

Nayara Braga Emidio*
Arthur Girardi Carpanez**
Leonardo Ramos Quellis*
Priscila Silva Farani*
Eveline Gomes Vasconcelos*
Priscila Faria-Pinto*

RESUMO

As proteínas desempenham a maior parte das funções fisiológicas das células, constituindo também importantes alvos farmacológicos e biomarcadores de doenças. A pesquisa qualitativa, quantitativa e a elucidação estrutural destas moléculas são fundamentais para a compreensão do funcionamento dos sistemas biológicos, bem como na aplicação destas para o desenvolvimento de novos métodos diagnóstico. O estudo do proteoma nos permite identificar as proteínas que estão sendo expressas em um determinado momento, quantificá-las e observar suas modificações pós-transducionais. Dessa maneira, a análise proteômica fornece informações mais abrangentes e que não podem ser inferidas a partir das informações obtidas através da análise genômica. Este tipo de estudo envolve etapas como: extração e tratamento da amostra, separação das proteínas e/ou peptídeos, espectrometria de massas e análise dos dados usando ferramentas de bioinformática. O presente trabalho faz uma revisão narrativa sobre as principais técnicas aplicadas desde o preparo de amostras até a identificação das proteínas.

Palavras-chave: Proteômica, Espectrometria de Massas, Peptídeos, Proteínas, Biomarcadores.

1 INTRODUÇÃO

Proteoma designa o conjunto de proteínas que estão sendo expressas por uma célula, tecido ou organismo em um determinado momento. De forma distinta, a análise proteômica consiste no estudo do proteoma utilizando técnicas de separação e identificação, tais como eletroforese, cromatografia, espectrometria de massas e bioinformática (HEIN et al., 2013).

As proteínas são moléculas com as mais distintas funções celulares, sendo capazes de orquestrar expressão de genes, catalisar reações metabólicas, além de compor a parte estrutural da célula. Modificações genéticas que promovam a ausência de algumas proteínas ou defeitos em sua estrutura (afetando sua função), podem acarretar doenças ou ser marcadores destas, como nos casos da fenilcetonúria (doença causada pela atividade reduzida ou inexistente da fenilalanina hidroxilase) (ARN, 2014) e a anemia falciforme (causada pela substituição de um resíduo de glutamato por uma valina na posição

seis das cadeias betas de globina, promovendo uma mudança importante na estrutura terciária da proteína e da morfologia da célula) (WEATHERALL, 2013). A importância das proteínas pode ser ressaltada quando observamos que parte significativa dos fármacos disponíveis comercialmente, são direcionados a alvos proteicos, como os medicamentos da família das estatinas que inibem a ação da enzima HMG-CoA redutase, envolvida na síntese do colesterol, e são empregadas no tratamento de dislipidemias; os inibidores da enzima conversora da angiotensina (ECA), utilizados para tratamento da hipertensão e o antibiótico trimetoprim que inibe a ação da enzima dihidrofolato redutase de bactérias (IMMING; SINNING; MEYER, 2006). Assim, o estudo do proteoma, bem como a elucidação da função das proteínas no contexto em que são expressas são de grande relevância para o prognóstico e terapia de doenças (BAKER et al., 2012; HEIN et al., 2013).

As informações provenientes do genoma promoveram um grande progresso no campo das ciências biológicas. Por ser relativamente estável

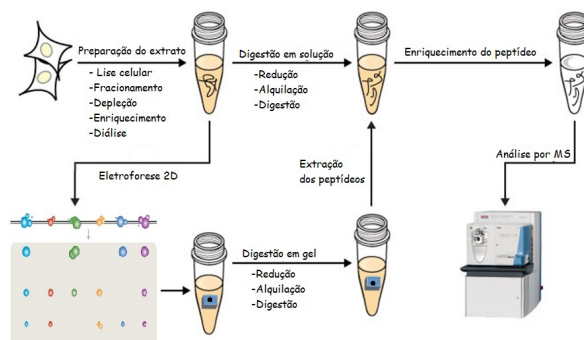
* Universidade Federal de Juiz de Fora, Instituto de Ciências Biológicas, Departamento de Bioquímica, Laboratório de Estrutura e Função de Proteínas - Juiz de Fora, MG.

** Universidade Federal de Juiz de Fora, Instituto de Ciências Exatas, Departamento de Química - Juiz de Fora, MG.

não variar seu código informacional de acordo com o tipo celular ou estímulo hormonal, o genoma não é capaz de revelar informações relativas aos níveis de expressão proteicas e em andamento as modificações pós-transducionais (PTMs) (PAULO et al., 2012; CATHERMAN; SKINNER; KELLEHER, 2014; OVERALL, 2014). A análise do transcriptoma permite enxergar as partes do genoma que estão sendo expressas em um dado momento, porém ainda não há uma correlação entre os níveis de expressão de proteínas e o nível de mRNA, justificada por eventos como degradação, splicing alternativo, secreção, formação de complexos e outros (GHAZALPOUR et al., 2011).

A análise proteômica nos permite a identificação, a quantificação e avaliação do tipo de PTM sofrida pelas proteínas envolvidas em vários processos, permitindo a descoberta de novos alvos terapêuticos e moléculas bioativas. O caráter dinâmico do proteoma oferece como vantagem a identificação de variações ocasionais neste conteúdo sob ação de estímulos hormonais, utilização de drogas e/ou exposição a patógenos (AVEZUM et al., 2004; CARVALHO, 2006).

A análise proteômica apresenta maior número de variáveis em relação a análise genômica e do transcriptoma devido à grande diversidade química das proteínas e interconectividade destas em complexos e redes de sinalização, as quais podem variar muito de acordo com o tempo e espaço (ALTELAAR; MUNOZ; HECK, 2013). O estudo das proteínas requer ferramentas analíticas com alta seletividade, resolução e sensibilidade (POMASTOWSKI; BUSZEWSKI, 2014). Em uma das abordagens, as proteínas do extrato total são pré-fracionadas por eletroforese e então digeridas. Alternativamente, na estratégia shotgun do tipo “bottom-up” digere-se o extrato total (Fluxograma 1) e analisa-se o mesmo por cromatografia líquida diretamente acoplada ao espectrômetro de massas. Dependendo do objetivo do estudo e das técnicas disponíveis, diferentes tipos de abordagens e modificações nas etapas citadas são necessárias (LIU; LIN; YATES, 2002; MALLICK; KUSTER, 2010; YATES; RUSE; NAKORCHEVSKY, 2009; ZHANG et al., 2013).



Fluxograma 1- Etapas comumente seguidas em uma análise proteômica “bottom-up”: preparação do extrato, digestão e análise por espectrometria de massa.

Fonte - Adaptado de: <http://www.piercenet.com/method/sample-preparation-mass-spectrometry>, 2014.

2 DESENVOLVIMENTO

Neste tópico iremos fazer uma revisão das principais etapas envolvidas na análise proteômica, como a separação de proteínas, tratamento da amostra, análise por espectrometria de massa e bioinformática.

2.1 Métodos de separação de proteínas

A separação de amostras complexas é um ponto importante na análise proteômica (RABILLOUD; LELONG, 2011). A eletroforese, especialmente a eletroforese bidimensional (eletroforese 2D), é um método comumente utilizada para análise de proteínas. Esta técnica consiste na aplicação de um campo elétrico para separação destas macromoléculas, primeiramente, de acordo com o ponto isoelétrico e, posteriormente, por volume molecular.

A primeira dimensão da corrida ocorre em um gel com gradiente de pH, em que as proteínas migram até alcançar seu ponto isoelétrico, isto é, até que sua carga líquida seja igual a zero. A segunda dimensão consiste em um gel de poliacrilamida-SDS, no qual as proteínas revestidas pelo detergente SDS migram através da rede entrelaçada formada pela acrilamida, sendo separadas, basicamente, por sua massa molecular (O'FARRELL, 1975). A análise do gel obtido pela eletroforese bidimensional já permite uma análise preliminar do conteúdo proteico majoritariamente expresso na amostra. Com auxílio de softwares, pode-se observar variações na expressão de determinadas proteínas ao comparar a posição e a intensidade de bandas que foram identificadas em géis que correram amostras preparadas pelo mesmo protocolo (RABILLOUD; LELONG, 2011). A eletroforese se configura como uma etapa preliminar a espectrometria de massas, servindo inclusive como pré-fracionamento da amostra. As bandas identificadas no gel podem ser excisadas e após tratamento e extração, analisadas por

um espectrômetro de massas para identificação e/ou caracterização das proteínas.

Apesar de muitos estudos utilizarem a estratégia de separação de proteínas por eletroforese 2D, onde em seguida se faz a excisão de um “spot” para identificar as proteína(s) nele contido, observa-se uma tendência em utilizar a abordagem conhecida como “shotgun proteomics”, a qual se baseia na digestão de todo extrato proteico (sem um pré-fracionamento por gel) e separação dos peptídeos por cromatografia líquida multidimensional (MCDONALD; YATES, 2003; YATES, 2004; YATES; RUSE; NAKORCHEVSKY, 2009).

A separação de proteínas e peptídeos por cromatografia líquida se baseia nos fenômenos de partição e/ou adsorção à fase estacionária; este fenômeno é dependente do tamanho da cadeia polipeptídica do analito. As colunas mais empregadas na separação de peptídeos possuem fase estacionária apolar (octadecilsilano-C18), permitindo assim, que compostos mais apolares fiquem retidos na coluna por um tempo maior (CATHERMAN; SKINNER; KELLEHER, 2014). Na proteômica “shotgun”, devido a alta complexidade da amostra, comumente se faz uma separação multidimensional utilizando métodos ortogonais para promover uma melhor separação, como por exemplo, no método conhecido como MudPIT (Multidimensional Protein Identification Technology) que intercala cromatografia de troca catiônica forte com a de fase reversa (WOLTERS; WASHBURN; YATES, 2001; LIU; LIN; YATES, 2002; DI PALMA et al., 2012).

2.2 Tratamento da amostra

A proteômica utiliza principalmente duas abordagens de análise: “Bottom-Up” e “Top-Down”. A mais utilizada, bottom-up, é composta de processo onde as proteínas são hidrolisadas por peptidases e seguidamente analisadas por cromatografia acoplada ao espectrômetro de massas. A abordagem top-down, por sua vez, não utiliza a digestão prévia das proteínas a serem analisadas. Apesar de parecer uma estratégia menos complexa, esta técnica apresenta barreiras técnicas e um maior custo (CATHERMAN; SKINNER; KELLEHER, 2014; KELLIE et al., 2010).

Ao utilizar a abordagem bottom-up, previamente à digestão das proteínas algumas etapas devem ser cumpridas. A primeira etapa consiste na redução das pontes dissulfeto entre os resíduos de cisteína, sendo o reagente mais utilizado o ditiotreitol (DTT). Em seguida, para impedir uma nova estruturação das pontes dissulfeto, realiza-se a alquilação dos enxofres, geralmente utilizando iodoacetamida (ZHANG et al.,

2013). Após a eliminação da estrutura terciária das proteínas, estas deverão ser digeridas por um processo químico ou enzimático. O método mais empregado é a digestão enzimática por tripsinização, cujos pontos de hidrólise são os resíduos de lisina e arginina pelo lado C-terminal, caso o próximo resíduo de aminoácido não for uma prolina. Quando a sequência da proteína investigada é conhecida, é possível prever os fragmentos gerados pela digestão, sendo o MS-Digest – uma das inúmeras ferramentas oferecidas pelo ProteinProspector (<http://prospector.ucsf.edu/prospector/mshome.htm>) - um dos softwares mais utilizados para este objetivo.

2.3 Espectrometria de massas (mass spectrometry – MS)

A espectrometria de massas é uma técnica que data do início do século passado (THOMPSON, 1913) e se baseia na formação de íons na fase gasosa (carregados positivamente ou negativamente) que podem ser detectados baseado na sua razão massa/carga (m/z) - formalmente, os espectrômetros de massas deveriam ser chamados de espectrômetros de massa/carga (EL-ANEED; COHEN; BANOUB, 2009). A habilidade da espectrometria de massas em analisar proteínas e extratos biológicos advém dos grandes avanços obtidos com o desenvolvimento das técnicas de ionização branda, como a técnica de ionização eletrospray (ESI-MS) e a ionização/dessorção a laser assistida por matriz (MALDI), que são técnicas capazes de transformar macromoléculas em íons (TANAKA et al., 1988; FENN et al., 1989; KARAS). O termo ionização branda significa que uma mínima quantidade de energia é transmitida para o analito durante o processo de ionização. Independentemente da técnica de ionização utilizada, a sensibilidade e a acurácia do espectrômetro de massas está diretamente relacionada com o tipo de analisador de massa empregado na separação dos íons obtidos.

2.3.1 Espectrometria de massas com ionização por “eletrospray” - ESI-MS.

A técnica de ionização por eletrospray foi desenvolvida para o uso em sistemas biológicos por Fenn (FENN et al., 1989). As amostras são dissolvidas em um tampão ou solvente que são bombeadas a um fluxo de microlitros por minuto através de uma agulha hipodérmica que está em uma alta voltagem para dispersar eletrostaticamente, ou eletrospray, gotas de tamanho micrométrico, que são rapidamente evaporadas e transmitem a sua carga para o analito. Duas teorias tentam explicar este fenômeno: “Charged-residue model” e “Ionevaporation model” (GASKELL, 1997). Este processo de ionização ocorre em condições ambientes

e portanto é muito suave, ou seja, não ocorre a fragmentação dos íons do analito na fase gasosa (MANN; HENDRICKSON; PANDEY, 2001).

Uma ampla gama de compostos pode ser analisada por ESI-MS; o único requisito é que esta molécula possa ser solubilizada em um solvente polar para permitir a incorporação da carga. A intensidade do sinal em um espectro obtido pelo eletrospray é proporcional a concentração do analito até a saturação do sinal. Íons grandes são tipicamente e possuem mais de uma carga (proteínas e peptídeos através da adição de prótons no modo positivo e da abstração de prótons no modo negativo), o que leva a formação dos típicos envelopes isotópicos multicarregados, uma vez que o espectrômetro de massas responde a medida da massa/carga (HO et al., 2003).

2.3.2 Ionização/dessorção a laser assistida por matriz - MALDI

A técnica de ionização/dessorção a laser assistida por matriz foi desenvolvida por KaraseHellenkamp no final da década de 80, mas o grande avanço para a ionização de biomoléculas foi introduzido por Tanaka e colaboradores no mesmo período ao publicarem um trabalho mostrando a ionização de proteínas com até 35KDa (KARAS; HILLENKAMP, 1988; TANAKA et al., 1988).

A formação dos íons em fase gasosa é obtida através da co-cristalização de um excesso de matriz com o analito em uma placa de metal. A matriz, em geral pequenas moléculas orgânicas, é capaz de absorver o comprimento de onda emitido pelo laser. O laser atinge os cristais de matriz e analito formados na placa, levando a absorção dessa energia pela matriz e a subsequente dessorção e ionização dos analitos presentes na amostra. Durante este processo de ionização íons, em sua grande maioria

sendomonocarregados, são comumente formados, todavia o processo de formação destes íons ainda não está completamente compreendido, as teorias que tentam explicar tal fenômeno são: “Charge transfer” e “Lucky survivor”(EDMOND DE HOFFMANN, 2001; GLISH; VACHET, 2003). As matrizes diferem na quantidade de energia que elas conseguem transmitir para as biomoléculas durante o processo de dessorção e ionização e, conseqüentemente o grau de fragmentação que elas podem causar (COHEN; CHAIT 1996).

A intensidade do sinal nos espectros de massas da técnica MALDI dependem de alguns fatores como: (I) incorporação dos analitos no cristal, (II) probabilidade de capturar e/ou reter prótons durante o processo de dessorção, incluindo efeitos de supressão da ionização, o qual pode ser provocado pela presença de SDS. Por essas razões é difícil correlacionar a intensidade do pico no cromatograma com a quantidade de analito presente na amostra (KNOCHENMUSS; ZENOBI, 2003).

2.3.3 Comparação entre as técnicas MALDI e ESI-MS

Ambas as técnicas, MALDI e ESI-MS, são eficientes para proteômica e podem ser aplicadas a analitos em concentrações mais baixas que picomols. Uma das maiores diferenças entre as técnicas é o estado em que os analitos são introduzidos no ionizador. Apesar da técnica de eletrospray ser capaz de reproduzir melhor os dados que a técnica de MALDI, deve-se atentar para o fato de que a abundância relativa dos íons gerados no ESI e MALDI não é uma representação real da concentração da amostra. Conseqüentemente, um padrão interno deve ser utilizado, preferencialmente um análogo isotópico do seu analito, para propósitos de quantificação. A

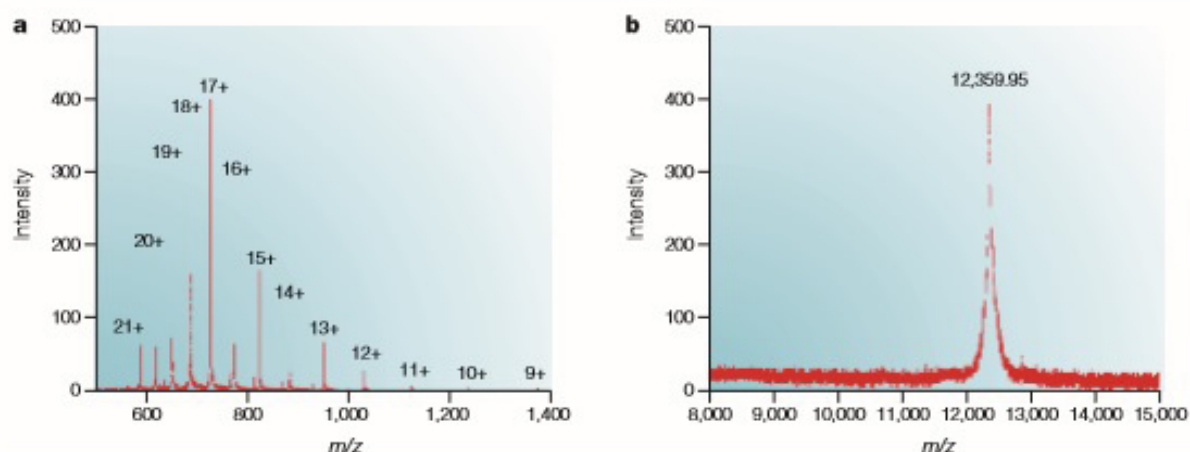


Ilustração 1- Comparação do espectro de massas do citocromo c utilizando a técnica de eletrospray e MALDI: podemos observar a produção de íons multicarregados utilizando a técnica ESI e monocarregado utilizando MALDI

Fonte- Adaptado de GLISH; VACHET, 2003.

heterogeneidade do cristal formado na técnica de MALDI impossibilita a utilização da mesma para propósitos de quantificação. Na verdade, a qualidade do espectro de massas é influenciada pela posição do laser e o operador deve identificar um “sweet spot” – um ponto ideal – na amostra cristalizada que possa gerar o espectro com maior número de informações (EL-ANEED; COHEN; BANOUB, 2009).

A técnica de ESI-MS tende a produzir espécies multicarregadas para proteínas e peptídeos, já a de MALDI produz íons monocarregados como exemplificados na Ilustração 1 (GLISH; VACHET, 2003; HAN; ASLANIAN; YATES, 2008).

2.4 Analisadores de massa

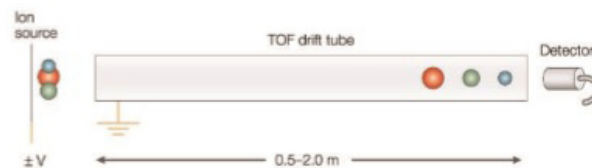
O analisador de massas é a parte do instrumento que realiza a separação dos íons gerados através da sua razão massa/carga. Assim como os processos de ionização, tem-se um enorme leque de opções que podem realizar esta separação (EL-ANEED; COHEN; BANOUB, 2009). Três diferentes princípios podem ser aplicados para alcançar a separação das massas: separação baseada no tempo de voo (TOF MS), separação através de campos elétricos gerados por hastes metálicas (quadrupolo MS), ou a separação pela ejeção seletiva de íons de um campo de aprisionamento de íons tridimensional (aprisionamento de íons ou transformada de Fourier ciclotron íon MS). Para análises estruturais, como o sequenciamento de peptídeos, duas etapas de MS podem ser realizadas em sequência (espectrometria de massas em sequência ou MS/MS), empregando o mesmo princípio de separação duas vezes ou através da combinação de dois princípios de separação (MCLAFFERTY, 1981).

2.4.2 Tempo de voo, TOF.

O analisador de tempo de voo separa os íons baseando-se na velocidade dentro do tubo de voo. Teoricamente, os íons são formados no mesmo lugar ao mesmo tempo na fonte de íons e então acelerados a um potencial fixo para dentro do “TOF drift tube” – tubo de desvio TOF (Esquema 1). Uma vez que todos os íons formados com a mesma carga apresentam a mesma energia potencial elétrica quando expostos ao campo, pode-se inferir a razão massa carga dos mesmos, pois esta energia será convertida em energia cinética que é uma função, entre outras coisas, da massa da molécula; consequentemente, íons com menor valor de m/z alcançarão uma maior velocidade que os íons de maior m/z .

Após os íons serem acelerados, eles viajam por uma distância fixa, livres de campo elétrico, geralmente de 0,5 a 2,0 metros, antes de chegarem ao detector.

Portanto, através da medida do tempo gasto pelos íons para chegar no detector após a sua formação na fonte de íons, pode-se determinar a razão m/z do mesmo (GLISH; VACHET, 2003).



Esquema 1 - Analisador de massas do tipo TOF (tempo de voo)
Fonte- Adaptado de GLISH; VACHET, 2003.

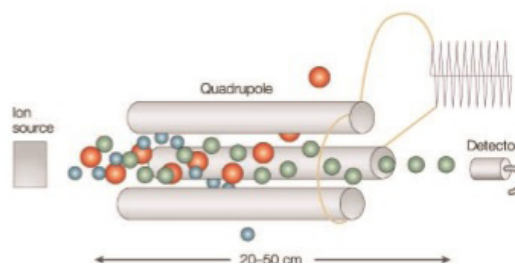
2.4.2 Quadrupolo

A separação dos íons em um instrumento do tipo quadrupolo é o resultado da mobilidade do íon em um campo elétrico dinâmico (rádio frequência), e está diretamente relacionada com a m/z dos íons em questão. Desta maneira, pode-se pensar o detector do tipo quadrupolo como um filtro de massa, que consiste de quatro hastes em que um campo elétrico oscilante é aplicado e apenas certos valores de m/z conseguem alcançar o detector.

Os íons formados na câmara de ionização são “guiados” para o quadrupolo através de um campo elétrico. Um íon carregado positivamente, por exemplo, mover-se para a direção da haste carregada negativamente. Entretanto, quando a polaridade é trocada, o íon mudará sua trajetória antes de colidir com a haste. Deste modo, os íons passarão por oscilações complexas (trajetórias) e com valores apropriados de potencial (V), frequência (ω), apenas uma estreita faixa de íons chegará ao detector (Esquema 2). Os íons que possuem trajetórias instáveis colidirão com as hastes e não serão detectados. Através da variação do campo elétrico e da detecção dos íons formados, tem-se o espectro de massas.

2.4.3 Instrumentos de aprisionamento de íons (Ion - Trapping).

O analisador de massas quadrupolo com aprisionamento de íons está muito relacionado com o



Esquema 2- Analisador de massas do tipo quadrupolo.
Fonte- Adaptado de GLISH; VACHET, 2003.

analisador do tipo quadrupolo. Enquanto o quadrupolo apresenta campos elétricos em duas dimensões (eixos x e y) e os íons movem-se perpendiculares ao campo, isto é no eixo z , o analisador do tipo aprisionamento de íons possui campos elétricos em todas as três dimensões, o que possibilita o aprisionamento dos íons no campo (Esquema 3) (GLISH; VACHET, 2003). Estas armadilhas de íons capturam o feixe contínuo de íons até atingirem o número máximo de íons que podem ser introduzidos no instrumento sem que haja a distorção do campo aplicado. Estes íons são então submetidos a radiofrequências, que ejetam um íon com m/z próximos por vez e estes são então detectados, gerando assim o espectro de massas. Para um experimento de massas em tandem (i.e., MS/MS), todos os outros íons, menos o de interesse, são ejetados. Então o íon de interesse é fragmentado e os seus íons produto são analisados. Vários experimentos de massas em sequência podem ser realizados, em princípio, permitindo assim um detalhado estudo de fragmentação. Diferentemente dos equipamentos de quadrupolo, os analisadores de aprisionamento de íons não apresentam uma diminuição da sensibilidade para valores elevados de m/z (GLISH; VACHET, 2003).

Uma segunda versão do aprisionamento de íons está baseada no princípio da transformada de Fourier ciclotron íon MS. Nesta técnica, os íons são capturados através de uma combinação de campos elétricos e um forte campo magnético (GLISH; VACHET, 2003).

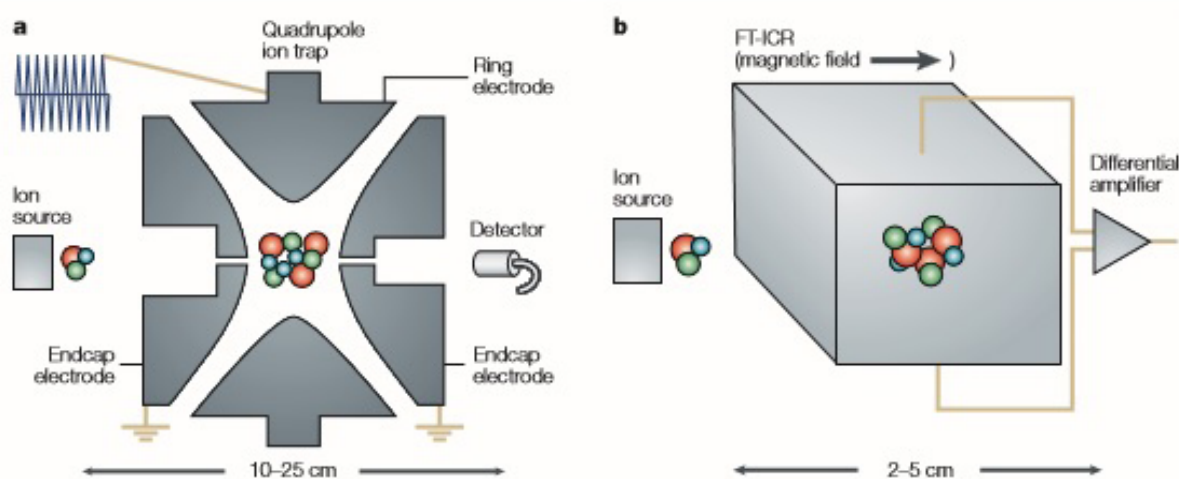
O analisador de massas Orbitrap é considerado um dos mais novos analisadores em espectrometria de massas (MAKAROV, 2000), entretanto suas raízes datam de 1923, quando Kingdon propôs o princípio do “orbital trapping” (KINGDON, 1923). Ele é

composto por um eletrodo central fusiforme e um eletrodo externo em forma de barril. Os íons são injetados no volume entre os dois eletrodos. Com uma voltagem aplicada entre o eletrodo central e o eletrodo mais externo, um campo elétrico radial distorce a trajetória dos íons em direção ao eletrodo central, enquanto a velocidade tangencial cria uma força centrípeta (Hu et al., 2005; PERRY et al., 2008). Ajustando os parâmetros do equipamento, os íons ficarão em um espiral quase circular dentro da armadilha. O campo elétrico axial causado pela forma cônica dos eletrodos força os íons para a parte mais larga da armadilha iniciando assim as oscilações harmônicas axiais. Eletrodos externos são então utilizados para detectar a corrente destas oscilações axiais. A corrente obtida no domínio do tempo, é transformada, através da transformada de Fourier, para o domínio da frequência e então convertida no espectro de massas (ZUBAREV; MAKAROV, 2013).

2.5 Espectrometria de massas sequencial - tandem MS/MS.

Como o próprio nome sugere, as técnicas de MS/MS envolvem dois estágios de MS. De forma que, no primeiro estágio, os íons de interesse são isolados dos outros íons gerados na fonte de íons. Este íon isolado – nomeado íon precursor – pode sofrer então subsequentes dissociações gerando assim fragmentos – denominados íons produto (Esquema 4).

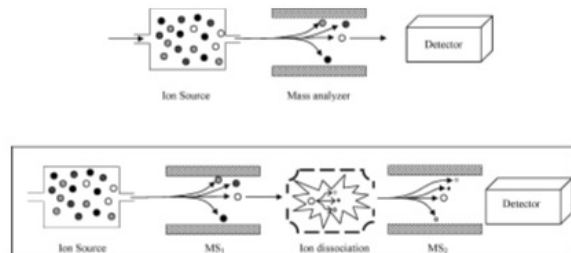
O modo de alcançar este objetivo é através do uso de analisadores de massas conectados em séries (tandem); de modo que o isolamento do íon precursor possa ser realizado no primeiro analisador de massas, seguido então da sua fragmentação, gerando os íons produtos que são então separados com base na



Esquema 3 - Analisadores de massas do tipo aprisionamento de íons. A- Quadrupolo com aprisionamento de íons. B- Transformada de Fourier ciclotron íon MS.

Fonte- Adaptado de GLISH; VACHET, 2003.

sua massa/carga (GLISH; VACHET, 2003; HAN; ASLANIAN; YATES, 2008).



Esquema 4- Comparação entre um experimento de massas convencional (em cima) com o experimento em massas sequencial (abaixo).

Fonte- Adaptado de EL-ANEED; COHEN; BANOUB (2009)

Os tipos de instrumento ilustrado acima são considerados sequenciais no espaço, pois as análises são realizadas em diferentes analisadores em diferentes “espaços”. Espectrometria de massas sequenciais em tempo no entanto são realizadas com instrumentos do tipo aprisionamento de íons, onde todos os íons são ejetados exceto o de interesse que sofre então a fragmentação “no mesmo espaço”. Estes tipos de equipamentos são capazes de realizar múltiplos experimentos de massas, (MS)_n, sendo assim uma poderosa ferramenta para análise estrutural (GLISH; VACHET, 2003; HAN; ASLANIAN; YATES, 2008).

2.6 Bioinformática aplicada a proteômica

A partir do espectro gerado pela fragmentação de um precursor, é possível reconstruir o peptídeo fragmentado (sequenciamento de novo), entretanto na prática, é inviável analisar um espectro de cada vez, sendo esta tarefa abreviada pela aplicação de algoritmos (HEIN et al., 2013). Grande parte dos softwares existentes realizam comparações entre o conjunto de dados gerados por peptidomass “fingerprint” (PMF) e MS/MS com os espectros obtidos da fragmentação teórica de proteínas presentes nos bancos de dados (ZHANG et al., 2013). Após esta comparação, o software retorna as prováveis proteínas contidas na amostra com um grau de confiança que varia de acordo com a cobertura alcançada (o número de peptídeos identificados na sequência), dentre outros parâmetros. Entre os softwares utilizados para identificação baseando-se em bancos de dados, podemos citar: MASCOT (<http://www.matrixscience.com/>), ProLuCID/SEQUEST (<http://fields.scripps.edu/proluclid/>), X!tandem (<http://www.thegpm.org/tandem/>), Comet (ENG; JAHAN; HOOPMANN, 2013; <http://cometms.sourceforge.net/>), MSFIT (<http://prospector.ucsf.edu/prospector/cgibin/msform.cgi?form=msfitstandard>). Atualmente existem

pipelines para a análise dos dados proteômicos, entre estes podemos citar o PatternLab for proteomics que contém uma variedade de ferramentas que permitem, por exemplo, assinalar proteínas ou peptídeos que são diferentemente expressos e interpretar os dados de acordo com o “Gene Ontology”. Outros exemplos são o Transproteomic pipeline e o OpenMS (KOHLBACHER et al., 2007; CARVALHO; YATES III; BARBOSA, 2010; DEUTSCH et al., 2010).

No caso de organismos que não tenham o genoma sequenciado e, portanto, poucas proteínas constam nos bancos de dados, a abordagem de sequenciamento de novo é essencial. Exemplos de algoritmos que realizam sequenciamento de novo são UniNovo, PepNovo, pNovo+ePEAKS (MA et al., 2003; FRANK; PEVZNER, 2005; CHI et al., 2013; JEONG; KIM; PEVZNER, 2013). Neste caso, também existem softwares que auxiliam na interpretação dos resultados dos algoritmos de sequenciamento de novo realizando alinhamentos contra bancos de dados contendo sequências de organismos homólogos, como o PepExplorer e o MSBlast (SHEVCHENKO et al., 2001; LEPREVOST et al., 2014)

3 DISCUSSÃO

A espectrometria de massas permite o trabalho de identificação qualitativa e quantitativa de proteínas e fragmentos protéicos (peptídeos). Podendo ser uma técnica aplicada a identificação de outras macromoléculas tais como glicoconjugados, lipídeos, DNA e RNA. A massa molecular obtida através desta técnica possui grande acurácia, especificidade e reprodutibilidade. Tais características permitem a aplicação de softwares capazes de comparar os dados dos espectros em estudo com aqueles presentes nos bancos de dados. Esta ferramenta pode ser empregada em amostras diversas com distintas finalidades, sendo as mais utilizadas: identificação e quantificação de proteínas e peptídeos, análise de modificações pós-traducionais, mapeamento de epítomos antigênicos, análise de biomarcadores.

A análise proteômica tem se tornado uma ferramenta valiosa na busca de novas moléculas de aplicação clínica. Tais biomarcadores podem ter diversas aplicações no prognóstico, diagnóstico, monitoramento e direcionamento de tratamentos. Em muitas doenças, como no câncer, o diagnóstico precoce e exato da doença é essencial para a escolha de um tratamento efetivo. Entretanto na maioria das vezes, isto não é possível devido à ausência de biomarcadores já descritos (OVERALL, 2014) e a

baixa sensibilidade dos métodos utilizados (CHO, 2014).

A avaliação dos perfis proteicos busca então identificar comparativamente ou quantitativamente em amostras clínicas, biomarcadores que tenham níveis alterados de expressão, modificações pós-traducionais ou formas variantes das proteínas. A identificação de um conjunto de marcadores tem um maior potencial de alcançar maior especificidade e sensibilidade (LOPEZ et al., 2012). O “SelectReactionMonitoring” (SRM) é uma abordagem que permite a análise de várias proteínas simultaneamente com alta especificidade e está crescendo como uma forma de interligar a descoberta de biomarcadores e sua validação para utilização na rotina clínica (HÜTTENHAIN et al., 2012; CERCIELLO et al., 2013; CHO, 2014).

Uma abordagem rápida e simples da espectrometria de massas na microbiologia clínica consiste na aplicação de colônias isoladas em placas de MALDI, seguida pela matriz. Os espectros gerados são comparados com as bases de dados contendo microrganismos de relevância clínica. Além da identidade, a análise dos picos gerados pelo MALDI-TOF/MS pode fornecer informações como susceptibilidade a drogas (PATEL, 2013).

Atualmente, um dos grandes desafios encontra-se em correlacionar descobertas científicas que sejam aplicáveis em diagnósticos clínicos. Um entrave para este desenvolvimento é a necessidade de se fazer os

testes com uma população amostral maior e validação destes dados (SANCHEZ, 2013).

4 CONCLUSÃO

A tecnologia para sequenciamento genético alcançou alto nível de maturidade, permitindo a obtenção de genomas inteiros em pouco tempo (PAREEK; SMOCZYNSKI; TRETYN, 2011). Entretanto, apenas as informações do genoma são insuficientes para entender a relação entre fenótipo e genético, visto que aproximadamente 100 000 proteínas são expressas por cerca de 20235 genes em humanos, mostrando a complexidade do estudo do proteoma (BENSIMON; HECK; AEBERSOLD, 2012; ZHANG et al., 2013).

A identificação e análise de proteínas utilizando métodos baseados em espectrometria de massas tem se tornado uma abordagem interessante no campo da biologia, sendo um complemento as técnicas tradicionalmente utilizadas em pesquisas bioquímicas e de biologia molecular (BENSIMON; HECK; AEBERSOLD, 2012; ZHANG et al., 2013).

Além disso, proteômica também vem sendo utilizada para a busca de novos biomarcadores de doenças, entretanto sua utilização no diagnóstico clínico ainda é limitada, visto a dificuldade de se validar estes métodos em uma população amostral significativa (FUZERY et al., 2013; ZHANG et al., 2013).

INTRODUCTION TO METHODS AND APPLICATIONS IN PROTEOMICS

ABSTRACT

Proteins play most of cell physiological functions, being important pharmacological targets and disease related biomarkers. The investigation (quantitative, qualitative and structural) of those molecules is essential to understand how the biological systems work and has potential application in clinical diagnosis. The proteomics study allow us to identify proteins being expressed at a certain time, measure them and detect post translational modification in their structure. Hence, the proteome provides information that might not be inferred from genome. Proteomics comprises the following steps: sample extraction and treatment, protein and/or peptide separation, mass spectrometry and analysis using bioinformatics tools. This review aims to describe the main techniques applied in proteomics from the sample preparation to protein identification.

Key-words: Proteomics, Mass Spectrometry, Peptides, Proteins, Biomarkers.

REFERÊNCIAS

ALTELAAR, A. F.; MUNOZ, J.; HECK, A. J. Next-generation proteomics: towards an integrative view of proteome dynamics. *Nature Reviews Genetics*, London, v. 14, no. 1, p. 35-48, Jan 2013.

ARN, P. H. Phenylketonuria (PKU). In: Aminoff, M. J. e Daroff, R. B. *Encyclopedia of the Neurological Sciences* (Second Edition). Oxford: Academic Press, 2014, p.887-889.

AVEZUM, Á. et al. III Diretriz sobre tratamento do infarto agudo do miocárdio. *Arquivos Brasileiros de Cardiologia*, Rio de Janeiro, v. 83, n.4, p. 1-86, Set 2004.

- BAKER, E. S. et al. Mass spectrometry for translational proteomics: progress and clinical implications. *Genome Medicine*, London, v. 4, no. 8, p. 63, Sep 2012.
- BENSIMON, A.; HECK, A. J.; AEBERSOLD, R. Mass spectrometry-based proteomics and network biology. *Annual Review of Biochemistry*, Palo Alto, v. 81, no. 1, p. 379-405, Jul 2012.
- CARVALHO, P. C.; FISHER, J. S. G.; WIM M. DEGRAVE, W. M.; CARVALHO, M. G. C.. Marcadores séricos e espectrometria de massa no diagnóstico do cancer. *Jornal Brasileiro de Patologia e Medicina Laboratorial*, Rio de Janeiro, v. 42, n. 6, p. 6, Dez 2006.
- CARVALHO, P. C.; YATES III, J. R.; BARBOSA, V. C. Analyzing Shotgun Proteomic Data with PatternLab for Proteomics. In: (Ed.). *Current Protocols in Bioinformatics*: John Wiley & Sons, Inc., 2010.
- CATHERMAN, A. D.; SKINNER, O. S.; KELLEHER, N. L. Top Down proteomics: Facts and perspectives. *Biochemical and Biophysical Research Communications*, New York, v. 445, no. 4, p. 683-693, Mar 2014.
- CERCIELLO, F.; CHOI, M.; NICASTRI, A.; BAUSCH-FLUCK, D.; ZIEGLER, A.; VITEK, O.; FELLE-BOSCO, E.; STAHEL, R.; AEBERSOLD, R.; WOLLSCHIED, B. Identification of a seven glycopeptide signature for malignant pleural mesothelioma in human serum by selected reaction monitoring. *Clinical Proteomics*, London, v. 10, no. 1, p. 1-12, Nov 2013.
- CHEVALIER, F. Highlights on the capacities of "Gel-based" proteomics. *Proteome Science*, London, v. 8, no. 23, Apr 2010.
- COHEN, S. L.; CHAIT, B. T. Influence of matrix solution conditions on the MALDI-MS analysis of peptides and proteins. *Analytical Chemistry*, Washington, v. 68, no. 1, p. 31-37, Jan 1996.
- CHI, H.; CHEN, H.; HE, K.; WU, L.; YANG, B.; SUN, R. X.; LIU, J.; ZENG, W. F.; SONG, C. Q.; HE, S. M.; DONG, M. Q. pNovo+: de novo peptide sequencing using complementary HCD and ETD tandem mass spectra. *Journal of proteome research*, Washington, v. 12, no. 2, p. 615-625, Dec 2013.
- CHO, W. C. Proteomics in translational cancer research: biomarker discovery for clinical applications. *Expert Review of Proteomics*, London, v. 11, no. 2, p. 131-133, Apr 2014.
- DEUTSCH, E. W.; MENDOZA, L.; SHTEYNBERG, D.; FARRAH, T.; LAM, H.; TASMAN, N.; SUN, Z.; NILSSON, E.; PRATT, B.; PRAZEN, B.; ENG, J. K.; MARTIN, D. B.; NESVIZHSKII, A. I.; AEBERSOLD, R. A guided tour of the Trans-Proteomic Pipeline. *Proteomics*, London, v. 10, no. 6, p. 1150-1159, Mar 2010.
- DI PALMA, S. et al. Recent advances in peptide separation by multidimensional liquid chromatography for proteome analysis. *Journal of Proteomics*, London, v. 75, no. 13, p. 3791-3813, Jul 2012.
- EDMOND DE HOFFMANN, V. S. **Mass Spectrometry: Principles and Applications**: Wiley, 2001
- EL-ANEED, A.; COHEN, A.; BANOUB, J. Mass Spectrometry, Review of the Basics: Electrospray, MALDI, and Commonly Used Mass Analyzers. *Applied Spectroscopy Reviews*, v. 44, no. 3, p. 210-230, Oct 2009.
- ENG, J. K.; JAHAN, T. A.; HOOPMANN, M. R. Comet: an open-source MS/MS sequence database search tool. *Proteomics*, Weinheim, v. 13, no. 1, p. 22-24, Jan 2013.
- FRANK, A.; PEVZNER, P. PepNovo: de novo peptide sequencing via probabilistic network modeling. *Analytical Chemistry*, Washington, v. 77, no. 4, p. 964-973, Jan 2005.
- FENN, J. B.; MANN, M.; MENG, C. K.; WONG, S. F.; WHITEHOUSE, C. M. Electrospray ionization for mass spectrometry of large biomolecules. *Science*, New York, v. 246, no. 4926, p. 64-71, Oct 1989.
- FUZERY, A. K.; LEVIN, J.; CHAN, M. M.; CHAN, D. W. Translation of proteomic biomarkers into FDA approved cancer diagnostics: issues and challenges. *Clinical Proteomics*, London, v. 10, no. 1, p. 13, Oct 2013.
- GASKELL, S. J. Electrospray: Principles and Practice. *Journal of Mass Spectrometry*, Chichester, v. 32, no. 7, p. 677-688, Jul 1997.
- GHAZALPOUR, A. et al. Comparative analysis of proteome and transcriptome variation in mouse. *PLoS Genetics*, San Francisco, v. 7, no. 6, Jun 2011.
- GLISH, G. L.; VACHET, R. W. The basics of mass spectrometry in the twenty-first century. *Nature Reviews Drug Discovery*, London, v. 2, no. 2, p. 140-150, Feb 2003.
- HAN, X.; ASLANIAN, A.; YATES, J. R., 3RD. Mass spectrometry for proteomics. *Current Opinion in Chemical Biology*, London, v. 12, no. 5, p. 483-490, Aug 2008.
- HEIN, M. Y. et al. Chapter 1 - Proteomic Analysis of Cellular Systems. In: Walhout, A. J. M., Vidal, M., et al. *Handbook of Systems Biology*. San Diego: Academic Press, 2013, p.3-25.
- HO, C. S. et al. Electrospray Ionisation Mass Spectrometry: Principles and Clinical Applications. *The Clinical Biochemist Reviews*, Chippendale, v. 24, no. 1, p. 3-12, Feb 2003.
- HU, Q. et al. The Orbitrap: a new mass spectrometer. *Journal of mass spectrometry*, Chichester, v. 40, no. 4, p. 430-43, Apr 2005.

- HUGHES, C.; MA, B.; LAJOIE, G. A. De novo sequencing methods in proteomics. *Methods in molecular biology*, Totowa, v. 604, no., p. 105-121, 2010.
- HÜTTENHAIN, R.; SOSTE, M.; SELEVSEK, N.; RÖST, H.; SETHI, A.; CARAPITO, C.; FARRAH, T.; DEUTSCH, E. W.; KUSEBAUCH, U.; MORITZ, R. L.; NIMÉUS-MALMSTRÖM, E.; RINNER, O.; AEBERSOLD, R. Reproducible quantification of cancer-associated proteins in body fluids using targeted proteomics. *Science translational medicine*, Washington, v. 4, no. 142, p. 142ra194, Sep 2012.
- IMMING, P.; SINNING, C.; MEYER, A. Drugs, their targets and the nature and number of drug targets. *Nature Reviews Drug Discovery*, London, v. 5, no. 10, p. 821-834, Oct 2006.
- JEONG, K.; KIM, S.; PEVZNER, P. A. UniNovo: a universal tool for de novo peptide sequencing. *Bioinformatics*, Oxford, v. 29, n. 16, p. 1953-1962, Aug 2013.
- KARAS, M.; HILLENKAMP, F. Laser desorption ionization of proteins with molecular masses exceeding 10,000 daltons. *Analytical Chemistry*, Washington, v. 60, no. 20, p. 2299-2301, Oct 1988.
- KELLIE, J. F.; TRAN, J. C.; LEE, J. E.; AHLF, D. R.; THOMAS, H. M.; NTAI, I.; CATHERMAN, A. D.; DURBIN, K. R.; ZAMDBORG, L.; VELLAICHAMY, A.; THOMAS, P. M.; KELLEHER, N. L. The emerging process of Top Down mass spectrometry for protein analysis: Biomarkers, protein-therapeutics, and achieving high throughput. *Molecular BioSystems*, Cambridge, v. 6, no. 9, p. 1532-1539, Sep 2010.
- KINGDON, K. H. A Method for the Neutralization of Electron Space Charge by Positive Ionization at Very Low Gas Pressures. *Physical Review*, Schenectady v. 21, nO. 4, p. 408-418, Jan 1923.
- KNOCHENMUSS, R.; ZENOBI, R. MALDI ionization: the role of in-plume processes. *Chemical Reviews*, Washington, v. 103, no. 2, p. 441-452, Feb 2003.
- KOHLBACHER, O.; REINERT, K.; GROPL, C.; LANGE, E.; PFEIFER, N.; SCHULZ-TRIEGLAFF, O.; STURM, M. TOPP- the OpenMS proteomics pipeline. *Bioinformatics*, Oxford, v. 23, no. 2, p. e191-197, Jan 2007.
- LIU, H.; LIN, D.; YATES, J. R., 3RD. Multidimensional separations for protein/peptide analysis in the post-genomic era. *Biotechniques*, London, v. 32, no. 4, p. 898-902, Apr 2002.
- LEPREVOST, F. V.; VALENTE, R. H.; LIMA, D. B.; PERALES, J.; MELANI, R.; YATES, J. R., 3RD; BARBOSA, V. C.; JUNQUEIRA, M.; CARVALHO, P. C. PepExplorer: a similarity-driven tool for analyzing de novo sequencing results. *Molecular & cellular proteomics : MCP*, Bethesda v. 13, no. 9, p. 2480-2489, Sep 2014.
- LOPEZ, E.; MADERO, L.; LOPEZ-PASCUAL, J.; LATTERICH, M. Clinical proteomics and OMICS clues useful in translational medicine research. *Proteome Science*, London, v. 10, no. 1, p. 35, May 2012.
- MA, B.; ZHANG, K.; HENDRIE, C.; LIANG, C.; LI, M.; DOHERTY-KIRBY, A.; LAJOIE, G. PEAKS: powerful software for peptide de novo sequencing by tandem mass spectrometry. *Rapid communications in mass spectrometry : RCM*, Chichester, v. 17, no. 20, p. 2337-2342, Oct 2003.
- MCDONALD, W. H.; YATES, J. R., 3RD. Shotgun proteomics: integrating technologies to answer biological questions. *Current opinion in molecular therapeutics*, London, v. 5, no. 3, p. 302-309, Jun 2003.
- MCLAFFERTY, F. "Tandem mass spectrometry." *Science*, New York, v. 214, no. 4518, p. 280-287, 1981.
- MAKAROV, A. Electrostatic Axially Harmonic Orbital Trapping: A High-Performance Technique of Mass Analysis. *Analytical Chemistry*, Manchester v. 72, no. 6, p. 1156-1162, Mar 2000.
- MALLICK, P.; KUSTER, B. Proteomics: a pragmatic perspective. *Nature Biotechnology*, New York, v. 28, no. 7, p. 695-709, Jul 2010.
- MANN, M.; HENDRICKSON, R. C.; PANDEY, A. Analysis of proteins and proteomes by mass spectrometry. *Annual Review of Biochemistry*, Palo Alto, v. 70, no. 1, p. 437-473, Jul 2001.
- O'FARRELL, P. H. High Resolution Two-Dimensional Electrophoresis of Proteins. *The Journal of biological chemistry*, Baltimore, v. 250, no. 10, p. 4007-4021, May 1975.
- OVERALL, C. M. Can proteomics fill the gap between genomics and phenotypes? *Journal of Proteomics*, Amsterdam, v. 100, p. 1-2, Apr 2014.
- PAREEK, C. S.; SMOCZYNSKI, R.; TRETYN, A. Sequencing technologies and genome sequencing. *Journal of Applied Genetics*, Poznań, v. 52, no. 4, p. 413-435, Nov 2011.
- PATEL, R. MALDI-TOF mass spectrometry: transformative proteomics for clinical microbiology. *Clinical Chemistry*, Washington, v. 59, no. 2, p. 340-342, Feb 2013.
- CARVALHO, P. C.; FISHER, J. S. G.; WIM M. DEGRAVE, W. M.; CARVALHO, M. G. C.. Marcadores séricos e espectrometria de massa no diagnóstico do cancer. *Jornal Brasileiro de Patologia e Medicina Laboratorial*, Rio de Janeiro, v. 42, n. 6, p. 6, Dez 2006.
- PAULO, J. A.; KADIYALA, V.; BANKS, P. A.; STEEN, H.; CONWELL, D. L. Mass spectrometry-based proteomics for translational research: a technical overview. *Yale Journal of Biology and Medicine*, New Haven, v. 85, no. 1, p. 59-73, Mar 2012.

- PERRY, R. H.; COOKS, R. G.; NOLL, R. J. Orbitrap mass spectrometry: instrumentation, ion motion and applications. *Mass spectrometry reviews*, New York v. 27, no. 6, p. 661-99, Nov-Dec 2008.
- POMASTOWSKI, P.; BUSZEWSKI, B. Two-dimensional gel electrophoresis in the light of new developments. *Trends in Analytical Chemistry*, Amsterdam, v. 53, p. 167-177, Jan 2014.
- RABILLOUD, T.; LELONG, C. Two-dimensional gel electrophoresis in proteomics: A tutorial. *Journal of Proteomics*, Amsterdam, v. 74, no. 10, p. 1829-1841, Jun 2011.
- SANCHEZ, J.-C. The art of proteomics translation. *Translational Proteomics*, Amsterdam, v. 1, no. 1, p. 1-2, 2013.
- SHEVCHENKO, A.; SUNYAEV, S.; LOBODA, A.; SHEVCHENKO, A.; BORK, P.; ENS, W.; STANDING, K. G. Charting the proteomes of organisms with unsequenced genomes by MALDI-quadrupole time-of-flight mass spectrometry and BLAST homology searching. *Analytical chemistry*, Washington, v. 73, no. 9, p. 1917-1926, May 2001.
- TANAKA, K.; WAKI, H.; IDO, Y.; AKITA, S.; YOSHIDA, Y.; YOSHIDA, T.; MATSUO, T. Protein and polymer analyses up to m/z 100 000 by laser ionization time-of-flight mass spectrometry. *Rapid Communications in Mass Spectrometry*, London, v. 2, no. 8, p. 151-153, Aug 1988.
- THOMPSON, J.J.; Rays of Positive Electricity and their Application to Chemical Analysis. London: Longmans, Green and Co. Ltd., 1913.
- WEATHERALL, D. J. Sickle Cell Anemia. In: Hughes, S. M. *Brenner's Encyclopedia of Genetics (Second Edition)*. San Diego: Academic Press, 2013, p.429-431.
- WOLTERS, D. A.; WASHBURN, M. P.; YATES, J. R., 3RD. An automated multidimensional protein identification technology for shotgun proteomics. *Analytical chemistry*, Washington, v. 73, no. 23, p. 5683-5690, Dec 2001.
- YATES, J. R.; RUSE, C. I.; NAKORCHEVSKY, A. Proteomics by Mass Spectrometry: Approaches, Advances, and Applications. *Annual Review of Biomedical Engineering*, Palo Alto, v. 11, no. 1, p. 49-79, Apr 2009.
- YATES, J. R., 3RD. Mass spectral analysis in proteomics. *Annual review of biophysics and biomolecular structure*, Palo Alto, v. 33, no., p. 297-316, Jun 2004.
- ZHANG, Y.; FONSLow, B. R.; SHAN, B.; BAEK, M. C.; YATES, J. R., 3RD. Protein analysis by shotgun/bottom-up proteomics. *Chemical Reviews*, Washington, v. 113, no. 4, p. 2343-2394, Apr 2013.
- ZUBAREV, R. A.; MAKAROV, A. Orbitrap Mass Spectrometry. *Analytical Chemistry*, Washington v. 85, no. 11, p. 5288-5296, Apr 2013.

Enviado em 26/01/2015

Aprovado em 23/04/2015