

pyQms enables universal and accurate quantification of mass spectrometry data

Johannes Leufken^{1, 2}, Anna Niehues¹, L. Peter Sarin², Florian Wessel^{1, 3}, Michael Hippler¹, Sebastian A. Leidel^{2, 4, 5*} and Christian Fufezan^{1, 6*}

¹ Institute of Plant Biology and Biotechnology, University of Muenster, Schlossplatz 8, 48143 Muenster, Germany

² Max Planck Research Group for RNA Biology, Max Planck Institute for Molecular Biomedicine, Von-Esmarch-Strasse 54, 48149 Muenster, Germany

³ Deutsches Krebsforschungszentrum, G181 DKFZ-Bayer Joint Immunotherapy Laboratory, 69120 Heidelberg, Germany

⁴ Cells-in-Motion Cluster of Excellence, University of Muenster, 48149 Muenster, Germany

⁵ Faculty of Medicine, University of Muenster, Albert-Schweitzer-Campus 1, 48149 Muenster, Germany

⁶ Cellzome A GSK Company, Meyerhofstrasse 1, 69117 Heidelberg, Germany

Running title: Universal quantification tool for mass spectrometry

Abbreviations

DDA	data dependent acquisition
DIA	data independent acquisition
FASP	filter-aided sample preparation
FDR	false discovery rate
FN	false negative
FP	false positive
FPR	false positive rate
LC-MS/MS	liquid chromatography coupled tandem mass spectrometry
MGF	mascot generic format
MIC	matched isotope pattern chromatograms
MS	mass spectrometry
PEP	posterior error probability
PSM	peptide spectrum match
ROC	receiver-operating characteristic
RT	retention time
SILAC	stable isotope labelling with amino acids in cell culture
TAP	tris-acetate-phosphate
TN	true negative
TP	true positive
TPR	true positive rate

Summary

Quantitative mass spectrometry (MS) is a key technique in many research areas (1), including proteomics, metabolomics, glycomics, and lipidomics. Because all of the corresponding molecules can be described by chemical formulas, universal quantification tools are highly desirable. Here we present pyQms, an open-source software for accurate quantification of all types of molecules measurable by MS. pyQms uses isotope pattern matching which offers an accurate quality assessment of all quantifications and the ability to directly incorporate mass spectrometer accuracy. pyQms is, due to its universal design, applicable to every research field, labeling strategy, and acquisition technique. This opens ultimate flexibility for researchers to design experiments employing innovative and hitherto unexplored labeling strategies. Importantly, pyQms performs very well to accurately quantify partially labeled proteomes in large-scale and high-throughput, the most challenging task for a quantification algorithm.

Introduction

Current mass spectrometric workflows use a plethora of labeling strategies (Fig. 1). Established examples are label-free quantification (Fig. 1A) and metabolic labeling with partially or fully enriched isotopes (Fig. 1B and C). Furthermore, labeled molecules can be added to the culture, as in SILAC (2), or can be introduced *in vitro* by chemical tagging, e.g. TMT10 (3, 4) (Fig. 1D). For certain research areas and labeling strategies, powerful data analysis tools are well established (5–14).

However, existing software solutions are generally not universal as they have been tailored to specific research fields and are often restricted to defined experimental protocols (see supplemental Table S1 for a summary). For example, some software can quantify peptides with artificial isotope distributions, but not metabolites or can quantify molecules when they are metabolically labeled with nitrogen isotope ^{15}N but not if labeled with carbon isotope ^{13}C . Finally, some labeling strategies can currently not be combined within the same experiment (Fig. 1E and F).

These limitations, however, are artificial since all quantified entities are molecules defined by chemical formulas with known isotope distributions and masses. pyQms takes advantage of this knowledge and treats all molecules as formulas to calculate accurate isotope patterns that are based on the labeling strategy. This liberates the algorithm to perform analyses irrespective of the type of molecule (protein, metabolite, lipid, glycan, etc.), the type of label (metabolic or fixed) or MS level. Accordingly, there is no restriction in quantifying any combination of labels within the same experiment, paving the way for innovative experimental designs that would be precluded with most quantification tools. pyQms has been evaluated for accuracy and sensitivity in label-free proteomics (supplemental Fig. 1 A –C) and for pulse (chase) metabolic labeling data analysis using a novel partially labelled proteome gold standard data set (supplemental Fig. 1 D–J).

Experimental procedures

Partially labelled proteome gold standard data set sample preparation

Chlamydomonas reinhardtii (strain CW15) cells were grown in photoheterotrophic conditions in tris-acetate-phosphate (TAP) medium (15) at a light intensity of 50 μE $\text{m}^{-2} \text{s}^{-1}$ at 22 °C on a rotary shaker at 120 rpm or on TAP-agar plates containing 1.5% agar at a light intensity of 40-50 μE $\text{m}^{-2} \text{s}^{-1}$. Metabolic labeling was performed by mixing unlabeled TAP medium with fully labeled TAP medium containing 100% ^{15}N at different proportions (0, 20, 40, 60, 80 and 100% (w/w)). Fully labeled medium was created using 99.4% ^{15}N enriched $^{15}\text{NH}_4\text{Cl}$ (Cambridge Isotope Laboratories, Tewksbury, MA, USA). Cells were grown several generations on ^{15}N containing agar plates to ensure complete metabolic labeling with the defined ^{15}N proportion. Cells were then grown for three days in liquid medium, maintaining the labeling proportion and harvested at 5000 x g (Beckmann Coulter J 20 XP), suspended in H6 buffer (5 mM HEPES, pH 7.5, 10 mM EDTA) and stored at -80 °C. Protein samples were digested with trypsin using a modified filter-aided sample preparation (FASP) protocol (16) as described in Barth *et al.* (17) with the following modifications: Samples were mixed based on equal chlorophyll content (6.25 μg) and washing steps were repeated 4 times.

LC-MS/MS measurement

Liquid chromatography coupled tandem mass spectrometry (LC-MS/MS) measurements were done employing a Dionex Ultimate 3000 UPLC system (Thermo Scientific, Bremen, Germany) and Q Exactive Plus (Thermo Scientific, Bremen, Germany) instrument. Software versions, which were used are: Exactive Series (Tune) 2.3 Build 1765 and Xcalibur 3.0.63. Peptides were separated by reversed phase chromatography. Peptide samples were loaded on a trap column (Acclaim PepMap100, 300 μm x 5 mm, 5 μm particle size, 100 Å pore size; Thermo Scientific, Bremen, Germany). Samples were desalted utilizing a flow rate of 5 $\mu\text{L}/\text{min}$ for 5 min using 2% (v/v) acetonitrile/0.05% (v/v) trifluoroacetic acid in ultrapure water. Peptide separation was done using a mobile phase composed of 0.1% (v/v) formic acid in ultrapure water (A) and 80% (v/v) acetonitrile/0.1% (v/v) formic acid in ultrapure water (B). The trap column was switched for peptide elution in-line with a C18 capillary column (Acclaim PepMap 100, 75 μm x 150 mm, 2 μm particle size, 100 Å pore size, Thermo Scientific, Bremen, Germany). The gradient used was: 2.5–35% B (90 min), 35–99% B (5min), 99% B (5 min). Ions were generated by electrospray ionization (ESI). For full scans a resolution of 70,000 at m/z 200 was used (maximum injection time: 50 ms, AGC target: 1e^6 , range: 400–1600 m/z). For each full scan the 12 most abundant precursor ions (charge 2+ to 7+) were selected for fragmentation (MS^2) by higher energy c-trap dissociation (HCD). For MS^2 scans a resolution of 17,500 at m/z 200 was used (maximum injection time: 50 ms, AGC target: 5e^4 , underfill ratio: 1%). A dynamic exclusion of 60s for fragmented precursor ions was used. In total 35 LC-MS/MS runs were recorded consisting of 636,910 MS^1 and 1,258,099 MS^2 scans.

Data analysis

LC-MS/MS files in RAW format were converted to mzML (18, 19) using Proteome Discoverer (version 1.4.0.0). Subsequently, pymzML (20) was used to convert the mzML files to the mascot generic format (mgf), if required. It is noteworthy that although MS1 scans were recorded in profile mode, all spectra used for quantification in pyQms need to be transformed to a centroided format, which can be done automatically during conversion to mzML or alternatively if mzML parsing is performed using pymzML (20).

All peptide identification algorithms were executed using Ursgal (21), a Python framework for performing peptide identifications, statistical post-processing and data visualization using unified parameters. Briefly, peptides were identified using the algorithms OMSSA (version 2.19, (22)), X!Tandem (version 2013.09.01, (23)), MS-GF+ (version 9979, (24)) and MyriMatch (version 2.1.138, (25)). Default values were used for most of the search parameters. Precursor mass accuracy was set to 5 parts per million (ppm), fragmentation mass accuracy was set to 20 ppm. Trypsin was defined as protease. A shuffled-peptide based target-decoy database (conserving trypsin cleavage sites) was generated as described previously (17) using Ursgal. The *Chlamydomonas reinhardtii* database from the Joint Genome Institute (JGI) version 5.3.1/236 (26) with Augustus 11.6 (27), the chloroplastic (28) and mitochondrial proteome as well as a contaminant database (cRAP, (29)) were used for the generation of the target decoy database containing in total 19,537 (target) protein sequences (see supplemental material). Further search parameters were variable modifications: oxidation of methionine (+15.9949 Da) and acetylation of the N-terminus (+42.0106 Da). No fixed modifications were defined. Two missed

cleavage sites were permitted. Database searches were conducted for ^{14}N and ^{15}N labeling separately. The posterior error probability (PEP) was determined for each Peptide-spectrum match (PSM) using Percolator (version 2.08, (30, 31)). All PSMs with PEPs $\leq 1\%$ at the level of the database search engine were used for all subsequent analyses. A total of 19,976 unique peptides were identified, mapping to 18,285 unique chemical formulas. The mass spectrometry proteomics data have been deposited to the ProteomeXchange (32) via the PRIDE partner repository with the dataset identifier PXD003236. Venn diagrams of identified peptides and overlap of the different samples can be found in supplemental Fig. 2 and 3, respectively. Spectrum annotations of proteins with one PSM or one distinct peptide can be found in the supplemental material. All identified proteins and peptides including sequence coverage can be found in supplemental Table S2 and S3.

Peptides were quantified using pyQms (v0.5.0). Retention-time (RT) alignment and enhancement defining RT windows for all peptides (supplemental Table S4) was carried out using piqDB as described earlier (17).

pyQms scoring

The pyQms matching score (mScore) is based on the work of Gower (33). The matching and scoring is performed on the m/z values and the *intensity* values independently yielding two scores, i.e. S^{mz} and $S^{intensity}$. In both cases, each peak k is scored, comparing the measured value i with the calculated value j (equation 1), whereas a perfect match is 1. Each peak of the isotopologue that has a relative intensity (relative to the maximum intensity isotope peak) r_k above the matching

threshold (by default 1% of the maximum intensity isotope peak) is matched and scored.

$$s_{ijk} \in [0, 1] \quad (1)$$

The m/z score: S^{mz}

For each peak k , the m/z similarity between measured value i and the calculated value j is defined as:

$$s_{ijk}^{mz} = 1 - \left(\frac{\delta_{ijk}^{mz}}{\alpha} \right) \quad (2)$$

Whereas δ_{ijk}^{mz} the difference in ppm between measured m/z_{ik} and calculated m/z_{jk} and α defines the range in ppm, in which the score decreases from 1 to 0 in a linear fashion. In principle, α is equal to the precision of the measurement defined by the user (pyQms parameter “REL_MZ_RANGE”, default 5 ppm, <http://pyqms.readthedocs.io/en/latest/params.html>). For example, if the difference between measured and theoretical m/z values would be 2.5 ppm, then the s_{ijk}^{mz} score for this peak k would be 0.5.

The total m/z score for all peaks termed S^{mz} is the weighted sum of all single similarity m/z scores s_{ijk}^{mz} (equation 3). The weighting is defined by the theoretical intensity of the peak k relative to the highest peak in the theoretical isotope pattern, termed r_k .

$$S^{mz} = \frac{\sum^k s_{ijk}^{mz} r_k}{\sum^k r_k} \quad (3)$$

The intensity score: $S^{intensity}$

Prior to intensity scoring, the scaling factor σ is calculated by comparing the intensities of the measured i and calculated j intensities for all peaks k within the matching threshold (see above). This scaling factor is calculated by dividing the

weighted sum of the measured intensity by the weighted sum of the theoretical intensities (equation 4).

$$\sigma = \frac{\sum^k \text{intensity}_{ik} r_k}{\sum^k \text{intensity}_{jk} r_k} \quad (4)$$

Using this scaling factor, which is equal to the abundance of the measured molecule, one can calculate $\delta_{ijk}^{intensity}$, which is the relative intensity error between measured and theoretical intensity for each peak k (equation 5).

$$\delta_{ijk}^{intensity} = \frac{|\text{intensity}_{ik} - \sigma * \text{intensity}_{jk}|}{\sigma * \text{intensity}_{jk}} \quad (5)$$

The intensity score of peak k is then defined (equation 6).

$$s_{ijk}^{intensity} = 1 - \left(\frac{\delta_{ijk}^{intensity}}{1 - r_k + \epsilon} \right) \quad (6)$$

In analogy to the m/z score (s_{ijk}^{mz}), the denominator defines the range in which the peak based intensity score decreases from 1 to 0. However, in contrast to the m/z score, the intensity error has to be weighted by the abundance of each peak ($1 - r_k$) as more abundant peaks can be measured more accurately than smaller peaks. Additionally, we introduced ϵ (pyQms parameter “REL_I_RANGE”, default 0.2), which represents the most conservative relative error applied to the most precisely measured peak ($r_k = 1$). Thus, the overall relative error (denominator) will increase with lower peaks (see online documentation <http://pyqms.readthedocs.io/en/latest>).

The total intensity score $S^{intensity}$ is the weighted sum of all similarity scores k in analogy to the S^{mz} score:

$$S^{intensity} = \frac{\sum^k s_{ijk}^{intensity} r_k}{\sum^k r_k} \quad (7)$$

The combined final score: mScore

The final score is termed mScore and is a sum of S^{mz} and $S^{intensity}$. However, because some machines can measure m/z much more accurately than intensities, we introduced ξ to allow for flexibilities depending on the type of mass spectrometer used. ξ (the pyQms parameter “MZ_SCORE_PERCENTILE”, default 0.4) is the fraction the S^{mz} score is weighted into the sum. Thus, the final mScore is defined as:

$$mScore = \xi S^{mz} + (1 - \xi) S^{intensity} \quad (8)$$

Statistical evaluation of the gold standard data set

Statistical evaluation was performed by reducing and grouping the matched isotope pattern chromatograms (MICs). A MIC is defined as all spectra matches of a peptide with a certain charge state in one LC-MS/MS run. As the name suggests, a MIC is similar to a XIC but based on all matched isotope patterns instead of a single m/z value, thus containing an additional data dimension, that is the mScore. Grouping of the MICs was based on the chemical formula and charge of the respective peptide. Firstly, each MIC was reduced to one single match, which showed the highest mScore, yielding two values per MIC (i.e. score and intensity). Secondly, the reduced MICs belonging to one chemical formula and charge state were grouped depending on their ^{15}N labeling into one of 11 label percentile bins ranging from 0-5, 6-15, 16-25, ..., 96-100%. Thirdly, each bin was reduced to a single match based on the maximum score. Finally, for each mixed sample we considered only one ground truth

at a time. Given this setup, we reduced the evaluation of a quantified chemical formula and charge combination for each ground truth in each sample to ten bins, for which the true and false positives/negatives were counted at different score thresholds. For this, MICs (with a certain charge in one MS run) were grouped as following: for true positives (TP) the expected labeling percentile bin has matches and no other bin has matches; for false positives (FP) the expected bin has no matches, but another unexpected bin has matches, for false negative (FN) the expected bin has no match at all and for true negative (TN) the unexpected bin has no match at all. False discovery rate (FDR) was defined as $FP/(FP+TP)$, true positive rate (TPR, sensitivity) as $TP/(TP+FN)$ and false positive rate (FPR) as $FP/(FP+TN)$. In total, $13.9e^6$ MICs were grouped according to molecular formula, charge state and LC-MS/MS run.

Label-free quantification

The data set from Bruderer *et al.* (34) was used to evaluate label-free peptide quantification performance and quality of pyQms for data dependent acquisition (DDA) and data independent acquisition (DIA) mode. RAW mass spectrometry files were obtained from peptidealtdas.org and converted into mzML using msconvert, which is part of Proteowizard (version 3.0.7408, (35)). MaxQuant/Andromeda (version 1.4.1.2, (5)) peptide identification results were used from Bruderer *et al.* (34) and stored in piqDB. All peptides were subsequently quantified by pyQms using default parameters except an adjustment of the machine offset in ppm to correct for measuring error (21, 34). Carbamidomethylation of cysteine was defined as fixed modification. RT alignment (supplemental Table S5) and enhancement were carried

out using piqDB as described earlier (17, 36). Intensity alignment (supplemental Table S6) of all samples was performed as described earlier (37). The signal intensity of a given peptide (i.e. peptide charge combination) within each MS run was defined as the maximum intensity over all spectra within the predefined retention time window. RT alignment functions and raw peptide amounts for the DDA data set can be found in supplemental Table S5 and S7, respectively. The linear correlation of peptide concentrations and their matched peptide intensities over a wide concentration range (0.8 fmol/ μ L – 819.2 fmol/ μ L) can be found in supplemental Fig. 4 and 5.

Peptide ratios (\log_2) between LC-MS/MS runs were calculated using these raw quantification data (mScore ≥ 0.8). MaxQuant raw peptide quantification results were obtained from the supplement material of Bruderer *et al.* (34). Correlation plots between the MaxQuant and pyQms amounts for the spiked-in proteins and for the human background proteins can be found in supplemental Fig. 6 and 7, respectively. In order to compare pyQms with MaxQuant, the peptide (peptide charge combinations) ratios were equally calculated using the raw peptide amounts reported by Bruderer *et al.* (34).

A similar procedure was applied to the DIA measurements of the same samples. For all peptides that are proteotypic to the spiked-in proteins, all fragment ions were determined and ions suitable for quantification were determined using an algorithm that will be described elsewhere. These selected fragment ions were quantified in the cycle window corresponding to their peptide precursor m/z. A complete table of all ions used for quantification can be found in supplemental Table S8. DIA runs were intensity aligned with the same method as for the DDA runs as described earlier

(37). DIA peptide quantification was based on summing up all fragment ion intensities (maximum intensity in each MS run within the retention time window). Only ions that could be quantified in both samples were taken into account for the ratio calculation. All pyQms peptide quantification results for the DIA data set can be found in supplemental Table S9. Spectronaut DIA raw peptide amounts were taken from the supplement material of Bruderer *et al.* (34) and those were equally summed up at the peptide level. Finally, to ensure comparability, peptide ratios (log2) were calculated similar as for the pyQms results.

Peptide ratio evaluation

The data set published by Bruderer *et al.* (34) contains 12 non-human proteins spiked in a background of human HEK293 cells. The setup contained three different master mixes (MMs) spanning several orders of magnitudes of various protein concentration (MM1: 1.1 to 13.33 fold, MM2: 1 to 200 fold, MM3: 1 to 16384 fold; eight concentrations in total for each MM). Three technical replicates for each of the eight different samples containing a distinct concentration of each master mix were measured by LC-MS/MS (supplemental Fig. 1). pyQms was used to quantify the peptides belonging to the spiked-in proteins. Subsequently, the quantification was used to calculate ratios for all peptide charge combinations between all samples (see above) including all technical replicates. This resulted in 276 potential ratios for each peptide (defined by all combinations of the 24 measurements). All peptides belonging to one spiked-in protein were used to evaluate the observed and calculated log2 ratio against the known ground-truth using a two-sided t-test implemented in Scipy (www.scipy.org, (38)). At least three peptides (or peptide

charge combination ratios) were required per sample comparison. We grouped the calculated p-values into four bins that classify the quality of the quantification, i.e. the similarity to the expected ground truth: a p-value 1) above 0.05 was classified as ‘not significant different’ and colored blue, 2) 0.05 - 0.01 (*) colored green, 3) 0.01 - 0.001 (**) colored yellow and 4) smaller than 0.001 (***) colored red). This color scheme is consistent throughout this work (Fig. 2 and supplemental Fig. 8-10).

Liquid chromatography mass spectrometry of ribonucleosides

Chemically synthesized MS-grade adenosine ($C_{10}O_4N_5H_{13}$; Carbosynth Ltd, Berkshire, UK) was dissolved in 5 mM ammonium formate pH 5.3 at a concentration of 50 ng/ μ L and further diluted to yield the following concentrations; 1000 pg/ μ L, 500 pg/ μ L, 250 pg/ μ L, 125 pg/ μ L, 62.5 pg/ μ L, 31.25 pg/ μ L, 20 pg/ μ L, 10 pg/ μ L, 5 pg/ μ L, 2.5 pg/ μ L, 1 pg/ μ L, 0.5 pg/ μ L, 0.2 pg/ μ L, 0.1 pg/ μ L, 0.05 pg/ μ L, 0.02 pg/ μ L, and 0.01 pg/ μ L. 4 μ L of each dilution was subjected to reversed phase LC-MS analysis on a self-packed 75 μ m \times 500 mm porous graphitic carbon column connected to a Q Exactive mass spectrometer (Thermo Scientific, Bremen, Germany). Full MS spectra (m/z 100-700) were recorded with three technical replicates for each sample, and the resulting LC-MS runs were analyzed by pyQms. Full details of the chromatography and mass spectrometry conditions will be published elsewhere.

Experimental Design and Statistical Rationale

The pulse (-chase) gold standard data set contains in total six samples (0, 20, 40, 60, 80 and 100% ^{15}N labelling) representing six biological replicates. These samples

were mixed (0|100, 20|100, 40|100, 0|60 and 0|80), resulting in three technical replicates for the 0% sample and in two technical replicates for the 100% sample. All other samples have no technical replicate within these mixes. Furthermore each of the seven FASP fractions (16) was measured individually in order to achieve higher sensitivity. In total 35 LC-MS/MS runs including 636,910 MS¹ and 1,258,099 MS² scans of the pulse gold standard data set were evaluated. Please refer to the methods section ‘Statistical evaluation of the gold standard data set’ for all statistical tests used for the analysis of this data set.

Requirements, availability and documentation

pyQms requires Python 3.4+ and is platform independent (OS X, macOS, Linux, and Windows). The module is freely available on <https://github.com/pyQms/pyqms> or pypi, published under Massachusetts Institute of Technology (MIT) license and requires no additional modules to be installed. We recommend pymzML (20), for fast access to spectra from mzML files. To run example scripts it is necessary to install pymzML or to change the code for alternated spectra access. Some scripts also require the openpyxl or rpy2 modules.

The documentation of pyQms including parameter description (<http://pyqms.readthedocs.io/en/latest/params.html>), a quick-start tutorial (http://pyqms.readthedocs.io/en/latest/quick_start.html) and example scripts (http://pyqms.readthedocs.io/en/latest/example_scripts.html) can be found online. pyQms can be run on standard Desktop computers. For further hardware requirements please refer to the online documentation.

Results

Algorithm design

The core of pyQms is a standalone open source Python module that can be incorporated easily into any analytical workflow, can be run on standard desktop computers and that benefits from the resources of a rich scientific computing community (20, 21, 39–42). The workflow of pyQms can be divided into two steps. First, an isotope pattern library is built, based on user-defined molecules (Fig. 1G). These molecules are specified as chemical formulas or peptide sequences, optionally with modifications following the unimod standard. The metabolic isotopic distribution and the artificial isotopic distribution of the fixed label are taken into account. pyQms uses this information to calculate an accurate isotope pattern for each molecule based on its elemental composition, the isotope distributions and the respective masses. Second, the predicted isotope patterns are compared to the MS measurements (Fig. 1H) and scored by calculating a similarity coefficient (33) optimized for mass spectrometry data yielding a similarity match score, termed mScore. The matching and scoring algorithm initially uses the m/z values and the intensity values independently and combines both later, offering the possibility to adjust the scoring algorithm to the machine type, as some instruments can measure m/z values much more accurately than intensities. Furthermore, the similarity between each measured and calculated isotope pattern peak is weighted by its relative abundance. Thus, the more abundant peaks contribute more to the mScore (see M&M for details). We developed this similarity matching approach as the basis for pyQms since it provides several advantages: i) The isotope pattern contains

multiple peaks that are all used for quantification, thus increasing the robustness over using only a single value. ii) The matching approach tags each quantification with a quality assessment reflected by its mScore. iii) Co-eluting molecules can either be distinguished reliably since the isotope patterns differ or they receive a low mScore. iv) The matching algorithm can adapt to any mass spectrometer accuracy. Thus, future technological advancement in machine accuracy will automatically translate into higher matching sensitivity without the need to adjust the algorithm. Overall, any MS data analysis workflow that relies on single *m/z* observations will benefit from the incorporation of isotope pattern matching and our scoring methodology.

Furthermore, pyQms accepts so called evidence files as input (e.g. peptide identifications from Ursgal (21) or manually curated data). These files allow molecules with the identical chemical formula but different identities to be distinguished by associating their identity with a given retention-time window. pyQms offers a modular system to use custom functions that can be used to determine the abundance of a given molecule. The basic built-in function determines abundance by the maximum intensity within a retention time window that is defined in the evidence files. However, pyQms was primarily build for bioinformaticians, thus functions to define retention windows or to determine abundances can easily be incorporated if required.

pyQms provides quantification with high accuracy and sensitivity

To assess the accuracy and sensitivity of pyQms for metabolomics, we analyzed chemically synthesized MS-grade adenosine in a dilution series ranging from 0.04-4000 pg/ μ L (Fig. 2A). Importantly, the quantification remained linear over a

broad range of concentrations (4-5 orders of magnitude) with an R^2 of 0.998, showing that pyQms can be used for the sensitive quantification of metabolites.

Next, to compare pyQms against established quantification tools in proteomics, we used the data sets of Bruderer *et al.* (34) that contain eight biological replicates of human HEK293 cells, each measured three times in data dependent (DDA) and data independent acquisition (DIA) mode. We compared the published quantification results of the DDA data set analyzed with MaxQuant (5) and of the DIA data set analyzed with Spectronaut (43) to our results obtained with pyQms (Fig. 2B).

We found that 17.4% of the comparisons reflect the ground truth in the DDA samples. However, concentrations below 4 fmol/ μ L could not be quantified confidently, which is similar to the results reported for MaxQuant (5) (21.3%, supplemental Fig. 8, 9I2). Furthermore, we found that 82.2% of the ratios obtained with pyQms reflect the ground truth for the DIA data set (Fig. 2B, bottom right half), similar to what was reported for Spectronaut (43) (76.8%, supplemental Fig. 8, 9I3). These results are comparable to the remaining spiked-in proteins (Fig. 2C). In summary, pyQms quantifications reflected 71.2%/85.2% ($n=1380$ ratios) of the ground truth in the DDA/DIA runs for master mix 1 (MM1), 34.9%/66.3% ($n=1380$ ratios) for MM2 and 20.8%/52.5% ($n=552$ ratios) for MM3 (Fig. 2C), which is similar to the results reported for MaxQuant and Spectronaut (34) (MM1: 62.8%/78.8%, MM2: 23.1%/68.5%, MM3: 15%/72.1%) (supplemental Fig. 8-11, supplemental Table S10 and S11). This shows that pyQms covers the detection limits of data acquired in DDA mode or DIA modes.

Quantification of pulse (-chase) samples

In order to evaluate the quantification performance of pyQms for pulse (-chase) samples, we create a gold-standard data set of well-defined partially labeled proteomes. This data set can be used to benchmark the accuracy of any quantitative software aiming at analyzing partially labeled proteomes. We cultivated the green algae *Chlamydomonas reinhardtii* in media containing 0 to 100% ^{15}N metabolic label in 20% increments. Extracted proteins were combined to create five mixes: 0I100, 20I100, 40I100, 0I60 and 0I80 (Fig. 3A), which were analyzed by LC-MS/MS. Subsequently, we quantified all identified peptides, allowing all possible combinations of charge states and all ^{15}N enrichment percentiles to be matched in all MS¹ scans. This summed up to a total of 5.8e^{12} performed matches, which resulted in 1.6e^8 positive matches with mScores ≥ 0.5 . We grouped matches with an identical chemical formula, charge state and labeling percentile from one LC-MS/MS run into matched isotope pattern chromatograms (MICs). Altogether, we found 1.4e^7 MICs. For each sample, we assessed the matching quality by aggregating the single matches of all MICs into one heat map (Fig. 3D, supplemental Fig. 12). We used the ^{15}N -labeling percentile and the mScore of each match as coordinates for the heat map bin and added the abundance of that match to this bin (including peptide abundances of all spectra assembled in a MIC). While the 0I100 mix can be separated reliably (supplemental Fig. 12A), this was not the case for the other mixes (supplemental Fig. 12C,E and G). Even though it was possible to identify the ground truth, the correct result was obscured by false positives leading to a high background signal that was distributed over all labeling percentiles.

We statistically evaluated our results by creating receiver-operating characteristics (ROC) of the MICs against the ground truth for each mixed sample (Fig. 3E). The ROC curves show that pyQms reliably matches the 0I100 mix (71.1% and 76.9% true positive rate (TPR) at 99% specificity, blue and red circles). In contrast, the partially labeled samples result in TPRs between 12.7% and 21.5% at 99% specificity. The fully labeled (100%) or unlabeled samples (0%), if mixed with a partially labeled sample, have lower ROC curves compared to their counterpart in the 0I100 mix (30.9% - 35.4% TPR at 99% specificity). The mixes containing partially labeled samples show a high FDR even at very high mScores (e.g. the 0% and 60% samples show FDRs of 4.5% and 29.3%, respectively at mScores ≥ 0.9 in the 0I60 mix; Fig. 3F). Only in the 0I100 mix we detected FDRs of 2.7% and 2.1% for the 0% and 100% samples, respectively, at mScores ≥ 0.7 .

Decreasing the FDR in pulse (-chase) experiments

To obtain a lower FDR, we limited quantification to stringent (2 - 2.5 min) retention time (RT) windows (supplemental Fig. 13, supplemental Table S4) based on the RT alignment strategy described earlier (17). This strategy significantly reduced false positive matches, as shown by reduced noise in the heat map (Fig. 3G). Furthermore, the ROC curves improve significantly for all five mixes (Fig. 3H). Finally, the mScore-dependent FDR (Fig. 3I) shows rates of 3.2% and 3.8% for the 40% (40I100 mix) and 60% (0I60 mix) samples, respectively, at very stringent mScores of ≥ 0.9 (Fig. 3I). This shows, that applying RT windows can significantly reduce the FDR in pulse (-chase) experiments.

Discussion

pyQms performs state-of-the-art DDA and DIA quantification

The presented results emphasize that pyQms performs optimally within the detection limits of data acquired in DDA mode or DIA modes similar to what is regarded as state-of-the art in proteomics. However, pyQms combines the functionality to quantify both DDA and DIA data in a single software. Furthermore, pyQms uniquely offers a highly accessible bioinformatics library so other packages can incorporate its isotope matching procedure to increase their matching quality with ease. Ultimately, isotope pattern matching will replace approaches relying on single *m/z* values, especially in the advent of the broad availability of high-resolution mass spectrometers.

Benchmarked partial label quantification

To demonstrate that pyQms goes beyond the current state-of-the-art, we analyzed a data set of differentially labeled proteomes, like observed during pulse- or pulse-chase experiments. This reflects the ultimate challenge for any quantification algorithm, since each molecule has different enriched isotope incorporation levels depending on the time of synthesis. Thus, one single labeling state cannot be observed, complicating analyses for several reasons: First, the isotope distribution is different for each labeling state; yet, it is difficult to distinguish closely related labeling states. Each isotope pattern must therefore be evaluated independently or in a mixed model. Second, partially labeled molecules broaden the isotope pattern and lead to more peaks, which are consequently less intense (supplemental Fig. 14). This effect

leads to a loss of signal and quantification accuracy. Finally, the number of isotope patterns that need to be matched increases by two orders of magnitude when compared to label-free quantification, thus requiring significant computational resources.

Our pulse or partially labeled gold standard data set revealed a low TPR for samples containing partially labeled peptides at high specificities in contrast to a sample containing only unlabeled or fully labeled proteins (Fig. 3E). This confirms that the complexity of the sample and the number of detected peaks strongly influences the classification. Similarly, this is reflected in the mScore-dependent FDR (Fig. 3F). These results illustrate the challenge of assessing partially labeled molecules in general. These difficulties are rarely addressed as they only become apparent by benchmarking against a defined partially labeled sample. Our gold standard data set can therefore be used to benchmark new software tools for the quantification of pulse- and pulse-chase data.

The benchmarking results of pyQms against the gold standard data set underlines the universal application and high quality of pyQms which goes beyond what is currently available in single algorithms.

pyQms is suitable for large scale pulse (-chase) studies

The approach to lower the FDR and increase the TPR by applying stringent RT windows demonstrates that pyQms allows accurate large-scale high-throughput quantification of pulse (-chase) MS experiments to be performed. Interestingly, the sample resembling the most common proteomic samples (O1100 mix) shows a FDR of only 1.2% at mScores ≥ 0.7 using accurate isotope pattern matching in

combination with RT windows (Fig. 3I, blue and red circles). This indicates that pyQms could be sufficient for MS¹-based identifications and quantification in proteomics. Importantly, this should also pave the way to reliably overcome the under-sampling issue in proteomics (44), especially if one takes the mScore as a quality criteria into account. In fact, if the Bruderer *et al.* (34) data set is evaluated with respect to mScore, this effect becomes very obvious (supplemental Fig. 15).

Conclusions

In conclusion, we have demonstrated that pyQms is a powerful software to quantify any molecule measured by mass spectrometry, independent of the type of molecule, label, research field or acquisition mode. These unique features enable novel experimental designs and multiplexing strategies, which are urgently required in the era of personalized medicine in order to simultaneously compare hundreds of clinical samples. Using pyQms, researchers will be able to quantify mass spectrometry data from proteomics, metabolomics, lipidomics, glycomics or other research fields with a single, universal software solution. pyQms is open source and freely available via <https://github.com/pyQms/pyqms> under an MIT license.

Authors contributions

C.F. designed the algorithm. J.L., A.N. and C.F. programmed and optimized the algorithm. J.L and L.P.S. performed experiments and sample preparation. J.L., L.P.S and C.F. analyzed samples and data. S.A.L., M.H. and C.F. supervised the project. S.A.L., J.L. and C.F. wrote the paper with input from all authors.

Acknowledgements

We thank Dr. R.P. Zahedi, Dr. M. Stitt, Dr. K. Bartscherer, Dr. G. Bentley and the members of the Leidel, Hippler and Fufezan groups and the eBio group of Cellzome for their valuable feedback and suggestions. Dr. H. Drexler and Dr. M. Scholz are acknowledged for the measurement of the LC-MS samples.

Data availability

The mass spectrometry proteomics data have been deposited to the ProteomeXchange (32) via the PRIDE partner repository (<http://www.ebi.ac.uk/pride/archive/>) with the data set identifier PXD003236

References

1. Yates, J. R., Ruse, C. I., and Nakorchevsky, A. (2009) Proteomics by Mass Spectrometry: Approaches, Advances, and Applications. *Annu. Rev. Biomed. Eng.* 11, 49–79
2. Ong, S.-E. (2002) Stable Isotope Labeling by Amino Acids in Cell Culture, SILAC, as a Simple and Accurate Approach to Expression Proteomics. *Mol. Cell. Proteomics* 1, 376–386
3. Filla, L. A., Yuan, W., Feldman, E. L., Li, S., and Edwards, J. L. (2014) Global metabolomic and isobaric tagging capillary liquid chromatography-tandem mass spectrometry approaches for uncovering pathway dysfunction in diabetic mouse aorta. *J. Proteome Res.* 13, 6121–34
4. Werner, T., Sweetman, G., Savitski, M. F., Mathieson, T., Bantscheff, M., and Savitski, M. M. (2014) Ion Coalescence of Neutron Encoded TMT 10-Plex Reporter Ions. *Anal. Chem.* 86, 3594–3601
5. Cox, J., and Mann, M. (2008) MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. *Nat. Biotechnol.* 26, 1367–1372
6. Park, S. K., Venable, J. D., Xu, T., and Yates, J. R. (2008) A quantitative analysis software tool for mass spectrometry-based proteomics. *Nat. Methods* 5, 319–322
7. Sturm, M., Bertsch, A., Gröpl, C., Hildebrandt, A., Hussong, R., Lange, E., Pfeifer, N., Schulz-Trieglaff, O., Zerck, A., Reinert, K., and Kohlbacher, O. (2008) OpenMS – An open-source software framework for mass spectrometry.

8. Sachsenberg, T., Herbst, F.-A., Taubert, M., Kermer, R., Jehmlich, N., Bergen, M. Von, Seifert, J., and Kohlbacher, O. (2015) MetaProSIP: Automated Inference of Stable Isotope Incorporation Rates in Proteins for Functional Metaproteomics. *J. Proteome Res.* 14, 619–627
9. Tsou, C.-C., Avtonomov, D., Larsen, B., Tucholska, M., Choi, H., Gingras, A.-C., and Nesvizhskii, A. I. (2015) DIA-Umpire: comprehensive computational framework for data-independent acquisition proteomics. *Nat Methods* 12, 258–264
10. Röst, H. L., Rosenberger, G., Navarro, P., Gillet, L., Miladinović, S. M., Schubert, O. T., Wolski, W., Collins, B. C., Malmström, J., Malmström, L., and Aebersold, R. (2014) OpenSWATH enables automated, targeted analysis of data-independent acquisition MS data. *Nat. Biotechnol.* 32, 219–223
11. MacLean, B., Tomazela, D. M., Shulman, N., Chambers, M., Finney, G. L., Frewen, B., Kern, R., Tabb, D. L., Liebler, D. C., and MacCoss, M. J. (2010) Skyline: An open source document editor for creating and analyzing targeted proteomics experiments. *Bioinformatics* 26, 966–968
12. Reiter, L., Rinner, O., Picotti, P., Hüttenhain, R., Beck, M., Brusniak, M.-Y., Hengartner, M. O., and Aebersold, R. (2011) mProphet: automated data processing and statistical validation for large-scale SRM experiments. *Nat. Methods* 8, 430–435
13. Mitchell, C. J., Kim, M.-S., Na, C. H., and Pandey, A. (2016) PyQuant: A Versatile Framework for Analysis of Quantitative Mass Spectrometry Data.

14. Palmer, A., Phapale, P., Chernyavsky, I., Lavigne, R., Fay, D., Tarasov, A., Kovalev, V., Fuchser, J., Nikolenko, S., Pineau, C., Becker, M., and Alexandrov, T. (2016) FDR-controlled metabolite annotation for high-resolution imaging mass spectrometry. *Nat. Methods* 14, 1–8
15. Harris, E. H. (1989) *The Chlamydomonas Sourcebook*
16. Wiśniewski, J. R., Zougman, A., Nagaraj, N., and Mann, M. (2009) Universal sample preparation method for proteome analysis. *Nat. Methods* 6, 359–62
17. Barth, J., Bergner, S. V., Jaeger, D., Niehues, A., Schulze, S., Scholz, M., and Fufezan, C. (2014) The Interplay of Light and Oxygen in the Reactive Oxygen Stress Response of *Chlamydomonas reinhardtii* Dissected by Quantitative Mass Spectrometry. *Mol. Cell. Proteomics* 13, 969–989
18. Martens, L., Chambers, M., Sturm, M., Kessner, D., Levander, F., Shofstahl, J., Tang, W. H., Römpf, A., Neumann, S., Pizarro, A. D., Montecchi-Palazzi, L., Tasman, N., Coleman, M., Reisinger, F., Souda, P., Hermjakob, H., Binz, P.-A., and Deutsch, E. W. (2011) mzML--a community standard for mass spectrometry data. *Mol. Cell. Proteomics* 10, R110.000133
19. Deutsch, E. (2008) mzML: A single, unifying data format for mass spectrometer output. *Proteomics* 8, 2776–2777
20. Bald, T., Barth, J., Niehues, A., Specht, M., Hippler, M., and Fufezan, C. (2012) pymzML--Python module for high-throughput bioinformatics on mass spectrometry data. *Bioinformatics* 28, 1052–1053

21. Kremer, L. P. M., Leufken, J., Oyunchimeg, P., Schulze, S., and Fufezan, C. (2016) Ursgal, Universal Python Module Combining Common Bottom-Up Proteomics Tools for Large-Scale Analysis. *J. Proteome Res.* 15, 788–794
22. Geer, L. Y., Markey, S. P., Kowalak, J. A., Wagner, L., Xu, M., Maynard, D. M., Yang, X., Shi, W., and Bryant, S. H. (2004) Open Mass Spectrometry Search Algorithm. *J. Proteome Res.* 3, 958–964
23. Craig, R., and Beavis, R. C. (2003) A method for reducing the time required to match protein sequences with tandem mass spectra. *Rapid Commun. Mass Spectrom.* 17, 2310–6
24. Kim, S., Mischerikow, N., Bandeira, N., Navarro, J. D., Wich, L., Mohammed, S., Heck, A. J. R., and Pevzner, P. a (2010) The Generating Function of CID, ETD, and CID/ETD Pairs of Tandem Mass Spectra: Applications to Database Search. *Mol. Cell. Proteomics* 9, 2840–2852
25. Tabb, D. L., Fernando, C. G., and Chambers, M. C. (2008) MyriMatch:highly accurate tandem mass spectral peptide identificaiton by multivariate hypergeometric analysis. *J Proteome Res.* 6, 654–661
26. Merchant, S. S., Prochnik, S. E., Vallon, O., Harris, E. H., Karpowicz, S. J., Witman, G. B., Terry, A., Salamov, A., Fritz-Laylin, L. K., Marechal-Drouard, L., Marshall, W. F., Qu, L.-H., Nelson, D. R., Sanderfoot, A. a, Spalding, M. H., Kapitonov, V. V., Ren, Q., Ferris, P., Lindquist, E., Shapiro, H., Lucas, S. M., Grimwood, J., Schmutz, J., Cardol, P., Cerutti, H., Chanfreau, G., Chen, C.-L., Cognat, V., Croft, M. T., Dent, R., Dutcher, S., Fernandez, E., Fukuzawa, H., Gonzalez-Ballester, D., Gonzalez-Halphen, D., Hallmann, A., Hanikenne, M.,

- Hippler, M., Inwood, W., Jabbari, K., Kalanon, M., Kuras, R., Lefebvre, P. a,
Lemaire, S. D., Lobanov, A. V., Lohr, M., Manuell, A., Meier, I., Mets, L., Mittag,
M., Mittelmeier, T., Moroney, J. V., Moseley, J., Napoli, C., Nedelcu, A. M.,
Niyogi, K., Novoselov, S. V., Paulsen, I. T., Pazour, G., Purton, S., Ral, J.-P.,
Riano-Pachon, D. M., Riekhof, W., Rymarquis, L., Schroda, M., Stern, D.,
Umen, J., Willows, R., Wilson, N., Zimmer, S. L., Allmer, J., Balk, J., Bisova,
K., Chen, C.-J., Elias, M., Gendler, K., Hauser, C., Lamb, M. R., Ledford, H.,
Long, J. C., Minagawa, J., Page, M. D., Pan, J., Pootakham, W., Roje, S.,
Rose, A., Stahlberg, E., Terauchi, A. M., Yang, P., Ball, S., Bowler, C.,
Dieckmann, C. L., Gladyshev, V. N., Green, P., Jorgensen, R., Mayfield, S.,
Mueller-Roeber, B., Rajamani, S., Sayre, R. T., Brokstein, P., Dubchak, I.,
Goodstein, D., Hornick, L., Huang, Y. W., Jhaveri, J., Luo, Y., Martinez, D.,
Ngau, W. C. A., Otiliar, B., Poliakov, A., Porter, A., Szajkowski, L., Werner, G.,
Zhou, K., Grigoriev, I. V., Rokhsar, D. S., and Grossman, A. R. (2007) The
Chlamydomonas Genome Reveals the Evolution of Key Animal and Plant
Functions. *Science* (80-). 318, 245–250
27. Stanke, M., and Waack, S. (2003) Gene prediction with a hidden Markov
model and a new intron submodel. *Bioinformatics* 19, ii215-ii225
28. Maul, J. E., Lilly, J. W., Cui, L., Claude, W., Miller, W., Harris, E. H., and Stern,
D. B. (2002) The Chlamydomonas reinhardtii Plastid Chromosome : Islands of
Genes in a Sea of Repeats. 14, 2659–2679
29. Mellacheruvu, D., Wright, Z., Couzens, A. L., Lambert, J., St-Denis, N. A., Li,
T., Miteva, Y. V., Hauri, S., Sardiu, M. E., Low, T. Y., Halim, V. A., Bagshaw, R.

- D., Hubner, N. C., Al-Hakim, A., Bouchard, A., Faubert, D., Fermin, D.,
Dunham, W. H., Goudreault, M., Lin, Z.-Y., Badillo, B. G., Pawson, T.,
Durocher, D., Coulombe, B., Aebersold, R., Superti-Furga, G., Colinge, J.,
Heck, A. J. R., Choi, H., Gstaiger, M., Mohammed, S., Cristea, I. M., Bennett,
K. L., Washburn, M. P., Raught, B., Ewing, R. M., Gingras, A.-C., and
Nesvizhskii, A. I. (2013) The CRAPome: a contaminant repository for affinity
purification–mass spectrometry data. *Nat. Methods* 10, 730–736
30. Käll, L., Canterbury, J. D., Weston, J., Noble, W. S., and MacCoss, M. J.
(2007) Semi-supervised learning for peptide identification from shotgun
proteomics datasets. *Nat. Methods* 4, 923–925
31. Kall, L., Storey, J. D., and Noble, W. S. (2008) Non-parametric estimation of
posterior error probabilities associated with peptides identified by tandem
mass spectrometry. *Bioinformatics* 24, i42–i48
32. Vizcaíno, J. A., Deutsch, E. W., Wang, R., Csordas, A., Reisinger, F., Ríos,
D., Dianes, J. A., Sun, Z., Farrah, T., Bandeira, N., Binz, P.-A., Xenarios, I.,
Eisenacher, M., Mayer, G., Gatto, L., Campos, A., Chalkley, R. J., Kraus, H.-J.,
Albar, J. P., Martínez-Bartolomé, S., Apweiler, R., Omenn, G. S., Martens, L.,
Jones, A. R., and Hermjakob, H. (2014) ProteomeXchange provides globally
coordinated proteomics data submission and dissemination. *Nat. Biotechnol.*
32, 223–226
33. Gower, J. C. (1971) A General Coefficient of Similarity and Some of Its
Properties. *Biometrics* 27, 857
34. Bruderer, R., Bernhardt, O. M., Gandhi, T., Miladinović, S. M., Cheng, L.-Y.,

- Messner, S., Ehrenberger, T., Zanotelli, V., Butscheid, Y., Escher, C., Vitek, O., Rinner, O., and Reiter, L. (2015) Extending the Limits of Quantitative Proteome Profiling with Data-Independent Acquisition and Application to Acetaminophen-Treated Three-Dimensional Liver Microtissues. *Mol. Cell. Proteomics* 14, 1400–1410
35. Kessner, D., Chambers, M., Burke, R., Agus, D., and Mallick, P. (2008) ProteoWizard: open source software for rapid proteomics tools development. *Bioinformatics* 24, 2534–6
36. Hoehner, R., Barth, J., Magneschi, L., Jaeger, D., Niehues, A., Bald, T., Grossman, A., Fufezan, C., and Hippler, M. (2013) The metabolic status drives acclimation of iron deficiency responses in *Chlamydomonas reinhardtii* as revealed by proteomics based hierarchical clustering and reverse genetics. *Mol. Cell. Proteomics*, 1–59
37. Bergner, S. V., Scholz, M., Trompelt, K., Barth, J., Gäbelein, P., Steinbeck, J., Xue, H., Clowez, S., Fucile, G., Goldschmidt-Clermont, M., Fufezan, C., and Hippler, M. (2015) STATE TRANSITION7-Dependent Phosphorylation Is Modulated by Changing Environmental Conditions, and Its Absence Triggers Remodeling of Photosynthetic Protein Complexes. *Plant Physiol.* 168, 615–634
38. Oliphant, T. E. (2007) Python for Scientific Computing. *Comp Sci Eng* 9, 10–20
39. Röst, H. L., Schmitt, U., Aebersold, R., and Malmström, L. (2014) pyOpenMS: A Python-based interface to the OpenMS mass-spectrometry algorithm library.

40. Kiefer, P., Schmitt, U., and Vorholt, J. A. (2013) EMZed: An open source framework in Python for rapid and interactive development of LC/MS data analysis workflows. *Bioinformatics* 29, 963–964
41. Perez-Riverol, Y., Wang, R., Hermjakob, H., Müller, M., Vesada, V., and Vizcaíno, J. A. (2014) Open source libraries and frameworks for mass spectrometry based proteomics: A developer's perspective. *Biochim. Biophys. Acta - Proteins Proteomics* 1844, 63–76
42. Goloborodko, A. a, Levitsky, L. I., Ivanov, M. V, and Gorshkov, M. V (2013) Pyteomics--a Python framework for exploratory data analysis and rapid software prototyping in proteomics. *J. Am. Soc. Mass Spectrom.* 24, 301–4
43. Reiter, L., Rinner, O., Picotti, P., Hüttenhain, R., Beck, M., Brusniak, M.-Y., Hengartner, M. O., and Aebersold, R. (2011) mProphet: automated data processing and statistical validation for large-scale SRM experiments. *Nat. Methods* 8, 430–5
44. Nilsson, T., Mann, M., Aebersold, R., Yates, J. R., Bairoch, A., and Bergeron, J. J. M. (2010) Mass spectrometry in high-throughput proteomics: ready for the big time. *Nat. Methods* 7, 681–685
45. Yao, X., Freas, a., Ramirez, J., Demirev, P. a., and Fenselau, C. (2001) Proteolytic ¹⁸O labeling for comparative proteomics: Model studies with two serotypes of adenovirus. *Anal. Chem.* 73, 2836–2842

Figure descriptions

Figure 1

Labeling strategies employed in mass spectrometry separated into metabolic (left) and fixed labels (right). Metabolic labeling (e.g. ^{15}N salt or ^{13}C sugar) is metabolized in the cell and incorporated into newly synthesized molecules. The isotope distribution of the labeled element can be natural (A, white circles), partially enriched by an isotope (e.g. ^{15}N , light blue color represents e.g. an average labeling of 60 %, i.e. 3 of 5 nitrogens are ^{15}N) as observed during pulse- or pulse-chase experiments (B, light blue) or fully enriched (C, dark blue). Fixed labels are incorporated into or attached to the molecule during or after the synthesis steps. Fixed labeling can be performed *in vivo* (e.g. SILAC incorporation (2)) or *in vitro* (e.g. digestion in ^{18}O labeled water (45)). In both cases the element isotope distributions of the label are independent of the cellular distributions and are thus treated as different element pools (D-F). Combining different labeling strategies permits novel multiplexing strategies. Only pyQms can be used to quantify all six cases (A-F) in all variations and combinations irrespective of the label or the molecule type and to, most importantly, score the quantifications. The metabolic isotopic distribution (left isotopologue) and the artificial isotopic distribution (right isotopologue of a potential fixed label is used to calculate an accurate isotope pattern for each molecule (G). These patterns are compared to the MS measurements (H). Matches are evaluated providing the similarity match score (mScore). Black bars, measured peaks; green triangles, matched peaks; x-axes, m/z value; y-axes, intensity. For considerations on the terms fixed and metabolic labeling, please refer to the online methods.

Figure 2

Label-free pyQms quantification of metabolites and peptides. A) Dilution series of chemically synthesized adenosine (axes plotted in log scale), each dilution comprising three technical replicates. Line shows the linear regression function ($R^2 = 0.998$). Inset shows axes plotted in linear scale; x-axes, dilution of nucleoside; y-axes, maximum intensity. B) Example of the statistical evaluation of pyQms peptide quantifications of the Bruderer *et al.* data sets (data dependent acquisition (DDA), top left; data independent acquisition (DIA), bottom right). Shown is the heat map for spiked-in protein P61823 (ribonuclease pancreatic, *Bos Taurus*, master mix 2). The colors reflect the p-value for the two-sided t-test obtained after testing whether the calculated peptide ratios between two samples (x & y-axis, eight samples, ticks represent each three technical replicates) differ significantly from the ground truth ratios. Histograms at the axes show the spiked-in protein concentrations for the eight samples (master mix 2 , for sample setup and other master mix composition refer to supplemental Fig. 1 and Bruderer *et al.* (34), x-axes, sample; y-axes, concentration of spiked-in proteins. The number of peptide charge combination ratios (for DDA samples) or peptide ratios (DIA) used for the t-test are shown as numbers in the bins. C) Stacked bar plot for relative occurrences of the p-values over all 12 spiked-in proteins across all three master mixes (MM1, MM2 & MM3). Legend: x-axis, relative occurrence of p-value; y-axis, master mix and data acquisition method combination (DDA or DIA). P-value legend for B) and C), blue: reflecting the ground truth; green (p-value ≤ 0.05), yellow (p-value ≤ 0.01) and red (p-value ≤ 0.001): not reflecting the ground truth.

Figure 3

The partially labeled proteome gold standard. A) Cultures of *Chlamydomonas reinhardtii* were grown on medium containing 20, 40, 60, 80 and 100% ^{15}N . Each partially labeled proteome was mixed with an unlabeled (^{14}N) or fully labeled (^{15}N) proteome, yielding five mixed samples (0l100, 0l60, 0l80, 20l100 and 40l100). B) Unbiased quantification workflow. All identified peptides (19976 peptides, 18285 distinct chemical formulas) were quantified in all MS1 spectra (6.3e^5) of all LC-MS/MS runs in 5 charge states performing in total 5.8e^{12} matches. Matches were filtered using an mScore threshold of 0.5 and assembled into 1.4e^7 matched isotope pattern chromatograms (MICs). Example isotopologues for 0% and 60% ^{15}N incorporation are shown on the left. C) 3D visualization of a typical MIC, colors indicate mScores for each match within the MIC; x-axis, m/z; y-axis, spectrum id; z-axis, intensity. D) Visual evaluation of all matches in the 0l60 sample, x-axis labeling percentile, y-axis mScore, heat equals summed up intensities of matches in all MS¹ spectra in a given bin (representing all identified and quantified peptides in all matched charge states). E) Receiver operating characteristic (ROC) curves of pyQms performance in all samples shown as specificity (x-axis, log scale) vs. true positive rate (y-axis). F) mScore-dependent false discovery rate; x-axis, mScore; y-axis, false discovery rate (log scale). G-I) as D-F but quantifications are limited to retention time windows. Legends (D, E, G, H): 0% (blue), 100% (red), 20% (green), 40% (purple), 60% (orange), 80% (brown); mixtures: 0l100 (circles), 0l60 (triangles), 20l100 (diamonds), 0l80 (squares), 40l100 (reverse triangles).

Figure 1

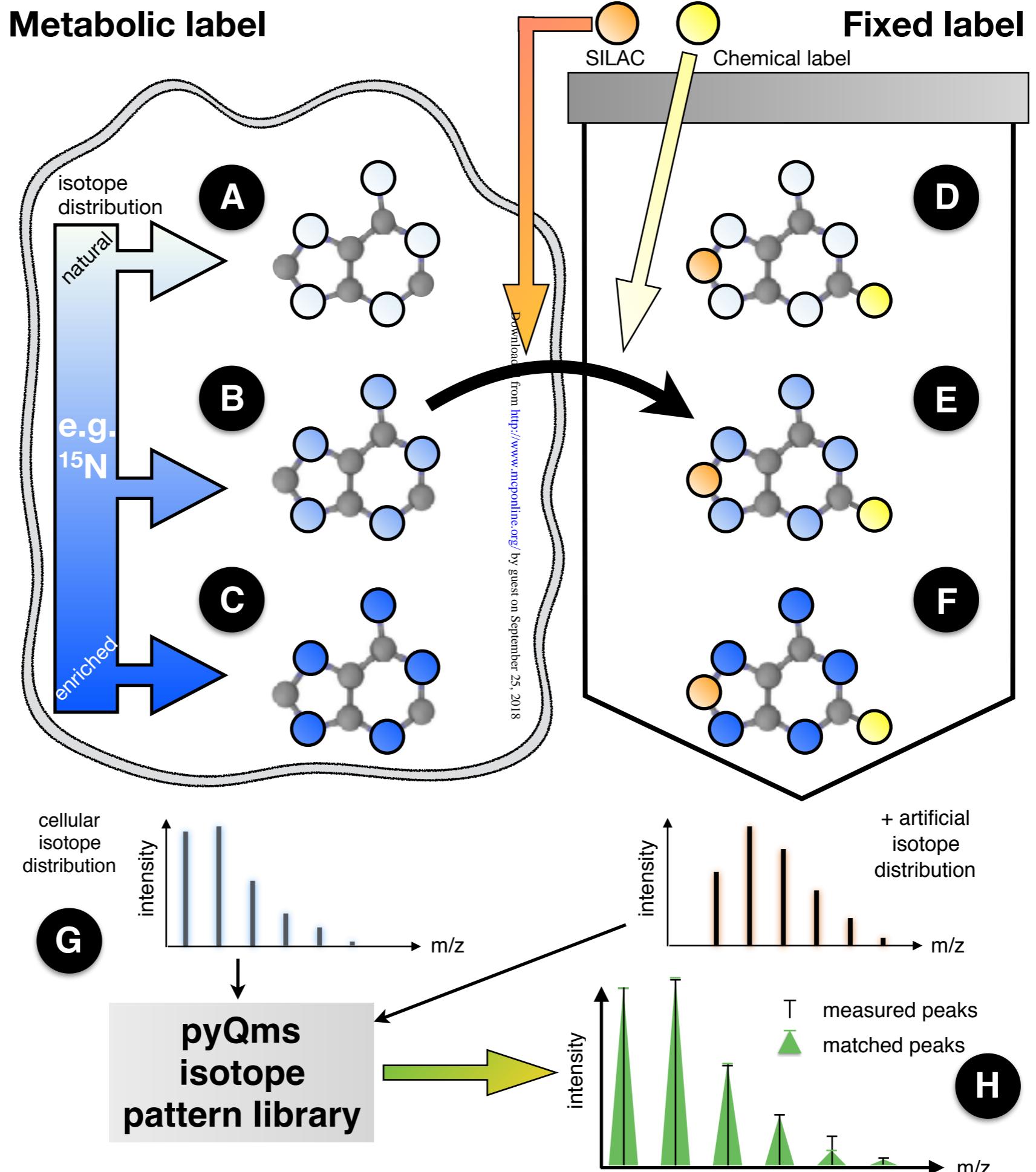


Figure 2