Journal of the History of Biology (2010) 43:1–16
DOI 10.1007/s10739-009-9204-1

The Embryo Project: An Integrated Approach to History, Practices, and Social Contexts of Embryo Research

JANE MAIENSCHEIN and MANFRED D. LAUBICHLER
Center for Biology and Society
Arizona State University
Tempe, AZ 85287-4501
USA
E-mail: maienschein@asu.edu

MANFRED D. LAUBICHLER
E-mail: Manfred.Laubichler@asu.edu

Abstract. This essay describes the approach and early results of the collaborative Embryo Project and its on-line encyclopedia (http://embryo.asu.edu). The project is based on a relational database that allows federated searches and inclusion of multiple types of objects targeted for multiple user groups. The emphasis is on the history and varied contexts of developmental biology, focusing on people, places, institutions, techniques, literature, images, and other aspects of study of embryos. This essay introduces the ways of working as well as the long-term goals of the project. We invite others to join the effort, both in this particular project and in joining together in digital collection, archiving, and knowledge generation at the borders of biology and history.

Keywords: embryo, developmental biology, Embryo Project, history, embryology, databases, digital humanities

One of us (JM) has not yet gotten tired of pointing out to stem cell researchers excited about the prospects for regenerative medicine that we are actually already entering the second century of such research. Someday, however, this will become tiresome because people really should have learned that they do not exist in an ahistorical world. Meanwhile both authors continue to argue that awareness of history is an integral part of scientific progress, and this message has recently been finding an increasingly favorable audience.[1]

Perspective is useful for science. In fact, the first stem cell research was

[1] Laubichler, 2007a, b; Laubichler and Davidson, 2008; Maienschein et al. 2008; Maienschein and Smith, 2008.

Ross Granville Harrison's 1907 study of nerve fiber outgrowth in frogs. Alexis Carrell's research building on the new techniques of tissue culture introduced regenerative clinical applications.[2] Few of today's researchers eager to capitalize on new funding and promises of translational results recognize this century-old research as directly relevant to their own. Few ask what they might learn from reflecting on how we got where we are today. Few policy makers or funders ask about the larger context in which the work they are being asked to support has arisen or about the constraints or opportunities created by the historical contingencies. But, optimists or fools that we are, we believe that this is not due to an intrinsic hostility towards the messy complexity of history; rather, we are convinced that this is primarily due to a lack of awareness. In other words, it is an education and communication problem. It is up to historians to stimulate the imaginations of our would-be expanded audiences.

It is, however, difficult for one or a few historians of embryology to talk individually to all these interested stem cell researchers and "set them straight," even if we wanted to do that or thought we could. Even through publications we can only reach a fraction of the possibly interested audiences. And any author will only see some things and not others and can only tell some stories and not all. To overcome these limitations, a group of us are building a collaborative international network of research scholars and students and are putting together an interactive relational database on the science, history and various contexts of embryo research.

As co-directors of this large funded project, we speak for the group in describing the project and strongly inviting further participation. It is appropriate to issue the invitation through this traditional as well as through new electronic formats since the entire project depends on managing electrons and scholars in productive ways. This is why we have an Embryo Project. As a former editor of the *Journal of the History of Biology*, one of us (JM) is grateful to the current editor Paul Farber for initiating the provocative essay series – to which we are pleased to contribute with the goal of reaching the wider audience that the journal addresses.

What: The Embryo Project

A collaborative interdisciplinary team based at Arizona State University in Tempe, Arizona, started with a major grant through the Human and Social Dynamics Initiative at the U.S. National Science Foundation. This grant has funded the research and discovery component of the

[2] Harrison, 1907; Landecker, 2007.

project. In addition, the Arizona State University library researchers led by Associate Dean John Howard and his team of programmers, database experts and information systems managers have been adapting the library's Fedora Commons-based information systems to work for the multiple challenges of our project. A series of focused workshops has added to that development as well as to our ability to share tools and trials with other similar projects and like-minded project leaders.

Furthermore, we work with an international team of scholars and collaborating institutions through workshops and virtual collaboratory meetings focused on well-defined and structured scholarly projects. We began in partnership with the Max Planck Institute for the History of Science in Berlin, as we were inspired by their Virtual Laboratory (http://vlp.mpiwg-berlin.mpg.de/index_html), and the Marine Biological Laboratory-Woods Hole Oceanographic Institution Library in Woods Hole, Massachusetts, because of their many digital projects in related areas such as biodiversity. It quickly became clear that for our project, it made sense to build on and expand the infrastructure developed by these existing projects. The librarians at ASU, the MBL, and the MPI recommended the Fedora system (http://www.fedora-commons.org/) and have been extremely generous in contributing to the research to make the project work effectively.

The mission of the project is to document, collect, and compile materials in digital form, and interpret "everything" related to embryo research and the multiple contexts in which it occurs. Study of embryos has taken place through the fields of philosophy, embryology, and developmental biology, and the relevant contexts bring in history, religious studies, bioethics, law, and studies of the material practice of science including history of technology. Of course it is ambitious to suggest that we want to capture it all, but we indeed maintain that as our goal even while realizing that we will succeed best by starting with focal projects. It is the beauty of electronic media that they enable new dimensions of collaborative research.

To facilitate collaborative activities we have been building an editorial process so that interested scholars from anywhere can submit materials that we review, edit, and add to the database. This is essentially a peer reviewed publishing project complete with an ISSN # to prove that we are "real" by traditional standards. But even that extended review process and the additional help of an extended editorial team are not enough, since even a small group can only do so much and the editorial service can still be a bottleneck.

This is where the object related databases and the library's collaboration become essential. In addition to developing our own new materials, we can discover, retrieve and include materials stored in other open access databases through robust searching of metadata. This digital federation is critically important to the ability to accumulate a significant project effectively. To this end we have been following the model of the Bioinformatics revolution in the biomedical sciences. In these areas the huge data-gathering project, such as the various – *omics* projects, have led to an overabundance of data, which are now being annotated, cross-linked, and otherwise curated. For these various needs, many tools for data-mining and data management have been developed. We and our partners, including many of the other digital HPS projects, which have in the meantime formed a consortium for Digital HPS (www.digitalhps.org) have been adapting both strategies and tools from bioinformatics, as well as developing our own applications (see the Digital HPS website for a list of tools and further links to the various projects). We are also working on the integration between these projects and a platform for the distribution of shared tools that will facilitate new ways of research and collaboration in HPS related areas.

For our intellectual goals we are taking scientific research on embryos (or embryo research) as a case study of the way that science changes over time. It is a case of considerable importance for both science and for society. Embryo research changes for many complex reasons and because of diverse interacting agents of change. We are identifying, documenting, and exploring these factors and interplay among them that lead to changes over time in both science and the relevant social contexts (see Figure 1). Thus, within the context of the NSF initiative that initially funded the project, we are examining a case of Human and Social Dynamics of rapidly changing science within radically contested contexts.

Embryo research is a fascinating object of study, with significant changes in factors such as choices of which organisms to study (mostly animals until recently when human embryology began); techniques to use (including experimental techniques, equipment, and the way results are presented in publications and presentations and represented in images and models); and which actors, laboratory settings, institutions, and local contexts are important for each episode of scientific and technical work. As with all of science but sometimes in exaggerated and unique ways, embryo research is also always embedded in webs of unsettled ethical, legal, political, religious, cultural, and social negotiations that shape the conduct of science, its diverse meanings, and the

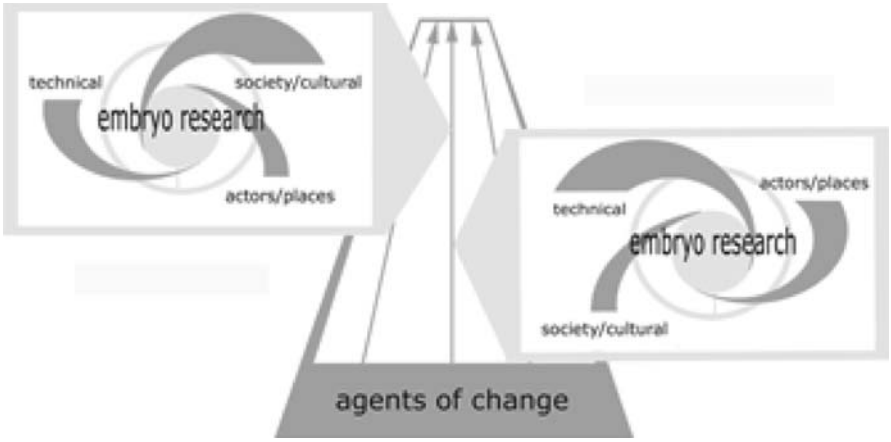

Figure 1. Agents of change shaping the course of embryo research through time.

spectrum of decisions built upon such understandings. To organize and make sense of the swirl of all these factors that affect scientific change, we have organized them into the analytic categories of People, Places, Practices, Concepts, Contexts, Images, and Literature, although the actual coding structure of the relationships between objects in the database is much more fine-grained and specific (see Figure 2).

Who: The Network

The project brings together researchers from the multiple disciplines needed to make sense of such a complex subject. Individuals can submit any materials for consideration, but the project is also organized into defined Projects with leaders from our international network. Our network involves about 30 researchers from at least seven countries and we expect it to grow considerable as more individual projects are added to the larger context of the Embryo Project. A current and updated list of members can be found at (http://embryo.asu.edu/about/network.php).

At ASU the network also includes groups of graduate and undergraduate, especially honors, students who work on different elements of the projects through directed study and seminars. These groups add substantially to the database as students find and annotate materials, which are then reviewed and edited through the editorial process of the Embryo Project. As new generations of students have become involved with the Embryo project we are also benefitting from their perspectives towards the web, digital media, and novel forms of communications,

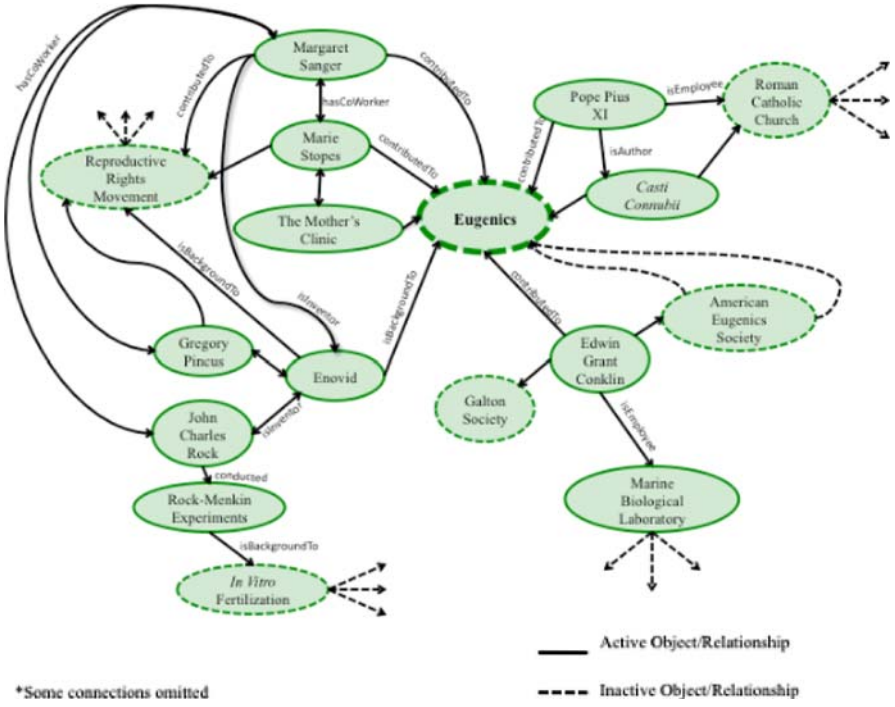

Figure 2. An illustrative small subset of relationships among objects in the database based on the connection of the Eugenics movement with embryo research. Fedora stores both objects and the relationships, so that both can be retrieved through searches and other web-applications (with thanks to Cera Lawrence).

such as social networking sites and communication tools (Twitter, etc.). The overall Embryo Project is thus adapting to a Web 2.0 framework, which, interestingly, also helps us see traditional scholarly sources and analytical categories in a new light. Most prominently, our focus on multiple relationships and their visualization (see Figure 2) is greatly aided by this new focus on ''connectivity.''

How: Defined Projects

We mentioned above that the project mainly grows through specific research projects and questions. These defined projects generally take a particular locus of embryo research and document the people, places, practices, images, literature, and concepts of that research and their interactions. We then look at the relevant contexts for broader perspective. These focused projects grow out of workshops or individual research efforts, so that anyone interested in the history or current issues

related to embryo research can propose a project and apply for assistance in establishing the necessary groundwork to take advantage of the Embryo project infrastructure.

A list of ongoing research projects (see also Figure 3) can be found at (http://embryo.asu.edu/about/projects.php). These include projects on individual scientists and their associated research programs (such as Victor Hamburger or John Bonner and morphogenesis), research institutions (such as the Marine Biological Laboratory), or broader topics (such as Regenerative Biomedicine, the legal status of embryos, theories of development in the 20th century, or cloning and stem cell research and policies). All the projects are collaborative and most bring together developmental biologists with historians, philosophers and science studies scholars.

For example, the project on theories of development in the 20th century has generated a close collaboration between Manfred Laubichler and Caltech developmental biologist Eric Davidson that has resulted in both scholarly/historical and theoretical scientific products thus transcending the disciplinary silos of science and history in multiple ways.[3] Another example is the collaboration of Jane Maienschein and Manfred Laubichler with William C. Aird on the latter's foundational textbook on Endothelial Biomedicine.[4] Our initial collaboration on the history of this emerging field of biomedicine, especially on the importance of a developmental dimension in early research on the endothelium (the inner lining of the blood vessels and the lymphatic system), has led Bill Aird to redefine the theoretical foundation of his (1800 + page) multi-authored volume in such a way that developmental and evolutionary/comparative considerations have become much more central.

As mentioned, our goal from the beginning has been to study all aspects of embryo research, by all people in all times and places, from Aristotle to tomorrow. Of course we cannot do that alone or easily, and we expect this aspect of the project to grow tremendously well into the future, with input from an expanding network of researchers. We have been focusing on western research and really in the U.S. and Europe, drawing primarily on English language sources at first for practical reasons related to searching across databases. But clearer definitions of metadata and additional tools developed in the context of other digital HPS projects, such as ontologies and dictionaries, are helping us to overcome these restrictions (see also the digital HPS site for some concrete examples).

[3] Laubichler 2007a, b; Laubichler and Davidson 2008.

[4] Laubichler et al. 2007.

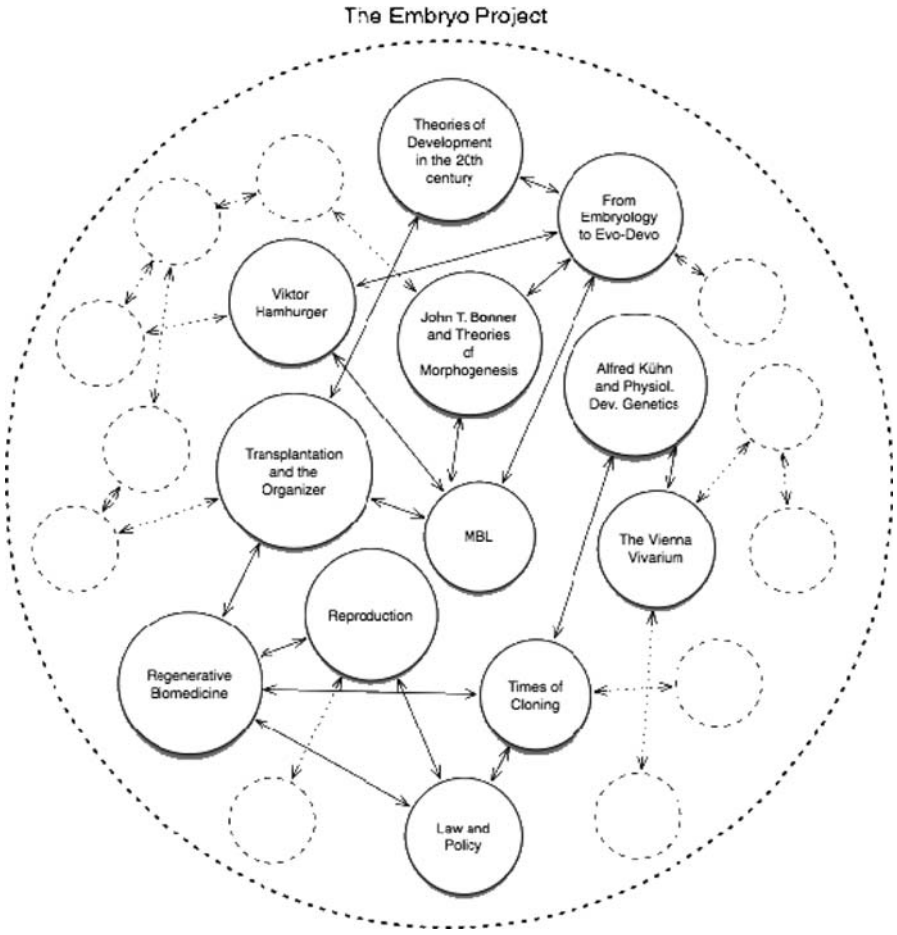

Figure 3. An illustrative set of some current projects and their connections.

Also, as more individual projects generate lists of objects and their relationships within the broad categories of the project, we are already seeing an exponential increase in the number of often hidden or surprising relationships among objects.

Our database is thus not primarily a repository but rather a tool of discovery that adds new dimensions to traditional scholarship in the humanities and sciences. The fact, that we do not just present the final interpretation as a scholarly result (with sources and references buried in footnotes), but also make available all the research objects and their relationships transforms the practice of the history of science and allows it to interface much better with current science and science policy.

As with all databases, our approach requires a certain amount of standardization of metadata and relationships, for reasons of internal consistency and also external compatibility with other databases. Yet our descriptors of the data are not completely static, but evolve together with the insights gained by the project. Our system's architecture then allows for sophisticated searches that form the foundations of more traditional interpretations of the history, the contexts, and agents that have shaped the many transformations of embryo research. Basic and important relationships that are hard to discover without such a system include such things as:

Who was a student of whom

Who worked at a particular place? With particular organisms? Using particular equipment or techniques?

What concepts were associated with what practices or what organisms? For example regeneration, transplantation, stem cells, induction, etc.

What definitions of the embryo have driven what research and what interpretations? How does seeing the embryo as, for example, a set of gene regulatory networks, make a difference rather than, say, seeing it as a cluster of differentiating cells undergoing morphogenesis?

What contextual factors were connected to any particular type of experiment? Which were medically relevant? Part of a comparative project? Interpreted in evolutionary terms?

For example, one fundamental problem of historical analysis is tracing connections on a large scale, especially within the ever-growing body of scientific literature. Most off the shelf tools, such as citation analysis, establish a link (i.e. a citation), but do not easily allow further exploration in form of a semantic or content analysis (i.e. is the citation affirmative or does it refute the claims of the previous paper?). To overcome these limitations we need to be able to analyze the connections at a more fine-grained scale. Computationally this is a problem of data mining, functional annotations, and the application of tools for semantic analysis. Such tools exist in various parts of Bioinformatics and can be adapted for use in digital HPS.

We have tested these possibilities in the context of Medline abstracts. Medline contains text (abstracts as well as full text and meta-data) of a majority of the biomedical research literature of the last decades (more than 14 million entries so far). To test our system we initially selected a set of over 1,000 Medline abstracts that used the purple sea urchin as a model organism. Model organism is, of course, a relevant category for historical analysis. We then created a set of simple ontologies that each

represented a category of interest. These included people – we mined the author list to populate the ontology of all people that published on the purple sea urchin, places – we used the institutions of these authors, concepts – we developed an ontology of concepts relevant to research in developmental biology, experimental methods, etc. All these ontologies were then converted into and added to already existing ontologies relevant to our project, such as species names and gene names.

Next we could then analyze the abstracts based on connections between different instances, and quickly answer questions such as: Eric Davidson publishes on the purple sea urchin, he is at Caltech – who else at Caltech publishes on the purple sea urchin, who uses the same methods, who discusses the same concepts, in what papers are certain methods and concepts connected, what are all the genes discussed in papers on the purple sea urchin that deal with a specific developmental process such as the formation of the larval skeleton, etc.?

The list of queries that can be formulated and addressed is huge and the analyses can be done immediately. Based on the results we were then able to refine the underlying ontologies, formulate new queries and expand the search to a larger set of Medline entries or additional digital materials. To list just one concrete example of how our analysis of this sample set has changed ongoing research (within the context of Laubichler's project on Theories of Development and the EP): Sea urchins have been a one of the major model organisms in experimental embryology for over a century. However, as developmental biology began its gradual transformation towards developmental genetics, sea urchins were dropped as a "model organism" as it was not possible to perform standard genetic analysis on them. They were re-introduced within the context of molecular developmental biology as soon as it was possible to clone and isolate individual genes and within a specific conceptual program of studying gene action and gene regulation in differentiation. We were able to establish an initial history (timeline) of molecular techniques connected with this gradual re-introduction of the sea urchin into developmental biology and are now exploring how these techniques were connected to a new set of questions and related concepts, such as gene regulatory networks, that emerged in the course of this novel research trajectory.

The much more difficult understanding of interactions among all the factors requires just the sort of organized and managed multi-disciplinary and interdisciplinary research strategies and methods that we offer with the Embryo Project. The Project therefore provides a basis for new scholarly interpretations. Yet this is not all. We also explicitly

intend to provide materials for multiple user groups, including research materials for scholars in multiple disciplines and also educational materials and interpretations for teachers, the general public, and public figures such as Congressmen and judges. While the collection of materials for scholars is more familiar, the development of interpretive essays for the multiple user groups and multiple purposes takes more effort and has important broader implications. The four types of contributions, outlined below, help show how the complex project works to meet all these objectives.

How: The Embryo Project Encyclopedia Database

This part of the project involves developing a database, with the formally recognized name of the Embryo Project Encyclopedia. The Encyclopedia consists of three different types of items linked through coded relationships within the system and through standardized metadata to the full databases of our and other library networks. The four types of objects included are:

1. Short descriptions of the objects in the database. These objects may be examples of those people, places, practices, and so on. The descriptions are typically derivative, relying on trusted secondary information sources. They are submitted for review and editing, and then checked and coded by an editorial team that identifies and codes from lists of desired relationships and metadata. To date, student researchers have contributed most of these entries, after they are trained in the Embryo Project research seminar for undergraduates and graduate students and led by the faculty and a postdoctoral fellow. The items of this type are formally attributed to the author, and we expect many other authors to begin contributing now that we have the system developed and working, and given that these descriptive entries provide the entry point for linking to other materials the author may be interested in seeing in the Encyclopedia. But these entries are not just mere encyclopedic descriptions or mimicking already existing sources; rather they are written and marked up in such a way that they help populating the database with additional objects that have interesting and relevant relationships to the object of the entry. The encoded relationships and metadata also provide visible hyperlinks that add functionality for users seeking primarily factual information in the database.

2. Found objects. This category includes materials already "out there." This includes such things as published literature – both primary and secondary, images of people or places or practices, video, interviews – both tape and transcripts, lab notebooks, archival materials, and anything available digitally. Our library collaborators have been developing several programs that allow us to find already existing objects in a number of ways. Therefore, we can link different forms of data from widely divergent types and locations. Prototype functionalities currently allow us for instance to harvest digitally available literature written by any identified developmental biologist in the database that is scattered across multiple journals and libraries and resolve access based on the physical location and privileges of the person sending the query.

Since some of these materials are behind various firewalls or are password protected, we are following the librarians' lead in how best to carry out these linkages to provide access. We have also been developing an image archive and are working with the Marine Biological Laboratory in Woods Hole to capture the thousands of relevant practice-based materials and images they have made available to us. We hope other collections can also be linked for open source use in ways that can be discovered and accessed by a much wider range of users. All of our own material is fully open access and the Embryo Project supports all initiatives that aim to make as much of the scientific heritage freely available to all researchers.

3. Created objects. Based on the ever-growing set of web-applications and tools for digital HPS we can also add created objects, to our repository. One example is timelines, such as the one we created for the history of the embryology course at the MBL or for Ross Harrison's biography (see Figures 4 and 5). Many of these tools are dynamic in the sense that they can easily incorporate additions to the database.
4. Interpretive essays through an online scholarly journal dedicated to studies of embryo research and its multiple contexts. We have begun to accept articles submitted for this purpose, which will be fully peer reviewed and edited. All articles will also be processed so that all objects and relationships between them will become part of the database. It is important that the result be open access and help users find links from materials and questions that they will find familiar to other materials and articles that are related but perhaps addressing different questions or different aspects of a subject that

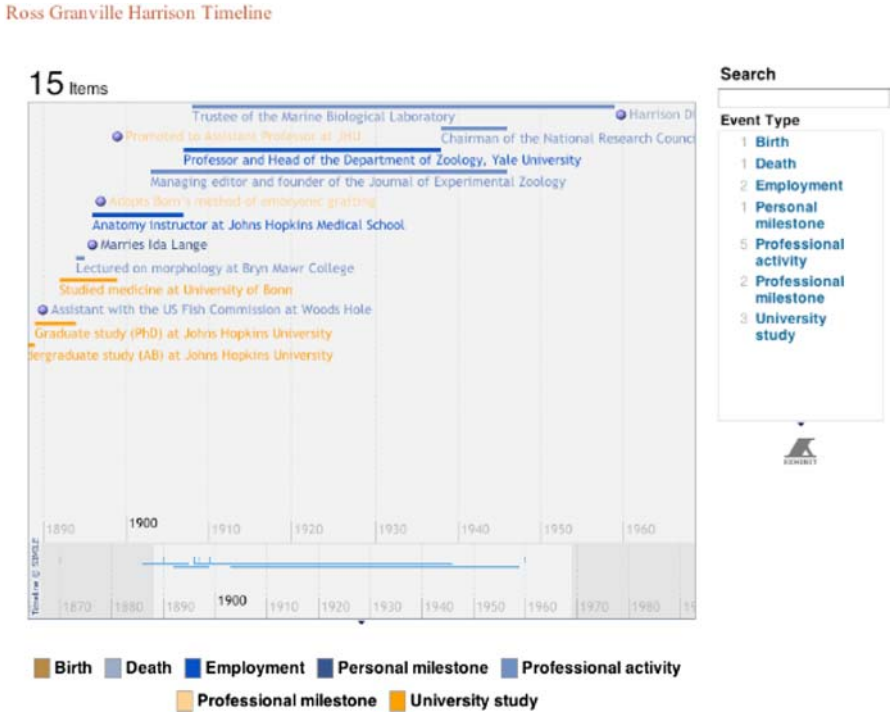

Figure 4. A timeline generated form material stored in our database related to Ross Harrison's career. As more materials are added, the timeline can be populated with these events.

they would not have thought to ask about. Our tools and applications are designed to aid researchers in finding new and unexpected connections, in our case related to embryo research. For example, embryologists rarely know how to carry out legal research, and yet they might well want to know what the regulatory climate is before they set up a stem cell research unit. Or American historians might be interested in learning about the rich history of embryo research to include a unit alongside discussions of genetics and evolution as important biological contributions. Historians of technology might start with an interest in microscopes and move to understanding of the concepts associated with those technologies. All these various researchers can query the database and not only find "hits" as in a traditional Google search, but also find relationships and links that are scholarly verified.

Setting up a new on-line journal that is scholarly, respected, listed in the relevant professional databases and fully searchable is challenging,

Figure 5. A short excerpt of a timeline of the history of the Marine Biological Laboratory at Woods Hole showing various events between 1890 and 1920. Each events links to various materials in the database, such as biographies or images (shown here for Frank R, Lillie and thanks to Grant Yamashita).

as we are learning about different kinds of scholarly publications and all the details about intellectual property, ownership, setting high scholarly and editorial standards, and the full range of publishing details. We are working with librarians on this project, along with our editorial team. We welcome suggestions as well as submissions.

Why Participate?

The Embryo Project is driven by the scholarly community and depends on the contributions of interested parties to help it grow and become more useful in future. Current funding allows for workshops and to hire student workers. Most of the programming has been an investment in digital research infrastructure by the libraries at ASU, the MPI and the MBL. We have a reasonable system in place to sustain the project, and are also actively seeking additional support. But we hope that the scholarly community in the history of biology will take advantage of

this new approach and contribute by creating entries and providing new materials, which in turn will support everyone's scholarship as well as reaching beyond to have extended broader impacts for multiple user groups.

Our growing research networks that work across what are normally different disciplinary boundaries have already led to new projects and collaborations and we have begun a series of workshops, to which Michael Dietrich has been an especially valuable contributor. Topics include the history of morphogenesis (with the active participation of noted developmental biologist John Bonner), the role of the decades old embryology course at the MBL (bringing together all living directors of this course for an oral history project) or the history of developmental genetics and cloning and stem cell research (with an initial workshop organized by Christina Brandt and Giuseppe Testa at the MPI in Spring 07). Many other projects are also being developed, including several close collaborations between historians and developmental biologist, such as those between Manfred Laubichler and Eric Davidson on theories of development.

The scientific community clearly benefits from these synergies, as do students, who early on become active participants in the many different aspects of the scientific process, from data collection, annotation to scholarly interpretation, editing and publishing. The open access philosophy and the database organization of the Embryo Project also allows us to reach out to multiple user groups – the general public, policy makers, different scholarly communities and also scientists who increasingly see the benefit of having historical information available in a form that is not completely foreign to them. The Embryo Project is thus an example of an emerging research network that spans multiple disciplines and countries, with a virtual collaboratory, a repository/scholarly publishing product that is openly accessible, linked sources, and interpretations for multiple user groups.

Acknowledgments

This project has benefited tremendously from grant support from the National Science Foundation (SES 0623176, 0914069, 0645729) and Arizona State University, and from the support and guidance of our institutional partners at the Max Planck Institute for the History of Science in Berlin and the Marine Biological Laboratory-Woods Hole Oceanographic Institution Library. Felicity Snyder, Andrew

Hamilton, and Grant Yamashita have provided far more in the way of intellectual leadership than can be indicated on the website, as have Marie Glitz, Mary Sunderland, and Cera Lawrence. Grant Yamashita has developed the timelines and has made possible the intellectual content as well as the images in the other figures. ASU librarians John Howard and Philip Konomos have played indispensable intellectual and implementation roles. MBL-WHOI Library Director Cathy Norton and MBL Director Gary Borisy have provided tremendous support and encouragement, as has the MBL technical team.

References

Harrison, Ross Granville. 1907. "Observations on the Living Developing Nerve Fiber." *Anatomical Record* 1: 116–118.

Landecker, Hannah. 2007. *Culturing Life. How Cells Became Technologies*. Cambridge, MA: Harvard University Press.

Laubichler, Manfred, D. 2007a. "The Regulatory Genome: Eric Davidson at 70." *Bioessays* 29: 937–939.

—— 2007b. "The Specter of the Past. What the History of Theoretical Biology Means Today." *Biological Theory* 2: 131–133.

Laubichler, Manfred, D. and Eric, H. Davidson. 2008. "Boveri's Long Experiment: Sea urchin Merogons and the Establishment of the Role of Nuclear Chromosomes in Development." *Developmental Biology* 314: 1–11.

Laubichler, Manfred, D., Aird, W. and Maienschein, J. 2007. "The Endothelium in History." W. Aird (ed.), *Endothelial Biomedicine*. Cambridge: Cambridge University Press, pp. 5–22.

Maienschein, Jane, Laubichler, Manfred, D. and Loettgers, Andrea. 2008. "How Can History of Science Matter to Scientists?" *Isis* 99: 341–349.

Maienschein, Jane. and Smith, George. 2008. "What Difference does History of Science Make Anyway?" *Isis* 99: 318–321.