

# Preparing a Minimum Information about a Flow Cytometry Experiment (MIFlowCyt) Compliant Manuscript Using the International Society for Advancement of Cytometry (ISAC) FCS File Repository (FlowRepository.org)

Josef Spidlen,<sup>1</sup> Karin Breuer,<sup>2</sup> and Ryan Brinkman<sup>1</sup>

<sup>1</sup>BC Cancer Agency, Terry Fox Laboratory, Vancouver, British Columbia, Canada

<sup>2</sup>Department of Molecular Biology and Biochemistry, Simon Fraser University, Burnaby, British Columbia, Canada

## ABSTRACT

FlowRepository.org is a Web-based flow cytometry data repository provided by the International Society for Advancement of Cytometry (ISAC). It supports storage, annotation, analysis, and sharing of flow cytometry datasets. A fundamental tenet of scientific research is that published results should be open to independent validation and refutation. With FlowRepository, researchers can annotate their datasets in compliance with the Minimum Information about a Flow Cytometry Experiment (MIFlowCyt) standard, thus greatly facilitating third-party interpretation of their data. In this unit, we will mainly focus on the deposition, sharing, and annotation of flow cytometry data. *Curr. Protoc. Cytom.* 61:10.18.1-10.18.26. © 2012 by John Wiley & Sons, Inc.

Keywords: data repository • data sharing • minimum information • annotation • flow cytometry

## INTRODUCTION

In the last decade, advances in flow cytometry instrumentation and reagent technologies have enabled simultaneous single-cell measurement of surface and intracellular markers, including cellular activation markers, intracellular cytokines, immunological signaling, and cytoplasmic and nuclear cell cycle and transcription factors, thus positioning cytometry to play an even larger role in health care and medical research. However, the rapid expansion of flow cytometry applications has outpaced the development of tools for storage, analysis, and data representation (Bocsi and Tárnok, 2008; Bashashati and Brinkman, 2009; Lugli et al., 2010). In addition, datasets are usually not stored with an explicit link to experimental metadata, such as data analysis procedures, experimental conditions used, or information about the samples processed. This link between data and metadata is crucial, as it facilitates better understanding of analysis approaches and reproducible research. When datasets are linked to figures and summaries through a detailed explanation of the sample processing and analysis pipeline, other scientists can increase the amount of detailed analysis that can be conducted on already collected data, re-examine and expand on biological results, and better reproduce and understand published findings.

We developed FlowRepository.org, a public flow cytometry data repository, to address this shortcoming by facilitating the integrated review of datasets referenced in published literature with its associated metadata. Through FlowRepository, users can access, review,

download, deposit, annotate, share, and analyze flow cytometry datasets. Since proper annotation is key to third-party interpretation, FlowRepository supports all annotation concepts specified by the Minimum Information about a Flow Cytometry Experiment (MIFlowCyt, Lee et al., 2008).

MIFlowCyt establishes criteria to describe flow cytometry experiments that provide enough information to allow for correct interpretation and review of experimental details including samples, analysis, and results. It enumerates all the items that are required to be reported, and the depth of detail. Therefore, it minimizes problems such as alternative interpretations resulting from nonuniform and inconsistent reporting, which are critical issues for public data repositories (Burgoon, 2006). Since MIFlowCyt does not specify the format or the location of the required information, data submitters are not forced to provide all of their annotations on line. Information that exists in a manuscript does not necessarily need to be repeated in FlowRepository; however, the more information that is integrated with the dataset on line, the easier it is to locate and interpret the data and experimental findings by independent researchers.

The protocols in this unit explain the use of FlowRepository to explore, review, and retrieve public datasets (Basic Protocol 1), and to deposit and annotate data in compliance with the MIFlowCyt standard (Basic Protocol 2). Since FlowRepository is built on top of the Cytobank platform, the majority of Cytobank's functionality is also available in FlowRepository. Exceptions include the ability to clone experiments, publish duplicated datasets, or keep experiments private for an unlimited time. Despite these minor restrictions, additional information including protocols to analyze data on line within FlowRepository and creating illustrations and reports can be easily adapted from Cytobank's guidelines (*UNIT 10.17*).

## **BASIC PROTOCOL 1**

### **FlowRepository PUBLIC DATA RETRIEVAL**

Basic Protocol 1 explains how to access FlowRepository and to locate and review public datasets and download data, reports, and images from these datasets. Datasets in FlowRepository are organized using the notion of experiments; an experiment is a set of data files accompanied by annotations, which may include details about data analysis and its results. In this unit, the terms “experiment” and “dataset” are used as synonyms.

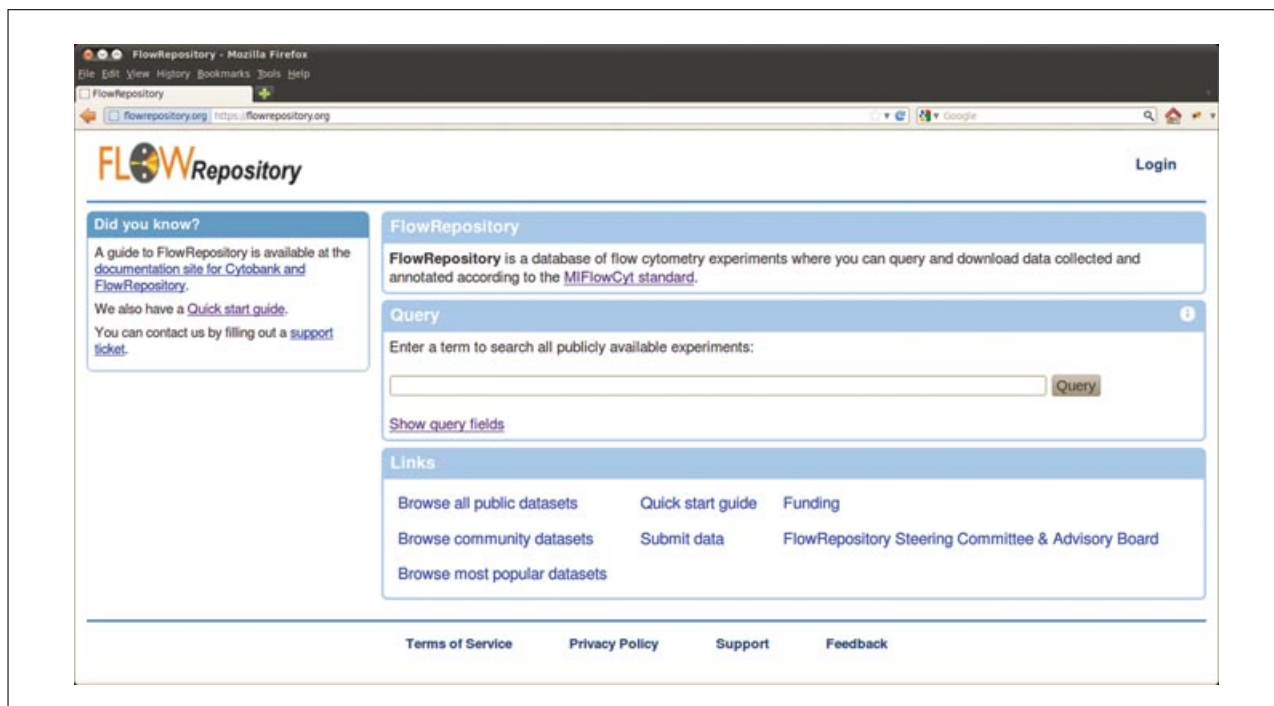
#### **Materials**

A computer with an Internet connection and a supported Web browser: Firefox (<http://www.mozilla.org>) is recommended; some other major Web browsers, such as Safari, Internet Explorer, or Chrome can also be used; however, their use may result in suboptimal visual appearance or limited functionality of the application

Ability to run Java applets in your Web browser: the Java version should be version 1.5.1 or higher (Java version 1.6 or higher is recommended, preferably provided by Sun Microsystems Inc.); navigate your browser to

<http://javatester.org/version.html> to verify your ability to run Java and your Java version. If needed, Mac users should use the Apple software update feature to install Java; Windows users can get the latest Java from a Java download page (<http://www.java.com>), and Linux users should preferably obtain Java and the Firefox Java plugin directly from their distribution.

A Google ID, Yahoo! ID, or another OpenID, optional: the ID is not needed for basic anonymous read-only access to FlowRepository, but it is required in order to deposit and annotate datasets, analyze data online, and perform additional advanced tasks. You may have an OpenID already if you have an account with certain Web-based service providers, such as Google, Yahoo!, Flickr, AOL, Orange, Blogger, WordPress, or MySpace. Visit <http://openid.net/get-an-openid/> for further details.



**Figure 10.18.1** FlowRepository—landing page. The landing page of FlowRepository allows users to quickly browse or query public datasets.

### ***Access FlowRepository and open a public view of a public dataset***

1. Navigate your Web browser to <https://flowrepository.org>.

*Alternatively, you could also use the <http://flowrepository.org> Web site, since FlowRepository is available using Hypertext Transfer Protocol (HTTP) directly, as well as its secure version—HTTPS—which provides encrypted communication between your computer and the FlowRepository server. Using HTTPS is typically preferable.*

2. Take a moment to review the FlowRepository landing page (Fig. 10.18.1).

*The Web site is horizontally divided into three sections: the FlowRepository logo and main menu in the header at the top, the main section in the middle, and a few static links (i.e., Terms of Service, Privacy Policy, Support, and Feedback) in the footer at the bottom. If you are accessing FlowRepository anonymously, then the main menu contains the Login link only; other items will be displayed if you are logged in. The main section is vertically divided, with some quick links, references, and contextual menu items on your left and the main contents stretching from the center to the right. The main content is commonly organized into boxes; these are typically blue to display requested information or to point to available actions. On the landing page, there is a box with basic information about FlowRepository, another box to query the repository, and finally a box with various useful links, such as for browsing public datasets or submitting data. On other pages, colors, such as green or red, are used for additional tips (green) and notifications about a positive (green) or negative (red) result of an attempted action. An orange color is also used to either indicate a warning or to provide a context-aware link backwards. Some boxes have an “i” in a circle at the top right corner of their title bar. Mousing over this “i” displays an additional informative floating box. Boxes having an “x” in the top right corner of their title bar can be closed by clicking on the “x” symbol. Finally, boxes having a small triangular arrow in the top left corner of their title bar may be collapsed (if the arrow is pointing down) or restored from a collapsed state (if the arrow is pointing right).*

3. Search for the experiment FR-FCM-ZZZ3 (Identification of B cells through negative gating; Blimkie et al., 2010). You can search by typing in the Query box and pressing the Query button. For example, you can type Identification of B cells or FR-FCM-ZZZ3 to locate this experiment.

1 matching experiment found.

Search:

Experiment Name	Repository ID	Pubmed ID(s)	FCS Files	Illustrations	Primary Researcher	Project	MIFlowCyt Score	Updated
Identification of B cells through negative gating	<a href="#">FR-FCM-ZZZ3</a>	<a href="#">[20131398]</a>	284	3	Karin Breuer		<div><div></div></div>	04:45 AM

Showing 1 to 1 of 1 entries

**Figure 10.18.2** Result of a query. Repository queries return a list of matching datasets.

*You may also click on the Show query fields link to show and select which query fields should be used for your search (e.g., experiment names, repository identifiers, keywords, researcher names, sample annotations, PubMed IDs, etc.).*

- Click on the Query button, which will direct you to a page with a table of experiments matching your search criteria (Fig. 10.18.2). Review the results of your query.

*The result table may list a single experiment if your query was specific enough. Alternatively, several experiments may be shown. You may filter these further by typing in the Search field. In addition, you may sort the table by clicking on the up or down arrow in any of the table headings (i.e., the experiment name, the repository identifier, the PubMed identifiers of related publications, the number of FCS files, the number of illustrations, the name of the primary researcher, the name of the project, MIFlowCyt score, and the date of the last update of the experiment).*

- Click on FR-FCM-ZZZ3 to open the public view of this experiment.

*FR-FCM-ZZZ3 is the repository identifier of this experiment. Each repository identifier is in the form of FR-FCM-xxxx, where xxxx is a sequence of four alphanumeric characters (case sensitive). Knowing the repository identifier, a public view of an experiment can be accessed directly by a URL in the form of <https://flowrepository.org/id/FR-FCM-xxxx>, e.g., <https://flowrepository.org/id/FR-FCM-ZZZ3>.*

- Take a moment to review the annotation of this experiment.

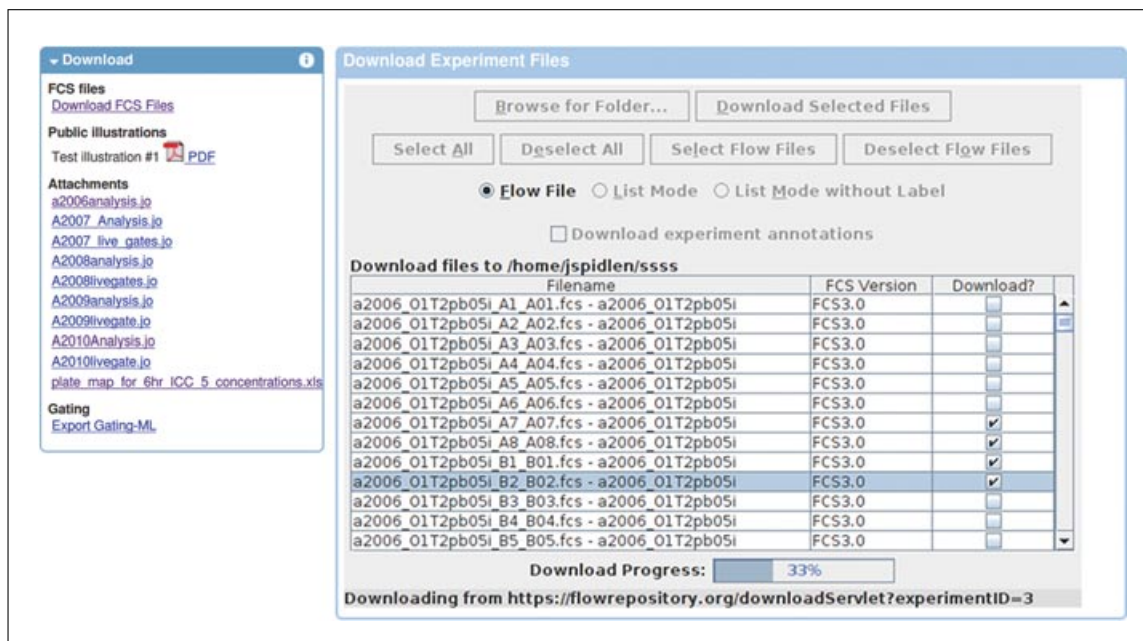
*The public experiment header at the top displays the experiment name, repository identifier, primary researcher, and its MIFlowCyt score—a value between 0% and 100% that approximates the level of compliance with the MIFlowCyt standard (Lee et al., 2008). The public view of this experiment is shown in four boxes corresponding to the four major sections of MIFlowCyt: experiment overview, flow sample/specimen details, flow cytometer details, and data analysis details. You can collapse and restore these boxes by clicking on the arrow in the left top corner. You can also select a single FCS file in the Flow Sample/Specimen Details box and display relevant sample details by clicking on the Display button. Finally, you may review instrumentation details for a specific FCS file by clicking on the details link in Flow Cytometer Details.*

### Download data from a public dataset

- Follow the Download FCS Files link in the Download box on the left (Fig. 10.18.3). This will start a Java applet on your Web page. Depending on your browser security settings, you may need to explicitly confirm that you allow this applet to run.

*If you do not see anything in the Download Experiment Files box, then your browser did not execute the applet. Please verify that your browser is able to run Java applets (see Necessary Resources).*

*Follow the instructions in the green information box to download the FCS files. If you are only trying to get familiar with the repository, you may not need to download all 284 FCS files (2.7 GB); possibly select just a few.*



**Figure 10.18.3** Dataset download. The download box includes links for downloading FCS files, illustrations, attachment, and gating. Downloading FCS files starts a Java applet where users select the download location (i.e., a folder locally on their computer) and which data files to download.

- Download individual attachments (FlowJo workspaces, MS Excel spreadsheet) by clicking on the appropriate links and saving the files on your file system.

*In addition, you may export Gating-ML (Spidlen et al., 2008) files, download PDF files with illustrations, or download the MIFlowCyt report. Note that rendering of the PDFs may take a few minutes due to the size of the actual report.*

### **Explore advanced options for reviewing a public dataset**

- Log in to FlowRepository by following the Login link in the top right corner of your screen; you will need an OpenID in order to log in (see Necessary Resources). If it is your first time logging in, then you will need to register for a new account with FlowRepository by clicking on the registration link in the New Users box.

*For registration, fill out your name and e-mail address and select your OpenID provider. Optionally, you may specify your location and organization. Finally, you will have to agree to the FlowRepository Terms of Service in order to proceed with the registration by clicking on the Register button.*

- Locate and select the FR-FCM-ZZZ3 (Identification of B cells through negative gating) experiment in your inbox.

*You can filter the experiments by starting to type in the Search field. You can also sort the Inbox table by any of the columns by clicking on the up or down arrow with any of the column headings.*

- Take a moment to get familiar with the advanced experiment view.

*This is your personal view of this experiment, which is similar to the view displayed when you create your own experiments. However, in this case you do not have write permissions for this experiment; therefore, no editing actions are offered.*

*You may click on FR-FCM-ZZZ3 in the experiment header to quickly switch to the experiment public view discussed earlier. Since you are now logged in, you may click on "Identification of B cells through negative gating" in the header, to return to your personal view of this experiment.*



FCS Files (284)						
Download Files Upload More Files De-identify FCS Files Review Keywords in FCS files						
File Name	Sample	Tube Name	Experiment Variables	Panel	Events	Size
<a href="#">a2006_O1T2pb05i_A1_A01.fcs details</a>	<a href="#">show sample description</a>	A1	a2006, A, 1, a2006 O1T2pb05i	Panel 1	109046	9.2 MB
<a href="#">a2006_O1T2pb05i_A2_A02.fcs details</a>	<a href="#">show sample description</a>	A2	a2006, A, 2, a2006 O1T2pb05i	Panel 1	150000	12.6 MB
<a href="#">a2006_O1T2pb05i_A3_A03.fcs details</a>	<a href="#">show sample description</a>	A3	a2006, A, 3, a2006 O1T2pb05i	Panel 1	150000	12.6 MB
<a href="#">a2006_O1T2pb05i_A4_A04.fcs details</a>	<a href="#">show sample description</a>	A4	a2006, A, 4, a2006 O1T2pb05i	Panel 1	150000	12.6 MB
<a href="#">a2006_O1T2pb05i_A5_A05.fcs details</a>	<a href="#">show sample description</a>	A5	a2006, A, 5, a2006 O1T2pb05i	Panel 1	150000	12.6 MB
<a href="#">a2006_O1T2pb05i_A6_A06.fcs details</a>	<a href="#">show sample description</a>	A6	a2006, A, 6, a2006 O1T2pb05i	Panel 1	150000	12.6 MB
<a href="#">a2006_O1T2pb05i_A7_A07.fcs details</a>	<a href="#">show sample description</a>	A7	a2006, A, 7, a2006 O1T2pb05i	Panel 1	150000	12.6 MB
<a href="#">a2006_O1T2pb05i_A8_A08.fcs details</a>	<a href="#">show sample description</a>	A8	a2006, A, 8, a2006 O1T2pb05i	Panel 1	150000	12.6 MB
<a href="#">a2006_O1T2pb05i_B1_B01.fcs details</a>	<a href="#">show sample description</a>	B1	a2006, B, 1, a2006 O1T2pb05i	Panel 1	150000	12.6 MB
<a href="#">a2006_O1T2pb05i_B2_B02.fcs details</a>	<a href="#">show sample description</a>	B2	a2006, B, 2, a2006 O1T2pb05i	Panel 1	150000	12.6 MB

**Figure 10.18.4** List of experiment FCS files. The advanced experiment view lists FCS files including details, such as staining panel assignment, experiment variables, tube identification, number of events, and size of the data file. Additional details may be displayed by clicking on the details link near each of the FCS file name.

- Review the details of the `a2006_O1T2pb05i_A1_A01.fcs` FCS file by clicking on the appropriate link in the FCS Files box (Fig. 10.18.4).

*The first box shows general information about the FCS file. The second box displays details about instrumentation, the third about laser configuration, and the fourth information about channels. The last box shows a raw FCS file header (Spidlen et al., 2010). An alternative method to show FCS keywords and their values in a more readable way is by clicking on the Review Keywords in FCS files link in the Actions box when on the experiment summary page.*

- Go back to experiment summary and review the sample description for the `a2006_O1T2pb05i_A1_A01.fcs` FCS file by clicking on the appropriate “show sample description” link.

*A pop-up window will display the description of the sample, sample source, sample characteristics, and treatment and staining details.*

### Review MIFlowCyt compliance of a public dataset

- Review the MIFlowCyt score details by clicking on the MIFlowCyt Score bar on the right side of the experiment header.

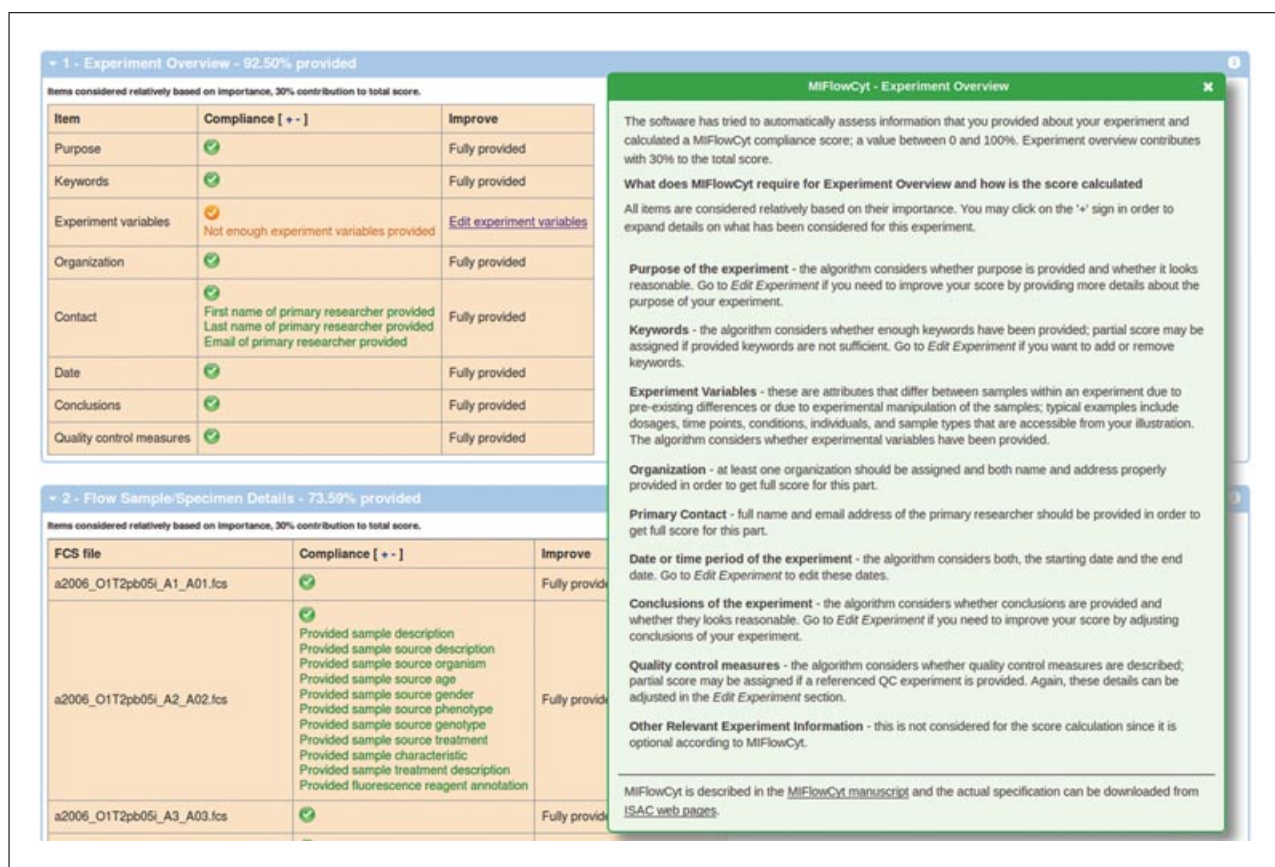
*Four collapsible boxes will show MIFlowCyt compliance details of this experiment annotation. There are green, orange, and red check marks indicating provided, partially provided, and missing information. You may expand additional details about any particular item by clicking on the check mark symbol. Alternatively, you may expand or close all details in a particular section by clicking on the + or – symbols near the Compliance caption in the table header.*

- Review details of how various annotation items contribute to the MIFlowCyt score by moving your mouse over the “i” symbol on the right side of the caption bar of each of the four boxes.

*An information pop-up box will appear containing information specific for the selected MIFlowCyt section (Fig. 10.18.5). The box may be closed by clicking on the “x” symbol in its caption bar.*

- Preview a print view of the MIFlowCyt score and compliance report by clicking on the Print View link in the MIFlowCyt box on the left.

*You may also download this report in a PDF form by clicking on the nearby PDF link.*



**Figure 10.18.5** MIFlowCyt compliance details. Provided and missing information are summarized for each section of MIFlowCyt; additional details may be expanded by clicking on the check mark symbols. Mousing over the “i” in the top right corner displays more information about MIFlowCyt score calculation.

## FlowRepository DATA UPLOAD AND ANNOTATION

Basic Protocol 2 demonstrates how to upload datasets to FlowRepository and how to annotate these to provide all the information as required by the MIFlowCyt standard.

### Materials

Necessary resources are the same as for Basic Protocol 1, with the exception of OpenID, which is required for Basic Protocol 2, whereas it is optional for most of Basic Protocol 1

### Upload data to FlowRepository

1. Log in to FlowRepository by navigating your Web browser to <https://flowrepository.org> and following the Login link.

*Additional details are discussed in step 9 of Basic Protocol 1.*

2. Start a new data submission by following the Create a new experiment link in your inbox view (Fig. 10.18.6).

*Alternatively, if you are in the Public view (landing page), then you may start a new data submission by following the Submit data link. The system will ask you to log in if you are not currently logged.*

*Information about your available credits is shown in the top left corner of your Inbox view. Credits are used to control the number of unpublished datasets that a user is allowed to upload to FlowRepository.org, as FlowRepository.org is not intended to be a personal archival system. Typically, you will have 50 credits to begin with, and you may also request*

## BASIC PROTOCOL 2


### Data Processing and Analysis

## 10.18.7

**Bank New Experiment**

\* Experiment Name

Project

\* Primary Researcher  Josef Spiden

\* PI Manager  [invite a new user](#)

Allow PI Manager to have full access to experiment ☒

\* Starting Date  (yyyy-mm-dd)

End Date  (yyyy-mm-dd)

(optional)

\* Purpose

Conclusion (optional)

Comments (optional)

Quality Control Measures (optional)

Quality Control Experiment (optional)

Keywords (optional)

Organizations (optional)   
 BC Cancer Agency, Terry Fox Lab  
 Child & Family Research Institute, Department of Pediatrics  
 University of Washington Medical Center, Department of Immunology  
 LSUHSC, MEDICINE

[Add new organization](#)

Pubmed IDs (optional)

\* required field

[Create Experiment](#)

**Figure 10.18.6** New experiment form. Basic experimental metadata should be provided for each new dataset as required by MIFlowCyt.

*additional credits if needed. You will need one credit to create a new experiment. This credit will be returned to your account once your experiment becomes public. Note that experiments in FlowRepository can only be kept private for up to 1 year after the first creation of the experiment or until an associated manuscript referencing the experiment data is published.*

3. Fill in the experiment name in the Bank New Experiment form.

*Try to stay relatively concise while being as descriptive as possible. Choosing a good title will help others locate your experiments in the future. If you intend to publish a manuscript related to this dataset, using an experiment name that is the same or similar to the intended manuscript title is usually a good choice.*

4. Keep Project as None.

*Projects allow you to group experiments; if you expect having many experiments in FlowRepository and you wish to use these additional means of organization, you may create a project by following the Manage My Projects link in your Inbox view. Your experiments can be assigned to your projects once projects are created.*



5. Fill out the Primary Researcher and PI/Manager fields.

*The Primary Researcher is usually the person who performed the experiment and is also typically submitting the data. The PI/Manager may be the same person or it may be a person responsible for a related higher-level project (e.g., a senior scientist, professor, or group leader). Both of them need to have an account with FlowRepository; you may invite new users to FlowRepository at this point.*

6. Fill out the date (range) of the experiment by entering the starting and end date fields.

*Use the date-picker component to provide the dates. The end date of the experiment is optional, to account for cases where an experiment has not been finished yet; however, the MIFlowCyt score will be negatively affected if the end date is not provided.*

7. State the purpose, conclusions, and additional comments about the experiment.

*While only a brief description of the purpose of the experiment is technically sufficient to upload a dataset, it is recommended (and required by MIFlowCyt) that you provide additional details. More information will help you reach a higher MIFlowCyt compliance score.*

8. Describe your quality-control measures and/or reference a quality-control experiment.

*As per MIFlowCyt (section 1.8), a description of the quality-control measures used, such as replicates, calibrations, or control assays, should be provided. If another experiment was performed for the purpose of establishing quality-control standards, then that experiment may be referenced.*

9. Provide a few keywords characterizing the experiment.

*Auto-complete, based on keywords describing existing experiments, is offered.*

10. Specify the organization that performed the experiment. You may need to use the “Add new organization” button if the appropriate organization does not exist in the FlowRepository database.

*If you are entering a new organization, please provide the name and department (if applicable). Continue by entering your postal code (zip). Once you leave this field, the system will try to automatically resolve the city and state/province (if applicable) based on the postal code, assuming that you are specifying an address in the same country from which you are currently connected. If this is not the case, you may specify the country field first, and then the postal code. Postal code lookup is currently supported for approximately 60 countries. Finish by completing and adjusting your address as appropriate.*

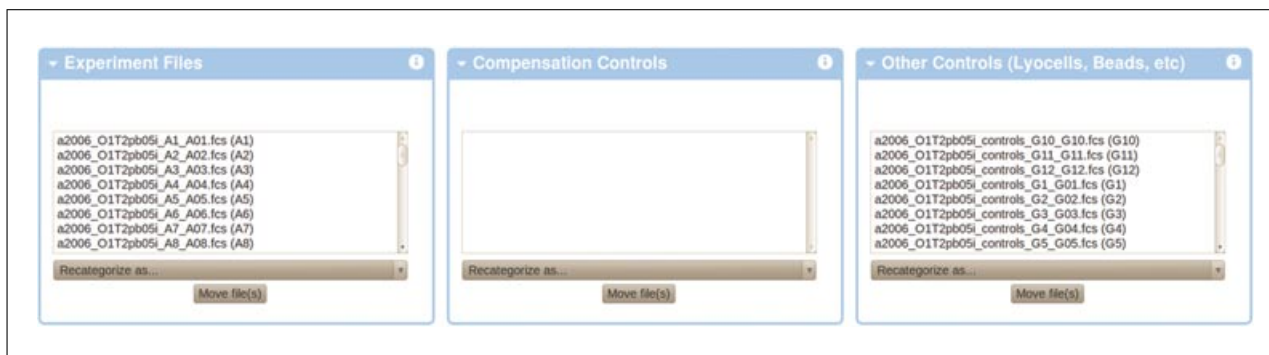
11. If there are one or more publications associated with this experiment, then specify their PubMed IDs. You may provide a single numeric PubMed ID or a comma-separated list of multiple PubMed IDs.

*Related publications may also be linked to the experiment later on, which is commonly the case since data are typically shared (at least with the reviewers) prior to the publication of the related manuscript.*

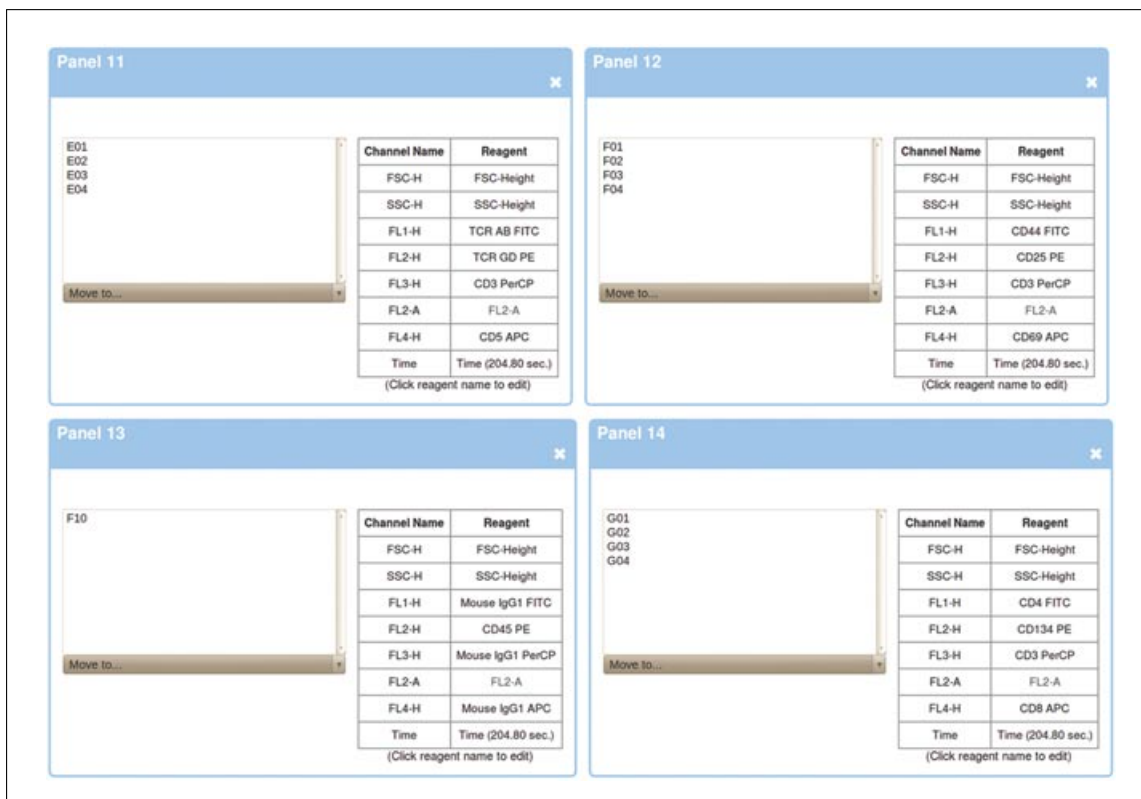
12. Finalize the submission of general metadata by pressing the Create Experiment button. This will direct you to a data upload page.

13. Organize a folder containing all the files to be uploaded on your personal computer.

*Note that the file/folder structure from your local file system will not be replicated on the FlowRepository server, where files are organized based on detected characteristics and additional annotations (e.g., staining panels, experimental variables, etc.). Consequently, you cannot upload multiple files with an identical file name even if these reside in different folders on your local file system. Please make sure that your files have unique names, or rename them in order to avoid an automated renaming mechanism that would be applied on duplicated file names within a single experiment on the FlowRepository server.*



**Figure 10.18.7** FCS file categorization. FlowRepository will automatically assign FCS files into experimental files, compensation controls, and other controls. Users may adjust the categorization if needed.



**Figure 10.18.8** FCS panel assignment. FlowRepository will assign FCS files to staining panels based on distinct groups of markers detected in the data. Users may adjust these assignments and rename, add, or delete panels if needed.

- Follow the instructions on your screen and select data (FCS files) as well as any additional files (figures, MS Word documents, MS Excel spreadsheets, FlowJo workspaces, FCS Express project files, FACS Diva files, etc.) that you wish to upload. Start by pressing the Browse for Folder button to choose the appropriate folder on your local computer. Make sure the “Upload?” check box is selected for all the files you wish to upload, and then click on the Upload Selected Files button.

*Upload will begin, and you should see the progress bar start to fill toward 100%. You will be directed to an upload status page once the upload is finished.*

- Verify that all of your files were uploaded correctly. FCS files should be automatically categorized into experimental files, compensation controls, and other controls (Fig. 10.18.7). In addition, they should be automatically assigned to staining panels based on markers detected in the FCS header (Fig. 10.18.8). Please verify

correct assignment or re-categorize files using the Channel annotation page link if needed.

*In addition, you should check that all files are correctly readable. The File Inventory box should list all experimental data files with a small scatter-plot thumbnail preview for each of them. The creation of these thumbnails may take some time, especially for multiple or large data files.*

*FlowRepository will refuse the submission of data files that exist in the repository already (e.g., uploaded by any other user). The detection of duplicate files is based on MD5 checksums, and therefore files are compared irrespectively of their names. You may contact the administrators using the Support link in the footer if you feel that the refusal of your data files is not justified.*

16. Review channel naming by following the experiment summary page, and subsequently all different staining panels as listed in the FCS files box.

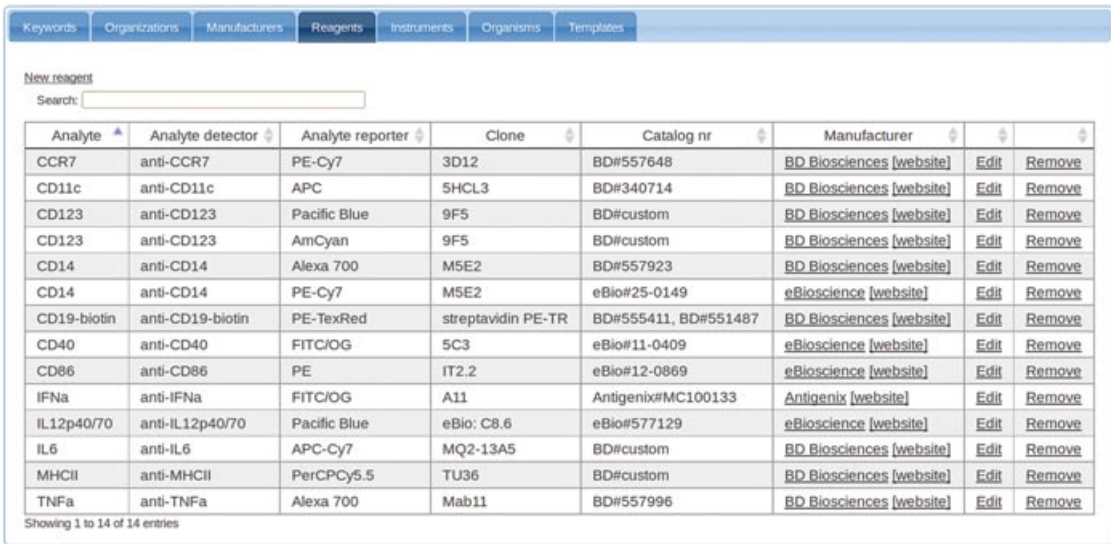
*Additional details, including guidelines to rename channel descriptions, rename, add or delete panels, change panel assignments, etc., can be found in UNIT 10.17, Basic Protocol 1, steps 15 to 18.*

### **Prepare annotation items and create templates**

17. Follow the Annotation Data link in the header of the page.

*Some annotations may be reusable among multiple experiments: FCS files, sample sources, etc. Therefore, FlowRepository allows you to create a super-set of various annotation concepts that may be reused to describe your experiment. It is therefore recommended that you follow the Annotation Data link prior to annotating your FCS files and samples. Specifically, you should create templates for samples and sample sources, import organisms as necessary, and specify reagents. Providing appropriate details in the Annotation Data section will significantly simplify annotation of your FCS files later. You may also want to review the information in the Getting Started with MIFlowCyt Annotations box.*

18. Select the Reagents tab (Fig. 10.18.9).



The screenshot shows the 'Reagents' tab in the FlowRepository interface. At the top, there are navigation tabs: Keywords, Organizations, Manufacturers, Reagents (selected), Instruments, Organisms, and Templates. Below the tabs is a 'New reagent' section with a search bar. The main part of the interface is a table listing various reagents. Each row includes columns for Analyte, Analyte detector, Analyte reporter, Clone, Catalog nr, and Manufacturer. To the right of the Manufacturer column are 'Edit' and 'Remove' links for each entry. The table lists 14 reagents, including CCR7, CD11c, CD123, CD14, CD19-biotin, CD40, CD86, IFNa, IL12p40/70, IL6, MHCII, and TNFa. The bottom of the table indicates 'Showing 1 to 14 of 14 entries'.

Analyte	Analyte detector	Analyte reporter	Clone	Catalog nr	Manufacturer		
CCR7	anti-CCR7	PE-Cy7	3D12	BD#557648	BD Biosciences [website]	Edit	Remove
CD11c	anti-CD11c	APC	5HCL3	BD#340714	BD Biosciences [website]	Edit	Remove
CD123	anti-CD123	Pacific Blue	9F5	BD#custom	BD Biosciences [website]	Edit	Remove
CD123	anti-CD123	AmCyan	9F5	BD#custom	BD Biosciences [website]	Edit	Remove
CD14	anti-CD14	Alexa 700	M5E2	BD#557923	BD Biosciences [website]	Edit	Remove
CD14	anti-CD14	PE-Cy7	M5E2	eBio#25-0149	eBioscience [website]	Edit	Remove
CD19-biotin	anti-CD19-biotin	PE-TexRed	streptavidin PE-TR	BD#555411, BD#551487	BD Biosciences [website]	Edit	Remove
CD40	anti-CD40	FITC/OG	5C3	eBio#11-0409	eBioscience [website]	Edit	Remove
CD86	anti-CD86	PE	IT2.2	eBio#12-0869	eBioscience [website]	Edit	Remove
IFNa	anti-IFNa	FITC/OG	A11	Antigenix#MC100133	Antigenix [website]	Edit	Remove
IL12p40/70	anti-IL12p40/70	Pacific Blue	eBio: C8.6	eBio#577129	eBioscience [website]	Edit	Remove
IL6	anti-IL6	APC-Cy7	MQ2-13A5	BD#custom	BD Biosciences [website]	Edit	Remove
MHCII	anti-MHCII	PerCPy5.5	TU36	BD#custom	BD Biosciences [website]	Edit	Remove
TNFa	anti-TNFa	Alexa 700	Mab11	BD#557996	BD Biosciences [website]	Edit	Remove

Showing 1 to 14 of 14 entries

**Figure 10.18.9** Reagents available for annotation. Users describe available reagents in the annotation data section. Later on, samples can be annotated with these reagents.

Typically, you do not need to edit anything in the first three tabs (Keywords, Organizations, and Manufactures).

The table of reagents contains a list of reagents and their characteristics as required by MIFlowCyt. You can filter the reagents shown by starting to type in the Search field.

19. Add all reagents used in your experiment (if not listed already) by invoking the “New reagent” form (Fig. 10.18.10) repeatedly. Later on, you will be able to assign reagents to FCS files or to combine them into reusable staining cocktails.

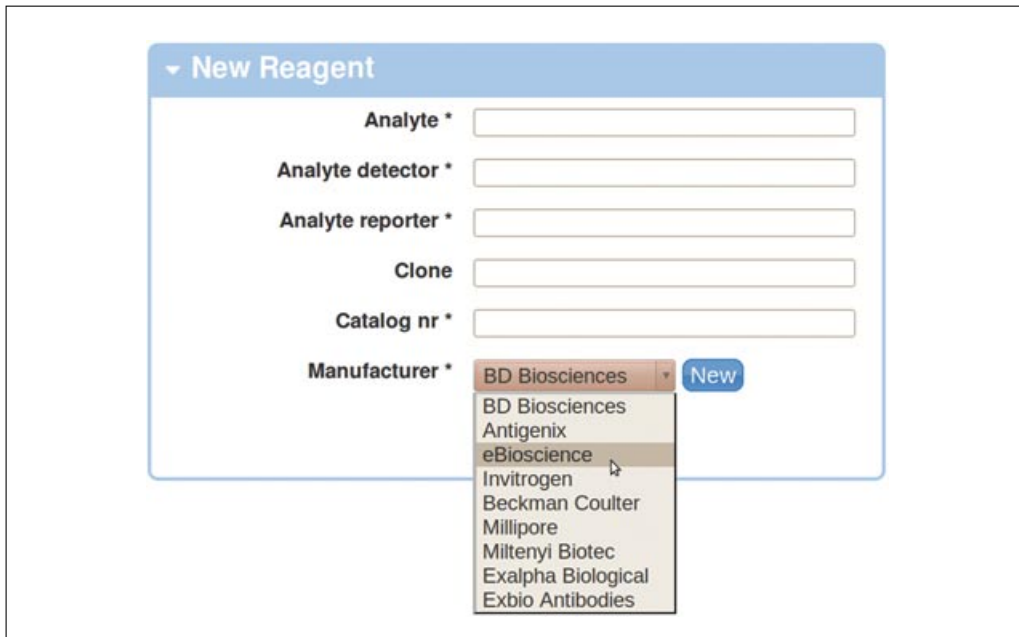
*For each new reagent, fill out the analyte, analyte detector, analyte reporter, clone, and catalog number as applicable. See MIFlowCyt section 2.4 for details (Lee et al., 2008). Select the reagent manufacturer from the list. In case your reagent manufacturer is not listed, then follow the New button to create and save a new Manufacturer.*

20. Click on the Annotation Data link, Instruments tab, to ensure that the make and model of the flow cytometer(s) used in your experiment is (are) listed among the known instruments.

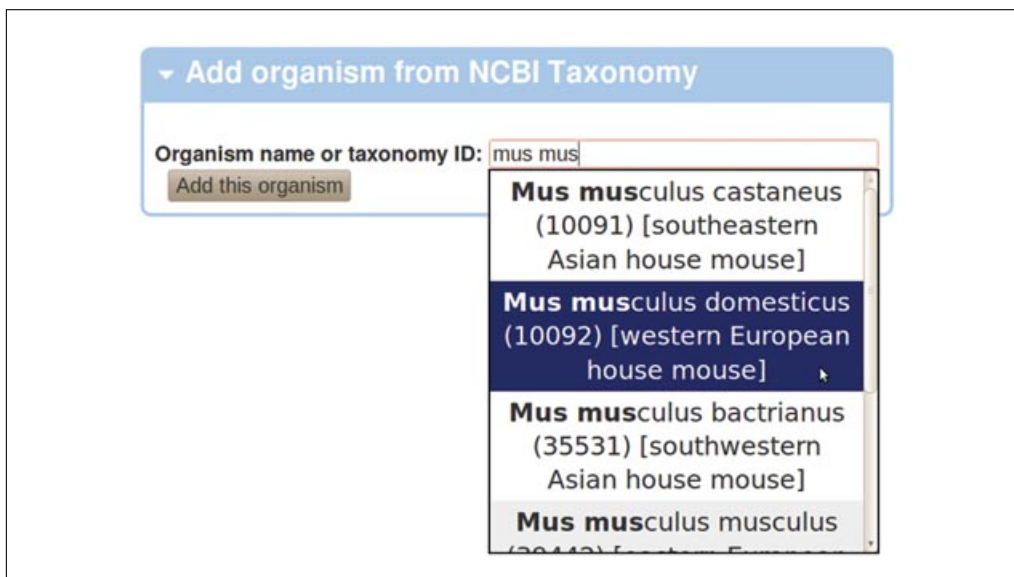
*You can filter the instruments shown by starting to type in the Search field. The majority of existing instruments are in the system already. Follow the “New instrument” link and provide the model and manufacturer in case your instrument is not listed. Click the “New manufacturer” button in case the manufacturer of your instrument is not known to FlowRepository.*

21. Click on the Annotation Data link, Organisms tab, to ensure that the organism(s) used in your experiment is (are) listed among the known organisms.

*You can filter the organisms shown by starting to type in the Search field. The NCBI taxonomy contains hundreds of thousands of organisms. In order to simplify the selection, we have imported only a few that are most commonly used in flow cytometry. If your biological sample is derived from an organism that is not listed, follow the “Add organism” link, and you will be able to add it from the NCBI taxonomy (Fig. 10.18.11). Start typing either the Latin or the English name; the auto-complete will show suggestions after the first three characters are typed. Select the appropriate organism and add it by pressing the Add this organism button.*



**Figure 10.18.10** New reagent form. Reagent details should be provided as required by MIFlowCyt.



**Figure 10.18.11** Adding an organism from the NCBI taxonomy. Suggestions will be displayed after three or more characters are typed.

22. Click on the Annotation Data link, Templates tab, and add templates for your sample sources by following the New annotation template for sample sources link.

*Annotation templates may be created in order to simplify and speed up MIFlowCyt annotation. Templates contain common annotations that can be applied on specific samples. For biological sample sources, MIFlowCyt requires details such as description, organism, age, gender, phenotype, genotype, and treatment. For example, let us assume that your sample sources are human donors that differ in their gender and age, but everything else is assumed to be the same. You can create an appropriate annotation template, fill out all the common details, and leave the gender and age out of the template. Later on, when annotating the FCS files of your experiment, you will be able to easily create sample sources based on this template, and you will only have to provide the gender and age that differ among the subjects.*

23. Click on the “From new” button in the New Annotation Template box with Sample source listed as the template type (Fig. 10.18.12). Select the sample source type and fill out all the fields as required by MIFlowCyt. You may enter ? or one of the common values in the fields that are variable among your sample sources; you will be rewriting this value when applying the template to specify actual sample sources. Click on the Save button to close the Details for sample source template pop-up window. Finally, name the template (you may use the default name as suggested based on your description) and click on the Create button.

*MIFlowCyt section 2.1 requires sample materials to be described depending on type: biological, environmental, or other. The form will dynamically adjust its fields once you have selected your sample source type. For biological samples, you will need to state the description, organism, age (with units), gender, phenotype, genotype, and treatment. A description and location is required for environmental sample sources, and description only for other sample-source types. Additional information may be provided in the other field (optional).*

24. Click on the Annotation Data link, Templates tab, and add templates for your samples by following the New annotation template for sources link.
25. Create your sample templates in a similar fashion to that used to create the sample source template earlier. Click on the “From new” button in the New Annotation



**Figure 10.18.12** Creation of a biological sample source template. Required fields will change dynamically based on the type of sample source specified. Values provided in a template may be adjusted when the template is used to annotate specific samples.

Template box (now with Sample listed as the template type). Fill out the description, sample characteristics, and treatment as required by MIFlowCyt sections 2.2 to 2.3.

*You will need to create new entries by following the New buttons in order to describe sample treatment and characteristics. Describe these accordingly, save them, and then select in the drop-down selection boxes.*

26. Select existing staining reagents and/or cocktails or create new entries (Fig. 10.18.13) based on reagents specified in step 19. For each reagent used, follow the New button, fill in the characteristics being measured (see MIFlowCyt section 2.4), select the reagent, and save it with the Save button. If a staining cocktail is used, specify its name, catalog number, and manufacturer (if applicable), and select (create) all fluorescence reagents used in the cocktail.

*Make sure the appropriate staining reagents and/or cocktails are selected before saving the sample template form. You may select/deselect multiple entries by pressing Ctrl or Shift while you click on an item.*

*Specifying staining details in your sample template is only useful if the same staining is used across multiple samples. Otherwise, you may skip this step and leave the staining unspecified in the template; you will still be able to specify staining details for each FCS file individually (step 30 below).*

27. Click on the Save button to close the Details for sample template pop-up window. Finally, name the template (you may use the default name as suggested based on your description) and click on the Create button.

**Figure 10.18.13** Application of a reagent. The characteristic being measured should be specified when samples (or templates) are annotated with reagents, since the same reagent could be used for slightly different purposes in different experiments.

### Annotate FCS files

28. Return to your experiment overview by clicking on the Inbox link in the header and then the name of your experiment.

*You may use the Search field if the size of your Inbox does not allow for the experiment to be located easily.*

29. Follow the “describe sample” link of the first FCS file in the FCS Files box at the end of the experiment overview page. Confirm that you have prepared your annotation data and that you wish to continue with sample description.

30. Apply the sample template that characterizes the sample for this FCS file and adjust the description as appropriate.

*Applying a template will automatically fill in the sample description, characteristics, treatment, and staining. You may adjust any of these the same way as when defining your sample template in steps 25 to 26.*

*Details about the sample description, characteristics, and treatment may also be extracted from a spreadsheet rather than provided manually; you may skip these at this point and see steps 34 to 38 for instructions how to create CSV files with annotations.*

31. Define the sample source for this sample by clicking on the new sample source button, selecting a template, and adjusting as appropriate.

*Selecting an appropriate sample source template will fill in the sample source type and all the specific details as provided when defining the template in steps 22 to 23. Details that are specific for this particular FCS file should be corrected accordingly. In addition, you should change the description to unambiguously identify this particular sample source before closing the New sample source details pop-up window by clicking on the Save button.*

*Details about the sample source description, location, age, gender, phenotype, genotype, and treatment, as well as other information, may also be extracted from a spreadsheet rather than provided manually; you may skip these at this point and see steps 34 to 38 for instructions how parse CSV files with annotations.*

32. Verify that the correct sample source is selected and that all other details are correct on the FCS sample information page before clicking on Save and proceeding to the next FCS file button.

*Alternatively, you may apply the same template to all FCS files and then extract and overwrite sample and sample source details from a CSV file. If you intend to do so, press the “Apply to all FCS files” button and see steps 34 to 38 for instructions how parse CSV files with annotations.*

33. Repeat the same process to annotate all FCS files.

*If the same sample source has been used for multiple samples (FCS files), then you should select it from the sample source drop-down selection rather than creating a new sample source from a template.*

*You will be automatically directed back to your experiment overview page once you have annotated all FCS files. The links in the sample columns of the FCS Files box will change from describe sample to show sample description for all FCS files that have been annotated.*

#### ***Extract sample and sample source information from a spreadsheet (optional)***

Steps 34 to 38 are applicable if you opted for extracting some of your sample and/or sample source annotations from a CSV file (spreadsheet) rather than providing these details by filling out on-line forms. Many researchers find this a very useful option, especially if the experiment contains many FCS files. Support for additional file formats may be developed as per community requirements; please see documentation provided on-line at the FlowRepository Web site for updated information.

34. Use your favorite spreadsheet tool to prepare a single table spreadsheet with sample and/or sample source annotations. Provide column headings in the first row. The spreadsheet should contain a column named FCS file. In addition, there should be one or more of the following columns: sample description, sample characteristic, sample treatment, sample source description, sample source treatment, age, age unit, gender, phenotype, genotype, location, or other sample source information. The location is applicable for environmental samples only. Annotation details should be provided starting from the row number 2. The FCS file column should contain file names corresponding to the data files in your experiment. Other cells should contain values that you wish to assign to samples and sample sources associated with the particular FCS file.

*The comma-separated values (CSV) file format is a simple spreadsheet file format supported by virtually all spreadsheet tools. You may use MS Excel, OpenOffice spreadsheet (Calc), a Google Doc spreadsheet, or any other CSV-compatible program.*

35. Save the spreadsheet to your local computer as a CSV file. Make sure the field delimiter is set to comma (.). The file should be saved with the .csv file name extension.

*The spreadsheet can typically be saved as a CSV file by selecting the CSV option in the File/Save, File/Export, or File/Download as menu item. Files from MS Excel and Google Doc are typically compatible without any further tweaks. With OpenOffice, some versions tend to use semicolon as the default delimiter character. If this is the case, then you will need to change it to comma by selecting the Edit filter settings in the OpenOffice Save dialog and choosing “comma” as the field delimiter (keep UTF-8 as the character set and quotation marks as the text delimiter).*

36. Use the Browse button within the Attachments box of your experiment view to locate your annotations CSV file, and upload it as attachment to your experiment by pressing the Upload button.

37. Parse the attachment to extract and update information by clicking on the parse icon (P in circle) next to the CSV attachment listed in the Attachments box.

*You will need to confirm that you understand the consequences of performing this action. Specifically, it will replace some of the annotations saved in FlowRepository with values specified in your CSV file, without the option of going back.*

*Since the MIFlowCyt-required annotation depends on the sample type, you can only apply annotations from a CSV file to samples that have the sample type assigned already. This means that the sample must have been annotated previously, as specified in steps 28 to 33. This can be verified by reviewing the Sample column in the FCS Files box in your experiment view below, which should be displaying the “show sample description” link rather than the “describe sample” link.*

38. Review the results of extracting information from the spreadsheet by inspecting the information box that will temporarily appear on your experiment view.

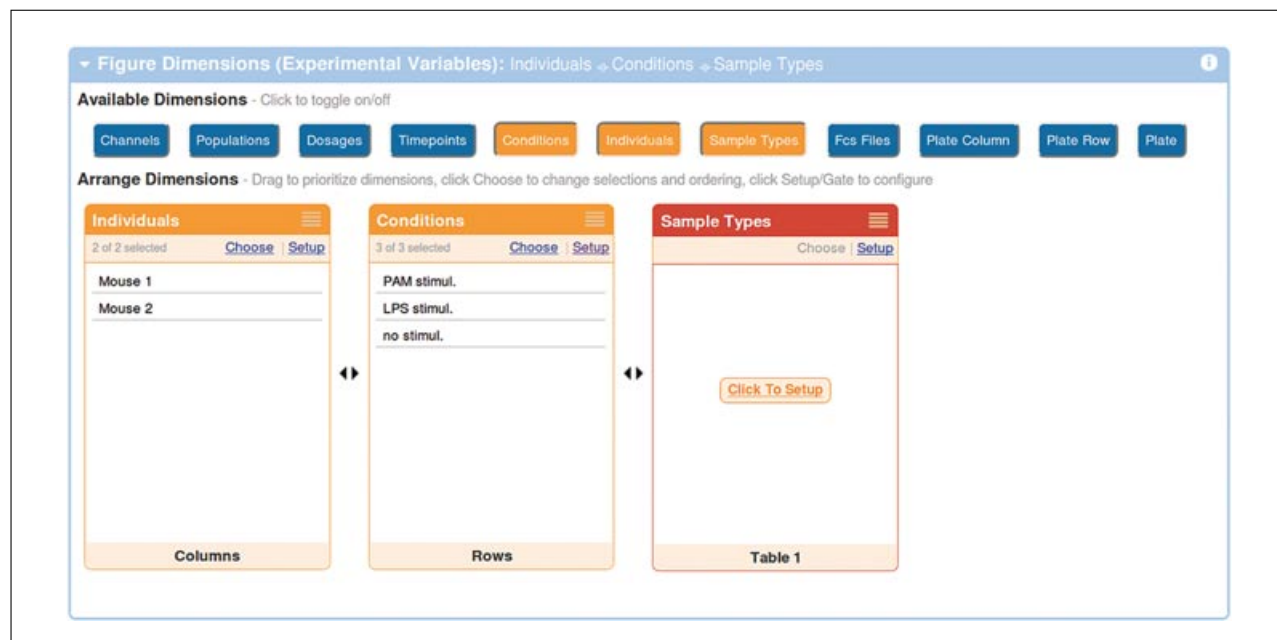
*The information box will list annotation details than have been applied, as well as any unrecognized items or possible errors. The notation of File[item]=value is used. For example, 152597.fcs[age]=21 indicates that the age field of the sample source associated with FCS file 152597.fcs has been set to the value of 21.*

### **Specify experiment variables**

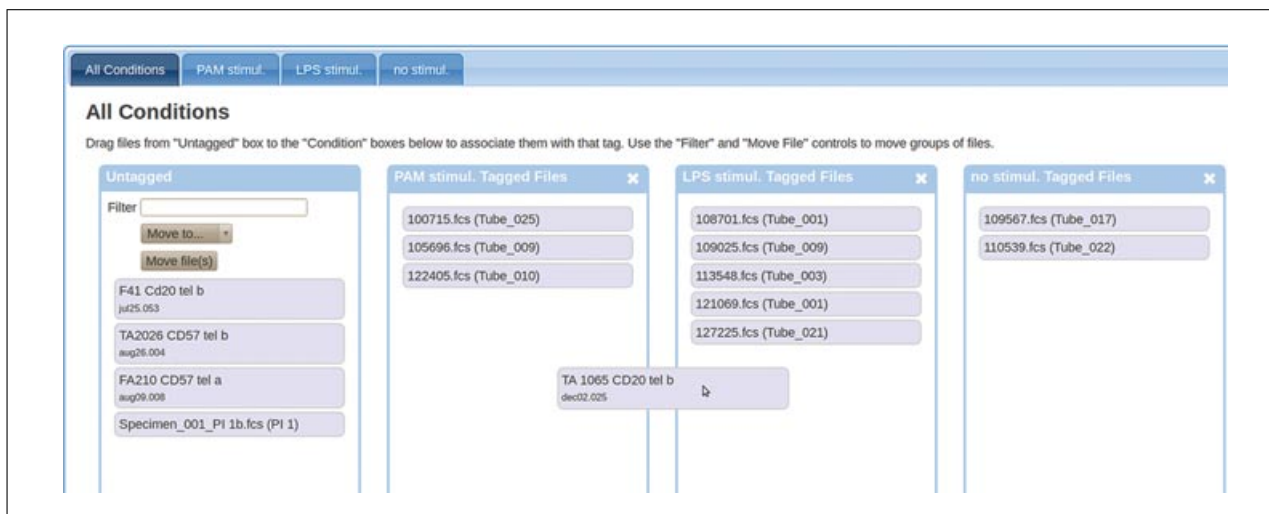
39. Follow the Review experiment variables link in the green information box or click on your working illustration in the blue Illustrations box.

*Experiment variables are attributes that differ between samples within an experiment due to pre-existing differences in sample states or due to experimental manipulation of the samples (e.g., smoker versus non-smoker, IL-2 treatment versus no treatment, knockout versus wild-type, or variable number of transplanted stem cells).*

40. Click on the appropriate experiment variable type (Fig. 10.18.14, dosages, time points, conditions, individuals, etc.).



**Figure 10.18.14** Setting up experimental variables. Experimental variables are defined and assigned in the illustration creation section.



**Figure 10.18.15** Conditions assignment. Experimental variables (e.g., sample stimulation conditions) can be assigned to FCS files by dragging and dropping the file labels to appropriate boxes.

*Experiment variables are commonly used in Figures as dimensions, since researchers are typically interested in visual inspection of the effects of different values of a variable. Therefore, experimental variables are defined within the illustration section of FlowRepository. However, you may still define variables and assign their values to FCS files without using these in any of your illustrations.*

*For the purpose of this example, we will assume that you have selected Conditions as the appropriate variable type for your experiment; a red box titled Conditions will appear.*

41. Click on the Setup link inside the Conditions box.
42. Enter a comma-separated list of conditions to add in the Add conditions box (e.g., PAM stimul., R-FSL stimul., pIC stimul., LPS stimul., no stimul.), and click on the Add Conditions button.

*This will create additional boxes (containers) for your FCS files, one box for each of the conditions specified.*

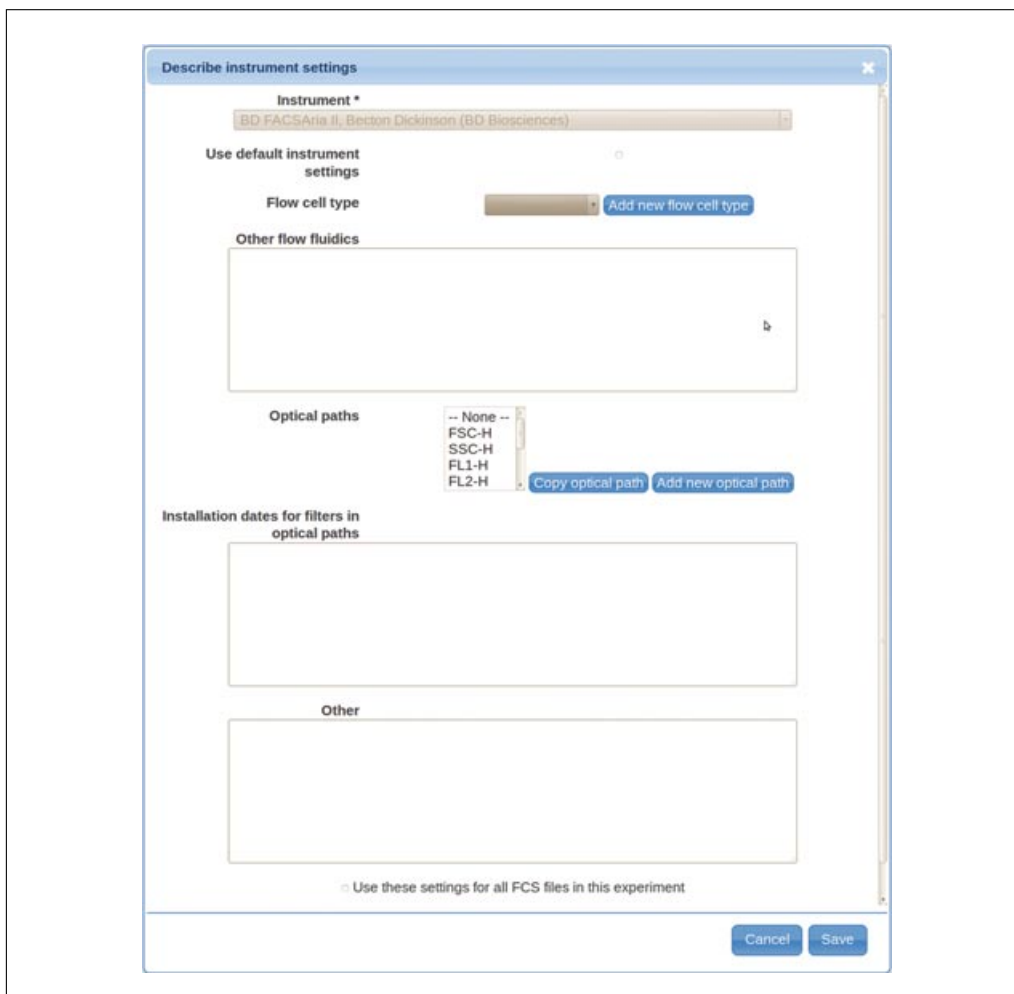
43. Assign FCS files to the appropriate conditions (Fig. 10.18.15) by dragging them and dropping in the appropriate box. Click on Return to Illustration once all FCS files are correctly assigned (tagged).
44. Repeat steps 40 to 43 for other types of experimental variables you wish to define. The appropriate boxes, buttons, etc. will be named based on the type of variable you are defining.

*You may toggle the boxes with experimental variables on and off (i.e., include/exclude from a working illustration) by clicking on the variable type in the row of Available Dimensions. You need to display a variable in order to define the values (assign FCS files), but you can safely hide it afterwards.*

### **Describe instrumentation settings**

45. Return to your experiment overview by clicking on the Back to Experiment link.
46. Follow the details link near the first of your FCS files listed in the FCS Files box at the bottom of the page.
47. Click on the “Describe instrument settings” button inside the FCS File Instrument Settings box





**Figure 10.18.16** Description of instrumentation settings. The dialog window is asking for details as required by MIFlowCyt. Consequently, it will be significantly simplified if default instrumentation settings are used.

48. From the drop-down selection box, select the instrument used for acquisition of the particular FCS file.

*See step 20 above in case your instrument is not included in the selection.*

*The dialog window will expand to include the instrumentation settings description once the instrument has been selected (Fig. 10.18.16).*

49. If you have not altered the instrument by customizing optical filters, light sources, or other hardware components, then select the “Use default instrument settings” check box and skip to step 65 below.

*As per MIFlowCyt section 3, using default configuration will significantly minimize the effort needed for proper instrumentation annotation. Steps 50 to 64 are included to cover cases when an instrument has been customized.*

50. Select the flow cell type of your instrument from the drop-down selection box. Use the “Add new flow cell type” button and enter the appropriate type in case your flow cell type is not listed in the drop-down selection box.
51. Provide additional flow fluidics information in the “Other flow fluidics” field (optional).

*As per MIFlowCyt section 3.3.1.2., this may include the flow cell manufacturer name and model number, flow cell outside shape and dimensions, flow channel shape and dimensions, sheath flow velocity, sheath flow pump type, sample flow velocity, and sample flow pump and other relevant details.*

52. Click on “Add new optical path” in order to describe the optical path for a channel.

*You will need to describe the optical path for each channel.*

53. Use the Parameter drop-down selection box to select the FCS parameter (channel) for which you are describing the optical path.

54. Specify the basic components of the optical path by dragging and dropping the framed components (i.e., light source, filter, optical detector, and others) into the sorted list in the yellow box. You may also drag the labels around inside the box to re-order them, or remove them by dragging them over to the trash bin symbol.

*Initially, the list will contain three items creating a simple optical path to begin with—a light source, a filter, and an optical detector. You can always reset to this stage by clicking on the garbage bin symbol.*

55. Double-click on the Light Source label in the ordered list (typically, item 1) to provide details about the light source. There are two ways to specify the details, either by defining a new light source (click on the Create new light source button) or by selecting a previously described light source.

*For the first optical path, you will likely have to define a new light source. However, you will be able to reuse previously defined light sources by selecting them from the “Select existing light source” drop-down selection box when defining an additional optical path.*

56. Select the light source type from the drop-down selection box. Specify the excitatory wavelength, power, polarization (optional), beam characteristics (optional), and other relevant information (optional), and save the information by clicking on the “Add component” button.

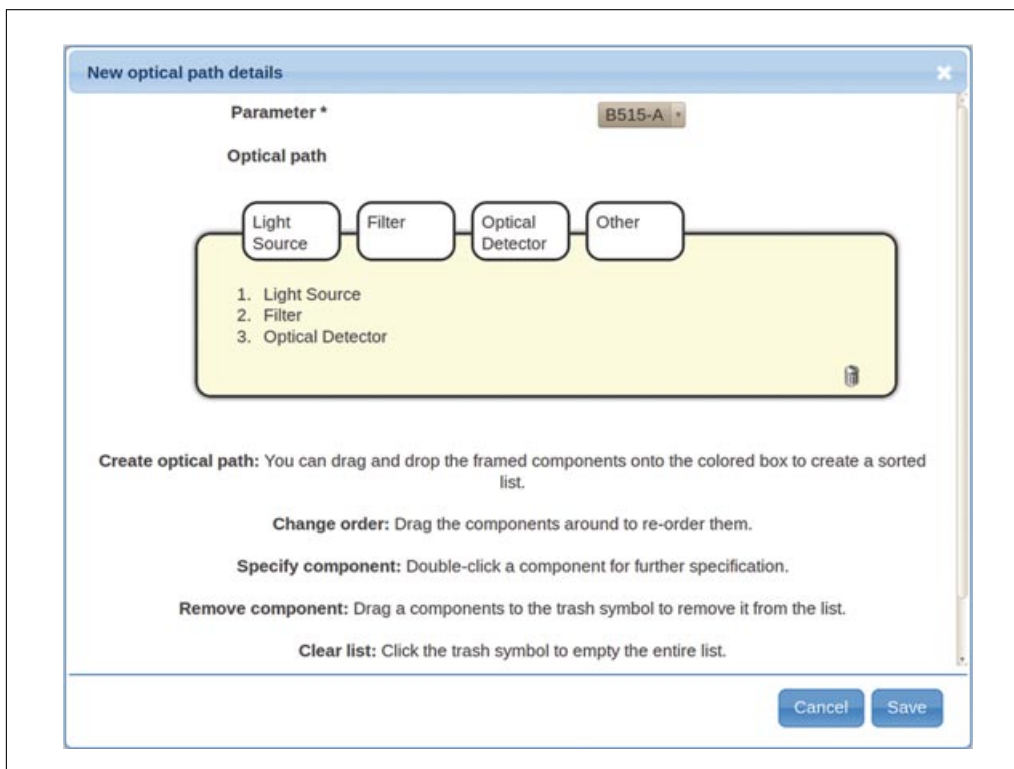
*Specified details are as required by MIFlowCyt section 3.3.2. The label in the ordered list will change from light source to light source—edited (Fig. 10.18.17) to indicate that details about this particular component have been provided already. Now, you will need to describe the other components.*

57. Double-click on the Filter label in the ordered list (typically, item 2) to provide details about the optical filter. Again, there are two ways to specify the details—either by defining a new filter (click on the Create new filter button) or by selecting a previously described optical filter.

*You will have to define a new optical filter the first time you need to use it; however, you will be able to reuse previously defined filters by selecting them from the Select existing filter drop-down selection box if these filters are used in multiple optical paths.*

58. Select the filter type from the drop-down selection box. Specify the transmitted wavelengths, model (optional), manufacturer (optional), and other relevant information (optional). Follow the New link near the manufacturer field if the manufacturer of your optical filter is not included in the list. Provide the name, Web site (optional), phone number (optional), and other important information (optional) about the manufacturer, and save it by clicking the Save button. Save the new optical filter by clicking on the “Add component” button.

*Specified details are as required by MIFlowCyt section 3.3.4 except for optical filter installation dates (MIFlowCyt section 3.3.4.3), which will be addressed in step 65 below. The label in the ordered list will change from filter to filter—edited to indicate that details about this particular component have been provided already.*



**Figure 10.18.17** Definition of optical path. An optical path can be defined by dragging and dropping its components into an ordered list in the yellow box. Details about each component need to be provided prior saving the optical path.

59. Repeat steps 57 to 58 to describe all other optical filters that are on the same optical path.
60. Double-click on the Optical Detector label in the ordered list (typically the last item) to provide details about the optical detector. As before, there are two ways to specify the details, either by defining a new optical detector (click on the Create new detector button) or by selecting a previously described optical detector.

*You will have to define a new optical detector the first time you need to use it; however, you will be able to reuse previously defined detectors by selecting them from the “Select existing detector” drop-down selection box if these detectors are used in multiple optical paths (e.g., if multiple signal characteristics, such as the height and area, of a certain signal are captured).*

61. Select the optical detector type from the drop-down selection box. Specify the name (e.g., PMT A; typically, identifiers assigned by instrument manufacturers make up for an appropriate name), amplification type (optional), and other relevant information (optional). Save the new optical detector by clicking on the Add component button.

*Specified details are as required by MIFlowCyt section 3.3.5, except for optical detector voltages (MIFlowCyt section 3.3.5.3), which are typically included in the optional \$PnV keywords in the FCS file (Spidlen et al., 2010). Use the field for “other relevant information” to provide the detector voltage in case it is not included in your FCS file. The label in the ordered list will change from optical detector to optical detector—edited to indicate that details about this particular component have been provided already.*

62. If there are other optical components (beam splitters, spectrometer, etc.) on the optical path, describe these similarly to how you just described the light source, optical filters, and the optical detector. Double click on the Other label in the ordered

list, provide the name, type (optional), and description (optional), and save the component by clicking on the Add component button.

63. Save the optical path by clicking on the Save button on the New optical path details.

*The newly created optical path will get added among the existing Optical paths on the Describe instrument settings dialog window.*

64. Repeat steps 52 to 63 and add the optical path for each measured parameter where applicable (i.e., excluding time).

*Make sure all the relevant optical paths are selected in the “Optical paths” selection box; you may add/remove an item from the current selection by holding Ctrl while clicking on the item.*

65. Provide installation dates for filters in optical paths and other relevant details (optional) in the “Describe instrument settings” dialog.

*You may list specific dates for each filter or include a generic statement, such as “all filters came with the instrument purchased on February 26, 2010.”*

66. Check the “Use these settings” for all FCS files in this experiment check box if you have used the same instrument to acquire all your data files. Close the instrument settings dialog using the Save button.

*If you used multiple instruments to acquire your data, then you will have to describe these separately by repeating steps 45 to 66 for each of the FCS files.*

#### **Provide your analysis details**

67. Return to the experiment illustrations section by following the Back to Experiment Summary link, and then the “My working illustration” link. Gate and analyze your data online (optional) and create custom illustrations (optional) by following the data analysis protocols described in *UNIT 10.17*.

*Alternatively, you may upload third-party analysis files (e.g., FlowJo workspaces, FCS Express project files, FACS Diva files, etc.) and relevant figures, summary statistics, etc., as attachments if you prefer using your own tools for data analysis. Your MIFlowCyt score will not be negatively affected if you decide to provide third-party analysis files rather than analyzing data online.*

#### **Review and improve your MIFlowCyt compliance**

68. Review your MIFlowCyt score details by following the MIFlowCyt Score compliance progress bar on the right side of your experiment header.





*At this point, your MIFlowCyt compliance score should be close to 100% if you followed all previously described steps and provided all the information required by MIFlowCyt.*

69. Review the details of your annotation compliance for each MIFlowCyt section and provide additional information as needed to improve your score.

*There are four tables detailing the compliance of your annotation with the MIFlowCyt standard, one for each section of MIFlowCyt. In each of these tables, the compliance column shows a green (fully compliant), orange (partially compliant), or red (information missing) checkmark. You may expand details by clicking on the checkmark symbol. All details may be expanded or collapsed by clicking on the plus or minus sign in the table heading, respectively. The “improve” column provides links for you to use to improve your annotation and increase your score (Fig. 10.18.18).*

#### **Share your dataset with reviewers**

70. Once you are satisfied with your annotations, return to the experiment summary by following the Back to Experiment Summary link.

2 - Flow Sample/Specimen Details - 73.42% provided		
Items considered relatively based on importance, 30% contribution to total score.		
FCS file	Compliance [ + - ]	Improve
452 WT_001.fcs	 Provided sample description Provided sample source description Provided sample source organism Provided sample source age Provided sample source gender Provided sample source phenotype Provided sample source genotype Provided sample source treatment Provided sample characteristic Missing sample treatment description Provided fluorescence reagent annotation	<a href="#">[ View   Edit ]</a> sample description
454 KO_001.fcs		Fully provided
451 KO_001.fcs		Fully provided
Partec Example.FCS		<a href="#">Describe sample</a>

**Figure 10.18.18** MIFlowCyt compliance review of own experiments. An additional column with references to improve the annotation is provided if applicable.

71. Click on the Share with Reviewers button inside the Sharing Permissions box. Confirm the message (press OK) notifying you that sharing this experiment with reviewers will also lock your experiment.

*As a result, the experiment will be locked and a URL with a secret access code will be generated (e.g., <https://flowrepository.org/id/RvFra4aTsCvHN0Yxrqfc-NrwOX5DvygvxQGCbly2QqZpKbYXrSrpWNswo7XXVood>).*

72. Share the secret URL with your reviewers.

*For example, you may copy and paste the URL to the cover letter of your manuscript. With this URL, the reviewers will be able to obtain anonymous read-only access to the Public View of your experiment. Once your manuscript is published, the experiment will become publicly available without the need for a secret access code.*

*It is recommended to also include your repository identifier (FR-FCM-xxxx as discussed in Basic Protocol 1, step 5) in the text of the related manuscript. Placing the repository identifier in the manuscript text will help readers locate the related dataset. Depending on the journal, your data may automatically become publicly available once the manuscript is published. Please contact FlowRepository.org administrators if unsure.*

*If reviewers ask for changes to your dataset and/or annotation prior publication of the manuscript, the editor will typically arrange for the experiment to be unlocked for you to edit. Should this not be the case, please contact FlowRepository administrators by following the Support link in the footer of each page.*

## COMMENTARY

### Background Information

Transparency and public availability of protocols, data, analyses, and results are crucial in order to make sense of the complex biology of human disease, and research efforts should be integrated across teams in an open environment. Private funding agencies, regulatory agencies, publishers, and the scien-

tific community have all recognized the importance of protecting cumulative data outputs to accelerate subsequent exploitation through the community-based development of public data repositories (Field et al., 2009). In a recent re-evaluation of 18 peer-reviewed *Nature Genetics* microarray articles, the inability of researchers to reproduce analyses was



directly linked to data unavailability, incomplete data annotation, or specification of data processing and analysis (Ioannidis et al., 2009). Data sharing allows scientists to expedite the translation of research results into knowledge, products, and procedures to improve human health. Sharing data reinforces open scientific inquiry, encourages diversity of analysis and opinion, promotes new research, makes possible the testing of new or alternative hypotheses and methods of analysis, supports studies on data-collection methods and measurement, facilitates the education of new researchers, enables the exploration of topics not envisioned by the initial investigators, and permits the creation of new datasets when data from multiple sources are combined.

The International Society for Advancement of Cytometry (ISAC), founded in 1979 and the largest international flow cytometry community with approximately 2000 members, recognized the need for a public, community-based data repository. FlowRepository was developed under the auspices of ISAC by extending and adapting Cytobank—an online tool for collaborative analysis of cytometric data to address this need.

To efficiently exchange information, the community must first agree on the kind of information they want to store, though the precise content of documentation will vary by scientific area, study design, type of data collected, and characteristics of the dataset. In flow cytometry, this need was addressed by MIFlowCyt (Lee et al., 2008), which is accepted as a standard by the society. The journal publishers Wiley-Blackwell and Nature Publishing Group reference MIFlowCyt in their instructions to authors. MIFlowCyt is also a component of the Minimum Information for Biological and Biomedical Investigations (MIBBI) set of minimal informational standards (Taylor et al., 2008), and it is endorsed by the Data Interoperability Steering Committee of the Division of Allergy, Immunology and Transplantation within the National Institute of Allergy and Infectious Diseases. Based on the microarray repository experience, the establishment of MIFlowCyt as a minimum information standard was the first crucial step toward developing universal solutions for representing, collecting, annotating, archiving, analyzing, and disseminating flow cytometry data through a public repository. Consequently, support for all MIFlowCyt compatible annotation is a crucial requirement

for FlowRepository, and has influenced many of its design aspects and decisions. Although the effort spent providing MIFlowCyt annotations may seem like a burden to data submitters, it will pay off by facilitating third-party understanding of the data and conclusions reached, and by allowing for comparisons to experiments performed by different laboratories. It will foster scientific collaboration and preserve the integrity and quality of cytometry data, which will contribute to the advancement of biomedical sciences by enabling scientists to build upon previous findings.

### Critical Parameters

The speed of your Internet connection may become a bottleneck, especially when uploading large datasets consisting of gigabytes of data files. We suggest that you upload your data overnight in this case (you do not need to be present by the computer while the upload is in progress).

In clinical environments, users may need to de-identify FCS files in order to be able to share the dataset while conforming to applicable privacy rule requirements, such as the Health Insurance Portability and Accountability Act (HIPAA, Harrelson et al., 2007). FCS data files contain a segment with keyword/value pairs. Most of these keep basic information essential for the interpretation of the raw data matrix of fluorescence expression levels and other values, such the number of acquired parameters, their names, detector voltage settings, etc. However, some of these keywords may include values that could be used to identify an individual and should therefore be removed. Currently, it may be challenging for clinicians to edit FCS files locally, although there are some tools available that can help with that task (e.g., a free De-Identification Utility offered by Tree Star, Inc., available from <http://www.flowjo.com/utilities/>). In the future, we intend to simplify this process by integrating the FCS de-identification functionality directly in the data upload mechanism.

### Troubleshooting

Cytobank troubleshooting tips, common issues, and questions can be found at <http://docs.cytobank.org/wagn/Troubleshooting>. FlowRepository-specific inquiries and comments may be submitted via the support link in the footer of each FlowRepository page or directly via an email to [flowrepository-admin@isac-net.org](mailto:flowrepository-admin@isac-net.org).

## Time Considerations

The time and effort required to upload and properly annotate a dataset is proportional to the size and complexity of your experiment. It can reach anywhere from ~30 min for simple cases with a handful of FCS files acquired on a single instrument with default instrumentation settings, to several hours for very large and complicated experiments with hundreds of data files, several experimental variables, and many samples.

## Acknowledgements

Development of FlowRepository was supported by the International Society for Advancement of Cytometry, the Terry Fox Research Institute, the Terry Fox Foundation, and the Wallace H. Coulter Foundation. The authors would also like to thank Cytobank, Inc. and its development team (especially Nikesh Kotecha and Chad Rosenberg) for providing the underlying technological platform as free, open-source software, and for their assistance throughout the project, as well as initial hosting of the FlowRepository application. Ryan R. Brinkman was supported by a Michael Smith Foundation for Health Research Scholar Award.

## Literature Cited

- Bashashati, A. and Brinkman, R.R. 2009. A survey of flow cytometry data analysis methods. *Adv. Bioinformatics* 584603.
- Blimkie, D., Fortuno, E.S. 3rd, Thommai, F., Xu, L., Fernandes, E., Crabtree, J., Rein-Weston, A., Jansen, K., Brinkman, R.R., and Kollmann, T.R. 2010. Identification of B cells through negative gating: An example of the MIFlowCyt standard applied. *Cytometry A* 77:546-551.
- Bocsi, J. and Tárnok, A. 2008. Toward automation of flow data analysis. *Cytometry A* 73:679-680.
- Burgoon, L.D. 2006. The need for standards, not guidelines, in biological data reporting and sharing. *Nat. Biotechnol.* 24:1369-1373.
- Field, D., Sansone, S., Collis, A., Booth, T., Dukes, P., Gregurick, S.K., Kennedy, K., Kolar, P., Kolker, E., Maxon, M., Millard, S., Mugabushaka, A., Perrin, N., Remacle, J.E., Remington, K., Rocca-Serra, P., Taylor, C.F., Thorley, M., Tiwari, B., and Wilbanks, J. 2009. 'omics data sharing. *Science* 326:234-236.
- Harrelson, J.M. and Falletta, J.M. 2007. The privacy rule (HIPAA) as it relates to clinical research. *Cancer Treat. Res.* 132:199-207.
- Ioannidis, J.P., Allison, D.B., Ball, C.A., Coulibaly, I., Cui, X., Culhane, A.C., Falchi, M., Furlanello, C., Game, L., Jurman, G., Mangion, J., Mehta, T., Nitzberg, M., Page, G.P., Petretto, E., and van Noort, V. 2009. Repeatability of published microarray gene expression analyses. *Nat. Genet.* 41:149-55.

Lee, J.A., Spidlen, J., Boyce, K., Cai, J., Crosbie, N., Dalphin, M., Furlong, J., Gasparetto, M., Goldberg, M., Goralczyk, E.M., Hyun, B., Jansen, K., Kollmann, T., Kong, M., Leif, R., McWeeney, S., Moloshok, T.D., Moore, W., Nolan, G., Nolan, J., Nikolich-Zugich, J., Parrish, D., Purcell, B., Qian, Y., Selvaraj, B., Smith, C., Tchuvatkina, O., Wertheimer, A., Wilkinson, P., Wilson, C., Wood, J., Zigon, R., Scheuermann, R.H., and Brinkman, R.R. 2008. MIFlowCyt: The Minimum Information about a Flow Cytometry Experiment. *Cytometry A* 73:926-930.

Lugli, E., Roederer, M., and Cossarizza, A. 2010. Data analysis in flow cytometry: The future just started. *Cytometry A* 77:705-713.

Spidlen, J., Leif, R.C., Moore, W., Roederer, M., and Brinkman, R.R. 2008. Gating-ML: XML-based gating descriptions in flow cytometry. *Cytometry A* 73:1151-1157.

Spidlen, J., Moore, W., Parks, D., Goldberg, M., Bray, C., Bierre, P., Gorombey, P., Hyun, B., Hubbard, M., Lange, S., Lefebvre, R., Leif, R., Novo, D., Ostruszka, L., Treister, A., Wood, J., Murphy, R.F., Roederer, M., Sudar, D., Zigon, R., and Brinkman, R.R. 2010. Data File Standard for Flow Cytometry, version FCS 3.1. *Cytometry A* 77:97-100.

Taylor, C.F., Field, D., Sansone, S.A., Aerts, J., Apweiler, R., Ashburner, M., Ball, C.A., Binz, P.A., Bogue, M., Booth, T., Brazma, A., Brinkman, R.R., Michael Clark, A., Deutsch, E.W., Fiehn, O., Fostel, J., Ghazal, P., Gibson, F., Gray, T., Grimes, G., Hancock, J.M., Hardy, N.W., Hermjakob, H., Julian, R.K. Jr., Kane, M., Kettner, C., Kinsinger, C., Kolker, E., Kuiper, M., Le Novère, N., Leebens-Mack, J., Lewis, S.E., Lord, P., Mallon, A.M., Marthandan, N., Masuya, H., McNally, R., Mehrle, A., Morrison, N., Orchard, S., Quackenbush, J., Reecy, J.M., Robertson, D.G., Rocca-Serra, P., Rodriguez, H., Rosenfelder, H., Santoyo-Lopez, J., Scheuermann, R.H., Schober, D., Smith, B., Snape, J., Stoeckert, C.J. Jr, Tipton, K., Sterk, P., Untergasser, A., Vandesompele, J., and Wiemann, S. 2008. Promoting coherent minimum reporting guidelines for biological and biomedical investigations: The MIBBI project. *Nat. Biotechnol.* 26:889-896.

## Key References

Lee et al., 2008.

*This reference contains the full normative text of the MIFlowCyt specification, which specifies the required minimal annotation for publishing flow cytometry experiments. Reviewing MIFlowCyt will provide readers with additional clarification of what metadata to provide when describing their datasets.*

Kotecha, N., Krutzik, P.O., and Irish, J.M. 2010. Web-based analysis and publication of flow cytometry experiments. *Curr. Protoc. Cytom.* 53:10.17.1-10.17.24.

*This reference describes several protocols to publish, analyze, and share flow cytometry data using Cytobank. Since FlowRepository is built on top of the Cytobank platform, most of the procedures*

described in these protocols are also applicable for FlowRepository. Therefore, this reference provides additional resources and guidelines for FlowRepository users.

### **Internet Resources**

<https://flowrepository.org/>

*This is the home page of FlowRepository, with links for browsing and querying public datasets, logging in, registration, user's guides and recent documentation.*

<http://www.mozilla.org/>

*This is the home page of Mozilla with links to obtain the latest version of Firefox. Firefox is the recommended Web browser for FlowRepository.*

<http://javatester.org/version.html>

*Java Tester is a Web page that helps you verify your ability to run Java applets in your Web browser, as well as the version of Java that your Web browser*

*is using. The ability to run Java version 1.5.1 or higher is a requirement for accessing FlowRepository. Java version 1.6 or higher is recommended.*

<http://www.java.com/>

*Java download page; Windows users should use this page to obtain or update Java if needed.*

<http://openid.net/get-an-openid/>

*An OpenID Web site helps to determine whether you have an OpenID already and suggests options in case you need to obtain one. FlowRepository requires an OpenID in order to deposit and annotate datasets, analyze data online, and perform additional advanced tasks.*

<http://docs.cytobank.org/>

*The documentation site for Cytobank; most of the Cytobank documentation is also applicable for FlowRepository since FlowRepository is built on top of the Cytobank platform.*