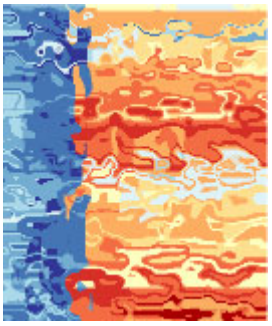


MIMAS

User Manual

Version 3.0
January 2009

Table of Contents

1 Welcome to MIMAS

Welcome to MIMAS, a multiomics information management and annotation system. MIMAS is a fully MGED/MIAME standards compliant system and database repository designed to describe, store and retrieve experiments using Affymetrix GeneChip® and Illumina BeadArray® microarray technologies as well as Illumina Genomic Analyzer UHTS technology. MIMAS uses a web browser based graphical user interface, allowing access from any place where there is an internet connection. The MIMAS system provides a secure and private place for users to store and completely describe their experiments and then retrieve them easily for public repository submission (ArrayExpress, GEO) and publication.

2 MIMAS Home, Registration and Login

The MIMAS home page is a place where relevant general information about MIMAS and its underlying services is displayed. The top navigation bar, found in the top right-hand corner of the web page and shown in Figure 1, contains important links for MIMAS navigation. From the MIMAS home area, one can go to the registration or login pages using the navigation bar.

Figure 1. MIMAS Top Navigation Bar

Before using MIMAS, a user must first visit the registration page to submit their registration information in order to receive a login. The registration page is broken into two parts, shown in Figures 2 and 3. First is the user's personal information followed by the user's laboratory and organization information. Required fields are followed by a red asterisk *. If the laboratory connected to the principal investigator email address already exists in MIMAS, this information will be provided in the lower registration section. The following are some important field notes:

Field	*Notes*
Login/Username	A case-sensitive user identifier (i.e. "test" is different from "Test") with a maximum of 15 characters
Password	A sufficiently complex password using at least 6 characters with a mix of numbers and letters, uppercase and lowercase
Phone Number	Telephone number of the form: +country code area code phone number (e.g. +41 61 555 5555)
Fax Number	Fax number must be of the form: +country code area code fax number (e.g. +41 61 444 4444)
Email Address	The main and valid user email address (used by MIMAS to notify the user of important alerts and information)
Principal Investigator Email	The main email address of the head of the user's laboratory
Laboratory/Group Name	The official name of the user's laboratory or, if this does not exist, the research concentration of the laboratory. Both should be followed by the principal investigator name and short from of the institution name to easily discern between different laboratories with the same name (e.g. "Genome Bioinformatics, Michael Primig SIB Biozentrum Basel")
Institution	The official name of the institution which the laboratory belongs to

Figure 2. MIMAS Registration Part I

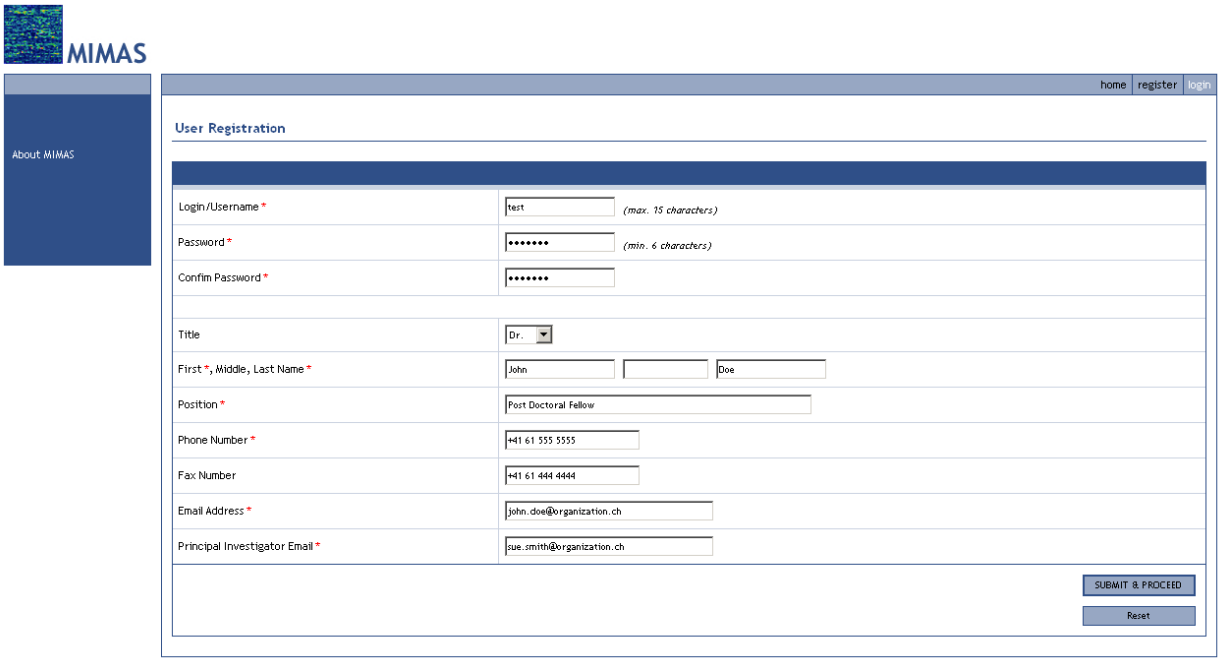
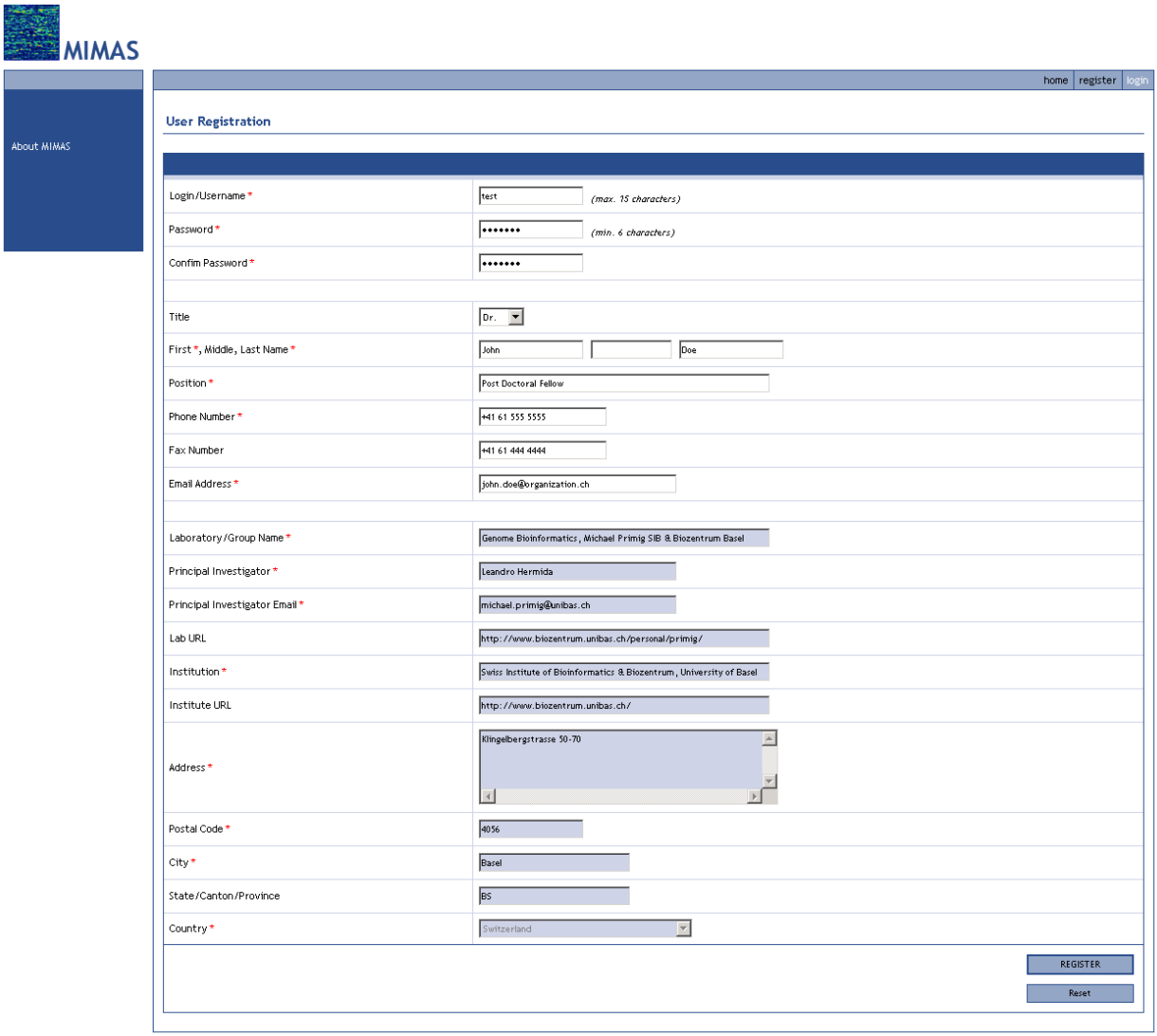


Figure 3. MIMAS Registration Part II



After successfully submitting your registration, MIMAS will reply to you via email to alert you that your registration has been accepted and that you may start using your login. Simply follow the link located on the top navigation bar to find the login page.

3 MIMAS User Home

After successfully logging into MIMAS, you will be immediately forwarded to your personal user home page, shown in Figure 4. On this page MIMAS will display information and alerts relevant to you. The top navigation bar now has a "*Logout*" link which you should make sure to click on whenever you decide to leave MIMAS. You will also notice that in addition to the top navigation bar you now have available to you the main internal navigation area on the left side of the web page. The main navigation area, shown in detail in Figure 5, is broken into two parts, the main menu on the top and detail menu on the bottom. The main menu has links to the four major areas inside MIMAS: 1) User Information and Management "*User Home*", 2) Microarray Experiment Data Uploads and Annotation "*Experiment Uploads*", and 3) Microarray Data Search and Retrieval "*Search Repository*". Each main menu contains a corresponding detail menu with additional links.

Figure 4. MIMAS User Home

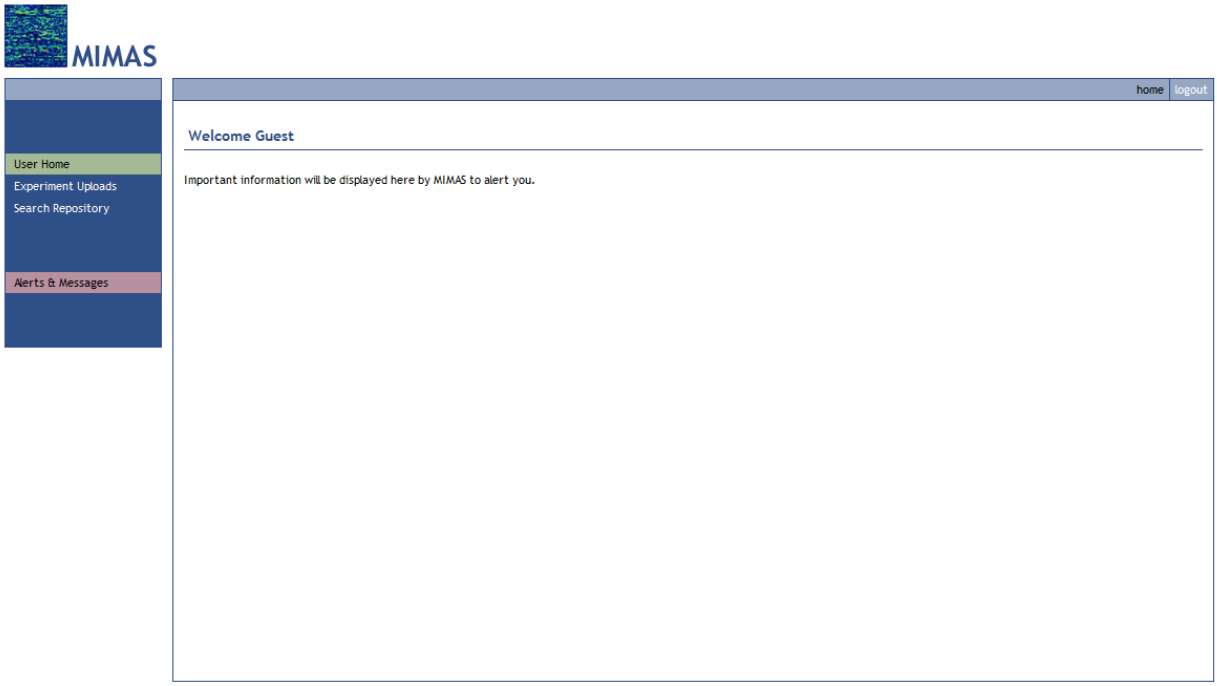
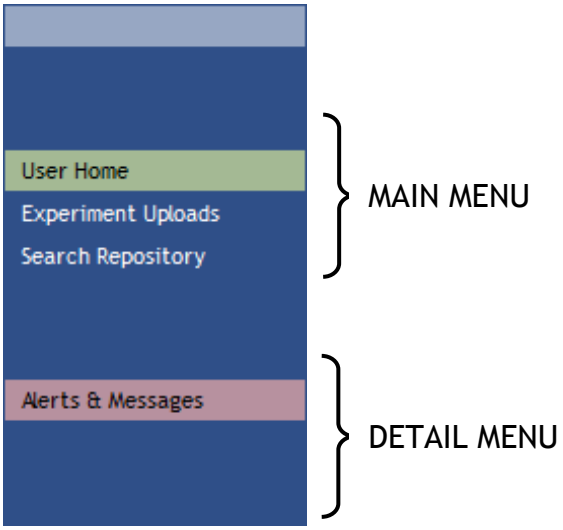
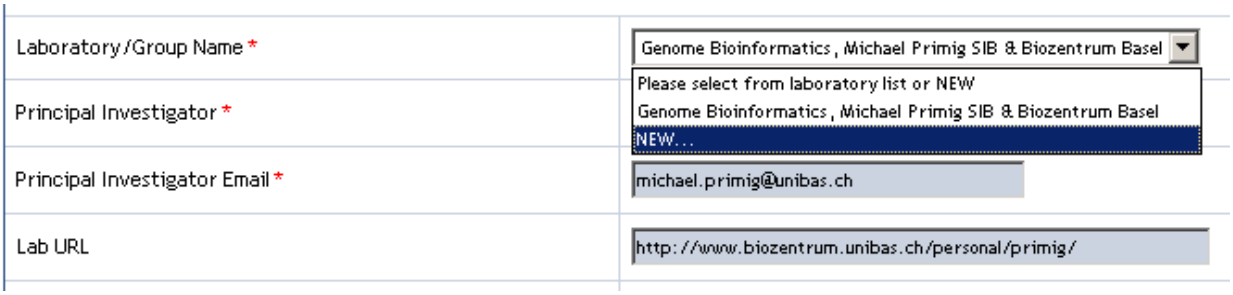


Figure 5. MIMAS Main Navigation Area



You can see that when you log into MIMAS you are directed to your user "*Alerts & Messages*" page which is in the User Information and Management section. By clicking on the "*Personal Information*" link in the detail menu you can update all of the user, laboratory and organization information you entered during your registration. If you make any mistake and would like to start over, simply click on the "*Reset*" button to repopulate the page with user information that is in the database. You will notice that unlike in the registration page the "*Laboratory/Group Name*" field now has a drop down menu showing other laboratories that exist in MIMAS. If you have changed laboratories and your laboratory does not exist in this list MIMAS provides a "***NEW…***" option at the bottom of this drop down menu as shown in Figure 6. Upon selecting the "***NEW…***" option you will be prompted for the name of this new laboratory. MIMAS will then unlock all of the other laboratory and organization fields so that you can complete them with the information for the new lab. For already existing laboratories, only if you are the principal investigator of this laboratory will MIMAS have the fields unlocked to allow you to change the laboratory and organization information.

Figure 6. Creating a New Laboratory

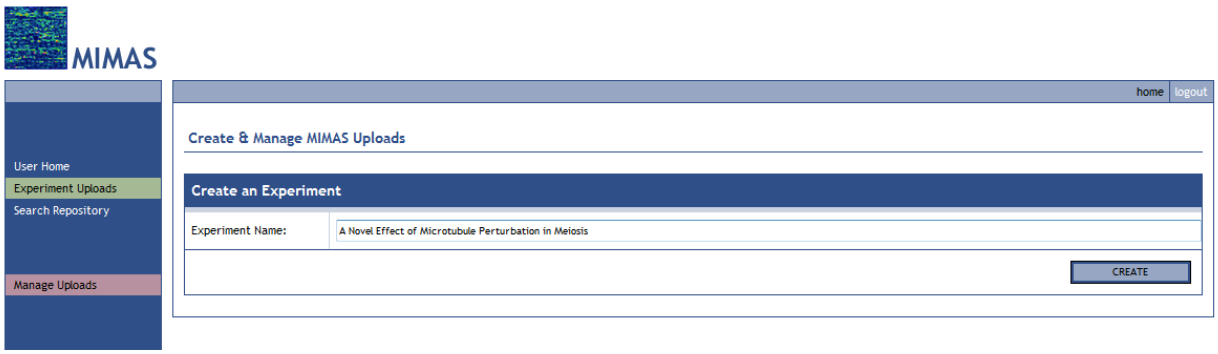


4 MIMAS Microarray Experiment Uploads and Annotation

Upload Management

To submit microarray experiment data to MIMAS, click on the "*Experiment Uploads*" main menu link found in the main navigation area. You will be forwarded to your own personal upload management page, shown in Figure 7. On the upload management page you can create new upload experiments, rename existing experiments, view experiment upload status, and remove experiments not yet submitted to the MIMAS repository.

Figure 7. MIMAS Upload Management: Create a New Experiment

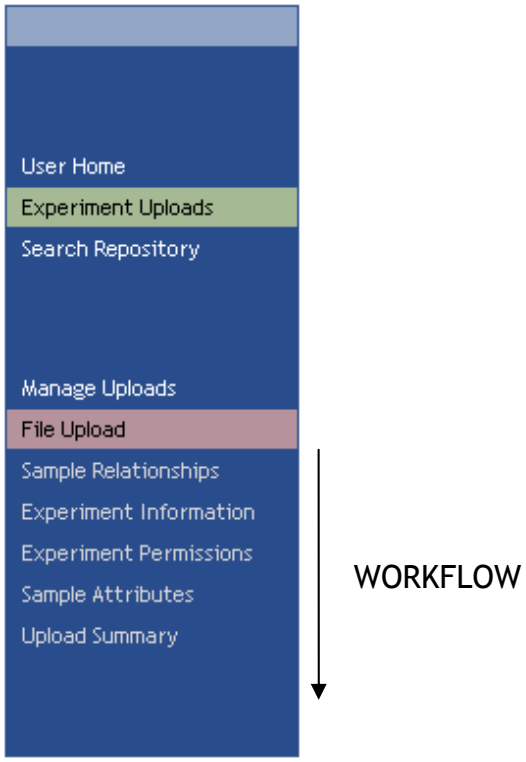


You can create and simultaneously work on as many upload experiments as you would like in MIMAS. MIMAS saves all of the data and the state of your working experiments indefinitely, so if you log out and would like to come back to an experiment at a later date everything will be there for you just as you had left it. To begin a new microarray experiment upload, simply type in a one line descriptive name which summarizes the experiment. The experiment name can be phrased like a journal paper title (e.g. "Study of GAL-SPO13 Overexpression during Mitosis in S. cerevisiae").

Upon creation of the experiment, you will be forwarded to the file upload page (Figure 8). You will now notice that in the detail menu section, shown in Figure 9, a set of links appears describing the required steps of an experiment upload. This is the workflow used by MIMAS to obtain and annotate your experiment.

Figure 8. MIMAS File Upload: Upload a File Archive

Figure 9. MIMAS Experiment Upload Workflow



The order of the experiment upload workflow is determined by the interdependency of each of the workflow steps. For example, information obtained in the "*Sample Relationships*" step is then needed by the "*Experiment Information*" and "*Sample Attributes*" steps. You can see in Figure 9 that when certain steps in the workflow are grayed out and unclickable that information needed upstream has yet to be completed. With the experiment upload workflow you can always go back and forth between any steps to make changes, additions and deletions. When changes, additions or deletions are made MIMAS will automatically know whether downstream parts of the workflow are affected and you may see that steps you have already visited will need to be reviewed.

If you click on the "*Manage Uploads*" detail menu link to return to your upload management page, you will see, as shown in Figure 10, that you now have a "*Manage Experiments*" section with one experiment in it. An important feature of this section is the "*Status*" designation for each experiment. This tells you what stage in the MIMAS upload process the experiment is in. For experiments that are still a work in progress, you can click on the "*Working*" button to go to that experiment. To rename an experiment, hover over and click on the experiment name. A prompt window will then open to allow you to rename it, as shown in Figure 11. If you are using Internet Explorer 7 or later, by default allowing scripted windows will be disabled and you will get a warning. In order to enable scripted windows, go to Tools -> Internet Options -> Security and add the URL for your MIMAS installation into the "Trusted Sites" group.

Figure 10. MIMAS Upload Management: Manage Experiments

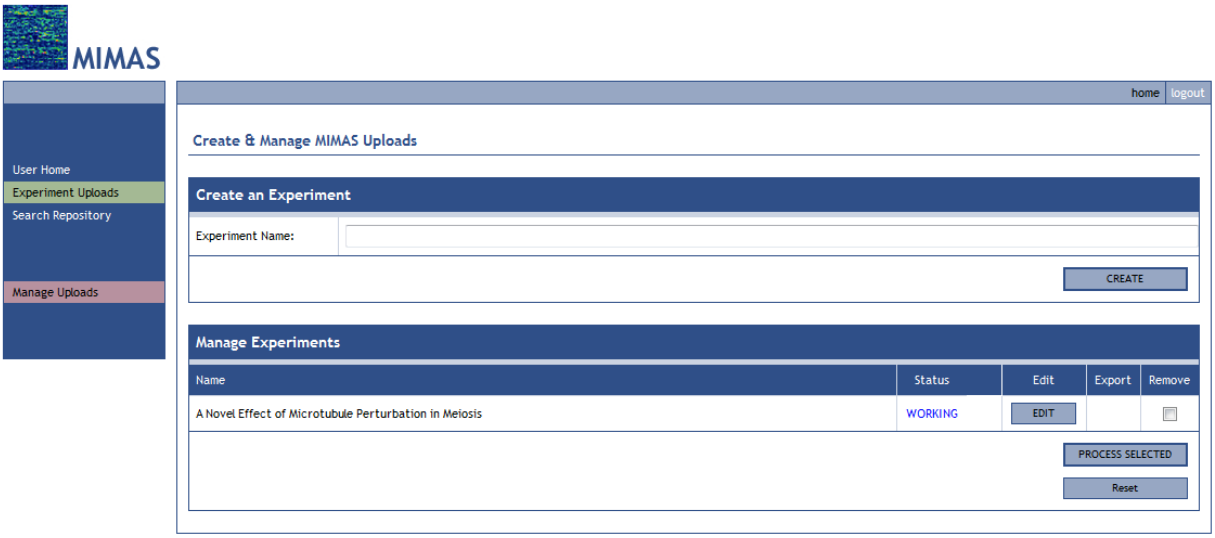
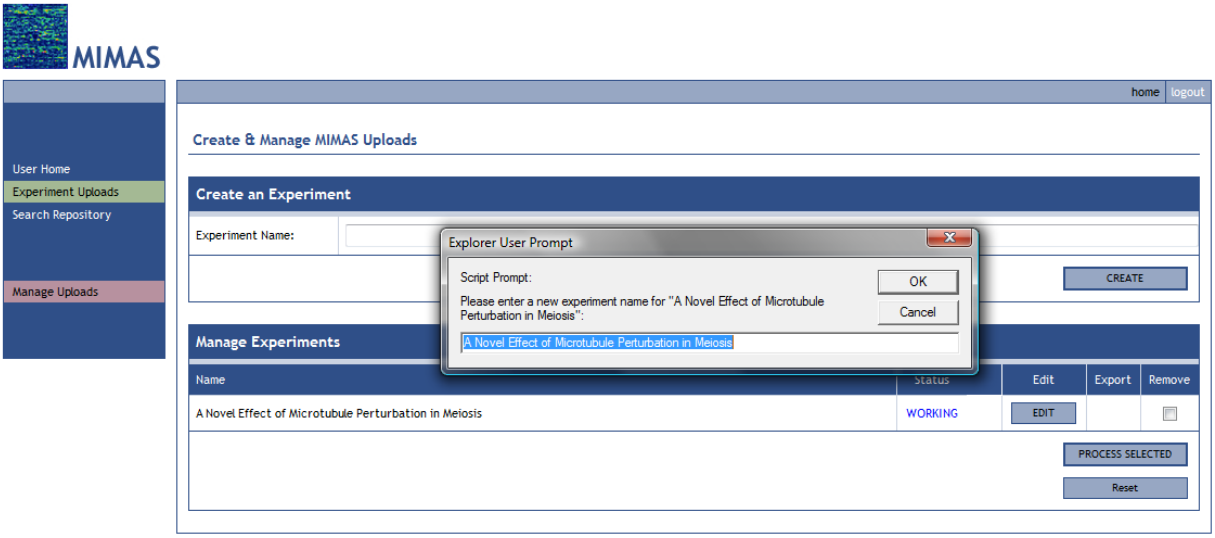


Figure 11. MIMAS Upload Management: Rename Experiment



After clicking OK on the rename prompt, you must click the "*Process Selected*" button to submit the change to MIMAS. The same holds true if you want to remove an experiment in your upload area. Simply click the check boxes for the experiments you want removed and then click the "*Process Selected*" button and you will be asked if you are sure you want to remove the experiments in question. Remember that when you remove an experiment from your upload area it is permanently deleted and cannot be recovered.

File Upload

When a new upload experiment is created in MIMAS, the first step in the workflow is the upload of your experiment data files.

Affymetrix GeneChip® array data

The Affymetrix scanner software (MAS4, MAS5 or GCOS) used to process Affymetrix GeneChip® samples produces a set of files associated with each hybridization. Each hybridization file set contains the same name and each of the following suffixes: DAT, CEL, CHP, EXP, RPT and TXT. MIMAS requires that you upload the CEL (probe-level raw data) file. Additionally, you may upload files with the same sample name and the following extensions:

- TXT (probeset-level data calculated using the Affymetrix probe-level analysis algorithm)
- RMA (probeset-level data calculated using the RMA probe-level analysis algorithm)
- GMA (probeset-level data calculated using the GC-RMA probe-level analysis algorithm)

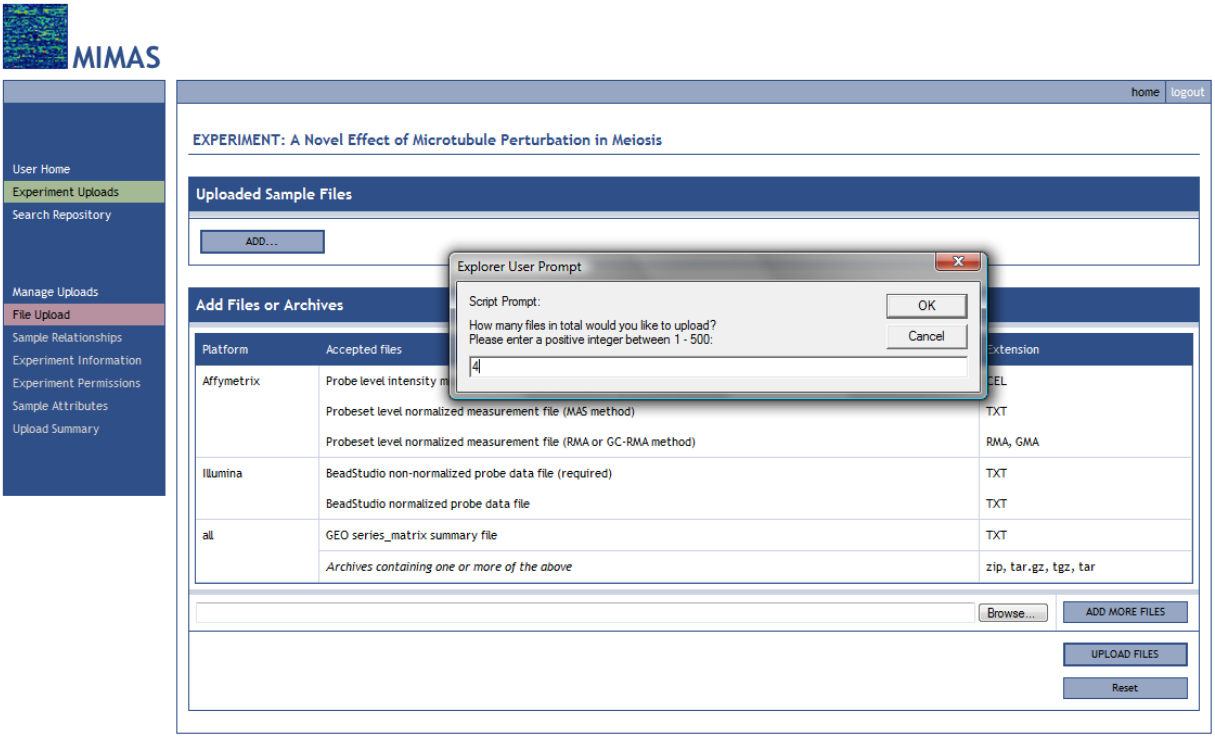
Please see Appendix B: prompted to type in the number of total files you would like to upload, shown in Figure 12. *RMA and GC-RMA File Creation Using GeneSpring* on how to create these files. MIMAS recommends that you create these files for upload at this step so that analysis and file management will be easier later. But if this is not possible then you can omit the RMA and GMA files. In Figure 8 you will see that you have two options by which you can upload your data files into MIMAS: 1) using a file archive ZIP or TAR.GZ/TGZ file or 2) uploading each file individually. Method 1 is the easiest method: simply create a ZIP or TAR.GZ/TGZ archive of your CEL, TXT (and RMA, GMA) files. Then browse to locate this file and the path to the archive will automatically be entered into the file upload field. If you choose method 2 first click on "*Add More Files*" and you will be

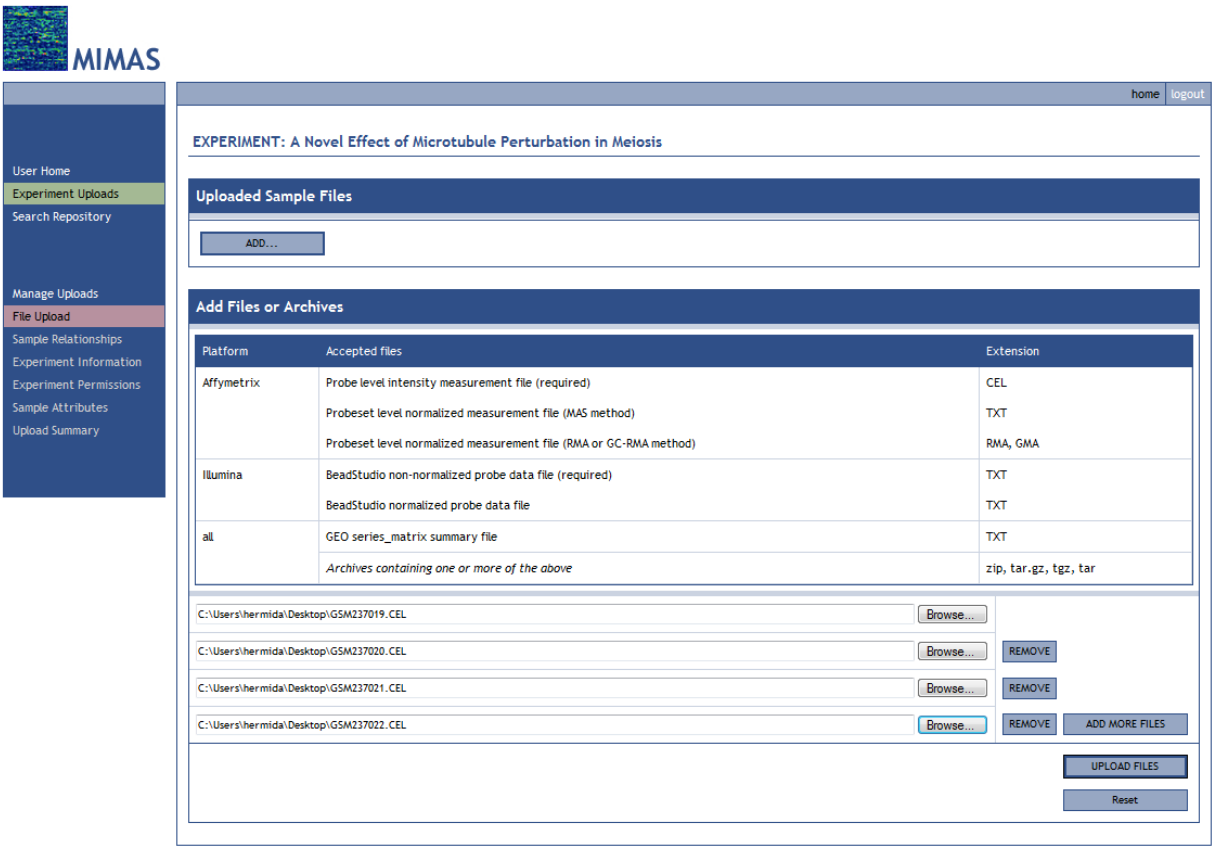
Illumina BeadArray® data

The Illumina BeadStudio software should be used to analyze and summarize probe data. MIMAS requires that you upload a TXT file produced by BeadStudio with the `normalization` option set to `none`. This file will be shown in MIMAS as a file in 'Illumina' format. Additionally, you may upload a normalized measurement file in one of these formats:

- TXT file produced by BeadStudio with normalization
- TXT file produced by a custom method in the same format as that produced by BeadStudio
- TXT files produced by a custom method in the same format as that produced by GeneSpring. In that case each file must be named with the sample name in the non-normalized BeadStudio file and the TXT extension (e.g. '1735640015_A.TXT').

Figure 12. MIMAS File Upload: Upload Individual Data Files





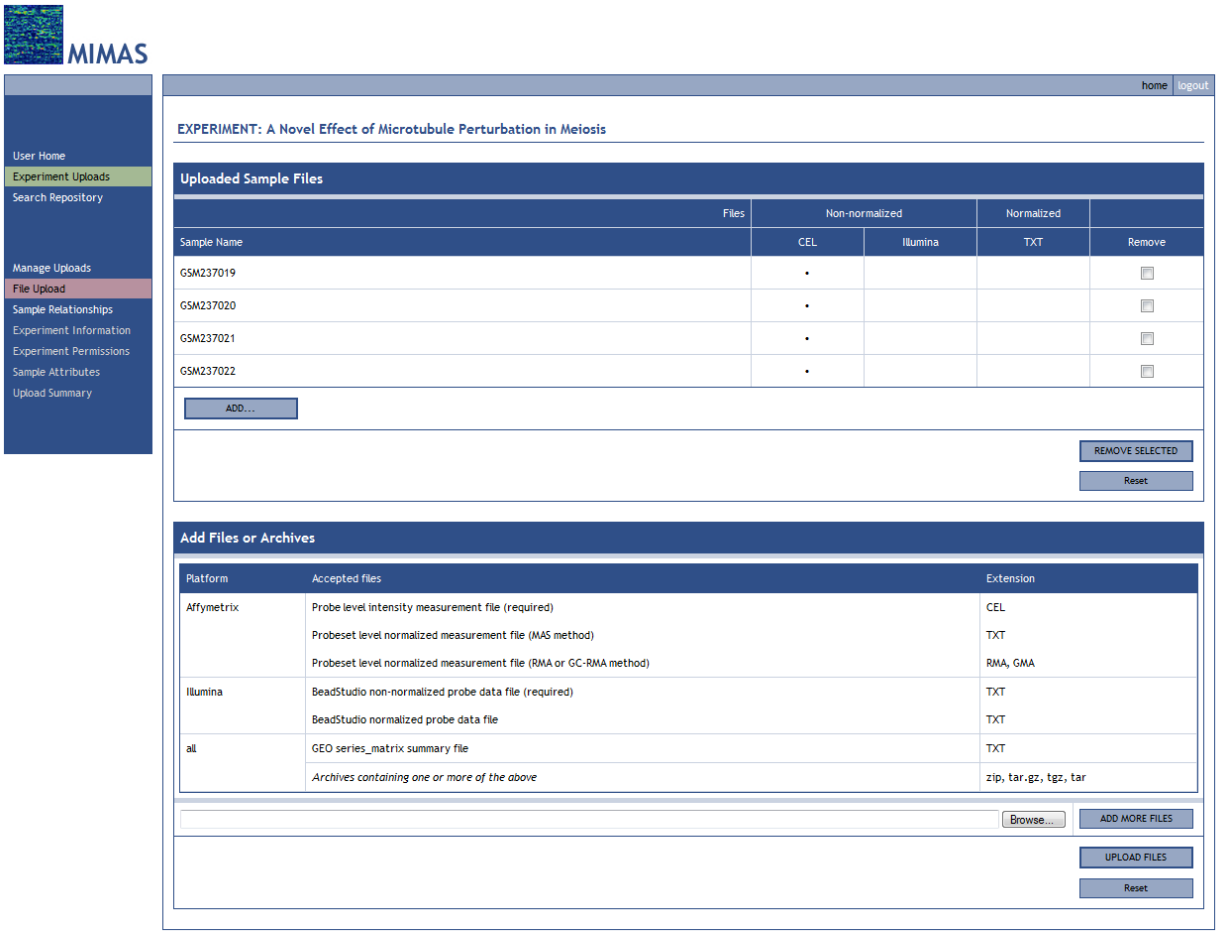
Do not worry if you accidentally type in the wrong number, MIMAS only uses it to generate the initial number of individual file upload fields you will need. If you accidentally type in the wrong number you can adjust the number of file upload fields by using the "*Add More Files*" or "*Remove*" buttons. For each individual file upload field, browse for each of the data files you would like to upload.

When you have completed telling MIMAS the location of your data file archive or individual data files click on either the "*Upload Archive*" or "*Upload Files*" button depending on which method you have used. You will then be presented with a progress screen, shown in Figure 13, while the file(s) are being uploaded to the MIMAS server and processed.

Figure 13. MIMAS File Upload Progress

Make certain to not to click on any other links on the page and do not close the browser window while the upload and processing are in progress. While this is happening you may minimize the MIMAS browser window if you choose and continue to use your computer and other browser windows as normal. Depending on the number of data files, the data file chip type and the present load on the MIMAS server, this process can take from a minute or two to even half an hour. When the upload and processing is complete you will automatically be forwarded to the next step in the workflow, the "*Sample Relationships*" page. You will see that the detail menu on the main navigation area has shifted down and that you are free to go back to the "*File Upload*" step. If you return to the "*File Upload*" page you will see that now you have an "*Uploaded Sample Files*" section, shown in Figure 14.

Figure 14. MIMAS File Upload: Manage Uploaded Files



If you have the case where you would like to begin annotating your samples but don't yet have data files you can create sample using the ADD button in the "*Uploaded Sample Files*" pane, shown in Figure 15. The sample name should be that of the data file that you will upload later (e.g. GS237025 for data file GS237025.CEL). You will see that the sample is created with a reminder that the data file is missing and should be uploaded before finishing your experiment.

Figure 15a. MIMAS File Upload: Creating Samples w/o Data Files

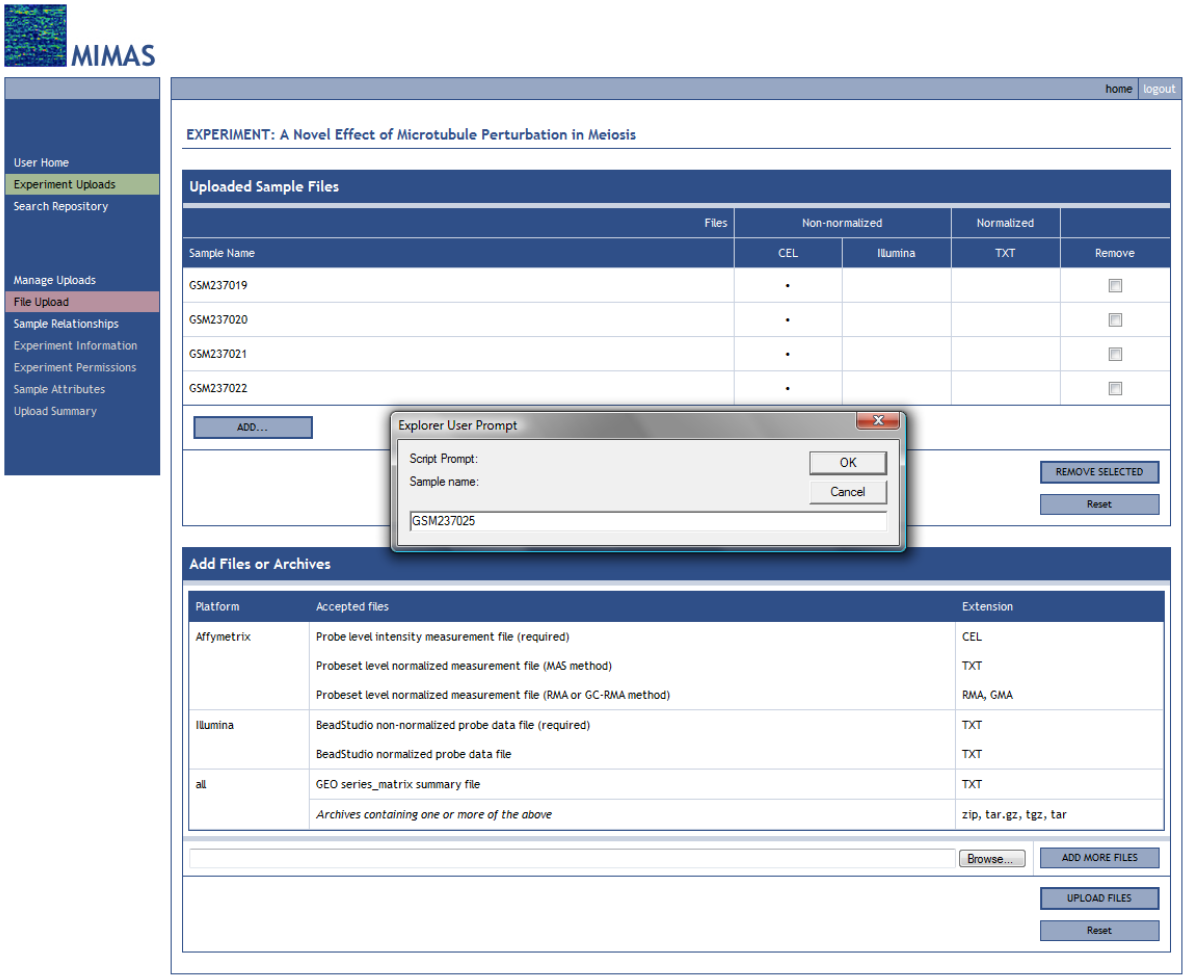
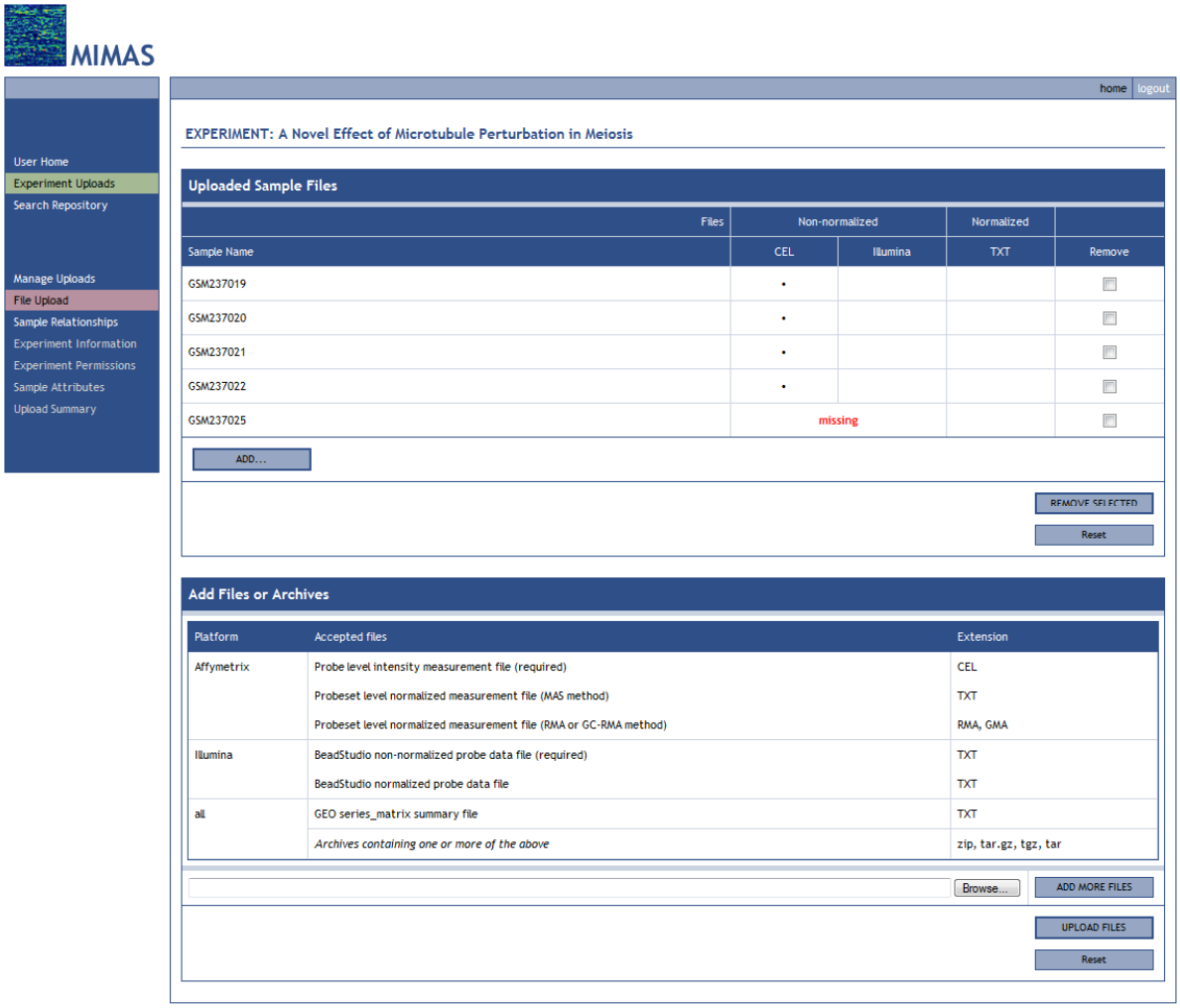


Figure 15b. MIMAS File Upload: Creating Samples w/o Data Files



Sample files can be removed by checking the appropriate check boxes and clicking on the "*Remove Selected*" button. If you need to upload additional files use the upload archive or upload individual files tools located below the file list on this page.

Experimental Conditions and Sample Relationships

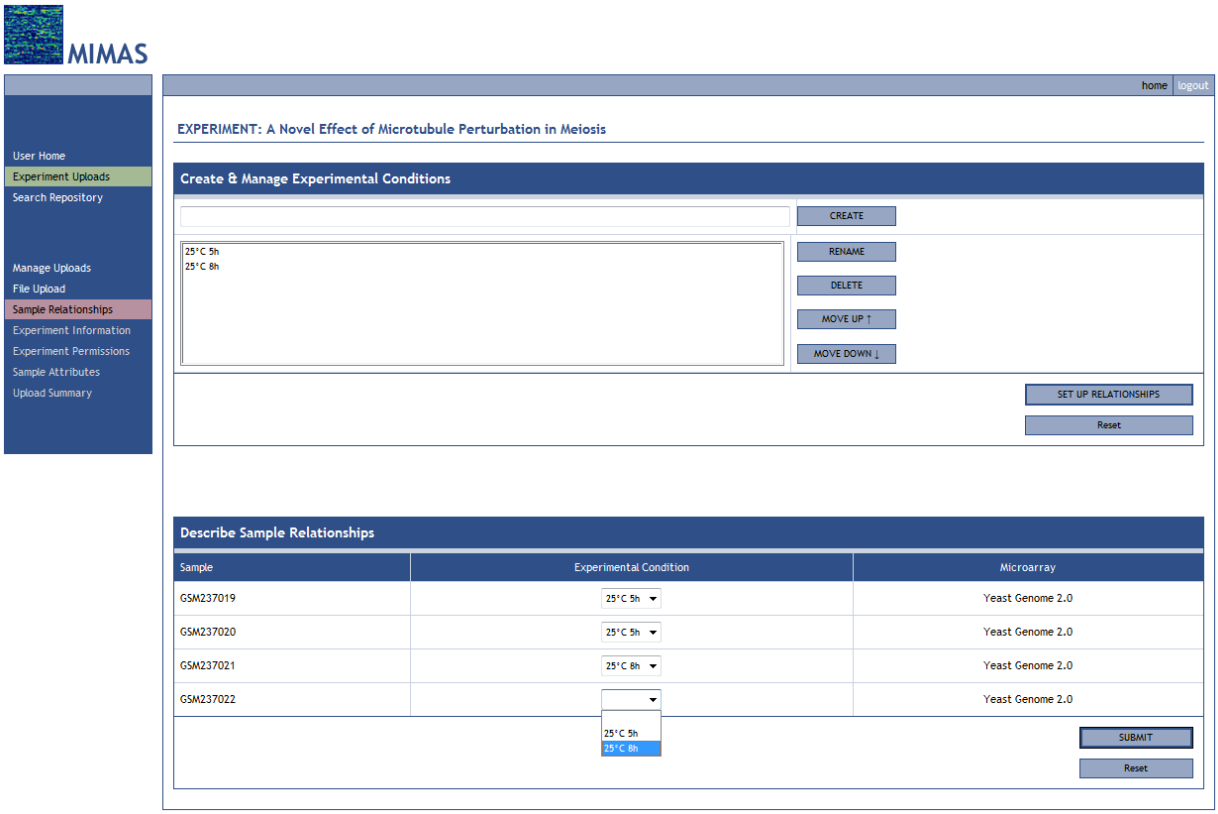
After successfully uploading a file archive or individual files, the next step in the workflow is to define experimental conditions and relating them to your samples. As shown in Figure 16, you first give descriptive names for each of the experimental conditions present in your experiment. These names will appear in the experimental condition box as you create them. You may then rename, delete, or change the order of your experimental conditions. The experimental condition order displayed in this section is the same order used to annotate your experimental conditions and samples later in the workflow. Then click on the "*Set up Relationships*" button.

Alternatively, there is a button labeled *"Automatically create one condition per sample"* that will create as many experimental conditions as there are samples. The experimental conditions can then be individually renamed. This is useful in cases when there are many conditions with no replicates, e.g. a time series experiment. The use of this button is normally **not recommended**.

Figure 16. MIMAS Sample Relationships: Add Experimental Conditions

After submission you will see a new section called "*Describe Sample Relationships*", shown in Figure 17. In this section the goal is to map the samples in your experiment to their appropriate experimental condition.

Figure 17. MIMAS Sample Relationships: Map Samples to Conditions

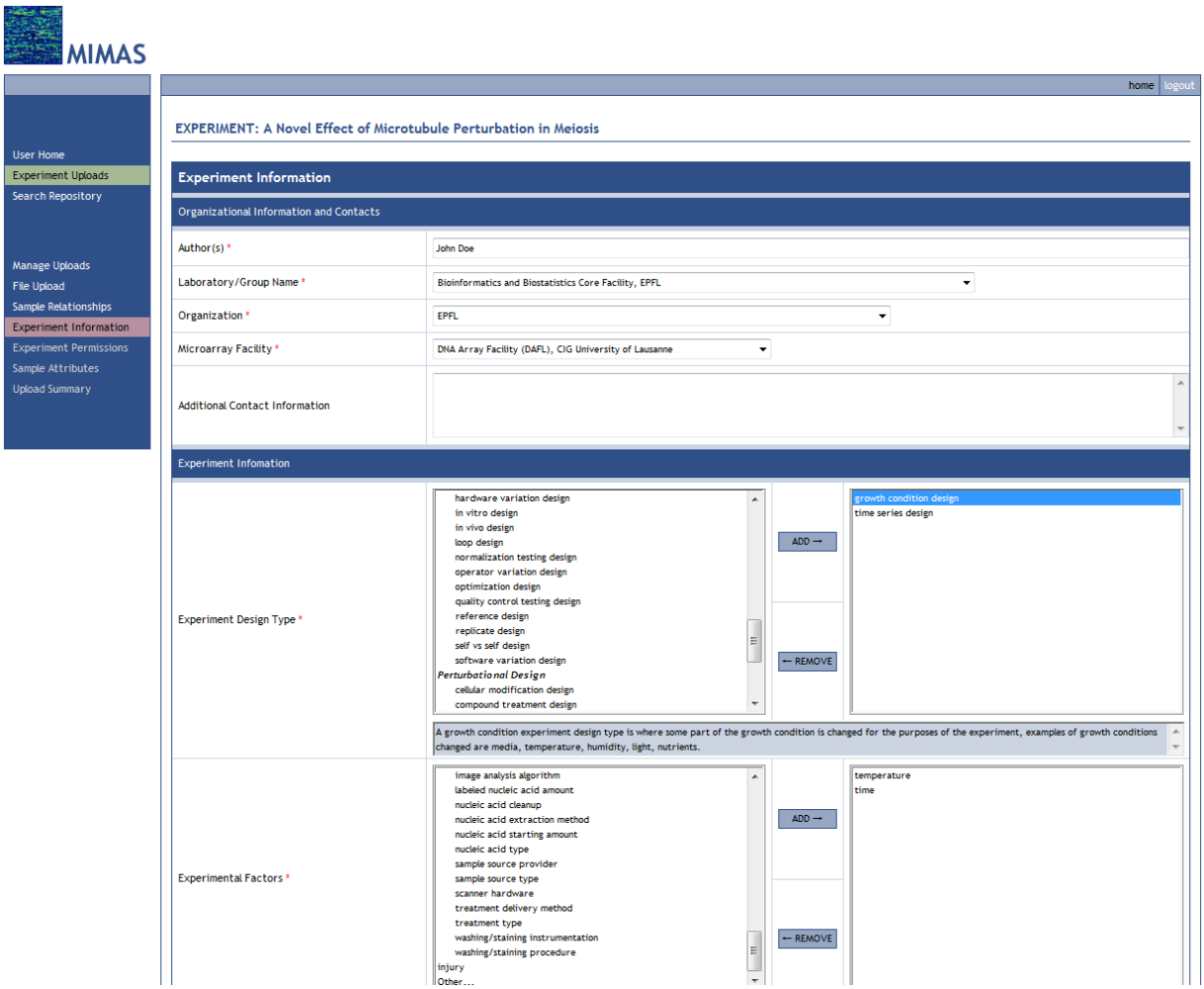


In general, you should have the same number of samples (replicates) in each experimental condition. MIMAS will warn you if you are breaking this rule but in the rare case where the numbers are not the same you may override this warning.

Experiment Information

The "*Experiment Information*" page, shown in Figure 18, is the first page in the workflow where you really start to describe the MIAME (Minimum Information About Microarray Experiments) attributes for your experiment. The Microarray Gene Expression Data Society (MGED - for more information, please see http://www.mged.org/), the organization which created MIAME, is focused on establishing standards for microarray data annotation and exchange. In MIMAS, MGED/MIAME annotation can be broken up into three levels of coverage: 1) experiment, 2) experimental condition, and 3) sample/hybridization. The "*Experiment Information*" page is where you complete the experiment-level attributes, those attributes which apply to all of the samples in your experiment. The "*Experiment Information*" page is functionally divided into three sections: 1) *Organization Information and Contacts*, 2) *Experiment Design Information*, and 3) *Microarray Technology and Quality Control*.

Figure 18. MIMAS Experiment Information



Required fields are followed by a red asterisk *. By default your MIMAS personal information is used to fill in the "*Author(s)*", "*Laboratory/Group Name*", and "*Organization*" fields but this may be changed to the appropriate values. You will also notice that some attribute fields are already filled in for you and locked. These fields use information you already filled in on previous pages in the workflow so if you then later on change something upstream they will be properly propagated onto these fields by MIMAS. A short definition of each attribute is available if you hover over the attribute name, as shown in Figure 19. Please see Appendix A: *MIMAS MGED/MIAME Attribute Reference*, for more complete definitions and instructions regarding MGED/MIAME attributes.

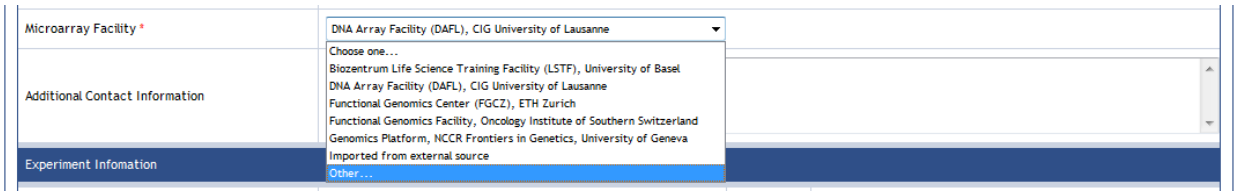
Figure 19. MGED/MIAME Definitions Popup

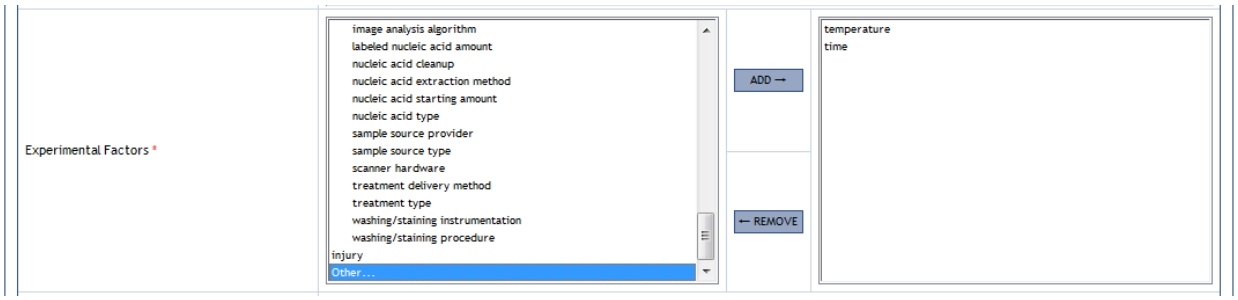
In the MIMAS system, MGED/MIAME attributes are divided into the following categories:

- Controlled vocabulary fields allowing a single selection (select-one drop down menu)
- Controlled vocabulary fields allowing multiple selection (select-multiple add/remove menu)
- Free text fields allowing short phrases, terms, identifiers, names, etc. (one line text box)
- Free text fields allowing free-form paragraph entry (multiple line text box)
- Numeric fields with or without controlled vocabulary units (number box with units drop down menu)

Controlled vocabulary drop down and add/remove menus use the latest MGED Ontology where appropriate (for more information, see http://mged.sourceforge.net/ontologies/index.php). Since MIMAS has an extensible vocabulary system, certain controlled vocabulary menus will have an "***Other…***" option located at the end of the menu, as shown in Figure 20.

Figure 20. MIMAS Menus with Extensible Vocabulary

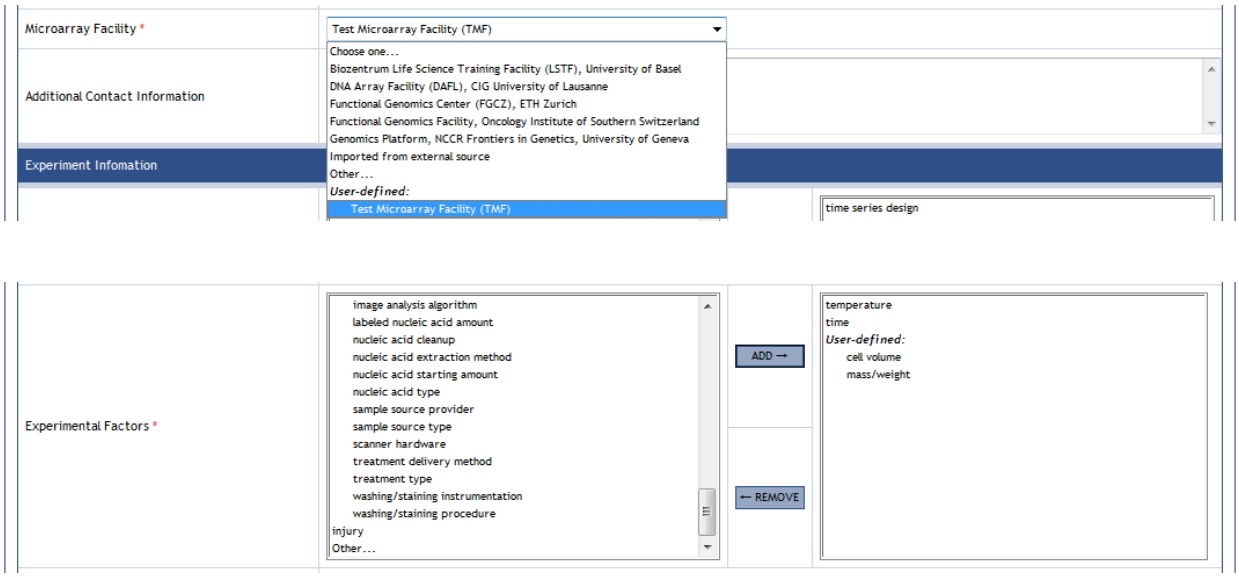




If the appropriate attribute value(s) is not located in the controlled vocabulary list, you may add user-defined values by selecting the "***Other…***" option. A prompt window will appear asking you for the user-defined value. Please make certain to read through the entire list of controlled vocabulary to make sure a valid attribute value does not exist before attempting to add a user-

defined value. Once you have added a user-defined value you will see it at the end of the menu list as shown in Figure 21.

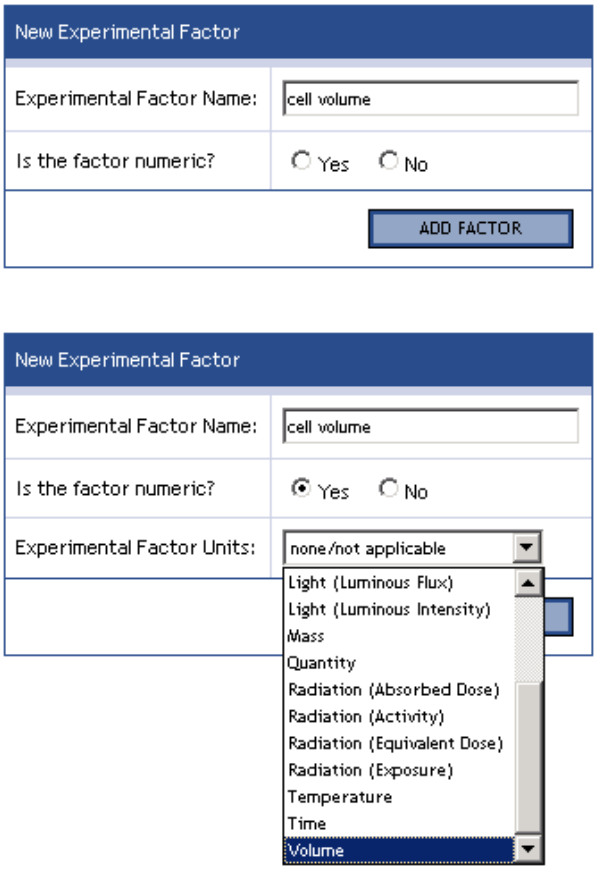
Figure 21. MIMAS User-defined Attribute Values



PLEASE NOTE: user-defined values, upon submission of your experiment upload, will be curated as best as possible and, if approved, will be added to the global controlled vocabulary library in MIMAS. If rejected, you will be notified and asked to revise or replace the attribute value. Please keep in mind that you are creating a piece of new vocabulary for ALL MIMAS users to eventually use if your user-defined attribute value is approved. Therefore, please use the standard and/or official scientific terminology for what you are trying to describe whenever it is possible.

The "*Experimental Factors*" attribute menu requires special attention when creating a user-defined value by selecting the "***Other…***" option. A special MIMAS window will appear as shown in Figure 22. You will be asked for the experimental factor name and whether it is numeric or not. If the factor is not numeric then the form is complete and you may click the "*Add Factor*" button to create the user-defined experimental factor. If the factor is numeric a new menu will appear asking whether the experimental factor has units. MIMAS gives you a choice of different unit groups based on different standard scientific measurement types. By default the "*none/not applicable*" option is chosen if the user-defined experimental factor is numeric but has no units. Otherwise, choose the appropriate unit group for your user-defined factor and click the "*Add Factor*" button.

Figure 22. MIMAS User-defined Experimental Factor Creation



Certain controlled vocabulary menus also have a "*none/not applicable*" option available for selection when it is appropriate. If the attribute is required (*) then the "*none/not applicable*" option is located at the end of the menu list, just before the "***Other…***" option if one exists. For controlled vocabulary menus that are not required the "*none/not applicable*" is selected by default.

Experimental Condition and Sample Attributes

The "*Sample Attributes*" page, shown in Figure 23, is the page where you complete the experimental condition-level, attributes which apply to an entire experimental condition, and sample/hybridization-level, attributes which apply to only specific sample/hybridization. The "*Sample Attributes*" page is functionally divided into six sections: 1) *Microarray Technology and Quality Control*, 2) *Biomaterial Characteristics*, 3) *Hybridization Protocol*, 4) *Image Analysis and Data Processing*, 5) *Experimental Factor Details* (if appropriate), and 6) *Cross-references*.

Figure 23a. MIMAS Experiment Condition and Sample Attributes

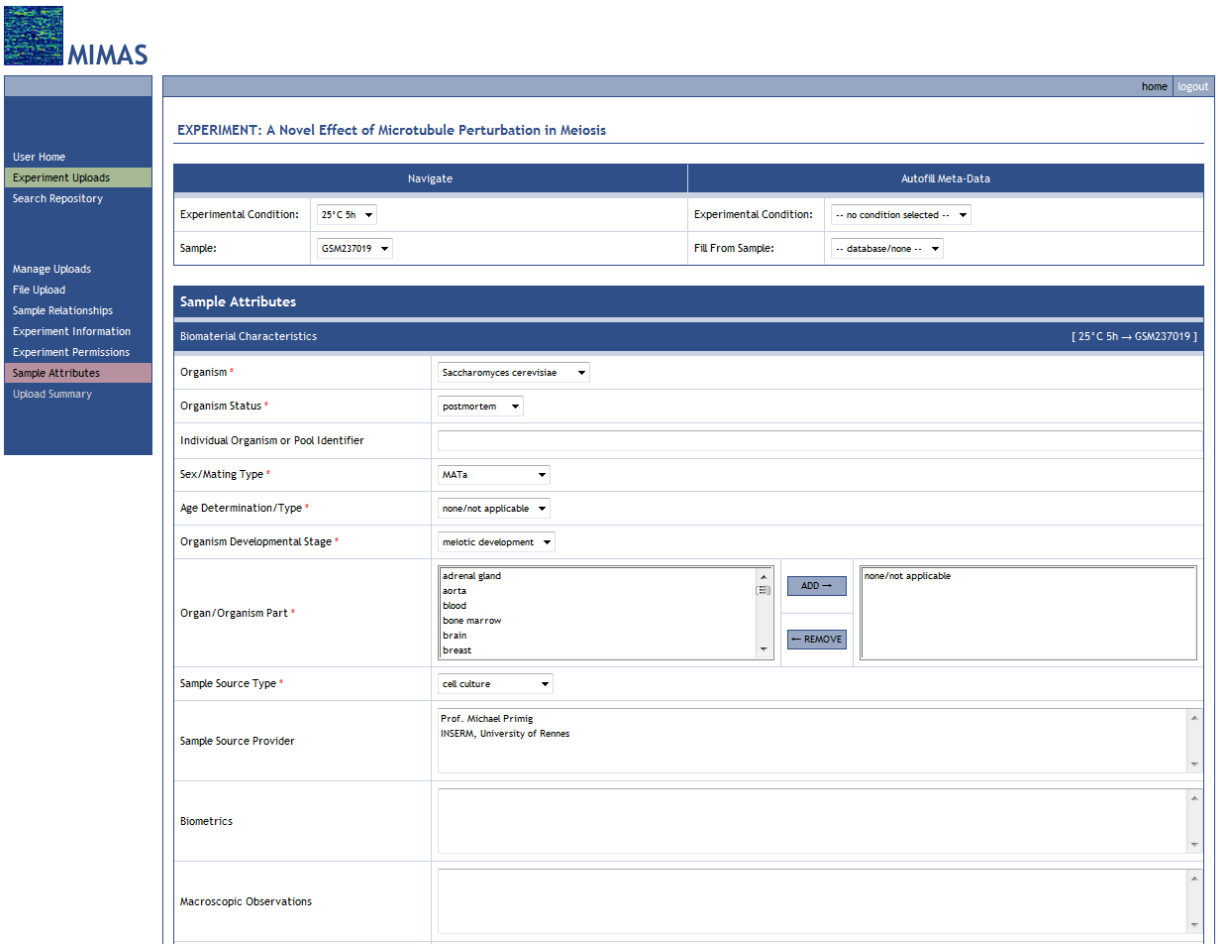


Figure 23b. MIMAS Experiment Condition and Sample Attributes

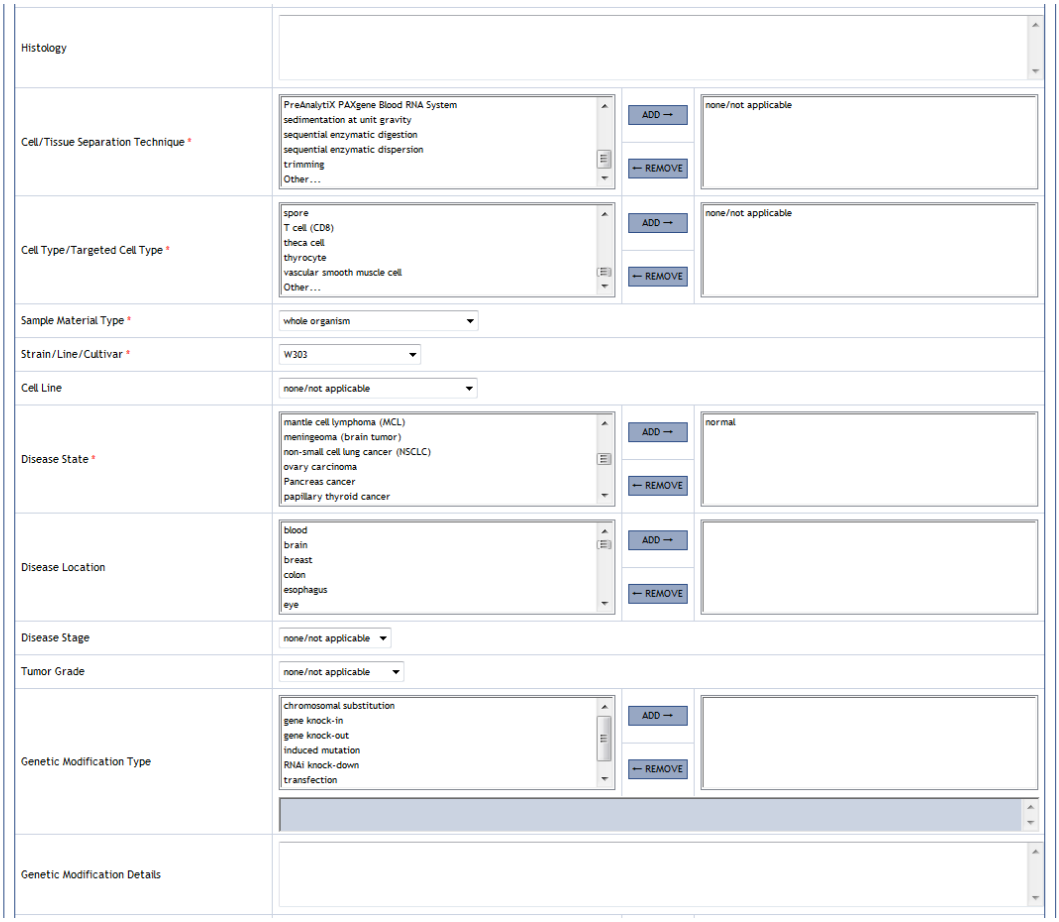


Figure 23c. MIMAS Experiment Condition and Sample Attributes

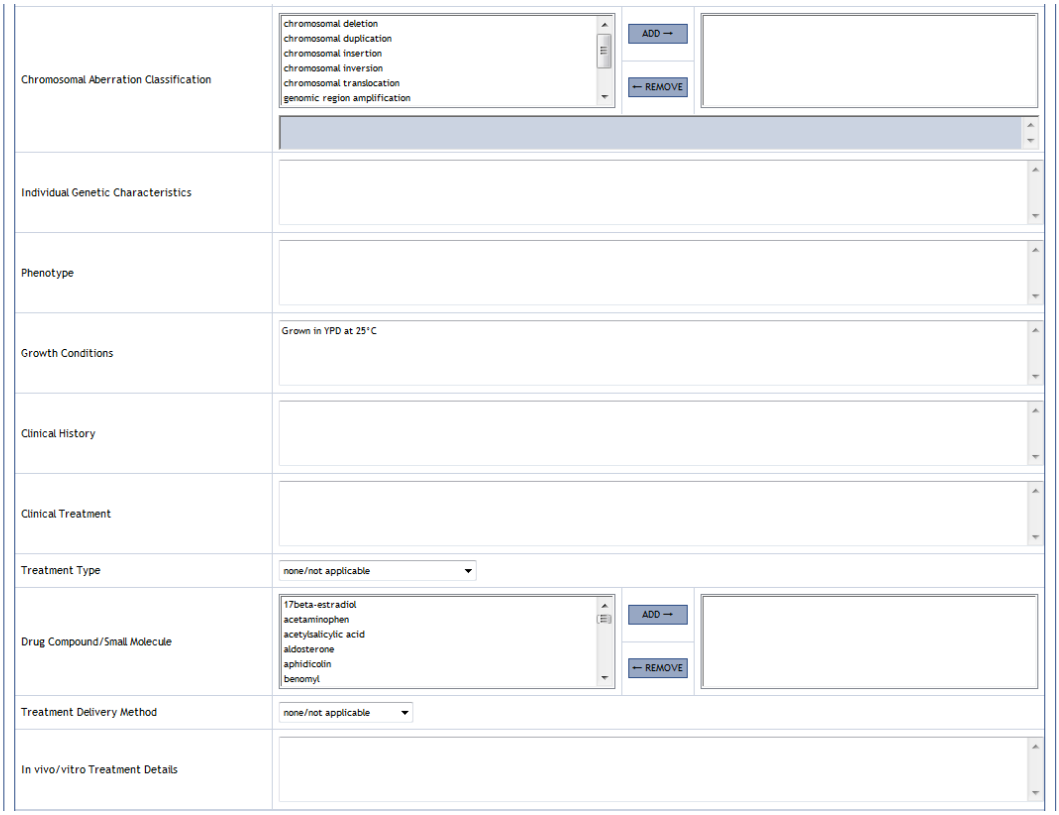


Figure 23d. MIMAS Experiment Condition and Sample Attributes

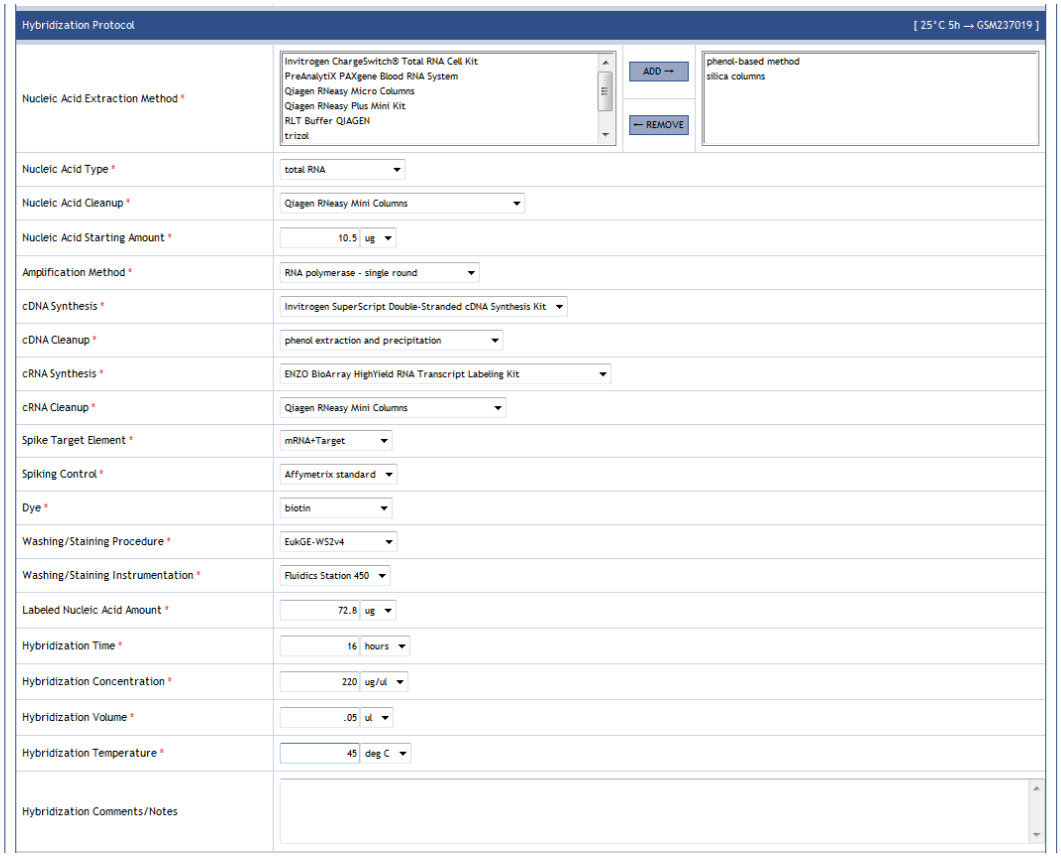
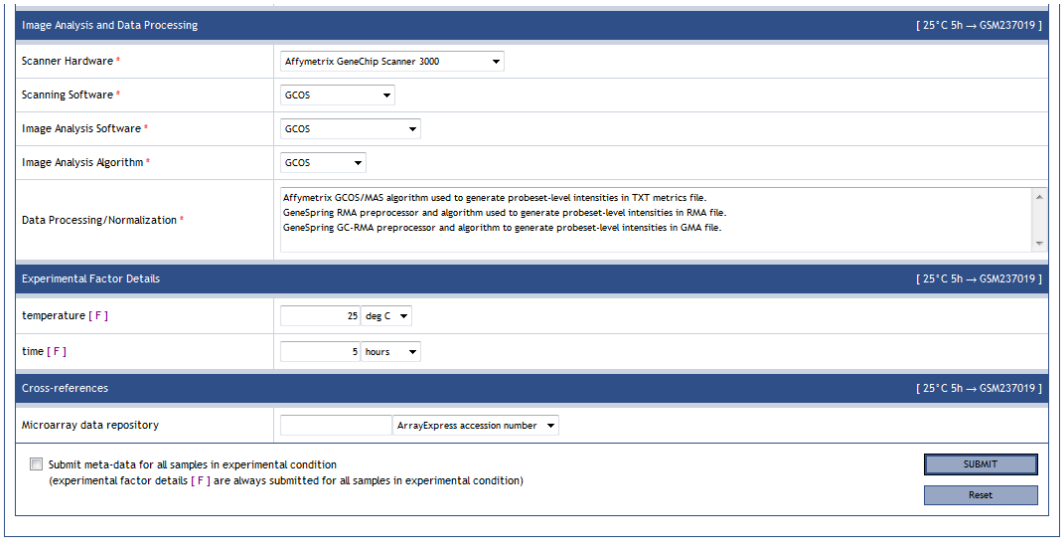


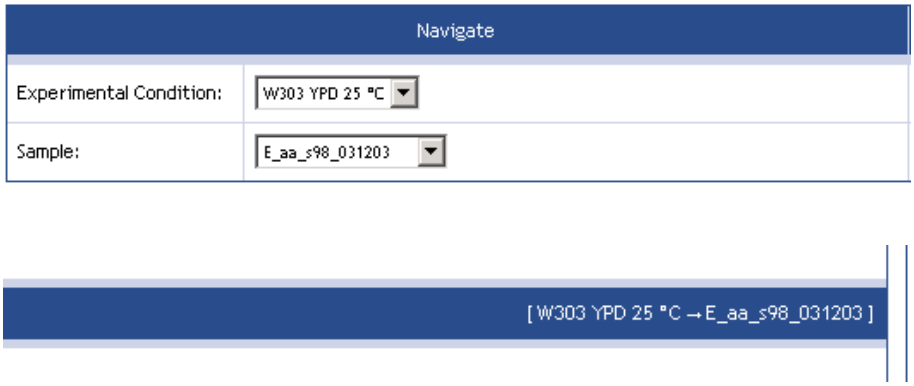
Figure 23e. MIMAS Experiment Condition and Sample Attributes



At the top of the "*Sample Attributes*" page, you will see two areas: 1) Navigate, as shown in Figure 24, and 2) Autofill Meta-Data. We will discuss the Navigate menu first and the Autofill menu and tool later. The Navigate menu tells you what experimental condition you are in and exactly which sample you working on. On the right-hand side of each section header you will also see the experimental condition and sample you are presently working so that if you are scrolled down on the page you can quickly find this information. As its name states, with Navigate menu you can move between different experimental conditions and samples. The experimental condition drop down menu gives you a list of your experimental conditions in the order that you defined them on

the "*Sample Relationships*" page. If you change the order it will be automatically reflected here. The sample drop down menu gives you an alphabetical list of samples in the experimental condition selected. When you come to the "*Sample Attributes*" page either by completing the "*Experiment Information*" page or clicking on the main navigation area link, MIMAS takes you to the first sample alphabetically in the first experimental condition in your ordered list. Work then proceeds through each sample in that experimental condition and then to the first sample alphabetically in the next experimental condition in your ordered list and so on.

Figure 24. MIMAS Experimental Condition and Sample Attributes Navigation



Required fields are followed by a red asterisk *. You will notice that again MIMAS has already filled in and locked some attribute fields based on information in previous pages in the workflow. So if you go back to these pages and change something they will be properly propagated onto these fields by MIMAS. As in the "*Experiment Information*" page, a short definition of each attribute is available if you hover over the attribute name, as was shown previously in Figure 19. Please see Appendix A: *MIMAS MGED/MIAME Attribute Reference* for more complete definitions and instructions regarding MGED/MIAME attributes.

As shown previously in Figure 20, controlled vocabulary drop down and add/remove menus will also have an "***Other…***" option located at the end of the menu if extensible vocabulary is allowed for this attribute. Please make certain to read through the entire list of controlled vocabulary to make sure a valid attribute value does not exist before attempting to add a user-defined value.

PLEASE NOTE: user-defined values, upon submission of your experiment upload, will be curated as best as possible and, if approved, will be added to the global controlled vocabulary library in MIMAS. If rejected, you will be notified and asked to revise or replace the attribute value. Please keep in mind that you are creating a piece of new vocabulary for ALL MIMAS users to eventually use if your user-defined attribute value is approved. Therefore, please use the standard and/or official scientific terminology for what you are trying to describe whenever it is possible.

Certain controlled vocabulary menus also have a "*none/not applicable*" option available for selection when it is appropriate. If the attribute is required (*) then the "*none/not applicable*" option is located at the end of the menu list, just before the "***Other…***" option if one exists. For controlled vocabulary menus that are not required the "*none/not applicable*" is selected by default.

As shown in Figure 25, certain attributes will have an **[F]** flag located next to them. This means that you selected these fields as experimental factors in the "*Experiment Information*" page and that you need to pay special attention because the experimental factor detail information you put here tells MIMAS what you varied in your experiment. If the "*Experimental Factor Details*" section exists for your experiment, every attribute field there will have this flag. Some experimental factors, like "*Strain/Line/Cultivar*" shown in Figure 24, are also standard MGED/MIAME fields and will have an **[F]** flag but remain in their own section and not in the "*Experimental Factor Details*" section. When filling in fields with an **[F]** flag, please note that you should leave them blank if you

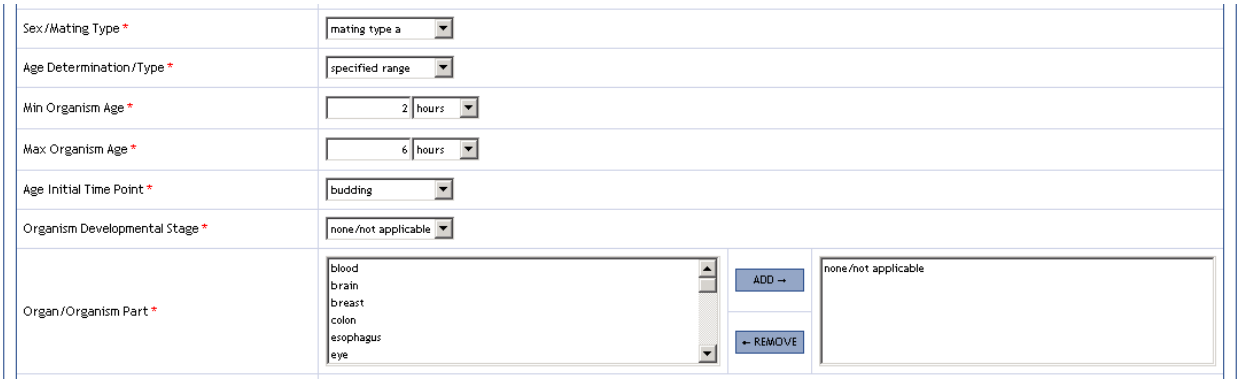
did not vary this experimental factor in the experimental condition for the pages you are working on. For example, if "*dose*" was selected as an experimental factor in the "*Experiment Information*" page, it will appear on the "*Sample Attributes*" page in the "*Experimental Factor Details*" section as a numeric box with a units drop down menu. You would leave this field blank if you were working on the "*Sample Attributes*" pages for the experimental condition where you gave no drug to the organism (i.e. the control group).

Figure 25. Experimental Factor Flag

A special required field on the "*Sample Attributes*" page is the "*Age Determination/Type*" field, shown in Figure 26a. Depending on which option you pick in the drop down menu, other attribute fields may appear which ask for additional information. If the "*specified range*" option is selected, "*Min Organism Age*" and "*Max Organism Age*" fields will appear below it, as shown in Figure 26b. If the "*specified single*" option is selected, then the "*Organism Age*" field will appear. In addition, the for either the "specified range" or "specified single" options, the "*Age Initial Time Point*" option will appear. Make sure to complete all of these fields if they automatically appear.

Figure 26a. MIMAS Sample Attribute "*Age Determination/Type*" Selection

Figure 26b. MIMAS Sample Attribute "*Age Determination/Type*" Additional Fields



When you have completed all of the required and other relevant attributes on the "*Sample Attributes*" page, you have two options by which you may submit the data to MIMAS. The first option is to simply click on the "*Submit*" button, then the data will be submitted only for the sample you are working on (shown in the Navigate menu and on the section headers, Figure 24) and

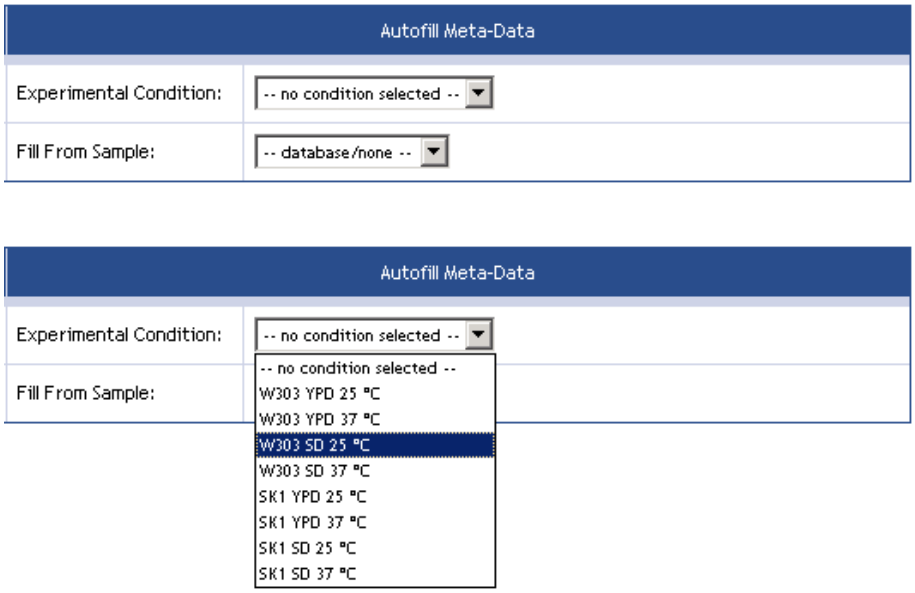
MIMAS will take you to the next sample in that experimental condition or, if you are on the last sample in that condition, the first sample of the next experimental condition. The second option uses the "*Fill Meta-Data for Entire Condition*" checkbox at the bottom of the "*Sample Attributes*" page, shown previously in Figure 23e. If you click on this checkbox before clicking the "*Submit*" button, then you are telling MIMAS to submit the data on this page for ALL of the samples in the present experimental condition. This is an extremely useful feature for quickly propagating sample attribute data since samples within the sample experimental condition have very similar annotation. You can then navigate to these samples later to make any appropriate changes. **PLEASE NOTE:** if data already exists for other samples in the experimental condition, they will be COMPLETELY overwritten. After clicking the "Submit" button with the "*Entire Condition*" option turned on, MIMAS will then forward you to the first sample of the next experimental condition, skipping the rest of the samples in the experimental condition that were automatically completed. The "*Entire Condition*" option will remain on and MIMAS uses the following events to determine how this feature stays on:

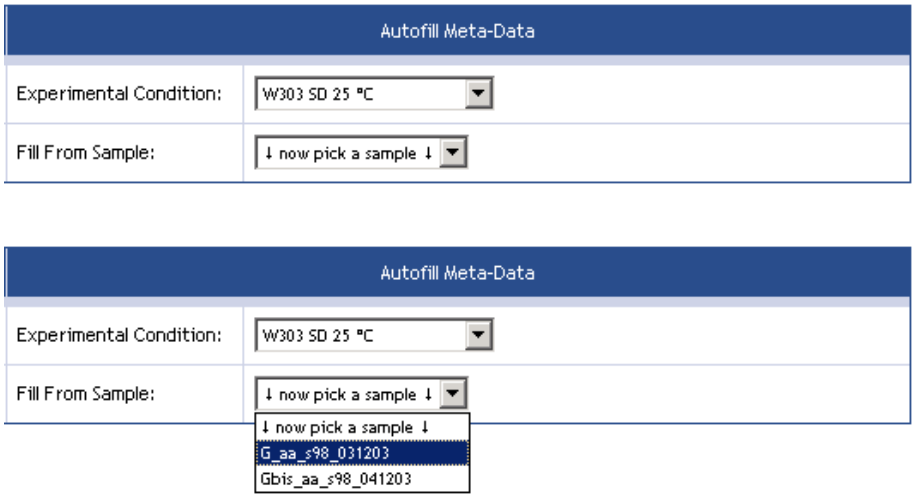
- If the "*Entire Condition*" checkbox is turn ON, it will stay ON for every "*Sample Attributes*" page until you decide to turn if OFF or leave the "*Sample Attributes*" pages to go to another area of MIMAS
- If the "*Entire Condition*" checkbox is OFF or turned OFF, it will stay OFF for every "*Sample Attributes*" page until you decide to turn it ON
- If you leave the "*Sample Attributes*" pages and then return, the "*Entire Condition*" checkbox will always be OFF

For experimental factor attribute fields which have an **[F]** flag, data from this field is ALWAYS submitted for all of the samples in the experimental condition regardless of whether you have the "*Entire Condition*" checkbox turned on or off.

After submitting sample attribute data and being forwarded to the next "*Sample Attributes*" page, MIMAS has another useful tool to speed up completion of the sample attribute information. The "*Autofill*" tool, shown in Figure 27, allows you to take sample attribute data from any other sample in any experimental condition in your experiment and fill it into the fields on the page for the present sample. The tool does not submit the page for you, thus allowing you to then make any appropriate changes after it has filled in the other samples data. To use the "Autofill" tool, simply select the source experimental condition from the drop down menu and then select the source sample in that condition from the sample drop down menu. If you already know the name of the source sample you want to use, you can omit first selecting the source experimental condition and just select the sample to fill from using just the sample drop down menu.

Figure 27. MIMAS Experimental Condition and Sample Attributes Autofill Tool





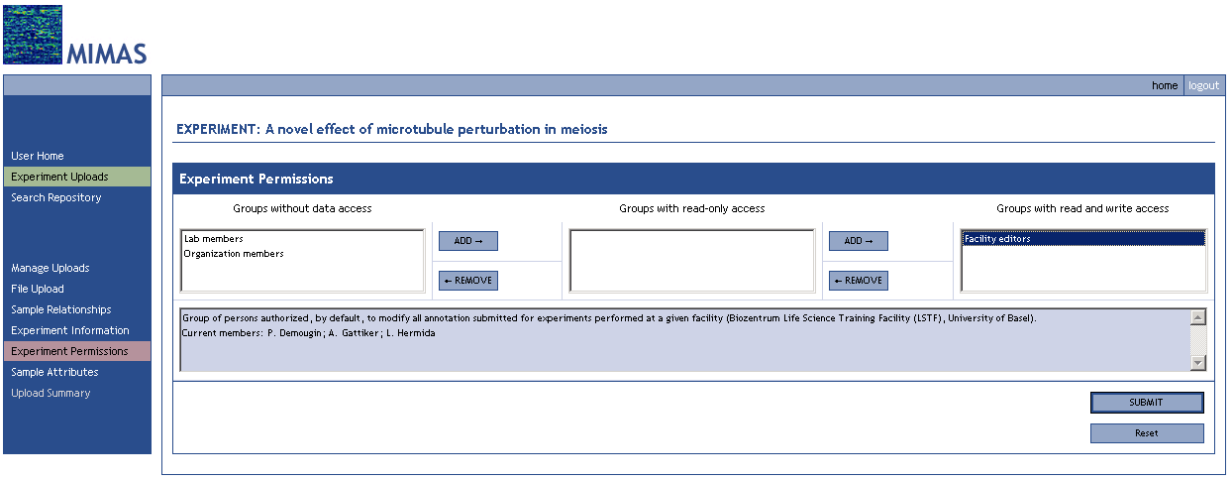
By using a combination of the "*Entire Condition*" and "*Autofill*" tools, you can quickly and accurately complete all of the experimental condition and sample attributes for your experiment. MIMAS transparently keeps a flag of which "*Sample Attributes*" pages you have personally completed or asked MIMAS to automatically complete (using the "*Entire Condition*" tool) and when a completion flag exists for every sample in your experiment MIMAS will forward you to the "*Upload Summary*" page. If you visit pages upstream in the experiment upload workflow and make changes which affect the "*Sample Attributes*" pages, MIMAS will then reset these completion flags which then forces you to revisit and submit these pages to make sure any propagated changes are what you have intended. It is always a good idea to review each and every "*Sample Attributes*" page to make sure the data you have submitted is correct.

Experiment Permissions

The "*Experiment Permissions*" page, shown in Figure 28, allows setting which other users of the MIMAS system are allowed to view and/or modify the annotation of your experiment. This allows a smoother annotation process where different people (sample provider, platform technician, data analyst) may supply different parts of the complete annotation information.

Permissions are assigned at the level of groups (rather than individual users). Clicking on the name of a groups brings up a description of the group and a list of its current members. Groups may be moved between the sections "Groups without data access", "Groups with read-only access" and "Groups with read and write access" by means of the ADD and REMOVE buttons.

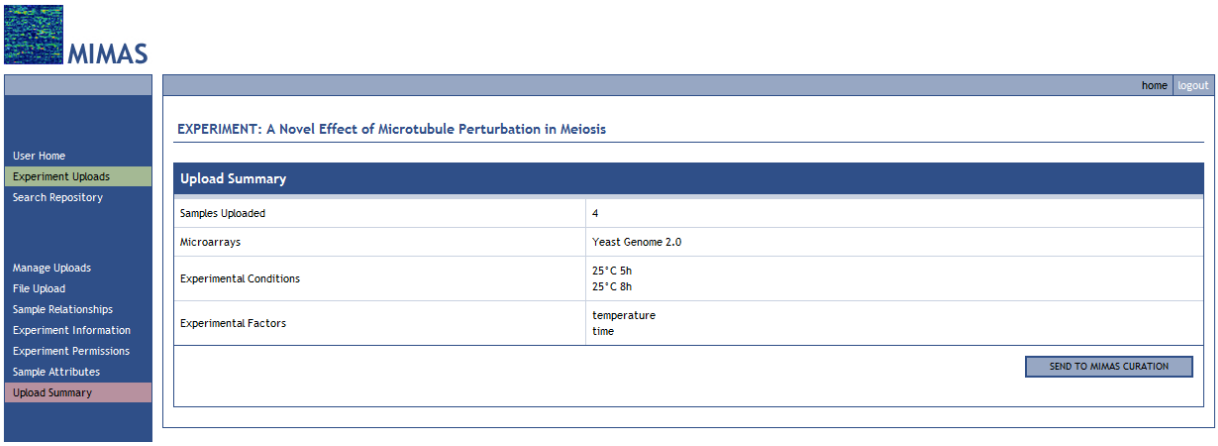
Figure 28. MIMAS Experiment Permissions



Experiment Upload Summary

The "*Upload Summary*" page, shown in Figure 29, marks the completion of the microarray experiment data upload and annotation process. Summary information about your experiment is listed on this page for your review. Before submitting your experiment to MIMAS curation, go back through each of the pages in the workflow and review all of the data to make sure your experiment upload and annotation is complete and accurate. Remember that MIMAS saves you upload experiments in the exact state that you left them indefinitely, so if you choose to log out and review everything at a later date it will be waiting for you.

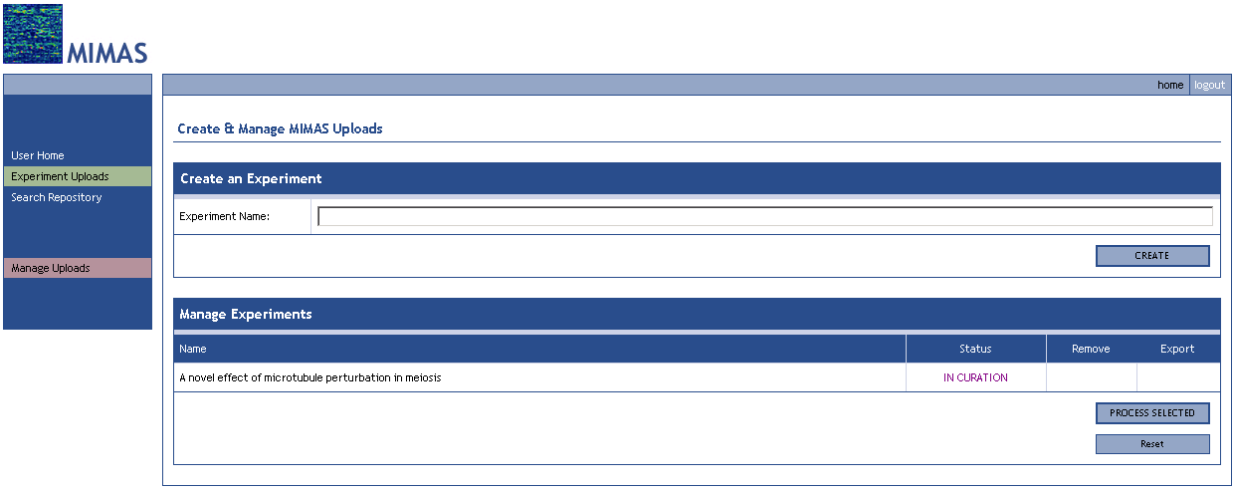
Figure 29. MIMAS Experiment Upload Summary



Curation and Repository Submission

After clicking the "*Send to MIMAS Curation*" button, your experiment will be sent to a MIMAS Curator for expedited review and you will be returned to your personal upload management page, shown in Figure 30. You will notice that the status of your experiment is now "*In Curation*" and you cannot edit the experiment or remove it. If your experiment is curated and approved it will show as "*SUBMITTED*". Remember that you are free to start any number of other experiments while this process is underway.

Figure 30. MIMAS Upload Management: Experiment in Curation



When MIMAS curators have completed the curation process, you will be notified via email. If your submission failed and you need to correct something with your experiment upload, you will also be notified and you will see the experiment in your "*Manage Experiments*" listing with a "*Working*" status.

ArrayExpress Data Export

Experiments that have been fully annotated and that have gone through the curation process are marked with the "SUBMITTED" status in the *"Manage Uploads"* page (Figure 30). Pressing the "EXPORT" will generate an archive containing experiment annotation and files. It takes a few minutes to generate the file, which is then made available in the *"Alerts & Messages"* section of the *"User Home"* page (Figure 31). Press the "Retrieve" link to download the archive.

Figure 31. MIMAS User Home: Download Request Job Complete Message

The zip-formatted archive will contain the following files:

- Experiment files in CEL, TXT, RMA and/or GMA format, as submitted into MIMAS.
- A file named 'tab2mage_log.txt' that should be carefully reviewed, as it may contain errors or missing information that should be added to the spreadsheet (see below) before submission.
- A file named 'spreadsheet.txt' that should be opened in Microsoft Excel or another spreadsheet program. This is the actual file that contains the annotation that will be used in ArrayExpress to describe your samples. Carefully review and complete the annotation as necessary. If you modify the spreadsheet file, make sure to insert the modified file back into the zip file, overwriting the old spreadsheet file.

To upload your file to ArrayExpress, follow the guidelines at http://tab2mage.sourceforge.net/docs/aesubs.html.

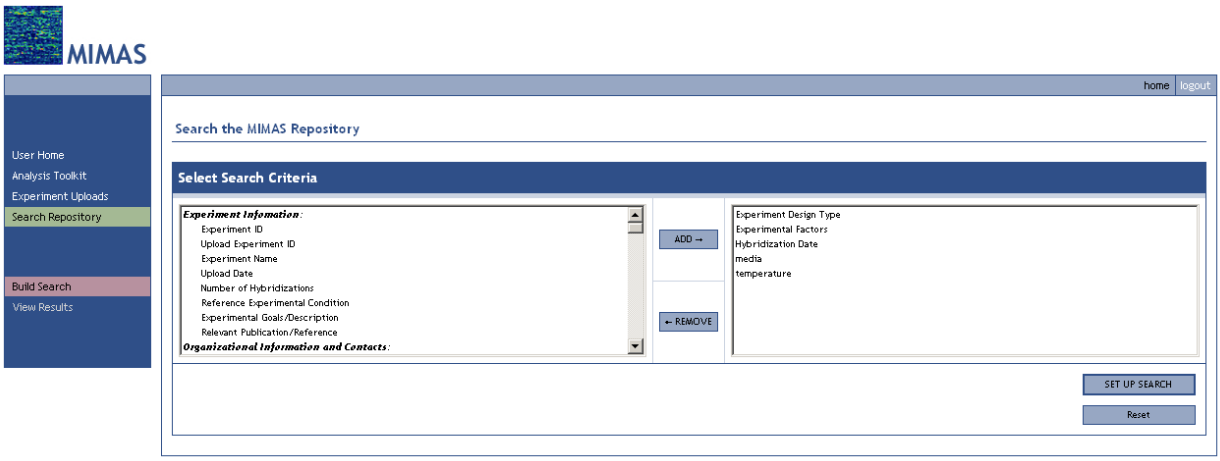
For more information about the spreadsheet format, refer to http://tab2mage.sourceforge.net/docs/spreadsheet.html.

The procedure may be quite complex. Do not hesitate to contact the MIMAS administrator for assistance. The submission can also be done for you by the MIMAS administrator.

5 MIMAS Repository Searching and Data Retrieval

To search for microarray data in the MIMAS repository, click on the "*Search Repository*" main menu link found in the main navigation area and you will be forwarded to the main search page, shown in Figure 32. By default, you can only search for samples that you own or were given access to. Other users would have to give you access to their data in order for you to search for it. To begin a search, you need to first select what general criteria you would like to search by. All of the MGED/MIAME and additional attribute fields you saw in the experiment upload section are available to you as search criteria and they are divided into the same functional groups: 1) *Experiment Information*, 2) *Organization Information and Contacts*, 3) *Microarray Technology and Quality Control*, 4) *Sample Information* (new), 5) *Biomaterial Characteristics*, 6) *Hybridization Protocol*, 7) *Image Analysis and Data Processing*, and 8) *Experimental Factor Details*. After selecting from these criteria, click the "*Set up Search*" button.

Figure 32. MIMAS Search Page: Select Search Criteria

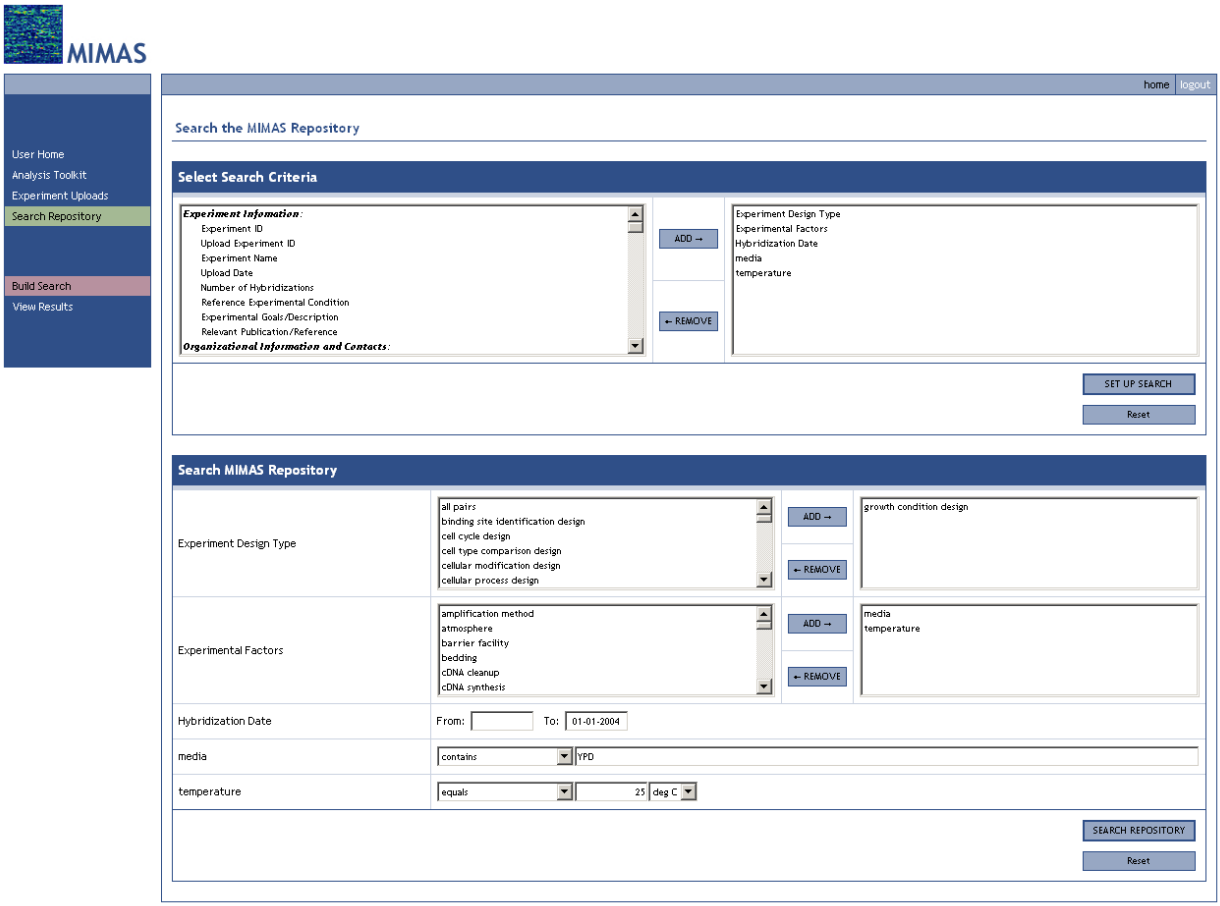


As shown in Figure 33, MIMAS will then produce a new section called "*Search MIMAS Repository*" based on your search criteria. Using this search form you can create your detailed query. The type of detailed interface you are given for each criterion depends on what type of field it is:

- Controlled vocabulary fields give a select-multiple add/remove menu
- Free text fields give a string search box with string search modifiers (contains, equals, starts with, ends with, does not contain)
- Numeric fields give a numeric search box with numeric search modifiers (equals, less than, less than or equal to, greater than, greater than or equal to, does not equal) and if they require units then the controlled vocabulary units drop down menu
- Date fields give a from/to date range set of search boxes

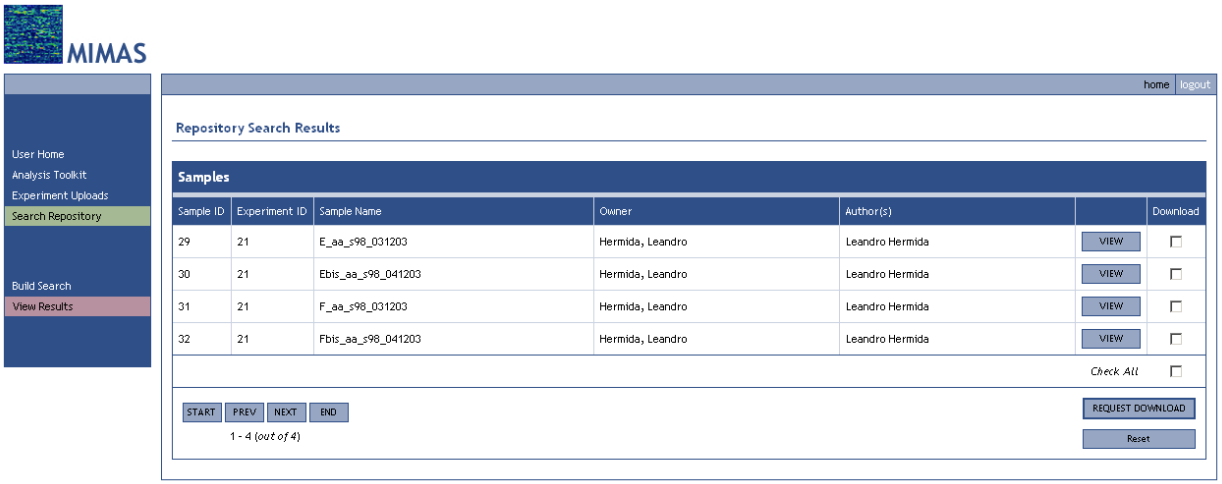
If you selected multiple criteria, you have effectively asked MIMAS to find samples which match ALL of the criteria, not ANY of them. For example, the query built in Figure 33 asks MIMAS to find samples which you own AND are a "*growth condition design*" experiment design type AND have "*media*" and "*temperature*" as experimental factors AND have a hybridization date on or before January 1, 2004 AND contained "*YPD*" as a growth media AND used 25 °C as the growth temperature. If you want to find samples which match ANY of your selected criteria please make separate queries for each criterion. An improvement in the search interface is in development to allow for more flexible search queries.

Figure 33. MIMAS Search Page: Build Detailed Search Query



If you want to clear just the detailed search query you can do so by clicking the "*Reset*" button in the lower detailed search form. If you would like to completely start over with a new query and set of criteria, click the "*Reset*" button in the "*Select Criteria*" search form. After clicking the "*Search Repository*" button, MIMAS will search the database to find samples and forward you to the search results page, shown in Figure 34. Here you have a summary table of matching samples. If there are many samples which match your search criteria, the summary table may extend across multiple web pages. By using the "*Start*", "*Prev*", "*Next*" and "*End*" buttons you can navigate through the different pages of the summary table.

Figure 34. MIMAS Search Results Page



You can view the detailed description for a particular sample by clicking the "*View*" button on the row of that sample. This will take you to the "*View Sample*" page, shown in Figure 35. From this page you can navigate to the other sample description pages by using identical "*Start*", "*Prev*", "*Next*" and "*End*" buttons located at the top of the page. A short definition of each attribute is available if you hover over the attribute name, as shown in Figure 18. Please see Appendix A: *MIMAS MGED/MIAME Attribute Reference* for more complete definitions and instructions regarding MGED/MIAME attributes.

Figure 35a. MIMAS View Sample Page

Figure 35b. MIMAS View Sample Page

Sample Source Type	cell culture
Cell/Tissue Separation Technique	Filtered spores and mycelium
Cell Type/Targeted Cell Type	mycelium spore
Sample Material Type	organ/organism part (mixed cell types)
Strain/Line/Cultivar	Delta l Delta t
Disease State	normal
Hybridization Protocol	
Nucleic Acid Extraction Method	phenol-based method silica columns
Nucleic Acid Type	total RNA
Nucleic Acid Cleanup	Qiagen RNeasy MinElute
Nucleic Acid Starting Amount	2 ug
Amplification Method	RNA polymerase - single round
cDNA Synthesis	Affymetrix GeneChip One-Cycle cDNA Synthesis Kit
cDNA Cleanup	Affymetrix GeneChip Sample Cleanup Module
cRNA Synthesis	Affymetrix GeneChip Expression 3'-Amplification Kit for IVT Labeling
cRNA Cleanup	Affymetrix GeneChip Sample Cleanup Module
Spike Target Element	mRNA+Target
Spiking Control	Affymetrix standard
Dye	biotin
Washing/Staining Procedure	FS450_0004
Washing/Staining Instrumentation	Fluidics Station 450
Labeled Nucleic Acid Amount	102 ug
Hybridization Time	16 hours
Hybridization Concentration	.05 ug/ul
Hybridization Volume	200 ul
Hybridization Temperature	45 deg C
Image Analysis and Data Processing	
Scanner Hardware	Affymetrix GeneChip Scanner 3000
Scanning Software	GCOS
Image Analysis Software	GCOS
Image Analysis Algorithm	GCOS
Data Processing/Normalization	Affymetrix GCOS/MAS algorithm used to generate probeset-level intensities in TXT metrics file. GeneSpring RMA preprocessor and algorithm used to generate probeset-level intensities in RMA file. GeneSpring GC-RMA preprocessor and algorithm to generate probeset-level intensities in GMA file.
Experimental Factor Details	
time	9 hours
Cross-references	

From the search results page you can choose samples to download by checking the download check box in the appropriate sample row or by checking the "*Check All*" checkbox to get all of the samples on the page. Click the "*Request Download*" button and MIMAS will fetch the sample data you requested. Because this process can take some time, MIMAS will run the process in the background and you will receive at Job ID for the download request. You can continue working as usual while this process is running.

Figure 36. MIMAS Sample Download Request

When MIMAS has completed the download request you will see an alert and link on your personal user home page, shown in Figure 37. Click on the >>Retrieve<< link to obtain the compressed archive of the requested sample data and annotations.

Figure 37. MIMAS User Home: Download Retrieval

A MIMAS MGED/MIAME Attribute Reference

Attribute	*Definition*	*Notes*
Author(s)	Person(s) conducting microarray experiment.	***Default Entry**: User Full Name* *Please change default entry where appropriate.*
Laboratory/Group Name	Primary laboratory of person(s) conducting microarray experiment.	***Default Entry**: User Laboratory* *Please change default entry where appropriate.*
Organization	Institute or organization where laboratory resides.	***Default Entry**: User Organization* *Please change default entry where appropriate.*
Microarray Facility	Facility in charge of microarray services.	
Additional Contact Information	Additional contact details of person(s) conducting microarray and/or related experiment(s).	*Full Name(s) and Address(es)*
Experiment ID	MIMAS Repository Experiment ID.	*Automatically generated.*
Upload Experiment ID	MIMAS Upload Experiment ID.	*Automatically generated.*
Experiment Name	Experiment name/short description.	*A single statement complete description of the experiment. Name could be in the same form and style as a journal paper title describing experiment.* ***Example**: "Study of GAL-SPO13 Overexpression during Mitosis in S. cerevisiae"*
Experiment Design Type	High-level classification of type of experiment being performed.	
Upload Date	Date experiment was uploaded into MIMAS.	
Experimental Factors	The factors in the study that are experimental parameters or regarded as influencing the experimental results.	
Number of Hybridizations	Number of hybridizations performed in the experiment.	*Automatically calculated from File Upload.*
Reference Experimental Condition	If exists, the name of the reference experimental condition to which all other experimental conditions are compared.	*If exists, the biological reference experimental condition/control group, like normal in normal vs. diseased, wild-type in wild-type vs. mutation, untreated in treated vs. untreated, etc.*
Experimental Goals/Description	Detailed description of experiment and relevant goals.	*Please describe in detail your experiment goals, description, any relevant background information, etc. This is a very important field and should be much more than one sentence!*

Attribute	*Definition*	*Notes*
Relevant Publication/Reference	Relevant publications associated with experiment.	*Author List. Title. Journal. Date;Volume(Issue):Pages. Epub Date (if applicable). PMID: Number.* ***Example**: Miyake T, Reese J, Loch CM, Auble DT, Li R. Genome-wide analysis of ARS (autonomously replicating sequence) binding factor 1 (Abf1p)-mediated transcriptional regulation in Saccharomyces cerevisiae. J Biol Chem. 2004 Aug 13;279(33):34865-72. Epub 2004 Jun 10. PMID: 12200417.*
Relevant Publication / Pubmed ID	Relevant publications associated with experiment.	
Relevant Web Site / URL	Relevant web site URL associated with experiment.	
Array Series	Manufacturer given array series name.	*Automatically selected from Sample Relationships.*
Array Design Name	Manufacturer given unique array name.	*Automatically selected from Sample Relationships.*
Array Platform Type	Name of manufacturer or technology type used to create array.	*Automatically selected from Sample Relationships.*
Array Provider	The primary contact (manufacturer) for the information on array design.	*Automatically selected from Sample Relationships.*
Number of Replicates	Number of replicates performed for this sample.	*Automatically calculated from Sample Relationships.*
Replicate Type	Experimental stage at which replication was performed.	
Replicate Comments/Notes	Details about the type of replication used in the experiment.	*Describe any non-standard procedures or if you have any special notes specific to the replication used.*
Sample ID	MIMAS Repository Sample ID.	*Automatically generated.*
Sample Owner	MIMAS User who owns sample.	
Sample Name	Original user sample name.	*Original file name from File Upload.*
Experimental Condition Name	User-supplied name of experimental condition to which sample belongs.	*Mapped to samples in Sample Relationships.*
Hybridization Date	Date hybridization was performed.	*Automatically extracted from data files.*
Organism	The genus and species (and subspecies) of the organism that the biomaterial is derived from.	
Organism Status	The stage, premortem or postmortem, at which the sample was processed for extraction of biomaterial.	*If the organism was alive (premortem) or dead (postmortem) when the sample was processed for extraction of biomaterial.*

Attribute	*Definition*	*Notes*
Individual Organism or Pool Identifier	Identifier or name of the individual organism or pool of organisms that the biomaterial is derived from. Use the same name in different samples if they are from the same organism or pool.	*Possible bar code number, serial number, or other individual organism or pool identifier for the organism that the biomaterial was derived from.*
Sex/Mating Type	Term applied to any organism able to undergo sexual reproduction in order to differentiate the individuals or types involved. Sexual reproduction is defined as the ability to exchange genetic material with the potential of recombinant progeny.	
Age Determination/Type	Whether age can be determined for the organism(s) that the biomaterial is derived from and if so, whether age is a single value or a range or values.	
Organism Age	The time period elapsed since an identifiable point in the life cycle of an organism. If a developmental stage is specified, the identifiable point would be the beginning of that stage. Otherwise the identifiable point must be specified.	
Min Organism Age	The minimum time period elapsed since an identifiable point in the life cycle of an organism. If a developmental stage is specified, the identifiable point would be the beginning of that stage. Otherwise the identifiable point must be specified.	
Max Organism Age	The maximum time period elapsed since an identifiable point in the life cycle of an organism. If a developmental stage is specified, the identifiable point would be the beginning of that stage. Otherwise the identifiable point must be specified.	
Age Initial Time Point	The identifiable point in the life cycle of an organism from which age measurements were taken.	
Organism Developmental Stage	The developmental stage of the organism's life cycle during which the biomaterial was extracted.	
Organ/Organism Part	The part of organism's anatomy or substance arising from an organism from which the biomaterial was derived - excludes cells.	
Sample Source Type	The original form in which the biomaterial was obtained/maintained.	
Sample Source Provider	The contact details of the resource (e.g. person, company, hospital, geographical location) used to obtain or purchase the biomaterial.	*Full Name(s) and Address(es)*
Biometrics	Important physical properties of the biomaterial, e.g. mass or height.	
Macroscopic Observations	Details of the macroscopic examination of the biomaterial.	

Attribute	*Definition*	*Notes*
Histology	Details of the microscopic morphology of the tissues that the biomaterial is derived from.	
Cell/Tissue Separation Technique	General technique used to separate tissues or cells from a heterogeneous sample.	
Cell Type/Targeted Cell Type	The type of cell used in the experiment if non-mixed, if mixed then the targeted cell type. The target cell type is the cell of primary interest. The biomaterial may be derived from a mixed population of cells although only one cell type is of interest.	
Sample Material Type	Whether the sample contain a purified cell type of a mixture of cell types.	
Strain/Line/Cultivar	For animals or plants, these are offspring that have a single ancestral breeding pair or parent as a result of brother x sister or parent x offspring matings. For microbes, these are isolates derived from nature or in the laboratory.	
Cell Line	The identifier for the established culture of a metazoan cell if one was used as a biomaterial.	
Disease State	The name of the pathology diagnosed in the organism that the biomaterial is derived from. The disease state is normal if no disease has been diagnosed.	
Disease Location	Anatomical location(s) of disease in the organism that the biomaterial is derived from.	
Disease Stage	The stage or progression of a disease in the organism that the biomaterial is derived from. Includes pathological staging of cancers and other disease progression.	
Tumor Grade	A descriptor used in cancer biology to describe abnormalities of tumor cells.	
Genetic Modification Type	A genetic modification synthetically introduced into the organism that the biomaterial is derived from.	
Genetic Modification Details	Genetic modification description and details.	
Chromosomal Aberration	An irregularity in the number or structure of chromosomes, usually in the form of a gain (duplication), loss (deletion), exchange (translocation), or alteration in sequence (inversion) of genetic material. Excludes simple changes in sequence such as mutations, and is usually detectable by cytogenetic and microscopic techniques.	

Attribute	*Definition*	*Notes*
Individual Genetic Characteristics	The genotype of the individual organism that the biomaterial is derived from. Individual genetic characteristics include polymorphisms, disease alleles, and haplotypes.	
Phenotype	The observable form taken by some character (or group of characters) in an individual or an organism, excluding pathology and disease. The detectable outward manifestations of a specific genotype.	
Growth Conditions	A description of the conditions used to grow organisms or parts of the organism. This includes isolated environments such as cultures and open environments such as field studies.	
Clinical History	Relevant clinical history and information about the individual organism that the biomaterial is derived from.	
Clinical Treatment	The current clinical treatment(s) of the patient that the biomaterial is derived from.	
Treatment Type	The type of manipulation applied to the biomaterial for the purposes of generating one of the variables under study.	
Drug Compound/Small Molecule	Name of drug, compound, solvent, chemical, etc., used in treatment.	
Treatment Delivery Method	Treatment delivery method.	
In vivo/vitro Treatment Details	Details and steps of manipulation applied to the biomaterial for the purposes of generating one of the variables under study.	
Nucleic Acid Extraction Method	Protocol used to extract nucleic acids from the sample.	
Nucleic Acid Type	The type of nucleic acid extracted.	
Nucleic Acid Cleanup	The method used to clean up nucleic acids.	
Nucleic Acid Starting Amount	Amount of nucleic acid used to start hybridization protocol.	
Amplification Method	The method used to amplify the nucleic acid extracted.	
cDNA Synthesis	The method used to synthesize cDNA.	
cDNA Cleanup	The method used to clean up cDNA.	
cRNA Synthesis	The method used to synthesize cRNA.	
cRNA Cleanup	The method used to clean up cRNA.	
Spike Target Element	Type of spike control used.	

Attribute	*Definition*	*Notes*
Spiking Control	Spike control qualifiers -- concentration, expected ratio.	
Dye	Name of dye used for labeling the extract.	
Washing/Staining Procedure	Washing and staining procedure used in hybridization protocol.	
Washing/Staining Instrumentation	Washing instrumentation hardware.	
Labeled Nucleic Acid Amount	Amount of labeled nucleic acid obtained after target synthesis.	
Hybridization Time	Hybridization time.	
Hybridization Concentration	Hybridization concentration.	
Hybridization Volume	Hybridization volume.	
Hybridization Temperature	Hybridization temperature.	
Hybridization Comments/Notes	Additional notes or comments on the hybridization procedure.	*Describe any non-standard procedures or if you have any special notes specific to this hybridization.*
Scanner Hardware	Scanner hardware model.	
Scanning Software	Scanning software used.	
Image Analysis Software	Image analysis software used.	
Image Analysis Algorithm	Image analysis algorithm used.	
Data Processing/Normalization	Details on any data processing and normalization strategies.	***Default Entry:*** *"Affymetrix GCOS/MAS algorithm used to generate probeset-level intensities in TXT metrics file. GeneSpring RMA preprocessor and algorithm used to generate probeset-level intensities in RMA file. GeneSpring GC-RMA preprocessor and algorithm to generate probeset-level intensities in GMA file."* *Please change default entry if appropriate.*

B RMA and GC-RMA File Creation Using GeneSpring

1. Open GeneSpring and the genome of interest (the array your CEL files are generated from).
2. Go to *File -> Import Data*. Browse to your CEL file directory.
3. Pick one CEL file from the list.
4. On the "*File Format*" popup, GeneSpring should have automatically picked "*AffyMetrix **Your Genome Name Here** CEL File*" and the correct genome should already be selected. Click "*Next*".
5. Select the preprocessor (RMA or GC-RMA). Remember which one you picked!
6. On the "*Selected Files*" popup, add all the rest of the CEL files from your experiment.
7. On the "*Sample Attributes*" popup, leave it blank and click "*Next*".
8. On the question "*N samples have been created. Would also you like to create an experiment from these samples?*", click "*NO*".
9. On the "*Sample Inspector*" popup, click on the "*Associated Files*" tab. Click on the *sample_name.txt* data file line and then click on the "*Extract File*" button. Do not accidentally extract the CEL file (we don't need that).
10. When asked where to save the extracted file, GeneSpring will automatically start in the GeneSpring data directory. Do not save the file here. Create a new folder in this directory called "*extracted files*" (if it doesn't already exist) by using the "*Create New Folder*" icon. Go into this directory. Change the extension of the file from *.txt* to *.RMA* (if you used RMA preprocessing) or *.GMA* (if you used GC-RMA preprocessing). Click "*Save*".
11. Back in the "Sample Inspector" popup, click the >> button in the lower right hand corner to go to the next sample.
12. Repeat steps 9-11 until you have finished all of the samples in your experiment.
13. When you have finished, repeat the entire process but use the other preprocessor that you didn't use the first round.