Published in final edited form as:
Anal Chem. 2012 July 17; 84(14): 5845–5850. doi:10.1021/ac300629p.

QuaMeter: Multivendor Performance Metrics for LC–MS/MS Proteomics Instrumentation

Ze-Qiang Ma[†], **Kenneth O. Polzin**[‡], **Surendra Dasari**[†], **Matthew C. Chambers**[†], **Birgit Schilling**[§], **Bradford W. Gibson**[§], **Bao Q. Tran**[⊥,‖], **Lorenzo Vega-Montoto**[†], **Daniel C. Liebler**[¶], and **David L. Tabb**[*,†,‖,¶]

[†]Department of Biomedical Informatics, Vanderbilt University Medical Center, Nashville, Tennessee 37232–8575, United States

[‡]Chemical and Physical Biology Admissions Program, Vanderbilt University Medical Center, Nashville, Tennessee 37232–0301, United States

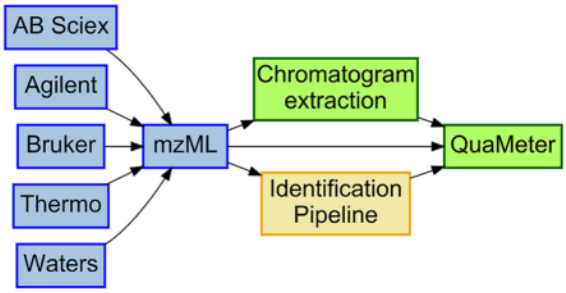
[§]Buck Institute for Research on Aging, Novato, California 94945, United States

[⊥]Department of Medicine, Vanderbilt University Medical Center, Nashville, Tennessee 37232–2358, United States

[‖]Mass Spectrometry Research Center, Vanderbilt University Medical Center, Nashville, Tennessee 37232–8575, United States

[¶]Department of Biochemistry, Vanderbilt University Medical Center, Nashville, Tennessee 37232–6350, United States

Abstract



LC-MS/MS-based proteomics studies rely on stable analytical system performance that can be evaluated by objective criteria. The National Institute of Standards and Technology (NIST) introduced the MSQC software to compute diverse metrics from experimental LC-MS/MS data, enabling quality analysis and quality control (QA/QC) of proteomics instrumentation. In practice, however, several attributes of the MSQC software prevent its use for routine instrument monitoring. Here, we present QuaMeter, an open-source tool that improves MSQC in several aspects. QuaMeter can directly read raw data from instruments manufactured by different vendors. The software can work with a wide variety of peptide identification software for improved reliability and flexibility. Finally, QC metrics implemented in QuaMeter are rigorously defined

[*]Corresponding Author: Phone: 615–936–0380. Fax: 615–343–8372. david.l. tabb@vanderbilt.edu.

The authors declare no competing financial interest.

Supporting Information
Additional information as noted in text. This material is available free of charge via the Internet at http://pubs.acs.org.

and tested. The source code and binary versions of QuaMeter are available under Apache 2.0 License at http://fenchurch.mc.vanderbilt.edu.

A shotgun proteomics experiment contains a chain of complex processing steps: cell or tissue lysis; protein denaturation, alkylation and digestion; liquid chromatographic (LC) separation of peptides, followed by electrospray ionization; and tandem mass spectrometry. Variation in the performance for any of these elements may impact proteomic identification and repeatability. The publication of LC-MS/ MS quality metrics by Paul Rudnick et al. at NIST, working in collaboration with the NCI CPTAC network, introduced a set of metrics that span this complex process,[1] enabling recognition of components that were operating at variance with their typical performance. The strategy makes use of defined quality control samples that are periodically analyzed between experimental samples in a queue for the mass spectrometer. The QC metrics are encoded in the MSQC computational pipeline from NIST for public use.

The NIST MSQC pipeline relies on a complex set of algorithms. Data from Thermo RAW files are first transcoded to mzXML, MS1, and MGF formats for subsequent processing. The MS1 files enable peptide precursor ion chromatograms to be assessed in the NIST ProMS software. The tandem mass spectra present in the data file are identified by either the SpectraST spectral library search engine[2] or the OMSSA database search algorithm.[3] The MSQC pipeline can then match precursor ion chromatograms with peptide identifications to compute its set of metrics and report them to a text file. Despite the need for tools to support QC in LC-MS/MS, MSQC has yet to achieve widespread use, in part because of important technical limitations. Its reliance on a modified ReAdW tool for reading raw data initially limited its application to instruments manufactured by Thermo Fisher Scientific (Waltham, MA). Its dependency on coordination among several different software packages may lead to misassociation of peptide identifications and tandem mass spectra when alternative file formats or high scan rate instrumentation are employed. Finally, adapting the pipeline for site-specific identification workflows (such as a different database search engine) is a nontrivial task.

In this work, we present the QuaMeter tool that encapsulates the key capabilities of MSQC while making several important additions. QuaMeter can read files from most mass spectrometry vendors via the ProteoWizard[4] library and does not lose time trancoding to other formats. The software accepts identification data from IDPicker,[5] so any identification engine that performs a target-decoy search and produces standard identification files (pepXML or mzIdentML) can be used. We demonstrate the generality of QuaMeter with data collected from instruments of three different vendors. We also examine the impact of peptide identification software choice on QC metrics. The improvements in QuaMeter make it a robust and flexible quality metric assessor with open source.

EXPERIMENTAL SECTION

Overview of QuaMeter

Figure 1 illustrates the QuaMeter workflow for computing QC metrics from LC-MS/MS experiments. The software accepts a spectrum file and corresponding filtered identifications file as inputs. QuaMeter reads the spectrum files in a variety of instrument-native and derived formats via the ProteoWizard library (see Figure 1 for a partial list). Translating to open formats, such as mzML, is generally unnecessary if peptide identification does not require it. Filtered identifications are read from IDPicker (version 3.0) report files. For each LC-MS/MS experiment, QuaMeter writes the QC metrics to a tab-delimited file. Scripts for combining multiple metric files are provided in the Supporting Information. QuaMeter was

written in C++, and its multithreaded architecture can take advantage of multicore CPUs. Binaries and source code for the software are available for download from http://fenchurch.mc.vanderbilt.edu.

QuaMeter starts by connecting the peptide identifications to their corresponding MS/MS scans in the spectrum file. For this, the software uses the instrument-assigned unique spectrum identifiers (NativeIDs). Both the spectral data and the IDPicker report associate each spectrum with a NativeID. Next, the software groups the identifications by sequence (including modifications) and charge state. Chromatographic peak information corresponding to each unique interpretation is extracted from the raw file using the CRAWDAD[6] module that is a part of the ProteoWizard library. Finally, the software computes a total of 42 QC metrics for the raw file, each metric measuring a different aspect of data acquisition. An overview and definitions for all implemented metrics are provided in Supporting Information Table S-1.

The peptide identification tools MyriMatch[7] (database search), TagRecon[8] (sequence tag-based database search), and Pepitome[9] (spectral library search) all incorporate ProteoWizard for both data import and identification export via pepXML and mzIdentML formats. Figure 1 highlights Pepitome because spectral library search is particularly appropriate for repeat identification of QC standard samples. In addition, both Pepitome and QuaMeter can process native file formats directly, thus eliminating the data conversion step. Raw identifications from this step are filtered within the IDPicker protein assembler. This allows the software to accept search results from any software that performs target-decoy searches and produces standard pepXML or mzIdentML formatted identification files.

To start quality assessment, QuaMeter requires the same spectral file that the identification tool was supplied, and it associates identified peptides with their corresponding MS/MS scans using NativeIDs. It should be noted that the propagation of the raw spectral designations is not always reliable during the identification process. Pepitome, MyriMatch, and TagRecon all support NativeIDs in their search results to ensure proper propagation of raw spectral designations. Some identification tools do not contain NativeIDs in their search outputs. If so, NativeIDs can be manually added to the search output before IDPicker process. We have tested QuaMeter to work with Sequest,[10] X!Tandem,[11] and Mascot[12] search results (data not shown). Example scripts to manually add NativeIDs to these search results are provided in the Supporting Information.

QuaMeter can read native file formats from multiple instrument vendors for quality assessment. However, if the identification tool does not support native file formats, it will be necessary to convert raw files to open formats. Since QuaMeter uses NativeIDs to connect identified peptides and their corresponding MS/MS scans, we recommend using the mzML standard format if data conversion is necessary. This reduces the risk of breaking the associations between peptide identifications and spectral data. The MSConvert tool in the ProteoWizard library supports the conversion of multiple vendor formats to mzMLs.

Data Sources and Bioinformatics

We tested QuaMeter on several data sets spanning six different mass spectrometers. Table 1 summarizes all the data sets used in this study. In brief, we analyzed three different samples (bovine serum albumin, β-galactosidase, and yeast) on six different platforms. Each sample was analyzed in replicates spread over time to perform instrument quality control and maintenance. Full experimental details of these data sets are available in the Supporting Information.

The MS/MS scans present in all data sets were identified using MyriMatch or Pepitome. MyriMatch is a database search engine, whereas Pepitome is a spectral library search engine. Table 1 summarizes the sequence databases and mass tolerances used in all searches. Detailed configuration parameters for all searches are listed in the Supporting Information. MyriMatch was configured to derive semitryptic peptides from the protein database while using carbamidomethylation of cysteine (+57.0125 Da), oxidation of methionine (+15.996 Da), and formation of N-terminal pyroglutamine (−17.0265 Da) as variable modifications. Pepitome was configured to consider only fully tryptic and semitryptic peptides from the NIST ion trap spectral library (http://peptide.nist.gov). All search engines produced identifications in pepXML format.

IDPicker software filtered the peptide identifications from all search engines at a false discovery rate (FDR) of 5% unless otherwise stated. For MyriMatch, the software automatically combined the MVH and XCorr scores for FDR filtering. IDPicker was configured to use HGT, Kendall-Tau, and mzFidelity scores for filtering Pepitome results. Peptides passing the FDR thresholds were assembled into protein identifications following parsimony rules, and proteins with at least two distinct peptide identifications were considered for further analysis. QuaMeter software processed the raw files and corresponding IDPicker identifications to produce QC metrics for each data file. Since we wanted to compare the QC metrics generated from QuaMeter to that of MSQC software, we made the IDPicker identifications accessible to the MSQC software via an AWK script. We also created scripts in the R statistical programming language for combining multiple metric files to perform variability analysis. Detailed software configurations, scripts, and data processing methods are presented in the Supporting Information.

RESULTS AND DISCUSSION

Comparison of QuaMeter and NIST MSQC Metrics

For its initial release, QuaMeter implemented the same metrics as in the MSQC package to facilitate comparison. Data from bovine serum albumin (BSA) QC samples analyzed on a Thermo Fisher LTQ-XL mass spectrometer were employed to evaluate how code differences altered values reported by these two tools. Pepitome identified peptides from the data files, and IDPicker filtered the results at 2% FDR. For each data file, QuaMeter computed the QC metrics using IDPicker results. We modified the MSQC software to accept the same IDPicker results to compute the QC metrics.

Figure 2 illustrates the correspondence between QuaMeter and MSQC outputs for a set of representative metrics. Median precursor m/z error for +2 peptides (MS1–5A in NIST nomenclature) is shown in the top-left panel as a representative of metrics with very good agreement between both implementations. Most metrics representing peptide identifications (such as P-2A, number of MS/MS spectra identifying tryptic peptide ions; P-2B, number of tryptic peptide ions identified; P-2C, number of unique tryptic peptides identified; and P-3, ratio of semitryptic/tryptic peptides) yielded similar results.

The key C-2A metric was a note of discord between QuaMeter and MSQC. This metric describes the time period in which the middle 50% of peptides are identified. Despite measures taken to produce the metrics as comparably as possible, QuaMeter disagreed with MSQC (top-right panel in Figure 2). A code inspection revealed that MSQC vacillates in whether modifications or precursor charge differentiate identifications. Because C-2A plays a role in the computation of many other metrics, the QuaMeter implementation was changed to a “distinct match” rule (under which either a sequence difference, a modification, or a precursor charge difference results in the identification counting as a new peptide). This

change leads to a more representative metric because modifications to peptides alter their chromatographic behavior.

MSQC and QuaMeter detect chromatographic peaks using distinct tools. Hence, differences in metrics involving MS1 peak intensity and peak width are not surprising. The DS-3B metric evaluates the maximum intensity versus the intensity at the time when MS/MS was triggered for the 50% of peptides with the least-intense trigger intensities (see bottom-left panel in Figure 2). The MSQC software estimated far lower peak intensity maxima than expected from manual inspection, resulting in little correlation for this metric. This effect propagated through metrics describing the chromatographic process as well as dynamic sampling. The C-4C metric (lower-right panel in Figure 2) reports the median peak width (full width at half max; fwhm) for identified peptides in the median retention decile. QuaMeter consistently reports lower peak widths compared with MSQC. We note, however, that these comparisons were made with an early version of MSQC that uses a modified ReAdW tool for chromatogram extraction. The latest MSQC software substitutes ProMS for chromatogram analysis to produce more reliable chromatographic data (personal communications). We could not, however, acquire a new build of MSQC for evaluation in time for publication.

Chromatographic peak detection plays a vital role in computing QC metrics. QuaMeter uses the CRAWDAD module for detecting peaks from raw data acquired on a variety of instruments. The software provides an option to export the peak detection results in mz5 format[13] that can be visualized by the SeeMS tool in ProteoWizard. Figure 3 illustrates the extracted ion chromatograms (XIC) of experimentally measured peptide intensities and CRAWDAD-modeled peaks of representative peptides from three instrument platforms. For high abundance peptides that were identified with many MS/MS scans, CRAWDAD modeled peaks that closely resemble experimental measurements (top panels in Figure 3). In addition, CRAWDAD also showed excellent performance for low abundance peptides with noisy experimental XIC or interfering peaks (bottom panels in Figure 3). Although it is hard to assess the success rate for CRAWDAD's peak detection, these example traces imply that it is able to yield good results across a wide range of peptide abundances.

Interpretation of QuaMeter results is not a trivial task. A large number of metrics are generated for each LC-MS/MS experiment. The analysis of instrument stability may start from several key metrics, such as MS1–5A (median real value of precursor errors), MS2–3 (median peak count in the MS/MS scans), P-2A (number of MS/MS spectra identifying tryptic peptide ions), P-2B (number of tryptic peptide ions identified), C-2A (time period over which the middle 50% of peptides were identified), C-3A (median peak widths for all identified unique peptides), DS-2A (number of MS1 scans taken over C-2A), DS-2B (number of MS/MS scans taken over C-2A), and IS-2 (median m/z value for all identified peptides). The meanings and interpretations of these QC metrics have been described in the MSQC publication.[1]

To determine whether the instrument has stable performance, traditional process control methods, such as Shewhart charts, can be applied to visualize the variation of individual metrics across a number of LC-MS/MS experiments; however, this is not an ideal approach because metrics contain mutual information, and some metrics may be more informative than others in measuring instrument stability. Hence, it will be invaluable to develop multivariate statistical methods to handle such data appropriately. Ideally, these methods would be able to summarize QuaMeter metrics to a single quality score that could enable on-the-fly instrument QC.

Multivendor Compatibility

We tested QuaMeter's compatibility with six different mass spectrometers: Thermo Fisher LTQ-XL, LTQ-Orbitrap, LTQ-Velos, Bruker Daltonics HCT Ultra, AB SCIEX QSTAR Elite, and TripleTOF 5600. Instrument raw files from Thermo and Bruker were converted to mzML[14] format using the MSConvert tool in ProteoWizard. AB SCIEX data were converted to mzML format using the AB SCIEX MS Converter (version 1.2) because peak centroiding was not yet accessible in ProteoWizard. All data were searched by MyriMatch, and search results were processed by IDPicker. Filtered identifications were then processed by QuaMeter to compute QC metrics.

QuaMeter produces the 42 metrics described in the NIST MSQC publication. These metrics are intended to measure the stability of an instrument on six different axes: chromatography, ion source, dynamic sampling, MS1 signals, MS/MS signals, and peptide identification. For instance, Figure 4 illustrates a set of four metrics that summarize chromatographic data and variability for five instruments. The C-2A metric, which measures the time frame during which the middle 50% of peptides are identified, is very small for the TripleTOF and QSTAR data sets. These values correlate with the rapid LC gradient employed by this laboratory for QC. The C-2A metric from the HCT data set showed greater variability because different BSA samples for this instrument were separated by different HPLC columns and gradients.

We also observed significant variations in other QC metrics computed for this data set, such as the number of identifications (Supporting Information Figure S-1). The C-4A, C-4B, and C-4C metrics report the median peak width for peptides identified in the early, late, and middle retention time periods, respectively (Figure 4). We observed that peak widths of peptides vary with their elution time. For instance, late-eluting peptides tend to have wider peaks than early-eluting peptides in several data sets (Figure 4). These plots demonstrate two important aspects of our QC workflow: QuaMeter enjoys cross-instrument compatibility, and the QC metrics can be used to detect trends in the data. A complete plot of all computed metrics for five instruments is available in the Supporting Information, Figure S-1.

QuaMeter metrics are also useful for detecting abnormal performance of an instrument. For instance, an early analysis of TripleTOF data flagged six data files with low identification rates compared with other QC experiments. A close examination of their QuaMeter metrics and instrument log revealed a systematic mass accuracy drift due to variation in the ambient temperature. The instrument operator later identified a failed air handler in the laboratory as the root cause of this problem. We recalibrated the ion masses to account for the drift. Reprocessing the new data significantly improved the identification rates for the problematic data files (Supporting Information Figure S-2).

Impact of Peptide Identification on QC Metrics

QuaMeter relies on identified peptides to compute QC metrics. Hence, changes in the identification methods may ripple into the QC metrics. To evaluate this, we employed a yeast lysate LC-MS/MS data set with five technical replicates acquired on a Thermo Fisher LTQ-Velos mass spectrometer. Spectra were identified using two different methods: MyriMatch database search and Pepitome spectral library search. Filtered identi-fications by IDPicker were read into QuaMeter for QC evaluation.

Figure 5 illustrates a handful of QC metrics that are idiosyncratic to the search method. For instance, thePepitome library search identified 15% more spectra than MyriMatch database search (the P-2A metric in Figure 5). Changes in the identification methods, however, did not produce substantial changes in most metrics. In addition, the variation observed for a QC

metric did not depend significantly on the search method (Figure 5). Hence, it is likely that the identification method has a limited effect on the evaluation of analytical system performance and measurement of its variability. Because spectral library searches are faster than database searches, we recommend coupling QuaMeter with Pepitome for routine analysis of QC standards.

This test also demonstrates that QuaMeter works well not only for simple samples such as BSA and beta-galactosidase but also for complex mixtures. As mass spectrometer scan rates increase, samples of greater complexity will be necessary to test their capabilities.

CONCLUSIONS

We present an open-source tool that computes objective metrics for the evaluation of shotgun proteomics instrumentation performance. QuaMeter represents an advance over the previous MSQC tool by supporting most mass spectrometer vendors via the use of the ProteoWizard library. The ability to work with IDPicker identification data allows it to be incorporated in any identification workflow that produces pepXML or mzIdentML files. The improvements in QuaMeter make it a reliable and flexible tool for shotgun proteomics QC analysis.

Shotgun proteomics laboratories will benefit handsomely from tools to automate QA/QC. QuaMeter represents a significant step in the right direction. The software accepts raw data and identification data from a variety of sources to produce QC metrics. Drawing conclusions from QuaMeter output is currently less established, although statistical efforts will clarify interpretation of the metrics. A subset of key metrics should be determined to evaluate the analytical systems in routine practice. Statistical methods that recognize mutual information in the metrics will be valuable in support of on-the-fly instrument QC.

Future directions for QuaMeter include a number of goals. First, recording metrics for experiments to a database rather than a collection of text files will greatly improve the production utility of the software. Second, incorporating assessments of MS/MS quality[15] would be much faster and more adaptable than incorporating peptide identifications. Optimizing the strategies by which metric values can be evaluated to diagnose sources of instrument variability will be essential. As these techniques mature, QC metrics promise to automate recognition of instrument inconsistency in real time before critical samples are wasted.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

The authors thank Stephen Stein and Paul Rudnick for sharing source code of the NIST MSQC software. This work is funded via the National Cancer Institute Grants CA126470 and CA159988 to D. C. Liebler, CA126218 to D. L. Tabb, and CA126477 to B. W. Gibson (subcontract to UCSF). B. Schilling and B. W. Gibson acknowledge the support of instrumentation from the National Center for Research Resources shared instrumentation Grant S10 RR024615 and AB SCIEX for evaluation of the TripleTOF 5600 system at the Buck Institute. We thank the anonymous reviewers for their invaluable comments and suggestions.

References

1. Rudnick PA, Clauser KR, Kilpatrick LE, Tchekhovskoi DV, Neta P, Blonder N, Billheimer DD, Blackman RK, Bunk DM, Cardasis HL, Ham AJL, Jaffe JD, Kinsinger CR, Mesri M, Neubert TA, Schilling B, Tabb DL, Tegeler TJ, Vega-Montoto L, Variyath AM, Wang M, Wang P, Whiteaker

JR, Zimmerman LJ, Carr SA, Fisher SJ, Gibson BW, Paulovich AG, Regnier FE, Rodriguez H, Spiegelman C, Tempst P, Liebler DC, Stein SE. Mol Cell Proteomics. 2010; 9:225–241. [PubMed: 19837981]

2. Lam H, Deutsch EW, Eddes JS, Eng JK, King N, Stein SE, Aebersold R. Proteomics. 2007; 7:655–667. [PubMed: 17295354]
3. Geer LY, Markey SP, Kowalak JA, Wagner L, Xu M, Maynard DM, Yang X, Shi W, Bryant SH. J Proteome Res. 2004; 3:958–964. [PubMed: 15473683]
4. Kessner D, Chambers M, Burke R, Agus D, Mallick P. Bioinformatics. 2008; 24:2534. [PubMed: 18606607]
5. Ma ZQ, Dasari S, Chambers MC, Litton MD, Sobecki SM, Zimmerman LJ, Halvey PJ, Schilling B, Drake PM, Gibson BW, Tabb DL. J Proteome Res. 2009; 8:3872–3881. [PubMed: 19522537]
6. Finney GL, Blackler AR, Hoopmann MR, Canterbury JD, Wu CC, MacCoss MJ. Anal Chem. 2008; 80:961–971. [PubMed: 18189369]
7. Tabb DL, Fernando CG, Chambers MC. J Proteome Res. 2007; 6:654–661. [PubMed: 17269722]
8. Dasari S, Chambers MC, Slebos RJ, Zimmerman LJ, Ham AJL, Tabb DL. J Proteome Res. 2010; 9:1716–1726. [PubMed: 20131910]
9. Dasari S, Chambers MC, Martinez MA, Carpenter KL, Ham AJL, Vega-Montoto LJ, Tabb DL. J Proteome Res. 2012; 11:1686–1695. [PubMed: 22217208]
10. Eng JK, McCormack AL, Yates JR III. J Am Soc Mass Spectrom. 1994; 5:976–989.
11. Craig R, Beavis RC. Bioinformatics. 2004:921.
12. Perkins DN, Pappin DJ, Creasy DM, Cottrell JS. Electrophoresis. 1999; 20:3551–3567. [PubMed: 10612281]
13. Wilhelm M, Kirchner M, Steen JAJ, Steen H. Mol Cell Proteomics. 2012:11.
14. Deutsch E. Proteomics. 2008; 8:2776–2777. [PubMed: 18655045]
15. Ma ZQ, Chambers MC, Ham AJL, Cheek KL, Whitwell CW, Aerni HR, Schilling B, Miller AW, Caprioli RM, Tabb DL. J Proteome Res. 2011; 10:2896–2904. [PubMed: 21520941]

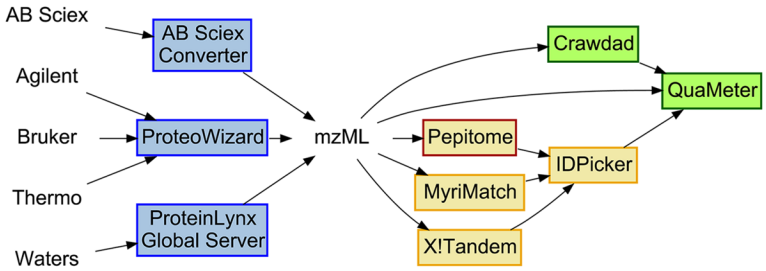


Figure 1.
Workflow diagram for QuaMeter operation.

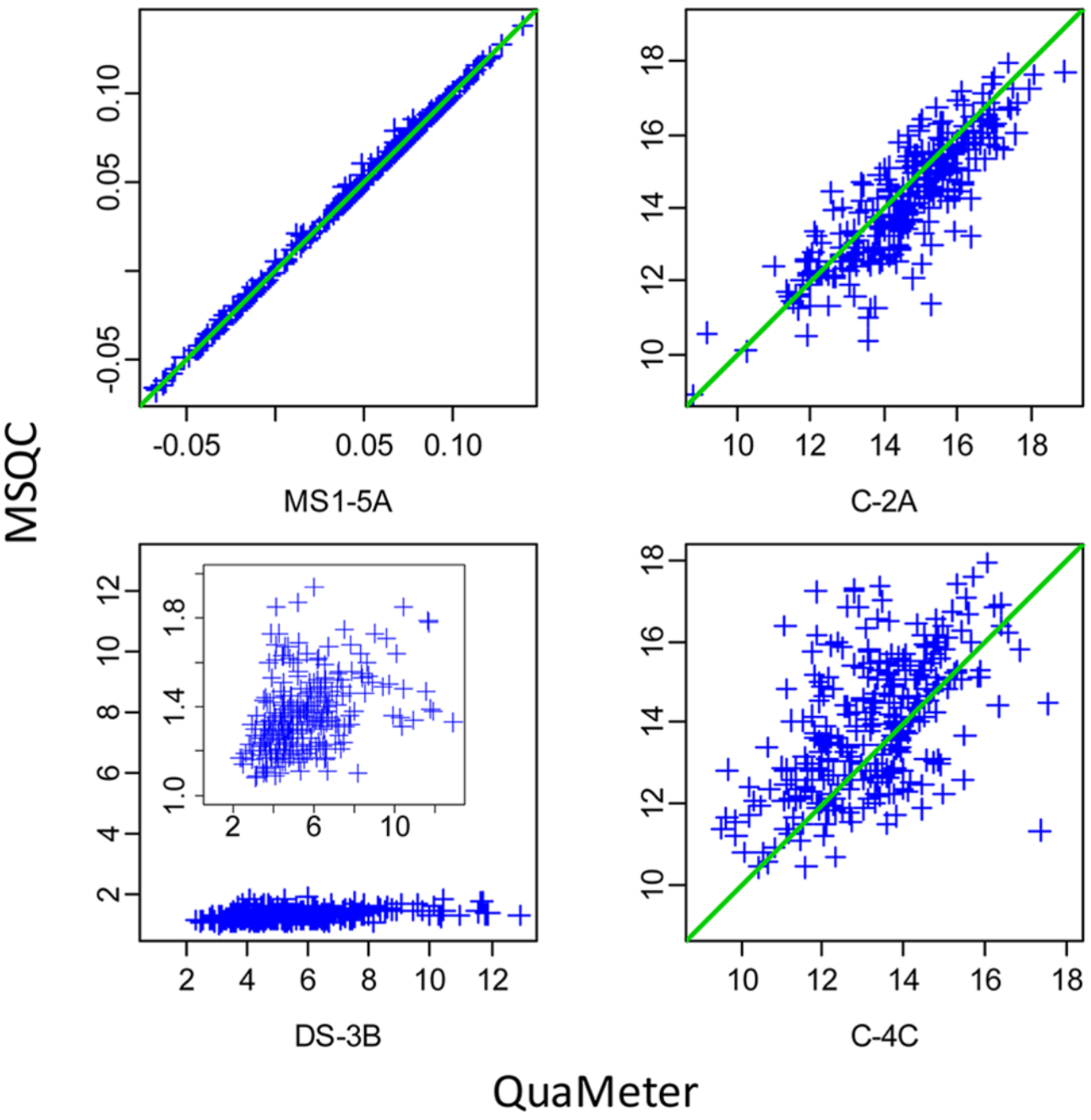


Figure 2.
QuaMeter generates metrics similar to MSQC with the exception of chromatographic metrics due to the use of distinct chromatogram extraction tools. Metrics were generated from BSA QC experiments collected on a Thermo Fisher LTQ-XL mass spectrometer. The subgraph in bottom-left panel is a rescaled plot for DS-3B metric.

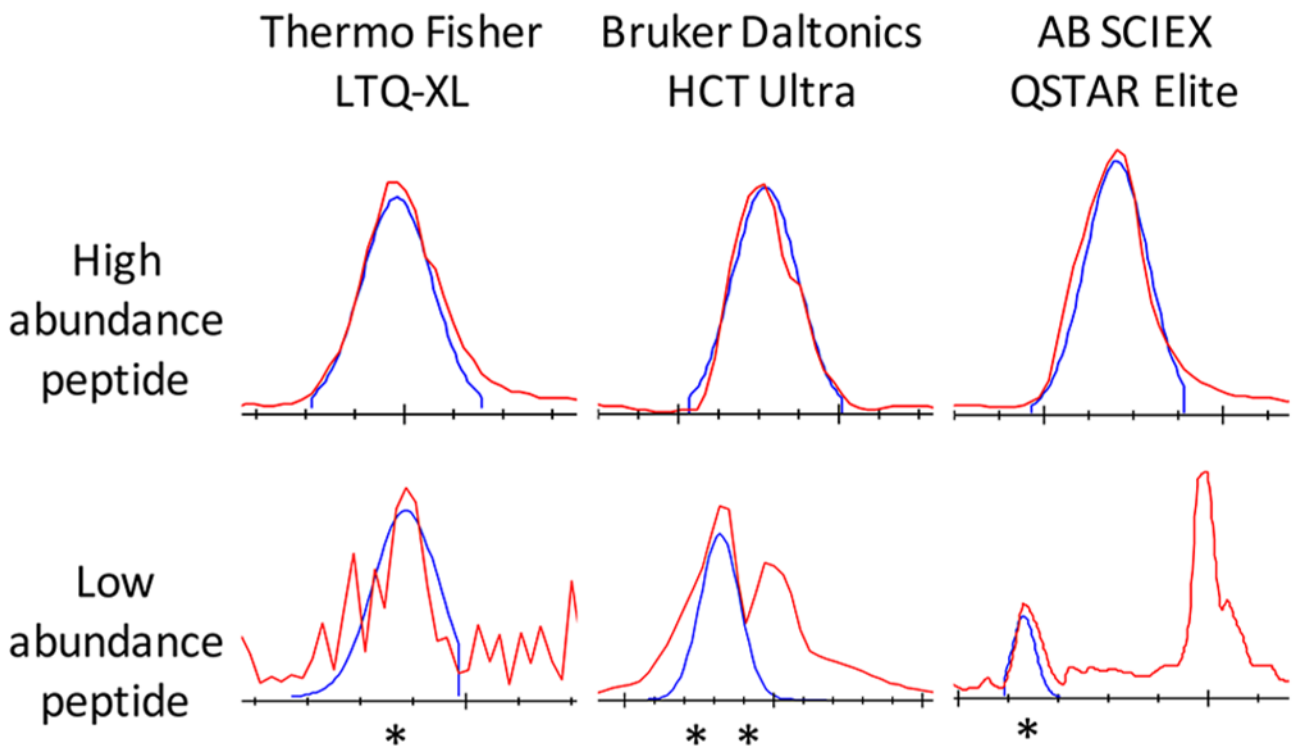


Figure 3.
QuaMeter detects chromatographic peaks from instruments via the CRAWDAD module in ProteoWizard. Red lines represent experimentally measured intensities in MS, and blue lines are extracted ion chromatograms generated by CRAWDAD. Asterisks for the low abundance peptides signify the acquisition times for identified MS/MS scans.

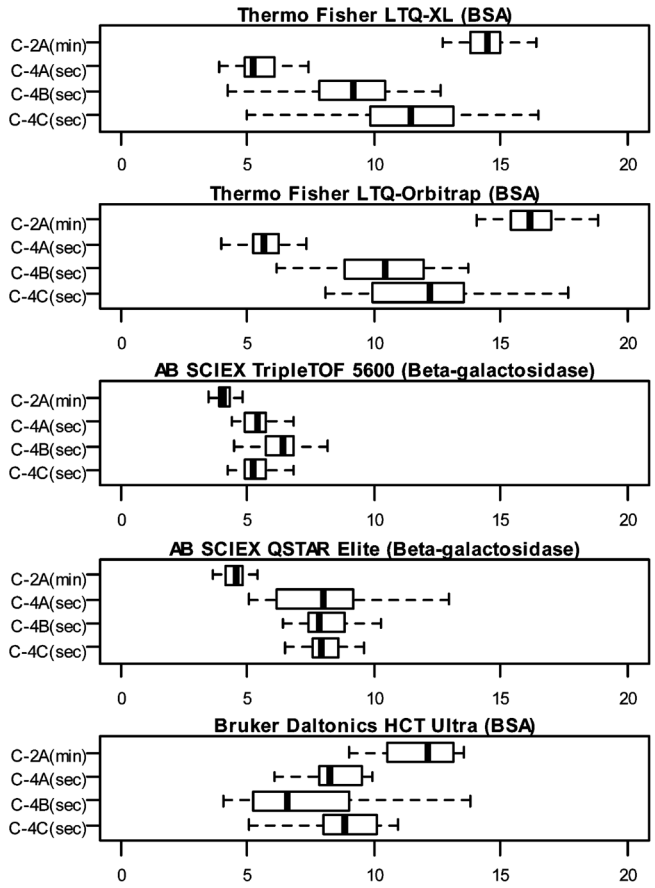


Figure 4.

QuaMeter computes QC metrics for multiple instrument platforms. Standard samples, such as BSA or β-galactosidase, were analyzed for routine instrument evaluation. C-2A: time period over which middle 50% of peptides were identified. C-4A, C-4B, C-4C: median peak width for identified peptides in first, last, and median RT decile.

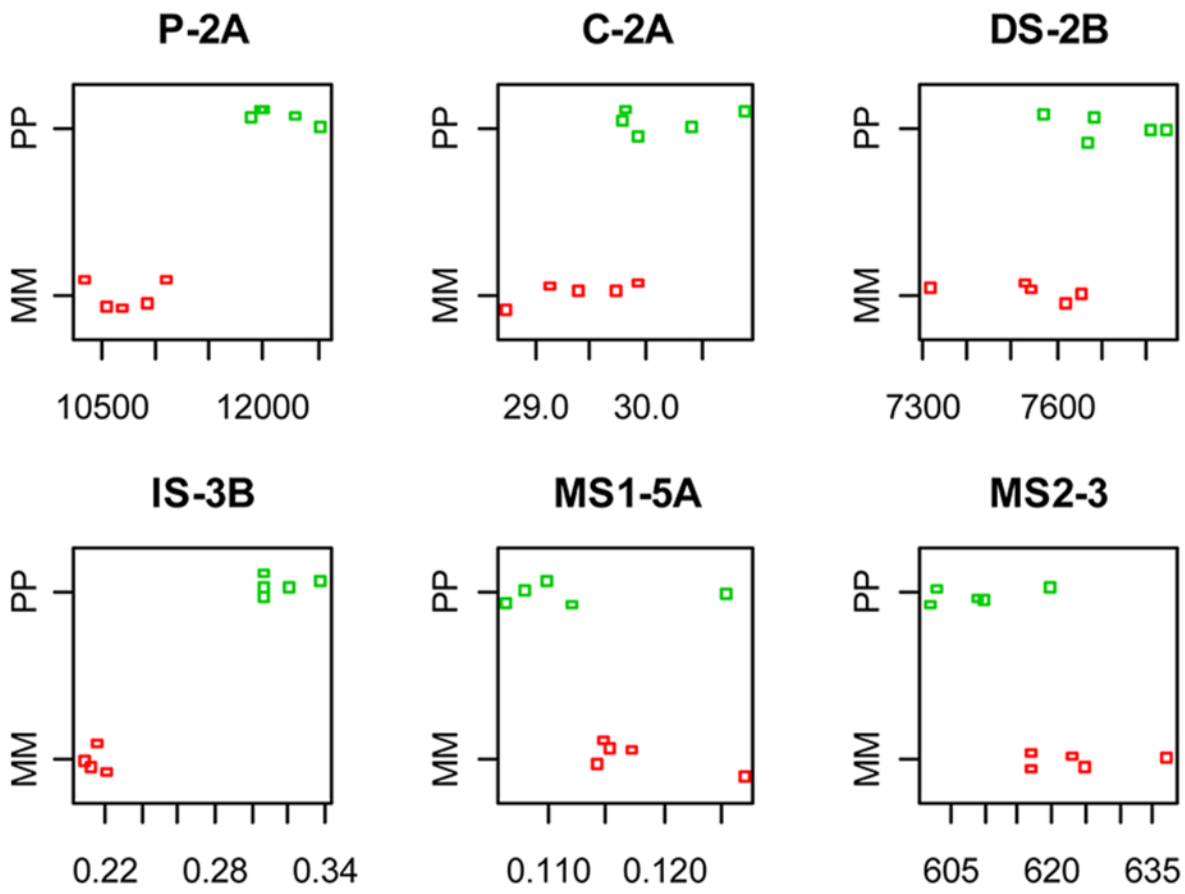


Figure 5.

Distinct peptide identification methods produce different QC metrics with similar variation. Five technical replicates of yeast lysate samples were analyzed on a Thermo Fisher LTQ-Velos mass spectrometer. Spectra were identified by a MyriMatch (MM) database search and a Pepitome (PP) spectral library search. Identifications from each search engine were used to compute QC metrics. P-2A, number of MS2 spectra identifying tryptic peptide ions; C-2A, time period over which middle 50% of peptides were identified; DS-2B, number of MS2 scans taken over C-2A; IS-3Bn number of 3+ peptides over 2+ peptides; MS1–5A, median real value of precursor errors; MS2–3, median number of peaks in MS2 scans.

Table 1

Experimental Data Sets Summary[a]

instrument	vendor	sample[a]	no. files	precursor/fragment *m/z* tolerance[b]	databases used for search
LTQ-XL	Thermo Fisher	BSA	280	1.25/0.5	RefSeq.BOVINE
LTQ-Orbitrap	Thermo Fisher	BSA	53	10*/0.5	RefSeq.BOVINE
HCT Ultra	Bruker Daltonics	BSA	24	1.25/0.5	RefSeq.BOVINE
QSTAR Elite	AB SCIEX	Bgal	23	100*/0.4	UniProt.ECOLI
TripleTOF 5600	AB SCIEX	Bgal	60	50*/50*	UniProt.ECOLI
LTQ-Velos	Thermo Fisher	yeast	5	1.25/0.5	SGD.orf_trans_all

[a]BSA (bovine serum albumin) and Bgal (*Escherichia coli* β-galactosidase) were purchased as protein standards. Yeast reference materials were generated as part of the NCI CPTAC program. All samples were analyzed via LC-MS/MS experiments.

[b]Asterisk denotes mass tolerances in parts-per-million.