Design, Implementation and Multisite Evaluation of a System Suitability Protocol for the Quantitative Assessment of Instrument Performance in Liquid Chromatography-Multiple Reaction Monitoring-MS (LC-MRM-MS)*⑤

**Susan E. Abbatiello‡[g], D. R. Mani‡, Birgit Schilling§, Brendan MacLean¶,
Lisa J. Zimmerman‖, Xingdong Feng**, Michael P. Cusack§, Nell Sedransk**,
Steven C. Hall‡‡, Terri Addona‡, Simon Allen‡‡, Nathan G. Dodder§§,
Mousumi Ghosh¶¶, Jason M. Held§, Victoria Hedrick‖‖, H. Dorota Inerowicz‖‖,
Angela Jackson[a], Hasmik Keshishian‡, Jong Won Kim[b], John S. Lyssand[c],
C. Paige Riley‖‖, Paul Rudnick§§, Pawel Sadowski[c], Kent Shaddox‖, Derek Smith[a],
Daniela Tomazela¶, Asa Wahlander[c], Sofia Waldemarson[c], Corbin A. Whitwell‖,
Jinsam You[b], Shucha Zhang[d], Christopher R. Kinsinger[e], Mehdi Mesri[e],
Henry Rodriguez[e], Christoph H. Borchers[a], Charles Buck‖‖, Susan J. Fisher‡‡,
Bradford W. Gibson§, Daniel Liebler‖, Michael MacCoss¶, Thomas A. Neubert[c],
Amanda Paulovich[d], Fred Regnier‖‖, Steven J. Skates[f], Paul Tempst¶¶, Mu Wang[b],
and Steven A. Carr‡**

From the ‡Broad Institute of MIT and Harvard, Cambridge, Massachusetts 02142, USA; §Buck Institute for Research on Aging, Novato, California 94945, USA; ¶Department of Genome Sciences, University of Washington School of Medicine, Seattle, Washington 98195, USA; ‖Vanderbilt University, Nashville, Tennessee 37232, USA; **National Institute of Statistical Sciences, Research Triangle Park, North Carolina 27709, USA; ‡‡Department of Obstetrics, Gynecology, and Reproductive Sciences, University of California, San Francisco, San Francisco, California 94143, USA; §§National Institute of Standards and Technology, Gaithersburg, Maryland 20899, USA; ¶¶Memorial Sloan-Kettering Cancer Center, New York, New York 10065, USA; ‖‖Purdue University, West Lafayette, Indiana 47907, USA; [a]University of Victoria-Genome BC Proteomics Centre, Victoria, British Columbia V8Z 7X8 CAN; [b]Department of Biochemistry and Molecular Biology, Indiana University School of Medicine, Indianapolis, Indiana 46202, USA; [c]Kimmel Center for Biology and Medicine at the Skirball Institute and Department of Pharmacology, New York University School of Medicine, New York, New York 10016, USA; [d]Fred Hutchinson Cancer Research Center, Seattle, Washington 98109, USA; [e]National Cancer Institute, National Institutes of Health, Bethesda, Maryland 20892, USA; [f]Biostatistics Center, Massachusetts General Hospital, Boston, Massachusetts 02114

Multiple reaction monitoring (MRM) mass spectrometry coupled with stable isotope dilution (SID) and liquid chromatography (LC) is increasingly used in biological and clinical studies for precise and reproducible quantification of peptides and proteins in complex sample matrices. Robust LC-SID-MRM-MS-based assays that can be replicated across laboratories and ultimately in clinical laboratory settings require standardized protocols to demonstrate that the analysis platforms are performing adequately. We developed a system suitability protocol (SSP), which employs a predigested mixture of six proteins, to facilitate performance evaluation of LC-SID-MRM-MS instrument platforms, configured with nanoflow-LC systems interfaced to triple quadrupole mass spectrometers. The SSP was designed for use with low multiplex analyses as well as high multiplex approaches when software-driven scheduling of data acquisition is required. Performance was assessed by monitoring of a range of chromatographic and mass spectrometric metrics including peak width, chromatographic resolution, peak capacity, and the variability in peak area and analyte retention time (RT) stability. The SSP, which was evaluated in 11 laboratories on a total of 15 different instruments, enabled early diagnoses of LC and MS anomalies that indicated suboptimal LC-MRM-MS performance. The observed range in variation of each of the metrics scrutinized serves to define the criteria for optimized LC-SID-MRM-MS platforms for routine use, with pass/fail criteria for system suitability performance measures defined as peak area coefficient of variation <0.15, peak width coefficient of variation <0.15, standard deviation of RT <0.15 min (9 s), and the RT drift <0.5min (30 s). The deleterious effect of a marginally performing LC-SID-MRM-MS system

on the limit of quantification (LOQ) in targeted quantitative assays illustrates the use and need for a SSP to establish robust and reliable system performance. Use of a SSP helps to ensure that analyte quantification measurements can be replicated with good precision within and across multiple laboratories and should facilitate more widespread use of MRM-MS technology by the basic biomedical and clinical laboratory research communities. *Molecular & Cellular Proteomics 12: 10.1074/mcp.M112.027078, 2623–2639, 2013.*

Targeted analysis by liquid chromatography-stable isotope dilution-multiple reaction monitoring-MS (LC-SID-MRM-MS)[1] (also referred to as LC-SID-SRM-MS) has experienced rapid expansion over the last several years for precise relative quantification of peptides in the context of basic biological studies (1–4) and for verification of candidate biomarkers in clinical applications (5–10). As a quantitative proteomics tool, LC-SID-MRM-MS offers numerous benefits. First, the overall precision of analyte quantification, taking into account all sample processing steps from digestion through data acquisition, is often in the range of 5–10% (CV) above the limit of quantification (LOQ) and ≤25% at the LOQ when stable isotope labeled internal standards are employed. The excellent precision is inherent to LC-SID-MRM-MS experimental design and how data are acquired. For example, these assays employ classical SID methodology in which synthetic peptide standards incorporating one or more labeled amino acids (^{13}C, ^{15}N or a combination thereof) are spiked at known amounts into the samples thereby enabling the endogenous peptide concentration to be determined (1–3). Confidence in detection and quantification of analytes in SID-MRM-MS is based on multiple orthogonal measurements, specifically (1) the labeled internal standard and analyte must co-elute chromatographically, (2) the heavy and light peptides fragment identically (yielding fragment ions that either have identical m/z values or are shifted upward in mass if they contain the label), and (3) the fragment ions have the same relative abundance in both the analyte and internal standard channels. Second, analyte detection sensitivities on the order of ELISA assays (low ng/ml) are achievable when either sample fractionation or stable isotope standards with capture by antipeptide antibodies (SISCAPA) are incorporated into the assay (5, 11–13). Finally, several hundred analytes can be multiplexed into a single LC-SID-MRM-MS experiment using scheduling software that segregates collection of data from subsets of the analyte peptides into different time periods in an LC-MRM-MS analysis (10, 14, 15). This feature enables implementation of extremely efficient and high throughput quanti-

tative protein assays. Accordingly, these attributes make quantitative LC-SID-MRM-MS workflows attractive for adaptation into a clinical setting (16, 17). However, before clinical adoption becomes a reality, robust procedures must be in place to ensure that optimal performance of an LC-SID-MRM-MS instrument platform is maintained throughout the progression of a particular assay (18).

To date, little work has been presented on the development of standardized protocols that quantitatively assess the "suitability" (19) of nanoflow (typically 150–300 nL/min) high performance liquid chromatography (HPLC) systems interfaced to triple quadrupole mass spectrometers during the course of an LC-SID-MRM-MS protein assay. Although promising LC-SID-MRM-MS biomarker verification studies were recently reported for cardiovascular disease (8), cancer (7, 10), and other disorders (9), this approach is still in its infancy. Development of quantitative protein/peptide LC-SID-MRM-MS experiments is challenged by many of the same hurdles as those for small molecules and as this technology evolves, other unique issues relevant to separation and quantification of peptides are being realized (20). In the most systematic and thorough interlaboratory evaluation of LC-SID-MRM-MS to date, eight laboratories followed a standard operating procedure (SOP) and analyzed identical sample sets to determine the limits of quantification and measures of variation using a set of 10 tryptic peptides generated from seven proteins spiked into human plasma at known concentrations (21). The results from this study detailed the range of reproducibility attainable for individual laboratories performing an identical experimental protocol, based on increasing complexity of the sample workflow (21). Although the focus of this study was on the contributions of sample handling and processing to the intra- and interlaboratory precision (CV) of the measurements, it also uncovered several common issues that contributed to intralaboratory variability. Most notably, HPLC-related problems resulting in peak tailing or random peak shape deformities caused inconsistencies in peak area integration and subsequent quantitative calculations. These performance degradations often went unnoticed by the laboratory that generated the data, and were only discovered when all of the data were evaluated by a small team of expert analysts. In addition, this study confirmed that use of stable isotope labeled peptides as internal standards for each peptide analyte helps to dampen measurement imprecision introduced by HPLC and MS-associated problems, but did not eliminate them, nor did it improve interlaboratory accuracy (22).

Recently, measures have been proposed to quantitatively monitor aspects of discovery-based proteomics approaches to better understand technical variability associated with chromatography, dynamic sampling, ion source configuration, signal intensity of MS and MS/MS scans, and peptide identification for data-dependent HPLC-MS/MS acquisitions (23–26). Likewise, there is a critical need for standardized methods to demonstrate that LC-SID-MRM-MS analysis plat-

---

[1] The abbreviations used are: LC-SID-MRM-MS, Liquid chromatography-stable isotope dilution-multiple reaction monitoring-MS; CV, coefficient of variation; LOQ, limit of quantification; SSP, system suitability protocol; FWHM, full width half maximal; RT, retention time.

forms are performing optimally. However, the approaches developed for discovery LC-MS/MS platform evaluation are not directly adaptable to LC-SID-MRM-MS systems, which have their own unique requirements for system performance assessment. In this context, we now describe the development and evaluation of an easy to implement system suitability protocol (SSP) to assess performance metrics of triple quadrupole-based nanoLC-SID-MRM-MS instrument configurations. In this investigation, 11 laboratories comprising 15 individual nanoLC-SID-MRM-MS platforms (from 4 different vendors, including 8 different models of mass spectrometers) participated in development and evaluation of an SSP, including peptide selection and platform-specific LC and MRM-MS optimization. Specific chromatographic and MS metrics including peak width, chromatographic resolution, peak capacity and the variability in peak area and retention time stability were monitored to assess reproducibility of replicate injections of a commercially available peptide mixture generated from trypsin digestion of 6 bovine proteins. Our study demonstrates that the SSP facilitated rapid detection, diagnosis and correction of system problems that were a source of performance degradation in terms of precision (CVs) and sensitivity (limits of detection and quantification, LOD/LOQ). A critical part of the study involved optimization and use of vendor-neutral data analysis tools, including Skyline (27), for rapid assessment of reproducibility and data quality in individual laboratories. These tools were then implemented across multiple sites for comparison of variability among different platform configurations. The SSP was also shown to be of value in understanding sources of variation across multiple laboratories performing an identical experimental protocol as well as for tracking an individual instrument over an extended period of time. The effect of a marginally performing LC-SID-MRM-MS system on assay sensitivity, determined by calculating LOQ values for peptide targets, is illustrated and appropriate limits for all system suitability parameters monitored are proposed.

MATERIALS AND METHODS

Materials—An equimolar predigested "Bovine 6 Protein Mix" (P/N PTD/00001/63) was purchased from Bruker-Michrom, Inc. (Auburn, CA). ReproSil-Pur C18-AQ resin (3 μm particle size) was purchased from Dr. Maisch, GmbH (Ammerbuch-Entringen, Germany). Picofrit self-pack columns, 75 μm ID, 10 μm ID tip, were purchased from New Objective (Woburn, MA).

Study Samples and Sample Preparation—The commercial "Bovine 6 Protein Mix" referred to as 6ProteinMix-QC is a trypsin predigested mixture containing beta lactoglobulin, lactoperoxidase, carbonic anhydrase, glutamate dehydrogenase, alpha casein, and serum albumin with each protein at an equimolar amount of 100 pmol per commercial vial (iodoacetic acid was used by the vendor to alkylate cysteine residues). The 6ProteinMix-QC aliquots were centrally prepared at Vanderbilt University by generating stock solutions with concentrations of 1 pmol/μl per protein using an aqueous solution of 30% acetonitrile (v/v) and 0.1% formic acid (v/v) for dissolution. The stock solutions were stored as 10 μl aliquots (at 1 pmol/μl concentration) at −80 °C, and shipped frozen on dry ice to the participating laborato-

ries. As described in detail in the SOP (see supplementary Methods), before analysis individual sites further diluted the stock solution to a working solution of 50 fmol/μl using an aqueous diluent containing 3% acetonitrile (v/v) and 0.1% formic acid (v/v).

MRM Assay Method Development - Skyline—MRM-MS transition lists were developed that could be applied to all 15 participating triple quadrupole mass spectrometers spanning four different vendors (AB Sciex, ThermoFisher Scientific, Waters and Agilent, for details see below). The selection of MRM transitions for each peptide was performed independently for each platform and aided by the use of Skyline software. Skyline was used to build spectral libraries from data dependent acquisitions and peptide search engine results using the BiblioSpec library builder (28) so that MRM transitions could be selected based on previously acquired discovery platform data (for AB Sciex and Waters instruments). Alternatively (for Thermo and Agilent instruments), SRM Refinement approaches (29) were used on a triple quadrupole MS to determine optimal MRM transitions for each peptide.

To select the five best transitions per peptide, spectral libraries for different instrument platform data formats were generated in Skyline, such as an "ABI spectral library" (obtained by data acquisition of the 6 ProteinMix-QC in data dependent IDA mode on a 4000 QTRAP), and a "Waters spectral library" (obtained by data acquisition of the 6 ProteinMix-QC in data dependent mode on a Waters QTOF Premier). For the QTRAP 5500 all transitions were selected to be below m/z 1,250 due to the instrument's upper mass limit of m/z 1250 in Q1 and Q3. For the ThermoFisher Scientific and Agilent platforms, optimal peptides and their corresponding transitions were determined by predicting all tryptic peptides from the 6 ProteinMix-QC and by monitoring all y-ions from y_3 to y_{n-1} on the triple quadrupole TSQ Quantum-Ultra and Agilent 6460 (29). The raw data were imported into Skyline and results were refined to select the 5 most abundant transitions for all detectable peptides. From this refined peptide list, 22 peptides that were readily detectable between all MS platforms were selected for the final MS method. One peptide was included with an additional charge state, resulting in a total of 115 transitions.

The final list of 22 bovine 6 Protein Mix signature peptides is listed in the SOP (see supplementary Methods). Finally, five different Skyline "Instrument Method templates" were generated, including all specific MRM transition information: (1) Study9S_ABI_MichromMix_template.sky (for 4000 QTRAP instruments), (2) Study9S_ABI5500_MichromMix_template.sky (for QTRAP 5500 instrument), (3) Study9S_Thermo_MichromMix_template.sky, (4) Study9S_Waters_MichromMix_template.sky, and v) Study9S_Agilent_MichromMix_template.sky. All Skyline Method templates were uploaded onto Panorama (panoramaweb.org, additional details below) and further details can be viewed in supplemental Table S1. Skyline Instrument Method templates were distributed to all sites and were used to export a transition list that could be directly imported into the mass spectrometer vendor method set-up (for detailed description see supplementary Methods).

Reversed Phase Nanoflow High Performance Liquid Chromatography (nanoHPLC)—Peptide mixtures were separated by on-line reversed phase nanoHPLC systems equipped with autosamplers: two NanoLC-1D Plus systems, seven NanoLC-2D systems and two NanoLC_Ultra systems, one NanoLC_Ultra 1D Plus and one NanoLC_Ultra 2D Plus from Eksigent Technologies (Dublin, CA); one Ultimate 3000 system from Dionex (Sunnyvale, CA); one nanoAquity system (Waters, Milford, MA); and two 1100 series systems (Agilent, Santa Clara, CA). PicoFrit® (New Objective, Woburn, MA) columns, 75 μm ID x 120 mm, 10 μm ID tip, were self-packed with ReproSil-Pur C18-AQ (3 μm particle size and 120 Å pore size) at each of the sites. Separations were performed at mobile phase flow rates of 300 nL/min using 0.1% (v/v) formic acid in water (mobile phase A) and 90% (v/v) acetonitrile with 0.1% (v/v) formic acid (mobile phase B). One microliter injections

of the peptide digestion mixture were separated using a gradient of 3–15% B in 3 min, 15–35% B in 34 min, 35–90% B in 3 min, and at 90% B for 4 min. Each site, except the Agilent ChipCube-LC instruments, acquired sample data with a PicoFrit/direct injection HPLC hardware configuration set-up. Details including plumbing configurations and autosampler injection routines are described in the SOP (see supplementary Methods and supplemental Table S2).

Mass Spectrometer Operating Parameters—General instrument operating parameters were kept as uniform as possible across the different MS platforms, except where noted in the individual sections below (additional details found in SOP, supplementary Methods and supplemental Table S2A). All MRM transitions are listed in supplemental Table S1. A total of 115 MRM transitions were monitored during an individual sample analysis with Q1 and Q3 set to unit resolution and the dwell time set to 10 msec for each transition.

ABI 4000 QTRAP/QTRAP 5500 Mass Spectrometer—Nine 4000 QTRAP hybrid triple quadrupole/linear ion trap mass spectrometers (AB Sciex, Foster City, CA) located at different sites were used to acquire MRM-MS data for this study. Ion spray voltages of 2200 ± 200 V, curtain gas 20, nebulizer gas (GS1) 5 ± 2, and interface heater temperature (IHT) 150 °C were used. Declustering potential (DP), and collision energy (CE) were calculated in Skyline using the following regression equations: $DP = 0.0729^* m/z + 31.117$ and $CE = 0.0431^* m/z + 4.7556$ (from personal communication with Dr. Jeffrey Whiteaker). Interscan delays were set to 5 msec and cycle times were 1.73 s. The QTRAP 5500 instrument was operated in a similar manner (for differences see supplemental Table S2A, and supplementary Methods). The following CE linear equations were used: $CE = (0.036^* m/z + 8.857)$ for z = 2 and $CE = (0.0544^* m/z - 2.41)$ for z = 3. For the QTRAP 5500 an interscan delay of 3 msec was used for all transitions, resulting in slightly shorter cycle times of 1.495 s.

ThermoFisher TSQ Quantum Ultra/TSQ Vantage Mass Spectrometer—One TSQ Quantum Ultra triple quadrupole mass spectrometer (ThermoFisher Scientific, San Jose, CA) was used to acquire MRM data for this study. Representative instrument operating parameters for the TSQ Quantum Ultra included spray voltage of 2300 V, capillary offset voltage of 35 V, and capillary temperature of 275 °C. Tube lens voltages used for all peptides, which were based on values generated during the automatic tuning and calibration process, were not individually optimized. The collision energy (CE) was calculated with Skyline for all individual precursors using the CE linear equation $CE = 0.034^* m/z + 3.314$ for doubly-charged peptides. Collision cell argon gas pressure was 1.5 mTorr, scan width of 0.002 m/z, (for more details see supplemental Table S2A) and the total cycle time was 1.48 s. The TSQ Vantage instrument was operated in a similar manner: spray voltage of 1300 V, and capillary temperature of 210 °C. Q1 and Q3 unit resolution of 0.7 FWHM, Q2 Argon gas pressure of 1.5 mTorr, scan width of 0.005 m/z, and total cycle time of 1.32 s. TSQ Vantage 2.0.0 and Xcalibur 2.0.7 SP1 operating software were used. On the TSQ Vantage, the Skimmer Offset and Tube Lens were replaced by the S-Lens (for typical S-Lens settings see Supplementary Methods).

Waters Xevo TQ Mass Spectrometer—One Waters Xevo TQ triple quadrupole mass spectrometer (Milford, MA) was used to acquire MRM data for this study. The collision energy (CE) for Waters instruments was calculated within Skyline using the linear equation $CE = 0.034^* m/z + 3.314$. In the tune page the capillary voltage was set to 2.2 kV, the collision gas to 4 ml/min, and the Cone Voltage to 37 V. MS Inter-scan time and the Inter-channel Delay were both set to 3.5 msec with a total cycle time of 1.55 s. MassLynx version 4.1 SCM725 was used.

Agilent ChipCube Mass Spectrometer—One Agilent 6410 and one 6460 triple quadrupole mass spectrometers (Santa Clara, CA) coupled to ChipCube interfaces were used to acquire MRM data. The CE for Agilent instruments was calculated in Skyline using the linear equation $CE = 0.036^* m/z - 4.8$. The ESI Voltage was 1750 V with a gas flow of 2.5 L/min, source temperature of 325 °C and Fragmentor Voltage of 125 V. The total cycle time was 1.485 s. MassHunter version 4.01 was used.

System Suitability LC-MRM-MS Data Acquisition—Mass spectrometric data were acquired for (1) a column conditioning study, comprised of ten subsequent HPLC-MS-MRM analyses of 50 fmol/μl of the 6ProteinMix-QC to assess how rapidly a newly packed reversed phase nanoHPLC column stabilizes and conditions, and (2) a "Phase I" Study comprised of ten subsequent HPLC-MS-MRM analyses of 50 fmol/μl of the 6ProteinMix-QC to assess system suitability metrics. Chromatographic carry-over effects were assessed in three blank injections which followed the Phase I runs. Chromatographic hardware setup for Phase I at all sites was required to be a PicoFrit/direct injection configuration with a 1 μl loop, except for the Agilent instruments, which used a ChipCube interface. Further study details, such as sample setup, sample and file naming nomenclature etc. are described in the accompanying SOP (supplementary Methods).

Concentration Curve for 10 Analyte Peptides: LC-MRM-MS Data Acquisition—To evaluate the relationship between system suitability performance and instrument sensitivity in a quantitative assay, the SSP was used in the context of a larger study in which 125 analyte peptides were monitored by a scheduled MRM-MS method at a number of the participating sites. A nine-point calibration curve was prepared in human K_2EDTA plasma that had been delipidated and depleted of the high abundant proteins using the Multiple Affinity Removal System (MARS-14, Agilent) by PPD Biomarker Discovery Sciences (Menlo Park, CA). The plasma was denatured, reduced and alkylated, digested with Lys-C and trypsin and desalted according to a standard operating procedure. The depleted plasma digest was spiked with a total of 125 synthetic ^{12}C/^{14}N peptides and the corresponding ^{13}C/^{15}N-isotopically labeled internal standard peptides. The standard curve consisted of the on-column concentration range from 1 amol/μl to 100 fmol/μl (with 8 concentration points prepared in quadruplicate) prepared by serial dilutions in a 0.5 μg/μl background of the depleted plasma digest and a constant concentration of 10 fmol/μl of all isotopically labeled peptides. All sample preparations were performed at a central location before distribution of samples to selected sites.

Of the 125 peptides spiked into the depleted plasma digest, 115 represented potentially cancer relevant peptide targets and were used in the context of a larger multiplexed study (30). The remaining 10 peptides, derived from seven target proteins (aprotinin: AGLCQTFVYGGCR; C-reactive protein: ESDTSYVSLK and GYSIFSYATK; horseradish peroxidase: SSDLVALSGGHTFGK; leptin: INDISHTQSV-SAK; myelin basic protein: HGFLPR and YLASASTMDHAR; myoglobin: LFTGHPETLEK; and prostate specific antigen: LSEPAELTDAVK and IVGGWECEK), had previously been well-characterized and assayed in Addona *et al.* (21), and were used to assess the limits of detection and quantification as related to SSP performance.

Timed-acquired multiplexed LC-MRM-MS was employed based on the retention times of each peptide being monitored. Using the isotopically labeled peptides, a series of six to seven unscheduled LC-MRM-MS analyses were performed with each run represented by a separate LC-MRM-MS method. Retention times of each peptide were determined and used to generate a single scheduled method using a 2-min retention time window. The SSP was applied before the analysis of the standard curve samples using an unscheduled LC-MRM-MS method to condition the column and assess instrument and LC-performance over five replicate injections. To track the instrument performance throughout the acquisition of the standard curve, the 6ProteinMix-QC was analyzed after every six to eight runs in the sample queue.

Data Analysis Platform - Skyline—Due to the demands of comparing results from 12 laboratories using instruments from 4 different vendors, we used the software application Skyline which allowed for a common data analysis platform to be used across sites that fostered sharing of data during acquisition as well as in final reports from each laboratory. Native instrument data files were imported into Skyline using the ProteoWizard Data Access Library. Skyline stores the data it needs from these files in a high-performance data cache file, which simplified sharing data between laboratories during the acquisition phase.

The MRM transitions for each peptide were integrated with Skyline to generate ion current peak areas representing each of the fragment ion signals. All data analysis and peak area integrations were initially performed at each of the sites; however, all data sets were further inspected at one central site for uniform data analysis and uniform determination of outlier peaks. Processed MRM data, such as peak areas, peak heights, chromatographic parameters, etc. were directly exported from Skyline as customized report tables for further statistical analyses (as .csv files). The Skyline report wizard was used to create a uniform report template with specific data fields of interest. These reports were then further processed with Excel, R and our own cross-site validation tools to generate further statistical results and graphics.

Supporting Data Analysis Software "Retention Time Viewer Program" (RT-Viewer)—The Retention Time Viewer program is one example of how exported Skyline reports were used for cross-site data analysis (available for download at http://gibsonproteomics.org/resources/rt-viewer). It was developed to easily visualize and troubleshoot chromatographic properties of identical samples analyzed between multiple laboratories or across all participating sites.

Statistical Methods—Data from MRM-MS experiments were preprocessed and integrated in Skyline, and exported as customized report tables in comma-separated value format (as .csv files). A uniform processing and export template was used for all MS instruments, resulting in output that was instrument independent. Only a limited set of key fields in Skyline were used in the statistical calculations to ensure that special features unique to Skyline were not used to keep the calculations free from unwarranted software dependences. To this end, the following fields were extracted from Skyline: sample name, peptide sequence, replicate name, retention time, area, full-width-half-maximum (FWHM) peak width, fragment ion, precursor charge and product m/z. These extracted fields were then processed using custom developed code written in the R statistical programming language (31).

The following calculations and analyses were performed with this data.

● *Determine Number of Transitions to Use for the Analysis*—A preliminary analysis was performed to evaluate whether the use of three, four, or five transitions resulted in a more robust measure of the various peptide parameters including retention time, peak width and peak area, specifically focusing on peak area. The coefficient of variation for peak area was very similar irrespective of how many transitions were used (supplemental Fig. S1). Based on these results, all five measured transitions have been retained for further analysis.

● *Choose a Representative Subset of Peptides*—The retention time, peak area and peak width of the 22 targeted peptides were monitored across different triple quadrupole instruments, sites and HLPC systems to choose a robust subset. The peptides in this subset were required to span the entire retention time range in a relatively uniform fashion, and have a reasonably good and consistent response. Based on the results displayed in supplemental Fig. S2 and Supplemental Table S3, the nine peptides highlighted in red in Supplemental Table S3 were selected and were used for subsequent analyses for evaluation of system suitability.

● *Combine Information From all Five Transitions for Each Peptide*—The retention time of the most intense transition was taken to represent the peptide retention time for that specific replicate. The maximum FWHM peak width over the 5 transitions was defined as the peptide FWHM peak width, and the peak area for the peptides was obtained by summing the peak areas for the 5 transitions.

● *Normalized Peak Area*—Not only do MS instruments from different vendors provide different measured responses for a given peptide in terms of peak area (or intensity), even identical instrumentation from the same vendor provides varying responses. To account for this arbitrary variation in peak area, and to render the measurements of the same peptide comparable from site to site, we calculated the normalized peak area for the best nine peptides. The peak area for each of the 9 peptides was represented as a fraction of the total peak area for the nine peptides, resulting in a normalized peak area ranging between 0 and 1 for all peptides and all sites (supplemental Fig. S3).

● *Chromatographic Resolution*—The ability of chromatography to separate peptides in time is represented by resolution. For a given pair of peptides, resolution was calculated as the difference in retention time for the two peptides divided by the sum of the FWHM peak widths for the peptides [Resolution = (RT2-RT1)/(FWHM1+FWHM2)]. In this study, peptides were sorted in order of increasing average retention time (over all the sites), and resolution was calculated between prespecified adjacent pairs of peptides (supplemental Fig. S4).

● *Peak Capacity*—Peak Capacity was determined for each sample acquisition replicate. First, the RT spread for the best 9 peptides was calculated subtracting the RT of the first eluting peptide (for all sites RT of peptide TAA) from the RT of the last eluting peptide (for all sites RT of peptide FFV) (supplemental Fig. S5). Peak Capacity was then calculated by dividing the determined retention time spread by average peak width at base line for best nine peptides [Peak Capacity = [RT(FFV) − RT(TAA)]/(average peak width at base line for best nine peptides) (32) (supplemental Table S4).

● *Peak Tailing*—Peak tailing was calculated as full peak width at baseline divided by (factor 2 multiplied with front half width of the peak), with front half width defined as "from the start of where the peak elutes at baseline to the apex." Peak tailing = AC/(2ˆAB), in which AC is the width of the peak at baseline and AB is the front half-width at baseline, (A: start eluting of peak at baseline, B: apex, and C end eluting of peak at baseline (supplemental Fig. S6, supplemental Table S5).

● *Retention Time Drift*—The retention time drift is determined for each peptide over the course of 10 replicate injections (RT drift = RT_{max} − RT_{min}) (33) (Fig. 2*A*, Supplemental Table S6).

● *Coefficient of Variation (CV)*—CV values, calculated as the ratio of the standard deviation to the mean, estimate variation for parameters of interest including retention time, peak area, normalized peak area, FWHM peak width, chromatographic resolution and peak capacity. In addition, CV reports were generated directly out of Skyline for the best nine peptides using a specific "CV results report" Skyline report template (supplemental Table S7).

● *Column Conditioning and Carryover*—Ten replicate injections of the 6ProteinMix-QC were made during the "column conditioning" phase of this experiment and data were analyzed to determine the optimal number of injections that resulted in achieving more stable retention times, peak areas, peak widths, etc. (data shown in supplemental Fig. S7). Although there did not seem to be agreement across sites on a precise number of injections that resulted in minimized variances for each of the measured metrics, it was clear that the first injection for most sites was the most variable, specifically in terms of retention time.

For carryover analysis the peak area from the last three replicate runs of the 6ProteinMix-QC, R08-R10 were compared with three post-sample blank injections as regular and logarithmic plots, respec-

tively (supplemental Fig. S8). For more specific carryover calculations, for each peptide the peak area of each blank, Bl-1, Bl-2, and Bl-3 was divided by the mean of the peak area of sample replicates R8-R10 (supplemental Table S8). Most peptides appeared to have a similar extent of carryover at a given site with a few exceptions. The absolute intensity of the carryover correlated to the absolute intensity of the peptides in the preceding injections, in that more intense peptides displayed a larger signal in the blanks. Each successive blank decreased the overall signal from each peptide with the peak areas of the third replicate equaling about 0.1% of the peak areas of the 50 fmol injection.

Graphical Methods—Data from MRM experiments exported from Skyline and analyzed using the custom R code described above are further explored using a variety of plots, also generated using the R statistical computing environment (http://www.R-project.org) (29), specifically the Lattice Graphics package (34). Comprehensive plots were made to analyze not only the overall statistics such as CV, but also understand more detailed variation of the LC and MS parameters and measurements from replicate to replicate, between sites and across peptides. See Fig. 3 and supplemental Figs. S4–S6.

Calculation of LOD for 10 Peptides From Nine-point Curve—Limits of detection for 10 peptides from a 9-point response curve in 0.5 μg/μl plasma were determined as previously described (21).
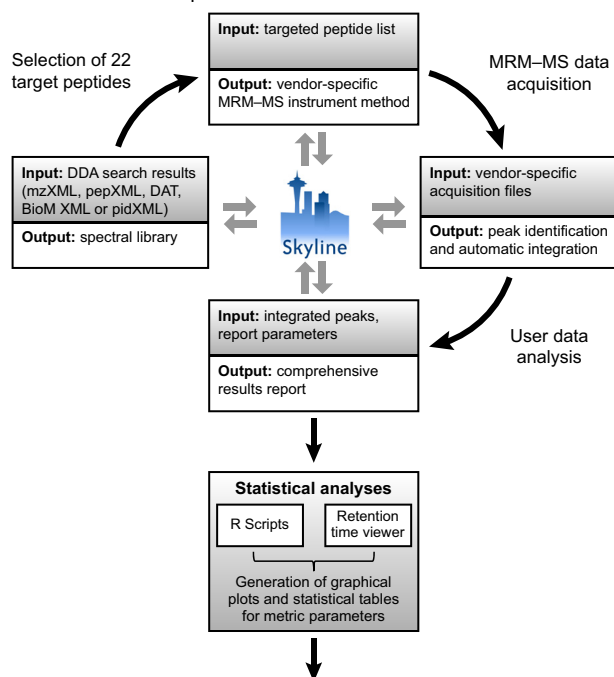
Assessment of Pass/Fail Metrics for System Suitability MRM-MS Assays—The following metrics were used for assessing LC-MRM-MS instrument performance: i.) RT Drift, ii.) Standard Deviation of RT, iii.) CV of FWHM, and iv.) CV of Peak Area. The median value for each peptide across 14 of the 15 instruments participating in this study were determined and reported along with the range of values. One of the 15 instruments was excluded due to HPLC-related issues that caused many values to be inflated compared with all other sites.

Data Availability on Panorama—All data have been uploaded to PanoramaWeb, an open-access web site, hosted by the MacCoss lab at the University of Washington, running on the open-source Panorama software (35) created as part of the LabKey Server (36). Data can be found at http://proteome.gs.washington.edu/software/panorama/abbatiello_study9s.html.

<div align="center">RESULTS</div>

Initial Analysis of the Protein Digest Standard: Selection of Peptides, Transitions, and Performance Metrics to Monitor— Our first task was to select the appropriate tryptic peptides to monitor from the six protein digest mixture and identify metrics for assessing reproducibility and performance integrity of the nanoflow HPLC-MRM-MS systems during routine analysis at 11 sites comprising 15 individual systems. Initial experiments conducted at a subset of the laboratories generated preliminary qualitative data from which 22 peptides (23 precursors, one peptide was monitored using two precursors) were selected for monitoring using 5 transitions each, facilitated through the use of Skyline (Fig. 1). The LC-MRM-MS method was designed to be analyzed in a retention time independent manner (not time-targeted), so that shifts in peptide elution would be observed. Each site acquired 10 replicate injections of 50 fmol (on column) of the SS sample to condition the freshly packed Picofrit columns, followed by 10 additional injections for metric assessment. Lastly, three injections of "null matrix" (mobile phase A, 0.1% formic acid) were acquired to determine the extent of carryover. Although 22 peptides were monitored in the final method, we illustrate
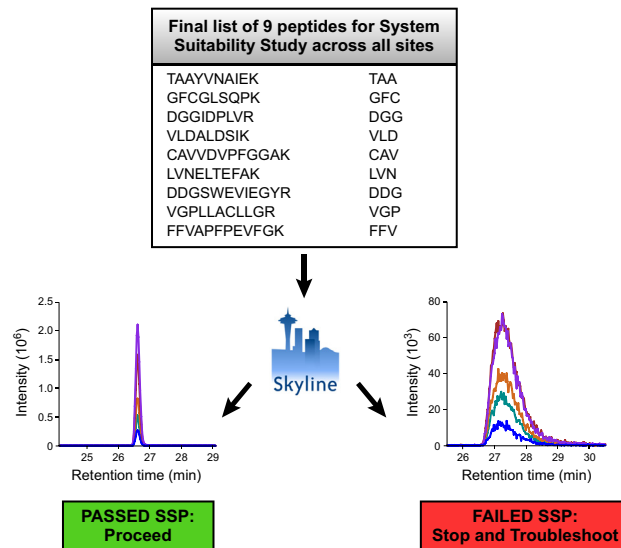

FIG. 1. **Workflow for the development, evaluation and use of a system suitability protocol (SSP) for system performance.** *A*, scheme of the method development process for picking peptides and transitions from a predefined sample, including generation of spectral libraries, generation of vendor-specific transition lists, data processing, and report generation in Skyline. Exported Skyline data reports were further analyzed to select the nine most appropriate peptides (based on criteria outlined in text) for use in the SSP for all sites. *B*, use of the SSP for evaluating system performance. The final nine peptides are acquired and the data analyzed in Skyline, from which the data metrics (peak area CV, normalized peak area CV, RT drift, RT standard deviation, and FWHM CV) can be directly exported in a report format to observe Pass/Fail status.

key results using nine peptides spanning a range of electrospray response, all of which were detected equally well across all instruments and that had retention times spanning the entire gradient (supplemental Figs. S1, S2, S9 and Supplemental Table S3). The workflow representing peptide selection and method development across the various instrument platforms is shown in Fig. 1*A*. The goal of developing a SSP for its use in multiple laboratories was to define and implement Pass/Fail criteria for evaluating whether or not a LC-MRM-MS instrumental configuration is suitable for performing a preconfigured assay. Accordingly, applying the SSP would prompt a site to proceed with the assay if the SSP criteria were passed or pause and troubleshoot the LC-MRM-MS system if it failed to meet all or part of the assessment (Fig. 1*B*). The specific parameters that we selected to monitor and develop pass/fail metrics and the rationale for their selection are described below.

Retention Time Variability—Retention time (RT) stability is one of the more important criteria for multiplexed MRM-MS-based assays. With modern triple quadrupole instruments it is quite straightforward to monitor ca. 20 different peptides in a 1 s scan cycle using three transitions/peptide (*i.e.* 3 MRM transitions $\times$ 20 peptides, with 10 msec dwell time/transition and 5 msec interscan delay for each transition = cycle time of 0.9 s). This data acquisition rate is compatible with typical nanoflow chromatographic peak widths and the need to sample multiple instances across a peak. However, it is both possible and desirable to configure assays to monitor much larger numbers of peptides in a single LC-MRM-MS analysis, for example 100 analyte peptides and 100 corresponding stable isotope-labeled internal peptides with three transitions monitored each, for a 200-plex assay. To achieve this multiplex level requires that the MS instrument is programmed to detect only selected peptides during predefined retention time windows. This is generically referred to as "scheduling" and has been implemented by all vendors of triple quadrupole instruments. Scheduled MRM can greatly increase the number of analytes that can be targeted and detected in a single LC-MRM-MS analysis. However, for it to work as desired the retention time stability is crucial and the target peptides must elute in the predefined RT windows, otherwise compromised signal or no signal will be detected. Therefore, assessing RT variability is an important parameter to define and monitor in the SSP.

The average RT and its variability over ten injections were calculated to determine if there were large shifts in RT of individual peptides over the course of the experiment, and if the variability in the RT was random or systematic. The results are visualized in Fig. 2. A peptide's RT drift was defined as the range of RTs from a given site over the course of 10 replicate injections of the SS standard (RT drift = $RT_{max} - RT_{min}$, for a given peptide at an individual site). In Fig. 2*A* the RT drift ranges for 9 of the monitored peptides is shown for each of the participating sites, and are grouped according to the

model of nanoHPLC system used. Ideally the observed RT drift at an individual site should be minimal; however, certain sites experienced RT drift of up to 1 min for several peptides during the 10 replicate injections.

The RT drift of each peptide was also evaluated within and across sites for the 10 replicate injections using "Retention Time Viewer" (RTV) software (http://gibsonproteomics.org/resources/rt-viewer). The results for one of the peptides, CAVVDVPFGGAK (CAV; from glutamate dehydrogenase) for each technical replicate at eight of the sites are shown in Fig. 2*B* (see also supplemental Fig. S10 and supplemental Table S6). RT drift at many of the sites was random (*e.g.* Sites 6, 8, 9, and 10), whereas more systematic drift toward longer retention was evident at other locations (*e.g.* Sites 12, 14, and 15), possibly indicating slightly decreasing flow rates or a systematic change in mobile phase composition over time for these systems. In addition, for assessment of RT reproducibility at a single site, Skyline provided a peptide replicate view, allowing the RTs of each transition for a given peptide (CAV in this example, Fig. 2*C*) to be compared across replicate injections. This view was critical to the instrument operators because it quickly indicated gross shifts in RT, which can be a symptom of column degradation or ambient temperature fluctuations, and allowed for rapid troubleshooting of the LC-MRM-MS system. The total spread in absolute RT between sites for this peptide was ~5 min and was explained by differences in capillary length used to plumb the nanoflow HPLC (supplemental Table S2*B*), and was also attributed to differences in the self-packed columns used at each site (target length, 12 cm $\pm$ 1 cm).

Raw Peak Area and Normalized Peak Area—To assess MS detector response reproducibility, raw peak areas were monitored for each peptide over replicate injections. Due to the variety of triple quadrupole (TQ) MS instruments employed in this study and the difficulty in comparing MS detector intensities across platforms, the raw peak area values for each peptide were normalized to the total peak area detected for the 9 peptides shown (Fig. 3 and supplemental Fig. S3*A* and S3*B*) in each separate LC-MRM-MS run, as described in the Methods section. The observed normalized peak areas for the majority of peptides varied over a narrow range (less than 3-fold) across sites and were largely independent of differences in instrument configurations across sites (Fig. 3*A*). In general, the CV of the normalized peak areas was <0.15 for all sites, except for a few later eluting peptides (DDG, VGP, FFV) that had somewhat higher CVs. The sites that observed higher normalized peak area CVs for later eluting peptides also had more variable raw peak areas for these peptides, in some cases related to electrospray stability (supplemental Fig. S11) as well as lower mean peak areas (Site 5, Fig. 3*A*). The median CV of normalized peak area for 14 of the instruments ranged from 0.04–0.08 across the 9 peptides, with the maximum CV of 0.36 for peptide FFV, which tended to have a large range of peak areas across instruments (Table I). Site 9 (purple symbol,

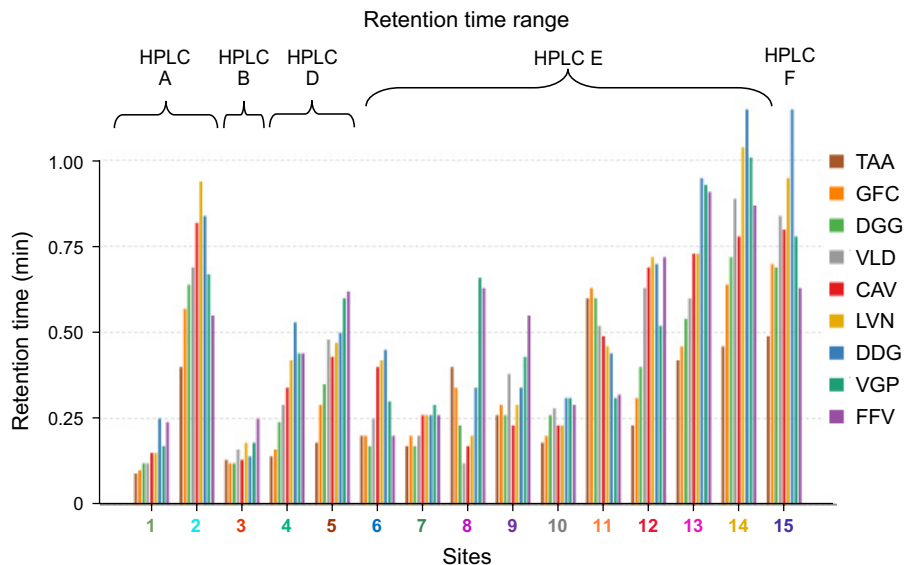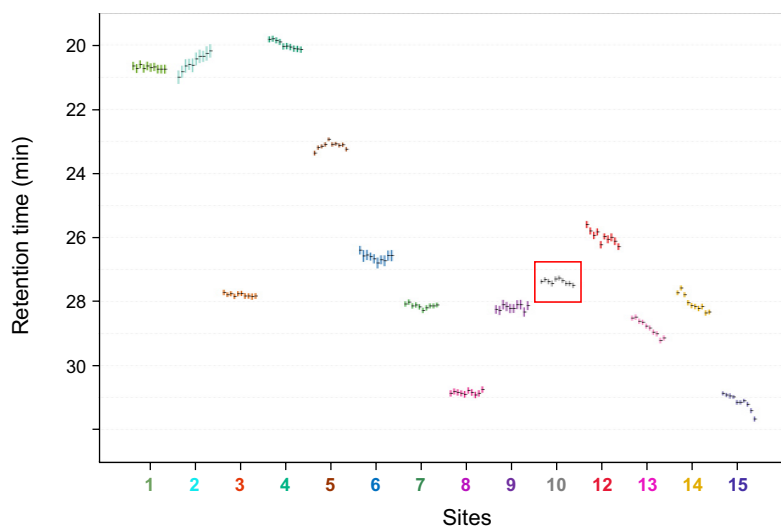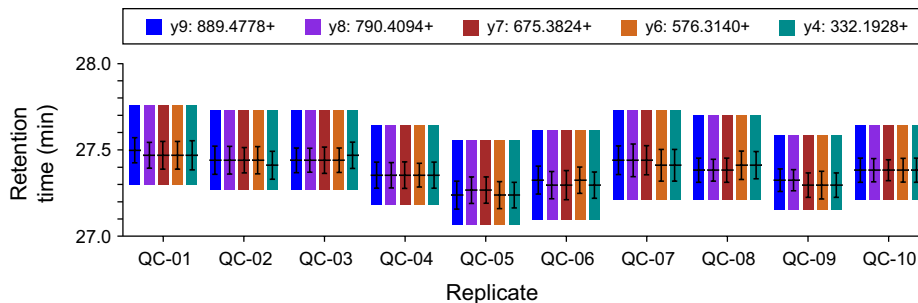## A Retention time drift and range at all sites for 10 replicate injections

B RT variability for peptide CAV within and across 14 selected sites

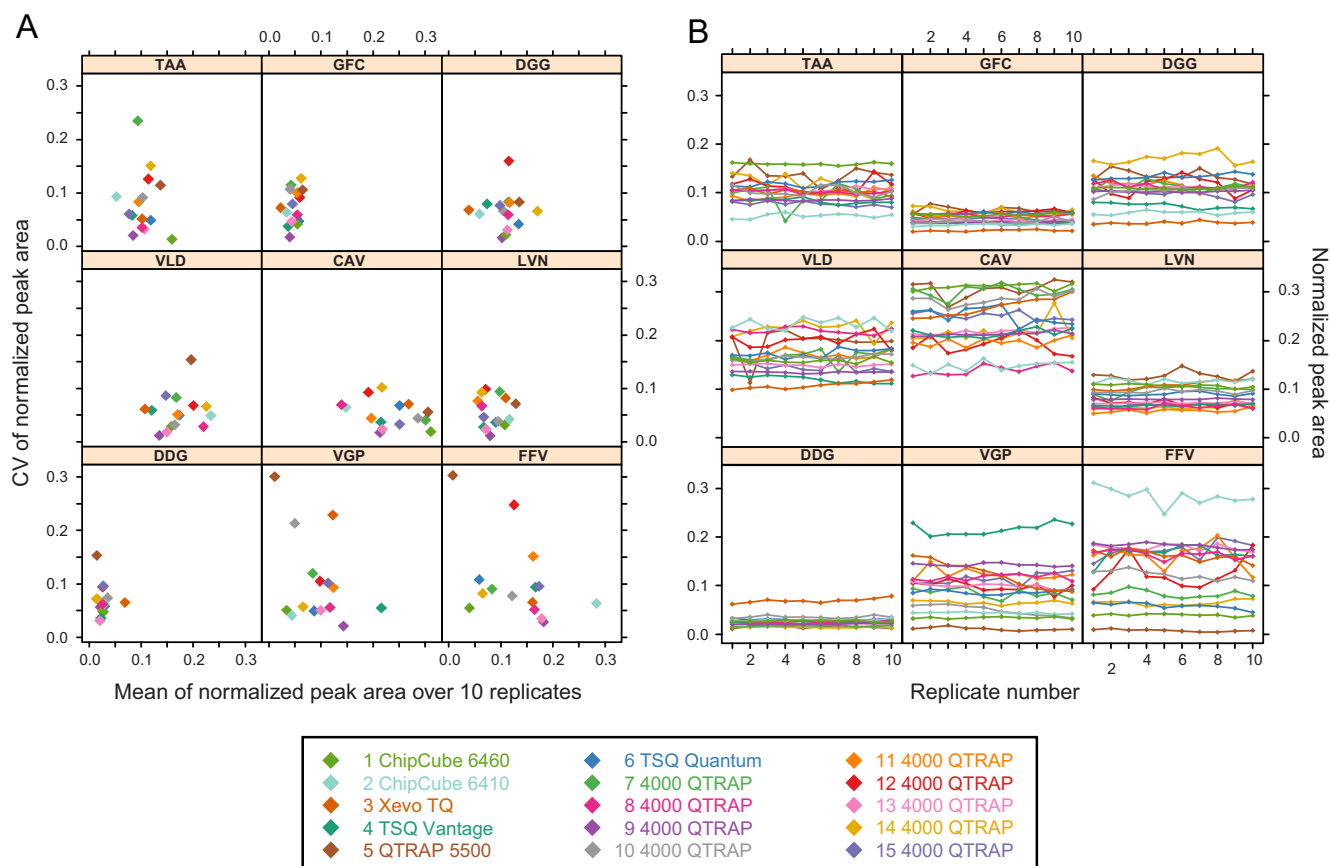## C RT variability for peptide CAV on inject-to-inject basis at single site

FIG. 2. **Evaluation of retention time stability from the SSP.** *A*, RT drift observed for the 9 peptides on each of the 15 instrument platforms, grouped by LC manufacturer. Peptides are sorted in order of increasing retention time drift at each site. *B*, RT Viewer display of the injection-to-injection RT drift for peptide CAV (CAVVDVPFGGAK from glutamate dehydrogenase) at 14 selected sites. Sites 12, 14, and 15 show the largest drift in RT whereas sites 3, 8 and 10 show the smallest. Site 11 is not displayed as results were off scale due to problems with flow calibration. *C*, Skyline display of injection-to-injection variability in RT for one peptide at Site 10.

FIG. 3. **Comparison of normalized peak area in the system suitability sample across 15 sites.** *A*, CV of normalized peak area *versus* the mean of the normalized peak area for each site for the nine final peptides (each box represents one peptide). *B*, injection-to-injection variability of normalized peak area. The Site Legend identifies sites by color and the type of MS instrument used at each site.

TABLE I

The median and range of SS metrics is shown for the 9 SS peptides for 14 instruments. Median and range values were determined from 10 replicate injections at each site

	TAA	GFC	DGG	VLD	CAV	LVN	DDG	VGP	FFV
RT drift (min)	0.25	0.30	0.26	0.34	0.37	0.42	0.45	0.44	0.50
	0.09–0.60	0.10–0.72	0.12–0.72	0.12–0.89	0.15–0.89	0.15–1.04	0.14–1.15	0.17–1.01	0.20–0.91
Std Dev of RT (min)	0.08	0.10	0.09	0.11	0.12	0.12	0.13	0.14	0.14
	0.03–0.19	0.03–0.22	0.03–0.23	0.04–0.28	0.05–0.26	0.05–0.36	0.04–0.43	0.05–0.36	0.06–0.30
CV of FWHM	0.07	0.06	0.08	0.06	0.05	0.07	0.07	0.06	0.06
	0–0.19	0–0.14	0.02–0.12	0.03–0.19	0.03–0.12	0–0.24	0.02–0.16	0.02–0.22	0.03–0.20
CV of peak area	0.10	0.08	0.07	0.07	0.09	0.08	0.08	0.11	0.10
	0.04–0.27	0.04–0.28	0.04–0.28	0.03–0.25	0.03–0.26	0.04–0.25	0.03–0.23	0.04–0.26	0.04–0.36

Fig. 3*A*), consistently had the lowest normalized peak area CVs for each of the peptides monitored despite using an older generation system, the AB Sciex 4000 QTRAP. Although certain peptides exhibited a wider and higher range of CVs appearing to correlate with greater variation in their mean normalized peaks areas across sites (*e.g.* compare peptides LVN with peptide FFV, and early eluting peptides TAA and DGG with peptides eluting in the middle of the gradient), most of the variation observed in normalized peak area and peak width did not appear to be site specific. We investigated this phenomenon further and found that the more hydrophobic peptides YLGYLEQLLR, VGPLLACLLGR and FFVAPFPEVFGK (YLG, VGP, and FFV, respectively) exhibited decreasing raw peak area with increasing storage time in the autosampler (4 °C at 50 fmol/μl in 0.1% formic acid). Given that these peptides were at low concentration in a relatively noncomplex sample matrix it is likely that they adsorbed to either the sample vial walls or, in one case, to the coated metal autosampler needle during sample aspiration. Monitoring the intensities of these peptides over

subsequent injections indicated an issue through decreasing or nonexistent signal.

The normalized peak area for replicate injections, plotted individually for each peptide by site is shown in Fig. 3*B*. This view of the data facilitates observation of site-specific trends such as systematic variation in normalized peak area as well as anomalies occurring in any single replicate. For example, peptide VLDALDSIK (VLD; carbonic anhydrase) decreased ~2-fold in injection number 2 at Site 5 relative to all other measurements for this peptide at that site (Fig. 3*B*) resulting in a higher CV for this peptide. On further analysis of the data, the observed decrease in normalized peak area for this injection was due to unstable electrospray whereas the peptide was eluting, which was also the cause for the sites that had CVs higher than 0.15 for peptides TAA and DGG. Additional examples of performance deviation are detailed below. Although normalized peak area was employed in this study for cross-site comparison, we recommend simply using raw peak area for single LC-MRM-MS configuration evaluation since this metric is a quantitative output of Skyline software, which allows easy assessment of system performance (Fig. 4).

Peak Width (FWHM)—The peak width at FWHM metric allows comparison of a fundamental aspect of chromatography over time on one instrument and between individual sites. Overall, peptide peak widths within a site were highly reproducible with CVs generally ≤0.2 and often ≤0.15 in which the FWHM values for individual peptides varied between sites from ~0.1 min to 0.3 min. Later eluting peptides exhibited somewhat larger variation in peak width across sites suggesting a gradient-dependent effect. The individual FWHM values for each replicate are plotted in supplemental Fig. S12, where the increasing range of FWHM values between sites for later eluting peptides is more apparent. The four sites with different HPLCs [supplemental Fig. S12, Sites 3 (Waters nanoAcquity), 4 (Eksigent Ultra 1D), 5 (Eksigent Ultra 2D), and 15 (Dionex U-3000)] did not stand out from the laboratories using Eksigent 1D Plus and 2D nanoHPLCs. This indicates that the major sources of variability observed were more likely due to variations in how the LCs were plumbed (supplemental Table S2*B*), and dissimilarities in the self-packed columns rather than major differences in performance between HPLCs from different vendors or separate models of HPLC from the same vendor.

Chromatographic Resolution, Peak Capacity, and Peak Tailing—Chromatographic resolution (R_s) was calculated for adjacent eluting peptides to observe the variation in this parameter during column conditioning and over the 10 subsequent replicate injections (supplemental Fig. S4). Although some sites exhibited more consistent R_s values over the 10 injections, the trends observed between the sites were very similar in that R_s values did not change substantially with peptides maintaining elution order and spacing. Peaks exhibiting the largest variation in R_s also happened to have the lowest

calculated R_s values at each site indicating small changes in closely eluting peptides. Eleven of the 15 sites maintained CVs less than 0.15 for almost all peaks. The remaining 2 sites with CVs above 0.15 (Sites 3 and 6) also had slightly more variable FWHM values, which would affect R_s.

Peak capacity (n_c) was evaluated to determine if the retention time span over which the peptides eluted changed during the replicate injections, as well as to compare across sites (Supplemental Fig. S5, supplemental Table S4). Although the peak capacity values ranged from 40 to 75 between the various sites, the reproducibility within each site was excellent with CVs < 0.075. This divergence in peak capacity values across the sites may be explained by differences in the capillary tubing used to plumb the nanoHPLC, the dissimilar column length, or the efficiency of the nanoHPLC pumps in delivering the proper gradient conditions.

Chromatographic peak shape was evaluated by calculating the Peak Tailing Factor (PTF) in the Retention Time Viewer Program. All sites observed PTF very close to 1 for every peptide monitored, which indicated overall suitable peak shape (supplemental Fig. S6, supplemental Table S5).

Software Tools for SSP Development and Evaluation—A critical requirement of implementing a LC-MRM-MS SSP is the capability to rapidly determine if the instrument platform is performing within specifications. We used several features in Skyline to view LC-MRM-MS data from replicate injections that enabled quick assessment of reproducibility (Fig. 4 and supplemental Fig. S13). First, peptides were sorted in order of their RTs and the mean peak areas for the 10 replicates were plotted with error bars representing one standard deviation (Fig. 4*A*). Second, changing the view to plot peak area CV on the *y* axis allowed for a quick evaluation of the overall reproducibility of the injections for all peptides, in which the goal was for the CV to be as low as possible and less than 0.2 (Fig. 4*B*). Third, a replicate view allowed for inspection of the total raw peak area (summed from all 5 transitions monitored per peptide, in this case, TAA, Fig. 4*C*) for each replicate injection. In cases in which unusually large CVs were observed it was possible to quickly determine which injection was the outlier (Fig. 5). Using these Skyline tools allowed operators to quickly assess the quality of data over replicate injections of the SS standard and determine whether the system was performing normally. If irreproducibility was observed, then immediate troubleshooting of the system was performed to identify and rectify the problem. Finally, all SS sample runs generated during a given timeframe were saved into a Skyline file to track system performance over time.

The RTV program provided additional tools to quickly assess chromatographic behavior by comparing across multiple sites or instruments (supplemental Fig. S14). RTV is an interactive, graphical Java program that imports a set of Skyline reports, one per site. The program visualizes retention times of peptides across the sites for either all 10 HPLC replicate runs or a subset of them. Features can be visual-
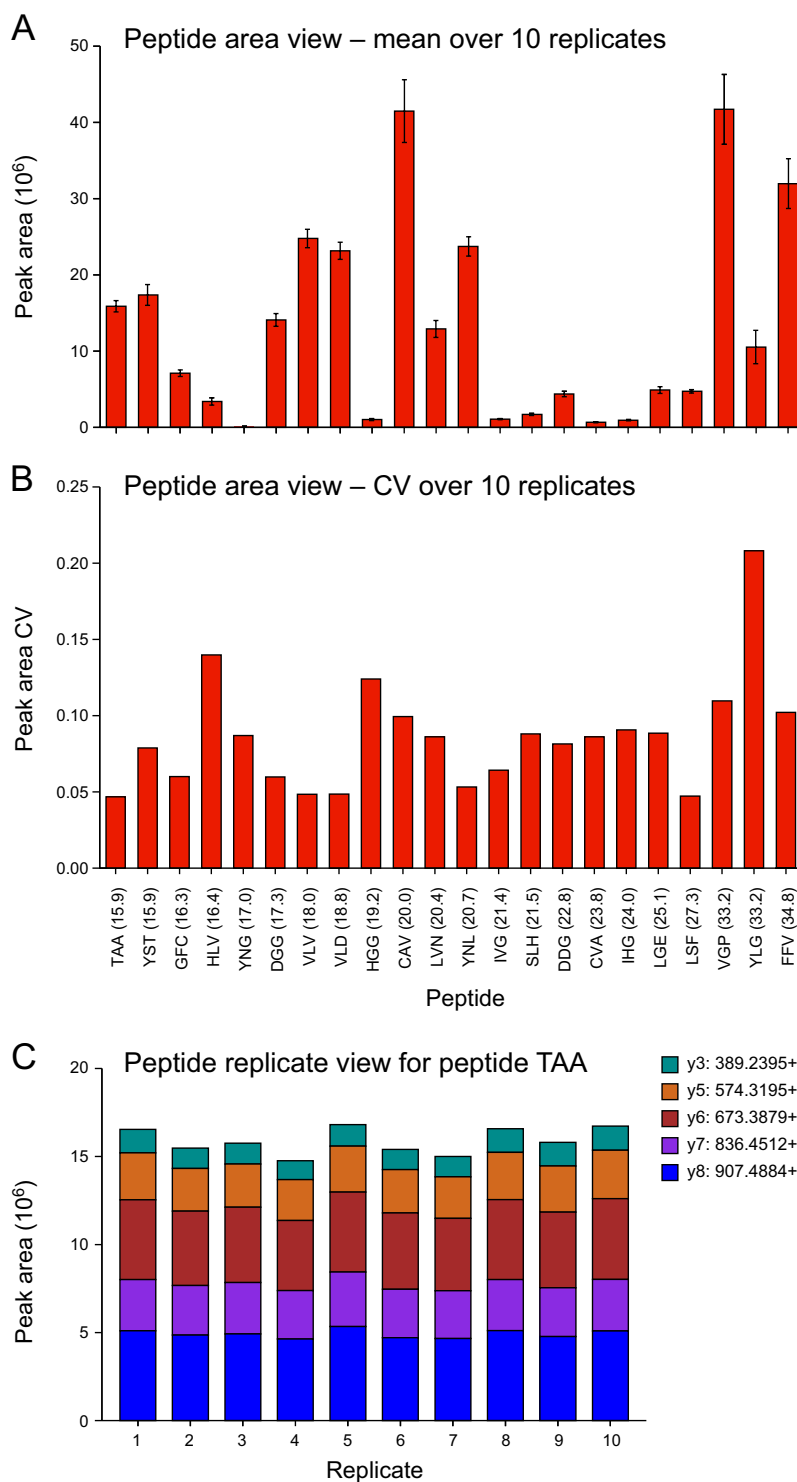
FIG. 4. **Skyline tools developed for rapid assessment of system suitability performance at a single site.** Panel *A*: average peak area across 10 injections for each of the original 22 peptides monitored in the SSP (error bars represent 1 standard deviation). Peptides are sorted in order of retention time with retention time shown in the *x* axis. Panel *B*: CV of the peak areas for each peptide across 10 injections. Panel C: view of results for the raw peak area observed for peptide TAA for 10 replicate injections; each color representing the signal from the individual transitions.

ized for all sites at the same time or for a subgroup of selected sites. For any single peptide, chromatographic information similar to what Skyline displays for a single site in the Retention Time Replicate chart can now also be easily compared among many sites using the RTV program including features, such as peptide retention time, start and stop elution times, retention time stability or drift over sample replicates, peak width (FWHM), peak tailing, etc. Average values of all chromatographic properties over 10 replicate HPLC runs were automatically calculated and displayed for each CPTAC site, and all chromatographic data can be exported as tabulated text.
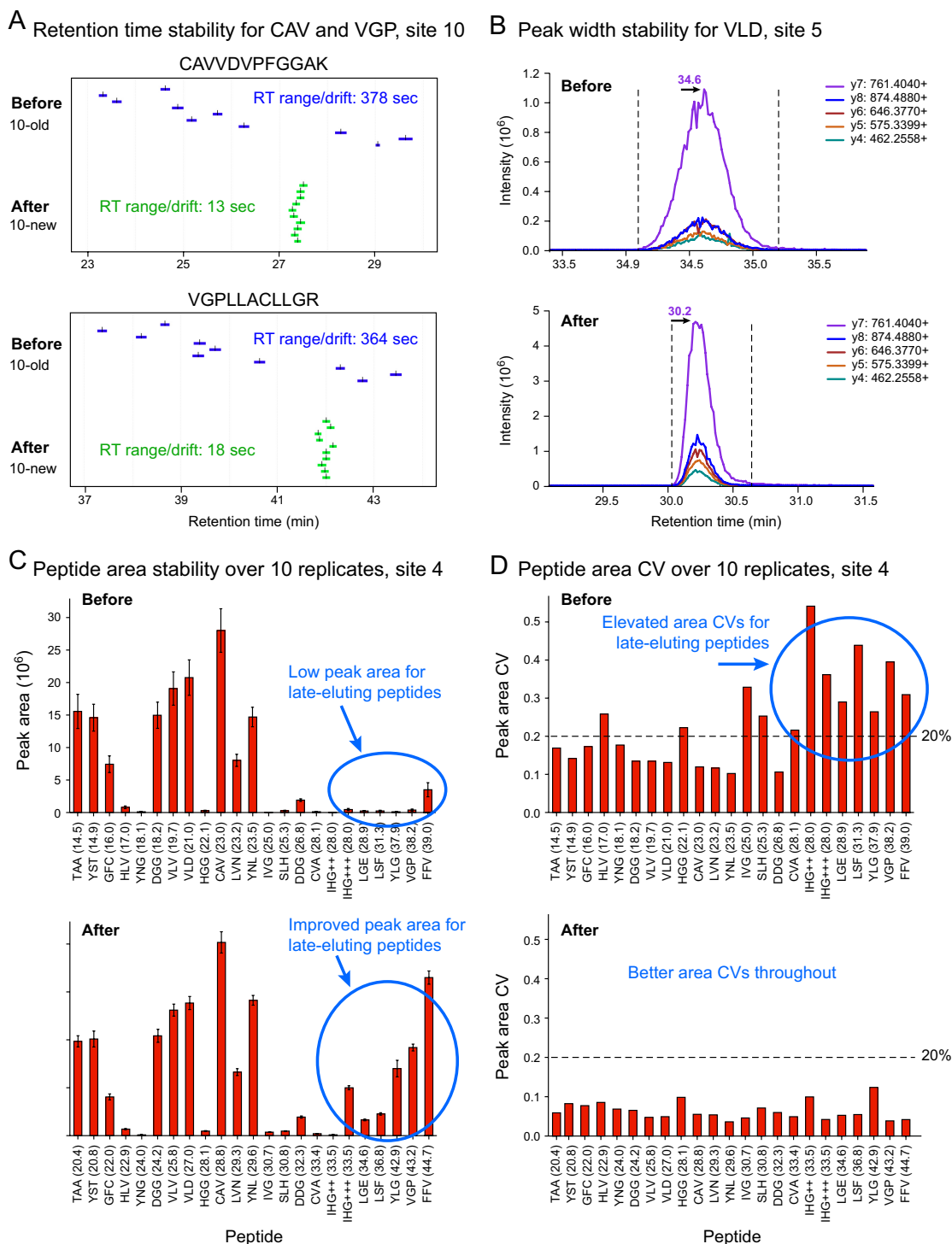
FIG. 5. **Before and After Plots of SSP Data from Individual Sites with System Performance Problems.** Panel *A*: data from two SSP peptides (CAVVDVPFGGAK and VGPLLACLLGR) that exhibited large retention time shifts (in blue), before room temperature stabilization at Site 10. When the sample was re-run, the "after" retention time shifts, shown in green, are more representative of the retention time variability observed at other sites. Panel *B*: unusually wide FWHM for peptide VLDALDSIK at Site 5, indicating a problem with the HPLC plumbing. After the connections were fixed, Site 5 obtained narrower FWHM more typical of the other sites. Panel *C*: before and after view of the Peak Area view in Skyline for Site 4. The last 5 peptides of the SS sample showed very low signal as compared with the other sites involved in the study. When Site 4 re-calibrated the flow meter on their organic solvent pump, the peak areas for the later eluting peptides increased and were observed to be more similar to the other sites. Panel *D*: same data from Site 4 as in Panel *C* but using the Peak Area CV view in Skyline. The later eluting peptides all have elevated CV values for Peak Area in the "before" case. When the system was fixed, all CVs dropped, but the effect was largest on the later eluting peptides.

Use of the SSP to Recognize Out of Bounds Variability in System Performance—During the course of this investigation, there were incidents of hardware issues and/or environmental conditions that resulted in dramatically shifting retention times, irreproducible peak areas, and peak tailing that caused excessively large CVs ($\gg$0.2). The SSP was instrumental in the early detection of these complications. Fig. 5 illustrates three "before and after" examples of problems encountered at various sites and their remedies. In the first example, during initial SSP test acquisitions, site 10 experienced wildly shifting retention times for all peptides during the ten replicate injections of the SS standard. For example, peptides CAVVDVPF-GGAK and VGPLLACLLGR demonstrated RT ranges of 378 s (6.3 min) and 364 s (6.1 min), respectively (Fig. 5*A*). This site identified the potential source of the RT fluctuation to be systematic changes in the ambient temperature of the laboratory over the course of a 24 h period. Once the temperature was stabilized, the site re-ran the sample and obtained a very stable retention time range of 13 and 18 s for the same peptides (0.2 and 0.3 min, Fig. 5*A*). Example 2 illustrates peak tailing, which is another indication of suboptimal chromatography. At one point during the course of this study, site 7 began to experience peak widths of ~0.5 min FWHM, whereas this site previously had peak widths of 0.25 min FWHM or less as did all other sites in the study (Fig. 5*B*). On closer inspection of their nanoHPLC hardware, adjustment of the liquid junctions in the autosampler valve to reduce dead volume resulted in the peak widths returning to <0.25 min FWHM. Finally, example 3 illustrates a problem with the later eluting peptides at Site 4 that the SSP uncovered (Figs. 5*C* and 5*D*). The last six peptides detected in the gradient were exhibiting lower than average peak areas, with correspondingly high CVs for those peptides (Fig. 5*D*, top panel). On measuring the flow rates from the two nanoflow pumps delivering mobile phases it was found that mobile phase B was being delivered at a flow rate lower than expected, resulting in inconsistent peptide detection during the latter part of the gradient. After flow rate calibration, the peptides were more reproducibly detected and their peak area CVs returned to ≤0.1 (Fig. 5*D*, bottom panel). Additional similar Skyline plots are shown in supplemental Fig. S15. In each case shown, report templates were developed in Skyline to output the CVs of the following metrics: retention time, FWHM, total area and normalized area (shown in supplemental Table S9) in a tabular format.

Acceptable Range of Variability in SSP Metrics—The two goals of our study were to develop a SSP and data analysis toolkit for routine use on individual LC-MRM-MS platforms, and to attempt to define the acceptable range of variation in each of the MS and LC-based parameters selected as performance metrics. The data described to this point were acquired on LC-MRM-MS systems that were known to be functioning properly and that produced data with variability of less than 20% for the various parameters measured. Therefore we believe that the observed ranges in variation of each of the parameters monitored serves to define what well performing nanoLC-MRM-MS platforms should be able to achieve during routine peptide-based assays. Table I presents the interlaboratory median and range for all metrics monitored during the SSP study for each of the 9 final peptides. Although the magnitude of each parameter (*e.g.* peak area, FWHM, retention time) may differ somewhat from the values shown depending on the actual flow rate, column packing material, LC plumbing and MS instrument used, the CVs of each of these metrics should remain low (<0.2 for single peptides and < 0.15 for median of all peptides measured) assuming the system is in good operating condition. In fact, when assessing 14 "normally functioning" systems, the median CVs for peak area and FWHM did not exceed 0.11 for any peptide. Inconsistent performance at certain sites is evident in the maximum values for FWHM and peak area shown in Table I.

Effect of System Variability on Limit of Detection in Quantitative Assays—The SSP was incorporated into a quantitative study designed to target 10 analyte peptides and their corresponding $^{13}C/^{15}N$ analogs in depleted and digested plasma (21) to determine if there was a relationship between any of the SS metrics and limits of detection (LOD). As described above SS replicates were acquired at 8 sites, followed by generating a response curve for the 10 analyte peptides spiked into depleted human plasma at concentration levels ranging from 0.001 to 100 fmol/μl. Additional LC-MRM-MS analyses of the SS standard were interspersed between (every 8–12 injections) and following acquisition of the response curve data. The 4 main metrics for the SSP (peak area CV, FWHM CV, retention time standard deviation and retention time drift) were calculated for the 9 SS peptides and the median generated for each site. These values were plotted against the median LOD for the 10 target peptides monitored in the response curve, during which the SS runs were interspersed. We observed that, although the correlations determined across 8 individual sites were weak (supplemental Fig. S16), the metric with the strongest correlation was peak area CV. This is further illustrated when comparing the performance of several instrument platforms for the same quantitative assay (Fig. 6). As the system suitability metrics, specifically peak area, begin to deteriorate during the assay, it adversely affects determination of LOD, resulting in higher LODs than what could otherwise be achieved. In all cases, each site passed the preassay SS metrics (at or below median values shown in Table I). However, Site 9 displayed a dramatic increase in peak area CVs in subsequent SS samples throughout the quantitative assay, accompanied by a greatly elevated LOD as compared with the other 4 sites. Further investigation of Site 9's data revealed a gradual decrease in peak area for each injection, causing irreproducibility in peak area, increased CV values and poorer overall sensitivity.

Longitudinal Assessment of SSP—An important benefit of using a SSP is the tracking of system performance over
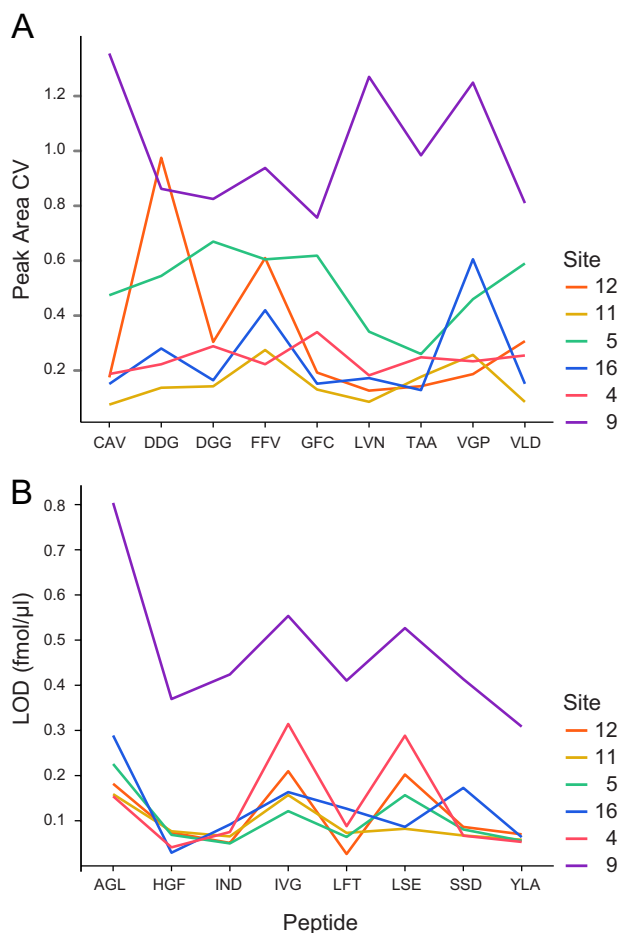
FIG. 6. **Correlation of Peak Area CV of SS peptides and LOD of peptides from a quantitative SID-MRM-MS assay.** Six sites acquired a 9-point response curve to determine LOD of 10 target peptides. The SSP was acquired at the beginning of the study and periodically throughout the response curve. Panel *A*: peak area CV of the SS peptides for all injections of the SS sample throughout the study for each site. Panel *B*: plot of the calculated LOD of 8 out of the 10 peptides from the quantitative study. The two peptides from CRP were not included as endogenous amounts were detected in the blank plasma matrix and LODs could not be accurately determined.

extended periods of time, before and after changes to the hardware (either LC or MS), acquisition software, preventative maintenance, or routine maintenance, such as cleaning of the MS source. Software such as Skyline facilitates the capture and storage of SSP data and provides longitudinal snapshots of system performance as a function of time, use, and maintenance. An example of SSP performance over the course of a 10–12 day experiment is shown supplemental Fig. S17, in which the SS sample was injected at least every 8–10 h at each of two laboratories. The SSP is sensitive enough to detect deteriorating system performance and shows improvement after corrective actions (cleaning the MS source) were taken at one site (Site 5), whereas system performance was relatively stable at the other (Site 11). Associated SS metrics are provided in supplemental Table S10.

DISCUSSION

Before initiation of a quantitative LC-SID-MRM-MS peptide based study, it is important to confirm that the nanoHPLC and mass spectrometer are optimized to achieve the performance level expected to detect and quanify the analytes of interest. Although individual labs may have their own set of standards and methods through which this is accomplished, cross-laboratory comparison of LC-SID-MRM-MS assays is difficult unless the same sample and methods are used. In addition, until a lab has become accustomed to running a SSP, it is difficult to know what levels of variability to expect from a particular nanoHPLC-MRM-MS system. Here we described the development, assessment and application of a SSP for nanoHPLC-MRM-MS peptide based studies. The SSP consists of a defined SS standard sample, for which well characterized performance metrics under defined operating conditions were established, and the software tools necessary to facilitate data extraction for swift assessment of instrument platform performance. The SSP developed allows for rapid and thorough analysis of raw data derived at an individual site and for cross-site comparison between different vendor platforms. Based on these analyses, we recommend target CVs for raw peak area of 0.15, with 0.25 defined as the maximum allowable peak area CV. Recommended CVs for FWHM can be lower, at a target of 0.10 and a maximum of 0.25. Target RT drift values should be <0.5 min with a maximum of 1 min, and RT standard deviation target of 0.15 min and maximum of 0.5 min. These recommendations are based on the performance of a wide range of instruments and laboratories and serve to define what well performing nanoHPLC-MRM-MS platforms should be able to achieve before commencing a peptide-based quantitative study. Inflation of these values, specifically peak area CV was observed in the context of a quantitative assay ("throughout assay" SSP, Fig. 6), over the span of 100's of injections, and is not uncommon. However, keeping these metrics at or below the maximum recommended value will improve the chances of maintaining the best assay sensitivity possible. Implementation of a SSP at key points in an assay (before method development, periodically during method development and quantitative assay acquisition, and after any HPLC or MS hardware has been replaced, cleaned or adjusted) enables monitoring of specific metrics that may influence the day to day and month to month performance reproducibility of the LC-MRM-MS system, and ultimately impacts the overall precision and sensitivity of the method.

The SSP we developed uses a nanoLC-MRM-MS method that monitors 9 tryptic peptides from a commercially available predigested, equimolar mixture of 6 bovine proteins. This standard is inexpensive (<$0.10/injection), readily available and easily prepared with minimal sample handling. This method has been transferred to 4 different vendor platforms in 11 different laboratories across a total of 15 individual instrument configurations. All data were analyzed through Skyline

and Retention Time Viewer, both of which are publicly available and vendor neutral software packages. Transition lists for data acquisition as well as report templates are available on-line for use with Skyline. The methods described for peptide and fragment ion selection could readily be applied to any mixture of peptides, with the reporting templates adjusted accordingly. In selection of a SS sample, we recommend choosing something similar to the analytes of interest and stable for at least 1 week in the autosampler (4 °C). During the course of this study, several of the hydrophobic peptides monitored were found to decrease in intensity after 1 week at 4 °C and when stored at low (≤25 fmol/uL) concentrations in aqueous conditions (3% acetonitrile/0.1% formic acid, data not shown). Ideally, a SS sample would be as stable as the analytes targeted in a quantitative assay, yet sensitive enough to changes in the LC-MS system to indicate potential issues that would influence sensitivity or reproducibility. From a practical standpoint, it is convenient if the sample requires minimal preparation before use and is made reproducibly so batch-to-batch variability is much smaller than the variability in instrument performance.

Our study has defined a number of factors that can adversely affect both the precision and accuracy of quantification in LC-SID-MRM-MS assays. During the course of this study, it became apparent that the hardware configuration of the nanoHPLC system and its fluidic connections had the largest impact on overall system performance and reproducibility. Therefore, many of the metrics chosen for evaluation are chromatography-based measures including RT, peak width (FWHM), chromatographic resolution, column peak capacity, and peak tailing. RT reproducibility is a key factor when developing an assay to target many (>20) analytes and their corresponding internal standards. For highly multiplexed scheduled MRM-MS, the RT of each peak must be accompanied by a scheduling RT window, allowing for small shifts in the RT over the course of many acquisitions. If there are large shifts in the RT of peptides then the RT window in the scheduling software must be increased to accommodate the large shifts in chromatography. Opening up the RT window reduces the number of concurrent transitions that may be acquired within a reasonable scan cycle time and using appropriate dwell times (e.g. ~10 msec). Furthermore, if the RT window is set too narrow, peptide peaks may drift outside of the acquisition window and no data or partial data will be collected. It is for these reasons that RT stability is a key aspect in this SSP. Although the data shown do not represent the RT shifts that might be observed over hundreds of injections or large sample loads, it does provide a baseline of performance that can be used to compare subsequent RT drift.

Variability in peak area and width can also produce imprecise and inaccurate quantitative results in a LC-SID-MRM-MS assay (20, 21). Stable isotope-labeled internal standard peptides help to reduce the variable effects of chromatography and ionization. However, they cannot compensate for low signal-to-noise resulting from an autosampler malfunction or from erratic signal due to electrospray instability or signal loss due to a dirty ion source - factors that all adversely affect achievable LOD and LOQ values. Such events often occur after multiple (tens to hundreds) injections on the same nanocolumn, causing residue build-up on the nanospray tip or ion optics in the source, column degradation, or air in the autosampler. Identification of recurring events such as these becomes more feasible when actively monitoring for effects on the chromatography or MS signal using an SSP. The most influential factor on peak area and width in this study was quality of the fluidic connections associated with the nanoHPLC. Introduction of even a small void in an autosampler valve or fitting may introduce 1–2 μl of dead volume, which translates to several minutes of delay when flowing at 300 nL/min. In addition, dead volumes also introduce new sources of mixing that can cause chromatographic tailing and irreproducible peak shapes (illustrated in Fig. 5B). We demonstrated that 15 individual LC-MRM-MS instrument platforms were able to achieve very similar FWHM and normalized peak areas for the same sample. These results provide a baseline for the metrics evaluated that any lab can target when configuring their own hardware and assessing its performance. Additional metrics such as chromatographic resolution and peak capacity are useful ways to monitor both environmental conditions (ambient temperature) and column life during the course of an assay that may cause subtle deleterious performance. Both parameters had very small variances over 10 replicate injections for the 15 instruments involved and together with RT drift or FWHM variation, can provide insight to column degradation.

Interestingly, our study revealed that the SSP performance metrics obtained on previous generation triple quadrupole mass spectrometers (*i.e.* 4000 QTRAP) were comparable to newer instruments (*i.e.* QTRAP 5500). These results point to an obvious conclusion: best performance is not guaranteed by using the latest generation of instruments, but rather, by optimizing each component of the platform. The weakest component will often be the limiting factor in terms of overall data quality.

Assessment and tracking of longitudinal system performance is another benefit of employing a SSP for LC-MRM-MS instrumentation. The use of software with simple visual and tabular data outputs, such as Skyline and Retention Time Viewer, allow easy access to the history of system performance, and can even provide a tool for on-the-fly evaluation of changes to the LC or MS hardware or acquisition parameters. Further, data can be uploaded to PanoramaWeb for easy access to and storage of system performance.

In conclusion, assessment of the performance and reproducibility of nanoHPLCs and triple quadrupole platforms using a SSP before the start of a study, and periodically throughout the course of a study will help ensure that the quality of quantitative data generated is reliable and reproducible. Im-

plementation of a SSP as a standard procedure as part of targeted quantitative peptide-based MRM-MS assays enables potential problems to be identified early and addressed. We believe that a SSP should be incorporated into all LC-MRM-MS studies, whether the focus of the study is biological or clinical, to allow laboratories to easily monitor their own instrumental performance and to understand and minimize variables affecting assay data quality. Incorporation of a robust SSP will help to ensure that analyte quantification measurements can be replicated with good precision within and across laboratories, and should facilitate widespread use of this technology by the basic biomedical and clinical laboratory research communities.

Public Access to the Data—An ftp server at the National Institute of Standards (NIST) was used by the CPTAC teams for initiating uploads and downloads of large data files. The data associated with this manuscript may be found at http://proteome.gs.washington.edu/software/panorama/abbatiello_study9s.html. Accessible information includes processed data results.

AUTHOR CONTRIBUTIONS: The CPTAC Network contributed collectively to this study. The following CPTAC Network investigators made significant intellectual contributions to the work described:

Study design and SOP development: S. E. Abbatiello, T. Addona, S. A. Carr, S. C. Hall, B. Schilling, N. Sedransk, L. J. Zimmerman.

Sample preparation: K. Shaddox, C. Whitwell, L. J. Zimmerman.

Data acquisition and processing: S. E. Abbatiello, S. Allen, N. G. Dodder, M. Ghosh, J. Held, V. Hedrick, H. D. Inerowicz, A. Jackson, H. Keshishian, J. W. Kim, J. S. Lyssand, C. P. Riley, P. Sadowski, B. Schilling, D. Smith, D. Tomazela, Å. Wahlander, S. Waldemarson, J. R. Whiteaker, J. You, L. J. Zimmerman.

Bioinformatics and statistical analysis: M. P. Cusack, X. Feng, B. MacLean, D. Mani, N. Sedransk.

Manuscript writing and preparation: S. E. Abbatiello, S. A. Carr, S. C. Hall, D.R. Mani, B. Schilling, L. J. Zimmerman.

General: C. Buck, C. H. Borchers, S. J. Fisher, B. W. Gibson, C. R. Kinsinger, D. Liebler, M. MacCoss, M. Mesri, T. Neubert, A. G. Paulovich, F. E. Regnier, H. Rodriguez, P. Rudnick, S. J. Skates, P. Tempst, M. Wang.

REFERENCES

1. Gerber, S. A., Rush, J., Stemman, O., Kirschner, M. W., and Gygi, S. P. (2003) Absolute quantification of proteins and phosphoproteins from cell lysates by tandem MS. *Proc. Natl. Acad. Sci. U.S.A.* **100,** 6940–6945
2. Barnidge, D. R., Dratz, E. A., Martin, T., Bonilla, L. E., Moran, L. B., and Lindall, A. (2003) Absolute quantification of the G protein-coupled receptor rhodopsin by LC/MS/MS using proteolysis product peptides and synthetic peptide standards. *Anal. Chem.* **75,** 445–451
3. Barr, J. R., Maggio, V. L., Patterson Jr., D. G., Cooper, G. R., Henderson, L. O., Turner, W. E., Smith, S. J., Hannon, W. H., Needham, L. L., and Sampson, E. J. (1996) Isotope-dilution mass spectrometric quantification of specific proteins: model application with apolipoprotein A-1. *Clin. Chem.* **42,** 1676–1682
4. Kuzyk, M. A., Smith, D., Yang, J., Cross, T. J., Jackson, A. M., Hardie, D. B., Anderson, N.L., and Borchers, C.H. (2009) Multiple reaction monitoring-based, multiplexed, absolute quantitation of 45 proteins in human plasma. *Mol. Cell. Proteomics* **8,** 1860–1877
5. Keshishian, H., Addona, T. A., Burgess, M., Kuhn, E., and Carr, S. A. (2007) Quantitative, multiplexed assays for low abundance proteins in plasma by targeted mass spectrometry and stable isotope dilution. *Mol. Cell. Proteomics* **6,** 2212–2229
6. Abbatiello, S. E., Pan, Y. X., Zhou, M., Wayne, A. S., Veenstra, T. D., Hunger, S. P., Kilberg, M. S., Eyler, J. R., Richards, N. G., and Conrads, T. P. (2008) Mass spectrometric quantification of asparagine synthetase in circulating leukemia cells from acute lymphoblastic leukemia patients. *J. Proteomics* **71,** 61–70
7. Hoofnagle, A. N., Becker, J. O., Wener, M. H., and Heinecke, J. W. (2008) Quantification of thyroglobulin, a low-abundance serum protein, by immunoaffinity peptide enrichment and tandem mass spectrometry. *Clin. Chem.* **54,** 1796–1804
8. Keshishian, H., Addona, T., Burgess, M., Mani, D., Shi, X., Kuhn, E., Sabatine, M., Gerszten, R., and Carr, S. (2009) Quantification of cardiovascular biomarkers in patient plasma by targeted mass spectrometry and stable isotope dilution. *Mol. Cell. Proteomics* **8,** 2339–2349
9. Kumar, V., Barnidge, D. R., Chen, L. S., Twentyman, J. M., Cradic, K. W., Grebe, S. K., and Singh, R. J. (2010) Quantification of serum 1–84 parathyroid hormone in patients with hyperparathyroidism by immunocapture in situ digestion liquid chromatography-tandem mass spectrometry. *Clin. Chem.* **56,** 306–313
10. Huttenhain, R., Soste, M., Selevsek, N., Rost, H., Sethi, A., Carapito, C., Farrah, T., Deutsch, E. W., Kusebauch, U., Moritz, R. L., Nimeus-Malmstrom, E., Rinner, O., and Aebersold, R. (2012) Reproducible quantification of cancer-associated proteins in body fluids using targeted proteomics. *Sci. Transl. Med.* **4**
11. Kuhn, E., Addona, T. A., Keshishian, H., Burgess, M., Mani, D. R., Lee, R. T., Sabatine, M. S., Gerszten, R. E., and Carr, S. A. (2009) Developing multiplexed assays for troponin I and interleukin-33 in plasma by peptide immunoaffinity enrichment and targeted mass spectrometry. *Clin. Chem.* **55,** 1108–1117

12. Anderson, N. L., Anderson, N. G., Haines, L. R., Hardie, D. B., Olafson, R. W., and Pearson, T. W. (2004) Mass spectrometric quantitation of peptides and proteins using Stable Isotope Standards and Capture by Anti-Peptide Antibodies (SISCAPA). *J. Proteome Res.* **3,** 235–244

13. Whiteaker, J. R., Zhao, L., Anderson, L., and Paulovich, A. G. (2010) An automated and multiplexed method for high throughput peptide immunoaffinity enrichment and multiple reaction monitoring mass spectrometry-based quantification of protein biomarkers. *Mol. Cell. Proteomics* **9,** 184–196

14. Stahl-Zeng, J., Lange, V., Ossola, R., Eckhardt, K., Aebersold, R., and Domon, B. (2007) High sensitivity detection of plasma proteins by multiple reaction monitoring of N-glycosites. *Mol. Cell. Proteomics* **6,** 1809–1817

15. Escher, C., Reiter, L., MacLean, B., Ossola, R., Herzog, F., Chilton, J., MacCoss, M. J., and Rinner, O. (2012) Using iRT, a normalized retention time for more targeted measurement of peptides. *Proteomics* **12,** 1111–1121

16. Hammett-Stabler, C. A., and Garg, U. (2010) The evolution of mass spectrometry in the clinical laboratory. *Methods Mol. Biol.* **603,** 1–7

17. Strathmann, F. G., and Hoofnagle, A. N. (2011) Current and future applications of mass spectrometry to the clinical laboratory. *Am. J. Clin. Pathol.* **136,** 609–616

18. Percy, A. J., Chambers, A. G., Smith, D. S., and Borchers, C. H. (2013) Standardized Protocols for Quality Control of MRM-based Plasma Proteomic Workflows. *J. Proteome Res.* **12,** 222–233

19. Heller, D. N. (2003) Guidance for Industry: Mass Spectrometry for Confirmation of the Identity of Animal Drug Residues, C.f.V.M. U.S. Department of Health and Human Services, ed., pp. 1–11

20. Abbatiello, S. E., Mani, D. R., Keshishian, H., and Carr, S. A. (2010) Automated detection of inaccurate and imprecise transitions in peptide quantification by multiple reaction monitoring mass spectrometry. *Clin. Chem.* **56,** 291–305

21. Addona, T. A., Abbatiello, S. E., Schilling, B., Skates, S. J., Mani, D. R., Bunk, D. M., Spiegelman, C. H., Zimmerman, L. J., Ham, A. J. L., Keshishian, H., Hall, S. C., Allen, S., Blackman, R. K., Borchers, C. H., Buck, C., Cardasis, H. L., Cusack, M. P., Dodder, N. G., Gibson, B. W., Held, J. M., Hiltke, T., Jackson, A., Johansen, E. B., Kinsinger, C. R., Li, J., Mesri, M., Neubert, T. A., Niles, R. K., Pulsipher, T. C., Ransohoff, D., Rodriguez, H., Rudnick, P. A., Smith, D., Tabb, D. L., Tegeler, T. J., Variyath, A. M., Vega-Montoto, L. J., Wahlander, A., Waldemarson, S., Wang, M., Whiteaker, J. R., Zhao, L., Anderson, N. L., Fisher, S. J., Liebler, D. C., Paulovich, A. G., Regnier, F. E., Tempst, P., and Carr, S. A. (2009) Multi-site assessment of the precision and reproducibility of multiple reaction monitoring-based measurements of proteins in plasma. *Nat. Biotechnol.* **27,** 633–641

22. Xia, J.Q., Sedransk, N., and Feng, X. (2011) Variance Component Analysis of a Multi-Site Study for the Reproducibility of Multiple Reaction Monitoring Measurements of Peptides in Human Plasma. *PLoS ONE* **6,** e14590

23. Rudnick, P. A., Clauser, K. R., Kilpatrick, L. E., Tchekhovskoi, D. V., Neta, P., Blonder, N., Billheimer, D. D., Blackman, R. K., Bunk, D. M., Cardasis, H. L., Ham, A. J. L., Jaffe, J. D., Kinsinger, C. R., Mesri, M., Neubert, T. A., Schilling, B., Tabb, D. L., Tegeler, T. J., Vega-Montoto, L., Variyath, A. M., Wang, M., Wang, P., Witeaker, J. R., Zimmerman, L. J., Carr, S. A., Fisher, S. J., Gibson, B. W., Paulovich, A. G., Regnier, F. E., Rodriguez, H., Spiegelman, C., Tempst, P., Liebler, D. C., and Stein, S. E. (2010) Performance metrics for liquid chromatography-tandem mass spectrometry systems in proteomics analyses. *Mol. Cell. Proteomics* **9,** 225–241

24. Bell, A. W., Deutsch, E. W., Au, C. E., Kearney, R. E., Beavis, R., Sechi, S., Nilsson, T., Bergeron, J. J., and Group, H. T. S. W. (2009) A HUPO test sample study reveals common problems in mass spectrometry-based proteomics. *Nat. Methods* **6,** 423–430

25. Burkhart, J. M., Premsler, T., and Sickmann, A. (2011) Quality control of nano-LC-MS systems using stable isotope-coded peptides. *Proteomics* **11,** 1049–1057

26. Briscoe, C. J., Stiles, M. R., and Hage, D. S. (2007) System Suitability in bioanalytical LC/MS/MS. *J. Pharmaceutical Biomed. Anal.* **44,** 484–491

27. 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

28. Frewen, B., and MacCoss, M. J. (2007) Using BiblioSpec for Creating and Searching Tandem MS Peptide Libraries. In Current Protocols in Bioinformatics (John Wiley & Sons, Inc.)

29. Bereman, M. S., MacLean, B., Tomazela, D. M., Liebler, D. C., and MacCoss, M. J. (2012) The development of selected reaction monitoring methods for targeted proteomics via empirical refinement. *Proteomics* **12,** 1134–1141

30. Abbatiello, S. E., Schilling, B., MacLean, B., Sadowski, P., Jackson, A. M., Ghosh, M., Keshishian, H., Addona, T., Whiteaker, J., Allen, S., Burgess, M., Sedransk, N., Mani, D. R., Hall, S. C., Carr, S. A., Network, C., Zimmerman, L., and Tomazela, D. (2011) Meeting the design, development and implementation challenges of >100-plex quantitative assays for proteins in plasma: A large-scale, NCI-CPTAC interlaboratory study. paper presented at: American Society for Mass Spectrometry (Denver, Colorado)

31. Team, R. D. C. (2008) R: A language and environment for statistical computing. (Vienna, Austria: R Foundation for Statistical Computing)

32. Giddings, J. C., and Dahlgren, K. (1971) Resolution and Peak Capacity in Equilibrium-Gradient Methods of Separation. *Separation Sci.* **6,** 345–356

33. Radulovic, D., Jelveh, S., Ryu, S., Hamilton, T. G., Foss, E., Mao, Y., and Emili, A. (2004) Informatics Platform for Global Proteomic Profiling and Biomarker Discovery Using Liquid Chromatography-Tandem Mass Spectrometry. *Mol. Cell. Proteomics* **3,** 984–997

34. Sarkar, D. (2008) Lattice: Multivariate Data Visualization with R. (Springer)

35. Sharma, V., MacLean, B., Eckels, J., Stergachis, A. B., Jaffe, J. D., and MacCoss, M. J. (2012) Panorama: A repository of targeted proteomics assays for Skyline. Paper presented at: American Society for Mass Spectrometry (Vancouver, British Columbia: Springer)

36. Nelson, E. K., Piehler, B., Eckels, J., Rauch, A., Bellew, M., Hussey, P., Ramsay, S., Nathe, C., Lum, K., Krouse, K., Stearns, D., Connolly, B., Skillman, T., and Igra, M. (2011) LabKey Server: An open source platgorm for scientific data integration, analysis and collaboration. *BMC Bioinformatics* **12,** 1–23