Fig. 8. Vocal pathways of songbird brain. White lines indicate the anterior vocal pathway involved in vocal learning and social context production; black lines indicate the posterior vocal pathway involved in production of learned vocalizations; dashed lines indicate connections between the two pathways. Vocal nuclei are highlighted in light green, purple, and gray. Av, avalanche; DM, dorsal medial nucleus; DLM, dorsal lateral nucleus of the medial thalamus; E, entopallium; B, basorostralis; HVC, high vocal center or alternatively a letter based name; LLD, lateral lemniscus, dorsal nucleus; LMAN, lateral magnocellular nucleus of the anterior nidopallium; LAreaX, lateral AreaX of the striatum; LMO, lateral oval nucleus of the mesopallium; NIf, interfacial nucleus of the nidopallium; nXIIts, nucleus XII, tracheosyringeal part; PAm, paraambiguus; RAm, retroambiguus; RA, robust nucleus of the arcopallium; Uva, nucleus uvaeformis. Adapted from Jarvis ED, Gunturkun O, Bruce L, Csillag A, Karten H, Kuenzel W, Medina L, Paxinos G, Perkel DJ, Shimizu T, *et al.* (2005) *Nat Rev Neurosci* 6:151–159.

Fig. 9. Benefit of full-length clones for cross-species hybridizations, using the *Foxp1* gene as an example. (*Left*) Zebra finch *FoxP1* DNA and protein coding sequence showing 3' partial (blue box) and full-length (red box) cloned regions. (*Right*) *In situ* hybridizations (film autoradiographs) to parrot brain (frontal sections, left hemisphere) with riboprobes made from zebra finch 3' *FoxP1* partial and full-length cDNA clones. The 3' partial *FoxP1* clone did not cross-hybridize even though it contains the 3' region of the highly conserved protein coding sequence and the same probe hybridized to zebra finch brain sections (data not shown). The full-length zebra finch *FoxP1* clone cross-hybridized to parrot brain with results similar to that reported in Haesler S, Wada K, Nshdejan A, Morrisey EE, Lints T, Jarvis ED, Scharff C (2004) *J Neurosci* 24:3164–3175, showing, relative to hyperpallium (H) and nidopallium (N), enriched expression in the mesopallium (M; upper white labeled region) and striatum (St; lower white labeled region).

Fig. 10. Example sonograms of songs from birds used for the rapid vocal learning brain states (Table 1). (*Top*) Songs of an untutored juvenile bird, which is relatively variable. (*Middle*) Songs of a tutored bird, showing change over 8 days where the song becomes more stereotyped and begins to match the tutored bird (the four song motifs at PH48). (*Bottom*) The tutor's song, SAMBA (4.5 s, *x* axis; *y* axis is frequency in kilohertz).

Fig. 11. Distribution of cDNA sizes among libraries. For each library, 70–90 randomly selected PCR products were run on ethidium bromide-stained 1% agarose gels. The sizes in kilobases were then measured and plotted against percent clones of each size in 0.5-kb bins.

Fig. 12. Example of a clone page from the songbird brain transcriptome database. Shown are parts of the page for H3 histone, Family 3B, a singing-regulated gene. (*A*) Cluster, subcluster, and clone hierarchy. (*B*) Primer annotation information summed from all clones. (*C*) Graphical displays of the three top BLAST hits among the different databases searched. (*D*) National Center for Biotechnology Informatio (NCBI) Unigene/Gene homolog annotations and the gene ontology functions based on the BLAST results. (*C*) Consensus clone sequence from the 5' and 3' reads. (*E*) Phred quality scores across the consensus clone sequence (color-coded as in Table 2).

Fig. 13. Homology analyses of zebra finch cDNAs to another avian species (chicken) and a mammal (human). The *y* axis shows the percent of clusters that have at least one clone with identity hits of 100-bp or more, at >70%, 85%, and 95% identity, of 4,738 clusters total. These numbers are underestimates, because the coding sequence of cDNAs >1.8 kb will be underrepresented (we sequenced ≈0.9 kb of the ends), and coding sequence usually has higher identities across species. The database sources are NCBI nucleotide, Biotechnology and Biological Sciences Research Council (BBSRC) chicken transcriptome, and the Institute for Genome Research (TIGR) chicken EST (URLs in *Supporting Text* section 3.6.6).

Fig. 14. Comparison of microarray signals of singing-regulated genes under various conditions. (*A*) Brain slice, 120 μm, before (*Aa*) and after (*Ba*) dissection of vocal nuclei. (*Ab*) Two quadrants of a microarray slide hybridized to probes from the entire forebrain of a singing (Cy5, red) versus silent (Cy3, green) animal. (*Bb*) The same quadrants on duplicate slides but hybridized with probes made from HVCs of a singing and a silent animal. For the microarray of *Ab*, the entire telencephalon, not just a section, was used as a source of probe for the hybridization. Arrowheads indicate genes up-regulated by singing; arrows indicate genes down-regulated by singing. The down-regulated ones were the result of individual differences in HVC, but not other brain areas, among animals, as verified by *in situ* hybridization (data not shown and see Fig. 15). (*C*) Apparent ratio (fold difference) in singing-driven gene expression in HVC of immediate early genes (*egr-1*, *c-fos*, *c-jun*, *BDNF*, and *Arc*) as a function of their DNA concentration on the microarray slide. The higher the DNA concentration, the higher the ratio difference of probes from the same animal, indicating that DNA concentration artificially affects the fold induction

measurement. (*D*) Ratio differences of dye signal on 3X SSC spots without DNA, showing that replicate hybridizations (*n* = 3) have variations. Ideally, each replicate should have SSC ratios (Cy3 to Cy5) equally distributed around 1. The replicate 3X SSC spots did not result in significant differences of *P* < 0.2, *t* test.

Fig. 15. Sample *in situ* hybridization of a false-positive gene (*Sfrs1*) that showed differences in AreaX among birds but not between silent and singing (0.5–3 h) groups. Shown are inverse images of film autoradiographs from eight different animals. White arrow indicates AreaX with lower expression; yellow arrow indicates AreaX with higher expression. The source of the individual animal variation is not known as yet.

Fig. 16. Time course of singing-regulated expression of all 33 genes in four song nuclei. The genes are categorized into one of six types of temporal expression patterns based on expression peak time (0.5, 1, or 3 h) and expression profile. Expression levels were measured as pixel density on x-ray film minus background on the film. The criterion for including a gene as singing-regulated was that it had to have a significant difference at one or more time points relative to silent controls (0 h; ANOVA by post hoc probable least-squares difference (PSLD) test. *, *P* < 0.05; **, *P* < 0.01; *n* = 4 each time point and *n* = 5 at 0.5 h). The criterion for placing genes in a category was that the gene had to have or not have significant differences in expression between the 1–3-h time points (ANOVA; *P* < 0.05, specific values not shown); this second analysis revealed six such categories.

Fig. 17. Viral vector expression. (*A*) Viral vector used to express full-length zebra finch cDNAs *in vivo*. The gene of interest was amplified in a nested PCR reaction with primers that recognize both viral vector and zebra finch sequences. The 5' UTR and 3' UTR regions were not included in the amplified region, because we needed to keep the insert size as small as possible. Large inserts (>2 kb) are thought not to work well in lentiviral vectors. The particular vector shown is that with the Ubiquitin C promoter, obtained from Pavel Osten (Max Planck Institute, Heidelberg, Germany), which was modified from that of Carlos Lois (Massachusetts Institute of Tecnology, Cambridge, MA). (*B*) Western blots of protein from transfected HEK293T cells reacted with either anti-FLAG-tag or Met enkephalin antibodies. The FLAG tag of the enkephalin insert was apparently cleaved off in the processing of proenkephalin to the mature molecule. The HEK293T cells do not (−) express either of these molecules. (*C*) Table of animals (21 total) and number of injections performed (42 total) of different lentiviral constructs. First column, different constructs

tested; second column, number of animals tested with each construct; remaining columns, the site of injection and time of sacrifice after injection. (*D*) UbiC-GFP lentivirus expression in AreaX at different time points after injection (sagittal sections, anterior is to the right). (*E*) Quantification of GFP fluorescent intensity in 60 cells of AreaX per brain (*n* = 2, 3, and 1 animal). There were no significant differences across time (*P* > 0.05, ANOVA by Fisher PSLD post hoc test).

Fig. 18. Example of cross-species microarray hybridization. Shown are hybridizations to the zebra finch microarrays with brain probes from male (Cy5, red) versus female (Cy3, green) zebra finches or canaries. The identified genes (arrows) will be reported on separately.

Fig. 19. Outline for our high-throughput molecular neuroethological approach. Prior knowledge is used to include brain states that induce behaviorally regulated genes. cDNA libraries and sequence information are used to generate cDNA arrays and an annotated database, which are then used to select candidate cDNAs for gene manipulation and thus to determine their causal and functional roles in cells and in behavior, respectively. Arrows indicate the direction of resource use.

Fig. 20. Plasmid vector used for zebra finch brain cDNA libraries. pFLC-I means plasmid full-length cloning vector I. The forward (Fwd) and reverse (Rev) primer sites were used for sequencing the 5' and 3' ends, respectively, of the inserted cDNA and also for amplifying inserts for microarray production. The T7 and T3 promoters were used to make sense and antisense RNA. The SfiI restriction site is rare and can thus be used to cut out the insert with a low probability of cutting into the insert. The avian library 5' primer is specific to the libraries made with avian RNA at RIKEN. The NNNNNN has the unique 3' primer ID for each bird. More details are in Jarvis ED, Smith VA, Wada K, Rivas MV, McElroy M, Smulders TV, Carninci P, Hayashizaki Y, Dietrich F, Wu X, *et al*. (2002) *J Comp Physiol A* 188:961–980.

Fig. 21. Examples of loss of plasmids with full-length cDNAs that depend on the amount of growth time in primary and secondary cultures. Shown is an ethidium bromide-stained agarose gel. Although most clones have vector$^+$ inserts after 20 h of primary culture growth, 8–12 h of primary growth prevents loss of clones on secondary inoculations (examples *A*, *B*, and *C*). Not shown is that some clones were even lost after 20 h of primary culture growth. M, 1-kb molecular weight marker. Plasmid samples were restriction-digested with BamHI.