2020 *Xenopus* Community White Paper

Table of Contents:

Section I. Executive Summary

A. *Xenopus* is 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.

B. The NIH has made a substantial recent investment in biomedical research using *Xenopus*. The NIH's stated mission is "*to seek fundamental knowledge about the nature and behavior of living systems and to apply that knowledge to enhance health, lengthen life, and reduce illness and disability*". As evidence of the key role that *Xenopus* research plays in the pursuit of this mission and achieve these goals, The NIH RePorter database identified:

- **304 active grants** using the search term *Xenopus* in **2019-2020** alone.
- These grants were funded by **18 different Institutes** plus the Office of the Director.
- These grants totaled over **$126.7 million**.
- This investment by the NIH has been **sustained:** In fiscal years **2014-2018**, the NIH invested over **$560 million** in this model system.

Importantly, this investment has not been limited to investigator-driven projects, but also includes substantial investment in community-wide resources for *Xenopus* research, including:

- Establishment and recent renewal of the *National Xenopus Resource* at the MBL in Woods Hole (https://www.mbl.edu/xenopus/).
- Establishment and continued support of *Xenbase*, the *Xenopus* model organism database (www.Xenbase.org).
- Development of a *Xenopus* ORFeome (https://www.ncbi.nlm.nih.gov/pubmed/31403250).
- Projects related to sequencing and assembly of the *Xenopus* genomes (e.g., Session et al., *Nature* 2016 (https://www.ncbi.nlm.nih.gov/pubmed/27762356).

The outstanding return on this investment is evident from the large number of published research contributions that have utilized *Xenopus*.

C. The NIH investment in *Xenopus* has been well justified by discoveries in basic biology. In the last ten years alone over 12,500 papers are retrieved from PubMed using the search term *Xenopus*, and *Xenopus* continues to play a crucial role in elucidating the fundamental molecular, cellular and biochemical mechanisms that govern biological systems. Only a small fraction of the discoveries made using *Xenopus* in the past two years are highlighted here:

Molecular and cell biological basis of embryonic development and differentiation:
Briggs, *Science* 2019. https://www.ncbi.nlm.nih.gov/pubmed/29700227
Chen, *Developmental Cell*, 2019. https://www.ncbi.nlm.nih.gov/pubmed/31211992
Yan, *Science* 2018. https://www.ncbi.nlm.nih.gov/pubmed/30467143
Mena, *Science* 2018. https://www.ncbi.nlm.nih.gov/pubmed/30190310

Molecular Neurobiology:
Cioni, *Cell* 2019. https://www.ncbi.nlm.nih.gov/pubmed/30612743
Cagneta, *Molecular Cell* 2019. https://www.ncbi.nlm.nih.gov/pubmed/30595434
Koser, *Nature Neuroscience* 2018. https://www.ncbi.nlm.nih.gov/pubmed/27643431

Genome organization:
Ly, *Nature Genetics* 2019. https://www.ncbi.nlm.nih.gov/pubmed/30833795
Gibeaux, *Nature* 2018. https://www.ncbi.nlm.nih.gov/pubmed/29320479

DNA damage repair:
Sparks, *Cell* 2019. https://www.ncbi.nlm.nih.gov/pubmed/30595447
Wu, *Nature* 2019. https://www.ncbi.nlm.nih.gov/pubmed/30842657

Vertebrate tissue regeneration:
Aztekin, *Science* 2019. https://www.ncbi.nlm.nih.gov/pubmed/31097661

Intracellular scaling:
Browlee, *Cell* 2019. https://www.ncbi.nlm.nih.gov/pubmed/30639102

Tissue Biomechanics:
Thompson, *eLife* 2019. https://www.ncbi.nlm.nih.gov/pubmed/30642430
Barriga, *Nature* 2018. https://www.ncbi.nlm.nih.gov/pubmed/29443958

Collective cell movement:
Shellard, *Science* 2018. https://www.ncbi.nlm.nih.gov/pubmed/30337409

Cytoplasmic organization:
Cheng, *Science* 2019. https://www.ncbi.nlm.nih.gov/pubmed/31672897
Huizar, *eLife* 2018. https://www.ncbi.nlm.nih.gov/pubmed/30561330
Hayes, *eLife* 2018. https://www.ncbi.nlm.nih.gov/pubmed/30015615

Cytoskeletal organization:
Thawani, *Nature Cell Biology* 2018. https://www.ncbi.nlm.nih.gov/pubmed/29695792
Shrank, *Nature* 2018. https://www.ncbi.nlm.nih.gov/pubmed/29925947

D. NIH investment in *Xenopus* has led to significant advances in understanding human disease. In addition to this exceptional work in basic biology, *Xenopus* has also made substantial and important contributions to directed studies of human pathology. Again, a small number of representative contributions from the *Xenopus* community in the past two years include:

Acute Febrile Encephalopathy
Fitchman, 2019, *Am. J. Human Genetics*. https://www.ncbi.nlm.nih.gov/pubmed/31178128

Floating-Harbor syndrome (FHS)
Greenberg, *Cell* 2018. https://www.ncbi.nlm.nih.gov/pubmed/31491386

Frontotemporal lobar degeneration
Qamar, *Cell* 2018. https://www.ncbi.nlm.nih.gov/pubmed/29677515

Heart Disease:
Federspiel, *PLoS Biology* 2019. https://www.ncbi.nlm.nih.gov/pubmed/31490923
Deniz, *Front. Phys*. 2019. https://www.ncbi.nlm.nih.gov/pubmed/31620018

Immunodeficiency–centromeric instability–facial anomalies (ICF) syndrome:
Jenness, 2018, *Proc. Nat. Acad. Sci.* https://www.ncbi.nlm.nih.gov/pubmed/29339483

Nephrotic syndromes:
Braun, *J. Clinical Inv.* 2018. https://www.ncbi.nlm.nih.gov/pubmed/30179222

Primary Ciliary Dyskinesia
Bustamante-Marin 2019, *Am. J. Hum. Genet.* https://www.ncbi.nlm.nih.gov/pubmed/30665704

Tetra-amelia syndrome:
Szenker-Ravi, *Nature* 2018. https://www.ncbi.nlm.nih.gov/pubmed/29769720

Treacher Collins syndrome
Calo, *Nature* 2018. https://www.ncbi.nlm.nih.gov/pubmed/29364875

E. There is a broad consensus within the *Xenopus* community regarding our most pressing infrastructure needs.

As outlined in Section II (pg. 5), the *Xenopus* Community undertook informal discussions at the *17th International Xenopus Conference in August 2018* as well as a more formal process at the *Xenopus Resources and Emerging Technologies (XRET) Meeting* in October 2019. Over 100 *Xenopus* PIs and senior trainees reached a consensus concerning the most crucial needs of our community. As described in detail in Sections III (pg. 6) and IV (pg. 10), these were further prioritized into two Immediate Needs and one Essential Resource.

F. The *Xenopus* community has identified two Immediate Needs:

1. Renewal of Xenbase, the *Xenopus* model organism database. Revealing the deep underlying principles of biological structure and function requires not simply experimentation across the tree of life, but also harmonization of tremendously disparate data. As recognized be the NIH Common Fund's recent "Big Data to Knowledge" Initiative, model organism databases (MODs) serve as crucial conduits for translating the vast wealth of data into a meaningful biological synthesis. The *Xenopus* model organism bioinformatic database, Xenbase, is therefore an essential tool for integrating *Xenopus* research with other model organisms and with clinical efforts. Continued funding for this critical resource is the Community's top priority.

2. Enhancement of efficient gene editing approaches and generation of mutant lines. Advances in biomedical research have centered on the ability to modify the genomes of organisms. For example, generating loss of function alleles has been essential across model organisms in order to understand gene function including creating models of human disease. The extraordinary advances in genome editing technologies (e.g., CRISPR) has permitted the simple and efficient generation of new genetic mutants in *Xenopus* in a truly transformative way. However, this is only the beginning. Exploiting CRISPR and other gene modification techniques, it is now possible to *precisely* modify the genome, for example to modify a single base pair, a series of base pairs, or knock-in specific sequences into the genome. These technologies coupled with the experimental advantages of *Xenopus* have extraordinary potential for the modeling and understanding of human disease. Therefore, these gene editing technologies must continue to be optimized, standardized and disseminated for *Xenopus*. Advancements in the application of these technologies in *Xenopus* to generate mutant lines, increase the efficiency of homology directed

repair, and generate new transgenic lines, including knock-ins, will dramatically expand the repertoire of experiments that can be done in this system. *These approaches are essential for fully exploiting the Xenopus system to model human disease.*

G. The *Xenopus* community has identified one additional Essential Resource:

XenCAt: The single cell atlas of *Xenopus* development: *Science* magazine identified *in toto* -omic analysis of embryos at the single cell level -including in *Xenopus*- as the Breakthrough of the Year for 2018. In this fast-accelerating field, so called "single cell atlases" have already been compared to a "Facebook for cells" – providing a "social" data structure that organizes gene expression of individual cells according to bipartite similarity across cells and genes. The value of such a resource to a wide research community is comparable to whole genome sequencing and annotation earlier in the century. Both demand and opportunity make the *Xenopus* Cell Atlas (XenCAt) an essential resource for the community that is intimately connected to both of the aforementioned "Immediate Needs". On the one hand, *Xenbase* is a natural knowledge portal to collect and host the XenCAt effort and eventually to provide a sophisticated browsing and interrogation environment. On the other, genome editing efforts will be central for construction of the atlas (through generation of tools for lineage tracing, for example) but will also benefit from such an atlas as a detailed spatial and temporal reference for gene expression and function.

Section II: Consensus building process for the 2020 *Xenopus* White Paper:

The *Xenopus* community strongly believes that the pace of discovery using this powerful model system could be further accelerated if additional resources were available. Thus, the community formed the *International Xenopus Board* (IXB) a 501(c)3 non-profit whose mission is to formalize the development of community-wide resources and to support research with *Xenopus* generally (https://www.xenbase.org/entry/doNewsRead.do?id=249).

To identify resources needed by the broad and diverse *Xenopus* community, *Xenopus* researchers met at two IXB-supported meetings during the last 18 months. First, supported by 1R13HD096856, the *17th International Xenopus Conference* brought together 243 scientists representing 28 U.S. states (and 17 countries worldwide) in Seattle in August 2018. In a series of informal discussions, the progress made on the goals of the previous 2016 *Xenopus* Community White Paper were discussed. Second, 95 PIs and senior postdocs assembled at the *Xenopus Resources and Emerging Technologies (XRET) Meeting*, held by the *National Xenopus Resource* in Woods Hole, MA in October 2019. Formal discussions held at the XRET led to a clear consensus concerning the most pressing needs for continued success with *Xenopus* in biomedical research. The community further prioritized resources into two categories: Immediate Needs and Essential Resources.

The goal of the following sections is to outline these needs, justify them, and provide a preliminary plan on how to obtain them.

Section III: The Immediate Needs identified by the *Xenopus* community:

1. Renewal of Xenbase (P41HD064556-10)

Impact:

Comparative functional studies are the foundation of translational research, as hypotheses generated from clinical findings are thoroughly explored in experimentally tractable model organisms so that deeper mechanistic insights can be returned to the physicians and their patients. Accordingly, the careful aggregation and curation of data from model organisms, and its harmonization to human data is essential if we are to fully leverage the power of these diverse experimental systems. As recognized by the NIH Common Fund's recent "Big Data to Knowledge" Initiative, model organism databases (MODs) serve as crucial conduits for translating the vast body of diverse data types into a meaningful biological synthesis.

Xenbase, the *Xenopus* model organism bioinformatic knowledge base, is an essential web-accessible resource that integrates all the diverse biological, genomic, genotype and phenotype data available from *Xenopus* research. Xenbase enhances the value of *Xenopus* data through high quality curation, by providing bioinformatics tools optimized for *Xenopus* experiments, and through the integration of diverse data types in a way not available at any other single resource. Xenbase inter-relates *Xenopus* genomic, epigenetic, gene and protein expression and phenotype data as well as physical reagents such as antibodies, morpholinos, CRISPR-mutant and transgenic *Xenopus* lines.

Xenbase also provides data sharing infrastructure for many other NIH-funded projects enabling access to a number of unique large-scale datasets not available elsewhere including: the pre-release *Xenopus* genomes (HD080708; Rokhsar), the *Xenopus* proteome (GM103785/HD091846; Kirschner), the ORFeome (OD023697; Stukenberg), *Xenopus* antibody resources (OD021485; Alfandari), the International *Xenopus* stock centers (OD10997; Horb, AI059830; Robert) and many omic-based studies (e.g.: HD073104, Kirshner; DE026434; Moody; GM126395, Cho; GM127069, Harland; NS099124, Wills; HL135007, Conlon). Importantly Xenbase links all this *Xenopus* data sets to human genes and diseases and makes *Xenopus* data accessible to the broader biomedical community through reciprocal data sharing with many other resources such as NCBI, UCSC, UniProtK, Ensembl and other MODs (e.g.: MGI, ZFIN, FlyBASE, Echinobase).

Xenbase content, tools and usage have grown tremendously in the last funding cycle. Xenbase now has over 300 unique visitors a day and is the central hub for the research community. It hosts over 16,500 gene pages, supports the newest *Xenopus tropicalis* and *Xenopus laevis* genomes, contains >50,000 *Xenopus* publications, and offers >70,000 gene expression images. Xenbase has added new tools and data types including: 1) support for human disease models with links to OMIM and the human Disease Ontology; and 2) integration of all the RNA-seq and ChIP-seq data available in GEO. Importantly Xenbase processes these data with a single standardized bioinformatic pipeline enabling users to compare different studies with genome browser tracks, downloadable data tables and heatmaps all linked to the other functional data in Xenbase. Xenbase is the only vertebrate model organism database to offer this to date and other MODs (MGI and ZFIN) are in the process of adopting this approach. In sum, Xenbase is the single most important clearinghouse for *Xenopus* data.

Why renewing Xenbase is essential: Maintenance of Xenbase is essential to protect and maximize the NIH's >$120 million annual investment in *Xenopus* research. Xenbase plays a critical role in making data from this research accessible in a meaningful way. Without Xenbase much of the *Xenopus* data generated from R01-funded studies would remain largely buried in the scientific literature. Large-scale

data sets and reagents from resource projects such as the genome (HD080708, Rokhsar), ORFeome (OD023697) and various omic-studies (HD073104, Kirshner; GM126395, Cho; GM127069, Harland) would be largely unavailable. Virtually all research projects using *Xenopus* depend on the unique tools and highly curated data in Xenbase on a daily basis and without these most of this research would effectively grind to a halt. Moreover, integration of data across model organisms is a cornerstone of modern biology, and without Xenbase, the enormous wealth of data from *Xenopus* would no longer be readily available to the larger biomedical community. Xenbase maintains the official gene nomenclature and ensures that it is in register with human and other organisms. As there are constant, weekly changes to gene/RNA/protein names and models, the reciprocal data sharing with other resources such as NCBI would quickly breakdown and *Xenopus* gene centric data would rapidly become out of date.

In order to maintain these essential services, renewed funding for Xenbase is an immediate top priority of the *Xenopus* research community. Xenbase has become a virtual focal point for the *Xenopus* community with over 1800 registered users. It hosts a website of community announcements, protocols, educational material and community resources. It has become clear that to keep up with technological advances and increasing pace of *Xenopus* research, the community needs Xenbase to develop new infrastructure, data pipelines, bioinformatics tools and annotation teams to support the exponential growth in single cell omics data and the increasing use of *Xenopus* phenotypes as a tool to model human diseases.

How should we proceed? With unanimous support from the community, Drs. Zorn and Vize are submitting a competitive renewal for Xenbase in early 2020. In order for this to be received effectively it is critical that the NIH re-issue a PAR such as the PA-08-180 “Resource Program Grants in Bioinformatics” that funded the previous Xenbase P41 grant and which specifically addresses model organism databases. This is essential because Resource initiatives such as Xenbase cannot be properly reviewed in study sections in comparison to hypothesis driven R01 applications.

The community also strongly encourages other NIH Institutes to participate along with NICHD in funding this widely used resource. An analysis of the NIH RePorter demonstrates that the NIH currently funds 238 grants that use *Xenopus* from 19 different Institutes. The top funders are: NIGMS (74), NICHD (37), NHLBI (23), NIDDK (16), NINDS (15), NIDCR (15), NCI (13), NEI (10), NIMH (10), NIAID (7), and OD (6).

The current plans for renewal of Xenbase, which are based on feedback of the community, are:

- Aim 1: Maintain and enhance Xenbase infrastructure, tools and curated content.
- Aim 2: Develop the next generation of Integrated Omics support, linking transcriptomics, epigenetics, proteomics, gene/protein regulatory networks, high content imaging and genotype-phenotype data at the single cell resolution.
- Aim 3: Develop support for *Xenopus* phenotype annotation and human disease modeling

Renewal of Xenbase is essential for all *Xenopus* research. Xenbase will accelerate the ability of investigators using this preeminent experimental system to discover fundamental new knowledge about the nature of biological systems and Xenbase will facilitate the application of that knowledge to improve human health.

2. Enhancement of efficient gene editing approaches and mutant lines

Impact: In the last 5 years the use of gene editing technologies, especially CRISPR-Cas, has become commonplace for the generation of loss-of-function mutations in almost every organism. In *Xenopus*, these technologies have been easily adapted as they only require microinjection into early embryos, which has been a mainstream technique in *Xenopus* for decades. Gene function can be initially assessed with bi-allelic gene targeting in F0 embryos and tadpoles but generating homozygous mutants requires more extensive breeding and additional space and equipment. Most *Xenopus* labs, however, are not experienced with breeding *Xenopus* and increased investment for infrastructure expansion requires additional funding, which would delay the creation of these mutant lines. The National *Xenopus* Resource (NXR), however, has the experience and infrastructure to generate these mutant lines and has been generating mutant lines for the community. This project has generated over 116 mutants, providing tremendous opportunities for the entire *Xenopus* community; however, the project is ongoing and requires continued funding to fully capitalize on mutant generation.

In addition to loss-of-function, CRISPR-Cas gene editing can be used for homology directed repair (HDR) for insertion of epitope tags or creating precise missense mutations. Nonetheless, to realize their full potential, these technologies need additional optimization in *Xenopus*. Several reports have demonstrated that insertion of small fragments is feasible, but enhanced reliability and efficiency is required. Improvements in knock-in technology in *Xenopus* would be extremely advantageous for cell biological and biochemical approaches that are common in *Xenopus*. The knock-in of tags would facilitate an extraordinary host of experiments including, but not limited to, FAC sorting of cells, live imaging of proteins without overexpression, biochemical assays (including ChIP) for which antibodies are not available, and precise mutation modeling of human disease. In addition, advances in knock-in technologies for manipulating cell lines will also be beneficial, especially with the recent creation of several new euploid cell lines in both *X. tropicalis* and *X. laevis*. Lastly, production of transgenic lines that express fluorescently labelled proteins at native concentrations or express Cas9 ubiquitously or in a tissue specific fashion would provide tremendous benefit to the community.

What is needed to enhance genome editing technologies in *Xenopus*?

Rapid and straightforward genome-editing technologies will transform the use of *Xenopus* in the biomedical research community by providing means to knock in/out genes of interest and to generate targeted mutations in genes/genomic regions to model human diseases. While the use of TALENs/CRISPRs are routine for the rapid generation of mutant F0 *Xenopus*, the methodologies need to be optimized and standardized for the generation of germline mutants in both *X. tropicalis* and *X. laevis*. In addition, methods to knock-in tags including fluorescent proteins need to be optimized and exploited. This will increase the use of mutant/transgenic lines in both fundamental biological and disease-related research. Approaches to be developed and/or refined include:

- Continued generation of germline mutants in both *X. laevis* and *X. tropicalis* as a community resource; these mutant lines have been requested *directly* by the community for immediate use.
- Generation of advanced tools for establishing knock-in strategies in *Xenopus*, including gene replacement, gene editing, targeted insertions, tagging proteins and creation of dominant active or repressive genes.
- Creation of new transgenic lines that express different Cas9 variants (e.g. dCas9, hfCas9) ubiquitously, especially in the egg.
- Establishment of tissue-specific transgenic lines expressing Cas9 allowing for generation of mutants in a cell type-specific fashion.

- Generation of new transgenic lines that label different cell types, signaling pathways and cellular structures.
- Development of new cell lines that express tagged fluorescent proteins for cell biological studies.

One of the difficulties faced by the *Xenopus* community in using mutant lines is the increased housing costs and housing space required to raise and breed the lines. This difficulty is currently being addressed by encouraging labs to interact directly with the NXR to create and analyze specific mutants through constant communication and short-term visits to the NXR. The NXR is also helping to coordinate the efforts to generate mutations in several hundred key genes that are studied by many investigators. Mutant sperm could easily be stored frozen and distributed to the community through the NXR. This would reduce duplication of effort and generate a resource similar to JAX for mice and ZIRC for fish, which have proven to be very successful. With this goal in mind, the NXR is in the ideal position to provide value for money for such an approach. Together the proposed generation, distribution, and data collection would be a modest, but highly productive, investment relative to the already substantial NIH investment. Another difficulty faced is expanding from *X. laevis* to *X. tropicalis*, since these animals require different housing and handling. This problem could be addressed via administrative supplements to existing *X. laevis* grants to cover the cost of new *X. tropicalis* housing systems.

Community recommendations: The *Xenopus* community identified the further development of genome-editing technologies as an Immediate Need.

- Continued generation of new mutant lines. The NXR has been generating mutants for the community over the last four years and recently submitted an R24 to continue the generation of new mutant lines. The NXR works closely with individual researchers and will continue to generate new mutant lines for the community.
- Establishment of new methods for HDR that would improve its efficiency in *Xenopus*. As new technologies continue to be developed, they must be tested and standardized in *Xenopus*.
- “Resource” RFAs/PARs of R01/R21/R24 level projects should be issued to generate advanced genome editing tools.
- Increased support for resource grants (R24) for the *Xenopus* community, because these projects have a broad and long-lasting impact on the entire biomedical community.
- Supplements to existing grants should be available to allow individual laboratories to utilize mutant lines in their research projects and/or to establish *X. tropicalis* colonies.

Conclusion: Generating and improving genetic resources for *Xenopus* is seen as an immediate priority by community. These new mutant and transgenic lines will provide a rich source of material that can be used by everyone but require appropriate investment. As gene editing technologies continue to rapidly evolve it is imperative that these are developed for *Xenopus*, which is an incredibly valuable model for biomedical research. Improving gene editing resources will help achieve the overarching goal of optimizing the utility of *Xenopus* to study human disease and improve human health.

Section IV: An Essential Resource needed by the *Xenopus* Community:

XenCAt: A single cell atlas of *Xenopus* development:

Impact: The genetic causes of human diseases are rapidly being identified thanks to a revolution in human personal genomics. True progress, however, requires further analysis of underlying developmental, cellular and molecular mechanisms, and establishment of predictive disease models to test therapeutic options. Ultimately, genes do not function in isolation; they are grouped spatially and temporally at multiple nested levels, the most salient functional unit being the single cell. Indeed, observing biological systems at the cellular level provides an unprecedented opportunity to define functional modularity and combinatorial interactions of genes in various physiological contexts. Many of these contexts are conserved in evolution, thus defining a baseline, deviations from which will produce malformations and disease. Accordingly, a Human Cell Atlas is already being built with the hope that it will form a core of this new single-cell perspective. Parallel work in model organisms will be crucial, and cell atlases are being constructed currently, for example, in mouse and zebrafish. Despite the widespread use of *Xenopus*, there are as yet no plans domestically or worldwide to establish a single cell atlas for this important model system.

Every NIH Institute funds *Xenopus*-based projects, all of which will benefit from the XenCAt. Thus, its impact is in many ways similar to the earlier impact of NIH funding for sequencing and building the *Xenopus* genome. By organizing information and coordinating efforts at multiple scales, XenCAt will enhance the value of the unique methods already available in *Xenopus*. Moreover, it will be critical in complementing other emerging approaches. For example, as genome edited *Xenopus* mutants are generated, they can be characterized in development and adult function at the single cell level. Moreover, though lagging behind single cell transcriptomics, the large cell size of amphibians has already made single cell proteomics possible in *Xenopus*, well ahead of other organisms. *Xenopus* is thus the natural choice for spearheading the inevitable - and ultimately necessary - shift from single cell genomics to single cell proteomics. Thus, data generated during the assembly of the XenCAt will begin to forge the links between single cell transcriptomics of today with the single cell proteomics of tomorrow, thereby opening new vistas of biological insight.

What is needed to create XenCAt:

Administrative

- Substantial initial investment to jump-start the Cell Atlas portal at Xenbase
- Connection and collaboration with other species efforts to bootstrap the *Xenopus* Atlas
- Organize workshops for training in single cell methods

Instrumentation and Technology

- Strategies to genetically engineer animals to facilitate cell lineage tracing
- Strategies to create 3D representation of tissues and anatomy using technology such as micro-CT
- Pilot effort in emerging in-situ sequencing technologies that link anatomy with gene expression

Protocols

- A set of best practices for dissociation of various organs, tissues and developmental stages
- Methods for barcoding specific tissues and developmental stages, and establishing standards for using the technology, including the numbers of cells required
- Methods for sequencing and standards for sequencing technology and parameters

Bioinformatics

- Reusable pipelines for mid-level bioinformatics of read alignment, reference set of sequences, including tools for troubleshooting / quality controlling the data
- A platform that federates the data from different projects and enables one to contrast and compare cells across experiments using state-of-the-art computational approaches
- Support for both *X. laevis* and *X. tropicalis* analysis, being aware of homeologues and cross-species gene homology. This should include support for mapping across the evolving annotation of genomes in related species for contrast data of the same cell type across related organisms (e.g., mouse to frog and fish to frog)
- Automated way to identify cell types from a curated set of *Xenopus* marker genes

Community recommendations: The *Xenopus* community identified the development of XenCAT as an Essential Resource that requires:

- Issue of RFAs/PARs for R01/R21 level projects to generate advanced methods for lineage tracing and in-situ sequencing tools in *Xenopus*, as well as 3D scanning/mapping tools and efforts.
- Support for specialized single cell transcriptomics workshops at which *Xenopus* community members would be trained in best practices of experiment design, dissociation, barcoding, sequencing and expression analysis, as well as high level bioinformatics tools.
- Support for resource grants (R24) for broad single cell profiling efforts in the *Xenopus* community, since such projects will generate knowledge that has a broad and long-lasting impact on the entire biomedical community.
- Supplements to existing grants to allow individual *Xenopus* laboratories specializing in particular organs or tissues or embryonic stages to rigorously characterize the normal biological state of their respective system of interest, following a standardized set of quality requirements and contributing the data towards comprehensive XenCAT warehousing.
- Establishment of a XenCAT working group/committee that includes leaders in the *Xenopus* community to further detail the priorities and coordinate and cross-pollinate the effort with the two existing *Xenopus* centers – Xenbase and NXR.

Conclusion: The Human Cell Atlas Project was initiated to harness the potential impact of understanding the combinatorial interactions of genes in various physiological contexts and better understand how deviations produce malformations and disease. Similar efforts have been initiated in mouse and fish in recognition of the importance of creating similar atlases for the major biomedical model systems. Since *Xenopus* is one of these key systems, the community expressed high enthusiasm to create a *Xenopus* Cell Atlas that can efficiently bootstrap to our mature community resources (Xenbase, NXR) and take advantage of novel methods in single cell sequencing that can be scaled up and used in a robust, reproducible fashion in *Xenopus*-specific projects. Thus, it is most timely to create a host of best practices and protocols, as well as crucial infrastructure, to ensure that *Xenopus* single cell profiling comes together into a unified, high-quality, long-lasting body of knowledge, rather than a dispersed set of miscellaneously obtained datasets that rapidly loose significance. Building such resource will dramatically improve the value of the current NIH funding for *Xenopus* by informing and enriching most of current research projects and will open new translational potential for using *Xenopus* to model human disease.

Section V: General Conclusions:

The unifying objective of the *Xenopus* research community is to accelerate the use of this animal as a model for understanding fundamental biology as well as human disease. From Sir John Gurdon's Nobel Prize winning discovery of nuclear reprogramming, through Tim Hunt's identification of the cyclins, to the myriad recent discoveries outlined in Section I, *Xenopus* has been and remains at the forefront of basic biology and biomedical research. With the proliferation of -omics datasets, an increased understanding of gene-environment interactions, and the increasing use of patient-specific genome sequencing for precision medicine, the data on gene variants related to human disease are increasing exponentially. The major challenge now is deciphering, validating, and functionally characterizing specific gene variants/mutations and the impact they have on disease. In many cases, the necessary experimentation cannot be accomplished in tissue culture. The recent advances in genome editing and high throughput analysis complement the already strong experimental repertoire of this model organism, and together, these assets make *Xenopus* an indispensable and cost-effective platform for the rapid identification, validation, and characterization of genes involved in human disease.

Accordingly, all of the initiatives presented in this White Paper will help achieve the overarching goal of optimizing the utility of *Xenopus* to study human disease and improve human health.

Section VI: Appendices

Appendix A – Authors of the 2019 *Xenopus* Community White Paper

John Wallingford (University of Texas at Austin)
Marko Horb (National *Xenopus* Resource, Marine Biological Laboratory)
Aaron Zorn (Xenbase, Cincinnati Children's Hospital)
Leonid Peshkin (Harvard Medical School)
Mustafa Khokha (Yale Medical School)

Appendix B – The Cold Spring Harbor Laboratory course on "Cell & Developmental Biology of *Xenopus*"

1993-1996: Co-Directors: Hazel Sive (MIT), Rob Grainger (UVa), Richard Harland (UC Berkeley)

1997-2000: Co-Directors: Paul Krieg (UT Austin), Sally Moody (GWU)

2001-2004: Co-Directors: Ken Cho (UC Irvine), Jan Christian (OHSU)

2005-2007: Co-Directors: Janet Heasman (CCHMC), Chris Wiley (CCHMC)

2008-2010: Co-Directors: Ray Keller (Univ. Virginia), Kristen Kroll (Washington Univ.)

2011-2014: Co-Directors: Amy Sater (Univ. Houston), Gerald Thomsen (Stony Brook Univ.)

2015-2017: Co-Directors: Karen Liu (King's College London), Mustafa Khokha (Yale Univ.)

2018-Present: Co-Directors: Chenbei Chang (University of Alabama), Lance Davidson (U. Pitt.)

Supported by NIH grants and by the Howard Hughes Medical Institute

Appendix C – The International *Xenopus* Conference

The goal of the International *Xenopus* Conference is to provide a venue for sharing of cutting edge research advances using *Xenopus* as a model system.

1984: Organizational workshop at Airlie House, Warrenton, Virginia, USA

1986: University of East Anglia, UK

1988: Tulane University, New Orleans, LA, USA

1990: Les Diablerets Conference Center, Switzerland

1992: Asilomar Conference Center, California, USA

1994: Congress Center De Branding, The Netherlands

1996: YMCA of the Rockies, Estes Park, Colorado, USA

1998: Sardinia, Italy

2000: YMCA of the Rockies, Estes Park, Colorado, USA

2002: Cambridge University, Cambridge, UK

2004: Marine Biology Laboratory, Woods Hole, MA, USA

2006: Kazusa Akademia Park, Japan

2008: Leiwen, Germany

2010: Lake Louise, Canada

2012: Giens Peninsula, France

2014: Asilomar Conference Center, California, USA

2016: Orthodox Academy of Crete, Chania, Greece

2018: University of Washington, Seattle

Supported by NIH R13 grants and the International *Xenopus* Board.

Appendix D – The *Xenopus* Resources and Emerging Technologies (*XRET*) Meeting

The goal of the *Xenopus* Resources and Emerging Technologies (*XRET*) Meeting is to provide an informal venue for discussions about new and upcoming technologies, establishing infrastructure priorities for the community, and allowing for more in depth discussions of individual research topics. These have been held in alternate years from the International *Xenopus* Conference at the Marine Biological Laboratory:

2011: Co-organizers: Mustafa Khokha (Yale) and Marko Horb (MBL)

2013: Co-organizers: Brian Mitchell (Northwestern) and Marko Horb (MBL)

2015: Co-organizers: Laura Lowery (Boston College), Brian Mitchell (Northwestern) and Marko Horb (MBL)

2017: Co-organizers: Andrea Wills (University of Washington), Matt Good (Univ. Pennsylvania) and Marko Horb (MBL)

2019: Co-organizers: Rachel Miller (Univ. Texas at Houston), Gregory Weber (Univ. Indianapolis), Helen Willsey (UCSF) and Marko Horb (MBL)

Supported by the National *Xenopus* Resource and the International *Xenopus* Board.