ABRF 2019 ANNUAL MEETING

30 years of challenging the limits of science and technology, opening doors for the future

Sunday
2:30 pm
Rm 214D

# The PRG DIA Study: Goals, study design and participation

ABRF Proteomics Research Group

https://abrf.org/research-group/proteomics-research-group-prg

# Goal of Study

Our goal was to eliminate barriers to adoption and demonstrate where DIA is across platforms and cores.

# The Proteomics Research Group

*Current members:*

- Pratik Jagtap (Chair) — University of Minnesota
- Laura Herring — University of North Carolina at Chapel Hill
- Joanna Kirkpatrick — Leibniz Institute on Aging, Germany
- LeRoy Martin III — Waters Corporation
- Mukul Midha — Institute for Systems Biology
- Benjamin Neely — National Institute of Standards and Technology
- Brett Phinney — University of California Davis
- Baozhen (Paul) Shan — Bioinformatics Solutions, Inc.
- Paul Stemmer — Wayne State University
- Yan Wang — University of Maryland
- Allis Chien (EB-liaison) — Stanford University

- *Contact prg.abrf@gmail.com*

# Goal of Study

## Specific Goals

- Provide baseline methods across platforms
- Create and distribute a test sample that can benchmark performance
- Recruit diverse platforms and skill levels
- Collect data with goal of making public
- Analyze data along with industry partners

# Intended Benefits of Study

## Participant Benefits

- Develop working DIA methods
- Defined sample for self-evaluation
- Use available software to process data
- Compare your work anonymously to others in the area

## Community Benefits

- Multi-platform multi-laboratory DIA data on the same sample set
- Anonymous DIA data made publicly available to help algorithm, workflow, and application development and benchmarking
- Serve as demo dataset for new-comers

# ABRF Study Timeline

## Sample

Blank 25 fmol 100 fmol

HeLa digest spiked with four non-endogenous proteins and iRT

## Data Acquisition

- Acquisition parameters were provided
- Total DIA time < 24 h
- Given enough sample to generate library OR spectral library available for download

## Software

Skyline

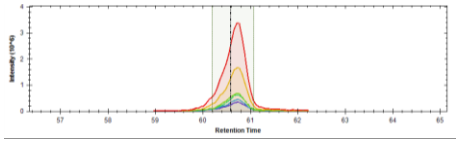
PEAKS

OpenSWATH

TPP

Participants had the option of using trial licenses from Spectronaut, Scaffold-DIA, and PEAKS DIA

## Data analysis & interpretation



- PRG members analyzed data
- Participants encouraged to analyze as well

October 2018 | November 2018 | December 2018 | January 2019 | February 2019 | March 2019

# Spiked proteins

**Sample**

Blank 25 fmol 100 fmol

HeLa digest spiked with four non-endogenous proteins and iRT

ABRF-1: beta-galactosidase; 1024 aa; 116.5 kDa

ABRF-2: lysozyme C; 147 aa; 16.2 kDa

ABRF-3: glucoamylase; 640 aa; 68.3 kDa

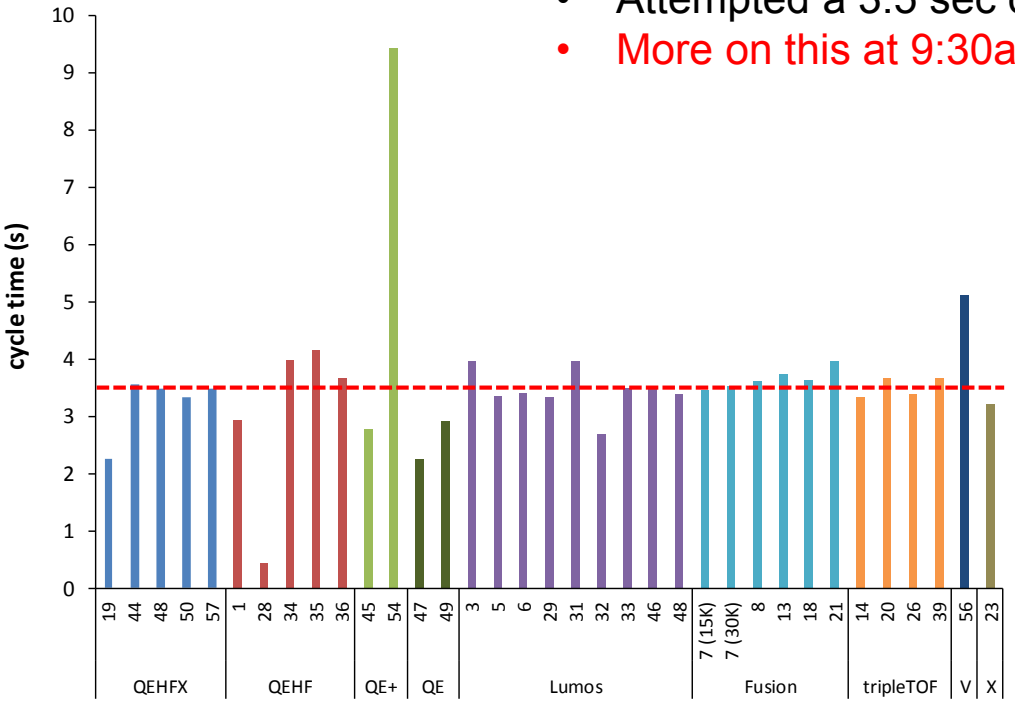
ABRF-4: Protein G; 185 aa; 20.1kDa

Sample A: 25 fmol/µg HeLa digest
Sample B: 100 fmol/µg HeLa digest
Sample C: blank (just HeLa digest)

# Data Acquisition

**Data Acquisition**

- Acquisition parameters were provided
- Total DIA time < 24 h
- Given enough sample to generate library OR spectral library available for download




- An “adequate” base DIA method
- Attempted a 3.5 sec cycle
- More on this at 9:30am

*available at https://github.com/neely/PRG2018 or https://www.lcmsmethods.org/*

# Software and Data Analysis


Skyline
Scaffold DIA
PEAKS
OpenSWATH
TPP

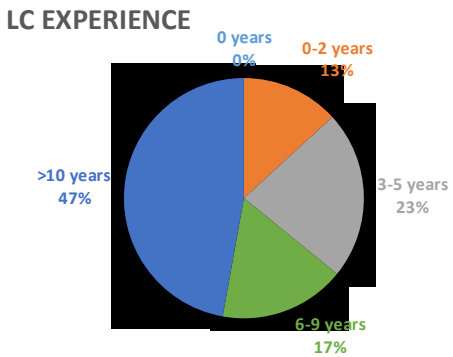
Participants uploaded data to MassIVE

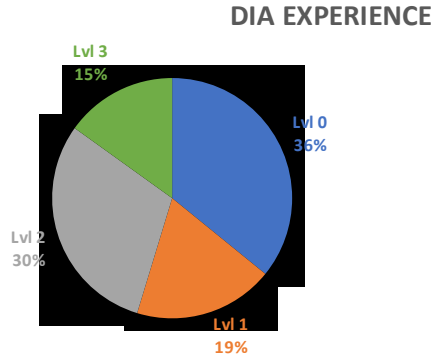
Participants were eligible for extended licenses to commercial software and non-commercial is always available

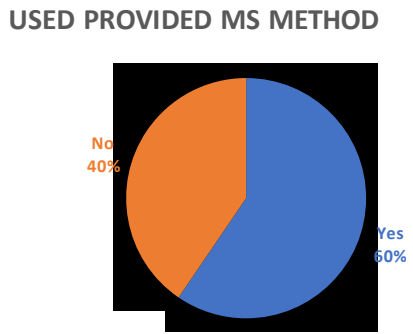
# Study Participants: 63 labs, 20 countries, 16 US States

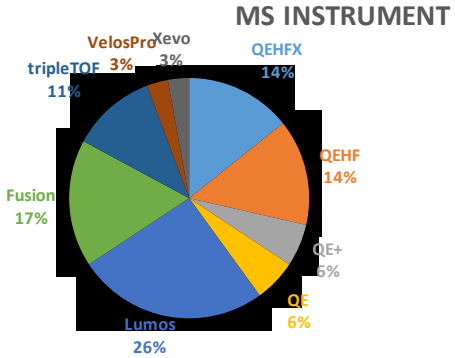
# Study Participants

- 40 (63%) participants deposited data
- 35 data sets used for prelim analysis
- 53 survey responses (84%)
- Experience was broad
- Majority used provided acquisition method
- Majority of MS platforms were Thermo



# Future Plans and Data Availability

- Continue to look into the nuanced results of the study
- Some trends are consistent within platforms, and that users were consistent across performance metrics
  - if you had “good” DPPP and “good” high protein IDs, then you likely did well at everything else
- Summarize results into manuscripts
- Anonymize raw data and make available before June via MassIVE
- Alert software makers (commercial and non-commercial) to data availability to help with development and education