

**a. Importance of *Xenopus tropicalis* to biomedical and biological research**

*Xenopus laevis* is a unique resource for two critical areas in vertebrate biology: early embryonic development and cell biology. In the former, *Xenopus laevis* has led the way in identifying the mechanisms of early fate decisions, patterning of the basic vertebrate body plan, and early organogenesis. Contributions in cell biology and biochemistry include seminal work on chromosome replication, chromatin and nuclear assembly, control of the cell cycle, in vitro reconstruction of cytoskeletal dynamics, and signaling pathways. In fact, *Xenopus* has become a major vertebrate model for the cellular and developmental biology research that is supported by most of the Institutes of the NIH. Information amassed from these studies provides a strong underpinning for future work, and, although *Xenopus laevis* is superb for characterizing the activities of particular genes, only a tiny fraction have as yet been assayed. A major goal now is to examine the expressed genome in the context of this whole range of biological processes using genomic technology, such as ESTs and full-length cDNA libraries. This information will provide an extraordinary resource for a clearer evaluation of the cellular and developmental processes for which *Xenopus* has been so useful, for example by permitting the production of microarrays to be used for systematically evaluating gene expression.

We are now well placed to use *Xenopus tropicalis* to complement studies of *Xenopus laevis* and to extend research into new areas. Two exciting areas that will become increasingly larger foci for researchers are morphogenesis and organogenesis. Both areas are acknowledged by cell and developmental biologists to be key problems about which we know very little regarding fundamental molecular mechanisms, and both areas impinge in important medical problems. To study these areas effectively will require the ability to target gene activities to specific populations of cells, in the former case, and to relatively late stages of development in the latter. This has not been feasible in the past, but now that one can make, with ease, transgenic lines in *Xenopus tropicalis*—because of the efficient technique for transgenesis in *Xenopus*, and the rapid development of this species—the prospects for new insights in these areas are dramatically improved.

Two new areas concerning genomic issues are obvious: First, *Xenopus tropicalis* offers a relatively simple genome among the amphibians. When combined with the mass of research on *Xenopus laevis*, *Xenopus tropicalis* arguably has the most interesting genome yet to be sequenced. The amphibians occupy a key phylogenetic position between those genomes that are, or have been sequenced, namely the mammals and the fish. As comparative genomics develops, the *Xenopus tropicalis* genome will allow insights into the similarities in development and physiology of the vertebrates, as well as the important differences in the use of regulatory DNA sequences. For example, Amphibians have a limb skeleton that clearly resembles that of the mammals, but is radically different from the fin of the fish. Although similar to the mammalian limb, the *Xenopus* limb is different in the numbers of digits, and fusion of skeletal elements. We expect that these similarities and differences will ultimately be understood by a comparative analysis, and lead to a deep understanding of developmental mechanisms. When combined with the ease of analysis of regulatory information in transgenic animals (see below), we expect many new insights and principles to emerge from study of *Xenopus tropicalis*.

As researchers begin to use genome information to study large-scale genome organization issues, *Xenopus* offers another important resource for studying genome evolution. Whole genome duplications, followed by regression back to the diploid state, are acknowledged to be a widespread mechanism for genome evolution. However, there are few extant animal families in which there remain “polyploid” species. The *Xenopus* genus, however, is a very useful

exception, because it includes tetraploid, octoploid and dodecaploid species in addition to the diploid *X. tropicalis*. In the polyploid species many duplicated genes are silenced, for example only one set of genes at the Major Histocompatibility Complex is active. Recent evidence reveals an unexpected and fascinating finding: that in *Xenopus* this silencing occurs non-syntenically, that is, that some genes are silenced from one copy of the complex, and others, from the duplicated copy. The opportunity to study mechanisms for genome-wide partitioning and silencing of gene activities will offer a phenomenal resource for genome researchers.

The development of genetic analysis will allow a quantum leap forward in the potential of *Xenopus*. As a potential genetic system *X. tropicalis* presents very important advantages. When compared to *Xenopus laevis* it develops to sexual maturity in 1/3 the time, has 1/2 the genome size (and is believed to be diploid), and requires 1/5 the housing space. It nonetheless shares the embryological advantages of *Xenopus laevis* and gene sequences are sufficiently similar that probes cross-react between the two species. Haploid and gynogenetic diploid embryos are easy to generate; this greatly reduces the time (and space) needed to make inbred lines and to perform genetic screens when compared, for example, to the resources needed for mouse genetics. As a vertebrate genetic model, this system offers significant advantages over the mouse system, which costs far more to house and maintain, and has raised more controversy as a model because it is a mammal. The evolutionary conservation of developmental processes revealed in recent years among all vertebrates gives us confidence that information revealed from studies of lower vertebrates will, however, at least in large part, apply to mammalian systems.

The utility of a genetic approach in frogs over mice is highlighted in other ways as well. Simply having the large numbers of eggs in a typical *Xenopus* spawning is also of great value. Of course, ability to manipulate the frog embryo is significantly better; an advantage that is underscored when one considers the possibility of making genetic chimeras. In frogs it should be possible to make chimeras of mutant and wildtype tissue to be able to very readily study the effects of gene manipulation on particular developmental events by obviating early, more general lethal effects that have diminished the ability to interpret many mouse knockouts. The urgent need for developing a genetic component to *Xenopus* research now emerges when one considers the implications of recent breakthroughs arising from this system. Very novel findings, for example, regarding signal transduction (e.g. concerning the Wnt and BMP pathways) have resulted from analysis in *X. laevis*. However, as we try to integrate the many components of these pathways the ability of reagents traditionally used in this system (e.g. antisense and dominant negatives) are very limited when compared to the prospect of having mutations in these genes for epistasis studies. The same principle applies to tissue and organ formation, where, for example, in the eye, brain and heart, many regulatory genes have been identified but the subtleties of gene interaction remain poorly understood because there are no developed genetic tools. To take full advantage of the *Xenopus* system the genetic dimension must be developed. Over a dozen mutations are already under study in the Grainger laboratory. The recent RFA on *X. tropicalis* genetics will allow a significant number of labs to develop large numbers of mutations using complementary strategies in the next several years.

Some may ask whether the development of a new lower vertebrate genetics system, in light of recent successes with zebrafish, is warranted. There are a number of points that provide strong support for the new venture. First, the *Xenopus* system has an enormous wealth of information about a multitude of developmental processes and gene activities that act as an unparalleled resource to complement genetic studies. Second, amphibian embryos are generally

much more amenable to embryological manipulations than are fish embryos, particularly at gastrula and neurula stages, an issue which becomes important for example, in making genetic chimeras. Third, at least at present, transgenic technology is more advanced in the frog system; one can readily generate, at low cost, hundreds of transgenic animals per day. Finally, the zebrafish genome is now believed to have undergone a duplication event, so that the genome is tetraploid relative to mammals, diminishing at least to some extent its potential as a genetic system. There is overwhelming evidence that *X. tropicalis* is a diploid species and efforts in progress in many laboratories will establish this fact unequivocally.

**b. Uses to which the BAC library would be put, in addition to genomic sequencing:**

1. Promoter analysis of genes – a growing number of *Xenopus* genes are being shown to contain regulatory regions that extend more than 100 kb. The only effective way to obtain such clones is from large insert genomic libraries. BAC library plus sequence will greatly speed up the process of identification and isolation of genetic control elements (promoters, enhancers, silencers etc). The highly efficient transgenic system in *Xenopus* provides an important complementary technology for evaluating promoter function.
2. Making transgenic lines to target gene activities (and modified genes, e.g. dominant negative constructs) to particular tissues. As noted earlier, targeting of gene activities to particular groups of cells and tissues will be vital for elucidating key mechanisms of morphogenesis and organogenesis. Using promoters isolated from BAC clones will vastly increase the prospects for doing this kind of genetic intervention.
3. Mapping RH panel mutants to genomic sequences. An *X. tropicalis* RH panel is now under construction for mapping purposes.
4. Cloning genomic loci corresponding to mutations identified from the mutagenesis grants funded by the recent RFA.
5. Identifying *Xenopus* homologs of interesting human genes for functional analysis.
6. Rapid isolation of *X. tropicalis* homologs of *X. laevis* genes.

**c. The size and interest of the research community who will use the proposed BAC library**

NIH currently funds 402 research grants utilizing *Xenopus* with a total award amount of \$86,289,711. In addition, NSF funds an additional 13 awards in the amount of \$2,000,000. The current mailing list for researchers interested in *Xenopus* research numbers 328 researchers worldwide. The *Xenopus* community has been canvassed by e-mail regarding the BAC library and support was uniformly enthusiastic.

**d. Whether the organism will be, or has been, proposed for BAC-based sequencing and the status of that request.**

Prof. Rob Grainger has given a presentation at Washington University of St. Louis regarding the benefits, use and feasibility of sequencing the *Xenopus tropicalis* genome. It is anticipated that a white paper will follow targeted at the next deadline for submissions.

**e. Other genomic resources that are available that will complement this resource.**

*Xenopus laevis* and *tropicalis* EST projects are currently underway with 118,996 ESTs available for *Xenopus laevis* and 6,632 ESTs from *Xenopus tropicalis*. The limiting step for *X.*

*tropicalis* has been the availability of high quality cDNA libraries. Bruce Blumberg at the University of California, Irvine has recently received a grant from NCRR to generate eighteen normalized, full-length enriched cDNA libraries from *Xenopus tropicalis* over the next three years. These will be made available to the community at large without restriction and could serve as an important genomic resource for functional studies and working out the details of transcription units from the BACs and genomic sequence when it becomes available.

**f. The strain of organism proposed and the rationale for its selection**

A number of strains of *X. tropicalis*, confirmed to have a diploid genome by karyotyping, are now in labs in the United States, with two large populations, one from Nigeria (N) and one from the Ivory Coast (IC), comprising the preponderance of animals used in most labs. Both N and IC strains have been inbred in the Grainger lab during the past several years. The N strain is now in the seventh generation of inbreeding and the IC strain in its sixth generation. AFLP analysis of polymorphisms in these strains and their crosses, as well as in wild-caught animals, indicate that the N line has by far the fewest polymorphisms. Seven or eight generations of brother-sister inbreeding is often used as the standard for defining an “inbred” line in many organisms. Seventh generation animals are available now and eighth generation animals will be mature within two months. These should serve as a good source for BAC library DNA.

**g. The genome size.**

*X. tropicalis* has a haploid genome size of  $1.7 \times 10^9$  base pairs.

**h. The availability of source DNA and evidence of its quality for this purpose**

As indicated in section f, highly inbred *X. tropicalis* animals are available that should serve as an excellent source of DNA. Like all amphibians, *Xenopus* has nucleated red blood cells. These metabolically quiescent cells, with compacted DNA, and absence of nucleases, are the ideal source for preparation of genomic DNA of the highest quality (Sive et al., 1999).

**i. Specifications for the library (depth, BAC insert size) and rationale**

Representatives of the *Xenopus* community have met to discuss genomics issues on three occasions. The consensus from these meetings was that at least two large insert (100kb) 10 x coverage genomic libraries should be prepared for promoter analysis and mapping purposes. Subsequent conversations with Sandy Clifton at Washington University in St. Louis have revealed that the minimum acceptable size for BACs to be used for genomic sequencing is 150 kb and that two such libraries, each made with a different restriction enzyme, are required to obtain good genomic sequence coverage. Researchers who work with BAC libraries on a regular basis also tell us that it can be very challenging in some cases to isolate a particular gene region from a given library, thus providing another strong reason to have two independent libraries for the community. Consequently we strongly recommend, if funds are available, that two BAC libraries from *X. tropicalis* be made now.

**j. Time frame in which library is needed**

The overwhelming consensus was that the libraries are needed now for promoter mapping and gene identification studies and relatively soon should the genome sequencing become a reality.

**k. Other support that is available or has been requested for the construction of the requested library**

None

**l. The need for an additional BAC library if one or more already exists.**

There are three libraries that have been prepared that should be discussed. The first was an “*X. tropicalis*” library available from Research Genetics. The animals used to make this library were subsequently karyotyped in the Grainger laboratory and found to be from a tetraploid species, i.e., not *X. tropicalis*, and thus this library is not useful for the community. The second library is one that was initiated as a collaborative project between the Grainger lab and Shizhen Qin from the Institute for Systems Biology in Seattle. This library has been only recently completed. It has relatively small inserts (about 75kb) with about 10X coverage, and though some assays to examine the quality of the library have been performed by Dr. Qin, the library has not been evaluated by the community. There is no mechanism for distributing the library to the community at present, though this is being pursued. In addition we have learned that Prof. Naoto Ueno and colleagues in Japan have generated a moderate size and coverage library (~100,000 clones of 100kb) that could be useful for promoter mapping studies when appropriate distribution arrangements can be made. However in this case as well, no evaluation of the library has been made by users in the community.

The advice we have received from numerous sources is that it will be essential for us to have two very large insert libraries for genome sequencing (greater than 150 kb inserts). Such large inserts will also be imperative for evaluation of the larger promoters associated with some genes. In addition, libraries need to be evaluated by a thorough quality-control process. The two existing libraries do not meet these criteria, though after further evaluation, they may be of some value to the community. In summary, we believe that there remains a strong justification for the large-insert, high quality *X. tropicalis* libraries that could be generated by the NIH program.

**m. Any other relevant information.**

This document was written in the context of the trans-NIH initiative on *Xenopus*, in which members of the *Xenopus* community were polled, and then met, to lay out genetic and genomic priorities for the community. Creation of BAC libraries was mentioned by nearly every researcher as a very high priority. In addition, during the past month we have again polled the *Xenopus* community, and other research communities (e.g. those involved in the zebrafish genome initiative, and researchers working on genome projects) for new information to add to the arguments developed earlier. The substance of this white paper is the result of these discussions.

**Submitted by:**

**Rob Grainger, University of Virginia**

**Bruce Blumberg, University of California, Irvine**

**Richard Harland, University of California, Berkeley**

**Chris Amemiya, Benaroya Research Institute, Virginia Mason Research Center**