Singing is silver, hearing is gold: impacts of local FoxP1 knockdowns on auditory perception and gene expression in female zebra finches

Heim, F.D.

Chapter 4

Transcriptomic investigations of age- and region-specific knockdowns in female zebra finches identify potential downstream networks of FoxP1

Chapter 4: Transcriptomic investigations of age- and region-specific knockdowns in female zebra finches identify potential downstream networks of FoxP1

Fabian Heim, Katharina Riebel, Carel ten Cate, Constance Scharff, Simon E. Fisher

Abstract

FOXP1 is a highly conserved transcription factor that regulates the expression of target genes in diverse species. Evidence from multiple sources suggests that FOXP1 is important for aspects of brain development and function. For example, humans with rare heterozygous disruptions of *FOXP1* have been diagnosed with intellectual disability and/or autism spectrum disorder, as well as speech and language deficits. The avian ortholog, *FoxP1*, is highly expressed in a subset of song-related nuclei in the brains of songbirds, and prior studies have employed experimental knockdowns of this gene in Area X, HVC and CMM of male or female zebra finches, to investigate potential links to behaviour. In particular, in the work described in earlier parts of this thesis, female zebra finches were injected with lentiviral knockdown constructs in HVC or CMM during two different developmental stages. The present chapter sought to use tissue samples from the targeted brain areas in these birds to identify putative molecular targets and pathways that lie downstream of FoxP1, via a transcriptomic approach based on next-generation RNA-sequencing. Differentially expressed genes between control and knockdown groups were analysed across the different brain areas and the different developmental stages of genetic manipulations. With the exception of *FoxP1* itself, no individual gene showed significant differences in expression in all groups of this study. Nonetheless, data from the different groups on differentially expressed genes, enriched GO terms, gene sets and local networks together point to possible links of *FoxP1* to retinoic acid signaling or SLIT-ROBO pathways, among others. Moreover, differentially expressed genes associated with FoxP1 knockdown showed an overrepresentation of candidate loci involved in autism spectrum disorder and intellectual disability, based on analyses of independent databases that collated likely risk genes. The expression profiling data from this study can offer new insights into neurogenetic networks that may be regulated by FoxP1, suggesting hypotheses for future investigation in a range of species and model systems.

Introduction

Studies in multiple species indicate roles of the FOXP1 transcription factor in aspects of brain development and function. In humans, rare variants that disrupt FOXP1 result in a neurodevelopmental syndrome involving a range of features including intellectual disability, autism spectrum disorder, and impairments in speech and language (Sollis *et al.*, 2016; Siper *et al.*, 2017). Orthologues of *FOXP*1 have been identified in highly similar form in many different vertebrate and invertebrate species where they are thought to regulate the expression of downstream target genes in the brain and other tissue (Mazet *et al.*, 2003; Haesler *et al.*, 2004; Teramitsu *et al.*, 2004; Hannenhalli and Kaestner, 2009; Lawton *et al.*, 2014; Viscardi *et al.*, 2017). The high degree of homology has made it possible to investigate implications of FOXP1 dysfunction in animal models (Takahashi *et al.*, 2009; Scharff and Petri, 2011; Deriziotis and Fisher, 2017; Co *et al.*, 2020a), often with a focus on consequence for vocal behaviours (Fröhlich *et al.*, 2017; Norton *et al.*, 2019). For example, heterozygous deletions of *Foxp1* disrupt mouse vocalisations (Araujo *et al.*, 2015), and forebrain-specific knockouts of the gene result in perturbed isolation calls of mouse pups (Usui *et al.*, 2017a). Consistent with these findings, other studies have shown that brain-wide homozygous deletions of mouse *Foxp1* reduce social interactions (Bacon *et al.*, 2015) and the rate of pup isolation calls upon removal of the mother (Fröhlich *et al.*, 2017).

As described in earlier Chapters of this thesis, and other recent studies, the contributions of FoxP1 to vocal behaviours have also been investigated in the zebra finch (*Taeniopygia guttata*), a songbird in which males learn their vocalisations by imitation of a tutor (Immelmann, 1962; Zann, 1997; Tchernichovski *et al.*, 2001)(Doupe and Kuhl, 1999; Scharff and Petri, 2011; Bruno *et al.*, 2021). It is of particular interest for this thesis that *FoxP1* is expressed in distinct nuclei in the brains of songbirds, including those known to be important for vocal learning, such as Area X in the striatum. Compared to the surrounding tissue, *FoxP1* expression is also elevated in the robust nucleus of the arcopallium (RA, a motor nucleus), the premotor area HVC, and the entire mesopallium, including the caudomedial mesopallium (CMM), which is a secondary auditory area (Haesler *et al.*, 2004; Teramitsu *et al.*, 2004; Mendoza *et al.*, 2015). With the exception of RA, *FoxP1* seems to be highly expressed in brain areas associated with tasks related to auditory perception and feedback. Notably, and despite the absence of Area X and a negligible state of RA in female zebra finches

(Nottebohm and Arnold, 1976), neural *FoxP1* expression patterns are highly similar between sexes of this species (Teramitsu *et al.*, 2004).

Genetic manipulations that reduce expression levels of *FoxP1* in the basal ganglia Area X of juvenile male zebra finches lead to impaired song learning, as do knockdowns of paralogues FoxP2 and FoxP4, albeit with differences in the nature of impairments (Haesler *et al.*, 2007; Murugan *et al.*, 2013; Norton *et al.*, 2019). FoxP2 overexpression in Area X of adult zebra finch males also alters song by increasing its variability (Day *et al.*, 2019a).

Knockdown of *FoxP1* expression levels in HVC of male juvenile zebra finches has also been shown to impair song learning, but only if the knockdown occurs prior to song exposure (Garcia-Oscos *et al.*, 2021). The experiments described in prior Chapters of this thesis showed that *FoxP1* knockdown in HVC of adult female zebra finches may disrupt rewarding qualities of conspecific song (Chapter 2), while the same manipulations made during earlier developmental stages, prior to song preference establishment, do not affect preference strength in adults or preference establishment for familiar song (Chapter 3). *FoxP1* knockdown in CMM of juvenile or adult females did not affect the birds' ability to establish or maintain a song preference (Chapter 2). Knockdowns of *FoxP1* in HVC or CMM in juvenile females prior to the onset of the sensory phase, or in adults well after closure of the vocal learning period, did not alter the birds' ability to discriminate different conspecific song stimuli or categorise altered versions of them (Chapter 3). These studies characterise behavioural consequences of manipulating *FoxP1* expression in particular brain structures, but they do not give information about the neurogenetic pathways that are regulated by the transcription factor. To gain insights at that level, it is necessary to integrate the employed knockdown strategy with a molecular screening technique, which is the subject of the current Chapter.

As a transcription factor, FoxP1 acts by forming homo- or heterodimers and multimers with other FoxP molecules (Li *et al.*, 2004; Sin *et al.*, 2014; Castells-Nobau *et al.*, 2019). Together, these complexes bind to DNA and modify expression levels of other genes – its downstream targets. Identification of target genes and their respective functions can provide insight into the molecular and cellular pathways that a transcription factor regulates. For transcription factors, like FoxP1, that have highly similar orthologues in multiple species (Hannenhalli and Kaestner, 2009), model systems can be used to

facilitate the identification of target genes. Depending on the underlying motivation of a study, experiments employ model systems ranging from cell-culture to analyses of tissue obtained from genetically modified animals, such as knockout mice, or zebra finches which underwent localised knockdowns.

To identify genes and pathways which are regulated by transcription factors such as FoxP1, multiple alternative molecular approaches are available. Chromatin immunoprecipitation (ChIP) methods make it possible to define the genomic interaction sites of a DNA-binding protein (Buck and Lieb, 2004; Park, 2009). These techniques involve cross-linking of DNA-binding molecules to the DNA, followed by immunoprecipitation of linked protein-DNA complexes with antibodies that specifically recognize the protein of interest. The precipitated complexes are then treated to remove the crosslinks, and the extracted DNA is analysed, for example by screening with arrays (ChIP-chip, Buck and Lieb, 2004) or via sequencing (ChIP-seq, Furey, 2012), to identify which genomic regions are enriched in the immunoprecipitated samples. The process yields knowledge of the genomic binding sites of the protein, and this information can be used to determine the identities of candidate target genes that it may regulate. However, the success of ChIP-based assays depends critically on the reliability of the antibody used for immunoprecipitation, and levels of enrichment can be subtle, making it difficult to identify differences between datasets. Additionally, the target epitope of the antibodies employed may be blocked by additional proteins.

For studies in which genetic manipulations are used to alter the levels of an important regulatory molecule, comparison of expression profiles represents a valuable strategy for characterizing downstream pathways. While ChIP uncovers primary targets of a transcription factor, expression profiling can identify both direct and indirectly regulated targets. These methods do not rely on binding sites of one specific protein but consider changes throughout the transcriptome (Pollack *et al.*, 1999). RNA is extracted, reverse transcribed, fragmented, amplified and can then be analysed with a number of techniques. In microarray-based expression profiling, the amplified fragments are tagged and applied onto a chip which contains complementary fragments based on the transcriptome of the species being studied. Tagged fragments binding their complementary strands on the chip are used to characterize the levels of expression of genes in the sample. However, due to the necessary preselection of complementary probe fragments, microarray analyses may be biased and are not optimal for covering

the entire transcriptome. These limitations have been largely overcome with the application of next-generation sequencing techniques to expression profiling (RNA-sequencing) allowing for quantification of transcript levels across the transcriptome in a less biased manner than arrays (Cloonan *et al.*, 2008; Wang *et al.*, 2009).

Genome wide changes due to knockdowns or other genetic manipulations could also be assessed by the identification of chromatin state changes. Chromosome conformation changes can provide insight about the status of DNA within the chromatin complex and possible differences between samples (Schmitt *et al.*, 2016). Based on the accessibility of a genomic region, putative target genes and genomic regions with increased or decreased accessibility due to a gene specific knockdown (for example) can be deduced. However, this approach requires relatively large amounts of DNA from the target tissue.

In mice, shRNA mediated *Foxp1* knockdown has been found to yield changes in the expression of genes with GO terms associated to neurogenesis, regulation of synapse organisation and nervous system development, in addition to the Notch signalling pathway which in turn might contribute to impaired differentiation of neural stem cells to astrocytes and neurons *in utero* and *in vivo* (Braccioli *et al.*, 2017). In mice with heterozygous knockout of *Foxp1* in medium spiny neurons, the development of subtype composition of these neurons is altered as determined by single cell RNAseq. *Foxp1* knockout reduces the occurrence of indirect pathway spiny neurons when compared to wildtype controls, possibly due to differential regulation of genes specific to spiny neuron subtypes. Additionally, *Foxp1* knockouts result in differential regulation of genes associated to autism spectrum disorders (ASD, Anderson *et al.*, 2020). Global *Foxp1* heterozygous knockout mice show significant overlap of differentially expressed genes to human neural progenitor cells (NPC) overexpressing *FOXP1*. When compared to gene expression in striatal or hippocampal tissue from mice with global heterozygous *Foxp1* knockouts, *FOXP1* overexpressing human NPCs show a larger overlap in gene expression with striatal tissue than with hippocampus as shown by module preservation in weighted gene coexpression network analyses. Gene expression in hippocampus of mice with global heterozygous *Foxp1* knockouts also highlights pathways linked to long term potentiation, synaptic signalling and spatial memory which are all relevant for learning (Araujo *et al.*, 2015). These findings are interesting given that global heterozygous *Foxp1* knockout mice demonstrated poor

learning during Morris water maze trials, less successful performance on T-Maze tasks, and reduced maintenance of long term potentiation, assessed via slice electrophysiology (Araujo *et al.*, 2017).

In this Chapter, I took advantage of the availability of zebra finches with *FoxP1* knockdowns in selective brain regions and distinct developmental stages, as generated in my prior thesis work, to help identify *in vivo* networks downstream of this transcription factor. I employed RNA sequencing analyses (RNAseq), since that made it possible to analyse expression levels of distinct genes and pathways directly with relatively low amounts of material at a larger dynamic range, covering the entire transcriptome and with improved detection of weakly expressed genes (Wang *et al.*, 2009; Zhao *et al.*, 2014). In the process of RNAseq, total RNA from different samples is extracted and reverse transcribed into cDNA which is further fragmented and purified prior to library generation. Following amplification of these libraries, the fragmented cDNA strands are sequenced and aligned to a reference genome. Based on the alignment, each fragment can be assigned to a coding or non-coding region of the target organism's genome. The number of assignments can then be used to determine and compare gene expression levels based on the number of fragments assigned to a specific region (= counts).

In this study, total RNA from the targeted brain areas of the different groups was extracted. Genes and pathways associated with knockdowns of *FoxP1* in general were investigated with the aim to identify affected genes and molecular pathways in the brain in general, but also how these genes and pathways are differentially affected in the context of age or developmental status (juvenile and adult groups) and local brain areas (HVC and CMM). Following the identification of unique and overlapping genes and pathways, the findings were compared to previous mouse studies focusing on downstream targets of Foxp1, and potential physiological and behavioural consequences that have been related to dysfunction of this important transcription factor.

Material and Methods

Test subjects

Subjects were 96 female zebra finches from the breeding colony at the Freie Universität Berlin. At Leiden University, birds were housed in groups until behavioural testing was started (Chapters 2 and 3 of this thesis). The four treatment groups were defined by when (as juveniles < 25 days post hatch (dph) or adults > 90 dph) and where (HVC or CMM) they received the *FoxP1* knockdown and labelled accordingly: HVC juvenile, HVC adult, CMM juvenile, CMM adult (for details see Method sections of Chapters 2 and 3). Each corresponding knockdown and control group consisted of 12 females.

Viral particles and injection

Viral particles were produced at the Freie Universität Berlin as described in Chapters 2 and 3. Birds were injected with one of two shRNA constructs complementary to *FoxP1* mRNA to knock down *FoxP1* expression. Both knockdown constructs also contained a GFP sequence to label successfully transduced cells. The two different knockdown constructs were employed to obtain an opportunity to filter for putative off-target effects induced by either one of the shRNAs (Song *et al.*, 2015). A similar construct which contained the sequence for GFP but no *FoxP1* targeting shRNA was used for control animals. The sequence of the two short-hairpin constructs was as follows (Norton *et al.*, 2019):

	Hairpin sequence
Construct 1	5'-CCCCTATGCAAGCAATGCACCCAGTGCATGTCAAAGAAGAACCATTAGACCCAGATGAAA-3'
Construct 2	5'-CCAGATGAAAATGAAGGCCCACTATCCTTAGTGACAACAGCCAACCACAG-3'

Viral particles were produced in seven batches for each knockdown construct, and five batches of control virus. Each virus batch was injected into both hemispheres of on average 4 birds (range 2 – 6). This corresponds to on average 6 different batches per treatment group (range 3 – 9) including matched controls. By merging samples from birds which received injections from different batches into larger control or knockdown groups, it was possible to control for batch-specific effects due to differences in titre or transduction-efficiency (for details see extended data of Chapter 2, Table 2-1). The injection procedure is described in detail in Chapters 2 and 3.

Briefly, viral constructs were injected bilaterally in one of the two target areas of a juvenile or adult female (see Chapter 2 Table 1 for injection coordinates in reference to the bifurcation of the midsagittal sinus). The injection site was closed with previously removed bone tissue, and the skin was sealed. After the surgical procedure, the birds were returned to their respective housing cages.

Brain extraction

After completion of behavioural experiments (preference tests and Go/Nogo tasks, described in Chapters 2 and 3 respectively), females were housed in their home cages with other familiar females for at least one week. Between 3 – 5 pm on the day before brain extraction, birds were individually transferred into familiar sound attenuated chambers used during the prior behavioural tests. In order to minimise activity-dependent expression changes, birds were sacrificed with an overdose of isoflurane gas before light onset on the next morning (6:30 – 6:50 AM). Birds of the juvenile groups were 179 – 210 days old and those of the adult groups 165 – 579 days old, respectively. Note that juvenile and adult refers to the developmental stage the birds received lentiviral injections while all behavioural experiments and subsequent tissue extractions where conducted in adult birds. Fresh hemispheres were separated along the midline and frozen in Tissue Tek Optimal Cutting Temperature Compound (OCT, Sakura, Leiden) on dry ice and stored at -80°C at the Language and Genetics Department at the Max Planck Institute in Nijmegen, the Netherlands.

Validation of injected area and extracted tissue

The injection site was validated immunohistochemically (Chapter 2) by staining with antibodies against FoxP1 and GFP, and counterstaining of nuclei with Hoechst (Thermo Fisher Scientific, Waltham USA). The target areas, HVC or CMM were extracted with biopsy punches of frozen brain slices. Correct placement of the biopsy punch site in HVC or CMM was validated visually under a stereomicroscope, and GFP-based fluorescence was documented in the extracted tissue punches.

RNA extraction

GFP-positive biopsy punches were submerged in RNAlater (Qiagen, Hilden) and pooled by hemisphere for each bird. At least 12hrs after punching, a column-based

RNA extraction kit was used to extract and purify total RNA according to the manufacturer's protocol (RNeasy micro plus, Qiagen, Hilden). RNA quality and concentration were determined with a Bioanalyser RNA kit (Biorad, Hercules). Extracted RNA was stored at -80°C until transcriptome sequencing.

Total RNA sequencing

RNA sequencing was performed in three batches. The first batch contained only samples from adult HVC knockdowns and their respective controls. The second batch included adult HVC and adult CMM knockdown and control samples, while the last batch consisted of knockdown and control samples from all targeted areas and ages (see Supplementary Table 1). Only a subset of samples from all birds (N = 104/192, 54%) fulfilled the necessary quality criteria for sequencing (> 0.4 µg total RNA, RNA integrity index > 7.2). Drop-out rate was distributed evenly among groups resulting in 5 to 9 samples per group (see Supplement). These drop-outs can be partly attributed to cases where tissue punches did not show fluorescence under the stereomicroscope or misplaced punching sites which led to fewer biopsy punches for RNA extraction. In a recent study on gene expression differences in brain nuclei of different birds, unrelated to the present work, no samples had to be dropped. However, because that study was limited to microarray analyses, the experiments only required one fourth of the RNA amount necessary for total RNAseq analyses (Ko *et al.*, 2021). As punched tissue samples in the current study were further preselected based on the correct site of the biopsy punch and presence of fluorescence, it is not possible to make meaningful comparisons of RNA yields to those in prior work. Library preparation and sequencing was performed by Novogene Co., Ltd. (Beijing). After enrichment with oligo(dT) beads and random fragmentation, libraries were constructed with 150 – 200 base-pair (bp) inserts. cDNA was synthesised using random hexamers and reverse transcriptase. The second strand was completed by nick-translation with a custom second-strand synthesis buffer provided by Illumina (San Diego, USA) containing dNTPs, RNAse H and *Escherichia coli* polymerase I. cDNA libraries then underwent purification, terminal repair, A-tailing, ligation of sequencing adapters, size selection and PCR enrichment. Illumina HiSeq 2500 sequencers were used to produce > 20 million single-end, 50-bp reads.

Table 1: *Summary of all samples suitable for RNA sequencing. Samples are grouped in columns by treatment, targeted area, age during injection and the hemisphere from which RNA was extracted.*

Treatment	Control								Knockdown							
Target area	HVC				CMM				HVC				CMM			
Age group	Juvenile		Adult		Juvenile		Adult		Juvenile		Adult		Juvenile		Adult	
Hemisphere	L	R	L	R	L	R	L	R	L	R	L	R	L	R	L	R
Samples	7	6	6	8	6	9	7	5	7	5	7	8	5	6	6	7

Gene expression analyses

Read counts ranged from 17.9 – 39.9 million reads per sample (mean: 24.2 ± 4 million reads per sample). One adult HVC control sample with significantly lower reads (bird ID: 5424; 2.48 million reads) was excluded from further data analyses. Quality control was conducted with FastQC (v0.11.9, Babraham Bioinformatics). Reads were aligned to the zebra finch Blue55 reference genome (NCBI assembly ID 5966711) using Rstudio (v1.3.1093) and the Rsubread package (v2.4.3) with standard settings (exception: indels = 10, count exon junctions). A BAM file was produced which included mapped (90.7 – 95.9 %, mean: 93.7 ± 1.2 %) and unmapped reads (see Figure 1a) to d) for controls and i) to l) for knockdowns). Multi-mapped reads were included in the analyses to cover potential splice variants. Rsubread was used to assign mapped reads (89.6 – 95.5 %, mean: 93.0 ± 1.4 %, see Figure 1e) to h) for controls and m) to p) for knockdowns) according to published annotations for the female zebra finch Blue55 reference genome. Counts were subsequently calculated according to protein-coding genes of the annotation file using standard settings of Rsubread with the exception of enabled counts of exon-exon junctions.

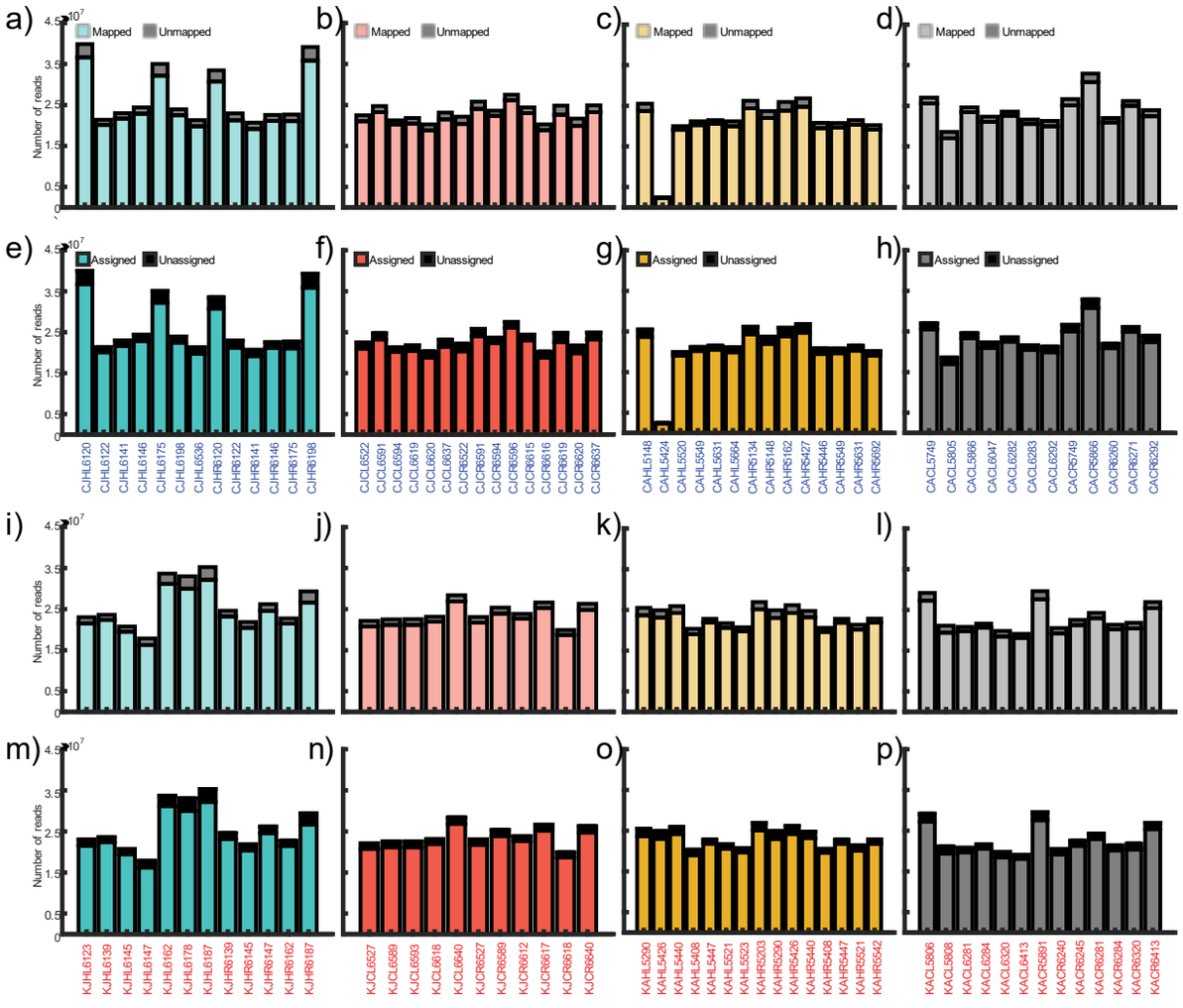


Figure 1: *Mapped and assigned reads per age group, injected area and treatment. Mapped and unmapped reads of juvenile control samples taken from HVC (a) or CMM (b) and adult control samples from HVC (c) or CMM (d). Note that one adult HVC control sample (CAHL5424) contained fewer reads than any other sample and was thus excluded from further analyses. Mapped and unmapped reads from knockdown samples taken from juvenile HVC (i), juvenile CMM (j), adult HVC (k) and adult CMM (l). The number of assigned reads is shown for control samples from juvenile HVC (e), juvenile CMM (f), adult HVC (g) and adult CMM (h). Assigned reads of knockdown samples is shown in the same order from m) to p). Individual sample IDs are indicated at the bottom and consist of a four-letter, four-digit code that is structured as follows: first letter = C for control or K for knockdown; second letter = J for juvenile or A for adult; third letter = H for HVC or C for CMM; fourth letter = L for left hemisphere or R for right hemisphere. Four numbers indicate the individual bird ID the sample was taken from. Hemispheres were not analysed separately during further analyses.*

Gene counts were normalised using reads per kilobase per million mapped reads (RPKM, Mortazavi *et al.*, 2008). Differentially expressed genes were limited to occurrences of >1 read per million mapped in at least two samples and visualised based on K-means clustering analyses.

Gene expression comparisons were conducted in Matlab release 2020a (Mathworks, Natick, USA) with the bioinformatics toolbox. Variance of read counts was identified by plotting the dispersion against the mean of the respective sample group. To determine statistical significance of gene expression differences, negative binomial models of the normalised read counts were conducted assuming a Poisson distribution, a constant variance link and a locally regressed non-parametric smooth function of the mean. Locally regressed modelling provided the best fit and was chosen for further analyses. To account for multiple testing of differentially expressed genes between control and knockdown samples, p-values were adjusted according to the Benjamini-Hochberg method (Benjamini and Hochberg, 1995) considering a 10% false positive rate. Differences in gene expression were visualised as a Venn diagram to identify overlapping and exclusive genes identified for each group. As the female reference genome has a slightly lower coverage (82.5) than the latest reference genome of the male zebra finch (88.2, NCBI assembly ID 10005361), highly significant but unannotated loci were referenced to the male reference genome to identify potentially unannotated regions in the female genome. Additionally, annotations from an Affymetrix array containing predicted exon sequences from previous incomplete zebra finch genome assemblies (MPIO-ZF1s520811, Dittrich *et al.*, 2014) were used to cross-reference loci that were not annotated in the female reference genome.

Relative expression levels of *FoxP1*, based on RPKM differences between control and knockdown samples, were further compared with those determined during qPCR analyses of the same samples. Gene ontology enrichment (GO) and local network analyses were conducted based on avian/mammalian gene orthologues in STRING v11 (Szklarczyk *et al.*, 2019) considering only categories with a p-value <0.05 and corrected for false discovery rate (Benjamini and Hochberg, 1995). All assigned genes across all control and knockdown groups from this study were used as the background reference to control for brain- and area-derived enrichments. Unbiased gene set enrichment analyses was conducted with GSEA v4.1.0 (Subramanian *et al.*, 2005) including only genes with >1 read per million mapped in at least two control and

knockdown samples, respectively. Standard settings remained unchanged and the cut-off was set to a gene set size of 15. The gene set database was based on a complete list of human gene symbols (c2.all.v7.4.symbols) and data were permutated 1000 times by phenotype (control or knockdown). Annotation data of the identified genes were based on a chip annotation database integrated in the GSEA programme to identify human orthologues of the counts mapped in the samples of this study (Human Gene Symbol with Remapping MSigDB.v7.4). The overlaps of differentially expressed genes of this study and databases were tested for significance using a Chi-Square test including Yate's correction for continuity.

Results

Modelling of normalised read counts and identification of genes showing significant differences in expression

Read counts of this study were modelled best by a negative binomial distribution based on a local regression for control and knockdown samples respectively taken from juvenile HVC (Fig. 2a, b), juvenile CMM (Fig 2e, f), adult HVC (Fig 2i, j) or adult CMM (Fig. 2m, n). The relationship between local dispersion and means for samples from different regions and developmental stages did not differ between control (blue) and knockdown samples (red) from birds injected as juveniles in HVC (Fig. 2c), as juveniles in CMM (Fig. 2g), and as adults in CMM (Fig. 2o). However, dispersion of normalised reads of HVC from birds which received a control construct was positively biased (Fig. 2k) and thus these may have a possible underrepresentation of weakly expressed genes in comparison to their respective knockdown samples. The distribution of differentially expressed genes in knockdown samples is shown as volcano plots for samples from juvenile HVC (Fig 2d), juvenile CMM (Fig 2h), adult HVC (Fig 2l) and adult CMM samples (Fig 2p). The distribution of log2 fold changes of individual control (filled circles) and knockdown samples (open circles) across all groups (Fig. 2q) does not indicate a bias in fold change of one specific subgroup or treatment and thus comparable levels of gene expression changes between all groups. The relative expression of *FoxP1* in target areas of knockdown birds was comparable to qPCR data used previously to identify knockdown efficiency (Chapter 2). However, as determined by a two-way ANOVA, the variance differed between assessment methods ($p < 0.0001$), while no significant differences were evident for area and age during

treatment ($p > 0.05$) or the identified knockdown efficiency within individual groups ($p > 0.05$).

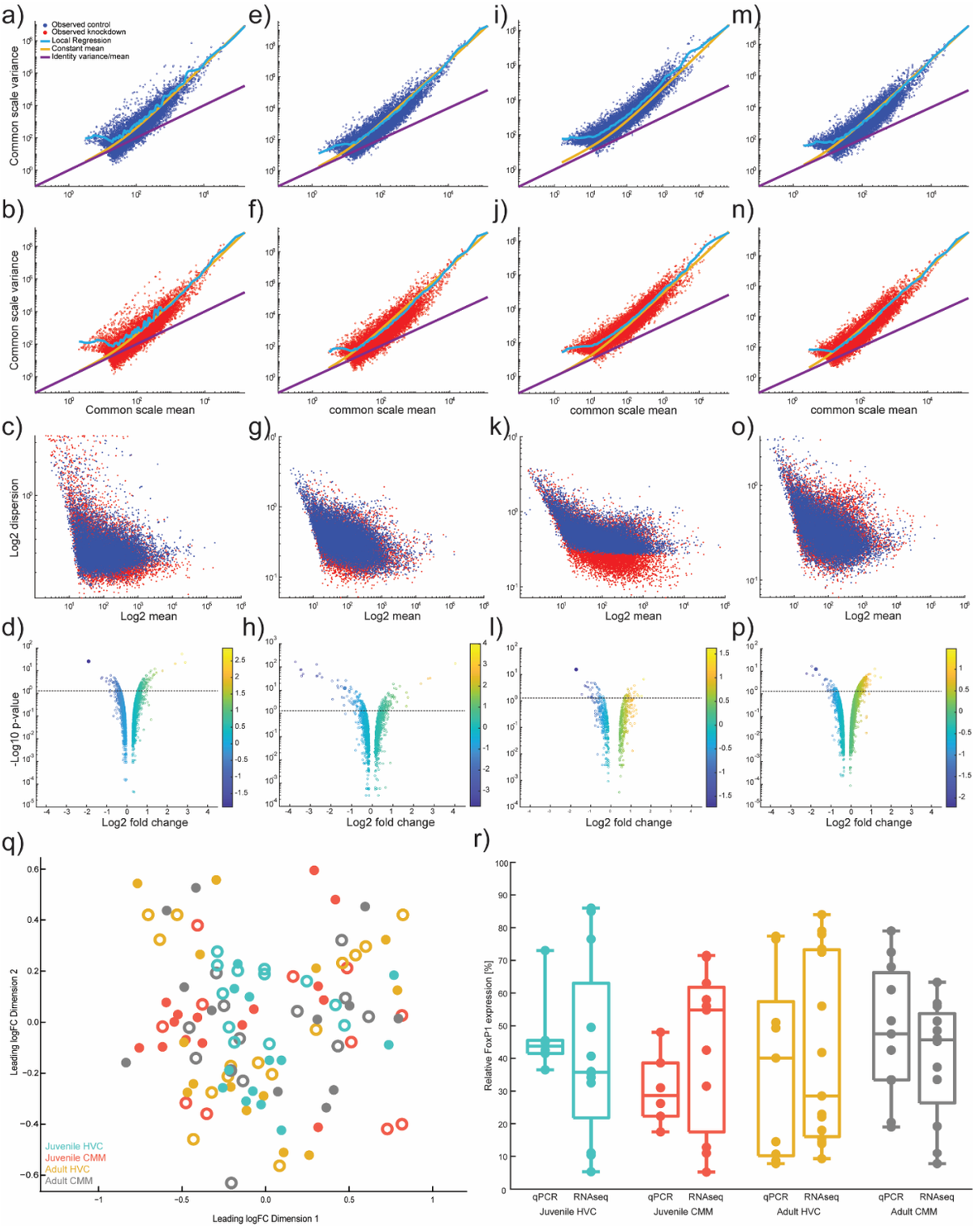


Figure 2: *Feature count properties, differentially expressed genes and their feature count distribution, and comparison between relative FoxP1 expression shown by qPCR and RNAseq. Each column symbolises one subgroup. a) – d) represent the data*

from juvenile HVC birds, e) – h) from juvenile CMM birds, i) – l) from adult HVC birds and m) – p) from adult CMM birds. a), e), i) and m) show the relation between a subgroups variance and mean for control samples, in blue. The same data are shown for knockdown samples in b), f), j) and n), in red. In order to specify the linkage type between the variance and mean, three approaches were taken. The purple line shows the results of assuming a linear correlation between variance and mean. The yellow curve shows the results of assuming the variance is a sum of the mean and a constant multiplied by the squared mean (yellow curve). Considering the variance as the mean read count variability, and applying a locally regressed smoothing function, results in the best fitting correlation (light blue curve). c), g), k) and o) show scatter plots of the log2 transformed dispersion and means of each subgroup's feature counts for control samples in blue and knockdown samples in red. In the subgroup of birds which received a control construct as adults in HVC, the dispersion dominates over the mean which suggests that more genes are highly expressed in the control when compared to the knockdown samples. d), h), l) and p) show volcano plots of all differentially expressed genes in the different subgroups. The dotted line indicates the significance threshold at an adjusted p-value of 0.05. Filled and enlarged circles indicate the differential expression strength of FoxP1. Colourbars symbolise the log2 fold change for each differentially expressed gene. q) shows two-dimensional scaling based on overall log2 fold changes of genes across all samples of this study, separated by colour. Filled circles indicate control samples, while open circles indicate knockdown samples. r) shows the relative FoxP1 expression levels in knockdown samples compared to control samples for a qPCR- and an RNAseq-based approach for all different subgroups. Relative expression levels differ significantly between results obtained by qPCR and RNAseq across all groups but neither between the target areas and developmental stages nor the assessment methods within each area and developmental stage (two-way ANOVA + TukeyHSD, method ($F=43.81$, $p < 0.0001$), subgroup ($F=2.65$, $p > 0.05$), method x subgroup($F=1.12$, $p > 0.05$).

Mapped genes across different groups

The number of identified genes based on read assignments varied between groups (Table 2, range 10,923 juvenile HVC control – 12,642 Adult HVC control).

Table 2: *Number of protein-coding genes per group and treatment (Control = Ctrl, Knockdown = KD) which were identified based on sequenced, mapped and assigned transcripts from total RNA sequencing data.*

Group	Juvenile HVC		Juvenile CMM		Adult HVC		Adult CMM	
Treatment	Ctrl	KD	Ctrl	KD	Ctrl	KD	Ctrl	KD
# of identified genes	10,923	11,068	11,439	11,140	12,642	12,530	11,841	11,409

In summary, counts were assigned to 13,695 protein-coding, annotated or predicted genes (Supplementary Table 2). This number corresponds to 84.5% of all 16,197 annotated or predicted protein-coding genes of the female zebra finch reference genome that was used for annotation. When non-coding genes and pseudogenes are also included, transcripts from 63.57% of 21,543 annotated segments of the reference genome were mapped and assigned in this study.

Dendrograms of differentially expressed genes

Hierarchical clustering of normalised reads from all genes identified across all samples of all groups results in a dendrogram which clusters all samples by age during treatment and the targeted area (Fig 3a). Samples generated from HVC of birds treated as adults were allocated in the most distant cluster in relation to the other three groups. Among the remaining three main clusters, juvenile CMM samples were the most distant, and juvenile HVC and adult CMM samples the closest. However, neither treatment nor hemisphere of the samples segregated in the overall hierarchical dendrogram, which is why samples from each subgroup of this analyses were then clustered separately. Samples from HVC of birds injected as juveniles form two main clusters to which both controls and knockdowns across both hemispheres contribute similarly (Fig. 3b). Samples from birds injected in CMM as juveniles cluster in two groups based on the normalised counts of all identified genes irrespective of hemisphere or treatment (Fig. 3c). Hierarchical clustering of normalised gene counts

of samples obtained from birds injected as adults in HVC (Fig. 3d) or CMM (Fig. 3e) results in three main clusters which neither segregate by hemisphere nor by treatment. In summary, hierarchical clustering suggests that local knockdowns during different developmental stages contribute differently to the transcriptome since the four treatment groups were well separated in the dendrogram that included all samples. However, when the groups are clustered separately, interindividual gene expression differences outweigh the effects caused by local knockdowns, as no clear discrimination of the used construct is evident at the transcriptome-wide level.

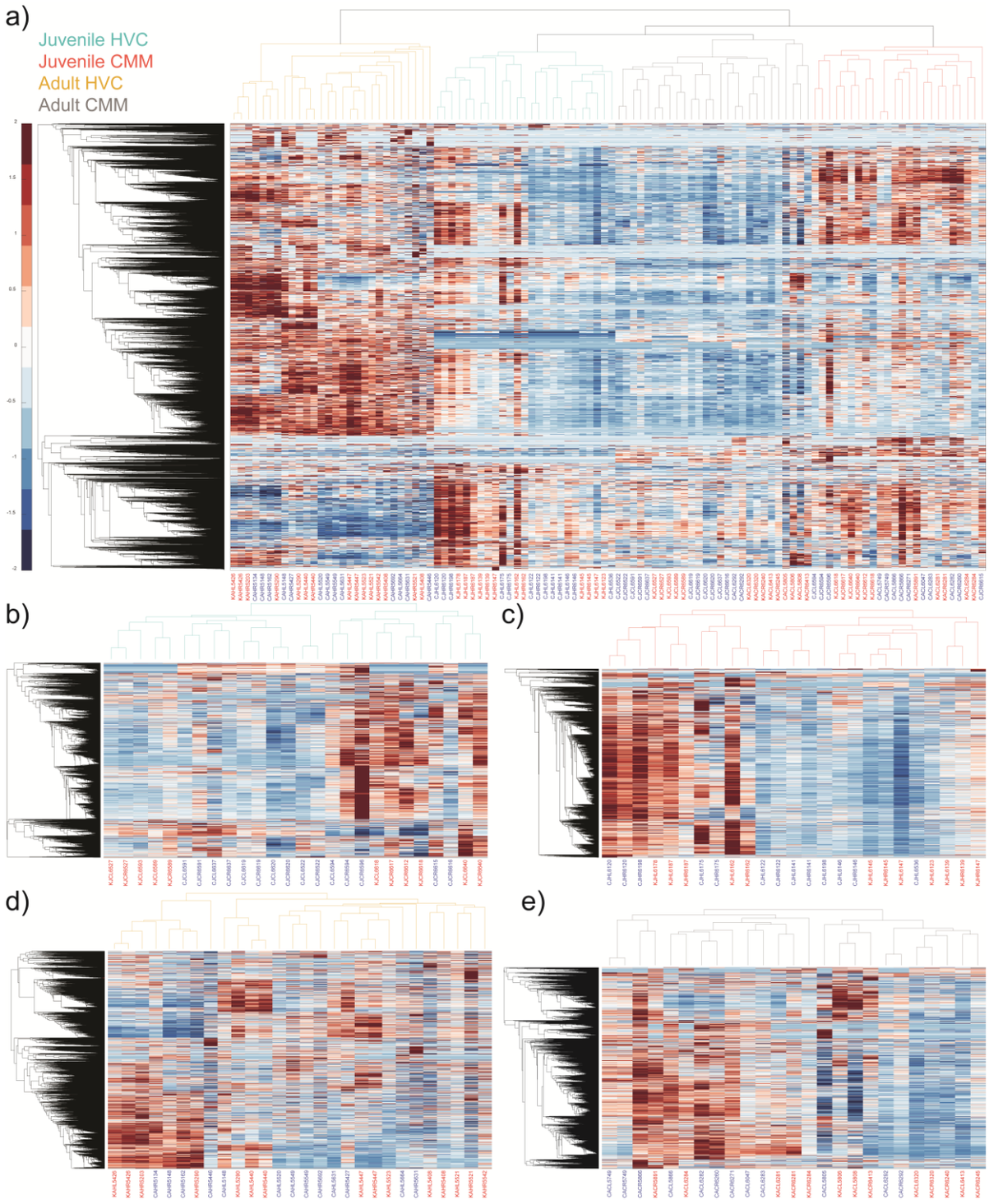


Figure 3: *Hierarchically clustered dendrograms of gene counts. Columns are clustered by sample ID, rows by gene ID. A heatmap encodes the standardised counts of the genes present in the dataset in relation to a mean of 0 and a standard deviation of 1. a) The dendrogram of the entire RNAseq data set shows clear clustering of knockdowns at different ages and in different areas as indicated by the colours of the dendrogram. b) to e) show individual dendrograms for each subgroup and indicate that,*

considered at the transcriptome-wide level, the within group effects of interindividual variability appear larger than effects of knockdowns. This is underlined by the close clustering of both hemispheres of one bird whenever data from both hemispheres were available. Even though some control and knockdown samples form subclusters, there is no clear overall distinction between treatments (i.e. control versus knockdown) within the dendrograms.

Differentially expressed genes

The number of genes which showed higher expression in response to *FoxP1* knockdown varied for the different subgroups, ranging from 26 (adult HVC) to 268 (adult CMM). (Figure 4a). Only a few such genes overlapped between different treatment groups, and there was no gene that overlapped between all groups (Figure 4a and b). The overlap was largest between the juvenile groups, where ten genes showed higher expression in response to *FoxP1* knockdown in both HVC and CMM. Four of these genes were annotated in the female zebra finch reference genome (*EXTL2*, *ASS1*, *THSD4*, *RP2*), while one gene was unannotated in the female reference genome (LOC100230755) yet recently annotated as coding for *ADAM33* in the male zebra finch reference genome. The last annotated gene overlapping between juvenile HVC and CMM samples codes for a tRNA (TRNAG-GCC). Four unannotated loci are also upregulated in both juvenile knockdown groups. Three loci, all unannotated, showed higher expression in response to *FoxP1* knockdown in both adult groups of this study, including one putative orthologue (LOC100222415) of *CYP2D14*, a cytochrome oxidase. One unannotated gene each overlapped between juvenile and adult HVC (LOC116806907) or CMM (LOC100230293) samples. According to previous microarray analyses (Dittrich *et al.*, 2014), the unannotated gene identified in both juvenile and adult CMM samples might be an orthologue of *CYB561*, a cytochrome oxireductase. One gene showed consistently elevated expression in response to *FoxP1* knockdown in three of the four subgroups (i.e. all except the adult HVC samples), annotated as NEK5, NIMA (never in Mitosis Gene A)-Related Kinase 5, encoding a serine/threonine-protein kinase.

Of the 26 genes that exclusively showed elevated expression in adult HVC knockdowns, *RPE65*, *PLXNB1* and *CUTA* may be of special interest, considering the prior literature. *RPE65* codes for retinoid isomerohydrolase; this protein is involved in

retinoic acid signalling, a pathway that has been linked to FOXP2 regulatory networks in previous studies (Van Rhijn and Vernes, 2015). *PLXNB1* expression has been reported in HVC (Lovell *et al.*, 2008) and the gene is implicated in the SLIT-ROBO signalling pathway (Xu and Fan, 2008; Hirschberg *et al.*, 2010; Schiweck *et al.*, 2015); variants of genes in this pathway have been associated with developmental dyslexia, expressive vocabulary in human infants, and performance on non-word repetition tasks in some studies (Hannula-Jouppi *et al.*, 2005; Bates *et al.*, 2010; Pourcain *et al.*, 2014; Mozzi *et al.*, 2016). Deletions of a human genomic region encompassing the ortholog *CUTA*, a CutA Divalent Cation Tolerance Homolog have been identified in humans with intellectual disability, hearing loss and delayed speech development (Writzl and Knegt, 2013) or absence of language (Zollino *et al.*, 2010).

Table 3 gives information on the ten most significant genes with elevated expression in each subgroup of this study. The entire set of significant genes is shown in Supplementary Table 3. In juvenile HVC knockdowns, *RGR, RLBP1* and *TUBAL3* showed significantly elevated expression as compared to controls. *RGR* and *RLBP1* encode members of the retinoid cycle (Saari *et al.*, 2001; Maeda *et al.*, 2003), which could be interesting in light of the putative link proposed between the dimerising partner of FoxP1, FoxP2 and retinoid related processes (Van Rhijn and Vernes, 2015). *TUBAL3* codes for a tubulin that may interact with the SLIT-ROBO signalling pathway according to the PathCards database (OMICS_07645). Among genes showing elevated expression in juvenile CMM knockdowns are *RP2*, encoding a protein implicated in retinitis pigmentosa in humans (Veltel and Wittinghofer, 2009), *IFI6*, encoding an interferon inducible protein, and *ROBO2*. In adult CMM knockdowns, genes with increased expression are associated to mitochondria, the respiratory complex and ribosomal actions as indicated by the top enriched cellular component GO terms (see Table 5, next section). Among these genes is *NDUFB1*, an oxireductase that shows reduced expression in the blood of early stage Alzheimer's disease patients (Lunnon *et al.*, 2017).

The number of genes showing significantly reduced expression in knockdown birds also varied across groups, ranging from 34 (adult HVC) to 120 (juvenile HVC) (Fig. 4c and d, Supplementary Table 3). Few genes overlapped between the treatment areas/stages, and *FoxP1* was the only to show significantly reduced expression across all groups (Fig. 4d). Knockdowns in both juvenile groups resulted in lower expression

for three protein-coding genes (*EIF2B5*, *SPOCK2*, *B2M*) and one unannotated locus (LOC10228369). Adult knockdowns of either area did not share any differentially expressed genes. Juvenile and adult knockdowns in HVC resulted in reduced expression of *IFI6*, encoding an interferon inducible protein and *SPINT1*, encoding a serine peptidase inhibitor. Notably, *IFI6* expression was significantly elevated in juvenile CMM knockdowns, raising the possibility that it may be differentially regulated by FoxP1 in different brain areas. In juvenile HVC and adult CMM two protein-coding genes (*ETFB*, *DHX33*) and four unannotated loci (LOC101234199, LOC100229421, LOC100224927, LOC116807667) overlapped between samples. *ETFB* is potentially linked to ASD according to the SFARI database which lists genes associated with this group of developmental disabilities. Although officially unannotated, it is thought that LOC100229421 may code for interferon-induced protein IFIT5 (Scalf, 2018), while LOC100224927 has been annotated in the male zebra finch genome as *OASL*, an oligoadenylate synthetase. All groups, with the exception of adult HVC knockdowns, overlapped in showing significantly reduced expression of *JCHAIN*, an immunoglobulin, as well as an unannotated locus LOC116809013 (Fig 4d).

Considering current knowledge on cognitive phenotypes associated with genetic manipulations of *FoxP1* in animals or disruptive variants in humans, some of the genes showing reduced expression in only one treatment group may be of special interest (Table 4 and Supplementary Table 3). In juvenile HVC knockdowns, such genes include the chromatin remodelling gene *ACTL6B*, human mutations of which result in intellectual disability, absence of speech or limited vocabulary (Bell *et al.*, 2019; Fichera *et al.*, 2019), and *DBN1,* encoding an actin-binding protein implicated in Alzheimer's disease in humans, and neuronal migration and synaptic plasticity in animal models (Shirao *et al.*, 2017). In juvenile CMM knockdowns, genes with significantly reduced expression included *SEMA3E,* encoding a semaphorin protein which forms complexes with plexins to regulate neuronal development, possibly via the SLIT-ROBO pathway (Xu and Fan, 2008; Schiweck *et al.*, 2015; Mata *et al.*, 2018). Genes exclusively downregulated in birds which received a *FoxP1* knockdown as adults in HVC are e.g. *PNMT*, *HRH1*, *TMEM233* and *TUBAL3* which is upregulated in birds of the juvenile HVC group. *PNMT* and *HRH1* are both implicated in the catecholamine pathway (Marley *et al.*, 1991; Kubovcakova *et al.*, 2004). In mouse astrocytes, homozygous knockouts of *HRH1* resulted in reduced anxiety and impaired novel object recognition

memory (Kárpáti *et al.*, 2019). Knockdown of *FoxP1* in CMM of adults resulted in reduced expression of *CHRNA10*, which encodes a subunit of a nicotinic acetylcholine receptor, and is implicated in auditory olivocochlear system development and function in mice (Vetter *et al.*, 2007). Deletion of the interferon regulatory factor *IRF1* in mice leads to cognitive impairments as demonstrated by reduced performance during water maze tasks (Mogi *et al.*, 2018). However, no cognitive impairments in female zebra finches which received a *FoxP1* knockdown could be identified during behavioural experiments of this thesis.

Table 3: *Top ten genes with most significantly elevated expression following local FoxP1 knockdowns in each subgroup. Gene symbols are followed by their average counts in knockdown and control samples, their respective log2 fold change (log2FC) and p-values adjusted by false discovery rate. In the last column, the gene name and an associated function are indicated. For loci which are not annotated in the reference genome, putative orthologues based on previous micro-array data* (Dittrich *et al.*, 2014) *are indicated.*

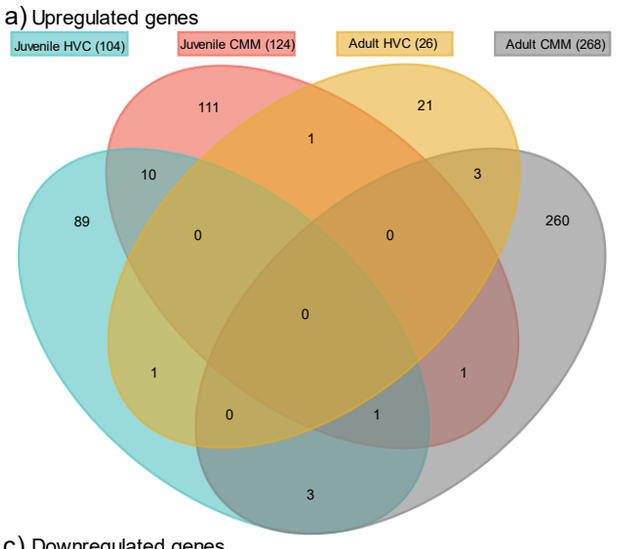
Juvenile HVC	**Counts KD**	**Counts Ctrl**	**log2FC**	**p adj.**	**Name and putative function**
FANCM	382.521	23.5654	4.0208	2.8E-142	FA complementation group M, cytogenetic instability, increased chromosomal breakage
RGR	80.3917	11.5054	2.80474	9.48E-34	Retinal g protein coupled receptor, retinaldehyde binding, retinitis pigmentosa
RLBP1	77.675	11.95	2.70044	4.27E-32	Retinaldehyde binding protein 1, component of visual cycle
TRNAE-UUC-2	60.2292	11.75	2.3578	9.45E-20	Transfer RNA glutamic acid
LOC100220024	25.6817	7.74308	1.72976	1.34E-05	Folate receptor gamma, FOLR3, cancer, innate immune system, endocytosis
LOC100230755	38.0108	12.0385	1.65876	1.22E-08	ADAM33, disintegrin and metalloproteinase domain-containing protein 33, cell-cell and cell-matrix interactions, neurogenesis
CHST9	18.9017	6.01692	1.65142	0.008931	Carbohydrate sulfotransferase 9, protein modification in golgi membrane, cell-cell interaction, signal transduction
TUBAL3	136.321	53.99	1.33624	9.41E-16	Tubulin alpha like 3, development slit-robo signalling, GTP binding
NEK5	27.3525	11.9677	1.19253	0.000712	Nima related kinase 5, transferase activity
LOC100228510	45.4667	21.2508	1.09729	8.49E-06	Sodium/hydrogen exchanger 2, ion exchange, cell volume regulation
Juvenile CMM					
LOC100224927	45.9546	6.20333	2.88909	4.82E-24	2’5’-oligoadenylate synthase 1, antiviral enzyme, also apoptosis, cell growth, differentiation, gene regulation
IFI6	1187.18	179.073	2.72892	2.91E-55	Interferon alpha inducible protein 6, apoptosis regulation

THSD4	58.2682	21.1693	1.46073	7.8E-09	Thrombospondin Type 1 Domain Containing 4, protein metabolism, metalloendopeptidase
TRNAG-GCC	167.5	64.1333	1.38502	1E-12	Transfer RNA glycine
ERP27	54.8536	21.5347	1.34893	1.72E-07	Endoplasmic reticulum protein, putative chaperone?
RP2	58.0673	23.0547	1.33267	1.4E-07	Rp2 activator of ARL3 GTPase, retinitis pigmentosa, folding of neuron specific tubulin isoforms, GTP binding
LOC101233947	19.3309	8.45533	1.19298	0.013185	Protein tilB homolog, LRRC6, DNAAF11, cilia motility
LOC100230755	29.2209	13.2607	1.13985	0.000164	ADAM33, disintegrin and metalloproteinase domain-containing protein 33, cell-cell and cell-matrix interactions, neurogenesis
ROBO2	392.557	178.265	1.13888	4.02E-09	Roundabout guidance receptor 2, axon guidance, cell migration, expressive language vocabulary in infants
AGR3	30.9036	15.0093	1.04192	0.000666	Anterior gradient 3, ER protein, protein folding, ciliary beat frequency
Adult HVC					
BGLAP	19.1887	7.42571	1.36965	0.030324	Bone gamma carboxyglutamate protein, osteoblasts, energy metabolism, calcium binding
CUTA	21.37	8.61643	1.31043	0.018327	Cuta divalent cation tolerance homolog, deafness
LPL	80.1173	35.7721	1.16328	0.000138	Lipoprotein lipase, heart, muscle and adipose tissue, receptor mediated lipoprotein uptake,
MMP2	36.538	16.6279	1.1358	0.009484	Matrix metallopeptidase 2, cleave ecm components, signal transduction
ENPP6	41.626	21.2114	0.97264	0.0305	Ectonucleotide pyrophosphatase, phosphodiesterase, neuropathy,
LOC100226434	59.9187	31.395	0.93247	0.017886	Hes5 like, transcription factor, brain development process, notch signalling
LOC100222415	56.8087	30.7921	0.88355	0.041241	Cytochrome p450 like, oxyreductase
VAMP1	126.771	69.4379	0.86843	0.003487	Vesicle associated membrane protein, synaptic vesicle docking, fusion presynaptic, spastic ataxia
PLXNB1	64.6407	35.5671	0.8619	0.022988	Plexin B1, slit robo signalling, semaphorine receptor
RPE65	70.1247	39.4414	0.83021	0.033594	Retinal pigment epithelium specific 65kda protein, vision
Adult CMM					
ND3	916.392	490.076	0.90296	2.94E-08	Mitochondrial NADH dehydrogenase, neurometabolic disorders
LOC100190731	445.739	243.702	0.87108	5.14E-08	Metallothionein, zinc ion binding
EPSTI1	41.7846	23.9875	0.80069	0.0007	Epithelial stromal interaction, cancer, lupus
LOC100190094	548.732	319.727	0.77926	2.22E-07	Metallothionein-i-like, MT4(?), zinc and copper ion binding, differentiation of stratified epithelia
COX7C	1101.5	646.551	0.76864	4.3E-06	Cytochrome oxidase subunit 7c, neurodegeneration
ATP5MPL	732.368	433.533	0.75643	1.53E-06	ATP synthase membrane subunit j
RPS29	280.552	166.643	0.75151	1.18E-06	Ribosomal protein s29, protein metabolism
HBAD	202.328	120.857	0.7434	1.81E-06	Hemoglobin subunit alphaD
NDUFB1	292.978	176.772	0.7289	2.32E-06	NADH Ubqiquinone Oxireductase Subunit 1, neuropathy, dysarthria
DIO2	87.9962	53.1717	0.72678	0.000033	Iodothyronine deiodinase 2, thyroid hormone metabolism

Table 4: *Top ten genes with most significantly reduced expression following local FoxP1 knockdowns in each subgroup. Gene symbols are followed by their average counts in knockdown and control samples, their respective log2 fold change (log2FC) and p-values adjusted by false discovery rate. In the last column, the gene name and an associated function are indicated. For loci which are not annotated in the reference genome, putative orthologues based on previous micro-array data* (Dittrich *et al.*, 2014) *are indicated.*

Juvenile HVC	**Counts KD**	**Counts Ctrl**	**log2FC**	**p adj.**	**Name and putative function**
JCHAIN	53.9292	584.981	-3.43925	8.15E-78	IgA and IgM factor, immune system
IFI6	578.643	3745.34	-2.69435	2.3E-130	Interferon Alpha Inducible Protein 6, apoptosis, innate immune system
LOC100229421	22.0233	138.795	-2.65585	1.83E-42	IFIT5, Interferon induced protein with tetratricopeptide repeats 5, tRNA binding, innate immune response
LOC100224927	16.8525	91.3369	-2.43824	1.07E-36	2'-5'oligoandeylate synthase 1, interferone induced, RNA degradation, reduced gene expression
LOC100224071	6.48667	24.6092	-1.92365	8.16E-05	Ovostatin-like(?), proteinase inhibitor
RSAD2	11.6642	30.1862	-1.3718	2.01E-05	Radical s adenosyl methionine domain containing 2, antiviral protein
FOXP1	55.3808	139.545	-1.33328	4.93E-13	Forkhead box transcription factor P1, intellectual disability, autism spectrum disorder
ETFB	208.16	445.645	-1.0982	4.25E-11	Electron transfer flavoprotein subunit beta, beta polypeptide, electron shuttling
ACTL6B	39.3208	84.0692	-1.09628	4.98E-09	Actin like 6b, intellectual developmental disorder with severe speech and articulation defects, cytoskeleton
DBN1	97.3458	201.215	-1.04754	1.1E-14	Drebrin 1, neuronal growth, Alzheimer, down syndrome
Juvenile CMM					
FOXP1	42.16	157.591	-1.90224	1.39E-26	Forkhead box transcription factor P1, intellectual disability, autism spectrum disorder
JCHAIN	10.6082	24.8373	-1.22733	7.51E-06	Joining of multimeric IgA and IgM
TNFSF13B	10.1591	20.71	-1.02756	0.000465	TNF superfamily, tumor necrosis factor, signalling receptor binding
TRNAE-CUC2	14.5455	29.25	-1.00787	2.14E-05	RNAgene, transfer RNA glutamic acid
SEMA3E	12.0636	22.1753	-0.87829	0.003138	Semaphorin 3E, axon guidance
LOC100232025	33.4455	56.0313	-0.74442	0.000387	Extracellular fatty acid-binding protein-like, immune?
SOX2	30.7546	49.03	-0.67287	0.001287	Stem-cell development
B2M	263.276	411.456	-0.64417	8.3E-06	Beta2Microglobulin, MHC heavy chain, antibacterial activity in amniotic fluid
FGFBP3	24.7791	38.1753	-0.62352	0.008708	Fibroblast growth factor binding protein 3, gpcr signalling
DPY19L3	35.2927	53.3673	-0.59659	0.003742	DPY-19 like C-mannosyltransferase 3, spermatogenic failure, podoconiosis
Adult HVC					
FOXP1	36.1647	119.288	-1.72179	3.49E-16	Forkhead box transcription factor P1, intellectual disability, autism spectrum disorder

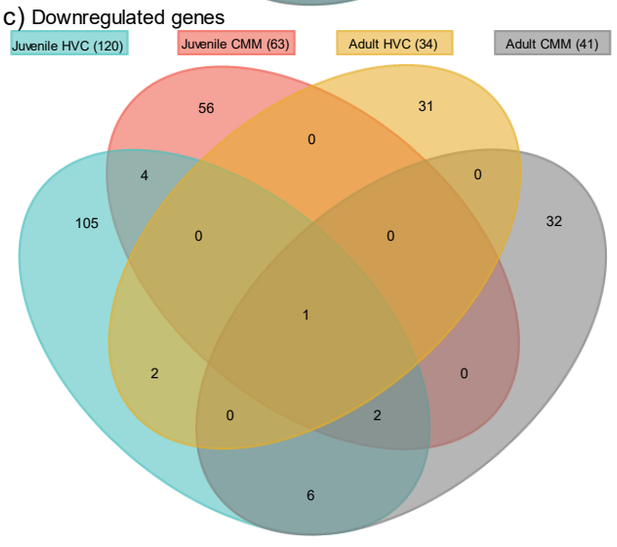
PNMT	5.44333	14.7936	-1.44241	0.016428	Phenylethanolamine n methyltransferase, catecholamine pathway
FOXJ1	10.7327	26.1686	-1.28583	0.000166	Motile cilia tf, left/right asymmetry, lupus
TMEM233	18.9307	42.3543	-1.16178	7.17E-06	Transmembrane protein 233, interferon induced,
DHTKD1	13.726	30.3729	-1.14587	0.000473	Dehydrogenase e1 and transketolase domain containing 1, mitochondrial, charcot marie tooth
MROH1	14.3007	31.1921	-1.1251	0.000494	Maestro heat like repeat family member 1, binding
C3	14.0433	28.5521	-1.02371	0.002424	Complement c3, inflammation and antimicrobial activity, gpcr signalling
PPL	37.25	72.9143	-0.96896	1.41E-05	Periplakin, desmosome component, cell growth
HRH1	9.35267	18.2821	-0.96699	0.035042	Histamine receptor h1, messenger, catecholamine release, neurotransmission, memory and learning
ZNFX1	9.10733	17.6557	-0.95503	0.047669	Zinc finger nfxtype containing 1, cancer, parotid disease
Adult CMM					
LOC100224927	6.35	29.3792	-2.20996	1.34E-08	2'-5'-oligoadenylate synthase-like protein 1
JCHAIN	45.4785	176.443	-1.95595	2.17E-17	Joining of multimeric IgA and IgM
LOC100229421	8.09	30.145	-1.89771	3.23E-07	Putative retinoic Acid and Interferon inducible Protein, viral RNA sensor
FOXP1	57.7454	194.766	-1.75396	2.31E-13	Forkhead box transcription factor P1, intellectual disability, autism spectrum disorder
TRNAG-GCC-2	7.30077	19.2958	-1.40217	0.002437	Glycine Transfer RNA
IRF1	21.58	53.4717	-1.30908	8.99E-05	Interferon regulating factor, viral response
TGM4	17.5331	42.8817	-1.29028	4.46E-05	Transglutaminase4, seminal tract in mammals
TRNAG-GCC	30.25	69.3958	-1.19791	0.000037	Glycine Transfer RNA
CHRNA10	8.62846	19.3392	-1.16435	0.046732	Neuronal acetylcholine receptor subunit alpha10, nAchR, olivocochlear sytem in auditory system
TNNI1	11.7777	26.14	-1.1502	0.020979	Troponin1, slow skeletal Muscle, but also corpus callosum in mice



b)

Juvenile HVC ∩ juvenile CMM
LOC116806890
LOC116806735
EXTL2
ASS1
LOC116807081
THSD4
LOC116807587
TRNAG-Gcc
RP2
LOC100230755
Juvenile HVC ∩ adult HVC
LOC116806907
Adult HVC ∩ adult CMM
LOC115490588
LOC115497561
LOC100222415
Juvenile CMM ∩ adult CMM
LOC100230293
Juvenile HVC ∩ juvenile CMM ∩ adult CMM
NEK5

Adult HVC exclusive
LOC100225683
BSG
LOC101233698
LOC100228369
CAVIN2
PHYHD1
RHBDL1
RPE65
LOC116806758
PLXNB1
VAMP1
LOC116807507
LOC100226434
ENPP6
MMP2
LPL
LOC115494492
LOC115491277
CUTA
BGLAP
LOC100218227



d)

Juvenile HVC ∩ juvenile CMM
EIF2B5
SPOCK2
LOC100228369
B2M
Juvenile HVC ∩ adult HVC
IFI6
SPINT1
Juvenile HVC ∩ adult CMM
LOC101234199
LOC100229421
LOC100224927
LOC116807667
ETFB
DHX33
Juvenile HVC ∩ juvenile CMM ∩ adult CMM
LOC116809013
JCHAIN
Juvenile HVC ∩ juvenile CMM ∩ adult HVC ∩ adult CMM
FoxP1

Adult HVC exclusive
PNMT
FOXJ1
LOC115495014
TMEM233
DHTKD1
MROH1
LOC115494495
C3
PPL
HRH1
ZNFX1
OLFM4
LOC100229516
TUBAL3
LOC115497572
ADSS1
LOC115495360
CYP19A1
LOC115495051
ENKD1
ME2
GPR26
C2CD4C
NRBP1
GPR78
KL
BOP1
VIPR1
LOC115497743
KCNK5
LOC101233348

Figure 4: *Overlapping and exclusive differentially expressed genes across all subgroups. a) shows a Venn diagram of genes with significantly elevated expression which are exclusively represented in one or shared across multiple subgroups. b) shows the list of genes corresponding to a). As a behavioural phenotype was identified for birds which received a knockdown in HVC as adults (see Chapter 2 of this thesis), genes which exclusively show elevated expression in these birds are listed separately. Similar to a), c) shows a Venn diagram of genes with significantly reduced expression, while d) lists the genes represented in c).*

GO terms and local network clusters

Due to the low number of overlapping genes which showed significantly increased or reduced expression across all experimental groups, GO and network terms were analysed separately for each region and developmental stage. The ten most enriched GO terms of each category and network terms are shown for genes with increased (Table 5) and reduced (Table 6) expression associated with *FoxP1* knockdown. In birds which received a knockdown as juveniles in HVC, no GO terms were significantly enriched, and only one network cluster associated with retinol metabolic process and retinol binding (ID: 9606_CL_25005) was enriched. Genes with elevated expression in juvenile CMM knockdown samples are represented by multiple GO terms of the biological processes (BP) and cell cycle (CC) categories. All three significantly enriched BP terms are related to cell adhesion. Seven of the ten most enriched GO terms of the CC category are related to plasma membrane or the extracellular matrix. Neuron projection (GO:0043005), postsynapse (GO:0098794) and syntrophin complex (GO:0016013) are also significantly enriched terms. Multiple enriched local network clusters based on genes with elevated expression in juvenile CMM samples represent functions related to cell adhesion, and there was also enrichment for networks related to voltage gated potassium channels (9606_CL_8930), interneuron migration (9606_CL_230) and growth response (9606_CL:2481). In adult HVC samples, enrichment was seen for one BP term related to bone trabecula morphogenesis (GO:0061430) and one term of the molecular function (MF) category corresponding to the membrane protein phosphatidylserine (GO:0001786). In this group of genes showing increased expression local network clusters are related to retinol metabolic processes and retinol-binding (9606_CL:24001), similar to juvenile HVC samples. Additionally, a local cluster implicated in mixed processes such as matrix metalloproteinases (9606_CL:907) is enriched across genes with elevated expression in adult HVC knockdown samples.

The top GO terms of genes showing increased expression in adult CMM knockdowns are related to ribosomal or mitochondrial processes across all three GO categories. The most enriched BP terms are linked to processes from e.g. translation (GO:0006412) to protein targeting to the endoplasmatic reticulum (GO:0045047) and protein localisation (GO:0072594). Six of ten CC terms are related to ribosomes, and four are implicated in mitochondrial processes. Out of ten MF terms, two are related to

ribosomal functions, six represent mitochondrial functions and cellular respiration, and the remaining two relate to structural molecule activity (GO:0005198) and proton transmembrane transporter activity (GO:0015078). Local clusters enriched in genes with increased expression following knockdown in adult CMM represent similar functions to the enriched GO terms, including e.g. ribosomal activities such as peptide chain elongation (9606_CL:14976) or mitochondrial complexes like the respirasome (9606_CL:22328). As the number of significantly enriched GO terms based on upregulated genes in adult CMM knockdowns was larger than in all other groups, the entire set of GO terms of this group is listed in Supplementary Table 4.

Compared to GO terms and local network clusters enriched in genes with increased expression in specific knockdown groups, few terms and networks were enriched in genes with reduced expression (Table 6). In this case, for juvenile HVC knockdowns there was enrichment for none of the GO terms and only one local cluster, with mixed associations including axonal growth inhibition (9606_CL:616). Genes with reduced expression in juvenile CMM knockdowns were enriched for one BP term on antibacterial humoral response (GO:0019731), CC terms related to mitochondrial processes such as mitochondrial proton-transporting ATP synthase complex (GO:0005753), and organelle membrane related processes such as e.g. organelle envelope (GO:0031967). Two enriched local clusters were related to oxidative phosphorylation (9606_CL:22327) and proton-transporting ATP synthase complex (9606_CL:22571). For genes with knockdown-related reductions of expression in birds that had been injected as adults, there were no enriched GO terms or local clusters.

In addition, GO terms and network clusters were assessed for genes exclusively regulated in birds which received a local *FoxP1* knockdown in HVC as adults, since this subgroup had shown a behavioural phenotype in Chapter 2 of this study. In this case, for genes with significantly increased expression, enrichment was seen for GO:0061430 which is associated to bone trabecula morphogenesis, and two local network clusters: 9606_CL_24001 retinol metabolic process, and retinol binding and 9606_CL_907 mixed, incl. activation of matrix metalloproteinases and dissolution of fibrin clot. For genes with significantly reduced expression, no GO terms or clusters were enriched. The only detected enrichment in this set of downregulated genes was attributed to a UniProt keyword (KW-0297) associated to G-protein coupled receptor

which is further linked to two MF GO terms G protein-coupled receptor activity (GO:0004930) and G protein-coupled receptor signalling pathway (GO:0007186).

Table 5: *Significantly enriched gene ontology (GO) terms and local clusters based on genes with increased expression following local FoxP1 knockdowns in each subgroup. GO terms and local clusters are based on the human orthologues of the genes found to be differentially expressed in this study. Analyses are based on data deposited in the string database (v11.0). The maximum ten most significant GO terms are clustered based on their affiliation to cellular components (CC), molecular functions (MF) or biological processes (BP). Each term is followed by the number of genes contributing to it as well as the total number of genes represented by each term and its respective false discovery rate (FDR).*

Juvenile HVC			
Local cluster	**Description**	**Gene counts**	**FDR**
9606_CL:24005	retinol metabolic process, and Retinol-binding	3/12	0.032
Juvenile CMM			
BP term			
GO:0098742	cell-cell adhesion via plasma-membrane adhesion molecules	11/101	1.32E-06
GO:0007156	homophilic cell adhesion via plasma membrane adhesion molecules	8/65	6.65E-05
GO:0007155	cell adhesion	13/480	0.038
CC term			
GO:0005886	plasma membrane	38/2816	0.0177
GO:0016010	dystrophin-associated glycoprotein complex	3/16	0.0177
GO:0016021	integral component of membrane	36/2716	0.0177
GO:0031224	intrinsic component of membrane	38/2792	0.0177
GO:0045211	postsynaptic membrane	8/186	0.0177
GO:0016013	syntrophin complex	2/4	0.0216
GO:0016020	membrane	53/4968	0.0216
GO:0043005	neuron projection	16/869	0.0216
GO:0098794	postsynapse	9/334	0.0282
GO:0031226	intrinsic component of plasma membrane	15/866	0.0452
Local cluster			
9606_CL:2481	bZIP transcription factor, and Early growth response, N-terminal	3/12	0.02
9606_CL:6791	mixed, incl. Adherens junctions interactions, and Alpha-catenin	4/25	0.02
9606_CL:6796	Adherens junctions interactions	3/10	0.02
9606_CL:6813	Cadherin cytoplasmic region	2/2	0.021
9606_CL:8930	Voltage gated Potassium channels, and Phase 1 - inactivation of fast Na+ channels	3/17	0.0228
9606_CL:230	chemorepulsion involved in interneuron migration from subpallium to cortex and ovarian cumulus expansion	2/4	0.0289
Adult HVC			
BP term			
GO:0061430	bone trabecula morphogenesis	2/9	0.0345
MF term			
GO:0001786	phosphatidylserine binding	2/28	0.0442
Local cluster			

9606_CL:24001	retinol metabolic process, and Retinol-binding	2/18	0.0111
9606_CL:907	mixed, incl. Activation of Matrix, and Dissolution of Fibrin Clot Metalloproteinases	2/21	0.0111
Adult CMM			
BP term			
GO:0006614	SRP-dependent cotranslational protein targeting to membrane	42/80	2.24E-40
GO:0045047	protein targeting to ER	43/88	2.24E-40
GO:0000184	nuclear-transcribed mRNA catabolic process, nonsense-mediated decay	42/100	6.12E-38
GO:0006412	translation	56/293	3.72E-36
GO:0043043	peptide biosynthetic process	57/308	3.72E-36
GO:0006413	translational initiation	42/120	1.77E-35
GO:0000956	nuclear-transcribed mRNA catabolic process	44/158	9.39E-34
GO:0090150	establishment of protein localization to membrane	45/173	1.62E-33
GO:0006605	protein targeting	48/264	7.66E-30
GO:0072594	establishment of protein localization to organelle	48/311	5.43E-27
CC term			
GO:0044391	ribosomal subunit	42/125	1.14E-34
GO:0005840	ribosome	44/165	3.83E-33
GO:0022626	cytosolic ribosome	33/57	
GO:0005743	mitochondrial inner membrane	46/331	3.87E-24
GO:0098798	mitochondrial protein complex	38/201	6.53E-24
GO:0031966	mitochondrial membrane	52/489	8.31E-23
GO:0098800	inner mitochondrial membrane protein complex	28/94	4.88E-22
GO:0022627	cytosolic small ribosomal subunit	20/31	7.05E-21
GO:0015935	small ribosomal subunit	23/56	1.11E-20
GO:0070469	respirasome	23/64	1.19E-19
MF term			
GO:0003735	structural constituent of ribosome	45/126	2.33E-38
GO:0005198	structural molecule activity	49/405	3.86E-23
GO:0008137	NADH dehydrogenase (ubiquinone) activity	15/41	3.59E-12

GO:0016651	oxidoreductase activity, acting on NAD(P)H	17/70	8.04E-12
GO:0016491	oxidoreductase activity	33/437	4.46E-10
GO:0015078	proton transmembrane transporter activity	12/68	6.34E-07
GO:0019843	rRNA binding	9/48	2.43E-05
GO:0004129	cytochrome-c oxidase activity	5/6	2.48E-05
GO:0009055	electron transfer activity	9/61	9.60E-05
GO:0008121	ubiquinol-cytochrome-c reductase activity	3/4	4.40E-03
Local cluster			
9606_CL:14976	Peptide chain elongation	45/72	7.28E-47
9606_CL:14978	Peptide chain elongation	43/70	5.48E-45
9606_CL:14966	GTP hydrolysis and joining of the 60S ribosomal subunit, and Protein export	48/122	1.75E-43
9606_CL:14980	Peptide chain elongation	41/66	3.64E-43
9606_CL:14967	GTP hydrolysis and joining of the 60S ribosomal subunit, and Protein export	46/117	1.16E-41
9606_CL:14982	Peptide chain elongation	38/63	1.11E-39
9606_CL:14983	Peptide chain elongation	33/55	3.08E-34
9606_CL:14985	Viral mRNA Translation	30/50	4.73E-31
9606_CL:22327	Oxidative phosphorylation	33/109	1.54E-26
9606_CL:22328	respirasome	24/64	6.08E-21

Table 6: *Significantly enriched GO terms and local clusters based on genes with decreased expression following local FoxP1 knockdowns in each subgroup. GO terms and local clusters are based on the human orthologues of the genes found to be differentially expressed in this study. Analyses are based on data deposited in the string database (v11.0). GO terms are clustered based on their affiliation to cellular components (CC), molecular functions (MF) or biological processes (BP). Note that no significantly enriched GO terms or local clusters were identified in each of the adult subgroups. Each term is followed by the number of genes contributing to it as well as the terms' size and its respective false discovery rate (FDR).*

Juvenile HVC			
Local cluster	**Description**	**Counts/Size**	**FDR**
9606_CL:616	mixed, incl. Axonal growth inhibition, (RHOA activation) and Intermediate filament head, DNA-binding domain	4/28	0.0225
Juvenile CMM			
BP term			
GO:0019731	antibacterial humoral response	3/10	0.0216
CC term			
GO:0005753	mitochondrial proton-transporting ATP synthase complex	3/18	0.0165
GO:0005743	mitochondrial inner membrane	6/331	0.0417
GO:0019866	organelle inner membrane	7/370	0.0417
GO:0031966	mitochondrial membrane	8/849	0.0417
GO:0031967	organelle envelope	10/830	0.0417
GO:0098800	inner mitochondrial membrane protein complex	4/94	0.0417
GO:0070469	respirasome	3/64	0.0417
Local cluster			
9606_CL:22327	Oxidative phosphorylation	6/109	0.0009
9606_CL:22571	proton-transporting ATP synthase complex	3/20	0.0057
Adult HVC			
no enriched GO terms or local clusters			
Adult CMM			
no enriched GO terms or local clusters			

Gene set enrichment analyses

Due to the low number of overlapping differentially expressed genes between groups which suggests large variability, gene set enrichment analyses (GSEA) was performed to allow next to GO term analyses for an additional, less biased perspective on the putative implications of all genes which were identified based on the mapped and assigned reads of each sample. During GSEA analyses, normalised counts of all assigned genes filtered for low expression were ranked and weighted based on their log2fold-change. Subsequently the association of a gene to a specific pathway elevated this pathways' normalised enrichment score (NES), while no known pathway contribution of a gene respectively lowered it. As samples segregated by age during injection and injected area during hierarchical clustering, GSEA was performed separately for each group (Fig. 5, Supplementary Table 5).

For each GSEA, the 50 genes with increased or decreased expression which contribute most to the outcome of the gene set enrichment analyses are shown in matrix plots for juvenile HVC (Fig. 5a), juvenile CMM (Fig. 5b), adult HVC (Fig. 5c), and adult CMM samples (Fig. 5d). Note that the downregulated *FoxP1* in knockdowns of this study contributes most to gene sets enriched in these samples across all groups.

In samples from juvenile HVC, gene set enrichment scores are bimodally distributed (Fig. 5e). This distribution indicates comparable numbers of gene sets with high and low enrichment scores where high scores correspond to enrichments in control samples and low scores suggest enrichment in knockdown samples, respectively. However, no gene set was enriched significantly in this group. In samples from juvenile CMM (Fig. 5f) enrichment scores are negatively biased with three significantly enriched gene sets (FDR < 0.25) in knockdown samples. The significantly enriched gene sets consist of genes which are upregulated in an epithelial cell line after stimulation with serum (FDR = 0.22, NES = -2.04, Amit *et al.*, 2007), genes which are upregulated in a cell line derived from colon cancer after expression of FOXO3 (FDR = 0.21, NES = -1.95, Delpuech *et al.*, 2007) and genes which are downregulated in amyloidosis plasma cells in comparison to multiple myeloma cells (FDR = 0.23, NES = -1.93, Abraham *et al.*, 2005). No gene sets are enriched significantly in any of the adult groups, yet gene sets of adult HVC samples are biased towards negative enrichment scores (Fig. 5g) whereas adult CMM samples indicate a bias towards positive gene set enrichment scores (Fig. 5h).

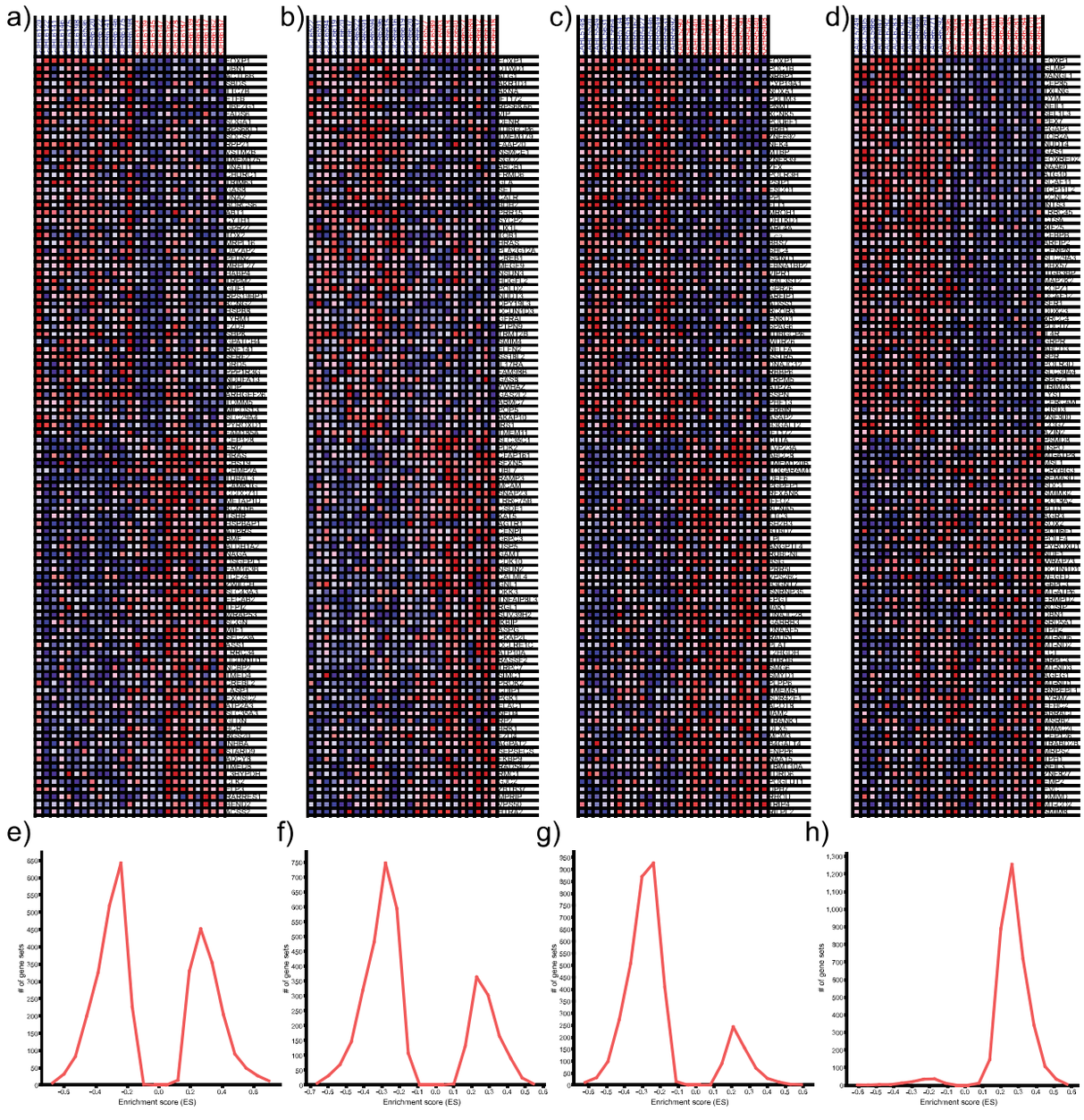


Figure 5: *Gene set enrichment analyses (GSEA) for each of the investigated subgroups. a) – d) show heatmaps of differentially expressed genes as they were ranked by GSEA based on their contribution to gene sets in samples from a) juvenile HVC, b) juvenile CMM, c) adult HVC and d) adult CMM (d). Each row represents one gene as indicated by its symbol on the right. Each column consists of the data generated from one sample as indicated at the top. Control samples are labelled in blue, knockdown samples in red, respectively. e) – h) show the respective enrichment scores for the number of gene sets identified in each subgroup consisting of samples from e) juvenile HVC, f) juvenile CMM, g) adult HVC or h) adult CMM. Negative scores*

represent enrichment in the knockdown samples while positive scores represent enrichment in controls.

Differentially expressed genes overlapping with a previous study of the striatum in heterozygous *FoxP1* knockout mice

64.5% (324) of all genes with increased expression and 46.1% (111) of all genes with decreased expression in this study (Supplementary Table 6) overlap significantly ($\chi^2(1, 104) = 1643.25$, $p < 0.001$) with genes previously identified in expression profiling experiments during a prior investigation of brain tissue samples from heterozygous *Foxp1* knockout mice (Araujo *et al.*, 2015). The authors of this study analysed differential gene expression in the striatum, the hippocampus and the neocortex of *Foxp1* heterozygous mice and compared the findings to consequences of *FOXP1* overexpression in human neural progenitor cells. The largest overlap of differentially expressed genes was identified between samples from mouse striatum and neural progenitors which suggests a higher level of module preservation in the striatum of mammals. Considering the direction of change, 21 genes consistently showed significant increases in expression and 18 genes consistently showed significant decreases in expression as a consequence of *FoxP1* knockout/knockdown across both studies (Table 7). However, in contrast to the overlap of genes irrespective of their direction of differential expression, neither upregulated ($\chi^2(1, 104) = 2.46$, $p > 0.05$) nor downregulated ($\chi^2(1, 104) = 0.93$, $p > 0.05$) genes of both studies significantly overlapped when analysed separately.

One of these overlapping genes shows significant increases in expression in both juvenile knockdown groups of this study and codes for Argininosuccinate synthetase 1 (*ASS1*). At least one human patient with a mutation in *ASS1* has also presented with speech delay (Lin *et al.*, 2019). Overlapping genes which show significant decreases in expression in both juvenile groups of this study are those coding for Beta-2-Microglobulin (*B2M*) and Testican-2 (*SPOCK2*). *B2M* expression increases with age in humans and mice, and artificially increased levels result in impaired performance in radial arm water mazes in mice whereas absence of B2M in mice leads to increased performance in the same type of maze (Smith *et al.*, 2015). SPOCK2 is a proteoglycan that is responsive to retinoic acid signalling in mice (Wei *et al.*, 2016). Another overlapping gene is *RPE65* of the retinoic acid signalling pathway which shows reduced expression in juvenile HVC knockdowns. Finally, two overlapping genes

(*MGST1*, *PTGR1*) contributing to the prostaglandin synthase pathway (Kelner *et al.*, 2000; Dick *et al.*, 2001) which in turn affects retinoic acid signalling (Ziboh *et al.*, 1975; Stock *et al.*, 2011) show increased expression in adult CMM knockdowns.

Table 7: *Genes with significant expression changes in this study which overlap in their direction with the findings of a previous study on differentially expressed genes in the striatum of heterozygous Foxp1 knockout mice (Araujo et al. 2015).*

Group	Regulation	Gene ID	Name and putative function
juvenile HVC & CMM	up	ASS1	Argininosuccinate Synthase 1, Citrullinemia
juvenile HVC	up	CDKN1A	Cyclin Dependent Kinase Inhibitor 1A, tissue regeneration
juvenile HVC	up	GFRA1	GDNF Family Receptor Alpha 1, neuron differentiation
juvenile HVC	up	OAF	Out at first homolog, Spondylocarpotarsal Synostosis Syndrome
juvenile HVC	up	PDLIM4	PDZ and LIM Domain4, bone development, osteoporosis
juvenile HVC	up	PERP	P53 apoptosis effector related to PMP22, Keratinization, desmosome junctions
juvenile HVC	up	SCGN	Secretagogin, calcium binding
juvenile CMM	up	ENOX1	Ecto-NOX Disulfide-Thiol Exchanger 1, plasma membrane electron transport
juvenile CMM	up	MID1	Midline 1, multiprotein formation, midline abnormalities
juvenile CMM	up	PCDH7	Protocadherin 7, cell-cell recognition and adhesion
adult CMM	up	DIO2	Iodothyronine Deiodinase 2, thyroid hormone pathway
adult CMM	up	FBLIM1	Filamin Binding LIM Protein 1, cell adhesion to actin
adult CMM	up	MGST1	Microsomal Glutathione S-Transferase 1, prostaglandin and inflammation
adult CMM	up	PFDN1	Prefoldin subunit1, chaperone
adult CMM	up	PTGR1	Prostaglandin Reductase 1, inflammation
adult CMM	up	RPL22L1	Ribosomal protein l22 like1, sarcoma
adult CMM	up	RPL37A	Ribosomal protein l37a, 60S subunit part
adult CMM	up	SDC1	Syndecan1, cell binding and signalling
adult CMM	up	ABRACL	ABRA C-Terminal Like, cleft lip
adult CMM	up	AQP1	Aquaporin 1, ocular fluid movement
juvenile HVC & CMM	down	B2M	Beta2Micoglobulin, MHC complex
juvenile HVC & CMM	down	SPOCK2	SPARC osteonectin, extracellular matrix, calcium binding
juvenile HVC	down	ACBD7	Acyl-CoA Binding Domain Containing 7, lipid metabolism
juvenile HVC	down	CD59	CD59 molecule, cell lysis

juvenile HVC	down	CRISPLD1	Cysteine rich secretory protein lccl domain containing 1
juvenile HVC	down	CYTH1	Cytohesin 1, membrane trafficking
juvenile HVC	down	MOK	MOK protein kinase, cell growth and differentiation
juvenile HVC	down	PLP1	Proteolipid protein 1, myelin component, oligodendrocyte development, axonal survival
juvenile HVC	down	PROCA1	Protein interacting with cyclin a1, calcium ion binding, gaucher disease
juvenile HVC	down	PRR16	Proline rich 16, cardiomyopathy
juvenile HVC	down	RPE65	Retinoid isomerohydrolase RPE65
juvenile HVC	down	RPS6KL1	Ribosomal protein S6 Kinase Like 1, transferase
juvenile CMM	down	FGFBP3	Fibroblast growth factor binding protein 3, gpcr signalling
juvenile CMM	down	HOPX	HOP homeobox, cardiac development
juvenile CMM	down	PTPN9	Protein tyrosine phosphatase non-receptor-type9, cell growth and differentiation
juvenile CMM	down	SOX2	SRY-box transcription factor 2, cell fate determination
adult HVC	down	FOXJ1	Cilia production, left right asymmetry
adult CMM	down	CHSY3	Chondroitin sulfate synthase 3, glucosyl metabolism

Differentially expressed genes overlapping with the SFARI database on putative ASD risk genes

Across all age-groups and regions, six genes which showed significantly increased expression and 27 genes which showed significantly decreased expression in response to *FoxP1* knockdown are listed as putative ASD risk genes in the SFARI database (Table 8) resulting in a significant overlap ($\chi^2(1, 104) = 3.96$, $p < 0.05$) between the differentially expressed genes of this study and putative ASD risk genes. At the time of this study, the SFARI database listed 1011 genes which are scored at four different levels based on the available evidence of a gene's relevance for ASD ranging from S (syndromic, highest) to 3 (suggestive evidence, lowest). The overlapping genes with increased expression were only identified in one of the tested groups of this study. Next to *FoxP1* which (as expected) showed reduced expression in all groups, one gene, *ETFB* (Electron transfer flavoprotein subunit beta) had significantly lower levels in juvenile HVC and adult CMM samples. As an electron transfer protein, *ETFB* is involved in the energy metabolism in mitochondria and mutations are linked to multiple acyl-CoA dehydrogenase deficiencies (Schiff *et al.*, 2006) which can result in slight speech delay (Chautard *et al.*, 2020) and neurodevelopmental disorder (Pollard *et al*., 2010) in infants.

Next to these genes overlapping between the SFARI database and more than one subgroup of these analyses, genes found to be differentially expressed in one of this studies' subgroups include *ROBO2*, *PLXNB1* and two glutamate-receptor interacting proteins *GRIP1* and *GRID1*. Neuronal homozygous deletions of *GRIP1* in mice impair synaptic plasticity and inhibitory avoidance learning and memory (Tan *et al.*, 2020). *GRID1* homozygous knockout mice demonstrate decreased social novelty preference of conspecifics and impaired memory in context specific fear learning, as well as lowered motivation in a forced swim test when compared to controls (Nakamoto *et al.*, 2020).

Table 8: *Genes with significant expression changes in this study which overlap with genes listed in the SFARI gene database focused on autism candidate genes. According to SFARI the listed genes are associated with the listed phenotypes: neurodevelopmental disorder (NDD), epilepsy (EP), autism spectrum disorder (ASD), intellectual disability (ID), attention deficit hyperactivity disorder (ADHD), developmental delay (DD), schizophrenia (SCHZ), bipolar disorder (BIP), mental retardation (MR), Down syndrome (DS). SFARI provides a score for each gene, ranging from 3 (suggestive evidence) to 2 (strong candidate) and 1 (high confidence) up to S (syndromic).*

Group	Regulation	Gene ID	Association	Score
juvenile HVC	up	SATB1	ND, EP	1
juvenile HVC	up	HRAS	ASD	1
juvenile CMM	up	SOX5	ASD, ID, ADHD	1
juvenile CMM	up	CNTN5	ASD, ID, ADHD, DD, NDD	2
juvenile CMM	up	AGMO	ASD	3
juvenile CMM	up	CDH10	ASD	3
juvenile CMM	up	DLGAP1	ASD	2
juvenile CMM	up	RORA	ASD, EP	S
juvenile CMM	up	KCND2	ASD	3
juvenile CMM	up	GRIP1	ASD	2
juvenile CMM	up	MYO16	ASD	3
juvenile CMM	up	CDH9	ASD	3
juvenile CMM	up	LRRC4C	ASD	1
juvenile CMM	up	CSMD1	ASD, SCHZ, BIP	3
juvenile CMM	up	DMD	ID, ADHD, EP, ASD,	S
juvenile CMM	up	GRID1	ASD	2
juvenile CMM	up	AUTS2	ASD, MR, ADHD, EP	1
juvenile CMM	up	GATM	ID, EP, ASD	S
juvenile CMM	up	GPC6	ASD	3
juvenile CMM	up	DSCAM	ID, down syndrome	1
juvenile CMM	up	IL1RAPL1	ID, DD, EP, ASD	3

juvenile CMM	up	ATP10A	Conflicting reports	2
juvenile CMM	up	PCDH9	ID, ASD	3
adult HVC	up	ROBO2	DD, ASD	2
adult HVC	up	PLXNB1	ASD	2
adult CMM	up	AP1S2	EP, ASD, MR	S
adult CMM	up	NDUFA5	ASD	3
juvenile HVC & CMM, adult HVC & CMM	down	FOXP1	ID, ASD, MR	1
juvenile HVC & CMM	down	ETFB	ASD	2
juvenile HVC	down	ACTL6B	ASD, EP, ID, DD	S
juvenile HVC	down	BCAS1	ASD	3
juvenile HVC	down	ATP1A1	ASD, DD, ID	3S
adult CMM	down	BICDL1	ASD	3

Differentially expressed genes overlapping with the SysID database on putative risk genes for intellectual disability

Due to implications of human FOXP1 malfunctions in intellectual disability, differentially expressed genes of this study were also compared with the SysID dataset (Kochinke *et al.*, 2016) which collects genes associated to intellectual disability in humans. The database consisted of 2778 human genes at the time of this study of which 75 genes (see Table 9) overlapped significantly ($\chi^2(1, 104) = 26.63$, $p < 0.0001$) with differentially expressed genes in females which received local *FoxP1* knockdowns. Among the overlapping genes were *ACTL6B*, *ASS1*, *PLXNB1*, *RGR*, *SEMA3E* and *TUBAL3* which were previously listed in this Chapter as potentially interesting candidate genes that might contribute to the phenotypical consequences following FoxP1 malfunctions or altered expression levels. In addition to these previously mentioned genes, *ILRAPL1* is an overlapping gene which is significantly upregulated in juvenile CMM knockdowns. This gene encodes an interleukin 1 receptor accessory protein and its homozygous knockout in mice results in reduced dendritic spine density in cortical layer 2/3 and CA1 of the hippocampus. The same mice also show impaired spatial reference memory, working memory, fear learning and motor learning while they simultaneously present with increased social interaction when compared to controls (Yasumura *et al.*, 2014).

Table 9: *Genes with significant expression changes in this study which overlap with the SysID database on genes mutated in intellectual disability.*

Group	Regulation	Gene ID	Name and putative function
juvenile HVC	up	ALDH1A2	Aldehyde Dehydrogenase 1 Family Member A2, Retinoic acid synthesis
juvenile HVC	up	GLS	Glutaminase, Glutamate synthesis
juvenile HVC	up	HRAS	HRas Proto-Oncogene GTPAse, cell division
juvenile HVC	up	SATB1	Special AT-rich sequence-binding protein-1, Chromatin accessibility
juvenile HVC	up	UBR7	Ubiquitin Protein Ligase E3 Component N-Recognin 7, Ubiquitinylation
juvenile HVC	up	WRAP53	WD Repeat Containing Antisense To TP53, telomere synthesis
juvenile HVC	up	CHMP2A	Charged Multivesicular Body Protein 2A, chromatin modification
juvenile HVC	up	FH	Fumarate Hydratase, tricarboxylic acid cycle
juvenile HVC	up	RGR	Retinal G Protein Coupled Receptor, retinal conversion
juvenile HVC	up	SCGN	Secretagogin, calcium binding
juvenile HVC*	up	TUBAL3	Tubulin Alpha Like 3, SLIT-ROBO signalling
juvenile HVC	up	WWP2	WW Domain Containing E3 Ubiquitin Protein Ligase 2, ubiquitination
juvenile CMM	up	ADGRB3	Adhesion G Protein-Coupled Receptor B3, angiogenesis
juvenile CMM	up	AGMO	Alkylglycerol Monooxygenase, Kleefstra Syndrome 2
juvenile CMM	up	AUTS2	Activator Of Transcription And Developmental Regulator AUTS2, ASD
juvenile CMM	up	CNTN5	Contactin 5, nervous system development
juvenile CMM	up	DLGAP1	DLG Associated Protein 1, protein-protein interaction at synapses
juvenile CMM	up	DMD	Dystrophin, cytoskeleton
juvenile CMM	up	DSCAM	DS Cell Adhesion Molecule, nervous system development
juvenile CMM	up	GATM	Glycine Amidinotransferase, creatine biosynthesis
juvenile CMM	up	IL1RAPL1	Interleukin 1 Receptor Accessory Protein Like 1, synapse formation
juvenile CMM	up	KCND2	Potassium Voltage-Gated Channel Subfamily D Member 2, potassium channel
juvenile CMM	up	LRP1B	LDL Receptor Related Protein 1B, cellular metabolism
juvenile CMM	up	MID1	Midline 1, multiprotein structure formation
juvenile CMM	up	NPAS3	Neuronal PAS Domain Protein 3, transcription factor
juvenile CMM	up	NRG3	Neuregulin 3, tyrosine kinase receptor
juvenile CMM	up	PDE10A	Phosphodiesterase 10A, nucleotide phosphodiesterase
juvenile CMM	up	RORA	RAR Related Orphan Receptor A, nuclear hormone receptor
juvenile CMM	up	SLC35C1	Solute Carrier Family 35 Member C1, GDP-fucose transporter
juvenile CMM	up	SNTG1	Syntrophin Gamma 1, gamma-enolase trafficking to plasma membrane
juvenile CMM	up	SOX5	SRY-Box Transcription Factor 5, embryonic development
juvenile HVC, juvenile CMM	up	ASS1	Argininosuccinate Synthase 1, Citrullinemia
adult HVC	up	PLXNB1	Plexin B1, axon guidance
adult CMM	up	AP1S2	Adaptor Related Protein Complex 1 Subunit Sigma 2, clathrin recruitment
adult CMM	up	ASPA	Aspartoacylase, white matter maintenance
adult CMM	up	ATP5F1E	ATP Synthase F1 Subunit Epsilon, mitochondrial ATP synthase
adult CMM	up	ATP5PF	ATP Synthase Peripheral Stalk Subunit F6, mitochondrial ATP synthase
adult CMM	up	BOLA3	BolA Family Member 3, mitochondrial respiratory chain complex assembly
adult CMM	up	COX7B	Cytochrome C Oxidase Subunit 7B, mitochondrial respiratory chain
adult CMM	up	EEF1B2	Eukaryotic Translation Elongation Factor 1 Beta 2, guanine nucleotide exchange

adult CMM	up	GCSH	Glycine Cleavage System Protein H, methylamine group transfer
adult CMM	up	ISCA1	Iron-Sulfur Cluster Assembly 1, iron-sulfur cluster biogenesis
adult CMM	up	LYRM7	LYR Motif Containing 7, mitochondrial respiratory chain
adult CMM	up	MICOS13	Mitochondrial Contact Site And Cristae Organizing System Subunit 13, oxidative phosphorylation
adult CMM	up	NDUFA12	NADH:Ubiquinone Oxidoreductase Subunit A12, mitochondrial membrane respiratory chain
adult CMM	up	NDUFA2	NADH:Ubiquinone Oxidoreductase Subunit A2, mitochondrial membrane respiratory chain
adult CMM	up	NDUFA4	NADH:Ubiquinone Oxidoreductase Subunit A4, mitochondrial membrane respiratory chain
adult CMM	up	NDUFB8	NADH:Ubiquinone Oxidoreductase Subunit B8, mitochondrial membrane respiratory chain
adult CMM	up	NDUFS4	NADH:Ubiquinone Oxidoreductase Subunit S4, mitochondrial membrane respiratory chain
adult CMM	up	NPRL3	NPR3 Like, GATOR1 Complex Subunit, epilepsy
adult CMM	up	PSMG4	Proteasome Assembly Chaperone 4, chaperone
adult CMM	up	RPLP1	Ribosomal Protein Lateral Stalk Subunit P1, ribosome component
adult CMM	up	RPS23	Ribosomal Protein S23, ribosome component
adult CMM	up	SNAPIN	SNAP Associated Protein, vesicle docking and fusion
adult CMM	up	SOX2	SRY-Box Transcription Factor 2, embryonic development
adult CMM	up	SVBP	Small Vasohibin Binding Protein, neurodevelopmental disorder
juvenile HVC	down	ACTL6B	Actin Like 6B, cytoskeleton
juvenile HVC	down	ATP1A1	ATPase Na+/K+ Transporting Subunit Alpha 1, cation transportin ATPase
juvenile HVC	down	HEATR5B	HEAT Repeat Containing 5B
juvenile HVC	down	NDP	Norrin Cystine Knot Growth Factor NDP, Wnt/beta-catenin pathway activation
juvenile HVC	down	NDUFA1	NADH:Ubiquinone Oxidoreductase Subunit A1, mitochondrial membrane respiratory chain
juvenile HVC	down	NDUFA13	NADH:Ubiquinone Oxidoreductase Subunit A13, mitochondrial membrane respiratory chain
juvenile HVC	down	PIGP	Phosphatidylinositol Glycan Anchor Biosynthesis Class P, down syndrome
juvenile HVC	down	PLP1	Proteolipid Protein 1, oligodendrocyte development and axonal survival
juvenile HVC	down	PROCA1	Protein Interacting With Cyclin A1, calcium ion binding
juvenile HVC	down	SBDS	SBDS Ribosome Maturation Factor, ribosome biogenesis
juvenile CMM	down	ALG1	ALG1 Chitobiosyldiphosphodolichol Beta-Mannosyltransferase, oligosaccaride biosnythesis
juvenile CMM	down	SEMA3E	Semaphorin 3E, axon guidance ligand
juvenile HVC & CMM	down	EIF2B5	Eukaryotic Translation Initiation Factor 2B Subunit Epsilon, vanishing white matter
adult HVC	down	DHTKD1	Dehydrogenase E1 And Transketolase Domain Containing 1, amino acid degradation
adult HVC	down	FOXJ1	Forkhead Box J1, transcription factor
adult CMM	down	SPR	Sepiapterin Reductase, DOPA-responsive dystonia
juvenile HVC & CMM, adult HVC & CMM	down	FOXP1	Forkhead Box P1, transcription factor

Discussion

This study aimed to investigate the transcriptional differences following local *FoxP1* knockdowns in HVC and CMM of juvenile and adult female zebra finches. Even though not all the generated samples matched the quality criteria, the number of mapped and assigned reads was comparable between groups. Variance, mean and distribution of read counts as well as log 2fold changes did not differ visibly between the different treatment groups of this study. However, the dispersion of log2 fold changes across all genes identified in samples from birds injected in HVC as adults differed visibly between controls and knockdowns. Knockdown samples were more dispersed when compared to controls, which could be the result of variable knockdown efficiency across samples from this particular group. Different log2 dispersion does not result in read count bias when replicates within a group consist of unrelated or genetically distant samples as was the case in this study. However, dispersion is also affected by the presence of a large number of genes with a low count, which could be the case in this group as samples taken from adult HVC yielded the highest number of assigned genes which might be represented by a low number of counts (Yoon and Nam, 2017). Even though *FoxP1* knockdown efficiency varied across samples, *FoxP1* was the only gene to show significant reductions in expression in knockdowns of all groups when compared to their matched controls. Across all groups tested during this study, knockdown efficiency differed between previous qPCR analyses and the results from RNAseq, but this difference was not significant on the level of different subgroups. The assessment of relative expression levels during qPCR and total transcript counts during RNAseq analyses might account for this methodological difference. Besides *FoxP1*, no gene showed significantly altered expression across all the different knockdown groups. This suggests that region-specific but probably also interindividual differences outweigh common transcriptional changes across knockdowns. Substantive interindividual differences are also supported by the results from hierarchical clustering of samples, where samples clustered according to age during treatment and injected area but no further segregation between controls and knockdowns was visible on group level. Interindividual variability cannot be explained by activity-regulated genes as all samples were obtained in silence prior to light onset early in the morning, excluding immediate effects on different gene expression levels. However, variable knockdown efficiency and general variability in gene expression

levels between individuals with different degrees of relatedness could account for large interindividual differences.

Even though little overlap occurred between the groups of this study regarding differentially expressed genes, some GO terms and local network clusters were enriched in multiple groups. Additionally, genes with comparable functional implications were found to be differentially expressed in knockdowns injected in different areas during different developmental stages.

In all groups except for adult CMM, genes related to retinoic acid signalling, synthesis or other retinal proteins were among the genes showing the most significant increases in expression in response to *FoxP1* knockdown. Among the genes showing the most significant decreases in adult CMM knockdowns, one unannotated locus LOC100229421 is suspected to code for IFIT5 (Scalf, 2018), a retinoic acid and interferon inducible protein. Moreover, local network clusters related to retinoic acid signalling were enriched among genes showing increased expression in juvenile or adult HVC knockdowns. Taken together, these findings suggest that FoxP1 might be linked to retinoic acid signalling, possibly as a heterodimer with FoxP2 (Li *et al.*, 2004; Roeske *et al.*, 2014; Mendoza and Scharff, 2017) which is only weakly expressed in songbird HVC (Teramitsu *et al.*, 2004; Mendoza *et al.*, 2015) but has been shown to interact with the retinoic acid signalling pathway and thereby helps regulate neuronal differentiation (Devanna *et al.*, 2014; Van Rhijn and Vernes, 2015; Negwer and Schubert, 2017).

In addition to the putative gene *IFIT5* showing elevated expression in adult CMM knockdowns, other interferon-regulated genes are differentially expressed across knockdowns of all groups. Interferon signalling can be related to retinoic acid signalling as both pathways are linked and possibly potentiate each other (Pelicano *et al.*, 1997; Chelbi-Alix and Pelicano, 1999). Transcripts of interferon-related genes are among the most significantly reduced by *FoxP1* knockdown in samples of all groups except juvenile CMM where transcripts of one gene coding for an interferon-inducible protein is among the transcripts showing most significant increases. Next to the interferon signalling pathway, two more genes which show elevated expression in adult CMM knockdowns (*MGST1*, *PTGR1*) might indirectly contribute to retinoic acid related processes via prostaglandin signalling (Kelner *et al.*, 2000; Dick *et al.*, 2001). Prostaglandins have been shown to inhibit neuronal correlates of mate calling in frogs

(Schmidt and Kemnitz, 1989) and interact with retinoic acid signalling by suppressing retinoic acid synthesis (Stock *et al.*, 2011) which in turn stimulates prostaglandin production (Kim *et al.*, 2008).

Even though differentially expressed genes related to retinoic acid signalling or connected pathways were detected across all treatment groups, the behavioural changes following lentiviral *FoxP1* knockdowns were limited to adult HVC (Chapter 2). Perhaps FoxP1 and its contributions to retinoic acid signalling may be especially impactful in this area and developmental stage, as the retinoic acid synthesising enzyme zRalDH is highly expressed in HVC but not in CMM (Denisenko-Nehrbass *et al.*, 2000; Olson *et al.*, 2011) where only retinoic acid receptors are expressed (Roeske *et al.*, 2014). Dietary supplementation of retinoic acid (Wood *et al.*, 2008) or blockage of retinoic acid synthesis in HVC (Denisenko-Nehrbass *et al.*, 2000) during the critical learning phase of juvenile male zebra finches leads to more variable songs in adults.

Another possibly relevant group of genes showing differential expression in knockdown samples consists of loci related to SLIT-ROBO signalling. Among the most significant increases in expression in both juvenile and adult HVC knockdown samples was at least one gene associated to SLIT-ROBO signalling, and one gene of this pathway is also among the those showing the most significant reductions in juvenile CMM knockdown samples. Proteins of the SLIT-ROBO signalling pathway have been identified as downstream targets of human FOXP2 in vitro (Vernes *et al.*, 2007a; Konopka *et al.*, 2009), and binding partners of FoxP1 in zebra finches (Mendoza and Scharff, 2017). The SLIT-ROBO signalling pathway has been implicated in human language-related impairments (Hannula-Jouppi *et al.*, 2005; Bates *et al.*, 2010; Suda *et al.*, 2011; Pourcain *et al.*, 2014; Mozzi *et al.*, 2016) and its proteins show convergent substitutions and expression levels in vocal learning mammals (Wang *et al.*, 2015). Genes related to SLIT-ROBO signalling are also enriched in HVC of juvenile (45 days post hatch) and adult (100 days post hatch) male zebra finches (Shi *et al.*, 2021). Convergence between the avian and human orthologs of this pathway has been suggested based on differential regulation of *SLIT1* in RA of zebra finches and human laryngeal motor cortex (Pfenning *et al.*, 2014).

In addition to individual genes of specific pathways, the significant overlap with gene expression data from striatal neurons in mice with heterozygous knockout of *Foxp1* (Araujo *et al.*, 2015) further emphasizes that this transcription may regulate similar

molecular and cellular mechanisms in different species. Even though mice do not need to learn how to produce their vocalisations (Hammerschmidt *et al.*, 2012; Screven and Dent, 2019), female mice can discriminate contextual differences of male song (Hammerschmidt *et al.*, 2009; Chabout *et al.*, 2015) and develop preferences for specific songs by imprinting (Asaba *et al.*, 2014).

Next to overlaps with differentially expressed genes in mice following genetic *Foxp1* manipulations, significant subsets of genes which were differentially expressed in groups of this study are also listed as putative risk genes involved in autism spectrum disorder in the SFARI database or the SysID database on genes mutated in intellectual disability. This pattern is consistent with the involvement of *FOXP1* in phenotypes related to ASD and ID (Sollis *et al.*, 2016; Co *et al.*, 2020a).

Taken together, the potential regulation of genes related to retinoic acid, interferon, prostaglandin, SLIT-ROBO signalling, and orthologues of putative genes related to ASD-risk genes by FoxP1 in female zebra finches might enhance our understanding of the in vivo functions of this transcription factor in the songbird brain. The data presented here could be helpful for gaining new insights into how FoxP1 contributes to song motor control and auditory perception and memory in different brain areas during song production learning in male zebra finches (Norton *et al.*, 2019; Garcia-Oscos *et al.*, 2021) and perception in females (Chapter 2).

Another possibly relevant gene which shows significantly reduced expression in adult CMM knockdowns is *ETFB*. The protein that this gene encodes is implicated in energy metabolism of mitochondria, which could perhaps be related to the large amount of differentially expressed genes related to mitochondrial processes next to genes implicated in ribosomal processes in this group. This pattern of findings could be the result of biased knockdown-specific effects in CMM of adult birds, since differentially expressed genes related to energy metabolism in the mitochondria or ribosomal activity were also present in the other groups albeit at a smaller rate. Another possible contributory factor might be the different amounts of tissue that went into the RNA preparations of different groups. Tissue punches for HVC were placed at the dorsal edge of the brain resulting in lower amounts of tissue compared to CMM samples, where a biopsy punch was taken more centrally. However, as both juvenile and adult treated birds were sacrificed as adults, both groups should result in comparable differentially expressed genes unless the birds' age during the *FoxP1* knockdown

affects mitochondrial and ribosomal-related gene expression differentially. Interestingly, a recent study shows mitochondrial dysfunction in the striatum of heterozygous *Foxp1* knockout mice (Wang *et al.*, 2021) suggesting a possible contribution of altered energy supply and oxidative stress to *FoxP1*-related phenotypes.

In summary, gene expression analyses of samples generated from birds which received *FoxP1* knockdowns in HVC or CMM during different developmental stages show interesting convergences with previous studies on transcriptional differences following manipulations of this gene in other species and pathways relevant to FoxPs and vocalisation behaviours. Even though female zebra finches do not learn to produce a song of their own, FoxP1 might be implicated in similar pathways and mechanisms in both sexes. To further validate potential contributions of FoxP1 to pathways identified in this study, putative regulation of target genes should be experimentally verified. Overall, the expression profiling data from this study provide a valuable resource for further deciphering conserved roles of FoxP1 in vocalisation (and related) behaviours in diverse species, ranging from vocal learning in songbirds to speech and language in humans.

Conclusion

The data from this chapter suggest that, despite large interindividual and group-based differences in gene expression, the contributions of *FoxP1* to the regulation of specific pathways show some intriguing overlaps across the targeted brain regions and ages during treatment. The knockdown target itself, *FoxP1*, was the only individual gene to show significantly different expression across all the groups studied. However, analyses of differentially expressed genes with respect to enrichment of GO terms, gene sets and local networks identified a number of processes that had been previously associated directly or indirectly to FoxP1. Highlighted pathways include retinoic acid signaling or SLIT-ROBO signaling. A significant number of differentially expressed genes overlapped between this research and a study on striatal gene expression in FoxP1 knockout mice. Further, differentially expressed genes identified in this chapter overlap with databases on genes implicated in autism spectrum disorders or intellectual disability which are both associated with human FOXP1 mutations. Taken together, the results from this study can contribute to the

understanding of downstream effects which are influenced by *FoxP1* across different species and may also help to understand the molecular underpinnings of vocal learning at the basis of human speech and language.

Appendix Chapter 4

Supplementary Table 1: *Samples contributing to this analysis where sufficient RNA was obtained for RNAseq analyses. Individual bird ID encodes treatment group, target area, age group and hemisphere RNA was obtained from. shRNA type identifies the shRNA which was virally transduced in each bird. Batch date corresponds to the date each virus batch was produced and Seq. batch indicates the batch in which each sample was sent for total RNA sequencing.*

Bird ID	Treatment group	Age group	Target area	Hemi-sphere	shRNA type	Batch date	Seq. Batch
CJHL6120	Control	Juvenile	HVC	Left	shCtrl	13.1.2017	3
CJHL6122	Control	Juvenile	HVC	Left	shCtrl	13.1.2017	3
CJHL6141	Control	Juvenile	HVC	Left	shCtrl	13.1.2017	3
CJHL6146	Control	Juvenile	HVC	Left	shCtrl	3.2.2017	3
CJHL6175	Control	Juvenile	HVC	Left	shCtrl	3.2.2017	3
CJHL6198	Control	Juvenile	HVC	Left	shCtrl	3.2.2017	3
CJHL6536	Control	Juvenile	HVC	Left	shCtrl	3.2.2017	3
CJHR6120	Control	Juvenile	HVC	Right	shCtrl	13.1.2017	3
CJHR6122	Control	Juvenile	HVC	Right	shCtrl	13.1.2017	3
CJHR6141	Control	Juvenile	HVC	Right	shCtrl	13.1.2017	3
CJHR6146	Control	Juvenile	HVC	Right	shCtrl	3.2.2017	3
CJHR6175	Control	Juvenile	HVC	Right	shCtrl	3.2.2017	3
CJHR6198	Control	Juvenile	HVC	Right	shCtrl	3.2.2017	3
KJHL6123	Knockdown	Juvenile	HVC	Left	shKRAK	13.1.2017	3
KJHL6139	Knockdown	Juvenile	HVC	Left	shKRAK	13.1.2017	3
KJHL6145	Knockdown	Juvenile	HVC	Left	shY	13.1.2017	3
KJHL6147	Knockdown	Juvenile	HVC	Left	shKRAK	13.1.2017	3
KJHL6162	Knockdown	Juvenile	HVC	Left	shY	13.1.2017	3
KJHL6178	Knockdown	Juvenile	HVC	Left	shY	13.1.2017	3
KJHL6187	Knockdown	Juvenile	HVC	Left	shY	13.1.2017	3
KJHR6139	Knockdown	Juvenile	HVC	Right	shKRAK	13.1.2017	3
KJHR6145	Knockdown	Juvenile	HVC	Right	shY	13.1.2017	3
KJHR6147	Knockdown	Juvenile	HVC	Right	shKRAK	13.1.2017	3
KJHR6162	Knockdown	Juvenile	HVC	Right	shY	13.1.2017	3
KJHR6187	Knockdown	Juvenile	HVC	Right	shY	13.1.2017	3
CAHL5148	Control	Adult	HVC	Left	shCtrl	10.10.2014	1
CAHL5424	Control	Adult	HVC	Left	shCtrl	10.10.2014	1
CAHL5520	Control	Adult	HVC	Left	shCtrl	10.10.2014	2
CAHL5549	Control	Adult	HVC	Left	shCtrl	8.8.2015	2
CAHL5631	Control	Adult	HVC	Left	shCtrl	19.4.2013	3
CAHL5664	Control	Adult	HVC	Left	shCtrl	19.4.2013	3
CAHR5134	Control	Adult	HVC	Right	shCtrl	10.10.2014	1
CAHR5148	Control	Adult	HVC	Right	shCtrl	16.5.2014	1

CAHR5162	Control	Adult	HVC	Right	shCtrl	16.5.2014	1
CAHR5427	Control	Adult	HVC	Right	shCtrl	16.5.2014	1
CAHR5446	Control	Adult	HVC	Right	shCtrl	19.4.2013	1
CAHR5549	Control	Adult	HVC	Right	shCtrl	8.8.2015	3
CAHR5631	Control	Adult	HVC	Right	shCtrl	19.4.2013	3
CAHR5692	Control	Adult	HVC	Right	shCtrl	19.4.2013	3
KAHL5290	Knockdown	Adult	HVC	Left	shKRAK	10.10.2014	1
KAHL5426	Knockdown	Adult	HVC	Left	shKRAK	16.5.2014	1
KAHL5440	Knockdown	Adult	HVC	Left	shY	6.6.2014	1
KAHL5408	Knockdown	Adult	HVC	Left	shKRaK	10.10.2014	2
KAHL5447	Knockdown	Adult	HVC	Left	shY	10.10.2014	2
KAHL5521	Knockdown	Adult	HVC	Left	shKRAK	10.10.2014	2
KAHL5523	Knockdown	Adult	HVC	Left	shKRAK	10.10.2014	2
KAHR5203	Knockdown	Adult	HVC	Right	shRKAK	16.5.2014	1
KAHR5290	Knockdown	Adult	HVC	Right	shKRAK	10.10.2014	1
KAHR5426	Knockdown	Adult	HVC	Right	shKRAK	16.5.2014	1
KAHR5440	Knockdown	Adult	HVC	Right	shY	6.6.2014	1
KAHR5408	Knockdown	Adult	HVC	Right	shKRaK	10.10.2014	2
KAHR5447	Knockdown	Adult	HVC	Right	shY	10.10.2014	2
KAHR5521	Knockdown	Adult	HVC	Right	shKRAK	10.10.2014	2
KAHR5542	Knockdown	Adult	HVC	Right	shY	10.10.2014	3
CJCL6522	Control	Juvenile	CMM	Left	shCtrl	3.2.2017	3
CJCL6591	Control	Juvenile	CMM	Left	shCtrl	21.7.2017	3
CJCL6594	Control	Juvenile	CMM	Left	shCtrl	3.2.2017	3
CJCL6619	Control	Juvenile	CMM	Left	shCtrl	3.2.2017	3
CJCL6620	Control	Juvenile	CMM	Left	shCtrl	21.7.2017	3
CJCL6637	Control	Juvenile	CMM	Left	shCtrl	21.7.2017	3
CJCR6522	Control	Juvenile	CMM	Right	shCtrl	3.2.2017	3
CJCR6591	Control	Juvenile	CMM	Right	shCtrl	21.7.2017	3
CJCR6594	Control	Juvenile	CMM	Right	shCtrl	3.2.2017	3
CJCR6596	Control	Juvenile	CMM	Right	shCtrl	3.2.2017	3
CJCR6615	Control	Juvenile	CMM	Right	shCtrl	21.7.2017	3
CJCR6616	Control	Juvenile	CMM	Right	shCtrl	21.7.2017	3
CJCR6619	Control	Juvenile	CMM	Right	shCtrl	3.2.2017	3
CJCR6620	Control	Juvenile	CMM	Right	shCtrl	21.7.2017	3
CJCR6637	Control	Juvenile	CMM	Right	shCtrl	21.7.2017	3
KJCL6527	Knockdown	Juvenile	CMM	Left	shKRAK	21.7.2017	3
KJCL6589	Knockdown	Juvenile	CMM	Left	shKRAK	21.7.2017	3
KJCL6593	Knockdown	Juvenile	CMM	Left	shY	21.7.2017	3
KJCL6618	Knockdown	Juvenile	CMM	Left	shY	21.7.2017	3
KJCL6640	Knockdown	Juvenile	CMM	Left	shY	21.7.2017	3
KJCR6527	Knockdown	Juvenile	CMM	Right	shKRAK	21.7.2017	3
KJCR6589	Knockdown	Juvenile	CMM	Right	shKRAK	21.7.2017	3
KJCR6612	Knockdown	Juvenile	CMM	Right	shY	21.7.2017	3

KJCR6617	Knockdown	Juvenile	CMM	Right	shKRAK	21.7.2017	3
KJCR6618	Knockdown	Juvenile	CMM	Right	shY	21.7.2017	3
KJCR6640	Knockdown	Juvenile	CMM	Right	shY	21.7.2017	3
CACL5749	Control	Adult	CMM	Left	shCtrl	21.7.2017	2
CACL5805	Control	Adult	CMM	Left	shCtrl	14.2.2016	2
CACL5866	Control	Adult	CMM	Left	shCtrl	3.2.2016	2
CACL6047	Control	Adult	CMM	Left	shCtrl	21.7.2017	3
CACL6282	Control	Adult	CMM	Left	shCtrl	3.2.2017	3
CACL6283	Control	Adult	CMM	Left	shCtrl	3.2.2017	3
CACL6292	Control	Adult	CMM	Left	shCtrl	21.7.2017	3
CACR5749	Control	Adult	CMM	Right	shCtrl	21.7.2017	2
CACR5866	Control	Adult	CMM	Right	shCtrl	3.2.2016	2
CACR6260	Control	Adult	CMM	Right	shCtrl	3.2.2017	3
CACR6271	Control	Adult	CMM	Right	shCtrl	21.7.2017	3
CACR6292	Control	Adult	CMM	Right	shCtrl	21.7.2017	3
KACL5806	Knockdown	Adult	CMM	Left	shKRAK	10.10.2014	2
KACL5808	Knockdown	Adult	CMM	Left	shY	16.6.2014	2
KACL6281	Knockdown	Adult	CMM	Left	shY	21.7.2017	3
KACL6284	Knockdown	Adult	CMM	Left	shY	3.2.2017	3
KACL6320	Knockdown	Adult	CMM	Left	shKRAK	3.2.2017	3
KACL6413	Knockdown	Adult	CMM	Left	shKRAK	21.7.2017	3
KACR5891	Knockdown	Adult	CMM	Right	shY	21.7.2017	2
KACR6240	Knockdown	Adult	CMM	Right	shY	21.7.2017	3
KACR6245	Knockdown	Adult	CMM	Right	shY	21.7.2017	3
KACR6281	Knockdown	Adult	CMM	Right	shY	21.7.2017	3
KACR6284	Knockdown	Adult	CMM	Right	shY	3.2.2017	3
KACR6320	Knockdown	Adult	CMM	Right	shKRAK	3.2.2017	3
KACR6413	Knockdown	Adult	CMM	Right	shKRAK	21.7.2017	3

Supplementary Table 2: *Normalised read counts of all mapped genes across all samples of this study* [access via: https://doi.org/10.17026/dans-xux-y5ja].

Supplementary Table 3: *Differentially expressed genes for all subgroups of this study* [access via: https://doi.org/10.17026/dans-zg3-qvba].

Supplementary Table 4: *Extended list of enriched GO terms and local clusters based on significantly upregulated genes in adult CMM knockdowns* [access via: https://doi.org/10.17026/dans-zr3-eutj].

Supplementary Table 5: *Gene set enrichment analyses data for all subgroups of this study based on GSEA 4.11* [access via: https://doi.org/10.17026/dans-23p-7626].

Supplementary Table 6: *Differentially expressed genes across all subgroups of this study that overlap with a previous study by Araujo et al. (2015)* [access via: https://doi.org/10.17026/dans-2bj-ks3v].