The Characterization of Genetic Risk Factors Associated with Autism

By

Cora Sin

A THESIS SUBMITTED TO THE FACULTY OF GRADUATE STUDIES IN PARTIAL FULFILMENT OF THE REQUIREMENTS FOR THE DEGREE OF MASTER OF SCIENCE

GRADUATE PROGRAM IN KINESIOLOGY AND HEALTH SCIENCE

YORK UNIVERSITY

TORONTO, ONTARIO

OCTOBER, 2012

ABSTRACT

Autism is a severe neurodevelopmental disorder. Development of a molecular diagnostic screen is an imperative step towards personalized treatments. Gene expression profiling using buccal samples was employed to identify susceptibility genes and dysregulated signaling pathways. Analyses of differentially regulated genes revealed numerous genes that were associated with development and function of the nervous and immune systems, circadian rhythm, and ERBB signaling. Amongst the affected participants there was a patient with a 3p14.1-p13 deletion, where *FOXP1* is located. *FOXP2* mutations are responsible for human speech and language disorders. Since FOXP1, FOXP2, and FOXP4 require dimerization for transcriptional activity, investigating the FOXP1/2/4 molecular network provides insight into the neural mechanisms behind language impairments in autism. HEK293 cells were transfected with *FOXP1/2/4* constructs. QRT-PCR was used to evaluate mRNA expression of *FOXP2* target genes. Results suggest that specific combinations of FOXP1/2/4 dimers may influence the transcription of target genes involved in language acquisition.

ACKNOWLEDGEMENTS

Foremost, I would like to thank my supervisor Dr. Dorota Crawford, whose expertise and guidance has added considerably to my graduate school experience. I would also like to acknowledge my committee members, Dr. Caroline Davis and Dr. Amro Zayed, for their contributions toward the development of my projects.

These studies were supported by the generosity of the Alva Foundation and Dean's Health Research Catalyst Award (York University). I would like to recognize contributions from the Children's Treatment Network for helping with participant recruitment and the University Health Network Microarray Centre for performing microarray experiments and data analysis.

I would like to express my most sincere appreciation to my colleagues – it would not have been the same without their invaluable assistance and warm friendship. I am thankful to Dr. Hongyan Li for her vast reserve of patience and knowledge. I also wish to thank Christine Wong and Ravneet Bhogal for their immense moral support whenever I was in need.

The care and encouragement of many friends has been indispensable. A very special thank you goes to Dmitry Verniba; without whose affection and encouragement this work would not have been possible.

Finally, I take this opportunity to express the profound gratitude to my family, especially my parents, for their love and continuous support. Thank you for providing me with a loving environment in which to learn and grow.

TABLE OF CONTENTS

TABLE OF CONTENTS CONTINUED

TABLE OF CONTENTS CONTINUED

LIST OF TABLES

LIST OF FIGURES

LIST OF FIGURES CONTINUED

ABBREVIATIONS

ADI-R: Autism Diagnostic Interview-Revised

ADOS: Autism Diagnostic Observation Schedule

ANOVA: Analysis of variance

bHLH: Basic helix loop helix

BLAST: Basic local alignment search tool

BMP: Bone morphogenetic proteins

bp: Base pairs

BSA: Bovine serum albumin

CARS: Childhood Autism Rating Scale

ChIP-chip: Chromatin immunoprecipitation coupled with microarray technology

CNS: Central nervous system

CS: *FOXP2* affected patient

DAVID: Database for Annotation, Visualization and Integrated Discovery

DCs: Dendritic cells

DMEM: Dulbecco's Modified Eagle's Medium

DSM-IV: *Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition*

DVD: Developmental verbal dyspraxia

DZ: Dizygotic

E: Embryonic

FBS: Fetal bovine serum

FC: Fold change

FDCs: Follicular dendritic cells

FITC: Fluorescein isothiocyanate

KE: Family in which the *FOXP2* human mutation was found

LB: Luria-Bertani

LPS: Lipopolysaccharide

MGC: Mammalian Gene Collection

Mbp: Megabase pair

MZ: Monozygotic

NCBI: National Center for Biotechnology Information Reference Sequence

NICD: Notch intracellular domain

ORF: Open reading frame

P: Postnatal

PBS: Phosphate buffered saline

PCR: Polymerase chain reaction

PDD: Pervasive developmental disorders

PDD-NOS: Pervasive developmental disorder–not otherwise specified

Poly(I:C): Polyinosinic:polycytidylic acid

qRT-PCR: Quantitative real-time polymerase chain reaction

RIN: RNA Integrity Number

RQ: Relative quantity

RRBs: Restricted or repetitive behaviours

RT: Reverse transcription

RT-PCR: Reverse transcription polymerase chain reaction

SAM: Significance Analysis of Microarrays

SEM: Standard error of the mean

SLI: Specific language impairment

TR: Texas Red

TRH: Thyrotropin-releasing hormone

TSH: Thyroid stimulating hormone

CHAPTER 1: INTRODUCTION AND LITERATURE REVIEW

1.1. Introduction to Autism

1.1.1. Autism Overview

Autism is a devastating childhood developmental disorder that features social interaction deficits, language impairment, and behavioural abnormalities. This disorder is one of the most prevalent neurological disorders of childhood, affecting 1 in every 88 children with a four to one ratio of males to females.[1,2] Autism affects all races, ethnic groups, and socioeconomic levels.[1] It is a lifelong condition that generates enormous financial and personal costs to families and society.

Autism is not a modern medical disorder, although it has just recently gained vast recognition. Swiss psychiatrist Eugen Bleuler introduced the term "autism" in 1911 to describe a subset of symptoms of schizophrenia.[3] Derived from the Greek word "autos" meaning "self", it referenced a persona of extreme withdrawal from social interaction.[4] Subsequently, the definition of autism was largely redefined by Dr. Leo Kanner and Dr. Hans Asperger. Kanner published his original account of "early infantile autism" while at John Hopkins University in 1943. He conducted a qualitative case study of 11 children with common characteristics of social interaction and communication impairments, as well as restricted interests and repetitive behaviours.[3] Kanner became convinced that these behavioural features reflected a syndrome – a specific disorder with a characteristic set of symptoms. In the following year at the University of Children's Clinic in Vienna, Asperger independently identified a similar condition in four boys that he called "autistic psychopathy". Asperger noted that the children, all of whom were male, exhibited a lack

of empathy, little ability to form friendships, one-sided conversations, intense absorption in specific interests, and clumsy movements.[5] Unfortunately, Asperger's findings were largely ignored due to his untranslated work, which was mostly written in German. Dr. Lorna Wing is credited with widely popularizing the term "Asperger's disorder" in the English-speaking medical community and also challenged the previously accepted model of autism with the concept of an autism spectrum.[6] In recent years, the idea of an autism syndrome has been elaborated to allow for a spectrum – a continuum of disorders reflecting shared core characteristics but with slightly different patterns of symptoms. The terms 'autism' and 'autism spectrum disorders' are used interchangeably as generic descriptions of this spectrum.

In the past two decades, there has been a dramatic worldwide increase in diagnosed autism. Ten years ago, autism had an estimated prevalence of 1 in 250; at present, it is believed to affect 1 in 88 children.[7] Multiple studies have attributed it to greater societal awareness of the condition, the broadening of diagnostic criteria, advanced parental age, and more frequent concurrent diagnosis of children with mental retardation as also having autism.[7] However, these factors only partially account for the rise in diagnosed autism. The unexplained increase has mobilized researchers to gain a better understanding of the etiology of this disorder in order to develop effective diagnostic tools, intervention programs, and treatments. The specific causes of autism are unknown, although with further research, genetic factors are becoming more prevalent, among other possible causes. There is no cure for autism, but there is evidence to suggest

that with early intervention, some children may function independently and successfully in a variety of environments.[8]

1.1.2. Symptomatology of Autism

Autism is an enduring condition that begins in early childhood and generally persists throughout the life course.[9] It is distinguished not by a single symptom but by a characteristic triad of symptoms: deficits in reciprocal social interactions, impairments in verbal and nonverbal communication, and restricted or repetitive behaviours (RRBs).[10,11] Children on the autism spectrum demonstrate the same core deficits, although there is a marked variability in the severity of symptoms across affected individuals and over time.[9,11]

Retrospective literature based on parental reports and observations of home videotapes indicate that symptoms of autism exist very early in life.[12,13] One such symptom is observable impairment in reciprocal social interactions. In the first year of life, infants in whom autism was later diagnosed are more likely to fail in developing interest in reciprocal interactions and forming attachments to their primary caregivers.[14] During preschool years, affected children continue to demonstrate little interest in their peers and are frequently described as emotionally remote or disconnected.[14,15] They show flat facial affect, do not consistently respond to their own name being called, demonstrate lack of showing or pointing out objects of interest to others, and have abnormalities in gaze monitoring (i.e., does not direct one's gaze to where someone else is looking.)[13] Other behavioural factors associated with deficits in reciprocal social interactions include

poor eye contact, fewer social interactions, decrease use of gestures, and lack of spontaneous or imaginative play.[11-14]

Another hallmark of autism is difficulty with communication. In approximately one-third of children with autism, speech never develops beyond the occasional word.[14,16] Moreover, this language deficit is not accompanied by attempts to compensate through alternative modes of communication, such as gesture, signing, or eye contact.[11] The remaining two-thirds of affected children develop a varying amount of speech, ranging from stereotyped phrases to speech that is normal in structure, although pragmatics in language remain impaired.[14,16] The language of children with autism includes more egocentric speech, consisting of echolalia, self-repetition, thinking aloud, and purposeless remarks.[17] Their speech tends to have a monotone quality, without the natural lilt that aids in communication.[14] In contrast to this picture of pragmatic deficits, the language of typically developing children shows pragmatic competence from two years of age.[17] Furthermore, this deficit in pragmatics is a language abnormality unique to autism. This notion is supported by the fact that non-autistic developmentally delayed clinical groups appear to possess pragmatic competence, in comparison to their autistic counterparts.[17] This has led to the prevailing view that children with autism use language instrumentally but not communicatively.

The third feature that defines autism is RRBs. Children with autism demonstrate inflexible adherence to specific non-functional routines or rituals and often become upset at being interrupted. Also included in this category are stereotyped and repetitive motor mannerisms, such as hand or finger flapping, toe walking, or complex whole body

movements. Other behaviours include overly focused, preoccupying interests, such as schedules, calendars, or weather forecasts.[11]

Other associated symptoms, such as gastrointestinal problems, immune system irregularities, motor impairments, and sleep disturbances are common but not essential for a diagnosis.[18,19] The heterogeneous presentation of symptoms emerges gradually and with inconsistent patterns. In general, symptoms appear to manifest in one of three patterns: early onset (i.e., at or around the first birthday), late onset (i.e., after the first birthday), or regression (i.e., a period of normal development followed by a loss of previously mastered skills).[20,21] In addition, the co-occurrence of autism with another medical condition is quite frequent. For example, 1-4% of children with autism have tuberous sclerosis and 1-2% have fragile X syndrome.[22,23] Thus, the symptoms of autism may also coexist with symptoms that result from co-morbid disorders.

1.1.3. Screening and Diagnosis of Autism

A diagnosis of autism is based on descriptions and observations of behaviour. The fourth edition of the *Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition* (DSM-IV), published by the American Psychiatric Association, is an international classification guide, which lists the diagnostic criteria for autism. According to DSM–IV, pervasive developmental disorders (PDD) is the diagnostic umbrella classification for "autism-like" behaviours, which includes the prototypic autistic disorder, Asperger's disorder, Rett syndrome, and pervasive developmental disorder–not otherwise specified (PDD–NOS).[10] The distinction between these diagnoses hinge on subtle differences in age of onset, severity of symptoms, and association (or lack thereof)

with language delay and intellectual disability.[15] Autistic disorder is defined by onset prior to three years of age and includes, but is not limited to, deficits in the three core symptom domains.[11] The qualitative impairments in social interaction and RRBs in Asperger's disorder are identical to those for autistic disorder; however, language development and nonverbal intelligence are nearly normal.[24] Asperger's disorder tends to be diagnosed later than autistic disorder.[10] Because children with Asperger's disorder do not have delayed early language or problems with cognitive development, there are few early signs of abnormality. Difficulties in social interaction may become more apparent in the context of school when the child is seen with peers in a more structured social setting. As a result, a diagnosis for Asperger's disorder is most commonly made between the ages of four and eleven.[24] Additionally, delays in motor development or clumsiness may often be noted in children with Asperger's disorder.[10] Rett syndrome is characterized by normal early growth and development followed by a deceleration of head growth between ages 5 and 48 months, loss of previously acquired purposeful hand skills, loss of social engagement, impaired expressive and receptive language development, psychomotor retardation, and stereotyped hand movements.[10] Unlike the other autism conditions, Rett syndrome has a clear genetic cause (mutation of the *MECP2* gene) and affects girls almost exclusively. PDD-NOS is a "sub-threshold" condition in which there is marked impairment of social interaction, communication, and/or RRBs, but when the full features for another explicitly defined PDD are not met. Typically, the symptoms of PDD-NOS should be recognized before a child is three years old.[10] In order to identify autism, several diagnostic instruments are available. Two commonly used diagnostic

tools in autism research are the Autism Diagnostic Interview-Revised (ADI-R), a semi-structured parent interview, and the Autism Diagnostic Observation Schedule (ADOS), which uses observation and interaction with the child.[25] The Childhood Autism Rating Scale (CARS) is also widely use in clinical settings to assess severity of autism based on observation of the child.[26]

1.1.4. Genes and Autism

Autism is a neurodevelopmental disorder that has strong genetic underpinnings. Despite this, finding the cause has been daunting for scientists because of the genetic complexity and wide phenotypic variation associated with this disorder. Initial evidence for a genetic etiology in autism came from twin and family studies. Bailey *et al.* found that 60% of monozygotic (MZ) twins were concordant for autism versus no dizygotic (DZ) twins; furthermore, 92% of MZ twins were concordant for a broader spectrum of related cognitive or social abnormalities compared to 10% of DZ twins.[27] Additional evidence comes from the 2-6% recurrence rate in siblings of affected children, which is much higher than that approximate rate of 0.1-0.15% in the general population.[28] Studies of parents and siblings have further shown that several characteristics are more prevalent in the family members of affected children than that of typically developing children. These characteristics, including social reticence, language and communication difficulties, preference for routines, and difficulty with change, are qualitatively similar to the traits comprising the criteria for autism, but are milder and have come to be known as the broad autism phenotype.[14,29]

The genetics of autism remains an extremely active area of research. Combined with rapidly advancing technology, an enormous amount of data attempting to explain the influence of genetics over autism has emerged. Patterns of inheritance in multiplex families (i.e., those families in which there are more than one child affected with the disorder), and the fact that some individuals demonstrate subtle related symptoms, suggest that autism is caused by multiple genetic abnormalities, with each abnormality contributing a small increment of risk to that individual. These observations have led to a theory that the liability to autism is due, in large part, to genetic inheritance in which combinations of susceptibility alleles contribute to the autism phenotype.[30] Those individuals who inherit many of the genetic abnormalities will exhibit more severe symptoms, whereas those individuals who inherit fewer genetic abnormalities will be only slightly affected for autism. However, in spite of the evidence supporting this theory, many cases of autism are the result of a *de novo* mutation (i.e., a genetic mutation that the parents neither possessed nor transmitted). Furthermore, autism is only attributed to a recognized genetic cause in less than 10% of cases, most commonly with tuberous sclerosis, fragile X syndrome, and other medical conditions.[31] With regard to identifying genetic abnormalities associated with autism, there are three traditional approaches: genetic linkage studies, cytogenetic studies, and candidate gene studies.

Given the neurobiological complexity of autism, genome-wide methods such as genetic linkage analysis have been utilized in an attempt to identify multiple interacting chromosome regions (loci), each of which may have a small effect on an individual's overall susceptibility to autism. Many researchers have focused on multiplex families to

perform linkage analysis. Genetic linkage analysis involves systematically testing markers in multiplex families to assess whether loci that seem to segregate with the disorder can be identified. Thus far, data from the analysis of genome scans have implicated as many as 20 genes in the causation of autism.[32] Consistent evidence for linkage has emerged for chromosome 7q (the long arm of chromosome 7). Other regions of interest, verified by several studies, include 2q (the long arm of chromosome 2), 16p (short arm of chromosome 6), and 17q (the long arm of chromosome 17).[31,32]

Cytogenetic studies have also been very revealing by way of pinpointing relevant inherited or *de novo* chromosomal abnormalities in affected individuals. High-resolution cytogenetic scans in families with affected individuals have aided the location of specific genes or chromosomal loci potentially associated with autism. Chromosome aberrations, including breakpoints, translocations, deletions, duplications, and inversions, have been visualized with the use of various stains. With the exception of chromosomes 14 and 20, cytogenetic abnormalities of virtually all the chromosomes have been found to be associated with autism.[31] Cytogenetic abnormalities found at the 15q11-q13 locus are reported most frequently, representing up to 1-4% of patients.[33] Various case reports have described duplications at this locus, where maternal transmission of the defect is implicated.[34,35] Chromosome translocations have also implicated the 7q22-33 locus.[31] Reports of cytogenetic abnormalities coinciding with genetic linkage findings have been rare; however, Ashley-Koch and colleagues have identified a multiplex family in which three siblings had inherited from their mother a paracentric inversion on chromosome 7q.[36] In this family, the two male siblings have autistic disorder, while the female sibling

has expressive language disorder. With regard to mode of inheritance, this suggests that the disorder is not inherited in a simple Mendelian fashion. Multiple interacting genes, as well as environmental factors, are likely to be involved.

Candidate genes have been chosen on the basis of their position near known chromosomal aberrations, location near linkage signals, or knowledge about their biological function. The numerous reports of linkage findings and cytogenetic abnormalities on chromosomes 7 and 15 have indicated that these regions may harbour autism susceptibility genes.[37,38] Several genes have been identified within these regions and are now considered potential candidate genes of autism. The candidate genes on chromosome 7 include *WNT2* and *FOXP2*.[39,40] *WNT2* was investigated as a putative risk factor of autism because members of the WNT (Wingless) gene family influence the development of numerous organs and systems, including the central nervous system (CNS).[41] It has been reported that variations in the *WNT2* gene are found more often in affected individuals, especially in a subset of people with severe language delays.[39] One of the most interesting candidate genes of chromosome 7 is the *FOXP2* gene. Speech and language impairments have been documented in two different multigenerational families with point mutations of *FOXP2*, as well as in several cases of chromosomal rearrangements that disturb the integrity of the *FOXP2* locus.[42,43] Thus far, the association of *FOXP2* with autism has been met with mixed results, although two studies have reported positive findings.[40,44] Additionally, genes such as *UBE3A* and *GABRB3* have been identified as potential candidate genes because they map onto the common duplication site on chromosome 15q11-13. *UBE3A* is also linked to Angelman

Syndrome, a genetic disorder that shares many similar characteristics with autism. Individuals with Angelman syndrome have mental retardation, usually do not possess speech, may flap their arms and hands, and have poor muscle tone and inability to coordinate muscle movements.[32] *GABRB3* codes for GABA receptor subunits. GABA is the principal inhibitory neurotransmitter in the CNS, controlling excitability in the adult brain. There is evidence to support a role for genetic variants within *GABRB3* in autism.[45] Taken together, it is apparent that one prominent factor in the development of autism is a genetic predisposition.

1.1.5. Environment and Autism

The disparity among MZ twin pairs, who share 100% of their genes and are non-concordant for diagnosis, indicates that there are other factors that contribute to the development of autism. Although autism is believed to be primarily genetic in origin, the role of environmental factors is an area of growing interest. The developing brain undergoes its most rapid development from the prenatal period through to early postnatal life. This critical period is most vulnerable to environmental insults because it depends on the temporal and regional status of developmental processes. Some aspects of neural development continue into adolescence, and even into early adulthood (approximately 25 years of age).[46] Neurodevelopment encompasses the process of neurogenesis, proliferation, migration, differentiation, synapse formation, and myelination. Synaptic pruning and reduction of excess neurons are also critical aspects of brain development. The sequential unfolding of form and function involves tightly regulated gene expression changes. Disruption of this sequence at any point during development can have long-

lasting consequences for brain function later in life. Each of these developmental events represents a substrate for environmentally induced dysregulation, alone or in concert with genetic susceptibility, to influence the risk for autism.

There is evidence to suggest that environmental exposure to chemical and biological agents during critical periods of gestation and/or early postnatal development can increase the likelihood of the development of autism.[47] Figure 1.1 depicts the critical periods during prenatal human development that, following exposure to teratogens, can lead to developmental abnormalities.[48] For instance, maternal use of alcohol during pregnancy can result in fetal alcohol syndrome, birth defects, and cognitive or behavioural problems.[49] Several studies have described children with co-morbid diagnoses of both fetal alcohol syndrome and autism, suggesting that in some cases alcohol may be responsible for autism as a result of early gestational exposures.[50,51] Likewise, valproic acid, an anti-epileptic medication, has been linked with cases of autism, as well as developmental delays or neurological abnormalities in a number of exposed children.[52,53] Thalidomide is another notable association of autism with a pharmaceutical agent. Thalidomide, advertised in the 1950s as a safe drug for morning sickness, caused developmental disabilities and severe limb defects in the offspring of some women who took it while pregnant. The risk of autism was particularly increased when the exposure to the drug occurred early in pregnancy, specifically between the 20th and 24th day of gestation, lending support to the idea that autism reflects a disruption of neurological events early in the embryological period.[47] More recently, a connection between misoprostol exposure during the first trimester of pregnancy and the

development of autism-related features in Möbius syndrome has been established.[54] Collectively, these findings suggest that the window of susceptibility for autism induction may be very early in gestation. Fortunately, there is a growing children's environmental health movement that emphasizes the unique vulnerabilities of fetuses and children to environmental exposures and points out the increasing body of evidence for adverse effects on biological growth and development.[55]

Maternal infection and prenatal viral exposure is considered yet another non-genetic cause of autism. Many studies over the year have presented evidence for the association of autism with viral infections. The strongest association of a viral infection to autism is congenital rubella.[56] The prevalence rates were 412 children meeting full criteria of autism, with an additional 329 children displaying a significant number of autism-related symptoms, for a total of 741 affected children per 10,000 children having congenital rubella. Other viruses have also been reported to be associated with autism. Most of these reports are case studies linking viruses such as cytomegalovirus and herpes simplex to autism.[57,58]

A large body of evidence indicates that the immune system appears to play an important role in the development and progression of neurological disorders.[59] More specifically, disturbances of the blood-brain barrier system resulting from inflammation have been implicated as one of the leading causes. Infection-associated immunological events in early fetal life may play a role in the initiation of neurological disorders due to long-term alterations in the vasculature and function of the blood-brain barrier. A dysfunctional blood-brain barrier could lead to immediate cell death during susceptible

periods of brain development, thereby predisposing the developing nervous system to additional failures in subsequent cell migration and synapse maturation, or may increase the susceptibility of the brain to further exposures to toxins.

Recent studies have shown that children diagnosed with autism display sustained inflammation, as evidenced by an increased expression of pro-inflammatory cytokines in the brain.[60] The transport or synthesis of cytokines in the brain facilitates neuroinflammation and possible neurotransmitter imbalances.[61] This can have profound effects on many aspects of brain function including neurotransmitter metabolism, neuroendocrine function, synaptic plasticity, and neurocircuits that regulate mood, motor activity, motivation, anxiety, and alarm.[62] In a mouse experimental system, it has been shown that the induction of prenatal inflammation using poly(I:C) during gestation led to disturbances of brain maturation associated with cognitive and behavioural defects, which were comparable to some of the symptoms observed in autism.[63] Specifically, the relative expression of pro- and anti-inflammatory cytokines in the fetal brain in response to maternal immune challenge was deemed an important determinant for the precise pathological profile emerging in later life. The mechanisms of fetal neurotoxicity are complex and still not fully understood.

Autism is also associated with some obstetric risk factors, including advanced parental age, low birth weight, gestation duration, and hypoxia during childbirth.[64,65] A variety of postnatal contributors to autism have been proposed as well, including allergies, drug exposure, vaccines, post-natal infections, certain foods, and heavy

metals.[66] Thus far, the evidence for these risk factors is anecdotal and has not been confirmed by reliable studies.

A primary area of interest that holds great importance for discerning the environmental contributions to autism is whether regressive autism reflects a distinct etiology from non-regressive autism. There has been significant public concern that exposure to harmful environmental agents during the early postnatal period triggers autism. The prospect that regressive autism is provoked by an environmental exposure suggests the need to carefully examine a range of postnatal exposures in addition to possible *in utero* exposures. On the other hand, it should be noted that the occurrence of developmental regression does not necessarily implicate environmental causation. There are examples of disorders, such as Rett syndrome, that are known to be linked to a single gene mutation that features a period of normal development followed by regression. Instead, regressive autism may reflect a latent effect of much earlier exposures.

The rare instances of environmental agents acting alone to cause autism suggests the probable interaction of environmental influences with autism susceptibility genes. Interactions of genes and environment might occur at different levels. Genetic influences may act synergistically with environmental exposures as primary triggers of autism. Alternatively, a genetic predisposition may lead to the development of autism while environmental factors act to modify the phenotypic expression of the disorder. Research to date supports the theory that many different patterns of genetic and environmental influences act on a common set of neurodevelopmental processes to affect the risk of autism or its phenotypic expression.[67]

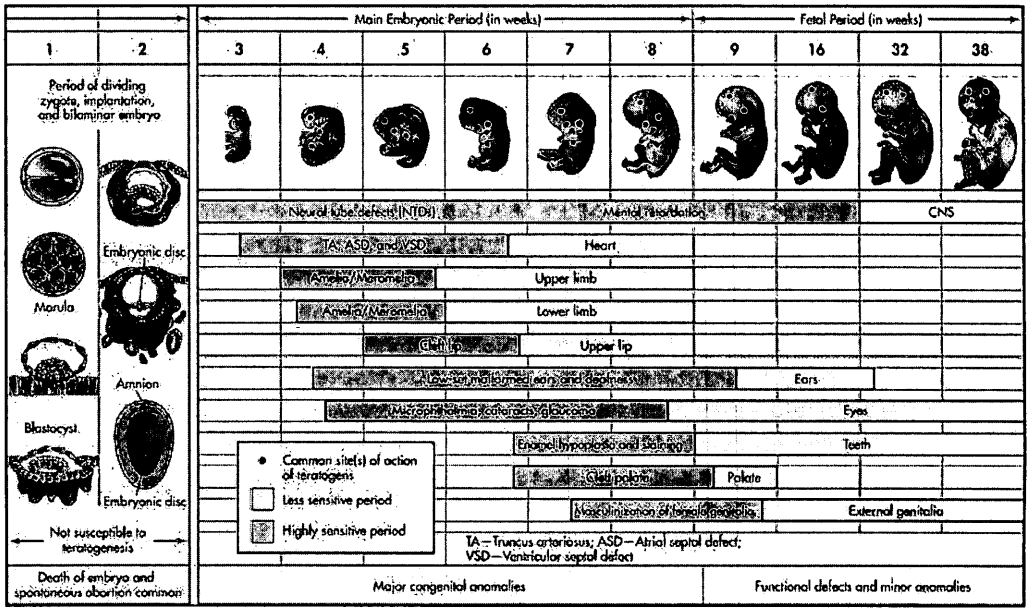

Figure 1.1. Critical periods of human development. Prenatal development is divided into three stages: periconception (generally the 2 weeks following fertilization), embryonic (3 to 8 weeks), and fetal (9 to 38 weeks). During the periconceptual period, hazardous environmental exposures usually cause fetal death. Disruption during the embryonic period can result in major physical malformations. Both the organ affected by exposure during this period and the resulting type of anomaly are highly dependent on both the agent and gestational age at which the exposure occurs. During the later stages of prenatal development, environmental exposures can result in impaired growth, physiological defects, or functional deficiencies. Figure adapted from Moore and Persaud. 1998.[48]

1.2. Early Detection of Autism

1.2.1. Eluding Early Diagnosis: Diagnosis Based on Behaviour Alone

As with any child with special needs, early identification is crucial in order for interventions to be implemented sooner. With regard to autism, there is an urgent need for early identification before deviation from the normal trajectory of development has progressed too far. As many as 50% of parents suspect a problem in the first year of life and most parents express concern to their paediatrician by the time their child is 18 months of age.[12] Yet, autism is not commonly identified before two to three years of age.[68] Many clinicians hesitate to ascribe a diagnosis of autism even when some symptoms may be evident due to concerns about family distress, the possibility of being incorrect, or hope that the symptoms will reverse over time. In reality, the advantages of early diagnosis are numerous and include better preparation for educational planning and treatment, provisions for family support, reduction of family stress and anguish, and delivery of appropriate medical care to the child.[11] The available evidence from a number of programs and studies suggest that early intervention leads to better outcomes. The fact that children demonstrate greater gains when they enter a program at a younger age further emphasizes the need for diagnostic instruments that are effective before two years of age.[69] However, given the variability of symptoms, differences in age of onset, severity of symptoms, and possible regression of previously acquired skills for some children, autism may be especially difficult to diagnosis early and accurately based on clinically observed behavioural signs alone.

1.2.2. The Search for Biomarkers: Indicators of Risk for Autism

The main impediment to early identification efforts resides in the fact that research on autism generally does not begin until a child has been identified with the disorder, which in the best scenario may not be until two years of age. Thus, knowledge about the infancy and toddler periods in autism is rare and exclusively based on retrospective sources of information gathered well after a diagnosis is made. Although parent reports and home videotapes have provided crucial information concerning the early behavioural phenotype of autism, this method has limited experimental control. One promising approach that is making headway in the field of autism is the supplementation of behavioural criteria with biological markers (or simply, biomarkers). Biomarkers have potential for improving diagnostic efficiency. The identification of a reliable biomarker for autism could overcome the limitations of the existing clinical behavioural methods. The ability to assess an infant's risk status based on a biomarker that is expressed before the onset of symptoms could eliminate the need to wait for behavioural criteria to be met before beginning treatment. This would likely improve the prognosis of affected individuals. A promising area of biomarker research focuses on genes and their encoded products. It is clear that genes play a large role in the causation of autism; thus, the pursuit of gene-based biomarkers seems to be a plausible approach to the research.

1.2.3. Limitations of Traditional Methods of Studying Autism

The identification of genetic linkages and chromosomal abnormalities, in combination with candidate gene studies, has aided in the elucidation of the neurobiological underpinnings of autism. Genetic linkage studies have even shown

progression by expanding to include environmentally responsive genes in traditional linkage searches.[70] However, genetic studies to date have not yet uncovered any genes of strong effect. Given the multi-factorial nature of autism, studying a single candidate gene at a time is not likely to provide a comprehensive explanation of the phenotypic variation, which is believed to result from dysregulation of multiple genes. Furthermore, another aspect of studying genetic abnormalities that limits their utility as biomarkers is that they are static factors that do not change over the life course. In comparison, gene expression is highly flexible and can vary over time within the same individual. Gene expression can also reflect genetic or environmental influences, or both. These qualities afford gene expression greater potential to serve as useful biomarkers for various stages of autism, such as the asymptomatic at-risk period, early periods of subtle deficits, or responsiveness to treatment.[71]

1.2.4. The New Frontier: Gene Expression Profiling

The emergence of microarray technology has allowed many researchers to monitor the activity of a multitude of genes simultaneously to create a global picture of cellular function. DNA microarrays are ordered arrangements of nucleic acid fragments derived from individual genes located at defined positions on a solid support, enabling the investigation of thousands of genes in parallel by specific hybridization.[72] There are two types of DNA microarrays depending on the material arrayed: cDNA or oligonucleotide. Whether spotting cDNA or synthesized oligonucleotides, the nucleic acid fragments are deposited onto substrate-coated glass slides using highly precise robotic systems to produce dense arrays. This confers the advantage of smaller amounts

of sample for analysis because the arrays have small volumes for hybridization. The study of gene expression using microarrays is based on the competitive hybridization of differently labelled populations of cDNAs. This technique, called gene expression profiling, provides the opportunity to evaluate possible deviations in autism at the mRNA level. Since early development is largely dependent upon intrinsic gene expression patterns, altered mRNA levels during critical periods may significantly impact brain development. This method allows the evaluation of a wide range of candidate genes and how they fit into common molecular pathways. This approach may identify novel genes which are dysregulated in autism, as well as provide the basis for the development of more sensitive diagnostic tools for this disorder. The experimental strategy used in gene expression studies is to tease out differences in mRNA levels between affected and unaffected individuals. Beyond this working model, it may also address the phenotypic heterogeneity of autism, which may potentially map onto heterogeneity in biomarker profiles. Thus, as influential etiological factors and their associated biomarkers are identified, different phenotypic profiles of autism may be established based on similar gene expression profiles obtained between individuals that display overlapping symptoms (e.g., regressive versus early onset of symptoms). The subsequent detection of both environmental and biological influences on autism and biomarkers within these subgroups can then be used to begin appropriate, personalized interventions.

1.2.5. Biological Materials for Gene Expression Profiling

Several different types of biological materials have been used to differentiate gene expression patterns between affected individuals and healthy controls: post-mortem adult

or fetal brain tissues and white blood cells either in the form of peripheral mononuclear blood cells or immortalized lymphoblastoid cell lines.[73,74] Post-mortem brain tissue confers the advantage of exploring gene expression alterations in tissues most directly implicated in this disorder. Brain tissue is presumed to be the ideal study material for the mechanistic basis of autism but it is not without its limitations. The scarce availability of donated brain tissue samples for adequate sample sizes, difficulties in preserving the RNA, and potentially confounding factors related to the cause of death and tissue modifications during the post-mortem interval represent major drawbacks for this tissue source. More importantly, it is not an appropriate target for diagnostic assays. Researchers have turned instead to RNA extracted from white blood cells drawn from blood samples. It has been hypothesized that autism might arise through dysregulation of specific neuronal genes and that expression differences between affected and unaffected individuals might be present in tissues other than the brain.[75] While conceivably more distant from the pathophysiology of this disorder, it provides access to larger quantities of RNA and it is easier to obtain, which ensures larger sample sizes.[76] However, the development of a non-invasive method of sample collection is of foremost importance considering that blood collection is a significant barrier to study participation, especially in studies involving young children.

Buccal cells, also known as mouth epithelial cells, are easily collected with minimal discomfort and present an alternative sample material for biomarker testing. The rationale behind the use of buccal cells is that since epithelial and nervous tissue are derived from the same embryonic germ layer – the ectoderm – insults that occurred

during early development would be common to both cell types. Collection of buccal cells has proven to be an effective and painless procedure, and can be obtained using a variety of different methods, including cytobrushes, sterile swabs, mouthwash, and saliva.[77] The pursuit of epithelium-based biomarkers is still in its infancy but early successes provide the impetus for further investigation in this area. A few studies pertaining to oral cancer address the efficacy of using buccal cells for microarray gene expression analysis.[78,79] An initial study comparing blood and buccal cells from the same individuals looked at the relative amounts of four genes and determined that the RNA isolated from buccal cells was of sufficient quality to be used to detect expression differences.[78] Buccal cell RNA has also been used for microarray-based differential gene expression studies by comparing gene expression between smokers and non-smokers. The results revealed several groups of genes that were either over- or under-expressed in smokers and which could be used to predict smoking exposure.[79] To date, buccal cells have not been used in a transcriptomic approach to investigate gene expression differences in autism. A successful study of this type would support the notion that buccal cells have efficacy as source material for biomarker discovery in autism.

1.3. *FOXP* Gene Studies

1.3.1. Discovery of *FOXP2* as a 'Language Gene'

As mentioned previously, *FOXP2* was brought to the attention of the scientific community when it was identified as a gene that predisposes individuals to a rare speech and language disorder. In 1990, Hurst *et al.* reported a case of a three-generation pedigree, known as the KE family, half of whose members had developmental verbal dyspraxia (DVD) that is inherited in a pattern consistent with an autosomal dominant penetrance (Figure 1.2).[42,80] DVD is a complex condition that has its onset early in childhood and includes elements of impairment in speech articulation and other linguistic skills, as well as broader intellectual and physical problems. Using data from the KE family, Fisher and colleagues performed a genome-wide search for the candidate gene and finally mapped the gene locus to chromosome 7q.[81] In 2001, analysis of patient CS, who had a translocation between chromosome 5 and 7 that caused almost identical language deficits, helped to identify *FOXP2* as the gene responsible for this disorder.[82] These findings were further supported by the discovery of a *FOXP2* nonsense mutation at arginine 328 in another family, who affected members had orofacial dyspraxia.[43]

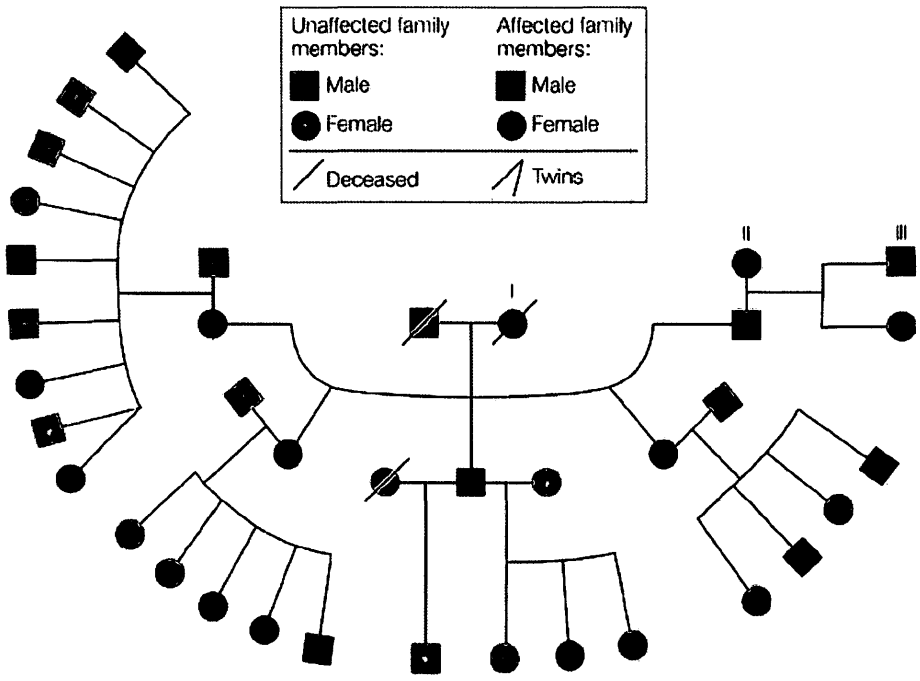

Figure 1.2. Inheritance pattern of the *FOXP2* mutation in the KE family. The KE family consists of three generations, half of whose members (15) are affected by developmental verbal dyspraxia (DVD). The 16 unaffected family members and 6 spouses who married into the family are also displayed in the KE family pedigree. Squares represent males and circles show females. Dark blue colour shows affected members and light blue colour unaffected. Image adapted from Vergha-Kadehm *et al.* 2005.[80]

1.3.2. *FOXP2* in Relation to Specific Language Impairment and Autism

Specific language impairment (SLI) is defined as a failure to acquire normal expressive or receptive language despite adequate environmental influences and in the absence of other medical genetic conditions. Although autism has a clear diagnostic criterion from SLI, there is an interesting phenotypic overlap between these disorders. This idea is best demonstrated in instances where features of autism and SLI cannot be distinctly separated. There are reported cases of individuals who were diagnosed with language disorders during childhood whom nevertheless show significant social difficulties as adults.[83] Given the report of linkage to 7q in the KE family and a positive association in another sample of children with SLI, the Collaborative Linkage Study of Autism carried out a genome screen on individuals with autism from multiplex families. An estimated 29% of the families were linked to a specific region on chromosome 7.[84] This region is the same as the locus that Fisher *et al.* had previously implicated in SLI. Additional evidence comes from the prevalence rates of these disorders. Just as there is a higher prevalence of language difficulties in families of autistic individuals, there is also a higher prevalence of autism in the siblings of probands with SLI as compared to the general population. Furthermore, the prevalence of autism among the siblings of SLI probands is approximately the same as that seen in the siblings of autistic probands.[85] This suggests that the two conditions are pathogenetically linked and may have some causative genes in common, such as *FOXP2*. Chromosome 7q31, in which *FOXP2* is located, has been considered to be a susceptibility locus for the language deficits in SLI and autism. Thus far, association and mutation screening analyses of *FOXP2* have been

performed in these disease groups with mixed results. Researchers have yet to identify a mutation or association with *FOXP2* within SLI patients and, although two studies show a positive association between *FOXP2* and autism, the majority of the literature indicates a negative association.[40,44,86,87] However, the *FOXP2* gene is exceptionally large in size and novel exons have recently been identified.[43] Further studies will need to be conducted to determine the relationship between *FOXP2*, SLI, and autism.

1.3.3. *FOXP* Sub-Family Genes

The forkhead box (*FOX*) gene family encodes a large group of transcription factors that share a common DNA binding domain of sequences called the forkhead. Many members of the *FOX* family are involved in embryonic morphogenesis and mutations in the *FOX* genes have been linked to human developmental disorders including immune, skeletal, circulatory, and craniofacial defects.[88]

The *FOXP* sub-family of genes were recently recognized to be members of the *FOX* gene family. In addition to a forkhead domain, *FOXP* sub-family members contain a zinc finger domain and a leucine zipper motif.[43] The *FOXP* sub-family comprises four members: *FOXP1-4* (Figure 1.3).[89] The expression patterns of *FOXP* genes in fetal mouse, rat, nonhuman primate, and human brains show striking homology. Therefore, it stands to reason that findings from animal models would translate to humans. Here, the standard accepted nomenclature to refer to genes in different species has been adopted: *FOXP* for humans and *Foxp* for mice. Three of the *FOXP* subfamily members (*FOXP1*, *FOXP2*, and *FOXP4*) are abundantly expressed throughout the developing brain. Conversely, *FOXP3* is expressed predominantly in the immune system.[43] *FOXP1/2/4* are

highly homologous and show more than 60% identity at the amino acid level and around 80% identity at their forkhead domains at the amino acid level.[43] Although *FOXP1/2/4* genes are expressed in the same anatomical regions, there is a distinct expression pattern for each of these genes in some neuronal cell types.

FoxP1, the member that was isolated first, is a regulator of mouse lung, heart, brain, testis, kidney, and gut development.[90] *Foxp1* null embryos have severe defects in the cardiovascular system, including defects in ventricular sepatation, cardiac outflow tract, and myocyte maturation and proliferation.[91] *Foxp1* null embryos have also shown abnormalities in motor neuronal identity in the spinal cord.[92] Since *Foxp1* null embryos die at embryonic day 14.5, the role of *Foxp1* during the later stages of brain development has yet to be elucidated.[91] *Foxp2* is involved in lung, heart, gut, and brain development.[93] Within the brain, *Foxp2* is expressed in several structures during development, including the cortical plate, basal ganglia, thalamus, inferior olives, and cerebellum.[94] Functional analyses of *FOXP2* in multiple species have revealed intriguing aspects of brain development and function. In humans, as discussed previously, *FOXP2* is implicated in the biological mechanisms specifically involved in speech and language development. In mice, heterozygous mutations of *Foxp2* display abnormal synaptic plasticity in striatal and cerebellar brain circuits, accompanied by significant deficits in motor skill learning.[95] Homozygous *Foxp2* mutations yield more severe neural effects – altered ultrasonic vocalizations, cerebellar abnormalities, motor impairment, reduced postnatal weight gain, and early postnatal lethality.[95-97] Additionally, selective knockdown of the avian *FoxP2* gene leads to incomplete and inaccurate imitation of tutor song, suggesting its

requirement in normal vocal learning in songbirds.[98] *Foxp4* is expressed in the developing heart, kidney, liver, spleen, and brain.[99] It has been revealed that in *Foxp4* mutant mice, *Foxp4* is essential for cardiac morphogenesis as these mice develop abnormally with two complete hearts and died in the embryonic stage.[100] Moreover, *Foxp4* is involved in the maintenance of Purkinje cells in the cerebellum.[101]

Members of the FOX family of proteins bind to target DNA as monomers. FOXP1/2/4 proteins have the unique ability to dimerize for DNA binding and their transcriptional activity is regulated by homo- and heterodimerization, which is dependent on the zinc finger domain and leucine zipper motif.[102] The precise combination of homo- and heterodimers of different FOXP1/2/4 proteins within the same neurons may modulate the transcription of downstream gene targets. With regard to brain development, changes in the expression of FOXP1/2/4 transcriptional factors may result in premature or delayed activation of multiple downstream target genes and force morphogenetic processes out of step.

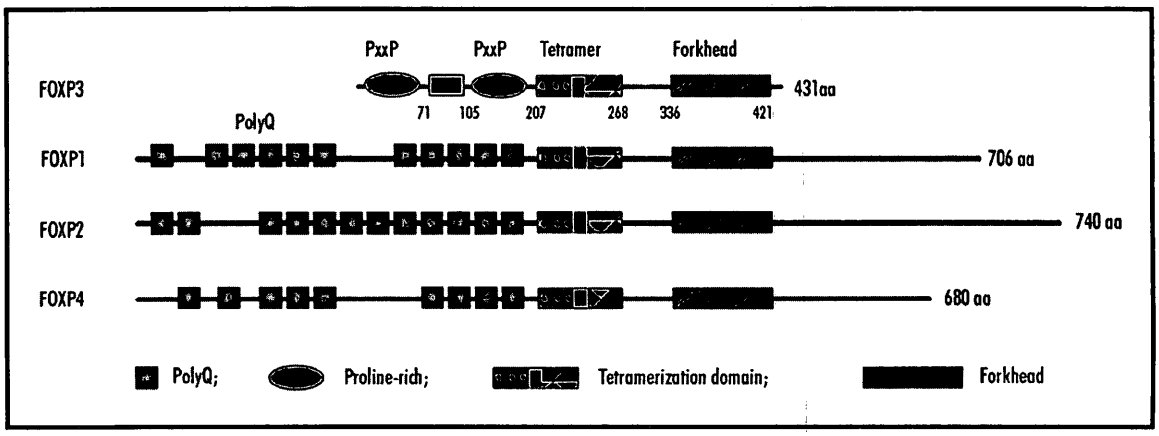

Figure 1.3. Primary domain structure of FOXP subfamily members. The N-terminal amino acid sequence of FOXP3 is different from other FOXP subfamily members, including FOXP1, FOXP2, and FOXP4. Image adapted from Li and Greene. 2007.[89]

1.3.4. Defining Downstream Pathways Regulated by FOXP2

Since *FOXP2* encodes a transcription factor, functional genomics have been employed to investigate other key factors in the signaling pathways of which it participates. A signaling pathway involves mediators at multiple levels, including those that act upstream of *FOXP2* to regulate its expression, elements that modify the intracellular localization or activity of the FOXP2 protein, co-factors which interact with the FOXP2 protein, and downstream targets that it activates or represses.[103] Although mutations in the *FOXP2* gene are uncommon, it is reasonable to expect that the downstream targets in which it regulates in the brain are more broadly implicated in subtle defects of speech and language. High-throughput searches for genes regulated by FOXP2 have been carried out in human fetal brain samples and neuronal-like cell models. Two groups have identified several hundred putative gene targets; half of these recognized gene targets are overlapped between the two studies, thus strengthening the findings.[104,105] The expression of a majority of the target genes was repressed by FOXP2 while the expression of a minority of the targets was activated by FOXP2. Thus, FOXP2 can act as both a repressor and an activator under certain circumstances, possibly depending on co-factors such as FOXP1 or FOXP4. The analyses of co-factors that interact with FOXP2 is a gap in the existing literature that needs to be addressed, especially since it might provide further clues to the involvement of FOXP2 in speech and language functions. Studies of *FOXP2* will not only clarify fundamental questions regarding the origins of human speech and language, but also its contribution to the etiology of neurodevelopmental disorders such as autism.

CHAPTER 2: GENE EXPRESSION PROFILING IN INDIVIDUALS WITH AUTISM

2.1. Introduction

Autism is dramatically increasing in incidence. Yet, to date there are no definitive biomarkers that can be reliably used to diagnose autism. The majority of diagnoses are dependent on DSM-IV behavioural guidelines, using diagnostic tools such as the Autism Diagnostic Interview-Revised (ADI-R) or the Autism Diagnostic Observation Schedule (ADOS).[25] Although the guidelines are well-defined, the rater's perception of the evaluated behaviour allows for ambiguity since the interpretation of the symptoms might vary among different individuals (e.g. parents, teachers, or clinicians). Furthermore, a diagnosis for Asperger's Syndrome is usually not made until well after the child starts school and, even then, the child is often diagnosed with common learning disorders before Asperger's Syndrome is considered.[106] Therefore, there is a strong need for biomarkers that can be used consistently in a clinical setting to facilitate the diagnosis of autism, thereby allowing appropriate intervention and effective educational programs to commence at an earlier age. Moreover, it is important to identify genes that belong to the same molecular pathways and are associated with distinct autism phenotypes in order to design pharmaceutical treatments targeted to specific individuals.

The advent of microarray technology, which allows for simultaneous comparison of thousands of transcripts, has provided the opportunity to investigate susceptibility genes and molecular pathways in psychiatric disorders. The advantage of gene expression

profiling using microarrays is the ability to study multiple genes in the context of gene regulatory networks within a living cell. Since autism is a multi-factorial disorder that is believed to be caused by a combination of genetic and environmental factors, using a global approach to study this complex disorder is well-suited. Thus far, the application of microarray technology to autism research has been limited to the use of post-mortem brain tissue, peripheral mononuclear blood cells, or lymphoblastoid cell lines.[73,74] Given that autism has an early childhood onset, a less invasive method of sample collection would be of foremost importance. Exfoliated cytologic specimens from the buccal mucosa (mouth epithelium) have been shown to contain viable cells for gene expression assays, permitting non-invasive measurement of gene-environment interactions.[78] Buccal cells are easily accessible and are derived from the same germ layer as neuronal cells (ectoderm) and thus may be used as a proxy to represent the neural transcriptome. Currently, there is no evidence that dysregulated gene expression or biomarkers exist in buccal cells of autism patients. However, it has been hypothesized that autism might arise through dysregulated expression of specific neuronal genes and that expression differences between affected and unaffected individuals might be present in tissues other than the brain.

2.2. Hypotheses

In this study, buccal samples were utilized for microarrays to investigate gene expression differences between typically developing children and children with autism. I hypothesize that:

1) RNA derived from buccal cells is a reliable genetic source for detection of gene expression profiles (mRNA levels);
2) RNA derived from buccal cells can be used for detection of autism susceptibility genes that are neuronally expressed;
3) Identified candidate susceptibility genes participate in common signaling pathways that are critical to the development and function of the central nervous system (CNS);
4) Identified candidate susceptibility genes are expressed during early embryonic development; and
5) There is a relationship between severity of autism symptoms exhibited by the probands and the expression level of the most differentially expressed genes.

2.3. Methods

2.3.1. Collection of Buccal Samples

Buccal samples were obtained from individuals that were 25 years of age or under at the time of sample collection. The rationale behind the age range is that this study is focused on investigating the developmental changes in autism and some aspects of neural development continue on into early adulthood (approximately 25 years of age).[46] The collection protocol was reviewed and approved by the Research Ethics Board of York University, Toronto. All parents or legal guardians provided their written informed consent to participate. Human HT-12 v4.0 BeadChips, which require a minimum of three control samples for assays, was used for global gene expression analysis. There were 9

probands (8.7 ± 2.5 years old; 8 males and 1 female) and 4 healthy controls (9.3 ± 5.0 years old; 3 males and 1 female) recruited to this study through participant recruitment advertisement at York University and collaboration with the Children's Treatment Network. All probands met the ADI-R, ADOS, or CARS criteria for autism as determined by an experienced child psychologist or a multidisciplinary team which included a paediatrician, occupational therapist, speech-language pathologist, and social worker. Participants were defined as healthy controls if the individual had no history of psychiatric disorders and met milestones for social, emotional, and communication development at appropriate ages. The number of male to female participants in this study reflects the skewed sex ratio indicative of a greater preponderance of males over females in autism. Full demographic details of the participants can be found in Table 2.1. The symptomatology (phenotype profile) of the affected participants will be discussed in section 2.4.5.

Buccal samples were collected using sterile Cytobrush Plus® (Cooper Surgical). The subjects were instructed not to consume any foods or drinks other than water at least 30 minutes prior to the sample collection. The brush was inserted into the subject's mouth while holding the chin firmly with one hand. The brush was pressed against the buccal mucosa and swiped in a single direction at least 15 times. The brush was then rotated and the procedure was repeated with the other side of the brush. Duplicate brushes were collected from each subject, one from each side of the inner cheek, at a single time point. There were no significant episodes of irritation or bleeding. The brushes were immediately stored in 2 mL microcentrifuge tubes containing 400 μL of

room temperature RNAlater solution (Clonetech) to prevent post-sampling degradation of the RNA. The brush ends were cut off with sterile surgical scissors so that the tubes could be capped.

Table 2.1. Demographic details of the participants with RNA Integrity Number (RIN).

Participant ID	Sex	Age	Ethnicity	Status	RIN
C8	Male	10	CAU	Healthy Control	6.5
C9	Male	23	ASN	Healthy Control	5.9
D8	Male	1	ASN	Healthy Control	6.9
D3	Female	3	CAU	Healthy Control	8.2
A6	Male	3	CAU	Proband	7.2
B1	Male	25	CAU	Proband	6.7
B8	Male	2	CAU	Proband	6.6
B10	Male	14	CAU	Proband	7.2
C7	Male	4	HIS	Proband	6.6
C10	Male	10	CAU	Proband	6
C11	Male	5	ASN / HIS	Proband	9.4
C12	Male	4	CAU	Proband	7.1
D5	Female	11	CAU	Proband	7.5

Shown is the demography of the study population used in the microarray study. CAU – Caucasian; ASN – Asian; HIS – Hispanic.

2.3.2. Total RNA Isolation

Total RNA was extracted from the buccal cells within 24 hours after sample collection. The duplicate brushes from each subject were processed together. Total RNA was isolated from buccal cells using the NucleoSpin RNA XS kit (Clontech). The extraction protocol was followed as outlined by manufacturer's instructions. RNA yield was determined on a NanoDrop ND-1000 Spectrophotometer. The RNA samples were subsequently sent to the University Health Network Microarray Center for quality control and microarray data analysis.

2.3.3. Quality Control and Microarray Data Analysis

Total RNA quality was assessed using the Agilent Bioanalyzer 2100. Electropherograms generated on the Agilent Bioanalyzer 2100 provide information about RNA concentration, allow visual inspection of RNA integrity, and automatically generates the ratio of 18S to 28S ribosomal subunits. RNA degradation is a gradual process where, as degradation proceeds, there is a decrease in the 18S to 28S ribosomal band ratio and an increase in the baseline signal between the two ribosomal peaks and the lower marker. The important elements of an electropherogram are depicted in Figure 2.1. The RNA Integrity Number (RIN) software algorithm allows for the classification of eukaryotic total RNA, based on a numbering system from 1 to 10, where 1 represents the most degraded profile and 10 being the most intact.

Illumina gene expression protocols feature a first- and second-strand reverse transcription step, followed by *in vitro* transcription amplification that incorporates biotin-labelled nucleotides. For global gene expression studies, 50 ng of each RNA

sample was labelled using Illumina TotalPrep RNA Amplification kit Lot #1107026 (Ambion) as per amplification protocol. 750 ng of cRNA generated from amplification and labelling were randomized and hybridized onto Human HT-12 v4.0 BeadChip arrays featuring 12 arrays per slide. Each Human HT-12 v4.0 BeadChip targets 47,231 probes derived from the National Center for Biotechnology Information Reference Sequence (NCBI) RefSeq Release 38 (November 7, 2009) and other sources. The BeadChip array was incubated at 58°C, with rotation speed 5 for 18 hours for hybridization. The BeadChip array was washed and stained as per Illumina protocol and scanned on the iScan (Illumina). The data files were quantified in GenomeStudio Version 2011.1 (Illumina). Differential expression comparing probands to healthy controls was performed with Significance Analysis of Microarrays (SAM). *P*-values were corrected using Benjamini and Hochberg False Discovery Rate method. Gene Enrichment Analysis was used to test the array data for enrichment of differentially expressed genes.

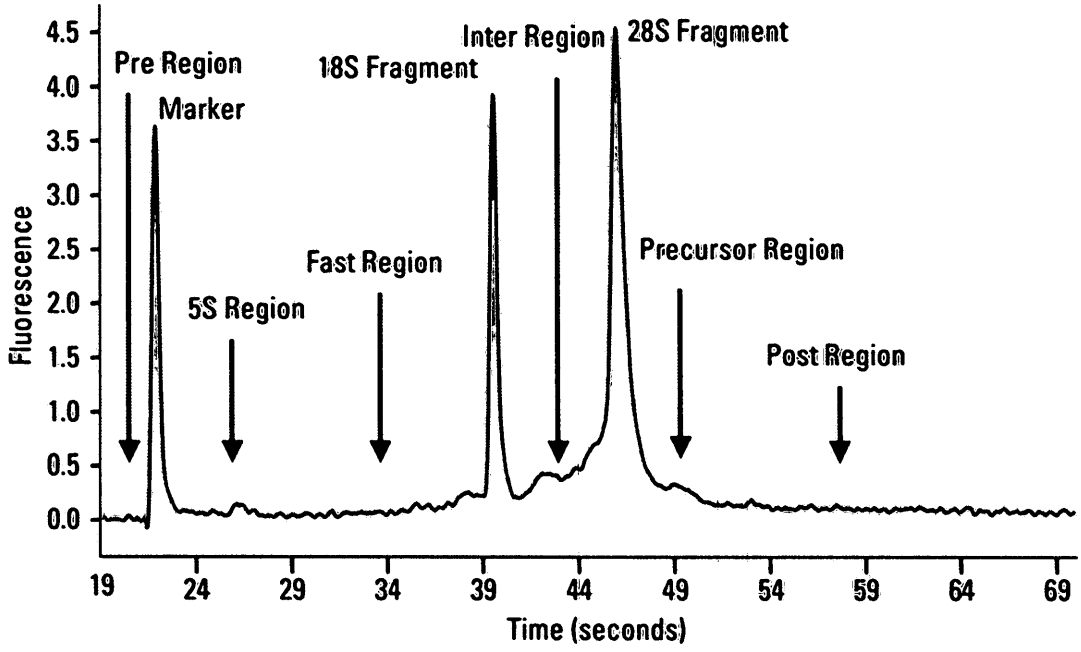

Figure 2.1. Electropherogram detailing the regions that are indicative of RNA quality. The segment preceding the lower marker is designated the pre region. The marker region coincides with the area occupied by the marker peak. The 5S region covers the small ribosomal RNA fragments (5S and 5.8S ribosomal RNA, and transfer RNA). The fast region lies between the 5S and the 18S peak. The 18S and 28S peaks correspond to the 18S and 28S ribosomal RNA species, respectively. The inter region lies between the 18S and 28S peaks. The precursor region covers the precursor RNA following the 28S peak and the post region lies beyond the precursor region. As RNA degradation proceeds, there is a decrease in the 18S to 28S ribosomal band ratio and an increase in the baseline signal between the two ribosomal peaks and the lower marker. Image adapted from Agilent Application Note, Publication Number 5989-1165EN. 2004.

2.3.4. Function and Pathway Analysis

The dataset of differentially regulated genes between probands and healthy controls were analyzed using The Database for Annotation, Visualization and Integrated Discovery (DAVID) Bioinformatics Resources (http://david.abcc.ncifcrf.gov/) to identify molecular functions and biological processes, as well as signaling pathways. Statistically significant functions were determined using the Fisher Exact test. To compare expression data at the transcriptomic level to protein-level expression, the Human Protein Atlas (http://www.proteinatlas.org/) was used to find out if the differentially regulated genes are expressed in the CNS. The Human Protein Atlas is a publicly available database cataloging the distribution of proteins in different normal human tissues, cancer types, and cell lines via validated antibody analysis. The data provided include immunohistochemistry and Western blot analysis. Establishing the protein expression profile of the differentially regulated genes helped to determine if genes expressed in peripheral buccal cells may also have a role in the CNS.

2.3.5. Reverse Transcription Polymerase Chain Reaction and Gel Electrophoresis

Based on the microarray results, differentially regulated genes with fold change (FC) above 1.5 were further characterized using mouse as an experimental model system. The mouse is often considered a good model system for investigating human development because the developmental sequences of the organ systems in the two species closely resemble one another. Human gestation lasts for approximately 40 weeks and corresponds to 19 days of gestation in mice. As well, a mouse model was used because 99% of mouse genes have analogues in humans. Three mouse developmental

stages were investigated – embryonic day 16 (E16), embryonic day 19 (E19), and postnatal day 8 (P8) – as well as the adult stage (Figure 2.2). The mouse brain samples for E16 and E19 comprise pooled RNA from 11 and 13 embryos, respectively. The P8 sample consists of pooled RNA from 6 pups, and the adult sample is from 5 animals.

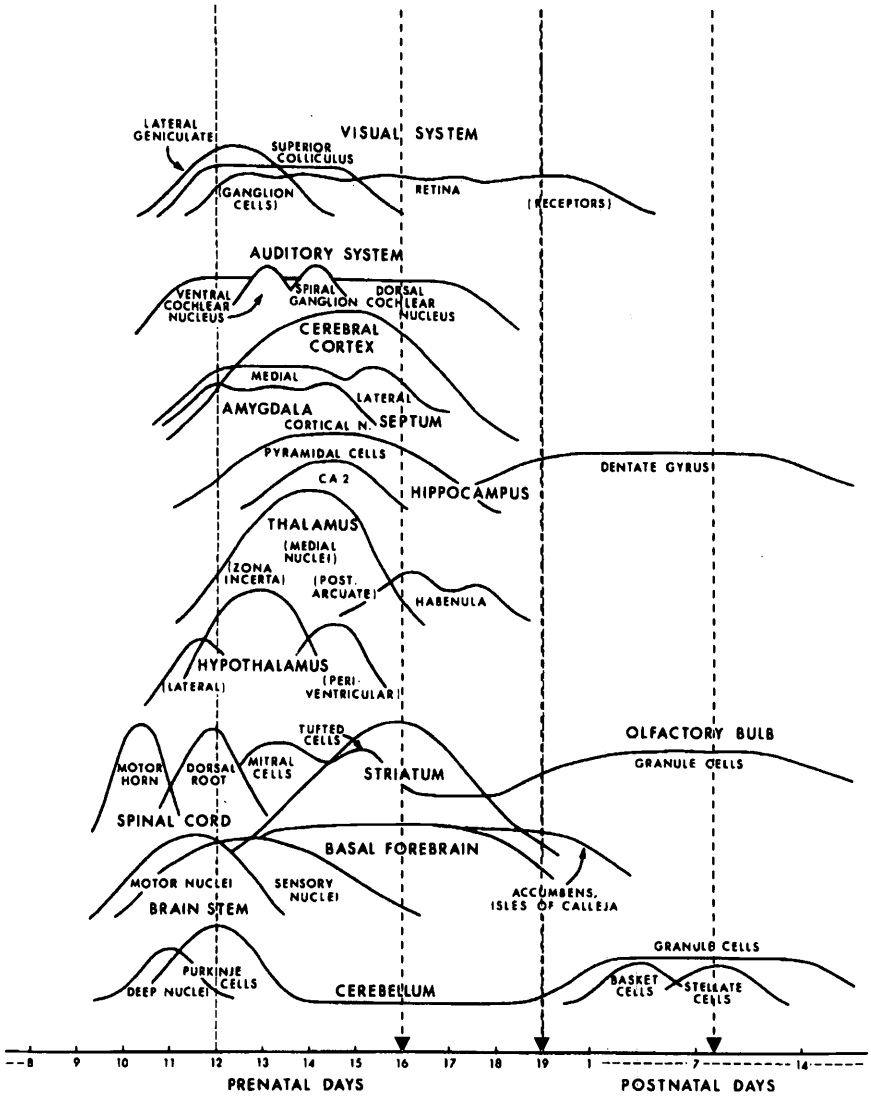

Figure 2.2. Chronology of neurogenesis in mouse brain structures. The developmental timeline of neurogenesis in mouse brain structures is shown. The x-axis depicts the days of gestation. The horizontal slopes correspond to the period of neurogenesis of different brain regions, with the peaks representing the time of maximum neuron production. The black vertical line on E12 (embryonic day 12) represents the last day of gestation when gross external malformations can be caused by interference with cell proliferation. The red vertical line on E16 represents a stage of prenatal development when neurogenesis is occurring in many brain structures. The second red vertical line on E19 represents the approximate day of birth. The third red vertical line on P8 (postnatal day 8) corresponds to a stage of postnatal development when neurogenesis is over in most brain regions. Image adapted from Rodier. 1980.[107]

Reverse transcription polymerase chain reaction (RT-PCR) was used to determine if the gene of interest was expressed in the wild-type mouse brain. Reverse transcription (RT) was performed with M-MuLV Reverse Transcriptase (New England Biolabs) in accordance with the manufacturer's instructions. 2 μg of RNA was used in a reaction with 10x RT Buffer, Oligo(dT) primer (0.25 μg/μL), dNTP mix (10 mM), and M-MuLV Reverse Transcriptase (200 U/μL). This mixture was incubated at 42°C for one hour and then at 90°C for 10 minutes to deactivate the enzyme. The success of the RT reaction was confirmed by PCR with *Gapdh* primers. The PCR reaction contained cDNA, 5X One*Taq* Standard Reaction buffer, forward and reverse primers (10 μM), dNTP mix (10 mM), and OneTaq DNA Polymerase (0.025 U/μL) (New England Biolabs). The thermal cycler was programmed for 30 cycles of PCR amplification at 94°C for 30 seconds for denaturation; 68°C for 2 minutes for annealing; and 68°C for 30 seconds for primer extension. The PCR products were then verified by gel electrophoresis. A gel concentration of 1.2% agarose in 1X TAE buffer with 6% ethidium bromide was used. The samples were mixed with 6X nucleic acid loading buffer and loaded into the wells of the gel. Electrophoresis was conducted at 100V on a PowerPac Basic Power Supply. A UV Transilluminator was used to visualize the size of the DNA bands. A 100 base pair DNA standard ladder was used to estimate the size of the DNA bands by comparison. Following the confirmation of cDNA synthesis, PCR was performed using primers specifically designed for the genes identified from the microarray analysis. All PCR primers were designed using the Primer3 program (http://frodo.wi.mit.edu/primer3/) and synthesized by Sigma-Aldrich. PCR primer sequences for microarray genes of interest are shown in Table 2.2.

Table 2.2. PCR primer sequences for microarray genes (FC > 1.5).

Gene	Primer Sequences (5' → 3')	Amplicon Size (base pair)
Gapdh	F: TTGTGATGGGTGTGAACCAC R: GTCTTCTGGGTGGCAGTGAT	169
C4orf7	F1: GGAGCAGAGTGGAGAGTTTCA R1: CCACGGGTAGCCTTGATTTA	220
	F2: CGATAGCAGTATCAGCTAG R2: CTCTCTTAACAGGAATATTGATC	404
	F3: GTCCGCGATAGCAGTATCAG R3: CAGTTAAGGATCTGCAGTTGGA	312
Csf3r	F: TGTGCCCCAACCTCCAAACCA R: GCTAGGGGCCAGAGACAGAGACAC	236
Cd14	F: GCTCAACTTTTCCTGCGAAC R: GCAAAGCCAGAGTTCCTGAC	542
Krt80	F: AGCTCTCGAAGCCTCACAAG R: AGCATCCAGGTCCTTCTTCA	459
Arhgef5l	F: CGAAAGGCACTGGTCTCTTC R: ACATTGCTGTTGCAGTCTCG	686
Klf4	F: CCACCAGGACTACCCCTACA R: CTGTGTGAGTTCGCAGGTGT	330
Sh3d19	F: TCCTGCATCCAAATCATCAA R: GTCTTTGGTGGCACCTTTGT	885
Cdkn1b	F: CAGAATCATAAGCCCCTGGA R: TCTGACGAGTCAGGCATTTG	224
Erbb3	F: GCGAGAAGGCAAACAAAGTC R: AGGTCATCAACTCCCAAACG	642
Arid5b	F: GTTGTGGGAAGAAAGGACCA R: TCCCAGTTTTTGAGCAGCTT	916
Rbm47	F: GAGGGCGTACTCAACGTGAT R: GCCACTGTGCCAGGTTTAAT	815
Clpx	F: GCGCAGAGCTCCTCTTAGAA R: CTTCTCAGCCTCTGCTTGCT	505
Mal2	F: TACTCCGGAGCTTTCGTCTG R: CAGACCCAAACTGCAACCAT	405

2.3.6. Quantitative Real-Time Polymerase Chain Reaction (qRT-PCR)

The expression level of a select few microarray genes of interest, chosen based on functional relevance in the CNS, was quantified using the 7500 Fast Real-Time PCR System with SYBR Green reagent (Applied Biosystems). The comparative Ct method was used to analyze the data. Raw Ct values were normalized using the housekeeping genes *Hprt* (hypoxanthine phosphoribosyl transferase) and *Sdha* (succinate dehydrogenase) to obtain the ΔCt values. The ΔCt values of the samples were compared to a calibrator (adult stage sample). The relative quantity (RQ) value represents the fold change expression of each target sample compared to the reference. Reported RQ ratios are the mean of comparisons between three qRT-PCR experiments. QRT-PCR primers were designed using Primer Express 3.0 (Applied Biosystems) and synthesized by Sigma-Aldrich. QRT-PCR primer sequences for microarray genes of interest are shown in Table 2.3. Statistical significance was determined by performing a one-way analysis of variance (ANOVA) followed by Student's *t*-test.

Table 2.3. QRT-PCR primer sequences for microarray genes of interest.

Gene	Primer Sequences (5' → 3')	Amplicon Size (base pair)
Gapdh	F: CGGCCGCATCTTCTTGTG R: ACACCGACCTTCACCATTTTG	62
Pgk1	F: CAGTTGCTGCTGAACTCAAATCTC R: GCCCACACAATCCTTCAAGAA	65
Csf3r	F: GACCCCATGGATGTTGTGAAA R: GAGACTACATCAGGGCCAATGTC	83
Cd14	F: GGAAGCCAGAGAACACCA R: CTTGTTGCTGTTGCTTCTGG	130
Klf4	F: ACTATGCAGGCTGTGGCAAAA R: CCGTCCCAGTCACAGTGGTA	100
Cdkn1b	F: GGTGCCTTTAATTGGGTCTCA R: CCGGATTGTCTGACGAGTCA	80
Erbb3	F: GAGGATGGCAATGGTTATGTCA R: CCCGGGAAGAGGATGTACCT	61
Clpx	F: CCAAGATGCCAATTACAATGTAGAG R: GCACACTGCCAATCTTATCTACTT	80

2.4. Results

2.4.1. Quality Assessment of RNA Samples

In order to use buccal mucosa samples for microarray studies, total RNA was isolated from buccal cells and purified. Total RNA yield was initially assessed on the NanoDrop ND-1000 Spectrophotometer. RNA quality was then measured on the Agilent Bioanalyzer 2100 at the University Health Network Microarray Center to generate an estimate of the quality of RNA samples. The RNA Integrity Number (RIN) is a quality metric used to aid in assessment of RNA samples independent of the relative concentration. The RIN system evaluates the ribosomal RNA (rRNA) species, due to the abundance and high stability of rRNA transcripts in the cells, which is assumed to be a good indication of the mRNA species (the RNA species of interest). Even though it is a reasonable assumption, it is not yet possible to determine ahead of time whether a microarray experiment will work or not. Rather, the RIN offers repeatability and quality control to determine whether to use a sample or to discard it. Conventionally, a RIN of 7 or greater is required for a sample to be used for microarray studies.[108] However, previous studies have used samples with RIN values as low as 5.2 in generating microarray data of reasonable quality.[109]

The RIN for each buccal mucosa sample is shown in Table 2.1. The mean RIN for all experiments is 7.1 (range, 5.9 to 9.4). The results indicate that the quality of RNA derived from buccal cells varies between different samples (see example in Figure 2.3). Some samples (from participants C8, C9, D8, B1, B8, C7, and C10) did not meet the threshold (RIN value of 7) and were deemed of "borderline quality", but were considered

to be of sufficient quality to be included in microarray experiments. The samples were fluorescently labelled and hybridized onto the BeadChip array. A list of genes showing differential expression in probands compared to healthy controls resulted from statistical analysis of the array. In this sense, microarray experiments were successfully performed with the buccal RNA from all participants listed in Table 2.1.

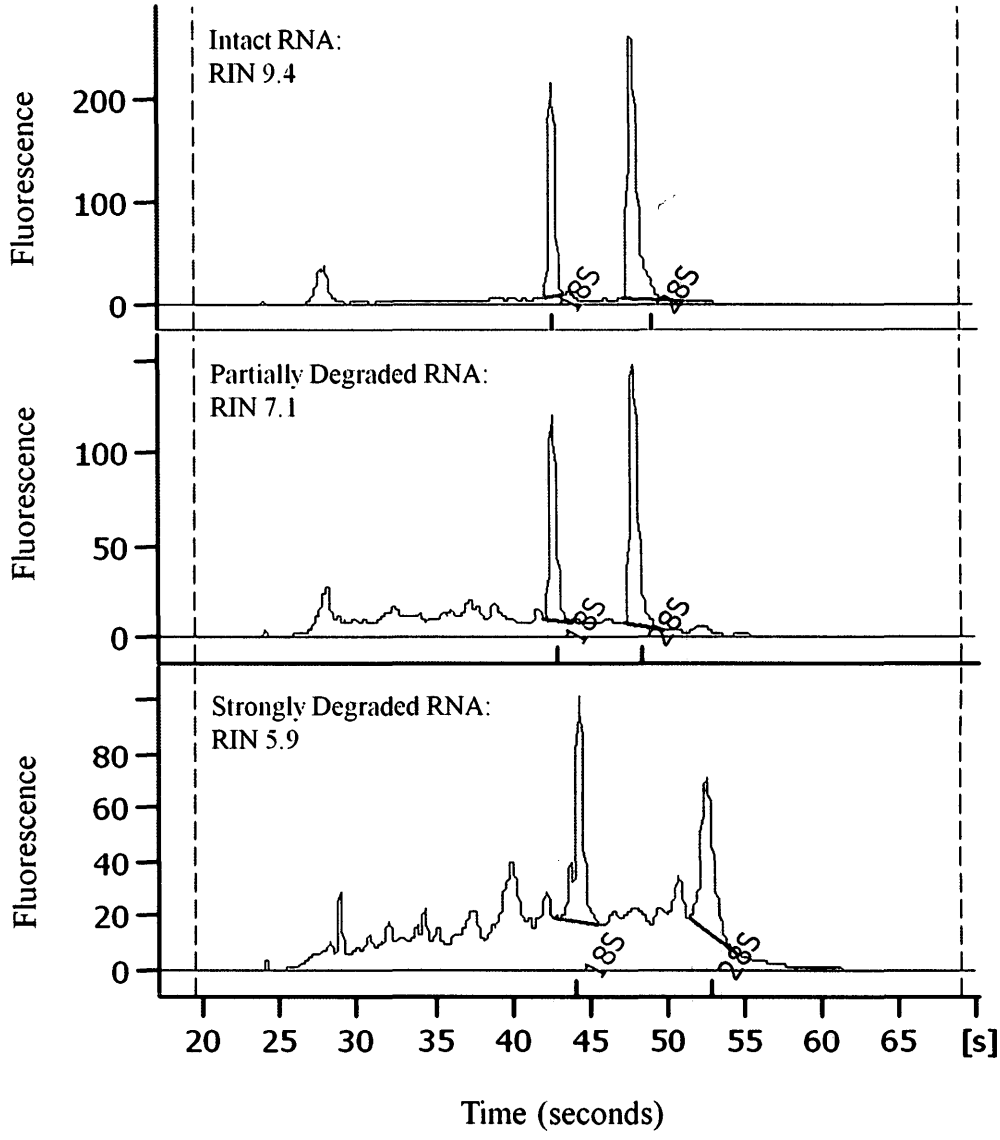

Figure 2.3. Quality assessment of buccal RNA samples. The integrity of buccal cell-derived RNA from the study participants was found to be highly variable. Three RNA samples at varying stages of intactness are shown in this figure. The top electropherogram indicates that buccal RNA can be of high quality (RIN 9.4), or as the middle electropherogram shows, of acceptable quality (RIN 7.1). Conversely, the bottom electropherogram reveals that buccal RNA from a different sample was strongly degraded (RIN 5.9) since there is a series of peaks shown at low retention times, while the two ribosomal RNA peaks are of lower intensity. RIN – RNA Integrity Number.

2.4.2. Expression of Differentially Regulated Genes in the Central Nervous System

Learning of the expression of the differentially regulated genes in the central nervous system (CNS) is of importance because autism is a neurodevelopmental disorder – a dysregulation of these genes in peripheral tissues would imply a similar deviation in the brain. The dataset of differentially regulated genes (see Table 2.4) between probands and healthy controls was examined to determine the percentage of genes that are expressed in the CNS. In addition to reviewing the existing literature pertaining to the differentially regulated genes, the Human Protein Atlas was used to establish protein expression profiles as well. Of the entire list of differentially regulated genes, 66.3% are expressed in the CNS and only 6.2% are expressed exclusively in peripheral tissues (Figure 2.4). There is no information about the remaining 27.5% of the differentially regulated genes either because they have yet to be characterized or have not been studied in the CNS. The fact that the majority of these genes are expressed in the CNS indicates that RNA derived from buccal cells may be used to study the CNS.

Table 2.4. Differentially regulated genes between probands and healthy controls.

Probe ID	Gene Symbol	P-value	Corrected P-value	Average Fold Change
ILMN_1725276	C4orf7	4.42×10^{-5}	2.85×10^{-2}	25.28
ILMN_2371280	CSF3R	2.62×10^{-4}	4.37×10^{-2}	6.18
ILMN_2396444	CD14	1.25×10^{-4}	3.43×10^{-2}	5.14
ILMN_1705814	KRT80	4.7×10^{-5}	2.89×10^{-2}	-2.6
ILMN_2214473	ARHGEF5L	1.55×10^{-4}	3.82×10^{-2}	-1.85
ILMN_2137789	KLF4	8.01×10^{-5}	3.2×10^{-2}	-1.76
ILMN_1653133	SH3D19	5.58×10^{-5}	2.99×10^{-2}	-1.75
ILMN_2196347	CDKN1B	1.76×10^{-4}	4.01×10^{-2}	-1.74
ILMN_1751346	ERBB3	2.05×10^{-4}	4.2×10^{-2}	-1.72
ILMN_1721626	ARID5B	1.93×10^{-4}	4.09×10^{-2}	-1.64
ILMN_3224926	RBM47	1.13×10^{-4}	3.43×10^{-2}	-1.63
ILMN_1709894	CLPX	8.04×10^{-5}	3.2×10^{-2}	-1.6
ILMN_1770653	MAL2	6.61×10^{-5}	2.99×10^{-2}	-1.52
ILMN_1768867	AP3B1	1.87×10^{-4}	4.01×10^{-2}	-1.49
ILMN_1741176	CHMP2B	6.15×10^{-5}	2.99×10^{-2}	-1.48
ILMN_1790549	TSPAN3	1.62×10^{-5}	1.98×10^{-2}	-1.46
ILMN_2043109	TOP1P2	1.94×10^{-5}	1.98×10^{-2}	-1.42
ILMN_1760593	CRY1	1.35×10^{-6}	6.48×10^{-2}	-1.41
ILMN_1849013	†	1.13×10^{-4}	3.43×10^{-2}	-1.39
ILMN_2319913	DGKA	3.35×10^{-5}	2.62×10^{-2}	-1.38
ILMN_1688295	ZNF219	2.17×10^{-4}	4.2×10^{-2}	-1.36
ILMN_1714364	PTK2	7.76×10^{-5}	3.2×10^{-2}	-1.36
ILMN_2045911	FBXO28	2.2×10^{-5}	1.98×10^{-2}	-1.36
ILMN_1718070	CASP9	1.65×10^{-4}	4×10^{-2}	-1.35
ILMN_2100689	MAP2K4	9.99×10^{-6}	1.69×10^{-2}	-1.34
ILMN_1701413	PIGQ	2.56×10^{-4}	4.37×10^{-2}	-1.34
ILMN_3225300	LOC728532	1.67×10^{-5}	1.98×10^{-2}	-1.33
ILMN_1702837	PSMD1	8.85×10^{-6}	1.69×10^{-2}	-1.33
ILMN_1679322	SH2D4A	8.11×10^{-7}	6.48×10^{-2}	-1.33
ILMN_1767747	HDAC2	2.02×10^{-4}	4.2×10^{-2}	-1.31
ILMN_1696028	ETNK1	2.32×10^{-4}	4.3×10^{-2}	-1.31
ILMN_1774427	CALCOCO1	2.37×10^{-4}	4.3×10^{-2}	-1.29
ILMN_2073289	MTSS1	1.23×10^{-4}	3.43×10^{-2}	-1.28
ILMN_2404917	AFAP1L2	1.25×10^{-4}	3.43×10^{-2}	-1.27
ILMN_2413615	LLGL2	6.22×10^{-5}	2.99×10^{-2}	-1.27
ILMN_1678957	WDR55	1.19×10^{-4}	3.43×10^{-2}	-1.26
ILMN_1722276	PAFAH1B1	1.07×10^{-4}	3.43×10^{-2}	-1.26
ILMN_2341548	MYO5B	6.85×10^{-5}	2.99×10^{-2}	-1.26
ILMN_1778236	PTPN11	1.44×10^{-6}	6.48×10^{-3}	-1.26
ILMN_1657470	YTHDF3	6.74×10^{-5}	2.99×10^{-2}	-1.25
ILMN_1789283	PPP2R5C	1.8×10^{-4}	4.01×10^{-2}	-1.24
ILMN_1812776	FBXO28	2.24×10^{-4}	4.27×10^{-2}	-1.24
ILMN_3245452	FAM149B1	9.66×10^{-5}	3.43×10^{-2}	-1.24
ILMN_2336186	LCMT1	1.86×10^{-4}	4.01×10^{-2}	-1.24
ILMN_2090123	DHX29	2.76×10^{-6}	9.32×10^{-3}	-1.24
ILMN_2246510	TSC1	1.44×10^{-4}	3.69×10^{-2}	-1.24
ILMN_1717154	AQR	1.29×10^{-4}	3.43×10^{-2}	-1.23
ILMN_1730995	AFAP1L2	8.04×10^{-6}	1.69×10^{-2}	-1.23

ILMN_3283592	LOC442609	8.53×10^{-5}	3.29×10^{-2}	-1.22
ILMN_1718813	CNKSR1	1.04×10^{-4}	3.43×10^{-2}	-1.22
ILMN_1805999	MYO1D	2.39×10^{-4}	4.3×10^{-2}	-1.22
ILMN_1729019	SEPT7	2.16×10^{-4}	4.2×10^{-2}	-1.22
ILMN_1788778	SEPT11	1.87×10^{-4}	4.01×10^{-2}	-1.22
ILMN_1676241	BCOR	1.36×10^{-4}	3.53×10^{-2}	-1.21
ILMN_1765558	NPAS2	2.66×10^{-4}	4.37×10^{-2}	-1.21
ILMN_2052717	GRAMD1C	6.65×10^{-5}	2.99×10^{-2}	-1.2
ILMN_2399877	COG5	4.69×10^{-6}	1.27×10^{-2}	-1.2
ILMN_1651278	SNIP1	2.64×10^{-4}	4.37×10^{-2}	-1.19
ILMN_3251550	PHLDA1	2.1×10^{-5}	1.98×10^{-2}	-1.19
ILMN_2113362	ARL6IP1	2.13×10^{-4}	4.2×10^{-2}	-1.19
ILMN_2149766	APPBP2	3.57×10^{-5}	2.62×10^{-2}	-1.19
ILMN_1660270	C7orf28A	1.16×10^{-4}	3.43×10^{-2}	-1.18
ILMN_1724410	USP46	5.87×10^{-5}	2.99×10^{-2}	-1.18
ILMN_1656293	GOSR2	2.58×10^{-4}	4.37×10^{-2}	-1.18
ILMN_2186137	RRAD	2.42×10^{-4}	4.3×10^{-2}	-1.17
ILMN_1758337	ZNF213	1.95×10^{-5}	1.98×10^{-2}	-1.17
ILMN_2258774	MRPL43	8.76×10^{-5}	3.29×10^{-2}	-1.17
ILMN_1697460	REEP6	2.17×10^{-4}	4.2×10^{-2}	-1.16
ILMN_2362346	TRPV4	2.68×10^{-4}	4.37×10^{-2}	-1.16
ILMN_1720476	PHF2	3.64×10^{-5}	2.62×10^{-2}	-1.16
ILMN_1652407	ZMYND8	1.51×10^{-4}	3.79×10^{-2}	-1.16
ILMN_1795704	KIAA0232	4.13×10^{-5}	2.8×10^{-2}	-1.14
ILMN_2294784	PRDM1	1.4×10^{-5}	1.98×10^{-2}	-1.14
ILMN_2192032	SRP19	2.41×10^{-4}	4.3×10^{-2}	-1.14
ILMN_3211488	LOC400061	2.67×10^{-4}	4.37×10^{-2}	-1.14
ILMN_1658759	PEX19	1.29×10^{-4}	3.43×10^{-2}	-1.13
ILMN_1750482	LOC285307	1.8×10^{-4}	4.01×10^{-2}	-1.13
ILMN_2395204	SLTM	1.1×10^{-4}	3.43×10^{-2}	-1.13
ILMN_1740213	ELOVL7	1.1×10^{-4}	3.43×10^{-2}	-1.13
ILMN_1731268	GAK	1.28×10^{-4}	3.43×10^{-2}	-1.11
ILMN_1754772	LOC650698	1.72×10^{-4}	4.01×10^{-2}	-1.1
ILMN_1720438	LOC653147	3.68×10^{-5}	2.62×10^{-2}	-1.1
ILMN_1779356	TP53	6.07×10^{-5}	2.99×10^{-2}	-1.08

Differentially regulated genes (adjusted *p*-values < 0.05) with fold changes are displayed in this table. The table is sorted by descending average fold change values. The genes that show at least 1.5-fold change (listed above the double line shown) were examined in further detail as potential autism susceptibility genes. † One differentially regulated gene does not have an official gene symbol and is recognized by probe ID (ILMN_1849013).

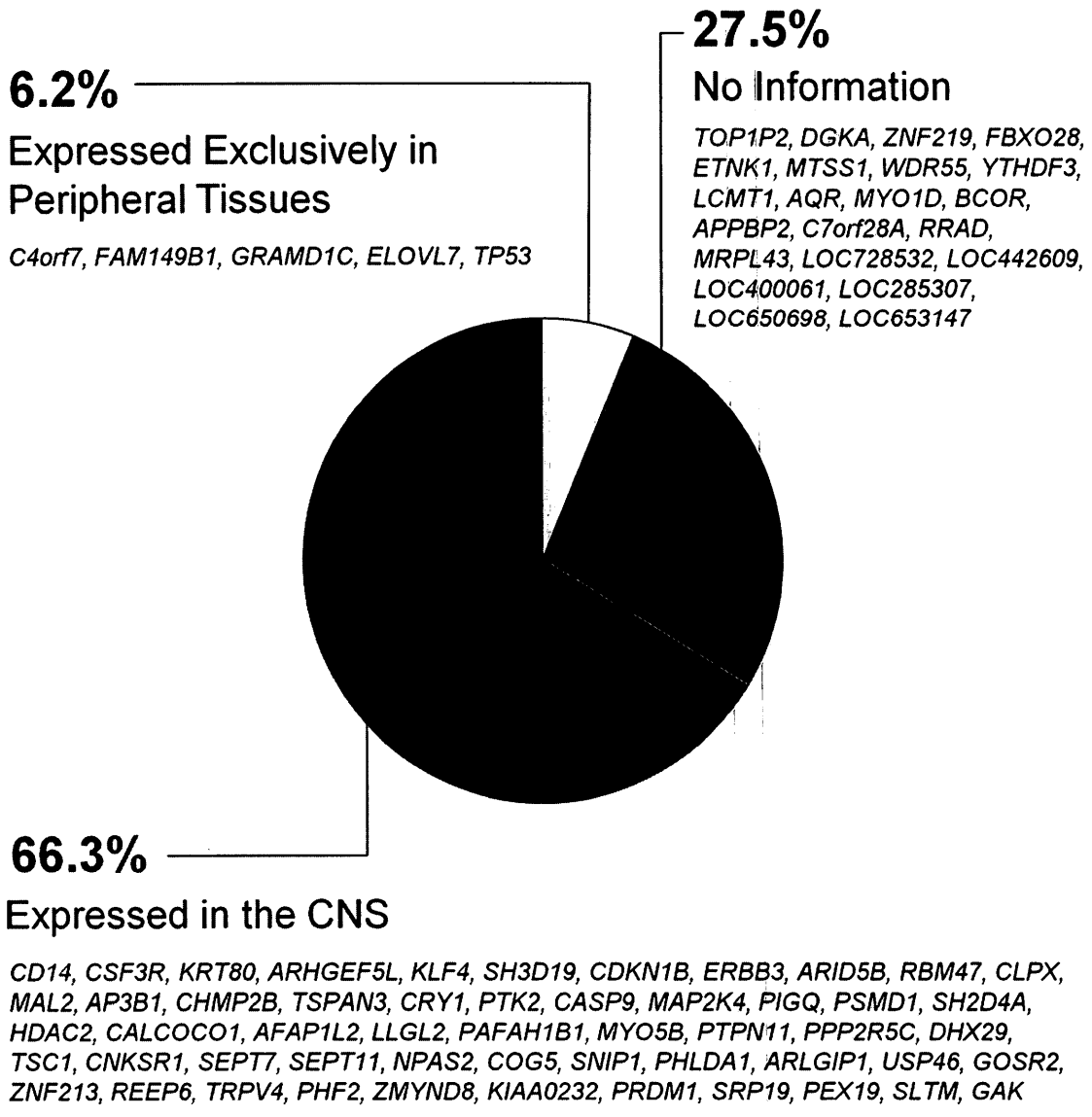

Figure 2.4. Tissue distribution of the differentially regulated genes. The entire list of differentially regulated genes is shown to be categorized into three groups: genes that are expressed in the CNS, genes that are expressed exclusively in peripheral tissues, and genes in which there is not sufficient information to determine its localization. A large majority of the identified genes has expression in the CNS, which indicates a role for these genes outside of peripheral tissues. CNS – Central Nervous System.

2.4.3. Differentially Upregulated Genes in Autism

The list of differentially regulated genes revealed three genes that were upregulated among the probands as compared to the healthy controls. These three genes are *C4orf7* (chromosome 4 open reading frame 7), *CSF3R* (colony stimulating factor 3 receptor), and *CD14* (CD14 molecule) with approximately 25-, 6-, and 5-average fold change upregulation, respectively. Surprisingly, the three upregulated genes had the highest average fold changes.

Individual Fold Changes of Differentially Upregulated Genes

Gene expression is the most fundamental level at which the genotype gives rise to the phenotype. Although these genes are uniformly upregulated among the probands, the mRNA levels were found to vary considerably from individual to individual (Figure 2.5). For instance, even though *C4orf7* had a 25-fold change average, participant A6 demonstrated extremely high expression of *C4orf7* (approximately 73-fold change) where as participant C10 showed only a slight upregulation (approximately 5-fold change) (Figure 2.5A). For *CSF3R*, the probands showed a 6-fold change average. Participant B8 demonstrated the greatest expression of this gene (17-fold change) whereas participant C11 only showed a 3-fold change difference (Figure 2.5B). *CD14* has a 5-fold change average; participant B8 once again showed the greatest expression of this gene (13-fold change) and participant C11 showed the lowest (almost 2-fold change) (Figure 2.5C). This suggests that the differences in gene expression could potentially be represented in behavioural outcomes (see section 2.4.5 for further discussion).

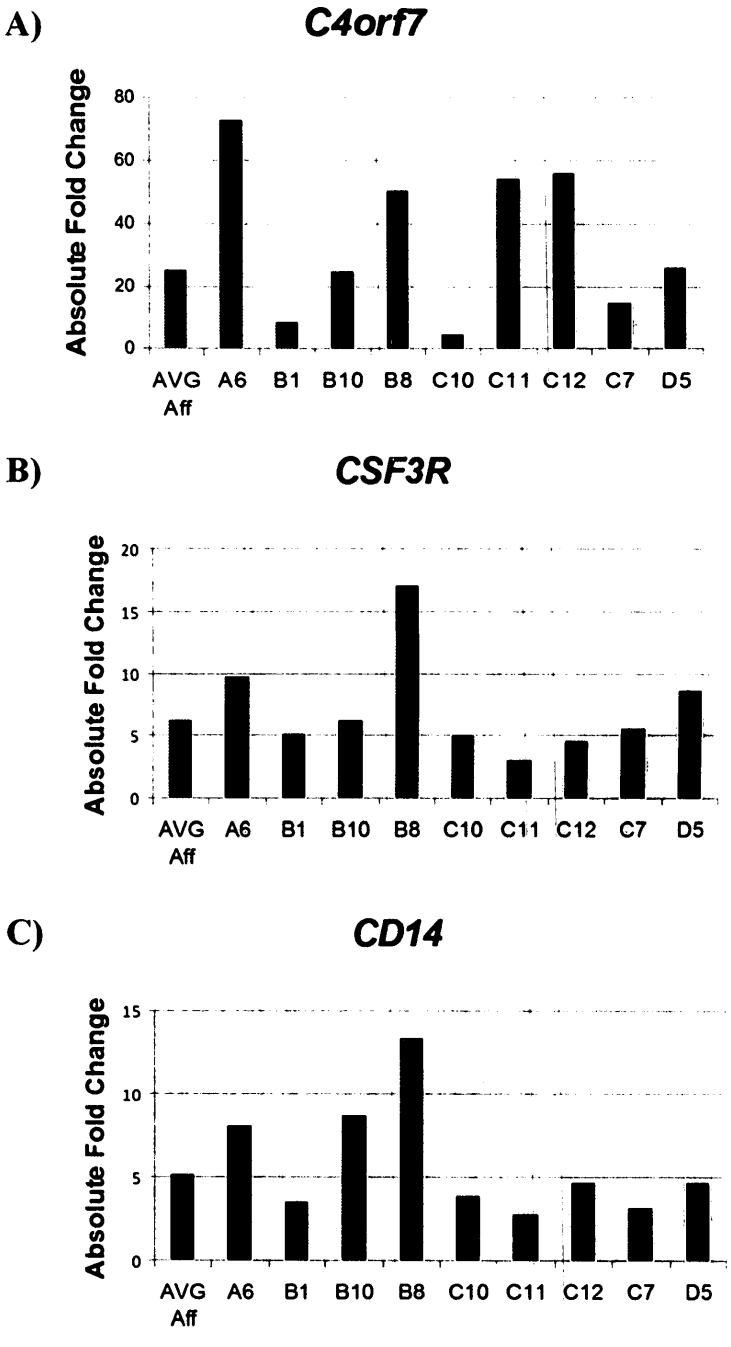

Figure 2.5. Differentially upregulated genes in autism. The relative gene expression levels of the three upregulated genes **A)** *C4orf7*, **B)** *CSF3R*, and **C)** *CD14* are shown. The red bars represent the average absolute fold change and the black bars represent the individual fold change for each proband relative to healthy controls. The participants involved in this study are identified as A6, B1, B10, B8, C10, C11, C12, C7, and D5. AVG Aff – Average of affected participants.

Molecular Functions and Biological Processes Associated with Differentially Upregulated Genes

DAVID Bioinformatics Resources revealed molecular functions and biological processes that are associated with these differentially upregulated genes (Table 2.5). The molecular function and biological process of *C4orf7*, renamed *FDC-SP* (follicular dendritic cell secreted protein), has yet to be clearly defined. However, C4orf7 is known to be expressed by follicular dendritic cells and activates leukocytes during an ongoing immune response.[110] The activated leukocytes then bind onto the surface of B-lymphoma cells, spurring speculation that C4orf7 is a modulator of B cell activity. *CSF3R* encodes a protein that is the receptor for colony stimulating factor 3, a cytokine that regulates the production, differentiation, and function of granulocytes.[111] CD14 is a surface antigen that is expressed on monocytes and macrophages. It coordinates with other proteins to mediate the innate immune response to bacterial lipopolysaccharide and also upregulates adhesion molecules at the cell surface.[112] The fact that these upregulated genes all play a role in some aspect of the immune response supports the notion of immune system irregularities in individuals with autism.

Table 2.5. Molecular functions and biological processes associated with differentially upregulated genes.

Gene Symbol	Fold Change	Molecular Function	Biological Process
C4orf7 Chromosome 4 open reading frame 7	25.3	None found	None found
CSF3R Colony stimulating factor 3 receptor	6.2	Receptor activity	Defense response Signal transduction Cell adhesion
CD14 CD14 molecule	5.1	Receptor activity Lipopolysaccharide binding	Signal transduction Response to heat Immune response Cytokine secretion

DAVID Bioinformatics Resources analysis revealed molecular functions and biological processes associated with the differentially expressed genes.

2.4.4. Differentially Downregulated Genes in Autism

The differentially downregulated genes which showed at least 1.5-fold change were investigated further to establish functional significance. There were 10 genes that exceeded the 1.5-fold change cut-off: *KRT80* (keratin 80), *ARHGEF5L* (rho guanine nucleotide exchange factor 5-like), *KLF4* (Kruppel-like factor 4), *SH3D19* (SH3 domain containing 19), *CDKN1B* (cyclin-dependent kinase inhibitor 1B), *ERBB3* (v-erb-b2 erythroblastic leukemia viral oncogene homologue 3), *ARID5B* (AT rich interactive domain 5B), *RBM47* (RNA binding motif protein 47), *CLPX* (ClpX caseinolytic peptidase X homolog), and *MAL2* (Mal, T-cell differentiation protein 2).

Individual Fold Changes of Differentially Downregulated Genes

For the downregulated genes, the average fold change among the affected group and individual fold change for each proband are graphed in Figure 2.6. Although the downregulated genes did not demonstrate fold change values as large as that seen for the upregulated genes, the mRNA levels were still found to fluctuate among probands. *KRT80* had a -2.6-fold change average among affected participants; participant B8 showed the greatest downregulation of this gene (approximately -3.6-fold change) while participant B10 had the least (approximately -1.8-fold change) (Figure 2.6A). *ARHGEF5L* had a -1.9-fold change average; participant B8 showed the greatest downregulation of this gene (-2.3-fold change) whereas participant B1 had the least (-1.3-fold change) (Figure 2.6B). For *KLF4*, the affected participants had a -1.8-fold change average; participant A6 had -2.4-fold change and participant C7 had -1.5-fold change of this gene, representing the highest and lowest downregulation of this gene, respectively

(Figure 2.6C). *SH3D19* had -1.8-fold change; participant B8 had the greatest downregulation and participant C7 showed the least downregulation of this gene, with -2.3- and -1.5-fold change, respectively (Figure 2.6D). *CDKN1B* had a -1.7-fold change average; participant B8 showed the greatest downregulation of this gene (-2.3-fold change) and participant B1 showed the least downregulation (-1.4-fold change) (Figure 2.6E). *ERBB3* had a -1.7-fold change; participant A6 showed the greatest downregulation (-1.9-fold change) and participant C10 showed the least downregulation (-1.4-fold change) (Figure 2.6F). *ARID5B* had a -1.6-fold change average; participant B8 had the greatest downregulation (-2.2-fold change) and participant D5 had the least (-1.3-fold change) (Figure 2.6G). *RBM47* had a -1.6-fold change average; participant B1 showed the greatest downregulation of this gene (-2.2-fold change) whereas participant D5 demonstrated the least downregulation (-1.3-fold change) (Figure 2.6H). *CLPX* had a -1.6-fold change average; participant B8 showed the greatest downregulation of this gene (-1.9-fold change) and participant D5 showed the least downregulation (-1.2-fold change) (Figure 2.6I). *MAL2* had a -1.5-fold change average; participant C11 demonstrated the greatest downregulation (-1.7-fold change) and participant B10 showed the least (-1.2-fold change) (Figure 2.6J). As with the upregulated genes, the differences in gene expression among probands could be represented in behavioural outcomes, such as the severity of autism symptoms (see section 2.4.5 for further discussion).

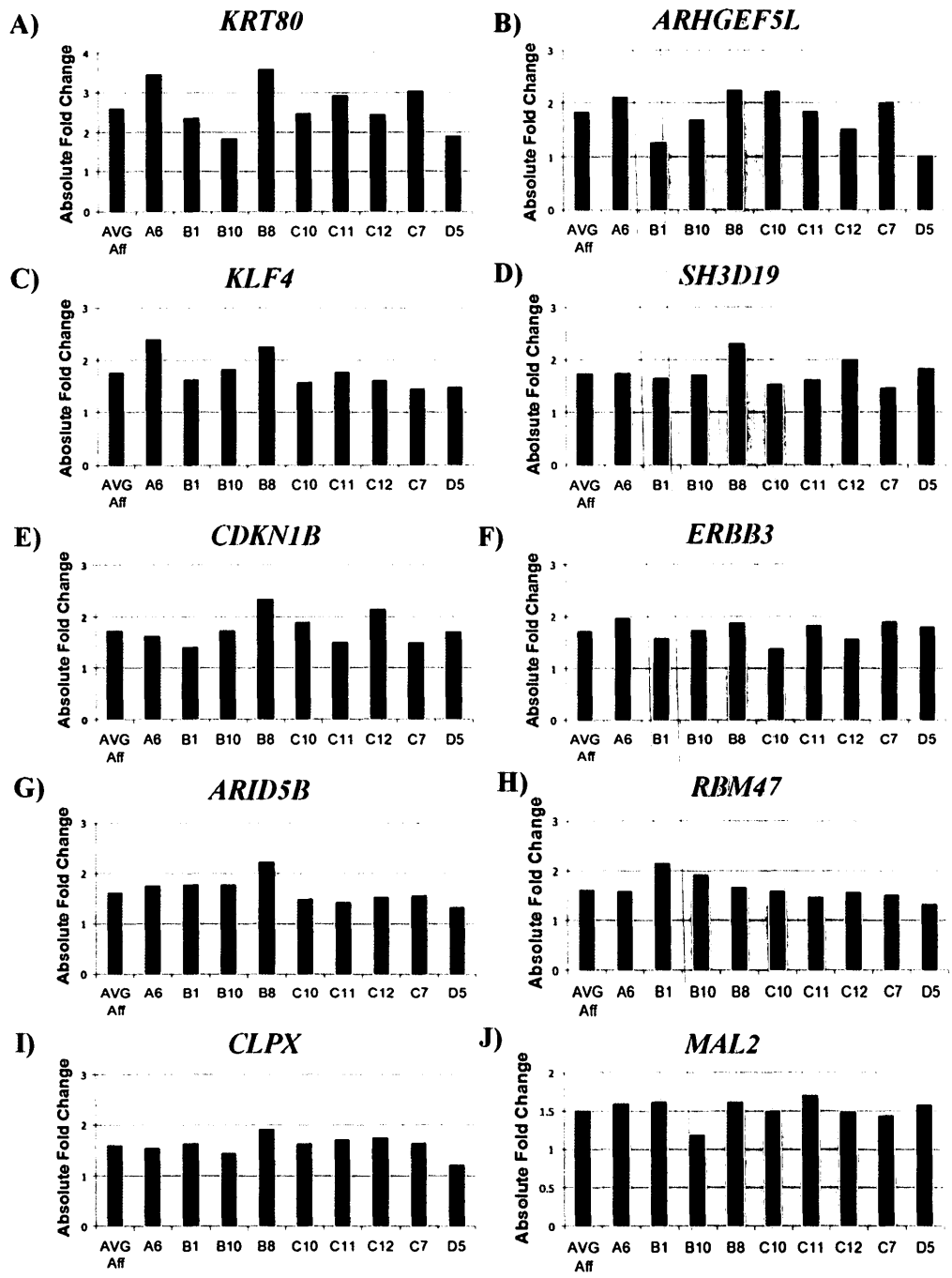

Figure 2.6. Differentially downregulated genes in autism. The relative gene expression levels of the downregulated genes that also show at least 1.5-absolute fold change. The down regulated genes are **A)** *KRT80*, **B)** *ARHGEF5L*, **C)** *KLF4*, **D)** *SH3D19*, **E)** *CDKN1B*, **F)** *ERBB3*, **G)** *ARID5B*, **H)** *RBM47*, **I)** *CLPX*, and **J)** *MAL2*. The green bars represent the average absolute fold change and the black bars represent the individual fold change for each proband relative to healthy controls. The participants involved in this study are identified as A6, B1, B10, B8, C10, C11, C12, C7, and D5. AVG Aff – Average of affected participants.

Molecular Functions and Biological Processes Associated with Differentially Downregulated Genes

DAVID Bioinformatics Resources was used to identify molecular functions and biological processes that are associated with the differentially downregulated genes meeting expression cut-off of 1.5-fold change (Table 2.6). *KRT80* encodes intermediate filament proteins that are responsible for the structural integrity of epithelial cells.[113] The molecular function and biological process of *ARHGEF5L*, renamed *ARHGEF35* (rho guanine nucleotide exchange factor 35), is not known yet though it is structurally similar to *ARHGEF5*. By structural similarity, the encoded protein may form a complex with G proteins and stimulate Rho-dependent signals. Rho GTPases play a role in numerous cellular processes initiated by extracellular stimuli that work though G protein coupled receptors. *KLF4* acts as a transcriptional activator or repressor depending on the promoter context and cooperation with other transcription factors. KLF4 regulates the expression of key transcriptional factors during embryonic development and also plays a role in preventing the differentiation of embryonic stem cells.[114] *SH3D19* plays a role in the regulation of cell morphology.[115] *CDKN1B* is an important regulator of cell cycle progression.[116] *ERBB3* encodes a member of the ERBB family of receptor tyrosine kinases. Four members of the ERBB family have been identified: ERBB1, ERBB2, ERBB3, and ERBB4. These receptor proteins, through homo- and heterodimerization, convey extracellular signals into the cell through protein phosphorylation.[117] Following activation, a signaling cascade is initiated that drives many cellular responses, including cell proliferation, differentiation, migration, and survival or apoptosis. *ARID5B* is a

transcriptional regulator that plays a role in adipogenesis and organ development.[118] RBM47 is known to bind to RNA but the biological process remains unknown. *CLPX* encodes an ATP-dependent specificity component of the Clp protease that directs the protease to specific substrates for proteolysis.[119] *MAL2* is gene that encodes a multispan transmembrane protein that is required for transcytosis, an intracellular transport pathway used to deliver membrane-bound proteins and exogenous cargos from the basolateral to apical surface.[120]

Table 2.6. Molecular functions and biological processes associated with differentially downregulated genes meeting expression cut-off of 1.5-fold change.

Gene Symbol	Fold Change	Molecular Function	Biological Process
KRT80 Keratin 80	-2.6	Structural molecule	None found
ARHGEF5L Rho guanine nucleotide exchange factor 5	-1.9	None found	None found
KLF4 Kruppel-like factor 4	-1.8	DNA and ion binding Transcription activity	Regulation of: Transcription Cell proliferation Inflammatory response
SH3D19 SH3 domain containing 19	-1.8	Protein binding	Cell morphogenesis Membrane organization Intracellular transport
CDKN1B cyclin-dependent kinase inhibitor 1B	-1.7	Protein binding Kinase activity Receptor activity	Regulation of: Cell proliferation Signal transduction Microtubule polymerization DNA damage response
ERBB3 v-erb-b2 erythroblastic leukemia viral oncogene homologue 3	-1.7	Protein binding Kinase activity Receptor activity Protein dimerization activity	Signal transduction Organ morphogenesis Schwann cell differentiation Neuron apoptosis Axonogenesis
ARID5B AT rich interactive domain 5B	-1.6	Protein binding DNA binding Transcription activity	Regulation of transcription Organ morphogenesis Signal transduction
RBM47 RNA binding motif protein 47	-1.6	RNA and nucleotide binding	None found
CLPX ClpX caseinolytic peptidase X homolog	-1.6	Protein binding Peptidase activity	Protein folding Proteolysis
MAL2 Mal, T-cell differentiation protein 2	-1.6	Protein binding	None found

DAVID Bioinformatics Resources analysis revealed molecular functions and biological processes associated with the differentially expressed genes.

2.4.5. Genotype-Phenotype Profiles of Probands

Although clinical assessments of the probands were performed by different professionals, thereby adding variability to the available behavioural data, an attempt to establish the genotype-phenotype profiles of the probands was made. The goal was to determine if there are any correlations between the level of mRNA of dysregulated genes and severity of autism symptoms on an individual basis. The genotype information was obtained from the microarray results. The expression levels of the three upregulated genes (*C4orf7*, *CSF3R*, and *CD14*) and 10 downregulated genes (*KRT80*, *ARHGEF5L*, *KLF4*, *SH3D19*, *CDKN1B*, *ERBB3*, *ARID5B*, *RBM47*, *CLPX*, and *MAL2*) comprise the genotype profile of probands. The probands that had an above-average expression of the upregulated and downregulated genes were denoted with red and green markings, respectively (Figure 2.7). The phenotype profile of the probands was established based on the obtained clinical assessment reports (with the exception of participant D5). Reported symptoms were organized into the three main domains of autism: speech impairments, social interaction deficits, and restricted and repetitive behaviours (RRBs). Other characteristics that are commonly associated with autism were also taken note of, such as motor impairment, developmental regression, resistance to change, and anxiety. The combined genotype-phenotype profile is shown in Figure 2.6. Based on observations, the data suggest a possible relationship between the number of symptoms in the speech and social interaction domains and the mRNA levels of the upregulated genes, especially with *C4orf7*. For instance, participants A6, B8, C11, and C12 demonstrated an above-average expression of *C4orf7* and displayed at least five symptoms of deficits in social

interaction. Additionally, participant A6 revealed an approximately 73-fold change upregulation of *C4orf7* and demonstrated the most speech impairments in comparison to all the probands. In contrast, participants B1 and B10 have a sub-average expression of *C4orf7* and showed less than five symptoms of deficits in social interaction. However, due to the variability between clinical assessment reporting formats and type of reported symptoms, as well as lack of assessment report from participant D5, this observation remains tentative and preliminary.

PHENOTYPE PROFILE	A6	B1	B10	B8	C10	C11	C12	C7	D5†
Speech Impairments					S				
Language Delay	X	X	X			X	X	X	
Echolalia	X	X	X	X			X	X	
Odd Intonation	X						X		
Babbles / Use of jargon	X					X		X	
Use of scripted language	X	X						X	
Social Interaction Deficits					S				
Weak eye contact	X	X		X		X	X	X	
Flat facial affect	X	X				X	X	X	
Limited use of gestures	X			X				X	
Limited peer interaction	X	X	X			X	X	X	
Failed orientation response to name	X			X		X	X	X	
No spontaneous / imaginative play	X	X		X		X			
Deficits in joint attention	X			X		X	X	X	
Restricted and Repetitive Behaviours									
RRB with objects									
Lines up / Stacks objects	X					X		X	
Spins / Wobbles	X					X	X	X	
Stereotyped routines with objects	X								
RRB with body			X						
Rocks / Spins	X					X	X	X	
Flaps / Wiggles		X				X	X		
Toe walking				X					
Sensory behaviour									
Gaze fixation		X				X			
Sensitivity to sounds			X					X	
Other Symptoms									
Motor Impairment		X	X			X	X		
Developmental Regression	X	X				X	X	X	
Resistant to Change	X	X	X	X	X			X	
Anxiety		X				X			

GENOTYPE PROFILE	A6	B1	B10	B8	C10	C11	C12	C7	D5
C4orf7									
CSF3R									
CD14									
KRT80									
ARHGEF5L									
KLF4									
SH3D19									
CDKN1B									
ERBB3									
ARID5B									
RBM47									
CLPX									
MAL2									

Figure 2.7. Genotype-phenotype profiles of probands. The symptoms are organized into speech impairments, social interaction deficits, RRBs, as well as other symptoms common to autism. The genotype profile is shown for 13 differentially regulated genes (≥1.5-absolute fold change). The red and green rectangles represent above-average expression. S – Severe impairment; RRB – Restricted and Repetitive Behaviours. † Clinical assessment report was not obtained from this participant.

2.4.6. Signaling Pathways Dysregulated in Autism

The dataset of differentially regulated genes between probands and healthy controls was analyzed using DAVID Bioinformatics Resources to identify statistically enriched signaling pathways. The significance of the association between the dataset and the signaling pathway was measured in two ways: i) a ratio of the number of molecules from the dataset that map to the signaling pathway divided by the total number of molecules that map to the signaling pathway and ii) Fisher's exact test was used to calculate a *p*-value determining the probability that the association between the genes in the dataset and the signaling pathway is explained by chance alone.[121] A *p*-value of less than 0.05 was considered significant. Two signaling pathways of interest were revealed: the ERBB signaling pathway and circadian rhythm. The downregulated genes *ERBB3*, *PTK2*, *MAP2K4*, and *CDKN1B* were found to participate in the ERBB signaling pathway (Figure 2.8). The downregulated genes *CRY1* and *NPAS2* are involved in circadian rhythm (Figure 2.9).

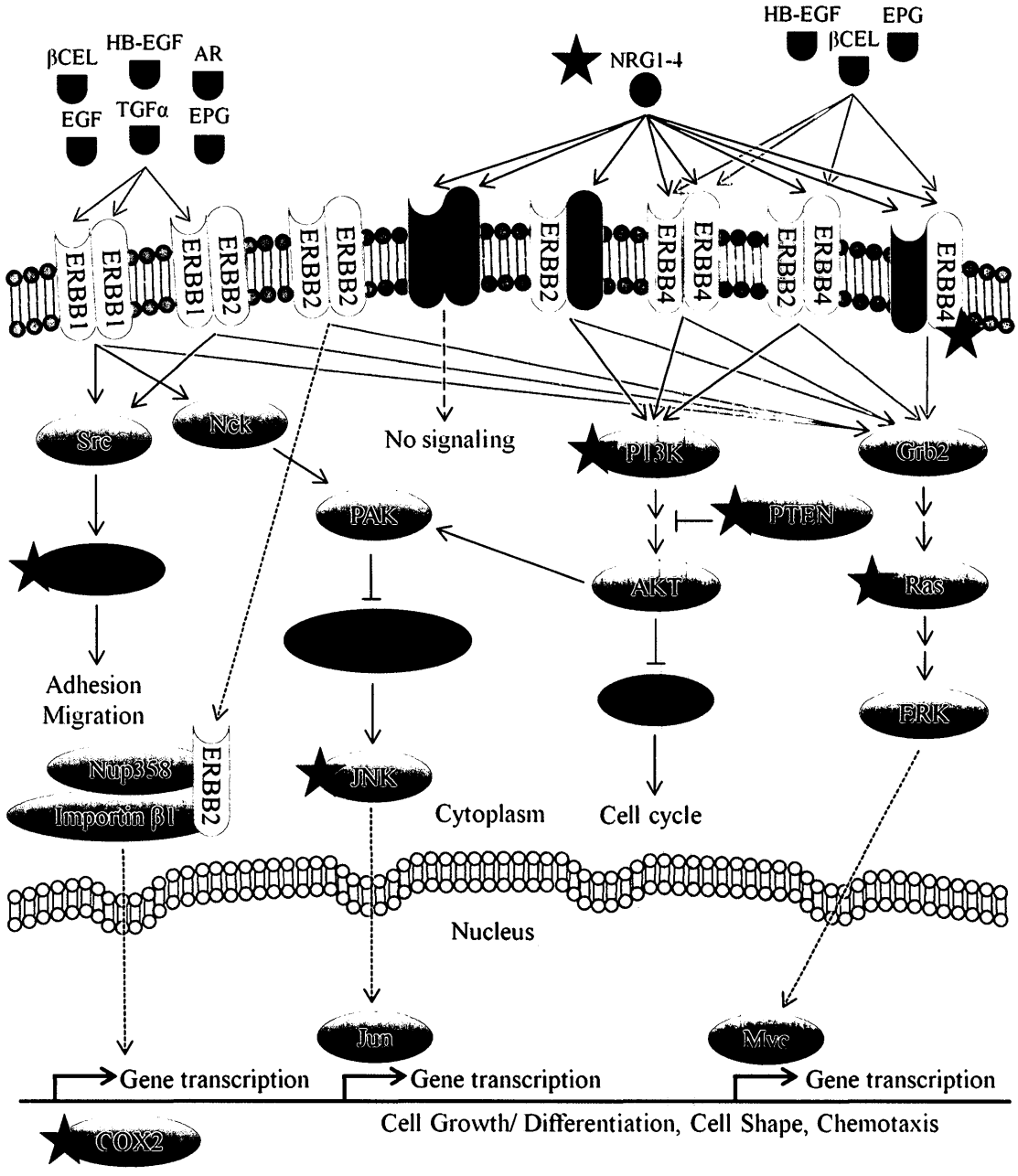

Figure 2.8. Dysregulation of the ERBB signaling pathway in autism. The ERBB receptor tyrosine kinase family (ERBB1-4) are activated following ligand binding and receptor dimerization. Ligands can display receptor specificity or bind to one or more receptors. ERBB2 lacks a known ligand. ERBB3 lacks several key conserved amino acid residues but has been shown to possess catalytic activity through transautophosphorylation. ERBB proteins are also present in the nucleus as transcription regulators. For instance, ERBB2 interacts with Nup358 and importin β1 and translocates to the nucleus where it modulates the transcription of multiple downstream genes including cyclooxygenase 2 (COX2). The differentially downregulated genes *CDKN1B*, *ERBB3*, *PTK2*, and *MAP2K4* participant in this pathway (denoted in green). Red stars denote previously described neuropsychiatric susceptibility genes.

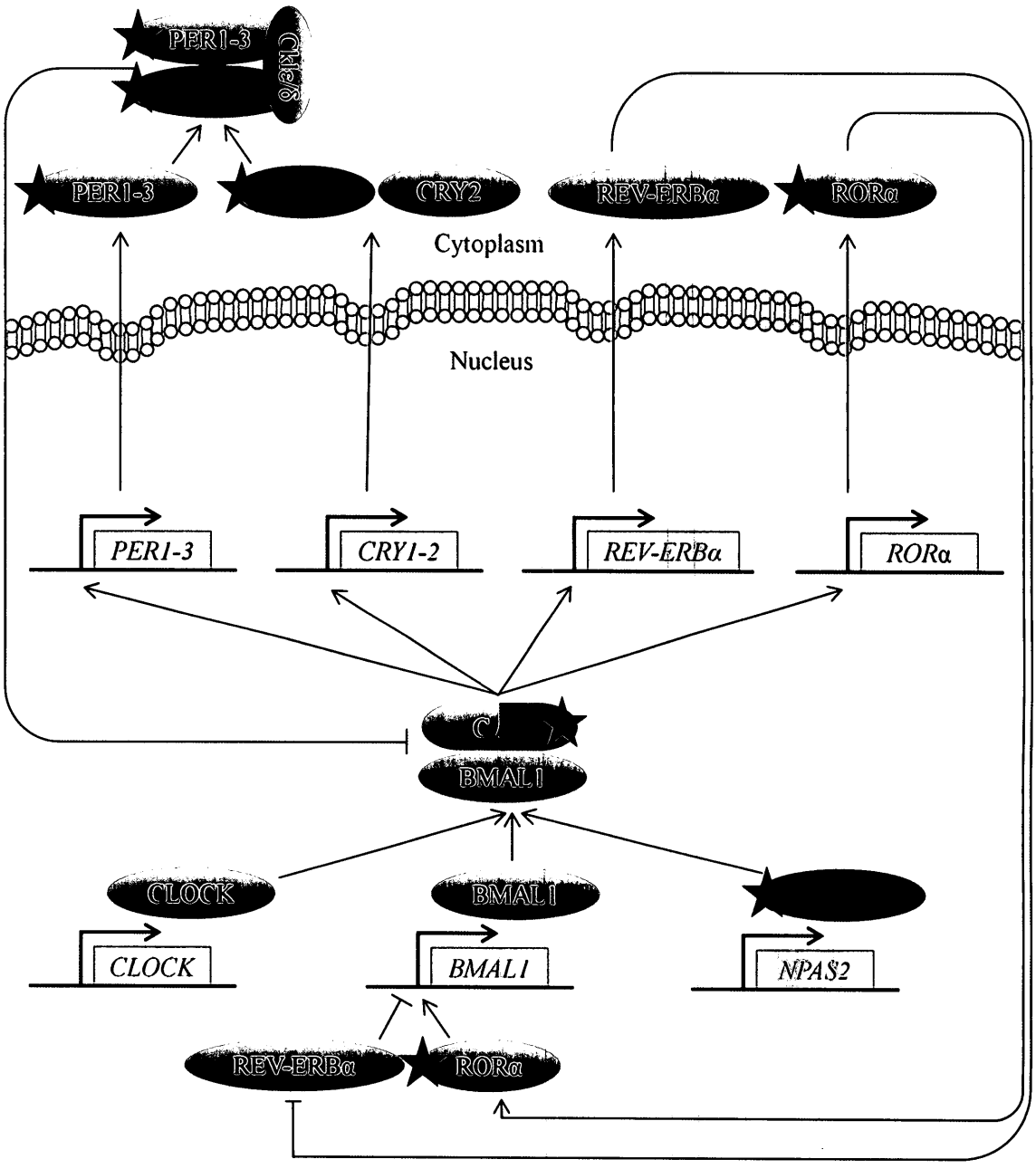

Figure 2.9. Dysregulation of circadian rhythm in autism. Circadian rhythm involves a molecular network of genes and proteins that control biological timing. CLOCK and NPAS2 heterodimerize to BMAL1 to activate the transcription of genes, including *PER1-3*, *CRY1-2*, *REV-ERBα*, and *RORα*. The gene products of *PER*, *CRY*, and *REV-ERBα* operate via a PER/CRY heterodimer that inhibits the effects of the CLOCK/BMAL1 or NPAS2/BMAL1 heterodimer while REV-ERBα represses the expression of BMAL1. Protein turnover eventually releases the genes from repression and the cycle starts over. The differentially downregulated genes CRY1 and NPAS2 participate in this pathway (denoted in green). Red stars denote previously described autism susceptibility gene candidates.

2.4.7. Temporal Expression of Potential Susceptibility Genes for Autism

The temporal expression level of identified potential susceptibility genes for autism was evaluated in the wild-type mouse. The differentially regulated genes meeting the 1.5-absolute fold change threshold were initially evaluated as potential susceptibility genes for autism. Polymerase chain reaction (PCR) was used to assess the expression of *C4orf7*, *Csf3r*, *Cd14*, *Krt80*, *Arhgef5l*, *Klf4*, *Sh3d19*, *Cdkn1b*, *Erbb3*, *Arid5b*, *Rbm47*, *Clpx*, and *Mal2* in the mouse brain during embryonic day 16 (E16), embryonic day 19 (E19), postnatal day 8 (P8), and adult stage. With the exception of *C4orf7*, all of the genes were found to have positive expression in the mouse brain (Figure 2.10). The absence of *C4orf7* expression in the brain is supported by the existing literature.[110] Six genes (*Csf3r*, *Cd14*, *Klf4*, *Cdkn1b*, *Erbb3*, and *Clpx*) were then chosen as potential susceptibility genes for autism because of their functional relevance in the nervous system. As previously discussed, *Csf3r* is a cytokine that regulates the production, differentiation, and function of granulocytes. *Cd14* is a surface marker present on monocytes and plays a role in triggering cytokine secretion. *Klf4* is a regulator of key transcription factors during embryonic development. *Cdkn1b* is a cell cycle regulator. *Erbb3* is involved in biological processes such as Schwann cell differentiation and axonogensis. *Clpx* directs the Clp protease to specific substrates for proteolysis and is also a putative target of *Foxp2*, an autism susceptibility gene.

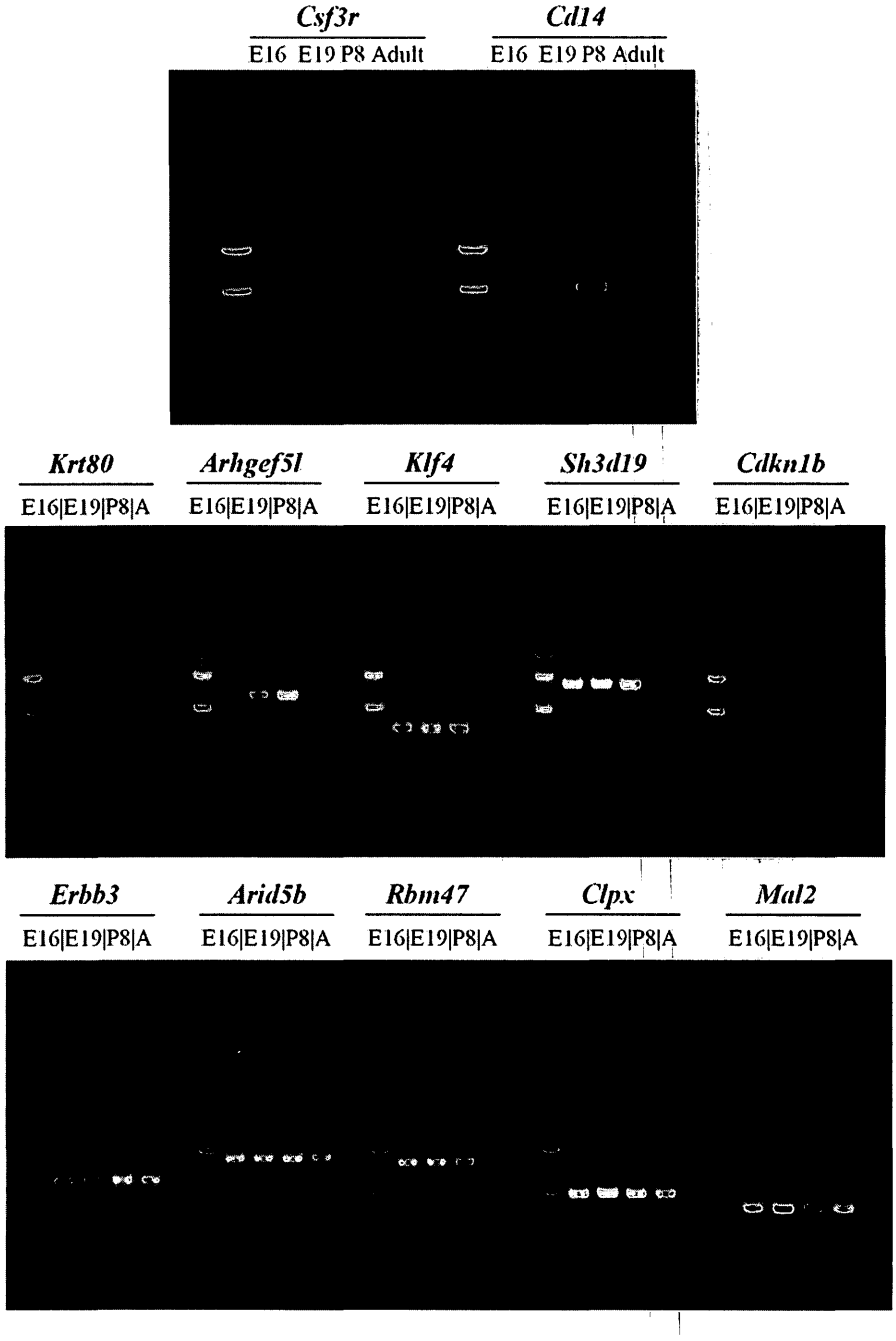

Figure 2.10. Positive expression of twelve differentially regulated genes (≥1.5-fold change) in the wild-type mouse brain. Polymerase chain reaction followed by agarose gel electrophoresis was used to assess expression. *Csf3r* (236 bp), *Cd14* (542 bp), *Krt80* (459 bp), *Arhgef5l* (686 bp), *Klf4* (330 bp), *Sh3d19* (885 bp), *Cdkn1b* (224 bp), *Erbb3* (642 bp), *Arid5b* (916 bp), *Rbm47* (815 bp), *Clpx* (505 bp), and *Mal2* (405 bp) demonstrate positive expression in the mouse brain during embryonic day 16 (E16), embryonic day 19 (E19), postnatal day 8 (P8), and (A) adult stage.

Quantitative real-time polymerase chain reaction (qRT-PCR) was used to quantify the expression of *Csf3r*, *Cd14*, *Klf4*, *Cdkn1b*, *Erbb3*, and *Clpx* to determine whether they are more highly expressed during early or late development. The adult sample was used as a reference to compare expression levels. As seen in Figure 2.11, *Csf3r* had the highest expression during the adult stage, followed by P8, and then E19 and E16. The level of gene expression was found to be statistically significant between all stages except E19 and P8. *Cd14* had the highest expression during E16 which gradually decreased through E19 to the adult stage. The expression levels of *Cd14* were found to be significantly different between E16 and P8, as well as E16 and adult. *Klf4* had the highest expression during E16 and E19, which were not significantly different from each other. The expression of *Klf4* declined as development progressed. The expression of *Klf4* at E16 and E19 was found to be significantly different from that at P8 and adult. *Cdkn1b* is most highly expressed at E16 with declining expression through E19 and P8. The expression levels of *Cdkn1b* were significantly different from each other except between E19 or P8 and adult. *Erbb3* was found to be most highly expressed during adulthood, which was significantly different from E16, E19, and P8. *Clpx* had the highest expression during E16 which gradually declined from E16 through adulthood. All stages were found to be statistically significant except P8 and adult. In summary, it appears that *Cd14*, *Klf4*, *Cdkn1b*, and *Clpx* have clear roles during prenatal development, as illustrated by the higher levels of expression during E16 and E19 relative to the postnatal stages. Although *Erbb3* seems to be more highly expressed in the adult stage, the E16 stage has the second highest level of expression, which might indicate some functional significance for brain

development at E16. *Csf3r* has the lowest expression during E16 and demonstrates gradual increase in expression after this stage, which suggests a greater need for this gene as maturation of the progeny proceeds.

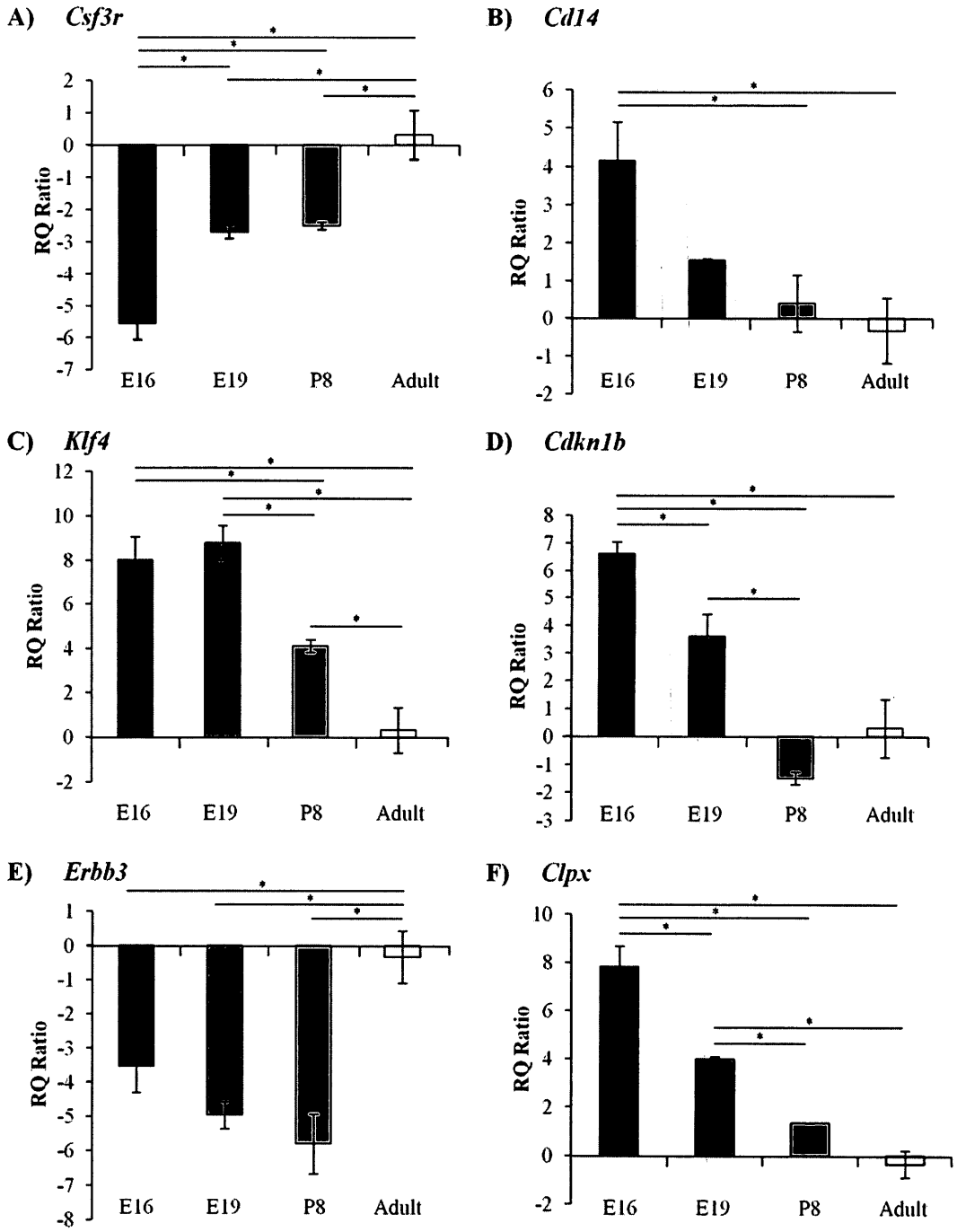

Figure 2.11. Expression of identified potential susceptibility genes for autism during prenatal and postnatal development in mouse brain. Quantitative real-time polymerase chain reaction was used to assess expression. **A)** *Csf3r* and **E)** *Erbb3* expression is highest in the adult stage. **B)** *Cd14*, **D)** *Cdkn1b*, and **F)** *Clpx* demonstrate the highest expression at E16. **C)** *Klf4* shows equally high expression during E16 and E19. *p<0.05; Error Bars: ±SEM.

2.5. Discussion

2.5.1. Quality Assessment of RNA Samples

The results presented here are an evaluation of buccal RNA samples as an efficient and non-invasive means of sample collection for evaluating relative gene expression changes. The quality of RNA from buccal cells was found to vary between samples even though swabbing pressure and technique was perceived to be consistent. In some cases, the RNA was more prone to degradation although the same procedure was precisely followed during each sample collection. By way of explanation, perhaps it is because the cells recovered from the mouth are superficial ones that are in the process of apoptosis, which may cause damage to the genetic material. The existing literature indicates that approximately 25% of buccal cells collected from healthy subjects with non-inflammatory mucosa are apoptotic.[122] Another retrospective consideration is that lifestyle habits can cause differences in exfoliation of oral mucosa. The oral flora is extremely diverse and subject-specific based on individual habits, including diet, eating mannerisms, and brushing habits.[123] All of these factors can lead to unequal rates of DNA/RNA damage. Nevertheless, despite the inconsistency in RNA quality, total RNA from buccal cells was found to be of sufficient quality to be used in successful microarray experiments and a list of genes showing differential expression between probands and healthy controls was produced (Table 2.4).

2.5.2. Expression of Differentially Regulated Genes in the Central Nervous System

Further analysis of the dataset of differentially regulated genes in autism revealed that the majority of these genes (66.3%) are expressed in the CNS (Figure 2.4). This is a strong indication that RNA from buccal cells can be used to study the human brain and also strengthens the idea that autism can be studied using easily accessible peripheral tissues, which may reflect abnormalities that are also present in the brain. A few differentially regulated genes have restricted expression in peripheral tissues, including *C4orf7*, a 25-fold upregulated gene in the autism group. Yet, despite the absence of *C4orf7* expression in the CNS, there is increasing evidence for the involvement of other body systems in autism, such as the gastrointestinal and immune systems, which suggests that peripheral tissues may also be affected. Identification of consistent autism susceptibility genes from buccal cells could aid in the efforts of learning more about the related physiological symptoms associated with autism, such as gastrointestinal irregularities and immune system abnormalities.

2.5.3. Differentially Upregulated Genes in Autism: Evidence of Altered Immune Response and Possible Genotype-Phenotype Correlations

Emergence of a common biological theme among the differentially upregulated genes affirms reports of immune response abnormalities in autism and lends insight into possible genotype-phenotype relationships.

The dataset of differentially regulated genes revealed three upregulated genes (*C4orf7*, *CSF3R*, and *CD14*) that were all found to play a role in the immune system.

Dysregulated immune function is a recurrent finding in autism. Results from numerous studies include evidence of brain reactive antibodies; skewed T cells numbers; altered cytokine levels in the brain, cerebrospinal fluid, and periphery; and altered function of innate immune cells such as monocytes and NK cells.[124-128] Here, the results of the present study also support the notion of immune system irregularities in autism.

C4orf7 (Chromosome 4 open reading frame 7)

Elevated C4orf7 mRNA levels may be triggered by the greater quantities of activated B cells in autism and provides further evidence that the activation status of immune cells is altered in autism.

Recent findings on C4orf7 (renamed FDC-SP, but hence forward still referred to as C4orf7), a secreted protein expressed by follicular dendritic cells (FDCs), reveal a distinctive expression pattern within the immune system and the ability to bind specifically to activated B cells.[110] FDCs are the stromal cells located in the germinal centre, a dynamic site within lymph nodes where antigen-activated B cells rapidly proliferate and differentiate, generating plasma cells (which produce antibodies) and memory B cells. A prior study examining the characteristics of whole blood immune cell subpopulations found that B cells were 20 – 25% higher in children with autism compared with typically developing children.[129] Notably, there were significant differences in activated B cell subsets in autism, while the number of immature B cells was not different. These data suggest that B cells are increased in autism, and furthermore, it is preferentially the activated and mature B cells which differ from controls. To put the findings of the present study into perspective with the existing

literature, it would be reasonable to expect an increase in expression of *C4orf7* in response to greater quantities of activated B cells, which would be consistent with its function as a modulator of B cell activity.[130]

A connection exists between tumour necrosis factor alpha (TNF-α), C4orf7 mRNA levels, and behavioural outcomes in autism.

As stated in the results section, participants A6, B8, C11, and C12 had an above-average expression of *C4orf7* and also demonstrated at least five symptoms indicative of social interaction deficits. Notably, participant A6 showed a 73-fold change upregulation of *C4orf7* and displayed the most number of symptoms in the speech and social interaction domains, in addition to developmental regression.

Brief exposure of FDC-like cell lines to tumour necrosis factor alpha (TNF-α) has been shown to induce high-level expression of C4orf7.[110] Elevation of CSF levels of TNF-α have been significantly higher in patients with autism demonstrating regression of previously acquired language skills, in comparison to typically developing or neurological disease controls.[126,131] Among the probands of this study, participants A6, B1, C11, C12, and C7 exhibited developmental regression and of which, participants A6, C11, and C12 had an above-average upregulation of *C4orf7* (Figure 2.7). It is an intriguing observation that participants A6, C11, and C12 also showed the highest fold changes in *C4orf7* expression, which were approximately 73-, 54-, and 56-fold change, respectively (Figure 2.5A). The results are suggestive of a connection between C4orf7 mRNA levels and developmental regression in autism.

The findings of this study suggest that elevated levels of TNF-α, which are associated with developmental regression of linguistic skills in autism, could trigger high-level expression of *C4orf7*. In turn, *C4orf7* upregulation could be correlated with more general aberrant behaviours relating to autism, in addition to developmental regression. Admittedly, conclusions cannot be reached based on this study alone, due to a lack of unified reporting format for the clinical assessments. The connection between *C4orf7* expression and behavioural outcomes in autism warrants further investigation; however, if such a relationship exists, *C4orf7* may be potentially used a biomarker for autism symptoms.

CSF3R (Colony stimulating factor 3 receptor)

Higher levels of colony stimulating factor 3 (CSF) production is associated with better cognitive and adaptive function in children with autism but higher expression of its receptor, CSF3R, is not correlated with more favourable developmental outcomes.

The current study demonstrated that *CSF3R* expression is greatly increased in children with autism compared to healthy counterparts. CSF3R is the receptor for colony stimulating factor 3 (CSF3), a hematopoietic growth factor and an inflammatory cytokine that stimulates growth and differentiation of myeloid progenitor cells into monocytes and dendritic cells (DCs).[132] In addition to playing a role in increasing innate immune cell numbers during infections, CSF3 has been shown to cross the blood-brain barrier to stimulate neurite outgrowth and promote axonal regeneration.[133,134]

CSF3 has been shown to be greatly increased in individuals with autism.[60,135] Ashwood and colleagues found that the increased production of CSF3 was associated

with improved developmental outcomes and less aberrant behaviour in autism, and suggested a role of CSF3 in the improvement of cognition and adaptive function in autism.[135] In the present study, an upregulation was seen not with *CSF3*, but with its receptor *CSF3R*. An above-average over-expression of *CSF3R* exhibited by the probands of the current study did not seem to correlate with better socialization skills or fewer behavioural abnormalities and, in fact, seem to suggest the opposite.

Mechanisms of chronic inflammation in the brain may be mediated by upregulation of CSF3R.

The observation that CSF3R expression by the probands did not correlate with better developmental outcomes may be explained by the fact that CSF3 activation is also responsible for the production of reactive glial cells, astrocytes, and microglia. An activated profile of these cells is considered to be a barrier to neuroprotection.[136] For instance, in CNS diseases such as Alzheimer's disease and multiple sclerosis, increased production of astrocytes and microglia contribute to increased inflammation, and subsequent neuronal damage and neuronal loss.[137] In autism, the presence of activated microglia has been described in post-mortem brain specimens.[60,138]

The findings of this study suggest that the mechanisms of chronic inflammation could be mediated by upregulation of *CSF3R*, which would provide more receptor bindings sites for CSF3 to exert its actions as a pro-inflammatory cytokine. For example, CSF3 is a potent enhancer of TNF-α production.[139] The production of cytokines from microglia could influence neuronal survival, function, and synaptic plasticity.[140] Although CSF3 activation is involved in several beneficial effects, including neurite

outgrowth and axonal regeneration, which could lead to better developmental outcomes as explained by Ashwood and colleagues; hyperstimulation of CSF3-CSF3R signaling is likely to have deleterious effects on neurons.[135]

The blockade of CSF3R has been shown to suppress microglia activity.[141] Since microglia cell activation is a prominent feature of autism, blocking CSF3R, and thus microglia activity, may have important anti-inflammatory effects and implications for therapy. Further investigation of the role of CSF3R in immune responses and neurodevelopment in autism is warranted.

CD14 (CD14 molecule)

Dysregulated CD14 expression in children with autism may render them more susceptible to bacterial dysbiosis, and possibly result in gastrointestinal irregularities and worsen disease prognosis.

CD14 is essential for the recognition of bacterial lipopolysaccharide (LPS) present in Gram-negative bacteria.[112] It acts as a pattern recognition protein which accepts LPS from LPS-binding protein to elicit the endotoxin cellular response by reacting with various Toll-like receptors (TLRs).

The present study found an upregulation of *CD14* in autism compared to controls. This finding conflicts with the existing literature seeing that the association between autism and CD14 has been previously refuted. CD14 was investigated for its possible role in autism after reports of intestinal bacterial dysbiosis in affected children.[142] Emanuele and colleagues investigated whether adult patients with severe autism have differences in circulating endotoxin, its soluble receptor CD14, and markers of immune-inflammatory

activation (IL-1β, IL-6, and IL-10) compared with healthy controls and found no significant elevation of CD14.[143]

However, CD14 exists in two forms – either in a soluble form or anchored to the membrane of monocytes by a glycosylphosphatidylinositol tail.[144] Monocytes, precursors of macrophages and microglia, are robust cytokine producers that profoundly influence the activity of neighbouring immune cells. A prior investigation indicated significantly increased monocyte counts in children with autism compared with typically developing children.[145] It should be noted that there are three types of monocytes in human blood: the classical monocytes are characterized by high-level expression of the CD14 cell surface receptor ($CD14^{++}CD16^{-}$); the non-classical monocytes show low-level expression of CD14 and with additional co-expression of the CD16 receptor ($CD14^{+}CD16^{++}$); and the intermediate monocytes with high-level expression of CD14 and low-level of CD16 ($CD14^{++}CD16^{+}$).[146] These monocyte subsets are physiologically different and changes in their numbers have been observed in various disease conditions. In stroke patients, the number of circulating monocyte subsets is associated with disease prognosis: $CD14^{++}CD16^{-}$ monocyte numbers are positively associated with poor outcome, early clinical worsening, and higher mortality.[147] It has been suggested that in the presence of CNS inflammation, peripheral blood monocyte numbers and functions may be associated with the degree of disease severity.[18]

Furthermore, $CD14^{++}CD16^{-}$ monocytes express high levels of chemokine receptors CCR1, CCR2, CXCR2, and respond to the chemokine ligands CCL2 and CCL7.[148] In a study of transcript profiles of peripheral blood monocytes in children with

autism who exhibit aberrant behavioural symptoms following infection and other immune insults, those with persistent gastrointestinal symptoms revealed altered regulation of CCL2 and CCL7.[18] The data indicate that in a subset of affected children with gastrointestinal irregularities, dysregulated immune responses may render them more vulnerable to common bacterial infection or dysbiosis.

Unfortunately, the clinical assessment reports obtained from the probands of the present study did not place emphasis on gastrointestinal irregularities in affected individuals. Based on the results of the present study, however, a resultant hypothesis is formed: probands exhibiting an above-average expression of *CD14* (participants A6, B10, and B8) have gastrointestinal symptoms. Whether increased CD14 may contribute to the pathophysiology of inflammation and autism symptoms should be further explored in future studies.

The relationship between CSF3R and CD14 in neutrophil maintenance support previous reports of dysregulated immune responses and enhanced immune activity in autism.

Mobley and colleagues investigated the differences between the monocytes subsets to provide clues for possible distinctive functions and also revealed a relationship between CD14 and CSF3R.[149] At the earliest stage of the innate immune response, $CD14^{++}CD16^{+}$ monocytes respond directly to invading microorganisms by releasing a multifunctional cytokine that initiates the acute phase immune response and directs the production of CSF3. CSF3 then promotes the production of neutrophils and instigates $CD14^{++}CD16^{-}$ monocytes that remove them once their job is finished. Differential

expression of *CSF3R* by the CD14++CD16- monocytes suggests that the induction of this subset could be correlated with the CSF3-dependent production of neutrophils. This safety mechanism would ensure that whenever neutrophils are released, the cells required to dispose of those neutrophils are also produced. Neutrophils are capable of producing potent bactericidal agents, including reactive oxygen species, nitric oxide, and proteases that not only kill the invading pathogen, but can also inflict severe collateral damage on nearby host tissues. In the present study, all the probands (A6, B10, and B8) but one (D5) that demonstrated an above-average upregulation of *CSF3R* also had an above-average up-regulation of CD14 (Figure 2.7). Thus, it seems that the expression of *CSF3R* and *CD14* might be related. The current findings are consistent with previous reports of enhanced immune activity in autism and further indicate that a dysfunctional innate immune response may occur in a number of individuals with autism.

The data indicate that altered immune function in children with autism may be linked to disturbances in behaviour and developmental functioning.

Taken together, the data present compelling evidence for immune dysfunction in autism and may indicate that the dysregulation of genes involved in the immune response, such as *C4orf7*, *CSF3R*, and *CD14*, can modulate behaviours relating to autism. Although the data should be treated with caution until further validation can be performed, it should be noted that a number of other studies have also shown an association between altered immune profiles and autism symptoms. [128,135,150,151] More extensive longitudinal studies will be required to characterize the immunological changes associated with autism and the severity of behavioural symptoms.

2.5.4. Differentially Downregulated Genes in Autism: Evidence of Dysregulated Signaling Pathways in Autism

The differentially downregulated genes meeting expression cut-off of 1.5-fold change (*KRT80*, *ARHGEF5L*, *KLF4*, *SH3D19*, *CDKN1B*, *ERBB3*, *ARID5B*, *RBM47*, *CLPX*, and *MAL2*) demonstrated common molecular functions including DNA binding, protein binding, and transcriptional activity, and also shared common biological processes such as signal transduction and regulation of cell proliferation and transcription. However, these genes did not reveal a unifying biological theme, unlike the differentially upregulated genes which were all found to be involved in the immune response. Nevertheless, pathway analysis revealed two signaling pathways of interest from the entire list of differentially regulated genes (Table 2.4). Furthermore, it was several of these downregulated genes, and not the upregulated genes, that were found to be involved in dysregulated signaling pathways associated with autism.

The importance of identifying dysregulated signaling pathways in autism.

Many genetic linkage, cytogenetic, and candidate gene studies have implicated a multitude of loci, chromosomal hot spots, and genetic variants in autism. Despite substantial efforts to unravel the genetic basis of autism, a detectable genetic aberration is recognized in less than 10% of cases.[31] The remaining 90% of idiopathic cases represent a broad clinical spectrum that has proven exceedingly difficult to define. Yet, regardless of the wide genetic heterogeneity, there are core phenotypic features common to all cases of autism and it is conceivable to think that there are common biological pathways that are dysregulated in autism. A proposed theory posits that distinct genetic variants, each of

which occurs rarely among autism cases, disrupt the function of a core group of genes, of which perturbation to these genes is necessary to cause autism. Such a functional disruption may explain how multiple genetic factors could result in a single disorder with central neural network dysfunction.[152] The ascription of differentially regulated genes to gene regulatory networks may provide further insights into potential mechanisms contributing to autism, whereas individual gene studies are unlikely to provide a comprehensive explanation of the phenotypic variation that is characteristic of autism.

ERBB Signaling Pathway

The results from the pathway analysis show that the differentially downregulated genes *ERBB3*, *CDKN1B*, *PTK2*, and *MAP2K4* belong to the ERBB signaling pathway (Figure 2.8).

The ERBB family of receptor tyrosine kinases includes four members: ERBB1 to ERBB4.[117] These receptor proteins, through homodimerization and heterodimerization, convey extracellular signals into the cell through protein phosphorylation. Following activation, a signaling cascade is initiated that drives many cellular responses, including cell proliferation, differentiation, migration, and survival or apoptosis. The signaling diversity originating from the ERBB family of receptor tyrosine kinases is generated by the repertoire of ERBB ligands and the combinatorial properties of receptor dimers. As well, ERBB proteins may also act as transcriptional regulators. For instance, ERBB2 can interact with Nup358 and importin β1 to translocate to the nucleus where it modulates the transcription of multiple genes including cyclooxygenase 2 (COX2), which has been previously associated with autism.[153,154]

Alterations in ERBB signaling have been implicated in the underlying pathology of neuropsychiatric disorders.

Neregulins (NRGs) are growth factors that signal through ERBB receptors. NRG1-ERBB signaling is known to play crucial roles in the development and patterning of the cerebral cortex and cerebellum.[155-157] Moreover, NRG1-ERBB4 interactions can modulate gamma oscillations in the brain, which is perceived to be critical for normal cognitive functions.[158] Collectively, these studies imply that NRG1-ERBB signaling defects may alter the generation or function of neurons in the developing and adult brain. Several genes in the NRG-ERBB signaling pathway, including NRG1, NRG3, and ERBB4, have been implicated in the genetic predisposition to schizophrenia.[159-161] The symptoms of schizophrenia include altered social behaviours, which is an area of phenotypic overlap with autism. Resultant changes in neural circuitry and electrophysiology in the cerebral cortex may lead to neuropsychiatric disorders such as schizophrenia, as well as autism.

The present study revealed that ERBB3 was downregulated among the probands. ERBB3 has an inactive kinase domain but has been shown to possess catalytic activity through transautophosphorylation, thus indicating it acts as a co-receptor.[117] Given that ERBB3 and ERBB4 interact to form receptors together, it would be important to further investigate the function of ERBB3 in neurodevelopmental events.

Major signaling pathways activated by ERBB receptors.

Activated ERBB tyrosine kinase domains generate intracellular docking sites for adaptor proteins, such as Grb2 and Nck; and kinases, such as Src and

phosphatidylinositol 3-kinase (PI3K), which in turn activates the Ras/ERK, PI3K/AKT, Src/PTK2, and Nck/PAK signaling pathways.[117]

The Ras/ERK signaling pathway is initiated through ERBB receptor activation via Grb2. There are genetic mutations that have been found to increase the risk of autism and that have potentially strong links to the Ras/ERK pathway. The Ras/ERK signaling pathway has been shown to play dual roles by promoting both cell development and cell death. Zou and colleagues detected significantly enhanced Ras/ERK signaling activities in the frontal cortex of people with autism, as well as in the brain of BTBR mice that model autism.[162]

Dysregulation of the PI3K/AKT signaling is currently thought to be a contributing factor of autism. The PI3K/AKT signaling pathway (via the p85 regulatory subunit) is known to be involved in cellular functions such as cell growth, proliferation, differentiation, motility, and survival.[163] Although there is a great deal of overlap in the signaling pathways activated by the four ERBB receptors, there are also examples of preferential activation of specific pathways. For example, ERBB3 is thought to be the most efficient activator of PI3K due to the presence of multiple binding sites for p85.[164] Another component of the PI3K/AKT signaling pathway is PTEN (phosphatase and tensin homologue on chromosome 10). PTEN functions primarily as a phosphatase to antagonize PI3K signaling. Removal of PTEN results in the upregulation of P13K signaling, through the increased phosphorylation of P13K effectors such as AKT. PTEN is a tumour suppressor gene mutated in many human cancers, and has also been associated with autism.[165,166] A recent study found that within a small subset of subjects

with autism and macrocephaly (larger head circumference), 17% had mutations in the *PTEN* gene.[167] In this context, it is interesting to note that CDKN1B is a downstream target of AKT. CDKN1B was found to be downregulated among the probands of this study. In reinforcing the view that autism involves abnormal regulation of brain growth during development, Courchesne and colleagues have hypothesized that CDKN1B may be associated with early brain overgrowth, which could disrupt the organization and connectivity of the cortex and alter the balance of excitatory and inhibitory signals in the brain.[168] Furthermore, other related neurological disorders, which also harbour mutations in the PI3K pathway, are associated with autism. For instance, individuals with mutations in the tuberous sclerosis complex 1/2 (TSC1/2) – a direct substrate of AKT activity – develop tuberous sclerosis and, with a high incidence of 25-50%, also exhibit autism.[166]

Recent studies suggest that reduced cell migration may be involved in the pathogenesis of autism.[169] PTK2 has a vital functional role in neural migration, dendritic morphology, axonal branching, and synapse formation. Protein expression levels of Src and PTK2 have been shown to be significantly decreased in lymphoblasts from subjects with autism.[170] In addition, these lymphoblasts exhibit decreased migration, have increased adhesion properties, and impaired IgG production. These results suggest that Src-PTK2 signaling activity is reduced in autism; a finding that validates the current study results, as PTK2 mRNA levels were found to be significantly reduced among affected subjects.

Lastly, *MAP2K4* (mitogen-activated protein kinase 4) is another differentially downregulated gene belonging to the ERBB signaling pathway. MAP2K4 mediates

responses to various cellular stresses and inflammatory cytokines. Downstream of Nck/PAK signaling, MAP2K4 phosphorylates and activates JNK. An *in silico* study of the gene expression profile from autism implicated genes found that JNK signaling was a central hub for immune signaling in autism.[152] This work provides integrated evidence that autism susceptibility genes and other neuropsychiatric-implicated genes may converge onto central signaling pathways.

Circadian Rhythm

Another significantly enriched pathway identified from the microarray data analysis was circadian rhythm. Circadian rhythm involves a molecular network of genes and proteins that control biological timing. The identified differentially downregulated genes *CRY1* and *NPAS2* participate in this pathway (Figure 2.9).

In the generation of circadian rhythm, the positive elements of the system, CLOCK and NPAS2, heterodimerize to BMAL1 and activate the transcription of *PER1-3*, *CRY1-2*, *REV-ERBα*, and *RORα*. The encoded gene products of *PER*, *CRY*, and *REV-ERBα*, the negative elements of the system, operate via a PER/CRY heterodimer that inhibits the effects of the CLOCK/BMAL1 or NPAS2/BMAL1 heterodimer while REV-ERBα represses the expression of BMAL1. Protein turnover eventually releases the genes from repression and the cycle starts over. The time taken for the migration of proteins and mRNA to and from the nucleus broadly defines the ~ 24 hour period of this oscillatory system.

Circadian rhythm has been proposed to have multiple effects on neurological, hormonal, and gastrointestinal functions commonly dysregulated in autism. Atypical

sleep architecture has been reported in many individuals with autism.[171] Circadian hormone (melatonin) abnormalities have been found as well.[172] Dysregulation of genes involved in the circadian rhythm mechanism could have a systemic impact on individuals with autism, causing many of the related symptons that are commonly found in autism. Genetic variations in the circadian genes have been previously investigated in individuals with autism and showed a positive association with *NPAS2* and *PER1*.[173] As well, both *CRY1* and *NPAS2* have been found to be differentially regulated in severe cases of autism in a microarray study using lymphoblastoid cell lines.[174] In a different study employing lymphoblastoid cells, DNA methylation differences in RORα between subjects with autism and controls were demonstrated.[175] In addition to its role in circadian rhythm, RORα transcriptionally regulates aromatase, an enzyme that converts testosterone to estrogen. The aromatase protein is significantly reduced in the frontal cortex of autism subjects relative to controls, and is strongly correlated with RORα protein levels in the brain.[176] The results indicate that RORα, through its transcriptional target aromatase, could provide insight into the sex bias in autism. Collectively, these studies pertaining to dysregulation of circadian rhythm in autism further validate the results of the current study.

2.5.5. Temporal Expression of Potential Susceptibility Genes for Autism

Evaluation of the temporal expression level of the differentially upregulated genes Csf3r and Cd14 in wild-type mouse and speculated roles in brain development.

In this study, the expression of *Csf3r* and *Cd14* was further examined in the mouse brain throughout three stages of development (E16, E19, and P8) and the adult stage to elucidate its role in the developing CNS.

Csf3r was found to be most highly expressed in the adult stage and had lowest expression at E16 (Figure 2.11A). Similarly, Kirsch and colleagues studied the expression pattern of *Csf3r* in the mouse embryo to determine whether its expression in the adult nervous system had any correlation to the embryonic expression.[177] The authors found that in adults, *Csf3r* is expressed in neurons of many brain areas, including neurogenic regions such as the subventricular zone and the subgranular layer of the dentate gyrus. However, the expression of *Csf3r* in the developing CNS is most prominent in radial glial cells. Radial glial cells are multifunctional cells involved in many aspects of brain and spinal cord development.[178] Most intriguingly, radial glial cells have been recently perceived as the neuronal progenitors that generate neurons in the cerebral cortex.[179,180] These findings implicate the importance of *Csf3r* embryonic expression and a likely significance in the early development of the nervous system, despite the low levels of expression at E16.

On the contrary, *Cd14* was found to be most highly expressed in the mouse brain at E16 with expression gradually decreasing throughout development (Figure 2.11B). The CD14/TLR4 complex has been identified as a pathogen recognition receptor that responds to LPS, a cell wall component of gram-negative bacteria.[112] Binding of LPS to its receptor complex activates the transcription factor nuclear factor kappa B (NFκB), leading to robust increases in production and release of pro-inflammatory cytokines,

including IL-1β and TNF-α.[181] In turn, increases in pro-inflammatory cytokines activate inducible cyclooxygenases (COX) to convert arachidonic acid into prostaglandins that influence the inflammatory response and physiological outcomes.[182,183] There are an increasing number of hypotheses that try to explain the immune abnormalities in autism, including maternal infections during pregnancy, congenital viral infections, or xenobiotic-induced abnormalities.[184,185] A previous report indicated that in the first month of life, children that later develop autism have more infections than their counterparts.[186] These findings, in conjunction with the observation that children with autism are more responsive to signaling via TLRs, suggest that aberrant signaling through CD14/TLRs may participate in this disorder.[187] Inappropriate stimulation of immune responses during critical neurodevelopmental periods could contribute to alterations that are characteristic of autism.

Evaluation of the temporal expression level of the differentially downregulated genes Klf4, Cdkn1b, Erbb3, and Clpx in wild-type mouse and speculated roles in brain development

The 10 differentially downregulated genes that met the expression cut-off 1.5-fold change were found to have positive expression in the brain. Based on functional relevance in the CNS, four genes were further studied in the mouse brain throughout development: *Klf4*, *Cdkn1b*, *Erbb3*, and *Clpx*.

Klf4 was most highly expressed during prenatal development, which includes the neurogenesis period (Figure 2.11C). Embryonic neurogenesis is regulated by the activation of specific genetic programs. In the hypothalamus, a brain region that has been

previously implicated in autism, discrete populations of neurons play essential roles in homeostasis by regulating hormone secretion from the pituitary. Among these, neuronal thyrotropin-releasing hormone (TRH) populations control essential physiological roles, including energy homeostasis, autonomic function, and coordinating the hypothalamus-pituitary-thyroid axis function. The transcription factor Klf4 has been shown to be enriched in TRH neurons during hypothalamic development.[188] In murine hypothalamic cells, it has been demonstrated that Klf4 regulates Trh activity through motifs present on the Trh gene promoter.[189] Accordingly in *Klf4* null mice, hypothalamic Trh expression was downregulated at E15; resulting in diminished bioactive peptide levels. The present study examined the expression of Klf4 within a similar period of embryonic development (E16), and is able to confirm that Klf4 is a critical factor in brain development during this stage, given the high-level expression exhibited (Figure 2.10C). At the neonatal stage, although the Trh transcript levels of the *Klf4* null mice were observed to be normal, reduced peptide levels persisted.[189] TRH simulates the release of thyroid-stimulating hormone (TSH) and prolactin. TSH and prolactin responses to TRH have been compared among individuals with autism, neurological control groups, and healthy controls. Interestingly, mean TSH basal and peak levels were significantly lower in individuals with autism, suggesting hypothalamic dysfunction in autism.[190] A previous report describing the characterization of deficiencies in the TRH receptor (TRH-R1) in mice found that TRH-R1 knockout mice displayed increased anxiety-like behaviours.[191] The behavioural changes in TRH-R1 knockout mice provide evidence for the involvement of the TRH/THR-R1 system in regulating mood and shed insight into the participation of

particular circuits in autism related phenotypes. Thus, the results indicate that Klf4 is part of the differentiation program for maturation of TRH expression within the hypothalamus. Further, dysregulation of Klf4 during critical periods of prenatal development may contribute to hypothalamic dysfunction, possibly leading to altered behavioural outcomes.

Cdkn1b was found to be most highly expressed during prenatal development as well, particularly at E16 (Figure 2.11D). Some studies have reported that children with autism have larger brains than controls, but only up to approximately four years of age.[168,192] The findings provide evidence that increased brain volume results from an increased rate of brain growth occurring prior to four years of age. In cross-sectional analysis of two year olds, it was reported that children with autism had significantly enlarged gray and white matter volumes compared to the developmentally delayed group, but only white matter volumes were enlarged compared to the typically developing group.[193] Because genetic factors are most likely to underlie the abnormal regulation of brain growth during early life, Courchesne and colleagues have suggested that the linkage between autism and genes that are known to alter white matter development be further explored.[168] *CDKN1B* was proposed as a candidate gene because it controls proliferation of glial and neural cells and affects myelin formation by oligodendrocytes. A remarkable increase of glial cells in the cerebellum and neuronal cells in the hippocampus has been shown in mice knock out studies of *CDKN1B*.[194] Dysregulation of the CDKN1B gene could lead to additional glial and neuronal cell divisions before withdrawal from cell proliferative cycles and subsequently contribute to brain overgrowth. It has been

hypothesized that increased head brain size in autism has its onset around 12 months of age.[193] Longitudinal behavioural studies of infants at high genetic risk for autism, who are later diagnosed with autism at 36 months, report marked deficits in reciprocal social interaction observed by 12 months of age.[195] The temporal relationship between the onset of both autism symptoms and brain overgrowth at approximately 12 months suggest a relationship between these two events. Although evidence of brain overgrowth does not become apparent until postnatal life, brain development begins prenatally and is achieved through a delicate balance between cell proliferation and subsequent cell cycle withdrawals. Abnormalities in early brain development may set the stage for additional failures in postnatal life. It is plausible that brain overgrowth, through dysregulation of *CDKN1B*, results in the development of autism phenotypes through physical disruption of neural circuitry.

The developmental expression of *Erbb3* was studied and the highest expression of this gene was found during the adult stage (Figure 2.11E). This gene encodes a member of the Erbb family of receptor tyrosine kinases. The importance of Erbb receptors in development has been demonstrated from analyses of genetically modified mice. Loss of *Erbb1* leads to perinatal lethality in mice showing abnormalities in multiple organs, including the brain, skin, lung, and gastrointestinal tract.[196] *Erbb2* null mice die before E11 due to cardiac malformations and affected neural development.[197] Sharing a similar phenotype, *Ebb4* null mice die during mid-embryogenesis from cardiac malformations and alterations in innervation of the hindbrain in the CNS.[198] In the case of *Erbb3*, it is also expressed throughout development; however, a detailed analysis of its function has

yet to be reported. The present study found that *Erbb3* has predominant expression in the adult stage; nevertheless, its expression during embryonic development appears to be critical. *Erbb3* knockout mice die by E13.5 and display cardiac defects. Additionally, these animals lack Schwann-cell precursors and Schwann cells that accompany peripheral axons of sensory and motor neurons.[199] From these data, it is clear the *ERBB3* plays a critical role in modulating specific aspects of embryogenesis, although its exact function in CNS development remains to be elucidated.

Finally, *Clpx* revealed the highest expression at E16, which gradually decreased throughout development (Figure 2.11F). Clpx recognizes peptide sequences in protein substrates, unfolds stable tertiary structures in the protein, and translocates the unfolded polypeptide chain into a sequestered proteolytic compartment in Clpp for degradation into small peptide fragments.[119] Clpx was investigated as a gene of interest because according to a study conducted by Vernes and colleagues, it may be a putative downstream target of FOXP2.[105] In humans, mutations of the *FOXP2* gene are the only known cause of developmental speech and language disorders.[81] Studying the downstream targets of FOXP2 can provide the opportunity to investigate the language deficits in autism (discussed further in Chapter 3). However, in spite of the high-level expression at E16, evidence showing the CLPX contributes to neuronal physiology within the CNS is scarce.

2.6. Conclusions

An overview of autism research emphasizes the genetic endeavours in this area as well as the diversity of the methods used. Whereas cytogenetic studies search for inherited or *de novo* genetic abnormalities on an individual basis, genetic linkage studies are performed to examine if the inheritance of a particular allele occurs more frequently in affected members of multiplex families than would be expected by chance. Additional investigations, such as candidate gene studies, are chosen *a priori* based on a plausible pathogenetic model of autism. The purpose of these techniques is to identify genetic mutations that predispose an individual to autism. However, it is becoming evident that autism may result from the disrupted function of a core group of genes that cause neural network dysfunction, which could occur at the level of mRNA transcriptional regulation. In this case, transcriptomic data are required to bridge together distinct genetic variants. Microarray studies are beginning to provide important insights into gene expression alterations and dysregulated signaling pathways in autism. In addition, this method presents a valuable framework for accounting for environmental influences on gene expression.

The current work demonstrating the efficacy of buccal samples in gene expression assays provides further impetus for using peripheral tissues to identify potential autism susceptibility genes that are also dysregulated in the brain. The results of this study revealed an upregulation of genes participating in the immune response – a finding that was initially described in brain transcriptome studies. Thus, the present findings illustrate the reproducibility of gene expression changes associated with immune system

dysfunctions in autism when using samples outside of disease-relevant tissues. Collectively, these data support the notion that an increased immune response is involved in the pathogenesis of autism. Circadian rhythm, an identified dysregulated signaling pathway in this study, represents another finding that is in line with previous autism candidate gene studies as well as gene expression studies using lymphoblastoid cell lines. Other key study findings include evidence for *CDKN1B* and *CD14* dysregulation in autism. Indication of *CDKN1B* downregulation supports a previously formulated hypothesis regarding brain overgrowth in the first few years of autism and proof of *CD14* dysregulation may re-open debate over the involvement of the CD14 molecule in autism, as it has been refuted.

At the same time, true to the discovery purposes of microarray experiments, this study also revealed new genes and molecular mechanisms underlying gene expression patterns in autism. Analysis of the differentially regulated genes identified several novel dysregulated genes, including *CSF3R*, *KLF4*, *ERBB3*, and *CLPX*. As evidenced by the positive expression during brain development in a mouse model system, all of these genes are likely to be involved in some aspect of nervous system development and function. Accordingly, focused investigations on the promising role of these genes in autism development can be particularly helpful in this regard. Another novel dysregulated gene identified by this study that did not show positive expression in the brain is *C4orf7*. However, the fact that autism seems to affect multiple organs, and not exclusively the brain, suggests that peripheral tissues may also be affected. Studying *C4orf7*, a 25-fold change upregulated gene in autism, affords the opportunity to learn more about the

physiological symptoms associated with autism. Another identified dysregulated pathway, ERBB signaling, appears to play an important role in the migration and differentiation of neuronal and glial precursors early in brain development. Further, ERBB signaling is speculated to play a significant role in social behaviours and has been implicated in the pathogenesis of schizophrenia. Perturbations of the ERBB signaling pathway could lead to a variety of abnormal neurodevelopmental events, which could have implications for neurodevelopmental disorders characterized by social deficits, including schizophrenia and autism. These results support the notion that various genetic factors underlie the development of autism, and that these factors can be converging at or diverging from central networks. Lastly, the findings of this study enabled the identification of potential susceptibility genes that may underlie brain development and other physiological symptoms connected to autism. Ultimately, identified biomarkers can facilitate the process of early diagnosis and aid in the development of effective and personalized intervention programs and treatments. Moreover, buccal cell collection is a non-invasive method that would be ideal for screening initiatives as part of medical diagnoses.

2.7. Limitations and Future Areas of Study

The results suggest that buccal RNA is a genetic source that could be used to detect differential gene expression between affected and unaffected individuals. However, the inter-individual variation of buccal RNA quality and reduced RNA yield (in comparison to lymphoblastoid cell lines) are limiting factors for use as a source

material in expression studies. One method to overcome these limitations is to amplify cell numbers from buccal swabs by growth in tissue culture. This experimental system would offer high quality RNA yields above the minimal threshold required for genome-wide expression assays and additional DNA/RNA recoveries for future studies. Therefore, the aim of future experiments would be to generate immortalized cell cultures for propagation of human buccal epithelium.

A review of the literature indicates that it is possible to grow buccal cells in tissue culture. Transfection with the SV40T (simian virus 40 T) antigen has been commonly used to transform cells *in vitro*. Stable integration of SV40T into human buccal cells has been successfully used to produce an immortalized cell line, which retained non-malignant properties upon extended culture.[200] Recently, Ramirez and colleagues found a technique that reproducibly generates immortalized cells from human bronchial epithelium without the use of viral oncoproteins.[201] By expressing two genes, *hTERT* (human telomerase reverse transcriptase) and *Cdk4* (cyclin-dependent kinase 4), it was demonstrated that the resulting immortal human bronchial epithelial cell lines do not have a malignant phenotype and have few genetic alterations. Using the methods described by Ramirez and colleagues, another research group has accomplished the establishment of immortalized buccal cell lines.[202] This method could be used in future gene expression profiling studies of children with autism.

Another study limitation was the inconsistency between different professionals in the evaluation of autism symptoms. This represented a major drawback in making genotype-phenotype correlations. Future research collaborations with organizations

committed to the treatment and education of children with autism, such as New Haven Learning Centre, may help in establishing a more consistent approach. The application process of New Haven Learning Centre involves a specific behavioural assessment that confirms the primary diagnosis of autism in a strict, uniform manner. Systematic clinical assessments with identical reporting formats would lead to more accurate investigations of the relationship between genetic information and behavioural symptoms.

CHAPTER 3: TRANSCRIPTIONAL REGULATION BY FOXP1/2/4 DIMERIZATION: IMPLICATIONS FOR SPEECH AND LANGUAGE DEFICITS IN AUTISM

3.1. Introduction

The existing literature draws a connection between autism and the neural mechanisms behind the development of speech and language. Genome-wide screens using multiplex families to identify autism susceptibility loci revealed that chromosome 7q31-33 harboured the most consistent results.[203,204] Mutations of the *FOXP2* gene, located on chromosome 7q31, are responsible for a severe developmental disorder of verbal communication, suggesting that it may be related to the speech and language abnormalities of autism. It has been proposed than an insufficient dosage of FOXP2 at a key stage of embryogenesis may result in pathology of neural structures that are important for speech development.[82] *FOXP2*, which encodes a transcription factor, has been implicated not only in speech and language disorders, but also as a regulator of genes with a role in autism.[205] Therefore, genetic association studies with *FOXP2* and autism have been conducted. Although a few studies found negative results, two studies have shown a positive association.[40,44,87,206] Moreover, another member of the *FOXP* gene subfamily, *FOXP1*, has been recently linked to autism as well.[207,208] Amongst the probands that participated in the "Gene Expression Profiling in Individuals with Autism" study (see Chapter 2), there was a patient with a 5.3 Mbp deletion of the 3p14.1-p13 region, where *FOXP1* is located; hence, further interest in the *FOXP* genes was fostered. Studies in mice have demonstrated that the neural expression pattern of *Foxp2* and its

closest homologues, *Foxp1* and *Foxp4*, are partly overlapping and abundantly expressed in the developing brain.[209] Given that the transcriptional activity of FOXP2 requires dimerization, either with itself or with FOXP1 and FOXP4, the specific combination of FOXP1/2/4 homodimers and heterodimers may regulate the transcription of downstream target genes that may be involved in language acquisition. Studies of FOXP1/2/4 will not only clarify the origins of human speech and language, but also its contributions to the etiology of neurodevelopmental disorders such as autism.

3.2. Hypothesis

It has been previously shown that various FOXP transcription factors can form homodimers (i.e., FOXP1/P1, FOXP2/P2, and FOXP4/P4) or heterodimers (i.e., FOXP1/P2, FOXP1/P4, and FOXP2/P4).[102] Although the gene targets of various FOXP1/2/4 heterodimers are unknown, high-throughput searches for the direct neural targets of *FOXP2* have revealed hundreds of genes with roles in the nervous system, such as modulation of synaptic plasticity, neurodevelopment, and neural transmission.[104,105]

The objective of this study was to determine whether different combinations of FOXP1/2/4 proteins will lead to altered gene dosages of downstream targets of *FOXP2*. Ten known neural targets of *FOXP2* were chosen for further examination in this study, including *CER1*, *NCOR2*, *NEUROD2*, *PAX3*, *SNW1*, *SFRP4*, *WISP2*, *PRICKLE1*, *EFNB3*, and *SLIT1*. These genes were selected based on their function during embryonic development. *CER1*, *SFRP4*, *WISP2,* and *PRICKLE1* participate in the WNT (Wingless) signaling pathway and *NCOR2* and *SNW1* belong to the Notch signaling pathway.[210-215]

Both WNT and Notch signaling pathways are essential for regulation of embryonic development.[216] Moreover, *NEUROD2*, *PAX3*, *EFNB3*, and *SLIT* are involved in the development of the nervous system.[217-220]

It was hypothesize that the formation of different FOXP1/2/4 homodimers or heterodimers will lead to differential expression of targets genes that may contribute to altered neurodevelopment. To test this hypothesis, quantitative real-time polymerase chain reaction (qRT-PCR) was used to quantify the expression level of target genes in HEK293 cells that were transfected with various FOXP1/2/4 constructs (i.e., FOXP1/P1, FOXP2/P2, and FOXP4/P4 homodimers or FOXP1/P2, FOXP1/P4, and FOXP2/P4 heterodimers).

3.3. Methods

3.3.1. Molecular Cloning of *Homo sapiens FOXP1*, *FOXP2*, and *FOXP4* cDNA

The *FOXP1* (GenBank: BC152752), *FOXP2* (GenBank: BC126104), and *FOXP4* (GenBank: BC052803) cDNA was obtained from the DNASU Plasmid Repository or the Mammalian Gene Collection (MGC). The open reading frames (ORFs) were amplified by polymerase chain reaction (PCR) using forward and reverse primers incorporating *Kpn*I and *Xho*I restriction sites. Following PCR amplification and restriction digest with *Kpn*I and *Xho*I, the digested fragments were subject to gel purification and clean-up with the EZ-10 Spin Column DNA Gel Extraction Kit (Bio Basic Inc.). The 2.0 kb FOXP1 fragment was ligated into the *Kpn*I/*Xho*I cloning sites of mammalian expression vector pcDNA3-Flag with T4 DNA ligase. The 2.2 kb FOXP2 fragment was ligated into the

*Kpn*I/*Xho*I cloning sites of the mammalian expression vector pcDNA3/*myc*-His(-A) with T4 DNA ligase. The 2.0 kb FOXP4 fragment was ligated into the *Kpn*I/*Xho*I cloning sites of both pcDNA3-Flag and pcDNA3/*myc*-His A expression vectors. The resulting FOXP1/pcDNA3-Flag, FOXP2/pcDNA3.1-*myc*-His, FOXP4/pcDNA3-Flag, and FOXP4/pcDNA3.1-*myc*-His constructs were transformed into NEB 5-alpha Competent *E. coli* cells to be amplified and then digested with the restriction enzymes *Kpn*I and *Xho*I to confirm the appropriate length of the insert. The FOXP1/2/4 gene inserts were sequenced with forward and reverse primers, as well as an internal primer, at the York University Core Molecular Facility. The sequences were confirmed using the multiple sequence alignment tool NCBI BLAST (http://www.ncbi.nlm.nih.gov/BLAST/). All primers were designed using the Primer3 program (http://frodo.wi.mit.edu/primer3/) and produced by Sigma-Aldrich. Primer sequences are shown in Table 3.1.

Table 3.1. Primer sequences for *FOXP1*, *FOXP2*, and *FOXP4* cDNA for PCR and sequencing.

Gene	Designation Vector	Primer Sequences (5' → 3')	Amplicon size (base pair)
FOXP1	pcDNA3-Flag	F: AGCTTGGTACCGATGATGCAAGAATCTGGGACT R: TGCCGCTCGAGCCTCCATGTCCTCGTTTACTGG I: AACCACAGGCAACAATCACA	2036
FOXP2	pcDNA3.1/ *myc*-His A	F: GCTTGGTACCATGATGACTCCCCAGGTGATCA R: CCGCTCGAGGTTCCAGATCTTCAGATAAAGGCTC I: TTCCTCCTCGACTACCTCCT	2205
FOXP4	pcDNA3.1/ *myc*-His A	F: GCTTGGTACCGAATGATGGTGGAATCTGCCTCG R: CCGCTCGAGGGACAGTTCTTCTCCCGGCA I: CACCGCTACCTCGTTTGC	2061
FOXP4	pcDNA3-Flag	F: AGCTTGGTACCGATGATGGTGGAATCTGCCTCG R: CGTACTCGAGCGGACAGTTCTTCTCCCGGCA I: CACCGCTACCTCGTTTGC	2064

3.3.2. Polymerase Chain Reaction (PCR)

PCR was used to amplify *FOXP1*, *FOXP2*, and *FOXP4* ORFs for the purpose of molecular cloning into mammalian expression vectors. The 50 µL PCR reaction mixture contained a cDNA template, the forward and reverse primer at 0.2 µM, 200 µM dNTPs, 5x One*Taq* Hot Start Standard Reaction buffer, and 0.025 U/µL of the One*Taq* Hot Start DNA Polymerase (New England Biolabs). The thermal cycler was programmed for 30 cycles of PCR amplification at 94°C, 30 seconds for denaturation; 68°C, 2 minutes for annealing; and 68°C, 30 seconds for primer extension. The presence of a PCR product of the appropriate size was verified by agarose gel electrophoresis.

3.3.3. Agarose Gel Electrophoresis

Agarose gel electrophoresis was performed to verify the size of PCR products and restriction digest fragments. A gel concentration of 1% agarose in 1X TAE buffer with 10% GelGreen (VWR International) was used. The samples were mixed with 6X nucleic acid loading buffer and loaded into the wells of the agarose gel. Electrophoresis was conducted at 100V on a PowerPac Basic Power Supply. A UV Transilluminator was used to visualize the size of the DNA bands. A 1 kb DNA standard ladder (Sigma) was used to estimate the sizes of the DNA bands by comparison.

3.3.4. Restriction Digests

Restriction digests were performed to produce DNA fragments for directional insertion into mammalian expression vectors and to confirm whether plasmid construction was successful. Restriction digests were performed for 1-2 hours at 37°C using 20 U/µL of restriction enzyme (New England Biolabs). The appropriate buffer

(NEBuffer 1-4) and bovine serum albumin (BSA) were added to the reaction mixture at 10-fold dilution. For sequential restriction digests, the digested DNA fragments were purified using the EZ-10 Spin Column PCR Products Purification Kit (Bio Basic Inc.) prior to performing another digest on the same fragments. Subsequently, the resulting restriction fragments were separated according to their lengths by agarose gel electrophoresis. A UV Transilluminator was used to visualize the DNA. Gel slices containing the desired DNA fragments were excised and purified using the EZ-10 Spin Column DNA Gel Extraction Kit (Bio Basic Inc.).

3.3.5. Ligation of DNA Inserts into a Plasmid

DNA fragments with cohesive ends were ligated into complimentary cohesive end vectors of choice using T4 DNA Ligase (New England Biolabs). The ligation mixture included 400 U/μL T4 DNA Ligase, 10X T4 DNA Ligase Reaction Buffer (New England Biolabs), vector DNA, and DNA insert. A molar ratio of 1:3 vector to insert was used for the ligation reaction. This mixture was incubated at 16°C overnight prior to use in *E. coli* transformation.

3.3.6. Transformation Using Chemically Competent *E. coli* Cells

NEB 5-alpha Competent *E. coli* cells (New England Biolabs) or One Shot TOP10 Chemically Competent *E. coli* cells (Invitrogen) were used to generate large quantities of the desired plasmid. After thawing a tube of NEB 5-alpha Competent *E. coli* cells on ice for 10 minutes, 5 μL of plasmid DNA was added to the competent cell mixture, and placed on ice for 30 minutes. After this incubation period, the mixture was then heat shock at 42°C for exactly 30 seconds and placed back on ice for 5 minutes. Room

temperature SOC Outgrowth Medium (New England Biolabs) was pipetted into this mixture, which was shaken vigorously at 250 rpm at 37°C for 1 hour. For One Shot TOP10 Chemically Competent *E. coli* cells, 5 μL of plasmid DNA was added to the thawed competent cell mixture, and incubated on ice for 30 minutes. The mixture was then heat shocked for 30 seconds at 42°C and quickly placed back in ice for 2 minutes. Pre-warmed S.O.C. Medium (Invitrogen) was added to each vial, which was shaken at 225 rpm at 37°C for 1 hour. The cells were plated onto warmed selection plates and incubated overnight at 37°C. Single bacterial colonies were inoculated into 5 ml of Luria-Bertani (LB) broth supplemented with the appropriate antibiotic in 15 mL snap-cap tubes and shaken (200 rpm) at 37°C. Incubation of the cultures did not exceed 24 hours to ensure that the plasmids are produced in healthy cells. The plasmids were purified using the EZ-10 Spin Column Plasmid DNA Minipreps Kit (Bio Basic Inc.). The results of a diagnostic restriction digest were analyzed by agarose gel electrophoresis to ensure that the DNA fragments were of the expected lengths. A glycerol stock solution containing 800 μL of *E. coli* solution in 50% glycerol was created and stored at -80°C for future purposes.

3.3.7. Luria-Bertani Plates and Medium

Luria-Bertani (LB) plates and media were used to grow clone plasmids and transformed competent *E. coli* cells. The LB media (1% tryptone, 0.5% yeast extract, and 1% NaCl in deionized H_2O) was autoclaved for 20 minutes and left to cool at room temperature. The appropriate selection antibiotic (50 μg/mL of ampicillin, kanamycin, or streptomycin) was added prior to use. For LB plates, agar was added to LB media (1%

tryptone, 0.5% yeast extract, 1% NaCl, and 1.5% agar in deionized H_2O) before autoclaving. After the addition of the antibiotic, the LB media was poured into 10 cm plates and left to solidify at room temperature before storage at 4°C.

3.3.8. Plasmid Preparation of pcDNA3.1/myc-His(-A) and pcDNA3-Flag

The mammalian expression vectors pcDNA3.1/*myc*-His(-A) (Invitrogen) and pcDNA3-Flag (Invitrogen) were streaked onto ampicillan plates and left to incubate overnight at 37°C. The single colonies that formed on the ampicillan plates were then inoculated into 5mL cultures of LB ampicillan broth. The culture was incubated at 37°C with shaking (200 rpm) overnight. Incubation of the cultures did not exceed 24 hours. Plasmid DNA was generated from this culture using a miniprep DNA isolation kit (EZ-10 Spin Column Plasmid DNA Minipreps Kit, Bio Basic Inc.). Restriction digest with *Kpn*I and *Xho*I was performed in preparation for directional insertion of a desired gene of interest. A 5.5 kb digestion product was produced.

3.3.9. HEK293 Cell Culture and Lipfectamine Transfection

HEK293 cells, obtained from the American Tissue Culture Collection (ATCC), were maintained in a 37°C incubator containing 5% CO_2. HEK293 cells were grown in Dulbecco's Modified Eagle's Medium (DMEM) (Invitrogen) supplemented with 10% fetal bovine serum (FBS). Lipofectamine 2000 (Invitrogen) was used for transfection of HEK293 cells. One day prior to transfection, cells were plated in 6-well plates (10 cm^2/well) or 12-well plates (4 cm^2/well) in 2.5 mL or 1.5 mL of growth medium without antibiotics, respectively. Cells were allowed to reach 90-95% confluency at the time of transfection. DNA was diluted in Opti-MEM Reduced Serum Medium without serum and

mixed gently. Opti-MEM Reduced Serum Medium was also used to dilute Lipofectamine 2000 prior to complexing with DNA. The diluted DNA and Lipofectamine 2000 were combined and incubated at room temperature for 20 minutes to promote DNA-Lipofectamine 2000 complex formation. The ratio of LiopfectAMINE 2000 to DNA used was 3:1. This mixture was subsequently added to each well containing cells and media. The cells where incubated at 5% CO_2 in a 37°C incubator for 6 hours before changing to 2 mL complete growth medium with antibiotics. The cells were incubated for another 18 hours to allow protein expression. 1000 µg/mL Geneticin (G418) was used for selecting cells transfected with appropriate FOXP1/2/4 constructs.

3.3.10. Immunocytochemistry

Immunocytochemistry was performed to assess the presence of a particular protein by use of a specific antibody, thereby allowing visualization and examination under a microscope. HEK293 cells were seeded in 35 mm dishes with cover slips previously treated with poly-D-lysine hydrobromide for 18-24 hours. For fixation, the cells were incubated for 20 minutes at 4°C with phosphate buffered saline (PBS) containing 4% paraformaldehyde. For immunocytochemistry, the cells were first incubated for 10 minutes at room temperature with 0.5% Triton X-100 in PBS. The primary antibodies used were mouse monoclonal anti-Myc (1:200; #sc-40, Santa Cruz Biotechnology) or rabbit polyclonal anti-His tag (1:400; #ab78159, Abcam) or both for double staining. The incubation period was for one hour at room temperature. The cells were rinsed and then repeatedly washed with PBS-T for 15 minute intervals before an hour-long incubation with secondary antibodies. The secondary antibodies used were fluorescein isothiocynate

(FITC)-conjugated anti-rabbit IgG (1:300; #111-095-144, Jackson ImmunoResearch Laboratories) and Texas Red (TR) dye-conjugated anti-mouse IgG (1:300; #115-0750146, Jackson ImmunoResearch Laboratories). The cells were rinsed, washed four times with PBS-T for 5 minute intervals, incubated with 4'-6-diamidino-2-phenylindole (DAPI) (0.05% DAPI in PBS-T) for 20 minutes, then washed three times with PBS-T for 5 minute intervals. The coverslips were mounted onto microscope slides with Vectashield Mounting Medium (Vector Laboratories) and visualized using a Nikon Eclipse 80i microscope.

3.3.11. Western Blot Analysis

Western blot analysis was used in order to determine whether HEK293 cells expressed FOXP1/2/4 proteins. Protein samples were mixed with 1x Laemmli sample buffer (BioRad), heated for 5 minutes in boiling water, cooled on ice and loaded on the gel alongside the molecular marker, Precision Plus Protein Dual Colour Standards (BioRad). The samples were run at 90 V for approximately 10 minutes until samples reached the bottom of the stacking gel and then at 100 V through the separating gel in 1X running buffer until the samples had migrated to the bottom of the separating gel. After transferring the gel onto a 0.2 μM nitrocellulose membrane, proteins were blocked in 5% milk in TBS-T for two hours. The membranes were incubated overnight at 4°C with mouse monoclonal anti-Myc (1:200; #sc-40, Santa Cruz Biotechnology) or rabbit polyclonal anti-His tag (1:300; #ab78159, Abcam), and then incubated with anti-mouse horseradish peroxidase-conjugated secondary antibody (1:10,000; #sc-2060 Santa Cruz Biotechnology) or anti-rabbit horseradish peroxidase-conjugated secondary antibody

(1:10,000; #sc-2054 Santa Cruz Biotechnology) for one hour at room temperature. The membranes were washed four times in five minute intervals with TBS-T and incubated with Amersham ECL Prime Western Blotting Detection Reagent (#RPN2232sk, GE Healthcare). The bound antibodies were detected by Geliance 600 Imaging System (PerkinElmer).

3.3.12. Reverse Transcription

RNA from transfected HEK293 cells was isolated using the NucleoSpin RNA/Protein kit (Machery-Nagel). Reverse transcription was performed with M-MuLV Reverse Transcriptase (New England Biolabs) in accordance with the manufacturer's instructions. 1 μg of RNA was used in a reaction with 10x RT Buffer, OligodT primer, dNTP mix (2.5 mM each), and M-MuLV Reverse Transcriptase (200 U/μL). This mixture was incubated at 42°C for one hour and then at 90°C for 10 minutes to inactivate the enzyme. The success of the reaction was confirmed by polymerase chain reaction with *GAPDH* primers. The subsequent PCR products were then verified by agarose gel electrophoresis.

3.3.13. Quantitative Real-Time Polymerase Chain Reaction (qRT-PCR)

Primers specific for candidate genes were designed using Primer Express 3.0 (Applied Biosystems) and synthesized by Sigma-Aldrich (Table 3.2). QRT-PCR was performed using the 7500 Fast Real-Time PCR System. The qRT-PCR reaction mixture included SYBR Green reagent, 2 μM each of forward and reverse primers, and cDNA template. The comparative Ct method was used to analyze the data. Raw Ct values were normalized to the geometric mean of the housekeeping genes *HPRT* (hypoxanthine

phosphoribosyl transferase) and *GAPDH* (glyceraldehyde 3-phosphate dehydrogenase) to obtain the ΔCt values. The ΔCt values of the samples were compared to a calibrator (empty vector control). The FOXP1-Flag sample, derived from HEK293 cells transfected with FOXP1/pcDNA3-Flag, was compared to a control derived from HEK293 cells transfected with the pcDNA3-Flag vector only. The FOXP2-Myc and FOXP4-Myc samples, derived from HEK293 cells transfected with FOXP2/pcDNA3.1-*myc*-His or FOXP4/pcDNA3.1-*myc*-His, were compared with HEK293 cells transfected with the pcDNA3.1/myc-His vector. The double transfection samples were compared to the average of both empty vector controls. Reported relative quantification (RQ) ratios are the mean of comparisons of three biological replicates. Three separate passages of each cell line were used as biological replicates. Statistical significance was determined by performing a one-way analysis of variance (ANOVA) followed by Student's *t*-test.

Table 3.2. Primer sequences for downstream genes of *FOXP2* for qRT-PCR.

Gene	Primer Sequences (5' → 3')	Amplicon size (base pair)
CER1	F: CAGAGTTCTCTTTCCCCCGTACT R: TTCTCCTCAGCTTCCTCATGGT	82
NCOR2	F: TCAGCGGAGTGAAGCAGGAG R: TCGATGCTGGCTGAGGAGAT	493
NEUROD2	F: CAAGCGGCCAGACCTAGTGT R: CTGCGACAGACCCTTGCA	55
PAX3	F: GAGAGAACCCGGGCATGTT R: CCGCGTCCTTGAGTAATTTGTC	57
SNW1	F: GGCAGAAGCCCTCTACATTGC R: ACTTGGGCACGCATTTCC	63
SFRP4	F: CCCGGAGGATGTTAAGTGGAT R: GTCAACATCAAGAGGCCTTTCC	70
WISP2	F: GGTCGCAGTCCACAAAACAG R: CACCGTGTCCCCATTCC	57
PRICKLE1	F: TCACTGTGGCAGGCACCAT R: CTCAGCTTCTGTGCACTCATCAG	91
EFNB3	F: TCGGCGAATAAGAGGTTCCA R: GTCCCCGATCTGAGGGTACA	60
SLIT1	F: GGAGGCCACTGGGATGTTTA R: CCATCTTCAATTTCTGACACCTTGT	84

3.4. Results

3.4.1. Development of FOXP1/2/4 Constructs

Homo sapiens FOXP1, *FOXP2*, and *FOXP4* ORFs were amplified by PCR using forward and reverse primers incorporating *Kpn*I and *Xho*I restrictions sites. After restriction digest with the restriction enzymes *Kpn*I and *Xho*I, the resulting *FOXP1* and *FOXP2* fragments were ligated into the *Kpn*I/*Xho*I cloning sites of the pcDNA3-Flag and pcDNA3.1/*myc*-His vector, respectively. *Homo sapiens FOXP4* ORF was inserted into both mammalian expression vectors (pcDNA3-Flag and pcDNA3.1/*myc*-His) at the *Kpn*I/*Xho*I cloning sites, for the purpose of co-transfections. Following transformation into NEB 5-alpha Competent *E. coli* cells for amplification, FOXP1/pcDNA3-Flag, FOXP2/pcDNA3.1-*myc*-His, FOXP4/pcDNA3-Flag, and FOXP4/pcDNA3.1-*myc*-His were digested with the restriction enzymes *Kpn*I and *Xho*I to ensure the integrity of the constructs. It was verified that the FOXP1/2/4 gene inserts that dropped out from the vector were of appropriate size (Figure 3.1). The FOXP1/2/4 gene inserts were also sequenced with forward, internal, and reverse primers for further confirmation.

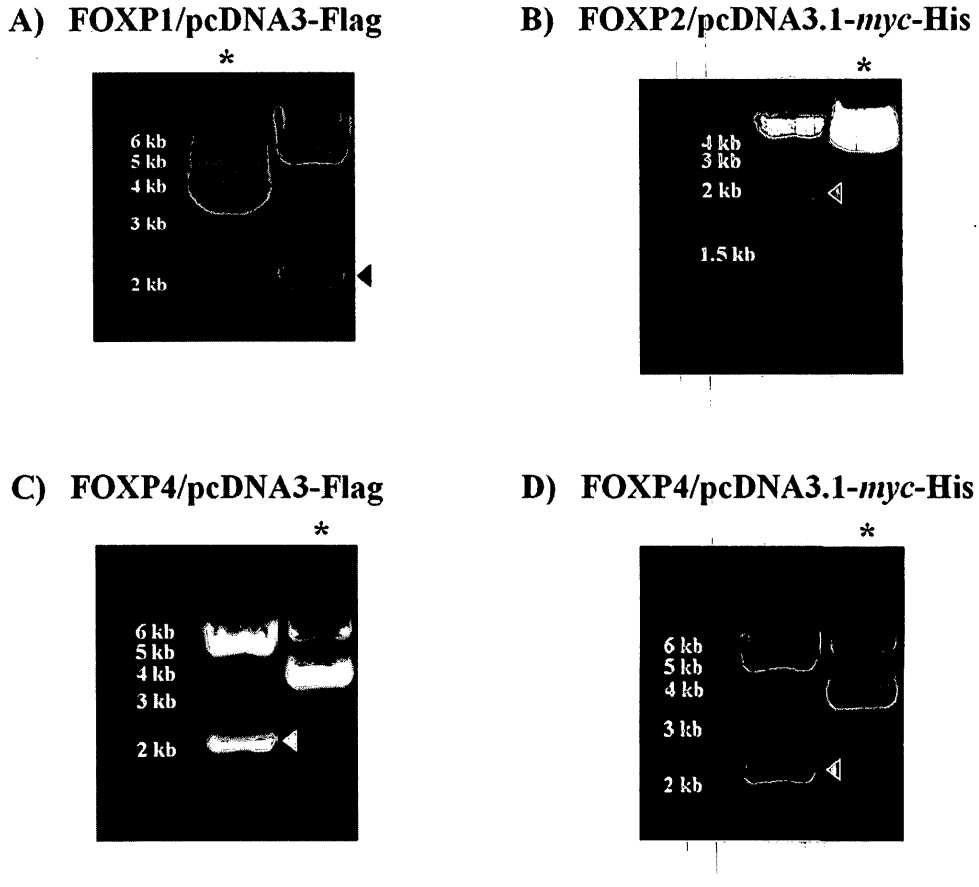

Figure 3.1. Restriction digest of FOXP1/2/4 constructs with *Kpn*I and *Xho*I to confirm insert drop-out. (A) FOXP1/pcDNA3-Flag, **(B)** FOXP2/pcDNA3.1-*myc*-His, **(C)** FOXP4/pcDNA3-Flag, and **(D)** FOXP4/pcDNA3.1-*myc*-His were digested with *Kpn*I and *Xho*I to determine whether plasmid construction was successful. The DNA band sizes for *FOXP1*, *FOXP2*, and *FOXP4* are approximately 2 kb, as indicated by the arrowheads. The asterisk (*) represents the corresponding undigested FOXP1/2/4 construct.

3.4.2. Subcellular Localization of FOXP1/2/4 Proteins

In order to determine the protein expression and the subcellular localization of FOXP1, FOXP2, and FOXP4, immunocytochemistry was performed on HEK293 cells over-expressing 3'-Flag-tagged FOXP1 (FOXP1-Flag), 3'-*myc*-His-tagged FOXP2 (FOXP2-Myc), 3'-Flag-tagged FOXP4 (FOXP4-Flag), or 3'- *myc*-His-tagged FOXP4 (FOXP4-Myc) alone or in combination. The subcellular localization of FOXP1-Flag and FOXP4-Flag were determined using anti-Flag antibodies and detected with FITC-conjugated secondary antibodies (Figure 3.2A and Figure 3.2C, respectively). FOXP2-Myc and FOXP4-Myc localization were determined using anti-Myc antibodies and detected with TR-conjugated secondary antibodies (Figure 3.2B and Figure 3.2D, respectively). FOXP1-Flag and FOXP2-Myc co-localization in HEK293 cells that were transfected with both FOXP1/pcDNA3-Flag and FOXP2/pcDNA3.1-*myc*-His was determined using anti-Flag and anti-Myc antibodies and detected with FITC-conjugated and TR-conjugated secondary antibodies (Figure 3.3A). Similarly, the co-localization of FOXP1-Flag and FOXP4-Myc, as well as FOXP2-Myc and FOXP4-Flag, were determined using both anti-Flag and anti-Myc antibodies, and detected with FITC-conjugated and TR-conjugated secondary antibodies (Figure 3.3B and Figure 3.3C). The results show that the over-expressed FOXP1/2/4 proteins were localized predominantly in the nucleus to various levels in different cells, which was confirmed through its co-localization with DAPI. DAPI is known to form fluorescent complexes with double-stranded DNA in the nucleus.

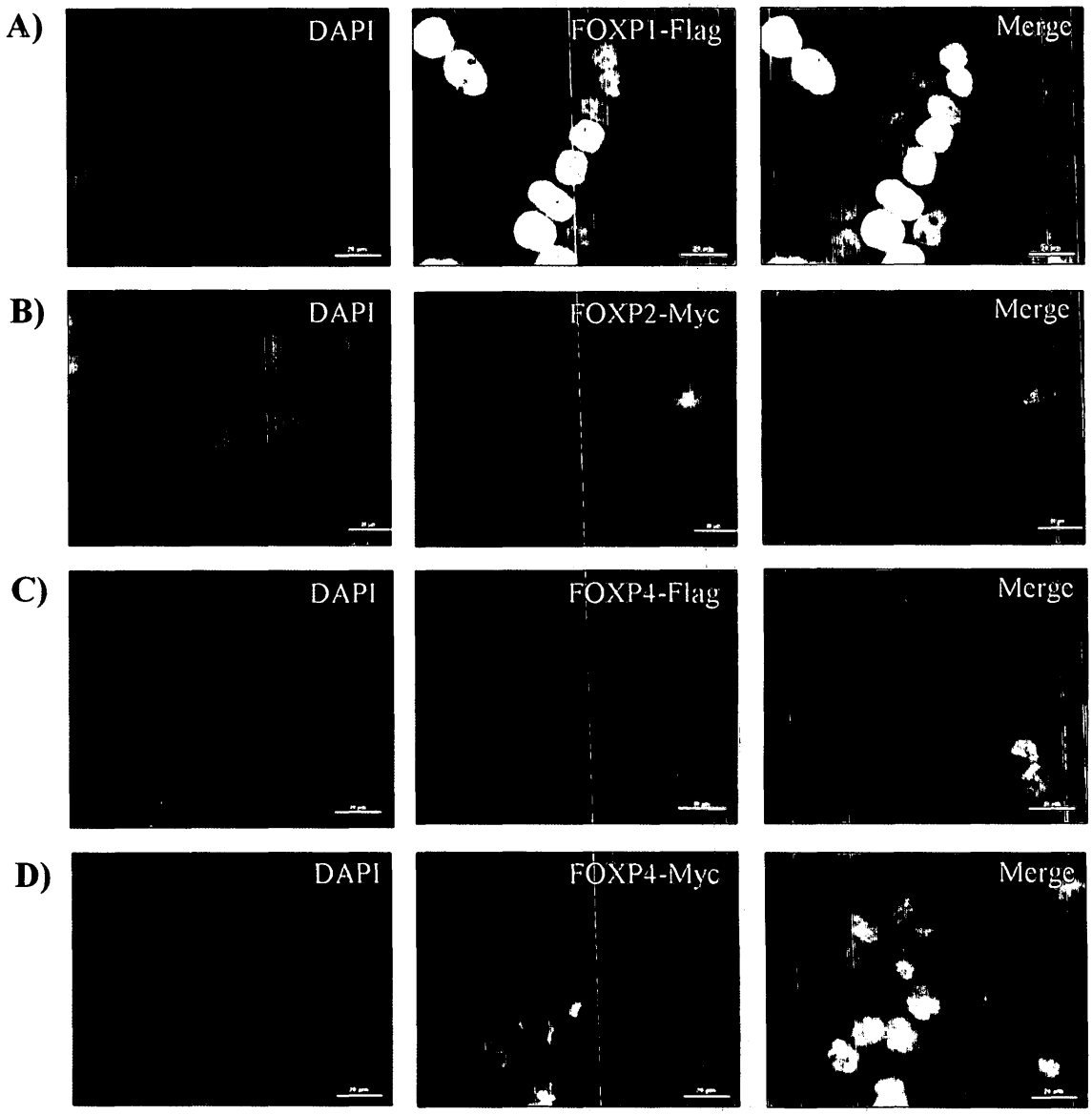

Figure 3.2. Immunocytochemical localization of FOXP1/2/4 homodimers. In HEK293 cells, subcellular localization of **A)** FOXP1-Flag was determined using anti-Flag antibodies and detected with FITC-conjugated secondary antibodies. **B)** Localization of FOXP2-Myc was determined using anti-Myc antibodies and TR-conjugated secondary antibodies. **C)** Localization of FOXP4-Flag was determined using anti-Flag antibodies and FITC-conjugated secondary antibodies. **D)** FOXP4-Myc was detected using anti-Myc antibodies and TR-conjugated secondary antibodies. The FOXP1/2/4 proteins are localized predominantly in the nucleus, which is confirmed with its co-localization with DAPI.

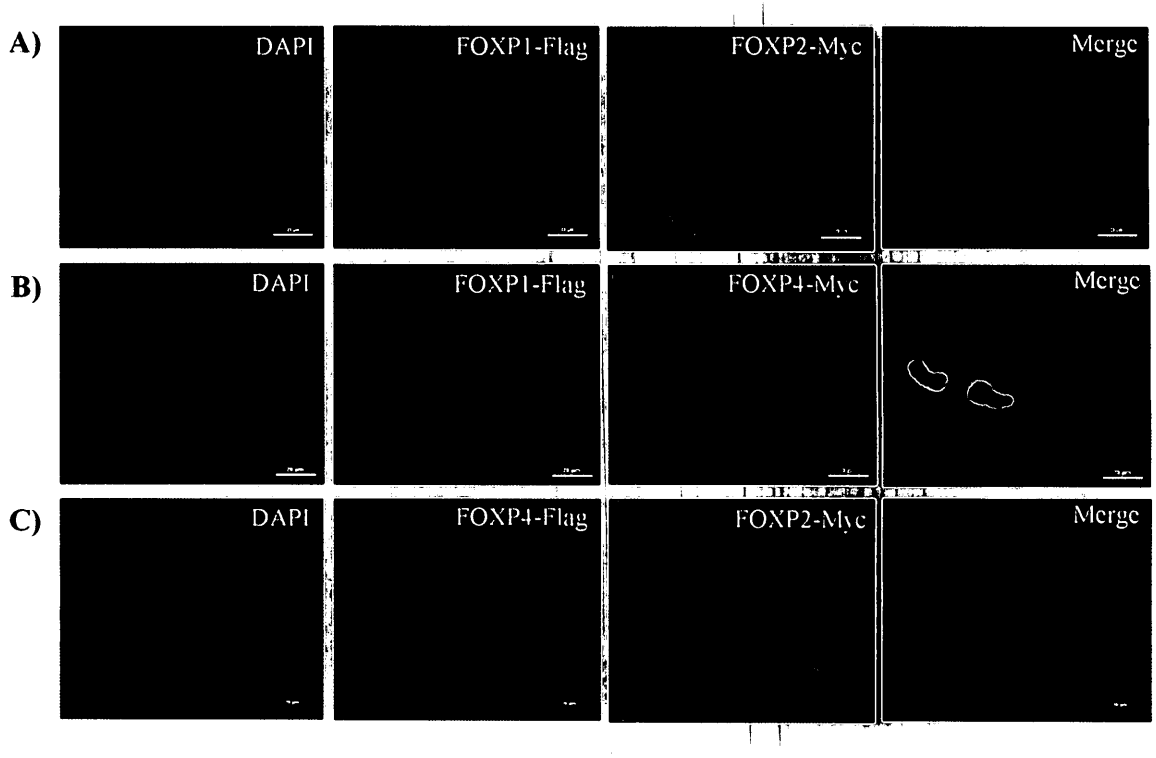

Figure 3.3. Immunocytochemical localization of FOXP1/2/4 heterodimers. The subcellular co-localization of **A)** FOXP1-Flag and FOXP2-Myc, **B)** FOXP1-Flag and FOXP4-Myc, and **C)** FOXP2-Myc and FOXP4-Flag in HEK293 cells was determined using anti-Flag and anti-Myc antibodies and detected with FITC-conjugated and TR-conjugated secondary antibodies. The FOXP1/2/4 proteins are localized predominantly in the nucleus, as determined by co-staining with DAPI.

3.4.3. Western Blot Analysis of FOXP1/2/4 Protein Expression

The expression of FOXP1-Flag and was confirmed by western blot analysis using anti-Flag antibody. An 82 kDa band corresponding to the size of *Homo sapiens* FOXP1 was detected in HEK293 cells transfected with 3'-Flag-tagged FOXP1 but not in untransfected HEK293 cells or cells that were transfected with the pcDNA3-Flag empty vector (Figure 3.4A). The expression of FOXP2-Myc and FOXP4-Myc was verified using anti-Myc antibody, where an 82 kDa band was seen to represent the *Homo sapiens* FOXP2 and FOXP4 protein but not in untransfected HEK293 cells or cells transfected with the pcDNA3.1/*myc*-His empty vector (Figure 3.4B). The co-expression of FOXP1-Flag and FOXP2-Myc, FOXP1-Flag and FOXP4-Myc, and FOXP2-Myc and FOXP4-Flag in HEK293 was confirmed as well (Figure 3.4A-B). The confirmation of protein expression in the transfected cells means that the RNA from these cell lines may be used for quantitative real-time polymerase chain reaction.

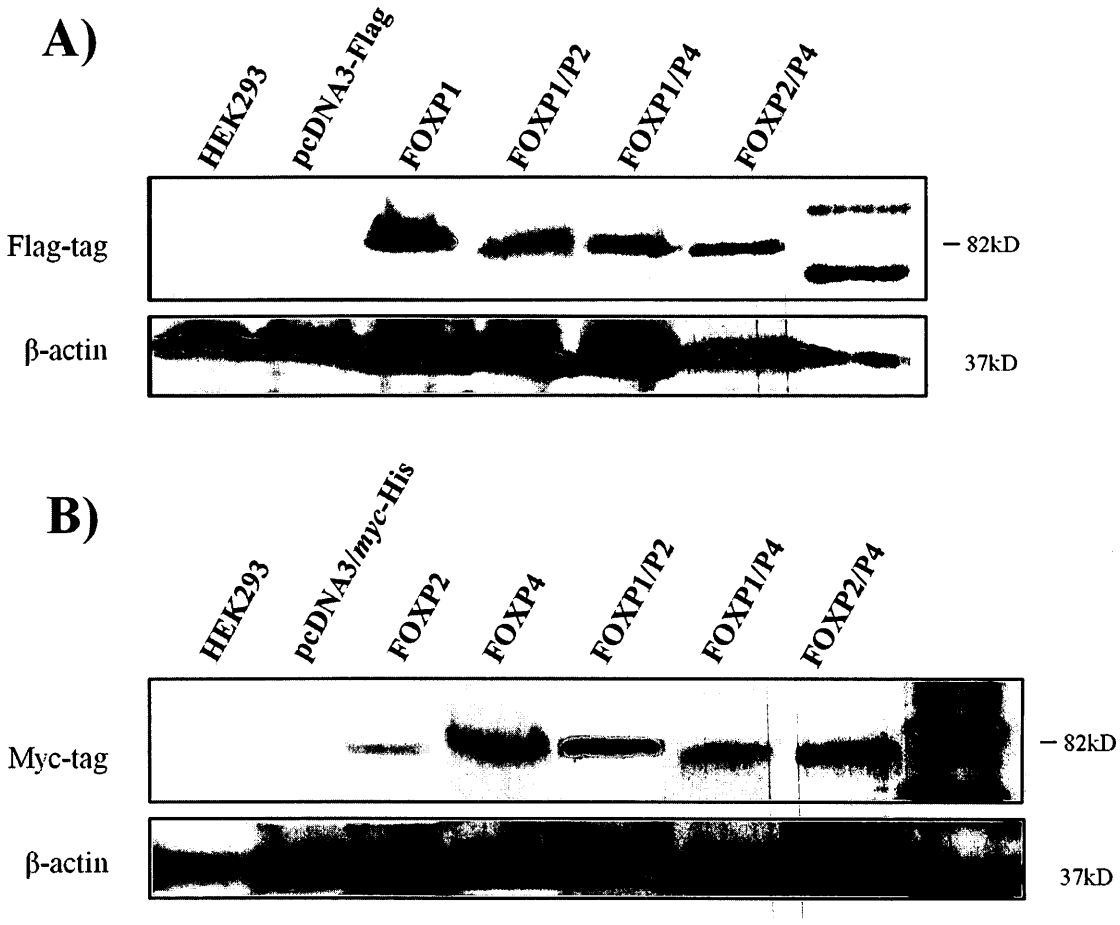

Figure 3.4. Expression of FOXP1/2/4 protein in HEK293 cells using western blot analysis. Antibodies against the **A)** 3'-Flag or **B)** 3'-Myc tag revealed protein band sizes for FOXP1/2/4 proteins to be 82 kDa. The membranes were re-probed with antibody against β-actin to ensure equal loading.

3.4.4. Quantification of the Expression of Downstream FOXP2 Targets

Quantitative real-time polymerase chain reaction (qRT-PCR) was used to assess gene expression differences due to the combinatorial actions of FOXP1/2/4 proteins on the downstream targets of FOXP2. The downstream targets of FOXP2 (*CER1*, *NCOR2*, *NEUROD2*, *PAX3*, *SNW1*, *SFRP4*, *WISP2*, *PRICKLE1*, *EFNB3*, and *SLIT1*) were chosen based on their role during embryonic development. Using mRNA from transfected HEK293 cell lines, the expression level of the target genes was quantified relative to the geometric means of the housekeeping genes *HPRT* and *GAPDH* and compared to cells transfected with the respective pcDNA3-Flag and pcDNA3.1-*myc*-His empty vectors.

Statistically significant differences of expression changes were found amongst the various combinations of FOXP1/2/4 homodimers and heterodimers for eight of the target genes (*CER1*, *NCOR2*, *NEUROD2*, *SNW1*, *SFRP4*, *WISP2*, *PRICKLE1*, and *EFNB3*). The various dimeric combinations of FOXP1/2/4 proteins had no effect on *PAX3* and *SLIT1* expression. Figure 3.5 depicts the relative quantification (RQ) of downstream *FOXP2* target genes activated by FOXP1 and FOXP2 homodimers and heterodimers. Expression changes caused by FOXP1 and FOXP4 homodimers and heterodimers are shown in Figure 3.6. The effects of FOXP2 and FOXP4 homodimers and heterodimers on downstream target genes are shown in Figure 3.7. A direct comparison of the differential expression induced by FOXP1/2/4 heterodimers is shown in Figure 3.8. The effect of transcriptional regulation by FOXP1/2/4 dimerization on each target gene is described in the following sections.

CER1 (Cerberus 1)

The expression of *CER1* was repressed by all FOXP1/2/4 combinations. A one-way analysis of variance (ANOVA) was conducted to compare the effects of FOXP1/2/4 dimerization on target gene expression. There was a significant main effect of FOXP1/2/4 dimerization on *CER1* gene expression ($F(5, 12) = 9.87$, $p = 0.0006$). *Post hoc* comparisons indicated that the RQ ratio for FOXP2 homodimers (-2.28 ± 0.35; mean ± SEM) was not significantly different than FOXP1 homodimers (-1.12 ± 0.01) (Figure 3.5). The RQ ratio for FOXP1 homodimers was found to be significantly ($p = 0.0004$) different from FOXP1/P2 heterodimers (-2.59 ± 0.11) (Figure 3.5). However, differential expression was not seen between FOXP2 homodimers and FOXP1/P2 heterodimers; thus, indicating that FOXP1 does not seem to influence the expression of *CER1* via a FOXP1/P2 heterodimeric combination. The RQ ratio for FOXP4 homodimers (-2.73 ± 0.14) was found to be significantly ($p = 0.0002$) different from FOXP1 homodimers (Figure 3.6). The expression level of *CER1* brought about by FOXP1/P4 heterodimers (-2.18 ± 0.35) appeared to be intermediary of that by the FOXP1 and FOXP4 homodimers, although this was not statistically significant (Figure 3.6). The RQ ratio for FOXP2/P4 heterodimers (-1.28 ± 0.02) was smaller than its corresponding FOXP2 and FOXP4 homodimeric counterparts ($p = 0.0064$ and $p = 0.0005$, respectively), suggesting that FOXP2/P4 heterodimers are weaker transcriptional repressors than FOXP2 and FOXP4 homodimers (Figure 3.7). Moreover, the FOXP2/P4 heterodimer was determined to be a significantly weaker transcriptional repressor than the FOXP1/P2 and FOXP1/FOXP4 heterodimers ($p = 0.001$ and $p = 0.012$, respectively) (Figure 3.8).

NCOR2 (Nuclear receptor co-repressor 2)

The results show that the transcriptional outcomes of *NCOR2* differed significantly across the six FOXP1/2/4 protein combinations ($F(5, 12) = 32.17$, $p < 0.0001$). *Post hoc* analyses indicated that while the RQ ratios of FOXP1 (1.31 ± 0.18) and FOXP2 homodimers (1.44 ± 0.11) were not different, FOXP1/P2 heterodimers (3.79 ± 0.41) produced a significantly higher upregulation of *NCOR2* than its homodimeric counterparts ($p = 0.0018$ and $p = 0.0026$, respectively) (Figure 3.5). FOXP1 and FOXP4 homodimers (0.54 ± 0.78) were not statistically significant (Figure 3.6). FOXP1/P4 heterodimers (-1.24 ± 0.16), which caused transcriptional repression, had a significantly different effect on *NCOR2* expression than FOXP1 and FOXP4 homodimers ($p = 0.0014$ and $p = 0.0143$, respectively), which resulted in transcriptional activation (Figure 3.6). FOXP2 and FOXP4 homodimers led to *NCOR2* upregulation with no significant difference; although comparisons with FOXP2/P4 heterodimers (5.79 ± 0.55) revealed a significantly larger RQ ratio by the heterodimeric combination ($p < 0.0001$ and $p < 0.0001$, respectively) (Figure 3.7). The expression levels of *NCOR2* were significantly different between FOXP1/P2 and FOXP1/P4 ($p < 0.0001$), FOXP1/P2 and FOXP2/P4 ($p = 0.0072$), and FOXP1/P4 and FOXP2/P4 ($p < 0.0001$) heterodimeric combinations (Figure 3.8).

NEUROD2 (Neurogenic differentiation factor 2)

Statistical analysis showed significant differences ($F(5, 12) = 26.25$, $p < 0.0001$) across FOXP1/2/4 protein combinations on *NEUROD2* expression. *Post hoc* tests revealed that the RQ ratios of FOXP1 (-2.43 ± 0.37) and FOXP2 homodimers (1.40 ±

0.10) were significantly ($p < 0.0001$) different; whereas FOXP1 homodimers caused downregulation, FOXP2 homodimers led to upregulation of *NEUROD2* (Figure 3.5). The RQ ratio of FOXP1/P2 heterodimers (-2.10 ± 0.53) was significantly ($p < 0.0001$) different from FOXP2 homodimers but was nearly equivalent to that of FOXP1 homodimers (Figure 3.5). Hence, it does not seem likely that the regulation of *NEUROD2* involves FOXP1/P2 heterodimers. The RQ ratio of FOXP4 homodimers (1.13 ± 0.06) was significantly different from FOXP1 homodimers and FOXP1/P4 heterodimers (-1.39 ± 0.23) ($p < 0.0001$ and $p = 0.0001$, respectively) (Figure 3.6). While FOXP4 homodimers led to *NEUROD2* upregulation, FOXP1 homodimers and FOXP1/P4 heterodimers caused *NEUROD2* downregulation (Figure 3.6). In addition, the RQ ratio of FOXP1 homodimers was not significantly different from that of FOXP1/P4 heterodimers (Figure 3.6). This demonstrates that FOXP4 is not likely to influence *NEUROD2* expression through interactions with FOXP1. The RQ ratio of FOXP2/P4 heterodimers (-1.51 ± 0.38) was significantly different from that of both FOXP2 and FOXP4 homodimers at the $p < 0.0001$ level (Figure 3.7). Even though FOXP2 and FOXP4 homodimers caused transcriptional activation, the heterodimeric combination of FOXP2 and FOXP4 brought about the transcriptional repression of *NEUROD2*. All FOXP1/2/4 heterodimeric combinations caused transcriptional repression of *NEUROD2* although they were not found to be significantly different from one another (Figure 3.8).

PAX3 (Paired box 3)

The expression of PAX3 was uniformly upregulated by FOXP1/2/4 homodimers

and heterodimers (Figure 3.5-3.8). No significant difference was found between FOXP1/2/4 dimeric combinations in *PAX3* expression ($F(5, 12) = 1.14, p = 0.39$).

SNW1 (SNW domain containing 1)

One-way ANOVA revealed significant differences in *SNW1* expression between FOXP1/2/4 protein combinations ($F(5, 12) = 190.85$, $p < 0.0001$). *Post hoc* analyses revealed that the RQ ratio of FOXP1 homodimers (1.22 ± 0.08) was significantly different from FOXP2 homodimers (-1.33 ± 0.13, $p < 0.0001$) and FOXP1/P2 heterodimers (-1.55 ± 0.04, $p < 0.0001$) (Figure 3.5). FOXP1 homodimers caused the transcriptional activation of *SNW1* but FOXP2 homodimers and FOXP1/P2 heterodimers led to transcriptional repression. Since FOXP2 homodimers and FOXP1/P2 heterodimers did not have a significantly different effect on *SNW1* expression, it seems unlikely that FOXP1 regulates the expression of *SNW1* in concert with FOXP2 (Figure 3.5). The RQ ratios of FOXP1 and FOXP4 homodimers (1.27 ± 0.06) were not significantly different (Figure 3.6). On the other hand, *SNW1* expression as influenced by FOXP1/P4 heterodimers (-1.20 ± 0.16) was significantly different from both FOXP1 and FOXP4 homodimers at the $p < 0.0001$ level and caused downregulation instead (Figure 3.6). The expression level of *SNW1* mediated by FOXP4 homodimers was found to be significantly different from that of FOXP2 homodimers and FOXP2/P4 heterodimers (-1.18 ± 0.02) at the $p < 0.0001$ level, although the latter two combinations were not significantly different from one another (Figure 3.7). The data indicates that FOXP4 is also unlikely to heterodimerize with FOXP2 to modulate the expression of *SNW1*. The FOXP1/2/4 heterodimers were observed to cause downregulation of *SNW1* expression (Figure 3.8).

SFRP4 (Secreted frizzled-related protein 4)

One-way ANOVA showed a significant main effect of FOXP1/2/4 dimerization on *SFRP4* gene expression ($F(5, 12) = 24.00$, $p < 0.0001$). *Post hoc* tests did not indicate a significant difference between FOXP1 homodimers (-1.64 ± 0.57) and FOXP2 homodimers (-1.55 ± 0.15) (Figure 3.5). Although FOXP1 and FOXP2 homodimers demonstrated nearly equal ability to repress *SFRP4* expression, FOXP1/P2 heterodimers (-2.60 ± 0.06) produced significantly greater transcriptional repression than its homodimeric counterparts ($p = 0.0205$ and $p = 0.0131$, respectively) (Figure 3.5). FOXP4 homodimers (1.07 ± 0.02) had an opposite effect on *SFRP4* expression and caused transcriptional activation; this was found to be significantly ($p < 0.0001$) different from FOXP1 homodimers (Figure 3.6). However, the RQ ratio of FOXP1/P4 heterodimers (-1.62 ± 0.10) was not significantly different from FOXP1 homodimers, suggesting that FOXP4 does not regulate *SFRP4* expression via FOXP1/P4 heterodimers (Figure 3.6). The RQ ratio of FOXP4 homodimers was significantly different from FOXP2 homodimers, as well as FOXP2/P4 heterodimers (-1.80 ± 0.18) at the $p < 0.001$ level (Figure 3.7). The RQ ratio of FOXP2 homodimers differed significantly ($p < 0.001$) from FOXP2/P4 heterodimers (Figure 3.7). A comparison of the effects of FOXP heterodimers on *SFRP4* expression revealed that these dimer combinations cause transcriptional repression (Figure 3.8). In particular, the FOXP1/P2 heterodimers caused the strongest repression, which was significantly different from FOXP1/4 and FOXP2/4 ($p = 0.0187$ and $p = 0.0475$, respectively).

WISP2 (WNT1 inducible signaling pathway protein 2)

One-way ANOVA showed significant differences in *WISP2* expression across FOXP1/2/4 protein combination ($F(5, 12) = 9.49$, $p = 0.0007$). *Post hoc* tests indicated that the RQ ratios of FOXP1 homodimers (-1.49 ± 0.08) and FOXP2 homodimers (-2.00 ± 0.27) were not significantly different (Figure 3.5). FOXP1 and FOXP2 homodimers were responsible for the transcriptional repression of *WISP2*. The RQ ratio of FOXP1/P2 heterodimers (-3.76 ± 0.47) was significantly greater than FOXP1 and FOXP2 homodimers and caused even further transcriptional repression ($p = 0.0029$ and $p = 0.0127$, respectively) (Figure 3.5). The RQ ratios of FOXP1 and FOXP4 homodimers (-1.39 ± 0.19) were not significantly different from one another; however, FOXP1/P4 heterodimers (0.32 ± 0.87) caused a significantly different effect than FOXP1 and FOXP4 homodimers ($p = 0.0116$ and $p = 0.0160$, respectively) (Figure 3.6). It is interesting to note that although FOXP1 and FOXP4 homodimers were comparable transcriptional repressors of *WISP2*, FOXP1/P4 heterodimers seem to nullify the transcriptional activity of its respective homodimers. There were no significant differences between FOXP2 homodimers, FOXP4 homodimers, and FOXP2/P4 heterodimers (-1.11 ± 0.05) in *WISP2* expression (Figure 3.7). Each of the FOXP heterodimeric combinations had a distinct influence over *WISP2* expression (Figure 3.8). FOXP1/P2 heterodimers are significantly ($p = 0.0009$) greater transcriptional repressors of *WISP2* than FOXP2/P4 heterodimers. In contrast, the FOXP1/P4 heterodimers are transcriptional activators.

PRICKLE1 (Prickle homolog 1)

The effect of FOXP1/2/4 dimerization on *PRICKLE1* expression was significant ($F(5, 12) = 7.65$, $p = 0.0019$). *Post hoc* analyses indicated that the RQ ratio of FOXP1 homodimers (-1.46 ± 0.23) was significantly ($p = 0.0003$) different from FOXP2 homodimers (1.68 ± 0.34) (Figure 3.5). FOXP1 and FOXP2 homodimers had opposing effects on PRICKLE1 expression: FOXP1 homodimers caused transcriptional repression but FOXP2 homodimers caused transcriptional activation. FOXP1/P2 heterodimers (0.23 ± 1.00) exhibited a slight upregulation of *PRICKLE1* that was at an intermediary level between the respective homodimers, but was not found to be statistically significant (Figure 3.5). Comparisons of RQ ratios between FOXP1 homodimers, FOXP4 homodimers, and FOXP1/P4 heterodimers did not reveal any significant differences (Figure 3.6). Although the expression level of *PRICKLE1* by FOXP2 homodimers was significantly different from FOXP4 homodimers and FOXP2/P4 heterodimers (-1.22 ± 0.10) at the $p = 0.0006$ level, the latter two dimer combinations demonstrated identical expression levels of *PRICKLE1* (Figure 3.7). Therefore, it does not appear that FOXP2 regulates *PRICKLE1* through heterodimeric interactions with FOXP4. The RQ ratios of FOXP1/P4 and FOXP2/P4 heterodimers were not significantly different (Figure 3.8). The RQ ratio of FOXP1/P2 heterodimers differed significantly ($p = 0.0405$) from FOXP2/P4 heterodimers.

EFNB3 (Ephrin-B3)

Statistical analysis revealed significant differences in *EFNB3* expression across FOXP1/2/4 protein combinations ($F(5, 12) = 7.32$, $p = 0.0023$). *Post hoc* tests showed

that the RQ ratio of FOXP1 homodimers (1.31 ± 0.05) was significantly different from FOXP2 homodimers (-1.74 ± 0.4, $p = 0.0028$) and FOXP1/P2 heterodimers (-2.04 ± 0.37, $p = 0.0014$), although the latter two protein combinations were not significantly different (Figure 3.5). The RQ ratio of FOXP1/P4 heterodimers (-2.39 ± 0.46) had a significantly different effect than FOXP1 and FOXP4 homodimers (0.30 ± 0.92) ($p = 0.0007$ and $p = 0.0064$, respectively) (Figure 3.6). *EFNB3* expression is differentially regulated between FOXP2 and FOXP4 homodimers ($p = 0.0275$) (Figure 3.7). The RQ ratio of FOXP2/P4 heterodimers (0.47 ± 0.79) was found to be significantly ($p = 0.185$) different from FOXP2 homodimers, but not FOXP4 homodimers (Figure 3.7). The expression of *EFNB3* activated by FOXP2/P4 heterodimers was found to be significantly different from FOXP1/P2 or FOXP1/P4 heterodimers ($p = 0.0094$ and $p = 0.0043$, respectively) (Figure 3.8). FOXP2/P4 heterodimers caused *EFNB3* upregulation while FOXP1/P2 and FOXP1/P4 heterodimers caused downregulation.

SLIT1 (Slit homolog 1)

The expression of SLIT1 was unanimously downregulated by FOXP1/2/4 homodimers and heterodimers (Figure 3.5-3.8). There was no significant main effect of FOXP1/2/4 dimeric combinations on *SLIT1* expression ($F(5, 12) = 1.82, p = 0.18$).

In summary, this study was able to confirm the functional regulation of downstream FOXP2 targets in a cell-based model system and found that FOXP1/2/4 transcription factors can act both as a repressor and an activator. The findings from the current study suggest that the relative levels of FOXP2 and its interacting co-factors, such as FOXP1 and FOXP4, affect its ability to act as a transcriptional activator or repressor.

Furthermore, this study provides evidence of the complexity of regulatory mechanisms mediated by the FOXP subgroup.

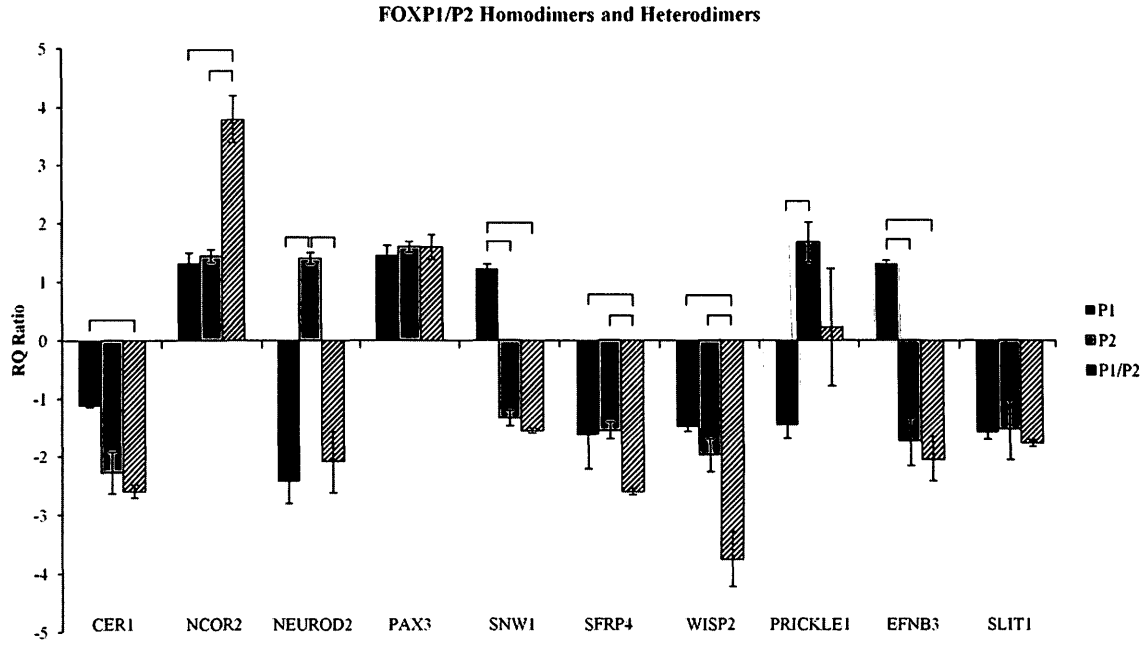

Figure 3.5. Differential expression of FOXP2 downstream genes mediated by FOXP1/P2 homodimers and heterodimers. FOXP1 (P1), FOXP2 (P2), or FOXP1/P2 (P1/P2) over-expression revealing functional regulation of downstream targets (*CER1*, *NCOR2*, *NEUROD2*, *PAX3*, *SNW1*, *SFRP4*, *WISP2*, *PRICKLE1*, *EFNB3*, and *SLIT1*) by P1 and P2 homodimers and P1/P2 heterodimers. Quantitative real-time polymerase chain reaction was performed using cDNA prepared from transfected HEK293 cell lines. Expression changes (Y-axis) are given as the mean of three cell passages from cells transfected with *FOXP1* and/or *FOXP2* compared with cells transfected with an empty vector control and are normalized for equal expression of the internal control, *HPRT* and *SDHA*. Genes with significant difference in expression between control and FOXP over-expressing cells are indicated with an asterisk (*). *$p<0.05$; Error Bars: ±SEM.

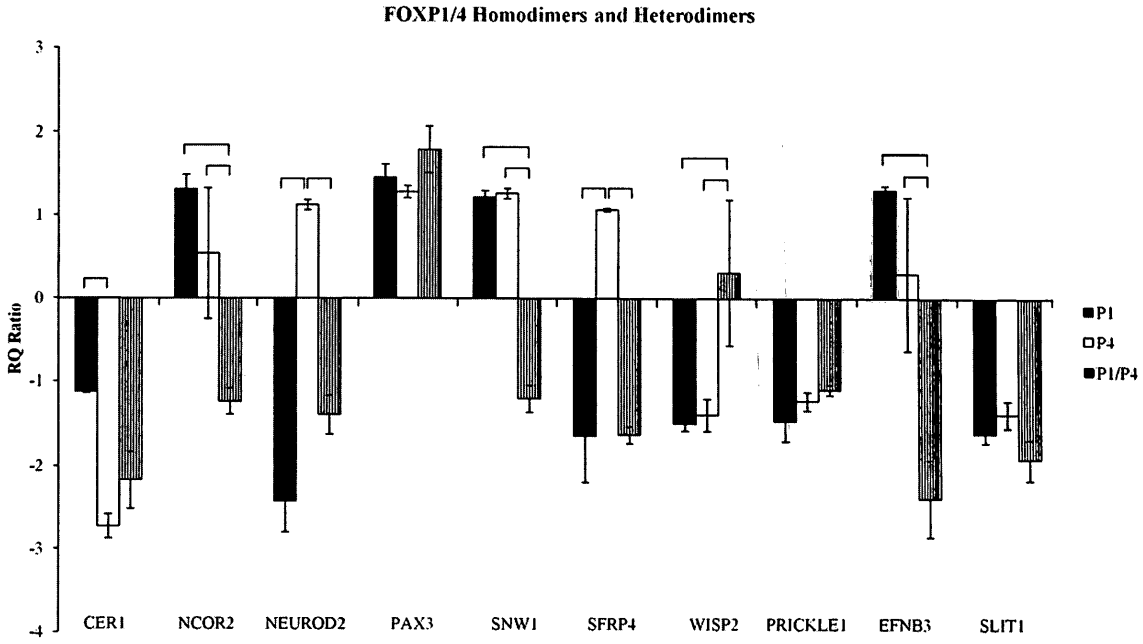

Figure 3.6. Differential expression of FOXP2 downstream genes mediated by FOXP1/P4 homodimers and heterodimers. FOXP1 (P1), FOXP4 (P4), or FOXP1/P4 (P1/P4) over-expression revealing functional regulation of downstream targets (*CER1*, *NCOR2*, *NEUROD2*, *PAX3*, *SNW1*, *SFRP4*, *WISP2*, *PRICKLE1*, *EFNB3*, and *SLIT1*) by P1 and P4 homodimers and P1/P4 heterodimers. Quantitative real-time polymerase chain reaction was performed using cDNA prepared from transfected HEK293 cell lines. Expression changes (Y-axis) are given as the mean of three cell passages from cells transfected with *FOXP1* and/or *FOXP4* compared with cells transfected with an empty vector control and are normalized for equal expression of the internal control, *HPRT* and *SDHA*. Genes with significant difference in expression between control and FOXP over-expressing cells are indicated with an asterisk (*). *$p<0.05$; Error Bars: ±SEM.

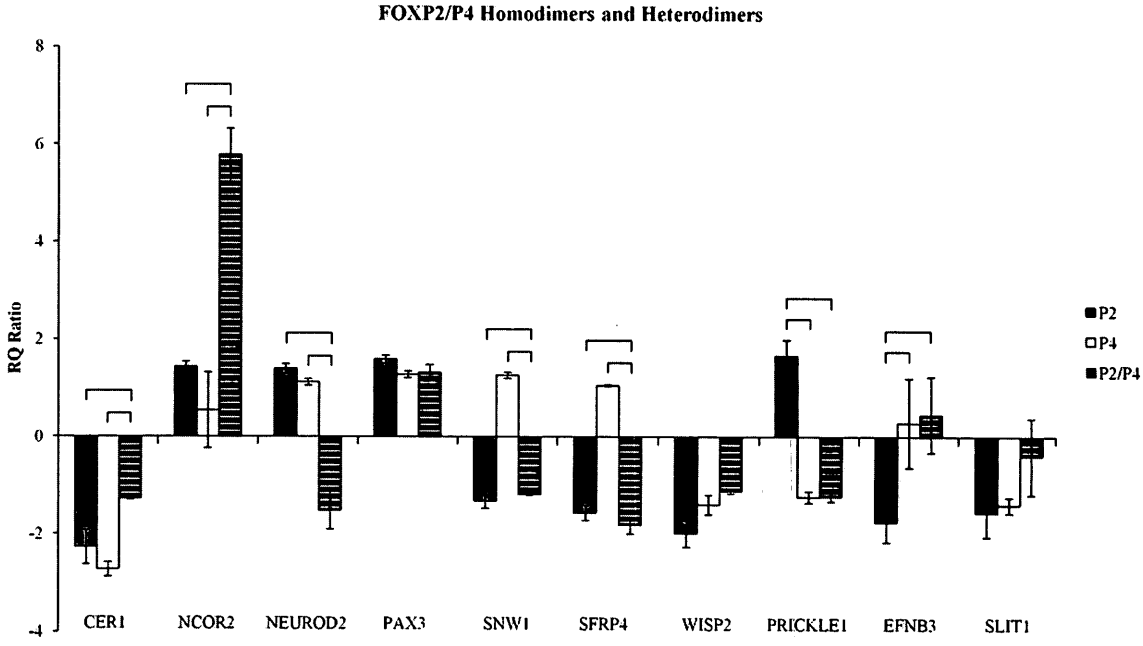

Figure 3.7. Differential expression of FOXP2 downstream genes mediated by FOXP2/P4 homodimers and heterodimers. FOXP1 (P2), FOXP4 (P4), or FOXP2/P4 (P2/P4) over-expression revealing functional regulation of downstream targets (*CER1*, *NCOR2*, *NEUROD2*, *PAX3*, *SNW1*, *SFRP4*, *WISP2*, *PRICKLE1*, *EFNB3*, and *SLIT1*) by P2 and P4 homodimers and P2/P4 heterodimers. Quantitative real-time polymerase chain reaction was performed using cDNA prepared from transfected HEK293 cell lines. Expression changes (Y-axis) are given as the mean of three cell passages from cells transfected with *FOXP2* and/or *FOXP4* compared with cells transfected with an empty vector control and are normalized for equal expression of the internal control, *HPRT* and *SDHA*. Genes with significant difference in expression between control and FOXP over-expressing cells are indicated with an asterisk (*). *$p<0.05$; Error Bars: ±SEM.

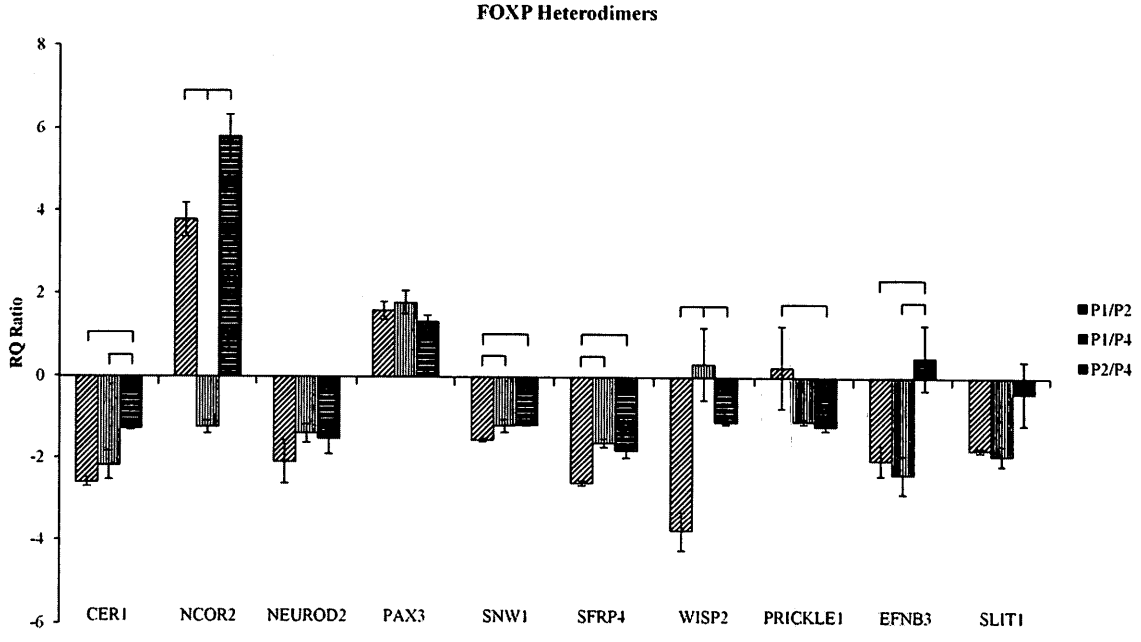

Figure 3.8. Differential expression of FOXP2 downstream genes mediated by FOXP1/2/4 heterodimers. FOXP1/P2 (P1/P2), FOXP1/P4 (P1/P4), or FOXP2/P4 (P2/P4) over-expression in HEK293 cell lines confirm that the combinatorial actions of FOXP1/2/4 transcription factors coordinately regulate the expression of downstream targets (*CER1*, *NCOR2*, *NEUROD2*, *PAX3*, *SNW1*, *SFRP4*, *WISP2*, *PRICKLE1*, *EFNB3*, and *SLIT1*). This graph shows the FOXP1/2/4 heterodimers together without the corresponding homodimeric combinations. Genes with significant difference in expression between control and FOXP over-expressing cells are indicated with an asterisk (*). *$p<0.05$; Error Bars: ±SEM.

3.5. Discussion

It is well-recognized that genetic mutations of *FOXP2* are associated with abnormalities in speech development and linguistic skills. Although *FOXP2* mutations are rare, it has been speculated that the downstream pathways it regulates could hold relevance for common language impairments.[221] However, since its discovery as a 'language gene', its precise functions from molecular and cellular perspectives have remained largely elusive. Previous studies have utilized a functional genomics approach to identify other elements of the pathways in which it participates, including targets that it regulates.[104,105] The identified putative targets of *FOXP2* represent a molecular entry point into the neural mechanisms underlying human speech and language.

The current study was undertaken to address the need for additional studies to determine which targets are most relevant for mediating the effects of *FOXP2* on speech and language. This study helped to elucidate the contributions of *FOXP1* and *FOXP4* in *FOXP2*-mediated regulation and the influence that FOXP1/2/4 dimerization has over the putative targets of *FOXP2*. In addition, this study investigated the downstream neural targets of *FOXP2* that are important for embryonic development, in order to examine the broader impact on language deficits and bridge together clinically distinct disorders on a molecular level, such as autism and 9p deletion syndrome.

3.5.1. Several Selected Gene Targets are Differentially Regulated by FOXP1/2/4 Homodimers and Heterodimers

The present study used HEK293 cell lines for stable transfections with FOXP1, FOXP2, and FOXP4 in all combinations or an empty vector control, and involved quantitative assessments of mRNA expression on a gene-by-gene basis with use of qRT-PCR. In line with previous studies suggesting that FOXP2 proteins act as repressors, six of the targets tested (*CER1*, *SNW1*, *SFRP4*, *WISP2*, *EFNB3*, and *SLIT1*) demonstrated reduced expression in the presence of FOXP2. Four genes (*NCOR2*, *NEUROD2*, *PAX3*, and *PRICKLE1*) displayed increases in transcription in response to FOXP2 over-expression in HEK293 cell lines, suggesting that the protein may also act as an activator for a subset of direct targets. The results indicating that FOXP2 has dual functionality, acting to either repress or activate gene expression, are parallel to that reported by Spiteri and Vernes and their colleagues.[104,105] Although activation of transcription has long been recognized as an essential component of gene regulation during development, the role of transcriptional repression programs is equally as crucial. Of the ten downstream targets investigated, eight genes (*CER1*, *NCOR2*, *NEUROD2*, *SNW1*, *SFRP4*, *WISP2*, *PRICKLE1*, and *EFNB3*) displayed significant effects of altered gene expression due to FOXP1/2/4 interaction. The ability of FOXP1/2/4 proteins to dimerize in many different combinations could be a way in which transcriptional regulation is fine tuned in a cell-specific manner. The results of this study did not demonstrate a clear synergistic pattern of FOXP1/2/4 dimerization preference across gene targets but rather, revealed that the

target genes have a preferred protein combination for transcriptional activation and/or repression that differed on a gene-to-gene basis.

3.5.2. WNT Signaling Modulation by FOXP1/2/4 Dimerization May Have Implications for Human Speech and Language

The WNT family of signaling proteins participate in multiple developmental events during embryogenesis and has also been implicated in adult tissue homeostasis.[222] WNT signaling is initiated by the binding of WNT ligands to a transmembrane FZD/LRP (frizzled/lipoprotein receptor-related protein) receptor complex, causing the activation of the cytoplasmic protein DVL (dishevelled). DVL inactivates GSK3β (glycogen synthase kinase-3β), which results in the accumulation and subsequent nuclear translocation of CTNNB1 (β-catenin). In the presence of appropriate co-activators, CTNNB1 binds to the TCF/LEF (T-cell factor/lymphoid enhancer factor) family of DNA-binding proteins and upregulates the transcription of various target genes that drive proliferation and cell fate determination. In the absence of WNT signaling, the activated GSK3β forms an intracellular complex that targets CTNNB1 for degradation via the ubiquitin-proteosome pathway.

CER1 (Cerberus 1)

The present study showed that the expression of *CER1* is repressed not just by FOXP2, but by other members of the FOXP subfamily as well. The results show that FOXP1 does not seem to influence *CER1* expression in a heterodimeric combination with FOXP2 or FOXP4. The data indicate that FOXP2 and FOXP4 homodimers were strong

transcriptional repressors of *CER1* (Figure 3.7). Conversely, FOXP2/P4 heterodimers are significantly weaker transcriptional repressor than its corresponding FOXP2 and FOXP4 homodimeric counterparts. Moreover, the FOXP2/P4 heterodimer was determined to be a significantly weaker transcriptional repressor than the FOXP1/P2 and FOXP1/FOXP4 heterodimers (Figure 3.8).

The *CER1* gene encodes a multifunctional antagonist that binds to Wnt, as well as Nodal and bone morphogenetic proteins (BMP), thus inhibiting the activity of these proteins in the extracellular space during embryogenesis.[210] In *Xenopus* embryos, *cerberus* has the unique property of inducing ectopic head formation without trunk formation.[223] Trunk formation relies on Nodal and Wnt signaling, whereas head induction requires inhibition of Wnt and BMP signaling. Studies of a mouse *CER1*-related gene, designated as *Cerr1*, have reported that *Cerr1* may play a role in anterior neural induction and somite formation during mouse development.[224]

Recently, Swinkels and colleagues reviewed the clinical features, developmental characteristics, and cytogenetic aberrations in 13 patients presenting with a 9p deletion, caused by a constitutional monosomy of part of the short arm of chromosome 9.[225] The deletion 9p syndrome is clinically characterized by dysmorphic facial features (i.e. trigonocephaly), hypotonia, and mental retardation. The authors of this paper had narrowed down the critical region for the deletion 9p syndrome to an area approximately 300 kb size, of which the *CER1* gene is located just distally and appeared to be a prime candidate for mutation analysis. Intriguingly, speech development was delayed in all patients, with test records from speech therapists showing speech disorders ranging from

mild to severe. Other behavioural problems included low concentration span, temper tantrums with head banging, sleep problems, and delayed motor development. These behaviours have been commonly described in patients with autism. Indeed, two patients showed definitive characteristics within the autism spectrum. Although *CER1* did not reveal any causative mutations, it may be hypothesized that altered gene dosages of *CER1* can still impact developmental processes that require spatial and temporal sensitivity. As demonstrated by this study, the imprecise combination of FOXP1/2/4 proteins can have indirect effects on the expression level of *CER1*.

To further cement the association of *FOX*-related genes in speech and language delays, another study pertaining to deletion 9p syndrome implicated the involvement of *FOXD4*, which was found to be deleted only amongst patients with speech and language deficits.[226] Notably, in addition to significant speech delays, the 9p deletion patients also exhibited autism-related characteristics. The finding that these 9p deletion patients have a heterozygous deletion of *FOXD4* implies that this gene may also have a role in speech and language development. The *FOXD4* gene belongs to the same winged helix/forkhead transcription factor gene family as *FOXP2* and shares a crucial forkhead box (FOX) domain, which is indicative of DNA binding.[227] Haploinsufficiency of *FOX* gene family members, such as *FOXP2* and *FOXD4*, may cause dosage imbalance in its target genes, which leads to abnormal neurological development.

With regard to brain development, altered expression of FOXP1/2/4 transcription factors may result in premature or delayed activation of multiple downstream targets that force biological processes out of step. Given the importance of WNT signaling during

embryogenesis, FOXP1/2/4 regulatory mechanisms that go awry during embryogenesis might lead to changed transcriptional outcomes that are potentially responsible for speech and language impairments.

SFRP4 (Secreted frizzled-related protein 4)

The results indicate that the expression of *SFRP4* was downregulated by FOXP1 and FOXP2 homodimers (Figure 3.5). Although FOXP1 and FOXP2 demonstrated equal ability to repress *SFRP4* expression, FOXP1/P2 heterodimers proved to be an even stronger transcriptional repressor. Moreover, a distinct role for FOXP4 homodimers in *SFRP4* expression was discovered. FOXP4 homodimers may be the only FOXP1/2/4 protein combination that causes *SFRP4* upregulation. Further, the results suggest that FOXP4 proteins are unlikely to regulate *SFRP4* expression via heterodimerization with FOXP1 (Figure 3.6). FOXP2/P4 heterodimers produced a different transcriptional effect and caused increased repression in comparison to FOXP2 and FOXP4 homodimers (Figure 3.7).

SFRP4 is another protein that has been demonstrated to be a Wnt antagonist, as well as a putative downstream target of *FOXP2*. SFRPs (secreted frizzled-related proteins) are structurally similar to FZD receptors but lack the transmembrane segment necessary for signal transduction.[228] SFRPs inhibit the Wnt signaling pathway by binding to either the Wnt ligand or FZD receptor.[229] The expression patterns of several SFRPs are complementary to that of Wnts when examined in developing mouse embryos.[41] One emerging explanation is that SFRPs facilitate boundary definition by limiting the range of Wnt activity. The expression of *SFRP4* in the embryo, particularly in the brain, makes

this an intriguing molecule to investigate.[228] Given the diversity of Wnt signaling effects on the developing embryo, defining the expression patterns of *SFRP4* and its modulator, FOXP2, is of potential significance in brain development. Dysregulation of *SFRP4* by FOXP1/2/4 proteins is likely to impact Wnt signaling and lead to alterations in cell fate and migration.

WISP2 (WNT1 inducible signaling pathway protein 2)

The results of this study suggest that various FOXP1/2/4 protein combinations have different regulatory roles in *WISP2* expression.For instance, even though FOXP1 and FOXP2 homodimers caused downregulation of *WISP2* expression, FOXP1/P2 heterodimers proved to be stronger transcriptional repressors (Figure 3.5). The results indicate that although FOXP1 and FOXP4 homodimers were comparable transcriptional repressors of *WISP2*, FOXP1/P4 heterodimers lead to *WISP2* expression levels that more closely resembled the baseline (Figure 3.6). This may represent a molecular mechanism whereby the strength of transcriptional repression is self-modulated. There were no differences between FOXP2 homodimers, FOXP4 homodimers, and FOXP2/P4 heterodimers; perhaps indicating the functional redundancy of some FOXP1/2/4 dimer combinations in the regulation of *WISP2* expression (Figure 3.7).

WISP2, another component of the Wnt signaling pathway, belongs to the *CCN* gene family. The acronym *CCN* is from the first three reported members of the family, namely *CYR61* (*cysteine rich 61*), *CTGF* (*connective tissue growth factor*), and *NOV* (*nephroblastoma overexpressed*), but has since grown to include the *WISP* genes (*WISP1*, *WISP2*, and *WISP3*).[230] Over-expression of Wnt1, either through retroviral infection or

transgenic mouse over-expression, was found to upregulate WISPs at both the mRNA and protein levels.[231] Although two independent experimental systems demonstrated that WISP2 induction was correlated with the expression of Wnt1, it is not clear whether WISP2 is directly or indirectly induced by the downstream components of the Wnt1 signaling pathway. Thus, *WISP2* expression could result from Wnt1 signaling directly through CTNNB1 transcription factor regulation or alternatively through Wnt1 signaling activating a transcription factor, which in turn regulates *WISP2*.[231] It is possible that the signaling machinery of the Wnt1 cascade may be engaged by FOXP1/2/4 transcription factors to modulate *WISP2* expression. In support of this hypothesis, it has been demonstrated in a fish experimental system that lef1, a member of the Tcf/Lef family of transcription factors activated by Wnt signaling, regulates foxP2 expression during CNS development.[232] Additionally, as previously mentioned, *WISP2* has been identified as a potential transcriptional target of FOXP2.

The present work shows that various FOXP protein combinations have different regulatory roles in *WISP2* expression. The diverse roles played by the different combinations of FOXP proteins in *WISP2* expression points towards a complex regulatory mechanism.

PRICKLE1 (Prickle homolog 1)

The results indicate that FOXP2 homodimers have differential effects on *PRICKLE1* expression than FOXP1 or FOXP4 homodimers (Figure 3.5 & 3.7). *PRICKLE1* was found to be transcriptionally activated by FOXP2 homodimers but transcriptionally repressed by FOXP1 and FOXP4 homodimers. The results also indicate

that FOXP2 does not appear to regulate *PRICKLE1* expression through heterodimeric interactions with FOXP4.

PRICKLE1 is a downstream Wnt signaling molecule that regulates gastrulation in the early embryonic stage.[213] Mouse *Prickle1*, and its related gene *Prickle2*, are abundantly expressed in the developing brain and may play important roles in the positive regulation of neurite outgrowth.[233] Previous investigations using chromatin immunoprecipitation coupled with hybridization to promoter microarrays (ChIP-chip) to identify direct targets of FOXP2 have found that neurite outgrowth and synaptic plasticity are the most prominent biological themes associated with Foxp2 function in the embryonic CNS.[234] These findings represent a starting point into the functional characterization of FOXP1/2/4 regulatory mechanisms, in particular the modulation of networks involved in neurite outgrowth and other core aspects of neurodevelopment.

3.5.3. Notch Signaling Modulation by FOXP1/2/4 Dimerization May Have Implications for Human Speech and Language

Notch-mediated intracellular signal transduction, an evolutionary conserved pathway, is reported to play seminal roles in cell fate control, influencing differentiation, proliferation, and apoptotic events, at all stages of development.[235] Notch is a transmembrane protein containing multiple epidermal growth factor (EGF)-like repeats in the extracellular domain and Ankyrin repeats in the intracellular domain.[236] Upon ligand binding, proteolytic cleavage of Notch near the transmembrane domain occurs, resulting in release of the Notch intracellular domain (NICD).[237] NICD is then released into the

cytoplasm and translocated into the nucleus to regulate transcription of Notch target genes. Upon translocation into the nucleus, NICD bind to RBPJ (recombining binding protein suppressor of hairless) which is a constitutive repressor of Notch signaling. RBPJ represses Notch target gene expression by recruiting a co-repressor complex, which includes NCOR1, NCOR2, SNW1, CIR, HDAC1, SPEN, FHL1, and SAP30. NICD binding to RBPJ replaces the co-repressor complex with a co-activator complex, which includes MAML1-3 and EP300. Transcriptional activation and repression mediated by sequence-specific DNA-binding factors underlie the binary decisions necessary for progression of cell determination events, and it is becoming increasingly clear that cofactors, including NCOR2 and SNW1, are indispensible for proper development in a variety of developmental systems.

NCOR2 (Nuclear receptor co-repressor 2)

The present study revealed that the expression of *NCOR2* is activated not only by FOXP2, but also by other members of the FOXP subfamily. FOXP1/P2 heterodimers were found to elicit an upregulation of *NCOR2* that far surpassed the expression levels produced by its respective homodimers (Figure 3.5). Likewise, FOXP2/P4 heterodimers also gave rise to an upregulation of *NCOR2* that exceeded the expression levels produced by its respective homodimers (Figure 3.7). Interesting, FOXP1/P4 heterodimers had opposing actions and instead caused the downregulation of *NCOR2* (Figure 3.6). These findings suggest that the relative levels of FOXP1/2/4 proteins, and thus the different combinations of protein dimers that may be formed, determine its ability to act as an activator or repressor.

A link between *NCOR2* and Rett syndrome, a neurodevelopmental disorder included in the autism spectrum, has been previously established. Rett syndrome, caused by mutations in the methyl-CpG binding protein MeCP2, is characterized by reduced brain growth, regressive effects on motor, speech, and cognitive development, and stereotyped hand-wringing behaviour.[238] A recent report demonstrates that MeCP2 cooperates with the NCOR2 co-repressor complex to inhibit the expression a *hairy*-related gene, a known target of the Notch signaling pathway that has significant effects on neurogenesis.[239] Notch signaling can regulate many aspects of neuronal maturation, including dendritic branching. In Rett syndrome, decreased dendritic branching in neurons has been found, leaving open the possibility that subtle alterations in Notch signaling may contribute to the Rett syndrome phenotype.[240]

Intriguingly, even as *NCOR2* is identified as a potential downstream target of FOXP2, a functional interaction between NCOR2 and FOXP1 has been reported.[241] NCOR2/FOXP1 protein complexes have been found to regulate a program of gene repression essential to proper myocardial development. Moreover, NCOR2/FOXP1 co-repressor complexes appear to mediate a more general strategy since NCOR2 and FOXP1 proteins also appear to be a component of *c-fms* regulation in monocytes. The data indicate that NCOR2-mediated co-repression may be a common mechanism by which FOXP1 and other FOXP proteins regulate gene expression programs in development of target organs.

The present study found that FOXP2 homodimers led to the upregulation of *NCOR2* and that a clear synergy was seen with FOXP1/P2 and FOXP2/P4 heterodimers,

where the cells expressing these combinations had a stronger NCOR2 induction than cells expressing them separately (Figure 3.5 & 3.7). Differential expression of *NCOR2* in the developing brain may have profound consequences. NCOR2 mRNA is primarily expressed in the ventricular zone region, where multipotent neural precursors reside.[242] NCOR2 has also been implicated in regulation of embryonic neural stem cell proliferation and differentiation, in which it may have an important role in controlling neural stem cell maintenance and lineage decisions, and plays a critical role in forebrain development.[243] The data suggest that NCOR2 and FOXP1/2/4 proteins may comprise a functional biological unit required to orchestrate specific programs critical for mammalian CNS development.

SNW1 (SNW domain containing 1)

In this study, FOXP1/2/4 heterodimers were observed to cause SNW1 downregulation but the homodimers demonstrated mixed regulatory roles. The data indicate that FOXP2 homodimers lead to the downregulation of *SNW1* and that FOXP2 proteins are unlikely to heterodimerize with FOXP1 and FOXP4 in *SNW1* regulation (Figure 3.5 & 3.7). FOXP1 and FOXP4 homodimers were found to upregulate the expression of *SNW1* to nearly equivalent levels (Figure 3.6). On the contrary, *SNW1* expression as influenced by FOXP1/P4 heterodimers was significantly different and caused downregulation instead.

In addition to participating in the Notch signaling pathway, SNW1 has been identified as a key regulator of BMP activity in vertebrate embryos.[244] BMPs are a family of morphogens that help direct early embryonic patterning.[245] BMP gradients provide

positional information to direct cell fate specification, such as patterning of the vertebrate ectoderm into the neural plate, neural crest, and epidermal tissues, with precise borders segregating these domains. In *Xenopus*, it was demonstrated that SNW1 is required for neural crest specification.[246] Further *in vivo* and *in vitro* assays using cell culture and tissue explants concluded that SNW1 acts upstream of the BMP receptors. BMP activity is normally concentrated in a tight horseshoe-shaped band that runs along the anterior edge and sides of the neural plate. Loss of SNW1 expression in the neural plate caused a strong drop in BMP signaling levels within this band of cells. When BMP is lost, the border between the neural plate and neighbouring tissues was found to be blurred and the neural crest tissue did not form. Therefore, SNW1 is an essential factor in determining proper embryonic development. Although the exact mechanisms by which SNW1 modulates BMP activity as yet remains unclear, the identification of this connection opens new avenues to better understand embryonic development.

3.5.4. FOXP1/2/4 Dimerization Regulates Nervous System Development and May Have Implications for Human Speech and Language

NEUROD2 (Neurogenic differentiation factor 2)

This study showed that the expression of *NEUROD2* was upregulated by FOXP2 and FOXP4 homodimers but downregulated by FOXP1 homodimers. The fact that *NEUROD2* expression by FOXP1 homodimers was significantly different from that of both FOXP2 and FOXP4 homodimers indicates a distinct role for FOXP1 homodimers in the regulation of *NEUROD2*. Although the results indicate that FOXP2 does not

heterodimerize with FOXP1, FOXP2 does appear to heterodimerize with FOXP4 where FOXP2/P4 heterodimers appear to cause transcriptional repression of *NEUROD2* (Figure 3.5 & 3.7). Additionally, FOXP1 is not likely to influence *NEUROD2* expression through interactions with FOXP4 (Figure 3.6). The results indicate that the relative levels of FOXP2 and its interacting co-factors determine the transcriptional activity (either upregulation or downregulation) of downstream FOXP2 targets, such as *NEUROD2*.

Transcription factors of the basic helix loop helix (bHLH) class have been shown to be essential for normal brain development in vertebrate species.[247] The members of the NEUROD subset of neurogenic bHLH transcription factors are characterized as neuronal differentiation genes because they induce cell cycle arrest in neural precursors and induce transcription of genes that contribute to mature neuronal phenotype.[248,249] Previous studies have shown that NeuroD2-null mice have defects in thalamocortical development in that axon terminals fail to segregate, barrel organization is disrupted, and synaptic transmission is defective.[217] Other studies have found that mice deficient for NeuroD2 experience excessive apoptosis in the CNS populations that normally express NeuroD2; in particular, immense granule cell loss was seen in postnatal cerebellar development.[249] These mice exhibit ataxia, motor deficits, postnatal premature death, and reduced seizure threshold. Mice that were heterozygous for NeuroD2 exhibited the same deficits as NeuroD2-null mice, though not as severe. This haploinsufficiency phenotype in mice raised the possibility that *NEUROD2* expression levels may influence human brain development and function. Since NeuroD2 participates in genesis of neurons involved in

learning, memory, motor coordination, and cranial nerve function, it is important to understand how this gene is regulated in the CNS.

FOXP2 transcription during CNS development is found predominantly in neural circuits that have been implicated in motor control, including the basal ganglia, the thalamus, the inferior olives, and cerebellum.[94] These brain structures are intricately interconnected to subserve motor-related functions: the basal ganglia modulate activity of the premotor and prefrontal cortical areas through connections projecting through the globus pallidus, substantia nigra, and thalamus, while the cerebellum plays a role in regulating motor coordination, receiving input from the inferior olives.[250] The data implicating *FOXP2* in the development of motor-related circuits during embryogenesis may account for the oromotor difficulties of humans with *FOXP2* mutations. It is thought that the accompanying linguistic and grammatical impairments observed in individuals with FOXP2 mutations are secondary consequences of basic deficits in motor planning and sequencing.[94] There is an intriguing level of concordance between *NeuroD2* and *FOXP2* mutations in that both are associated with significant structural abnormalities in the cerebellum. Studies of patients with acquired lesions have highlighted a cerebellar role in procedural learning, particularly in linguistic functions.[251] Therefore, the expression data of the present study draws attention to a possible link between the olivocerebellar system and FOXP2-NEUROD2 regulation.

PAX3 (Paired box 3)

In this study, FOXP2 homodimers caused the upregulation of *PAX3*, as did the FOXP1 and FOXP4 homodimers, although there was no significant difference in effect.

The FOXP1/2/4 heterodimers also did not demonstrate significantly different transcriptional outcomes of *PAX3*, perhaps indicating the functional redundancy of FOXP1/2/4 proteins in *PAX3* regulation.

PAX3 encodes a transcription factor belonging to a family of paired-domain containing proteins and is specifically expressed in the dorsal and posterior neural tube.[218] The importance of PAX3 in dorsal neural tube and neural crest patterning and differentiation is evident in human disorders associated with *PAX3* mutations, namely Waardenburg syndromes type I and type III.[252] Waardenburg syndrome is the most common inherited form of congenital deafness and has been associated with cognitive delay and other neurological abnormalities as well.[253,254] A case report of two individuals with Waardenburg syndrome highlights the presence of moderate to severe intellectual disability, history of developmental delay, and autism.[254] It remains to be conclusively determined whether there is a higher prevalence of autism in individuals with Waardenburg syndrome. If so, this association would be of interest in relation to the autism susceptibility gene *FOXP2*.

EFNB3 (Ephrin-B3)

The results of this study indicate that FOXP1/2/4 homodimers and heterodimers have differential effects on EFNB3 that, depending on the dimer combination, may lead to transcriptional activation or repression. Whereas FOXP2 homodimers caused transcriptional repression, FOXP1 and FOXP4 homodimers were responsible for transcriptional activation (Figure 3.5 & 3.7). FOXP1/P4 heterodimers led to the

downregulation of *EFNB3* (Figure 3.6). The results suggest that FOXP2 is unlikely to influence *EFNB3* expression through heterodimerization (Figure 3.5 & 3.7).

The Eph receptors are a family of growth and guidance molecules that are widely expressed in the developing nervous system. These receptor proteins and their ephrin ligands have well-known roles in the establishment of neuronal connectivity by guiding axons to the appropriate targets and regulating the formation of synaptic connections.[255] EphB receptors are known to be expressed in dendrites and transducer postsynaptic forward signals that promote the assembly, maturation, and plasticity of synapses.[256] Ephrin-B (EFNB) proteins are expressed in axons and dendrites and function to transducer reverse signals important for axon guidance and pruning, presynaptic development, and synaptic plasticity.[219]

It has been proposed that at around the time of language emergence in humans, the amino acid composition in the human variant of *FOXP2* has undergone accelerated evolution.[257] Because FOXP2 has an important role in speech and language in humans, the identified differentially regulated targets between humans and chimpanzees may have a critical function in the development and evolution of language circuitry in humans. Of particular relevance, *EFNB2* is a well-validated direct target of FOXP2 that has been identified as one of the genes that is differentially regulated between human and chimpanzee *FOXP2* orthologues.[234] Now, the related gene *EFNB3* is proposed to be a downstream FOXP2 target as well.[104,105] Examining *EFNB3* is a matter of particular interest because dysfunctional axon guidance signaling is thought to underlie the microstructural abnormalities in the brains of individuals affected with autism.[234] The

mRNA level of *EFNB3* has been found to be significantly lower in the anterior cingulated cortex of brains from individuals with autism compared to controls; this region has previously been implicated in the pathophysiology of autism.[258] As mentioned previously, studies using ChIP-chip to identify direct targets of FOXP2 found that neurite outgrowth and synaptic plasticity are strongly represented biological themes associated with FOXP2 function in the CNS.[234]

The present study found that FOXP1/2/4 homodimers and heterodimers have differential effects on *EFNB3* that, depending on the dimer combination, may lead to transcriptional activation or repression. Genes controlling neurite outgrowth or axon guidance during embryonic development have crucial roles in the maturation and stabilization of synaptic connectivity. Hence, the strong impact of FOXP2 on *EFNB3*, a gene that influences neurite outgrowth and axon guidance, might reflect major FOXP2 functions that are relevant to development and maintenance of neuronal networks. Dysregulation of this regulatory network might have implications for language abilities.

SLIT1 (Slit homolog 1)

The current study revealed that FOXP1/2/4 proteins transcriptionally repressed the expression of *SLIT1*. There were no significant differences in transcriptional outcomes between FOXP1/2/4 dimer combinations, suggesting that there is functional redundancy of these proteins in *SLIT1* regulation.

To achieve the connectivity underlying normal brain function, axons grow to form pathways, taking directional cues at specific points in space and time.[259] These decisions depend on contacts of growing axons with guidance molecules, which are tightly

controlled. Slit proteins are secreted guidance molecules produced by midline glial cells that function as a chemorepellant to prevent the re-crossing of commissural axons.[260] This effect is mediated by the Robo (Roundabout) receptors. In addition to acting as a chemorepellant for cortical axons, SLIT1 has been reported to regulate dendritic development.[220] As mentioned previously, genes that control axon guidance during embryogenesis have essential roles in the development of neural circuits.

3.6. Conclusions

The development and functioning of brain circuits underlying human speech result from complex sequential processes that are tightly regulated. Deciphering the neural pathways that are associated with speech-related circuits of the brain has been a challenge considering that speech is a unique human trait. Nevertheless, progress has been made by studying families where impairment of speech-associated processes is inherited in a monogenic fashion. A gene that predisposes individuals to speech and language disorders had not been conceived of until *FOXP2* was discovered.[82] Thus, the opportunity to pinpoint genes that signify key entry points into the molecular networks underlying speech-related processes was made possible. In addition to the characterization of *FOXP2* in various animal models, subsequent studies have focused on the identification of FOXP2-targeted promoter regions.[104,105] Given that *FOXP2* encodes a transcription factor, the emerging picture is that some putative target genes in which it regulates in the brain are likely to represent important players in speech-associated pathways.

This study has provided novel information on the potential role of FOXP2, in relation to its cofactors FOXP1 and FOXP4, in speech and language functions. Although the existing literature provides some evidence for the specific dimerization between FOXP proteins, this is the first study showing how dimerization of FOXP1/2/4 proteins can influence the expression of various target genes. Studies regarding FOXP2 and its cofactors will help uncover the regulatory influence of FOXP1/2/4 over genes that are potentially involved in the molecular mechanism underlying language acquisition and contribute to the understanding of language-related deficits in neurodevelopmental disorders such as autism. Furthermore, understanding the genes and pathways that are regulated by FOXP1/2/4 might lead to the discovery of candidate autism susceptibility genes.

3.7. Limitations and Future Areas of Study

The present study used an over-expression system to investigate the regulatory functions of FOXP1/2/4. Thus, the reported results only pertain to the regulation of downstream target genes under artificial conditions of high-level expression of FOXP1/2/4 proteins. This may represent a study limitation because FOXP1/2/4 proteins may demonstrate preferential binding partners under varying levels of FOXP1, FOXP2, or FOXP4 expression. For instance, high levels of FOXP2 may lead to homodimerization with itself but with decreasing amounts of FOXP2, there may be competition of FOXP2 homodimers with FOXP1 or FOXP4, possibly leading to changed preferential binding

partners and altered transcriptional outcomes. This speculation warrants further investigation.

For future experiments, it would be interesting to determine the localization of downstream targets of FOXP2 in the developing brain to see if its expression coincides with that of FOXP1/2/4 transcription factors. The evidence obtained will likely continue to provide new insights into the neural mechanisms behind speech and language.

REFERENCES

1. CDC. Autism Spectrum Disorders - Data and Statistics. 2011; http://www.cdc.gov/ncbddd/autism/data.html. Accessed December 20, 2011.
2. Fombonne E. Epidemiological surveys of autism and other pervasive developmental disorders: an update. *J Autism Dev Disord.* Aug 2003;33(4):365-382.
3. Kanner L. The birth of early infantile autism. *J Autism Child Schizophr.* Apr-Jun 1973;3(2):93-95.
4. Williams R. Autism Through Ages Baffles Science. 2000; http://www.pediatricservices.com/prof/prof-26.htm. Accessed December 20, 2011.
5. Attwood T. *Asperger's Syndrome: A Guide for Parents and Professionals* London: Jessica Kingsley Publishers; 1998.
6. Wing L. Asperger's syndrome: a clinical account. *Psychol Med.* Feb 1981;11(1):115-129.
7. Weintraub K. The prevalence puzzle: Autism counts. *Nature.* Nov 3 2011;479(7371):22-24.
8. Rogers SJ. Brief report: early intervention in autism. *J Autism Dev Disord.* Apr 1996;26(2):243-246.
9. Seltzer MM, Krauss MW, Shattuck PT, Orsmond G, Swe A, Lord C. The symptoms of autism spectrum disorders in adolescence and adulthood. *J Autism Dev Disord.* Dec 2003;33(6):565-581.
10. Association AP. *Diagnostic and statistical manual of mental disorders* 4ed. Washington, DC: American Psychiatric Association; 1994.
11. Filipek PA, Accardo PJ, Baranek GT, et al. The screening and diagnosis of autistic spectrum disorders. *J Autism Dev Disord.* Dec 1999;29(6):439-484.
12. Osterling JA, Dawson G, Munson JA. Early recognition of 1-year-old infants with autism spectrum disorder versus mental retardation. *Dev Psychopathol.* Spring 2002;14(2):239-251.
13. Mars AE, Mauk JE, Dowrick PW. Symptoms of pervasive developmental disorders as observed in prediagnostic home videos of infants and toddlers. *J Pediatr.* Mar 1998;132(3 Pt 1):500-504.
14. Folstein SE, Rosen-Sheidley B. Genetics of autism: complex aetiology for a heterogeneous disorder. *Nat Rev Genet.* Dec 2001;2(12):943-955.
15. DiCicco-Bloom E, Lord C, Zwaigenbaum L, et al. The developmental neurobiology of autism spectrum disorder. *J Neurosci.* Jun 28 2006;26(26):6897-6906.
16. Rapin I, Dunn M. Language disorders in children with autism. *Semin Pediatr Neurol.* Jun 1997;4(2):86-92.
17. Baron-Cohen S. Social and pragmatic deficits in autism: cognitive or affective? *J Autism Dev Disord.* Sep 1988;18(3):379-402.

18. Jyonouchi H, Geng L, Streck DL, Toruner GA. Children with autism spectrum disorders (ASD) who exhibit chronic gastrointestinal (GI) symptoms and marked fluctuation of behavioral symptoms exhibit distinct innate immune abnormalities and transcriptional profiles of peripheral blood (PB) monocytes. *J Neuroimmunol.* Sep 15 2011;238(1-2):73-80.
19. Maski KP, Jeste SS, Spence SJ. Common neurological co-morbidities in autism spectrum disorders. *Curr Opin Pediatr.* Dec 2011;23(6):609-615.
20. Landa RJ, Holman KC, Garrett-Mayer E. Social and communication development in toddlers with early and later diagnosis of autism spectrum disorders. *Arch Gen Psychiatry.* Jul 2007;64(7):853-864.
21. Werner E, Dawson G, Munson J, Osterling J. Variation in early developmental course in autism and its relation with behavioral outcome at 3-4 years of age. *J Autism Dev Disord.* Jun 2005;35(3):337-350.
22. Smalley SL. Autism and tuberous sclerosis. *J Autism Dev Disord.* Oct 1998;28(5):407-414.
23. Bailey A, Bolton P, Butler L, et al. Prevalence of the fragile X anomaly amongst autistic twins and singletons. *J Child Psychol Psychiatry.* Jul 1993;34(5):673-688.
24. McPartland J, Klin A. Asperger's syndrome. *Adolesc Med Clin.* Oct 2006;17(3):771-788; abstract xiii.
25. de Bildt A, Sytema S, Ketelaars C, et al. Interrelationship between Autism Diagnostic Observation Schedule-Generic (ADOS-G), Autism Diagnostic Interview-Revised (ADI-R), and the Diagnostic and Statistical Manual of Mental Disorders (DSM-IV-TR) classification in children and adolescents with mental retardation. *J Autism Dev Disord.* Apr 2004;34(2):129-137.
26. Mesibov GB, Schopler E, Schaffer B, Michal N. Use of the childhood autism rating scale with autistic adolescents and adults. *J Am Acad Child Adolesc Psychiatry.* Jul 1989;28(4):538-541.
27. Bailey A, Le Couteur A, Gottesman I, et al. Autism as a strongly genetic disorder: evidence from a British twin study. *Psychol Med.* Jan 1995;25(1):63-77.
28. London EA. The environment as an etiologic factor in autism: a new direction for research. *Environ Health Perspect.* Jun 2000;108 Suppl 3:401-404.
29. Hurley RS, Losh M, Parlier M, Reznick JS, Piven J. The broad autism phenotype questionnaire. *J Autism Dev Disord.* Oct 2007;37(9):1679-1690.
30. Szatmari P, Paterson AD, Zwaigenbaum L, et al. Mapping autism risk loci using genetic linkage and chromosomal rearrangements. *Nat Genet.* Mar 2007;39(3):319-328.
31. Muhle R, Trentacoste SV, Rapin I. The genetics of autism. *Pediatrics.* May 2004;113(5):e472-486.
32. Spence SJ. The genetics of autism. *Semin Pediatr Neurol.* Sep 2004;11(3):196-204.
33. Avdjieva-Tzavella D. Genetic and Non Genetic Aspects of Autism Spectrum Disorders. *Balk J Med Gen.* 2008;11(1):3-10.

34. Gurrieri F, Battaglia A, Torrisi L, et al. Pervasive developmental disorder and epilepsy due to maternally derived duplication of 15q11-q13. *Neurology.* May 12 1999;52(8):1694-1697.
35. Delahanty RJ, Kang JQ, Brune CW, et al. Maternal transmission of a rare GABRB3 signal peptide variant is associated with autism. *Mol Psychiatry.* Jan 2011;16(1):86-96.
36. Ashley-Koch A, Wolpert CM, Menold MM, et al. Genetic studies of autistic disorder and chromosome 7. *Genomics.* Nov 1 1999;61(3):227-236.
37. Cukier HN, Skaar DA, Rayner-Evans MY, et al. Identification of chromosome 7 inversion breakpoints in an autistic family narrows candidate region for autism susceptibility. *Autism Res.* Oct 2009;2(5):258-266.
38. Schroer RJ, Phelan MC, Michaelis RC, et al. Autism and maternally derived aberrations of chromosome 15q. *Am J Med Genet.* Apr 1 1998;76(4):327-336.
39. Wassink TH, Piven J, Vieland VJ, et al. Evidence supporting WNT2 as an autism susceptibility gene. *Am J Med Genet.* Jul 8 2001;105(5):406-413.
40. Gong X, Jia M, Ruan Y, et al. Association between the FOXP2 gene and autistic disorder in Chinese population. *Am J Med Genet B Neuropsychiatr Genet.* May 15 2004;127B(1):113-116.
41. Kim AS, Lowenstein DH, Pleasure SJ. Wnt receptors and Wnt inhibitors are expressed in gradients in the developing telencephalon. *Mech Dev.* May 2001;103(1-2):167-172.
42. Hurst JA, Baraitser M, Auger E, Graham F, Norell S. An extended family with a dominantly inherited speech disorder. *Dev Med Child Neurol.* Apr 1990;32(4):352-355.
43. Takahashi H, Takahashi K, Liu FC. FOXP genes, neural development, speech and language disorders. *Adv Exp Med Biol.* 2009;665:117-129.
44. Li H, Yamagata T, Mori M, Momoi MY. Absence of causative mutations and presence of autism-related allele in FOXP2 in Japanese autistic patients. *Brain Dev.* Apr 2005;27(3):207-210.
45. Buxbaum JD, Silverman JM, Smith CJ, et al. Association between a GABRB3 polymorphism and autism. *Mol Psychiatry.* 2002;7(3):311-316.
46. Smit DJ, de Geus EJ, van de Nieuwenhuijzen ME, et al. Scale-free modulation of resting-state neuronal oscillations reflects prolonged brain maturation in humans. *J Neurosci.* Sep 14 2011;31(37):13128-13136.
47. Rodier P, Hyman S. Early Environmental Factors in Autism. *Ment Retard Dev D R.* 1998;4:121–128.
48. Moore K, Persaud T. *The Developing Human: Clinically Oriented Embryology.* sixth ed. Philadephia: W.B. Saunders; 1998.
49. Chaudhuri JD. Alcohol and the developing fetus--a review. *Med Sci Monit.* Sep-Oct 2000;6(5):1031-1041.
50. Nanson JL. Autism in fetal alcohol syndrome: a report of six cases. *Alcohol Clin Exp Res.* Jun 1992;16(3):558-565.

51. Ortega Garcia JA, Angulo MG, Sobrino-Najul EJ, et al. Prenatal exposure of a girl with autism spectrum disorder to 'horsetail' (Equisetum arvense) herbal remedy and alcohol: a case report. *J Med Case Reports.* 2011;5(1):129.
52. Christianson AL, Chesler N, Kromberg JG. Fetal valproate syndrome: clinical and neuro-developmental features in two sibling pairs. *Dev Med Child Neurol.* Apr 1994;36(4):361-369.
53. Williams PG, Hersh JH. A male with fetal valproate syndrome and autism. *Dev Med Child Neurol.* Sep 1997;39(9):632-634.
54. Bandim JM, Ventura LO, Miller MT, Almeida HC, Costa AE. Autism and Mobius sequence: an exploratory study of children in northeastern Brazil. *Arq Neuropsiquiatr.* Jun 2003;61(2A):181-185.
55. Suk WA, Murray K, Avakian MD. Environmental hazards to children's health in the modern world. *Mutat Res.* Nov 2003;544(2-3):235-242.
56. Chess S. Autism in children with congenital rubella. *J Autism Child Schizophr.* Jan-Mar 1971;1(1):33-47.
57. Yamashita Y, Fujimoto C, Nakajima E, Isagai T, Matsuishi T. Possible association between congenital cytomegalovirus infection and autistic disorder. *J Autism Dev Disord.* Aug 2003;33(4):455-459.
58. Ghaziuddin M, Al-Khouri I, Ghaziuddin N. Autistic symptoms following herpes encephalitis. *Eur Child Adolesc Psychiatry.* Jun 2002;11(3):142-146.
59. de Vries HE, Kuiper J, de Boer AG, Van Berkel TJ, Breimer DD. The blood-brain barrier in neuroinflammatory diseases. *Pharmacol Rev.* Jun 1997;49(2):143-155.
60. Vargas DL, Nascimbene C, Krishnan C, Zimmerman AW, Pardo CA. Neuroglial activation and neuroinflammation in the brain of patients with autism. *Ann Neurol.* Jan 2005;57(1):67-81.
61. Cohly HH, Panja A. Immunological findings in autism. *Int Rev Neurobiol.* 2005;71:317-341.
62. Capuron L, Miller AH. Immune system to brain signaling: neuropsychopharmacological implications. *Pharmacol Ther.* May 2011;130(2):226-238.
63. Meyer U, Feldon J, Schedlowski M, Yee BK. Immunological stress at the maternal-foetal interface: a link between neurodevelopment and adult psychopathology. *Brain Behav Immun.* Jul 2006;20(4):378-388.
64. Pinto-Martin JA, Levy SE, Feldman JF, Lorenz JM, Paneth N, Whitaker AH. Prevalence of autism spectrum disorder in adolescents born weighing <2000 grams. *Pediatrics.* Nov 2011;128(5):883-891.
65. Kolevzon A, Gross R, Reichenberg A. Prenatal and perinatal risk factors for autism: a review and integration of findings. *Arch Pediatr Adolesc Med.* Apr 2007;161(4):326-333.
66. Rutter M. Incidence of autism spectrum disorders: changes over time and their meaning. *Acta Paediatr.* Jan 2005;94(1):2-15.
67. Lawler CP, Croen LA, Grether JK, Van de Water J. Identifying environmental contributions to autism: provocative clues and false leads. *Ment Retard Dev Disabil Res Rev.* 2004;10(4):292-302.

68. Lord C, Risi S, DiLavore PS, Shulman C, Thurm A, Pickles A. Autism from 2 to 9 years of age. *Arch Gen Psychiatry.* Jun 2006;63(6):694-701.

69. Corsello C. Early Intervention in Autism. *Infant Young Child.* 2005;18(2):74-85.

70. Herbert MR, Russo JP, Yang S, et al. Autism and environmental genomics. *Neurotoxicology.* Sep 2006;27(5):671-684.

71. Pierce K, Glatt SJ, Liptak GS, McIntyre LL. The power and promise of identifying autism early: insights from the search for clinical and biological markers. *Ann Clin Psychiatry.* Jul-Sep 2009;21(3):132-147.

72. Arcellana-Panlilio M, Robbins SM. Cutting-edge technology. I. Global gene expression profiling using DNA microarrays. *Am J Physiol Gastrointest Liver Physiol.* Mar 2002;282(3):G397-402.

73. Purcell AE, Jeon OH, Zimmerman AW, Blue ME, Pevsner J. Postmortem brain abnormalities of the glutamate neurotransmitter system in autism. *Neurology.* Nov 13 2001;57(9):1618-1628.

74. Baron CA, Liu SY, Hicks C, Gregg JP. Utilization of lymphoblastoid cell lines as a system for the molecular modeling of autism. *J Autism Dev Disord.* Nov 2006;36(8):973-982.

75. Lintas C, Sacco R, Persico AM. Genome-wide expression studies in autism spectrum disorder, Rett syndrome, and Down syndrome. *Neurobiol Dis.* Jan 2012;45(1):57-68.

76. Seno MM, Hu P, Gwadry FG, et al. Gene and miRNA expression profiles in autism spectrum disorders. *Brain Res.* Mar 22 2011;1380:85-97.

77. Ehli EA, Lengyel-Nelson T, Hudziak JJ, Davies GE. Using a commercially available DNA extraction kit to obtain high quality human genomic DNA suitable for PCR and genotyping from 11-year-old saliva saturated cotton spit wads. *BMC Res Notes.* 2008;1:133.

78. Kupfer DM, White VL, Jenkins MC, Burian D. Examining smoking-induced differential gene expression changes in buccal mucosa. *BMC Med Genomics.* 2010;3:24.

79. Smith RV, Schlecht NF, Childs G, Prystowsky MB, Belbin TJ. Pilot study of mucosal genetic differences in early smokers and nonsmokers. *Laryngoscope.* Aug 2006;116(8):1375-1379.

80. Vargha-Khadem F, Gadian DG, Copp A, Mishkin M. FOXP2 and the neuroanatomy of speech and language. *Nat Rev Neurosci.* Feb 2005;6(2):131-138.

81. Fisher SE, Vargha-Khadem F, Watkins KE, Monaco AP, Pembrey ME. Localisation of a gene implicated in a severe speech and language disorder. *Nat Genet.* Feb 1998;18(2):168-170.

82. Lai CS, Fisher SE, Hurst JA, Vargha-Khadem F, Monaco AP. A forkhead-domain gene is mutated in a severe speech and language disorder. *Nature.* Oct 4 2001;413(6855):519-523.

83. Folstein SE, Mankoski RE. Chromosome 7q: where autism meets language disorder? *Am J Hum Genet.* Aug 2000;67(2):278-281.

84. Barrett S, Beck JC, Bernier R, et al. An autosomal genomic screen for autism. Collaborative linkage study of autism. *Am J Med Genet.* Dec 15 1999;88(6):609-615.

85. Tomblin JB, Hafeman LL, O'Brien M. Autism and autism risk in siblings of children with specific language impairment. *Int J Lang Commun Disord.* Jul-Sep 2003;38(3):235-250.

86. Wassink TH, Piven J, Vieland VJ, et al. Evaluation of FOXP2 as an autism susceptibility gene. *Am J Med Genet.* Jul 8 2002;114(5):566-569.

87. Marui T, Koishi S, Funatogawa I, et al. No association of FOXP2 and PTPRZ1 on 7q31 with autism from the Japanese population. *Neurosci Res.* Sep 2005;53(1):91-94.

88. Carlsson P, Mahlapuu M. Forkhead transcription factors: key players in development and metabolism. *Dev Biol.* Oct 1 2002;250(1):1-23.

89. Li B, Greene MI. FOXP3 actively represses transcription by recruiting the HAT/HDAC complex. *Cell Cycle.* Jun 15 2007;6(12):1432-1436.

90. Santos ME, Athanasiadis A, Leitao AB, DuPasquier L, Sucena E. Alternative splicing and gene duplication in the evolution of the FoxP gene subfamily. *Mol Biol Evol.* Jan 2011;28(1):237-247.

91. Wang B, Weidenfeld J, Lu MM, et al. Foxp1 regulates cardiac outflow tract, endocardial cushion morphogenesis and myocyte proliferation and maturation. *Development.* Sep 2004;131(18):4477-4487.

92. Rousso DL, Gaber ZB, Wellik D, Morrisey EE, Novitch BG. Coordinated actions of the forkhead protein Foxp1 and Hox proteins in the columnar organization of spinal motor neurons. *Neuron.* Jul 31 2008;59(2):226-240.

93. Marcus GF, Fisher SE. FOXP2 in focus: what can genes tell us about speech and language? *Trends Cogn Sci.* Jun 2003;7(6):257-262.

94. Lai CS, Gerrelli D, Monaco AP, Fisher SE, Copp AJ. FOXP2 expression during brain development coincides with adult sites of pathology in a severe speech and language disorder. *Brain.* Nov 2003;126(Pt 11):2455-2462.

95. Groszer M, Keays DA, Deacon RM, et al. Impaired synaptic plasticity and motor learning in mice with a point mutation implicated in human speech deficits. *Curr Biol.* Mar 11 2008;18(5):354-362.

96. French CA, Groszer M, Preece C, Coupe AM, Rajewsky K, Fisher SE. Generation of mice with a conditional Foxp2 null allele. *Genesis.* Jul 2007;45(7):440-446.

97. Shu W, Cho JY, Jiang Y, et al. Altered ultrasonic vocalization in mice with a disruption in the Foxp2 gene. *Proc Natl Acad Sci U S A.* Jul 5 2005;102(27):9643-9648.

98. Haesler S, Rochefort C, Georgi B, Licznerski P, Osten P, Scharff C. Incomplete and inaccurate vocal imitation after knockdown of FoxP2 in songbird basal ganglia nucleus Area X. *PLoS Biol.* Dec 2007;5(12):e321.

99. Teufel A, Wong EA, Mukhopadhyay M, Malik N, Westphal H. FoxP4, a novel forkhead transcription factor. *Biochim Biophys Acta.* Jun 19 2003;1627(2-3):147-152.

100. Li S, Zhou D, Lu MM, Morrisey EE. Advanced cardiac morphogenesis does not require heart tube fusion. *Science.* Sep 10 2004;305(5690):1619-1622.
101. Tam WY, Leung CK, Tong KK, Kwan KM. Foxp4 is essential in maintenance of Purkinje cell dendritic arborization in the mouse cerebellum. *Neuroscience.* Jan 13 2011;172:562-571.
102. Li S, Weidenfeld J, Morrisey EE. Transcriptional and DNA binding activity of the Foxp1/2/4 family is modulated by heterotypic and homotypic protein interactions. *Mol Cell Biol.* Jan 2004;24(2):809-822.
103. Fisher SE, Scharff C. FOXP2 as a molecular window into speech and language. *Trends Genet.* Apr 2009;25(4):166-177.
104. Spiteri E, Konopka G, Coppola G, et al. Identification of the transcriptional targets of FOXP2, a gene linked to speech and language, in developing human brain. *Am J Hum Genet.* Dec 2007;81(6):1144-1157.
105. Vernes SC, Spiteri E, Nicod J, et al. High-throughput analysis of promoter occupancy reveals direct neural targets of FOXP2, a gene mutated in speech and language disorders. *Am J Hum Genet.* Dec 2007;81(6):1232-1250.
106. Coo H, Ouellette-Kuntz H, Lam M, et al. Correlates of age at diagnosis of autism spectrum disorders in six Canadian regions. *Chronic Dis Inj Can.* Mar 2012;32(2):90-100.
107. Rodier PM. Chronology of neuron development: animal studies and their clinical implications. *Dev Med Child Neurol.* Aug 1980;22(4):525-545.
108. Thompson KL, Pine PS, Rosenzweig BA, Turpaz Y, Retief J. Characterization of the effect of sample quality on high density oligonucleotide microarray data using progressively degraded rat liver RNA. *BMC Biotechnol.* 2007;7:57.
109. Diaz E, Barisone GA. DNA microarrays: sample quality control, array hybridization and scanning. *J Vis Exp.* 2011(49).
110. Marshall AJ, Du Q, Draves KE, Shikishima Y, HayGlass KT, Clark EA. FDC-SP, a novel secreted protein expressed by follicular dendritic cells. *J Immunol.* Sep 1 2002;169(5):2381-2389.
111. Park LS, Friend D, Gillis S, Urdal DL. Characterization of the cell surface receptor for human granulocyte/macrophage colony-stimulating factor. *J Exp Med.* Jul 1 1986;164(1):251-262.
112. Zanoni I, Ostuni R, Marek LR, et al. CD14 controls the LPS-induced endocytosis of Toll-like receptor 4. *Cell.* Nov 11 2011;147(4):868-880.
113. Schweizer J, Bowden PE, Coulombe PA, et al. New consensus nomenclature for mammalian keratins. *J Cell Biol.* Jul 17 2006;174(2):169-174.
114. Drews K, Jozefczuk J, Prigione A, Adjaye J. Human induced pluripotent stem cells--from mechanisms to clinical applications. *J Mol Med (Berl).* Jul 2012;90(7):735-745.
115. Nguyen TT, Ma LN, Slovak ML, Bangs CD, Cherry AM, Arber DA. Identification of novel Runx1 (AML1) translocation partner genes SH3D19, YTHDf2, and ZNF687 in acute myeloid leukemia. *Genes Chromosomes Cancer.* Oct 2006;45(10):918-932.

116. Polyak K, Kato JY, Solomon MJ, et al. p27Kip1, a cyclin-Cdk inhibitor, links transforming growth factor-beta and contact inhibition to cell cycle arrest. *Genes Dev.* Jan 1994;8(1):9-22.

117. Olayioye MA, Neve RM, Lane HA, Hynes NE. The ErbB signaling network: receptor heterodimerization in development and cancer. *EMBO J.* Jul 3 2000;19(13):3159-3167.

118. Whitson RH, Tsark W, Huang TH, Itakura K. Neonatal mortality and leanness in mice lacking the ARID transcription factor Mrf-2. *Biochem Biophys Res Commun.* Dec 26 2003;312(4):997-1004.

119. Baker TA, Sauer RT. ClpXP, an ATP-powered unfolding and protein-degradation machine. *Biochim Biophys Acta.* Jan 2012;1823(1):15-28.

120. de Marco MC, Martin-Belmonte F, Kremer L, et al. MAL2, a novel raft protein of the MAL family, is an essential component of the machinery for transcytosis in hepatoma HepG2 cells. *J Cell Biol.* Oct 14 2002;159(1):37-44.

121. Hosack DA, Dennis G, Jr., Sherman BT, Lane HC, Lempicki RA. Identifying biological themes within lists of genes with EASE. *Genome Biol.* 2003;4(10):R70.

122. Rudney JD, Chen R. The vital status of human buccal epithelial cells and the bacteria associated with them. *Arch Oral Biol.* Apr 2006;51(4):291-298.

123. Livy A, Lye S, Jagdish CK, et al. Evaluation of Quality of DNA Extracted from Buccal Swabs for Microarray Based Genotyping. *Indian J Clin Biochem.* 2011.

124. Dalton P, Deacon R, Blamire A, et al. Maternal neuronal antibodies associated with autism and a language disorder. *Ann Neurol.* Apr 2003;53(4):533-537.

125. Yonk LJ, Warren RP, Burger RA, et al. CD4+ helper T cell depression in autism. *Immunol Lett.* Sep 1990;25(4):341-345.

126. Jyonouchi H, Sun S, Le H. Proinflammatory and regulatory cytokine production associated with innate and adaptive immune responses in children with autism spectrum disorders and developmental regression. *J Neuroimmunol.* Nov 1 2001;120(1-2):170-179.

127. Zimmerman AW, Jyonouchi H, Comi AM, et al. Cerebrospinal fluid and serum markers of inflammation in autism. *Pediatr Neurol.* Sep 2005;33(3):195-201.

128. Enstrom AM, Lit L, Onore CE, et al. Altered gene expression and function of peripheral blood natural killer cells in children with autism. *Brain Behav Immun.* Jan 2009;23(1):124-133.

129. Ashwood P, Corbett BA, Kantor A, Schulman H, Van de Water J, Amaral DG. In search of cellular immunophenotypes in the blood of children with autism. *PLoS One.* 2011;6(5):e19299.

130. Al-Alwan M, Du Q, Hou S, et al. Follicular dendritic cell secreted protein (FDC-SP) regulates germinal center and antibody responses. *J Immunol.* Jun 15 2007;178(12):7859-7867.

131. Chez MG, Dowling T, Patel PB, Khanna P, Kominsky M. Elevation of tumor necrosis factor-alpha in cerebrospinal fluid of autistic children. *Pediatr Neurol.* Jun 2007;36(6):361-365.

132. Zhan Y, Xu Y, Lew AM. The regulation of the development and function of dendritic cell subsets by GM-CSF: More than a hematopoietic growth factor. *Mol Immunol.* Aug 2012;52(1):30-37.

133. McLay RN, Kimura M, Banks WA, Kastin AJ. Granulocyte-macrophage colony-stimulating factor crosses the blood--brain and blood--spinal cord barriers. *Brain.* Nov 1997;120 (Pt 11):2083-2091.

134. Kannan Y, Moriyama M, Sugano T, et al. Neurotrophic action of interleukin 3 and granulocyte-macrophage colony-stimulating factor on murine sympathetic neurons. *Neuroimmunomodulation.* 2000;8(3):132-141.

135. Ashwood P, Krakowiak P, Hertz-Picciotto I, Hansen R, Pessah IN, Van de Water J. Altered T cell responses in children with autism. *Brain Behav Immun.* Jul 2011;25(5):840-849.

136. Franzen R, Bouhy D, Schoenen J. Nervous system injury: focus on the inflammatory cytokine 'granulocyte-macrophage colony stimulating factor'. *Neurosci Lett.* May 6 2004;361(1-3):76-78.

137. Dhib-Jalbut S, Arnold DL, Cleveland DW, et al. Neurodegeneration and neuroprotection in multiple sclerosis and other neurodegenerative diseases. *J Neuroimmunol.* Jul 2006;176(1-2):198-215.

138. Morgan JT, Chana G, Pardo CA, et al. Microglial activation and increased microglial density observed in the dorsolateral prefrontal cortex in autism. *Biol Psychiatry.* Aug 15 2010;68(4):368-376.

139. Bundschuh DS, Barsig J, Hartung T, et al. Granulocyte-macrophage colony-stimulating factor and IFN-gamma restore the systemic TNF-alpha response to endotoxin in lipopolysaccharide-desensitized mice. *J Immunol.* Mar 15 1997;158(6):2862-2871.

140. Hanisch UK. Microglia as a source and target of cytokines. *Glia.* Nov 2002;40(2):140-155.

141. Reddy PH, Manczak M, Zhao W, et al. Granulocyte-macrophage colony-stimulating factor antibody suppresses microglial activity: implications for anti-inflammatory effects in Alzheimer's disease and multiple sclerosis. *J Neurochem.* Dec 2009;111(6):1514-1528.

142. Parracho HM, Bingham MO, Gibson GR, McCartney AL. Differences between the gut microflora of children with autistic spectrum disorders and that of healthy children. *J Med Microbiol.* Oct 2005;54(Pt 10):987-991.

143. Emanuele E, Orsi P, Boso M, et al. Low-grade endotoxemia in patients with severe autism. *Neurosci Lett.* Mar 8 2010;471(3):162-165.

144. Lloyd-Jones KL, Kelly MM, Kubes P. Varying importance of soluble and membrane CD14 in endothelial detection of lipopolysaccharide. *J Immunol.* Jul 15 2008;181(2):1446-1453.

145. Sweeten TL, Posey DJ, McDougle CJ. High blood monocyte counts and neopterin levels in children with autistic disorder. *Am J Psychiatry.* Sep 2003;160(9):1691-1693.

146. Zawada AM, Rogacev KS, Rotter B, et al. SuperSAGE evidence for CD14++CD16+ monocytes as a third monocyte subset. *Blood.* Sep 22 2011;118(12):e50-61.
147. Urra X, Villamor N, Amaro S, et al. Monocyte subtypes predict clinical course and prognosis in human stroke. *J Cereb Blood Flow Metab.* May 2009;29(5):994-1002.
148. Serbina NV, Jia T, Hohl TM, Pamer EG. Monocyte-mediated defense against microbial pathogens. *Annu Rev Immunol.* 2008;26:421-452.
149. Mobley JL, Leininger M, Madore S, Baginski TJ, Renkiewicz R. Genetic evidence of a functional monocyte dichotomy. *Inflammation.* Dec 2007;30(6):189-197.
150. Ashwood P, Enstrom A, Krakowiak P, et al. Decreased transforming growth factor beta1 in autism: a potential link between immune dysregulation and impairment in clinical behavioral outcomes. *J Neuroimmunol.* Nov 15 2008;204(1-2):149-153.
151. Enstrom A, Krakowiak P, Onore C, et al. Increased IgG4 levels in children with autism disorder. *Brain Behav Immun.* Mar 2009;23(3):389-395.
152. Ziats MN, Rennert OM. Expression profiling of autism candidate genes during human brain development implicates central immune signaling pathways. *PLoS One.* 2011;6(9):e24691.
153. Giri DK, Ali-Seyed M, Li LY, et al. Endosomal transport of ErbB-2: mechanism for nuclear entry of the cell surface receptor. *Mol Cell Biol.* Dec 2005;25(24):11005-11018.
154. Yoo HJ, Cho IH, Park M, et al. Association between PTGS2 polymorphism and autism spectrum disorders in Korean trios. *Neurosci Res.* Sep 2008;62(1):66-69.
155. Anton ES, Marchionni MA, Lee KF, Rakic P. Role of GGF/neuregulin signaling in interactions between migrating neurons and radial glia in the developing cerebral cortex. *Development.* Sep 1997;124(18):3501-3510.
156. Rio C, Rieff HI, Qi P, Khurana TS, Corfas G. Neuregulin and erbB receptors play a critical role in neuronal migration. *Neuron.* Jul 1997;19(1):39-50.
157. Anton ES, Ghashghaei HT, Weber JL, et al. Receptor tyrosine kinase ErbB4 modulates neuroblast migration and placement in the adult forebrain. *Nat Neurosci.* Dec 2004;7(12):1319-1328.
158. Fisahn A, Neddens J, Yan L, Buonanno A. Neuregulin-1 modulates hippocampal gamma oscillations: implications for schizophrenia. *Cereb Cortex.* Mar 2009;19(3):612-618.
159. Kao WT, Wang Y, Kleinman JE, et al. Common genetic variation in Neuregulin 3 (NRG3) influences risk for schizophrenia and impacts NRG3 expression in human brain. *Proc Natl Acad Sci U S A.* Aug 31 2010;107(35):15619-15624.
160. Zhang HX, Li WQ, Zhang HS, et al. [Expressional changes of neuregulin-1 gene mRNA in peripheral blood from schizophrenia patients]. *Zhonghua Yi Xue Yi Chuan Xue Za Zhi.* Dec 2011;28(6):620-624.

161. Silberberg G, Darvasi A, Pinkas-Kramarski R, Navon R. The involvement of ErbB4 with schizophrenia: association and expression studies. *Am J Med Genet B Neuropsychiatr Genet.* Mar 5 2006;141B(2):142-148.

162. Zou H, Yu Y, Sheikh AM, et al. Association of upregulated Ras/Raf/ERK1/2 signaling with autism. *Genes Brain Behav.* Jul 2011;10(5):615-624.

163. Brunet A, Datta SR, Greenberg ME. Transcription-dependent and -independent control of neuronal survival by the PI3K-Akt signaling pathway. *Curr Opin Neurobiol.* Jun 2001;11(3):297-305.

164. Prigent SA, Gullick WJ. Identification of c-erbB-3 binding sites for phosphatidylinositol 3'-kinase and SHC using an EGF receptor/c-erbB-3 chimera. *EMBO J.* Jun 15 1994;13(12):2831-2841.

165. Ali IU, Schriml LM, Dean M. Mutational spectra of PTEN/MMAC1 gene: a tumor suppressor with lipid phosphatase activity. *J Natl Cancer Inst.* Nov 17 1999;91(22):1922-1932.

166. Zhou J, Parada LF. PTEN signaling in autism spectrum disorders. *Curr Opin Neurobiol.* Jun 2 2012.

167. Butler MG, Dasouki MJ, Zhou XP, et al. Subset of individuals with autism spectrum disorders and extreme macrocephaly associated with germline PTEN tumour suppressor gene mutations. *J Med Genet.* Apr 2005;42(4):318-321.

168. Courchesne E, Karns CM, Davis HR, et al. Unusual brain growth patterns in early life in patients with autistic disorder: an MRI study. *Neurology.* Jul 24 2001;57(2):245-254.

169. Persico AM, Bourgeron T. Searching for ways out of the autism maze: genetic, epigenetic and environmental clues. *Trends Neurosci.* Jul 2006;29(7):349-358.

170. Wei H, Malik M, Sheikh AM, Merz G, Ted Brown W, Li X. Abnormal cell properties and down-regulated FAK-Src complex signaling in B lymphoblasts of autistic subjects. *Am J Pathol.* Jul 2011;179(1):66-74.

171. Glickman G. Circadian rhythms and sleep in children with autism. *Neurosci Biobehav Rev.* Apr 2010;34(5):755-768.

172. Melke J, Goubran Botros H, Chaste P, et al. Abnormal melatonin synthesis in autism spectrum disorders. *Mol Psychiatry.* Jan 2008;13(1):90-98.

173. Nicholas B, Rudrasingham V, Nash S, Kirov G, Owen MJ, Wimpory DC. Association of Per1 and Npas2 with autistic disorder: support for the clock genes/social timing hypothesis. *Mol Psychiatry.* Jun 2007;12(6):581-592.

174. Hu VW, Sarachana T, Kim KS, et al. Gene expression profiling differentiates autism case-controls and phenotypic variants of autism spectrum disorders: evidence for circadian rhythm dysfunction in severe autism. *Autism Res.* Apr 2009;2(2):78-97.

175. Nguyen A, Rauch TA, Pfeifer GP, Hu VW. Global methylation profiling of lymphoblastoid cell lines reveals epigenetic contributions to autism spectrum disorders and a novel autism candidate gene, RORA, whose protein product is reduced in autistic brain. *FASEB J.* Aug 2010;24(8):3036-3051.

176. Sarachana T, Xu M, Wu RC, Hu VW. Sex hormones in autism: androgens and estrogens differentially and reciprocally regulate RORA, a novel candidate gene for autism. *PLoS One.* 2011;6(2):e17116.
177. Kirsch F, Kruger C, Schneider A. The receptor for granulocyte-colony stimulating factor (G-CSF) is expressed in radial glia during development of the nervous system. *BMC Dev Biol.* 2008;8:32.
178. Campbell K, Gotz M. Radial glia: multi-purpose cells for vertebrate brain development. *Trends Neurosci.* May 2002;25(5):235-238.
179. Noctor SC, Flint AC, Weissman TA, Dammerman RS, Kriegstein AR. Neurons derived from radial glial cells establish radial units in neocortex. *Nature.* Feb 8 2001;409(6821):714-720.
180. Malatesta P, Hartfuss E, Gotz M. Isolation of radial glial cells by fluorescent-activated cell sorting reveals a neuronal lineage. *Development.* Dec 2000;127(24):5253-5263.
181. Quan N, He L, Lai W, Shen T, Herkenham M. Induction of IkappaBalpha mRNA expression in the brain by glucocorticoids: a negative feedback mechanism for immune-to-brain signaling. *J Neurosci.* Sep 1 2000;20(17):6473-6477.
182. Yamagata K, Matsumura K, Inoue W, et al. Coexpression of microsomal-type prostaglandin E synthase with cyclooxygenase-2 in brain endothelial cells of rats during endotoxin-induced fever. *J Neurosci.* Apr 15 2001;21(8):2669-2677.
183. Ikeda-Matsuo Y, Ikegaya Y, Matsuki N, Uematsu S, Akira S, Sasaki Y. Microglia-specific expression of microsomal prostaglandin E2 synthase-1 contributes to lipopolysaccharide-induced prostaglandin E2 production. *J Neurochem.* Sep 2005;94(6):1546-1558.
184. Patterson PH. Maternal infection and immune involvement in autism. *Trends Mol Med.* Jul 2011;17(7):389-394.
185. Engman ML, Lewensohn-Fuchs I, Mosskin M, Malm G. Congenital cytomegalovirus infection: the impact of cerebral cortical malformations. *Acta Paediatr.* Sep 2010;99(9):1344-1349.
186. Rosen NJ, Yoshida CK, Croen LA. Infection in the first 2 years of life and autism spectrum disorders. *Pediatrics.* Jan 2007;119(1):e61-69.
187. Enstrom AM, Onore CE, Van de Water JA, Ashwood P. Differential monocyte responses to TLR ligands in children with autism spectrum disorders. *Brain Behav Immun.* Jan 2010;24(1):64-71.
188. Guerra-Crespo M, Perez-Monter C, Janga SC, et al. Transcriptional profiling of fetal hypothalamic TRH neurons. *BMC Genomics.* 2011;12:222.
189. Perez-Monter C, Martinez-Armenta M, Miquelajauregui A, et al. The Kruppel-like factor 4 controls biosynthesis of thyrotropin-releasing hormone during hypothalamus development. *Mol Cell Endocrinol.* Feb 20 2011;333(2):127-133.
190. Hashimoto T, Aihara R, Tayama M, Miyazaki M, Shirakawa Y, Kuroda Y. Reduced thyroid-stimulating hormone response to thyrotropin-releasing hormone in autistic boys. *Dev Med Child Neurol.* Apr 1991;33(4):313-319.
191. Zeng H, Schimpf BA, Rohde AD, Pavlova MN, Gragerov A, Bergmann JE. Thyrotropin-releasing hormone receptor 1-deficient mice display increased

depression and anxiety-like behavior. *Mol Endocrinol.* Nov 2007;21(11):2795-2804.

192. Hazlett HC, Poe MD, Gerig G, et al. Early brain overgrowth in autism associated with an increase in cortical surface area before age 2 years. *Arch Gen Psychiatry.* May 2011;68(5):467-476.

193. Hazlett HC, Poe M, Gerig G, et al. Magnetic resonance imaging and head circumference study of brain size in autism: birth through age 2 years. *Arch Gen Psychiatry.* Dec 2005;62(12):1366-1376.

194. Casaccia-Bonnefil P, Tikoo R, Kiyokawa H, Friedrich V, Jr., Chao MV, Koff A. Oligodendrocyte precursor differentiation is perturbed in the absence of the cyclin-dependent kinase inhibitor p27Kip1. *Genes Dev.* Sep 15 1997;11(18):2335-2346.

195. Zwaigenbaum L, Bryson S, Rogers T, Roberts W, Brian J, Szatmari P. Behavioral manifestations of autism in the first year of life. *Int J Dev Neurosci.* Apr-May 2005;23(2-3):143-152.

196. Threadgill DW, Dlugosz AA, Hansen LA, et al. Targeted disruption of mouse EGF receptor: effect of genetic background on mutant phenotype. *Science.* Jul 14 1995;269(5221):230-234.

197. Lee KF, Simon H, Chen H, Bates B, Hung MC, Hauser C. Requirement for neuregulin receptor erbB2 in neural and cardiac development. *Nature.* Nov 23 1995;378(6555):394-398.

198. Gassmann M, Casagranda F, Orioli D, et al. Aberrant neural and cardiac development in mice lacking the ErbB4 neuregulin receptor. *Nature.* Nov 23 1995;378(6555):390-394.

199. Riethmacher D, Sonnenberg-Riethmacher E, Brinkmann V, Yamaai T, Lewin GR, Birchmeier C. Severe neuropathies in mice with targeted mutations in the ErbB3 receptor. *Nature.* Oct 16 1997;389(6652):725-730.

200. Kulkarni PS, Sundqvist K, Betsholtz C, et al. Characterization of human buccal epithelial cells transfected with the simian virus 40 T-antigen gene. *Carcinogenesis.* Oct 1995;16(10):2515-2521.

201. Ramirez RD, Sheridan S, Girard L, et al. Immortalization of human bronchial epithelial cells in the absence of viral oncoproteins. *Cancer Res.* Dec 15 2004;64(24):9027-9034.

202. Chakravarti N, Kadara H, Yoon DJ, et al. Differential inhibition of protein translation machinery by curcumin in normal, immortalized, and malignant oral epithelial cells. *Cancer Prev Res (Phila).* Mar 2010;3(3):331-338.

203. A full genome screen for autism with evidence for linkage to a region on chromosome 7q. International Molecular Genetic Study of Autism Consortium. *Hum Mol Genet.* Mar 1998;7(3):571-578.

204. An autosomal genomic screen for autism. *Am J Med Genet.* Dec 8 2001;105(8):609-615.

205. Poot M, Beyer V, Schwaab I, et al. Disruption of CNTNAP2 and additional structural genome changes in a boy with speech delay and autism spectrum disorder. *Neurogenetics.* Feb 2010;11(1):81-89.

206. Gauthier J, Joober R, Mottron L, et al. Mutation screening of FOXP2 in individuals diagnosed with autistic disorder. *Am J Med Genet A.* Apr 15 2003;118A(2):172-175.

207. Pariani MJ, Spencer A, Graham JM, Jr., Rimoin DL. A 785kb deletion of 3p14.1p13, including the FOXP1 gene, associated with speech delay, contractures, hypertonia and blepharophimosis. *Eur J Med Genet.* Mar-Jun 2009;52(2-3):123-127.

208. Hamdan FF, Daoud H, Rochefort D, et al. De novo mutations in FOXP1 in cases with intellectual disability, autism, and language impairment. *Am J Hum Genet.* Nov 12 2010;87(5):671-678.

209. Takahashi K, Liu FC, Hirokawa K, Takahashi H. Expression of Foxp4 in the developing and adult rat forebrain. *J Neurosci Res.* Nov 1 2008;86(14):3106-3116.

210. Piccolo S, Agius E, Leyns L, et al. The head inducer Cerberus is a multifunctional antagonist of Nodal, BMP and Wnt signals. *Nature.* Feb 25 1999;397(6721):707-710.

211. Muley A, Majumder S, Kolluru GK, et al. Secreted frizzled-related protein 4: an angiogenesis inhibitor. *Am J Pathol.* Mar 2010;176(3):1505-1516.

212. Katsube K, Sakamoto K, Tamamura Y, Yamaguchi A. Role of CCN, a vertebrate specific gene family, in development. *Dev Growth Differ.* Jan 2009;51(1):55-67.

213. Carreira-Barbosa F, Concha ML, Takeuchi M, Ueno N, Wilson SW, Tada M. Prickle 1 regulates cell movements during gastrulation and neuronal migration in zebrafish. *Development.* Sep 2003;130(17):4037-4046.

214. Kao HY, Ordentlich P, Koyano-Nakagawa N, et al. A histone deacetylase corepressor complex regulates the Notch signal transduction pathway. *Genes Dev.* Aug 1 1998;12(15):2269-2277.

215. Zhou S, Fujimuro M, Hsieh JJ, et al. SKIP, a CBF1-associated protein, interacts with the ankyrin repeat domain of NotchIC To facilitate NotchIC function. *Mol Cell Biol.* Apr 2000;20(7):2400-2410.

216. Hayward P, Kalmar T, Arias AM. Wnt/Notch signalling and information processing during development. *Development.* Feb 2008;135(3):411-424.

217. Ince-Dunn G, Hall BJ, Hu SC, et al. Regulation of thalamocortical patterning and synaptic maturation by NeuroD2. *Neuron.* Mar 2 2006;49(5):683-695.

218. Bang AG, Papalopulu N, Kintner C, Goulding MD. Expression of Pax-3 is initiated in the early neural plate by posteriorizing signals produced by the organizer and by posterior non-axial mesoderm. *Development.* May 1997;124(10):2075-2085.

219. Xu NJ, Sun S, Gibson JR, Henkemeyer M. A dual shaping mechanism for postsynaptic ephrin-B3 as a receptor that sculpts dendrites and synapses. *Nat Neurosci.* Nov 2011;14(11):1421-1429.

220. Whitford KL, Marillat V, Stein E, et al. Regulation of cortical dendrite development by Slit-Robo interactions. *Neuron.* Jan 3 2002;33(1):47-61.

221. Vernes SC, Fisher SE. Unravelling neurogenetic networks implicated in developmental language disorders. *Biochem Soc Trans.* Dec 2009;37(Pt 6):1263-1269.

222. Nelson WJ, Nusse R. Convergence of Wnt, beta-catenin, and cadherin pathways. *Science.* Mar 5 2004;303(5663):1483-1487.

223. Bouwmeester T, Kim S, Sasai Y, Lu B, De Robertis EM. Cerberus is a head-inducing secreted factor expressed in the anterior endoderm of Spemann's organizer. *Nature.* Aug 15 1996;382(6592):595-601.

224. Shawlot W, Deng JM, Behringer RR. Expression of the mouse cerberus-related gene, Cerr1, suggests a role in anterior neural induction and somitogenesis. *Proc Natl Acad Sci U S A.* May 26 1998;95(11):6198-6203.

225. Swinkels ME, Simons A, Smeets DF, et al. Clinical and cytogenetic characterization of 13 Dutch patients with deletion 9p syndrome: Delineation of the critical region for a consensus phenotype. *Am J Med Genet A.* Jun 1 2008;146A(11):1430-1438.

226. Hauge X, Raca G, Cooper S, et al. Detailed characterization of, and clinical correlations in, 10 patients with distal deletions of chromosome 9p. *Genet Med.* Aug 2008;10(8):599-611.

227. Kaestner KH, Knochel W, Martinez DE. Unified nomenclature for the winged helix/forkhead transcription factors. *Genes Dev.* Jan 15 2000;14(2):142-146.

228. Rattner A, Hsieh JC, Smallwood PM, et al. A family of secreted proteins contains homology to the cysteine-rich ligand-binding domain of frizzled receptors. *Proc Natl Acad Sci U S A.* Apr 1 1997;94(7):2859-2863.

229. Bafico A, Gazit A, Pramila T, Finch PW, Yaniv A, Aaronson SA. Interaction of frizzled related protein (FRP) with Wnt ligands and the frizzled receptor suggests alternative mechanisms for FRP inhibition of Wnt signaling. *J Biol Chem.* Jun 4 1999;274(23):16180-16187.

230. Brigstock DR, Goldschmeding R, Katsube KI, et al. Proposal for a unified CCN nomenclature. *Mol Pathol.* Apr 2003;56(2):127-128.

231. Pennica D, Swanson TA, Welsh JW, et al. WISP genes are members of the connective tissue growth factor family that are up-regulated in wnt-1-transformed cells and aberrantly expressed in human colon tumors. *Proc Natl Acad Sci U S A.* Dec 8 1998;95(25):14717-14722.

232. Bonkowsky JL, Wang X, Fujimoto E, Lee JE, Chien CB, Dorsky RI. Domain-specific regulation of foxP2 CNS expression by lef1. *BMC Dev Biol.* 2008;8:103.

233. Okuda H, Miyata S, Mori Y, Tohyama M. Mouse Prickle1 and Prickle2 are expressed in postmitotic neurons and promote neurite outgrowth. *FEBS Lett.* Oct 2 2007;581(24):4754-4760.

234. Vernes SC, Oliver PL, Spiteri E, et al. Foxp2 regulates gene networks implicated in neurite outgrowth in the developing brain. *PLoS Genet.* Jul 2011;7(7):e1002145.

235. Artavanis-Tsakonas S, Rand MD, Lake RJ. Notch signaling: cell fate control and signal integration in development. *Science.* Apr 30 1999;284(5415):770-776.

236. Blaumueller CM, Qi H, Zagouras P, Artavanis-Tsakonas S. Intracellular cleavage of Notch leads to a heterodimeric receptor on the plasma membrane. *Cell.* Jul 25 1997;90(2):281-291.

237. Kandasamy K, Mohan SS, Raju R, et al. NetPath: a public resource of curated signal transduction pathways. *Genome Biol.* 2010;11(1):R3.

238. Amir RE, Van den Veyver IB, Wan M, Tran CQ, Francke U, Zoghbi HY. Rett syndrome is caused by mutations in X-linked MECP2, encoding methyl-CpG-binding protein 2. *Nat Genet.* Oct 1999;23(2):185-188.

239. Vetter ML. Methylation gets SMRT. Functional insights into Rett syndrome. *Dev Cell.* Sep 2003;5(3):359-360.

240. Armstrong DD, Dunn K, Antalffy B. Decreased dendritic branching in frontal, motor and limbic cortex in Rett syndrome compared with trisomy 21. *J Neuropathol Exp Neurol.* Nov 1998;57(11):1013-1017.

241. Jepsen K, Gleiberman AS, Shi C, Simon DI, Rosenfeld MG. Cooperative regulation in development by SMRT and FOXP1. *Genes Dev.* Mar 15 2008;22(6):740-745.

242. Perissi V, Jepsen K, Glass CK, Rosenfeld MG. Deconstructing repression: evolving models of co-repressor action. *Nat Rev Genet.* Feb 2010;11(2):109-123.

243. Jepsen K, Solum D, Zhou T, et al. SMRT-mediated repression of an H3K27 demethylase in progression from neural stem cell to neuron. *Nature.* Nov 15 2007;450(7168):415-419.

244. Wu MY, Ramel MC, Howell M, Hill CS. SNW1 is a critical regulator of spatial BMP activity, neural plate border formation, and neural crest specification in vertebrate embryos. *PLoS Biol.* 2011;9(2):e1000593.

245. Kishigami S, Mishina Y. BMP signaling and early embryonic patterning. *Cytokine Growth Factor Rev.* Jun 2005;16(3):265-278.

246. Sedwick C. SNW1 orchestrates BMP signaling in early embryonic patterning. *PLoS Biol.* 2011;9(2):e1001018.

247. Kageyama R, Ohtsuka T, Hatakeyama J, Ohsawa R. Roles of bHLH genes in neural stem cell differentiation. *Exp Cell Res.* Jun 10 2005;306(2):343-348.

248. Lee JE, Hollenberg SM, Snider L, Turner DL, Lipnick N, Weintraub H. Conversion of Xenopus ectoderm into neurons by NeuroD, a basic helix-loop-helix protein. *Science.* May 12 1995;268(5212):836-844.

249. Olson JM, Asakura A, Snider L, et al. NeuroD2 is necessary for development and survival of central nervous system neurons. *Dev Biol.* Jun 1 2001;234(1):174-187.

250. Vargha-Khadem F, Watkins KE, Price CJ, et al. Neural basis of an inherited speech and language disorder. *Proc Natl Acad Sci U S A.* Oct 13 1998;95(21):12695-12700.

251. Murdoch BE. The cerebellum and language: historical perspective and review. *Cortex.* Jul-Aug 2010;46(7):858-868.

252. Demirci GT, Atis G, Altunay IK. Waardenburg Syndrome type 1: A case report. *Dermatol Online J.* 2011;17(11):3.

253. Kawabata E, Ohba N, Nakamura A, Izumo S, Osame M. Waardenburg syndrome: a variant with neurological involvement. *Ophthalmic Paediatr Genet.* Nov 1987;8(3):165-170.

254. Kiani R, Gangadharan SK, Miller H. Case Report: Association of Waardenburg Syndrome with Intellectual Disability, Autistic Spectrum Disorder, and Unprovoked Aggressive Outbursts: A New Behavioral Pehnotype? . *Brit J Dev Disabil.* January 2007 2007;53(104):53-62.

255. Pasquale EB. Eph-ephrin bidirectional signaling in physiology and disease. *Cell.* Apr 4 2008;133(1):38-52.

256. Klein R. Bidirectional modulation of synaptic functions by Eph/ephrin signaling. *Nat Neurosci.* Jan 2009;12(1):15-20.

257. Konopka G, Bomar JM, Winden K, et al. Human-specific transcriptional regulation of CNS development genes by FOXP2. *Nature.* Nov 12 2009;462(7270):213-217.

258. Suda S, Iwata K, Shimmura C, et al. Decreased expression of axon-guidance receptors in the anterior cingulate cortex in autism. *Mol Autism.* 2011;2(1):14.

259. Mire E, Mezzera C, Leyva-Diaz E, et al. Spontaneous activity regulates Robo1 transcription to mediate a switch in thalamocortical axon growth. *Nat Neurosci.* Jul 8 2012.

260. Hammond R, Vivancos V, Naeem A, et al. Slit-mediated repulsion is a key regulator of motor axon pathfinding in the hindbrain. *Development.* Oct 2005;132(20):4483-4495.