

Table 1. Animal states used to construct libraries

Bird ID	Individual animal brain states	3' ID	Category	Used for libraries
1	E10 male	ACCGTT	Embryonic development (sex verified by PCR)	0057
2	E10 female	TATCTT		0057
3	E15 male	TTGATT		0057
4	E15 female	CACGGT		0057
5	PH1 male	GGTCGT		0057
6	PH1 female	GCACGT		0057
7	PH5 male	CTGAGT	Post hatch development (sex verified by PCR)	0058, 0060, 0071
8	PH5 female	GTCGCT		0058, 0060, 0071
9	PH10 male	GAGCCT		0058, 0060, 0071
10	PH10 female	ATTCAAT		0058, 0060, 0071
11	PH15 male	CTCTTG		0058, 0060, 0071
12	PH15 female	AAGGTG		0058, 0060, 0071
13	PH25 male silence	GGACTG	Sensory acquisition phase	0058, 0060, 0071
14	PH25 female silence	ACTATG		0058, 0060, 0071
15	PH25 male hearing 2hr	AGCTGG		0058, 0060, 0071
16	PH25 female hearing 2hr	TAACGG		0058, 0060, 0071
17	PH35 female hearing in aviary 2hr	GTTAGG	Subsong phase (verified subsong prior day)	0058, 0060, 0071
18	PH35 (subsong singer) male silence	TCCTCG		0058, 0060, 0071
19	PH38 male undirected subsong & hearing adult song 2hr	ATACCG	Early plastic song phase (verified plastic song prior day)	0058, 0060, 0071
20	PH55 female hearing in aviary 2hr	CATACG		0058, 0060, 0071
21	PH55 (early plastic song singer) male silence	CGGACG		0058, 0060, 0071
22	PH54 male early plastic undirected singing & hearing adult song 2hr	GACTAG	Late plastic song phase (verified plastic song prior day)	0058, 0060, 0071
23	PH71 female hearing in aviary 2hr	CCACAG		0058, 0060, 0071
24	PH77 (late plastic song singer) male silence	TGTAAG		0058, 0060, 0071
25	PH76 male late plastic undirected singing & hearing adult song 2hr	ATGGTC		0058, 0060, 0071
26	PH88 (puberty) female hearing in aviary 2hr	TACCTC	Crystallization phase (verified crystallized song prior day)	0058, 0060, 0071, 0067, 0061
27	PH86 (crystallization-puberty) male silence	ACACTC		0058, 0060, 0071, 0066, 0061, 0062, 0064
28	PH86 male crystallized undirected singing & hearing adult song 2hr	AGTATC		0058, 0060, 0071
29	Adult female silence, 2hr morning in soundbox	TGATGC	Adult vocal communication (experienced breeding birds)	0058, 0060, 0071
30	Adult male silence in soundbox 2hr morning	TCTAGC		0058, 0060, 0071
31	Adult male undirected singing & hearing adult song 30 min	CTTACC		0058, 0060, 0071
32	Adult male directed singing & hearing adult song 30 min	GAGTTA		0058, 0060, 0071
33	Adult female calling & hearing adult male song 2hr	ACGGTA		0058, 0060, 0071
34	Adult male undirected singing & hearing adult song 2hr	AATTGA		0058, 0060, 0071, 0068, 0062, 0065
35	Adult male undirected singing & hearing adult song 6hr	TGGTGA		0058, 0060, 0071, 0068, 0062, 0065
36	Adult 3 years old male singing undirected song 1hr	TAGAAT		0058, 0060, 0071
37	Adult male sleeping 2hr	ATCGAA	Old Sleep	0058, 0060, 0071
38	Adult male kainate treatment, 30min post 1st seizure	CTAGGA	Seizure and receptor activation	0058, 0060, 0071
39	Adult male kainate treatment, 2hr post 1st seizure	CGTTCA		0058, 0060, 0071
40	Adult male kainate treatment, 6hr post 1st seizure	TACGCA		0058, 0060, 0071
41	Adult male apomorphin treatment, 30min post 1st seizure	TCACCA		0058, 0060, 0071
42	Adult male apomorphin treatment, 2hr post 1st seizure	GCTTAA		0058, 0060, 0071
43	Adult male deafened	CAGCAA	Deaf	0058, 0060, 0071
44	Adult 2.9 years old male isolated from PHD24 until adult	AGACAA	Isolation	0058, 0060, 0071
45	PH67 male isolated from PHD25, silence	CCTAGT		0058, 0060, 0071
46	PH67 male isolated from PHD25, hearing song first time, 30 min	TCGTAT		0058, 0060, 0071
47	PH120 male visually isolated from PHD25	CAGTCT		0058, 0060, 0071, 0064
48	PH40 male rapid vocal learning-singing, afternoon of 1st day	GGCACA	Rapid vocal learning (verified vocal learning in real time)	0058, 0060, 0071
49	PH52 male rapid vocal learning-singing, morning of 4th day	GTAGAC		0058, 0060, 0071
50	PH47 male rapid vocal learning-singing, evening of 2nd day	TTAGCC		0058, 0060, 0071, 0069, 0064, 0063, 0065
51	PH38 male no motor learning yet, 1st day	CAAGCG		0058, 0060, 0071
52	PH64 male rapid vocal learning-singing, afternoon of 7th day	GCCTGT		0058, 0060, 0071
53	PH48 male rapid vocal learning-singing, midday of 8th day	GCGAGA		0058, 0060, 0071
54	PH51 male rapid vocal learning-singing, evening of 2nd day	GCACAC		0058, 0060, 0071
55	PH44 male rapid vocal learning-singing, evening of 3rd day	CTCCTA		0058, 0060, 0071
56	PH37 male no training, singing isolate subsong	GTCTCC	0058, 0060, 0071, 0070	
57	Adult male silence overnight plus 1hr in morning	AGGTGT	Silent and singing (used for initial library test)	0031, 0054
58	Adult male directed singing to females 30min	AGGTGT		0033
59	Adult male directed singing to females 30min	AGGTGT		0033
60	Adult male undirected singing in aviary 40min	AGGTGT		0052, 0056

The numbers in the left-most column indicate the number ID for each bird. Red text, age; blue text, sex; green text, amount of time stimulated or performing the specified behavior; other colored text, categories of brain states. The unique 3'-ID is the 6-bp sequence of the 3' primer used for each individual bird; the exception is that samples from birds 57-60 were used with the same primer.

Table 2. Zebra finch brain cDNA libraries

	<b>Lib</b>	<b>T</b>	<b>Condition</b>	<b>Bird ID</b>
1	0031	N	Silent male	57
2	0033	N	Directed singing male	58
3	0052	N	Undirected singing males	59
4	0054	A	Silent male	57
5	0055	A	Directed singing male	58
6	0056	A	Undirected singing males	59
7	0057	N	Embryonic males and females	1 to 6
8	0058	N	Juveniles/adults 50 states pooled	7 to 56
9	0059	N	Rapid vocal learning males	48 to 56
10	0060	A	Juveniles/adults 50 states pooled	7 to 56
11	0061	S	Silent male-Hearing female	27-26
12	0062	S	Adult singing males-Silent male	34+35-27
13	0063	S	Rapid vocal learning-Untutored	50-56
14	0064	S	Rapid vocal learning-Adult silent	50-27
15	0065	S	Rapid vocal learning-Adult singing	50-34+35
16	0066	R	Silent male	27
17	0067	R	Hearing female	26
18	0068	R	Adult singing males	34+35
19	0069	R	Rapid vocal learning male	50
20	0070	R	Untutored male	56
21	0071	R	Juveniles/adults 50 states pooled	7 to 56

Library (Lib) ID designated at the Duke University Center for Genome Technology. Libraries types (T) are N, normalized; A, abundant; S, subtracted; and R, regular. Bird ID, brain samples (Table 1) used to make the libraries. Silent condition means bird was not singing. Hearing condition means bird heard song playbacks.

Table 3. Quality distribution of a total of the 17,490,205 sequenced bp in the database

<b>Phred score</b>	<b>Error probability</b>	<b>Accuracy, %</b>	<b>No. of bp in database</b>	<b>Percentage of total</b>
<b>1-10</b>	<b>1 in 10</b>	<b>90</b>	377,214	<b>2.1</b>
<b>11-20</b>	<b>1 in 100</b>	<b>99</b>	1,108,410	<b>6.3</b>
<b>21-30</b>	<b>1 in 1,000</b>	<b>99.9</b>	1,704,845	<b>9.6</b>
<b>31-40</b>	<b>1 in 10,000</b>	<b>99.99</b>	2,992,367	<b>11.6</b>
<b>41-50</b>	<b>1 in 100,000</b>	<b>99.999</b>	5,171,875	<b>30.4</b>
<b>&gt;51</b>	<b>1 in 1,000,000</b>	<b>99.9999</b>	6,135,494	<b>40.1</b>
<b>Total</b>			17,490,205	<b>100</b>

Color-coding is the same as quality graphs in the database (Fig. 12E). Only 2.1% of the sequences are low in quality (scores 1-10), and these passed trimming because they were within stretches of regions of high quality or were at the 5' and 3' primer ends, where trimming would have shortened the zebra finch-specific sequences.

Table 4. Microarray hybridizations

Song nucleus	Replicate 1	Replicate 2 (swap)	Replicate 3	Replicate 4
HVC	Green/red Silent/singing (~202 bouts), $n = 3$ each group, pooled	Red/green Silent/singing (~202 bouts), $n = 3$ each group, pooled	Red/green Silent/singing (208 bouts), $n = 1$ each group	Poor-quality hyb
RA	Green/red Silent/singing (~202 bouts), $n = 3$ each group, pooled	Red/green Silent/singing (~202 bouts), $n = 3$ each group, pooled	Green/red Silent/singing (208 bouts), $n = 1$ each group	Red/green Silent/singing (151 bouts), $n = 1$ each group
LMAN	Green/red Silent/singing (~202 bouts), $n = 3$ each group, pooled	Red/green Silent/singing (~202 bouts), $n = 3$ each group, pooled	Poor-quality hyb	Red/green Silent/singing (151 bouts), $n = 1$ each group
LAreaX	Green/red Silent/singing (~202 bouts), $n = 3$ each group, pooled	Red/green Silent/singing (~202 bouts), $n = 3$ each group, pooled	Red/green Silent/singing (208 bouts), $n = 1$ each group	Green/red Silent/singing (151 bouts), $n = 1$ each group

Different dye combinations were used for each of the 14 array hybridizations. In replicate 1 three animals per group were used, where dissections of silent and singing animals were separately pooled; in replicate 2, a dye swap of replicate 1 was performed; in replicates 3 and 4 one animal per group.  $n = 5$  silent and  $n = 5$  singing animals (1 h). Green, Cy3 label; Red, Cy5. Hyb, hybridization.

Table 5. List of 43 genes examined (41 + *egr-1* and *c-fos*), color-coded according to Fig. 2

	Gene name	Clone ID
	<b>Positives:</b>	
1	<b>Early growth response 1 (<i>egr-1</i>)</b>	Positive control
2	<b>Cellular-fos (<i>c-fos</i>), variant 2</b>	0033P0001H05
3	<b>Cellular-jun (<i>c-jun</i>), variant 1</b>	0063P0031G06
4	<b>Similar to jun B (<i>sim junB</i>)</b>	0058P0018G01
5	Activating transcription factor 4 ( <i>ATF4</i> ), variant 1	0065P0006B05
6	<b>Heat shock protein binding protein 1 (<i>Hsbp1</i>), variant 2</b>	0065P0001A03
7	Ubiquitin-conjugating E2 variant 1 ( <i>Ube2v1</i> )	0061P0003C04
8	Heterogeneous nuclear ribonucleoprotein H3 ( <i>HnrpH3</i> ), variant 1	0064P0008B02
9	H3 histone, family 3B ( <i>H3f3B</i> ), variant 1	0058P0006E06
10	H2A histone, family X ( <i>H2AFX</i> ), variant 1	0061P0021G07
11	Split hand/foot deleted gene 1, short clone ( <i>Shfdg1</i> ), polyA	0064P0001B06
12	Mothers against DPP homolog 2 ( <i>Madh2</i> )	0064P0014A01
13	Activity regulated cytoskeletal-associated protein ( <i>Arc</i> )	0065P0027D05
14	<b>Similar to formin-like protein (<i>fmln</i>), variant 2</b>	0058P0025D01
15	<b>Transgelin 2 (<i>Tagln2</i>)</b>	0064P0018E08
16	$\beta$ -actin, variant 1	0061P0009E09
17	<b>Cdc42 guanine nucleotide exchange factor 9 (<i>ARHGEF9</i>), variant 1</b>	0063P0031G05
18	Protein kinase, cAMP-dependent regulatory 1 $\alpha$ ( <i>Prkar1<math>\alpha</math></i> ), variant 3	0058P0061C03
19	<b>Growth arrest &amp; DNA-damage-inducible 45<math>\beta</math> (<i>Gadd45<math>\beta</math></i>), variant 1</b>	0061P0027C11
20	<b>Heat shock protein 40 (<i>Hsp40</i>)</b>	0058P0018H06
21	<b>Heat shock 70kDa protein 8 (<i>Hsp70-8</i>), variant 2</b>	0065P0027D04
22	<b>Heat shock protein 90<math>\alpha</math> (<i>Hsp90<math>\alpha</math></i>)</b>	0065P0018D07
23	<b>Heat shock protein 25 (<i>Hsp25</i>)</b>	0061P0030G03
24	<b>Calcyclin binding protein (<i>Cacybp</i>)</b>	0061P0010H02
25	START domain containing 7 ( <i>Stard7</i> ), variant 1	0058P0032C10
26	Synaptotagmin IV ( <i>Syt4</i> )	0065P0009F07
27	ATPase, H <sup>+</sup> transporting, V1 subunit B, isoform 2 ( <i>Atp6v1b2</i> )	0061P0014D07
28	NADH dehydrogenase (ubiquinone), 1 alpha subcomplex, 5 ( <i>Ndufa5</i> )	0061P0030A03
29	Estrogen receptor binding site associated, antigen, 9 ( <i>Ebag9</i> ), variant 1	0063P0010C08
30	<b>Jun-suppressed chemokine (<i>JSC</i>)</b>	0063P0015E10
31	Brain derived neurotrophic factor ( <i>BDNF</i> )	0065P0005E12
32	<b>Proenkephalin (<i>Penk</i>)</b>	0064P0014A11
33	Similar to NPD014 protein ( <i>sim NPD014</i> ), variant 2	0064P0014F06
	<b>Varied across birds</b>	
34	Heat shock protein 108 ( <i>Hsp108</i> )	0065P0027B02
35	Similar to stem cell antigen 2, variant 2	0058P0051B11
36	Engulfment and cell motility 1, variant 2	0063P0024G05
37	Similar to ribonuclease P	0058P0022A07
38	Vesicle associated membrane protein ( <i>Vamp</i> ), variant 3	0058P0035G01
39	Splicing factor, arginine/serine-rich 1 ( <i>Sfrs1</i> ), variant 1	0062P0003B01
	<b>False Positives:</b>	
40	Heat shock 70kDa protein 5 ( <i>Hsp70-5</i> ), variant 1	0058P0059A06
41	Similar to corticotropin releasing hormone	0058P0017H12
42	Zinc finger, MYND domain containing 11 ( <i>Zmynd11</i> )	0065P0026G12
43	Synaptosomal associated protein-25 ( <i>Snap25</i> )	0064P0015A03

Bold texts represent those clones that show concordance of the *in situ* hybridization results with the reanalyzed microarray data using GeneSpring *t* test with a P value cut-off of 0.2.

Table 6. Proportional representation of singing-regulated genes among the various libraries

mRNA	Clone	Norm, %	Subtr, %	0061, %	0062, %	0063, %	0064, %	0065, %
egr-1	1	0	100	0	0	100	0	0
c-fos	4	100	0	0	0	0	0	0
<b>c-jun, v1</b>	<b>10</b>	<b>10</b>	<b>90</b>	<b>10</b>	<b>20</b>	<b>40</b>	10	10
sim junB	1	100	0	0	0	0	0	0
<b>Atf4, v1</b>	<b>14</b>	<b>0</b>	<b>100</b>	<b>36</b>	0	<b>29</b>	14	21
Hspb1, v2	2	0	100	0	0	0	0	100
UbE2v1	2	0	100	100	0	0	0	0
<b>HnrpH3, v1</b>	<b>6</b>	33	66	17	0	17	17	17
<b>H3f3B, v1</b>	<b>11</b>	45	55	27	0	18	<b>0</b>	9
<b>H2AfX</b>	<b>10</b>	<b>20</b>	<b>80</b>	<b>60</b>	0	<b>0</b>	<b>0</b>	20
<b>Shfdg1</b>	3	33	66	0	0	0	33	33
Madh2	1	0	100	0	0	0	100	0
Arc	1	0	100	0	0	0	0	100
Fmn1, v2	1	100	0	0	0	0	0	0
Tagln2	1	0	100	0	0	0	100	0
$\beta$ -actin	3	0	100	66	0	0	0	33
ARHGEF9	3	0	100	33	0	66	0	0
Prkar1 $\alpha$ , □□	2	50	50	0	0	0	0	50
Gadd45 $\beta$ , □□	1	0	100	100	0	0	0	0
Hsp40	3	66	33	33	0	0	0	0
<b>Hsp70-8, v2</b>	<b>20</b>	<b>50</b>	<b>50</b>	<b>5</b>	<b>20</b>	10	5	10
<b>Hsp90<math>\alpha</math></b>	<b>32</b>	<b>0</b>	<b>100</b>	22	0	<b>31</b>	<b>22</b>	<b>25</b>
<b>Hsp25</b>	<b>13</b>	46	54	<b>54</b>	0	<b>0</b>	<b>0</b>	<b>0</b>
<b>Cacybp</b>	<b>12</b>	42	58	<b>33</b>	0	8	8	8
Stard7, v1	4	100	0	0	0	0	0	0
Syt4	1	0	100	0	0	0	0	100
<b>Atp6v1b2</b>	<b>10</b>	40	60	20	0	10	<b>20</b>	10
Ndufa5	2	50	50	50	0	0	0	0
Ebag9	5	80	20	0	0	20	0	0
<b>JSC</b>	<b>8</b>	<b>0</b>	<b>100</b>	<b>12.5</b>	0	<b>62.5</b>	12.5	<b>12.5</b>
BDNF	1	0	100	0	0	0	0	100
Penk	2	50	50	0	0	0	50	0
<b>sim NPD014, v2</b>	<b>8</b>	<b>12.5</b>	<b>87.5</b>	<b>0</b>	0	<b>37.5</b>	<b>25</b>	<b>25</b>
Observed (avg 198), %		<b>31</b>	<b>69</b>	21	1	14	13	21
Expected if random, %		41	59	20	3	15	10	12
Total clones	13,269	5,413	7,856	2,608	349	2,054	1,309	1,536

Clones: number of mRNA/cDNA clones from each gene in the database among 13,269 clones with good sequence quality in the libraries listed. Percentage is the percent distribution of these clones among normalized (0058) versus subtracted (0061-0065) libraries combined (fourth column) or distribution among the subtracted libraries 0061-0065 individually (fifth and sixth columns). The expected if random percentage is the percentage that is expected to occur if cDNAs were represented randomly among libraries. Red text, mRNAs (subclusters) with  $n = 6$  or more clones in the database. Bold text, Percentage of clones isolated from a given library (left columns) or set of libraries (subtracted or normalized) that had greater than +10% or -10% difference from the expected percentage.