

Human Proteome Project Knowledge Base Pillar Current Plans and Future

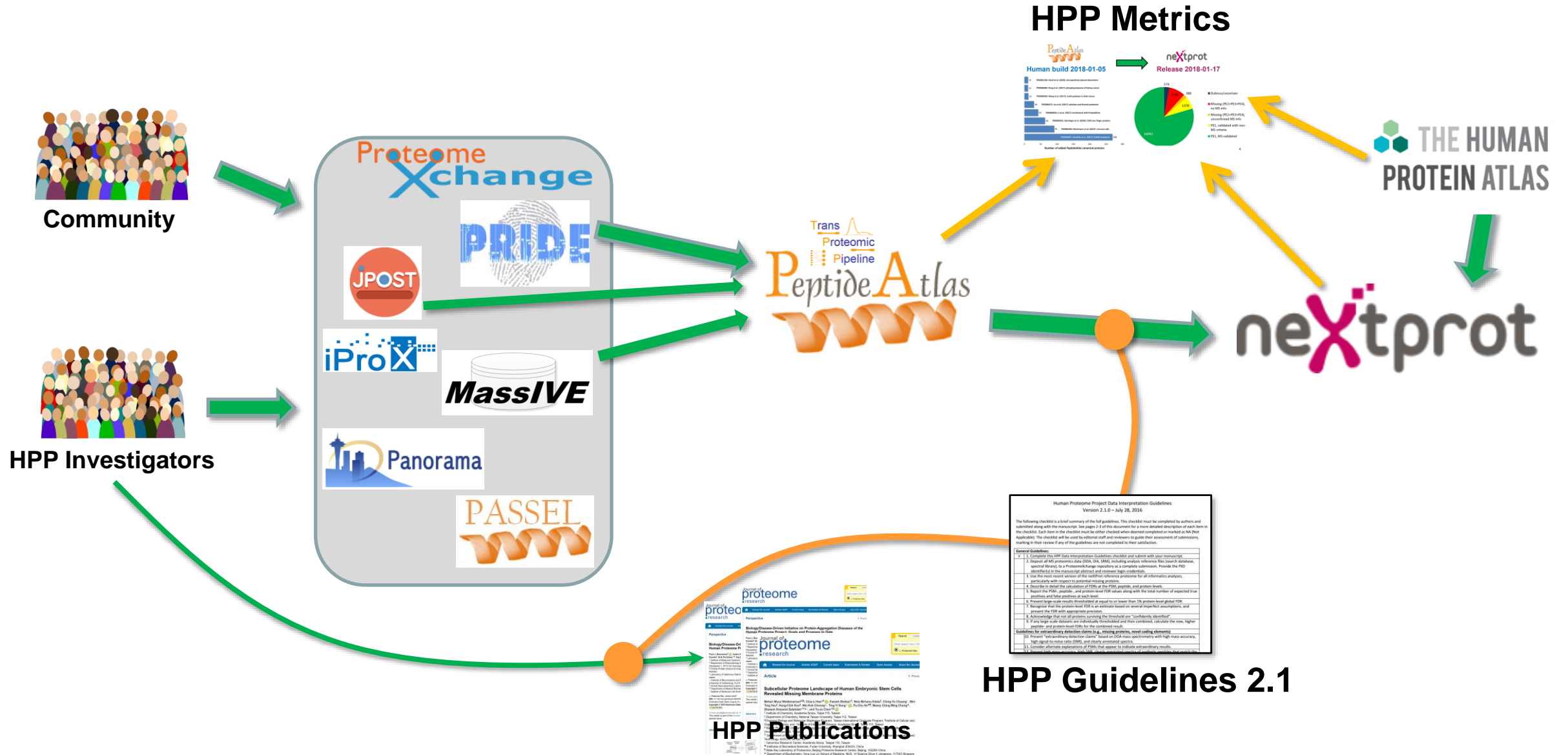
Eric Deutsch

Institute for Systems Biology

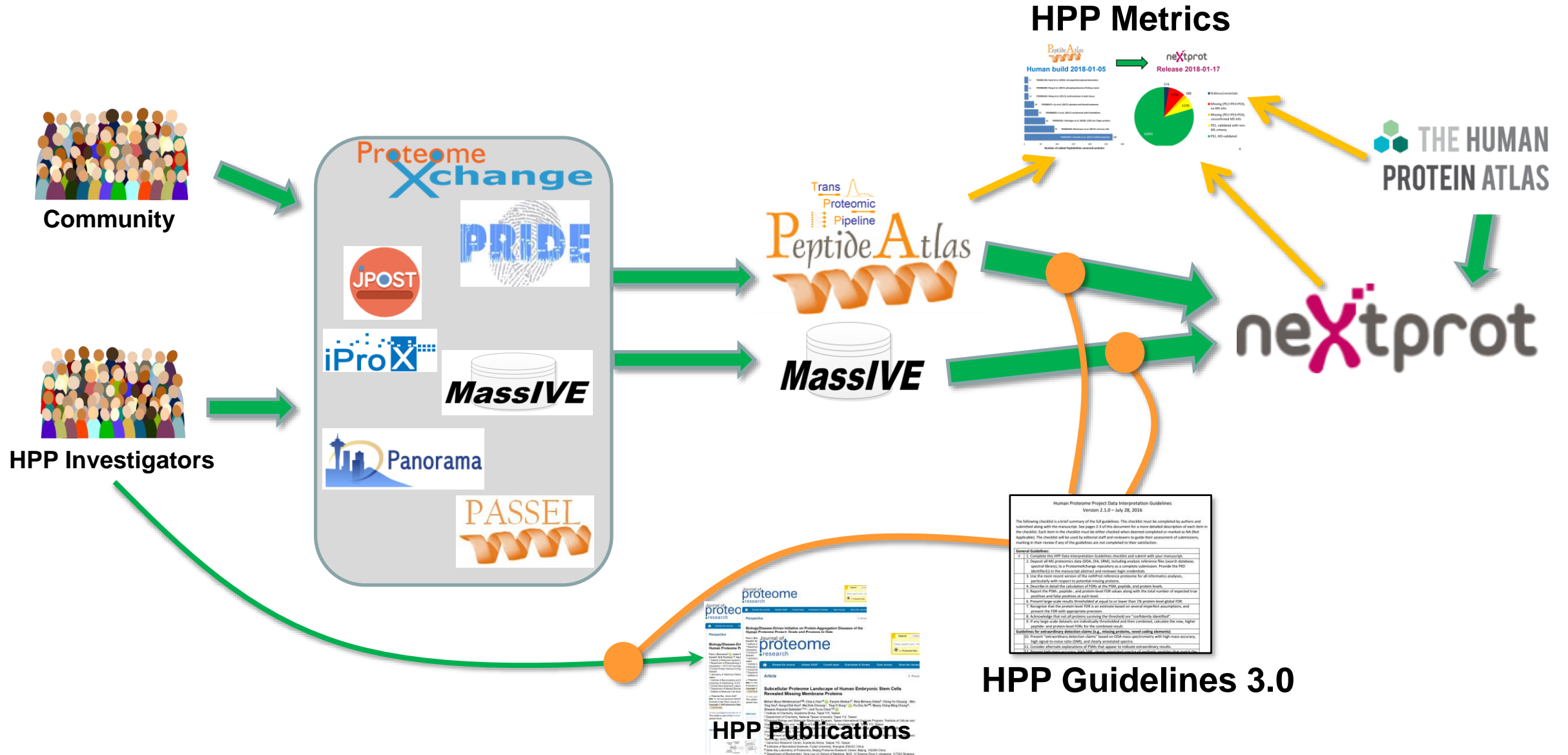
Lydie Lane

Swiss Institute of Bioinformatics

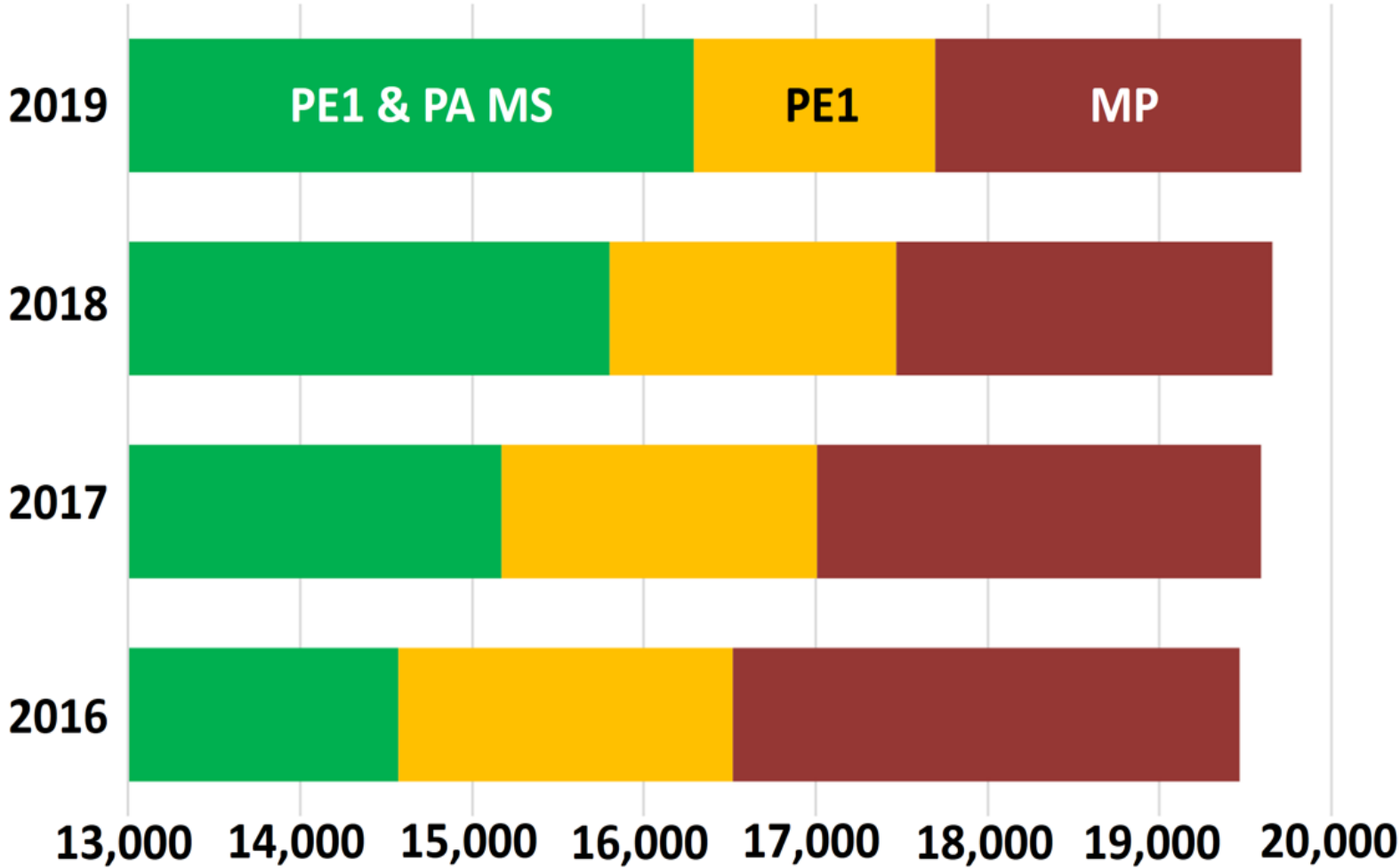
High-level HPP MS Data Workflow 2019



High-level HPP MS Data Workflow 2020 and on

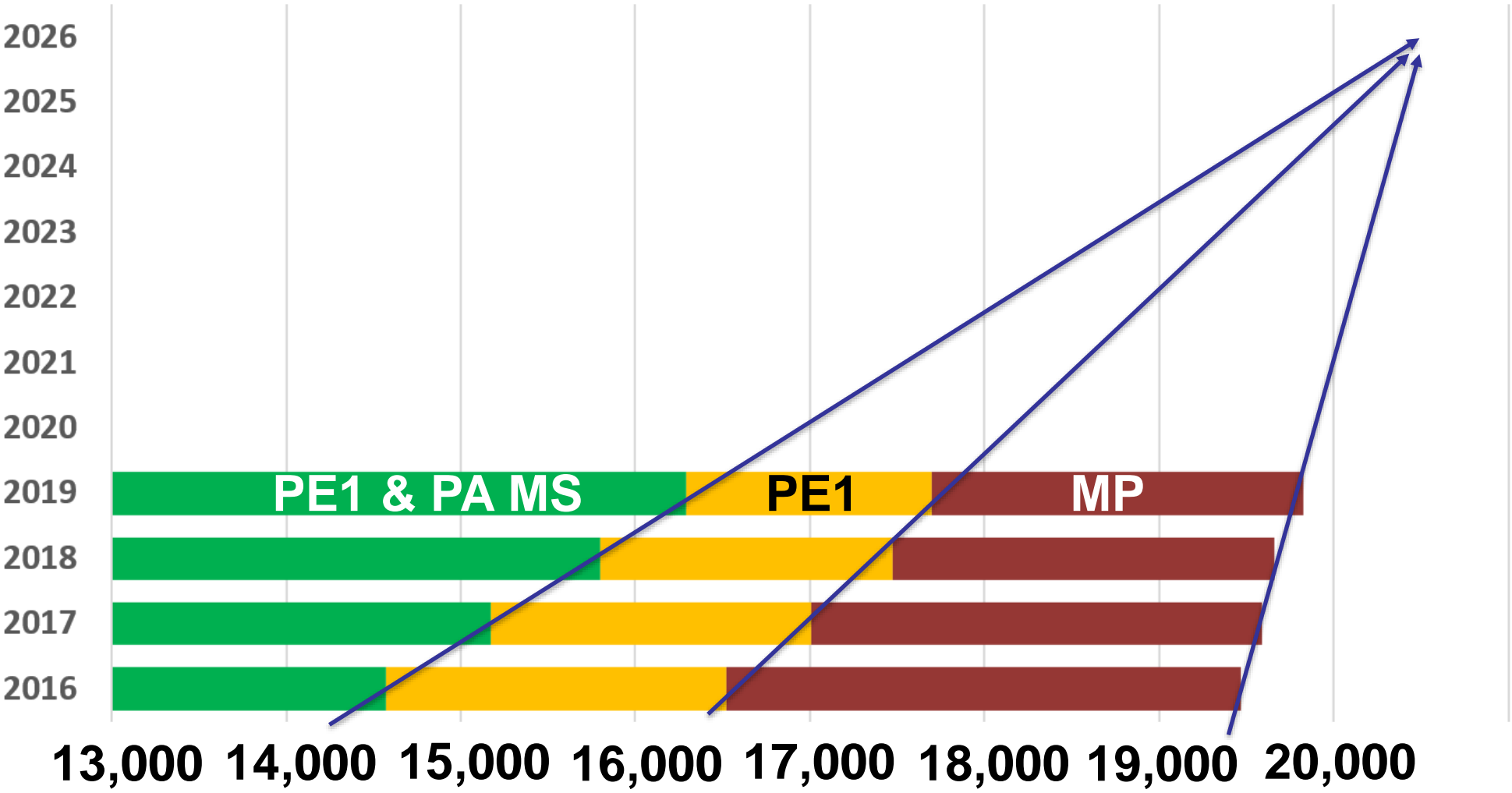


HPP Knowledge Base Metrics



Wild Extrapolation with Linear Trendlines

Can we maintain the trend? Accelerate the trend?



A synthetic peptide spectrum match for all crucial evidence

- Synthetic peptide spectra are a requirement for the published claims of all new published claims of new PE1 protein detection claims via MS
- However, many PE1 proteins in PeptideAtlas have peptides that meet guidelines except that they have no corresponding synthetic peptides
- We plan to leverage all existing synthetic peptide spectra from SRMAtlas, ProteomeTools, etc. to provide comparison spectra in PeptideAtlas, MassIVE, neXtProt for all crucial evidence peptides
- Hundreds (thousands?) of crucial peptides will be missing synthetic partners. Who will generate them?

A special panel within the KB Pillar

- The remaining MPs are becoming progressively more difficult to detect
- Some MPs already have convincing evidence, but do not meet the guidelines as written
- The guidelines implement a crude rule: 2x 9 AA uniquely mapping, non-nested peptides (& 18 AA coverage)
- But for uniform confidence, larger proteins should require more evidence than very small ones, since they are more likely to gather false positives
- Subset of the KB Pillar will form a special panel to review proteins for which existing evidence is quite good but yet do not meet the letter of the guidelines
- Manually add and remove PE1 status for some proteins after manual review
- Learn how to set better thresholds and rules for acceptance

HPP Knowledge Base Pillar Outlook

- PeptideAtlas will continue to progress human datasets and release every January
- neXtProt aiming to continue to release every Feb plus 1-2 other times per year
- *if* funding can be maintained

- Planned inclusion of MassIVE ProteinExplorer into KB workflow for 2020 and beyond

- New 2020 HPP MS Data Interpretation Guidelines and continued refinement

- Obtain synthetic peptide spectrum matches for *all* crucial evidence

- Subset of the KB Pillar will form a special panel to review borderline proteins and learn how to set better thresholds and rules for acceptance