Input Module Tutorial #1: Using the *Images* Module

This tutorial describes how to use the **Images** module for preparing images for processing in CellProfiler. Setting this module correctly is the first step in creating a new project in CellProfiler.

The **Images** module specifies the location of image files to be analyzed by your pipeline. These files can be located on your hard drive, on a networked computer elsewhere, or accessible with a URL. You can also provide rules to specify only those files that you want analyzed out of a larger collection (for example, from a folder containing both images for analysis and non-image files that should be disregarded).

For the purposes of this tutorial, you can use a set of files from a translocation assay which are available from http://www.cellprofiler.org/linked_files/TranslocationActivity/TranslocationData.zip.
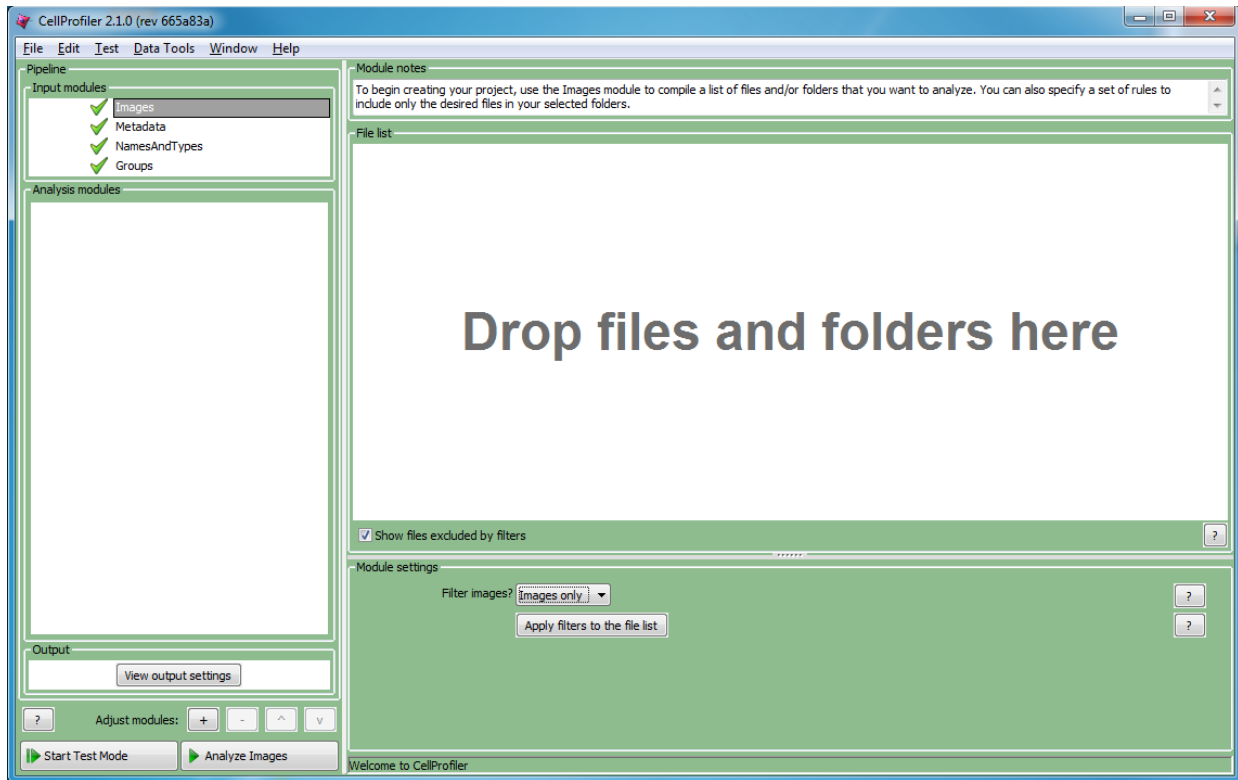
What is a *digital* image?

A digital *image* is a set of numbers arranged into a two-dimensional format of rows and columns; a pixel refers to the row/column location of a particular point in the image. Pixels in grayscale or monochrome (black/white) images contain a single intensity value, whereas in color images, each pixel contains a red, green, and blue (RGB) triplet of intensity values. Additionally, the term *image* can be used as short-hand for an image sequence, that is, an image collection such as a time-lapse series (2-D + *t*), confocal Z-stacks (3-D), etc.

CellProfiler can read a wide variety of image formats by using a library called Bio-Formats; see here for the formats available. Some image formats are better than others for use in image analysis. Some are "lossy" (information is lost in the conversion to the format) like most JPG/JPEG files; others are "lossless" (no image information is lost). For image analysis purposes, a lossless format like TIF or PNG is recommended.
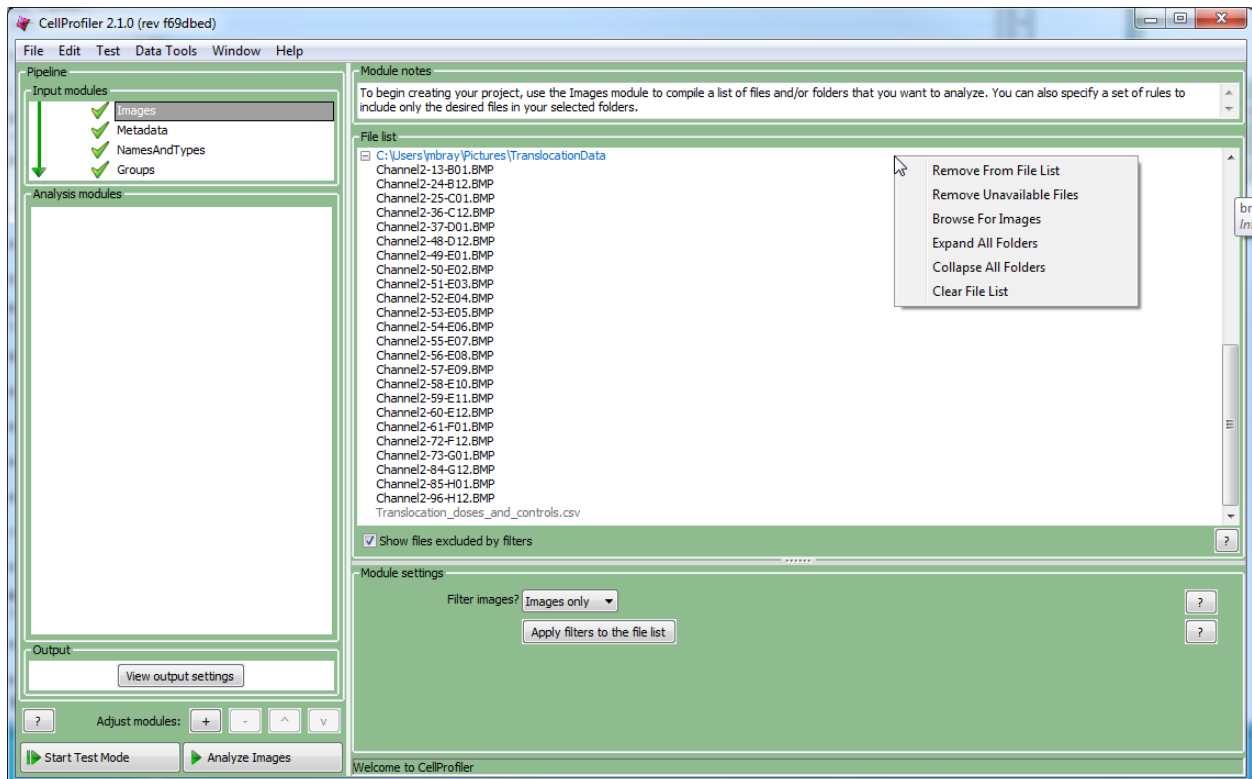
How do I start using this module?

When CellProfiler is opened from the Start Menu (Windows) or from the Applications folder (Mac), the **Images** module is the first module that is selected, as indicated by the highlighted module name in the "Input modules" panel on the upper left. Selecting this module will display the *File list panel*, as a large white space on the right containing the words "Drop files and folders here".

Using the file explorer tool of your choice (e.g., Explorer in Windows, Finder in Mac), you can drag-and-drop individual files and/or entire folders into this panel. You can also right-click in the File list panel to bring up a file selection window to browse for individual files; on the Mac, folders can be drag-and-dropped from this window and you can select multiple files using Ctrl-A (Windows) or Cmd-A (Mac).

Dragging an entire folder into this panel will update the panel similar to the figure below:

This panel shows the files as a list under the containing folder(s). You can expand or collapse the file display by clicking the ⊟ icon to the left of the folder name.

When this panel contains a selection of files/folders, you can perform a variety of actions to refine your list:

- Right-clicking on a file or set of selected files will bring up a context menu (see figure above) which allows you to change the display of the folder list, remove files from the list, clear the entire list, etc.
- Double-clicking on an image will display that image in a window.
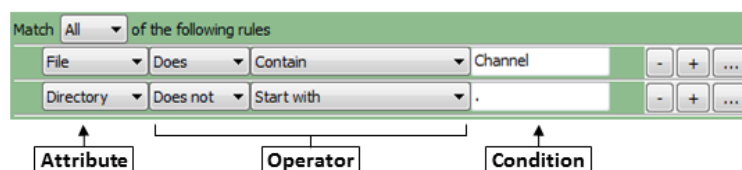
What should I do with this list of files?

In some cases, you may want to remove a subset of files. For example, CellProfiler requires that only image files be processed in a pipeline, so non-image files in your list would need to be excluded. You may also want to remove files from the list if you have dragged a folder of images onto the file list panel that contains the images from one experiment that you want to process along with images from another experiment that you want to ignore for now.

To remove a particular file or folder from the File list panel manually, select it and click the Delete key, or right-click and select the "Remove" option.

Another way to remove files is to filter them out using the module settings below the File list panel.

- The default filtering behavior is to retain only image files. By pressing the "Apply" button underneath the filter, you will see that any non-image files are grayed out (see figure above; the file "Translocation_doses_and_controls.csv" is grayed out).
- To have finer control over the filtering, select "Custom" from the filtering drop-down to display a set of control menus to specify which files should be retained.

Clicking the rule menus shows you all the file *attributes*, *operators* and *conditions* that you can specify to narrow down the image list.

1. For each rule, first select the *attribute* from the drop-down menu that the rule is to be based on. For example, you can select "File" to define a rule that will filter files on the basis of their filename.
2. The *operator* drop-down menu is then updated with operators applicable to the attribute you selected. For example, if you select "File" as the attribute, the operator menu includes text operators such as "Contain" or "Starts with". On the other hand, if you select "Extension" as the attribute, you can choose the logical operators "Is" or "Is not" from the menu.
3. In the *operator* menu, select the operator you want to use. For example, if you want to match data exactly, you may want the "Exactly match" or the "Is" operator. If you want the condition to be less strict and match anywhere in the attribute, select an operator such as "Contains", "Start with" or "End with".
4. Use the *condition* box to type the condition you want to match. The more you type, the more specific the condition is.

As an example, if you create a new filter and select "File" as the attribute, then select "Does" and "Contain" as the operators, and type "Channel" as the condition, the filter finds all files that include the text "Channel", such as "Channel1.tif" "Channel2.jpg", "1-Channel-A01.BMP" and so on.

If you select "Does" and "Start with" as the operators and "Channel1" in the Condition box, the rule will includes such files as "Channel1.tif" "Channel1-A01.png", and so on.

You may specify as many rules as necessary to define the desired list of images. At any point, you can click the "Apply" button to see the effect of your rules to the files in the File list panel.

What should I have when I'm done with this module?

The final product of the **Images** module is a file list in which any files that are not intended for further processing have been removed, whether manually or using filtering. This list will be used when collecting metadata (if desired) and when assembling the image sets in **NamesAndTypes**. The list can be filtered further in **NamesAndTypes** to specify, for example, that a subset of these images represents a particular wavelength.