From the DEPARTMENT OF BIOSCIENCES AND NUTRITION
Karolinska Institutet, Stockholm, Sweden

THE ROLE OF RFX TRANSCRIPTION FACTORS IN NEURONS AND IN THE HUMAN BRAIN

Debora Sugiaman-Trapman

Karolinska Institutet

Stockholm 2018

All previously published papers, figures and tables were reproduced with permission.

Cover illustration by the author, titled “How I Imagine”. Some cartoon figures were taken from the internet (Accessed 7 May 2018), namely the human brain (Shutterstock 69563776), the Hedgehog signaling (Richard Wheeler, Calculated Images), and the Monas Tower of Jakarta (Freepik, Flaticon).

Published by Karolinska Institutet
Printed by E-Print AB 2018

ISBN 978-91-7831-043-2

The role of RFX transcription factors in neurons and in the human brain

THESIS FOR DOCTORAL DEGREE (Ph.D.)

By

Debora Sugiaman-Trapman

Friday the 28th of September, 2018 at 09:30
Lecture Hall 9Q Månen, Alfred Nobels allé 8, Huddinge
Campus Flemingsberg

Principal Supervisor:
Prof. Juha Kere
Karolinska Institutet
Department of Biosciences and Nutrition

Co-supervisor:
Dr. Peter Swoboda
Karolinska Institutet
Department of Biosciences and Nutrition

Opponent:
Prof. Andrew Jarman
University of Edinburgh
Centre for Discovery Brain Sciences

Examination Board:
Prof. Mattias Mannervik
Stockholm University
Department of Molecular Biosciences
The Wenner-Gren Institute

Prof. Maria Eriksson
Karolinska Institutet
Department of Biosciences and Nutrition

Assoc. Prof. Helena Karlström, Docent
Karolinska Institutet
Department of Neurobiology, Care Sciences and Society

TO MY FATHER,

DJONADY SUGIAMAN

ABSTRACT

RFX transcription factors (TFs) are conserved in animals, fungi and some amoebae, but not in algae, plants and protozoan species. The conservation is based on the protein sequence of the DNA binding domain (DBD). The RFX DBD recognizes and binds to a DNA sequence motif called the X-box. In addition to the DBD, most RFX TFs have a Dimerization domain (DIM). The DIM enables RFX TFs to form homo- or heterodimers in detecting the X-box motif, rendering the X-box often described as an imperfect palindromic sequence of two 6-bp half-sites with variable spacers.

So far, RFX TFs are known to regulate gene transcription in cell cycle, DNA repair, immune response, collagen transcription, insulin production, spermatogenesis and hearing. In animals, the most common feature of RFX TFs is their regulation of ciliogenesis and the maintenance of specialized functions of ciliated cells. Cilia are hair-like cell protrusions. They are present in all animals but absent in many species of fungi, amoebae and flowering plants. Based on the inner structure, cilia can be divided into two types, the primary cilia (one cilium per cell) and the motile cilia (either as mono-cilia or multiple-cilia per cell). The primary cilia are less understood despite being present on nearly every cell in the human body.

Humans have eight *RFX* genes (*RFX1-8*) which are expressed in diverse tissues and cell types. This thesis serves to expand knowledge of the RFX TF family in humans and their role in primary cilia and neurons, with interest in human brain development and function. We used databases (Paper I), human cell lines (Papers I and II) and the worm *C. elegans* (Paper III) as our materials for experimentation.

In **Paper I**, we performed an extensive survey of *RFX1-8* expression by transcription start site (TSS) counts from the FANTOM5 database. *RFX1-4* and *RFX7* are prominently expressed in different brain tissues and spinal cord, making them the reference RFX TFs for neurons and the human brain. Furthermore, we predicted the regulation preference of RFX TFs based on co-clustering expression analysis with known RFX target genes. We also analyzed the positioning of the X-box motifs in the human genome and uncovered potential upstream regulators of *RFX* genes.

In **Paper II**, we explored the role of RFX TFs in the context of developmental dyslexia, a developmental disorder of the human brain. The dyslexia candidate genes *DYX1C1*, *DCDC2* and *KIAA0319* have functional X-box motifs in their promoter regions, as shown by luciferase reporter assay of wild-type *versus* mutated X-boxes. By siRNA knockdowns of *RFX1-3*, we showed a complex regulatory mechanism among RFX1-3 in regulating *DYX1C1* and *DCDC2*. Additionally, both DYX1C1 and DCDC2 localize to the primary cilia.

In **Paper III**, we performed microarray analysis of target genes of DAF-19, the sole RFX TF of *C. elegans*, at three developmental stages (3-fold embryo, L1-larvae and adult). At all stages, DAF-19-regulated target genes were significantly enriched in neurons. Using transcriptional GFP reporter constructs, we observed that DAF-19-dependent target genes (both activated and repressed) affected only neurons, both ciliated and non-ciliated.

Altogether, we provided insight into the role of RFX TFs for primary cilia and neurons. We speculate that RFX TFs and primary cilia continue to play a defined role for mature neuron function in the human brain.

LIST OF ORIGINAL PUBLICATIONS

The thesis is based on the following articles or papers, which are referred in the text as Paper I, Paper II and Paper III, respectively.

I. Sugiaman-Trapman D, Vitezic M, Jouhilahti EM, Mathelier A, Lauter G, Misra S, Daub CO, Kere J, Swoboda P. **Characterization of the human RFX transcription factor family by regulatory and target gene analysis**. *BMC Genomics.* 2018 Mar 6;19(1):181.

II. Tammimies K*, Bieder A*, Lauter G*, Sugiaman-Trapman D, Torchet R, Hokkanen ME, Burghoorn J, Castrén E, Kere J, Tapia-Páez I, Swoboda P. **The ciliary dyslexia genes *DYX1C1* and *DCDC2* are regulated by RFX transcription factors through X-box promoter motifs**. *FASEB Journal*. 2016 Oct;30(10):3578-3587.

III. De Stasio EA, Mueller KP, Bauer RJ, Hurlburt AJ, Bice SA, Scholtz SL, Prasad Phirke, Sugiaman-Trapman D, Stinson LA, Olson HB, Vogel SL, Ek-Vazquez Z, Esemen Y, Korzynski J, Wolfe K, Arbuckle BN, Zhang H, Lombard-Knapp G, Piasecki BP, Swoboda P. **An expanded role for the RFX transcription factor DAF-19, with dual functions in ciliated and non-ciliated neurons**. *Genetics.* 2018 Mar;208(3):1083-1097.

*Authors contributed equally

CONTENTS

LIST OF ABBREVIATIONS

AD	activating domain
B	B domain
bp	base pair
C	C domain
CAGE	Cap Analysis of Gene Expression
ChIP-seq	Chromatin immunoprecipitation followed by sequencing
cKO	Conditional knockout
Daf-c	Dauer formation constitutive
DBD	DNA binding domain
DIM	Dimerization domain
FANTOM5	Functional Annotation of Mammalian Genome 5
GFP	Green Fluorescent Protein
Hh / Shh	Hedgehog / Sonic hedgehog
IFT	Intraflagellar transport
qRT-PCR	Quantitative Reverse Transcriptase Polymerase Chain Reaction
RFX	Regulatory Factor binding to the X-box
siRNA	short interference RNA
TF	transcription factor
TFBS	transcription factor binding site
TPM	tags per million
TSS	transcription start site

1 INTRODUCTION

1.1 RFX TRANSCRIPTION FACTORS

Transcription factors regulate gene expression by binding to specific DNA sequences located in gene regulatory regions [1]. The gene regulatory region around a transcription start site (TSS) is called the promoter region. The promoters have typically two distinct regions; the proximal promoter-region that is located from tens to a few hundred bases upstream of the TSS and contains binding sites for sequence specific TFs; and the promoter core-region that is located directly over the TSS and contains the general promoter sequence elements, such as the TATA-box. The proximal promoter-region may contain tens of target sites for multiple different TFs, and the regulatory outcome is determined by the combination of TFs that bind to it.

Some TFs recruit chromatin modifying protein complexes that remodel the local chromatin towards a more open or closed state, and others control the level of gene expression by recruiting components of the general transcriptional machinery to TSSs [2]. A single gene has often multiple promoters and which of them is used to regulate transcription varies between cell types and conditions [3]. In addition to promoters, TFs can also bind to enhancers which are positioned further away from the TSS. Enhancers are located typically tens to hundreds of thousands of bases either up- or downstream of the TSS of their target gene(s). An enhancer stimulates promoter activity via direct physical interaction, i.e. by looping, folding and compacting the intermediate region of the chromatin, so that these two elements reside close to each other [4].

Apart from TFs, there are other factors that regulate gene expressions. The genomic DNA is packed in multiprotein complexes called histones. Epigenetic control of gene expression works by changing the accessibility of DNA for transcription through histone modifications or methylation of DNA. Histones are modified by additions of chemical groups (e.g. methylation, acetylation, or phosphorylation), whereas DNA methylation is a covalent addition of a methyl group to a C nucleotide that is next to a G nucleotide (CpG). TFs and epigenetic marks influence one another in gene regulation and add to the complexity of cell type specificity. In addition, different non-coding RNA transcripts (enhancer RNA, miRNA, long non-coding RNA) have shown activating roles in gene expression [5, 6].

Transcriptional regulatory systems control many biological processes, from the constant cell cycle progression to maintenance of specific cell functions. Scores of diseases are the result of a faulty regulatory system: many TFs are important to suppress cancer and attribute to proper development in humans. In 2009, Vaquerizas *et al*. performed a census of human TFs, where there are about 1,400–1,600 different TFs with several hundreds of them expressed in most cell types [7]. Each of these TFs plays important roles in human development, including the recently discovered TFs that triggered the embryo genome activation [8].

This thesis serves to expand knowledge in the RFX (Regulatory Factor binding to the X-box) TF family in humans and their role in nervous system development and function, using human cell lines and the worm *C. elegans* as model systems.

1.1.1 RFX TFs are evolutionarily conserved

Members of the RFX TF family share and are defined by a conserved winged-helix DNA binding domain (DBD), which facilitates DNA sequence recognition by contacting the minor groove [9]. *RFX* genes have been identified in the genomes of all animals, various fungi and even amoebozoan species, together representing the unikont branch of eukaryotes, while they are excluded from algae, plants and various protozoan branches [10].

The exact evolutionary origin of RFX TFs is unknown. The RFX TF winged-helix DBD is structurally very similar to a protein domain of the origin of replication complex (ORI), which regulates DNA replication and has been conserved from archaea to eukaryotes [11, 12]. RFX and ORI genes may, therefore, have evolved from a common ancestor [10].

Single-celled unikonts tend to possess a single RFX TF, e.g. Sak1 and Crt1 are RFX TFs found in the yeasts *S. pombe* [13] and *S. cerevisiae* [14], respectively. Metazoan genomes encode one to several *RFX* genes. The worm *C. elegans* possesses a single *RFX* gene called *daf-19* [15], the fruit fly *D. melanogaster* has two [16], while the chordate sea squirt *C. intestinales* has four and mammals have eight *RFX* genes ([17], www.ensembl.org ENSG00000196460). Outside of the DBD, RFX TFs may contain other conserved domains, such as the activation domain (AD), the dimerization domain (DIM), and the B and C domains of unknown function [18-20].

1.1.2 RFX TFs recognize and bind X-box motifs

The RFX DBD recognizes a DNA sequence, the X-box motif, to which it binds [9]. The first consensus X-box motif sequence was described by Emery *et al*. [21] as an imperfect palindromic sequence (GTNRCC/N-$N_{0–3}$-RGYAAC), whereby the letters follow the nucleic acid notation of the IUPAC nomenclature [22] **(Figure 1)**.

The X-box motif consists of two 6-bp half-sites separated by a spacer region that is variable in length (0 to 3 bp – preferentially 1 or 2 bp), typically with a strict dependence on the RGYAAC sequence. The RFX binding profiles have been validated by microarrays [23] and SELEX ChIP-seq [1] which show a consistent model of RFX dimeric binding. Most RFX TFs form homo- or heterodimers and bind to the full X-box motif [24]. For RFX5 [25-27] and RFX7 [28] which lack the dimerization (DIM) domain, they bind to a half-site X-box motif **(Figure 1D)**.

Little is known about the dimerization preference among RFX members. RFX4 has been reported to homodimerize and to heterodimerize with RFX2/3 but not with RFX1 [29]. Mouse Rfx6 was shown to dimerize with Rfx3 for islet development but not in cilia formation [30].

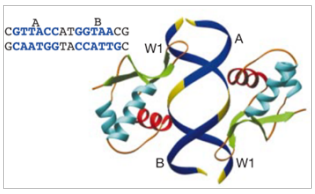
Computational X-box searches have focused on upstream promoter sequences, either upstream of the first exon or of the start codon, as functional X-box motifs are often located close to gene start sites [31]. A genome-wide association study revealed significant overlap of the X-box motifs to risk alleles in type 2 diabetic patients [32].

A

			LEFT HALF-SITE						SPACER 0-3 N	RIGHT HALF-SITE							
A	5	10	6	3	11	13	3	4		11	0	3	30	32	0	11	11
C	5	4	4	2	10	1	25	14		3	0	16	0	0	32	9	3
G	6	8	22	7	2	17	1	7		17	32	3	1	0	0	7	5
T	8	6	0	20	9	1	3	7		1	0	10	1	0	0	4	10
TOTAL	24	28	32	32	32	32	32	32		32	32	32	32	32	32	31	29
CONSENSUS	N	N	G	T	N	R	C	C/n		R	G	Y	A	A	C	N	N

Source: Emery et al. 1996, Mol Cell Biol.

B



Source: Gajiwala et al. 2000, Nature Letters

C

Symbol	Mnemonic	Translation
A		A (adenine)
C		C (cytosine)
G		G (guanine)
T		T (thymine)
U		U (uracil)
R	puRine	A or G (purines)
Y	pYrimidine	C or T/U (pyrimidines)
M	aMino group	A or C
K	Keto group	G or T/U
S	Strong interaction	C or G
W	Weak interaction	A or T/U
H	not G	A, C or T/U
B	not A	C, G or T/U
V	not T/U	A, C or G
D	not C	A, G or T/U
N	aNy	A, C, G or T/U

Source: Johnson 2010, Bioinformatics

D

RFX	JASPAR ID	Sequence Logo
RFX2	MA0600.2	
RFX4	MA0799.1	
RFX5	MA0510.1	

Source: JASPAR 2018, http://jaspar.genereg.net/

Figure 1: (A) The consensus X-box motif for RFX1 from Emery *et al*. [21], showing the full-site motif sequence with preference for the right (3') RGYAAC half-site. (B) Ribbon diagram of the human RFX1 – DNA 2:1 complex, whereby the wings W1of the RFX1 homodimer bind to the imperfect palindromic X-box sequence [9]. (C) The IUPAC nomenclature for nucleic acid notation [22]. (D) Examples of the human full-site X-box (RFX2/4) and the half-site-X-box (RFX5) motif sequence logos from the JASPAR database version 2018.

1.1.3 RFX target gene modules

Target genes of RFX TFs have been identified and studied in a variety of different organisms. Different experimental approaches have been used to demonstrate the physical interaction of RFX TFs with target gene promoters (e.g. ChIP, gel retardation assays, reporter assays following X-box mutagenesis) and to show the dependence of target gene expression on RFX function (e.g. expression analysis after loss of RFX function). In addition, bioinformatic approaches that screen entire genome sequences for conserved X-box promoter motifs have contributed substantially to a growing list of RFX candidate target genes [10, 17, 33-40].

RFX TFs regulate gene transcription in seemingly diverse cellular and developmental processes [18], from basic functions (ciliogenesis, cell cycle, DNA repair) to specialized functions (immune response, collagen transcription, insulin production, spermatogenesis and hearing).

An RFX TF was first described in the context of a human autoimmune disease [25]. RFX5 as part of a nuclear complex exerts DNA binding activity. This RFX complex consists of three subunits; RFX5, RFX associated protein (RFXAP) and RFX anchoring protein (RFXANK), which coordinates the assembly of a multiprotein "enhanceosome" that serves as a docking site for the binding of co-factors and thereby regulates the expression of major histocompatibility class II (MHC-II) genes. Mutations in the human *RFX5* gene [41] result in Bare Lymphocyte Syndrome (BLS) (reviewed in [26]). A similar RFX enhanceosome is also recruited to the promoters of MHC-I genes, where it acts as a transcriptional enhancer [42, 43]. The role of RFX TFs in the immune system is further exemplified in RFX1-2 being regulators of the interleukin-5 receptor α chain (*IL5RA*) [44], *CD11α* (or *ITGAL*) and *CD70* [45].

The ciliogenic role of RFX TFs is by far the most studied. One reason is connected to cilia being present on most cells in the human body, resulting in increased interest in understanding a rare disease class called ciliopathies. *RFX* mutants in different model organisms have been characterized by structural and functional abnormalities in cilia (**Table 1**). The RFX/DAF-19 regulation in *C. elegans'* ciliogenesis [15] led the search for ciliary RFX target genes by genome-wide X-box promoter motif scanning. Indeed, at least half of the candidate genes with functional X-box motifs in *C. elegans* and *D. melanogaster* were confirmed to be ciliary genes [34, 36].

In vertebrates only a few RFX target genes have so far been identified, owing to greater diversity of RFX isoforms and cell types. Most studies in vertebrates point to the RFX2/3 and FOXJ1 TF cooperation in motile cilia biogenesis [20, 46, 47], for instance, in the nervous system ependyma [48] and the lung airway epithelia [49].

Bilateria	Species	RFX TF	Ciliary RFX target gene	Reference
Invertebrates	Worm *C. elegans*	DAF-19	*che-2, osm-1, osm-5, osm-6, bbs-2, xbx-1, xbx-2, etc*	[34, 50]
	Fruit fly *D. melanogaster*	dRFX	*CG6129, CG31036, CG13125, etc* See (i)	[36] [51]
Vertebrates	Zebrafish *D. rerio*	rfx2 rfx2 rfx2	See (ii) See (iii) See (iv)	[52] [53] [54]
	Frog *X. laevis*	Rfx2 Rfx4, Rfx7	*TTC25* See (v)	[55] [56]
	Mouse *M. musculus*	Rfx1, Rfx3 Rfx2 Rfx2 Rfx3 Rfx4	See (vi) *Lrd* *Ccdc65, Ttll1* *Dync2li1, Bbs4* *Ift172*	[57] [54] [58] [46] [59]
	Human *H. sapiens*	RFX1 RFX1, RFX2 RFX1, RFX3 RFX4	*KIF3A* *ALMS1* *MAP1A* *TMEM216, TMEM138*	[60] [61] [62] [63]

Table 1: Examples for the roles RFX TFs play in cilia function in different animal species, whereby either the RFX TF is proven to regulate transcription of ciliary genes or the *RFX* mutant is characterized by dysfunctional cilia.

(i) *Rfx* mutant flies had shorter or missing cilia in embryos and at the dendritic ends of type I neurons.
(ii) *rfx2* morphants exhibited absence of motile pronephric cilia.
(iii) *rfx2* morphants showed reduction in length and number of primary cilia 24 hours post fertilization embryos.
(iv) *rfx2* morphants exhibited shorter Kupffer's vesicle cilia.
(v) *Rfx4* and *Rfx7* morphants exhibited loss of cilia in neural tube at stage 23.
(vi) RFX target genes were highly enriched in the mouse auditory hair cell transcriptome by their X-box signature. Hair cells regeneration and kinocilia structure were disrupted in the *Rfx1/3* cKO mice.

At the same time, RFX TFs regulate the cell cycle exit and entry into the stationary phase (G0) and DNA repair pathways [13, 14] from observations in yeasts as non-ciliated organisms. In addition, human RFX1 and *Drosophila* dRFX2 regulate transcription of the gene *PCNA* [16, 64] with many roles in eukaryotic DNA replication [65]. The crucial role of

human RFX1 in cell proliferation was shown by RFX1 repressing *c-Myc* (*MYC*) [66], silencing *TGFβ2* [67] and regulating *SHP1* (*PTPN6*) [68] in various human cancer cell lines. Embryonic lethality in *Rfx1-/-* mice further suggests that Rfx1 function is vital for survival and cannot be compensated for by other Rfx members [69].

Moreover, RFX TFs were implicated in specific cell fate determinations. Rfx2 directly regulates the expression of genes required for mouse spermatogenesis (*H1t*) and sperm function (*Spag6*) [70-72]. RFX6 is required for the differentiation of pancreatic progenitors and insulin production [30, 73-75] as well as for regulating the gene *PDX1* [76], the TF important for pancreatic and duodenal development. Mutations in the human *RFX6* gene cause Mitchell-Riley Syndrome, characterized by neonatal diabetes and pancreas, intestines and gallbladder deformities [77, 78]. A group of RFX TFs has also been shown to regulate collagen transcription [79] and be essential for hearing [57].

In summary, RFX TFs seem to cause a series of anti-proliferative effects, while they promote the expression of genes necessary to produce or maintain the function of specific, differentiated cell types.

1.1.4 Regulators of *RFX* genes

Knowledge of the upstream regulators of RFX TFs is often revealed in a specific cell fate determination. In the invertebrate *Drosophila*, the TF *atonal* directly regulates *Rfx* for ciliogenesis of the chordotonal neurons [80].

In vertebrate multi-ciliated cells, the transcription factors Multicilin and E2Fs, together with DP1 complex co-factors, work upstream of RFX and FOXJ1 (reviewed in [81]). RFX and FOXJ1 themselves are able to cross-regulate each other [46, 49, 53, 82]. For embryonic nodal ciliogenesis, the homeobox protein Noto works upstream of Rfx3 and Foxj1 [83]. In mammalian differentiation of pancreatic islet cells, Ngn3 works upstream of *Rfx6* [73].

In terms of autoregulation, vertebrate RFX1 seems to be only RFX member that exhibits self-inhibition [84]. The *S. cerevisiae* homologue, Crt1, is also self-inhibited by an upstream phosphorylation cascade [85]. *C. elegans* DAF-19 may also be self-regulated as it harbors the X-box motif in the promoter region [50].

1.2 CILIA AND CILIOPATHIES

Cilia and flagella are cellular organelles which look like hair protrusions. The word flagella is used interchangeably with motile cilia because of their identical structure. The flagellar proteins in single-cell alga *C. reinhardtii* have a high degree of sequence conservation to ciliary proteins in multicellular organisms [86, 87]. However, cilia and flagella are absent in many species of fungi, amoebae and flowering plants [10, 17].

Cilia are generally classified into two categories, motile and primary cilia [88]. Motile cilia can be mono-ciliated (one cilium per cell) or multi-ciliated (multiple cilia per cell). In mammals, motile cilia are present **(i)** in the nodes of developing embryos for determining left-right asymmetry, **(ii)** along epithelial cell surfaces of lung trachea, fallopian tube and brain ventricles to generate fluid movement, and **(iii)** on sperm for cell locomotion.

Primary cilia are less well understood as compared to motile cilia. Unlike motile cilia, primary cilia are always mono-ciliated and primarily associated with sensory functions. The

primary cilium is found on the apical, polarized surface of the majority of cells in human body. However, they may differ in length, shape, molecular contents and functions depending on the organ systems and cell types.

Ciliopathies are a class of genetic disorders attributed to defects in the function of cilia and cilia-associated genes. Because of the widespread distribution of cilia in the human body, ciliopathies have broad pathologies and encompass most human organ systems [89, 90] (**Figure 2**). The current number of established human ciliopathies is 35 with at least 180 genes implicated in the diseases, including the genes involved in the early formation of the cilium, the basic ciliary structure and the ciliary trafficking processes (reviewed in [91]).

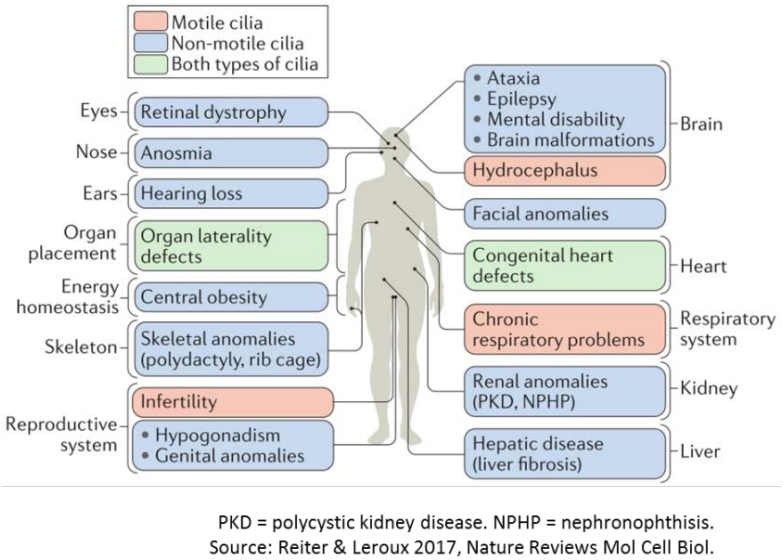


Figure 2: Dysfunctions in cilia cause ciliopathies that encompass most human organ systems [91].

1.2.1 Cilia architecture and intraflagellar transport (IFT)

The basic architecture of the cilium consists of **(i)** the basal body, **(ii)** the transition zone, **(iii)** the axoneme, consisting of: the middle segment and the distal segment (=ciliary tip), **(iv)** the ciliary membrane with enriched G-protein coupled receptors (GPCRs) and specialized channels, and **(v)** the ciliary proteins, which are transported along the axoneme by intraflagellar transport (IFT) particles (reviewed in [88]).

The axoneme is a microtubule-based structure that originates from the basal body and has a certain optimum length. Primary cilia typically display a ring of nine microtubular doublets (9+0 configuration) while motile cilia contain an extra pair of microtubules in the middle of the ring (9+2 configuration). The transport of ciliary proteins along the axoneme is mediated by IFT particles (or trains) in bidirectional movement; that is, from the cytoplasm to the ciliary tip (anterograde IFT), and from the tip back to the cytoplasm (retrograde IFT) (**Figure 3**).

IFT was first described in the flagella of the single-cell alga *C. reinhardtii* [92]. The IFT trains consist of the IFT motor proteins (kinesin and dynein), the IFT complex proteins (IFT complex A, IFT complex B), and the IFT accessory proteins (BBS proteins, or collectively termed as BBSome). The ciliary “cargo” proteins that the IFT particles transport include axonemal components, ciliary membrane proteins and signal transduction proteins. The list of

ciliary and cilia-associated proteins and their functions continues to increase with a growing number of studies on the IFT-cargo protein interactions in different species [93-96].

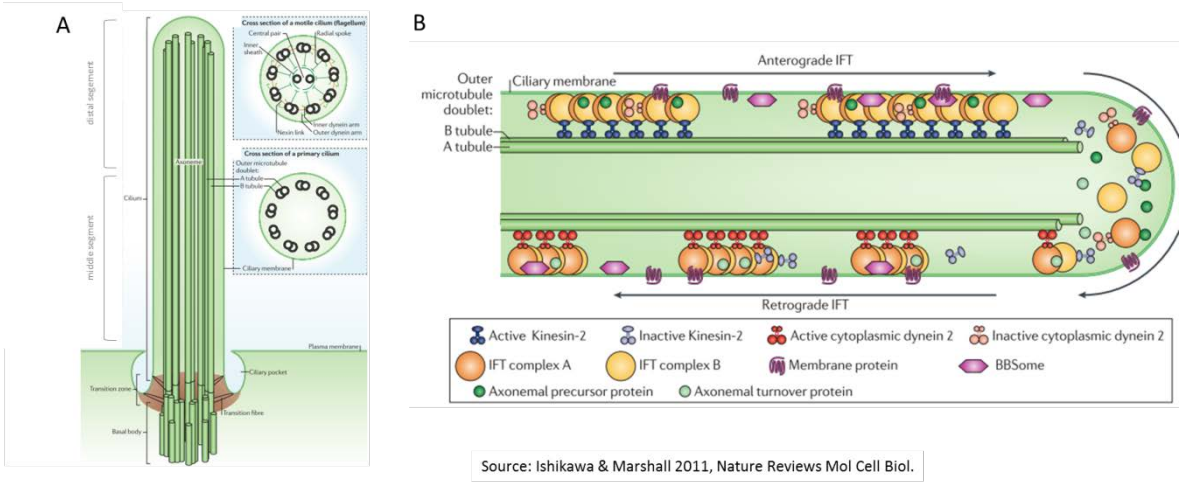


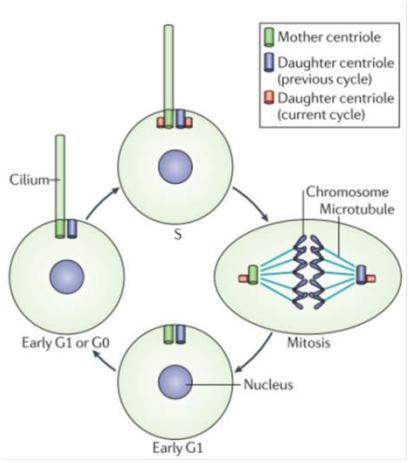
Figure 3: (A) The basic cilia architecture consisting of the basal body, transition zone and the microtubule-based axoneme, with cross sections of the microtubule doublets in motile *versus* primary cilium. (B) The intraflagellar transport (IFT) particles in bidirectional movement (anterograde *versus* retrograde) within the ciliary axoneme [88].

1.2.2 Primary cilia, centrioles and cell cycle progression

The term primary in primary cilium means that the basal body comes from the mother centriole inherited during the previous mitosis, as opposed to secondary cilia which form post-mitotically. The primary cilia assembly and disassembly are, therefore, connected to the cell cycle. In proliferating vertebrate cells, the same microtubules that make up the ciliary axoneme are recycled to make cytoplasmic centrosomes, including the mother and daughter centrioles in cell division.

This dichotomy in the role of centrosomes in dividing cells (as mitotic spindle poles organizers and as templates for ciliogenesis) is regulated by cell cycle-dependent post-translational modifications [97]. In other respects, centrioles are not essential for cell cycle progression in invertebrates but they are necessary for ciliogenesis. The fruit fly *Drosophila* needs centrioles to form mechano- and chemosensory cilia. In another case, the flatworm *S. mediterranea* assembles thousands of cilia from centrioles for gliding motility [98].

The primary cilium extends out during the G1 or G0 phase and appears to be resorbed throughout G1/S phase with centrosome duplication occurring during the S phase of the cell cycle (**Figure 4**). As a result, primary cilia are usually absent throughout mitosis and would easily be missed in rapidly dividing cells. This cycle, however, does not occur in multi-ciliated cells as they are terminally differentiated and no longer undergo cell division [88]. The involvement of cilia in cell cycle progression and in turn, the RFX TFs as known regulators of ciliogenesis, makes them subject to cancer studies as well [99].



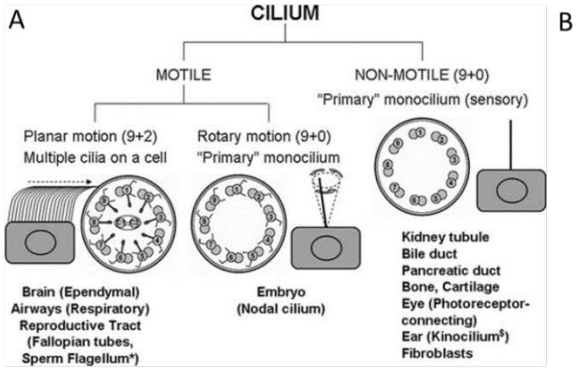
Source: Ishikawa & Marshall 2011, Nature Reviews Mol Cell Biol

Figure 4: Primary cilia assembly and disassembly in vertebrate cell cycle phases. The primary cilium extends out during the G1 or G0 phase and appears to be resorbed throughout the G1/S phase with centrosome duplication occurring during the S phase of the cell cycle [88].

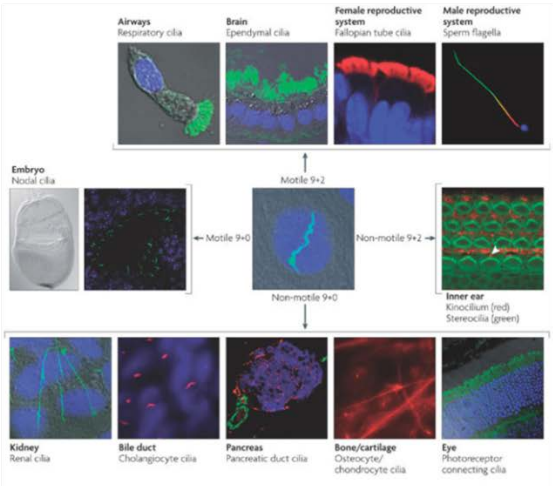
1.2.3 Cilia in vertebrates

In humans, most cells can form primary cilia, while some cells have motile cilia that are involved in sensory and motility functions. The functional specializations of human ciliated cells are manifested in the diverse ciliary axonemal structures present in various cell types [89, 100] (**Figure 5**).

While it is generally believed that hematopoietic cells (particularly erythrocytes) do not have primary cilia, there has been some evidence now for primary cilia in blood cells [101, 102]. So far, the role of primary cilia in the immune system is the least well characterized. Interestingly, in T-cells which lack cilia, IFT proteins were required for the construction of the immune synapse where the centriole and associated Golgi apparatus polarize in response to antigen presentation [103].



Source: Leigh et al. 2009, Genetics In Medicine.



Source: Fliegauf et al. 2007, Nature Reviews Mol Cell Biol.

Figure 5: (A) Diversity of human ciliated cells and their axonemal structures [100]. (B) Different shapes of cilia in human cell types [89].

The importance of the primary cilium in vertebrate Hedgehog (Hh) signaling was first revealed in a mouse embryo study [104]. This was unexpected since Hh signaling in invertebrates does not require cilia. The Hh signaling pathway is required for proper vertebrate embryonic development, such as organogenesis and stem cell maintenance. The binding of the Hh ligand to the Patched receptor (PTCH1) stimulates Smoothened (SMO) to move to the ciliary membrane and the kinesin motor KIF7 to the ciliary tip, allowing GLI-SUFU dissociation and subsequent GLI activation [105]. GLI transcription factors are then transported out of the cilium by the dynein motor and IFT particles and into the nucleus as effectors of Hh signaling pathway [106]. Hh signaling and primary cilia continue to be important for adult tissue homeostasis and repair [107] including the maintenance of adult neural stem cells [108].

In addition to Hh signaling, vertebrate embryonic development is also influenced by unique motile primary cilia which are present at the node (e.g. posterior notochord in mammals, gastrocoel roof plate in *Xenopus*, Kupffer's vesicle in *D. rerio*) that emerges during embryonic gastrulation. These nodal cilia produce nodal flow which induces left-right asymmetry of the organs [109], with RFX2/3 reported to be responsible for nodal cilia differentiation [54, 83, 110]. The mouse homeobox TF Noto has been shown to be the key regulator of posterior notochord ciliogenesis and regulation of *Rfx3* and *Foxj1* [83]. The synergy between RFX and FOXJ1 transcription factors is often associated with motile ciliogenesis [20, 111].

1.2.4 Ciliopathies in the nervous system and the brain

Ciliopathies affect many organ systems including the nervous system [91] **(Figure 2)**. The neural ciliopathic symptoms include brain malformations, cognitive deficits, hydrocephalus, ataxia, anosmia, and hearing loss. The classic ciliopathies with strong neural defects include Alström Syndrome, Bardet-Biedl Syndrome (BBS), and Joubert Syndrome [112, 113]. Both BBS and Alström Syndrome cause obesity, which suggests the role of neuronal primary cilia in the regulation of metabolism and energy homeostasis [114, 115].

More recently, ciliopathies have been linked to complex mental disorders such as developmental dyslexia, autism spectrum disorders, bipolar disorder and schizophrenia [114], mainly based on the connection of their risk genes to cilia [116]. An *in vitro* experiment by Muñoz-Estrada *et al.* [117] tested that in schizophrenia patients' olfactory neuronal precursor cells, primary cilia were diminished and the cytoplasmic microtubules were disorganized. It is thought that neural ciliopathies (i.e. primary cilia dysfunctions of the neural cell types) cause defective neuronal migration and result in aberrant connectivity during the early human brain development [114].

In this thesis, we use developmental dyslexia as the context of the human brain developmental disorder (Paper II). Developmental dyslexia is a complex hereditary reading disorder, characterized by literacy difficulties despite normal vision, hearing, IQ and access to education. The first, and so far, most successful attempts to identify genetic susceptibility loci are based on genetic linkage mapping in large families with dominant inheritance patterns. The Human Gene Nomenclature Committee has named the dyslexia genetic loci DYX1 to DYX9, which are based on families from different countries (and thus different languages) [118]. The dyslexia susceptibility (or candidate) genes which are identified in these loci are continuously expanding [119]. To date, *DYX1C1*, *DCDC2* and *KIAA0319* are the three most replicated and best-studied dyslexia candidate genes.

Dyslexia candidate genes were first connected to cilia when the gene *DCDC2* was found to influence rat neuronal cilia length and the morphology of *C. elegans* ciliated sensory neuron [120]. *DYX1C1*, *DCDC2* and *KIAA0319* genes were later found to be differentially regulated in an *in silico* [121] and an *in vitro* [122] analysis of motile, multi-ciliated cells. In zebrafish, morpholino knockdown of *Dyx1c1* and *Dcdc2* resulted in typical ciliopathy phenotypes such as hydrocephalus, *situs inversus* and kidney cysts [123, 124]. More recently, human *DCDC2* mutations were identified in nephronophthisis (renal-hepatic ciliopathy) [124] and hearing loss patients [125]. Furthermore, DCDC2 protein localized to the ciliary axoneme and to mitotic spindle fibers in a human retinal pigmented epithelial cell line, hTERT-RPE1 [125].

Deficiency in *Dyx1c1* causes phenotypes consistent with motile cilia defects in mouse and zebrafish. Human individuals carrying biallelic *DYX1C1* mutations had classical symptoms of primary ciliary dyskinesia, a multi-systemic motile ciliopathy. Since DYX1C1 was discovered in the ciliary axonemal dynein assembly, DYX1C1 was given a new name, DNAAF4 (Dynein Axonemal Assembly Factor 4) [126]. *KIAA0319*, on the other hand, is not yet observed to localize to cilia and is specifically expressed in the nervous system, unlike *DCDC2* and *DYX1C1*. Deletion of *Kiaa0319* in mice and rats led to defective auditory function and lack of control of axonal regeneration [127-129].

Whether dyslexia and other complex mental disorders are neural ciliopathies remains an open question. The experimental connection between dyslexia candidate genes and cilia may be strong, but the connection between cilia and dyslexia remains debatable. The loss-of-function mutations in ciliary dyslexia candidate genes are often found in ciliopathy patients without dyslexia [126]. Finding the molecular functions of dyslexia candidate genes constitutes only one piece of the puzzle in understanding the cellular etiology of developmental dyslexia.

1.3 RFX TFS AND THE VERTEBRATE NERVOUS SYSTEM

Finding the homologues of human *RFX* genes in other animal species is often based on the fundamental shared feature, the winged helix DNA binding domain (DBD) [10, 17, 19]. Human *RFX1-7* have corresponding, and functionally validated homologues in other vertebrates [130] with the exception of two additional putative *RFX* genes in fish, *Rfx8-9* [17]. According to HomoloGene (NCBI) which analyses the whole protein sequence, human *RFX3* is most conserved down, even in insect and worm, whereas *RFX5* is only present in mammals and *RFX8* is only present in mammals and birds (**Table 2**).

H. sapiens	Protein product conservation	Animals with the human *RFX* homologue
RFX1	Euteleostomi (bony vertebrates)	monkey, dog, mouse, rat, bird, frog, fish
RFX2	Euteleostomi (bony vertebrates)	monkey, dog, mouse, rat, bird, frog, fish
RFX3	Bilateria (bilateral symmetry animals)	monkey, dog, mouse, rat, bird, frog, fish, insect, worm
RFX4	Euteleostomi (bony vertebrates)	monkey, dog, mouse, rat, bird, frog, fish
RFX5	Boreoeutheria (placental mammals)	monkey, dog, mouse, rat
RFX6	Euteleostomi (bony vertebrates)	monkey, dog, mouse, rat, bird, frog, fish
RFX7	Euteleostomi (bony vertebrates)	monkey, dog, mouse, rat, bird, frog, fish
RFX8	Amniota (amniotic egg vertebrates)	monkey, dog, mouse, rat, bird

Table 2: The homologues of human *RFX1-8* genes according to HomoloGene (NCBI): https://www.ncbi.nlm.nih.gov/homologene, accessed in February 2018.

Studies in different vertebrate animals suggest that RFX1-4 and -RFX7 are the relevant RFX TFs in the vertebrate nervous system. RFX1 was demonstrated to be important in neuronal function and specification [60, 62, 131, 132]. Moreover, being identified as the homologue of yeast Crt1 and Sak1 [84, 133], RFX1 seems to have a fundamental role in cell cycle regulation. RFX2 is broadly required for ciliogenesis and development [55] while RFX3 plays a more specific role in ependymal cilia function for proper cerebrospinal fluid flow [48].

RFX4 is by far the most specific RFX TF in nervous system development [82, 134, 135]. According to functional investigation of RFX7 in *Xenopus* [56], both RFX4 and RFX7 contribute to ciliogenesis and the early development of the brain and spinal cord.

Interestingly, both RFX2 and RFX4 are highly expressed in brain and testis [19] and could form heterodimers [29]. A rat study suggested that Rfx4 could enhance Rfx2 binding to the X-box but it was itself weakly bound to the X-box of the testis-specific linker histone *H1t* [71, 136]. This indicates that RFX2 has a more prominent function in the male reproductive system [136, 137].

1.3.1 Expression of *Rfx* in the mouse, frog and fish nervous systems

Insight into the role of the human RFX TFs in the vertebrate nervous system has been derived mainly from studies in mice (**Table 3**), followed by the frog *Xenopus* and zebrafish *D. rerio*. In rats, Rfx studies have focused on Rfx2/4 and their role in spermatogenesis in isolated rat testis cells [71, 136]. Based on the mouse SAGE libraries data, human *RFX1-7* genes (except *RFX6*) were predicted to be ubiquitously expressed, with relatively high expression in the brain. *RFX6* was strikingly absent in almost all tissues except the pancreas [19].

Mouse Rfx1 proteins were expressed in the cerebral cortex, hippocampus and olfactory bulb. They were mainly found in the nuclei of neurons and microglial cells but not in astrocytes [131, 138]. Both *Rfx2* and *Rfx3* were detected in the mouse node posterior notochord responsible for left-right organ asymmetry [54]. In addition, *Rfx3* was also expressed in the multi-ciliated ependymal cells lining the cerebral ventricles after birth [46, 48, 139]. Interestingly, *Rfx2* deficient mice did not have severe defects in neural tube closure or in organ *situs*, but males were infertile [58, 130, 140].

Mouse *Rfx4* expression was restricted to neural tissue, including the embryonic forebrain, neural tube and spinal cord [59]. From a single-cell RNA-sequencing profiling of the ventral midbrain cells in mouse and human embryos, mouse *Rfx4* was detected in the neuro-progenitor cell, radial glia-like cell and ependymal cell types, whereas human *RFX4* was detected only in a subset of the radial glia-like cell types [141]. In another study, the mRNA levels of the mouse *Rfx1-7* except *Rfx6* were detected by qRT-PCR in the embryonic forebrain, and that *Rfx1*, *Rfx3*, *Rfx4* and *Rfx7* were strongly expressed (with *Rfx4* being the highest) in the neural stem cells taken from the cerebral cortices [142].

In the frog *Xenopus*, *Rfx2* was expressed broadly in ciliated tissues of the neural tube, the node gastrocoel roof plate, the epidermal multi-ciliated cells, auditory vesicles, and kidneys [55]. In contrast, *Rfx4* and *Rfx7* genes were expressed in the nervous system, whereby *Rfx4* was more specifically expressed in the brain and spinal cord, and *Rfx7* was more broadly detected in the brain, spinal cord, eyes and auditory vesicles. Neither *Rfx4* nor *Rfx7* were expressed in the node. Morpholino knockdown of both *Rfx4* and *Rfx7* resulted in failure in neural tube closure caused by the loss of cilia [56].

In zebrafish *D. rerio*, *rfx2* was expressed in the node Kupffer's vesicle. Morpholino knockdown of *rfx2* resulted in primary cilia defects [53] and *situs inversus* [54]. In contrast, *rfx4* was expressed in the developing neural tube but not in the node [135].

1.3.2 Phenotypes of *Rfx* deficient mice

The various phenotypes observed in *Rfx* mouse mutants reflect the important roles of the RFX TFs in vertebrate development. Phenotypes could be mild, but mostly severe and embryonic lethal when both copies of the *RFX* alleles were absent (**Table 3**).

Mouse Rfx	***Rfx* mutation**	**Phenotype**	**Reference**
Rfx1	*Rfx1*$^{+/-}$ (het)	Normal and fertile.	[69]
	Rfx1$^{-/-}$ (hom)	All embryos died before E2.5	
Rfx2	*Rfx2*$^{gt/gt}$ (hom) gt = gene trap insertion	Normal, no *situs inversus.* Females were fertile while males were sterile.	[140]
	Rfx2$^{-/-}$ (hom)	Approximately 25% showed severe growth retardation with age and died before 2 months of age. Females were fertile while males were sterile.	[58]
Rfx3	*Rfx3*$^{-/-}$ (hom)	Approximately half of the embryos died before E12 and two-thirds died at birth. The pups were smaller and approximately 6% had *situs inversus.*	[110]
		In addition to [110], nearly all pups that survived had hydrocephalus with variable severity.	[48]
		Most died pre- and post-natal from defects in left-right body patterning. Approximately 25% survived until E19. In addition, the reduction in insulin associated with defect in pancreatic islet cell development was highly significant among the viable adult -/- mice.	[143]
		At E18.5, approximately half exhibited partial to complete agenesis of the brain corpus callosum, with few or no callosal axons crossing the midline linking the two cerebral hemispheres.	[144]
		At E18.5, axon guidance defects were observed resulting in abnormality of the brain patterning.	[145]
Rfx4	*Rfx4_v3*$^{-/-}$ (hom) v3 = brain specific (v1-2 = testis specific)	Significantly smaller litter size and substantial loss of the -/- pups compared to +/-. All of the -/- pups born died within 1 hour of birth. The brains and spinal cord at E16.5 and at birth were grossly dysmorphic, while the anatomy of the rest of the body appeared normal.	[134]
	Rfx4_v3$^{+/-}$ (het) *Rfx4_v3*$^{-/-}$ (hom)	Deformed subcommissural organ, resulting in congenital hydrocephalus. Fatal failure of dorsal midline brain structure formation. **Additional note**: Microarray validation of differentially expressed genes (+/- versus -/-) confirmed Rfx4 regulation of genes involved in brain morphogenesis by the X-box promoter motif binding.	[146]
	Rfx4^{L298P} (hom)	All died shortly after birth, with approximately two-	[59]

	Point mutation by ENU-induced mutagenesis	thirds inclining toward the severe defects that closely resembled -/-, such as the distinct defects in the patterning of the ventral spinal cord and telencephalon due to abnormal Sonic hedgehog (Shh) signaling.	
	Rfx4$^{+/-}$ (het) *Rfx4*$^{-/-}$ (hom) cKO = conditional knockout	Almost half developed hydrocephalus within 8 weeks of age for both male and female, due to underdeveloped subcommissural organ and patchy motile cilia. Many embryos did not survive to E14.5. The developing central nervous system was affected as seen by a single ventricle in the forebrain, and severe dorsoventral patterning in the telencephalon and midbrain at E12.5. The phenotypes resembled human holoprosencephaly. No anatomical abnormalities were noted outside the brain in either +/- or -/-.	[82]
Rfx5	*Rfx5*$^{+/-}$ (het) *Rfx5*$^{-/-}$ (hom)	Indistinguishable from wild type mice. Mice were healthy and reproduce normally as long as they were kept pathogen-free. They retained expression of MHC-II in thymic medulla, mature dendritic cells, and activated B cells, but not in thymic cortex, resting B cells and resident or IFNγ-activated macrophages. They also failed in positive selection of CD4^{+} T cells.	[147]
Rfx6	*Rfx6*$^{eGFPcre/eGFPcre}$(hom) enhanced GFP (eGFP) – cre fusion	Mice died within P2 for being unable to feed normally due to gross bowel obstruction. Some had reduced pancreas size and ultra-analysis revealed the lack of all endocrine cells except pancreatic-polypeptide-producing cells.	[30]
	Rfx6$^{\Delta Endo}$ (hom) cKO endocrine lineage specific	Pups were diabetic and died by P3. Newborns almost entirely lacked insulin-positive cells.	[74]

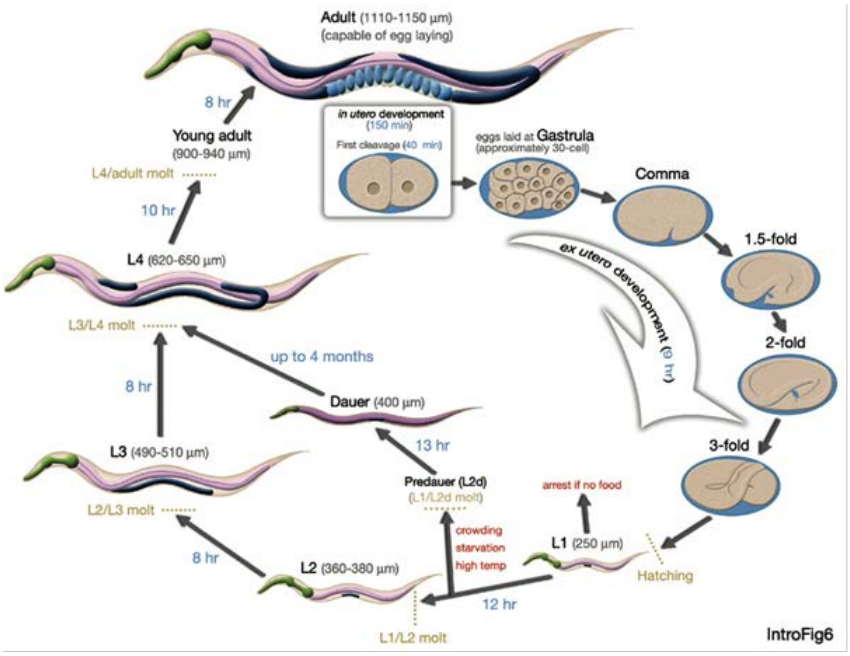
Table 3: Phenotypes of mice deficient in *Rfx1-6*. All the phenotypes described in this table agree with the corresponding Rfx studies in rats, *Xenopus*, and *D. rerio*. Rfx2 was examined in isolated rat testis cells [71, 136] and confirmed the mice *Rfx2* mutant phenotype of male sterility. To date, no *Rfx7-8* mutant mice have been generated or characterized. Rfx7 has been studied by morphant knockdown of *Xenopus* embryos [56].

het = heterozygous, hom = homozygous, E = embryonic day, P = postnatal day (E0-E19 precedes P0).

Multiple *Rfx* genes cKO studies in mice had been done by Elkon and colleagues [57] where they observed hearing loss in the *Rfx1/3* double cKO mice. Interestingly, the hearing ability was unaffected in the single *Rfx1* cKO and *Rfx3* cKO mice.

1.4 DAF-19, THE SOLE RFX TF IN *C. ELEGANS*

The only *RFX* gene in *C. elegans* is named *daf-19* (daf = dauer formation) because mutations were first identified in screens for constitutive dauer larva formation. Wild-type *C. elegans* enter an arrested larval stage called dauer under duress (**Figure 6**). In contrast, *daf-19* mutants formed many dauers even in optimal growth conditions [148]. The *daf-19* gene was later characterized to encode an RFX TF of which the dysfunction affected sensory cilium development and function in both the sexes, hermaphrodites and males [15, 149].



Source: Wormatlas, http://www.wormatlas.org

Figure 6: The 3-day life cycle of the *C. elegans* hermaphrodite at 22^{o}C, where 0 min is fertilization. After the embryo stages, there are four larval stages (L1-L4) before it matures to an adult. An adult worm lays about 300 progeny in 3-5 days and lives for another 2-3 weeks. The animal enters an arrested, alternative L2 stage, called the dauer stage when the environmental conditions are harsh.

There are five known *daf-19* protein-coding transcript isoforms, of which the annotations are different between Wormbase (https://wormbase.org) and the literature. The five transcript isoforms consist of: *daf-19a* and *daf-19b*, which are expressed broadly across development [150]; *daf-19c*, which is the ciliogenic isoform expressed only in early development (between the 3-fold embryo and L1 stage) and has two alternative exon start sites [150, 151]; and *daf-19m*, which is required for male mating behaviour [149, 152] (**Figure 7**).

A Northern blot analysis of embryonic total RNA revealed *daf-19a* to be the most abundant, followed by *daf-19c* and then *daf-19b* as the least present [151]. The isoform *daf-19m* is present in both sexes but the function was primarily analyzed in the male worms [149].

The isoforms of *daf-19* are generated from distinct promoters, enhancer elements and post-transcriptional silencing factors for spatial and temporal regulation. SL1 *trans*-splice leaders were observed in Exons 1 and 4 of *daf-19* [153]. A post-transcriptional silencing agent piRNA (piwi-interacting RNA) gene, *21ur-13428*, is present around exon 5 of *daf-19*. The CEM enhancer and HOB and RnB elements contribute to functional *daf-19m* isoform [149].

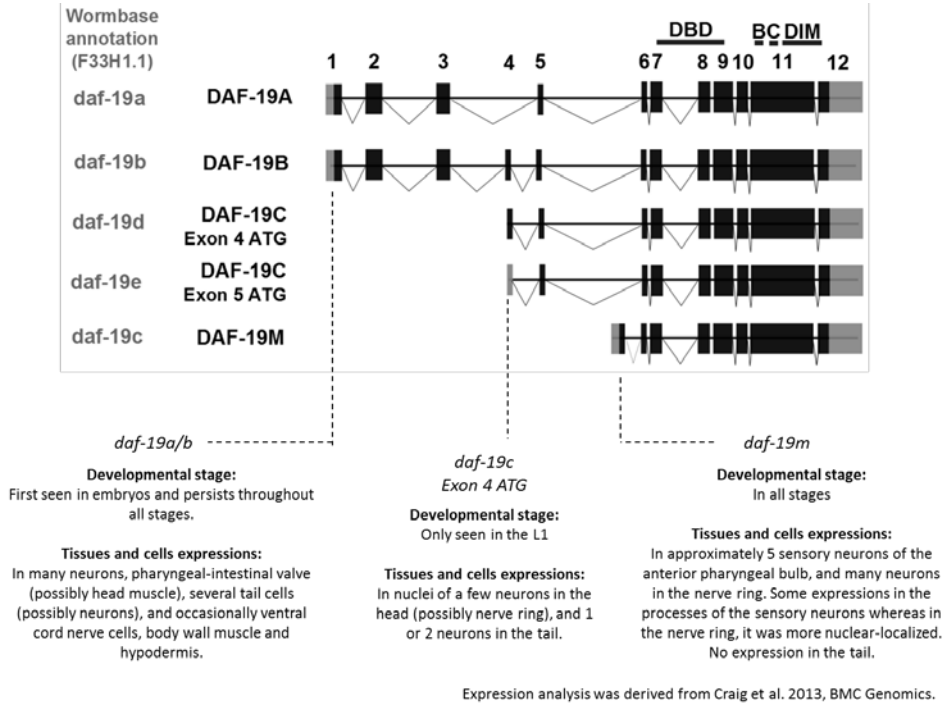


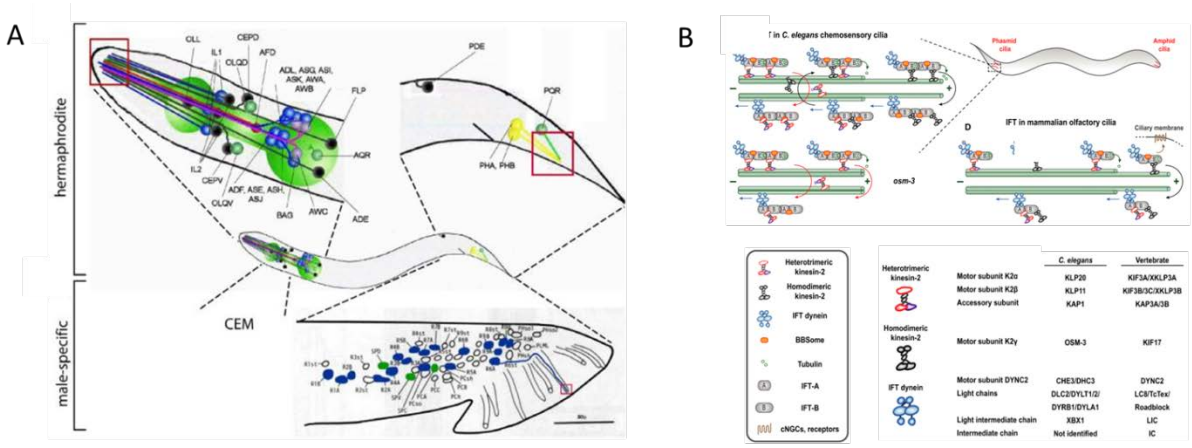
Figure 7: The five known DAF-19 isoforms in *C. elegans*. All the isoforms have the conserved protein domains DBD, B, C and DIM. The numbers on top of the boxes refer to the exon numbers, whereby the black region is coding region and the grey region is non-coding region. *daf-19* isoform-specific transcript expression patterns at all stages of wild-type *C. elegans* hermaphrodites were obtained from Craig *et al*. [150].

1.4.1 Cilia in the invertebrate *C. elegans*

In *C. elegans*, cilia are only present at the dendritic ends of ciliated sensory neurons which function as both vestibulary (mechanosensory) and olfactory (chemosensory) organs. Hermaphrodites have 60 ciliated sensory neurons, while males have an additional 48 ciliated neurons out of the 87 male-specific neurons. A reference table of the ciliated neurons can be found in Wormatlas (http://www.wormatlas.org/postembryonicneurons.htm).

The majority of the ciliated neurons are located in the head of the worm (=amphid) and act primarily as a nose. In the tail part of the worm (=phasmid), there are only two types of ciliated sensory neurons in the hermaphrodite responsible for chemorepulsion, while most of the male-specific ciliated sensory neurons are concentrated in the specialized male tail sexual organ (reviewed in [154]). *C. elegans* that lack functional DAF-19, the sole RFX TF of *C. elegans*, were found to be missing all ciliary structures and therefore exhibited severe sensory defects [15].

All the *C. elegans* cilia are classified as primary non-motile cilia with a 9+0 axonemal configuration. Anterograde IFT in *C. elegans* is driven by two types of kinesin-2 motors, heterotrimeric kinesin-II and homodimeric OSM-3, while retrograde IFT is driven by dyneins which recycle the kinesin-2 motors. This is similar to vertebrate olfactory cilia whereby the heterotrimeric kinesin-2 builds the axoneme, the homodimeric kinesin-2 (KIF17, homologous to *C. elegans* OSM-3) delivers signaling molecules such as cyclic nucleotide-gated channels (cNGCs) and dopamine receptors to the ciliary membrane, and the IFT dyneins drive transport back toward the basal body. However, the distal segment is built only by OSM-3 in *C. elegans* [96] (**Figure 8**).



Source: Bae & Barr 2008, Frontiers in bioscience.

Source: Prevo et al. 2017, The FEBS Journal.

Figure 8: (A) The positions of ciliated sensory neurons in the two sexes of *C. elegans*, where each neuron name consists of typically three uppercase letters, indicating the class. The locations of ciliary endings are indicated in red boxes [154]. (B) Comparison of the intraflagellar transport (IFT) in *C. elegans* and vertebrate olfactory cilia [96].

1.4.2 DAF-19 in *C. elegans* neurons

Out of 959 somatic cells in a *C. elegans* hermaphrodite, 302 are neurons and 56 are neuronal support cells. All the neurons can be sub-divided into sensory, motor and interneuron classes. All the motor and interneurons in *C. elegans* are non-ciliated. The sensory neurons are mostly ciliated except 2 head sensory neurons and 8 touch sensory neurons which have specialized microtubules instead of cilia. The neurons of *C. elegans* form several ganglia in the head and tail and into a spinal cord-like ventral nerve cord. The majority of the neurons are located in the head, surrounding the pharynx (the feeding organ) and forming a nerve ring akin to the primitive brain for the animal.

Serial electron microscopy has generated detailed information about the anatomy, position and connectivity of each neuron [155] and a complete list of the neurons, their lineage and descriptions can be found in Wormatlas (http://www.wormatlas.org/) and Wormbook (http://www.wormbook.org/). Compared to the vertebrate (human) brain with approximately 80 billion neurons and 10 times more glial cells, the simplicity of the *C. elegans* nervous system makes it a useful system to study the role of RFX TFs in neurons and its impact on *in vivo* whole-organism developmental (dauer formation) and behavioral (foraging, locomotion) output.

DAF-19 was first discovered to be the regulator of ciliogenesis of 60 ciliated sensory neurons [15]. Thereafter, the majority of DAF-19 target genes were found to be either X-box motif or DAF-19-dependent [33-35]. Subsequent genome-wide validations of DAF-19 target genes – by microarray [38] and enhanced yeast one-hybrid (eY1H) assays [50] – used animals at the 3-fold stage embryo when ciliogenesis occurs [156].

The role of DAF-19 expands outside the ciliated sensory neurons as different DAF-19 hermaphrodite isoforms have been characterized. The larger DAF-19A/B isoforms maintain the synaptic protein homeostasis in roughly 200 non-ciliated neurons of different neuronal classes [151]. Unlike the X-box motif transcriptional regulation in ciliated neurons, the mode of DAF-19 synaptic maintenance regulation in non-ciliated neurons is still unclear. Strikingly, synaptic defects in *daf-19* mutants become stronger as the animals grow older,

displaying parallels to the synaptic decline observed in human neurodegenerative disorders [151].

The expression patterns of *daf-19* itself are mainly neuronal, with occasional expressions in the body wall muscle, hypodermis and intestinal cells [15, 150] (**Figure 7**). In parts the significance of these non-neuronal expressions is still unclear, with the exception of the intestinal expression pattern of *daf-19* which fits with DAF-19 regulation of innate immune response upon ingestion of pathogenic bacteria [157].

1.4.3 Phenotypes of *daf-19* mutants

The *m86* allele [15] is the main reference allele for *daf-19* mutant. This null allele exerts a nonsense amino acid substitution in the shared coding exon 7 upstream of the DBD (**Figure 7**), thereby affecting all the DAF-19 protein isoforms.

The *daf-19(m86)* mutant has a temperature-sensitive, severe constitutive dauer phenotype (Daf-c), and chemosensory defects, such as dye-filling defective (Dyf) and osmotic-avoidance defective (Osm) [15]. The chemosensory deficiencies are caused by the absence of ciliated endings in the ciliated sensory neurons. The exact cause of the Daf-c phenotype, however, is still unknown. In gene expression studies at multiple life stages, a *daf-12(sa204)* allele background is required to fully suppress the Daf-c phenotype conferred by *daf-19(m86)* while the cilia remain completely absent [158].

The *daf-19(m86)* mutant does not have obvious impaired movement and can move as fast as wild type. However, they have abnormal dwelling/roaming behavior and are resistant to aldicarb and levamisole [151]. In the dwelling/roaming assay, a single wild-type worm put on a fresh plate seeded with bacteria will quickly cover the entire bacterial lawn with tracks. In contrast, the *daf-19(m86)* mutant covers only a small area and spends more time feeding locally (= increased dwelling behavior).

The paralysis aldicarb/levamisole assay is used to test synaptic transmission in *C. elegans* [159]. Aldicarb and levamisole are acetylcholine esterase inhibitor and acetylcholine receptor agonist, respectively. In the presence of these compounds, acetylcholine accumulates and causes persistent muscle contraction and eventual paralysis in wild-type animals. The *daf-19(m86)* mutant with impaired synaptic transmission is thus more resistant to the paralyzing effect.

In addition to neuronal impairments, *daf-19(m86)* also confers deficient innate immunity phenotype as shown by lower survival rates in *P. aeruginosa* PA14 pathogenesis assays and reduced immune gene markers in the intestine compared to wild-type [157]. Similarly, there is an increased susceptibility of *daf-19* RNAi worms to pathogen *S. flexneri* infection [160]. Although the exact mechanism is unclear, the mode of regulation is reminiscent of the RFX5-cofactors partnership for human MHC II gene expression [157].

Studies on *daf-19* isoform specific mutants have been driven by the discovery of the *daf-19m* isoform (**Figure 7)**. Two mutants, *daf-19(n4132)* and *daf-19(sm129),* disrupt only the *daf-19m* isoform and affect the male specific and core IL2 ciliated sensory neurons [149, 152]. Mutant males were inefficient in mating behavior in response to hermaphrodite contact (Lov = location-of-vulva defective) whereas mutant hermaphrodites seemed normal. Interestingly, the ciliated sensory neurons are intact, showing that the *daf-19m* isoform does not interfere with ciliogenesis, but ensures proper specialized function of the cilia.

2 AIMS

The overall aim of this thesis is to investigate the role of RFX transcription factors (TFs) during development, with a focus on neurons and the human brain function, and to understand the molecular etiology of the brain disorders that arise in the absence of functional RFX TFs.

The specific aim of Paper I is to characterize the eight human RFX TFs, by charting the expression patterns of all their protein-coding transcript isoforms from the FANTOM5 database, inferring correlations and X-box motif positioning from their known target genes, and investigating their upstream regulators.

The specific aim of Paper II is to explore the hypothesis that developmental dyslexia as a human brain disorder may be associated to ciliogenic RFX TFs by testing the regulation by RFX TFs of three dyslexia candidate genes through their X-box promoter motifs.

The specific aim of Paper III is to substantiate the role of DAF-19, the sole RFX TF in the invertebrate animal model *C. elegans*, and its contributions in the regulation of target genes which are expressed in various types of neurons throughout development and adulthood.

3 RESULTS

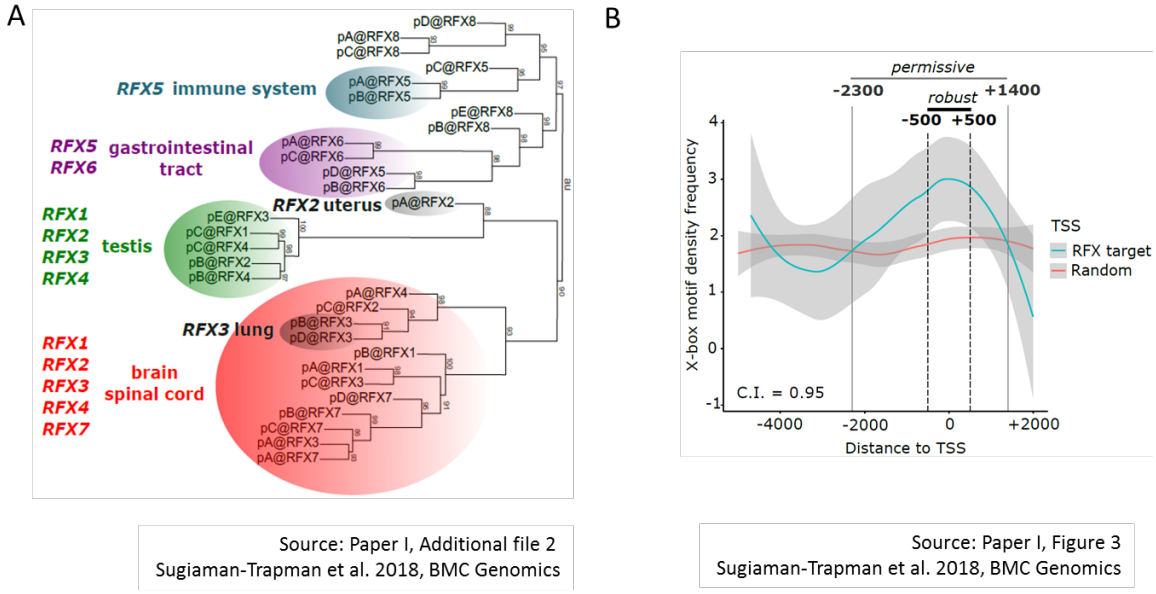
3.1 PAPER I

A survey of the human *RFX* genes, their target genes and upstream regulators.

Human *RFX1-8* are expressed in diverse human tissues and cell types. Our extensive survey of *RFX* expression by transcription start site (TSS) counts based on data from the FANTOM5 database allowed us to rank and cluster the human *RFX* genes in terms of their expression level in various tissues, primary cells and cell lines. We were able to connect most of the *RFX* TSS locations to protein-coding transcripts in Ensembl and experimentally validated novel TSS locations by RT-PCR from either brain or testis RNA. This included validation of the previously undescribed *RFX8*.

Further examination of the protein products based on the position of the start codon and functional domains of the *RFX* transcripts indicate that the majority of the *RFX* transcripts (of the same gene) encode the same protein. The exceptions are: **(i)** testis-specific RFX1 novel protein with a shorter N-terminal, **(ii)** three different RFX4 proteins, and **(iii)** three different RFX8 proteins. We also corroborated that *RFX4* and *RFX8* encode isoforms with and without the DBD, although their significance is not yet understood.

We performed hierarchical clustering of the human *RFX1-8* TSS expression based on 135 tissues present in FANTOM5 and found four major and two minor tissue clusters (**Figure 9A**). The four major tissue clusters were *RFX1-4* and *-RFX7* in the central nervous system (brain and spinal cord), *RFX1-4* in the testis, *RFX5-6* in the gastrointestinal tract, and *RFX5* in the immune system. The two minor tissue clusters were *RFX2* in the uterus and *RFX3* in the lung. *RFX8* was not assigned to any cluster because all *RFX8* TSSs were lowly expressed. In addition to the minimal *RFX8* expression, functional characterization of *RFX4* and *RFX6* would present a challenge because of their tissue specificity. Very few human cell lines would express the minimum detectable level of *RFX4*, *RFX6* and *RFX8*.



Source: Paper I, Additional file 2
Sugiaman-Trapman et al. 2018, BMC Genomics

Source: Paper I, Figure 3
Sugiaman-Trapman et al. 2018, BMC Genomics

Figure 9: (A) Hierarchical clustering based on the tissue expression of human *RFX1-8*. (B) The X-box motif positioning within human RFX target gene promoters with respect to the TSS.

RFX TFs typically bind to the X-box motif in the promoter region of their target genes and different members of RFX TFs may interact with one another in their target gene regulation. In order to predict the groups of RFX TFs working in different tissue clusters, we extracted the TSS expression data from FANTOM5 for the RFX target genes and performed hierarchical co-clustering with the *RFX* TSSs.

This pointed to conclusions on the mode of RFX TF regulation of their target genes; **(i)** for a given target gene, RFX TFs can act as activators or as inhibitors, **(ii)** there is a strong difference between the immune system and the nervous system group of RFX TFs, and **(iii)** RFX2 is an outlier within the nervous system group and clusters more closely to the reproductive system. In addition, we uncovered the X-box motif positioning of human RFX target genes with the most likely (robust) position of -500 to +500 bp, and otherwise a permissive window of -2300 to +1400 bp, with respect to the TSS (**Figure 9B**).

Without focusing on any specific cell differentiation context, we wanted to find all the potential upstream regulators of *RFX* genes using an extensive computational TFBS sequence analyses of the human *RFX1-8* promoter and enhancer regions. We found 19 over-represented TFBS profiles from the JASPAR database which includes TFs involved in neural development (SP2, ESR1, Creb5, SOX21), neurite outgrowth (KLF16, EGR3) and cell cycle control (E2F4). We validated SP2 as a repressor of *RFX7* and ESR1 as an inhibitor of *RFX2*, *-3*, *-5* and *-7* in the human MCF7 breast cancer cell line.

3.2 PAPER II

RFX TFs regulate dyslexia candidate genes through X-box promoter motifs.

Developmental dyslexia is considered a developmental disorder of the human brain characterized by reading deficits despite normal IQ, normal senses and learning opportunity. The candidate genes of dyslexia were obtained primarily from genetic linkage family studies. The three most replicated genes; *DYX1C1* (or *DNAAF4*), *DCDC2* and *KIAA0319* were associated to cilia and neuronal function. We were interested in testing if these genes are regulated by RFX TFs.

We looked for candidate X-box motifs in the promoter regions (-3000 bp with respect to the TSS) of *DYX1C1*, *DCDC2* and *KIAA0319*, starting with the principal transcript isoforms. We developed a stringent computational approach using seven different X-box consensus motifs; five from published work and two from our own studies (**Table 4**). We found a total of seven candidate X-boxes from the secondary transcript isoform of *DYX1C1* and the principal transcript isoforms of *DCDC2* and *KIAA0319.*

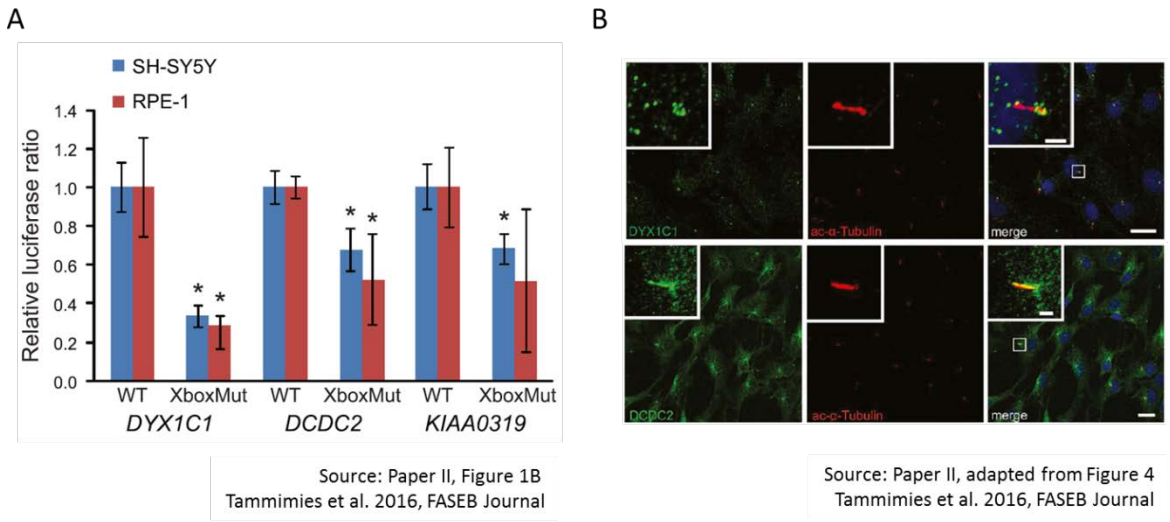
No.	X-box motif consensus (5' → 3')			Organism	Reference
1	GTNRCC	(0–3N)	RGYAAC	*H. sapiens*	[66]
2	GTNRC(C/N)	(0–3N)	RGYAAC	*H. sapiens*	[21, 132]
3	RYYNYY	WW	RRNRAC	*C. elegans*	[34]
4	RTHNYY	WT	RRNRAC	*C. elegans*	[34]
5	GTHNYY	AT	RRNAAC	*C. elegans*	[34]
6	GTTNCC	NN(0-1N)	GGHVAC	*H. sapiens*	This study
7	RBNNYY	NH	RGHAAC	*H. sapiens*	This study

Table 4: X-box motif sequences used for scanning the promoter regions with the EMBOSS fuzznuc tool.

The seven candidate X-boxes of *DYX1C1*, *DCDC2* and *KIAA0319* were further short-listed by their sequence conservation with other vertebrate species. Thus, we selected the X-box motifs of *DYX1C1*, *DCDC2* and *KIAA0319* positioned at -19 bp, -2110 bp and -70 bp, respectively, for functional validation by luciferase reporter assays. Indeed, these X-box motifs were functional as demonstrated by lower luciferase promoter activity in the case of a mutated X-box *versus* a wild-type X-box (**Figure 10A**). Furthermore, the positioning of the functional X-box motifs is concordant with our results in Paper I.

We used two human cell lines to study RFX TFs regulation of *DYX1C1*, *DCDC2* and *KIAA0319*; the retinal pigmented epithelial cells hTERT-RPE1 and the neuroblastoma cells SH-SY5Y. Endogenous *RFX1-7* expression was tested by qRT-PCR. All except *RFX4* and *-6* were undetectable, which confirmed our results from Paper I. Thereafter, we focused on RFX1-3 in the context of cilia induction in hTERT-RPE1 by serum starvation after 24 hours.

To test the regulation by RFX TFs of the dyslexia candidate genes, we performed single, double and triple knockdowns of *RFX1-3* in ciliated hTERT-RPE1 cells and measured the expression levels of *DYX1C1*, *DCDC2*, *KIAA0319*. The results suggested that **(i)** RFX1 acts as a repressor for *DYX1C1* and *DCDC2*, **(ii)** RFX2 and RFX3 act as an activator complex for *DCDC2*, and **(iii)** *KIAA0319* is not significantly regulated by RFX1-3. Furthermore, we corroborated the ciliary role of the RFX TFs by the endogenous protein localization of DYX1C1 and DCDC2 to primary cilia basal body and axoneme, respectively (**Figure 10B**).



Source: Paper II, Figure 1B
Tammimies et al. 2016, FASEB Journal

Source: Paper II, adapted from Figure 4
Tammimies et al. 2016, FASEB Journal

Figure 10: (A) Luciferase reporter assays confirmed the functionality of the X-box motifs (WT = wild type X-box, XboxMut = mutated X-box). (B) Endogenous protein localization of DYX1C1 and DCDC2 primary cilia basal body and axoneme, respectively, in hTERT-RPE1 cells.

Altogether, we proved the complex regulatory mechanism among the different RFX TFs, which may become a challenge in validating the roles of RFX TFs in humans and other vertebrates with multiple members of RFX TFs. Nonetheless, we presented the first evidence for a connection between RFX TFs and cilia within the context of a human brain developmental disorder.

3.3 PAPER III

The *C. elegans* RFX TF DAF-19 regulates target genes in both ciliated and non-ciliated (non-sensory) neurons during development and adulthood.

Using the invertebrate animal model *C. elegans*, we characterized the target genes of RFX TF DAF-19, by RT-PCR and microarray analysis. The comparisons were performed between *daf-12(sa204)* [*daf-19*(+/+)] versus *daf-19*(*m86*); *daf-12(sa204)* [*daf-19*(*null*)] mutants, and at three developmental stages; 3-fold embryo, L1-larvae and adult stages of hermaphrodites. We found that **(i)** the breadth of DAF-19-regulated target genes changed from embryos to L1-larvae to adults, **(ii)** the top score and thus the type of DAF-19-regulated gene ontology clusters changed from cilium assembly/dauer at embryos, to cuticle/collagen at L1-larvae, to signal peptide at adults, and **(iii)** at all three stages, DAF-19-regulated target genes were significantly enriched in neurons.

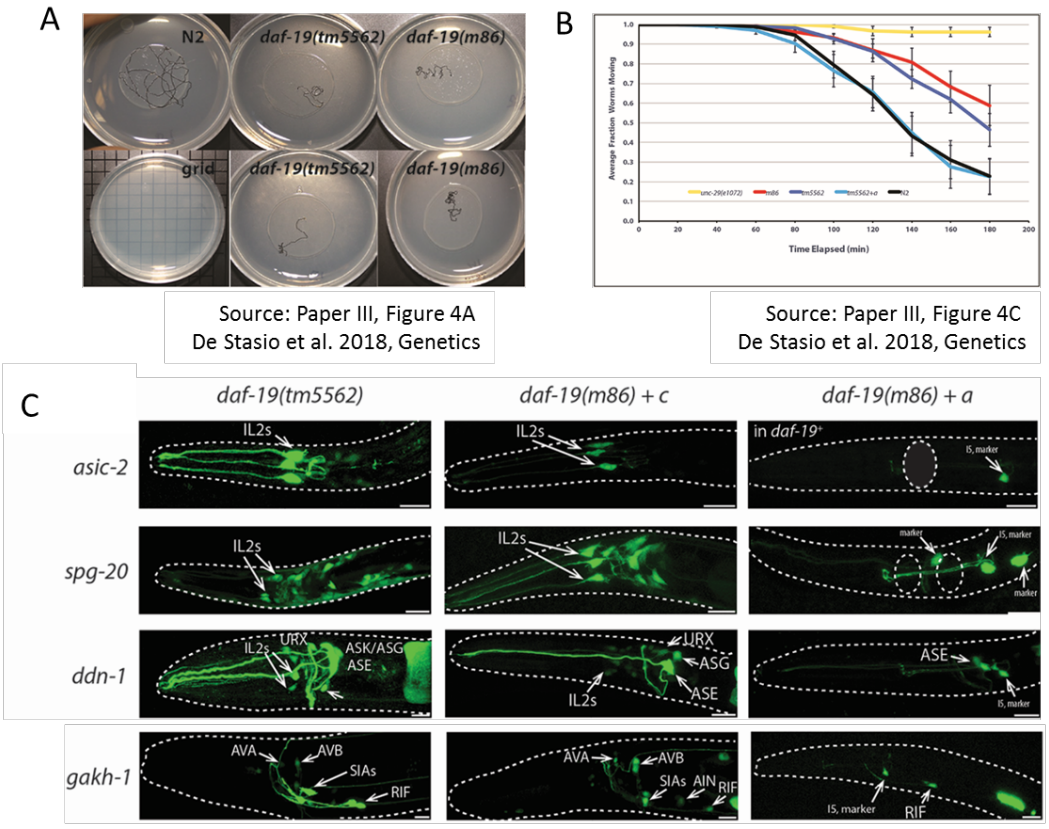
In order to examine in which way DAF-19-regulated target genes were DAF-19-dependent, we assessed the difference in expression patterns of 33 target genes using transcriptional GFP fusions between *daf-19*(+/+) and *daf-19*(*null*) backgrounds. We concluded that **(i)** the inner labial (IL2) ciliated sensory neurons were a common site of DAF-19 activation; **(ii)** the target genes that were DAF-19-dependent (both activated and repressed) affected only neurons (**Table 5**) and that two genes in particular, *gakh-1* and *del-4*, exhibited age-dependent neuronal expressions, and finally **(iii)** the target genes that were DAF-19-independent were predominantly expressed in neurons and the intestine.

Target gene	DAF-19 isoform-specific regulation	Neuronal expression	Neuronal cell type
asic-2 (T28F4.2) *spg-20* (F57B10.9)	DAF-19C activates and DAF-19A overexpression represses	IL2s	Ciliated sensory neurons
ddn-1 (B0507.10)	DAF-19C activates	URX, ASK, AFD, *etc* I5 M5	Ciliated sensory neurons Interneuron Motor neuron
gakh-1 (F46G11.3)	DAF-19A represses	AVA, AVB, SIA, *etc* M4	Interneurons Motor neuron

Table 5: Selected DAF-19-dependent target genes from Paper III.

To dissect functional differences between the DAF-19A/B and C isoforms, we characterized a new isoform-specific *daf-19* allele, *tm5562*, which is a loss-of-function allele affecting both *daf-19a* and *b* isoforms, but not the *c* isoform. We noted the following phenotypes of *daf-19*(*tm5562*) mutants; **(i)** they have cilia and are not Daf-c, **(ii)** they develop more slowly through the L3-larvae stage compared to wild-type N2, and **(iii)** in both roaming and aldicarb assays, the behavior was either indistinguishable from *daf-19*(*null*) or displayed an intermediate phenotype to wild-type N2 (**Figure 11A-B).**

Moreover, *tm5562* phenotypes were rescued by a translational fusion construct of *daf-19a*, suggesting that DAF-19A plays an important and unique role in neuronal function (**Figure 11B**).



Source: Paper III, Figure 4A
De Stasio et al. 2018, Genetics

Source: Paper III, Figure 4C
De Stasio et al. 2018, Genetics

Source: Paper III, adapted from Figure 5
De Stasio et al. 2018, Genetics

Figure 11: *C. elegans* assays and transcriptional GFP-fusion gene expression. (A) Roaming assay (1-hour): *daf-19(tm5562)* worms roam less than wild type N2. (B) Aldicarb assay (500µM): *daf-19(tm5562)* worms are more resistant to aldicarb than wild type N2, and a translational *daf-19(a)* fusion construct rescues the *tm5562* phenotype. (C) The GFP expression pattern of selected DAF-19-dependent target genes (see Table 5) in *daf-19(tm5562)* and in *daf-19(null)* rescued with either *daf-19a* or *daf-19c* constructs.

Subsequently, we used the *daf-19(tm5562)* genetic background to determine whether differential target gene expression required functional DAF-19A. We compared the expression patterns between *daf-19(tm5562)* and *daf-19(null)* mutants for target genes that were DAF-19-dependent. In addition, we examined differences upon genetic rescue with either *daf-19a* or *daf-19c* constructs (**Figure 11C**).

Altogether, we found that **(i)** DAF-19-activated target genes required DAF-19C, but not DAF-19A, **(ii)** DAF-19-repressed target genes appeared to depend primarily on the action of DAF-19A, and **(iii)** a particular amount (stoichiometry) of DAF-19 proteins is important for correct gene expression in some type of neurons (**Table 5**).

4 DISCUSSION

In humans and other vertebrates, multiple members of RFX TFs and the RFX heterodimer interactions present a challenge in dissecting their roles. Our work (Paper I) can be used as a guide to select the RFX TFs in the organ system and cell type of interest. It also highlights the utility of large, systematic data sets to analyze complex regulatory systems in detail in an unbiased manner. This allowed us to provide the most likely X-box motif positioning of the RFX target genes in the human genome. This is helpful in selecting the X-box motifs targeted for mutational analyses.

In addition to the positioning, applying a more stringent X-box motif consensus and checking for the sequence conservation with other vertebrate species will narrow down the candidate X-boxes further (Paper II). The RFX target gene regulation by X-box motif binding is very likely true for the vast majority of RFX-dependent genes. However, finding RFX target genes outside the X-box motif regulation would require transcriptomics (and proteomics) profiling of RFX mutant *versus* wild type cases at different developmental stages (e.g. Paper III).

Human *RFX1-4* and *-RFX7* are prominently expressed in different brain tissues and spinal cord (Paper I), making them the reference RFX TFs in neurons and the nervous system. We confirmed RFX1-3 regulation of the dyslexia candidate genes *DYX1C1* and *DCDC2* in the context of mono-ciliated, non-neuronal, hTERT-RPE1 cells (Paper II).

In the case of *KIAA0319*, the X-box motif was functional but it is likely regulated by other members of RFX TFs. *KIAA0319* is highly expressed in brain tissues [119] and is associated to multi-ciliated cells [122], hearing [128] and axonal regeneration [129]. This ascertains the role of RFX TFs in the specialized function of cilia and neurons in adulthood, similar to the invertebrate animal models DAF-19M in *C. elegans* male sensory neurons [149] and Rfx in *Drosophila* chordotonal neurons [161].

By analyzing the co-clustering patterns between validated RFX target genes and *RFX1-8* expression profiles, we noted an outlier in the nervous system cluster of RFX TFs, namely RFX2 (Paper I). The phenotypes of mice deficient in *Rfx1-4* and *Rfx7* (**Table 3**) support the notion that RFX2 plays a crucial role in the male reproductive system, and may act as a co-factor or form heterodimers with other RFX TFs in other organ systems (e.g. Paper II), including the nervous system.

The extent of RFX member dimerization and their regulatory effect on their target genes are likely dependent on the cellular context. In *C. elegans*, DAF-19 isoforms have different and complementary functions in neurons and their stoichiometry matters for correct target gene expression (Paper III). Therefore, the extent of knockdown and overexpression of the RFX TFs can be assessed in observing the impact on their target genes and the RFX TFs themselves. Our results point toward a need to analyze RFX effects systematically in cell models where the factors are co-overexpressed in pairs, to elucidate the functional effects of possible heterodimer formation.

In analyzing *RFX1-8* promoter and enhancer regions, we found 19 over-represented TFBS profiles (Paper I). The two highest-scoring TFBS profiles were associated to the TFs SP2 (*specificity protein 2*) and E2F4 (*E2 factor 4*). Both SP2 [162, 163] and E2F4 [164, 165] are

involved in aspects of cell cycle progression and development-specific pathways, such as neural development and multi-ciliated cell formation.

We validated by siRNA knockdown and qRT-PCR, SP2 and ESR1 as upstream inhibitors of *RFX2*, *-3*, *-5* and *-7* (Paper I), and *MZF1* as an activator of *RFX3* and *-5* (unpublished work). In all cases, *RFX1* exhibited high standard deviation in the fold-change expression level. Given the self-inhibitory mechanism of RFX1 [84] and its unique role as a tumor suppressor [66], it might be better to study RFX1 in isolation.

Transcriptome profiling of the *RFX*/*daf-19* mutant across *C. elegans* development confirmed the prominent embryonic role of RFX TFs in the development of ciliated sensory neurons (Paper III). Furthermore, the target genes of RFX/DAF-19 and their neuronal expression shed light upon the developmental (Daf-c) and behavioural (foraging defect) output of *RFX*/*daf-19* mutants. As expected, *daf-19* alters gene expression in ciliated sensory neurons, e.g. labial neurons (IL2), dauer-inhibiting neurons (ADF, ASI and ASG) and neurons involved in foraging and locomotion (AIY, ASE, ASI, and BAG).

One of the novel DAF-19-dependent target genes is *gakh-1*. The cyclin G-associated kinase *gakh-1* was repressed by DAF-19A in motor and interneurons (AVA, AVB, SIAD/V, AIN and M4). An ortholog of *gakh-1* is the human *GAK* gene which has been identified in GWAS risk loci for Parkinson's disease [166, 167]. In addition, GAK functions in the endocytosis of synaptic vesicles in mouse neurons [168, 169]. This agrees with the previously hypothesized role of the DAF-19A in synaptic vesicle maintenance in adult worms [151]. Altogether, we provided insight into the role of RFX TFs in both ciliated and non-ciliated neurons, and in both embryonic development and later in adulthood.

5 METHODOLOGICAL CONSIDERATIONS

5.1 DATABASES OF TRANSCRIPTOME AND TF-BINDING PROFILES

Transcriptome and TF-binding profiling technologies are tools to explore genome-wide transcript expressions and TF-DNA binding sites applied in different cellular contexts. The data generated are nowadays stored in large databases which can then be further mined, assessed by bioinformatics tools and subsequently validated experimentally, applying model systems and gene expression analysis.

5.1.1 FANTOM5 database (Paper I)

Based at RIKEN (Yokohama, Japan), the Functional Annotation of The Mammalian Genome 5 (FANTOM5) (http://fantom.gsc.riken.jp/5/) project provided comprehensive expression profiles and functional annotation of human and mouse cell-type-specific transcriptomes, using single-molecule sequencing [170]. The technique Cap Analysis of Gene Expression (CAGE) [171, 172] was applied across more than 800 human and about 400 mouse samples, including tissues, primary cells and cancer cell lines.

The strength of CAGE is in the high-resolution single bp TSS peak. CAGE libraries were sequenced to roughly a depth of 4 million mapped tags per sample. In addition to analysis of TSS (promoter) regions, the technique allowed the identification of an RNA signature for enhancers [173], as well as finding miRNAs [174] and long non-coding RNAs ([175], http://fantom.gsc.riken.jp/6/) as the distal elements which support promoter activity.

5.1.2 JASPAR database (Paper I)

JASPAR (http://jaspar.genereg.net) is the leading open-access database of matrix profiles describing the DNA-binding patterns of TFs and other proteins interacting with DNA in a sequence-specific manner. A given TF has a known set of transcription factor binding sites (TFBSs). These TFBSs are represented in an occurrence table called the position frequency matrix (PFM), which summarizes the number of each nucleotide observed at each position of the aligned TFBSs [176, 177].

The JASPAR database holds collections of PFM nucleotide profiles based on published experiments from ChIP-seq data collections in PAZAR database, ENCODE and modENCODE consortia for *H. sapiens*, *M. musculus*, *D. melanogaster* and *C. elegans*, as well as protein-binding microarrays (PBM) and high-throughput systematic evolution of ligands by exponential enrichment (HT-SELEX) studies.

The database continues to be updated since its inception in 2004 [178], with inclusion of new TF binding profiles every two years, with the most recent version in 2018 [179]. The version used in Paper I was JASPAR 2016 [180].

5.2 CELL LINES AND ANIMAL MODEL

5.2.1 Human cell lines (Papers I and II)

Typically *in vitro* cell cultures are not normal cells as they grow on plastic or unnatural matrices (such as Matrigel), out of their normal developmental context, and are adapted to such conditions. They are often hyper-malignant [181], karyotypically abnormal and have

high copy number variation. Nevertheless, cell cultures are well suited for molecular pathway studies, particularly the human-derived cells.

The MCF7 breast cancer line (Paper I) was a gift from Karin Dahlman-Wright, Karolinska Insitutet. The retinal pigmented epithelial hTERT-RPE1 and neuroblastoma SH-SY5Y cell lines (Paper II) were obtained commercially from the American Type Culture Collection (ATCC: https://www.atcc.org/). Like many immortalized cell lines today, they are de-identified with the consent of the donors.

MCF7 cells were derived from the mammary gland of a 69-year-old female Caucasian and are hypertriploid to hypotetraploid (modal chromosome number = 82, range = 66 to 87). hTERT-RPE1 cells are chromosomally female near-diploid in 90% of the cells counted with additional material at the X-chromosome. SH-SY5Y cells were derived from the bone marrow of a 4-year-old female with neuroblastoma and have a modal chromosome number of 47 with a trisomy of chromosome 1. These cell lines were established in the early 1970's.

hTERT-RPE1 is a common human ciliated cell line model because the proportion of quiescent, ciliated cells increases considerably upon serum starvation [182]. In addition, the epithelial morphology makes the cilia more stretched out and easily visible. For comparison, the length of *C. reinhardtii* motile cilia is 10-14 µm [183], *C. elegans* amphid cilia is 5-10 µm and the serum-starved (=ciliated) hTERT-RPE1 primary cilia is 2-3 µm. One can reasonably expect that primary cilia in most non-serum starved human cell lines are of similar length to cilia from hTERT-RPE1 cells. In Paper II, we did not test whether SH-SY5Y cells were ciliated, only that they were non-serum starved.

5.2.2 *C. elegans* animal model (Paper III)

Neuronal function and circuitry is best studied *in vivo* in whole organisms. Mice and rats are the most widely used animal models in neurobiology. However, rodent brains are not gyrated like humans' and they still lack the predictive power for human neural diseases (e.g. for treatment of Parkinson's disease). Invertebrate animal models, such as *D. melanogaster* and *C. elegans*, have been used to model neuronal function, synaptogenesis and to an extent, neural circuitry. They are preferred model systems at the embryonic and organogenesis stage because of the short propagation time.

The nematode worm *C. elegans* has two sexes: a self-fertilizing hermaphrodite (XX) and a male (XO). Males arise naturally from spontaneous non-disjunction in the hermaphrodite germ line at a low rate of 0.1%. The male occurrence can be increased up to 50% through setting up mating crosses. In standard culture condition, a *C. elegans* hermaphrodite has a short 3-day life cycle and lays approximately 300 eggs (**Figure 6**). Transgenic worms for multi-generational gene expression analyses were generated by microinjection of fluorescence-marker-encoding reporter plasmids.

Resources for the analysis of *C. elegans* anatomy, biology and genome (including genomic variations and their designated allele names) can be viewed in Wormatlas (http://www.wormatlas.org/), Wormbook (http://www.wormbook.org/) and Wormbase (http://www.wormbase.org/), respectively. The wild-type *C. elegans* strain, used in effectively all *C. elegans* laboratories today, is the N2 strain derived from Bristol, United Kingdom.

*In C. elegans, t*he genome editing technique CRISPR-Cas9 has been implemented (reviewed in [184]). We generated frameshift mutations by CRISPR-Cas9 which led to premature stop codons in *daf-19* exons 1 and 4 to create *daf-19* isoform-specific mutants (unpublished work). However, the phenotypes of these worms were not sufficiently solid for further analyses.

Instead, we found *daf-19(tm5562)* from the National BioResource Project (Mitani Laboratory, Japan) and obtained the worms from the Caenorhabditis Genetics Center (CGC). To ensure that the mutants were genetically unambiguous, we outcrossed the mutants six times with wild-type N2 and confirmed the mutations by single multi-worm PCR, RT-PCR and sequencing.

It is convenient that *C. elegans* has only one RFX TF compared to humans which have eight members of RFX TFs. A limitation is that the *RFX/ daf-19(m86)* mutant constitutively enters the dauer-larvae stage, a phenotype which is unique to the nematodes species. However, the molecular pathways identified in dauer formation were conserved endocrine pathways such as the insulin/IGF, TGF-β, serotonergic, and steroid hormone signal transduction pathways [185].

6 CONCLUDING REMARKS AND FUTURE PERSPECTIVES

Our results and other published work support a broad niche of RFX TFs in ciliogenesis, cell cycle and cell fate determination of early development. A smaller niche of RFX TFs is also evident in the maintenance of specialized functions of cells in adulthood, including neurons. Based on RFX target genes, we speculate a common theme of the RFX TFs in inducing a cell cycle arrest and promoting further cell differentiation, and that the RFX targets localize to polarized cell surfaces (e.g. cilia, immune synapse).

Accordingly, the expression level of RFX TFs can be monitored in various cell cycle phases, in cancer *versus* non-cancer cells, and in proliferating *versus* differentiating cells. The molecular and cellular contributions of RFX TFs in the nervous system can thus be explored in two different, but interconnected systems; the developing brain (e.g. neural tube formation, neuronal migration) and the (more) mature brain (e.g. synaptogenesis, synaptic transmission, axon guidance, neuronal circuitry). All in all, we aim for a better understanding of the cellular and molecular etiology of the nervous system phenotypes observed in the absence of functional RFX TFs.

The ciliogenic role of RFX TFs is largely connected to the increasingly recognized importance of primary and motile cilia function in the brain [114, 186, 187]. In particular, neuronal primary cilia are linked to memory of avoidance behaviour in mouse [188] and to increased risk of obesity [189] and type 2 diabetes [190]. Interestingly, while primary cilia are typically associated with proliferating cells [88], they are present in many adult neurons in mice brains around P60 [191-194]. Accordingly, neuronal primary cilia could assume neuronal specific functions such as extra-synaptic integration [195] and connectivity [194, 196].

Out of the eight *RFX* genes in humans, *RFX1-4* and -*RFX7* clustered in various tissues of the brain and spinal cord. We can test the RFX heterodimer specificities for *RFX1-4* (with DIM) by co-immunoprecipitation in the neural tube, adult neurons, the spinal cord and other specific brain tissues. The target genes of RFX TFs can be investigated, either by generating RFX knockouts *in vitro* in human neuronal cells or *in vivo* in vertebrate models, followed by RNA-sequencing to analyze the transcriptomic profiles. In addition, yeast two hybrid systems for large-scale RFX proteins interactions can be performed, particularly for RFX7 (without DIM).

The investigation of the RFX/DAF-19 in the invertebrate *C. elegans* revealed shifting DAF-19 target genes during development and that multiple DAF-19 isoforms can influence one another. The target genes of DAF-19 and their expression patterns serve as a guide for RFX TFs in vertebrate organisms. The role of DAF-19 changes from ciliogenesis (embryonic stage) to neuronal function by synaptic maintenance (adult stage). This confirms the role of RFX TF in the adult neuron function. At the same time, many of the sensory neurons in *C. elegans* are neuroendocrine cells which receive sensory input and produce hormones. The equivalents in vertebrates, the brain hypothalamus and pituitary gland, would thus be interesting brain parts to check in connection to the role of RFX TFs in neurons and the neuroendocrine cells.

Brain development begins with the neural tube formation at the embryonic stage. The neural tube contains neural progenitor cells which will mature into neurons and glial cells. Primary cilia ensure proper patterning and maturation of the neural progenitors through Sonic hedgehog (Shh) signaling [108, 197]. *Rfx4* mutants in mouse and zebrafish produced less Shh signaling because Rfx4 regulated the key components in cilia [59, 135]. As expected, the

primary and motile cilia were also structurally abnormal [59, 82]. The role of RFX4 in brain development may be more specific to the primary cilia-based Shh signaling and to the motile cilia-cerebrospinal fluid flow. Interestingly, the former corresponds to homozygous *Rfx4*$^{-/-}$ [59, 82, 134] and the latter corresponds to heterozygous *Rfx4*$^{+/-}$ phenotypes [82, 146].

Another aspect of brain development is the migration of inhibitory interneurons [198]. The cellular etiology of neuronal migration is mainly explained by centrosomal defects, cell cycle delays and spindle disorientations [187, 199-201]. In addition, primary cilia as signaling hubs could promote tangentially migrating neurons. This was documented by the interplay between Shh signaling and the ciliary genes *Arl13b*, *Ift88* and *Kif3a* [202, 203]. The ciliogenic *Rfx3*$^{-/-}$ mice mutants exhibited both motile and primary ciliopathy phenotypes, as well as axon guidance defects and agenesis of the brain corpus callosum [48, 144, 145]. Other than Shh signaling, FGF8 signaling was also implicated in *Rfx3*$^{-/-}$ mutants [144].

The proper function of neurons and their connectivity are supported by neuronal support (glia) cells. Based on expression values in primary cells, *RFX1-4* and -*RFX7* were expressed in neural stem cells, neurons and astrocytes (Paper I). While the vast majority of cilia in the brain extend from neurons, some astrocytes are also ciliated [191, 193]. One of the roles of astrocytes is to prune the synaptic circuit [204, 205]. It would thus be interesting to analyze RFX TF functions in astrocytes or in a co-culture of neurons and astrocytes.

We hypothesize that developmental dyslexia is a rather mild phenotype among the human-brain-specific neural ciliopathies based on the established links between dyslexia candidate genes to neurons and cilia. Furthermore, the regulation of these genes by ciliogenic RFX TFs prompts the need to further characterize the role of neuronal (or neural) primary cilia. For instance, it is still unclear why mature neurons would require the primary cilia if they can send signals (neurotransmitters) more efficiently through the synapses. Would primary cilia detect slow-acting neuromodulators instead? What makes a neuron ciliated and the other non-ciliated? The studies from these dyslexia candidate genes can be a platform for insight into other, more severe, human brain ciliopathies.

Based on the severe phenotypes observed in *Rfx1-6* knockout mice (**Table 3**), we would expect that most individuals with null mutations in *RFX1-6* would be spontaneously aborted or die shortly after birth (e.g. Mitchell Riley syndrome). RFX1-6 have conserved DBD and DIM, and technically should be able to form heterodimers or compensate for one another, especially since *RFX1-3* and -*RFX7* are broadly expressed (Paper I). It would be insightful to solve how these RFX TFs regulate or compete with one another, perhaps by analyzing the other possible functional domains (B and C) of RFX TFs, as of yet uncharacterized.

7 POPULAR SCIENCE SUMMARY

Imagine that our DNA consists of the 26-letter English alphabet. Having the letters themselves, however, does not mean anything. For example, if the DNA were letters A to Z repeated over millions of times, we would need an instructor to arrange specific letters to be read as sensible words of instructions. These instructors are called the transcription factors, and the words of instructions are what scientists collectively term “gene expression”, when DNA is transcribed into RNA (and some RNA is further translated to protein). While every cell in our body has the same DNA, they can look and behave very differently (skin cells, brain cells, *etc*) as a result of differences in gene expression. The transcription factors themselves are proteins acting on the DNA in the cell nucleus, and serve as an example of gene expression.

Different transcription factors bring about different gene expressions. While it is important that a particular transcription factor executes the correct gene expression, it is equally interesting to understand what this expression means to the cell. The RFX transcription factors are generally known to instruct the cell to make a structure called the cilium (plural = cilia). This is especially true for the worm *C. elegans*, which has only one RFX transcription factor called DAF-19. In humans, RFX transcription factors have diverged to multiple members and functions in different cell types.

We provide an exhaustive and updated survey of the eight members of the RFX transcription factor family in humans (Paper I). Human RFX1, RFX2 and RFX3 regulate the expression of the genes inherited in families with developmental dyslexia, and some of these genes become proteins that are parts of the primary cilia structure (Paper II). Furthermore, digging deeper into the gene expression of the DAF-19/RFX transcription factor in *C. elegans* worms, we discovered a novel role outside cilia, but within neurons in general (Paper III).

So what are cilia and what happens when they are not made properly? Cilia are hair-like cell protrusions which can be – per cell – single-hair-like or multiple-hair-like. For example, sperm uses their single-cilia as beating tails to move, while our lung airways are lined with multiple cilia hairs per cell to enable the flow of mucus. Here, we are interested in one type of cilia which is less understood despite being present on almost every cell of our body. It is called the primary cilium. A stationary single-cilium, it behaves like the cell’s antenna that can act as a signaling hub (including the Hedgehog – not the animal). Diseases caused by malfunctions of any type of cilia are termed “ciliopathies”. As expected, ciliopathies are complex and vary in severity depending on the cell types affected.

In short, we know that RFX transcription factors are prominent in the human brain and are important for cilia and neurons. Follow-up questions would be: Why do the neurons have primary cilia (= antennae) if they can communicate way more efficiently with one another through a large number of dedicated direct contact points (= synapses) that they construct along their large meshwork of neurites and processes (= cables)? Can we prove that developmental dyslexia is a mild form of a human brain ciliopathy? And, a relevant technical question: How do we model the human brain? All kinds of models will fall short, but they are nevertheless insightful as we tap into the wonder of how our brains are able to process information such as reading.

8 ACKNOWLEDGEMENTS

I thank my supervisors, Peter Swoboda and Juha Kere, and their complementary roles in supervising my PhD project.

Peter has been instrumental in getting me to Sweden in 2011 after I completed my Bachelor education in Australia. He gave me the once-in-a-lifetime opportunity to work as a Master-level research student and later accepted me as a PhD student. In good spirit, there were times that he wanted to postpone my PhD graduation indefinitely and also times when he wanted me out as soon as possible. All in all, he ensures that I finish the work well while giving me much room to be independent. He is tough, careful and a role model for being a scientist. Thank you, Peter for all these important lessons. I wish you all the best in your research and your family life.

When I first met Juha in 2013, he was affiliated to Karolinska Institutet and University of Helsinki. By the time I wrote this thesis in 2018, he acquired a third affiliation in King's College, London and had just completed a 10-month-fellowship in Japan. I do not think many people would manage four different time zones like Juha did while staying sane and positive. I believe one key to his success is that he considers the people to be the most important resource. Thank you, Juha for being a role model in your work ethic, for your optimistic energy in leadership and for giving me valuable career advice.

I thank the current and previous group members of Peter Swoboda, Juha Kere, Thomas Bürglin, Carsten Daub and Lennart Nilsson; Gilbert, Xin Xuan, Andrea C., Deffiz, Flavie, Shalini, Mariangela, Anastasia, Juan-Carlos, Agatha, Tiina, Andrea B., Beta, Shintaro, Hong, Kaarel, Masahito, Gayathri, Morana, Eeva-Mari, Nancy, Isabel, Myriam, Hans, Elo, Jingwen, Lisa, Cilla, Nathalie, Lovisa, Ingegerd, Kristiina, the MAF team, Johan H., Michel, Umesh, Lois, Enrichetta, Mattias, Kadir, Amitha, Alessandra, Marzieh, Vera, Yossa, You, Zohreh, Arzu and Eva. Thank you for the scientific discussion and the experimental help, and for creating a great atmosphere in the workplace. I have also had fun in conversations with many of you at the lunch table, during conferences, and activities together outside work. I wish you all the success in your present and future endeavors.

Through collaboration with Peter, I worked for some time in Junho Lee's research group at the Seoul National University. Junho was like a third supervisor for me and I felt an instant sense of belonging to his group. I want to thank members of Junho's group; Hyesook, Dongjun, Myungkyu, Jihyun, Soungyub, Daehan, Boyun, Son, Hanee, Sanghyun, Daisy, Beomseok, Chuna, Naree, Jun, Heesung, Soo, Sungjong, Eunkyeong and Imu. I hope to see you again in Korea to share a meal of samgyeopsal with makgeolli.

Carolyn Behm practically introduced me to research. Thanks to Carol, Peter was convinced to accept me all the way from down under. She introduced me to the world of *C. elegans* as there is much to learn from this tiny worm. I also want to thank all my collaborators worldwide with whom I have had the privilege to work with. A special mention goes out to Beth De Stasio, who has included me in the long-term DAF-19 collaboration with Peter and offered many kind words of encouragement.

I want to thank the administration team in the department, with special thanks to Monica Ahlberg and Eva Nordlander. You have both been very helpful and efficient in maneuvering through the changing systems.

I want to thank the following people, with whom I could have a thriving social life outside work.

To the Olssons: Helene, Jan, the late Snoddas, Jonesy (and hopefully one day I meet Grace) and to the Engbloms: Eva, Hans, Cecilia, Camilla, the late Mr Krum, and Sir Kasper. You have given me a wonderful first and lasting impression of being Swedish. Thank you for offering your home when I need it and for sharing your lives with us. To my brothers and sisters in Christ at the Immanuel Church (Immanuelskyrkan) and the Indonesian Reformed Evangelical Fellowship (PRII) in Stockholm, you are my family away from home. Thank you for the expression of love, strength, courage, wisdom and hope through the fellowships and prayers. I am thankful for Jingwen and Yossa as well, for being my sister and brother in the workplace.

To the Trapman family, Pa Piet and Ma Nel, Mirjam, Jos and little Pim, Nelly and Bernard: you have lovingly included me as daughter, sister and *tante* in-law. To my sister, Elhana Sugiaman and her boyfriend, Tom Frame: for showing support throughout my PhD project. To my father, Djonady Sugiaman, you have sacrificed so much. With sweat and tears, you work very hard to make sure that your daughters obtain the best education. This thesis is yours Papa, because my success is also your success. To my mother, Elly Sudjana-Sugiaman, I feel very blessed to have a kind and wise mother. Both of you were very close to getting PhD titles in Sweden (my father at the Stockholm University and my mother at Karolinska Institutet). Unfortunately, difficult circumstances compelled you to give them up and settled for the Licentiate degrees instead. Nevertheless, I am grateful that you stay together for your children's sake and that both of you thrive today in your professional lives, thus encouraging me not to glorify the title of a PhD.

I would not be able to go through the PhD education, literally from start to finish, without the support of my husband. My dear Pieter, you are the kindest, wisest and sweetest person I know. Thank you for asking my hand in marriage and for being my secure, *gezellig* home. I look forward to spending many more years ahead with you and hopefully our dog Kerberos "Kerbie", as my two smiling Dutchmen.

9 REFERENCES

1. Jolma A, Yan J, Whitington T, Toivonen J, Nitta KR, Rastas P, Morgunova E, Enge M, Taipale M, Wei G *et al*: **DNA-Binding Specificities of Human Transcription Factors**. *Cell* 2013, **152**(1–2):327-339.

2. Alberts B, Johnson A, Lewis J, Morgan D, Raff M, Roberts K, Walter P: **Molecular Biology of the Cell. 4th edition**. New York: Garland Science; 2002.

3. Roy AL, Singer DS: **Core promoters in transcription: old problem, new insights**. *Trends in Biochemical Sciences* 2015, **40**(3):165-171.

4. Chepelev I, Wei G, Wangsa D, Tang Q, Zhao K: **Characterization of genome-wide enhancer-promoter interactions reveals co-expression of interacting genes and modes of higher order chromatin organization**. *Cell Research* 2012, **22**:490–503.

5. Kim T-K, Shiekhattar R: **Architectural and Functional Commonalities between Enhancers and Promoters**. *Cell* 2015, **162**(5):948-959.

6. Holoch D, Moazed D: **RNA-mediated epigenetic regulation of gene expression**. *Nature Reviews Genetics* 2015, **16**:71–84.

7. Vaquerizas JM, Kummerfeld SK, Teichmann SA, Luscombe NM: **A census of human transcription factors: function, expression and evolution**. *Nature Reviews Genetics* 2009, **10**(4):252-263.

8. Töhönen V, Katayama S, Vesterlund L, Jouhilahti E-M, Sheikhi M, Madissoon E, Filippini-Cattaneo G, Jaconi M, Johnsson A, Bürglin TR *et al*: **Novel PRD-like homeodomain transcription factors and retrotransposon elements in early human development**. *Nature Communications* 2015, **6**:8207.

9. Gajiwala KS, Chen H, Cornille F, Roques BP, Reith W, Mach B, Burley SK: **Structure of the winged-helix protein hRFX1 reveals a new mode of DNA binding**. *Nature* 2000, **403**:916-921.

10. Piasecki BP, Burghoorn J, Swoboda P: **Regulatory Factor X (RFX)-mediated transcriptional rewiring of ciliary genes in animals**. *Proceedings of the National Academy of Sciences* 2010, **107**(29):12969-12974.

11. Gaudier M, Schuwirth BS, Westcott SL, Wigley DB: **Structural Basis of DNA Replication Origin Recognition by an ORC Protein**. *Science* 2007, **317**(5842):1213-1216.

12. Duncker BP, Chesnokov IN, McConkey BJ: **The origin recognition complex protein family**. *Genome Biology* 2009, **10**(3):214.

13. Garg A, Futcher B, Leatherwood J: **A new transcription factor for mitosis: in *Schizosaccharomyces pombe*, the RFX transcription factor Sak1 works with forkhead factors to regulate mitotic expression**. *Nucleic Acids Research* 2015, **43**(14):6874-6888.

14. Zaim J, Speina E, Kierzek AM: **Identification of New Genes Regulated by the Crt1 Transcription Factor, an Effector of the DNA Damage Checkpoint Pathway in *Saccharomyces cerevisiae***. *Journal of Biological Chemistry* 2005, **280**(1):28-37.

15. Swoboda P, Adler HT, Thomas JH: **The RFX-Type Transcription Factor DAF-19 Regulates Sensory Neuron Cilium Formation in *C. elegans***. *Molecular Cell* 2000, **5**(3):411-421.

16. Otsuki K, Hayashi Y, Kato M, Yoshida H, Yamaguchi M: **Characterization of dRFX2, a novel RFX family protein in *Drosophila***. *Nucleic Acids Research* 2004, **32**(18):5636-5648.

17. Chu J, Baillie D, Chen N: **Convergent evolution of RFX transcription factors and ciliary genes predated the origin of metazoans**. *BMC Evolutionary Biology* 2010, **10**(1):130.

18. Emery P, Durand B, Mach B, Reith W: **RFX proteins, a novel family of DNA binding proteins conserved in the eukaryotic kingdom**. *Nucleic Acids Research* 1996, **24**:803-807.

19. Aftab S, Semenec L, Chu J, Chen N: **Identification and characterization of novel human tissue-specific RFX transcription factors**. *BMC Evolutionary Biology* 2008, **8**(1):226.

20. Choksi SP, Lauter G, Swoboda P, Roy S: **Switching on cilia: transcriptional networks regulating ciliogenesis**. *Development* 2014, **141**(7):1427-1441.

21. Emery P, Strubin M, Hofmann K, Bucher P, Mach B, Reith W: **A consensus motif in the RFX DNA binding domain and binding domain mutants with altered specificity**. *Molecular and Cellular Biology* 1996, **16**(8):4486-4494.

22. Johnson AD: **An extended IUPAC nomenclature code for polymorphic nucleic acids**. *Bioinformatics* 2010, **26**(10):1386-1389.

23. Badis G, Berger MF, Philippakis AA, Talukder S, Gehrke AR, Jaeger SA, Chan ET, Metzler G, Vedenko A, Chen X *et al*: **Diversity and Complexity in DNA Recognition by Transcription Factors**. *Science* 2009, **324**(5935):1720-1723.

24. Reith W, Ucla C, Barras E, Gaud A, Durand B, Herrero-Sanchez C, Kobr M, Mach B: **RFX1, a transactivator of hepatitis B virus enhancer I, belongs to a novel family of homodimeric and heterodimeric DNA-binding proteins**. *Molecular and Cellular Biology* 1994, **14**(2):1230-1244.

25. Reith W, Satola S, Sanchez CH, Amaldi I, Lisowska-Grospierre B, Griscelli C, Hadam MR, Mach B: **Congenital immunodeficiency with a regulatory defect in MHC class II gene expression lacks a specific HLA-DR promoter binding protein, RF-X**. *Cell* 1988, **53**(6):897-906.

26. Reith W, Steimle V, Mach B: **Molecular defects in the bare lymphocyte syndrome and regulation of MHC class II genes**. *Immunology Today* 1995, **16**(11):539-546.

27. Chakraborty M, Sengupta A, Bhattacharya D, Banerjee S, Chakrabarti A: **DNA binding domain of RFX5: Interactions with X-box DNA and RFXANK**. *Biochimica et Biophysica Acta (BBA) - Proteins and Proteomics* 2010, **1804**(10):2016-2024.

28. Nie J, Xu C, Jin J, Aka JA, Tempel W, Nguyen V, You L, Weist R, Min J, Pawson T *et al*: **Ankyrin Repeats of ANKRA2 Recognize a PxLPxL Motif on the 3M Syndrome Protein CCDC8**. *Structure* 2015, **23**(4):700-712.

29. Morotomi-Yano K, Yano K-i, Saito H, Sun Z, Iwama A, Miki Y: **Human Regulatory Factor X 4 (RFX4) Is a Testis-specific Dimeric DNA-binding Protein That Cooperates with Other Human RFX Members**. *Journal of Biological Chemistry* 2002, **277**(1):836-842.

30. Smith SB, Qu H-Q, Taleb N, Kishimoto NY, Scheel DW, Lu Y, Patch A-M, Grabs R, Wang J, Lynn FC *et al*: **Rfx6 directs islet formation and insulin production in mice and humans**. *Nature* 2010, **463**(7282):775-780.

31. Burghoorn J, Piasecki BP, Crona F, Phirke P, Jeppsson KE, Swoboda P: **The *in vivo* dissection of direct RFX-target gene promoters in *C. elegans* reveals a novel cis-regulatory element, the C-box**. *Developmental Biology* 2012, **368**(2):415-426.

32. Varshney A, Scott LJ, Welch RP, Erdos MR, Chines PS, Narisu N, Albanus RDO, Orchard P, Wolford BN, Kursawe R *et al*: **Genetic regulatory signatures underlying islet gene expression and type 2 diabetes**. *Proceedings of the National Academy of Sciences* 2017, **114**(9):2301–2306.

33. Blacque OE, Perens EA, Boroevich KA, Inglis PN, Li C, Warner A, Khattra J, Holt RA, Ou G, Mah AK *et al*: **Functional genomics of the cilium, a sensory organelle**. *Current Biology* 2005, **15**:935-941.

34. Efimenko E, Bubb K, Mak HY, Holzman T, Leroux MR, Ruvkun G, Thomas JH, Swoboda P: **Analysis of xbx genes in *C. elegans***. *Development* 2005, **132**(8):1923-1934.

35. Chen N, Mah A, Blacque OE, Chu J, Phgora K, Bakhoum MW, Hunt Newbury CR, Khattra J, Chan S, Go A *et al*: **Identification of ciliary and ciliopathy genes in *Caenorhabditis elegans* through comparative genomics**. *Genome Biology* 2006, **7**:R126.

36. Laurencon A, Dubruille R, Efimenko E, Grenier G, Bissett R, Cortier E, Rolland V, Swoboda P, Durand B: **Identification of novel regulatory factor X (RFX) target genes by comparative genomics in *Drosophila* species**. *Genome Biology* 2007, **8**(9):R195.

37. Xie X, Mikkelsen TS, Gnirke A, Lindblad-Toh K, Kellis M, Lander ES: **Systematic discovery of regulatory motifs in conserved regions of the human genome, including thousands of CTCF insulator sites**. *Proceedings of the National Academy of Sciences* 2007, **104**(17):7145-7150.

38. Phirke P, Efimenko E, Mohan S, Burghoorn J, Crona F, Bakhoum MW, Trieb M, Schuske K, Jorgensen EM, Piasecki BP *et al*: **Transcriptional profiling of *C. elegans* DAF-19 uncovers a ciliary base-associated protein and a CDK/CCRK/LF2p-related kinase required for intraflagellar transport**. *Developmental Biology* 2011, **357**(1):235-247.

39. Chu JSC, Tarailo-Graovac M, Zhang D, Wang J, Uyar B, Tu D, Trinh J, Baillie DL, Chen N: **Fine tuning of RFX/DAF-19-regulated target gene expression through binding to multiple sites in *Caenorhabditis elegans***. *Nucleic Acids Research* 2012, **40**(1):53-64.

40. Henriksson J, Piasecki BP, Lend K, Bürglin TR, Swoboda P: **Chapter Sixteen - Finding Ciliary Genes: A Computational Approach**. In: *Methods in Enzymology*. Edited by Wallace FM, vol. 525. Amsterdam, Netherlands: Academic Press, Elsevier; 2013: 327-350.

41. Steimle V, Durand B, Barras E, Zufferey M, Hadam MR, Mach B, Reith W: **A novel DNA-binding regulatory factor is mutated in primary MHC class II deficiency (bare lymphocyte syndrome)**. *Genes & Development* 1995, **9**(9):1021-1032.

42. Rousseau P, Masternak K, Krawczyk M, Reith W, Dausset J, Carosella ED, Moreau P: ***In vivo*, RFX5 binds differently to the human leucocyte antigen-E, -F, and -G gene promoters and participates in HLA class I protein expression in a cell type-dependent manner**. *Immunology* 2004, **111**(1):53-65.

43. Meissner TB, Liu Y-J, Lee K-H, Li A, Biswas A, van Eggermond MCJA, van den Elsen PJ, Kobayashi KS: **NLRC5 Cooperates with the RFX Transcription Factor Complex To Induce MHC Class I Gene Expression**. *The Journal of Immunology* 2012, **188**(10):4951-4958.

44. Iwama A, Pan J, Zhang P, Reith W, Mach B, Tenen DG, Sun Z: **Dimeric RFX Proteins Contribute to the Activity and Lineage Specificity of the Interleukin-5 Receptor α Promoter through Activation and Repression Domains**. *Molecular and Cellular Biology* 1999, **19**(6):3940-3950.

45. Zhao M, Sun Y, Gao F, Wu X, Tang J, Yin H, Luo Y, Richardson B, Lu Q: **Epigenetics and SLE: RFX1 downregulation causes CD11a and CD70 overexpression by altering epigenetic modifications in lupus CD4+ T cells**. *Journal of Autoimmunity* 2010, **35**(1):58-69.

46. El Zein L, Ait-Lounis A, Morlé L, Thomas J, Chhin B, Spassky N, Reith W, Durand B: **RFX3 governs growth and beating efficiency of motile cilia in mouse and controls the expression of genes involved in human ciliopathies**. *Journal of Cell Science* 2009, **122**(17):3180-3189.

47. Chung M-I, Kwon T, Tu F, Brooks ER, Gupta R, Meyer M, Baker JC, Marcotte EM, Wallingford JB, Bronner ME: **Coordinated genomic control of ciliogenesis and cell movement by RFX2**. *eLife* 2014, **3**:e01439.

48. Baas D, Meiniel A, Benadiba C, Bonnafe E, Meiniel O, Reith W, Durand B: **A deficiency in RFX3 causes hydrocephalus associated with abnormal differentiation of ependymal cells**. *European Journal of Neuroscience* 2006, **24**:1020-1030.

49. Didon L, Zwick R, Chao IW, Walters M, Wang R, Hackett N, Crystal R: **RFX3 Modulation of FOXJ1 regulation of cilia genes in the human airway epithelium**. *Respiratory Research* 2013, **14**(1):70.

50. Fuxman Bass JI, Pons C, Kozlowski L, Reece-Hoyes JS, Shrestha S, Holdorf AD, Mori A, Myers CL, Walhout AJ: **A gene-centered *C. elegans* protein–DNA interaction network provides a framework for functional predictions**. *Molecular Systems Biology* 2016, **12**(10):884.

51. Dubruille R, Laurencon A, Vandaele C, Shishido E, Coulon-Bublex M, Swoboda P, Couble P, Kernan M, Durand B: ***Drosophila* regulatory factor X is necessary for ciliated sensory neuron differentiation**. *Development* 2002, **129**:5487-5498.

52. Liu Y, Pathak N, Kramer-Zucker A, Drummond IA: **Notch signaling controls the differentiation of transporting epithelia and multiciliated cells in the zebrafish pronephros**. *Development* 2007, **134**(6):1111-1122.

53. Yu X, Ng CP, Habacher H, Roy S: **Foxj1 transcription factors are master regulators of the motile ciliogenic program**. *Nature Genetics* 2008, **40**(12):1445-1453.

54. Bisgrove BW, Makova S, Yost HJ, Brueckner M: **RFX2 is essential in the ciliated organ of asymmetry and an RFX2 transgene identifies a population of ciliated cells sufficient for fluid flow**. *Developmental Biology* 2012, **363**(1):166-178.

55. Chung M-I, Peyrot SM, LeBoeuf S, Park TJ, McGary KL, Marcotte EM, Wallingford JB: **RFX2 is broadly required for ciliogenesis during vertebrate development**. *Developmental Biology* 2012, **363**(1):155-165.

56. Manojlovic Z, Earwood R, Kato A, Stefanovic B, Kato Y: **RFX7 is required for the formation of cilia in the neural tube**. *Mechanisms of Development* 2014, **132**:28-37.

57. Elkon R, Milon B, Morrison L, Shah M, Vijayakumar S, Racherla M, Leitch CC, Silipino L, Hadi S, Weiss-Gayet M *et al*: **RFX transcription factors are essential for hearing in mice**. *Nature Communications* 2015, **6**:8549.

58. Wu Y, Hu X, Li Z, Wang M, Li S, Wang X, Lin X, Liao S, Zhang Z, Feng X *et al*: **Transcription Factor RFX2 Is a Key Regulator of Mouse Spermiogenesis**. *Scientific Reports* 2016, **6**:20435.

59. Ashique AM, Choe Y, Karlen M, May SR, Phamluong K, Solloway MJ, Ericson J, Peterson AS: **The Rfx4 Transcription Factor Modulates Shh Signaling by Regional Control of Ciliogenesis**. *Science Signaling* 2009, **2**(95):ra70.

60. Kang GM, Han YM, Ko HW, Kim J, Oh BC, Kwon I, Kim M-S: **Leptin Elongates Hypothalamic Neuronal Cilia via Transcriptional Regulation and Actin Destabilization**. *Journal of Biological Chemistry* 2015, **290**(29):18146-18155.

61. Purvis TL, Hearn T, Spalluto C, Knorz VJ, Hanley KP, Sanchez-Elsner T, Hanley NA, Wilson DI: **Transcriptional regulation of the Alström syndrome gene *ALMS1* by members of the RFX family and Sp1**. *Gene* 2010, **460**:20-29.

62. Nakayama A, Murakami H, Maeyama N, Yamashiro N, Sakakibara A, Mori N, Takahashi M: **Role for RFX Transcription Factors in Non-neuronal Cell-specific Inactivation of the Microtubule-associated Protein MAP1A Promoter**. *Journal of Biological Chemistry* 2003, **278**(1):233-240.

63. Lee JH, Silhavy JL, Lee JE, Al-Gazali L, Thomas S, Davis EE, Bielas SL, Hill KJ, Iannicelli M, Brancati F *et al*: **Evolutionarily Assembled cis-Regulatory Module at a Human Ciliopathy Locus**. *Science* 2012, **335**(6071):966-969.

64. Liu M, Lee BH, Mathews MB: **Involvement of RFX1 Protein in the Regulation of the Human Proliferating Cell Nuclear Antigen Promoter**. *Journal of Biological Chemistry* 1999, **274**(22):15433-15439.

65. Boehm EM, Gildenberg MS, Washington MT: **The many roles of PCNA in eukaryotic DNA replication**. *The Enzymes* 2016, **39**:231-254.

66. Chen L, Smith L, Johnson MR, Wang K, Diasio RB, Smith JB: **Activation of Protein Kinase C Induces Nuclear Translocation of RFX1 and Down-regulates c-myc via an Intron 1 X Box in Undifferentiated Leukemia HL-60 Cells**. *Journal of Biological Chemistry* 2000, **275**(41):32227-32233.

67. Feng C, Zuo Z: **Regulatory Factor X1-induced Down-regulation of Transforming Growth Factor β2 Transcription in Human Neuroblastoma Cells**. *Journal of Biological Chemistry* 2012, **287**(27):22730-22739.

68. Amin S, Kumar A, Nilchi L, Wright K, Kozlowski M: **Breast Cancer Cells Proliferation Is Regulated by Tyrosine Phosphatase SHP1 through c-jun N-**

Terminal Kinase and Cooperative Induction of RFX-1 and AP-4 Transcription Factors. *Molecular Cancer Research* 2011, **9**(8):1112-1125.

69. Feng C, Xu W, Zuo Z: **Knockout of the regulatory factor X1 gene leads to early embryonic lethality**. *Biochemical and Biophysical Research Communications* 2009, **386**(4):715-717.

70. Wolfe SA, Wilkerson DC, Prado S, Grimes SR: **Regulatory factor X2 (RFX2) binds to the H1t/TE1 promoter element and activates transcription of the testis-specific histone H1t gene**. *Journal of Cellular Biochemistry* 2004, **91**(2):375-383.

71. vanWert JM, Wolfe SA, Grimes SR: **Binding of RFX2 and NF-Y to the testis-specific histone H1t promoter may be required for transcriptional activation in primary spermatocytes**. *Journal of Cellular Biochemistry* 2008, **104**(3):1087-1101.

72. Horvath GC, Kistler MK, Kistler WS: **RFX2 is a candidate downstream amplifier of A-MYB regulation in mouse spermatogenesis**. *BMC Developmental Biology* 2009, **9**:63-63.

73. Soyer J, Flasse L, Raffelsberger W, Beucher A, Orvain C, Peers B, Ravassard P, Vermot J, Voz ML, Mellitzer G *et al*: **Rfx6 is an Ngn3-dependent winged helix transcription factor required for pancreatic islet cell development**. *Development* 2010, **137**(2):203-212.

74. Piccand J, Strasser P, Hodson DJ, Meunier A, Ye T, Keime C, Birling M-C, Rutter GA, Gradwohl G: **Rfx6 Maintains the Functional Identity of Adult Pancreatic β Cells**. *Cell Reports* 2014, **9**(6):2219-2232.

75. Chandra V, Albagli-Curiel O, Hastoy B, Piccand J, Randriamampita C, Vaillant E, Cavé H, Busiah K, Froguel P, Vaxillaire M *et al*: **RFX6 Regulates Insulin Secretion by Modulating Ca^{2+} Homeostasis in Human β Cells**. *Cell Reports* 2014, **9**(6):2206-2218.

76. Zhu Z, Li QV, Lee K, Rosen BP, González F, Soh C-L, Huangfu D: **Genome Editing of Lineage Determinants in Human Pluripotent Stem Cells Reveals Mechanisms of Pancreatic Development and Diabetes**. *Cell Stem Cell* 2016, **18**(6):755-768.

77. Spiegel R, Dobbie A, Hartman C, de Vries L, Ellard S, Shalev SA: **Clinical characterization of a newly described neonatal diabetes syndrome caused by *RFX6* mutations**. *American Journal of Medical Genetics Part A* 2011, **155**(11):2821-2825.

78. Concepcion JP, Reh CS, Daniels M, Liu X, Paz VP, Ye H, Highland HM, Hanis CL, Greeley SAW: **Neonatal diabetes, gallbladder agenesis, duodenal atresia, and intestinal malrotation caused by a novel homozygous mutation in *RFX6***. *Pediatric Diabetes* 2014, **15**(1):67-72.

79. Xu Y, Sengupta PK, Seto E, Smith BD: **Regulatory Factor for X-box Family Proteins Differentially Interact with Histone Deacetylases to Repress Collagen α2(I) Gene (*COL1A2*) Expression**. *Journal of Biological Chemistry* 2006, **281**(14):9260-9270.

80. Cachero S, Simpson TI, zur Lage PI, Ma L, Newton FG, Holohan EE, Armstrong JD, Jarman AP: **The Gene Regulatory Cascade Linking Proneural Specification with Differentiation in *Drosophila* Sensory Neurons**. *PLOS Biology* 2011, **9**(1):e1000568.

81. Spassky N, Meunier A: **The development and functions of multiciliated epithelia**. *Nature Reviews Molecular Cell Biology* 2017, **18**:423–436.

82. Xu P, Morrison JP, Foley JF, Stumpo DJ, Ward T, Zeldin DC, Blackshear PJ: **Conditional ablation of the RFX4 isoform 1 transcription factor: Allele dosage effects on brain phenotype**. *PLOS One* 2018, **13**(1):e0190561.

83. Beckers A, Alten L, Viebahn C, Andre P, Gossler A: **The mouse homeobox gene Noto regulates node morphogenesis, notochordal ciliogenesis, and left–right patterning**. *Proceedings of the National Academy of Sciences* 2007, **104**(40):15765-15770.

84. Lubelsky Y, Reuven N, Shaul Y: **Autorepression of Rfx1 Gene Expression: Functional Conservation from Yeast to Humans in Response to DNA Replication Arrest**. *Molecular and Cellular Biology* 2005, **25**(23):10665-10673.

85. Huang M, Zhou Z, Elledge SJ: **The DNA Replication and Damage Checkpoint Pathways Induce Transcription by Inhibition of the Crt1 Repressor**. *Cell* 1998, **94**(5):595-605.

86. Ibañez-Tallon I, Heintz N, Omran H: **To beat or not to beat: roles of cilia in development and disease**. *Human Molecular Genetics* 2003, **12**:R27-R35.

87. Carvalho-Santos Z, Azimzadeh J, Pereira-Leal JB, Bettencourt-Dias M: **Tracing the origins of centrioles, cilia, and flagella**. *The Journal of Cell Biology* 2011, **194**(2):165-175.

88. Ishikawa H, Marshall WF: **Ciliogenesis: building the cell's antenna**. *Nature Reviews Molecular Cell Biology* 2011, **12**(4):222-234.

89. Fliegauf M, Benzing T, Omran H: **When cilia go bad: cilia defects and ciliopathies**. *Nature Reviews Molecular Cell Biology* 2007, **8**(11):880-893.

90. Waters AM, Beales PL: **Ciliopathies: an expanding disease spectrum**. *Pediatric Nephrology* 2011, **26**(7):1039-1056.

91. Reiter JF, Leroux MR: **Genes and molecular pathways underpinning ciliopathies**. *Nature Reviews Molecular Cell Biology* 2017, **18**(9):533-547.

92. Kozminski KG, Johnson KA, Forscher P, Rosenbaum JL: **A motility in the eukaryotic flagellum unrelated to flagellar beating**. *Proceedings of the National Academy of Sciences* 1993, **90**(12):5519-5523.

93. Pedersen LB, Rosenbaum JL: **Chapter Two Intraflagellar Transport (IFT): Role in Ciliary Assembly, Resorption and Signalling**. In: *Current Topics in Developmental Biology*. vol. 85: Academic Press; 2008: 23-61.

94. Sung C-H, Leroux MR: **The roles of evolutionarily conserved functional modules in cilia-related trafficking**. *Nature Cell Biology* 2013, **15**:1387–1397.

95. Lechtreck KF: **IFT-Cargo Interactions and Protein Transport in Cilia**. *Trends in Biochemical Sciences* 2015, **40**(12):765-778.

96. Prevo B, Scholey JM, Peterman EJG: **Intraflagellar transport: mechanisms of motor action, cooperation, and cargo delivery**. *The FEBS Journal* 2017, **284**(18):2905-2931.

97. Avidor-Reiss T, Gopalakrishnan J: **Cell cycle regulation of the centrosome and cilium**. *Drug Discovery Today: Disease Mechanisms* 2013, **10**:e119-e124.

98. Conduit PT, Wainman A, Raff JW: **Centrosome function and assembly in animal cells**. *Nature Reviews Molecular Cell Biology* 2015, **16**(10):611-624.

99. Walentek P: **Ciliary transcription factors in cancer – how understanding ciliogenesis can promote the detection and prognosis of cancer types**. *The Journal of Pathology* 2016, **239**(1):6-9.

100. Leigh MW, Pittman JE, Carson JL, Ferkol TW, Dell SD, Davis SD, Knowles MR, Zariwala MA: **Clinical and genetic aspects of primary ciliary dyskinesia/Kartagener syndrome**. *Genetics in Medicine* 2009, **11**:473–487.

101. Prosser SL, Morrison CG: **Centrin2 regulates CP110 removal in primary cilium formation**. *The Journal of Cell Biology* 2015, **208**(6):693-701.

102. Singh M, Chaudhry P, Merchant AA: **Primary cilia are present on human blood and bone marrow cells and mediate Hedgehog signaling**. *Experimental Hematology* 2016, **44**(12):1181-1187.

103. Finetti F, Paccani SR, Rosenbaum J, Baldari CT: **Intraflagellar transport: a new player at the immune synapse**. *Trends in Immunology* 2011, **32**(4):139-145.

104. Huangfu D, Liu A, Rakeman AS, Murcia NS, Niswander L, Anderson KV: **Hedgehog signalling in the mouse requires intraflagellar transport proteins**. *Nature* 2003, **426**:83-87.

105. He M, Agbu S, Anderson KV: **Microtubule Motors Drive Hedgehog Signaling in Primary Cilia**. *Trends in Cell Biology* 2017, **27**(2):110-125.

106. Goetz SC, Anderson KV: **The primary cilium: a signalling centre during vertebrate development**. *Nature Reviews Genetics* 2010, **11**(5):331-344.

107. Petrova R, Joyner AL: **Roles for Hedgehog signaling in adult organ homeostasis and repair**. *Development* 2014, **141**(18):3445-3457.

108. Han Y-G, Spassky N, Romaguera-Ros M, Garcia-Verdugo J-M, Aguilar A, Schneider-Maunoury S, Alvarez-Buylla A: **Hedgehog signaling and primary cilia are required for the formation of adult neural stem cells**. *Nature Neuroscience* 2008, **11**:277–284.

109. Satir P, Christensen ST: **Overview of Structure and Function of Mammalian Cilia**. *Annual Review of Physiology* 2007, **69**(1):377-400.

110. Bonnafe E, Touka M, AitLounis A, Baas D, Barras E, Ucla C, Moreau A, Flamant F, Dubruille R, Couble P: **The transcription factor RFX3 directs nodal cilium development and left-right asymmetry specification**. *Molecular and Cellular Biology* 2004, **24**:4417 - 4427.

111. Quigley IK, Kintner C: **Rfx2 Stabilizes Foxj1 Binding at Chromatin Loops to Enable Multiciliated Cell Gene Expression**. *PLOS Genetics* 2017, **13**(1):e1006538.

112. Louvi A, Grove EA: **Cilia in the CNS: The Quiet Organelle Claims Center Stage**. *Neuron* 2011, **69**(6):1046-1060.

113. Lauter G, Swoboda P, Tapia-Páez I: **Chapter 1: Cilia in Brain Development and Disease, pp. 1 - 35.** In: *Cilia: Development and Disease.* Edited by Goggolidou P. Boca Raton, Florida, USA: CRC Press, Taylor and Francis Publishers; 2018.

114. Guemez-Gamboa A, Coufal NG, Glesson JG: **Primary Cilia in the Developing and Mature Brain**. *Neuron* 2014, **82**(3):511-521.

115. Vaisse C, Reiter JF, Berbari NF: **Cilia and Obesity**. *Cold Spring Harbor Perspectives in Biology* 2017, **9**(7):a028217.

116. Marley A, von Zastrow M: **A Simple Cell-Based Assay Reveals That Diverse Neuropsychiatric Risk Genes Converge on Primary Cilia**. *PLOS One* 2012, **7**(10):e46647.

117. Muñoz-Estrada J, Lora-Castellanos A, Meza I, Alarcón Elizalde S, Benítez-King G: **Primary cilia formation is diminished in schizophrenia and bipolar disorder: A possible marker for these psychiatric diseases**. *Schizophrenia Research* 2017, **195**:412-420.

118. Kere J: **The molecular genetics and neurobiology of developmental dyslexia as model of a complex phenotype**. *Biochemical and Biophysical Research Communications* 2014, **452**(2):236-243.

119. Einarsdottir E, Peyrard-Janvid M, Darki F, Tuulari JJ, Merisaari H, Karlsson L, Scheinin NM, Saunavaara J, Parkkola R, Kantojärvi K *et al*: **Identification of *NCAN* as a candidate gene for developmental dyslexia**. *Scientific Reports* 2017, **7**(1):9294.

120. Massinen S, Hokkanen M-E, Matsson H, Tammimies K, Tapia-Páez I, Dahlström-Heuser V, Kuja-Panula J, Burghoorn J, Jeppsson KE, Swoboda P *et al*: **Increased Expression of the Dyslexia Candidate Gene *DCDC2* Affects Length and Signaling of Primary Cilia in Neurons**. *PLOS One* 2011, **6**(6):e20580.

121. Ivliev AE, 't Hoen PAC, van Roon-Mom WMC, Peters DJM, Sergeeva MG: **Exploring the Transcriptome of Ciliated Cells Using In Silico Dissection of Human Tissues**. *PLOS One* 2012, **7**(4):e35618.

122. Hoh RA, Stowe TR, Turk E, Stearns T: **Transcriptional Program of Ciliated Epithelial Cells Reveals New Cilium and Centrosome Components and Links to Human Disease**. *PLOS One* 2013, **7**(12):e52166.

123. Chandrasekar G, Vesterlund L, Hultenby K, Tapia-Páez I, Kere J: **The Zebrafish Orthologue of the Dyslexia Candidate Gene *DYX1C1* Is Essential for Cilia Growth and Function**. *PLOS One* 2013, **8**(5):e63123.

124. Schueler M, Braun DA, Chandrasekar G, Gee HY, Klasson TD, Halbritter J, Bieder A, Porath JD, Airik R, Zhou W *et al*: **DCDC2 Mutations Cause a Renal-Hepatic Ciliopathy by Disrupting Wnt Signaling**. *The American Journal of Human Genetics* 2015, **96**(1):81-92.

125. Grati Mh, Chakchouk I, Ma Q, Bensaid M, Desmidt A, Turki N, Yan D, Baanannou A, Mittal R, Driss N *et al*: **A missense mutation in DCDC2 causes human recessive deafness *DFNB66*, likely by interfering with sensory hair cell and supporting cell cilia length regulation**. *Human Molecular Genetics* 2015, **24**(9):2482-2491.

126. Tarkar A, Loges NT, Slagle CE, Francis R, Dougherty GW, Tamayo JV, Shook B, Cantino M, Schwartz D, Jahnke C *et al*: **DYX1C1 is required for axonemal dynein assembly and ciliary motility**. *Nature Genetics* 2013, **45**:995-1003.

127. Centanni TM, Booker AB, Sloan AM, Chen F, Maher BJ, Carraway RS, Khodaparast N, Rennaker R, LoTurco JJ, Kilgard MP: **Knockdown of the Dyslexia-Associated Gene *Kiaa0319* Impairs Temporal Responses to Speech Stimuli in Rat Primary Auditory Cortex**. *Cerebral Cortex* 2014, **24**(7):1753-1766.

128. Guidi LG, Mattley J, Martinez-Garay I, Monaco AP, Linden JF, Velayos-Baeza A, Molnár Z: **Knockout Mice for Dyslexia Susceptibility Gene Homologs *KIAA0319***

and *KIAA0319L* have Unaffected Neuronal Migration but Display Abnormal Auditory Processing. *Cerebral Cortex* 2017, **27**(12):5831-5845.

129. Franquinho F, Nogueira-Rodrigues J, Duarte JM, Esteves SS, Carter-Su C, Monaco AP, Molnár Z, Velayos-Baeza A, Brites P, Sousa MM: **The Dyslexia-susceptibility Protein KIAA0319 Inhibits Axon Growth Through Smad2 Signaling**. *Cerebral Cortex* 2017, **27**(3):1732-1747.

130. Kistler WS, Baas D, Lemeille S, Paschaki M, Seguin-Estevez Q, Barras E, Ma W, Duteyrat J-L, Morlé L, Durand B *et al*: **RFX2 Is a Major Transcriptional Regulator of Spermiogenesis**. *PLOS Genetics* 2015, **11**(7):e1005368.

131. Ma K, Zheng S, Zuo Z: **The Transcription Factor Regulatory Factor X1 Increases the Expression of Neuronal Glutamate Transporter Type 3**. *Journal of Biological Chemistry* 2006, **281**(30):21250-21255.

132. Hsu Y-C, Liao W-C, Kao C-Y, Chiu I-M: **Regulation of FGF1 Gene Promoter through Transcription Factor RFX1**. *Journal of Biological Chemistry* 2010, **285**(18):13885-13895.

133. Katan Y, Agami R, Shaul Y: **The transcriptional activation and repression domains of RFX1, a context-dependent regulator, can mutually neutralize their activities**. *Nucleic Acids Research* 1997, **25**(18):3621-3628.

134. Blackshear PJ, Graves JP, Stumpo DJ, Cobos I, Rubenstein JL, Zeldin DC: **Graded phenotypic response to partial and complete deficiency of a brain-specific transcript variant of the winged helix transcription factor RFX4**. *Development* 2003, **130**:4539 - 4552.

135. Sedykh I, Keller AN, Yoon B, Roberson L, Moskvin OV, Grinblat Y: **Zebrafish Rfx4 controls dorsal and ventral midline formation in the neural tube**. *Developmental Dynamics* 2018, **247**(4):650-659.

136. Wolfe SA, vanWert JM, Grimes SR: **Transcription factor RFX4 binding to the testis-specific histone H1t promoter in spermatocytes may be important for regulation of H1t gene transcription during spermatogenesis**. *Journal of Cellular Biochemistry* 2008, **105**(1):61-69.

137. Kistler WS, Horvath GC, Dasgupta A, Kistler MK: **Differential expression of Rfx1-4 during mouse spermatogenesis**. *Gene Expression Patterns* 2009, **9**(7):515-519.

138. Feng C, Li J, Zuo Z: **Expression of the transcription factor regulatory factor X1 in mouse brain**. *Folia histochemica et cytobiologica / Polish Academy of Sciences, Polish Histochemical and Cytochemical Society* 2011, **49**(2):344-351.

139. Spassky N, Merkle FT, Flames N, Tramontin AD, García-Verdugo JM, Alvarez-Buylla A: **Adult Ependymal Cells Are Postmitotic and Are Derived from Radial Glial Cells during Embryogenesis**. *The Journal of Neuroscience* 2005, **25**(1):10-18.

140. Shawlot W, Vazquez-Chantada M, Wallingford JB, Finnell RH: **Rfx2 is required for spermatogenesis in the mouse**. *Genesis: The Journal of Genetics and Development* 2015, **53**(9):604-611.

141. La Manno G, Gyllborg D, Codeluppi S, Nishimura K, Salto C, Zeisel A, Borm LE, Stott SRW, Toledo EM, Villaescusa JC *et al*: **Molecular Diversity of Midbrain Development in Mouse, Human, and Stem Cells**. *Cell* 2016, **167**(2):566-580.

142. Kawase S, Kuwako K, Imai T, Renault-Mihara F, Yaguchi K, Itohara S, Okano H: **Regulatory Factor X Transcription Factors Control Musashi1 Transcription in Mouse Neural Stem/Progenitor Cells**. *Stem Cells and Development* 2014, **23**(18):2250-2261.

143. Ait-Lounis A, Baas D, Barras E, Benadiba C, Charollais A, Nlend Nlend R, Liegeois D, Meda P, Durand B, Reith W: **Novel function of the ciliogenic transcription factor RFX3 in development of the endocrine pancreas**. *Diabetes* 2007, **56**:950-959.

144. Benadiba C, Magnani D, Niquille M, Morlé L, Valloton D, Nawabi H, Ait-Lounis A, Otsmane B, Reith W, Theil T *et al*: **The Ciliogenic Transcription Factor RFX3 Regulates Early Midline Distribution of Guidepost Neurons Required for Corpus Callosum Development**. *PLOS Genetics* 2012, **8**(3):e1002606.

145. Magnani D, Morle L, Hasenpusch-Theil K, Paschaki M, Jacoby M, Schurmans S, Durand B, Theil T: **The ciliogenic transcription factor Rfx3 is required for the formation of the thalamocortical tract by regulating the patterning of prethalamus and ventral telencephalon**. *Human Molecular Genetics* 2015, **24**(9):2578-2593.

146. Zhang D, Stumpo DJ, Graves JP, DeGraff LM, Grissom SF, Collins JB, Li L, Zeldin DC, Blackshear PJ: **Identification of potential target genes for RFX4_v3, a transcription factor critical for brain development**. *Journal of Neurochemistry* 2006, **98**(3):860-875.

147. Clausen BE, Waldburger J-M, Schwenk F, Barras E, Mach B, Rajewsky K, Förster I, Reith W: **Residual MHC Class II Expression on Mature Dendritic Cells and Activated B Cells in RFX5-Deficient Mice**. *Immunity* 1998, **8**(2):143-155.

148. Perkins LA, Hedgecock EM, Thomson JN, Culotti JG: **Mutant sensory cilia in the nematode *Caenorhabditis elegans***. *Developmental Biology* 1986, **117**(2):456-487.

149. Wang J, Schwartz HT, Barr MM: **Functional Specialization of Sensory Cilia by an RFX Transcription Factor Isoform**. *Genetics* 2010, **186**(4):1295-1307.

150. Craig HL, Wirtz J, Bamps S, Dolphin CT, Hope IA: **The significance of alternative transcripts for *Caenorhabditis elegans* transcription factor genes, based on expression pattern analysis**. *BMC Genomics* 2013, **14**(1):249.

151. Senti G, Swoboda P: **Distinct Isoforms of the RFX Transcription Factor DAF-19 Regulate Ciliogenesis and Maintenance of Synaptic Activity**. *Molecular Biology of the Cell* 2008, **19**(12):5517-5528.

152. Wells KL, Rowneki M, Killian DJ: **A splice acceptor mutation in *C. elegans* daf-19/Rfx disrupts functional specialization of male-specific ciliated neurons but does not affect ciliogenesis**. *Gene* 2015, **559**(2):196-202.

153. Allen MA, Hillier LW, Waterston RH, Blumenthal T: **A global analysis of *C. elegans* trans-splicing**. *Genome Research* 2011, **21**(2):255-264.

154. Bae Y-K, Barr MM: **Sensory roles of neuronal cilia: Cilia development, morphogenesis, and function in *C. elegans***. *Frontiers in bioscience : a journal and virtual library* 2008, **13**:5959-5974.

155. White J. G. , Southgate E., Thomson J. N., S. B: **The structure of the nervous system of the nematode *Caenorhabditis elegans***. *Philosophical Transactions of the Royal Society of London B, Biological Sciences* 1986, **314**(1165):1-340.

156. Jensen VL, Carter S, Sanders AAWM, Li C, Kennedy J, Timbers TA, Cai J, Scheidel N, Kennedy BN, Morin RD *et al*: **Whole-Organism Developmental Expression Profiling Identifies RAB-28 as a Novel Ciliary GTPase Associated with the BBSome and Intraflagellar Transport**. *PLOS Genetics* 2016, **12**(12):e1006469.

157. Xie Y, Moussaif M, Choi S, Xu L, Sze JY: **RFX Transcription Factor DAF-19 Regulates 5-HT and Innate Immune Responses to Pathogenic Bacteria in *Caenorhabditis elegans***. *PLOS Genetics* 2013, **9**(3):e1003324.

158. Senti G, Ezcurra M, Löbner J, Schafer WR, Swoboda P: **Worms With a Single Functional Sensory Cilium Generate Proper Neuron-Specific Behavioral Output**. *Genetics* 2009, **183**(2):595-605.

159. Mahoney TR, Luo S, Nonet ML: **Analysis of synaptic transmission in *Caenorhabditis elegans* using an aldicarb-sensitivity assay**. *Nature Protocols* 2006, **1**:1772–1777.

160. George DT, Behm CA, Hall DH, Mathesius U, Rug M, Nguyen KCQ, Verma NK: ***Shigella flexneri* Infection in *Caenorhabditis elegans*: Cytopathological Examination and Identification of Host Responses**. *PLOS One* 2014, **9**(9):e106085.

161. Newton FG, zur Lage PI, Karak S, Moore DJ, Göpfert MC, Jarman AP: **Forkhead Transcription Factor Fd3F Cooperates with Rfx to Regulate a Gene Expression Program for Mechanosensory Cilia Specialization**. *Developmental Cell* 2012, **22**(6):1221-1233.

162. Baur F, Nau K, Sadic D, Allweiss L, Elsässer H-P, Gillemans N, de Wit T, Krüger I, Vollmer M, Philipsen S *et al*: **Specificity Protein 2 (Sp2) Is Essential for Mouse Development and Autonomous Proliferation of Mouse Embryonic Fibroblasts**. *PLOS One* 2010, **5**(3):e9587.

163. Liang H, Xiao G, Yin H, Hippenmeyer S, Horowitz JM, Ghashghaei HT: **Neural development is dependent on the function of specificity protein 2 in cell cycle progression**. *Development* 2013, **140**(3):552-561.

164. Chen H-Z, Tsai S-Y, Leone G: **Emerging roles of E2Fs in cancer: an exit from cell cycle control**. *Nature Reviews Cancer* 2009, **9**(11):785-797.

165. Ma L, Quigley I, Omran H, Kintner C: **Multicilin drives centriole biogenesis via E2f proteins**. *Genes & Development* 2014, **28**(13):1461-1471.

166. Pankratz N, Wilk JB, Latourelle JC, DeStefano AL, Halter C, Pugh EW, Doheny KF, Gusella JF, Nichols WC, Foroud T *et al*: **Genomewide association study for susceptibility genes contributing to familial Parkinson disease**. *Human Genetics* 2009, **124**(6):593-605.

167. Nagle MW, Latourelle JC, Labadorf A, Dumitriu A, Hadzi TC, Beach TG, Myers RH: **The 4p16.3 Parkinson Disease Risk Locus Is Associated with *GAK* Expression and Genes Involved with the Synaptic Vesicle Membrane**. *PLOS One* 2016, **11**(8):e0160925.

168. Yim Y-I, Sun T, Wu L-G, Raimondi A, De Camilli P, Eisenberg E, Greene LE: **Endocytosis and clathrin-uncoating defects at synapses of auxilin knockout mice**. *Proceedings of the National Academy of Sciences* 2010, **107**(9):4412-4417.

169. Park B-C, Yim Y-I, Zhao X, Olszewski MB, Eisenberg E, Greene LE: **The clathrin-binding and J-domains of GAK support the uncoating and chaperoning of clathrin by Hsc70 in the brain**. *Journal of Cell Science* 2015, **128**(20):3811-3821.

170. FANTOM Consortium, RIKEN PMI, CLST (DGT): **A promoter-level mammalian expression atlas**. *Nature* 2014, **507**(7493):462-470.

171. Kodzius R, Kojima M, Nishiyori H, Nakamura M, Fukuda S, Tagami M, Sasaki D, Imamura K, Kai C, Harbers M *et al*: **CAGE: cap analysis of gene expression**. *Nature Methods* 2006, **3**(3):211-222.

172. Takahashi H, Lassmann T, Murata M, Carninci P: **5' end-centered expression profiling using cap-analysis gene expression and next-generation sequencing**. *Nature Protocols* 2012, **7**(3):542-561.

173. Andersson R, Gebhard C, Miguel-Escalada I, Hoof I, Bornholdt J, Boyd M, Chen Y, Zhao X, Schmidl C, Suzuki T *et al*: **An atlas of active enhancers across human cell types and tissues**. *Nature* 2014, **507**(7493):455-461.

174. de Rie D, Abugessaisa I, Alam T, Arner E, Arner P, Ashoor H, Åström G, Babina M, Bertin N, Burroughs AM *et al*: **An integrated expression atlas of miRNAs and their promoters in human and mouse**. *Nature Biotechnology* 2017, **35**(9):872–878.

175. Hon C-C, Ramilowski JA, Harshbarger J, Bertin N, Rackham OJL, Gough J, Denisenko E, Schmeier S, Poulsen TM, Severin J *et al*: **An atlas of human long non-coding RNAs with accurate 5′ ends**. *Nature* 2017, **543**(7644):199-204.

176. Wasserman WW, Sandelin A: **Applied bioinformatics for the identification of regulatory elements**. *Nature Reviews Genetics* 2004, **5**(4):276-287.

177. Stormo GD: **DNA binding sites: representation and discovery**. *Bioinformatics* 2000, **16**(1):16-23.

178. Sandelin A, Alkema W, Engström P, Wasserman WW, Lenhard B: **JASPAR: an open-access database for eukaryotic transcription factor binding profiles**. *Nucleic Acids Research* 2004, **32**(D1):D91-D94.

179. Khan A, Fornes O, Stigliani A, Gheorghe M, Castro-Mondragon JA, van der Lee R, Bessy A, Chèneby J, Kulkarni SR, Tan G *et al*: **JASPAR 2018: update of the open-access database of transcription factor binding profiles and its web framework**. *Nucleic Acids Research* 2018, **46**(D1):D260-D266.

180. Mathelier A, Fornes O, Arenillas DJ, Chen C-y, Denay G, Lee J, Shi W, Shyr C, Tan G, Worsley-Hunt R *et al*: **JASPAR 2016: a major expansion and update of the open-access database of transcription factor binding profiles**. *Nucleic Acids Research* 2016, **44**(D1):D110-D115.

181. Lukk M, Kapushesky M, Nikkilä J, Parkinson H, Goncalves A, Huber W, Ukkonen E, Brazma A: **A global map of human gene expression**. *Nature Biotechnology* 2010, **28**:322–324.

182. Pitaval A, Tseng Q, Bornens M, Théry M: **Cell shape and contractility regulate ciliogenesis in cell cycle–arrested cells**. *The Journal of Cell Biology* 2010, **191**(2):303-312.

183. Tuxhorn J, Daise T, Dentler WL: **Regulation of flagellar length in *Chlamydomonas***. *Cell Motility* 1998, **40**(2):133-146.

184. Xu S: **The Application of CRISPR-Cas9 Genome Editing in *Caenorhabditis elegans***. *Journal of Genetics and Genomics* 2015, **42**(8):413-421.

185. Fielenbach N, Antebi A: ***C. elegans* dauer formation and the molecular basis of plasticity**. *Genes & Development* 2008, **22**(16):2149-2165.

186. Spassky N: **Motile Cilia and Brain Function: Ependymal Motile Cilia Development, Organization, Function and Their Associated Pathologies**. In: *Cilia and Nervous System Development and Function.* Edited by Tucker KL, Caspary T. Dordrecht: Springer Netherlands; 2013: 193-207.

187. Youn YH, Han Y-G: **Primary Cilia in Brain Development and Diseases**. *The American Journal of Pathology* 2018, **188**(1):11-22.

188. Berbari NF, Malarkey EB, Yazdi SMZR, McNair AD, Kippe JM, Croyle MJ, Kraft TW, Yoder BK: **Hippocampal and Cortical Primary Cilia Are Required for Aversive Memory in Mice**. *PLOS One* 2014, **9**(9):e106576.

189. Siljee JE, Wang Y, Bernard AA, Ersoy BA, Zhang S, Marley A, Von Zastrow M, Reiter JF, Vaisse C: **Subcellular localization of MC4R with ADCY3 at neuronal primary cilia underlies a common pathway for genetic predisposition to obesity**. *Nature Genetics* 2018, **50**(2):180-185.

190. Grarup N, Moltke I, Andersen MK, Dalby M, Vitting-Seerup K, Kern T, Mahendran Y, Jørsboe E, Larsen CVL, Dahl-Petersen IK *et al*: **Loss-of-function variants in ADCY3 increase risk of obesity and type 2 diabetes**. *Nature Genetics* 2018, **50**(2):172-174.

191. Bishop GA, Berbari NF, Lewis J, Mykytyn K: **Type III adenylyl cyclase localizes to primary cilia throughout the adult mouse brain**. *Journal of Comparative Neurology* 2007, **505**(5):562-571.

192. Green JA, Mykytyn K: **Neuronal Primary Cilia: An Underappreciated Signaling and Sensory Organelle in the Brain**. *Neuropsychopharmacology* 2014, **39**(1):244-245.

193. Kasahara K, Miyoshi K, Murakami S, Miyazaki I, Asanuma M: **Visualization of Astrocytic Primary Cilia in the Mouse Brain by Immunofluorescent Analysis Using the Cilia Marker Arl13b**. *Acta Medica Okayama* 2014, **68**(6):317-322.

194. Guo J, Otis JM, Higginbotham H, Monckton C, Cheng J, Asokan A, Mykytyn K, Caspary T, Stuber GD, Anton ES: **Primary Cilia Signaling Shapes the Development of Interneuronal Connectivity**. *Developmental Cell* 2017, **42**(3):286-300.

195. Kumamoto N, Gu Y, Wang J, Janoschka S, Takemaru K-I, Levine J, Ge S: **A role for primary cilia in glutamatergic synaptic integration of adult-born neurons**. *Nature Neuroscience* 2012, **15**(3):399-405.

196. Guo J, Higginbotham H, Li J, Nichols J, Hirt J, Ghukasyan V, Anton ES: **Developmental disruptions underlying brain abnormalities in ciliopathies**. *Nature Communications* 2015, **6**:7857.

197. Sarkisian MR, Guadiana SM: **Influences of Primary Cilia on Cortical Morphogenesis and Neuronal Subtype Maturation**. *The Neuroscientist* 2015, **21**(2):136-151.

198. Paredes MF, James D, Gil-Perotin S, Kim H, Cotter JA, Ng C, Sandoval K, Rowitch DH, Xu D, McQuillen PS *et al*: **Extensive migration of young neurons into the infant human frontal lobe**. *Science* 2016, **354**(6308):aaf7073.

199. O'Neill RS, Schoborg TA, Rusan NM: **Same but different: pleiotropy in centrosome-related microcephaly**. *Molecular Biology of the Cell* 2018, **29**(3):241-246.

200. Arquint C, Gabryjonczyk A-M, Nigg EA: **Centrosomes as signalling centres**. *Philosophical Transactions of the Royal Society B: Biological Sciences* 2014, **369**(1650):20130464.

201. Nigg EA, Raff JW: **Centrioles, Centrosomes, and Cilia in Health and Disease**. *Cell* 2009, **139**(4):663-678.

202. Higginbotham H, Eom T-Y, Mariani LE, Bachleda A, Hirt J, Gukassyan V, Cusack CL, Lai C, Caspary T, Anton ES: **Arl13b in Primary Cilia Regulates the Migration and Placement of Interneurons in the Developing Cerebral Cortex**. *Developmental Cell* 2012, **23**(5):925-938.

203. Baudoin J-P, Viou L, Launay P-S, Luccardini C, Espeso Gil S, Kiyasova V, Irinopoulou T, Alvarez C, Rio J-P, Boudier T *et al*: **Tangentially Migrating Neurons Assemble a Primary Cilium that Promotes Their Reorientation to the Cortical Plate**. *Neuron* 2012, **76**(6):1108-1122.

204. Schafer DP, Stevens B: **Phagocytic glial cells: sculpting synaptic circuits in the developing nervous system**. *Current Opinion in Neurobiology* 2013, **23**(6):1034-1040.

205. Chung W-S, Clarke LE, Wang GX, Stafford BK, Sher A, Chakraborty C, Joung J, Foo LC, Thompson A, Chen C *et al*: **Astrocytes mediate synapse elimination through MEGF10 and MERTK pathways**. *Nature* 2013, **504**:394–400.