PTMProphet + Magnum: Localization of Open-Mass Search Adducts and Mass Differences in Search Engine Results

David D. Shteynberg1, Michael R. Hoopmann1, Eric W. Deutsch2, Luis Mendoza1, Michael Riffle2, Zhi Sun1, Alex Zelter2, Nina Isoherranen2, and Robert L. Moritz1

1Institute for Systems Biology, Seattle WA  2University of Washington, Seattle WA

Overview

• Search algorithms often excel at assigning the peptide sequence, though not always the correct position of post-translational modifications (PTMs) and other mass adducts contained in the peptide
• Magnum is a search engine optimized for open-mass searching that can identify peptides containing unknown adducts and unlocalized mass differences within the user defined mass range
• Trans-Proteomic Pipeline (TPP) software tool, PTMProphet, can be applied toward the localization of known and unknown PTM modifications within a given peptide-spectrum-match (PSM) as identified and annotated by Magnum

PTMProphet Methodology

PTMProphet starts with PSMs collected in a pepXML file. User defined variable modifications are specified to PTMProphet at runtime. The MASSDIFFMODE option tells PTMProphet to localize mass differences and any annotated adducts anywhere on the peptide. PTMProphet computes the Bayesian probability that the PTM occurs at that site in the PSM given the available fragment peak evidence. All probabilities are re-normalized so the sum of the probabilities for each type of PTM equals the total number of modifications of that type in the PSM. For each type of PSM evaluated, PTMProphet computes the information content of the localization statistics and estimates the number of PTMs of each type that can be localized with certainty in the PSM.

ALGORITHM:
1. Compute Probability for each potential PTM site for each PTM type
2. Apply Expectation / Maximization until probabilities remain constant
3. Normalize all probabilities by the number of modifications in the peptide
4. Record the output in the pepXML

Support & Information

Support provided by:
• NIH, NIGMS grants: R01GM087221, R24GM127667
• NHLBI grant: R21HL133335
• NHLBI grant: R01HL133135

TPP + Open Mod Supported By Limelight:
https://limelight.ms.org/
TPP Resources:
http://www.tppms.org

Magnum: Trans-Proteomic Pipeline Open Search

MassDiffs and Adducts
• Magnum encodes massdiffs and mass adducts in pepXML
• mass adducts are massdiffs assigned to an amino acid
• When a PSM has an adduct it absorbs the massdiff
• mass of adduct becomes 0
• MassProphet has been updated such that when it sees an annotated mass adduct it uses that in place of the massdiff for the PSM for the purposes of modeling the mass differences
• PTMProphet has been updated to localize mass diffs or adducts together with variable PTMs specified by the users for localization
• Massdiffs and adducts are localized to each possible amino acid on the peptide
• The "EXCLUDEMASSDIFFMIN=", "EXCLUDEMASSDIFFMAX=" options together specify a window of mass diffs to exclude from localizing
• mass diffs that are the result of measurement or isotope error can be excluded
• PTMProphet annotates the localized masses using pepXML source= attribute as one of: "param", "massdiff" or "adduct"

PTMProphet + Magnum: Localization of Open-Mass Search Adducts and Mass Differences in Search Engine Results

David D. Shteynberg1, Michael R. Hoopmann1, Eric W. Deutsch2, Luis Mendoza1, Michael Riffle2, Zhi Sun1, Alex Zelter2, Nina Isoherranen2, and Robert L. Moritz1

1Institute for Systems Biology, Seattle WA  2University of Washington, Seattle WA

Outline:

Overview

PTMProphet Methodology

ALGORITHM:
1. Compute Probability for each potential PTM site for each PTM type
2. Apply Expectation / Maximization until probabilities remain constant
3. Normalize all probabilities by the number of modifications in the peptide
4. Record the output in the pepXML

Support & Information

Support provided by:
• NIH, NIGMS grants: R01GM087221, R24GM127667
• NHLBI grant: R21HL133335
• NHLBI grant: R01HL133135

TPP + Open Mod Supported By Limelight:
https://limelight.ms.org/
TPP Resources:
http://www.tppms.org

Magnum: Trans-Proteomic Pipeline Open Search

MassDiffs and Adducts
• Magnum encodes massdiffs and mass adducts in pepXML
• mass adducts are massdiffs assigned to an amino acid
• When a PSM has an adduct it absorbs the massdiff
• mass of adduct becomes 0
• MassProphet has been updated such that when it sees an annotated mass adduct it uses that in place of the massdiff for the PSM for the purposes of modeling the mass differences
• PTMProphet has been updated to localize mass diffs or adducts together with variable PTMs specified by the users for localization
• Massdiffs and adducts are localized to each possible amino acid on the peptide
• The "EXCLUDEMASSDIFFMIN=", "EXCLUDEMASSDIFFMAX=" options together specify a window of mass diffs to exclude from localizing
• mass diffs that are the result of measurement or isotope error can be excluded
• PTMProphet annotates the localized masses using pepXML source= attribute as one of: "param", "massdiff" or "adduct"

PTMProphet + Magnum: Localization of Open-Mass Search Adducts and Mass Differences in Search Engine Results

David D. Shteynberg1, Michael R. Hoopmann1, Eric W. Deutsch2, Luis Mendoza1, Michael Riffle2, Zhi Sun1, Alex Zelter2, Nina Isoherranen2, and Robert L. Moritz1

1Institute for Systems Biology, Seattle WA  2University of Washington, Seattle WA

PTMProphet Methodology

PTMProphet starts with PSMs collected in a pepXML file. User defined variable modifications are specified to PTMProphet at runtime. The MASSDIFFMODE option tells PTMProphet to localize mass differences and any annotated adducts anywhere on the peptide. PTMProphet computes the Bayesian probability that the PTM occurs at that site in the PSM given the available fragment peak evidence. All probabilities are re-normalized so the sum of the probabilities for each type of PTM equals the total number of modifications of that type in the PSM. For each type of PSM evaluated, PTMProphet computes the information content of the localization statistics and estimates the number of PTMs of each type that can be localized with certainty in the PSM.

ALGORITHM:
1. Compute Probability for each potential PTM site for each PTM type
2. Apply Expectation / Maximization until probabilities remain constant
3. Normalize all probabilities by the number of modifications in the peptide
4. Record the output in the pepXML

Support & Information

Support provided by:
• NIH, NIGMS grants: R01GM087221, R24GM127667
• NHLBI grant: R21HL133335
• NHLBI grant: R01HL133135

TPP + Open Mod Supported By Limelight:
https://limelight.ms.org/
TPP Resources:
http://www.tppms.org

Magnum: Trans-Proteomic Pipeline Open Search

MassDiffs and Adducts
• Magnum encodes massdiffs and mass adducts in pepXML
• mass adducts are massdiffs assigned to an amino acid
• When a PSM has an adduct it absorbs the massdiff
• mass of adduct becomes 0
• MassProphet has been updated such that when it sees an annotated mass adduct it uses that in place of the massdiff for the PSM for the purposes of modeling the mass differences
• PTMProphet has been updated to localize mass diffs or adducts together with variable PTMs specified by the users for localization
• Massdiffs and adducts are localized to each possible amino acid on the peptide
• The "EXCLUDEMASSDIFFMIN=", "EXCLUDEMASSDIFFMAX=" options together specify a window of mass diffs to exclude from localizing
• mass diffs that are the result of measurement or isotope error can be excluded
• PTMProphet annotates the localized masses using pepXML source= attribute as one of: "param", "massdiff" or "adduct"