Journal of Biology

MINIREVIEW

Genome of a songbird unveiled

Raphael Pinaud*

Abstract

An international collaborative effort has recently uncovered the genome of the zebra finch, a songbird model that has provided unique insights into an array of biological phenomena.

The study of songbirds has revealed a variety of fundamental properties of biological systems. In particular, neurobiological studies carried out in songbirds have revealed the presence of newly born neurons in the adult brain, how steroid hormones affect brain development, the neural and mechanistic bases of vocalizations, and how experience modifies neuronal physiology. More evidently, however, songbirds have been extensively used as a model for imitative vocal learning, a behavior thought to be a substrate for speech acquisition in humans [1,2].

Now an international consortium has unveiled the genome of the zebra finch (*Taeniopygia guttata,* Figure 1), along with a multi-layered analysis of its sequence [3]. Sequencing the zebra finch genome was initiated in 2005 under the Large Scale Genome Sequencing Program of the National Human Genome Research Institute [4], leveraging prior work in the research community characterizing the zebra finch brain transcriptome [5-7]. These initiatives, along with new zebra finch genome sequences, have resulted in the complete genome sequenced with 17,475 protein-coding genes identified, as well as regulatory regions and non-coding RNAs. The annotation and sequence coverage of the zebra finch genome will certainly be refined in the years to come, but the initial endeavor is expected to provide a unique platform for modern genomics research in this organism. Furthermore, this initial snapshot of the songbird genome should provide critical insights into fundamental

scientific questions, including an array of physiological and evolutionary processes. Here, I review some of the most exciting findings of this pioneering effort.

Development and the brain

The zebra finch genome project [3] has revealed that nearly 10,000 genes are expressed in the forebrain of juvenile birds (50 days after hatching; within the critical period for vocal learning) and adult birds, with an overlap of approximately 91% across these age groups. These results indicate that up to 60% of the genes in the genome are expressed in the brain at any one given time. These findings also suggest that a significant fraction of protein-coding genes (9%) are developmentally regulated in the songbird brain, consistent with previous observations obtained with forward genetic approaches.

Sensory- and motor-regulated transcripts

Auditory experience, a fundamental consequence of social interactions within and across songbird species, had been previously shown to strongly affect gene regulatory events in the auditory forebrain [8]. It was found [3] that in the auditory forebrain of animals in silent conditions, approximately 40% of the detected transcripts are non-coding, indicating that regulatory microRNAs may have a central role in brain homeostasis. When birds were stimulated with playbacks of recorded song, thousands of transcripts were upregulated or downregulated [9], and analyses of their genomic sequences revealed that roughly two-thirds of the down-regulated transcripts were non-coding RNAs. Furthermore, known and novel microRNAs were found to be expressed in the auditory forebrain, and their binding sites were detected in the untranslated regions of regulated genes.

Singing behavior also drives robust gene expression programs in structures of the song control system, a specialized brain network required for sensorimotor integration and vocal output [5,10]. By using a microarray platform with oligonucleotides generated as part of this project, the songbird genome consortium [3] was able to uncover a series of transcriptional regulators whose

*Correspondence: pinaud@bcs.rochester.edu
Department of Brain and Cognitive Sciences, 117 Meliora Hall, River Campus, University of Rochester, Rochester, NY 14627, USA

Figure 1. Zebra finches (*Taeniopygia guttata*): an adult female (left) and an adult male (right).

expression was modulated by the act of singing. Changes in transcription factor expression that occurred early after singing were strongly correlated with later modifications in the expression patterns of groups of their predicted target genes [3]. In fact, many of these targets have been identified for the first time and will now enable researchers to develop testable hypotheses about the gene regulatory interactions that are induced during a learned behavior.

Overall, these findings [3] highlight the role of micro-RNAs and non-coding RNAs in the control of gene expression in the songbird brain, in addition to the active regulation of transcription factors and their respective target genes. When comparing hearing-driven transcripts with genes thought to have been positively selected in songbirds, a significant over-representation of genes encoding ion channels was uncovered [3], consistent with robust and complex expression patterns of ion channel-associated transcripts in stations of the song-control circuit [11,12].

Genes gained, genes lost

The unveiling of the zebra finch genome also provides exciting insights into the evolution of avian and mammalian species. As detailed by the consortium authors [3], the genome lacks genes that encode milk, salivary and vomeronasal receptor proteins, similarly to what has been documented for the chicken, a non-vocal-learning avian species whose genome was uncovered 6 years ago [13]. Curiously, similarly to chickens, zebra finches lack the synapsin I gene, which encodes a phosphoprotein involved in the regulation of neurotransmitter vesicle availability in pre-synaptic membranes. This finding suggests that the synaptic transmission machinery differs between mammalian and avian species, although it is not clear if such molecular changes translate into functional modifications at the systems level.

Duplications of a variety of genes relative to chickens or humans, including growth hormone and caspase-3, the latter of which is associated with the induction of apoptosis, and gene family expansions, including of the *PAK3* and *PHF7* genes, which are involved in dendritic plasticity and transcriptional regulation, respectively, were also found in the zebra finch genome [3]. Interestingly, multiple duplications of the *PHF7* gene seem to have occurred independently in zebra finches and chickens, suggesting that some aspects of transcriptional regulation may have been under evolutionary pressure in avian species. Whereas these avian lineages have groups of 17 and 18 *PHF7* genes, respectively, the human genome has been shown to contain only one *PHF7* gene [3].

Finally, the zebra finch genome was found to have a significant fraction of transcribed mobile elements and a higher degree of intrachromosomal rearrangement relative to chicken. An example detailed by the consortium authors [3] refers to genes of the major histocompatibility complex, which are scattered across several chromosomes in the zebra finch genome; in the chicken and human genomes, such genes have a well established syntenic organization. Despite these informative species-specific differences, the population of coding genes and the syntenic organization of the zebra finch genome were found to be highly similar to that of the chicken and, in many respects, to that of humans.

An exciting future for songbird biological studies

As a result of the pioneering efforts of the consortium and in addition to the sequencing and annotation of the genome, an array of publicly available resources and tools has been developed for songbird studies. These include normalized and subtracted cDNA libraries and bacterial artificial chromosome libraries, a largely complete set of annotated expressed sequence tags, and a microarray platform [3,6,7,11]. Such tools have enabled multiple research groups, independently and in collaboration, to systematically study the functional organization of the songbird brain and its genomic response to a variety of conditions, including sensory experience, hormonal manipulations and sensory-motor learning. The outcomes of this research have been blossoming into exciting recent advances, including, but not limited to, insights into the estradiol-synthetic pathway [14], the repertoire of proteases (the 'degradome') [15] and the collection of neuropeptide prohormones and their processed peptides (the neuropeptidome) [16]. These efforts have revealed key information on genes related to steroid receptors and estrogen biosynthesis [14], insights into how proteases may shape neuronal functional, immunological and developmental processes [15], as well as the identity and expression patterns of an array of

neuropeptides thought to be involved in the development and functionality of brain circuits involved in vocal communication [16].

Over the next few years these efforts will contribute to an integrative understanding of how the songbird genomic machinery responds to environmental and physiological challenges and, more broadly, how the songbird brain is functionally organized. In addition, active research in these areas is expected to shed light on basic biological and evolutionary principles in vertebrates. The importance of a complete understanding of the songbird transcriptome is highlighted by ongoing, contiguous research ventures aimed at creating a songbird gene expression brain atlas. Finally, the study of songbird biology is reaching an exciting era with the convergence of the genomic resources detailed above and the successful development of transgenic zebra finches using a lentiviral-vector approach [17]. This interface will provide a unique opportunity for songbird biologists to test causal relationships between the induction of gene expression programs, altered cellular physiology and their behavioral correlates.

Resources for exploring the sequence and annotation data are available on browser displays at UCSC [18], Ensembl and the NCBI and at [19].

References

1. Doupe AJ, Kuhl PK: **Birdsong and human speech: common themes and mechanisms.** *Annu Rev Neurosci* 1999, **22**:567-631.
2. Jarvis ED: **Learned birdsong and the neurobiology of human language.** *Ann N Y Acad Sci* 2004, **1016**:749-777.
3. Warren WC, Clayton DF, Ellegren H, Arnold AP, Hillier LW, Kunstner A, Searle S, White S, Vilella AJ, Fairley S, Heger A, Kong L, Ponting CP, Jarvis ED, Mello CV, Minx P, Lovell P, Velho TAF, Ferris M, Balakrishnan CN, Sinha S, Blatti C, London SE, Li Y, Lin Y-C, George J, Sweedler J, Southey B, Gunaratne P, Watson M, *et al.*: **The genome of a songbird.** *Nature* 2010,**463**:536-539.
4. Proposal to Sequence the Genome of the Zebra Finch (Taeniopygia guttata) [http://www.genome.gov/Pages/Research/Sequencing/SeqProposals/ZebraFinchSeq2.pdf]
5. Wada K, Howard JT, McConnell P, Whitney O, Lints T, Rivas MV, Horita H, Patterson MA, White SA, Scharff C, Haesler S, Zhao S, Sakaguchi H, Hagiwara M, Shiraki T, Hirozane-Kishikawa T, Skene P, Hayashizaki Y, Carninci P, Jarvis ED: **A molecular neuroethological approach for identifying and characterizing a cascade of behaviorally regulated genes.** *Proc Natl Acad Sci USA* 2006, **103**:15212-15217.
6. Li X, Wang XJ, Tannenhauser J, Podell S, Mukherjee P, Hertel M, Biane J, Masuda S, Nottebohm F, Gaasterland T: **Genomic resources for songbird research and their use in characterizing gene expression during brain development.** *Proc Natl Acad Sci USA* 2007, **104**:6834-6839.
7. Replogle K, Arnold AP, Ball GF, Band M, Bensch S, Brenowitz EA, Dong S, Drnevich J, Ferris M, George JM, Gong G, Hasselquist D, Hernandez AG, Kim R, Lewin HA, Liu L, Lovell PV, Mello CV, Naurin S, Rodriguez-Zas S, Thimmapuram J, Wade J, Clayton DF: **The Songbird Neurogenomics (SoNG) Initiative: community-based tools and strategies for study of brain gene function and evolution.** *BMC Genomics* 2008, **9**:131.
8. Mello CV, Velho TA, Pinaud R: **Song-induced gene expression: a window on song auditory processing and perception.** *Ann N Y Acad Sci* 2004, **1016**:263-281.
9. Dong S, Replogle KL, Hasadsri L, Imai BS, Yau PM, Rodriguez-Zas S, Southey BR, Sweedler JV, Clayton DF: **Discrete molecular states in the brain accompany changing responses to a vocal signal.** *Proc Natl Acad Sci USA* 2009, **106**:11364-11369.
10. Jarvis ED, Nottebohm F: **Motor-driven gene expression.** *Proc Natl Acad Sci USA* 1997, **94**:4097-4102.
11. Wada K, Sakaguchi H, Jarvis ED, Hagiwara M: **Differential expression of glutamate receptors in avian neural pathways for learned vocalization.** *J Comp Neurol* 2004, **476**:44-64.
12. Lovell PV, Clayton DF, Replogle KL, Mello CV: **Birdsong "transcriptomics": neurochemical specializations of the oscine song system.** *PLoS ONE* 2008, **3**:e3440.
13. International Chicken Genome Sequencing Consortium: **Sequence and comparative analysis of the chicken genome provide unique perspectives on vertebrate evolution.** *Nature* 2004, **432**:695-716.
14. London SE, Clayton DF: **Genomic and neural analysis of the estradiol-synthetic pathway in the zebra finch.** *BMC Neurosci* 2010, **11**:46.
15. Quesada V, Velasco G, Puente XS, Warren WC, Lopez-Otin C: **Comparative genomic analysis of the zebra finch degradome provides new insights into evolution of proteases in birds and mammals.** *BMC Genomics* 2010, **11**:220.
16. Xie F, London SE, Southey BR, Annangudi SP, Amare A, Rodriguez-Zas SL, Clayton DF and Sweedler JV: **The zebra finch neuropeptidome: prediction, detection and expression.** *BMC Biol* 2010, **8**:28.
17. Agate RJ, Scott BB, Haripal B, Lois C, Nottebohm F: **Transgenic songbirds offer an opportunity to develop a genetic model for vocal learning.** *Proc Natl Acad Sci USA* 2009, **106**:17963-17967.
18. **UCSC Genome Bioinformatic** [http://genome.ucsc.edu]
19. **Avian Genomes** [http://aviangenomes.org]