ELSEVIER

Mechanisms of Development 74 (1998) 111–120

An internet-accessible database of mouse developmental anatomy based on a systematic nomenclature

Jonathan B.L. Bard[a,*], Matthew H. Kaufman[a], Christophe Dubreuil[b], Renske M. Brune[a], Albert Burger[c], Richard A. Baldock[b], Duncan R. Davidson[b]

[a]*Centre for Developmental Biology and Department of Anatomy, Medical School, Edinburgh University, Edinburgh EH8 9AG, UK*
[b]*MRC Human Genetics Unit, Western General Hospital, Edinburgh EH4 2XU, UK*
[c]*Department of Computing and Electrical Engineering, Heriot–Watt University, Edinburgh EH14 4AS, UK*

Received 24 February 1998; revised version received 21 April 1998; accepted 21 April 1998

Abstract

This paper reports an internet-accessible database of mouse developmental anatomy (DMDA) that currently holds a hierarchy of the names and synonyms of the tissues in the first 22 Theiler stages of development (E1–E13.5), together with other appropriate information. The purposes of the database are to provide, first, a nomenclature for analyzing normal and mutant mouse anatomy, and second, a language for inputting, storing and querying gene-expression and other spatially organized data. DMDA currently contains some 6900 named and staged tissues (e.g. 360 and 1161 tissues in Theiler stage (TS) 14 (E9) and TS22 (E13.5) embryos). DMDA will be extended to include further lineage and other data when it becomes available. The database can be interactively accessed over the internet using either a Java or a non-Java WWW browser at http://genex.hgu.mrc.ac.uk/.

Keywords: Anatomical database; Bioinformatics; Database; Mouse anatomy; Mouse development

1. Introduction

This paper describes a database of the major named tissues in the developing mouse embryo to provide what could be a standard nomenclature for analyzing both normal and mutant tissue anatomy. The production of the database derives from an appreciation of the very large amounts of gene-expression data that are being produced for the mouse. Such is the sheer volume of this data that it is hard to keep up with the flow, and the obvious solution is to complement the standard literature with a database of gene-expression. Storing this information does however demand a complete listing of the tissues present at each stage of mouse development so that there are unambiguous names for inputting and accessing domains of expression. This requirement has also been recognized by the neuroscience community with the development of the *NeuroNames Hierarchy* (Bowden and Martin, 1995) which is an analogous nomenclature for the human adult brain, and the brain maps of the mouse (Jacobowitz and Abbott, 1997) and rat (Alvarez-Bolado and Swanson, 1996).

This database is the first step in a four-stage enterprise intended to facilitate the study of the genetic basis of mouse development (Baldock et al., 1992). The second is the construction of a further text database for accepting and accessing gene-expression data, and this database, currently being developed at the Jackson Laboratory (Ringwald et al., 1994, 1997) will use the anatomy database outlined here for its terminology. The next stage is to superimpose the text descriptions of the developmental anatomy onto high resolution 3D embryo reconstructions that are being made in Edinburgh. Finally, a fully graphical database will be assembled to link gene-expression data with these embryo reconstructions and so provide a direct mapping of spatial data onto the embryo (Davidson et al., 1997). This is required to describe the many gene-expression domains in embryos that do not respect anatomically-defined tissue boundaries.

* Corresponding author. Tel.: +44 131 6503107; fax: +44 131 6506545.

0925-4773/98/$19.00
PII S0925-4773(98)00069-0

2. The anatomy database

This database was designed to provide both an anatomical hierarchy for the mouse that could be used as a standard reference and as an annotation system for a gene-expression database. For the former, we had to decide which tissues were present at each stage, and to arrange them in a hierarchical tree. For the latter, the spatial domains associated with these names had to fill the 3D volume of the embryo.

This last requirement raises the problem that many domains in an embryo do not have a formal name, either because it has not differentiated (e.g. regions of mesenchyme in the early embryo) or because a notable part of a tissue has a name but the rest has not. In the left atrium, for example, the auricular region is named, but the rest of this tissue is not. To meet the second requirement, 'pseudo names' such as 'unnamed part of left atrium' need to be included to ensure volumetric completeness, and the database has a facility for doing this.

2.1. Tissue identification

The lists of the tissues present in each Theiler stage were made on the initial basis of the index constructed for the *Atlas of Mouse Development*, although histological analysis of sectioned material showed that many corrections and additions were necessary (Figs. 1–3). Most names refer to space-filling components, although a few refer to landmarks or features (e.g. the sulcus limitans). The general working criterion for identifying a tissue was that it was recognizable morphologically under 100× magnification – exceptions were the neural crest and the somite derivatives. In constructing the stage lists, careful attention has been paid to ensure that differentiating tissues were given their new names as and when appropriate. Extraembryonic tissues have been excluded after TS12 because they are usually dissected away at this stage in order to expose the embryo.

The tissue lists for each stage are long (even for an early stage, see Fig. 2) as each somite and its derivatives are mentioned separately, and the names of all major blood vessels are given, where they can be distinguished. As it is expected to be used for storing data on signalling genes, the lists often include subordinate cell types within a tissue (e.g. the epithelial ridge and the underlying mesnchyme of the early limb bud). The tissue lists are thus far more extensive than those in the index of the *Atlas of Mouse* Development (Kaufman, 1994). One simplification is that, where there is an obvious right/left symmetry (e.g. limbs, the somites, ganglia, etc.), only a single name is given for the two parts. Exceptions to anatomical completeness include the smaller muscles and bones as it is often hard to determine when their condensations first form.

The lists cannot be comprehensive as any tissue may include smaller unnamed domains and subdivisions of different cell types. To allow for the inclusion of such detail, a datafile is associated with each tissue in which can be stored additional information about cell types, subordinate tissues and tissue architecture (we plan to include such descriptions).

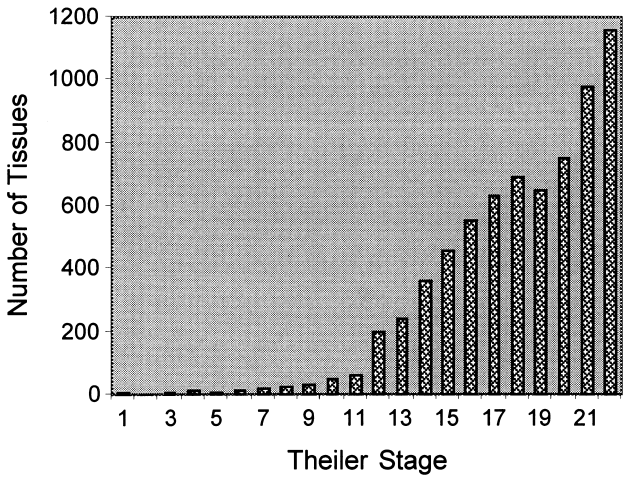

Fig. 1. The number of tissues present at each Theiler stage. The slight drop in number that occurs at TS19 is due to the loss of the main body somites and their dermomyotomes.

2.2. Naming

As different embryologists and anatomists have used different terms for the same tissue (e.g. *branchial*, *aortic* and *pharyngeal* arch), we have used what seems to be the most generally accepted anatomical term as the prime identifier, but have also included a list of synonyms that can be searched if a user cannot find their choice of tissue name. Where boundaries are imprecise (mainly in early embryos), we have named regions by their fate (e.g. future forebrain region). Moreover, some tissues change their names as they develop (*ectoderm* becomes *epithelium* and may become *skin*) and perhaps the most difficult of these problems is with the *mesoderm* that mainly becomes *mesenchyme*; we have taken a radical solution here, and from TS12 onwards, have abandoned the term mesoderm in favour of mesenchyme, even in the case of somites that are transitionally epithelial. Where changes of name are not obvious, we have included a note saying 'future <new name>' or 'previously <old name>' (these notes will be available in future releases of the database) and these comments are most important in the context of lineage (see below).

2.3. Tissue hierarchy

We have emphasized obvious components (the branchial arch system, the muscles, skeleton and glands) and the main organ systems (e.g. neural, vascular and visceral), with their components being organized within a 'parts of' hierarchy. Here, each tissue is assigned a unique name that starts with the Theiler stage, and goes through successive subdivisions of the embryo so that, for example, the superior glossophar-

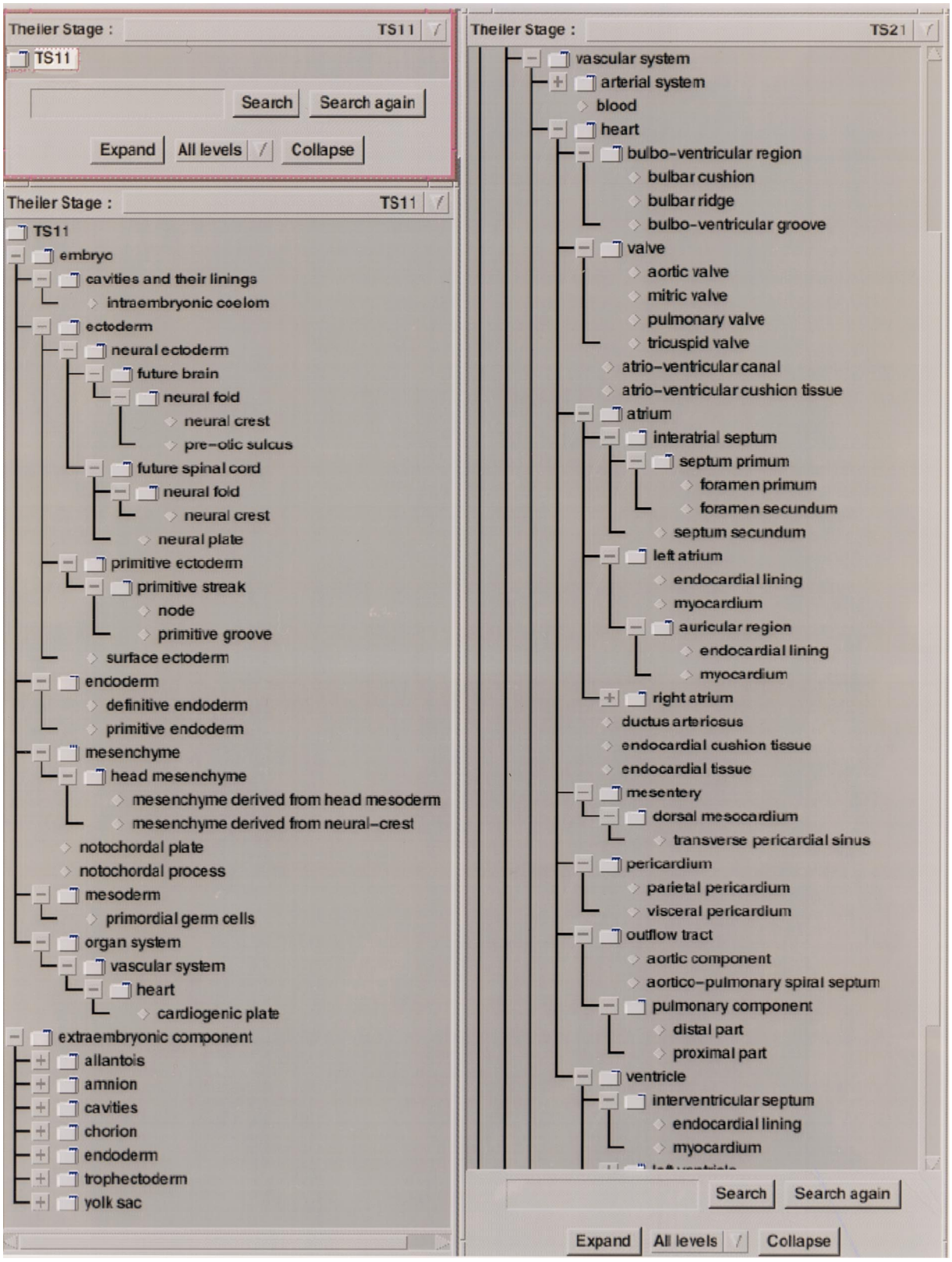

Fig. 2. Three views of the WWW Java interface for the anatomy database. Top-left and left show unexpanded and expanded versions of the anatomical nomenclature for the TS11 embryo, respectively. The figure on the right shows and an expanded view of the TS21 heart illustrating the depth and detail of the nomenclature. In each diagram a box with a plus sign indicates that that node can be expanded further.

Table 1
Expanded Theiler staging for mouse embryos

Theiler stage	Embryonic age (range)[a]	Somite no. (range)[b]	Cell no.	(C57BLxCBA)F1 mice[c]	PO mice[d]
1	0–0.9 (0–2.5)		1	**One-cell egg**	
2	1 (1—2.5)		1–4	**Dividing egg**	
3	2 (1–3.5)		2–16	**Morula**	
4	3 (2–4)		18–40	**Blastocyst**, inner cell mass apparent	
5	4 (3–5.5)			**Blastocyst** (zona-free)	
6	4.5 (4–5.5)			**Attachment of blastocyst**, primary endoderm covers blastocoelic surface of inner cell mass	
7	5 (4.5–6)			**Implantation and formation of egg cylinder** ectoplacental cone appears, enlarged epiblast, primary endoderm lines mural trophectoderm	
8	6 (5–6.5)			**Differentiation of egg cylinder**, implantation sites 2 × 3 mm, ectoplacental cone region invaded by maternal blood, Reichert's membrane and proamniotic cavity form	
9	6.5 (6.25–7.25)			**Pre-streak (PS)**, advanced endometrial reaction, ectoplacental cone invaded by blood, extraembryonic ectoderm, embryonic axis visible	PS
9a				**Early streak (ES)**, gastrulation starts, first evidence of mesoderm	ES
10	7 (6.5–7.75)			**Mid streak (MS)**, amniotic fold starts to form	MS
10a				**Late streak**, no bud (LSOB), exocoelom	LS
10b				**Late streak**, early bud (LSEB), allantoic bud first appears, node, amnion closing	
11	7.5 (7.25–8)			**Neural plate (NP)**, head process developing, amnion complete	OB
11a				**Late neural plate (LNP)**, elongated late allantoic bud	EB/LB
11b				**Early head fold (EHF)**	EHF
11c				**Late head fold (LHF)**, foregut invagination	LHF
12	8 (7.5–8.75)	1–4		**1–4 somites**, allantois extends, 1st branchial arch, heart starts to form, foregut pocket visible, preotic sulcus (at 2–3 somite stage), cephalic neural crest starts to migrate	
12a		5–7		**5–7 somites**, allantois contacts chorion at the end of TS12 ***absent*** *2nd arch, >7 somites*	
13	8.5 (8–9.25)	8–12		**Turning of the embryo**, 1st branchial arch has maxillary and mandibular components, 2nd arch present ***absent*** *3rd arch, >12 somites*	

14	9 (8.5–9.75)	13–20	**Formation and closure of anterior neuropore**, 13-–20 somite pairs, otic pit indented but not closed, 3rd branchial arch visible ***absent*** *forelimb bud*
15	9.5 (9–10.25)	21–29	**Formation of posterior neuropore**, forelimb bud, 21–29 somites, forebrain vesicle subdivides ***absent*** *hindlimb bud, Rathke's pouch*
16	10 (9.5–10.75)	30–34	**Posterior neuropore closes**, formation of hindlimb and tail buds, lens plate, Rathke's pouch; the indented nasal processes start to form ***absent*** *thin and long tail*
17	10.5 (10–11.25)	35–39	**Deep lens indentation**, tail elongates and thins, umbilical hernia starts to form ***absent*** *nasal pits*
18	11 (10.5–11.25)	40–44	**Closure of lens vesicle**, nasal pits, cervical somites no longer visible ***absent*** *auditory hillocks, handplate*
19	11.5 (11–12.25)	45–47	**Hand plate but no foot plate**, lens vesicle completely separated, auditory hillocks first visible ***absent*** *retinal pigmentation and sign of fingers*
20	12 (11.5–13)	48–51	**Earliest sign of fingers**, splayed-out digits, foot plate apparent, retina pigmentation apparent, tongue well-defined, brain vesicles clear ***absent*** *five rows of whiskers, indented anterior footplate*
21	13 (12.5–14)	52–55	**Anterior footplate indented**, elbow and wrist identifiable, five rows of whiskers, umbilical hernia now clearly apparent ***absent*** *hair follicles, fingers separate distally*
22	14 (13.5–15)	56– ~60	**Fingers separate distally**, only indentations between toes, long bones of limbs present, hair follicles in pectoral, pelvic and trunk regions ***absent*** *open eyelids, hair follicles in cephalic region*
23	15		**Toes separate**, hair follicles also in cephalic region but not at periphery of vibrissae, eyelids open ***absent*** *nail primordia, fingers two to five parallel*
24	16		**Reposition of umbilical hernia**, eyelids closing, fingers two to five parallel, nail primordia visible on toes ***absent*** *wrinkled skin, fingers and toes joined together*
25	17		**Skin is wrinkled**, eyelids are closed, umbilical hernia is gone ***absent*** *ear extending over auditory meatus, long whiskers*
26	18		**Long whiskers**, eyes barely visible through closed eyelids, ear covers auditory meatus

[a]Days post conception, with the morning after the vaginal plug is found being designated 0.5 dpc (or E0.5). For detailed discussion see Kaufman (1994).
[b]The figure given refers to the number of the most caudal somite. No account is taken of somites partitioning into dermomyotomes and sclerotomes, nor of their subsequent differentiation.
[c]Adapted from Theiler (1989) and Kaufman (1994); detailed staging for Theiler stages 9–12 courtesy of K. Lawson (pers. commun.)
[d]From Downes and Davies (1993).

yngeal ganglion of the Theiler stage 22 embryo has the full description (with synonyms in brackets): *TS22 (E13.5)* ← *embryo* ← *organ system* ← *nervous system* ← *central nervous system (CNS)* ← *ganglion* ← *cranial* ← *glossopharyngeal IX* ← *superior* where ← implies the relationship 'part of' or 'member of'. The developing embryo is thus described as a tree of anatomical structures whose root is the whole mouse embryo and that becomes successively finely divided into non-overlapping named parts that accordingly give ever more detailed information (see Fig. 2).

There is however no complete partition into head and body as there is no explicit, anatomical boundary between them, particularly in the early stages of development. The user interested in the former will have to consult separately the appropriate parts of the central nervous system, the sense organs, the skeleton and the musculature. One advantage of working within a database, however, is that the system allows users some freedom in constructing their own organization from the data using the notions of *groups* (see below).

2.4. Temporal development

To partition development, we have used the Theiler system, which bases its stages on developmental morphology rather than embryonic age, and, of course, imposes arbitrary temporal boundaries on what is a continuous process (see Theiler, 1989, and the anatomy web page). The list for each stage has therefore to contain all the components found at any point during that stage, and some tissues do of course appear or disappear mid-stage. Where such is the case (e.g. the otocyst (otic vesicle) is first apparent during late stage 15), a note to this effect is added to the tissue name (see Fig. 2).

The Theiler system bases its stages separations on the velocity of development: where this is slow, stages are at daily intervals, where it is fast, stages are separated by 12 h. The speed of development is particularly fast around gastrulation and the stages there have been further partitioned (see Table 1). For users unaccustomed to the Theiler system, basic information is given in Table 1, while more detailed morphology is available on the website.

While the Theiler staging system is well-established and useful (with the caveat that tissues in certain mouse strains exhibit their own temporal idiosyncrasies (e.g. at around gastrulation), see Table 1), it is wrong to assume that every part of an embryo develops at the same pace in every individual. Moreover, it is not always easy to recognize the exact stage at which a particular tissue can first be discerned. Users should thus be aware that the tissues identifiable in a particular staged embryo may differ slightly from those used for this database, although it is unlikely that the timing of the appearance or disappearance of a given tissue will be out by more than a single stage.

While developmental timing in the database is routinely assayed through Theiler stages, other systems have been used (e.g. days post-coitum and somite count), although only some span the full range of development. While the database uses the specific descriptions for Theiler stages, we recognize that expression data may be included that may not map exactly onto these stages, particularly in early development. To handle this difficulty, we have implemented a special staging object which will interpret any staging criterion in terms of a notional 'floating-point' Theiler stage. Table 1 shows the relationships between the different staging systems with some indication of the range of each alternative system with respect to the fixed Theiler stage.

2.5. Groups

Although the tissues at a given stage form a unique tree with branches and leaves (end points), the underlying organ–system-based hierarchy may not be the most convenient format for the viewer. First, the interface may contain more information than a user may want: a person interested in muscles, for example, would like access to all the muscles present at a given stage, without having the interface complicated by non-muscle tissues. Second, the hierarchy may not mesh with the needs of a user: we have, for example, defined the deltoid muscle as being part of the 'pectoral girdle', but it might more usefully be assigned to 'muscles' or the 'arm' or the 'shoulder region'.

To handle problems of this type, we are providing a *groups* facility, with groups being defined as collections of linked components from a particular stage. Some of these links are already included within the full name used for the hierarchy (e.g. ganglion, see Fig. 3), and it is planned to provide a facility whereby others can be added by the user. In this way, a tissue can be accessed under various headings: if the deltoid muscle is given additional links, it can be accessed under arm, shoulder and muscle group headings without generating any ambiguity.

2.6. Lineage

The database schema allows links between the stages and these permit a user to query the system about progenitor and derivative tissues. To facilitate this process, we have tried to keep the hierarchy description as consistent as possible across stages, although, as development proceeds, tissues change their name (e.g. the neural tube becomes the spinal cord). Where the name stays the same, the lineage relationships are included automatically. Where the name changes, the lineage has to be defined 'by hand' and much of this work still has to be done.

Currently, the database software allows the user to follow the progression of a tissue in two ways. First, the *search* facility allows the user to display on the screen, all entries in the database whose full description contains a specific name. Second, the user can simultaneously display windows

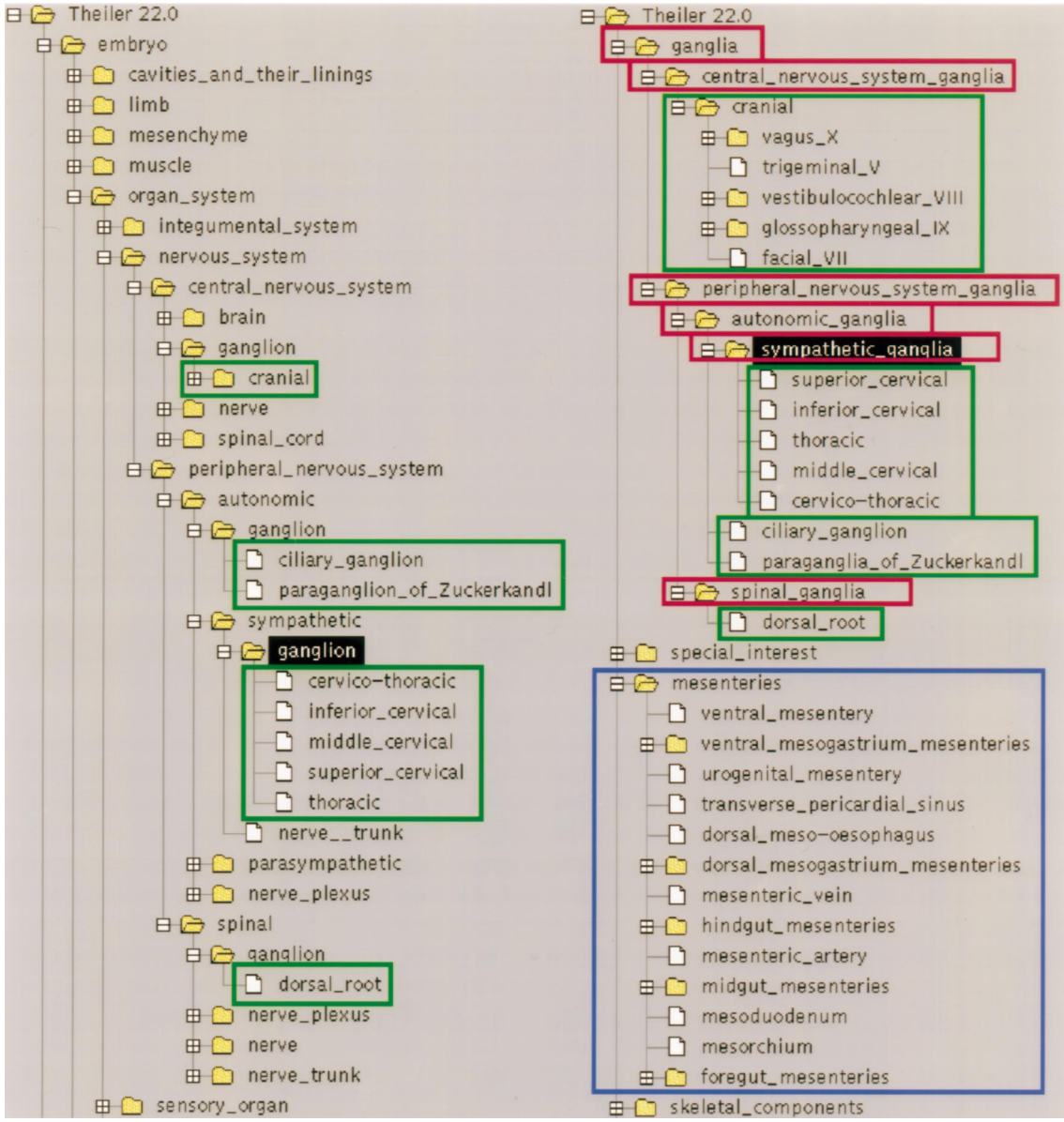

Fig. 3. Example groups defined for TS22. The left diagram shows the original tree with components highlighted by green boxes that are the components of the *ganglia* group and sub-groups. The corresponding original objects are similarly marked on the right hand side. The right figure highlights in red the corresponding *ganglia* groups with a structure intended to match that of the original nomenclature. Note the component *cranial* included in the group implies that all sub-parts of *cranial* are also part of the group. The blue highlight shows an example of a group, *mesenteries*, in which the original nomenclature hierarchy is not preserved.

for the hierarchies of several Theiler stages on the screen, and use these to follow the development of a particular organ system.

3. The WWW interface

Entry to the database interface is through the internet, through a server maintained at the MRC Human Genetics Unit (Edinburgh, UK). The website is intended to be a general tool for information on mouse embryogenesis and currently includes access to the following items.

3.1. The anatomy database with Java – an interactive tree

This starts with a menu of Theiler stages, each of which brings up 'buttons' for displaying and expanding an anatomy tree (see Fig. 2). Clicking a button associated with an open tree causes the tree to close, and the screen to be redrawn so that the vacated space is reused (this is important as the fully opened tree for a late stage covers many screensworth of space). This format will also be used for accessing alternate views of the hierarchies, secondary information, and lineage data. The software allows the user to display as

many stage windows as are needed so that lineage can be followed.

3.2. The anatomy database without Java – structured lists

This basic version of the database allows the user to browse through the complete mouse developmental anatomy at any stage.

3.3. Search

This displays on the screen, all cases in the database where a specific name is mentioned. The search can be restricted to a range of Theiler stages.

3.4. Notes on the standardized nomenclature

These include details of how the databases were assembled and copyright restrictions.

3.5. Staging criteria

This gives a page containing the data included in Table 1.

3.6. Diagrams and descriptions of embryos at each Theiler stage

This includes a page of drawings of staged embryos, each of which can be expanded to give structural and temporal information.

3.7. Current synonyms defined in the database

This page lists common alternatives for a particular anatomical term, and can be searched.

4. Discussion

4.1. Limitations on the anatomical data

One limitation of the tissue database is that only haemotoxylin- and eosin-stained or Toluidine Blue-stained material has been used for analyzing when tissues first appear, and the data is thus based on simple morphological criteria rather through the use of molecular markers. Although tissues generally show molecular changes before they become morphologically distinct, we have chosen simple morphology as our yardstick as, first, it is relatively uncontentious; second, it is the basis for tissue naming; third, it can readily be checked by any user; and finally, the morphology is time invariant, and not liable to regular review as new markers are discovered.

4.2. The advantages of a database format

It should be emphasized that storing the anatomical hierarchy in a database rather than in a list provides benefits beyond the presentational advantages given by the Java viewer. First, it allows the user to choose which parts of the data should be viewed; second, it enables the components to be re-grouped and so provides alternative views of the hierarchy; and third, the database can be programmed to display lineage relationships. In addition, further data (e.g. on lineage, cell types, tissue architecture, etc.) can readily be incorporated into the database. The use of the web interface brings the advantage that peripheral information on mouse development in any format that becomes available can readily be included through additional hypertext links.

Perhaps the most interesting aspect of the database, however, will be through interoperability with other, related databases, and for this we have adopted (see below) the CORBA standard. The IDL defined for this database could act as a prototype standard for anatomical nomenclatures for other species and provides the basic mechanism for true database interoperability.

4.3. Future versions

The database is being released in an incomplete state, partly because the anatomy lists up to TS22 (E13.5) covering the key stages needed by most mouse developmental biologists are now usable, partly because colleagues who know of the work have asked for access, and partly because the database is the anatomical component of the text-based, gene-expression database (GXD) being made at the Jackson Laboratory (Ringwald et al., 1994; Ringwald et al., 1997).

Two sorts of upgrades are planned, an expansion of the information within the database, and the establishment of links between it and other databases. The expansion will include the tissues from TS23 to TS26 (this will involve more than doubling the size of the current database), together with the list of attributes (e.g. lineage, cell type, images, subcellular and extracellular details) associated with each tissue, although collecting and entering such data will take time, and material will have to be added piecemeal (a link in the front page will inform users of any recent upgrades). We will provide links to other relevant databases through the CORBA interface as this interoperability between databases will enable the full value of the data collected in the myriad of database systems to be realized.

4.4. Further uses of the database

It was mentioned in the Section 1 that this database is the first part of a programme to make a text and graphical gene-expression database of mouse development. For the text component, a standard and precise language is needed to describe the relationship between gene expression and developmental anatomy so that there is an appropriate means for inputting and storing this information in the databases, as well as a set of terms for retrieving information.

The graphical component will require 3D reconstructions (e.g. Kaufman et al., 1996) with their tissues delineated, and the terms in the anatomical database are being matched to these tissues.

It is also likely that this database will have other uses in, for example, analyzing the phenotypes of mutant mice by providing a checklist of anatomical keywords that should be present at particular stages. The terminology can also be used for searching the literature and other databases, for constructing similar anatomical databases for other vertebrates, and attaching any anatomically-defined data (e.g. about developmental mechanisms) to databases. We are therefore happy to allow free access within the copyright limitations given in the database.

In short, this database has been designed as a tool for the developmental biology community and we hope that they will both use it, and enjoy the richness of developing vertebrate anatomy. It is however inevitable that the database will contain errors of omission, category and timing and the authors would appreciate feedback both about these and about any new features that users would like to see incorporated into future releases of the database. Comments on these or any other matters should be emailed to the database curator: j.bard@ed.ac.uk.

5. Experimental procedures

5.1. The anatomical data

The mouse embryos used for this work were isolated from (C57BLxCBA) F1 hybrid females previously mated to similar F1 hybrid males. Older embryos were routinely fixed, processed to wax, serially sectioned at 7 μm and stained with haemotoxylin and eosin (Kaufman, 1994), and were those previously analyzed in *The Atlas of Mouse Development* (Kaufman, 1994). Analyses of the early embryos (up to TS12) are based on plastic-embedded embryos that had been serially sectioned at 2 μm and stained with Toluidine Blue.

5.2. The database

This database is designed to be able to stand alone, but has been constructed on the basis that it will provide the reference, indexing and querying mechanism for a stage-dependent gene-expression database that includes 3D image data as well as purely textual entries. We have therefore adopted an object-oriented (OO) design (Booch, 1994 implemented within the commercial database management system *ObjectStore* (Object Design, Bracknell, UK) as it allows the complex structure of the anatomy and gene-expression data to be readily matched by the database schema. In such a database, each entry (which can, for example, be an anatomical component or a 3D binary image) is viewed as a free-standing object that has links to other objects rather than represented within a set of relational tables that the traditional relational database system uses.

The nomenclature has the structure of a complex semantic network with spatial, temporal, physiological and cell morphology links. The first 22 Theiler stages currently include about 10 000 components with about 2500 spatial links and 6000 temporal (lineage) links.

5.3. Database access

The anatomy database is a key part of the Mouse Gene Expression Information Resource (MGEIR) and its tissue names will not only be referenced for this purpose, but are also likely to be used for linking data to other key species (Davidson and Baldock, 1997; Davidson et al., 1997). To facilitate this interoperability, we have incorporated the common objects request broker architecture (CORBA) protocols (OMG, 1995; Orfali et al., 1996), an industrial standard for database interoperability (Booch, 1994), and one that has also been adopted by the European Bioinformatics Institute. CORBA defines access to the database via a mechanism that is independent of the machine architecture (type and operating system) and database management system.

With the CORBA interface, the database is made available to a user as a set of ‘objects’ which comprise the underlying data and as a set of operations (methods) that can be performed on the data (e.g. queries of the dataset). The structure of these objects and the methods that can be applied are defined in a published interface specification which is written in COBRA’s interface definition language (IDL). Given this IDL specification, a user is then able to define a new query interface to the anatomy database or to link it to other information (or databases).

A key data-member of each object (e.g. anatomical component) in the database is the unique identifier (UID). This identifier can be used by other database systems as a reference ID, thus enabling the data to be exported and so allows a measure of interoperability.

5.4. Database organization

The information for each tissue is split into two parts, one of which holds all stage-independent information (e.g. the tissue name itself and its sub parts), while the other holds the time-dependent information (e.g. when that component first appears and its tissue attributes (e.g. cell types)). The time-independent part is a hierarchy holding all the anatomy of all stages of development and is referred to as the *abstract mouse*. Within the database, each component in the anatomy hierarchy holds a list of object links, or *relationships*, which provide direct access to the other higher and lower components. Furthermore, each time-independent component is linked to one or more time-dependant components which also holds links to progenitor and deri-

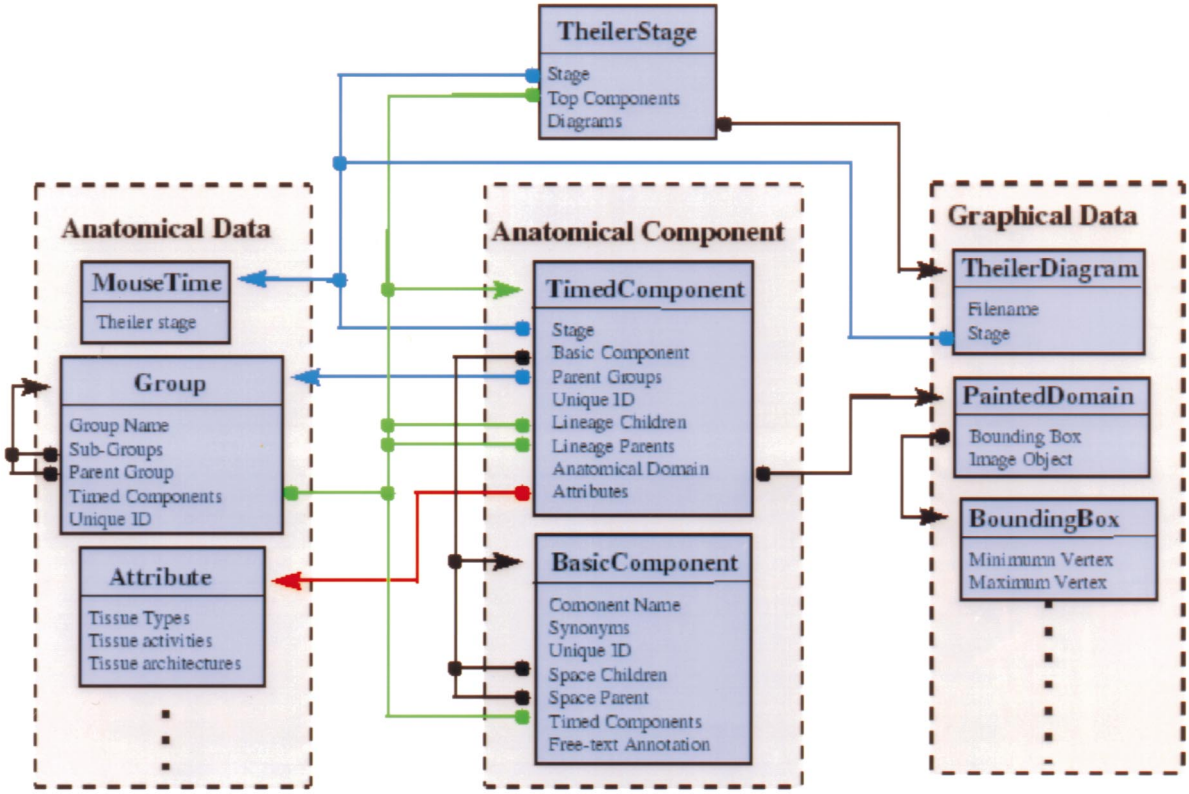

Fig. 4. The main objects in the anatomy database. A diagrammatic representation of the database structure and of how the different objects are linked (this logical view of the database structure shows a selection of relationships and the main objects, but does not represent the OO schema). The boxes with solid outlines are part of the database schema, while those with dashed outlines show how the data can be partitioned into the anatomical nomenclature (components), supplementary anatomical information and graphical data. Most of the data members of each object are complex data types as indicated by the connecting lines.

vative tissues (i.e. lineage data), and other data discussed below (see Fig. 4).

Acknowledgements

We thank our colleagues Kirstie Lawson, Margaret Stark, Allison Ross and Bill Hill for help. We also thank Martin Ringwald, Janan Eppig, Alex Smith and Joel Richardson of the Bioinformatics Laboratory at The Jackson Laboratory, for many useful discussions. We are grateful to the BBSRC for supporting part of this work.

References

Alvarez-Bolado, G., Swanson L.W., 1996. Developmental Brain Maps: Structure of the Embryonic Rat Brain. Elsevier, Amsterdam.

Baldock, R., Bard, J., Kaufman, M.H., Davidson, D., 1992. A real mouse for your computer. BioEssays 14, 501–502.

Booch, G., 1994. Object-Oriented Analysis and Design with Applications second ed. Benjamin Cummings, Menlo Park, CA.

Bowden, D.M., Martin, R.F., 1995. Neuronames brain hierarchy. Neuroimage 2, 63–83.

Davidson, D.R. and Baldock, R.A., 1997. A 3D atlas and gene expression database of mouse development: implications for a database of human development. In: Strachan, T., Lindsay, S., Wilson, D. (Eds.), Molecular Genetics of Early Human Development. BIOS Scientific, Oxford.

Davidson, D., Bard, J., Brune, B., Burger, A., Dubreuil, C., Hill, W., Kaufman, M., Quinn, J., Stark, M., Baldock, R., 1997. The mouse atlas and graphical gene-expression database. Semin. Cell Dev. Biol. 8, 489–498.

Downes, K.M., Davies, T., 1993. Staging of gastrulating mouse embryos by morphological landmarks in the dissecting microscope. Development 118, 1255–1266.

Jacobowitz, D.M., Abbott, L.C., 1997. Chemiarchitectonic atlas of the developing mouse brain. CRC Press, Boca Raton FL.

Kaufman, M.H., 1994. An index of mouse developmental anatomy. The Atlas of Mouse Development, Academic Press, London, pp. 515–525.

Kaufman, M.H., Brune, R.M., Baldock, R.A., Bard, J.B.L., Davidson, D., 1996. Computer-aided 3D reconstruction of serially-sectioned mouse embryos: its use in integrating anatomical organization. Int. J. Dev. Biol. 41, 223–233.

OMG, 1995. The Common Object Request Broker Architecture and Specification. Object Management Group Publications, Framingham, MA.

Orfali, R., Harkey, D., Edwards, J., 1996. The Distributed Objects Survival Guide. Wiley, New York.

Ringwald, M., Baldock, R., Bard, J., Kaufman, M., Eppig, J., Richardson, J.E., Nadeau, J., Davidson, D., 1994. A. database for mouse development. Science 265, 2033–2044.

Ringwald, R., Davis, G. L, Smith, A.G., Trepanier, L.E., Begley, D.A., Richardson, J.E., Eppig, J.T., 1997. The mouse gene-expression database GXD. Semin. Cell Dev. Biol. 8, 509–518.

Theiler, K., (1989). The House Mouse: Atlas of Embryonic Development. Springer–Verlag, New York.