analytical chemistry

Article

pubs.acs.org/ac

MultiNotch MS3 Enables Accurate, Sensitive, and Multiplexed Detection of Differential Expression across Cancer Cell Line Proteomes

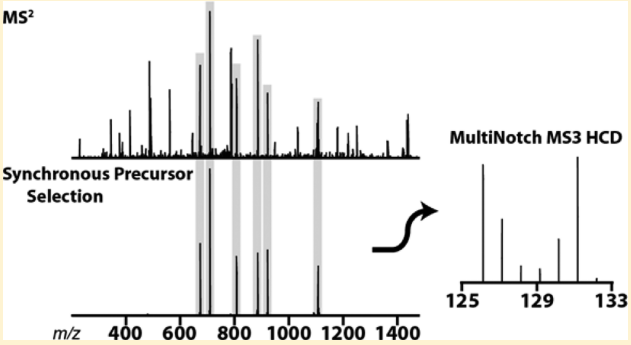
Graeme C. McAlister,[†] David P. Nusinow,[†,§] Mark P. Jedrychowski,[†,‡,§] Martin Wühr,[†,‡] Edward L. Huttlin,[†] Brian K. Erickson,[†] Ramin Rad,[†] Wilhelm Haas,[†] and Steven P. Gygi*[,†]

[†]Harvard Medical School, Department of Cell Biology, Boston, Massachusetts 02115, United States

[‡]Harvard Medical School, Department of Systems Biology, Boston, Massachusetts 02115, United States

Supporting Information

ABSTRACT: Multiplexed quantitation via isobaric chemical tags (e.g., tandem mass tags (TMT) and isobaric tags for relative and absolute quantitation (iTRAQ)) has the potential to revolutionize quantitative proteomics. However, until recently the utility of these tags was questionable due to reporter ion ratio distortion resulting from fragmentation of coisolated interfering species. These interfering signals can be negated through additional gas-phase manipulations (e.g., MS/MS/MS (MS3) and proton-transfer reactions (PTR)). These methods, however, have a significant sensitivity penalty. Using isolation waveforms with multiple frequency notches (i.e., synchronous precursor selection, SPS), we coisolated and cofragmented multiple MS2 fragment ions, thereby increasing the number of reporter ions in the MS3 spectrum 10-fold over the standard MS3 method (i.e., MultiNotch MS3). By increasing the reporter ion signals, this method improves the dynamic range of reporter ion quantitation, reduces reporter ion signal variance, and ultimately produces more high-quality quantitative measurements. To demonstrate utility, we analyzed biological triplicates of eight colon cancer cell lines using the MultiNotch MS3 method. Across all the replicates we quantified 8 378 proteins in union and 6 168 proteins in common. Taking into account that each of these quantified proteins contains eight distinct cell-line measurements, this data set encompasses 174 704 quantitative ratios each measured in triplicate across the biological replicates. Herein, we demonstrate that the MultiNotch MS3 method uniquely combines multiplexing capacity with quantitative sensitivity and accuracy, drastically increasing the informational value obtainable from proteomic experiments.



Mass spectrometry (MS) based quantitative proteomics has traditionally been limited to binary and ternary comparisons (e.g., SILAC based quantitation).[1−4] As such, proteomics has trailed behind the technologies employed in transcriptome analysis. Multiplexed quantitation via isobaric chemical tags (e.g., tandem mass tags (TMT) and isobaric tags for relative and absolute quantitation (iTRAQ)) provide an avenue for greater parallelization of quantitative mass spectrometry.[5−7]

Identical peptides, derived from different samples, and labeled with different versions of the isobaric tags, are indistinguishable in their intact form. However, upon isolation and fragmentation in the mass spectrometer, each peptide variant produces a unique reporter ion. Multiplexing quantitative analyses through multichannel isobaric tagging shows great promise for its ability to (1) improve throughput, (2) increase the breadth of coverage by avoiding missing values, and (3) deepen analysis by simplifying complex chromatograms that are typically populated by multiple forms of the same peptide.[8]

In theory, the abundance of the isobaric tag reporter ions should be directly proportional to the relative amount of each precursor in each sample. In practice, however, coisolation and cofragmentation of interfering ions results in distorted TMT ratios. Reporter ions originating from the isobaric tags of the target population are indistinguishable from reporter ions originating from any interfering ions. Therefore, any coisolated interfering precursor ions will contribute to the final reporter ion population in an unpredictable manner that obfuscates the true reporter ion intensities.

This interference phenomenon was detailed recently in a series of publications.[9−15] Using a two-proteome model, we showed that nearly all the measurements obtained with standard tandem MS (i.e., MS2) were distorted by interfering ions. We also demonstrated that an MS3 spectrum based on one of the TMT labeled MS2 fragment ions can mitigate the

Received: June 3, 2014
Accepted: June 13, 2014
Published: June 13, 2014

   dx.doi.org/10.1021/ac502040v | *Anal. Chem.* 2014, 86, 7150−7158

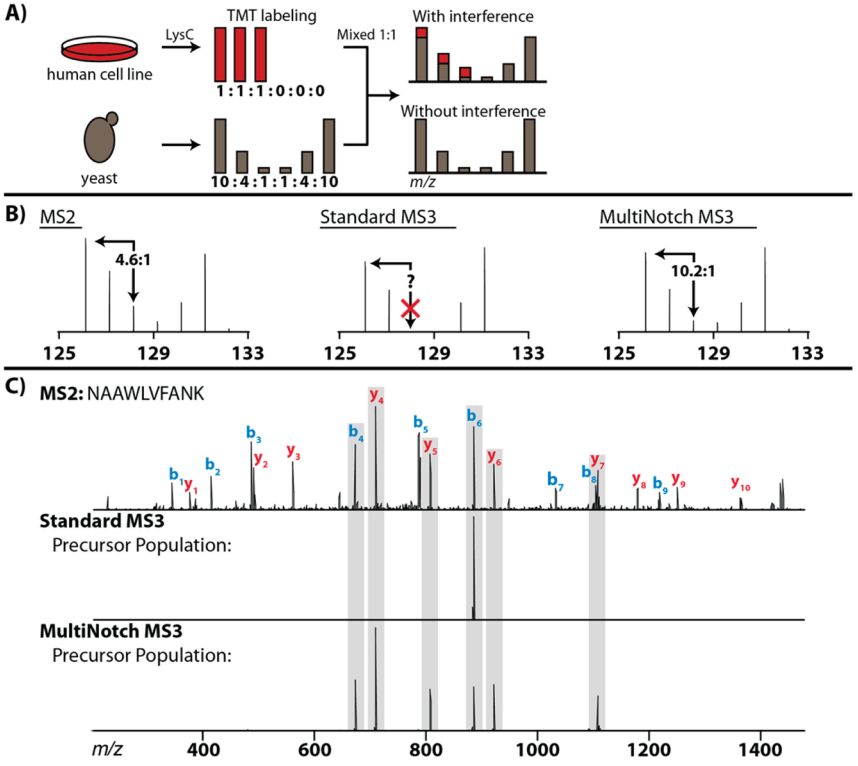


Figure 1. (A) Yeast was digested with LysC and labeled with TMT (10:4:1:1:4:10). That sample was combined with a TMT labeled HeLa sample (1:1:1:0:0:0). (B) A TMT-labeled, yeast peptide (NAAWLVFANK) was interrogated in back-to-back scans using (left spectrum) MS2, where the MS1 precursor was fragmented using HCD. (Middle) MS3, where the MS1 precursor was fragmented with CID, and a single MS2 product ion was isolated and fragmented using HCD. And, (right) MultiNotch MS3, where multiple MS2 product ions were simultaneously isolated and fragmented. (C) The precursor populations of the standard and MultiNotch MS3 scans used to generate the reporters above (middle and bottom spectra, respectively). For reference, we also include the ITMS2 spectrum prior to MS3 precursor isolation (top).

negative impact of these interfering signals.[14] Alternative methods that utilize ion−ion chemistry in place of energetic fragmentation have been demonstrated.[15] While the additional round of gas-phase selectivity provided by these MS3 methods dramatically reduces the contribution of any interfering signals, it also reduces overall sensitivity. By dividing the initial precursor signal among all the possible product ions and selecting only a single product ion for subsequent interrogation, only a small percentage of MS1 precursor ions are converted into the MS3 reporter ions.

Herein, we describe a solution to the sensitivity limitations of the MS3 method, in which we use isolation waveforms with multiple frequency notches for synchronous precursor selection (SPS) of multiple MS2 fragment ions. We fragment the aggregate MS3 precursor population (i.e., MultiNotch MS3) to produce a reporter ion population that is far more intense than the population we would have produced had we only fragmented a single MS2 ion. At the same time, we maintain the selectivity of the standard MS3 method by carefully defining the isolation notches of the SPS isolation waveform to ensure high isolation specificity of the target MS2 fragment ions. In summary, we introduce a new quantitative proteomic method, which provides a unique combination of multiplexing capacity, high sensitivity, and quantitative accuracy.

METHODS

Two-Proteome Interference Model. The two-proteome interference model was prepared as previously.[14,16] HeLa S3 cells were grown in suspension to 1×10^6 cells/mL. Yeast cells were grown to an OD of 1.0. Cells were lysed in 6 M guanidiumthiocyanate, 50 mM Hepes (pH 8.5, HCl). Protein content was measured using a BCA assay (Thermo Scientific), disulfide bonds were reduced with dithiothreitol (DTT), and cysteine residues were alkylated with iodoacetamide as previously described.[17] Protein lysates were cleaned with methanol−chloroform precipitation.[18] The samples were redissolved in 6 M guanidiumthiocyanate, 50 mM Hepes pH 8.5, and diluted to 1.5 M guanidium thiocyanate, 50 mM Hepes (pH 8.5). Both lysates were digested overnight with Lys-C (Wako) in a 1/50 enzyme/protein w/w ratio. Following digestion, the sample was acidified with TFA to a pH < 2 and subjected to C_{18} solid-phase extraction (SPE, Sep-Pak, Waters).

The TMT reagents were dissolved in 40 μL of acetonitrile, and 10 μL of the solution was added to 100 μg of peptides dissolved in 100 μL of 50 mM HEPES (pH 8.5). After incubating for 1 h at room temperature (22 °C), the reaction was quenched by adding 8 μL of 5% w/v hydroxylamine. Following labeling, the sample was combined in desired ratios. Yeast aliquots were mixed at 10:4:1:1:4:10, and HeLa was mixed at 1:1:1:0:0:0 (Figure 1A). Those two samples were then mixed at a 1/1 w/w ratio and subjected to C_{18} solid-phase extraction.

Colorectal Cancer Cell Culture, Sample Preparation, and TMT Labeling. Colo205, LoVo, DLD-1, SW48, HT-29, HCT-15, HT55, and HCT-116 cells were cultured in 15 cm plates containing RPMI, 10% FBS, penicillin, and streptomycin. Each cell line was grown in 10% CO_2 to ~90% confluence. Cells were starved for 4 h in RPMI, washed 3 times with 15 mL of cold PBS, harvested into an Eppendorf tube, and snap frozen in liquid nitrogen. Following harvesting, the remaining sample

dx.doi.org/10.1021/ac502040v | Anal. Chem. 2014, 86, 7150−7158

steps in the sample preparation process closely matched the steps of the two-proteome sample (Supporting Information).

Basic-pH Fractionation and Low-pH Reverse Phase LC Analysis. Human colorectal peptides were subjected to basic-pH reverse-phase HPLC fractionation. Mixed and labeled peptides were solubilized in buffer A (5% ACN, 10 mM ammonium bicarbonate, pH 8.0) and separated on an Agilent 300 Extend C18 column (5 μm particles, 4.6 mm i.d., and 20 cm in length). Using an Agilent 1100 binary pump equipped with a degasser and a photodiode array (PDA) detector, a 50 min linear gradient from 18% to 45% acetonitrile in 10 mM ammonium bicarbonate pH 8 (flow rate of 0.8 mL/min) separated the peptide mixture into a total of 96 fractions. The 96 fractions were consolidated into 24 samples, acidified with 10% formic acid, and vacuum-dried. Each sample was redissolved with 5% formic acid/5% ACN, desalted via StageTip, dried via vacuum centrifugation, and reconstituted for LC–MS/MS analysis.

All LC–MS experiments were performed on a Velos-Orbitrap Elite mass spectrometer (Thermo Fischer Scientific) coupled to a Proxeon nLC-1000 (Thermo Fisher Scientific) ultra high-pressure liquid chromatography (UPLC) pump. Peptides were separated on a 75 μm inner diameter microcapillary column. The tip for the column was pulled in-house and the column was packed with approximately 0.5 cm of Magic C4 resin (5 μm, 100 Å, Michrom Bioresources) followed by 25 cm of Sepax Technologies GP-C18 resin (1.8 μm, 120 Å). Separation was achieved by applying a 3–22% ACN gradient in 0.125% formic acid over 165 min at ~300 nL/min. Electrospray ionization was enabled by applying a voltage of 2.0 kV through an IDEX high-pressure fitting at the inlet of the microcapillary column. In the case of the two-proteome mixture, the linear gradient was shortened to 70 min.

Implementation of the MultiNotch MS3 Method. The implementation of the isolation waveforms for synchronous precursor selection was the greatest hurdle toward implementing the MultiNotch MS3 method. All instrument modifications needed to enable this method were performed in-house and entailed changes to the instrument control code. The fundamental technologies of this method are ion traps and notched isolation waveforms. The idea of concurrently isolating multiple discrete ion populations with isolation waveforms that have multiple frequency notches was developed in the mid-1980s, when much of the original work was done in the laboratories of Marshall[19,20] and McLafferty.[21] Since then Cooks, McLuckey, and others have contributed to the field.[22,23] Yet, this is the first time this technology has ever been used in an online, large-scale proteomics experiment with multiplexed quantitation. Briefly, these waveforms isolate the ions of interest by ejecting unwanted ions through energetic excitation. The frequencies comprising the isolation waveform excite the trapped ions through on-resonance excitation. However, where a notch exists in the list of frequencies, the ions remain stably trapped (Figure S1 in the Supporting Information).

All of the initial development work for this method was focused on defining equations that could accurately describe the size and location for any possible notch in the SPS waveform. To this end, we injected mixtures of ions with known m/z values into the instrument, and we isolated those ions using a series of SPS isolation waveforms. During these experiments, we varied the notch width and location, and the isolation efficiency was recorded as a function of those parameters. Figure 2A shows the data generated by performing

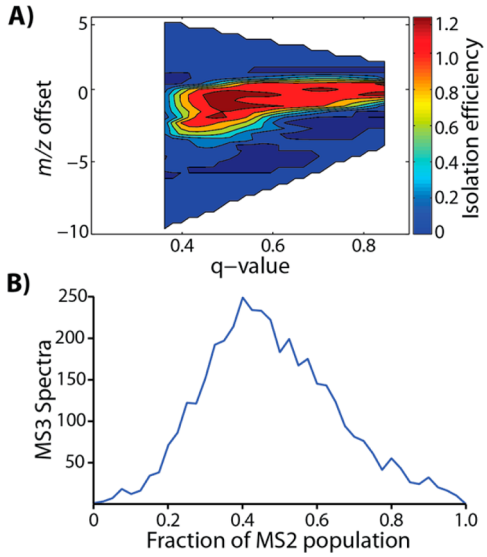


Figure 2. (A) While infusing an ion with an m/z ratio of 989, we varied the isolation notch width and location. We recorded the isolation efficiency as a function of those parameters. This analysis was repeated for the series of ions, and the resulting data set was fitted using linear regression. (B) During a 90 min MultiNotch LC–MS2/MS3 analysis of the yeast/human two-proteome sample, we isolated the MS3 precursor population without any subsequent fragmentation. We then calculated the fraction of MS2 ions retained in the MS3 precursor population.

this analysis on an ion with 989 m/z. As the q-value of the ion varies, so does the optimal position and width of the isolation notch. This type of analysis was repeated for many ions, and the resulting aggregate data sets were fitted using linear regression (Figure S2 in the Supporting Information).

With these equations in hand, we could define isolation waveforms for any ensemble of ions. Typically we captured 10 MS2 fragment ions in our MS3 precursor population. Surprisingly, by simply isolating 10 MS2 fragment ions, we are able to retain ~40% of the MS2 total ion current in the MS3 precursor population (Figure 2B). Coupling in the losses from CID, we typically retained ~25% of our initial MS1 population in the MS3 precursor population. This was a vast improvement over the traditional MS3 approach, where only ~5% of the initial precursor population was retained.

In addition to the waveform declaration algorithms, we employed functions that intelligently scaled the precursor injection time based upon the expected TMT signal (Supporting Information). Coupled together, these features for selecting precursors, declaring the MultiNotch MS3 isolation waveform, and intelligently scaling ion injection times resulted in a 10-fold increase in TMT reporter ion intensities with the MultiNotch MS3 approach, relative to the standard MS3 scan type (Figure 3A).

Mass Spectrometry. The mass spectrometer was operated in data-dependent mode for both the MS2 and MS3 methods. For both methods we collected a survey scan of 300–1 500 m/z in the Orbitrap at a resolution of 60 000 (FTMS1) and an AGC target of 1×10^6. We selected the 10 most intense ions for MS analysis. Precursor ions were filtered according to charge state (required >1z), dynamic exclusion (40 s with a ±10 ppm window), and monoisotopic precursor selection.

During the MS2 analyses, precursors were fragmented by high-energy collision induced dissociation (HCD) followed by

dx.doi.org/10.1021/ac502040v | Anal. Chem. 2014, 86, 7150–7158

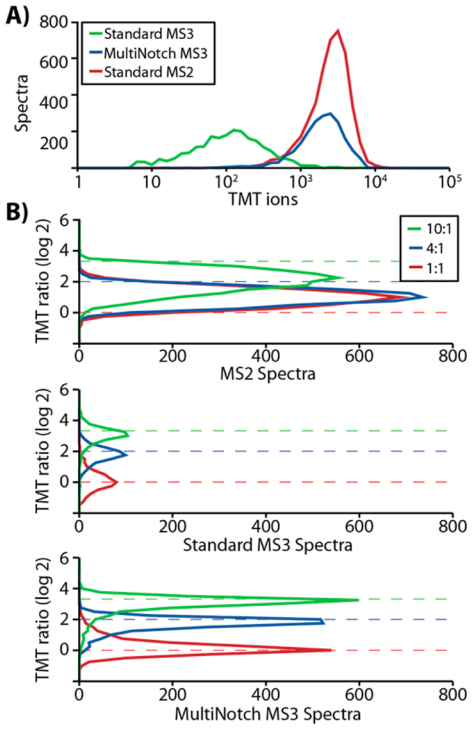


Figure 3. TMT-labeled, two-proteome mixture (yeast/human) was analyzed by LC−MS2, standard MS3, and MultiNotch MS3. (A) We distributed the quantitative spectra by the number of TMT reporter ions. (B) We also distributed the quantitative spectra using three select TMT ratios (channels 126:128, 127:128, and 128:129, i.e., 10:1, 4:1, and 1:1). The expected ratios are denoted using the dashed lines.

Orbitrap analysis (FTMS2). FTMS2 precursors were isolated using a width of 2.0 m/z and fragmented with a normalized collision energy of 40. Precursors were accumulated to an AGC target of 5×10^4 or a maximum injection time of 250 ms.

During the MS3 analyses, the MS1 precursors were first interrogated by ITMS2 using CID. Precursors were isolated using a 1.2 m/z isolation window. They were accumulated to an AGC target of 5 000 or a maximum injection time of 125 ms. This ITMS2 spectrum was used to determine the conditions of the MS3 analysis (e.g., which fragments to interrogate). For the MS3 scan, the MS1 precursor was isolated using a 2.5 m/z wide window and fragmented with CID. Following fragmentation, the MS3 precursor population was isolated using the SPS waveform and then fragmented by HCD. The HCD normalized collision energy was set to 50. The m/z value used in the NCE calculation was the weighted average of all the MS3 precursor ions. During the MS3 analysis we used an online isolation specificity filter (Supporting Information).

MS2 Spectra Assignment, Data Processing, and Protein Assignment. Following data acquisition, Thermo RAW files were processed using a series of software tools that were developed in-house. First the RAW files were converted to mzXML using a custom version of ReAdW.exe (http://sashimi.svn.sourceforge.net/viewvc/sashimi/) that had been modified to export ion accumulation times and FT peak noise. During this initial processing we also corrected any erroneous assignments of monoisotopic m/z. Using Sequest,[24] MS2 spectra were searched against the human UniProt database (downloaded on 08/02/2011), supplemented with the sequences of common contaminating proteins such as trypsin. This forward database was followed by a decoy component, which included all target protein sequences in reversed order.

Searches were performed using a 50 ppm precursor ion tolerance.[25] When searching Orbitrap MS2 data, we used 0.02 Th fragment ion tolerance. The fragment ion tolerance was set to 1.0 Th when searching ITMS2 data. Only peptide sequences with both termini consistent with the protease specificity of LysC were considered in the database search, and up to two missed cleavages were accepted. TMT tags on lysine residues and peptide N-termini (+ 229.162932 Da) and carbamidomethylation of cysteine residues (+ 57.02146 Da) were set as static modifications, while oxidation of methionine residues (+ 15.99492 Da) was treated as a variable modification. An MS2 spectral assignment false discovery rate of less than 1% was achieved by applying the target-decoy strategy.[26] Filtering was performed using linear discriminant analysis as described previously[27] to create one composite score from the following peptide ion and MS2 spectra properties: Sequest parameters XCorr and unique ΔCn, peptide length and charge state, and precursor ion mass accuracy. The resulting discriminant scores were used to sort peptides prior to filtering to a 1% FDR, and the probability that each peptide-spectral-match was correct was calculated using the posterior error histogram.

Following spectral assignment, peptides were assembled into proteins and proteins were further filtered based on the combined probabilities of their constituent peptides to a final FDR of 1%. In cases of redundancy, shared peptides were assigned to the protein sequence with the most matching peptides, thus adhering to principles of parsimony.[28]

Quantitative Data Analysis and Data Presentation. When analyzing TMT reporter ion signals, we used the ratio between the reporter ion intensity and the peak noise. This ratio has been shown to scale quite well with the number of ions in the Orbitrap peak, i.e., ∼5 charges are equal to a S/N ratio of 1.[29] Hence, this ratio is more meaningful when making qualitative judgments about the ion statistics of a given reporter signal than the injection time scaled intensity. The isotopic impurities of the TMT reagent were corrected using the values specified by the manufacturer.[30]

When processing the colorectal cancer data, we filtered our data to only include quantitative spectra that possessed a summed TMT reporter S/N of 100. On the basis of previous work and simulation, this was determined to be the minimum number of reporter ions necessary to ensure that ion statistics would not be a source of significant reporter variance (Figure S6 in the Supporting Information).[31]

Using our protein-level peptide grouping as a guide, we summed the TMT signals from the peptide level quantitative spectra to produce our protein level quantitative data. These quantitative signals were summed across all quantified proteins, and then these sums were normalized across all TMT reporter channels, hence, correcting for minor mixing errors and reflecting equal total protein content per cell line. These normalized intensities were then scaled across each protein to a net total of 100.

To look for proteins that were significantly altered in at least one colorectal cancer cell line, we used a one-way analysis of variance (ANOVA) with Welch's correction to control for unequal variances between TMT channels.[32] An false discovery rate (FDR) of 1% was controlled using the Benjamini−Hochberg method. The ReactomeFI analysis was performed using the Cytoscape plug-in provided by the developers.[33] Principal components analysis was performed using *R* (R Core

dx.doi.org/10.1021/ac502040v | *Anal. Chem.* 2014, 86, 7150−7158

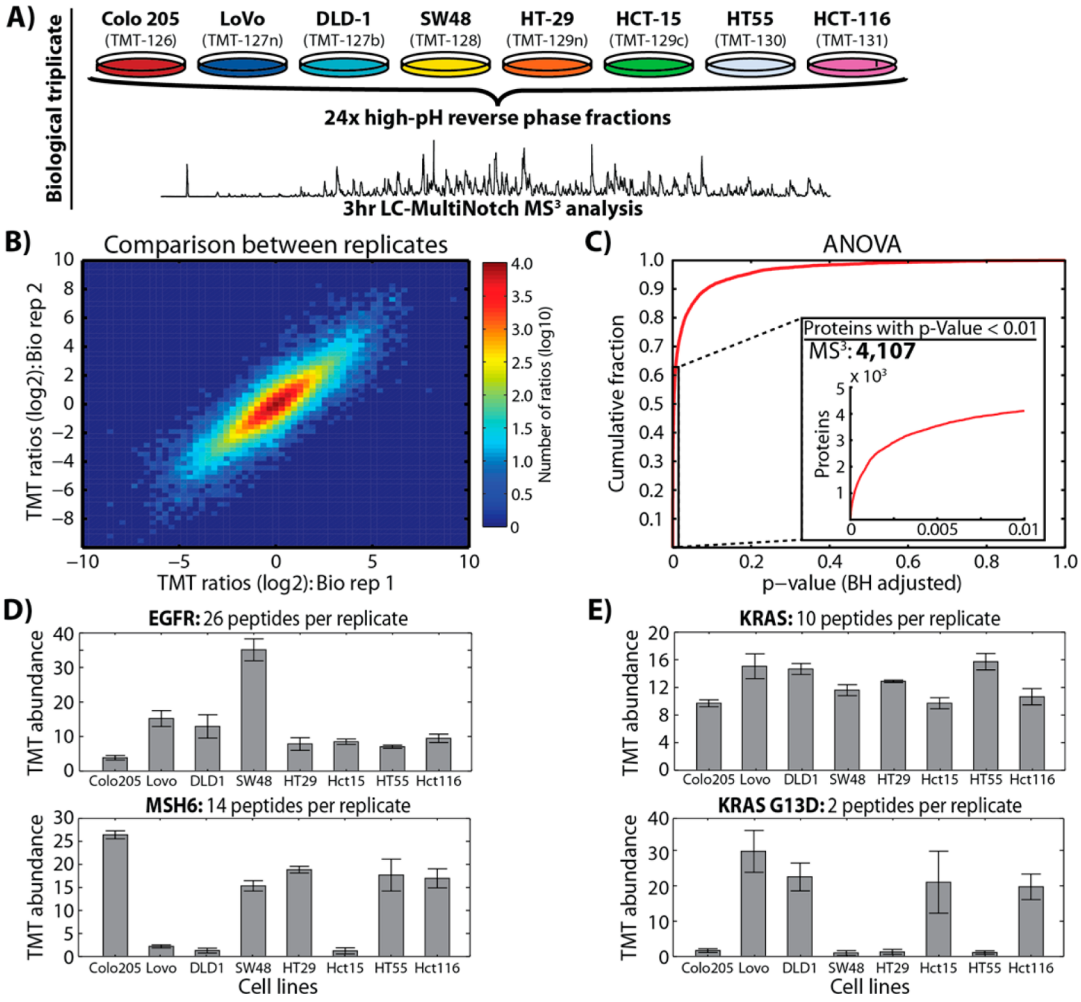


Figure 4. (A) Eight colorectal cancer cell lines were grown in biological triplicate. Each replicate was digested with LysC, labeled with TMT, fractionated, and analyzed using MultiNotch MS3 (3-h LC gradients). (B) All protein ratios from replicates 1 and 2 were plotted against each other. In total this represents 172 704 quantitative ratios. (C) Across the three replicate we performed a one way ANOVA with Welch's correction. (D) We highlighted the protein expression profile for two commonly studied proteins, EGFR and MSH6, and (E) the WT and mutant (G13D) forms of KRAS.

Team, Vienna, Austria, http://www.R-project.org). PC1 loadings were extracted and sorted by absolute value to choose the top contributors. All proteins from the Vogelstein et al. gene set that were quantified in the MultiNotch MS3 data set were extracted and mean expression values across all replicates were used to calculate a Euclidean distance to SMAD4 and IDH2. Proteins were then sorted by this distance to produce the nearest expression profiles.

RESULTS

Coisolating and Cofragmenting Multiple MS2 Fragment Ions. Using isolation waveforms with multiple frequency notches, we coisolated and cofragmented multiple MS2 fragment ions during our MS3 analyses (MultiNotch MS3). We implemented this method with the aim of increasing the number of TMT reporter ions in the resulting quantitative MS3 spectrum. However, with the MultiNotch MS3 method we employ multiple isolation notches, and some of those notches tend to be larger than the isolation notches of the standard MS3 method (see Methods); hence, we lose some selectivity with the MultiNotch method compared to the standard. With these concerns in mind, we benchmarked the MultiNotch MS3 method using a two-proteome model to measure ratio distortion and sensitivity.[14]

We digested a yeast lysate with LysC, labeled separate aliquots using TMT, and then mixed those aliquots at 10:4:1:1:4:10 (Figure 1A). We also digested a human lysate using LysC, labeled three aliquots, and mixed those aliquots at ratios of 1:1:1. Finally, the labeled human and yeast peptide mixtures were combined at a 1:1 ratio.

For all subsequent analyses, we treated the yeast peptides as the target population, and the human peptides as the interfering ions. In the absence of any interfering human signals, yeast peptides should produce a reporter ion distribution that corresponds to the mixing ratios (i.e., 10:4:1:1:4:10). However, with the MS2 method we generally observed a distortion of these intensities. This is exemplified by the left spectrum of Figure 1B, where a ratio that should have been 10:1 has been distorted to 5:1 by the presence of interfering reporter ions.

When the same yeast peptide precursor ion was interrogated using a standard MS3 scan (Figure 1B, middle spectrum), due to poor sensitivity we were unable to detect the lowest abundance channels. In comparison, when we used the MultiNotch MS3 method to coisolate and cofragment multiple peaks from the MS2 spectrum, even the lowest abundance

dx.doi.org/10.1021/ac502040v | Anal. Chem. 2014, 86, 7150−7158

channels were detected and the ratios remained accurate (Figure 1B, right spectrum).

In Figure 1C, we provide spectra of the MS3 precursor populations used to generate the reporter ion distributions of Figure 1B. The middle spectrum presents analysis of the standard MS3 precursor population, while the bottom spectrum presents the MultiNotch MS3 precursor population. For reference we also include the spectrum on top, in which we analyzed the MS2 fragment ion population prior to isolating the MS3 precursor populations. By increasing the number of precursor from 1 to 6, we are able to substantially increase the final reporter ion population. At the same time the notched isolation waveforms preserve the specificity of the MS3 scan by effectively removing unwanted ions.

Analyzing the Entire Two-Proteome Data Set. In Figure 3, we provide a more complete picture of the difference in performance between the three methods. In separate 90 min LC−MS analyses, we interrogated the two-proteome sample using the MS2, standard MS3, and MultiNotch MS3 methods. In Figure 3A, we show the distribution of TMT ions. As noted earlier, the sensitivity of the standard MS3 method is quite poor, and on average we only detected ∼200 TMT reporter ions per MS3 spectrum. The MultiNotch MS3 scan typically produced ∼2 500 reporter ions per MS3 spectrum. This is over an order of magnitude more reporter ions per quantitative spectrum, and it is also close to the number of reporter ions we produced on average using the MS2 method (∼3 100). Though, in the case of the MS3 measurements, we utilized higher AGC targets and maximum injection times than for the MS2 measurements (see Methods).

In Figure 3B we compare how accurately the three methods measured the reporter ion ratios. We focused on the 10:1 and 4:1 ratios, where all the channels have interfering signals (TMT-126/128 and TMT-127/128, respectively). We also included a 1:1 ratio, which we derived from the two channels with the lowest intensity (TMT-128/129). With this ratio one of the channels is shared with an interfering human reporter ion (TMT-128), while the other channel is pure (TMT-129). We included this last ratio to highlight how interfering signals can distort measured ratios upward as well and downward. In the standard MS2 data, the 10:1 and 4:1 ratios distorted downward such that the typical values across the entire LC−MS analysis were 5:1 and 2:1, respectively. Even more distressing, the 1:1 ratio distorted upward such that typical values were 2:1. Because of the coisolation and cofragmentation of interfering ions during an MS2 analysis, we could no longer confidently measure the difference between ratios that should be 1:1 and 4:1. In contrast, with the MS3 methods we typically recorded the expected ratios (i.e., 10:1, 4:1, and 1:1). For this analysis, we required that all six reporter channels were present before analyzing a given spectrum, which resulted in a more than 4-fold increase of the number of quantified yeast spectra from the MultiNotch MS3 method (1794), compared to the standard MS3 method (438).

Analyzing an 8-Plex Sample of Colon Cancer Cell Lines. Beyond demonstrating the technical capabilities of the MultiNotch MS3 method with the two-proteome model, we sought to demonstrate the practicality of the method using a large-scale proteomics experiment. To this end, we prepared a TMT 8-plex sample that consisted of eight different colorectal cancer cell lines: Colo-205, LoVo, DLD-1, SW48, HT-29, HCT-15, HT-55, and HCT-116 (Figure 4A). We grew all eight cell lines in biological triplicate, harvested the proteins, and digested the resulting proteome samples with LysC. Following digestion, we labeled the samples with the TMT reagents, mixed the labeled peptides, and fractionated the mixtures using offline basic-pH reverse phase HPLC. We collected 24 fractions, which we then analyzed using a 3-h LC−MS3 method. To analyze each biological replicate required 3 days of analysis time, and to collect the entire data set required 9 days.

Across the three biological replicates, the MultiNotch MS3 method quantified 8 378 proteins in at least one of the replicates and 6 168 in all three. In the latter case, this represents eight quantitative measurements per protein per biological replicate. Considering that 8 TMT channels allow the determination of 28 binary comparisons, the quantitative breadth of these analyses entails 172 704 quantitative protein abundance ratios measured in all three biological replicates.

Next, we compared the quantitative reproducibility between replicate measurements. In Figure 4B, we show all the quantitative ratios from the first biological replicate plotted against the ratios from the second. We observed high correlation between the replicate measurements (Pearson correlation coefficient of 0.86), which confirms the high reproducibility of the MultiNotch MS3 method, as well as of the TMT-based workflow, and of quantitative proteomics in general. After all, the variation in this plot is the summation of all the variances across the entire experiment, from cell culture to mass spectrometer analysis.

Using data from the biological triplicate measurements, we also performed a one way ANOVA with Welch's correction for unequal variance. We adjusted the resulting p-values using the Benjamini−Hochberg multiple test correction (Figure 4C). Across the biological data set, we detected 4 107 proteins where the cell-line specific expression profile varied in a statistically significant manner (adjusted p-value <0.01). These data demonstrate strikingly different expression levels for thousands of proteins. Nearly 10% of the quantified proteins showed a 10-fold difference somewhere in their expression profiles, and ∼75% demonstrated at least one 2-fold difference.

As examples, we highlight two proteins known to often play central roles in colorectal tumorigenesis: EGFR and MSH6 (Figure 4D). EGFR was quantified based on an average of 26 peptides per biological replicate. MSH6 was quantified based on an average of 14 peptides/replicate. Indeed, the average number of peptides used for quantitation in each replicate was ∼88 000. This good protein coverage, and the high accuracy of the MultiNotch MS3 method, translates into high reproducibility between replicates.

In Figure 4E, we highlight the expression profile of KRAS, a protein tightly associated with many cancers. As with EGFR and MSH6, we observed good protein coverage (>10 peptides per replicate) and high reproducibility between the biological replicates. In the bottom panel of Figure 4E, we highlight the detection and quantitation of a peptide from a mutated form of KRAS. Four of the cell lines examined are known to harbor this KRAS-activating mutation (DLD-1, HCT-116, HCT-15, and Lovo). This mutation involves a substitution of an aspartic acid (D) for the glycine (G) at the 13th position. Because this mutation entails a single amino acid substitution, to unambiguously quantify the mutated form of the protein, the MS3 method must successfully interrogate a single peptide in the pool of hundreds of thousands that comprise the sample. In spite of this hurdle, the MultiNotch MS3 quantified the mutated form in all three biological replicates.

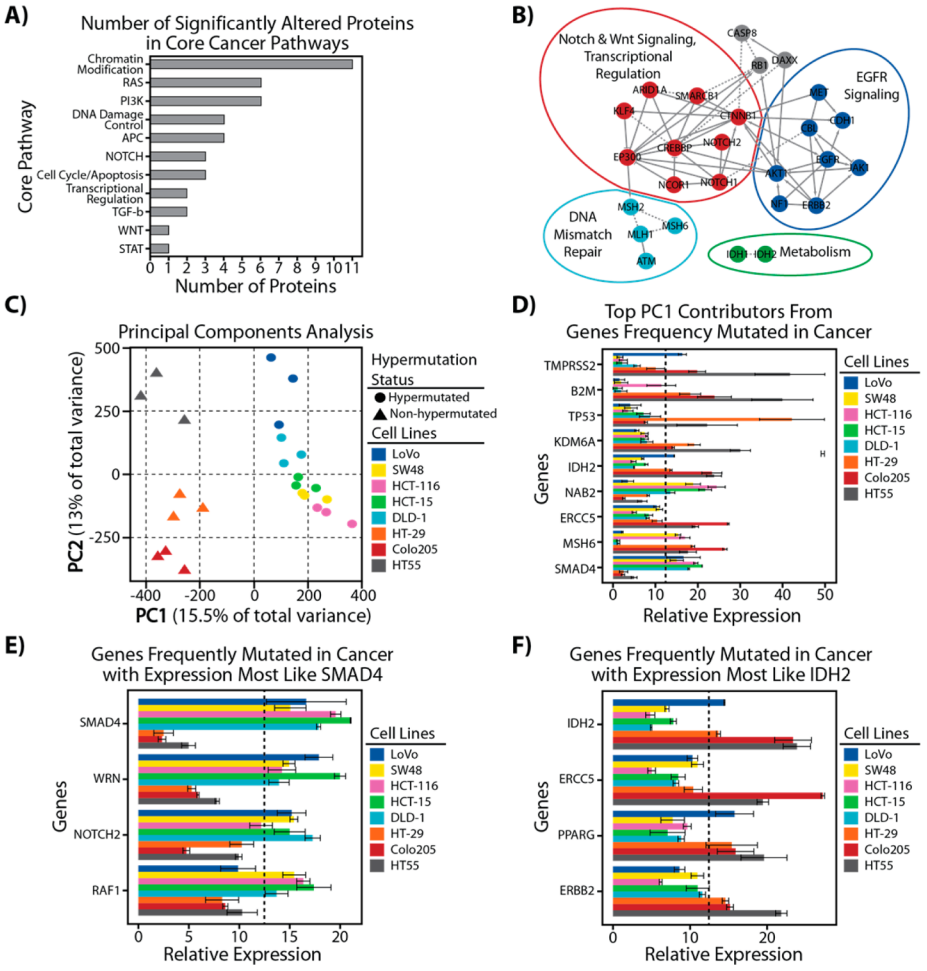


Figure 5. (A) Number of members of the core pathways annotated by Vogelstein et al. showing significantly different expression. (B) The lists of proteins with altered expression were mapped on to the Reactome Pathway Database and clustered into graph modules. (C) Principal component analysis of the quantified proteome shows that PC1 distinguishes hypermutated from nonhypermutated cell lines. (D) The top contributors to PC1 among the Vogelstein gene set. SMAD4 and IDH2 are preferentially expressed in the hypermutated and nonhypermutated lines, respectively. (E,F) The Vogelstein gene set proteins with expression profiles most similar to SMAD4 (E) and IDH2 (F).

Quantitative Proteomics Reveals Cellular Pathways Deregulated in the Cancer Cell Lines. A recent review by Vogelstein and colleagues detailed a set of genes frequently mutated across many cancers and are thus likely to be drivers of the disease.[34,35] The MultiNotch MS3 method detected many differentially expressed proteins in this gene set across several biological processes (Figure 5A). These included chromatin modification, PI3K and RAS signaling, and the cell cycle. Mapping these proteins onto the ReactomeFI network (Figure 5B)[33] revealed four modules that correlate with known colorectal cancer biology. Notably, DNA mismatch repair defects are a primary driver of mutational burden in heritable colorectal cancers.[36] Similarly, signaling downstream of EGFR and WNT are often a critical drivers in many colorectal cancers.[37]

Principal component analysis on all quantified proteins in the entire MultiNotch MS3 data set (Figure 5C)[38] revealed that PC1 clearly separated hypermutated from nonhypermutated cell lines. This division is prognostic for colorectal cancers, as patients with hypermutated tumors fare better.[36] Several of the primary contributors to PC1 are also mutated at a high rate in colorectal cancers (e.g., MSH6 and P53, Figure 5D).

The MultiNotch MS3 data reproduced a recent report that SMAD4 is highly expressed in hypermutated tumors (Figure 5D), correlating with a positive prognosis.[39] That work did not quantify the difference between the two phenotypes, and here we report an average 5.5-fold up-regulation of SMAD4 protein in the hypermutated cell lines.

Using the SMAD4 expression profile as a prototype, we looked for other proteins encoded by the Vogelstein gene set that showed similar expression patterns (Figure 5E). The top hits for nearby profiles were not for the BMP/TGF-Beta signaling pathways but rather for telomere maintenance (WRN), Notch (NOTCH2), and MAPK signaling (RAF1). WRN and NOTCH2 have been previously linked to colorectal cancer.[40,41] In contrast to BRAF, RAF1 (CRAF) has not been linked to colorectal cancer, but these results suggest that it might play a role in hypermutated tumors.

Mutations in IDH1 and IDH2 are frequently found in a number of cancer types.[34,42] Surprisingly, higher IDH2 expression was one of the best predictors of the non-hypermutated phenotype in our cell lines (Figure 5D) despite not being traditionally associated with colorectal cancer. As with SMAD4, we used IDH2 as a prototype to look for similar expression profiles (Figure 5F). Other metabolic proteins, however, were not the closest profiles to IDH2. Rather, excision repair (ERCC5), PPAR gamma signaling (PPARG), and EGFR signaling (ERBB2) were closer matches. Both PPARG and

dx.doi.org/10.1021/ac502040v | Anal. Chem. 2014, 86, 7150−7158

ERBB2 have been well tied to colorectal cancer.[36,43] In contrast, ERCC5 has not been associated with colorectal cancer. The up-regulation suggests a possible role for DNA excision repair in nonhypermutated tumors

DISCUSSION

The combination of quantitative isobaric reagents and the MultiNotch MS3 method facilitated the quantitation of 8 378 proteins across biological triplicates. Taking into account that every quantified protein contains expression information for eight different cells lines, and limiting our scope to only measurements that were made in all three biological replicates, we reproducibly quantified 172 704 protein abundance changes between individual cell lines. To achieve similar breadth of quantitative analysis, including eight samples and three biological replicates, would require 24 separate experiments using a simple duplex quantitative method (e.g., SILAC). In contrast we only need three separate experiments using the TMT 8-plex workflow.

Though isobaric reagents for relative quantitation (e.g., TMT and iTRAQ) always had the potential to increase sample throughput, the actual utility of the tags was questionable due to ratio distortion caused by interfering ions. Though the standard MS3 method successfully countered the detrimental effects of the interfering ions, this method entailed a substantial sensitivity penalty. Herein, we overcome this limitation by coisolating multiple MS2 fragment ions using isolation waveforms with multiple notches (i.e., SPS), thereby converting far more MS1 precursor ions into MS3 TMT reporter ions.

It is remarkable how much of the MS2 signal can be captured in the MS3 precursor population using a MultiNotch MS3 scan. Typically 10 MS2 fragment ions were targeted for synchronous precursor selection, which translated into ∼40% of the MS2 total ion current (Figure 2). Taking into account the fragmentation efficiency of CID, we retained ∼25% of our initial MS1 precursor population in the MS3 precursor population. This is a vast improvement over the traditional MS3 approach, where ∼5% of the initial precursor population was retained in the MS3 precursor population. Coupled together with some additional algorithms for selecting precursors, setting collision energies, and intelligently scaling ion injection times, we see a 10-fold increase in TMT reporter ion intensities with the MultiNotch MS3 approach, relative to the standard MS3 scan type (Figure 3).

On the basis of the success of this work, Thermo Scientific has implemented a MultiNotch MS3 method. This method also captures multiple MS2 fragment ions in the MS3 precursor population using isolation waveforms with multiple frequency notches (i.e., synchronous precursor selection). They have implemented this method on the Orbitrap Fusion mass spectrometer. Preliminary data collected using the Orbitrap Fusion matches with the trends presented in this manuscript, i.e., that MS2-based measurements of TMT ratios are inaccurate due to the coisolation and cofragmentation of interfering ions and that the MultiNotch MS3 method effectively counters these interfering signals (Figure S7 in the Supporting Information). We also observed that improvements in parallel ion processing and active ion management dramatically reduce the scan overhead associated with MS3 analysis. We are currently preparing a manuscript that further details the performance characteristics of this instrument.

ASSOCIATED CONTENT

Supporting Information

Two supplementary tables: Table S1, the protein level quantitation information; Table S2, peptide level quantitation information. The raw data files produced by the mass spectrometer are available upon request. This material is available free of charge via the Internet at http://pubs.acs.org.

AUTHOR INFORMATION

Corresponding Author

*Phone: 1-617-432-3155. E-mail: steven_gygi@hms.harvard.edu.

Author Contributions

§D.P.N. and M.P.J. contributed equally to this work.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

We thank all members of the Gygi lab for many helpful discussions. We thank Vlad Zabrouskov, Ian Jardine, Mike Senko, Philip Remes, and Jesse Canterbury for implementation advice. This work was funded in part by grants from the NIH (Grants GM67945 and HG6673) to S.P.G. G.C.M. was partially funded by an industry-sponsored research project from Thermo Fisher Scientific to S.P.G. M.P.J. and M.W. were supported by NIH Grants R01GM103785 and R01HD073104.

REFERENCES

(1) Aebersold, R.; Mann, M. *Nature* **2003**, *422*, 198−207.
(2) Jiang, H.; English, A. M. *J. Proteome Res.* **2002**, *1*, 345−350.
(3) Ong, S. E.; Blagoev, B.; Kratchmarova, I.; Kristensen, D. B.; Steen, H.; Pandey, A.; Mann, M. *Mol. Cell. Proteomics* **2002**, *1*, 376−386.
(4) Oda, Y.; Huang, K.; Cross, F. R.; Cowburn, D.; Chait, B. T. *Proc. Natl. Acad. Sci. U.S.A.* **1999**, *96*, 6591−6596.
(5) Dayon, L.; Hainard, A.; Licker, V.; Turck, N.; Kuhn, K.; Hochstrasser, D. F.; Burkhard, P. R.; Sanchez, J.-C. *Anal. Chem.* **2008**, *80*, 2921−2931.
(6) Ross, P. L.; Huang, Y. L. N.; Marchese, J. N.; Williamson, B.; Parker, K.; Hattan, S.; Khainovski, N.; Pillai, S.; Dey, S.; Daniels, S.; Purkayastha, S.; Juhasz, P.; Martin, S.; Bartlet-Jones, M.; He, F.; Jacobson, A.; Pappin, D. J. *Mol. Cell. Proteomics* **2004**, *3*, 1154−1169.
(7) Thompson, A.; Schafer, J.; Kuhn, K.; Kienle, S.; Schwarz, J.; Schmidt, G.; Neumann, T.; Hamon, C. *Anal. Chem.* **2003**, *75*, 1895−1904.
(8) Mertins, P.; Udeshi, N. D.; Clauser, K. R.; Mani, D. R.; Patel, J.; Ong, S. E.; Jaffe, J. D.; Carr, S. A. *Mol. Cell. Proteomics* **2012**, *11*, M111 014423.
(9) Bantscheff, M.; Boesche, M.; Eberhard, D.; Matthieson, T.; Sweetman, G.; Kuster, B. *Mol. Cell. Proteomics* **2008**, *7*, 1702−1713.
(10) Karp, N. A.; Huber, W.; Sadowski, P. G.; Charles, P. D.; Hester, S. V.; Lilley, K. S. *Mol. Cell. Proteomics* **2010**, *9*, 1885−1897.
(11) Lu, R.; Markowetz, F.; Unwin, R. D.; Leek, J. T.; Airoldi, E. M.; MacArthur, B. D.; Lachmann, A.; Rozov, R.; Ma'ayan, A.; Boyer, L. A.; Troyanskaya, O. G.; Whetton, A. D.; Lemischka, I. R. *Nature* **2009**, *462*, 358−U126.
(12) Ow, S. Y.; Salim, M.; Noirel, J.; Evans, C.; Rehman, I.; Wright, P. C. *J. Proteome Res.* **2009**, *8*, 5347−5355.
(13) Shirran, S. L.; Botting, C. H. *J. Proteomics* **2010**, *73*, 1391−1403.
(14) Ting, L.; Rad, R.; Gygi, S. P.; Haas, W. *Nat. Methods* **2011**, *8*, 937−940.
(15) Wenger, C. D.; Lee, M. V.; Hebert, A. S.; McAlister, G. C.; Phanstiel, D. H.; Westphall, M. S.; Coon, J. J. *Nat. Methods* **2011**, *8*, 933−935.

(16) Wuhr, M.; Haas, W.; McAlister, G. C.; Peshkin, L.; Rad, R.; Kirschner, M. W.; Gygi, S. P. *Anal. Chem.* **2012**, *84*, 9214−9221.
(17) Villen, J.; Gygi, S. P. *Nat. Protoc.* **2008**, *3*, 1630−1638.
(18) Wessel, D.; Flugge, U. I. *Anal. Biochem.* **1984**, *138*, 141−143.
(19) Chen, L.; Marshall, A. G. *Int. J. Mass Spectrom. Ion Processes* **1987**, *79*, 115−125.
(20) Marshall, A. G.; Wang, T. C. L.; Ricca, T. L. *J. Am. Chem. Soc.* **1985**, *107*, 7893−7897.
(21) McLafferty, F. W.; Stauffer, D. B.; Loh, S. Y.; Williams, E. R. *Anal. Chem.* **1987**, *59*, 2212−2213.
(22) Goeringer, D. E.; Asano, K. G.; McLuckey, S. A.; Hoekman, D.; Stiller, S. W. *Anal. Chem.* **1994**, *66*, 313−318.
(23) Julian, R. K.; Cooks, R. G. *Anal. Chem.* **1993**, *65*, 1827−1833.
(24) Eng, J. K.; McCormack, A. L.; Yates, J. R. *J. Am. Soc. Mass Spectrom.* **1994**, *5*, 976−989.
(25) Beausoleil, S. A.; Villen, J.; Gerber, S. A.; Rush, J.; Gygi, S. P. *Nat. Biotechnol.* **2006**, *24*, 1285−1292.
(26) Elias, J. E.; Gygi, S. P. *Nat. Methods* **2007**, *4*, 207−214.
(27) Huttlin, E. L.; Jedrychowski, M. P.; Elias, J. E.; Goswami, T.; Rad, R.; Beausoleil, S. A.; Villen, J.; Haas, W.; Sowa, M. E.; Gygi, S. P. *Cell* **2010**, *143*, 1174−1189.
(28) Chvatal, V. *Math. Oper. Res.* **1979**, *4*, 233−235.
(29) Makarov, A.; Denisov, E. *J. Am. Soc. Mass Spectrom.* **2009**, *20*, 1486−1495.
(30) Werner, T.; Becher, I.; Sweetman, G.; Doce, C.; Savitski, M. M.; Bantscheff, M. *Anal. Chem.* **2012**, *84*, 7188−7194.
(31) McAlister, G. C.; Huttlin, E. L.; Haas, W.; Ting, L.; Jedrychowski, M. P.; Rogers, J. C.; Kuhn, K.; Pike, I.; Grothe, R. A.; Blethrow, J. D.; Gygi, S. P. *Anal. Chem.* **2012**, *84*, 7469−7478.
(32) Welch, B. L. *Biometrika* **1951**, *38*, 330−336.
(33) Wu, G.; Stein, L. *Genome Biol.* **2012**, *13*, R112.
(34) Garraway, L. A.; Lander, E. S. *Cell* **2013**, *153*, 17−37.
(35) Vogelstein, B.; Papadopoulos, N.; Velculescu, V. E.; Zhou, S.; Diaz, L. A., Jr.; Kinzler, K. W. *Science* **2013**, *339*, 1546−1558.
(36) The Cancer Genome Atlas Network. *Nature* **2012**, *487*, 330−337.
(37) Fearon, E. R. *Annu. Rev. Pathol* **2011**, *6*, 479−507.
(38) Raychaudhuri, S.; Stuart, J. M.; Altman, R. B. *Pac. Symp. Biocomput.* **2000**, 455−466.
(39) Isaksson-Mettavainio, M.; Palmqvist, R.; Dahlin, A. M.; Van Guelpen, B.; Rutegard, J.; Oberg, A.; Henriksson, M. L. *Int. J. Cancer* **2012**, *131*, 779−788.
(40) Agrelo, R.; Cheng, W. H.; Setien, F.; Ropero, S.; Espada, J.; Fraga, M. F.; Herranz, M.; Paz, M. F.; Sanchez-Cespedes, M.; Artiga, M. J.; Guerrero, D.; Castells, A.; von Kobbe, C.; Bohr, V. A.; Esteller, M. *Proc. Natl. Acad. Sci. U.S.A.* **2006**, *103*, 8822−8827.
(41) Kloosterman, W. P.; Hoogstraat, M.; Paling, O.; Tavakoli-Yaraki, M.; Renkens, I.; Vermaat, J. S.; van Roosmalen, M. J.; van Lieshout, S.; Nijman, I. J.; Roessingh, W.; van't Slot, R.; van de Belt, J.; Guryev, V.; Koudijs, M.; Voest, E.; Cuppen, E. *Genome Biol.* **2011**, *12*, R103.
(42) Parsons, D. W.; Jones, S.; Zhang, X.; Lin, J. C.; Leary, R. J.; Angenendt, P.; Mankoo, P.; Carter, H.; Siu, I. M.; Gallia, G. L.; Olivi, A.; McLendon, R.; Rasheed, B. A.; Keir, S.; Nikolskaya, T.; Nikolsky, Y.; Busam, D. A.; Tekleab, H.; Diaz, L. A., Jr.; Hartigan, J.; Smith, D. R.; Strausberg, R. L.; Marie, S. K.; Shinjo, S. M.; Yan, H.; Riggins, G. J.; Bigner, D. D.; Karchin, R.; Papadopoulos, N.; Parmigiani, G.; Vogelstein, B.; Velculescu, V. E.; Kinzler, K. W. *Science* **2008**, *321*, 1807−1812.
(43) Dai, Y.; Wang, W. H. *World J. Gastrointest. Oncol.* **2010**, *2*, 159−164.