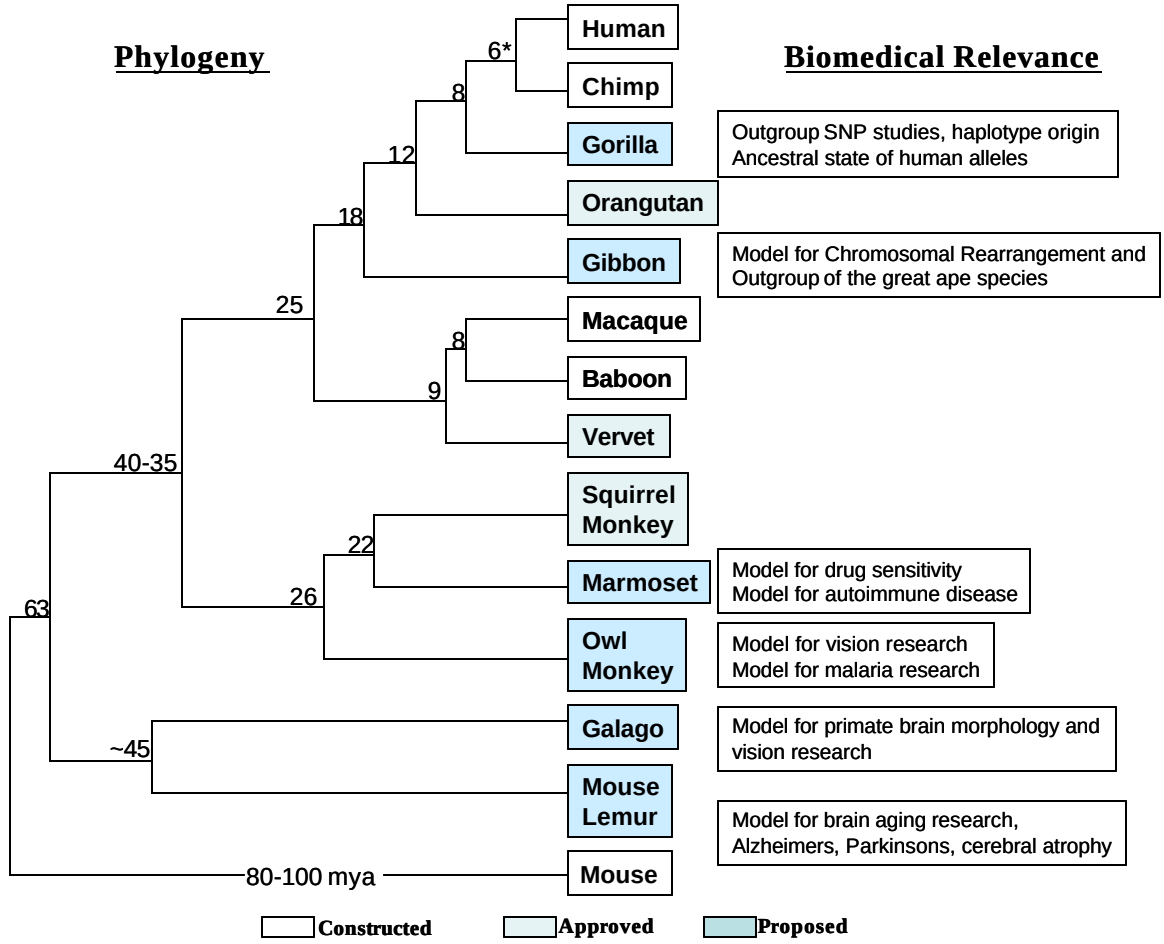
Re: Proposal for Construction a Primate BAC Library Resource
Date: Feb. 10th, 2002
From: Evan Eichler, Ph.D., Dept. of Genetics, Case Western Reserve University
To: BAC Library Resource Network, National Human Genome Research Institute

Importance: The study of the pattern and nature of sequence variation and its association with phenotype is central to the field of genetics. The human and mouse genomes are anticipated to be among the most completely and well-annotated genomes to be sequenced, thus providing important groundwork for the future genetic and genomic studies (Collins et al. 1998). The mouse and human are estimated to have diverged 80-100 million years ago (mya) (Kumar and Hedges 1998). Between these two species, most coding regions show a high degree of sequence similarity such that the majority of orthologous genes can readily be identified (Makalowski et al. 1996). In contrast, most non-coding regions have diverged rapidly due to the simple decay of nucleotide sequence not under selective constraint and processes of duplication, deletion and retrotransposition. The former processes are essential to understanding the nature and pattern of single nucleotide polymorphism in contemporary human populations (Kaessmann et al. 1999; Mathews et al. 2001). Other changes (large-scale events) are important sources for sporadic and recurrent forms of genetic and genomic disease (Emanuel and Shaikh 2001; Stankiewicz and Lupski 2002). The molecular bases and proclivity of specific regions subject to these complex genomic mutational events are not well-understood. The comparison of mouse and human genomes, alone, will not be sufficient to understand the origin and evolution of the complex series of events (Dehal et al. 2001) that have been responsible for this variation. An assessment of genomic variation among a panel of non-human primate species is, therefore, required for understanding the bulk of evolutionary change between these two species and its phenotypic relevance.

I propose the construction of series of BAC libraries from a collection of primate species that have diverged from human at defined intervals over evolutionary time. The principal objective will be for comparative sequencing of regions of high, biological interest (rapidly evolving genes and biomedically relevant regions) and for understanding the pattern of genomic mutation / change over evolutionary time with respect to the human genome. Two criteria were considered in the selection of primate species: 1) its position within the species phylogeny of primates in order to provide a temporal view of genomic mutational change and 2) the relevance of the species to the biomedical research community (Fig. 1). Representatives of phylogenetic anchor species that are relevantto the biomedical research community were given precedent. As part of the formulation of this proposal, a large cross-section of more than one dozen researchers was contacted in order to provide consensus on the most appropriate species for this resource. This included: Jeffrey Rogers (SouthWest Regional Primate Center, non-human primate models of disease), Chi-hua Chiu (Rutgers University, primate molecular evolution), Maynard Olson (University of Washington, structure and function of the human genome), Larry Williams (University of Alabama, comparative medicine), Joe Erwin (Bioqual Inc, primate neurobiology), Mariano Rocchi (University of Bari, comparative primate cytogenetics), Vivien Casagrande (Vanderbilt University, function and evolution of the human visual system), Oliver Ryder (Center for Reproduction of Endangered Species, primate species conservation), John Allman (Caltech Univeristy, primate brain aging research), Cathy Williams (Duke University Primate Center, prosimian evolution and research) and a variety of members from the La Jolla Initiative on

Human Origins (Kenn Kidd, Ajit Varki, etc.) During the writing of this proposal, due consideration was given to existing (chimpanzee, rhesus monkey, baboon) or approved (squirrel, vervet and orangutan) BAC libraries. As such, if a suitable phylogenetic anchor species already existed for a particular primate model, it was not included.

Figure 1: Existing, Approved and Proposed Primate BAC Libraries. Generally accepted phylogeny, divergence times *in millions of years (Goodman 1999) and biomedical relevance of proposed libraries (shaded blue) are summarized.



Usage:

General Considerations

1) Non-human primates share closer behavioural and genetic kinship with humans than any other species. Concomitantly, they are more suitable biomedical models for a variety of human diseases including models of aging, stroke, heart disease, behaviour, drug sensitivity and susceptibility to parasitic infection than any other mammal (Austad 1997; Rogers and VandeBerg 1998).
2) With few exceptions, non-human primates are not classical genetic organisms. Low reproduction rates, long developmental periods, the exorbitant costs of handling and the animal rights movement have curtailed research. However, despite these impediments over 30 species of non-human primates (from an estimated 302) have emerged as

preferred targets for biomedical research (Austad 1997; Rogers and VandeBerg 1998). These represent a refined set of model species that have been selected due to their direct applicability to a particular disease or human condition. The availability of a well-defined human reference sequence provides a unique opportunity for an explosion of research to investigate the genetic basis of these conditions. For example, the existence of large-insert libraries allows orthologous loci to be readily identified through hybridization experiments. The limited sequence divergence among primates (<20% at the nucleotide level) facilitates the rapid construction of physical comparative maps based on comparison of large-insert end-sequence to a human reference sequence. Thus, genes and genetic markers can be recovered readily from model non-human primate species and used for further refined association studies within limited primate breeding programs. **A genomic-based approach involving large-insert libraries provides the most practical means to interrogate the genetic bases for primate traits relevant to disease and evolution.**

3) The paleontological record of mankind and its relatives is among the most complete (Goodman et al. 1998). Over 100 years of anthropological research has provided independent estimates of divergence among the various superfamilies of primates. Comparative sequencing of targeted regions among a panel of non-human primate species provides the opportunity to directly estimate rates of single-base pair change, retroposition, duplication and deletion as a function of evolutionary time (Chen and Li 2001; Li 1997).
4) Comparative sequencing of genomic regions among non-human primates (Chen and Li 2001; Goodman 1999) is essential for testing models of selection. Among immunologists, for example, comparative sequencing between human and non-human primates has been used to provide compelling evidence for models of balancing selection regarding genes associated with human blood group antigens (Grimsley et al. 1998; O'HUigin et al. 2000). Recently published SNP studies emphasize the value of genomic sequence from non-human primates to determine the ancestral and derived status of human alleles (Chen and Li 2001; Kaessmann et al. 2001). Closely related species to human (chimp, gorilla and bonobo) are particularly valuable to eliminate ambiguity with respect to the ancestral status of a common human polymorphism. Sequence from these species provides a critical backdrop for testing the impact of genetic drift and rapid expansion on the frequency and structure of contemporary human haplotypes.
5) Studies of karyotype evolution require the development of a series of large-insert libraries. To date, most studies have been limited to gross cytogenetic observations which likely masquerade the complexity of underlying genomic events. The number of rearrangement events has only superficially been surveyed (Haig 1999; Muller and Wienberg 2001; Yunis and Prakash 1982) and the identification of underlying genetic mechanisms which produce such rearrangements requires breakpoint characterization at the molecular level.
6) Specific regions of the hominoid genome evolve much more rapidly than “generic” DNA and therefore require a series of primate outgroup species in order to resolve the complexity of these regions. Processes such as Y chromosome evolution, pericentromeric duplication, subtelomeric rearrangements and centromere repositioning necessitate the construction of these libraries. As an example, primate genomes are frequently used to determine the timing and movement of recent segmental duplications associated with

chromosomal rearrangement disorders (Velocardiofacial / DiGeorge, Prader-Willi Syndrome, Smith Magenis, etc), pericentromeric duplications and subtelomeric rearrangements (Kuroda-Kawaguchi et al. 2001; Mefford and Trask 2002; Samonte and Eichler 2002; Stankiewicz and Lupski 2002). These regions comprise an estimated 5-7% of the human genome and exhibit accelerated rates of evolutionary turnover (Bailey et al. 2001). Targeted analysis of these regions in outgroup primates has been used to reconstruct the ancestral origin of several segmental duplications and to infer the series of events that have created this duplication architecture in humans and other primates (Chiu et al. 1996; Eichler et al. 1996; Jackson et al. 1999; Johnson et al. 2001; Trask et al. 1998). The construction of a BAC library is necessary in order to survey the structure and organization of these regions over large expanses of genomic sequence (many of the duplications or sites of rearrangement are in excess of 100 kb). Comparative sequencing will provide insight into the underlying mechanisms that have predisposed to duplication-mediated rearrangements associated with human genetic disease.

Summary: The primary use of non-human primate BAC libraries would be for comparative sequencing and mapping analyses of targeted genomic regions. It is anticipated that select regions of high biological / biomedical interest (immunological genes, genes under positive Darwinian selection, regions of rapid genomic rearrangement, haplotype characterization, etc) would be primary targets. Due to the relatively high degree of neutral genomic sequence identity between human and our closest relatives (1.5-20%) it is unlikely that many of these libraries would be used for a complete genomic sequencing effort. The availability of a high quality human reference sequence, however, in combination with a high quality, genomic BAC libraries is essential for cross-genome comparisons (BAC-end sequencing or fingerprint overlay against human) in order to identify regions of hypervariability or regions containing genes of biomedical interest. It should be noted that a whole-genome shotgun approach (whose inserts are relatively limited in size <10 kb) would be ineffective in the resolution of complex genomic regions as outlined above. Furthermore, despite the low level of sequence divergence (<20%), sufficient variation exists within "generic" DNA to thwart the development of contiguous sequence over large genomic DNA using PCR methodology. For example, an attempt to analyze eight regions from the orangutan X chromosome (each 100 kb in length) by PCR amplification from human sequence showed 15% failure after two rounds of oligonucleotide design and amplification based on human reference sequence (Eichler and Chakravarti, unpublished). Similarly, attempts to use cross-species PCR of smaller microsatellite marker amplicons has been problematic when sequence divergence exceeds ~5.5% (>25 mya of separation) (Rogers et al. 2000) and Jeffrey Rogers, *personal communication*). Consequently, anchor species from model organisms are required to facilitate both genetic and genomic studies in various branches of the primate order. Large-insert BAC libraries from a total of 6 primate species at critical points in the primate phylogeny are requested. Specific details regarding biomedical relevance and usage, DNA source material and strain selection is briefly summarized for each species.

1. Gorilla (*Gorilla gorilla*)

***Relevance*:** The gorilla is now recognized as an outgroup to human and chimpanzee having diverged 1-2 mya prior to the separation of these sister taxa (Goodman et al. 1998) .The principal use of a gorilla BAC library would be for the purpose of comparative sequencing in order to

determine the ancestral state of single nucleotide polymorphisms. This is particular relevant in regions of unusual selection, i.e. HLA antigen loci, where comparative sequencing has been used to resolve the evolution of immune-related genes. Many other molecular evolution studies require a third organism to root trees that include the human and chimpanzee comparison (Kaessmann et al. 2001; Mathews et al. 2001). Long-range PCR amplification of genomic regions has proven difficult in this regard with many amplicons >10 kb in length (~10%) failing to PCR. Finally, areas of rapid evolutionary turnover (subtelomeric, pericentromeric, Y chromosome and large-low copy repeats) can not be adequately assessed unless a large insert library of this species becomes available.

Source: Three subspecies of gorilla are recognized (Western lowland, Eastern Lowland and Mountain Gorilla). Western Lowland gorillas are the most common, least endangered and most suitable for the purposes of BAC library construction. A male blood donor has been identified (Frank, *Gorilla gorilla gorilla*, Lincoln Park Zoo). The request for blood has been approved by local IUCAC (Lisa Faust, Lincoln Park Zoo) and should become available between March and April, 2003, when the animal is scheduled for its next physical examination.

2. Gibbon (*Hylobates concolor*).

Phylogenetic rationale: The black gibbon is one of seven representative species of the lesser apes *(Family Hylobatidae*). It represents a phylogenetic link between the great apes and the Old World monkeys. It provides a unique view of genomic temporal change between 15-20 mya of species separation (human and gibbon).

Biomedical rationale: This organism demonstrates an accelerated rate of karyotype evolution--compared to other primate and most mammals (Muller et al. 1997; Muller and Wienberg 2001). Comparative studies indicate an unusually large number (n>45) of chromosomal rearrangements when compared to hominoid species. The black gibbon (*Hylobates concolor)* demonstrates the largest number of such derivative rearrangement events. Unlike most hominoids, these karyotypes have been subjected to a large number of fission events. Comparative sequencing of BACs would be used to understand the molecular basis for chromosomal rearrangements—i.e. the transition region and sequences that may have predisposed to such events. Detection and sequence characterization of such large-scale rearrangements require large-insert libraries to satisfactorily traverse regions enriched in common and low-copy repeats sequences. Information obtained from such studies could provide valuable insight into both germline and somatic chromosomal instability associated with chromosomal rearrangement.

Source: There are at least five different species or subspecies belonging to the genus Hylobates. *Hylobates concolor* has been selected because it shows the greatest amount of karyotype variation (65 conserved linkage groups) when compared to the hominoid ancestral state (n=23) (Burt et al. 1999). This species is also the most commonly held in captivity (World Conservation Union Report, 1996). A biomaterials request for male blood will be made through the Hylobatid Species Survival Plan coordinator, Dr. Alan Varsik at the Santa Barbara Zoo, CA.

Two BAC libraries (Macaque and baboon) representing the Old World Monkeys have been constructed. A third (Cercopithecus aethiops) has been approved. Although the colobine branch of Cercopithecoids is not represented, no biomedically relevant species could be identified. Therefore, there are no additional requests for this group.

3. Marmoset (*Callithrix jacchus*).

***Phylogenetic Rationale*:** This organism is a member of the New World Monkeys (Superfamily Ceboidea), estimated to have diverged from the anthropoid common ancestor (35-40 mya). It is an anchor species of the callitrichine clade, one of seven anciently separated New World monkey clades that diverged from each other at least 18 mya (Chiu et al. 1996). A BAC library for the squirrel monkey, another represenative of the seven ancient clades, has already been approved. Combined (squirrel monkey, marmoset and owl monkey), these three species would provide a reasonable sampling of genomic diversity among the New World monkeys.

***Biomedical Rationale*:** This species is a key organism for studies related to immunity, drug sensitivity and brain function. Its small size, fecundity and inexpensive handling make it one of the non-human primate models of choice. This species is commonly used to assess the toxicological effects of various drugs and has, on occasion, been shown to be a more appropriate model than rodents in which to test adverse drug reaction or long-term side effects (Carey et al. 1992; Jackh et al. 1984). Immunological studies have shown that the marmoset immune system is particularly good model when compared to other primates for testing antibody specificity and recognition. Marmosets have been used to develop models of multiple sclerosis, an autoimmune disease of the central nervous system (Genain and Hauser 2001; t Hart et al. 2000) as well as autoimmune colitis and thyroiditis. Cloning and comparative sequencing of gene clusters associated with drug-detoxification (i.e. cytochrome P450genes) and immune response (T cell receptor, immunoglobulin genes) are essential for providing an understanding of the genetic basis for these events (Mankowski et al. 1999; Schulz et al. 2001; von Budingen et al. 2001). In addition to its primary use in immunological and toxicological studies, the marmoset has been used to develop non-human primate models of coronary heart disease, stroke and reproductive disease (Charnock and Poletti 1994; Marshall et al. 2000). In the case of the latter, considerable effort has been placed on cloning, sequencing and development of expression constructs associated with hormones and their receptors(chorionic gonadotropin, oestrogen receptor, gonadotropin-releasing hormone, hydroxysteroid dehydrogenase, prolactin receptor) (Dalrymple and Jabbour 2000; Husen et al. 2001; Millar et al. 2001; Saunders et al. 2001).

Source: The most commonly used marmoset subspecies in research is *Callithrix jacchus jacchus*. The animal is not endangered. Several large colonies exist within the United States including 235 animals at the Wisconsin Regional Primate Center and ~80 animals located at the Southwest Regional Primate Center. For the purpose of the large insert BAC library construction, liver tissue material immersed in saline solution from a sacrificed male will be obtained from Dr. Jeffrey Rogers, SouthWest Regional Primate Center. Animals are routinely euthanized and there is no difficulty in obtaining fresh tissue source material for library construction. Due to the small size of this animal, limited blood tissue can be obtained from a living marmoset.

4. Owl Monkey (*Aotus trivirgatus*)

Phylogenetic Rationale: It is a member of the New World Monkey family Cebidae and has diverged ~18-20 mya from the squirrel monkey. Sufficient genetic distances separates these two species (esimated 7% nucleotide divergence) complicating cross-species PCR amplification. The biomedical interest in this species and its divergence justify it as a separate anchor species for BAC library construction.

***Biomedical Rationale*:** In the last 10 years, the owl monkey (*Aotous trivirgatus*) has emerged as an important model for studying malaria drug and vaccine development associated with *Plasmodium* infections. Extreme variability in susceptibility to sporozoite infection has been observed among different subspecies of this monkey and these differences have been exploited to

test the efficacy of various malaria treatment regimens. Identification and characterization of wide variety of immunological genes in this species has been a focus of recent research in an attempt to understand the genetic basis for this susceptibility to sporozoite infections (Diaz et al. 2000; Nino-Vasquez et al. 2000; Villinger et al. 2001). In addition to its role as a model of infectious disease, this species has also been used as a model to study brain structure and morphology. The main value of the owl monkey from a neurobiological perspective is that more CNS structures have been mapped electrophysiologically in them than in any other primate except the rhesus macaque. They have also been used extensively in studies of adult cortical plasticity. Considerable research regarding cortical regions, especially the sensory areas of the brain has been published (Ding and Casagrande 1997). Finally due to the unique life history of these primates (e.g monogamy and the disparity in longevity between males and females), it is anticipated that there will be a continued interested in genomic studies of these animals.
Source: *Aotus trivirgatus* is the most commonly used species of owl monkey used in biomedical research. Small colonies are maintained in several regional primate facilities. The animals are not endangered and are easily bred in captivity. For the purpose of large insert BAC library construction, liver tissue material immersed in saline solution will be obtained from a euthanized male (provided by Dr. Vivien Casagrande, Primate Center, Vanderbilt University when one becomes available). A small colony of over 25 individuals is maintained at this facility. In the event that tissue material does not become available in a reasonable time frame (1 year), cell line material (Oliver Ryder / Coriell Institute) or consecutive small blood draws (5 mls) from the same individual (Dr. Casagrande) may be considered as an alternative source.

No biomedically relevant representative of the tarsiers was identified.

5. Malagasy gray mouse lemur (*Microcebus murinus*)
***Phylogenetic rationale*:** The primate order may be divided into two major divisions: prosimians and anthropoids. Ancestral prosimians diverged ~60 mya from the primate lineage leading to the ancestors of New World monkeys, Old World monkeys, apes, and humans. Despite a massive extinction of prosimian species in the late Eocene (50 mya), remarkable diversity still exists (43 species are currently identified). Although several species have acquired adaptive specializations to specific ecological niches, prosimian features are generally regarded as more primitive. The prosimians occupy a unique position both morphologically and phylogenetically in the primate lineage (Goodman et al. 1998). Evolutionarily, they are regarded as the outgroup of all simian species and the link to more "primitive" mammalian orders (Insectivora and Chiroptera). From the perspective of molecular evolution studies, a very strong argument can be made for the construction of BAC libraries from this group. There are at least 2 major divisions of prosimians (galago and lemur). *Microcebus murinus* is a biomedically representative of the latter.
***Biomedical Rationale*:** Over the last 10 years, this species has emerged as a model for aging research. The organisms are small (50-80 g), short lived, fecund (2-3 offspring per year) and reach sexual maturity at a young age (10 months). *Microcebus* shows stereotypical signs of aging such as susceptibility to blindness due to lens opacity, increased frequency of tumour formation, stereotypic geriatric behavioural changes and brain lesions similar to those associated with Alzheimer's disease (Austad 1997). Histological examination of mouse lemur brains have identified the accumulation of A beta deposits within the blood vessel walls of the cortical parenchyma similar those observed in human Alzheimer patients (Gilissen et al. 1999). Several molecular studies have been initiated to recover genes associated with AD and brain aging (Bons

et al. 1995; Calenda et al. 1998). Interestingly, the life span of mouse lemurs is dependent on the number of annual photocycles that the animal experiences. The average life span is 5 annual photocycles. If the photocycle is accelerated to 8 months in duration, the mouse lemur still lives only 5 cycles on average. These observations suggest that they will become important models for other studies related to the molecular mechanisms of aging (Perret 1997). Finally, in recent years, data has emerged that suggest this species may serve as a useful model for bovine spongiform encephalopathy infection (Bons et al. 1999). Its fecundity, propinquity to humans and its usefulness in brain aging research has lead to a dramatic surge in biomedical research on this species. Large research centers with 200-300 individuals are maintained particularly in Europe (Brunoy and Paris, France).

***Source*:** Duke University Primate Center is "an international center for research on living and fossil primates". They have the largest and most diversified collection of lemurs in the U.S which includes a small cohort of mouse lemurs. We have directly contacted Dr. Cathy Williams, director of the DUPC and discussed obtaining access to tissue material. The difficulty with *Microcebus* similar to most prosimians is there size (<100 g) which makes obtaining sufficient blood material (25 ml) in one draw impossible unless the animal is exsanguinated. The DUPC maintains a tissue bank of frozen tissue material from *Microcebus*. Dr. Pieter DeJong has agreed to test the usefulness of frozen tissue for extraction of high quality DNA. If this fails, DNA will be extracted from lymphoblastoid transformed cell lines obtained from Dr Cathy Williams. In addition to this source, Dr. John Allman (Caltech University) has a small colony of *Microcebus*. Two animals are currently scheduled to be euthanized within the next 6 months and he as agreed to donate freshly collected liver material from these organisms for this purpose.

6. Galago. (*Otolemur garnetti*).

***Phylogenetic Rationale*:**Galagos represent the second major division of the prosimians and are estimated to have diverged from the lemurs approximately 43 million years ago (Martin 1990). From a phylogenetic perspective, at least two different species of prosimian should be considered for BAC library construction to control for genomic / genic idiosyncrasies in either lineage and because the lineages leading to extant lemurs and galagos are anciently separated. Furthermore, the degree of neutral sequence divergence (>10%) and extensive karyotype variability requires at least two different anchor species for this suborder of primates.

***Biomedical Rationale*:** Bush babies are prosimian primates; as such, they occupy a unique position in primate evolution, since it has been argued that ancestral prosimians gave rise to both New and Old World simians (Martin 1990). Furthermore, contemporary prosimians are believed to possess more primitive morphological and developmental characteristics. From this perspective, bush baby visual system represents a more basic plan from which specializations in other primate lines arose. Bush babies have several advantages over other primates in proposed, as well as future studies, of the organization of the visual system. Parallel visual pathways are well segregated in bush babies and the koniocellular lateral geniculate nucleus (LGN) pathway is well-developed. There are several unique features of the bush baby particularly suitable including: 1) a relatively small lissencephalic brain which can easily be unfolded and flattened to reveal almost the full extent of all brain areas in single sections, 2) cytoarchitecturally well_defined and segregated LGN layers, 3) well developed LGN parallel pathways (M, P, K), distinct visual (V1) cortical layers, and clear CO blobs, and 4) distinct boundaries between V1 and extrastriate area V2 as well as areas DM and MT (Yamada et al. 1998). Finally, there is the practical consideration that the brains of bush babies have been very well studied. Bush babies

are small primates that breed well in captivity. They are easier and cheaper to house and maintain than their larger cousins. They are not endangered.
***Sources*:** Dr. Casagrande maintains a breeding colony of these animals at the Vanderbilt Primate Center that will supply the tissue needed for these experiments. The source of *O. garnetti* DNA for constructing the BAC genomic library organism will be obtained from leukocytes of freshly drawn blood from a male donor. *O. garnetti* are among the largest galagos (~1.2 kg). Several blood draws from the same individual will be needed in order to create a sufficient amount of material for 10X library construction. DNA plugs will be accumulated over time. High-molecular weight DNA stores well over time without significant degradation of DNA quality (DeJong, personal communication).

Research Community: Interest in the development of an non-human primate libraries is broad (immunology, human genetics, genomics, chromosomal evolution, biomedical research, species conservation etc). In addition to specific researchers developing non-human primate models of disease (see above), there is considerable interest in comparative sequencing within the genetics / genomics community. This includes researchers interested in recapitulating the evolutionary history of human haplotypes (eg. Aravinda Chakravarti, Svante Paabo, Maynard Olson and Wen-Hsiung Li), those studying recently duplicated regions associated with human genomic disorders and rearrangements (eg. Jim Lupski, Evan Eichler, Barbara Trask and Tamim Shaikh) and those interested in large-scale primate comparative genomic sequencing efforts (eg. Eric Green, Ken Dewar, Bruce Roe and Shaying Zhao). Within the field of immunology, there is a long-standing interest in the development of these resources for the characterization of regions associated with blood group antigens or "hypervariable" genes of the immune system (eg. Peter Parham, Ron Bontrop and Eric Long). Those interested in chromosomal evolution (centromere repositioning, Y chromosome evolution and comparative primate cytogenetics) (eg. Mariano Rocchi, David Page, Johannes Wienberg and Malcolm Ferguson-Smith) would benefit from such a resource. Finally, a strong interest has been expressed by those in the area of species conservation (Oliver Ryder and Lori Perkins) that such a resource and the concomitant sequence data generated from it would enhance their ability to manage and monitor genetic diversity within wild and captive populations of this endangered species.

Has the organism been proposed to NHGRI or another publicly funded agency for BAC-based genomic sequence? No formal request, to my knowledge, has been made for any of the organisms in this proposal. A previous proposal was submitted by Dr. Chi-hua Chiu for the galago library which was rejected due to incomplete biomedical justification. I contacted Dr. Chiu who agreed to include allow the organism to be resubmitted as part of this proposal. An independent lemur library (*Lemur catta)* was initiated from cell line material by Jan-Fang Cheng from Lawrence Berkeley Laboratory. Unfortunately due to a lack of funding, the library was terminated early. It is only available for screening at a depth of 5.8X coverage. We've have screened this library extensively and have been unable to recover approximately 30% of targeted regions.

Complementary Genomic Resources. Corresponding lymphoblastoid or primary fibroblast cell lines for these species exists either at Coriell, the Frozen Zoo (Center for Endagnered Species) or as part of private collections eg. Dr. Oliver Ryder. Although not a first choice for preparation of high molecular weight DNA, these materials may be used in the event that high molecular weight DNA of sufficient quality can not be prepared from primary source material (see above).

In addition, cDNA and genomic material are now available from the Center for Reproduction of Endangered Species (Oliver Ryder) for these species.

Proposed Strain Selection. See above for details pertaining to each proposed organism.

Genome Size: Genome size for each of these proposed species is comparable to human, 3 Gb.

Source DNA: See above for details pertaining to each proposed organism.

Specifications: For each species, a random BAC library (*EcoR*1 partial digest, insert size (not less than 150 kb), 10-fold genomic redundancy) from a male individual should be considered using standard cloning vector (i.e. BACe3.6). Libraries significantly less than 150 kb insert will complicate the analysis of duplicated regions as these regions are often in excess of 100 kb. Regions of rearrangement when compared to a human reference sequence require multiple clones to eliminate the possibility of artifacts due to chimeric BACs. Consequently sufficient depth (10X) is requested.

Time frame: It is important that many of the most obvious lacunae within the primate phylogeny be filled as soon as possible. It is particularly essential to generate representatives from the major branches of primates (i.e. lemur, galago, lesser apes, etc). Many analyses of complex regions of the human genome *already* require the existence of this library. Some investigators have been forced to screen inferior cosmid libraries from more distant species (i.e. Gibbon, *Hylobates klossi* library) to recover genes and to understand the evolution and implications of these regions (Johnson et al. 2001). The need for these libraries is immediate and should be considered moderate to high priority. One important consideration is the accessibility of tissues. Due to the nature of tissue collection, the preparation of high molecular weight material for some of these organisms requires significant time (i.e. multiple blood draws from smaller primates or approval of biomaterials request for endangered species—the latter takes ~ 4 months). From a practical perspective, tissue material from gorilla and marmoset can be obtained immediately. Tissue material from other organisms will require 4-12 months to complete. Irrespective of set priorities, tissue acquisition plans must be initiated immediately once the organisms are approved.

Other Support: No other support has been requested or is available for the construction of these libraries.

Other Relevant Information: I have worked previously with Dr. Pieter deJong in obtaining tissue material from primates (orangutan and chimpanzee) for BAC library construction. Dr. DeJong has expressed an interest in continuing to develop a BAC primate resource as well as continuing to explore the feasibility of extracting high molecular weight DNA from both frozen as well as necropsied tissue. Such experimental aspects of BAC-library production are particularly germane when endangered species are being considered and / or the organisms are particularly small in size. In addition, relevant coordinators of primate research centers (Drs. Jeff Rogers, Vivien Casagrande, Larry Williams, Cathy Williams) have all been contacted during the preparation of this proposal and have expressed their support for this endeavor, as well as offered their assistance in obtaining material. The consortium of researchers and BAC library producers is critical to the successful development of this valuable biomedical research resource.

References.

Austad SN (1997) Small Nonhuman Primates as Potential Models of Human Aging. Ilar J 38:142-147

Bailey JA, Yavor AM, Massa HF, Trask BJ, Eichler EE (2001) Segmental duplications: organization and impact within the current human genome project assembly. Genome Res 11:1005-17.

Bons N, Jallageas V, Silhol S, Mestre-Frances N, Petter A, Delacourte A (1995) Immunocytochemical characterization of Tau proteins during cerebral aging of the lemurian primate Microcebus murinus. C R Acad Sci III 318:77-83

Bons N, Mestre-Frances N, Belli P, Cathala F, Gajdusek DC, Brown P (1999) Natural and experimental oral infection of nonhuman primates by bovine spongiform encephalopathy agents. Proc Natl Acad Sci U S A 96:4046-51

Burt DW, Bruley C, Dunn IC, Jones CT, Ramage A, Law AS, Morrice DR, et al (1999) The dynamics of chromosome evolution in birds and mammals. Nature 402:411-3

Calenda A, Mestre-Frances N, Czech C, Pradier L, Petter A, Perret M, Bons N, et al (1998) Cloning of the presenilin 2 cDNA and its distribution in brain of the primate, Microcebus murinus: coexpression with betaAPP and Tau proteins. Neurobiol Dis 5:323-33

Carey GJ, Costall B, Domeney AM, Jones DN, Naylor RJ (1992) Behavioural effects of anxiogenic agents in the common marmoset. Pharmacol Biochem Behav 42:143-53

Charnock JS, Poletti VM (1994) Dietary lipids and adipose tissue fatty acids in the marmoset monkey (Callithrix jacchus). Comp Biochem Physiol Comp Physiol 108:445-9

Chen FC, Li WH (2001) Genomic divergences between humans and other hominoids and the effective population size of the common ancestor of humans and chimpanzees. Am J Hum Genet 68:444-56.

Chiu CH, Schneider H, Schneider MP, Sampaio I, Meireles C, Slightom JL, Gumucio DL, et al (1996) Reduction of two functional gamma-globin genes to one: an evolutionary trend in New World monkeys (infraorder Platyrrhini). Proc Natl Acad Sci U S A 93:6510-5

Collins FS, Patrinos A, Jordan E, Chakravarti A, Gesteland R, Walters L (1998) New goals for the U.S. Human genome project: 1998-2003. Science 282:682-9

Dalrymple A, Jabbour HN (2000) Localization and signaling of the prolactin receptor in the uterus of the common marmoset monkey. J Clin Endocrinol Metab 85:1711-8

Dehal P, Predki P, Olsen AS, Kobayashi A, Folta P, Lucas S, Land M, et al (2001) Human chromosome 19 and related regions in mouse: conservative and lineage specific evolution. Science in press

Diaz D, Naegeli M, Rodriguez R, Nino-Vasquez JJ, Moreno A, Patarroyo ME, Pluschke G, et al (2000) Sequence and diversity of MHC DQA and DQB genes of the owl monkey Aotus nancymaae. Immunogenetics 51:528-37

Ding Y, Casagrande VA (1997) The distribution and morphology of LGN K pathway axons within the layers and CO blobs of owl monkey V1. Vis Neurosci 14:691-704

Eichler EE, Lu F, Shen Y, Antonacci R, Jurecic V, Doggett NA, Moyzis RK, et al (1996) Duplication of a gene-rich cluster between 16p11.1 and Xq28: a novel pericentromeric-directed mechanism for paralogous genome evolution. Hum Molec Genet 5:899-912

Emanuel BS, Shaikh TH (2001) Segmental duplications: an 'expanding' role in genomic instability and disease. Nat Rev Genet 2:791-800

Genain CP, Hauser SL (2001) Experimental allergic encephalomyelitis in the New World monkey Callithrix jacchus. Immunol Rev 183:159-72

Gilissen EP, Jacobs RE, Allman JM (1999) Magnetic resonance microscopy of iron in the basal forebrain cholinergic structures of the aged mouse lemur. J Neurol Sci 168:21-7

Goodman M (1999) The genomic record of Humankind's evolutionary roots. Am J Hum Genet 64:31-9

Goodman M, Porter CA, Czelusniak J, Page SL, Schneider H, Shoshani J, Gunnell G, et al (1998) Toward a phylogenetic classification of Primates based on DNA evidence complemented by fossil evidence. Mol Phylogenet Evol 9:585-98

Grimsley C, Mather KA, Ober C (1998) HLA-H: a pseudogene with increased variation due to balancing selection at neighboring loci. Mol Biol Evol 15:1581-8

Haig D (1999) A brief history of human autosomes. Philos Trans R Soc Lond B Biol Sci 354:1447-70.

Husen B, Adamski J, Rune GM, Einspanier A (2001) Mechanisms of estradiol inactivation in primate endometrium. Mol Cell Endocrinol 171:179-85

Jackh R, Rhodes C, Grasso P, Carter JT (1984) Genotoxicity studies on di-(2-ethylhexyl) phthalate and adipate and toxicity studies on di-(2-ethylhexyl) phthalate in the rat and marmoset. Food Chem Toxicol 22:151-5

Jackson M, Rocchi M, Hearn T, Crosier M, Guy J, Viggiano L, Piccininni S, et al (1999) Characterisation of the heterochromatin/euchromatin boundary at 10q11 and identification of novel transcripts by repeat induced instability. Am J Hum Genet 65 (Supplement):A56

Johnson ME, Viggiano L, Bailey JA, Abdul-Rauf M, Goodwin G, Rocchi M, Eichler EE (2001) Positive selection of a gene family during the emergence of humans and African apes. Nature 413:514-9.

Kaessmann H, Heissig F, von Haeseler A, Paabo S (1999) DNA sequence variation in a non-coding region of low recombination on the human X chromosome. Nat Genet 22:78-81.

Kaessmann H, Wiebe V, Weiss G, Paabo S (2001) Great ape DNA sequences reveal a reduced diversity and an expansion in humans. Nat Genet 27:155-6.

Kumar S, Hedges SB (1998) A molecular timescale for vertebrate evolution. Nature 392:917-20

Kuroda-Kawaguchi T, Skaletsky H, Brown LG, Minx PJ, Cordum HS, Waterston RH, Wilson RK, et al (2001) The AZFc region of the Y chromosome features massive palindromes and uniform recurrent deletions in infertile men. Nat Genet 29:279-86

Li W (1997) Molecular Evolution. Sinauer Associates, Sunderland, MA

Makalowski W, Zhang J, Boguski MS (1996) Comparative analysis of 1196 orthologous mouse and human full-length mRNA and protein sequences. Genome Res 6:846-57

Mankowski DC, Laddison KJ, Christopherson PA, Ekins S, Tweedie DJ, Lawton MP (1999) Molecular cloning, expression, and characterization of CYP2D17 from cynomolgus monkey liver. Arch Biochem Biophys 372:189-96

Marshall JW, Cross AJ, Jackson DM, Green AR, Baker HF, Ridley RM (2000) Clomethiazole protects against hemineglect in a primate model of stroke. Brain Res Bull 52:21-9

Martin RD (1990) Primate Origins and Evolution: A Phylogenetic Reconstruction. University Press, Princeton

Mathews DJ, Kashuk C, Brightwell G, Eichler EE, Chakravarti A (2001) Sequence variation within the fragile X locus. Genome Res 11:1382-91

Mefford HC, Trask BJ (2002) The complex structure and dynamic evolution of human subtelomeres. Nat Rev Genet 3:95-102

Millar R, Lowe S, Conklin D, Pawson A, Maudsley S, Troskie B, Ott T, et al (2001) A novel mammalian receptor for the evolutionarily conserved type II GnRH. Proc Natl Acad Sci U S A 98:9636-41

Muller S, O'Brien PC, Ferguson-Smith MA, Wienberg J (1997) A novel source of highly specific chromosome painting probes for human karyotype analysis derived from primate homologues. Hum Genet 101:149-53

Muller S, Wienberg J (2001) "Bar-coding" primate chromosomes: molecular cytogenetic screening for the ancestral hominoid karyotype. Hum Genet 109:85-94.

Nino-Vasquez JJ, Vogel D, Rodriguez R, Moreno A, Patarroyo ME, Pluschke G, Daubenberger CA (2000) Sequence and diversity of DRB genes of Aotus nancymaae, a primate model for human malaria parasites. Immunogenetics 51:219-30

O'HUigin C, Satta Y, Hausmann A, Dawkins RL, Klein J (2000) The implications of intergenic polymorphism for major histocompatibility complex evolution. Genetics 156:867-77

Perret M (1997) Change in photoperiodic cycle affects life span in a prosimian primate (Microcebus murinus). J Biol Rhythms 12:136-45

Rogers J, Mahaney MC, Witte SM, Nair S, Newman D, Wedel S, Rodriguez LA, et al (2000) A genetic linkage map of the baboon (Papio hamadryas) genome based on human microsatellite polymorphisms. Genomics 67:237-47

Rogers J, VandeBerg JL (1998) Gene Maps of Nonhuman Primates. Ilar J 39:145-152

Samonte RV, Eichler EE (2002) Segmental duplications and the evolution of the primate genome. Nat Rev Genet 3:65-72

Saunders PT, Sharpe RM, Williams K, Macpherson S, Urquart H, Irvine DS, Millar MR (2001) Differential expression of oestrogen receptor alpha and beta proteins in the testes and male reproductive system of human and non-human primates. Mol Hum Reprod 7:227-36

Schulz TG, Thiel R, Neubert D, Brassil PJ, Schulz-Utermoehl T, Boobis AR, Edwards RJ (2001) Assessment of P450 induction in the marmoset monkey using targeted anti-peptide antibodies. Biochim Biophys Acta 1546:143-55

Stankiewicz P, Lupski JR (2002) Genomic architecture, rearrangements and genomic disorders. Trends Genet 18:74-82

t Hart BA, van Meurs M, Brok HP, Massacesi L, Bauer J, Boon L, Bontrop RE, et al (2000) A new primate model for multiple sclerosis in the common marmoset. Immunol Today 21:290-7

Trask B, Friedman C, Martin-Gallardo A, Rowen L, Akinbami C, Blankenship J, Collins C, et al (1998) Members of the olfactory receptor gene family are contained in large blocks of DNA duplicated polymorphically near the ends of human chromosomes. Hum. Molec. Genet. 7:13-26

Villinger F, Bostik P, Mayne A, King CL, Genain CP, Weiss WR, Ansari AA (2001) Cloning, sequencing, and homology analysis of nonhuman primate Fas/Fas-ligand and co-stimulatory molecules. Immunogenetics 53:315-28

von Budingen HC, Hauser SL, Nabavi CB, Genain CP (2001) Characterization of the expressed immunoglobulin IGHV repertoire in the New World marmoset Callithrix jacchus. Immunogenetics 53:557-63

Yamada ES, Marshak DW, Silveira LC, Casagrande VA (1998) Morphology of P and M retinal ganglion cells of the bush baby. Vision Res 38:3345-52

Yunis JJ, Prakash O (1982) The origin of man: a chromosomal pictorial legacy. Science 215:1525-30.