The LumiR User Guide

Renan Sauteraud*

February 22, 2013

Contents

1 Introduction

1.1 Technology

Luminex xMAP technology is a multiplex assay using flow cytometry to concurrently measure up to 500 analytes in a single reaction volume.

---

*rsautera@fhcrc.org

1.2 LumiR

`LumiR` is an R package that provides data structure and functions to read, store and analyze Luminex xMAP experiment.

As with any R package, it should first be loaded in the session

```
> library(LumiR)
```

2 Softwares

Even though Luminex is the company delivering the beads, a wide range of partners offer customizable kits and solutions for acquisition and/or analysis of the data. While the technology is identical, the output depends greatly on the software used for the acquisition. In its current version, `LumiR` can read the data from three different vendors:

2.1 Luminex

Luminex's own acquisition software xPONENT produces one csv file per well. These files contain the raw bead level information: bead id, fluorescence measured as well as the fluorescences used to map the bead. `LumiR` can read data from xPONENT versions 1.x and 3.x.

2.2 MiraiBio

MiraiBio also developped a set of tools for the acquisition and analysis of Luminex xMAP platform. The acquisition software 'MasterPlex CT' currently in version 1.2 creates one file per well. These binary files with a .lxb extension are based on the format FCS 3.0 and contain the bead level information. Along with these files comes a summary in xml format with a .lxd extension. It contains some information regarding the setup used for the experiment, the matching of the bead id with the analyte name and some basic calculations such as the MFI or the bead count in each well for each analyte.

2.3 BIO-RAD

Bio-Plex Manager Software yields files with a proprietary format. However, it allows the user to exports the results in XML files that can be processed by `LumiR`. The output is a single XML file per experiment gathering all informations available.

3 Requirements

In order to read the data into R, the package require the data to be organized in a specific file tree. Additional user-provided information is necessary to run most of the functions in this package.

3.1 Folders architecture

Regardless of which software produced the data, the required organization of the folder remains the same. The experiment should be located in a folder (root) with the raw bead level files for each plate in a subfolder named after the plate name or ID. The mapping files should be placed at the root.

3.2 Analyte mapping file

In the raw data, the beads are referenced using a unique ID that is specific to a kit and a vendor. In order to use the actual analyte name, the package will look for a file 'analyte.csv' that will map each ID to a name. It should contain two columns **analyte** and **bid**. If this file is not submitted by the user, LumiR will still be able to read the data but will display the bead ID instead of the actual name. Alternately, with MasterPlex CT data, the reading function will look for the .lxd file and attempt to extract this information.

3.3 Phenotype mapping file

This user provided file should contain information to match a sample (defined by a 'plate', a 'filename' and a 'well') and its phenotype information, including a 'group_name' a 'sample_type' (standard, case or control) as well as complementary details like standards 'expected_concentration', or cases matching 'control_idx'. The package can still read the data and format it into a blum. However, this information is critical and the user will be asked to add it prior to any analysis such as standard curve fitting.

    The minimum requirements for the mapping file are the three columns **plate**, **filename** and **well**. If no file is provided, the reading function will try to guess these informations based on the structure of the experiment folder. Additional information regarding the phenotype can be passed in the form of additional columns and should be provided in order to run any analysis such as standard curve fitting. It is recommended to include the following columns:

A **group_name** and a **sample_type** (standard, case or control) as well as complementary details like the standards expected **concentration**, or the case's matching **control_idx**.

    Additional optional info but if not provided, no analysis can be done. If nothing is provided, we can still guess

An example of each mapping file is provided in the inst folder.

```
> path <- system.file("extdata", package = "LumiR")
> analyte <- read.csv(file.path(path, "analyte.csv"))
> head(analyte)

  analyte bid
1  CHEX.1  97
2  CHEX.2  98
3  CHEX.3  99
4  CHEX.4 100
5 EOTAXIN  52
6   G.CSF  46

> phenotype <- read.csv(file.path(path, "phenotype.csv"))
> head(phenotype, 2)

   plate well sample_type         group_name sample_name   filename
1 plate1   A1     control       wt_10^3_24h      mouse1 Run001.csv
2 plate1   B1     unknown L.p.FlaA_10^3_24h      mouse1 Run002.csv
  concentration control_idx
1             1           2
2             1           2
```

4 Data structure

`LumiR` uses two classes to store different level of information.

4.1 blum

`blum` objects store the bead level information: bead id, the two fluorescence levels used for the bead characterization and the measured fluorescence for the analyte.

The reading function `read.experiment` only argument path takes the pathname of the root folder of an experiment and returns a `blum` object.

4.2 bsum

`bSummarize` is used to create a `bsum`. It calculates the MFI associated with each analyte in each well and uses the given formula to fit the standard curves based on the standard wells.

5 Setters

At some points (either setter on blum/bsum and/or ags in bSummarize) the user will have to add the potentialy missing phenotype information.

6 Plate layout

The function `plot_layout` makes use of the package `ggplot2` to display a representation of a plate stored in a `blum` or `bsum` object.

7 Standard curve fitting

One of the most common tools for analysis of xMAP data is the Standard Curve. While this package offers new tools for finding significant difference in expression levels between samples, it also implements a standard curve fitting method.

`geom_sc` is a method that extends `ggplot2`'s `geom_line` method. Therefore, it should be added Here we plot the curves for each analyte of the experiment. As seen, multiple plates can be shown at the same time.

8 SAxCyB

Saxcyblablabalbal