD patients (van Steensel et al., 2011). This has been observed to amplify existing ASD behaviour such as deficits in social skills and repetitive behaviour (Zaboski & Storch, 2018). Anxiety in mice can be studied using tests such as elevated plus maze (EPM) and open field (OF) tests.

Since the sociability studies I previously performed at P21 and P40 showed a pronounced phenotype in male as opposed to females, I conducted anxiety tests primarily in male mice. For EPM, I computed an anxiety index by comparing time spent in the open (potentially unsafe) and closed (protected) arms of the maze (Figure 3.21A&B). Although I observed no significant differences between the two genotypes in their anxiety index, I found APLP1 -/- mice to show a tendency for increased anxiety compared to control mice (Figure 3.21C, p-value = 0.0557). In the OF test, I observed that APLP1-/- mice spent significantly less time in the centre of the open field (potentially unsafe) and more time along the edges (protected) compared to their WT counterparts (Figure 3.21 D-F)

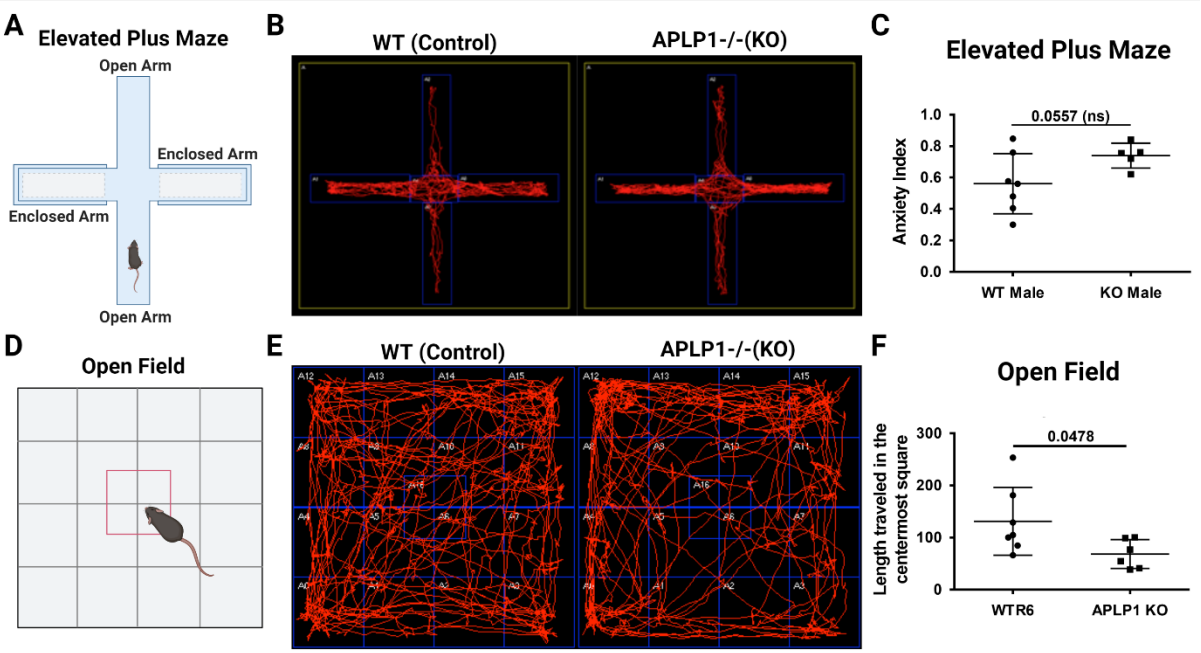


**Figure 3.21:** ***Elevated plus maze and open field test in APLP1-/- vs WT mice*** *A Schematic Overview of the Elevated Plus Maze (EPM) for anxiety (Created with BioRender.com). B: Sample tracking data (Sygnis Tracker) of mice recorded over a period of 10 mins showing movement within the enclosed & open arms for WT (left) & APLP1-/- (right). C: Plots showing EPM anxiety index estimated for Aplp1-/- & WT mice in males. D: Schematic Overview of the Open Field (OF) for anxiety (Created with BioRender.com). E: Sample tracking data (Sygnis Tracker) of mice recorded over a period of 10 mins showing movement within a square open field for WT (left) & APLP1-/- (right). F: Plots showing length travelled by test mice in the centre square which is estimated for Aplp1-/- & WT mice in males.*

In summary, I was able to establish that APLP1-/- mice showed ASD like behaviour namely; alterations in USV call patterns at P5 and P8, altered sociability at P21, reduced social novelty at P40 and anxiety-like behaviour in adult mice compared to controls. These observed changes in behaviour altering communication and sociability are common symptoms for across different neurodevelopmental disorders in humans especially in ASD (Kazdoba et al., 2016; Sungur et al., 2016, 2017).

### 3.5.Strong phenotype of 3'UTR lengthening in most genes & cell types of ASD vs. controls

Given that CPE/CPEB4/APLP1 dysregulation resulted in an altered APA molecular phenotype as well as behavioural ASD phenotype, APA dysregulation in human ASD was of interest. To this end the previously published NucSeq data by Velmeshev et al, generated from post-mortem human ASD patients and healthy controls, was re-analysed for APA changes in different neuron types of the prefrontal cortex and anterior cingulate (Figure 3.22A) (Velmeshev et al., 2019b).

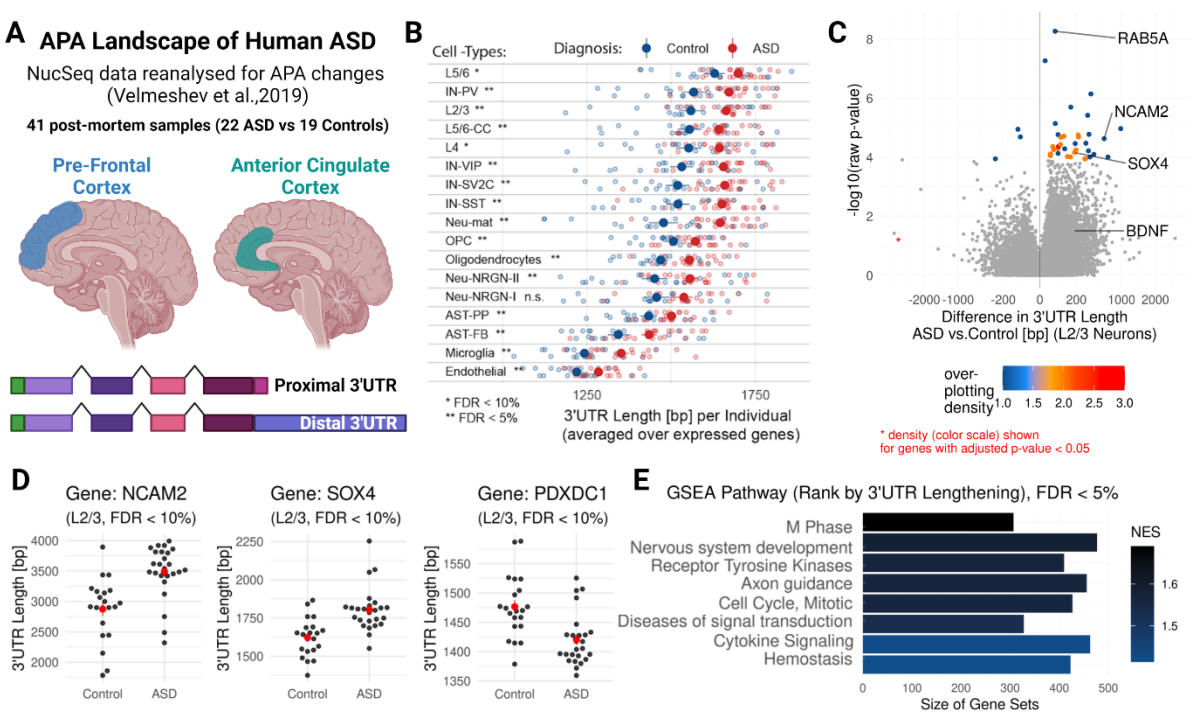


**Figure 3.22:** ***APA changes in ASD brains compared to healthy controls*** *A, Schematic representation of Velmeshev et al NucSeq data used for reanalysis to look for APA changes between ASD patients & controls* (Velmeshev et al., 2019c). *A linear model was used to test the impact of 3'UTR length on diagnosis (ASD vs Control) for each gene and within the different cell types studied, excluding the effects of sex and brain region (Created with BioRender.com).* ***B****, Plot showing the average 3'UTR length of all genes per cell type for each sample; ASD (red) Controls (blue). The darker dots represent group means for each cell type showing a clear lengthening trend for ASD compared to controls. A multiple two-sided t-tests was performed for each cell-type with Benjamini-Hochberg FDR correction.* ***C****, Volcano plot showing differences in 3'UTR lengths between ASD & Controls for each gene in layer 2/3 excitatory neurons. Positive value implies the 3'UTR gets longer in ASD while negative implies lengthening in controls. Overplotting density colour scale assigned to genes with an adjusted p-value < 0.05.* ***D****, Examples of genes from C showing differences in 3'UTR length (ASD vs Control), each point represents mean UTR length for all L2/3 neurons per sample, group means are indicated by red dots; NCAM2 (neural cell adhesion molecule 2) & SOX4 (SRY-box transcription factor 4), get longer in ASD while PDXDC1 gets longer in controls.* ***E****, A gene set enrichment analysis (GSEA), of all genes expressed that are ranked by their trend for 3'UTR lengthening in ASD vs. controls (from C). Bars coloured according to normalized*

*enrichment score (NES), higher the NES higher the association of the gene sets to the pathway. Figure B modified and taken from* (Göpferich, 2020)

A striking observation was the transcriptome-wide lengthening trend of 3'UTRs in ASD compared to controls (Figure 3.22B). This trend was consistent across samples irrespective of the cell type and brain region (Figure 3.22B). For example, a closer look at individual cell types like in L2/3 neurons showed that a there was a large proportion of the genes with the same lengthening trend (Figure 3.22C). This is demonstrated in more detail when looking at individual genes such as NCAM2 & SOX4 (Figure 3.22D, left & middle). Despite some genes such as PDXDC1 showing the opposite trend most genes illustrate a lengthening trend in ASD compared to controls (Figure 3.22D, right). A GSEA analysis revealed that the genes with the lengthening trend were enriched for pathways involved in cell cycle, axon guidance, receptor tyrosine kinase signalling & diseases of signal transduction (Figure 3.22E). Previous studies have shown that abnormal or dysfunctional connectivity in the brain & synaptic signalling or axonal pathfinding deficits are characteristic of impaired social interaction and repetitive behaviours: the primary indicators of ASD (Gilbert & Man, 2017b; Greco et al., 2013). Despite a few isolated studies showing the involvement of 3'UTRs in ASD, this is the first study that has shown a global regulation of 3'UTRs length in ASD using scRNA-Seq (Szkop et al., 2017).

## 3.6. Investigating the potential of APA changes as a marker for ASD diagnosis from Blood

Accurate and early diagnosis of ASD is essential for providing affected individuals with the necessary treatment (Hus & Segal, 2021; Myers et al., 2007). Studies have reported the association of early diagnosis with improved prognosis of ASD (reviewed in Volkmar, 2014) Especially medical and social interventions before the age of 4 years have been reported to have improved their language difficulties, cognitive skills, IQ and adaptive behaviour (Dawson et al., 2010; Devescovi et al., 2016; Vivanti & Dissanayake, 2016). Therefore, the need for a simple and reliable diagnostic tool is warranted. Owing to a strong phenotype of APA in human ASD across different cell types of the brain including endothelial cells & microglia, I wanted to investigate if this phenotype was represented across other tissues of the body such as blood. If this phenotype were tissue-independent, tissues such as blood could be used to develop molecular tests for ASD diagnosis.

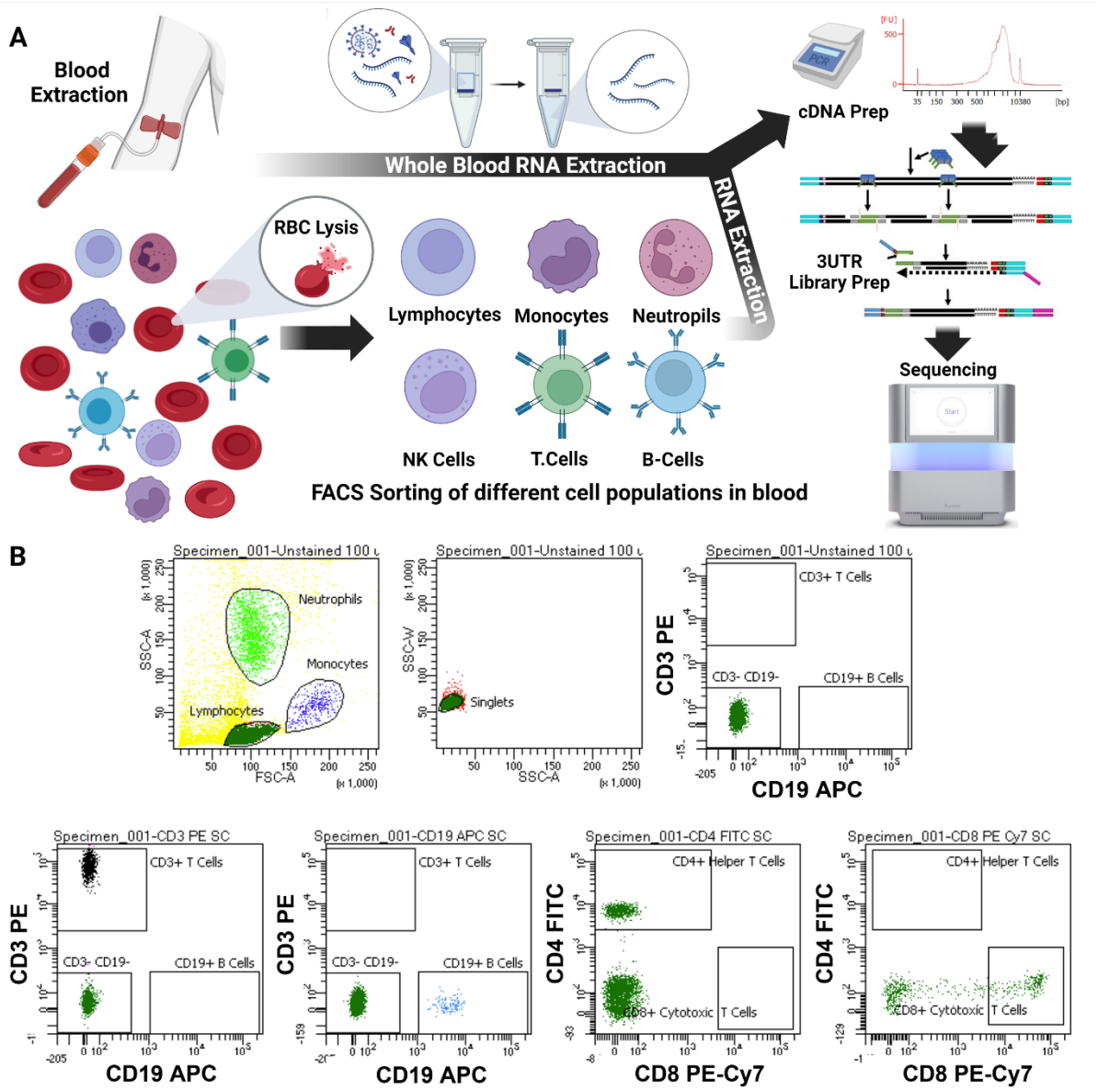


**Figure 3.23:** ***Overview: Generating & sequencing of 3UTR libraries from both whole blood and individual blood cell types to compare difference in APAome.*** *A: Schematic overview. Whole blood was extracted from healthy test subjects and either directly processed for whole blood RNA extraction or processed for FACS sorting to isolate different peripheral blood mono nuclear cells (PBMCs). B: The FACS sorting strategy used to isolate different PBMC populations. FSC vs SSC allows to distinguish lymphocytes, monocytes, and neutrophils. Antibody staining to sort $CD19^{+}CD3^{-}$ B-cells, $CD3^{+}CD4^{+}$ Helper T cells & $CD3^{+}CD8^{+}$ Cytotoxic T-cells.*

### 3.6.1. Investigating APA differences between blood cell populations

To be able to use APA as a diagnostic marker for ASD, the simplest method was to use whole blood from patients as test material. However, as different cell types or activation states have been previously shown to have altered APA, any observable differences may be obscured by such variations in whole blood (differences in cell types) (Figure 3.23A). I wanted to

investigate if such differences existed, and therefore set out to sequence the APAome of whole blood and compare it to FACS sorted individual PBMC cell types (Figure 3.23).

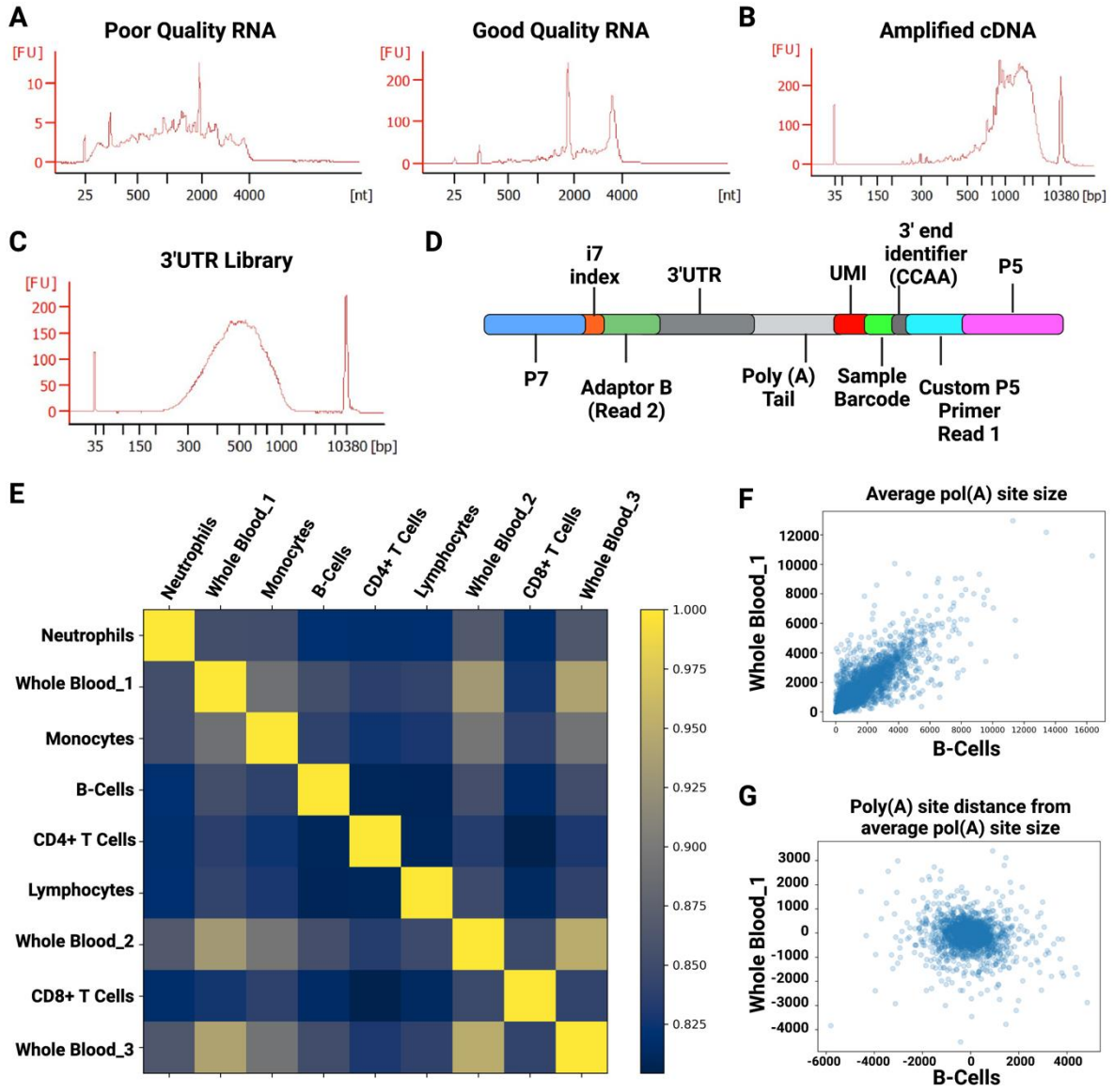


**Figure 3.24:** ***Comparison of whole blood APA to individual cell types in healthy individuals.*** *A: RNA quality assessed by bioanalyser profiles estimating RIN values for each RNA sample run showing either poor quality RNA (top) or good quality RNA (bottom; RIN >6.5) B: Bioanalyser profile of reverse transcribed cDNA after amplification. C: Bioanalyser profile of tagmented 3'UTR bulk libraries ready for sequencing. D: Schematic representation of 3'UTR libraries used for sequencing (Hennig et al., 2018). E: Heat map showing polyA site size (APA) correlations between different samples (whole blood and individual cell types). The correlation index shown in the heatmap ranges from 0.8 (Deep blue) to 1 (Yellow). F-G: Individual sample correlations (Whole Blood_1 vs B-Cells) showing F: comparison of average poly(A)site sizes G: Distance of poly(A) site size from average size.*

Firstly, I extracted RNA from FACS sorted individual PBMC populations as well as whole blood samples from healthy controls. I then selected samples with good quality RNA (RIN>6.5). Later, I processed these RNA samples further for reverse transcription (RT) and

cDNA Amplification (Figure 3.24 A-B). To prepare tagmented 3'end enriched libraries, I adapted a modified version of the tagmentation protocol previously described by Henning et al. I worked on the optimization of this protocol, to successfully generate 3'UTR bulk libraries from whole blood and individual PBMC cell types (Figure 3.24 C-D, Methods) (Hennig et al., 2018). I then sequenced these libraries to generate the APA landscape from both. After sequencing, the 3'UTRs were mapped to the genome and an analysis of the poly-A site sizes across all samples was performed. The APA landscape correlated quite well across all samples (Figure 3.24E). While the 3 replicates of the whole blood samples correlated the most with each other (> 0.9), the correlation between individual cell types and whole blood was also quite high (> 0.8) (Figure 3.24E). For example, the comparison of the average PAS size between samples Whole Blood_1 and B cells showed a good correlation (Figure 3.24F). The correlation of the distance of individual PAS size for an individual sample to the average PAS size for both samples combined showed a very symmetric distribution (Figure 3.24G). Based on these results, I hypothesized that the whole blood APAome may not be obscured by individual variations within cell types.

### 3.6.2. Investigating APA differences from whole blood in ASD vs healthy controls

As I had not observed any huge variations between individual blood cell types I then wanted to investigate APA differences between ASD patients & healthy controls in blood (Figure 3.25A). In order to do this, I processed a total of 83 samples comprising of 27 ASD children, 50 healthy parents, 5 healthy children using the previously described 3'UTR bulk protocol. Additionally, I also included for this analysis an adult male (22 yrs.) diagnosed with ASD reported to harbour an APLP1 mutation. Upon processing the blood samples, I observed a striking difference between the cDNA prepared from freshly isolated (test) & frozen patient (ASD & Controls) samples. I observed a sharp peak at around 650-700 bp and a small shoulder to the right of the peak in the cDNA BioAnalyser profiles of frozen patient samples. The sharp peak corresponded to the size of RBC transcripts (HBB & HBA1/2) previously categorized and the left shoulder appeared to be PBMC cDNA. I therefore hypothesized that inefficient RBC lysis resulted in this overrepresentation of RBC transcripts observed in the cDNA profiles (Figure 3.25B&C). Given that the RBC transcripts were within a narrow size range, I performed an SPRI bead purification step (0.5X; size select > 600 bp) both after RT & cDNA amplification steps (Figure 3.25D). Using this method, I was able to significantly deplete the RBC transcript peak in the purified cDNA. With this selective size depletion, I was able to recover and amplify PBMC cDNA to a significant extent (Figure 3.25D). I further processed

the resulting cDNA for Tagmentation to prepare 3'UTR bulk libraries (Figure 3.25E). I finally sequenced these libraries to generate the APA landscape of ASD vs control samples.

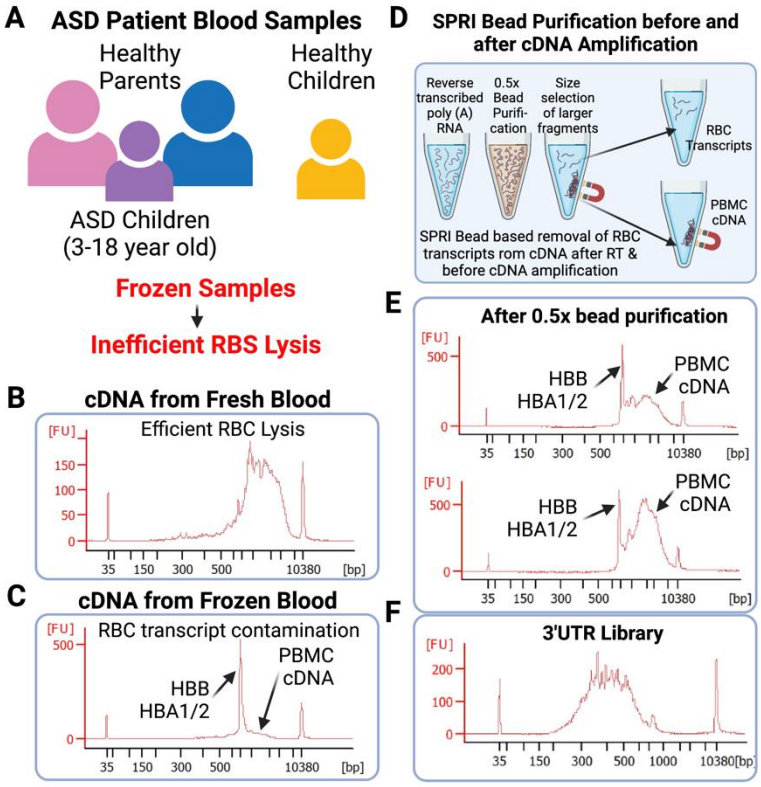


**Figure 3.25:** ***Experimental setup to study APA from blood of ASD patients and healthy controls.*** *A: Overview of samples used for analysis of APA differences in blood as a diagnostic tool for ASD. B: BioAnalyser profile of cDNA from fresh blood after efficient RBC lysis. The profile shows a broader peak between 400-10000 bp. C: BioAnalyser profile of cDNA from frozen blood of ASD/control samples showing RBC transcript contamination owing to inefficient RBC lysis. The profile shows a sharp peak around 600-700bp (RBC Transcripts) with a tiny shoulder to the right (PBMC cDNA). D: Schematic representation of SPRI bead-based size selection before and after cDNA amplification to remove contaminating RBC transcripts (HBB & HBA1/2) that are smaller in fragment size than most of the PBMC cDNA. E: Bioanalyser profiles of 2 samples after the 0.5X bead purification. The RBC transcript (HBB & HBA1/2) peak was greatly reduced and the PBMC cDNA peak (right shoulder) is more pronounced. F: Bioanalyser profile of the final 3'UTR library after Tagmentation and library amplification ready to be sequenced.*

Despite the SPRI bead-based depletion of RBC transcripts a large percentage of reads were taken up by the RBC transcripts. Despite this affecting the total number of reads mapped to 3UTR's from PBMCs, the data was sufficient for a preliminary analysis of the APA landscape in ASD vs controls. Unlike the brain, there was no observable differences in average 3'UTR lengths between ASD and healthy controls in whole blood (Figure 3.26A). However, with deeper look at individual candidate genes, 3'UTRs with altered APA trends were observed between ASD and healthy controls. Both lengthening & shortening trends were observed as opposed to the clear lengthening phenotype in the brain (Figure 3.26B). For some candidates

a clear separation was observed in their preferred PAS usage (3'UTR length) between ASD and controls with low variance among individual samples, while some others did have a huge variance between the individual samples in one or both diagnoses (Figure 3.26B). Overall, these results suggested that the data was over-dispersed, i.e., the variance observed in the data was larger than what the model anticipated. After correction for over dispersion using a Dirichlet-Multinomial regression (MNR) 4204 genes were identified with significant APA changes between ASD and controls (Figure 3.26C).

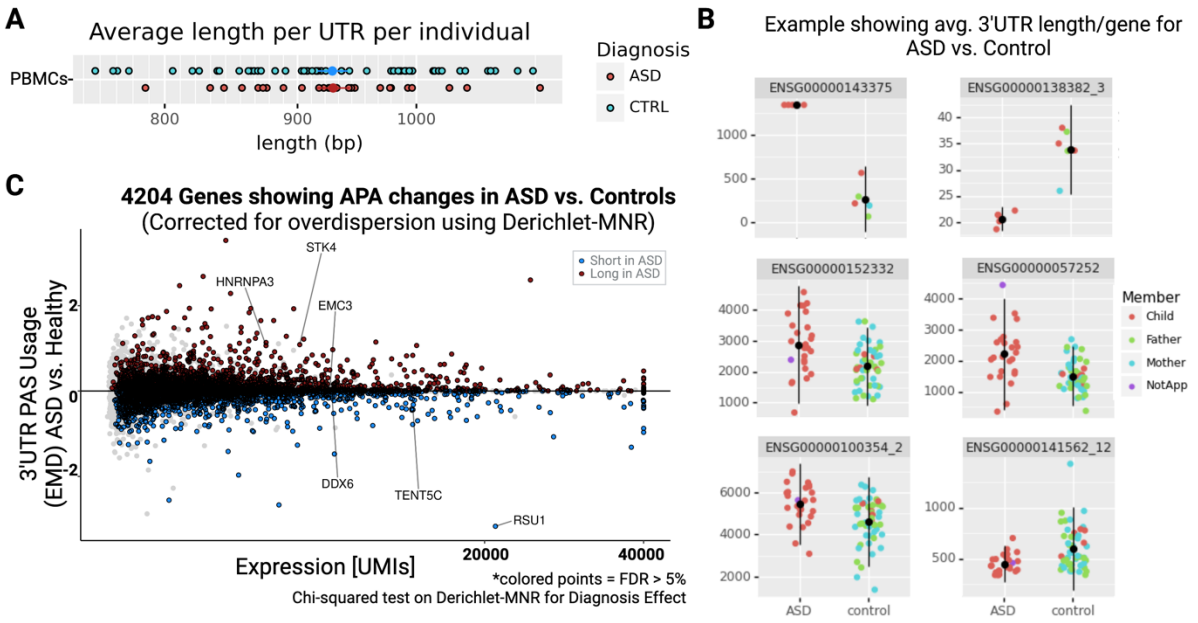


**Figure 3.26:** ***Analysis and comparison of APA from blood of ASD patients and healthy controls*** *A. Average 3'UTR lengths for all genes/sample in ASD compared to healthy controls. Bright dots in the middle are average across all samples for each diagnosis. B: Examples showing average 3'UTR lengths of all transcripts for a gene in ASD vs. Control. Colours represent different members of the family (father, mother & child). NotApp is an adult male (22yrs) diagnosed with ASD harbouring a mutation in APLP1. C: Differential 3'UTR usage between ASD & controls using earth mover's distance (EMD). Positive and negative EMDs are indicative of the trend showing 3'UTR lengthening & shortening in ASD compared to controls. A Dirichlet multinomial regression (MNR) was performed to correct for overdispersion. A Chi-squared test on the Dirichlet-MNR provided 4204 significant hist. Coloured points show significant genes with APA differences; ASD vs. Controls (FDR < 5%).*

### 3.6.3. Using APA differences in whole blood to classify ASD & controls

To build the classifier for ASD diagnosis the top 10 significant hits from the Dirichlet-MNR that fulfilled minimum criteria (Figure 3.27) were chosen as potential candidates. Upon testing different classifiers, a Gaussian Naïve Bayes classifier performed the best among others (logistic regression, SVC, & nearest neighbour) and was used for subsequent testing. A Receiver Operator Characteristic (ROC) curve for the Gaussian NB classifier had an area under the curve (AUC) of 0.69. A precision-recall curve showed an average precision (AP) of 0.78.

Other metrics for the classifier's performance were a) accuracy, a metric for correctly classified samples, which was 50% and b) F1 score, a measure of incorrectly classified samples, was 0.583 (Figure 3.27). The classifier seemed to perform well in classifying controls but require improvement overall in detecting ASD.

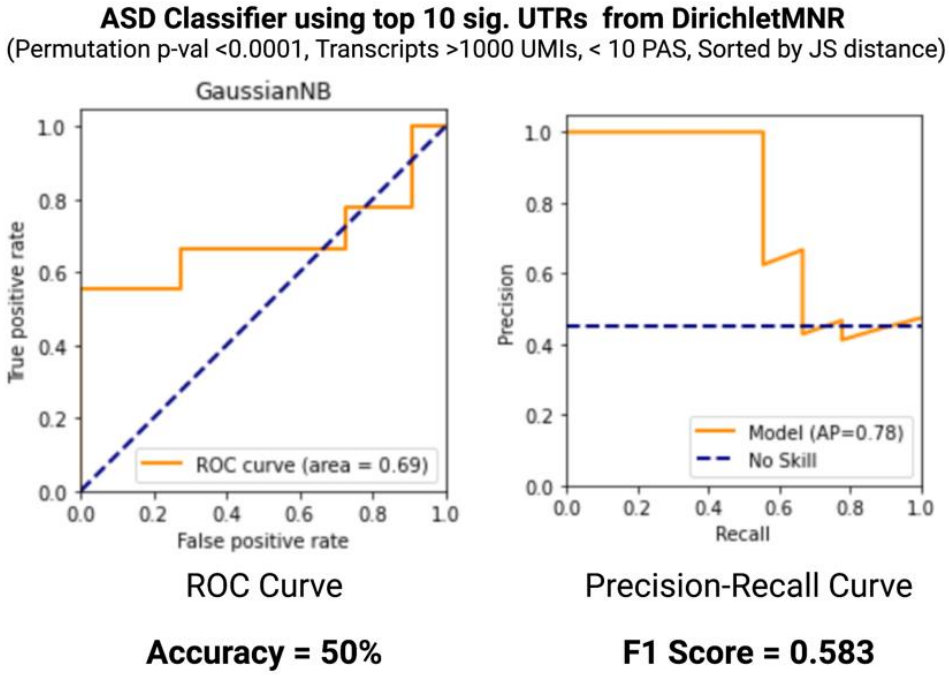


**Figure 3.27:** ***Performance of an ASD classifier in diagnosing ASD*** *built using the top 10 significant UTRs identified using the Dirichlet Multinomial. Left: ROC curve for a Gaussian Naïve Bayes classifier. Right: Precision-Recall Curve for the same classifier. All UTRs used for the classifier were required to have a Permutation p-value < 0.00001, Transcripts > 1000 UMIs, Total number of PAS per gene/3'UTR < 10, Sorted by Jenson-Shannon (JS) distance.*

### 3.6.4. Improving the 3'UTR library quality by targeted RBC depletion

Additionally, inefficient RBC lysis and the resulting contaminating transcripts seem to affect the sequencing depth. The variance in the levels of contaminating RBC transcripts may also account for the observed variance between individual samples.

To improve sample quality and increase sequencing depth, I replaced the SPRI bead-based RBC transcript depletion by an LNA primer-based depletion (Figure 3.28A). I speculated that this would provide a more uniform means of exclusive depletion of the RBC transcripts while still preserving the smaller mRNA transcripts and enriching the overall PBMC transcripts (Figure 3.28A&B). In order to optimize this protocol, I tested three different primer concentrations (2.5 µM, 1 µM & 0.25 µM). I observed significant depletion of the RBC transcripts during RT & cDNA amplification irrespective of primer concentration (Figure 3.28B). Therefore, I used 0.25 µM, the lowest working concentration, for all further experiments. Using this modification, I processed a total of 47 samples. These samples

included 33 new samples (19 ASD & 14 controls) as well as 14 old samples (5 ASD & 9 controls; Figure 3.28). I was able to compare the efficiency of both protocols, by resequencing 14 of the old samples with the new protocol. I observed an improvement in the average number of mapped 3'UTRs from around 11000 (processed with SPRI bead-based depletion methods) to around 14000 (primer-based depletion improved the average number of (Figure 3.28D). Additionally, I observed an improvement in the proportion of PBMC transcripts over my previous sequencings (Figure 3.28E). However, like the previous sequencing, I observed no significant differences in the average 3'UTR length between ASD and controls (Figure 3.28C).

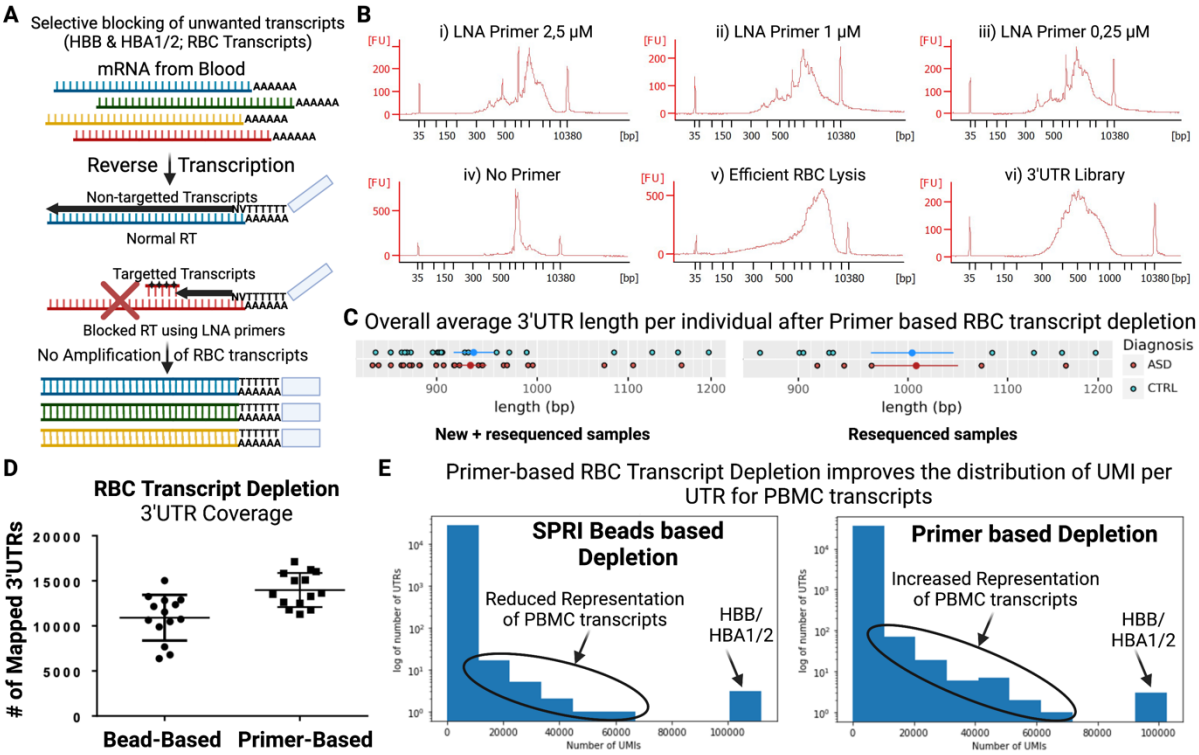


**Figure 3.28:** ***RBC transcript depletion to improve sequencing quality****. A. Overview of the method for selective depletion of RBC transcripts (HBB & HBA1/2) using LNA primers that block transcript amplification during RT and cDNA amplification. B: BioAnalyser profiles of samples depleted using this method. Primer concentrations tested for optimization i) 2,5 µM ii) 1 µM iii) 0,25 µM iv) No primer. Additionally for comparison cDNA of sample with efficient RBC lysis is shown (v). The final 3'UTR library after Tagmentation and amplification (vi). C. Average 3'UTR length of all genes per individual using the RBC transcript depletion strategy shown for ASD and control samples. Left: shown for all 47 samples sequenced (33 new + 14 re-sequenced) Right: shown for only the 14 re-sequenced samples. Bright dots in the middle represent sample means. D: Comparison of 3'UTR coverage of the 14 samples RBC transcript depletion with SPRI Beads-based (size selection) vs Primer-based methods. E: Plots showing distribution of UMI per UTR for all transcripts. Left: For samples that have not been depleted for RBC transcripts*

### 3.6.5. Testing the ASD classifier

The new samples were tested with the previously built Gaussian NB classifier. Despite an improvement in the accuracy (54%) and F1 scores (0.71) from before, the classifier performed poorly when it came to AUC (0.56) & average precision (0.59) for ROC & Precision-Recall curves respectively (Figure 3.29). However, a new classifier developed on the improved samples could help identify better candidates as well as improve the performance of the classifier. Ultimately, by adding more samples to build and test the classifier as well as accounting for diagnosis severity and comorbidities of ASD more meaningful candidates could be used to produce a more robust classifier for ASD diagnosis (Figure 3.30).

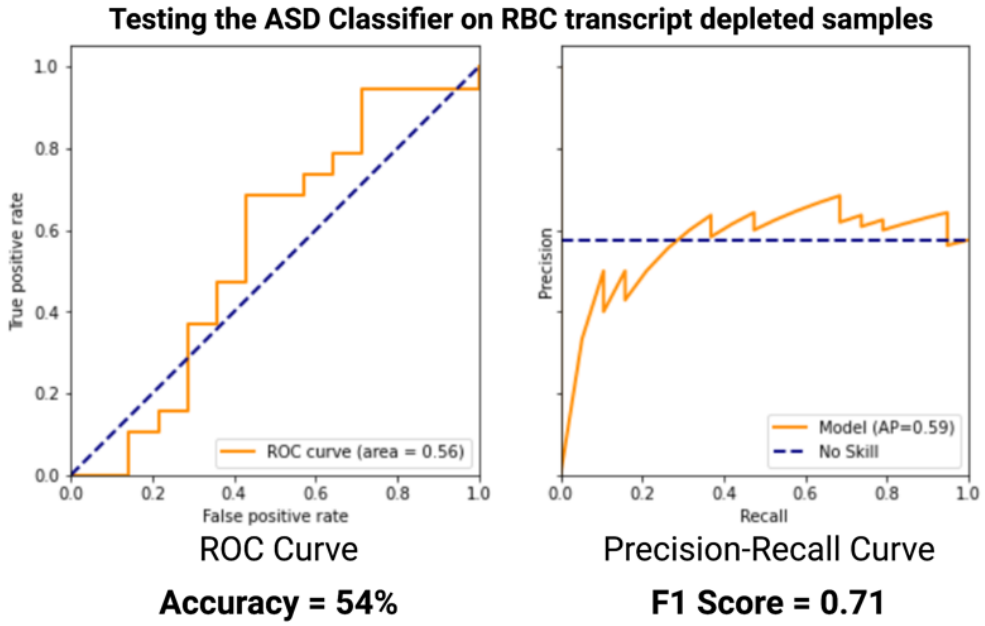


**Figure 3.29:** ***Testing the ASD classifier on samples after primer-based RBC transcript depletion.*** *Left: Receiver operating characteristic (ROC) curve for a Gaussian Naïve Bayes classifier. Right: Precision-Recall Curve using the same classifier.*

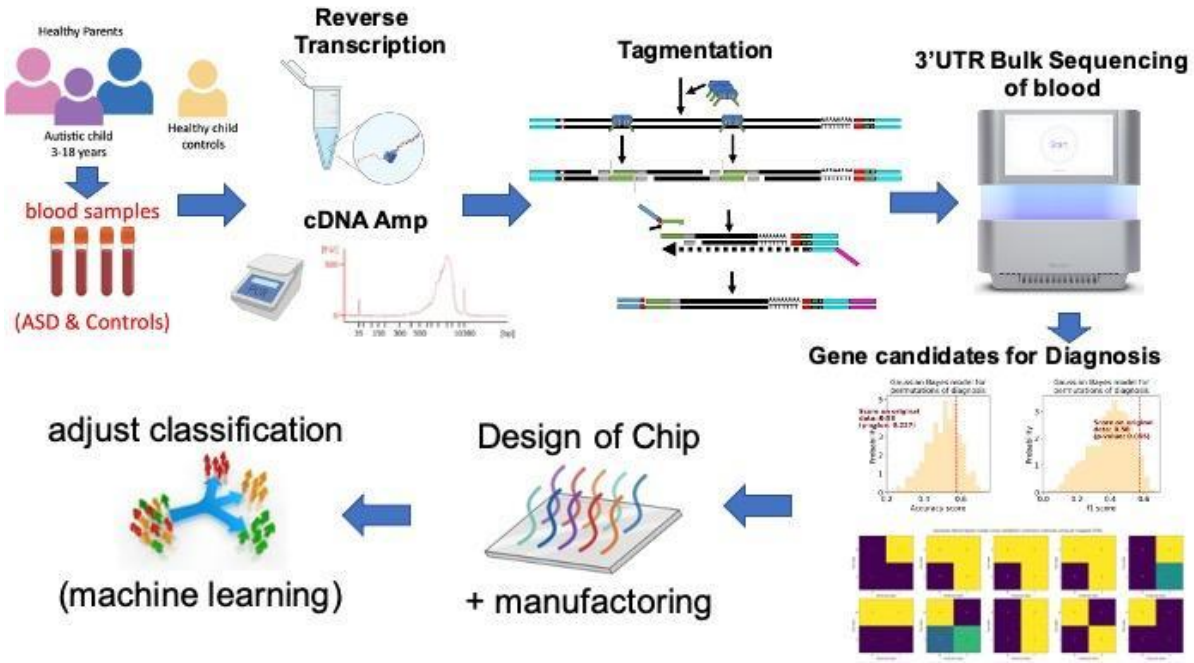


**Figure 3.30:** ***Schematic overview of 3'UTR Bulk sequencing*** *from blood to identify gene candidates to be used for diagnosis of ASD*

# 4. Discussion

## 4.1. The modified 10X 3'scRNASeq pipeline for APA in the NSC Lineage.

While there are many computational methods that can repurpose traditional RNASeq databases by trying to identify PAS within the sequenced 3'UTR, such methods do eliminate transcripts with insufficient read length that reach a credible 3'UTR end (detection a PAS in proximity of the end) (Reviewed by Shah et al., 2021; Zhang et al., 2021). Although the 3'seq methods currently used to study APA allow for the enrichment of the 3' ends of transcripts, these methods are used for bulk RNASeq on samples with prior knowledge of cell/tissue identity. In this approach (Göpferich & George et al., 2020), a modified 10x genomics 3'scRNAseq pipeline to study APA, was used to sequence the libraries with a 100bp paired end (PE) instead of the traditional (R1+R2; 26+98) sequencing read length. This allowed us to read through the 30bp of poly(A), right into the ends of the 3'UTR. Thus, by reading through the poly(A) into the 3'UTR, our method (Göpferich et al., 2020) allows for the identification of 3'UTR ends in an unbiased manner and thereby even identify novel 3'UTRs. The use of the 10X genomics 3'scRNASeq platform allows us to have APA information at single cell resolution. The pros and cons of the bioinformatic pipeline has been discussed in detail by Manuel Göpferich in his PhD thesis and would not be discussed here (Göpferich, 2020).

## 4.2. APA genes in the NSC lineage are ASD risk

As a part of this thesis, the analysis of APA in the SVZ-NSC lineage included 4 additional replicates to the ones previously reported and published (2 replicates) (Göpferich, 2020; Göpferich & George et al., 2020). The previous analysis by Manuel Göpferich used the scRNASeq data from Kalamakis et al had both old and young SVZ-NSC lineages. Despite the difference in age between the replicates the APA finding were still valid irrespective of age. However, the new NSC lineage replicates (n=4) from mostly younger mice (3 reps; 4 months + 1 rep 18 months) were added in order to increase the robustness of our findings. While, the addition of the newer replicates increased the number of APA genes observed along the NSC lineage, it appears that there was a trend towards 3'UTR shortening and not lengthening as previously reported (Göpferich, 2020; Göpferich et al., 2020). Firstly, the NSC lineage consists of different activation states and the observed trend is over the entire lineage (qNSCs to NBs). As the 6 replicates are from different batches and from slightly different age groups, it is likely that the composition of the NSC pool maybe altered. Secondly the method used to highlight

3’UTR trend is different between the previous report (n=2) to the current study. While the current method also modelled the data using multinomial regression of APA genes, it took advantage of the increased number of replicates and performed a chi-squared test on the likelihood ratio of the pseudotime effect for each replicate to determine significant APA changes. While the APA changes over the entire pseudotime is important to explore, it would be also interesting to determine APA changes related individual lineage transitions such as activation and differentiation, as they may have different trends. Global shortening of 3’UTRs have been reported in processes such as activation (T-cells) (A. R. Gruber et al., 2014; Sandberg et al., 2008), induction of reprogramming of somatic cells (Ji & Tian, 2009) and cancer (Mayr & Bartel, 2009). Shorter 3’UTRs accommodate fewer regulatory elements and can provide a surge in translation potential required during processes such as activation and proliferation (Mayr, 2019). On the other hand, studies have reported a widespread lengthening of 3’UTRs in the brain (Miura et al., 2013, 2014) which may be essential for a more nuanced post-transcriptional and/or translational control. Despite these above-mentioned differences, it was reassuring to see that the gene sets enriched using a disease ontology database were still risk genes for neurodevelopmental disorders, predominantly ASD. Additionally, the GSEA for biological processes improved from more general processes such as central nervous system development (Göpferich, 2020; Göpferich & George et al., 2020) to more resolved terms such as GO terms such as synaptic transmission, dendritic spine development, organization of dendritic spines, synapse & neuronal projections to name a few. Also, as previously stated deficits in such biological processes are typical for neurodevelopmental disorders (Lo & Lai, 2020; Nakai et al., 2018; Pan et al., 2019; Zieger & Choquet, 2021). Therefore, it was interesting that already in NSC lineage these APA genes showed association with biological processes related to neuron function in addition to generally regulating mRNA stability. This suggested that a dysregulation of APA in these genes could potentially result in the production of neurons that may have defects in homeostatic function and could potentially result in neurodevelopmental disorders like ASD.

### 4.3. The *in vitro* system for active and quiescent NSCs

The use of an in vitro system provided the flexibility in the setup required for some of the experiments. The ability to carry out experiments that required a larger starting material, a challenge with *in vivo* samples, helped us investigate translation efficiency of APA affected genes by combing scRNASeq and proteomics. With over 3200 genes correlating in their

expression between *in vitro* and *in vivo* NSC lineages (scRNASeq), the BMP4-treatemnt based *in vitro* system mimicked the cells of their *in vivo* counterparts quite well. While scRNASeq was not as sensitive to initial cell numbers, proteomics was certainly challenging. In the proteomics data from (50K) *in vitro* NSCs I detected around 2700 proteins per replicate, while our attempts do the same with *in vivo* samples were not as successful. Primarily, in order to get at least 7000-8000 aNSCs (limiting cell type) per replicate I needed to use at least 12 mice to achieve those numbers. In addition, cell loss during washing steps reduced the final cell numbers usable for the mass-spect analysis. This significantly reduced the number of proteins identified (< 1000). Additionally, for SILAC experiments (de novo proteomes) this number reduced down to around 100 proteins per replicate. Most of these proteins were highly expressed proteins like housekeeping genes that did not result in any significant observations in our APA study. Therefore, the use of an *in vitro* system mimicking active and quiescent NSCs was warranted. Despite the similarities between the two systems in their transcriptome, the APA landscape did not correlate quite well between the two. However, as reported in this thesis the genes showing differential APA between the two systems were predominantly associated with gene categories for response to stimuli, cell surface etc. These appear to be a feature of the culture conditions themselves or the relevant external stimuli that the NSC niche provides *in vivo* which is absent *in vitro*. Overall, putting these differences aside there is substantial evidence for similarity that the *in vitro* system can be used a proxy for its in vivo counterparts to study APA changes.

### 4.4. Regulation of the APLP1-CPEB4 axis

In the search for regulators of the APA changes along the NSC lineage, CPEB4 and APLP1 were identified as potential co-regulators. CPEB4 was previously reported to regulate poly(A) tail lengthening and was also reported as an ASD-risk gene (Parras et al., 2018a). They had reported that mis-splicing of CPEB4 in both human and mouse neurons altered the protein expression of ASD-risk genes (Parras et al., 2018a). While this study reported a global increase in poly(A) tail length as a consequence of CPEB4 mis-splicing, the regulation of APA by CPEB4 was unexplored. CPEB1 & 4 have been studied in the context of poly(A) tail length regulation, however, only CPEB1 has previously associated with APA (Cao et al., 2005). Additionally, our study has shown that PAS selection is regulated by the presence of an upstream CPE motif that are primarily CPEB4 binders in the NSC lineage. A dual role for CPEB4 in modulating both APA (primarily nuclear) as well as poly(A) tail length (primarily

cytoplasmic) seems a likely explanation. The presence of a co-regulator that could balance nuclear APA and cytoplasmic polyadenylation by regulating the sub-cellular localization of CPEB4 seems warranted. This study identified APLP1, a transmembrane protein, to be a potential co-regulator of APA in the NSC lineage working in tandem with CPEB4. The altered APA upon APLP1 knockout, more specifically altering the expression of CPE containing transcripts along the NSC lineage adds more credibility to the hypothesis that APLP1 may be acting in a CPEB4/CPE dependent manner. In addition to the APA dysregulation upon APLP1 knockout these mice also showed an ASD-like behaviour phenotype. Interestingly, CPEB4 mutant mice also showed ASD-related electrophysiological, anatomical, and/or behavioural defects (Parras et al., 2018a). Furthermore, an overlap of altered APA genes upon APLP1 knockout to those altered in human ASD (Göpferich, 2020) does suggest the relevance of the APLP1-CPEB4 axis in ASD. Further, splice variants especially related to microexons have been previously reported in ASD patients (Irimia et al., 2014) and the CPEB4 microexon (exon4) is one such example. The CPEB4 transcript lacking a 24-nucleotide neuron-specific microexon has been reported in ASD (Parras et al., 2018a). This encodes an 8-amino-acid peptide on the CPEB4 protein that allows for post-translational regulation via motifs for phosphorylation by AKT, S6K, PKA or PKC. These motifs may be essential in the interaction with APLP1 in healthy neurons while such an interaction could be dysregulated in ASD patients.

### 4.5. Regulation of translation & local translation via APA

While the impact of APA on translation for a subset of transcripts has been shown, this is not true for all APA affected genes. It is known that APA does not just affect translation & stability but also localization (Mayr, 2019; Tian & Manley, 2017). Therefore, it is likely that for those transcripts where APA does not affect protein outcome, the consequence of APA could alter transcript compartmentalization. As previously mentioned, transcripts with CPE motifs have been detected in dendrites and their binding to CPEB proteins regulates local translation at synapses (Fernandez-Moya et al., 2014; Groisman et al., 2006; Weill et al., 2012). Therefore, regulation of translation and sub-cellular localization are both crucial processes required during neurogenesis, maturation and eventually neuron function. For example, receptor-ligand interactions between NSCs and their niche that maintain quiescence in NSCs (Llorens-Bobadilla et al., 2015) may likely be the consequence of localized rapid protein production at the site of interaction (plasma membrane). Aberrant local translation could lead to alterations in the NSC's response to niche signals. In APLP1-/- mice where such receptor-ligand

interactions are likely disrupted and thereby affect the maintenance of quiescence I did observe a reduction in the NSC pool. The possible increase in the activation rate in APLP1-/- does not seem to translate into the production of functional OB neurons. APLP1-/- mice do exhibit autistic-like behaviour and could be the consequence of aberrant neuronal signalling caused by dysregulated regulation of neuronal communication genes. Vnencak et al reported that the deletion of the APLP1 increased excitatory synaptic transmission and reduced network inhibition in the dentate gyrus (Vnencak et al., 2015). Further, Schilling et al reported a reduction in basal synaptic transmission resulting from reduced spine density in aged APLP1-/- mice (Schilling et al., 2017).

### 4.6. Global dysregulation of APA is a phenotype of ASD

The regulation of APA has been documented in different diseases such as cancer, autoimmune disorders, and neurodevelopmental disorders. A global trend of 3'UTR shortening has been documented in different malignancies (López De Silanes et al., 2007). A vast majority of studies that have investigated causes of ASD have focused on single gene or even polygenic variants resulting from loss of function mutations in the coding regions of ASD-risk genes (Reviewed in Lim et al., 2022). However, in the recent years there is mounting evidence that implicated non-coding regions alterations that contribute to the ASD phenotype (Szkop et al., 2017; Wanke et al., 2018). While all of these studies focused on non-coding regions alterations for single genes (as summarized by Szkop et al., 2017; Wanke et al., 2018) no global mechanism of APA dysregulation has been investigated in ASD. Such a mechanism has only been hypothesised by Szkop et al's APA analysis of publicly available bulk RNASeq datasets. Consistent with this study, they reported a trend for longer 3'UTRs in mature neurons of ASD patients compared to their healthy counterparts in a subset of genes (Szkop et al., 2017). Further, iPSC derived NSCs from ASD patients resulted in a skewed glial fate, impaired maturation in excitatory neurons and altered calcium-signalling compared to healthy controls (Lam et al., 2019). Brumbaugh et al had reported a knockdown of NUDT21, a core member of the APA machinery, in iPSCs also resulted in altered neuronal fate choice (Brumbaugh et al., 2018). Reduced expression of NUDT21 also resulted in APA dysregulation in hundreds of genes that resulted in learning deficits in mice. A significant fraction of these APA dysregulated genes (when mutated) are risk genes for intellectual disability, a common co-morbidity of ASD (Alcott et al., 2020). This study therefore highlights the significance of APA being tightly regulated in neurogenesis already at the level of progenitors. Any dysregulation of APA could potentially result in altered fate choices and persist within neurons to impair its function

consequentially resulting in NDDs such as ASD. Such a phenotype of global dysregulation of APA persistent across many different cell types and likely tissues warrants its use for ASD diagnosis.

An important fact to be considered is that, in order to study APA dysregulation in ASD, a publicly available RNASeq dataset was used (Velmeshev et al., 2019c). A limitation of such a study in the human brain was the use of post-mortem tissues and its difficulties in having intact cells and so Velmeshev et al were restricted to performing single nuclear RNASeq as opposed to single cell transcriptomics. Therefore, the RNASeq information and by extension our APA analysis gives us only half the picture, i.e., of what APA looks like in the nucleus. Distal PAS is usually the strongest and is most selected for, at least during the initial phase of mRNA processing within the nucleus. This longer isoform serves as a substrate for further processing in the later phase of APA (Tang et al., 2022). This might explain why there is a general trend in long 3'UTRs in the RNASeq data from Velmeshev et al. Despite the fact that only the nuclear snippet of the APA events was observed in this dataset, the findings are still valid as both the healthy and ASD patient samples have been processed alike. It does appear that there may be defects in the APA machinery in ASD patients that might lead to improperly processed mRNA transcripts. A consequence of such a processing could alter the retention of transcripts with such longer 3'UTRs (L. L. Chen & Carmichael, 2009; Jacq et al., 2020; Tushev et al., 2018).

### 4.7. Neurological deficits of ASD

Deficits in social communication and restricted and repetitive behaviours which are characteristic symptoms of ASD have been shown to result from excitatory-inhibitory (E-I) imbalances in the neural networks (Rubenstein & Merzenich, 2003). For example, optogenetic studies in the prefrontal cortex have showed an association of increased E-I ratio with social and behavioural deficits as seen in ASD (Yizhar, 2012; Yizhar et al., 2011). ASD has been characterized as a developmental disconnection syndrome with local hyperconnectivity along with long-range hypoconnectivity to the rest of the brain (Courchesne & Pierce, 2005; Peters et al., 2013; Tye & Bolton, 2013). While, single gene origins of ASD (syndromic ASD) are much rarer, such as tuberous sclerosis complex (TSC) 1 & 2, Fragile X syndrome variants, the bulk of ASD is highly polygenic (Gaugler et al., 2014; Robinson et al., 2016). Studies have shown that most ASD manifest without an apparent cause; referred to as idiopathic ASD. A recent exome sequencing study of ASD patients reported that the bulk of the identified ASD

risk gene are associated with regulating gene expression or neuronal communication that either indirectly or directly affect neurodevelopment and induce neurophysiological changes (Satterstrom et al., 2020). Maturing excitatory and inhibitory neurons in the human cortex appear to have an enrichment of many different ASD risk genes. This dysregulation of gene expression consequentially may contribute to the E-I imbalances observed in ASD (Satterstrom et al., 2020). Therefore, a more global mechanism in determining the cause of ASD is merited, where a large cluster of smaller changes across many different genes eventually add up, leading to the observed neurogenic deficits.

### 4.8. Potential upstream regulators of APA

The role of the APLP1/CPEB4 axis in APA regulation along the NSC lineage and its likely dysregulation in ASD has been explored in this thesis. This does open-up potential avenues for targeting the APA machinery that could thereby regulate process like exit from quiescence (cancer therapy) or even potentially ameliorate the impact of APA dysregulation in ASD. For example, quiescent cancer cells (QCCs) are more resistant to traditional treatment method that primarily target actively proliferating cells (Endo & Inoue, 2019; Nabil et al., 2021; Phan & Croucher, 2020). By targeted regulation of the APA repertoire of the QCCs they could potentially be activated and thereby allow such resistant cancers to be treated. While the APA phenotype in ASD might be beneficial in the diagnosis of ASD and treatment of ASD may be much more challenging. Our current understanding of a master regulator of the APA is not very clear. Szkop et al had previously hypothesized that APA dysregulation in ASD, is the due to disruptions in the calcium signalling networks that could modify enzymatic kinetics of proteins such as RNA polymerase II (RNAL Pol II) consequently affecting PAS selection (Szkop et al., 2017). Calcium, an essential secondary messenger, has been implicated in various signalling cascades, both intracellular and extracellular. Calcium signalling has been known to interact with reactive oxygen species (ROS) and regulating cellular processes such as growth and cell death (Görlach et al., 2015). Additionally, alterations in ROS signalling have been reported in ASD patients. Further, ASD patients are more susceptible to experiencing oxidative stress and are highly prone to ROS-mediated damages leading to neuronal toxicity (Ohja et al., 2018; Pangrazzi et al., 2020; Zeidán-Chuliá et al., 2014). ROS levels have been shown to affect APA and splicing, favouring proximal PAS and usage of the terminal exon (Fontana et al., 2017). The overall levels of ROS and calcium signalling vary between the different cell type of the NSC lineage (Maffezzini et al., 2020; Toth et al., 2016) and any changes thereof could

consequentially alter PAS usage and effectively the cell state. If such a master regulator is identified, it would indeed open up possibilities for APA modulation and rectify any dysregulation for example seen in ASD.

### 4.9. APA as a diagnostic tool for ASD

While exploring treatment avenues for ASD is essential and an interesting area of research, focus on early diagnosis of ASD can help provide timely intervention to the affected child. In the absence of a single gene biomarker, the diagnosis of ASD currently relies heavily on clinical tests that assess behaviour and cognitive function in addition to the developmental history of the child (American Psychiatric Association, 2013; Charman & Gotham, 2013; Hus & Segal, 2021). As ASD is polygenic in nature it is likely that a more global mechanism like APA can characterize the disease better. While screening for genetic markers can be easily carried out using any tissue as the genetic makeup is consistent across all cells of the body, process such as APA are unique to cell and tissue types. Our challenge was to identify differences in APA between different cell types. While I did find that there was a general agreement between APA of different blood cells, there definitely existed some difference between them. While the genes that do not vary much between the different cell types could be used, there is a likelihood that some of the more potent candidates specific to a certain cell type may not be a good predictor when using whole blood for ASD diagnosis. Sorting of individual blood populations and looking for cell type specific APA changes can be laborious and cost extensive. Despite these challenges this approach was able to identify candidate from whole blood that can be used for ASD diagnosis.

Additionally, the whole blood samples that I used for this study were from a genetic study (Satterstrom et al., 2020) that used parents as healthy controls. Therefore, I additionally included 5 age matched healthy controls to have more appropriate controls. More of these age matched controls are currently being acquired that can serve as better controls for the analysis. Further the parents used in the genetic study were not necessarily healthy controls. The clinical information from the ASD patients (which was acquired only recently) does provide us a brief history of the parents. Despite not being diagnosed with ASD, some of the parents did have a history of social communication and learning deficits, intellectual disability or even extensive family history of ASD and other co-morbidities. Currently reanalysing the data after removing parents with questionable clinical history is being carried out.

While the potential candidates identified were predictive of the diagnosis, the variation within the patient samples was quite high. While there were some candidates that were quite nicely separated between the ASD and controls there were only few samples showing this effect while the rest were dropouts. This could be consequence of sequencing itself or even limited sequencing depth. For some other samples where the coverage good there seemed to be a high overlap in the 3'UTR length distribution between healthy and ASD samples. This was likely the consequence of the high variance in the data as mentioned before. ASD is usually manifest along with other co-morbidities such and ADHD and ID and other neurodevelopmental disorders. Additionally, ASD manifests with varying degrees of severity. This may explain the high variance in the data. Combining all of these different ASD categories into one group 'ASD' would make it harder to identify clear candidates for ASD diagnosis. The best approach would be to have a well curated list of patients classified according to their clinical information and find individual marker candidates that could allow patient stratification. While this is a better approach indeed, in order to be able to have sufficient sample representation for each subcategory for ASD, I would need to use a much larger sample set to achieve such a patient stratification. Nevertheless, APA does seem to be a great predictor for ASD and could likely be used for ASD diagnosis in the near future.

## 5. Bibliography


Aase, J. M. (Jon M. (1990). *Diagnostic dysmorphology*. Plenum Medical Book Co. https://link.springer.com/book/9780306434440

Alcott, C. E., Yalamanchili, H. K., Ji, P., van der Heijden, M. E., Saltzman, A., Elrod, N., Lin, A., Leng, M., Bhatt, B., Hao, S., Wang, Q., Saliba, A., Tang, J., Malovannaya, A., Wagner, E. J., Liu, Z., & Zoghbi, H. Y. (2020). Partial loss of CFIM25 causes learning deficits and aberrant neuronal alternative polyadenylation. *ELife*, *9*. https://doi.org/10.7554/eLife.50895

Alonso-Gonzalez, A., Calaza, M., Rodriguez-Fontenla, C., & Carracedo, A. (2019). Novel gene-based analysis of ASD GWAS: Insight into the biological role of associated genes. *Frontiers in Genetics*, *10*(JUL), 733. https://doi.org/10.3389/FGENE.2019.00733/BIBTEX

Altman, J., & Das, G. D. (1965). Autoradiographic and histological evidence of postnatal hippocampal neurogenesis in rats. *Journal of Comparative Neurology*, *124*(3), 319–335. https://doi.org/10.1002/CNE.901240303

Alvarez-Buylla, A., & García-Verdugo, J. M. (2002). Neurogenesis in Adult Subventricular Zone. *Journal of Neuroscience*, *22*(3), 629–634. https://doi.org/10.1523/JNEUROSCI.22-03-00629.2002

American Psychiatric Association. (2013). *Diagnostic and Statistical Manual of Mental Disorders*. https://doi.org/10.1176/APPI.BOOKS.9780890425596

Aßfalg, M. (2021). *Elucidating the role of key regulators of alternative polyadenylation in neurogenesis using the CRISPR/Cas system*. Ruprecht-Karls-Universität Heidelberg.

Baio, J., Wiggins, L., Christensen, D. L., Maenner, M. J., Daniels, J., Warren, Z., Kurzius-Spencer, M., Zahorodny, W., Rosenberg, C. R., White, T., Durkin, M. S., Imm, P., Nikolaou, L., Yeargin-Allsopp, M., Lee, L. C., Harrington, R., Lopez, M., Fitzgerald, R. T., Hewitt, A., … Dowling, N. F. (2018). Prevalence of Autism Spectrum Disorder Among Children Aged 8 Years — Autism and Developmental Disabilities Monitoring Network, 11 Sites, United States, 2014. *MMWR Surveillance Summaries*, *67*(6), 1. https://doi.org/10.15585/MMWR.SS6706A1

Barnes, I. H. A., Ibarra-Soria, X., Fitzgerald, S., Gonzalez, J. M., Davidson, C., Hardy, M. P., Manthravadi, D., Van Gerven, L., Jorissen, M., Zeng, Z., Khan, M., Mombaerts, P., Harrow, J., Logan, D. W., & Frankish, A. (2020). Expert curation of the human and mouse olfactory receptor gene repertoires identifies conserved coding regions split across two exons. *BMC Genomics*, *21*(1), 1–15. https://doi.org/10.1186/S12864-020-6583-3/FIGURES/5

Bartocci, M., Winberg, J., Ruggiero, C., Bergqvist, L. L., Serra, G., & Lagercrantz, H. (2000). Activation of Olfactory Cortex in Newborn Infants After Odor Stimulation: A Functional Near-Infrared Spectroscopy Study. *Pediatric Research 2000 48:1*, *48*(1), 18–23. https://doi.org/10.1203/00006450-200007000-00006

Baser, A., Skabkin, M., Kleber, S., Dang, Y., Gülcüler Balta, G. S., Kalamakis, G., Göpferich, M., Ibañez, D. C., Schefzik, R., Lopez, A. S., Bobadilla, E. L., Schultz, C., Fischer, B., & Martin-Villalba, A. (2019). Onset of differentiation is post-transcriptionally controlled in adult neural stem cells. *Nature*, *566*(7742), 100–104. https://doi.org/10.1038/s41586-019-0888-x

Bashaw, G. J., & Klein, R. (2010). Signaling from Axon Guidance Receptors. *Cold Spring Harbor Perspectives in Biology*, *2*(5), 1941–1942. https://doi.org/10.1101/CSHPERSPECT.A001941

Bauman, M. L., & Kemper, T. L. (2005). Neuroanatomic observations of the brain in autism: a review and future directions. *International Journal of Developmental Neuroscience*,

*23*(2–3), 183–187. https://doi.org/10.1016/J.IJDEVNEU.2004.09.006
Bava, F.-A., Eliscovich, C., Ferreira, P. G., Miñana, B., Ben-Dov, C., Guigó, R., Valcárcel, J., & Méndez, R. (2013). CPEB1 coordinates alternative 3'-UTR formation with translational regulation. *Nature*, *495*(7439), 121–125. https://doi.org/10.1038/nature11901
Baxter, A. J., Brugha, T. S., Erskine, H. E., Scheurer, R. W., Vos, T., & Scott, J. G. (2015). The epidemiology and global burden of autism spectrum disorders. *Psychological Medicine*, *45*(3), 601–613. https://doi.org/10.1017/S003329171400172X
Beaudoing, E., Freier, S., Wyatt, J. R., Claverie, J. M., & Gautheret, D. (2000). Patterns of variant polyadenylation signal usage in human genes. *Genome Research*, *10*(7), 1001–1010. https://doi.org/10.1101/GR.10.7.1001
Benjamini, Y., & Hochberg, Y. (1995). Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society: Series B (Methodological)*, *57*(1), 289–300. https://doi.org/10.1111/J.2517-6161.1995.TB02031.X
Betancur, C. (2011). Etiological heterogeneity in autism spectrum disorders: more than 100 genetic and genomic disorders and still counting. *Brain Research*, *1380*, 42–77. https://doi.org/10.1016/J.BRAINRES.2010.11.078
Blomfield, I. M., Rocamonde, B., del Mar Masdeu, M., Mulugeta, E., Vaga, S., van den Berg, D. L. C., Huillard, E., Guillemot, F., & Urbán, N. (2019). Id4 promotes the elimination of the pro-activation factor Ascl1 to maintain quiescence of adult hippocampal stem cells. *ELife*, *8*. https://doi.org/10.7554/ELIFE.48561
Bozza, T., McGann, J. P., Mombaerts, P., & Wachowiak, M. (2004). In Vivo Imaging of Neuronal Activity by Targeted Expression of a Genetically Encoded Probe in the Mouse. *Neuron*, *42*(1), 9–21. https://doi.org/10.1016/S0896-6273(04)00144-8
Brock, O., Keller, M., Veyrac, A., Douhard, Q., & Bakker, J. (2010). Short term treatment with estradiol decreases the rate of newly generated cells in the subventricular zone and main olfactory bulb of adult female mice. *Neuroscience*, *166*(2), 368–376. https://doi.org/10.1016/J.NEUROSCIENCE.2009.12.050
Brumbaugh, J., Di Stefano, B., Wang, X., Borkent, M., Forouzmand, E., Clowers, K. J., Ji, F., Schwarz, B. A., Kalocsay, M., Elledge, S. J., Chen, Y., Sadreyev, R. I., Gygi, S. P., Hu, G., Shi, Y., & Hochedlinger, K. (2018). Nudt21 Controls Cell Fate by Connecting Alternative Polyadenylation to Chromatin Signaling. *Cell*, *172*(1–2), 106-120.e21. https://doi.org/10.1016/j.cell.2017.11.023
Buck, L., & Axel, R. (1991). A novel multigene family may encode odorant receptors: A molecular basis for odor recognition. *Cell*, *65*(1), 175–187. https://doi.org/10.1016/0092-8674(91)90418-X
Burek, M. J., Nordeen, K. W., & Nordeen, E. J. (1995). Estrogen promotes neuron addition to an avian song-control nucleus by regulating post-mitotic events. *Developmental Brain Research*, *85*(2), 220–224. https://doi.org/10.1016/0165-3806(94)00215-L
Cao, Q., Huang, Y.-S., Kan, M.-C., & Richter, J. D. (2005). Amyloid precursor proteins anchor CPEB to membranes and promote polyadenylation-induced translation. *Molecular and Cellular Biology*, *25*(24), 10930–10939. https://doi.org/10.1128/MCB.25.24.10930-10939.2005
Casanova, M. F., Casanova, E. L., Frye, R. E., Baeza-Velasco, C., LaSalle, J. M., Hagerman, R. J., Scherer, S. W., & Natowicz, M. R. (2020). Editorial: Secondary vs. Idiopathic Autism. *Frontiers in Psychiatry*, *11*. https://doi.org/10.3389/FPSYT.2020.00297
Cecchi, G. A., Petreanu, L. T., Alvarez-Buylla, A., & Magnasco, M. O. (2001). Unsupervised Learning and Adaptation in a Model of Adult Neurogenesis. *Journal of Computational Neuroscience*, *11*, 175–182.

Charlesworth, A., Cox, L. L., & MacNicol, A. M. (2004). Cytoplasmic polyadenylation element (CPE)- and CPE-binding protein (CPEB)-independent mechanisms regulate early class maternal mRNA translational activation in Xenopus oocytes. *The Journal of Biological Chemistry*, *279*(17), 17650–17659. https://doi.org/10.1074/jbc.M313837200

Charman, T., & Gotham, K. (2013). Measurement Issues: Screening and diagnostic instruments for autism spectrum disorders – lessons from research and practise. *Child and Adolescent Mental Health*, *18*(1), 52–63. https://doi.org/10.1111/J.1475-3588.2012.00664.X

Chen, C. Y., Chen, S. T., Juan, H. F., & Huang, H. C. (2012). Lengthening of 3'UTR increases with morphological complexity in animal evolution. *Bioinformatics (Oxford, England)*, *28*(24), 3178–3181. https://doi.org/10.1093/BIOINFORMATICS/BTS623

Chen, J., & Li, H. (2013). VARIABLE SELECTION FOR SPARSE DIRICHLET-MULTINOMIAL REGRESSION WITH AN APPLICATION TO MICROBIOME DATA ANALYSIS. *The Annals of Applied Statistics*, *7*(1), 418–442. https://doi.org/10.1214/12-AOAS592

Chen, L. L., & Carmichael, G. G. (2009). Altered nuclear retention of mRNAs containing inverted repeats in human embryonic stem cells: functional role of a nuclear noncoding RNA. *Molecular Cell*, *35*(4), 467–478. https://doi.org/10.1016/J.MOLCEL.2009.06.027

Chong, E., Moroni, M., Wilson, C., Shoham, S., Panzeri, S., & Rinberg, D. (2020). Manipulating synthetic optogenetic odors reveals the coding logic of olfactory perception. *Science*, *368*(6497). https://doi.org/10.1126/SCIENCE.ABA2357/SUPPL_FILE/ABA2357-CHONG-SM.PDF

Cohen, M. M. (2003). Mental deficiency, alterations in performance, and CNS abnormalities in Overgrowth syndromes. *American Journal of Medical Genetics Part C: Seminars in Medical Genetics*, *117C*(1), 49–56. https://doi.org/10.1002/AJMG.C.10013

Corbett, B. A., Schwartzman, J. M., Libsack, E. J., Muscatello, R. A., Lerner, M. D., Simmons, G. L., & White, S. W. (2021). Camouflaging in Autism: Examining Sex-Based and Compensatory Models in Social Cognition and Communication. *Autism Research : Official Journal of the International Society for Autism Research*, *14*(1), 127–142. https://doi.org/10.1002/AUR.2440

Courchesne, E., & Pierce, K. (2005). Why the frontal cortex in autism might be talking only to itself: local over-connectivity but long-distance disconnection. *Current Opinion in Neurobiology*, *15*(2), 225–230. https://doi.org/10.1016/J.CONB.2005.03.001

Crow, A. J. D., Janssen, J. M., Vickers, K. L., Parish-Morris, J., Moberg, P. J., & Roalf, D. R. (2020). Olfactory Dysfunction in Neurodevelopmental Disorders: A Meta-Analytic Review of Autism Spectrum Disorders, Attention Deficit/Hyperactivity Disorder and Obsessive-Compulsive Disorder. *Journal of Autism and Developmental Disorders*, *50*(8), 2685. https://doi.org/10.1007/S10803-020-04376-9

Dantonel, J. C., Murthy, K. G. K., Manjey, J. L., & Tora, L. (1997). Transcription factor TFIID recruits factor CPSF for formation of 3' end of mRNA. *Nature*, *389*(6649), 399–402. https://doi.org/10.1038/38763

Dawson, G., Rogers, S., Munson, J., Smith, M., Winter, J., Greenson, J., Donaldson, A., & Varley, J. (2010). Randomized, controlled trial of an intervention for toddlers with autism: the Early Start Denver Model. *Pediatrics*, *125*(1). https://doi.org/10.1542/PEDS.2009-0958

De Rubeis, S., & Buxbaum, J. D. (2015). Genetics and genomics of autism spectrum disorder: embracing complexity. *Human Molecular Genetics*, *24*(R1), R24–R31. https://doi.org/10.1093/HMG/DDV273

Derti, A., Garrett-Engele, P., MacIsaac, K. D., Stevens, R. C., Sriram, S., Chen, R., Rohl, C.

A., Johnson, J. M., & Babak, T. (2012). A quantitative atlas of polyadenylation in five mammals. *Genome Research*, *22*(6), 1173–1183. https://doi.org/10.1101/GR.132563.111

Devescovi, R., Monasta, L., Mancini, A., Bin, M., Vellante, V., Carrozzi, M., & Colombi, C. (2016). Early diagnosis and Early Start Denver Model intervention in autism spectrum disorders delivered in an Italian Public Health System service. *Neuropsychiatric Disease and Treatment*, *12*, 1379–1384. https://doi.org/10.2147/NDT.S106850

Devlin, B., & Scherer, S. W. (2012). Genetic architecture in autism spectrum disorder. *Current Opinion in Genetics & Development*, *22*(3), 229–237. https://doi.org/10.1016/J.GDE.2012.03.002

Di Giammartino, D. C., Nishida, K., & Manley, J. L. (2011). Mechanisms and consequences of alternative polyadenylation. *Molecular Cell*, *43*(6), 853–866. https://doi.org/10.1016/J.MOLCEL.2011.08.017

Dobin, A., Davis, C. A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M., & Gingeras, T. R. (2013). STAR: ultrafast universal RNA-seq aligner. *Bioinformatics*, *29*(1), 15. https://doi.org/10.1093/BIOINFORMATICS/BTS635

Doernberg, E., & Hollander, E. (2016). Neurodevelopmental Disorders (ASD and ADHD): DSM-5, ICD-10, and ICD-11. *CNS Spectrums*, *21*(4), 295–299. https://doi.org/10.1017/S1092852916000262

Doetsch, F., Caillé, I., Lim, D. A., Manuel García-Verdugo, J., & Alvarez-Buylla, A. (1999). Subventricular Zone Astrocytes Are Neural Stem Cells in the Adult Mammalian Brain. *Cell*, *97*, 703–716.

Doty, R. L. (2012). Olfaction in Parkinson's disease and related disorders. *Neurobiology of Disease*, *46*(3), 527–552. https://doi.org/10.1016/J.NBD.2011.10.026

Eckmann, C. R., Rammelt, C., & Wahle, E. (2011). Control of poly(A) tail length. *Wiley Interdisciplinary Reviews. RNA*, *2*(3), 348–361. https://doi.org/10.1002/WRNA.56

Endo, H., & Inoue, M. (2019). Dormancy in cancer. *Cancer Science*, *110*(2), 474. https://doi.org/10.1111/CAS.13917

Eriksson, P. S., Perfilieva, E., Björk-Eriksson, T., Alborn, A. M., Nordborg, C., Peterson, D. A., & Gage, F. H. (1998). Neurogenesis in the adult human hippocampus. *Nature Medicine 1998 4:11*, *4*(11), 1313–1317. https://doi.org/10.1038/3305

Ester, M., Kriegel, H.-P., Sander, J., & Xu, X. (1996). *A Density-Based Algorithm for Discovering Clusters in Large Spatial Databases with Noise*. www.aaai.org

Everaert, C., Verwilt, J., Verniers, K., Vandamme, N., Rubio, A. M., Vandesompele, J., & Mestdagh, P. (2022). Blocking unwanted sequences by target specific high-affinity binding oligonucleotides during transcriptome library preparation. *BioRxiv*, 2022.03.11.483910. https://doi.org/10.1101/2022.03.11.483910

Farrants, H., Tarnawski, M., Müller, T. G., Otsuka, S., Hiblot, J., Koch, B., Kueblbeck, M., Kräusslich, H. G., Ellenberg, J., & Johnsson, K. (2020). Chemogenetic Control of Nanobodies. *Nature Methods 2020 17:3*, *17*(3), 279–282. https://doi.org/10.1038/s41592-020-0746-7

Fernandez-Moya, S. M., Bauer, K. E., & Kiebler, M. A. (2014). Meet the players: local translation at the synapse. *Frontiers in Molecular Neuroscience*, *7*, 84. https://doi.org/10.3389/fnmol.2014.00084

Fontana, G. A., Rigamonti, A., Lenzken, S. C., Filosa, G., Alvarez, R., Calogero, R., Bianchi, M. E., & Barabino, S. M. L. (2017). Oxidative stress controls the choice of alternative last exons via a Brahma-BRCA1-CstF pathway. *Nucleic Acids Research*, *45*(2), 902–914. https://doi.org/10.1093/nar/gkw780

Gaugler, T., Klei, L., Sanders, S. J., Bodea, C. A., Goldberg, A. P., Lee, A. B., Mahajan, M., Manaa, D., Pawitan, Y., Reichert, J., Ripke, S., Sandin, S., Sklar, P., Svantesson, O.,

Reichenberg, A., Hultman, C. M., Devlin, B., Roeder, K., & Buxbaum, J. D. (2014). Most genetic risk for autism resides with common variation. *Nature Genetics*, *46*(8), 881–885. https://doi.org/10.1038/ng.3039

Gheusi, G., Cremer, H., McLean, H., Chazal, G., Vincent, J. D., & Lledo, P. M. (2000). Importance of newly generated neurons in the adult olfactory bulb for odor discrimination. *Proceedings of the National Academy of Sciences of the United States of America*, *97*(4), 1823–1828. https://doi.org/10.1073/PNAS.97.4.1823/ASSET/D88308AD-F939-4F9F-87D1-191591728013/ASSETS/GRAPHIC/PQ0205227006.JPEG

Gheusi, G., Lepousez, G., & Lledo, P. M. (2012). Adult-born neurons in the olfactory bulb: Integration and functional consequences. *Current Topics in Behavioral Neurosciences*, *15*, 49–72. https://doi.org/10.1007/7854_2012_228/FIGURES/3

Gilbert, J., & Man, H. Y. (2017a). Fundamental Elements in Autism: From Neurogenesis and Neurite Growth to Synaptic Plasticity. *Frontiers in Cellular Neuroscience*, *11*. https://doi.org/10.3389/FNCEL.2017.00359

Gilbert, J., & Man, H. Y. (2017b). Fundamental Elements in Autism: From Neurogenesis and Neurite Growth to Synaptic Plasticity. *Frontiers in Cellular Neuroscience*, *11*. https://doi.org/10.3389/FNCEL.2017.00359

Göpferich, M. (2020). *Single cell 3'UTR analysis identifies changes in alternative polyadenylation throughout neuronal differentiation and in autism*. University of Heidelberg.

Göpferich, M., George, N. O., Muelas, A. D., Bizyn, A., Pascual, R., Fijalkowska, D., Kalamakis, G., Müller, U., Krijgsveld, J., Mendez, R., Fariñas, I., Huber, W., Anders, S., & Martin-Villalba, A. (2020). Single cell 3'UTR analysis identifies changes in alternative polyadenylation throughout neuronal differentiation and in autism. *BioRxiv*, 2020.08.12.247627. https://doi.org/10.1101/2020.08.12.247627

Görlach, A., Bertram, K., Hudecova, S., & Krizanova, O. (2015). Calcium and ROS: A mutual interplay. *Redox Biology*, *6*, 260–271. https://doi.org/10.1016/j.redox.2015.08.010

Granlund, M., Imms, C., King, G., Andersson, A. K., Augustine, L., Brooks, R., Danielsson, H., Gothilander, J., Ivarsson, M., Lundqvist, L. O., Lygnegård, F., & Almqvist, L. (2021). Definitions and Operationalization of Mental Health Problems, Wellbeing and Participation Constructs in Children with NDD: Distinctions and Clarifications. *International Journal of Environmental Research and Public Health 2021, Vol. 18, Page 1656*, *18*(4), 1656. https://doi.org/10.3390/IJERPH18041656

Greco, B., Managò, F., Tucci, V., Kao, H. T., Valtorta, F., & Benfenati, F. (2013). Autism-related behavioral abnormalities in synapsin knockout mice. *Behavioural Brain Research*, *251*, 65. https://doi.org/10.1016/J.BBR.2012.12.015

Grelat, A., Benoit, L., Wagner, S., Moigneu, C., Lledo, P. M., & Alonso, M. (2018). Adult-born neurons boost odor–reward association. *Proceedings of the National Academy of Sciences of the United States of America*, *115*(10), 2514–2519. https://doi.org/10.1073/PNAS.1716400115/SUPPL_FILE/PNAS.201716400SI.PDF

Groisman, I., Ivshina, M., Marin, V., Kennedy, N. J., Davis, R. J., & Richter, J. D. (2006). Control of cellular senescence by CPEB. *Genes & Development*, *20*(19), 2701–2712. https://doi.org/10.1101/gad.1438906

Gruber, A. J., Schmidt, R., Gruber, A. R., Martin, G., Ghosh, S., Belmadani, M., Keller, W., & Zavolan, M. (2016). A comprehensive analysis of 3' end sequencing data sets reveals novel polyadenylation signals and the repressive role of heterogeneous ribonucleoprotein C on cleavage and polyadenylation. *Genome Research*, *26*(8), 1145–1159. https://doi.org/10.1101/GR.202432.115

Gruber, A. J., & Zavolan, M. (2019). Alternative cleavage and polyadenylation in health and disease. *Nature Reviews Genetics 2019 20:10*, *20*(10), 599–614. https://doi.org/10.1038/s41576-019-0145-z

Gruber, A. R., Martin, G., Müller, P., Schmidt, A., Gruber, A. J., Gumienny, R., Mittal, N., Jayachandran, R., Pieters, J., Keller, W., van Nimwegen, E., & Zavolan, M. (2014). Global 3' UTR shortening has a limited effect on protein abundance in proliferating T cells. *Nature Communications*, *5*, 5465. https://doi.org/10.1038/ncomms6465

Haghverdi, L., Büttner, M., Wolf, F. A., Buettner, F., & Theis, F. J. (2016). Diffusion pseudotime robustly reconstructs lineage branching. *Nature Methods 2016 13:10*, *13*(10), 845–848. https://doi.org/10.1038/nmeth.3971

Harvey, J. D., & Heinbockel, T. (2018). Neuromodulation of Synaptic Transmission in the Main Olfactory Bulb. *International Journal of Environmental Research and Public Health 2018, Vol. 15, Page 2194*, *15*(10), 2194. https://doi.org/10.3390/IJERPH15102194

Heber, S., Herms, J., Gajic, V., Hainfellner, J., Aguzzi, A., Rulicke, T., Kretzschmar, H., Von Koch, C., Sisodia, S., Tremml, P., Lipp, H. P., Wolfer, D. P., & Muller, U. (2000). Mice with Combined Gene Knock-Outs Reveal Essential and Partially Redundant Functions of Amyloid Precursor Protein Family Members. *The Journal of Neuroscience*, *20*(21), 7951. https://doi.org/10.1523/JNEUROSCI.20-21-07951.2000

Heberden, C. (2017). Sex steroids and neurogenesis. *Biochemical Pharmacology*, *141*, 56–62. https://doi.org/10.1016/J.BCP.2017.05.019

Hennig, B. P., Velten, L., Racke, I., Tu, C. S., Thoms, M., Rybin, V., Besir, H., Remans, K., & Steinmetz, L. M. (2018). Large-scale low-cost NGS library preparation using a robust Tn5 purification and tagmentation protocol. *G3: Genes, Genomes, Genetics*, *8*(1), 79–89. https://doi.org/10.1534/G3.117.300257/-/DC1

Hidalgo, A., Barami, K., Iversen, K., & Goldman, S. A. (1995). Estrogens and non-estrogenic ovarian influences combine to promote the recruitment and decrease the turnover of new neurons in the adult female canary brain. *Journal of Neurobiology*, *27*(4), 470–487. https://doi.org/10.1002/NEU.480270404

Hornix, B. E., Havekes, R., & Kas, M. J. H. (2019). Multisensory cortical processing and dysfunction across the neuropsychiatric spectrum. *Neuroscience & Biobehavioral Reviews*, *97*, 138–151. https://doi.org/10.1016/J.NEUBIOREV.2018.02.010

Huang, Y. S., Carson, J. H., Barbarese, E., & Richter, J. D. (2003). Facilitation of dendritic mRNA transport by CPEB. *Genes & Development*, *17*(5), 638–653. https://doi.org/10.1101/GAD.1053003

Huilgol, D., & Tole, S. (2016). Cell migration in the developing rodent olfactory system. *Cellular and Molecular Life Sciences*, *73*(13), 2467. https://doi.org/10.1007/S00018-016-2172-7

Hus, Y., & Segal, O. (2021). Challenges Surrounding the Diagnosis of Autism in Children. *Neuropsychiatric Disease and Treatment*, *17*, 3509. https://doi.org/10.2147/NDT.S282569

Iossifov, I., O'Roak, B. J., Sanders, S. J., Ronemus, M., Krumm, N., Levy, D., Stessman, H. A., Witherspoon, K. T., Vives, L., Patterson, K. E., Smith, J. D., Paeper, B., Nickerson, D. A., Dea, J., Dong, S., Gonzalez, L. E., Mandell, J. D., Mane, S. M., Murtha, M. T., … Wigler, M. (2014). The contribution of de novo coding mutations to autism spectrum disorder. *Nature*, *515*(7526), 216. https://doi.org/10.1038/NATURE13908

Irimia, M., Weatheritt, R. J., Ellis, J. D., Parikshak, N. N., Gonatopoulos-Pournatzis, T., Babor, M., Quesnel-Vallières, M., Tapial, J., Raj, B., O'Hanlon, D., Barrios-Rodiles, M., Sternberg, M. J. E., Cordes, S. P., Roth, F. P., Wrana, J. L., Geschwind, D. H., & Blencowe, B. J. (2014). A highly conserved program of neuronal microexons is

misregulated in autistic brains. *Cell*, *159*(7), 1511. https://doi.org/10.1016/J.CELL.2014.11.035

Ismail, F. Y., & Shapiro, B. K. (2019). What are neurodevelopmental disorders? *Current Opinion in Neurology*, *32*(4), 611–616. https://doi.org/10.1097/WCO.0000000000000710

Ivshina, M., Lasko, P., & Richter, J. D. (2014). Cytoplasmic polyadenylation element binding proteins in development, health, and disease. *Annual Review of Cell and Developmental Biology*, *30*, 393–415. https://doi.org/10.1146/annurev-cellbio-101011-155831

Jacq, A., Becquet, D., Boyer, B., Guillen, S., Bello-Goutierrez, M.-M., Blanchard, M.-P., Villard, C., Belghazi, M., Torres, M., Franc, J.-L., François-Bellan, A.-M., & Crossover, C. (2020). A Sequence determinant in 3'UTR of mRNAs for Nuclear Retention by Paraspeckles. *BioRxiv*, 2020.07.19.206417. https://doi.org/10.1101/2020.07.19.206417

Ji, Z., & Tian, B. (2009). Reprogramming of 3' untranslated regions of mRNAs by alternative polyadenylation in generation of pluripotent stem cells from different cell types. *PloS One*, *4*(12). https://doi.org/10.1371/JOURNAL.PONE.0008419

Jung, H. J., Shin, I. S., & Lee, J. E. (2019). Olfactory function in mild cognitive impairment and Alzheimer's disease: A meta-analysis. *The Laryngoscope*, *129*(2), 362–369. https://doi.org/10.1002/LARY.27399

Jurkowski, M. P., Bettio, L., K. Woo, E., Patten, A., Yau, S. Y., & Gil-Mohapel, J. (2020). Beyond the Hippocampus and the SVZ: Adult Neurogenesis Throughout the Brain. *Frontiers in Cellular Neuroscience*, *14*, 293. https://doi.org/10.3389/FNCEL.2020.576444/BIBTEX

Kaitz, M., Good, A., Rokem, A. M., & Eidelman, A. I. (1987). Mothers' recognition of their newborns by olfactory cues. *Developmental Psychobiology*, *20*(6), 587–591. https://doi.org/10.1002/DEV.420200604

Kalamakis, G., Brüne, D., Ravichandran, S., Bolz, J., Fan, W., Ziebell, F., Stiehl, T., Catalá-Martinez, F., Kupke, J., Zhao, S., Llorens-Bobadilla, E., Bauer, K., Limpert, S., Berger, B., Christen, U., Schmezer, P., Mallm, J. P., Berninger, B., Anders, S., … Martin-Villalba, A. (2019). Quiescence Modulates Stem Cell Maintenance and Regenerative Capacity in the Aging Brain. *Cell*, *176*(6), 1407-1419.e14. https://doi.org/10.1016/j.cell.2019.01.040

Katoh, H., Shibata, S., Fukuda, K., Sato, M., Satoh, E., Nagoshi, N., Minematsu, T., Matsuzaki, Y., Akazawa, C., Toyama, Y., Nakamura, M., & Okano, H. (2011). The dual origin of the peripheral olfactory system: Placode and neural crest. *Molecular Brain*, *4*(1), 1–16. https://doi.org/10.1186/1756-6606-4-34/FIGURES/5

Kaufmann, I., Martin, G., Friedlein, A., Langen, H., & Keller, W. (2004). Human Fip1 is a subunit of CPSF that binds to U-rich RNA elements and stimulates poly(A) polymerase. *The EMBO Journal*, *23*(3), 616–626. https://doi.org/10.1038/SJ.EMBOJ.7600070

Kazdoba, T. M., Leach, P. T., Yang, M., Silverman, J. L., Solomon, M., & Crawley, J. N. (2016). Translational Mouse Models of Autism: Advancing Toward Pharmacological Therapeutics. *Current Topics in Behavioral Neurosciences*, *28*, 1–52. https://doi.org/10.1007/7854_2015_5003

Kendrick, K. M., Da Costa, A. P. C., Broad, K. D., Ohkura, S., Guevara, R., Lévy, F., & Keverne, E. B. (1997). Neural Control of Maternal Behaviour and Olfactory Recognition of Offspring. *Brain Research Bulletin*, *44*(4), 383–395. https://doi.org/10.1016/S0361-9230(97)00218-9

Kenworthy, L., Case, L., Harms, M. B., Martin, A., & Wallace, G. L. (2010). Adaptive Behavior Ratings Correlate with Symptomatology and IQ Among Individuals with High-Functioning Autism Spectrum Disorders. *Journal of Autism and Developmental Disorders*, *40*(4), 416. https://doi.org/10.1007/S10803-009-0911-4

Khan, N. Z., Gallo, L. A., Arghir, A., Budisteanu, B., Budisteanu, M., Dobrescu, I., Donald, K., El-Tabari, S., Hoogenhout, M., Kalambayi, F., Kawa, R., Espinoza, I. L., Lowenthal, R., Malcolm-Smith, S., Montiel-Nava, C., Odeh, J., de Paula, C. S., Rad, F., Tarpan, A. K., … Elsabbagh, M. (2012). Autism and the Grand Challenges in Global Mental Health. *Autism Research*, *5*(3), 156–159. https://doi.org/10.1002/AUR.1239

Khurshid, K., Crow, A. J. D., Rupert, P. E., Minniti, N. L., Carswell, M. A., Mechanic-Hamilton, D. J., Kamath, V., Doty, R. L., Moberg, P. J., & Roalf, D. R. (2019). A Quantitative Meta-analysis of Olfactory Dysfunction in Epilepsy. *Neuropsychology Review*, *29*(3), 328. https://doi.org/10.1007/S11065-019-09406-7

Kirkovski, M., Enticott, P. G., & Fitzgerald, P. B. (2013). A Review of the Role of Female Gender in Autism Spectrum Disorders. *Journal of Autism and Developmental Disorders 2013 43:11*, *43*(11), 2584–2603. https://doi.org/10.1007/S10803-013-1811-1

Klingler, E. (2017). Development and Organization of the Evolutionarily Conserved Three-Layered Olfactory Cortex. *ENeuro*, *4*(1). https://doi.org/10.1523/ENEURO.0193-16.2016

Kogan, M. D., Vladutiu, C. J., Schieve, L. A., Ghandour, R. M., Blumberg, S. J., Zablotsky, B., Perrin, J. M., Shattuck, P., Kuhlthau, K. A., Harwood, R. L., & Lu, M. C. (2018). The prevalence of parent-reported autism spectrum disorder among US children. *Pediatrics*, *142*(6), 20174161. https://doi.org/10.1542/PEDS.2017-4161/-/DCSUPPLEMENTAL

Lam, M., Moslem, M., Bryois, J., Pronk, R. J., Uhlin, E., Ellström, I. D., Laan, L., Olive, J., Morse, R., Rönnholm, H., Louhivuori, L., Korol, S. V, Dahl, N., Uhlén, P., Anderlid, B.-M., Kele, M., Sullivan, P. F., & Falk, A. (2019). Single cell analysis of autism patient with bi-allelic NRXN1-alpha deletion reveals skewed fate choice in neural progenitors and impaired neuronal functionality. *Experimental Cell Research*, *383*(1), 111469. https://doi.org/10.1016/j.yexcr.2019.06.014

Lau, B. W. M., Yau, S. Y., Lee, T. M. C., Ching, Y. P., Tang, S. W., & So, K. F. (2011). Effect of Corticosterone and Paroxetine on Masculine Mating Behavior: Possible Involvement of Neurogenesis. *The Journal of Sexual Medicine*, *8*(5), 1390–1403. https://doi.org/10.1111/J.1743-6109.2010.02081.X

Lenroot, R. K., & Yeung, P. K. (2013). Heterogeneity within autism spectrum disorders: What have we learned from neuroimaging studies? *Frontiers in Human Neuroscience*, *0*(OCT), 733. https://doi.org/10.3389/FNHUM.2013.00733/BIBTEX

Li, W. L., Chu, M. W., Wu, A., Suzuki, Y., Imayoshi, I., & Komiyama, T. (2018). Adult-born neurons facilitate olfactory bulb pattern separation during task engagement. *ELife*, *7*. https://doi.org/10.7554/ELIFE.33006

Lim, D. A., & Alvarez-Buylla, A. (2016). The Adult Ventricular–Subventricular Zone (V-SVZ) and Olfactory Bulb (OB) Neurogenesis. *Cold Spring Harbor Perspectives in Biology*, *8*(5), a018820. https://doi.org/10.1101/CSHPERSPECT.A018820

Lim, M., Carollo, A., Dimitriou, D., & Esposito, G. (2022). Recent Developments in Autism Genetic Research: A Scientometric Review from 2018 to 2022. *Genes*, *13*(9), 1646. https://doi.org/10.3390/GENES13091646

Llorens-Bobadilla, E., Zhao, S., Baser, A., Saiz-Castro, G., Zwadlo, K., & Martin-Villalba, A. (2015). Single-Cell Transcriptomics Reveals a Population of Dormant Neural Stem Cells that Become Activated upon Brain Injury. *Cell Stem Cell*, *17*(3), 329–340. https://doi.org/10.1016/J.STEM.2015.07.002

Lo, L. H. Y., & Lai, K. O. (2020). Dysregulation of protein synthesis and dendritic spine morphogenesis in ASD: Studies in human pluripotent stem cells. *Molecular Autism*, *11*(1), 1–9. https://doi.org/10.1186/S13229-020-00349-Y/METRICS

Logan, D. W., Brunet, L. J., Webb, W. R., Cutforth, T., Ngai, J., & Stowers, L. (2012).

Learned Recognition of Maternal Signature Odors Mediates the First Suckling Episode in Mice. *Current Biology*, *22*(21), 1998–2007. https://doi.org/10.1016/J.CUB.2012.08.041

Lois, C., & Alvarez-Buylla, A. (1994). Long-Distance Neuronal Migration in the Adult Mammalian Brain. *Science*, *264*(5162), 1145–1148. https://doi.org/10.1126/SCIENCE.8178174

Loomes, R., Hull, L., & Mandy, W. P. L. (2017). What Is the Male-to-Female Ratio in Autism Spectrum Disorder? A Systematic Review and Meta-Analysis. *Journal of the American Academy of Child and Adolescent Psychiatry*, *56*(6), 466–474. https://doi.org/10.1016/J.JAAC.2017.03.013

López De Silanes, I., Paz Quesada, M., & Esteller, M. (2007). Aberrant regulation of messenger RNA 3'-untranslated region in human cancer. *Cellular Oncology : The Official Journal of the International Society for Cellular Oncology*, *29*(1), 1–17. https://doi.org/10.1155/2007/586139

Lord, C., Elsabbagh, M., Baird, G., & Veenstra-Vanderweele, J. (2018). Autism spectrum disorder. *Lancet (London, England)*, *392*(10146), 508. https://doi.org/10.1016/S0140-6736(18)31129-2

Lu, J. X., Wang, Y., Zhang, Y. J., Shen, M. F., Li, H. Y., Yu, Z. Q., & Chen, G. (2021). Axonal mRNA localization and local translation in neurodegenerative disease. *Neural Regeneration Research*, *16*(10), 1950–1957. https://doi.org/10.4103/1673-5374.308074

Lyons-Warren, A. M., Herman, I., Hunt, P. J., & Arenkiel, B. (2021). A systematic-review of olfactory deficits in neurodevelopmental disorders: from mouse to human. *Neuroscience and Biobehavioral Reviews*, *125*, 110. https://doi.org/10.1016/J.NEUBIOREV.2021.02.024

MacDonald, C. C., Wilusz, J., & Shenk, T. (1994). The 64-kilodalton subunit of the CstF polyadenylation factor binds to pre-mRNAs downstream of the cleavage site and influences cleavage site location. *Molecular and Cellular Biology*, *14*(10), 6647–6654. https://doi.org/10.1128/MCB.14.10.6647-6654.1994

Maffezzini, C., Calvo-Garrido, J., Wredenberg, A., & Freyer, C. (2020). Metabolic regulation of neurodifferentiation in the adult brain. *Cellular and Molecular Life Sciences : CMLS*, *77*(13), 2483–2496. https://doi.org/10.1007/s00018-019-03430-9

Malvaut, S., & Saghatelyan, A. (2016). The role of adult-born neurons in the constantly changing olfactory bulb network. *Neural Plasticity*, *2016*. https://doi.org/10.1155/2016/1614329

Mandel, C. R., Kaneko, S., Zhang, H., Gebauer, D., Vethantham, V., Manley, J. L., & Tong, L. (2006). Polyadenylation factor CPSF-73 is the pre-mRNA 3'-end-processing endonuclease. *Nature*, *444*(7121), 953–956. https://doi.org/10.1038/NATURE05363

Martynoga, B., Mateo, J. L., Zhou, B., Andersen, J., Achimastou, A., Urbán, N., van den Berg, D., Georgopoulou, D., Hadjur, S., Wittbrodt, J., Ettwiller, L., Piper, M., Gronostajski, R. M., & Guillemot, F. (2013). Epigenomic enhancer annotation reveals a key role for NFIX in neural stem cell quiescence. *Genes & Development*, *27*(16), 1769–1786. https://doi.org/10.1101/gad.216804.113

Matoulkova, E., Michalova, E., Vojtesek, B., & Hrstka, R. (2012). The role of the 3' untranslated region in post-transcriptional regulation of protein expression in mammalian cells. *RNA Biology*, *9*(5), 563–576. https://doi.org/10.4161/RNA.20231

Mayr, C. (2019). What Are 3′ UTRs Doing? *Cold Spring Harbor Perspectives in Biology*, *11*(10). https://doi.org/10.1101/CSHPERSPECT.A034728

Mayr, C., & Bartel, D. P. (2009). Widespread shortening of 3'UTRs by alternative cleavage and polyadenylation activates oncogenes in cancer cells. *Cell*, *138*(4), 673–684. https://doi.org/10.1016/j.cell.2009.06.016

McCarthy, M. T., & O'Callaghan, C. A. (2014). PeaKDEck: a kernel density estimator-based peak calling program for DNaseI-seq data. *Bioinformatics*, *30*(9), 1302–1304. https://doi.org/10.1093/BIOINFORMATICS/BTT774

McFadden, K., & Minshew, N. J. (2013). Evidence for dysregulation of axonal growth and guidance in the etiology of ASD. *Frontiers in Human Neuroscience*, *7*(OCT). https://doi.org/10.3389/FNHUM.2013.00671

Mefford, H. C., Batshaw, M. L., & Hoffman, E. P. (2012). Genomics, Intellectual Disability, and Autism. *The New England Journal of Medicine*, *366*(8), 733. https://doi.org/10.1056/NEJMRA1114194

Merks, J. H. M., Özgen, H. M., Koster, J., Zwinderman, A. H., Caron, H. N., & Hennekam, R. C. M. (2008). Prevalence and Patterns of Morphological Abnormalities in Patients With Childhood Cancer. *JAMA*, *299*(1), 61–69. https://doi.org/10.1001/JAMA.2007.66

Miles, J. H., & Hillman, R. E. (2000). Value of a clinical morphology examination in autism. *American Journal of Medical Genetics*, *91*(4), 245–253.

Miles, J. H., Takahashi, T. N., Bagby, S., Sahota, P. K., Vaslow, D. F., Wang, C. H., Hillman, R. E., & Farmer, J. E. (2005). Essential versus complex autism: definition of fundamental prognostic subtypes. *American Journal of Medical Genetics. Part A*, *135*(2), 171–180. https://doi.org/10.1002/AJMG.A.30590

Ming, G. L., & Song, H. (2005). ADULT NEUROGENESIS IN THE MAMMALIAN CENTRAL NERVOUS SYSTEM. *Http://Dx.Doi.Org/10.1146/Annurev.Neuro.28.051804.101459*, *28*, 223–250. https://doi.org/10.1146/ANNUREV.NEURO.28.051804.101459

Mira, H., Andreu, Z., Suh, H., Lie, D. C., Jessberger, S., Consiglio, A., San Emeterio, J., Hortigüela, R., Marqués-Torrejón, M. A., Nakashima, K., Colak, D., Götz, M., Fariñas, I., & Gage, F. H. (2010). Signaling through BMPR-IA regulates quiescence and long-term activity of neural stem cells in the adult hippocampus. *Cell Stem Cell*, *7*(1), 78–89. https://doi.org/10.1016/j.stem.2010.04.016

Mirzadeh, Z., Doetsch, F., Sawamoto, K., Wichterle, H., & Alvarez-Buylla, A. (2010). The Subventricular Zone En-face: Wholemount Staining and Ependymal Flow. *Journal of Visualized Experiments : JoVE*, *39*, 1938. https://doi.org/10.3791/1938

Miura, P., Sanfilippo, P., Shenker, S., & Lai, E. C. (2014). Alternative polyadenylation in the nervous system: to what lengths will 3' UTR extensions take us? *BioEssays : News and Reviews in Molecular, Cellular and Developmental Biology*, *36*(8), 766–777. https://doi.org/10.1002/bies.201300174

Miura, P., Shenker, S., Andreu-Agullo, C., Westholm, J. O., & Lai, E. C. (2013). Widespread and extensive lengthening of 3' UTRs in the mammalian brain. *Genome Research*, *23*(5), 812–825. https://doi.org/10.1101/gr.146886.112

Mombaerts, P. (1999). Seven transmembrane proteins as odorant and chemosensory receptors. *Science*, *286*(5440), 707–711. https://doi.org/10.1126/SCIENCE.286.5440.707/ASSET/D88D3712-B7DC-401F-915A-C38BF61713F0/ASSETS/GRAPHIC/SE4197930001.JPEG

Mombaerts, P. (2001). How smell develops. *Nature Neuroscience 2001 4:11*, *4*(11), 1192–1198. https://doi.org/10.1038/nn751

MOSIMANN, J. E. (1962). On the compound multinomial distribution, the multivariate β-distribution, and correlations among proportions. *Biometrika*, *49*(1–2), 65–82. https://doi.org/10.1093/BIOMET/49.1-2.65

Moss, J., Richards, C., Nelson, L., & Oliver, C. (2013). Prevalence of autism spectrum disorder symptomatology and related behavioural characteristics in individuals with Down syndrome. *Autism : The International Journal of Research and Practice*, *17*(4), 390–404. https://doi.org/10.1177/1362361312442790

Mullin, A. P., Gokhale, A., Moreno-De-Luca, A., Sanyal, S., Waddington, J. L., & Faundez, V. (2013). Neurodevelopmental disorders: mechanisms and boundary definitions from genomes, interactomes and proteomes. *Translational Psychiatry*, *3*(12), e329. https://doi.org/10.1038/TP.2013.108

Myers, S. M., Johnson, C. P., Lipkin, P. H., Cartwright, J. D., Desch, L. W., Duby, J. C., Elias, E. R., Levey, E. B., Liptak, G. S., Murphy, N. A., Tilton, A. H., Lollar, D., Macias, M., McPherson, M., Olson, D. G., Strickland, B., Skipper, S. M., Ackermann, J., Del Monte, M., … Yeargin-Allsopp, M. (2007). Management of children with autism spectrum disorders. *Pediatrics*, *120*(5), 1162–1182. https://doi.org/10.1542/PEDS.2007-2362

Nabil, W. N. N., Xi, Z., Song, Z., Jin, L., Zhang, X. D., Zhou, H., De Souza, P., Dong, Q., & Xu, H. (2021). Towards a Framework for Better Understanding of Quiescent Cancer Cells. *Cells*, *10*(3), 1–19. https://doi.org/10.3390/CELLS10030562

Nakai, N., Takumi, T., Nakai, J., & Sato, M. (2018). Common defects of spine dynamics and circuit function in neurodevelopmental disorders: A systematic review of findings from in vivo optical imaging of mouse models. *Frontiers in Neuroscience*, *12*(JUN), 412. https://doi.org/10.3389/FNINS.2018.00412/BIBTEX

O'Reilly, C., Lewis, J. D., & Elsabbagh, M. (2017). Is functional brain connectivity atypical in autism? A systematic review of EEG and MEG studies. *PLoS ONE*, *12*(5). https://doi.org/10.1371/JOURNAL.PONE.0175870

O ' Roak, B. J., Vives, L., Girirajan, S., Karakoc, E., Krumm, N., Coe, B. P., Levy, R., Ko, A., Lee, C., Smith, J. D., Turner, E. H., Stanaway, I. B., Vernot, B., Malig, M., Baker, C., Akey, J. M., Borenstein, E., Rieder, M. J., Nickerson, D. A., … Eichler, E. E. (2012). Sporadic autism exomes reveal a highly interconnected protein network of de novo mutations. *Nature*, *485*(7397), 246–250. https://doi.org/10.1038/NATURE10989

Ohja, K., Gozal, E., Fahnestock, M., Cai, L., Cai, J., Freedman, J. H., Switala, A., El-Baz, A., & Barnes, G. N. (2018). Neuroimmunologic and Neurotrophic Interactions in Autism Spectrum Disorders: Relationship to Neuroinflammation. *NeuroMolecular Medicine 2018 20:2*, *20*(2), 161–173. https://doi.org/10.1007/S12017-018-8488-8

Okumura, T., Kumazaki, H., Singh, A. K., Touhara, K., & Okamoto, M. (2020). Individuals With Autism Spectrum Disorder Show Altered Event-Related Potentials in the Late Stages of Olfactory Processing. *Chemical Senses*, *45*(1), 37–44. https://doi.org/10.1093/CHEMSE/BJZ070

Ozgen, H., Hellemann, G. S., De Jonge, M. V., Beemer, F. A., & Van Engeland, H. (2013). Predictive Value of Morphological Features in Patients with Autism versus Normal Controls. *Journal of Autism and Developmental Disorders*, *43*(1), 147. https://doi.org/10.1007/S10803-012-1554-4

Ozgen, H., Hop, J. W., Hox, J. J., Beemer, F. A., & Van Engeland, H. (2010). Minor physical anomalies in autism: a meta-analysis. *Molecular Psychiatry*, *15*(3), 300–307. https://doi.org/10.1038/MP.2008.75

Ozgen, Heval, Hellemann, G. S., Stellato, R. K., Lahuis, B., Van Daalen, E., Staal, W. G., Rozendal, M., Hennekam, R. C., Beemer, F. A., & Van Engeland, H. (2011). Morphological Features in Children with Autism Spectrum Disorders: A Matched Case–Control Study. *Journal of Autism and Developmental Disorders*, *41*(1), 23. https://doi.org/10.1007/S10803-010-1018-7

Pan, Y. H., Wu, N., & Yuan, X. B. (2019). Toward a Better Understanding of Neuronal Migration Deficits in Autism Spectrum Disorders. *Frontiers in Cell and Developmental Biology*, *7*(SEP). https://doi.org/10.3389/FCELL.2019.00205

Pangrazzi, L., Balasco, L., & Bozzi, Y. (2020). Oxidative Stress and Immune System Dysfunction in Autism Spectrum Disorders. *International Journal of Molecular*

*Sciences*, *21*(9), 3293. https://doi.org/10.3390/IJMS21093293
Parikshak, N. N., Luo, R., Zhang, A., Won, H., Lowe, J. K., Chandran, V., Horvath, S., & Geschwind, D. H. (2013). Integrative functional genomic analyses implicate specific molecular pathways and circuits in autism. *Cell*, *155*(5), 1008. https://doi.org/10.1016/J.CELL.2013.10.031
Parras, A., Anta, H., Santos-Galindo, M., Swarup, V., Elorza, A., Nieto-González, J. L., Picó, S., Hernández, I. H., Díaz-Hernández, J. I., Belloc, E., Rodolosse, A., Parikshak, N. N., Peñagarikano, O., Fernández-Chacón, R., Irimia, M., Navarro, P., Geschwind, D. H., Méndez, R., & Lucas, J. J. (2018a). Autism-like phenotype and risk gene mRNA deadenylation by CPEB4 mis-splicing. *Nature*, *560*(7719), 441–446. https://doi.org/10.1038/s41586-018-0423-5
Parras, A., Anta, H., Santos-Galindo, M., Swarup, V., Elorza, A., Nieto-González, J. L., Picó, S., Hernández, I. H., Díaz-Hernández, J. I., Belloc, E., Rodolosse, A., Parikshak, N. N., Peñagarikano, O., Fernández-Chacón, R., Irimia, M., Navarro, P., Geschwind, D. H., Méndez, R., & Lucas, J. J. (2018b). Autism-like phenotype and risk gene-RNA deadenylation by CPEB4 mis-splicing. *Nature*, *560*(7719), 441. https://doi.org/10.1038/S41586-018-0423-5
Pereira-Castro, I., & Moreira, A. (2021). On the function and relevance of alternative 3′-UTRs in gene expression regulation. *Wiley Interdisciplinary Reviews: RNA*, *12*(5), e1653. https://doi.org/10.1002/WRNA.1653
Peters, J. M., Taquet, M., Vega, C., Jeste, S. S., Fernández, I. S., Tan, J., Nelson, C. A., Sahin, M., & Warfield, S. K. (2013). Brain functional networks in syndromic and non-syndromic autism: a graph theoretical study of EEG connectivity. *BMC Medicine*, *11*(1), 54. https://doi.org/10.1186/1741-7015-11-54
Phan, T. G., & Croucher, P. I. (2020). The dormant cancer cell life cycle. *Nature Reviews. Cancer*, *20*(7), 398–411. https://doi.org/10.1038/S41568-020-0263-0
Pinto, D., Delaby, E., Merico, D., Barbosa, M., Merikangas, A., Klei, L., Thiruvahindrapuram, B., Xu, X., Ziman, R., Wang, Z., Vorstman, J. A. S., Thompson, A., Regan, R., Pilorge, M., Pellecchia, G., Pagnamenta, A. T., Oliveira, B., Marshall, C. R., Magalhaes, T. R., … Scherer, S. W. (2014). Convergence of genes and cellular pathways dysregulated in autism spectrum disorders. *American Journal of Human Genetics*, *94*(5), 677–694. https://doi.org/10.1016/J.AJHG.2014.03.018
Piqué, M., López, J. M., Foissac, S., Guigó, R., & Méndez, R. (2008). A combinatorial code for CPE-mediated translational control. *Cell*, *132*(3), 434–448. https://doi.org/10.1016/j.cell.2007.12.038
Poliak, S., Morales, D., Croteau, L. P., Krawchuk, D., Palmesino, E., Morton, S., Cloutier, J. F., Charron, F., Dalva, M. B., Ackerman, S. L., Kao, T. J., & Kania, A. (2015). Synergistic integration of Netrin and ephrin axon guidance signals by spinal motor neurons. *ELife*, *4*(DECEMBER2015). https://doi.org/10.7554/ELIFE.10841
Ponti, G., Farinetti, A., Marraudino, M., Panzica, G. C., & Gotti, S. (2018). Sex Steroids and Adult Neurogenesis in the Ventricular-Subventricular Zone. *Frontiers in Endocrinology*, *9*(APR), 156. https://doi.org/10.3389/FENDO.2018.00156
Proudfoot, N. J., & Brownlee, G. G. (1976). 3’ non-coding region sequences in eukaryotic messenger RNA. *Nature*, *263*(5574), 211–214. https://doi.org/10.1038/263211A0
Raznahan, A., Wallace, G. L., Antezana, L., Greenstein, D., Lenroot, R., Thurm, A., Gozzi, M., Spence, S., Martin, A., Swedo, S. E., & Giedd, J. N. (2013). Compared to what? Early brain overgrowth in autism and the perils of population norms. *Biological Psychiatry*, *74*(8), 563–575. https://doi.org/10.1016/j.biopsych.2013.03.022
Reynolds, B. A., & Weiss, S. (1992). Generation of Neurons and Astrocytes from Isolated Cells of the Adult Mammalian Central Nervous System. *Science*, *255*(5052), 1707–

1710. https://doi.org/10.1126/SCIENCE.1553558
Richards, L. J., Kilpatrick, T. J., & Bartlett, P. F. (1992). De novo generation of neuronal cells from the adult mouse brain. *Proceedings of the National Academy of Sciences of the United States of America*, *89*(18), 8591. https://doi.org/10.1073/PNAS.89.18.8591
Richter, J. D. (2007). CPEB: a life in translation. *Trends in Biochemical Sciences*, *32*(6), 279–285. https://doi.org/10.1016/j.tibs.2007.04.004
Robinson, E. B., St Pourcain, B., Anttila, V., Kosmicki, J. A., Bulik-Sullivan, B., Grove, J., Maller, J., Samocha, K. E., Sanders, S. J., Ripke, S., Martin, J., Hollegaard, M. V, Werge, T., Hougaard, D. M., Neale, B. M., Evans, D. M., Skuse, D., Mortensen, P. B., Børglum, A. D., … Daly, M. J. (2016). Genetic risk for autism spectrum disorders and neuropsychiatric variation in the general population. *Nature Genetics*, *48*(5), 552–555. https://doi.org/10.1038/ng.3529
Rodemer, W., Gallo, G., & Selzer, M. E. (2020). Mechanisms of Axon Elongation Following CNS Injury: What Is Happening at the Axon Tip? *Frontiers in Cellular Neuroscience*, *14*, 177. https://doi.org/10.3389/FNCEL.2020.00177/XML/NLM
Rubenstein, J. L. R., & Merzenich, M. M. (2003). Model of autism: increased ratio of excitation/inhibition in key neural systems. *Genes, Brain, and Behavior*, *2*(5), 255. https://doi.org/10.1034/J.1601-183X.2003.00037.X
Sahin, M., Jones, S. R., Sweeney, J. A., Berry-Kravis, E., Connors, B. W., Ewen, J. B., Hartman, A. L., Levin, A. R., Potter, W. Z., & Mamounas, L. A. (2018). Discovering translational biomarkers in neurodevelopmental disorders. *Nature Reviews Drug Discovery 2021 18:4*, *18*(4), 235–236. https://doi.org/10.1038/d41573-018-00010-7
Sandberg, R., Neilson, J. R., Sarma, A., Sharp, P. A., & Burge, C. B. (2008). Proliferating cells express mRNAs with shortened 3' untranslated regions and fewer microRNA target sites. *Science (New York, N.Y.)*, *320*(5883), 1643–1647. https://doi.org/10.1126/science.1155390
Sanders, S. J., Murtha, M. T., Gupta, A. R., Murdoch, J. D., Raubeson, M. J., Willsey, A. J., Ercan-Sencicek, A. G., Di Lullo, N. M., Parikshak, N. N., Stein, J. L., Walker, M. F., Ober, G. T., Teran, N. A., Song, Y., El-Fishawy, P., Murtha, R. C., Choi, M., Overton, J. D., Bjornson, R. D., … State, M. W. (2012). De novo mutations revealed by whole exome sequencing are strongly associated with autism. *Nature*, *485*(7397), 237. https://doi.org/10.1038/NATURE10945
Sanes, J. R., & Zipursky, S. L. (2020). Synaptic Specificity, Recognition Molecules, and Assembly of Neural Circuits. *Cell*, *181*(3), 536–556. https://doi.org/10.1016/J.CELL.2020.04.008
Satterstrom, F. K., Kosmicki, J. A., Wang, J., Breen, M. S., De Rubeis, S., An, J. Y., Peng, M., Collins, R., Grove, J., Klei, L., Stevens, C., Reichert, J., Mulhern, M. S., Artomov, M., Gerges, S., Sheppard, B., Xu, X., Bhaduri, A., Norman, U., … Buxbaum, J. D. (2020). Large-Scale Exome Sequencing Study Implicates Both Developmental and Functional Changes in the Neurobiology of Autism. *Cell*, *180*(3), 568. https://doi.org/10.1016/J.CELL.2019.12.036
Scandurra, V., Emberti Gialloreti, L., Barbanera, F., Scordo, M. R., Pierini, A., & Canitano, R. (2019). Neurodevelopmental Disorders and Adaptive Functions: A Study of Children With Autism Spectrum Disorders (ASD) and/or Attention Deficit and Hyperactivity Disorder (ADHD). *Frontiers in Psychiatry*, *10*, 673. https://doi.org/10.3389/FPSYT.2019.00673/BIBTEX
Schäfer, P., Tüting, C., Schönemann, L., Kühn, U., Treiber, T., Treiber, N., Ihling, C., Graber, A., Keller, W., Meister, G., Sinz, A., & Wahle, E. (2018). Reconstitution of mammalian cleavage factor II involved in 3' processing of mRNA precursors. *RNA (New York, N.Y.)*, *24*(12), 1721–1737. https://doi.org/10.1261/RNA.068056.118

Schilling, S., Mehr, A., Ludewig, S., Stephan, jonathan, Zimmermann, M., August, A., Strecker, P., Korte, M., Koo, E. H., Müller, U. C., Kins, S., & Eggert, S. (2017). APLP1 Is a Synaptic Cell Adhesion Molecule, Supporting Maintenance of Dendritic Spines and Basal Synaptic Transmission. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *37*(21), 5345–5365. https://doi.org/10.1523/JNEUROSCI.1875-16.2017

Seng, C., Luo, W., & Földy, C. (2022). Circuit formation in the adult brain. *European Journal of Neuroscience*. https://doi.org/10.1111/EJN.15742

Shah, A., Mittleman, B. E., Gilad, Y., & Li, Y. I. (2021). Benchmarking sequencing methods and tools that facilitate the study of alternative polyadenylation. *Genome Biology*, *22*(1), 1–21. https://doi.org/10.1186/S13059-021-02502-Z/FIGURES/6

Shao, Z., Liu, S., Zhou, F., Puche, A. C., & Shipley, M. T. (2019). Reciprocal Inhibitory Glomerular Circuits Contribute to Excitation–Inhibition Balance in the Mouse Olfactory Bulb. *ENeuro*, *6*(3). https://doi.org/10.1523/ENEURO.0048-19.2019

Shi, Y., Di Giammartino, D. C., Taylor, D., Sarkeshik, A., Rice, W. J., Yates, J. R., Frank, J., & Manley, J. L. (2009). Molecular architecture of the human pre-mRNA 3' processing complex. *Molecular Cell*, *33*(3), 365–376. https://doi.org/10.1016/J.MOLCEL.2008.12.028

Smear, M., Resulaj, A., Zhang, J., Bozza, T., & Rinberg, D. (2013). Multiple perceptible signals from a single olfactory glomerulus. *Nature Neuroscience 2013 16:11*, *16*(11), 1687–1691. https://doi.org/10.1038/nn.3519

Soucy, E. R., Albeanu, D. F., Fantana, A. L., Murthy, V. N., & Meister, M. (2009). Precision and diversity in an odor map on the olfactory bulb. *Nature Neuroscience 2009 12:2*, *12*(2), 210–220. https://doi.org/10.1038/nn.2262

Subramanian, A., Tamayo, P., Mootha, V. K., Mukherjee, S., Ebert, B. L., Gillette, M. A., Paulovich, A., Pomeroy, S. L., Golub, T. R., Lander, E. S., & Mesirov, J. P. (2005). Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles. *Proceedings of the National Academy of Sciences of the United States of America*, *102*(43), 15545–15550. https://doi.org/10.1073/PNAS.0506580102/SUPPL_FILE/06580FIG7.JPG

Sungur, A. Ö., Jochner, M. C. E., Harb, H., Kılıç, A., Garn, H., Schwarting, R. K. W., & Wöhr, M. (2017). Aberrant cognitive phenotypes and altered hippocampal BDNF expression related to epigenetic modifications in mice lacking the post-synaptic scaffolding protein SHANK1: Implications for autism spectrum disorder. *Hippocampus*, *27*(8), 906–919. https://doi.org/10.1002/hipo.22741

Sungur, A. Ö., Schwarting, R. K. W., & Wöhr, M. (2016). Early communication deficits in the Shank1 knockout mouse model for autism spectrum disorder: Developmental aspects and effects of social context. *Autism Research : Official Journal of the International Society for Autism Research*, *9*(6), 696–709. https://doi.org/10.1002/aur.1564

Szkop, K. J., Cooke, P. I. C., Humphries, J. A., Kalna, V., Moss, D. S., Schuster, E. F., & Nobeli, I. (2017). Dysregulation of alternative poly-adenylation as a potential player in autism spectrum disorder. *Frontiers in Molecular Neuroscience*, *10*, 279. https://doi.org/10.3389/FNMOL.2017.00279/BIBTEX

Tang, P., Yang, Y., Li, G., Huang, L., Wen, M., Ruan, W., Guo, X., Zhang, C., Zuo, X., Luo, D., Xu, Y., Fu, X. D., & Zhou, Y. (2022). Alternative polyadenylation by sequential activation of distal and proximal PolyA sites. *Nature Structural & Molecular Biology 2022 29:1*, *29*(1), 21–31. https://doi.org/10.1038/s41594-021-00709-z

Tian, B., & Manley, J. L. (2017). Alternative polyadenylation of mRNA precursors. *Nature Reviews. Molecular Cell Biology*, *18*(1), 18. https://doi.org/10.1038/NRM.2016.116

Tonacci, A., Billeci, L., Tartarisco, G., Ruta, L., Muratori, F., Pioggia, G., & Gangemi, S. (2017). [Formula: see text]Olfaction in autism spectrum disorders: A systematic review. *Child Neuropsychology : A Journal on Normal and Abnormal Development in Childhood and Adolescence*, *23*(1), 1–25. https://doi.org/10.1080/09297049.2015.1081678

Toth, A. B., Shum, A. K., & Prakriya, M. (2016). Regulation of neurogenesis by calcium signaling. *Cell Calcium*, *59*(2–3), 124–134. https://doi.org/10.1016/j.ceca.2016.02.011

Treloar, H. B., Feinstein, P., Mombaerts, P., & Greer, C. A. (2002). Specificity of Glomerular Targeting by Olfactory Sensory Axons. *Journal of Neuroscience*, *22*(7), 2469–2477. https://doi.org/10.1523/JNEUROSCI.22-07-02469.2002

Tushev, G., Glock, C., Heumüller, M., Biever, A., Jovanovic, M., & Schuman, E. M. (2018). Alternative 3′ UTRs Modify the Localization, Regulatory Potential, Stability, and Plasticity of mRNAs in Neuronal Compartments. *Neuron*, *98*(3), 495-511.e6. https://doi.org/10.1016/J.NEURON.2018.03.030

Tuszynski, M. H., & Steward, O. (2012). Concepts and Methods for the Study of Axonal Regeneration in the CNS. *Neuron*, *74*(5), 777–791. https://doi.org/10.1016/J.NEURON.2012.05.006

Tye, C., & Bolton, P. (2013). Neural connectivity abnormalities in autism: Insights from the Tuberous Sclerosis model. *BMC Medicine*, *11*(1), 55. https://doi.org/10.1186/1741-7015-11-55

Uzunova, G., Pallanti, S., & Hollander, E. (2016). Excitatory/inhibitory imbalance in autism spectrum disorders: Implications for interventions and therapeutics. *The World Journal of Biological Psychiatry : The Official Journal of the World Federation of Societies of Biological Psychiatry*, *17*(3), 174–186. https://doi.org/10.3109/15622975.2015.1085597

Vaglio, S. (2009). Chemical communication and mother-infant recognition. *Communicative & Integrative Biology*, *2*(3), 279. https://doi.org/10.4161/CIB.2.3.8227

van Steensel, F. J. A., Bögels, S. M., & Perrin, S. (2011). Anxiety Disorders in Children and Adolescents with Autistic Spectrum Disorders: A Meta-Analysis. *Clinical Child and Family Psychology Review*, *14*(3), 302. https://doi.org/10.1007/S10567-011-0097-0

Varghese, M., Keshav, N., Jacot-Descombes, S., Warda, T., Wicinski, B., Dickstein, D. L., Harony-Nicolas, H., De Rubeis, S., Drapeau, E., Buxbaum, J. D., & Hof, P. R. (2017). Autism spectrum disorder: Neuropathology and animal models. *Acta Neuropathologica*, *134*(4), 537. https://doi.org/10.1007/S00401-017-1736-4

Velmeshev, D., Schirmer, L., Jung, D., Haeussler, M., Perez, Y., Mayer, S., Bhaduri, A., Goyal, N., Rowitch, D. H., & Kriegstein, A. R. (2019a). Single-cell genomics identifies cell type–specific molecular changes in autism. *Science*, *364*(6441), 685–689. https://doi.org/10.1126/SCIENCE.AAV8130

Velmeshev, D., Schirmer, L., Jung, D., Haeussler, M., Perez, Y., Mayer, S., Bhaduri, A., Goyal, N., Rowitch, D. H., & Kriegstein, A. R. (2019b). Single-cell genomics identifies cell type–specific molecular changes in autism. *Science*, *364*(6441), 685–689. https://doi.org/10.1126/SCIENCE.AAV8130

Velmeshev, D., Schirmer, L., Jung, D., Haeussler, M., Perez, Y., Mayer, S., Bhaduri, A., Goyal, N., Rowitch, D. H., & Kriegstein, A. R. (2019c). Single-cell genomics identifies cell type–specific molecular changes in autism. *Science*, *364*(6441), 685–689. https://doi.org/10.1126/science.aav8130

Vivanti, G., & Dissanayake, C. (2016). Outcome for Children Receiving the Early Start Denver Model Before and After 48 Months. *Journal of Autism and Developmental Disorders*, *46*(7), 2441–2449. https://doi.org/10.1007/S10803-016-2777-6

Vnencak, M., Paul, M. H., Hick, M., Schwarzacher, S. W., Del Turco, D., Müller, U. C., Deller, T., & Jedlicka, P. (2015). Deletion of the amyloid precursor-like protein 1

(APLP1) enhances excitatory synaptic transmission, reduces network inhibition but does not impair synaptic plasticity in the mouse dentate gyrus. *The Journal of Comparative Neurology*, *523*(11), 1717–1729. https://doi.org/10.1002/cne.23766

Volkmar, F. R. (2014). Editorial: the importance of early intervention. *Journal of Autism and Developmental Disorders*, *44*(12), 2979–2980. https://doi.org/10.1007/S10803-014-2265-9

Wang, R., Zheng, D., Yehia, G., & Tian, B. (2018). A compendium of conserved cleavage and polyadenylation events in mammalian genes. *Genome Research*, *28*(10), 1427–1441. https://doi.org/10.1101/GR.237826.118

Wanke, K. A., Devanna, P., & Vernes, S. C. (2018). Understanding Neurodevelopmental Disorders: The Promise of Regulatory Variation in the 3'UTRome. *Biological Psychiatry*, *83*(7), 548–557. https://doi.org/10.1016/J.BIOPSYCH.2017.11.006

Weill, L., Belloc, E., Bava, F.-A., & Méndez, R. (2012). Translational control by changes in poly(A) tail length: recycling mRNAs. *Nature Structural & Molecular Biology*, *19*(6), 577–585. https://doi.org/10.1038/nsmb.2311

Weitlauf, A. S., McPheeters, M. L., Peters, B., Sathe, N., Travis, R., Aiello, R., Williamson, E., Veenstra-VanderWeele, J., Krishnaswami, S., Jerome, R., & Warren, Z. (2014). Therapies for Children With Autism Spectrum Disorder: Behavioral Interventions Update. *AHRQ Publication No. 14-EHC036-EF. Rockville, MD: Agency for Healthcare Research and Quality*, *137*, 120. http://europepmc.org/books/NBK241444

Wheeler, A. C., Mussey, J., Villagomez, A., Bishop, E., Raspa, M., Edwards, A., Bodfish, J., Bann, C., & Bailey, D. B. (2015). DSM-5 changes and the prevalence of parent-reported autism spectrum symptoms in Fragile X syndrome. *Journal of Autism and Developmental Disorders*, *45*(3), 816–829. https://doi.org/10.1007/S10803-014-2246-Z

Wilkins, J., & Matson, J. L. (2009). A comparison of social skills profiles in intellectually disabled adults with and without ASD. *Behavior Modification*, *33*(2), 143–155. https://doi.org/10.1177/0145445508321880

Willsey, A. J., Sanders, S. J., Li, M., Dong, S., Tebbenkamp, A. T., Muhle, R. A., Reilly, S. K., Lin, L., Fertuzinhos, S., Miller, J. A., Murtha, M. T., Bichsel, C., Niu, W., Cotney, J., Ercan-Sencicek, A. G., Gockley, J., Gupta, A. R., Han, W., He, X., … State, M. W. (2013). Coexpression networks implicate human midfetal deep cortical projection neurons in the pathogenesis of autism. *Cell*, *155*(5), 997. https://doi.org/10.1016/J.CELL.2013.10.020

Winberg, J., & Porter, R. H. (1998). Olfaction and human neonatal behaviour: clinical implications. *Acta Pædiatrica*, *87*(1), 6–10. https://doi.org/10.1111/J.1651-2227.1998.TB01376.X

Winner, B., Cooper-Kuhn, C. M., Aigner, R., Winkler, J., & Kuhn, H. G. (2002). Long-term survival and cell death of newly generated neurons in the adult rat olfactory bulb. *European Journal of Neuroscience*, *16*(9), 1681–1689. https://doi.org/10.1046/J.1460-9568.2002.02238.X

Yang, Q., Coseno, M., Gilmartin, G. M., & Doublié, S. (2011). Crystal structure of a human cleavage factor CFI(m)25/CFI(m)68/RNA complex provides an insight into poly(A) site recognition and RNA looping. *Structure (London, England : 1993)*, *19*(3), 368–377. https://doi.org/10.1016/J.STR.2010.12.021

Yang, Q., Gilmartin, G. M., & Doublié, S. (2010). Structural basis of UGUA recognition by the Nudix protein CFI(m)25 and implications for a regulatory role in mRNA 3' processing. *Proceedings of the National Academy of Sciences of the United States of America*, *107*(22), 10062–10067. https://doi.org/10.1073/PNAS.1000848107

Yizhar, O. (2012). Optogenetic Insights into Social Behavior Function. *Biological Psychiatry*, *71*(12), 1075–1080. https://doi.org/10.1016/J.BIOPSYCH.2011.12.029

Yizhar, O., Fenno, L. E., Prigge, M., Schneider, F., Davidson, T. J., Ogshea, D. J., Sohal, V. S., Goshen, I., Finkelstein, J., Paz, J. T., Stehfest, K., Fudim, R., Ramakrishnan, C., Huguenard, J. R., Hegemann, P., & Deisseroth, K. (2011). Neocortical excitation/inhibition balance in information processing and social dysfunction. *Nature*, *477*(7363), 171. https://doi.org/10.1038/NATURE10360

Zaboski, B. A., & Storch, E. A. (2018). Comorbid autism spectrum disorder and anxiety disorders: a brief review. *Future Neurology*, *13*(1), 31. https://doi.org/10.2217/FNL-2017-0030

Zeidán-Chuliá, F., Salmina, A. B., Malinovskaya, N. A., Noda, M., Verkhratsky, A., & Moreira, J. C. F. (2014). The glial perspective of autism spectrum disorders. *Neuroscience & Biobehavioral Reviews*, *38*, 160–172. https://doi.org/10.1016/J.NEUBIOREV.2013.11.008

Zhang, Y., Liu, L., Qiu, Q., Zhou, Q., Ding, J., Lu, Y., & Liu, P. (2021). Alternative polyadenylation: methods, mechanism, function, and role in cancer. *Journal of Experimental & Clinical Cancer Research : CR*, *40*(1). https://doi.org/10.1186/S13046-021-01852-7

Zieger, H. L., & Choquet, D. (2021). Nanoscale synapse organization and dysfunction in neurodevelopmental disorders. *Neurobiology of Disease*, *158*, 105453. https://doi.org/10.1016/J.NBD.2021.105453

Zou, D. J., Feinstein, P., Rivers, A. L., Mathews, G. A., Kim, A., Greer, C. A., Mombaerts, P., & Firestein, S. (2004). Postnatal refinement of peripheral olfactory projections. *Science*, *304*(5679), 1976–1979. https://doi.org/10.1126/SCIENCE.1093468/SUPPL_FILE/ZOU_SOM.PDF

# 6. Appendix

## 6.1. List of Abbreviations

| | |
|---|---|
| 3'UTR | 3' untranslated region |
| ADHD | Attention deficit hyperactivity disorder |
| aNSCs | Active NSCs |
| APA | Alternative polyadenylation |
| APLP1 | Amyloid Beta Precursor Like Protein 1 |
| ASD | Autism Spectrum Disorder |
| AUC | Area under the curve |
| BMP-4 | Bone morphogenic protein-4 |
| bp | Basepair |
| cDNA | Complimentary DNA |
| CFI | Cleavage factor I |
| CFII | Cleavage factor II |
| CNS | Central nervous system |
| Co-IP | Co-immunoprecipitation |
| CPE | Cytoplasmic Polyadenylation Element |
| CPEB4 | Cytoplasmic Polyadenylation Element Binding Protein 4 |
| CRISPR | Clustered regularly interspaced short palindromic repeats |
| CSPF | Cleavage and polyadenylation factor |
| CSTF | Cleavage stimulating factor |
| DG | Dentate gyrus |
| DNA | Deoxy-ribonucleic acid |
| DO | Disease Ontology |
| DSM-5 | Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition |
| E/I | Excitatory/Inhibitory |
| ED | Euclidean distances |
| EGF | Epidermal growth factors |
| EMD | Earth movers' distance |
| EPM | Elevated plus maze |
| EtOH | Ethanol |
| FACS | Fluorescence-activated single cell sorting |
| FDR | False discover rate |
| FGF | Fibroblast growth factors |
| FMO | Fluorescent minus one |
| FSC | Forward scatter |
| GFP | Green fluorescent protein |
| GSEA | Gene set enrichment analysis |
| KI | Knockin |
| KO | Knockout |
| LAMA | Ligand-modulated antibody fragments |
| LM | Linear model |

| | |
|---|---|
| LNA | Locked nucleic acids |
| LR | Linear regression |
| miRNAs | Micro RNAs |
| MNR | Multinomial regression |
| mRNA | Messenger RNA |
| NBs | Neuroblasts |
| NDDs | Neurodevelopmental disorders |
| NES | Normalized enrichment scores |
| NSCs | Neural stem cells |
| NTDK | Neural tissue dissociation kit |
| OB | Olfactory bulb |
| OF | Open field |
| PABPN1 | poly(A) binding protein 1 |
| PAP | poly(A) polymerase |
| PAS | poly(A) signal |
| PBMCs | Peripheral blood mononuclear cells |
| PBS | Phosphate buffer saline |
| PEG | Polyethelyene Glycol |
| PMI | Postmortem interval |
| qNSCs | Quiescent NSCs |
| RBCs | Red blood cells |
| RIN | RNA integrity number |
| RMS | Rostral migratory stream |
| RNA | ribonucleic acid |
| ROC | Receiver operating characteristic |
| RT | Reverse transcription/room temperature |
| SCC | Single colour controls |
| scRNASeq | Single cell RNA sequencing |
| SCT | Social chamber test |
| SI | Social Interaction |
| SSC | Side scatter |
| TAPs | Transit amplifying progenitors |
| TMP | Trimethoprim |
| TSO | Template switching oligo |
| UMI | Unique molecular identifier |
| USVs | Ultrasonic vocalizations |
| V-SVZ | Ventricular subventricular zone |
| WHO | World health organization |
| WT | Wildtype |

## 6.2. List of Figures

## 6.3. List of Tables

## 6.4.Acknowledgement

*I am grateful for this opportunity that was indeed challenging, and my journey thus far would not have been possible without the support and encouragement of mentors, family, friends and colleagues to whom I extend my deepest gratitude.*

*Firstly, I would like to thank Prof. Dr. Ana Martin-Villalba who provided the opportunity to of pursue my PhD in the department of molecular neurobiology at the German Cancer Research Centre (DKFZ). I thank her for the last five and a half years for providing me with her guidance during the pursuit of this project. I thank her for pushing me whenever necessary to be able to keep moving forward and I also thank her for the feedback that she has given me over the years. I am grateful for the funding she provided through the department budget after the first three years of my PhD that allowed me to focus on the research.*

*I am also grateful to my TAC members, Prof. Dr. Hellmut Augustin and Prof. Claudio Joazeiro who provided critical feedback during my thesis advisory committee meetings that helped with the overall project. I also thank them for their encouragement that kept me motivated to pursue my PhD. I am grateful to Prof. Dr. Henrik Kaessmann, my second supervisor for agreeing to be a part of my defense committee. I would also like to thank Prof. Dr. Lazaro Centanín for also being a part of my defence committee and for all the nice discussions during my talk at the SFB 873 seminars. I am grateful to all the member of the defense committee that have agreed to share their valuable time and be a part of my defense.*

*I would like to thank Dr. Manuel Göpferich, my first tandem partner on the project for the bioinformatics work he did. I would like to specially thank Dr. André Lopes Martins Macedo, my current tandem partner on the project, firstly for the bioinformatics analysis that he did during the latter part of the project. Secondly, I really appreciate all the fruitful discussions that helped with the project overall. Thirdly, I thank him for the time he would always take to explain the bioinformatics background and the rationale behind the methods applied that provided me a clearer understanding of the final analysis.*

*I am also grateful to our technicians Aylin Korkmaz & Mohammad Eid Alshukairi who helped with experiments whenever necessary. I also thank my colleagues Dr. Santiago Cerrizuela and Oguzhan Kaya for their support both on a professional and personal level and for the nice moments we had both inside and outside the lab. I would also like to thank Damian, Alena, Leo, Lukas and Andres with whom I've had both, fruitful scientific discussion and general conversations that made my stay in the lab productive and fun. I also thank all other colleagues for creating a friendly atmosphere that helped cope with being far away from home.*

*Finally, I am grateful to my family for the love, encouragement and support they have provided me almost daily despite being far away in India. I am grateful to my parents who supported me when I came to Germany for my masters and continue to do so. Last but not the least, I am grateful to the love of my life, my wife Rhea. She has been a constant source of encouragement that has helped me navigate some of the most challenging times of my life.*