Supporting Information

SI Materials and Methods

Dissections and RNA extraction

We dissected the following regions from the brains of eight adult (postnatal day 56) male mice (C57BL/6J strain): dorsal cortex layers I-VIb and lateral cortex layers I-VIb (both are the samples reported in Belgard *et al.* (1)), hippocampus, claustrum-endopiriform complex, pallial amygdala (basolateral and basomedial nuclei), and striatum (**Fig. S1**). The mice were killed by cervical dislocation according to approved schedule one UK Home Office guidelines (Scientific Procedures Act, 1986). The mice were decapitated, the skull opened in the midline and the brain removed. Dissected brains were rinsed in RNAse free PBS, submerged in ice-cold RNAlater (Ambion) for 24 hours and stored at -20°C in RNAlater (Ambion). Whole brains were embedded in 5% agarose (low melting, Bioline) and sectioned using a vibrating microtome (Leica, VT1000S) into 200 µm coronal sections using a chilled solution of 1:1 mixture of RNAlater and PBS. Samples were dissected out using microsurgical scalpels (Weck, USA) under visual guidance, using transillumination on a dissecting microscope (MZFLIII, Leica) and stored separately in RNAlater at -80°C until all microdissection was complete.

We dissected the following regions from the brains of eighteen twelve-week-old chickens of mixed sex: arcopallium, dorsolateral cortex (dorsolateral corticoid area), hippocampus, mesopallium, nidopallium, hyperpallium, and striatum (formerly the

paleostriatum augmentatum; not including the globus pallidus) (Fig. S2) using a stererotaxic atlas (2). We removed the brains from freshly decapitated heads of mixed-sex, organic, free-range chickens obtained from a slaughterhouse engaged in its normal operations. Brains for *in situ* hybridization were rinsed with chilled RNAse free PBS, then embedded in OCT and frozen on dry ice. Brains destined for immunohistochemistry were fixed in 4% paraformaldehyde. Brains for RNA extraction were rinsed in RNAse-free PBS, submerged in ice-cold RNAlater (Ambion), transported to the laboratory and stored at -20˚C in RNAlater (Ambion). Whole brains were embedded in 5% agarose and sectioned using a vibrating microtome (Leica, VT1000S) into 200 μm coronal sections using a 1:1 mixture of RNAlater and PBS. The samples were dissected out under visual guidance, using transillumination on a dissecting microscope (MZFLIII, Leica) and stored separately in RNAlater at -80°C until all microdissection was complete.

For RNA extraction, samples from individual regions from the eighteen chickens and the eight mice, respectively, were combined and all tissue samples were processed concurrently. We extracted total RNA using the RNeasy Lipid Tissue Mini kit (QIAGEN), following the manufacturer's instructions and using the on-column DNase digest. The silica column is reported by the manufacturer to deplete transcripts smaller than 200 nt. RNA quantity was assessed using a NanoDrop 1000 spectrophotometer (ThermoScientific) (Table S1; Table S2), and, for the chicken samples, RNA quality and integrity assessed using a Bioanalyzer (Agilent Laboratories) (Table S2).

Sequencing and gene expression quantification

We prepared the chicken samples for paired-end sequencing on Illumina's Genome Analyzer IIx by following the standard Illumina RNA-seq library preparation protocol with one round of poly(A) selection. We did the same for mouse, but with two rounds of poly(A) selection. Both used the standard Illumina library preparation protocol.

Next, we filtered and mapped the reads. We computationally trimmed all reads to 50 nt for analysis and discarded three lanes with wildly aberrant GC content. The internal insert size and standard deviation were empirically estimated for each library as described by Belgard *et al.* (1) (Table S3). We required both read pairs to pass Illumina's chastity filter to be used in alignment. Using this insert size and standard deviation, each lane was separately mapped with tophat (3) v1.2.0 to the reference genome of mouse (NCBIM37, downloaded from Ensembl, masking the Y PAR but not simple repeats) or chicken (WASHUC2, downloaded from Ensembl, not masking simple repeats) as appropriate, using the --GTF option with corresponding GTF files from Ensembl release 61 (Table S3) (ENSMUST00000127664 was manually removed from the mouse file as it is absurdly long and interferes with mapping and quantification). The indel search was enabled for chicken, with 3 as the maximum insertion or deletion size, to allow more accurate quantification of genes containing indel polymorphisms. This was not necessary with the laboratory mice, since they all belonged to the C57BL/6J inbred strain that was used for the mouse genome reference. The minimum isoform fraction filter was disabled, as was the search for novel junctions.

For quantification purposes, and to test for potential confounds, we estimated the pre-mRNA fraction in each sample (Table S9). To do this, we calculated the proportion of

read fragments overlapping more than 3 nt of an intron for every Ensembl protein-coding gene with >200 nt exonic length having one and only one annotated transcript, and then took the median of these values for which there were at least 100 reads mapping in the locus. We used the same pre-mRNA fraction for the cufflinks quantification in each species: 10% for chicken quantifications and 5% for mouse quantifications, since those correspond to the highest values for each species rounded up to the nearest percent. The lower value in mouse may be attribuTable to the fact that the mouse samples underwent two rounds of poly(A) selection, while the chicken samples only underwent one, or to the relative incompleteness of the chicken gene models.

To normalize across libraries within a species, and to equalize the variance for genes of similar expression levels, the total number of sequenced fragments was then downsampled as follows: for each protein-coding gene model (having at least one transcript with total exonic length >200 nt, as the experimental methods selected against these), the number of fragments overlapping any of its exons (by at least one base) were summed separately for each library. This included 16,734 and 22,670 protein-coding genes with a transcript >200 nt in chicken and mouse, respectively. Then, for each protein-coding gene model >200 nt total exonic length in each library, we calculated the ratio of (fragments overlapping that model's exons in the library):(total number of fragments overlapping that model's exons across all libraries). Using only genes having at least one read in each library (14,109 in chicken, 16,329 in mouse), the median ratio across all protein-coding gene models >200 nt exonic length was found for each library. Finally, for each library, the total number of fragments overlapping an exon of a protein-coding transcript with >200 nt exonic length were summed, and we calculated the

proportions of fragments within genes to the overall number of mapped fragments in each library (Table S10). We then randomly selected read fragments without replacement to downsample each library to the following level: (total number of mapped fragments in the library)*(the minimum median of all samples)/(median ratio of that sample) (Table S10).

Next we calculated FPKMs for all protein-coding genes (>200 nt exonic length) in each library using cufflinks (4) v0.9.3 with the –GTF option and corresponding GTF files from Ensembl release 61. Additionally, rRNAs, tRNAs and all mitochondrial transcripts (as annotated in Ensembl release 61) were explicitly masked from being used in the denominator of the normalization. The reference sequence was provided to cufflinks to improve quantification accuracy. Upper-quartile normalization was enabled to improve FPKM robustness when comparing across libraries.

Quality control for batch effects

To assess possible batch effects correlated with library, flow cell or lane (5, 6), we also calculated FPKM as described above but on a lane-by-lane basis. We then performed, separately for each species, principal component analysis on FPKM values from all lanes in this project on all flow cells, including only genes that appeared in the top 4,000 most highly expressed in at least one library in that species. We used the R function prcomp on the transpose of the standard gene by sample matrix to cluster samples rather than genes, first centering the variables (samples) on zero and scaling the samples to have unit variance. For each species, we constructed biplots (using the R function biplot) covering every principal component that explained more than 2% of the variance in expression.

For chicken, this consisted of the first six principal components, cumulatively explaining 96% of the variance across the 17 lanes (Table C.6). For mouse, this consisted of the first eight principal components, cumulatively explaining 86% of the variance across the 31 lanes (Table C.6). None of these principal components separated mainly resequenced libraries on different lanes, suggesting that flowcell and lane batch effects had no major contributions to the observed differences amongst these libraries (Fig. S23 and S24).

In addition to looking at lane or flowcell effects (which implicitly takes into account the run date), we also looked explicitly in these principal components for batch effects at the level of libraries due to median pre-mRNA fraction (Table S9), internal insert size (Table S3), RIN (chick only; Table S2), 260:280 nm absorbance ratios (Table S1; Table S2), 260:230 nm absorbance ratios (Table S1; Table S2), and the percentage of unique mapped fragments in protein coding genes (Table S10). For this we performed a principal component analysis of the transpose of the gene by sample matrix considering specificities of genes in the downsampled libraries rather than the FPKMs of each individual lane using the R function prcomp (no scaling to unit variance; specificities were zero-centred for each sample). We only included genes that were used in the cross-species comparison (see below). We then built biplots with the first nine components in mouse (explaining 96% of the variance; see Table S12) and the first five components in chicken (explaining 97% of the variance; see Table S12). We used both the library labels and the technical variable labels to identify differences in gene expression that could be driven by batch effects (Fig. S25-S33). Some of these technical variables were correlated with the first principal component of both chicken and mouse (Fig. S34-S37), suggesting that RNA Integrity Number, pre-mRNA fraction, and the percentage of

uniquely mapped fragments in protein-coding genes significantly affected the quantifications. In the absence of a well-validated method to control for continuous, as opposed to categorical (7), confounds, and not having biological replicates, we took these batch effects into consideration when interpreting the results of subsequent analyses.

Since samples underwent poly(A) selection and had varying degrees of RNA quality, we also tested for any transcript 3′ bias in read distribution. Note that this is affected both by the efficiency of the poly(A) selection process and by RNA degradation. We considered Ensembl-annotated protein-coding transcripts for which the isoform fraction (as identified by cufflinks) was greater than or equal to 95% and the FPKM was greater than 20. We measured read coverage at each position, starting from the 3′ annotated end and working to the 5′ annotated end. We normalized read coverage in each transcript such that, on average, coverage would equal one at each position of the transcript. Finally, we assessed the mean and standard deviation at each position having greater than 10 overlapping transcripts. This was done separately for each library (Fig. S38-S39). We then looked for possible batch effects by plotting on the major principal components the maximum density within 2000 nt upstream of the 3′ end, based on 9,024-9,488 transcripts in chicken and 6,373-7,900 transcripts in mouse (Fig. S40-S41). This revealed that the first principal component of the chicken samples and the first two principal components of the mouse samples were correlated with this proxy for RNA quality (Fig. S42-S44). We propose that this measure of 3′ bias reflects RNA quality since it is highly and significantly negatively correlated with RIN in chick (r=-0.93, 95% confidence interval [-0.48, -0.99]; two-tailed P=0.0072), and there is a ready theoretical

explanation for why it should reflect RNA quality: the more degraded an RNA sample is, the more 3′ bias should result from an equally efficient poly(A) selection. Differences in 3′ bias may also be caused by differences in efficiency of the poly(A) selection process, but that is unlikely to be the sole explanation given the strong correlation with RIN observed in the chick samples.

To ensure that the final gene co-expression modules were not tainted by batch effects, we tested all resulting co-expressed gene sets for Pearson's correlations of their module eigengenes (see below) with these potential confounding technical variables, assessing significance using the online calculator at http://faculty.vassar.edu/lowry/ch4apx.html. We used the online calculator at http://faculty.vassar.edu/lowry/rdiff.html to assess the significance of the difference between two correlation coefficients.

Determining correspondence of the mouse cortical layer dissections

To determine the true correspondence of laminarly dissected samples to known cytoarchitectural layers, we created the heatmaps in Fig. S3 as follows: if genes g_l are annotated (8) as being preferentially *and specifically* expressed in layer l and f_{sgl} is the fractional expression of gene g_l in sample s relative to expression in all samples, then t_{sl} is the median of (f_{sgl}) over genes g_l. The heat map intensities represent [t_{sl} – $\text{mean}_{\text{all layers } l}$ (t_{sl})] / [$\text{mean}_{\text{all layers } l}$ (t_{sl})].

Confirming accuracy of the mouse dissections in other regions

To confirm the accuracy of the non-cortical dissections in the mouse, we manually reviewed images series of *in situ* hybridizations, where present, from the Allen Mouse

Brain Atlas (8) of the three genes most specific to each non-cortical library amongst genes that were also used for the cross species comparisons (see below). We defined specificity as the fraction of total expression across all libraries that occurred in the library of interest. Fig. S4-S7 display an example from the top three most specific genes of each non-cortical sample, which were all consistent with the intended dissections.

Confirming accuracy of the chicken dissections

To confirm the accuracy of the chicken dissections, we performed *in situ* hybridizations in 20 μm coronal sections of adult chicken brains for the most specific genes to seven chicken regions dissected (hyperpallium, mesopallium, nidopallium, striatum, hippocampus, dorsolateral corticoid area and arcopallium), using a similar digoxigenin-based method to that described by Wang *et al.* (9) (Fig. S8). Species-specific riboprobes were synthesized from respective cDNAs. Total RNA was extracted from brains of individual species and the first strand cDNA was synthesized using Superscript III reverse transcriptase together with random hexamers (Invitrogen, Paisley, UK) following the manufacturer's instructions. **Table S4** lists the forward and reverse primers used to generate gene specific cDNA fragments using polymerase chain reaction (PCR). The resulting PCR products were individually ligated into the pST-Blue 1 plasmid (Novagen, Nottingham, UK) and confirmed by sequencing. The antisense and sense (a negative control) cRNA probes were transcribed using T7 and SP6 RNA polymerase with digoxigenin (DIG)-labelled RNA mixture, respectively (Roche, Penzberg, Germany). The *in situ* hybridizations were performed as previously described(9). Fresh-frozen brains were sectioned to 20 μm coronally on a cryostat (Jung CM3000; Leica, Germany).

Frozen sections were fixed with 4% PFA in PBS for 30 min, de-proteinized with 0.1N HCL for 5 min, acetylated with acetic anhydride (0.25% in 0.1M triethanolamine hydrochloride) and pre-hybridized at RT for at least 1hr in a solution containing 50% formamide, 10mM Tris, pH7.6, 200mg/ml E. coli tRNA, 1x Denhardt's solution, 10% dextran sulphate, 600 mM NaCl, 0.25% SDS and 1mM EDTA. The sections were hybridized in the same buffer with the DIG-labelled probes overnight at 66-68°C. After hybridization, sections were washed to a final stringency of 30mM NaCl/3mM sodium citrate at 66-68°C and detected by anti-DIG-alkaline phosphatase antibody in conjunction with a mixture of nitroblue tetrazolium (NBT) and 5-bromo-4-chloro-3-indolyl-phosphate (BCIP) (Roche, Penzberg, Germany).

Selecting genes for the cross-species comparison

For the evolutionary comparisons, we only considered one-to-one protein-coding orthologs with >200 nt exonic length in both mouse and chicken as identified by Ensembl (release 61)(10). This comprised 11,033 genes in total. We then required a gene to be expressed in the top 5,000 genes in at least one library in one species and to have a non-zero variance in both species. This left 5,170 genes ranging in expression from 115 FPKM to 106,823 FPKM or greater in at least one library.

Identifying gene co-expression modules

We next identified gene co-expression modules using weighted gene co-expression network analysis (11-13) with the set of highly expressed one-to-one orthologs described

above. To identify the appropriate soft threshold (power) for each species in turn, we plotted the R^2 of the scale-free topology model fit against the soft threshold for powers 1-10, 12, 14, 16, 18, 20, 22, 24, 26, 28, and 30 (**Table S13**;

Table S14; **Fig. S45**-S46). We chose 10 for mouse and 20 for chicken for the following reasons: the R^2 was high (>0.8) in both species so they both approximated a scale-free topology while simultaneously producing networks having high mean connectivity, they both had negative slopes (without which the network has no clear biological interpretation), and they produced the most preserved modules (number of modules with combined Zpreservation>3) when comparing between the two species for several combinations (these comparisons are described below; Table S15).

Once the soft threshold was chosen for each species, we determined the general network properties of the two sets using the softConnectivity R function(12), comparing both ranked levels of expression and ranked connectivity of orthologous genes in the two species with the verboseScatterplot R function (Fig. S9). This demonstrated that, while there was significant conservation between mouse and chicken at the level of expression ($P<10^{-200}$), the position of a gene in a network (as a 'hub' or at the periphery) was not generally conserved.

We built networks as described above. Briefly, we computed an adjacency matrix whose entries a_{ij} correspond to $|0.5+\text{cor}(x_i,x_j)/2|^{\beta}$ where β is the soft threshold chosen above and cor is Pearson's *rho*. The weighted adjacency between two genes *i* and *j* is thus proportional to their correlation on a logarithmic scale, and includes information on positive or negative correlation. We then set the diagonal of the adjacency matrix to zero and used it to calculate the topological overlap matrix using the signed TOMsimilarity R function. We subtracted the resulting topological overlap matrix from a matrix of ones and hierarchically clustered the result (with the flashClust R function and the average

agglomeration method) to produce the dendrograms in Fig. S47 and S48.

Next, we defined modules using the dynamic tree cutting algorithm as implemented in the Dynamic Tree Cut R package (14) to produce five sets of modules of different sizes for both chicken and mouse (Fig. S47 and S48). We used a cut height of 0.99 and a minimum cluster size of 30-3*ds where ds is an integer in the closed set [0,4]. We chose the third split in chicken and the fourth split in mouse for the same reasons the soft threshold was chosen.

We then determined the "module eigengene" for each module (15). This corresponds to the first principal component of the expression matrix of genes in the module, and is thus similar to a sort of weighted average expression profile for the module of co-expressed genes. These were hierarchically clustered to visualize similarities of different modules (Fig. S49 and S50).

Comparing gene co-expression modules between species

To visualize how well individual modules were preserved between species, we projected the colours of the final mouse modules onto the chicken dendrogram (Fig. S10) and the colours of the final chicken modules onto the mouse dendrogram (Fig. S11) using the R function plotDendroAndColors.

To quantify how well individual modules were preserved between species, we calculated several module preservation statistics using the modulePreservation R function (16) with the following parameters: "signed" network, 30 permutations, max gold module (module of random genes) size of 30 and max module size of 400 (Table S5;

Table S6). The grey module consists of genes that were not assigned to a coexpressed module. These statistics fall into two broad categories: density statistics and connectivity preservation statistics, which are averaged ([*Zdensity.pres*+ *Zconnectivity.pres*]/2) into one final Z score of cross-species module preservation *Zsummary.pres*. Modules with *Zsummary.pres* scores between 2 and 10 are generally considered weakly to moderately preserved (16).

Module density statistics, summarized by *Zdensity.pres*, reflect how densely connected genes in the module of a reference network are in the test data. This summary statistic is defined as the median of several module density metrics: *meanSignAwareCorDat* (the mean correlation density of the module q, calculated as $\text{mean}\left\{vectorizeMatrix\left(sign(r_{ij}^{[ref](q)})r_{ij}^{[test](q)}\right)\right\}$); *propVarExplained* (proportion of the variance explained by the module eigengene in the test data, calculated as the mean squared *kME*); *meanSignAwareKME* (measures the mean module membership where genes whose module memberships in the reference and test networks have opposite signs contribute negatively; is calculated as $mean_{i\in \mathrm{M}q}\left\{sign(kME_i^{[ref][q]})kME_i^{[test](q)}\right\}$); and *meanAdj* (mean adjacency of genes in the module, calculated as $mean\left(vectorizeMatrix(A^{[test](q)})\right)$). Other density statistics not used to create the summary statistic, but which may nevertheless be informative, are *meanMAR* (mean Maximum Adjacency Ratio, $mean\left(MAR_i^{[test]}=\frac{\sum_{j\neq i}a_{ij}^2}{\sum_{j\neq i}a_{ij}}\right)$, of genes i in the module) and *meanClusterCoeff* (mean clustering coefficient of genes i in the module where

$$mean\left(clusterCoef_i^{[test]} = \frac{\sum_{j\neq i}\sum_{m\neq j.i} a_{ij}a_{jm}a_{mi}}{\left(\sum_{j\neq i} a_{ij}\right)^2 - \sum_{j\neq i} a_{ij}^2} \right).$$

Connectivity pattern preservation statistics, summarized by *Zconnectivity.pres*, reflect how well the co-expression connections within a module are preserved. This summary statistic is defined as the median of several connectivity pattern metrics: *cor.cor* (the correlation between pairwise correlations of expression of module genes); *cor.kME* (the correlation of correlations of genes in the module to the module eigengene in the test set); and *cor.kIM* (the correlation of intramodular connectivity for module *q* $kIM_i = \sum_{\substack{j\in \mathrm{M}_q \\ j\neq -1}} a_{ij}^q$, which quantifies if hub genes *i* of module *q* in the reference network are also hub genes of module *q* in the test network). Other connectivity statistics not used to create the summary statistic, but which may nevertheless be informative, are *cor.kMEall* (the correlation of correlations of all genes to the module eigengene in the test set); *cor.MAR* (correlation of Maximum Adjacency Ratios, $MAR_i = \frac{\sum_{j\neq i} a_{ij}^2}{\sum_{j\neq i} a_{ij}}$, for genes *i* in the module); and *cor.clusterCoeff* (correlation of the clustering coefficients, $clusterCoef_i = \frac{\sum_{j\neq i}\sum_{m\neq j.i} a_{ij}a_{jm}a_{mi}}{\left(\sum_{j\neq i} a_{ij}\right)^2 - \sum_{j\neq i} a_{ij}^2}$, for genes *i* in the module).

We also compared module preservation by considering the module eigengene, as previously described (17). Briefly, this involved taking the gene module as defined in the first (reference) species and finding the first principal component of the expression

matrix of their orthologs in the second (test) species. Once this "module eigengene" was defined in the new species, we calculated the correlation of the expression of every module gene with the expression of the module eigengene (kME) in samples from the first (reference) species and then in the second (test) species. We then plotted these correlations in each species against one another using the verboseScatterplot R function to find Pearson's correlation coefficient of the two correlations plotted against one another (along with its statistical significance). Next, we did the same, but instead of only correlating module genes to the module eigengene under consideration, we correlated all genes to the module eigengene in both species.

To compare the modules in the chicken and mouse networks we used the userListEnrichment R function to identify the modules in each species that were significantly overlapping. For every chicken module, this tested for significant enrichment in every mouse module using a hypergeometric distribution, and subsequently applied a Bonferroni correction for every pairwise comparison made (taking into account both lists).

For those modules having significant cross-species overlap, we identified genes likely to be hubs in both networks by finding the genes (in or not in the module) having the highest average kME rank (correlation with the respective module eigengene) using the modules defined in the original networks and not projecting one onto the other.

Annotating gene co-expression modules

Using the same collection of databases as described by Belgard *et al.* (1), we performed conditional one-sided Fisher's exact tests of enrichment for both the mouse modules and for the chick modules (using the mouse orthologs). The background consisted of all genes used to build modules for both chick and mouse that had any kind of annotation in the given database. To reduce the number of underpowered tests or results having limited biological interpretability, we only tested terms that (1) included at least 2 genes in the background and (2) included a sufficient number of genes in the background such that a *p*-value of 0.0005 would be theoretically possible for the given module size.

For all modules discussed here, we provide the top 10 genes most correlated with the module eigengene, their median quantile-normalized FPKM expression levels, and relative fold expression with respect to this median value in Table S7 and Table S8

Graphical depictions of gene co-expression networks

Topological overlap dissimilarity matrices and pairwise gene correlations from chicken and mouse striatal, hippocampal and layer IV/nidopallium modules were imported into Cytoscape 2.8.2. Species-specific networks were constructed using a JGraph spring embedding layout and merged with the other species using the "union" operation. Duplicated edges were removed and intersected genes between chicken and mouse were identified using the "intersection" operation and highlighted with bigger size and pink colour in the network. Correlations across nodes were visualized as weighted thickness in the edges using NetworkAnalyzer.

Network analysis removing low-RIN samples

We repeated the network analysis removing chicken arcopallium and hyperpallium, and overlapped the new chicken modules with both the old mouse modules and new mouse modules built using different parameters. In the new analysis a soft power threshold of 7 was chosen for mouse and 12 for chick. We used the R function blockwiseModules with a maxBlockSize parameter that would accommodate all genes in the analysis. The networks were signed with a minimum module size of 20, a reassign threshold of 0, and a mergeCutHeight of 0.25. We then used the userListEnrichment function to identify modules with cross-species overlaps.

Expression specificity analysis

Defining specificity (S) of the expression of gene G in region R of species Sp as the quantile-normalized FPKM of gene G in region R of species S divided by the sum of the quantile-normalized FPKMs of gene G in all regions of species Sp.

One vector was formed for each region in each species, and its values populated with specificities for each gene. The dot product of each region in one species was computed against vectors with the same ordering of orthologous genes in the other species, yielding (Specificity(Region 1, Gene 1) x Specificity(Region 2, Gene 1) + Specificity(Region 1, Gene 2) x Specificity(Region 2, Gene 2) + ...). To account for the fact that the proportions of highly specific genes varied considerably between regions, a Z score was

computed based on an empirical distribution of dot products in which genes were permuted (e.g. the order was shuffled). For example, a simulated dot product could be (Specificity(Region 1, Gene 1) x Specificity(Region 2, Gene 1038) + Specificity(Region 1, Gene 2) x Specificity(Region 2, Gene 8420) + ...).

The simulations of similarity were done as follows:

1. An empirical null distribution of dot products is constructed as above.

2. A simulated distribution of more similar vectors is constructed as follows: a specified percentage of similarity X (where $0\%<X\leq 50\%$) of abs[Specificity(Region 1, Gene 1) - Specificity(Region 2, Gene 1038)] is added to min[Specificity(Region 1, Gene 1), Specificity(Region 2, Gene 1038)] and subtracted from max[Specificity(Region 1, Gene 1), Specificity(Region 2, Gene 1038)] to arrive at a dot product computed as {min[Specificity(Region 1, Gene 1), Specificity(Region 2, Gene 1038)] + X%*abs[Specificity(Region 1, Gene 1) - Specificity(Region 2, Gene 1038)]}x{max[Specificity(Region 1, Gene 1), Specificity(Region 2, Gene 1038)] - X%*abs[Specificity(Region 1, Gene 1) - Specificity(Region 2, Gene 1038)]} + ...

Where the specificities are equal, that term of the dot product is equivalent for all X.

3. For each cross-species pairing of regions, the smallest X was determined for which a significant Bonferroni-corrected (adjusted for the number of pairwise cross-species regional comparisons) difference would be called 80% of the time. If the smallest X that can be called is low, that suggests the test is well powered to detect even very small similarities beyond chance.

Marker gene analysis

The samples were quantile normalized to one another as follows: first, genes were sorted in order of post-downsampling FPKMs (as described above) in every sample; second, 0.001 was added to each FPKM to avoid artifacts arising from genes having an FPKM of 0; third, the geometric mean of these new FPKMs was taken at every position of the list (e.g. 1st position of sample A, 1st position of sample B, etc.) and the genes in this position were assigned this resulting geometric mean as the new expression level.

The normalized FPKMs were used to identify candidate marker genes of each region. To be initially defined as a 'strict' marker of a region, a gene must have the following properties: (1) normalized FPKM in the region must be at least 50% higher than in the region of second highest expression, (2) normalized FPKM in the region must be at least 100% higher than in the region of third highest expression, (3) normalized FPKM in the region must be at least 300% higher than the mean FPKM, and, (4) at least three samples must have a nonzero FPKM.

Numbers of 'strict' marker genes varied: in chick there were 20 in hippocampus, 11 in striatum, 2 in mesopallium, 1 in arcopallium and hyperpallium and none in dorsolateral cortex and nidopallium; in mouse there were 51 in striatum, 22 in hippocampus, 3 in pallial amygdala, 2 in dorsal cortex A, 1 in lateral cortex A, E, F and dorsal cortex F and none in the claustrum-endopiriform complex, dorsal cortex B, C, D, E and lateral cortex B, C, and D. Note that fewer markers meet these strict criteria in the case of similar

and/or overlapping dissections, as is found in the laminar dissections of the two cortical areas. If a region did not contain at least ten of these 'strict' marker genes, additional genes were added in descending order of specificity until the region was associated with ten candidate markers. (Specificity is defined as FPKM in that region divided by the sum total of FPKM across all regions.) These candidate markers also had to be more highly expressed in the marked region than in any other region. Each region in both species was thus associated with at least 10 candidate markers (mouse Ensembl gene IDs or the corresponding one-to-one mouse ortholog for chicken samples):

mouse striatum
ENSMUSG00000024077
ENSMUSG00000070687
ENSMUSG00000048251
ENSMUSG00000027210
ENSMUSG00000032625
ENSMUSG00000046922
ENSMUSG00000029755
ENSMUSG00000029754
ENSMUSG00000045534
ENSMUSG00000021180
ENSMUSG00000023868
ENSMUSG00000021990
ENSMUSG00000019990
ENSMUSG00000068696
ENSMUSG00000044288
ENSMUSG00000070720
ENSMUSG00000051650
ENSMUSG00000051111
ENSMUSG00000027849
ENSMUSG00000034472
ENSMUSG00000042604
ENSMUSG00000042453
ENSMUSG00000020953
ENSMUSG00000049511
ENSMUSG00000041762
ENSMUSG00000061762
ENSMUSG00000031906
ENSMUSG00000030220
ENSMUSG00000055639
ENSMUSG00000030222
ENSMUSG00000031112
ENSMUSG00000038718
ENSMUSG00000044167
ENSMUSG00000063446
ENSMUSG00000033007
ENSMUSG00000021478
ENSMUSG00000027827
ENSMUSG00000021379
ENSMUSG00000054162
ENSMUSG00000036095
ENSMUSG00000032698
ENSMUSG00000032259
ENSMUSG00000017491
ENSMUSG00000030854
ENSMUSG00000024524
ENSMUSG00000027203
ENSMUSG00000022840
ENSMUSG00000031837
ENSMUSG00000071234
ENSMUSG00000020121
ENSMUSG00000020723

mouse hippocampus
ENSMUSG00000027971
ENSMUSG00000047712
ENSMUSG00000049420
ENSMUSG00000039037
ENSMUSG00000049281
ENSMUSG00000049892
ENSMUSG00000037984
ENSMUSG00000038463
ENSMUSG00000031618
ENSMUSG00000028004
ENSMUSG00000055078
ENSMUSG00000028341
ENSMUSG00000009075
ENSMUSG00000028532
ENSMUSG00000055540
ENSMUSG00000041225
ENSMUSG00000035168
ENSMUSG00000052087
ENSMUSG00000048218
ENSMUSG00000036111
ENSMUSG00000039358
ENSMUSG00000040372

mouse lateral cortex A
ENSMUSG00000091207
ENSMUSG00000070803
ENSMUSG00000053166
ENSMUSG00000032033
ENSMUSG00000029122
ENSMUSG00000017754
ENSMUSG00000055407
ENSMUSG00000039323
ENSMUSG00000058498
ENSMUSG00000049796

mouse dorsal cortex A
ENSMUSG00000048583
ENSMUSG00000030218
ENSMUSG00000031490
ENSMUSG00000029661
ENSMUSG00000019929
ENSMUSG00000000753
ENSMUSG00000041559
ENSMUSG00000021032
ENSMUSG00000035783
ENSMUSG00000061878

mouse lateral cortex B

ENSMUSG00000046593
ENSMUSG00000039057
ENSMUSG00000040164
ENSMUSG00000023192
ENSMUSG00000021193
ENSMUSG00000044071
ENSMUSG00000036192
ENSMUSG00000039137
ENSMUSG00000035202
ENSMUSG00000037843

mouse dorsal cortex B
ENSMUSG00000026019
ENSMUSG00000061559
ENSMUSG00000063626
ENSMUSG00000024798
ENSMUSG00000034755
ENSMUSG00000029815
ENSMUSG00000040998
ENSMUSG00000014232
ENSMUSG00000039470
ENSMUSG00000018634

mouse lateral cortex C
ENSMUSG00000020396
ENSMUSG00000033615
ENSMUSG00000061086
ENSMUSG00000039620
ENSMUSG00000040111
ENSMUSG00000017417
ENSMUSG00000022940
ENSMUSG00000025781
ENSMUSG00000030869
ENSMUSG00000021743

mouse dorsal cortex C
ENSMUSG00000037492
ENSMUSG00000027977
ENSMUSG00000024256
ENSMUSG00000027423
ENSMUSG00000023927
ENSMUSG00000042115
ENSMUSG00000030103
ENSMUSG00000046079
ENSMUSG00000022419
ENSMUSG00000003923

mouse lateral cortex D
ENSMUSG00000050447
ENSMUSG00000048004
ENSMUSG00000026610
ENSMUSG00000037426
ENSMUSG00000028613
ENSMUSG00000053141
ENSMUSG00000047976
ENSMUSG00000017978
ENSMUSG00000026384
ENSMUSG00000035236

mouse dorsal cortex D
ENSMUSG00000006800
ENSMUSG00000049422
ENSMUSG00000026307
ENSMUSG00000050608
ENSMUSG00000060402
ENSMUSG00000069769
ENSMUSG00000026427
ENSMUSG00000020526
ENSMUSG00000015002
ENSMUSG00000064357

mouse lateral cortex E
ENSMUSG00000025723
ENSMUSG00000027787
ENSMUSG00000041540
ENSMUSG00000034009
ENSMUSG00000006005
ENSMUSG00000020955
ENSMUSG00000035033
ENSMUSG00000031879
ENSMUSG00000028559
ENSMUSG00000028132

mouse dorsal cortex E
ENSMUSG00000039068
ENSMUSG00000022602
ENSMUSG00000010721
ENSMUSG00000038602
ENSMUSG00000002205
ENSMUSG00000028843
ENSMUSG00000046546
ENSMUSG00000032532
ENSMUSG00000050663
ENSMUSG00000047842

mouse lateral cortex F
ENSMUSG00000045636
ENSMUSG00000033717
ENSMUSG00000078591
ENSMUSG00000026347
ENSMUSG00000022246
ENSMUSG00000021645
ENSMUSG00000019997
ENSMUSG00000056296
ENSMUSG00000051920
ENSMUSG00000021189

mouse dorsal cortex F
ENSMUSG00000028883
ENSMUSG00000026830
ENSMUSG00000041377
ENSMUSG00000029563
ENSMUSG00000022306
ENSMUSG00000036777
ENSMUSG00000038668
ENSMUSG00000026519
ENSMUSG00000016918
ENSMUSG00000019888

mouse claustrum endopiriform
ENSMUSG00000025969
ENSMUSG00000026826
ENSMUSG00000031654
ENSMUSG00000029101
ENSMUSG00000027669
ENSMUSG00000027978
ENSMUSG00000045648
ENSMUSG00000035513
ENSMUSG00000039579
ENSMUSG00000025905

mouse pallial_amygdala
ENSMUSG00000046523
ENSMUSG00000049630
ENSMUSG00000041380
ENSMUSG00000026344
ENSMUSG00000053819
ENSMUSG00000074575
ENSMUSG00000025370
ENSMUSG00000034796
ENSMUSG00000071379
ENSMUSG00000040856

chicken striatum
ENSMUSG00000026930
ENSMUSG00000032259
ENSMUSG00000090223
ENSMUSG00000067578
ENSMUSG00000021948
ENSMUSG00000029754
ENSMUSG00000027347
ENSMUSG00000023868

ENSMUSG00000061762
ENSMUSG00000045573
ENSMUSG00000039474

chicken hyperpallium
ENSMUSG00000068220
ENSMUSG00000020309
ENSMUSG00000043259
ENSMUSG00000029822
ENSMUSG00000034647
ENSMUSG00000014763
ENSMUSG00000044716
ENSMUSG00000000753
ENSMUSG00000049744
ENSMUSG00000021395

chicken dorsolateral cortex
ENSMUSG00000027584
ENSMUSG00000024501
ENSMUSG00000026826
ENSMUSG00000029673
ENSMUSG00000049630
ENSMUSG00000050511
ENSMUSG00000064293
ENSMUSG00000041559
ENSMUSG00000022103
ENSMUSG00000030905

chicken mesopallium
ENSMUSG00000038331
ENSMUSG00000033717
ENSMUSG00000045731
ENSMUSG00000035513
ENSMUSG00000035033
ENSMUSG00000032532
ENSMUSG00000019880
ENSMUSG00000005672
ENSMUSG00000009075
ENSMUSG00000022861

chicken hippocampus
ENSMUSG00000048583
ENSMUSG00000025969
ENSMUSG00000024112
ENSMUSG00000005958
ENSMUSG00000066687
ENSMUSG00000025020
ENSMUSG00000006476
ENSMUSG00000034796
ENSMUSG00000025582
ENSMUSG00000050069
ENSMUSG00000026278
ENSMUSG00000070720
ENSMUSG00000070866
ENSMUSG00000073680
ENSMUSG00000027313
ENSMUSG00000032402
ENSMUSG00000024479
ENSMUSG00000037679
ENSMUSG00000019772
ENSMUSG00000029135

chicken nidopallium
ENSMUSG00000048004
ENSMUSG00000038055
ENSMUSG00000046321
ENSMUSG00000032452
ENSMUSG00000059742
ENSMUSG00000032625
ENSMUSG00000056306
ENSMUSG00000016918
ENSMUSG00000050663
ENSMUSG00000036192

chicken arcopallium
ENSMUSG00000045636
ENSMUSG00000052229
ENSMUSG00000030317
ENSMUSG00000063531
ENSMUSG00000030342
ENSMUSG00000021217
ENSMUSG00000036777
ENSMUSG00000040452
ENSMUSG00000026519
ENSMUSG00000019888

Finally, the highly expressed one-to-one orthologs of candidate markers for a region identified in the 'reference' species were projected into regions of the 'target' species, and their expression pattern was analyzed in the target species. We calculated the specificity of each putative marker across regions in the target species, and then summed these specificities.

Empirical *p*-values were calculated to account for the fact that genes with more variance in expression in the one species tended to be more variable in expression in the other species as well. For example, the median standard deviation of specificities of markers of chick regions were in the 99th percentile for chick regions and their orthologs in the

93rd percentile for mouse regions. Likewise, the median standard deviation of specificities of markers in mouse regions were in the 98th percentile for mouse regions and their orthologs in the 84th percentile for chick regions. For each species separately, we plotted all genes compared between species where the independent axis denoted the standard deviation of specificities of that gene in the target species and the dependent variable was either 1 or 0 to reflect that the gene either was or was not a marker in the reference species. A LOESS curve was constructed through these points using a smoothing parameter of 0.5 and the Excel plugin available from http://peltiertech.com/WordPress/loess-utility-awesome-update/. Thus every gene had an assigned probability of having been a marker in the reference species based on the standard deviation of its expression in the target species. A small number of points that were interpolated to have slightly negative probabilities were reset to zero probability. Subsequent random samplings were weighted by normalized probabilities. To calculate empirical p-values, we sampled with replacement, using these weightings, from genes in our set of highly expressed one-to-one orthologs. The size of this resampled set was equal to the larger of either ten or the number of strict markers defined for the region in the reference species. The specificities of the randomly drawn genes were then summed separately for each region of the target species. This was repeated 100,000 times to determine the observed total specificity's place on the null distribution. Position on the null distribution was generally transformed into a *p*-value by taking two times the smaller of either the observed percentile or 1.0 minus the observed percentile. Where the observed total specificity was greater than or less than all total specificities in the simulated null distribution, a *p*-value of 2x(1/100,001) was conservatively assigned for

purposes of calculating *q*-values. *Q*-values were calculated from these empirical p-values using qvality v1.11. Only results with $q<0.05$ were considered significant.

SI Results

We quantified gene expression in 16 regions pooled from eight adult mice and 7 regions pooled from eighteen adult chickens, and noted specific expression trends that could be tainted by batch effects. Laminar marker genes confirmed dissections of cortical layers in mouse. Non-cortical dissections in mouse were confirmed by pre-existing *in situ* hybridizations of genes we found to be among the most specific to each structure. All the completed *in situ* hybridizations in chicken are concordant with the sequencing-based predictions. Ranked expression levels between the two species were correlated with one another (Spearman's *rho*, r_S=0.49; $P<10^{-200}$), but connectivity rank was not (r_S=0.0071; P=0.61). (However, the latter may be artificially depressed by variable RNA quality.) We then constructed modules of co-expressed genes in each species.

Functional annotation enrichments amongst co-expressed genes in mouse and chick

There were 15 distinct modules in chick and 49 in mouse (we do not include in this count the grey modules, which consists of genes that were not assigned to a module). There are at least two likely contributors to the greater number of mouse modules: first, the minimum size allowed for a module in mouse was smaller than for chick; second, there were many more samples sequenced from mouse which allowed for a greater diversity of clearly delineated gene expression patterns.

We tested for functional enrichments from several different annotation sources. When we applied a Bonferroni correction to account for all tests performed across all modules for a single annotation source in a single species, no modules were significantly enriched in either species or among any of the annotation sources with a FWER<0.05. (The

FWER, or familywise error rate, is the probability of making at least one type I error in a set of hypotheses being tested.) When we instead applied the Bonferroni correction to account only for all tests performed from a single annotation source in a single module, there were some 'significant' enrichments as shown in Table S18.

Three mouse modules – brown, magenta and tan – were not evenly expressed across samples and differed from one another (Fig. S16). (We do not discuss the fourth module, violet, here since the less-than-two-fold enrichment – 'membrane' – is not easily biologically interpreTable.) Mouse brown was specific to striatum. Mouse magenta genes tended to be expressed in striatum, hippocampus and deep cortical layers. Mouse tan was primarily found in upper layers of lateral cortex.

Two of these three mouse modules – magenta and tan – were significantly correlated with technical variables. Mouse magenta was negatively correlated with both 3′ bias (r=-0.67; two-tailed P=0.0048) and pre-mRNA fraction (r=-0.86; two-tailed P=0.000016). In contrast, mouse tan was *positively* correlated with 3′ bias (r=0.63; two-tailed P=0.0083).

Like the mouse modules, the four chick modules – brown, cyan, green and red – were not evenly expressed across samples and differed from one another (Fig. S17). Chick brown was found in dorsolateral cortex, mesopallium, striatum, and hippocampus. Chick cyan was mostly in arcopallium, but in nidopallium to a lesser extent. Chick green was specific to striatum. Chick red was especially low in hippocampus.

Two of the four chick modules – brown and green – were significantly correlated with technical variables. Chick brown was positively correlated with both RNA Integrity Number (r=0.92; two-tailed P=0.0097) and pre-mRNA fraction (r=0.82; two-tailed

P=0.025) and negatively correlated with the percentage of uniquely mapped reads falling into protein-coding models (*r*=-0.90; two-tailed *P*=0.0052). Green was also positively correlated with pre-mRNA fraction (*r*=0.79; two-tailed *P*=0.033).

Two of the chick modules with functional enrichments (green and cyan), significantly overlapped two mouse modules that also exhibited functional enrichments (brown and magenta, respectively). The first of these overlaps – chick green and mouse brown – also shared two significant functional enrichments: "signal transduction" (GO BP) and "adrenoceptor activity" (GO MF). The other overlapping modules – chick cyan and mouse magenta – did not have overlapping functional enrichments after Bonferroni correction (chick cyan had only one significant enrichment), but all the mouse magenta annotations were significant in chick cyan as a single test. The most significantly enriched functional annotation for chick cyan in the GO BP database was "myelination" (*P*=0.0039), with a $\log_2$(fold difference) of 4.6. The most significantly enriched functional annotation in the GO MF database was "structural constituent of the myelin sheath" (*P*=0.00028), with a $\log_2$(fold difference) of 6.9. The second most significantly enriched annotation in the GO CC database was "compact myelination" (*P*=0.0043), with a $\log_2$(fold difference) of 6.6. The final annotation enriched in mouse magenta, "myelin sheath" (GO CC), was enriched in chick cyan with a $\log_2$(fold difference) of 5.2 (*P*=0.033).

Limited and weak conservation of gene co-expression patterns between mouse and chick

We found only limited and weak conservation of gene co-expression patterns between

mouse and chick. Of the 49 modules in mouse, only two (mouse salmon and mouse steelblue) were significantly preserved in chick at a Bonferroni-corrected $P<0.05$ (per the summary score *Zpreservation*). The first of these, mouse salmon, significantly overlapped a chick module (chick turquoise). Although what appear to be housekeeping genes dominate the top five 'consensus' genes for the mouse-chick overlap (Table S19), mouse salmon was not evenly expressed across samples (**Fig. S18**). The other, mouse steelblue, did not significantly overlap a chick module. It was preferentially expressed in striatum (**Fig. S51**).

Mouse steelblue may, however, be tainted by batch effects. Its module eigengene was negatively correlated with both a metric of 3′ bias (r=-0.63; P=0.0092) and pre-mRNA fraction (r=-0.77; P=0.00044).

By the same criteria, only three of the fifteen modules in chick (chick green, chick greenyellow and chick cyan) were significantly preserved in mouse. All three of these modules significantly overlapped a mouse module. Chick green and chick greenyellow genes were preferentially expressed in striatum, while chick cyan genes tended to be oligodendrocyte markers and expressed in arcopallium (Table S19; **Fig. S19**).

Two of these three chick modules were correlated with technical variables. Chick green was positively correlated with pre-mRNA fraction (r=0.79; two-tailed P=0.033), though chick greenyellow (similar to chick green, see Fig. S49) was not significantly correlated with pre-mRNA fraction (r=0.26; two-tailed P=0.57). However, the difference between the correlation coefficients of chick green and chick greenyellow with pre-mRNA fraction was not significant (two-tailed P=0.25). Chick cyan was the second module correlated with a technical variable. Its module eigengene was positively correlated with

the percentage of uniquely mapped fragments in protein-coding genes (r=0.87; two-tailed P=0.011).

Matched co-expression modules in chick and mouse

Most of the significantly overlapping modules included well-known marker genes for specific cell types or organelles (Table S19) or were specifically expressed in each species, usually in regions known to be homologous to one another (Fig. S18-S19). Some of these modules were also correlated with technical variables in one or both species.

Modules mark chick and mouse striatum

Genes in the chicken green/greenyellow modules and the corresponding mouse brown module were preferentially expressed in striatum (Fig. S18 and S19). The chicken green & greenyellow modules cluster closely in the dendrogram of all chick modules, suggesting they are very similar to one another (Fig. S49). As discussed above, chick green was positively correlated with pre-mRNA fraction (r=0.79; two-tailed P=0.033), though there was no evidence that greenyellow was correlated with any technical variable.

We confirmed these predictions in mouse by examining *in situ* hybridizations(8) of the top five consensus genes from both the chick green/mouse brown and the chick greenyellow/mouse brown overlaps (Fig. S52). Seven of the nine unique genes were enriched in striatum; the probes covering the other two genes were too lightly stained to resolve expression patterns.

Modules mark chick and mouse hippocampus

Both chick magenta and mouse black were specific to hippocampus (relative, of course, to the dissected samples). We confirmed these predictions in mouse by examining *in situ* hybridizations (8) of the top five consensus genes from the chick magenta and mouse black overlap. In each image, the dissected portion of the hippocampal formation expressed the gene as highly if not higher than the rest of the section, and, of regions included in the bioinformatic analyses across all sections, none expressed the gene more highly than hippocampus (Fig. S53). However, not all regions are included in this automated analysis, and the expression intensity may be saturated in some sections. Neither chick magenta nor mouse black was significantly correlated with any technical variables.

Modules mark functionally analogous thalamic recipients in chick and mouse

According to their module eigengenes, chick black and mouse orange were primarily expressed in nidopallium and layer IV neocortex (both dorsal and lateral regions), respectively. The Allen Mouse Brain Atlas includes *in situ* hybridizations for three of the top five consensus genes. All three were highly expressed outside of neocortex; indeed, two of the three (*Fam19a2* and *Dctn3*) were considerably more highly expressed outside neocortex (in regions that were not dissected).

In situ hybridizations from the Allen Developing Mouse Brain Atlas demonstrate that *Rorb* is expressed in many places throughout development. At various times in development, *Rorb* is found in the eye, spinal cord (laminae 1-6 and the gray matter), hindbrain, midbrain, and forebrain. In adult, *Rorb* is expressed in layer
IV of neocortex, but is also found in several additional regions (Fig. S20). The specificity of *Rorb* to chick nidopallium was also confirmed by *in situ* hybridization (Fig.

3).

The extra-neocortical expression patterns of *Fam19a2* and *Dctn3* overlapped in layer 2 of piriform cortex (Fig. S21). Beyond that, *Fam19a2* was expressed predominantly in layer 2a of entorhinal cortex, taenia tecta and the anterior olfactory nucleus; and *Dctn3* was expressed in hippocampus, motor nucleus of trigeminal, substantia nigra (compact part), pontine gray, paragigantocellular reticular nucleus, pons and spinal cord (see Allen Mouse Brain Atlas) (8).

The module eigengene of mouse orange (but not of chick black) was positively correlated both with 3′ bias (r=0.56; two-tailed P=0.025) and with pre-mRNA fraction (r=0.56; two-tailed P=0.025).

Modules mark oligodendrocytes in chick and mouse

All of the top five genes most correlated with the module eigengenes of chicken cyan and mouse magenta were associated with oligodendrocytes. Two are known oligodendrocyte markers: *Tspan2*(18) and *Mbp*(19). All five were significantly more highly enriched in oligodendrocytes than in neurons or astrocytes in a microarray-based experiment in adult mouse forebrain(20): *Bcas1* (34-fold), *Gab1* (4.9-fold), *Anln* (18-fold), *Tspan2* (45-fold), *Mbp* (49-fold).

Both mouse magenta and chick cyan were correlated with technical variables. The module eigengene of mouse magenta was negatively correlated both with pre-mRNA fraction (r=-0.86; two-tailed P=0.000016) and with 3′ bias (r=-0.67; two-tailed P=0.0048). Chick cyan may also be negatively correlated with pre-mRNA fraction, though the test was underpowered to be conclusive (r=-0.74; two-tailed P=0.060). Chick

cyan was, however, positively correlated with the percentage of uniquely mapped fragments in protein-coding genes (r=0.87; two-tailed P=0.011).

Batch effects confound interpretation of the final module overlap

The final overlap involved chick turquoise and mouse salmon. Many of the top five genes correlated with these modules had housekeeping functions. *Cisd1* encodes mitoNEET, an outer mitochondrial membrane protein that regulates maximal mitochondrial respiratory rate (21). *Atp5f1* (ATP synthase, H+ transporting, mitochondrial F0 complex, subunit B1) encodes a member of the ATP synthase complex (22) consistently found in mouse mitochondria (23). *Rpl22l1* (ribosomal protein L22 like 1) contains a ribosomal protein L22e domain, and is thus annotated by the MGI curatorial staff as being a ribosomal subunit. The protein encoded by the human ortholog of *2310003C23Rik*, Twa1, is localized in the nucleus and forms a protein complex with Ran-binding protein in microtubule organising centre (RanBMP), and binder of the small GTPase Ran that is involved in nucleocytoplasmic transport of both RNA and proteins (24-26). *Ndufa5* (NADH dehydrogenase [ubiquinone] 1 alpha subcomplex, 5) is a mitochondrial gene encoding a protein found in mitochondria of every one of fourteen mouse tissues assessed in a previous study (23).

Despite this, the module eigengenes of both chick turquoise and mouse salmon were differentially expressed across samples (Fig. S18 and S19). *In situ* hybridizations further confirmed that some of these genes were differentially expressed across samples (Fig. S54).

However, chick turquoise (by far the largest module in chick) was highly and significantly correlated with technical variables. Chick turquoise was negatively

correlated with RNA Integrity Number (r=-0.92; two-tailed P=0.010) and positively correlated both with the percentage of reads uniquely mapping in protein-coding genes (r=0.87; two-tailed P=0.010) and with the 3′ bias metric (r=0.79; two-tailed P=0.035). In contrast, mouse salmon was not significantly correlated with any technical variables we considered in this analysis (r=-0.090, two-tailed P=0.74 for 3′ bias metric; r=-0.29, two-tailed P=0.27 for pre-mRNA fraction). In fact, its correlation with the 3′ bias metric differed significantly from the correlation of chick turquoise (two-tailed P=0.0424).

Marker gene analysis

Over a dozen results were significant at q<0.05. For chicken markers in mouse regions, chick striatal markers were more likely to have their orthologs in mouse striatum and less likely to be in dorsal cortex C & D (layers IV/V); and chick mesopallium markers were less likely to be in striatum. These all reflect pallial-subpallial differences. Such a test would be well powered in such a marker gene analysis because there was only a single subpallial dissection.

Orthologs of markers of chicken nidopallium were 40% less likely to be in mouse pallial amygdala. This is interesting since they are both ventral pallial derivatives, and is probably not explained by striatal components in pallial amygdala since the striatum was third from bottom with thirteen pallial regions above it.

Chicken arcopallium markers were more likely to be found in dorsal cortex sample F (layer VIb). These samples are especially high in oligodendrocyte genes.

As before, we found pallial-subpallial differences when projecting mouse markers to chick. Mouse striatal markers were higher in chick striatum and lower in hippocampus.

Lateral cortex layer IV was higher in nidopallium, reflecting the coexpression module overlap. However, the contributing genes were the same as in the network analysis.

Dorsal cortex layer VIb genes were higher in arcopallium and nidopallium and lower in hippocampus and dorsolateral cortex. Many of these genes are oligodendrocyte markers.

Reanalysis removing low RIN samples

A manual functional and expression analysis was performed for the genes in each module (16 in total) derived from WGCNA reanalysis after removing the samples with poor RIN (hyperpallium and arcopallium) (Table S2). The ten genes most correlated with the module eigengene for each of the overlaps were reviewed in Pubmed, the Allen Mouse Brain Atlas, and our own database. We summarize representative literature of function, anatomy and pathology (emphasizing, but not restricting to the nervous system) for each gene. We also compared our chicken brain expression patterns to those of the literature. Because current literature only has information on 1.3% of the screened genes, we extended the search to any avian species to attain a 3.8% of coverage. In contrast, 88% of these mouse genes had expression information, largely due to the Allen Mouse Brain Atlas.

Mouse modules from the second WGCNA network

Mouse 1 is a hippocampal-enriched module that overlapped significantly with four chicken modules involving hippocampal genes (chicken 2; Bonferroni-corrected P=0.009, chicken, 3; Bonferroni-corrected $P=9.6x10^{-8}$, chicken, 5; Bonferroni-corrected $P=8.1x10^{-9}$ and chicken 15; Bonferroni-corrected P=0.05, described below). Functional analysis confirms that genes of these modules are associated with hippocampal molecular and cellular neurobiological function at developmental and adult stages.

Zfpm2 is downregulated during development by *Zbtb20,* a cell fate determinant for CA1 hippocampal pyramidal neurons. *Zfpm2* is released from this repression in other hippocampal and cortical regions; and at adult stages (27). *Spast* mutations cause abnormal neuron morphology, dystrophic neurites, and axonal growth defects in hippocampus and spastic paraplegia, an axonopathy associated with degeneration of long spinal neurons (28). *Cbln1* participates in synapse formation in different brain regions (29). *Dner* participates in cell-fate assignation during nervous system development and in cerebellar maturation (30). *Slit1* gene product participates in axon guidance and migration and morphological differentiation of forebrain interneurons (31). *Dlc1* is a tumour suppressor; its mice mutants do not survive beyond embryonic day 10.5 (32). The protein encoded by *Vps26a* participates in intracellular receptor sorting (33). The functions of other gene members of this module (*Cdk14*, *Ccdc82* and *Fam169a*) are poorly characterized.

Mouse 2 genes, which present a broad expression patterned module without preferential enrichment, overlaps significantly with chicken module 1 (Bonferroni-corrected $P=1.5x10^{-6}$; described below). This module is associated with broad expression with some enrichment in the corticoid dorsolateral area. Mouse 2 gene members participate in structural and functional development of neurons. *Eif5* encodes a translation initiation factor with synaptic expression and cell cycle control participation (34, 35). *Adcyap1r1* product mediates antiapoptotic and neuroprotective functions (36). *Ilk* is associated with neuronal polarity, and its deletion from mouse cortex results in cortical lamination defects (37, 38). *Psmd9* is linked to bipolar disorder, depression, and type 2 diabetes (39). Supt5h participates in neuronal development (40). Other members have functions which are not restricted to neurons: *Cnot8* is related to antiproliferative activity (41),

Stk11 participates in organ development regulating cell polarity (42) and division and is highly associated with cancer (43, 44). Other gene members (*Mms19*, *Tbc1d9b* and *Btbd2*) are poorly characterized.

Mouse 3 genes have a broad expression pattern without clear enrichment for any dissected structure; this module overlaps significantly with chicken modules 1 and 5 (Bonferroni-corrected P=0.03; described below). Chicken module 1 presents a broad expression with some enrichment in the corticoid dorsolateral area and chicken module 5 presents enrichment in the nidopallium and striatum. Genes from this module have contrasting functions: specifically, axon guidance, neural induction, organelle enzymes and enzymes for cell membrane synthesis. *Srd5a3* encodes an enzyme necessary for the reduction of polyprenol to dolichol, the lipid anchor for N-glycosylation in the endoplasmic reticulum (45). *Srd5a3* defect is linked to cerebellar ataxia, mental retardation and ophthalmologic defects (46). *St3gal5* encodes an enzyme that catalyses the initial step in the biosynthesis of most complex gangliosides, found predominantly in the nervous system (47). *Unc5d* belongs to the Unc5 family of netrin receptors that participate in axon guidance, cell migration, and cell survival. *Uncd5* is primarily expressed by layer 4 cells in the primary sensory areas of the developing neocortex and may mediate the effect of netrin-4 on cortical cell survival in a lamina-specific manner (48). *Nog* inhibits TGF-β signal transduction and promotes organizing centres of forebrain development in the mouse (49). *Acot11* supports the transition of adipose tissue towards increased metabolic activity(50). *Arhgap22* determines different modes of tumour cell movement (51). *Vegfa* encodes a member of the PDGF/VEGF growth factor family, and is a mitogen that specifically acts on endothelial cells and has various effects, including mediating increased vascular permeability, angiogenesis, vasculogenesis,

endothelial cell growth, promoting cell migration, and inhibiting apoptosis (52). *Csdc2* encodes a key protein in controlling the recruitment of mRNA to the translational machinery, in response to environmental cues, both in development and in differentiated cells (53). *Smyd2* encodes a protein lysine methyltransferase that catalyses the transfer of methyl groups from S-adenosylmethionine (AdoMet) to acceptor lysine residues on histones and other proteins (54). *Tmem63c* is a functionally poorly characterized gene.

Mouse 5 genes show a generally broad expression with some enrichment in the hippocampus. This module is significantly overlapped with the chicken module 5 (Bonferroni-corrected P=0.0002; described below), which is associated with enrichment in the nidopallium and striatum. Genes from this module participate in cortical development, neurite extension, synaptic physiology and more general cell physiology. *Camk2d* is involved in light-induced phase delays (55) and vascular smooth muscle cell migration (56). A polymorphism in this gene is associated with seizure susceptibility in rats (57). *Efnb1* is highly expressed in cortical progenitors (58) and its overexpression stimulates cell division of neighbouring cells. *Efnb1* mutations cause craniofrontonasal syndrome(59). *Cadm1* product is an Ig superfamily member expressed on superior cervical ganglion neurites and it mediates cell-cell adhesion by trans-homophilic binding; a spliced form, sCADM1, appears to be involved in directional neurite extension (60). Cadm1-expressing synapses on Purkinje cell dendrites are involved in mouse ultrasonic vocalization activity (61) and its specific neuronal isoform enhances nerve-mast cell interaction (62). A *Cadm1* mutation has been identified in people with autism spectrum disorder who have impaired speech and language. *Syn2* is a member of two neuron-specific phosphoproteins of small synaptic vesicles (63). *Syn2* deletion extensively impaired various aspects of social behaviour and memory, altered

exploration of a novel environment and increased self-grooming (64); inducible expression of this gene has been reported in zebra finch (65). *Lnp*, limb and neural patterns is a gene that exhibits limb and central nervous system expression. *Lnp* mutation alters patterning of the appendicular but not the axial skeleton (66, 67). *Ranbp9* is associated with Alzheimer's disease, and its protein binds low-density lipoprotein receptor-related protein, amyloid precursor protein, and BACE1 and robustly increased Aβ generation (68). Other genes from this module participate in more general cell functions such as repairing and translation. *Cetn4* together with other centrins, participates in nucleotide excision repair (69). *Lsm12* is a candidate for translation-machinery-associated proteins (70). Other members of this module *(Fam189a1*, *Tmem130*) are not functionally characterized.

Mouse 7 genes show broad expression in mouse brain and overlaps significantly with chicken module 17 (Bonferroni-corrected P=0.04; described below), which presents a broad expression with some enrichment in the nidopallium. Genes from this module are associated with neurodevelopment and neurotransmission, but most are poorly characterized. *Syne1* participates in neurogenesis and neuronal migration (71) and its defects lead to a recessive form of cerebellar ataxia (72). *Coro2b* is associated with neuronal cell motility and growth cone advance (73). *Dtnbp1* is associated with glutamatergic neurotransmission and is a leading susceptibility gene candidate in schizophrenia (74). The functions of other members of this module (*Acsbg1*, *Fam195a*, *Prmt8*, *Snrpa1*, *Mrpl12*, *Uqcrh* and *Timp4*) have not been characterized yet.

Mouse 11 genes have a broad expression pattern and overlap significantly with chicken module 18 (Bonferroni-corrected P=0.04; described below) whose genes' expression is enriched in the nidopallium and dorsolateral corticoid area. Most genes from this module

are poorly functionally characterized. *Grin2b* encodes an excitatory neurotransmitter receptor involved in neuronal development (75). *Grin2b* has been associated with mental retardation (76), schizophrenia (77) and autism (78). Other genes are associated with more general functions: *Tmem138* is involved in ciliogenesis (79), and in Joubert syndrome, which is characterized by absence or underdevelopment of the cerebellar vermis and a malformed brain stem (80). *Fermt2* is implicated in integrin activation (81). *C1qb* has been associated with immune response to trigeminal pain (82). Other genes are poorly characterized: *Anapc16*, *Zcchc9*, *Nfkbia*, *Fmod*, *Sumf2* and *Dnajc1*.

Mouse modules from the first WGCNA network

Mouse turquoise genes are associated with broad expression with some enrichment in the hippocampus and overlaps significantly with chicken module 2 (Bonferroni-corrected P=0.002, described below), whose genes' expression is enriched in the striatum, mesopallium and hippocampus. Genes from this module are associated with general cell functions that are not restricted to the nervous system. *Map2k2* encodes a protein known to play a critical role in mitogen growth factor signal transduction and participates in regulating gliogenesis in the developing cerebral cortex (83). Mutations in this gene cause cardiofaciocutaneous syndrome, a disease characterized by heart defects, mental retardation, and distinctive facial features (84). *Ctsd* encodes a protease similar to pepsin A, and mutations within it are involved in the pathogenesis of several diseases, including breast cancer and possibly Alzheimer's disease (85). *Tpcn1* encodes voltage-gated Ca^{2+} and Na^{+} channel subunits (86). *Stk11* participates in organ development regulating cell polarity and division (42); its variants are also highly associated with cancer (43, 44). The *Akap8l* gene product facilitates constitutive transport that likely interacts with cellular export proteins (87). *Fbxw5* encodes a member of the F-box

protein family that function in phosphorylation-dependent ubiquitination (88). *Gna11* has been associated with melanoma (89, 90).

Mouse brown is a striatal module (see Main Text) that overlaps significantly with chicken modules 1 (Bonferroni-corrected P=0.01), 6 (Bonferroni-corrected P=$1.9x10^{-5}$) and 14 (Bonferroni-corrected P=0.02; described below). Chicken modules 6 and 14 are striatal modules and chicken 1 genes exhibit expression enrichment in the dorsolateral corticoid area.

Chicken modules from the second WGCNA network

Chicken 1 genes show broad expression with some enrichment in the dorsolateral corticoid area. Functional analysis shows contrasting associations with growth and cellular biology, but most of these genes are poorly functionally characterized. *Ptk2* pathways mediate cortical dendrite arborization (91). *Negr1* is associated with obesity and body mass phenotype (92). *C1galt1c1*, also known as Cosmc, is required for T-synthase folding, transport to the Golgi and its enzyme activation (93, 94). *Polr2f*, among other polymerases, is responsible for synthesizing messenger RNA in eukaryotes (95). *Erlec1* influences cellular stress-response pathways to promote metastasis (96). *Dpcd* is deleted in primary ciliary dyskinesia, involving hydrocephalus, sinusitis, and male infertility (97). *Cenpv*, *Alg14*, *Sepx1* and *Utp11l* are poorly functionally characterized.

Chicken 2 presents a broad gene expression profile, with preferential expression in the striatum, mesopallium and hippocampus; this module significantly overlaps with mouse module 1. Functional analysis indicates a neurophysiological role in synapses, cellular physiology, and cell contact. *Cdh2* encodes a protein that participates in cell interactions,

mediates synaptic adhesion complex and is linked to multiple cytoskeletal elements in hippocampal neurons (98). The general expression pattern of *Cdh2* has been described in chicken brain as showing similarity to the transcriptomic profile presented in this study (99). *Pcdha1* is involved in establishment of axonal projection and has been associated to learning and memory functions (100). Ankrd17 is involved in protein-protein interactions and *Ankrd17* mutant mice embryos die between embryonic day (E) 10.5 and E11.5 due to cardiovascular defects (101). *Tnrc6a* encodes an Argonaute protein, having a catalytic component of the RNA-induced silencing complex that participates in post-transcriptional gene silencing through RNA interference (RNAi) and microRNA pathways (102). *Ormdl3* has been strongly linked with asthma and *in silico* has been associated with glioma risk (103). Other genes from this module (*Sfrs18, Gpr137c, Wsb1 and Ralgapb*) are poorly functionally characterized.

Chicken 3 is a hippocampal-enriched module. Genes from this module show functional association with neurodevelopment, neurite outgrowth, myelination and neuronal physiology support. Alterations in these genes are associated with a wide spectrum of neuropathologies. *Fa2h* participates in fatty acid 2-hydroxylation in postnatal mouse brain (104) and in formation of alpha-hydroxylated galactosylceramide in myelin (105). This gene is associated with neurodegeneration and neurodysfunction (106) and a form of hereditary spastic paraplegia (107). *Tcof1* encodes a centrosome- and kinetochore-associated protein that is critical for spindle fidelity and mitotic progression, and is critically required for proper cortical neurogenesis (108). *Arl6ip5* disruption in mouse results in increased neuronal glutathione content, neuroprotection against oxidative stress and an improved performance in motor/spatial learning and memory tests compared with wild-type mice(109). *Acap2* encodes a protein that has a dual function during neurite

outgrowth participating as a *Rab35* effector and as an Arf6-GTPase-activating protein(110). *Rapgef5* has been associated with telencephalic neurogenesis, particularly in the development of GABAergic interneurons (111). Other genes from this module are associated with more general biological functions. *Rnf139* (112) and *Senp5* participate in protein degradation and recycling (113, 114). *Cbx7* controls cellular lifespan (115) and embryonic stem cell fate commitment (116) and participates as a tumour suppressor (117). *Slco3a1* is poorly functionally characterized.

Chicken 5 is a nidopallium and striatum enriched module. Genes from this module participate in diverse functions. *Ctnnb1* encodes a cytoplasmic component of the classical cadherin adhesion complex that forms the adherens junction in epithelia and mediates cell-cell adhesion in many other tissues; it is also a key signaling molecule in the canonical Wnt signaling pathway that controls cell growth, migration and differentiation during both normal development and tumorigenesis (118). *Faim* encodes an apoptotic inhibitor (119) and is a neurite growth factor (120). Different spliced forms are expressed differentially: FAIM-L is dominantly expressed in the brain whereas FAIM-S is widely expressed in many tissues (119). *Sdc4* encodes one of the principal heparan sulfate-carrying proteins on the cell surface; it is involved in cell-matrix adhesion, migration, neuronal development, and inflammation (121). *Clcn3* mice mutation results in hippocampal degeneration (122). *Txndc12* encodes a protein that has roles in redox regulation, defense against oxidative stress, refolding of disulfide-containing proteins, and regulation of transcription factors (123). The *Dram2* product regulates p53-mediated cell death by autophagy (124). The functions of other gene members (*Spats2l*, *Cstf1*, *Wdr70* and *Clip4*) are not characterized in detail.

Chicken 6 module genes show highest levels of expression in the striatum followed by the dorsolateral corticoid area. Genes from this module are involved in protection, proliferation, motility and synaptic transmission and some of them are associated with syndromes. *Drg1* encodes a cytoplasmic protein involved in stress responses, hormone responses, cell growth, and differentiation (125). *Cfl2* encodes a member of F-actin depolymerizing factors essential for actin driven motility, by accelerating actin filament turnover (126). In mammalian development, regulation of the actin cytoskeleton by the F-actin depolymerizing factor n-cofilin is critical for the epithelial-mesenchymal transition as well as cell proliferation (126). The *Nsg1* encoded protein is the most important early endosomal protein in receptor recycling. *Nsg1* may participate in neurons and germ-cell chemotaxis and endocytosis machinery (127) and plays a critical role in apoptosis as a mediator of p53 (128). *Sv2c* encodes a member of integral proteins localized on the surface of synaptic vesicles in all neurons. SV2 proteins appear to play an important role in synaptic vesicle exocytosis and neurotransmitter release may contribute to the regulation synaptic transmission in the basal ganglia (including cholinergic striatal interneurons and nigro-striatal/mesolimbic dopamine neurons) and is likely to be the receptor of the botulinum neurotoxin A (129). *Lancl1* encodes a glutathione-binding protein, its overexpression inhibits cystathionine β-synthase and its downregulation reduces H_2O_2 or glutamate-induced neuronal damage (130). *Jag2* encodes a notch signalling pathway ligand that is expressed in virtually all postnatal neurons and is transiently expressed in cerebral blood vessels and choroid plexus during the first postnatal week (131). *Jag2* plays an essential role during limb, craniofacial, and thymic development in mice (132). The *Mrpl3* protein product has been associated with Tourette syndrome (133). *Mpped2* encodes an enzyme whose activity and its human orthologue is deleted in patients with WAGR (Wilms tumor, aniridia, genitourinary

anomalies, and mental retardation) syndrome (134). *Unc119b* encodes a protein that acts as a chaperone co-factor in the transport of myristoylated G protein α-subunits and src-type tyrosine kinases (135).

Chicken 14 is a striatal-enriched module. Some genes in this module are known to have enriched expression in mouse striatum, including known striatal markers such as dopamine receptor (*Drd2*) and other novel markers without described function such as 2610019F03Rik. The top 10 genes from this module are generally associated with differentiation, neuronal development, survival, synaptic plasticity and functional networks. *Plxdc2* encodes a protein that has been associated with induction of proliferation in neurons during mouse and chick development (136). *Csnk1d* encodes a clock protein of the mammalian circadian oscillator (137). Drd2 is a G protein-coupled dopamine receptor that inhibits adenylyl cyclase activity, is highly expressed in the striatum and is significantly associated with schizophrenia and its phenotypes (138). *Rasgrp1* is considered as a striosome marker, participates in Ras signal transduction pathways that are involved in cellular differentiation, neuronal survival and synaptic plasticity, and its expression is not detected in glial cells (139). *Dapk1* encodes a protein that binds the NMDA receptor mediating NMDA functions and brain damage in stroke (140). *Mbnl2* mutant mice develop myotonia and have skeletal muscle pathology consistent with human muscular dystrophy (141). Possible functions of other gene members (*Tmem41a*, *2610019F03Rik*, *Tmem90b* and *Ptpdc1*) are poorly characterized.

Chicken 15 is enriched in the hippocampus and striatum. Genes from this module are associated with diverse neuronal functions, such as neurite growth and neurogenesis and some are associated with neural syndromes. *Sfswap* encodes a factor that regulates splicing of CD45 and fibronectin, and its own transcripts (142), Sfswap contains

domains that are required for regulation of exon 10 splicing of the Tau protein, and mutations in this domain cause inherited frontotemporal dementia (143). *Kras* encodes a member of Ras signaling proteins that when up-regulated in neurons promotes gliosis and astrocytoma formation (144). *Trio* encodes a protein that participates in motor axon guidance and dendritic morphogenesis in Drosophila. *Trio* knock-out mice show defective cerebella with no granule cells in the internal granule cell layer due to aberrant granule cell migration as well as abnormal neurite growth and presents severe signs of ataxia cerebella (145). The *Ash1l* gene product regulates neurogenesis in zebrafish and exhibits a conserved expression pattern in mouse, zebrafish and Xenopus (146). *Prpf6* mutations are associated with retinitis pigmentosa (147). *Wnk1* defect is the cause of hereditary sensory and autonomic neuropathy type II (HSANII), an early-onset autosomal recessive disorder characterized by loss of perception to pain, touch, and heat due to a loss of peripheral sensory nerves and is additionally associated with hypertension disease (pseudohypoaldosteronism II) (148). This gene have been considered to be widely expressed, with highest levels in the testis, heart, kidney, and skeletal muscle and lower expression in brain (149) - nevertheless we observe its high expression in the mouse brain that is supported by the Allen Brain Atlas. *Gdap1* encodes a protein located in the outer mitochondrial membrane and it seems that may be related with mitochondrial network dynamics (149). Its mutations are the cause of Charcot-Marie-Tooth type 4A disease (CMT4A) (150). Other members (*Plekhj1*, *Gm5567* and *Stxbp5l*) of this module are poorly characterized.

Chicken 17 genes exhibit broad expression with some enrichment in the nidopallium. Genes from this module tend to participate in proliferation, transmission, and cell physiology. *Tfdp1* encodes a protein that binds to the promoters of various cell cycle

genes and coordinates events in the cell cycle with transcription by its cyclical interactions with important regulators of cellular proliferation such as the retinoblastoma tumour-suppressor gene product (Rb) and the Rb-related protein (151-153). *Gria4* encodes a glutamate receptor that participates in excitatory transmission and is thought to be involved in synaptic plasticity and the development of functional neural circuitry through the recruitment of other AMPA receptor subunits. *Gria4* has been implicated in schizophrenia and cognitive disorders (154). *Trnt1* encodes a protein that generates and maintains tRNA CCA-termini (155) whose modification plays a pivotal role in the fidelity of the translational process (156) in the cell and mitochondria. The protein encoded by *Khdrbs2* is highly expressed in the brain and may function as an adaptor protein, regulating tissue-specific splicing of various proteins (157). *Stx7* encodes a SNARE protein enriched in cells of the immune system (158). *Rasl11b* encodes a small GTPase belonging to a Ras subfamily of putative tumor suppressor genes (159). *Magi3* encodes a protein that interacts directly with lysophosphatidic acids and regulates its ability to activate Erk and RhoA (160). *Rtn1* encodes a transmembrane protein that participates in maintaining the tubular structure of the endoplasmic reticulum (161). Other members (*Ephx4* and *1110057K04Rik*) of this module are poorly characterized.

Chicken 18 genes exhibit a tendency for expression enrichment in nidopallium and dorsolateral corticoid area. Genes from this module are associated with diverse functions. *Spred1* encodes a protein whose deficiency produces defects in short- and long-term synaptic hippocampal plasticity and deficits in hippocampus-dependent learning related to hyperactivation of the Ras/ERK pathway (162). *Eif3b* encodes a protein that is essential for initiation of protein synthesis, and promotes the proliferation of glioblastoma cells (163). *Smyd2* encodes a protein that catalyses the transfer of methyl

groups to acceptor lysine residues on histones and other proteins (54). *Fabp5* encodes a protein that is part of the fatty acid binding protein family that participates in fatty acid uptake, transport, and metabolism. In humans mutations in this gene have been associated with psoriasis and type 2 diabetes. In mouse deficiency of this gene in combination with a deficiency in *Fabp4* confers protection against atherosclerosis, diet-induced obesity, insulin resistance and experimental autoimmune encephalomyelitis (the mouse model for multiple sclerosis) (164). *Rcn2* encodes a protein localized to the endoplasmic reticulum and belongs to the CREC protein family; it is involved in various diseases, abnormal cell behaviour and may have roles in immunity, redox homeostasis, cell cycle regulation and coagulation (165). *Rnf11* encodes a member of the A20 ubiquitin-editing protein complex and a negative regulator of the NF-κB pathway implicated in immune response (166). Other gene members (*Polr3f*, *Smim4*, *Ccdc34* and *BC013529*) of this module are poorly characterized.

Gene connectivity is less conserved than median expression level

This work is in accordance with previous studies comparing humans and chimps (167) and mice and humans (168) that demonstrated that gene connectivity is less conserved than median expression level. Unsurprisingly, the correlation of median expression level between mice and chickens (r_S=0.49) was significantly lower than between mice and humans (r_S=0.60) (two-tailed $P<2x10^{-14}$). The apparent lack of correlation in gene connectivity in this study, considerably lower than the r_S=0.27 observed between mice and humans (two-tailed $P=5x10^{-40}$), may be due in part to batch effects.

The study comparing transcriptional networks in mice and humans found that most, but not all, co-expression modules were shared by both species. In contrast, only a handful

of modules overlapped in this study. Possible contributors to there being considerably less overlap include (1) the greater phylogenetic divergence between mice and chickens, (2) the fact that most of the anatomical structures included in this study simply do not exist in a recognisable functional and homologous form in the other species (while humans and mice share a much more similar basic layout), and (3) less power in this study to construct robust modules owing to fewer available samples.

Functional enrichments guide module interpretation

No modules were significantly enriched for functional annotations following a Bonferroni correction that accounted both for the number of tests performed given a fixed module and fixed annotation source and for the number of modules within a species, but not for the number of species or for the number of annotation sources. Ideally, all of these would be taken into account; however, a more stringent Bonferroni correction would necessarily produce no significant results as well. From this study, one cannot then say with confidence that genes co-expressed across these samples share some common functions more than expected by chance. However, this has been previously demonstrated in mammalian brain, where the number of dissections was larger, and the expected 'noise' in co-expressed modules correspondingly lower (167, 168).

It is quite conservative to control the familywise error rate in high-throughput studies since its null hypothesis is that there are no true differences. Furthermore, the assumption of independence is violated with functional annotations: all genes annotated with '(in) plasma membrane' should also be annotated with '(in) membrane'. Considerable overlap is also to be expected between some annotation sources (for

example, ‘myelination’ in GO biological process, ‘myelin sheath’ in GO cellular component and ‘structural constituent of myelin sheath’ in GO molecular function were all enriched in the mouse magenta module), while less overlap might be expected between others (for example, between MGI knockout phenotypes and GO cellular component). This dependence will make a Bonferroni-corrected familywise error rate more conservative still. By relaxing the boundaries within which the correction is applied, one might instead identify suggestive but nevertheless inconclusive trends of functional clustering. Thus, we instead applied the Bonferroni correction to take into account the number of tests performed given a fixed module and fixed annotation source, but not for the number of modules within a species, the number of species or the number of annotation sources.

Three mouse modules and four chick modules were significantly enriched for one or more functional annotations following a Bonferroni correction using this relaxed criterion. While interpreting the results of these enrichments one must be mindful that many of the annotations were generated through the study of humans or common model organisms such as mouse, fly, worm or yeast (and carried over to orthologs or proteins sharing the same characterised domain). Few if any annotations would be derived from chick. We would expect a gene product’s involvement in an ancestral biochemical process (which may only require one or two interactions) to be more conserved than its higher-order physiological effects (which may depend on highly complex signalling pathways or a similar response across multiple tissues to a physiological stressor).

We will first discuss enrichments in modules that did not significantly overlap a module in the other species. We shall discuss the other functional enrichments in context of the observed cross-species overlaps.

Mouse tan has an intriguing expression profile, being found primarily in upper layers of lateral, but not dorsal, cortex. Unfortunately, this same lateral-dorsal difference is also seen in measures of 3′ bias, proposed above as a proxy for RNA quality (Fig. S41). While correlation with this metric of 3′ bias does not necessarily imply that this apparent gene co-expression module is driven in part by differences in RNA quality, the possibility cannot be excluded. It may contribute nothing to the module, or it may contribute entirely to the module. Likewise, the apparent functional enrichments could be driven by differences in genes affected by RNA quality. The largest functional enrichment (by fold difference) was "respiratory chain". Many things could affect the apparent expression of genes for mitochondrial proteins encoded in the nuclear genome. For example, if transcripts encoding mitochondrial proteins were shorter than other transcripts, they would appear to be more highly expressed in degraded, poly(A) selected RNA. Thus, without additional confirmation, one cannot attribute this result to biology with any certainty. The Allen Mouse Brain Atlas might have provided such confirmation, but the top genes in the tan module are not represented by high-quality coronal images that would allow one to confidently compare dorsal and lateral cortex.

There were two chick modules that had functional enrichments but did not have a significant overlap with a mouse module. Chick brown is plagued by very high correlations with numerous technical variables, including RNA Integrity Number. Thus, without biological replicates, it is difficult to separate biological from technical effects. In contrast, there was no evidence that the intriguing expression profile of chick red (being particularly low in hippocampus) was tainted by correlation with technical variables. The gene symbol of the mouse ortholog of the most correlated gene with the chick red eigengene is *Cul4b*, whose human ortholog is an X-linked mental retardation

gene that (in mouse) targets ubiquitylation of the H3K4 methyltransferase component WDR5 and regulates neuronal gene expression (169). In contrast to chick, *Cul4b* was not downregulated in mouse hippocampus, as confirmed in the AMBA (8).

Some modules are significantly conserved

Here we used two approaches to identify co-expression modules preserved in the other species. Since the specific modules produced are dependent on parameters selected, the first approach calculates various network preservation metrics in a second species of the genes in modules built in the first species (16). The second approach identifies statistically significant overlap of co-expression modules based on orthologous genes within those modules.

The first approach indicated two mouse modules that were preserved in chick and three chick modules that were preserved in mouse. Only two modules – mouse salmon and chick greenyellow – were not correlated with technical variables. (These two modules had significant cross-species matches in the next analysis as well.) In this type of analysis, correlation with technical variables is of particular concern because there is not a corresponding module in the other species that one could test for correlation with technical variables. Nevertheless, the significance of these two modules after a proper Bonferroni correction is strong evidence that some modules are conserved. Note that the near-zero correlation of connectivity observed might increase with more samples, which would reduce the noise in ranking connectivity in both species.

The second approach yielded five to six overlapping modules between the two species, which we will now discuss in turn.

Conserved striatal genes are functionally important in striatum

The chick greenyellow/mouse brown overlap strongly marked striatum in both species, and neither module was correlated with technical variables. A similar module to chick greenyellow, chick green, also overlapped with mouse brown. This module may be partially tainted by technical artefacts (perhaps this accounts for its split from chick green). It is not likely to be driven entirely by technical artefacts for several related reasons. First, green chick and mouse brown both shared two significant functional enrichments: "signal transduction" and "adrenoceptor activity". Indeed, both α and β adrenoceptors are highly expressed in striatum (170-174). Second, chick green and chick greenyellow have similar expression patterns. If the overlap were purely due to technical artefacts, there would be no reason chick greenyellow would have to mark striatum in chick. Third, *in situ* hybridizations in mouse generally confirmed the expression of the top consensus markers from both overlaps (Fig. S52). Fourth, homeobox transcription factor *Dlx6*, the top consensus marker for chick green/mouse brown, is a known marker of subpallium (175), and striatum was the only subpallial dissection in either species. Other chick green/mouse brown consensus markers including transcription factor *Foxo1* (176), phosphodiesterase *Pde10a* (177), and *Drd1a* (178) are notably enriched in striatum. This result is as expected, since the striatum in both species is of uncontroversial homology.

One might expect the genes that are conserved in expression to be particularly important in striatal function, and indeed they are. *Dlx6* regulates molecular properties of the striatum (179). Inhibiting *Pde10a* changes the physiology of striatal neurons (180). The knockout of *Drd1a*, or the pharmacological inactivation of its gene product, changes the expression of G-protein signalling transcripts in striatum (181). Striatal expression of

Adcy5 differed between mice selectively bred for excessive exercise or obesity (182). *Pde7b* had altered striatal expression in a mouse model of Parkinson's disease (183). *Rps6ka5* knockout mice show decreased locomotor sensitisation to repeated injections of cocaine; and the protein it encodes, MSK1, is a striatal kinase acting downstream of ERK that phosphorylates CREB and histone H3 and induces expression of c-Fos and dynorphin (184). Mouse knockouts for *Tac1*, which encodes substance P, show reduced sensitivity to intense pain (185) (striatum is particularly important in pain modulation (186)). Perhaps the most direct evidence for a supraspinal-substance P behavioural link comes from a study demonstrating an altered stress-induced analgesia and reduced aggressiveness in response to a territorial challenge in homozygous mouse knockouts for the receptor of substance P (187). Substance P induces dopamine release in striatum, and also speeds dopamine clearance (188).

It also suggests new directions for future research. For example, mouse knockouts of *Dgkb* have impaired spatial and long-term memory, and hippocampal neurons have reduced branch- and spine-formation (189). However, to the best of our knowledge, nobody has examined the striatum of these mice. Furthermore, the subcellular localisation (and co-localisation) of the enzyme encoded by *Dgkb* suggests a role in the Gq-coupled receptor signalling pathway in the medium spiny neurons that dominate the striatum (190).

Many conserved hippocampal genes are presently of unknown function

The second and final positive control dissection, hippocampus, also gave the expected results. Chick magenta and mouse black marked hippocampus in each species, and there was no evidence that either module was tainted by technical artefacts. In mouse, this

was further confirmed with the Allen Mouse Brain Atlas.

While many of these genes have been described as being in hippocampus, there has generally been little work on their potential hippocampal functions or behavioural effects. *Abi1* knockout mice show embryonic lethality and malformations in their developing brains and hearts (191). Uniprot annotations suggest *Ptk2b* is most abundant in the brain, especially hippocampus and amygdala, but also found in the kidney, spleen and lymphocytes. The protein encoded by *Prickle2* interacts with PSD-95 and NMDA receptors at the hippocampal postsynaptic density (192), and regulates neurite outgrowth in mouse neuroblastoma Neuro2a cells (193). Suggestively, mutations in *Prickle2* orthologs cause seizures in flies, mice and humans (194). (Note that hippocampus has an important role in epilepsy (195)) The fly mushroom body, which is field homologous to the vertebrate pallium (196), is involved in modulating sensitivity to seizures in flies (197). It is noTable that hippocampal lesions cause a highly similar constellation of impairments in birds and mammals (198).

Shared expression in functionally analogous areas may be evidence of convergent evolution

Perhaps the most intriguing overlap associates chick nidopallium (chick black) with mouse layer IV neocortex (mouse orange). This is interesting because chick nidopallium derives from the (*Emx1*-negative) ventral pallium, while mouse neocortex is generally thought to derive from the (*Emx1*-positive) dorsal and/or lateral pallium. (It is nevertheless possible that there remains an undiscovered tangential migration in mouse from ventral pallium into neocortex, or that there is an overreliance the expression pattern of *Emx1* to best define the ventral pallium lineage in chick and that other, yet

undiscovered, developmental markers may suggest nidopallium and neocortex share the same progenitors as defined by other developmental markers.)

However, we do not believe this makes a strong argument for one-to-one homology of these two regions, for several regions. First, there were only five genes that explained this particular significant overlap (0.5 expected, yielding a ten-fold enrichment). Gene co-expression modules can sometimes mark a larger trend, since some of the genes within them should tend (by mathematical necessity) to be relatively specific to a cell type or to a process, or even to physically co-located cell types or processes (the genes needn't necessarily be regulated in the same cell type if they are expressed in cell types that co-occur). However, the significant overlap of two modules need not *necessarily* imply that there are many more genes under similar regulatory control if there is a plausible alternative explanation.

In this case, an alternative explanation could be convergence. The three genes for which the Allen Mouse Brain Atlas had *in situ* hybridization images are frequently expressed in brain regions involved in processing or analysing streams of sensory input. For example, *Fam19a2* is found predominantly in layer 2a of entorhinal cortex, layer 2 of piriform cortex, taenia tecta and the anterior olfactory nucleus. All of these regions except the taenia tecta receive input from the lateral olfactory tract; and the taenia tecta receives input from the medial olfactory tract (199). These regions, which were not included in the dissections, express *Fam19a2* even more highly than does layer IV of neocortex. Some of the regions in which *Dctn3* and *Rorb* were expressed have similar sensory processing or relay functions: for example, layer 2 of piriform cortex for *Dctn3* and layer IV neocortex (especially primary somatosensory cortex), preoptic area, medial preoptic area, nucleus of the lateral olfactory tract, lateral geniculate nucleus, lateral amygdaloid

nucleus, superior colliculus, laminae 1-6 of the spinal cord and the ventral postero-lateral thalamic nucleus for *Rorb*. There is also some functional evidence consistent with a sensory processing explanation of potential convergence: *Rorb* protein overexpression is sufficient to produce 'barrel'-like clustering of layer IV neurons (which may assist in processing large streams of information), even in the absence of thalamic innervation (200). *Fam19a2* is hypothesized to have a role in axonal guidance (201). (Both chick nidopallium and mouse layer IV are thalamic recipients.) Finally, if these genes indicated homologous regions from an as-yet-undiscovered tangential migration of neurons into layer IV of neocortex, one might also expect these genes to be found in other ventral pallial derivatives, namely pallial amygdala. This, however, was not the case.

Since these genes have little overlap with one another in mouse outside the dissected regions, they may not be under co-ordinated regulatory control. If this were the case, genes might be recruited one at a time in a scenario of convergence. The small number of overlapping genes increases the plausibility of such a process.

A possible, though unlikely, explanation is technical confounds. Mouse orange was moderately positively correlated with 3′ bias and pre-mRNA fraction (both r=0.56; two-tailed P=0.025). Note that a perfect correlation would imply highest expression in dorsal and lateral cortex sections A, while the module eigengene expression in sample A is instead middling. Chick black was not significantly correlated with any technical variable. Chick black's 3′ bias correlation was -0.35 (two-tailed P=0.44), which may differ from mouse orange (two-tailed P=0.080). Note that, of all the chick module eigengenes (not just those which had a corresponding partner in the other species), chick black was the most specific to nidopallium. Likewise, mouse orange was the most

specific to dorsal and lateral neocortex section B (layer IV). We consider a possible correlation with RNA quality to be a less satisfactory explanation for this overlap since it would require the chick orthologs of mouse genes whose transcripts happen to degrade faster than others to be preferentially specific to chick nidopallium. Moreover, a correlation at this level of significance may happen by chance since we conservatively did not apply a multiple-testing correction while looking for potential confounds. Thus the apparent correlation with mouse orange may not be significant and is unlikely to explain this particular result satisfactorily.

With the exception of *Rorb*, little is known about the functions of these genes. This observation is thus suggestive for future research.

Conserved amniote markers of oligodendrocytes

The fourth overlapping set (chick cyan and mouse magenta) unambiguously picks out oligodendrocyte markers. All of the top five consensus genes were oligodendrocyte markers. Mouse magenta was highly and significantly enriched for several myelin-related annotations after a Bonferroni correction, and chick cyan was also highly enriched (120-fold for "structural constituent of the myelin sheath") for these same annotations (these were significant at a single test level). Furthermore, many of the annotations from one functional source that were enriched in mouse were the most enriched annotation for that functional source in chick. Finally, the module eigengene expression in mouse is as one would expect from oligodendrocytes, being found primarily in the lowest neocortical laminar sections. While both modules were correlated with technical variables this is unlikely to drive this particular overlap unless oligodendrocyte-specific transcripts globally happen to degrade *much* more slowly than

other transcripts. Since less functional work has been done on oligodendrocytes than on neurons, these genes, which are specific to oligodendrocytes over long evolutionary timescales, are of particular interest for functional studies.

Batch effects confound interpretation of the final module

The final overlap, chick turquoise and mouse salmon, remains an enigma. Mouse salmon was one of only two mouse modules significantly preserved in chick by the first approach to preservation. Chick turquoise was very highly correlated with technical variables, while mouse salmon was notably not, and even differed significantly from chick turquoise. While the top five consensus genes were generally housekeeping, both modules were differentially expressed across samples, with chick turquoise found predominantly in hyperpallium and (to a lesser degree) arcopallium and mouse salmon found in striatum, hippocampus, and *lateral* neocortex. Little is known about these genes in brain. Given the extremely high correlation of chick turquoise with technical variables and the lack of relevant functional information amongst the consensus gene set, it is difficult to interpret this result without more data. Note that chick turquoise was the extremely large set in chick. It could be that the significant overlap is a true overlap, and that the consensus genes are still driven by batch effects.

Genes specifically expressed in mouse but not in chicken that have spatially-relevant functional consequences

Several genes are specific to a structure in mouse, and functional in that structure despite being low in the corresponding homologous structure in chick. For example, *Rarb*, a retinoic acid receptor, is specific to mouse striatum but expressed at a very low level in chick striatum (Fig. S22). Nevertheless, *Rarb* plays an important role in mouse striatum,

as *Rarb* null mutant mice have defective striatal dopaminergic neurogenesis, resulting in deficits in motor behaviour (202). Like *Rarb*, *Gpr6* is specific to mouse striatum but rare in chick striatum. *Gpr6*-deficient mice produce less cAMP in their striatum and have enhanced instrumental conditioning (203). In another example, *Npy2r*, high in mouse hippocampus but low in chick hippocampus, is thought to mediate neuropeptide Y activity in rat hippocampus, with effects on hippocampal-based spatial discrimination learning (204).

Table S1: RNA concentration and quality metrics of mouse samples.

Sample	Abbreviation	ng/µL	260:280	260:230
Dorsal cortex A (layers I-III)	MDC_A	51.8	2.03	0.87
Dorsal cortex B (layer IV)	MDC_B	128.7	1.96	1.11
Dorsal cortex C (layers IV-V)	MDC_C	84.6	1.92	0.61
Dorsal cortex D (layer V)	MDC_D	127.4	1.91	0.64
Dorsal cortex E (layer VI)	MDC_E	64.9	2.04	0.73
Dorsal cortex F (layers VI-VIb)	MDC_F	298.8	1.7	0.57
Lateral cortex A (layers I-III)	MLC_A	67.7	2.02	1.07
Lateral cortex B (layer IV)	MLC_B	128.8	2.06	1.23
Lateral cortex C (layers IV-V)	MLC_C	31.5	2.01	0.66
Lateral cortex D (layers V-VI)	MLC_D	79.4	2.04	0.76
Lateral cortex E (layer VI)	MLC_E	55.5	1.95	1.03
Lateral cortex F (layer VIb)	MLC_F	13.7	1.93	0.46
Hippocampus	MHipp	285.3	2.03	1.2
Claustrum/endopiriform	MCE	82	1.92	1.1
Pallial amygdala	MPA	125.8	1.92	1.14
Striatum	MS	99.8	2.05	1.41

Table S2: RNA concentration, quality metrics and RIN of chicken samples.

Sample	Abbreviation	ng/µL	260:280	260:230	RIN
arcopallium	CA	203.7	2.11	2	6.9
dorsolateral cortex	CDLC	255.1	2.11	2.01	8.3
hippocampus	CHipp	490.1	2.12	1.83	7.6
mesopallium	CM	266.5	2.1	1.97	8.3
nidopallium	CN	696.6	2.14	1.95	not called
hyperpallium	CH	232.2	2.13	1.76	5.3
striatum	CS	174.1	2.09	1.9	8.4

Table S3: Insert sizes and mapping information for lanes from all libraries.

Species	Region	Date YY-MM-DD	Sequencer	Flowcell	Ln	IIS	ISS std dev	good reads (M)	alignments (M)	properly paired	both mates mapped
Chicken	Arcopallium	091029	DOGFISH	42WDDAAXX	1	52	31	28.1	24.7	15.0 (61%)	22.4 (91%)
		091120	DOGFISH	42VFJAAXX	6	52	31	34.7	30.4	18.4 (61%)	27.3 (90%)
		091124	ALPACA	42UY8AAXX	1	55	33	30.2	26.8	17.2 (64%)	24.5 (91%)
	Dorsolateral cortex	091029	DOGFISH	42WDDAAXX	2	52	31	25.8	20.5	13.3 (65%)	18.4 (89%)
		091120	DOGFISH	42VFJAAXX	7	52	31	32.9	26.3	16.8 (64%)	23.2 (88%)
		091124	ALPACA	42UY8AAXX	2	53	32	29.0	23.5	15.5 (66%)	21.0 (89%)
	Hippocampus	091029	DOGFISH	42WDDAAXX	3	90	30	27.9	21.8	11.7 (54%)	19.1 (88%)
		091120	DOGFISH	42VFJAAXX	8	90	30	32.6	25.7	13.9 (54%)	22.4 (87%)
		091124	ALPACA	42UY8AAXX	3	90	30	29.7	23.2	12.5 (54%)	20.3 (88%)

	Mesopallium	091029	DOGFISH	42WDDAAXX	4	89	28	28.8	23.4	14.1 (61%)	20.6 (89%)
		091124	ALPACA	42UY8AAXX	4	89	28	30.1	24.2	14.9 (61%)	21.5 (89%)
	Nidopallium	091120	DOGFISH	42VFJAAXX	3	120	31	37.8	31.4	18.1 (58%)	27.4 (87%)
		091124	ALPACA	42UY8AAXX	5	121	31	30.9	26.6	15.8 (60%)	23.8 (90%)
	Hyperpallium	091120	DOGFISH	42VFJAAXX	2	53	30	29.7	26.3	16.0 (61%)	23.5 (90%)
		091124	ALPACA	42UY8AAXX	7	53	31	26.4	23.7	14.7 (62%)	21.5 (91%)
	Striatum	091120	DOGFISH	42VFJAAXX	1	62	31	32.7	27.2	17.4 (64%)	24.5 (90%)
		091124	ALPACA	42UY8AAXX	6	63	32	27.2	23.0	15.2 (66%)	21.1 (92%)
Mouse	Dorsal cortex A (layers I-III)	100720	PLATYPUS	621GFAAXX	1	287	30	31.0	30.0	18.9 (63%)	27.4 (91%)
		100805	LEOPARD	6232KAAXX	1	287	31	60.4	59.8	38.9 (65%)	55.9 (93%)
	Dorsal cortex B (layer IV)	100720	PLATYPUS	621GFAAXX	2	290	28	58.3	55.3	34.5 (62%)	51.0 (92%)
		100805	LEOPARD	6232KAAXX	2	290	28	69.4	66.2	42.1 (64%)	61.5 (93%)
	Dorsal cortex C (layers IV-V)	100720	PLATYPUS	621GFAAXX	3	291	28	52.3	50.7	30.2 (59%)	47.3 (93%)

		100805	LEOPARD	6232KAAXX	3	291	28	62.3	60.7	36.6 (60%)	56.6 (93%)
	Dorsal cortex D (layer V)	100720	PLATYPUS	621GFAAXX	4	269	30	58.9	58.1	34.7 (60%)	54.9 (94%)
		100805	LEOPARD	6232KAAXX	4	268	30	66.6	65.4	38.9 (59%)	61.2 (93%)
	Dorsal cortex E (layer VI)	100720	PLATYPUS	621GFAAXX	5	284	31	56.8	56.1	33.2 (59%)	53.0 (95%)
	Dorsal cortex F (layers VI-VIb)	100720	PLATYPUS	621GFAAXX	6	279	21	53.6	52.8	29.9 (57%)	49.8 (94%)
		100810	MARMOSET	62328AAXX	5	279	21	61.6	60.6	34.4 (57%)	56.7 (94%)
	Lateral cortex A (layers I-III)	100720	PLATYPUS	621GFAAXX	7	281	40	40.5	39.0	25.3 (65%)	36.1 (93%)
		100810	MARMOSET	62328AAXX	6	281	40	50.7	48.9	32.2 (66%)	45.3 (93%)
	Lateral cortex B (layer IV)	100720	PLATYPUS	621GFAAXX	8	275	32	53.3	52.1	34.3 (66%)	48.7 (93%)
		100831	PLATYPUS	627RLAAXX	5	275	32	60.5	60.0	39.7 (67%)	56.0 (94%)
	Lateral cortex C (layers IV-V)	100720	MARMOSET	621UPAAXX	1	281	40	60.8	60.0	36.8 (61%)	55.7 (93%)
		100810	MARMOSET	62328AAXX	8	283	42	53.0	52.3	31.8 (61%)	48.6 (93%)
	Lateral cortex D (layers V-VI)	100720	MARMOSET	621UPAAXX	2	267	29	47.8	47.1	29.1 (62%)	44.0 (94%)
		100810	WALLABY	62344AAXX	5	268	29	46.8	46.4	29.0 (63%)	43.7 (94%)

	Lateral cortex E (layer VI)	100720	MARMOSET	621UPAAXX	3	268	38	42.3	40.8	23.7 (58%)	37.8 (93%)
		100810	WALLABY	62344AAXX	6	268	38	46.0	44.7	26.1 (58%)	41.6 (93%)
	Lateral cortex F (layer VIb)	100720	MARMOSET	621UPAAXX	4	273	34	47.2	46.2	26.8 (58%)	43.3 (94%)
		100810	WALLABY	62344AAXX	7	272	34	51.7	50.9	29.8 (59%)	48.0 (94%)
	Hippocampus	100720	MARMOSET	621UPAAXX	7	259	34	49.2	49.4	29.9 (61%)	45.8 (93%)
		100810	WALLABY	62344AAXX	4	259	34	52.4	53.1	32.6 (61%)	49.8 (94%)
	Claustrum/ endopiriform	100720	MARMOSET	621UPAAXX	6	276	28	56.3	55.8	33.8 (61%)	52.1 (93%)
		100810	WALLABY	62344AAXX	3	276	28	55.6	55.7	34.2 (61%)	52.6 (95%)
	Pallial amygdala	100720	MARMOSET	621UPAAXX	5	283	32	58.2	55.4	33.1 (60%)	50.6 (91%)
		100810	WALLABY	62344AAXX	2	284	33	51.6	49.5	29.7 (60)	45.7 (92%)
	Striatum	100720	MARMOSET	621UPAAXX	8	283	35	56.3	55.6	34.1 (61%)	51.2 (92%)
		100810	WALLABY	62344AAXX	8	284	35	47.2	47.3	29.2 (62%)	44.5 (94%)

Table S4. Primer sequences used for the generation of species-specific riboprobes.

Target structure	Gene	Probe (bp)	Forward primer	Reverse primer
Nidopallium Layer IV module	*ggRorß*	606	GGCTTGGGCAATCTGAATAA	GAACATCTGCATCCCTCCAT
	mmRorß	609	TGGACATGACTGGGATCAAA	CTCGGTCTGGGGATATCAGA
	ggFam19a2	239	GCATCTGGAGCAAATCATCA	CATTCCTCCCCTTCAAGACA
	mmFam19a2	622	TTGCAGAAAGCAACACAAGG	TGTTGTGAGCCAAGCGTTAG
	ggMyo16	641	TGTGCTCGCTACGATAATGC	TCCTCCCAAACAATTTCAGC
	mmMyo16	644	AGTGCTATCCTGCTGGCCTA	TTGCTTGTCATGCTGGTCTC
Hippocampus module	*ggAbi1*	585	GAGGAGGAAATCCCTTCTGG	TGGCTCCAGGGTTTTGTAAG
	mmAbi1	577	CGGACTCAGCTTCCTCTGTC	TCTCGACAATGTGCCAGTTC
	ggNr3c2	683	CCCATATGACCAGCAAAACC	TGGAAGCAGGGAATCCTAGA
	mmNr3c2	675	GGTCACAGGTCCTCCACACT	AACGTCGTGAGCACCTTTCT
Striatum module	*ggPenk*	637	CCGCTGGCATGTACACTAGA	TCTCTGGGACCTCTTTGGAA
	mmPenk	653	GACAGCAGCAAACAGGATGA	AGCACAAAGCAGCATGTGAC
	ggDrd2	688	GTGCACTGCCAGTATCCTCA	CTTGCCATTAGGCATGGACT
	mmDrd2	665	TGCCTTCGTGGTCTACTCCT	AGAGGACTGGTGGGATGTTG
Hyperpallium	*ggCpne8*	634	CATATGGCATGGCACTGAAG	GGTGCAGGTGATGGCTTTAT
	mmCpne8	675	GGCACTGAAGGCTGTAGGAG	CGTGGTCATATTTGCGTCTG
Mesopallium	*ggRspo3*	630	CCAAGGAGGGTGTGCTACAT	TTGTTTTCCCTCTGCTCTCG
	mmRspo3	603	GCTGTGCAACGTGTTCAGAT	CAATGCTGGACTCCAAACCT
Nidopallium	*ggKcnh7*	607	AGCTGTCTGGGACTGGCTAA	GGATGGTCCAGAACTTGCAT
	mmKcnh7	677	CAGCCTAGCCAGTGTTCTCC	TAGCCACATTCTCGCCTTTT
Dorsolateral corticoid area	*ggOprl1*	684	ACAAGGCCAAAGTGGTGAAC	GCTCTGATGTGGGAGTGGTT
	mmOprl1	639	CAGGCACACCAAGATGAAGA	ACACCCAGTCCTTGAACCAG
Arcopallium	*ggTshz3*	634	CTCACGACACTTTGCAGGAA	GTTTTTGGCACAGGTGAGGT
	mmTshz3	615	GCTCACATGATGGTCACTGG	AAAGTTTGTCTTGGGCATGG

Table S5: Preservation statistics of mouse modules in chicken sorted by overall preservation.

module name	size	Zsummary.pres	Zdensity.pres	Zconnectivity.pres	Z.propVarExplained.pres	Z.meanSignAwareKME.pres	Z.meanSignAwareCorDat.pres	Z.meanAdj.pres	Z.meanClusterCoeff.pres	Z.meanMAR.pres	Z.cor.kIM	Z.cor.kME	Z.cor.kMEall	Z.cor.cor	Z.cor.clusterCoeff	Z.cor.MAR
salmon	61	3.42	5.78	1.05	3.26	4.80	10.13	6.77	4.17	3.84	1.05	-0.32	11.40	2.38	0.46	0.30
steelblue	31	3.23	4.18	2.28	1.04	4.15	6.75	4.21	1.61	1.19	1.54	2.28	14.05	3.39	2.08	1.46
yellow	170	1.99	3.53	0.45	1.61	3.20	4.68	3.85	2.92	2.64	0.45	-1.13	16.75	1.05	0.81	-0.19
grey	400	1.88	1.40	2.36	0.16	1.91	5.18	0.90	1.24	1.02	1.50	2.36	15.68	4.28	2.34	1.13
lightcyan1	26	1.29	4.00	-1.42	2.81	3.13	4.86	4.96	3.04	3.05	-1.42	1.70	15.56	-3.32	-1.15	-1.01
tan	62	1.27	2.65	-0.10	2.76	1.78	2.53	4.50	3.86	4.18	-1.23	-0.10	4.80	0.39	-1.20	-0.68
red	123	1.04	2.06	0.03	2.07	1.42	2.05	2.10	1.55	1.66	0.34	-0.46	15.95	0.03	0.69	0.23
grey60	45	0.93	1.54	0.32	1.48	1.56	1.52	1.80	0.93	1.01	0.32	-0.62	16.25	0.50	0.10	-0.58
white	32	0.91	3.78	-1.95	1.36	3.78	6.74	3.77	0.94	0.84	-1.95	1.16	14.30	-2.62	-0.35	-1.77
plum1	27	0.90	1.78	0.02	0.02	2.36	2.53	1.19	0.88	0.79	0.02	-0.66	16.33	0.25	0.63	-0.12
plum2	21	0.88	2.56	-0.81	1.37	2.78	4.24	2.34	0.66	0.63	-0.81	1.17	16.94	-1.80	-1.47	-0.99
turquoise	212	0.81	-0.32	1.94	1.22	-1.12	-0.85	0.21	0.81	0.37	2.46	1.33	14.31	1.94	-0.25	2.00
blue	191	0.81	2.74	-1.12	1.08	2.90	4.33	2.58	1.16	1.23	-1.12	0.42	13.48	-1.90	0.18	-0.83
darkmagenta	29	0.74	0.37	1.10	0.82	0.40	-0.07	0.35	0.09	0.21	-0.11	1.54	7.99	1.10	-1.36	-0.35
lightgreen	44	0.71	1.33	0.08	2.19	0.47	0.37	3.14	2.69	2.40	0.62	-0.64	10.17	0.08	-0.34	0.00
skyblue3	28	0.64	0.27	1.01	0.15	0.46	0.39	-0.26	-0.87	-0.69	1.01	0.84	8.52	1.03	1.88	0.27
green	140	0.61	0.54	0.69	-0.39	1.24	2.17	-0.17	-1.17	-1.38	0.15	0.69	7.50	1.37	-1.70	-0.45
cyan	50	0.58	3.41	-2.26	1.12	3.02	3.80	4.02	1.78	1.80	-2.26	1.90	10.81	-2.81	-2.24	-2.34
midnightblue	47	0.52	1.66	-0.61	0.05	1.75	2.00	1.58	0.25	-0.32	-1.02	-0.61	17.30	1.65	1.20	-1.04
darkgreen	37	0.47	3.49	-2.55	1.39	3.66	6.07	3.33	1.06	0.97	-2.55	3.28	14.93	-4.67	-2.26	-1.88
darkorange	33	0.33	1.08	-0.41	0.93	1.18	1.24	0.97	0.22	0.21	-0.41	-1.15	10.70	1.55	0.37	0.38
darkgrey	34	0.31	-0.47	1.09	0.08	-1.70	-0.92	-0.03	0.36	0.41	1.09	-1.29	6.02	1.09	0.36	0.90
lightsteelblue1	27	0.26	0.76	-0.23	0.65	0.80	0.72	0.93	1.79	1.66	0.03	-1.36	6.69	-0.23	-1.91	-1.64

module name	size	Zsummary.pres	Zdensity.pres	Zconnectivity.pres	Z.propVarExplained.pres	Z.meanSignAwareKME.pres	Z.meanSignAwareCorDat.pres	Z.meanAdj.pres	Z.meanClusterCoeff.pres	Z.meanMAR.pres	Z.cor.kIM	Z.cor.kME	Z.cor.kMEall	Z.cor.cor	Z.cor.clusterCoeff	Z.cor.MAR
ivory	25	0.15	0.55	-0.25	1.06	0.17	-0.40	0.92	1.06	1.02	-0.25	-0.07	15.79	-0.30	-0.49	-0.49
sienna3	28	0.14	-0.65	0.92	-3.47	0.56	1.20	-1.86	-2.37	-2.22	0.90	0.92	14.00	2.33	1.53	1.32
skyblue	32	0.10	0.61	-0.41	1.25	0.01	0.53	0.69	0.67	0.83	-1.44	-0.41	16.01	0.63	-2.79	-2.40
floralwhite	24	0.09	-0.85	1.03	-1.43	-0.20	-0.27	-1.50	-1.82	-1.66	1.04	0.47	7.83	1.03	-0.17	0.21
greenyellow	62	0.04	-0.18	0.25	-1.31	0.46	1.06	-0.82	-0.58	-0.46	-0.18	0.25	14.84	0.88	0.43	-0.98
magenta	64	0.00	0.83	-0.84	-1.12	1.54	3.88	0.11	-0.95	-0.21	-0.84	-0.86	14.15	2.43	-1.94	-0.86
paleturquoise	31	-0.04	-0.44	0.36	-0.63	0.12	-0.31	-0.57	-0.83	-0.80	0.59	-1.53	10.33	0.36	0.45	1.21
lightyellow	44	-0.04	-0.48	0.41	0.39	-0.91	-0.54	-0.43	-0.35	-0.07	-1.35	1.44	3.40	0.41	-1.95	-1.43
bisque4	23	-0.05	-0.85	0.75	-0.87	-1.57	0.66	-0.83	-0.99	-0.76	1.03	-1.09	7.53	0.75	1.10	1.34
yellowgreen	28	-0.07	0.74	-0.87	-0.61	1.62	1.76	-0.14	-0.43	-0.56	-0.21	-0.87	14.27	-1.25	0.67	0.24
gold	100	-0.08	-0.45	0.30	0.28	-0.71	-0.42	-0.49	-0.61	-0.61	0.49	0.30	15.40	0.17	-0.08	0.04
darkolivegreen	30	-0.11	-1.05	0.83	-1.64	-0.85	-0.15	-1.25	-0.92	-0.93	-0.12	2.27	3.73	0.83	0.14	-0.40
mediumpurple3	27	-0.17	-0.06	-0.27	-0.24	0.13	1.10	-0.44	-0.64	-0.60	-0.28	-0.27	5.25	1.22	1.26	0.35
orangered4	27	-0.18	0.04	-0.39	-0.27	0.45	0.35	-0.64	-1.09	-0.75	-0.39	0.42	8.68	-0.40	-1.70	-1.09
saddlebrown	32	-0.20	-1.34	0.95	-1.94	-1.00	-0.79	-1.68	-1.76	-1.82	0.44	1.15	4.49	0.95	0.36	0.80
pink	94	-0.24	0.51	-0.99	0.52	0.50	0.79	0.26	-0.26	-0.16	-1.58	-0.99	14.38	2.13	-0.59	-1.83
darkturquoise	36	-0.30	0.59	-1.19	1.17	0.01	-0.29	1.30	0.75	0.94	-2.20	0.88	9.38	-1.19	-0.47	-1.68
darkred	40	-0.33	-0.07	-0.59	-0.10	-0.56	0.12	-0.03	0.36	0.59	-1.03	-0.59	9.58	1.73	0.15	-0.59
darkorange2	24	-0.38	-0.70	-0.06	-3.06	0.82	0.61	-2.02	-2.64	-2.48	-0.06	-0.97	9.03	0.79	-0.23	-0.48
royalblue	42	-0.39	-0.28	-0.50	1.12	-0.58	-0.52	-0.04	-0.69	-0.64	-0.81	-0.50	-0.15	0.46	-0.68	-1.53
brown4	24	-0.51	-0.44	-0.59	-1.30	0.21	0.45	-1.09	-1.38	-1.34	-0.59	-1.02	6.05	0.37	-0.10	-0.61
orange	33	-0.57	-1.30	0.16	-3.64	-0.47	-0.26	-2.13	-2.09	-2.06	0.16	-0.91	9.39	0.23	-1.04	-0.48
violet	30	-0.76	-1.47	-0.05	-3.04	-0.92	-0.51	-2.01	-3.04	-2.79	0.06	-0.05	8.95	-0.71	0.24	-0.19

module name	size	Zsummary.pres	Zdensity.pres	Zconnectivity.pres	Z.propVarExplained.pres	Z.meanSignAwareKME.pres	Z.meanSignAwareCorDat.pres	Z.meanAdj.pres	Z.meanClusterCoeff.pres	Z.meanMAR.pres	Z.cor.kIM	Z.cor.kME	Z.cor.kMEall	Z.cor.cor	Z.cor.clusterCoeff	Z.cor.MAR
purple	64	-0.84	-0.94	-0.75	-0.29	-1.24	-0.80	-1.08	-1.10	-1.44	-1.53	-0.75	13.78	1.03	-0.97	-1.31
lightcyan	47	-0.91	-1.33	-0.48	-5.28	0.17	0.73	-2.84	-2.38	-1.82	-0.48	-1.52	1.74	1.56	-1.52	-2.21
darkslateblue	22	-0.97	-1.13	-0.82	-1.88	-0.73	-0.28	-1.53	-1.53	-1.56	-1.16	-0.82	6.12	0.28	-0.31	-1.14
black	118	-1.19	-1.78	-0.61	-3.31	-1.31	-1.06	-2.24	-2.14	-1.60	-2.47	0.29	17.35	-0.61	-1.18	-2.24
brown	177	-1.26	-4.38	1.85	-9.80	-2.86	3.37	-5.89	-7.25	-5.48	1.85	-0.36	9.99	4.99	-0.29	0.77

Table S6: Preservation statistics of chicken modules in mouse sorted by overall preservation.

module name	size	Zsummary.pres	Zdensity.pres	Zconnectivity.pres	Z.propVarExplained.pres	Z.meanSignAwareKME.pres	Z.meanSignAwareCorDat.pres	Z.meanAdj.pres	Z.meanClusterCoeff.pres	Z.meanMAR.pres	Z.cor.kIM	Z.cor.kME	Z.cor.kMEall	Z.cor.cor	Z.cor.clusterCoeff	Z.cor.MAR
green	107	3.33	4.56	2.10	3.45	3.61	5.51	14.08	12.26	9.75	2.10	1.46	8.29	2.64	1.37	2.62
greenyellow	54	3.33	3.86	2.80	2.84	3.42	4.30	8.74	8.34	7.71	2.00	2.80	16.57	3.83	-0.02	2.25
cyan	33	3.01	5.02	1.01	3.72	3.70	6.31	14.79	7.92	6.67	1.01	-0.04	7.89	1.03	0.45	0.95
tan	38	2.06	3.32	0.80	1.47	2.70	3.93	4.30	3.20	2.58	0.80	0.55	5.82	1.23	1.23	0.48
red	89	1.22	2.58	-0.13	1.74	2.36	2.80	5.22	3.95	4.35	-0.13	-0.20	2.86	0.73	0.64	1.24
turquoise	400	1.13	0.93	1.33	0.34	1.38	1.63	0.49	-1.14	-1.85	1.33	-0.72	7.62	2.66	-1.60	-0.50
salmon	37	0.88	2.27	-0.51	1.43	2.21	2.41	2.33	0.34	0.46	0.83	-1.25	8.88	-0.51	1.54	0.66
purple	57	0.60	0.19	1.01	-0.44	0.29	1.44	0.09	-0.50	-0.30	1.01	0.70	3.24	1.48	0.62	0.62
black	74	0.41	0.27	0.55	0.97	-0.95	-0.01	0.56	0.91	1.14	2.44	-0.99	0.82	0.55	2.24	1.58
yellow	181	0.32	0.30	0.34	1.17	-1.42	-0.57	2.15	2.90	2.83	1.04	-0.39	4.78	0.34	-0.42	0.60
pink	65	0.06	0.83	-0.72	-2.30	1.43	1.43	0.24	0.91	0.82	-0.72	-1.09	5.51	0.20	-0.70	-0.93
midnightblue	31	-0.09	0.19	-0.38	-0.03	0.33	0.53	0.05	0.26	0.23	-0.38	-1.21	0.53	0.58	0.76	-0.51
gold	100	-0.12	-0.72	0.49	-2.21	0.34	0.54	-1.78	-1.97	-1.81	0.49	0.99	7.93	0.47	-0.33	0.09
blue	400	-0.38	-2.39	1.63	-1.39	-6.80	-2.52	-2.26	-0.98	-0.88	-0.63	1.63	10.15	1.68	0.25	-1.38
magenta	57	-0.44	-0.15	-0.72	-0.04	-0.27	-0.76	0.45	1.43	1.48	-1.37	-0.72	8.66	0.21	-0.17	-1.20
brown	322	-0.71	-1.33	-0.10	2.55	-3.48	-3.23	0.57	0.78	0.22	-0.10	-0.99	8.82	0.84	-1.09	-0.73
grey	400	-1.12	-2.56	0.31	-2.51	-3.32	0.15	-2.61	-1.81	-1.27	-1.74	2.18	10.05	0.31	-1.70	-1.78

Table S7: Top ten correlated genes with each chicken module eigengene discussed, the median quantile-normalized FPKM for that gene, and the fold difference from the median expression in each chicken sample. DLC, dorsolateral cortex; HP, hyperpallium; MP, mesopallium; AP, arcopallium; NP, nidopallium; HIPP, hippocampus; STR, striatum.

green	107 genes		fold difference from median						
Gene symbol	Gene ID	median normalized FPKM	DLC	HP	MP	AP	NP	HIPP	STR
Kcna4	ENSMUSG00000042604	93	0.86	1.20	0.99	1.01	1.00	0.81	2.29
Smpd3	ENSMUSG00000031906	227	1.00	0.81	0.96	0.82	1.13	1.08	2.06
Foxo1	ENSMUSG00000044167	44	0.94	1.12	1.00	0.92	1.14	0.57	3.57
Pde10a	ENSMUSG00000023868	48	0.74	1.75	0.86	2.24	1.00	0.52	10.16
Drd1a	ENSMUSG00000021478	147	1.62	0.85	0.92	1.00	1.22	0.03	7.09
Tmem90a	ENSMUSG00000071234	89	1.51	0.56	2.20	0.79	1.00	0.41	4.67
Tac1	ENSMUSG00000061762	248	1.47	1.42	0.00	1.00	0.45	0.00	9.03
Bcr	ENSMUSG00000009681	106	0.91	0.81	1.03	0.82	1.00	1.14	2.09
Bcl11b	ENSMUSG00000048251	69	2.23	0.67	1.62	0.46	0.75	1.00	4.80
Adcy5	ENSMUSG00000022840	93	1.15	0.69	0.97	1.62	0.59	1.00	6.47

greenyellow	54 genes		fold difference from median						
Gene symbol	Gene ID	median normalized FPKM	DLC	HP	MP	AP	NP	HIPP	STR
Kdm3a	ENSMUSG00000053470	76	0.69	1.54	0.85	1.37	1.00	0.63	1.28
Mctp1	ENSMUSG00000021596	200	0.79	1.16	1.00	1.12	0.49	0.39	2.01

Serinc5	ENSMUSG00000021703	44	0.90	1.97	0.66	1.61	1.00	0.73	2.07
Mapk1	ENSMUSG00000063358	1408	1.00	0.59	1.12	0.72	1.29	0.86	2.17
Pppde1	ENSMUSG00000026502	37	0.85	1.99	0.89	1.74	1.00	0.77	1.93
Rnf103	ENSMUSG00000052656	176	0.99	0.95	1.00	1.05	1.13	0.87	1.41
BC031353	ENSMUSG00000034858	143	0.63	1.21	0.71	1.16	1.00	0.74	1.36
Ptpn11	ENSMUSG00000043733	192	0.75	1.12	0.68	1.00	1.05	0.66	1.26
Fam105a	ENSMUSG00000056069	151	0.80	1.30	0.65	2.84	0.93	1.00	4.52
Ano3	ENSMUSG00000074968	91	0.26	1.79	0.77	2.29	1.00	0.21	2.95

magenta	57 genes		fold difference from median						
Gene symbol	Gene ID	median normalized FPKM	DLC	HP	MP	AP	NP	HIPP	STR
Slco3a1	ENSMUSG00000025790	63	1.00	0.91	0.73	0.97	1.08	2.38	1.13
Slc16a1	ENSMUSG00000032902	41	1.03	0.80	1.43	0.82	0.81	2.03	1.00
Gnai3	ENSMUSG00000000001	835	1.04	0.40	1.04	0.48	0.89	1.38	1.00
Mtmr9	ENSMUSG00000035078	351	1.03	0.49	1.00	0.67	0.64	1.37	1.07
Agfg1	ENSMUSG00000026159	245	1.00	0.51	1.03	0.58	0.70	1.31	1.00
Uso1	ENSMUSG00000029407	191	1.19	0.68	1.18	0.73	0.90	1.63	1.00
Wasl	ENSMUSG00000029684	82	1.34	0.64	1.00	0.53	0.55	3.22	1.75
Kras	ENSMUSG00000030265	168	0.96	0.67	1.34	0.59	1.00	2.04	1.45
Ndst1	ENSMUSG00000054008	182	1.27	0.30	1.10	0.34	0.55	1.74	1.00
Rac1	ENSMUSG00000001847	1751	1.30	0.55	1.34	0.62	0.84	1.94	1.00

black	74 genes		fold difference from median						
Gene symbol	Gene ID	median normalized FPKM	DLC	HP	MP	AP	NP	HIPP	STR
Dgke	ENSMUSG00000000276	104	1.00	0.75	1.30	0.67	1.34	0.96	1.33
Crk	ENSMUSG00000017776	187	1.00	0.85	0.98	0.89	1.22	1.00	1.00
Klhl24	ENSMUSG00000062901	67	1.00	1.06	0.95	0.91	1.13	1.07	0.87
Dnajb14	ENSMUSG00000074212	653	0.98	0.61	1.06	0.73	1.19	1.12	1.00
1810013L24Rik	ENSMUSG00000022507	341	1.02	0.68	1.03	0.77	1.29	1.00	0.97
Klhl7	ENSMUSG00000028986	297	1.02	1.11	1.00	0.95	1.62	0.74	0.86
Rnf11	ENSMUSG00000028557	394	1.06	0.71	1.02	1.00	1.14	0.85	0.95
Chmp7	ENSMUSG00000034190	480	1.00	0.58	1.12	0.68	1.14	0.98	1.15
Fam172a	ENSMUSG00000064138	86	1.21	0.90	0.95	1.00	1.17	0.99	1.05
Klhl11	ENSMUSG00000048732	146	0.90	0.55	1.19	0.80	1.22	1.00	1.10

cyan	33 genes		fold difference from median						
Gene symbol	Gene ID	median normalized FPKM	DLC	HP	MP	AP	NP	HIPP	STR
Bcas1	ENSMUSG00000013523	229	0.59	1.32	0.67	2.70	1.42	1.00	0.61
Spg20	ENSMUSG00000036580	113	0.61	1.68	0.79	2.31	1.61	1.00	0.94
Pllp	ENSMUSG00000031775	180	0.70	0.51	0.85	2.34	1.70	1.18	1.00
Gab1	ENSMUSG00000031714	108	0.67	1.09	1.00	2.60	1.68	0.94	0.98
Anln	ENSMUSG00000036777	34	0.02	2.64	0.57	5.30	2.75	1.00	0.81
Elovl7	ENSMUSG00000021696	27	0.03	3.61	1.00	4.72	2.05	0.85	0.84
Slain1	ENSMUSG00000055717	113	0.90	1.03	0.82	1.13	1.01	1.00	0.90

Tspan2	ENSMUSG00000027858	153	0.70	1.02	0.73	2.46	1.54	0.75	1.00
Reep3	ENSMUSG00000019873	117	0.64	1.59	0.82	2.43	1.48	0.72	1.00
Mbp	ENSMUSG00000041607	8687	0.50	1.00	0.71	3.19	1.94	0.77	1.00

turquoise	2452 genes		fold difference from median						
Gene symbol	Gene ID	median normalized FPKM	DLC	HP	MP	AP	NP	HIPP	STR
Gosr1	ENSMUSG00000010392	119	0.96	1.96	0.92	1.73	1.16	0.90	1.00
Ubqln1	ENSMUSG00000005312	292	0.89	1.55	1.00	1.40	1.25	0.91	0.92
March7	ENSMUSG00000026977	107	1.00	2.22	0.94	1.91	1.33	0.87	1.00
Necap1	ENSMUSG00000030327	949	0.77	1.19	0.88	1.11	1.02	0.78	1.00
Ap1g1	ENSMUSG00000031731	344	0.91	1.43	0.95	1.27	1.09	0.81	1.00
Tmx3	ENSMUSG00000024614	184	0.84	2.61	0.92	2.26	1.40	0.78	1.00
AI597479	ENSMUSG00000010290	60	0.99	3.26	0.94	2.63	1.72	1.00	0.87
Eif4e	ENSMUSG00000028156	420	1.00	1.86	0.95	1.68	1.26	0.91	0.93
Dnajc3	ENSMUSG00000022136	233	0.94	2.15	1.00	1.93	1.53	0.94	1.00
Heatr5b	ENSMUSG00000039414	169	0.90	2.03	0.89	1.77	1.39	1.00	0.92

brown	322 genes		fold difference from median						
Gene symbol	Gene ID	median normalized FPKM	DLC	HP	MP	AP	NP	HIPP	STR
Rab28	ENSMUSG00000029128	285	1.14	0.54	1.10	0.60	0.81	1.00	1.17
Myo9a	ENSMUSG00000039585	167	1.00	0.22	1.08	0.22	0.45	1.13	1.46

Ccny	ENSMUSG00000024286	206	1.11	0.48	1.06	0.43	0.72	1.00	1.11
Kpna1	ENSMUSG00000022905	174	1.17	0.46	1.00	0.55	0.63	1.06	1.06
Rprd2	ENSMUSG00000028106	173	1.16	0.36	1.07	0.33	0.57	1.21	1.00
Cpsf6	ENSMUSG00000055531	111	1.21	0.54	1.00	0.48	0.66	1.13	1.00
Pogz	ENSMUSG00000038902	74	1.03	0.05	1.03	0.39	0.37	1.00	1.43
Ythdf2	ENSMUSG00000040025	101	1.18	0.21	1.02	0.26	0.37	1.23	1.00
Sgip1	ENSMUSG00000028524	235	1.19	0.19	1.37	0.16	0.42	1.41	1.00
Kpnb1	ENSMUSG00000001440	228	1.13	0.38	1.09	0.40	0.63	1.01	1.00

red	89 genes		fold difference from median						
Gene symbol	Gene ID	median normalized FPKM	DLC	HP	MP	AP	NP	HIPP	STR
Cul4b	ENSMUSG00000031095	117	0.87	1.00	1.07	0.95	1.14	0.84	1.25
Samd8	ENSMUSG00000021770	163	0.72	1.00	1.25	0.91	1.52	0.84	1.37
Vcpip1	ENSMUSG00000045210	83	0.68	1.04	1.00	1.00	1.02	0.86	1.00
Atp13a3	ENSMUSG00000022533	72	0.73	1.08	0.85	1.00	1.02	0.61	1.16
Rprd1a	ENSMUSG00000040446	145	1.00	0.83	1.39	0.81	1.07	0.93	1.43
Lmbrd2	ENSMUSG00000039704	97	0.82	1.00	1.04	0.85	1.02	0.70	1.21
B230219D22Rik	ENSMUSG00000045767	455	1.00	0.60	1.19	0.61	1.27	0.82	1.50
Gopc	ENSMUSG00000019861	79	0.66	1.00	1.24	0.80	1.13	0.67	1.00
Ulk2	ENSMUSG00000004798	318	1.00	0.91	1.24	0.83	1.20	0.50	1.16
Cul5	ENSMUSG00000032030	125	0.66	1.00	1.04	0.92	1.09	0.79	1.05

Table S8: Top ten correlated genes with each mouse module eigengene discussed, the median quantile-normalized FPKM for that gene, and the fold difference from the median expression in each mouse sample. DC, dorsal cortex; LC, lateral cortex; HIPP, hippocampus; C/E, claustrum/endopiriform complex; PA, pallial amygdala; STR, striatum.

brown	177 genes		fold difference from median															
Gene symbol	Gene ID	median normalized FPKM	DC A	DC B	DC C	DC D	DC E	DC F	LC A	LC B	LC C	LC D	LC E	LC F	HIPP	C/E	PA	STR
Me2	ENSMUSG00000024556	16	0.83	0.86	0.91	1.14	0.85	1.04	0.82	0.75	0.67	0.96	1.30	1.24	1.24	1.30	1.30	3.64
4922501L14Rik	ENSMUSG00000042943	12	0.92	0.72	1.01	0.72	1.08	0.99	1.52	0.85	0.59	1.22	1.32	0.94	0.86	1.89	2.85	3.79
Pik3c2b	ENSMUSG00000026447	9	1.00	0.76	0.72	0.79	1.12	1.62	1.09	1.00	0.83	1.00	1.08	1.51	1.25	0.13	0.81	3.66
Adcy3	ENSMUSG00000020654	43	0.98	0.76	0.87	0.96	0.85	1.01	1.31	0.88	1.12	0.99	1.06	1.16	0.93	1.23	1.37	1.62
Mgea5	ENSMUSG00000025220	180	0.87	1.09	1.09	0.99	0.99	0.97	0.70	0.89	0.92	1.01	1.03	1.06	0.96	1.10	1.25	1.55
Wscd2	ENSMUSG00000063430	49	1.70	1.02	0.98	0.83	0.86	0.83	1.67	1.30	1.12	0.89	0.97	0.90	1.98	0.96	1.15	3.06
Cacna2d2	ENSMUSG00000010066	39	0.51	0.69	1.42	1.45	1.11	1.02	0.55	0.98	1.24	1.31	0.74	0.64	0.42	0.81	1.50	2.69
Tnrc18	ENSMUSG00000039477	21	0.81	0.99	0.97	0.92	0.95	1.00	1.46	1.52	0.91	1.20	1.09	1.00	1.21	0.89	1.13	1.84
Sik3	ENSMUSG00000034135	113	1.01	1.06	0.98	0.99	0.95	0.98	1.13	1.07	0.99	0.92	0.85	0.91	1.49	1.01	1.12	1.99
Phf15	ENSMUSG00000020387	50	0.80	0.83	1.12	1.09	1.05	1.04	0.81	0.92	1.03	0.97	0.90	0.85	1.12	0.94	1.19	1.69

magenta	64 genes		fold difference from median															
Gene symbol	Gene ID	median normalized FPKM	DC A	DC B	DC C	DC D	DC E	DC F	LC A	LC B	LC C	LC D	LC E	LC F	HIPP	C/E	PA	STR

Lamp1	ENSMUSG00000031447	292	0.76	0.87	1.01	0.94	0.97	1.47	0.89	1.18	1.13	1.04	0.99	1.18	1.09	0.92	0.90	1.16
Rdx	ENSMUSG00000032050	59	0.91	0.82	0.80	0.96	0.94	1.42	0.80	0.85	1.00	1.00	1.09	1.31	1.29	1.04	1.15	1.38
Npc1	ENSMUSG00000024413	39	1.01	0.89	0.86	1.07	1.19	1.94	0.83	0.93	1.02	0.84	0.99	1.37	1.26	0.94	0.61	1.41
Dram2	ENSMUSG00000027900	23	0.69	0.81	0.73	0.80	1.22	2.75	0.96	0.75	0.71	1.17	1.07	1.04	1.46	1.21	0.84	1.84
Ndrg1	ENSMUSG00000005125	155	0.42	0.36	0.84	1.22	1.26	2.48	0.37	0.47	0.68	1.02	0.98	1.51	1.40	0.72	1.04	1.47
Rnf13	ENSMUSG00000036503	101	0.52	0.75	0.98	1.02	1.04	1.78	0.44	0.76	0.94	1.07	1.02	1.26	1.30	0.82	0.78	1.47
Cln8	ENSMUSG00000026317	23	0.79	0.81	1.00	1.36	1.14	1.64	0.79	0.93	1.00	1.00	1.02	1.04	1.08	0.72	0.77	1.48
Elovl7	ENSMUSG00000021696	11	0.74	0.73	0.94	1.20	1.82	2.98	0.69	0.71	0.74	1.06	1.30	2.07	1.61	0.57	0.53	2.51
Bmpr1a	ENSMUSG00000021796	39	0.60	0.65	0.86	1.01	1.01	1.37	0.44	0.60	0.66	1.04	0.97	1.11	1.37	0.99	1.13	1.31
Itgb1	ENSMUSG00000025809	43	0.85	0.82	1.02	1.06	1.05	1.41	0.67	0.84	0.98	0.97	0.97	1.13	1.08	0.96	1.02	1.26

tan	62 genes		fold difference from median															
Gene symbol	Gene ID	median normalized FPKM	DC A	DC B	DC C	DC D	DC E	DC F	LC A	LC B	LC C	LC D	LC E	LC F	HIPP	C/E	PA	STR
Pdcd6	ENSMUSG00000021576	54	1.08	1.27	0.99	0.85	0.69	0.65	1.45	1.56	1.18	1.07	0.95	0.65	1.00	0.75	1.00	1.76
Atox1	ENSMUSG00000018585	543	1.30	1.02	0.98	1.13	0.83	0.32	1.61	2.22	0.77	1.44	0.77	0.46	1.32	0.96	0.89	1.17
Srp19	ENSMUSG00000014504	105	1.16	1.09	0.70	0.74	0.65	1.00	1.59	1.19	1.07	1.00	0.83	0.72	1.09	0.84	1.14	0.79
H47	ENSMUSG00000075701	74	1.01	1.25	1.02	0.69	0.61	0.60	1.42	1.52	1.16	0.89	0.78	0.75	1.04	0.62	0.99	1.31
Acsbg1	ENSMUSG00000032281	158	1.24	1.06	1.02	0.87	0.66	0.64	1.36	1.48	1.18	0.98	0.68	0.58	1.03	0.58	0.78	1.03
Rpl22l1	ENSMUSG00000039221	87	0.57	1.17	0.49	0.65	0.50	0.47	2.28	2.71	0.82	2.06	1.26	0.59	1.40	0.83	1.40	2.76
Ndufb10	ENSMUSG00000040048	409	1.20	0.93	0.79	0.76	0.70	0.63	1.50	1.36	1.15	1.07	1.10	0.89	1.12	0.82	0.88	1.35
Serf1	ENSMUSG00000021643	21	1.43	1.17	0.73	0.60	0.64	0.42	1.57	1.85	0.95	1.05	1.23	0.71	1.14	0.94	1.19	0.83
4632415K11Rik	ENSMUSG00000034105	16	1.46	1.02	0.96	0.97	0.87	1.01	2.03	1.09	1.23	0.99	0.96	0.98	1.04	0.61	0.81	1.48
Park7	ENSMUSG00000028964	277	1.30	1.18	1.00	0.84	0.89	0.61	1.70	1.42	1.01	0.99	0.99	0.71	1.06	0.90	1.00	1.05

black	118 genes		fold difference from median														
Gene symbol	Gene ID	median normalized FPKM	DC A	DC B	DC C	DC D	DC E	DC F	LC A	LC B	LC C	LC D	LC E	LC F	HIPP	C/E	PA	STR
Tars	ENSMUSG00000022241	22	0.95	1.00	0.99	1.00	0.96	0.95	1.15	1.13	1.14	1.08	1.06	0.95	1.60	0.98	0.91	1.13
Rnf2	ENSMUSG00000026484	29	0.93	1.20	0.88	0.85	1.00	1.06	0.84	0.98	0.91	1.00	0.98	1.05	2.23	1.16	1.17	1.09
Rwdd1	ENSMUSG00000019782	23	0.99	0.72	0.81	1.15	1.17	1.04	1.01	0.58	0.98	0.80	0.94	1.17	1.94	1.05	1.54	0.70
Ankrd46	ENSMUSG00000048307	363	0.92	1.11	1.03	0.90	0.93	1.00	0.75	0.90	0.92	0.94	1.00	1.05	1.21	1.14	1.10	1.03
Copb2	ENSMUSG00000032458	67	0.87	1.01	0.89	1.11	1.06	0.89	0.84	0.99	0.87	1.01	1.03	0.95	1.62	1.14	1.16	0.85
Acadl	ENSMUSG00000026003	23	0.89	1.01	0.87	1.10	0.94	0.88	0.90	1.17	0.95	1.09	0.79	1.00	1.91	1.00	1.50	1.48
Etfdh	ENSMUSG00000027809	19	0.99	1.11	0.85	1.19	1.01	0.87	0.78	0.99	1.11	1.16	0.96	1.03	1.62	0.96	0.92	1.31
Smu1	ENSMUSG00000028409	39	0.89	0.87	0.77	0.94	0.91	0.88	1.02	0.96	0.98	1.11	1.25	1.12	1.60	1.07	1.22	1.39
Slc1a3	ENSMUSG00000005360	368	1.30	1.10	1.01	0.95	0.93	0.98	0.90	0.99	1.04	0.95	0.79	0.85	1.78	1.03	1.47	1.33
Plrg1	ENSMUSG00000027998	25	0.84	0.98	0.83	1.18	0.89	1.18	1.01	1.09	0.82	0.97	0.99	0.99	1.72	1.15	1.36	1.22

orange	33 genes		fold difference from median														
Gene symbol	Gene ID	median normalized FPKM	DC A	DC B	DC C	DC D	DC E	DC F	LC A	LC B	LC C	LC D	LC E	LC F	HIPP	C/E	PA	STR
ORF61	ENSMUSG00000013858	208	1.03	1.44	1.15	1.03	1.07	0.88	1.44	1.21	1.24	0.89	0.94	0.97	0.81	0.95	0.77	0.86
Ggt7	ENSMUSG00000027603	148	1.29	1.17	1.15	0.87	0.82	0.87	1.37	1.37	1.20	0.99	1.05	1.01	0.73	0.72	0.80	0.61
Atg13	ENSMUSG00000027244	98	1.04	1.08	1.16	0.97	1.00	1.00	1.05	1.15	1.14	1.01	0.97	0.96	0.77	0.79	0.74	0.87
Fbxo34	ENSMUSG00000037536	103	1.15	1.48	1.33	1.03	0.80	0.67	1.25	1.53	1.37	0.97	0.80	0.68	1.45	0.73	0.90	0.53
Zdhhc16	ENSMUSG00000025157	30	1.12	1.55	1.01	1.04	0.79	0.83	1.59	1.40	1.08	1.04	0.99	0.95	0.85	0.79	0.98	0.96
Spsb3	ENSMUSG00000024160	234	1.33	1.55	1.27	1.04	0.98	1.02	1.45	1.36	1.42	0.98	0.90	0.89	0.74	0.82	0.61	0.95
Rapgef1	ENSMUSG00000039844	90	1.48	1.27	1.19	0.99	0.99	0.88	1.44	1.39	1.25	1.01	1.03	0.92	0.81	0.99	0.87	0.94
Clip2	ENSMUSG00000063146	134	1.31	1.10	1.21	1.00	0.91	1.00	1.44	1.60	1.36	1.08	0.90	0.98	0.78	0.88	0.69	0.52

Klhl18	ENSMUSG00000054792	33	1.06	1.02	1.18	1.04	0.93	0.87	1.16	1.30	1.18	1.01	0.90	0.82	0.76	0.99	0.86	0.98
Dnajb5	ENSMUSG00000036052	308	1.04	1.04	1.20	0.98	1.01	0.91	1.28	1.32	1.27	0.99	0.96	0.93	1.03	0.81	0.65	0.64

salmon	61 genes		fold difference from median															
Gene symbol	Gene ID	median normalized FPKM	DC A	DC B	DC C	DC D	DC E	DC F	LC A	LC B	LC C	LC D	LC E	LC F	HIPP	C/E	PA	STR
Mrps33	ENSMUSG00000029918	35	0.86	1.45	0.95	0.87	0.72	0.67	1.05	1.23	1.28	0.86	1.23	0.95	1.44	0.85	1.28	1.57
Pdcd6	ENSMUSG00000021576	54	1.08	1.27	0.99	0.85	0.69	0.65	1.45	1.56	1.18	1.07	0.95	0.65	1.00	0.75	1.00	1.76
Cisd1	ENSMUSG00000037710	177	0.97	1.00	0.91	0.98	0.80	0.85	1.08	1.26	1.22	1.22	1.00	0.95	1.10	0.91	1.12	1.25
2310003C23Rik	ENSMUSG00000027573	71	0.84	0.97	0.77	0.78	0.73	0.69	1.05	1.21	1.05	1.12	1.01	0.91	0.99	1.03	1.03	1.17
Atp5f1	ENSMUSG00000000563	206	0.55	1.03	0.90	1.03	0.74	0.66	0.87	1.37	1.10	1.37	1.01	0.78	1.21	0.79	0.99	1.55
Rpl22l1	ENSMUSG00000039221	87	0.57	1.17	0.49	0.65	0.50	0.47	2.28	2.71	0.82	2.06	1.26	0.59	1.40	0.83	1.40	2.76
Cops4	ENSMUSG00000035297	79	0.91	1.07	0.85	0.90	0.76	0.74	1.10	1.24	1.08	1.00	1.00	0.85	1.32	0.98	1.14	1.01
Uqcrb	ENSMUSG00000021520	69	0.74	1.13	0.80	0.89	0.68	0.53	0.79	1.21	1.11	1.06	0.94	0.69	1.47	1.12	1.16	1.35
Atp5j	ENSMUSG00000022890	320	0.68	1.04	0.80	0.91	0.50	0.65	1.00	0.87	1.43	1.11	1.09	0.64	1.82	1.00	1.23	1.78
Ndufa5	ENSMUSG00000023089	391	0.68	0.87	0.70	0.64	0.58	0.35	1.01	1.53	0.99	2.03	1.42	0.76	1.68	1.02	1.54	2.33

Table S9: Estimated pre-mRNA fraction in mouse and chick samples.

Mouse sample	Median pre-mRNA fraction	Chicken sample	Median pre-mRNA fraction
Dorsal cortex A (layers I-III)	4.49% (N=2972)	arcopallium	6.95% (N=7237)
Dorsal cortex B (layer IV)	4.20% (N=3150)	dorsolateral cortex	8.16% (N=7608)
Dorsal cortex C (layers IV-V)	3.50% (N=3102)	hippocampus	8.19% (N=7761)
Dorsal cortex D (layer V)	3.39% (N=3148)	mesopallium	8.05% (N=7048)
Dorsal cortex E (layer VI)	2.92% (N=2705)	nidopallium	7.07% (N=7095)
Dorsal cortex F (layers VI-VIb)	2.55% (N=3150)	hyperpallium	7.31% (N=5911)
Lateral cortex A (layers I-III)	4.86% (N=2943)	striatum	9.24% (N=7144)
Lateral cortex B (layer IV)	4.11% (N=3093)		
Lateral cortex C (layers IV-V)	3.37% (N=3097)		
Lateral cortex D (layers V-VI)	3.22% (N=3023)		
Lateral cortex E (layer VI)	2.63% (N=2975)		
Lateral cortex F (layer VIb)	2.55% (N=3066)		
Hippocampus	2.35% (N=2980)		
Claustrum/endopiriform	3.13% (N=3119)		
Pallial amygdala	3.29% (N=3136)		
Striatum	2.31% (N=2993)		

Table S10: Downsampling read fragments.

Species	Region	# unique fragments mapped (% fragments sequenced)	# (%) unique mapped fragments in pc genes >200nt	Ratio to downsample	# unique fragments sampled
Chicken	Arcopallium	43,356,930 (93%)	29,563,619 (68%)	0.488717	21,189,264
	Dorsolateral cortex	37,306,331 (85%)	21,573,838 (58%)	0.537091	20,036,913
	Hippocampus	38,116,880 (85%)	22,922,907 (60%)	0.509703	19,428,271
	Mesopallium	25,280,073 (86%)	15,302,760 (61%)	0.724450	18,314,158
	Nidopallium	31,066,189 (90%)	19,310,314 (62%)	0.653362	20,297,472
	Hyperpallium	26,299,836 (94%)	17,071,439 (65%)	N/A	26,299,836
	Striatum	26,246,353 (88%)	15,125,692 (58%)	0.752200	19,742,508
Mouse	Dorsal cortex A	43,539,003 (96%)	28,760,754 (66%)	0.749822	32,646,495
	Dorsal cortex B	59,885,178 (94%)	41,072,917 (69%)	0.508724	30,464,999
	Dorsal cortex C	54,582,567 (95%)	41,863,139 (77%)	0.502602	27,433,302
	Dorsal cortex D	60,080,966 (96%)	46,445,743 (77%)	0.464065	27,881,490
	Dorsal cortex E	27,363,620 (96%)	20,950,956 (77%)	N/A	27,363,620
	Dorsal cortex F	56,037,833 (97%)	44,312,232 (79%)	0.454334	25,459,896
	Lateral cortex A	43,682,519 (96%)	26,523,474 (61%)	0.775756	33,886,992
	Lateral cortex B	54,747,441 (96%)	36,283,334 (66%)	0.564627	30,911,895
	Lateral cortex C	55,058,857 (97%)	41,437,719 (75%)	0.506372	27,880,256
	Lateral cortex D	45,732,069 (97%)	34,588,590 (76%)	0.605109	27,672,889
	Lateral cortex E	42,428,722 (96%)	33,439,586 (79%)	0.613748	26,040,543
	Lateral cortex F	48,260,801 (98%)	37,984,095 (79%)	0.522717	25,226,725
	Hippocampus	49,280,437 (97%)	39,544,659 (80%)	0.582436	28,702,719
	Claustrum/ endopiriform	54,403,180 (97%)	41,710,672 (77%)	0.501311	27,272,906
	Pallial amygdala	52,006,334 (95%)	38,990,285 (75%)	0.519640	27,024,585
	Striatum	49,859,361 (96%)	39,701,019 (80%)	0.572858	28,562,311

Table S11: Variance explained by the first 15 components of a lane-wise principal component analysis.

principal component	mouse proportion of variance explained	mouse cumulative proportion of variance	chicken proportion of variance explained	chicken cumulative proportion of variance
1	32.8%	32.8%	65.8%	65.8%
2	15.9%	48.8%	10.7%	76.5%
3	11.2%	60.0%	6.3%	82.8%
4	9.2%	69.1%	5.5%	88.3%
5	6.5%	75.6%	5.1%	93.4%
6	5.3%	80.1%	3.1%	96.5%
7	2.7%	83.6%	0.7%	97.2%
8	2.5%	86.1%	0.4%	97.6%
9	1.9%	88.0%	0.4%	98.0%
10	1.8%	89.8%	0.3%	98.3%
11	1.5%	91.3%	0.3%	98.7%
12	1.1%	92.4%	0.3%	99.0%
13	1.0%	93.4%	0.3%	99.2%
14	0.8%	94.1%	0.3%	99.5%
15	0.7%	94.8%	0.3%	99.8%

Table S12: Variance explained by the first 15 components of a library-wise principal component analysis.

principal component	mouse proportion of variance explained	mouse cumulative proportion of variance	chicken proportion of variance explained	chicken cumulative proportion of variance
1	34.7%	34.7%	68.8%	68.8%
2	22.6%	57.3%	13.3%	82.1%
3	15.8%	73.0%	7.4%	89.5%
4	7.6%	80.6%	4.6%	94.1%
5	5.8%	86.4%	3.3%	97.4%
6	3.9%	90.3%	2.6%	100%
7	2.2%	92.5%	0%	100%
8	1.7%	94.2%		
9	1.5%	95.7%		
10	1.1%	96.7%		
11	0.8%	97.6%		
12	0.7%	98.3%		
13	0.7%	99.0%		
14	0.5%	99.5%		
15	0.5%	100%		

Table S13: Scale-free topology (SFT) approximation and connectivity (k) in chicken for different values of the soft threshold.

Power	SFT R^2	slope	mean k	median k	max k
1	0.632	1.53	3170	3560	3960
2	0.36	0.589	2360	2760	3380
3	0.153	0.27	1890	2230	3010
4	0.0311	0.102	1580	1850	2730
5	0.000157	-0.00626	1360	1560	2510
6	0.0308	-0.0814	1190	1330	2330
7	0.103	-0.144	1050	1140	2170
8	0.204	-0.197	941	993	2040
9	0.332	-0.243	850	868	1930
10	0.425	-0.279	774	764	1820
12	0.599	-0.343	652	600	1650
14	0.712	-0.393	560	479	1510
16	0.797	-0.437	488	386	1390
18	0.846	-0.473	430	318	1290
20	0.878	-0.505	383	264	1200
22	0.907	-0.533	344	222	1120
24	0.932	-0.562	311	189	1060
26	0.943	-0.587	284	160	998
28	0.953	-0.605	260	137	949
30	0.961	-0.63	239	119	905

Table S14: Scale-free topology (SFT) approximation and connectivity (k) in mouse for different values of the soft threshold.

Power	SFT R^2	slope	mean k	median k	max k
1	0.562	2.06	1860	1860	2710
2	0.0996	0.35	942	907	1790
3	0.159	-0.353	557	511	1300
4	0.511	-0.711	362	313	991
5	0.649	-0.941	249	203	786
6	0.721	-1.07	180	139	640
7	0.749	-1.19	134	97.8	530
8	0.781	-1.25	103	70.7	447
9	0.796	-1.3	80.3	52.1	381
10	0.805	-1.35	64.1	39.1	328
12	0.824	-1.42	42.5	22.8	249
14	0.835	-1.46	29.6	14.1	194
16	0.846	-1.48	21.4	9.1	155
18	0.856	-1.49	15.9	6.09	125
20	0.866	-1.49	12.1	4.2	102
22	0.874	-1.5	9.38	2.96	85
24	0.876	-1.49	7.42	2.14	71.1
26	0.885	-1.47	5.95	1.57	60
28	0.893	-1.46	4.85	1.18	51.1
30	0.919	-1.41	3.99	0.89	43.7

Table S15: Number of preserved modules (Zpres > 3) with different soft thresholds and deep split parameters

Mouse Power	Chicken Power	Species	First split	Second split	Third split	Fourth split	Fifth split
10	24	Mouse	0	1	0	2	0
		Chicken	1	1	1	1	1
10	10	Mouse	0	1	0	2	0
		Chicken	1	1	1	2	0
10	15	Mouse	0	1	0	2	0
		Chicken	1	1	2	1	2
15	24	Mouse	0	1	1	2	1
		Chicken	1	1	1	1	1
15	15	Mouse	0	1	1	2	1
		Chicken	1	1	2	1	2
10	20	Mouse	0	1	0	2	0
		Chicken	1	1	3	1	1

Table S16: Chicken modules significantly overlapping mouse modules

chicken module	mouse module	Bonferroni-corrected *p*-value
green	brown	8.7×10^{-21}
cyan	magenta	3.0×10^{-18}
greenyellow	brown	1.4×10^{-4}
turquoise	salmon	1.6×10^{-3}
magenta	black	3.6×10^{-3}
black	orange	4.6×10^{-3}

Table S17: Functional enrichments in mouse modules having a Bonferroni-corrected $P<0.05$. Modules correlated with a technical variable (two-tailed $P<0.05$) are indicated with an asterisk. Annotation sources are drawn from the databases described in Belgard *et al*.**(1)** – GO biological process (GO BP), GO cellular component (GO CC), GO molecular function (GO MF), and Kyoto Encyclopaedia of Genes and Genomes Pathways (KEGG).

module	single test *p*-value (one-tailed)	annotation (source)	$\log_2$ (fold diff-erence)	# genes with this annotation in the module	# genes with any annotation in the module
brown	$2.1x10^{-6}$	signal transduction (GO BP)	1.4	28	140
brown	$3.9x10^{-6}$	G-protein coupled receptor protein signaling pathway (GO BP)	1.7	21	140
brown	$1.5x10^{-6}$	G-protein coupled receptor activity (GO MF)	2.5	13	154
brown	$4.3x10^{-5}$	adrenoceptor activity (GO MF)	4.2	5	154
brown	$6.6x10^{-5}$	signal transducer activity (GO MF)	1.6	17	154
brown	$1.2x10^{-5}$	Neuroactive ligand-receptor interaction (KEGG)	2.1	13	64
magenta*	$3.5x10^{-6}$	myelination (GO BP)	4.7	5	52
magenta*	$1.6x10^{-6}$	myelin sheath (GO CC)	6.3	4	57
magenta*	$1.6x10^{-5}$	integral to membrane (GO CC)	0.94	33	57
magenta*	$3.7x10^{-5}$	compact myelin (GO CC)	6.3	3	57
magenta*	$3.5x10^{-5}$	structural constituent of myelin sheath (GO MF)	6.3	3	55
tan*	$9.1x10^{-6}$	ribonucleoprotein complex (GO CC)	2.8	9	54
tan*	$7.6x10^{-5}$	respiratory chain (GO CC)	3.6	5	54
tan*	$5.9x10^{-7}$	Huntington's disease (KEGG)	3.1	9	20
tan*	$1.4x10^{-4}$	Parkinson's disease (KEGG)	2.9	6	20
tan*	$1.5x10^{-4}$	Oxidative phosphorylation (KEGG)	2.9	6	20
tan*	$2.2x10^{-4}$	Alzheimer's disease (KEGG)	2.8	6	20
violet	$9.6x10^{-5}$	membrane (GO CC)	0.95	20	25

Table S18: Functional enrichments in chick modules having a Bonferroni-corrected $P<0.05$. Modules correlated with a batch variable (two-tailed $P<0.05$) are indicated with an asterisk. Annotation sources are drawn from the databases described in Belgard *et al.***(1)**– GO biological process (GO BP), GO cellular component (GO CC), GO molecular function (GO MF), and Kyoto Encyclopaedia of Genes and Genomes Pathways (KEGG).

module	single test *p*-value (one-tailed)	annotation (source)	$\log_2$ (fold diff-erence)	# genes with this annotation in the module	# genes with any annotation in the module
brown*	$7.5\text{x}10^{-5}$	actin cytoskeleton (GO CC)	2.0	13	272
brown*	$2.7\text{x}10^{-4}$	Spliceosome (KEGG)	1.8	12	91
cyan*	$4.6\text{x}10^{-5}$	integral to membrane (GO CC)	1.1	20	30
green*	$5.9\text{x}10^{-6}$	signal transduction (GO BP)	1.6	20	87
green*	$3.1\text{x}10^{-5}$	calcium ion binding (GO MF)	1.6	18	96
green*	$8.0\text{x}10^{-5}$	potassium channel activity (GO MF)	3.2	6	96
green*	$8.5\text{x}10^{-5}$	adrenoceptor activity (GO MF)	4.5	4	96
red	$5.0\text{x}10^{-5}$	transferase activity (GO MF)	1.4	20	71

Table S19: Top candidates for cross-species hub genes in the overlapping modules (mouse gene symbols), and descriptors of the modules; asterisk indicates possible technical confound for that module.

rank	chick green*, mouse brown	chick greenyellow mouse brown	chick magenta, mouse black	chick black, mouse orange*	chick cyan*, mouse magenta*	chick turquoise*, mouse salmon
	striatum	striatum	hippocampus	layer IV/ nidopallium	oligo-dendrocytes	?
1	*Dlx6*	*Pde10a*	*Abi1*	*Fam19a2*	*Bcas1*	*Cisd1*
2	*Foxo1*	*Pde7b*	*Ptk2b*	*Myo16*	*Gab1*	*2310003C23Rik*
3	*Pde10a*	*Dgkb*	*2010011I20Rik*	*Dctn3*	*Anln*	*Atp5f1*
4	*Drd1a*	*Rps6ka5*	*Pcdh1*	*Rorb*	*Tspan2*	*Rpl22l1*
5	*Adcy5*	*Tac1*	*Prickle2*	*Fam73b*	*Mbp*	*Ndufa5*

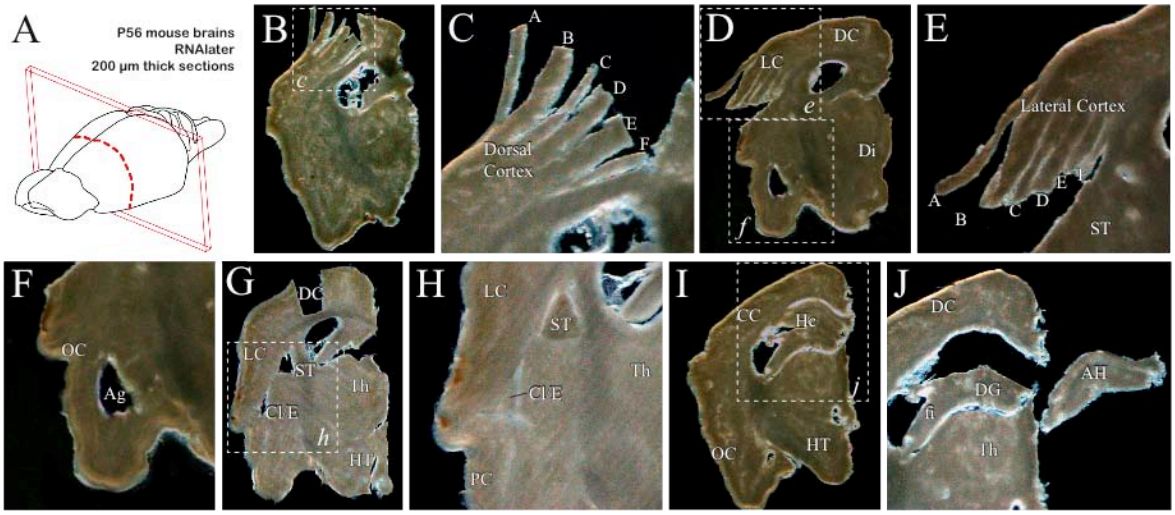


Fig. S1: Dissection of murine tissue samples. (A) Adult (P56) male mouse brains were immersed in RNAlater, coronally sectioned (200 µm) and dissected in a chilled solution of 1:1 mixture of RNAlater:PBS. (B) Dissection of dorsal cortical layers on a rostral section. (C) Magnification of the rectangle present in *B*, showing the cuts of the blade separating the layers. (D) Dissection of lateral cortical layers and pallial amygdala. (E) Magnification of the top rectangle present in *D*, showing the dissected cortical layers. (F) Magnification of the bottom rectangle depicted in *D*, following dissection of the pallial amygdala (Ag). (G) Dissection of the striatum (ST) and the claustrum/endopiriform complex (Cl/E). Layers of dorsal cortex were already dissected from a sector (indicated with DC) in this image. (H) Higher magnification of the rectangle in *G*, where both striatum (ST) and the claustrum/endopiriform complex (CLE) have already been cut around, but not removed. (I) Dissection of hippocampal (Hc) tissue samples from caudal sections. (J) Magnification from rectangle present in *I*, showing the removal of the Ammon's horn (AH) from the dentate gyrus (DG) and the fimbria (fi); the hippocampal samples included both AH and DG. Other abbreviations used are cerebral cortex (CC), diencephalon (Di), hypothalamus (HT), lateral cortex (LC), olfactory cortex (OC), piriform cortex (PC) and thalamus (Th).

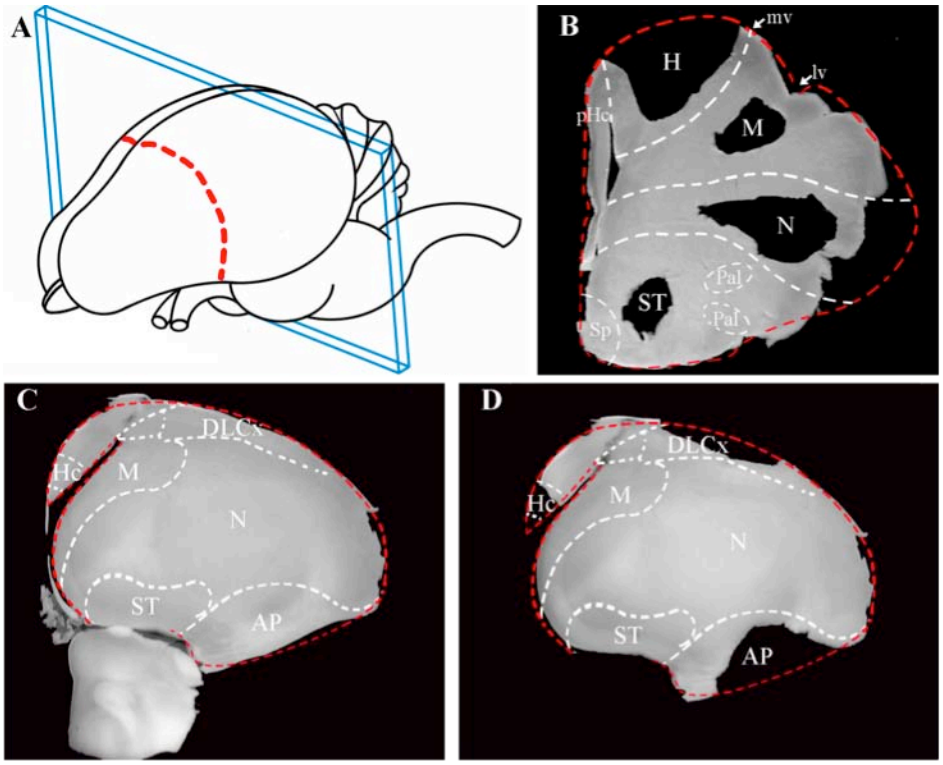


Fig. S2: Chicken brains were coronally sectioned as depicted in (A). Representative sections from which hyperpallium (H), mesopallium (M), nidopallium (N), arcopallium (AP), striatum (ST), hippocampus (Hc) and dorsolateral cortex (DCLx) were dissected are shown in (B; anterior level) and (D; posterior level). (C) An undissected posterior section with approximate boundaries marked.

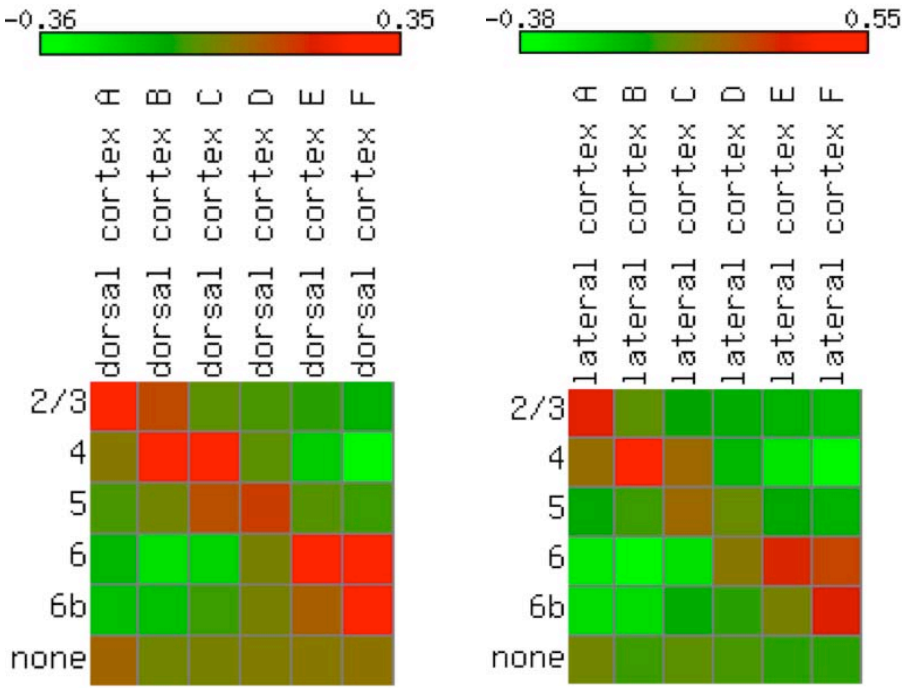


Fig. S3: Dissected samples correspond to specific neocortical layers. Heatmaps were created using Matrix2png (http://www.bioinformatics.ubc.ca/matrix2png/) and reflect the calculation described above.

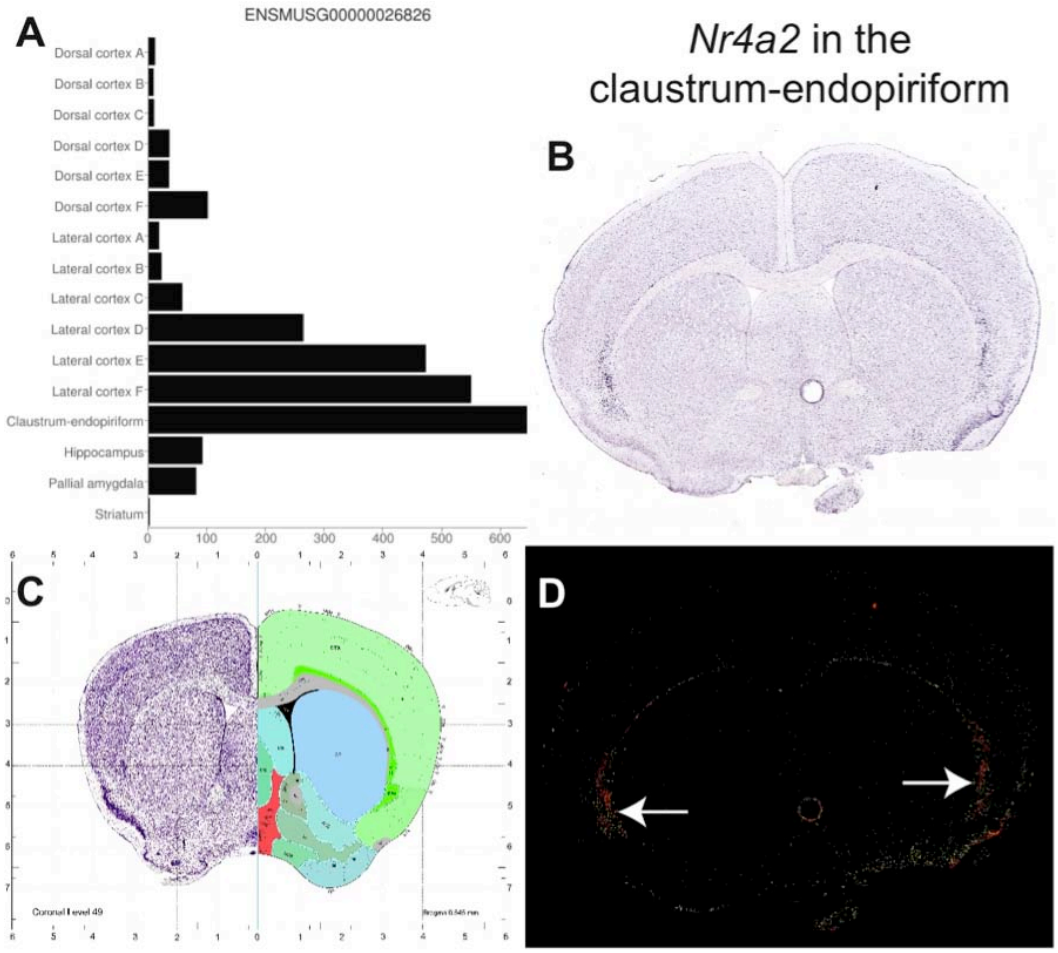


Fig. S4: *Nr4a2* (also known as *Nurr1*), one of the three most specific marker genes in assessed regions for the claustrum-endopiriform as assessed by RNA-seq, was also specific to this region (among regions assessed) by *in situ* hybridization. (A) gives the FPKM values of the gene in each dissected mouse region as assessed by RNA-seq, (B) is an *in situ* hybridization image of the gene, (C) gives the corresponding coronal section of the reference atlas and (D) gives the expression analysis of the *in situ* hybridization image with the region of interest indicated with arrows. (B)-(D) are reproduced from the Allen Mouse Brain Atlas.

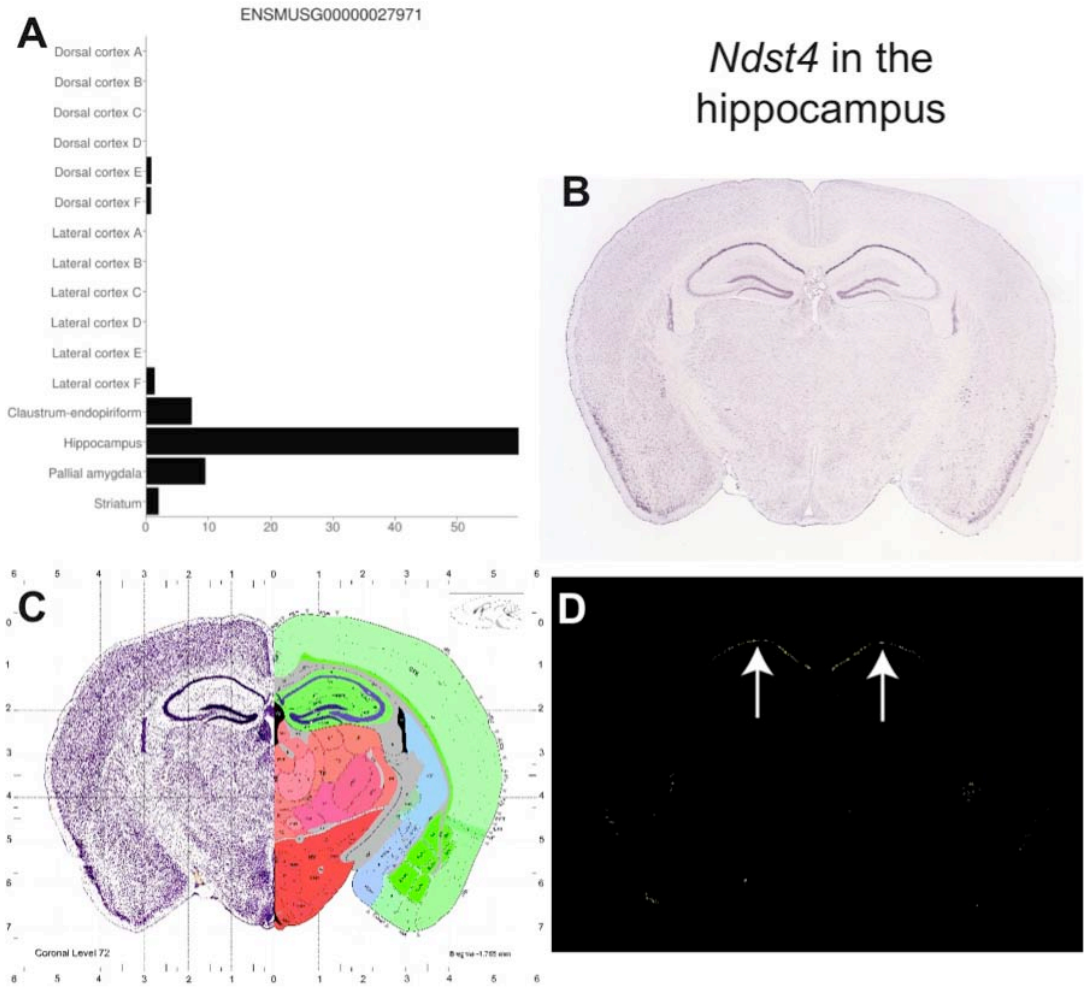


Fig. S5: *Ndst4*, one of the three most specific marker genes in assessed regions for the hippocampus as assessed by RNA-seq, was also specific to this region (among regions assessed) by *in situ* hybridization. (A) gives the FPKM values of the gene in each dissected mouse region as assessed by RNA-seq, (B) is an *in situ* hybridization image of the gene, (C) gives the corresponding coronal section of the reference atlas and (D) gives the expression analysis of the *in situ* hybridization image with the region of interest indicated with arrows. (B)-(D) are reproduced from the Allen Mouse Brain Atlas.

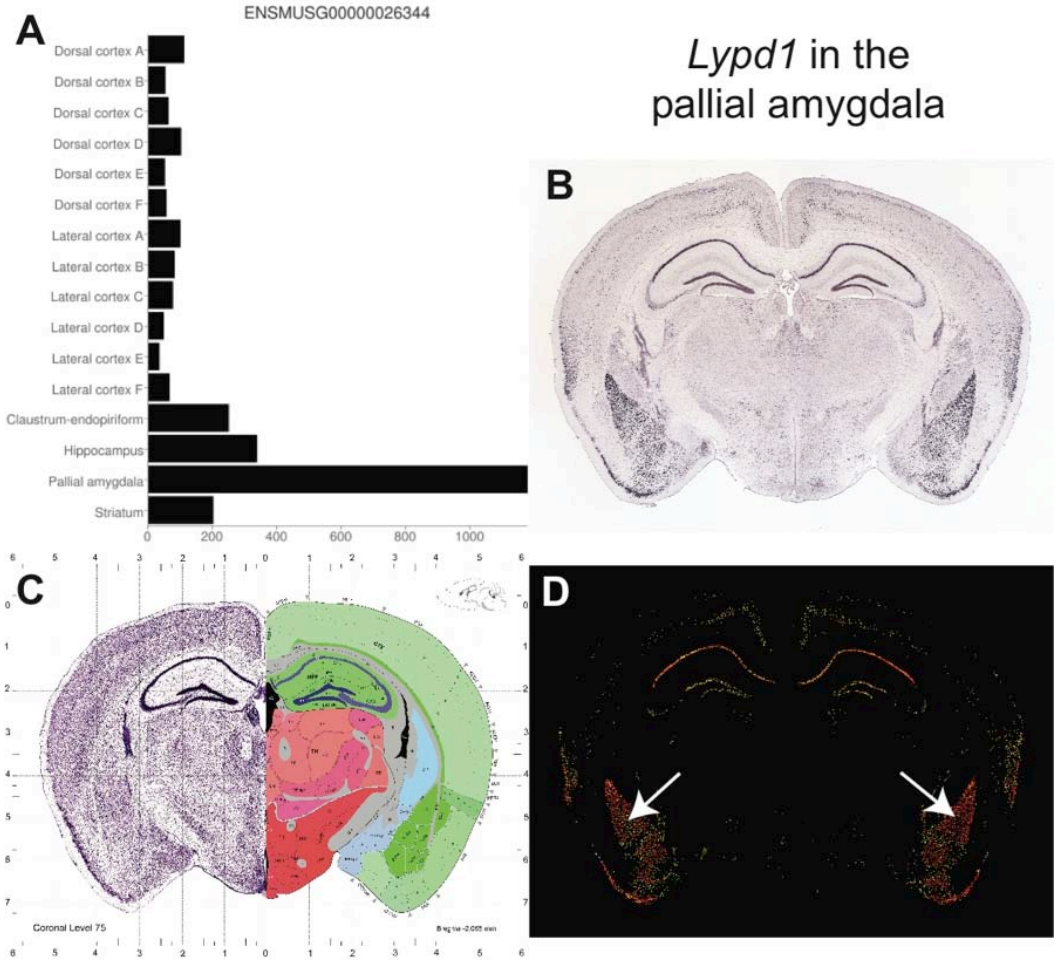


Fig. S6: *Lypd1*, one of the three most specific marker genes in assessed regions for the pallial amygdala as assessed by RNA-seq, was also specific to this region (among regions assessed) by *in situ* hybridization. (A) gives the FPKM values of the gene in each dissected mouse region as assessed by RNA-seq, (B) is an *in situ* hybridization image of the gene, (C) gives the corresponding coronal section of the reference and (D) gives the expression analysis of the *in situ* hybridization image with the region of interest indicated with arrows. (B)-(D) are reproduced from the Allen Mouse Brain Atlas.

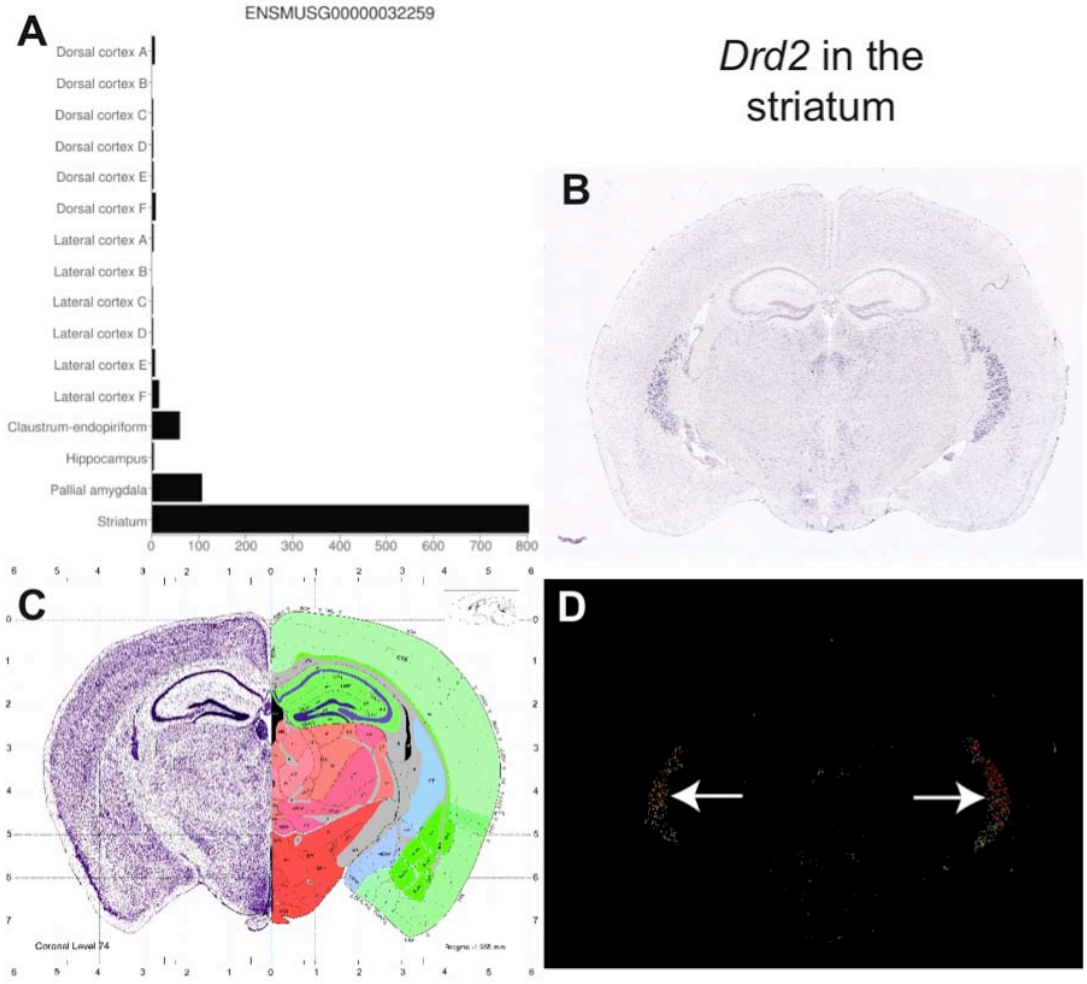


Fig. S7: *Drd2*, one of the three most specific marker genes in assessed regions for the striatum as assessed by RNA-seq, was also specific to this region (among regions assessed) by *in situ* hybridization. (A) gives the FPKM values of the gene in each dissected mouse region as assessed by RNA-seq, (B) is an *in situ* hybridization image of the gene, (C) gives the corresponding coronal section of the reference atlas and (D) gives the expression analysis of the *in situ* hybridization image with the region of interest indicated with arrows. (B)-(D) are reproduced from the Allen Mouse Brain Atlas.

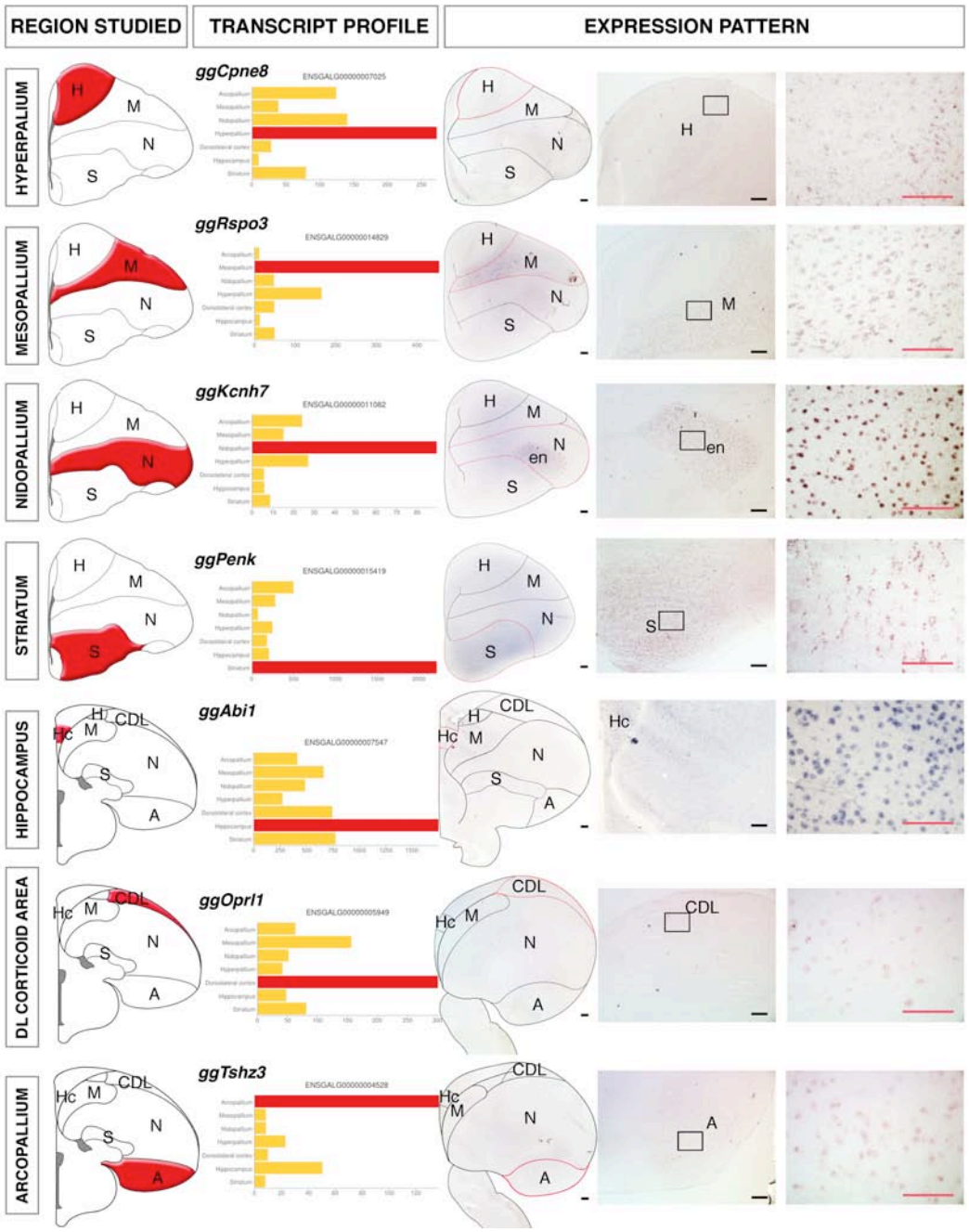


Fig. S8. ISH validations of the chicken dissections. Coronal chicken brain sections showing the telencephalic region studied (red area), expression levels in FPKM highlighting the enriched marker gene (red bar) and the corresponding expression pattern. A, arcopallium; CDL, corticoid dorsolateral corticoid area; en, entopallium; H, hyperpallium; Hc, hippocampus; M, mesopallium; N, nidopallium; S, striatum. Scale bars: black=500μm, red= 200μm.

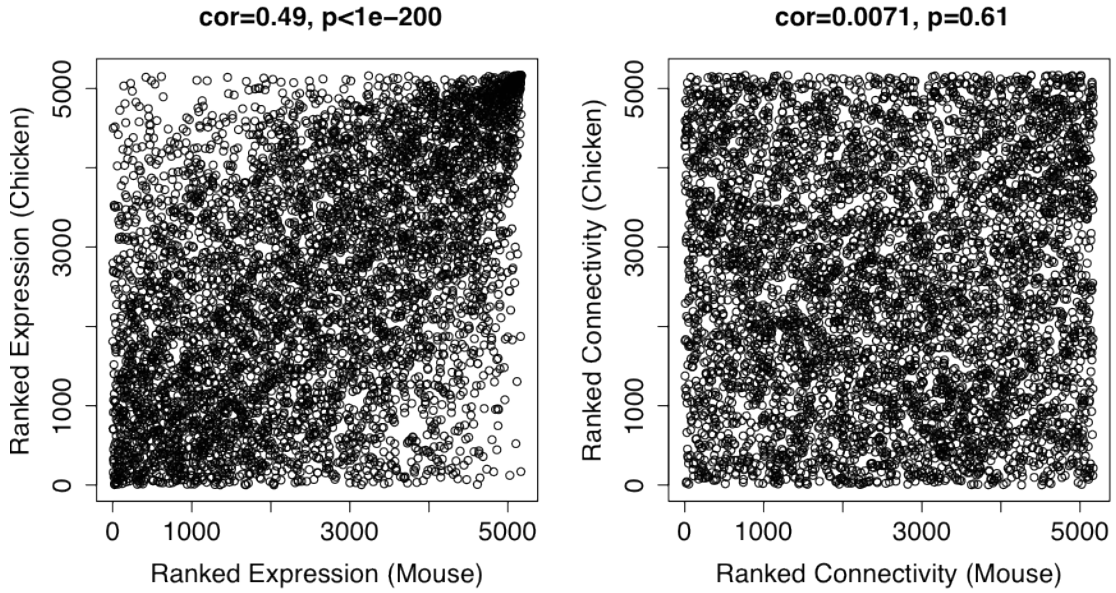


Fig. S9: Plots of the ranked expression of orthologs in both species, and of the ranked connectivity. Correlation coefficients above represent Spearman's rank correlation coefficient of the original variables (Pearson's *rho* of the ranked variables) and *p*-values are two-tailed.

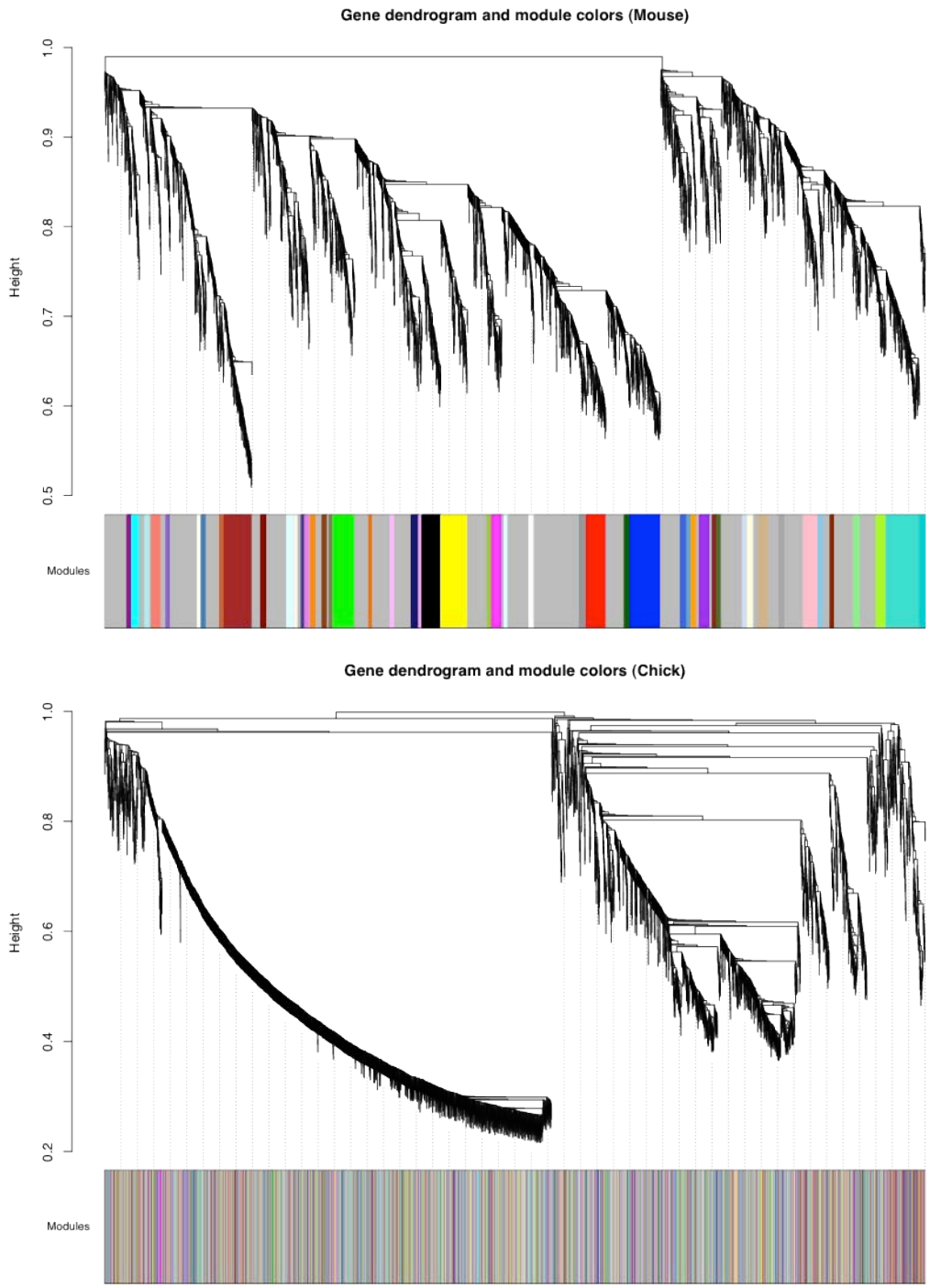


Fig. S10: Mouse modules and dendrogram of topological dissimilarity (top) and mouse modules projected onto the corresponding chicken dendrogram (bottom).

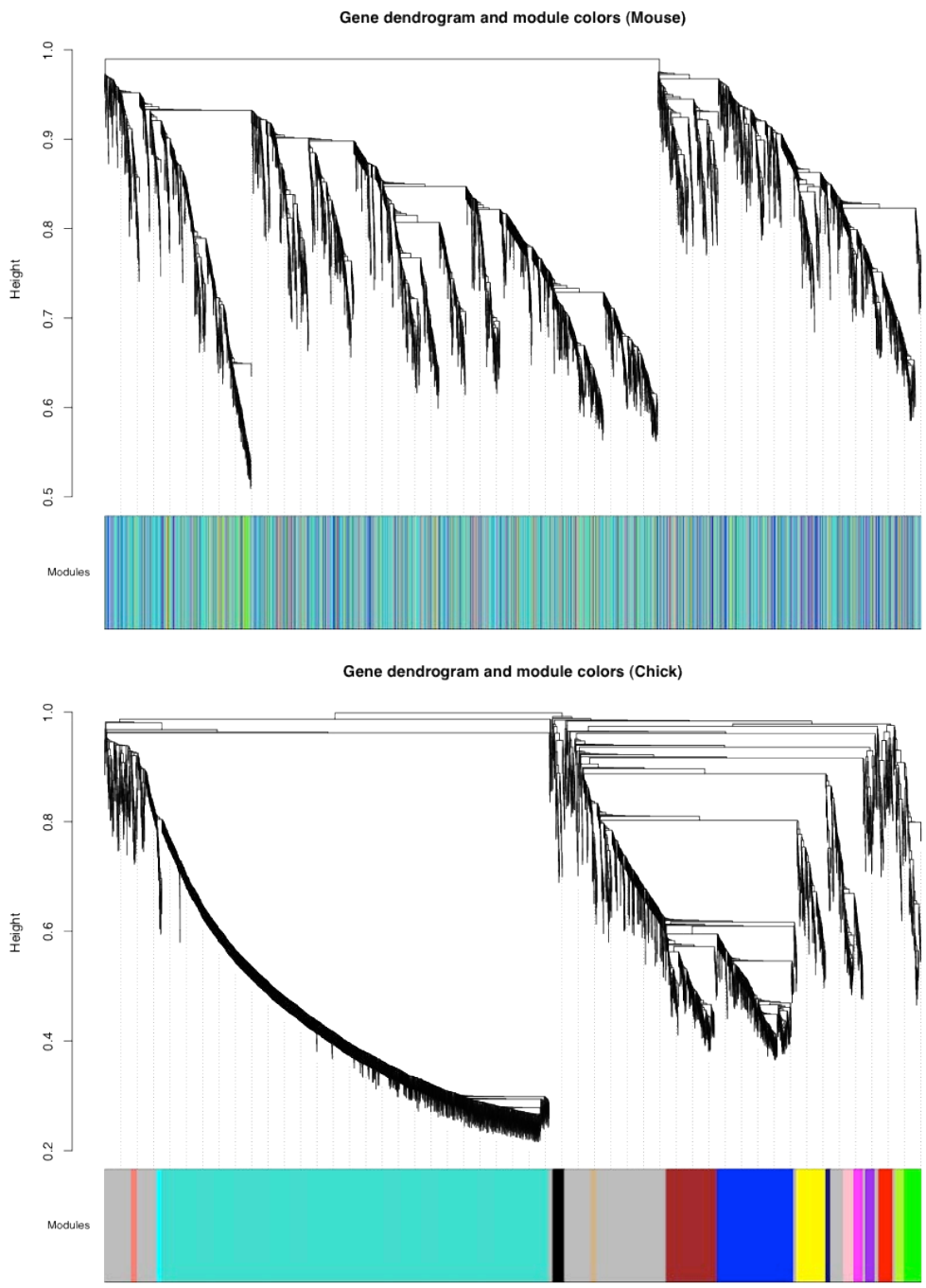


Fig. S11: Chicken modules and dendrogram of topological dissimilarity (bottom) and chicken modules projected onto the corresponding mouse dendrogram (bottom).

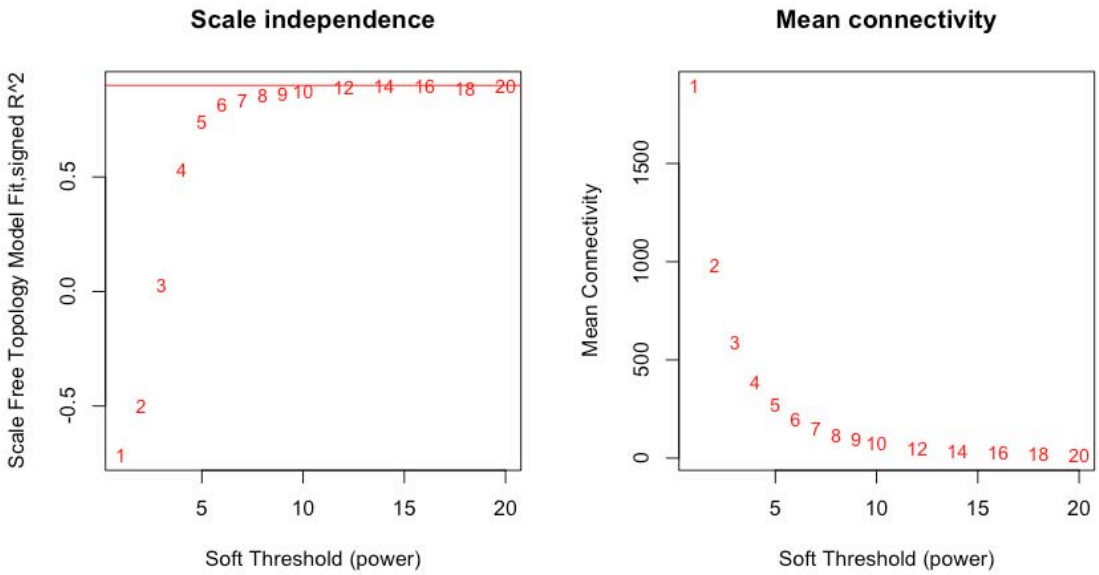


Fig. S12: Parameters for choosing the soft threshold in the mouse reanalysis.

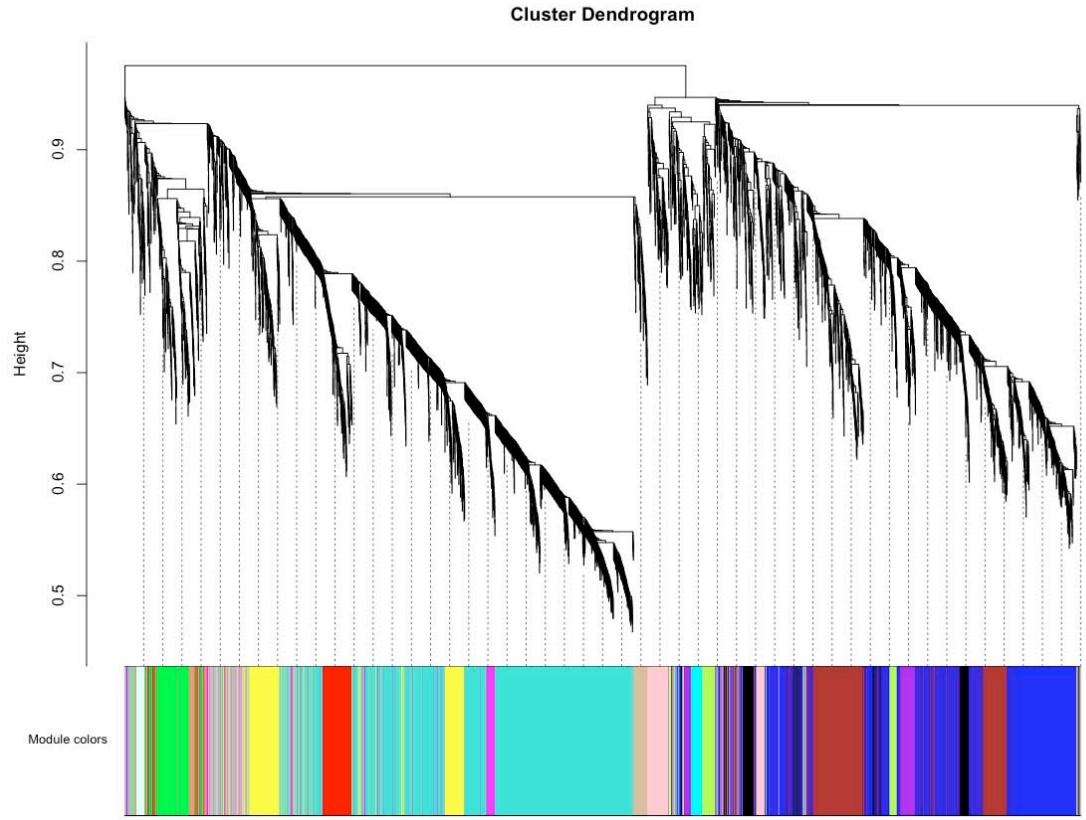


Fig. S13: Topological dissimilarity dendrogram and coexpressed mouse modules in the reanalysis.

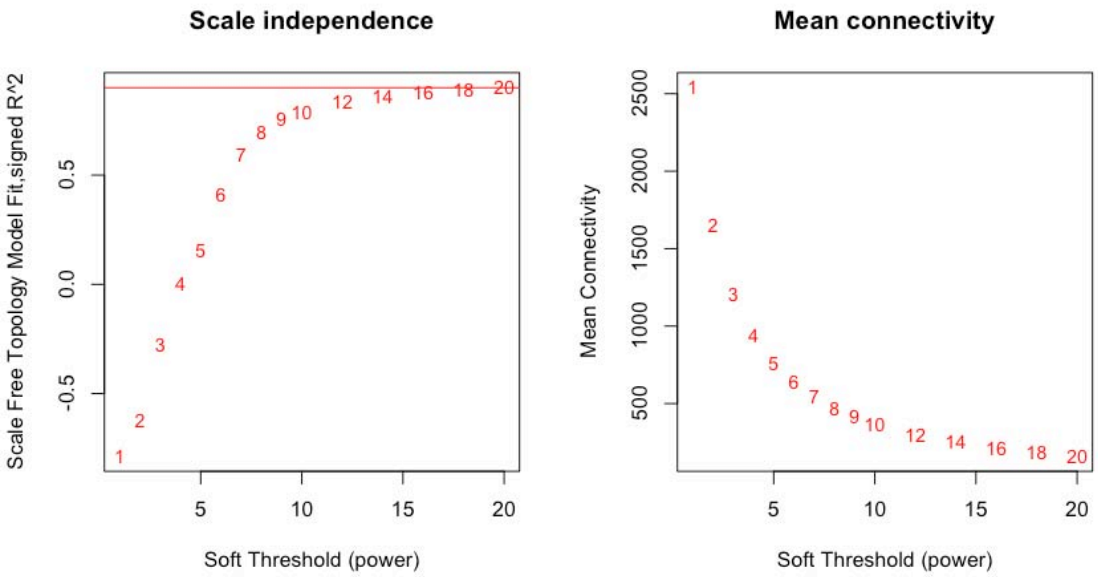


Fig. S14: Parameters for choosing the soft threshold in the chicken reanalysis.

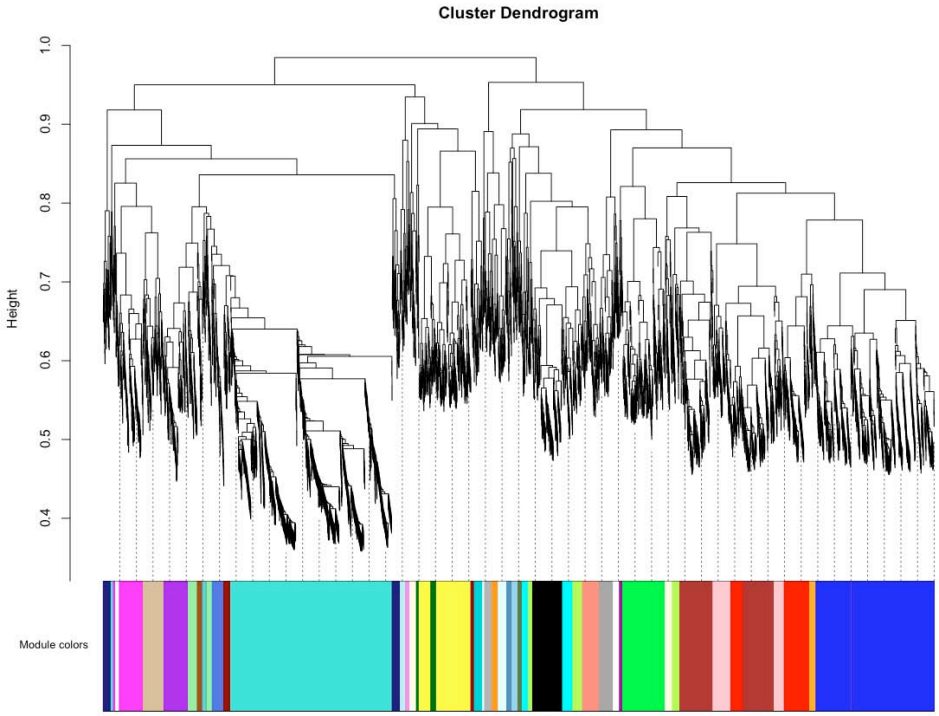


Fig. S15: Topological dissimilarity dendrogram and coexpressed chicken modules in the reanalysis.

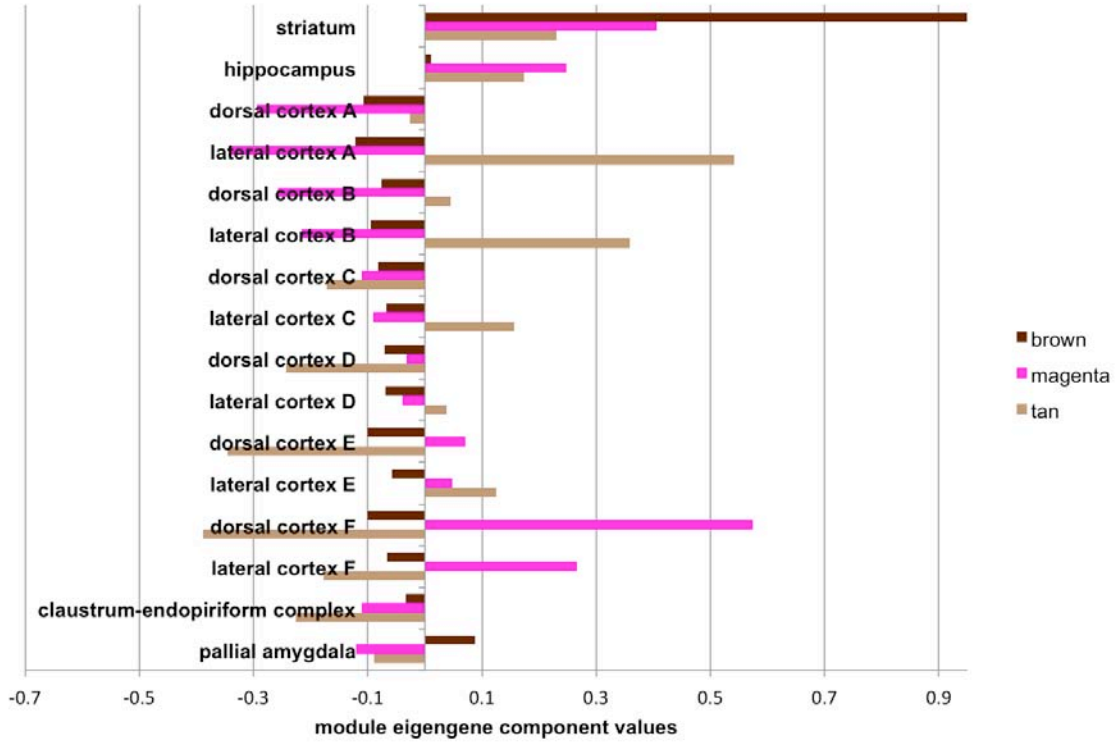


Fig. S16: Module eigengene expression in different samples of the mouse modules with significant functional enrichments.

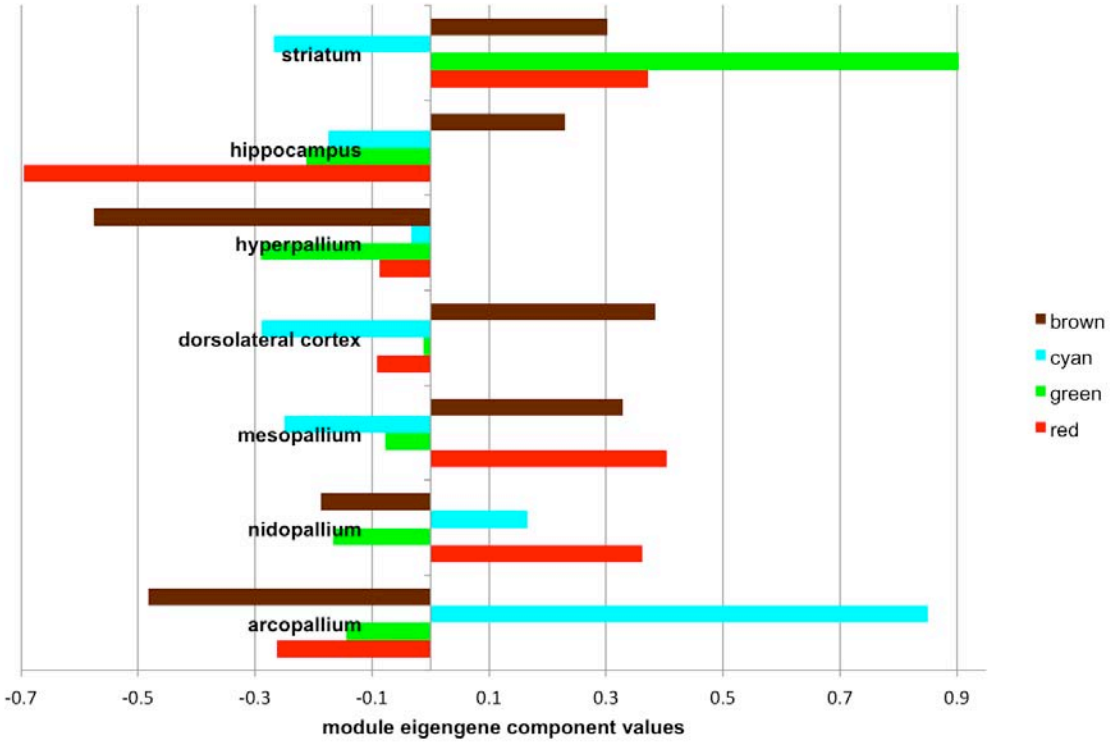


Fig. S17: Module eigengene expression in different samples of the chick modules with significant functional enrichments.

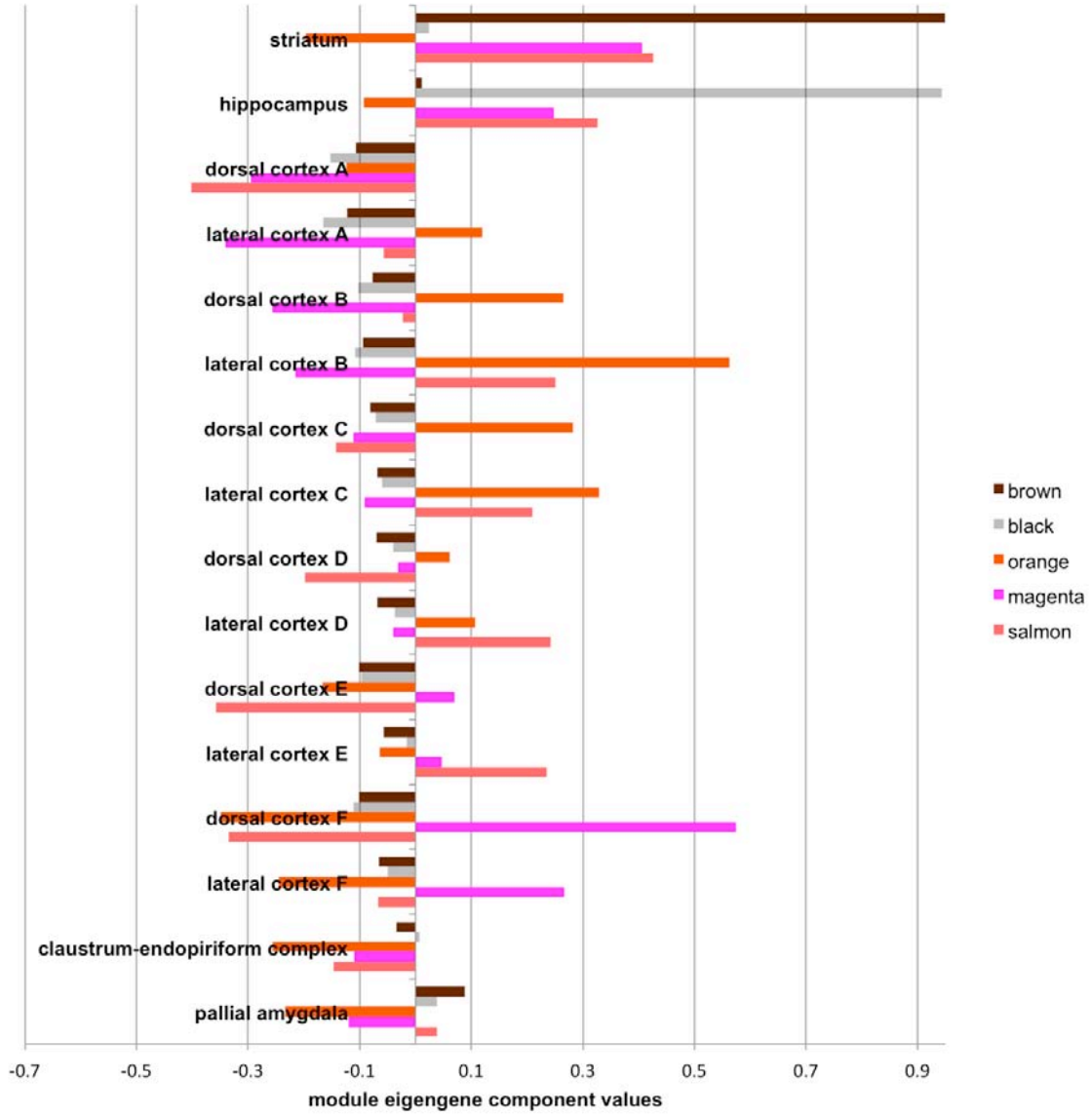


Fig. S18: Module eigengene expression of the mouse modules that overlapped chick modules.

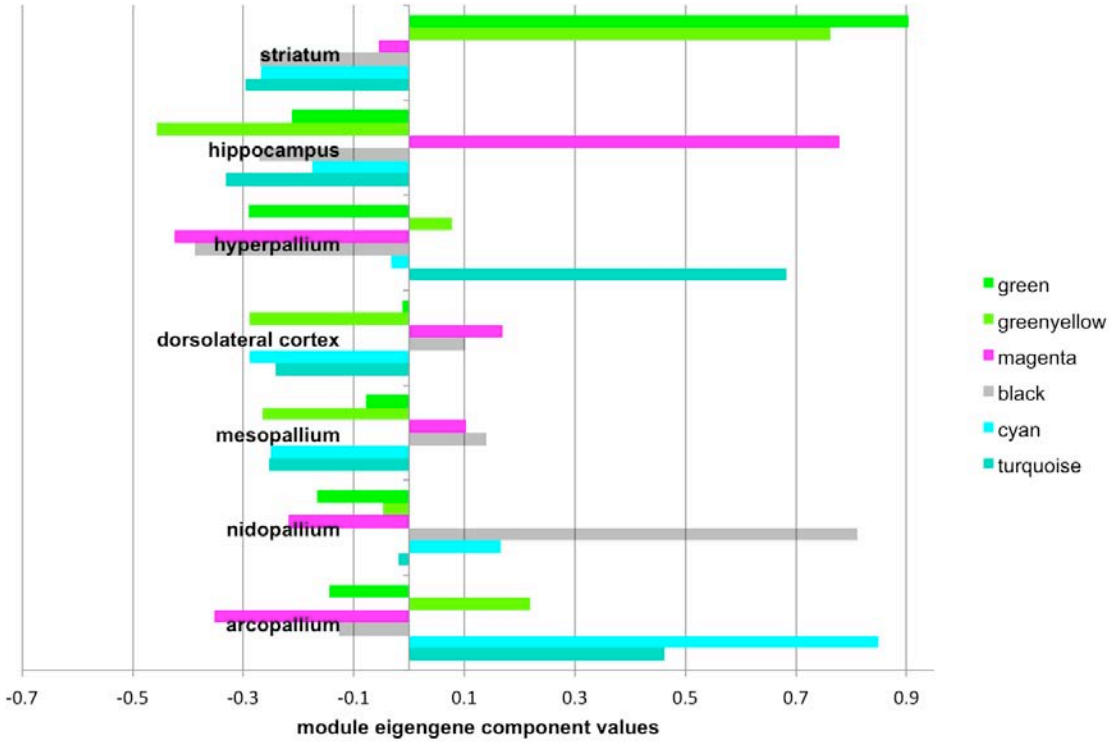


Fig. S19: Module eigengene expression of the chicken modules that overlapped mouse modules.

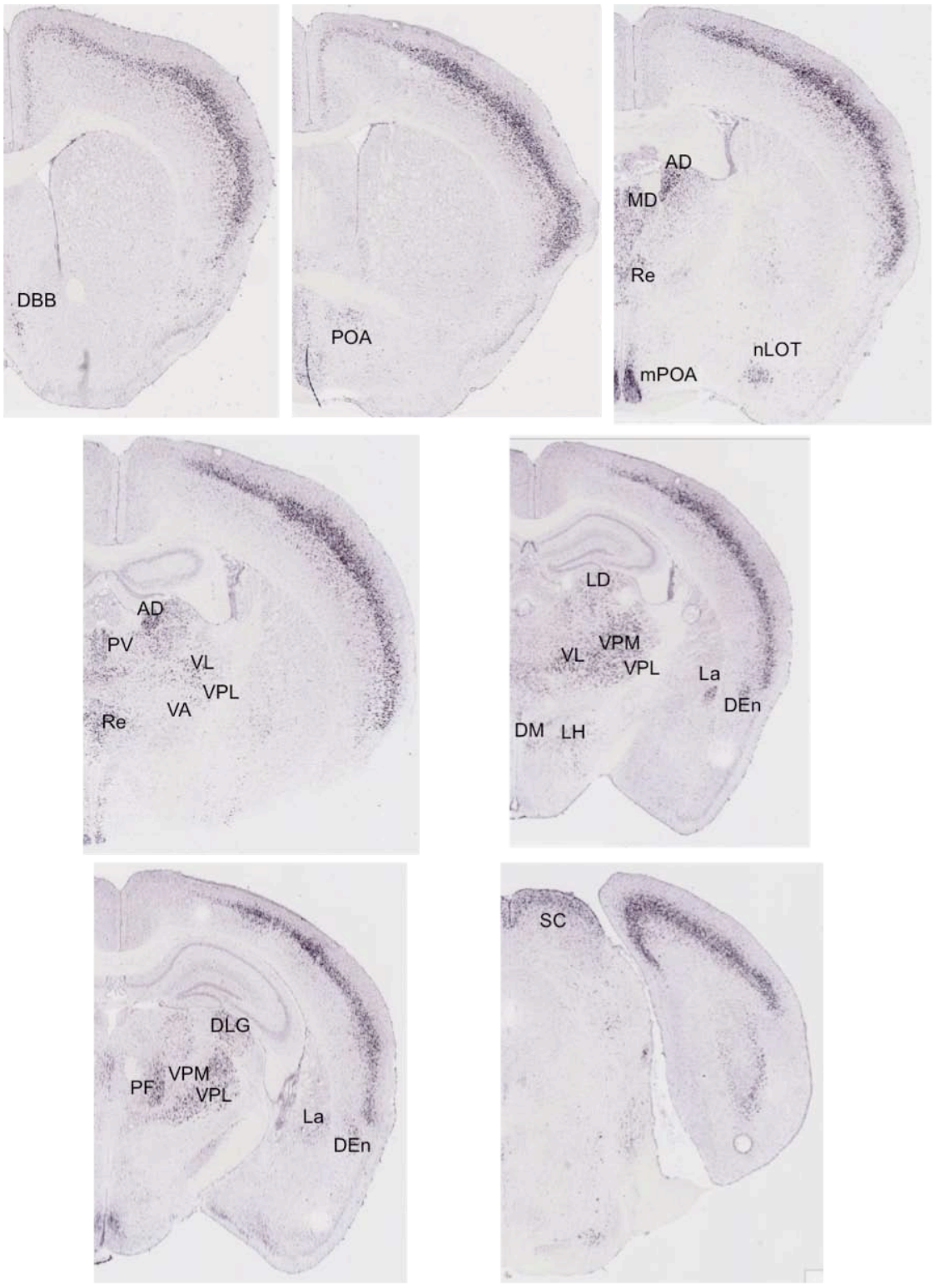


Fig. S20: *Rorb* expression in adult mouse from the Allen Mouse Brain Atlas (8). AD: Antero dorsal Thalamic nucleus; DBB: Diagonal band of Broca; Den: Dorsal Endopiriform nucleus; DLG: Dorsal lateral Geniculate nucleus; DM: Dorso medial

Hypothalamic nucleus; La: Lateral amygdaloid nucleus; LD: Latero dorsal Thalamic nucleus; LH: Lateral Hypothalamic nucleus; MD: Medio dorsal Thalamic nucleus; mPOA: medial Preoptic area; nLOT: nucleus of the lateral olfactory tract; PF: Parafascicular nucleus of the Thalamus; POA: Preoptic area; PV: Paraventricular thalamic nucleus; Re: Reuniens nucleus; SC: Superior colliculus; VA: Ventral anterior Thalamic nucleus; VL: Ventral lateral Thalamic nucleus; VPL: Ventral postero-lateral Thalamic nucleus.

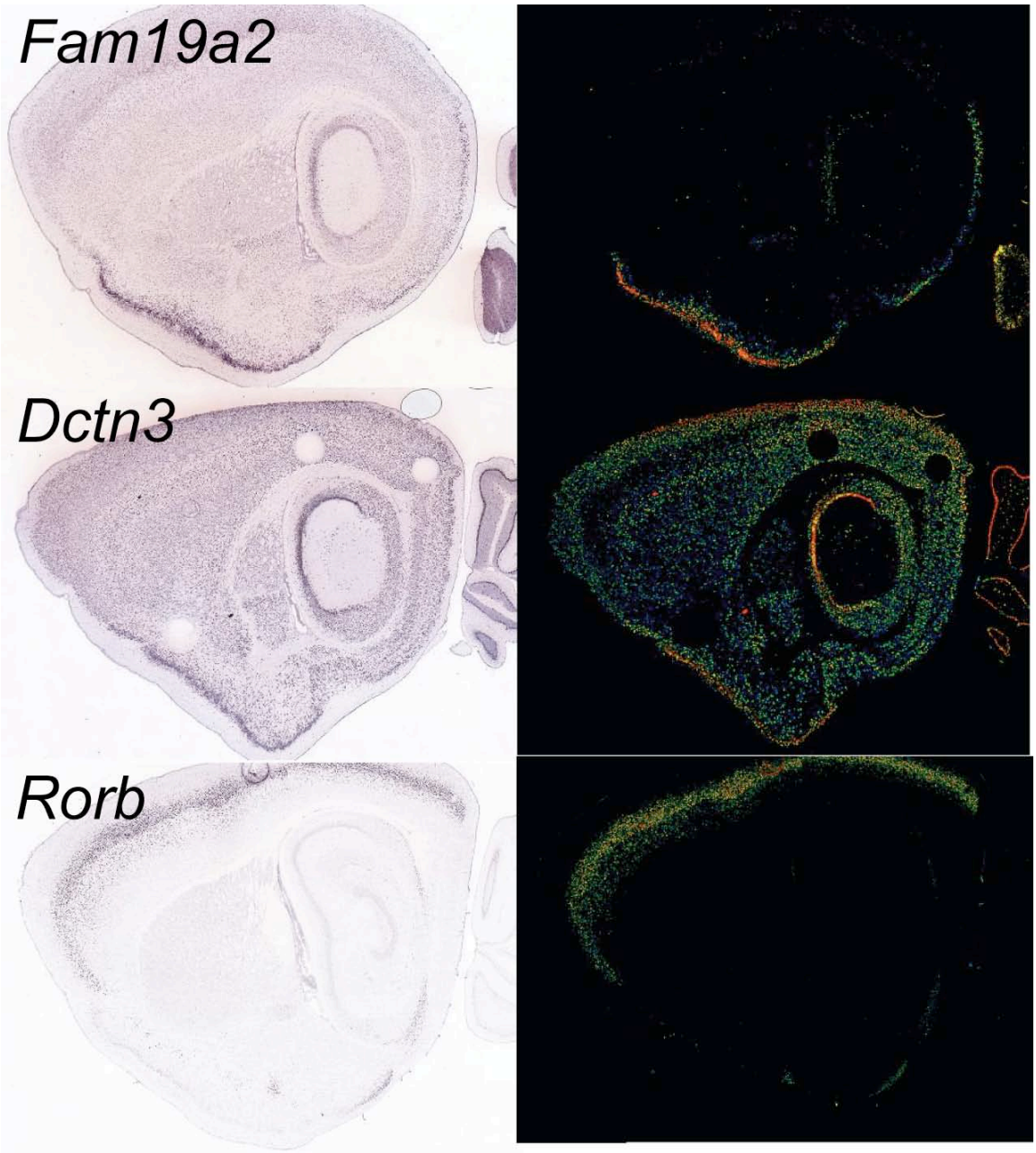


Fig. S21: *Fam19a2* and *Dctn3* were both highly expressed in layer 2 of piriform cortex in the Allen Mouse Brain Atlas (8). *Rorb*, in contrast, was not particularly highly expressed in piriform cortex.

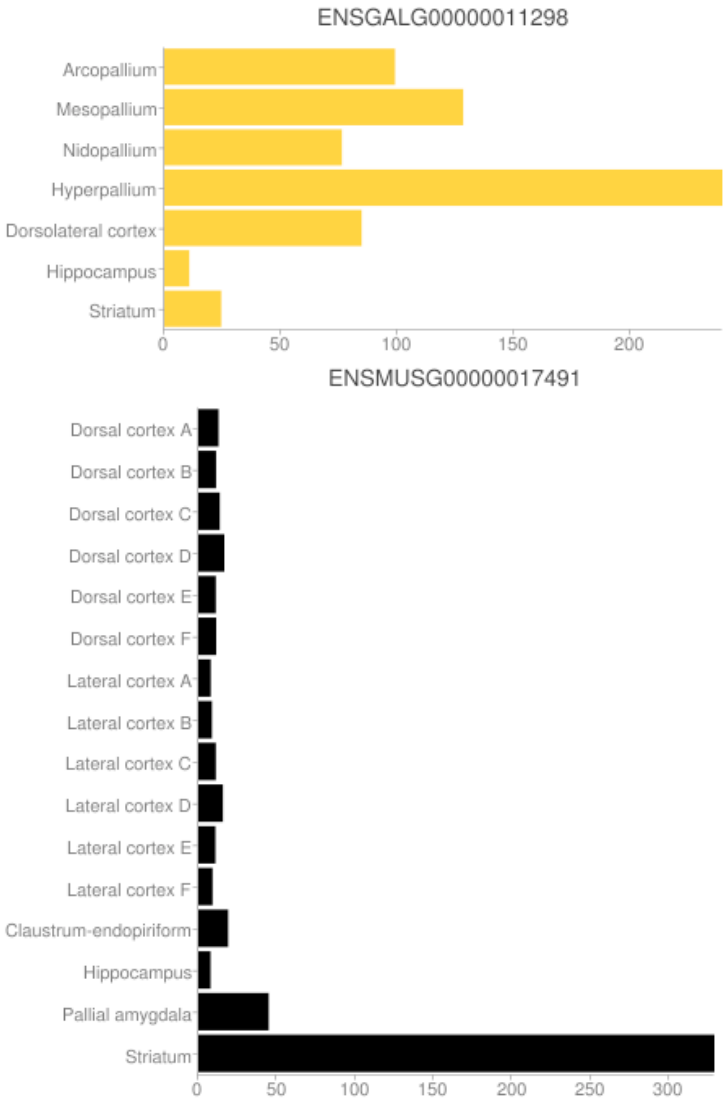


Fig. S22: *Rarb* is specific to mouse striatum (below) but relatively rare in chick striatum (above). X-axis represents FPKM.

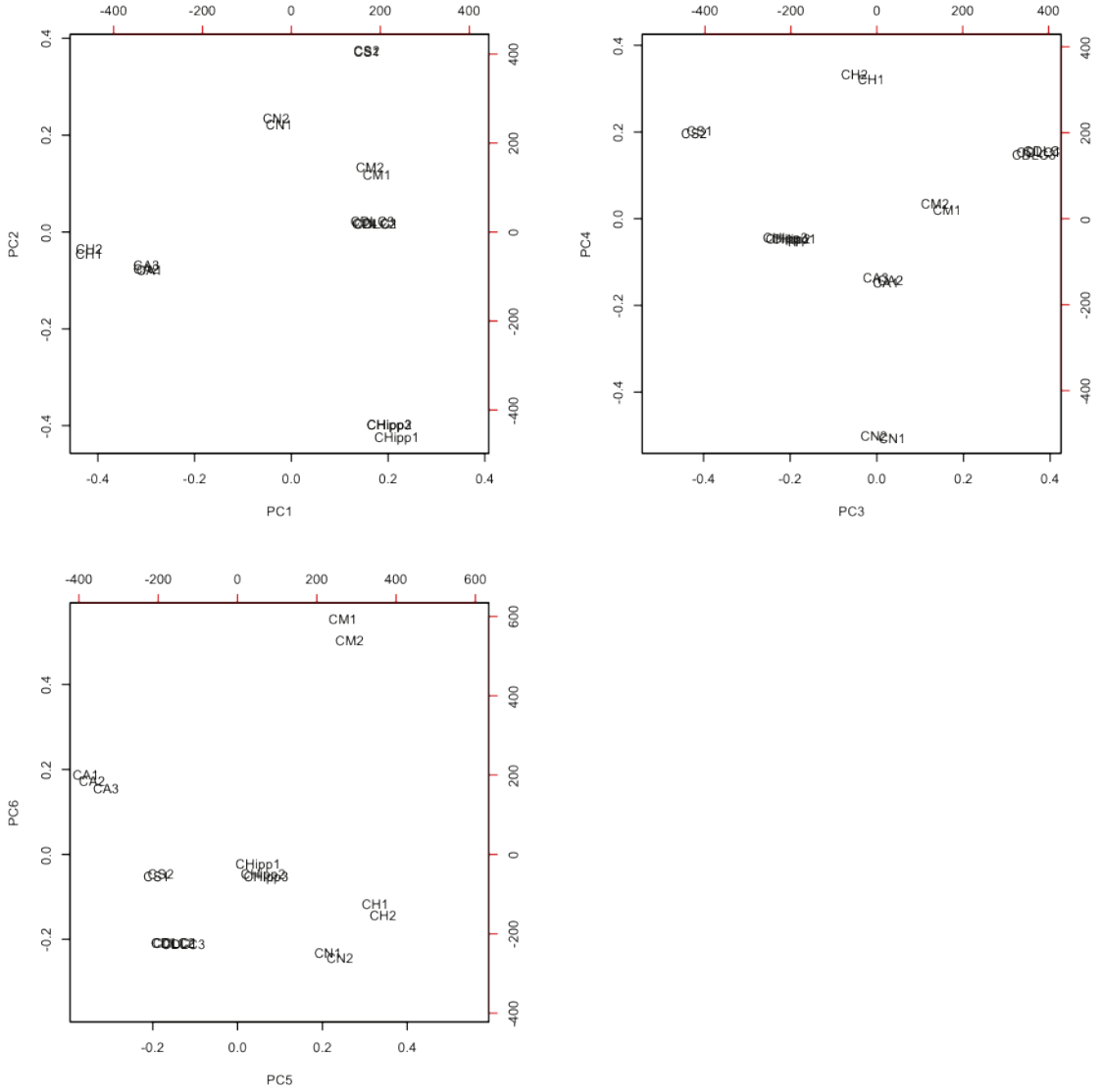


Fig. S23: Different lanes and flowcells of chicken libraries plotted on the first six principal components.

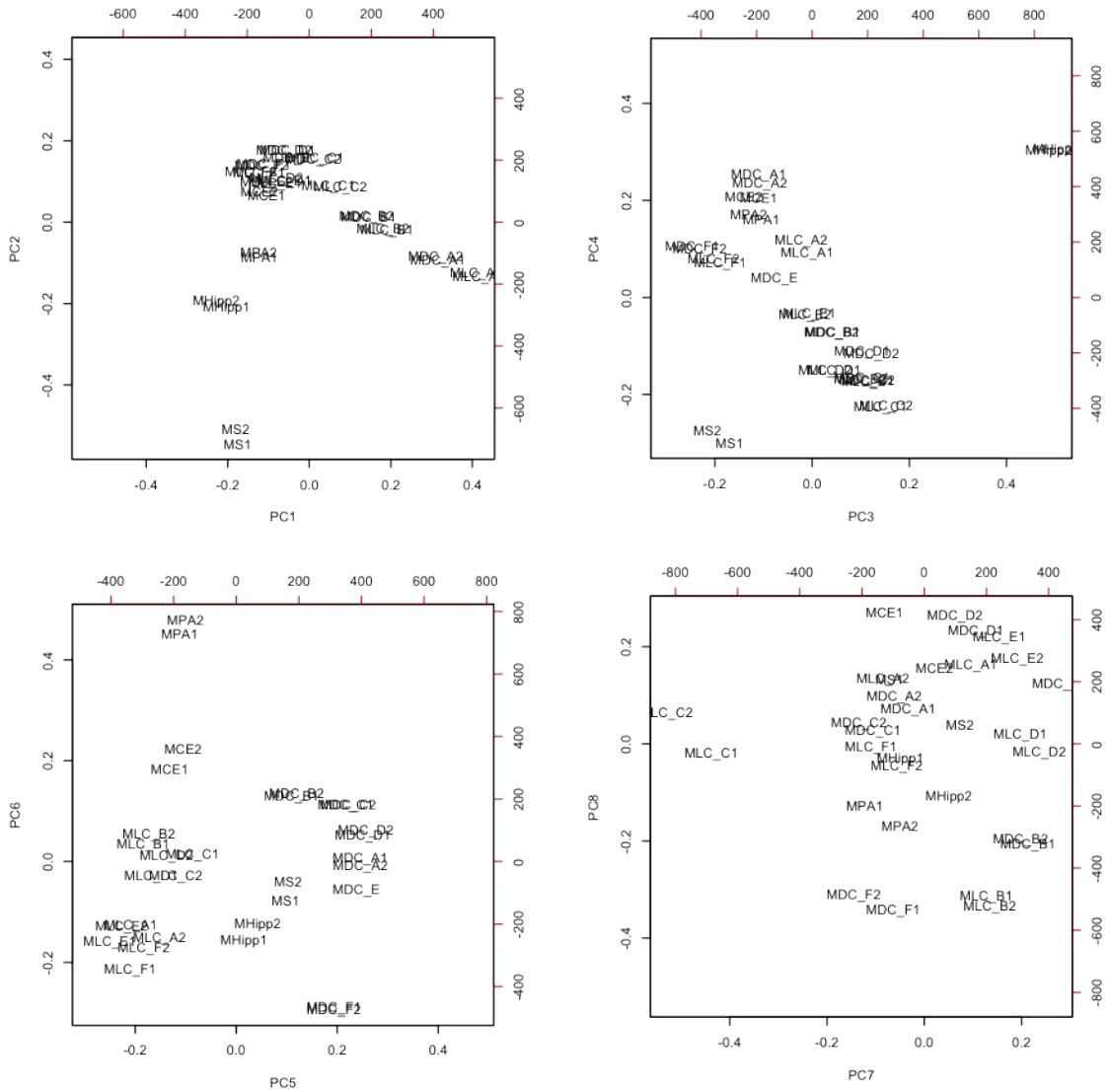


Fig. S24: Different lanes and flowcells of mouse libraries plotted on the first eight principal components.

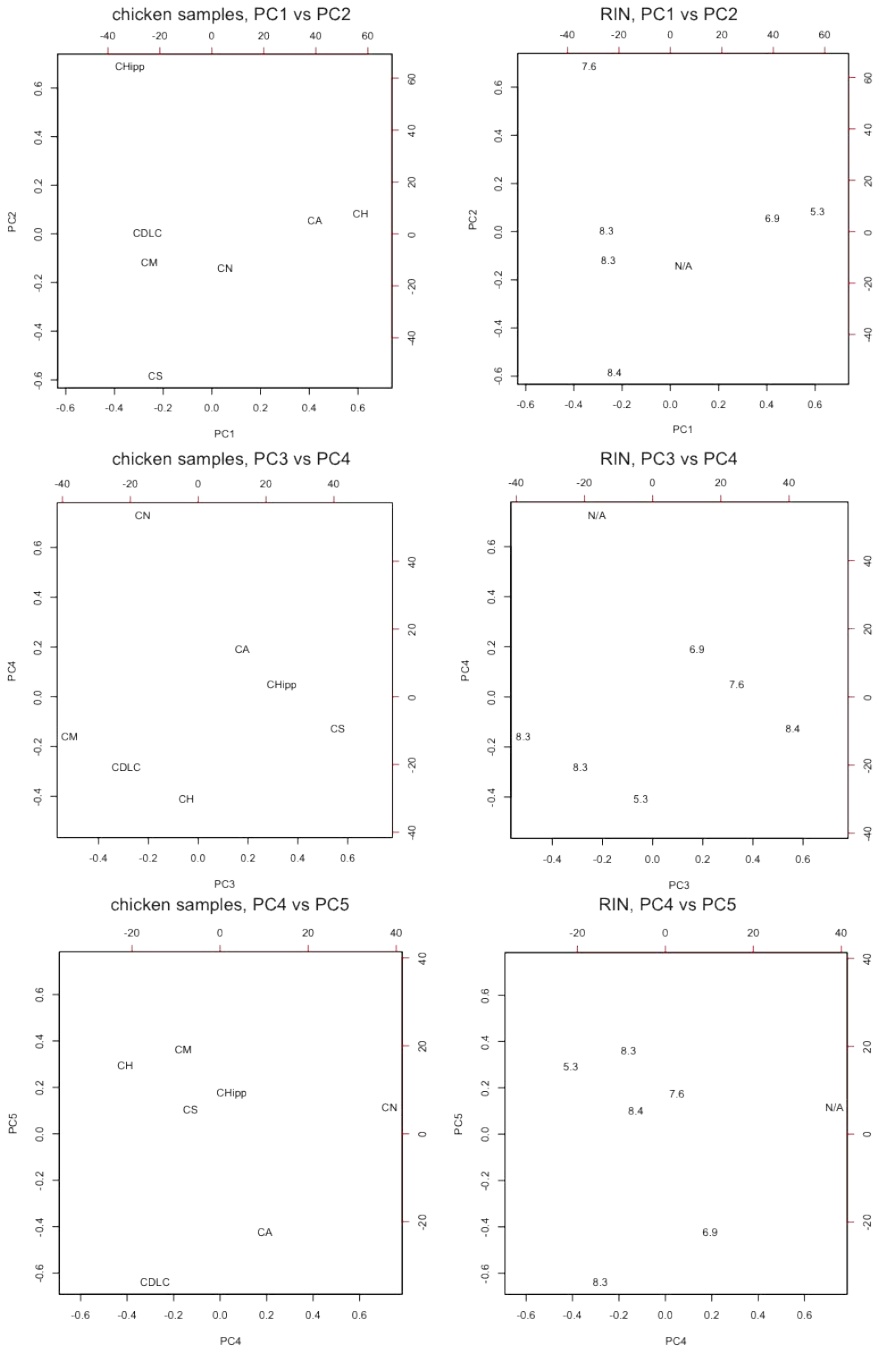


Fig. S25: RNA Integrity Numbers (RINs) of chicken samples plotted on chicken principal components.

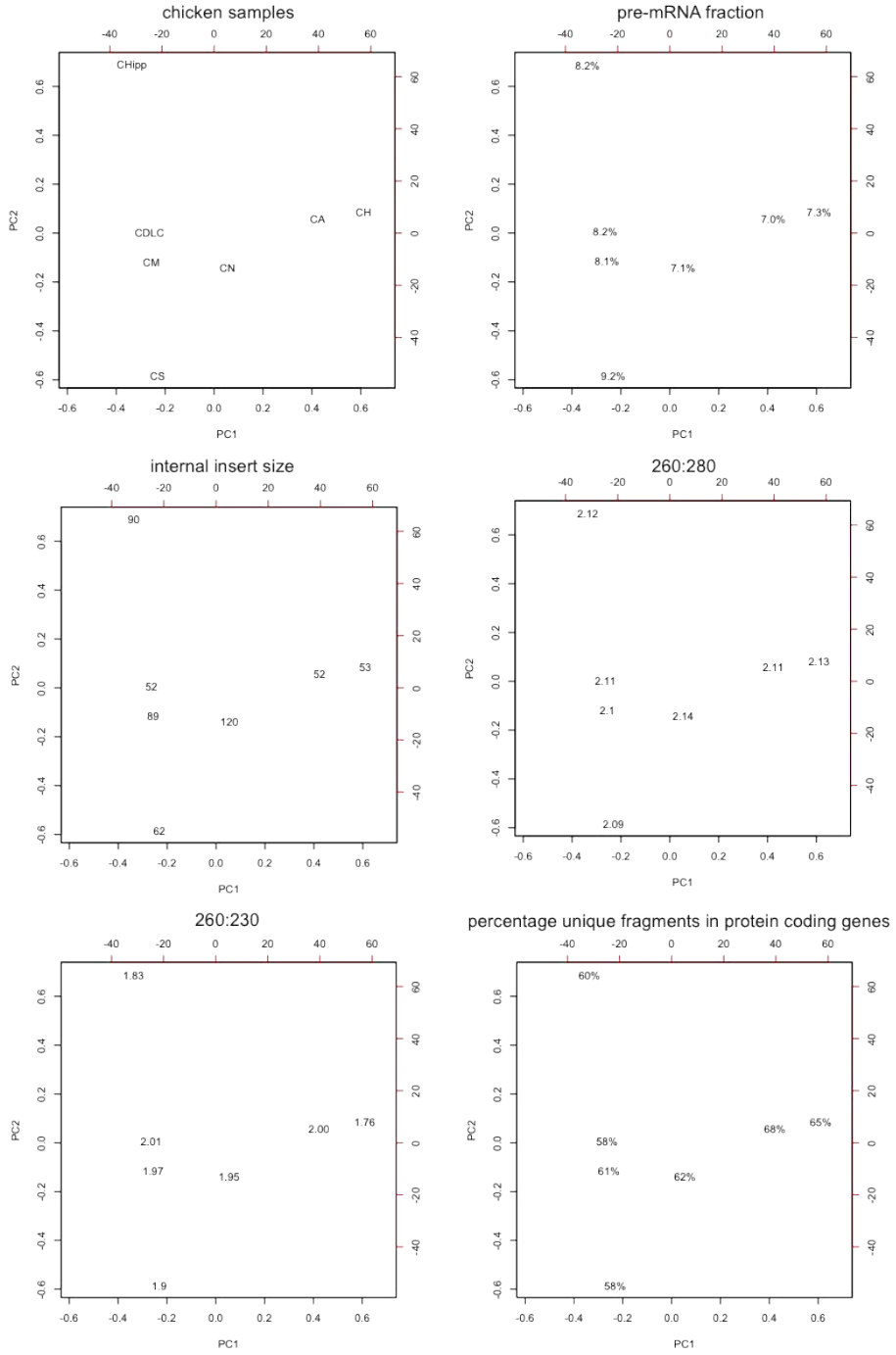


Fig. S26: Technical variables plotted on the first two chicken principal components.

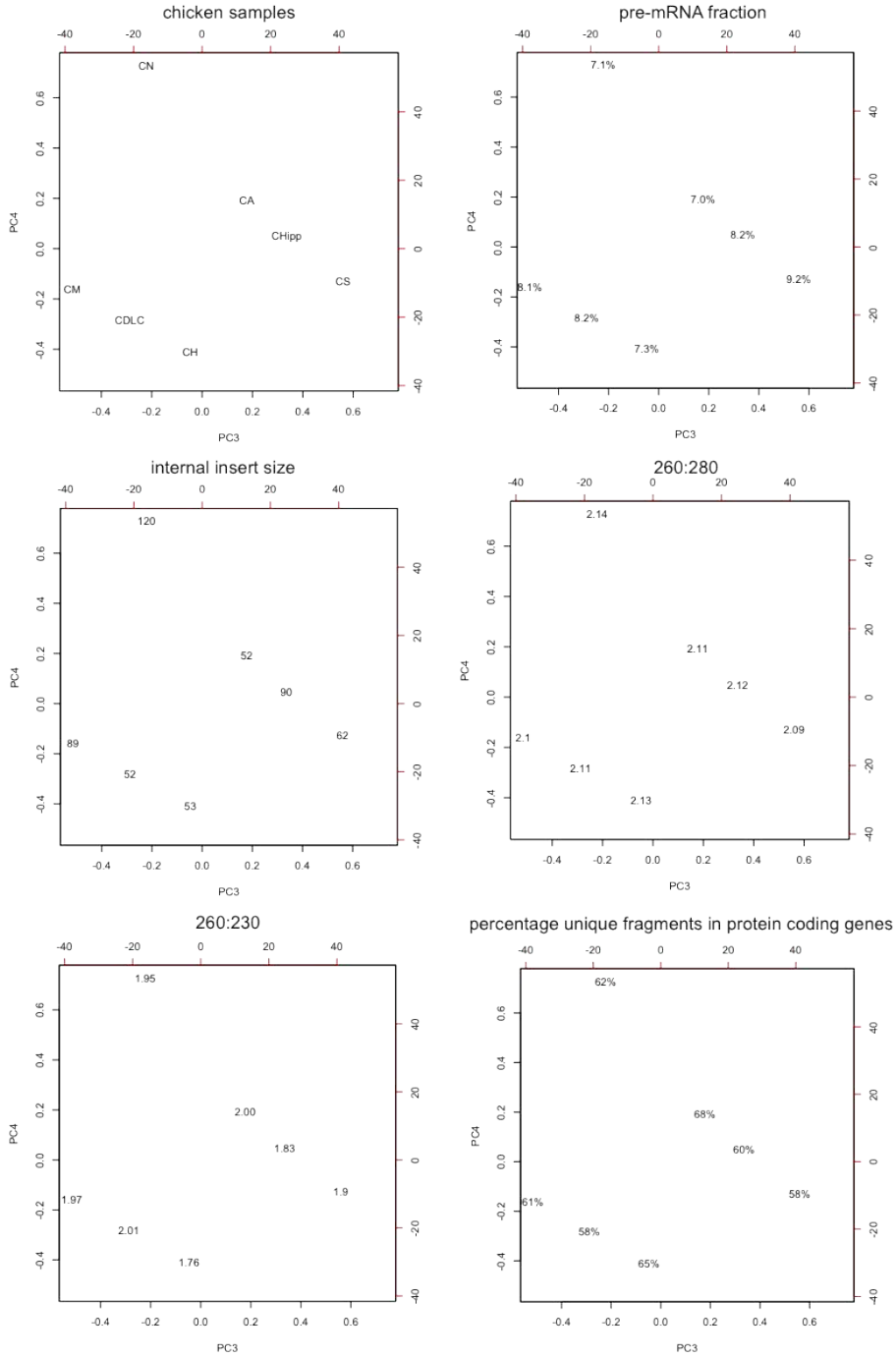


Fig. S27: Technical variables plotted on the third and fourth chicken principal components.

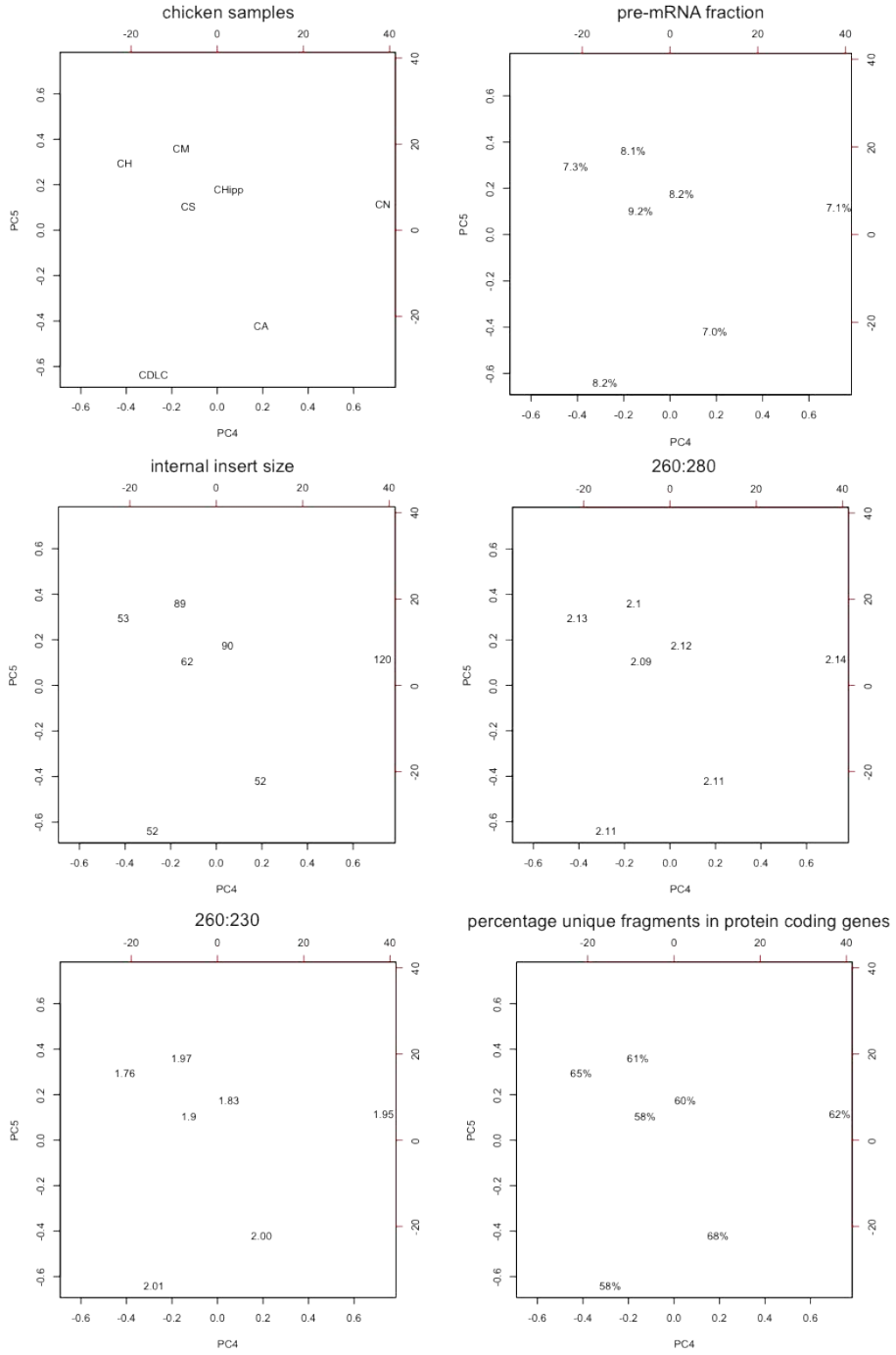


Fig. S28: Technical variables plotted on the first and fifth chicken principal components.

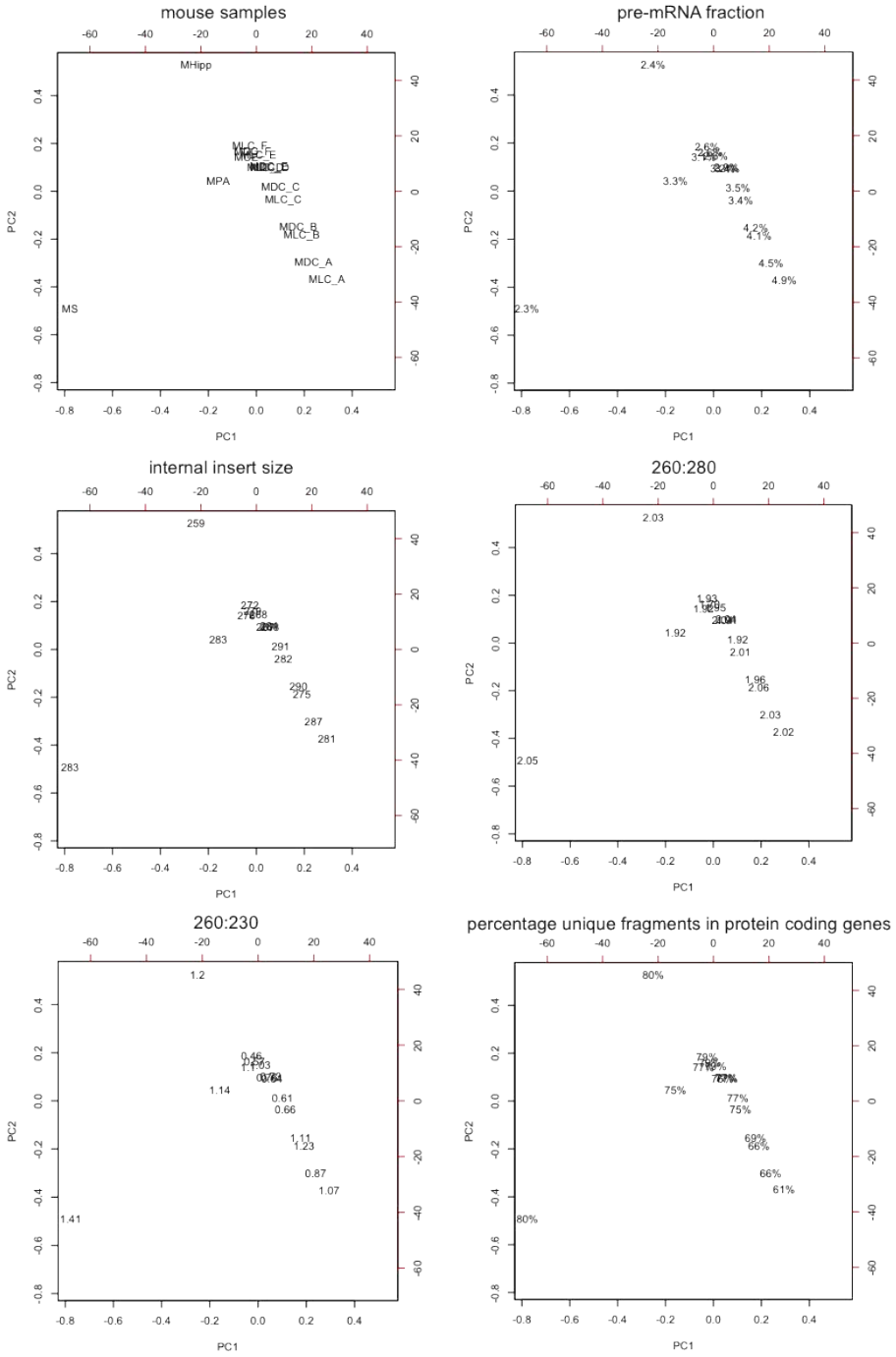


Fig. S29: Technical variables plotted on the first and second mouse principal components.

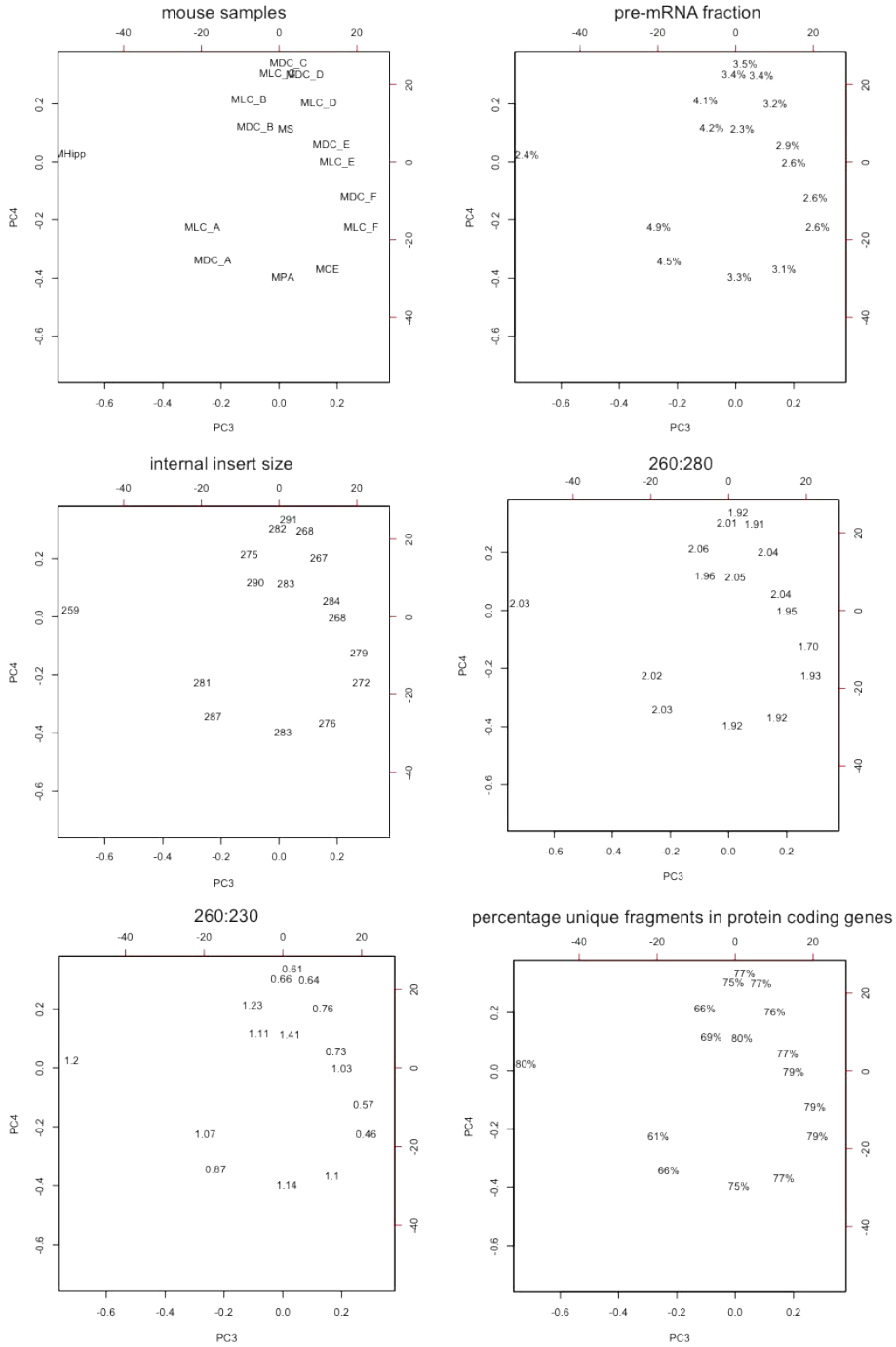


Fig. S30: Technical variables plotted on the third and fourth mouse principal components.

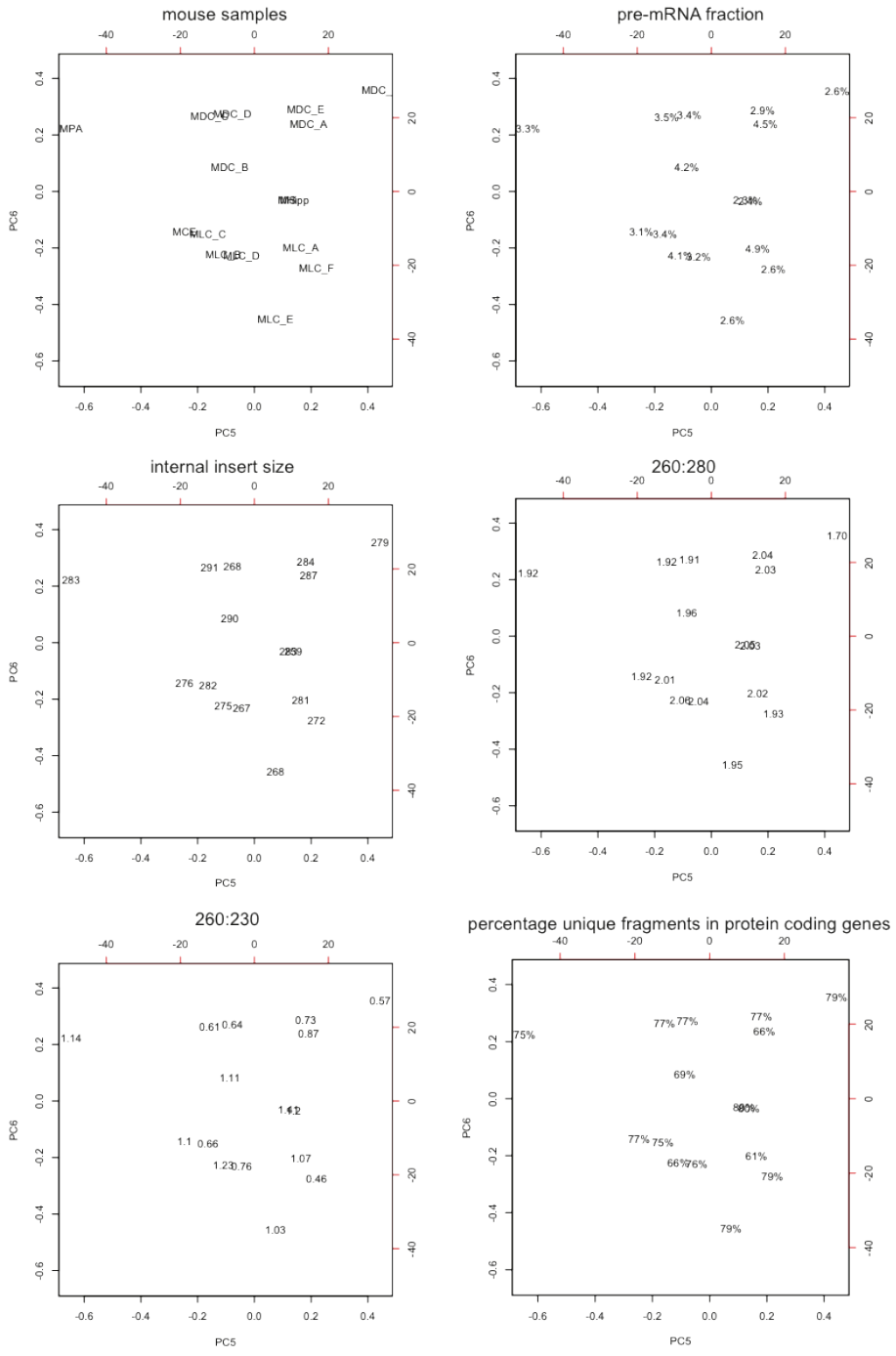


Fig. S31: Technical variables plotted on the fifth and sixth mouse principal components.

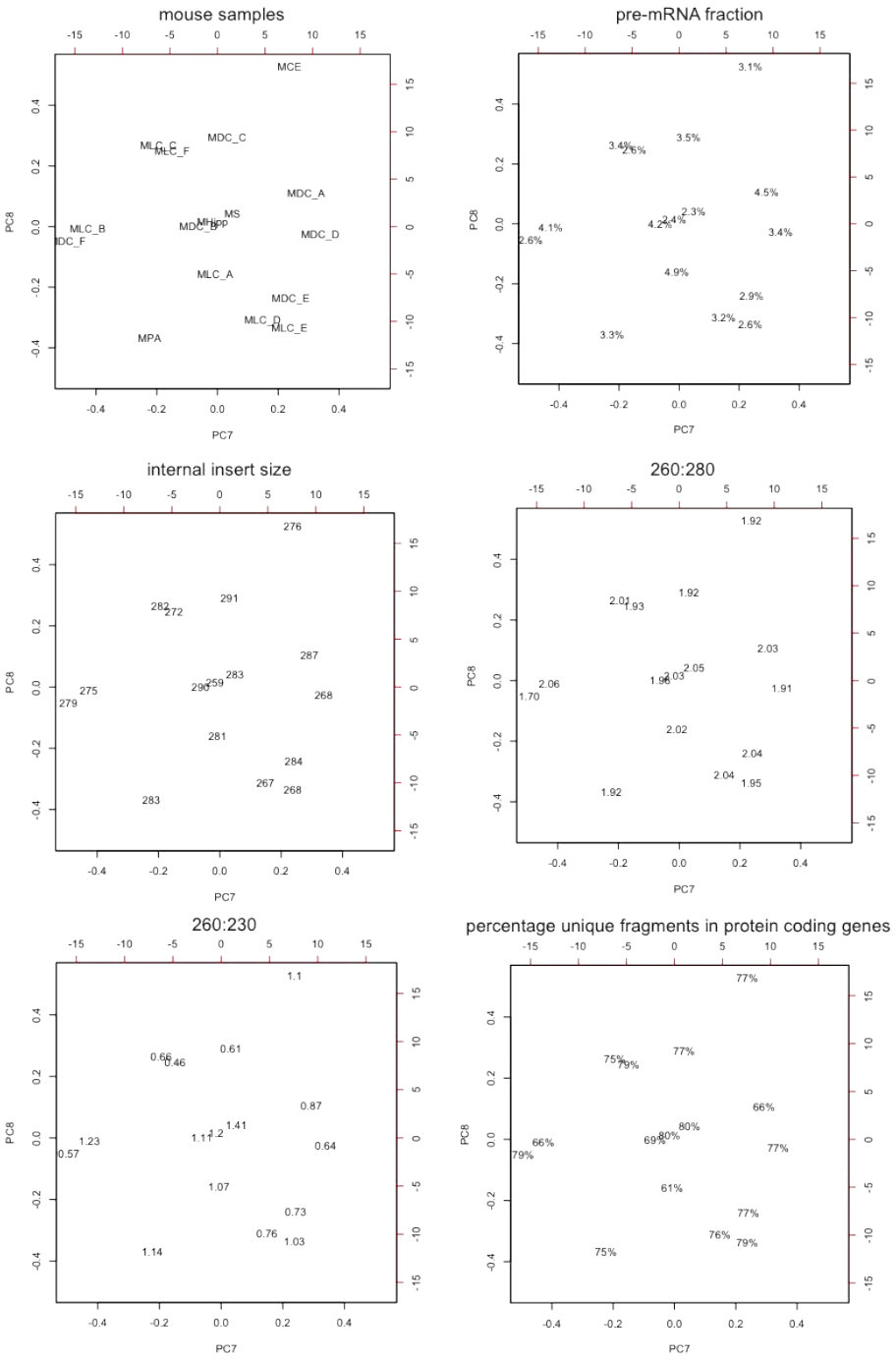


Fig. S32: Technical variables plotted on the seventh and eighth mouse principal components.

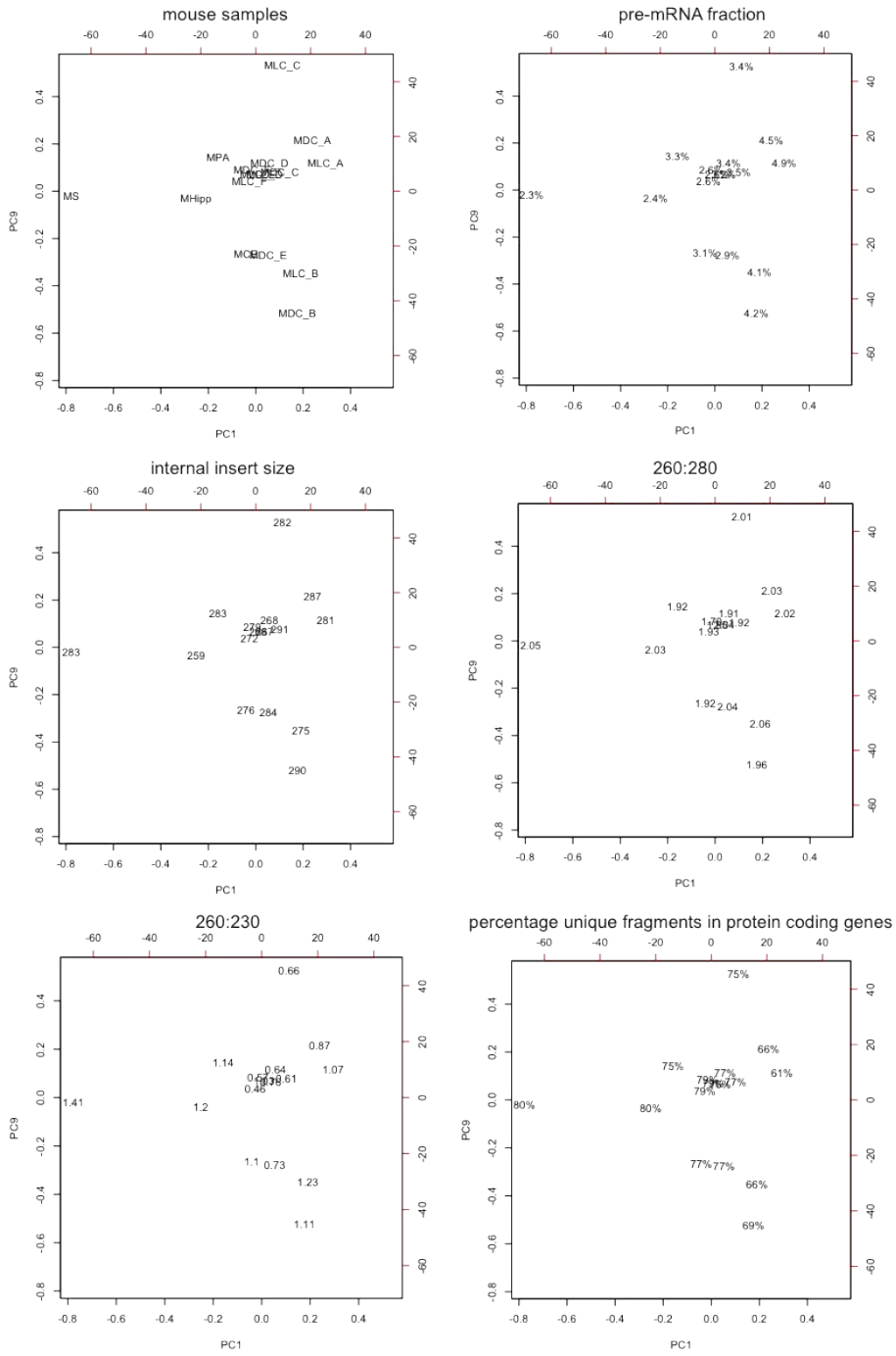


Fig. S33: Technical variables plotted on the first and ninth mouse principal components.

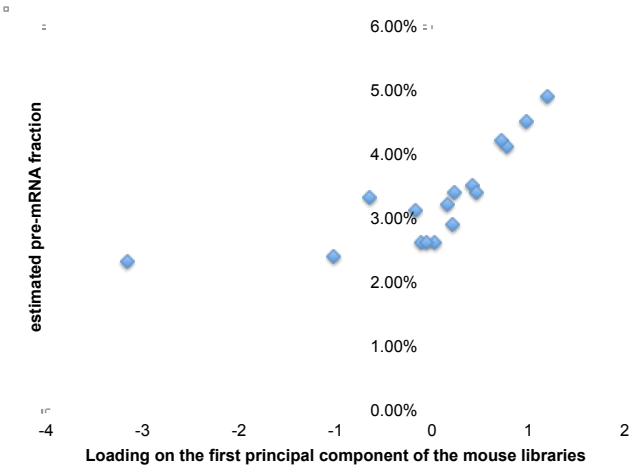


Fig. S34: The first principal component of the mouse libraries was correlated (r = 0.73; two-tailed p = 0.0014) with the pre-mRNA fraction.

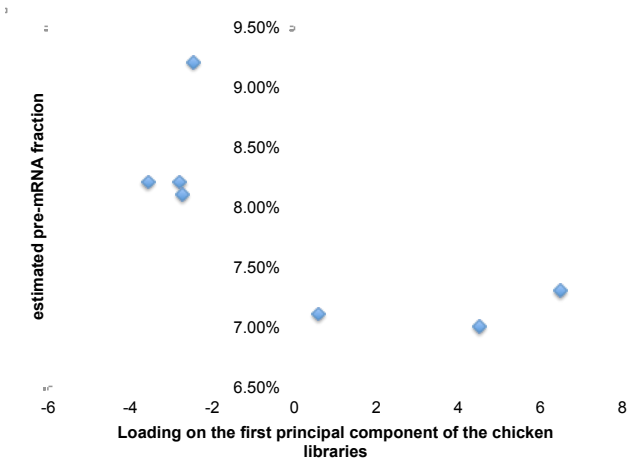


Fig. S35: The first principal component of the chicken libraries may also be correlated (r = -0.74; two-tailed p = 0.056) with the pre-mRNA fraction.

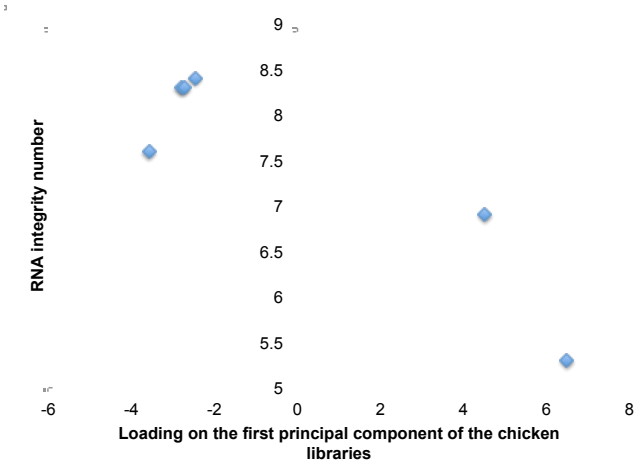


Fig. S36: The first principal component of the chicken libraries was correlated (r = -0.90; two-tailed p = 0.013) with RNA Integrity Number.

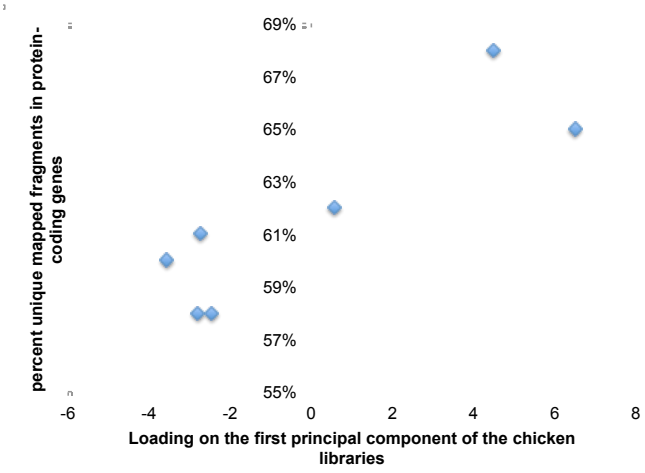


Fig. S37: The first principal component of the chicken libraries was correlated (r = 0.87; two-tailed p = 0.011) with the percentage of uniquely mapped fragments in protein-coding genes.

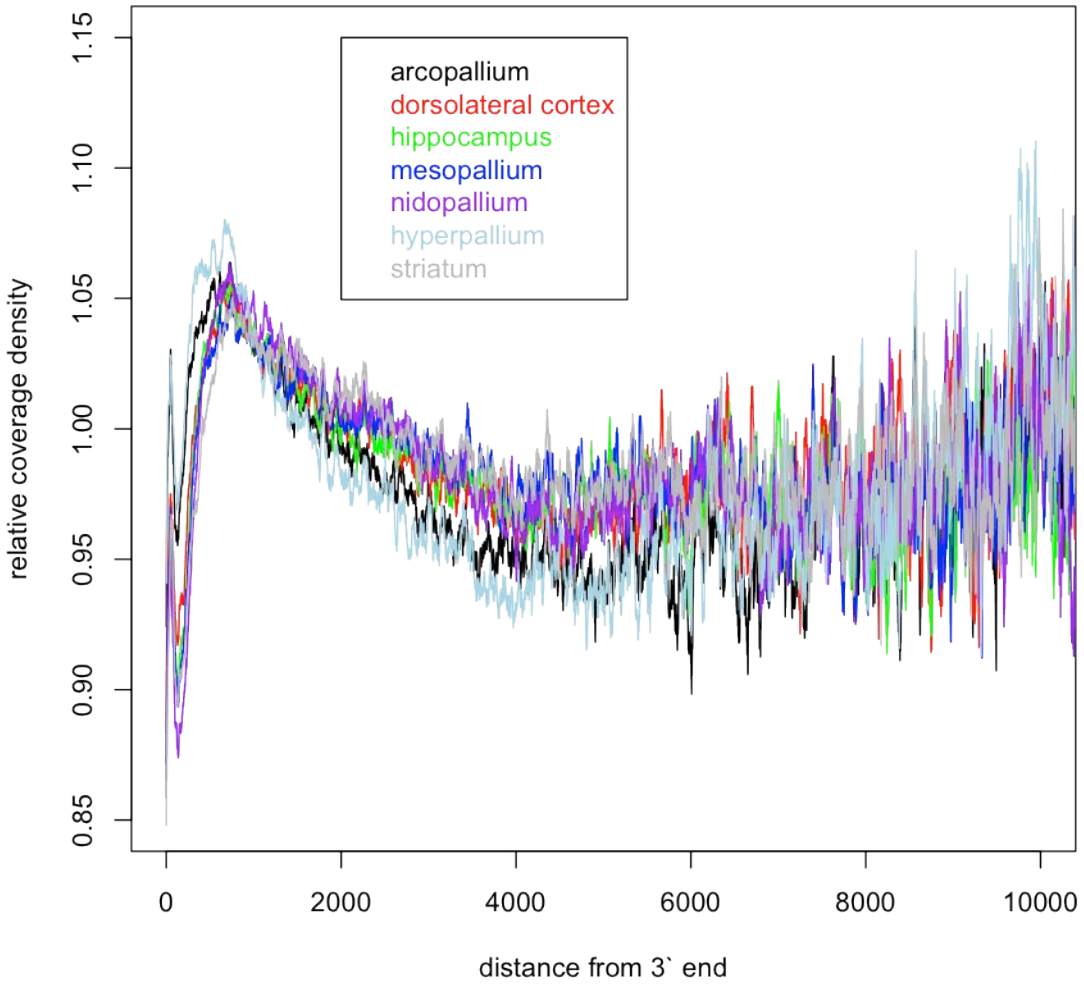


Fig. S38: Relative read coverage density in major chicken transcripts with respect to distance from the 3´ end.

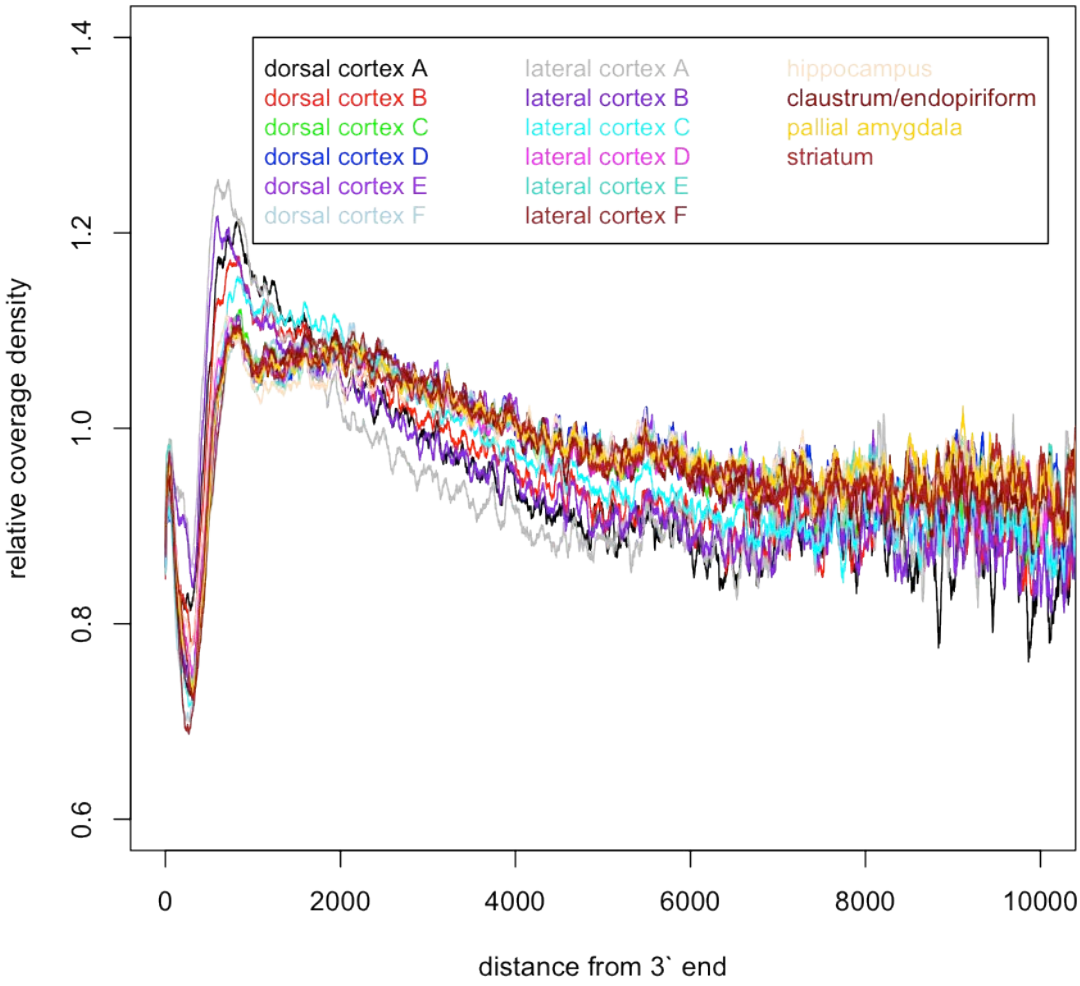


Fig. S39: Relative read coverage density in major mouse transcripts with respect to distance from the 3´ end.

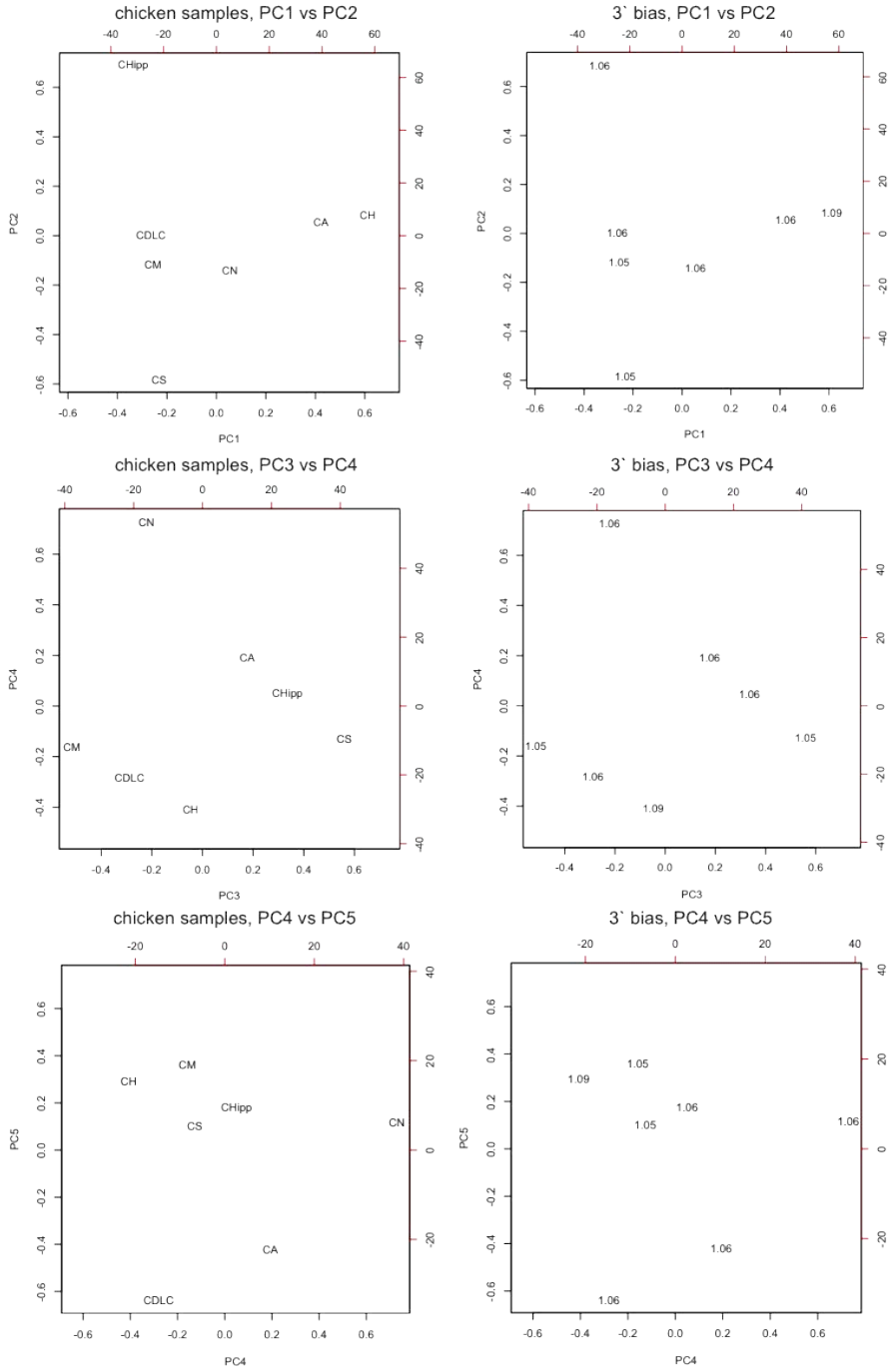


Fig. S40: Measures of 3´ bias of chicken samples plotted on chicken principal components.

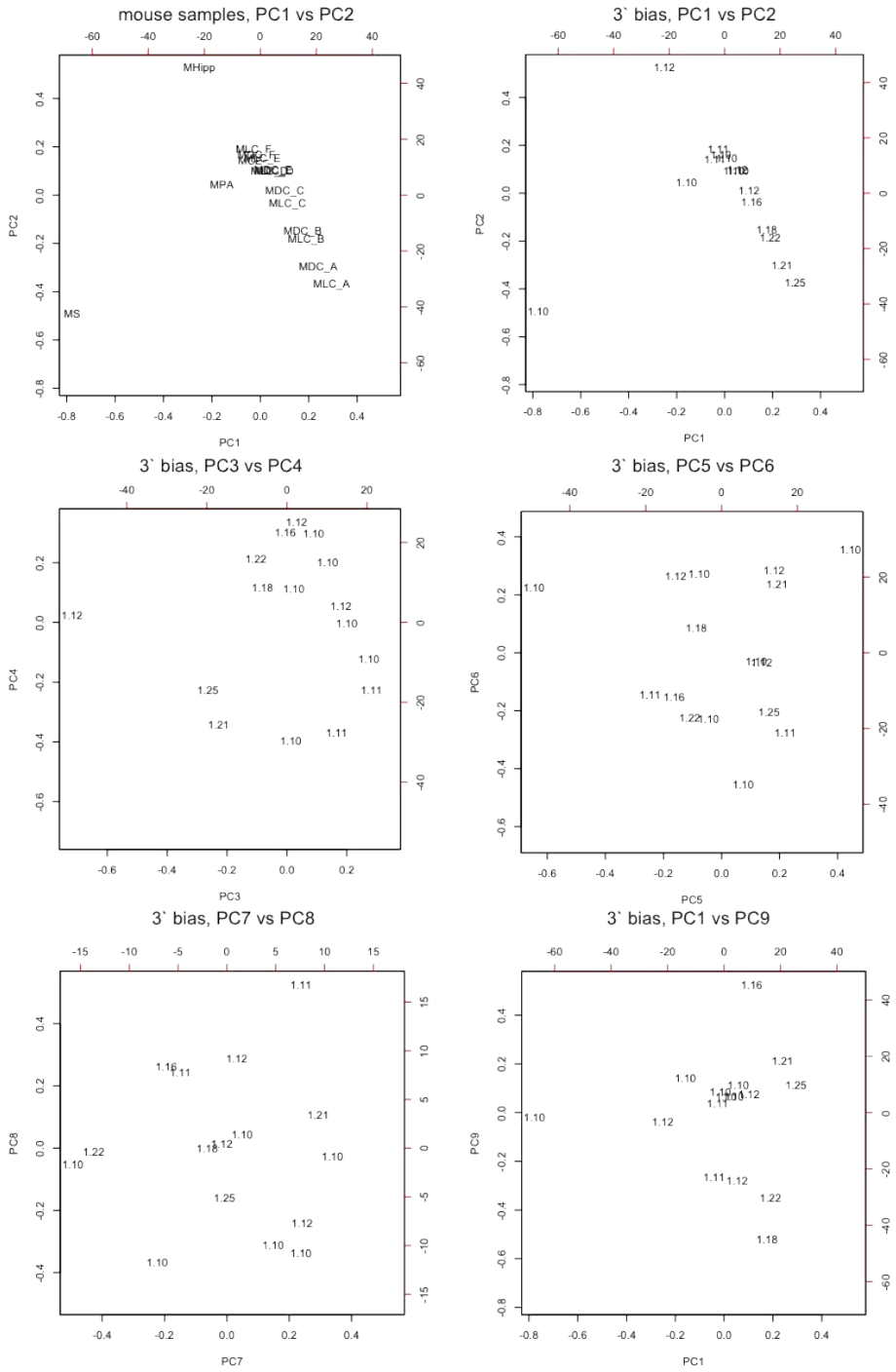


Fig. S41: Measures of 3´ bias of mouse samples plotted on mouse principal components.

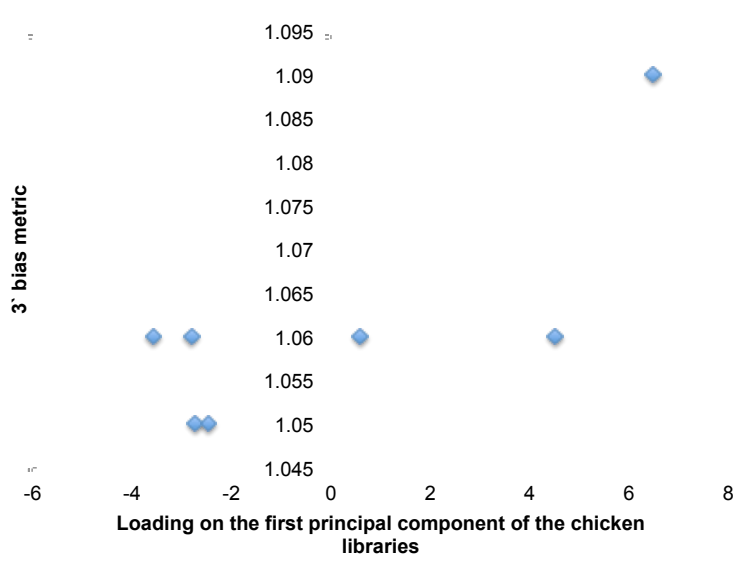


Fig. S42: The first principal component of the chicken libraries was correlated (r = 0.76; two-tailed p = 0.047) with a metric of 3′ bias.

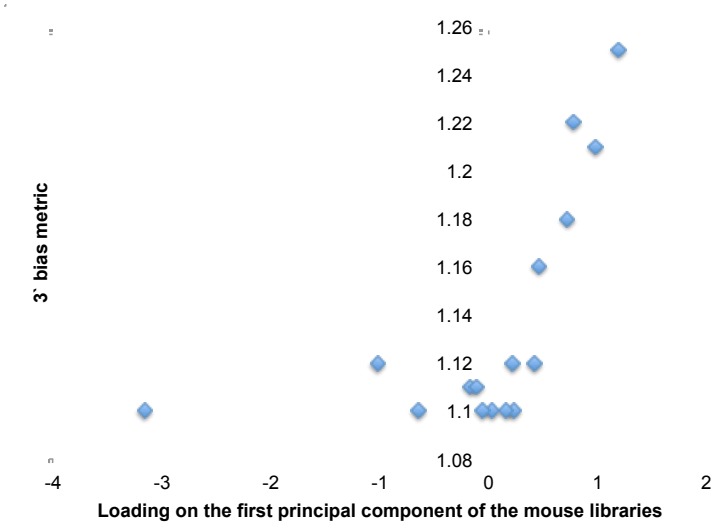


Fig. S43: The first principal component of the mouse libraries was correlated (r = 0.59; two-tailed p = 0.016) with a metric of 3′ bias.

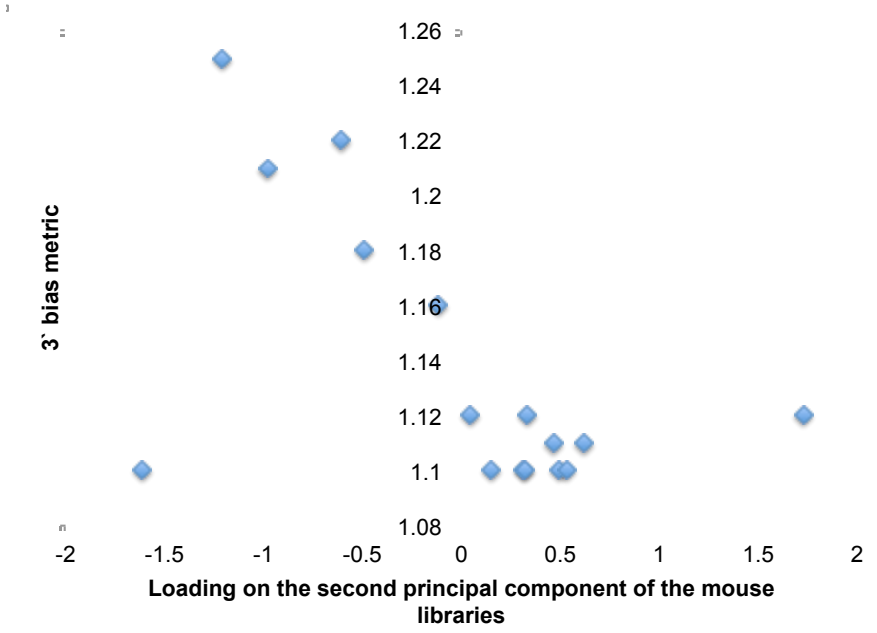


Fig. S44: The second principal component of the mouse libraries was correlated (r = -0.57; two-tailed p = 0.020) with a metric of 3´ bias.

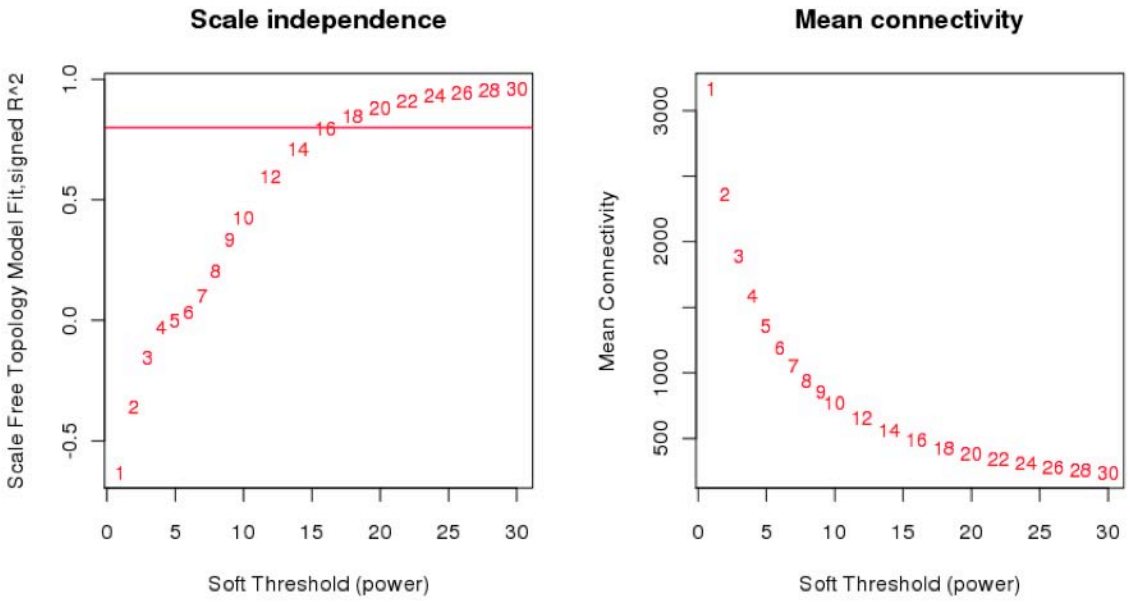


Fig. S45: Plots showing the relationships between chicken soft threshold and both fit to a scale free topology model and mean connectivity. The horizontal line on the left indicates an R^2 of 0.8.

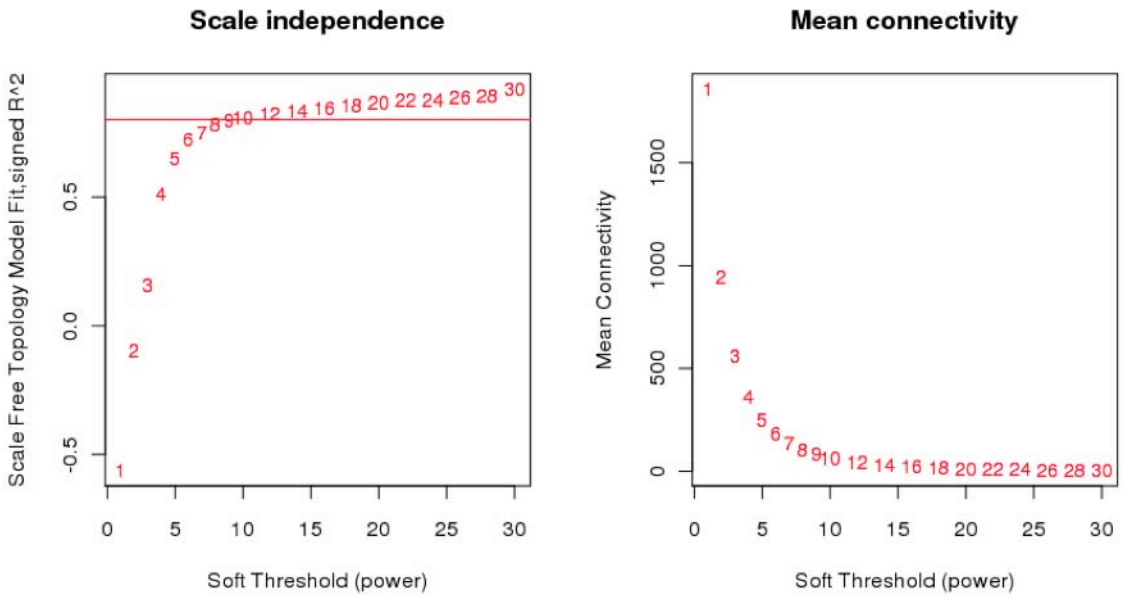


Fig. S46: Plots showing the relationships between mouse soft threshold and both fit to a scale free topology model and mean connectivity. The horizontal line on the left indicates an R^2 of 0.8.

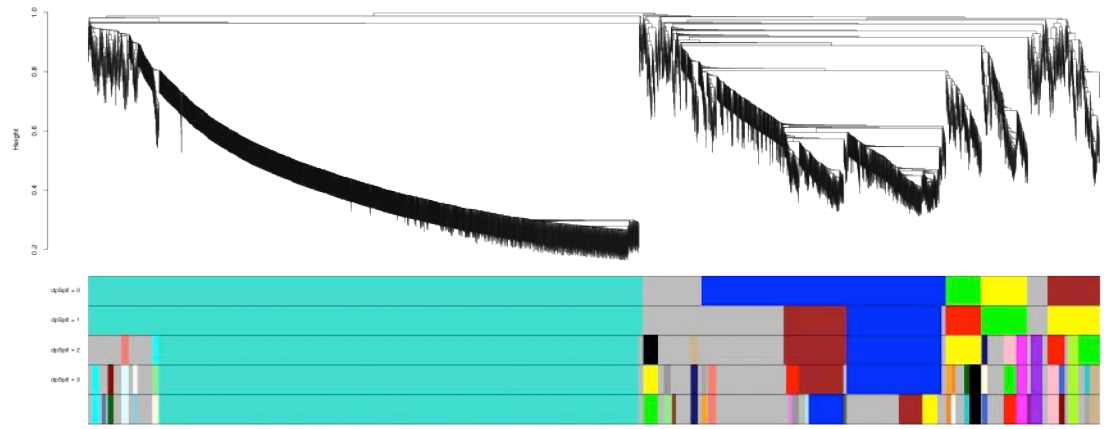


Fig. S47: The chicken dendrogram is shown at top with the module choices for different splitting parameters at bottom.

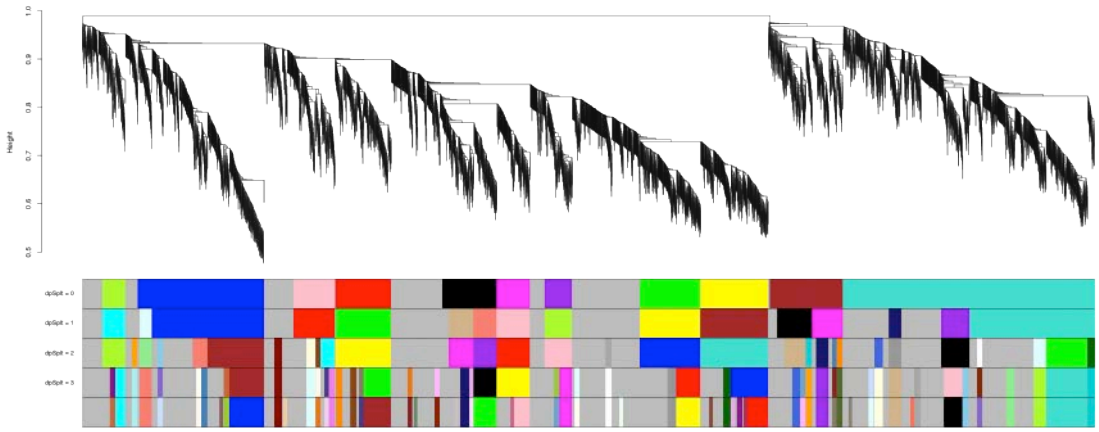


Fig. S48: The mouse dendrogram is shown at top with the module choices for different splitting parameters at bottom.

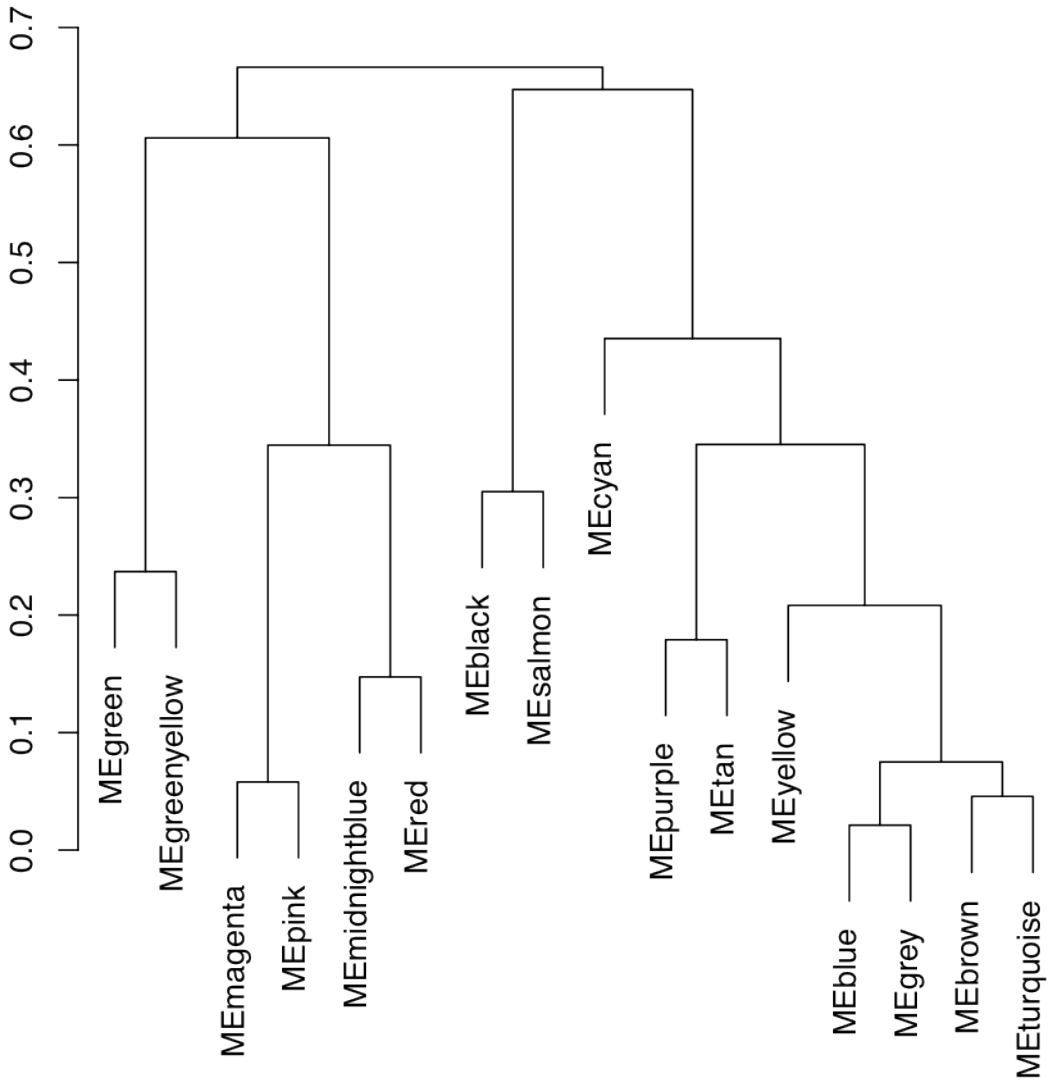


Fig. S49: Hierarchical clustering of the chicken module eigengenes using the hclust R function (agglomerating on the average).

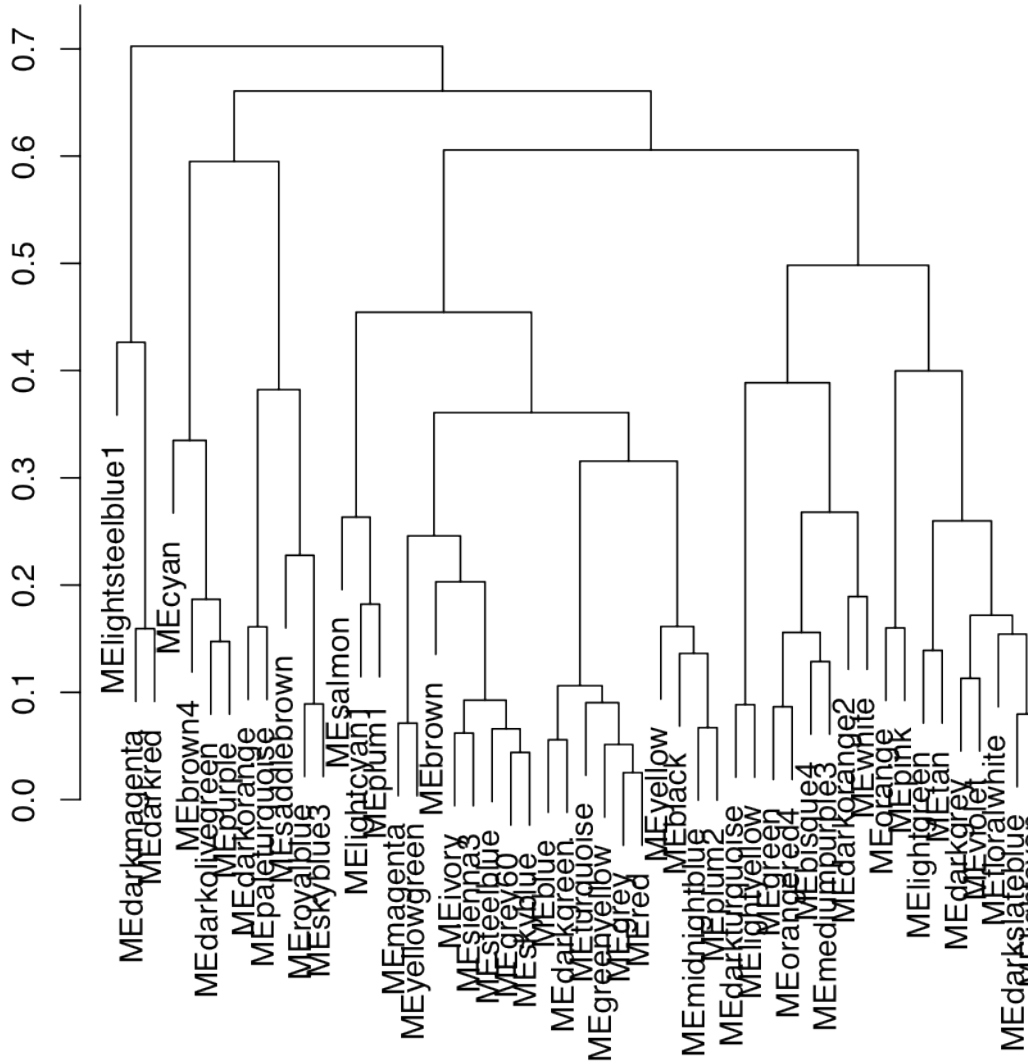


Fig. S50: Hierarchical clustering of the mouse module eigengenes using the hclust R function (agglomerating on the average).

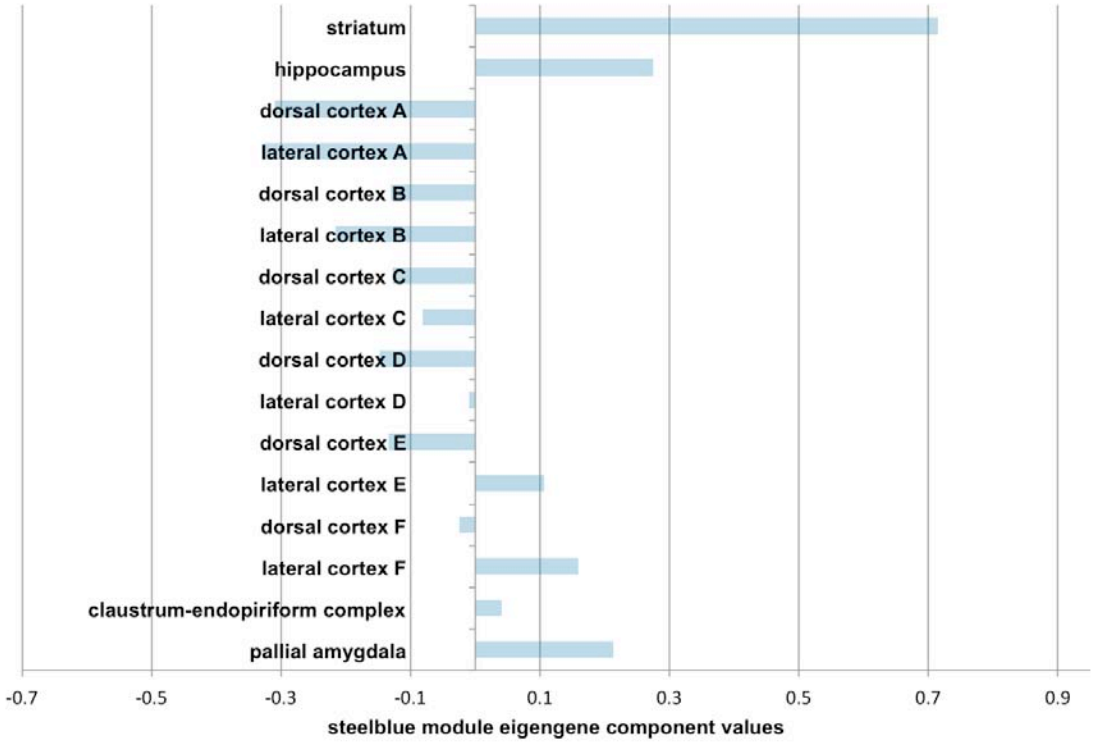
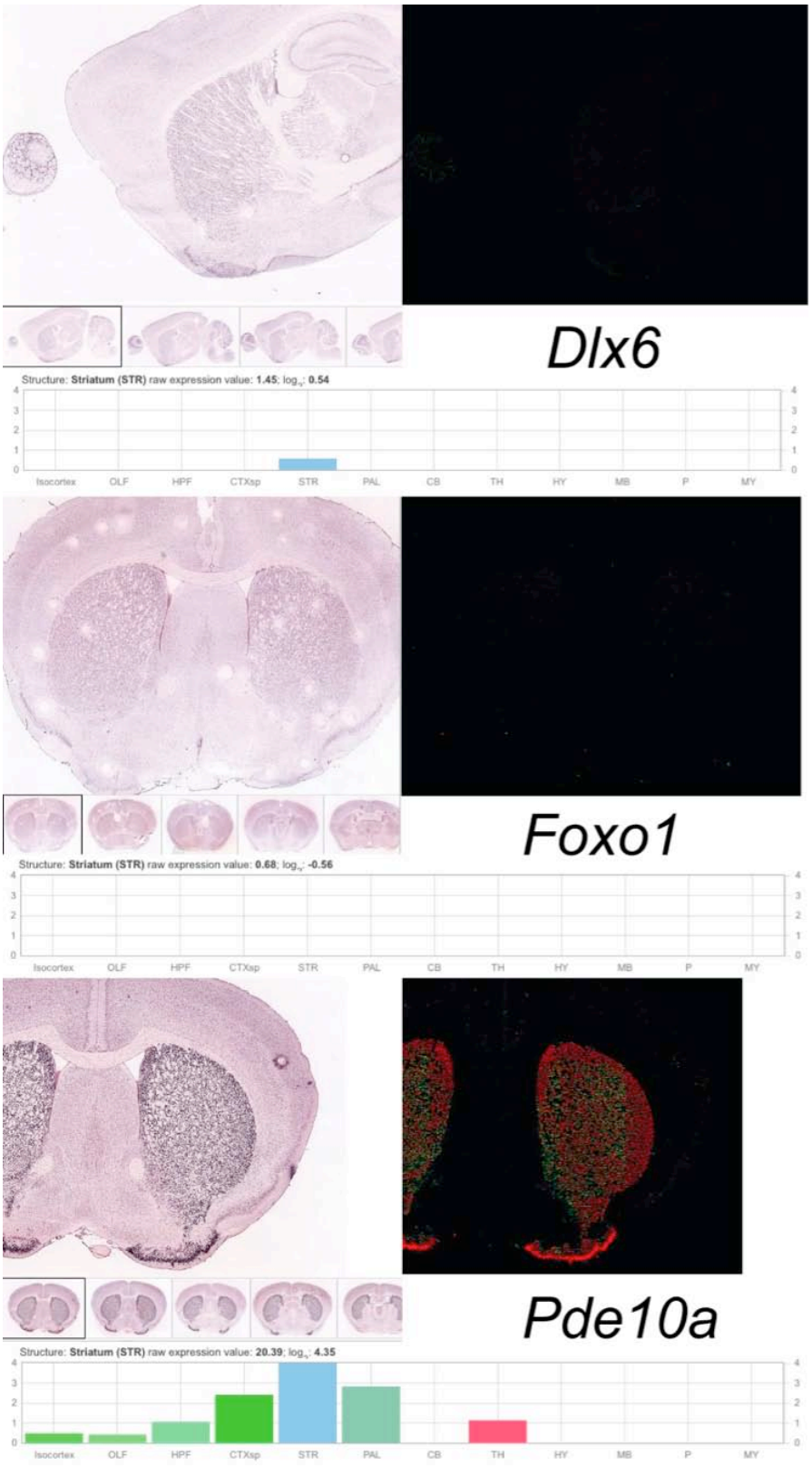
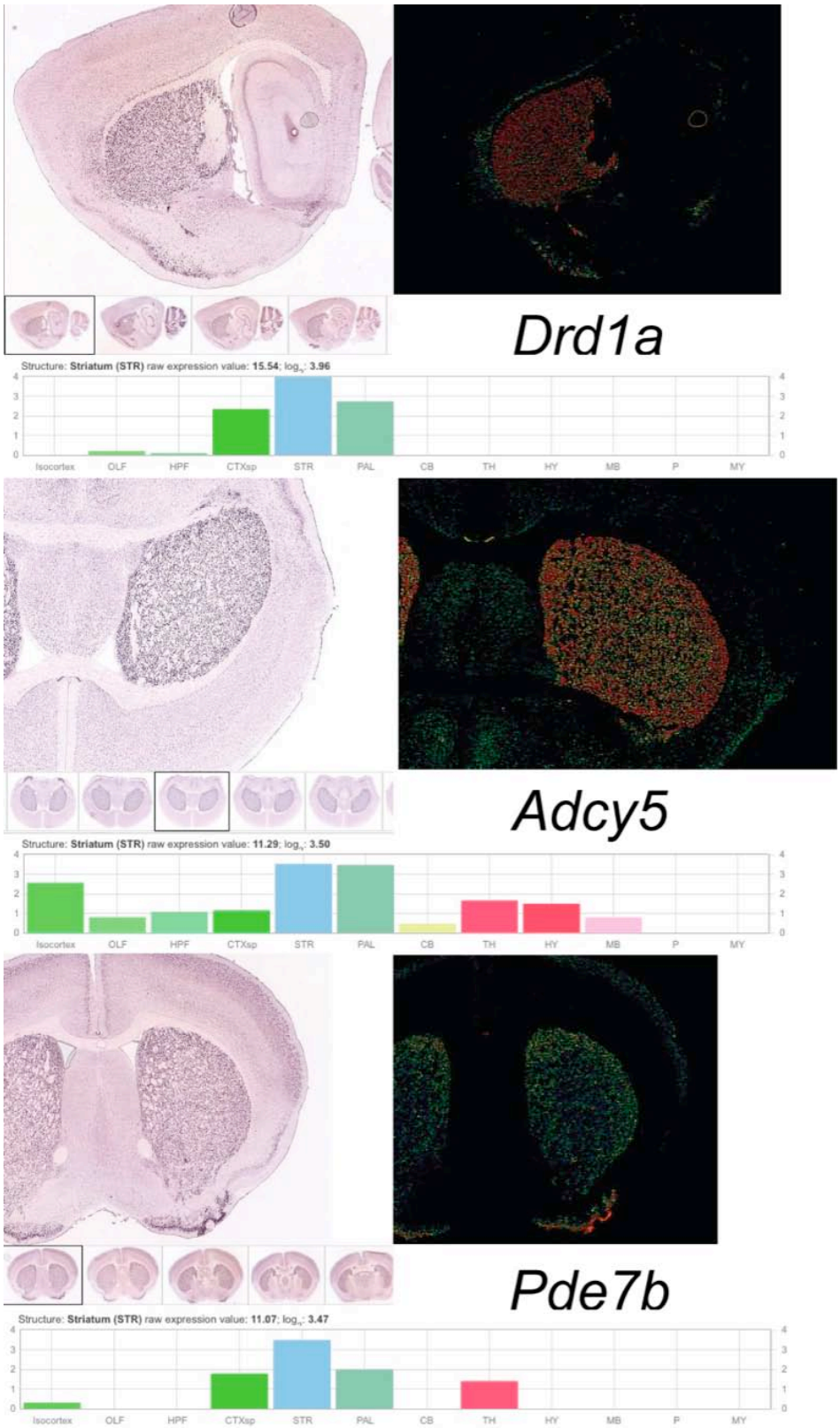


Fig. S51: The steelblue mouse module eigengene is highly expressed in striatum. The x-axis represents the values of the components of the module eigengene.


Dlx6
Structure: Striatum (STR) raw expression value: 1.45; log2: 0.54
Isocortex OLF HPF CTXsp STR PAL CB TH HY MB P MY
Foxo1
Structure: Striatum (STR) raw expression value: 0.68; log2: -0.56
Isocortex OLF HPF CTXsp STR PAL CB TH HY MB P MY
Pde10a
Structure: Striatum (STR) raw expression value: 20.39; log2: 4.35
Isocortex OLF HPF CTXsp STR PAL CB TH HY MB P MY


Drd1a
Structure: Striatum (STR) raw expression value: 15.54; log2: 3.96
Isocortex OLF HPF CTXsp STR PAL CB TH HY MB P MY
Adcy5
Structure: Striatum (STR) raw expression value: 11.29; log2: 3.50
Isocortex OLF HPF CTXsp STR PAL CB TH HY MB P MY
Pde7b
Structure: Striatum (STR) raw expression value: 11.07; log2: 3.47
Isocortex OLF HPF CTXsp STR PAL CB TH HY MB P MY

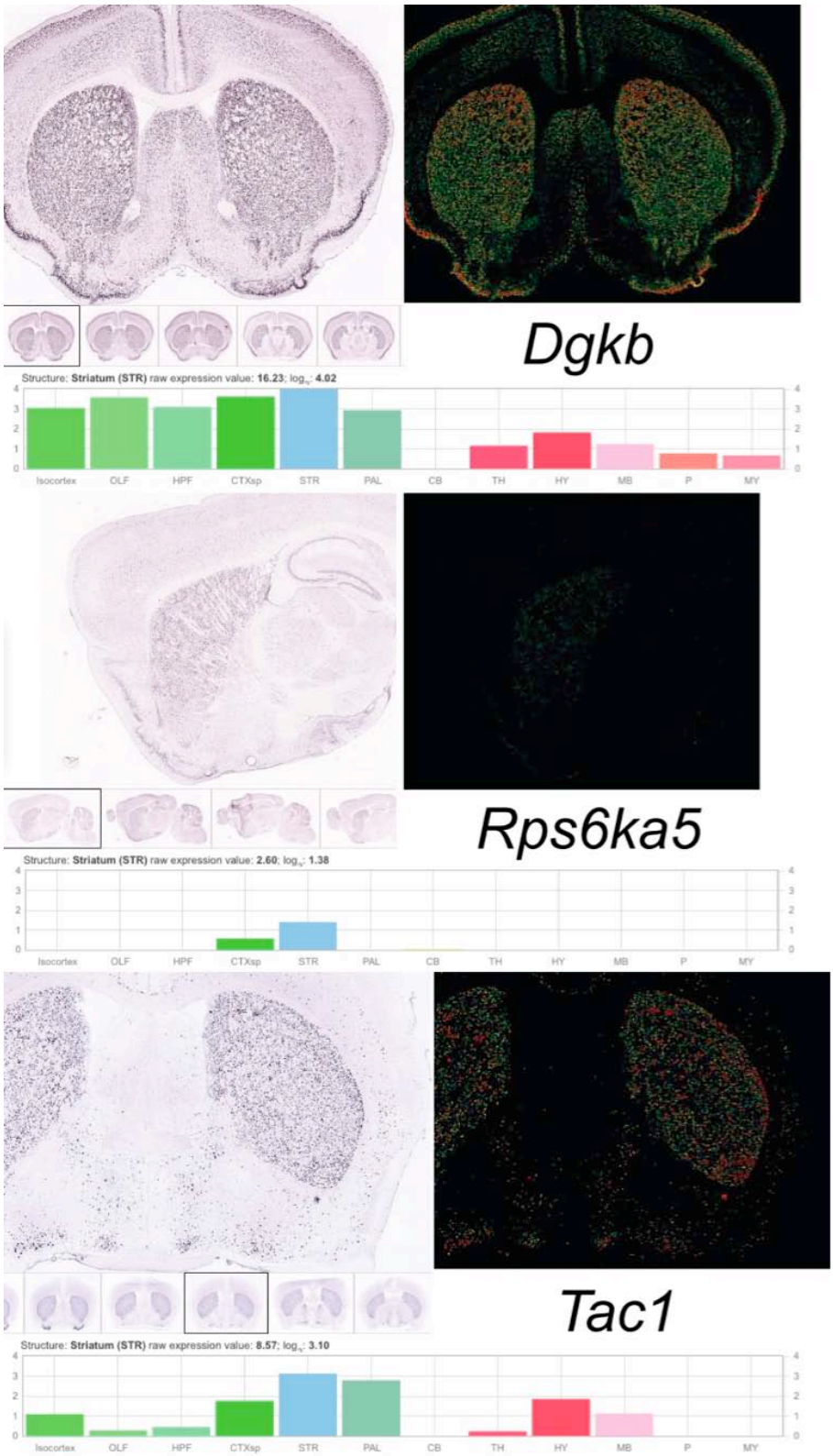
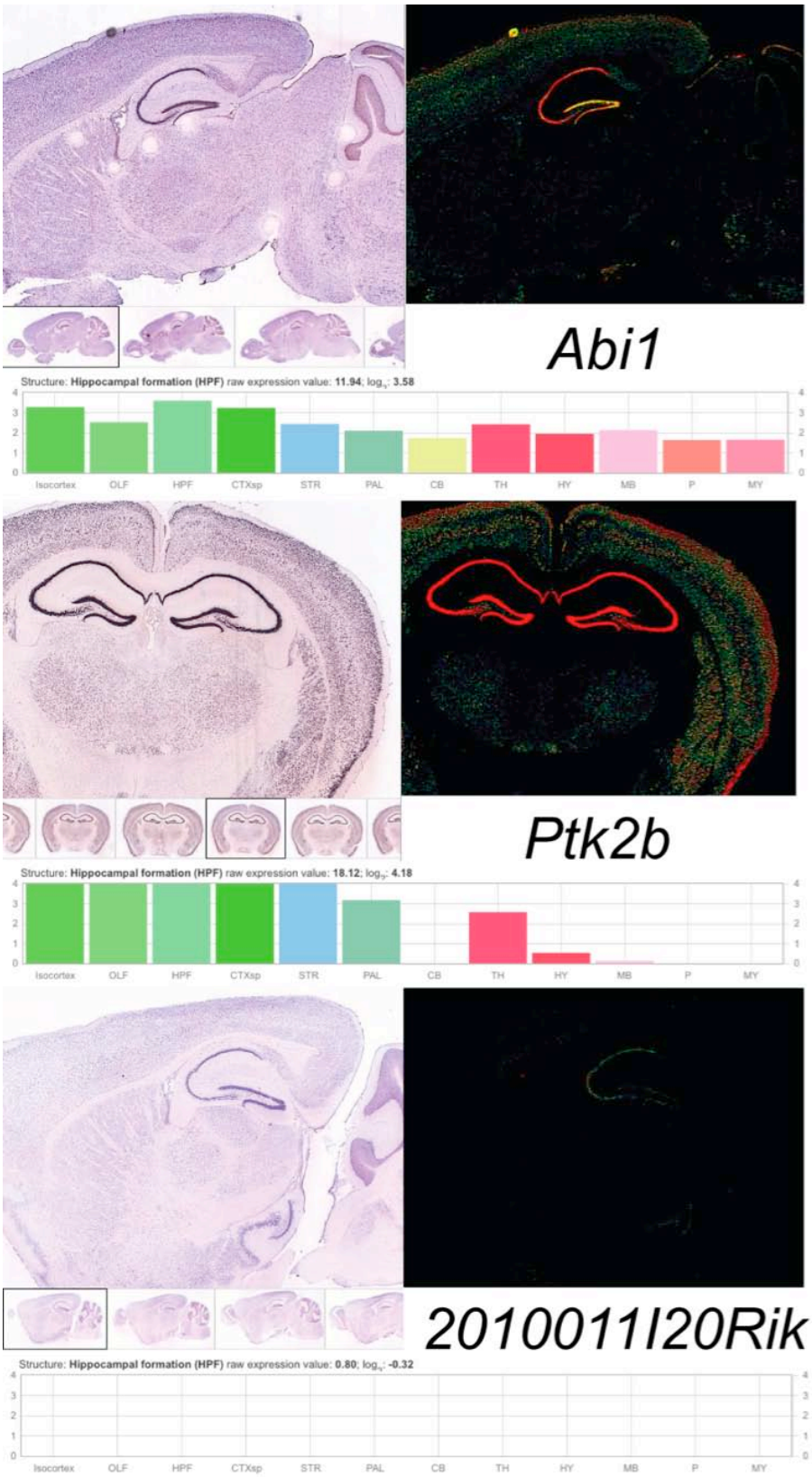


Fig. S52: *In situ* hybridizations, expression analyses of the hybridizations, and automated expression analyses across several major regions (STR corresponds to

striatum) from the Allen Mouse Brain Atlas (8) showing the nine unique genes from the top five in either the chick green/mouse brown overlap or the chick greenyellow/mouse brown overlap.


Abi1
Structure: Hippocampal formation (HPF) raw expression value: 11.94; log2: 3.58
Isocortex OLF HPF CTXsp STR PAL CB TH HY MB P MY
Ptk2b
Structure: Hippocampal formation (HPF) raw expression value: 18.12; log2: 4.18
Isocortex OLF HPF CTXsp STR PAL CB TH HY MB P MY
2010011I20Rik
Structure: Hippocampal formation (HPF) raw expression value: 0.80; log2: -0.32
Isocortex OLF HPF CTXsp STR PAL CB TH HY MB P MY

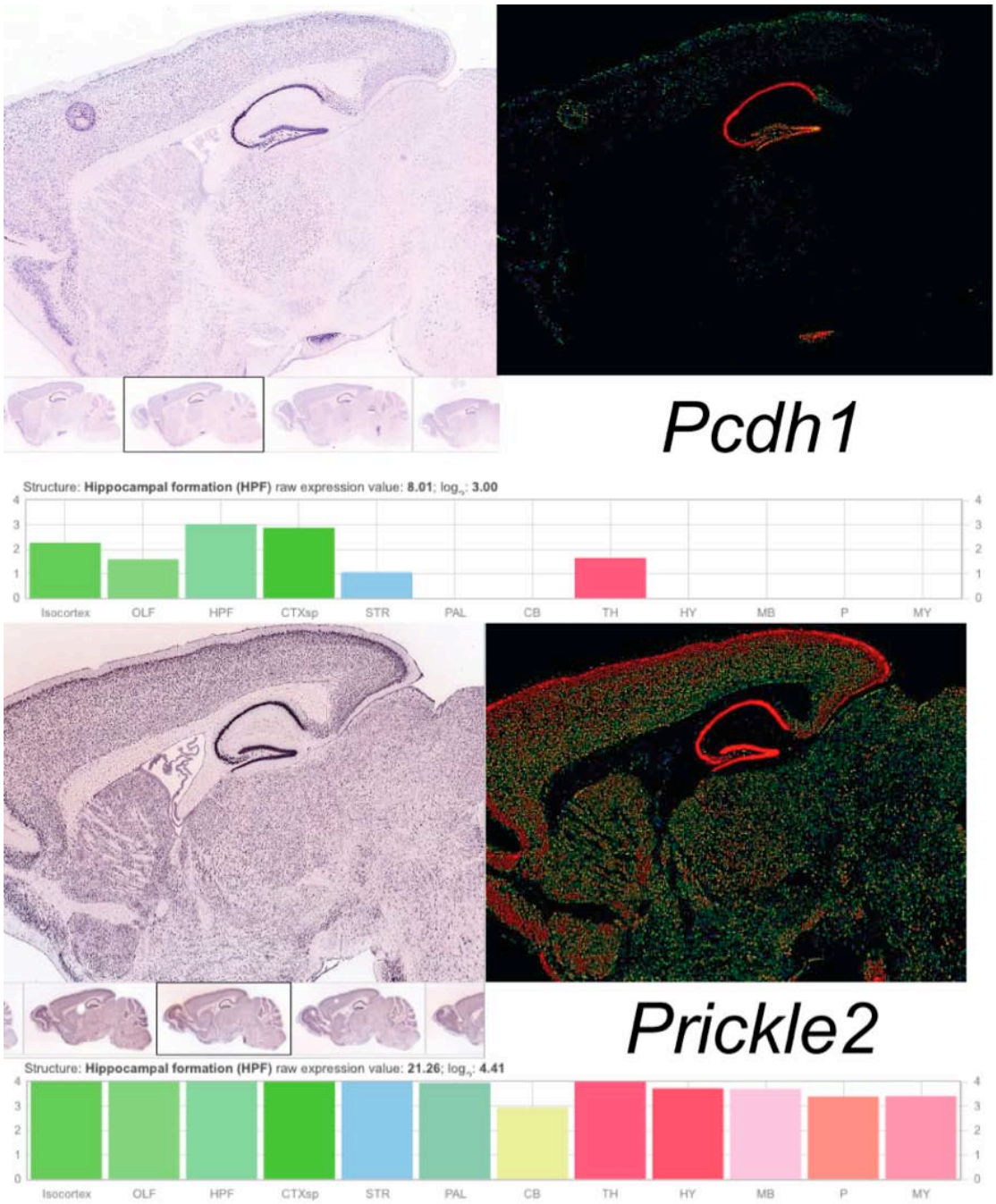


Fig. S53: *In situ* hybridizations, expression analyses of the hybridizations, and automated expression analyses across several major regions (HPF corresponds to the hippocampal formation) from the Allen Mouse Brain Atlas (8) showing the top five genes in the chick magenta/mouse black overlap.

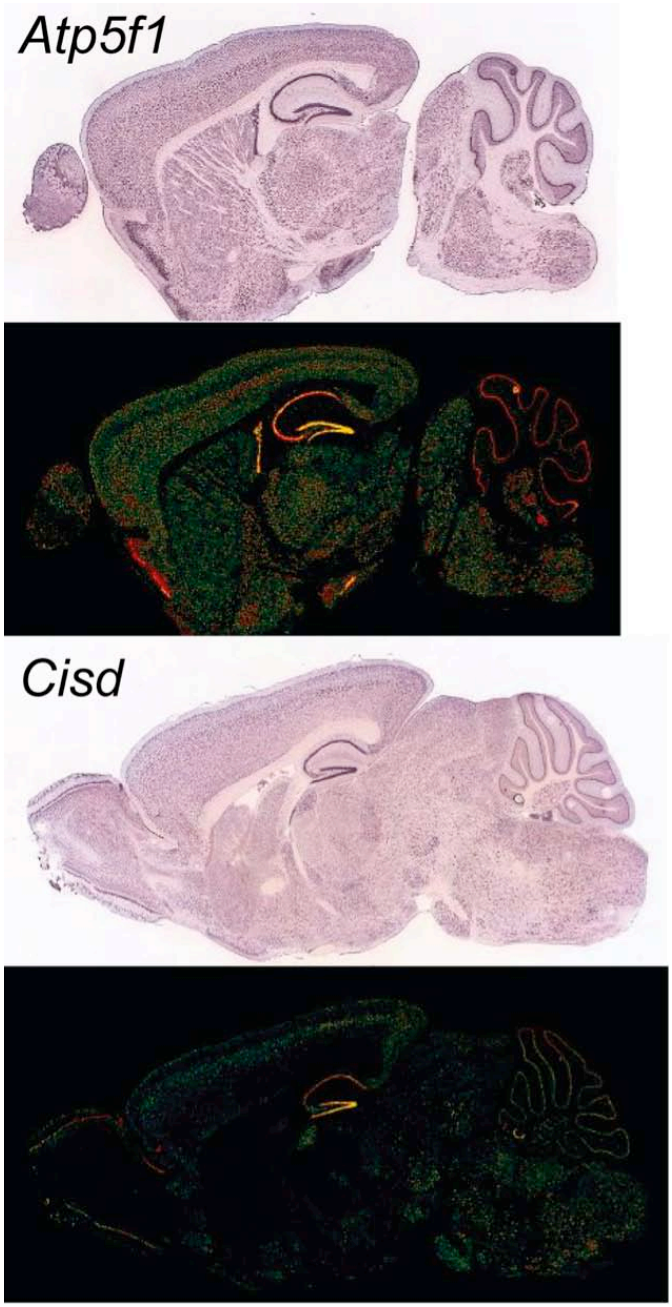


Fig. S54: Two genes correlated with both chick turquoise and mouse salmon – *Atp5f1* and *Cisd* – were differentially expressed across different brain regions in the Allen Mouse Brain Atlas (8).

Dataset S1. Excel spreadsheet of the module assignments in the original chicken WGCNA, module assignments in the original mouse WGCNA, module assignments in chicken reanalysis WGCNA (without the low-RIN hyperpallium & arcopallium samples) and module assignments in mouse reanalysis WGCNA.

SI References

1. Belgard TG et al. (2011) A transcriptomic atlas of mouse neocortical layers. *Neuron* 71:605–616.

2. Puelles L, Martinez-de-la-Torre M, Paxinos G, Watson C, Martinez S (2007) *The Chick Brain in Stereotaxic Coordinates: An Atlas featuring Neuromeric Subdivisions and Mammalian Homologies* (Academic Press). 1st Ed.

3. Trapnell C, Pachter L, Salzberg SL (2009) TopHat: discovering splice junctions with RNA-Seq. *Bioinformatics* 25:1105–1111.

4. Trapnell C et al. (2010) Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. *Nat Biotechnol* 28:511–515.

5. Leek JT et al. (2010) Tackling the widespread and critical impact of batch effects in high-throughput data. *Nat Rev Genet* 11:733–739.

6. Marioni JC, Mason CE, Mane SM, Stephens M, Gilad Y (2008) RNA-seq: an assessment of technical reproducibility and comparison with gene expression arrays. *Genome Res* 18:1509–1517.

7. Johnson WE, Li C, Rabinovic A (2007) Adjusting batch effects in microarray expression data using empirical Bayes methods. *Biostatistics* 8:118–127.

8. Lein ES et al. (2007) Genome-wide atlas of gene expression in the adult mouse brain. *Nature* 445:168–176.

9. Wang WZ et al. (2011) Comparative aspects of subplate zone studied with gene expression in sauropsids and mammals. *Cereb Cortex* 21:2187–2203.

10. Flicek P et al. (2011) Ensembl 2011. *Nucleic Acids Res* 39:D800–6.

11. Langfelder P, Horvath S (2008) WGCNA: an R package for weighted correlation network analysis. *BMC Bioinformatics* 9:559.

12. Zhang B, Horvath S (2005) A general framework for weighted gene co-expression network analysis. *Stat Appl Genet Mol Biol* 4:Article17.

13. Horvath S et al. (2006) Analysis of oncogenic signaling networks in glioblastoma identifies ASPM as a molecular target. *Proc Natl Acad Sci USA* 103:17402–17407.

14. Langfelder P, Zhang B, Horvath S (2008) Defining clusters from a hierarchical cluster tree: the Dynamic Tree Cut package for R. *Bioinformatics* 24:719–720.

15. Langfelder P, Horvath S (2007) Eigengene networks for studying the relationships between co-expression modules. *BMC Syst Biol* 1:54.

16. Langfelder P, Luo R, Oldham MC, Horvath S (2011) Is my network module preserved and reproducible? *PLoS Comput Biol* 7:e1001057.

17. Horvath S, Dong J (2008) Geometric interpretation of gene coexpression network analysis. *PLoS Comput Biol* 4:e1000117.

18. Birling MC, Tait S, Hardy RJ, Brophy PJ (1999) A novel rat tetraspan protein in cells of the oligodendrocyte lineage. *J Neurochem* 73:2600–2608.

19. Boggs JM (2006) Myelin basic protein: a multifunctional protein. *Cell Mol Life Sci* 63:1945–1961.

20. Cahoy JD et al. (2008) A transcriptome database for astrocytes, neurons, and oligodendrocytes: a new resource for understanding brain development and function. *J Neurosci* 28:264–278.

21. Wiley SE, Murphy AN, Ross SA, van der Geer P, Dixon JE (2007) MitoNEET is an iron-containing outer mitochondrial membrane protein that regulates oxidative capacity. *Proc Natl Acad Sci USA* 104:5318–5323.

22. Higuti T et al. (1991) Molecular cloning of cDNA for the import precursor of human subunit B of H(+)-ATP synthase in mitochondria. *Biochem Biophys Res Commun* 178:1014–1020.

23. Pagliarini DJ et al. (2008) A mitochondrial protein compendium elucidates complex I disease biology. *Cell* 134:112–123.

24. Umeda M, Nishitani H, Nishimoto T (2003) A novel nuclear protein, Twa1, and Muskelin comprise a complex with RanBPM. *Gene* 303:47–54.

25. Kobayashi N et al. (2007) RanBPM, Muskelin, p48EMLP, p44CTLH, and the armadillo-repeat proteins ARMC8alpha and ARMC8beta are components of the CTLH complex. *Gene* 396:236–247.

26. Wennerberg K, Rossman KL, Der CJ (2005) The Ras superfamily at a glance. *J Cell Sci* 118:843–846.

27. Nielsen JV, Thomassen M, Møllgård K, Noraberg J, Jensen NA (2013) Zbtb20 Defines a Hippocampal Neuronal Identity Through Direct Repression of Genes That Control Projection Neuron Development in the Isocortex. *Cereb Cortex*.

28. Riano E et al. (2009) Pleiotropic effects of spastin on neurite growth depending

on expression levels. *J Neurochem* 108:1277–1288.

29. Ito-Ishida A et al. (2012) Presynaptically released Cbln1 induces dynamic axonal structural changes by interacting with GluD2 during cerebellar synapse formation. *Neuron* 76:549–564.

30. Saito S-Y, Takeshima H (2006) DNER as key molecule for cerebellar maturation. *Cerebellum* 5:227–231.

31. Barber M et al. (2009) The role of Robo3 in the development of cortical interneurons. *Cereb Cortex* 19 Suppl 1:i22–31.

32. Sabbir MG et al. (2010) Identification and characterization of Dlc1 isoforms in the mouse and study of the biological function of a single gene trapped isoform. *BMC Biol* 8:17.

33. Worby CA, Dixon JE (2002) Sorting out the cellular functions of sorting nexins. *Nat Rev Mol Cell Biol* 3:919–931.

34. Homma MK, Homma Y (2008) Cell cycle and activation of CK2. *Mol Cell Biochem* 316:49–55.

35. Asaki C, Usuda N, Nakazawa A, Kametani K, Suzuki T (2003) Localization of translational components at the ultramicroscopic level at postsynaptic sites of the rat brain. *Brain Res* 972:168–176.

36. Falluel-Morel A et al. (2008) Interactions of PACAP and ceramides in the control of granule cell apoptosis during cerebellar development. *J Mol Neurosci* 36:8–15.

37. Guo W, Jiang H, Gray V, Dedhar S, Rao Y (2007) Role of the integrin-linked kinase (ILK) in determining neuronal polarity. *Dev Biol* 306:457–468.

38. Naska S et al. (2006) An essential role for the integrin-linked kinase-glycogen synthase kinase-3 beta pathway during dendrite initiation and growth. *J Neurosci* 26:13344–13356.

39. Gragnoli C (2012) Proteasome modulator 9 and depression in type 2 diabetes. *Curr Med Chem* 19:5178–5180.

40. Guo S et al. (2000) A regulator of transcriptional elongation controls vertebrate neuronal development. *Nature* 408:366–369.

41. Aslam A, Mittal S, Koch F, Andrau J-C, Winkler GS (2009) The Ccr4-NOT deadenylase subunits CNOT7 and CNOT8 have overlapping roles and modulate cell proliferation. *Mol Biol Cell* 20:3840–3850.

42. Brenman JE (2007) AMPK/LKB1 signaling in epithelial cell polarity and cell division. *Cell Cycle* 6:2755–2759.

43. Gurumurthy S et al. (2008) LKB1 deficiency sensitizes mice to carcinogen-

induced tumorigenesis. *Cancer Res* 68:55–63.

44. Lo B et al. (2012) Lkb1 regulates organogenesis and early oncogenesis along AMPK-dependent and -independent pathways. *J Cell Biol* 199:1117–1130.

45. Cantagrel V et al. (2010) SRD5A3 is required for converting polyprenol to dolichol and is mutated in a congenital glycosylation disorder. *Cell* 142:203–217.

46. Morava E et al. (2010) A novel cerebello-ocular syndrome with abnormal glycosylation due to abnormalities in dolichol metabolism. *Brain* 133:3210–3220.

47. Simpson MA et al. (2004) Infantile-onset symptomatic epilepsy syndrome caused by a homozygous loss-of-function mutation of GM3 synthase. *Nat Genet* 36:1225–1229.

48. Takemoto M et al. (2011) Laminar and areal expression of unc5d and its role in cortical cell survival. *Cereb Cortex* 21:1925–1934.

49. Anderson RM, Lawrence AR, Stottmann RW, Bachiller D, Klingensmith J (2002) Chordin and noggin promote organizing centers of forebrain development in the mouse. *Development* 129:4975–4987.

50. Adams SH et al. (2001) BFIT, a unique acyl-CoA thioesterase induced in thermogenic brown adipose tissue: cloning, organization of the human gene and assessment of a potential link to obesity. *Biochem J* 360:135–142.

51. Sanz-Moreno V et al. (2008) Rac activation and inactivation control plasticity of tumor cell movement. *Cell* 135:510–523.

52. Claesson-Welsh L, Welsh M (2013) VEGFA and tumour angiogenesis. *J Intern Med* 273:114–127.

53. Raimondi L, D'Asaro M, Proia P, Nastasi T, Di Liegro I (2003) RNA-binding ability of PIPPin requires the entire protein. *J Cell Mol Med* 7:35–42.

54. Wu J et al. (2011) Biochemical characterization of human SET and MYND domain-containing protein 2 methyltransferase. *Biochemistry* 50:6488–6497.

55. Nomura K, Takeuchi Y, Yamaguchi S, Okamura H, Fukunaga K (2003) Involvement of calcium/calmodulin-dependent protein kinase II in the induction of mPer1. *J Neurosci Res* 72:384–392.

56. Mercure MZ, Ginnan R, Singer HA (2008) CaM kinase II delta2-dependent regulation of vascular smooth muscle cell polarization and migration. *Am J Physiol Cell Physiol* 294:C1465–75.

57. Rijkers K et al. (2010) Polymorphisms in CACNA1E and Camk2d are associated with seizure susceptibility of Sprague-Dawley rats. *Epilepsy Res* 91:28–34.

58. North HA et al. (2009) Promotion of proliferation in the developing cerebral cortex by EphA4 forward signaling. *Development* 136:2467–2476.

59. Wieland I et al. (2004) Mutations of the ephrin-B1 gene cause craniofrontonasal syndrome. *Am J Hum Genet* 74:1209–1215.

60. Hagiyama M, Ichiyanagi N, Kimura KB, Murakami Y, Ito A (2009) Expression of a soluble isoform of cell adhesion molecule 1 in the brain and its involvement in directional neurite outgrowth. *Am J Pathol* 174:2278–2289.

61. Fujita E, Tanabe Y, Imhof BA, Momoi MY, Momoi T (2012) Cadm1-expressing synapses on Purkinje cell dendrites are involved in mouse ultrasonic vocalization activity. *PLoS ONE* 7:e30151.

62. Hagiyama M et al. (2011) Enhanced nerve-mast cell interaction by a neuronal short isoform of cell adhesion molecule-1. *J Immunol* 186:5983–5992.

63. Li X, Rosahl TW, Südhof TC, Francke U (1995) Mapping of synapsin II (SYN2) genes to human chromosome 3p and mouse chromosome 6 band F. *Cytogenet Cell Genet* 71:301–305.

64. Greco B et al. (2012) Autism-related behavioral abnormalities in synapsin knockout mice. *Behav Brain Res* pii: S0166-4328(12)00797-8.

65. Velho TAF, Mello CV (2008) Synapsins are late activity-induced genes regulated by birdsong. *J Neurosci* 28:11871–11882.

66. Spitz F, Gonzalez F, Duboule D (2003) A global control region defines a chromosomal regulatory landscape containing the HoxD cluster. *Cell* 113:405–417.

67. Peichel CL, Prabhakaran B, Vogt TF (1997) The mouse Ulnaless mutation deregulates posterior HoxD gene expression and alters appendicular patterning. *Development* 124:3481–3492.

68. Lakshmana MK et al. (2012) Role of RanBP9 on amyloidogenic processing of APP and synaptic protein levels in the mouse brain. *FASEB J* 26:2072–2083.

69. Dantas TJ, Wang Y, Lalor P, Dockery P, Morrison CG (2011) Defective nucleotide excision repair with normal centrosome structures and functions in the absence of all vertebrate centrins. *The Journal of Cell Biology* 193:307–318.

70. Fleischer TC, Weaver CM, McAfee KJ, Jennings JL, Link AJ (2006) Systematic identification and functional screens of uncharacterized proteins associated with eukaryotic ribosomal complexes. *Genes Dev* 20:1294–1307.

71. Zhang X et al. (2009) SUN1/2 and Syne/Nesprin-1/2 complexes connect centrosome to the nucleus during neurogenesis and neuronal migration in mice. *Neuron* 64:173–187.

72. Gros-Louis F et al. (2007) Mutations in SYNE1 lead to a newly discovered form

of autosomal recessive cerebellar ataxia. *Nat Genet* 39:80–85.

73. Nakamura T et al. (1999) A neurally enriched coronin-like protein, ClipinC, is a novel candidate for an actin cytoskeleton-cortical membrane-linking protein. *J Biol Chem* 274:13322–13327.

74. Carlson GC et al. (2011) Dysbindin-1 mutant mice implicate reduced fast-phasic inhibition as a final common disease mechanism in schizophrenia. *Proc Natl Acad Sci USA* 108:E962–70.

75. Ohno T et al. (2010) Specific involvement of postsynaptic GluN2B-containing NMDA receptors in the developmental elimination of corticospinal synapses. *Proc Natl Acad Sci USA* 107:15252–15257.

76. Endele S et al. (2010) Mutations in GRIN2A and GRIN2B encoding regulatory subunits of NMDA receptors cause variable neurodevelopmental phenotypes. *Nat Genet* 42:1021–1026.

77. Ayalew M et al. (2012) Convergent functional genomics of schizophrenia: from comprehensive understanding to genetic risk prediction. *Mol Psychiatry* 17:887–905.

78. Tarabeux J et al. (2011) Rare mutations in N-methyl-D-aspartate glutamate receptors in autism spectrum disorders and schizophrenia. *Transl Psychiatry* 1:e55.

79. Lee JH et al. (2012) Evolutionarily assembled cis-regulatory module at a human ciliopathy locus. *Science* 335:966–969.

80. Parisi M, Glass MD (1993) in eds Pagon RA, Bird TD, Dolan CR, Stephens K, Adam MP (GeneReviews).

81. Lai-Cheong JE, Parsons M, McGrath JA (2010) The role of kindlins in cell biology and relevance to human disease. *Int J Biochem Cell Biol* 42:595–603.

82. Poh K-W, Yeo J-F, Stohler CS, Ong W-Y (2012) Comprehensive gene expression profiling in the prefrontal cortex links immune activation and neutrophil infiltration to antinociception. *J Neuroscience* 32:35–45.

83. Li X et al. (2012) MEK Is a Key Regulator of Gliogenesis in the Developing Brain. *Neuron* 75:1035–1050.

84. Allanson JE et al. (2011) Cardio-facio-cutaneous syndrome: does genotype predict phenotype? *Am J Med Genet C Semin Med Genet* 157:129–135.

85. Majores M et al. (2002) Cathepsin D: screening for new polymorphisms using single-strand conformation polymorphism analysis. *Int J Mol Med* 9:185–187.

86. Ishibashi K, Suzuki M, Imai M (2000) Molecular cloning of a novel form (two-repeat) protein related to voltage-gated sodium and calcium channels. *Biochem Biophys Res Commun* 270:370–376.

87. Yang JP, Tang H, Reddy TR, Wong-Staal F (2001) Mapping the functional domains of HAP95, a protein that binds RNA helicase A and activates the constitutive transport element of type D retroviruses. *J Biol Chem* 276:30694–30700.

88. Kanei-Ishii C, Nomura T, Egoh A, Ishii S (2012) Fbxw5 suppresses nuclear c-Myb activity via DDB1-Cul4-Rbx1 ligase-mediated sumoylation. *Biochem Biophys Res Commun* 426(1):59-64.

89. Murali R, Wiesner T, Rosenblum MK, Bastian BC (2012) GNAQ and GNA11 mutations in melanocytomas of the central nervous system. *Acta Neuropathol* 123:457–459.

90. Van Raamsdonk CD et al. (2010) Mutations in GNA11 in uveal melanoma. *N Engl J Med* 363:2191–2199.

91. Garrett AM, Schreiner D, Lobas MA, Weiner JA (2012) γ-protocadherins control cortical dendrite arborization by regulating the activity of a FAK/PKC/MARCKS signaling pathway. *Neuron* 74:269–276.

92. Winter Y, Sankowski R, Back T (2013) Genetic determinants of obesity and related vascular diseases. *Vitam Horm* 91:29–48.

93. Aryal RP, Ju T, Cummings RD (2012) Tight complex formation between Cosmc chaperone and its specific client non-native T-synthase leads to enzyme activity and client-driven dissociation. *J Biol Chem* 287:15317–15329.

94. Wang Y et al. (2010) Cosmc is an essential chaperone for correct protein O-glycosylation. *Proc Natl Acad Sci USA* 107:9228–9233.

95. Lenhard B, Sandelin A, Carninci P (2012) Metazoan promoters: emerging characteristics and insights into transcriptional regulation. *Nat Rev Genet* 13:233–245.

96. Yanagisawa K et al. (2010) Novel metastasis-related gene CIM functions in the regulation of multiple cellular stress-response pathways. *Cancer Res* 70:9949–9958.

97. Vogel P et al. (2010) Situs inversus in Dpcd/Poll-/-, Nme7-/- , and Pkd1l1-/- mice. *Vet Pathol* 47:120–131.

98. Tanaka H et al. (2012) Linkage of N-cadherin to multiple cytoskeletal elements revealed by a proteomic approach in hippocampal neurons. *Neurochem Int* 61:240–250.

99. Redies C, Medina L, Puelles L (2001) Cadherin expression by embryonic divisions and derived gray matter structures in the telencephalon of the chicken. *J Comp Neurol* 438:253–285.

100. Noguchi Y et al. (2009) Total expression and dual gene-regulatory mechanisms maintained in deletions and duplications of the Pcdha cluster. *J Biol Chem*

284:32002–32014.

101. Hou S-C et al. (2009) Ankrd17, an ubiquitously expressed ankyrin factor, is essential for the vascular integrity during embryogenesis. *FEBS Lett* 583:2765–2771.

102. Nishi K, Nishi A, Nagasawa T, Ui-Tei K (2013) Human TNRC6A is an Argonaute-navigator protein for microRNA-mediated gene silencing in the nucleus. *RNA* 19:17–35.

103. Dobbins SE et al. (2011) Allergy and glioma risk: test of association by genotype. *Int J Cancer* 128:1736–1740.

104. Alderson NL, Maldonado EN, Kern MJ, Bhat NR, Hama H (2006) FA2H-dependent fatty acid 2-hydroxylation in postnatal mouse brain. *J Lipid Res* 47:2772–2780.

105. Eckhardt M, Yaghootfam A, Fewou SN, Zöller I, Gieselmann V (2005) A mammalian fatty acid hydroxylase responsible for the formation of alpha-hydroxylated galactosylceramide in myelin. *Biochem J* 388:245–254.

106. Kruer MC et al. (2010) Defective FA2H leads to a novel form of neurodegeneration with brain iron accumulation (NBIA). *Ann Neurol* 68:611–618.

107. Dick KJ et al. (2010) Mutation of FA2H underlies a complicated form of hereditary spastic paraplegia (SPG35). *Hum Mutat* 31:E1251–60.

108. Sakai D, Dixon J, Dixon MJ, Trainor PA (2012) Mammalian neurogenesis requires Treacle-Plk1 for precise control of spindle orientation, mitotic progression, and maintenance of neural progenitor cells. *PLoS Genet* 8:e1002566.

109. Aoyama K et al. (2012) Increased neuronal glutathione and neuroprotection in GTRAP3-18-deficient mice. *Neurobiol Dis* 45:973–982.

110. Kobayashi H, Fukuda M (2012) Rab35 regulates Arf6 activity through centaurin-β2 (ACAP2) during neurite outgrowth. *J Cell Sci* 125:2235–2243.

111. Bithell A, Alberta J, Hornby F, Stiles CD, Williams BP (2003) Expression of the guanine nucleotide exchange factor, mr-gef, is regulated during the differentiation of specific subsets of telencephalic neurons. *Brain Res Dev Brain Res* 146:107–118.

112. Gimelli S et al. (2009) The tumor suppressor gene TRC8/RNF139 is disrupted by a constitutional balanced translocation t(8;22)(q24.13;q11.21) in a young girl with dysgerminoma. *Mol Cancer* 8:52.

113. Zunino R, Braschi E, Xu L, McBride HM (2009) Translocation of SenP5 from the nucleoli to the mitochondria modulates DRP1-dependent fission during mitosis. *J Biol Chem* 284:17783–17795.

114. Yun C et al. (2008) Nucleolar protein B23/nucleophosmin regulates the vertebrate SUMO pathway through SENP3 and SENP5 proteases. *J Cell Biol* 183:589–595.

115. Gil J, Bernard D, Martínez D, Beach D (2004) Polycomb CBX7 has a unifying role in cellular lifespan. *Nat Cell Biol* 6:67–72.

116. Morey L, Aloia L, Cozzuto L, Benitah SA, Di Croce L (2013) RYBP and Cbx7 define specific biological functions of polycomb complexes in mouse embryonic stem cells. *Cell Rep* 3:60–69.

117. Forzati F, Federico A, Pallante P, Fedele M, Fusco A (2012) Tumor suppressor activity of CBX7 in lung carcinogenesis. *Cell Cycle* 11:1888–1891.

118. Pöschl J, Grammel D, Dorostkar MM, Kretzschmar HA, Schüller U (2013) Constitutive activation of β-catenin in neural progenitors results in disrupted proliferation and migration of neurons within the central nervous system. *Dev Biol* 374:319–332.

119. Zhong X, Schneider TJ, Cabral DS, Donohoe TJ, Rothstein TL (2001) An alternatively spliced long form of Fas apoptosis inhibitory molecule (FAIM) with tissue-specific expression in the brain. *Mol Immunol* 38:65–72.

120. Sole C et al. (2004) The death receptor antagonist FAIM promotes neurite outgrowth by a mechanism that depends on ERK and NF-kapp B signaling. *J Cell Biol* 167:479–492.

121. Zhang Y, Li J, Partovian C, Sellke FW, Simons M (2003) Syndecan-4 modulates basic fibroblast growth factor 2 signaling in vivo. *Am J Physiol Heart Circ Physiol* 284:H2078–82.

122. Riazanski V et al. (2011) Presynaptic CLC-3 determines quantal size of inhibitory transmission in the hippocampus. *Nat Neurosci* 14:487–494.

123. Wei J, Ji H, Guo M, Qin Q (2012) Isolation and characterization of a thioredoxin domain-containing protein 12 from orange-spotted grouper, Epinephelus coioides. *Fish Shellfish Immunol* 33:667–673.

124. Yoon J-H, Her S, Kim M, Jang I-S, Park J (2012) The expression of damage-regulated autophagy modulator 2 (DRAM2) contributes to autophagy induction. *Mol Biol Rep* 39:1087–1093.

125. Zhou RH et al. (2001) Characterization of the human NDRG gene family: a newly identified member, NDRG4, is specifically expressed in brain and heart. *Genomics* 73:86–97.

126. Gurniak CB, Perlas E, Witke W (2005) The actin depolymerizing factor n-cofilin is essential for neural tube morphogenesis and neural crest cell migration. *Dev Biol* 278:231–241.

127. Rengaraj D et al. (2011) The distribution of neuron-specific gene family member 1 in brain and germ cells: Implications for the regulation of germ-line

development by brain. *Dev Dyn* 240:850–861.

128. Ohnishi S et al. (2010) Identification of NEEP21, encoding neuron-enriched endosomal protein of 21 kDa, as a transcriptional target of tumor suppressor p53. *Int J Oncol* 37:1133–1141.

129. Dardou D et al. (2011) Distribution of SV2C mRNA and protein expression in the mouse brain with a particular emphasis on the basal ganglia system. *Brain Res* 1367:130–145.

130. Zhong W-X et al. (2012) Lanthionine synthetase C-like protein 1 interacts with and inhibits cystathionine β-synthase: a target for neuronal antioxidant defense. *J Biol Chem* 287:34189–34201.

131. Irvin DK, Nakano I, Paucar A, Kornblum HI (2004) Patterns of Jagged1, Jagged2, Delta-like 1 and Delta-like 3 expression during late embryonic and postnatal brain development suggest multiple functional roles in progenitors and differentiated cells. *J Neurosci Res* 75:330–343.

132. Jiang R et al. (1998) Defects in limb, craniofacial, and thymic development in Jagged2 mutant mice. *Genes Dev* 12:1046–1057.

133. Guo Y et al. (2012) Mutation screening of the HTR2B gene in patients with Tourette syndrome. *Neurosci Lett* 526:150–153.

134. Dermol U, Janardan V, Tyagi R, Visweswariah SS, Podobnik M (2011) Unique utilization of a phosphoprotein phosphatase fold by a mammalian phosphodiesterase associated with WAGR syndrome. *J Mol Biol* 412:481–494.

135. Constantine R, Zhang H, Gerstner CD, Frederick JM, Baehr W (2012) Uncoordinated (UNC)119: coordinating the trafficking of myristoylated proteins. *Vision Res* 75:26–32.

136. Miller-Delaney SFC, Lieberam I, Murphy P, Mitchell KJ (2011) Plxdc2 is a mitogen for neural progenitors. *PLoS ONE* 6:e14565.

137. Etchegaray J-P et al. (2009) Casein kinase 1 delta regulates the pace of the mammalian circadian clock. *Mol Cell Biol* 29:3853–3866.

138. Chien Y-L et al. (2013) DRD2 haplotype associated with negative symptoms and sustained attention deficits in Han Chinese with schizophrenia in Taiwan. *J Hum Genet* doi: 10.1038/jhg.2012.157.

139. Pierret P, Dunn RJ, Djordjevic B, Stone JC, Richardson PM (2000) Distribution of ras guanyl releasing protein (RasGRP) mRNA in the adult rat central nervous system. *J Neurocytol* 29:485–497.

140. Tu W et al. (2010) DAPK1 interaction with NMDA receptor NR2B subunits mediates brain damage in stroke. *Cell* 140:222–234.

141. Hao M et al. (2008) Muscleblind-like 2 (Mbnl2) -deficient mice as a model for

myotonic dystrophy. *Dev Dyn* 237:403–410.

142. Sarkissian M, Winne A, Lafyatis R (1996) The mammalian homolog of suppressor-of-white-apricot regulates alternative mRNA splicing of CD45 exon 4 and fibronectin IIICS. *J Biol Chem* 271:31106–31114.

143. Wang J et al. (2004) Tau exon 10, whose missplicing causes frontotemporal dementia, is regulated by an intricate interplay of cis elements and trans factors. *J Neurochem* 88:1078–1090.

144. Ryu M-J et al. (2012) Oncogenic Kras expression in postmitotic neurons leads to S100A8-S100A9 protein overexpression and gliosis. *J Biol Chem* 287:22948–22958.

145. Peng Y-J et al. (2010) Trio is a key guanine nucleotide exchange factor coordinating regulation of the migration and morphogenesis of granule cells in the developing cerebellum. *J Biol Chem* 285:24834–24844.

146. Osório J, Mueller T, Rétaux S, Vernier P, Wullimann MF (2010) Phylotypic expression of the bHLH genes Neurogenin2, Neurod, and Mash1 in the mouse embryonic forebrain. *J Comp Neurol* 518:851–871.

147. Tanackovic G et al. (2011) A missense mutation in PRPF6 causes impairment of pre-mRNA splicing and autosomal-dominant retinitis pigmentosa. *Am J Hum Genet* 88:643–649.

148. McCormick JA, Ellison DH (2011) The WNKs: atypical protein kinases with pleiotropic actions. *Physiol Rev* 91:177–219.

149. Pedrola L et al. (2008) Cell expression of GDAP1 in the nervous system and pathogenesis of Charcot-Marie-Tooth type 4A disease. *J Cell Mol Med* 12:679–689.

150. Cuesta A et al. (2002) The gene encoding ganglioside-induced differentiation-associated protein 1 is mutated in axonal Charcot-Marie-Tooth type 4A disease. *Nat Genet* 30:22–25.

151. Li B, Zhang S, Li M, Hertz L, Peng L (2009) Chronic treatment of astrocytes with therapeutically relevant fluoxetine concentrations enhances cPLA2 expression secondary to 5-HT2B-induced, transactivation-mediated ERK1/2 phosphorylation. *Psychopharmacology (Berl)* 207:1–12.

152. Girling R et al. (1993) A new component of the transcription factor DRTF1/E2F. *Nature* 365:468.

153. Wu CL, Classon M, Dyson N, Harlow E (1996) Expression of dominant-negative mutant DP-1 blocks cell cycle progression in G1. *Mol Cell Biol* 16:3698–3706.

154. Sagata N et al. (2010) Comprehensive behavioural study of GluR4 knockout mice: implication in cognitive function. *Genes Brain Behav* 9:899–909.

155. Reichert AS, Thurlow DL, Mörl M (2001) A eubacterial origin for the human tRNA nucleotidyltransferase? *Biol Chem* 382:1431–1438.

156. Yan Q, Guan M-X (2004) Identification and characterization of mouse TRMU gene encoding the mitochondrial 5-methylaminomethyl-2-thiouridylate-methyltransferase. *Biochim Biophys Acta* 1676:119–126.

157. Stoss O et al. (2004) p59(fyn)-mediated phosphorylation regulates the activity of the tissue-specific splicing factor rSLM-1. *Mol Cell Neurosci* 27:8–21.

158. Offenhäuser C et al. (2011) Syntaxin 11 binds Vti1b and regulates late endosome to lysosome fusion in macrophages. *Traffic* 12:762–773.

159. Pézeron G et al. (2008) Rasl11b knock down in zebrafish suppresses one-eyed-pinhead mutant phenotype. *PLoS ONE* 3:e1434.

160. Zhang H, Wang D, Sun H, Hall RA, Yun CC (2007) MAGI-3 regulates LPA-induced activation of Erk and RhoA. *Cell Signal* 19:261–268.

161. Hu J et al. (2008) Membrane proteins of the endoplasmic reticulum induce high-curvature tubules. *Science* 319:1247–1250.

162. Denayer E et al. (2008) Spred1 is required for synaptic plasticity and hippocampus-dependent learning. *J Neuroscience* 28:14443–14449.

163. Liang H et al. (2012) Knockdown of eukaryotic translation initiation factors 3B (EIF3B) inhibits proliferation and promotes apoptosis in glioblastoma cells. *Neurol Sci* 33:1057–1062.

164. Bagheri R et al. (2010) Relation of plasma fatty acid binding proteins 4 and 5 with the metabolic syndrome, inflammation and coronary calcium in patients with type-2 diabetes mellitus. *Am J Cardiol* 106:1118–1123.

165. Ludvigsen M, Jacobsen C, Maunsbach AB, Honoré B (2009) Identification and characterization of novel ERC-55 interacting proteins: evidence for the existence of several ERC-55 splicing variants; including the cytosolic ERC-55-C. *Proteomics* 9:5267–5287.

166. Pranski EL et al. (2012) Neuronal RING finger protein 11 (RNF11) regulates canonical NF-κB signaling. *J Neuroinflammation* 9:67.

167. Oldham MC, Horvath S, Geschwind DH (2006) Conservation and evolution of gene coexpression networks in human and chimpanzee brains. *Proc Natl Acad Sci USA* 103:17973–17978.

168. Miller JA, Horvath S, Geschwind DH (2010) Divergence of human and mouse brain transcriptome highlights Alzheimer disease pathways. *Proc Natl Acad Sci USA* 107:12698–12703.

169. Green EM, Gozani O (2011) CUL4B: trash talking at chromatin. *Mol Cell* 43:321–323.

170. Hara M et al. (2010) Role of adrenoceptors in the regulation of dopamine/DARPP-32 signaling in neostriatal neurons. *J Neurochem* 113:1046–1059.

171. Nicholas AP, Pieribone V, Hökfelt T (1993) Distributions of mRNAs for alpha-2 adrenergic receptor subtypes in rat brain: an in situ hybridization study. *J Comp Neurol* 328:575–594.

172. Pisani A et al. (2003) Activation of beta1-adrenoceptors excites striatal cholinergic interneurons through a cAMP-dependent, protein kinase-independent pathway. *J Neurosci* 23:5272–5282.

173. Paschalis A et al. (2009) beta1-Adrenoceptor distribution in the rat brain: an immunohistochemical study. *Neurosci Lett* 458:84–88.

174. Rommelfanger KS, Mitrano DA, Smith Y, Weinshenker D (2009) Light and electron microscopic localization of alpha-1 adrenergic receptor immunoreactivity in the rat striatum and ventral midbrain. *Neuroscience* 158:1530–1540.

175. Watson C, Paxinos G, Puelles L (2011) *The Mouse Nervous System* (Academic Press).

176. Hoekman MFM, Jacobs FMJ, Smidt MP, Burbach JPH (2006) Spatial and temporal expression of FoxO transcription factors in the developing and adult murine brain. *Gene Expr Patterns* 6:134–140.

177. Xie Z et al. (2006) Cellular and subcellular localization of PDE10A, a striatum-enriched phosphodiesterase. *Neuroscience* 139:597–607.

178. Matamales M et al. (2009) Striatal medium-sized spiny neurons: identification by nuclear staining and study of neuronal subpopulations in BAC transgenic mice. *PLoS ONE* 4:e4770.

179. Wang B, Lufkin T, Rubenstein JLR (2011) Dlx6 regulates molecular properties of the striatum and central nucleus of the amygdala. *J Comp Neurol* 519:2320–2334.

180. Threlfell S, Sammut S, Menniti FS, Schmidt CJ, West AR (2009) Inhibition of Phosphodiesterase 10A Increases the Responsiveness of Striatal Projection Neurons to Cortical Stimulation. *J Pharmacol Exp Ther* 328:785–795.

181. Stanwood GD, Parlaman JP, Levitt P (2006) Genetic or pharmacological inactivation of the dopamine D1 receptor differentially alters the expression of regulator of G-protein signalling (Rgs) transcripts. *Eur J Neurosci* 24:806–818.

182. Mathes WF et al. (2010) Dopaminergic dysregulation in mice selectively bred for excessive exercise or obesity. *Behav Brain Res* 210:155–163.

183. Kurz A et al. (2010) A53T-alpha-synuclein overexpression impairs dopamine signaling and striatal synaptic plasticity in old mice. *PLoS ONE* 5:e11464.

184. Brami-Cherrier K et al. (2005) Parsing molecular and behavioral effects of cocaine in mitogen- and stress-activated protein kinase-1-deficient mice. *J Neurosci* 25:11444–11454.

185. Cao YQ et al. (1998) Primary afferent tachykinins are required to experience moderate to intense pain. *Nature* 392:390–394.

186. Barceló AC, Filippini B, Pazo JH (2012) The striatum and pain modulation. *Cell Mol Neurobiol* 32:1–12.

187. De Felipe C et al. (1998) Altered nociception, analgesia and aggression in mice lacking the receptor for substance P. *Nature* 392:394–397.

188. Tang FI, Chiu TH, Wang Y (1998) Electrochemical studies of the effects of substance P on dopamine terminals in the rat striatum. *Exp Neurol* 152:41–49.

189. Shirai Y et al. (2010) Essential role of neuron-enriched diacylglycerol kinase (DGK), DGKbeta in neurite spine formation, contributing to cognitive function. *PLoS ONE* 5:e11602.

190. Hozumi Y et al. (2008) Diacylglycerol kinase beta accumulates on the perisynaptic site of medium spiny neurons in the striatum. *Eur J Neurosci* 28:2409–2422.

191. Dubielecka PM et al. (2011) Essential role for Abi1 in embryonic survival and WAVE2 complex integrity. *Proc Natl Acad Sci USA* 108:7022–7027.

192. Hida Y et al. (2011) Prickle2 is localized in the postsynaptic density and interacts with PSD-95 and NMDA receptors in the brain. *J Biochem* 149:693–700.

193. Okuda H, Miyata S, Mori Y, Tohyama M (2007) Mouse Prickle1 and Prickle2 are expressed in postmitotic neurons and promote neurite outgrowth. *FEBS Letters* 581:4754–4760.

194. Tao H et al. (2011) Mutations in prickle orthologs cause seizures in flies, mice, and humans. *Am J Hum Genet* 88:138–149.

195. Kuruba R, Hattiangady B, Shetty AK (2009) Hippocampal neurogenesis and neural stem cells in temporal lobe epilepsy. *Epilepsy Behav* 14 Suppl 1:65–73.

196. Tomer R, Denes AS, Tessmar-Raible K, Arendt D (2010) Profiling by image registration reveals common origin of annelid mushroom bodies and vertebrate pallium. *Cell* 142:800–809.

197. Hekmat-Scafe DS et al. (2010) Seizure Sensitivity Is Ameliorated by Targeted Expression of K+–Cl− Cotransporter Function in the Mushroom Body of the Drosophila Brain. *Genetics* 184:171–183.

198. Colombo M, Broadbent N (2000) Is the avian hippocampus a functional homologue of the mammalian hippocampus? *Neurosci Biobehav Rev* 24:465–

484.

199. Neafsey EJ, Hurley-Gius KM, Arvanitis D (1986) The topographical organization of neurons in the rat medial frontal, insular and olfactory cortex projecting to the solitary nucleus, olfactory bulb, periaqueductal gray and superior colliculus. *Brain Res* 377:261–270.

200. Jabaudon D, Shnider SJ, Tischfield DJ, Galazo MJ, Macklis JD (2012) RORβ induces barrel-like neuronal clusters in the developing neocortex. *Cereb Cortex* 22:996–1006.

201. Tom Tang Y et al. (2004) TAFA: a novel secreted family with conserved cysteine residues and restricted expression in the brain. *Genomics* 83:727–734.

202. Liao W-L et al. (2008) Modular patterning of structure and function of the striatum by retinoid receptor signaling. *Proc Natl Acad Sci USA* 105:6765–6770.

203. Lobo MK, Cui Y, Ostlund SB, Balleine BW, Yang XW (2007) Genetic control of instrumental conditioning by striatopallidal neuron-specific S1P receptor Gpr6. *Nat Neurosci* 10:1395–1397.

204. Sørensen AT et al. (2008) NPY gene transfer in the hippocampus attenuates synaptic plasticity and learning. *Hippocampus* 18:564–574.