Proposal to Sequence the Genome of the Zebra Finch (*Taeniopygia guttata*)

Submitted to NHGRI by the Zebra Finch Genome Consortium*, 10 July 2005
(contact: David Clayton, dclayton@uiuc.edu, 217-244-3668)

Summary: We enthusiastically advocate the zebra finch, a songbird, as a target of highest priority for whole genome sequencing. The zebra finch is an established research organism with particular relevance to human neuroscience and a proven track record of influential ground breaking discoveries that have informed human biology. Evolutionary considerations, such as distance from other sequenced genomes, and conserved and convergent functions, put the zebra finch in a unique position to inform human biology even more deeply. Zebra finches and other songbirds are well-observed by humans in both natural and domestic contexts, so their study contributes also to environmental and population biology. Full application of this powerful biological model to human biology has been hampered primarily by the lack of genetic/genomic tools. It is time to correct this deficiency. Sequence investment here will be rewarded in spades, by advancing a powerful model system that illuminates both basic biological processes and larger integrative issues (evolution, communication, environment) relevant to human health and disease. A collaborative, international group of capable and highly motivated investigators has come together to press this scientific agenda forward.

Table of Contents

A. Specific biological/biomedical rationales for utility of new sequence data

1. Improving human health.

Fundamental discoveries in songbirds have already influenced and informed human disease research. Four examples illustrate the potential impact of a zebra finch whole genome sequence.

a. Adult neurogenesis.

In 1984 Goldman and Nottebohm showed that the adult songbird brain makes new neurons, in contrast to the universally held belief that neurogenesis does not occur in adults. This discovery catalyzed a reexamination of the dogma, with the result that it is now realized that specific regions of the mammalian (including human) brain also make a considerable number of neurons in adulthood. It is fair to say that this discovery has led to a large shift in the field of neurology. The promise of stem cell biology for treatment of neurological diseases was foreshadowed by Nottebohm's 1985 book, *Hope for a New Neurology.* The adult songbird telencephalon remains one of the few places to study the functional significance and control of adult neurogenesis because of well-defined circuits linked to specific behaviors (Nottebohm, 2004). The challenge now is to understand the high-level biological processes that orchestrate and target the generation and replacement of neurons into functional circuits. For this we must think beyond the simple identification of individual proteins, and consider more broadly the integrative mechanisms that work through regulatory elements, organization of chromosomes, non-coding RNAs and the like – all of which require a high quality genome sequence.

b. Steroid hormone synthesis in brain.

Estrogen is normally thought of as a gonadal steroid, but in zebra finches it is also synthesized actively in the brain. This was convincingly demonstrated in a recent study that showed estrogen of neural origin causes masculine patterns of neural development (Holloway and Clayton, 2001), the first time that sex steroids of neural origin have been directly related to sex differences in neural development (Schlinger et al., 2001). Moreover, the zebra finch brain has all of the enzymes needed for *de novo* synthesis of testosterone and estradiol from cholesterol, similar to steroidogenic cells of the gonads. The exceptional steroid synthetic capacity of the finch brain makes this system a unique resource for understanding the roles of brain-derived steroids and for exploring the therapeutic possibilities of manipulating neurosteroid production (e.g., as a strategy for preventing or ameliorating age-related neurodegenerative disease). High quality whole-genome sequence will help us understand the essential multigenic orchestration of the various synthetic enzymes and receptor/response systems involved.

c. Avian virology and host defense mechanisms.

There is considerable interest in the evolution of immunoglobulin genes in songbirds (Edwards et al., 1999). Moreover, avian virology is assuming new importance for humans with the realization that critical human viruses are carried also by birds. These include influenza A viruses, Japanese encephalitis (MacKenzie et al., 2004; Hollenbeck, 2005), and West Nile virus in North America (Reed et al., 2003). Songbirds may spread some of these viruses because of natural migration (Reed et al., 2003). In North America, songbird populations are a barometer of West Nile infection rates. In Asia, the bird-human connection has resulted not only in lethal infection of humans, but also has had a huge economic impact because of the widespread slaughter of poultry based on a fear of virus co-infection. Avian viruses that normally do not infect humans may serve as a reservoir for generation of new strains via genetic reassortment with human viruses in intermediary species (Kida et al., 1994), which could result in novel and highly virulent strains infecting humans. Sequencing of the zebra finch genome will provide valuable information about the evolution of host defense mechanisms and infectious agents that challenge those mechanisms.

d. The relation of ecology to human health.

Songbirds are conspicuous, diurnal species that are amenable to study in the wild. Songbirds hold a prominent place in studies of reproduction, behavioral and population ecology, and perturbations of the environment. Songbird research played a pivotal role in realizing the impact of endocrine disrupters such as DDT (Carson, 1962). With drastic environmental change arguably representing one of the greatest threats to human health, genetic and other monitoring of natural populations of songbirds has assumed greater importance. Like the canary in a mine, songbirds provide an early warning system for the negative effects of environmental pollutants on human heath. Sequencing of the zebra finch genome will give environmental biologists a host of new genetic tools for monitoring populations of songbirds, and thus the environment.

2. Informing human biology.

How the zebra finch genomic sequence will lead to a better understanding of biological function in the human:

a. The complex nervous system.

For a long time, birds in general were underappreciated for their sophistication and their potential relevance to human behavioral and neural biology. Birds and mammals evolved brains that are superficially different in organization, especially as regards the laminar structure of the mammalian cortex versus the nuclear organization of the avian forebrain. Modern neurochemical, embryological and molecular data, however, now indicate a much greater functional similarity to the mammalian brain than previously suspected. This realization led recently to a revised avian brain nomenclature, which now emphasizes and clarifies these similarities (Jarvis et al., 2005). Among birds, songbirds show the greatest neural complexity and the most interesting parallels to the human. By at least one important measure (telencephalization), the zebra finch is significantly more human-like than is chicken or pigeon – and more so even than a rat or a mouse (Johnson and Whitney, 2005). Thus, songbird research is now much more relevant than it seemed a few decades ago for understanding the human brain. The zebra finch genome sequence will help identify genomic factors to underlie complex brain evolution and function.

b. Speech and language.

The complexity of human language is unparalleled among biological communication systems, and there are no completely suitable animal models. However, songbirds are considered the best model as they are unique among non-humans in their combination of vocal sophistication and experimental accessibility. Like humans, songbird vocalizations are learned. The process of song learning has close similarities to human speech acquisition (Marler, 1970; Doupe and Kuhl 1999; Konishi, 2004; Prather and Mooney, 2004; Jarvis, 2004). Recent discoveries even indicate that songbirds communicate symbolic information about the environment through their vocalizations (Templeton, 2005). Perhaps not so surprisingly, the first gene to date linked to speech disorders in humans (FoxP2) is also expressed in the neural song circuit, especially at times of vocal learning (Haesler et al. 2004; Teramitsu et al., 2004). where it can be studied much more fruitfully than in humans. **There is no other organism that rivals the songbird** for experimental study of genomic mechanisms underlying evolution and expression of vocal learning, speech communication and related cognitive processes.

The conservation or convergence of vocal learning in humans and songbirds across ~310 MY may be useful in identifying key genetic elements that underlie evolution of systems for vocal learning and communication. Moreover, nature has already done a rich experiment for us if only we can develop the genomic tools to read and interpret the results. About half of all avian species are songbirds, and this radiation developed explosively in a monophyletic lineage over the past 100 MY. All 4000+ songbirds share vocal learning circuits and skills that are NOT found in the other half of the avian world (including chickens) – and this despite enormous variation among songbirds in most other physical attributes. With the chicken as an outgroup and thousands of songbird species to contrast, the potential for comparative genomic analysis is rich, in a way that may complement comparison of chimpanzee and human genomes to understand evolution of speech and language.

c. Learning and forgetting.

In both humans and songbirds, vocal skills are developed during a critical period in juvenile life, through a process of trial and error learning (Marler, 1970, 1997). In the zebra finch, further vocal learning ceases completely by adulthood and the “bird’s own song” is tenaciously remembered and performed with minimal variation for the rest of the bird’s life. In other songbird species, such as the canary, song elements are forgotten and new ones are added in adulthood. Sequencing of the zebra finch genome will be the starting point for answering questions like these: How does the genome contribute to the onset and closure of critical learning periods? What genomic elements underlie species differences in learning, remembering and forgetting? How are these elements conserved in the human genome?

d. Increasing the power of comparative genomics.

Study of the zebra finch genome will provide new information about the events and processes that helped mold the structure and function of other vertebrates, including humans. The sequencing of the genome of a second avian species after the chicken will help confirm or reject the many hypotheses that emerged from the primary comparison of human and chicken genomes (ISGSC, Nature 2004). Zebra finch and chicken are sufficiently distant themselves (~100 MYA) to contribute significantly to a three-point comparison with humans. This

comparative analysis will reveal new information regarding which genomic sequences have been subject to positive, neutral, or negative selection. The addition of a new avian genome sequence will increase the power of multiple sequence alignments, capturing more information on which non-coding genomic regions are conserved between birds and humans and therefore deserve further scrutiny for functional roles in the human genome (Margulies et al., 2005). It will allow sharper definition of evolutionary rates and functional constraints on proteins, pinpointing fast evolving proteins of great interest for studies of reproduction, host defense, and behavior. It will show patterns of evolution and rearrangement of avian chromosomes, in particular the sex chromosomes, from which will come insight into mechanisms of sex determination in vertebrates: are they common or diverse?

e. Population and ecological biology.

Birds have been enormously informative for generating ideas about human population dynamics and environmental biology. The new sequence information will provide genetic markers and probes for determining relatedness of individuals in wild populations, greatly improving studies of population biology. Moreover, the selective pressures on bird populations will be studied much more efficiently, improving our understanding of how specific traits evolve in response to selective forces.

3. Expanding our understanding of basic biological processes relevant to human health.

Basic research on songbirds has had a profound impact on concepts of brain development and function in humans because concepts developed first from discoveries on songbirds have been found subsequently to be generally true also in mammals. In addition to the ones already described above, other seminal discoveries made in songbirds include the following:

a. Large sex differences in neural structure and function.

Large sex differences in the brain of vertebrates were first discovered from study of the song control nuclei (Nottebohm and Arnold, 1976), catalyzing subsequent discoveries of morphological sexual dimorphisms in the brain of mammals including humans (e.g., DeLacoste-Utamsing and Holloway, 1982). Sex steroid hormones cause changes in the synaptic organization of the adult neural song circuit, a phenomenon first discovered in songbirds (Nottebohm, 1981; DeVoogd and Nottebohm, 1981) and subsequently found in other circuits in mammals (e.g., Kurz et al., 1986). More recently, studies in songbirds have revealed unexpected evidence for additional non-gonadal factors in sexual differentiation of the brain (Agate et al, 2003; Holloway and Clayton, 2001). The songbird has led to a reconsideration of the forces that lead to sex differences in physiology and disease in mammals (Arnold, 2004). The 2001 National Institute of Medicine report *Exploring the Biological Contributions to Human Health: Does Sex Matter?* (http://www.iom.edu/report.asp?id=5437) argues for gender-specific approaches to medicine and cites the role of songbird studies in understanding the cellular and molecular forces that shape sex-specific development.

b. Neural mechanisms of representation and integration.

Songbirds use complex perceptual mechanisms to interpret sounds, and they match their vocal output to these sounds. They also produce complex sequences of movements that are amenable to study because of the close mapping of movements and sounds produced. These traits make songbirds attractive for computational physiologists who study how the brain processes sensory stimuli (e.g., Gentner and Margoliash, 2003), represents a sequence of fine motor actions (Hahnloser et al., 2002, Leonardo and Fee, 2005), integrates motor commands with perceptual feedback (Margoliash, 1997), and modulates its function through processes like sleep (Deregnaucourt, 2005)

c. The "Genomic Action Potential."

The sound of song triggers a transient wave of gene expression in higher auditory centers of the songbird forebrain (Mello et al., 1992). The act of singing triggers a similar genomic response in song motor control centers (Jarvis & Nottebohm, 1997). The pattern of activation differs depending on the social context in which the song is produced – directed to a female, or undirected solo rehearsal (Jarvis et al., 1998). These observations stand as some of the clearest demonstrations that natural experience actively engages the genome and modulates its expression in the brain. This pulse of gene expression has been likened to the electrophysiological action potential, as a mechanism for temporal and functional integration except occurring on a different time scale (Clayton 2000). Whole genome sequence is needed to characterize this genomic response more thoroughly, identify gene sequences that mediate it, and analyze conservation with other organisms including humans.

4. Providing additional surrogate systems.

As the previous sections should have demonstrated, the zebra finch is already a robust model system in health-related research. With a generation time close to a mouse, the zebra finch could potentially be used much more widely for tests of various therapeutic approaches, such as pharmacological or transplant manipulation of neurogenesis or other aspects of brain repair. Arguably, the single biggest barrier is the lack of zebra finch sequence data needed to assess relationships between the zebra finch and human proteomes.

5. Facilitating the ability to do experiments.

a. Tools for songbird research.

Whole genome sequence will provide critical tools for songbird research, enhancing its impact and relevance to human health and disease. We expect to identify the **coding sequences** of ~90% of the genes of the zebra finch genome, including some genes still missing from the chicken genome. The coding sequences will immediately allow the synthesis of nucleotide probes and PCR primers for diagnostic use in research in neurobiology, immunology, behavior, ecology, evolution, etc. This information will save many person-years of work to clone genes of interest one by one. We also expect to identify ~90% of non-transcribed **regulatory regions.** These sequences will allow investigators to find promoters that drive genes of interest, to study the regulation of specific genes as a function of interesting physiological states. For example, the genome sequence will immediately allow studies of methylation of selected promoters, an emerging mechanism for regulation of gene expression especially as a function of experience and developmental stage. The identification of promoters is a prerequisite for sophisticated manipulation of the genome. An integrated, mapped zebra finch genome sequence will replace costly BAC sequencing, which is currently the method of choice for determining the sequence of putative regulatory regions. In combination with the physical map of the genome and other resources now under development (above and below), the sequence will also provide **genetic markers** that will revolutionize the genetic study of songbirds. These markers will make it much easier to perform linkage studies of specific traits, studies of population genetics on wild species, and evaluation of informative mutants.

b. Enhancing and leveraging the investment in chicken genomics

The chicken was the first avian genome sequenced, recognizing its enormous importance as both a food source and a classical model for experimental biology. The zebra finch genome sequence will directly enhance the value of the existing chicken genome sequence. The two genomes should be relatively easy to align, given fewer chromosome rearrangements in birds (e.g., Itoh and Arnold, 2005) than has occurred in mammals. Comparative sequence analysis with the chicken will aid in the assembly and gene identification in both species, clarifying the location of possible sequencing gaps and defining real gene deletions/insertions. The zebra finch and chicken are at a good evolutionary distance (100 MYA) for defining conserved regulatory regions. Biological information gained in one species can be transferred to the other (genes, proteins, QTLs, genetics, expression, etc). Together, the chicken and zebra finch genomes will constitute a powerful resource for comparative studies.

c. Recruitment

Finally, the zebra finch genome sequence will attract molecular and cellular biologists to the study of birds. Although songbird biology has much to offer, especially in fields such as neurobiology and behavior, its utility to humans has been retarded by the difficulty of performing cutting-edge molecular biology. The relatively easy availability of genetic sequences will help make this biological model attractive to young investigators.

B. Strategic issues in acquiring new sequence data

1. Demand for the new sequence data.

a. Size of the research community.

Several hundred laboratories around the world study songbirds in the context of neurobiology, physiology, field biology, evolution, and ecology. Three members of the US National Academy of Sciences (Peter Marler, Fernando Nottebohm, and Masakazu “Mark” Konishi) have built their careers on the study of song behavior and the neural circuit controlling song and vocal learning. Letters of support from each are appended. Other national academy members (e.g., Gordon Orians) have studied songbird behavior and ecology.

A search of the NIH CRISP database in December 2004 indicated that, in the area of songbird research, NIH was funding 46 grants including 37 R01 grants; 13 F30, F31, F32 or K02 awards; 5 R03, R21 or

R37 awards; and 4 T32s. A conservative estimate of the yearly funding for these grants is more than $10 million. The numbers of years of support previously awarded just for these funded grants is 359 grant-years. Other support comes from the NSF and non-US granting agencies in Europe and Japan. These funded projects will be profoundly impacted by the sequencing of the zebra finch genome. Had such a wealth of genetic information been available ten years ago, there is no doubt that these funded proposals would have been enormously more productive. More importantly, they would have made a different set of novel discoveries that we can only imagine.

This grant count is a serious underestimate of the size of the research community that will be affected by sequencing of the zebra finch genome. Just in the last two years, the advent of two EST databases and the zebra finch BAC library has sparked many new collaborations and research projects among investigators who are not in the list of NIH investigators. Dr. Eric Greene's lab (NIH) sequenced numerous BAC clones as part of a comparative genomic analysis. Drs. Burt and Ellegren, members of this consortium, have changed research paths based on the availability large quantities of songbird coding sequence. Whole genome sequencing will have a much larger impact.

Zebra finches have emerged as one of the top few avian species used as models for biomedical research. The table below shows the annual numbers of papers published on several comparison species. The amount of research on songbirds has increased dramatically in recent years and now surpasses any other avian taxon except for chickens. Songbirds offer an entirely new set of biological phenomena that are not amenable to study in galliforms such as quail and chickens.

PubMed Search Terms	**2002**	**2003**	**2004**
zebra finch OR taeniopygia OR birdsong OR songbird	285	339	420
chicken OR chick OR Gallus	3409	3622	3575
Meleagris [turkey genus]	16	18	8
quail OR coturnix	212	245	218
pangolin	7	2	9
hyrax	3	6	11
sloth	8	6	8

b. Community's enthusiasm for having the sequence:

The initiative to sequence the zebra finch genome had its roots in informal meetings in the Fall of 2001 at the SFN meeting, beginning with Arnold, Clayton, Jarvis, and Mello and expanding outward from there. The immediate outcome was a successful BAC construction white paper (2002), the Songbird Neurogenomics Initiative RO1 for ESTs and microarrays (a community collaboration, funded beginning 2003, centered at the University of Illinois), the songbird transcriptome project at Duke University, and a successful white paper to construct a physical map of the genome (2005). With the latter we began an ongoing collaboration with Wes Warren and the Washington University Genome Sequencing Center, and our range began to widen further to solicit input and collaboration with chicken genome researchers. These activities were routinely discussed openly over email listservs and the proposals have been designed to make the resources available quickly to other researchers. Dr. Arnold attended the CSHL Chicken genome meeting and made further efforts to discuss and coordinate our efforts with chicken genome researchers. At the Society for Neuroscience annual meeting in 2004, we held an open satellite event on the topic of "songbird neurogenomics," which was attended by ~50 songbird researchers, and an associated minisymposium attended by over 400 researchers, which reinforced the sense of broad community enthusiasm and support for the eventual goal of a whole genome sequence. To the best of our knowledge, the Zebra Finch Genome Consortium (the submitters of the present white paper) represent the broad consensus of the entire songbird research community.

Interest in a zebra finch genome sequence is not limited to songbird biologists and has been expressed by scientists of many stripes, as indicated in letters attached from three Nobel laureates (Tonegawa, Wiesel, Kandel) and five comparative biologists (Edwards, Ellegren, Zhang, Burt, Schmidt).

2. Suitability of the Zebra Finch for experimentation.

By far, the most thoroughly studied single songbird species is the zebra finch (*Taeniopygia guttata*, family Estrildidae, Suborder Oscines, Order Passeriformes). Zebra finches are small, have short generation time (4 months) for a complex vertebrate, and breed easily in captivity. Each zebra finch male sings a unique learned song as part of the courtship ritual and to maintain a monogamous bond with his mate. To develop a normal song, the young male must hear both an adult tutor (typically his father) and his own vocal performance, during a critical period in adolescence. Once the song is learned, it is sung stably throughout adult life, and new learning ceases. The existence of many studies of zebra finches, and the existence of other genomic and genetic resources for this species make the zebra finch a natural choice among songbirds for genome sequencing.

a. Genetic Strains and Pedigrees.

Morphs have been defined based mostly on color (e.g., *http://zebrafinch.info/colours/gentech.asp*), but there are no fully inbred strains available yet. Informative mutants have been described, including a half-male half-female lateral gynandromorph, which allowed novel conclusions about sexual differentiation of the brain (Agate et al., 2003). The Arnold lab is currently studying other informative mutants such as other gynandromorphs and hermaphrodites. These individual birds offer novel insights into the role of the sex chromosomes in tissue development. Genetic analysis of the mutants is rapidly becoming more feasible because of the introduction of techniques for producing metaphase chromosomes from adult tissues (Itoh and Arnold, 2005), and the BAC and genomic resources described in the next few paragraphs. A whole genome sequence would revolutionize this type of analysis with new genetic probes, PCR primers, and a detailed genome map.

Since the 1980s Professor T.R. Birkhead at the University of Sheffield (U.K.) has developed a comprehensive 18-generation genealogy (pedigree) of zebra finches with blood/tissue samples for about 1,500 birds from the most recent generations. The goal is to perform detailed analysis of the quantitative genetics of reproductive traits, in particular to explain the genetic basis for the considerable inter-male variation in sperm phenotypes, and the maternal effects on sperm morphology and function. These relate to his established research on post-copulatory sexual selection, sperm competition and cryptic female choice (Forstmeier et at., 2004; Pizzari et al 2003; Birkhead & Pizzari 2003). For the 1,500 birds with DNA samples, data exist for other traits including beak color, body mass, body size and immune function. Dr. Birkhead also has three lines selected for sperm traits. He plans to make a zebra finch genome linkage map by analyzing SNPs, which will dovetail and reinforce with the current NHGRI-funded BAC-based effort to construct a physical map of the zebra finch genome (see next paragraph).

b. BAC library and genome maps.

Through the NHGRI White Paper mechanism, a zebra finch BAC library (TG_Ba) was constructed by the Arizona Genome Institute (*http://www.genome.arizona.edu/* ; Luo et al., submitted). The library has an average insert size of 134kb (genome size: 1200 Mb) covering ~16 genome equivalents. It contains 147,456 clones. The NHGRI also announced funding in June 2005 for a proposal to construct a physical map of the zebra finch genome, and minimum BAC tiling path of the genome, based on BAC end sequencing, BAC fingerprinting, overgo hybridization, and some BAC sequencing. That effort is being conducted by Dr. Wes Warren at Washington University Genome Sequencing Center. At the same time, Dr. Hans Ellegren of Uppsala University (Sweden) has been developing a linkage map of the genome of the collared flycatcher (Ficedula albicollis) in Sweden. The map is based on analysis of SNPs spaced approximately every 10MB throughout the genome, analyzed in a population of flycatchers of known genetic relationships. Drs. Ellegren and Warren will coordinate their linkage and BAC-based mapping strategies for maximum benefit to the research community. A linkage map of the zebra finch will also likely be constructed from Dr. Birkhead's DNA resources (see last paragraph).

c. EST databases and microarrays.

Two large-scale, collaborative transcriptome/EST efforts are underway in the zebra finch research community, both with NIH support. Both efforts are accessible via a common web URL: *http://www.songbirdgenome.org*

One project, at the Duke University Medical Center, has produced 21 full-length, normalized and subtracted cDNA libraries enriched in sex-specific, developmental, and behaviorally regulated genes (Jarvis et al., 2002). From these libraries, 18,000 cDNA clones have been 5'-and 3' end-sequenced and spotted on a microarray. These cluster into ~9000 unique transcripts, which have been manually curated with detailed annotation at *http://songbirdtranscriptome.net/.* Many of the Duke clones are essentially full-length (based on the

protocol of Carninci 2003) and should be especially useful for analysis of splicing variants and for gene over-expression.

The other project, at the University of Illinois (Keck Center for Comparative and Functional Genomics), has so far produced ~34,000 5'-end-sequenced ESTs from a normalized brain cDNA library representing both sexes and a mixture of ages. These cluster into ~18,000 unique sequences, which have been spotted onto a microarray. The microarray is being distributed to the songbird research community via an organized proposal-and-review process (*http://titan.biotec.uiuc.edu/songbird/*). The Illinois database is searchable via an online software interface developed at the Keck Center, called ESTIMA (EST Information Management and Annotation tool). Via ESTIMA, one can retrieve EST sequence files and annotations by sequence ID, direct BLAST search, Gene Ontology terms, or keywords from description fields imported from the external databases during the annotation process.

The Illinois group has also performed a master assembly of the all ESTs combined from both groups (Duke + Illinois). In combination, ~60,000 ESTs cluster into ~23,000 unique sequences which appear to represent ~16,000 unique genes (accounting for non-contiguous ESTs from common transcripts). The Duke and Illinois EST collections are highly complementary, even for transcripts commonly represented in both, because their different cloning and sequencing strategies resulted in many staggered overlaps that facilitate contig assembly. This combined assembly is now publicly available (*http://titan.biotec.uiuc.edu/songbird/*) and has been machine-annotated by BLAST sequence similarity searches against four external databases: TIGR Gallus gallus (chicken) EST, NCBI chicken unigene, Swissprot, NR.aa. Approximately 76% of these zebra finch ESTs have highly significant hits against the chicken EST collection, and ~72% align to the full chicken genome database (ENSEMBL) by BLASTN alignment. At this writing, work is underway to create a custom Track on the UCSC Genome Browser, using the chicken genome as the alignment reference. About half of the unique transcripts in the combined assembly are represented by only a single EST in the current collection, indicating a high probability that substantial numbers of brain-expressed sequences remain yet unidentified. Both Duke and Illinois are continuing to produce new ESTs and are working together to develop a single master oligonucleotide array that will represent the evolving combined assembly.

An initial comparative genomic hybridization experiment underscores the potential for using zebra finch sequence as a basis for comparisons across songbirds (SoNG Initiative, ms in prep). Using genomic DNA from zebra finch, canary, starling and two different sparrows to hybridize to the Illinois cDNA microarray, a main effect of species is observed for less than a third of the probes on the array (FDR-corrected p-value <.05), and less than 2% of the probes give substantially (>2x) lower signals in any non-zebra finch species. A few probes on the array show evidence of various "species-specific" hybridization patterns, such as the ortholog for myosin-6 which gives a much higher signal in white-crowned sparrows than in any of the other species; interestingly, chicken myosin-6 is annotated as associated with sound perception.

These EST sequencing and microarray production efforts have stimulated numerous labs to incorporate molecular analysis into their arsenal of tools to study brain and behavior. The EST databases have also been mined by avian groups such as Ellegren's and Burt's, for studies of chicken reproduction and avian molecular evolution. This process will be accelerated considerably if the genes and their regulatory sequences become available from whole genome sequencing.

d. Gene transfer technologies.

Although transgenesis and germline knock-outs have not yet been broadly successful in birds, several labs are working on this problem (Sang, 2004). In the meantime in vivo gene manipulation has been carried out successfully in chick embryos using RNAi techniques (Krull, 2004; Bron et al., 2004). These techniques should be easily adapted to songbirds. The most immediate route for genetic manipulation in the zebra finch may be via use of genetic expression vectors and/or RNAi targeted to specific song control nuclei. These techniques can be applied in adult birds to influence function acutely. Because much of brain development occurs after hatch, gene manipulation is also feasible at many stages of brain development, including during learning. Preliminary studies in several labs have demonstrated the feasibility of several such approaches in the zebra finch, including injection of naked antisense DNA, lentiviral vectors, adenoviral vectors, novel nanoparticle carriers, and RNAi. This strategy builds on a long history of targeted pharmacological manipulations in the song system (e.g., steroids), but will ultimately benefit from, if not require, much more extensive genomic information (e.g., complete sequence from genes of interest, including regulatory regions from gene promoters).

e. Cytological maps of the zebra finch genome.

Dr. Darren Griffin of the University of Kent (UK) has submitted an application to the UK Biotechnology and Biological Science Research Council in January 2005, part of which proposes to map BAC clones to zebra finch metaphase chromosomes using FISH. We intend to coordinate our efforts with Dr. Griffin. Other FISH mapping of BACs is occurring on a smaller scale in other labs.

3. Rationale for the complete sequence of the Zebra Finch.

A good assembly of the whole genome sequence is essential as many of the uses outlined above will be focused on identification of gene regulatory regions and comparative analyses of genome structure. Non-coding sequences will also allow direct study of epigenetic influences on brain gene expression, such as those mediated by methylation of promoter sequences. Recent evidence shows that processes such as methylation can themselves be regulated by an animals' early experience. This is an important issue, because such influences on the expressed genome can be fundamental to understanding modulations of brain function such as critical periods for learning and hormone effects. The genome sequence will also be the initial basis for SNP identification, which is needed to catalyze further development of studies relating genetic variation to phenotype. There are no alternatives to whole genome sequencing to get all this information.

4. Cost of sequencing the genome, state of readiness of organism's DNA for sequencing.

a. Size of the genome, repeat structure.

The zebra finch genome is similar to the chicken in size and a low density of repeats. Preliminary studies suggest high conservation of the physical map of the genome in birds. When chicken single chromosome paints are used in other avian species including the zebra finch (Itoh and Arnold, 2005), they typically hybridize to one (rarely two) chromosome(s), indicating that despite millions of years of separation, only minor chromosomal rearrangement has occurred between the chicken and zebra finch lineages. Moreover, when individual BAC clones from the zebra finch are sequenced, they align very closely with homologous regions of the chicken genome, both within and between exons (Luo et al., submitted).

b. Quality of sequence product needed.

We propose a whole genome shotgun, with the coverage strategy outlined below. The physical map and BAC end sequences will allow us to scaffold these shotgun read contigs together.

c. Sequencing strategy.

Sequencing strategies continue to evolve toward more efficient, less costly methods. For the zebra finch genome we propose to generate 6X whole genome shotgun sequence coverage using 3730xl instruments of which a yet to be determined coverage of reads comes from the pyrosequencing method developed by 454 Life Sciences. Experience and insights gained recently from our 454 sequencing instrument indicate that this platform could contribute to the sequence of the zebra finch genome. Furthermore, since the genome architecture of the zebra finch should be relatively similar to chicken, we are confident that this modified assembly approach will produce a high quality end product. The WUGSC will assemble these combined reads with the PCAP assembly software following the production phase. The preliminary assembly will be released shortly after internal quality review. The main impetus for proposing this mixed read source approach is that its cost savings is significant; by our estimates a potential savings of ~$3M would be realized if we could successfully substitute a 2X coverage of 454-generated contigs into the assembly of the zebra finch genome in place of 2X traditional 3730xl coverage.

Prior to the availability of the zebra finch whole genome sequence assembly a comprehensive BAC physical map with BAC end sequence links will be available for sequence assembly validation. This is a significant advantage to gauge whole genome assembly quality at early stages of genome coverage and an early resource not available for previous whole genome sequencing projects. A fingerprint map will be developed at Washington University at a targeted 12x whole genome coverage utilizing the agarose fingerprinting method.

d. DNA source considerations.

In order to generate diploid coverage of the Z sex chromosome, it would be necessary to sequence a ZZ homogametic (male) zebra finch genome. However, to be consistent with the representation of the Z and W chromosomes in the chicken genome assembly and the zebra finch BAC library source (female), we suggest

the DNA sequencing source be a female. We recognize the under-representation of the Z and W chromosome but feel this is a better link to existing efforts in multiple labs to further characterize the Z and W in zebra finch. We propose to utilize the existing DNA for the currently available BAC library (www.genome.arizona.edu) or a domesticated zebra finch source within Dr. Art Arnold's laboratory, the same population of zebra finches that were sampled to develop the currently available female BAC library. WUGSC will assay the heterozygosity rate of these samples by PCR-based resequencing of 100 random loci. Our analysis will indicate if the proposed sample is adequate for whole genome sequencing. Based on previous experience with several organisms we expect heterozygosity rates of better than 1 SNP in every 800 bp to be a threshold that should not adversely affect WGS sequence assembly. Furthermore, zebra finch repeat content is expected to be similar to chicken at ~10% and should make sequence assembly straightforward.

e. EST sequencing.

Since gene structure and organization are critical in evidence-based gene prediction algorithms we recognize the ongoing funded effort to generate additional zebra finch EST sequences to be a valuable complement to the proposed whole genome sequencing. The current initiatives led by Drs. Clayton and Jarvis have already accomplished much with over 60,000 ESTs produced from brain cDNA libraries. Since these initiatives have focused on brain, we feel it is important to sample other tissues for a more complete gene set. We propose to sequence 10,000 ESTs from each of six tissue-specific cDNA libraries derived from non-brain tissues.

5. Are there other (partial) sources of funding available or being sought for this sequencing project?

There are no other sources available or being sought for this sequencing project.

Zebra Finch Genome Consortium (see also http://songbirdgenome.org)

David Clayton, Ph.D., Professor
Cell & Structural Biology, Neuroscience
Member, Institute for Genomic Biology
B107 Chemical & Life Sciences Laboratory
University of Illinois
Urbana IL 61801, phone 217-244-3668, fax 503-213-6611 dclayton@uiuc.edu
http://www.life.uiuc.edu/clayton/

Arthur P. Arnold, Ph.D., Distinguished Professor and Chair
Department of Physiological Science
University of California, Los Angeles
621 Charles E. Young Drive South, Room 4117
Los Angeles CA 90095-1606, phone 310-825-2169, fax 310-825-8081, arnold@ucla.edu
http://www.physci.ucla.edu/html/arnold.htm

Wes Warren, Ph.D., Assistant Director
Washington Univ. Genome Sequencing Center
Washington University School of Medicine
4444 Forest Park Blvd
St. Louis MO
wwarren@watson.wustl.edu
Phone: (314) 286-1899

Erich Jarvis, Ph.D.
Assistant Professor
Department of Neurobiology,
Duke University Medical Center

Claudio Mello, M.D., Ph.D.
Assistant Professor
Neurological Sciences Institute
Oregon Health Sciences University

David W. Burt, Professor
Dept. of Genetics & Genomics
Roslin Institute (Edinburgh)
Roslin, Midlothian EH25 9PS, UK
tel: +44-131-527-4218
fax: +44-131-440-0434
e-mail: Dave.Burt@bbsrc.ac.uk
http://www.chicken-genome.org

Hans Ellegren, Professor
Evolutionary Biology Centre
Dept of Evolutionary Biology
Uppsala University
Norbyvägen 18D
SE-752 36 Uppsala
Phone: +46-18-4716460
Fax: +46-18-4716310
Email: Hans.Ellegren@ebc.uu.se
Web: http://www.ebc.uu.se/

REFERENCES

Arnold AP 2004 Sex chromosomes and brain gender. Nat Rev Neurosci. 5:701-8.

Arnold KE, Griffiths R, Stevens DJ, Orr KJ, Adam A, Houston DC. 2003 Subtle manipulation of egg sex ratio in birds. Proc R Soc Lond B Biol Sci. 270 Suppl 2:S216-9.

Birkhead TR, Pizzari T. 2002 Postcopulatory sexual selection. Nat Rev Genet. 3:262-73.

Birkhead TR. 1998 Sperm competition in birds. Rev Reprod. 3:123-9.

Brainard MS, Doupe AJ. 2002 What songbirds teach us about learning. Nature. 417:351-8.

Bron R, Eickholt BJ, Vermeren M, Fragale N, Cohen J. 2004 Functional knockdown of neuropilin-1 in the developing chick nervous system by siRNA hairpins phenocopies genetic ablation in the mouse. Dev Dyn. 230:299-308.

Carninci, P. 2003. Generation of full-length libraries. Bowtell, D., Sambrook, J. (eds). DNA Microarrays: A Molecular Cloning Manual. Cold Spring Harbor. NY. 647-670.

Carson R. 1962. Silent Spring. Boston: Houghton Mifflin

Clayton DF, George JM 1998 The synucleins: a family of proteins involved in synaptic function, plasticity, neurodegeneration and disease. Trends Neurosci 21, 249-254.

Clayton DF. 2000 The Genomic Action Potential. Neurobiol Learn Mem 74:185-216.

DeLacoste-Utamsing C, Holloway RL 1982 Sexual dimorphism in the human corpus callosum. Science 216: 1431-1432.

Denisenko-Nehrbass NI, Jarvis Ed, Scharff C, Nottebohm F, Mello CV. 2000 Neuron 27, 359-370.

Deregnaucourt S, Mitra PP, Feher O, Pytte C, Tchernichovski O. 2005 How sleep affects the developmental learning of bird song. Nature 433:710-716.

DeVoogd TJ, Nottebohm F 1981 Gonadal hormones induce dendritic growth in the adult avian brain. Science 214: 202-204.

Doupe A, Kuhl P 1999 Birdsong and human speech: common themes and mechanisms. Annu Rev Neurosci. 22:567-631.

Edwards SV, Kingan SB, Calkins JD, Balakrishnan CN, Jennings WB, Swanson WJ, Sorenson MD. 2005 Speciation in birds: genes, geography, and sexual selection. Proc Natl Acad Sci USA. 102 Suppl 1:6550-7.

Edwards SV, Hess CM, Gasper J, Garrigan D. 1999 Toward an evolutionary genomics of the avian Mhc. Immunol Rev. 167:119-32.

Forstmeier W, Coltman DW, Birkhead TR. 2004 Maternal effects influence the sexual behavior of sons and daughters in the zebra finch. Evolution Int J Org Evolution. 58:2574-83.

Gardiner JS, Schroeder S, Polacco ML, Sanchez-Villeda H, Fang Z, Morgante M, Landewe T, Fengler K, Useche F., Hanafey M, Tingey S, Chou H, Wing R, Soderlund S and Coe EH, Jr. 2004. Anchoring 9,371 maize expressed sequence tagged unigenes to the bacterial artificial chromosome contig map by two-dimensional overgo hybridization. Plant Physiol. 134:1317-1326.

Gee JM, Craig-Veit CB, Millam JR. 2004 Posthatch methoxychlor exposure adversely affects reproduction of adult zebra finches, Taeniopygia guttata. Bull Environ Contam Toxicol. 4:607-12.

Gentner TQ, Margoliash D. 2003 Neuronal populations and single cells representing learned auditory objects. Nature 424:669-674.

Goldman SA, Nottebohm F. 1983 Neuronal production, migration, and differentiation in a vocal control nucleus of the adult female canary brain. PNAS 80:2390-4.

Grant PR, Grant BR, Markert JA, Keller LF, Petren K. 2004 Convergent evolution of Darwin's finches caused by introgressive hybridization and selection. Evolution Int J Org Evolution. 58:1588-99

Hahnloser RH, Kozhevnikov AA, Fee MS. 2002 An ultra-sparse code underlies the generation of neural sequences in a songbird. Nature. 419:65-70

Hambly C, Harper EJ, Speakman JR. 2004 The energetic cost of variations in wing span and wing asymmetry in the zebra finch Taeniopygia guttata. J Exp Biol.207:3977-84.

Hollenbeck JE. 2005 An Avian Connection as a Catalyst to the 1918-1919 Influenza Pandemic.

Holloway CC, Clayton DF 2001 Estrogen synthesis in the male brain triggers development of the avian song control pathway in vitro. Nature Neuroscience 4: 1-7.

Int J Med Sci. 2(2):87-90.

International Chicken Genome Sequencing Consortium. 2004. Sequence and comparative analysis of the chicken genome provide unique perspectives on vertebrate evolution. Nature 432:695-716.

Itoh Y, Arnold AP. 2005 Chromosomal polymorphism and comparative painting analysis in the zebra finch. Chromosome Research 13:1-10.

Jarvis, ED, Nottebohm F. 1997. Motor-driven gene expression. Proc. Natl. Acad. Sci. 94:4097-4102.

Jarvis ED, Scharff C, Grossman MR, Ramos JA, Nottebohm F. 1998. For whom the bird sings - context-dependent gene expression. Neuron 21:775-788

Jarvis ED, Gunturkun O, Bruce L, Csillag A, Karten H, Kuenzel W, Medina L, Paxinos G, Perkel DJ, Shimizu T, Striedter G, Wild JM, Ball GF, Dugas-Ford J, Durand SE, Hough GE, Husband S, Kubikova L, Lee DW, Mello CV, Powers A, Siang C, Smulders TV, Wada K, White SA, Yamamoto K, Yu J, Reiner A, Butler AB; Avian Brain Nomenclature Consortium. Avian brains and a new understanding of vertebrate brain evolution. Nat Rev Neurosci. 2005 Feb;6(2):151-9.

Jarvis ED, Smith VA, Wada K, Rivas MV, McElroy M, Smulders TV, Carninci P, Hayashizaki Y, Dietrich F, Wu X, McConnell P, Yu J, Wang PP, Hartemink AJ, Lin S. 2002. A framework for integrating the songbird brain. J Comp Physiol A. 188:961-980.

Jarvis ED. 2004 Learned birdsong and the neurobiology of human language. Ann N Y Acad Sci. 1016:749-77.

Johnson, F, Whitney, O. 2005. Singing-driven gene expression in the developing songbird brain. Physiol Behav (in press).

Kellner, W.A., R.T. Sullivan, B.H. Carlson and J.W. Thomas. 2004 Uprobe: A genome-wide universal probe resource for comparative physical mapping in vertebrates. Genome Res. Dec 8; [Epub ahead of print]

Kida H, Ito T, Yasuda J, Shimizu Y, Itakura C, Shortridge KF, Kawaoka Y, Webster RG. Potential for transmission of avian influenza viruses to pigs. J Gen Virol. 1994 75:2183-2188.

Konishi M. 2004 The role of auditory feedback in birdsong. Ann N Y Acad Sci. 1016:463-75.

Krull CE 2004 A primer on using in ovo electroporation to analyze gene function. Dev Dyn. 229:433-9

Kruse AA, Stripling R, Clayton DF. 2004 Context-specific habituation of the zenk gene response to song in adult zebra finches. Neurobiol Learn Mem. 82:99-108

Kurz EM, Sengelaub DR, Arnold AP 1986 Androgens regulate dendritic length of sexually dimorphic mammalian motoneurons in adulthood. Science 232: 395-398.

Leonardo A, Fee MS. 2005 Ensemble coding of vocal control in birdsong. J Neurosci 25:652-661

Luo M., Yu Y, Kim H, Kudrna D, Itoh Y, Agate RJ, Goicoechea JL, Talag J, Mueller C, Wang W, Currie J, Sisneros NB, Wing RA, Arnold AP. 2005 Construction and characterization of a BAC library of the zebra finch genome, and its utilization to determine the structure of an avian androgen receptor genomic region, submitted.

Mackenzie JS, Gubler DJ, Petersen LR. 2004 Emerging flaviviruses: the spread and resurgence of Japanese encephalitis, West Nile and dengue viruses. Nat Med. 10(12 Suppl):S98-109.

Margoliash D. 1997 Functional organization of forebrain pathways for song production and perception. J Neurobiol. 33:671-93.

Margulies EH, Vinson JP, Miller W, Jaffe DB, Lindblad-Toh K, Chang JL, Green ED, Lander ES, Mullikin JC, Clamp M; NISC Comparative Sequencing Program. 2005. An initial strategy for the systematic identification of functional elements in the human genome by low-redundancy comparative sequencing. Proc Natl Acad Sci U S A. 102(13):4795-800.

Marler P. 1970 Birdsong and speech development: could there be parallels? Am Sci. 58:669-73

Marler P. 1997 Three models of song learning - evidence from behavior. J Neurobiol 33:501-516..

Mello CV, Nottebohm F, Clayton D. 1995 Repeated exposure to one song leads to a rapid and persistent decline in an immediate early gene's response to that song in zebra finch telencephalon. J Neurosci. 15:6919-25.

Mello CV, Vicario D, Clayton D 1992 Song presentation induces gene expression in the songbird forebrain. PNAS 89: 6818-6822.

Mello CV. 2002 Mapping vocal communication pathways in birds with inducible gene expression. J Comp Physiol A Neuroethol Sens Neural Behav Physiol. 188:943-59.

Nottebohm F 1981 A brain for all seasons: cyclic anatomical changes in song control nuclei of the canary brain. Science 214: 1368-1370.

Nottebohm F 1985 Hope for a New Neurology. New York: New York Academy of Science.

Nottebohm F, Arnold AP 1976 Sexual dimorphism in vocal control areas of the song bird brain. Science 194: 211-213.

Nottebohm F, Stokes TM, Leonard CM 1976 Central control of song in the canary (Serinus canarius). J Comp Neurol 165: 457-486.

Nottebohm F. 2004 The road we travelled: discovery, choreography, and significance of brain replaceable neurons. Ann N Y Acad Sci. 1016:628-58.

Pizzari T, Cornwallis CK, Lovlie H, Jakobsson S, Birkhead TR. 2003 Sophisticated sperm allocation in male fowl. Nature 426:70-4.

Prather JF, Mooney R. 2004 Neural correlates of learned song in the avian forebrain: simultaneous representation of self and others. Curr Opin Neurobiol. 14:496-502.

Reed KD, Meece JK, Henkel JS, Shukla SK. 2003 Birds, Migration and Emerging Zoonoses: West Nile Virus, Lyme Disease, Influenza A and Enteropathogens. Clin Med Res. 2003 1:5-12.

Romanov M.N., J.A. Price, and J.B. Dodgson. 2003. Integration of animal linkage and BAC contig maps using overgo hybridization. Cytogenetics and Genome Research 102:277-281.

Ross MT, LaBrie S, McPherson J, Stanton VP: Screening large-insert libraries by hybridization, in Dracopoli NC, Haines JL, Korf BR, Moir DT, Morton CC, Seidman CE, Seidman JG, Smith DR (eds): Current Protocols in Human Genetics, pp 5.6.1-5.6.52 (John Wiley and Sons, New York, 1999).

Saetre GP, Borge T, Lindell J, Moum T, Primmer CR, Sheldon BC, Haavie J, Johnsen A, Ellegren H. 2001 Speciation, introgressive hybridization and nonlinear rate of molecular evolution in flycatchers. Mol Ecol. 0:737-49.

Sang H 2004 Prospects for transgenesis in the chick. Mech Devel. 121:1179-1186.

Schlinger BA, Soma KK, London SE. 2001 Neurosteroids and brain sexual differentiation. Trends Neurosci. 24:429-31.

Sibley CG, Ahlquist JE. 1990. Phylogeny and Classification of Birds: A Study in Molecular Evolution. Yale University Press, New Haven.

Snoeijs T, Dauwe T, Pinxten R, Darras VM, Arckens L, Eens M. 2005 The combined effect of lead exposure and high or low dietary calcium on health and immunocompetence in the zebra finch. EnvironPollut. 134:123-32.

Templeton CN, Greene E, Davis K. 2005 Allometry of alarm calls: black-capped chickadees encode information about predator size. Science 308:1934-1937.

Thomas, JW, Prasad AB, Summers TJ, Lee-Lin S-Q, Maduro VVB, Idol JR, Ryan JF, Thomas PJ, McDowell JC and Green ED 2002 Parallel construction of orthologous sequence-ready clone contig maps in multiple species. Genome Res. 12:1277-1285.

Wade J, Peabody C, Coussens P, Tempelman RJ, Clayton D, Liu L, Arnold AP, Agate R. 2004 A cDNA microarray from the telencephalon of juvenile male and female zebra finches. J Neurosci Methods 138:199-206.

Wallis JW, Aerts J, Groenen MA, Crooijmans RP, Layman D, Graves TA, Scheer DE, Kremitzki C, Fedele MJ, Mudd NK, Cardenas M, Higginbotham J, Carter J, McGrane R, Gaige T, Mead K, Walker J, Albracht D, Davito J, Yang SP, Leong S, Chinwalla A, Sekhon M, Wylie K, Dodgson J, Romanov MN, Cheng H, de Jong PJ, Osoegawa K, Nefedov M, Zhang H, McPherson JD, Krzywinski M, Schein J, Hillier L, Mardis ER, Wilson RK, and Warren WC. 2004. A physical map of the chicken genome. Nature. 432:761-764.

Wingfield JC, Jacobs J, Hillgarth N. 1997 Ecological constraints and the evolution of hormone-behavior interrelationships. Ann N Y Acad Sci. 807:22-41

Wingfield JC, Sapolsky RM 2003 Reproduction and resistance to stress: when and howY J Neuroendocrinol. 15:711-24.

Index of Supporting Letters (following)

Page
##

June 22, 2005

Dear Art,

I write with strong support for the application to sequence the zebra finch genome. This single project will likely catalyze rapid progress on numerous fronts related to understanding human health and disease.

The brain of songbirds has become fertile material for the study of basic processes of sexual differentiation, ageing, learning and brain repair. It was in songbirds that the phenomenon of spontaneous neuronal replacement in adult brain was discovered, changing 100 years of dogma about what adult brains could not do and songbird studies provided, too, the first identification of neuronal stem cells in adult brain, the manner of migration of young neurons and the behavioral conditions and trophic factors that promoted new neuron survival. Importantly, all these insights generalized very well to mammals, in which adult neurogenesis has now become a major field of study. Birds, however, retain an advantage, because the vocal control network in which neuronal replacement was first studied is a gold mine in many other ways. These pathways are sexually dimorphic, have sensitive periods for learning and a narrowly defined anatomy and function. In this single system comparisons of the expression profile of replaceable and non-replaceable neurons now promis to yield basic insights on the mechanisms for long-term memory storage, neuronal ageing and the basic biology of neurodegenerative disorders. Yet this material has not been able to realize its full potential because of insufficient molecular tools. One of the requirements for unlocking the secrets of cells and cell circuits is to be able to look up the nucleotide sequences of any gene of interest and its genomic regulatory regions. This information, which will be available once the zebra finch genome is fully sequenced, will greatly accelerate our ability to develop markers for specific neuronal populations or neuronal functions and measure and manipulate gene expression. With these tools in hand, further work with the songbird brain is bound to yield new conceptual insights on sexual differentiation of brain and behavior and on ageing and tools for brain rejuvenation, brain repair following lesion and neuronal replacement following neurodegenerative disorders in humans. Songbirds are just superb material for many issues of basic medical interest and their full use has been held back by a lack of the full breadth of molecular tools enjoyed by those working in mice. It is worth noting, for those not familiar with zebra finches, that these little birds breed in the laboratory all year and reach sexual maturity at 90 days. Unlike mice, though, they can live up to 8 years, thus having the advantages of quick generation turnover as well as being suitable for the study of brain and memory changes over a number of years.

For all of these reasons it is urgent that there be a fully sequenced and annotated genome of a songbird.

Cordially, Fernando

Dr. Fernando Nottebohm
Professor, Director
Rockefeller University Field Research Center
495 Tyrrel Road
Millbrook, New York 12545

Susumu Tonegawa
Picower Professor of Biology and Neuroscience
Director, Picower Center for Learning and Memory
Massachusetts Institute of Technology

June 23, 2005

Endorsement for Dr. Clayton and his associate's research proposal on genomics of songbirds.

I write to endorse the proposal to sequence the genome of the zebra finch, a songbird. Songbirds have occupied a special place in neurobiology, because males learn their song from others during critical periods of development. The study of songbird learning has led to some exciting and novel findings, including the ability of the adult brain to make new neurons and insert them into functional pathways. However, one effective approach for identifying underlying mechanisms for learning and memory in the songbird system that has been missing in the songbird research is genetic manipulation. My own research on learning has benefited from this genetic approach. This proposal is a significant attempt to introduce this powerful approach to the songbird system. Study of diverse model systems and species is required for a full understanding of how the human brain acquires and stores information, and retrieves it. I believe that comparative genetic and functional studies of this group of animals will have a long-term positive impact on our understanding of human biology.

Sincerely,
Susumu Tonegawa

21 June 2005

Dear Art,

I write to offer my wholehearted support to the application to fund whole genome sequencing of the zebra finch. As the second bird species to have its genome sequenced, the access to the zebra finch genome would reveal the strengths and power of comparative genomics. Specifically, the possibility to align and compare the zebra finch and chicken genomes will allow the identification of genetic characteristics underlying traits common to different birds but also to uncover the genetic background to traits that differ among birds. Moreover, from learning about the degree of sequence conservation in non-coding DNA of different avian lineages, we shall be able to learn if conserved segments are shared between birds and mammals. The zebra finch genome sequence would be of great benefit to many avian biologists, but also to those who use the zebra finch as a model for neurobiological and cognitive studies.

My primary interest in the zebra finch genome would be to perform molecular evolutionary analysis of gene sequence data. Specifically, by aligning chicken and zebra finch orthologs it will be possible to search for genes showing evidence of adaptive evolution, i.e. being subject to positive selection. Such genes are candidates for traits that have been under strong directional selection during bird evolution, and that may be associated with phenotypic evolution. From an evolutionary biology perspective, this will also allow addressing the link between sexual selection and molecular evolution.

The possibility to align orthologous sequences from zebra finch and chicken will also allow detailed studies of mutation rate evolution in birds. Avian genomes are characterised by the presence of macro- and microchromosomes and there is indication that microchromosomal genes mutate more frequently but are subject to higher selective constraint, than macrochromosomal genes, (Axelsson 2005 Genome Research 15:120-125). We would like to study the causes of this peculiar form of genome evolution, and the zebra finch genome would be critical for such studies. The same applies to studies of sex chromosome evolution in birds.

Another area of interest to our group is the issue of how domestication has affected the sequence evolution in chicken. Artificial selection imposed by humans during domestication has resulted in some of the most dramatic and rapid phenotypic changes in nature. The comparison of chicken and zebra finch orthologs will help in identifying those genes that have been under selection during domestication, and may help to uncover the genetics of production traits and disease resistance. The latter is particularly important in light of the risk for spread of infectious disease by birds.

It should finally be mentioned that as a representative of the bird order Passeriformes (passerine birds, songbirds), the access to the zebra finch genome would be of outmost importance to ecologists and evolutionary biologists studying aspects such as life history evolution, mating systems, sexual selection and population dynamics. Species from Passeriformes are the most common model species in evolutionary ecological studies and the general trend of integration genetic analysis in ecology and evolutionary biology will be facilitated by sequence information from a songbird.

Yours sincerely,

Hans Ellegren
Professor and Chair in Evolutionary Biology
Dept of Evolutionary Biology
Uppsala University
Norbyvägen 18D
SE-752 36 Uppsala

HARVARD UNIVERSITY
DEPARTMENT OF ORGANISMIC AND
EVOLUTIONARY BIOLOGY
26 OXFORD STREET
Cambridge, Massachusetts 02138 USA

20 June 2005

Arthur P. Arnold, Ph.D.
Professor and Chair
Department of Physiological Science
UCLA
621 Charles E. Young Drive South, Room 4117
Los Angeles CA 90095-1606

Dear Art:

It's a pleasure to write in support of a proposal to conduct whole-genome sequencing on the Zebra Finch. As a molecular evolutionist with interests in comparative genomics and nonmodel species, I am, frankly, even more excited by a zebra finch genome than by the chicken genome. That said, having two avian genomes would be an incredible boon to my research and to the research in the community of evolutionary biologists. As you know, we have already begun to study the population genetics of zebra finches and have been focusing on population variation in genes of the major histocompatibility complex (Mhc), as well as population variation in a number of ESTs expressed in the brain and curated by Dr. Erich Jarvis. A full genome sequence would immediately allow us to bring such studies to a much higher level of sophistication and rigor. In particular we would immediately know details of the regulatory structures and genomic context surrounding each gene and would therefore be able to interpret variation in those genes and the functional consequences of such variation in a much more informed way.

I should also add that I view the proposal for a zebra finch genome as genuinely complementary to a second initiative, to sequence the genome of a non-avian reptile (Anolis). The two genomes would be mutually informative; indeed the greatest gain would be realized by completing both genomes.

Here are just a few of the ways in which a zebra finch genome would benefit my current research and inspire new research directions in my laboratory:

– Adaptive evolution of zebra finch genes in nature: A zebra finch genome would allow access to the many genes that are evolving under natural selection in the wild and would permit truly multilocus searches for adaptive compartments of the genome

– Rates of molecular evolution in birds: With the availability of a second avian genome sequence, we would be in a position to ask whether passerine birds exhibit faster rates of evolution than do non-passerines. This hypothesis has received some support from a limited sampling of genes, and now we would be able to address this on a truly genomewide scale.

– Behavioral genetics of zebra finches: Many behavioral ecologists, including myself, have been searching for genes underlying natural mating preferences in birds. Because the zebra finch is such an excellent laboratory model for mate choice and phenotypic plasticity, a full genome sequence would allow us to immediately access specific loci that might underlie patterns of mate choice, secondary sexual traits and inbreeding avoidance, all of which appear to have a genetic component.

– Molecular basis of phenotypic evolution: Because there are a number of pedigrees and mapping studies available for zebra finches, a full genome sequences would allow us to immediately pinpoint genes underlying variation in phenotypes such as bill color,

plumage traits, song differences between species and, indeed, speciation itself.

As you can see, a zebra finch genome would be a remarkable additional resource that we could access and exploit to bring avian evolutionary studies to a new level. This is a fantastic opportunity and I very much hope that it gets funded and given a high priority.

Sincerely,
Scott Edwards
Alexander Agassiz Professor of Zoology
Curator of Ornithology, Museum of Comparative Zoology

Date: Tue, 28 Jun 2005 17:19:24 +0200
To: arnold@ucla.edu
From: Torsten WIESEL <tWiesel@hfsp.org>
Subject: Zebra finch
Cc: prmarler@ucdavis.edu
X-NAS-MessageID: 53968

Dear Art,

Peter Marler just sent me a message about the possibility that, if funding can be obtained, the zebra finch genome might be sequenced , which is wonderful news. Peter also asked if I could send you a note about the potential value of such a study.

For a long time the beautiful story about vocal learning in this bird, first from Peter's laboratory and then by his students Konishi and Nottebohm, is ideal for a genetic analysis. Actually, in my presidential speech at the Society of Neuroscience in 1979, I made this very point emphasizing that this approach had the possibility to open the door to a better understanding of the basis of learning of songs in the bird, but also most likely give clues to the process of language development in children.

A short story about the way David Baltimore, then president of the Rockefeller University, reacted at our joint visit to the very impressive bird colony at the RU Field Center. Seeing cage after cage with zebra finches David shouted "we must find out about the genetics on this bird", fully realizing the potential of such an approach.

If I can be of additional help to indicate my support of the proposed sequencing of the zebra finch genome your should feel free to contact me.

Best regards,

Torsten

Date: Tue, 28 Jun 2005 07:36:00 -0700
To: Art Arnold <arnold@ucla.edu>
From: Peter Marler <prmarler@ucdavis.edu>
Subject: Re: update on genome application
X-NAS-MessageID: 53960

Dear Art and David
I was excited to hear that NHGRI would consider funding the project to sequence the zebra finch genome. I hope and pray that it comes to pass. Now that research on avian vocal learning is accepted as mainstream neuroscience it is crucial that study of the genetic underpinnings of this unique and highly sophisticated behavior, and the
special brain circuitry that makes it possible, becomes a top priority. The progress with the chicken genome is a major step forward, setting the stage for comparing in detail the genomes of the two species, one a vocal learner, the other not. The scientific importance of a project that is focused on this profound behavioral contrast cannot be overstated. The fact that the chicken and the zebra finch have so much basic neuroanatomy in common should bring into relief the specific genomic determinants of the brain circuitry required for vocal learning-mechanisms about which a great deal is already known. From a neurophysiological point of view I am confident in predicting that this would be a very high yield project., for geneticists and behavioral neurobiologists alike. I wish you every success.
With best wishes
Peter Marler
Distinguished Professor Emeritus
University of California, Davis

June 27, 2005

To whom it may concern:

I support the proposal to sequence the zebra finch's genome. Young zebra finches memorize the song of their father and use this memory to guide the development of their own song. My laboratory identified a nuclear protein that is found predominantly in the brain areas that are involved in song learning. However, the bases on which molecular genetic manipulations are carried out are missing in zebra finches in contrast with Drosophila, C. elegans, zebra fish, and mice in which molecular genetics revolutionized the way we study brain and behavior. Genomic information in zebra finches will open up a new vista in such areas as the study of vocal development, brain gender differences and neuroendocrinology.

Sincerely yours

Signed
Masakazu Konishi
Bing Professor of Behavioral Biology
Division of Biology, California Institute of Technology

ROSLIN INSTITUTE
(EDINBURGH)

Dept. Genomics and Genetics
Roslin Institute (Edinburgh)
Roslin,
Midlothian EH25 9PS, UK

Dr. Wesley Warren
Assistant Director
Genome Sequencing Center
Washington University School of Medicine
St Louis, MO 63108, USA

29th June 2005

Dear Wes and Colleagues,

RE: Genome sequencing of the Zebrafinch genome

I am writing to support you and your colleagues in the NIH/NHGRI research proposal "on sequencing the Zebrafinch genome". Last year we worked very successfully together and the Chicken Genome Consortium on the first draft of the chicken genome (Nature, 2004). This was a major landmark for the chicken and has been a tremendous resource for research in agriculture, medicine and basic biology. As part of the analysis we were able to compare the organization of the chicken genome (genes, proteins, gene orders, etc) with those determined for human and other model organisms. However the phylogenetic distance between the chicken and mammals does have its limitations, mostly due to the saturation of sequence divergence between the genome sequences. The availability of the Zebrafinch genome sequence will compensate for this deficiency and will aid the analysis of both avian and the mammalian genomes, here are few examples of what I mean. The selection constraints that limit the divergence of amino acid sequences (often expressed as the Ka/Ks ratio) is being used more and more these days to highlight classes of rapidly evolving proteins e.g. involved in behaviour (e.g. recent work on brain size in man), reproduction and host defence (e.g. so important these days in both avian and humans as we combat diseases like SARs and AVI). We can also use such comparisons to define specific domains within proteins themselves. However to do this we need to compare species within the same clade i.e. birds with birds and mammals with mammals. Another application would be the identification of regulatory sequences, for example, involved in the regulation of genes during chick development. Our analysis has shown that 30-40% of such signals can still be detected between human and chickens. But the recent work by Eliot Margulies has shown that increasing the number of species can increase this value. In addition a further avian species will help to define the rapidly evolving signals for example use din development.

Yours truly,

Prof David W. Burt

The University of Michigan

Department of Ecology and Evolutionary Biology

830 N. University • Ann Arbor, Michigan 48109-1048
734-615-4917 • FAX 734 763 0544
http://www.eeb.lsa.umich.edu/

June 29, 2005

To whom it may concern,

I am extremely excited to learn that Dr. Arnold and his colleagues are proposing to sequence the complete genome of the zebra finch, a Passerine songbird. This project will be of exceptional scientific value to many fields of biology such as ornithology, evolutionary biology, systematics, genetics and genomics, and neurobiology. So far, there is only one bird (chicken) that has its genome completely sequenced. Many evolutionary studies of birds are limited because of this fact. If the zebra finch genome sequence is available, I would like to use it to address these questions.

1. What is the rate and pattern of genomic evolution in birds and how are they different from what we have learnt in mammals.
2. The chicken genome has some unexpected features that need explanation. For example, chicken does not have the sweet taste receptor gene that other vertebrates have. Is it a unique phenomenon to chicken or universal to birds.
3. Birds have many microchromosomes. The evolution of these microchromosomes can be studied if we can compare chicken with zebra finch.
4. It would be interesting to study the expression and function of zebra finch genes that are homologous to those human genes known to be important in speech/language.
5. Darwin's finches of the Galapagos Island have been used as one of the best examples and models in evolutionary biology. These finches and zebra finch are relatively closely related. Having the zebra finch genome sequence will dramatically improve the evolutionary genetic studies of Darwin's finches, which will substantially help us understand evolutionary processes.

In short, I think the zebra finch genome sequence would be an extremely useful resource for many biologists. I strongly support the project. Please feel free to contact me should have any questions.

Sincerely yours,

Jianzhi Zhang, PhD
Assistant Professor
Department of Ecology and Evolutionary Biology
University of Michigan
1075 Natural Science Building
830 North University Avenue
Ann Arbor, MI 48109-1048
Phone (office): 734-763-0527
Fax: 734-763-0544
Email:jianzhi@umich.edu

U N I V E R S I T Ä T W Ü R Z B U R G
I N S T I T U T F Ü R H U M A N G E N E T I K

EXPERIMENTELLE ZYTOGENETIK
BIOZENTRUM AM HUBLAND
97074 WÜRZBURG
GERMANY
TEL. 0931/8884078
FAX: 0931/8884058

Institut für Humangenetik,Biozentrum,Am Hubland, 97074 Würzburg

July 7, 2005

To,
Prof. Arthur P. Arnold
Distinguished Professor and Chair
Department of Physiological Science
UCLA
621 Charles E. Young Drive South,
Room 4117
Los Angeles CA 90095-1606
USA

Dear Dr. Arnold,

I was delighted to know that your group and collaborators are approaching the U.S. National Human Genome Research Institute (NHGRI) to support your proposal to sequence the genome of the zebrafinch (*Taengiopygia guttata*). I fully agree with your objectives to go for the whole genome sequencing of the zebrafinch. Your work documenting that sex differences in the song circuit of zebrafinches is the result of sex chromosome genes acting locally in the brain is highly admired by the scientific community. Obviously zebrafinch offers an ideal model system to study the role of sex chromosome-linked genes in sex specific neural and behavioral differentiation. Differential action of genes on sex chromosomes in sexually dimorphic brain development was evident from mammals which have XX/XY sex chromosomes. On the contrary, birds have a different sex chromosome system in which the male bird is homogametic (ZZ) and the female represents the heterogametic (ZW) sex. Furthermore, the Z chromosome does not show genetic inactivation in the homogametic males. Therefore, genome sequencing of the zebrafinch will contribute to our knowledge about specific genes on sex chromosomes in brain cells and information from birds should complement our knowledge accumulated on the effects of sex chromosome constitution in the mammalian brain.

Recently, myself and Dr. Dave Burt from Roslin institute (UK) edited a *Report on Chicken Genes and Chromosomes* which has been well received by the avian community. However, many laboratories have suggested that there is a need to include the developments going on in other bird genomes. Therefore, your initiative to sequence the genome of a second avian species is a significant step forward beyond the current state of avian comparative genomics and should stimulate significant contacts between different laboratories. My laboratory has been extensively involved in physical mapping Z-linked genes to chicken chromosomes. In case there is a need for physical mapping of BAC clones or genomic probes to zebrafinch or chicken chromosomes we will be happy to cooperate with you.

I wish you every success in your application and hope to stay in contact to know about a potential mutual research activity.

Yours sincerely,

Michael Schmid

From: "Eric R. Kandel" <erk5@columbia.edu>
To: <arnold@ucla.edu>
Date: Mon, 11 Jul 2005 10:31:30 -0400
X-NAS-Validation: {CF759A5F-14C8-4C89-87D7-8857BB0DDBAD}

Dear Art:

This is simply to add my enthusiastic endorsement of your idea to sequence the genome of the zebra finch. This bird has been a marvellous experimental system for numerous studies on vocal learning. Having its genome would move this system into the genomic era and make it even more valuable.

Sincerely,
Eric

Eric R. Kandel, M.D.
University Professor
Fred Kavli Professor and Director, Kavli Institute for Brain Sciences
Senior Investigator, Howard Hughes Medical Institute
Center for Neurobiology and Behavior
Columbia University