This is an open access article published under an ACS AuthorChoice License, which permits copying and redistribution of the article or any adaptations for non-commercial purposes.

Journal of proteome research

Reviews

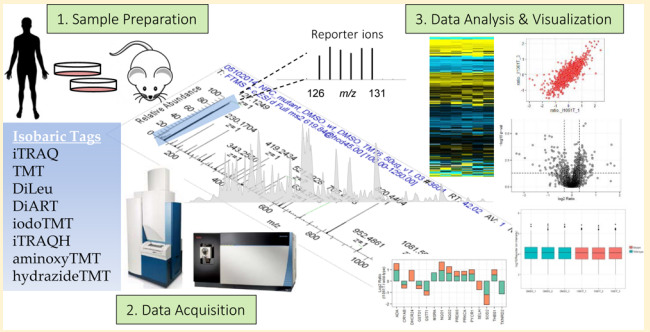
pubs.acs.org/jpr

Isobaric Labeling-Based Relative Quantification in Shotgun Proteomics

Navin Rauniyar and John R. Yates, III*

Department of Chemical Physiology, The Scripps Research Institute, 10550 North Torrey Pines Road, La Jolla, California 92037, United States

ABSTRACT: Mass spectrometry plays a key role in relative quantitative comparisons of proteins in order to understand their functional role in biological systems upon perturbation. In this review, we review studies that examine different aspects of isobaric labeling-based relative quantification for shotgun proteomic analysis. In particular, we focus on different types of isobaric reagents and their reaction chemistry (e.g., amine-, carbonyl-, and sulfhydryl-reactive). Various factors, such as ratio compression, reporter ion dynamic range, and others, cause an underestimation of changes in relative abundance of proteins across samples, undermining the ability of the isobaric labeling approach to be truly quantitative. These factors that affect quantification and the suggested combinations of experimental design and optimal data acquisition methods to increase the precision and accuracy of the measurements will be discussed. Finally, the extended application of isobaric labeling-based approach in hyperplexing strategy, targeted quantification, and phosphopeptide analysis are also examined.



KEYWORDS: *iTRAQ, isobaric tags for relative and absolute quantification, TMT, tandem mass tags, isobaric tags, isobaric labeling, quantitative proteomics, mass spectrometry*

Downloaded via CALIFORNIA INST OF TECHNOLOGY on April 10, 2019 at 20:30:42 (UTC).
See https://pubs.acs.org/sharingguidelines for options on how to legitimately share published articles.

1. INTRODUCTION

Mass spectrometry (MS) is a powerful tool to assess the relative abundance of proteins among biological samples. Numerous methodologies now support relative quantification measurements, providing a routine means to analyze protein expression patterns and post-translational modification states as a function of biological perturbation. One of the most popular methods for relative quantification through MS is stable isotope labeling of proteins in samples prior to analysis. Labeling can be achieved by the application of combinatorial heavy isotopologues of C, H, N, and O and can be introduced in proteins either by metabolic means or through chemical derivatization processes. In vivo metabolic labeling approaches include techniques such as stable isotope labeling in mammals (SILAM),[1] stable isotope labeling by amino acids in cell culture (SILAC),[2] and NeuCode (neutron encoding) SILAC.[3] The in vitro chemical derivatization processes include techniques such as isotope-coded affinity tags (ICAT),[4] dimethyl labeling,[5] isobaric mass tags,[6,7] and others.[8]

With the exception of isobaric mass tags, stable isotope derivatization methods introduce a small mass difference to identical peptides from two or more samples so that they can be distinguished in the MS1 spectrum. The relative-abundance ratio of peptides is experimentally measured by comparing heavy/light peptide pairs, and then protein levels are inferred from statistical evaluation of the peptide ratios. Isobaric tags, on the other hand, use a different concept for peptide quantification. In isobaric labeling-based quantification, each sample is derivatized with a different isotopic variant of an isobaric mass tag from a set, and then the samples are pooled and analyzed simultaneously in MS. Since the tags are isobaric, peptides labeled with isotopic variants of the tag appear as a single composite peak at the same m/z value in an MS1 scan with identical liquid chromatography (LC) retention time. The fragmentation of the modified precursor ion during MS/MS event generates two types of product ions: (a) reporter ion peaks and (b) peptide fragment ion peaks. The quantification is accomplished by directly correlating the relative intensity of reporter ions to that of the peptide selected for MS/MS fragmentation. The fragment ion peaks observed at higher m/z are specific for peptide amino acid sequence and are used for peptide identifications, which are eventually assigned to the proteins that they represent. Since every tryptic peptide can be labeled in an isobaric labeling method, more than one peptide representing the same protein may be identified, thereby increasing the confidence in both the identification and quantification of the protein. This technology has proved to be successful in numerous experimental contexts for comparative analysis upon perturbation. A general workflow of an isobaric labeling experiment is depicted in Figure 1.

Received: August 22, 2014
Published: October 22, 2014

 © 2014 American Chemical Society dx.doi.org/10.1021/pr500880b | J. Proteome Res. 2014, 13, 5293−5309

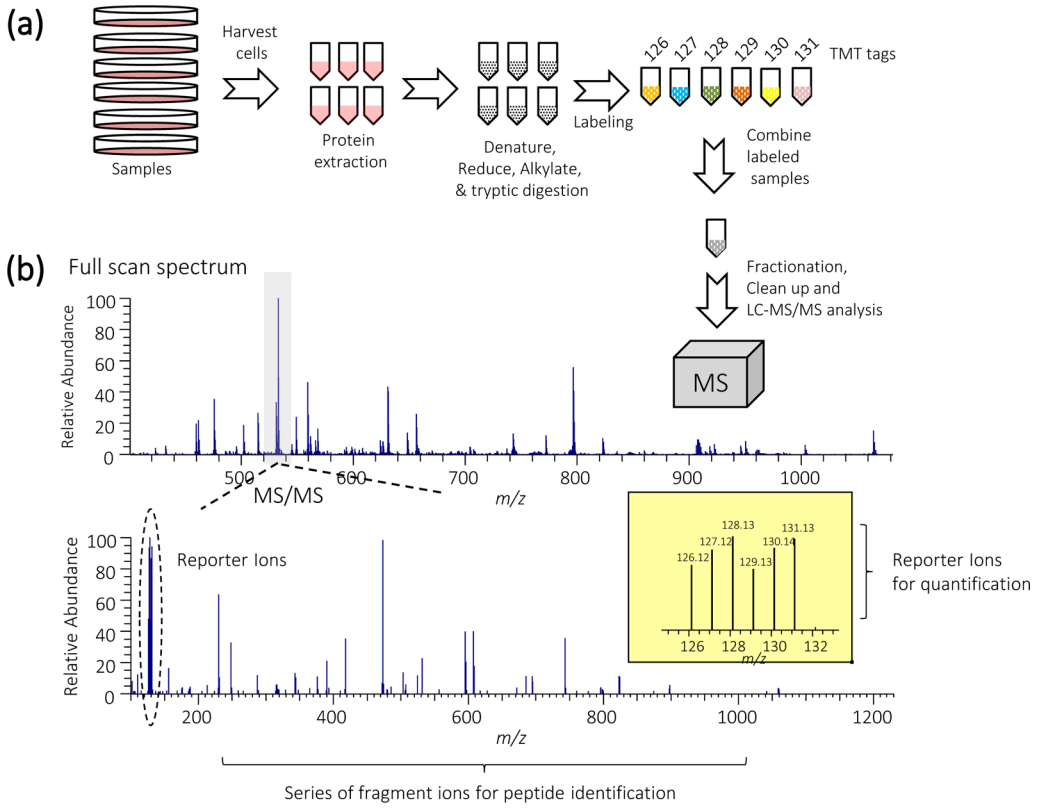


Figure 1. (a) General workflow of an isobaric labeling experiment. The protocol involves extraction of proteins from cells or tissues followed by reduction, alkylation, and digestion. In the case of TMT 6-plex, up to six samples can be labeled with the six isobaric tags of the reagent. Resulting peptides are pooled at equal concentrations before fractionation and clean up. The TMT-labeled samples are analyzed by LC–MS/MS. (b) In an MS1 scan, same-sequence peptides from the different samples appear as a single unresolved additive precursor ion. Following fragmentation of the precursor ion during MS/MS, the six reporter ions appear as distinct masses between m/z 126–131, and the remainder of the sequence-informative b- and y-ions remains as additive isobaric signals. The reporter ion intensity indicates the relative amount of peptide in the mixture that was labeled with the corresponding reagent.

2. ISOBARIC MASS TAGS

Isobaric mass tags include families of stable isotope chemicals that are used for labeling of peptides. They generate relative quantitative information in an isobaric labeling-based quantification strategy. Isobaric mass tags have identical overall mass but vary in terms of the distribution of heavy isotopes around their structure. The most common isobaric tag is amine-reactive, but tags that react with cysteine residues and carbonyl groups in proteins are also available. The amine specificity of the amine-reactive isobaric mass tags makes most peptides in a sample amenable to this labeling strategy. The tags employ *N*-hydroxysuccinimide (NHS) chemistry, and the structure consists of three functional groups: an amine-reactive group and an isotopic reporter group (*N*-methylpiperazine) linked by an isotopic balancer group (carbonyl) for the normalization of the total mass of the tags. The amine-reactive, NHS-ester-activated group reacts with N-terminal amine groups and ε-amine groups of lysine residues to attach the tags to the peptides. The labeling is efficient for all peptides regardless of protein sequence or proteolytic enzyme specificity. The labeling does not occur, however, if the primary amino groups are modified, such as when N-terminal glutamine or glutamic acid forms a ring (pyro-glutamic acid) or if the group is acetylated. The NHS-based isobaric tags may lead to acylation of side chain hydroxyl group of serine, threonine, and tyrosine residues under reaction conditions normally employed for the acylation of primary amines.[9] For successful quantification, labeling should be specific to the targeted residues (N-terminal amine and lysyl ε-amine groups in a peptide) and should proceed to completion. Reversal of peptide O-acylation reactions can be achieved by treatment with hydroxylamine that has no disruptive effect on acyl modifications on primary amines.[9]

The mass normalization group balances the mass difference among the reporter ion groups so that different isotopic variants of the tag have the same mass. The overall mass of reporter and balance components of the molecule are kept constant using differential isotopic enrichment with ^{13}C, ^{15}N, and ^{18}O atoms. The relative intensities of the reporter ion are used to derive quantitative information on the labeled peptides between the samples. Figure 2 shows chemical structure of commercially available isobaric mass tags: tandem mass tag (TMT) and isobaric tags for relative and absolute quantification (iTRAQ).

2.1. TMT and iTRAQ Isobaric Mass Tags

The application of isobaric tags for simultaneous determination of both the identity and relative abundance of peptide pairs was first demonstrated by Thompson et al. in 2003.[6] They synthesized peptides containing a tandem mass tag and showed that this strategy could be used to obtain relative quantification in MS/MS experiment. A year later, Ross et al. published a similar approach using the iTRAQ approach.[7] In this study, they demonstrated for the first time the application of isobaric mass tags with 4-fold multiplexing to identify global protein expression trends in a set of isogenic yeast strains. An 8-plex

dx.doi.org/10.1021/pr500880b | J. Proteome Res. 2014, 13, 5293–5309

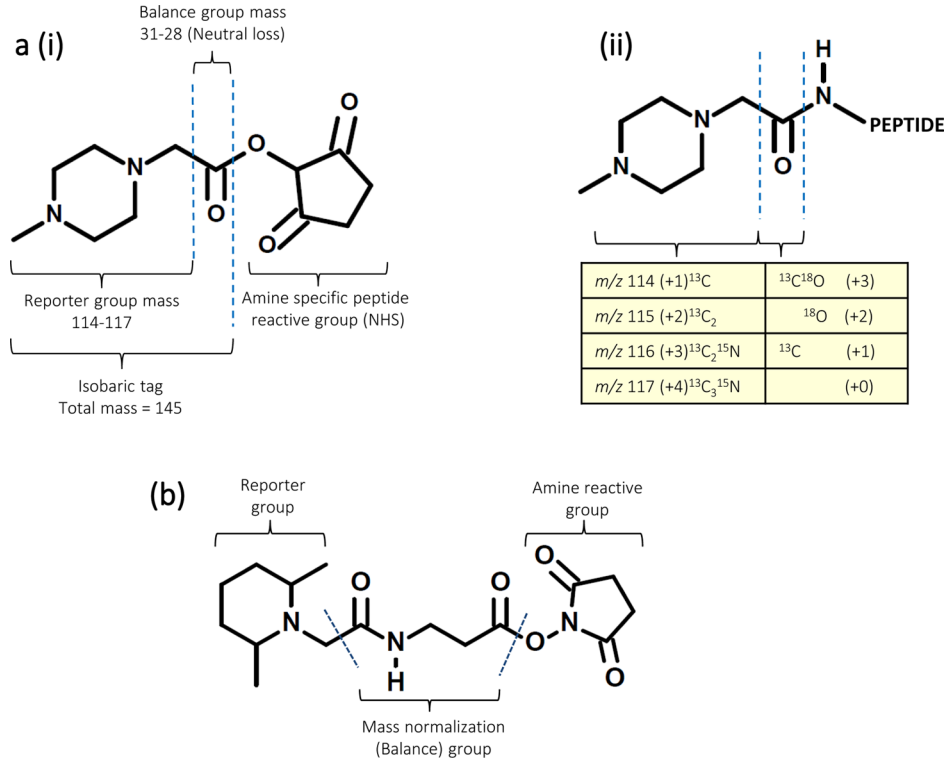


Figure 2. (a) (i) Chemical structure of iTRAQ 4-plex reagent.[7] The complete molecule consists of a reporter group (based on *N*-methylpiperazine), a mass balance group (carbonyl), and a peptide-reactive group (NHS ester). The overall mass of the reporter and balance components of the molecule are kept constant using differential isotopic enrichment with ^{13}C, ^{15}N, and ^{18}O atoms. The reporter group ranges in mass from m/z 114–117, whereas the balance group ranges in mass from 28 to 31 Da, such that the combined mass remains constant (145 Da) for each of the four reagents of the iTRAQ 4-plex set. (ii) The tag reacts with peptide N-terminus or ε-amino group of lysine to form an amide linkage that fragments in a similar fashion to that of backbone peptide bonds when subjected to CID. Following fragmentation of the tag amide bond, the balance (carbonyl) moiety is lost as neutral loss, whereas charge is retained by the reporter group. The number in parentheses in the table indicates the number of enriched centers in each section of the molecule.[7] (b) Chemical structure of a generic TMT reagent showing the three functional groups: an amine-reactive group that labels the N-terminus and ε-amine group of lysine in peptides, a mass normalization (balance) group that balances mass differences from individual reporter ions to ensure the same overall mass of the reagents, and a reporter group that provides the abundance of a peptide upon MS/MS in individual samples being mixed. The blue dashed lines indicate a cleavable linker that enables the release of the reporter ion from the whole tag upon MS/MS. The TMT reagent family consists of TMTzero, TMTduplex, TMT 6-plex, and TMT 10-plex sets, and each of them is based on the same chemical structure.

series of iTRAQ reagent performs similarly and increases throughput of analyses by a factor of 2 when compared to that of the 4-plex approach.[10] A few year later, Dayon et al.[11] showed the increased multiplexing capability of TMT tags and demonstrated its application by using 6-plex TMT reagents in relative quantification of standard protein mixtures at various concentrations. In this study, TMT 6-plex was also used to assess the differential protein abundance in post-mortem cerebrospinal fluid samples after brain injury vs antemortem samples.[11]

Isobaric reagents are commercially available through vendors such as AB Sciex (Framingham, MA, USA) and Thermo Scientific (Rockford, IL, USA). The iTRAQ reagents available from AB Sciex are set of 4-plex and 8-plex mass tags that can be used to label and derive quantitative information on up to four and eight different biological samples simultaneously. The 4-plex iTRAQ reagents have reporter ion masses at m/z 114–117 and a corresponding balancer group added to accommodate the extra isotopes has masses of 28–31 Da such that they sum to about 145 Da. The 8-plex reagents have reporter ion masses at m/z 113–119 and 121 with a balance group ranging from 24–31 Da. Mass 120 is omitted in iTRAQ 8-plex to avoid contamination from phenylalanine immonium ion (m/z 120.08). Thermo Scientific TMT reagents, available as TMTzero, TMT duplex, TMT 6-plex, and TMT 10-plex, share an identical structure with each other but contain different numbers and combinations of ^{13}C and ^{15}N isotopes in the mass reporter region. The identical structure of TMT reagents facilitates efficient transition from method development using TMTzero or TMT duplex to multiplex quantification using TMT 6-plex or TMT 10-plex. The chemical structure of the TMT tag enables the introduction of five heavy isotopes (^{13}C or ^{15}N) in the reporter group and five heavy isotopes (^{13}C or ^{15}N) in the balancer group to provide six isobaric tags (Figure 3a). Each of the six tags of TMT 6-plex has a specific reporter ion that appears at m/z 126, 127, 128, 129, 130, and 131. TMT 10-plex is an expansion of TMT 6-plex generated by combining current TMT 6-plex reagents with four isotope variants of the tag with 6.32 mDa mass differences between ^{15}N and ^{13}C isotopes.[12,13] Even though the mass difference between these reporter ion isotopologues is incredibly small, current generation high-resolution and high mass accuracy analyzers can resolve these ions. The seemingly miniscule difference is sufficient to achieve baseline resolution between the reporter ions when high resolving power is employed (30 K at m/z 400).[13] Figure 3b

dx.doi.org/10.1021/pr500880b | J. Proteome Res. 2014, 13, 5293–5309

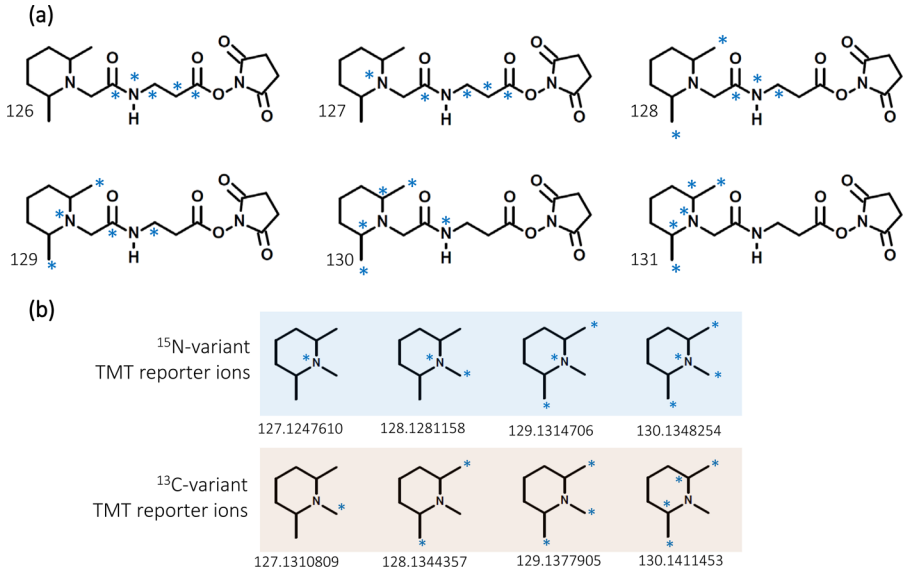


Figure 3. (a) Chemical structure of TMT 6-plex reagents with ^{13}C and ^{15}N heavy isotope positions (blue asterisks). The tags are isobaric, with a different distribution of isotopes between the reporter and mass normalization (balance) groups. (b) The substitution of ^{15}N for ^{13}C to generate new reporter ions that are 6.32 mDa lighter than the original forms used in TMT 6-plex.[12] The TMT 6-plex reagents in combination with four isotope variants of the tag with 6.32 mDa mass differences were used to generate TMT 10-plex reagent.

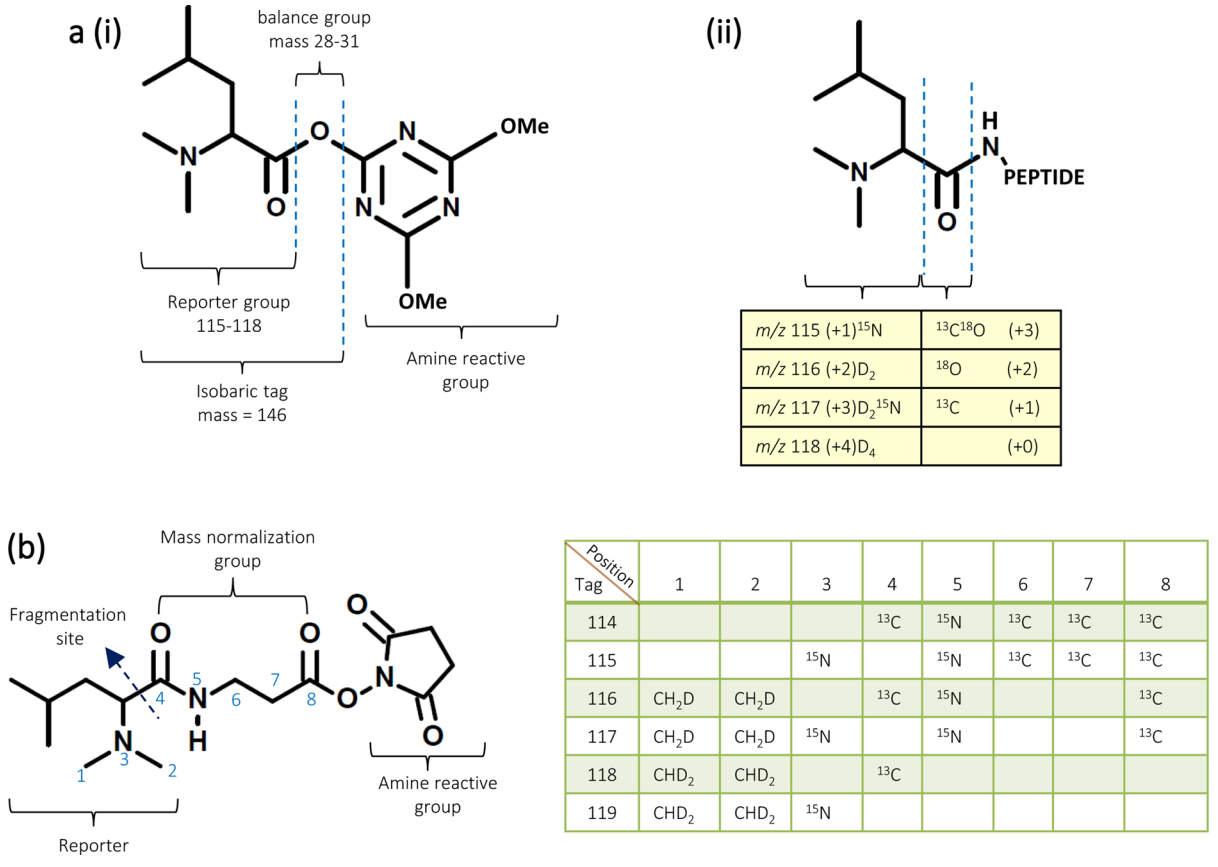
shows the substitution of ^{15}N for ^{13}C to generate new reporter ions that are lighter than the original forms used in TMT 6-plex. In cases where coalescence, fusion of the proximate reporter ion signals into a single measurable entity, phenomenon is observed, the artifact can be completely eliminated by lowering the maximum ion target for MS/MS spectra.[14] This modified setting does not result in any losses in identification depth or quantification quality of proteins.[14] The high-throughput TMT 10-plex reagent enables concurrent MS analysis and relative quantification of up to 10 different samples derived from cells, tissues, or biological fluids. The higher multiplexing potential also facilitates incorporation of replicates, providing additional statistical validation within any given isobaric labeling experiments.[15]

The numbers of identified peptides and proteins in shotgun proteomics experiments have been compared for the three commercially available isobaric mass tags: iTRAQ 4-plex, TMT 6-plex, and iTRAQ 8-plex.[16] Even though the number of identified proteins and peptides was largest with iTRAQ 4-plex, followed by TMT 6-plex, and smallest with iTRAQ 8-plex, the precision on the level of peptide−spectrum matches and protein level dynamic range was similar. The discrepancy in peptide identification observed with different *n*-plex isobaric mass tags was suggested to be due to combination of several factors, such as search algorithms and scoring functions, fragment ions derived from cleavage of the label itself or within the label from precursor ions, or disparate physiochemical properties conferred to the peptides depending on the type of isobaric mass tags used for their derivatization.[16] However, in a study by Pottiez et al. on comparison of quantitative measurements of proteins in human plasma samples by iTRAQ 4-plex versus 8-plex reagents, 8-plex tagging provided more consistent ratios than that with 4-plex without compromising protein identification.[17] The discrepancies in observations from Pichler et al. and Pottiez et al. could be due to different instruments (LTQ Orbitrap versus MALD-TOF/TOF 4800 platform) and search algorithms (Mascot and Proteome Discoverer software versus ProteinPilot 4.0 with Paragon Algorithm) that were used for the data acquisition and analysis.[16,17] Nevertheless, the obvious advantage of 8-plex tagging is that it allows investigation of eight experimental conditions in one analytical experiment. For example, a study of one control and seven experimental conditions can be performed in one 8-plex experiment but would require at least three 4-plex experiments (using the control and up to three experimental samples in each). The three 4-plex experiments would need more instrument time, likely introducing a source of variability, and would be more laborious.

2.2. DiLeu and DiART Isobaric Mass Tags

N,*N*-Dimethyl leucine (DiLeu) is an isobaric tandem mass tagging reagent that uses isotope-encoded dimethylated leucine as reporters and serves as attractive alternative for iTRAQ and TMT.[18] Labeling with DiLeu, however, requires activation of the reagents using 4-(4,6-dimethoxy-1,3,5-triazin-2-yl)-4-methylmorpholinium chloride (DMTMM)/*N*-methylmorpholine (NMM) in *N*,*N*-dimethylformamide (DMF). Each label can be freshly activated before use. The general structure of DiLeu resembles that of other isobaric mass tags, with an amine-reactive group (triazine ester) targeting the N-terminus and ε-amino group of the lysine side chain of a peptide, a balance group, and a reporter group.[18] A mass shift of 145.1 Da is observed for each incorporated label. By using DiLeu isobaric tags, up to four samples can be analyzed simultaneously at greatly reduced cost. The labeling efficiency of DiLeu is comparable to that of the iTRAQ reagents. However, DiLeu-labeled peptides undergo better fragmentation and hence generate higher reporter ion intensities than iTRAQ, thereby offering improved confidence for peptide identification and more reliable quantification.[18] Intense reporter ions (dimethylated leucine a1 ion) at *m*/*z* 115, 116, 117, and 118 are observed for all pooled samples upon MS/MS. Even though deuterium affects the retention time of small- to intermediate-sized peptides in reversed-phase chromatography,[19] the increased polarity of the amine group offsets the small

dx.doi.org/10.1021/pr500880b | J. Proteome Res. 2014, 13, 5293−5309



m/z 115 (+1)^{15}N	$^{13}C^{18}O$	(+3)
m/z 116 (+2)D_2	^{18}O	(+2)
m/z 117 (+3)$D_2{}^{15}$N	^{13}C	(+1)
m/z 118 (+4)D_4		(+0)

Tag \ Position	1	2	3	4	5	6	7	8
114				^{13}C	^{15}N	^{13}C	^{13}C	^{13}C
115			^{15}N		^{15}N	^{13}C	^{13}C	^{13}C
116	CH_2D	CH_2D		^{13}C	^{15}N			^{13}C
117	CH_2D	CH_2D	^{15}N		^{15}N			^{13}C
118	CHD_2	CHD_2		^{13}C				
119	CHD_2	CHD_2	^{15}N					

Figure 4. (a) (i) General structure of dimethyl leucine isobaric (DiLeu) mass tag.[18] Reporter ions range from *m/z* 115−118. (ii) Illustration of formation of new peptide bond at N-terminus or ε-amino group of the lysine side chain and isotope combination of isobaric tags (b) Chemical structure of DiART isobaric reagents.[21] Positions containing heavy stable isotopes are illustrated as numbers in the structure, and the table lists the elemental composition of the corresponding numbers. During MS/MS, the DiART-tagged peptides yield reporter ions ranging from *m/z* 114 to 119.

deuterium number difference in 4-plex DiLeu tags.[18] Figure 4a shows the chemical structure of a DiLeu tag.

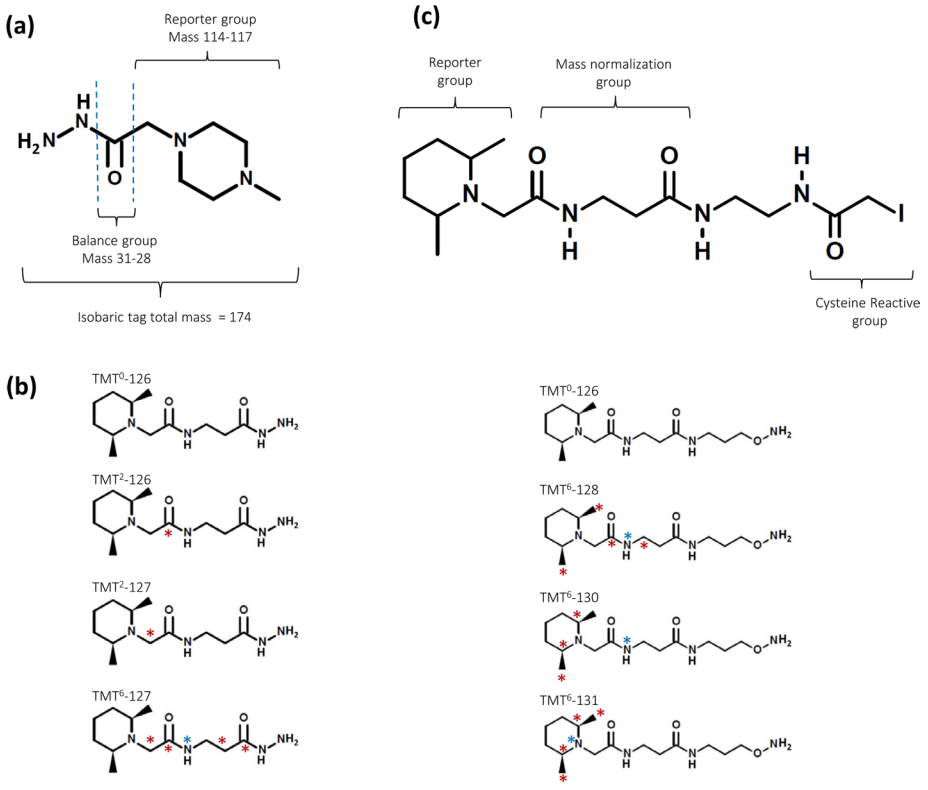
Deuterium isobaric Amine Reactive Tag (DiART) is another alternative to iTRAQ and TMT for isobaric tagging in quantitative proteomics.[20,21] Like iTRAQ or TMT, DiART reagents have three functional groups, an amine-reactive (NHS ester) group for coupling of peptides, a balancer, and a reporter (*N*,*N*′-dimethylleucine) with a *m/z* range of 114−119 (Figure 4b). Up to six samples can be labeled with DiART reagents and analyzed by MS.[21] DiART reagents have high isotope purity; hence, unlike that for iTRAQ, TMT, or DiLeu labeling, isotopic impurities correction is not required during data analysis of DiART-labeled samples.[21] The performances of DiART and iTRAQ, including their fragmentation mechanisms, the number of identified proteins, and the accuracy of quantification, have been compared.[20] Regardless of the peptide sequence, DiART tags generate high-intensity reporter ions compared to those with iTRAQ. Since quantification accuracy is dependent on the intensity of reporter ions,[22] as high-intensity reporter ions are less susceptible to underestimation effect,[23] DiART labeling quantifies more peptides, including low-abundance ones, and with reliable results.[20] While DiLeu uses a nontraditional activation chemistry (DMTMM/NMM in DMF) to label peptides,[18,21] DiART uses the same labeling protocol (NHS-ester-based peptide coupling chemistry) as that of TMT and iTRAQ, making it easy for users to switch between the techniques. However, unlike that for iTRAQ or TMT, DiART-labeled samples cannot be analyzed by the HCD-only instrument method due to easy fragmentation of its reporter ions.[20] Nevertheless, DiART and DiLeu serve as a cost-effective alternatives to TMT and iTRAQ with comparable labeling efficiency. DiART has been shown to be useful in labeling large quantities of proteins from cell lysates prior to TiO_2 enrichment in quantitative phosphoproteomics study.[24]

2.3. Post-translational Modification- and Cysteine-Specific Isobaric Mass Tags

Isobaric labeling-based quantification can also be used for differential quantification of various protein post-translational modifications. Isobaric mass tags are available that are especially designed to measure relative abundance of modified cysteine residues or carbonylated residues in protein.

2.3.1. Isobaric Reagents for Protein Carbonyl and Glycan Modifications. Carbonylation of proteins is caused by the reactive oxygen and carbonyl species generated as byproducts of lipid oxidation during oxidative stress.[25] iTRAQ hydrazide (iTRAQH) is a novel reagent for the selective labeling and relative quantitative analysis of carbonyl groups in proteins.[26] iTRAQH was synthesized from iTRAQ and an excess of hydrazine (Figure 5a). iTRAQH reacts with a carbonylated peptide, resulting in the formation of a hydrazone moiety. Consistent with the isobaric labeling approach, peptides labeled with different isotopic variants of iTRAQH reagents are indistinguishable in MS scan. However, the iTRAQH reporter ions in the low *m/z* region of the MS/MS spectrum provide the relative abundance information on the carbonylated proteins in the samples. The iTRAQH reporter ions have been used as

dx.doi.org/10.1021/pr500880b | J. Proteome Res. 2014, 13, 5293−5309



Hydrazide compounds	Reagent mass (Da)	Induced mass shift (Da)
TMT^{0}-126	256.189926	238.179361
TMT^{2}-126	257.193281	239.182716
TMT^{2}-127	257.193281	239.182716
TMT^{6}-127	261.20038	243.189816

Aminoxy compounds	Reagent mass (Da)	Induced mass shift (Da)
TMT^{0}-126	314.231791	296.221226
TMT^{6}-128	319.242245	301.231680
TMT^{6}-130	319.242245	301.231680
TMT^{6}-131	319.242245	301.231680

Figure 5. (a) General structure of iTRAQ hydrazide (iTRAQH) for relative quantitative analysis of carbonylation sites in proteins.[26] (b) Chemical structure of the carbonyl-reactive glyco-TMT compounds.[27] (Left) Hydrazide reagents; (right) aminoxy reagents. Red asterisks indicate ^{13}C, and blue asterisks, ^{15}N. The table below the compound structures shows isotope codes of the hydrazide- and aminoxy-functionalized glyco-TMT compounds. The carbonyl-reactive tags can be used to quantify a broad range of biologically important molecules including carbohydrates, steroids, or oxidized proteins. (c) Chemical structure of the cysteine-reactive Thermo Scientific iodoTMTzero isobaric mass tag. The iodoTMT reagents are iodoacetyl-activated isobaric mass tags for covalent, irreversible labeling of sulfhydryl (-SH) groups. IodoTMT 6-plex enable measurement of protein and peptide cysteine modifications (S-nitrosylation, oxidation, and disulfide bridges) by multiplex quantitative mass spectrometry. The workflow (not shown in the image) involves derivatization of modified peptides or proteins with the reagent, enrichment of TMT tagged peptide using anti-TMT antibody, and their subsequent elution. The eluent is analyzed by LC−MS/MS to determine the sites of modification and to measure their relative abundance across samples.

targets for precursor selection in precursor ion scan analysis, which allows selective acquisition of MS/MS spectra of only the carbonylated peptides.[26] This eliminates the need for the step involving enrichment of modified peptides prior to LC−MS/MS analysis.

On the basis of similar chemistry as that of iTRAQH, the stable isotope-labeled carbonyl-reactive tandem mass tags (glyco-TMTs) have been used for quantification of N-linked glycans.[27] Glyco-TMT reagents are derivatives of the original TMT compounds but are functionalized with carbonyl-reactive groups involving either hydrazide chemistry or aminoxy chemistry (Figure 5b). A study reported that aminooxy TMTs outperformed their hydrazide counterparts in labeling efficiency and quantification.[27] The glyco-TMT compounds are coded with stable isotopes and enable (i) isobaric quantification in MS/MS spectra and (ii) quantification in MS1 spectra using heavy/light pairs. Isobaric quantification using glyco-TMT can be achieved by using the aminoxy TMT^{6}-128 and TMT^{6}-131 as well as the hydrazide TMT^{2}-126 and TMT^{2}-127 reagents (Figure 5b). The MS1 level quantification is accomplished by the mass difference of 5.0105 Da between the light TMT^{0} and the heavy TMT^{6} reagents (Figure 5b) that is sufficient to separate the isotopic patterns of all commonly existing N-glycans. Glycan quantification using heavy and light glycol-TMTs provided more accurate quantification in MS1 spectra over a broad dynamic range compared with that from quantification based on the reporter ions generated in MS/MS spectra.[27] Glyco-TMTs with aminooxy-functionalized groups are available commercially from Thermo Scientific (Rockford, IL, USA) as aminoxyTMTzero and aminoxyTMT 6-plex reagents. Labeling with aminoxyTMT reagents involves treating intact proteins or proteolytic digests of proteins extracted from biological specimens with PNGase F/A glycosidases to release N-linked glycans. The free glycans are subsequently purified from protein or peptide matrix and labeled at the reducing end with the aminoxyTMT reagents.

dx.doi.org/10.1021/pr500880b | J. Proteome Res. 2014, 13, 5293−5309

The derivatized glycans from individual samples are then combined and analyzed in MS to identify glycoforms in the sample and to quantify reporter ion relative abundance at MS/MS level.

2.3.2. Isobaric Reagents for Tagging Cysteine Residues. Cysteine sulfhydryls in proteins are potential sites of reversible oxidative modification because of the unique redox chemistry of this amino acid.[28] S-Nitrosylation is a redox-based protein post-translational modification that occurs in response to nitric oxide signaling and is involved in a wide range of biological processes.[28] It involves addition of a nitric oxide (NO) group to a specific cysteine residue of a protein to form S-nitrosothiol. An analytical strategy to enrich and relatively quantify cysteine-containing peptides in complex mixtures has been reported.[29] In this strategy, cysteine residues in proteins are first derivatized with *N*-{2-((2-acryloyl)amino)ethyl-1,3 thiazolidine-4-carboxamide) (ATC) followed by labeling with amine-reactive TMT tags for relative quantification of the targeted peptides after the covalent capture. The workflow involves reduction, derivatization of cysteine residue in protein samples with ATC tag, digestion with trypsin, and differential labeling with TMT tags followed by pooling of the labeled samples. The ATC-derivatized cysteinyl peptides are subsequently isolated on an aldehyde resin through the covalent capture technique and analyzed with LC−MS/MS.

The cysteine-reactive TMT reagents allow measurement of S-nitrosylation occupancy and determination of individual protein thiol reactivity.[30,31] However, the disulfide linkage between the (reversible) cysteine-reactive TMT tag and protein thiol group cannot survive the strong reducing conditions normally used during enzymatic digestion for subsequent shotgun proteomic analysis.[32] An irreversible cysteine-reactive TMT reagent containing a sulfhydryl-reactive iodoacetyl reactive group called iodoTMT has been developed.[32] IodoTMT reagents such as iodoTMTzero and iodoTMT 6-plex are commercially available from Thermo Scientific (Rockford, IL, USA). Each isobaric iodoTMT 6-plex reagent within a set has the same nominal mass and consists of a thiol-reactive iodoacetyl functional group for covalent and irreversible labeling of cysteine, a balancer, and a reporter group. The quantification using iodoTMT tags is achieved by inspection of the reporter ion region in MS/MS spectra. The chemical structure of iodoTMTzero reagent is shown in Figure 5c. An iodoTMT switch assay uses an isobaric set of thiol-reactive iodoTMT 6-plex reagents to specifically detect and quantify protein S-nitrosylation.[32,33] The iodoTMT switch assay workflow includes irreversible labeling of S-nitrosylated cysteines followed by enrichment of S-nitrosylated peptides using high-affinity anti-TMT chromatography with competitive elution and finally multiplexed quantification of protein S-nitrosylation via six unique TMT reporter ions.[32,33]

3. BENEFITS OF ISOBARIC LABELING-BASED QUANTIFICATION STRATEGY

Isobaric labeling-based quantification has many advantages compared to other stable isotope labeling techniques, one of which is the ability to perform high-throughput quantification due to sample multiplexing. The ability to combine and analyze several samples within one experiment eliminates the need to compare multiple LC−MS/MS data sets, thereby reducing overall analytical time and run-to-run variation. Moreover, the information replication within LC−MS/MS experimental regimes provides additional statistical validation within any given experiment.[15] This is desirable in an analysis where conventional upregulation and downregulation measurements are not nearly as meaningful as obtaining temporal expression patterns of proteins throughout the experimental condition, such as in studies involving different stages of cell differentiation, comparisons of multiple drug treatments, identifications of protein−drug interactions,[34] measurement of inhibitor dose response, or time course comparisons.[35] When each sample is run separately or with limited multiplexing, as required in label-free, metabolic-labeling and other MS1-based quantification methods, an ion selected for fragmentation on one LC−MS/MS run may not be selected consistently in subsequent runs or spectra of suitable quality may not be acquired. This results in missing observations, affecting identification and quantification. The isobaric labeling strategy, however, is immune to the stochastic nature of data-dependent mass spectrometry because a common precursor ion is fragmented that corresponds to the same peptide species present in all of the labeled samples, yielding quantitative data across samples within an isobaric tagging experiment. Isobaric labeling has been shown to surpass metabolic labeling in quantification precision and reproducibility.[36]

Isobaric labeling exhibits a wide dynamic range in profiling both high- and low-abundance proteins and proteins with wide array of physiological properties.[37] It can be used to identify and quantify proteins across diverse molecular weight and pI ranges, functional categories, and cellular locations.[38,39] The isobaric mass tags do not interfere with peptide fragmentation, and the peptide length distribution profile and amino acid content of the isobarically derivatized peptides are similar to those obtained using other MS-based approaches.[38] In fact, isobaric tags have been reported to improve the efficiency of MS/MS fragmentation and result in increased signal intensities of native peptides in samples of human parotid saliva that, in general, lack the uniform architecture of tryptic cleavage products, e.g., a basic C-terminal amino acid residue.[40]

With an MS1-based quantification approach, the co-elution of light and heavy peptides can compromise sensitivity as the ion current is divided between multiple samples during MS analysis. Occurrence of multiple precursor ion species in the MS1 level can also create redundancy in MS/MS scanning events of the same peptide bearing different labels. This results in undersampling of the proteome. It is reported that up to 50% of MS/MS scans acquired during data acquisition can be redundant.[41] By contrast, labeling of samples by isobaric mass tags does not increase the sample complexity during chromatographic separation and MS analysis because they are isotope-coded molecules with the same chemical structure and molecular weight, thus eluting at the same chromatographic time and with the same peptide mass. In fact, since differentially labeled but identical peptides from multiple samples are efficaciously merged, an improvement in overall signal-to-noise ratios occurs, allowing good-quality MS/MS data to be acquired from low-copy-number proteins.[40,42] Moreover, the sequence informative b- and y-ions in MS/MS spectra also show this summed intensity, which aids sensitivity.[43]

The in vitro labeling procedure used for isobaric labeling-based quantification strategy is highly efficient and enables this method to be applicable to wide variety of samples such as cultured cells, human tissues and biofluids, and tissues from model animals. This technique has been successfully applied to various biological studies, demonstrating its validity and robustness for quantitative MS-based proteomics.[37,42,44−51]

dx.doi.org/10.1021/pr500880b | J. Proteome Res. 2014, 13, 5293−5309

Isobaric labeling, especially iTRAQ has been used in identifying and distinguishing protease-generated neo-N termini from N-termini of mature proteins by performing terminal amine isotopic labeling of substrates (TAILS).[52,53] After tryptic digestion of iTRAQ-labeled protein samples, N-terminal peptide separation is accomplished using a high-molecular-weight dendritic polyglycerol aldehyde polymer that binds internal tryptic and C-terminal peptides that now have N-terminal alpha amines. The unbound iTRAQ-labeled mature N-terminal and neo-N-terminal peptides and naturally blocked (acetylated, cyclized, and methylated) peptides are recovered by ultrafiltration and analyzed by mass spectrometry. The neo-N-terminal peptides specific to the protease of interest appear only in the protease-treated sample and therefore show a high protease/untreated iTRAQ reporter ion intensity ratio, thus differentiating them from trypsin cleavage products that are present in all samples in equal amounts and therefore have expected iTRAQ ratios of 1.[53] The applications of the isobaric labeling strategy have also been extended to studies involving the characterization of post-translational modifications such as phosphorylation[41,54−56] and other modifications (discussed in the section above).

4. INSTRUMENTATION AND DATA ACQUISITION METHODS FOR ISOBARICALLY LABELED SAMPLES

Many different mass spectrometers are capable of analyzing isobarically tagged peptides. Initially, isobaric labeling experiments were carried out on MALDI-TOF/TOF[57,58] and quadrupole time-of-flight (Q-TOF)[7,35] instruments. Quadrupole[59] and TOF instruments are capable of detecting low m/z fragment ions in the region where reporter ions are observed. However, the large ion selection window of the TOF/TOF instrument can result in a relatively high background of chemical noise for the reporter groups, compressing the dynamic range of the ratios significantly.[58] Quadrupole ion trap geometries generally produce suboptimal results because the reporter ions often lie below the stability limit, as dictated by the precursor peptide mass-to-charge ratio and pseudopotential well parameters used for activation (for example, activation q = 0.23).[22] The slow scanning Q-TOF instruments also have less sensitivity for complex mixtures compared to that of linear ion traps.[60]

Isobaric quantification using standard collision-induced dissociation (CID) conditions is not feasible using ion traps. The "$^1/_3$ rule" for ion-trap instruments restricts the analysis of product ions with m/z values less than 25−30% of the precursor ion. This low mass cutoff limitation also applies to hybrid instruments containing an ion-trap for fragmentation, such as the LTQ-FT and the LTQ-Orbitrap.[61] This limitation can, in principle, be overcome by pulsed Q dissociation (PQD).[62] PQD in the ion trap facilitates detection of low m/z reporter ions, bridging the gap between the linear ion trap with PQD and a quadrupole TOF instrument.[60] However, unlike conventional CID spectra, typical PQD spectra are dominated by the unfragmented precursor ion, indicating poor fragmentation efficiency and thus limiting its practical utility for quantification of peptides by iTRAQ or TMT approaches. Nevertheless, Bantscheff et al. and Griffin et al. have shown that by carefully optimizing instrument parameters such as collision energy, activation q, delay time, ion isolation width, number of microscans, repeat count, and number of trapped ions, low m/z fragment ion intensities can be generated that enable accurate peptide quantification.[60,63] A combined CID-PQD scan strategy exploits CID for efficient peptide identification and PQD for quantification.[49,64]

The development of higher energy collision-induced dissociation (HCD) in the LTQ-Orbitrap has also overcome the $^1/_3$ rule limitation. In an ion trap CID is a resonance-based process, whereas HCD is a beam-type CID event that results in a different fragmentation pattern. During HCD, ions are accelerated as they leave the C-trap and then are fragmented in the nitrogen-filled collision cell. The resulting fragments are returned to the C-trap and detected in the Orbitrap mass analyzer. This fragmentation technique allows analysis of the low m/z region of reporter ions in the Orbitrap mass analyzer since there is no mass cutoff for the multipole.[65] HCD enables efficient reporter ion generation with high mass accuracy detection, but, in general, it suffers from poor peptide sequence-ion recovery compared to that of the classical ion trap CID analysis. The combined use of CID and HCD for efficient identification and relative quantification of proteins with isobaric tags has been demonstrated.[61,66] In this dual-fragmentation method, HCD is used to derive the accurate quantitative information from the reporter ions, whereas CID provides identification of the corresponding peptides. This method alternates MS/MS spectra generated by CID fragmentation with MS/MS spectra obtained from the same precursor ion by HCD fragmentation. Since CID in the ion trap occurs in parallel to acquisition of HCD MS/MS spectra in the Orbitrap, the analysis duty cycle is unaffected. CID and HCD spectra are subsequently combined by merging the peptide sequence-ion m/z range from CID spectra and the reporter ion m/z range from HCD spectra. It should be noted that the extracted intensity values of the reporter ions from each HCD spectrum should be normalized to low ion counts when merging with the respective CID data, otherwise peptide scores can be significantly reduced.[61] The CID-HCD method was shown to be superior to HCD alone in terms of sensitivity and ability to identify proteins in complex mixtures.[61] However, a recent study has shown that with fine-tuning of the normalized collision energy values on Orbitrap Velos instruments, an HCD-only method can perform better than a CID-HCD dual-fragmentation method.[67] This is due to the implementation of the new HCD cell with an axial electric field to push the fragment ions into the C-trap and mounted on Orbitrap XL ETD and Orbitrap Velos instruments that allows an improvement in the analytical precision of the acquired reporter ions.[68] In addition, the redundancy in precursor selection in the dual CID-HCD method compared to that for the stand alone HCD method can result in a reduced number of total peptide and protein identifications.[67] The use of a stepped HCD scheme in Q Exactive instruments has been shown to enhance the intensity of reporter ions without adversely affecting peptide identifications.[69]

Another method for analyzing isobaric labeled samples is to use triple-stage mass spectrometry (MS3) in a hybrid ion trap-Orbitrap platform.[70] In this approach, a peptide precursor ion is isolated and fragmented with CID-MS/MS to generate a plurality of first-generation product ion species comprising different respective m/z ratios. The most intense product ion in MS/MS scan is then selected for HCD-MS3, yielding quantitative data. This method provides an experimental solution to remove interference, thus eliminating the ratio distortion problem (discussed in the next section). A variant of this method referred to as Multinotch MS3[71,72] involves selecting and co-isolating two or more of the first-generation

dx.doi.org/10.1021/pr500880b | J. Proteome Res. 2014, 13, 5293−5309

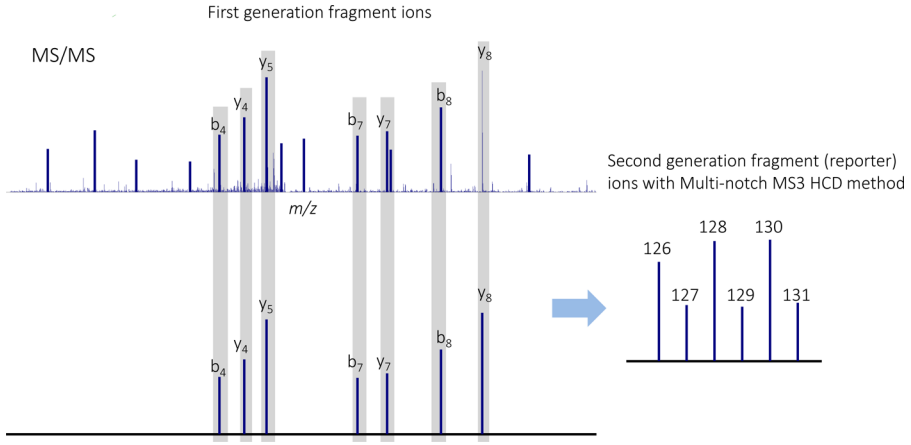


Figure 6. Multinotch MS3 involves selecting and co-isolating multiple MS/MS product ion and fragmenting them to generate a plurality of second-generation fragment ion species including released isobaric tags.[71,72] The method increases the sensitivity and quantitative accuracy achieved by isobaric labeling-based quantification approach.

product-ion species and fragmenting them to generate a plurality of second-generation fragment ion species including released isobaric tags (Figure 6). The Multinotch MS3 method significantly improves quantitative accuracy and increases the sensitivity of the MS experiment up to *n*-fold, where *n* is the number of MS fragments selected and simultaneously isolated.[71]

5. FACTORS AFFECTING QUANTIFICATION BY ISOBARIC LABELING: TECHNICAL AND BIOINFORMATICS ISSUES

The ratios of the intensity of the reporter ions reflect the relative abundance of the peptides from which they are derived. The integration of the relative quantification data for the peptides allows elucidation of relative protein expression levels. This section discusses the various aspects of data analysis in isobaric labeling-based quantification.

5.1. Evaluation of Labeling Efficiency and Isotope Impurity Correction

Isobaric labeling is usually very efficient; however, when primary amino groups are present elsewhere in the sample, they may interfere with the labeling reaction since they can react with the amine-reactive isobaric mass tags. Hence, proper sample preparation is imperative for the success of an isobaric labeling-based quantification technique and includes either avoiding the use of primary amine-containing buffers such as Tris and ammonium bicarbonate or performing sample cleanup prior to the isobaric labeling reaction.[73] To improve detection limits and achieve a reliable estimate of quantification, it is recommended that the labeling efficiency be determined for each isobaric labeling experiment. The labeling efficiency can be ascertained by searching the data separately against protein database using TMT and iTRAQ modifications as variable instead of fixed modifications. Using these parameters, both labeled and unlabeled peptides can be identified and used to calculate labeling efficiency, which is defined as the percent of labeled peptides among all identified peptides. The labeling efficiency can be estimated as

$$100\% \times (n_{ti} + n_{ki})/(n_{tt} + n_{kt})$$

where n_{ti} and n_{ki} are the number of isobaric tag-labeled N-termini and lysine residues, respectively, and, n_{tt} and n_{kt} are the total number of peptide N-termini and lysine residues, respectively.[74] Additionally, due to isotopic contamination in isobaric mass tags, the peaks for each reporter ion will have some contribution from adjacent reporter ions. Hence, prior to data analysis, each of the reporter ion peaks must be corrected to account for isotopic overlap (values reported in the manufacturer's instruction sheet) in order to achieve accurate quantification. The uncorrected data will appear distorted and confound the observed change in protein expression levels.[23] A detailed procedure to calculate true peak areas that account for overlapping isotopic contributions using the reagent purity values provided by the manufacturer is described elsewhere.[75]

5.2. Ratio Compression and Its Correction

In isobaric labeling-based experiments, accurate ratios can be determined only when a single precursor ion is selected for fragmentation during an MS/MS scan event. It has been observed that the presence of co-eluting peptides within the isolation window used for the selection and subsequent fragmentation of individual peptide ions typically results in an underestimation or compression of actual protein abundance differences in the analyzed samples.[23,76,77] This effect is ubiquitous and not dependent only on the instrument used to acquire the data.[77] The compression in relative abundance is based on the assumption that the vast majority of proteins in biological studies do not change significantly; therefore, when the peptides from these proteins co-fragment, the reporter ion intensity ratios generated will be less pronounced in terms of fold changes. Precursor ions of similar intensities can produce reporter ions that span over 2 orders of magnitude in intensity.[41] This means that very low intensity background ions can significantly contribute to reporter ion signals when they get co-fragmented with a selected precursor ion. Additionally, if the coeluting peptides display a nonequimolar distribution of reporter ions, then the net effect of this co-selection is the unpredictable and context-specific distortion of reporter ion intensities.[78] In addition to the distortion in quantification accuracy due to coselection phenomena, the source of quantification error can also be due to presence and interference from artifactual spectral peaks. The reporter ion region in Orbitrap HCD MS/MS spectra contains many signals that are nearly isobaric with reporter ions generated from isobaric mass tags. These signals do not correspond to any plausible chemical compositions and may, in part, be attributed to artifacts related to amplifying and processing the transient

dx.doi.org/10.1021/pr500880b | J. Proteome Res. 2014, 13, 5293−5309

signal of the Orbitrap.[79] Depending on the mass tolerance used for picking the reporter ion signals, the presence of these nonreporter ion signals may distort the quantification results.

Peptide abundance ratios are calculated by combining data from multiple fractions across MS runs and then averaging across peptides to give an abundance ratio for each parent protein. The measured relative abundance can be influenced by the separation (e.g., SCX) stage in which the MS/MS was acquired, a phenomenon termed as fraction effect.[77] Fraction effect for a given peptide is defined as a significant dependence between the measured ratio and the fraction in which the reading was taken. The error within a fraction group for a peptide is smaller than the error between fraction groups and arises from the additional variance from the repeated SCX separation stage.[77] The observation of fraction effect could be due to differences in a peptide's concentration across fractions that contribute to variability in precursor ion intensity measures and subsequent reporter ion peak areas.[80] In addition to fraction effect, the measured ratio is also dependent on the precursor ion (i.e., peptide) used to characterize a protein.[77,81] The measurement error within a peptide group for a protein was found to be smaller than the error between peptide groups.[77] This phenomenon is termed peptide effect. The difference in quantitative value from one peptide to another, even though belonging to the same protein, might result from factors such as post-translational modifications and/or splice variants,[80] tryptic digestion artifacts, peptide recovery, and stability.[81] Other factors of peptide effect include noise peaks with high signal-to-noise in the reporter ion region,[82] sequence of the peptide used for quantification and the possibility of interference from the immonium ion signals in the reporter ion region,[23,56] various charge states of the same peptide, and the number of isobaric tags per peptide.[73]

Since interference due to coisolation is dependent on sample complexity and the number of co-eluting peptides, the ratio compression can be partly alleviated by better fractionation of complex biological samples at the protein or peptide level.[83] Ratio compression was observed to be smaller for enriched phosphoproteome samples compared to that for whole proteome samples due to their overall lower sample complexity.[41] Another approach involves using an optimized (narrow) MS/MS isolation width setting so that fewer contaminant ions are present during precursor ion activation.[76] The high mass resolving power ($m/\Delta m$ > 15 000) in the reporter ion region also minimizes interference from potential contaminant species that may confound quantification data.[22] Delaying peptide selection and fragmentation until the apex of the chromatographic peak during LC−MS/MS analysis has been shown to reduce co-fragmentation by 2-fold.[76] With the delayed fragmentation approach, peptides were fragmented with 2.8-fold better signal-to-noise ratios, significantly improving the quantification.[76] A targeted mass spectrometric data acquisition methodology with reporter ion-based quantification has been shown to be useful in applications where it is essential to reidentify and requantify a defined set of target proteins in a complex mixture.[84] The gas-phase purification[85] and MS3[70,71] methods also eliminate interfering ions in complex mixtures. In Q-TOF instruments, ion mobility (IM) separations have the potential to mitigate quantitative inaccuracies caused by isobaric interference since IM-MS has the ability to separate ions based on charge, m/z, and collision cross section (shape and size).[86]

The ratio correction can also be achieved by various computational approaches post data acquisition. One of the strategies is to use an algorithm that corrects experimental ratios on the basis of determined peptide interference levels.[87] In this method, the measurement for spectrum purity in survey spectra (signal-to-interference measure) was used to improve the accuracy of protein quantification. Signal-to-interference at the time of an MS/MS event is calculated by dividing precursor abundance by the sum of all ion signals observed within the isolation window.[76,84] Consequently, values close to one indicate little and values close to zero indicate a high degree of interference caused by co-eluting components. Other informatics approaches include the intensity-based weighed average technique,[88] variance-stabilizing normalization,[77] and robust statistic-based metric called redescending M-estimator.[89] The interference from non-TMT signals can be eliminated by mass difference processing in which TMT reporter ions in HCD spectra are identified via accurate mass differences between TMT reporter ions present within the same tandem mass spectrum instead of applying fixed mass error tolerances for all tandem mass spectra.[90] This process leads to unambiguous reporter ion identification and eliminates all non-TMT ions from the spectra. Zhang et al. developed an error model that relates the variance of measured ratios to observed reporter ion intensity and provides a p value, q value, and confidence interval for every peptide identified.[22] The identification and exclusion of outlier data, with Grubb's and Rosner's tests, that alter or inappropriately skew the average observed expression ratios has shown to result in a more statistically robust estimation of relative protein abundance.[82] The ability to consider outlier data, however, can occur only for proteins in which there are more than three MS/MS measurements of protein expression.[82] In summary, even though all of the suggested strategies have merit, some techniques only partially remove the problem, and others come with decreased throughput or utilize specialized mass spectrometric instrumentation.

5.3. Reporter Ion Intensity Dynamic Range

Isobaric labeling-based quantification accuracy is also influenced by reporter ion signal intensity and may result in either an underestimation or overestimation of quantification ratio if the signal intensity is outside the detector's saturation point.[91,92] The reporter ions intensities will range between two extremes: the maximum intensity, which corresponds to saturation, and the minimum intensity, which corresponds to the lowest intensity detected. This range is known as the detection limits.[89] However, not all reporter ion intensity peaks will lead to accurate relative quantification. The peak intensities of high-abundance peptide ions may be underestimated by a saturation effect of the detector, which is instrument-dependent.[92] Nevertheless, high-intensity peptides convey more reliable quantitative information about the protein.[23] Larger variances of peptide ratios have been observed for reporter ions of lower intensity[93] because the noise associated with low-intensity reporters constitutes a major handicap in determining the statistical significance of the differential expression of a protein.[23] Therefore, peptides with higher reporter ion intensities should be given higher weight when used to calculate a protein's abundance.[36] Reporter ion signal intensity can be increased by increasing the MS/MS acquisition duration; however, this comes at the expense of decreased sampling, resulting in fewer protein identifications.[93] It is

dx.doi.org/10.1021/pr500880b | J. Proteome Res. 2014, 13, 5293−5309

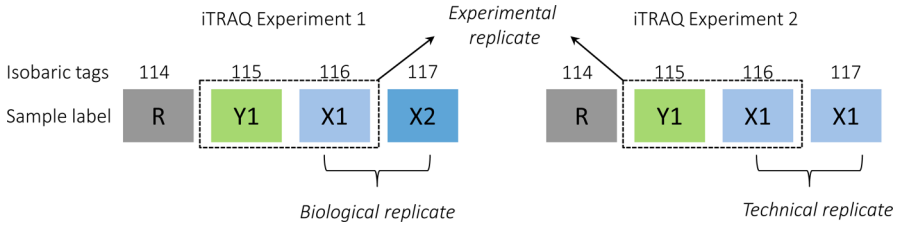


Figure 7. An example defining the relation among technical, experimental, and biological replicates in isobaric labeling (iTRAQ 4-plex in this example) experiments.[97] A biological replicate has two distinct biological samples (X1 and X2) from the same condition in an iTRAQ set, whereas a technical replicate has two identical samples (X1 and X1) from the same biological source in an iTRAQ set. An experimental replicate is the repetitive analysis of the setup to assess the variation of the identical sample in two different iTRAQ sets (Y1 and X1 in experiment 1 versus Y1 and X1 in experiment 2). R refers to a reference sample that can be an individual sample or a pooled sample and allows cross-set comparison.

therefore important to estimate the quantification limits of the instrument and the method used in order to assess the reliability of the obtained quantification measurement. This can be achieved by spiking samples with known quantities of reference proteins prior to analysis and confirming the expected protein ratio from the measured reporter ion intensity ratios.[89]

5.4. Effect of Unique and Shared Peptides in Inferring Protein Ratios

In isobaric labeling, peptide ratios are usually compiled to infer protein ratios. Significant quantification errors arise if a quantified peptide is not unique to its corresponding protein.[92] Hence, relative quantification based on shared peptides (i.e., peptides that match multiple proteins or protein isoforms) due to sequence homology should be interpreted with caution.[94] For a distinct peptide, its relative abundance ratio is a direct measure of the abundance ratio of its corresponding protein. In contrast, the relative abundance ratio of a shared peptide is a weighted average of the abundance ratios of all its corresponding proteins, with the weighting factors being determined by the absolute abundance of those proteins in the samples.[94]

Even though isobaric quantification is not dependent on the total number of spectra matching to each protein, a high number of relative abundance ratios obtained from multiple peptide/spectra increase the confidence in the observed protein ratios.[89] Both intact protein mass and abundance level influence the reliability of the quantification results since highly abundant proteins generate a larger number of peptides per protein.[95] More data, whether from multiple observations per protein or from increasing replication, increases the detection of real signal and reduces false positives.[95] Quantitative information derived from proteins identified with a single peptide lacks variance measurements. The identification of so-called one-hit wonders should be filtered intelligently based on the goal of the study.[96] These proteins deserve special attention if isobaric labeling is used as a screening tool since potentially important biological information or novel biomarkers may be discarded before they are even considered.

5.5. Estimation of Protein Fold Changes

Fold change has been shown to be a function of protein mass and abundance, with small, low-abundance proteins showing the largest variance.[95] A protein is considered to be differentially regulated if the measured fold change exceeds a certain threshold. The actual protein expression level is normally distorted by many factors, with biological variation being the most significant and which ultimately increases the cutoff point.[97] The cutoff point that defines significant differential protein regulation upon perturbation can be estimated by including sample replicates in the experiment.[73,98] The replicate samples can be technical, experimental, or biological. According to Gan et al.,[97] the definition of replicates in terms of isobaric labeling is as follows: a technical replicate will have two identical samples from the same biological source in an isobaric experiment set, whereas a biological replicate will have two distinct biological samples from same condition in an experiment set. The experimental replicate is the actual isobaric experiment replicate, the repetition of the same samples in two or more experimental sets, and they must have the same reference point or control. An illustration of the relationship among technical, experimental, and biological replicates in isobaric labeling experiments is depicted in Figure 7.

Typically, a technical replicate assesses possible errors contributed from sample preparation, also commonly known as the sample variance. Biological replicates are used to examine the variation of random biological effects. Biological variation is protein-, patient-, and disease-dependent.[99] An experimental replicate compares the variation of an identical sample in two different isobaric experiment sets. During analysis of replicate samples, the theoretical relative quantification ratio should be 1:1;[99] however, due to associated variations, the observed relative protein ratios might deviate from the theoretical value. The threshold should be chosen such that it encompasses the majority of technical and biological variation among the replicates. Since in isobaric labeling multiple samples are combined and run together, good quantification precision is observed. Hence, the ratio cutoff applied for significant protein change via the isobaric labeling-based quantification approach is lower than the cutoff applied for the label-free quantification approach;[100] however, the researcher will need to assess whether such a change is biologically significant.

5.6. Comparison of Multiple Isobaric Labeling Experiments

For comparing biological replicates with isobaric labeling in multiple experimental designs, it is recommended to include a reference sample in each experimental setup. The common reference sample among experiments will allow for cross-set comparison. This can be accomplished by first comparing protein ratios of each sample against its reference within individual experiments and then extending the information among multiple experiments. The reference can be an individual sample or a pooled sample prepared by mixing small aliquots of equal amounts of protein from different individual samples.[101] The composition of reference samples does not contribute to missing quantitative values, hence pooling to form a reference sample does not negatively impact the ability to quantitate peptides from comparative individual samples.[93,99] The random biological variation in a pooled

dx.doi.org/10.1021/pr500880b | J. Proteome Res. 2014, 13, 5293−5309

sample is generally lower, as the biological variation can be normalized by *n* samples before being introduced into the experiment.[97] Pooling provides a representative proteome of all of the samples that are detected in comparative samples and is needed for reliable quantification. It also provides sufficient reference material that can be used in many experiments. Herbrich et al., however, have shown that using a masterpool is counterproductive since the latter is also subjected to experimental noise and can result in highly variable estimates when ratios are calculated.[102] According to their study, more precise estimates of protein relative abundance can be obtained by using the available biological data.[102] When a reference sample is used, consistency of the reference is necessary throughout the entire experiment, otherwise even small changes to the reference sample are sufficient to alter the proteins that are reported as differentially expressed.[93]

Regardless of using an individual sample or a pooled sample as a reference, before employing isobaric quantification results for follow-up studies, it is imperative to determine that the data was normalized adequately and the shortlisted protein targets hold merit. Improper normalization might remove some of the biological effects, resulting in attenuated estimates of the protein fold change. Like any other quantification technique, isobaric labeling-based quantification is also biased toward identifying and quantifying a larger percentage of the more abundant proteins, such as ribosomal proteins, heat shock proteins, cytoskeletal proteins, transcription factors, and many others, and often with multiple peptides.[103,104] This is mainly due to the fact that their precursor ions have higher signal intensity. The greater signal intensity increases the likelihood that a given peptide will be selected for fragmentation during LC−MS/MS analysis. Since, in most cases, the expression levels of these house-keeping proteins remain unperturbed in related cell types or growth conditions, they can be used as an effective means to determine the reliability of data normalization.[104] Normalizing by total intensity is not appropriate when the amount of protein is different in the different quantitative sample such as samples that are enriched for certain proteins by pull-down experiments.

6. EXTENDED APPLICATIONS OF ISOBARIC LABELING-BASED QUANTIFICATION STRATEGY

Isobaric labeling experiments can be used for phosphopeptide quantification, and, in cases where the number of samples exceeds the number of isobaric mass tags available for labeling, the throughput can be increased by a hyperplexing method. Isobaric mass tags can also be used for targeted quantification.

6.1. Phosphopeptide Quantification Using Isobaric Mass Tags

Amine-reactive isobaric mass tags have successfully been used in the quantification of post-translational modification such as phosphorylation. Phosphopeptides exist in substoichiometric quantities, and because of the high background of non-phosphorylated peptides in a proteome digest, enrichment of phosphorylated peptides is necessary prior to introduction into the MS. Phosphopeptide enrichment can be performed on isobaric-labeled peptides,[24,54,55] or the phosphorylated peptide can be labeled upon enrichment.[105] Labeling before enrichment minimizes analytical variations caused by further sample manipulation of individual samples during enrichment, whereas labeling after enrichment might improve the yield of the labeling since the nonphosphorylated peptides would otherwise compete for isobaric reagent and interfere with the complete labeling of phosphopeptides.

During the CID-HCD dual method for the quantification and identification of isobarically tagged phosphopeptides, CID with detection in the linear ion trap provides better sensitivity and can be an advantage for low-abundance precursors such as phosphopeptides. However, quantitative information from the low mass region of subsequent HCD scans may not be available for all such CID scans since HCD scans requires higher ion counts.[41] Linke et al. and Wu et al. have examined the optimal fragmentation conditions using the CID-HCD method for iTRAQ-labeled synthetic phosphopeptides in a complex phosphopeptide mix[106] and phosphopeptides enriched from cells.[105] During CID-MS/MS, the spectra derived from phosphoserine- and phosphothreonine-containing peptides show facile fragmentation of the phosphate group and dominance of neutral phosphate losses from the precursor ions.[107] The neutral loss in HCD-MS/MS is much lower, and the sequence-specific fragments are significantly more abundant. With the increasing charge state of the precursor ions, the neutral loss in HCD-MS/MS becomes insignificant and is surpassed by the amide-bond cleavage.[105] However, the study of the effect of normalized collision energy on the HCD-MS/MS fragments and reporter ion abundances shows that the HCD identified phosphopeptides and the HCD spectra with reporter ion information are strongly dependent on precursor charge state.[105] The 2+ charged precursors are more sensitive to the applied normalized collision energy values than the 3+ charged precursor ions in HCD experiments.[106] Thingholm et al. have shown that derivatization with isobaric mass tags significantly increases the average ion charge state of phosphopeptides compared to that of nonlabeled peptides, resulting in a considerable reduction in the number of identified phosphopeptides.[108] Interestingly, it was demonstrated that adding a perpendicular flow of ammonia vapor between the needle and the MS orifice in LC−MS/MS analyses reduced the average charge state of isobaric labeled peptides and resulted in an increase in peptide identification. Thus, the application of isobaric labeling strategies for quantitative phosphopeptide analysis requires simultaneous monitoring of peptide backbone dissociation, loss of phosphoryl group, and the generation of reporter ions.

6.2. Hyperplexing with Isobaric Mass Tags

With existing isobaric mass tags, the maximum number of samples that can be combined and analyzed in a single LC−MS/MS experiment is eight in the case of iTRAQ and 10 with TMT. An effort to increase the multiplexing capacity by the combined use of metabolic and isobaric labeling has been demonstrated.[109] In this strategy, the mass separation of co-eluting intact peptides with the same sequence in an MS1 scan achieved by duplex (heavy and light) or triplex SILAC labeling was exploited to allow for the simultaneous quantification of multiple sets of TMT 6-plex isobaric labels in a single run. Using a 3 × 6 hyperplexing experiment that enables simultaneous quantification of 18 samples, yeast response to the immunosuppressant drug rapamycin, which inhibits the kinase target of rapamycin (TOR), was monitored by measuring the changes in their protein abundance.[109] In this study, three separate cultures of yeast cells grown in light, medium, or heavy SILAC culture medium were treated with 200 mM rapamycin, and samples were removed at 0, 30, 60, 120, and 180 min. A single 120 min sample was taken from

dx.doi.org/10.1021/pr500880b | J. Proteome Res. 2014, 13, 5293−5309

parallel cultures treated with DMSO. Equal amounts of peptides from each sample were labeled with 6-plex TMT reagents, mixed, and separated by SCX before LC−MS/MS analysis. The increased multiplexing capacity enabled analyses of multiple biological replicates of a time-course study in the same run, providing the statistical power required to identify significant trends. The hyperplexing technique with combined metabolic labeling and isobaric mass tags can also be extended to ^{15}N-labeled samples. Alternatively, the dimethyl chemical labeling technique can be combined with isobaric mass tags to increase the multiplexing capacity of quantitative proteomics. Theoretically, the combination of iTRAQ 8-plex or TMT 10-plex reagents and triplex SILAC would allow 24 or 30 channels to be monitored simultaneously.

6.3. Targeted Analysis with Isobaric Mass Tags

Proteins that are identified and quantified as differentially expressed can be used for subsequent targeted studies using the isobaric labeling technique to assess reproducibility of the entire procedure and to validate the observed differences in protein expression levels between samples. During biomarker discovery experiments, targeted investigations are necessary to verify proteins with higher variance in additional patient samples or to obtain greater statistical power. For successful targeted analysis, peptides that allow clear protein quantification and are also sufficiently intense should be selected as representative target peptides for validation. Isobaric mass tags are often used for discovery studies to reveal proteins being differentially expressed under any given conditions. However, Stella et al. have shown that isobaric mass tags can be used in combination with multiple reaction monitoring (MRM) for targeted quantification.[110] In this study, the instrument monitored the two reporter ions and three transitions for each peptide selected from the target membrane proteins. The relative quantification was achieved by comparing the intensities of the reporter ions generated from the labeled precursor peptide of two samples, wild-type (reporter ion m/z 129) and prion protein(PrP)-knockout (reporter ion m/z 131) cerebellar granule neurons.[110] Byers et al. used isotopic versions of TMT reagents for targeted quantification to verify protein regulations observed in a discovery study.[111] These isotopic sets of reagents are structurally identical to the isobaric ones but have different numbers of heavy isotopes incorporated and are referred to as light TMT and heavy TMT (Figure 8). The labeling of peptides by these reagents results in an increase in mass of 224 and 229 Da, respectively, per introduced tag.

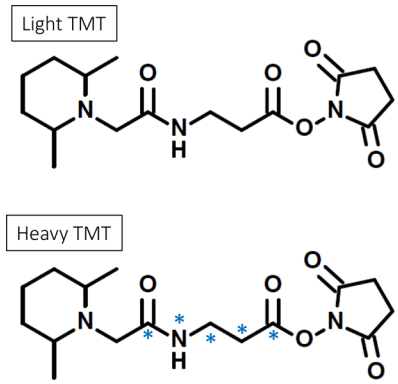


Figure 8. Chemical structure of isotopic reagents, light TMT and heavy TMT, used for targeted quantification.[111] The light reagent has no heavy isotope incorporated, whereas the heavy reagent has five heavy isotopes incorporated (4 × ^{13}C and 1 × ^{15}N). Labeling with these reagents introduces mass differences into the peptides from different samples. In targeted experiments, quantification is obtained by structural b and/or y ions generated after collision-induced dissociation.

7. SUMMARY

The isobaric labeling-based quantification technique has developed as a powerful tool for obtaining the relative expression level of proteins in quantitative proteomics studies. Moreover, the ability to multiplex with isobaric mass tags has expanded its applicability to a wide range of sample types. Isobaric mass tags are isotope-coded molecules with the same chemical structure and molecular weight that are used to differentially label peptides without introducing mass difference and sample complexity. The isotopically derivatized peptides display a single peak on an MS spectrum and yield a series of low-mass reporter ions for quantification upon fragmentation in tandem mass spectrometry. However, since peptide quantification ratios are measured to determine protein relative abundance, the variance in peptide ratio measurements will contribute into the protein-level variance, affecting the accuracy of the quantification. Herein, we have reviewed the studies of different aspects of an isobaric labeling-based quantification approach. This includes studies on different types of isobaric reagents and their applications, sources of variation that affect quantification, and the suggested combinations of experimental design and optimal data acquisition methods to increase the precision and accuracy of the measurements. We have also reviewed studies on challenges in data analysis and the proposed solutions for data processing to increase the confidence in the acquired data set.

AUTHOR INFORMATION

Corresponding Author

*E-mail: jyates@scripps.edu; Phone: 858-784-8862; Fax: 858-784-8883.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

We apologize to authors whose work we could not cite. We would like to thank Drs. Jolene Diedrich, Daniel McClatchy, Aaron Aslanian, and Claire Delahunty for helpful comments and suggestions about the manuscript. Financial support was provided by NIH grant nos. P41 GM103533, R01 HL079442-09, R01 MH067880, and UCLA/NHLBI Proteomics Centers HHSN268201000035C.

REFERENCES

(1) Wu, C. C.; MacCoss, M. J.; Howell, K. E.; Matthews, D. E.; Yates, J. R., III Metabolic labeling of mammalian organisms with stable isotopes for quantitative proteomic analysis. *Anal. Chem.* **2004**, *76*, 4951−4959.

(2) Ong, S. E.; Blagoev, B.; Kratchmarova, I.; Kristensen, D. B.; Steen, H.; Pandey, A.; Mann, M. Stable isotope labeling by amino acids in cell culture, SILAC, as a simple and accurate approach to expression proteomics. *Mol. Cell. Proteomics* **2002**, *1*, 376−386.

(3) Rose, C. M.; Merrill, A. E.; Bailey, D. J.; Hebert, A. S.; Westphall, M. S.; Coon, J. J. Neutron encoded labeling for peptide identification. *Anal. Chem.* **2013**, *85*, 5129−5137.

(4) Gygi, S. P.; Rist, B.; Gerber, S. A.; Turecek, F.; Gelb, M. H.; Aebersold, R. Quantitative analysis of complex protein mixtures using isotope-coded affinity tags. *Nat. Biotechnol.* **1999**, *17*, 994−999.
(5) Hsu, J. L.; Huang, S. Y.; Chow, N. H.; Chen, S. H. Stable-isotope dimethyl labeling for quantitative proteomics. *Anal. Chem.* **2003**, *75*, 6843−6852.
(6) Thompson, A.; Schafer, J.; Kuhn, K.; Kienle, S.; Schwarz, J.; Schmidt, G.; Neumann, T.; Johnstone, R.; Mohammed, A. K.; Hamon, C. Tandem mass tags: a novel quantification strategy for comparative analysis of complex protein mixtures by MS/MS. *Anal. Chem.* **2003**, *75*, 1895−1904.
(7) Ross, P. L.; Huang, Y. 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. Multiplexed protein quantitation in Saccharomyces cerevisiae using amine-reactive isobaric tagging reagents. *Mol. Cell. Proteomics* **2004**, *3*, 1154−1169.
(8) Yao, X.; Freas, A.; Ramirez, J.; Demirev, P. A.; Fenselau, C. Proteolytic ^{18}O labeling for comparative proteomics: model studies with two serotypes of adenovirus. *Anal. Chem.* **2001**, *73*, 2836−2842.
(9) Wiktorowicz, J. E.; English, R. D.; Wu, Z.; Kurosky, A. Model studies on iTRAQ modification of peptides: sequence-dependent reaction specificity. *J. Proteome Res.* **2012**, *11*, 1512−1520.
(10) Choe, L.; D'Ascenzo, M.; Relkin, N. R.; Pappin, D.; Ross, P.; Williamson, B.; Guertin, S.; Pribil, P.; Lee, K. H. 8-Plex quantitation of changes in cerebrospinal fluid protein expression in subjects undergoing intravenous immunoglobulin treatment for Alzheimer's disease. *Proteomics* **2007**, *7*, 3651−3660.
(11) Dayon, L.; Hainard, A.; Licker, V.; Turck, N.; Kuhn, K.; Hochstrasser, D. F.; Burkhard, P. R.; Sanchez, J. C. Relative quantification of proteins in human cerebrospinal fluids by MS/MS using 6-plex isobaric tags. *Anal. Chem.* **2008**, *80*, 2921−2931.
(12) 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. Increasing the multiplexing capacity of TMTs using reporter ion isotopologues with isobaric masses. *Anal. Chem.* **2012**, *84*, 7469−7478.
(13) Werner, T.; Becher, I.; Sweetman, G.; Doce, C.; Savitski, M. M.; Bantscheff, M. High-resolution enabled TMT 8-plexing. *Anal. Chem.* **2012**, *84*, 7188−7194.
(14) Werner, T.; Sweetman, G.; Savitski, M. F.; Mathieson, T.; Bantscheff, M.; Savitski, M. M. Ion coalescence of neutron encoded TMT 10-plex reporter ions. *Anal. Chem.* **2014**, *86*, 3594−3601.
(15) Zieske, L. R. A perspective on the use of iTRAQ reagent technology for protein complex and profiling studies. *J. Exp. Bot.* **2006**, *57*, 1501−1508.
(16) Pichler, P.; Kocher, T.; Holzmann, J.; Mazanek, M.; Taus, T.; Ammerer, G.; Mechtler, K. Peptide labeling with isobaric tags yields higher identification rates using iTRAQ 4-plex compared to TMT 6-plex and iTRAQ 8-plex on LTQ Orbitrap. *Anal. Chem.* **2010**, *82*, 6549−6558.
(17) Pottiez, G.; Wiederin, J.; Fox, H. S.; Ciborowski, P. Comparison of 4-plex to 8-plex iTRAQ quantitative measurements of proteins in human plasma samples. *J. Proteome Res.* **2012**, *11*, 3774−3781.
(18) Xiang, F.; Ye, H.; Chen, R.; Fu, Q.; Li, L. *N,N*-Dimethyl leucines as novel isobaric tandem mass tags for quantitative proteomics and peptidomics. *Anal. Chem.* **2010**, *82*, 2817−2825.
(19) Zhang, R.; Sioma, C. S.; Wang, S.; Regnier, F. E. Fractionation of isotopically labeled peptides in quantitative proteomics. *Anal. Chem.* **2001**, *73*, 5142−5149.
(20) Chen, Z.; Wang, Q.; Lin, L.; Tang, Q.; Edwards, J. L.; Li, S.; Liu, S. Comparative evaluation of two isobaric labeling tags, DiART and iTRAQ. *Anal. Chem.* **2012**, *84*, 2908−2915.
(21) Zhang, J.; Wang, Y.; Li, S. Deuterium isobaric amine-reactive tags for quantitative proteomics. *Anal. Chem.* **2010**, *82*, 7588−7595.
(22) Zhang, Y.; Askenazi, M.; Jiang, J.; Luckey, C. J.; Griffin, J. D.; Marto, J. A. A robust error model for iTRAQ quantification reveals divergent signaling between oncogenic FLT3 mutants in acute myeloid leukemia. *Mol. Cell. Proteomics* **2010**, *9*, 780−790.
(23) Ow, S. Y.; Salim, M.; Noirel, J.; Evans, C.; Rehman, I.; Wright, P. C. iTRAQ underestimation in simple and complex mixtures: "the good, the bad and the ugly". *J. Proteome Res.* **2009**, *8*, 5347−5355.
(24) Ramsubramaniam, N.; Tao, F.; Li, S.; Marten, M. R. Cost-effective isobaric tagging for quantitative phosphoproteomics using DiART reagents. *Mol. BioSyst.* **2013**, *9*, 2981−2987.
(25) Nystrom, T. Role of oxidative carbonylation in protein quality control and senescence. *EMBO J.* **2005**, *24*, 1311−1317.
(26) Palmese, A.; De Rosa, C.; Chiappetta, G.; Marino, G.; Amoresano, A. Novel method to investigate protein carbonylation by iTRAQ strategy. *Anal. Bioanal. Chem.* **2012**, *404*, 1631−1635.
(27) Hahne, H.; Neubert, P.; Kuhn, K.; Etienne, C.; Bomgarden, R.; Rogers, J. C.; Kuster, B. Carbonyl-reactive tandem mass tags for the proteome-wide quantification of N-linked glycans. *Anal. Chem.* **2012**, *84*, 3716−3724.
(28) Thomas, J. A.; Mallis, R. J. Aging and oxidation of reactive protein sulfhydryls. *Exp. Gerontol.* **2001**, *36*, 1519−1526.
(29) Giron, P.; Dayon, L.; Turck, N.; Hoogland, C.; Sanchez, J. C. Quantitative analysis of human cerebrospinal fluid proteins using a combination of cysteine tagging and amine-reactive isobaric labeling. *J. Proteome Res.* **2011**, *10*, 249−258.
(30) Kohr, M. J.; Aponte, A.; Sun, J.; Gucek, M.; Steenbergen, C.; Murphy, E. Measurement of S-nitrosylation occupancy in the myocardium with cysteine-reactive tandem mass tags: short communication. *Circ. Res.* **2012**, *111*, 1308−1312.
(31) Murray, C. I.; Uhrigshardt, H.; O'Meally, R. N.; Cole, R. N.; Van Eyk, J. E. Identification and quantification of S-nitrosylation by cysteine reactive tandem mass tag switch assay. *Mol. Cell. Proteomics* **2012**, *11*, M111.013441.
(32) Pan, K. T.; Chen, Y. Y.; Pu, T. H.; Chao, Y. S.; Yang, C. Y.; Bomgarden, R. D.; Rogers, J. C.; Meng, T. C.; Khoo, K. H. Mass spectrometry-based quantitative proteomics for dissecting multiplexed redox cysteine modifications in nitric oxide-protected cardiomyocyte under hypoxia. *Antioxid. Redox Signaling.* **2014**, *20*, 1365−1381.
(33) Qu, Z.; Meng, F.; Bomgarden, R. D.; Viner, R. I.; Li, J.; Rogers, J. C.; Cheng, J.; Greenlief, C. M.; Cui, J.; Lubahn, D. B.; Sun, G. Y.; Gu, Z. Proteomic quantification and site-mapping of S-nitrosylated proteins using isobaric iodoTMT reagents. *J. Proteome Res.* **2014**, *13*, 3200−3211.
(34) West, G. M.; Tucker, C. L.; Xu, T.; Park, S. K.; Han, X.; Yates, J. R., III; Fitzgerald, M. C. Quantitative proteomics approach for identifying protein-drug interactions in complex mixtures using protein stability measurements. *Proc. Natl. Acad. Sci. U.S.A.* **2010**, *107*, 9078−9082.
(35) Zhang, Y.; Wolf-Yadlin, A.; Ross, P. L.; Pappin, D. J.; Rush, J.; Lauffenburger, D. A.; White, F. M. Time-resolved mass spectrometry of tyrosine phosphorylation sites in the epidermal growth factor receptor signaling network reveals dynamic modules. *Mol. Cell. Proteomics* **2005**, *4*, 1240−1250.
(36) Li, Z.; Adams, R. M.; Chourey, K.; Hurst, G. B.; Hettich, R. L.; Pan, C. Systematic comparison of label-free, metabolic labeling, and isobaric chemical labeling for quantitative proteomics on LTQ Orbitrap Velos. *J. Proteome Res.* **2012**, *11*, 1582−1590.
(37) Bouchal, P.; Roumeliotis, T.; Hrstka, R.; Nenutil, R.; Vojtesek, B.; Garbis, S. D. Biomarker discovery in low-grade breast cancer using isobaric stable isotope tags and two-dimensional liquid chromatography-tandem mass spectrometry (iTRAQ-2DLC−-MS/MS) based quantitative proteomic analysis. *J. Proteome Res.* **2009**, *8*, 362−373.
(38) Aggarwal, K.; Choe, L. H.; Lee, K. H. Shotgun proteomics using the iTRAQ isobaric tags. *Briefings Funct. Genomics Proteomics* **2006**, *5*, 112−120.
(39) Trotter, M. W.; Sadowski, P. G.; Dunkley, T. P.; Groen, A. J.; Lilley, K. S. Improved sub-cellular resolution via simultaneous analysis of organelle proteomics data across varied experimental conditions. *Proteomics* **2010**, *10*, 4213−4219.
(40) Hardt, M.; Witkowska, H. E.; Webb, S.; Thomas, L. R.; Dixon, S. E.; Hall, S. C.; Fisher, S. J. Assessing the effects of diurnal variation on the composition of human parotid saliva: quantitative analysis of

native peptides using iTRAQ reagents. *Anal. Chem.* **2005**, *77*, 4947−4954.
(41) Mertins, P.; Udeshi, N. D.; Clauser, K. R.; Mani, D. R.; Patel, J.; Ong, S. E.; Jaffe, J. D.; Carr, S. A. iTRAQ labeling is superior to mTRAQ for quantitative global proteomics and phosphoproteomics. *Mol. Cell. Proteomics* **2012**, *11*, M111.014423.
(42) Choi, S.; Cho, K.; Kim, J.; Yea, K.; Park, G.; Lee, J.; Ryu, S. H.; Kim, J.; Kim, Y. H. Comparative proteome analysis using amine-reactive isobaric tagging reagents coupled with 2D LC/MS/MS in 3T3-L1 adipocytes following hypoxia or normoxia. *Biochem. Biophys. Res. Commun.* **2009**, *383*, 135−140.
(43) Unwin, R. D. Quantification of proteins by iTRAQ. *Methods Mol. Biol.* **2010**, *658*, 205−215.
(44) Garbis, S. D.; Tyritzis, S. I.; Roumeliotis, T.; Zerefos, P.; Giannopoulou, E. G.; Vlahou, A.; Kossida, S.; Diaz, J.; Vourekas, S.; Tamvakopoulos, C.; Pavlakis, K.; Sanoudou, D.; Constantinides, C. A. Search for potential markers for prostate cancer diagnosis, prognosis and treatment in clinical tissue specimens using amine-specific isobaric tagging (iTRAQ) with two-dimensional liquid chromatography and tandem mass spectrometry. *J. Proteome Res.* **2008**, *7*, 3146−3158.
(45) Huang, Z.; Wang, H.; Huang, H.; Xia, L.; Chen, C.; Qiu, X.; Chen, J.; Chen, S.; Liang, W.; Huang, M.; Lang, L.; Zheng, Q.; Wu, B.; Lai, G. iTRAQ-based proteomic profiling of human serum reveals down-regulation of platelet basic protein and apolipoprotein B100 in patients with hematotoxicity induced by chronic occupational benzene exposure. *Toxicology* **2012**, *291*, 56−64.
(46) Zhou, L.; Beuerman, R. W.; Chan, C. M.; Zhao, S. Z.; Li, X. R.; Yang, H.; Tong, L.; Liu, S.; Stern, M. E.; Tan, D. Identification of tear fluid biomarkers in dry eye syndrome using iTRAQ quantitative proteomics. *J. Proteome Res.* **2009**, *8*, 4889−4905.
(47) Cong, Y. S.; Fan, E.; Wang, E. Simultaneous proteomic profiling of four different growth states of human fibroblasts, using amine-reactive isobaric tagging reagents and tandem mass spectrometry. *Mech. Ageing Dev.* **2006**, *127*, 332−343.
(48) Zhong, J.; Krawczyk, S. A.; Chaerkady, R.; Huang, H.; Goel, R.; Bader, J. S.; Wong, G. W.; Corkey, B. E.; Pandey, A. Temporal profiling of the secretome during adipogenesis in humans. *J. Proteome Res.* **2010**, *9*, 5228−5238.
(49) O'Brien, R. N.; Shen, Z.; Tachikawa, K.; Lee, P. A.; Briggs, S. P. Quantitative proteome analysis of pluripotent cells by iTRAQ mass tagging reveals post-transcriptional regulation of proteins required for ES cell self-renewal. *Mol. Cell. Proteomics* **2010**, *9*, 2238−2251.
(50) Hirsch, J.; Hansen, K. C.; Choi, S.; Noh, J.; Hirose, R.; Roberts, J. P.; Matthay, M. A.; Burlingame, A. L.; Maher, J. J.; Niemann, C. U. Warm ischemia-induced alterations in oxidative and inflammatory proteins in hepatic Kupffer cells in rats. *Mol. Cell. Proteomics* **2006**, *5*, 979−986.
(51) Rauniyar, N.; Gao, B.; McClatchy, D. B.; Yates, J. R., III Comparison of protein expression ratios observed by sixplex and duplex TMT labeling method. *J. Proteome Res.* **2013**, *12*, 1031−1039.
(52) Kleifeld, O.; Doucet, A.; Prudova, A.; auf dem Keller, U.; Gioia, M.; Kizhakkedathu, J. N.; Overall, C. M. Identifying and quantifying proteolytic events and the natural N terminome by terminal amine isotopic labeling of substrates. *Nat. Protoc.* **2011**, *6*, 1578−1611.
(53) Prudova, A.; auf dem Keller, U.; Butler, G. S.; Overall, C. M. Multiplex N-terminome analysis of MMP-2 and MMP-9 substrate degradomes by iTRAQ-TAILS quantitative proteomics. *Mol. Cell. Proteomics* **2010**, *9*, 894−911.
(54) Hsu, P. P.; Kang, S. A.; Rameseder, J.; Zhang, Y.; Ottina, K. A.; Lim, D.; Peterson, T. R.; Choi, Y.; Gray, N. S.; Yaffe, M. B.; Marto, J. A.; Sabatini, D. M. The mTOR-regulated phosphoproteome reveals a mechanism of mTORC1-mediated inhibition of growth factor signaling. *Science* **2011**, *332*, 1317−1322.
(55) Iwai, L. K.; Benoist, C.; Mathis, D.; White, F. M. Quantitative phosphoproteomic analysis of T cell receptor signaling in diabetes prone and resistant mice. *J. Proteome Res.* **2010**, *9*, 3135−3145.
(56) Nilsson, C. L.; Dillon, R.; Devakumar, A.; Shi, S. D.; Greig, M.; Rogers, J. C.; Krastins, B.; Rosenblatt, M.; Kilmer, G.; Major, M.; Kaboord, B. J.; Sarracino, D.; Rezai, T.; Prakash, A.; Lopez, M.; Ji, Y.; Priebe, W.; Lang, F. F.; Colman, H.; Conrad, C. A. Quantitative phosphoproteomic analysis of the STAT3/IL-6/HIF1alpha signaling network: an initial study in GSC11 glioblastoma stem cells. *J. Proteome Res.* **2010**, *9*, 430−443.
(57) Wiese, S.; Reidegeld, K. A.; Meyer, H. E.; Warscheid, B. Protein labeling by iTRAQ: a new tool for quantitative mass spectrometry in proteome research. *Proteomics* **2007**, *7*, 340−350.
(58) Keshamouni, V. G.; Michailidis, G.; Grasso, C. S.; Anthwal, S.; Strahler, J. R.; Walker, A.; Arenberg, D. A.; Reddy, R. C.; Akulapalli, S.; Thannickal, V. J.; Standiford, T. J.; Andrews, P. C.; Omenn, G. S. Differential protein expression profiling by iTRAQ-2DLC-MS/MS of lung cancer cells undergoing epithelial-mesenchymal transition reveals a migratory/invasive phenotype. *J. Proteome Res.* **2006**, *5*, 1143−1154.
(59) Wolf-Yadlin, A.; Hautaniemi, S.; Lauffenburger, D. A.; White, F. M. Multiple reaction monitoring for robust quantitative proteomic analysis of cellular signaling networks. *Proc. Natl. Acad. Sci. U.S.A.* **2007**, *104*, 5860−5865.
(60) Griffin, T. J.; Xie, H.; Bandhakavi, S.; Popko, J.; Mohan, A.; Carlis, J. V.; Higgins, L. iTRAQ reagent-based quantitative proteomic analysis on a linear ion trap mass spectrometer. *J. Proteome Res.* **2007**, *6*, 4200−4209.
(61) Kocher, T.; Pichler, P.; Schutzbier, M.; Stingl, C.; Kaul, A.; Teucher, N.; Hasenfuss, G.; Penninger, J. M.; Mechtler, K. High precision quantitative proteomics using iTRAQ on an LTQ Orbitrap: a new mass spectrometric method combining the benefits of all. *J. Proteome Res.* **2009**, *8*, 4743−4752.
(62) Schwartz, J. C.; Syka, J. E. P.; Quarmby, S. T. Improving the fundamentals of MS^n on 2D linear ion traps: New ion activation and isolation techniques, Proc. 53rd ASMS Conference on Mass Spectrometry and Allied Topics, San Antonio, TX, June 5−9, 2005, oral presentation.
(63) Bantscheff, M.; Boesche, M.; Eberhard, D.; Matthieson, T.; Sweetman, G.; Kuster, B. Robust and sensitive iTRAQ quantification on an LTQ Orbitrap mass spectrometer. *Mol. Cell. Proteomics* **2008**, *7*, 1702−1713.
(64) Guo, T.; Gan, C. S.; Zhang, H.; Zhu, Y.; Kon, O. L.; Sze, S. K. Hybridization of pulsed-Q dissociation and collision-activated dissociation in linear ion trap mass spectrometer for iTRAQ quantitation. *J. Proteome Res.* **2008**, *7*, 4831−4840.
(65) Olsen, J. V.; Macek, B.; Lange, O.; Makarov, A.; Horning, S.; Mann, M. Higher-energy C-trap dissociation for peptide modification analysis. *Nat. Methods* **2007**, *4*, 709−712.
(66) Dayon, L.; Pasquarello, C.; Hoogland, C.; Sanchez, J. C.; Scherl, A. Combining low- and high-energy tandem mass spectra for optimized peptide quantification with isobaric tags. *J. Proteomics* **2010**, *73*, 769−777.
(67) Chiva, C.; Sabido, E. HCD-only fragmentation method balances peptide identification and quantitation of TMT-labeled samples in hybrid linear ion trap/orbitrap mass spectrometers. *J. Proteomics* **2014**, *96*, 263−270.
(68) Pichler, P.; Kocher, T.; Holzmann, J.; Mohring, T.; Ammerer, G.; Mechtler, K. Improved precision of iTRAQ and TMT quantification by an axial extraction field in an Orbitrap HCD cell. *Anal. Chem.* **2011**, *83*, 1469−1474.
(69) Diedrich, J. K.; Pinto, A. F.; Yates, J. R., III Energy dependence of HCD on peptide fragmentation: stepped collisional energy finds the sweet spot. *J. Am. Soc. Mass Spectrom.* **2013**, *24*, 1690−1699.
(70) Ting, L.; Rad, R.; Gygi, S. P.; Haas, W. MS3 eliminates ratio distortion in isobaric multiplexed quantitative proteomics. *Nat. Methods* **2011**, *8*, 937−940.
(71) Senko, M. W. Multinotch isolation for ms3 mass analysis. Patent WO/2013/112677, 2013.
(72) McAlister, G. C.; Nusinow, D. P.; Jedrychowski, M. P.; Wuhr, M.; Huttlin, E. L.; Erickson, B. K.; Rad, R.; Haas, W.; Gygi, S. P. MultiNotch MS3 enables accurate, sensitive, and multiplexed detection of differential expression across cancer cell line proteomes. *Anal. Chem.* **2014**, *86*, 7150−7158.

(73) Armenta, J. M.; Hoeschele, I.; Lazar, I. M. Differential protein expression analysis using stable isotope labeling and PQD linear ion trap MS technology. *J. Am. Soc. Mass Spectrom.* **2009**, *20*, 1287−1302.
(74) Liu, T.; Hu, J.; Li, H. iTRAQ-based shotgun neuroproteomics. *Methods Mol. Biol.* **2009**, *566*, 201−216.
(75) Shadforth, I. P.; Dunkley, T. P.; Lilley, K. S.; Bessant, C. i-Tracker: for quantitative proteomics using iTRAQ. *BMC Genomics* **2005**, *6*, 145.
(76) Savitski, M. M.; Sweetman, G.; Askenazi, M.; Marto, J. A.; Lang, M.; Zinn, N.; Bantscheff, M. Delayed fragmentation and optimized isolation width settings for improvement of protein identification and accuracy of isobaric mass tag quantification on Orbitrap-type mass spectrometers. *Anal. Chem.* **2011**, *83*, 8959−8967.
(77) Karp, N. A.; Huber, W.; Sadowski, P. G.; Charles, P. D.; Hester, S. V.; Lilley, K. S. Addressing accuracy and precision issues in iTRAQ quantitation. *Mol. Cell. Proteomics* **2010**, *9*, 1885−1897.
(78) Christoforou, A. L.; Lilley, K. S. Isobaric tagging approaches in quantitative proteomics: the ups and downs. *Anal. Bioanal. Chem.* **2012**, *404*, 1029−1037.
(79) Mathur, R.; O'Connor, P. B. Artifacts in Fourier transform mass spectrometry. *Rapid Commun. Mass Spectrom.* **2009**, *23*, 523−529.
(80) Hill, E. G.; Schwacke, J. H.; Comte-Walters, S.; Slate, E. H.; Oberg, A. L.; Eckel-Passow, J. E.; Therneau, T. M.; Schey, K. L. A statistical model for iTRAQ data analysis. *J. Proteome Res.* **2008**, *7*, 3091−3101.
(81) Dayon, L.; Turck, N.; Kienle, S.; Schulz-Knappe, P.; Hochstrasser, D. F.; Scherl, A.; Sanchez, J. C. Isobaric tagging-based selection and quantitation of cerebrospinal fluid tryptic peptides with reporter calibration curves. *Anal. Chem.* **2010**, *82*, 848−858.
(82) Aggarwal, K.; Choe, L. H.; Lee, K. H. Quantitative analysis of protein expression using amine-specific isobaric tags in Escherichia coli cells expressing rhsA elements. *Proteomics* **2005**, *5*, 2297−2308.
(83) Ow, S. Y.; Salim, M.; Noirel, J.; Evans, C.; Wright, P. C. Minimising iTRAQ ratio compression through understanding LC−MS elution dependence and high-resolution HILIC fractionation. *Proteomics* **2011**, *11*, 2341−2346.
(84) Savitski, M. M.; Fischer, F.; Mathieson, T.; Sweetman, G.; Lang, M.; Bantscheff, M. Targeted data acquisition for improved reproducibility and robustness of proteomic mass spectrometry assays. *J. Am. Soc. Mass Spectrom.* **2010**, *21*, 1668−1679.
(85) Wenger, C. D.; Lee, M. V.; Hebert, A. S.; McAlister, G. C.; Phanstiel, D. H.; Westphall, M. S.; Coon, J. J. Gas-phase purification enables accurate, multiplexed proteome quantification with isobaric tagging. *Nat. Methods* **2011**, *8*, 933−935.
(86) Sturm, R. M.; Lietz, C. B.; Li, L. Improved isobaric tandem mass tag quantification by ion mobility mass spectrometry. *Rapid Commun. Mass Spectrom.* **2014**, *28*, 1051−1060.
(87) Savitski, M. M.; Mathieson, T.; Zinn, N.; Sweetman, G.; Doce, C.; Becher, I.; Pachl, F.; Kuster, B.; Bantscheff, M. Measuring and managing ratio compression for accurate iTRAQ/TMT quantification. *J. Proteome Res.* **2013**, *12*, 3586−3598.
(88) Onsongo, G.; Stone, M. D.; Van Riper, S. K.; Chilton, J.; Wu, B.; Higgins, L.; Lund, T. C.; Carlis, J. V.; Griffin, T. J. LTQ-iQuant: a freely available software pipeline for automated and accurate protein quantification of isobaric tagged peptide data from LTQ instruments. *Proteomics* **2010**, *10*, 3533−3538.
(89) Burkhart, J. M.; Vaudel, M.; Zahedi, R. P.; Martens, L.; Sickmann, A. iTRAQ protein quantification: a quality-controlled workflow. *Proteomics* **2011**, *11*, 1125−1134.
(90) Pachl, F.; Fellenberg, K.; Wagner, C.; Kuster, B. Ultra-high intra-spectrum mass accuracy enables unambiguous identification of fragment reporter ions in isobaric multiplexed quantitative proteomics. *Proteomics* **2012**, *12*, 1328−1332.
(91) Hu, J.; Qian, J.; Borisov, O.; Pan, S.; Li, Y.; Liu, T.; Deng, L.; Wannemacher, K.; Kurnellas, M.; Patterson, C.; Elkabes, S.; Li, H. Optimized proteomic analysis of a mouse model of cerebellar dysfunction using amine-specific isobaric tags. *Proteomics* **2006**, *6*, 4321−4334.
(92) Lin, W. T.; Hung, W. N.; Yian, Y. H.; Wu, K. P.; Han, C. L.; Chen, Y. R.; Chen, Y. J.; Sung, T. Y.; Hsu, W. L. Multi-Q: a fully automated tool for multiplexed protein quantitation. *J. Proteome Res.* **2006**, *5*, 2328−2338.
(93) Song, X.; Bandow, J.; Sherman, J.; Baker, J. D.; Brown, P. W.; McDowell, M. T.; Molloy, M. P. iTRAQ experimental design for plasma biomarker discovery. *J. Proteome Res.* **2008**, *7*, 2952−2958.
(94) Nesvizhskii, A. I.; Aebersold, R. Interpretation of shotgun proteomic data: the protein inference problem. *Mol. Cell. Proteomics* **2005**, *4*, 1419−1440.
(95) Mahoney, D. W.; Therneau, T. M.; Heppelmann, C. J.; Higgins, L.; Benson, L. M.; Zenka, R. M.; Jagtap, P.; Nelsestuen, G. L.; Bergen, H. R.; Oberg, A. L. Relative quantification: characterization of bias, variability and fold changes in mass spectrometry data from iTRAQ-labeled peptides. *J. Proteome Res.* **2011**, *10*, 4325−4333.
(96) Veenstra, T. D.; Conrads, T. P.; Issaq, H. J. What to do with "one-hit wonders"? *Electrophoresis* **2004**, *25*, 1278−1279.
(97) Gan, C. S.; Chong, P. K.; Pham, T. K.; Wright, P. C. Technical, experimental, and biological variations in isobaric tags for relative and absolute quantitation (iTRAQ). *J. Proteome Res.* **2007**, *6*, 821−827.
(98) Glen, A.; Gan, C. S.; Hamdy, F. C.; Eaton, C. L.; Cross, S. S.; Catto, J. W.; Wright, P. C.; Rehman, I. iTRAQ-facilitated proteomic analysis of human prostate cancer cells identifies proteins associated with progression. *J. Proteome Res.* **2008**, *7*, 897−907.
(99) Zhou, C.; Simpson, K. L.; Lancashire, L. J.; Walker, M. J.; Dawson, M. J.; Unwin, R. D.; Rembielak, A.; Price, P.; West, C.; Dive, C.; Whetton, A. D. Statistical considerations of optimal study design for human plasma proteomics and biomarker discovery. *J. Proteome Res.* **2012**, *11*, 2103−2113.
(100) Wang, H.; Alvarez, S.; Hicks, L. M. Comprehensive comparison of iTRAQ and label-free LC-based quantitative proteomics approaches using two *Chlamydomonas reinhardtii* strains of interest for biofuels engineering. *J. Proteome Res.* **2012**, *11*, 487−501.
(101) Romero, R.; Kusanovic, J. P.; Gotsch, F.; Erez, O.; Vaisbuch, E.; Mazaki-Tovi, S.; Moser, A.; Tam, S.; Leszyk, J.; Master, S. R.; Juhasz, P.; Pacora, P.; Ogge, G.; Gomez, R.; Yoon, B. H.; Yeo, L.; Hassan, S. S.; Rogers, W. T. Isobaric labeling and tandem mass spectrometry: a novel approach for profiling and quantifying proteins differentially expressed in amniotic fluid in preterm labor with and without intra-amniotic infection/inflammation. *J. Matern.-Fetal Neonat. Med.* **2010**, *23*, 261−280.
(102) Herbrich, S. M.; Cole, R. N.; West, K. P., Jr.; Schulze, K.; Yager, J. D.; Groopman, J. D.; Christian, P.; Wu, L.; O'Meally, R. N.; May, D. H.; McIntosh, M. W.; Ruczinski, I. Statistical inference from multiple iTRAQ experiments without using common reference standards. *J. Proteome Res.* **2013**, *12*, 594−604.
(103) Seshi, B. An integrated approach to mapping the proteome of the human bone marrow stromal cell. *Proteomics* **2006**, *6*, 5169−5182.
(104) Lund, T. C.; Anderson, L. B.; McCullar, V.; Higgins, L.; Yun, G. H.; Grzywacz, B.; Verneris, M. R.; Miller, J. S. iTRAQ is a useful method to screen for membrane-bound proteins differentially expressed in human natural killer cell types. *J. Proteome Res.* **2007**, *6*, 644−653.
(105) Wu, J.; Warren, P.; Shakey, Q.; Sousa, E.; Hill, A.; Ryan, T. E.; He, T. Integrating titania enrichment, iTRAQ labeling, and Orbitrap CID-HCD for global identification and quantitative analysis of phosphopeptides. *Proteomics* **2010**, *10*, 2224−2234.
(106) Linke, D.; Hung, C. W.; Cassidy, L.; Tholey, A. Optimized fragmentation conditions for iTRAQ-labeled phosphopeptides. *J. Proteome Res.* **2013**, *12*, 2755−2763.
(107) Nilsson, C. L. Advances in quantitative phosphoproteomics. *Anal. Chem.* **2012**, *84*, 735−746.
(108) Thingholm, T. E.; Palmisano, G.; Kjeldsen, F.; Larsen, M. R. Undesirable charge-enhancement of isobaric tagged phosphopeptides leads to reduced identification efficiency. *J. Proteome Res.* **2010**, *9*, 4045−4052.

(109) Dephoure, N.; Gygi, S. P. Hyperplexing: a method for higher-order multiplexed quantitative proteomics provides a map of the dynamic response to rapamycin in yeast. *Sci. Signaling* **2012**, *5*, rs2.
(110) Stella, R.; Cifani, P.; Peggion, C.; Hansson, K.; Lazzari, C.; Bendz, M.; Levander, F.; Sorgato, M. C.; Bertoli, A.; James, P. Relative quantification of membrane proteins in wild-type and prion protein (PrP)-knockout cerebellar granule neurons. *J. Proteome Res.* **2012**, *11*, 523−536.
(111) Byers, H. L.; Campbell, J.; van Ulsen, P.; Tommassen, J.; Ward, M. A.; Schulz-Knappe, P.; Prinz, T.; Kuhn, K. Candidate verification of iron-regulated *Neisseria meningitidis* proteins using isotopic versions of tandem mass tags (TMT) and single reaction monitoring. *J. Proteomics* **2009**, *73*, 231−239.

Cell

Leading Edge
Review

Communication by Extracellular Vesicles: Where We Are and Where We Need to Go

Mercedes Tkach[1] and Clotilde Théry[1,*]
[1]Institut Curie, PSL Research University, INSERM U932, 75005 Paris, France
*Correspondence: clotilde.thery@curie.fr
http://dx.doi.org/10.1016/j.cell.2016.01.043

In multicellular organisms, distant cells can exchange information by sending out signals composed of single molecules or, as increasingly exemplified in the literature, via complex packets stuffed with a selection of proteins, lipids, and nucleic acids, called extracellular vesicles (EVs; also known as exosomes and microvesicles, among other names). This Review covers some of the most striking functions described for EV secretion but also presents the limitations on our knowledge of their physiological roles. While there are initial indications that EV-mediated pathways operate in vivo, the actual nature of the EVs involved in these effects still needs to be clarified. Here, we focus on the context of tumor cells and their microenvironment, but similar results and challenges apply to all patho/physiological systems in which EV-mediated communication is proposed to take place.

Introduction

Cells can communicate with neighboring cells or with distant cells through the secretion of extracellular vesicles (EVs). EVs are composed of a lipid bilayer containing transmembrane proteins and enclosing cytosolic proteins and RNA. Cells can secrete different types of EVs that have been classified according to their sub-cellular origin (Colombo et al., 2014). On one hand, EVs can be formed and released by budding from the cells' plasma membrane. These EVs display a diverse range of sizes (100–1,000 nm in diameter) and are generally known in the literature as microvesicles, ectosomes, or microparticles. Other types of vesicles, the exosomes, are generated inside multivesicular endosomes or multivesicular bodies (MVBs) and are secreted when these compartments fuse with the plasma membrane. Exosomes are vesicles smaller than 150 nm in diameter and are enriched in endosome-derived components. All EVs bear surface molecules that allow them to be targeted to recipient cells. Once attached to a target cell, EVs can induce signaling via receptor-ligand interaction or can be internalized by endocytosis and/or phagocytosis or even fuse with the target cell's membrane to deliver their content into its cytosol, thereby modifying the physiological state of the recipient cell.

In this Review, we highlight and discuss the more recent studies on cancer-derived EVs, with a special focus on the latest discoveries on the role of EVs in cancer metastasis. The term "exosomes" is often used in these articles to designate the EVs analyzed. However, we now know that the most popular exosome purification protocols used historically in the literature (differential ultracentrifugation, 220 nm filtration [Thery et al., 2006])—and the recently released commercial kits—co-isolate different types of EVs. Thus, the term exosomes is generally used to refer to a mixed population of small EVs (sEVs) without further demonstration of their intracellular origin. In fact, functions assigned to exosomes may either reflect generic EV activities or truly exosome-specific ones; however, the published data cannot be used to determine the precise specificity. We thus chose here to use the generic term EVs when vesicles are isolated without specific attention to their size or sEVs when the method used selects vesicles smaller than 200 nm, independent of the term used in the article referred to.

EV-Borne Proteins Promote Cancer Progression and Metastasis

EVs have been shown to participate in the dissemination of cancer cells, and many groups have described how tumor- and stroma-derived EVs are involved in the different steps of the metastatic cascade (Figure 1). Tumor sEVs can directly modify tumor cells' intrinsic motility and invasiveness capacity. In particular, sEVs can promote directional cell motility through ECM components, such as fibronectin, which bind to integrins present on sEVs and thus provide a substrate favoring cell adhesion and enhancing cell speed (Sung et al., 2015). Moreover, sEVs participate in the biogenesis and activity of an invasive structure called invadopodia through the MVB-dependent delivery of metalloproteinases such as MT1-MMP and other cargo molecules (Hoshino et al., 2013), thus promoting cell motility. EVs can directly contribute to extracellular matrix (ECM) degradation by spreading matrix metalloproteinases present either on sEVs (Yue et al., 2015) or in tumor-shed large EVs (Clancy et al., 2015). The latter also have been shown to facilitate amoeboid movement and facilitate invasion. Not only tumor-derived EVs, but also sEVs from cancer-associated fibroblasts can stimulate invasiveness of recipient breast cancer cells, in this case by activating the planar cell polarity signaling pathway (Luga et al., 2012).

Tumor sEVs can alter the cellular physiology of both surrounding and distant non-tumor cells to allow dissemination and growth of cancer cells, i.e., by triggering vascular permeability

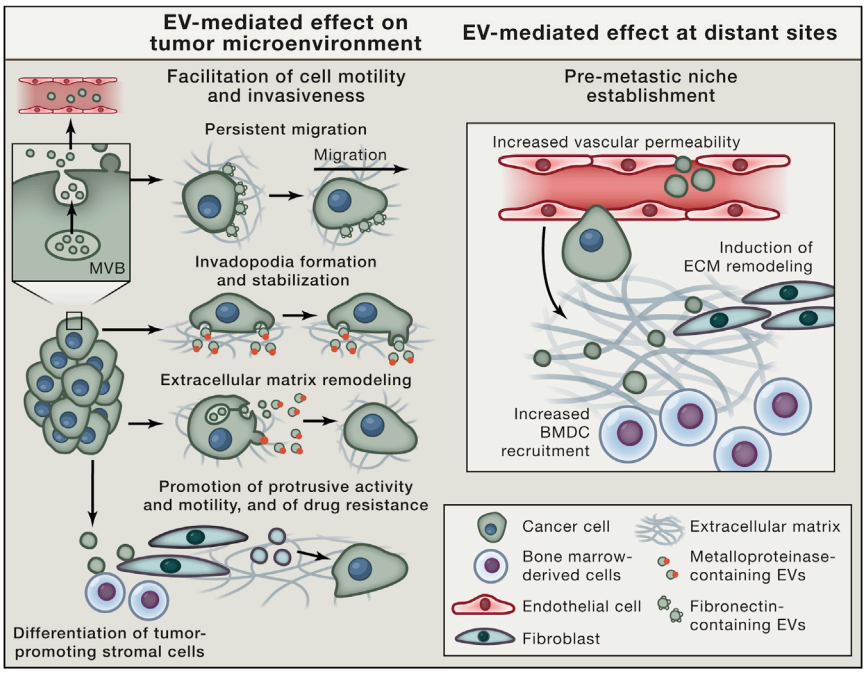


Figure 1. EV-Mediated Effects Promoting Tumor Growth, Invasiveness, and Metastasis

Tumor-derived EVs can have several effects on recipient cells. At the site of the primary tumor (left), EVs can enhance cancer cell motility by stabilizing cellular protrusions promoting an effective and directionally persistent migration via deposition of ECM cargoes, such as fibronectin, into sEVs. The secretion of EVs containing metalloproteinases also directly participates in ECM remodeling and promotes function of specialized cell protrusions endowed with degradative activity, called the invadopodia. ECM remodeling supports tumor cell motility through the tissues. EVs can also promote differentiation or recruitment of pro-tumoral stromal cells (fibroblasts and bone-marrow-derived cells). Reciprocally, tumor cell motility, but also acquisition of drug resistance, can be enhanced via a complex interplay with EVs secreted by surrounding fibroblasts. In addition, sEVs can enter the circulation and travel to distant sites from the primary tumor (right). Various sEV cargoes promote vascular permeability, and EVs can enter the distant tissue, where they may generate a pre-metastatic niche by inducing ECM remodeling and promoting the recruitment of bone-marrow-derived cells and eventually, tumor cells. This figure schematizes the effects of EVs demonstrated by mixed in-vivo-/in-vitro-based experiments. See the text for discussion on the evidence for fully physiological in vivo occurrence of these functions.

(Peinado et al., 2012; Zhou et al., 2014) or by conditioning pre-metastatic sites in distant organs (Costa-Silva et al., 2015; Hoshino et al., 2015; Peinado et al., 2012). In particular, melanoma tumor sEVs bearing a tyrosine-kinase receptor can promote migration of bone marrow progenitor cells to future sites of metastasis, whereas sEVs secreted by a less-aggressive version of the same tumor, devoid of the relevant receptor, do not display this effect (Peinado et al., 2012). Alternatively, sEVs from pancreatic cancer cells themselves migrate to distant organs and promote the formation of a pre-metastatic niche by creating a fibrotic environment enriched in TGFβ, fibronectin, and a macrophage-attracting chemokine (Costa-Silva et al., 2015). Interestingly, sEVs from different tumor types bear integrins (ITGs) that target these sEVs to specific organs and trigger signaling pathways, thereby initiating pre-metastatic niche formation (Hoshino et al., 2015). For example, sEVs expressing ITG$\alpha_v\beta_5$ bind specifically to Kupffer cells, mediating liver tropism, while ITG$\alpha_6\beta_4$ and ITG$\alpha_6\beta_1$ on sEVs bind to lung-resident fibroblasts and epithelial cells, leading to lung tropism (Hoshino et al., 2015). Modifications induced by sEVs in these distant organs then attract metastatic tumor cells.

This observation has been recently used in an innovative way to redirect tumor cell dissemination in a non-deleterious location (de la Fuente et al., 2015). An artificial pre-metastatic niche generated by embedding tumor sEVs in a 3D scaffold and then implanted in mouse peritoneum was able to capture ovarian tumor cells present in the peritoneum and divert them from their normal organ target for dissemination, resulting in strikingly increased survival of the animal. The possible application of this device in human patients could represent a very promising approach to suppress metastasis.

However, despite being extremely appealing, we must stress that the working model of circulating tumor-derived sEVs fostering pre-metastatic niche formation has not been demonstrated in a fully physiological in vivo context. In published articles to date, animals were subjected to sustained injections of in-vitro-purified tumor-derived sEVs, resulting in this enhanced metastasis. Whether sEV secretion in vivo by tumor cells is able to achieve this function is still not clear. One possible way to address this is by interfering in vivo with sEV biogenesis in cancer cells. Some studies have attempted to do this by inhibiting Ras-related RAB proteins. RAB27A or RAB35 have been first shown to be required for sEV secretion in HeLa cervical carcinoma (Ostrowski et al., 2010) and Oli-Neu oligodendroglial precursor cell lines (Hsu et al., 2010), respectively. Consistently, knocking down RAB27A in melanoma (Peinado et al., 2012), breast (Bobrie et al., 2012), fibrosarcoma (Sung et al., 2015), or prostate cancer cell lines (Webber et al., 2015) reduces the secretion of sEVs. Cells lacking RAB27A, when injected in vivo, displayed reduced local migration (Sung et al., 2015) or reduced growth due to impaired recruitment of bone-marrow-derived pro-tumoral immune cells (Bobrie et al., 2012), or impaired modification of co-injected fibroblasts into pro-tumoral myofibroblasts (Webber et al., 2015) (Figure 1). Lower incidence of metastasis was also observed (Bobrie et al., 2012; Peinado et al., 2012). However, RAB27A does not exclusively regulate EV secretion. Loss of the protein also decreases EV-independent secretion of soluble factors, such as some growth factors and metalloproteinases that are also involved in tumor metastasis (Bobrie et al., 2012; Peinado et al., 2012). The same problem has arisen with the other molecules proposed so far to regulate specifically sEV secretion, such as sphingomyelinases

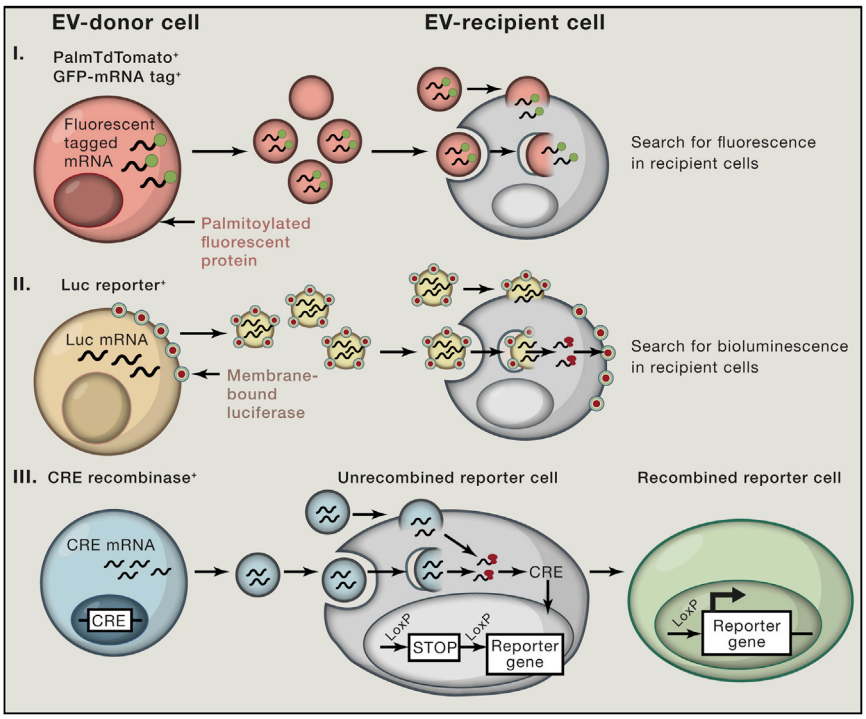


Figure 2. Approaches Used to Analyze EV-Mediated Transfer In Vivo

Recent novel approaches allowing the visualization of EV transfer to recipient cells in vivo involve genetic modification of the secreting cells, which then secrete EVs containing labeled components. (I) Genetic fusion of fluorescent proteins to a consensus palmitoylation sequence, enabling whole-cell membrane labeling, and to an mRNA allows tracking of cells with bound or internalized EVs by their newly acquired fluorescence. (II) Expression of a membrane-bound luciferase allows analysis of distant bioluminescent cells in vivo, resulting from EV-bound luciferase protein capture and/or luciferase mRNA neo-expression. (III) To demonstrate specifically neo-expression of an EV-associated mRNA, the CRE recombinase can be expressed in EV-secreting cells. Endogenous nuclear localization of CRE results in absence of the protein but presence of the mRNA in the secreted EVs. Transfer of the CRE mRNA contained within an EV to a cell carrying a fluorescent or enzymatic reporter gene expressed only after DNA recombination and excision of a STOP signal is detected by fluorescence or colorimetric changes.

(SMases) (Trajkovic et al., 2008). Knockdown or inhibition of SMases by small-molecule inhibitors, which results in impaired ceramide formation, has often been used to inhibit exosome/sEV secretion but without demonstration of the specificity of this effect for sEV secretion, as opposed to other secretions or other physiological features of the cells. Thus, to understand the functions of EVs in vivo, the development of complementary methodologies will be required.

Catching Communication in Action

A major challenge for the EV field and, more broadly, for understanding how EVs may support both physiological and pathophysiological processes is being able to demonstrate in vivo EV transfer between cells. To address this problem, a few groups have recently developed clever modifications of EVs, allowing tracking of their behavior and their target cells in vivo (Figure 2). Either proteins or mRNA cargoes of EVs have been thus modified. For instance, fusing a fluorescent protein to a palmitoylation sequence induces its localization at the plasma membrane as well as in secreted EVs of all sizes. The EVs could then be visualized in the tumor microenvironment by intravital microscopy on animals bearing tumors expressing this fusion protein (Lai et al., 2015). Fusion of luciferase to a protein transmembrane domain also allowed its secretion in EVs and subsequent enzymatic measurement of luciferase activity in distant cells (Lai et al., 2014). These authors also designed an intracellular probe to fluorescently label mRNA secreted in EVs and tools to measure EV-borne mRNA encoding luciferase signal (Lai et al., 2015). While these tools are promising, studies so far have only demonstrated EV-borne mRNA transfer between cultured cells in vitro.

Another elegant approach to address EV transfer has now demonstrated functional EV-mediated transfer in vivo of an mRNA into target cells in the absence of any ex vivo manipulation (Ridder et al., 2014). Transgenic mice expressing the CRE recombinase specifically in immune cells and a LacZ reporter gene, expressed only upon excision of a STOP sequence, were used for this purpose (Ridder et al., 2014). Ridder et al. showed that EVs containing the CRE mRNA, but devoid of the CRE protein, are present in blood circulation of these transgenic mice. Strikingly, LacZ expression was observed in some neurons and other non-immune cells throughout the animals. Leakiness of the immune cell-specific promoter and fusion between CRE-expressing and reporter-expressing cells were carefully excluded, leading to the conclusion that the CRE mRNA carried by immune cell EVs was transferred in the recipient non-immune cells and translated into functional CRE protein. Although the number of recombined neural cells was very low, suggesting a limited efficiency of this transfer in normal conditions, induction of systemic inflammation increased it, opening up the possibility that in vivo EV-mediated transfer may be particularly relevant to some pathological conditions.

This methodology has also been recently used to visualize cancer-derived EV transfer to other cancer cells (Zomer et al., 2015) and to immune cells (Ridder et al., 2015) in living mice and to study the effect of this transfer. In Zomer et al., the authors designed a reporter system based on the conversion of DsRed+ tumor cells to eGFP+ tumor cells upon uptake of tumor EVs containing CRE mRNA and analyzed the behavior of these latter cells without any manipulation of the EV-releasing cells. Using this system, intravital imaging revealed the EV transfer of functional mRNA from a malignant human tumor cell to a less malignant

one and demonstrated that the uptake of these EVs can alter the migratory behavior and metastatic capacity of the recipient cell (Zomer et al., 2015). It will be interesting to study this transfer of EVs among tumor cells in immunocompetent mice. In Ridder et al., a similar strategy highlighted the transfer of vesicle-enclosed CRE mRNA to non-tumor cells and showed that myeloid-derived suppressor cells that took up tumor EVs displayed increased immunosupressive functions (Ridder et al., 2015). In these two studies, the actual nature of the EVs involved in mRNA transfer was not investigated, nor was the transfer mechanism. It is therefore not clear whether transfer involves a direct fusion of EVs with the recipient cells or phagocytosis of live or apoptotic cell-derived EVs by the recipient cell. To more solidly support the pathway for EV-mediated transfer of CRE mRNA, inhibition of EV biogenesis in vivo would be ideal; however, as mentioned before, these experiments are quite difficult to accomplish and are not completely specific. Other transfer mechanisms, like formation of gap junctions or of nanotubes connecting two adjacent cells, were also not formally excluded. However, it was also observed that transfer occurred at a distance between two tumors localized in different parts of the animal, excluding the possibility of local communication between cells and supporting the idea of mRNA being transferred through a long-range extracellular carrier. Overall, this visualization system allows identification and isolation of cells that are targets of EVs in vivo and will be very important for understanding how tumor- and stromal-derived EVs affect their environment.

A Role for EV-Mediated Small RNA Transfer?

In addition to proteins and mRNAs, miRNAs and other non-coding RNAs are also possible active EV cargoes. The idea that miRNA secreted in sEVs can be functionally delivered to target cells, resulting in direct modulation of their mRNA targets, has become one of the most actively explored hypotheses in the EV field, especially in cancer. This idea was initially demonstrated for Epstein-Barr virus-infected cells, where secreted sEVs transferred viral miRNAs into neighboring non-infected cells, leading to repression of virus-target genes (Pegtel et al., 2010). Following this path, several groups have recently reported that EV-mediated secretion of a given miRNA in the tumor microenvironment is responsible for tumor metastasis. Various mechanisms have been proposed, involving either an effect on the local or distant tumor stroma (including modulating endothelial cell permeability, metabolism, or the pre-metastatic niche) or on the tumor cells themselves (by increasing oncogenic properties and/or invasiveness).

However, even if involvement of the described miRNA is generally well supported, direct demonstration that functional EV-mediated miRNA transfer is the relevant mechanism is still difficult to achieve. Importantly, carriers other than EVs could mediate miRNA transfer. All EV isolation techniques potentially co-isolate other RNA-binding structures, such as large protein complexes (Palma et al., 2012) and lipoproteins (Vickers et al., 2011). Lipoprotein-associated RNAs have been shown to be resistant to RNase treatment and to deliver miRNAs into host cells, and there is currently no reason for excluding that transfer of naked protein-miRNA complexes into the cytosol of host cells can occur. Therefore, as the International Society for Extracellular Vesicles recently highlighted (Lötvall et al., 2014), additional steps of separation of EVs from other structures, e.g., by floatation into density gradients or by immuno-isolation via specific antibodies, are necessary before claiming specific EV-mediated miRNA transfer.

The miRNA-dependent effect observed could, in fact, be mediated by induction of endogenous miRNA expression in the target cell by other EV components, rather than by the EV-enclosed miRNA. Such a mechanism can be excluded if the recipient cell is incapable of expressing the studied miRNA—for instance, when the transferred miRNA is encoded only by a foreign genome, e.g., viral (Pegtel et al., 2010) or parasitic (Buck et al., 2014). Similarly, if the recipient cell comes from a mouse engineered to lack the miRNA (Alexander et al., 2015), the dependence on EV-delivered miRNAs is more certain. However, this control has not yet been used to our knowledge in cancer studies. For example, recently, Zhang et al. have proposed that sEV-mediated transfer of astrocyte-derived miRNAs targeting PTEN leads to the loss of expression of this tumor suppressor in brain metastatic tumor cells, enhancing metastasis outgrowth (Zhang et al., 2015). Even though the hypothesis proposed in this work is very appealing, conclusive demonstration that miRNAs are being transferred through sEVs and that this is the mechanism responsible for PTEN downregulation in brain tumor metastasis is lacking. While the effects appear dependent on expression of the miRNAs in astrocytes and are lost upon global deletion of Rab27a in the brain, both controls potentially alter many other aspects than miRNA-containing sEV secretion by astrocytes; thus, specific EV-dependent transport of endogenous miRNA remains a model to be tested.

Other studies have explored similar questions, and it is clear that the field is working toward a suite of more definitive controls. For example, studies showing that co-treatment of recipient cells with an anti-miRNA together with miRNA-carrying EVs established that the analyzed miRNA is necessary for the functional effect observed (Le et al., 2014; Zhou et al., 2014). But since the anti-miRNA can inhibit expression of both the endogenous and the EV-transported molecule, this control does not actually demonstrate EV-mediated miRNA transfer. In this experimental setting, absence of upregulation of the endogenous pre- or pri-miRNA while the mature miRNA is increased in EV-recipient cells (Basu and Bhattacharyya, 2014; Le et al., 2014; Zhou et al., 2014) may be the strongest observation arguing for acquisition of the mature miRNA.

An important mechanistic aspect remains, however, mysterious. To achieve silencing of their mRNA targets, miRNAs must be associated with the RNA-induced silencing complex (RISC) containing the argonaute 2 (AGO2) endonuclease (Wilson and Doudna, 2013). For endogenous miRNAs, this association is formed following processing of the double-stranded pre-miRNA into miRNA/miRNA* duplex by DICER and subsequent incorporation into the AGO2-RISC-loading complex. Thus, how naked mature miRNAs brought in by EVs can associate with endogenous AGO2 and compete with an overwhelming amount of endogenous miRNA will remain unclear unless new molecular mechanisms of miRNA transfer into the RISC-loading complex are discovered. The form of miRNA present in EVs (i.e., as mature single-strand, miRNA/miRNA* duplex or other form,

naked or bound with AGO or other proteins) is generally not reported, and it will be important to answer this question in different cell types (e.g., tumoral or non-tumoral) and physiological contexts to elucidate the relevance of EVs as efficient miRNA carriers.

In that line, a recent study has brought an unexpected and possibly controversial turn to this field (Melo et al., 2014). Melo et al. have observed that pre-miRNAs loaded into the RISC machinery are secreted by tumor (but not by non-tumor) cell lines in sEVs and that miRNA maturation takes place extracellularly when sEVs are incubated at 37°C. The implication is that transferred AGO2-associated miRNA could thus be directly functional in a recipient cell. It will be interesting to see whether these results are borne out in other types of tumors. Furthermore, the actual nature of the RISC/miRNA carrier will have to be determined more precisely since another study reported that AGO2 secreted in tumor sEVs is not recovered in the same fractions as classical sEV markers (Van Deun et al., 2014). In conclusion, whether EV-mediated miRNA transfer is a functionally relevant communication mechanism in cancer, especially in vivo in the absence of artificial overexpression of a miRNA, is still an exciting but not yet fully demonstrated hypothesis.

In addition to miRNAs, EVs are now known to contain several other species of small non-coding RNA or RNA fragments (Nolte-'t Hoen et al., 2012). A recent report shows that miRNA may even be a minor form of RNA in all types of EVs, whereas tRNA fragments and Y-RNAs are specifically secreted in EVs (Tosar et al., 2015). It will be very interesting to determine now whether these comparatively poorly studied non-coding RNAs display some of the gene-regulatory functions so far attributed to miRNAs in EVs. Indeed, a recent report showed that tRNA fragments present in sperm regulate gene expression in the embryo and that the levels of these short tRNAs can be altered in response to paternal diet (Sharma et al., 2016). Interestingly, the tRNA fragments are not generated in spermatozoa themselves but are acquired during their transit through the epididymis, possibly via EVs named epididymosomes, which are secreted by cells forming the epithelium of this canal. The transfer of EVs to sperm had been previously documented in *Drosophila* in vivo (Corrigan et al., 2014) and in large mammals in vitro (Caballero et al., 2013). The novel idea that parental exposure can affect progeny through mechanisms involving transfer of information within EVs is a very promising and exciting hypothesis.

Ties to Immunity

A different consequence of EV-associated RNA transfer in the tumor microenvironment has been recently described (Boelens et al., 2014). EV-borne RNAs bearing 5′-triphosphate ends are recognized in the recipient cell cytosol by the RIG-I sensor, which induces development of an interferon response (including expression of genes like STAT1) similar to that induced upon viral infection. This response was shown to participate in cancer cell resistance to radiation or chemotherapy.

This observation provides one example of a growing appreciation of the similarities between sEVs and enveloped viruses (Assil et al., 2015). Indeed, enveloped viruses have recently been shown to carry a nucleotide-based compound, cGAMP, which can reach the cytosol of infected cells and induce an interferon response (Bridgeman et al., 2015; Gentili et al., 2015). In Gentili et al., we observed that cGAMP was present in non-viral sEVs, which could also transfer this signal to recipient cells, but only if they bore a fusogenic viral protein. This observation raises questions about the molecular mechanism involved in fusion of tumor sEVs with recipient cells: what surface molecules allow fusion with the membrane of recipient cells for delivery of the RNA or small-molecule content into the cytosol? Or can other mechanisms contribute, such as formation of channels between the apposed membranes of an EV and the recipient cell? These questions will be important to address.

Conclusions

With the functional implications proposed for EVs, it is now vital to understand these vesicles themselves. As mentioned above, most studies published so far analyze mixed EV populations, and we think that some of the most important steps the field must take are to comprehensively compare the different subtypes of EVs and to determine whether some of their functions are specific or prominent in a given subtype, e.g., exosomes, but not other EVs. This knowledge is necessary to identify which EVs should be targeted for any therapeutic approach. Indeed, EV research is now at the stage where the immunology field was in the 1950s. At that time, researchers could only claim that circulating white blood cells were capable of very different functions, such as killing other cells or making antibodies, simply because there were no means to distinguish what we know now as B versus T lymphocytes! For EVs, a recent article reported a calculation that the number of copies of a given miRNA present per EV, in a mixed sEV preparation, is below one (Chevillet et al., 2014), suggesting that either very few miRNA molecules are present within each sEV or, more likely, that only a restricted subtype of EVs contain significant amounts of miRNA molecules and thus are capable of transferring miRNA-based information.

This situation may apply to other transfer functions of EVs, whether they involve mRNA or proteins: if only a minor subtype of EVs carries the relevant activity, its actual efficiency will be difficult to detect, as it will be undermined by an abundance of non-functional EVs present in the same preparation. One of the challenges is therefore to re-define methods that allow discrimination between sEVs, exosomes, and other EVs. It is impossible to distinguish them on the basis of a single property, such as size, structure, buoyant density, or presence of a given protein, and the community is seeking novel methods of isolation leading to better enrichment of a specific subtype. For this, however, better knowledge of specific markers of EV subtypes is required. We have recently performed a quantitative comparison of the protein composition of several subtypes of EVs secreted simultaneously by human primary dendritic cells, which were separated by a combination of differential ultracentrifugation, floatation in a density gradient, and immuno-isolation (Kowal et al., 2016). Although a large majority of the proteins were shared between all isolated EV subtypes, including some generally used as "exosome markers" (e.g., heat shock proteins, flotillins, major histocompatibility complex molecules), we are able to propose a few new specific markers of medium and large EVs (e.g., actinins), of endosome-derived

exosomes (co-expressing three tetraspanins CD9/CD63/CD81 and including TSG101 and syntenin-1), and of non-endosomal sEVs (some ITGs), whose validity as specific markers can be tested in all EV sources, which will hopefully enable further functional studies. Indeed, knowing which EV markers to follow will eventually allow identification of molecular tools to specifically affect secretion of a given subtype of EVs and thus understand the patho/physiological function of a particular subtype. We hope that the near future will thus provide the necessary technical advances and subsequent understanding of the various and fundamental roles of each type of EV. The potential development of these delivery packets for efficient therapeutic strategies in cancer and in many other diseases depends on these next steps.

ACKNOWLEDGMENTS

Because of space restriction, only a few recent publications were discussed in this Review; we apologize to all our colleagues whose work has not been cited here. We thank Joanna Kowal, Claire Hivroz, and Lionel Navarro for critical reading of the manuscript. This work has received support under the program "Investissements d'Avenir" launched by the French Government and implemented by ANR with the references ANR-11-LABX-0043 and ANR-10-IDEX-0001-02 PSL and by Fondation ARC pour la Recherche sur le Cancer through SL220120605293 grant to C.T. and post-doctoral grant to M.T.

REFERENCES

Alexander, M., Hu, R., Runtsch, M.C., Kagele, D.A., Mosbruger, T.L., Tolmachova, T., Seabra, M.C., Round, J.L., Ward, D.M., and O'Connell, R.M. (2015). Exosome-delivered microRNAs modulate the inflammatory response to endotoxin. Nat. Commun. *6*, 7321.

Assil, S., Webster, B., and Dreux, M. (2015). Regulation of the Host Antiviral State by Intercellular Communications. Viruses *7*, 4707–4733.

Basu, S., and Bhattacharyya, S.N. (2014). Insulin-like growth factor-1 prevents miR-122 production in neighbouring cells to curtail its intercellular transfer to ensure proliferation of human hepatoma cells. Nucleic Acids Res. *42*, 7170–7185.

Bobrie, A., Krumeich, S., Reyal, F., Recchi, C., Moita, L.F., Seabra, M.C., Ostrowski, M., and Théry, C. (2012). Rab27a supports exosome-dependent and -independent mechanisms that modify the tumor microenvironment and can promote tumor progression. Cancer Res. *72*, 4920–4930.

Boelens, M.C., Wu, T.J., Nabet, B.Y., Xu, B., Qiu, Y., Yoon, T., Azzam, D.J., Twyman-Saint Victor, C., Wiemann, B.Z., Ishwaran, H., et al. (2014). Exosome transfer from stromal to breast cancer cells regulates therapy resistance pathways. Cell *159*, 499–513.

Bridgeman, A., Maelfait, J., Davenne, T., Partridge, T., Peng, Y., Mayer, A., Dong, T., Kaever, V., Borrow, P., and Rehwinkel, J. (2015). Viruses transfer the antiviral second messenger cGAMP between cells. Science *349*, 1228–1232.

Buck, A.H., Coakley, G., Simbari, F., McSorley, H.J., Quintana, J.F., Le Bihan, T., Kumar, S., Abreu-Goodger, C., Lear, M., Harcus, Y., et al. (2014). Exosomes secreted by nematode parasites transfer small RNAs to mammalian cells and modulate innate immunity. Nat. Commun. *5*, 5488.

Caballero, J.N., Frenette, G., Belleannée, C., and Sullivan, R. (2013). CD9-positive microvesicles mediate the transfer of molecules to Bovine Spermatozoa during epididymal maturation. PLoS ONE *8*, e65364.

Chevillet, J.R., Kang, Q., Ruf, I.K., Briggs, H.A., Vojtech, L.N., Hughes, S.M., Cheng, H.H., Arroyo, J.D., Meredith, E.K., Gallichotte, E.N., et al. (2014). Quantitative and stoichiometric analysis of the microRNA content of exosomes. Proc. Natl. Acad. Sci. USA *111*, 14888–14893.

Clancy, J.W., Sedgwick, A., Rosse, C., Muralidharan-Chari, V., Raposo, G., Method, M., Chavrier, P., and D'Souza-Schorey, C. (2015). Regulated delivery of molecular cargo to invasive tumour-derived microvesicles. Nat. Commun. *6*, 6919.

Colombo, M., Raposo, G., and Théry, C. (2014). Biogenesis, secretion, and intercellular interactions of exosomes and other extracellular vesicles. Annu. Rev. Cell Dev. Biol. *30*, 255–289.

Corrigan, L., Redhai, S., Leiblich, A., Fan, S.J., Perera, S.M., Patel, R., Gandy, C., Wainwright, S.M., Morris, J.F., Hamdy, F., et al. (2014). BMP-regulated exosomes from Drosophila male reproductive glands reprogram female behavior. J. Cell Biol. *206*, 671–688.

Costa-Silva, B., Aiello, N.M., Ocean, A.J., Singh, S., Zhang, H., Thakur, B.K., Becker, A., Hoshino, A., Mark, M.T., Molina, H., et al. (2015). Pancreatic cancer exosomes initiate pre-metastatic niche formation in the liver. Nat. Cell Biol. *17*, 816–826.

de la Fuente, A., Alonso-Alconada, L., Costa, C., Cueva, J., Garcia-Caballero, T., Lopez-Lopez, R., and Abal, M. (2015). M-Trap: Exosome-Based Capture of Tumor Cells as a New Technology in Peritoneal Metastasis. J. Natl. Cancer Inst. *107*, djv184.

Gentili, M., Kowal, J., Tkach, M., Satoh, T., Lahaye, X., Conrad, C., Boyron, M., Lombard, B., Durand, S., Kroemer, G., et al. (2015). Transmission of innate immune signaling by packaging of cGAMP in viral particles. Science *349*, 1232–1236.

Hoshino, D., Kirkbride, K.C., Costello, K., Clark, E.S., Sinha, S., Grega-Larson, N., Tyska, M.J., and Weaver, A.M. (2013). Exosome secretion is enhanced by invadopodia and drives invasive behavior. Cell Rep. *5*, 1159–1168.

Hoshino, A., Costa-Silva, B., Shen, T.L., Rodrigues, G., Hashimoto, A., Tesic Mark, M., Molina, H., Kohsaka, S., Di Giannatale, A., Ceder, S., et al. (2015). Tumour exosome integrins determine organotropic metastasis. Nature *527*, 329–335.

Hsu, C., Morohashi, Y., Yoshimura, S., Manrique-Hoyos, N., Jung, S., Lauterbach, M.A., Bakhti, M., Grønborg, M., Möbius, W., Rhee, J., et al. (2010). Regulation of exosome secretion by Rab35 and its GTPase-activating proteins TBC1D10A-C. J. Cell Biol. *189*, 223–232.

Kowal, J., Arras, G., Colombo, M., Jouve, M., Morath, J.P., Primdal-Bengtson, B., Dingli, F., Loew, D., Tkach, M., and Théry, C. (2016). Proteomic comparison defines novel markers to characterize heterogeneous populations of extracellular vesicle subtypes. Proc. Natl. Acad. Sci. USA *113*, E968–E977.

Lai, C.P., Mardini, O., Ericsson, M., Prabhakar, S., Maguire, C.A., Chen, J.W., Tannous, B.A., and Breakefield, X.O. (2014). Dynamic biodistribution of extracellular vesicles in vivo using a multimodal imaging reporter. ACS Nano *8*, 483–494.

Lai, C.P., Kim, E.Y., Badr, C.E., Weissleder, R., Mempel, T.R., Tannous, B.A., and Breakefield, X.O. (2015). Visualization and tracking of tumour extracellular vesicle delivery and RNA translation using multiplexed reporters. Nat. Commun. *6*, 7029.

Le, M.T., Hamar, P., Guo, C., Basar, E., Perdigao-Henriques, R., Balaj, L., and Lieberman, J. (2014). miR-200-containing extracellular vesicles promote breast cancer cell metastasis. J. Clin. Invest. *124*, 5109–5128.

Lötvall, J., Hill, A.F., Hochberg, F., Buzás, E.I., Di Vizio, D., Gardiner, C., Gho, Y.S., Kurochkin, I.V., Mathivanan, S., Quesenberry, P., et al. (2014). Minimal experimental requirements for definition of extracellular vesicles and their functions: a position statement from the International Society for Extracellular Vesicles. J. Extracell. Vesicles *3*, 26913.

Luga, V., Zhang, L., Viloria-Petit, A.M., Ogunjimi, A.A., Inanlou, M.R., Chiu, E., Buchanan, M., Hosein, A.N., Basik, M., and Wrana, J.L. (2012). Exosomes mediate stromal mobilization of autocrine Wnt-PCP signaling in breast cancer cell migration. Cell *151*, 1542–1556.

Melo, S.A., Sugimoto, H., O'Connell, J.T., Kato, N., Villanueva, A., Vidal, A., Qiu, L., Vitkin, E., Perelman, L.T., Melo, C.A., et al. (2014). Cancer exosomes perform cell-independent microRNA biogenesis and promote tumorigenesis. Cancer Cell *26*, 707–721.

Nolte-'t Hoen, E.N., Buermans, H.P., Waasdorp, M., Stoorvogel, W., Wauben, M.H., and 't Hoen, P.A. (2012). Deep sequencing of RNA from immune cell-derived vesicles uncovers the selective incorporation of small non-coding RNA biotypes with potential regulatory functions. Nucleic Acids Res. *40*, 9272–9285.

Ostrowski, M., Carmo, N.B., Krumeich, S., Fanget, I., Raposo, G., Savina, A., Moita, C.F., Schauer, K., Hume, A.N., Freitas, R.P., et al. (2010). Rab27a and Rab27b control different steps of the exosome secretion pathway. Nat. Cell Biol. *12*, 19–30, 1–13.

Palma, J., Yaddanapudi, S.C., Pigati, L., Havens, M.A., Jeong, S., Weiner, G.A., Weimer, K.M., Stern, B., Hastings, M.L., and Duelli, D.M. (2012). MicroRNAs are exported from malignant cells in customized particles. Nucleic Acids Res. *40*, 9125–9138.

Pegtel, D.M., Cosmopoulos, K., Thorley-Lawson, D.A., van Eijndhoven, M.A., Hopmans, E.S., Lindenberg, J.L., de Gruijl, T.D., Würdinger, T., and Middeldorp, J.M. (2010). Functional delivery of viral miRNAs via exosomes. Proc. Natl. Acad. Sci. USA *107*, 6328–6333.

Peinado, H., Alečković, M., Lavotshkin, S., Matei, I., Costa-Silva, B., Moreno-Bueno, G., Hergueta-Redondo, M., Williams, C., García-Santos, G., Ghajar, C., et al. (2012). Melanoma exosomes educate bone marrow progenitor cells toward a pro-metastatic phenotype through MET. Nat. Med. *18*, 883–891.

Ridder, K., Keller, S., Dams, M., Rupp, A.-K., Schlaudraff, J., Del Turco, D., Starmann, J., Macas, J., Karpova, D., Devraj, K., et al. (2014). Extracellular vesicle-mediated transfer of genetic information between the hematopoietic system and the brain in response to inflammation. PLoS Biol. *12*, e1001874.

Ridder, K., Sevko, A., Heide, J., Dams, M., Rupp, A.-K., Macas, J., Starmann, J., Tjwa, M., Plate, K.H., Sültmann, H., et al. (2015). Extracellular vesicle-mediated transfer of functional RNA in the tumor microenvironment. OncoImmunology *4*, e1008371.

Sharma, U., Conine, C.C., Shea, J.M., Boskovic, A., Derr, A.G., Bing, X.Y., Belleannee, C., Kucukural, A., Serra, R.W., Sun, F., et al. (2016). Biogenesis and function of tRNA fragments during sperm maturation and fertilization in mammals. Science *351*, 391–396.

Sung, B.H., Ketova, T., Hoshino, D., Zijlstra, A., and Weaver, A.M. (2015). Directional cell movement through tissues is controlled by exosome secretion. Nat. Commun. *6*, 7164.

Thery, C., Amigorena, S., Raposo, G., and Clayton, A. (2006). Isolation and characterization of exosomes from cell culture supernatants and biological fluids. Curr. Protoc. Cell. Biol. *Chapter 3*, Unit 3.22.

Tosar, J.P., Gámbaro, F., Sanguinetti, J., Bonilla, B., Witwer, K.W., and Cayota, A. (2015). Assessment of small RNA sorting into different extracellular fractions revealed by high-throughput sequencing of breast cell lines. Nucleic Acids Res. *43*, 5601–5616.

Trajkovic, K., Hsu, C., Chiantia, S., Rajendran, L., Wenzel, D., Wieland, F., Schwille, P., Brügger, B., and Simons, M. (2008). Ceramide triggers budding of exosome vesicles into multivesicular endosomes. Science *319*, 1244–1247.

Van Deun, J., Mestdagh, P., Sormunen, R., Cocquyt, V., Vermaelen, K., Vandesompele, J., Bracke, M., De Wever, O., and Hendrix, A. (2014). The impact of disparate isolation methods for extracellular vesicles on downstream RNA profiling. J. Extracell. Vesicles *3*, 24858.

Vickers, K.C., Palmisano, B.T., Shoucri, B.M., Shamburek, R.D., and Remaley, A.T. (2011). MicroRNAs are transported in plasma and delivered to recipient cells by high-density lipoproteins. Nat. Cell Biol. *13*, 423–433.

Webber, J.P., Spary, L.K., Sanders, A.J., Chowdhury, R., Jiang, W.G., Steadman, R., Wymant, J., Jones, A.T., Kynaston, H., Mason, M.D., et al. (2015). Differentiation of tumour-promoting stromal myofibroblasts by cancer exosomes. Oncogene *34*, 290–302.

Wilson, R.C., and Doudna, J.A. (2013). Molecular mechanisms of RNA interference. Annu. Rev. Biophys. *42*, 217–239.

Yue, S., Mu, W., Erb, U., and Zöller, M. (2015). The tetraspanins CD151 and Tspan8 are essential exosome components for the crosstalk between cancer initiating cells and their surrounding. Oncotarget *6*, 2366–2384.

Zhang, L., Zhang, S., Yao, J., Lowery, F.J., Zhang, Q., Huang, W.-C., Li, P., Li, M., Wang, X., Zhang, C., et al. (2015). Microenvironment-induced PTEN loss by exosomal microRNA primes brain metastasis outgrowth. Nature *527*, 100–104.

Zhou, W., Fong, M.Y., Min, Y., Somlo, G., Liu, L., Palomares, M.R., Yu, Y., Chow, A., O'Connor, S.T.F., Chin, A.R., et al. (2014). Cancer-secreted miR-105 destroys vascular endothelial barriers to promote metastasis. Cancer Cell *25*, 501–515.

Zomer, A., Maynard, C., Verweij, F.J., Kamermans, A., Schäfer, R., Beerling, E., Schiffelers, R.M., de Wit, E., Berenguer, J., Ellenbroek, S.I., et al. (2015). In vivo imaging reveals extracellular vesicle-mediated phenocopying of metastatic behavior. Cell *161*, 1046–1057.

Received: 24 May 2017 | Accepted: 26 September 2017
DOI: 10.1002/mas.21550

REVIEW ARTICLE

WILEY MassSpectrometryReviews

Analysis of the plasma proteome using iTRAQ and TMT-based Isobaric labeling

Robert Moulder[1] | Santosh D. Bhosale[1] | David R. Goodlett[2] | Riitta Lahesmaa[1]

[1] Turku Centre for Biotechnology, University of Turku and Åbo Akademi University, Turku, Finland

[2] University of Maryland, Baltimore, Maryland

Correspondence
Robert Moulder, Turku Centre for Biotechnology, Biocity (7th floor), Tykistökatu 6, Turku, Finland FI-20520
Email: robmou@utu.fi

Over the past decade, chemical labeling with isobaric tandem mass tags, such as isobaric tags for relative and absolute quantification reagents (iTRAQ) and tandem mass tag (TMT) reagents, has been employed in a wide range of different clinically orientated serum and plasma proteomics studies. In this review the scope of these works is presented with attention to the areas of research, methods employed and performance limitations. These applications have covered a wide range of diseases, disorders and infections, and have implemented a variety of different preparative and mass spectrometric approaches. In contrast to earlier works, which struggled to quantify more than a few hundred proteins, increasingly these studies have provided deeper insight into the plasma proteome extending the numbers of quantified proteins to over a thousand.

KEYWORDS
disease research, isobaric labeling, immunodepletion, peptide fractionation, plasma, protein biomarkers

1 | INTRODUCTION

With the emergence of proteomics technology that has enabled the detailed characterization of complex organelles and biological fluids, plasma proteomics has maintained attention as a source of biomarkers for disease status and prediction.[1] To date a wide body of research has indicated that there may be in the order of thousands of proteins detectable in plasma,[2] however, the number of proteins that have been detected in multiple data sets is substantially less. Moreover, beyond qualitative representation a quantitative component is essential for any comparison. Among the quantitative proteomic methods that have been developed and applied, chemical labeling with isobaric tandem mass tags[3,4] has been among the most popular methods of the past decade.[5,6] The iTRAQ method and isobaric labeling in general has been previously covered in detailed reviews including some indication of their usage in biomarkers studies, to which the reader is referred for additional information.[5–7] In this review we present a critical overview of the applications in which stable isotope isobaric labeling has been employed for serum or plasma analysis, which, unless otherwise indicated, refer to human samples, and on the basis of the literature identified describe applications with the commercially available TMT and iTRAQ reagents only. We will use plasma as the general collective term for serum and plasma and refer the reader to specific details of the given references.

Isobaric peptide labeling strategies have predominantly targeted the N terminus and lysine residues (there are also cysteine reagents), using an N-hydroxysuccinimide-reactive group that promotes peptide aminolysis with substitution of a bipartite adduct consisting of a mass balance and reporter.[3] The intensity of distinct fragment reporter ions at different m/z values, which arise from the different isobaric combinations of heavy and light C, N, and O isotopes in the reagent's alternative forms, are detected in the tandem mass spectra of the pooled labeled samples and used to determine the relative contribution/abundance of peptides and thereby the relative quantification of proteins. Reagents have been produced to permit 2, 4, 6, 8, 10, and 12-plex comparisons. Through the use of isotopologues, where the 6.32 mDa mass difference between positional variations of ^{13}C and ^{15}N substitutions are combined (as are currently used in the 10 and

 © 2017 Wiley Periodicals, Inc.

12-plex reagents), multiplexing up to 18-plex has been suggested as a future possibility.[8]

A striking characteristic of plasma proteome datasets, with or without labeling, has often been the relatively small number of proteins identified in comparison to analyses of cells or tissues.[9] Until recently many of the published studies in which iTRAQ or TMT reagents were used in the analysis of plasma have only reported several hundred proteins (identified with ≥2 unique peptides), and often fewer being detected consistently in multiple experiments. These failings can be partly attributed to the heterogeneity of plasma between individuals and the wide abundance range of plasma proteins. On account of the latter, while much of the methodology for labeling, subsequent fractionation, and mass spectrometry may be alike with other isobaric labeling applications, additional considerations have been made in terms of sample preparation. To provide the reader with a concise overview of the technology and the types of applications, this review is divided into sections first describing the methodology used and then subdivided by the classes of disease and health concerns studied.

2 | METHODOLOGY

2.1 | Labeling and general protocols

In keeping with other proteomic applications using isobaric labeling, the key steps in plasma preparation and analysis are the same, that is, denaturation, digestion, labeling, fractionation, and mass spectrometric analysis. However, most frequently some kind of pre-fractionation has been performed to remove the most abundant proteins (Figure 1). Among the literature several publications have been reported in the format of general protocols for iTRAQ analysis of plasma.[10–13] In the earliest of these, Song et al[8] used a workflow typical of its era in which immuno-affinity depletion was followed by digestion, labeling, fractionation by strong cation exchange chromatography (SCX) and then LC-MS/MS with electrospray ionization (ESI). In this example only 73 proteins (identified by at least two peptides each) were quantified in all eight iTRAQ experiments. This level of proteome coverage pales in comparison to recent work where improvements in many aspects of the analytical instrumentation, including depletion, UHPLC separation and innovations in mass spectrometry, has provided quantification of over a thousand proteins.[14] Keshishian et al have further presented this approach as a general protocol.[15] In the work of Dayon et al they evaluated the merits of label-based and label free analysis of depleted and undepleted sera,[16] and in a subsequent publication presented an automated set-up for depletion, buffer exchange, digestion and TMT labeling followed by 1D separation and ESI analysis with an Orbitrap Elite (Thermo Scientific).[17] In a large scale practical application with the latter system, 365 proteins were identified in 304 TMT 6-plex experiments, however, only 110 proteins were identified in all 1005 samples.[18] The absence of additional fractionation after labeling is a likely cause for the limited depth and consistency of these plasma profiles. Similarly, Keshishian et al recently reported that for non-fractionated plasma samples depleted of the 14 most abundant proteins, in the order of 300 proteins could be quantified. Notably, this was increased to 600 quantified proteins when the scope of depletion was extended to ~60 target proteins.[15]

Several technically orientated articles have considered the merits of the different multiplex options and the comparison of preparative steps, such as depletion of the abundant proteins and fractionation of the labeled mixtures.[19–22] For example, in the study of Keshishian et al they evaluated whether any fewer proteins were identified with 6 and 10-plex TMT reagents, when compared to 4-plex iTRAQ. They found that 9% fewer proteins were detected with a TMT 6-plex experiment and 17% fewer were detected in the TMT 10-plex experiment. Earlier studies similarly indicated that protein identification is more successful using ESI with 4-plex reagents than 6 or 8-plex,[21] although these were not tested with plasma. Pottiez et al, however, reported in their LC-MALDI based plasma analysis that the 8-plex reagents were preferable to 4-plex, providing more consistent ratios without compromising protein identification.[22] To date, 4-plex labeling has been the most frequently used approach, as can be seen in Figure 1B.

While primarily designed for peptide labeling, labeling at the protein level has also been demonstrated with TMT and iTRAQ reagents.[23,24] However, it has also been noted that labeling of lysine residues negatively affects trypsin digestion. Sinclair and Timms evaluated different digestion combinations, trypsin and chymotrypsin versus trypsin and Glu-C, with the aim to improve the digestion efficiency of TMT labeled proteins from depleted serum.[24] Using a series of temporal samples from women with pancreatic ductal adenocarcinoma they employed two separate work flows that differed by the depletion system and fractionation of the labeled proteins. Overall a larger percentage of the proteins digested with trypsin and chymotrypsin were quantified.

In all situations, particularly with small lists of detected proteins and the comparison of a few sample pools, normalization of the data is a concern. To achieve comparison of multiple labeling experiments many researchers have used a common pooled reference material that is labeled and included in each set of measurements.[6] As an alternative approach, Herbrich et al presented a method that used the biological data itself to achieve normalization.[25] The advantages of this approach include the provision of extra multiplexing capacity (ie, no reference channel is required), the increased scope for changing experimental design and scale of the study, and the removal of the need to find a suitable quantity of the reference sample and concerns about its stability and reproducible analysis. They reported that their method gave more precise estimates of protein abundance, and subsequently applied this normalization approach in a wide scale plasma proteomics study of nutrition.[26] Similarly, although not demonstrated with plasma proteomics, Maes and co-workers presented a normalization method for isobaric labeling that does not require a reference sample.[27] To our knowledge no direct comparison has been made between the relative performances of these methods, although it could be viable to reanalyze and compare data from studies made with a pooled reference with the results using normalization derived from a data driven approach.

Sample handling is a vital consideration in the search for differentially abundant proteins. Hassis et al used iTRAQ to evaluate

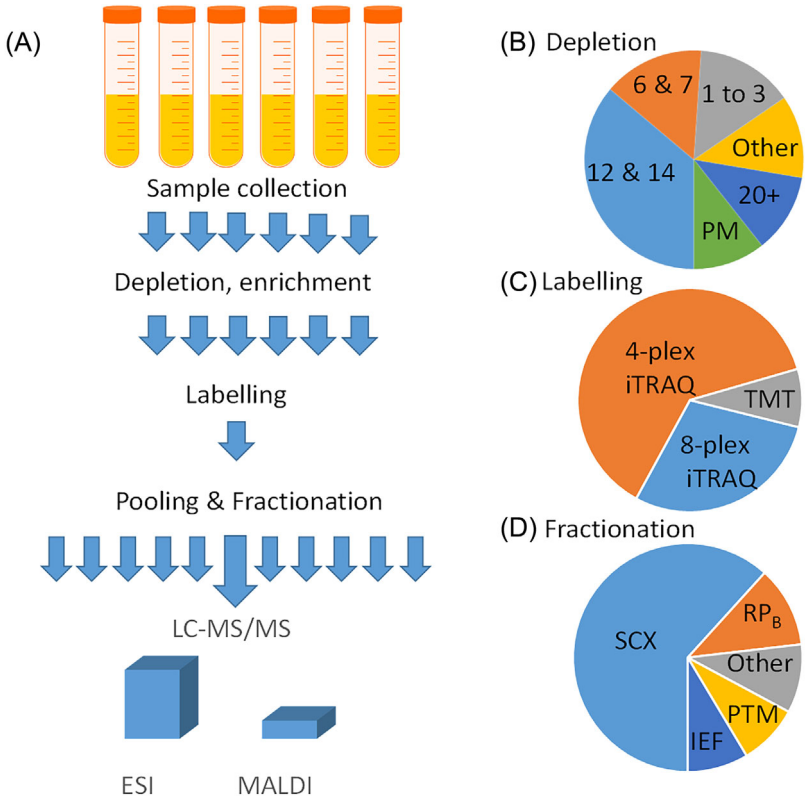


FIGURE 1 Plasma proteomics strategies and their usage. (A) A schematic is presented for a typical workflow for plasma proteomics with isobaric labeling. The accompanying pie charts, B, C, D, indicate the frequency that the methodologies have been used, and the bar chart indicates the relative usage of ESI and MALDI for mass spectrometry (~ 4:1), these are based on the information available from the literature reviewed in the manuscript (180–210 studies, depending on the information supplied). (B) Depletion strategies: the numerical values indicate the number of depletion targets, 20+ includes the both applications targeting the top 20 and the use of SuperMix columns, PM refers to ProteoMiner beads. (C) The figure indicates the type of reagents used, TMT combines applications with all forms of TMT reagents (including 2, 6 and10-plex). (D) The figure indicates the cumulative usage of the different fractionation strategies, IEF combines on and off-gel isoelectric focusing methods, AX combines different methods using anion exchange resins (i.e. HILIC & ERLIC) and RP_B refers to reversed phase fractionation at a basic pH

the effects of pre-analytical variables on the stability of the plasma proteome.[28] They determined the influence of the time delay (0.5, 6, and 96 h) and temperature (RT & 37°C) before processing, freeze thaw cycles and long term storage (14–17 years). They concluded that the time prior to processing was the most significant variable. Notably, although one of their tested pre-analytical storage conditions of four days at 37°C might seem somewhat unlikely, among the 83 differentially abundant proteins observed at this condition were many that have been frequently detected in biomarker studies. In contrast, on the basis of the results combined in the multi-centric study presented by Mateos et al,[29] comparing the effect of pre-analytical variables on the quality of plasma samples using different proteomic methods (including iTRAQ), they reported that delayed processing of blood samples and the number of freeze/thaw cycles had little effect on the integrity of proteins in the plasma samples. In another detailed study concerning sample handling, iTRAQ was used to evaluate the influence of serum clotting time upon protein abundance.[30] Here they found that in addition to changes with fibrinogen peptides, leucine-rich alpha-2-glycoprotein decreased with longer clotting times.

In summary for the general application of isobaric labeling, it has been reported that the lower the extent of multiplexing the larger number of proteins are identified and quantified (ie, 4-plex>6plex>8-plex). Further, depletion of the abundant proteins together with fractionation of the labeled peptides using an orthogonal separation method has so far been the only approach that has provided deeper plasma profiling.

2.2 | Depletion of abundant proteins

The plasma proteome spans a concentration range of 10 orders of magnitude[31]: from albumin, which represents around 50% of the protein mass, to low abundance cytokines, with the 22 most abundant proteins having been reported to account for up to 99%.[2] Subsequently, plasma proteomics data sets are frequently limited in the number of proteins identified and quantified, and although the disease state can be reflected by the moderately abundant proteins that are quite amenable to detection by general proteomics methodology, disease specific markers may be somewhat less abundant. To enable

the detection of lower abundance plasma proteins many researchers have used antibody-based affinity methods to remove (ie, immuno-deplete) the most abundant proteins. The depletion of the high abundance proteins has the advantage of reducing or avoiding overloading of the chromatographic and detection systems, and creating a sub-fraction of the plasma proteome with a reduced range of abundance.

Immuno-affinity based applications have ranged from removal of albumin only or albumin with IgG and evolved to the targeted removal of the top 3, 6, 7, 12, 14, or 20 most abundant proteins (Table 1). The frequency that these different approaches have been used in the literature presented in this review is indicated in Figure 1B. Among the major commercial products are the Multi Affinity Removal System, (MARS, Agilent) and the Seppro IgY products (developed by GenWay Biotech). The MARS systems use a mixture of polyclonal IgG antibodies attached to polymeric beads and the IgY products are derived from avian polyclonal immunoglobulin yolk (IgY) antibodies. In terms of their relative performance, lower cross-reactivity to non-target proteins and higher affinity for their targets have been reported for the IgY columns.[32] Tandem depletion methods have also been developed to extend the range of depletion targets to estimates of 60 and 155 proteins.[14,19] The more extensive depletion media have been developed by immunizing chickens to sera depleted of the top 12 or 14 proteins to produce antibodies that enable removal of a wider range of targets. These so called SuperMix columns (Sigma Aldrich) are used in combination with columns removing the original targets (ie, the IgY12 or IgY14). The enrichment achieved from these depletion steps has been estimated as 10-fold for the IgY14 alone, increasing to 100

TABLE 1 Depletion systems used in isobaric labeling proteomics analysis of plasma

Type	Targets	Examples
Agilent series (immuno-affinity)		
Mouse 3:	Albumin, IgG, transferrin	Jing et al[56], Chong et al[202]
MARS6	Albumin, transferrin, haptoglobin, IgG, IgA, and α1-antitrypsin	Boylan et al[202]
MARS7	MARS6 targets plus Fibrinogen	Luczak et al[42]
MARS14	MARS7 + α2-macroglobulin, α1-acid glycoprotein, complement C3, IgM, apolipoprotein AI, apolipoprotein AII, and transthyretin	Tremlett et al[147]
IgY series (Geneway) (immuno-affinity)		
IgY7	Albumin, IgG, fibrinogen, α1-antitrypsin, transferrin, haptoglobin and IgM	Takahashi et al[154]
IgY12	IgY7 + IgA, α2- macroglobulin, α1-Acid Glycoprotein, apolipoprotein A-I, apolipoprotein A-II	Boylan et al[202]
IgY14	IgY12 + Complement C3, Apolipoprotein B	Hollander et al[114]
IgY+SuperMix	IgY12/14 +60-150	Jones et al[34], Keshishian et al[14]
Other depletion systems (immuno-affinity)		
Proteo 20 (Sigma)	Albumin, IgG, transferrin, fibrinogen, IgA, α2- Marcroglobulin, IgM, α1- Antitrypsin, complement C3, haptoglobulin, apolipoprotein A1, apolipoprotein A2 and apolipoprotein B; α1- Acid Glycoprotein, ceruloplasmin, complement C4, C1q; IgD, transthyretin, and plasminogen.	Tonack et al[203]
Albumin Only	Human albumin	Mu et al[170]
	Bovine albumin	Faulkner et al[204]
Albumin & IgG	Albumin, IgG	Liu et al[60]
Not Immuno-affinity		
Poteominer- (hexapeptide beads)	Not applicable	Dwivedi et al[37]
Solvent precipitation	albumin	Seth et al[45]
SDS-PAGE		Zhang et al[43]
Size exclusion chromatography		Al-Daghri et al[39]
Undepleted		Song et al[41]

Plasma proteomics has mostly been performed with depletion of the most abundant proteins, typically using immuno-affinity resins. Examples of the type of immuno-affinity systems are listed with their targets, manufacturer and practical references. A number of these experiments involved xenographs (as indicated).

NOTE: α1- Acid Glycoprotein is often referred to as Orosomucoid, transthyretin as pre-albumin, Apolipoprotein B as the main component of LDL and apolipoprotein A-I, apolipoprotein A-II of the HDL.

fold for the combination.[32] Juhasz et al used iTRAQ together with a LC-MALDI MS/MS work flow and demonstrated that the SuperMix column provided consistent depletion and increased coverage of the plasma proteome.[33] They reported 870 proteins in total, with 480 detected in nine iTRAQ experiments. Patel et al used iTRAQ with ESI-LC-MS/MS and evaluated the off-target effects of two different depletion systems; one which removed albumin and immunoglobulins, and the other the SuperMix system, which they estimated removed 155 proteins.[19] They reported that the abundant proteins were quantified better with the removal of albumin and immunoglobulins only, while lower abundance proteins were better evaluated using the exhaustive immunodepletion. Although they only reported identification of 412 proteins in total (120 present in both the IgY12 and the IgY14 + SuperMix fractions), in studies using the same depletion strategy with more recent MS instrumentation, quantification of in the order of 1500 proteins has been reported.[34]

On the basis of gel-based comparison of the flow-through, bound fraction and untreated plasma, earlier publications have suggested that these depletion systems provide high efficiencies (ie, 99–100% removal). However, in our practical experience depletion is never complete and the depletion targets can still be detected by mass spectrometry with reasonable sequence coverage. The major depletion targets are biologically important plasma proteins, and their differential abundances often been reported in plasma labeling studies with and without depletion. Typical examples include haptoglobulin, alpha-1-antitrypsin, and alpha-2-macroglobulin. For example, Boichenko et al observed differential abundance of several immunodepletion targets (A1AT, HPT, TRFE, IGHA1, and IGHG3), for which the differences were confirmed for A1AT, HPT and TRFE in the undepleted serum.[35]

As an alternative to immuno-affinity depletion, Proteominer hexapeptide beads (BioRad) have been used.[20,36] This combinatorial library of hexapeptides theoretically presents a ligand for every protein. Here, by weight of numbers the sample is simplified as the more abundant proteins quickly saturate their targets and the excess pass through with the flow-through, while the lower abundance proteins are retained and subsequently released for analysis. ITRAQ-based comparisons of immuno-affinity depletion platforms and hexapeptide beads in the analysis of serum/plasma have reported similar numbers of proteins with both methods with some complementary and additional identifications.[37,38] There have, however, been some indications that the relative capture efficiencies of the hexapeptide beads can vary as much as two-fold with some proteins.[37]

Isobaric labeling analyses of plasma samples have also been made without immunodepletion or ligand affinity,[39–41] including direct labeling of undepleted plasma, as well as other modes of fractionation followed by labeling. For example, undepleted serum was TMT labeled, followed by on-line fractionation in a study of biomarkers of cardiac injury.[40,41] In another depletion-free approach, a three-dimensional fractionation strategy was used. This included protein fractionation by size exclusion chromatography, followed by iTRAQ peptide labeling, then high pH reverse phase fractionation and finally low pH (regular) reverse phase LC-MS/MS.[39] Here they indicated reproducible analysis of 2472 proteins, although they did not distinguish single peptide hits. Lukzac et al in their LC-MALDI-TOF/TOF based evaluation of the optimal methods for plasma sample pre-treatment prior to quantitative analysis using iTRAQ labeling, concluded that SCX chromatography without affinity depletion was the best pre-treatment method, reporting high reproducibility of the analysis of 1427 proteins.[42] SDS-PAGE has also been used as a pre-fractionation method prior to labeling.[43] As another alternative to change the range of protein abundance, solvent-based precipitation methods have been used.[44,45] In an alternative variant of solvent precipitation, Kodera et al used differential solubility to isolate native peptides from plasma. This was applied with 6-plex TMT reagents to compare pooled pre-surgery to post-surgery samples from renal cell carcinoma patients.[46]

In general, in order to profile the plasma proteome in depth it is necessary to reduce the range of protein abundance. To date, immuno-affinity depletion has provided the most convenient “off-the-shelf” method for this, and tandem depletion methods have been the most efficacious approach.

2.3 | Fractionation of labeled peptides

Following chemical labeling, sample fractionation is necessary to achieve accurate quantitative representation of the plasma proteome. This step increases the scope for identification and importantly can improve the accuracy of relative quantification by reducing the co-isolation of the parent ions for tandem mass spectrometry and the influence of chimeric spectra and mixed reporter ion signals.[5] The issue of ratio compression due to co-isolation of multiple precursors is among the key limitations of the use of isobaric labeling in quantification.[47] A number of solutions have been introduced to tackle this, including improved chromatographic fractionation,[48] real-time precursor filtering,[49] precursor charge manipulation,[49,50] and MS3 of the most intense precursor.[51]

Both off- and on-line fractionation with orthogonal separation methods have been employed to simplify labeled mixtures. Figure 1C indicates the relative frequency that the more common approaches have been used in the literature presented in this review. Although off-line strong cation exchange (SCX) fractionation has been used in the majority of studies and recommended in the manufacturer's instructions for both iTRAQ and TMT reagents, this is increasingly being replaced by other approaches, such as reversed-phase fractionation at a basic pH. For SCX fractionation, the labeled fractions are pooled, dried, acidified, either diluted or desalted, then eluted from a cationic stationary phase using a salt gradient and fractionated on the basis of size and charge. Non-volatile salts have been used for off-line fractionation (eg, KCl) and volatile salts are necessary for on-line systems (eg, ammonium formate). Removal of the labeling buffer (usually TEAB) prior to loading is important due to its cationic nature, as is the removal of the hydrolyzed reagents. Avoidance of salt concentrations greater than 10 mM is important to ensure recovery of the labeled peptides. The Polysulfoethyl columns from Poly-LC (Neste group) have been by far the most widely used column type. Notable examples where SCX fractionation has been

used include the study by Juhasz et al where it was used in combination with tandem depletion and LC-MALDI,[33] reporting 470 proteins detected in nine iTRAQ experiments, the ESI study of gastric adenocarcinoma by Subbannayya et al where they reported identification 643 proteins,[52] and the ESI & MALDI study of Zhang et al where they reported 1044 proteins.[53] SCX fractionation has mostly been performed off-line, necessitating additional concentration and desalting steps. Surprisingly, the on-line Multidimensional Protein Identification Technology (MudPIT) approach,[54] which is advantageous in terms of reduced sample handling and losses, has been less frequently employed. In the related studies of Song et al and Liu et al they used MudPIT in the analysis of TMT labeled undepleted serum[40,55] with which they reported identification of in the order of 380 proteins in both studies.

From a similar area of separation technology as SCX, anion exchange columns have been used to fractionate iTRAQ samples, including separation based on hydrophilic interaction chromatography (HILIC). Jing et al used both SCX and HILIC for pre-fractionation with analysis with ESI and LC-MALDI for an iTRAQ study of a mouse model of coronary heart disease.[56] While their data indicated that the two MS platforms gave complementary information, they combined the data from both fractionation methods and did not provide any comparison. Larkin et al implemented HILIC in a variation of the three dimensional depletion-free strategy initially described by Al-Daghri,[39] notably identifying in the order of a thousand proteins.[57] Another alternative using anion exchange columns for fractionation is Electrostatic Repulsion-Hydrophilic Interaction Chromatography (ERLIC), which was used by Datta et al in a study where they enriched plasma mircovesicles in a study of lacunar infarction.[58]

Reversed-phase fractionation at a basic pH (10–11) has emerged as popular alternative to SCX and was used in the work by Keshishian et al, in which in the order of 5000 proteins were reported from analysis of tandem depleted plasma.[14] The reversed phase separation and fractionation of peptides at a basic pH is orthogonal to the separation of ionic species at normal (acidic) pH. Typically a gradient of a basic aqueous solution of ammonium formate is run against an increasing percentage of acetonitrile. When performed with such a volatile elution buffer additional desalting can be omitted. In terms of published results the method is a superior alternative to SCX. As a further example, Liu et al used high pH reversed-phase fractionation for nine 10-plex TMT labeled plasma samples from ten subjects (10 TMT experiments), reporting identification of 1828 proteins with an impressive figure of 1009 proteins detected in all 90 samples. The samples were depleted of the fourteen most abundant proteins, and 24 fractions analyzed for each labeling experiment.

As a non-chromatographic option, peptide isoelectric focusing has been used with both immobilized pH gradients (IPG)[59–62] and off-gel approaches.[22,34,36,63,64] Using IPG strips, Pernemalm et al have demonstrated the potential of narrow-range peptide IEF for plasma analysis.[61] Among the examples of applications using off-gel IEF is the laboratory and instrument performance comparison by Jones et al in which plasma was first fractionated with tandem immuno-affinity columns, then the labeled peptides fractionated by IEF to create 30 fractions[34]; between 900 and 1300 proteins were quantified with the different ESI platforms.

Rather than fractionate on the basis of chromatographic or electrophoretic properties, Acosta-Martin et al used the data independent Precursor Acquisition Independent From Ion Count (PAcIFIC) approach, scanning selected mass ranges with a series of injections in the ESI analysis of plasma from abdominal aortic aneurysm (AAA) patients (n = 17) and controls (n = 17).[65] They use a combination of spectral counting and labeling with TMT reagents. A potential weakness in using DIA with labeled unfractionated samples is from the co-isolation of multiple precursors and hence mixed reporter ion signatures.

Overall, fractionation of the labeled samples is necessary to reduce ratio compression due to chimeric tandem mass spectra and enable a better representation of the plasma proteome. In the order of 20 to 30 fractions have been selected in the more successful applications. Here, among the orthogonal methods demonstrated, reversed-phase fractionation at a basic pH has emerged an alternative and improvement to the traditional SCX fractionation methods.

2.4 | Targeting post translational modifications

Most plasma proteins are glycoproteins and aberrant patterns of protein glycosylation have been implicated in many diseased states.[66] A number of research groups have used isobaric labeling to study these in cancer patients (*vide infra*). In these measurements the work flows have included immuno-depletion, digestion and labeling followed by affinity enrichment to capture the glycosylated peptides. Up to 350 glycosylated proteins have been identified including putative biomarkers, for example, of ovarian and pancreatic cancer.[67,68] Notably in the 2014 study by Nie et al TMT labeling was performed before digestion.[69] Also among these glycosylation targeting applications was the use of the Perkin-Elmer Exact-Tag 10-plex isobaric reagents in which fucosylated glycoproteins were identified in the serum of ovarian cancer patients.[68]

Kristensen et al used HILIC fractionation of iTRAQ labeled plasma followed by a titanium dioxide enrichment step that isolated both phosphopeptides and sialic acid-containing glycopeptides. With the combination of the identifications from the flow through and the enrichment, a total of 720 proteins were detected, including 220 represented by two or more peptides.

Protein carbonylation is another post-translational modification that has been investigated with isobaric labeling. This has been studied in plasma samples from breast cancer patients and in a study of T2D in a rat model.[70] To isolate proteins displaying this modification, biotin hydrazide was used to derivatize carbonyl groups and the biotinylated proteins subsequently enriched by avidin affinity, digested and labeled with iTRAQ reagents. In the study of carbonylation in breast cancer, the authors chose not to use immuno-depletion, in order to avoid unspecific binding, and reported identification and quantification of 160 carbonylated proteins.[70]

2.5 | Mass spectrometry and chromatography

The analyses of iTRAQ & TMT labeled samples have been conducted with both MALDI and ESI tandem mass spectrometry. Overall ESI has been used four times more frequently than MALDI (as indicated in Figure 1A). Initially, the analysis of iTRAQ labeled samples were made using instruments with time of flight mass analyzers (the Applied Biosystems instruments in particular), today, however, Orbitrap mass analyzers are clearly among the most common. For chromatographic separation mostly all the studies have used 75 μm i.d. packed capillary columns with nano-spray for ESI and spotting robots for MALDI. Earlier applications used 10 and 15 cm columns packed with 5 μm particles, although increasingly smaller (≤2 μm) UHPLC particles and longer columns (20–45 cm) are being used.

As a general indication of the capabilities of the recent generation of mass spectrometers for the ESI analysis of depleted and fractionated plasma samples, Jones et al compared the results of several typically equipped labs, including analyses using the Thermo Scientific Orbitrap-Velos and Q-Exactive and the AB Sciex TripleTOF 5600.[34] Based on the analysis of serum depleted of an estimated 155 proteins and fractionated by IEF (30 pooled fractions derived from 60 fractions), 1584 proteins were quantifiable from the combined data, of which 706 were common to the different platforms.[34] Keshishian et al exceeded this level of detail, reporting confident identification of over 5300 proteins, with 3390 quantified in multiple experiments.[14] In this study of microcardial infarction (in 16 subjects), tandem depletion of plasma samples (here they estimated removal of in the order of 60 proteins) together with high pH reverse phase fractionation of the labeled peptides were used. Furthermore, they reported that the coverage of the plasma proteome was five-fold better than their earlier label free approach.[71]

The number of proteins detected, quantified and reported is nevertheless dependent on the entire work flow from preparation and separation, the MS instrument and parameters used plus the choices made in data analysis (for example the database and filtering criteria). Without re-analyzing all these data with unified metrics, an accurate estimation of the relative success of these experiments is difficult. Notably, for example, in accompaniment to their recent TMT study of the maturing proteome, Liu et al re-analyzed Keshishian's data and stated that the breadth of the proteomic coverage was somewhat less, indicating that using their own parameters the extent of high confidence quantitative data was in the order of 1500–2500 proteins.[72] From their own 10-plex TMT data they reported 1800 proteins, with in the order of 1000 proteins detected in all reporter channels of their ten 10-plex TMT experiments.

The application of LC-MALDI for the analysis of isobarically labeled samples does not rely upon on-the-fly analysis, and thus pre- and post-defined acquisition methods can be used to avoid over-sampling. Also as the peptide ions generated are mostly singly charged, contrasting with the predominance of double and triple charged ions with ESI, crowding of m/z space and precursor overlap is reduced. Van der Greef et al reported clearly on the experimental design for iTRAQ experiments, alluding to datasets of 86 and 64 iTRAQ 4-plex experiments,[73] from which approximately 230 proteins were detected in more than half the samples. A notable example in terms of protein identification is the LC-MALDI work of Miike et al.[74] Using the combination of depletion of the six most abundant proteins, reversed phase protein fractionation, labeling and then strong cation exchange chromatography (SCX) of the labeled samples, they reported the identification of 4000 proteins and demonstrated gender differences (n = 12 vs 12) were detectable in this way. Whilst this was far ahead of the field in terms of the number of proteins detected at the time of publishing, notably there were two preparative steps before labeling and the labor intensive nature of the work could limit its applicability to multiple samples.

2.6 | Study size, discovery, and validation cohorts

The studies using isobaric labeling of plasma have tested the availability of detectable plasma biomarkers with the sample sizes ranging from a few individuals up to hundreds. Often the larger studies have involved the pooling of tens of samples, comparing the diseased or treated with pooled controls. Although the labeling necessitates fractionation, it is in this respect that when compared with label free methods isobaric labeling empowers biomarker discovery by permitting parallel processing and sub-fractionation of labeled samples. Here, the use of one of the reported ions as a reference channel overcomes issues associated with sample recovery and loading, and can accommodate changes in signal intensity. The analysis of suitably pooled samples can potentially deliver important information of lower abundance markers that necessitate follow up and validation. In view of the latter it is clear that the size of the discovery and validation cohorts is important.

Boichenko et al presented a discovery to validation pipeline for the identification of biomarkers from the serum of patients with cervical intraepithelial neoplasia (CIN) and squamous cell cervical cancer using iTRAQ.[35] In this they began with an iTRAQ-based discovery phase, followed by confirmation and verification with a label free approach, and finally validation with MRM. They equated this pipeline with the expectations of decreasing numbers of protein targets and increasing sample/subject numbers. Here, they identified six protein targets in the discovery phase (nine cervical cancer patients and 17 controls), then analyzed samples from 59 mixed cancer patients and 84 controls in verification measurements, and 180 mixed cancer patients and 50 controls in validation.

As another practical example, Gu et al investigated the availability of serum biomarkers of acute aortic dissection (AAD), comparing sera from 60 AAD patients with sera from 30 patients with acute myocardial infarction and 30 healthy volunteers.[75] On the basis of this data they identified lumican as a biomarker of vascular injury, which was supported by ELISA validation in the 120 subjects, although no independent cohort was tested. Similarly, Song et al used iTRAQ to compare plasma pools from patients with mild cognitive impairment (MCI, n = 261), Alzheimer's disease patients (n = 19) and cognitively normal subjects (n = 411). Western blotting was used to confirm that

afamin was decreased and IGHM increased in the sample pools from the same subjects.[55] In a later study they compared these data from an additional cohort of MCI subjects (n = 180) and controls (n = 153).

In their study of type 2 diabetes (T2D) Kaur et al used a randomized pooling strategy to compare T2D patients (n = 106) with controls (n = 76). They subsequently used Western blotting, SRM and qPCR to validate several differentially abundant proteins. Apolipoprotein A1, afamin, transthyretin (TTR), fibronectin, and vitamin D-binding protein (GC) were verified with MRM-based targeted proteomics in an independent cohort (n = 72).[76]

In a small scale study, Zhou et al considered the statistics and study design for biomarker discovery by plasma proteomics.[77] On the basis of their power calculations from the technical and biological variance of data from healthy volunteers and patients with pancreatic cancer, they suggested that six samples per group could provide sufficient statistical power for most of the proteins detected in their data with changes greater than two-fold.

2.7 | Differentially abundant proteins

A key aim of many of these studies has often been biomarker identification, for which lists of differentially abundant proteins have been reported that are typically in the order of ~15-20% of the detected proteome. For example, in many of the latter day applications, amounting to 35 out of 210 proteins (≥2 unique peptides), or presently in the order of hundreds of proteins and in instances with larger identification lists. As most frequently these have been cross sectional comparisons, that is, of a specific disease state versus control, it is noteworthy that in the analysis of the serum proteome in maturing children (~1-15 years of age), Liu et al observed temporal profiles for 900 hundred proteins,[72] underlining the importance of selecting suitable matched controls when comparing subjects.

While many plasma biomarker studies have strived to identify the lower abundance proteome, disease related patterns and differences in the moderately abundant proteome and depletion targets have frequently been reported. Moreover, since the fundamental processes moderated by the plasma proteome, such as inflammation and coagulation, are concomitant with a range of health disorders, common markers might be expected between similar or related diseases. Accordingly, the observation of the common proteins markers in different cancers and viral infections has been reported.[78–80] For example, alpha-1-acid glycoprotein 1 (ORM1) has been detected in non-small cell lung cancer[81] and gastric adenocarcinoma,[52] and apolipoprotein A-IV (APOA4) in ovarian[82] and pancreatic cancers.[67] Likewise, many studies have reported elevated inflammatory markers, such as C reactive protein (CRP) and/or Serum amyloid A-1 protein (SAA1).[81,83–85] For example, in plasma of patients with lung and liver cancer, and elsewhere in relation to circulation and cardiovascular risk,[41,86,87] including atherosclerosis and aortic aneurysms.

In Figure 2 we have summarized the top 20 most frequent markers on the basis of the available tabulated data from the articles included in this review. Examples among these include the differential abundance of plasma fibronectin, which has been reported in relation to hepatocellular carcinoma, abdominal aortic aneurysms, pulmonary tuberculosis, and conotruncal heart defect to name but a few. Similarly, differences in plasma gelsolin have been reported in comparisons of sepsis, aneurysmal subarachnoid hemorrhage, and post-stroke depression. The reoccurrence of the same markers in different diseases could be seen as a limitation to the utility of plasma proteomics. It should also be noted that since the representative peptides of each protein are often considered to represent a single entity that the details of potentially important proteoforms should not be overlooked.[88] Furthermore, the importance of lower abundant proteins may become more apparent as increasingly plasma protein lists grow from hundreds to thousands.[14,72]

In general, with the interpretation of plasma proteomics data it is important to recognize proteins frequently reported as putative biomarkers. Until validated these should be considered possible anomalies due to variations at the population level or from longitudinal changes or even from poorly characterized and unsuitable proteotypic peptides. Recent useful examples in the literature have considered changes with maturation[72,89] and hereditary differences on the basis of the studies of twins.[90] As a simpler directive, it is useful to recognize the depletion targets by their different aliases.

3 | APPLICATIONS OF ISOBARIC LABELING FOR SERUM AND PLASMA PROTEOMICS

It may be that having been the first commercially available variant of these isobaric labeling reagents that iTRAQ currently accounts for by far the largest proportion of the publications using this quantitative approach.[7] Collectively the applications of isobaric labeling of the plasma proteome have ranged from cancer to autoimmune diseases, as well as the influence of life style (smoking and nutrition) and environmentally influenced afflictions (occupational exposure). The references included in this review have not been restricted to human plasma, as for example ape models have been used to study the effect of viral infections,[91,92] xenographs have been made for tumor models,[93] and other animal-based disease models (rat or mouse). In addition to regular profiling of the plasma proteome, some studies have targeted glycosylation[94] and carbonylation.[70] These different fields of applications are summarized as follows and in Tables 2, 3, 4, and 5 and Figure 3.

3.1 | Cancer

There are more than 100 different types of cancer (named on the basis of the type of cell, organs or tissues where they occur), which account for in the order of 8 million deaths annually and have thus driven a large body of international research toward their prevention, diagnosis, and treatment. Accordingly there have been many cancer orientated proteomic studies that, on the basis of our literature review, cover one of the biggest subgroups of plasma proteomic isobaric labeling studies. The nature of these investigations is summarized in Table 3 and

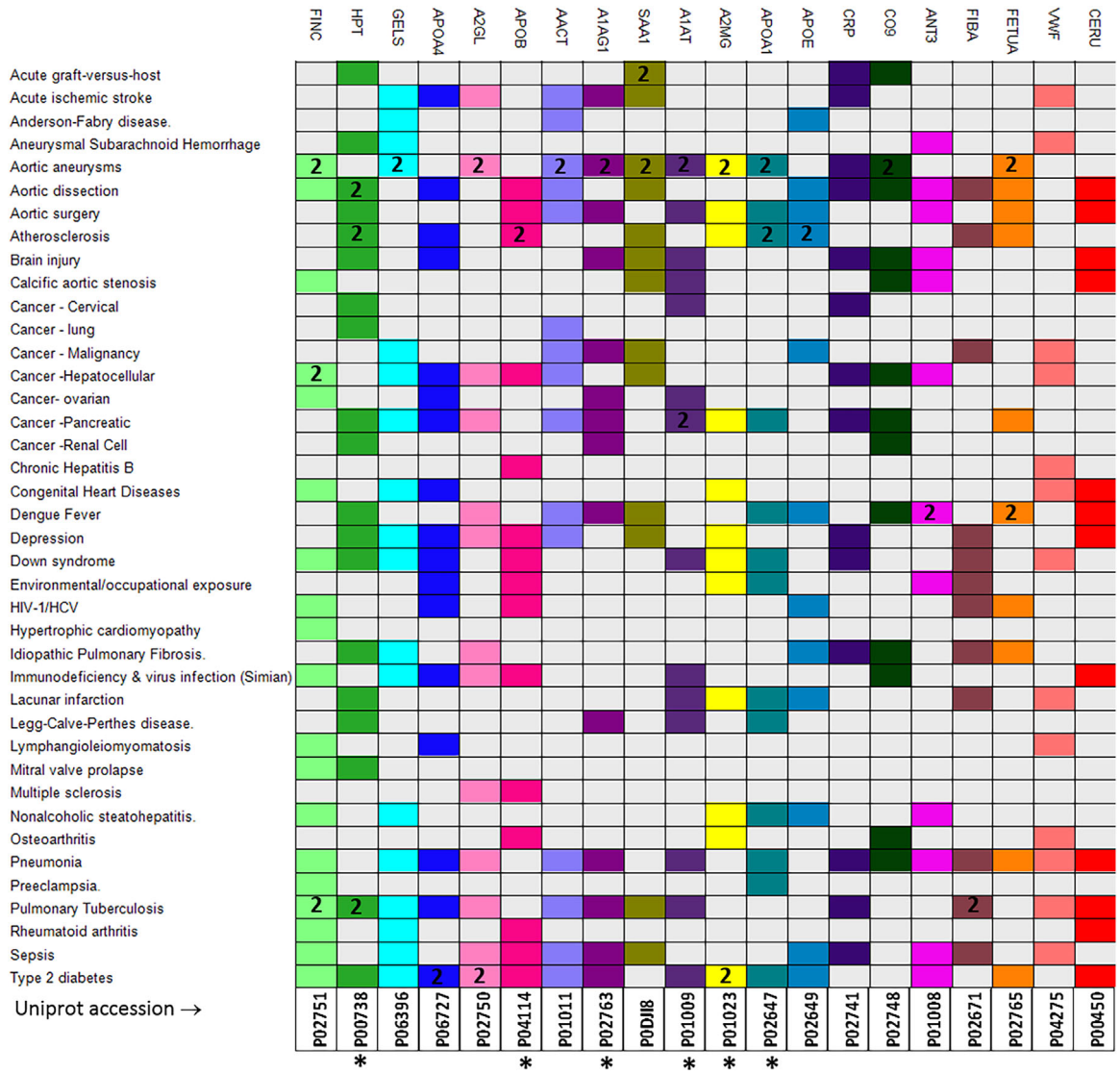


FIGURE 2 Candidate plasma protein biomarkers discovered in multiple diseases using isobaric labeling approach. Cumulatively there is a notable degree of reoccurrence of some proteins as biomarkers in different diseases and afflictions. As a representation of the top twenty most frequent biomarkers in each column each protein is represented by a unique color and mapped to diseases in which it has been reported as differentially abundant. The light grey color indicates no reported association, "2" indicates two publications with supporting evidence. The proteins are indicated with their UniProt accession numbers and IDs, where an asterisk (*) added to indicate proteins that are frequent immune-depletion targets

Figure 3. The analytical strategies that have been used have included most of the variants of sample preparation and analysis previously discussed, examples of these are indicated in Table 1. These have often been performed as pooled subject comparisons ranging from a few subjects to over 100, with validations in similar sized cohorts using ELISA, Western blot and SRM methods. Many of the validations have been healthy versus control, although some have included additional subgroups for cross validation. In such an example, Jenkinson et al used ESI and iTRAQ to search for serum biomarkers that might provide earlier diagnosis of pancreatic ductal adenocarcinoma (PDAC). Using a pooling strategy, iTRAQ enabled the comparisons of serum profiles from samples prior to and after PDAC diagnosis, chronic pancreatitis, and healthy controls (n = 150). MRM and/or Western blotting were used in validation using these and 322 additional human samples, including serum from patients with benign biliary disease, type 2 diabetes (T2D) and healthy controls. Lower levels of thrombospondin-1 (TSP-1) were identified prior to diagnosis of PDAC that together with a known cancer marker, CA19-9, provided improved risk classification.[95]

Evaluation of the conditions predisposing subjects or enhancing the risk of cancer have included the iTRAQ comparison of plasma from subjects with chronic hepatitis B virus infection—associated hepatocellular carcinoma, non-malignant cirrhosis, chronic hepatitis B, and healthy individuals.[60] In the latter study ESI analysis was used to compare labeled pools from 15 subjects of each category, with validations extending to a cohort of 310 subjects. Also studying

TABLE 2 Cancer studied using isobaric labeling of plasma

Cancer or target Type	References and subdivisions
Breast	Madian et al,[205] Meng et al,[206] Opstal-van Winden et al[207]
Blood	Human: Huang et al[208] mouse: Kristiansson et al[105]
Cervical Colorectal Polyposis Coli & Lynch syndrome	Boichenko et al[35] Zhang et al[209] Jones et al[34]
Gastric	Chong et al[93], Chong et al[210], Subbannayya et al[52] Xenograph: Chong et al[107]
Glioblastoma	Gautam et al[211]
Liver	He et al,[83] Huang et al,[212] Lee et al,[213] Li et al,[82] Qin et al,[101] Xenograph: Wang et al,[108] Hepatitis B: Liu et al,[60] Liu et al[96] Glycosylation: Yin et al[104]
Lung	Pernemalm et al,[61] Zhang et al,[81] Rice et al,[84] Jin et al,[214] Okano et al,[215]
Oral	Bijian et al,[106] Chai et al,[216] Yang et al[217] Zhang et al[43]
Ovarian	Human: Boylan et al,[202] Karabudak et al,[218] Li et al,[219] Shetty et al,[102] Wu et al,[68] Russell et al,[220] Wu et al,[97] Rat: Huang et al,[185]
Pancreatic	Nie et al,[67] Nie et al,[69] Sinclair and Timms,[24] Tan et al,[103] Tonack et al,[12] Tonack et al,[203] Zhou et al,[77] An et al,[98]
Prostrate	Rehman et al,[221] Larkin et al,[57] Zhang et al,[222]
Kidney	Zhang et al,[223]
Other: Chinese medicine, Endometrial hyperplasia vs. cancer, Glycosylation, Carbonylation	Upur et al,[224] Liu et al,[225] Wang et al,[159] Hu et al,[99] Nie et al,[100] Nie et al,[67] Qin et al[101] Shetty et al[102] Tan et al[103] Ueda et al[94] Wu et al[68] Yin et al[104] Liu et al[96] Madian et al[205]

Examples of the type of cancers studied with isobaric labeling and plasma proteomics are indicated with citations to published examples.

chronic Hepatitis B Virus-related hepatocellular carcinoma (HBV-related HCC), Liu et al used a glycoproteomic approach, indicating that Galectin 3 binding protein was a promising serum biomarker to identify HBV-related HCC.[96]

Studying the changes associated with response to treatment, Wu et al investigated proteomic and metabolic signatures of chemo-resistance in epithelial ovarian cancer, validating the differential abundance of FN1, SERPINA1, and ORM1.[97] Taking a somewhat more selective approach to study the effects of chemotherapy on pancreatic cancer, An et al isolated serum exosomes from patients and healthy controls. Using serum volumes of 4 mL per sample and ultracentrifugation they were able to isolate yields of exosomal proteins in the order of 0.2–0.6 μg.[98] After 4-plex iTRAQ labeling and LC-MS/MS with an Orbitrap Fusion they reported in the order of 700–800 proteins per sample (in total 1559 proteins were quantified), although it is notable that there was not any additional fractionation of the labeled mixtures.

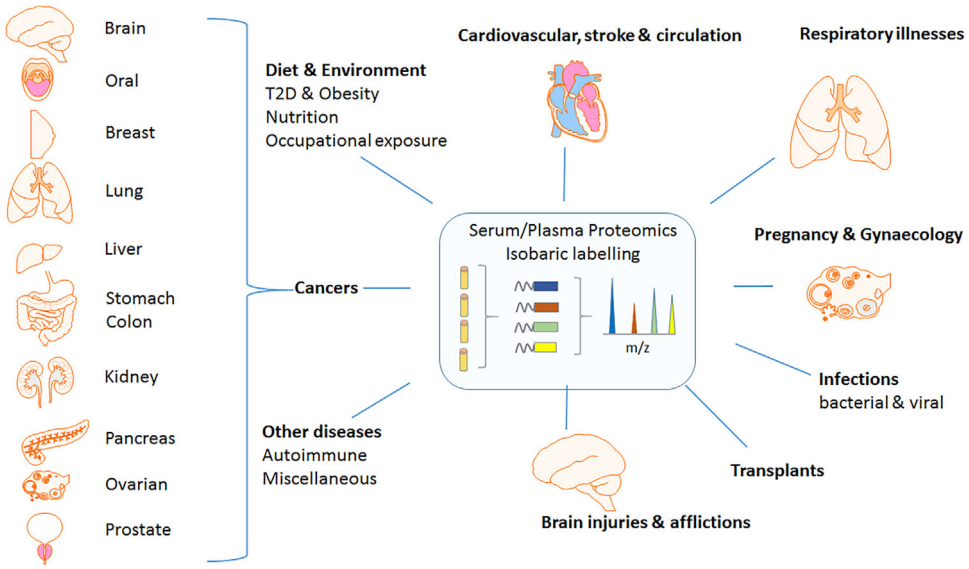


FIGURE 3 Subdivision of Isobaric labeling Plasma proteomic research by organ and disease. Plasma proteomics with isobaric labeling has been used to study diseases and conditions affecting a wide range of the body's organs. The figure provides a body-wise map of the different targets, including a separate subdivision for studies of cancer

TABLE 3 Examples of isobaric labeling plasma cancer studies

Author	Year	Depletion system	Fractionation	Spectrometer	Disease	Labeling	Proteins
Betancourt et al[184]	2014	IgYR7 (rat)	MudPIT	LTQ Orbitrap XL	Breast and prostate cancers	6-plex TMT	NA
Sinclair and Timms[24]	2011	Seppro 14 + SuperMix	SAX & 1D gel	LTQ Oribtrap XL	Pancreatic cancer	6-plex TMT	265
Subbannayya et al[52]	2015	Hu14	SCX	Orbitrap-Velos	Gastric adenocarcinoma	4-plex iTRAQ	643
He et al[83]	2014	Hu14	SCX,	MALDI, 5800	Hepatocellular carcinoma	4-plex iTRAQ	271
Meng et al[206]	2011	ProteoMiner	SCX	Agilent 6520 qTOF	Breast cancer	8-plex iTRAQ	397
Wang et al[108]	2011	MARS3 (mouse)	SCX	QSTAR XL TOF	Hepatocellular carcinoma	4-plex iTRAQ	554
Zhang et al[226]	2014	Hu14	High pH RP,	TripleTOF 5600	Lung cancer	4-plex iTRAQ	316
Liu et al[60]	2014	ALB & IgG	IEF	QSTAR TOF	Cancer/ hepatitis B virus	8-plex iTRAQ	328
Larkin et al[57]	2016	3D	HILIC	LTQ-Velos Pro Orbitrap Elite	prostate cancer	4-plex iTRAQ	1034
Zhang et al[222]	2016	ProteoMiner	SCX	TripleTOF 5600	prostate	4-plex iTRAQ	825
Walker et al[85]	2015	Hu14	High pH RP,	TripleTOF 5600	Non-small Cell Lung	8-plex iTRAQ	685

Examples are cited of plasma proteomic cancer studies indicating the approaches used and the scope of protein identification.

As an example of a rare cancer subtype, in the multilab ESI instrument benchmarking study of Jones et al, hereditary familial colorectal cancer has been studied, with samples from Familial Polyposis Coli syndrome patients ($n = 6$) and their controls ($n = 6$), and samples from Lynch syndrome and controls ($n = 5$).[34]

As indicated previously, not all these plasma orientated studies have been direct profiling measurements. Aberrant patterns of protein glycosylation have been implicated in carcinogenesis and a number of research groups have used isobaric labeling to study these in cancer patients.[68,69,94,99–104] These applications have targeted ovarian cancer, hepatocellular carcinoma and pancreatic cancer. Another post-translational modification considered has been carbonylation,[70] which was demonstrated in samples from breast cancer patients.

In addition to studies with samples from human subjects, animal models and other approaches have been employed. For example, Kristiansson et al used the immunodeficient SJL mouse to study B cell lymphomas[105] and rodent xenographs have been used is several instances,[93,106–108] including studies of gastric cancer and oral squamous cell carcinoma.

3.2 | Circulation and cardiovascular risk

Circulatory problems and the risk of stroke are a fundamental concern of the elderly, obese and increasingly with the young. Many aspects of circulation system and cardiovascular risk have been considered in terms of plasma proteomics with isobaric labeling (Table 5). These have included atherosclerotic disease,[86] abdominal and thoracic aneurysms,[65,87,109,110] arterial and pulmonary embolisms,[111] aortic dissection[75] haemorrhagic shock[112] and hypertrophic cardiomyopathy.[113] Investigating markers of cardiac injury, Liu and Song used an ape model together with isoproterenol treatment.[40,41]

Anderson-Fabry disease (AFD) is a rare genetic disease that is characterized by the lysosomal storage and is often manifested with heart related effects. Hollander et al compared AFD patients ($n = 32$) with healthy controls ($n = 14$) and reported gender-specific plasma protein biomarker panels that were specific and sensitive for the AFD phenotype.[114] Also, addressing a rare circulatory condition, Liu et al used iTRAQ to investigate the serum proteome of Legg-Calve-Perthes disease.[115] The panel of 26 differentially abundant proteins indicated that complement and coagulation cascades, and abnormal lipid metabolism might be involved in the pathogenesis of the disease. Western-blot was used to confirm the large differences in expression of S100-A8, apolipoprotein E (APOE) and two depletion targets ORM1 and HPT.

Several examples of studies of circulation were cited in the previous sections, such as the data independent PAcIFIC method in the TMT analysis of plasma from abdominal aortic aneurysm (AAA),[65] and the iTRAQ comparison of plasma proteins, sialic acid-containing glycopeptides and phosphopeptides in relation to atherosclerotic disease.[86]

Patients with chronic kidney disease (CKD) have a considerably higher risk of death due to cardiovascular causes. Luczak et al used

TABLE 4 Applications using isobaric labeling of plasma in the study of circulation, stroke, and cardiovascular risk

Circulation, stroke, and cardiovascular risk	
Aortic aneurysms, abdominal and thoracic aortic aneurysms:	Acosta-Martin et al,[65] Burillo et al,[109] Calvo et al,[110] Satoh et al,[87]
Aortic dissection:	Gu et al[75]
Deep vein thrombosis (microparticles):	Howes et al[111]
Haemorrhagic shock:	Li et al[112]
CVD/myocardial infarction	
CVD:	Juhasz et al[33]
Heart failure:	Lin et al[118]
Therapeutic, planned myocardial infarction:	Keshishian et al[14]
Mitral regurgitation:	Tan et al[122]
Coronary heart disease-human:	Yan et al[117]
Mouse model:	Jing et al[56]
Atherosclerosis	Kristensen et al,[86] Li et al[227]
Stroke	
Acute ischemic stroke.	Sharma et al[123]
Lacunar infarction:	Datta et al[58]
Post stroke depression:	Zhan et al[127]
Cerebral aneurysm:	Azurmendi et al[126] Saminathan et al[129]
Cardiac biomarker discovery using isoproterenol-treated nonhuman primates:	Song et al,[41] Liu et al[40]
Heart surgery	
Cardiopulmonary bypass surgery:	Oda et al[119]
Transplant:	Hollander et al,[191] Lin et al[121]
After valve replacement:	Satoh et al[120]
Legg-Calve-Perthes disease:	Liu et al[115]
Anders-Fabry's disease:	Hollander et al[114]

iTRAQ to investigate differences in the plasma of patients with CKD finding differences with proteins involved in lipid metabolism and atherosclerosis, including constituents of high-density lipoprotein (HDL) and low-density lipoprotein (LDL).[116] Notably they used a combination of analytical platforms (ESI & MALDI) and software, reporting 1038 proteins detected with two or more peptides from 10 iTRAQ experiments analyzing 180 samples with pooling.

Zhang et al investigated pulmonary arterial hypertension in relation to congenital heart diseases in a group of 140 children (3–11 years) with heart defects.[53] Detection and quantification of in the order of 1000 proteins were reported, 338 of which were detected with two or more unique peptides. One fifth of the proteins (~200) were differentially abundant among the six patient groups studied. ELISA was subsequently used to validate differences in carbamoyl-phosphate synthetase I (CPSI) and complement factor H-related protein 2 (CFHR2) in a cohort of 152 patients.

In the pursuit of pro-atherogenic HDL profiles in coronary heart disease patients, Yan et al progressed beyond direct plasma profiling and isolated the HDL fraction from plasma and depleted this of albumin and IgG prior to iTRAQ labeling.[53] With this targeted strategy they identified in the order of 200 proteins associated with the HDL fraction.

In relation to coronary artery disease (CAD), Jing et al used MALDI and ESI methods to compare the effects of dietary fat on the plasma proteome in an APOE knock out mouse model.[56] They were able to quantify more proteins with ESI (~400), and with the combined data from the platforms quantified in the order of 600 proteins, detecting differential abundance of several known CAD markers.

In studies on the effects of trauma upon the heart, researchers have probed for markers of myocardial infarction.[14,33,118] Included in these is the previously mentioned work of Juhasz and co-workers, in which they used the IgY and SuperMix depletion columns in series followed by off-line two-dimensional chromatography with fractions directly spotted on MALDI-TOF/TOF plates.[33] With this approach they used nine 8-plex iTRAQ experiments to compare pools (25 in each) from 252 subjects that experienced a myocardial infarction in the four subsequent years and 499 controls who suffered no event in the follow-up. In the order of 30 differential abundant proteins were detected, seven with known associations with CVD risk.

In relation to treatment and restoring the function of the heart, plasma profiles have been determined to compare different methods in aortic surgery,[119] to find markers of calcific aortic stenosis in patients after aortic valve replacement,[120] and find signatures of both cardiac allograft vasculopathy[121] and recovered heart function.[191] Other

TABLE 5 Illnesses, diseases, interventions and hazards studied by plasma proteomics with isobaric labeling:

A	
Respiratory-lung diseases and disorders	
Asthma/allergy:	Singh et al[130]
ARDS (acute respiratory distress syndrome):	Chen et al[59]
Obstructive sleep apnoea:	Jurado-Gamez et al[131]
Smoking:	Bortner et al[133]
Lymphangioleiomyomatosis (LAM):	Banville et al[132]
Autoimmune diseases	
Rheumatoid arthritis (RA):	Cheng et al,[140] Dwivedi et al,[141] Ortea et al,[142] Serada and Naka,[143] Yanagida et al[144]
RA & Crohns disease:	Serada et al[145]
Autoimmune hepatitis:	Li et al[152]
Multiple sclerosis:	Tremlett et al[147]
T1D:	Moulder et al,[149] Liu et al[72]
Diabetic nephropathy (T1D):	Overgaard et al,[150] Overgaard et al[151]
Other diseases and comparisons	
Osteoarthritis:	Fernandez-Puente et al[188]
Non-alcoholic fatty liver disease:	Miller et al[189]
Parkinson's disease:	Zhang et al[209]
Kidney-idiopathic nephrotic syndrome:	Andersen et al[190]
Anders-Fabry's disease:	Hollander et al[114]
Infection	
Sepsis-human:	Cao et al,[171] Cao et al[173]
Sepsis-rat:	Jiao et al[174]
Tuberculosis:	Li et al,[135] Wang et al,[136] Xu et al,[137] Xu et al[138]
Mycobacterial infections of cattle:	Seth et al[45]
Malaria:	Mu et al[170]
Viral infection	
Hepatitis B: hepatic fibrosis in hepatitis C:	Liu et al,[60] Peng et al,[176] Yang et al[178]
HIV:	Yang et al[177]
Dengue fever:	Shetty et al[179]
SIV (in apes):	Kumar et al,[64] Nhi et al[175]
SIV induced CNS disease:	Wiederin et al,[91] Wiederin et al[92]
	Pendyala et al[180]
B	
Gynaecological and pregnancy related	
Downs syndrome:	Kolla et al,[167] Sui et al[169]
Congenital heart defect:	Chen et al,[168] Zhang et al[81]
Preeclampsia:	Kolla et al[167]
Preeclampsia/intra-uterine growth restriction:	Auer et al,[164] Blankley et al[165]
Fertility (porcine and bovine):	Bijttebier et al,[228] Faulkner et al[204]
Endometrial hyperplasia:	Wang et al[159]
Endometriosis:	Xiaoyu et al[161]
Smoking and maternal serum:	Colquhoun et al[134]
Gestational diabetes:	Zhao et al[162]
T2D, obesity, diet, and nutrition	
T2D:	Kaur et al[76]

T2D + Very low calorie diet:	Sleddering et al[153]
Gastric bypass-related:	Jullig et al[156]
T2D/Obesity/Roux-en-Y gastric bypass:	Culnan et al[157]
T2D mouse:	Takahashi et al[154]
Diabetic Zucker rats (carbonylation):	Madian et al[205]
Diet and obesity:	Al-Daghri et al,[39] Cominetti et al[18]
Nutrition:	Cole et al[26]
Transplant	
Liver:	Parviainen et al[194]
Renal allograft rejection:	Freue et al,[192] Wu et al[193]
Heart/cardiac:	Hollander et al,[191] Lin et al[121]
Stem cell:	Lv et al,[195] Ye et al[38]
Diseases and afflictions of the brain	
Brain injury—human:	Hergenroeder et al[183]
Brain injury—mouse model:	Crawford et al[44]
Depression:	Xu et al[182]
Post stroke depression:	Zhan et al[127]
Alzheimer's versus mild cognitive impairment:	Muenchhoff et al,[181] Song et al[55]
Environmental exposure	
Rat exposure to bisphenol and genistein:	Betancourt et al[184]
Occupational health: benzene:	Huang et al[229]
Exposure to rare earth metals:	Liu et al[186]

investigations of the heart include a study of mitral valve prolapse, where Tan et al[122] evaluated whether plasma biomarkers could be used in the evaluation of mitral regurgitation in patients.

Considering the influence of cerebrovascular irregularities, the plasma proteomics of acute ischemic stroke[123] has been studied. Also regarding a subtype of the latter, that is, lacunar infraction, Datta et al enriched mircovesicles from plasma to determine whether differences in their proteomic profiles could be used to predict adverse outcome in the surviving patients.[58] The microvesicle fraction represents fragments of the plasma membrane that are shed from mostly all types of cells and include misfolded proteins, cytotoxic agents and metabolic waste. To enable the study of this informative sub-fraction, 5 mL of serum was processed. They used ERLIC to fractionate the labeled peptides. Whilst larger numbers of protein identifications might be expected from this sub-fraction,[124] the detail of their analysis was limited to several hundred proteins, although this was likely limited by their use of an early Applied Biosystems QSTAR TOF instrument.[58,125] In relation to stroke prognosis, Azurmendi and co-workers investigated biomarkers for survival from aneurysmal subarachnoid hemorrhage, comparing plasma from nosocomially infected (ie, pneumonia and urinary-tract) and non-infected patients.[126]

Beyond the effects more pertinent to circulation, Zhan et al considered post stroke depression.[127] Using a quite typical work flow, pooled sera was depleted of the top 14 proteins, labeled with 4-plex iTRAQ reagents and SCX fractionated prior to LC-MS/MS with an AB SCIEX TripleTOF 5600, in the order of 300 proteins were quantified. They observed that proteins involved in lipid metabolism and immunoregulation were differentially abundant. After validation by Western blotting concluded that the combination of increased levels of GSN accompanied by decreased HPT could be a promising as a plasma-based diagnostic biomarker panel post stroke depression.

Turning toward pharmaceutical toxicology, iTRAQ measurements of plasma were included with metabolomics methodologies in the study of drug toxicity in relation to the prevention and treatment of thromboembolic conditions.[128] An iTRAQ based LC-MS/MS pharmacoproteomics approach has also been used to analyze plasma protein profiles of 53 patients with a high (n = 28) and low dose of warfarin therapy (n = 25).[129]

3.3 | Respiratory illnesses

Among the studies of respiratory afflictions, researchers have considered asthma,[130] acute respiratory syndrome[59] and airway obstruction in relation to obstructive sleep apnoea.[131] In all three of these studies a MALDI/TOF/TOF approach was used with iTRAQ labeling. Banville et al used ESI with iTRAQ to search for serum biomarkers that could potentially distinguish lymphangioleiomyomatosis from other respiratory conditions.[132]

Considering illness from recreational drugs, Bortner and co-workers compared the plasma proteome of healthy non-smokers with

otherwise healthy cigarette smokers. Several of the differentially expressed proteins were already associated with tobacco related diseases, including lung cancer and chronic obstructive pulmonary disease.[133] Colquhoun et al investigated the effect of smoking on maternal serum from cord blood.[134] Although cord blood serum can be richer in intracellular proteins and more varied than matured serum, the study was conducted with older MS instrumentation (Applied Biosystems, QSTAR-Pulsar) and they reported quantification of in the order of 200 proteins. Other studies of respiratory afflictions have included tuberculosis[135–138] and Mycoplasma pneumoniae (MP) infections[139] (see below). Examples of these applications are indicated in Table 5A.

3.4 | Autoimmune diseases

Among the autoimmune diseases studied, rheumatoid arthritis (RA) has received some degree of attention.[140–145] Several of these RA studies have considered the effects of anti-TNF therapy using the drug infliximab. In comparison to healthy controls ($n = 50$), Serada et al observed elevated levels of Leucine-rich alpha-2-glycoprotein in RA ($n = 33$) and Crohn's disease patients ($n = 22$), and found these to decrease with infliximab therapy.[143–145] Fischer and co-workers used iTRAQ and metabolomics to identify biomarkers of ankylosing spondylitis.[146]

In a study of multiple sclerosis, Tremlett et al used a LC-MALDI approach to compare the serum of subjects with aggressive and benign multiple sclerosis, finding elevated leucine-rich alpha-2-glycoprotein in the aggressive state.[147]

Type 1 diabetes (T1D) is caused by the autoimmune destruction of the insulin producing cells in the pancreas, and is associated with genetic risk together with the factors from the environment. Research towards understanding the risk and causes of T1D has included several prospective sample collections from at-risk children.[148] In our own research we have used iTRAQ to compare the longitudinal serum proteome of T1D developing children ($n = 13$) with healthy controls ($n = 13$) matched by age, gender and risk, collected in the Finnish Diabetes Prediction and Prevention Study. The serum samples used ($n = 180$) spanned the period of early infancy through to seroconversion and on to diagnosis.[149] Based on 27 Labeling experiments, together with label free measurements from additional subjects ($n = 6$ vs 6, 86 samples), the data indicated that the moderately abundant serum proteome could be used to distinguish the subjects *en route* to developing T1D. Recently, Liu et al reported TMT analysis of plasma prospectively collected from at risk children enrolled with a T1D study in the USA, observing age related profiles for in the order of 900 proteins.[72] The latter analyses detailed the serum profiles of 10 healthy children from ~nine months of age to 15 years using 10 separate 10-plex TMT analyses of plasma depleted of the 14 most abundant proteins. Notably with their combination of high pH fractionation (off-line) and UHPLC analysis with a Q Exactive HF they detected 1800 proteins with quantitative information for ~1000 in all of the 90 samples measured. Other isobaric labeling studies related to T1D have considered complications associated with the disease state addressing diabetic nephropathy, in which samples from 123 subjects were used in the analysis of pooled samples.[150,151] This range of autoimmune diseases studied and others[152] are indicated in Table 5A.

3.5 | Type 2 diabetes, obesity, and diet

In contrast to T1D (*vide supra*), type 2 diabetes (T2D) is related to insulin resistance rather than the loss of the insulin producing cells. Although it is generally attributed to lifestyle, in particular diet and obesity, it is influenced by genetics, age and gender. Kaur et al (as cited earlier) used ESI analysis of iTRAQ labeled albumin and IgG depleted serum to compare pooled samples derived from T2D patients ($n = 70$) and controls ($n = 40$).[76] They used selected reaction monitoring (SRM) mass spectrometry to validate their finding in an independent cohort (36 vs 36). From their panel of validated markers both afamin and transthyretin provided area under the curve classification of 0.75, based on received operator characteristic curves.

Sleddering et al compared the influence of a very low calorie diet (VLCD) on obese T2D diabetes patients ($n = 27$) with some following an exercise program ($n = 13$) relative to lean and obese controls ($n = 2 \times 27$).[153] From eight 8-plex iTRAQ experiments the data indicated that 38% of the proteome was influenced by VLCD, from which several were selected for validation by SRM. They grouped and described the validated markers as a diet associated panel of apolipoproteins (ie, APOE, APO B-100, APOA-IV and APOC-III), T2D associated proteins (Fibrinogen alpha, beta and gamma chain and transthyretin) and obesity associated (complement 3). Other studies of T2D include the use of a mouse model[154] and plasma protein carbonylation in rats.[70]

Using samples collected from a large study cohort, Cominetti et al analyzed in the order of 1000 plasma samples from 525 individuals enrolled in a study related to diet, obesity and genetics (DiOGenes project).[18] Using a semi-automated platform and work flow (presented in an earlier publication by Dayon et al (*vide supra*),[17] a total of 1005 samples were analyzed. They observed that CRP and S100A9 were correlated with BMI and that pregnancy zone protein (PZP), SHBG and APOC-II provided potential gender discrimination. Using a mouse model to address β cell dysfunction and T2D, Kuo et al used iTRAQ to determine changes in the plasma proteome of mice with a genetic β-cell defect when subjected to an intravenous glucose/arginine challenge.[155]

In research related to T2D, Jullig et al used iTRAQ labeling of plasma to identify factors that favor resolution of T2D in subjects undergoing *roux-en-y* gastric bypass (GBP) surgery when compared with sleeve gastrectomy ($n = 15$).[156] Culnan and co-workers also considered changes in the plasma proteome associated with gastric bypass surgery.[157] Al-Daghri et al, in their 3D depletion free iTRAQ study (cited previously), demonstrated sexual dimorphism in obese subjects, comparing serum proteomes of nondiabetic overweight females ($n = 28$) and males ($n = 31$).[39]

Considering the socioeconomic limitations of nutrition, in their large scale isobaric labeling study, Cole et al used iTRAQ to identify

plasma proteins correlated to the micronutrient status in undernourished Nepalese children.[26] From a cohort of 1000, plasma from 500 children was used in iTRAQ analyses. Overall, seventy-three 8-plex iTRAQ labeling experiments were made. Of the 4705 proteins detected, 146 were quantified in all 500 children. From correlation analysis a number of nutrient-protein relationship associations were revealed. The established interrelationships observed could be useful in the interpretation of changes observed in other disease models, and in a follow up publication they reported on the correlation between biomarkers of inflammation, on the basis of correlations with α-1-acid glycoprotein.[158] The range of these studies related to T2D, obesity and diet are indicated in Table 5B.

3.6 | Pregnancy and gynecology

Isobaric labeling has been used to study the plasma proteome in the context of gynecological health, including the study of endometriosis, adenomyosis, endometrial hyperplasia and gestational diabetes (Table 5B). Wang et al compared serum from patients with different grades of endometrial hyperplasia and endometrial carcinoma in addition to healthy controls.[159] Xiaoyu et al used iTRAQ and ESI with an AB SCIEX TripleTOF 5600 to analyze the pooled serum depleted of albumin and IgG from women with adenomyosis (n = 20) and healthy controls (n = 20). Among the ~150 proteins quantified 25 were differentially abundant.[160] Later, Xiaoyu et al used the same iTRAQ approach to compare sera from adenomyosis and endometriosis patients (n = 20 vs 20).[161] They reported that proteins related to blood coagulation and complement activation were more abundant in adenomyosis, and those related to inflammatory response and the regulation of apoptosis were more abundant in endometriosis. The differential abundance of FN1, CD44, CFB, albumin (ALB) and fibrinogen alpha chain (FGA) were confirmed by Western-blotting.

Zhao et al have investigated biomarkers of gestational diabetes mellitus (GDM) using Proteominer depletion, iTRAQ and ESI (TripleTOF 5600). Samples from mothers with GDM were compared with normoglycemic pregnant controls (10 vs 10), and ELISA used for the validation of the observed differences (20 vs 20). The validated markers, APOE, F9, FGA, and IGFBP5, classified the subjects with high sensitivity and specificity.[162]

As well as mapping the changes in maternal serum during pregnancy,[163] the iTRAQ method has also been used to study complications during pregnancy, such as preeclampsia and intra-uterine growth restriction,[164–166] in addition to evaluating the effects of smoking on the cord blood serum proteome.[134] Auer and co-workers used iTRAQ to compare the plasma of women with preeclampsia and/or intra-uterine growth restriction with healthy controls (~40 subjects in total). Among the proteins differentially expressed with intra-uterine growth restriction were those from the complement and coagulation cascades.[164]

Analysis of the plasma proteome in relation to potential birth and genetic defects have included the iTRAQ-based comparisons of plasma from mothers bearing Down syndrome fetuses with plasma from uncomplicated pregnancies,[167] and serum of pregnant women carrying a fetus with conotruncal heart defects.[81,168] Sui et al compared umbilical cord blood from mothers carrying Down syndrome fetuses (n = 6) and healthy fetuses (n = 11). Using iTRAQ labeling, SCX fractionation and LC-MALDI they identified 505 proteins and reported five putative biomarkers.[169]

3.7 | Infection

Isobaric labeling-based studies of the plasma proteome in response to infection have ranged from bacterial, parasitic (ie, Malaria[170]) and viral encounters (see Table 5A). Bacterial studies include iTRAQ studies of pulmonary tuberculosis (PTB) in humans[135–138] and mycobacterial infections in cattle.[45] In the works of Xu et al and Wang et al, differences in serum proteome were studied in association with PTB.[137,138,171] Included in these was the use of a pooled iTRAQ strategy to compare treated, untreated and cured PTB patients with healthy controls (250 samples in total), in which markers of cured PTB were identified.[171] While the latter studies were performed using ESI (AB Sciex 5600), Li et al used MALDI-TOF/TOF (AB Sciex 5800) and iTRAQ to profile the serum of PTB patients, comparing sera from drug resistant, smear-positive and smear-negative pulmonary tuberculosis cases with normal controls and pneumonia patients (87 samples in iTRAQ discovery and 207 in the ELISA validations) and found that sex hormone binding globulin (SHBG) was significantly elevated in PTB.[135] Also considering bacterial infection and respiratory afflictions, Yu et al compared serum from healthy controls and patients with mycoplasma pneumoniae pneumonia.[139]

Cao and co-workers used iTRAQ labeling to analyze plasma from sepsis patients, including ten 4-plex iTRAQ comparisons of infected and uninfected subjects (n = 39 in total). These ESI data (Orbitrap-Velos) indicated age related differences in the incidence of infection.[172,173] In their study of sepsis in rats, Jiao et al correlated several biomarkers with diagnosis and others with prognosis.[174]

In terms of viral infections, Dengue fever,[64,175] Hepatitis B and C,[60,176–178] HIV[179] and Simian immunodeficiency virus (SIV)[91,92,180] have been studied using iTRAQ of plasma and serum samples. Liu et al compared plasma from individuals with chronic hepatitis B virus-associated hepatocellular carcinoma (HCC), non-malignant cirrhosis, chronic hepatitis B, and healthy individuals. They concluded that von Willebrand factor was a potential marker of HCC progression and viral infection.[60]

3.8 | Diseases, injuries and afflictions of the brain

The range of studies using isobaric labeling to study differences in the plasma proteome associated with afflictions of the brain is indicated in Table 5B. For example, Song et al performed iTRAQ based plasma profiling to identify markers that could provide a more specific diagnosis and prognosis of Alzheimer's disease (AD) and mild cognitive impairment (MCI), using plasma from 680 subjects.[55,181]

Turning from dementia to depression, Xu et al used iTRAQ of plasma samples to study major depressive disorder comparing a case to control group (n = 21 vs 21). In their validation experiments significant

differences were found in Apo-B100, A2M, VTDB, and Ceroplasmin.[182] Zhan et al used iTRAQ to study plasma profiles of subjects with post stroke depression, stroke, and healthy controls (n = 35 for each group).[127]

In an investigation of markers of traumatic brain injury, Hergenroeder et al compared serum protein profiles from patients with severe brain injury (n = 11) to age-, sex- and race-matched volunteers. SAA, CRP and RBP4 levels were identified and found to be robust indicators of injury.[183] Using instead a transgenic mouse model to identify plasma biomarkers of traumatic brain injury, Crawford et al used solvent based removal of albumin (chilled ethanol) prior to iTRAQ labeling and high pH reversed phase fractionation.[44] As this study was performed with an earlier Orbitrap mass spectrometer (XL-model), pulsed-Q dissociation (PQD) fragmentation was used to facilitate analysis of the iTRAQ reporter ions.

3.9 | Occupational exposure

A number of isobaric labeling studies have been made to evaluate the effects of occupational exposure on the plasma proteome. Using a rat model, Betancourt et al used TMT reagents to measure changes in the serum proteome when the animals were exposed to Bisphenol A and genistein.[184] The latter two are suspected to mimic endogenous hormones. A MudPIT approach was used in this study and a number of novel markers were identified and then cross validated by Western blot. Huang et al used iTRAQ to determine markers of benzene exposure in humans.[185] After ELISA validations they summarized that lowered expression of PBP and APOB served as potential biomarkers of chronic occupational benzene exposure. In an evaluation of occupational exposure of miners to rare earth elements Liu et al included iTRAQ analysis of serum.[186] They observed differential abundance of 29 proteins that could be related to the neurovirulence, hepatotoxicity, pathological fibrosis, osteoporosis, and anticoagulation caused by rare earth exposure.

3.10 | Other diseases and applications

ITRAQ and TMT analysis of plasma has been used to study number of other diseases, as are indicated in Table 5A These include Parkinson's disease,[187] osteoarthritis[188] and non-Alcoholic Fatty Liver Disease (NAFLD).[189] For example, Fernandez-Puente et al used iTRAQ with MALDI-TOF/TOF to profile sera obtained from 50 moderate osteoarthritis (OA) patients, 50 severe OA patients, and 50 non-symptomatic controls (5 pools of 10). Among the differentially abundant proteins detected were a number of complement components, lipoproteins, in addition to von Willebrand factor, tetranectin, and lumican.[188] Miller et al used both an iTRAQ and label free approach to study the serum proteome of sub classes of NFALD. Pooled samples made from controls (n = 10), simple steatosis (n = 7), nonalcoholic steatohepatitis (NASH) (n = 7), and NASH with fibrosis (n = 7).[189] Andersen et al studied both plasma and urine to identify biomarkers of idiopathic nephrotic syndrome.[190]

Some applications in which the analysis of the plasma proteome was made in relation to transplants were mentioned in the section on circulations and cardiovascular risk.[121,191] While in such specific instances it may be difficult to find sizeable study cohorts, liver, kidney, and stem cell transplantation have also been investigated. Freue et al used LC-MALDI-TOF/TOF and iTRAQ to identify plasma proteomic signatures of allograft rejection.[192] In this work they used 26 4-plex iTRAQ labeling experiments to analyze case and control samples (n = 11 vs 21). Of the 855 protein groups detected in the cumulative dataset, 144 were detected in two thirds of the samples. On the basis of subsequent ELISA measurements, four proteins, F9, SHBG, complement factor D (CFD), and LCAT, gave good classification of the subjects (tested in 10 patients and 19 controls). Wu et al used iTRAQ with ESI to compare depleted plasma from kidney transplant patients with allograft rejection patients (n = 5) against graft patients without rejection (n = 8). They reported that in the order of a third of the detected proteins (179 proteins detected) were two-fold different and that the dominant processes and responses were associated with inflammation and complement activation.[193] Parviainen et al have performed quantitative analysis of plasma from liver transplantation surgery.[194]

In their methodologically orientated study, Ye et al used SCX-LC-MALDI/TOF/MS/MS to analyze plasma samples from hematopoietic stem cell transplant patients.[38] They compared the plasma of these four patients at the time when they developed acute graft-versus-host disease (aGVHD) and after remission. Among the differentially abundant proteins CRP was observed with the largest fold difference. In a follow-up study from the same researchers using a larger study population they observed plasma ceruloplasmin as a potential plasma biomarker of aGVHD.[195] In this study they compared the technical performance of immuno-depletion and depletion using hexapeptide beads. Although the approaches produced complimentary data, more proteins were identified with immuno-depleted samples.

4 | DISCUSSION AND SUMMARY

During the past 15 years since their conception and commercial launch, isobaric labeling reagents have been used to analyze the plasma proteome in relation to many different diseases and afflictions. As indicated within this review, many of these have provided quantitative information for in the order of only a few hundred proteins and the same differentially abundant proteins can be common to several disease manifestations. Nevertheless, the information generated and the panels of markers can be important and beneficial within the specific contexts of these specific applications and there have been a growing number of more extensive proteomic characterizations.

In terms of competing methods, it is notable that in addition to throughput, the clear advantage of isobaric labeling at the time of its introduction was robustness to variations in sensitivity and chromatographic reproducibility, which are crucial for label free quantification. However, with the current generation of mass spectrometers

and chromatographs, label free profiling of undepleted plasma has been demonstrated with both DDA and DIA methods providing reasonable coverage and impressive throughput.[90,196,197] The use of labeling strategies nevertheless remains a powerful approach, facilitating the combined analysis of fractionated samples, particularly when executing complex preparative workflows. Alternatively, with the growing capabilities of SRM and PRM methods it is becoming increasingly possible to directly target wider panels of these proteins in non-depleted serum.[198–200] However, notably in the context of increasing the multiplexing capacity of such assays, Erickson et al have demonstrated the use of 10-plex TMT reagents with isotopically labeled synthetic peptide standards together with biological samples to improve the throughput of targeted assays.[201]

The throughput of labeling approaches is limited by whether depletion is performed and how digestion, labeling, fractionation (how many fractions?) and analysis (how many replicates, how much time is used to analyze each fraction?) are performed. Dayon et al demonstrated the scalability of depletion, TMT labeling and ESI analysis, with cited a throughput of 2 × 96 samples in 3 weeks and 1005 samples in 15 weeks.[17,18] Such measurements would allow plasma proteomics to catch up to some degree with genomics measurements. Although deeper profiling has yet to be demonstrated with the same throughput, it is notable that with the protein numbers reported in recent publications and the constant evolution of platforms for mass spectrometry, there are exciting new opportunities for biomarker discovery. Nevertheless, these measurements should be used in conjunction with current strategies to reduce compression of the determined protein abundance ratios. However, as can be seen from the earlier incremental improvements in protein identification from changes in depletion and fractionation strategies, and several generations of mass spectrometry platforms, plasma proteomics remains a formidable challenge.

ACKNOWLEDGMENTS

The National Technology Agency of Finland (Finland Distinguished Professor Programme, grant 40398/11), the Juvenile Diabetes Research Foundation (JDRF, 17-2011-586), the Academy of Finland (Centre of Excellence in Molecular Systems Immunology and Physiology Research (grant 250114), and Biocenter Finland (Turku Proteomics Facility) are thanked for their financial support.

REFERENCES

1. Pernemalm M, Lehtio J. Mass spectrometry-based plasma proteomics: state of the art and future outlook. *Expert Rev Proteomics*. 2014;11:431–448.
2. Nanjappa V, Thomas JK, Marimuthu A, et al. Plasma Proteome Database as a resource for proteomics research: 2014 update. *Nucl Acids Res*. 2014;42:D959–D965.
3. Ross PL, Huang YN, Marchese JN, et al. Multiplexed protein quantitation in Saccharomyces cerevisiae using amine-reactive isobaric tagging reagents. *Mol Cell Proteomics*. 2004;3:1154–1169.
4. Thompson A, Schafer J, Kuhn K, et al. Tandem mass tags: a novel quantification strategy for comparative analysis of complex protein mixtures by MS/MS. *Anal Chem*. 2003;75:1895–1904.
5. Evans C, Noirel J, Ow SY, et al. An insight into iTRAQ: where do we stand now? *Anal Bioanal Chem*. 2012;404:1011–1027.
6. Rauniyar N, Yates JR, 3rd. Isobaric labeling-based relative quantification in shotgun proteomics. *J Proteome Res*. 2014;13:5293–5309.
7. Westbrook JA, Noirel J, Brown JE, Wright PC, Evans CA. Quantitation with chemical tagging reagents in biomarker studies. *Proteomics Clin Appl*. 2015;9:295–300.
8. McAlister GC, Huttlin EL, Haas W, et al. Increasing the multiplexing capacity of TMTs using reporter ion isotopologues with isobaric masses. *Anal Chem*. 2012;84:7469–7478.
9. Tu C, Rudnick PA, Martinez MY, et al. Depletion of abundant plasma proteins and limitations of plasma proteomics. *J Proteome Res*. 2010;9:4982–4991.
10. Koutroukides TA, Jaros JA, Amess B, et al. Identification of protein biomarkers in human serum using iTRAQ and shotgun mass spectrometry. *Methods Mol Biol*. 2013;1061:291–307.
11. Song X, Bandow J, Sherman J, et al. ITRAQ experimental design for plasma biomarker discovery. *J Proteome Res*. 2008;7:2952–2958.
12. Tonack S, Aspinall-O'Dea M, Jenkins RE, et al. A technically detailed and pragmatic protocol for quantitative serum proteomics using iTRAQ. *J Proteomics*. 2009;73:352–356.
13. Tonack S, Neoptolemos JP, Costello E. Analysis of serum proteins by LC-MS/MS. *Methods Mol Biol*. 2010;658:281–291.
14. Keshishian H, Burgess MW, Gillette MA, et al. Multiplexed, quantitative workflow for sensitive biomarker discovery in plasma yields novel candidates for early myocardial injury. *Mol Cell Proteomics*. 2015;14:2375–2393.
15. Keshishian H, Burgess MW, Specht H, et al. Quantitative, multiplexed workflow for deep analysis of human blood plasma and biomarker discovery by mass spectrometry. *Nat Protoc*. 2017;12:1683–1701.
16. Dayon L, Kussmann M. Proteomics of human plasma: a critical comparison of analytical workflows in terms of effort, throughput and outcome. *EuPA Open Proteomics*. 2013;1:8–16.
17. Dayon L, Nunez Galindo A, Corthesy J, Cominetti O, Kussmann M. Comprehensive and scalable highly automated MS-Based proteomic workflow for clinical biomarker discovery in human plasma. *J Proteome Res*. 2014;13:3837–3845.
18. Cominetti O, Núñez Galindo A, Corthésy J, et al. Proteomic biomarker discovery in 1'000 human plasma samples with mass spectrometry. *J Proteome Res*. 2015;16:389–399.
19. Patel BB, Barrero CA, Braverman A, et al. Assessment of two immunodepletion methods: off-Target effects and variations in immunodepletion efficiency may confound plasma proteomics. *J Proteome Res*. 2012;11:5947–5958.
20. Pernemalm M, Orre LM, Lengqvist J, Wikstrom P, Lewensohn R, Lehtio J. Evaluation of three principally different intact protein prefractionation methods for plasma biomarker discovery. *J Proteome Res*. 2008;7:2712–2722.
21. Pichler P, Kocher T, Holzmann J, et al. Peptide labeling with isobaric tags yields higher identification rates using iTRAQ 4-plex compared to TMT 6-plex and iTRAQ 8-plex on LTQ Orbitrap. *Anal Chem*. 2010;82:6549–6558.
22. Pottiez G, Wiederin J, Fox HS, Ciborowski P. Comparison of 4-plex to 8-plex iTRAQ quantitative measurements of proteins in human plasma samples. *J Proteome Res*. 2012;11:3774–3781.
23. Wiese S, Reidegeld KA, Meyer HE, Warscheid B. Protein labeling by iTRAQ: A new tool for quantitative mass spectrometry in proteome research. *Proteomics*. 2007;7:340–350.
24. Sinclair J, Timms JF. Quantitative profiling of serum samples using TMT protein labelling, fractionation and LC-MS/MS. *Methods (San Diego, Calif.)* 2011;54:361–369.
25. Herbrich SM, Cole RN, West KP, Jr, et al. Statistical inference from multiple iTRAQ experiments without using common reference standards. *J Proteome Res*. 2013;12:594–604.

26. Cole RN, Ruczinski I, Schulze K, et al. The plasma proteome identifies expected and novel proteins correlated with micronutrient status in undernourished nepalese children. *J Nutr*. 2013;143:1540–1548.
27. Maes E, Hadiwikarta WW, Mertens I, Baggerman G, Hooyberghs J, Valkenborg D. CONSTANd: a normalization method for isobaric labeled spectra by constrained optimization. *Mol Cell Proteomics*. 2016;15:2779–2790.
28. Hassis ME, Niles RK, Braten MN, et al. Evaluating the effects of preanalytical variables on the stability of the human plasma proteome. *Anal Biochem*. 2015;478:14–22.
29. Mateos J, Carneiro I, Corrales F, et al. Multicentric study of the effect of pre-analytical variables in the quality of plasma samples stored in biobanks using different complementary proteomic methods. *J Proteomics*. 2017;150:109–120.
30. Govorukhina NI, de Vries M, Reijmers TH, Horvatovich P, van der Zee AG, Bischoff R. Influence of clotting time on the protein composition of serum samples based on LC-MS data. *J Chromatogr B Analyt Technol Biomed Life Sci*. 2009;877:1281–1291.
31. Anderson NL, Anderson NG. The human plasma proteome: history, character, and diagnostic prospects. *Mol Cell Proteomics*. 2002;1:845–867.
32. Shi T, Zhou J, Gritsenko MA, et al. IgY14 and SuperMix immunoaffinity separations coupled with liquid chromatography-mass spectrometry for human plasma proteomics biomarker discovery. *Methods (San Diego, Calif.)* 2011;56:246–253.
33. Juhasz P, Lynch M, Sethuraman M, et al. Semi-targeted plasma proteomics discovery workflow utilizing two-stage protein depletion and off-line LC-MALDI MS/MS. *J Proteome Res*. 2011;10:34–45.
34. Jones KA, Kim PD, Patel BB, et al. Immunodepletion plasma proteomics by tripleTOF 5600 and Orbitrap elite/LTQ-Orbitrap Velos/Q exactive mass spectrometers. *J Proteome Res*. 2013;12:4351–4365.
35. Boichenko AP, Govorukhina N, Klip HG, et al. A panel of regulated proteins in serum from patients with cervical intraepithelial neoplasia and cervical cancer. *J Proteome Res*. 2014;13:4995–5007.
36. Ernoult E, Bourreau A, Gamelin E, Guette C. A proteomic approach for plasma biomarker discovery with iTRAQ labelling and OFFGEL fractionation. *J Biomed Biotechnol*. 2010;2010:927917.
37. Dwivedi C, Krokhin OV, Cortens JP, Wilkins JA. Assessment of the reproducibility of random hexapeptide peptide library-based protein normalization. *J Proteome Res*. 2010;9:1144–1149.
38. Ye H, Sun L, Huang X, Zhang P, Zhao X. A proteomic approach for plasma biomarker discovery with 8-plex iTRAQ labeling and SCX-LC-MS/MS. *Mol Cell Biochem*. 2010;343:91–99.
39. Al-Daghri NM, Al-Attas OS, Johnston HE, et al. Whole serum 3D LC-nESI-FTMS quantitative proteomics reveals sexual dimorphism in the milieu interieur of overweight and obese adults. *J Proteome Res*. 2014;13:5094–5105.
40. Liu Y, Parman T, Schneider B, et al. Serum biomarkers reveal long-term cardiac injury in isoproterenol-treated African green monkeys. *J Proteome Res*. 2013;12:1830–1837.
41. Song B, Liu Y, Parman T, et al. Quantitative proteomics for cardiac biomarker discovery using isoproterenol-treated nonhuman primates. *J Proteome Res*. 2014;13:5909–5917.
42. Luczak M, Marczak L, Stobiecki M. Optimization of plasma sample pretreatment for quantitative analysis using iTRAQ labeling and LC-MALDI-TOF/TOF. *PLoS ONE*. 2014;9:e101694.
43. Zhang L, Jiang J, Arellano M, et al. Quantification of serum proteins of metastatic oral cancer patients using LC-MS/MS and iTRAQ labeling. *Open Proteomics J*. 2008;1:72–78.
44. Crawford F, Crynen G, Reed J, et al. Identification of plasma biomarkers of TBI outcome using proteomic approaches in an APOE mouse model. *J Neurotrauma*. 2012;29:246–260.
45. Seth M, Lamont EA, Janagama HK, et al. Biomarker discovery in subclinical mycobacterial infections of cattle. *PLoS ONE*. 2009;4:e5478.
46. Kodera Y, Hido Y, Kato R, et al. Establishment of a strategy for the discovery and verification of low-Abundance biomarker peptides in plasma using two types of stable-Isotope tags. *Mass Spectrom (Tokyo)*. 2014;3:S0044.
47. Ow SY, Salim M, Noirel J, Evans C, Rehman I, Wright PC. ITRAQ underestimation in simple and complex mixtures: "the good, the bad and the ugly. *J Proteome Res*. 2009;8:5347–5355.
48. Ow SY, Salim M, Noirel J, Evans C, Wright PC. Minimising iTRAQ ratio compression through understanding LC-MS elution dependence and high-resolution HILIC fractionation. *Proteomics*. 2011;11:2341–2346.
49. Wenger CD, Lee MV, Hebert AS, et al. Gas-phase purification enables accurate, multiplexed proteome quantification with isobaric tagging. *Nat Methods*. 2011;8:933–935.
50. Thingholm TE, Palmisano G, Kjeldsen F, Larsen MR. Undesirable charge-enhancement of isobaric tagged phosphopeptides leads to reduced identification efficiency. *J Proteome Res*. 2010;9:4045–4052.
51. Ting L, Rad R, Gygi SP, Haas W. MS3 eliminates ratio distortion in isobaric multiplexed quantitative proteomics. *Nat Methods*. 2011;8:937–940.
52. Subbannayya Y, Mir SA, Renuse S, et al. Identification of differentially expressed serum proteins in gastric adenocarcinoma. *J Proteomics*. 2015;127(Pt A):80–88.
53. Zhang X, Hou HT, Wang J, Liu XC, Yang Q, He GW. Plasma proteomic study in pulmonary arterial hypertension associated with congenital heart diseases. *Sci Rep*. 2016;6:36541.
54. Link AJ, Eng J, Schieltz DM, et al. Direct analysis of protein complexes using mass spectrometry. *Nat Biotech*. 1999;17:676–682.
55. Song F, Poljak A, Kochan NA, et al. Plasma protein profiling of Mild Cognitive Impairment and Alzheimer's disease using iTRAQ quantitative proteomics. *Proteome Sci*. 2014;12:5.
56. Jing L, Parker CE, Seo D, et al. Discovery of biomarker candidates for coronary artery disease from an APOE-knock out mouse model using iTRAQ-based multiplex quantitative proteomics. *Proteomics*. 2011;11:2763–2776.
57. Larkin SE, Johnston HE, Jackson TR, et al. Detection of candidate biomarkers of prostate cancer progression in serum: a depletion-free 3D LC/MS quantitative proteomics pilot study. *Br J Cancer*. 2016;115:1078–1086.
58. Datta A, Chen CP, Sze SK. Discovery of prognostic biomarker candidates of lacunar infarction by quantitative proteomics of microvesicles enriched plasma. *PLoS ONE*. 2014;9:e94663.
59. Chen X, Shan Q, Jiang L, Zhu B, Xi X. Quantitative proteomic analysis by iTRAQ for identification of candidate biomarkers in plasma from acute respiratory distress syndrome patients. *Biochem Biophys Res Commun*. 2013;441:1–6.
60. Liu Y, Wang X, Li S, et al. The role of von Willebrand factor as a biomarker of tumor development in hepatitis B virus-associated human hepatocellular carcinoma: a quantitative proteomic based study. *J Proteomics*. 2014;106:99–112.
61. Pernemalm M, De Petris L, Eriksson H, et al. Use of narrow-range peptide IEF to improve detection of lung adenocarcinoma markers in plasma and pleural effusion. *Proteomics*. 2009;9:3414–3424.
62. Pernemalm M, Lehtio J. A novel prefractionation method combining protein and peptide isoelectric focusing in immobilized pH gradient strips. *J Proteome Res*. 2013;12:1014–1019.
63. Chenau J, Michelland S, Sidibe J, Seve M. Peptides OFFGEL electrophoresis: a suitable pre-analytical step for complex eukaryotic samples fractionation compatible with quantitative iTRAQ labeling. *Proteome Sci*. 2008;6:9.

64. Kumar Y, Liang C, Bo Z, Rajapakse JC, Ooi EE, Tannenbaum SR. Serum proteome and cytokine analysis in a longitudinal cohort of adults with primary dengue infection reveals predictive markers of DHF. *PLoS Negl Trop Dis*. 2012;6:e1887.
65. Acosta-Martin AE, Panchaud A, Chwastyniak M, et al. Quantitative mass spectrometry analysis using PAcIFIC for the identification of plasma diagnostic biomarkers for abdominal aortic aneurysm. *PLoS ONE*. 2011;6:e28698.
66. Matei L. Plasma proteins glycosylation and its alteration in disease. *Rom J Intern Med*. 1997;35:3–11.
67. Nie S, Lo A, Wu J, et al. Glycoprotein biomarker panel for pancreatic cancer discovered by quantitative proteomics analysis. *J Proteome Res*. 2014;13:1873–1884.
68. Wu J, Xie X, Liu Y, et al. Identification and confirmation of differentially expressed fucosylated glycoproteins in the serum of ovarian cancer patients using a lectin array and LC-MS/MS. *J Proteome Res*. 2012;11:4541–4552.
69. Nie S, Yin H, Tan Z, et al. Quantitative analysis of single amino acid variant peptides associated with pancreatic cancer in serum by an isobaric labeling quantitative method. *J Proteome Res*. 2014;13:6058–6066.
70. Madian AG, Diaz-Maldonado N, Gao Q, Regnier FE. Oxidative stress induced carbonylation in human plasma. *J Proteomics*. 2011;74:2395–2416.
71. Addona TA, Shi X, Keshishian H, et al. A pipeline that integrates the discovery and verification of plasma protein biomarkers reveals candidate markers for cardiovascular disease. *Nat Biotechnol*. 2011;29:635–643.
72. Liu CW, Bramer L, Webb-Robertson BJ, Waugh K, Rewers MJ, Zhang Q. Temporal profiles of plasma proteome during childhood development. *J Proteomics*. 2017;152:321–328.
73. van der Greef J, Martin S, Juhasz P, et al. The art and practice of systems biology in medicine: mapping patterns of relationships. *J Proteome Res*. 2007;6:1540–1559.
74. Miike K, Aoki M, Yamashita R, et al. Proteome profiling reveals gender differences in the composition of human serum. *Proteomics*. 2010;10:2678–2691.
75. Gu G, Cheng W, Yao C, et al. Quantitative proteomics analysis by isobaric tags for relative and absolute quantitation identified Lumican as a potential marker for acute aortic dissection. *J Biomed Biotechnol*. 2011;2011:920763.
76. Kaur P, Rizk NM, Ibrahim S, et al. ITRAQ-Based quantitative protein expression profiling and MRM verification of markers in type 2 diabetes. *J Proteome Res*. 2012;11:5527–5539.
77. Zhou C, Simpson KL, Lancashire LJ, et al. Statistical considerations of optimal study design for human plasma proteomics and biomarker discovery. *J Proteome Res*. 2012;11:2103–2113.
78. Percy AJ, Chambers AG, Yang J, Borchers CH. Multiplexed MRM-based quantitation of candidate cancer biomarker proteins in undepleted and non-enriched human plasma. *Proteomics* 2013;13:2202–2215.
79. Polanski M, Anderson NL. A list of candidate cancer biomarkers for targeted proteomics. *Biomarker Insights* 2007;1:1–48.
80. Ray S, Patel SK, Kumar V, Damahe J, Srivastava S. Differential expression of serum/plasma proteins in various infectious diseases: specific or nonspecific signatures. *Proteomics Clin Appl*. 2014;8:53–72.
81. Zhang F, Lin H, Gu A, et al. SWATH- and iTRAQ-based quantitative proteomic analyses reveal an overexpression and biological relevance of CD109 in advanced NSCLC. *J Proteomics*. 2014;102:125–136.
82. Li L, Xu Y, Yu CX. Proteomic analysis of serum of women with elevated Ca-125 to differentiate malignant from benign ovarian tumors. *Asian Pac J Cancer Prev*. 2012;13:3265–3270.
83. He X, Wang Y, Zhang W, et al. Screening differential expression of serum proteins in AFP-negative HBV-related hepatocellular carcinoma using iTRAQ –MALDI-MS/MS. *Neoplasma*. 2014;61:17–26.
84. Rice SJ, Liu X, Miller B, et al. Proteomic profiling of human plasma identifies apolipoprotein E (APOE) as being associated with smoking and a marker for squamous metaplasia of the lung. *Proteomics*. 2015;15:3267–3277.
85. Walker MJ, Zhou C, Backen A, et al. Discovery and validation of predictive biomarkers of survival for non-small cell lung cancer patients undergoing radical radiotherapy: two proteins with predictive value. *EBioMedicine* 2015;2:839–848.
86. Kristensen LP, Larsen MR, Mickley H, et al. Plasma proteome profiling of atherosclerotic disease manifestations reveals elevated levels of the cytoskeletal protein vinculin. *J Proteomics*. 2014;101:141–153.
87. Satoh K, Maniwa T, Oda T, Matsumoto K. Proteomic profiling for the identification of serum diagnostic biomarkers for abdominal and thoracic aortic aneurysms. *Proteome Sci*. 2013;11:27.
88. Smith LM, Kelleher NL. Proteoform: a single term describing protein complexity. *Nat Methods*. 2013;10:186–187.
89. Bjelosevic S, Pascovici D, Ping H, et al. Quantitative age-specific variability of plasma proteins in healthy neonates, children and adults. *Mol Cell Proteomics*. 2017;16:924–935.
90. Liu Y, Buil A, Collins BC, et al. Quantitative variability of 342 plasma proteins in a human twin population. *Mol Syst Biol*. 2015;11:786.
91. Wiederin JL, Donahoe RM, Anderson JR, et al. Plasma proteomic analysis of simian immunodeficiency virus infection of rhesus macaques. *J Proteome Res*. 2010;9:4721–4731.
92. Wiederin JL, Yu F, Donahoe RM, Fox HS, Ciborowski P, Gendelman HE. Changes in the plasma proteome follows chronic opiate administration in simian immunodeficiency virus infected rhesus macaques. *Drug Alcohol Depend*. 2012;120:105–112.
93. Chong PK, Lee H, Loh MC, et al. Upregulation of plasma C9 protein in gastric cancer patients. *Proteomics*. 2010;10:3210–3221.
94. Ueda K, Takami S, Saichi N, et al. Development of serum glycoproteomic profiling technique; simultaneous identification of glycosylation sites and site-specific quantification of glycan structure changes. *Mol Cell Proteomics*. 2010;9:1819–1828.
95. Jenkinson C, Elliott V, Evans A, et al. Decreased serum thrombospondin-1 levels in pancreatic cancer patients up to 24 months prior to clinical diagnosis: association with diabetes mellitus. *Clin Cancer Res*. 2016;22:1734–1743.
96. Liu T, Liu D, Liu R, et al. Discovering potential serological biomarker for chronic Hepatitis B Virus-related hepatocellular carcinoma in Chinese population by MAL-associated serum glycoproteomics analysis. *Sci Rep*. 2017;7:38918.
97. Wu W, Wang Q, Yin F, et al. Identification of proteomic and metabolic signatures associated with chemoresistance of human epithelial ovarian cancer. *Int J Oncol*. 2016;49:1651–1665.
98. An M, Lohse I, Tan Z, et al. Quantitative proteomic analysis of serum exosomes from patients with locally advanced pancreatic cancer undergoing chemoradiotherapy. *J Proteome Res*. 2017;16:1763–1772.
99. Hu H, Gao L, Wang C, et al. Lower serum soluble-EGFR is a potential biomarker for metastasis of HCC demonstrated by N-glycoproteomic analysis. *Discov Med*. 2015;19:333–341.
100. Nie S, Lo A, Zhu J, Wu J, Ruffin MT, Lubman DM. Isobaric protein-level labeling strategy for serum glycoprotein quantification analysis by liquid chromatography-tandem mass spectrometry. *Anal Chem*. 2013;85:5353–5357.
101. Qin X, Chen Q, Sun C, et al. High-throughput screening of tumor metastatic-related differential glycoprotein in hepatocellular carcinoma by iTRAQ combines lectin-related techniques. *Med Oncol*. 2013;30:420.

102. Shetty V, Hafner J, Shah P, Nickens Z, Philip R. Investigation of ovarian cancer associated sialylation changes in N-linked glycopeptides by quantitative proteomics. *Clin Proteomics*. 2012;9:10.
103. Tan Z, Yin H, Nie S, et al. Large-scale identification of core-fucosylated glycopeptide sites in pancreatic cancer serum using mass spectrometry. *J Proteome Res*. 2015;14:1968–1978.
104. Yin H, Tan Z, Wu J, et al. Mass-Selected site-Specific core-Fucosylation of serum proteins in hepatocellular carcinoma. *J Proteome Res*. 2015;14:4876–4884.
105. Kristiansson MH, Bhat VB, Babu IR, Wishnok JS, Tannenbaum SR. Comparative time-dependent analysis of potential inflammation biomarkers in lymphoma-bearing SJL mice. *J Proteome Res*. 2007;6:1735–1744.
106. Bijian K, Mlynarek AM, Balys RL, et al. Serum proteomic approach for the identification of serum biomarkers contributed by oral squamous cell carcinoma and host tissue microenvironment. *J Proteome Res*. 2009;8:2173–2185.
107. Chong PK, Lee H, Zhou J, et al. ITIH3 is a potential biomarker for early detection of gastric cancer. *J Proteome Res*. 2010;9:3671–3679.
108. Wang C, Guo K, Gao D, et al. Identification of transaldolase as a novel serum biomarker for hepatocellular carcinoma metastasis using xenografted mouse model and clinic samples. *Cancer Lett*. 2011;313:154–166.
109. Burillo E, Lindholt JS, Molina-Sanchez P, et al. ApoA-I/HDL-C levels are inversely associated with abdominal aortic aneurysm progression. *Thromb Haemost*. 2015;113:1335–1346.
110. Calvo E, Martinez-Pinna R, Ramos-Mozo P, et al. Unraveling biomarkers of abdominal aortic aneurisms by iTRAQ analysis of depleted plasma. *Methods Mol Biol*. 2013;1000:157–166.
111. Howes JM. Proteomic profiling of plasma microparticles following deep-vein thrombosis. *Expert Rev Proteomics*. 2010;7:327–330.
112. Li Y, Liu B, Dillon ST, et al. Identification of a novel potential biomarker in a model of hemorrhagic shock and valproic acid treatment. *J Surg Res*. 2010;159:474–481.
113. Fucikova A, Lenco J, Tambor V, Rehulkova H, Pudil R, Stulik J. Plasma concentration of fibronectin is decreased in patients with hypertrophic cardiomyopathy. *Clin Chim Acta*. 2016;463:62–66.
114. Hollander Z, Dai DL, Putko BN, et al. Gender-specific plasma proteomic biomarkers in patients with Anderson-Fabry disease. *Eur J Heart Fail*. 2015;17:291–300.
115. Liu R, Fan L, Yin L, et al. Comparative study of serum proteomes in Legg-Calve-Perthes disease. *BMC Musculoskelet Disord*. 2015;16:281.
116. Luczak M, Formanowicz D, Marczak L, et al. ITRAQ-based proteomic analysis of plasma reveals abnormalities in lipid metabolism proteins in chronic kidney disease-related atherosclerosis. *Sci Rep*. 2016;6:32511.
117. Yan LR, Wang DX, Liu H, et al. A pro-atherogenic HDL profile in coronary heart disease patients: an iTRAQ labelling-based proteomic approach. *PLoS ONE*. 2014;9:e98368.
118. Lin D, Hollander Z, Meredith A, et al. Molecular signatures of end-stage heart failure. *J Card Fail*. 2011;17:867–874.
119. Oda T, Yamaguchi A, Yokoyama M, et al. Plasma proteomic changes during hypothermic and normothermic cardiopulmonary bypass in aortic surgeries. *Int J Mol Med*. 2014;34:947–956.
120. Satoh K, Yamada K, Maniwa T, Oda T, Matsumoto K. Monitoring of serial presurgical and postsurgical changes in the serum proteome in a series of patients with calcific aortic stenosis. *Dis Markers*. 2015;2015:694120.
121. Lin D, Hollander Z, et al. Plasma protein biosignatures for detection of cardiac allograft vasculopathy. *J Heart Lung Transplant*. 2013;32:723–733.
122. Tan HT, Ling LH, Dolor-Torres MC, Yip JW, Richards AM, Chung MC. Proteomics discovery of biomarkers for mitral regurgitation caused by mitral valve prolapse. *J Proteomics*. 2013;94:337–345.
123. Sharma R, Gowda H, Chavan S, et al. Proteomic signature of endothelial dysfunction identified in the serum of acute ischemic stroke patients by the iTRAQ-Based LC-MS approach. *J Proteome Res*. 2015;14:2466–2479.
124. Harel M, Oren-Giladi P, Kaidar-Person O, Shaked Y, Geiger T. Proteomics of microparticles with SILAC Quantification (PROMIS-Quan): a novel proteomic method for plasma biomarker quantification. *Mol Cell Proteomics*. 2015;14:1127–1136.
125. Datta A, Sze SK. Data for iTRAQ profiling of micro-vesicular plasma specimens: in search of potential prognostic circulatory biomarkers for Lacunar infarction. *Data Brief*. 2015;4:510–517.
126. Azurmendi L, Degos V, Tiberti N, et al. Measuring serum amyloid a for infection prediction in aneurysmal subarachnoid hemorrhage. *J Proteome Res*. 2015;14:3948–3956.
127. Zhan Y, Yang YT, You HM, et al. Plasma-based proteomics reveals lipid metabolic and immunoregulatory dysregulation in post-stroke depression. *Eur Psychiatry*. 2014;29:307–315.
128. Andersson U, Lindberg J, Wang S, et al. A systems biology approach to understanding elevated serum alanine transaminase levels in a clinical trial with ximelagatran. *Biomarkers*. 2009;14:572–586.
129. Saminathan R, Bai J, Sadrolodabaee L, et al. VKORC1 pharmacogenetics and pharmacoproteomics in patients on warfarin anticoagulant therapy: transthyretin precursor as a potential biomarker. *PLoS ONE*. 2010;5:e15064.
130. Singh A, Cohen Freue GV, Oosthuizen JL, et al. Plasma proteomics can discriminate isolated early from dual responses in asthmatic individuals undergoing an allergen inhalation challenge. *Proteomics Clin Appl*. 2012;6:476–485.
131. Jurado-Gamez B, Gomez-Chaparro JL, Munoz-Calero M, et al. Serum proteomic changes in adults with obstructive sleep apnoea. *J Sleep Res*. 2012;21:139–146.
132. Banville N, Burgess JK, Jaffar J, et al. A quantitative proteomic approach to identify significantly altered protein networks in the serum of patients with lymphangioleiomyomatosis (LAM). *PLoS ONE*. 2014;9:e105365.
133. Bortner JD, Richie JP, Das A, et al. Proteomic profiling of human plasma by iTRAQ reveals down-Regulation of ITI-HC3 and VDBP by cigarette smoking. *J Proteome Res*. 2011;10:1151–1159.
134. Colquhoun DR, Goldman LR, Cole RN, et al. Global screening of human cord blood proteomes for biomarkers of toxic exposure and effect. *Environ Health Perspect*. 2009;117:832–838.
135. Li C, He X, Li H, et al. Discovery and verification of serum differential expression proteins for pulmonary tuberculosis. *Tuberculosis (Edinb)*. 2015;95:547–554.
136. Wang C, Liu CM, Wei LL, et al. A group of novel serum diagnostic biomarkers for multidrug-Resistant tuberculosis by iTRAQ-2D LC-MS/MS and solexa sequencing. *Int J Biol Sci*. 2016;12:246–256.
137. Xu D, Li Y, Li X, et al. Serum protein S100A9, SOD3, and MMP9 as new diagnostic biomarkers for pulmonary tuberculosis by iTRAQ-coupled two-dimensional LC-MS/MS. *Proteomics*. 2015;15:58–67.
138. Xu DD, Deng DF, Li X, et al. Discovery and identification of serum potential biomarkers for pulmonary tuberculosis using iTRAQ-coupled two-dimensional LC-MS/MS. *Proteomics*. 2014;14:322–331.
139. Yu JL, Song QF, Xie ZW, et al. An iTRAQ-based quantitative proteomics study of refractory mycoplasma pneumoniae pneumonia patients. *Jpn J Infect Dis*. 2017;70:571–578.
140. Cheng Y, Chen Y, Sun X, et al. Identification of potential serum biomarkers for rheumatoid arthritis by high-resolution quantitative proteomic analysis. *Inflammation*. 2014;37:1459–1467.
141. Dwivedi RC, Dhindsa N, Krokhin OV, Cortens J, Wilkins JA, El-Gabalawy HS. The effects of infliximab therapy on the serum proteome of rheumatoid arthritis patients. *Arthritis Res Ther*. 2009;11:R32.

142. Ortea I, Roschitzki B, Ovalles JG, et al. Discovery of serum proteomic biomarkers for prediction of response to infliximab (a monoclonal anti-TNF antibody) treatment in rheumatoid arthritis: an exploratory analysis. *J Proteomics*. 2012;77:372–382.
143. Serada S, Naka T. Screening for novel serum biomarker for monitoring disease activity in rheumatoid arthritis using iTRAQ technology-based quantitative proteomic approach. *Methods Mol Biol*. 2014;1142:99–110.
144. Yanagida M, Kawasaki M, Fujishiro M, et al. Serum proteome analysis in patients with rheumatoid arthritis receiving therapy with tocilizumab: an anti-interleukin-6 receptor antibody. *BioMed Res Int*. 2013;2013:607137.
145. Serada S, Fujimoto M, Ogata A, et al. ITRAQ-based proteomic identification of leucine-rich alpha-2 glycoprotein as a novel inflammatory biomarker in autoimmune diseases. *Ann Rheum Dis*. 2010;69:770–774.
146. Fischer R, Trudgian DC, Wright C, et al. Discovery of candidate serum proteomic and metabolomic biomarkers in ankylosing spondylitis. *Mol Cell Proteomics*. 2012;11:M111.013904.
147. Tremlett H, Dai DL, Hollander Z, et al. Serum proteomics in multiple sclerosis disease progression. *J Proteomics*. 2015;118:2–11.
148. Moulder R, Bhosale SD, Lahesmaa R, Goodlett DR. The progress and potential of proteomic biomarkers for type 1 diabetes in children. *Expert Rev Proteomics*. 2017;14:31–41.
149. Moulder R, Bhosale SD, Erkkila T, et al. Serum proteomes distinguish children developing type 1 diabetes in a cohort with HLA-Conferred susceptibility. *Diabetes*. 2015;64:2265–2278.
150. Overgaard AJ, Hansen HG, Lajer M, et al. Plasma proteome analysis of patients with type 1 diabetes with diabetic nephropathy. *Proteome Sci*. 2010;8:4–5956-8-4.
151. Overgaard AJ, Thingholm TE, Larsen MR, et al. Quantitative iTRAQ-Based proteomic identification of candidate biomarkers for diabetic nephropathy in plasma of type 1 diabetic patients. *Clin Proteomics*. 2010;6:105–114.
152. Li H, Li G, Zhao X, et al. Complementary serum proteomic analysis of autoimmune hepatitis in mice and patients. *J Transl Med*. 2013;11:146.
153. Sleddering MA, Markvoort AJ, Dharuri HK, et al. Proteomic analysis in type 2 diabetes patients before and after a very low calorie diet reveals potential disease state and intervention specific biomarkers. *PLoS ONE*. 2014;9:e112835.
154. Takahashi E, Okumura A, Unoki-Kubota H, et al. Differential proteome analysis of serum proteins associated with the development of type 2 diabetes mellitus in the KK-A(y) mouse model using the iTRAQ technique. *J Proteomics*. 2013;84:40–51.
155. Kuo T, Kim-Muller JY, McGraw TE, Accili D. Altered plasma profile of antioxidant proteins as an early correlate of pancreatic beta cell dysfunction. *J Biol Chem*. 2016;291:9648–9656.
156. Jullig M, Yip S, Xu A, et al. Lower fetuin-A, retinol binding protein 4 and several metabolites after gastric bypass compared to sleeve gastrectomy in patients with type 2 diabetes. *PLoS ONE*. 2014;9:e96489.
157. Culnan DM, Cooney RN, Stanley B, Lynch CJ. Apolipoprotein A-IV, a putative satiety/antiatherogenic factor, rises after gastric bypass. *Obesity (Silver Spring)*. 2009;17:46–52.
158. Lee SE, West KP, Jr, Cole RN, et al. Plasma proteome biomarkers of inflammation in school aged children in Nepal. *PLoS ONE*. 2015;10:e0144279.
159. Wang YS, Cao R, Jin H, et al. Altered protein expression in serum from endometrial hyperplasia and carcinoma patients. *J Hematol Oncol*. 2011;4:15.
160. Xiaoyu L, Weiyuan Z, Ping J, Anxia W, Liane Z. Comparative serum proteomic analysis of adenomyosis using the isobaric tags for relative and absolute quantitation technique. *Fertil Steril*. 2013;100:505–510.
161. Xiaoyu L, Weiyuan Z, Ping J, Anxia W, Liane Z. Serum differential proteomic analysis of endometriosis and adenomyosis by iTRAQ technique. *Eur J Obstet Gynecol Reprod Biol*. 2014;182:62–65.
162. Zhao D, Shen L, Wei Y, et al. Identification of candidate biomarkers for the prediction of gestational diabetes mellitus in the early stages of pregnancy using iTRAQ quantitative proteomics. *Proteomics Clin Appl*. 2017;11: https://doi.org/10.1002/prca.201600152.
163. Scholl PF, Cole RN, Ruczinski I, et al. Maternal serum proteome changes between the first and third trimester of pregnancy in rural southern Nepal. *Placenta*. 2012;33:424–432.
164. Auer J, Camoin L, Guillonneau F, et al. Serum profile in preeclampsia and intra-uterine growth restriction revealed by iTRAQ technology. *J Proteomics*. 2010;73:1004–1017.
165. Blankley RT, Fisher C, Westwood M, et al. A label-free selected reaction monitoring workflow identifies a subset of pregnancy specific glycoproteins as potential predictive markers of early-onset pre-eclampsia. *Mol Cell Proteomics*. 2013;12:3148–3159.
166. Kolla V, Jeno P, Moes S, Lapaire O, Hoesli I, Hahn S. Quantitative proteomic (iTRAQ) analysis of 1st trimester maternal plasma samples in pregnancies at risk for preeclampsia. *J Biomed Biotechnol*. 2012;2012:305964.
167. Kolla V, Jeno P, Moes S, et al. Quantitative proteomics analysis of maternal plasma in Down syndrome pregnancies using isobaric tagging reagent (iTRAQ). *Biomed Res Int*. 2010;952047.
168. Chen L, Gu H, Li J, et al. Comprehensive maternal serum proteomics identifies the cytoskeletal proteins as non-invasive biomarkers in prenatal diagnosis of congenital heart defects. *Sci Rep*. 2016;6:19248.
169. Sui W, Zhang R, Chen J, et al. Quantitative proteomic analysis of Down syndrome in the umbilical cord blood using iTRAQ. *Mol Med Rep*. 2015;11:1391–1399.
170. Mu AK, Bee PC, Lau YL, Chen Y. Identification of protein markers in patients infected with Plasmodium knowlesi, Plasmodium falciparum and Plasmodium vivax. *Int J Mol Sci*. 2014;15:19952–19961.
171. Wang C, Wei LL, Shi LY, et al. Screening and identification of five serum proteins as novel potential biomarkers for cured pulmonary tuberculosis. *Sci Rep*. 2015;5:15615.
172. Cao Z, Yende S, Kellum JA, Angus DC, Robinson RA. Proteomics reveals age-related differences in the host immune response to sepsis. *J Proteome Res*. 2013;13:422–432.
173. Cao Z, Yende S, Kellum JA, Robinson RA. Additions to the human plasma proteome via a tandem MARS depletion iTRAQ-Based workflow. *Int J Proteomics*. 2013;2013:654356.
174. Jiao J, Gao M, Zhang H, et al. Identification of potential biomarkers by serum proteomics analysis in rats with sepsis. *Shock (Augusta, Ga.)*. 2014;42:75–81.
175. Nhi DM, Huy NT, Ohyama K, et al. A proteomic approach identifies candidate early biomarkers to predict severe dengue in children. *PLoS Negl Trop Dis*. 2016;10:e0004435.
176. Peng L, Liu J, Li YM, et al. Serum proteomics analysis and comparisons using iTRAQ in the progression of hepatitis B. *Exp Ther Med*. 2013;6:1169–1176.
177. Yang L, Rudser KD, Higgins L, et al. Novel biomarker candidates to predict hepatic fibrosis in hepatitis C identified by serum proteomics. *Dig Dis Sci*. 2011;56:3305–3315.
178. Yang J, Yang L, Li B, et al. ITRAQ-Based proteomics identification of serum biomarkers of two chronic hepatitis B subtypes diagnosed by traditional chinese medicine. *BioMed Res Int*. 2016;2016:3290260.
179. Shetty V, Jain P, Nickens Z, Sinnathamby G, Mehta A, Philip R. Investigation of plasma biomarkers in HIV-1/HCV mono- and coinfected individuals by multiplex iTRAQ quantitative proteomics. *Omics*. 2011;15:705–717.
180. Pendyala G, Trauger SA, Siuzdak G, Fox HS. Quantitative plasma proteomic profiling identifies the vitamin E binding protein afamin as

a potential pathogenic factor in SIV induced CNS disease. *J Proteome Res*. 2010;9:352–358.
181. Muenchhoff J, Poljak A, Song F, et al. Plasma protein profiling of mild cognitive impairment and Alzheimer's disease across two independent cohorts. *J Alzheimers Dis*. 2015;43:1355–1373.
182. Xu HB, Zhang RF, Luo D, et al. Comparative proteomic analysis of plasma from major depressive patients: identification of proteins associated with lipid metabolism and immunoregulation. *Int J Neuropsychopharmacol*. 2012;15:1413–1425.
183. Hergenroeder G, Redell JB, Moore AN, et al. Identification of serum biomarkers in brain-injured adults: potential for predicting elevated intracranial pressure. *J Neurotrauma*. 2008;25:79–93.
184. Betancourt A, Mobley JA, Wang J, et al. Alterations in the rat serum proteome induced by prepubertal exposure to bisphenol a and genistein. *J Proteome Res*. 2014;13:1502–1514.
185. Huang Y, Zhang X, Jiang W, et al. Discovery of serum biomarkers implicated in the onset and progression of serous ovarian cancer in a rat model using iTRAQ technique. *Eur J Obstet Gynecol Reprod Biol*. 2012;165:96–103.
186. Liu H, Wang J, Yang Z, Wang K. Serum proteomic analysis based on iTRAQ in miners exposed to soil containing rare earth elements. *Biol Trace Elem Res*. 2015;167:200–208.
187. Zhang X, Xiao Z, Liu X, et al. The potential role of ORM2 in the development of colorectal cancer. *PLoS ONE*. 2012;7:e31868.
188. Fernandez-Puente P, Mateos J, Fernandez-Costa C, et al. Identification of a panel of novel serum osteoarthritis biomarkers. *J Proteome Res*. 2011;10:5095–5101.
189. Miller MH, Walsh SV, Atrih A, Huang JT, Ferguson MA, Dillon JF. The serum proteome of nonalcoholic fatty liver disease: a multimodal approach to discovery of biomarkers of nonalcoholic steatohepatitis. *J Gastroenterol Hepatol*. 2014;29: 1839–1847.
190. Andersen RF, Palmfeldt J, Jespersen B, Gregersen N, Rittig S. Plasma and urine proteomic profiles in childhood idiopathic nephrotic syndrome. *Proteomics Clin Appl*. 2012;6:382–393.
191. Hollander Z, Lazarova M, Lam KK, et al. Proteomic biomarkers of recovered heart function. *Eur J Heart Fail*. 2014;16:551–559.
192. Freue GV, Sasaki M, Meredith A, et al. Proteomic signatures in plasma during early acute renal allograft rejection. *Mol Cell Proteomics*. 2010;9:1954–1967.
193. Wu D, Zhu D, Xu M, et al. Analysis of transcriptional factors and regulation networks in patients with acute renal allograft rejection. *J Proteome Res*. 2011;10:175–181.
194. Parviainen V, Joenvaara S, Tukiainen E, Ilmakunnas M, Isoniemi H, Renkonen R. Relative quantification of several plasma proteins during liver transplantation surgery. *J Biomed Biotechnol*. 2011;2011: 248613.
195. Lv M, Ye HG, Zhao XS, Zhao XY, et al. Ceruloplasmin is a potential biomarker for aGvHD following allogeneic hematopoietic stem cell transplantation. *PLoS ONE*. 2013;8:e58735.
196. Geyer PE, Kulak NA, Pichler G, Holdt LM, Teupser D, Mann M. Plasma proteome profiling to assess human health and disease. *Cell Systems*. 2016;2:185–195.
197. Geyer PE, Wewer Albrechtsen NJ, Tyanova S, et al. Proteomics reveals the effects of sustained weight loss on the human plasma proteome. *Mol Syst Biol*. 2016;12:901.
198. Chambers AG, Percy AJ, Yang J, Borchers CH. LC/MRM-MS enables precise and simultaneous quantification of 97 proteins in dried blood spots. *Mol Cell Proteomics*. 2015;14:3094–3104.
199. Domanski D, Percy AJ, Yang J, et al. MRM-based multiplexed quantitation of 67 putative cardiovascular disease biomarkers in human plasma. *Proteomics*. 2012;12:1222–1243.
200. Gallien S, Kim SY, Domon B. Large-Scale targeted proteomics using internal standard triggered-Parallel reaction monitoring (IS-PRM). *Mol Cell Proteomics*. 2015;14:1630–1644.
201. Erickson BK, Rose CM, Braun CR, et al. Strategy to combine sample multiplexing with targeted proteomics assays for high-Throughput protein signature characterization. *Mol Cell*. 2017;65: 361–370.
202. Boylan KL, Andersen JD, Anderson LB, Higgins L, Skubitz AP. Quantitative proteomic analysis by iTRAQ(R) for the identification of candidate biomarkers in ovarian cancer serum. *Proteome Sci*. 2010;8:31.
203. Tonack S, Jenkinson C, Cox T, et al. ITRAQ reveals candidate pancreatic cancer serum biomarkers: influence of obstructive jaundice on their performance. *Br J Cancer*. 2013;108:1846–1853.
204. Faulkner S, Elia G, Mullen MP, O'Boyle P, Dunn MJ, Morris D. A comparison of the bovine uterine and plasma proteome using iTRAQ proteomics. *Proteomics*. 2012;12:2014–2023.
205. Madian AG, Myracle AD, Diaz-Maldonado N, Rochelle NS, Janle EM, Regnier FE. Differential carbonylation of proteins as a function of *in vivo* oxidative stress. *J Proteome Res*. 2011;10:3959–3972.
206. Meng R, Gormley M, Bhat VB, Rosenberg A, Quong AA. Low abundance protein enrichment for discovery of candidate plasma protein biomarkers for early detection of breast cancer. *J Proteomics*. 2011;75:366–374.
207. Opstal-van Winden AW, Krop EJ, Karedal MH, et al. Searching for early breast cancer biomarkers by serum protein profiling of pre-diagnostic serum; a nested case-control study. *BMC Cancer*. 2011;11:381-2407-11-381.
208. Huang PY, Mactier S, Armacki N, et al. Protein profiles distinguish stable and progressive chronic lymphocytic leukemia. *Leuk Lymphoma*. 2016;57:1033–1043.
209. Zhang X, Yin X, Yu H, et al. Quantitative proteomic analysis of serum proteins in patients with Parkinson's disease using an isobaric tag for relative and absolute quantification labeling, two-dimensional liquid chromatography, and tandem mass spectrometry. *Analyst*. 2012;137: 490–495.
210. Chong PK, Lee H, Zhou J, et al. Reduced plasma APOA1 level is associated with gastric tumor growth in MKN45 mouse xenograft model. *J Proteomics*. 2010;73:1632–1640.
211. Gautam P, Nair SC, Gupta MK, et al. Proteins with altered levels in plasma from glioblastoma patients as revealed by iTRAQ-based quantitative proteomic analysis. *PLoS ONE*. 2012;7:e46153.
212. Huang C, Wang Y, Liu S, et al. Quantitative proteomic analysis identified paraoxonase 1 as a novel serum biomarker for microvascular invasion in hepatocellular carcinoma. *J Proteome Res*. 2013;12: 1838–1846.
213. Lee HJ, Na K, Choi EY, Kim KS, Kim H, Paik YK. Simple method for quantitative analysis of N-linked glycoproteins in hepatocellular carcinoma specimens. *J Proteome Res*. 2010;9:308–318.
214. Jin Y, Wang J, Ye X, et al. Identification of GlcNAcylated alpha-1-antichymotrypsin as an early biomarker in human non-small-cell lung cancer by quantitative proteomic analysis with two lectins. *Br J Cancer*. 2016;114:532–544.
215. Okano T, Seike M, Kuribayashi H, et al. Identification of haptoglobin peptide as a novel serum biomarker for lung squamous cell carcinoma by serum proteome and peptidome profiling. *Int J Oncol*. 2016;48:945–952.
216. Chai YD, Zhang L, Yang Y, et al. Discovery of potential serum protein biomarkers for lymph-node metastasis in oral cancer. *Head Neck*. 2016;38:118–125.
217. Yang Y, Huang J, Rabii B, Rabii R, Hu S. Quantitative proteomic analysis of serum proteins from oral cancer patients: comparison of two analytical methods. *Int J Mol Sci*. 2014;15:14386–14395.
218. Karabudak AA, Hafner J, Shetty V, et al. Autoantibody biomarkers identified by proteomics methods distinguish ovarian cancer from non-ovarian cancer with various CA-125 Levels. *J Cancer Res Clin Oncol*. 2013;139:1757–1770.

219. Li Y, Zhang QL, Yang XL, et al. Clusterin is a potential serum marker of hepatic carcinoma. *Zhonghua Gan Zang Bing Za Zhi=Zhonghua Ganzangbing Zazhi = Chinese Journal of Hepatology* 2012;20:275–279.
220. Russell MR, Walker MJ, Williamson AJ, et al. Protein Z: A putative novel biomarker for early detection of ovarian cancer. *Int J Cancer*. 2016;138:2984–2992.
221. Rehman I, Evans CA, Glen A, et al. ITRAQ identification of candidate serum biomarkers associated with metastatic progression of human prostate cancer. *PLoS ONE*. 2012;7:e30885.
222. Zhang M, Chen L, Yuan Z, et al. Combined serum and EPS-urine proteomic analysis using iTRAQ technology for discovery of potential prostate cancer biomarkers. *Discov Med*. 2016;22:281–295.
223. Zhang Y, Cai Y, Yu H, Li H. ITRAQ-Based quantitative proteomic analysis identified HSC71 as a novel serum biomarker for renal cell carcinoma. *BioMed Res Int*. 2015;2015:802153.
224. Upur H, Chen Y, Kamilijiang M, et al. Identification of plasma protein markers common to patients with malignant tumour and Abnormal Savda in Uighur medicine: a prospective clinical study. *BMC Complement Altern Med*. 2015;15:9–015-0526-6.
225. Liu C, Mao L, Ping Z, et al. Serum protein KNG1, APOC3, and PON1 as potential biomarkers for yin-Deficiency-Heat syndrome. *Evid Based Complement Alternat Med*. 2016;2016:5176731.
226. Zhang Y, Kang Y, Zhou Q, et al. Quantitative proteomic analysis of serum from pregnant women carrying a fetus with conotruncal heart defect using isobaric tags for relative and absolute quantitation (iTRAQ) labeling. *PLoS ONE*. 2014;9:e111645.
227. Li J, Liu X, Xiang Y, et al. Alpha-2-macroglobulin and heparin cofactor II and the vulnerability of carotid atherosclerotic plaques: an iTRAQ-based analysis. *Biochem Biophys Res Commun*. 2017;483:964–971.
228. Bijttebier J, Tilleman K, Dhaenens M, Deforce D, Van Soom A, Maes D. Comparative proteome analysis of porcine follicular fluid and serum reveals that excessive alpha(2)-macroglobulin in serum hampers successful expansion of cumulus-oocyte complexes. *Proteomics*. 2009;9:4554–4565.
229. Huang Z, Wang H, Huang H, et al. ITRAQ-based proteomic profiling of human serum reveals down-regulation of platelet basic protein and apolipoprotein B100 in patients with hematotoxicity induced by chronic occupational benzene exposure. *Toxicology*. 2012;291:56–64.

How to cite this article: Moulder R, Bhosale SD, Goodlett DR, Lahesmaa R. Analysis of the plasma proteome using iTRAQ and TMT-based Isobaric labeling. *Mass Spec Rev*. 2018;37:583–606. https://doi.org/10.1002/mas.21550

© 2018. Published by The Company of Biologists Ltd | Disease Models & Mechanisms (2018) 11, dmm032920. doi:10.1242/dmm.032920

REVIEW

SPECIAL COLLECTION: CANCER METABOLISM

The contributions of cancer cell metabolism to metastasis

Gloria Pascual[1,*], Diana Domínguez[1,*] and Salvador Aznar Benitah[1,2,‡]

ABSTRACT

Metastasis remains the leading cause of cancer-related deaths worldwide, and our inability to identify the tumour cells that colonize distant sites hampers the development of effective anti-metastatic therapies. However, with recent research advances we are beginning to distinguish metastasis-initiating cells from their non-metastatic counterparts. Importantly, advances in genome sequencing indicate that the acquisition of metastatic competency does not involve the progressive accumulation of driver mutations; moreover, in the early stages of tumorigenesis, cancer cells harbour combinations of driver mutations that endow them with metastatic competency. Novel findings highlight that cells can disseminate to distant sites early during primary tumour growth, remaining dormant and untreatable for long periods before metastasizing. Thus, metastatic cells must require local and systemic influences to generate metastases. This hypothesis suggests that factors derived from our lifestyle, such as our diet, exert a strong influence on tumour progression, and that such factors could be modulated if understood. Here, we summarize the recent findings on how specific metabolic cues modulate the behaviour of metastatic cells and how they influence the genome and epigenome of metastatic cells. We also discuss how crosstalk between metabolism and the epigenome can be harnessed to develop new anti-metastatic therapies.

KEY WORDS: Cancer, Metabolism, Metastasis, Epigenetics

Introduction

Metastasis is the leading cause of cancer-related deaths worldwide yet, at the cellular level, it is an inefficient process – only a small fraction of cells shed from a primary tumour into the bloodstream or lymphatic system will successfully complete all the sequential steps of the metastatic cascade (Oskarsson et al., 2014). The mechanisms by which some tumour cells detach from the primary lesion to colonize distant sites are beginning to be deciphered, providing new avenues for therapeutic intervention. For instance, pro-metastatic events common to most solid tumours include the reversible transition of tumour cells from an epithelial to a mesenchymal state [epithelial-to-mesenchymal transition (EMT); see Box 1 for a glossary of terms], as well as interactions with tumour-activated stromal cells, such as pericytes, fibroblasts, endothelial cells, adipocytes or immune cells (Calon et al., 2012, 2015; Cao et al., 2014; Chaffer and Weinberg, 2011; Chen et al., 2011; Goel and Mercurio, 2013; Kalluri and Zeisberg, 2006; Lu et al., 2011; McAllister and Weinberg, 2014; Nieman et al., 2011; Obenauf et al., 2015; Oskarsson et al., 2011, 2014; Paolino et al., 2014; Peinado et al., 2012; Sevenich et al., 2014; Ugel et al., 2015; Valiente et al., 2014; Wculek and Malanchi, 2015; Zhang et al., 2013). Tumours also secrete metastasis-promoting exosomes that contain various proteins, mRNAs and microRNAs, to establish a distant pro-metastatic niche (Costa-Silva et al., 2015; Ghajar et al., 2013; Obenauf et al., 2015; Peinado et al., 2012; Zhou et al., 2014; Zomer et al., 2015).

The field of cancer research historically posited that tumour progression entails the progressive accumulation of genetic mutations and the sequential selection of sub-clones – a process that culminates in metastasis as a clinical manifestation of late-stage disease (Merlo et al., 2006; Nowell, 1976). However, recent advances in whole-genome sequencing indicate that, at a very early stage, cancer cells harbour combinations of driver mutations (Box 1) that endow them with metastatic competency (Calon et al., 2012; Goel and Mercurio, 2013; Kalluri and Zeisberg, 2006; Nieman et al., 2011; Zhang et al., 2013). Furthermore, findings from metastatic assays performed *in vivo* suggest that metastatic cells reach distant organs early during primary tumour growth, yet can remain dormant (Box 1), and untreatable, for long periods of up to several years before generating metastases, which are often fatal (Cao et al., 2014; Kalluri and Zeisberg, 2006; Zhang et al., 2013). These studies have revealed that, early during tumorigenesis, the specific driver mutations that confer tumour cells with selective advantages might be the same mutations that provide them with the competency to metastasize (Jacob et al., 2015; Patel and Vanharanta, 2016; Vanharanta and Massagué, 2013). These findings indicate that tumour cells require additional local and systemic influences to metastasize, and imply that our lifestyle could impact tumour progression, which in turn suggests that such lifestyle factors could be modulated if understood. Nevertheless, we are only beginning to understand the nature of the factors that promote metastasis, their origin, and why not all tumour cells respond to them in the same way.

The recent and exciting identification of metastasis-initiating cells (MICs) in different types of tumours allows us to explore what distinguishes metastatic cells from their non-metastatic counterparts (Dieter et al., 2011; Hermann et al., 2007; Lawson et al., 2015; Pascual et al., 2017; Patrawala et al., 2006; Roesch et al., 2010; Wculek and Malanchi, 2015). One particularly interesting aspect of metastatic cells is that they seem to be strongly influenced by specific types of metabolism and their derived metabolites. For instance, lipid metabolism is emerging as an essential factor in tumour progression (Baenke et al., 2013; Pascual et al., 2017). Importantly, intracellular metabolic changes might establish and sustain transcriptional programmes required for metastatic competency, as exemplified by the strong link between specific metabolites and the epigenetic machinery that controls gene expression (Fan et al., 2015; Kinnaird et al., 2016).

[1]Institute for Research in Biomedicine (IRB Barcelona), Oncology Department, The Barcelona Institute of Science and Technology (BIST), 08028, Barcelona, Spain. [2]Catalan Institution for Research and Advanced Studies (ICREA), 08010, Barcelona, Spain.
*These authors contributed equally to this work

‡Author for correspondence (salvador.aznar-benitah@irbbarcelona.org)

S.A.B., 0000-0002-9059-5049

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/3.0), which permits unrestricted use, distribution and reproduction in any medium provided that the original work is properly attributed.

Disease Models & Mechanisms (2018) 11, dmm032920. doi:10.1242/dmm.032920

Box 1. Glossary

Anaplerosis: the replenishment of TCA cycle intermediates that have been used for biosynthetic processes.

Dormant/quiescent cancer cells: cells that remain in a state of dormancy or quiescence in response to intrinsic or extrinsic stimuli that arrest mitosis and cell growth.

Driver mutation: mutational events that confer cancer cells with either growth or survival advantages and therefore favour (drive) tumour development.

Enhancers: short DNA regions that regulate gene expression by recruiting transcription factors that establish and maintain cell-specific transcriptional signatures. They can be located far away from the genes they regulate.

Epithelial-to-mesenchymal transition (EMT): a process by which epithelial cells gain mesenchymal cell properties, including migratory and invasive traits. EMT occurs during embryogenesis, wound healing and during the malignant transformation of cancer cells.

Ketogenic diets: low-carbohydrate and high-fat diets that force the body to preferentially mobilize stored lipids for energy production, resulting in increased levels of ketone compounds.

Metformin: a drug commonly used to treat type-2 diabetes that inhibits OXPHOS and consequently reduces ATP production in mitochondria, favouring cellular ATP production via glycolysis.

Omental fat pad: an area of fat tissue in the abdominal cavity that surrounds the intestines.

Orthotopic model: an animal model, most commonly mouse, in which human tumour cells are injected or implanted into the equivalent organ or tissue that the human cancer originated from.

Oxidative phosphorylation (OXPHOS): this process is the most efficient way to produce ATP in eukaryotic cells via the electron flow that occurs in the mitochondrial inner membrane.

Pioneer transcription factor: a transcription factor that can bind compacted chromatin to recruit chromatin remodelling proteins and other transcription factors; these transcription factors are important for determining cell fate.

Preconditioned niche: an environment at a distance from the primary tumour that is modified and generates specific signals prior to the arrival of disseminated cancer cells; this environment can facilitate subsequent tumour cell infiltration and colonization.

Tricarboxylic acid (TCA) cycle: this metabolic pathway, also known as the Krebs cycle, produces electron carriers for ATP generation through the electron transport chain, which takes place in mitochondria.

Triple-negative breast cancer (TNBC): breast carcinoma histologically assessed as negative for estrogen receptors (ER–), progesterone receptors (PR–) and HER2 receptors (HER2–).

Warburg effect: process by which tumour cells increase their glucose uptake and convert pyruvate to lactate through aerobic glycolysis under normal conditions of oxygen and glucose availability, yielding ATP, increased NADPH reductive power and metabolic intermediates, instead of coupling glycolysis to the TCA cycle and OXPHOS pathways.

In this Review, we discuss recent insights into the metabolic plasticity of cancer cells and the way in which their metabolic processes can contribute to their metastatic transformation. We highlight the emerging role of lipid metabolism as an important source of cancer metabolic heterogeneity, provide an overview of the crosstalk that occurs between metabolic processes and the cancer cell epigenome, and examine how lifestyle influences, such as diet, might affect cancer progression. We also discuss the therapeutic potential of targeting metabolism during cancer progression, highlighting novel and experimental drugs currently under preclinical investigation.

Metabolic heterogeneity of cancer stem cells

Cancer stem cells (CSCs) sustain the growth of the tumour mass and are responsible for therapy failure and patient relapse (Blanpain, 2013). The identification and characterization of CSCs in a number of malignancies is paving the way towards developing novel CSC-targeted anti-cancer approaches (Collins et al., 2005; Eramo et al., 2008; Hermann et al., 2007; Kreso et al., 2013; Li et al., 2007; Prince et al., 2007; Ricci-Vitiani et al., 2007; Singh et al., 2004; Wu, 2008; Patrawala et al., 2006). One important conclusion of several of these studies is that CSCs display molecular and functional heterogeneity. Interestingly, this heterogeneity seems to be established early during tumorigenesis because genetically distinct CSC sub-clones are already present in primary tumours, some of which fade or become dominant during tumour progression and response to chemotherapy (Ben-David et al., 2017; Shlush et al., 2017; Zehir et al., 2017). However, it is still a matter of debate whether all cells capable of initiating and promoting primary tumour growth are equally competent to initiate metastasis. Substantial evidence suggests that only a few CSC clones present within a primary tumour possess the ability to behave as MICs (Campbell et al., 2010; Roesch et al., 2010; Wculek and Malanchi, 2015; Pascual et al., 2017). As these clones do not harbour new mutations relative to the primary tumour, non-genetic factors are likely to be required to promote their metastatic competency (Hansen et al., 2011). Thus, local and systemic signals might endow certain CSC clones with the ability to colonize distant organs, underlying the functional diversity of a population of genetically identical cancer clones (Kreso et al., 2013).

Intriguingly, the functional heterogeneity of CSCs might require them to use different types of metabolism. As early as 1926, Otto Warburg reported that tumour cells do not generally couple glycolysis to the tricarboxylic acid (TCA) cycle (Box 1) and oxidative phosphorylation (OXPHOS; Box 1), even under normal conditions of oxygen and glucose availability (Warburg, 1925). Instead, they convert pyruvate to lactate through aerobic glycolysis, to yield ATP, NADPH reductive power and metabolic intermediates for cellular biosynthesis (Koppenol et al., 2011; Vander Heiden et al., 2009). The Warburg effect (Box 1), although inefficient, allows for periods of increased biosynthetic demand (Vander Heiden and DeBerardinis, 2017). Nonetheless, cancer cells have functional mitochondria and can use glycolytic or oxidative metabolic programmes to obtain energy when confronted with different scenarios. This is particularly true for CSCs from different types of tumour, which generally and predominantly rely on OXPHOS to cope with their energetic demands, as compared to differentiated tumour cells (Lagadinou et al., 2013; Sancho et al., 2015; Viale et al., 2014) but can revert to glycolysis and enhanced glucose uptake under alternative microenvironments (Dong et al., 2013; Marin-Valencia et al., 2012). For instance, patient-derived $CD133^+$ pancreatic CSCs display enhanced mitochondrial respiration and impaired glycolytic plasticity relative to their $CD133^-$ non-CSC counterparts, rendering these cells more vulnerable to metformin (Box 1), which inhibits their mitochondrial respiration. Nevertheless, some CSCs within the CSC pool develop metformin resistance through an intermediate metabolic state that involves increased glycolysis and reduced mitochondrial oxygen consumption (Sancho et al., 2015).

Lipids and lipid metabolism have been also linked to CSC function. For instance, ovarian CSCs require *de novo* fatty acid (FA) synthesis and lipid desaturation, through the activity of stearoyl-coA desaturase 1 (SCD1), to promote tumour initiation in a nuclear factor kappaB (NFκB)-dependent manner (Li et al., 2017). In human breast cancer cell lines, expression of the mitochondrial protein lactamase beta (LACTB) is significantly downregulated; LACTB downregulation is required for the synthesis of two

Disease Models & Mechanisms (2018) 11, dmm032920. doi:10.1242/dmm.032920

phospholipids, phosphatidyl ethanolamine (PE) and LysoPE, that are essential for membrane biosynthesis (Keckesova et al., 2017). Thus, the ability of CSCs to adjust their metabolic state – their metabolic plasticity – likely influences how they contribute to tumour progression and relapse. The following section will address how CSCs modulate their metabolic activity to adapt to environmental cues when confronted with distinct conditions and microenvironments.

Tumour–microenvironment metabolic crosstalk in metastasis

How cancer cells engage in reciprocal communication with their tumour microenvironment (TME) affects tumour initiation (Oskarsson et al., 2014; Pein and Oskarsson, 2015; Shiozawa et al., 2011). However, we have only recently begun to understand the importance of the interactions between cancer cells and the TME in promoting metastatic initiation and macroscopic metastatic growth.

Metabolic preconditioning of the metastatic niche

An intriguing aspect of the metastatic process is that metastatic niches are preconditioned (Box 1) prior to the arrival of disseminated cancer cells from primary tumour sites (Kaplan et al., 2005; Peinado et al., 2017). Accumulating evidence indicates that specific metabolic changes are associated with the establishment of these niches, although whether these changes are tumour-specific or not remains unclear. For instance, factors secreted by the primary tumour can be transferred to neighbouring or distant cells via exosomes to metabolically modulate the TME and favour metastasis (Peinado et al., 2017) (Fig. 1). Thus, in a human xenograft model, circulating breast-cancer-derived vesicles that contain the microRNA miR-122 have been found to downregulate the expression of glucose transporter 1 (GLUT1) and pyruvate kinase M1/M2 (PKM1/2) isozymes, thereby reducing glucose uptake in lung and brain non-tumour cells. This in turn increases the local pool of glucose available for the incoming metastatic cancer cells (Fong et al., 2015). Similarly, in an *in vivo* experimental model of metastasis in mice, disseminated human colorectal cancer cells were found to secrete the protein creatine kinase brain-type (CKB) in the liver, where it catalyses the conversion of extracellular ATP and hepatic creatine into phosphocreatine, which is incorporated into metastatic cells via the creatine transporter, SLC6A8. Increased phosphocreatine availability in the extracellular space fuels the generation of ATP within metastatic cells, which is essential to meet the intense energetic requirements involved in colonizing the liver (Loo et al., 2015).

Hypoxia and cancer cell metabolism

The unique unstructured capillary network of an exponentially growing tumour generates areas of low oxygen diffusion and consumption, a phenomenon termed hypoxia. Upon arrival at distant organs, metastatic cells probably lack a proper vasculature to provide them with a plentiful oxygen supply. Importantly, signals triggered by hypoxic conditions are a feature of aggressive tumours, and are associated with poor prognosis and high levels of metastasis (Rankin and Giaccia, 2016; Semenza, 2016).

A key aspect of the hypoxic response is the metabolic adaptation of malignant cells to overcome cell death (Benjamin et al., 2012). The hypoxia-inducible transcription factor 1 alpha (HIF1-α, or HIF1a), is an important transducer of hypoxic signals and drives the expression of GLUTs and glycolytic enzymes, such as hexokinase 2 (HK2), phosphofructokinase 1 (PFK1) and lactate dehydrogenase A (LDHA), while inhibiting the expression of pyruvate dehydrogenase kinase 1 (PDK1). This response reduces the flux

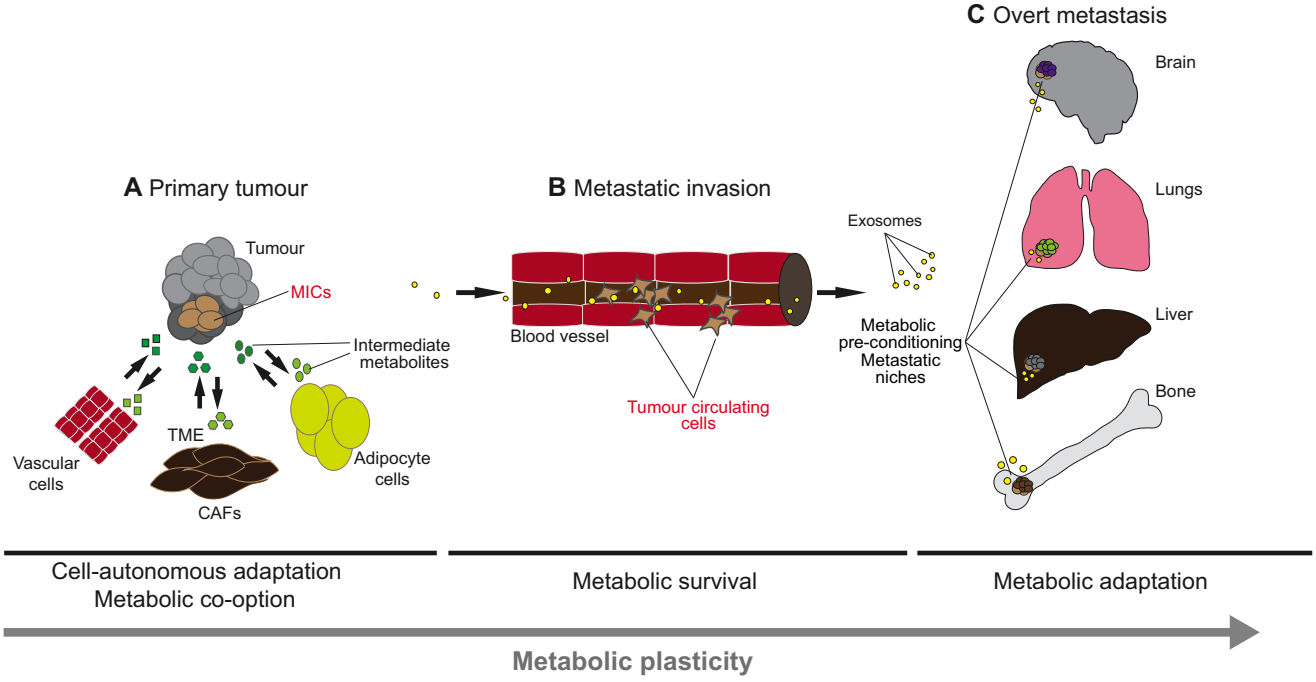


Fig. 1. Cancer metabolic plasticity contributes to metastatic disease. (A) Genetic mutations and epigenetic alterations in combination establish unique populations of tumour-initiating cells (TICs). Only certain TICs take advantage of the surrounding cells that constitute the tumour microenvironment (TME), such as vascular cells, cancer-associated fibroblasts (CAFs) and adipocytes, as well as their systemic environment, to exchange and hijack metabolites (shown in green) that support TIC survival. TICs hijack metabolites while egressing out of the primary lesion, thereby becoming metastasis-initiating cells (MICs). (B) As metastatic cells reach different distant organs via the vasculature, they adopt unique metabolic states and engage in further metabolic crosstalk with the (C) metastatic niches that form, for example, in the bone, lungs, liver and brain, ultimately supporting their survival. Tumour cells also secrete metastasis-promoting exosomes (yellow) that contain various proteins and RNAs that contribute to establish distant pro-metastatic niches.

of pyruvate into the TCA cycle, thereby decreasing the rate of OXPHOS and oxygen consumption (Semenza, 2010). In certain tumours, hypoxia also alters lipid metabolism by promoting the use of alternative carbon sources, such as acetate, to sustain FA synthesis and tumour growth. This lipid metabolic switch is associated with the transcriptional control of acetyl-CoA synthetase-2 (ACSS2) by HIF signalling, and by *ACSS2* copy-number gains. Interestingly, *ACSS2* expression is specifically increased in metastatic cells (Schug et al., 2015). In addition, hypoxic glioblastoma and mammary gland tumours accumulate lipid droplets (LDs) in a HIF1a-dependent manner, through the uptake of FAs via the fatty-acid-binding proteins 3 and 7 (FABP3 and FABP7). Lipid droplets, in turn, sustain cancer cell survival upon re-oxygenation by providing a lipid reservoir that ensures continued ATP production via FA β-oxidation and by protecting cells from reactive oxygen species (ROS) toxicity by reducing NADPH generation (Bensaad et al., 2014). However, whether and how these pathways contribute to the survival of metastatic cells upon their arrival at distant sites is currently unknown.

Metabolic coupling between tumour cells and stroma promotes cancer spread

One interesting aspect of tumour cells is their ability to shape the metabolism of their TME to ensure a plentiful supply of energy, through a process known as metabolic coupling. Metabolic coupling can occur either between tumour cells or between tumour and stromal cells. For instance, metastatic ovarian cancer cells induce the release of stored lipids from the omental fat pad (Box 1) they colonize, which guarantees their ability to obtain FAs for energetic and biosynthetic purposes (Nieman et al., 2011). Likewise, in a mouse model of chronic myeloid leukaemia (CML) and in primary human CML samples, a subpopulation of leukaemia stem cells (LSCs), called GAT-LSCs, associates with the gonadal adipose tissue (GAT) upon chemotherapy. GAT-LSCs remain quiescent and secrete pro-inflammatory cytokines, such as IL-1α and TNF-α, to stimulate lipolysis in the nearby GAT adipocytes, which facilitates the uptake and oxidation of free FAs (FFAs) captured by the FA receptor CD36 (cluster of differentiation 36). Interestingly, this metabolic coupling confers GAT-LSCs with the energetic supply to resist chemotherapy and is therefore essential for leukaemia relapse (Ye et al., 2016).

These findings suggest new therapeutic avenues by which to prevent the successful colonization of distant sites by metastatic cells by inhibiting the metabolic crosstalk between cancer cells and their environment. For instance, mouse models of both melanoma (B16) and Lewis lung carcinoma (LLC) have been shown to possess a hyper-glycolytic intra-tumour endothelium that favours metastatic spreading (Cantelmo et al., 2016). Interestingly, the inhibition of the glycolytic activator PFKFB3 (6-phosphofructo-2-kinase) in endothelial cells reduces cancer cell intravasation and metastasis by normalizing tumour vessel architecture (Cantelmo et al., 2016).

We are only just beginning to understand how different local metabolites, and the metabolic programmes they elicit, shape the ability of tumour cells to colonize different sites. In the next section, we discuss recent findings on how specific metabolic cues modulate the behaviour of tumour cells, with an emphasis on the role of FA and lipid metabolism in feeding metastatic progression.

Fuelling metastatic progression

Recent data suggest that the metabolic profiles of metastatic cells differ as they colonize different organs. For instance, in a murine orthotopic model (Box 1) of metastatic breast cancer, circulating tumour cells (CTCs) purified from blood exhibited increased transcription of the coactivator PPARGC1A (peroxisome proliferator-activated receptor gamma, coactivator 1 alpha; also known as PGC-1α) relative to the primary breast tumour or matched lung metastasis. PGC-1α was shown to determine the metastatic potential of CTCs through enhanced OXPHOS (LeBleu et al., 2014). By contrast, in a mouse model of prostate cancer, PGC-1α was reported to act as a metastasis-suppressing factor, further underscoring the relevance of cancer metabolic heterogeneity within different tumour microenvironments (Torrano et al., 2016).

Another study concluded that breast cancer cells with a broad metastatic potential (bone, lung and liver) could engage both OXPHOS and glycolysis-dependent metabolic strategies, whereas liver metastatic cells relied on glucose uptake and glycolysis, while lung metastatic cells relied on glutamine uptake and OXPHOS to metastasize (Dupuy et al., 2015).

Although these studies provide insights into the dynamic metabolic programmes of tumour cells during cancer progression, further studies will be necessary to mechanistically connect altered gene expression, and the associated bioenergetic behaviour of the cell, to each step of metastatic progression.

Why would metastatic cells that colonize different organs engage in different metabolic preferences? Although tumour cells can hijack nutrients from the cells they encounter as they arrive in an organ, nutrient availability might vary substantially between different organs. For instance, the availability of pyruvate in lungs activates pyruvate carboxylase (PC) in metastatic breast cancer cells, which in turn converts pyruvate to oxaloacetate, thereby enhancing anaplerosis (Box 1) (Christen et al., 2016). Additional studies carried out on *in vitro* three-dimensional (3D) models have revealed that breast cancer cells rely on the non-essential amino acid proline, through proline dehydrogenase enzyme (PRODH) activity, to form spheroids *in vitro* and lung metastases *in vivo* (Elia et al., 2017). Conversely, when breast cancer cells invade the interstitial space of the brain, a region with low glucose concentrations, they enhance gluconeogenesis and the oxidation of glutamine and branched-chain amino acids, which allows them to survive as they extravasate from the brain to blood vasculature (Chen et al., 2015). Cancer cells that metastasize to the brain also use acetate as an alternative substrate for the TCA cycle to support energy and biomass production. This process relies on the activity of ACSS2, which converts acetate to acetyl-CoA and fuels the TCA cycle; notably, *ACSS2* expression is associated with poor survival in patients with gliomas (Mashimo et al., 2014).

Thus, the availability of specific nutrients at different organs might impose unique metabolic states on metastatic cells. However, future studies are required to determine how specific metabolic routes sustain the high energy demand of metastatic cells while modulating the signalling pathways that allow metastatic growth.

Lipid metabolism in metastasis

Lipid metabolism is often altered in cancer cells. An increased pool of FAs might provide proliferating cancer cells with building blocks for new membranes, signalling metabolites, and substrates for FA oxidation that fulfil the increased energy demand associated with cancer progression. Although FA metabolism has been linked to the growth of primary lesions, including non-small-cell lung cancer and acute myeloid leukaemia (German et al., 2016; Svensson et al., 2016), recent studies suggest that it also associates closely with metastasis. For instance, the internalization of FAs is a main feature of quiescent MICs that eventually promotes their metastasis upon dissemination (Pascual et al., 2017). These MICs are characterized

by a high cell-surface expression of the FA receptor, CD36, and have been shown to be solely responsible for initiating metastasis in orthotopic models of human oral squamous cell carcinoma (OSCC), and in experimental metastasis models of human melanoma and of breast cancer (Pascual et al., 2017). Interestingly, human $CD36^+$ OSCC metastatic cells in mouse orthotopic models are exquisitely sensitive to circulating blood fat levels, and a high-fat diet or stimulation with palmitic acid strongly boosts these cells' metastatic potential (Pascual et al., 2017). Although the transcriptomic signature of $CD36^+$ cells does not strongly associate with EMT, FA uptake by CD36 and by the FA-binding proteins 1 and 4 (FABP1 and FABP4) induces EMT in liver cancer cells, thereby increasing their migration and invasion in *in vitro* assays (Nath et al., 2015).

Besides CD36, the enzyme monoacylglycerol lipase (MAGL), and the FFAs it produces, are elevated in aggressive human ovarian cancer cell lines and in primary ovarian tumours (Nomura et al., 2010). Interestingly, blocking MAGL impairs ovarian tumour growth *in vivo* and ovarian tumour cell migration *in vitro*, and both phenotypes are rescued by exogenous sources of FFAs, including a high-fat diet. These findings further underscore the importance of dietary lipids in promoting malignancy and the migratory capacity of cancer cells (Nomura et al., 2010).

The precise mechanisms by which FAs promote metastasis remain unknown. Louie et al. addressed this question by coupling an isotope-based FA-labelling strategy with metabolomic profiling of different, aggressive human tumour cells, including breast, ovarian, prostate and melanoma (Louie et al., 2013). Their results indicate that these cancer cells use exogenous FAs, such as palmitic acid, as structural lipids and to generate oncogenic signalling (Louie et al., 2013). Other studies, performed in highly aggressive, triple-negative breast cancers (TNBCs; Box 1) indicate that the mitochondrial β-oxidation of FAs meets the high energy demands of metastatic cells as they migrate from the primary tumour to the distant organ (Park et al., 2016). In this sense, acetyl-CoA generated from FA β-oxidation feeds into the TCA cycle to generate large quantities of ATP. Intriguingly, this metabolic pathway seems to be regulated by the proto-oncogenes *Myc* and tyrosine-protein kinase *Src*, two well-known signalling regulators of metastasis (Camarda et al., 2016; Park et al., 2016; Zhang et al., 2009).

Lipid metabolism adaptation by tumour cells has also been demonstrated to play a role in the mechanisms responsible for the adaptation and failure of anti-angiogenic therapies (Sounni et al., 2014). For instance, human and mouse *in vivo* cancer models of breast, colorectal and LLC show a metabolic shift towards *de novo* lipogenesis after therapy withdrawal, which is accompanied by tumour regrowth and a drastic increase in metastatic dissemination. Importantly, the pharmacological inhibition of lipogenesis prevents tumour relapse and metastases (Sounni et al., 2014).

Increased FA uptake and metabolism appears to be a general feature of metastatic cancers, irrespective of the tumour type studied, as shown by the strong correlation that exists between FA uptake and metabolism, and metastatic prevalence and poor patient survival (Nath and Chan, 2016; Pascual et al., 2017). Further studies are required to understand why it is that certain FAs, and their metabolism, associate so closely with metastatic behaviour. Nonetheless, our current knowledge highlights the perils of increased consumption of diets that are rich in added fats in industrialized countries. They also highlight a need for caution concerning ketogenic diets (Box 1), which some believe to be beneficial as an adjuvant cancer therapy. This is because the administration of the ketone compound 3-hydroxy-butyrate in MDA-MB-231 breast cancer xenograft models has been shown to increase and fuel tumour growth (Bonuccelli et al., 2010). Further studies are needed to investigate how diet and lifestyle influence the progression of cancer.

Metabolism-driven epigenetic alterations in cancer

Cancer is associated with genetic alterations (Curtis et al., 2010; Kreso and Dick, 2014). CSCs of different tumour types carry driver mutations that commonly target stemness-related genes and alter the expression signatures of these cells (Eppert et al., 2011; Merlos-Suárez et al., 2011; Pece et al., 2010). Recent genome-wide sequencing studies indicate that primary lesions and their matched metastases harbour the same set of driver mutations. Since very few cells that harbour these mutations become metastatic, non-genetic alterations must also contribute to the acquisition of metastatic traits (Makohon-Moore et al., 2017; Vogelstein et al., 2013). Indeed, distinct metastasis-associated molecular (transcriptional or proteomic) signatures exist in primary tumours (Nath and Chan, 2016; Pascual et al., 2017; Ramaswamy et al., 2003). In addition, single-cell gene-expression data have recently identified transcriptional signatures that evolve as the metastatic process progresses (Lawson et al., 2015). Thus, although the primary tumour cells harbour the necessary panoply of driver mutations to be metastatic, non-genetic factors must also be required for these cells to exert their full metastatic potential.

Metabolism shapes the tumour epigenome

Unlike stable genetic traits, epigenetic events are dynamic and reversible, and provide a considerable degree of functional plasticity (Avgustinova and Benitah, 2016). Importantly, large-scale epigenetic modifications have been detected in tumour cells as they colonize distant sites (Bell et al., 2016; Denny et al., 2016; McDonald et al., 2017; Roe et al., 2017). The epigenetic regulation of genes involved in EMT and in mesenchymal to epithelial transition (MET), such as cadherin 1 (*CDH1*), has been extensively documented in a number of malignancies, including breast and hepatocellular carcinomas (Choi et al., 2015; Zhang et al., 2016). Likewise, the acquisition of a neuronal programme that is associated with the metastasis of small-cell lung cancer (SCLC) occurs via the opening up of large-scale chromatin domains and is induced by the transcription factor NFIB (Denny et al., 2016; Minna and Johnson, 2016). Moreover, miRNAs that contribute to the acquisition of metastatic behaviour are controlled by changes in DNA methylation (Mudduluru et al., 2017; Rokavec et al., 2017).

Although the molecular mechanisms that underlie chromatin remodelling during metastasis are still largely unexplored, they are influenced by the local and systemic availability of metabolites (Fig. 2). In this sense, several of these metabolites function either as cofactors or substrates for enzymes involved in the deposition of epigenetic marks, thereby directly influencing the epigenetic landscape of tumour cells (Fan et al., 2015). It was recently suggested that enzymes producing some of these metabolic substrates could even colocalize with the epigenetic machinery in the nucleus to generate chromatin metabolic microdomains (Katada et al., 2012). These microdomains would, in turn, facilitate the epigenetic creation of a specific metabolic state in cancer cells.

Acetyl-CoA: the substrate for histone acetylation

Acetyl-CoA-derived acetyl moieties are transferred to lysine residues on histone tails by histone acetyltransferases (HATs). Histone acetylation leads to the opening up of compact chromatin and to increased transcription of the genes that are located within the acetylated chromatin region. Global transcriptional control through

Disease Models & Mechanisms (2018) 11, dmm032920. doi:10.1242/dmm.032920

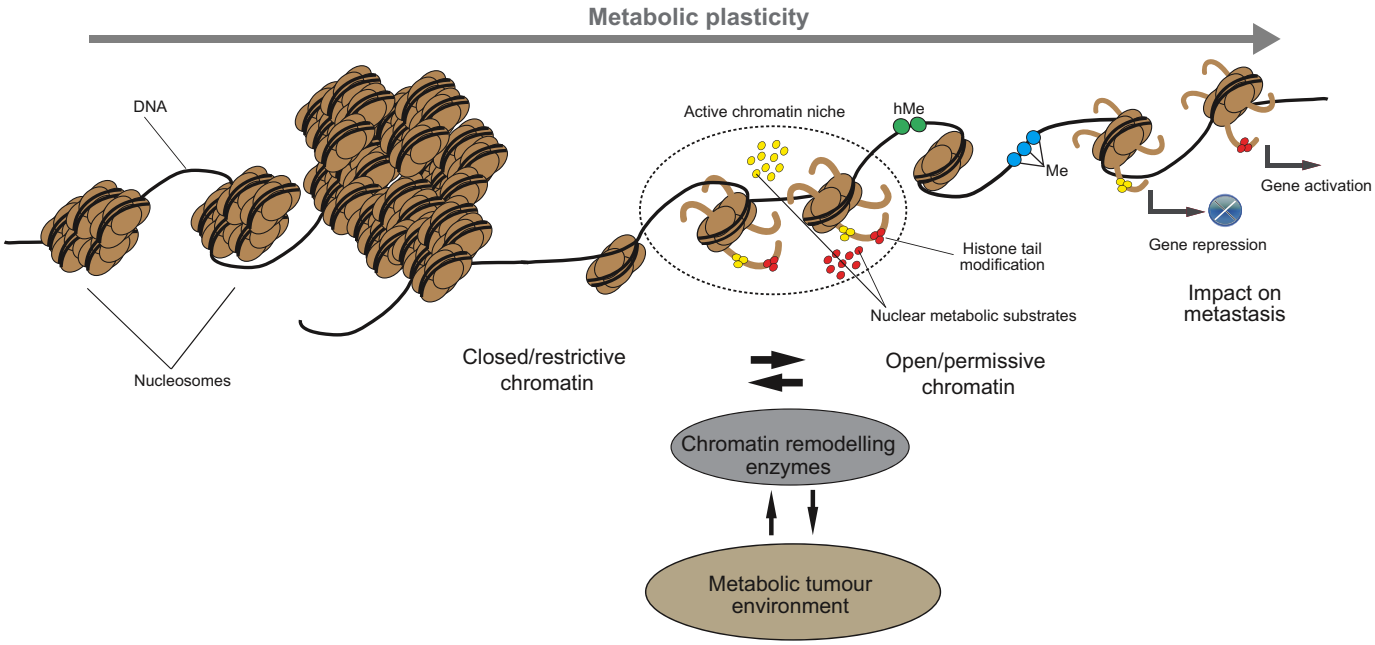


Fig. 2. Epigenetic factors integrate metabolic cues that boost metastatic transcriptional programmes in cancer cells. Yellow and red circles represent diet-derived metabolites that are utilized by several epigenetic factors in active chromatin niches to post-translationally modify histones and to methylate [blue circles are methyl groups (Me)] and hydroxymethylate [green circles are hydroxymethyl groups (hMe)] DNA. The metabolic status of a cell influences chromatin configuration, via histone tail modifications, to generate transcriptionally restrictive or permissive chromatin. A cell's metabolic status can also influence DNA methylation patterns to regulate gene transcription. This interaction between the cell metabolism and its epigenome can result in unique gene expression signatures that can contribute to the colonization of distant organs and tissues by cancer cells.

chromatin remodelling in the nucleus has been linked to the enzyme ATP citrate lyase (ACLY) (Wellen et al., 2009). In the cytosol, ACLY converts mitochondria-derived citrate, obtained from glucose oxidation in the TCA cycle, into acetyl-CoA. Importantly, ACLY localizes to the nucleus, where it produces acetyl-CoA that is then used to acetylate the chromatin that encompasses the genes involved in glucose metabolism, such as the glucose transporter 4 (*GLUT4*), *HK2*, *PFK1* and *LDHA* genes. A second enzyme, ACSS2, which is involved in controlling the pool of acetyl-CoA from acetate, translocates to the nucleus, where it can affect the incorporation of acetyl-CoA into histones (Mews et al., 2017). Although ACSS2 can positively or negatively impact tumour progression, depending on the tumour type, its production of acetyl-CoA is required for the growth and progression of liver cancer in murine models (Comerford et al., 2014). Acetyl-CoA generated by ACSS2 is also necessary to fuel breast tumour xenografts, as well as brain metastasis and glioblastoma growth in *in vivo* models (Mashimo et al., 2014; Schug et al., 2015). ACSS2 expression is enhanced in hypoxic and low-lipid environments, providing breast cancer cells with acetate metabolites required for lipid biomass production within the context of metabolic stress (Schug et al., 2015). ACSS2-derived acetate is also essential for energetic purposes because numerous primary and metastatic brain tumours can rely on acetate oxidation for energy, as recently demonstrated in orthotopic models of both human glioblastoma and brain metastasis (Mashimo et al., 2014; Schug et al., 2015). Additionally, the pyruvate dehydrogenase complex (PDC), which normally generates acetyl-CoA from glycolysis-derived pyruvate in the mitochondria, can also translocate to the nucleus to produce acetyl-CoA for histone acetylation at the loci of cell-cycle-progression genes (Sutendra et al., 2014).

Enhancer regions (Box 1) are a focus of current attention since they are highly deregulated in cancer and are the most frequently mutated non-coding regions in human tumours (Abraham et al., 2017; Bal et al., 2017; Bradner et al., 2017; Chapuy et al., 2013; Herz et al., 2014; Mansour et al., 2014; Michailidou et al., 2017; Oldridge et al., 2015; Puente et al., 2015). Interestingly, the acetylation of histone H3K27 at specific enhancers is also required to generate metastatic pancreatic ductal adenocarcinoma (PDA) cells, as seen in a PDA organoid culture system (Roe et al., 2017). The widespread rewiring of enhancers in PDA cells establishes a transcriptional network that aberrantly directs these cells towards embryonic endodermal fate, in a process that is driven by the pioneer transcription factor (Box 1) forkhead box A1 (FOXA1) (Roe et al., 2017).

Lipid metabolism can also contribute to the pool of cellular acetyl-CoA in a cancer-specific manner. When glucose-starved, ACLY-deficient cells are supplemented with FAs (the oxidation of which generates acetyl-CoA in the mitochondria), this supplementation fails to rescue the global loss of histone acetylation (Wellen et al., 2009). However, recent *in vitro* studies in various cancer cell lines have shown that acetyl-CoA obtained by FA oxidation provides acetyl groups that strongly activate a transcriptional programme involved in FA-related processes (McDonnell et al., 2016).

Taken together, these data suggest that multiple carbon sources control the pool of acetyl-CoA to regulate histone acetylation at specific loci in tumour cells. Therefore, acetyl-CoA might be a major force that establishes the transcriptional networks that shape the metabolic flexibility and heterogeneity that cancer cells exhibit as tumours progress.

Histone methylation

Histone methylation can occur at the arginine (R) and lysine (K) residues of histone 3 (H3) and histone 4 (H4), and is associated with transcriptional activation or repression, depending on the residue targeted and on the number of methyl groups deposited in the

N-terminal region of either histone. Histone methylation is tightly regulated by histone methyltransferases (HMTs) and histone demethylases. Similar to DNA methylation (discussed below), alterations in methionine metabolism and in the availability of S-adenosyl methionine (SAM) directly modulate histone methylation. For instance, the levels of histone H3K4me3, a histone mark typically associated with promoter activation, decrease upon *in vitro* methionine restriction in HCT116 cells (Mentch et al., 2015). Importantly, since methionine is an essential amino acid, the activities of HMTs, and of DNA methyltransferases (DNMTs), are strongly influenced by an organism's nutritional status. Indeed, diet-dependent changes in histone methylation have been observed in malignancies such as colorectal adenomas (Pufulete et al., 2005). Significantly, a correlation between enhanced metastatic spread and elevated levels of histone H3 trimethylation has been reported in patient-derived xenograft models of melanoma (Shi et al., 2017).

DNA methylation

The methylation of CpG islands generally impedes promoter activation and consequent gene expression, although it can also correlate with enhanced transcription factor binding (Domcke et al., 2015; Jones, 2012; Yin et al., 2017). DNA methylation also occurs at actively transcribed gene bodies and at active enhancers (Baubec et al., 2015; Charlet et al., 2016; Dhayalan et al., 2010; Rinaldi et al., 2016). DNMTs are responsible for catalysing the transfer of a methyl group to DNA. However, this modification can be eliminated by the action of ten-eleven translocation (TET) methylcytosine dioxygenases, which initiate the demethylation of DNA by hydroxylating 5-methylcytosines (Scott-Browne et al., 2017).

DNA methylation is vitally important for establishing stable epigenetic states during mammalian development (Rinaldi and Benitah, 2015; Seisenberger et al., 2012; Watanabe et al., 2002). However, in cancer cells, large regions of methylated DNA are hypermethylated, whereas others show widespread hypomethylation (Feinberg and Vogelstein, 1983). In addition, gene body methylation is also altered specifically at genes involved in tumorigenesis (Yan et al., 2015). Although the signals that specifically establish these regions of hyper- and hypomethylation in cancer are not known, it is likely that they are strongly influenced by metabolism. In this sense, the primary methyl group donor for DNA is SAM, which is generated in the methionine cycle from methionine and ATP. After losing the methyl group, SAM is converted to S-adenosylhomocysteine (SAH), which inhibits the activity of DNMTs; SAH can be recycled back to SAM, thus establishing a feedback loop that controls DNA methylation. Furthermore, the TCA cycle intermediate α-ketoglutarate is a co-substrate required for DNA demethylation through TET enzymatic activity, whereas the TCA-cycle-derived metabolites succinate, fumarate and 2-hydroxyglutarate act as competitive TET inhibitors.

How metabolic alterations converge with genetic and epigenetic changes in cancer is unknown. Nonetheless, the activation of the oncogene *KRAS* in pancreatic cancer results in tumours with increased glycolytic flux and serine biosynthesis, in which DNA methylation is fuelled by the increased availability of SAM. This results in the over-activation of DNMTs that control the transcriptional silencing of specific retrotransposons by regulating their methylation status. Hypermethylation of these repetitive elements might promote oncogenic transformation and growth through transcriptional control of host genes (Kottakis et al., 2016). In addition, epidermal stem cells that lack Dnmt3A and 3B upregulate a unique lipid transcriptional network that is associated with $CD36^{+}$ MICs, and are consequently much more predisposed to generating metastatic epidermal tumours (Pascual et al., 2017; Rinaldi et al., 2017). Considering that the genes encoding Dnmt3a and TET proteins are among the most frequently mutated in human cancers (Shlush et al., 2017; Zehir et al., 2017), it will be interesting to study how their mutations influence tumour metabolism, and *vice versa*, during metastatic progression.

Despite the increasing amount of data supporting the importance of metabolism–epigenetics crosstalk in malignancies, we still know very little about the specific impact that metabolism has on altering histone methylation and chromatin architecture in tumour cells. Future studies should aim to investigate how dietary habits affect histone acetylation and methylation in adult stem cells, and how dietary-induced epigenetic changes might predispose adult stem cells to acquire gene expression programmes that can promote metastatic disease.

Therapeutically targeting metastasis through metabolism

We have reviewed here the growing body of evidence showing that the metabolic plasticity of tumour cells plays an essential role during cancer progression. Tumour cells hijack metabolites from their surroundings and establish different metabolic states to cope with the challenging conditions they face as the tumour grows and colonizes distant sites. Metabolic heterogeneity endows cancer cells with the flexibility required for their growth and metastasis, and could constitute a potential Achilles heel for novel therapeutic strategies to prevent tumour growth and dissemination.

Several therapeutic strategies have already been proven successful in preclinical cancer models (Table 1). For instance, targeting leukotrienes through the inhibitor Zileuton strongly prevents the recruitment of neutrophils, which are required for the formation of a pro-metastatic niche for breast cancer cells in the lung (Wculek and Malanchi, 2015). Targeting lipid metabolism has also

Table 1. Compounds targeting cancer lipid metabolism*

Target	Drug	Affected process	Effect	References
CD36	Blocking/neutralizing CD36 antibodies	FA recognition and internalization	Diminishes metastasis initiating capacity *in vivo*	Pascual et al., 2017; Nieman et al., 2011
MAGL	JZL184	FA storage and mobilization	Promotes apoptosis	Nomura et al., 2010
ACS	Triacsin C	FA oxidation	Promotes apoptosis/impairs tumour growth	Cha and Lee, 2016
CPT1	Etomoxir	Mitochondrial B-oxidation	Impairs proliferation	Nieman et al., 2011
FASN	Orlistat	*De novo* FA synthesis	Promotes apoptosis/impairs tumour growth	Malvi et al., 2015; Seguin et al., 2012

*Possible anti-metastatic strategies that target different lipid metabolism components; examples of drugs that can alter cancer cell metabolism on inhibition of exogenous lipid utilization and *de novo* FA synthesis are provided. Please note that other inhibitors not listed in this Review have been reported to interfere with these processes.

CD36, cluster of differentiation 36; FASN, fatty acid synthase; MAGL, monoacylglycerol lipase; ACS, acyl-CoA synthetases; CPT1, carnitine palmitoyltransferase 1; FA, fatty acid.

produced promising results in preclinical cancer models. For example, in orthotopic models of oral squamous cell carcinoma, blocking the FA receptor CD36 *in vivo* with neutralizing antibodies at early time points of the disease completely prevented metastatic initiation, whereas metastatic regression was only partial when the treatment was initiated at the late stages of the disease (Pascual et al., 2017). In experimental mouse models of melanoma metastasis, the inhibition of *de novo* FA synthesis with Orlistat or thiazolidinediones, or by inhibiting mitochondrial FA β-oxidation with drugs such as etomoxir, a carnitine palmitoyltransferase 1A (CPT1) inhibitor, had anti-tumour effects (Malvi et al., 2015; Nieman et al., 2011; Seguin et al., 2012). Acetyl-CoA synthase (ACS) isoforms catalyse the conversion of long-chain FAs to acetyl-CoA; importantly, several isoforms of ACS show increased expression in different human tumours, including colorectal cancer (Cao et al., 2000). Consequently, the ACS inhibitor triacsin C is under clinical investigation for the treatment of ACS-dependent tumours (Cha and Lee, 2016).

As mentioned above, the activity of MAGL (which releases FAs from lipid reservoirs) is enhanced in a number of primary tumours, and its inhibition by the drug JZL184 results in reduced pathogenicity in murine models of melanoma and ovarian cancer (Nomura et al., 2010). Interestingly, a high-fat diet prevents the anti-tumour effect of MAGL inhibition, further emphasizing the key role played by exogenous dietary lipids in cancer, and highlighting the potential impact of dietary interventions as additional future therapeutic anti-tumour strategies. As an example of the benefits of dietary intervention, a recent study in both mouse models and humans has shown that, although obesity boosts the metastasis of breast cancer cells to the lung by recruiting neutrophils to the lung pre-metastatic niche, weight loss is sufficient to reverse this effect (Quail et al., 2017).

Finally, the intimate interplay between metabolism and epigenetic mechanisms in cancer opens the possibility of targeting both to obtain synergistic effects (Dawson and Kouzarides, 2012; Rinaldi et al., 2016). The recent development of epigenetic inhibitors, which are already being tested in clinical trials, offers hope in this direction (Avgustinova and Benitah, 2016; Tough et al., 2014).

Conclusions

Despite intense investigation, the identity of metastatic cells remains elusive, preventing the research community from developing effective anti-metastatic therapies. Several recent studies discussed in this Review indicate that metabolism plays a causal role in the different phenotypic states exhibited by cancer cells during cancer progression and metastasis. Given the plethora of metabolic mechanisms in each organ, it is not surprising that the CSCs that initiate and promote metastasis need to exhibit metabolic plasticity in order to hijack multiple metabolic pathways and to survive hostile microenvironments during the metastatic cascade. Future investigations are needed to characterize and identify the cellular metabolic requirements for initiating metastasis, and to profile the metabolome of pro-metastatic and anti-metastatic niches. Such studies should provide clues as to the requirements for treatment and might also identify a window of opportunity for treatment, even leading to disease prevention prior to clinical detection.

The idea that local and systemic factors might confer metastatic cells with enough plasticity to successfully colonize distant organs suggests that influences derived from lifestyles, such as our diet or other daily habits, exert a strong influence on tumour progression, and that such factors could be easily modulated if understood. Increasing data supports the importance of diet in the crosstalk between metabolism and epigenetic regulation (Stefanska et al., 2012), and the extent to which epigenomic changes occur in metastasis (Wong et al., 2017). In the future, it will be necessary to discern which particular components of a diet lead to chromatin changes that are permissive for metastasis. Additionally, the identification of such chromatin modifications and the chromatin-modulating enzymes responsible will point to new potential epigenetic targets for therapeutic intervention.

This article is part of a special subject collection 'Cancer Metabolism: models, mechanisms and targets', which was launched in a dedicated issue guest edited by Almut Schulze and Mariia Yuneva. See related articles in this collection at http://dmm.biologists.org/collection/cancermetabolism.

Acknowledgements

We thank Veronica Raker for editing the manuscript.

Competing interests

The authors declare no competing or financial interests.

Funding

Research in the lab of S.A.B. is supported by the European Research Council (ERC), the Government of Cataluña (Generalitat de Catalunya; SGR grant), the Government of Spain [Ministerio de Economía y Competitividad (MINECO)], the Worldwide Cancer Research Foundation, and the Fundación Botín and Banco Santander (Fundación Banco Santander), through Santander Universities. D.D. is the recipient of a La Caixa International ("la Caixa" Foundation) PhD Fellowship. IRB Barcelona is the recipient of a Severo Ochoa Award of Excellence from MINECO (Government of Spain).

References

Abraham, B. J., Hnisz, D., Weintraub, A. S., Kwiatkowski, N., Li, C. H., Li, Z., Weichert-Leahey, N., Rahman, S., Liu, Y., Etchin, J. et al. (2017). Corrigendum: small genomic insertions form enhancers that misregulate oncogenes. *Nat. Commun.* **8**, 15797.

Avgustinova, A. and Benitah, S. A. (2016). Epigenetic control of adult stem cell function. *Nat. Rev. Mol. Cell Biol.* **17**, 643-658.

Baenke, F., Peck, B., Miess, H. and Schulze, A. (2013). Hooked on fat: the role of lipid synthesis in cancer metabolism and tumour development. *Dis. Model. Mech.* **6**, 1353-1363.

Bal, E., Park, H.-S., Belaid-Choucair, Z., Kayserili, H., Naville, M., Madrange, M., Chiticariu, E., Hadj-Rabia, S., Cagnard, N., Kuonen, F. et al. (2017). Mutations in ACTRT1 and its enhancer RNA elements lead to aberrant activation of Hedgehog signaling in inherited and sporadic basal cell carcinomas. *Nat. Med.* **23**, 1226-1233.

Baubec, T., Colombo, D. F., Wirbelauer, C., Schmidt, J., Burger, L., Krebs, A. R., Akalin, A. and Schübeler, D. (2015). Genomic profiling of DNA methyltransferases reveals a role for DNMT3B in genic methylation. *Nature* **520**, 243-247.

Bell, R. E., Golan, T., Sheinboim, D., Malcov, H., Amar, D., Salamon, A., Liron, T., Gelfman, S., Gabet, Y., Shamir, R. et al. (2016). Enhancer methylation dynamics contribute to cancer plasticity and patient mortality. *Genome Res.* **26**, 601-611.

Ben-David, U., Ha, G., Tseng, Y.-Y., Greenwald, N. F., Oh, C., Shih, J., McFarland, J. M., Wong, B., Boehm, J. S., Beroukhim, R. et al. (2017). Patient-derived xenografts undergo mouse-specific tumor evolution. *Nat. Genet.* **49**, 1567-1575.

Benjamin, D. I., Cravatt, B. F. and Nomura, D. K. (2012). Global profiling strategies for mapping dysregulated metabolic pathways in cancer. *Cell Metab.* **16**, 565-577.

Bensaad, K., Favaro, E., Lewis, C. A., Peck, B., Lord, S., Collins, J. M., Pinnick, K. E., Wigfield, S., Buffa, F. M., Li, J.-L. et al. (2014). Fatty acid uptake and lipid storage induced by HIF-1alpha contribute to cell growth and survival after hypoxia-reoxygenation. *Cell Rep.* **9**, 349-365.

Blanpain, C. (2013). Tracing the cellular origin of cancer. *Nat. Cell Biol.* **15**, 126-134.

Bonuccelli, G., Tsirigos, A., Whitaker-Menezes, D., Pavlides, S., Pestell, R. G., Chiavarina, B., Frank, P. G., Flomenberg, N., Howell, A., Martinez-Outschoorn, U. E. et al. (2010). Ketones and lactate "fuel" tumor growth and metastasis: evidence that epithelial cancer cells use oxidative mitochondrial metabolism. *Cell Cycle* **9**, 3506-3514.

Bradner, J. E., Hnisz, D. and Young, R. A. (2017). Transcriptional addiction in cancer. *Cell* **168**, 629-643.

Calon, A., Espinet, E., Palomo-Ponce, S., Tauriello, D. V. F., Iglesias, M., Céspedes, M. V., Sevillano, M., Nadal, C., Jung, P., Zhang, X. H.-F. et al.

(2012). Dependency of colorectal cancer on a TGF-beta-driven program in stromal cells for metastasis initiation. *Cancer Cell* **22**, 571-584.
Calon, A., Lonardo, E., Berenguer-Llergo, A., Espinet, E., Hernando-Momblona, X., Iglesias, M., Sevillano, M., Palomo-Ponce, S., Tauriello, D. V. F., Byrom, D. et al. (2015). Stromal gene expression defines poor-prognosis subtypes in colorectal cancer. *Nat. Genet.* **47**, 320-329.
Camarda, R., Zhou, A. Y., Kohnz, R. A., Balakrishnan, S., Mahieu, C., Anderton, B., Eyob, H., Kajimura, S., Tward, A., Krings, G. et al. (2016). Inhibition of fatty acid oxidation as a therapy for MYC-overexpressing triple-negative breast cancer. *Nat. Med.* **22**, 427-432.
Campbell, P. J., Yachida, S., Mudie, L. J., Stephens, P. J., Pleasance, E. D., Stebbings, L. A., Morsberger, L. A., Latimer, C., McLaren, S., Lin, M.-L. et al. (2010). The patterns and dynamics of genomic instability in metastatic pancreatic cancer. *Nature* **467**, 1109-1113.
Cantelmo, A. R., Conradi, L.-C., Brajic, A., Goveia, J., Kalucka, J., Pircher, A., Chaturvedi, P., Hol, J., Thienpont, B., Teuwen, L.-A. et al. (2016). Inhibition of the glycolytic activator PFKFB3 in endothelium induces tumor vessel normalization, impairs metastasis, and improves chemotherapy. *Cancer Cell* **30**, 968-985.
Cao, Y., Pearman, A. T., Zimmerman, G. A., McIntyre, T. M. and Prescott, S. M. (2000). Intracellular unesterified arachidonic acid signals apoptosis. *Proc. Natl. Acad. Sci. USA* **97**, 11280-11285.
Cao, Z., Ding, B.-S., Guo, P., Lee, S. B., Butler, J. M., Casey, S. C., Simons, M., Tam, W., Felsher, D. W., Shido, K. et al. (2014). Angiocrine factors deployed by tumor vascular niche induce B cell lymphoma invasiveness and chemoresistance. *Cancer Cell* **25**, 350-365.
Cha, J.-Y. and Lee, H.-J. (2016). Targeting lipid metabolic reprogramming as anticancer therapeutics. *J. Cancer Prev.* **21**, 209-215.
Chaffer, C. L. and Weinberg, R. A. (2011). A perspective on cancer cell metastasis. *Science* **331**, 1559-1564.
Chapuy, B., McKeown, M. R., Lin, C. Y., Monti, S., Roemer, M. G. M., Qi, J., Rahl, P. B., Sun, H. H., Yeda, K. T., Doench, J. G. et al. (2013). Discovery and characterization of super-enhancer-associated dependencies in diffuse large B cell lymphoma. *Cancer Cell* **24**, 777-790.
Charlet, J., Duymich, C. E., Lay, F. D., Mundbjerg, K., Dalsgaard Sorensen, K., Liang, G. and Jones, P. A. (2016). Bivalent regions of cytosine methylation and H3K27 acetylation suggest an active role for DNA methylation at enhancers. *Mol. Cell* **62**, 422-431.
Chen, Q., Zhang, X. H.-F. and Massagué, J. (2011). Macrophage binding to receptor VCAM-1 transmits survival signals in breast cancer cells that invade the lungs. *Cancer Cell* **20**, 538-549.
Chen, J., Lee, H.-J., Wu, X., Huo, L., Kim, S.-J., Xu, L., Wang, Y., He, J., Bollu, L. R., Gao, G. et al. (2015). Gain of glucose-independent growth upon metastasis of breast cancer cells to the brain. *Cancer Res.* **75**, 554-565.
Choi, H.-J., Park, J.-H., Park, M., Won, H.-Y., Joo, H.-S., Lee, C. H., Lee, J.-Y. and Kong, G. (2015). UTX inhibits EMT-induced breast CSC properties by epigenetic repression of EMT genes in cooperation with LSD1 and HDAC1. *EMBO Rep.* **16**, 1288-1298.
Christen, S., Lorendeau, D., Schmieder, R., Broekaert, D., Metzger, K., Veys, K., Elia, I., Buescher, J. M., Orth, M. F., Davidson, S. M. et al. (2016). Breast cancer-derived lung metastases show increased pyruvate carboxylase-dependent anaplerosis. *Cell Rep.* **17**, 837-848.
Collins, A. T., Berry, P. A., Hyde, C., Stower, M. J. and Maitland, N. J. (2005). Prospective identification of tumorigenic prostate cancer stem cells. *Cancer Res.* **65**, 10946-10951.
Comerford, S. A., Huang, Z., Du, X., Wang, Y., Cai, L., Witkiewicz, A. K., Walters, H., Tantawy, M. N., Fu, A., Manning, H. C. et al. (2014). Acetate dependence of tumors. *Cell* **159**, 1591-1602.
Costa-Silva, B., Aiello, N. M., Ocean, A. J., Singh, S., Zhang, H., Thakur, B. K., Becker, A., Hoshino, A., Mark, M. T., Molina, H. et al. (2015). Pancreatic cancer exosomes initiate pre-metastatic niche formation in the liver. *Nat. Cell Biol.* **17**, 816-826.
Curtis, S. J., Sinkevicius, K. W., Li, D., Lau, A. N., Roach, R. R., Zamponi, R., Woolfenden, A. E., Kirsch, D. G., Wong, K.-K. and Kim, C. F. (2010). Primary tumor genotype is an important determinant in identification of lung cancer propagating cells. *Cell Stem Cell* **7**, 127-133.
Dawson, M. A. and Kouzarides, T. (2012). Cancer epigenetics: from mechanism to therapy. *Cell* **150**, 12-27.
Denny, S. K., Yang, D., Chuang, C.-H., Brady, J. J., Lim, J. S., Grüner, B. M., Chiou, S.-H., Schep, A. N., Baral, J., Hamard, C. et al. (2016). Nfib promotes metastasis through a widespread increase in chromatin accessibility. *Cell* **166**, 328-342.
Dhayalan, A., Rajavelu, A., Rathert, P., Tamas, R., Jurkowska, R. Z., Ragozin, S. and Jeltsch, A. (2010). The Dnmt3a PWWP domain reads histone 3 lysine 36 trimethylation and guides DNA methylation. *J. Biol. Chem.* **285**, 26114-26120.
Dieter, S. M., Ball, C. R., Hoffmann, C. M., Nowrouzi, A., Herbst, F., Zavidij, O., Abel, U., Arens, A., Weichert, W., Brand, K. et al. (2011). Distinct types of tumor-initiating cells form human colon cancer tumors and metastases. *Cell Stem Cell* **9**, 357-365.
Domcke, S., Bardet, A. F., Adrian Ginno, P., Hartl, D., Burger, L. and Schübeler, D. (2015). Competition between DNA methylation and transcription factors determines binding of NRF1. *Nature* **528**, 575-579.
Dong, C., Yuan, T., Wu, Y., Wang, Y., Fan, T. W. M., Miriyala, S., Lin, Y., Yao, J., Shi, J., Kang, T. et al. (2013). Loss of FBP1 by Snail-mediated repression provides metabolic advantages in basal-like breast cancer. *Cancer Cell* **23**, 316-331.
Dupuy, F., Tabariès, S., Andrzejewski, S., Dong, Z., Blagih, J., Annis, M. G., Omeroglu, A., Gao, D., Leung, S., Amir, E. et al. (2015). PDK1-dependent metabolic reprogramming dictates metastatic potential in breast cancer. *Cell Metab.* **22**, 577-589.
Elia, I., Broekaert, D., Christen, S., Boon, R., Radaelli, E., Orth, M. F., Verfaillie, C., Grunewald, T. G. P. and Fendt, S.-M. (2017). Proline metabolism supports metastasis formation and could be inhibited to selectively target metastasizing cancer cells. *Nat. Commun.* **8**, 15267.
Eppert, K., Takenaka, K., Lechman, E. R., Waldron, L., Nilsson, B., van Galen, P., Metzeler, K. H., Poeppl, A., Ling, V., Beyene, J. et al. (2011). Stem cell gene expression programs influence clinical outcome in human leukemia. *Nat. Med.* **17**, 1086-1093.
Eramo, A., Lotti, F., Sette, G., Pilozzi, E., Biffoni, M., Di Virgilio, A., Conticello, C., Ruco, L., Peschle, C. and De Maria, R. (2008). Identification and expansion of the tumorigenic lung cancer stem cell population. *Cell Death Differ.* **15**, 504-514.
Fan, J., Krautkramer, K. A., Feldman, J. L. and Denu, J. M. (2015). Metabolic regulation of histone post-translational modifications. *ACS Chem. Biol.* **10**, 95-108.
Feinberg, A. P. and Vogelstein, B. (1983). Hypomethylation of ras oncogenes in primary human cancers. *Biochem. Biophys. Res. Commun.* **111**, 47-54.
Fong, M. Y., Zhou, W., Liu, L., Alontaga, A. Y., Chandra, M., Ashby, J., Chow, A., O'Connor, S. T., Li, S., Chin, A. R. et al. (2015). Breast-cancer-secreted miR-122 reprograms glucose metabolism in premetastatic niche to promote metastasis. *Nat. Cell Biol.* **17**, 183-194.
German, N. J., Yoon, H., Yusuf, R. Z., Murphy, J. P., Finley, L. W. S., Laurent, G., Haas, W., Satterstrom, F. K., Guarnerio, J., Zaganjor, E. et al. (2016). PHD3 loss in cancer enables metabolic reliance on fatty acid oxidation via deactivation of ACC2. *Mol. Cell* **63**, 1006-1020.
Ghajar, C. M., Peinado, H., Mori, H., Matei, I. R., Evason, K. J., Brazier, H., Almeida, D., Koller, A., Hajjar, K. A., Stainier, D. Y. R. et al. (2013). The perivascular niche regulates breast tumour dormancy. *Nat. Cell Biol.* **15**, 807-817.
Goel, H. L. and Mercurio, A. M. (2013). VEGF targets the tumour cell. *Nat. Rev. Cancer* **13**, 871-882.
Hansen, K. D., Timp, W., Bravo, H. C., Sabunciyan, S., Langmead, B., McDonald, O. G., Wen, B., Wu, H., Liu, Y., Diep, D., et al. (2011). Increased methylation variation in epigenetic domains across cancer types. *Nat. Genet.* **43**, 768-775.
Hermann, P. C., Huber, S. L., Herrler, T., Aicher, A., Ellwart, J. W., Guba, M., Bruns, C. J. and Heeschen, C. (2007). Distinct populations of cancer stem cells determine tumor growth and metastatic activity in human pancreatic cancer. *Cell Stem Cell* **1**, 313-323.
Herz, H.-M., Hu, D. and Shilatifard, A. (2014). Enhancer malfunction in cancer. *Mol. Cell* **53**, 859-866.
Jacob, L. S., Vanharanta, S., Obenauf, A. C., Pirun, M., Viale, A., Socci, N. D. and Massagué, J. (2015). Metastatic competence can emerge with selection of preexisting oncogenic alleles without a need of new mutations. *Cancer Res.* **75**, 3713-3719.
Jones, P. A. (2012). Functions of DNA methylation: islands, start sites, gene bodies and beyond. *Nat. Rev. Genet.* **13**, 484-492.
Kalluri, R. and Zeisberg, M. (2006). Fibroblasts in cancer. *Nat. Rev. Cancer* **6**, 392-401.
Kaplan, R. N., Riba, R. D., Zacharoulis, S., Bramley, A. H., Vincent, L., Costa, C., MacDonald, D. D., Jin, D. K., Shido, K., Kerns, S. A. et al. (2005). VEGFR1-positive haematopoietic bone marrow progenitors initiate the pre-metastatic niche. *Nature* **438**, 820-827.
Katada, S., Imhof, A. and Sassone-Corsi, P. (2012). Connecting threads: epigenetics and metabolism. *Cell* **148**, 24-28.
Keckesova, Z., Donaher, J. L., De Cock, J., Freinkman, E., Lingrell, S., Bachovchin, D. A., Bierie, B., Tischler, V., Noske, A., Okondo, M. C. et al. (2017). LACTB is a tumour suppressor that modulates lipid metabolism and cell state. *Nature* **543**, 681-686.
Kinnaird, A., Zhao, S., Wellen, K. E. and Michelakis, E. D. (2016). Metabolic control of epigenetics in cancer. *Nat. Rev. Cancer* **16**, 694-707.
Koppenol, W. H., Bounds, P. L. and Dang, C. V. (2011). Otto Warburg's contributions to current concepts of cancer metabolism. *Nat. Rev. Cancer* **11**, 325-337.
Kottakis, F., Nicolay, B. N., Roumane, A., Karnik, R., Gu, H., Nagle, J. M., Boukhali, M., Hayward, M. C., Li, Y. Y., Chen, T. et al. (2016). LKB1 loss links serine metabolism to DNA methylation and tumorigenesis. *Nature* **539**, 390-395.
Kreso, A. and Dick, J. E. (2014). Evolution of the cancer stem cell model. *Cell Stem Cell* **14**, 275-291.

Kreso, A., O'Brien, C. A., van Galen, P., Gan, O. I., Notta, F., Brown, A. M. K., Ng, K., Ma, J., Wienholds, E., Dunant, C. et al. (2013). Variable clonal repopulation dynamics influence chemotherapy response in colorectal cancer. *Science* **339**, 543-548.

Lagadinou, E. D., Sach, A., Callahan, K., Rossi, R. M., Neering, S. J., Minhajuddin, M., Ashton, J. M., Pei, S., Grose, V., O'Dwyer, K. M. et al. (2013). BCL-2 inhibition targets oxidative phosphorylation and selectively eradicates quiescent human leukemia stem cells. *Cell Stem Cell* **12**, 329-341.

Lawson, D. A., Bhakta, N. R., Kessenbrock, K., Prummel, K. D., Yu, Y., Takai, K., Zhou, A., Eyob, H., Balakrishnan, S., Wang, C.-Y. et al. (2015). Single-cell analysis reveals a stem-cell program in human metastatic breast cancer cells. *Nature* **526**, 131-135.

LeBleu, V. S., O'Connell, J. T., Gonzalez Herrera, K. N., Wikman, H., Pantel, K., Haigis, M. C., de Carvalho, F. M., Damascena, A., Domingos Chinen, L. T., Rocha, R. M. et al. (2014). PGC-1alpha mediates mitochondrial biogenesis and oxidative phosphorylation in cancer cells to promote metastasis. *Nat. Cell Biol.* **16**, 992-1003, 1-15.

Li, C., Heidt, D. G., Dalerba, P., Burant, C. F., Zhang, L., Adsay, V., Wicha, M., Clarke, M. F. and Simeone, D. M. (2007). Identification of pancreatic cancer stem cells. *Cancer Res.* **67**, 1030-1037.

Li, J., Condello, S., Thomes-Pepin, J., Ma, X., Xia, Y., Hurley, T. D., Matei, D. and Cheng, J. X. (2017). Lipid desaturation is a metabolic marker and therapeutic target of ovarian cancer stem cells. *Cell Stem Cell* **20**, 303-314.e5.

Loo, J. M., Scherl, A., Nguyen, A., Man, F. Y., Weinberg, E., Zeng, Z., Saltz, L., Paty, P. B. and Tavazoie, S. F. (2015). Extracellular metabolic energetics can promote cancer progression. *Cell* **160**, 393-406.

Louie, S. M., Roberts, L. S., Mulvihill, M. M., Luo, K. and Nomura, D. K. (2013). Cancer cells incorporate and remodel exogenous palmitate into structural and oncogenic signaling lipids. *Biochim. Biophys. Acta* **1831**, 1566-1572.

Lu, X., Mu, E., Wei, Y., Riethdorf, S., Yang, Q., Yuan, M., Yan, J., Hua, Y., Tiede, B. J., Lu, X. et al. (2011). VCAM-1 promotes osteolytic expansion of indolent bone micrometastasis of breast cancer by engaging alpha4beta1-positive osteoclast progenitors. *Cancer Cell* **20**, 701-714.

Makohon-Moore, A. P., Zhang, M., Reiter, J. G., Bozic, I., Allen, B., Kundu, D., Chatterjee, K., Wong, F., Jiao, Y., Kohutek, Z. A. et al. (2017). Limited heterogeneity of known driver gene mutations among the metastases of individual patients with pancreatic cancer. *Nat. Genet.* **49**, 358-366.

Malvi, P., Chaube, B., Pandey, V., Vijayakumar, M. V., Boreddy, P. R., Mohammad, N., Singh, S. V. and Bhat, M. K. (2015). Obesity induced rapid melanoma progression is reversed by orlistat treatment and dietary intervention: role of adipokines. *Mol. Oncol.* **9**, 689-703.

Mansour, M. R., Abraham, B. J., Anders, L., Berezovskaya, A., Gutierrez, A., Durbin, A. D., Etchin, J., Lawton, L., Sallan, S. E., Silverman, L. B. et al. (2014). Oncogene regulation. An oncogenic super-enhancer formed through somatic mutation of a noncoding intergenic element. *Science* **346**, 1373-1377.

Marin-Valencia, I., Yang, C., Mashimo, T., Cho, S., Baek, H., Yang, X.-L., Rajagopalan, K. N., Maddie, M., Vemireddy, V., Zhao, Z. et al. (2012). Analysis of tumor metabolism reveals mitochondrial glucose oxidation in genetically diverse human glioblastomas in the mouse brain in vivo. *Cell Metab.* **15**, 827-837.

Mashimo, T., Pichumani, K., Vemireddy, V., Hatanpaa, K. J., Singh, D. K., Sirasanagandla, S., Nannepaga, S., Piccirillo, S. G., Kovacs, Z., Foong, C. et al. (2014). Acetate is a bioenergetic substrate for human glioblastoma and brain metastases. *Cell* **159**, 1603-1614.

McAllister, S. S. and Weinberg, R. A. (2014). The tumour-induced systemic environment as a critical regulator of cancer progression and metastasis. *Nat. Cell Biol.* **16**, 717-727.

McDonald, O. G., Li, X., Saunders, T., Tryggvadottir, R., Mentch, S. J., Warmoes, M. O., Word, A. E., Carrer, A., Salz, T. H., Natsume, S. et al. (2017). Epigenomic reprogramming during pancreatic cancer progression links anabolic glucose metabolism to distant metastasis. *Nat. Genet.* **49**, 367-376.

McDonnell, E., Crown, S. B., Fox, D. B., Kitir, B., Ilkayeva, O. R., Olsen, C. A., Grimsrud, P. A. and Hirschey, M. D. (2016). Lipids reprogram metabolism to become a major carbon source for histone acetylation. *Cell Rep.* **17**, 1463-1472.

Mentch, S. J., Mehrmohamadi, M., Huang, L., Liu, X., Gupta, D., Mattocks, D., Gómez Padilla, P., Ables, G., Bamman, M. M., Thalacker-Mercer, A. E. et al. (2015). Histone methylation dynamics and gene regulation occur through the sensing of one-carbon metabolism. *Cell Metab.* **22**, 861-873.

Merlo, L. M. F., Pepper, J. W., Reid, B. J. and Maley, C. C. (2006). Cancer as an evolutionary and ecological process. *Nat. Rev. Cancer* **6**, 924-935.

Merlos-Suárez, A., Barriga, F. M., Jung, P., Iglesias, M., Céspedes, M. V., Rossell, D., Sevillano, M., Hernando-Momblona, X., da Silva-Diz, V., Muñoz, P. et al. (2011). The intestinal stem cell signature identifies colorectal cancer stem cells and predicts disease relapse. *Cell Stem Cell* **8**, 511-524.

Mews, P., Donahue, G., Drake, A. M., Luczak, V., Abel, T. and Berger, S. L. (2017). Acetyl-CoA synthetase regulates histone acetylation and hippocampal memory. *Nature* **546**, 381-386.

Michailidou, K., Lindstrom, S., Dennis, J., Beesley, J., Hui, S., Kar, S., Lemacon, A., Soucy, P., Glubb, D., Rostamianfar, A. et al. (2017). Association analysis identifies 65 new breast cancer risk loci. *Nature* **551**, 92-94.

Minna, J. D. and Johnson, J. E. (2016). Opening a chromatin gate to metastasis. *Cell* **166**, 275-276.

Mudduluru, G., Ilm, K., Fuchs, S. and Stein, U. (2017). Epigenetic silencing of miR-520c leads to induced S100A4 expression and its mediated colorectal cancer progression. *Oncotarget* **8**, 21081-21094.

Nath, A. and Chan, C. (2016). Genetic alterations in fatty acid transport and metabolism genes are associated with metastatic progression and poor prognosis of human cancers. *Sci. Rep.* **6**, 18669.

Nath, A., Li, I., Roberts, L. R. and Chan, C. (2015). Elevated free fatty acid uptake via CD36 promotes epithelial-mesenchymal transition in hepatocellular carcinoma. *Sci. Rep.* **5**, 14752.

Nieman, K. M., Kenny, H. A., Penicka, C. V., Ladanyi, A., Buell-Gutbrod, R., Zillhardt, M. R., Romero, I. L., Carey, M. S., Mills, G. B., Hotamisligil, G. S. et al. (2011). Adipocytes promote ovarian cancer metastasis and provide energy for rapid tumor growth. *Nat. Med.* **17**, 1498-1503.

Nomura, D. K., Long, J. Z., Niessen, S., Hoover, H. S., Ng, S.-W. and Cravatt, B. F. (2010). Monoacylglycerol lipase regulates a fatty acid network that promotes cancer pathogenesis. *Cell* **140**, 49-61.

Nowell, P. C. (1976). The clonal evolution of tumor cell populations. *Science* **194**, 23-28.

Obenauf, A. C., Zou, Y., Ji, A. L., Vanharanta, S., Shu, W., Shi, H., Kong, X., Bosenberg, M. C., Wiesner, T., Rosen, N. et al. (2015). Therapy-induced tumour secretomes promote resistance and tumour progression. *Nature* **520**, 368-372.

Oldridge, D. A., Wood, A. C., Weichert-Leahey, N., Crimmins, I., Sussman, R., Winter, C., McDaniel, L. D., Diamond, M., Hart, L. S., Zhu, S. et al. (2015). Genetic predisposition to neuroblastoma mediated by a LMO1 super-enhancer polymorphism. *Nature* **528**, 418-421.

Oskarsson, T., Acharyya, S., Zhang, X. H.-F., Vanharanta, S., Tavazoie, S. F., Morris, P. G., Downey, R. J., Manova-Todorova, K., Brogi, E. and Massagué, J. (2011). Breast cancer cells produce tenascin C as a metastatic niche component to colonize the lungs. *Nat. Med.* **17**, 867-874.

Oskarsson, T., Batlle, E. and Massagué, J. (2014). Metastatic stem cells: sources, niches, and vital pathways. *Cell Stem Cell* **14**, 306-321.

Paolino, M., Choidas, A., Wallner, S., Pranjic, B., Uribesalgo, I., Loeser, S., Jamieson, A. M., Langdon, W. Y., Ikeda, F., Fededa, J. P. et al. (2014). The E3 ligase Cbl-b and TAM receptors regulate cancer metastasis via natural killer cells. *Nature* **507**, 508-512.

Park, J. H., Vithayathil, S., Kumar, S., Sung, P.-L., Dobrolecki, L. E., Putluri, V., Bhat, V. B., Bhowmik, S. K., Gupta, V., Arora, K. et al. (2016). Fatty acid oxidation-driven src links mitochondrial energy reprogramming and oncogenic properties in triple-negative breast cancer. *Cell Rep.* **14**, 2154-2165.

Pascual, G., Avgustinova, A., Mejetta, S., Martín, M., Castellanos, A., Attolini, C. S.-O., Berenguer, A., Prats, N., Toll, A., Hueto, J. A. et al. (2017). Targeting metastasis-initiating cells through the fatty acid receptor CD36. *Nature* **541**, 41-45.

Patel, S. A. and Vanharanta, S. (2016). Epigenetic determinants of metastasis. *Mol. Oncol.* **11**, 79-96.

Patrawala, L., Calhoun, T., Schneider-Broussard, R., Li, H., Bhatia, B., Tang, S., Reilly, J. G., Chandra, D., Zhou, J., Claypool, K. et al. (2006). Highly purified CD44+ prostate cancer cells from xenograft human tumors are enriched in tumorigenic and metastatic progenitor cells. *Oncogene* **25**, 1696-1708.

Pece, S., Tosoni, D., Confalonieri, S., Mazzarol, G., Vecchi, M., Ronzoni, S., Bernard, L., Viale, G., Pelicci, P. G. and Di Fiore, P. P. (2010). Biological and molecular heterogeneity of breast cancers correlates with their cancer stem cell content. *Cell* **140**, 62-73.

Pein, M. and Oskarsson, T. (2015). Microenvironment in metastasis: roadblocks and supportive niches. *Am. J. Physiol. Cell Physiol.* **309**, C627-C638.

Peinado, H., Alečković, M., Lavotshkin, S., Matei, I., Costa-Silva, B., Moreno-Bueno, G., Hergueta-Redondo, M., Williams, C., García-Santos, G., Ghajar, C. M. et al. (2012). Melanoma exosomes educate bone marrow progenitor cells toward a pro-metastatic phenotype through MET. *Nat. Med.* **18**, 883-891.

Peinado, H., Zhang, H., Matei, I. R., Costa-Silva, B., Hoshino, A., Rodrigues, G., Psaila, B., Kaplan, R. N., Bromberg, J. F., Kang, Y. et al. (2017). Pre-metastatic niches: organ-specific homes for metastases. *Nat. Rev. Cancer* **17**, 302-317.

Prince, M. E., Sivanandan, R., Kaczorowski, A., Wolf, G. T., Kaplan, M. J., Dalerba, P., Weissman, I. L., Clarke, M. F. and Ailles, L. E. (2007). Identification of a subpopulation of cells with cancer stem cell properties in head and neck squamous cell carcinoma. *Proc. Natl. Acad. Sci. USA* **104**, 973-978.

Puente, X. S., Beà, S., Valdés-Mas, R., Villamor, N., Gutiérrez-Abril, J., Martín-Subero, J. I., Munar, M., Rubio-Pérez, C., Jares, P., Aymerich, M. et al. (2015). Non-coding recurrent mutations in chronic lymphocytic leukaemia. *Nature* **526**, 519-524.

Pufulete, M., Al-Ghnaniem, R., Khushal, A., Appleby, P., Harris, N., Gout, S., Emery, P. W. and Sanders, T. A. (2005). Effect of folic acid supplementation on genomic DNA methylation in patients with colorectal adenoma. *Gut* **54**, 648-653.

Quail, D. F., Olson, O. C., Bhardwaj, P., Walsh, L. A., Akkari, L., Quick, M. L., Chen, I.-C., Wendel, N., Ben-Chetrit, N., Walker, J. et al. (2017). Obesity alters the lung myeloid cell landscape to enhance breast cancer metastasis through IL5 and GM-CSF. *Nat. Cell Biol.* **19**, 974-987.

Ramaswamy, S., Ross, K. N., Lander, E. S. and Golub, T. R. (2003). A molecular signature of metastasis in primary solid tumors. *Nat. Genet.* **33**, 49-54.

Rankin, E. B. and Giaccia, A. J. (2016). Hypoxic control of metastasis. *Science* **352**, 175-180.
Ricci-Vitiani, L., Lombardi, D. G., Pilozzi, E., Biffoni, M., Todaro, M., Peschle, C. and De Maria, R. (2007). Identification and expansion of human colon-cancer-initiating cells. *Nature* **445**, 111-115.
Rinaldi, L. and Benitah, S. A. (2015). Epigenetic regulation of adult stem cell function. *FEBS J.* **282**, 1589-1604.
Rinaldi, L., Datta, D., Serrat, J., Morey, L., Solanas, G., Avgustinova, A., Blanco, E., Pons, J. I., Matallanas, D., Von Kriegsheim, A. et al. (2016). Dnmt3a and Dnmt3b Associate with Enhancers to Regulate Human Epidermal Stem Cell Homeostasis. *Cell Stem Cell* **19**, 491-501.
Rinaldi, L., Avgustinova, A., Martin, M., Datta, D., Solanas, G., Prats, N. and Benitah, S. A. (2017). Loss of Dnmt3a and Dnmt3b does not affect epidermal homeostasis but promotes squamous transformation through PPAR-gamma. *eLife* **6**, e21697.
Roe, J. S., Hwang, C. I., Somerville, T. D. D., Milazzo, J. P., Lee, E. J., Da Silva, B., Maiorino, L., Tiriac, H., Young, C. M., Miyabayashi, K. et al. (2017). Enhancer reprogramming promotes pancreatic cancer metastasis. *Cell* **170**, 875-888.e20.
Roesch, A., Fukunaga-Kalabis, M., Schmidt, E. C., Zabierowski, S. E., Brafford, P. A., Vultur, A., Basu, D., Gimotty, P., Vogt, T. and Herlyn, M. (2010). A temporarily distinct subpopulation of slow-cycling melanoma cells is required for continuous tumor growth. *Cell* **141**, 583-594.
Rokavec, M., Horst, D. and Hermeking, H. (2017). Cellular model of colon cancer progression reveals signatures of mRNAs, miRNA, lncRNAs, and epigenetic modifications associated with metastasis. *Cancer Res.* **77**, 1854-1867.
Sancho, P., Burgos-Ramos, E., Tavera, A., Bou Kheir, T., Jagust, P., Schoenhals, M., Barneda, D., Sellers, K., Campos-Olivas, R., Graña, O. et al. (2015). MYC/PGC-1alpha balance determines the metabolic phenotype and plasticity of pancreatic cancer stem cells. *Cell Metab.* **22**, 590-605.
Schug, Z. T., Peck, B., Jones, D. T., Zhang, Q., Grosskurth, S., Alam, I. S., Goodwin, L. M., Smethurst, E., Mason, S., Blyth, K. et al. (2015). Acetyl-CoA synthetase 2 promotes acetate utilization and maintains cancer cell growth under metabolic stress. *Cancer Cell* **27**, 57-71.
Scott-Browne, J. P., Cw, L. and Rao, A. (2017). TET proteins in natural and induced differentiation. *Curr. Opin. Genet. Dev.* **46**, 202-208.
Seguin, F., Carvalho, M. A., Bastos, D. C., Agostini, M., Zecchin, K. G., Alvarez-Flores, M. P., Chudzinski-Tavassi, A. M., Coletta, R. D. and Graner, E. (2012). The fatty acid synthase inhibitor orlistat reduces experimental metastases and angiogenesis in B16-F10 melanomas. *Br. J. Cancer* **107**, 977-987.
Seisenberger, S., Andrews, S., Krueger, F., Arand, J., Walter, J., Santos, F., Popp, C., Thienpont, B., Dean, W. and Reik, W. (2012). The dynamics of genome-wide DNA methylation reprogramming in mouse primordial germ cells. *Mol. Cell* **48**, 849-862.
Semenza, G. L. (2010). Defining the role of hypoxia-inducible factor 1 in cancer biology and therapeutics. *Oncogene* **29**, 625-634.
Semenza, G. L. (2016). Targeting hypoxia-inducible factor 1 to stimulate tissue vascularization. *J. Investig. Med.* **64**, 361-363.
Sevenich, L., Bowman, R. L., Mason, S. D., Quail, D. F., Rapaport, F., Elie, B. T., Brogi, E., Brastianos, P. K., Hahn, W. C., Holsinger, L. J. et al. (2014). Analysis of tumour- and stroma-supplied proteolytic networks reveals a brain-metastasis-promoting role for cathepsin S. *Nat. Cell Biol.* **16**, 876-888.
Shi, X., Tasdogan, A., Huang, F., Hu, Z., Morrison, S. J. and DeBerardinis, R. J. (2017). The abundance of metabolites related to protein methylation correlates with the metastatic capacity of human melanoma xenografts. *Sci. Adv.* **3**, eaao5268.
Shiozawa, Y., Pedersen, E. A., Havens, A. M., Jung, Y., Mishra, A., Joseph, J., Kim, J. K., Patel, L. R., Ying, C., Ziegler, A. M. et al. (2011). Human prostate cancer metastases target the hematopoietic stem cell niche to establish footholds in mouse bone marrow. *J. Clin. Invest.* **121**, 1298-1312.
Shlush, L. I., Mitchell, A., Heisler, L., Abelson, S., Ng, S. W. K., Trotman-Grant, A., Medeiros, J. J. F., Rao-Bhatia, A., Jaciw-Zurakowsky, I., Marke, R. et al. (2017). Tracing the origins of relapse in acute myeloid leukaemia to stem cells. *Nature* **547**, 104-108.
Singh, S. K., Clarke, I. D., Hide, T. and Dirks, P. B. (2004). Cancer stem cells in nervous system tumors. *Oncogene* **23**, 7267-7273.
Sounni, N. E., Cimino, J., Blacher, S., Primac, I., Truong, A., Mazzucchelli, G., Paye, A., Calligaris, D., Debois, D., De Tullio, P. et al. (2014). Blocking lipid synthesis overcomes tumor regrowth and metastasis after antiangiogenic therapy withdrawal. *Cell Metab.* **20**, 280-294.
Stefanska, B., Karlic, H., Varga, F., Fabianowska-Majewska, K. and Haslberger, A. G. (2012). Epigenetic mechanisms in anti-cancer actions of bioactive food components–the implications in cancer prevention. *Br. J. Pharmacol.* **167**, 279-297.
Sutendra, G., Kinnaird, A., Dromparis, P., Paulin, R., Stenson, T. H., Haromy, A., Hashimoto, K., Zhang, N., Flaim, E. and Michelakis, E. D. (2014). A nuclear pyruvate dehydrogenase complex is important for the generation of acetyl-CoA and histone acetylation. *Cell* **158**, 84-97.
Svensson, R. U., Parker, S. J., Eichner, L. J., Kolar, M. J., Wallace, M., Brun, S. N., Lombardo, P. S., Van Nostrand, J. L., Hutchins, A., Vera, L. et al. (2016). Inhibition of acetyl-CoA carboxylase suppresses fatty acid synthesis and tumor growth of non-small-cell lung cancer in preclinical models. *Nat. Med.* **22**, 1108-1119.
Torrano, V., Valcarcel-Jimenez, L., Cortazar, A. R., Liu, X., Urosevic, J., Castillo-Martin, M., Fernández-Ruiz, S., Morciano, G., Caro-Maldonado, A., Guiu, M. et al. (2016). The metabolic co-regulator PGC1alpha suppresses prostate cancer metastasis. *Nat. Cell Biol.* **18**, 645-656.
Tough, D. F., Lewis, H. D., Rioja, I., Lindon, M. J. and Prinjha, R. K. (2014). Epigenetic pathway targets for the treatment of disease: accelerating progress in the development of pharmacological tools: IUPHAR Review 11. *Br. J. Pharmacol.* **171**, 4981-5010.
Ugel, S., De Sanctis, F., Mandruzzato, S. and Bronte, V. (2015). Tumor-induced myeloid deviation: when myeloid-derived suppressor cells meet tumor-associated macrophages. *J. Clin. Invest.* **125**, 3365-3376.
Valiente, M., Obenauf, A. C., Jin, X., Chen, Q., Zhang, X. H.-F., Lee, D. J., Chaft, J. E., Kris, M. G., Huse, J. T., Brogi, E. et al. (2014). Serpins promote cancer cell survival and vascular co-option in brain metastasis. *Cell* **156**, 1002-1016.
Vander Heiden, M. G. and DeBerardinis, R. J. (2017). Understanding the Intersections between Metabolism and Cancer Biology. *Cell* **168**, 657-669.
Vander Heiden, M. G., Cantley, L. C. and Thompson, C. B. (2009). Understanding the Warburg effect: the metabolic requirements of cell proliferation. *Science* **324**, 1029-1033.
Vanharanta, S. and Massagué, J. (2013). Origins of metastatic traits. *Cancer Cell* **24**, 410-421.
Viale, A., Pettazzoni, P., Lyssiotis, C. A., Ying, H., Sánchez, N., Marchesini, M., Carugo, A., Green, T., Seth, S., Giuliani, V. et al. (2014). Oncogene ablation-resistant pancreatic cancer cells depend on mitochondrial function. *Nature* **514**, 628-632.
Vogelstein, B., Papadopoulos, N., Velculescu, V. E., Zhou, S., Diaz, L. A., Jr. and Kinzler, K. W. (2013). Cancer genome landscapes. *Science* **339**, 1546-1558.
Warburg, O. (1925). The metabolism of carcinoma cells. *J. Cancer Res.* **9**, 148-163.
Watanabe, D., Suetake, I., Tada, T. and Tajima, S. (2002). Stage- and cell-specific expression of Dnmt3a and Dnmt3b during embryogenesis. *Mech. Dev.* **118**, 187-190.
Wculek, S. K. and Malanchi, I. (2015). Neutrophils support lung colonization of metastasis-initiating breast cancer cells. *Nature* **528**, 413-417.
Wellen, K. E., Hatzivassiliou, G., Sachdeva, U. M., Bui, T. V., Cross, J. R. and Thompson, C. B. (2009). ATP-citrate lyase links cellular metabolism to histone acetylation. *Science* **324**, 1076-1080.
Wong, C. C., Qian, Y. and Yu, J. (2017). Interplay between epigenetics and metabolism in oncogenesis: mechanisms and therapeutic approaches. *Oncogene* **36**, 3359-3374.
Wu, X.-Z. (2008). Origin of cancer stem cells: the role of self-renewal and differentiation. *Ann. Surg. Oncol.* **15**, 407-414.
Yan, X., Hu, Z., Feng, Y., Hu, X., Yuan, J., Zhao, S. D., Zhang, Y., Yang, L., Shan, W., He, Q. et al. (2015). Comprehensive genomic characterization of long non-coding RNAs across human cancers. *Cancer Cell* **28**, 529-540.
Ye, H., Adane, B., Khan, N., Sullivan, T., Minhajuddin, M., Gasparetto, M., Stevens, B., Pei, S., Balys, M., Ashton, J. M. et al. (2016). Leukemic stem cells evade chemotherapy by metabolic adaptation to an adipose tissue niche. *Cell Stem Cell* **19**, 23-37.
Yin, Y., Morgunova, E., Jolma, A., Kaasinen, E., Sahu, B., Khund-Sayeed, S., Das, P. K., Kivioja, T., Dave, K., Zhong, F. et al. (2017). Impact of cytosine methylation on DNA binding specificities of human transcription factors. *Science* **356**, eaaj2239.
Zehir, A., Benayed, R., Shah, R. H., Syed, A., Middha, S., Kim, H. R., Srinivasan, P., Gao, J., Chakravarty, D., Devlin, S. M. et al. (2017). Mutational landscape of metastatic cancer revealed from prospective clinical sequencing of 10,000 patients. *Nat. Med.* **23**, 703-713.
Zhang, X. H.-F., Wang, Q., Gerald, W., Hudis, C. A., Norton, L., Smid, M., Foekens, J. A. and Massagué, J. (2009). Latent bone metastasis in breast cancer tied to Src-dependent survival signals. *Cancer Cell* **16**, 67-78.
Zhang, X. H.-F., Jin, X., Malladi, S., Zou, Y., Wen, Y. H., Brogi, E., Smid, M., Foekens, J. A. and Massagué, J. (2013). Selection of bone metastasis seeds by mesenchymal signals in the primary tumor stroma. *Cell* **154**, 1060-1073.
Zhang, C., Li, J., Huang, T., Duan, S., Dai, D., Jiang, D., Sui, X., Li, D., Chen, Y., Ding, F. et al. (2016). Meta-analysis of DNA methylation biomarkers in hepatocellular carcinoma. *Oncotarget* **7**, 81255-81267.
Zhou, W., Fong, M. Y., Min, Y., Somlo, G., Liu, L., Palomares, M. R., Yu, Y., Chow, A., O'Connor, S. T. F., Chin, A. R. et al. (2014). Cancer-secreted miR-105 destroys vascular endothelial barriers to promote metastasis. *Cancer Cell* **25**, 501-515.
Zomer, A., Maynard, C., Verweij, F. J., Kamermans, A., Schäfer, R., Beerling, E., Schiffelers, R. M., de Wit, E., Berenguer, J., Ellenbroek, S. I. J. et al. (2015). In Vivo imaging reveals extracellular vesicle-mediated phenocopying of metastatic behavior. *Cell* **161**, 1046-1057.