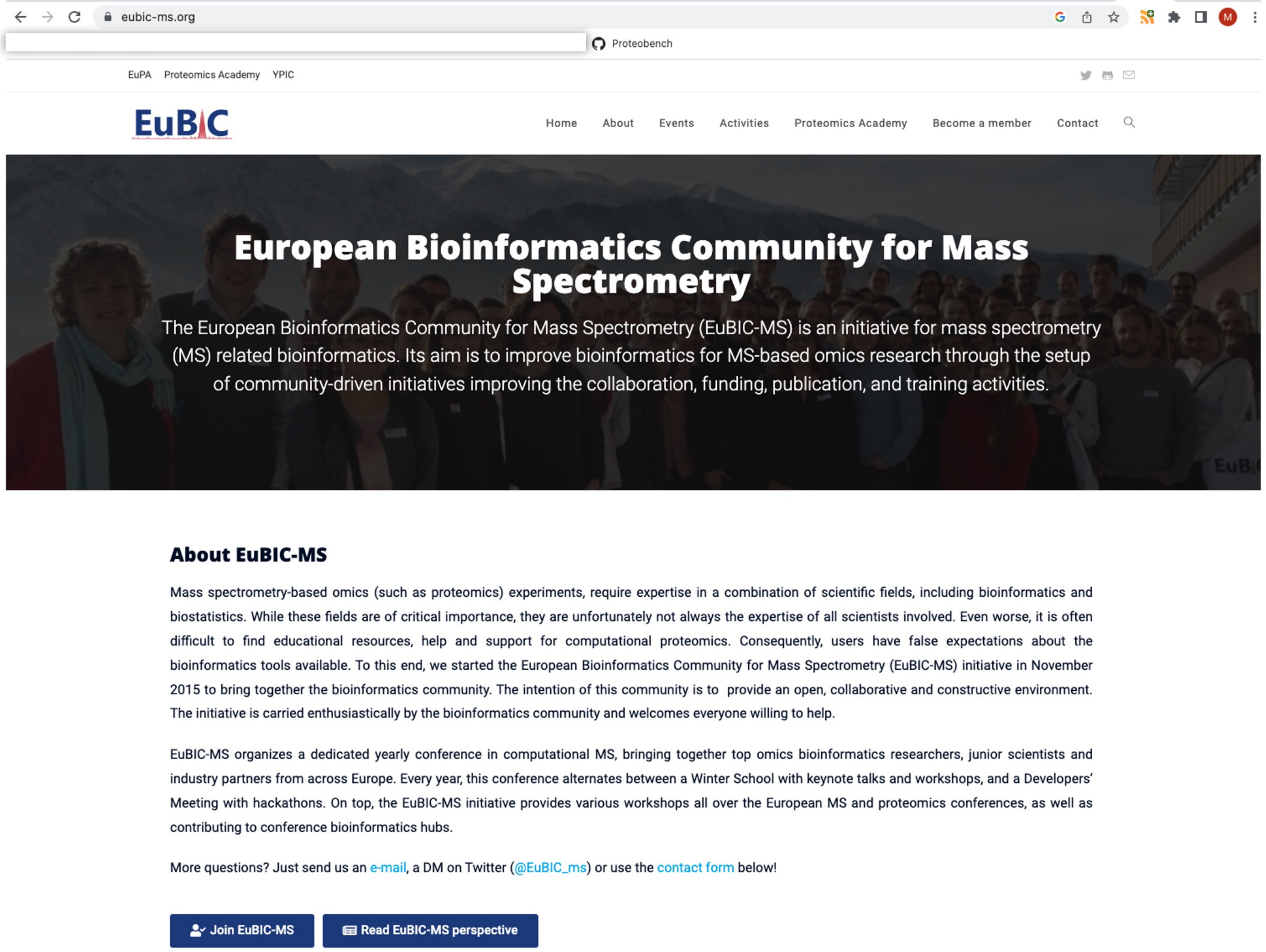
EuBIC

ProteoBench

Open platform for community-curated benchmarks for computational proteomics

ProFI webinar – 02/06/2025

EuPA
EUROPEAN PROTEOMICS ASSOCIATION



European Bioinformatics Community for Mass Spectrometry

The European Bioinformatics Community for Mass Spectrometry (EuBIC-MS) is an initiative for mass spectrometry (MS) related bioinformatics. Its aim is to improve bioinformatics for MS-based omics research through the setup of community-driven initiatives improving the collaboration, funding, publication, and training activities.

About EuBIC-MS

Mass spectrometry-based omics (such as proteomics) experiments, require expertise in a combination of scientific fields, including bioinformatics and biostatistics. While these fields are of critical importance, they are unfortunately not always the expertise of all scientists involved. Even worse, it is often difficult to find educational resources, help and support for computational proteomics. Consequently, users have false expectations about the bioinformatics tools available. To this end, we started the European Bioinformatics Community for Mass Spectrometry (EuBIC-MS) initiative in November 2015 to bring together the bioinformatics community. The intention of this community is to provide an open, collaborative and constructive environment. The initiative is carried enthusiastically by the bioinformatics community and welcomes everyone willing to help.

EuBIC-MS organizes a dedicated yearly conference in computational MS, bringing together top omics bioinformatics researchers, junior scientists and industry partners from across Europe. Every year, this conference alternates between a Winter School with keynote talks and workshops, and a Developers' Meeting with hackathons. On top, the EuBIC-MS initiative provides various workshops all over the European MS and proteomics conferences, as well as contributing to conference bioinformatics hubs.

More questions? Just send us an e-mail, a DM on Twitter (@EuBIC_ms) or use the contact form below!

eubic-ms.org

Next EuBIC-MS Winter School: 19-23 January 2026 in the Czech Republic

EUROPEAN PROTEOMICS ASSOCIATION

MS-based proteomics data analysis workflows

MS-based proteomics data analysis workflows

Many available workflows

High complexity

Evolves very fast

MS-based proteomics data analysis workflows

Many available workflows

High complexity

Evolves very fast

How do these compare?

🤨

MS-based proteomics data analysis workflows

Many papers comparing software tools

From developers

From users

MS-based proteomics data analysis workflows

Many papers comparing software tools

From developers

From users

New version?
New software?

Frozen

Outdated quite fast

MS-based proteomics data analysis workflows

Many papers comparing software tools

From developers

From users

New version?
New software?

Frozen

Outdated quite fast

Many groups compare new workflows in house

ProteoBench

Discuss how to compare software tool performances

Perform comparisons in an homogenized fashion

Share data

Continuous comparison of data analysis workflows used by the community

EuBIC

EuPA
EUROPEAN PROTEOMICS ASSOCIATION

CORE FOR LIFE

DDSA
Danish
Data Science
Academy

fwo

ipbs
CANCER
INFECTION
INFLAMMATION

CompOmics

VIB

Medical Bioinformatics
Ruhr University Bochum

The James Hutton Institute
Scientific Services

mpc
MEDIZINISCHES PROTEOM-CENTER

AD Lab

Metaproteomics Initiative
www.metaproteomics.org

cnrs

GHENT
UNIVERSITY

Novo Nordisk Foundation
CENTER FOR BASIC METABOLIC RESEARCH
UNIVERSITY OF COPENHAGEN

Novo Nordisk Foundation
Center for Protein Research

ProFI
PROTEOMICS

UNIVERSITY
OF BERGEN

MAX PLANCK INSTITUTE
OF BIOCHEMISTRY

DTU

University
of Antwerp

ProteoToul

Université
de Toulouse

PERUTZ
MASS
SPEC
FACILITY

PAPPSO

INSTITUT
Pasteur

SIB
Swiss Institute of
Bioinformatics

EMBL

ETH zürich

University of
Zurich UZH

SDU

dkfz.
GERMAN
CANCER RESEARCH CENTER
IN THE HELMHOLTZ ASSOCIATION
Research for a Life without Cancer

University
of Basel

BIOZENTRUM
The Center for
Molecular Life Sciences

netherlands
eScience center

RUHR
UNIVERSITÄT
BOCHUM
RUB

UNIVERSITÉ
LAVAL

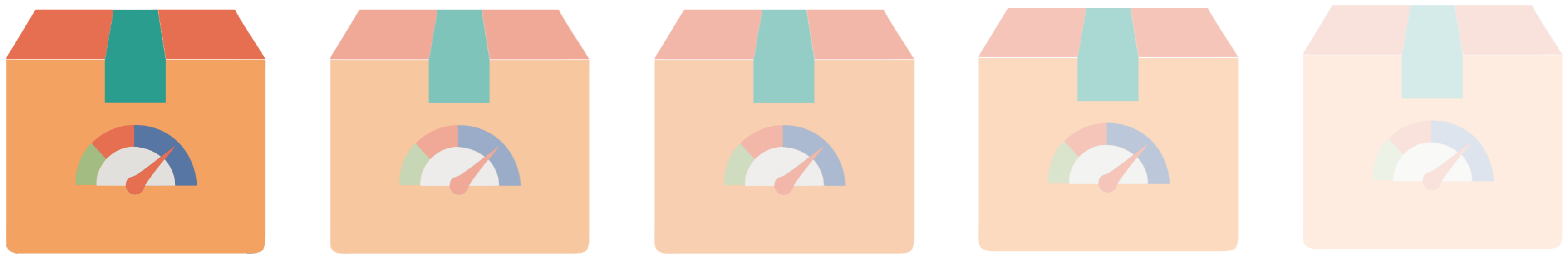
fgcz
functional
genomics
center
zurich

fhoö
UNIVERSITY
OF APPLIED SCIENCES
UPPER AUSTRIA

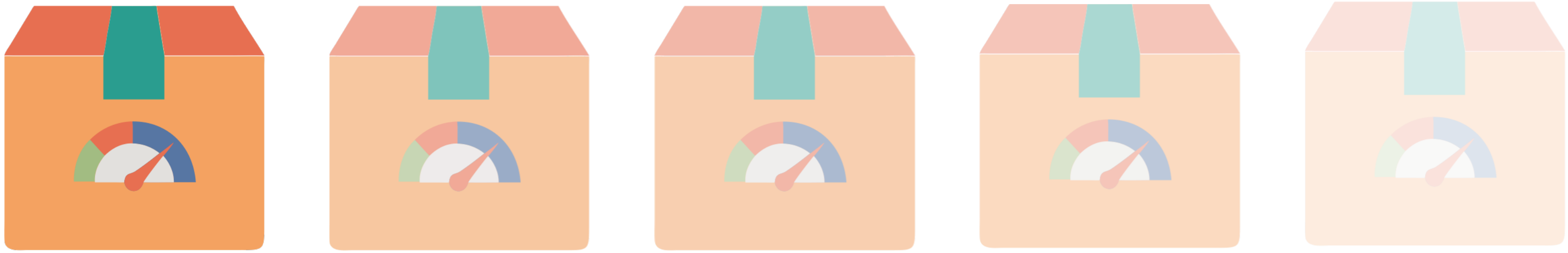
marie.locard-paulet@ipbs.fr
robbin.bouwmeester@ugent.be

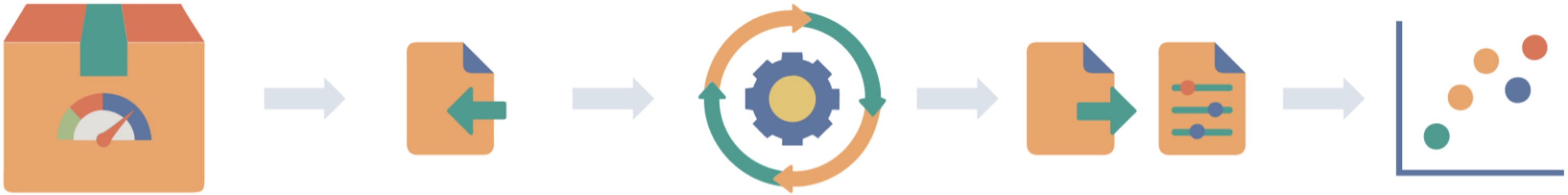
ProteoBench structure

One **module** per comparison (with its own dedicated data)



One **module** per comparison (with its own dedicated data)





choose a module

download the input data

run your workflow

upload the results

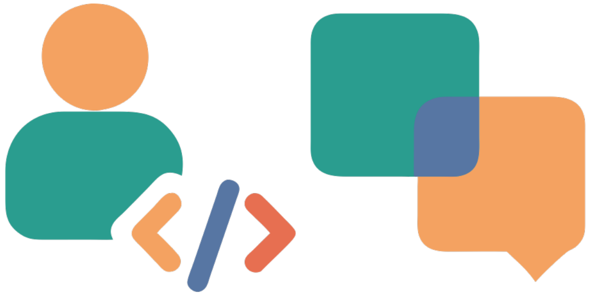
compare your workflow run

One data set

One main figure

Contribution of the entire community

(discussion on metrics, development, curation…)

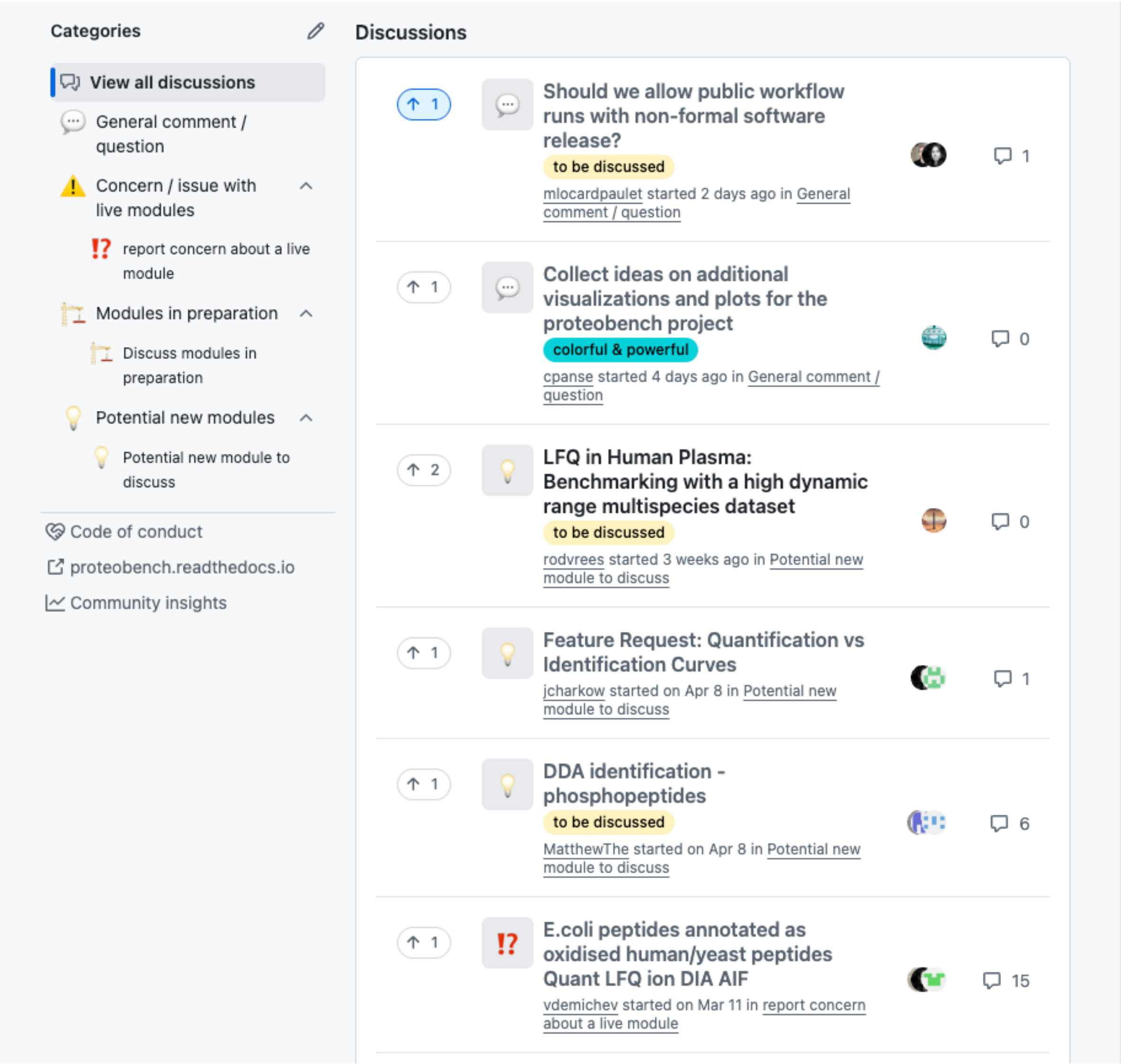


Add workflow outputs

Discuss how to make the comparisons

Design new modules

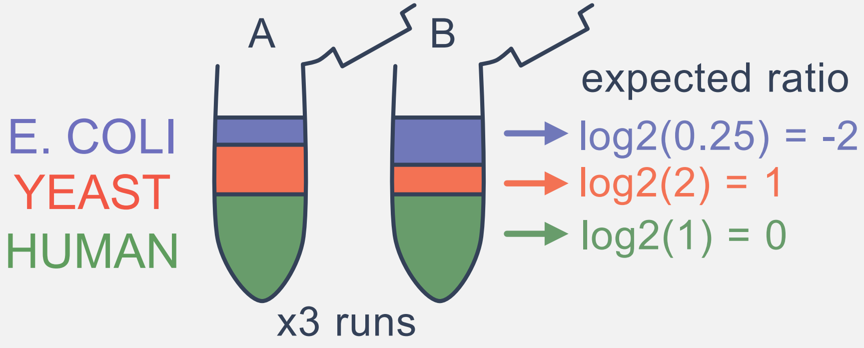
https://github.com/orgs/Proteobench/discussions



- ↑ 1 — **Should we allow public workflow runs with non-formal software release?** — to be discussed — mlocardpaulet started 2 days ago in General comment / question — 1
- ↑ 1 — **Collect ideas on additional visualizations and plots for the proteobench project** — colorful & powerful — cpanse started 4 days ago in General comment / question — 0
- ↑ 2 — **LFQ in Human Plasma: Benchmarking with a high dynamic range multispecies dataset** — to be discussed — rodvrees started 3 weeks ago in Potential new module to discuss — 0
- ↑ 1 — **Feature Request: Quantification vs Identification Curves** — jcharkow started on Apr 8 in Potential new module to discuss — 1
- ↑ 1 — **DDA identification - phosphopeptides** — to be discussed — MatthewThe started on Apr 8 in Potential new module to discuss — 6
- ↑ 1 — **E.coli peptides annotated as oxidised human/yeast peptides Quant LFQ ion DIA AIF** — vdemichev started on Mar 11 in report concern about a live module — 15

Precursor quantification with DDA

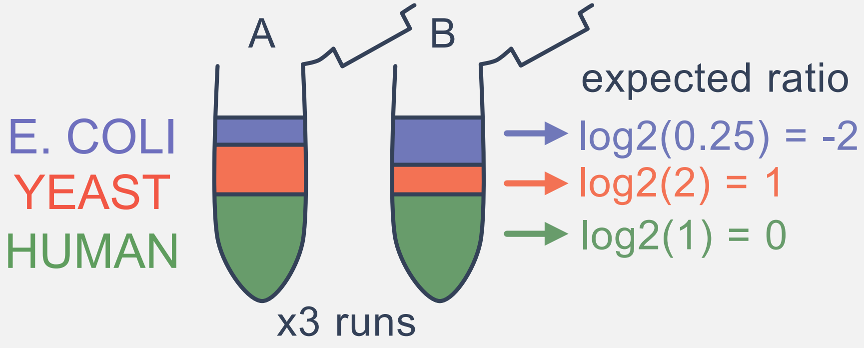
Precursor quantification with DDA



6 raw files*
FASTA file (3 species + contaminants**)

* Van Puyvelde et al., 2022, Scientific Data
** Frankenfield et al., 2022, J Proteome Research

Precursor quantification with DDA



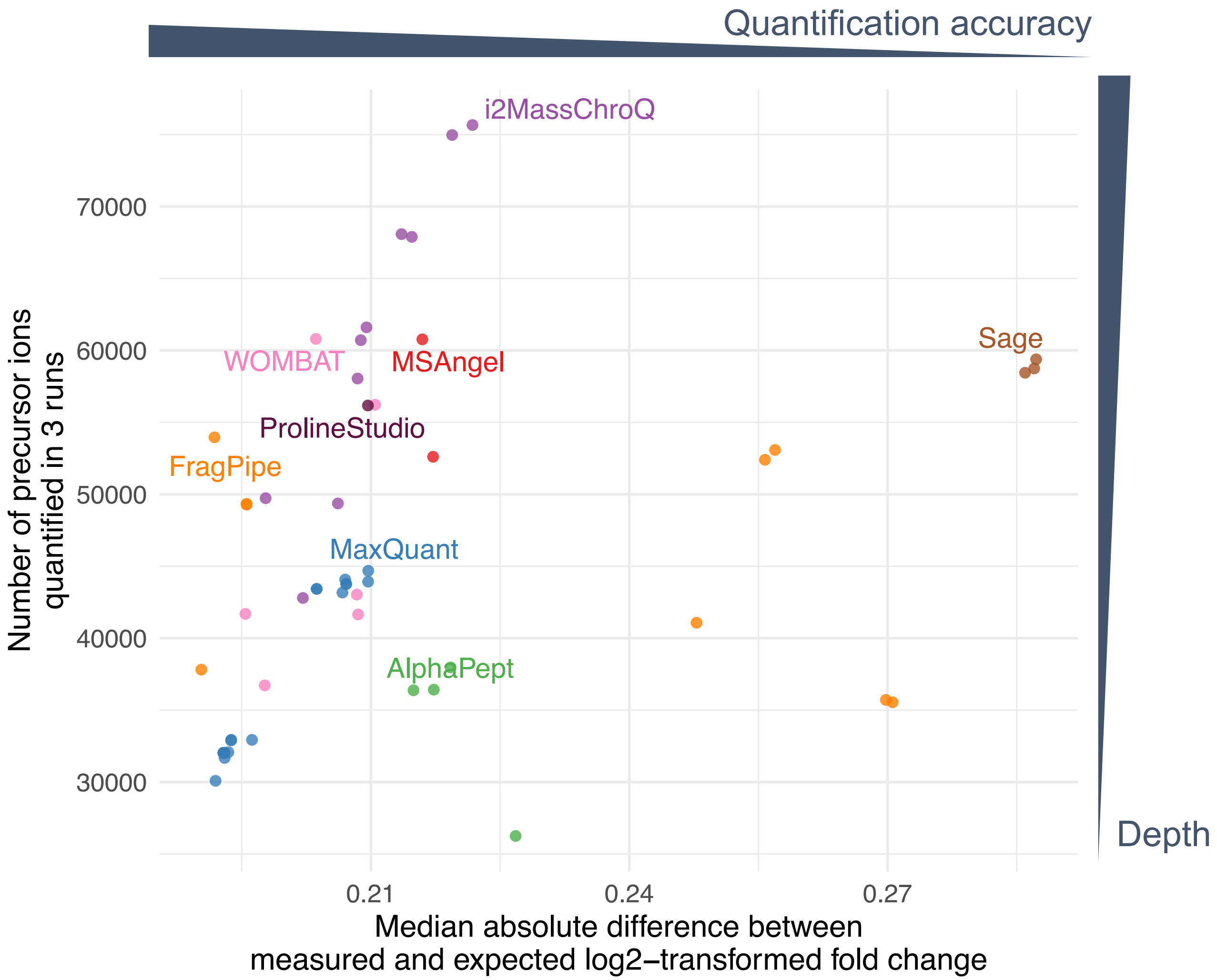
E. COLI
YEAST
HUMAN

A
B

expected ratio
log2(0.25) = -2
log2(2) = 1
log2(1) = 0

x3 runs

6 raw files*
FASTA file (3 species + contaminants**)



---

* Van Puyvelde et al., 2022, Scientific Data
** Frankenfield et al., 2022, J Proteome Research

ProteoBench

explore

validate

benchmark

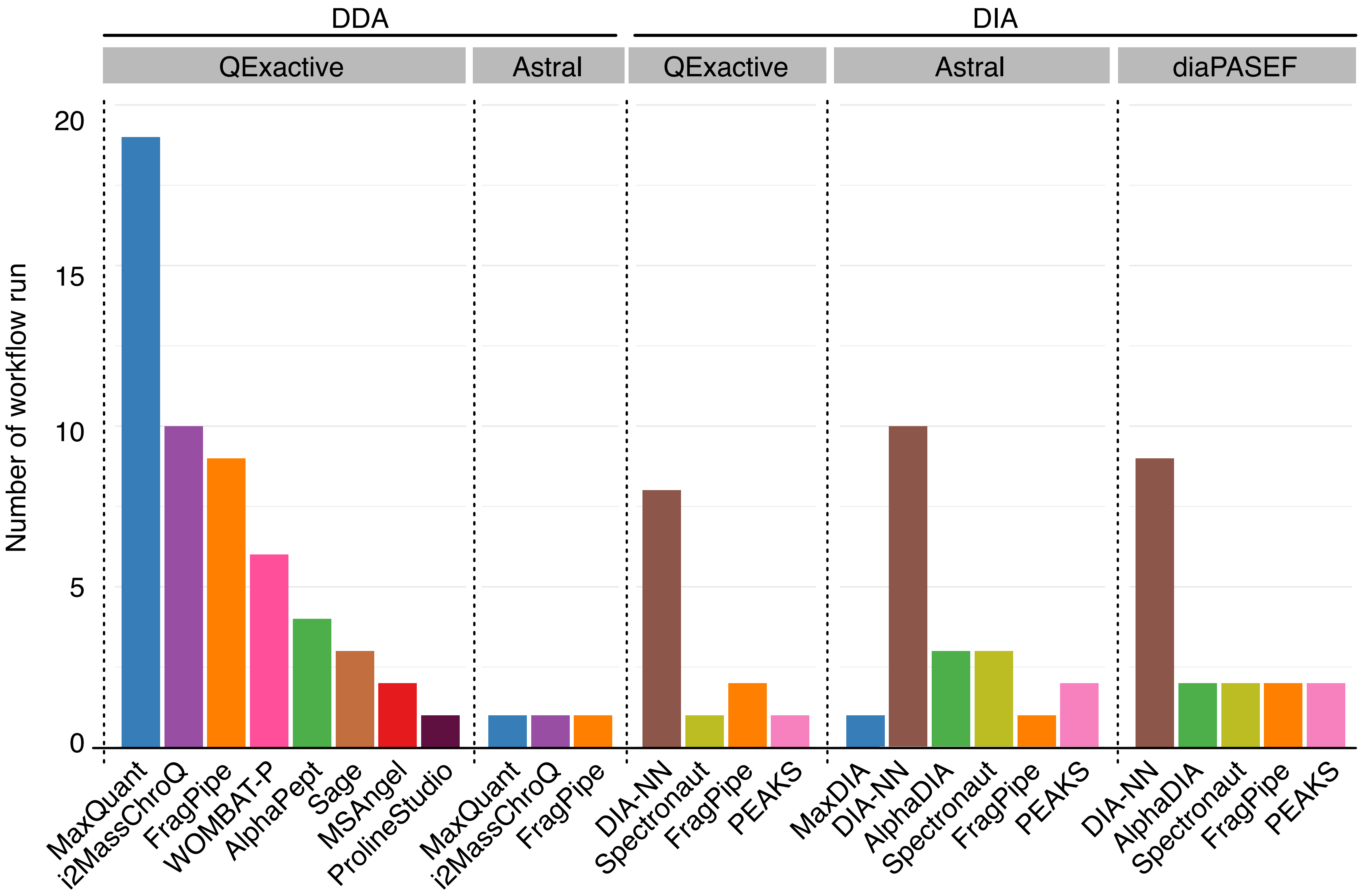
compare

collaborate

contribute

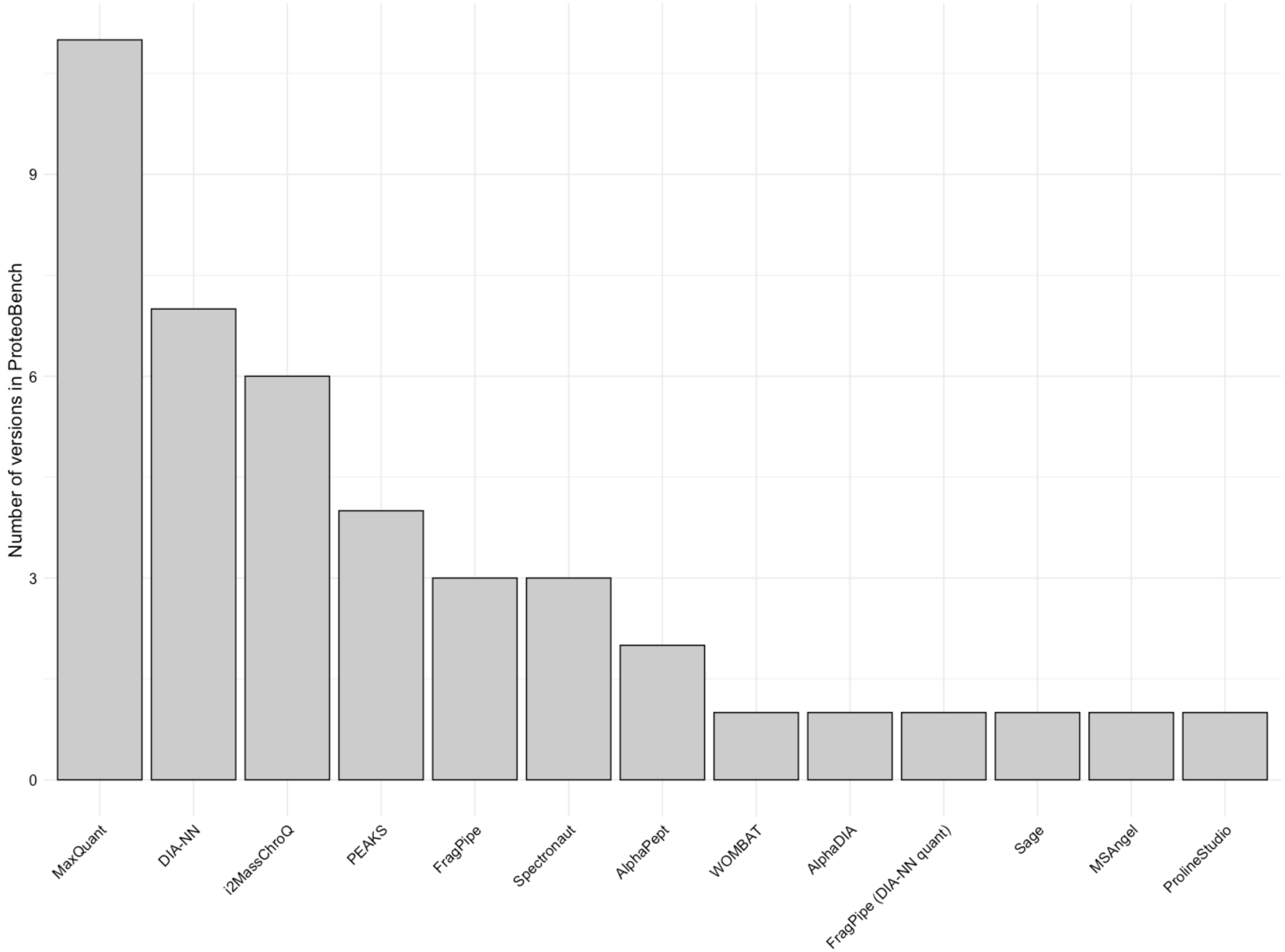
ProteoBench is still in development: there may be some disruption in the next few weeks.

Number of points: precursor quantification

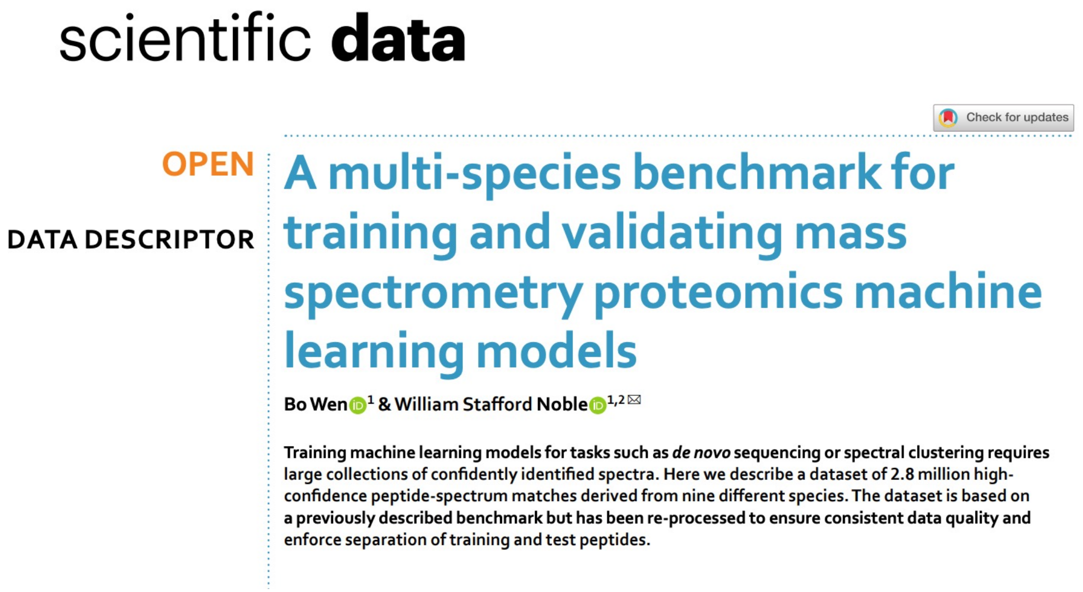


27/05/2025

Software versions



New modules in preparation



scientific **data**

Check for updates

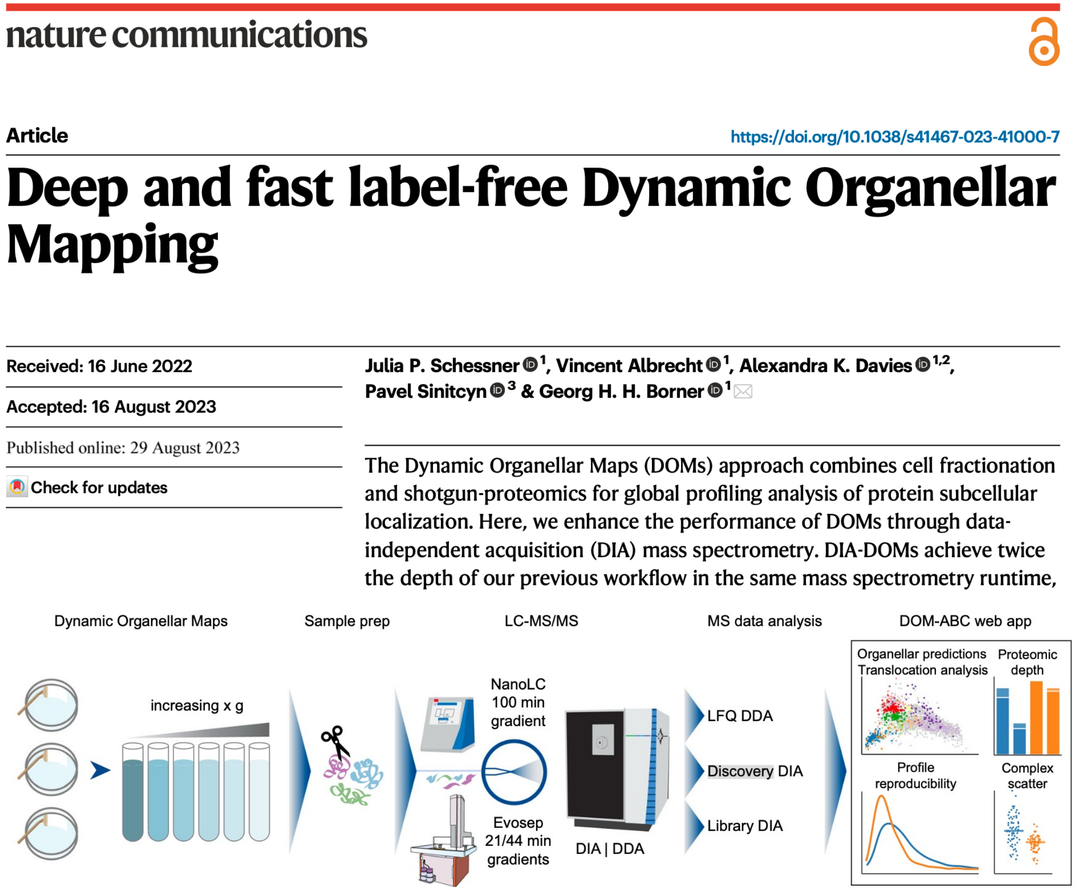
OPEN

DATA DESCRIPTOR

A multi-species benchmark for training and validating mass spectrometry proteomics machine learning models

Bo Wen[1] **& William Stafford Noble**[1,2]

Training machine learning models for tasks such as *de novo* sequencing or spectral clustering requires large collections of confidently identified spectra. Here we describe a dataset of 2.8 million high-confidence peptide-spectrum matches derived from nine different species. The dataset is based on a previously described benchmark but has been re-processed to ensure consistent data quality and enforce separation of training and test peptides.



nature communications

Article

https://doi.org/10.1038/s41467-023-41000-7

Deep and fast label-free Dynamic Organellar Mapping

Received: 16 June 2022

Accepted: 16 August 2023

Published online: 29 August 2023

Check for updates

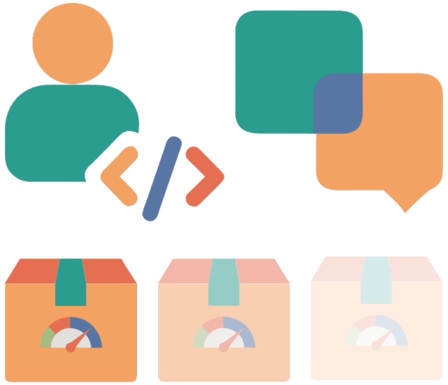
Julia P. Schessner[1], **Vincent Albrecht**[1], **Alexandra K. Davies**[1,2], **Pavel Sinitcyn**[3] **& Georg H. H. Borner**[1]

The Dynamic Organellar Maps (DOMs) approach combines cell fractionation and shotgun-proteomics for global profiling analysis of protein subcellular localization. Here, we enhance the performance of DOMs through data-independent acquisition (DIA) mass spectrometry. DIA-DOMs achieve twice the depth of our previous workflow in the same mass spectrometry runtime,

ProteoBench

Open public discussions for the choice of data (sample, instrument, acquisition methods, …) and metrics/plots.

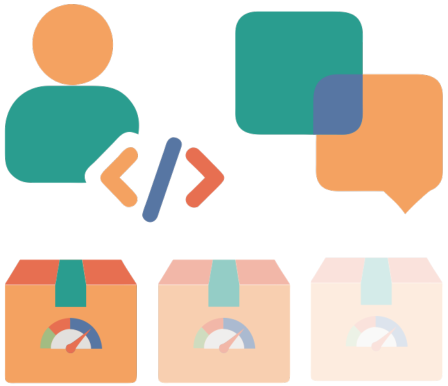
ProteoBench



Open public discussions for the choice of data (sample, instrument, acquisition methods, …) and metrics/plots.

Several modules allow complementary metrics and dataset properties

ProteoBench

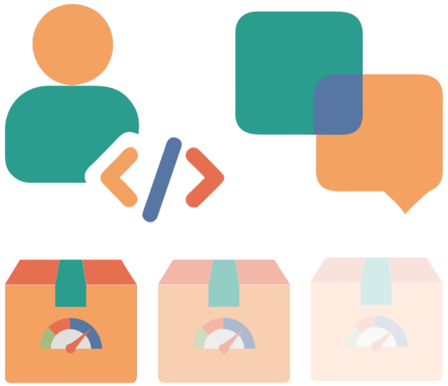


Open public discussions for the choice of data (sample, instrument, acquisition methods, …) and metrics/plots.

Several modules allow complementary metrics and dataset properties

Developers have access to raw data and other tools outputs

ProteoBench



Open public discussions for the choice of data (sample, instrument, acquisition methods, …) and metrics/plots.

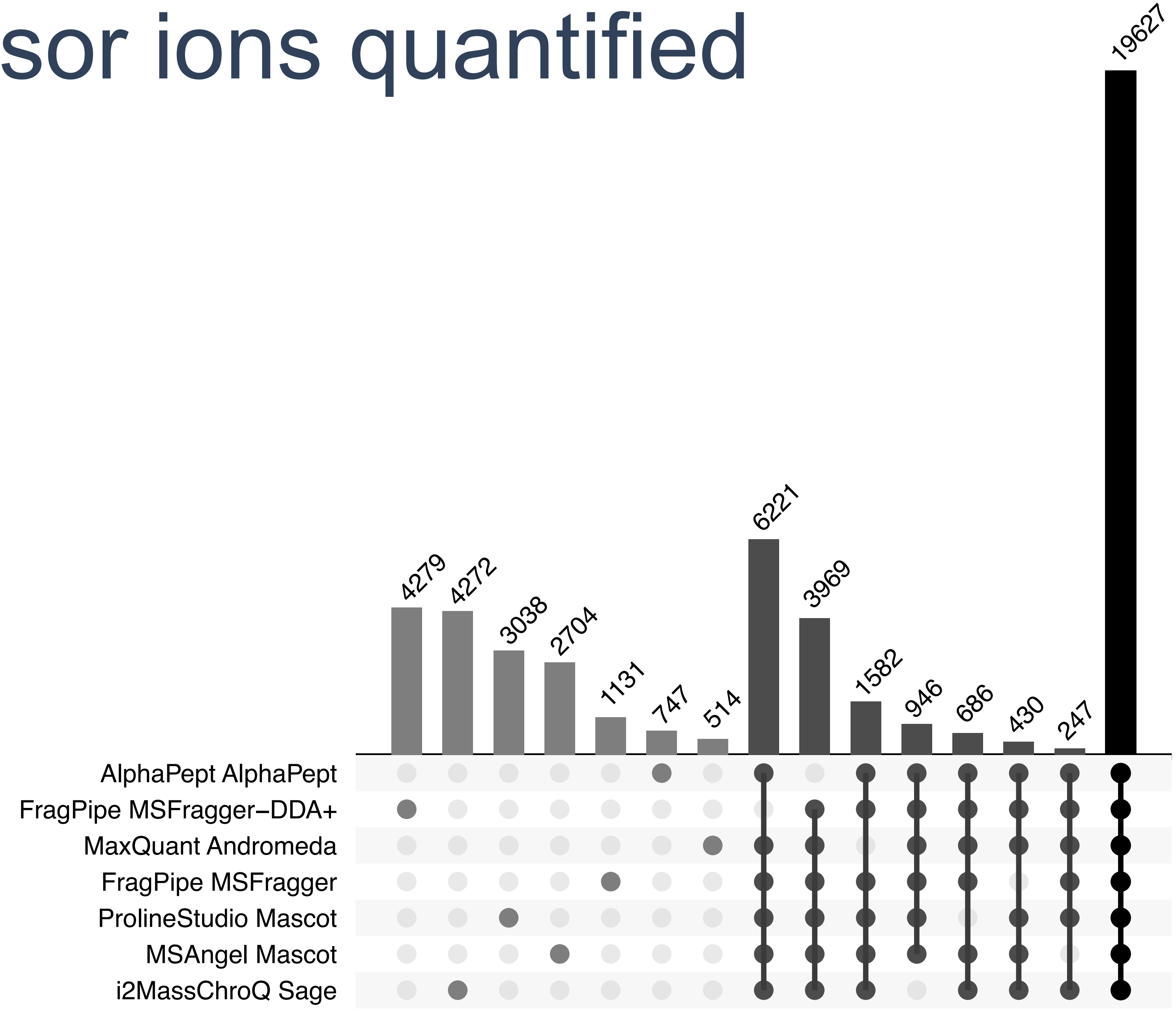
Several modules allow complementary metrics and dataset properties

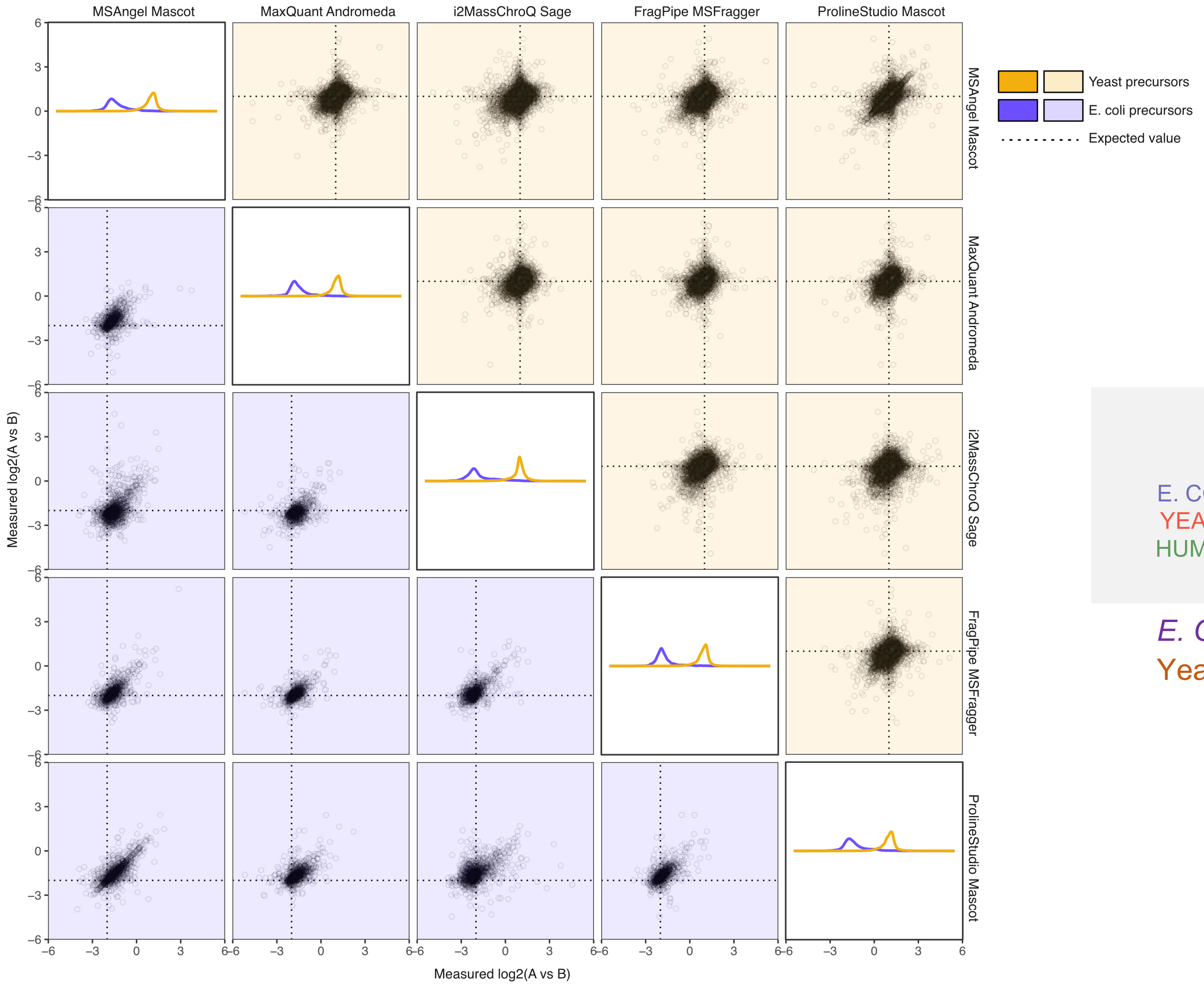
Developers have access to raw data and other tools outputs

Users can evaluate changes between versions, parameters…

Going deeper with the available data

Precursor ions quantified





A
B
expected ratio

E. COLI
YEAST
HUMAN

log2(0.25) = -2
log2(2) = 1
log2(1) = 0

x3 runs

E. Coli = 5% in A and 20% in B

Yeast = 30% in A and 15% in B

ProteoBench

EuBIC

EuPA
EUROPEAN PROTEOMICS ASSOCIATION

CORE FOR LIFE

Danish
Data Science
Academy

fwo

ipbs CANCER INFECTION INFLAMMATION
CompOmics
VIB
Medical Bioinformatics Ruhr University Bochum
The James Hutton Institute Scientific Services
mpc MEDIZINISCHES PROTEOM-CENTER
AD Lab

Metaproteomics Initiative www.metaproteomics.org
cnrs
GHENT UNIVERSITY
Novo Nordisk Foundation CENTER FOR BASIC METABOLIC RESEARCH
UNIVERSITY OF COPENHAGEN
Novo Nordisk Foundation Center for Protein Research
ProFI PROTEOMICS

UNIVERSITY OF BERGEN
MAX PLANCK INSTITUTE OF BIOCHEMISTRY
DTU
University of Antwerp

ProteoToul
Université de Toulouse
PERUTZ MASS SPEC FACILITY
PAPPSO
INSTITUT Pasteur

SIB Swiss Institute of Bioinformatics
EMBL
ETH zürich
University of Zurich UZH
SDU

dkfz. GERMAN CANCER RESEARCH CENTER IN THE HELMHOLTZ ASSOCIATION
Research for a Life without Cancer
University of Basel
BIOZENTRUM The Center for Molecular Life Sciences
netherlands eScience center

RUHR UNIVERSITÄT BOCHUM RUB
UNIVERSITÉ LAVAL
fgcz functional genomics center zurich
fhoö UNIVERSITY OF APPLIED SCIENCES UPPER AUSTRIA

marie.locard-paulet@ipbs.fr
robbin.bouwmeester@ugent.be