PiQuant: Software for Rapid and Accurate Quantification of Mass Spectrometry-based Proteomics Data

Pedro Moreno-Cardoso[1],David N. Perkins[1], Nazrath Nawaz[1], Weronika E. Borek[1], Josie A. Christopher[1], Amy E. Campbell[1], Andrew Williamson[1], Pedro R. Cutillas[1,2], Arran D. Dokal[1]

(1) Kinomica Ltd, Alderley Park, Macclesfield, United Kingdom; (2) Barts Cancer Institute, Queen Mary University of London, London, United Kingdom

Introduction

- Data-dependent acquisition (DDA) is widely used for the identification and label-free quantification of peptides and proteins
- We have developed the PiQuant software which has been demonstrated to be a reliable, fast and efficient tool to analyze and quantify DDA data
- Accurate and reliable quantification is extremely important for downstream phosphoproteomic analysis including biomarker identification using machine learning
- PiQuant has been extensively optimized and improved compared to previous versions

Here, we present the comparison of PiQuant results against other popular tools (Fragpipe and Maxquant) using an ideal benchmark dataset for phosphopeptide quantification

Methods

Major improvements to PiQuant software:

- New peptide grouping which more accurately represents PTM localisation
- Different methods of matching between runs (MBR) have been tested and optimised (*Fig 1-2*):
- Peak picking has been improved using isotopic contributions and optimized using observed fragmented peptides (*Fig 3*)
- Quantification techniques were improved to reliably quantify narrow peaks and allow more robust quantification at retention time (RT) extremes
- Quality control measures were added including isotopic contribution modeling, accuracy of RT alignment and XIC visualization

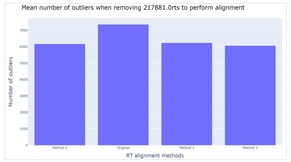


Fig 1- Number of outliers (> 1.5 min) for each RT alignment method

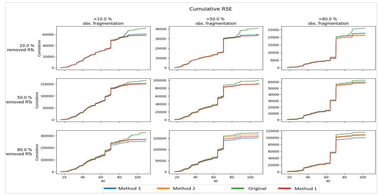


Fig 2- Residual sum of errors (RSE) for each RT alignment method with different fractions of observed RTs. Method 1 performs consistently better

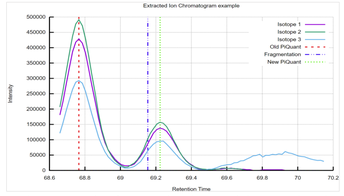
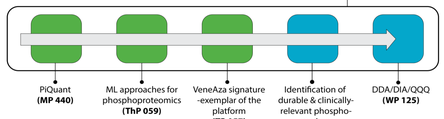


Fig 3- Example of an XIC showing much improved peak picking performance

Benchmark to evaluate PiQuant developments:

- 15 samples with different percentages of pervanadate-treated cells (*Fig 4)* underwent trypsin digestion and phosphopeptide enrichment using TIO_2
- Samples were analysed on a Thermo Obritrap Exploris 240 with Vanquish Neo using the DDA method
- Data was analysed using Fragpipe, Maxquant and PiQuant for comparison purposes
- Comparable settings were used in PiQuant, Fragpipe and Maxquant where appropriate, otherwise default settings were used

Other aspects of the full KScan® workflow are covered in **several other ASMS 2024 posters:**



Comparisons

a) Experimental Design

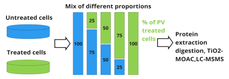


Fig 4 - Diagram of linearity experiment using different % of pervanadate .

- Pervanadate (PV) inhibits the de-phosphorylation of tyrosines
- Samples processed with 5 different concentrations of PV with 3 replicates each (15 samples total) were used in this study
- This is an ideal dataset for quantification comparisons as phosphotyrosine peptide behaviour is predictable

b) General Overview and Quantification Correlations

Pipeline	Runtime	Number of peptides	% Missing	Median CV
Fragpipe	65 mins	38161 (=27253)	61.51621	70.71445
Maxquant	19.30 hrs	33251 (=23751)	49.01979	70.56589
Mascot + Piquant	67 mins (46 + 21)	23019 (=23019)	15.64919	65.11044

Fig 5- General outputs

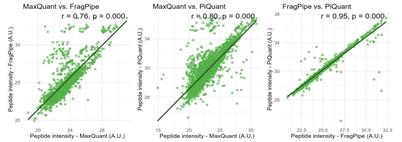


Fig 6- Correlation of shared peptide quantification between each software on all samples

- PiQuant and Fragpipe have similar run times whereas Maxquant is significantly slower (*Fig 5*)
- PiQuant MBR demonstrates much fewer missing values and lower co-variation due to its optimised RT alignment algorithms and peak picking (*Fig 5*)
- There is a high correlation between PiQuant and Fragpipe quantification whereas Maxquant shows more significant variation (*Fig 6*)

c) Sample Clustering and Differential Expression Analysis

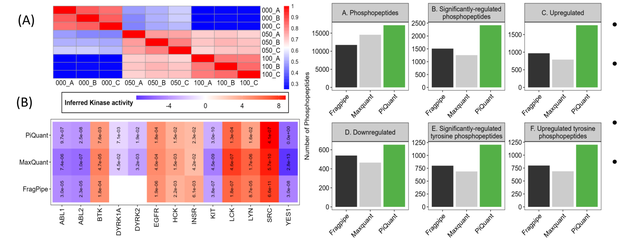


Fig 7 – (A) Clustering of samples by replicates and PV concentration (0%, 50% and 100%). (B) Tyrosine kinase-substrate enrichment activity (KSEA) analysis

Fig 8 - Total number (A) and differently regulated phosphopeptides (B-F) between 100% PV-treated and untreated samples

- Heatmaps show clear clustering of sample replicates by pervanadate concentration (*Fig 7A*)
- Tyrosine kinase-substrate enrichment activity (KSEA) is similar in both PiQuant and Fragpipe but Maxquant has a lower correlation (*Fig 7B*)
- PiQuant analysis shows a wider coverage of phosphopeptides (*Fig 8A)*
- A higher number of differentially regulated phosphotyrosines and phosphopeptides are identified by PiQuant compared to Maxquant and Fragpipe (*Fig 8B-F*)

d) Correlation of Tyrosine Sites with Pervanadate Concentration

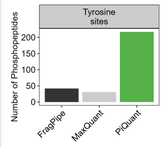


Fig 9- Correlation of phosphosite with PV concentration

- PiQuant identifies a much greater number of phosphosites that correlate with pervanadate concentration with r > 0.8 (*Fig 9*)
- PiQuant clearly identifies the expected correlation between tyrosine peptide quantification and pervanadate concentration (*Fig 10*)

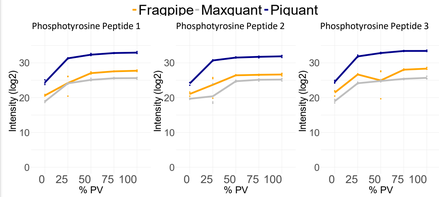


Fig 10- Examples of phosphotyrosine sites where quantification correlates with pervanadate (PV) concentration

e) PiQuant Viewer Allows the Visualisation of Quantified Peptides

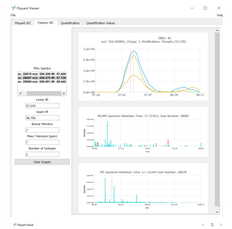


Fig 11- XIC of a PiQuant identified peptide along with the corresponding MS1 and MS2 spectra

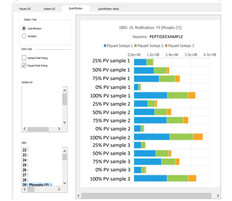


Fig 12- Representation of the contribution of each isotope to the total quantification for a given peptide in all samples

Biomarker Database

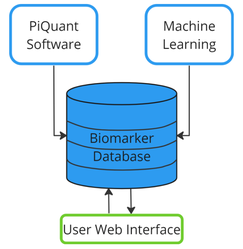


Fig 13- Diagram of biomarker database interactions

- Peptides identified by PiQuant are stored in the database along with their metadata (*Fig 13)*
- Users can search for specific biomarkers or machine learning parameters (*Fig 13*)
- The database allows the rapid identification of significant peptides (biomarkers) across different studies
- Full tracking of peptides from the initially collected data, through peptide identification to model building facilitates regulatory compliance

Conclusions

- We have demonstrated that the changes to PiQuant have made the algorithm more reliable, robust and accurate
- Comparisons with other software show that PiQuant has a superior phosphopeptide quantification performance than Fragpipe and Maxquant
- The MBR algorithms in PiQuant show much fewer missing values allowing greater reproducibility and lower co-variation
- PiQuant identifies a higher number of differentially regulated and PV correlated peptides (r > 0.8) than other packages
- The extra QC measures and visualisation tool allows for greater confidence in quantification and matching between runs
- A database permits the tracking and monitoring of identified potential biomarkers across different studies

References

- Casado, P. et al. Sci Signal. 2013 Mar 26;6(268).
- Cox, J. et al. Nat Biotechnol 26, 1367–1372 (2008)
- Kong, A. T. et al. Nature Methods, 14(5), 513-520.
- Perkins, DN. et al. Electrophoresis, 20(18) 3551-67 (1999).
- Wilkes, EH. et al. Mol Cell Proteomics. 2017 Sep;16(9):1694-1704

COI: Authors affiliated with Kinomica Ltd are employees, co-founders and/or shareholders of the company