EMAGE: database structure and rules of the annotation and querying language.
20 April 2005

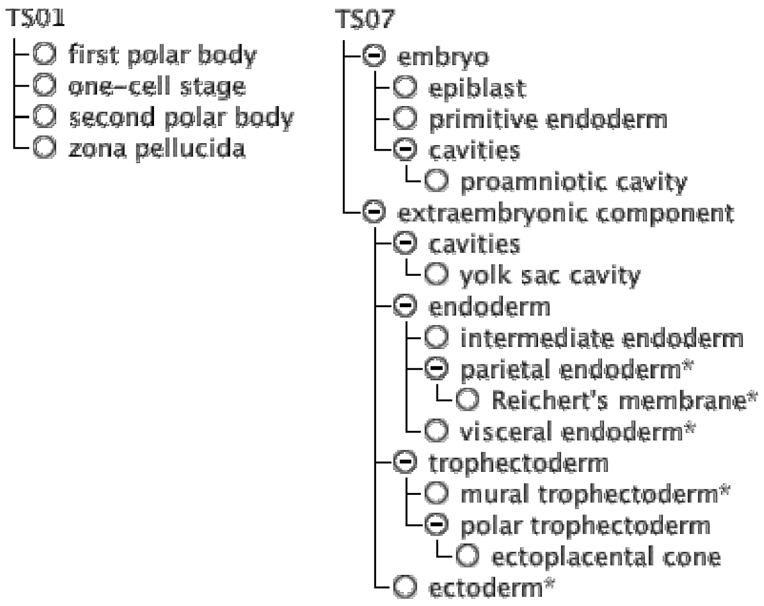
Section 1 Structure of the framework that houses EMAGE data.

The framework in which all EMAGE data is housed, is the EMAP Digital Atlas of mouse development and this has two parts: a text-based part (an ontology of anatomical terms) and a spatially based part (a set of 3D virtual embryo models).

1.1 Text based anatomy in the EMAP Digital Atlas.

The ontology of anatomical terms contains standardised words used to describe anatomical structures present throughout development. As of April 2005, there were approximately 26000 terms in the database, describing visible anatomical structures that are present from fertilisation right through to birth.

This list of ~26000 terms is subdivided into smaller, independent groupings for each of the 26 Theiler stages from fertilisation to birth. For example, at TS01 (the one cell stage embryo), there are 4 terms in the list because only 4 structures are visible at that stage: the first polar body, the second polar body, the zona pellucida and the one-cell stage embryo (see Fig 1).



```
TS01
├○ first polar body
├○ one-cell stage
├○ second polar body
└○ zona pellucida

TS07
├⊖ embryo
│ ├○ epiblast
│ ├○ primitive endoderm
│ └⊖ cavities
│   └○ proamniotic cavity
└⊖ extraembryonic component
  ├⊖ cavities
  │ └○ yolk sac cavity
  ├⊖ endoderm
  │ ├○ intermediate endoderm
  │ ├⊖ parietal endoderm*
  │ │ └○ Reichert's membrane*
  │ └○ visceral endoderm*
  ├⊖ trophectoderm
  │ ├○ mural trophectoderm*
  │ └⊖ polar trophectoderm
  │   └○ ectoplacental cone
  └○ ectoderm*
```

Fig1 Examples of the anatomy ontology at two different Theiler Stages.
The ontology is shown at TS01 and TS07. At TS01, four terms are in the list corresponding to the visible structures present. By TS07, the number of visible structures has increased considerably. The list is therefore presented in a structured tree format in order to organise the data.

By TS07, the embryo has developed to include several discernable structures: the embryonic part (which includes the epiblast, primitive endoderm, and proamniotic cavity) and an extra-embryonic part (which contains the extraembryonic ectoderm, the mural trophectoderm, polar trophectoderm, ectoplacental cone, intermediate endoderm, parietal endoderm, Reichert's membrane and the yolk sac cavity).

When the complexity of the embryo increases, such that one part has various sub-parts (eg at TS07 the embryonic part contains 3 subparts: the epiblast, primitive endoderm, and the proamniotic cavity), the terms are organised in a tree, with some components containing two or more sub-components (see Fig1).

In the anatomy terms database, each of the ~26000 terms (or 'components') has a unique ID. This ID is unique for the term at that stage (for example the term 'embryo' at TS07, has a different ID from the term 'embryo' at TS20 because these are different anatomical structures). In the underlying database, it is the relationships between these IDs that govern the tree format that the user sees.

1.1.1 Structure of the ontology trees.

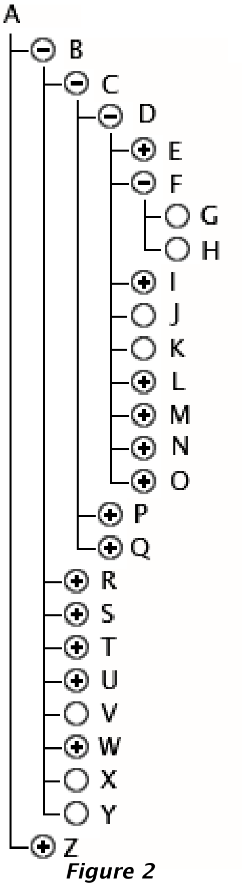


Figure 2

In Figure 2, a tree is displayed. 'A' is the highest level component in the ontology at this stage (eg. Stage ## mouse embryo) and it has two children: B and Z (eg. an embryonic part and an extra-embryonic part).

B has been 'opened' (denoted by the '-' symbol) to reveal its nine children (C, R, S, T, U, V, W, X and Y). These nine children of B are also the grandchildren of A. Of these nine, six (C, R, S, T, U and W) have children themselves (the '+' symbol denotes that it has children that can be viewed should that branch be 'opened' to reveal the next level down).
Going one level down the tree, C has three children: D, P and Q (which are also the grandchildren of B and great-grandchildren of A). D has been opened to reveal its nine children: E, F, I, J, K, L, M, N and O.
One more level down, F has been opened to reveal its two children: G and H. These are both at the end of a 'branch' of the tree (ie. they have no children) and are called 'leaves' (denoted by the the empty circle).

This arrangement of the components in the tree, shows ‘part-of’ relationships. (ie. G and H are ‘part-of’ F. F is ‘part-of’ D, D is ‘part-of’ C. Everything (apart from A itself) is ‘part-of’ A etc).

The complete set of sub-components (or children) of any component are intended to be *non-overlapping* and *complete*. For example, in Fig 1, epiblast, primitive endoderm and cavities are *non-overlapping* structures and when added together, constitute the parent term (embryo) *completely*.

Note: The current default structures of the trees, are the outcome of one line of thought on how the data should be presented. For example in Figure 1, the first branching is into two components: ‘embryo’ vs. ‘extra-embryonic’ with the next level down dividing each of these structures into ‘endoderm’ and ‘ectoderm’. An equally valid representation of the tree could be to first divide TS07 into ‘endoderm’ and ‘ectoderm’ and then divide each of these into ‘embryonic’ and ‘extra-embryonic’. To accommodate such other views of the tree, we are currently implementing a system of ‘groups’ in the anatomy ontology. In this case ‘ectoderm’ and ‘endoderm’ could be considered to be ‘groups’ of structures with both ‘embryonic’ and ‘extra-embryonic’ parts.

1.2 Structure based anatomy in the EMAP Digital Atlas.

The other part of the EMAP Digital Atlas is a set of 3D virtual embryo models. These are each made from 10's of 1000's of voxels (or 'volumetric pixels') that are stacked in 3D space. (See Fig 3 below).

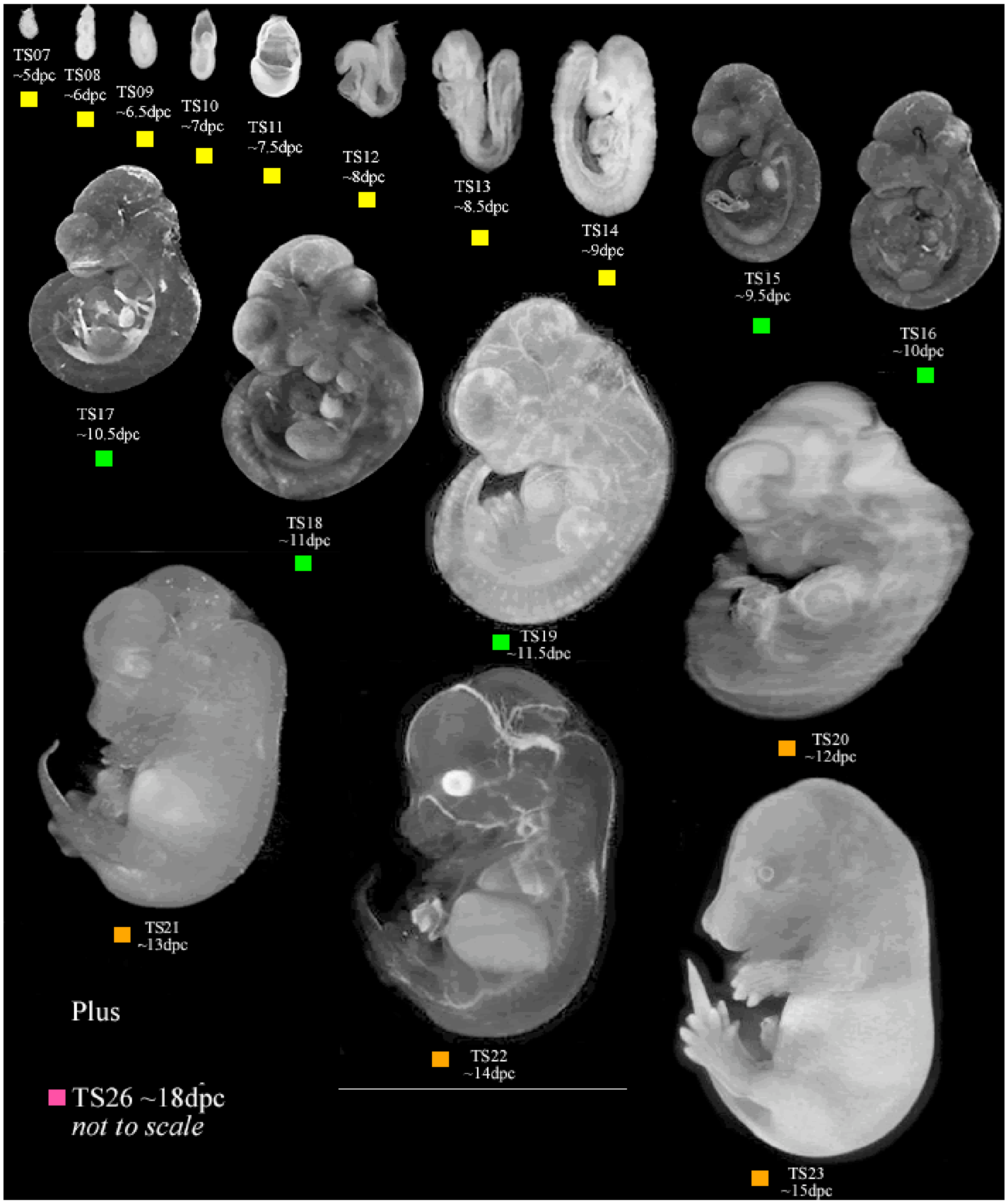


Figure 3 The EMAP 3D virtual embryo models.
Each of the embryo models are made of voxels, stacked in 3D space. Some of these embryo models have been incorporated into the framework for annotation of gene expression in EMAGE.
Key: Yellow – Embryo models with fully defined 3D anatomical domains. All of these are offered in EMAGE for 3D and wholemount spatial annotation. Green – embryo models without 3D anatomical domains defined within them. All of these are offered in EMAGE for both 3D and wholemount annotation. Orange – embryo models without 3D anatomical domains defined within them. These are not currently offered in EMAGE for any form of spatial annotation. Pink – embryo model without 3D anatomical domains defined within them and without a volume rendered view. This model is not currently offered in EMAGE for any form of spatial annotation.

Whereas the aforementioned anatomy ontology covers all of the 26 Theiler stages from fertilisation to birth, 3D embryo models have been produced for only a sub-set of these stages: TS07-23 and TS26 (ie. most post-implantation stages of development).

A further subset of these models has 3D anatomical domains defined within them (ie. TS07-14. These are denoted with a yellow square in Fig 3). These domains are also *non-overlapping* and *complete*. That is, each and every voxel of each of these 3D embryo models have been grouped into smaller groups or 3D regions that *do not overlap*, and when added together, make up the full 3D space *completely*.

For example, the TS07 model has a total volume of 251942 voxels. Contained within this volume are 9 smaller *non-overlapping* 3D regions, which *entirely* constitute the whole structure. These are 'embryonic ectoderm' (42755 voxels), 'extra-embryonic ectoderm' (47570 voxels), 'embryonic primitive endoderm' (36297 voxels), 'extra-embryonic visceral endoderm' (58461 voxels), 'proamniotic cavity' (2704 voxels), 'Reichert's membrane' (18915 voxels), 'parietal endoderm' (13917 voxels), 'mural trophectoderm' (6499 voxels) and 'polar trophectoderm' (24824 voxels) (see Fig 4).

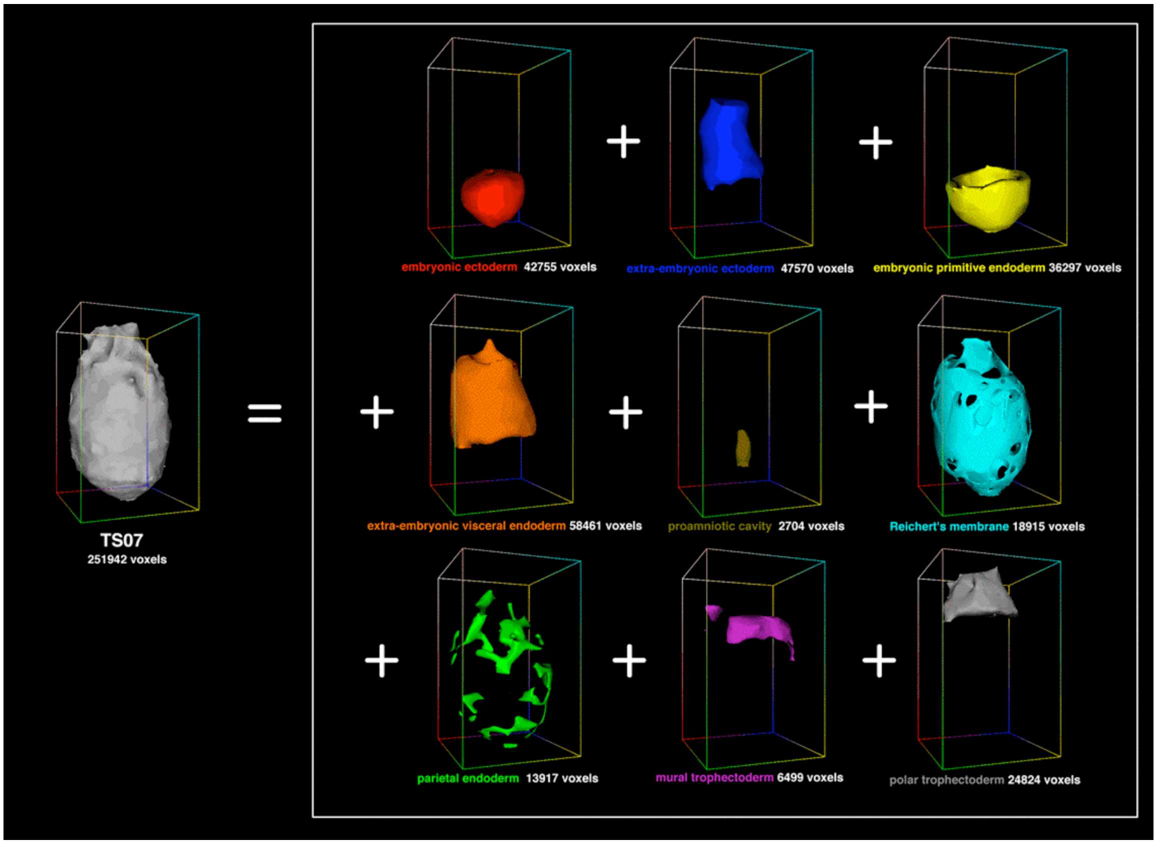

TS07
251942 voxels
=
embryonic ectoderm 42755 voxels
+
extra-embryonic ectoderm 47570 voxels
+
embryonic primitive endoderm 36297 voxels
+
extra-embryonic visceral endoderm 58461 voxels
+
proamniotic cavity 2704 voxels
+
Reichert's membrane 18915 voxels
+
parietal endoderm 13917 voxels
+
mural trophectoderm 6499 voxels
+
polar trophectoderm 24824 voxels

Fig 4 The constituent 3D anatomical domains of the TS07 embryo model
The nine constituent regions are non-overlapping and completely make up the 3D volume of the TS07 model.

1.3 Linking the text and structure-based anatomies.

The 3D regions in the virtual embryo models that have been defined as corresponding to a particular anatomical structure always correspond to one (and only one) term in the anatomy ontology for that stage. These can correspond to either a leaf in the tree (ie. a term with no children), or a term that has children.

In Fig 5 below, the 3D region that corresponds to component F is shown in red and the 3D region that corresponds to component R in yellow. In this case both of these terms in the tree have children (the two children of F (G and H) are shown).

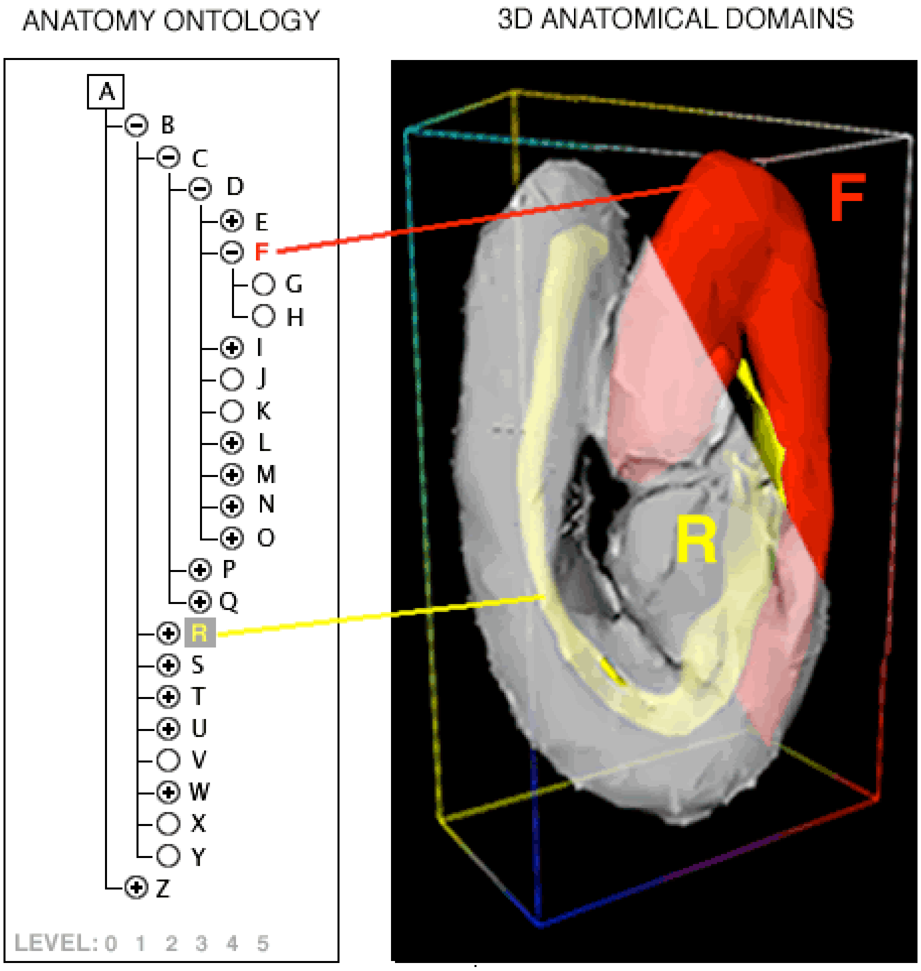


Fig5 The text and structure based anatomies are linked.

Note that the anatomical tree for any particular Theiler stage lists all the components that can be found over the interval represented by that stage. In contrast, each 3D model embryo, being based on one actual specimen represents only a single time-point during the stage. Thus some anatomical structures are not represented in the models. For example, while the Theiler Stage 14 anatomy tree contains both the otic placode and the otic pit that develops from it during TS 14, the 3D model has only an otic pit.

Section 2 Methods of annotating sites of gene expression in EMAGE into the EMAP Digital Atlas framework.

In EMAGE, sites of gene expression can be annotated into the EMAP Digital Atlas framework using either a text or spatial based method (or both).

2.1 Text annotation in EMAGE

This is performed manually by the person annotating and is based on the decision of the annotator as to what level the gene is expressed at. This can be performed for any Theiler stage from TS01-26.

2.1.1. Rules observed during text annotation

One of four levels of gene expression may be used to annotate any single component: ‘strong’, ‘moderate’, ‘weak’ and ‘possible’ (Note: ‘possible’ is used in circumstances when the annotator cannot decide if the expression is real or not). The colour convention used in all EMAGE annotation (be it text or spatial) to signify these levels are: red for ‘strong’, yellow for ‘moderate’, blue for ‘weak’ and green for ‘possible’. Terms that have been directly annotated (as text) by a person are shown in the interface in CAPITALS (see Fig6).

‘Strong’, ‘moderate’ and ‘weak’ together constitute ‘detected’ or ‘present’, whereas ‘possible’ is treated separately in the database.

The rule when using any of these levels to annotate is that they signify that expression is detected *somewhere* in the component ie. in at least one cell of the component. A term describing the pattern of expression can also be attached (homogeneous, graded, regional, spotted or single cell. ‘not applicable’ is the default). As the ontology is arranged in a hierarchical ‘part-of’ manner, when a component is annotated as having expression detected, by definition this means that expression is also detected in all of its parents, grandparents etc up to the top of the tree (see Fig6).

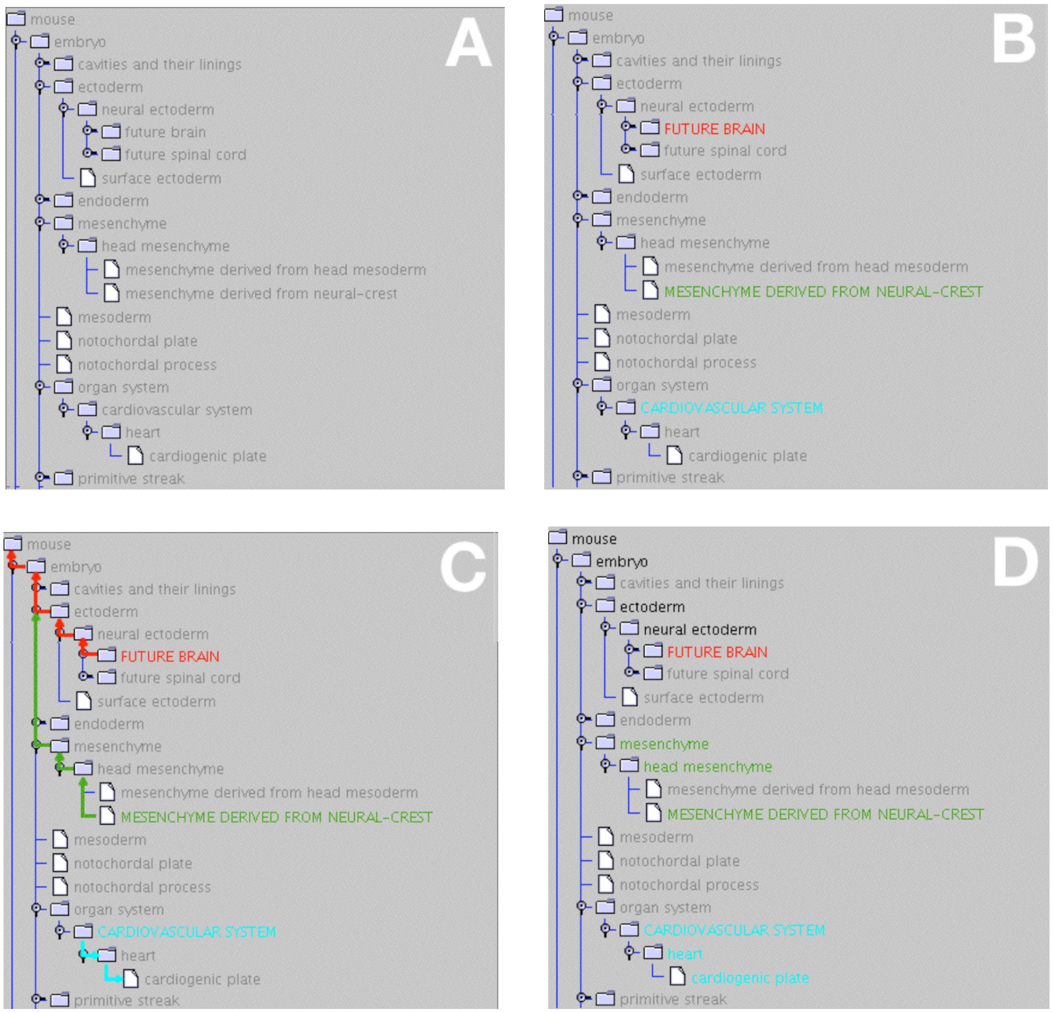


Fig6 Example of annotation of gene expression levels to the text ontology.
(A) Example of an un-annotated ontology tree. The terms are shown in grey. (B) The term 'future brain' has been directly annotated by a person as having strong expression (red), the 'mesenchyme derived from neural crest' as having possible expression (green) and the term 'cardiovascular system' as having no expression detected (cyan). Direct annotation is indicated by capital lettering. (C) As the tree is organised in a hierarchical manner with 'part-of' relationships, the annotation is propagated up or down the tree according to the general 'detected somewhere in' and 'not detected everywhere in' rules. The logic of this propagation is shown by the arrows. (D) The resultant annotation showing the inferred sites of expression. These are shown in lower-case coloured text. When the original term is manually annotated as having 'strong', 'moderate' or 'weak' expression (eg. future brain in this example) the inferred terms with expression 'present' include all the parents, grandparents etc of 'future brain' up to the top of the tree. These are shown in black (black = 'present'). When the original term is annotated as having 'possible' expression (eg. 'mesenchyme derived from neural crest' in this example) the inferred sites of having 'possible' expression include all parents, grandparents etc of 'mesenchyme derived from neural crest' up to the top of the tree. These are shown in green. Note that in the display 'present' is dominant to 'possible'. The inferred sites of 'expression not detected' include all the children, grandchildren etc of 'cardiovascular system' to the bottom of the tree. These are shown in cyan.

An annotation level of ‘not detected’ can also be attached to a term and the convention is to use cyan for denoting this in EMAGE (see also Fig 6). The rule in using ‘not detected’ for annotation is that it signifies that expression is not detected *everywhere* in the structure (ie. not detected in ANY cell of this anatomical structure), which also includes *everywhere in any children/descendants of this structure down to the bottom of the tree* (see also Fig6). Thus, when ‘not detected’ is used, the only pattern that can be applied to the annotation is ‘homogeneous’ and this is done by default.

2.1.2 Retrieving text annotated data from EMAGE

The general rules described above of ‘detected *somewhere* in’, ‘possibly detected *somewhere* in’ and ‘not detected *everywhere* in’ apply when interrogating the database to retrieve data that has been annotated using text. For example to retrieve a database entry that has an annotation of expression in the ‘future brain’ as shown in Fig6, the query term used to interrogate the database can be the term itself and anything higher than it in the tree (in this case, those terms shown in red or black).

Conversely, to retrieve the entry shown in Fig6 when asking “What genes are not detected *everywhere* in the component?”, the term used for query would have to be ‘cardiovascular system’ or anything lower than it in the tree (ie. those terms shown in this example in cyan).

2.2 Spatial annotation in EMAGE

There are two separate ways in which spatial annotation of sites of gene expression can be done in EMAGE – either to a wholemount view, or into the 3D space of one of the 3D virtual embryo models.

2.2.1 Spatial annotation of wholemount data

In EMAGE it is possible to depict regions of expression on a wholemount view of one of the virtual embryos. Only lateral views are offered and the annotation can be done on either a left or right view. As of April 2005, the embryo stages at which this type of annotation can be done are TS07 – TS19 inclusive. Raw data that is suitable for this type of annotation is a photograph of a wholemount stained embryo viewed from the same angle as the virtual model.

2.2.1.1 Rules observed during spatial annotation of wholemount data

Annotation is performed, by adding 2D areas (or domains) onto the 2D 'template' area of either a left, or right view[1]. These depict regions of strong, moderate or weak expression or an area of possible expression. In addition, regions can be annotated as having no expression detected. All of these domains are non-overlapping (ie each pixel of the template can only have one of these levels associated with it). The remainder of the 2D view is left un-annotated ('not examined') (see Fig 7 below).

Spatial annotation to wholemount views is performed independently of any text annotation that may accompany it in an EMAGE entry (ie. terms in the ontology are NOT automatically annotated based on wholemount spatial mapping).

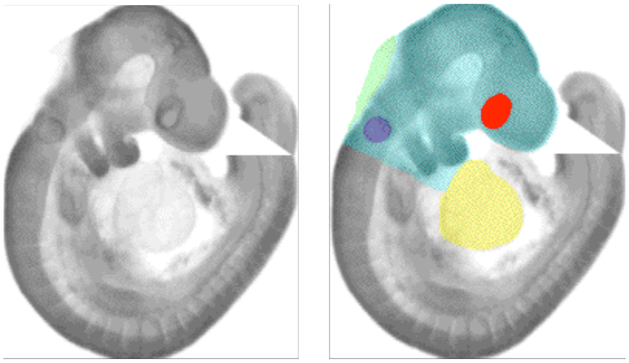
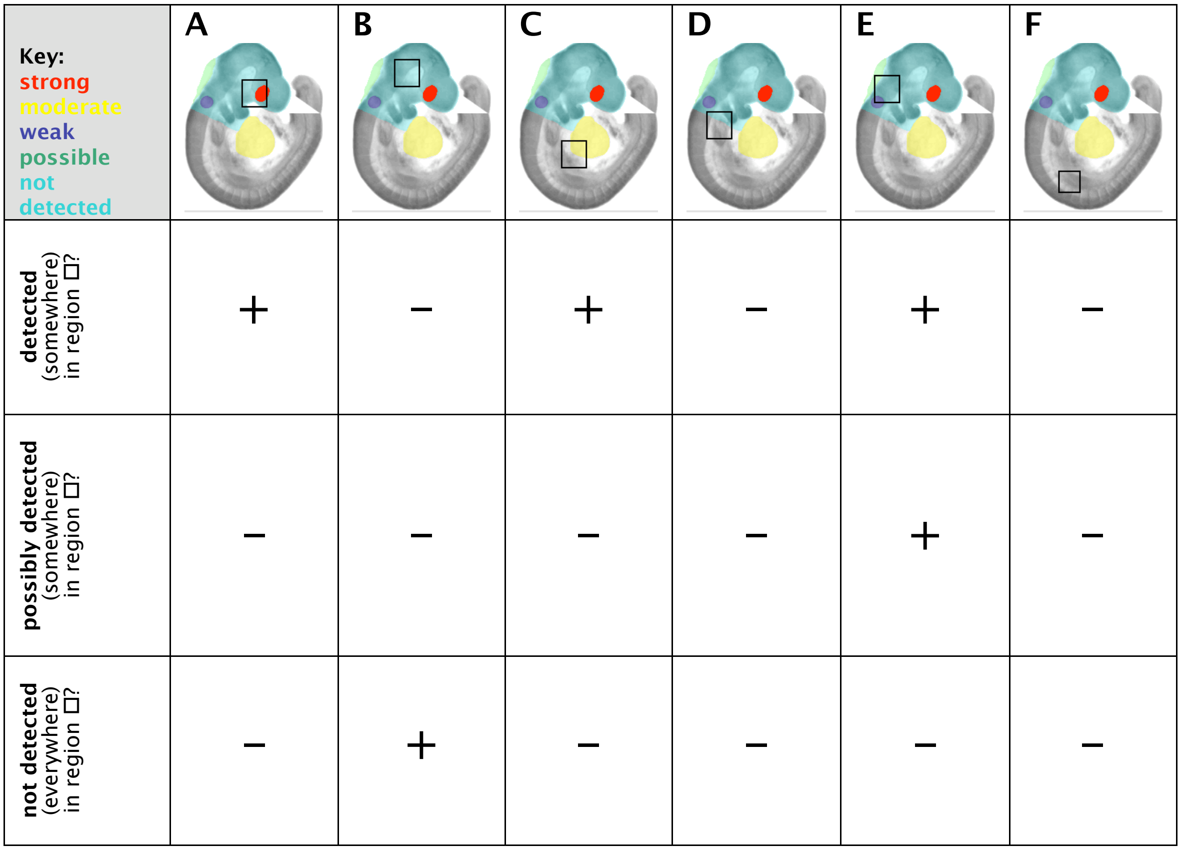


Fig 7 Example of whole-mount spatial annotation
The image on the left shows the right-hand view of the TS15 embryo model. This acts as a template on which regions are added that represent gene expression domains. The image on the right depicts several regions of expression for a particular gene: strong (red) in the eye, moderate (yellow) in the heart, weak (blue) in the otic vesicle and possible (green) in the roof of the hindbrain. Cyan represents regions where no gene expression is detected. The remainder of the image on the right (in grey) are regions that have not been examined and/or spatially annotated.

[1] The spatial annotation is done using a warping method. See separate notes on how to spatially map.

2.2.1.2 Retrieving wholemount data from EMAGE

The general rule of 'detected *somewhere* in' and 'not detected *everywhere* in' also applies when interrogating the database to retrieve data that has been annotated to wholemount views[2]. See Figure 8 below for an explanation.



Key: strong moderate weak possible not detected	A	B	C	D	E	F
detected (somewhere) in region □?	+	−	+	−	+	−
possibly detected (somewhere) in region □?	−	−	−	−	+	−
not detected (everywhere) in region □?	−	+	−	−	−	−

Fig 8 Logic used when querying wholemount data in EMAGE.
The same mapped pattern as shown in Figure 7 is shown. Six query regions are shown in A–F by the black framed square. Positive results are shown with '+' and negative with '–'.
1) The results of asking the question "Is this gene detected (somewhere) in the region specified?" is shown in the first row. Query regions A, C and E return positive results as a region of strong (red); moderate (yellow) and weak (blue) are respectively found somewhere in the query regions. Query regions B, D and F do not return results as no regions of strong, moderate or weak expression are found within them.
2) The results of asking the question "Is the gene possibly detected (somewhere) in the region specified?" is shown in the second row. Only query region E returns a positive result as it is the only one to contain any of the possible (green) expression domain.
3) The results of asking the question "Is the gene not detected (everywhere) in the region specified?" is shown in the last row. Only query region B returns a positive result as this is the only one that contains 100% 'not detected' (cyan).

[2] Note: A transformation between the left and right hand views allows the user to define their spatial query on one side of the embryo, yet retrieve data from the database that has been mapped to both sides.

2.2.2 Spatial annotation of 3D data (ie. sections, OPT data)

In EMAGE it is also possible to depict regions of expression within the 3D space of one of the virtual embryo models. As of March 2005, the embryo stages at which this type of annotation can be done are TS07 - TS19 inclusive. Raw data that is suitable for this type of annotation are photographs of one or more sections of *in situ* hybridisation or immunohistochemistry data, or data imaged using OPT (optical projection tomography).

2.2.2.1 Rules observed during spatial annotation of 3D data

Annotation is performed, by adding 3D areas (or domains) into the 3D space and (similarly to that done in 2D for wholemounts), and these depict regions of strong, moderate or weak expression or an area of possible expression[3]. In addition, regions can be annotated as having no expression detected. All of these domains are non-overlapping (ie. a single voxel of the 3D embryo template can only have one of these levels associated with it). The remainder of the 3D volume is left un-annotated ('not examined' - shown in grey) (see Fig 9).

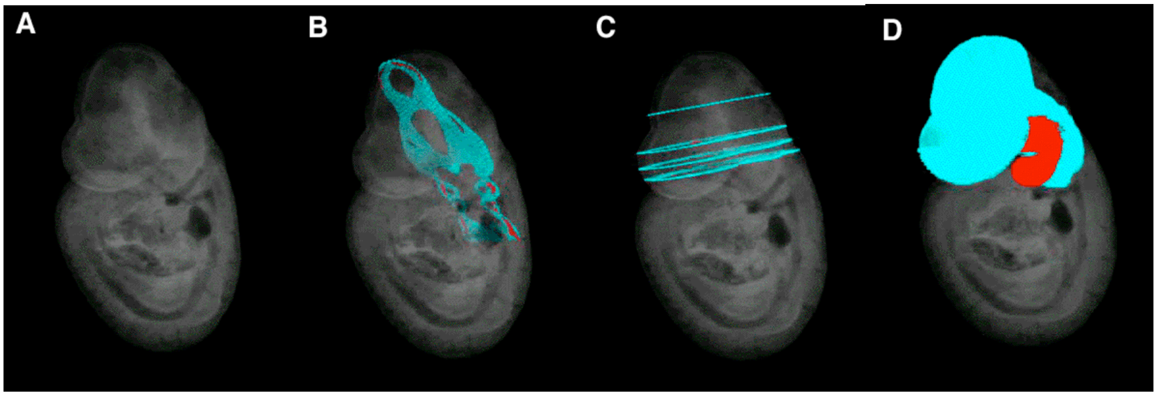


Fig 9 Examples of 3D spatial annotation in EMAGE
Image A shows the 3D TS15 embryo model. This acts as a 3D template in which regions are added that represent gene expression domains. B-D show three examples of spatial annotation into voxels within the 3D space of A. B shows a single section that has been spatially annotated, C shows a set of multiple sections that have been spatially annotated and D shows a full 3D region that has been spatially annotated. In each of these examples only regions depicting strong expression (red) and expression not detected (cyan) are shown. Un-annotated regions are grey.

[3] The spatial annotation is done using a warping method. See separate notes on how to spatially map.

In the embryo models that have full 3D anatomical domains defined within them (those denoted by yellow squares in Fig3), the anatomical domains are linked to the text anatomy ontology (as shown in Fig5). Thus, when spatial data is entered into the 3D space of these embryo models, it is known which anatomical domains the expression domain intersects with, and annotation to the text ontology is done automatically (see Fig 10).

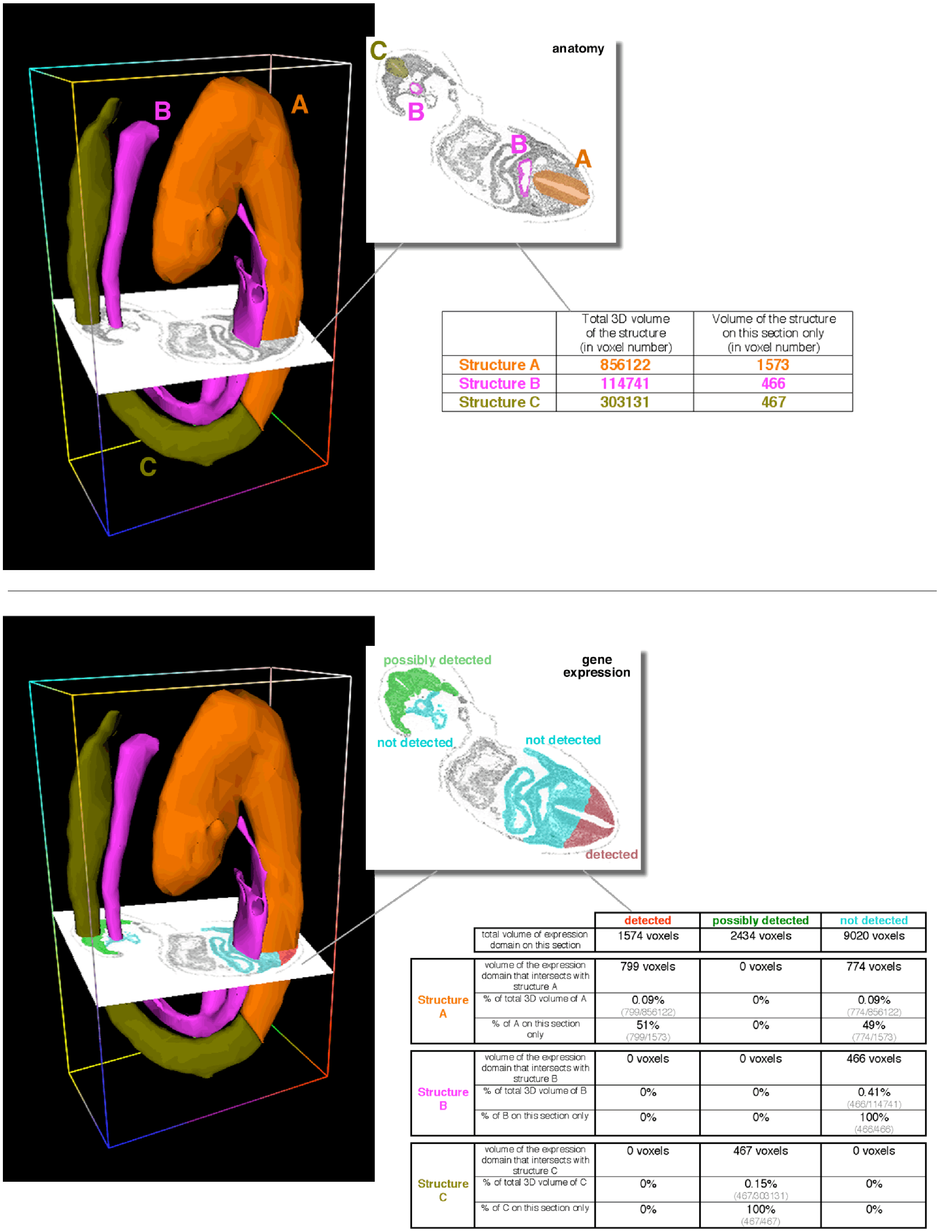
Note that annotation performed manually and automatically are denoted differently in the interface: *manual annotation* is depicted in CAPITALS, whereas *automatic annotation* is shown in non-grey lower case lettering and includes a list of voxel overlap numbers following the term.

The colour of automatically annotated terms reflects the spatial mapping:

In cases where at least one 'strong', 'moderate' or 'weak' voxel intersects with the 3D anatomical domain that corresponds to the text term, the term is shown in black, which represents that expression is 'present'. This adheres to the 'detected somewhere in' rule.

In cases where no 'strong', 'moderate' or 'weak' voxels intersect with the 3D anatomical domain that corresponds to the text term but at least one 'possible' voxel does, the term is shown in green, which represents that expression is 'possibly detected' in that component. This adheres to the 'possibly detected somewhere in' rule.

In cases when no 'strong', 'moderate', 'weak' or 'possible' voxels intersect with the 3D anatomical domain that corresponds to the text term, but 'not detected' voxels do, there are two possibilities. These are due to the 'not detected everywhere in' rule. The first is when the volume of 'not detected' voxels is greater than 100% of the total number of voxels in the anatomical domain that corresponds to the text term. Due to the 'not detected everywhere in' rule, in these cases the term is shown in cyan, which denotes that expression is not detected everywhere in the structure. The second case is when the volume of 'not detected' voxels is less than 100% of the total number of voxels in the anatomical domain that corresponds to the text term. Due to the 'not detected everywhere in' rule, in these cases the term is left unannotated (ie. shown in grey, but with the voxel information following) which is in accordance with the 'not detected everywhere in' rule.



	Total 3D volume of the structure (in voxel number)	Volume of the structure on this section only (in voxel number)
Structure A	856122	1573
Structure B	114741	466
Structure C	303131	467

		detected	possibly detected	not detected
	total volume of expression domain on this section	1574 voxels	2434 voxels	9020 voxels
Structure A	volume of the expression domain that intersects with structure A	799 voxels	0 voxels	774 voxels
	% of total 3D volume of A	0.09% (799/856122)	0%	0.09% (774/856122)
	% of A on this section only	51% (799/1573)	0%	49% (774/1573)
Structure B	volume of the expression domain that intersects with structure B	0 voxels	0 voxels	466 voxels
	% of total 3D volume of B	0%	0%	0.41% (466/114741)
	% of B on this section only	0%	0%	100% (466/466)
Structure C	volume of the expression domain that intersects with structure C	0 voxels	467 voxels	0 voxels
	% of total 3D volume of C	0%	0.15% (467/303131)	0%
	% of C on this section only	0%	100% (467/467)	0%

Fig10 Example of automatic text annotation in EMAGE from an entry with gene expression spatially mapped on one section.

Fig 10 – legend

Upper panel: *In the image on the left, the plane of a single section through a 3D embryo model is shown by the white rectangle. This intersects with multiple 3D anatomical domains in the embryo model, of which three (A, B and C) are shown. A 2D view of the section is shown on the top right depicting the 2D areas of intersection of this section plane with structures A, B and C. The table shows that this section plane intersects with structure A by 1573 voxels of its total 856122 voxel volume, with structure B by 466 voxels of its total 114741 voxel volume and with structure C by 467 voxels of its 303131 total volume.*
Lower panel: *Gene expression for three domains has been spatially annotated onto the section. This constitutes three domains: expression detected (red), possible expression (green) and expression not detected (cyan). This is shown both in the 3D view on the left and on the 2D section in the top right. The table on the lower right shows the total number of voxels that constitute each expression domain, and below this, the absolute numbers of voxels of these expression domains that intersect with each of structures A, B and C. Also shown are these amounts expressed as a percentage of the total volumes of A, B and C and the percentage of structures A, B and C that are present on the section.*

The resulting text annotation is: Gene X is 'detected' (somewhere) in Structure A and 'possibly detected' (somewhere) in structure C. Expression is not automatically text annotated as being 'not detected' in either structure A or B because according to the "not detected (everywhere) in" rule, the whole 3D volume of an anatomical structure has to be spatially annotated as 'not detected' in order to elicit a text annotation of 'not detected'. This is to safeguard from the scenario where the gene may be expressed in a structure in another part of the embryo.

2.2.2.2 Retrieving 3D data from EMAGE

There are two different ways to retrieve 3D mapped data and these depend on whether the data has been mapped into a model with 3D anatomical domains defined or not. In order to retrieve data from the database that has been mapped into the 3D space of one of the embryo models that DO NOT have anatomical domains defined within them (ie. those denoted with green squares in Fig 3), this is based on voxel intersection of the spatial domains (query vs. expression domain) alone. With data that has been mapped into a model that DOES have 3D anatomical domains defined within it (those models denoted with a yellow square in Fig3), this can be retrieved in the same way as just described (ie. intersection of the spatial query region with the expression domains), but because annotation to the text ontology is done automatically for this data, this type of data can also be retrieved by interrogating the text anatomy ontology.

When interrogating the database to retrieve data that has been mapped into the 3D space of one of the embryo models, the general rule of 'detected *somewhere* in' and 'not detected *everywhere* in' still applies.

When spatially interrogating a 3D model (either with or without anatomical domains defined within it), to retrieve any of the 3 entries shown in Fig 9 when asking "What genes are *detected* (somewhere) in the region?", the query region simply has to intersect somewhere with the gene expression domain (in the example cases in Fig 9 only 'strong' is annotated). The same holds true for asking "What genes are *possibly* *detected* (somewhere) in the region?", the query region simply has to intersect with a gene expression domain for possible. However, to retrieve any of these entries when asking "What genes are *not detected* (everywhere) in the region?", the 3D region used for query would have to be the same volume or larger to retrieve the data. The query logic for 3D cases such as these is identical to the 2D wholemount case shown in Fig8, however a third spatial dimension is simply added.

In order to retrieve data that has been spatially mapped into an embryo model with 3D anatomical domains defined within it (eg. the entry shown in Fig10) by interrogating the database using text and asking "What genes are detected (somewhere) in component X?", the text query term used in the question can be the term itself which resulted from the automatic text annotation or anything higher than it in the tree. This is the same logic as discussed in section 2.1.2.

An extension of this situation is that it is also possible to retrieve data that has been text annotated when spatially querying data in embryo models with full anatomical domains defined within them. For example, if the query formulated is such: "What genes are detected (somewhere) in 3D region X?", when region X is the same volume or larger than a 3D region defined for an anatomical structure (eg. region 'A' in Fig 10), all entries that have text annotation of 'detected' to the text component that corresponds to the 3D anatomical region, or its children in the ontology tree, are also returned. For example, if the 3D query region is larger in space than the entire domain for 'future brain' at TS12, this returns all entries in the database that have been text annotated to the term 'future brain' and any of its children in the ontology (eg. future midbrain), as well as those entries that have voxel intersections with the query region.

Because of the "not detected *everywhere* in" rule, in order for a term to be automatically annotated as having no expression detected, 100% of the entire 3D volume of an anatomical domain has be annotated as having 'no expression detected'. If this is the case, the term used for query when asking "What genes are not detected (everywhere) in component X?" can be the term itself which resulted from the automatic text annotation as 'not detected' and/or anything lower than it in the ontology tree. Once again, this logic is discussed in section 2.1.2.