

# 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

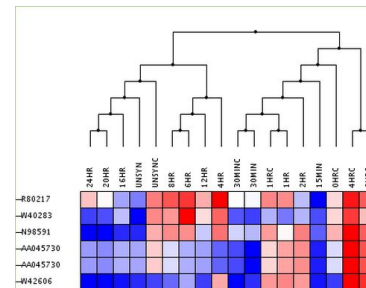
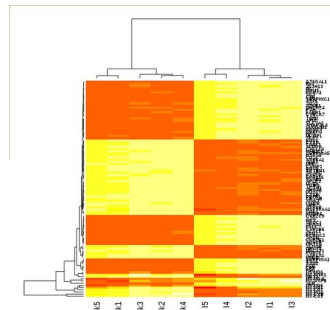
# Outline

**Design**

**Processing**

- **Differential expression**
- **Clustering**
- **Predictors**

**Functional profiling**



Any questions?



# Babelomics 5

## Babelomics 5

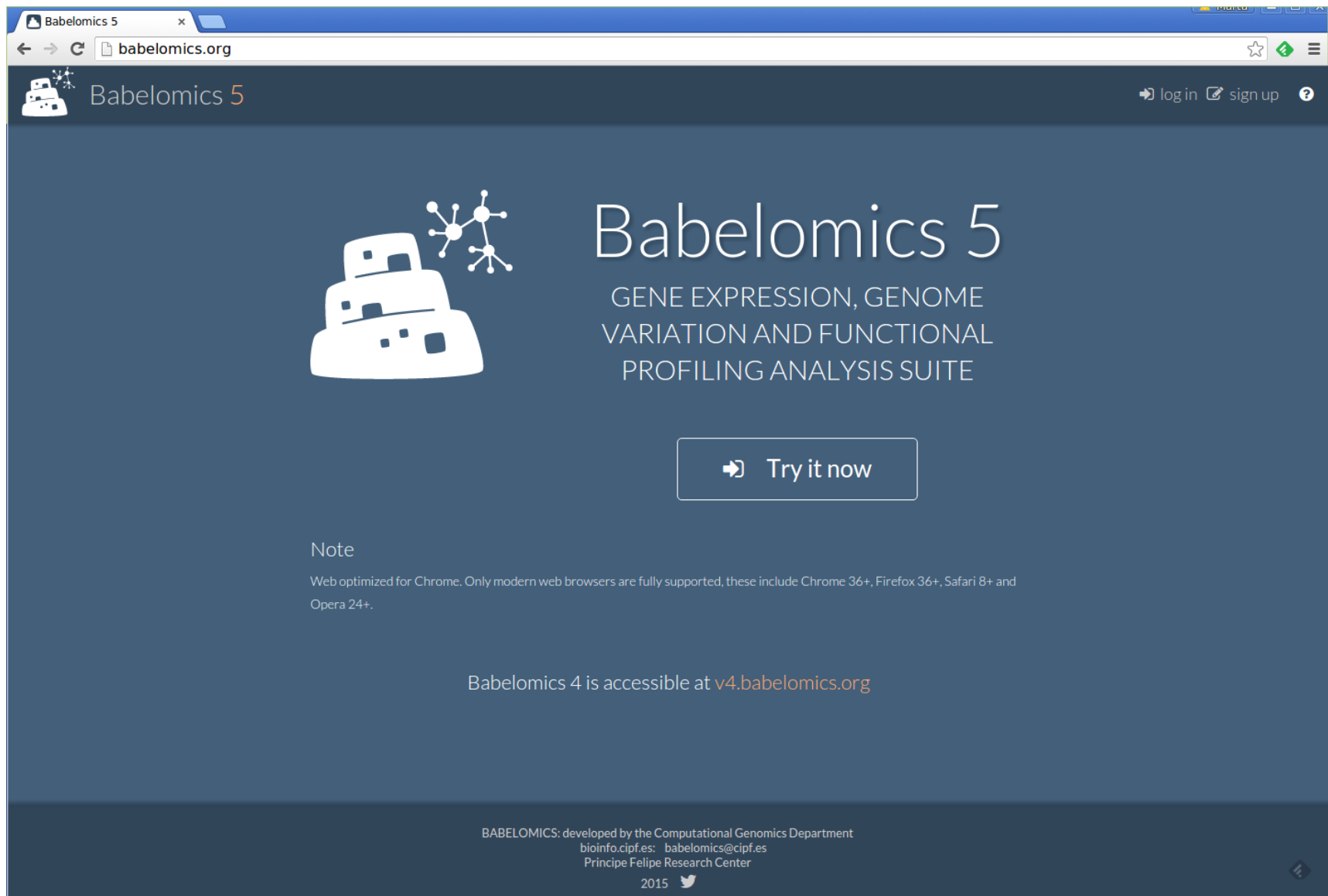
Integrative open source platform with advanced functional profiling for the analysis of

- Transcriptomics
- Proteomics
- Genomics

<http://www.babelomics.org>



# Babelomics 5



The screenshot shows the Babelomics 5 web application interface. The browser address bar displays `babelomics.org`. The page header includes the Babelomics 5 logo and navigation links for `log in`, `sign up`, and a help icon. The main content area features a large illustration of a microarray chip with a network diagram above it. The title **Babelomics 5** is prominently displayed, followed by the subtitle **GENE EXPRESSION, GENOME VARIATION AND FUNCTIONAL PROFILING ANALYSIS SUITE**. A central button labeled `Try it now` with a right-pointing arrow is present. 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+." A link indicates that Babelomics 4 is accessible at `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", the year "2015", and a Twitter icon.

Babelomics 5

GENE EXPRESSION, GENOME VARIATION AND FUNCTIONAL PROFILING ANALYSIS SUITE

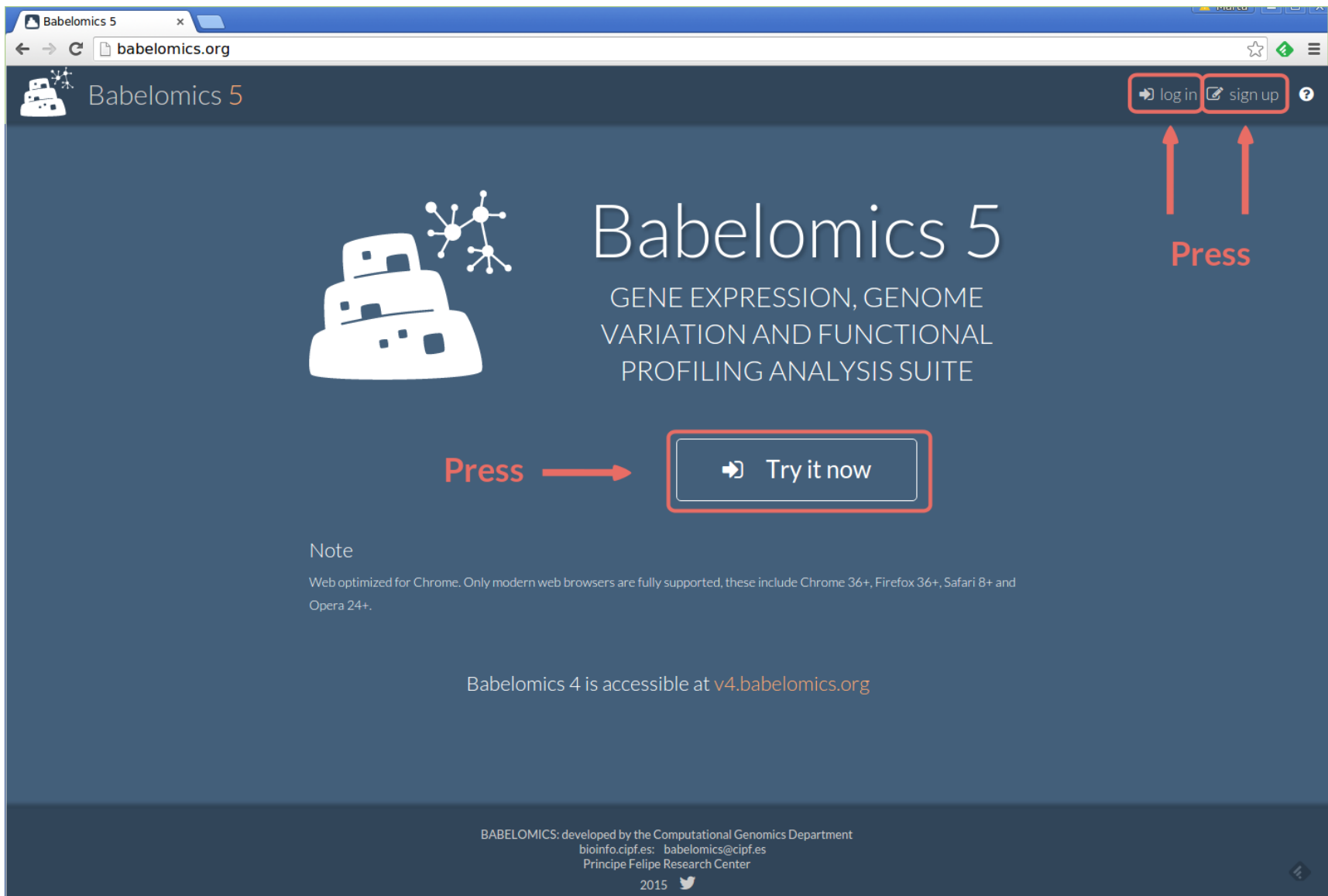
[Try it now](#)

**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's address bar displays 'babelomics.org'. The website header includes the 'Babelomics 5' logo and navigation links for 'log in' and 'sign up', both of which are highlighted with red boxes and red arrows pointing to them from the word 'Press' below. The main content area features a large illustration of a microscope and a network diagram, with the text 'Babelomics 5' and 'GENE EXPRESSION, GENOME VARIATION AND FUNCTIONAL PROFILING ANALYSIS SUITE'. A red arrow points from the word 'Press' to a 'Try it now' button, which is also highlighted with a red box. 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\")'. The footer contains contact information for the Computational Genomics Department at the Principe Felipe Research Center, the year 2015, and a Twitter icon.

Babelomics 5

GENE EXPRESSION, GENOME VARIATION AND FUNCTIONAL PROFILING ANALYSIS SUITE

Press → Try it now

Press ↑ log in sign up

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

- **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

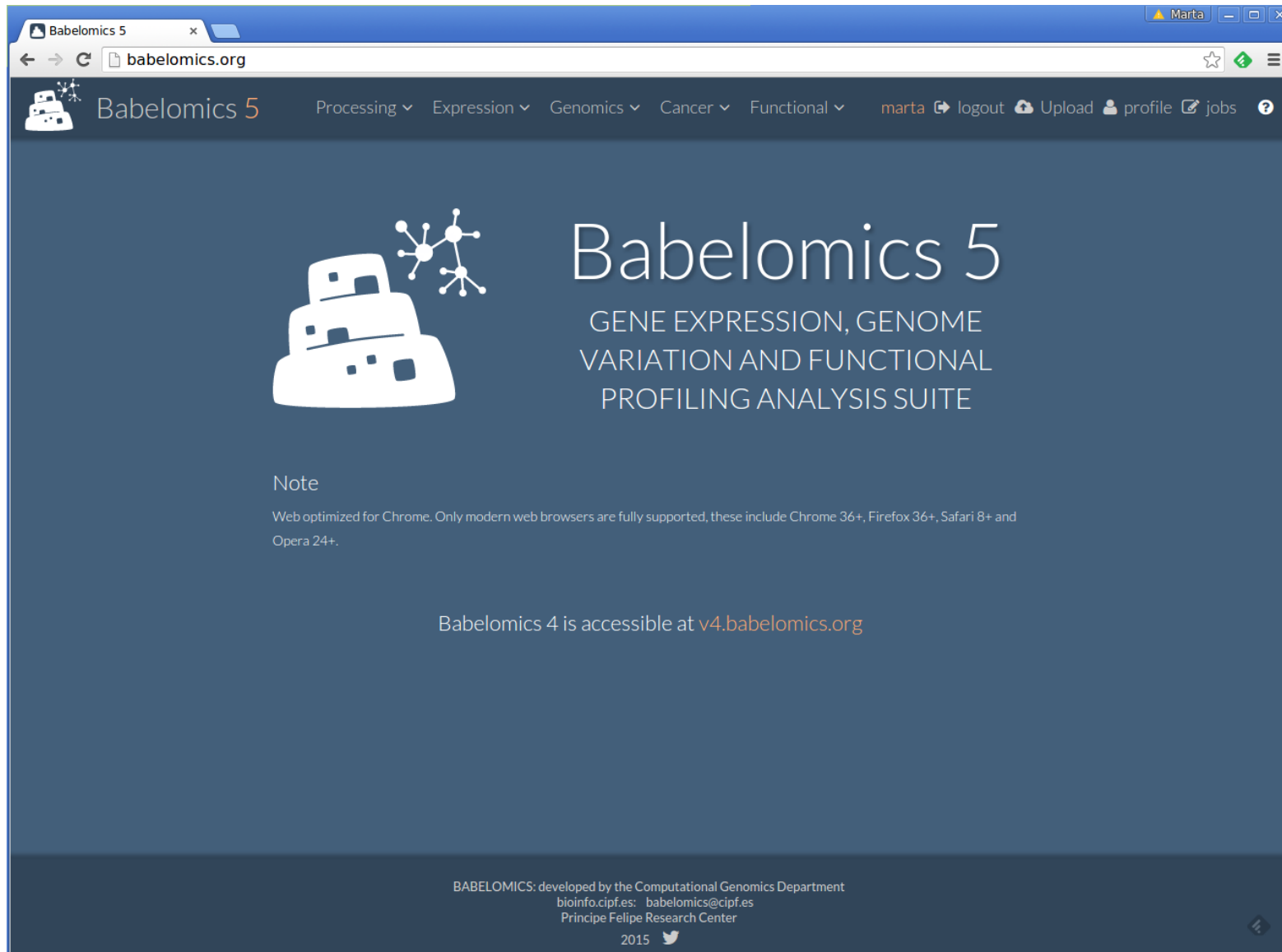
## Exercise

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 Babelomics 5 web application interface. The browser window has a single tab titled 'Babelomics 5' and the address bar shows 'babelomics.org'. The page features a dark blue header with a navigation menu including 'Processing', 'Expression', 'Genomics', 'Cancer', and 'Functional'. A user named 'marta' is logged in, with links for 'logout', 'Upload', 'profile', and 'jobs'. The main content area has a large white icon of a server stack with 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+'. Below this, it says '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', and a Twitter icon.

Babelomics 5

Processing ▾ Expression ▾ Genomics ▾ Cancer ▾ Functional ▾ marta logout Upload profile jobs ?


 Babelomics 5

GENE EXPRESSION, GENOME VARIATION AND FUNCTIONAL PROFILING ANALYSIS SUITE

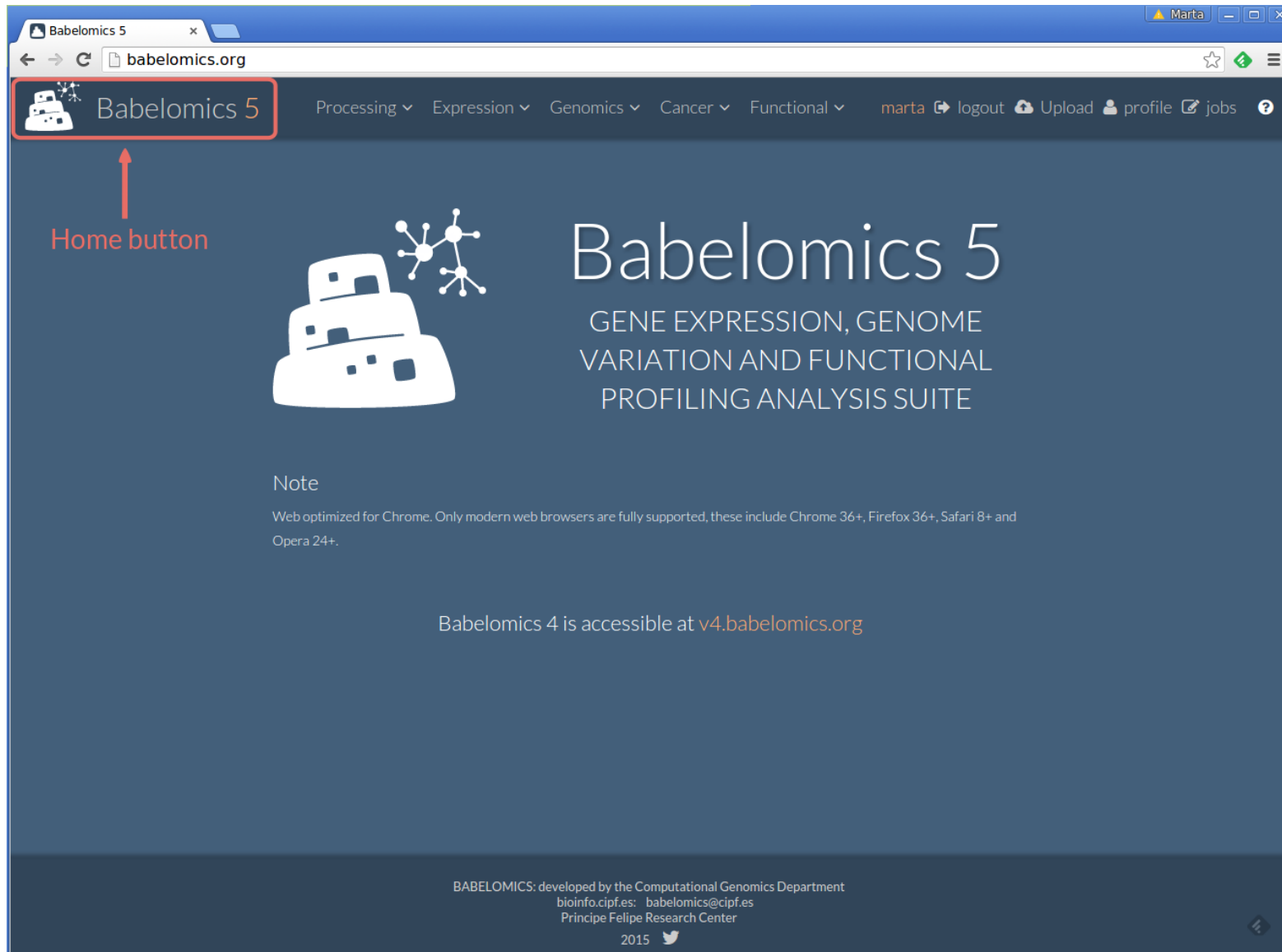
Note

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



# Web structure

The screenshot shows the Babelomics 5 web interface. The browser address bar displays 'babelomics.org'. The top navigation bar includes a 'Babelomics 5' logo (a stylized building icon) and a menu with categories: 'Processing', 'Expression', 'Genomics', 'Cancer', and 'Functional'. To the right of the menu are links for 'marta', 'logout', 'Upload', 'profile', 'jobs', and a help icon. A red arrow points from the text 'Home button' to the 'Babelomics 5' logo. The main content area features 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+'. Further down, it says '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', and a Twitter icon.

Babelomics 5

Processing ▾ Expression ▾ Genomics ▾ Cancer ▾ Functional ▾

marta logout Upload profile jobs ?

Home button

Analysis tools

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 stylized building icon) and a menu with 'Processing', 'Expression', 'Genomics', 'Cancer', and 'Functional'. A user profile 'marta' is logged in, with a 'logout' link. Other links include 'Upload', 'profile', 'jobs', and a help icon. The main content area features the title 'Babelomics 5' and the subtitle 'GENE EXPRESSION, GENOME VARIATION AND FUNCTIONAL PROFILING ANALYSIS SUITE'. A 'Note' section mentions browser compatibility. A link for 'Babelomics 4' is provided at the bottom. The footer contains development information and the year 2015.

Annotations on the screenshot:

- Home button:** An arrow points to the Babelomics 5 logo in the navigation bar.
- Analysis tools:** A red box highlights the menu items: Processing, Expression, Genomics, Cancer, and Functional.
- Logout:** A red box highlights the 'marta' user profile and the 'logout' link.

**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 top 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 boxes highlight the logo, the menu, and the user options. Red arrows point from the text 'Home button' to the logo, 'Analysis tools' to the menu, 'Logout' to the 'logout' link, and 'Upload' to the 'Upload' link. The main content area features a stylized server icon, the title 'Babelomics 5', and the subtitle 'GENE EXPRESSION, GENOME VARIATION AND FUNCTIONAL PROFILING ANALYSIS SUITE'. A 'Note' section mentions browser compatibility. A link for 'Babelomics 4' is provided at the bottom. The footer contains development information 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 top navigation bar includes a 'Babelomics 5' logo, a menu with 'Processing', 'Expression', 'Genomics', 'Cancer', and 'Functional', and user options: 'marta', 'logout', 'Upload', 'profile', 'jobs', and a help icon. Red arrows point from the following labels to their corresponding elements: 'Home button' points to the logo; 'Analysis tools' points to the menu; 'Logout' points to the 'logout' link; 'Upload' points to the 'Upload' link; and 'Settings' points to the 'profile' link. The main content area features a server icon, the title 'Babelomics 5', the subtitle 'GENE EXPRESSION, GENOME VARIATION AND FUNCTIONAL PROFILING ANALYSIS SUITE', a 'Note' about browser compatibility, and a link to 'v4.babelomics.org'. The footer contains development information and the year '2015'.

Babelomics 5

Processing ▾ Expression ▾ Genomics ▾ Cancer ▾ Functional ▾

marta logout Upload profile jobs ?

Home button

Analysis tools

Logout

Upload

Settings

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 top navigation bar includes a 'Babelomics 5' logo, a menu with 'Processing', 'Expression', 'Genomics', 'Cancer', and 'Functional', and user options: '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 a stylized server icon, the title 'Babelomics 5', and the subtitle 'GENE EXPRESSION, GENOME VARIATION AND FUNCTIONAL PROFILING ANALYSIS SUITE'. A 'Note' section mentions browser compatibility. A footer provides development information 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 top navigation bar includes a 'Babelomics 5' logo, a menu with 'Processing', 'Expression', 'Genomics', 'Cancer', and 'Functional', and user options: '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 icon, the title 'Babelomics 5', and the subtitle 'GENE EXPRESSION, GENOME VARIATION AND FUNCTIONAL PROFILING ANALYSIS SUITE'. A 'Note' section mentions browser compatibility. A footer provides 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

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)

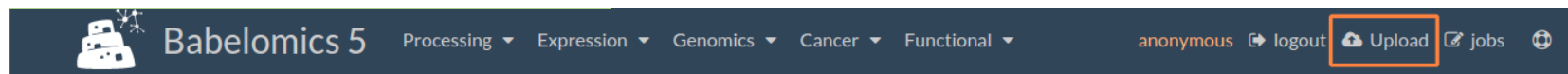
BABELOMICS: developed by the Computational Genomics Department  
bioinfo.cipf.es: babelomics@cipf.es  
Principe Felipe Research Center

2015



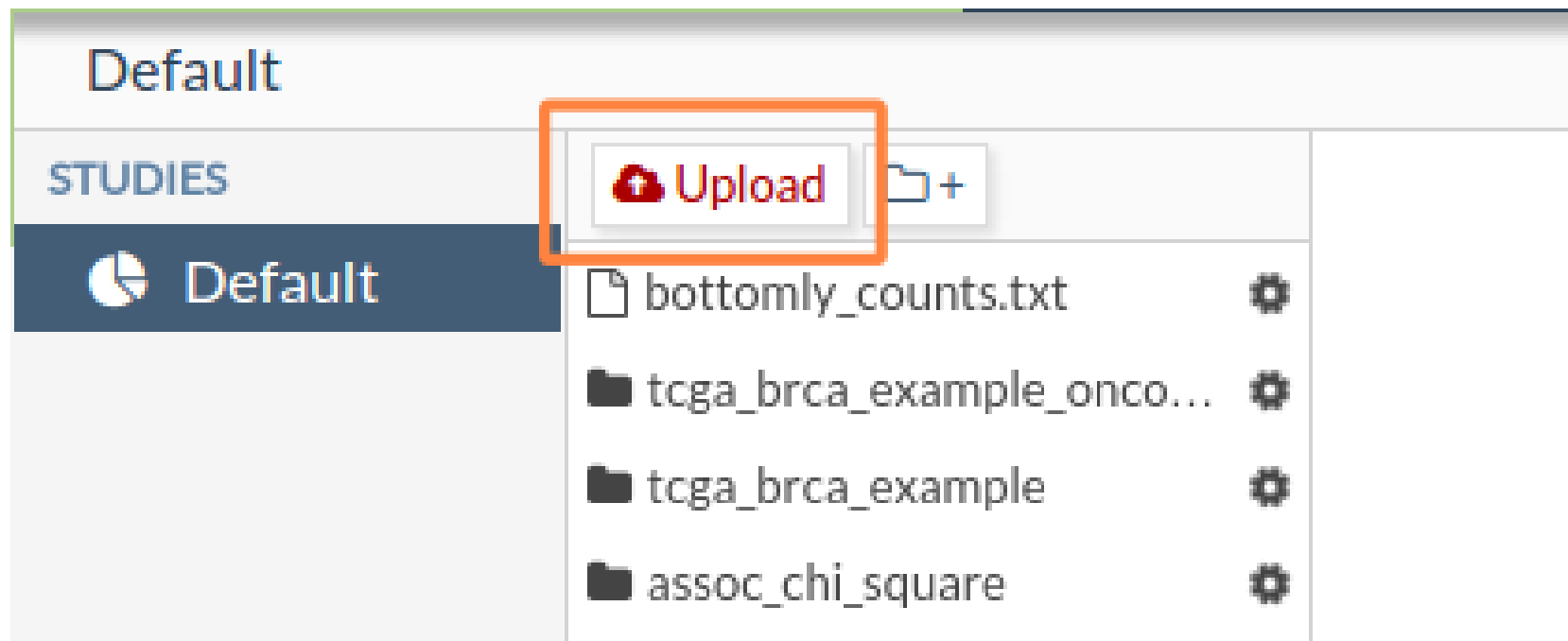
# Uploading data

## Uploading data tutorial



# Uploading data

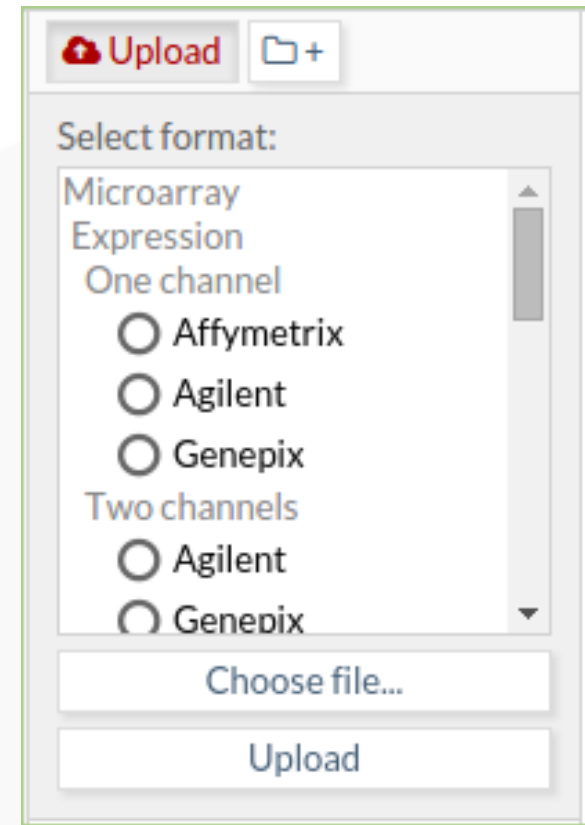
## Uploading data tutorial



# 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

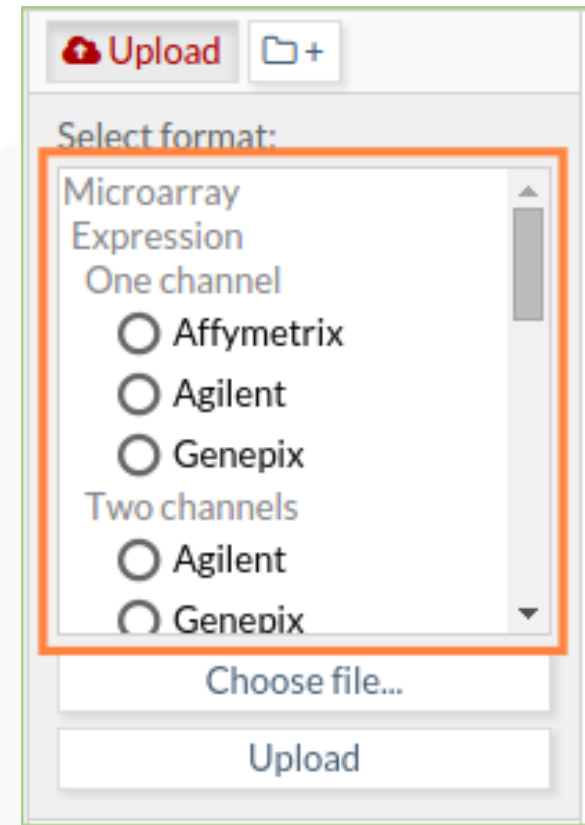


The screenshot shows a web-based upload interface. At the top, there is a red 'Upload' button with an upward arrow icon and a folder icon with a plus sign. Below this, a section titled 'Select format:' contains a scrollable list of options. The list is organized into two main categories: 'Microarray' and 'Expression'. Under 'Microarray', there is a sub-category 'One channel' with three radio button options: 'Affymetrix', 'Agilent', and 'Genepix'. Under 'Expression', there is a sub-category 'Two channels' with two radio button options: 'Agilent' and 'Genepix'. At the bottom of the form, there are two buttons: 'Choose file...' and '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

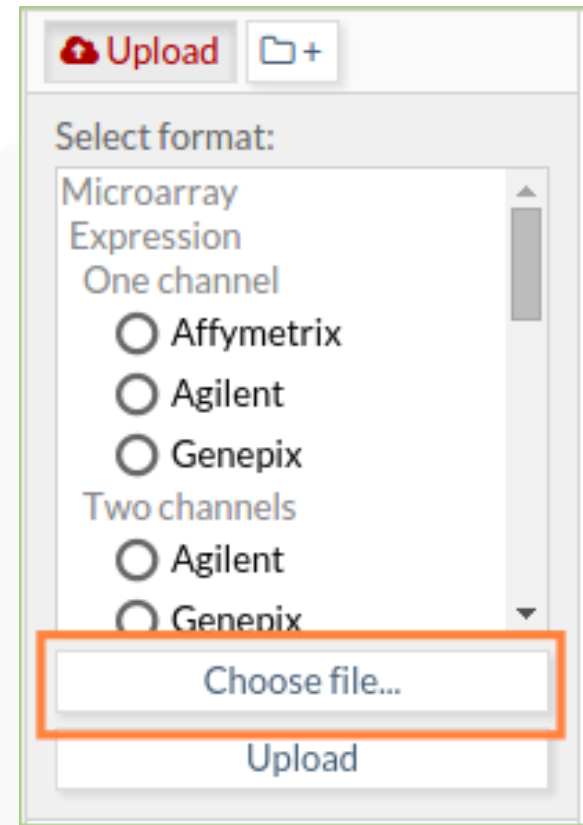


The screenshot shows a web interface for uploading data. At the top, there is a red 'Upload' button with an upward arrow icon and a folder icon with a plus sign. Below this is a 'Select format:' section. A list of options is shown, with 'Microarray' and 'Expression' as main categories. Under 'Expression', there is a sub-category 'One channel' which is highlighted with an orange border. Under 'One channel', there are three radio button options: 'Affymetrix', 'Agilent', and 'Genepix'. Below 'One channel' is another sub-category 'Two channels' with two radio button options: 'Agilent' and 'Genepix'. At the bottom of the form, there are two buttons: 'Choose file...' and '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

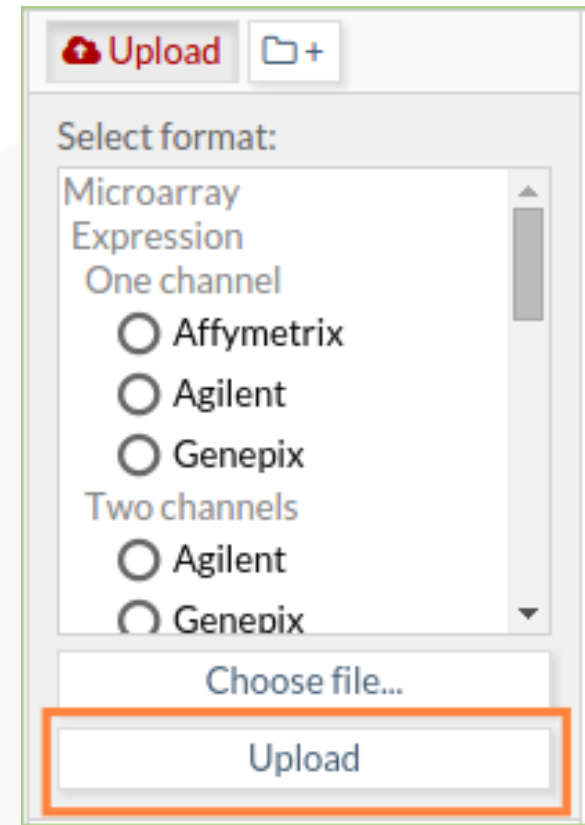


The screenshot shows a web interface for uploading data. At the top, there is a red 'Upload' button with an upward arrow icon and a folder icon with a plus sign. Below this is a 'Select format:' section. It contains a list of options: 'Microarray' (selected), 'Expression', and 'One channel'. Under 'Microarray', there are three radio button options: 'Affymetrix', 'Agilent', and 'Genepix'. Under 'Expression', there are two radio button options: 'Agilent' and 'Genepix'. At the bottom of the form, there is a 'Choose file...' button highlighted with an orange border, and an 'Upload' button below it.

# 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

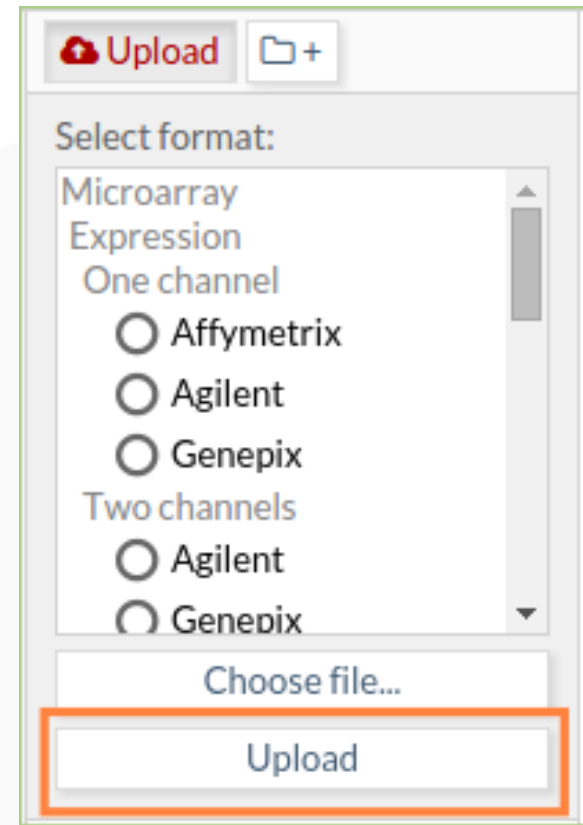


The screenshot shows a web interface for uploading data. At the top, there is a red 'Upload' button with an upward arrow icon and a folder icon with a plus sign. Below this is a 'Select format:' section. It contains a list of options: 'Microarray' (selected), 'Expression', and 'One channel'. Under 'One channel', there are three radio button options: 'Affymetrix', 'Agilent', and 'Genepix'. Below these, there is a 'Two channels' section with two radio button options: 'Agilent' and 'Genepix'. At the bottom of the form, there is a 'Choose file...' button and an 'Upload' button, which is highlighted with a red rectangular border.

# 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 a red 'Upload' button with an upward arrow icon and a folder icon with a plus sign. Below this is a 'Select format:' section. It contains a list of options: 'Microarray' (selected), 'Expression', and 'One channel'. Under 'Microarray', there are three radio button options: 'Affymetrix', 'Agilent', and 'Genepix'. Under 'Expression', there are two radio button options: 'Agilent' and 'Genepix'. Below the list is a 'Choose file...' button. At the bottom, there is a large 'Upload' button, which is highlighted with a red rectangular border.

# Uploading data

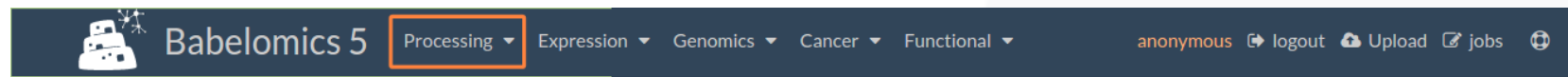
## Exercise

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





## Processing ▼

Expression ▼

Genomics ▼

Cancer ▼

Functional ▼



- ▶ Affymetrix

- Agilent

- Genepix

## ▶ Agilent

- Genepix

► Edit your uploaded data

- Pre-processing

# 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 ▼

+

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

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](#)