PNAS
www.pnas.org

Supplementary Information for

Acute social isolation alters neurogenomic state in songbird forebrain

Julia M George, Zachary W Bell, Dan Condliffe, Kirstin Dohrer, Teresa Abaurrea, Karen Spencer, Albertine Leitao, Manfred Gahr, Paul J Hurd, David F Clayton

Email: d.clayton@qmul.ac.uk

This PDF file includes:

Supplementary text
Figures S1 to S10
Table S1
Caption for Dataset S1
References for SI

Other supplementary materials for this manuscript include the following:

Dataset S1

www.pnas.org/cgi/doi/10.1073/pnas.1820841116

Supplementary Information Text

SI Materials and Methods.

Animals

In the USA (June, 2012), zebra finches were taken from a colony maintained at the Beckman Institute, University of Illinois at Urbana-Champaign. All procedures involving animals were conducted with protocols approved by the University of Illinois Institutional Animal Care and Use Committee. Male and female adult zebra finches (>120 days post-hatch) were housed in single sex flight cages within the same room ("group aviary"), so that animals of the same sex could see and hear one another but were physically separated. Birds were acclimated to these housing conditions for at least one week prior to initiation of experiments. On day 1 of the treatment, 3-6 birds were removed from the aviary between 2 and 4pm and placed individually in separate sound attenuation chambers in a different room. On day 2, birds were removed from the sound attenuation chambers and immediately killed by decapitation. An equal number of control animals of the same sex were immediately collected from the group aviary and also killed by the same method; all animals were killed on the same day between 11am and 12pm. The sequence was repeated as necessary to complete each cohort, except that on subsequent experimental days, the order of killing was alternated for experimental balance. Brains were removed and then either immediately dissected (auditory lobule dissection, below) followed by flash freezing in liquid nitrogen, or frozen in plastic molds (Peel-a-Way, Sigma, St. Louis, MO) with O.C.T. embedding medium (TissueTek, Sakura Finetek, Torrance, CA) in a slurry of ethanol and dry ice. All samples were frozen within 5 minutes of death and stored at -80°C until use.

For subsequent replications in the UK, birds were collected from a different colony established at Queen Mary University of London from UK stocks. Animal housing and welfare were in compliance with the European directives for the protection of animals used for scientific purposes (2010/63/EU) under Procedures Project License PPL70-8183. For the time course experiment (Figure 3, November-December 2016), aviary (0 day) and solo (1 day) groups were collected as before, but with an additional treatment group (2 day) that was housed in the chamber for 2 nights (9, 10, and 9 females per group, respectively). For the partner experiment (Figure 5, February-March 2016), aviary and solo groups were collected as before, with the addition of a group housed overnight with a partner of the same sex (females: 10 aviary, 10 solo, 11 duo; males: 12 aviary, 11 solo, 11 duo). For the primary RT-qPCR replication (Figure S4, January 2015), females (5 aviary, 5 solo) were housed similarly, but killed by isoflurane overdose and decapitation.

Animals for RRBS experiments were collected at Max Planck Institute for Ornithology in Seewiesen in August 2015. Animal housing and welfare were in compliance with the European directives for the protection of animals used for scientific purposes (2010/63/EU). Protocols were approved by the Government of Upper Bavaria. Male

zebra finches (range 147-170 days old, mean age in each treatment group 156 and 157 days, respectively) were collected either directly from a group aviary or placed in sound chambers between 4 and 5pm, and killed two days later. All animals (6 aviary, 6 solo) were killed by decapitation, between 3 and 4pm.

Auditory Lobule Dissection
RNA-seq, RT-qPCR and RRBS analyses were all based on tissues collected using the "auditory lobule" (AL) dissection of the caudomedial telencephalon (1). This dissection collects the major loci of the *zenk/egr1* response to song playback (the caudomedial nidopallium (NCM) and the caudomedial mesopallium (CMM)), along with non-responsive Field L2a.

RNA Sequencing
AL samples were processed by the Barts and the London Genome Centre (BLGC) at Queen Mary University of London, who performed RNA extraction (RNEasy Mini Kit, Qiagen), yielding 2-12 ug total RNA per sample, with RNA Integrity Number (RIN) > 9 for all samples. The Genome Centre also prepared indexed sequencing libraries (TruSeq, Illumina) and provided sequencing (HiSeq, Illumina) at an average read depth of 31.7 million 100bp paired-end reads per sample. Raw read files are deposited at European Nucleotide Archive Study Accession #PRJEB28085). Reads were processed with Trimmomatic (2) v. 0.30 with the flags: ILLUMINACLIP:TruSeq3-PE.fa:2:30:10 LEADING:3 TRAILING:3 SLIDINGWINDOW:4:15 MINLEN:36. Surviving paired reads were mapped to the Taeniopygia guttata transcriptome (GCF_000151805.1_Taeniopygia_guttata-3.2.4_rna.fna, downloaded from ftp://ftp.ncbi.nlm.nih.gov/genomes/all/GCF/000/151/805/GCF_000151805.1_Taeniopygia_guttata-3.2.4/GCF_000151805.1_Taeniopygia_guttata-3.2.4_rna.fna.gz) with Salmon (v 0.9.1) (3). Mapping efficiency ranged between 57.7 and 61.1%. Transcript read counts were collapsed to genes with the package tximport (v 1.6.0) (4) in R version 3.4.1 (https://www.R-project/org); counts and fragment lengths were imported into DESeq2 (v 1.18.1) (5) for analysis of differential gene expression. This package fit a generalized linear model to a binomial distribution and tested for differential expression, with a Wald test for the significance of GLM coefficients and independent filtering to optimize detection of genes below the target threshold. After this filtering step, adjusted p values (6) were returned for 14687 genes (among 16628 genes detected at minimal level of > 10 counts across all samples in this experiment). The function lfcShrink() was employed to moderate fold-changes of low abundance genes.
The R code for tximport and DESeq2 is below:

```
#terminal commands to get transcript IDs and gene IDs from transcriptome fasta

grep ">" GCF_000151805.1_Taeniopygia_guttata-3.2.4_rna.fna > headers.txt #extract fasta headers
cat headers.txt | cut -d " " -f1 |sed 's/^>//' > tx_id.txt #extract list of transcript IDs
sed 's/^.*(//' headers.txt | sed 's/).*$//' > gene_id.txt #extract list of gene IDs (symbols)

#load required R packages
library(dplyr)
library(tibble)
```

```r
library(magrittr)
library(tximport)
library(readr)
library(rjson)
library(DESeq2)
library(matrixStats)
library(ggplot2)
library(ggsci)
library(cowplot)

#make a data frame with transcript to gene mappings
tx2gene<-
data.frame(read.csv("tx_id.txt",header=FALSE),read.csv("gene_id.txt",header=FALSE))

#get sampleID
directory<-"quants"
files <- file.path(directory, list.files(directory), "quant.sf")
sampleID<-gsub("_quant","",list.files(directory))

#import transcript-level abundances
d <- tximport(files, type="salmon", tx2gene=tx2gene)

#add group information
samples<-
data.frame(sampleID=sampleID,condition=factor(rep(c("Avi","Avi","Avi","Iso","Iso","Iso"
),4)))

#create DESeq Dataset
ddsTxi <- DESeqDataSetFromTximport(d,
                                   colData = samples,
                                   design = ~ condition)
#explore sample quality
#minimal filtering to remove low-count genes
keep <- rowSums(counts(ddsTxi)) >= 10
DESeq2Table<-ddsTxi[keep,]
DESeq2Table <- estimateSizeFactors(DESeq2Table)

#variance stabilizing transformation of count data
rld <- rlogTransformation(DESeq2Table, blind=TRUE)

#Principal Components Analysis (select 500 genes with highest variance)
ntop = 500
Pvars <- rowVars(assay(rld))
select <- order(Pvars, decreasing = TRUE)[seq_len(min(ntop,
        length(Pvars)))]
PCA <- prcomp(t(assay(rld)[select, ]), scale = F)
percentVar <- round(100*PCA$sdev^2/sum(PCA$sdev^2),1)

dataGG = data.frame(PC1 = PCA$x[1:length(colnames(rld)),1], PC2 =
PCA$x[1:length(colnames(rld)),2],
                    PC3 = PCA$x[1:length(colnames(rld)),3], PC4 =
PCA$x[1:length(colnames(rld)),4],
                    PC5 = PCA$x[1:length(colnames(rld)),5],PC6 =
PCA$x[1:length(colnames(rld)),6],
                    condition = colData(rld)$condition,
                    sampleID=colData(rld)$sampleID)
summary(PCA)

## Importance of components%s:
##                             PC1    PC2     PC3     PC4     PC5     PC6
## Standard deviation     6.2867 3.8219 2.77869 2.68367 2.45412 2.23019
```

```
## Proportion of Variance 0.3239 0.1197 0.06327 0.05902 0.04935 0.04076
## Cumulative Proportion  0.3239 0.4436 0.50682 0.56584 0.61519 0.65595
##                            PC7     PC8     PC9    PC10    PC11    PC12
## Standard deviation     2.04817 2.00609 1.91510 1.82950 1.81062 1.75274
## Proportion of Variance 0.03438 0.03298 0.03005 0.02743 0.02686 0.02517
## Cumulative Proportion  0.69032 0.72330 0.75336 0.78078 0.80765 0.83282
##                           PC13    PC14    PC15    PC16    PC17   PC18
## Standard deviation     1.64250 1.61472 1.52631 1.44961 1.39435 1.3843
## Proportion of Variance 0.02211 0.02137 0.01909 0.01722 0.01593 0.0157
## Cumulative Proportion  0.85493 0.87629 0.89538 0.91260 0.92853 0.9442
##                           PC19    PC20    PC21    PC22    PC23      PC24
## Standard deviation     1.30001 1.22093 1.17186 1.09174 1.02931 9.573e-15
## Proportion of Variance 0.01385 0.01221 0.01125 0.00977 0.00868 0.000e+00
## Cumulative Proportion  0.95808 0.97030 0.98155 0.99132 1.00000 1.000e+00

#Plot Principal Components
PC1_2<-ggplot(dataGG,aes(PC1,PC2))+
  geom_text(aes(label=sampleID,colour=condition))+
  theme(legend.position="none")+
  scale_color_npg()+
  xlab(paste0("PC1: ",percentVar[1],"% variance")) +
  ylab(paste0("PC2: ",percentVar[2],"% variance"))
PC3_2<-ggplot(dataGG,aes(PC3,PC2))+
  geom_text(aes(label=sampleID,colour=condition))+
  theme(legend.position="none")+
  scale_color_npg()+
  xlab(paste0("PC3: ",percentVar[3],"% variance")) +
  ylab(paste0("PC2: ",percentVar[2],"% variance"))

plot_grid(PC1_2,PC3_2,labels=c("A", "B"), ncol = 2)

#Sample F02 is a clear outlier (Figure S1)
#Analyse with DESeq, omitting Sample F02
#import transcript-level abundances (without sample F02)
d <- tximport(files[-2], type="salmon", tx2gene=tx2gene)

#create DESeq Dataset (without sample F02)
ddsTxi <- DESeqDataSetFromTximport(d,
                                   colData = samples[-2,],
                                   design = ~ condition)

#minimal filtering to remove low-count genes (this is not necessary as it is done
automatically at a later step, but it may increase speed);
keep <- rowSums(counts(ddsTxi)) >= 10
dds <- ddsTxi[keep,]

#The function DESeq() estimates size factors, estimates dispersion, and carries out
negative binomial GLM fitting and Wald statistics
set.seed(1)
dds <- DESeq(dds) #default Wald test

#The function results() extracts a table of results, default pAdjustMethod="BH"
res <- results(dds, alpha=0.05, contrast=c("condition","Iso","Avi"),)
summary(res)

##
## out of 16621 with nonzero total read count
## adjusted p-value < 0.05
## LFC > 0 (up)     : 141, 0.85%
## LFC < 0 (down)   : 187, 1.1%
## outliers [1]     : 0, 0%
```

```
## low counts [2]   : 1941, 12%
## (mean count < 6)
## [1] see 'cooksCutoff' argument of ?results
## [2] see 'independentFiltering' argument of ?results

#The function lfcShrink() moderates the estimates of fold change to reduce the impact
of low count genes
resLFC <- lfcShrink(dds=dds, res=res, contrast=c("condition","Iso","Avi")) # shrinkage
estimator type="normal"

#filter for significance threshold and fold change (head()displays first 6 rows)
resLFC %>%  as.data.frame() %>% rownames_to_column(var = "gene") %>% filter(padj<.001)
%>% filter(abs(log2FoldChange)>.4) %>% arrange(log2FoldChange) %>% head()

##     gene   baseMean log2FoldChange      lfcSE      stat       pvalue
## 1   EGR1 2931.04884     -0.7180029 0.11006224 -6.528282 6.652847e-11
## 2   NWD2 2994.21645     -0.5845969 0.07062372 -8.277005 1.263117e-16
## 3  DUSP4   73.86684     -0.5801860 0.10558348 -5.489183 4.037981e-08
## 4   BDNF  157.28263     -0.5579260 0.08276232 -6.742003 1.562177e-11
## 5   SHC3  314.12519     -0.5282510 0.09053797 -5.827809 5.615986e-09
## 6 UBXN2A  354.03779     -0.5263443 0.08622529 -6.108470 1.005907e-09
##           padj
## 1 1.372583e-07
## 2 1.855140e-12
## 3 2.280993e-05
## 4 4.588739e-08
## 5 4.582333e-06
## 6 9.849171e-07

#visualize log2 fold changes versus mean counts
plotMA(resLFC, ylim=c(-1,1))
```

Gene Set Enrichment Analysis

A database of gene sets representing KEGG pathways (c2.cp.kegg.v6.1.symbols.gmt) was downloaded from the Molecular Signatures Database v6.1 (http://software.broadinstitute.org/gsea/msigdb/download_file.jsp?filePath=/resources/msigdb/6.1/c2.cp.kegg.v6.1.symbols.gmt). The gene set database was indexed against the gene symbols used as identifiers in the DESeq2 analysis. cameraPR from the R package limma (v 3.34.8) (7) was used to perform competitive gene set enrichment analysis, using the Wald statistic from DESeq2 to rank the gene list, and with an intergene correlation of 0.01. Competitive gene set enrichment analysis considers the rank order of expression intensity for each gene in a defined subset, compared to all other gene sets, and asks whether there are significant differences in the overall mean rank order for the set as a whole between treatment conditions. This approach avoids the severe correction for multiple testing required in evaluating individual gene effects in genome-wide datasets.

RT-qPCR

RNA was extracted from dissected AL brain tissue (RNeasy Plus Mini Kit, Qiagen), and cDNA was prepared by reverse transcription (iScript cDNA synthesis kit, Biorad). Primer sets for RT-qPCR were selected using NCBI Primer-BLAST, or from a study of reference genes for RT-qPCR (8) (Table S1), and purchased from Sigma or Eurofins Genomics. Assays were tested to confirm that they gave 95-100% amplification efficiency relative to a standard curve, and primers for assays using DNA-binding dyes (e.g. SYBR green) were confirmed to yield a single PCR product by melting curve analysis (Figure S3). RT-qPCR reactions were prepared with 15ng template cDNA per well and the appropriate supermix (Biorad iTaq Universal SYBR Green Supermix for primer-only assays, 5x HOT FIREPol Probe Universal qPCR Mix for assays with hydrolysis probes). Optimal annealing temperatures were between 60-61ºC for all assays. PCR reactions were run on a CFX Connect Real-Time PCR Detection System, except for assays in Fig. S4, which were run by staff at BLGC on an ABI 7900HT.

RT-qPCR data were analyzed with the R package MCMC.qpcr (v 1.2.3) (9). In this method, Cq data are converted to molecule counts, and gene expression is modeled across all genes and samples simultaneously in a joint GLMM. This model estimates fixed treatment effects along with unobserved random effects common to all genes in a sample (attributable e.g. to differences in sample loading or quality). This approach thus achieves the same objective as normalization to a stable "control" gene, without assumptions about the stability of such genes.

Corticosterone radioimmunoassay

Each animal was captured by hand, immediately decapitated, and the body placed into a 50 mL plastic tube containing 10 µL K_2EDTA (anticoagulant) to collect the trunk blood (~1 mL). Blood collection was completed within one minute of capture. The tube was swirled in order to mix the blood with the K_2EDTA, and 50 µL was centrifuged at 3000 x g for 10 min in a benchtop microcentrifuge and the plasma layer kept at -20º C for later radioimmunoassay (10). Corticosterone was extracted from 20 µl aliquots of plasma using diethyl ether, individual extraction efficiencies were calculated using radiolabeled CORT (3H-Cort, Perkin Elmer, UK) and ranged from 65 to 90%. Radioimmunoassay protocols quantified CORT in duplicate samples, using an anti-CORT antibody (1:100 dilution, MP Biomedicals) and dextran coated charcoal to separate bound and free portions. All samples were run in one assay with an intraassay CV of 8.5%, a detection limit of 0.04ng/ml, and a 50% binding level of 0.84 ng/ml. 5 of 71 samples (2 female aviary, 2 male aviary, and 1 female solo) were below detection and set to the detection limit for analysis purposes. We analyzed plasma CORT concentrations using GLM in R (version 3.4.1, https://www.R-project/org), using the glm function, with a Gamma family and an inverse link function. Likelihood ratio tests indicated that a model based on a single combined factor of sex and housing condition provided a good fit to plasma CORT concentration, when compared to the null model (Loglikelihood Ratio = 15.59073, delta AIC = -21.18146, df = 7, p = 8.155e-07). We then carried out post-hoc tests using the glht function in package multcomp (11) , using the leave-one-out correction for multiple comparisons. We examined whether there were differences in CORT levels among the

three housing conditions (aviary, duo, solo) within each sex, as well as aggregate effects between sexes.

In Situ Hybridization

With the goal of comparing the general anatomical patterns of the four RNAs of interest, 864 sections of adult female brains were examined from three replications of the aviary-solo experiment, each with 6 animals (3 aviary, 3 solo): USA 2012 (coronal sections); UK 2015 (sagittal sections); UK 2016 (2 coronal and 1 sagittal for each group). In the coronal plane, sections were collected from the caudal edge of the forebrain rostrally through the end of CMM. In the sagittal plane, sections were collected from the midline to 1.5 mm laterally. Our analysis focused primarily on sections containing elements of the auditory forebrain (NCM, CMM and Field L2a), but labeling patterns in adjacent parts of each section were also noted. Brains were sectioned at 10 µm with a Leica cryostat (Model CM 3050S) and collected onto Superfrost Plus slides. After sectioning, the tissue was fixed in fresh 3% paraformaldehyde in PBS (pH 7.4) for 5 minutes. Slides were then rinsed in PBS (pH 7.4), dehydrated in an ascending ethanol series (70%, 95%, 100%; 2 min each), air dried, and stored at -80^{o}C.

Riboprobe templates for EGR1 (GenBank CK308891.1) and UTS2B (GenBank DV945629.1), were used from the ESTIMA cDNA collection (12). cDNAs for FKBP5 and BDNF were first cloned via reverse transcription PCR. Forward and reverse primers (FKBP5 forward 5'- GTGTCCTGGCTGGAAATGGA -3', FKBP5 reverse 5'- TTGGCTTCCACGATCACACA -3'; BDNF forward 5'- TTGGCTTACCCAGGTCTTCG -3', BDNF reverse 5'- TTATGAAGCGCCAGCCAACT -3') were designed using Primer-BLAST (13). PCR was carried out with zebra finch forebrain cDNA and Taq DNA polymerase under standard conditions (New England Biolabs), products cloned into the vector pCRII using the TOPO-TA cloning kit (Life Technologies), and inserts confirmed by commercial sequencing (Eurofins). The 610bp FKBP5 cDNA insert corresponds to residues 962-1571 of the zebra finch FKBP5 transcript (NCBI RefSeq XM_002198980.3), while the 598bp BDNF cDNA insert aligns to residues 98 to 695 of the zebra finch BDNF transcript (NCBI RefSeq NM_001048255.1), which is common to all currently annotated transcripts of BDNF. *In vitro* transcription of digoxigenin- (DIG) labeled riboprobes and subsequent in situ hybridizations were carried out as described (14). Digital images were obtained for all *in situ* hybridized sections using a Nanozoomer whole slide scanner (Hamamatsu). These images were examined in NDP.view2 (Hamamatsu) and compared to histological reference images (15) available in the Zebra Finch Expression Brain Atlas (http://www.zebrafinchatlas.org). Cells were considered DIG-labeled if dark circles were observed indicating cytoplasmic labeling (16). For semi-quantitative comparisons, TIFF-formatted images were imported into ImageJ (17) and converted to 8-bit format. The threshold was set to include DIG-labeled cells and exclude unlabeled cells as described (18).

Reduced representation bisulfite sequencing (RRBS)

RRBS libraries were prepared as described previously (19), with 300ng input DNA digested with both Msp1 and Taq1 to increase genomic coverage, and sequenced by the Biomedical Sequencing Facility at CeMM, Vienna, on the Illumina HiSeq2000 platform

in 50bp single-read mode (European Nucleotide Archive Study Accession # PRJEB28656, sample accession numbers ERS2719622-ERS2719628 and ERS2719634-ERS2719639). Data were preprocessed with trimgalore v.0.3.3 (http://www.bioinformatics.babraham.ac. uk/projects/trim_galore/) for removal of adapters and low-quality sequence, as described (19). Reads were aligned to the zebra finch genome (taeGut1, WUSTL v3.2.4, downloaded from http://hgdownload.soe.ucsc.edu/goldenPath/taeGut1/chromosomes/ and concatenated to a single fasta file) with BSMAP v2.90 (20), using command: bsmap -a $input_fastq -d $ref_genome_fasta -o $output_bam -D C-CGG -D T-CGA -w 100 -v 0.08 -r 0 -p 4 -n 0 -s 12 -S 0 -f 5 -q 0 -u -V 2. Resulting BAM outputs were filtered to extract alignments corresponding to 328 differentially expressed genes (Dataset S1, q<0.05), including 10kb upstream of the 5'-most transcription start site of each with samtools v1.9, using the command: samtools view -b -L $genomic_intervals.bed $output_bam > $interval.bam. Methylated CpGs were extracted with the python script methratio.py distributed with BSMAP, with the command: methratio.py -o $methratio.txt -d $ref_genome_fasta -p -z -x CG $interval.bam. The R package methylKit v1.8.1 (21) was used to filter for coverage and assess differential methylation. The R code for methylKit is as follows:

```
library("methylKit","dplyr")

setwd("./methratio_files")
file.list = list("G1AV144_methratio.txt", "G1AV207_methratio.txt",
"G1AV234_methratio.txt", "G1SI152_methratio.txt", "G1SI199_methratio.txt",
"G1SI246_methratio.txt", "G2AV176_methratio.txt", "G2AV216_methratio.txt",
"G2AV229_methratio.txt", "G2SI146_methratio.txt", "G2SI188_methratio.txt",
"G2SI218_methratio.txt")

#import outputs of methratio.py, and specify groups
myobj=methRead(
file.list,pipeline=list(fraction=TRUE,chr.col=1,start.col=2,end.col=2,
coverage.col=6,strand.col=3,freqC.col=5 ),
sample.id=list("G1AV144",
"G1AV207","G1AV234","G1SI152","G1SI199","G1SI246","G2AV176","G2AV216",
"G2AV229", "G2SI146", "G2SI188",
"G2SI218"),assembly="taeGut1",treatment=c(0,0,0,1,1,1,0,0,0,1,1,1))

#filter for minimum coverage of 10 reads per CpG
filtered.myobj=filterByCoverage(myobj,lo.count=10,lo.perc=NULL,
                                hi.count=NULL,hi.perc=NULL)
#merge sites covered in all samples
meth=unite(filtered.myobj, destrand=TRUE)

#get a methylDiff object containing the differential methylation statistics
and locations for regions or bases
myDiff=calculateDiffMeth(meth)

## two groups detected:
## will calculate methylation difference as the difference of
## treatment (group: 1) - control (group: 0)

getData(myDiff)
```

```
##      chr     start       end strand       pvalue       qvalue  meth.diff
##1   chr1  26807844  26807844      + 2.100911e-06 5.749180e-05 -13.295020
##2   chr1  28674040  28674040      + 2.174991e-06 5.852703e-05  -7.422916
##3   chr1  28674068  28674068      + 7.902220e-07 2.667788e-05  -9.704433
##4   chr1  28674071  28674071      + 2.323728e-09 1.782083e-07 -15.074925
##5   chr1  58268045  58268045      + 2.196834e-08 1.182296e-06  19.397022
##6  chr11  11235339  11235339      + 4.882716e-15 1.313894e-12  22.076329
##7  chr12    824557    824557      + 4.827407e-14 9.742585e-12 -22.676898
##8  chr13   6719766   6719766      + 1.670352e-06 4.649758e-05 -17.737380
##9  chr13   6719775   6719775      + 2.748954e-09 1.849297e-07 -24.985700
##10 chr13  16606405  16606405      + 1.311896e-06 3.996444e-05   7.998255
##11 chr13  16607295  16607295      + 5.304617e-10 5.037968e-08   9.286151
##12 chr13  16607309  16607309      + 9.647587e-07 3.115294e-05   7.767296
##13 chr13  16607328  16607328      + 4.152742e-16 1.676198e-13  12.409930
##14 chr13  16607363  16607363      + 2.439995e-08 1.270801e-06  11.008230
##15 chr13  16607380  16607380      + 3.553113e-08 1.738383e-06   9.657910
##16 chr13  16607392  16607392      + 2.970682e-07 1.169830e-05   8.540993
##17 chr14   2250807   2250807      + 2.222056e-10 2.759701e-08 -33.323595
##18 chr14   2401154   2401154      + 1.720389e-07 7.507150e-06 -10.861558
##19 chr14   6267396   6267396      + 2.517082e-09 1.782083e-07  -6.245528
##20 chr14   6268911   6268911      + 2.042045e-07 8.580086e-06  12.006766
##21 chr15  11242323  11242323      + 2.463816e-10 2.841387e-08 -14.598691
##22 chr15  12567169  12567169      + 2.574194e-08 1.298798e-06  23.042776
##23 chr18   1566555   1566555      + 1.639625e-06 4.644296e-05 -13.843321
##24 chr19   6837757   6837757      + 2.369287e-06 6.271019e-05 -15.233379
##25 chr19   9531949   9531949      + 1.306709e-08 7.274971e-07  -9.285692
##26  chr2  26858137  26858137      + 7.931261e-07 2.667788e-05   8.225538
##27  chr2  26863165  26863165      + 5.174218e-13 9.282234e-11 -22.428964
##28  chr2  26868252  26868252      + 2.072554e-07 8.580086e-06   4.705882
##29  chr2 142482532 142482532      + 3.397552e-07 1.306072e-05  -6.267817
##30  chr2 155945864 155945864      + 5.147458e-07 1.888819e-05  17.589631
##31  chr2 155955302 155955302      + 1.147959e-10 1.684939e-08 -18.957080
##32 chr22   3233132   3233132      + 1.434724e-10 1.930356e-08  16.138862
##33 chr23   3996896   3996896      + 9.382528e-10 7.972904e-08 -28.139785
##34 chr24    321566    321566      + 5.844660e-07 2.051406e-05   9.714463
##35 chr24    346587    346587      + 5.275977e-09 3.407322e-07  22.222222
##36 chr24   1604965   1604965      + 4.721728e-07 1.772894e-05  19.309048
##37 chr27   1847413   1847413      + 3.843899e-06 9.851031e-05 -14.055267
##38 chr28    833898    833898      + 5.934702e-10 5.323248e-08  36.057692
##39 chr28    876003    876003      + 2.538666e-09 1.782083e-07  28.586171
##40 chr28    876015    876015      + 1.522277e-06 4.551455e-05  14.931258
##41 chr28   2223370   2223370      + 1.338362e-14 3.086923e-12 -14.776769
##42 chr28   4235123   4235123      + 3.774818e-06 9.830025e-05  10.434556
##43 chr28   4235977   4235977      + 1.107491e-06 3.506067e-05   9.016393
##44  chr3   1315369   1315369      + 2.480313e-07 1.001144e-05 -26.527200
##45  chr3   2071416   2071416      + 3.801128e-11 6.137094e-09 -17.346199
##46  chr3  22795202  22795202      + 3.200611e-15 1.033506e-12  -2.283380
##47  chr3  35440046  35440046      + 5.748626e-39 4.640709e-36 -18.699342
##48  chr3  35441117  35441117      + 3.476055e-10 3.507656e-08  -8.069717
##49  chr3  52533892  52533892      + 5.442362e-08 2.584395e-06   7.577038
##50  chr3 109033987 109033987      + 8.305338e-21 4.469782e-18 -29.939348
##51 chr4A  11827369  11827369      + 5.811654e-09 3.608915e-07 -16.550558
##52  chr5   1308194   1308194      + 1.050659e-75 1.696337e-72 -41.017145
##53  chr5   1328464   1328464      + 9.209430e-07 3.034499e-05  26.038462
##54  chr5  51663801  51663801      + 1.209249e-06 3.754592e-05   7.134168
```

```
##55  chr5  51664419  51664419      + 1.102218e-07 5.084515e-06  -1.988072
##56  chr6   2233280   2233280      + 1.619688e-06 4.644296e-05  11.574657
##57  chr7  10955391  10955391      + 7.880729e-09 4.544213e-07  -3.641457
##58  chr7  10955405  10955405      + 1.605702e-06 4.644296e-05  -3.401880
##59  chr7  10955420  10955420      + 1.296274e-07 5.813593e-06  -4.950676
##60  chr9   4565870   4565870      + 2.769982e-10 2.981508e-08  26.587302
##61  chr9   4565875   4565875      + 7.201156e-09 4.306146e-07 -22.346007
##62  chrZ   7304971   7304971      + 5.589131e-07 2.005312e-05  -8.252427
##63  chrZ  39325638  39325638      + 1.683187e-09 1.358791e-07 -13.607678
```

```
#these sites were annotated for nearest gene and plotted in Figure S10
```

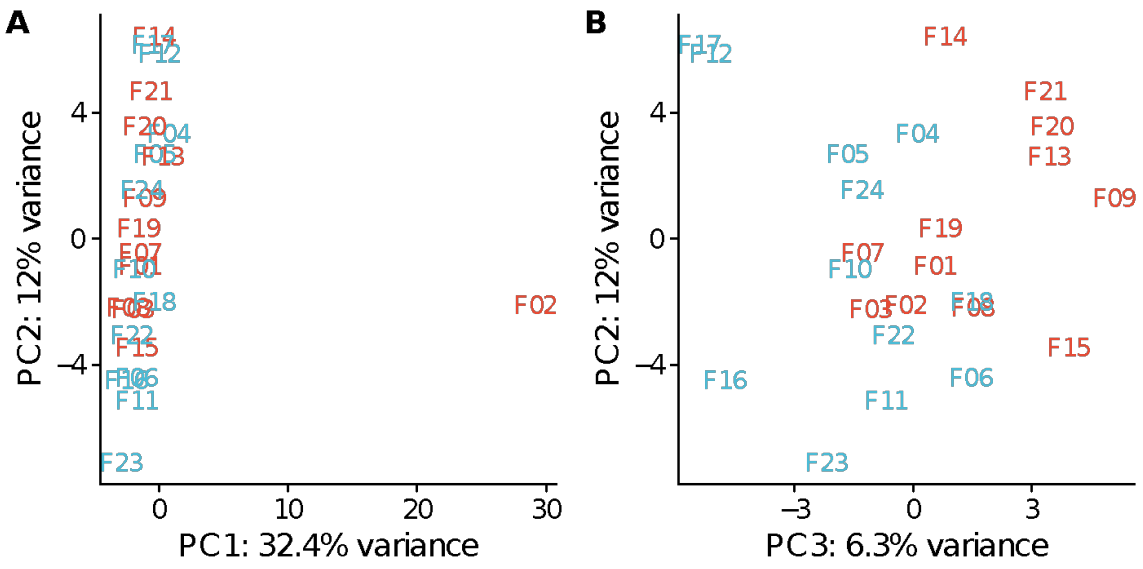

Figure S1. Principal components analysis of RNA-seq samples. Count data for the 500 most variant genes across all samples was subjected to principal components analysis. A) PC2 versus PC1. The single sample (F02) is an outlier with respect to all other samples in PC1. B) PC2 versus PC3. Samples cluster somewhat by treatment group; aviary (red) and solo (blue).

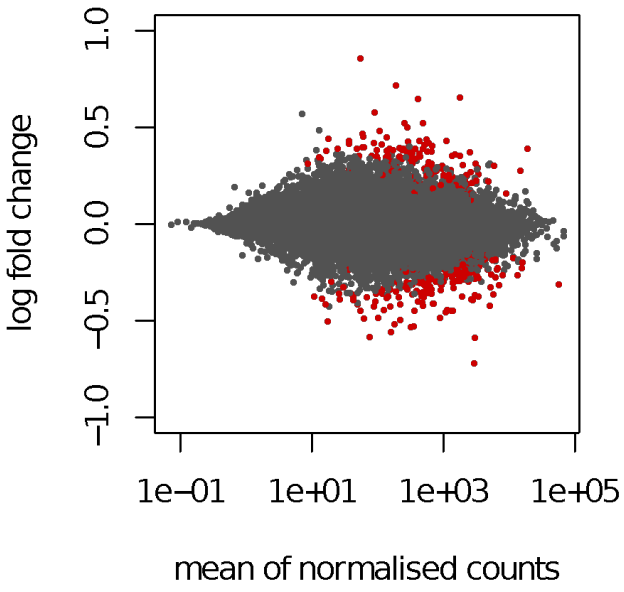

Figure S2. MA plot for RNA-seq in female auditory lobule after overnight solo isolation. Log2 fold changes in gene expression are plotted versus the mean of normalized counts. Fold changes are moderated with a shrinkage estimator to reduce noise arising from low-count genes. Points plotted in red represent genes with FDR < 0.05. The full results of the DESeq2 analysis are available in Dataset S1.

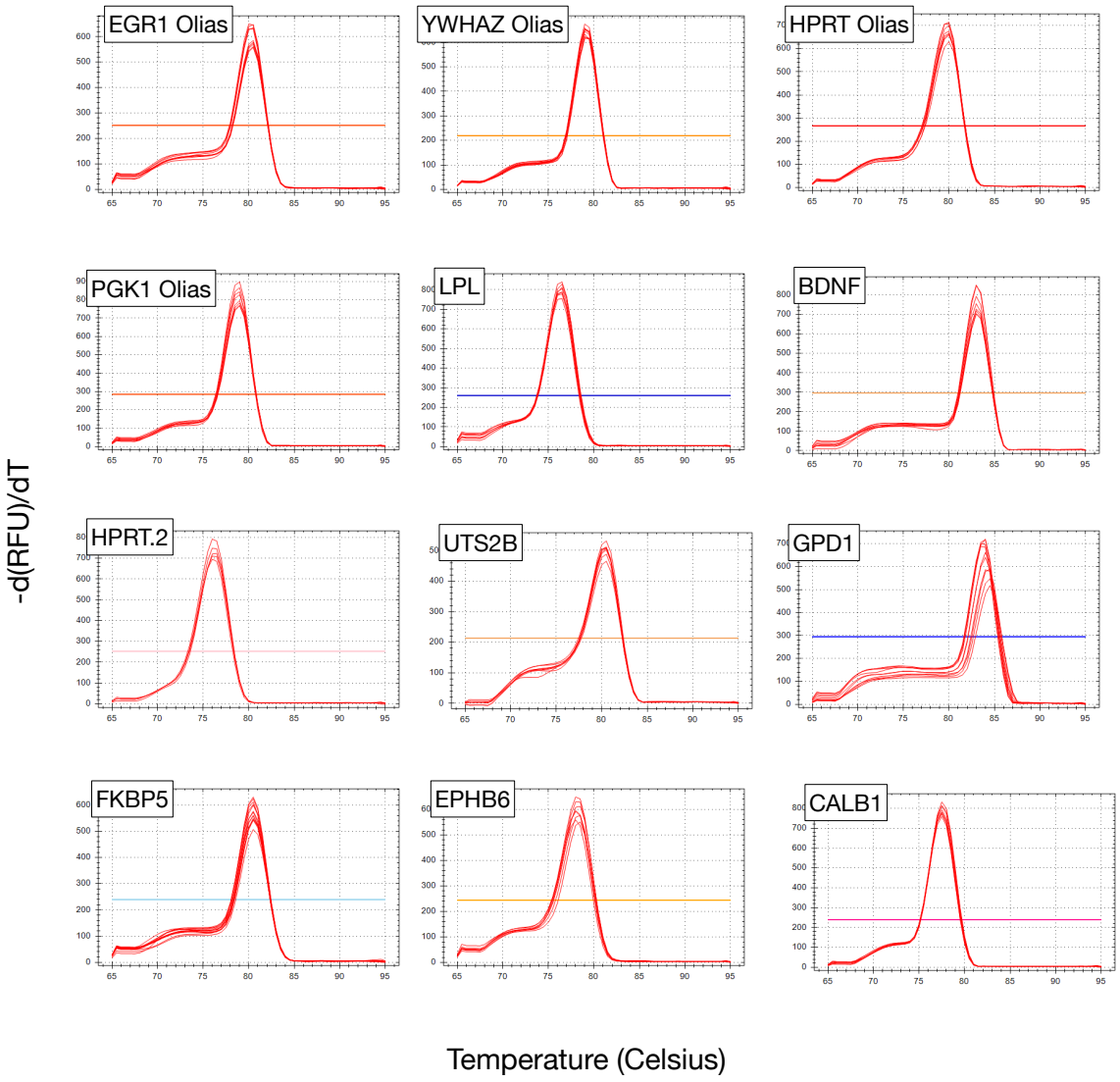

Figure S3. Melting curve analysis of RT-qPCR assays. To validate the RT-qPCR assays that depend upon DNA-intercalating dyes for detection of the amplicon, melting curve analysis was conducted on the final reaction product. After the reaction was complete, the temperature was raised from 65 to 95°C, and the fluorescence was measured at 0.5°C increments, to monitor denaturation of dsDNA, and the rate of denaturation plotted as a function of temperature. The presence of a single sharp peak indicates a single reaction product, with no evidence for secondary product or primer dimers.

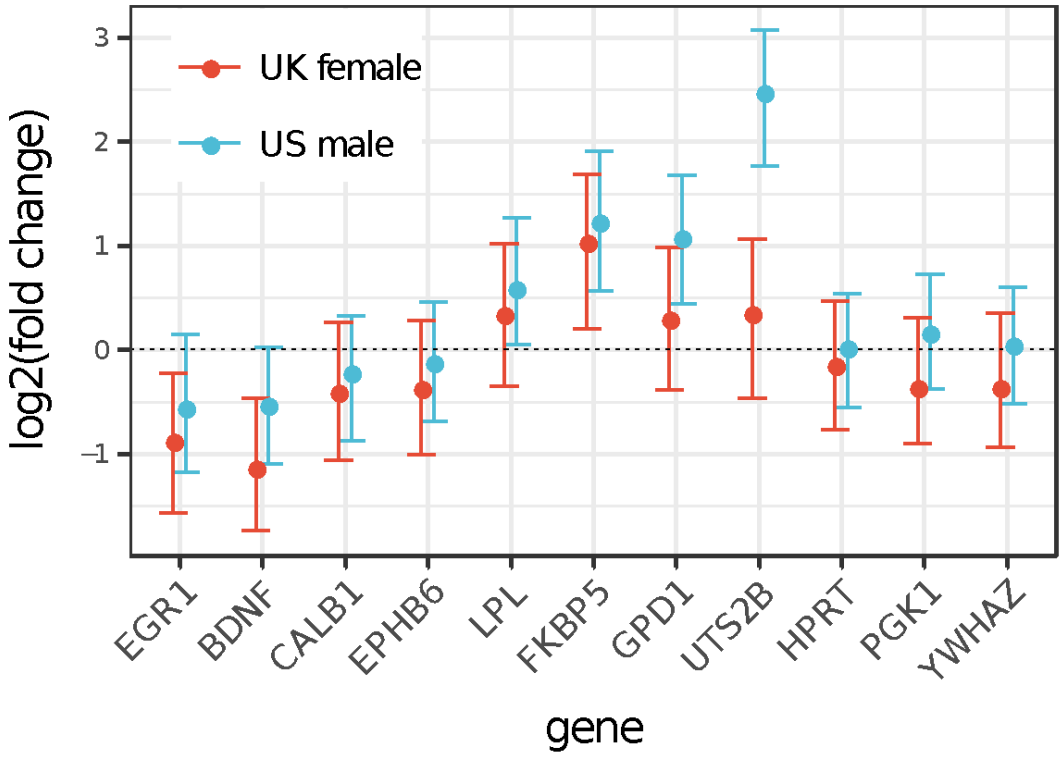

Figure S4. Validation of Primary RNA Sequencing Result by Quantitative RT PCR. Relative expression of specific RNAs in the AL from birds in isolation vs aviary conditions, for two separate experimental replications: (blue) from 12 males (6 aviary, 6 isolate) collected in the USA at the same time as the females used for the primary RNA-seq; (red) 10 females (5 aviary, 5 isolate) collected from a distinct population derived from stocks in the UK and maintained in a UK aviary. All gene targets were measured on replicate arrays of cDNA from these "US male" and "UK female" samples, and fold-changes in abundance were estimated by fitting a joint GLMM model over all targets using MCMC.qpcr (v 1.2.3) (9). Fixed effects of solo isolation (relative to aviary), with 95% credible intervals, are plotted from a naive model (no control genes specified). However, as expected, HPRT, PGK1, and YWHAZ are relatively stable in this experiment (fold-change near zero, with narrow credible intervals), while the targets selected for validation of solo response show response patterns consistent with the RNA-seq result, or trends in the same direction. It seems likely that the smaller sample size (n=5-6 per group) limited the power to confidently detect smaller fold changes and/or low abundance targets. We thus increased our group size and input cDNA quantities in subsequent experiments.

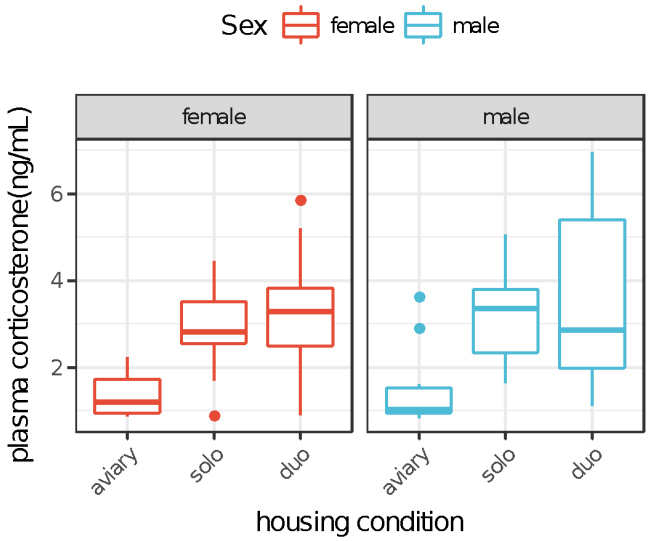

Figure S5. Blood corticosterone is increased after overnight isolation, with or without a partner. Radioimmunoassay was used to measure corticosterone in blood of birds (n=65) from the partner experiment (Figure 5). No difference was observed between males and females, or between solo and duo, but both solo and duo are elevated relative to aviary (female solo versus aviary, padj=0.012, female duo versus aviary, padj=0.0027; male solo versus aviary, padj=0.0084, male duo versus aviary, padj=0.0021).

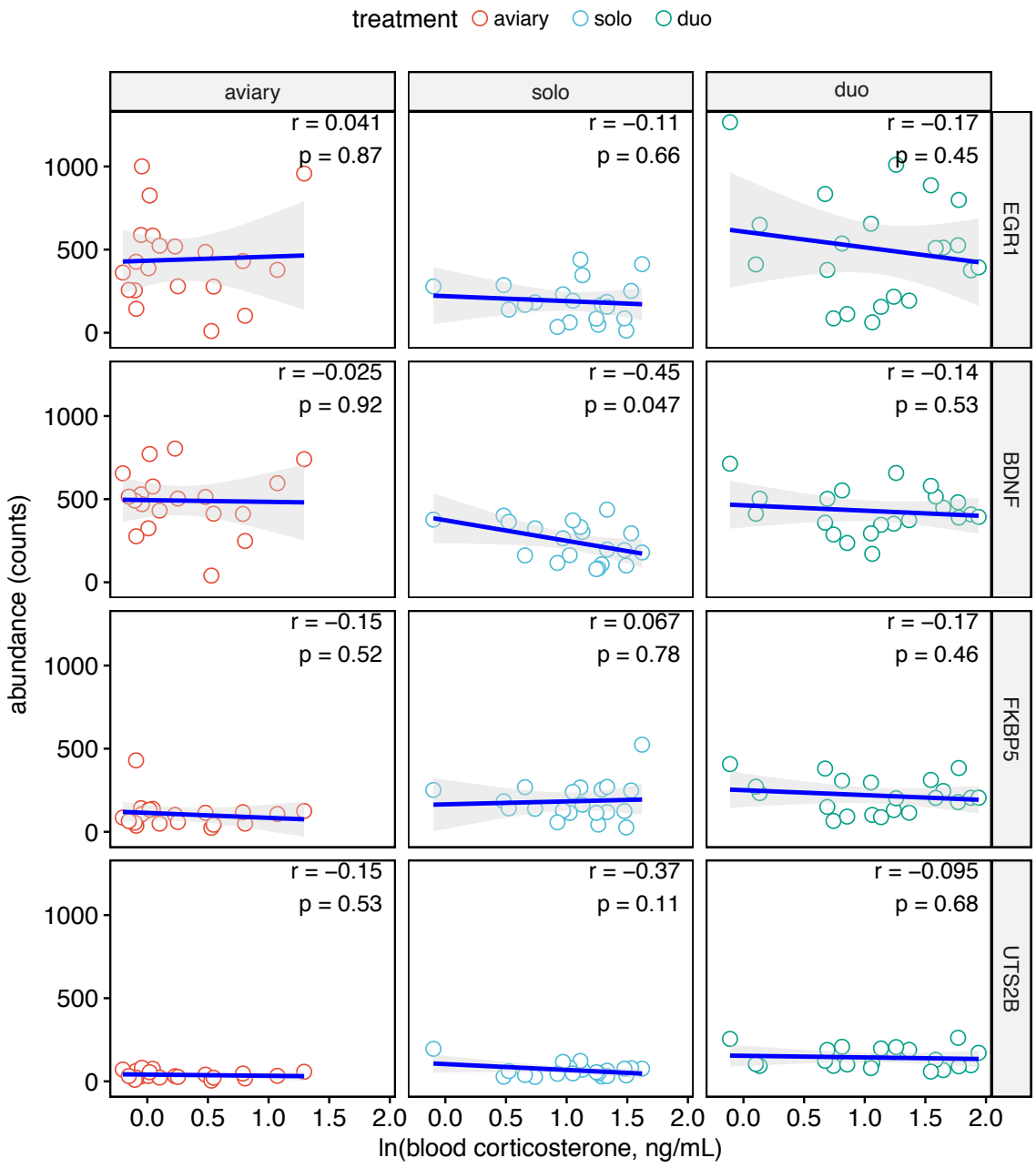

Figure S6. Blood corticosterone does not predict gene expression within treatment group. Gene expression of four focal genes is plotted versus plasma corticosterone, with Pearson coefficient (r) and corresponding p value for the correlation. CORT does not predict expression of FKBP5 or UTS2B within treatment groups, although all are elevated in solo and duo conditions relative to aviary. Plots were prepared with the R package ggpubr v. 0.1.7 (22).

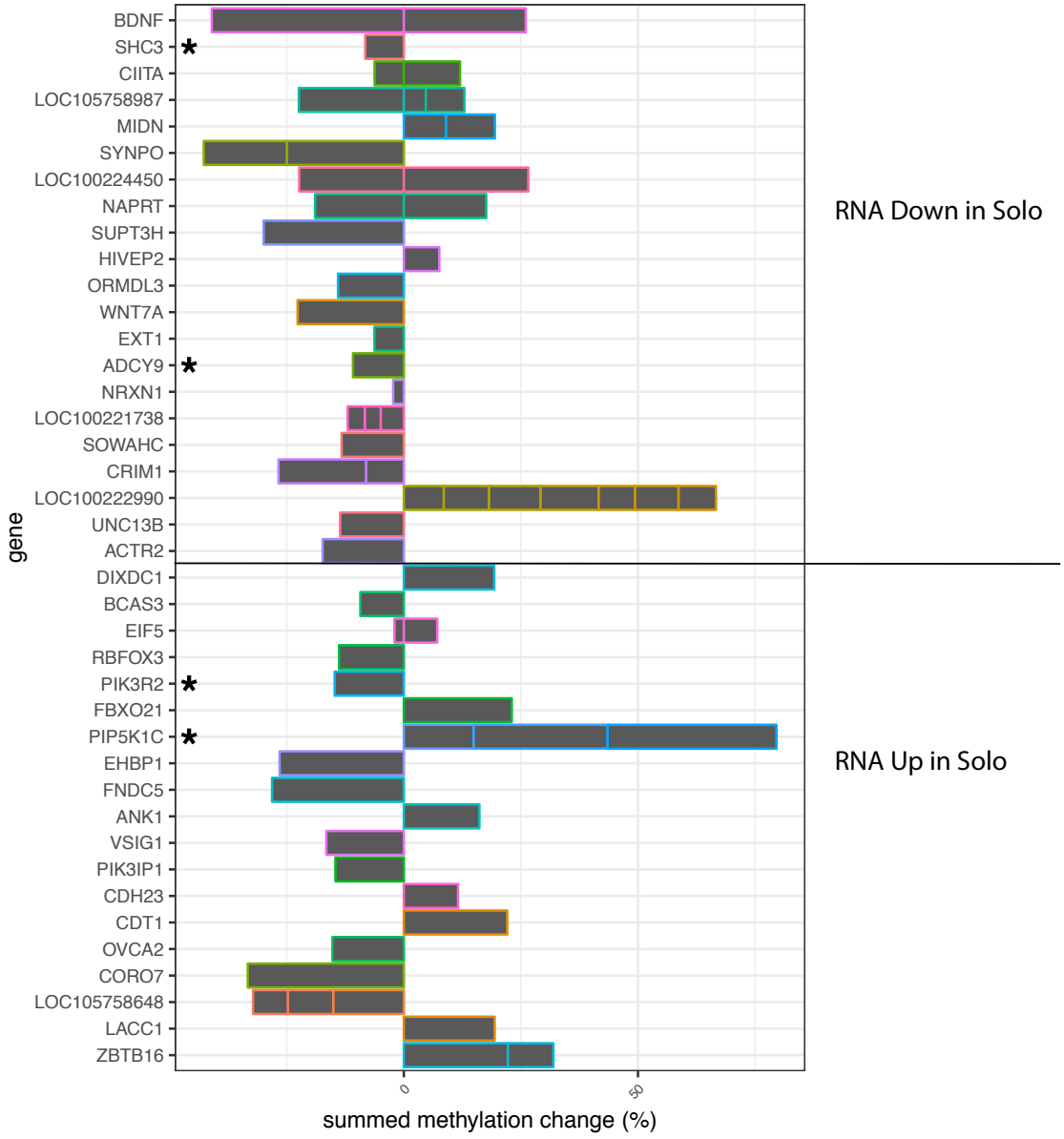

Figure S7. Differential Methylation of CpGs associated with Regulated Genes. Summary of significant CpG methylation (q<0.0001) for the set of genes determined to be differentially regulated (q<0.05) by RNA-Seq. Each differentially methylated CpG is depicted by a single block representing the percent methylation change at that site for solo versus aviary. Blocks are stacked to represent multiple significant CpGs per gene. Functional enrichment analysis was performed using g:Profiler (version e95_eg42_p13_f6e58b9) with human pathways as reference, applying g:SCS multiple testing correction method and significance threshold of 0.05 (23). Phospholipase D signaling pathway (KEGG:04072) was found to be enriched (padj=0.01); genes in this pathway are indicated by *.

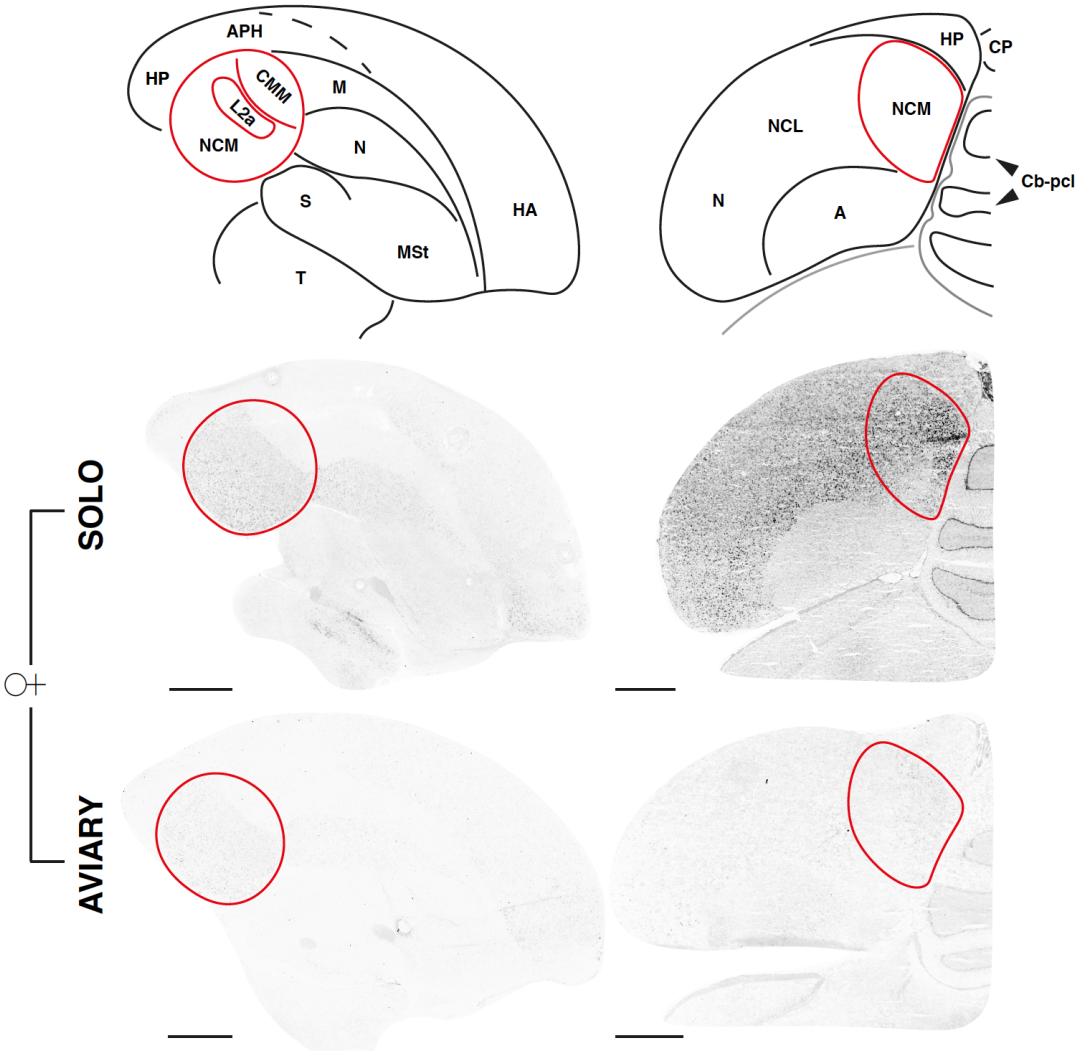

Figure S8. FKBP5 expression across the brain in SOLO vs AVIARY conditions. Shown are low power in situ hybridization images of adult female zebra finch brain sections in the sagittal (left column) and coronal (right column) planes, from birds after the AVIARY or SOLO conditions as in Figure 4. The area corresponding to AL is circled in red. Note generally higher expression in the SOLO condition throughout most of the forebrain (contrast with opposite response of EGR1 in Figure S5). Scale bar = 1 mm. NCM - caudomedial nidopallium; CMM - caudomedial mesopallium; L2a - L2a subfield of Field L; HP - hippocampus; APH - area parahippocampalis; M - mesopallium; N - nidopallium; S - septum; MSt - medial striatum; T - thalamus; HA – apical hyperpallium; NCL - caudolateral nidopallium; A - arcopallium; CP – choroid plexus; Cb-pcl - purkinje cell layer of the cerebellum

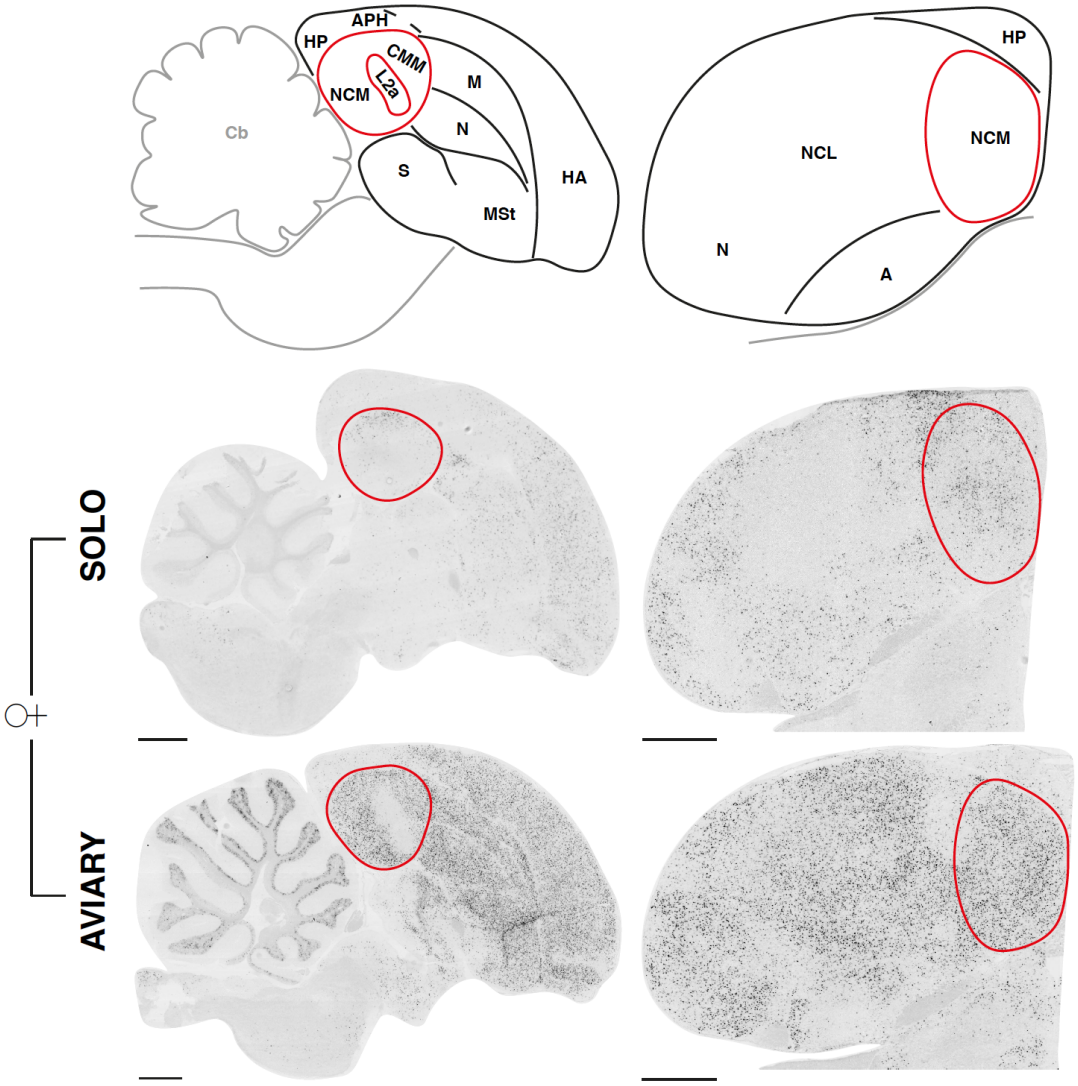

Figure S9. EGR1 expression across the brain in SOLO vs AVIARY conditions. Shown are low power in situ hybridization images of adult female zebra finch brain sections in the sagittal (left column) and coronal (right column) planes, from birds after the AVIARY or SOLO conditions as in Figure 4. The area corresponding to AL is circled in red. Note generally higher expression in the AVIARY condition throughout most of the forebrain (labeled in bold in the line drawings at top). Scale bar = 1 mm. NCM - caudomedial nidopallium; CMM - caudomedial mesopallium; L2a - L2a subfield of Field L; HP - hippocampus; APH - area parahippocampalis; M - mesopallium; N - nidopallium; S - septum; MSt - medial striatum; HA - apical hyperpallium; NCL - caudolateral nidopallium; A - arcopallium; Cb – cerebellum.

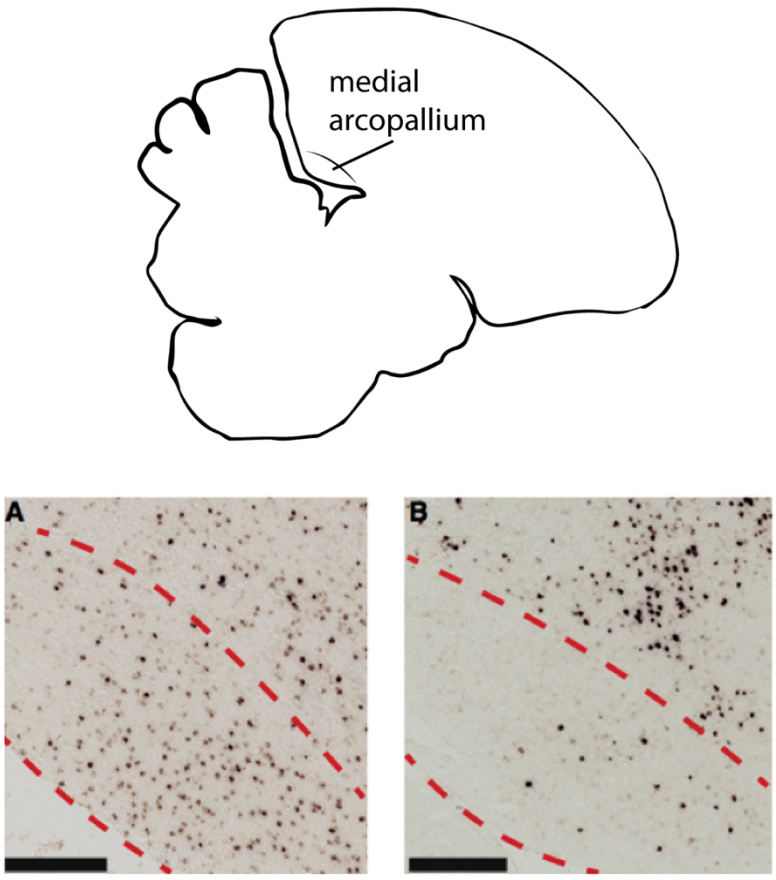

Figure S10. EGR1 mRNA expression in the medial arcopallium. In situ hybridization images of EGR1 expression for AVIARY (A) and SOLO (B) treatment conditions, in sagittal sections collected 1.2mm from the midline. The area inside the dashed lined corresponds to medial arcopallium, as described in (24), an area sharing molecular markers of mammalian cortex and amygdala. Scale bar = 250 μm.

Table S1. Primers and hydrolysis probes for RT-qPCR.

assays with DNA-binding dyes (SYBR-green)

primer set	Target	Forward primer	Reverse primer	Product length (bp)
EGR1 Olias	EGR1	ACTTCATCATCGCCATCCTC	TGGAATTGGGAAATGTTGGT	109
BDNF_1	BDNF	GGGTGACAGCAGCAGAGAAA	GACTGGGTAGTTCGGCACTG	199
CALB1_JG1	CALB1	TGGAGGAAATATGACAGCGACC	TGGCCAGTTCAGTAAGCTCC	187
EPHB6_JG2	EPHB6	CCATTGCCATCCTTGCCATC	GTGCACCGATATACTGCTGC	84
LPL_JG2	LPL	GTGCTCAGATGCCCTACAAAG	GATCAGGAATGGCTGGTTTGTC	95
FKBP5_JG2	FKBP5	GAGAAGGCCAAGGAGTCGTG	GCCCTTCTCCTTGACCACAG	75
GPD1_JG4	GPD1	AATGTCAAGTACCTGCCGGG	CATCCACGCCCTTGATGAGT	193
UTS2B_JG5	UTS2B	CTCAGCATCCCAGCCCAAAAA	TTCCAGATTAGCAGCACGATG	161
HPRT Olias	HPRT	GATGAACAAGGTTACGACCTGGA	TATAGCCACCCTTGAGTACACAGAG	181
YWHAZ Olias	YWHAZ	GTGGAGCAATCACAACAGGC	GCGTGCGTCTTTGTATGACTC	222 (1 mismatch)
PGK1 Olias	PGK1	AAAGTTCAGGATAAGATCCAGCTG	GCCATCAGGTCCTTGACAAT	167
HPRT.2	HPRT	GTTGGGTTGTGCTGTTTGTG	ATTACCACCCAAAGGTGCAA	222

assays with hydrolysis probes

primer set	Target	Forward primer	Reverse primer	Hydrolysis probe
BDNF_1	BDNF	GGGTGACAGCAGCAGAGAAA	GACTGGGTAGTTCGGCACTG	FAM-CCAAATGCAACCCCAAGGGG-BHQ1
UTS2B	UTS2B	CGTGGAGAAGATGTGGTCTG	GCACTCGGTTCTCTTGGAGT	FAM-CATGGCTCTGTGTGTCCCGTCA-TAMRA
PGK1 Olias	PGK1	AAAGTTCAGGATAAGATCCAGCTG	GCCATCAGGTCCTTGACAAT	HEX-GGTGGTGGGATGGCATTCACC-BHQ1
FKBP5_p2	FKBP5	CAAAGGCGAGGACTTGTT	CACGGTAGCACCTTCATTAG	FAM-CGGAGGATCAAGAGGAAGGGAGAAGGT -BHQ1
EGR1.2	EGR1	CTGGCTCCCAAGAACTGAAG	AAAGGTTTCTGTCCCGTGTG	HEX-GCGTCCCTATGCCTGCCCAG-BHQ1

Dataset S1 (separate file)

Excel spreadsheet with full results from DESeq2 analysis of differential gene expression in female auditory lobule after overnight isolation. This file contains data for 16628 genes which were initially detected at a threshold of >10 reads total across 23 samples. After independent filtering (conducted automatically in DESeq2 to increase statistical power), adjusted p-values (FDR) are reported for 14687 genes. baseMean represents expression in aviary samples. Log2FoldChange represents change in solo samples relative to baseMean, after shrinkage to moderate fold-changes of low abundance genes. lfcSE is the standard error, stat is the Wald statistic, False Discovery Rate (FDR) is calculated after the method of Benjamini and Hochberg (6).

.

References

1. Cheng H-Y, Clayton DF (2004) Activation and habituation of extracellular signal-regulated kinase phosphorylation in zebra finch auditory forebrain during song presentation. *Journal of Neuroscience* 24(34):7503–7513.

2. Bolger AM, Lohse M, Usadel B (2014) Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* 30(15):2114–2120.

3. Patro R, Duggal G, Love MI, Irizarry RA, Kingsford C (2017) Salmon provides fast and bias-aware quantification of transcript expression. *Nat Methods* 14:417.

4. Soneson C, Love MI, Robinson MD (2016) Differential analyses for RNA-seq: transcript-level estimates improve gene-level inferences. *F1000Res* 4. doi:10.12688/f1000research.7563.1.

5. Love MI, Huber W, Anders S (2014) Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol* 15(12):550.

6. Benjamini Y, Hochberg Y (1995) Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *J R Stat Soc Series B Stat Methodol* 57(1):289–300.

7. Ritchie ME, et al. (2015) limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res* 43(7):e47.

8. Olias P, Adam I, Meyer A, Scharff C, Gruber AD (2014) Reference genes for quantitative gene expression studies in multiple avian species. *PLoS One*. Available at: http://journals.plos.org/plosone/article?id=10.1371/journal.pone.0099678.

9. Matz MV, Wright RM, Scott JG (2013) No control genes required: Bayesian analysis of qRT-PCR data. *PLoS One* 8(8):e71448.

10. Wingfield JC, Farner DS (1975) The determination of five steroids in avian plasma by radioimmunoassay and competitive protein-binding. *Steroids* 26(3):311–321.

11. Hothorn T, Bretz F, Westfall P (2008) Simultaneous inference in general parametric models. *Biom J* 50(3):346–363.

12. Replogle K, et al. (2008) The Songbird Neurogenomics (SoNG) Initiative: community-based tools and strategies for study of brain gene function and evolution. *BMC Genomics* 9:131.

13. Ye J, et al. (2012) Primer-BLAST: a tool to design target-specific primers for polymerase chain reaction. *BMC Bioinformatics* 13:134.

14. Carleton JB, et al. (2014) An optimized protocol for high-throughput in situ hybridization of zebra finch brain. *Cold Spring Harb Protoc* 2014(12):1249–1258.

15. Karten HJ, et al. (2013) Digital atlas of the zebra finch (Taeniopygia guttata) brain: A high-resolution photo atlas. *J Comp Neurol* 521(16):3702–3715.

16. Olson CR, Hodges LK, Mello CV (2015) Dynamic gene expression in the song system of zebra finches during the song learning period. *Dev Neurobiol* 75(12):1315–1338.

17. Schneider CA, Rasband WS, Eliceiri KW (2012) NIH Image to ImageJ: 25 years of image analysis. *Nat Methods* 9(7):671–675.

18. Charlier TD, Ball GF, Balthazart J (2005) Sexual behavior activates the expression of the immediate early genes c-fos and Zenk (egr-1) in catecholaminergic neurons of male Japanese quail. *Neuroscience* 131(1):13–30.

19. Klughammer J, et al. (2015) Differential DNA methylation analysis without a reference genome. *Cell Rep* 13(11):2621–2633.

20. Xi Y, Li W (2009) BSMAP: whole genome bisulfite sequence MAPping program. *BMC Bioinformatics* 10:232.

21. Akalin A, et al. (2012) methylKit: a comprehensive R package for the analysis of genome-wide DNA methylation profiles. *Genome Biol* 13(10):R87.

22. Kassambara A (2017) ggpubr:"ggplot2" based publication ready plots. *R package version 0 1* 6.

23. Raudvere U, Kolberg L, Kuzmin I, Arak T, Adler P, Peterson H, Vilo J (2019). g:Profiler: a web server for functional enrichment analysis and conversions of gene lists. *Nuc Acids Research*. doi:10.1093/nar/gk369.

24. Mello CV, Kaser T, Buckner AA, Wirthlin M, Lovell PV (2019) Molecular architecture of the zebra finch arcopallium. *J Comp Neurol*. doi:10.1002/cne.24688.