

*An efficient GUI tool
for spectra processing
from 1D 1H-NMR metabolomics data*
<http://nmrprocflow.org/>



*combined with the computing power of
a Cloud Research Environment*

Daniel Jacob

Two major metabolomics approaches

Metabolic Fingerprinting

The complex data are directly and initially used for global multivariate statistical analysis.

Subsequently, **metabolite features that distinguish sample classes are identified** and then the structures of distinguishing metabolic features are established

Targeted Metabolomics

Quantitative approach wherein **a set of known metabolites are quantitated**.

The identities of metabolites were initially established based on the available databases and using standard compounds.

The identified metabolite peaks are then quantified based on internal or external reference compounds.

Standard Operating Procedures (SOP)

Harvest / Storage

Sample Preparation

Analytical Sample Preparation

Analytical

Metabolic Fingerprinting

Targeted Metabolomics



Pre-processing

Spectra processing

Statistical analysis

Quantification

Identification

Efficiency,
Repeatability



Need of
Traceability



Evidence,
Confidence

Science ⇒ Data ⇒ Experiments ⇒ Good practices

Galaxy – Weakness

“Standard workflow”



Given the nature of the 1D NMR spectra and due to the diversity of problems encountered during the various stages of processing:

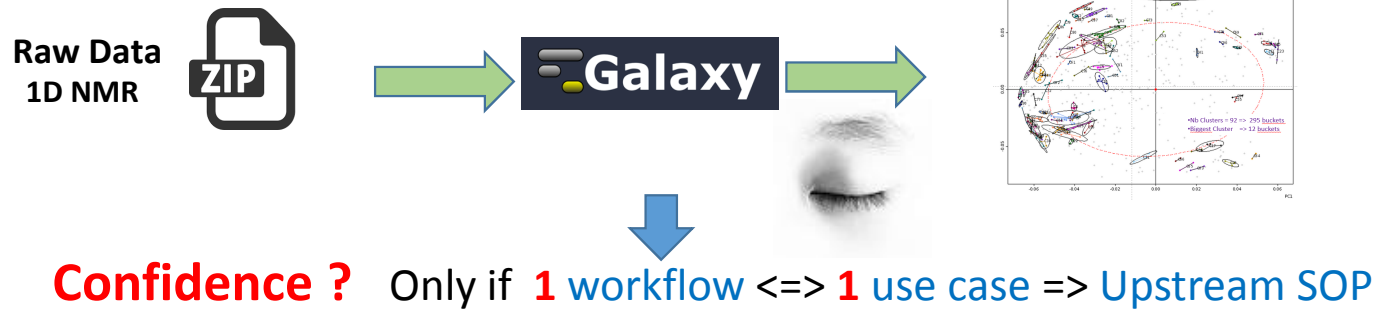
- *baseline correction,*
- *calibration ppm,*
- *removal of solvents and other contaminants,*
- *re-alignment of areas having high variations in chemical shifts between spectra, ...*

and depending on:

- *the biological context (humans, plants, micro-organisms),*
- *the type of tissue (plasma, urine, extract of fruit flesh ...),*
- *the analytical protocol (choice of the NMR sequence, additives for calibration and / or quantification, the buffer solution to stabilize the pH in the tube, etc ...).*

It is essential to process this type of data, with an interactive interface that enables the spectra visualization.

Galaxy – Recommendation



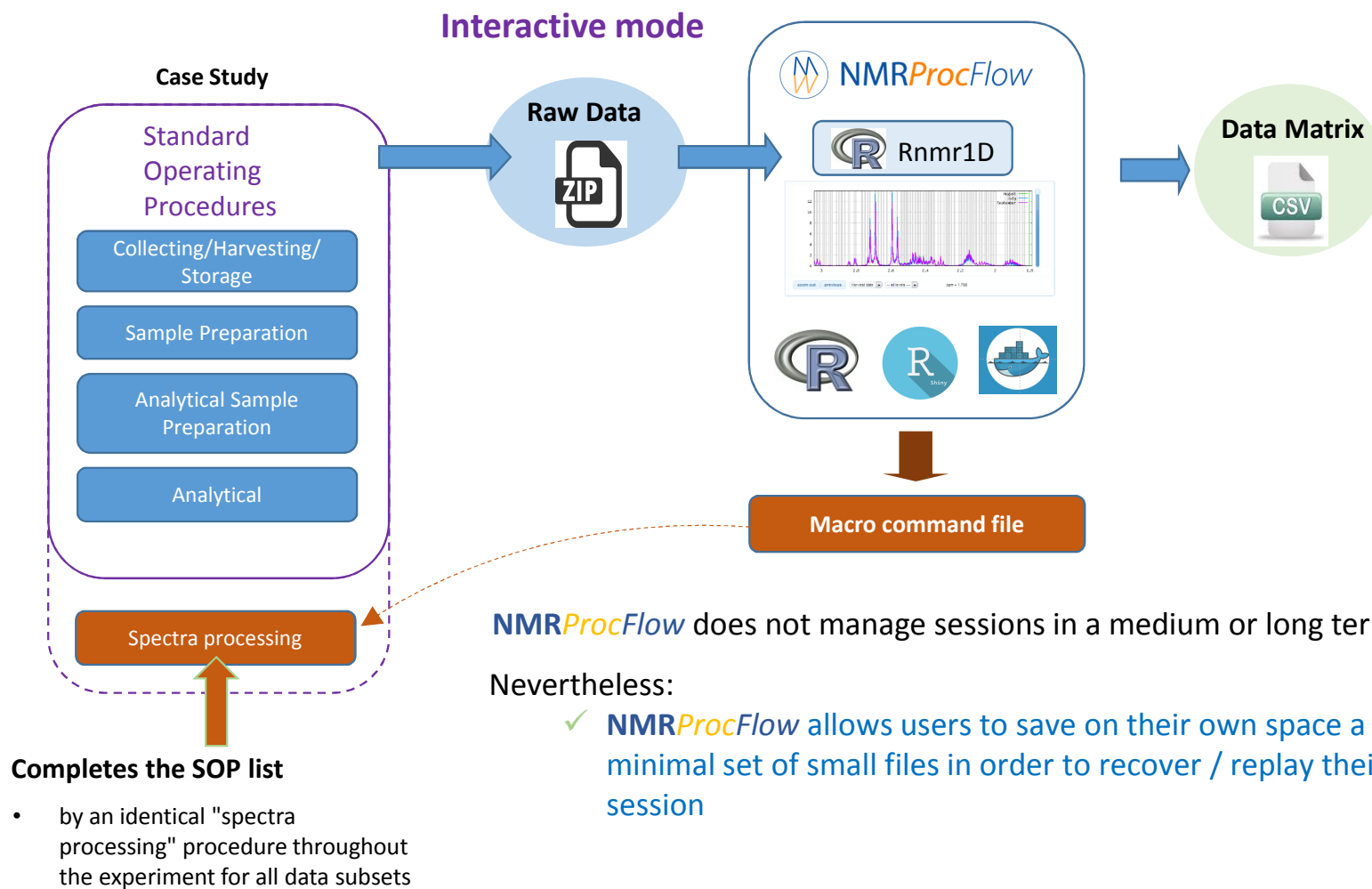
We consider the implementation of **NMR spectra processing workflows** executed in *batch mode* as relevant provided that we want to process in this way **very well-mastered** and **very reproducible use cases**, i.e. by applying the **same Standard Operating Procedures (SOP)**.

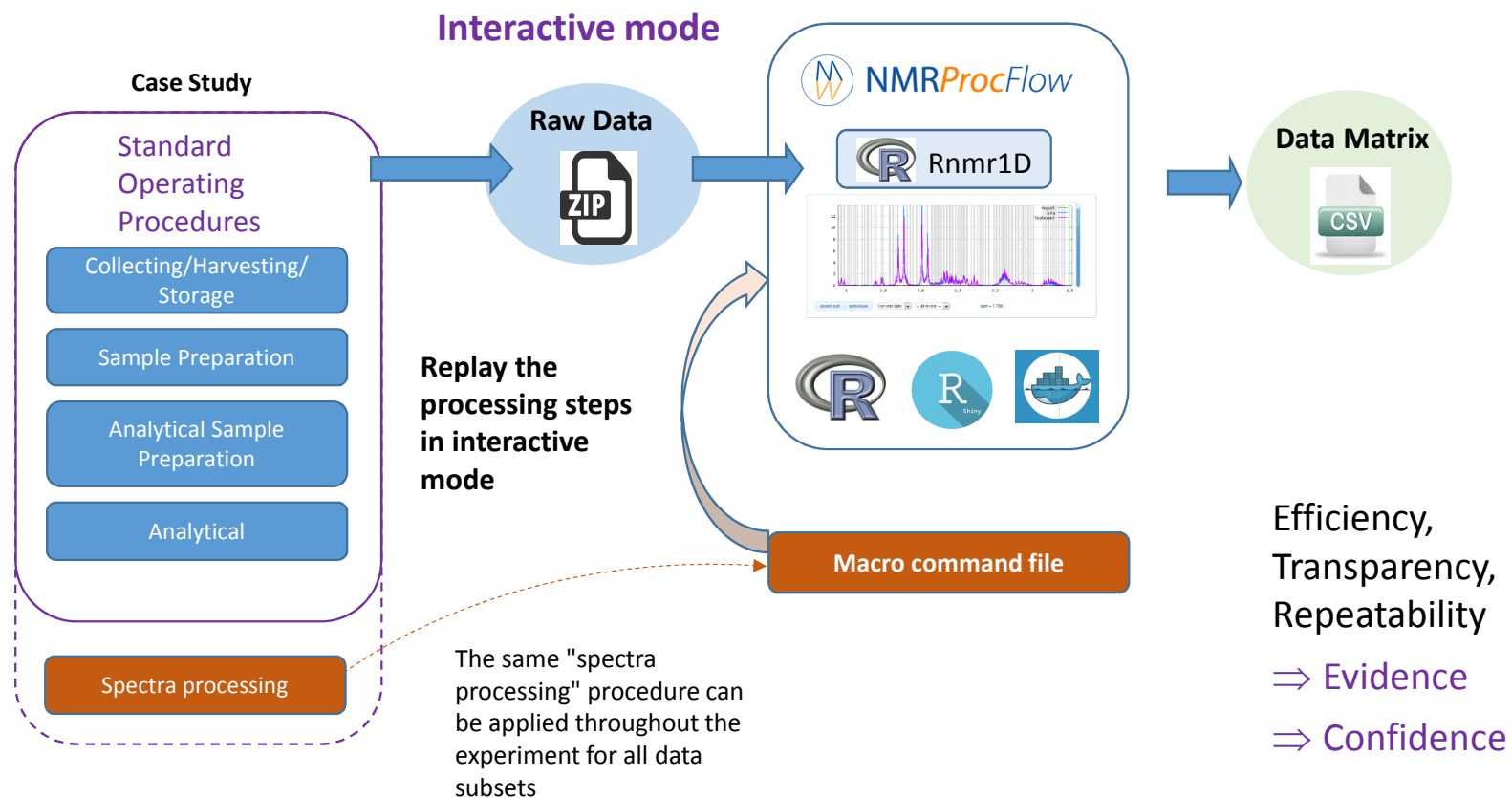
Galaxy – Strength

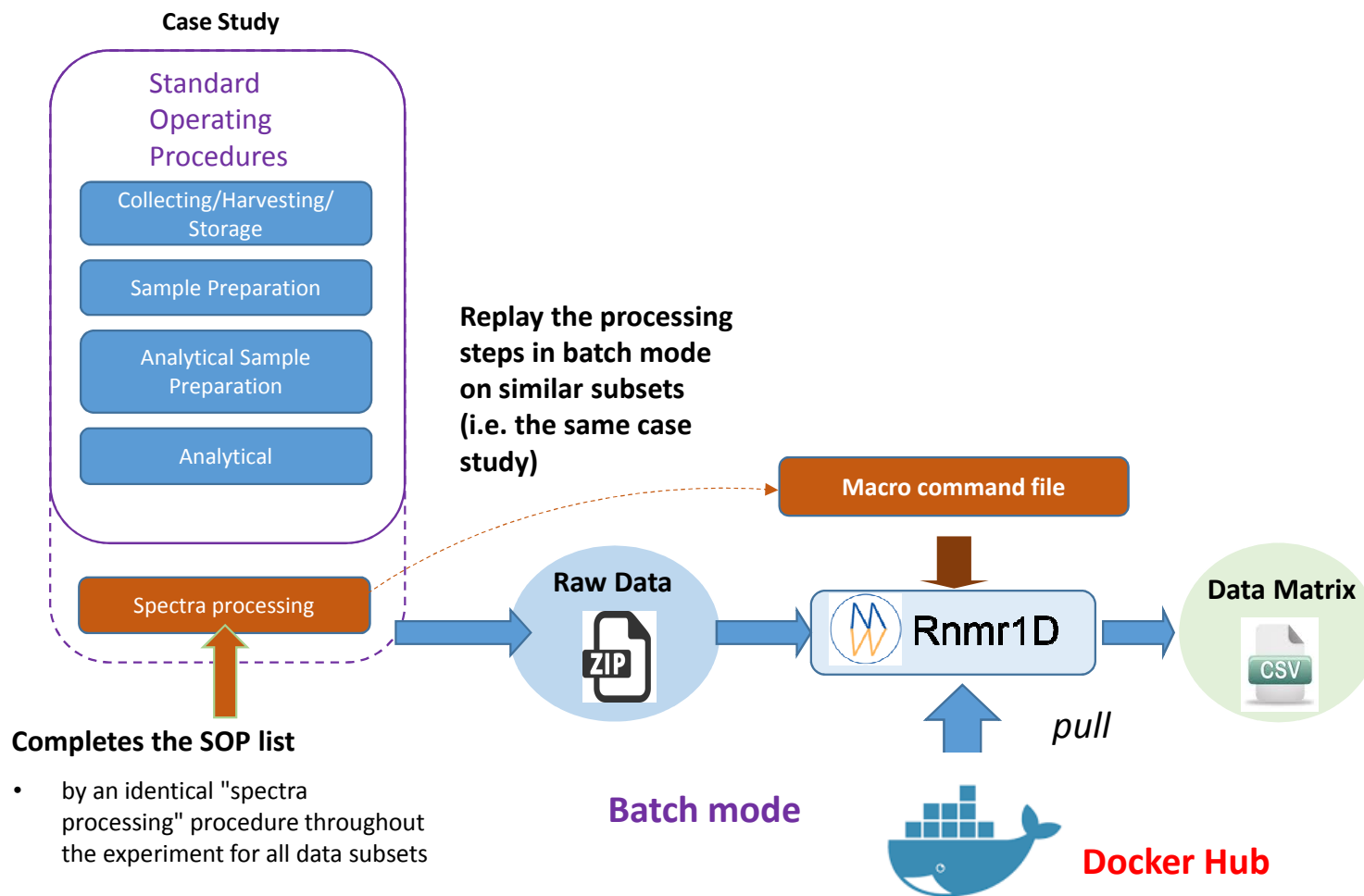


By taking into account the previous consideration,
the strenghtness of Galaxy is to :

- Allows to process large dataset sets (ie> 300)
- **Allows non-experts to handle their data**







<https://bitbucket.org/nmrprocflow/rnmr1d>


Atlassian Bitbucket Features Pricing Find a repository... English Sign up Log in

nmrprocflow / NMRProcFlow / Rnmr1D

Source

master Rnmr1D /

examples			
src			
Dockerfile	2.5 KB	2017-05-26	Rnmr1D v 1.2
LICENSE	35.1 KB	2017-05-30	Rnmr1D v1.2
README.md	4.0 KB	7 hours ago	update v1.2.8
logo_Rnmr1D.png	13.1 KB	2017-05-29	Rnmr1D v 1.2



Rnmr1D

Rnmr1D

Version 1.2.8

Rnmr1D is the main module in the NMRProcFlow web application (<http://nmrprocflow.org>) concerning the NMR spectra processing.

- Inside NMRProcFlow, Rnmr1D allows users to process their NMR spectra within a GUI application and thus the macro-command sequence coming from the GUI can be saved.
- Outside NMRProcFlow Rnmr1D become a CLI application allowing users to replay the macro-command sequence generated within NMRProcFlow. Moreover, without using NMRProcFlow, this module can also be used to replace any 'home-made script' by a macro-command sequence.

<https://hub.docker.com/r/nmrprocflow/rnmr1d/>

Search

PUBLIC REPOSITORY

nmrprocflow/rnmr1d

Last pushed: a month ago

Repo Info Tags Collaborators Webhooks Settings

Short Description

Short description is empty for this repo.

Full Description

Copyright 2015-2016

INRA UMR 1332 BFP, Bordeaux Metabolomics Facility
France

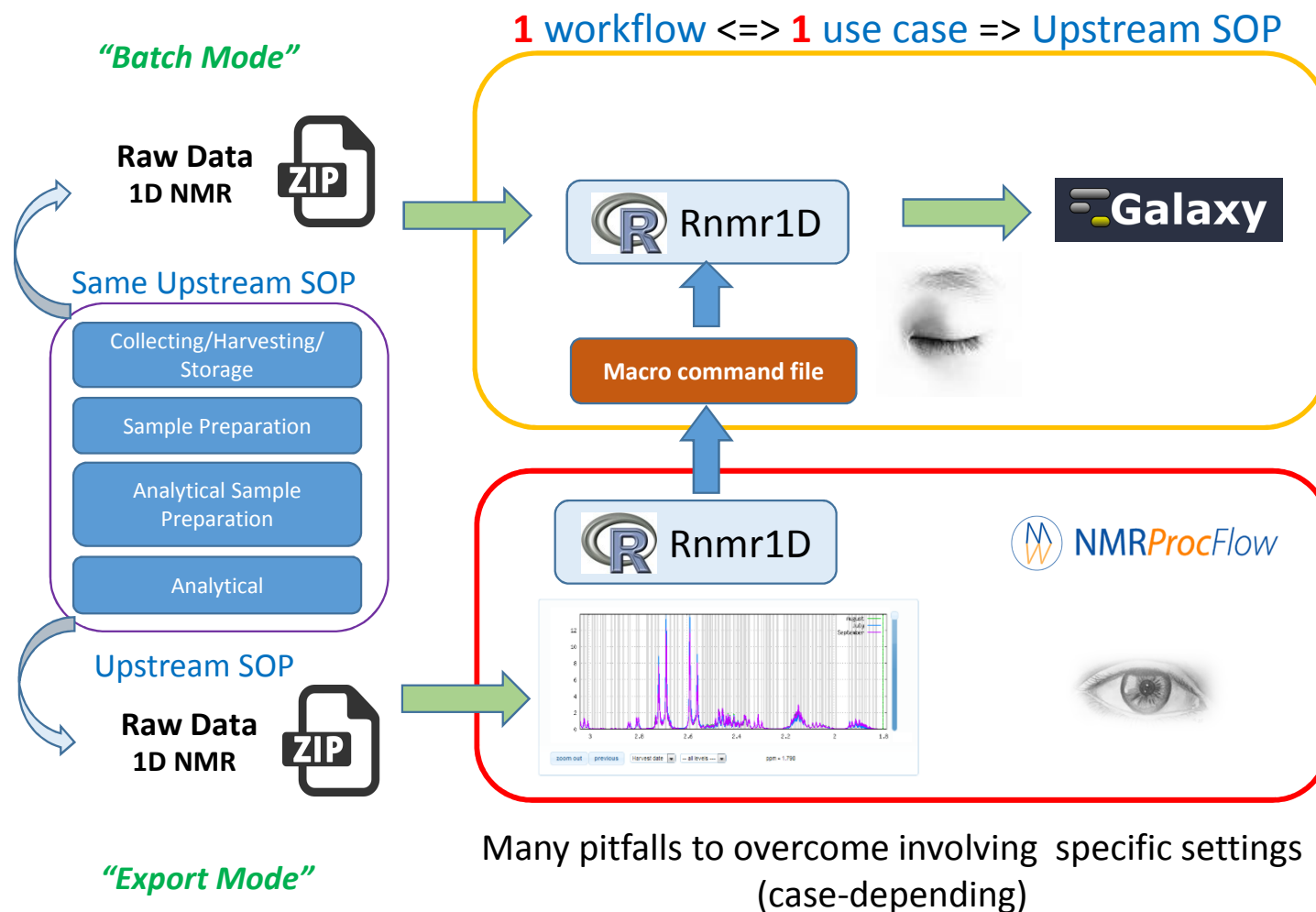
Funded by:

INRA UMR 1332 BFP, Bordeaux Metabolomics Facility
the ANR-11-INBS-0010 grant (MetaboHUB)

Docker Pull Command

```
docker pull nmrprocflow/rnmr1d
```

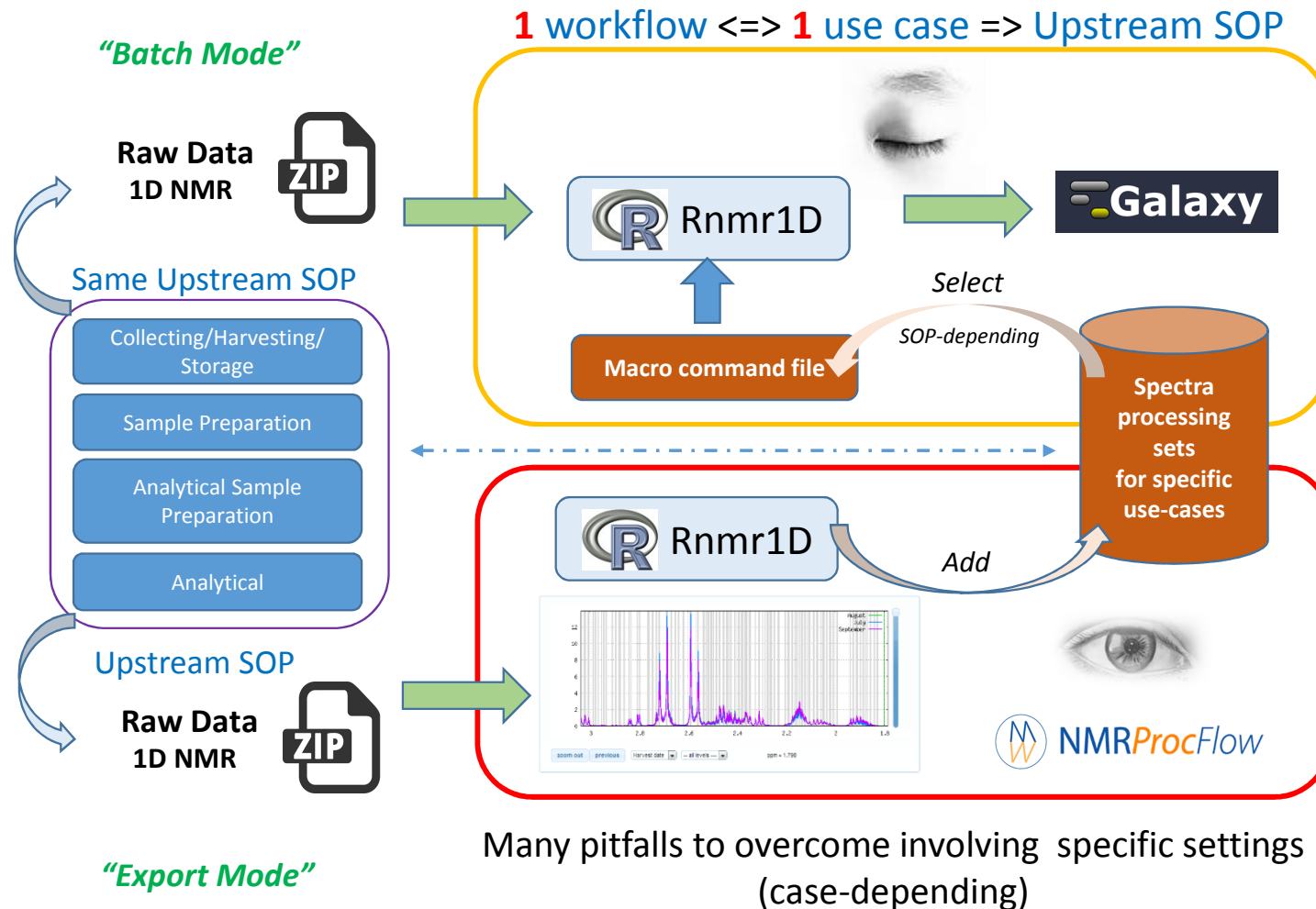
NMRProcFlow & Galaxy – Combine strength of each one



Motivation for using Galaxy (in that way)

- Tuning of the spectra processing applied on a subset of the NMR spectra (~ 100) within NMRProcFlow,
- Then applying the processing upon the whole set (i.e. > 300) within Galaxy

NMRProcFlow & Galaxy – Combine strength of each one



Motivation for using Galaxy (in that way)

- Expert settings checked and validated for each corresponding use case
- Allow non-experts to handle their data **with confidence and efficiency**

Work Package 9: Tools, Workflows, Audit and Data Management

WP9 (led by [IPB](#)) aims to develop and maintain the primary scientific- and technological tools and corresponding interfaces. We will support the data standards defined by WP8 and facilitate the interoperability of tools both within this consortium and those externally developed by the community. We will establish [distributed tools for phenomics, metabolomics and bioinformatics processing pipelines and workflows](#), including longitudinal primary research data management (continuous availability to avoid data lock-in) and data audit mechanisms, as well as quality assurance schemes. Thus, this work package will produce several tailored VMIs which will be the basis for the service activities in WP5.

Objectives:

- [Specify and integrate software pipelines and tools utilised in the PhenoMeNal e-Infrastructure into VMIs](#), adhering to data standards developed in WP8 and supporting the interoperability and federation middleware developed in WP5. Most tools will be already available (see table 1.1) and we will develop new applications to complete ‘missing links’ in pipelines. Although two explicit releases for VMIs are listed as deliverables below, we will use public repositories and continuous integration to always provide development snapshots of the infrastructure VMIs.
- [Develop methods to scale-up software pipelines for high-throughput analysis](#), supporting distributed execution on e.g. local clusters, private clouds, federated clouds, or GRIDs.
- [Add quality control and quality assurance to pipelines to ensure high quality and reliable data](#), keep an audit trail of intermediate steps and results.
- Develop methods to present and summarize the results of the pipelines in biomedical and disease contexts.

Work Package 9: Tools, Workflows, Audit and Data Management


Deliverables:

- Report on existing software tools, workflows and analytical pipelines initially supported in the PhenoMeNal grid – D9.1
- PhenoMeNal-Preprocess Virtual Machine Image to enable data producers to locally process raw data into standard formats supported in PhenoMeNal – D9.2.1
- PhenoMeNal-Data Virtual Machine image to enable sharing and dissemination of standardised and processed omics data to participating online repositories, like MetaboLights – D9.2.2
- Services Virtual Machine Image to facilitate the PhenoMeNal toolsets and pipelines, both locally and in the grid – D9.2.3
- Compute Virtual Machine Image to enable standardised compute capabilities for all the grid supplying partners – D9.2.4

Work Package Title	WP9 Tools, Workflows, Audit and Data Management
Deliverable Title	D9.2.1 <i>PhenoMeNal-Preprocess</i> Virtual Machine Image to enable data producers to locally process raw data into standard formats supported in PhenoMeNal
Delivery Date	M12
Work Package leader	IPB
Contributing Partners	EMBL-EBI, IPB, CRS4, UU, SIB, CEA, ICL, UB
Authors	Kristian Peters, Daniel Schober, Steffen Neumann, Pablo Moreno, Reza Salek, Luca Pireddu, Rico Rueedi, Roger Mallol, Ibrahim Karaman, Tim Ebbels, Pedro de Atauri, Etienne Thévenot, Pierrick Roger
Abstract: This deliverable report on the development of container images of tools to convert raw data into standard formats and their integration into the PhenoMeNal build and deployment infrastructure.	

D9.2.1 - 3.5.Integration of tools into Workflows





Analyze Data Workflow Shared Data Visualization Help Login or Register

Tools

search tools

Get Data

Text Manipulation

Filter and Sort

Join, Subtract and Group

Statistics

Graph/Display Data

PHENOMENAL H2020 TOOLS

Transfer

Fluxomics

NMR

mtbIs nmr raw importer

Imports RAW data for further conversion with nmrm1conv.

nmrm1conv

Converts vendor RAW NMR to nmrm1.

ZIP nmrm1 Collection

Create a zip archive of a nmrm1 collection.

nmrm1d


Preprocessing RAW NMR data with macro-commands.

nmrm1d-stacked-plot

Draw stacked plot of NMR FID.

BATMAN

BATMAN

















Large-Scale Computing for Medical Metabolomics

The [PhenoMeNal](#) Galaxy installation allows users to access all of the PhenoMeNal containerised tools through a workflow environment, on a scalable infrastructure that can be deployed to public and private cloud installations.

This [PhenoMeNal H2020](#) Galaxy instance, and all of its tools, run as containers on top of [Kubernetes](#), an open source container orchestrator system backed by Google. If you wish to deploy the [PhenoMeNal Galaxy installation](#) on top of your own Kubernetes instance, you can find instructions at our [wiki](#).

The [PhenoMeNal consortium](#) is driven by 14 European research groups with strong experience in the development of tools and methods for large data acquisition, integration and analysis for metabolic phenotypes, genome and cross-omics data.



PhenoMeNal is funded by European Commission's Horizon2020 programme, grant agreement number 654241. The [Galaxy Project](#) is supported in part by NHGRI, NSF, [The Huck Institutes of the Life Sciences](#), [The Institute for CyberScience at Penn State](#), and [Johns Hopkins University](#).

<http://public.phenomenal-h2020.eu/>

D9.2.1 - 3.5.NMR Workflow

Galaxy | Analyze Data | Workflow | Shared Data | Visualization | Help | Login or Register

Tools

search tools

Get Data

Text Manipulation

Filter and Sort

Join, Subtract and Group

Statistics

Graph/Display Data

PHENOMENAL H2020 TOOLS

Transfer

Fluxomics

NMR

[mtbls_nmr_raw_importer](#)
Imports RAW data for further conversion with nmrmlconv.

[nmrmlconv](#)
Converts vendor RAW NMR to nmrML.

[ZIP nmrML Collection](#)
Create a zip archive of a nmrML collection.

[rnmr1d](#)
Preprocessing RAW NMR data with macro-commands.

[rnmr1d-stacked-plot](#)
Draw stacked plot of NMR FID.

[BATMAN](#) BATMAN

PhenoMeNal
Large-Scale Computing for Medical Metabolomics

Workflow:

- mtbls_nmr_raw_dummy_importer** (Input: nmr_raw)
- nmrmlconv** (Input: Input file in zip format to convert; Output: outfile (txt))
- ZIP nmrML Collection** (Input: Collection data; Output: outfile (no_unzip.zip))
- rnmr1d** (Input: Zipped nmrML file collection, Macro-Command file; Output: outfile (txt))
- rnmr1d-stacked-plot** (Input: NMR data matrix; Output: outfile (pdf))

Macro command file (Input to rnmr1d)

nmrML
Converter for 1D NMR

Rnmr1D

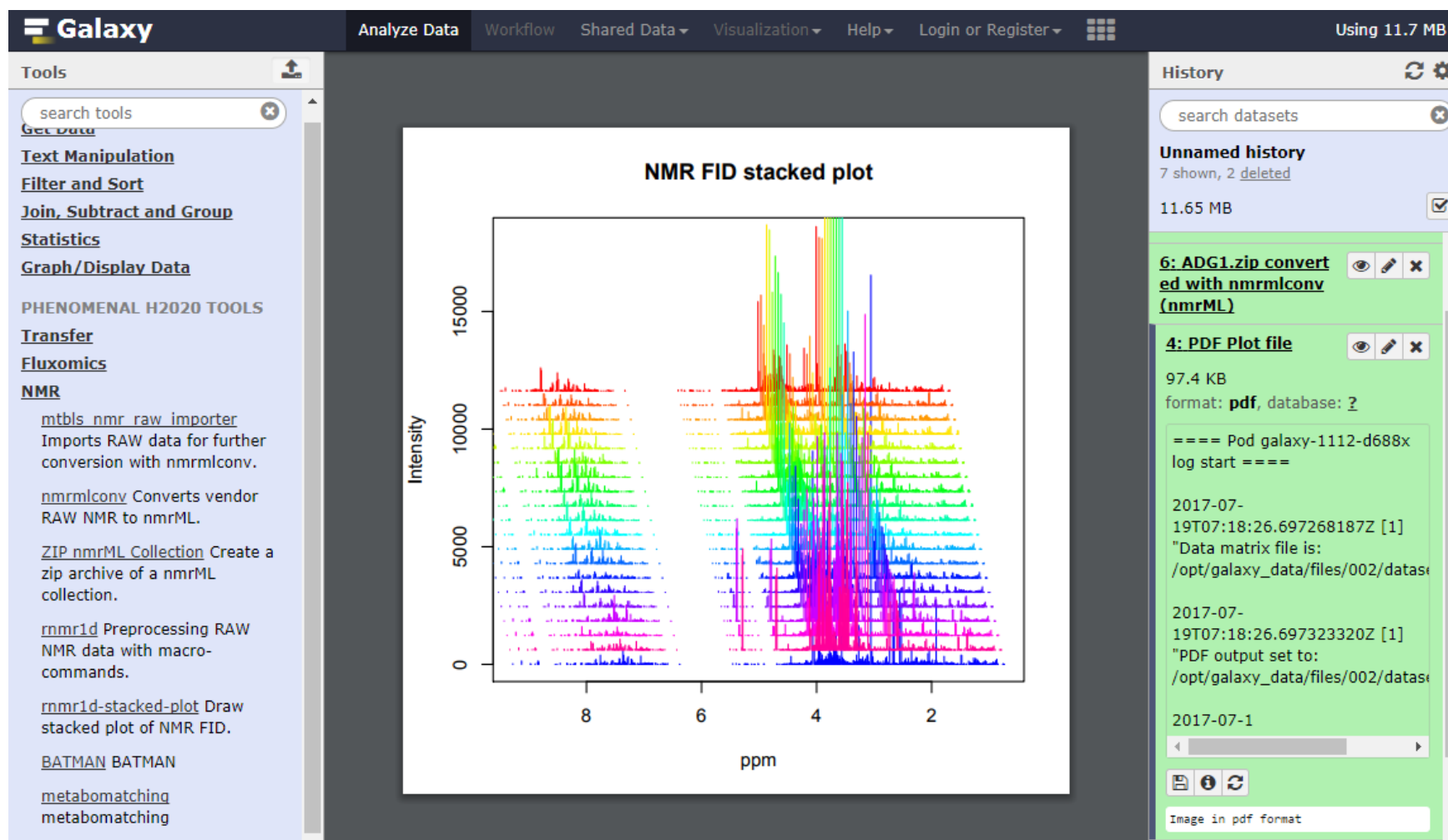
<https://github.com/phnmnl/container-rnmr1d>



Thanks to
Kristian Peters

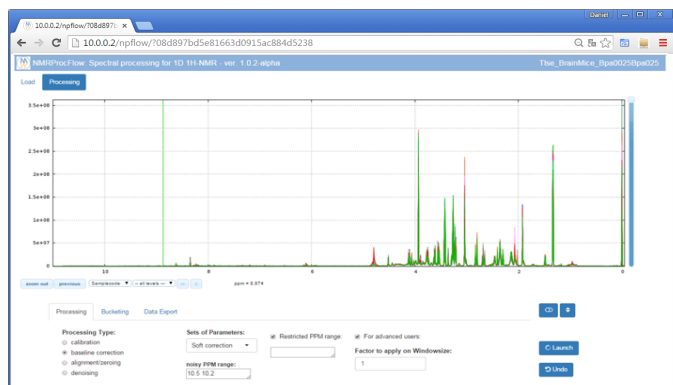


D9.2.1 - 3.5.NMR Workflow



Thanks to
Kristian Peters





*An efficient GUI tool
for spectra processing
from 1D 1H-NMR metabolomics data*
<http://nmrprocflow.org/>

Daniel Jacob



<http://phenomenal-h2020.eu/home/>

*combined with the computing power of
a Cloud Research Environment*

**Thank you for
your attention**

Thanks to
Kristian Peters

