Impact of the *Xenopus* system on the mission of the NHGRI

Gabriela Loots, PhD - Lawrence Livermore National Labs

The primary mission of NHGRI is to bring a genomic approach to the translation of genomic sequence information into health benefits. NHGRI has outlined a vision for the future of Genomic Research which encompasses three major themes: (I) Genomics to Biology; (II) Genomics to Health and (III) Genomics to Society. Each one of these themes further defines several grand challenges and research targets for the scientific community aimed at facilitating new achievements that would lead to substantial advances in genomic research and its applications to medicine. Several of the grand challenges outline the need to identify and catalog all the structural and functional components encoded in the human genome and to determine the organization of the genetic and protein networks. Comprehensively research aimed at understanding the building blocks of the human genome will eventually help us to understand how each component contributes to the cellular and organismal phenotype, and how evolutionary variation modifies phenotypes and contributes to susceptibility to disease.

Capabilities developed and optimized for model organisms will contribute substantially to efforts to catalogue, characterize and comprehend the entire set of functional elements encoded in the human genome, Compiling this genome 'parts list' represents an immense challenge that will preoccupy decades of research to come. Even the well-known classes of functional elements, such as protein-coding sequences, still cannot be accurately predicted from sequence information alone. Comparison of genome sequences from evolutionarily diverse species has emerged as a powerful tool for identifying functionally important genomic elements. Initial analyses of available vertebrate genome sequences have revealed many previously unidentified protein-coding sequences. Cross-species sequence comparisons have revealed large numbers of homologies outside of known or predicted protein regions, the majority of which are of unknown function. In particular, since *Xenopus* is a unique biological resource for cell and developmental biology, the advancement of genomic tools and resources for the frog genome will directly contribute to the identification and characterization of novel genes with as of yet unidentified function.

While funding has been allocated for the production of *Xenopus* expression tag sequences (ESTs), full-length *Xenopus* cDNA libraries and *Xenopus* microarrays, additional funding to generate a comprehensive *Xenopus*e ORFeome library will create a powerful resource that would benefit not only members of the *Xenopus* community but also members of the wider community of genomics researchers. The *Xenopus* model system has been at the forefront of expression cloning and functional analysis of protein function via gain-of-function experiments. To obtain insights into human gene function, similar assays can be employed to evaluate human transcript activity in *Xenopus* oocytes. Using evolutionary comparisons, a priority for funding would be for examining human transcripts that are highly orthologous in frog, and examine their putative roles during early embryonic development by gain- and loss- of function. Human and frog expression clones can be tested in parallel in gain-of-function experiments and *Xenopus* morpholinos can be subsequently tested in loss-of-function experiments to determine if such genes play critical roles during embryonic development.

Mammal-to-mammal sequence comparisons have revealed large numbers of homologies in non-coding regions, some of which may play important functions in transcriptional regulation. Functional diversification through transcriptional regulation represents one of the hypotheses for phenotypic differences among species. Comparisons of sequences derived from multiple species, especially those occupying

distinct evolutionary positions, could lead to significant refinements in our understanding of the functional importance of conserved sequences, in particular regarding to gene expression patterns. NHGRI has a strong interest in the development of novel tools and approaches for characterizing transcriptional regulatory elements. The recent successes in *Xenopus* transgenesis provide a unique opportunity for transforming the frog into a new inexpensive and efficient *in vivo* transgenic system that would complement, or even replace the current gold standard of mouse transgenesis. The relative large size of *Xenopus* embryos coupled with external development that allows one to monitor events that occur shortly after fertilization would permit the characterization of embryological events that are almost impossible to study in the mouse. In addition transgenesis will allow later embryological events, such as organogenesis to be amenable to molecular analysis in the frog and combine trangenesis with other molecular or embryological manipulations that are routine in the frog. Funding that would facilitate the development of high throughput transgenic technologies in the frog that increase reliable functional characterization of conserved non-coding elements would be of great value to the entire scientific community.

The *Xenopus* community has already greatly benefited from the recently emerging genetic and genomic resources made available for the *Xenopus Tropicalis* and *Laevis* genomes. Among the non-mammalian model organisms advocated for biomedical research, *Xenopus* continues to be underrepresented, despite its tremendous potential to contribute to the advancement of biomedical research. Future tools and resources will further improve *Xenopus'* ability to contribute to the elucidation of the cellular, molecular and genetic mechanisms that control embryonic development, in particular the following resources gaps would highly parallel and contribute to NHGRI's mission:

1. ORFeome: comprehensive catalog of all full length *Xenopus* transcripts that can be used in expression assays to determine function in *Xenopus* embryos.
2. Improving transgenic technologies: high throughput assays that can be used for robust regulatory element characterization
3. Chip-Seq technologies. Development of chromatin immunoprecipitation assays in *Xenopus* for identifying transcription factor DNA targets.
4. Develop novel methods for real-time measurement of transcripts and proteins. Improve the ability to monitor multiple protein interactions at the same time to aid in network elucidation and establish the temporal and cellular distribution of proteins.
5. *Xenopus laevis* and *tropicalis* comparisons provide unique opportunity to understand evolutionary variation between two closely related species both at protein and gene regulatory level. Lessons learned from frog could become paradigm for other types of evolutionary events that have separated other species.
6. Use the large emerging collection of mutant frogs to study effects of sequence variation and phenotyping impact. By combining mutagenesis with allelic series can be generated that would provide a valuable resource for the study of single nucleotide effects on potential disease genes.

References:

The repertoire of G-protein-coupled receptors in *Xenopus tropicalis*. Ji Y, Zhang Z, Hu Y. BMC Genomics. 2009 Jun 9; 10:263.

Chemical genomics identifies compounds affecting *Xenopus laevis* pigment cell development. Tomlinson ML, Rejzek M, Fidock M, Field RA, Wheeler GN. Mol Biosyst. 2009 Apr;5(4):376-84. Epub 2009 Feb 3.
Genomic regions with distinct genomic distance conservation in vertebrate genomes. Sun H, Skogerbø G, Zheng X, Liu W, Li Y. BMC Genomics. 2009 Mar 27;10:133.
Regulatory evolution of a duplicated heterodimer across species and tissues of allopolyploid clawed frogs (*Xenopus*). Anderson DW, Evans BJ. J Mol Evol. 2009 Mar; 68(3):236-47.
Identification of targets of the Wnt pathway destruction complex in addition to beta-catenin. Kim NG, Xu C, Gumbiner BM. Proc Natl Acad Sci U S A. 2009 Mar 31; 106(13):5165-70.
Kruppel-like factor 2 cooperates with the ETS family protein ERG to activate Flk1 expression during vascular development. Meadows SM, Salanga MC, Krieg PA. Development. 2009 Apr; 136(7):1115-25.
Xenbase: a *Xenopus* biology and genomics resource. Bowes JB, Snyder KA, Segerdell E, Gibb R, Jarabek C, Noumen E, Pollet N, Vize PD. Nucleic Acids Res. 2008 Jan;36(Database issue):D761-7.
Testis-derived microRNA profiles of African clawed frogs (*Xenopus*) and their sterile hybrids. Michalak P, Malone JH. Genomics. 2008 Feb; 91(2):158-64.
Inducible gene expression in transient transgenic *Xenopus* embryos. Wheeler GN, Lavery DL, Hoppler S. Methods Mol Biol. 2008; 469:431-49.
Large-scale appearance of ultraconserved elements in tetrapod genomes and slowdown of the molecular clock. Stephen S, Pheasant M, Makunin IV, Mattick JS. Mol Biol Evol. 2008 Feb; 25(2):402-8.
Multiple ING1 and ING2 genes in *Xenopus laevis* and evidence for differential association of thyroid hormone receptors and ING proteins to their promoters. Wagner MJ, Helbing CC. Biochim Biophys Acta. 2008 Mar; 1779(3):152-63.
Single-species microarrays and comparative transcriptomics. Chain FJ, Ilieva D, Evans BJ. PLoS One. 2008 Sep 25; 3(9):e3279.
Expression cloning in *Xenopus* identifies RNA-binding proteins as regulators of embryogenesis and Rbmx as necessary for neural and muscle development. Dichmann DS, Fletcher RB, Harland RM. Dev Dyn. 2008 Jul; 237(7):1755-66.
Loss of neurturin in frog--comparative genomics study of GDNF family ligand-receptor pairs. Hätinen T, Holm L, Airaksinen MS. Mol Cell Neurosci. 2007 Feb; 34(2):155-67.
Exploring nervous system transcriptomes during embryogenesis and metamorphosis in *Xenopus tropicalis* using EST analysis. Fierro AC, Thuret R, Coen L, Perron M, Demeneix BA, Wegnez M, Gyapay G, Weissenbach J, Wincker P, Mazabraud A, Pollet N. BMC Genomics. 2007 May 16; 8:118.
Anti-infection peptidomics of amphibian skin. Li J, Xu X, Xu C, Zhou W, Zhang K, Yu H, Zhang Y, Zheng Y, Rees HH, Lai R, Yang D, Wu J. Mol Cell Proteomics. 2007 May;6(5):882-94.
Evolution and comparative analysis of the MHC Class III inflammatory region. Deakin JE, Papenfuss AT, Belov K, Cross JG, Coggill P, Palmer S, Sims S, Speed TP, Beck S, Graves JA. BMC Genomics. 2006 Nov 2; 7:281.
Comparative genomics of vertebrate Fox cluster loci. Wotton KR, Shimeld SM. BMC Genomics. 2006 Oct 24;7:271.
SANTA domain: a novel conserved protein module in Eukaryota with potential involvement in chromatin regulation. Zhang D, Martyniuk CJ, Trudeau VL. Bioinformatics. 2006 Oct 15; 22(20):2459-62.

Cracking the genome's second code: enhancer detection by combined phylogenetic footprinting and transgenic fish and frog embryos. Allende ML, Manzanares M, Tena JJ, Feijóo CG, Gómez-Skarmeta JL. Methods. 2006 Jul; 39(3):212-9.
Genetic screens for mutations affecting development of *Xenopus tropicalis*. Goda T, Abu-Daya A, Carruthers S, Clark MD, Stemple DL, Zimmerman LB. PLoS Genet. 2006 Jun; 2(6):e91.
Using several pair-wise informant sequences for de novo prediction of alternatively spliced transcripts. Flicek P, Brent MR. Genome Biol. 2006; 7 Suppl 1:S8.1-9.
Identification of mutants in inbred *Xenopus tropicalis*. Grammer TC, Khokha MK, Lane MA, Lam K, Harland RM. Mech Dev. 2005 Mar; 122(3):263-72.
Highly efficient transgenesis in *Xenopus tropicalis* using I-SceI meganuclease. Ogino H, McConnell WB, Grainger RM. Mech Dev. 2006 Feb; 123(2):103-13.

Xenopus grants funded by the Institute:

According to NIH RePORTER Search Tool, in the fiscal year of 2011, the National Human Genome Research Institute (NHGRI) funded two grants for projects involving Xenopus. These grants total to $1,298,885.

2011 *Xenopus* White Paper - Community Needs:

Executive Summary

***Xenopus*: An essential vertebrate model system for biomedical research:**

Model animals are crucial to advancing biomedical research. Basic studies in vertebrate animals rapidly accelerate our understanding of human health and disease. Among the commonly used model animals, the frog, *Xenopus,* has great impact because of its close evolutionary relationship with mammals. Moreover, the remarkable experimental repertoire of the *Xenopus* system has made it a cornerstone of neurobiology, physiology, molecular biology, cell biology, and developmental biology.

Current NIH investment in research using *Xenopus*:

Consistent with its broad utility, the NIH has made a large and continuing investment in *Xenopus* research. Indeed, a search of the NIH rePORT database for R01 or equivalent grants using the search term "*Xenopus*" returned **678 grants for a total of over $217,000,000** for FY09-10. The NIH has also recently demonstrated its commitment to *Xenopus* community resources by approving $2.5 million to establish the National *Xenopus* Resource in Woods Hole, MA and a similar amount to maintain and expand Xenbase, the *Xenopus* Community's online database.

***Xenopus* as a model system for human disease gene function**

Given the tremendous power of the *Xenopus* system, the pace of new biological discovery by the *Xenopus* Community is vigorous. Using *Xenopus*, we have significantly improved our understanding of human disease genes and their mechanisms of action, justifying the NIH's investment. For example:

Xenopus embryos are used for *in vivo* analysis of gene expression and function:
Congenital Heart Disease – *PNAS* 2011. 108, 2915-2920
CHARGE Syndrome – *Nature* 2010. 463, 958-962.

Bardet-Biedl and Meckel-Gruber Syndromes – *Science* 2010. 329, 1337-1340.
Hereditary hypotrichosis simplex – *Nature* 2010. 464, 1043-1047.
Hutchison-Gilford Progeria – *Dev. Cell* 2010. 19, 413-25.
Cutis laxa – *Nat Genet.* 2009. 41, 1016-21.
Colorectal cancer – *Genome Res.* 2009. 19, 987-93.
Nephronophthisis – *Hum Mol Genet.* 2008. 17, 3655-62; *Nat Genet.* 2005. 37, 537-43.

Xenopus egg extracts are used for *in vitro* biochemical studies:

Fanconi Anemia – *Mol. Cell.* 2009. 35, 704-15; *Science.* 2009, 326, 1698-701.
C-myc oncogene – *Nature*. 2007. 448, 445-51.
BRCA1 – *Cell*. 2006. 127, 539-552

Xenopus oocytes are used to study gene expression and channel activity:

Rapid-onset dystonia-parkinsonsim – *Nature* 2010. 467, 99-102.
Trypanosome transmission – *Nature* 2009. 459, 213-217.
Epilepsy, ataxia, sensorineural deafness – *N Engl J Med.* 2009. 360, 1960-70.
Catastrophic cardiac arrhythmia (Long-QT syndrome) – *PNAS* 2009. 106,13082-7.
Megalencephalic leukoencephalopathy – *Hum Mol Genet.* 2008. 17, 3728-39.

Xenopus **as a model system for understanding basic biological processes:**

Xenopus also plays a crucial role in elucidating the basic cellular and biochemical mechanisms underlying the entire spectrum of human pathologies. Just a small fraction of the many recent discoveries are highlighted here:

Xenopus contributes to our understanding of vertebrate genome organization.
(*Science.* 2010*. 328, 633-636*).

Xenopus egg extracts reveal fundamental aspects of cell division.
(*Cell.* 2010. 140, 349-359; *Nature*. 2008. 453, 1132-6; *Science*. 2008. 319, 469-72).

Xenopus reveals new aspects of eukaryotic nuclear structure and function.
(*Cell*. 2010. 143, 288-98; *Science. 2010.* 318, 640-643).

Xenopus embryos are used for studies of Wnt and TGF-β signal transduction.
(*Science*. 2010. 327, 459-463; *Cell*. 2009. 136,123-35).

Xenopus embryos are used for studying mucociliary epithelia.
(*Nat Cell. Biol.* 2009 11 1225-32; *Nature*. 2007. 447, 97-101).

Xenopus embryos are used for studying development of the vasculature.
(*Cell*. 2008.135, 1053-64).

Xenopus egg extracts provide key insights into DNA damage responses.
(*Mol Cell.* 2009. 35,704-15; *Cell.* 2008.134, 969-80).

Xenopus embryos link telomerase to Wnt signaling.
(*Nature*. 2009. 460, 66-72).

Xenopus are used for small molecule screens to develop therapeutics.
(*Nat Chem Biol.* 2010. 6, 829-836; *Blood*. 2009. 114, 1110-22; *Nat Chem Biol.* 2008. 4, 119-25).

Despite its demonstrated utility and despite the recent investments by the NIH, *Xenopus* still lacks many resources that are considered entirely essential for other model systems. It is the consensus of the *Xenopus* community that their biomedical research could be greatly accelerated by the development of key resources of use to the entire *Xenopus* research community.

At the 2010 International *Xenopus* Conference, developmental, cell, and molecular biologists gathered to discuss the resources needed and the priority that should be assigned to each. There was broad community-wide consensus that eleven resources are currently needed, and these were prioritized into two categories: Immediate Needs and Essential Resources:

The Immediate Needs of the ***Xenopus*** **research community:**

1. Generation of the *Xenopus* ORFeome:

-Will enable genome-wide *in vivo* analyses of gene function.
-Will enable genome-wide *in vivo* analyses of protein localization.
-Will enable, when combined with transgenesis, the first large-scale biochemical determination of protein-protein interactions in specific tissues and at specific embryonic stages.
-Will facilitate more-rapid functional characterization of specific proteins.

2. Improvement of the *Xenopus* genome sequence:

-Will accelerate molecular studies by providing a complete catalogue of *Xenopus* genes.
-Will enable completion of the *Xenopus* ORFeomes.
-Will enable genomic analyses & systems biology approaches for novel gene discovery.
-Will facilitate proteomics approaches and peptide analysis.

Essential Resources for *Xenopus* research community:

In addition to these most-pressing needs, the community has identified nine other Essential Resources that should be developed as soon as possible, so that *Xenopus* biologists can more effectively fulfill the missions of the NIH. The *Xenopus* community considers all of these additional resources to be essential, but understands that priorities must be set, and therefore ranks these as indicated below:

3. **Improvement of long-range contiguity in the *Xenopus laevis* genome**
4. **Improvement of *Xenopus* antibody resources**
5. **Loss of function: Zinc Finger Nucleases/TILLING**
6. **Loss of function: Small inhibitory hairpin RNAs**
7. **Novel loss of function/knockdown/knockout technologies**
8. **Intergenic annotation of the *Xenopus* genome**
9. **Improvements of the *X. tropicalis* genome – long range contiguity**
10. **Additions and improvements to Xenbase: the *Xenopus* Model Organism Database**
11. **Frogbook: A comprehensive resource for methods in *Xenopus* biology**

Community Recommendations for Attaining Resources:

The *Xenopus* Community feels that in order to attain these much needed resources it will be imperative to renew the PAR-09-240/1: "Genetic and Genomic Analyses of *Xenopus*". This mechanism can help to direct funding to the establishment of resources that will accelerate research by the entire community. Development of research resources is essential to the NIH mission, but because such work is not hypothesis-driven, these proposals fare poorly in standard CSR study sections. Moreover, the standard study sections typically lack the depth of expertise that is needed to properly evaluate these proposals. The "Genetics and Genomic Analyses of *Xenopus*" PAR allows for a focused and expert review of resource development proposals, and its renewal will help to ensure a continuing return on the current NIH investment in biomedical research using *Xenopus*.

The *Xenopus* Community also feels that, given the ease with which massive amounts of biological samples can be obtained using this organism, a new PAR to support systems biology using *Xenopus* is warranted. A new PAR in this area would

allow all biomedical researchers to exploit the emerging genomic resources for *Xenopus* to perform systems-level analyses *in vivo*, in a vertebrate, and in a cost-effective manner. Such research would generate significant advances into the "New Biology" described below.

Anticipated Gains for Biomedical Research:

Xenopus as an animal model continues to have a broad impact for biomedical research. Given its already long history of large-scale screens of gene function and its broad use in molecular, cell, and developmental biology, the establishment of additional community-wide resources will greatly facilitate the impact of *Xenopus* as a premier vertebrate model for systems-level analyses.

The National Research Council and the National Academy of Sciences have recently called on the United States "to launch a new multiagency, multiyear, and multidisciplinary initiative to capitalize on the extraordinary advances recently made in biology". This report (http://www.nap.edu/catalog.php?record_id=12764) recommends the term "New Biology" to describe an approach to research where "physicists, chemists, computer scientists, engineers, mathematicians, and other scientists are integrated into the field of biology." The promise of systems-level analysis in *Xenopus*, combined with its already proven strengths, make *Xenopus* the ideal model organism for pursuing "New Biology."

Specifically, genome improvements will provide *Xenopus* researchers with the ability to perform genome-wide screens for biological activities that will in turn allow the rapid assembly and analysis of gene regulatory networks and their relationship to phenotypes. The ORFeome will greatly facilitate such genome-wide screening by allowing all ORFs to be rapidly analyzed or large numbers of proteins to be tagged for analysis of protein-protein interaction or for *in vivo* visualization. Using extracts and biochemical purification coupled with mass-spectrometry and genomic sequence, protein interactomes can be rapidly identified and validated. *Xenopus* offers a unique resource because it is the only *in vivo* vertebrate animal model that couples vast amounts of biological material and a sequenced genome, thus cell-type specific interactomes can also be identified. Large-scale genetic screens will identify important novel genes in developmental pathways, especially given the relatively simple genome of *X. tropicalis* compared to zebrafish. Finally, the flexibility of both *Xenopus* extracts and embryos make this system ideal for chemical biology screens.

Identifying gene-regulatory networks, interactomes, and novel genes will be only the first steps. The well-established power of *Xenopus* for rapid analysis of gene function will then allow deeply mechanistic analyses to complement the systems-level approaches described above. It is the combination of these characteristics that distinguishes *Xenopus* from other vertebrate model systems such as mouse and zebrafish and allows for a systems-level approach to understanding biological mechanisms. The tremendous impact of the *Xenopus* model cannot be realized, however, without the immediate development of community-wide research resources. This White Paper presents the needed resources, and we look to the NIH for guidance in how to best achieve these goals.

For complete details of the 2011 *Xenopus* White Paper, please visit
http://www.xenbase.org/community/xenopuswhitepaper.do

Appendix

Project Number	Project Title	Activity	Principal Investigator	Organization Name	Total Cost
2P01HG004120-04	HUMAN GENOME SRUCTURAL VARIATION	P01	EICHLER, EVAN E.	UNIVERSITY OF WASHINGTON	$1,022,726
5R01HG004359-04	DISCOVERY OF CIS-REGULATORY MODULES IN HUMAN GENOME	R01	LI, XIAOMAN	UNIVERSITY OF CENTRAL FLORIDA	$276,159
				Total	$1,298,885