

Proposal for Construction of a Physical Map of the Genome of the Zebra Finch (*Taeniopygia guttata*)

9 January 2005

Proposal submitted by:

| | |
|---|--|
| <p>David Clayton, Ph.D., Professor Cell & Structural Biology, Neuroscience Member, Institute for Genomic Biology B107 Chemical & Life Sciences Laboratory University of Illinois Urbana, Illinois 61801 phone 217-244-3668, fax 503-213-6611 dclayton@uiuc.edu http://www.life.uiuc.edu/clayton/</p> | <p>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</p> |
| <p>Wes Warren, Ph.D. Assistant Director Washington Univ. Genome Sequencing Center Washington University School of Medicine 4444 Forest Park Blvd wwarren@watson.wustl.edu Phone: (314) 286-1899</p> | <p>Jerry Dodgson, Ph.D. Department of Microbiology and Molecular Genetics 2209 Biomedical and Physical Sciences Building Michigan State University East Lansing, MI 48824 dodgson@msu.edu Phone: (517) 355-6463 ext 1556</p> |
| <p>Consultants Erich Jarvis, Ph.D., Department of Neurobiology, Duke University Medical Center Claudio Mello, M.D., Ph.D., Neurological Sciences Institute, Oregon Health Sciences University</p> | <p>Tim Birkhead, Ph.D., Department of Animal and Plant Sciences, University of Sheffield, UK Darren Griffin, Ph.D., Dept Biological Sciences, Brunel University, UK</p> |

I. Summary

We propose to develop a complete physical map of the zebra finch genome. The zebra finch (ZF) is a model organism for study of a range of fundamental issues of central relevance to human health and disease. The proposed genomic map will (1) accelerate and empower research on songbirds, an important surrogate system for testing the developmental, behavioral and neurological consequences of genetic variation, (2) enrich our understanding of genome evolution, and 3) enhance the comparative genomics of all avian species to deliver the full promise of the chicken genome sequence. This project will build on existing NIH-supported resources for ZF genomics, including a BAC library (Tg_Ba, www.genome.arizona.edu) and growing EST databases. A key element of our strategic rationale is the leveraging of the recently completed draft sequence of the chicken genome (International Chicken Genome Sequencing Consortium, 2004), the only avian species sequenced to date. The chicken genome shares sufficient synteny and sequence similarity (evolutionary distance ~90 MY) to serve as a template for facilitating the alignment and annotation of the ZF genome map. A high resolution ZF physical map, closely aligned with the chicken sequence, is cost-effective relative to whole genome sequencing, and will focus subsequent studies of high-interest regions of the ZF genome.

II. Importance of the ZF to biomedical and biological research

A great deal of biological research has centered on songbirds. Nearly half of all avian species are songbirds, and all songbirds are members of a single monophyletic Order, Passeriformes. This explosion of genetic and species diversity occurred relatively recently, after the divergence from non-songbirds such as chicken, quail, and pigeon. Songbirds are readily observed in the wild, and thus during the 20th Century a large amount of information became available about their behavior, population biology, ecology and evolution. Songbirds figure prominently in research on topics as varied as stress, reproduction, and endocrinology (Wingfield and Sapolsky, 2003; Wingfield et al., 1997); sperm competition and genetics of sperm morphology (Birkhead, 1998), evolution of reproductive strategies (K. Arnold et al., 2003), immune function and environmental toxicology (Snoeijs et al., 2005, Gee et al., 2004), flight physiology (Hambly et al., 2004), and population and behavioral ecology (Grant et al., 2004). Above all, their greatest importance as a model for biomedicine, however, derives from their ability to communicate via complex learned vocalizations, an ability not found in chickens. Indeed, songbirds are one of the few non-human animals that use auditory feedback to learn their

vocalizations, and are held to be the most tractable experimental model for human speech learning (Marler, 1970; Doupe and Kuhl 1999; Prather and Mooney, 2004; Jarvis, 2004).

By far, the most thoroughly studied single songbird species is the zebra finch (ZF; *Taeniopygia guttata*, family Estrildidae, Suborder Oscines, Order Passeriformes). ZFs are small, have short generation time (4 months) for a complex vertebrate, and breed easily in captivity. Each ZF 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.

In 1976 Nottebohm and colleagues described an interconnected circuit in the brain that controls song and song learning. This circuit comprises a set of large, discrete anatomical nuclei, and it evolved uniquely in songbirds. The identification of discrete brain nuclei clearly linked to a learned behavior led to a series of path-breaking discoveries that have strongly shaped the field of neurobiology ever since. Research on songbirds has had a dramatic impact on concepts of brain development and function in humans because concepts developed first from discoveries on songbirds have been found subsequently also to be generally true in mammals. The perspective offered from studies of songbirds has often had unexpected and major impact on understanding of human biology. Some of the seminal discoveries are the following:

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 disease 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 (Nottebohm, 2004).

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). The song system became a major model system for understanding sexual differentiation of the brain, especially the interaction of sex steroid hormones and non-gonadal factors (Agate et al, 2003; Holloway and Clayton, 2001). This system 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.

Influences of steroid hormones on neural networks. 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). The songbird is a model system for understanding changes in the adult brain induced by hormones.

Steroid hormone synthesis in brain. Estrogen is normally thought of as a gonadal steroid, but in ZFs it is synthesized actively in the brain as well. Recently, estrogen of neural origin has been implicated in causing 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 ZF 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 steroid production in the brain.

The neural basis for learning. The songbird has well-studied auditory and motor pathways adapted for vocal learning. These pathways are homologous or analogous to those in mammals, including the brain regions involved in learning of human language (Jarvis 2004). Unlike the human brain, these pathways are amenable to detailed experimental investigation. In one of the first applications of molecular genetics to songbird biology, Mello et al (1992) reported that the mere act of listening to tape-recorded birdsong induces a sharp wave of gene expression in brain regions associated with auditory perception. Moreover, this genomic response changes with song familiarity and context (Mello et al., 1995, Kruse et al., 2004, Mello, 2002). Meanwhile George et al. (1995) identified a novel gene expressed actively in the song system only during the critical period for song learning and found the expressed protein to be virtually identical to human alpha-synuclein, now implicated in Parkinson's, Alzheimer's and other neurodegenerative diseases (Clayton and George, 1998). Other studies on the song system include the first clear example that retinoic acid, the main metabolite of vitamin A and a morphogen with key roles in embryonic development, is active in the adult brain and regulates the maturation of a learned behavior (Denisenko-Nehrbass et al., 2000). The FoxP2 gene, critical for human language, is

expressed in the song learning pathway in ZFs at higher levels when learning occurs (Heasler et al. 2004; Teramitsu et al., 2004). Study of the neurons in the neural circuit offers tremendous advantages for understanding the cellular events that open critical periods of development and that underlie synaptic plasticity and learning (Brainard and Doupe, 2002). This system has therefore emerged as one of a small number of model systems for the study of the cellular and molecular basis of learning.

Complex auditory processing and auditory-motor 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 auditory physiologists who study how the brain processes sensory stimuli and integrates motor commands with perceptual feedback (Margoliash, 1997) and for those who study neural control of motor sequences (e.g., Hahnloser et al., 2002).

Diversity of songbird species allows for extensive comparative analysis. The great diversity of songbirds means that numerous traits differ in closely related taxa, allowing species comparisons to illuminate the genetic basis for differences in physiology, behavior, or population biology.

Thus, as a model for biomedical investigations, the ZF offers an unusual constellation of strengths:

- Discrete brain nuclei for biochemical analysis, infusions, manipulation via genetic vectors, etc.
- Unique behavioral readout relevant to human learning, speech, and auditory-motor integration (learned song production)
- Potential for neural repair through neurogenesis, steroid signaling, steroidogenesis
- Clear contrasts for study of regulatory biology (sex differences, steroid effects, critical learning periods)
- Interaction of social and environmental factors with neural and genomic responses
- Generation time (4 months) compatible with quantitative genetic analyses
- Opportunity for rich phylogenetic comparisons to many other closely related songbird species

III. Resources for further development of the zebra finch model

A. Genetic Strains and Pedigrees.

Zebra finches breed readily in captivity and reach sexual maturity at 4 months. 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 a bird with male sex chromosomes (ZZ), male brain, and female gonads, and a bird with male phenotype that carried a W chromosome normally found only in females (ZW). These individual birds offer novel insights into the role of the sex chromosomes in tissue development. Genetic analysis of the mutants has recently become much more feasible because of the introduction of techniques for producing metaphase chromosomes from adult tissues (Itoh and Arnold, 2005), and the availability of the ZF BAC library produced with NHGRI funding in 2002 (www.genome.arizona.edu). The study of individual informative mutants is currently extremely time consuming because the relevant BAC clones and probes must be isolated by hand, for use as probes in FISH (fluorescent in situ hybridization) studies of metaphase chromosomes or for identifying genes in areas of chromosome duplication, deletion, or translocation.

Since the 1980s Professor T.R. Birkhead at the University of Sheffield (U.K.), a consultant on this proposal, has developed a comprehensive 18-generation genealogy (pedigree) of ZFs 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, which relate to his established research on post-copulatory sexual selection (sperm competition and cryptic female choice; Forstmeier et al., 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. The value of these genetic resources will be considerably enhanced if the ZF genome is mapped, and if markers become available for linkage studies.

B. BAC library

Through the NHGRI White Paper mechanism, a ZF BAC library (TG_Ba) was constructed by the Arizona Genome Institute (<http://www.genome.arizona.edu/>). The library has an average insert size of 134kb (genome size: 1200 Mb) covering ~16 genome equivalents. It was constructed in the HindIII site of pCUGIBAC1 vector and contains 147,456 clones.

C. EST databases and microarrays

Two large scale songbird transcriptome/EST efforts are being conducted. With the support of NIH RO1 NS045264-03, a normalized cDNA library was prepared from ZF brain RNA (both sexes, multiple ages), and 34,000 clones were 5'-end-sequenced, at the University of Illinois (Keck Center for Comparative and Functional Genomics). Clustering and contig analysis place these cDNAs into ~18,000 non-redundant sequences. These gene products have been 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 ZF ESTs have highly significant hits against the chicken EST collection, and ~72% align to the full chicken genome database (ENSEMBL) by BLASTN alignment. High-quality trimmed sequences have been deposited in Genbank, and both raw and trimmed sequences are available publicly at http://titan.biotech.uiuc.edu/cgi-bin/ESTWebsite/estima_start?seqSet=songbird. The 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 18,000 non-redundant ESTs have been spotted on glass microarrays and are now being distributed to the songbird research community via an organized proposal-and-review process (<http://titan.biotech.uiuc.edu/songbird/>). A second resource at the Duke University Medical Center has produced 21 full-length and subtracted cDNA libraries enriched in sex-specific, developmental, and behaviorally regulated genes (Jarvis et al., 2002). From these libraries, 18,000 cDNA clones were 5'- and 3' end-sequenced, which cluster into ~9,000 unique transcripts. An integrative behavior-annotated cDNA database of these clones is available at <http://songbirdtranscriptome.net/>. These clones are also on cDNA microarrays. The full-length clones are useful for transcriptome analysis and for gene over-expression studies. The two cDNA resources will be integrated to form a superset of ZF ESTs.

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 ZF may be via use of genetic expression vectors and/or RNAi targeted to specific song control nuclei. These techniques can be applied in adult bird 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 ZF, 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 ZF genome

Dr. Darren Griffin of Brunel University (UK), a consultant on this application, will submit an application to the UK Biotechnology and Biological Science Research Council in January 2005, part of which proposes to map ~200 BAC clones to ZF metaphase chromosomes using FISH. We intend to coordinate our efforts with Dr. Griffin so that the FISH mapping can be used as an independent check of the physical map generated here, and to help resolve any difficult portions of the map.

IV. Size of the research community.

A search of the NIH CRISP database in December 2004 indicated that, in the area of songbird research, NIH is currently 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 for these currently funded grants is 359 grant-years. Other support comes from the NSF and non-US granting agencies in Europe and Japan. Because genetic and genomic information has been lacking, these research programs have been held back in research projects that could be improved through the use of genetic tools for measurement and manipulation of genes.

The ZF has emerged as one of the top few avian species used as models for biomedical research. A Medline search for "zebra finch OR taeniopygia OR birdsong OR songbird" retrieves about 2169 papers. In comparison, a search for "Meleagris", the genus of the turkey, retrieves about 377, and "quail OR Coturnix" retrieves 7352. The quail and chicken are quite closely related, and have similar utility in biological research. Songbirds offer an entirely new set of biological phenomena that are not amenable to study in galliforms such as quail and chickens.

We estimate that about 120-200 labs 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. Other national academy members (e.g., Gordon Orians) have studied songbird behavior and ecology. Recently we sent an email request to 70 researchers who work on songbirds or other birds, asking for expressions of interest in this proposal. Forty-seven responded with letters of support. Several are appended to this proposal. The interest in the ZF genome comes not only from those with specific research interests in songbirds, but from other avian biologists and comparative geneticists. For example, the letter from Prof. David W. Burt, of the Department of Genomics and Genetics at the Roslin Institute in Edinburgh (see below), offers to contribute two web resources: help in mapping ZF genes at the *ARK-Genomics* website (www.ark-genomics.org) and linking ZF maps with the chicken and other genomes at www.ensembl.org.

V. Aims, Methods, Budget, and Schedule

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 ZF (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 ZF lineages. Moreover, when individual BAC clones from the ZF are sequenced, they align very closely with homologous regions of the chicken genome, both within and between exons (Luo et al., in preparation). The similarity of the ZF and chicken genome suggests strongly that the chicken genome can be used as a template for construction of the physical map of the ZF genome.

Aim 1. Use BAC fingerprinting to develop a clone-based physical map of the ZF genome. We will use the ZF BAC library (TG_Ba) described above. The WUGSC mapping group uses a high throughput BAC-fingerprinting process based on restriction digest of BAC clones. This process involves DNA isolation, restriction enzyme digestion and high-resolution fragment size separation. All clones in the TG_Ba BAC library will be fingerprinted until an attempted total of 10x whole genome coverage is achieved. We will isolate BAC DNA robotically with a 96-well alkaline lysis protocol and employ high-resolution agarose gel separation of BAC DNA restriction enzyme cleavage products. A high stringency automated band identification of clone digests will be performed with a WUGSC-modified integrated suite of functions, written in MATLAB and collectively called BandLeader. Initial assembly parameters of individual clone fingerprints will be determined empirically. After an optimized initial assembly of the fingerprints is reached, we will refine clone order within contigs using the automated clone ordering program CORAL. Once clone order is established, potential joins between contigs are identified by querying the local database with clones at the extreme ends of each contig, at reduced fingerprint overlap stringency. Fingerprints of clones involved in potential joins will be visually inspected to confirm that all restriction fragments are logically consistent and the joins made are appropriate. Using this pathfinding process, the WUGSC has used clone-based physical maps to guide large genome assemblies for several important species, including *Homo sapiens*. At later stages of map editing the chicken genome assembly will be used to enhance the merging of fingerprint contigs.

Aim 2. Establish a minimum tiling path of clones representing the ZF genome. A minimum tiling path will be chosen using a clone selection application called Minilda. The Minilda method starts with an ordered fingerprint map and chooses clones to maximize the amount of unique content in each clone selection, minimize excessive overlap, avoid gaps between adjacent selections when possible, and ignore clones with unusual fingerprints. After clone selection, identified tile path clones will be rearranged from their original plates for distribution to the community through the Arizona Genome Institute (Rod Wing, Director), which made and distributes the BAC library. This rearranged clone set will all be end sequenced for more comprehensive end sequence coverage within each fingerprinted contig (Aim 3).

Aim 3. Perform BAC End Sequencing (BES). In project year two, at the close of map construction, BES will commence on all clones derived from the minimum tile path and selected targets. The predicted 12,000 BESs associated with the derived tile path will provide high resolution assessments of clone resources for BAC clone sequencing. In addition, the BES will provide an immediate resource to investigators for gene discovery, marker development and future linkage map generation. The sequencing process begins with high quality DNA prepared for fingerprinting as described above and follows these steps: (1) preparation of sequencing reactions; (2) sequence reaction loading and processing on ABI 3730 sequencing robots; (3) transfer of data to a Unix platform, where the runs enter two different automated queues, to complete the transfer step and run the XGASP script, the WUGSC pre-processing system; (4) automated transfer of the data to the WUGSC Oracle database to determine the appropriate destination for each trace; and (5) automated screening for vector.

Aim 4. Perform overgo hybridization and build a comparative chicken-ZF genome map. We will integrate the ZF BAC contig map with selected ESTs and build a comparative chicken-ZF genome map by mapping ZF ESTs, BAC end sequences (Aim 4), and chicken sequences using a high throughput "overgo mapping" strategy.

A. Overgo strategy and methodology. We will link the ZF fingerprint-based BAC contig map to ZF and chicken genes/sequences via high throughput hybridization using "overgo" probes (Ross et al., 1999). Overlapping synthetic oligonucleotides will be labeled by Klenow polymerase end-filling. The overgos will then be hybridized in pools to gridded BAC filters. Since the size and sequence of the probes are completely controlled, the hybridization temperature can be set to be nearly equal for all probes in any pool, and any known repetitive sequence can be eliminated in advance. We primarily employ a 4-dimensional pool approach (Romanov et al., 2003). A two-dimensional overgo strategy (e.g., Gardiner et al., 2004) presents a feasible alternative, likely to generate BAC-marker assignments more rapidly, but also more susceptible to errors, since it lacks the built-in redundancy of our strategy.

B. Overgo design. Overgos can be designed from most non-repetitive DNA sequences 200 bp or longer (for details see Romanov et al., 2003). We check all overgos for repetitive sequences using BLAST. Although we will not be able to detect ZF-specific repeats, given that there is very little ZF sequence in GenBank, most repeats will be filtered out by testing against the chicken genome. The majority of repetitive sequences in the chicken genome are ancient LINE-like CR1-related repeat families, which should also be present in the ZF genome. Alternatively, we will pre-screen overgo probes to eliminate repetitive ones in advance. We propose to use three types of overgo probes:

1. *ZF ESTs.* The ESTIMA and songbird transcriptome databases presently contain about 52,000 ZF ESTs combined, and more are being developed. We will use these sequences to design overgo mapping probes. Most of these identify a clear orthologue in the chicken genome sequence and can be localized on the chicken map. For ESTs, we will use sequence information from the 3' end, where available, as introns are extremely rare in 3' UTRs (introns that occur by chance within an overgo sequence are likely to prevent hybridization to genomic DNA), and this region is least likely to cross-hybridize to other members of a gene family. It will usually be possible to identify 3' UTR regions of ZF ESTs by comparison to homologous chicken ESTs.

2. *Comparative map chicken overgos.* As described below, a high frequency of overgos designed based on chicken sequences (we have over 1000 chicken overgos already available from previous work) show good hybridization to ZF BACs. These overgos were chosen from genes and markers already on the chicken linkage map or, in some cases, chosen to improve the human-chicken comparative map in regions of special interest. The recent availability of the chicken genome sequence will also allow us to design overgos for comparative mapping in a more systematic fashion, for example using the "universal" overgo approach (Thomas et al., 2002; Kellner et al., 2004). Many of the initial ZF fingerprint contigs will subsequently be anchored to the chicken sequence by BAC end sequence homologies. To provide an independent and complementary set of anchors, we will choose overgo probes spaced throughout the chicken genome. Overgos will be chosen either from coding or 3'UTR regions of genes, where possible. In "gene desert" regions, overgos will be chosen from non-coding conserved sequences shown to have very high avian-mammalian similarity (International Chicken Genome Sequencing Consortium, 2004). We estimate that at least 60-70% of our newly designed chicken overgos will detect a single ZF BAC fingerprint contig. This method will provide independent evidence anchoring many BAC contigs to the ZF-chicken comparative map. Since these markers would be evenly spaced, they will assist in anchoring BAC contigs for which there are fortuitously few BES anchors. More important, since many of these overgo markers will be gene-based rather than random sequence-based, they will specifically identify the BACs that contain genes of particular interest. Probes will be included for all specific genes of interest to the songbird community, along with chicken genes related to endocrinology, reproduction and behavior that are likely to be of interest in songbird research.

3. *Gap filling overgos from ZF BES.* In the final phase of this project, we will focus on filling gaps in the ZF BAC contig map. At this point, we expect to have a preliminary BAC contig map that is anchored to the chicken genome sequence by both BES matches and overgo analysis. Fingerprint-based contigs are overly conservative in detecting overlaps. (Two BAC inserts may overlap but generate few (or even no) common restriction fragments. Therefore, they will not be "called" as overlapping by the software program in order to mitigate against false positive overlap calls.) This is even true for the second generation chicken BAC contig map that contains only 260 contigs (Wallis et al., 2004). In many cases, alignment of the BAC contig map to the chicken genome sequence will predict that two adjacent contigs should either overlap or fall within <100kb of one another (smaller than the size of a typical BAC insert). In these cases, we will employ end sequences from the terminal BACs of the two contigs that are predicted to overlap or be nearby. Alternatively, we will use alignment to the chicken genome to predict the segment of chicken sequence that should lie

within the overlap or span the small gap. We will use either the ZF BES sequences or the predicted chicken sequences to design overgos. Hybridization of such overgos to ZF BAC filters may either confirm the suspected overlap or identify gap-spanning BACs that were not detected by fingerprinting alone. Since we will have the comparative alignment of our BAC contigs to the chicken genome, we will prioritize those gaps that appear mostly easily merged, that will merge the largest, most useful contigs or that lie in a particular region of interest within the chicken and ZF genomes.

C. Preliminary tests of chicken overgos for comparative mapping

The Dodgson lab has performed tests of two types of overgo probes on the TG_Ba BAC clone set. First are overgos designed using ZF EST sequences, many of which are homologous to chicken Z chromosome genes. As expected, these appear to be very reliable probes. Although we still have an additional set of hybridizations to do before full data analysis for all 216 probes can be completed, 72 pooled overgos based solely on ZF ESTs gave 830 positives, or about 11.5 per probe. Given that many of these overgos are likely on the ZF Z chromosome, and thus may be present at half the normal rate in the BAC library, this success rate actually exceeds what we had expected, based on the estimated size of the library. Complete analysis of all of our pool data will likely weed out some false positives in these raw numbers. The second probe type involves chicken overgo probes, again with almost all of these from the chicken Z chromosome. A group of 72 pooled chicken probes generated 658 positive ZF BACs or about 9 per overgo. This is about what we had expected based on the predicted library size. Although the failure rate will surely be a bit higher for chicken overgos than ZF EST-based overgos, the observed success rate indicates that both existing chicken overgos and future overgos designed based on the chicken genome sequence will be useful probes, especially in filling gaps in the comparative ZF-chicken BAC map. Thus, we are confident that chicken overgo probes hybridize well with ZF BACs despite the ~90 million years of divergent evolution.

Aim 5. Sequence 50 BAC clones of high biological interest. BACs that harbor genomic regions of biological interest will be selected for sequencing to adequate coverage (~6x). BAC clones will be selected after advertising the availability of this sequencing service internationally to songbird and other avian biologists, using established email listservs. A short application will be required from each applicant to explain the biological rationale for the BAC to be sequenced. A committee, comprising the individuals listed on page one of this proposal, will evaluate the proposals and prioritize the BACs for sequencing. The BAC sequence will be made available simultaneously to the proposing investigator and on Genbank. Although the selecting committee itself may propose specific BAC clones, in no case will more than 50% of the clones selected be those proposed solely by committee members. Our intent is to ensure that the regions of most interest to the scientific community, broadly defined, are quickly finished and annotated, with minimal duplication of effort.

Aim 6. Publication of ZF genome map. The WUGSC group has accepted guidelines for the release of all sequence and mapping data. Any sequence of ZF produced by WUGSC, and all subsequent annotated and improved versions of the sequences produced by the community, will be provided, without use restrictions, to the scientific community at large for any and all subsequent research purposes. The raw sequences produced by WUGSC will be deposited into the NCBI trace archives every night. The assembled and annotated sequences likewise will be available for each BAC assembly as soon as they have passed quality control. The public will have continuous access to the evolving annotation via the web. The physical map data will be placed on the WUGSC web site (<http://genome.wustl.edu>) for easy access via standard FPC interfaces.

Project Timelines We anticipate that the physical map project could start in September of 2005. Some parts of this proposed project could start earlier, such as the overgo oligo hybridizations. A map project start date of October 2005 would place total project completion at June of 2006. Therefore, we propose funding this project over a 2 year period.

Benchmark

Estimated Completion (months not cumulative)

Aim 1

| | |
|--|----------|
| Isolation of BAC DNAs | 4 months |
| Restriction digestion and sizing of BAC DNAs | 5 months |
| High stringency clone assembly | 4 months |

Aim 2

| | |
|---|---------|
| Minimum tiling path selection | 1 month |
| Minimum physical genome map BAC set rearranging and archiving | 1 month |

Aim 3

| | |
|---|---------|
| BAC end sequencing of 12,000 clones = 24,000 ends | 1 month |
|---|---------|

Aim 4

| | |
|---------------------------|----------|
| Overgo oligo selection | 1 month |
| Hybridization experiments | 4 months |
| Data analysis | 2 months |

Aim 5

| | |
|------------------------------------|----------|
| Selection of BACs by the community | 3 months |
| Sequencing of all selected BACs | 4 months |
| Sequence annotation of all BACs | 2 months |

Aim 6

| | |
|---|---------|
| Data accessibility for physical maps | 1 month |
| Data accessibility for sequences for BACs | 1 month |

Budget justification

| Activity | Unit Description | Total estimated costs (\$) |
|---------------------------|------------------|----------------------------|
| library acquisition | 1 BAC library | 4,608 |
| fingerprinting | 110,000 BACs | 302,500 |
| BAC end sequencing | 12,000 BACs | 75,120 |
| MAP construction | 4 months | 81,284 |
| Overgo hybridizations | 1000 probes | 61,000 |
| Library rearray | 2 months | 18,360 |
| Individual BAC sequencing | 50 BACs | 105,187 |
| Total Project Cost | | 648,059 |

VI. Expected benefits of this research

1. Identification of regulatory regions in the genome. Although two large cDNA and EST libraries have been generated and partially sequenced (http://titan.biotech.uiuc.edu/cgi-bin/ESTWebsite/estima_start?seqSet=songbird and <http://songbirdtranscriptome.net/>), neither provides any information on non-transcribed regulatory regions of the genome. The physical map proposed here will allow any investigator to quickly locate BAC clones encoding most genes of interest, for further sequencing to identify putative regulatory regions which must be identified for ultimate understanding of the molecular basis of phenotypes. This information will allow the development of reagents for studying and controlling gene expression, for driving specific expression of reporters such as Green Fluorescent Protein (GFP), and for marking cells by their molecular phenotype.

2. Identification of genetic markers. The present proposal will provide the first list of genetic markers for a songbird, and begin to make feasible linkage studies to study the genetic basis for neural and other traits influenced by mutations or genetic polymorphisms.

3. A comparative map of the chicken and zebra finch genome. The comparison of the ZF and chicken genome will allow songbird researchers to make detailed use of the information in the chicken genome sequence. The annotated chicken genome provides a wealth of candidate genes that may contribute to songbird traits of interest. The aligned physical map that we propose to generate immediately would allow songbird researchers access to the corresponding songbird genes, their flanking regulatory regions, and other nearby genes. The relevant ZF BACs can then be employed in a variety of molecular tests, including transcriptional profiling, chromatin structure and DNA methylation analyses, and studies of genetic polymorphism. Comparison of other species' genomes often has demonstrated interesting differences in the expansion or contraction of linked gene families. The ZF BAC map will facilitate similar comparisons with the chicken. For example, have gene families that contribute to song behavior undergone expansion and diversification in the ZF? Additional targeted sequencing of the ZF genome will then allow more detailed chicken/ZF comparisons to be made and new genetic hypotheses to be generated.

4. Contribution to the study of genome evolution. The comparison of the ZF genome map with other vertebrates will shed light on the evolution of genomes. In particular, ZF and chicken are phylogenetically distant within the Neognathae (about 90 My apart, similar to distantly related members of the eutherian mammals), and thus comparisons between these two should substantially contribute to our understanding of the evolution of most of the ~9600 avian species. As noted above, many of these species are of major importance to comparative physiology, ecology, evolution, behavior and endocrinology.

5. Provide a scaffold for subsequent high-throughput genome sequencing of selected regions. The studies outlined in items 1 to 4 are expected to elucidate larger genome regions that will become of special interest. A likely example might be those regions of the Z and W chromosome that contribute to sex determination and brain sexual differentiation (Arnold, 2004). At some point in the future, it likely will prove worthwhile to sequence these larger regions of the ZF

genome. The physical map allows for such an approach in a cost-effective manner, without the need for complete WGS sequencing of the whole ZF genome.

Other support

To our knowledge there are no other applications for producing a physical map of the ZF genome.

Reference List

- 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. 1998 Sperm competition in birds. *Rev Reprod.* 3:123-9.
- Birkhead TR, Pizzari T. 2002 Postcopulatory sexual selection. *Nat Rev Genet.* 3:262-73.
- 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.
- Clayton DF, George JM 1998 The synucleins: a family of proteins involved in synaptic function, plasticity, neurodegeneration and disease. *Trends Neurosci.* 21, 249-254.
- DeLacoste-Utamsing C, Holloway RL 1982 Sexual dimorphism in the human corpus callosum. *Science* 216: 1431-1432.
- DeVoogd TJ, Nottebohm F 1981 Gonadal hormones induce dendritic growth in the adult avian brain. *Science* 214: 202-204.
- Denisenko-Nehrbass NI, Jarvis Ed, Scharff C, Nottebohm F, Mello CV. 2000 *Neuron* 27, 359-370.
- Doupe A, Kuhl P 1999 Birdsong and human speech: common themes and mechanisms. *Annu Rev Neurosci.* 22:567-631
- Forstmeier W, Coltman DW, Birkhead TR. 2004 Maternal effects influence the sexual behavior of sons and daughters in the ZF. *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.
- 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.
- 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.
- 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. 2004 Learned birdsong and the neurobiology of human language. *Ann N Y Acad Sci.* 1016:749-77.
- 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.
- 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]
- 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.

- Margoliash D. 1997 Functional organization of forebrain pathways for song production and perception. *J Neurobiol.* 33:671-93.
- Marler P. 1970 Birdsong and speech development: could there be parallels? *Am Sci.* 58:669-73.
- 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.
- 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.
- 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. 2004 The road we travelled: discovery, choreography, and significance of brain replaceable neurons. *Ann N Y Acad Sci.* 1016:628-58.
- 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 traveled: 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.
- 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).
- 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.
- Snøeijts 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.
- 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.
- 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.
- 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.
- 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 how... *J Neuroendocrinol.* 15:711-24.

Letters of Support

To: Art Arnold <arnold@ucla.edu>

From: Fernando Nottebohm <nottebo@mail.rockefeller.edu>

Subject: A physical map of the zebra finch genome: a most welcome project!

Date: Mon, 20 Dec 2004 18:44:14 -0500

Dear Art:

I read with great interest the material you sent me and I am very pleased that you and David Clayton have taken on the challenge of putting together, with the help of others -- Drs. Wesley Warren and Jerry Dodgson -- a physical map of the zebra finch genome. As you know, my laboratory's work has taken a strong molecular bent as we have begun to grapple with the molecular events that underlie neuronal replacement in adult brain. Having the physical map on hand, will not only fuel much comparative work on the genetic evolution of birds and vertebrates in general, but will also make it easier to focus on specific genes and their promoters.

The brain of songbirds has become fertile material for the study of basic processes of learning and brain repair. In addition, birds have contributed much to basic issues in development and sexual differentiation. For all of these reasons it is urgent that there be a fully sequenced and annotated genome of a songbird and its physical layout in the chromosomes. I think that all of us working on birds will benefit much from the work that you and David Clayton propose to do and hope it will receive the highest funding priority. I trust that you and David will ensure that the results of this work are made available to all on prompt and fair terms.

Cordially, Fernando

Dr. Fernando Nottebohm

Professor, Director

Rockefeller University Field Research Center

495 Tyrrel Road

Millbrook, New York 12545

Date: Tue, 04 Jan 2005 12:41:29 -0800

To: Art Arnold <arnold@ucla.edu>

From: Peter Marler <prmarler@ucdavis.edu>

Subject: Re: request for letter

Dear Art and David

I was excited to hear about your plans for a grant request to the NHGRI to begin assembling a map of the zebra finch genome. I hope and pray that you are successful. Now that research on avian vocal learning has become 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. This could become a very high yield project., for geneticists and behavioral neurobiologists alike. I wish you every success with it. Do keep me posted on how it all develops.

With Best Wishes

Peter Marler

Distinguished Professor Emeritus

University of California, Davis

6 Jan 2005

Dear Art and David,

Thank you for your recent letter about your plans to submit a grant application to NHGRI to build a physical map of the zebra finch genome based on the available BAC library. I fully support these plans, which will complement and extend the usefulness of the recently completed draft of the chicken genome sequence. Based on Zoo-fish studies of colleagues it is clear that such a map will align with little difficulty with the chicken genome sequence, using the zebra finch BAC-end sequences. Such a comparison will provide more insight in to the exceptionally stable avian genome, when compared to the more dynamic mammalian genomes. So this project will add to our knowledge on vertebrate genome evolution. One of the limitations of the current chicken sequence is the lack of another closely related genome. For example, with a Zebra finch BAC map we could isolate and sequence genes orthologous to chicken genes, from

which we could identify critical amino acids that confer a positive advantage to a bird – through the application of PAML-like analyses of synonymous vs. non-synonymous changes. This is not possible at this time between chicken and mammals due to the 600 million years that separate these extant species. These are only a few of the possible applications of such a resource.

In addition, I can provide two other resources. *ARK-Genomics* (www.ark-genomics.org) is a UK-funded project and we could provide access to the chicken EST collection, which through hybridisation, provides a rapid means of mapping genes in the Zebra finch. Together with Manchester University and EBI (Hinxton), the Roslin Institute has funding from the BBSRC to support the chicken Ensembl database for at least 3 years. We would be able to link the Zebra finch maps with the chicken and other genomes, thus adding further value to all these projects.

Therefore, I fully support your proposal and wish you success with the application.

Yours sincerely,

Professor David W. Burt

Dept. Genomics and Genetics

Roslin Institute (Edinburgh)

Roslin Midlothian EH25 9PS, UK

E-mail: Dave.Burt@bbsrc.ac.uk

WWW site: <http://www.chicken-genome.org/>

From: John Wingfield <jwingfie@u.washington.edu>

Subject: Zebra finch genome

Date: Fri, 10 Dec 2004 14:53:56 -0800

To: Art Arnold <arnold@ucla.edu>, dclayton@uiuc.edu

Dear Art and David:

I am very happy to support your proposal to move forward with production of a physical map of the zebra finch genome, with support from the NHGRI. Not only will this will be a valuable resource for songbird researchers, it should make it easier to link functional studies in songbirds to annotation of related sequences in the human genome, and has the potential to transform research on how organisms deal with environmental change. In my own research, I anticipate that a physical map would help me in my studies of environmental control of transitions in life cycles. Of particular interest is how gene expression changes in response to perturbations of the environment, and how this translates into coping mechanisms for the organism in the real world. We know a great deal about specific hormonal responses to stress, but very little about the consequences of stress and how organisms (including humans) recover.

Your project will have a truly major impact on basic research at all levels.

Sincerely,

John C. Wingfield

Professor of Biology

Department of Biology, Box 351800

University of Washington

Seattle Washington 98195

<http://faculty.washington.edu/jwingfie>

5 January 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 and David,

I am delighted that you will be submitting a proposal to produce a physical map of the zebra finch genome from the NHGRI. Given that the draft of the chicken genome has just been completed, a physical map of a songbird would be an invaluable research tool for comparative genomics. I can imagine a number of ways in which such a map would benefit my own research. First, I have just submitted a NSRA postdoctoral application with Dr. Chris Balakrishnan to conduct population genetic analyses of a number of expressed and anonymous (noncoding) loci in Zebra Finches; a physical map would immediately indicate to us where in the genome these loci are located, and thus considerably improve the quality of

our analyses and our interpretation of correlations in patterns among loci. Part of this proposed research includes sequencing and population variation analysis of the major histocompatibility complex (MHC). It will be crucial to know whether the MHC is located on a macro- or microchromosome, and a physical map would be of immediate use in this regard. Finally, a physical map of zebra finches would allow us and other researchers to understand rates of chromosomal rearrangement in birds, and to compare these to rates in mammals so as to better understand the forces governing such rearrangements in different vertebrate groups. Thus a zebra finch map will not only benefit the extensive community of researchers working with birds, but should be an important advance for vertebrate (and human) genomics generally.

Your proposal has my highest enthusiasm and I hope NHGRI has the vision to support it.

Sincerely,

Scott Edwards

Alexander Agassiz Professor of Zoology

Curator of Ornithology, Museum of Comparative Zoology

Harvard University

Department of Organismic and Evolutionary Biology

26 Oxford Street

Cambridge, Massachusetts 02138 USA

sedwards@fas.harvard.edu

Date: Mon, 20 Dec 2004 12:04:37 +1100

Subject: Letter of Support

From: Andrew Sinclair <andrew.sinclair@mcri.edu.au>

To: Art Arnold <arnold@ucla.edu>

Dear Professor Arnold,

I'm writing in support of your proposal to expand genomic information on the zebra finch.

My research on sex determining genes in the chicken will benefit from having access to a physical map of the zebra finch genome. Primarily because the degree of similarity between the chicken and zebra finch genomes will allow easy identification of orthologous genes and most importantly will identify conserved regulatory regions. This type of comparative analysis is the most efficient way to identify promoter regions of chicken genes. Furthermore, such comparative analysis will provide insights into avian genome evolution.

Best wishes for this proposal.

A/Professor Andrew Sinclair

Dept of Paediatrics, University of Melbourne

Murdoch Childrens Research Institute

Royal Children's Hospital

Melbourne Vic 3052 Australia

From: Michael Brainard <msb@phy.ucsf.edu>

Subject: Re: request for letter December 2004

Date: Thu, 9 Dec 2004 11:47:54 -0800

To: Art Arnold <arnold@ucla.edu>, dclayton@uiuc.edu

Dear Art and David,

I am writing to enthusiastically support your proposal to map the zebra finch genome. As evidenced by the burgeoning number of songbird labs, songbirds have become a premier model system for studying a diverse set of important neurobiological and neuroendocrine questions (learning and memory, sensorimotor control, sexual dimorphism of brain and behavior, the role of new neurons, etc...). Zebra finches in turn are the canonical species for these studies. One of the great advantages of this system has been the ability to combine behavioral, neurophysiological, pharmacological and anatomical techniques to study a specialized neural system (the 'song system') that subserves a neuroethologically important behavior and that is readily accessible for measurement and manipulation. A number of labs are now increasingly excited about the prospects of applying genetic techniques to studying songbirds. My own lab, and many others are beginning to experiment with techniques for in vivo manipulation of genes, using viral infection or in vivo electroporation. Such experiments will potentially allow us to precisely manipulate molecular events and test their function in the context of a circuit with a well defined behavioral role. Moreover, monitoring of levels of expression of different genes in songbirds under differing conditions (stages of development, rearing conditions, etc) is likely to provide important insights into molecular events that contribute to nervous system function. All of these approaches will be greatly facilitated by a systematic mapping and subsequent sequencing of the zebra finch genome. Such an undertaking is hard for any individual laboratory to underwrite, but will be greatly beneficial for the entire birdsong community. It

seems like an ideal enterprise for support by the NHGRI, and I am excited at the prospect that this new resource may become available for general use. Please let me know if there is anything else that I can do provide help for this important undertaking.

best wishes,

Michael S. Brainard
Assistant Professor
Depts. Physiology and Psychiatry
University of California, San Francisco
Room S-762, 513 Parnassus ave.
San Francisco, CA 94143-0444

Date: Thu, 9 Dec 2004 09:25:24 -0500

To: Art Arnold <arnold@ucla.edu>

From: "John R. Kirn" <jrkn@wesleyan.edu>

Dear Art and David,

I strongly support your work both coordinating, and providing empirical data for a zebra finch genome project. With the continuing support of the NHGRI, I can see many important applications across avian biology. However, I see much broader ramifications to the work. Bird studies have impacted virtually all areas of biology (see Konishi et al., 1989 "Contributions of bird studies to biology", Science 246:465-72), and avian embryological work has laid the foundation for much of what we know, and continue to learn, about vertebrate development. One can only begin to imagine how powerful molecular tools, applied to avian biology, will advance science further. My own work centers on understanding adult neuronal addition and loss which occur throughout the telencephalon in birds, but not mammals. Identifying genes that regulate this widespread process has obvious implications for stem cell research and potential therapies for brain repair. Therefore, I see your continuing efforts being of paramount importance for biology and biomedical sciences alike. Sincerely, John Kirn, Professor, Wesleyan University

Date: Thu, 09 Dec 2004 13:20:11 -0500

To: Art Arnold <arnold@ucla.edu>

From: Gerry Borgia <borgia@mail.umd.edu>

Subject: Re: request for letter

Dear Art and David:

I am very pleased to support your proposal NHGRI to map the zebra finch genome. This will be a valuable resource for students of passerine birds and will provide a foundation for linking many years of behavioral, neurobiological and developmental work on these birds to modern genomics. The passerine /songbird radiation is of special comparative interest because it provides a another group to contrast with the mammals and particularly primates in which there has been rapid brain evolution leading to complex behaviors.

My own work is on sexual selection and mate choice and passerines have been a key group in which many of the most important discoveries have been made. Genomic information on a passerine will allow entirely new areas to open integrating genomic information with behavioral and evolutionary studies.

There is a fantastic opportunity here and I sincerely hope this research will be funded.

Best of luck in your endeavors!

Gerald Borgia, Professor
Department of Biology
University of Maryland
College Park, MD 20742-4415

Date: Thu, 9 Dec 2004 18:48:00 -0600

To: Art Arnold <arnold@ucla.edu>

From: Daniel Margoliash <dan@bigbird.uchicago.edu>

Subject: Re: request for letter

Dear Art and David:

I have great enthusiasm and strongly support your proposal to move forward with production of a physical map of the zebra finch genome, with support from the NHGRI. This will be a valuable resource for songbird researchers, and should make it easier to link functional studies in songbirds to annotation of related sequences in the human genome. Beyond any application to my own research questions, of which I envision several, I would point out that the song system has become possibly the premier animal model system for linking cellular and systems level questions in neurobiology with

cognitive and behavioral phenomena. What has been missing from this exceptional mix are the limitations imposed on molecular studies in an avian model. Your proposal is based on excellent scientific logic, and it also would serve to protect and enhance an enormous funding commitment the NIH has already made to song bird research.

I wish you the best of luck in your endeavors!

Sincerely,

Daniel Margoliash

Professor, University of Chicago

From: "Harvey J. Karten" <hjkarten@ucsd.edu>

To: "Art Arnold" <arnold@ucla.edu>

Subject: Re: request for letter

Date: Fri, 10 Dec 2004 06:22:05 -0800

Dear Art and David:

I support your proposal with greatest enthusiasm. Avian models of nervous system development have proven of vital importance in contemporary understanding of spinal cord development, visual system development and organization, adult neurogenesis and in a number of other critical areas of neurobiological research. The release of the first draft of the chicken genome in December will be of vital importance in clarifying fundamental issues in molecular genetics and the regulatory changes associated with all these areas of research. This single event will serve to accelerate research in developmental biology beyond all past accomplishments. The prospect of extending this to the Zebra Finch is very important and exciting. The Zebra Finch is emerging as an experimental subject of vital importance in understanding complex higher functions related to vocal learning. Recent discoveries in the songbird have revealed that similar genes in birds and humans are concerned with these functions. My own work on the avian visual and auditory systems will greatly benefit from this development and facilitate future progress.

I enthusiastically support your proposal to move forward with production of a physical map of the zebra finch genome, with support from the NHGRI. This will be a valuable resource for songbird researchers, and should make it easier to link functional studies in songbirds to annotation of related sequences in the human genome. In my own research, I anticipate that a physical map would help me in my studies of neurotransmitter regulation and morphogenesis in the auditory and vocal control systems.

Sincerely yours,

Harvey J. Karten, M.D.

Distinguished Professor, University of California

Dept. of Neurosciences UCSD

La Jolla, CA 92093

Date: Fri, 10 Dec 2004 09:46:24 -0600

From: Anton Reiner <areiner@utmem.edu>

Subject: Re: request for letter

To: Art Arnold <arnold@ucla.edu>, david clayton <dclayton@life.uiuc.edu>

Dear Art and David,

I thoroughly and enthusiastically support your proposed production of a physical map of the zebra finch genome, and your plan to seek funding for this endeavor from the NHGRI. Due to the well defined and circumscribed brain regions devoted to song learning and production in songbirds, songbirds have emerged as major tools for advancing understanding of the role of the cerebral cortex and basal ganglia in perception of complex sensory stimuli (such as song), in motor learning, and in the higher order neural coding involved in the production of complex motor patterns (such as song). Given that zebra finch are the most commonly studied songbirds, progress in defining the zebra finch genome will be of immense benefit for the full exploitation of the songbird model system, in two general ways. First, mechanistic understanding of the neural and developmental processes involved in perception, learning and motor control requires understanding of the molecules involved in the cellular processes underpinning these phenomena. Information on the zebra finch genome will provide the needed substrate for such molecularly based mechanistic studies. Secondly, the ability to generalize information from zebra finch song learning and production depends on knowing to what extent the same brain regions and same molecules are involved as might be involved in mammals. Insight into the zebra finch genome will, again, make it possible to determine if the molecular fingerprint of brain regions and neural functions devoted to song learning in zebra finch resembles those in specific brain areas and for specific cortical and basal ganglia functions in mammals. In my own research, I anticipate that a physical map would aid immensely in my studies of the role of the basal ganglia in song

learning and the resemblance of this to the role of the basal ganglia in motor learning in mammals. I look forward to progress in your efforts, and greatly appreciate what you are doing for the field of song learning.

Sincerely,

Anton Reiner, Ph.D. Professor
Department of Anatomy & Neurobiology
University of Tennessee Health Science Center
855 Monroe Avenue, Memphis, TN 38163

Date: Mon, 20 Dec 2004 12:04:31 +0000
From: "Katherine Buchanan" <BuchananKL1@Cardiff.ac.uk>
To: <arnold@ucla.edu>
Subject: Re: request for letter

Dear Profs Arnold and Clayton,

I am writing to express my enthusiastic support for your proposal to map the zebra finch genome, given support from the NHGRI. I don't have to point out to you what a model system the avian songbird brain is for posing both proximate and ultimate questions relating to neural development, learning processes, the mechanisms underlying neural function as well as the evolutionary processes shaping neural design. The zebra finch is without a doubt the best species for this work, being the avian 'lab-rat' of neurobiology. With the investment your labs have made in developing innovative molecular tools, your proposed approach seems the natural and exciting step in this process. I realise also that this work would potentially benefit a large number of overseas collaborators and so I wish you every success in your endeavours!

best wishes

Dr Katherine Buchanan
Cardiff School of Biosciences
Cardiff University
Main Building, Park Place
Cardiff CF10 3TL, UK

4 January 2005

Dear Art and David:

I enthusiastically support your proposal to move forward with production of a physical map of the zebra finch genome, with support from the NHGRI. It will be fantastic to be able to take advantage of genomic approaches in the songbird system in order to really move ahead on fundamental questions relating to molecular genetics of learned behaviors. This will be an invaluable resource for songbird researchers, and should make it easier to link functional studies in songbirds to annotation of related sequences in the human genome. In my own research, I anticipate that a physical map would help me in studies using stereotaxically targeted recombinant lentiviral infections to manipulate functional expression of neurotrophic factors such as BDNF.

Best of luck in your endeavors!

Regards,

Sarah Bottjer, Professor
Department of Biological Sciences
University of Southern California