ARGUMENTS FOR SEQUENCING THE GENOME OF THE SEA URCHIN *STRONGYLOCENTROTUS PURPURATUS*

Eric H. Davidson (California Institute of Technology)
R. Andrew Cameron (California Institute of Technology)

FOR THE SEA URCHIN GENOME ADVISORY GROUP:

Robert C. Angerer (University of Rochester)
Lynne Angerer (University of Rochester)
Roy J. Britten (California Institute of Technology)
James A. Coffman (Stowers Institute for Medical Research)
William H. Klein (M. D. Anderson Cancer Center)
Donal Manahan (University of Southern California)
David R. McClay (Duke University)
Jonathan P. Rast (California Institute of Technology)
Victor D. Vacquier (Scripps Institute of Oceanography)

FOR BCM, HUMAN GENOME SEQUENCING CENTER:

Richard A. Gibbs, Director
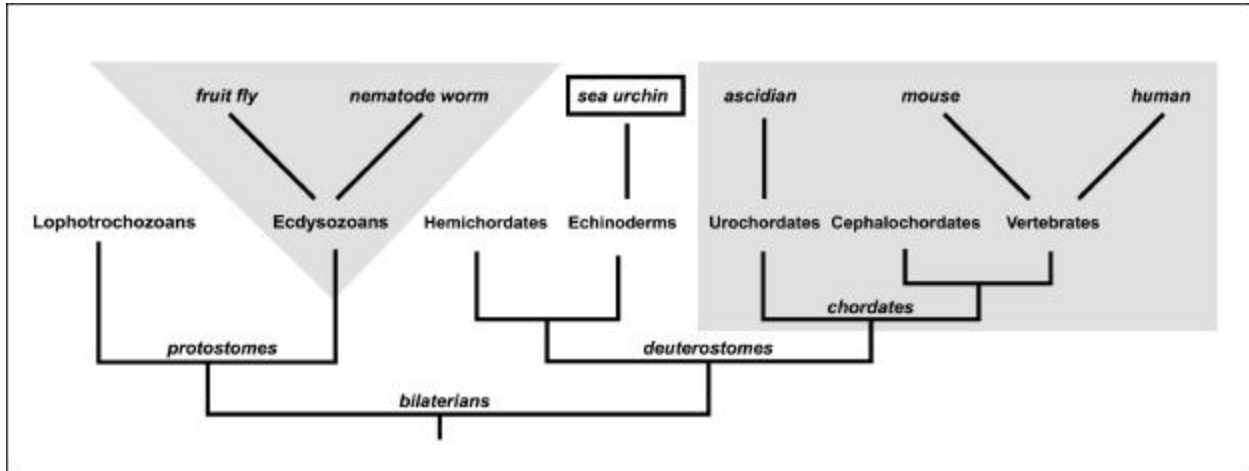**George Weinstock, Co-Director**

RATIONALE

1.   Introduction

There are reasons to sequence the genomes of many different animals: every animal genome holds secrets that when unlocked will yield invaluable mechanistic information that will in some measure illuminate not only our own biology but that of the rest of the world as well. However, in the case of most possible candidate genomes, it will be impossible to cash in on the potential value of the sequence without greatly augmenting scientific efforts on the species sequenced. Experimental systems that do not now exist or only barely exist in a few labs will have to be developed; comparative explorations of interspecific differences across phylogenetic distances that are often unclear will be required; and so will accumulation of repertoires of new molecular and genomic data. In contrast, to cash in on the value of a sea urchin genome sequence will require only the availability of the sequence: it will immediately be utilized by a large assemblage of active laboratories. The wise decision made to sequence the *Drosophila* and *C. elegan*s genomes was based on the intense research use of these organisms and the large store of knowledge already available. Of all remaining invertebrate genomes the same is most pointedly the case for a sea urchin genome. *En passant*, one might note that by comparison there is but a tiny handful of labs focusing on ascidians, although not one but two ascidian genomes are now being sequenced. The sea urchin research community is possibly 20-40 times as large as the ascidian community.

As detailed below, we have 75 letters from scientists in the US and elsewhere in support of an effort to sequence the genome of the sea urchin *Strongylocentrotus purpuratus*, so the opinions summarized in this document are not merely ours alone. *S. purpuratus* is the most widely used of the several sea urchin species in use the world around, particularly for the kinds

of molecular biology research that will be most immediately impacted by availability of genomic sequence. There are reciprocal ways to look at the effect of an *S. purpuratus* genomic sequencing project, and both are true: the large amount of funds and effort expended on sea urchin research will be leveraged enormously by the availability of genomic sequence; and equally, the value of the sequencing effort will be leveraged enormously by the already extant commitment of research effort to this model system.

An important aspect of the sea urchin model system is the phylogenetic position of these animals relative to ourselves. Figure 1 shows the relative positions of echinoderms, chordates, *C. elegans* and *Drosophila* in animal phylogeny; i.e., of echinoderms, relative to the four genera of animals from which genomes are so far sequenced.

These are all either chordates or ecdysozoans. The echinoderms and the hemichordates (a little known, though very interesting group of marine worms) are sister groups. Echinoderms and hemichordates are the only other living animal forms besides the chordates in the deuterostome subgroup of the Animal Kingdom. In other words the chordates (including us) share a common ancestor with echinoderms (including sea urchins). Therefore sea urchins are more closely related to all other deuterostomes (including us) than is any deuterostome to any other animal (e.g., flies or worms). So from the standpoint of phylogenetic position the sea urchin genome would provide an invaluable outgroup for assessment of what is ancient in the regulatory architecture and functionality of our own genome, what is chordate-specific, and the origin of that which has been modified in evolution. Furthermore there remain some deep functional mysteries about genomes that can only be solved by comparative examination across great distances. For example, we noticed molecular linkages of a number of genes in the *S. purpuratus* genome that are also linked in the mammalian MHC complex on a megabase scale, even though sea urchins lack MHC genes: there must be a functional meaning to the preservation of this system of linked genes over this huge evolutionary distance, but what is it? Among the nonchordate deuterostomes the sea urchins are obviously the primary target for investment in genomics since they are the only nonchordate deuterostomes that serve as major current research models.

The detailed reasons why the *S. purpuratus* genome should be sequenced are ultimately the same reasons that people work on this and other closely-related sea urchin species. In the following we summarize the size and activity of the sea urchin research community, and the major uses of this model system in central areas of biology, *viz* gene regulation molecular biology; the cell biology and biochemistry of eggs, embryos, and the fertilization process; and evolutionary biology. We also briefly review its medical relevance, and then propose a genomic

sequencing strategy, in collaboration with our partners in this enterprise, Richard Gibbs and George Weinstock of the Human Genome Sequencing Center, Baylor School of Medicine.

2.   The Sea Urchin Research Community

A minimum criterion of accountability in respect to the expenditure of effort and public resources needed to sequence an animal genome might be that there is a sufficient research community to properly and immediately utilize the sequence.

Table 1 provides an overview of the dimensions of the sea urchin research community. There are scores of sea urchin labs in the US, and sea urchins are also the research models in use in a number of long-standing labs in Italy, Japan, as well as other countries. Hundreds of papers are published each year describing results of research on sea urchins, and the US spends millions of dollars supporting this research each year. A small sample of the papers published in prominent scientific journals over the last two years is included at the end of this document as a bibliography: these publications display the role of this model system in many areas of biology. The sea urchin model system is most intensely used for research in gene regulatory molecular biology, molecular embryology, fertilization biology, cell biology, and evolutionary biology, but it also used for many other purposes such as marine population genetics, toxicology, nonadaptive immune system biology, and so forth.

Table 1. Parameters of the Sea Urchin Research Community
Average number of attendees at the sea urchin meeting: **150** (steady for the last several years).
The number of laboratories worldwide that use sea urchins as a primary research organism: **143**.
From the ISI database searched on the term "sea urchin", papers retrieved for 2000, 2001, 2002: **827**.
Total grant dollars for sea urchin-based research at the NIH for the last fiscal year: **$15M**

The sea urchin research community is united and enthusiastic about the need for *S. purpuratus* genomic sequence. As direct evidence for this statement, in Appendix 1 we have excerpted some of the letters in our files from scientists whose main research model is the sea urchin. Their names and affiliations of the authors of the 75 letters of support are also listed in Appendix 1. It is obvious to one and all that in every area of mechanistic bioscience genomic sequence is an invaluable resource, whether it be for study of gene function, or regulation thereof.

The sea urchin research community has held a major meeting focussed on sea urchin developmental cell and molecular biology, that lasts for about five days every year and a half since 1981. At present this meeting is attended by about 150 people. The issue of sea urchin genomics has been a prominent topic of discussion at the recent Sea Urchin Meetings, and the Sea Urchin Genome Advisory Group was set up in consequence. The following proposal takes into account the major needs of the community. These needs are focused directly on the requirement for genomic sequence, since there are already extensive sea urchin EST data, BAC libraries for a number of sea urchin species, arrayed cDNA libraries and other preparatory genomics researches, detailed below.

Uses of Sea Urchin Embryos for Gene Regulation Molecular Biology

The area in which sea urchin research has had its major impact on the state of knowledge in gene expression and gene regulation in development. This has been true for a long time: maternal mRNA was discovered in sea urchin eggs; the first measurements that established the complexity and prevalence distribution of mRNAs in any embryo were carried out on sea urchin

embryos; the first measurements of transcription rates and average and specific mRNA turnover rates as well as of protein synthesis rates in embryos were carried out on sea urchin embryos. This was all between about 1965 and the early 1980's. These foundations of sea urchin embryo molecular biology were in turn built on the century-long earlier history of experimental work on sea urchin embryos. For example, pronuclear fusion at fertilization was first recognized in sea urchin eggs in 1879 (by Fol), and the first realization that continuing gene expression is required for embryogenesis to occur followed from experiments of Boveri carried out in 1904.

Since the mid-1980's there has been a great focus of attention in the world of sea urchin developmental biology on the regulatory molecules that drive embryonic development: transcription factors and signaling components. Because the sea urchin embryo can literally be experimentally disassembled and reassembled, because the early cell lineage is invariant and embryonic cell interactions and cell fates have been so well worked out, the molecular biology of the embryo is uncommonly well integrated with knowledge of the process of embryogenesis. In consequence, our level of understanding of early development in this embryo has achieved a paradigmatic state.

There are four technological advances that place the sea urchin embryo at the forefront of developmental regulatory genomics.

•A relatively very high throughput gene transfer system. Thousands of eggs can be injected with expression constructs in a few hours, and the transcriptional readout obtained in spatial and/or quantitative terms within a day or two.

•Technology for obtaining stable nuclear extract. Enormous numbers of sea urchin embryos are available (literally trillions of embryo nuclei are extracted each year). These extracts are used for purification and microsequencing of transcription factors, given only a known genomic DNA target site.

•Use of morpholino substituted antisense oligonucleotides, and of mRNA encoding Engrailed domain fusions. These perturbation reagents permit shut down of any specific transcriptional regulatory process at will.

•Powerful and sensitive methods for whole mount *in situ* hybridization and immunocytology. By these means normal or perturbed patterns of gene expression can be visualized at single-cell resolution in sea urchin embryos.

As a result, we know a great deal about the signaling and transcription control processes leading to cell specification in this embryo. The best characterized of all developmentally active transcriptional control systems is a sea urchin *cis*-regulatory element (the *endo16 cis*-regulatory system). The most comprehensive image of how maternal spatial cues initiate differential transcription along the animal-vegetal axis has been constructed for sea urchin embryos. The first large-scale gene regulatory network for a major process of development has been worked out for sea urchin embryos (references included in bibliography). These are all genome-based studies, and the rate of advance in this essential area, which lies at the heart of functional genomics, will accelerate immediately as genomic sequence become available.

3. Uses of Sea Urchins for the Cell Biology and Biochemistry of Eggs, Embryos, and the Fertilization Process

There are many other areas in which sea urchin research has a very high connectivity with respect to the general state of knowledge. This has long been so: for example cyclins were first observed in sea urchin eggs; the role of cell adhesion in embryogenesis was first analyzed in

sea urchins; cytonemes were discovered in sea urchin embryos. Among the areas of general interest in which sea urchin research at present provides leading contributions are:

- •Ca^{+2}-mediated mechanisms of metabolic activation on fertilization.
- •Mechanisms of sperm activation.
- •Biochemistry underlying sperm flagellar motility.
- •Biochemical basis of sperm-egg recognition.
- •Structure/function analyses of cytoskeletal components, including microtubules, microfilaments, and membrane substructures.
- •Role of cytonemes in morphogenetic processes.
- •Structure and function of centrioles.
- •Response to metals and other toxic agents.
- •Function of extracellular matrix in development.
- •Biomineralization processes,
- •Cellular mechanisms of morphogenesis.

Each of these areas resolves (more or less directly) into studies of the function of given gene products. As for most areas of cell biology, progress will be greatly advanced when the perturbations and structure-function assays afforded by experimental control of the underlying molecular biology can be applied. The power of such approaches will in turn depend ultimately on availability of genomic sequence.

4.   Uses of Sea Urchins for Evolutionary Biology

There are three areas in which sea urchins are very important in evolutionary biology, all directly relevant to their genomics. These are, first, the amazing retention of syntenic relations with mammals already noted above, in connection with Fig. 1; second, marine population genetics, gene flow, and speciation; and third, comparative regulatory molecular biology.

With respect to the first of these, assembly of a genomic sea urchin sequence would de facto produce a whole new field of "distant syntenics." The result could be completely novel insights regarding the structure, evolution, and function of deuterostome genomes.

With respect to speciation and population structure, sea urchins have features that make them very different from any animals whose genomes have yet been sequenced. They develop indirectly by way of long-lived, feeding pelagic larvae. Therefore there is continuous intermixing of their gene pool: for example in *S. purpuratus*, which extends from Vancouver to Mexico, there is no greater difference between genomes from individuals collected at the extremes of their range than there is between the two haploid genomes of any given individual: probably typical of many invertebrate marine organisms, the *S. purpuratus* population consists of a huge, panmyctic gene pool. But it also has been a huge gene pool for a very long time, with the result that the *S. purpuratus* genome is about 10-20X more polymorphic than are the genomes of mammals. In other words, there are things to be learned from these animals here that will illuminate large aspects of the biology of oceans and of organismal divergence and speciation therein. Among the current fields of study are:

- •Microsatellite genotyping in the context of wild populations and laboratory inbred populations.
- •Evolutionary rates of divergence, e.g., in populations isolated by the rise of the Isthmus of Panama.
- •Mechanisms of gene pool isolation by evolution of sperm/egg recognition barriers.

In the area of comparative regulatory evolution, as in the area of "distant synteny," there is a direct and relevant link between understanding how a sea urchin genome works and how our own genomes work. It is likely to be possible to translate experimental evidence of the regulatory network features for given developmental processes in sea urchin to our genomes: once one knows what to look for, it is much easier to find it by comparative means alone. For this it may be useful to use a "stepping stone" approach, i.e, from sea urchins to ascidians to vertebrates. There is also an unlimited variety of fascinating comparative regulatory evolution problems with respect to other forms. Three that are being worked on now are comparisons of developmental regulatory processes between *S. purpuratus* and directly-developing sea urchins; comparison to a distant echinoderm, the starfish; and comparison to hemichordates (see Fig. 1). These kinds of studies will illuminate the ways in which developmental gene networks form and reform, the basic process of evolution.

Uses of Sea Urchins for Studies Relevant to Human Disease

Ongoing medically relevant basic research is based on use of sea urchin gametes and cells as model systems. For example, the membrane-bound receptor guanylate cyclase implicated in the important human disease, heat-stable enterotoxin dysentery, was first isolated from sea urchin sperm. The ubiquitous C^{+2} releasing second messenger, cyclic ADP ribose, was discovered in sea urchin eggs and subsequently found to be important in calcium release in the mammalian pancreas. Recently, a connection has been shown between β-catenin, a crucial molecule in early embryonic cell specification, and the differentiation of metastatic cancer cells. Human polycystic kidney disease leading to end-stage renal failure is the most frequent human genetic disease among whites of European extraction. The disease is caused by mutations in human polycystin, whose role in human physiology is unknown. The sea urchin sperm cell receptor for egg jelly (REJ) is the only known protein in GenBank with homology to human polycystin. It controls ion channel activity and by analogy polycystin may be an ion channel regulatory protein whose mis-regulation could be the basis for this human disease. In another area of cell biology, the sea urchin egg is the best model system for the fundamental cellular processes of exocytosis and endocytosis, processes that lie at the heart of synaptic function, insulin release in diabetes, rennin release in hypertension and immunoglobulin release in immune system function. Furthermore envelope viruses such as influenza, hepatitis C and HIV use these same cellular pathways to infect cells. In the sea urchin embryos these processes can be studied in isolation.

PROPOSAL

The *S. purpuratus* Genome and Current Status of its "Pre-genomics"

The genome of *S. purpuratus* is 800 mb in size. It is 39% GC, and consists of about 25% repetitive sequences, consisting of a complex set of diverse elements which occur at frequencies ranging up to tens of thousands per genome. The repetitive sequences have been unusually well studied, and all but the lowest frequency classes can be masked computationally by reference to a library of repeat sequences. The genome has a typical short period interspersed sequence organization, i.e., the single copy domains are punctuated by short repeat elements (a few hundred base pairs in length) every couple of kb or so. There are also some clustered long repeat domains, which include roughly a third of all the repeat sequence length: thus the vast majority of repeat sequence elements are short. There are also on average several insertions per 100 kb of

transposable element genes, usually the remains of reverse transcriptase genes. The genome contains an ordinary frequency of microsatellites as well.

In addition to these general characteristics a large amount of effort has been put into genomics in preparation for ultimate genome sequencing. The following has been achieved, largely as a result of a two-year, $4 million Sea Urchin Genome Project funded by the Stowers Institute for Medical Research, that ended in 2001:

•BAC libraries have been prepared and arrayed from *S. purpuratus* (17.5X coverage, 140 kb average length), and from three other sea urchin species at various evolutionary distances from *S. purpuratus* (*Paracentrotus lividus, Lytechinus variegatus*, and *Eucidaris tribuloides*). Further BAC libraries are in process. These libraries will enable interspecific sequence comparisons around any given gene at any desired distance that is likely to be useful: the range of divergence from *S. purpuratus* is from 50 my for *L. variegatus* (and less for *P. lividus*) to 250 my for *Eucidaris*.

•An *S. purpuratus* BAC-end sequencing project was carried out by Lee Hood and colleagues which has resulted in about 8×10^4 sequence tags, including about 5% of the genome.

•Over 30 BACs have been fully sequenced, resulting in about 4.6 mb of assembled genomic sequence.

•A custom-built sea urchin annotation program has been built, SUGAR (Sea Urchin Genome Annotation Resource). From analysis of the genomic sequence in hand, an estimate of about 22,000 genes is obtained with an average intergenic distance of 30 kb.

•About 16,000 *S. purpuratus* EST sequences exist (though they are not yet all organized in appropriate data bases). There may be an additional EST data set several times this large soon available from the Max-Planck Genome Institute in Berlin.

•Large, arrayed cDNA libraries have been built for every embryonic stage, for unfertilized eggs, and for a number of adult tissues. These libraries are stored at Caltech, and with the support of the NIH National Center for Research Resources, stamped high density filters bearing these libraries are sent on request to any laboratory wishing access to them. Data obtained from screening these libraries are accumulated on a central web site maintained at Caltech.

•Software for interspecies genomic sequence comparison for the purpose of identifying *cis*-regulatory sequence elements; for automated quantitation and comparison of arrayed filter screens; and for construction of regulatory networks, has been built and tested, and is being heavily used.

Whole Genome Sequencing Strategy

The methodology below is based on the approach being taken at the BCM-HGSC for the rat genome project. However it also incorporates a new strategy, clone array pooled shotgun sequencing (CAPSS) to introduce efficiencies and reduce costs. The overall approach is only a suggestion, but likely represents a description that is close to the actual method that could be used.

The sea urchin genome is about 800MB. There currently exists a BAC library of about 100,000 clones (average insert size 140 KB; 17.5x clone coverage). BAC end sequences (BES) have been generated for 38,000 of these but no fingerprints. These existing resources will be used to generate a 6x coverage draft sequence.

About 25000 of the clones (about 4x clone coverage), which have high confidence for correct BES pairing, will be lightly sequenced (0.75x coverage) and the BES will be used with

the resulting contigs to build a tiling path. In order to avoid having to make 25000 separate BAC DNA preparations and libraries, the CAPSS strategy will be used. Cell cultures for the clones will be distributed into a 158x158 array and the clones in rows and columns will be grouped into 316 pools. DNA preparations and shotgun libraries will be prepared from these pools and sequenced to an average of between 0.75x coverage per clone (a total of 3x average sequence coverage for the genome). The sequences from each row will be mixed with the sequences from each column and co-assembled. The assemblies will be analyzed to identify contigs containing both row and column reads, indicating the contigs corresponds to sequences from the BAC at the intersection of the row and column that were mixed, thus deconvoluting the array. This will save having to produce over 24000 BAC DNA preparations and shotgun libraries. It is possible that a more conservative pooling scheme will be used, such as using 250 arrays of 10x10 clones. This would save having to produce 20000 (25000 – 20x250) BAC DNA preparations and shotgun libraries. Initial results on 10x10 arrays from the rat genome project show that this deconvolution technique is successful. The current uncertainty lies with the technical issues in dealing with larger arrays.

Once sequence information is available for the 25000 BACs, the overlap analysis will follow the methodology currently being used at the BCM-HGSC for clone picking in the rat genome project. The contigs will be anchored to the ends of the BACs (by identifying read pairs where one read is in the vector and the other in a sequence contig) as well as linking contigs together into scaffolds based on read pair information. Each BES will be compared to the sequence contigs to find BACs that overlap and the size of the overlap region will be estimated based on where the BES matches the ordered, oriented, and anchored contigs in the scaffolds. This will be confirmed by comparing the sequence scaffolds in overlapping BACs. From this information a tiling path with minimal overlaps will be generated. The informatics pipeline for this method is currently in place at the BCM-HGSC and is being used to identify BACs to sequence for the rat genome project.

Sequences from the BACs that are not in the tiling path will be added to the tiling path and reassembled with the tiling path BACs. The tiling path that is produce will contain some gaps and these will be filled by two methods. First the remaining BES will be compared to the tiling path sequence to place the remaining BACs on the map. Candidates for gap filling will be lightly sequenced and this information used to determine if the gap is completely filled. Any gaps remaining after this process will be filled by screening the BAC library for gap fillers by hybridization, using probes based on sequences at the end of contigs flanking gaps.

The sequence will be brought up to 6x by doing whole genome shotgun sequencing to 3x coverage. The WGS reads will be binned to the proper BAC and assembled using the ATLAS whole genome assembler, the method developed at the BCM-HGSC for assembling the rat genome sequence from a mixed BAC-WGS approach. ATLAS initially finds overlaps between WGS reads, then assigns these groups of reads to BACs based on sequence comparison, and finally assembles the reads in each BAC. The software for clustering reads is also available on the BCM-HGSC web site, where it is called the BAC fisher. This tool allows any investigator who has some sequence information on a region of interest to pull out all reads of relevance before the whole project is over.

Overall this approach would require about 9.6 million successful reads for a 6x sequence coverage (500 bases per read), half in WGS and half in BAC sequencing. This is approximately 6 months sequencing at the BCM-HGSC if all capacity were directed at this project. In addition the project will require from a few hundred to about 5000 BAC DNA preparations and shotgun

libraries. This would take from a few weeks to few months if all capacity at the BCM-HGSC were focused on this project. The overall cost is estimated at about $30 million.

How the Sequence Will Be Used.

The sequences determined by the sequencing center and the preliminary assembly of them will be managed by the sequencing center. It is our aim to make further assembled and annotated sequences immediately accessible, in both practical and intellectual terms, to the community of experimentalists who will make use of them. These further assemblies and annotations will be posted on the web site connected with the Sea Urchin Genome Project. The sequence coverage for our proposed strategy will likely provide sufficient sequence for each BAC to make an ordered and oriented assembly. The assembled sequences can then be accessioned into our sequence database and cross-referenced to the macro-array location for the original BAC clone.

In order to make the process of analyzing sequences convenient we have installed a web-based set of programs, the Cartwheel Project which has a loosely-coupled architecture built on open source code. This system, designed by Titus Brown at Caltech, is in essence a bioinformatics infrastructure which allows the user to have complete control over the analytical process. They are currently supported at Caltech with funding from the NIH National Center for Research Resources. The analyses produced by Cartwheel are then viewable by programs such as SUGAR, the Sea Urchin Genome Annotation Resource, a viewer designed to concurrently represent a variety of genomic features on the sequence, including cDNA matches, repeat sequence matches and genes predicted by several different prediction programs. The analyses stored in Cartwheel can also be viewed with FamilyRelations, a graphical interface that shows large sequence comparisons focusing on conserved elements between two genomes. Because Cartwheel will adhere to a number of open protocols, most notably the Distributed Annotation System (DAS) and Distributed Authoring and Versioning (WebDAV), it is both extremely extensible and compatible with the many distinct formats and protocols used in bioinformatics today. In particular, DAS will allow collaborative field-wide annotation of genome projects based on data generated from and served by individual labs, a heretofore unprecedented ability.

We will immediately erect a community structure to conduct annotation and analysis procedures on the sea urchin genome based on the computational arrangements described above. We will apply for funds to expand the infrastructure to incorporate additional analysis systems as needed so that we are able to annotate the sequence and post the results as they are obtained. The Genome Advisory Group will carry oversight responsibility and install quality control standards for the annotation process. Furthermore, the members of the sea urchin community and other interested parties will be invited to join the annotation effort under the services provided by DAS and DAV.

APPENDIX

Recent Citations Illustrating the Variety of Research Studies that Utilize Sea Urchins

Ameye L, De Becker G, Killian C, Wilt F, Kemps R, Kuypers S, Dubois P. Proteins and saccharides of the sea urchin organic matrix of mineralization: Characterization and localization in the spine skeleton. J. Struct. Biol. 2001 Apr;134(1):56-66.

Angerer LM, Oleksyn DW, Levine AM, Li X, Klein WH, Angerer RC. Sea urchin *goosecoid* function links fate specification along the animal-vegetal and oral-aboral embryonic axes. Development. 2001 Nov;128(22):4393-404.

Arenas-Mena C, Cameron AR, Davidson EH. Spatial expression of *Hox* cluster genes in the ontogeny of a sea urchin. Development. 2000 Nov;127(21):4631-43.

Blank PS, Vogel SS, Malley JD, Zimmerberg J. A kinetic analysis of calcium-triggered exocytosis. J. Gen. Physiol. 2001 Aug;118:145-56.

Cameron RA, Mahairas G, Rast JP, Martinez P, Biondi TR, Swartzell S, Wallace JC, Poustka AJ, Livingston BT, Wray GA, Ettensohn CA, Lehrach H, Britten RJ, Davidson EH, Hood L. A sea urchin genome project: sequence scan, virtual map, and additional resources. Proc. Natl. Acad. Sci. USA. 2000 Aug 15;97(17):9514-8.

Chui KK, Rogers GC, Kashina AM, Wedaman KP, Sharp DJ, Nguyen DT, Wilt F, Scholey JM. Roles of two homotetrameric kinesins in sea urchin embryonic cell division. J. Biol Chem. 2000 Dec 1;275(48):38005-11.

Coffman JA, Davidson EH. Oral-aboral axis specification in the sea urchin embryo. I. Axis entrainment by respiratory asymmetry. Dev. Biol. 2001 Feb 1;230(1):18-28.

Croce J, Lhomond G, Lozano JC, Gache C. ske-T, a T-box gene expressed in the skeletogenic mesenchyme lineage of the sea urchin embryo. Mech. Dev. 2001 Sep;107(1-2):159-62.

D'Andrea-Winslow L, Strohmeier GR, Rossi B, Hofman P. Identification of a sea urchin Na(+)/K(+)/2Cl(-) cotransporter (NKCC): microfilament-dependent surface expression is mediated by hypotonic shock and cyclic AMP. J. Exp. Biol. 2001 ;204:147-56.

Davidson, EH, Rast, JP, Oliveri, P, Ransick, A, Calestani, C, Yuh, C-H, Minokawa, T, Amore, G, Hinman, V, Arenas-Mena, C, Otim, O, Brown, C T, Livi, C B , Lee, P Y, Revilla, R, Clarke, PJC, Arnone, MI, Rowen, L, Cameron, RA, McClay, DR, Hood, LE and Bolouri, H. A provisional regulatory gene network for specification of endomesoderm in the sea urchin embryo. Dev. Biol., in press.

Green GR, Ferlita RR, Walkenhorst WF, Poccia DL. Linker DNA destabilizes condensed chromatin. Biochem. Cell Biol. 2001;79(3):349-63.

Gross PS, Clow LA, Smith LC. SpC3, the complement homologue from the purple sea urchin, *Strongylocentrotus purpuratus,* is expressed in two subpopulations of the phagocytic coelomocytes. Immunogenetics. 2000 Oct;51(12):1034-44..

Holy J. Intermediate filament proteins in echinoderm coelomocytes. Comp Biochem Physiol B Biochem. Mol. Biol. 2000 Dec;127(4):491-504.

Huang L, Li X, El-Hodiri HM, Dayal S, Wikramanayake AH, Klein WH. Involvement of Tcf/Lef in establishing cell types along the animal-vegetal axis of sea urchins. Dev. Genes Evol. 2000 Feb;210(2):73-81.

Kinsey WH, Shen SS. Role of the Fyn kinase in calcium release during fertilization of the sea urchin egg. Dev. Biol. 2000 Sep 1;225(1):253-64.

Kumano M, Carroll DJ, Denu JM, Foltz KR. Calcium-mediated inactivation of the MAP kinase pathway in sea urchin eggs at fertilization. Dev. Biol. 2001 Aug 1;236(1):244-57.

Lessios HA, Kessing BD, Pearse JS. Population structure and speciation in tropical seas: global phylogeography of the sea urchin Diadema. Evolution Int. J. Org. Evolution. 2001 May;55(5):955-75.

Marsh AG, Maxson RE Jr, Manahan DT. High macromolecular synthesis with low metabolic cost in Antarctic sea urchin embryos. Science. 2001 Mar 9;291(5510):1950-2.

McDougall A, Shearer J, Whitaker M. The initiation and propagation of the fertilization wave in sea urchin eggs. Biol. Cell. 2000 Jul;92(3-4):205-14.

Mengerink KJ, Moy GW, Vacquier VD. suREJ3, a polycystin-1 protein, is cleaved at the GPS domain and localizes to the acrosomal region of sea urchin sperm. J. Biol. Chem. 2002 Jan 11;277(2):943-8.

Murray G, Reed C, Marsden M, Rise M, Wang D, Burke RD. The alphaBbetaC integrin is expressed on the surface of the sea urchin egg and removed at fertilization. Dev. Biol. 2000 227:633-47.

Rast JP, Amore G, Calestani C, Livi CB, Ransick A, Davidson EH. Recovery of developmentally defined gene sets from high-density cDNA macroarrays. Dev. Biol. 2000 Dec 15;228(2):270-86.

Robinson JJ. Effects of calcium and magnesium on a 41-kDa serine-dependent protease possessing collagen-cleavage activity. J. Cell Biochem. 2000 Sep 18;80(1):139-45.

Sanchez D, Labarca P, Darszon A. Sea urchin sperm cation-selective channels directly modulated by cAMP. FEBS Lett. 2001 Aug 10;503(1):111-5.

Sherwood DR, McClay DR. LvNotch signaling plays a dual role in regulating the position of the ectoderm-endoderm boundary in the sea urchin embryo. Development. 2001 Jun;128(12):2221-32.

Smith LC. The complement system in sea urchins. Adv. Exp. Med. Biol. 2001;484:363-72.

Stephens RE. Preferential incorporation of tubulin into the junctional region of ciliary outer doublet microtubules: a model for treadmilling by lattice dislocation. Cell Motil. Cytoskeleton. 2000 Oct;47(2):130-40.

Susan JM, Just ML, Lennarz WJ. Cloning and characterization of alphaP integrin in embryos of the sea urchin *Strongylocentrotus purpuratus.* Biochem. Biophys. Res. Commun. 2000 272(3):929-35.

Thomas JM, Masgrau R, Churchill GC, Galione A. Pharmacological characterization of the putative cADP-ribose receptor. Biochem. J. 2001 Oct 15;359(Pt 2):451-7.

Wessel GM, Conner S, Laidlaw M, Harrison J, LaFleur GJ Jr. SFE1, a constituent of the fertilization envelope in the sea urchin is made by oocytes and contains low-density lipoprotein-receptor-like repeats. Biol. Reprod. 2000 Dec;63(6):1706-12.

Yuh CH, Bolouri H, Davidson EH. *cis*-Regulatory logic in the *endo16* gene: switching from a specification to a differentiation mode of control. Development. 2001 Mar;128(5):617-29.

Zhu X, Mahairas G, Illies M, Cameron RA, Davidson EH, Ettensohn CA. A large-scale analysis of mRNAs expressed by primary mesenchyme cells of the sea urchin embryo. Development. 2001 Jul;128(13):2615-27.