MICRORNA GENE EXPRESSION IN THE ZEBRA FINCH BRAIN

BY

YA-CHI LIN

DISSERTATION

Submitted in partial fulfillment of the requirements
for the degree of Doctor of Philosophy in Cell and Developmental Biology
in the Graduate College of the
University of Illinois at Urbana-Champaign, 2012

Urbana, Illinois

Doctoral Committee:

Professor Mary A. Schuler, Chair
Professor David F. Clayton, Director of dissertation research
Professor Gene E. Robinson
Associated Professor Stephanie S. Ceman
Assistant Professor Kannanganattu V. Prasanth
ABSTRACT

Songbirds such as zebra finches communicate via learned vocalizations (songs) and studies have shown that experiencing song playback triggers complex genomic responses in the zebra finch auditory forebrain. MicroRNAs (miRNAs or miRs) are important regulators of gene expression which may coordinate complex biological processes through post-transcriptional mechanisms. This dissertation aims to explore the potential roles of miRs in the genomic response to song.

This study began with a bioinformatic analysis of published microarray data and qPCR analysis of a specific conserved miR (miR-124) in zebra finch auditory forebrain, elements of which contributed to the primary paper describing the zebra finch genome (Warren et al. 2010). These preliminary studies are described in the Introduction to this thesis. Chapter 2 then presents a full de novo characterization of miRs in the songbird brain and demonstrates that song exposure has effects on several. This has now been published as Gunaratne, Lin et al. 2011 (co-first authors). A significant outcome of Chapter 2 was the identification of a novel sex-linked miR, miR-2954. Chapter 3 describes the tissue, cellular and subcellular distribution of miR-2954 and localizes it to subsets of cells in the brain. An antisense inhibitor of miR-2954 was then applied to a zebra finch cell line followed by RNA-seq analysis to test the hypothesis that changes in miR-2954 levels lead to changes in the network of genes expressed. The results confirm this hypothesis and suggest that the initial song-induced decline in miR-2954 expression described in Chapter 2 may help reprogram gene expression networks to support the metabolic changes associated with song habituation (Dong et al., 2009). This thesis research helps better understand the transcriptome of songbird brain and establishes novel roles for microRNAs in song perception, discrimination and memory.
ACKNOWLEDGEMENTS

First, I would like to thank my advisor Dr. David Clayton for his guidance, support and encouragement on my thesis study. It is my significant fortune having the academic training under Dr. Clayton’s mentoring, and learning to think independently and to work cooperatively in the team work. I especially appreciate Dr. Clayton’s great patience to teach me knowledge about songbirds and assistance to make this thesis project always on the right track.

Thanks to my committee members Dr. Mary Schuler, Dr. Gene Robinson, Dr. Stephanie Ceman and Dr. Kannanganattu Prasanth for their advice, support and suggestions. I would also like to thank Dr. Sarah London and Dr. Chris Balakrishnan who always provide helpful suggestions and especially thank Dr. Preethi Gunaratne for providing help on RNA sequencing at the early stage of my thesis project. Thanks to Dr. Julie George and Lisa Stubbs for their technical support.

Thanks to all of my friends and family for being supportive over these important years in my life. I especially want to thank Dr. Sunny Sun for being both a mentor and friend who advises and encourages me all the way during this journey in America. Finally, I would like to show great appreciation to my family. This USA journey would not be completed without their support and encouragement. More importantly, thanks to my husband who has always been there for me no matter good or bad time.
TABLE OF CONTENTS

CHAPTER 1 INTRODUCTION ......................................................................................1
  NON-CODING RNAs .................................................................................................2
  MicroRNAs ..............................................................................................................4
  MicroRNA Targets ..................................................................................................8
  Sex-Linked MicroRNAs ..........................................................................................11
  Zebra Finch Sex Chromosomes and Sex Differences in Behavior ..........................13
  Preliminary Studies .................................................................................................14
  Goals of This Thesis ...............................................................................................15
  Figure ......................................................................................................................17

CHAPTER 2 Song Exposure Regulates Known and Novel MicroRNAs in the Zebra Finch Auditory Forebrain ..............18
  Abstract ..................................................................................................................18
  Background .............................................................................................................19
  Results ....................................................................................................................21
  Discussion ...............................................................................................................28
  Conclusions ............................................................................................................32
  Methods ..................................................................................................................33
  Authors' Contributions ..........................................................................................38
  Acknowledgements ................................................................................................38
  Note Added in Proof ................................................................................................39
  Tables and Figures ................................................................................................40

CHAPTER 3 A Gene Network Regulated by Mir-2954, a Song-Responsive Z-Linked MicroRNA Expressed in a Subset of Zebra Finch Brain Cells ..................................................56
  Abstract ..................................................................................................................56
  Introduction .............................................................................................................57
  Results .....................................................................................................................59
  Discussion ...............................................................................................................64
  Materials and Methods ..........................................................................................70
  Acknowledgements ................................................................................................74
  Tables and Figures ................................................................................................75
CHAPTER 1 INTRODUCTION

Zebra finches are songbirds, which are one of a few animals capable of vocal learning (Jarvis 2004; Williams 2004). Learned vocalizations in birds (birdsong) and in humans (speech) have many analogues and suggest similar neural mechanisms may have evolved in both animals (Doupe and Kuhl 1999; Jarvis 2004). Learning and memory of vocalizations (songs) are essential for songbirds to perform many social behaviors, such as mate recognition (Miller 1979) and territory defense (Clayton 1988). In the zebra finch, only the male learns to sing and each adult male has his unique song and uses it for vocal communication with other zebra finches, however both sexes process and discriminate specific songs (Clayton 1987; Clayton 1988; Brenowitz, Margoliash et al. 1997; Stripling, Kruse et al. 2001). Adult male zebra finches have been shown to discriminate songs from other conspecific individuals implying the formation of memories of particular songs (Mello, Nottebohm et al. 1995; Stripling, Milewski et al. 2003). The molecular and cellular mechanisms of song perception, discrimination and memory are of great interest as a general model for how the brain processes social experience.

Early studies using molecular biology identified several immediate early genes (IEGs) such as zenk and c-jun (Mello, Vicario et al. 1992; Mello and Clayton 1994; Nastiuk, Mello et al. 1994) that are induced by song stimulation in the auditory forebrain lobule (AL) of songbird brain. More recent efforts in the collaborative community of songbird neurogenomics (e.g., Songbird Neurogenomics (SoNG) Initiative, Replogle et al., Wada et al., Li et al.) have established resources to support high-throughput analyses on a larger scale, and are revealing complex interactions among many more genomic components engaged by song experience. Expressed sequence tags (ESTs) from the ESTIMA database were used to produce a cDNA microarray, which has been used to identify thousands of
song-regulated genes (Replogle, Arnold et al. 2008; Dong, Replogle et al. 2009). In the group of novel-song regulated ESTs, Gene Ontology (GO) analysis indicated significant enrichment for genes associated with transcription factor activity, carbohydrate metabolic process, RNA metabolic process and nuclease activity; ion channel activity and Notch signaling pathway are significant in the ESTs that are decreased by song (Dong, Replogle et al. 2009). However, a surprisingly large number of song-responsive ESTs could not be associated with any protein-coding gene based on the initial annotation of Replogle et al., which looked for similarities to chicken genes in the chicken International Protein Index (IPI) database (Replogle et al., 2008). This is especially true for the set of RNAs that decreased immediately after novel song exposure, as only 34% of them could be linked to a chicken ortholog (Dong et al., 2009).

These findings suggested that non-coding RNAs might play a major role in the brain’s response to novel experience, and provided the primary motivation for the research described in this thesis. In the remainder of this introduction, I first provide a general background on non-coding RNAs with specific emphasis on the subclass known as microRNAs. I then describe my own preliminary studies that validated the existence of song-responsive non-coding RNAs and microRNAs and contributed to the large multi-authored manuscript describing the zebra finch genome assembly (Warren et al., 2010). Finally I outline the specific goals of my own thesis research.

**NON-CODING RNAS**

Genes for non-coding RNA (ncRNA) encode functional RNA molecules instead of proteins, and produce a larger portion of total genomic output compared to protein-coding genes (Eddy 2001; Mattick 2001). NcRNAs can be classified as either house-keeping or
regulatory based on their functionality, or long versus short according to their size (Prasanth and Spector 2007; Brosnan and Voinnet 2009). Examples of the house-keeping ncRNAs include ribosomal RNAs (rRNAs), transfer RNAs (tRNAs), small nuclear RNAs (snRNAs) etc., which are constitutively expressed and highly abundant. Examples of regulatory ncRNAs include Xist (X-inactive-specific transcript) and aHIF (HIF-1α antisense transcript), and are typically expressed at specific developmental stages or in response to external stimuli. Another example of a regulatory ncRNA is encoded at the MHM (male hypermethylated) locus in the genomes of chicken and turkey; it is involved in the mechanism of local dosage compensation along the Z chromosome (Teranishi, Shimada et al. 2001; Bisoni, Batlle-Morera et al. 2005; Itoh, Kampf et al. 2011). Unlike in chicken and turkey, the MHM ncRNA sequence is absent in the genome assembly of zebra finch, suggesting that this regional Z-chromosome dosage compensation mechanism is not common to all bird species (Itoh, Replogle et al. 2010; Warren, Clayton et al. 2010).

Specific efforts have been directed to identify long conserved noncoding sequences (LCNS) that are in mammalian and avian genomes and expressed as RNAs (Sakuraba, Kimura et al. 2008; Janes, Chapus et al. 2011). LCNS are defined as sequences that are longer than 500 bp with more than 95% identity between two or more species. Zebra finch and chicken share far more LCNS than do human and mouse, however only a small fraction of the avian-specific LCNS (97/4294, 2.3%) overlaps with the brain expressed ESTs from zebra finch. 28 of the LCNS expressed in zebra finch brain are regulated in response to song playback stimuli (Dong, Replogle et al. 2009; Janes, Chapus et al. 2011). These song-regulated LCNS may play roles of binding for transcription factors when the zebra finch experiences different social behaviors, although this function remains hypothetical.
MICRORNAS

MicroRNAs (miRNAs or miRs) comprise a family of ncRNAs that are distinguished both by their distinctive small size and regulatory functions in many biological processes and are the primary focus of the present dissertation. The first miRNA, lin-4, was discovered by the group led by Victor Ambros in 1993 (Lee, Feinbaum et al. 1993). It negatively regulates the protein-coding gene lin-14 during the developmental stages of C. elegans, by binding (despite imperfect complementarity) to the 3’UTR of lin-14 mRNA as demonstrated by Gary Ruvkun and colleagues at that time (Wightman, Ha et al. 1993).

Lin-4 and lin-14 are conserved only in closely related nematodes and therefore the generality of lin-4-like small regulatory RNAs in other species was questioned. However, seven years after lin-4 was discovered, a second miRNA, let-7, was found also in C. elegans (Reinhart, Slack et al. 2000). Let-7 is highly conserved from worm all the way to human, and in the year following let-7 discovery, over 100 miRNAs were reported in C. elegans, Drosophila and human (Lagos-Quintana, Rauhut et al. 2001; Lau, Lim et al. 2001; Lee and Ambros 2001). Currently, the microRNA database miRBase Release 17 contains 16772 hairpin precursor miRNAs, expressing 19724 mature miRNA products, in 153 species (http://www.mirbase.org/).

Transcriptions of miRNA genes are carried out mainly by RNA polymerase II (Lee, Kim et al. 2004). The primary transcript (pri-miRNA) is formed, capped and polyadenylated in the nucleus and then processed by the RNase III endonuclease Drosha to generate ~60-70 nucleotides hairpin precursor (pre-miRNA). The pre-miRNA is transported from the nucleus to the cytoplasm by Ran-GTP and the export receptor Exportin-5, and then cleaved by the RNase III endonuclease Dicer to yield ~21-22 base pairs RNA-duplex. One strand (miRNA) of the RNA-duplex is selected to be incorporated
into the miRNA-induced silencing complex (miRISC, or called miRNP), and the other strand (miRNA*) generally is degraded (Bartel 2004; Schanen and Li 2011). The mature miRNA can guide the miRISC onto its target and regulate translation or mRNA turnover at the post-transcriptional level (Bartel 2004; Filipowicz, Bhattacharyya et al. 2008; Wu and Belasco 2008; Chekulaeva and Filipowicz 2009). The hypothesized models of miRNA-mediated translational repression include inhibition of translation initiation by competition between miRISC and eIF4E for cap binding or by preventing 60S ribosomal subunit from association with 40S ribosomal subunit; suppression of translation elongation by stalling the elongating ribosome or causing premature termination; and proteolytic digestion of nascent polypeptides (Filipowicz, Bhattacharyya et al. 2008; Wu and Belasco 2008; Chekulaeva and Filipowicz 2009). Although initial studies showed that miRNAs affected their targets primarily at the protein translation level, later microarray studies have revealed that some miRNAs can also accelerate degradation of their target mRNAs (Bagga, Bracht et al. 2005; Krutzfeldt, Rajewsky et al. 2005; Lim, Lau et al. 2005; Giraldez, Mishima et al. 2006). The mechanisms of miRNA-mediated mRNA instability include removal of 3’ poly(A) tail (deadenylation) followed by 5’ decapping then exonuclease decay when the miRNA is partially complementary with the target (mostly observed in animal cells); and endonucleolytic cleavage when the miRNA is near-perfectly or fully complementary with the target (Filipowicz, Bhattacharyya et al. 2008; Wu and Belasco 2008; Chekulaeva and Filipowicz 2009). Although miRNAs have been reported generally to repress their targets, recent results have shown that miRNAs can activate their targets under certain conditions, i.e. miRNAs act as translation enhancer under the condition of starvation (Vasudevan and Steitz 2007; Vasudevan, Tong et al. 2007; Henke, Goergen et al. 2008; Jopling 2008; Orom, Nielsen et al. 2008).


**Genomic studies of miRNAs**

The developments of cloning strategies, next generation sequencing and computational approaches make it easier to discover miRNA genes at the genomic scale (Ambros 2004; Bartel 2004; Berezikov, Cuppen et al. 2006). Some miRNAs are transcribed as polycistronic transcripts (clustered miRNAs, multiple miRNAs in one pri-miRNA transcript) and the frequency of such genes varies in different species (Griffiths-Jones, Saini et al. 2008). The miRNAs within the same cluster frequently belong to different miRNA families but have related functions by targeting the same gene or the genes in the same pathway (Kim and Nam 2006; Olena and Patton 2010). Relative to other known transcription units, most vertebrate miRNA genes (~50% of mammalian miRNAs) are located in the intronic regions in the sense orientation of the host pre-mRNA (Rodriguez, Griffiths-Jones et al. 2004; Kim and Nam 2006; Griffiths-Jones, Saini et al. 2008; Olena and Patton 2010). The intronic miRNAs and their host transcripts usually have highly-correlated expression patterns, implying that these miRNAs are originated from the same primary transcripts as their host genes and under control of the common promoter or regulatory elements (Rodriguez, Griffiths-Jones et al. 2004; Baskerville and Bartel 2005; Wang, Lu et al. 2009).

Identification of chicken miRNAs has focused on embryonic development, using bioinformatic prediction, direct cDNA cloning and deep sequencing approaches (Hillier, Miller et al. 2004; Glazov, Cottee et al. 2008; Shao, Zhou et al. 2008). A total of 121 miRNA gene families in the chicken genome were first identified by different computational methods including finding sequence similarity to known pre-miRNAs in human, mouse and rat, potential to form hairpin secondary structure and alignment to mature ~22 nucleotides miRNAs allowing 2 or fewer mismatches (Hillier, Miller et al.
Twenty-seven chicken miRNA clusters were determined and nearly all of them are conserved within vertebrates. The expression profiling of chicken miRNAs correlates well with the developmental stages; diverse chicken miRNAs are detected at later periods of embryogenesis and organogenesis, but almost no miRNAs are detected early in embryonic development (Shao, Zhou et al. 2008). Glazov et al. used deep sequencing to register 488 new miRNAs of the developing chicken embryo including 39 pre-miRNA/intron (mirtron) sequences which are pre-miRNAs in short introns generated from alternative miRNA biogenesis utilizing splicing enzyme instead of Drosha (Berezikov, Chung et al. 2007; Okamura, Hagen et al. 2007; Ruby, Jan et al. 2007; Glazov, Cottee et al. 2008). Only a small number of these newly identified chicken miRNAs are conserved in non-avian vertebrates, suggesting the majority are probably avian or chicken specific miRNAs and may be involved in regulation of lineage-specific developmental programs (Glazov, Cottee et al. 2008).

**MiRNAs in brain**

MiRNAs in brain have been documented and are especially diverse, comprising up to approximately 70% of experimentally detected miRNAs (Cao, Yeo et al. 2006). Even though many brain miRNA sequences are conserved between species, the expression patterns in different species are divergent (Ason, Darnell et al. 2006; Berezikov, Thuemmler et al. 2006; Bak, Silahtaroglu et al. 2008). The high variety of miRNAs and the distinct expression profiles in brain are suggested to contribute to the complex functions of brain. Brain expressed miRNAs are involved in numerous neural processes and brain diseases (Kuss and Chen 2008; Coolen and Bally-Cuif 2009; Saba and Schratt 2010). One of the signature miRNAs in the nervous system is miR-124 which sequence is conserved from *C. elegans* to humans and an abundant miRNA in brain (Lagos-Quintana, Rauhut et al.)
Mir-124 has been linked to neurogenesis, neuronal differentiation, and neurite outgrowth (Lim, Lau et al. 2005; Visvanathan, Lee et al. 2007; Yu, Chung et al. 2008; Cheng, Pastrana et al. 2009; Yoo, Staahl et al. 2009). More than 20 miRNAs have been identified to localize at synapses, and may regulate synaptic protein synthesis and function and thereby modulate synapse development and plasticity (Ashraf and Kunes 2006; Schratt 2009; Smalheiser and Lugli 2009). Extending previously mentioned functional involvements, miR-124 has been shown to target CREB1, regulating the switch between short-term and long-term synaptic plasticity in Aplysia (Rajasethupathy, Tuschi et al. 2009).

MICRORNA TARGETS

One essential step toward understanding the regulatory functions of miRNAs is to identify their targets. The first case of a confirmed miRNA and target pair was developed by molecular genetic analyses in C. elegans (Lee, Feinbaum et al. 1993; Wightman, Ha et al. 1993). However, hundreds of miRNAs in animals have been identified by cloning and computation, and only a few of the hundreds of miRNAs have been assigned to previously known functions. Finding targets for the large number of miRNAs cannot be done only by genetics, hence it is necessary for the systematic identification of miRNA targets with the help of bioinformatics. Unlike plant miRNAs, which usually regulate their targets through extensive base pairing with the target transcripts, the non-extensive complementarity between animal miRNAs and targets becomes the main challenge to predict the targets for animal miRNAs with high specificity or without too many false positives (Bartel 2009). To improve the feasibility to do the computational prediction with good sensitivity and specificity, diverse computational algorithms have been developed and the criteria for a
candidate miRNA target often include strong pairing at the seed region (defined in the next section), conservation of miRNA binding sites among different species, favorable minimum free energy and structural accessibility of mRNA.

**Interaction of miRNAs and targets**

Before bioinformatic prediction of miRNA targets, only three miRNAs with experimentally validated targets had been described: the *lin-4* miRNA targets *lin-14* and *lin-28*, *let-7* targets *lin-28* and *lin41*, and *bantam* targets *hid* (Stark, Brennecke et al. 2003; Bartel 2009). General observations from these miRNA and target pairs are the two blocks of short sequences of the mature miRNA with different levels of base paring to the 3’UTR of the target and the multiple miRNA binding sites on the 3’UTR of the target (Lee, Feinbaum et al. 1993; Wightman, Ha et al. 1993; Reinhart, Slack et al. 2000; Brennecke, Hipfner et al. 2003). The first block of sequences from the second to the eighth nucleotides at the 5’ end of the mature miRNA (2-8 nucleotides), called the seed region, mediates perfect and contiguous base pairing to the 3’UTR of the target. The second block at the 3’ half of the mature miRNA tolerates mismatches and bulges, although good base pairing between miRNA 13-16 nucleotides and the 3’UTR of the target determines target specificity within miRNA families and enhances efficacy of miRNA targeting (Brennecke, Stark et al. 2005; Grimson, Farh et al. 2007; Filipowicz, Bhattacharyya et al. 2008). Other site contexts on the 3’UTR of the target transcript that affect the interaction of miRNAs and targets include the sequence content near the target site, and the numbers and positions of miRNA binding sites along the 3’UTR. The features of the 3’UTR context that can increase site effectiveness include AU-rich nucleotide composition near the site or other factors which make the target region less structured and more accessible to miRNP recognition, proximity to sites for coexpressed miRNAs which leads to cooperative action, location
within 3’UTR at least 15 nucleotides away from the termination codon, and location away from the center of long UTRs (Grimson, Farh et al. 2007; Filipowicz, Bhattacharyya et al. 2008; Bartel 2009).

**Algorithms to identify miRNA targets**

The computational algorithms for predicting miRNA targets in worms, flies and mammals have been developed by four different groups: the EMBL algorithm by the Cohen group, the miRanda algorithm by the Enright group, the TargetScan/TargetScanS by the Bartel group, and the PicTar algorithm by the Rajewsky group (Enright, John et al. 2003; Lewis, Shih et al. 2003; Stark, Brennecke et al. 2003; Krek, Grun et al. 2005). Although all methods use similar criteria considering seed pairing, site conservation and thermodynamic stability, the predictions are not 100% overlapping. Small differences in the algorithms, slightly different parameters in the alignments, or the use of different 3’UTR or miRNA sequences can produce a great diversity in target predictions (Rajewsky 2006; Bartel 2009; Li, Xu et al. 2010). Along with the algorithms are web-based databases for target predictions in the genomes of commonly studied species (such as human, mouse, Drosophila, or C. elegans). The miRanda and TargetScan also provide customized Perl scripts for target predictions in the newly-sequenced genomes.

**Validation of predicted targets**

More than 60% of human protein coding genes are predicted to be miRNA targets (Friedman, Farh et al. 2009). Analysis in eight vertebrate genomes has predicted that each miRNA can target approximately 200 transcripts on average (Krek, Grun et al. 2005). Small number of predicted targets can be validated by comparing with the previously validated targets or reporter assays (Enright, John et al. 2003; Lewis, Shih et al. 2003; Stark, Brennecke et al. 2003); however to validate the predicted targets by high throughput means
was a challenge before advanced proteomic or transcriptomic technologies were developed (Bentwich 2005). The strategies of profiling the transcriptome or the proteome after miRNA depletion or overexpression have been used to validate targets for individual miRNAs. Microarray, mass spectrometry or SILAC (stable isotope labeling with amino acids in cell culture, which is a quantitative-mass-spectrometry-based approach) have been used for transcriptome or proteome profiling (Lim, Lau et al. 2005; Linsley, Schelter et al. 2007; Baek, Villen et al. 2008; Selbach, Schwanhausser et al. 2008). Other strategies are immunoprecipitation of the miRISC components then microarray profiling or deep sequencing to identify miRISC-associated miRNAs and mRNAs (Beitzinger, Peters et al. 2007; Easow, Teleman et al. 2007; Karginov, Conaco et al. 2007; Zhang, Hammell et al. 2009), or high-throughput sequencing of RNAs isolated by crosslinking immunoprecipitation (HITS-CLIP) to identify the protein-miRNA or protein-mRNA binding sites (Chi, Zang et al. 2009).

**SEX-LINKED MICRORNAS**

The miRNA cluster on the X chromosome in primates includes 10 miRNAs (spanning ~100 kb) which are specifically expressed in testis. The homologous clusters in dog, mouse and rat have less numbers of miRNAs implying this X-linked miRNA cluster has been through gene expansion during primate evolution (Bentwich, Avniel et al. 2005). Frequent copy number changes and sequence substitutions indicate rapid evolution of the X-linked miRNA cluster in primates (Zhang, Peng et al. 2007). Another miRNA cluster on the primate X chromosome, with 6 miRNAs (spanning ~33 kb) which are predominantly expressed in human epididymis, has also been identified and evolutionary analysis has revealed complex evolutionary dynamics on this miRNA cluster (Li, Liu et al. 2010).
above miRNA clusters have not been found in amphibian, fish and bird species suggesting they emerged relatively recently or the miRNA sequences have diverged too far to recognize orthology.

Using a PCR-based method to detect the expression of 77 X-linked miRNAs in 12 mouse organs, Song et al. reported that 58% of the X-linked miRNAs are expressed ubiquitously and the remaining portion are expressed preferentially or specifically in testis. They further assayed the expression levels of the X-linked miRNAs at all stages of spermatogenesis and found that nearly all X-linked miRNAs exhibit continuous expression during male meiosis, in contrast to the protein-coding genes on the X chromosome which generally lose transcription during male meiosis (Song, Ro et al. 2009). Both analyses of miRNA transcriptomes in mouse prepubertal testis and newborn ovary showed the enrichment of miRNAs encoded from chromosome 2 and X, but the significance of the preferential expression from certain chromosomes is not clear (Ahn, Morin et al. 2010; Buchold, Coarfa et al. 2010).

Mutation screening of 13 brain-expressed X-linked miRNA genes in patients with mental retardation has identified four rare nucleotide changes in three pre-miRNA sequences; all four changes are outside of the mature miRNA sequences and appear to be functionally neutral (Chen, Jensen et al. 2007). Another screening for alterations on 59 miRNA genes on the X chromosome in patients with schizophrenia has revealed the first statistically significant association between microRNA mutant alleles and schizophrenia (Feng, Sun et al. 2009).
ZEbra Finch Sex Chromosomes and Sex Differences in Behavior

In contrast to mammals, male birds have homogametic sex chromosomes (ZZ) while female birds have heterogametic chromosomes (ZW) and global dosage compensation is generally not present in birds (Ellegren, Hultin-Rosenberg et al. 2007; Itoh, Melamed et al. 2007; Itoh, Replogle et al. 2010; Naurin, Hansson et al. 2011; Wolf and Bryk 2011). Microarray studies in the brain of zebra finch have shown that most genes on the Z chromosome are expressed at higher levels in males than in females (Itoh, Melamed et al. 2007; Itoh, Replogle et al. 2010). The higher expression of the Z-linked genes in male versus female zebra finches suggests that the Z-linked genes might be more likely to have evolved a role in controlling sexual differentiation or contributing sexual dimorphism in their behaviors.

Both sexes of zebra finches produce calls that are used in different behavioral contexts such as nest building and sexual behavior. However, only male zebra finches produce learned vocalizations (songs) (Zann 1996). This behavioral sexual dimorphism is reflected in the difference of the forebrain neural song circuit which is much larger in males (Arnold 2004). Attempts to reverse sexual development by administering male hormones to females or blocking testicular hormones in males have only been partially successful (Arnold 1997; Holloway and Clayton 2001; Arnold 2004), suggesting other factors such as male-biased Z gene expression may be potential triggers of neural and behavioral sex differentiation.
PRELIMINARY STUDIES

MiRNAs in the zebra finch

Prior to the work of this thesis, knowledge about miRNAs in the brain of zebra finch was sparse. There was only one report, and it focused on songbird brain development (Li, Wang et al. 2007). Li et al. searched cDNA libraries from zebra finch brains at consecutive developmental stages from embryo to post hatching day 90, and found seven EST sequences which are highly homologous with five miRNAs previously known in other animals. Zebra finches reach sexual maturity after post hatching day 90. Therefore, knowledge about the miRNA expression in the brain of adult zebra finch was absent. Besides, only conserved miRNAs can be identified by looking for sequence similarity and novel miRNAs specifically in the lineage of zebra finch were unexamined.

In my own preliminary studies leading to this thesis research, I manually evaluated 20 of the unmapped “novel down” ESTs from the microarray study of Dong et al. (2009, discussed above), using the UCSC Genome browser to probe their position in the initial internal release of the zebra finch genome assembly. I was able to map 100% of these to the assembly and found that 11 of these fell within putative introns of conserved genes, and 9 of them appeared to be intergenic. These findings motivated a fuller analysis of the complement of non-coding RNAs in the zebra finch brain. My analysis contributed to Table 1 in Warren et al (2010), which formally described the large number of non-coding RNAs (ncRNAs) expressed in the brain and regulated by song from the microarray study of Dong et al. (2009). These findings also motivated a direct PCR analysis of a candidate miRNA, miR-124, following song stimulation; this experiment (which I performed) is described in Figure 3 in Warren et al (2010), and is presented here as Figure 1.1 in this thesis. This experiment showed that song exposure can suppress miR-124 expression in the
auditory forebrain lobule of zebra finch, and validated the comprehensive analyses of song-responsive miRNAs described in this thesis.

GOALS OF THIS THESIS

The key hypothesis of my thesis is that miRNAs play a role in coordinating the brain’s response to experience, using the model of song recognition learning in the zebra finch. The main question this study intends to answer is this: are miRNAs involved in the genomic response to song stimulation, and if so, how? To test the hypothesis of the present thesis, the following three specific aims are investigated:

1. **To develop a comprehensive identification of the microRNAs expressed in the zebra finch auditory forebrain.** In the collaborative study in which I have participated (Gunaratne, Lin et al. 2011), we have cataloged conserved and putative novel miRNA sequences expressed in the auditory forebrain lobule of adult zebra finch’s brain by RNA sequencing and bioinformatic analyses.

2. **To test whether any miRNAs change in response to song playback, and thus may participate in the complex transcriptional network that is modulated by song playback.** Additional measurements by performing TaqMan qPCR have validated the differential expression of the song-responsive miRNAs from RNA sequencing analysis. For one novel miRNA of great interest, tgu-miR-2954, we have predicted its candidate targets and to validate these predicted targets is included in the next functional study.

3. **To test for the functional significance of song-responsive miR-2954 by assessing expression in the different sexes, mapping expression at both regional and intracellular levels, and testing for effects of targeted manipulation on gene expression using cultured cells.** The miRNA tgu-miR-2954 located on Z chromosome is
intriguing. Does miR-2954 express in a sex-biased pattern similar to many Z chromosome genes in zebra finch? Is this miRNA expressed in tissues other than the auditory forebrain? Does it express in specific cell type of brain or in particular subcellular localization? The experiments in Chapter 3 aim to investigate the sex differences, tissue specificity and subcellular distribution of miR-2954 expression. TaqMan qPCR and fluorescence in situ hybridization are applied for obtaining the miRNA expression pattern.

What biological function does miR-2954 have? What are the target mRNAs of miR-2954? Although I performed a bioinformatic analyses to predict mRNAs targeted by miR-2954, I encountered a major limitation in that the zebra finch genome lacks complete 3’UTR annotation (Chapter 2). Nevertheless, through the strategy of considering conservation between chicken and zebra finch, I was able to predict eight putative miR-2954 target candidates (details in Chapter 2). We probably have many false negative targets given the hypothesis that one miRNA is estimated to target hundreds of mRNAs in other vertebrates. Can we identify the false negative targets? The experiments of RNA sequencing (RNA-seq) to measure mRNA changes after manipulating miR-2954 expression in zebra finch cell lines have been carried out. I envision these experiments as a precursor to the longer term goal of manipulating miRNAs in the whole animal.
Figure 1.1. Song response of miR-124 by TaqMan qPCR.
Measurement of miR-124 expression has been done in (A) two pool samples and (B) 12 individual samples. Silence samples are in open bar; samples from birds hearing 30 minutes song are in black filled bars. Error bars show SEM in technical triplicates. Comparison of silence and song in 12 individual samples by Wilcoxon paired t test shows significant change of miR-124 expression (p value = 0.03). Parallel TaqMan analyses of the small RNA RNU6B were performed with all samples and showed no significant effect of treatment for this control RNA.
CHAPTER 2 SONG EXPOSURE REGULATES KNOWN AND NOVEL MICRORNAS IN THE ZEBRA FINCH AUDITORY FOREBRAIN

Reprinted, with permission, from BMC Genomics 2011, 12:277; Preethi H Gunaratne†,
Ya-Chi Lin†, Ashley L Benham, Jenny Drnevich, Cristian Coarfa, Jayantha B Tennakoon,
Chad J Creighton, Jong H Kim, Aleksandar Milosavljevic, Michael Watson, Sam
Griffiths-Jones and David F Clayton*

* Correspondence: dclayton@uiuc.edu
† Contributed equally

ABSTRACT

Background: In an important model for neuroscience, songbirds learn to discriminate
songs they hear during tape-recorded playbacks, as demonstrated by song-specific
habituation of both behavioral and neurogenomic responses in the auditory forebrain. We
hypothesized that microRNAs (miRNAs or miRs) may participate in the changing pattern
of gene expression induced by song exposure. To test this, we used massively parallel
Illumina sequencing to analyse small RNAs from auditory forebrain of adult zebra
finches exposed to tape-recorded birdsong or silence.

Results: In the auditory forebrain, we identified 121 known miRNAs conserved in other
vertebrates. We also identified 34 novel miRNAs that do not align to human or chicken
genomes. Five conserved miRNAs showed significant and consistent changes in copy
number after song exposure across three biological replications of the song-silence
comparison, with two increasing (tgu-miR-25, tgu-miR-192) and three decreasing
(tgu-miR-92, tgumiR-124, tgu-miR-129-5p). We also detected a locus on the Z sex
chromosome that produces three different novel miRNAs, with supporting evidence from Northern blot and TaqMan qPCR assays for differential expression in males and females and in response to song playbacks. One of these, tgu-miR-2954-3p, is predicted (by TargetScan) to regulate eight song-responsive mRNAs that all have functions in cellular proliferation and neuronal differentiation.

**Conclusions:** The experience of hearing another bird singing alters the profile of miRNAs in the auditory forebrain of zebra finches. The response involves both known conserved miRNAs and novel miRNAs described so far only in the zebra finch, including a novel sex-linked, song-responsive miRNA. These results indicate that miRNAs are likely to contribute to the unique behavioural biology of learned song communication in songbirds.

**BACKGROUND**

Songbirds are important models for exploring the neural and genomic mechanisms underlying vocal communication, social experience and learning (reviewed in (Clayton, Balakrishnan et al. 2009)). Songbirds communicate using both innate calls and learned vocalizations (songs), and unique specializations of the brain evolved to support this behavior (reviewed in (Jarvis 2004)). In the zebra finch, only the male produces songs, although both sexes process and discriminate specific songs (Miller 1979; Miller 1979; Clayton 1988; Stripling, Kruse et al. 2001). The genome is actively engaged by song communication, as first shown in an early demonstration of how gene responses in the brain discriminate among different song stimuli (Mello, Vicario et al. 1992). The genomic response is not a simple correlate of neural activity and it can vary significantly according to the salience and behavioral context of the experience (Mello, Nottebohm et al.
1995; Jarvis, Scharff et al. 1998; Clayton 2000; Kruse, Stripling et al. 2004; Vignal, Andru et al. 2005; Woolley and Doupe 2008). Recent studies using microarray technology have now shown that song exposure affects the expression of thousands of genes in the auditory forebrain (Dong, Replogle et al. 2009; London, Dong et al. 2009). Repeated exposure to one song leads to an altered gene expression profile, correlated with habituation of both the behavioral and immediate genomic responses to that specific song. These observations suggest the involvement of large and dynamic transcriptional network in the recognition and memory of complex vocal signals (Dong, Replogle et al. 2009).

MicroRNAs (miRNAs or miRs) are emerging as potential control points in transcriptional networks, and may be particularly important for the evolution of brain and behavior. Many miRNAs are expressed in the brain (Cao, Yeo et al. 2006), often in different patterns in different species (Ason, Darnell et al. 2006; Berezikov, Thuemmler et al. 2006; Bak, Silahtaroglu et al. 2008). Brain miRNAs undergo dramatic changes in expression during development (Krichevsky, King et al. 2003; Miska, Alvarez-Saavedra et al. 2004; Sempere, Freemantle et al. 2004) and aging (Li, Bates et al. 2009) and have been functionally implicated in neurological disease (Schratt 2009). They may also function in the normal physiological operation of the nervous system as suggested by evidence for involvement of miR-132 and miR-219 in circadian clock regulation (Cheng, Papp et al. 2007) and miR-134 in control of dendritic translation (Schratt, Tuebing et al. 2006; Fiore, Khudayberdiev et al. 2009).

Here we apply massively parallel Illumina sequencing to probe the involvement of miRNAs in the processing of song experience in the zebra finch auditory forebrain. We begin by identifying 155 different miRNA sequences and the genomic loci of their precursor sequences in the zebra finch genome, including 34 miRNA genes that have not
been detected in the genomes of other species. We then ask whether the miRNA content changes after song exposure and find robust evidence of miRNA responses to song playbacks. We also assess correlations between expression changes of a novel miRNA and its predicted target mRNAs during song habituation. The results indicate an active role for miRNAs in the neural processing of a natural perceptual experience – hearing the sound of another bird singing.

RESULTS

*The miRNAs of the zebra finch auditory forebrain*

We carried out Illumina small RNA sequencing (RNA-seq) on the small RNA (~18-30 nucleotides) fraction of total RNA isolated from adult zebra finch auditory forebrain. Ultimately, we performed 6 Illumina runs on 6 different RNA samples, to assess the effects of song exposure (next section). First we describe the overall small RNA profile obtained by combining the results of all the runs, representing 36 adult zebra finches (equal numbers of males and females). A total of 20 million reads were obtained (Table 2.1) and aligned to reference miRNA sequences from other species (miRBase version 13.0). Overall we identified 107 non-redundant miRNAs representing 52% of sequences that have been previously identified in chicken, rodent and human. The remaining sequences mapping to the piRNA database were denoted as piRNA reads (~30%) (Additional File 1, Table 2.5).

Reads that did not align to known RNAs were assessed for miRNA potential through a novel miRNA discovery pipeline described by Creighton et al. (Creighton, Reid et al. 2009) which tests for properties that are characteristic of known miRNAs. These properties include the following: 1) The mature sequence must map to the stem region of the hairpin sequence of the putative precursor extracted from the zebra finch genome. 2) The mature
miRNA sequence must map to the precursor such that it can be processed following the Drosha processing rules (Reid, Nagaraja et al. 2008). All novel miRNA candidates that map to the loop region and/or lack appropriate Drosha processing sites are failed. 3) Known miRNAs have stable 5'-ends that vary at the most by +/- 1 nucleotide. 4) By contrast the 3'-ends of miRNAs are highly heterogeneous in length due to imprecise Dicer processing (Morin, O'Connor et al. 2008; Reid, Nagaraja et al. 2008) and exhibit non-templated nucleotide sequence changes due to RNA editing (Landgraf, Rusu et al. 2007; Morin, O'Connor et al. 2008; Reid, Nagaraja et al. 2008). 5) Consequently, the putative precursor must give a strong signal of sequence alignments in a tight area of 18-25 nucleotides. Small RNA sequences that are distributed fairly evenly along the entire length of the precursor are rejected since they likely represent degraded products of a large RNA. The candidates that also demonstrate the presence of the miRNA star sequence (miR*) mapping on the opposite side of the mature miRNA and occurring at a lower abundance in the deep sequencing data are considered to be confirmed novel miRNAs in zebra finch.

Using this pipeline (Figure 2.1) we discovered 48 putative novel miRNAs that map on the zebra finch genome to a stem loop structure that folds with a minimum free energy of < -20 kcal/mol (Lee and Ambros 2001). The complete analysis and mapping information for all the novel miRNA candidates is given in Additional File 1, Tables S2 and S3.

All novel miRNA candidates were mapped to genomic loci in the zebra finch genome assembly (Warren, Clayton et al. 2010), and also to human and chicken genomes using the BLAT function of the UCSC Genome Browser (Additional File 1, Table 2.6). In the zebra finch genome, the loci include both annotated exons and introns as well as unannotated intergenic regions. Thirty-four (34) novel microRNAs uncovered from zebra finch are not presently detected in the human or chicken genome assemblies. Eleven (11) map to
genome positions in chicken, and six to positions in the human (with three of these found in human but not chicken assemblies). Tgu-mir-2976 maps to three loci in the finch and 14 in the chicken, indicating a probable expansion of this miRNA in the chicken lineage. This putative novel miRNA is not currently detected in the human assembly HG18.

Tgu-mir-2985 is intriguing as it is located within two stem loops within the introns of two functionally related genes: the glutamate receptor subunits GRIA2 and GRIA4 in all three genomes.

**MiRNA responses to song exposure**

When zebra finches are exposed to playback of a song they have not heard recently, changes occur in the expression of many different mRNAs as detected 30 min after stimulus onset (Dong, Replogle et al. 2009). To determine whether specific miRNAs also change in expression, we counted the Illumina reads in samples of RNA pooled from the auditory forebrain of birds either 30 min after onset of song playback (Song group) or from matched controls (Silence group). In our first such experiment, the birds in both groups were all males (n = 6 each). The read count for each miRNA in each sample was normalized to the total number of usable reads mapped in that sample. We then calculated the ratio of the normalized count in the Song-stimulated condition compared to the Silence condition and performed a Fisher's exact test (with correction for multiple testing) to evaluate whether the ratio differed significantly from the range of expected values at a 95% confidence interval. In the initial experiment with males, 49 of the known conserved miRNAs showed a significant difference, with 28 decreasing and 21 increasing in the group exposed to song (Additional File 1, Table 2.8).

To address the biological reproducibility of the miRNA responses to song more broadly, we then repeated the small RNA-seq comparison two additional times using new
groups of birds. In the second experiment, we used only females, and in the third we used
an equal mix of males and females. In total, therefore, we performed three independent
"song-silence" pairwise comparisons by small RNA-seq, with an overall sex balance but
different sex ratios in each individual comparison. These second and third experiments
were done six months after the first and Illumina technology had improved by this time so
that we obtained twice as many read counts (Table 2.1) - but again we normalized to the
total mapped read number in each individual sample for our statistic analyses. As in the
first experiment, we again observed differential read counts for roughly a third of the
miRNAs, but the identities of the miRNAs affected were somewhat different in each
comparison. This is summarized graphically as a Venn diagram (Additional File 2, Figure
2.4), and comprehensive read count data are presented in Additional File 1, Table 2.8.
Across all three experiments, five conserved miRNAs showed changes that were both
significant and in same direction in all comparisons (Table 2.2). For a number of other
miRNAs, including let-7f, an apparent effect of song exposure was measured in all three
experiments but the direction of change was not consistent (Additional File 1, Table 2.8).

We performed TaqMan assays on RNA from additional birds, probing for eleven of
the "significantly affected" miRNAs, and obtained fluorescent signals in PCR for ten. In
nine out of ten cases, we observed the same direction of song response by TaqMan as in the
small RNA-seq experiment, although the P-value by TaqMan was below 0.05 in only five
cases (tgu-miR-124, tgu-miR-29a, tgu-miR-92, tgu-129-5p, and tgu-miR-2954-3p,
Additional File 1, Table 2.8). The lack of statistical significance in the TaqMan assay for
the others could reflect differences in the sensitivity and resolution of Illumina vs. TaqMan
assays, or the operation of other uncontrolled factors in our experiments that lead to
variability in the expression of some miRNAs.
The transcriptional response in the auditory forebrain of zenk and other mRNAs is specific to song relative to non-song auditory stimuli (Mello, Vicario et al. 1992; Stripling, Kruse et al. 2001; Park and Clayton 2002; Bailey and Wade 2003). To test for song-specificity of the miRNA response, we conducted a further TaqMan experiment assessing the levels of six miRNAs (tgu-miR-124, tgu-miR-92, tgu-miR-129-5p, and three miRNAs derived from the tgu-miR-2954 locus, next section), in birds who had heard either a normal song or a carefully matched non-song acoustic stimulus, "song enveloped noise" (SEN). SEN has the same amplitude envelope as the song from which it is derived but spectral content has been randomized so it does not sound like a song (Park and Clayton 2002). By TaqMan PCR, we confirmed that normal song induced a larger increase in zenk mRNA in these birds than did SEN (Additional File 2, Figure 2.6 panel D). In these same animals, normal song, but not SEN, triggered a significant decrease in the levels of tgu-miR-124, tgu-mir-129-5p, tgu-miR-92 and tgu-miR-2954-3p (Additional File 2, Figure 2.6 panels A-C, H). Thus we conclude that there is indeed a unique miRNA response in the auditory forebrain that is selective for song over non-song acoustic stimuli.

A complex sex-linked miRNA locus in zebra finch and other birds

The novel miRNA, tgu-mir-2954, that was detected most frequently in our Illumina assays maps to the sense strand of an intron in the XPA gene, on the Z chromosome (Figure 2.2A). The precursor hairpin contains reads from both arms, thus meeting our bioinformatic criteria for a confirmed miRNA (Figure 2.2B). By contrast to most known miRNAs, the numbers of reads from both 5' and 3' arms were found at similar copy numbers, suggesting that both arms may make functional mature miRNAs. BLAST analysis of the mir-2954 hairpin precursor sequence against the NCBI nr database identified a putative mature miRNA in chicken (gi|145279910|emb|AM691163.1), and
BLAT analysis of a collection of transcripts from crocodile and 11 other bird species (Kunstner, Wolf et al. 2010) detected mir-2954 transcripts in 2 non-passerine species (two hummingbirds) and 3 passerine species (the American crow, the pied flycatcher, and the golden collared manakin) (Additional File 2, Figure 2.5). There was no BLAT hit in the crocodile, the remaining 3 non-passerine birds (Emu, budgerigar, and ringneck dove), and 3 passerine species (collared flycatcher, blue tit and Eastern phoebe). The lack of a hit does not necessarily mean absence of the gene as these datasets represent incomplete transcriptomes derived by 454 sequencing (Kunstner, Wolf et al. 2010). These results clarify that the sequence is not unique to the zebra finch or passerines, but may nevertheless have a restricted distribution within birds.

To validate the existence of these two miRNAs in zebra finch, we performed TaqMan analyses for both, using their reverse complements as controls. Interestingly, we got significant expression values not only for the predicted miRNAs but also for one of the reverse-complement miRNAs (tgu-miR-2954R-5p) although no significant song regulation for miR-2954R-5p was found (Additional File 2, Figure 2.6 panels I-J). With respect to the XPA gene within which this locus is embedded (Figure 2.2A), these data suggest that precursor-miRNA-stem loops are produced from both the sense (same orientation as XPA) and antisense strands. The stem loop precursor processed by Drosha from the sense RNA (tgu-mir-2954) generates two active miRNAs from its both arms (tgu-miR-2954-3p and tgu-miR-2954-5p). The stem loop precursor processed by Drosha from the antisense RNA (tgu-mir-2954R) generates at least one active miRNA (tgu-miR-2954R-5p) from its 5' end sequence.

We carried out Northern analysis on tgu-miR-2954-3p, which is the miRNA that has the highest number of read counts detected in our Illumina assays among the three
miRNAs from the tgu-mir-2954 locus. A robust signal at ~22 nucleotides is evident in mixed-sex pools of RNA from birds hearing either song or silence, and a weaker signal is also detectable in two female-only pools of RNA (Figure 2.2C). Greater expression in males is consistent with the ZZ genotype of males and the lack of efficient sex chromosome dosage compensation in the zebra finch (Itoh, Melamed et al. 2007; Itoh, Replogle et al. 2010).

By TaqMan as well as by Illumina, we observed an apparent sex difference in the direction of the response of tgu-miR-2954-3p to song - up in males and down in females (Figure 2.3 and Additional File 1, Table 2.8). This suggests this locus may be under complex regulation, integrating information about sex, auditory or social experience and perhaps also other factors related to XPA gene expression.

To gain insight into the potential functional role of tgu-miR-2954-3p in the response to song, we used a conservative strategy to predict gene targets that are both conserved in birds and responsive to song exposure in the zebra finch. Potential targets of miRNAs are described as mRNAs that have sequences that can undergo Watson-Crick base pairing with the 5'-seed (nucleotide 2-7) of the miRNA (Bartel 2009). For target prediction we applied the TargetScan (5.1) algorithm using the chicken genome as an initial reference, and then confirmed presence of the target sequence in the zebra finch. For evidence of song responsiveness, we used the data set of Dong et al. (Dong, Replogle et al. 2009). Eight genes met all these criteria (Table 2.3) and are thus both song-responsive and also subject to regulation by tgu-miR-2954-3p. These genes all have functions in control of cell proliferation or neurite outgrowth (see below).
DISCUSSION

Here we show that a natural perceptual experience, hearing the sound of another bird singing, alters the profile of miRNAs in parts of the songbird brain responsible for auditory perception, integration and memory. The song-regulated population includes both known (conserved) and novel miRNAs. We highlight one sex-linked song-responsive miRNA and identify mRNAs that are potential targets of its action during song exposure. Thus miRNAs may have roles in the information processing functions of the brain, in addition to their roles in brain development and evolution.

To demonstrate this, we first catalogued the miRNAs expressed in the adult zebra finch auditory forebrain. We used massively parallel Illumina sequencing of small RNAs to perform this cataloguing efficiently. In addition to known conserved miRNAs, our analysis identified 48 small RNA sequences that meet the structural criteria for miRNAs but had not been described in miRBase in any organism at the time of our analysis. Fourteen of these are detected in the chicken or human genome assemblies and may give rise to miRNAs that have not yet been described elsewhere due to low copy number, restricted tissue distribution or other factors. The remaining novel miRNAs, 34 in number, may be unique to the zebra finch or the songbird lineage. Few studies have attempted de novo identification of miRNAs from the brain (Berezikov, Thuemmler et al. 2006) and ours is the first to report direct sequencing of songbird brain miRNAs. A previous study did identify precursor sequences for five conserved miRNAs in the developing zebra finch brain (Li, Wang et al. 2007). Also, in parallel with our own Illumina analysis, Li and her colleagues used 454 sequencing to identify miRNAs in the brain and liver of adult zebra finches. These different sets of annotations are compared and collated in a supplement to the analysis of the zebra finch genome assembly (Warren, Clayton et al. 2010).
By comparing birds hearing novel song playbacks or silence, we found evidence for experience-dependent fluctuations in large numbers of miRNAs in the auditory forebrain. We performed three separate pairwise comparisons by Illumina, where all aspects of the experimental conditions were carefully counterbalanced between the two groups in each comparison. The three comparisons were not direct replications of each other, as each had a different sex ratio. Our reasons for varying the sex ratio were partly pragmatic (limited numbers of birds of the same sex that could be removed from our aviary) and partly analytical (males and females have different behavioral responses to songs). Some of the differences between the three sets of results may reflect real biological differences in the responses of males and females. Indeed, our Northern analysis of the tgu-miR-2954-3p confirms a sex difference in expression of this Z-linked miRNA gene. This is especially intriguing because we also obtained TaqMan evidence for both sense and antisense transcripts of this miRNA. One can imagine scenarios where different ratios of sense and antisense transcription occur in males (two copies of the gene) and females (one copy of the gene) with different consequences on the transcriptional networks affected by song exposure in the two sexes.

Ignoring the potential effects of sex, we identified five miRNAs that showed significant and consistent changes in response to song across all three Illumina comparisons. Three miRNAs consistently decreased after song (tgu-miR-92, tgu-miR-124, tgu-miR-129-5p) and two increased (tgu-miR-25, tgu-miR-192). The down-regulated miRNAs are at much higher abundance (> 1000 reads in each run) and perhaps for this reason we were more successful at detecting them and replicating their song regulation by TaqMan assay in subsequent experiments with additional groups of birds. The most abundant miRNA in our regulated set, tgu-miR-124, consistently met the statistical test for
significant down-regulation by song, in each of six separate experiments (three Illumina comparisons, two TaqMan analyses in Additional File 1, Table 2.8, and the TaqMan comparison of song vs. SEN in Additional File 2, Figure 2.6).

In studies in other species, miR-124 has been linked to brain plasticity and development in several contexts. Chronic cocaine administration results in down-regulation of miR-124 in the rodent mesolimbic dopaminergic system (Chandrasekar and Dreyer 2009). In the developing chick neural tube, miR-124a is a component of a regulatory network that controls the transition between neural progenitors and post-mitotic neurons (Visvanathan, Lee et al. 2007). miR-124 also regulates adult neurogenesis, and its overexpression promotes neuronal differentiation (Yu, Chung et al. 2008; Cheng, Pastrana et al. 2009) and neurite outgrowth (Yu, Chung et al. 2008).

Intriguingly, in songbirds neurogenesis continues in the forebrain throughout adulthood, from a population of precursor cells that line the walls of the lateral ventricles and have the characteristics of neural stem cells (Goldman and Nottebohm 1983; Alvarez-Buylla, Theelen et al. 1990; Alvarez-Buylla and Kirn 1997). The net rate of neuronal addition and loss in the adult songbird has been shown to depend on social and environmental influences (Nottebohm, O'Loughlin et al. 1994; Wilbrecht, Crionas et al. 2002; Barnea 2009; Kirn 2010). Perhaps tgu-miR-124 is a regulatory link between experience and neurogenesis - further study of this fascinating possibility is clearly warranted.

Although miRNAs can have diverse functions, they often act by altering the concentrations of specific mRNAs they target via complementary base pairing. We used the TargetScan algorithm (Lewis, Burge et al. 2005) to predict binding sites of tgu-miR-2954-3p in chicken genes, and then we confirmed the presence of the same conserved target sequence in the zebra finch genome assembly. We found eight targets that
met these criteria and were also regulated by song in the Dong et al. microarray data (Dong, Replogle et al. 2009). These eight genes have a provocative coherence in their function, as they are all implicated in control of cell proliferation and neuronal differentiation. Six operate by affecting gene expression and chromatin remodeling as we briefly review here. ELAVL2 is a member of a protein family that binds AU-rich regions in the 3'UTR of genes such as c-fos and promotes the shift from cell proliferation into cellular differentiation (Levine, Gao et al. 1993; Abe, Yamamoto et al. 1996; Ma, Chung et al. 1997; Akamatsu, Okano et al. 1999; Hambardzumyan, Sergent-Tanguy et al. 2009). TLK2 is a kinase tightly associated with DNA replication during cell division (Sillje, Takahashi et al. 1999). At least one of its targets, the histone chaperone Asf1, controls chromatin assembly, thus TLK2 activity can regulate transcription and elongation (Sillje and Nigg 2001; Blackwell and Walker 2003; Carrera, Moshkin et al. 2003). BTG1 is also regulated during the cell cycle (Rouault, Puisieux et al. 1997). It acts as a cofactor for Hoxb9, a transcription factor that controls cell proliferation and differentiation, and BTG1 reduces rates of cell proliferation (Rouault, Puisieux et al. 1997; Corjay, Kearney et al. 1998; Li, Liu et al. 2009). CHD2 can potentially affect transcription of many genes by remodeling chromatin (Hall and Georgel 2007; Marfella and Imbalzano 2007); disruption of CHD2 has profound consequences for development and is implicated in many human diseases (Bandres, Malumbres et al. 2007; Kulkarni, Nagarajan et al. 2008; Nagarajan, Onami et al. 2009). HMGB1 is another DNA binding protein that facilitates transcription by altering chromatin structure to ease promoter binding (Bustin and Reeves 1996; Grasser 1998; Hall, Thomas et al. 2001; Bustin 2002). Some of the genes regulated by HMGB1 may play a role in cell proliferation and migration (Guazzi, Strangio et al. 2003; Bassi, Giussani et al. 2008). Neuronal migration and neurite outgrowth are affected by CRKL, a transcriptional activator that is a
component of the reelin pathway (Ballif, Arnaud et al. 2004; Yip, Kronstadt-O'Brien et al. 2007; Matsuki, Pramatarova et al. 2008; Hubbard, Aken et al. 2009). Unlike the other six genes, NEGR1 and LINGO2 do not seem to alter transcription but they do have established roles in neuronal differentiation. NEGR1 affects cell-cell adhesion to modulate neurite outgrowth and synapse formation (Brennan, Schellinck et al. 1999; Schafer, Brauer et al. 2005; Hashimoto, Yamada et al. 2008). LINGO2 is one member of a family of transmembrane proteins that are involved in neural and axonal regeneration (Ishii, Wanaka et al. 1996; Bormann, Roth et al. 1999). The function of LINGO2 is untested, but expression of a related protein, LINGO1, is attenuated in cortical areas deprived of sensory input and is a partner in a signaling pathway that correlates with neuronal activity during a learning paradigm (Josephson, Trifunovski et al. 2003; Endo, Spenger et al. 2007).

CONCLUSIONS

In conclusion, these data reveal a network of miRNAs in the zebra finch's auditory forebrain, responsive to the experience of hearing another bird sing. The network includes well-characterized conserved miRNA known to have roles in neuronal differentiation (miR-124), and novel miRNAs that can target genes that control neuronal differentiation (tgu-miR-2954-3p). Our data suggest this miRNA network may influence the fundamental shift we have observed in the transcriptional and metabolic state of the auditory forebrain during the process of song-specific habituation (Dong and Clayton 2009; Dong, Replogle et al. 2009). Further study of song responses in the zebra finch may reveal general insights into the neurogenomic mechanisms that underlie learning, memory and the ongoing adaptation to experience.
METHODS

Song stimulation and brain dissections

Zebra finches were obtained from aviaries maintained at the University of Illinois. All procedures involving animals were conducted with the approval of the University of Illinois Institutional Animal Care and Use Committee. The birds were raised in a standard breeding aviary and were tutored under normal social conditions (i.e., by their parents or other adult birds in the breeding colony). All birds used in this study were adults (older than 90 days after hatching). The song playback procedures and brain dissections were performed exactly as in previous microarray analyses, using the same equipment (Replogle, Arnold et al. 2008; Dong, Replogle et al. 2009). Briefly, each bird was put individually into a sound isolation chamber for 18 hours on the first day, and on the second day those in the song group heard 30 minutes of a song not heard previously ("novel song"). Matched controls collected in parallel heard no song playback ("silence"). Birds were sacrificed in song-silence pairs, so that 5 minutes before the end of the song playback to one bird, a bird in the silence group was sacrificed and its auditory forebrain was dissected and frozen in dry ice. Then the auditory forebrain of the song-stimulated bird was dissected and frozen in dry ice. The auditory forebrain dissection (also referred to as auditory lobule) is described in (Cheng and Clayton 2004) and collects NCM (caudomedial nidopallium), CMM (caudomedial mesopallium) and the enclosed Field L subregions. At the end of the song stimulation procedure, all auditory forebrains were transferred and stored at -80C until RNA isolation. For the comparison of responses after overnight isolation to song versus SEN (Additional File 2, Figure 2.6), we used two matched stimuli derived from bird "C7" as previously described (Park and Clayton 2002).
**RNA samples**

For Illumina analyses: Total RNA was extracted using the mirVana miRNA Isolation Kit (Ambion) from three pairs of pooled auditory forebrain samples. 1) Males (samples S7 and S8): 6 birds per pool, collected in November 2008. 2) Females (samples S1 and S2): 6 birds per pool, collected in May 2009. 3) Mixed (samples S3 and S4), 3 males and 3 females each pool, collected in May 2009. Samples with odd numbers were from birds hearing song, and even number hearing silence.

For Northern analysis: Auditory forebrains of 22 birds (12 females and 10 males) were collected in April 2009, and total RNA was extracted by Tri-Reagent (Ambion). Male and female samples were pooled after extraction.

For TaqMan analysis: Analyses were performed on total RNA extracted either by mirVana or Tri-Reagent (Ambion), from the auditory forebrains of individual males or females, collected in April-August 2009, March 2010 or December 2010.

**Illumina small RNA sequencing and novel miRNA discovery**

Fifteen micrograms of total RNA from auditory forebrain of song bird samples described above were gel-fractionated to isolate 18-30 nt small RNAs. 3' and 5' adapters were ligated to the small RNAs and constructs amplified following RT-PCR following the conditions specified in the small RNA kit (FC-102-1009, Illumina) protocol. The small RNA library was sequenced using a Solexa/Illumina GA-1 Genome analyzer. Small RNA sequences were analyzed through a high-throughput computational pipeline described by (Nagaraja, Andreu-Vieyra et al. 2008; Reid, Nagaraja et al. 2008; Creighton, Reid et al. 2009; Ma, Buchold et al. 2009). To identify zebra finch miRNAs that are also conserved in chicken, human and mouse, we performed a local Smith-Waterman alignment of each unique sequence read against each of the mature miRNAs in miRBase version 13.0 for
each of these species. We allowed for a 3 base overhang on the 5’ end and a 6 base overhang on the 3’ end. In the case of redundantly aligning reads, mature miRNA sequences were equally apportioned among each of the hairpins. For each sample, all sequence reads were aligned to a reference set of precursor miRNAs from miRBase version 13.0. The reads that did not align to any known miRNA were passed to our novel miRNA discovery platform as previously described (Creighton, Reid et al. 2009). Briefly, each sequence is first mapped to the reference genome sequence (WUGSC 3.2.4) and 200 bases of flanking sequence are extracted to further define the putative hairpin. This extracted sequence is then folded using the Vienna RNA folding package (Schuster, Fontana et al. 1994) and those sequences that form a plausible hairpin are selected as potential novel miRNA hairpins. These candidates are filtered through a set of three Ambros criteria: 1) the mature putative miRNA sequence must rest on one side of a single hairpin; 2) the putative miRNA sequence must bind relatively tightly within the hairpin stem containing no large or energetically unfavorable loops; and 3) the putative hairpin must have a miRNA-appropriate energy (free energy below -20 kcal/mol). All sequences that passed were then carefully curated to determine if Drosha and Dicer processing could yield the resulting mature sequence from the predicted hairpin. These candidates are then divided into four different categories: "not likely", "potential", "high confidence", and "confirmed" (as in red, gray, blue and green colors in Additional File 1, Tables S2 and S3). Candidates that are flagged red as "not likely" either failed to map in a pile of sequences in a very tight space of 15-25 nt of the predicted hairpin (e.g. were scattered evenly across the full length of the hairpin), mapped within the loop of the hairpin, or mapped to known tRNAs or rRNAs. Candidates that passed all of the above criteria, and also mapped within a hairpin with predicted Drosha and Dicer cut sites were categorized as "high confidence" (blue
annotation in Additional File 1, Tables S2 and S3). All high confidence candidates for which we detected both the mature sequence and the putative star sequence from the same hairpin we categorized as "confirmed" (green annotation in Additional File 1, Table 2.7). In addition to miRNA precursors, the reads were also mapped to the reference zebra finch genome using the Pash software package (Kalafus, Jackson et al. 2004; Coarfa and Milosavljevic 2008), and uploaded to the Genboree platform (http://www.genboree.com website) to identify potential mappings to piRNAs, snoRNAs and other annotations in addition to miRNAs (data shown in Additional File 1, Table 2.5). PiRNAs (i.e., Piwi-interacting RNAs) have a central role in the maintenance of the integrity of genomes through the silencing of transposable elements (Thomson and Lin 2009). SnoRNAs (small nucleolar RNAs) function in site-specific ribosomal RNA modification, rRNA processing and more recently have been found to guide alternate splicing and RNA editing of mRNA transcripts (Royo and Cavaille 2008).

**TaqMan qPCR**

To measure the mature miRNA, the TaqMan MicroRNA Assay Kit (Applied Biosystems) was used according to the manufacturer's instructions. Probe sequences used for each target miRNA are given in Table 2.4.

**Northern blot analysis**

Northern blotting to confirm novel miRNA tgu-miR-2954-3p was performed by modifying the protocol of (Gu, Reid et al. 2008). 2 μg of total RNA was heated at 65°C for 5 min with 2X loading dye (Ambion), quenched on ice, and loaded on a 15% TBE Urea gel (Invitrogen). Total RNA was separated by electrophoresis at 200V for 50 min. The gel was stained with with EtBr in 1x TBE (4 μL of 10 mg/ml EtBr per 100 ml of 1x TBE) for 3 minutes with gentle shaking and transferred to nylon membrane for 90 min at 200V using...
1X TBE buffer at room temperature. The membrane was cross-linked at 1200 kJ for 45 seconds. RNA probes were synthesized for tgu-miR-2954-3p probe 5' - UGCUAGGAGUGGAAUGGGGAU G - 3' by Integrated DNA Technologies. Radio labeling was carried out in a reaction of 12.0ul dH2O + 2.0ul PNK buffer + 1.0ul (100ng/ul) probe + 1.0ul PNK polymerase (Promega) + 4.0ul P32-gamma-ATP (10mCi/ml) (PerkinElmer). The reaction was incubated at 37°C for 1 hour and inactivated at 65°C for 10 min. The probe was purified using Nick columns from GE following manufacturer's instructions. The membranes were pre-hybridized for 30 min with 20 ml of pre-hybridization buffer (5X SSC + 20 mM NaPO4 + 7X SDS + 2X Denhardt (pre warmed) at 60°C) in a rotating hybridization oven. Hybridization was carried out at 50°C in a rotating incubator for 24h. The membranes were washed for 10 min at 50°C with 20-30mL of wash buffer (2X SSC + 0.5% SDS). When background was ~0.5 cpm, the membranes were wrapped in saran wrap and exposed at -80°C for ~72h.

Additional Files

Additional file 1. Supplemental tables.xls. This one file contains all four Supplemental Tables, each as a separate worksheet. Table 2.5 ("1 overview") is a summary of Illumina sequence read alignments for six pools of RNA from zebra finch auditory forebrain responding to song versus silence, and shows the distribution of sequence reads in relation to multiple genomes and multiple annotations in the current genomic databases. Table 2.6 ("2 novel hairpins") gives detailed alignments of putative pre-miRNAs and read sequences. Table 2.7 ("3 novel genes") shows annotations of all novel miRNA loci mapped in genome assemblies of zebra finch, chicken or human. Table 2.8 ("4 all read counts") gives read counts and current annotation in miRBase of all conserved and novel miRNAs,
with statistics. File can be viewed from
www.biomedcentral.com/content/supplementary/1471-2164-12-277-s1.xls

Additional file 2. Supplemental figures.doc. This one file contains all three supplemental figures. Figure 2.4 is a Venn diagram of numbers of miRNAs with significant differential expression in response to novel song in three Illumina experiments. Figure 2.5 shows a comparative mapping in other avian transcriptomes of tgu-mir-2954. Figure 2.6 demonstrates the song-specificity of the miRNA response, using TaqMan to compare the levels of specific miRNAs in animals from groups that heard song, matching song-enveloped noise, or silence. File also can be viewed from
www.biomedcentral.com/content/supplementary/1471-2164-12-277-s2.doc

AUTHORS' CONTRIBUTIONS

PHG coordinated the work of Illumina RNA-seq and prepared the manuscript. YL conducted the song exposure experiments and subsequent dissections and RNA extractions, performed TaqMan qPCR, analyzed differentially expressed miRNAs and participated in drafting the manuscript. ALB performed Illumina RNA-seq and Northern blot. JD helped analyze expression data of Illumina RNA-seq and TaqMan qPCR. CC, JBT, CJC, JHK, and AM participated in mapping and analyzing Illumina RNA-seq data. MW and SGJ helped with miRNA sequence annotation. DFC designed and coordinated the study and drafted the manuscript. All authors read and approved the manuscript.

ACKNOWLEDGEMENTS

We thank Sarah London for useful discussions and contributions to the text. Supported by NIH RO1 NS045264 and RO1 NS051820 (to D.F.C).
NOTE ADDED IN PROOF

The novel miRNA referred to here as "miR-2954-3p" is now identified in miRBase as "miR-2954". The novel miRNA referred to here as "miR-2954-5p" is now identified in miRBase as "miR-2954*".
**TABLES AND FIGURES**

Table 2.1. Summary statistics for the read alignments.
Six different pools of auditory forebrain were analyzed independently by Illumina small RNA sequencing, as described in the text.

<table>
<thead>
<tr>
<th></th>
<th>Male silence</th>
<th>Male song</th>
<th>Female silence</th>
<th>Female song</th>
<th>Mix silence</th>
<th>Mix song</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Total Reads</strong></td>
<td>2,704,778</td>
<td>2,056,391</td>
<td>3,173,108</td>
<td>3,546,038</td>
<td>3,962,050</td>
<td>4,738,528</td>
</tr>
<tr>
<td><strong>Total Usable Reads</strong></td>
<td>1,179,330</td>
<td>1,155,168</td>
<td>2,244,376</td>
<td>2,498,648</td>
<td>2,249,188</td>
<td>2,950,398</td>
</tr>
<tr>
<td><strong>Reads aligning with known miRNA</strong></td>
<td>401,934</td>
<td>209,944</td>
<td>1,638,528</td>
<td>1,755,748</td>
<td>1,348,109</td>
<td>2,113,006</td>
</tr>
<tr>
<td><strong>Fraction</strong></td>
<td>34%</td>
<td>18%</td>
<td>73%</td>
<td>70%</td>
<td>60%</td>
<td>72%</td>
</tr>
</tbody>
</table>
Table 2.2. Conserved miRNAs with consistent responses to song exposure.

Shown are the Illumina read data for the five miRNAs that show a consistent response to song (same direction of change, significant in all three comparisons). “Song” and “Silence” list raw counts from the Illumina read analysis (Additional File 1, Table 2.8). “Fold Change” is the ratio of Song versus Silence read counts, after the raw counts were normalized within each run to the sum of mapped reads for that sample. Thus a value of >1 indicates a relative increase in the group exposed to song, and <1 indicates a decrease. “FDR-P” indicates the result of the Fisher’s exact test (FDR adjusted) for this comparison. See Additional File 1, Table 2.8 for full list of values for all miRNAs, and associated TaqMan values for a subset of these miRNAs (measured in a different set of males and females).

<table>
<thead>
<tr>
<th></th>
<th>Male</th>
<th></th>
<th>Female</th>
<th></th>
<th>Mix</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Silence</td>
<td>Song</td>
<td>Fold Change</td>
<td>FDR-P</td>
<td>Silence</td>
<td>Song</td>
</tr>
<tr>
<td>Increasing</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>tgu-miR-25</td>
<td>227</td>
<td>423</td>
<td>3.57</td>
<td>1.65-27</td>
<td>55</td>
<td>212</td>
</tr>
<tr>
<td>tgu-miR-152</td>
<td>26</td>
<td>69</td>
<td>5.08</td>
<td>1.27-06</td>
<td>36</td>
<td>90</td>
</tr>
<tr>
<td>Decreasing</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>tgu-miR-22</td>
<td>539</td>
<td>100</td>
<td>0.53</td>
<td>1.1E-04</td>
<td>5479</td>
<td>5398</td>
</tr>
<tr>
<td>tgu-miR-124</td>
<td>24624</td>
<td>7056</td>
<td>0.55</td>
<td>2.1E-251</td>
<td>56802</td>
<td>46434</td>
</tr>
<tr>
<td>tgu-miR-129-5p</td>
<td>2020</td>
<td>602</td>
<td>0.57</td>
<td>4.05-19</td>
<td>9778</td>
<td>7272</td>
</tr>
</tbody>
</table>
**Table 2.3. Song-regulated targets of tgu-miR-2954-3p.**

We used TargetScan to find binding sites of tgu-miR-2954-3p on eight chicken genes and here are listed the information of their homologous genes in the zebra finch genome including Ensembl IDs, Gene Symbols, EST (Accession numbers of song-regulated EST identified in the previous microarray study) and Gene Names (or aliases in parenthesis).

<table>
<thead>
<tr>
<th>Ensembl ID</th>
<th>Gene Symbol</th>
<th>EST</th>
<th>Gene Name</th>
</tr>
</thead>
<tbody>
<tr>
<td>ENSTGUG00000001349</td>
<td>ELAVL2</td>
<td>CK313262 B(HuB)(ELAV-like neuronal protein 1)(Nervous system-specific RNA-binding protein Hel-N1)</td>
<td>ELAV-like protein 2 (Hu-antigen) (HuB)</td>
</tr>
<tr>
<td>ENSTGUG00000001404</td>
<td>LINGO2</td>
<td>DV957508</td>
<td>Leucine-rich repeat and immunoglobulin-like domain-containing nogo receptor-interacting protein 2 Precursor (Leucine-rich repeat neuronal protein 6C)(Leucine-rich repeat neuronal protein 3)</td>
</tr>
<tr>
<td>ENSTGUG00000003073</td>
<td>TLK2</td>
<td>CK305975 (EC 2.7.11.1)(Tousled-like kinase 2)(PKU-alpha)</td>
<td>Serine/threonine-protein kinase tousled-like 2</td>
</tr>
<tr>
<td>ENSTGUG00000008207</td>
<td>BTG1</td>
<td>CK303273</td>
<td>Protein BTG1 (B-cell translocation gene 1 protein)</td>
</tr>
<tr>
<td>ENSTGUG00000008540</td>
<td>CHD2</td>
<td>DV958991 2 (EC 3.6.1.)(ATP-dependent helicase CHD2)</td>
<td>Chromdomain-helicase-DNA-binding protein</td>
</tr>
<tr>
<td>ENSTGUG00000010181</td>
<td>XP_002196848.1</td>
<td>CK304764</td>
<td>crk-like protein (v-crk avian sarcoma virus CT10 oncogene homolog-like) (CRKL)</td>
</tr>
<tr>
<td>ENSTGUG00000010364</td>
<td>NEGR1</td>
<td>DV954047</td>
<td>Neuronal growth regulator 1 Precursor</td>
</tr>
<tr>
<td>ENSTGUG00000011700</td>
<td>HMGB1</td>
<td>CK314519</td>
<td>High mobility group protein B1 (High mobility group protein 1)(HMG-1)</td>
</tr>
</tbody>
</table>
Table 2.4. Probes used for Taqman analysis of specific miRNA sequences.

<table>
<thead>
<tr>
<th>miRBase name</th>
<th>Company name</th>
<th>Sequence detected</th>
</tr>
</thead>
<tbody>
<tr>
<td>tgu-let-7a</td>
<td>let-7a</td>
<td>5’-UGAGGUAGGUAGGUAGGUAGUAGU-3’</td>
</tr>
<tr>
<td>tgu-let-7f</td>
<td>let-7f</td>
<td>5’-UGAGGUAGGUAGGUAGGUAGUAGU-3’</td>
</tr>
<tr>
<td>tgu-miR-124</td>
<td>miR-124</td>
<td>5’-UAAGGCACGCGUGGAAUGCC-3’</td>
</tr>
<tr>
<td>tgu-miR-9</td>
<td>miR-9</td>
<td>5’-UCUUUGGUAAUCUAGCUUGUAUGA-3’</td>
</tr>
<tr>
<td>tgu-miR-129-5p</td>
<td>miR-129-5p</td>
<td>5’-CUUUUUGCGGUCUGCGCUUGC-3’</td>
</tr>
<tr>
<td>tgu-miR-129-3p</td>
<td>miR-129-3p</td>
<td>5’-AAGGCCUUAACCCAAAAAGCAU-3’</td>
</tr>
<tr>
<td>tgu-miR-29a</td>
<td>miR-29c</td>
<td>5’-UAGCACCAUUUGAAUCCGGU-3’</td>
</tr>
<tr>
<td>tgu-miR-92</td>
<td>miR-92a</td>
<td>5’-UAUUGCAGCAGGUGRGCCUGU-3’</td>
</tr>
<tr>
<td>tgu-miR-25</td>
<td>miR-25</td>
<td>5’-CAUUGACAGGUGUCUCGUGU-3’</td>
</tr>
<tr>
<td>RNU6B</td>
<td>RNU6B</td>
<td>5’-CGCAAGGGAUCAAGCGCAACUUCGUGAGCGUUCAUUUU-3’</td>
</tr>
<tr>
<td>tgu-miR-2954-5p</td>
<td>novel51F-5p</td>
<td>5’-GCUGAGAGGGCUUGGGGAGAGGA-3’</td>
</tr>
<tr>
<td>tgu-miR-2954-3p</td>
<td>novel51F-3p</td>
<td>5’-CAAUCUCCAUCCACUCCUGAGCA-3’</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(Northern validated)</td>
</tr>
<tr>
<td>tgu-miR-2954R-5p</td>
<td>novel51R-5p</td>
<td>5’-UCGUAGGAGUAGAAUGGGGAUG-3’</td>
</tr>
<tr>
<td>tgu-miR-2954R-3p</td>
<td>novel51R-3p</td>
<td>5’-UCCUCUCCCAAGCCUCCUCUCAGC-3’</td>
</tr>
</tbody>
</table>
Figure 2.1. Pipeline with yields for analysis of putative novel miRNAs.
Figure 2.1. (cont.) Pipeline with yields for analysis of putative novel miRNAs.
52 small RNA sequences did not align to miRBase reference sequences and were assessed for miRNA potential. 48 sequences passed the minimum criteria and were categorized into three groups according to strength of evidence (sequences are color-coded in Additional File 1, Table 2.7, as indicated). Seven (7) are confirmed novel miRNAs since they had all the characteristics of known miRNAs and in addition also had a less abundant miR* sequence that maps on the opposite side of the stem from the putative novel miRNA. These are labelled green in Additional File 1, Table 2.7. Twenty-one (21) putative novel miRNAs are highly confident (labelled blue) since they also shared characteristics of known miRNAs but no sequence was found aligning to the miR* region. Given that the miR and miR* sequences for most known miRNAs have vastly different copy numbers such that the miR* sequence is sometimes not found, the highly confident candidates are also highly likely to be genuine novel miRNAs. Twenty (20) candidates (labelled grey) had a subset of the characteristics of known miRNAs but not all and therefore were deemed potential candidates that require more evidence.
Figure 2.2. The genome locus for tgu-mir-2954 produces three different miRNAs. (A) Alignments via the UCSC Genome Browser of the three detected miRNAs to the intron of the zebra finch XPA gene. (B) Hairpin precursors for the three miRNAs. (C) Northern blot analysis using an RNA probe complementary to novel miRNA tgu-miR-2954-3p.
Figure 2.3 Analysis of miRNAs produced at the tgu-mir-2954 locus.
Figure 2.3. (cont.) Analysis of miRNAs produced at the tgu-mir-2954 locus.
TaqMan and Illumina RNA-seq data generated from independent sets of birds (n = 6 in each data set) for expression from the tgu-mir-2954 locus. (A) TaqMan results, where the relative gene expression of each individual bird (open circle) was obtained by using the $2^{\Delta\Delta C_{T}}$ method [98]; the relative gene expression of either Silence (white bar) or Song (gray bar) group was the mean of six individuals; the P value was calculated by paired t test since each song stimulated animal was explicitly paired with a silence control animal collected simultaneously. (B) Read counts from the Illumina RNA-seq for miR-2954-3p and miR-2954-5p (also shown in the Additional File 1, Table 2.8).
Figure 2.4. Venn diagram of numbers of miRNAs with significant differential expression in response to novel song in three Illumina experiments. (A) Increased in Song compared to Silence. (B) Decreased in Song compared to Silence.
Figure 2.5. Comparative mapping in other avian transcriptomes of tgu-mir-2954.
Figure 2.5. (cont.) Comparative mapping in other avian transcriptomes of tgu-mir-2954.
Figure 2.5. (cont.) Comparative mapping in other avian transcriptomes of tgu-mir-2954.
cDNA from other species (34) was sequenced (Roche 454) and alignments of tgu-2954 are shown at two levels of magnification: (A) the nucleotide level for the mature miRNA sequence. (B) the level of the larger XPA gene in which the miRNA is embedded.
Figure 2.6. Song-specificity of the miRNA response.
ANOVA $p = 0.04574 ^ *$
Post hoc analyses after ANOVA
C7noise vs. Silence $p = 0.8724359$
C7song vs. Silence $p = 0.0494895$
C7song vs. C7noise $p = 0.1228024$

ANOVA $p = 2.696 \times 10^{-08} ^ {***}$
Post hoc analyses after ANOVA
C7noise vs. Silence $p = 0.0020331$
C7song vs. Silence $p = 0.0000063$
C7song vs. C7noise $p = 0.0000000$

ANOVA $p = 0.01287 ^ *$
Post hoc analyses after ANOVA
C7noise vs. Silence $p = 0.9140920$
C7song vs. Silence $p = 0.0372510$
C7song vs. C7noise $p = 0.0168752$

ANOVA $p = 1.251 \times 10^{-08} ^ {***}$
Post hoc analyses after ANOVA
C7noise vs. Silence $p = 0.1521374$
C7song vs. Silence $p = 0.0000002$
C7song vs. C7noise $p = 0.0000000$

Figure 2.6. (cont.) Song-specificity of the miRNA response.
Figure 2.6. (cont.) Song-specificity of the miRNA response.

TaqMan was used to compare the levels of specific miRNAs in animals from three treatment groups. One group heard silence, another heard the normal song of bird C7 (C7song), and the third group heard a matched non-song stimulus derived by randomizing the spectral content of C7 (C7noise) (Park and Clayton 2002). Relative gene expression in each individual bird was obtained by using the 2^−ddCt method and presented as a dot in the plot. The relative gene expression mean of each group of birds is shown by the bar. Results of statistical analyses (ANOVA followed by Tukey’s HSD posthoc tests) for each miRNA are listed below each plot.
CHAPTER 3 A GENE NETWORK REGULATED BY MIR-2954, A SONG-RESPONSIVE Z-LINKED MICRORNA EXPRESSED IN A SUBSET OF ZEBRA FINCH BRAIN CELLS

Manuscript in preparation by
Ya-Chi Lin¹,², Christopher N. Balakrishnan³ and David F. Clayton¹,²,⁴

¹Institute for Genomic Biology, University of Illinois, Urbana-Champaign, IL 61801
²Department of Cell and Developmental Biology, University of Illinois, Urbana-Champaign, IL 61801
³Department of Biology, East Carolina University, Greenville, NC 27858
⁴Beckman Institute, University of Illinois Urbana-Champaign, IL 61801

ABSTRACT

Natural experience can cause large and rapid changes in gene expression in brain centers for cognition and perception. MicroRNAs (miRNAs or miRs) could have a role in organizing experience-responsive gene expression networks. A previous study identified miRs that increase or decrease in the zebra finch auditory forebrain in response to song playbacks. Among these, miR-2954 is of special interest as its gene maps to the Z sex chromosome and is expressed differently in males and females. To assess the functional significance of miR-2954 regulation, we first characterized its broader expression pattern across various tissues of the zebra finch and in two recently developed zebra finch cell lines. With fluorescence in situ hybridization and immunohistochemistry, we show that miR-2954 is present in subsets of cells in the major brain regions involved in song production and perception. Using Illumina RNA sequencing (RNA-seq), we show that
pharmacological inhibition of miR-2954 in the G266 (male) zebra finch cell line leads to
significant changes in expression of thousands of mRNAs two days later, increasing
MAPK signaling and decreasing ribosomal and mitochondrial gene expression.
Functionally similar transcriptional changes occur in the auditory forebrain after song
exposures that also suppress miR-2954. Thus a microRNA may have a primary role in
coordinating the system of genes involved in the neural response to a perceptual
experience.

INTRODUCTION

Zebra finches are songbirds that communicate using learned vocalizations
(Immelmann 1969; Miller 1979; Miller 1979; Clayton 1988), and have become important
model organisms for studying the neural and genomic mechanisms of social learning,
memory and behavior (Robinson, Fernald et al. 2008; Clayton, Balakrishnan et al. 2009).
Previous genomic studies have shown that playback of recorded vocalizations (songs)
triggers a complex genomic response in the zebra finch brain, with changes in expression
of thousands of RNAs including both mRNAs and noncoding RNAs (ncRNAs) (Replogle,
of the expressed gene network shifts as a song becomes familiar (“song habituation”),
favoring the suppression of genes associated with energetics and macromolecular synthesis
(Dong, Replogle et al., 2009). What mechanisms coordinate this major change in gene
expression?

Of particular functional interest are the song-responsive microRNAs (Gunaratne, Lin
et al. 2011). MicroRNAs (miRNAs or miRs) comprise a family of ncRNAs that are
distinguished both by their distinctive small size and their potential to regulate the
expression of other genes via direct interactions with mRNAs. The population of miRs expressed in brain is complex and highly diverse (Cao, Yeo et al. 2006) with great variation in expression patterns among different species (Ason, Darnell et al. 2006; Berezikov, Thuemmler et al. 2006; Bak, Silahtaroglu et al. 2008), suggesting a potential role in brain evolution and divergence. Specific brain expressed miRs have now been linked to numerous neural processes and brain diseases (Kuss and Chen 2008; Coolen and Bally-Cuif 2009; Saba and Schratt 2010). In the unique neurobiology of songbirds, miRs have been proposed as potential regulators of brain circuit development, brain sexual dimorphism, and the perceptual learning of social information mediated by song (Li, Wang et al. 2007; Warren, Clayton et al. 2010; Gunaratne, Lin et al. 2011). Using Illumina-based RNA sequencing (RNA-seq), we defined the population of miRs expressed in the zebra finch auditory forebrain and showed that six of these undergo rapid changes in abundance immediately following song playbacks (Gunaratne, Lin et al. 2011).

Here we have focused on one of these song-responsive miRs, miR-2954, which is distinctive for several reasons. The primary mature miR-2954 product is predicted to target the 3’UTR of NR4A3, one of the most robust song-responsive mRNAs (Warren, Clayton et al. 2010). The gene for miR-2954 is on the avian Z chromosome and produces at least three different products from both strands, with significantly higher expression in males (ZZ) compared to females (ZW) (Gunaratne, Lin et al. 2011). Moreover, evidence suggests that miR-2954 may respond to song differently in the two sexes, clearly decreasing in females but increasing somewhat in males. So far, miR-2954 has been found only in avian species, including chickens, hummingbirds, crows, flycatchers, manakins, and zebra finches (Zhao, McBride et al. 2010; Gunaratne, Lin et al. 2011). These observations suggest that
miR-2954 may have evolved to regulate some of the distinctive sex differences notable in avian vocal communication.

To evaluate the plausibility of this hypothesis, we first set out to confirm the sex difference in expression and determine the tissue and subcellular distribution of the mature miR-2954 product. We then tested whether targeted manipulation of miR-2954 using sequence specific inhibitors (Meister, Landthaler et al. 2004; Schratt, Tuebing et al. 2006) would alter the expression of any mRNAs; for this purpose we took advantage of the recent developments of male and female zebra finch cell lines (Itoh and Arnold 2011) coupled with RNA-seq methodology (Balakrishnan, Lin et al. 2012). We find that suppressing miR-2954 causes broad changes in gene expression that are functionally similar to those observed during song habituation (Dong, Replogle et al. 2009), suggesting a primary role for this miRNA in organizing a gene expression network that adapts to perceptual experience.

RESULTS

Variation in miR-2954 expression by sex and tissue

Using RT-qPCR, we measured miR-2954 in two non-neural zebra finch cell lines and nine zebra finch tissues, and found it to vary by only ± 2-fold among the tissues (Figure 3.1). A sex-biased expression pattern is evident in all tissues, with lower expression in females by 10-fold or more, and roughly 100-fold lower expression in the female ZFTMA cell line (derived from a tumor found on the thigh) compared to the male G266 line (derived from a tumor found on the forehead). Ubiquitous expression pattern in different tissues and higher expression levels in males is consistent with observations made in chicken (gga-miR-2954) by Northern blot (Zhao, McBride et al. 2010).
Cellular and subcellular localization in male song nuclei

To map miR-2954 expression within zebra finch brain at cellular and subcellular levels, we used fluorescence in situ hybridization (FISH) with a specific locked nucleic acid (LNA) probe. The sections were then double-stained with an antibody to the neuronal marker, NeuN, to distinguish neurons from non-neurons. A scrambled LNA sequence was used as negative control and generated no fluorescence signals. We detected expression broadly throughout the brain, and focused our analysis on the major regions of the telencephalon involved in song production and perception: HVC (letters used as proper name), the striatal nucleus Area X, the lateral magnocellular nucleus of the anterior nidopallium (LMAN), the robust nucleus of the arcopallium (RA), and the caudomedial nidopallium (NCM). HVC, Area X, LMAN and RA are responsible for song production in males and are absent or much reduced in females whereas NCM (in the auditory forebrain) is morphologically similar in both sexes.

In males, HVC is readily distinguished from the surrounding nidopallium by its concentration of large neurons, evident at low magnification by Nissl staining (not shown) or by immunocytochemistry for NeuN (Figure 3.2A). The density of miR-2954-positive cells is similar in both the HVC interior and the surrounding nidopallium (Figure 3.2A). At higher magnification, the miR-2954 signal is typically concentrated in small domains that appear to lie within or immediately adjacent to DAPI-positive nuclei (Figure 3.2B, compare blue DAPI and red miR-2954) inside larger cell bodies that are co-labeled for NeuN (Figure 3.2B, compare green NeuN, and also merged image). Only about half of the NeuN-positive figures are double-labeled for miR-2954, however, and conversely, a few miR-2954-positive figures can be seen that are clearly not associated with NeuN staining. This indicates that miR-2954 is present in subsets of both neurons and glial cells. HVC has
two classes of projection neurons and the larger neurons project to Area X (Nixdorf, Davis et al. 1989; Fortune and Margoliash 1995; Dutar, Vu et al. 1998); based on the large soma size of miR-2954-positive/NeuN-positive figures, we suggest neuronal expression may be limited to the X-projecting subset although additional evidence (e.g., retrograde tracer analysis) would be needed to confirm this.

Area X itself is characterized by a dense population of relatively small neurons, and miR-2954 is present generally in this population (Figures 3.2A and 3.2B). In contrast, LMAN is characterized by a core region of large neuronal cell bodies, and this core region contains little or no miR-2954 (Figure 3.2A). Figure 3.2B confirms absence of miR-2954 in the large central core neurons of LMAN. Some labeling is evident in the surrounding shell region, possibly including some non-neuronal cells (one apparent example of non-neuronal labeling in the LMAN shell can be seen in Figure 3.2B). Nucleus RA is also characterized by a central population of large neurons, and these also show relatively little labeling for miR-2954 (Figures 3.2A and 3.2B).

In the auditory region NCM, miR-2954 labeling is apparent throughout, though somewhat less robust in the ventral region immediately beneath the overlying ventricle (Figure 3.2A). Cells in NCM are often arranged in grape-like clusters in which neurons and glia are hard to resolve; some of these clusters show little or no labeling, and in other cases only one or two cells of a cluster appear to express mir-2954 (Figure 3.2B). We also considered whether miR-2954 localization in NCM might change as a result of song playbacks or vary with tissue sex. We did not observe any evidence for this by FISH, however (Figures 3.4 and 3.5).

In sum, miR-2954 labeling is most apparent within cell nuclei where it is often further concentrated in puncta or subregions suggestive of nucleoli. It is found in distinct neuronal...
subpopulations in the song system (e.g., present in HVC and Area X, absent from the large neurons of LMAN core and RA), and occasionally detected in cells that appear to be non-neuronal by absence of NeuN immunostaining.

**Consequences of miR-2954 knock-down on target gene expression**

Many miRs function by binding to complementary sequences typically found in the 3’-UTRs of target mRNAs, triggering degradation of these mRNAs and thus suppressing the expression of the target genes at a post-transcriptional level. We previously noted the presence of predicted binding sites for miR-2954 in the 3’-UTR of the song-regulated NR4A3 mRNA (Warren, Clayton et al. 2010), and we identified eight other song-regulated mRNAs that also carry predicted miR-2954 binding sites in their 3’-UTRs (Gunaratne, Lin et al. 2011). To test whether changes in miR-2954 expression affect the levels of these or other mRNAs, we explored the use of a synthetic sequence-specific miR inhibitor for transfection into a new cell line (G266) recently developed from a male zebra finch (Itoh and Arnold 2011; Balakrishnan, Lin et al. 2012). We used a miR inhibitor prepared by Dharmaco RNAi Technology, which is a single-stranded RNA molecule bearing a proprietary chemical modification and secondary structures in the flanking region to suppress the function of miR-miRISC complex through reverse complementary to miR sequence (Vermeulen, Robertson et al. 2007). Endogenous miR-2954 is abundant in male G266 cells (Figure 3.1), and so we transfected G266 cells with miR-2952 inhibitor and confirmed that this results in a significant reduction in endogenous miR-2954 measured 48h later (Figure 3.3A). We also confirmed that this reduction leads to an increase in a major predicted target mRNA (Warren, Clayton et al. 2010; Gunaratne, Lin et al. 2011), NR4A3 (Nuclear receptor subfamily 4 group A member 3) (Figure 3.3B). Transfection with a control pseudo-inhibitor had no effect on either miR-2954 (Figure 3.3A) or NR4A3
mRNA (Figure 3.3B). We also tested for effects of mir-2954 inhibition on the eight other targets predicted in Gunaratne, Lin et al (2011), but did not detect significant expression changes for any of these.

To test more generally for genome-wide effects of manipulating miR-2954 in zebra finch cells, we applied next generation sequencing (RNA-seq) to compare the population of mRNAs in the G266 cell line transfected with either the miR-2954 inhibitor or the control pseudo-inhibitor, which is predicted not to interact with any known zebra finch transcript. We performed 3 replications of the experimental comparison, yielding a total of six RNA libraries. RNAs from 14188 of 17475 Ensembl-annotated genes were detected with at least one read in at least one of the libraries, while 3287 genes gave no evidence of expression (Supplemental Table 3.3). Using DESeq analysis (Anders and Huber 2010), we detected 2923 genes that were differentially expressed between the two treatments at an adjusted p value less than 0.05. Filtering the differentially expressed genes by magnitude of effect, we arrived at lists of 974 up-regulated (>1.5-fold after inhibitor) and 979 down-regulated (<0.75-fold after inhibitor) genes (Supplemental Table 3.4).

We then performed GO and KEGG analyses by comparing these lists of significantly regulated genes against the entire population of 14188 whose expression was detected in at least one library. The functional representations of the regulated genes are listed in Tables 3.1 and 3.2. Among the up-regulated genes (i.e., those that increase following miR-2954 inhibition and therefore must be directly or indirectly suppressed by mir-2954 expression), genes associated with serine/threonine and tyrosine protein kinase signaling function are strongly over-represented (with adjusted p-values of <.0001, Table 3.1A). KEGG pathway analysis (Table 3.2A) specifically implicates the MAPK pathway in this gene set (Figure 3.6). Conversely, the G-protein coupled receptor protein signaling pathway (GO:0007186)
is significantly underrepresented in the gene set that increases after miR-2954 inhibition (Table 3.1A).

A very different functional profile is seen in the set of genes that decrease when miR-2954 is inhibited. Terms for ribosome, translation and mitochondria are all profoundly over-represented in this gene set, whereas terms for membrane, signal transduction and transcription factor activities are all under-represented (Table 3.1B). KEGG analysis is consistent with this and also indicates an underrepresentation of MAPK, Wnt signaling and ubiquitin-mediated proteolysis pathways in the genes that decrease upon mir-2954 signaling.

Together, this suggests that normal expression of miR-2954 in the G266 cells supports energetic growth functions (ribosomes and mitochondria) and suppresses protein kinase but not G-protein-coupled receptor signaling pathways. A reduction of miR-2954 reverses these effects, shifting the gene expression network away from energy metabolism and increasing functions associated with MAPK signaling in particular.

**DISCUSSION**

Natural experience elicits complex changes in gene expression in parts of the brain associated with perception and cognition. Though early studies focused on just a few genes that are broadly responsive to cellular signals (the IEGs), high throughput techniques have now revealed that thousands of genes may vary in their expression depending on the experience, environment, brain system and species (Mello, Vicario et al. 1992; Clayton 2000; Dong, Replogle et al. 2009; London, Dong et al. 2009; Mukai, Replogle et al. 2009). By what logic are these different gene networks organized, and how do they evolve? Regulation through microRNAs represents one mechanism that may coordinate the
expression of many genes at once and allow the rapid evolution of new patterns of coordinate gene regulation. We have begun to explore the relevance of this to brain function and psychological experience, using the zebra finch song response as a case study. Here we have characterized a novel song-responsive miRNA, miR-2954, and shown how it can influence the expression (mRNA levels) of thousands of genes. There are strong, non-random functional signatures in the population of genes influenced by miR-2954. These signatures are similar to ones observed in the brain after song habituation. As the miR is expressed differently in males and females and found in subsets of neurons and other cell types, we suggest that the evolution and regulation of miR-2954 itself may have shaped the way the songbird brain responds to song experience.

The mir-2954 gene lies within an intron of the XPA gene. We do not yet know whether there is functional or regulatory significance to this relationship. Many miRs are processed from introns of other genes, potentially affording multiple levels of control over production of the miR. Here we focused only on the primary mir-2954 sequence, but note that other sense and antisense sequences also arise from the mir-2954 gene locus (Figure 2 in [Gunaratne, Lin et al. 2011]) and could have functional significance though we have not yet explored that possibility.

Though discovered in the context of brain genes, we find that miR-2954 is expressed broadly in many tissues. Yet in the brain its expression is not ubiquitous, as it shows discrete localization to subsets of brain cells. We paid special attention to its expression in the major well-defined, sexually dimorphic nuclei of the song control system. There we observed predominant but not exclusive expression in neurons. The miR is found in different neuronal subsets in the different song nuclei, being conspicuously absent from the large magnocellular neurons of lMAN and RA yet present in large neurons in HVC. In
general the miR is expressed at similar levels within the song nuclei compared to surrounding brain regions. Hence it could clearly play a role in vocal behaviors but is also likely to have other roles as well.

The mir-2954 gene maps to the Z chromosome and thus is present in twice as many copies in males (ZZ) than in females (ZW). In birds, due to incomplete dosage compensation, Z-linked genes typically give rise to somewhat more RNA in males than in females (Ellegren, Hultin-Rosenberg et al. 2007; Itoh, Melamed et al. 2007). Higher expression of miR-2954 in males has previously been shown for brain tissue (Gunaratne, Lin et al. 2011) and in chicken for various other tissues (Zhao, McBride et al. 2010), and here we also measured higher male expression in various tissues of the zebra finch. We found the sex difference in miR-2954 to be much more than 2-fold in most tissues, however, indicating that factors other than gene dosage must amplify expression specifically in males or suppress it specifically in females. For example, expression of the gene could be sensitive to circulating gonadal steroids or affected by various rates of transcription, processing of pri-miRNA and pre-miRNA, or stability of mature miRNA in different tissues. However, we detected an expression difference of 100-fold in two zebra finch cell lines that differ in chromosomal sex. In this case, the expression difference must be due to cell-autonomous factors and not to the influence of extrinsic regulatory signals. This suggests that miR-2954 participates in a transcriptional regulatory network that amplifies intrinsic sex differences in gene dosage for some genes.

Prior quantitative analyses showed that, in the auditory forebrain lobule (AL), miR-2954 responds differently to song exposure in the two sexes (Gunaratne, Lin et al. 2011). We did not explicitly test for that effect again here, as the methods we used for localization in the brain (FISH) are not ideal for quantification. However, we did not see
any obvious difference in localization pattern by FISH, either between the sexes or in response to song playback (Figure 3.5). It is interesting to speculate that a sex difference in the quantitative response of this miR to song experience may lead to a broader sex difference in the brain mRNA networks after song exposure. Different transcriptional responses to song in the two sexes could contribute to the different developmental and behavioral functions of song in males and females. A direct comparison of the transcriptional response to song in males and females has yet to be reported, although indirect evidence (Drnevich, Replogle et al. submitted) and studies of individual genes (Bailey and Wade 2003) suggest differences do exist. MiRs are generally considered to function as post-transcriptional regulators of gene expression, altering either the stability (amount) or functional readout (translation) of specific target mRNAs. To test hypotheses about miR-2954 mechanism and function, we used a new zebra finch cell line which, although derived from non-neural tissue, does express a large number of neural genes (Balakrishnan, Lin et al. 2012). Observations made in this cell line may not necessarily generalize to the intact brain, but cell line studies represent a tractable first step towards more ambitious future research to manipulate miRs in the zebra finch brain itself. Using a commercially supplied sequence-specific miR-2954 inhibitor, we were successful in reducing the endogenous expression of miR-2954 in the G266 cell line (Figure 3.3A). This resulted in a significant increase in expression of one of the predicted targets of miR-2954, the transcription factor gene NR4A3 (Figure 3.3B) (Warren, Clayton et al. 2010). However, we did not observe an effect on several other predicted targets. This could reflect inadequacies of current miR prediction algorithms, especially as applied to a non-traditional species with poorly annotated 3’ UTR regions, as is the zebra finch. Or it
could be that predicted interactions occur but the effects are seen at the level of translation and not mRNA abundance.

To test more generally for consequences of miR-2954 inhibition on cell function, we employed RNA-seq technology (Mortazavi, Williams et al. 2008; Singh, Orellana et al. 2011; Tarazona, Garcia-Alcalde et al. 2011; Zhou, Xia et al. 2011) to quantitatively assess the complete profile of mRNAs in the G266 cell line after exposure to the miR-2954 inhibitor or a control sequence. RNA-seq was recently used in the primary characterization of both of the two new zebra finch cell lines (Balakrishnan, Lin et al. 2012). We observed robust, specific and functionally significant effects of miR-2954 inhibition on the mRNAs present in the G266 cells. Multiple mRNAs encoding components of the MAPK signaling system increased, and this effect was quite specific to the MAPK pathway in comparison to G-protein signaling pathway RNAs which did not change. Conversely, multiple genes involved in ribosomes and energy metabolism specifically decreased in expression. This implies that under the basal conditions of this cell line, endogenous miR-2954 expression is suppressing MAPK gene expression and supporting the expression of genes involved in energetics and protein biosynthesis. Note that these effects of miR-2954 on the gene expression networks do not have to be direct – the point is that a change in miR-2954 leads (whether directly or indirectly or both) to a complex but specific set of changes in the structure of the cell’s transcriptional network.

Song exposure also causes changes in miR-2954 expression in the brain (AL region), and a comparison with the cell line data here is intriguing. In the AL, novel song exposure initially causes a decrease in miR-2954 (Gunaratne, Lin et al. 2011). A day after song exposures, the functional profile of gene expression has changed profoundly, with marked decline in mRNAs annotated for ribosomes and mitochondrial function (the “Habituate
“Down” profile in (Dong, Replogle et al. 2009)). This appears functionally similar to the suppression of ribosomal and mitochondrial genes observed in the cell line after two days of exposure to the miR-2954 inhibitor. This comparison is at best only suggestive, of course, as there are as yet no data to define the timecourses of miR-2954 suppression in the brain after song, nor is there much information about the timecourses for emergence of mRNA changes after either song or miR inhibitor exposures. Song exposure also causes changes in at least five other miRs in addition to miR-2954, so any functional consequences of song will necessarily be some integration of multiple signals, pathways and mechanisms.

Nevertheless the parallels suggest some broad conclusions and important directions for future investigation. A single targeted miR manipulation can bring about large changes in ribosomal and mitochondrial gene expression – and in a pattern which shows similarities to the effects of perceptual experience in the brain. This suggests that there may be an integrated molecular system or pathway that links an organisms’ perceptual experience to broad regulation of energy metabolism – something that could clearly have long-term adaptive value to the organism. MiR-2954 may represent an important node in such a network, at least in birds. It would be interesting in the future to test specifically for changes in energetics (e.g., using fMRI or specific metabolic indicators) following either miR manipulation or song exposure. (A prior study has already confirmed that song exposures do lead to changes in mitochondrial energetics a day later, (Dong, Replogle et al. 2009)). Our results also suggest that the overwhelming complexity of gene regulation after natural experience might be reducible to a smaller subset of functional modules, which might be isolated and studied even in non-neural cell lines. Finally, our results here support the drive towards systems biology approaches even in non-traditional model
organisms emphasizing evolving networks of interacting genes as opposed to the
deterministic linear cascade models that have dominated molecular neurobiology in the
past.

MATERIALS AND METHODS

Animals

All procedures involving animals were conducted with protocols approved by the
University of Illinois Institutional Animal Care and Use Committee. Zebra finches used in
this study were adults (older than 90 days after hatching) and obtained from aviaries
maintained at the University of Illinois. The birds were raised in a standard breeding aviary
and were tutored under normal social conditions (i.e., by their parents or other adult birds
in the breeding colony).

For collecting different tissues, six birds including three females and three males
directly from our aviary were sacrificed by rapid decapitation. A total of nine tissues
including whole brain, muscle, heart, liver, lung, spleen, gonad, kidney and adrenal gland
were dissected from each bird, frozen on dry ice and stored in -80 freezer until RNA
purification. For collecting brains to map miR-2954 expression in zebra finch song nuclei,
another seven male and six female birds were used.

The song playback procedures and brain dissections were performed as described in
previous studies, using the same equipment (Dong, Replogle et al. 2009; Gunaratne, Lin et
al. 2011). Each bird was put individually into a sound isolation chamber for 18 hours on the
first day, and on the second day those in the song group heard 30 minutes of a song not
heard previously ("novel song"). Matched controls collected in parallel heard no song
playback ("silence"). Birds were sacrificed in song-silence pairs, so that 5 minutes before
the end of the song playback to one bird, a bird in the silence group was sacrificed by rapid decapitation, and its whole brain was removed, placed in Tissue-Tek O.C.T. compound, and rapidly frozen on dry ice. Then the brain of the song-stimulated bird was dissected and frozen in dry ice. At the end of the song stimulation procedure, all brains were transferred and stored at -80°C until sectioning process on cryostat.

**Cell culture and transfection**

Two cell lines of zebra finch were obtained from the laboratory of Dr. Arthur Arnold at University of California Los Angeles. The cell culture protocols for growing two zebra finch cell lines have been described previously and were followed in this study (Itoh and Arnold 2011). The miR-2954 sequence was submitted to a commercial supplier (Dharmacon) for designing and manufacturing the miRNA inhibitors. The inhibitors are chemically modified and synthetic single-stranded RNA molecules which can correspondingly suppress the effect of the endogenous miRNA on the target through complementary pairing to the active miRNA. The miR-2954 inhibitors and inhibitor negative controls were transfected into cells with Oligofectamine (Invitrogen) according to the manufacturer’s manual.

**Reverse transcription real-time quantitative PCR**

RNA samples were extracted from zebra finch tissues or cell lines by TRI reagent (Ambion), treated with DNase (Ambion). The MicroRNA Assay Kit (Applied Biosystems) was used for reverse transcription and real-time qPCR of miR-2954; the procedure for relative quantification of miR-2954 expression was described in the previous study (Gunaratne, Lin et al. 2011). The principle of the microRNA RT-qPCR was described in the paper of Chen et al. 2005 (Chen, Ridzon et al. 2005). The RETROscript Kit (Ambion) and SYBR Green (04913850001; Roche) qPCR was used to measure mRNA expression; The
primers for SYBR Green qPCR were designed by the Primer3 software (Rozen and Skaletsky 2000).

**Fluorescence in situ hybridization and immunohistochemistry**

Unlabeled locked nucleic acid (LNA) oligonucleotides (purchased from Exiqon) were labeled with digoxigenin (DIG) by the DIG Oligonucleotide Tailing Kit (Roche) and the labeling efficiency was estimated using the procedure of dot blot described in the manual. For negative control, an LNA-modified, DIG-labeled scramble sequence was used. The protocol of using LNA probes and tyramide signal amplification (TSA) to detect microRNA in frozen tissue sections (Silahtaroglu, Nolting et al. 2007) was followed and modified with staining of protein marker. Briefly, parasagittal sections, 10 μm in thickness, were collected beginning from midline to ~4.1 mm and alternatively placed on either a glass slide and stored at −80°C. Sections were removed from −80°C storage, fixed with 4% (wt/vol) paraformaldehyde (pH 7.6) for 5 minutes, treated with 0.25% (vol/vol) acetic anhydride/0.1 M triethanolamine for 10 minutes, and permeabilized with 0.2% Triton X-100 for 15 minutes. In situ hybridization with 2.5 pmol DIG-labeled LNA probe in hybridization buffer (50% (vol/vol) formamide, 5X SSC, 500 μg/ml yeast tRNA, 1X Denhard's solution and DEPC treated water) was carried out for 3 hours at 52°C followed by serial washes with saline-sodium citrate (SSC) buffers at 62°C. Sections were treated with 3% hydrogen peroxide to block endogenous peroxidase, washed with TN buffer (0.1 M Tris-HCl, pH 7.5 and 0.15 M NaCl) for three times and then incubated in blocking buffer (0.1 M Tris-HCl, pH 7.5, 0.15 M NaCl, 0.5% (wt/vol) blocking reagent and 0.5% (wt/vol) BSA) for 30 minutes at room temperature or overnight at 4°C.

Following blocking, IHC detection was used to detect microRNA signals and neuronal markers. In brief, sections were incubated by peroxidase-conjugated anti-DIG
antibody (1:400 in blocking buffer; Roche) for 45 minutes and then washed with TNT buffer (0.1 M Tris-HCl, pH 7.5, 0.15 M NaCl, and 0.3% Triton-X-100) for three times. After washes, the DIG signals were amplified by the Cy5-tyramide Plus Kit (1:100; PerkinElmer) for 10 minutes. To label the neurons, sections were incubated with the primary antibodies against the neuronal marker, NeuN (1:500, MAB377; Millipore) for 1 hour. After 1 hour, tissues were washed with TNT buffer (5 min, 6 times) and incubated 1 hour with the secondary antibodies Alexa 488 (1:500, A-21202; Invitrogen) and followed by TNT washes (5 min, 6 times). Slides were then dried and cover slip mounted with ProLong Gold Antifade Reagent with DAPI (P-36931; Invitrogen) for staining of cell nuclei. All IHC incubations were done at room temperature in a humidity chamber. Slides were imaged under a Zeiss LSM700 confocal microscope (Carl Zeiss Microimaging, Inc.) fluorescence microscope at the core facilities of Institute for Genomic Biology, UIUC.

**RNA sequencing**

Total RNA samples were purified 48 hours after transfection of miR-2954 inhibitor. The purified RNA samples were analyzed on Bioanalyzer (Agilent) to ensure adequate quality and quantity of RNA. In total, six RNA-seq libraries were constructed with Illumina’s TruSeq RNAseq Sample Prep kit following manufacturer’s instructions. The libraries were quantitated by qPCR, pooled and sequenced on one lane for 100 cycles on an Illumina HiSeq2000 using a TruSeq SBS sequencing kit version 3 and analyzed with Casava1.8 (pipeline 1.8). Above procedures of library construction and sequencing were done at the University of Illinois Roy J. Carver Biotechnology Center.

Reads were mapped using Tophat version 1.4.1 using the public instance of the Galaxy Server. We mapped reads using the -g 1 option in Tophat to identify uniquely mapping reads and we based differential expression tests on this subset of reads. We also
used an anchor length of 8 base pairs with no mismatches within the anchor regions. The result of Tophat mapping is a BAM file that we converted to SAM format using SAMtools. The read mapping profile was then converted into a table read counts per gene using the resulting SAM file, Ensembl gene annotations, and HT-seq. We tested for differential expression using default settings in DE-seq. The test compared the mapping profile of three replicate control libraries and three libraries that received the inhibitor treatment. To examine the functional representation of the differentially expressed transcripts, we used Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) analyses (http://bioinformatics.iah.ac.uk/tools/GOfinch; http://bioinformatics.iah.ac.uk/tools/KEGGfinch; (Wu and Watson 2009)).

**Additional Files**

Table 3.3. Results of DESeq analysis. File can be viewed at https://netfiles.uiuc.edu/ylin29/shared/TableS1_DESeq.xls. Table 3.4. Differentially expressed genes with fold changes >1.5-fold after inhibitor and <0.75-fold after inhibitor. File can be viewed at https://netfiles.uiuc.edu/ylin29/shared/TableS2_UpDownGenes_CNB.xls.

**ACKNOWLEDGEMENTS**

This study was supported by NIH grant 1RC1GM091556. We thank Lisa Stubbs and Hui Lui for helping on improvement of zebra finch 3’UTR annotation.
Table 3.1. Over- or under-represented Gene Ontology (GO) terms associated with (A) up-regulated genes and (B) down-regulated genes following miR-2954 inhibition.

(A)

<table>
<thead>
<tr>
<th>GO</th>
<th>go_description</th>
<th>total</th>
<th>exp.</th>
<th>obs.</th>
<th>adj.hyper</th>
<th>adj.fisher</th>
</tr>
</thead>
<tbody>
<tr>
<td>Up-Regulated Genes:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GO:0004674</td>
<td>protein serine/threonine kinase activity</td>
<td>428</td>
<td>31</td>
<td>66</td>
<td>2.50E-06</td>
<td>3.70E-06</td>
</tr>
<tr>
<td>GO:0004713</td>
<td>protein tyrosine kinase activity</td>
<td>403</td>
<td>29</td>
<td>63</td>
<td>2.50E-06</td>
<td>3.70E-06</td>
</tr>
<tr>
<td>GO:0004672</td>
<td>protein kinase activity</td>
<td>441</td>
<td>32</td>
<td>63</td>
<td>5.40E-05</td>
<td>6.80E-05</td>
</tr>
<tr>
<td>GO:0006468</td>
<td>protein amino acid phosphorylation</td>
<td>502</td>
<td>36</td>
<td>68</td>
<td>1.00E-04</td>
<td>0.00012</td>
</tr>
<tr>
<td>GO:0046777</td>
<td>protein amino acid autophosphorylation</td>
<td>35</td>
<td>3</td>
<td>13</td>
<td>0.00012</td>
<td>0.00012</td>
</tr>
<tr>
<td>GO:0007243</td>
<td>protein kinase cascade</td>
<td>30</td>
<td>2</td>
<td>11</td>
<td>0.00077</td>
<td>9.00E-04</td>
</tr>
<tr>
<td>GO:0008270</td>
<td>zinc ion binding</td>
<td>1126</td>
<td>82</td>
<td>121</td>
<td>0.00077</td>
<td>0.0011</td>
</tr>
<tr>
<td>GO:0007186</td>
<td>G-protein coupled receptor protein signaling</td>
<td>371</td>
<td>27</td>
<td>9</td>
<td>1</td>
<td>0.0094</td>
</tr>
<tr>
<td>GO:0005097</td>
<td>Rab GTPase activator activity</td>
<td>37</td>
<td>3</td>
<td>10</td>
<td>0.028</td>
<td>0.028</td>
</tr>
<tr>
<td>GO:0032313</td>
<td>regulation of Rab GTPase activity</td>
<td>37</td>
<td>3</td>
<td>10</td>
<td>0.028</td>
<td>0.028</td>
</tr>
<tr>
<td>GO:0005515</td>
<td>protein binding</td>
<td>3113</td>
<td>226</td>
<td>271</td>
<td>0.025</td>
<td>0.034</td>
</tr>
<tr>
<td>GO</td>
<td>go_description</td>
<td>total</td>
<td>exp.</td>
<td>obs.</td>
<td>adj.hyper</td>
<td>adj.fisher</td>
</tr>
<tr>
<td>-------------</td>
<td>---------------------------------------------</td>
<td>-------</td>
<td>-----</td>
<td>-------</td>
<td>-----------</td>
<td>------------</td>
</tr>
<tr>
<td>GO:0005840</td>
<td>ribosome</td>
<td>142</td>
<td>10</td>
<td>82</td>
<td>1.10E-54</td>
<td>1.10E-54</td>
</tr>
<tr>
<td>GO:0003735</td>
<td>structural constituent of ribosome</td>
<td>134</td>
<td>10</td>
<td>77</td>
<td>3.00E-51</td>
<td>3.00E-51</td>
</tr>
<tr>
<td>GO:0006412</td>
<td>translation</td>
<td>198</td>
<td>14</td>
<td>88</td>
<td>9.30E-47</td>
<td>9.30E-47</td>
</tr>
<tr>
<td>GO:0005739</td>
<td>mitochondrial</td>
<td>371</td>
<td>26</td>
<td>78</td>
<td>1.70E-16</td>
<td>1.70E-16</td>
</tr>
<tr>
<td>GO:0016020</td>
<td>membrane</td>
<td>1247</td>
<td>88</td>
<td>33</td>
<td>1.30E-10</td>
<td>1.30E-10</td>
</tr>
<tr>
<td>GO:0022627</td>
<td>cytosolic small ribosomal subunit</td>
<td>11</td>
<td>1</td>
<td>9</td>
<td>4.30E-07</td>
<td>3.50E-07</td>
</tr>
<tr>
<td>GO:0005743</td>
<td>mitochondrial inner membrane</td>
<td>131</td>
<td>9</td>
<td>28</td>
<td>1.70E-05</td>
<td>1.50E-05</td>
</tr>
<tr>
<td>GO:0007165</td>
<td>signal transduction</td>
<td>466</td>
<td>33</td>
<td>8</td>
<td>1.80E-05</td>
<td>1.80E-05</td>
</tr>
<tr>
<td>GO:0015935</td>
<td>small ribosomal subunit</td>
<td>10</td>
<td>1</td>
<td>7</td>
<td>0.00013</td>
<td>9.80E-05</td>
</tr>
<tr>
<td>GO:0005839</td>
<td>proteasome core complex</td>
<td>14</td>
<td>1</td>
<td>8</td>
<td>0.00014</td>
<td>0.00012</td>
</tr>
<tr>
<td>GO:0004298</td>
<td>threonine-type endopeptidase activity</td>
<td>14</td>
<td>1</td>
<td>8</td>
<td>0.00014</td>
<td>0.00012</td>
</tr>
<tr>
<td>GO:0006468</td>
<td>protein amino acid phosphorylation</td>
<td>502</td>
<td>36</td>
<td>12</td>
<td>0.00024</td>
<td></td>
</tr>
<tr>
<td>GO:0004672</td>
<td>protein kinase activity</td>
<td>441</td>
<td>31</td>
<td>10</td>
<td>0.00053</td>
<td></td>
</tr>
<tr>
<td>GO:000276</td>
<td>mitochondrial proton-transporting ATP</td>
<td>6</td>
<td>0</td>
<td>5</td>
<td>0.00085</td>
<td>0.00064</td>
</tr>
<tr>
<td>GO:0005762</td>
<td>mitochondrial large ribosomal subunit</td>
<td>6</td>
<td>0</td>
<td>5</td>
<td>0.00085</td>
<td>0.00064</td>
</tr>
<tr>
<td>GO:0006414</td>
<td>translational elongation</td>
<td>6</td>
<td>0</td>
<td>5</td>
<td>0.00085</td>
<td>0.00064</td>
</tr>
<tr>
<td>GO:0004674</td>
<td>protein serine/threonine kinase activity</td>
<td>428</td>
<td>30</td>
<td>10</td>
<td>0.00075</td>
<td></td>
</tr>
<tr>
<td>GO:0003700</td>
<td>transcription factor activity</td>
<td>452</td>
<td>32</td>
<td>11</td>
<td>0.00075</td>
<td></td>
</tr>
<tr>
<td>GO:0004713</td>
<td>protein tyrosine kinase activity</td>
<td>403</td>
<td>29</td>
<td>9</td>
<td>0.00089</td>
<td></td>
</tr>
<tr>
<td>GO:0004129</td>
<td>cytochrome-c oxidase activity</td>
<td>14</td>
<td>1</td>
<td>7</td>
<td>0.0015</td>
<td>0.00098</td>
</tr>
<tr>
<td>GO:0016021</td>
<td>integral to membrane</td>
<td>894</td>
<td>63</td>
<td>35</td>
<td>0.002</td>
<td></td>
</tr>
<tr>
<td>GO:003565</td>
<td>sequence-specific DNA binding</td>
<td>305</td>
<td>22</td>
<td>6</td>
<td>0.0043</td>
<td></td>
</tr>
<tr>
<td>GO:0043234</td>
<td>protein complex</td>
<td>49</td>
<td>3</td>
<td>12</td>
<td>0.0086</td>
<td>0.0052</td>
</tr>
<tr>
<td>GO:0015078</td>
<td>hydrogen ion transmembrane transporter activity</td>
<td>24</td>
<td>2</td>
<td>8</td>
<td>0.01</td>
<td>0.0066</td>
</tr>
<tr>
<td>GO:0030529</td>
<td>ribonucleoprotein complex</td>
<td>24</td>
<td>2</td>
<td>8</td>
<td>0.01</td>
<td>0.0066</td>
</tr>
<tr>
<td>GO:0005677</td>
<td>chromatin silencing complex</td>
<td>9</td>
<td>1</td>
<td>5</td>
<td>0.01</td>
<td>0.0069</td>
</tr>
<tr>
<td>GO:0015986</td>
<td>ATP synthesis coupled proton transport</td>
<td>46</td>
<td>3</td>
<td>11</td>
<td>0.015</td>
<td>0.011</td>
</tr>
<tr>
<td>GO:0003746</td>
<td>translation elongation factor activity</td>
<td>6</td>
<td>0</td>
<td>4</td>
<td>0.017</td>
<td>0.012</td>
</tr>
<tr>
<td>GO:0003723</td>
<td>RNA binding</td>
<td>186</td>
<td>13</td>
<td>27</td>
<td>0.015</td>
<td>0.014</td>
</tr>
<tr>
<td>GO:0007242</td>
<td>intracellular signaling cascade</td>
<td>223</td>
<td>16</td>
<td>4</td>
<td>0.018</td>
<td></td>
</tr>
<tr>
<td>GO:0015934</td>
<td>large ribosomal subunit</td>
<td>7</td>
<td>0</td>
<td>4</td>
<td>0.036</td>
<td>0.024</td>
</tr>
<tr>
<td>GO:0006364</td>
<td>rRNA processing</td>
<td>23</td>
<td>2</td>
<td>7</td>
<td>0.036</td>
<td>0.025</td>
</tr>
<tr>
<td>GO:0006355</td>
<td>regulation of transcription, DNA-dependent</td>
<td>558</td>
<td>40</td>
<td>21</td>
<td>0.028</td>
<td></td>
</tr>
<tr>
<td>GO:0006811</td>
<td>ion transport</td>
<td>132</td>
<td>9</td>
<td>1</td>
<td>0.03</td>
<td></td>
</tr>
<tr>
<td>GO:0048147</td>
<td>negative regulation of fibroblast proliferation</td>
<td>4</td>
<td>0</td>
<td>3</td>
<td>0.048</td>
<td>0.034</td>
</tr>
<tr>
<td>GO:0005744</td>
<td>mitochondrial inner membrane presequence translocase complex</td>
<td>4</td>
<td>0</td>
<td>3</td>
<td>0.048</td>
<td>0.034</td>
</tr>
<tr>
<td>GO:0006563</td>
<td>L-serine metabolic process</td>
<td>4</td>
<td>0</td>
<td>3</td>
<td>0.048</td>
<td>0.034</td>
</tr>
<tr>
<td>GO:0006476</td>
<td>protein amino acid deacetylation</td>
<td>8</td>
<td>1</td>
<td>4</td>
<td>0.048</td>
<td>0.034</td>
</tr>
<tr>
<td>GO:0017136</td>
<td>NAD-dependent histone deacetylase activity</td>
<td>8</td>
<td>1</td>
<td>4</td>
<td>0.048</td>
<td>0.034</td>
</tr>
<tr>
<td>GO:0005740</td>
<td>mitochondrial envelope</td>
<td>8</td>
<td>1</td>
<td>4</td>
<td>0.048</td>
<td>0.034</td>
</tr>
<tr>
<td>GO:0008121</td>
<td>ubiquinol-cytochrome-c reductase activity</td>
<td>8</td>
<td>1</td>
<td>4</td>
<td>0.048</td>
<td>0.034</td>
</tr>
<tr>
<td>GO:0006284</td>
<td>base-excision repair</td>
<td>13</td>
<td>1</td>
<td>5</td>
<td>0.048</td>
<td>0.034</td>
</tr>
</tbody>
</table>
Table 3.2. Significant representation of KEGG pathways associated with the differentially expressed genes.

(A)

<table>
<thead>
<tr>
<th>pathway_id</th>
<th>pathway_title</th>
<th>total</th>
<th>exp.</th>
<th>obs.</th>
<th>adj.hyper</th>
<th>adj.fisher</th>
</tr>
</thead>
<tbody>
<tr>
<td>gga00564</td>
<td>Glycerophospholipid metabolism</td>
<td>27</td>
<td>2</td>
<td>10</td>
<td>0.00045</td>
<td>0.00045</td>
</tr>
<tr>
<td>gga04010</td>
<td>MAPK signaling pathway</td>
<td>121</td>
<td>9</td>
<td>22</td>
<td>0.00061</td>
<td>0.00061</td>
</tr>
</tbody>
</table>

(B)

<table>
<thead>
<tr>
<th>pathway_id</th>
<th>pathway_title</th>
<th>total</th>
<th>exp.</th>
<th>obs.</th>
<th>adj.hyper</th>
<th>adj.fisher</th>
</tr>
</thead>
<tbody>
<tr>
<td>gga03010</td>
<td>Ribosome</td>
<td>54</td>
<td>6</td>
<td>49</td>
<td>7.90E-44</td>
<td>7.90E-44</td>
</tr>
<tr>
<td>gga00190</td>
<td>Oxidative phosphorylation</td>
<td>69</td>
<td>7</td>
<td>26</td>
<td>2.60E-08</td>
<td>2.60E-08</td>
</tr>
<tr>
<td>gga04010</td>
<td>MAPK signaling pathway</td>
<td>121</td>
<td>13</td>
<td>1</td>
<td>1</td>
<td>0.00044</td>
</tr>
<tr>
<td>gga04310</td>
<td>Wnt signaling pathway</td>
<td>85</td>
<td>9</td>
<td>1</td>
<td>1</td>
<td>0.026</td>
</tr>
<tr>
<td>gga04120</td>
<td>Ubiquitin mediated proteolysis</td>
<td>77</td>
<td>8</td>
<td>1</td>
<td>1</td>
<td>0.044</td>
</tr>
<tr>
<td>gga03050</td>
<td>Proteasome</td>
<td>26</td>
<td>3</td>
<td>8</td>
<td>0.087</td>
<td>0.044</td>
</tr>
</tbody>
</table>
Figure 3.1. Expression profile of miR-2954 in nine major tissues and two cell lines of zebra finch.

Each bar represents the log2 mean with the standard error of mean (SEM) of relative miR expression normalized to the value in female cell line because of its lowest expression level (Ct mean: 27 cycles). Each tissue value represents the mean of samples from three separate birds for each sex. Each cell line value represents the mean of three separate culture dishes.
Figure 3.2. Mapping of miR-2954 expression in zebra finch song nuclei.
Figure 3.2. (cont.) Mapping of miR-2954 expression in zebra finch song nuclei.
Figure 3.2. (cont.) Mapping of miR-2954 expression in zebra finch song nuclei.
(A) Overview of miR-2954 expression in HVC, Area X, the lateral magnocellular nucleus of the anterior nidopallium (LMAN), the robust nucleus of the arcopallium (RA) and the caudal medial nidopallium (NCM). Scale bar represents 0.2 mm. (B) Zoom in images in the song nuclei show the overlapping of miR-2954 and NeuN. Scale bar represents 0.02 mm.
Figure 3.3. Effects of miR-2954 inhibitor.
(A) Dose response of miR-2954 inhibitor. (B) qPCR validated NR4A3 induction after knocking down miR-2954.
Figure 3.4. Overview of miR-2954 expression in the auditory forebrain of zebra finch. Tiling images of parasagittal sections under the 10X objective are shown.
Figure 3.5. Cellular localization of miR-2954 expression in the auditory forebrain of zebra finch.

Images taken using 60X objective are shown.
Figure 3.6. Up-regulated genes by miR-2954 inhibitor are enriched in MAPK pathway and colored by the online tool of KEGG Mapper. The homologous genes found in the genome of zebra finch are in green boxes; the up-regulated genes by miR-2954 inhibitor are in red boxes.
CHAPTER 4 SUMMARY AND DISCUSSION

OVERALL SUMMARY

The ability to use learned vocalizations (song) for communication make zebra finches great model animals to study gene, brain and behavior. Previous studies have shown that song exposure alters the network of genes expressed in the auditory forebrain and that ncRNAs are major components of the song-regulated gene network (Dong et al., 2009). In this thesis, I intended to test the hypothesis that miRs, a particular type of ncRNAs, have important functions for coordinating the complex experience-dependent gene network in the zebra finch brain. I accomplished this by: 1) characterizing the miR population expressed in zebra finch auditory forebrain; 2) identifying specific miRs regulated by song stimulation; 3) demonstrating the regulatory consequences of perturbing the expression of one unusual song-regulated miR, miR-2954. I will briefly describe the major findings of this thesis and discuss their importance in the following section.

THE SIGNIFICANCE OF MAJOR FINDINGS

MiR expression and song regulation in zebra finch auditory forebrain

The rapid advances in sequencing technology have improved sensitivity to detect small RNA sequences. Using RNA-seq, the collaborators and I have reported 155 miR sequences identified from zebra finch auditory forebrain and submitted them to the miR database (Gunaratne, Lin et al, 2011). These provide molecular candidates for studying miR functions associated with songbird brain processes. Another interesting finding from our primary RNA-seq analysis is that many of the novel miR sequences have not been found in chicken. As more songbird genomes become available in the near future, it will be interesting and feasible to investigate whether the novel miRs have evolved specifically in
the songbird lineage and might that have roles in the evolution of vocal learning ability in this lineage.

Song regulation has been confirmed by qPCR for six miRs. More studies focusing on these song-responsive miRs would aid in comprehending the roles miRs play in perception, recognition or memorization of song. In addition, this thesis focuses very specifically on one brain region, the auditory forebrain of adult zebra finches. Whether similar or diverse miRs exist in other song nuclei or at different developmental stages remains to be examined.

**Expression of miR-2954 in zebra finch tissues, cells and song nuclei**

In examining the tissue specificity of miR-2954, I have found that miR-2954 is ubiquitously expressed in a variety of zebra finch tissues; and like most Z-linked genes, the sex-biased expression pattern has been validated. The presence of miR-2954 expression mostly in neurons of the song nuclei examined has been determined. This collected information provides a detailed spatial expression profile of miR-2954 and suggests it might be involved in various physiological functions. Some of these functions could be related to the sex differences in the way males and females respond to song exposure.

**MiR-2954 function associated with kinase and ribosome- or mitochondria-related genes**

Using a zebra finch cell line, I performed a second RNA-seq experiment in which miR-2954 expression was manipulated using an antisense inhibitor, to directly test whether miR-2954 expression affects endogenous gene expression at a global level. Additionally, pharmacological suppression of miR-2954 mimics the decline that occurs in the brain in response to song exposure. The RNA-seq results confirm that miR-2954 suppression significantly alters the landscape of gene expression. Intriguingly, the functional profile of
these gene expression changes resembles the profile that occurs after song training in the phenomenon of song habituation (Dong et al., 2009, Warren et al., 2010). The similarities include activation of transcription factors and down-regulation of ribosome- or mitochondria-related genes, implying that miR-2954 plays a role in cellular energetics and metabolism.

**FUTURE STUDIES**

The results here set the stage for a number of future investigations of fundamental interest to the neurobiology of behavior. If miR-2954 has an effect on gene expression networks as I have shown in Chapter 3, what about the roles of the other song-responsive miRs I identified in Chapter 2? Are their combined effects redundant, or additive, or more complex in some non-linear way? It will also be interesting in the future to test whether the effects of changing miR-2954 expression are different in females and males. This could be done using the female zebra finch cell line ZFTMA to complement the studies here using the male G266 line. Finally, it will be interesting to test the effects of manipulating miR-2954 expression in the behaving animal. For example, might miR-2954 inhibition in the song nuclei interfere with developmental song learning? Might inhibition in NCM in adults disrupt song recognition and habituation? Experiments like these may provide a deeper understanding of how specific molecular mechanisms underlie the evolution and expression of important brain functions and behaviors.
REFERENCES


