

科学研究費助成事業（科学研究費補助金）研究成果報告書

平成24年 6月15日現在

機関番号：82612

研究種目：基盤研究（C）

研究期間：2009～2011

課題番号：21500293

研究課題名（和文）バイオインフォマティクス手法による細胞発現糖鎖の網羅的解析と生物学的意義の検討

研究課題名（英文）Comprehensive and biological analysis of glycosphingolipid function expressed on cell using bioinformatics technique

研究代表者

中島 英規（NAKAJIMA HIDEKI）独立行政法人国立成育医療研究センター・小児血液・腫瘍研究部・研究員

研究者番号：30450620

研究成果の概要（和文）：質量分析装置（MS）と液体クロマトグラフィー（LC）を組み合わせた LC-MS のデータをバイオインフォマティクス手法により解析するソフトウェアの開発に協力し市販されるに至った。このソフトウェアを用いて様々な病態の小児白血病細胞を分析したところ、統計学的にそれぞれの病態で異なった発現糖鎖パターンを示すことを明らかにした。特に乳児性白血病ではこれまで報告のない糖脂質 Ln3Cer が発現していることを見いだした。

研究成果の概要（英文）：The data processing software of liquid chromatography - mass spectrometer (LC-MS) was developed and the software was applied for the analysis of glycolipid expression in some different types of infantile leukemia cell. Each different type of leukemia cells was categorized using the software by the statistical analysis of glycosphingolipid expression. Especially, new glycosphingolipid marker, Ln3Cer could be discovered in some type of leukemia cell using the software analysis.

交付決定額

（金額単位：円）

	直接経費	間接経費	合計
2009年度	1,800,000	540,000	2,340,000
2010年度	800,000	240,000	1,040,000
2011年度	800,000	240,000	1,040,000
年度			
年度			
総計	3,400,000	1,020,000	4,420,000

研究分野：総合領域

科研費の分科・細目：情報学・生体生命情報学

キーワード：糖鎖・情報工学・マイクロアレイ・生体分子・細胞・組織

1. 研究開始当初の背景

糖鎖は、核酸、タンパク質に続く第三の生命鎖と呼ばれ、発現が遺伝子に直接支配されない機能分子の一つとして細胞機能に重要な役割を發揮することが知られている。糖鎖はデンプンやグリコーゲンのように単糖が直鎖状に連なった単純ポリマー以外に、タンパク質や脂質に結合した形の複合糖質、すなわち「糖タンパク質」、「糖脂質」、「プロテオグリカン」として存在する。核酸やタンパク質

鎖が直鎖状であることに対して、複合糖質の糖鎖は構成単糖が多様であること、結合位置・様式が多様であることに加え、分枝構造が存在するなど構造が非常に複雑である。細胞に発現する複合糖質の糖鎖発現パターンは細胞の種類によって異なることに加え、成熟やがん化に伴って複雑に変化するため、これら細胞の機能に何らかの役割を担っていると考えられているが、詳細については不明な点も多い。そのためこれまでゲノミクス分

野で行われている網羅的な発現遺伝子解析と、細胞機能の関連を調べるようなバイオインフォマティクスのアプローチを発現糖鎖解析に応用することで、細胞の機能や働きと発現する糖鎖の機能解明に利用できる可能性がある。

複合糖質のうち糖脂質の分析ではこれまで、薄層クロマトグラフィー (TLC) が頻繁に使用されてきたが、感度が低いため比較的多く発現している糖脂質しか観察できなかった。糖タンパク質糖鎖の分析については、タンパク質に結合したままでは分析が困難なため、通常糖鎖のみを化学的もしくは酵素により遊離させる必要がある。遊離させた糖鎖はそれ自身、特異的な吸収や蛍光を持たず検出できないため、分析には糖鎖の還元末端残基を利用して蛍光標識して分析することが多く行われる。特にピリジルアミノ化 (PA 化) した糖鎖は原理の異なる 2 種類の高速液体クロマトグラフィー (HPLC) で分離し同定する技術が確立され、現在では 2 次元 HPLC 糖鎖マップとして数百の N 結合型糖鎖の分離同定、精製が可能となっている。しかしながら HPLC やカラムなど研究室ごとで微妙に条件が異なることもあり、完全な糖鎖同定や定量分析を行うためには PA 化標準糖鎖を多数用意しなければならなかった。近年高感度な MS が開発され微量な糖鎖の存在確認や構造解析に頻繁に利用されている。MS において糖鎖イオンの断片化を起こさせると、糖鎖の配列情報や結合位置に関する情報も得られるが、アノメリック位の立体配置の決定は不可能である。しかし上述の 2 次元 HPLC 糖鎖マップで行われているような液体クロマトグラフィー (LC) と MS を組み合わせた LC-MS では、予め LC でそれら質量数が同じアノメリック異性体を分離することができれば、標準糖鎖と比較することで質量数の同じ糖鎖であっても構造異性体を分離同定することは可能である。N 結合型糖鎖に関しては現在、多段階 MS (MSⁿ)、高分解能 MS 及びスペクトルデータベースが構築されつつあり、それが利用できれば細胞に発現する微量糖鎖を含め、網羅的な発現糖鎖情報を得ることが可能になる。

MS では構造解析が可能なこと、高感度であることが非常に有用な点であるが、一方で分子の種類によってイオン化効率やシグナル強度が異なるため、たとえある糖鎖で別の糖鎖の 2 倍のシグナル強度が得られたとしても単純に 2 倍になっているというような量的な比較することはできない。そのため MS のシグナル強度から試料の量を直接「定量」することは予め量の分かった比較標品がない限りは不可能である。細胞に含まれる糖鎖は多様であり、その全ての標品を用意することは事実上困難であるため、MS による異なる細

胞同士の定量的糖鎖発現パターン比較は難しかった。近年、表面プラズモンとエバネッセント波を応用したレクチンアレイが開発され、検体中に含まれる糖鎖の情報を高感度に検出することが可能になった。このアレイに固定化されたレクチンの結合特異性と標準糖鎖間の結合定数が精密に決定されているため、アレイから得られた結果をバイオインフォマティクスの手法で評価し、複数の試料中に含まれる糖鎖情報を比較することに利用されている。しかし質量分析装置のように一つの試料中に含まれる多数の糖鎖の構造情報を詳細に得ることはできない。

DNA マイクロアレイは近年の技術革新により、数万の遺伝子でもたった 1 枚のアレイで解析することが可能になっている。このような膨大な情報を扱うためにはバイオインフォマティクスのアプローチが不可欠である。DNA マイクロアレイで得られる結果も MS の場合と同様に、あるスポットが別のスポットの 2 倍のシグナルが得られたとしても、プローブと試料中の遺伝子間の相互作用がそれぞれ異なるため、実際量が 2 倍になっているとは限らない。また複数のマイクロアレイで全く同じ試料を解析した場合においても同じ遺伝子スポットで全く同じ強度のシグナルになることはない。そのためマイクロアレイを用いた実験では各マイクロアレイ間で比較が可能になるように、データの正規化を行い有意差解析、クラスタリング等の統計学的な解析が行われている。このような手法は LC-MS で得られた結果の解析にも適用可能であり、様々な生産現場などにおいて製品の品質管理に応用されつつある。LC-MS は試料中に含まれる成分を LC で分離しながら順次マススペクトルを取得する装置である。すなわち LC からのある溶出時間において、マススペクトルが一つずつ得られることになる。つまり LC-MS で得られる結果には、(溶出時間, m/z 値, シグナル強度) の 3 つのファクターが存在する。マイクロアレイの結果も同様に、(スポットの行, スポットの列, シグナル強度) の 3 つのファクターでデータを表すことが可能である。このことは LC-MS の結果もマイクロアレイと同様な統計学的解析を行うことが可能であることを意味する。LC-MS では同種の試料であっても、統計学的解析の結果各試料間の細分類を行うことが可能である。また、それらの分類において特徴的な成分を見いだすことができたなら、MS/MS 解析を行うことでその成分を同定することも可能である。このようなアプローチは近年、品質管理以外にメタボロミクスや薬物代謝においても適用され、METLIN Search (<http://metlin.scripps.edu>) を利用することで化合物の同定も容易になりつつある。糖鎖は従来の分析方法では分析対象の

数が限られていたため、このような膨大なデータを扱うバイオインフォマティクス的手法はあまり応用されてこなかった。また1種類の糖鎖のみに着目した生物機能について調べられることはあったが、数種類の糖鎖の複合的な機能について検討することは困難であった。しかし近年MSが糖鎖の分析に利用されるようになって発現糖鎖の網羅的解析が可能になり、得られる情報量が膨大になった。従って、DNAマイクロアレイの場合と同様にバイオインフォマティクスを駆使した複数のターゲット糖鎖を複合的に解析することは非常に有効であると考えられる。

2. 研究の目的

本研究は、高感度で構造解析が可能な質量分析装置 (MS) と液体クロマトグラフィー (LC) を組み合わせた LC-MS を使用して様々な細胞が発現する糖鎖を分析し、得られた発現糖鎖情報をマイクロアレイで培われたバイオインフォマティクス的手法を応用して各試料間のデータの正規化、有意差解析、クラスタリング等、統計学的データ解析を行って発現糖鎖パターンによる細胞の判別や診断等に応用する基礎技術を確立するとともに、様々な細胞の発現糖鎖をこの手法で解析して細胞機能と発現糖鎖の関連性を調べることを目的とする。

3. 研究の方法

MSを用いたグライコムクス研究においてバイオインフォマティクス的手法を適用するためにはLC-MSデータを統計解析ソフトウェアに読み込ませる必要がある。そのため、

(1) LC-MSの結果を統計解析するためのデータの書き出し及び読み込み方法を最適化する(2) LCの溶出時間は厳密に各実験で一致しないため、ある程度許容を持たせる(3) MSで検出するイオンには、「価数」が異なるものがあり、同じ分子であってもスペクトルが単一にならない場合があるため、LC-MSデータにおける全てのスペクトルを単一の価数で表示させる(デコンボリューション)

(4) 他の分子(イオン)が糖鎖のイオンに結合(アダクト)することがあるため、単一の糖鎖であっても複数種類のイオンが現れるという問題を解決する、等の問題点を解決する必要がある。LC-MSデータについては、MSサプライヤー間の統一されたデータ書き出しフォーマット(mzxml file)ができ、それを取り込むことができる統計解析ソフトウェア GeneSpring MS が市販化されたため、それを使用すれば読み込みは可能になったが、

(1)～(4)の問題点を含め、その他いくつかの問題が解決されていないため、現状では単純にファイル変換、読み込みを行っただけでは信頼できる統計的解析できない。近年

MS サプライヤー各社自身から統計解析ソフトウェアが発表されつつあるが未だ洗練されておらず、GeneSpringほど高度な統計処理は困難である。そのため双方のソフトウェアやExcelのマクロ機能等を共用して統計解析が可能になるようにデータファイルの調整を行い、最終的にはGeneSpringMSで統計的解析を行うことを目指す。また、細胞に含まれる糖鎖を含む生体分子の種類は膨大であるため、網羅的な解析系での精度を上げるため、各試料に対してある程度細胞の機能に基づいた分画法を構築する。

(1) LC-MSでの糖鎖分析実験の最適化

① 糖脂質

市販標準糖脂質を様々な組み合わせ・濃度でLC-MSで分析し、その結果を統計解析ソフトウェアに取り込んで検定・クラスタリングを行い、組み合わせ・濃度に応じた結果が得られるか検証する。その際、MSサプライヤー(Bruker社)のソフトウェアDataAnalysis, ProfileAnalysis及び表計算ソフトウェア等を用いてデータの最適化を行い、最終的にGeneSpring MSを用いて解析するための手法を確立する。

糖脂質に関しては、既に細胞からの固相抽出法を用いた分画法、LC-MSによる分析法を確立しており、MS-MSスペクトルのライブラリーを有している。そこで順次様々な細胞株の糖脂質をLC-MSで分析した結果を統計的に解析し、クラスタリングを行い細胞の種類・由来に応じた分類と一致するのか、また特徴的な糖鎖を見つけ出すことができるのか検証する。

② N結合型糖タンパク質糖鎖

市販N結合型糖鎖標準品を用いてLC-MSの分離・分析条件を最適化する。最適化された分析条件において、糖脂質の場合と同様に標準糖鎖混合物を用いてGeneSpring MSによる統計的解析が行えるのか検証を行う。

組成が複雑な細胞由来のN結合型糖鎖での分析を行う前に、N結合型糖鎖を持ち、既に個体差等で糖鎖構造が変化することが報告されている精製糖タンパク質を用いてヒドラジン分解を行って糖鎖を遊離させ、固相抽出法により遊離糖鎖を部分精製する。糖鎖は2-アミノピリジンにより標識(PA化)してLC-MSで分析する。精製糖タンパク質としてははじめに、市販品及びロットの異なった動物由来血清から精製したイムノグロブリンG(IgG)を用いて統計解析が可能か検証する。

(2) 各種複合糖質糖鎖の細胞など生体試料からの分画・調製法の確立

細胞には膨大な種類のN結合型糖鎖が発現し

ている。網羅的な解析系での精度を上げるため、各試料に対してある程度細胞の機能に基づいた分画法を構築する。近年、界面活性剤等を利用した簡便な細胞分画法が開発されているので、その分画法に従って(i) 細胞質、(ii) オルガネラ、(iii) 核、(iv) 細胞骨格の4つのフラクションに分画する。分画した試料をそれぞれヒドラジン分解でN結合型糖鎖を遊離させ、固相抽出法を組み合わせて糖鎖を部分精製し、PA化してLC-MSによる分析を行う。

この分画法以外に細胞表面のみの膜を分画し、同様にPA化糖鎖を調製して分析を行う。これらの結果に対して統計学的解析を試みる。

(3) O結合型糖タンパク質糖鎖

ヒドラジン分解の条件を変えることで糖タンパク質N結合型糖鎖より優位にO結合型糖鎖を遊離させることができることが知られている。市販品で入手可能なフェチュイン、アジアロフェチュインはN結合型糖鎖とO結合型糖鎖の双方を持っていることが知られている。フェチュイン、アジアロフェチュインを様々な条件でヒドラジン分解し、遊離糖鎖をPA化してLC-MSで分析してO結合型糖鎖に最適なヒドラジン分解の条件を決定する。その際統計的手法を活用する。条件が決定後、順次上記細胞試料に対してO結合型糖鎖の調製を行い、LC-MSによる分析、統計解析を行う。

(4) LC-MSによる発現糖鎖情報と発現糖鎖制御遺伝子発現との相関

糖鎖の発現は主に糖転移酵素、糖鎖分解酵素により制御されている。LC-MSによる糖鎖発現情報及び試料間での統計的解析を行った検体に関し、DNAマイクロアレイ並びにリアルタイムPCRを行って糖転移酵素、糖鎖分解酵素の発現を調べ遺伝子発現と糖鎖産物との間の相関について検証する。

(5) 細胞発現糖鎖の網羅的及び統計学的解析結果についての生物学的意義の検討

① DNAマイクロアレイデータとの統計的解析

LC-MSによる糖鎖発現情報及び試料間での統計的解析を行った検体に関し、DNAマイクロアレイを行い、細胞の発現糖鎖パターンと発現相関のある遺伝子群をピックアップする。その中から糖鎖機能と関連があると報告のある遺伝子をピックアップし、糖鎖とその生物学的機能について検証を行うとともにパスウェイ解析を行う。顕著な動きのある遺伝子に関しては、その遺伝子ないしは糖転移酵素強制発現系を構築し、生物学的機能についての検証を行う。

4. 研究成果

研究開始当初、LC-MSのデータをバイオインフォマティクス手法により解析するためのアライメント調製ソフトウェアはほとんど存在せず、当初使用予定であったGene Spring MSやProfile Analysisは使用に耐えうるほどの精度はなかった。特に国産のもので市販されていたものは皆無であった。現在、このようなソフトウェアは海外メーカーのものを中心に主に質量分析装置メーカーからいくつか市販されるようになった。この現状を考えると、研究開始時点における着眼点は良かったものと考えている。このようなソフトウェアのうち精度がよいと判断されたものは、海外メーカーのものではProgenesis、国内メーカーではライフイクス社のSignPostである。本研究では当初、データを取得するブルカー社製質量分析装置との連携を考え、ブルカー社との共同研究でこのようなソフトウェアを開発する予定であったが、SignPostのアライメント調製精度やアルゴリズムが優秀であること、将来的に統計学的解析が可能ないようにデザインされていることなどから、このソフトウェア開発に協力することで本研究への無償使用という形の共同研究となった。このソフトウェアは現在市販され、医療分野ばかりでなく薬物動態や工業化学製品・食品の品質管理などのLC-MSが頻繁に使用される分野で活躍するに至っている。

http://www.reifycs.com/japanese/SPMSoutline_jp.html

これまで、LC-MSにより取得されたデータをライフイクス社より開発されたLC-MS用データ解析ソフトウェアSignpost MSにより読み込ませ、LC-MSによる各分析で起こるアライメントのズレを補正することができた。さらに様々な細胞から抽出した糖脂質をこの手法で解析してそれぞれの細胞に発現する糖脂質の発現パターンの相違について統計学的な解析を行った。特に、様々な病態の小児白血病の細胞検体を分析したところ、統計学的にもそれぞれの病態で異なった発現糖鎖パターンを示すことが明らかになった(図1)。その結果を詳細に検証すると、過去の報告とも一致することに加え、乳児性白血病にはこれまで報告のない中性糖脂質Ln3Cerが発現していることを見いだすことができた(図2)。今後更に解析する検体数を増やし、実際の診断への有用性を示そうと考えている。

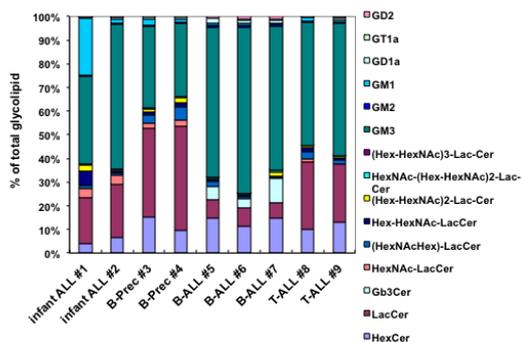


図1 小児白血病における発現糖脂質パターン

Ln3Cer

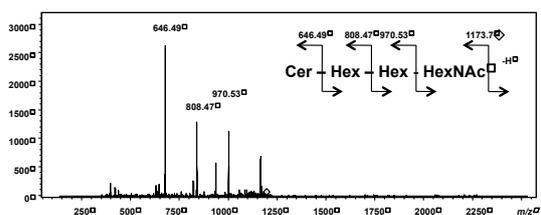


図2 小児白血病に発現するLn3CerのMS/MSスペクトル

5. 主な発表論文等

(研究代表者、研究分担者及び連携研究者には下線)

[雑誌論文] (計0件)

[学会発表] (計4件)

- (1) 中島英規, 藤本純一郎, 奥山虎之・瀧紙血検体を用いたライソゾーム病の欠損酵素活性測定・日本マススクリーニング学会・2010年8月28日・神奈川県 ワークピア横浜
- (2) 中島英規, 藤本純一郎, 奥山虎之・タンデムマス質量分析器を用いたポンペ病の診断法の開発・日本先天代謝異常学会総会、アジア先天性代謝異常症シンポジウム・2010年10月21日・大阪府 大阪国際会議場
- (3) 尾島琢磨, 中島英規, 梅澤明広・ヒト人工多能性幹(iPS)細胞における糖脂質の比較解析・日本再生医療学会・2011年3月1日・東京都 京王プラザホテル
- (4) 中島英規・LC-MS/MSによる代謝異常症の高感度検出・日本電気泳動学会シンポジウム・臨床プロテオーム研究会連合大会・2011年5月10日・山口県 山口大学

[図書] (計0件)

[産業財産権]

○出願状況 (計0件)

○取得状況 (計0件)

[その他]

ホームページ等

http://www.reifycs.com/japanese/SPMSoutline_jp.html

6. 研究組織

(1) 研究代表者

中島 英規 (NAKAJIMA HIDEKI)

国立成育医療研究センター・研究員

研究者番号：30450620

(2) 研究分担者

()

研究者番号：

(3) 連携研究者

()

研究者番号：