



## Electronic Data Submission: making a local database and submitting data.

After downloading the EMAGE Java Program, you can create a private local database. Any number of local databases can be created - e.g. each member of a laboratory can create a database to manage their own data.

The software allows the user to store the name of the experimenter, the details of the probe and specimen, experimental notes, original data images, notes that are useful for interpretation of the data, and data that has been spatially mapped into the framework of the EMAP mouse atlas. This is the same type and format of information you see in the central database.

In a local database, the user assigns their own names (the "LocalID") for each entry. Searching your own local, private database allows only text searching on the query. None of the spatial information is searchable in this database. For this, the central database must be queried.

- Start by making a database to enter this data into. Select the "Local DB" menu and the "Make New DB" option. A "Create" window will pop up. Call your new database EMAGE\_Course and click on the "Create" button. Your database name will automatically be given the standard .gxdb file extension. By default your database will have been created in your home directory.

Once you have created your database, it is opened automatically. If you want to check that it is open, select "Status" from the "EMAGE" menu. You will see the name of the "Current Local Database" (if no database is open it will say 'none').

- Select from the "Local DB" menu, the "Make New Entry" option. A text field will pop up, asking you to enter the LocalID for your submission. Type your ID for this entry into this field, and click the "OK" button. Whatever you called this entry is now the local ID for this submission and is the name of the submission window on your screen.

To use for your own laboratory requirements, how you choose to fill out this local database submission form is entirely up to you. However, to submit data to the public EMAGE database there are several fields that must be completed. These are highlighted in the interface in red.

- Name and contact details of at least one submitting author
- Name and contact details of the principal investigator of the group.
- Probe or antisera name and a definition of it
- Name of the gene/protein being assayed (MGD nomenclature committee approved)

- Theiler stage for the input data
- at least one original data image (digital photograph)

And to ensure your data is mapped as accurately as possible for the EMAGE central database, we would also like you to include data mapped spatially to an EMAP embryo model.

Database entries in the EMAGE Java Program appear in tabulated format. The tabs are: person, probe, specimen, data annotation, links and references. Instructions how to fill out each follows:

## Person Tab

This tab is where to enter details to identify the people involved in the experiment.

The easiest way to enter the details of people involved is to first use the Person Manager. This allows you to store name and contact details of as many people as you want. To launch the Person manager click on the “Add new person to Local Database” button.

- Click on the “Add new person to Local Database” button and fill in the details. Click on the “Save” button at the bottom of the window and this information is automatically entered into the Person page of the submission.

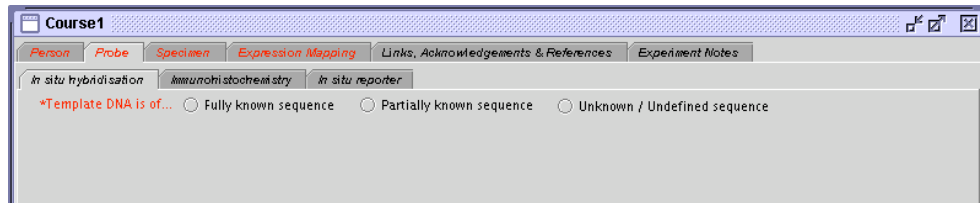
Whenever the EMAGE\_Course database is opened, the personal details of those entered into the Person Manager will be available. Access to the Person Manager is also available under the “Tools” menu.

- If author names have already been entered in the Person Manager, they may be selected from the drop down menu on the right hand side of the ‘authors’ field and added using the “Add to author list” button. New author names may also be typed directly into the text field.

## Probe Tab

This tab is where to enter details to uniquely identify the detection reagent you have used. Similar to the Person Manager, a Probe Manager can be used to enter probe details. This is of particular use if multiple submissions will be made for one probe. It can also be found under the “Tools” menu on the EMAGE window toolbar.

The probe tab has three sub-tabs to choose from: *in situ* hybridisation, Immunohistochemistry and *in situ* reporter.



### For *in situ* hybridisation probes:

“Fully known sequence” describes cases where every nucleotide of the template DNA sequence used to generate the probe is known and can be defined by referral to (part or all of) a single EMBL/Genbank accession number.

“Partially known sequence” describes cases where only part of the nucleotide sequence of the template DNA used to generate the probe is known and can be defined by referral to at least one GenBank/EMBL sequence. Examples of this type would be DNA fragments that are only EST tagged at one or both ends.

“Unknown/Undefined sequence” describes cases where no specific nucleotide sequence of the template DNA used to generate the probe is known. Examples of this type are DNA fragments of a known gene but where the user does not know any exact sequence that the probe corresponds to. Another example is unsequenced, randomly chosen clones as used in *in situ* hybridisation screening projects.

### For immunohistochemistry antibodies:

“Fully known sequence” describes cases where every amino acid of the target epitope sequence used to generate the antibody is known and can be defined by referral to a single EMBL/Genbank accession number.

“Partially/unknown sequence” describes cases where only part or none of the amino acid sequence of the target epitope used to generate the antibody is known.

### For *in situ* reporter assays:

Please use the notes to describe the experimental methods and where possible add references fully outlining the generation of the line.

The detailed examples here refer to the in situ hybridisation probe example, however the behaviours of the forms are similar for immunohistochemistry and in situ reporter.

#### *Name of DNA clone used to generate the probe*

If the probe was transcribed from a cloned DNA fragment, the name of that clone is required. If the probe was generated from a PCR product, a name for that PCR product is required.

If the clone has not been formally named but has previously been indexed by the MGI Molecular Probes and Segments Database, please use the MGI identifier number (entered as MGI:#) as the name of the clone.

#### *Name of probe*

This is a name that uniquely identifies the probe region relative to the template clone. If you have previously entered details of the probe in the Probe Manager, use the pull down menu to enter the details, alternatively, type the name in.

If the probe corresponds to the whole insert of the template clone, use the same name as the template clone. If the template clone was internally cut for probe generation, please give a name that reflects the cut site, or the region of the template that has been used as probe (e.g. "clonenameEcoRI" or "clonename3").

#### *Probe corresponds to cDNA or genomic DNA?*

Click on the appropriate option. If the probe is transcribed from a cDNA fragment, indicate whether it contains any 5'UTR, ORF or 3'UTR sequence.

#### *Nucleotide sequence of probe*

Define the exact nucleotide sequence of the probe by entering an EMBL/GenBank/DDBJ accession number of a sequence that covers the probe in a single continuous stretch and indicate the nucleotide positions at which the probe begins and ends. The version number of the sequence should also be included. This will be suffixed onto the accession number by a dot. Searching the EMBL, GenBank or DDBJ databases is possible by selecting the database using the pull down Search menu - you will be linked automatically to the appropriate website.

If the nucleotide sequence has not been submitted to GenBank/EMBL/DDBJ, but appears in a published paper, the positions of the clone ends can be defined with reference to the numbering of the sequence in that paper. This information should be entered in the 'additional notes' box and the bibliographic details of the paper entered in the Acknowledgements and Citations page.

#### *Gene name and symbol*

EMAGE has been designed such that the gene name and symbol that is currently approved by the Mouse Genome Database (MGD) Nomenclature Committee is used when

the data is entered but synonyms (eg. older or officially withdrawn names) are still supported for searching. When a gene symbol is entered in the interface in the "Search MGI by gene symbols/Accession IDs" field and the corresponding 'Go' button is pressed, a list is returned of current gene symbols along with any accompanying withdrawn synonyms.

#### *Additional notes*

Add any extra relevant information about the probe here.

#### *Probe Origin*

This refers to the origin of the clone that was used to generate the probe.

#### *Probe Type*

Enter information on the type of probe used for the experiment.



The screenshot shows a form titled "Probe Type" with the following fields:

- DNA
- RNA
- Type:
- Labelled with:
- Visualisation Method:

#### DNA vs. RNA

Type - Select sense or antisense from the pull down menu.

Labelled with - Choose the label from the pull down menu or type in the box

Visualisation Method - Choose the visualisation method from the pull down menu or type in the box.

- Once all the details have been entered for your probe, click on the "save these probe details to your local database" button at the bottom of the window. The name can then be selected from the "Probe Name" drop down menu to automatically fill out the relevant details in future entries in this private database.

## Specimen Tab

This page is used to record information about the specimen. Included are the Stage, Strain, Genotype, Preparation Details and Original Images of the sample embryo.

The screenshot shows a web application window titled "Course1" with a "Specimen" tab selected. The form includes fields for "Theiler Stage Value", "Theiler Staging Criteria", "Other Staging system & value", "Strain", "Search MGI Strain Database", "Sex", "Genotype" (with radio buttons for "Wild Type" and "Non-wild type"), "Specimen Preparation" (with radio buttons for "Whole Mount", "Section", and "Whole Mount Subsequently Sectioned"), "Discussion Notes / Critical Notes for Interpreting Results", "Original Images of Source Embryo(s) / (Sections(s))" (with "Add Image" and "Remove Image" buttons), and "Image Notes". A red note states "\*At least one original image is required".

### *Theiler Stage Value*

Select the Theiler Stage of the data embryo from the pull down menu. This is used to select the relevant model for subsequent mapping. The EMAP mouse embryo staging guide can be accessed by clicking on the 'Theiler Staging Criteria' button. If another staging system is known (e.g. dpc, somite number), enter by selecting the alternate staging system from the pull down menu and its value by typing in the adjacent box.

### *Strain*

Enter the strain of the data embryo by typing directly in the box or by searching the MGI Strain Database (by clicking on the 'Search MGI Strain Database' button).

### *Sex*

Select the sex of the sample (if known) from the pull down menu. If not an XX female or XY male, enter the appropriate details in the 'Genotype' or 'Notes Critical for Interpreting The Results' box on this page.

### *Genotype*

The default value is wild-type. If the sample embryo is not wild-type, select 'non wild-type' and enter details of any mutations or insertions.

### *Specimen Preparation*

Please indicate whether the sample is a whole mount or sectioned sample. For all sample types, please enter the fixative used (use the pull down menu or type in directly) and for sectioned samples, the embedding procedure used (use the pull down menu or type in directly).

### *Discussion Notes/Notes Critical For Interpreting the Data*

Enter any relevant details here that would aid someone browsing through your entry to interpret the data correctly. For example, a note to indicate if there is trapping in the neural tube.

### *Original Images of Source Embryo(s) or Section(s)*

Add images in jpg, or gif format by clicking on the 'Add Image' button. In order to add the image, it must have first have been saved on your computer. For the purposes of this course, Sox10.jpg has been saved for you in your home directory. If more than one image is entered, the first image entered will be displayed as a thumbnail image in returned entries from central EMAGE database searches. If you have more than one image, please add them in order of relevance to the annotation.

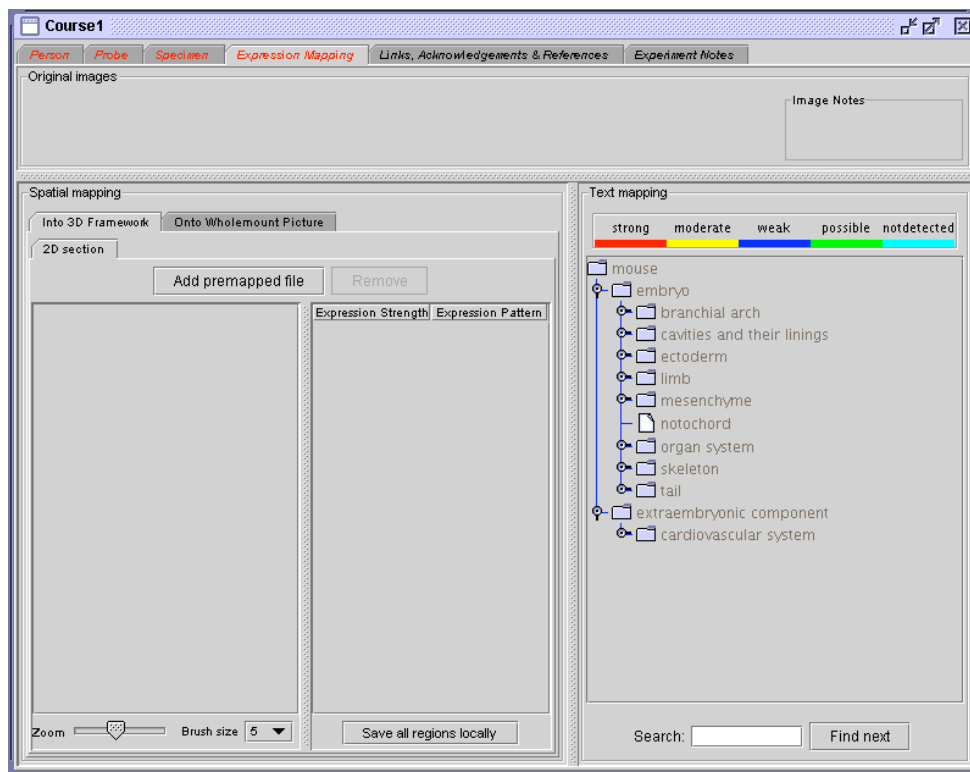
A short note can be added for each image and will appear in the interface when the appropriate thumbnail image is selected. Double clicking on a thumbnail image will open the image in full-size in another window that allows the user to zoom in and out using the slider on the left hand side. Any number of relevant images can be added.

## Expression Mapping Tab

This page is for denoting where the gene is expressed in the sample. If you try to select it before choosing a Theiler stage you will be asked to choose the stage.

At the top of this page appear the original data images that were entered on the previous page. Image notes appear to the right of any thumbnail image that is selected. The lower half of the window is for data mapping (left hand side for spatial mapping and right hand side for text mapping).

Note: this page can be enlarged and the dividers between the original images, spatial mapping and text mapping regions can be moved with your cursor.



### *Spatial mapping*

Further information on spatial annotation can be found elsewhere.

### *Text mapping*

The defaults open the anatomical nomenclature tree for the Theiler stage that was chosen on the previous page to three levels. The number of levels displayed can be changed using the pull down menu at the top of this panel and you can browse through the tree by opening and closing branches by clicking on the toggles that appear to the left of the names. If a toggle does not appear next to the name, this signifies that the structure is at the lowest level of the tree.

The tree can be searched using a text term by typing in the search box and clicking on 'find next'. When there is a positive result, the tree will open and the term will be highlighted. Click on 'find next' to display any more entries corresponding to the text term. If there are no subsequent entries to where you currently are in the tree, a dialogue box will appear with the message 'No matching component. Start search from top?'. Click on yes to search from the top of the list again. In future versions, more sophisticated searches will be possible, including synonym searching.

Additional notes can also be added. Click on OK to save the changes and dismiss this window.

The colour of the anatomical component will change in the tree to reflect the level of expression that has been mapped to it. The colours used are the same as those for the spatial mapping (i.e. red for strong, yellow for moderate, blue for weak, green for possible, cyan for not detected and the default brown/grey for not examined).

## Links, Acknowledgements and References Tab

This page contains acknowledgements you may wish to give and also the details of any references to the probe, technique employed etc. If you have used a probe of unknown or undefined sequence, you must include the paper or the source of this information.

The screenshot shows a software window titled 'Course1' with several tabs: 'Person', 'Probe', 'Specimen', 'Expression Mapping', 'Links, Acknowledgements & References', and 'Experiment Notes'. The 'Links, Acknowledgements & References' tab is active. It contains three main sections: 1. 'Related data in other database' with 'Add', 'Remove', 'Edit', and 'Show' buttons and a table with columns 'Database', 'ID', and 'Type'. 2. 'Acknowledgements' with 'Add', 'Remove', and 'Edit' buttons and a 'Project Name' input field. 3. 'References' with 'Add', 'Remove', and 'Edit' buttons and a 'Title List' input area. At the bottom, there is a 'Default details' dropdown menu and a 'Store Details' button.

### Links

Add links to information held in other databases by clicking on the 'Add' button in the upper left hand 'Links' panel. A window called 'Links' will open

A list of databases which you can link to are contained in the drop-down menu next to 'Database'. This list will include the gxdb database that you currently have open and this linkage method can be used to link two of your own entries to each other when submitting data to the public EMAGE database.

Add the identifier (ID) of the link (e.g. if linking to MGI or EMAGE, just type the ID number of the linked info (no need to include the MGI:' or 'EMAGE:' prefix). If linking to another entry in your local database, enter its Local ID. Define the type of link using the pull down menu next to 'Type' and then click on 'OK'.

### Acknowledgements

Add new acknowledgments using the 'Add' button. The name, address and URL (web address) of the acknowledged person and the acknowledgment can be entered.

## *References*

Add a new reference by clicking on the 'Add' button. A window will open entitled 'New Publication Details'.

It is possible to search the PubMed database for an article and the PubMed ID number (PMID) by clicking on the 'Complete using PMID' button. Alternatively, if you already know the PubMed ID number, you can fill out the form by typing this number into the box beside the 'Complete using PMID' button and then clicking on that button. The details will be automatically entered. Click on OK to enter these details into the EMAGE interface.

## Experimental Notes Tab

This page only appears in your local database - it will not appear in the central EMAGE database. It can be used as part of your local data management strategy.

The screenshot shows a web application window titled "Course1". The window has several tabs: "Person", "Probe", "Specimen", "Expression Mapping", "Links, Acknowledgements & References", and "Experimental Notes". The "Experimental Notes" tab is selected. The form contains the following fields:

- Experiment ID:
- Slide/Specimen ID:
- Embryo ID:
- Investigator:
- Date: 14 Oct 2002 (with dropdown menus for day, month, and year)
- Comments:
- Additional Info1:
- Additional Info2:
- Additional Info3:

- When you have completed a local database entry, save the submission to the local database using the "Local DB" menu and the "Save to Local DB" option. Alternatively, click on the crossed square at the top right hand corner of the entry window. A dialog box will appear, and ask if you wish to save changes. Save changes to your local database before closing.

## Submission to the Public EMAGE database

Following data entry into your local database, any number of specific entries can then be submitted to the EMAGE Editorial Office for inclusion in the public database.

Registration is required for submitting data to the public EMAGE database and is used to protect access to the data associated with each submission.

Data received electronically over the web by EMAGE is automatically given a temporary ID number (the "TempID") and sent to the Editorial Office for curation. Following curation, an editor will contact the submitter to check any queries and to finalise the submission prior to deposition into the public EMAGE database.

When a submission is deposited in the public EMAGE database, the TempID number is replaced by a permanent EMAGE ID number ("EMAGE:ID").

It is possible to organise a publication embargo date with an editor to ensure that data does not appear in EMAGE prior to scientific publication in a journal.

Information held in EMAGE is owned by the submitter.

■ To submit an entry to the central EMAGE database, open it and then select from the "Local DB" menu, "Submit to Central DB".

If you have previously registered to submit data, you would normally enter your username and password into the resulting dialog box.

In normal circumstances, if you are not registered, type anything into this box and click on "Submit". A second pop up window will say you are not registered and direct you to a web page. Follow this link, fill out the information requested and submit your details. Once you have done this, you will be sent an email directing you to another page to select and confirm your own password for future entry. Following this you will be sent another email confirming your user name and password. You are strongly encouraged to retain a copy of this email as you will need to remember your username and password should you wish to amend any of your entries held in the central database at a later date.