The e-Mouse Atlas Project (1992-) [1]

By: Miller, Shawn A. Keywords: Gene Expression [2] EMAP [3] EMAGE [4] Duncan Davidson [5] Richard Baldock [6]

The Edinburgh Mouse Atlas, also called the e-Mouse Atlas Project (EMAP), is an online resource comprised of the e-Mouse Atlas (EMA), a detailed digital model of mouse [7] development, and the e-Mouse Atlas of Gene Expression (EMAGE), a database that identifies sites of gene expression in mouse [7] embryos. Duncan Davidson and Richard Baldock founded the project in 1992, and the Medical Research Council (MRC) in Edinburgh, United Kingdom, funds the project. Davidson and Baldock announced the project in an article titled "A Real Mouse for Your Computer," citing the need to manage and analyze the volume of data that overwhelmed developmental biologists. Though EMAP resources were distributed via CD-ROM in the early years, the project moved increasingly online by the early 2000s. Into the early decades of the twenty-first century, it was in active development. EMAP can be utilized as a developmental biology teaching resource and as a research tool that enables scientists to explore annotated 3D virtual mouse [7] embryos. EMAP's goal is to illuminate the molecular basis of tissue differentiation [8].

The project founders, Baldock and Davidson, both studied bioinformatics. In the 1990s, they began to argue that developmental biologists faced a data-management problem similar to the one that molecular biologists grappled with in curating DNA sequence data. Davidson pursued bioinformatics to study the representation of the dynamics of gene expression, cell lineage [9], morphogenesis, and differentiation [8]. Baldock received a PhD in Theoretical Physics in 1980, when at the Australian National University [10] in Canberra, Australia, and he then completed a postdoc at the University of Oxford [11], in Oxford, England, before specializing in biomedical imaging. Davidson and Baldock argued that information should be stored digitally because of the ubiquity and relatively low cost of hard disk drive storage devices, which are relatively easy to copy, a feature that helped researchers share information. The move to digitize gene expression patterns was also a response to what Davidson and Baldock identified as the problem with the standard methods of publication, wherein much of the raw data on which published results were based were infrequently made available except in summary form. Embryo images are also often photographed in non-standardized ways by different laboratories.

In 1991, The Ciba Symposium called "Postimplantation Development in the Mouse" first met to address these challenges. Jeremy Green, who specialized in molecular signals at King's College London, in London, England, and Peter Rigby, a molecular biologist trained at the University of Cambridge, in Cambridge, England, and at Stanford University [12], in Stanford, California, organized the symposium. Both organizers were part of the UK's National Institute for Medical Research (MRC), headquartered in London, England. During the symposium, scientists agreed on a pilot project based on the 9-day mouse [7] embryo. According to Matthew Kaufman, author of 1992's *Atlas of Mouse Development*, mice were a suitable subject because they had been widely used by developmental biologists researching molecular mechanisms responsible for mammalian development and differentiation [8].

The scientists at the symposium determined three requirements for the pilot project. First, they sought to digitally represent mouse [7] anatomy at the level of tissues. Next, they aimed to store molecular expression data that typically does not comport with anatomical structures. Lastly, they expected to create an interface for users to interact with the database. In the 1990s, researchers developed these technologies at the MRC Human Genetics Unit, Western General Hospital, in Edinburgh, UK, where both Davidson and Baldock, who had attended the Ciba Symposium, held appointments in biomedical systems analysis.

By 1994, EMAP expanded its institutional relationships beyond MRC to involve the University of Edinburgh [13], in Edinburgh, UK. Baldock held an honorary professorship at the University of Edinburgh [13]. In 1999, EMAP expanded again, with the help of another honorary professor at the University of Edinburgh [13], computational biologist Martin Ringwald. Ringwald worked at Jackson Laboratory, an independent, nonprofit organization [14] in the US studying mammalian genetics. The Jackson Laboratory had been developing the Gene Expression Database (GXD), which stores and integrates textual gene expression data of various kinds, emphasizing in particular endogenous gene expression during mouse [7] development. The collaboration with the Jackson Laboratory resulted in incorporating EMAP database nomenclature into the GXD.

Portions of EMAP were originally designed as object-oriented databases, but researchers later converted those portions to a relational architectures to accommodate the data format standards set in the "Minimum Information Specification For *In Situ* Hybridization and Immunohistochemistry Experiments (MISFISHIE)," which the Microarray Gene Expression Data Society set in 2008. This document enacted guidelines and standards that governed the format of visual data produced in experiments investigating gene expression in tissue. Davidson and Baldock argued that standardizing and structuring forms of data in developmental biology is necessary for the online resources because those standards create a common language that facilitates the sharing of results. Also to facilitate research and communication, researchers licensed all content of EMAP under a Creative Commons Attributed 3.0 License and is thus an Open Access resource.

The EMA portion of EMAP provides 3D volumetric models, which detail the gross anatomy, shape, and detailed tissue structures

of mouse [7] embryos. For standardization, EMA uses Theiler Stages, which divide the mouse [7] development process into 26 prenatal and 2 postnatal stages. Karl Theiler's 1972 *The House of Mouse: Atlas of Embryonic Development* provides the basis for those stages, which became the standard stages among mouse [7] researchers. Theiler had received his PhD in 1945 from the University of Zurich [15], in Zurich, Switzerland. He then researched at both Columbia University [16] New York, in New York, and the Jackson Laboratory from 1953 to 1954 before returning to the University of Zurich [15] as a professor of anatomy. Theiler Stages were too broad to distinguish certain important phases of early embryological development. Therefore, EMAP researchers supplemented Theiler's descriptions with information such as cell number, somite number, or other characteristics. Karen Downs, a professor at the University of Wisconsin, in Madison, Wisconsin, and Timothy Davies, a researcher at the University of Oxford [11], detailed this method in a 1993 article titled "Staging of Gastrulating Mouse Embryos by Morphological Landmarks in the Dissecting Microscope."

The EMAP anatomy ontology consists of a controlled vocabulary of terms that detail, in text form, the names and structural relationships among parts of a developing mouse [7], from fertilization [17] of an egg [18] through birth of a pup. In EMA, researchers can map a term from the ontology onto images of embryos at various stages of development. To do so, they label locations deemed significant on mouse [7] embryo images, similar to the way a road atlas labels areas of interest on a map, such as cities, state parks, major highways, and so forth.

EMAGE, the second component of EMAP, consists of two elements. The first is a database of gene expression data for mouse [7] embryos. The second is a suite of digital tools that enable researchers to query and analyze the database, which a staff curates. The database is populated by scientific research on mouse [7] embryology [19] from numerous sources. Curators select data from academic journals with which EMAP has legal agreements or Creative Commons licensing arrangements. The GXD curating staff also identifies, compiles, and annotates mouse [7] gene expression data from hundreds of journals, and they collect that data in the Gene Expression Literature Index. Additional data comes from individual research laboratories and large-scale gene expression screening projects such as the European Union-funded EURExpress and FaceBase, which provides 3D images of craniofacial development. FaceBase is run by Mike Dixon at the University of Manchester, in Manchester, England, along with David FitzPatrick at the MRC Human Genetics Unit.

EMAGE describes gene expression patterns using textual annotation and spatial annotation tools. Curators annotate the EMAP anatomy ontology to indicate sites of gene expression, based on key words in journal articles. For instance, if a journal article describes the development of the cardiovascular system, curators can translate the textual description in the article into terms of the EMAP controlled vocabulary, which has associated four digit numbers, and which enables researchers to organize their annotations in hierarchies. Database entries for such an article might include labeling areas responsible for the development of, in order of increasing specificity:

```
organ system (EMAP:2220)
cardiovascular system (EMAP:2388)
        arterial system (EMAP:2389)
                branchial arch [20] artery (EMAP:2390)
                        1st arch artery (EMAP:2391)
                        2nd arch artery (EMAP:2392)
                        3rd arch artery (EMAP:2393)
                                4th arch artery (EMAP:2394)
                                6th arch artery (EMAP:2395)
```

Often the terms used in an article will correspond to the EMAP ontology, but not always.

To spatially annotate images of developing mice, researchers use data from the e-Mouse Atlas as a spatial framework, in which standard images indicate where in normal mouse [7] embryos genes [21] make proteins or other products. Unlike a natural language, which must conform to certain syntactic structures, data in images are subject to no such constraints. Spatial annotation takes this unstructured data and maps it onto a spatially standardized description to be stored in a database. The e-Mouse Atlas thus functions as a tool that helps identify and name parts of the embryo. EMAGE uses a bespoke program called MAPaint for spatial annotation. EMAGE also stores spatially mapped data produced using automated methods, though such automation requires data that has been produced in standardized ways. A goal of EMAP is developing computational tools capable of further automation.

EMAGE provides a variety of search and analysis tools for investigating data. Users can search the online resources by embryo space, anatomical structure, or by gene/protein. EMAP has developed JAtlasViewer, a Java program that allows viewing virtual sections of three-dimensional images, such as mouse [7] embryo models. They have also developed WlzIIP server and viewer, providing access to large three-dimensional volumetric data sets, as well as the woolz libraries and binaries for image processing. Another goal of the project is to manage these data sets and to make them available online. High-resolution sections and three-dimensional reconstructions of mouse [7] embryos can create files in excess of ten gigabytes, which are too large for many computers to handle. These large files led Baldock and his collaborators to create computer server software that attempts to provide access to multidimensional embryological maps, which retain much of the information, without placing burdens on network bandwidth and users' computational resources.

Researchers at EMAP worked to determine how best to represent data-dense, three-dimensional images in digital databases

that utilize data formats that researchers can store, analyze, and search. Other genomic database projects, such as GenBank, provide a contrast in that they focus on structural relations among the four nucleotides found in DNA: adenine, thymine, cytosine, and guanine, which researchers represent in computationally manageable ways, as A, T, C, and G. Developmental biologists' data, however, is more complex, including lists of RNAs, proteins, and other molecules synthesized during embryogenesis [22]. Davidson and Baldock noted at EMAP's inception that explicating the functional roles of regulatory molecules in embryogenesis [22] requires researchers to compare and interpret, across developmental stages [23] and across species, when and where in developing organisms genes [21] make proteins and other products. Such comparisons in turn require researchers to generate long lists that link gene expression data to specific areas of anatomy during development. These lists are what researchers store in the databases.

Sources

1. Baldock, Richard, and Albert Burger. "Anatomical Ontologies: Names and Places in Biology." *Genome Biology 6* (2005): 108–8. http:// genomebiology.com/2005/6/4/108 [24] (Accessed February 28, 2014).
2. Baldock, Richard, Jonathan Bard, Kaufman Matthew, and Duncan Davidson. "A Real Mouse for Your Computer." *BioEssays* 14 (1992): 501–2.
3. Bard, Jonathan, Richard Baldock and Duncan Davidson. "Elucidating the Genetic Networks of Development: A Bioinformatics Approach." *Genome Research* 8 (1998): 859–63. http://genome.cshlp. org/content/8/9/859.full [25] (Accessed February 28, 2014).
4. Bard, Jonathan, Matthew Kaufman, Christophe Dubreuil, Renske M. Brune, Albert Burger, Richard Baldock, and Duncan Davidson. "An Internet-accessible Database of Mouse Developmental Anatomy Based on a Systematic Nomenclature." *Mechanisms of Development* 74 (June 1998): 111–20. http://dx.doi.org /10.1016/S0925-4773(98)00069-0 [26] (Accessed February 28, 2014).
5. Davidson, Duncan, and Richard Baldock. "Bioinformatics Beyond Sequence: Mapping Gene Function in the Embryo." *Nature Reviews Genetics* 2 (2001): 409–17.
6. Davidson, Duncan, Jonathan Bard, Matthew H. Kaufman, and Richard Baldock. "The Mouse Atlas Database: A Community Resource for Mouse Development." *TRENDS in Genetics* 17 (2001): 49–51.
7. Davidson, Duncan, Jonathan Bard, Renske Brune, Albert Burger, Christophe Dubreuil, Bill Hill, Matthew Kaufman, Jane Quinn, Margaret Stark, and Richard Baldock. "The Mouse Atlas and Graphical Gene-Expression Database." *Seminars in Cell & Developmental Biology* 8 (1997): 509–17.
8. Deutsch, Eric W., Catherine A. Ball, Jules J. Berman, Steven G. Bova, Alvis Brazma, Roger E. Bumgarner, David Campbell et al. "Minimum Information Specification for In Situ Hybridization and Immunohistochemistry Experiments (MISFISHIE)." *Nature Biotechnology* 26 (2008): 305–12. http://www.nature.com/nbt/journal/v26/n3/full/nbt1391.html [27] (Accessed June 11, 2014).
9. Downs, Karen M., and Timothy Davies. "Staging of Gastrulating Mouse Embryos by Morphological Landmarks in the Dissecting Microscope." *Development* 118 (1993): 1255–66. http://dev. biologists.org/content/118/4/1255.full.pdf [28] (Accessed February 28, 2014)
10. e-Mouse Atlas. http://www.emouseatlas.org/ [29] (Accessed June 11, 2014).
11. Gkoutos, Georgios V., Jeffery E. Green, Ann-Marie Mallon, John M. Hancock, and Duncan Davidson. "Building Mouse Phenotype Ontologies." *Pacific Symposium on Biocomputing* 9 (2004): 178–89.
12. Husz, Zsolt L., Nicholas Burton, Bill Hill, Nestor Milyaev, and Richard Baldock. "Web Tools for Large-scale 3D Biological Images and Atlases." *BMC Bioinformatics* 13 (2012): 122. http://www. biomedcentral.com/1471-2105/13/122 [30] (Accessed February 28, 2014).
13. Kaufman, Matthew H. *Atlas of Mouse Development.* London: Academic Press, 1992.
14. Kaufman, Matthew H. "Postimplantation Development in the Mouse. Ciba Foundation Symposium No. 165." *Journal of Anatomy* 181 (1992): 170–71. http://www.ncbi.nlm.nih.gov/pmc/articles/PMC1259769/ [31] (Accessed June 11, 2014).
15. Richardson, Lorna, Shanmugasundaram Venkataraman, Peter Stevenson, Yiya Yang, Nicholas Burton, Jianguo Rao, Malcolm Fisher, Richard A. Baldock, Duncan R. Davidson and Jeffrey H. Christiansen. "EMAGE Mouse Embryo Spatial Gene Expression Database: 2010 Update." *Nucleic Acids Research* 38 (2010): 703–9. http://www. ncbi.nlm.nih.gov/pmc/articles/PMC2808994 [32] (Accessed February 28, 2014).
16. Theiler, Karl. *The House Mouse: Atlas of Embryonic Development.* New York: Springer, 1989.

The Edinburgh Mouse Atlas, also called the e-Mouse Atlas Project (EMAP), is an online resource comprised of the e-Mouse Atlas (EMA), a detailed digital model of mouse development, and the e-Mouse Atlas of Gene Expression (EMAGE), a database that identifies sites of gene expression in mouse embryos. Duncan Davidson and Richard Baldock founded the project in 1992, and the Medical Research Council (MRC) in Edinburgh, United Kingdom, funds the project. Davidson and Baldock announced the project in an article titled A Real Mouse for Your Computer, citing the need to manage and analyze the volume of data that overwhelmed developmental biologists. Though EMAP resources were distributed via CD-ROM in the early years, the project moved increasingly online by the early 2000s, and into the early decades of the twenty-first century, was in active development. EMAP can be utilized as a developmental biology teaching resource and as a research tool that enables scientists to explore annotated 3D virtual mouse embryos. EMAP's goal is to illuminate the molecular basis of tissue differentiation.

Subject

[Embryos](#) [33] [Databases, Genetic](#) [34] [Development](#) [35]

Topic

[Technologies](#) [36]

Publisher

Arizona State University. School of Life Sciences. Center for Biology and Society. Embryo Project Encyclopedia.

Rights

Format

[Articles](#) [37]

Last Modified

Tuesday, July 3, 2018 - 21:40

DC Date Accessioned

Thursday, June 12, 2014 - 00:15

DC Date Available

Thursday, June 12, 2014 - 00:15

DC Date Created

2014-06-11

- [Contact Us](#)

- The Embryo Project at Arizona State University, 1711 South Rural Road, Tempe Arizona 85287, United States

---

Source URL: https://embryo.asu.edu/pages/e-mouse-atlas-project-1992

Links
[1] https://embryo.asu.edu/pages/e-mouse-atlas-project-1992
[2] https://embryo.asu.edu/keywords/gene-expression
[3] https://embryo.asu.edu/keywords/emap
[4] https://embryo.asu.edu/keywords/emage
[5] https://embryo.asu.edu/keywords/duncan-davidson
[6] https://embryo.asu.edu/keywords/richard-baldock
[7] https://embryo.asu.edu/search?text=mouse
[8] https://embryo.asu.edu/search?text=differentiation
[9] https://embryo.asu.edu/search?text=cell%20lineage
[10] https://embryo.asu.edu/search?text=Australian%20National%20University
[11] https://embryo.asu.edu/search?text=University%20of%20Oxford
[12] https://embryo.asu.edu/search?text=Stanford%20University
[13] https://embryo.asu.edu/search?text=University%20of%20Edinburgh
[14] https://embryo.asu.edu/search?text=organization
[15] https://embryo.asu.edu/search?text=University%20of%20Zurich
[16] https://embryo.asu.edu/search?text=Columbia%20University
[17] https://embryo.asu.edu/search?text=fertilization
[18] https://embryo.asu.edu/search?text=egg
[19] https://embryo.asu.edu/search?text=embryology
[20] https://embryo.asu.edu/search?text=branchial%20arch
[21] https://embryo.asu.edu/search?text=genes
[22] https://embryo.asu.edu/search?text=embryogenesis
[23] https://embryo.asu.edu/search?text=developmental%20stages
[24] http://genomebiology.com/2005/6/4/108
[25] http://genome.cshlp.org/content/8/9/859.full

[26] http://dx.doi.org/10.1016/S0925-4773(98)00069-0

[27] http://www.nature.com/nbt/journal/v26/n3/full/nbt1391.html

[28] http://dev.biologists.org/content/118/4/1255.full.pdf

[29] http://www.emouseatlas.org/

[30] http://www.biomedcentral.com/1471-2105/13/122

[31] http://www.ncbi.nlm.nih.gov/pmc/articles/PMC1259769/

[32] http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2808994

[33] https://embryo.asu.edu/medical-subject-headings/embryos

[34] https://embryo.asu.edu/medical-subject-headings/databases-genetic

[35] https://embryo.asu.edu/medical-subject-headings/development

[36] https://embryo.asu.edu/topics/technologies

[37] https://embryo.asu.edu/formats/articles