

The Mouse Atlas Database: a community resource for mouse development

Biologists who study mouse development need to access and analyse large amounts of diverse information relating to gene function in the embryo. Gene expression patterns, mutant phenotypes, experimental results and descriptions of normal development all need to be integrated. The data that developmental biologists use in their analyses come in many different forms (tables, images, text) making them difficult to combine. To address this problem, we are developing the *Mouse Atlas Database* as a bioinformatics framework for gene-function data at the organism level¹. This public resource currently comprises two components: a controlled vocabulary to describe the developing anatomy and a series of digital 3D model embryos. Later the system will include a database of graphical gene-expression data. Building the system is an ongoing project, but, with the publication on the Internet of the controlled anatomical vocabulary, and the first model embryos on CD, there is now a resource that biologists can begin to use and comment on.

The *Mouse Atlas* CD can be used with Netscape 4.5 or Internet Explorer 4.0 (5.0 on Macs) (Boxes 1 and 2). The *Atlas* and other resources are available on the Internet at <http://genex.hgu.mrc.ac.uk/>, though the CD will work faster.

Controlled vocabulary for anatomy

The anatomical vocabulary² (with synonyms) includes all major structures that are visible in conventionally stained histological sections (i.e. sections from paraffin-embedded material stained with haematoxylin and eosin), but not domains that are only delineated by gene expression. The ~8000 items in the vocabulary cover all 26 Theiler developmental stages³ (E0–E17.5) and are organized in the form of a hierarchy of structures that specify 'part of' relationships. This is implemented in a web-accessible database with the hierarchy displayed as a series of lists, one for each Theiler stage (Fig. 1, Box 1).

Box 1. The Anatomy Database (<http://genex.hgu.mrc.ac.uk/Databases/Anatomy>)

The top-level page provides a description of the controlled vocabulary. This is available in XML, including unique identifiers for use as keywords, cross-referencing and linking. The data is also accessible through CORBA, which provides a powerful program interface. The primary links from this page are:

- Staging criteria: a detailed description of how the Theiler stages are defined and how we have extended the system to include other timing measures. The relationships between the systems are defined.
- Anatomy database: a controlled vocabulary for each stage. This is available as a text table that provides links to diagrams and text describing the stage. The controlled vocabulary can also be viewed using a Java browser, which displays an interactive

tree of the nomenclature with options for search and further detail. The viewers can be used to obtain both the name and the accession number for each component for precise cross-referencing.

- Search: enter a text sub-string or accession number to get details of matching components in the anatomy database.
- Notes: more detail on how the vocabulary was developed and the general principles of naming and grouping of terms.

The minimum requirement for viewing the database is a World-WideWeb browser. The Java interfaces require a browser that can include the Java 2 plug-in [available for MS Windows and Unix platforms now, Mac OS in due course (see <http://www.apple.com/java/>)]

Although the present version of the controlled vocabulary list covers all major tissues, researchers will probably want to make additions to the database or use different hierarchical groups. The authors are keen to include this expert knowledge and to make additions and corrections accordingly. Indeed, because no single hierarchical arrangement of parts will satisfy all requirements, software is being implemented that will allow the researcher to reorganize the same basic components into alternative hierarchies and groups. The authors welcome specific suggestions for such alternative views of the vocabulary.

3D model embryos

The second component of the *Atlas* is a series of digital 3D models of embryos. The aim is to provide at least one model for each successive Theiler stage (TS). At present, models for each post-implantation stage up to 9 days of development (TS7–14) are available on CD (Box 2). Later stages will be added as they are completed: models of embryos at 9.5 days (TS15), 12.5 days (TS20) and 17.5 days (TS26) of development are in progress (regularly updated draft material will be available at the web site through the online CD-ROM page). Each model shows the stained, histological

structure of the embryo with the major anatomical parts delineated^{4,5}. For embryos at 8.5 days of development (TS13) and older, the models will be based on the original paraffin-embedded specimens illustrated in *The Atlas of Mouse Development*⁶. For younger embryos, new plastic-embedded specimens have been used to achieve higher resolution. Each model embryo comprises an array of picture elements (voxels) whose dimensions define the spatial resolution; $2 \times 2 \times 2 \mu\text{m}$ for early stages (fertilization to 8 days TS1–12) and $\sim 4 \times 4 \times 7 \mu\text{m}$ for later stages. As a community resource, the CD is available to academic sites at nominal cost (see website). Because the *Atlas* aims to be an authoritative description of normal mouse development, feedback will be welcomed and anyone with expertise in a particular aspect of development who would like to contribute to the continuing process of annotating the models is encouraged to contact the authors.

Using the Mouse Atlas

The *Atlas* can be used as a laboratory bench tool. Together with the supporting staging guide, the digital models and controlled vocabulary can be used to help determine developmental stage and identify anatomical structures. The

Box 2. The CD-ROM (http://genex.hgu.mrc.ac.uk/CDROM_Copyrights/)

The CD is available to non-profit research labs at minimal cost (select 'Order a CD').

Browsing the CD on PCs, Macs or UNIX workstations requires a Java-enabled www browser such as Netscape 4.5 or Internet Explorer 4.0 (5.0 on Macs). These include, by default, the required version of Java (1.1.5). A simple help page is available for first-time users.

To use the advanced software, *SectionView* and *MAPaint*, a UNIX machine with X11 is necessary. The program will run on Sun Workstations (Solaris 2.6), Silicon Graphics Workstations (IRIX 5 or IRIX6) or PC compatibles (RedHat Linux 6.2 or Solaris 8)

A test page is available on the CD to help users see what can be run on their particular set-up.

models can also be used to interpret one's results by exploring the 3D relationship between parts of the embryo and by following their development from one stage to the next. For this, two software tools are available on the CD; these allow

the same models to be viewed either as whole embryos or as a series of histological sections. *SectionBrowser* provides a selection of pre-calculated section planes (Fig. 1), and *EmbryoView* allows the whole embryo to be visualized as if it were lying

in a dish illuminated from below (Fig. 2). Other basic features allow the user to pan through all sections in the three cardinal planes (although without anatomical annotation) and to view the original images of the transverse sections from which the models were built. These provide a higher-resolution view of the histology in each part of the embryo. These facilities are introduced in a simple help page for first-time users.

The facility to view the histological structure of a stained specimen in a variety of different ways with the aid of high-resolution anatomical annotation provides a powerful new tool for developmental biologists. Suppose, for example, one wishes to identify tissues labelled by a reporter gene in a whole-mount preparation. By choosing the appropriate-stage model in *EmbryoView*, interactive 3D views of the reference embryo that highlight landmark structures can be selected. These views can then be used to help determine which structures might be labelled in the specimen. For histological material *SectionBrowser* can be used to choose the cutting plane that corresponds to the relevant sections. As the cursor is moved over the section, structures and tissues are highlighted with their names displayed above the image. Clicking on a structure causes its position in the anatomical hierarchy to be highlighted. These features help identification of labelled expression domains, structures affected in mutants, and so on. In this context, the *Atlas* will also be useful in interpreting information from other databases. The obvious example is the GXD, the textual database that links mouse gene expression data to developmental anatomy⁷. As the *Mouse Atlas* and the GXD use the same anatomical nomenclature, movement from one to the other is straightforward. Thus, for example, the *Atlas* can be used to explore the spatial and developmental relationships between named tissues that express different genes as recorded in the GXD.

The *Atlas* will also be useful in writing papers, and we encourage developmental biologists to use the anatomical nomenclature to describe their results, in particular in tables, abstracts or as keywords. By doing so, researchers will automatically provide search terms for future database queries, and to facilitate this the *Atlas* database provides each

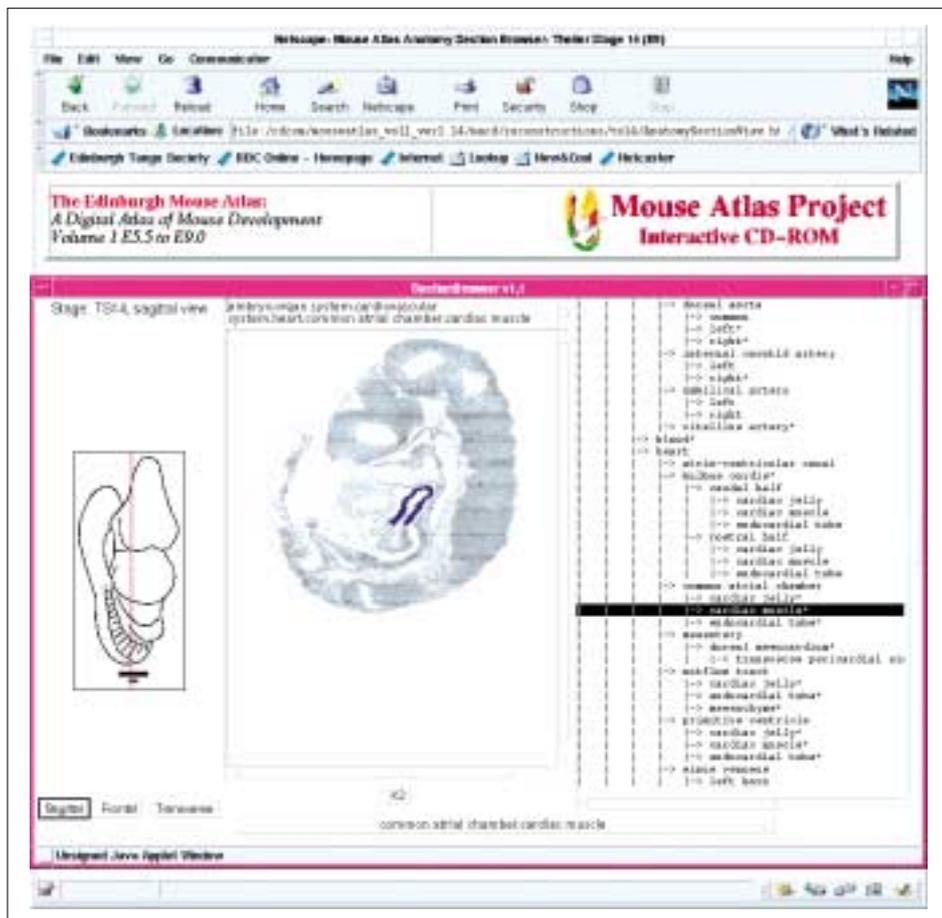


Fig. 1. The anatomy *SectionBrowser* allows the user to select digital sections through each embryo (Theiler Stages 7-14). This resource is designed to help the researcher explore the histological structure of the embryo and identify anatomical parts. For this a view is selected (lower left) and the red line is dragged to the desired position. The section is displayed in the left window and the anatomy database on the right. By moving the mouse cursor across the section, anatomical structures can be highlighted with their names appearing at the bottom of the window. Clicking on the section highlights the name of the corresponding structure in the database and vice versa. Structures that have been delineated in the model embryos and that appear in the current section are indicated by an asterisk in the database display. A simple list of these structures can be displayed by clicking on the 'select component' box at the top of the section window and individual structures selected for highlighting. Instructions are on the same page, below the part shown here.

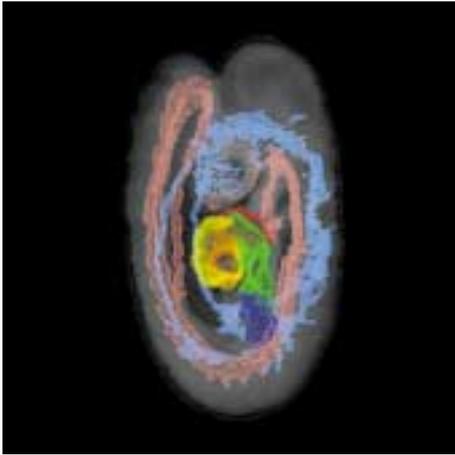


Fig. 2. An *EmbryoView* image of the Theiler stage-14 model enables the *Atlas* model to be displayed as if it were a whole embryo in a dish, but with key tissues highlighted; blue, venous system; pink, arterial system; red, outflow tract; yellow, primitive ventricle; green, common atrial chamber; purple, sinus venosus. This image can be rotated by moving the mouse cursor over the image. Both uncoloured and coloured versions are provided for each model embryo, the latter display selected anatomical structures. It is useful to have an *EmbryoView* window displayed while using the *SectionBrowser* shown in Fig. 1.

anatomical component with both a standard name and a unique identifier.

An additional role for the *Atlas* database will be to help integrate information from future studies of gene function. In particular, new databases relating to gene function in the embryo – for example, gene-expression results from microarrays and databases of mutant phenotypes – could use the standard anatomical vocabulary and model embryos to annotate results. The authors welcome suggestions for adaptations to the *Atlas* data that might be needed for these applications.

Scientists who want to use the CD as a more sophisticated research aid can use two advanced tools that run in a variety of UNIX environments (Box 2) and that can load and manipulate an entire model embryo. *SectionView* allows the user to cut sections interactively in any plane, for

example to find the best match of histological material from a particular experiment. The program allows the major anatomical structures to be highlighted in different colours in these sections and enables simultaneous exploration of several sections through the same model. *MAPaint* allows researchers to paint gene-expression patterns or other spatial domains directly into the model embryo, section by section, and to save these 3D domains as files that can be reloaded later or shared between labs. These programs can also be downloaded from the genex web page (<http://genex.hgu.mrc.ac.uk/>), but the model embryos are very large, and are therefore available only on the CD. Additional UNIX software is being developed that will allow data to be mapped into the *Atlas* models directly from images of experimental results. A preliminary version of this software can be obtained from the authors.

Future applications

The *Atlas* is designed to provide the bioinformatics framework for community databases of mouse development. To facilitate this, we are building a graphical database that can hold 3D information (e.g. data on gene expression, cell properties, mutant phenotype, etc. from whole-mounts or sections) spatially mapped onto the digital models by using the array of voxels as 'pigeon holes' in the database. (Guidelines on preparing gene-expression data for entry into this, and similar, databases can be found in Ref. 8.) Because the digital models and anatomical vocabulary are linked, it will be possible to query the database using combinations of spatial and textual criteria. It will thus be possible to find genes with expression patterns that are (as defined by user) similar, adjacent or complementary to the expression domain of a given gene, and to display their expression patterns graphically on the digital models.

We also plan to facilitate links with other databases, in particular the large amounts of textual gene-expression data in GXD, developed by our colleagues at the Jackson Laboratory. Indeed, we are collaborating with the Jackson Laboratory to create a 'mouse gene-expression information resource' that will provide easy and transparent access to both the *Mouse Atlas Database* and the GXD⁹. The facility to visualize and combine experimental data with information from this resource will, we hope, provide a powerful analytical tool for gene-function studies.

References

- Baldock, R.A.B. *et al.* (1999) The Edinburgh Mouse Atlas: basic structure and informatics. In *Bioinformatics: Databases and Systems*. (Letowsky, S. I., ed.), pp. 129–140, Kluwer Academic Publishers
- Bard, J.B.L. *et al.* (1998) An internet-accessible database of mouse developmental anatomy based on a systematic nomenclature. *Mech. Dev.* 74, 111–120
- Theiler, K. (1989) *The House Mouse: Atlas of Mouse Development*, Springer-Verlag
- Brune, R.M. *et al.* (1999) A three-dimensional model of the mouse at embryonic day 9. *Dev. Biology* 216, 457–468
- Kaufman, M.H. *et al.* (1998) Computer-aided 3-D reconstruction of serially sectioned mouse embryos. *J. Anat.* 193, 323–336.
- Kaufman, M. (1992) *The Atlas of Mouse Development*, Academic Press
- Ringwald, M. *et al.* (2000) GXD: integrated access to gene expression data for the laboratory mouse. *Trends Genet.* 16, 188–190
- Davidson, D. *et al.* (1998) Gene expression databases. In *In Situ Hybridization, a Practical Approach*, pp. 189–214, IRL Press (Oxford University Press)
- Ringwald, M. *et al.* (1994) A database for mouse development. *Science* 265, 2033–2034

Duncan Davidson* †

Jonathan Bard ‡

Matthew Kaufman ‡

Richard Baldock †

†MRC Human Genetics Unit, Western General Hospital, Edinburgh, UK EH4 2XU.

‡Dept of Biomedical Sciences, Hugh Robson Building, George Square, Edinburgh, UK EH8 9XD.

*e-mail: Duncan.Davidson@hgu.mrc.ac.uk

Letters to the Editor

We welcome letters on any topic of interest to geneticists and developmental biologists.

Please write to:

The Editor

Trends in Genetics, Elsevier Science London, 84Theobald's Road, London UK WC1X 8RR.

Tel: +44 (0)20 7611 4400 Fax: +44 (0)20 7611 4470; e-mail: tig@current-trends.com