

ELISA-BASE User Guide

Amanda M. White, James R. Collett, Richard C. Zangar
Pacific Northwest National Laboratory
Richland, WA

<http://www.pnl.gov/statistics/ProMAT/ELISA-BASE.stm>

Table of Contents

Table of Contents	2
Introduction.....	2
Installation.....	2
Install BASE	3
Install R.....	3
Install ELISA-BASE Plugins.....	3
Configure ELISA-BASE Plugins	4
Install ELISA-BASE AnnotationTypes.....	5
Using ELISA-BASE	5
Antigen/Detection Mix Concentration Importer.....	7
Sample Importer.....	8
ELISA Experiment Importer.....	8
ProMAT Plugin.....	11
Example	15
Configure GPR File Format.....	15
Import reporters	15
Create an experiment and import the microarray data.....	16
Set up antigen mixture and detection antibody mixture	16
Run ProMAT	17
References.....	17

Introduction

ELISA-BASE is a database system for the management of ELISA microarray assays for biomarker screening and protein expression analysis. It includes features for microarray experiment tracking, materials and methods documentation, and data quality assurance. The analytical engine for ELISA-BASE is the Protein Microarray Analysis Tool (ProMAT), our statistical program for the prediction of protein concentrations in biological samples. ELISA-BASE is an extension of the BioArray Software Environment v.2 (BASE2), an open source Java application server for DNA microarray data management. Please see [1] for more background information about this tool.

BASE [2] is a MIAME-compliant open source database designed for organizing DNA microarray data. Although ELISA microarray experiments bear many similarities to DNA microarray experiments, there are crucial differences between them that make using BASE for ELISA data non-trivial. ELISA-BASE extends BASE using plugins and annotations to create a system for tracking ELISA microarray data. Because we have not changed the core of the BASE system, our tool can continue to be used with new versions of BASE as they are released. In addition, ELISA data and DNA data can be stored in a single BASE system if desired.

Installation

This section describes the steps to installing and setting up ELISA-BASE. ELISA-BASE has been developed and tested on systems running Red Hat Enterprise Linux 4, Java 1.6.0, BASE version 2.7.0 and Apache Tomcat server version 6.0.16.

Install BASE

Before installing the ELISA-BASE tools, you must first install BASE version 2.7.0 or higher. The latest version of BASE as well instructions for installation as may be found here: <http://base.thep.lu.se/wiki/DownloadPage>.

Install R

R is a statistical computing language that is used to implement some of the algorithms in ELISA-BASE. R is free and open-source, and is available for Linux, Windows and Mac. Download and install R from <http://www.r-project.org>. (See Download on the left side.) Version 2.5.1 or higher is required.

After installing R, two R libraries must be installed: MASS and mgcv. These can be installed by running the script `install_R_libraries.sh` that is distributed with the ELISA-BASE files. This should be run from the command line by root or by a user with sudo privileges:

```
[sudo] bash install_R_libraries.sh
```

This script will download and install the two libraries from the Comprehensive R Archive Network (<http://cran.r-project.org>) so an internet connection is required.

Install ELISA-BASE Plugins

The following files are needed to run the ELISA-BASE plugins:

1. `elisa-base.jar`
2. `ProMAT_Library.RData`
3. `Base_Import_Lib.RData`
4. `log4j-1.2.9.jar`

Place these files in a directory on your system that BASE can access. It is recommended that the files be placed in a subdirectory of the designated BASE plugins directory specified in `<base-home>/www/WEB-INF/classes/base.config` specified by the `plugins.dir` variable. We will refer to the location of the ELISA-BASE plugin files as `<elisa-base-home>` henceforth. Create a subdirectory under `<elisa-base-home>` that will serve as a working directory for the plugins to create temporary files. Make sure the permissions are set so that BASE will be able to read to and write from the working directory.

In your web browser, log into BASE as an administrator, then go to `Administrate` → `Plugins` → `Definitions` and click the `New Button`. If you placed the ELISA-BASE plugin files under the plugins directory in your `base.config`, then you can choose to install the plugins automatically and click the `Next` button. With automatic installation, BASE will search for all plugins under the given directory, and present you with a list of available plugins. Simply choose `Yes` from the dropdown list next to each plugin you want to install. With manual installation, you will have to enter the class name and jar file (`<elisa-base-home>/elisa-base.jar`) of each plugin to install, which is shown below.

Plugin	Class Name
ELISA Experiment Import	gov.pnl.elisabase.importers.experiment.ElisaExperimentImport
ProMAT Plugin	gov.pnl.elisabase.promat.PromatPlugin
ELISA-BASE AnnotationType Importer	gov.pnl.elisabase.importers.annotations.ElisaBaseAnnotationTypeImporter
Antigen/Detection Mix Concentration Importer	gov.pnl.elisabase.importers.samples.AntigenAndDetectionMixImporter
Sample Importer	gov.pnl.elisabase.importers.samples.SampleImporter

Configure ELISA-BASE Plugins

Two of the plugins require configurations before they can be used. Go to `Administrat`→`Plugins`→`Definitions` to see the list of loaded plugins. You may need to click the refresh button at the top of the page to see the plugins just loaded. Choose the `ELISA Experiment Import` plugin by clicking on its name. Then click the `New Configuration` button. Give the configuration a name, and click `Save` and `configure`. You will then be prompted to enter the following parameters:

- **Log file:** an optional log file where error information will be stored. Leave blank for no log file.
- **R library directory:** the directory where `BASE_Import_Lib.RData` is located (`<elisa-base-home>`).
- **Working directory for intermediate files:** a directory where all users have read/write access. This directory will be used as a directory for temporary files during processing.
- **Default Label:** default label to use for experiments (user can change this value when running the plugin).
- **Default formula for GenePix data:** default formula to use for creating a root bioassay set when using GenePix data. This can be changed by the user when running the plugin.
- **Default formula for ScanArray csv 1 channel data:** default formula to use for creating a root bioassay set when using ScanArray csv 1 channel data. This can be changed by the user when running the plugin.
- **Give read/write permissions to this group by default:** read/write access will be given to the chosen group(s) for all database items created using this configuration. The user running the plugin is always given full access over the items created during importation.

Click `Next` to finish and the configuration will be saved. Scroll down to select the new configuration you created under `Configurations`. Click `Share` and set the permissions so that all desired users can `Read` and `Use` this configuration.

Return to `Administrate`→`Plugins`→`Definitions` and choose `ProMAT Plugin` from the list. Click `New Configuration` and give your configuration a name and click `Save` and `continue`. This configuration has two parameters which you must specify:

- **Working directory for intermediate files:** a directory where all users have read/write access. This directory will be used for exporting data from BASE to be read the ProMAT statistical routines which are written in R.
- **ProMAT directory:** the directory where the `ProMAT_Library.RData` file is located (`<elisa-base-home>`). This file is a binary version of the R algorithms for data analysis.

Click `Next` to finish and the configuration will be saved. Scroll down to select the new configuration you created under `Configurations`. Click `Share` and set the permissions so that all desired users can `Read` and `Use` this configuration.

The remaining ELISA-BASE plugins have the option of creating a configuration so that a log file may be specified. Configurations are not required for these plugins, however. If desired follow the instructions above to create a configuration for another plugin.

Install ELISA-BASE AnnotationTypes

The final step is to create the `AnnotationType` objects that the ELISA-BASE plugins require. This may be done by running the `ELISA-BASE AnnotationType Importer` plugin. This plugin should be run once only by a BASE administrator. Go to `Administrate`→`Types`→`AnnotationTypes` and click the `Import` button. Choose `ELISA-BASE AnnotationType Importer` from the drop-down list, click `Next` and `Next` again. This plugin will create the following `AnnotationTypes` and will give `Read` and `Use` permissions to everyone. (The user who runs the plugin will still have full permissions over all `AnnotationTypes` created.)

<code>AnnotationType</code>	<code>Item</code>
<code>AntigenConcentration</code>	<code>Sample</code>
<code>DetectionConcentration</code>	<code>Sample</code>
<code>dilution</code>	<code>Extract</code>
<code>Laser</code>	<code>Scan</code>
<code>PMT</code>	<code>Scan</code>

The table above shows the `AnnotationType` names and the item to which the `AnnotationType` applies.

Using ELISA-BASE

The three key differences between ELISA and DNA microarray experimental data which ELISA-BASE addresses are the use of purified antigen mixtures for generating standard curves, the use of detection antibody mixtures, and the use of multiple chips per slide.

Since purified antigen and detection antibody mixtures are applied to chips in place of or in addition to biological samples, these items are stored in the `Samples` table in BASE. ELISA-BASE uses BASE's ability to pool samples to track these mixtures back to their constituent parts, so the individual antigens and detection antibodies are also stored in the `Samples` table.

Figure 1 shows the organization of biomaterials in ELISA-BASE. The antigen and detection antibody mixtures are diluted to form extracts and then pooled before being applied to the chip. In the case where a biological sample is used instead of an antigen mixture, the sample and diluted sample would take the place of the antigen mix and diluted antigen mix on the left side. The white boxes show metadata that is stored in annotations. The ELISA-BASE plugins populate those annotations based on data the user provides when running one of the plugins. The plugins that import biomaterials data and construct the structure shown in Figure 1 are the Antigen/Detection Mix Concentration Importer, Samples Importer and the ELISA Experiment Importer which are discussed in detail below.

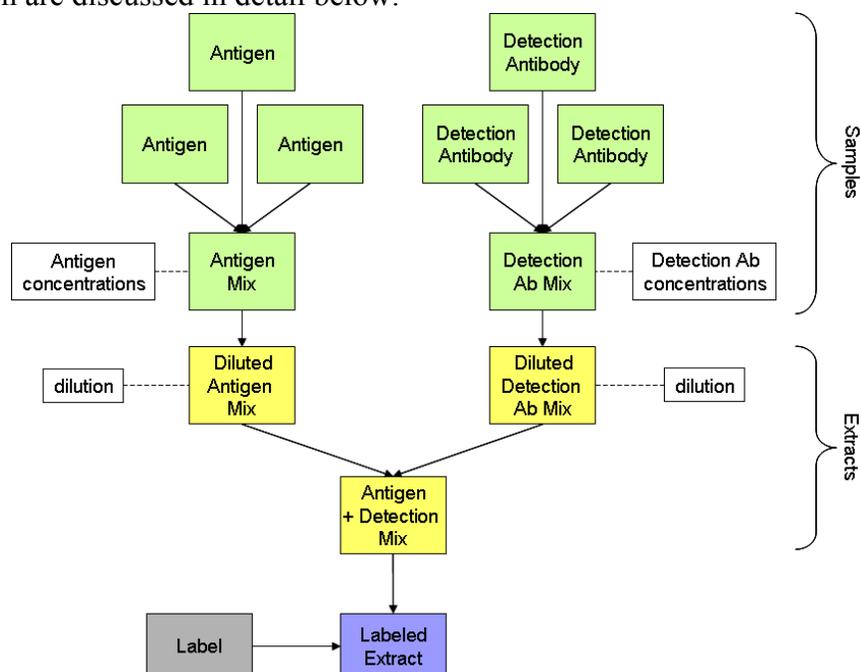


Figure 1: Antigen Standard and Detection Mixtures in ELISA-BASE. The figure shows how antigen standards and detection mixtures are organized in ELISA-BASE, with the arrows indicating parentage. The white items are annotations holding metadata. Reproduced from [1].

The other difference that ELISA-BASE addresses is the use of multiple chips per slide in ELISA microarray experiments. The underlying BASE data schema prescribes that one Raw Bioassay (one chip) is attached to one Scan object which is attached to one Hybridization¹ object, although a Hybridization may be linked to multiple Scans. This makes sense for DNA experiments where each slide is a single chip, but less so for multiple chips scanned together in an ELISA experiment. We chose to work within these restrictions and not change the data schema of BASE to ensure our tools will be able to be used with future versions of BASE, however this does create some minor redundancy in Scan data items in ELISA experiments.

In ELISA-BASE, each image analysis file is divided into chips, each of which becomes a Raw Bioassay. Each Raw Bioassay is linked to its own Scan which is linked to a Hybridization object.

¹ The BASE documentation and data schema refers to “hybridization” although for ELISA microarray experiments a better term would be “incubation.” For consistency, we use the term hybridization here for the equivalent step in an ELISA experiment since that is what BASE users will see when using the tool.

The Scan object has annotations for scanner settings and if a slide is scanned multiple times then multiple Scans will be linked to each Hybridization. Figure 2 shows the data organization. The ELISA Experiment Importer plugin is the tool that imports this data and organizes it into the structure shown. This plugin also creates a new Bioassay Set in the experiment with all the data imported.

Each of the ELISA-BASE plugins is discussed in the subsequent sections along with inputs, outputs and user parameters.

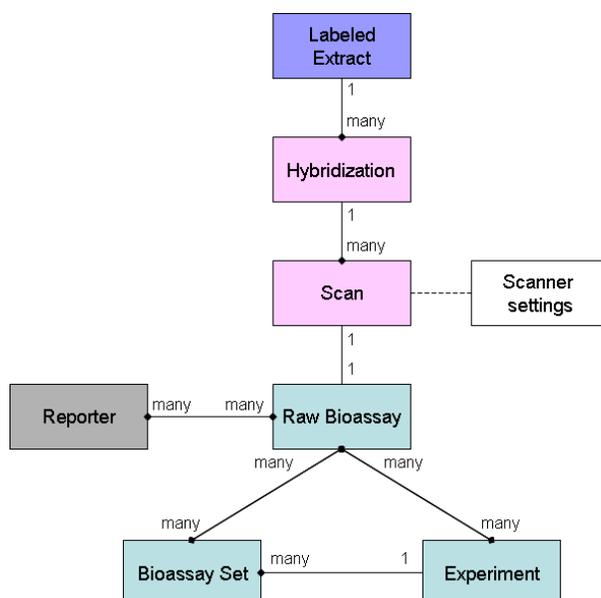


Figure 2: Organization of experimental data in BASE. The Labeled Extract box at the top connects to the biomaterials items as shown in Figure 1. Reproduced from [1].

Antigen/Detection Mix Concentration Importer

Antigen mixes and detection antibody mixes are mixtures of reagents that are applied to arrays. BASE provides the capability for easily tracking these mixtures and tracing them back to their original reagent batches, via pooled samples. Since antigens and detection antibodies are already stored as samples in BASE it is easy to create a pooled sample.

We also need to track the antigen or detection antibody concentrations, in order to create standard curves. To do this, we use two AnnotationsTypes for Samples: AntigenConcentration and DetectionConcentration. These annotations are used by the Antigen/Detection Mix Concentration Importer discussed below.

Note: Each antigen mix or detection mix must only be entered into BASE once. Thereafter, it can be used in as many experiments as necessary.

1. In BASE, go to View→Samples and find your antigen mix or detection mix sample (or click New to create a new sample). Click on the name to open that sample.
2. Click the Import button.

3. Choose the `Antigen/Detection Mix Concentration Importer` from the drop-down list then click `Next`.
4. Enter the plugin parameters:
 - **Type: Antigen or Detection Antibody Mix:** choose either `Antigen` or `Detection Antibody` depending on the type of mixture you are creating.
 - **File containing concentrations:** click `Browse...`, then click on the `Upload file...` button and upload a file from your computer that contains a list of antigen or detection IDs and concentrations. The file should be in CSV format where the first column is the antigen or detection antibody ID and the second is the concentration without any units, as shown below. The IDs must be the external IDs of antigens or detection antibodies in the `Samples` table. You may have column names in the first row but it is not necessary.
 - **Header:** choose `true` if the file has column names and `false` otherwise.
5. Click the `Next` then `Finish` buttons. A note will appear in the same window when the run is done telling you if the plugin finished successfully.

If the plugin was successful, the sample now shows that it was pooled from multiple samples, corresponding to the IDs provided. Also, it will have multiple values in either the `AntigenConcentration` or `DetectionConcentration` annotation, which have the form `[name]=[concentration]`.

The antigen or detection mix may now be used when importing experimental data. When creating the experiment info file for an experiment, use the external ID of an antigen mix where appropriate in the `sample.id` column and use the external ID of the appropriate detection antibody mix in the `detection` column.

Sample Importer

The Sample Importer plugin imports a list of samples from a text file. This plugin is provided as a convenient method to create many samples at once, and is used in the ELISA-BASE example below. The input to the Sample Importer is a tab-delimited text file with three columns: External ID, Name and Description, in that order. This file may be created in Excel by choosing `File→Save As` and choosing `Text (Tab delimited)` in the `Save as type` field. For each line of the file, the Sample Importer will create one sample with the specified name and external ID, and if a description is provided it will be applied to the sample created.

1. In BASE go to `View→Samples` and click `Import`.
2. Choose `Sample Importer` from the plugin drop-down list and click `Next`.
3. For the file parameter, provide the text file listing the samples.
4. Click `Next` and `Finish` to run the plugin.

An example input file may be found with the example files distributed with ELISA-BASE.

ELISA Experiment Importer

The ELISA Experiment Importer plugin imports a set of data files (the output from image analysis) and some meta-data about the experiment and creates all of the database objects and links between them to configure the entire experiment.

The ELISA Experiment Importer uses BASE's built in Raw Data Importer plugin. This plugin imports single file of microarray data based on a user-defined configuration. Thus, the user must create the correct configuration for their data type before using the ELISA Experiment Importer. This may be done by going to `Administratrate`→`Plugins`→`Plugin definitions` and choosing `Raw data importer` then clicking `New configuration...` The configuration tool will allow you to test your configuration against a data file to ensure it works.

The ELISA Experiment Importer also uses the `BatchDataImporter` plugin developed by Micha Bayer at the Scottish Crop Research Institute. The source was adapted for our purposes, and is included in `elisa-base.jar` in the `sbrn.base` package, with comments as to our changes. (The `BatchDataImporter` plugin does not need to be installed; the necessary code is included with ELISA-BASE.) The original `BatchDataImporter` plugin and source may be found here: <http://baseplugins.thep.lu.se/wiki/uk.ac.scri.batchimporter>.

Data Setup:

1. Combine the quantitative data output files from image analysis into a single `.zip` file. Files may be zipped using WinZip or WinRAR on Windows or via the `zip` command on the Windows or Linux command line.
2. Create a file that contains the following information about the experiment (which we will call the experiment info file).
 - This file must be in a comma-delimited (`*.csv`) format. This type of file may be created in Excel by choosing `File`→`Save As` and choosing `CSV` in the `Save as type` field.
 - The first row of the file should contain the column names, with the data in subsequent rows. The column names are shown below in bold and should be entered exactly as shown here.
 - Required columns:
 - **file.name**: the name of the data file (e.g. the appropriate file inside your zip archive)
 - **sample.id**: the external ID in BASE of the antigen mix used or the external ID of the sample used, whichever is appropriate
 - **detection**: the external ID in BASE of the detection mix used
 - **dilution**: the dilution factor of the sample or antigen mixture (e.g. 0.25 is a four-fold dilution)
 - positional columns to match to the data files (e.g. `Block` for GPR files or `Array Row` and `Array Column` for ScanArray files). These must appear exactly as they do in the data files.
 - Optional columns:
 - **Laser**: laser power of the scanner
 - **PMT**: PMT of the scanner
 - **Slide**: integer slide number
 - Laser and PMT can be extracted automatically by the import plugin from GPR and ScanArray files. Ensure that the appropriate setting is checked when running the ELISA Experiment Importer, which is described below. For other file types, the Laser and PMT should be entered in the experiment info file.

- See example experiment info file in Figure 3. It shows that the data in file “Slide 9.gpr” in Block 1 was treated with sample “sample 1” with a two-fold dilution, and was treated with detection mixture “DetMix_example.”

	A	B	C	D	E	F	G
1	file.name	Slide	Block	sample.id	dilution	Detection	
2	Slide 9.gpr	9	1	sample1	0.5	DetMix_example	
3	Slide 9.gpr	9	2	sample1	0.5	DetMix_example	
4	Slide 9.gpr	9	3	sample2	0.5	DetMix_example	
5	Slide 9.gpr	9	4	sample2	0.5	DetMix_example	
6	Slide 9.gpr	9	5	sample3	0.5	DetMix_example	
7	Slide 9.gpr	9	6	sample3	0.5	DetMix_example	
8	Slide 9.gpr	9	7	sample4	0.5	DetMix_example	
9	Slide 9.gpr	9	8	sample4	0.5	DetMix_example	
10	Slide 9.gpr	9	9	sample5	0.5	DetMix_example	
11	Slide 9.gpr	9	10	sample5	0.5	DetMix_example	
12	Slide 9.gpr	9	11	sample6	0.5	DetMix_example	
13	Slide 9.gpr	9	12	sample6	0.5	DetMix_example	
14	Slide 9.gpr	9	13	sample7	0.5	DetMix_example	
15	Slide 9.gpr	9	14	sample7	0.5	DetMix_example	
16	Slide 9.gpr	9	15	sample8	0.5	DetMix_example	

Figure 3: Example of an experiment info file.

Running the Plugin:

1. In BASE, choose your experiment by going to View→Experiments and clicking on the name of the desired experiment. (Or you can click New to create a new experiment.)
2. Click the Import... button and choose ELISA Experiment Importer from the plugin drop-down list. (Leave the File format field alone.)
3. Click Next to go to the next screen where the parameters are defined. Set the import parameters by clicking on each item in the list box on the left:
 - **Zip archive of data files:** Choose Browse... then upload the zip file of data files you created. Note that once you upload the archive, you must then click the button beside the name to select the archive for importation.
 - **Experiment info file:** Choose Browse... then upload the experiment info file. Again, be sure to select the file for importation after it appears in the list of files after uploading.
 - **Slide number column name:** Enter the header name in your experimental info file that defines the slide number column. The default value is "Slide". If you leave this value with the default setting, and your experimental info file uses another term, such as “slide number” the importer will issue an error and will not import your data.
 - **Use Laser and PMT values from data file headers:** Choose true to import the Laser and PMT values from your data file headers rather than looking in the experiment info file (default is true).
 - **Spot intensity value:** Choose a formula to define your spot intensity values.

- **Label:** What label was used in this experiment (i.e. cy3, cy5, etc)?
 - **Scan name prefix:** The scan items created during this import will start with this string (default is scan).
 - **Hybridization name prefix:** The hybridization items created during this import will start with this string (default is hybridization).
 - **Create samples:** Should new samples be created if the samples in the experiment info file do not already exist in the database? (Default is `true`.)
4. Once all the parameters have been specified, click `Next` then `Finish` to import the data.

When the importer is finished running, the experiment will include a Raw Bioassay for each chip (block) found in each data file in the zip archive. All of the Raw Bioassays will be grouped within a new root Bioassay Set that contains all the spots and the spot intensity value chosen in the parameters.

ProMAT Plugin

The Protein Microarray Analysis Tool (ProMAT) is a plugin for analyzing quantitative protein microarray data, which uses the R statistical programming language (<http://www.r-project.org>) to implement the computations. A description of the statistical methods implemented in this tool can be found in [3-4].

ProMAT takes as input microarray data (the output of microarray image analysis) where some chips are treated with standards of known antigen concentration and some chips are treated with samples of unknown concentration. ProMAT fits standard curves to the standards' data to relate spot fluorescence to concentration. ProMAT also calculates and outputs confidence bounds on these standard curves. The tool then uses the standard curves to predict antigen concentrations for the unknown samples, along with prediction intervals.

ProMAT can also incorporate multiple imager settings; if slides are scanned at multiple imager settings (e.g., laser power, PMT, etc) then a separate standard curve is created for each imager setting, and a separate concentration prediction is estimated for each spot at each imager setting. Then the protein concentration estimates from each spot are averaged by weighting according to the uncertainty of each estimate.

ProMAT produces diagnostic plots to aid the user in determining where there may be problems with the data. Figure 4 shows one such diagnostic plot. The lower right panel shows an example standard curve (in black) with its prediction intervals (in blue) plotted with the standard data. The gray region is the region of the curve where concentration predictions are the "best." The histogram in the lower left panel shows the sample spot values. The upper panel gives the coefficient of variation of the standard curve. The plot quickly shows the user how well the standard curve fits the data, how wide the prediction intervals are, and how well the range of the sample data matches up to the range of the standard curve. This type of plot is produced for each antigen.

ProMAT does not attempt to correct for variability in spot intensities between slides. We recognize that this is a problem that must be addressed to ensure accurate results, however no adequate solution has been found for low-spot-number slides of the type with which ProMAT

was designed to work. Thus the prediction uncertainty estimates are in fact under-estimates since they do not account for variability between slides.

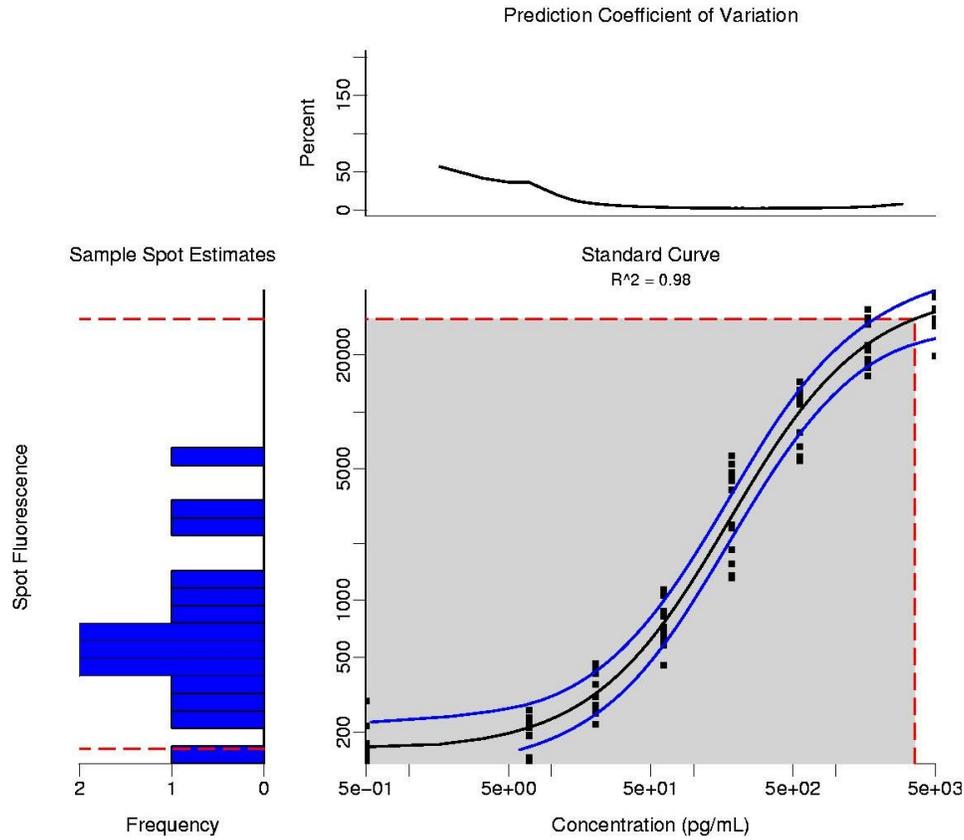


Figure 4: Example of prediction curve (black curve) with prediction bounds (blue). The histogram on the left shows the sample spot values for this antigen. The red bounds represent the region of best prediction on the curve; ideally sample spots should lie within these bounds. The curve at the top shows the percent coefficient of variation for the prediction curve.

Note: The ProMAT plugin will only work with bioassays comprised of ELISA microarray data that have been uploaded using the ELISA Experiment Importer or stored in the database in the same structure.

1. In BASE, select `View`→`Experiments` from the menu at the top of the page.
2. Click on an Experiment into which you have already imported ELISA microarray data using the ELISA Experiment Importer.
3. On the Experiment's page, click on the `Bioassay Sets` tab.
4. Click the name of the Bioassay Set that you wish to analyze with ProMAT. Only root bioassay sets that were created using the ELISA Experiment Importer (or their filtered child Bioassay Sets) may be analyzed with ProMAT.
5. On the chosen Bioassay Set's page, click the `Run Analysis` tab. A new dialog box will appear, as shown below. Select `ProMAT Plugin` in the `Plugin` drop-down box. Choose

the desired configuration if there is more than one listed in the Configuration to be used drop-down box. Click Next.

6. On the next screen, enter the ProMAT parameters.

Curve types

Choose one or more of the following standard curve types to fit. ProMAT will fit each type of curve you select to each antigen and then choose the curve that best fits the data.

- **Logistic curve/spline backup:** fit a four parameter logistic model. If and only if the logistic curve fit does not converge, it returns a spline curve.
- **Logistic curve:** fit a logistic curve
- **Power curve:** fit a power curve
- **Linear curve:** fit a linear curve
- **Spline curve:** fit a spline curve (this requires Monte Carlo bounds)

Methods for calculating bounds

Choose one method for calculating the upper and lower curve bounds.

- **Type of bounds:**
 - Analytic bounds: use propagation of error to calculate curve bounds and prediction intervals. For spline curves Monte Carlo bounds are always created unless “No bounds” is chosen. Analytic bounds are much quicker to calculate; however, due to the fact that a polynomial approximation of the curve is used to estimate the bounds, the intervals diverge near curve asymptotes (e.g., the upper and lower asymptote of a four-parameter logistic curve).
 - Monte Carlo bounds: use Monte Carlo simulation to estimate curve bounds and prediction intervals. Monte Carlo bounds give much more realistic intervals near curve asymptotes, but take more time to compute because many simulations are required for an accurate estimate. See [3] for a discussion of Monte Carlo versus propagation of error bounds.
 - No bounds: do not calculate curve bounds or prediction intervals.

Options

- **Standard curve data:** choose one or more antigen mixtures found in this experiment to create the standard curves.
- **Transform to apply to spot values:** choice of Identity or Natural Log. In most cases the Natural Log transform should be chosen, unless the data has been previously log transformed. (See [4] for a discussion of why the data is log transformed prior to standard curve estimation and prediction of concentrations.)
- **Transform to apply to concentration values:** see previous.
- **Create standard curves:** to simply plot the data without fitting standard curves, choose false (default is true). If false, ProMAT will simply plot the data without attempting to fit standard curves or make concentration predictions.
- **Identify sample outliers:** to identify and exclude outliers in the sample data, choose true (default is true). To identify outliers we calculate the spot replicate variance (one value for each curve) then for each set of replicate spots, values that are more than three standard deviations from the median value are determined to be outliers.

- **Treat standard data as sample data:** choose whether the standard data should be included with sample data for plotting and predicting concentrations.

Plotting options

- **x-axis label for plotting:** label that appears on the x (concentration) axis of the plots.
 - **y-axis label for plotting:** label that appears on the y (spot intensity) axis of the plots.
 - **Keep y-axis range constant in plots:** to keep the y-axis range constant in all plots for ease of comparison, choose true (default is false).
 - **Show standard data on plots:** choose true to show standard data points in the plots in black (default is true).
 - **Show replicate means on plots:** choose true to show means of replicate values on the plots in cyan (default is false).
 - **Show replicate standard deviations on plots:** show the one standard deviation above and below the replicate means on the plots in cyan (default is false).
 - **Show standard curves in plots:** show the fitted curves in the plots (default is true).
 - **Show upper and lower bounds in plots:** show the upper and lower confidence intervals in the plots (default is true).
7. Click `Next` and then `Finish` to begin ProMAT analysis of the Bioassay Set. The next dialog box that appears will apprise you of the progress being made while ProMAT runs.
 8. When ProMAT is finished, the dialog box will display the location a zip archive containing ProMAT output files, as shown below within the red ellipse. The zip archive may be found within the File management section of BASE. To reach the File items, select `View`→`Files` from the menu at the top of the page. From there, you will be able to download the zip archive of ProMAT results to your computer.

Figure 5: ProMAT successful completion screen showing the location of the results.

The zip file of ProMAT results contains four data tables: `predicted_concentrations.csv`, `standard_data.csv`, `sample_data.csv`, `standard_curve_statistics.csv` and `spot_intensity_variances.csv`.

- **predicted_concentrations.csv:** contains predicted concentrations and confidence bounds for each sample observation, where the within-chip replicates have been pooled prior to prediction. This file includes the antigen names, sample IDs and sample dilution plus the number of spots determined to be outliers which were excluded from the concentration prediction.
- **standard_data.csv:** contains the data used to fit the standard curves, and is provided for user reference. This file includes the antigen, sample, position on the slide, raw bioassay name, sample dilution, spot intensity and concentration.
- **sample_data.csv:** contains the sample data used for which protein concentrations are predicted, and is provided for user reference. This file includes the antigen, sample, position on the slide, raw bioassay name, sample dilution and spot intensity.

- **standard_curve_statistics.csv**: contains information about each standard curve including equation and goodness-of-fit statistics such as R^2 and mean-squared error. See the included file “Standard_Curve_Statistics.pdf” for a description of the data in this file.
- **spot_intensity_variances.csv**: contains estimates of the variance of spot replicates for both the standards and samples data for each curve. If you choose to log transform your data prior to curve fitting, then these variances are on the log scale.

The zip file also contains two types of diagnostic plots. The first plot type shows a standard curve with confidence bounds on curve fitting, and is named `standard_curve*` with one plot for each antigen. These plots give the user an idea of the variability in standard data and how well the estimated curves fit the data. The second type of diagnostic plot consists of the standard curves displayed with prediction intervals, which give the upper and lower confidence bounds on the predicted concentrations. These plots are named `prediction_curve*`. These plots also include a histogram of the sample data, which shows if the range of the standard curve encompasses to the sample data. Finally, the plot shows the percent coefficient of variation of the predictions, which is calculated from the distance between the concentration prediction and the prediction bounds divided by the prediction value. If the data is log transformed then this value is calculated on the log scale. An HTML file will also be created to allow the user to easily browse through the second type of diagnostic plots for each antigen.

Example

An example ELISA microarray dataset is distributed with ELISA-BASE to illustrate the use of the tools. The files for this example are found in the `example` director of the zipped archive of ELISA-BASE files. This section will describe how to import this data into the database and analyze the data using ProMAT.

It is recommended that you create a project for the example data to keep it separate from the rest of your data. To do this, go to `View`→`Projects` and click `New` to create a new project. After saving the project, be sure to click `Set active` to the right of the project name.

Configure GPR File Format

The example dataset consists of GPR files generated from ScanArray scanner software. Before the data can be imported, a configuration must be created so that BASE can parse the file. We have included a configuration file, `example.gpr.import.config.xml` with the example files. To create the configuration:

1. Go to `Administrative`→`Plugins`→`Configurations` and click `Import`.
2. Choose `Plugin configuration importer` from the `Plugin list` and click `Next`.
3. For `XML file`, provide the file `example.gpr.import.config.xml` distributed with the ELISA-BASE example.
4. Do not change the other parameters.
5. Click `Next` and `Finish` to create the new configuration.

Import reporters

First, importer the reporters from a GPR file:

1. Go to `View`→`Reporters` and click `Import`.
2. Choose `Reporter Importer` from the `Plugin` drop-down list.
3. Choose `Reporters` from `GenePix` file from `File format` list.
4. Click `Next`. For the file parameter, extract any GPR file from `ELISA-BASE_example_data.zip` and upload it.
5. Click `Next` and `Finish`. All the reporter names and IDs in the file will be imported.

Create an experiment and import the microarray data

1. To create an experiment, go to `View`→`Experiments` and click `New`.
2. Give the experiment a name and choose `GenePix` for the `Raw data type`. Click `Save`.
3. Click on the name of the new experiment. Click the `Import` button and choose the `ELISA Experiment Importer` plugin.
4. Plugin parameters:
 - a. Zip archive of data files: `ELISA-BASE_example_data.zip`
 - b. Experiment info file: `ELISA-BASE_example_experiment_info.csv`
 - c. Slide number column name: `Slide`
 - d. User Laser and PMT values from data file headers: `true`
 - e. Spot intensity value: `Median FG`
 - f. Label: choose any label
 - g. Scan name prefix: `anything`
 - h. Hybridization name prefix: `anything`
 - i. Create samples: `true`
5. Click `Next` and `Finish` to start the plugin. It will take a few minutes to complete. The plugin will create one `Raw Bioassay` for each chip on each slide in the dataset and will create the necessary biomaterials, scans and hybridizations to capture all the metadata.

Set up antigen mixture and detection antibody mixture

1. Go to `View`→`Samples` and click on `Import`.
2. Choose the `Samples Importer` plugin and click `Next`.
3. For file, provide the `samples_import.txt` file.
4. Click `Next` then `Finish` to run the plugin. This will create samples for each antigen and detection antibody.
5. When the plugin is finished, go to the `Samples list` (`View`→`Samples`) and click on `AgMix_example`.
6. Click the `Import` button and choose the `Antigen/Detection Mix Concentration Importer` plugin, and click `Next`.
7. Plugin parameters:
 - a. Type: `Antigen (default)`
 - b. File: `AgMix_concentrations.csv`
 - c. Separator: `,` (default)
 - d. Header: `true (default)`
6. Click `Next` and `Finish` to run the plugin. This will import the antigen concentrations and make `AgMix_example` a pooled sample of its constituent antibodies.
7. When the plugin is finished, return to the `Samples list` and click on `DetMix_example`.

8. Click the `Import` button and choose the `Antigen/Detection Mix Concentration Importer` plugin, and click `Next`.
9. Parameters:
 - a. Type: `Detection antibody`
 - b. File: `DetMix_concentrations.csv`
 - c. Separator: `,` (default)
 - d. Header: `true` (default)
10. Click `Next` and `Finish` to run the plugin. This will import the detection antibody concentrations and make `DetMix_example` a pooled sample of its constituent detection antibodies.

Run ProMAT

1. Go to `View`→`Experiments` and click on the experiment into which you imported the example data.
2. Click on the `Bioassay sets` tab at the top and click on the bioassay set created by the importer plugin, which is named `ELISA Bioassay Set: Median FG`.
3. Click `Run analysis` and choose the `ProMAT` plugin.
4. Parameters: it is suggested that you use the default parameters but you may adjust the options as desired. The parameters options are described in detail above in the `ProMAT Plugin` section.
5. Click `Next` and `Finish` to run the plugin. It will take several minutes to run.
6. When the plugin is finished, it will save a `.zip` file in your `Files/ProMAT Results` directory and will display a message showing you the file name. (The results file is named according to the experiment, the bioassay set and the timestamp when the file is uploaded.)
7. Go to `View`→`Files` and click the `ProMAT Results` directory to see the results file, which may be downloaded to your machine. When you download the file and unzip it you will see the results and figures as described above in the section describing the `ProMAT` plugin.

References

1. White, A.M., Collett, J.R., Seurnyck-Servoss, S.L., Daly, D.S., Zangar, R.C. “ELISA-BASE: An Integrated Bioinformatics Tool for Analyzing and Tracking ELISA Microarray Data.” Submitted to *Bioinformatics*.
2. Saal, L. H., Troein, C., Vallon-Christersson, J., Gruvberger, S., Borg, A., and Peterson, C. (2002) BioArray Software Environment (BASE): a platform for comprehensive management and analysis of microarray data, *Genome Biol* 3, SOFTWARE0003.
3. Daly D.S., Anderson, K.K., White, A.M., Gonzalez, R.M., Varnum, S.M., and Zangar, R.C. “Predicting protein concentrations with ELISA microarray assays, monotonic splines and Monte Carlo simulation.” *Statistical Applications in Genetics and Molecular Biology* 2008, 7(1): Article 21. <http://www.bepress.com/sagmb/vol7/iss1/art21>
4. Daly D.S., White, A.M., Varnum, S.M., Anderson, K.K., and Zangar, R.C. “Evaluating concentration estimation errors in ELISA microarray experiments.” *BMC Bioinformatics* 2005, 6:17. <http://www.biomedcentral.com/1471-2105/6/17>