

Biomarker visualization using Rshiny and Python

Xuan Yang

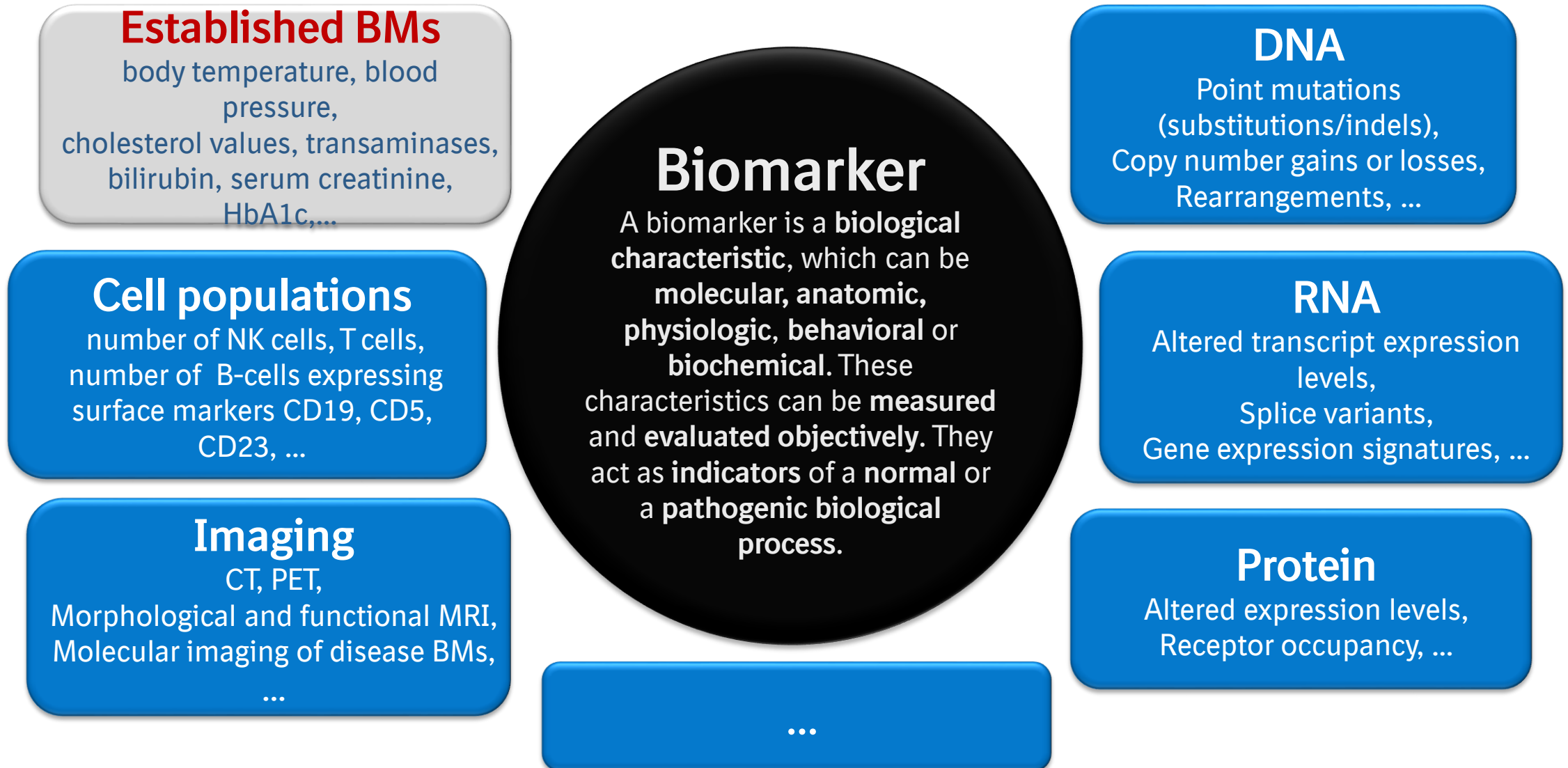
PHUSE SDE Shanghai, Nov 2019

Disclaimer

The views and opinions expressed in the following slides are those of the presenter and do not represent the views of Boehringer Ingelheim.

The visualizations were generated using simulated data.

Biomarkers in clinical trials



Biomarkers in drug development

Biomarkers can

- monitor safety and efficacy of a therapy
- define best doses for later clinical phases (PK/PD)
- get first impression on potential efficacy, thereby helping early go no-go decision making
- explore potential subgroups with differing response
- ...

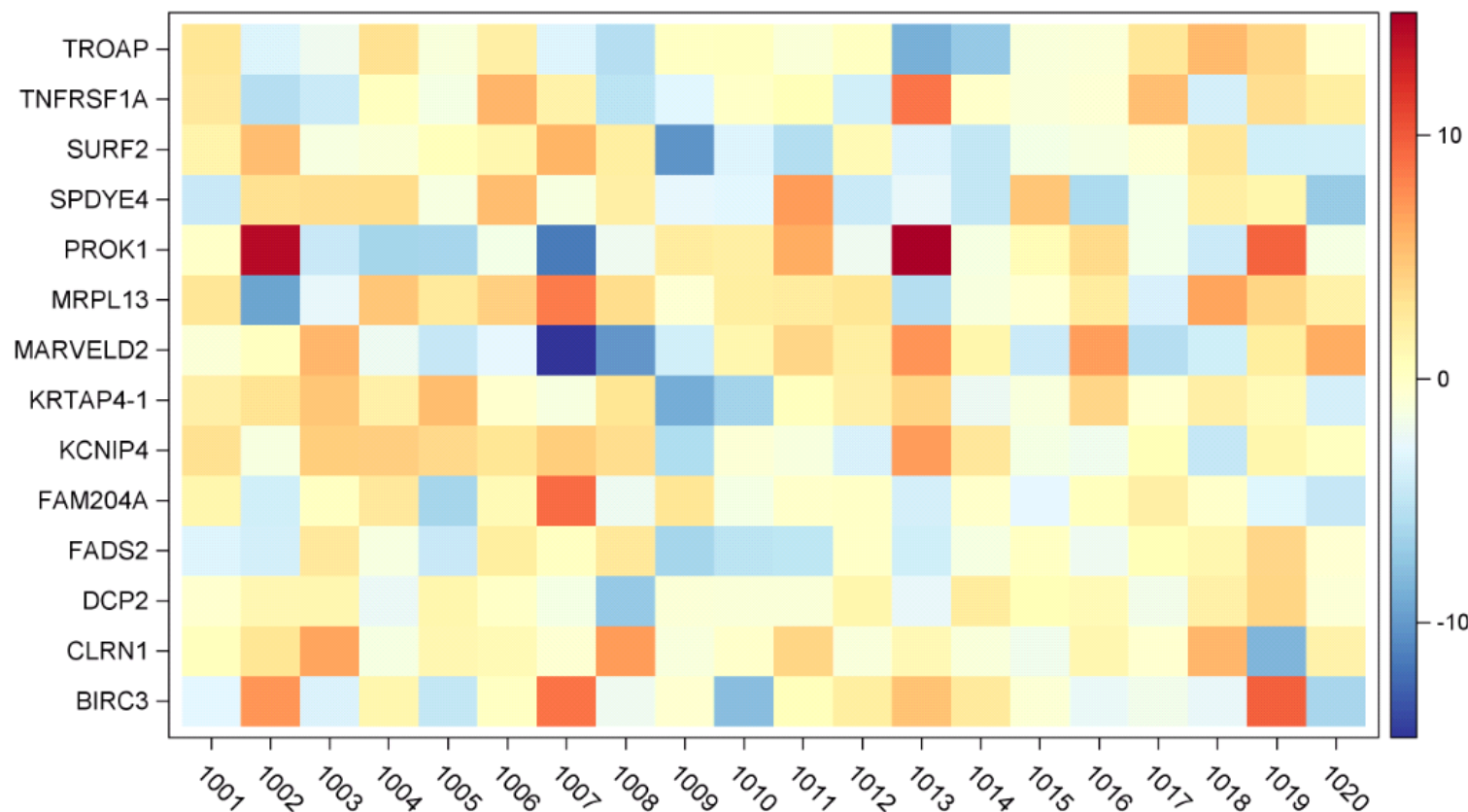
Visualization: heatmap

- A graphical representation of data that:
 - uses a system of color-coding
 - visualize values of different biomarkers in a single graphic
- Hierarchical cluster analysis of biomarkers
 - similarity of different biomarkers and subgroups could be shown visually
- Facilitate the identification of subgroups and trends

Challenges in programming

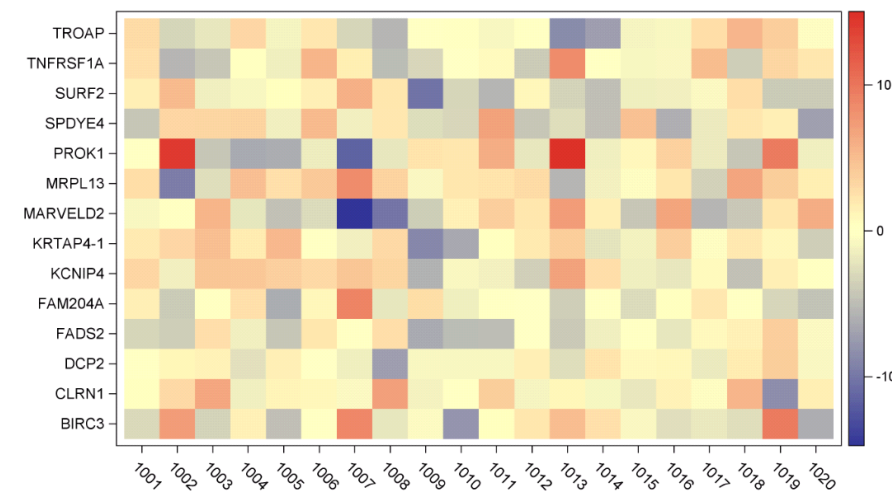
- Generation of heatmap based on selected biomarkers expression levels on specific cut-off point needs lots of programming efforts
- A tool to facilitate exploration is expected

Heatmap in SAS



Challenges:

- Distance
- Clustering
- Color tuning



`COLORMODEL=(cx313695 cx4575B4 cx74ADD1 cxABD9E9 cxE0F3F8 cxFFFFBF cxFEE090 cxFDAE61 cxF46D43 cxD73027 cxA50026)`

11 colors

VS

3 colors

`COLORMODEL=(cx313695 cxFFFFBF cxA50026)`

Import data

ADSL

Browse... No file sel

ADRNA

Browse... No file sel

Aggregated data

Browse... No file sel

View in Data panel:

ADSL

Data filter

Baseline

Visit 1

Target visit

Visit 2

Abs. LFC >=

0.2

Annotation

Row dendrogram

Column dendrogram

Distance method

Euclidean

Clustering linkage

Complete

Number of Clusters

1 2 15

Color

Dendrogram

Layout

Select Color Palette

RdYIBu

Set Number of Colors

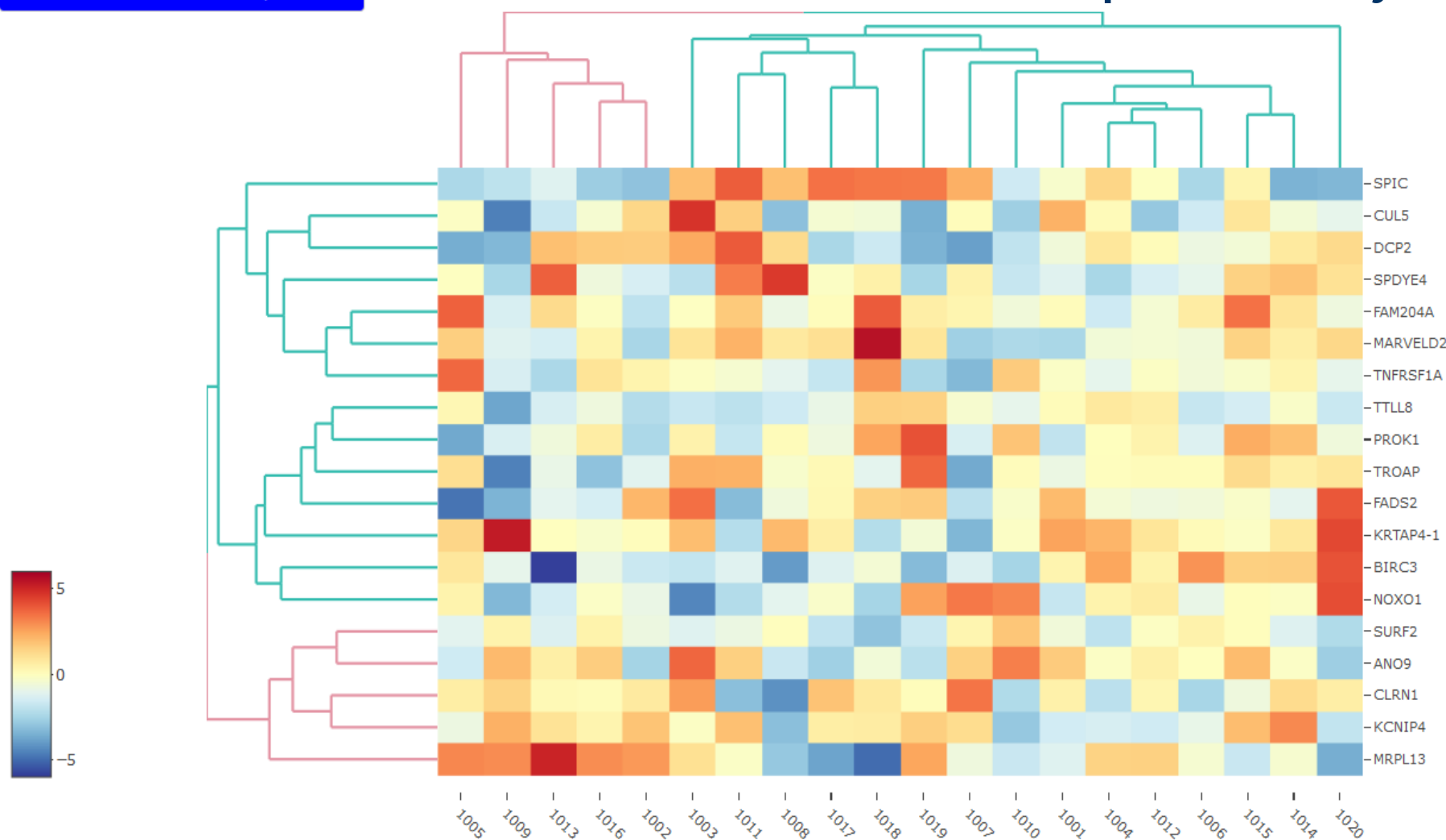
Heatmaply

Overall

Data

Download interactive Heatmap as HTML

Heatmap in RShiny



Key package: shinyHeatmaply
Shiny template developed by Jonathan Sidi and Tal Galili

Heatmap in RShiny

- Advantages
 - Highly flexible, great tool for exploring
 - ✓ choose any value as the cut-point
 - ✓ combination of different distance and clustering methods
 - ✓ display subgroup overview
 - ✓ give tables for input datasets
 - Interactive
 - ✓ show values when moving cursor
 - ✓ save modified plots and tables

Heatmap in Python

- Key libraries (packages)

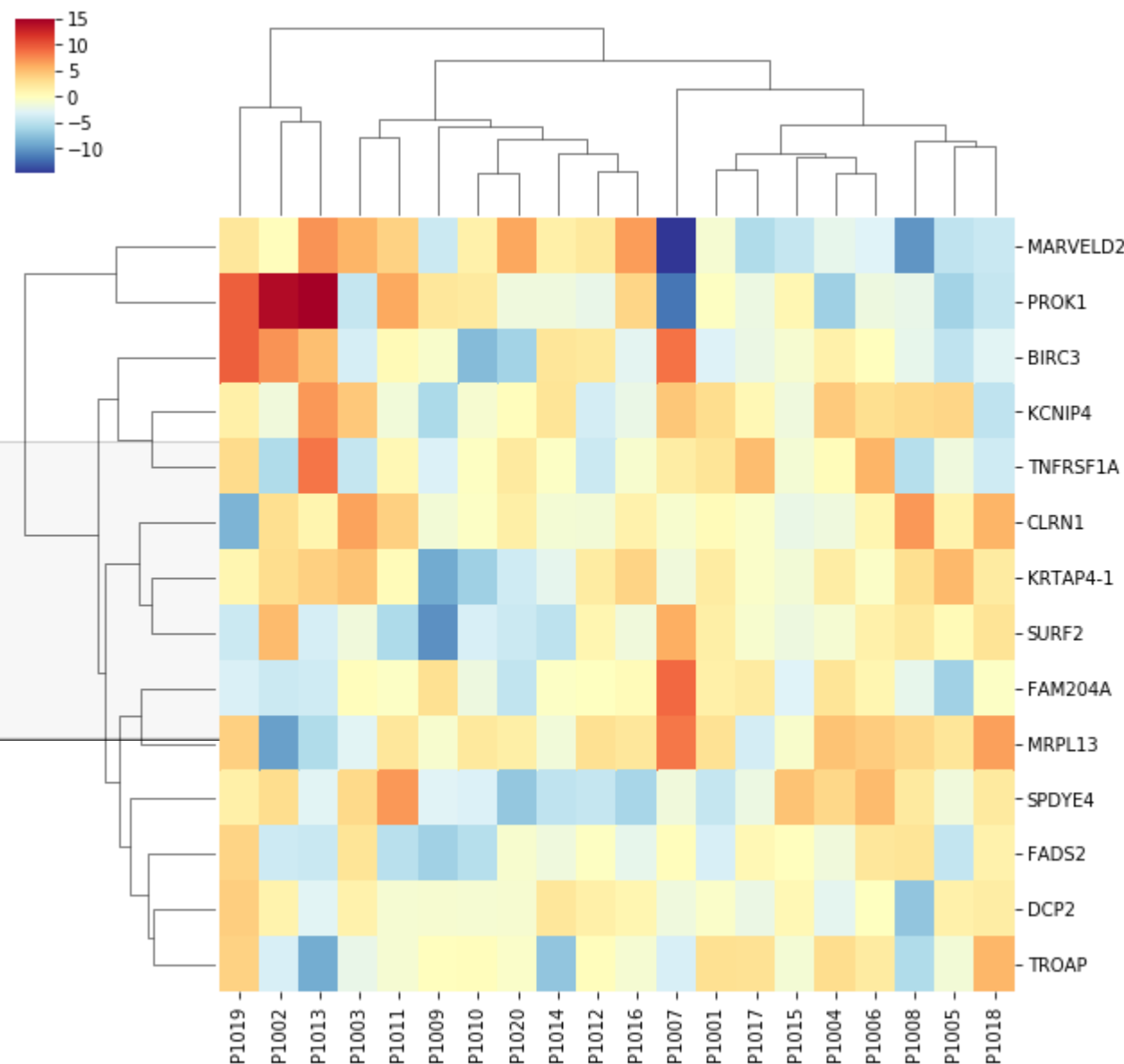
- pandas
- matplotlib
- seaborn

```
# Genomic data set
df = pd.read_csv(url)
df = df.iloc[:, 1:22]
df = df.set_index('gensmbl')
del df.index.name

# Set color palette
sns.clustermap(df, metric="euclidean", method="ward", cmap="RdYlBu_r")
plt.show()
```

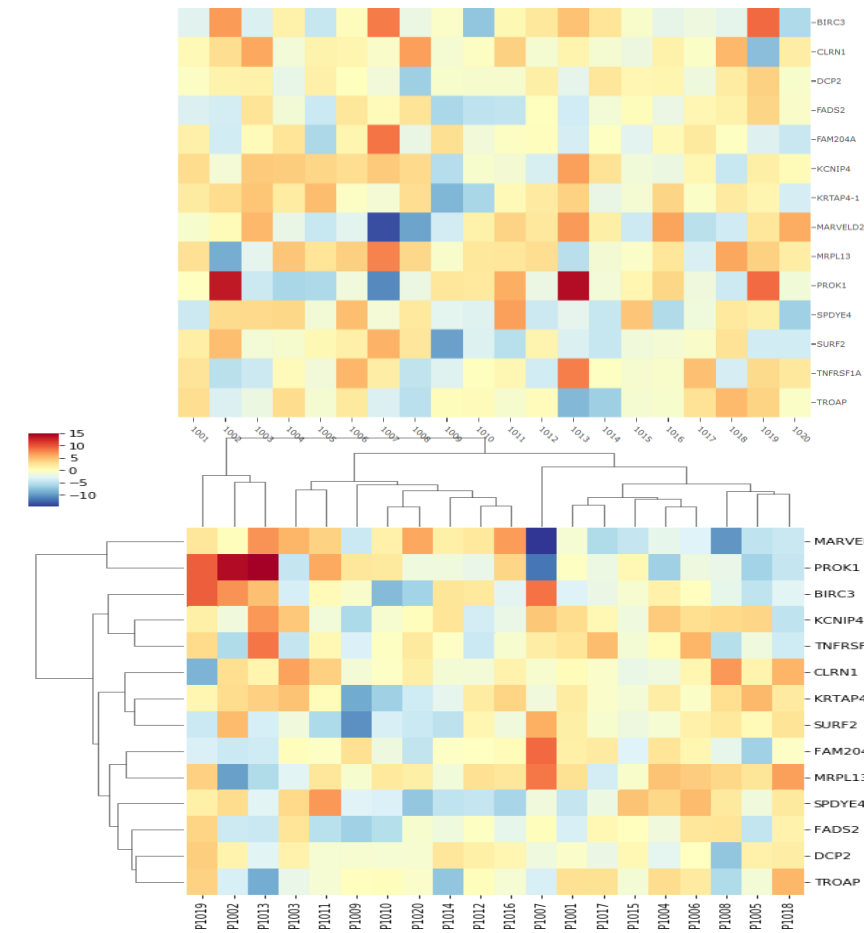
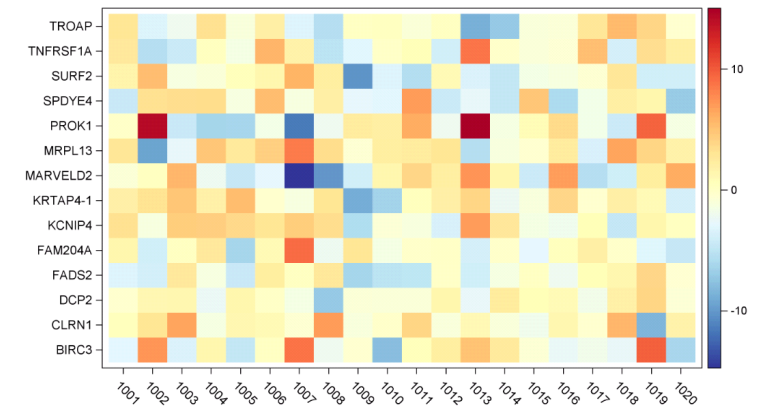
- Interactive plots potentials

- plotly
- matplotlib
- Bokeh



SAS vs. R vs. Python

- SAS
 - GTL
- R
 - Interactive
 - More functions/packages
- Python
 - More functions/libraries
 - Interactive
 - Speed
 - Machine learning



THANK YOU

Name: Xuan Yang

Organization: Boehringer Ingelheim (China) Investment Co., Ltd.

Address: Benben Building, No.300 Xikang Rd.

City, State ZIP: Shanghai, 200040

Work Phone: +86-21-5288 0026

E-mail: xuan.yang@boehringer-ingelheim.com