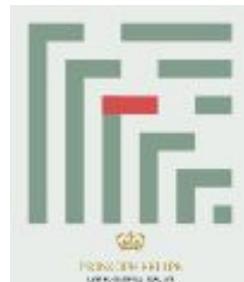


# Introduction to Babelomics 5

Máster en Bioinformática Aplicada a Medicina Personalizada y Salud

Francisco García García, fgarcia@cipf.es





# Outline

1. Introduction to Babelomics
2. Functional Profiling: over-representation methods
3. Unsupervised Classification: Clustering
4. Supervised Classification: Prediction

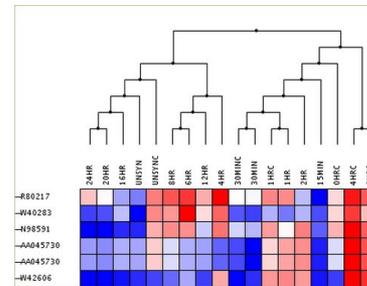
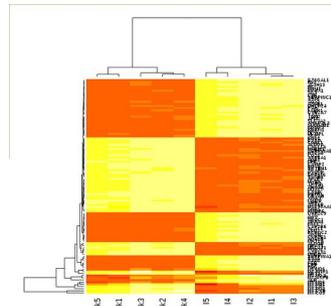
# Outline

Design

Processing

- Differential expression
- Clustering
- Predictors

Functional profiling



Any questions?

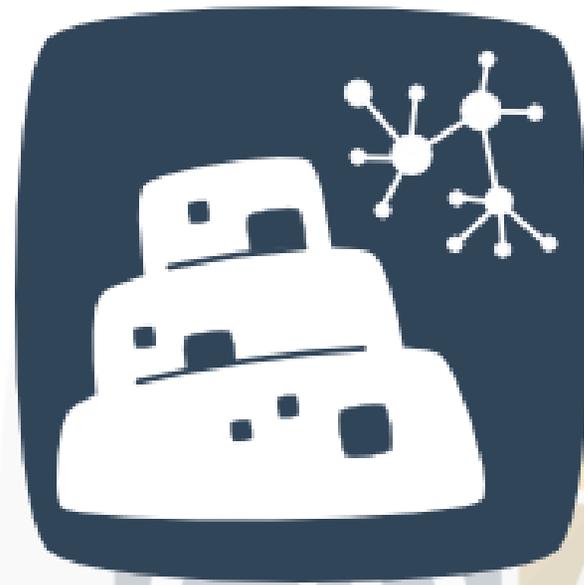


# Babelomics 5

## Babelomics 5

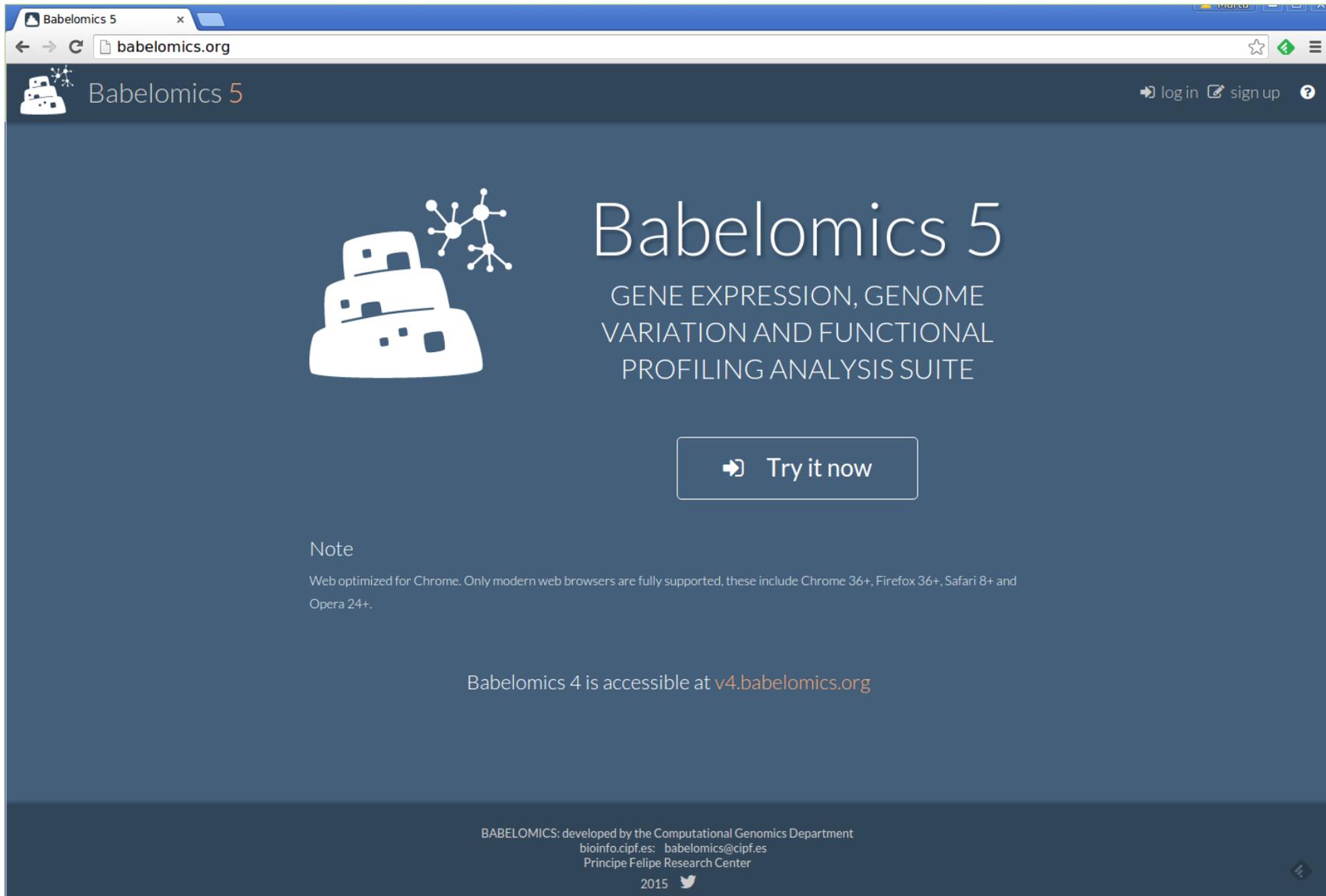
Integrative open source platform with advanced functional profiling for the analysis of

- Transcriptomics
- Proteomics
- Genomics



<http://www.babelomics.org>

# Babelomics 5



The screenshot shows the homepage of Babelomics 5. The browser address bar displays 'babelomics.org'. The page features a dark blue background with a white icon of a microarray chip and a network diagram. The main heading is 'Babelomics 5' followed by the subtitle 'GENE EXPRESSION, GENOME VARIATION AND FUNCTIONAL PROFILING ANALYSIS SUITE'. A prominent 'Try it now' button is centered below the text. A 'Note' section provides browser compatibility information. At the bottom, there is a footer with contact details and the year 2015.

Babelomics 5

log in sign up



# Babelomics 5

GENE EXPRESSION, GENOME VARIATION AND FUNCTIONAL PROFILING ANALYSIS SUITE

Try it now

Note

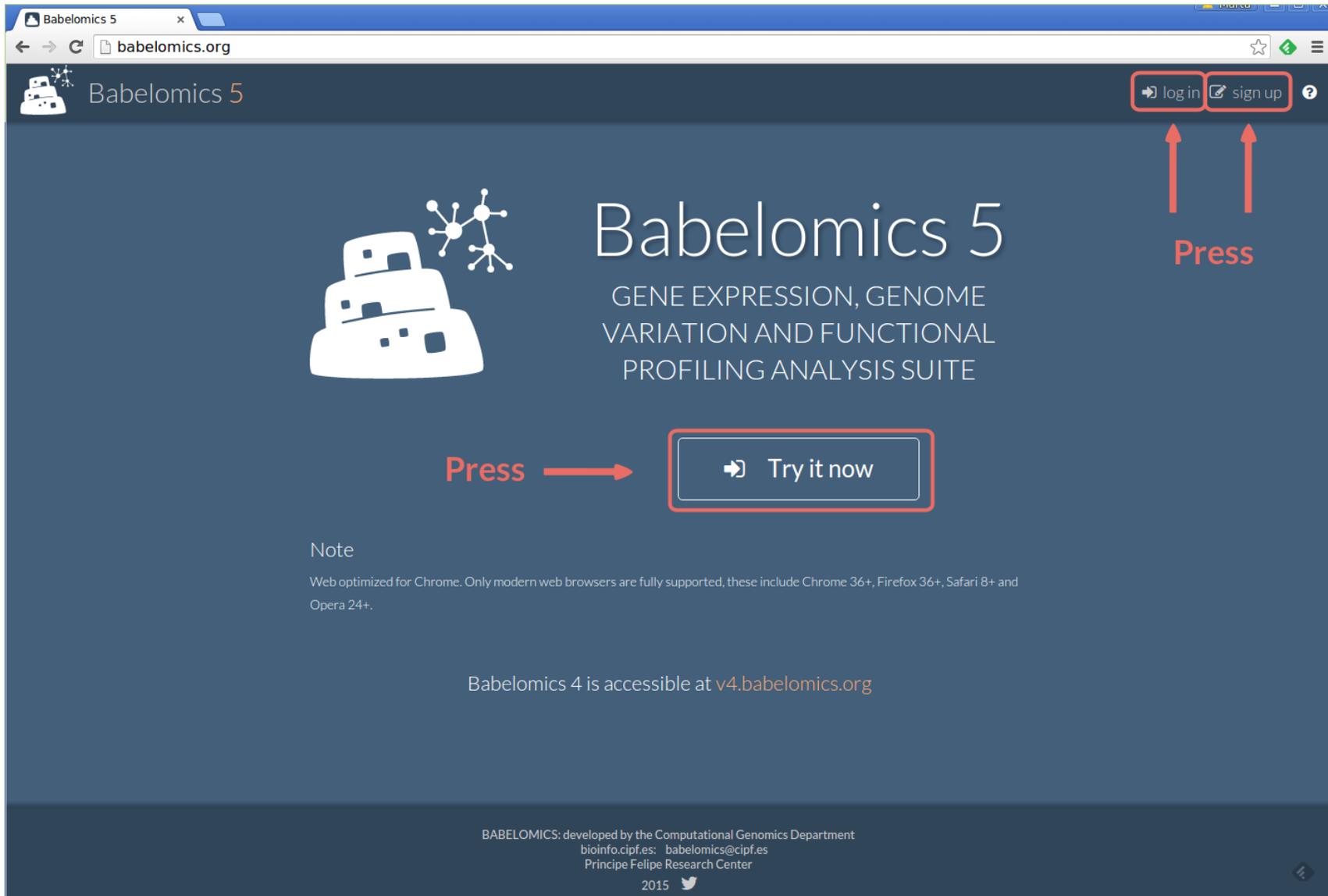
Web optimized for Chrome. Only modern web browsers are fully supported, these include Chrome 36+, Firefox 36+, Safari 8+ and Opera 24+.

Babelomics 4 is accessible at [v4.babelomics.org](http://v4.babelomics.org)

BABELOMICS: developed by the Computational Genomics Department  
bioinfo.cipf.es: babelomics@cipf.es  
Principe Felipe Research Center

2015

# Logging in



The screenshot shows the Babelomics 5 website interface. The browser address bar displays 'babelomics.org'. In the top right corner, the 'log in' and 'sign up' buttons are highlighted with red boxes, and two red arrows point to them from the word 'Press' below. In the center of the page, the 'Try it now' button is also highlighted with a red box, with a red arrow pointing to it from the word 'Press' to its left. The main content area features a stylized white icon of a microarray chip and a network diagram, followed by the text 'Babelomics 5' and 'GENE EXPRESSION, GENOME VARIATION AND FUNCTIONAL PROFILING ANALYSIS SUITE'. Below this, a 'Note' section states: 'Web optimized for Chrome. Only modern web browsers are fully supported, these include Chrome 36+, Firefox 36+, Safari 8+ and Opera 24+'. At the bottom, it mentions 'Babelomics 4 is accessible at [v4.babelomics.org](http://v4.babelomics.org)' and provides contact information for the Computational Genomics Department at the Principe Felipe Research Center, dated 2015.

# Logging in

- **Try it now:** Enter as anonymous user
- **Sign up:** Create an identified user
- **Log in:** Log in as identified user



# Logging in

- **Try it now:** Enter as anonymous user
- **Sign up:** Create an identified user
- **Log in:** Log in as identified user

## Anonymous user

- Can access all functionalities and tools
- Data will be lost

## Identified user

- Can access all functionalities and tools
- 10Gb of free disk

# Logging in

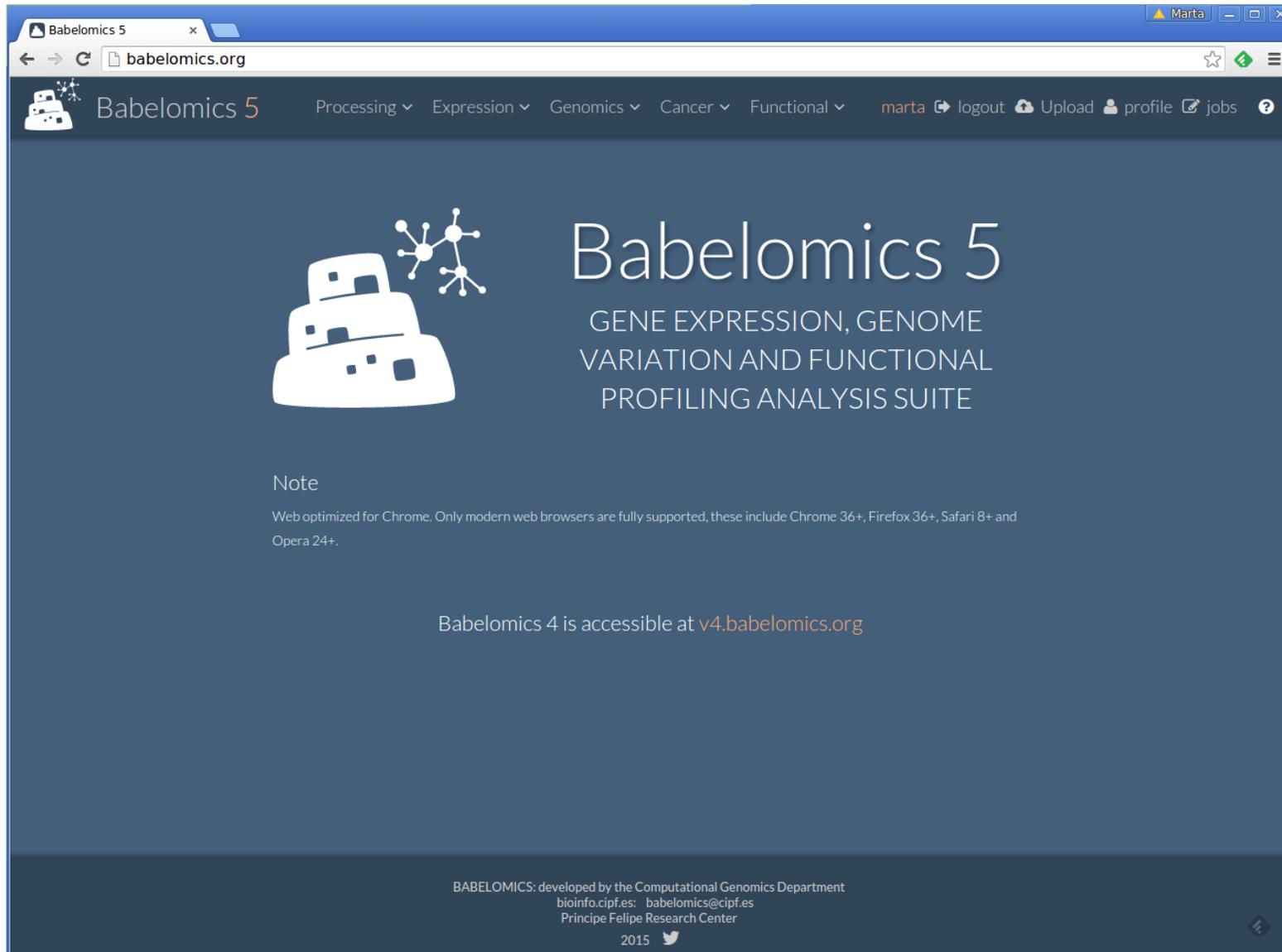
## Exercise

Create a new account for yourself

- 1 Go to [Babelomics 5](#)
- 2 Create a new user

For help, ask or visit the [logging in tutorial](#)

# Web structure



The screenshot shows the homepage of Babelomics 5. The browser address bar displays 'babelomics.org'. The navigation menu includes 'Processing', 'Expression', 'Genomics', 'Cancer', and 'Functional', along with user options for 'marta', 'logout', 'Upload', 'profile', and 'jobs'. The main content area features a stylized icon of a microarray chip and a network diagram, followed by the title 'Babelomics 5' and the subtitle 'GENE EXPRESSION, GENOME VARIATION AND FUNCTIONAL PROFILING ANALYSIS SUITE'. A 'Note' section states: 'Web optimized for Chrome. Only modern web browsers are fully supported, these include Chrome 36+, Firefox 36+, Safari 8+ and Opera 24+'. A link indicates 'Babelomics 4 is accessible at [v4.babelomics.org](http://v4.babelomics.org)'. The footer contains the text: 'BABELOMICS: developed by the Computational Genomics Department bioinfo.cipf.es: babelomics@cipf.es Principe Felipe Research Center 2015'.

# Web structure

The screenshot shows a web browser window with the address bar displaying 'babelomics.org'. The page header features a navigation menu with the following items: 'Babelomics 5' (highlighted with a red box and an arrow pointing to the text 'Home button'), 'Processing', 'Expression', 'Genomics', 'Cancer', 'Functional', 'marta', 'logout', 'Upload', 'profile', and 'jobs'. The main content area includes a large white icon of a microarray chip, the title 'Babelomics 5', and the subtitle 'GENE EXPRESSION, GENOME VARIATION AND FUNCTIONAL PROFILING ANALYSIS SUITE'. A 'Note' section below states: 'Web optimized for Chrome. Only modern web browsers are fully supported, these include Chrome 36+, Firefox 36+, Safari 8+ and Opera 24+'. A link is provided: 'Babelomics 4 is accessible at [v4.babelomics.org](http://v4.babelomics.org)'. The footer contains the text: 'BABELOMICS: developed by the Computational Genomics Department', 'bioinfo.cipf.es: babelomics@cipf.es', 'Principe Felipe Research Center', and '2015' with a Twitter icon.

# Web structure

The screenshot shows the Babelomics 5 website in a browser window. The browser's address bar displays 'babelomics.org'. The website's header features a logo on the left, a navigation menu with dropdowns for 'Processing', 'Expression', 'Genomics', 'Cancer', and 'Functional', and user options for 'marta', 'logout', 'Upload', 'profile', and 'jobs'. A red box highlights the logo and navigation menu, with an arrow pointing to the logo labeled 'Home button'. The main content area includes the heading 'Analysis tools', the title 'Babelomics 5', and the subtitle 'GENE EXPRESSION, GENOME VARIATION AND FUNCTIONAL PROFILING ANALYSIS SUITE'. Below this is a 'Note' section stating: 'Web optimized for Chrome. Only modern web browsers are fully supported, these include Chrome 36+, Firefox 36+, Safari 8+ and Opera 24+'. A link is provided: 'Babelomics 4 is accessible at [v4.babelomics.org](http://v4.babelomics.org)'. The footer contains the text: 'BABELOMICS: developed by the Computational Genomics Department bioinfo.cipf.es: babelomics@cipf.es Principe Felipe Research Center 2015'.

# Web structure

The screenshot shows the Babelomics 5 web interface. The browser address bar displays 'babelomics.org'. The navigation bar includes a 'Babelomics 5' logo, a menu with 'Processing', 'Expression', 'Genomics', 'Cancer', and 'Functional', and a user profile section for 'marta' with a 'logout' link. A red arrow points to the logo with the label 'Home button'. The main content area features the text 'Analysis tools' and 'Logout' in red, the title 'Babelomics 5', and the subtitle 'GENE EXPRESSION, GENOME VARIATION AND FUNCTIONAL PROFILING ANALYSIS SUITE'. A 'Note' section provides browser compatibility information. At the bottom, it states 'Babelomics 4 is accessible at [v4.babelomics.org](http://v4.babelomics.org)' and includes contact information for the Computational Genomics Department at the Principe Felipe Research Center, dated 2015.

Babelomics 5

Processing ▾ Expression ▾ Genomics ▾ Cancer ▾ Functional ▾

marta ↗ logout

Upload profile jobs

Home button

Analysis tools Logout

# Babelomics 5

GENE EXPRESSION, GENOME VARIATION AND FUNCTIONAL PROFILING ANALYSIS SUITE

Note

Web optimized for Chrome. Only modern web browsers are fully supported, these include Chrome 36+, Firefox 36+, Safari 8+ and Opera 24+.

Babelomics 4 is accessible at [v4.babelomics.org](http://v4.babelomics.org)

BABELOMICS: developed by the Computational Genomics Department  
bioinfo.cipf.es: babelomics@cipf.es  
Principe Felipe Research Center

2015

# Web structure

The screenshot shows the Babelomics 5 web interface. The browser address bar displays 'babelomics.org'. The navigation bar includes a 'Babelomics 5' logo, a menu with 'Processing', 'Expression', 'Genomics', 'Cancer', and 'Functional', and user options for 'marta', 'logout', 'Upload', 'profile', and 'jobs'. Red boxes highlight the logo, the menu, and the 'Logout' and 'Upload' buttons. Red arrows point from the text 'Home button', 'Analysis tools', 'Logout', and 'Upload' to their respective elements on the page. The main content area features a server rack icon, the title 'Babelomics 5', and the subtitle 'GENE EXPRESSION, GENOME VARIATION AND FUNCTIONAL PROFILING ANALYSIS SUITE'. A 'Note' section provides browser compatibility information, and a link for 'Babelomics 4' is provided at the bottom. The footer contains development credits and the year 2015.

Babelomics 5

Processing ▾ Expression ▾ Genomics ▾ Cancer ▾ Functional ▾

marta → logout Upload profile jobs ?

Home button

Analysis tools

Logout Upload

# Babelomics 5

GENE EXPRESSION, GENOME VARIATION AND FUNCTIONAL PROFILING ANALYSIS SUITE

Note

Web optimized for Chrome. Only modern web browsers are fully supported, these include Chrome 36+, Firefox 36+, Safari 8+ and Opera 24+.

Babelomics 4 is accessible at [v4.babelomics.org](http://v4.babelomics.org)

BABELOMICS: developed by the Computational Genomics Department  
bioinfo.cipf.es: babelomics@cipf.es  
Principe Felipe Research Center

2015

# Web structure

The screenshot shows the Babelomics 5 web interface. The browser address bar displays 'babelomics.org'. The navigation bar includes a 'Babelomics 5' logo, a menu with 'Processing', 'Expression', 'Genomics', 'Cancer', and 'Functional', and user options for 'marta', 'logout', 'Upload', 'profile', and 'jobs'. Red arrows point from the following text labels to their corresponding elements in the interface:

- Home button** points to the Babelomics 5 logo.
- Analysis tools** points to the main menu.
- Logout** points to the 'logout' link.
- Upload** points to the 'Upload' link.
- Settings** points to the 'profile' link.

The main content area features the Babelomics 5 logo (a stylized tower) and the text: 'Babelomics 5', 'GENE EXPRESSION, GENOME VARIATION AND FUNCTIONAL PROFILING ANALYSIS SUITE'. Below this is a 'Note' section stating: 'Web optimized for Chrome. Only modern web browsers are fully supported, these include Chrome 36+, Firefox 36+, Safari 8+ and Opera 24+'. A link is provided: 'Babelomics 4 is accessible at [v4.babelomics.org](http://v4.babelomics.org)'. The footer contains: 'BABELOMICS: developed by the Computational Genomics Department', 'bioinfo.cipf.es: babelomics@cipf.es', 'Principe Felipe Research Center', '2015', and a Twitter icon.

# Web structure

The screenshot shows the Babelomics 5 web interface. The browser address bar displays 'babelomics.org'. The navigation bar includes a 'Babelomics 5' logo, a menu with 'Processing', 'Expression', 'Genomics', 'Cancer', and 'Functional', and user options for 'marta', 'logout', 'Upload', 'profile', and 'jobs'. Red arrows point from text labels to these elements: 'Home button' to the logo, 'Analysis tools' to the menu, 'Logout' to the 'logout' link, 'Upload' to the 'Upload' link, 'Settings' to the 'profile' link, and 'Jobs panel' to the 'jobs' link. The main content area features the Babelomics 5 logo, the title 'Babelomics 5', and the subtitle 'GENE EXPRESSION, GENOME VARIATION AND FUNCTIONAL PROFILING ANALYSIS SUITE'. A 'Note' section provides browser compatibility information. A footer contains development details and the year 2015.

Babelomics 5

Processing ▾ Expression ▾ Genomics ▾ Cancer ▾ Functional ▾

marta logout Upload profile jobs

Home button

Analysis tools

Logout

Upload

Settings

Jobs panel

# Babelomics 5

GENE EXPRESSION, GENOME VARIATION AND FUNCTIONAL PROFILING ANALYSIS SUITE

Note

Web optimized for Chrome. Only modern web browsers are fully supported, these include Chrome 36+, Firefox 36+, Safari 8+ and Opera 24+.

Babelomics 4 is accessible at [v4.babelomics.org](http://v4.babelomics.org)

BABELOMICS: developed by the Computational Genomics Department  
bioinfo.cipf.es: babelomics@cipf.es  
Principe Felipe Research Center

2015

# Web structure

The screenshot shows the Babelomics 5 web interface. The browser address bar displays 'babelomics.org'. The navigation bar includes a 'Babelomics 5' logo, a menu with 'Processing', 'Expression', 'Genomics', 'Cancer', and 'Functional', and user options for 'marta', 'logout', 'Upload', 'profile', 'jobs', and a help icon. Red arrows point from text labels to these elements: 'Home button' to the logo, 'Analysis tools' to the menu, 'Logout' to the 'logout' link, 'Upload' to the 'Upload' link, 'Settings' to the 'profile' link, 'Jobs panel' to the 'jobs' link, and 'Help' to the help icon. The main content area features a server rack icon, the title 'Babelomics 5', and the subtitle 'GENE EXPRESSION, GENOME VARIATION AND FUNCTIONAL PROFILING ANALYSIS SUITE'. A 'Note' section provides browser compatibility information. A footer contains development credits and the year 2015.

Babelomics 5

Processing ▾ Expression ▾ Genomics ▾ Cancer ▾ Functional ▾

marta logout Upload profile jobs ?

Home button

Analysis tools

Logout

Upload

Settings

Jobs panel

Help

# Babelomics 5

GENE EXPRESSION, GENOME VARIATION AND FUNCTIONAL PROFILING ANALYSIS SUITE

Note

Web optimized for Chrome. Only modern web browsers are fully supported, these include Chrome 36+, Firefox 36+, Safari 8+ and Opera 24+.

Babelomics 4 is accessible at [v4.babelomics.org](http://v4.babelomics.org)

BABELOMICS: developed by the Computational Genomics Department  
bioinfo.cipf.es: babelomics@cipf.es  
Principe Felipe Research Center

2015

# Workflow

The screenshot displays the Babelomics 5.0 web interface with several workflow steps highlighted by numbered callouts:

- 1:** User profile and navigation menu (Marta, log out, Upload, jobs).
- 2:** Main navigation menu (Processing, Expression, Genomics, Cancer, Functional).
- 3:** Job type selection dropdown (Single enrichment, Gene set enrichment, Network enrichment, Gene set network enrichment).
- 4:** Selected job type: Gene set network enrichment.
- 5:** Examples section (Normalization example).
- 6:** Launch job button.
- 7:** Jobs list header (jobs < 1 of 7 >).
- 8:** Job list items (Differential expression - Class comparison demo (correlation.txt) class-comparison).
- 9:** Job details view (Differential expression - Class comparison demo (correlation.txt) class-comparison).

The main workflow area includes the following sections:

- Select your data:** File browser (default/)
- Select gene length file:** File browser (default/)
- Normalization method:** Choose automatically the normalization method (selected), Choose manually the normalization method (TMM, RPKM)
- Job information:** Output folder (File browser, default/), Job name (JobName), Description (Job info...)

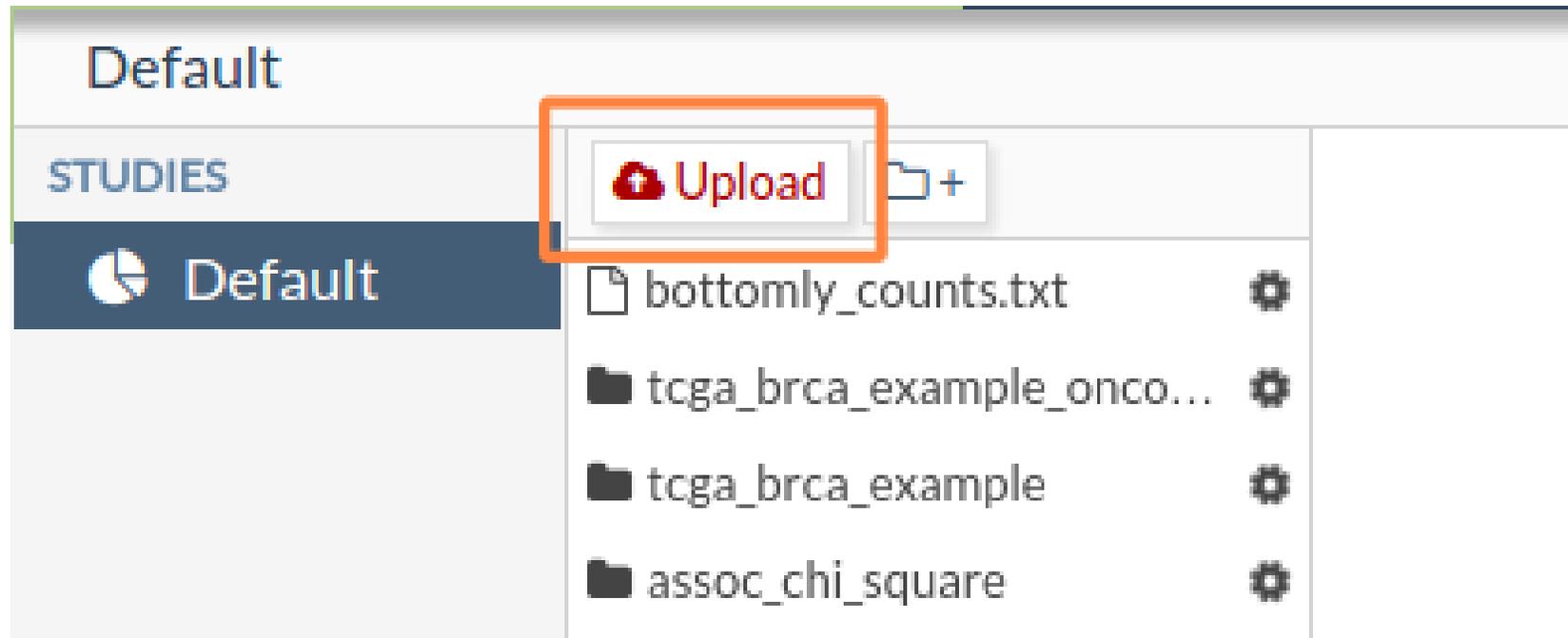
# Uploading data

## Uploading data tutorial



# Uploading data

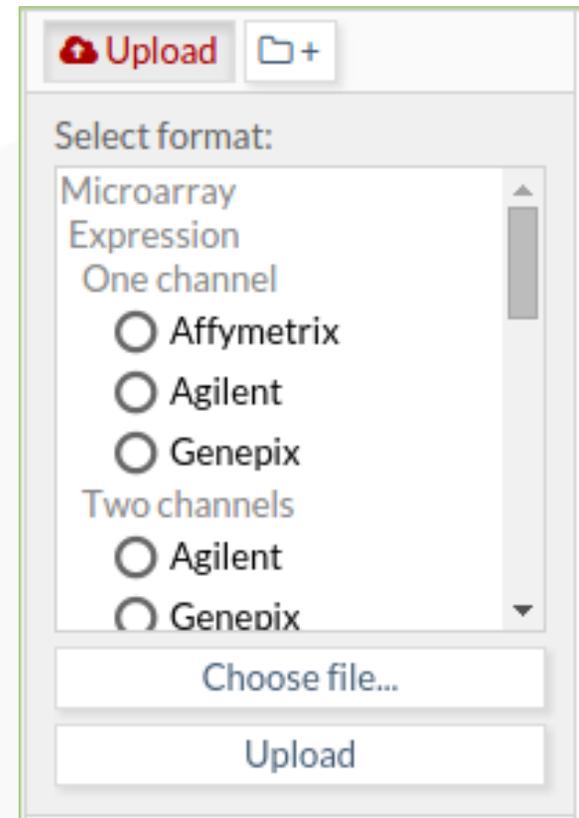
## Uploading data tutorial



# Uploading data

## Uploading data tutorial

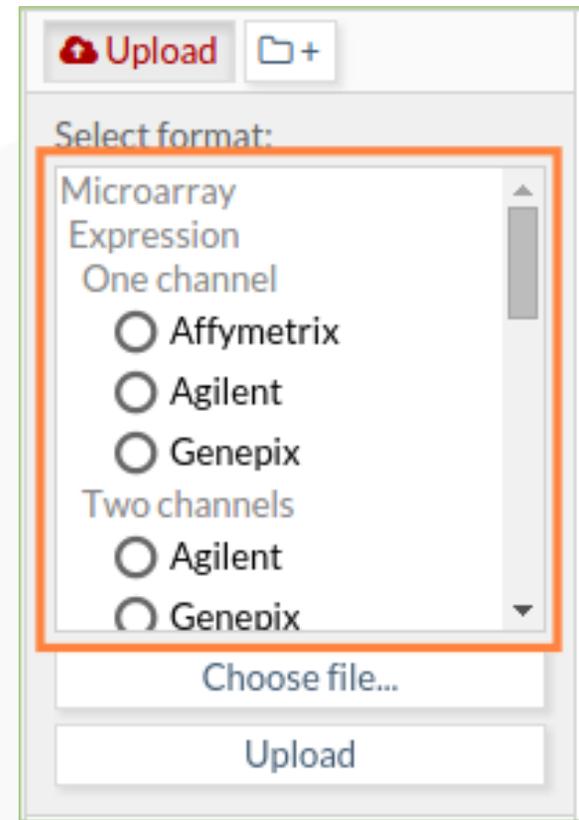
- Bioformat, [more info](#)
  - RNA-Seq: Data Matrix
- **IMPORTANT:** First row with rownames must begin with #NAMES
- We can remove files



# Uploading data

## Uploading data tutorial

- Bioformat, [more info](#)
  - RNA-Seq: Data Matrix
- **IMPORTANT:** First row with rownames must begin with #NAMES
- We can remove files

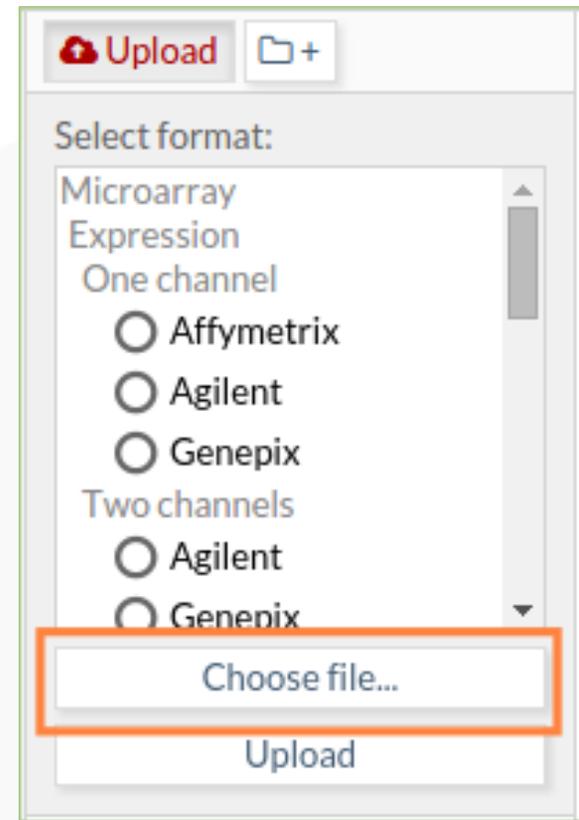


The screenshot shows a web interface for uploading data. At the top, there is an "Upload" button with a red cloud icon and a folder icon with a plus sign. Below this is a "Select format:" section. The options are grouped into "Microarray Expression" and "Two channels". Under "Microarray Expression", there are three radio button options: "Affymetrix", "Agilent", and "Genepix". Under "Two channels", there are two radio button options: "Agilent" and "Genepix". A red rectangular box highlights the "Microarray Expression" section. Below the format selection, there is a "Choose file..." button and an "Upload" button.

# Uploading data

## Uploading data tutorial

- Bioformat, [more info](#)
  - RNA-Seq: Data Matrix
- **IMPORTANT:** First row with rownames must begin with #NAMES
- We can remove files



# Uploading data

## Uploading data tutorial

- Bioformat, [more info](#)
  - RNA-Seq: Data Matrix
- **IMPORTANT:** First row with rownames must begin with #NAMES
- We can remove files

Upload

Select format:

Microarray Expression One channel

Affymetrix

Agilent

Genepix

Two channels

Agilent

Genepix

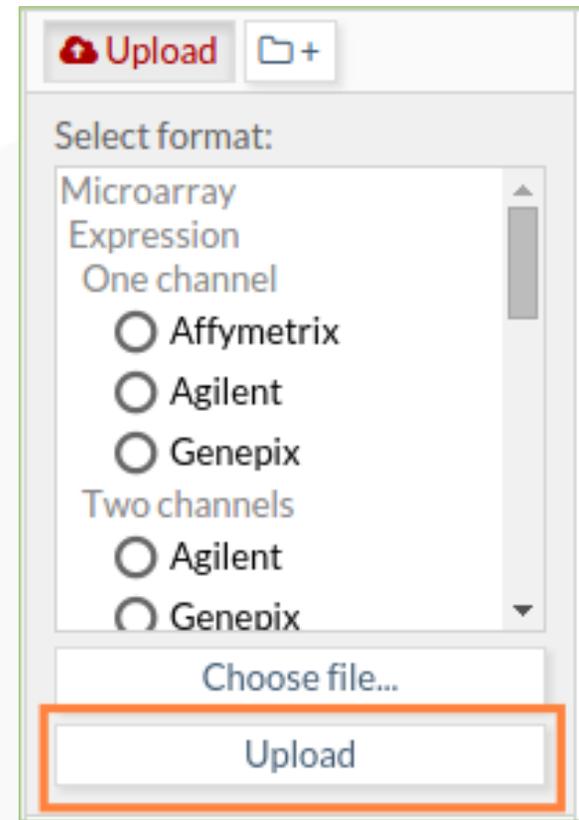
Choose file...

Upload

# Uploading data

## Uploading data tutorial

- Bioformat, [more info](#)
  - RNA-Seq: Data Matrix
- **IMPORTANT:** First row with rownames must begin with #NAMES
- We can remove files



# Uploading data

## Exercise

We will work with a breast cancer dataset from The Cancer Genome Atlas (TCGA). The expression matrix contains raw counts

- 1 Go to the [GDA 2016 wiki](#)
- 2 Download the *brca\_demo\_counts\_4babelomics.txt*
- 3 Upload this file to Babelomics 5

For help, ask or visit the [uploading data tutorial](#)

# Editing data

 Babelomics 5 Processing ▾ Expression ▾ Genomics ▾ Cancer ▾ Functional ▾ anonymous [logout](#) [Upload](#) [jobs](#) 





# Editing data

## Select your data

The files must be on the server to select them.

You can upload files using the button  inside file browser.

File browser

default/

# Editing data

**Edit attributes**

Add new attribute

Categorical ▼

Attribute name  +

Delete attribute

#NAMES ▼ ✖

#NAMES
A_1.bam
A_2.bam
B_1.bam
B_2.bam

# Editing data

**Edit attributes**

Add new attribute

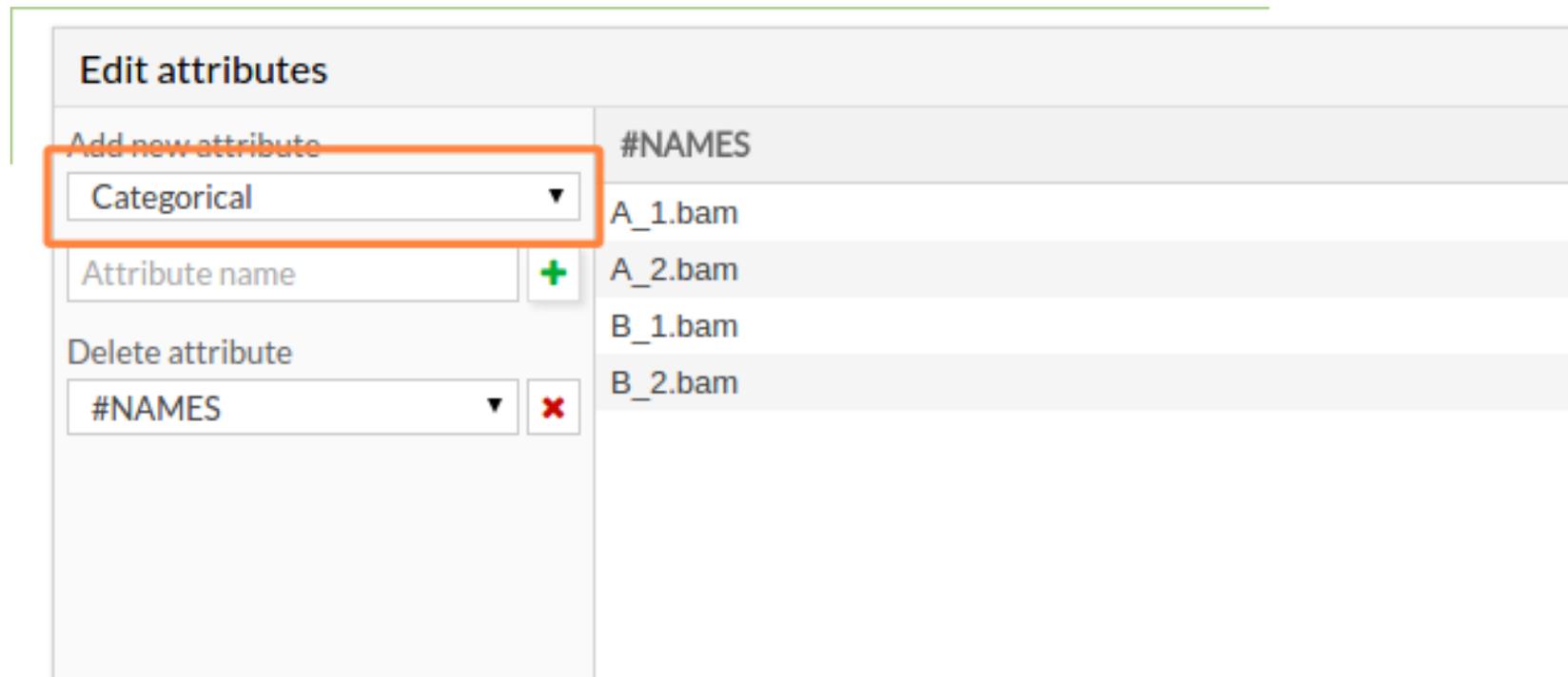
Categorical

Attribute name  +

Delete attribute

#NAMES  ✖

#NAMES
A_1.bam
A_2.bam
B_1.bam
B_2.bam



# Editing data

**Edit attributes**

Add new attribute

Categorical ▼

Attribute name +

Delete attribute

#NAMES ✖

#NAMES
A_1.bam
A_2.bam
B_1.bam
B_2.bam

# Editing data

**Edit attributes**

Add new attribute

Categorical ▼

group +

Delete attribute

#NAMES ✖

#NAMES
A_1.bam
A_2.bam
B_1.bam
B_2.bam

# Editing data

**Edit attributes**

Add new attribute  
Categorical

Attribute name  +

Delete attribute  
#NAMES

#NAMES	group
A_1.bam	
A_2.bam	
B_1.bam	
B_2.bam	

# Editing data

**Edit attributes**

Add new attribute  
Categorical

Attribute name  +

Delete attribute  
#NAMES ✖

#NAMES	group
A_1.bam	0
A_2.bam	0
B_1.bam	1
B_2.bam	



# Editing data

**Edit attributes**

Add new attribute

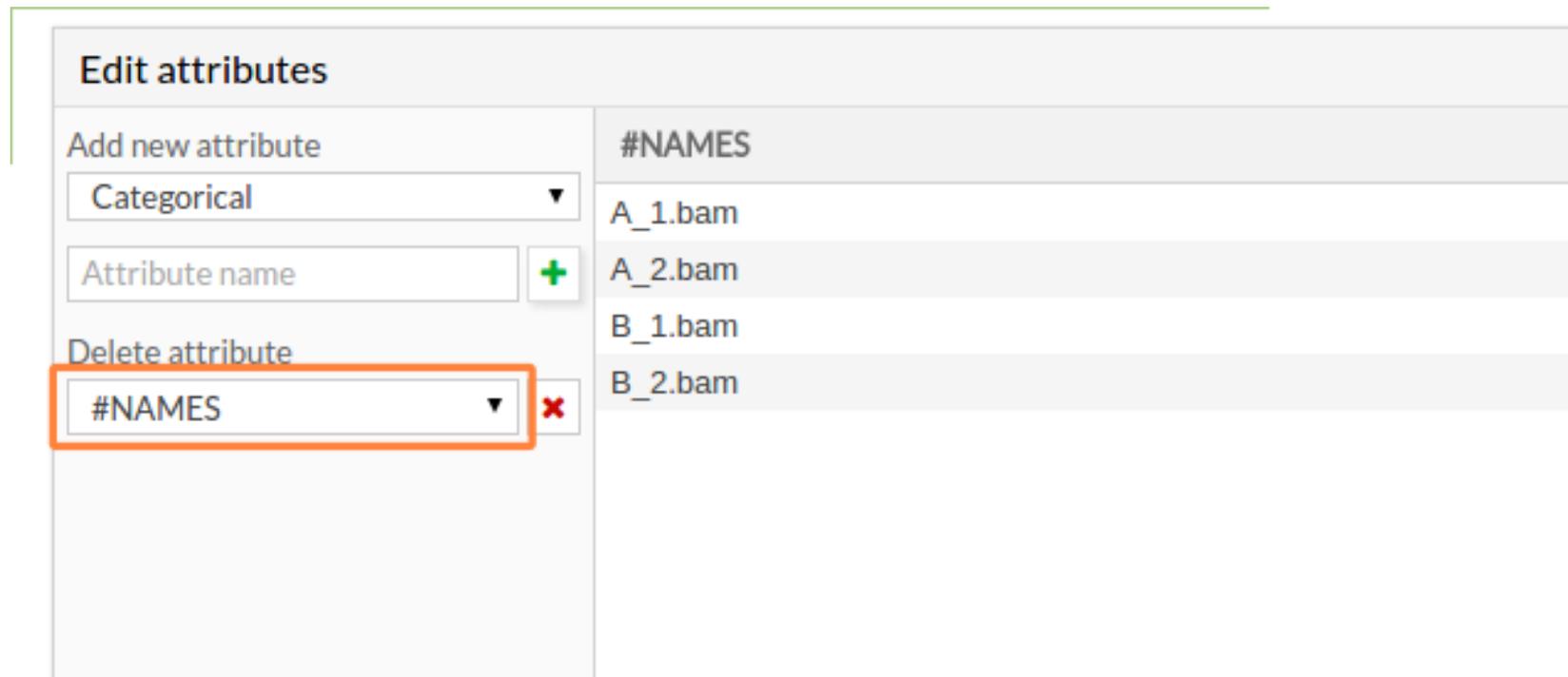
Categorical ▼

Attribute name  +

Delete attribute

#NAMES ▼ ✕

#NAMES
A_1.bam
A_2.bam
B_1.bam
B_2.bam

The image shows a software interface for editing data attributes. On the left, there is a control panel titled "Edit attributes". It includes a section for "Add new attribute" with a dropdown menu currently set to "Categorical" and a text input field for "Attribute name" with a green plus icon. Below that is a "Delete attribute" section with a dropdown menu showing "#NAMES" (highlighted with an orange border) and a red minus icon. To the right of this panel is a list of files under the heading "#NAMES". The list contains four entries: "A\_1.bam", "A\_2.bam", "B\_1.bam", and "B\_2.bam". The background of the interface is light gray with a faint keyboard graphic.



# Editing data

**Edit attributes**

Add new attribute  
Categorical

Attribute name  +

Delete attribute  
#NAMES ✖

#NAMES
A_1.bam
A_2.bam
B_1.bam
B_2.bam

4s < 1 of 1 >

Save attributes Reset changes

# Editing data

## Exercise

The dataset of breast cancer contains healthy and tumor samples. Annotate to which class belongs each sample

- 1 Go to the [GDA 2016 wiki](#)
- 2 Download the *BRCA\_Normal-Basal\_ED.txt*
- 3 Edit the former file with the data in this experimental design

For help, ask or visit the [editing data tutorial](#)