Fast and accurate MS² peak intensity prediction for multiple fragmentation methods, instruments and labeling techniques Ralf Gabriels

Ť \ \ / _







13/03/2019

MASSTRPLAN Final Conference, 2019

MS²PIP: MS² Peak Intensity Prediction



We are all trained to recognize people



However, recognizing penguins is not as easy



Unless you are a trained zookeeper



Specific fragmentation methods, instruments and labeling techniques are the proverbial penguins



TMT structure adapted from www.thermofisher.com/order/catalog/product/9040I

We retrained MS²PIP for these specific cases using publicly available datasets



Model performance is measured by calculating the Pearson correlations with empirical spectra

1.00 0.75 0.50 0.25 intensity 0.00 0.25 0.50 0.75 1.00 300 400 500 600 700 800 900 m/z

Prediction TMT model

Empirical TMT spectrum

 \rightarrow Pearson correlation = 0.85

1.00 0.75 0.50 0.25 0.00 0.25 0.50 0.75 1.00 300 400 500 600 700 800 900 m/z

Prediction HCD model

Empirical TMT spectrum

 \rightarrow Pearson correlation = 0.20

As expected, training data-specific models substantially improves the predictions



As expected, training data-specific models substantially improves the predictions



As expected, training data-specific models substantially improves the predictions



All models are available on iomics.ugent.be/ms2pip

MS²PIP SERVER

HOW TO RUN MS²PIP CONTACT

MS²PIP SERVER

MS² Peak Intensity Prediction

MS²PIP is a tool to predict MS² signal peak intensities from peptide sequences. It employs the XGBoost machine learning algorithm and is written in Python.

You can install MS²PIP on your machine by following our extended install instructions found on the <u>MS²PIP GitHub repository</u>. For a more user friendly experience, we created this web server. Below, you can easily upload a list of peptide sequences, after which the corresponding predicted MS² spectra can be downloaded in a CSV or MGF file format.

More advanced users can also access MS²PIP Server through our <u>RESTful API</u>. Swagger-generated documentation can be found <u>here</u> and an example Python script to access the API can be found <u>here</u>.

If you use MS²PIP for your research, please cite the following papers:

• Degroeve, S., Maddelein, D., & Martens, L. (2015). MS²PIP prediction server: compute and visualize MS2 peak intensity predictions for CID and HCD fragmentation. *Nucleic Acids Research*, 43(W1), W326-W330.

Try out MS²PIP yourself at iomics.ugent.be/ms2pip



















fwo















www.compomics.com @CompOmics