Justification for the Support of Echinobase and Associated Genome Resources: Enhancing the Impact of Echinoderm Model Systems for Biosciences

R. Andrew Cameron, California Institute of Technology
Charles Ettensohn, Carnegie Mellon University
Veronica Hinman, Carnegie Mellon University

1. PURPOSE

This document briefly outlines the current significance of echinoderm research models for understanding developmental mechanisms, with emphasis on their central importance in solving gene regulatory networks. It describes a set of genomic resources presently available to the echinoderm research community. Most importantly, it identifies the priorities of the research community concerning the future enhancement of these resources in support of core NIH missions.

2. IMPACT OF THE ECHINODERM COMMUNITY

In the late 1990s, the critical importance of the sea urchin as a research model spurred a sea urchin genome project. This NIH-funded project supported the sequencing and annotation of the genome of the purple sea urchin (*Strongylocentrotus purpuratus*) (Sea Urchin Sequencing Consortium, 2006a,b). Along with sequence information, many useful research materials such as macro-arrayed cDNA and genomic libraries were created. The growing popularity of other species of echinoderms as model systems, and the importance that comparative approaches have for illuminating genome and gene regulatory network (GRN) functions, have since prompted the inclusion of sequence data and library resources for several other echinoderm species. In addition, there have been continual improvements in the *S. purpuratus* genome assembly and gene annotations, and new research resources for this species have been produced (e.g., transcriptomes of many developmental stages and specific cell types). Today, a comprehensive web information system, Echinobase (http://echinobase.org), offers a fully indexed view of the sequences as well as annotations and gene expression data. In addition, a library resource provides materials to researchers.

(A) Echinoderms

Early Cambrian stem group echinoderms, identified by their calcite endoskeletons and water vascular systems, were survived by stem group lineages that gave rise to the five extant classes (Figure 1). One of the most striking aspects of echinoderm evolutionary history is that since the Ordovician, 495-440 million years (MY) ago, the characters that define the echinoderm classes have been completely static. Obviously their body plans work! As established by modern phylogenomics (Telford et al., 2014): (i) The echinoderms plus hemichordates are the

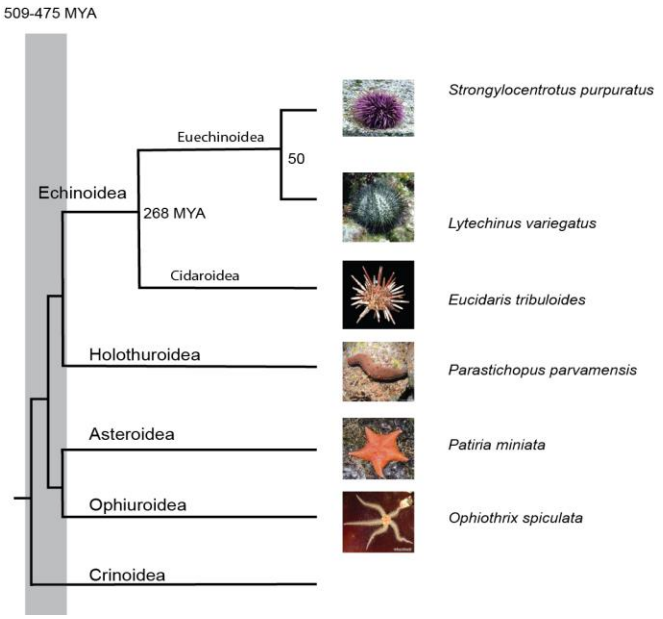

Fig. 1- Phylogenetic relationships, and times of divergence, of species of Echinoderms. The shaded region is the period of rapid divergence at the Cambrian boundary.

the sister clade to the chordates within deuterostomes; (ii) the primitively stalked crinoids are basal to the other four extant echinoderm classes, all of which are motile as adults; (iii) the sea stars and brittle stars are sister groups; (iv) the sea cucumbers and sea urchins are sister groups; (v) the sea urchins themselves diverged from common ancestors that last lived at the latest about 268 MY ago, i.e., 18 MY before the great end-Permian extinction, and which were likely already separated into the two extant subclasses (cidaroids, "pencil urchins", and euechinoids "modern sea urchins").

(B) The Echinoderm Research Community

The investigators who use sea urchins and other echinoderms as research models for development and cell biology are an active and intellectually important community of researchers. Currently, this community comprises about 150 investigators, as measured by the number of people who attend the sea urchin meeting (see below). There are about 50 laboratory directors on the current mailing list for this meeting. From a NIH Reporter search there are 12 active projects identified by the keywords: "sea urchin" and "FY2015". They reside in NICHD, NIGMS, NIEHS and OD institutes. These awards totaled $5.6 million for the one-year period. Likewise, NSF lists 42 active awards for "sea urchin" in cell and developmental biology programs, funded at $20.5 million. There are 158 active grants at NSF in all programs, funded at $67.7 million.

The echinoderm research community is a remarkably cooperative group. This collection of investigators has held an international meeting every 18 months for 35 years (since 1981) with only a rotating, ad hoc committee of organizers and no other official structure. Many investigators, graduate students, and post-doctoral scholars have spent time in echinoderm research laboratories other than their own for sabbaticals or short training experiences. Interdisciplinary work ranging from paleontology to molecular developmental biology is evident within the community. The exchange of unpublished sequence data and reagents occurs seamlessly among laboratories. This group of scientists exhibits "coopetitive" behavior, a mixture of competition and cooperation discussed in game theory and social organization

(Ghobadi and D'Ambra, 2012). In this context, the support of community resources returns much more than it costs while still supporting innovative efforts.

(C) Broader Impacts of the Echinoderm Community

Sea urchins have been an important research model for more than 150 years. Currently, studies with sea urchins and other echinoderms are making far-reaching contributions to many fields of biology. For example, a Google Scholar search reveals that the term, "sea urchin" has been used in the text of ~18,000 publications since 2011. Most significantly, echinoderms have from the beginning contributed uniquely to understanding the controlling role of the genome in the process of development. Still today, the best understood regulatory logic control processing is known for genes acting during sea urchin development. The strengths of this experimental model will continue to make it a preeminent system for the analysis of the genomic control of embryogenesis.

Research with echinoderms is currently impacting several major areas, including:

1. Gene Regulatory Networks: Presentation of the first detailed animal gene regulatory network model for development (Davidson et al., 2002) heralded the beginning of a new and ever more important research role for sea urchins. This work was tremendously augmented by the genome sequence announced in 2006 (Sea Urchin Genome Sequencing Consortium, 2006) and by the use of arrayed library resources. The sea urchin embryo is now the leading model system for analyzing the regulatory networks that underlie all animal development. Current work in this important field is expanding our understanding of the architecture of GRNs (Li et al., 2014; Andrikou et al., 2015), GRN evolution (Cheatle et al., 2014; Erkenbrack et al., 2015, linkages between GRNs and tissue morphogenesis (Rafiq et al., 2014; Saunders et al., 2014), the regulation of GRNs by intercellular signaling pathways (Cui et al., 2014; Sun and Ettensohn, 2014), and the potential utility of GRNs in re-engineering the process of embryogenesis (Damle and Davidson, 2012). This work has pioneered the use of systems and GRN approaches into many other models of biology (Dutkowski and Ideker, 2011; Sánchez Alvarado, 2012; Wilson et al., 2008; Zmasek and Godzik, 2013). As evidence of impact in this cardinal area, there are ~7700 Google Scholar citations using the terms "sea urchin gene regulatory network" since 2011.

Immunology: Sea urchins have potent non-adaptive immune systems which utilize hundreds of receptors of classes such as TLR receptors of which we have only a few members (Buckley and Rast, 2015). Elucidation of echinoderm immune function will not only produce startling insights but could well have major practical implications as the specificities of these receptors are determined. (Yue et al., 2014)

Cell Signaling. The genome sequence has also made possible important new insights regarding cell signaling processes in early development and their role in embryonic patterning (Halliot et al., 2015; Materna and Davidson, 2013; McIntyre et al, 2103; Sethi et al., 2012). Current research on detoxification biochemistry relies heavily on the sea urchin as a model (Bosnjak et al., 2013).

Both fertilization biology and cell cleavage processes continue to be informed by studies with echinoderms (Warner et al., 2014; Whittaker et al., 2006).

Neurobiology: In the realm of neurobiology, the genome sequence has uncovered a range of fascinating but little understood sensory organs and pathways (Yankura et al., 2013; Burke et al., 2014; Elphick, 2014).

Germ Cell Specification: The sea urchin has emerged as an important model for the study of germ cell specification by conditional mechanisms, work that has also been spurred by the availability of genomic data (Wessel et al., 2014).

Education and Outreach: The availability of sea urchin gametes and the ease of their manipulation has made the sea urchin a popular source of educational material for many years. There are two widely used and complementary educational web sites. “Sea Urchin Embryology” (http://web.stanford.edu/group/Urchin/) provides essential information concerning animal procurement and handling, gamete collection, and fertilization, as well as detailed protocols for simple wet-lab exercises related to fertilization and early development. “Virtual Urchin” (virtualurchin.stanford.edu) supports unique, interactive web-based educational modules related to sea urchin development, including a virtual lab bench for simulating complex experimental manipulations. “Embryology Experiment” kits are commercially available from Carolina Biological Supply Company and Gulf Specimen Marine Lab, attesting to the widespread use of sea urchin gametes and embryos as educational materials.

3. CURRENT RESOURCES THAT SUPPORT COMMUNITY ACTIVITIES

(A) Available Data for Echinoderm Research at NCBI

Number of search results for echinoderm terms in PubMed for 2015	631
Number of echinoderm nuclear genome projects registered at NCBI	7
Number of echinoderm transcriptome projects in the Short Read Archive at NCBI (34 species)	180

(B) Web Information System (Echinobase)

Number of species with assembled genomes in Echinobase	6
Number of species with skim genome sequencing in Echinobase	2
Number of bytes in the Echinobase web directories	316.0 Gb
Number of files in the Echinobase web directories	4,725,658
Number of bytes in 216 sequence download files	77.0 Gb

Since its creation, Echinobase (http://echinobase.org), the public portal to echinoderm genomic resources, has proved to be invaluable; for example, the remarkable progress made on developmental GRNs would not have been possible without it. Echinobase as it is now configured will serve as the foundation for future expansion in terms of additional data and new capabilities.

A total of 8 genome sequences for the echinoderms in various levels of draft assembly are housed at Echinobase (Table 1). There are assemblies from five sea urchin species (*S.*

purpuratus, Lytechinus variegatus, S. franciscanus, Allocentrotus fragilis and *Eucidaris tribuloides*) posted. Genomic data from a sea star (*Patiria miniata*), a sea cucumber (*Parastichopus parvamensis*) and a brittle star (*Ophiothrix spiculata*) are also available. Thus the genomics included in the web system samples the major evolutionary diversity extant in this phylum. It might be added that much recent work has shown that the ~50 MY divergence time separating Strongylocentrotus and Lytechinus turns out to be the "sweet spot" for identification of sequence patches conserved because of their cis-regulatory function (Cameron and Davidson, 2009).

Species	Status
Strongylocentrotus purpuratus v4.0	mature draft
Strongylocentrotus franciscanus	2x skim coverage
Allocentrotus fragilis	2x skim coverage
Lytechinus variegatus v2.2	improved draft
Eucidaris tribuloides v1.0	first draft
Patiria miniata v1.0	first draft
Parastichopus parvamensis v1.0	first draft
Ophiothrix spiculata v1.0	first draft

Table 1- Sequencing progress for echinoderm genomes. The eight genome projects begun at the Baylor College of Medicine, Human Genome Sequencing Center and the present stage of completion of each.

Transcriptomes for many embryonic stages and adult tissues of these and other species have been assembled into gene sets and accessioned into the database. Both gene annotation and Blast databases are available for these gene sets. Echinobase also includes a number of state-of-the-art bioinformatics resources and powerful search engines. Its JBrowse genome visualization tracks include transcript maps together with gene predictions, and for *S. purpuratus* the mounted gene models have been vastly improved by cross-reference analysis based on our dense developmental transcriptome studies (Tu et al, 2012; Tu et al, 2014). A continuing effort is made in Echinobase to improve genome curation by incorporation of new literature results, and great care has also been taken to update and improve the website which serves as the portal of Echinobase to the outside world. The broad range of functionalities currently incorporated into Echinobase is indicated in Fig. 2, a snapshot from the Navigation Guide on the website.

Echinobase is heavily used by many laboratories. Use can be measured by visits to the website, where a meaningful visit might be one in which the visitor stays to consult multiple website pages. By this criterion Echinobase receives about 4000 visits annually. Echinobase is used in many parts of the world (though mainly in the US); Fig. 3 is based on data retrieved from Google Analytics and filtered for users as stated above. But numbers of visits do not tell the whole story: for anyone who is working on genes or genomes of these model organisms, **the genomic database is an absolutely essential resource, and much of the most important research in the last few years in this field could not have been carried out without this resource**.

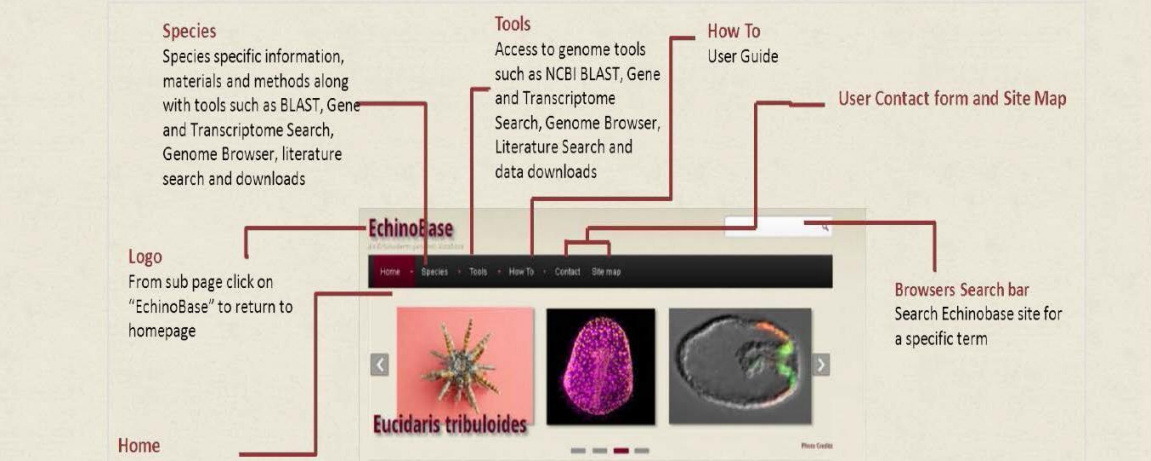

Fig. 2- Portals to various publicly accessible resources mounted in Echinobase.

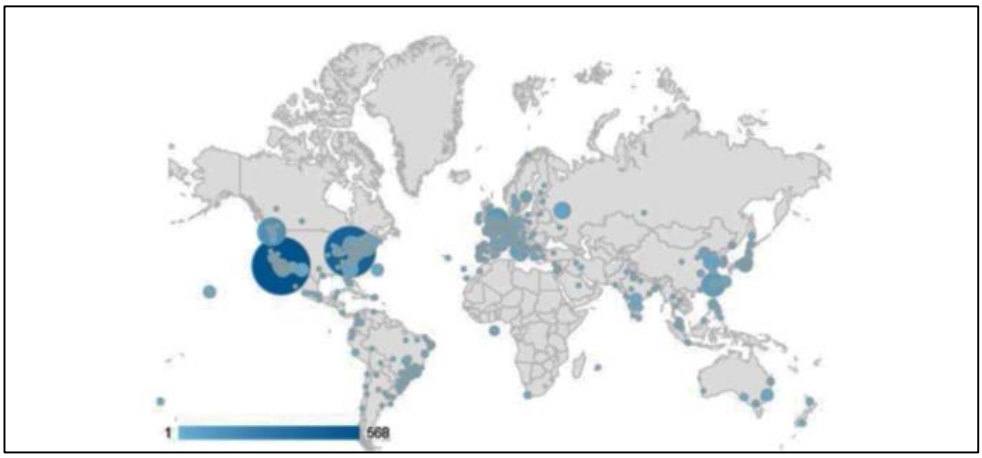

Fig. 3- Current world-wide use of Echinobase website; map shows hits in Taiwan, Japan, India, Australia, Canada, S. America, in addition to Europe and the U.S. (Data from Google Analytics.)

(C) Gene Library Resource

Number of macro-arrayed BAC libraries in the resources (7 species)	11
Number of cDNA libraries (2 species)	2
Number of identified BAC clones (7 species)	310
Number of S. purpuratus recombineered ABC clones	128
Number of recombineered BAC clones, other species	29

A major product of the Resource has been genomic BAC libraries and cDNA libraries arrayed in 384-well plates and spotted on nylon macro-array filters (Cameron et al., 2000; Rast *et al.*, 2000). The filters are approximately 22x22 cm and contain a total of 18,432 duplicated spots arranged in uniquely identifiable, geometrically predefined patterns in 4 x 4 blocks (Maier *et al.*,

1994.). The genomic BAC libraries are constructed in a common BAC vector at an average insert size of 140 Kb. Therefore, 100,000 clone libraries contain about 13x coverage of the 800 Mb sea urchin genome. We have in our arsenal both genomic and cDNA libraries from several species. In most cases we also have purified genomic DNA from the same animal used to make the genomic library. This latter feature obviates some of the problems associated with the highly polymorphic genomes of these animals, e.g., in primer design.

(D) BAC Recombineering

This has become an increasingly important focus of the Gene Library Resource. After much technological re-tuning, it is now possible to generate 10-20 recombinant BACs in 1-2 days bearing deletions, insertions or mutations, which is faster than they can be designed. In this manner, there have been generated over 100 recombineered BACs that bear regulatory genes expressed during sea urchin development, in which a gene encoding a fluorophore has been inserted shortly after the ATG. Since these BACs include the natural cis-regulatory systems controlling expression of the regulatory gene *in vivo*, when injected into sea urchin embryos they drive expression of the fluorophore accurately with respect to the parent regulatory gene (without affecting function of the latter). They are enormously useful: for cis-regulatory analyses, including a means of establishing the respective roles of diverse cis-regulatory modules; for isolation of cells expressing given regulatory states or cell types; as lineage markers; and for construction of re-engineering vectors. Recombineered BACs have been provided to many laboratories for use in their specific research projects.

4. PROPOSAL FOR ENHANCEMENT OF ECHINOBASE

At the recent Sea Urchin Developmental Biology XXIII meeting (October, 2015), a public forum was held to discuss the needs of the research community with respect to the genome website and other resources. Approximately 40 people attended the meeting. The group unanimously supported the continuation and expansion of Echinobase and the Gene Library Resource, as well as the development of additional research resources.

Based on this public discussion, the Scientific Advisory Board considers support for the following activities to be of high priority for maintaining and enhancing the research efforts of the community. **These recommendations address critically important needs identified by the community, seek to make best use of current resources, and are directed at enhancing the unique strengths of the echinoderm model system for the coming decade.**

Echinobase is a critically important source of genomic information for the Echinoderm Phylum. Without it, most of the important research resources that have been developed over the past decade (including the genome sequence itself) would be almost useless. The continual improvement of this vital resource is therefore of the highest priority. We propose the following enhancements to Echinobase:

1. To improve accessibility and ease of use of the web resource. As the types (e.g. annotation and expression) and the quantity of data expand, it becomes imperative to remodel the Echinobase web information system to ensure that it remains easily accessible to researchers, regardless of their experience with echinoderms. This will include use of uniform nomenclature, and searching tools, as well as intuitive links to resources and databases. Efforts will be made to seek input from researchers in other systems, and in particular other genomic web resource developers and to provide outreach efforts to service a wide community. The goal of this aim is to increase the impact of Echinobase and ensure that researchers from other communities can take advantage of the work done in echinoderms.

2. To improve genome assemblies. The genomes of echinoderms are large and polymorphic. Indeed, efforts to sequence and assemble them have often served as experiments for this kind of effort in general (English et al, 2012). A summary shown in Table 2 below indicates the state of the various genomes in our resource and the efforts needed to improve them to a state where they are useful for comparative genomics and the work of the echinoderm research community. The goal of this aim is to improve the assembly of these genomes and also to include new pages, for additional species of importance (e.g. crinoids) that might be generated by individual lab sequencing efforts.

3. To improve transcriptome consolidation and annotation. The number of user-generated echinoderm transcriptomes is ballooning. There are 180 echinoderm transcriptomes in Genbank for 34 species. Most of these were collected for an explicit experimental purpose and no consolidation has been undertaken. Thus a huge amount of data is lost to the experimentalist. This aim will collate transcriptomes from these many sources to provide high quality reference transcriptomes from multiple species, time points and tissue types. Since the utility of these many sequencing endeavors is to facilitate comparative genomics and GRN structures, it will be important to employ a standardized gene nomenclature so that researchers can effectively identify genes and gene annotations from many species through Echinobase. Efforts should be made to generally improve annotations; e.g., splicing isoforms, non-coding RNAs, translation starts sites, and UTRs. This will globally improve the utility of Echinobase for all researchers. This aim will also incorporate wiki approaches to allow registered members to annotate entries.

	Current Status			Proposed Status		
Genome	Contig N50 (kb)	Scaffold N50 (kb)	Sequencing	Additional Sequencing	Target Assembly Level	Functionality
S. purpuratus (v4.0)	17.6	431	40x Illumina 11x PacBio 18x SOLiD 8.3x Sanger	None	No change	-
P. miniata (v1.0)	9.5	52.6	70x Illumina 15x Roche 454	40x Illumina "rainbow" libraries 10x PacBio	100-200kb Scaffold N50	Improved gene prediction; Improved identification of regulatory ncDNA elements (e.g. ATAC-seq); Facilitate generation of molecular reagents
L. variegatus (v2.0)	9.7	46	21x Illumina 23x Roche 454 13x PacBio	40x Illumina "rainbow" libraries	100-200kb Scaffold N50	Improved gene prediction; Improved identification of regulatory ncDNA elements (e.g. ATAC-seq); Facilitate generation of molecular reagents
E. tribuloides	2.8	28.2	23x Illumina 23x Roche 454	10x PacBio	10kb contig N50	Improved gene prediction; Facilitate generation of accurate molecular reagents (in situ probes, PCR primers)
P. parvamensis	7.1	40	~140x Illumina	10x PacBio	10kb contig N50	Improved gene prediction; Facilitate generation of accurate molecular reagents (in situ probes, PCR primers)
O. spiculata	4.5	43	~160x Illumina	10x PacBio	10kb contig N50	Improved gene prediction; Facilitate generation of accurate molecular reagents (in situ probes, PCR primers)

Table 2- Current and proposed genome assembly and annotation statics for flagship species of Echinoderms.

4. To incorporate functional genomic data into Echinobase. New endeavors to include spatial expression should be included for *S. purpuratus* and other important experimental species. Significant individual lab efforts are directed at identifying spatial and quantitative gene expression profiles that can benefit the community as a whole. Perturbation data that have been generated for the sea urchin GRN will be also be included. Providing these data in a format that can readily be accessed and cross-linked from multiple species will aid comprehensive syntheses of gene regulatory network analyses from multiple species, including for researchers from other communities. These should, as much as possible, also follow standards for other models systems outside of the echinoderms to improve broader accessibility. Controlled vocabularies for developmental anatomy and developmental stages should be developed with the intent of coordinating with other taxa.

5. To provide options to include post genome-wide scans of features of a regulatory nature including ATAC-seq data for all species. Echinoderms are famous for the ease with synchronous embryo cultures can be obtained, suiting them perfectly for developmental profiling of chromatin architecture. Most importantly, echinoderm embryos are unusually well suited for functional cis-regulatory analyses of gene expression, an essential component of GRN studies. Such data are emerging from many labs and it will be crucial to establish tracks or other types of assemblies of these data onto Echinobase. This will greatly facilitate improved annotations of functional noncoding DNA and the use of echinoderms for regulatory functional genomics. Such data are also routinely needed by researchers from other models systems and in particular the

growing body of researchers performing comparative functional genomics that would like to use this major phylum in their analyses.

6. To continue to improve and extend gene annotation including NCBI data mining and data sources including literature. Efforts should be made assemble searchable databases of published echinoderm literature and to use these to improve genome annotation. This will dually increase the impact of this work as well as improve all types of annotations of genomic datasets.
.

5. PROPOSAL FOR ENHANCEMENT OF OTHER RESOURCES

The central goal of enhancements in research capabilities over the next 10 years should be to further exploit the unmatched utility of echinoderms as experimental models for dissecting the organization and function of developmental GRNs. To that end, we propose the following enhancements to the research resources of the community:

1. To develop improved methods for high throughput cis-regulatory analysis based on multiplexed BAC recombineering. Cis-regulatory analysis is "where the rubber meets the road" in regulatory molecular biology, but it has always been laborious and remains rate limiting. We should continue to develop new strategies for high throughput analysis of cis-regulatory modules.

2. To generate a community-wide resource of recombineered BACs. It would be extremely valuable to have publicly available a collection of recombineered reporter BACs that includes every regulatory gene expressed in sea urchins to the end of embryogenesis (~300 genes). These BACs would be useful for cis-regulatory analyses, for the isolation of cells expressing given regulatory states or cell types, as lineage markers; and for construction of vectors for the re-engineering of GRNs. By constructing versions of these BACs that are tagged with different fluorophores, it would be possible to visualize the expression of multiple reporters simultaneously in a single embryo.

3. To develop approaches for conditional gene perturbations in echinoderms. Gene perturbations are largely limited to injection of morpholinos (MOs) or dominant negative constructs into fertilized eggs. Conditional gene knockdown/knockouts would be a major technological advance and would allow the direct interrogation of network circuitry at late developmental stages and in targeted tissues. At least two strategies should be explored: a) the use of improved photoactivatable morpholinos (MOs) (Griepenburg et al., 2015) and b) the development of recombineered CRISPR/Cas9 BAC vectors that can direct genome editing in specific tissues and/or at specific stages, as directed by the cis-regulatory apparatus controlling Cas9 expression. It is important to note that, despite recent and well-publicized concerns regarding MOs, more recent studies have dispelled the most significant of these criticisms and there is widespread agreement that MOs are a indispensible tool when used with appropriate controls (Blum et al., 2015; Rossi et al., 2015). CRISPR/Cas9-mediated gene editing has recently been shown to be an effective tool in sea urchins (Lin and Su, 2015).

4. To extend the utility of high throughput, digital gene expression analysis. Quantitative assessment of gene expression is an essential aspect of the experimental analysis of GRNs. Although transcriptome profiling by RNA-seq is now routine, whole genome coverage is not required for many experiments and is cost prohibitive for studies involving many time points or many experimental conditions. Currently, the most cost effective approach to quantitative analysis of gene expression in large numbers of samples and with coverage of tens to hundreds of genes (i.e., exactly the scale of echinoderm GRNs) is nCounter analysis. The Davidson lab pioneered the use of this technology in GRN analysis, but the community will soon lose access to the nCounter at Caltech. An nCounter is currently housed in the Department of Biological Sciences at Carnegie Mellon, and we propose to expand community access to this unique technology. We will design and make available at reduced cost several nCounter probe sets that will allow rapid, digital analysis of the expression of all regulatory genes expressed during development, cell-type specific GRNs, and key signaling pathways, allowing researchers for the first time to quantitatively interrogate these gene sets under a wide variety of experimental conditions.

5. To generate cell-type specific chromatin accessibility maps to aid in regulatory network biology. ATAC-seq affords a means of mapping chromatin accessibility across the genome using relatively small numbers of cells; i.e., an order of magnitude fewer than required for DNase-seq or FAIRE. We should exploit methods for FACS-based isolation of cells expressing BAC reporters (Barsi et al., 2015) to generate genome-wide chromatin maps of specific cell types. The library of BAC reporters for all embryonically expressed transcription factors (Recommendation #2, above) will provide unparalleled access to any cell type in the embryo at any regulatory state.

6. To develop new capabilities for animal husbandry. Developmental research with echinoderms is affected by the seasonal availability of gravid adults. In addition, to exploit most effectively the genetic manipulations now possible with CRISPR/Cas9 technology and make these accessible to the community, the establishment of a limited number of transgenic strains will be required. Efforts should be made to improve methods for 1) long-term holding of gravid adults, effectively extending their reproductive season, and 2) egg-to-egg culture methods, opening the door to the production of transgenic strains. Significantly, these approaches should be developed for all species for which genomic data are available.

6. REFERENCES

Andrikou, C., Pai, C. Y., Su, Y. H., Arnone, M. I. (2015) Logics and properties of a genetic regulatory program that drives embryonic muscle development in an echinoderm. Elife 28: 4.

Barsi, J. C., Tu, Q., Calestani, C., Davidson, E. H. (2015) Genome-wide assessment of differential effector gene use in embryogenesis. Development 142: 3892-3901.

Bosnjak, I., Pleic, I. L.; Borra, M., Mladineo, I. (2013) Quantification and in situ localization of abcb1 and abcc9 genes in toxicant-exposed sea urchin embryos. Environmental Science and Pollution Research 20**:** 8600-8611.

Blum, M., De Robertis, E. M., Wallingford, J. B., Niehrs, C. (2015) Morpholinos: antisense and sensibility. Dev. Cell. 35:145-149.

Buckley, K. M., Rast, J. P. (2015) Diversity of animal immune receptors and the origins of recognition complexity in the deuterostomes. Develop. Comp. Immunol. 49**:** 179-189.

Burke, R. D., Moller, D. J., Krupke, O. A., Taylor, V. J. (2014) Sea urchin neural development and the metazoan paradigm of neurogenesis. Genesis 52**:** 208-221.

Cameron, R.A., Mahairas, G., Rast, J.P., Martinez, P. et al. (2000) A sea urchin genome project: sequence scan, virtual map, and additional resources. *PNAS*, 97:9514-9518.

Cameron R. A. & Davidson E. H. (2009) Flexibility of transcription factor target site position in conserved cis-regulatory modules. Dev. Biol. 336:122-135

Cheatle Jarvela, A. M., Brubaker, L., Vedenko, A, Gupta, A., Armitage, B. A., Bulyk, M. L., Hinman, V. F. (2014) Modular evolution of DNA-binding preference of. a Tbrain transcription factor provides a mechanism for modifying gene regulatory networks. Mol. Biol. Evol.. 31: 2672-2688.

Cui, M., Siriwon, N., Li, E., Davidson, E. H., Peter, I. S. (2014) Specific functions of the Wnt signaling system in gene regulatory networks throughout the early sea urchin embryo. Proc. Natl. Acad. Sci. U. S. A. 111: E5029-38.

Damle, S. S., Davidson, E. H. (2012) Synthetic in vivo validation of gene network circuitry. Proc. Natl. Acad. Sci. U. S. A. 109: 1548-53.

Dutkowski, J., and Ideker, T. (2011). Protein networks as logic functions in development and cancer. PLoS Comp. Biol. *7*, e1002180.

Elphick, M. R. 2014. SALMFamide salmagundi: The biology of a neuropeptide family in echinoderms. Gen. Comp. Endocrin. 205**:** 23-35.

English, A. C., Richards, S., Han, Y., Wang, M., Vee, V., Qu, J., Qin, X., Muzny, D. M., Reid, J. G., Worley, K. C., Richard A. Gibbs (2012) Mind the gap: upgrading genomes with Pacific Biosciences RS long-read sequencing technology. PLoS One 7: e47768.

Ettensohn, C. A. (2013) Encoding anatomy: developmental gene regulatory networks and morphogenesis. Genesis 51: 383-409.

Ghobadi, S., D'Ambra, J. (2012) Coopetitive relationships in cross-functional software development teams: How to model and measure? Journal of Systems and Software 85: 1096-1104.

Griepenburg, J. C., Rapp, T. L., Carroll, P. J., Eberwine, J., Dmochowski, I. J. (2015) Ruthenium-caged antisense morpholinos for regulating gene expression in zebrafish embryos. Chem. Sci. 6(4):2342-2346.

Haillot, E., Molina, M. D., Lapraz, F., Lepage, T. (2015) The maternal Maverick/GDF15-like TGF-β ligand Panda directs dorsal-ventral axis formation by restricting Nodal expression in the sea urchin embryo. PLoS Biol. 13: e1002247.

Li, E., Cui, M., Peter, I. S., Davidson, E. H. (2014) Encoding regulatory state boundaries in the pregastrular oral ectoderm of the sea urchin embryo. Proc. Natl. Acad. Sci. U. S. A. 111: E906-13.

Lin, C.-Y. Su, Y.-H. (2105) Genome editing in sea urchin embryos using a CRISPR/Cas9 system. Dev. Biol., in press.

Maier, E., Meier-Ewert, S., Ahmadi, A. R., Curtis, J., Lehrach, H. 1994. Application of robotic technology to automated sequence fingerprint analysis by oligonucleotide hybridization. J. Biotechnology 35, 191-203.

Materna, S. C., Davidson, E. H. 2012. A comprehensive analysis of Delta signaling in pre-gastrular sea urchin embryos Dev. Biol. 364**:** 77-87.

McIntyre, D. C., Seay, N. W, Croce, J. C and. McClay, D. R. 2013. Short-range Wnt5 signaling initiates specification of sea urchin posterior ectoderm. Development 140, 4881-4889.

Rafiq, K., Shashikant, T., McManus, C. J., Ettensohn, C. A. 2014. Genome-wide analysis of the skeletogenic gene regulatory network of sea urchins. Development. 141:950-61.

Rast, J. P., Amore, G., Calestani, C., Livi, C. B., Ransick, A., Davidson, E.H. 2000. Recovery of developmentally defined gene sets from high-density cDNA macroarrays. Dev. Biol. 228: 270-286.

Rossi, A., Kontarakis, Z., Gerri, C., Nolte, H., Hölper, S., Krüger, M., Stainier, D. Y. 2015. Genetic compensation induced by deleterious mutations but not gene knockdowns. Nature 524:230-3.

Sánchez Alvarado, A. (2012). Cellular hyperproliferation and cancer as evolutionary variables. Curr. Biol. 22: R772–R778.

Saunders, L. R., McClay, D. R. (2014) Sub-circuits of a gene regulatory network control a developmental epithelial-mesenchymal transition. Development 141: 1503-13.

Sea Urchin Genome Sequencing Consortium. (2006a) The genome of the sea urchin *Strongylocentrotus purpuratus*. Science 314: 941-952.

Sea Urchin Genome Sequencing Consortium. (2006b) Dev. Biol. 300(1) (special issue devoted to the sea urchin genome).

Sethi, A. J., Wikramanayake, R. M., Angerer, R. C., Range, R. C., Angerer, L. M. (2012) Sequential signaling crosstalk regulates endomesoderm segregation in sea urchin embryos. Science 335**:** 590-593.

Sun. Z, Ettensohn, C. A. (2014) Signal-dependent regulation of the sea urchin skeletogenic gene regulatory network. Gene Expr. Patterns 16: 93-103.

Telford, M. J., Lowe, C. J., Cameron, C. B., Ortega-Martinez, O., Aronowicz, J., Oliveri P., and. Copley. R. R. 2014. Phylogenomic analysis of echinoderm class relationships supports Asterozoa. Proc R. Sci. B 281: 20140479.

Tu, Q., Cameron, R. A., & Davidson, E.H. 2014. Quantitative developmental transcriptomes of the sea urchin *Strongylocentrotus purpuratus*. Dev. Bio.l 385:160-167.

Tu Q., Cameron R. A., Worley K. C., Gibbs, R. A., & Davidson, E. H. 2012. Gene structure in the sea urchin *Strongylocentrotus purpuratus* based on transcriptome analysis. Genome Res 22:2079-2087.

Warner, J. F., McCarthy, A. M., Morris, R. L., McClay, D. R. 2014. Hedgehog signaling requires motile cilia in the sea urchin. Mol. Biol. Evol. 31: 18-22.

Wessel, G. M., Brayboy, L,. Fresques, T., Gustafson, E. A., Oulhen, N., Ramos, I., Reich, A., Swartz, S. Z., Yajima, M., and Zazueta, V. (2014) The biology of the germ line in echinoderms. Mol Reprod. Dev. 81: 679-711.

Whittaker, C.A., Bergeron, K., Whittle, J., Brandhorst, B.P., Burke, R.D. and Hynes, R.O. (2006) The echinoderm adhesome. Developmental Biology 300: 252-266.

Wilson, M.D., Barbosa-Morais, N.L., Schmidt, D., Conboy, C.M., Vanes, L., Tybulewicz, V.L.J., Fisher, E.M.C., Tavaré, S., and Odom, D.T. (2008). Species-specific transcription in mice carrying human chromosome 21. Science 322, 434–438.

Yankura, K.A., Koechlein, C.S., Cryan, A.F., Cheatle, A., and Hinman, V.F. (2013). Gene regulatory network for neurogenesis in a sea star embryo connects broad neural specification and localized patterning. Proc. Natl. Acad. Sci. U. S. A. 110, 8591–8596.

Yue, F., Zhou, Z., Wang, L., Wang, M., and Song, L. (2014). A conserved zinc finger transcription factor GATA involving in the hemocyte production of scallop Chlamys farreri. Fish Shellfish Immunol. 39, 125–135.

Zmasek, C.M., and Godzik, A. (2013). Evolution of the animal apoptosis network. Cold Spring Harb. Perspect. Biol. 5, a008649.