

Title	GC/MSメタボロミクスにおける未同定代謝物同定の効率化および品質向上に向けた保持指標予測モデルの構築とその応用
Author(s)	松尾, 晃子
Citation	大阪大学, 2018, 博士論文
Version Type	VoR
URL	https://doi.org/10.18910/69533
rights	
Note	

The University of Osaka Institutional Knowledge Archive : OUKA

https://ir.library.osaka-u.ac.jp/

The University of Osaka

博士学位論文

GC/MS メタボロミクスにおける 未同定代謝物同定の効率化および品質向上に向けた 保持指標予測モデルの構築とその応用

松尾 晃子

2018年1月

大阪大学大学院 工学研究科 生命先端工学専攻

目次

略号	1
第一章 緒論	3
1.1. オミクス解析	3
1.2. メタボロミクス	4
1.3. ノンターゲットメタボロミクスに使用される測定手法	5
1.4. GC/MS ノンターゲットメタボロミクスの代謝物同定の現状	6
1.4. 本研究の戦略	9
第二章 保持指標予測モデルの構築	11
2.1. 緒論	11
2.2. 実験方法	12
2.2.1. 保持指標モデルの構築	
2.2.2. 保持指標の系統誤差	
2.3. 結果と考察	17
231 保持指標予測モデル	17
2.3.2 保持指標の系統誤差	
第三章 生薬センキュウ中の未同定代謝物	21
3.1. 諸言	21
3.2 実験材料および実験方法	22
3.2.1. 試薬およびカラム	22
3.2.2. 試料調製方法	23
3.2.3. 分析条件	24
3.2.4. データ解析	
3.2.4.1 データ行列の作成	
3.2.4.2 ピークの選択	
3.2.4.3 PCA	27

3.2.4.3 代謝物の同定とバリデーション	
3.2.4.4 休行相保と誤问足率の計算 3.3 結果と考察	28
 3.3.1. PCA の結果 3.3.2. 未同定代謝物の定性	29 30
第四章 総括と展望	
謝辞	
引用文献	
発表論文	45
本学位論文に関与する論文	45
その他の原著論文	46
総説等	46
学会発表	47
国際会議	47
国内会議	49
付録	50
1-選択された記述因子	50
2-モデル構築に使用された代謝物	51
3-生薬センキュウ中の代謝物	62

略号

ABF: analysis base framework

- CE: capillary electrophoresis
- DNA: deoxyribonucleic acid

EI: electron ionization

FT-NIR: Fourier transform-near-infrared spectroscopy

GC/MS: gas chromatography/mass spectrometry

LC/MS: liquid chromatography/mass spectrometry

MoNA: MassBank of North America

mRNA: messenger ribonucleic acid

RNA: ribonucleic acid

NIST: National Institute of Standards and Technology

NMR: nuclear magnetic resonance spectroscopy

PCA: principal component analysis

PC: principal component

QC: quality control

QSRR: quantitative structure-retention relationship

RI: retention index

RT: retention time

RSD: relative standard deviation

SDS-PAGE: sodium dodecyl sulfate- poly-acrylamide gel electrophoresis

SD: standard deviation

SDF: standard data file

SMILES: simplified molecular input line entry system

TMS: trimethylsilyl

UV: unit variance

第一章 緒論

1.1. オミクス解析

メタボロミクスとは、サンプル中に含まれる多種類の代謝物を一斉に測定する、オミクス解析の一種である.代謝物は生体内で生合成、または分解されることで生じた低分子化合物の総称であり、アミノ酸、糖類、脂質、有機酸などの様々な物理化学的性質を持つ化合物が含まれている.

生体中の代謝物はゲノムの転写・翻訳によって変化する. そのため, 代謝物の網羅的な 測定データはサンプルの詳細な表現型とみなされている. (Nicholson J. K., 1999, Fiehn O., 2000, Fiehn O., 2002) 代謝物だけでなく, 生体サンプル中の DNA (deoxyribonucleic acid) や mRNA (messenger ribonucleic acid), タンパク質もまた表現型と相関を持つことが報告さ れており, それぞれの物質でオミクス解析が行われている.

DNA はセントラルドグマのもっとも上流に位置する物質であり、五炭糖、リン酸、塩基から 成る核酸のポリマーである.遺伝情報はDNAの塩基配列に格納されている.塩基配列の網 羅的な分析および解析はゲノミクスと呼ばれている.次世代シークエンサーの発展によって、 膨大な数の塩基配列を短時間で解析できるようになった.(Koboldt D. C., 2013) 一方でゲノ ミクスの研究が進むにつれ、遺伝子の発現には DNA のメチル化やヒストンの構造などの要 因も関与していることが明らかとなった.エピジェネティクスでは上述のような遺伝子発現調 節機構を研究の対象とし、DNA の構造や化学装飾の迅速かつ網羅的な分析法の確立が 試みられている. (Rivera C. M., 2013)

トランスクリプトミクスでは、RNA (ribonucleic acid)の網羅的な分析を行う. RNAはDNAの 転写によって生成する物質であり、リボース、リン酸、塩基から成る核酸である.マイクロアレ イを用いたハイブリタイゼーション法や RNA-seq 法によってサンプル中の RNA を測定し、 (Wang Z., 2009) 異なる種類のサンプル間 (たとえば健常者と疾患者、自然株と変異株な ど) での発現量を比較する. しかし転写された mRNA の量が増加したにも関わらず, その mRNA にコードされているタンパク質は変化しない事例も報告されているため (Greif P. A., 2011),表現型との相関はやや間接的であると言える. また,トランスクリプトミクスではゲノミ クスのようにデータの解釈にあたって,その生物種のゲノム情報が必須であるため,ゲノムが 解読された生物種に適用されることが多い.

タンパク質はアミノ酸のペプチド結合から成るポリマーである.タンパク質の最適な抽出条件は種類に応じて大きく異なるため、サンプル中の全てのタンパク質を一度の分析で網羅することは現状では困難である. SDS-PAGE (sodium dodecyl sulfate- poly-acrylamide gel electrophoresis) などのゲルを用いた電気泳動 (Wrzesinska R. A., 2013),キャピラリー電気泳動 (Righetti P. G., 2013),質量分析器 (Aebersold R., 2016) など様々な分析法によって、表現型の違いによって有意に変動したタンパク質が探索される.測定されたタンパク質はアミノ酸配列に基づいて同定される.未知のタンパク質に対してはアミノ酸配列の類似性に基づいた生理機能の類推が行われているが、タンパク質の種類は数万種にも及ぶと言われており、類推不能なタンパク質も数多く存在している.タンパク質の立体構造や生理機能を効率に決定する手法は発展の途上にあり、その進捗に期待が寄せられている.

1.2. メタボロミクス

メタボロミクスでは代謝物を測定対象とする. 代謝物は遺伝情報が転写, 発現した結果で あるため, 表現型との相関はより直接的である. (Fuhrer T., 2015) 各生物種のゲノム情報は 代謝経路の推測などに有益であるが, データ解析に必須の情報ではない. また, 中央代謝 系に代表される一部の代謝経路は異なる生物間でも共通しているため, ある生物種のサン プルで構築した分析メソッドやデータ解析の手順を, まったく別の生物種へ応用することが 容易である. 以上の理由から, メタボロミクスは微生物学 (Xie G., 2013), 病態生理学 (Spratlin J. L., 2009), 食品工学 (Cevallos-Cevallos J. M., 2009), 栄養学 (Jones D. P.,

-4-

2012) など幅広い分野のサンプルへと応用されている.メタボロミクスの更なる普及に向け, 初心者にも扱いやすい客観的な測定・解析法の開発が急務となっている.

メタボロミクスの手法には、フィンガープリンティングとプロファイリングが存在する.フィンガ ープリンティングは異なるサンプル間の測定データを比較し、差異を示す領域を明らかにす る. (Pongsuwan W., 2007) データ解析において代謝物の同定・定量の手順を必要としない ため、フィンガープリンティングのスループットは高い.一方プロファイリングでは代謝物ごと の定量情報が得られるため、代謝経路に生じた変動の把握など、考察やフィードバックの深 めやすさにおいて利点がある.

プロファイリングは測定対象とする代謝物をあらかじめ決めるか否かによって,大きく二種 類に分けられる.ターゲットメタボロミクスでは,測定する代謝物をあらかじめ決定している. この手法は,特定酵素に関わる遺伝子組み換え体の評価など測定すべき代謝物群が明確 な研究に応用されている.一方,測定前に分析対象物を絞り込まずに可能な限り多くの代 謝物を一斉に測定する手法は,ノンターゲットメタボロミクスと呼ばれている.この手法では 予期せぬ変動を明らかにするため,バイオマーカー代謝物の探索などの研究で利用されて いる. (Beger R. D., 2012)

1.3. ノンターゲットメタボロミクスに使用される測定手法

代謝物の物理化学的な性質は,脂質のような疎水性の分子や糖,有機酸などの高極性の物質まで多岐に渡る.そのため,測定手法は重要な要素となる.

NMR (nuclear magnetic resonance spectroscopy) では磁場内におかれた原子に電磁波を 照射し, 生じた核磁気共鳴を検出する.¹H, ¹³C, ¹⁵N などの代謝物を構成する原子が核磁 気共鳴をするため, 多くの種類の代謝物を一斉に検出できる. 核磁気共鳴は非常に微細な 信号であるため, 高濃度かつ多量のサンプルが必要となる. (Smolinska A., 2012)

FT-NIR (Fourier transform-near-infrared spectroscopy) は分子が赤外線を吸収する性質 を利用している. 吸収されたエネルギーによって分子の振動や回転が励起状態となる. 励起

-5-

に必要なエネルギーは分子の構造に依るため,赤外吸収スペクトルは分子に固有のパターンを示す.しかし,メタボロミクスのような混合物のサンプルでは信号が重複してしまうため, 代謝物の同定や定量は困難である. (Wallace M., 2012)

質量分析ではサンプル中の代謝物をイオン化し、その強度を測定する. NMRと比べ感度 が高く、少量のサンプルで測定が可能である. 代謝物ごとの定量値を得るために、クロマトグ ラフィーや CE (capillary electrophoresis) などで代謝物を分離してから質量分析器で検出 することが多い. 代謝物の同定は、ピークが溶出した時間 (クロマトグラフィーでは保持時 間、CE では泳動時間と呼ばれる) と、質量分析から得られる m/z の情報に基づいて行 われる.

GC/MS (gas chromatography/mass spectrometry) は、気体を移動相としたクロマトグラフィーと質量分析を接続した機器で、ノンターゲットメタボロミクスで使用される機器の一つである. LC/MS (liquid chromatography/mass spectrometry) や CE/MS と比べ比較的安価で、技術習得が容易という特徴がある. (Kanani H., 2008) また気体の拡散係数は液体より高いため、一般に GC のカラム長は LC や CE より長い. そのため GC の理論段数は高く、多種類の代謝物を一斉に分離することに適している.

GC/MS では揮発性代謝物だけでなく, 誘導体化と呼ばれる前処理によって, 糖やアミノ 酸, 有機酸などの親水性代謝物も測定されている. 誘導体化では, 測定対象物と誘導体化 試薬を反応させ特定の官能基の構造を変化させる. この操作は測定対象物の沸点の低下, カラム吸着の低減, ピーク形状の改善などを目的としている. 操作の簡便性と汎用性の広さ から, GC/MS メタボロミクスでは, シリル化 (Kind T., 2009) とアルキル化 (Smart K. F., 2010) による誘導体化がよく使用される.

1.4. GC/MS ノンターゲットメタボロミクスの代謝物同定の現状

2007年に提案された Minimum reporting standards では純粋な標準品を使用し、複数

-6-

の測定法による確認を行った代謝物同定が最も確からしい結果とみなしている.しか しノンターゲットメタボロミクスでは,測定対象物をあらかじめ決定しない,同定すべき代謝 物が数百種類以上に及ぶなどの理由から,標準品の利用が困難である.そこで純粋な標準 物質の測定結果を格納した,ライブラリと呼ばれるデータベースの検索もまた,代謝 物同定の手法として認められている.(Sumner L. W., 2007)前節で述べた通り,GC/MS 代謝物の同定は各ピークの保持時間とマススペクトルに基づいて行われるため,保持 時間とマススペクトルのライブラリが必要となる.

GC/MS のイオン化部には主に EI 法 (electron ionization 法) が用いられる. EI 法に よるマススペクトル (以後, EI スペクトルと呼ぶ) はイオン源の構造に依存しないた めに再現性が高く, 大規模なライブラリが存在する. 例えば, MoNA (MassBank of North America) では 9003 種の分子に由来する 15,302 の EI スペクトルが一般に公開されて いる. 有償のライブラリでは, Wiley (第 10 版) では 242,477 種の分子から得られた 276,259 の EI スペクトルが, NIST14 では 583,059 種の分子から得られた 719,456 の EI スペクトルが登録されている. 代謝物が示す EI スペクトルを予測する手法も報告 され (Allen F., 2016), 利用可能な EI スペクトル情報は拡大し続けている.

EI 法ではフィラメントから生じた電子を分子に衝突させ、フラグメントイオンを生成する. 化学結合によって開裂に必要なエネルギーは異なるため、構造が類似した代謝物は類似した EI スペクトルとなる. また EI 法はハードイオン化法の一種であるため、分子イオンの強度はフラグメント化されたイオンと比べるとはるかに低く、観測されないこともしばしば生じている. 以上の理由から、EI スペクトルの類似性のみに基づいた代謝物の同定では、多くの偽陽性が生じる傾向にある. (Wagner C., 2003)

GC/MS の保持時間は LC や CE と比べ高い再現性を持ち, 代謝物の同定の重要な手掛 かりとなる.保持時間はカラム長によって左右されるため, 特定の物質を基準に補正を行う. 補正された値は保持指標 (Retention index) と呼ばれ, Kovats indices (Kovats E., 1958), Lee indices (Lee M. L., 1979), Fiehn indices (Kind T., 2009) などのいくつかの計算法が提 案されている. Fiehn 研究室が公開している BinBase (Wishart D. S., 2007) や Max Planck

-7-

研究所によって運営されている Golm (Strehmel N., 2008) などのライブラリには,各代 謝物の純粋な標準品から得られた保持指標の実測値が登録されている.しかし,保持 指標のライブラリに登録されているスペクトル数は2017年現在,Binbaseでは1021,Golmで は9156 にとどまっており, EI スペクトルのライブラリに比べて非常に少ない.

EI スペクトルのフラグメントの強度は主に、イオン化電圧によって決定される. EI スペクトル のライブラリに登録されているデータの大半は 70 eV のイオン化電圧で統一されているため、 異なる機器から得られた EI スペクトルの比較は比較的容易である. しかし保持指標はカラム の種類、昇温条件、機器の種類など複数の要因によって変動する. ライブラリの規模に加え、 全ての条件が合致したデータを探し出すことは、はるかに困難となっている.

ライブラリの検索によって得られた候補代謝物に対しては,偽陽性を取り除くため に純粋な標準品による確認が必須となる.そのため,一つの未同定ピークに対し多く の偽陽性が存在することは,スループットの低下とコストの増加の要因となってしま う.

さらに GC は分離能が高いため,一度の測定で数百から千のピークが得られる.これらの 中には,カラムの剥離した装飾基や未反応の誘導体化試薬に由来するピークも存在す る.このような代謝物に由来しないピークは,多変量解析の結果に考察に不要な影響 を与えてしまうため,解析に供するデータから取り除く必要がある.(Styczynski M. P., 2007) ピークの S/N (Signal/Noise) を閾値としたピーク選択や QC サンプル (Quality control;測定の安定性を評価するためのサンプル,全種類のサンプルを混合したもの などが使用される) との比較によるピーク選択 (Godzien J., 2015) などが行われてい る.しかし,これらの手法では適切な閾値の設定が経験則に依るため,解析者によっ て結果が変動してしまうという問題点がある.

GC/MS 測定データには数多くの未同定ピークが存在していることに加え,代謝物の保持 指標の情報が限られているためにピークごとに多くの偽陽性が存在しているために,GC/MS ノンターゲットメタボロミクスの代謝物同定には多くの労力と費用が必要となってしまっている. GC/MS メタボロミクスが持つ,堅牢かつ初心者にも扱いやすい測定法という利点を生かすた

-8-

めにも,客観的かつ効率的な代謝物同定法の確立が求められている.

1.4. 本研究の戦略

前節で述べた GC/MS ノンターゲットメタボロミクスにおける代謝物同定の問題点を解決するために、本研究では QC サンプル希釈系列に基づいたピークピッキング法と、保持指標予測モデルの構築とその活用を提案し、これらの組み合わせを実サンプルへと応用した.

第二章では保持指標予測モデルの構築と、構築したモデルによる代謝物同定の精度向 上について記述した. EI スペクトルの類似性に基づいて探索された、未同定代謝物の候補 (以後,候補化合物と呼ぶ)には、多くの偽陽性が含まれている.これらの候補化合物を更 に絞り込むためには、保持指標が重要な手掛かりとなるが、保持指標のライブラリの規模は EI スペクトルと比べて小規模であるため、保持指標ライブラリに登録されていない候補化合 物も多く存在している.そこで代謝物の保持指標と物理化学的性質の関係式を求める QSRR 法(quantitative structure-retention relationship)(Kaliszan R., 1992)を利用する. QSRR 法によって構築されたモデルから保持指標のライブラリに登録されていない候補化合 物の保持指標を予測し、同定するピークの保持指標と比較することで候補化合物をさらに 絞り込む.本研究では *in-silico* での誘導体化を行うソフトウェア"MetaboloDerivatizer"を利 用することで、誘導体化された代謝物群にも適用可能な保持指標予測モデルの構築法を 開発した.

第三章では QC サンプルの希釈系列ならびに多変量解析を利用したピークピッキング法と, 第二章で構築した保持指標予測モデルを漢方薬センキュウ (Condium officinale Makino, Ligusticum chuanxiong Hort) の GC/MS データへ適用した.

ピークピッキングでは第一段階として、QC サンプルの希釈系列に基づいたノイズ除去を 実施し、第二段階として PCA (principal component analysis) のローディングベクトルに行う 仮説検定に基づいた変数選択を行った.これらの手順によって、代謝物に由来しかつ統計 的に有意なピークを絞り込んだ.以上の手順によって選択された未同定ピークに対して、EI スペクトルの類似性と保持指標予測によって代謝物同定をした.

第四章には本研究で達成された事柄の総括と、今後の展望について記載した.

第二章 保持指標予測モデルの構築

2.1. 緒論

クロマトグラフィーの保持指標値は、化合物と固定相の相互作用によって決定する.相互 作用は分散力、水素結合などに影響され、これらは代謝の立体構造や帯びている電荷によ って変化する.このことを利用し、化合物の物性や構造情報から保持指標を計算する、 Quantitative structure-retention relationship (QSRR) が行われてきた. (Kaliszan R., 1992)

QSRR モデルの構築にあたっては、まずトレーニングセットとする保持指標情報を標準品の測定から取得する.次に、これらの分子種の物理化学的性質と保持指標との間で相関の高い変数を選択し、保持時間予測モデルを構築する.そして当該モデルを用いて、トレーニングセットに含まれない代謝物の保持指標を予測する.

GC/MSでは、多くの場合ある特定の化合物群を対象としたQSRRモデルの構築が数多く 報告されている. (Fatemi M. H., 2011, Hu R. J., 2005, Hemmateenejad B., 2007) しかしノン ターゲットメタボロミクスでは、可能な限り多くの種類の代謝物を測定することを目標としてい るため、代謝物の種類に関する前情報は少ない、そのため、これらのモデルの利用は難し い.

一部の研究では、代謝物群に制限がないモデルの構築も行われている. Stein S. E.らが報告した"group contribution method"では代謝物の部分構造を84種類に分け、それぞれの部分構造が保持指標に与える影響を係数として設定し、約35,000種の代謝物の保持指標を予測した. しかしこの手法は原理上、異性体の区別ができないという問題を抱えている. さらにシリル化などで誘導体化された代謝物に対しては、Kovat index による保持指標で300-400の誤差が生じたことが報告されている. (Stein S. E., 2007) Kumari S.らは TMS 化(trimethylsilyl) された代謝物の予測値に対し、一定の数値の加算、または実測値から得られた二次関数式による修正を提案し、Fiehn library に登録されている代謝物群への有用性

を示した. (Kumari S., 2011) しかし,この補正値はライブラリに登録された保持指標の実測 値に依存した数値であるため,その一般性については十分に担保されていない.

誘導体化はメタボロミクスで広く用いられている技術であり,汎用性と精度の高さを両立し たモデルの構築が望まれている.

本章では,誘導体化された代謝物に適用可能な保持指標予測モデルを構築し,その性 能について評価する.

2.2. 実験方法

2.2.1. 保持指標モデルの構築

GC/MS メタボロミクス用ソフトウェアである, Aloutput2 (Tsugawa H, 2011) に実装されてい る保持指標と EI スペクトルのライブラリを基に保持指標予測モデルを構築した. このライブラ リに格納されているデータは, 同一の研究室内で誘導体化, 測定された純粋な標準品から 得られたものである.

以下にモデル構築の手順を記載する.本節で使用したデータベース,ソフトウェアの URL は引用文献に記載した.

Aloutput2 のライブラリに登録されている代謝物のの構造ファイルを SDF 形式 (standard data file) で Pubchem Download Service から取得した. Pubchem Compound のデータベー スには、シリル化またはオキシム化によって誘導体された代謝物はほとんど登録されていない. そこで MetaboloDerivatizer によって *in silico* での誘導体化を行った.まず、表 2-1 で示した 6 種類の官能基に対して生じる誘導体化反応を MetaboloDerivatizer に設定し、先の手順でダウンロードした代謝物の構造ファイルを供し、誘導体化された代謝物の構造ファイルを得た.表 2-1 で示したようにカルボニル基や一級アミノ酸への誘導体化では、複数種の誘導体化生成物が生成する. これらを精査するため、Chemaxson Jchem Cxcalc (ver 6.3.0, 2014, ChemAxon) によって誘導体化された代謝物の分子量を計算した.誘導体化された

代謝物の分子量と,その代謝物が示す EI スペクトルの分子イオンが一致しない代謝物はラ イブラリから取り除いた.この作業の結果,ライブラリに含まれる保持指標-EI スペクトルのデ ータ数は 337 個となった.

誘導体化された代謝物の記述因子をPaDEL Descriptor (version 2.18) を用いて計算した. (Yap C. W., 2011) パラメータは表 2-2 に記載した. この設定では, 合計 4566 種の記述因子 が計算された. これらのうち, 保持指標との相関係数が 0.8 以下となった記述因子を取り除 いた. その結果, 80 個の記述因子が選択された. (付録 1)

以上の過程を経て,代謝物の記述因子と各代謝物の保持指標から成るデータセットを作成した.代謝物を保持指標が低いものから順に並べ替え,奇数の代謝物を set1,偶数を set2 として振り分けた.これは回帰分析の際,トレーニングセット (モデル構築用のデータ) とし,テストセット (確認用のデータ) が必要となるためである.モデル構築に使用したデー タセットは付録 2 に記載した.

さらにデータセットの X 変数に UV (Unit variance) 法を行い, 各変数 (記述因子, 保持指標) の分散を1, 平均値が0となるように変換した. その後, R (version 3.2.2) によって, 保持指標を目的変数, 記述因子を説明変数とした重回帰分析に供した. 赤池情報基準値とクロスバリデーションの組み合わせによって適切なモデルの選択を行い, Forward-step によって予測モデルに重要な変数の選択を行った.



表 2-1. MetaboloDerivatizer による変換

[※]R1, R2 は水素以外の原子.

パラメータ名	設定した項目,または数値		
Descriptors	1D&2D		
Descriptors	Fingerprints		
	Remove salt		
Standardiza	Detect aromaticity		
Standardize	Standardize nitrogroup		
	Retain 3D coordinates		
Advanced	Log		
Advanced	Retain molecules order		
Max. threads	-1		
Max. running time per molecule	-1		
Max. compounds per file	30000		
Finger prints	Pubchem Fingerprinter		

表 2-2. PaDEL Descriptor によるデータ処理条件

2.2.2. 保持指標の系統誤差

構築したモデルから得られる保持指標の予測値と、標準品の測定によって得られる実測 値に生じた誤差の妥当性を評価するために、異なる測定機関で測定された同一代謝物の 保持指標が示す誤差を求めた.代謝物同定において参照する保持指標ライブラリは、様々 な研究機関で取得されたデータから成る.保持指標の利用によってカラム長やキャリアガス の流速の違いなどに由来する誤差は補正されるが、測定データに含まれる系統誤差を全て、 なくすことはできない.そのため、保持指標ライブラリに基づいた代謝物同定においては、測 定機関の違いに由来する誤差は許容されている.構築したモデルの予測値と実測値が示 す誤差もまた、上述の誤差と同程度もしくはより少なくなることが望まれる.

図 2-1 は 5 つの機関が公開している保持指標-EI スペクトルのライブラリのベン図で ある. モデルの構築に利用した Aloutput2 のライブラリは Osaka University (D) のデー タに準拠している. A の Fiehn binbase は UC Davis で測定され, MoNA (MassBank of North America) に登録されている. 残り 3 つの測定データは MassBank に登録されて いる. B の GL-science は, MassBank に登録されているデータのうち, 接頭語に"GLS" が付けられている. RIKEN では PR が, Kazusa は KZ, Osaka University には OUF が それぞれ接頭語となっている. Others は MassBank と MoNA に登録されている化合物 のうち, 保持指標が登録されていないものである.

図 2-1 で示したように 70 種類の代謝物が 5 種類のライブラリで共通している. こ のうち EI スペクトルによる識別が困難である糖類 (計 14 種類) を取り除いた,計 56 種類の代謝物から測定機関の違いによる系統誤差を求めた. Osaka University で測定 された保持指標を基準とし,その他 4 種類のライブラリとの差分を代謝物ごとに計算 し,その標準偏差を求めた. (図 2-4) なお, (A) の Fiehn Binbase ライブラリでは, Fiehn index が使われているため,回帰式によって Kovats index に変換した. (Kind T., 2009)



A: Fienn BinBase (1,021 records, 503 unique structures) B: GL Sciences (494 records, 380 unique structures) C: Kazusa DNA (273 records, 150 unique structures) D: Osaka Univ. (430 records, 330 unique structures)

E: RIKEN (241 records, 182 unique structures)

図 2-1. 保持指標ライブラリの代謝物のベン図

2.3. 結果と考察

2.3.1. 保持指標予測モデル

図 2-2 は構築したモデルによる予測値と Aloutput2 のライブラリに登録されている実測値 に基づいて、予測モデルの基となったデータセットに登録されている代謝物をプロットした図 である. 黒色でプロットされた代謝物および記載された数値がトレーニングセットから得られ た結果となり、赤色で書かれたものはテストセットとなる. set1 と set2 でトレーニングセットとテ ストセットを入れ替えて行ったため、二種類のグラフが存在する.

いずれのモデルの結果においても R 二乗値は 0.93 となり, 予測値と実測値による代謝物 のプロットは高い直線性を示した.標準偏差値は 78 から 88 となった. Stein S. E.らが報告し た先行研究では TMS 基をもつ代謝物には 300-400 の誤差が生じていたが, 本モデルの利 用により, より精度の高い予測が可能となった.

どちらのモデルにおいても ATSc1, topoDiameter, MLFER_L, ETA_Beta が上位 4 つの重 要な記述因子として選択された. ATSc1 は Broto-Moreau autocorrelation method によって計 算される記述因子であり, 脂肪酸が持つ炭素鎖 (-(CH₂)_n-) や糖類のヒドロキシル基が付加 した炭素鎖 (-(CH(OTMS))_n-) など, 代謝物がもつ繰り返し構造を示す. topoDiameter は分 子の大きさを示す. MLFER_L はガス-ヘキサデカン間の分配係数を示し, 測定対象物とキ ャリアガス (He),カラムの疎水性を示す固相との相互作用を説明している. ETA_Beta は, 代謝物の電気的状態を示す記述因子であり, 固相との静電気的相互作用の影響を表す.

Lignoceric acid や behenic acid など、いくつかの脂肪酸の誤差は標準偏差の三倍以上と なる、大きな数値を示した. topoDimeter と保持指標による代謝物のプロットでは、脂肪酸は 分子の形が他の化合物と異なる傾向となった. (図 2-3) トレーニングセットに脂肪酸類が不 足していたことが、誤差の要因になったと考えられる.

本モデルではオキシム化によって生じる E 体, Z 体の異性体 (幾何異性体, geometrical

isomer) や糖類の立体異性体の予測が達成できなかった.これは,1,2 次元の記述因子や フィンガープリントでは立体構造の違いを表現することができなかったためである.本研究で モデル構築に使用した代謝物数は 337 種であり,先行研究と比べて非常に少ない.そのた め,計算に時間がかかる三次元の記述因子の応用が比較的容易であり,今後のモデル構 築の更なる発展が期待される.





トレーニングセット->テストセット, SD: Standard deviation





2.3.2. 保持指標の系統誤差

図 2-4 のグラフは、同一の代謝物を異なる機関で測定した際に生じた差分をプロットしたものである.

Kazusa DNA のライブラリと Aloutput2 のライブラリの間で生じた誤差の標準偏差は81となった, Fiehn Library とは42, RIKEN では40, GL Science では30となった.
保持指標予測モデルの標準偏差は78から88を示し, Kazusa DNA のライブラリが示した系統誤差と同程度となった.

Kazusa DNA ライブラリは Agilent 社 (Agilent Technologies Japan, Ltd., California, USA) の DB-17MS というカラム使用している. このカラムは 50%フェニル-メチルポ リシロキサンにほぼ同等の中極性のカラムである. これは, その他の4種のライブラ リで採用されている, 5%フェニル-メチルポリシロキサンの低極性カラムとは異なる 性質の固相である. 図 2-4 のプロットから保持指標による補正では, カラムの固相の 違いによる保持時間の変動は, 完全には取り除かれないことが示唆された. 保持指標 予測モデルの改善のため, トレーニングセットの数と種類を増やすことは有用である. しかし, 異なる固相のカラムから得られたデータを統合すると, 上述のような不要な変動の影 響を受けてしまう. この章で構築されたモデルのトレーニングセットは, 先行研究と比べ小 規模であるが, 異なる測定機関に由来する系統誤差と同程度の精度を達成したことから, ト レーニングセットの質もまたモデルの性能に深く関与していると言える.

同じ組成のカラムが示す系統誤差と比べ,本モデルが示した予測値の標準偏差,信頼区間はより大きな数値を示している.未同定代謝物の定性において,同一のカラムから得られた保持指標は,モデルから得られた予測値と比べ許容すべき誤差の範囲が狭いため,より正確な絞り込みが可能となり有利であると考えられる.保持指標における許容範囲の違いが代謝物同定の精度に与える影響について,次章 (3.3.3.保持指標の許容値と誤同定率)で議論した.

-19-



図 2-4. 各代謝物が示す機関間誤差のプロット

第三章 生薬センキュウ中の未同定代謝物

3.1. 諸言

本章では第二章で構築した保持指標予測モデルを応用し,生体サンプル中の未同定代 謝物の定性を試みた.また第一章の戦略で述べた,QC サンプルの希釈系列を利用したピ ークピッキング法も実施した.

本章において QC サンプルは, 測定対象とする全種類のサンプルを等量比で混合したものと定義する.この QC サンプルには原理上, サンプル中のすべての代謝物が含まれている. 本研究では QC サンプルの希釈系列を作製し, 各ピークのインテンシティと希釈濃度の相関に基づいて, サンプルに由来するピーク選択を行った.

その後, データ行列を作成し PCA に供した. PCA のローディング値を仮説検定に応用す ることで統計的に意味のあるピークを絞りこんだ. この手法は 2014 年に山本らが報告した手 法であり, ローディング値そのものに閾値を設定する場合と比べ, 信頼区間に基づいた議論 が可能になるという利点がある. (Yamamoto H., 2014)

上述のピーク選択と保持指標予測モデルの組み合わせによって, GC/MSによる生薬セン キュウの親水性代謝物測定データ中の未知代謝物を同定した. 植物サンプルには多くの二 次代謝物が含まれており, 効率的な代謝物の同定法が必要とされているためである.

センキュウは婦人病や冷え性に処方される漢方薬であり、その使用量は年々増加している. センキュウの供給の多くは中国からの輸入によって賄われている. 日本では Cnidium officinale Makino が、中国では Ligusticum chuanxiong Hort. がセンキュウとして扱われているため、これらの明確な区別が重要な課題となっている. センキュウは熟練した技術者の官能試験で分類されている. しかし技術者の育成は長期に渡ることから、定量マーカーによる識別などのより簡便な分類法の確立が望まれている. (Kobayashi S., 2012) 本研究では、品種間の区別に有用なバイオマーカー候補となりうる 3 種類の未同定ピークに対して、由来と

なった代謝物を同定した.これらの代謝物はセンキュウにおいて新規に報告されたものである.

さらに生薬センキュウ中に含まれる既知の代謝物群を用いて,異なる保持指標の許容範 囲による誤同定率の変動を調べ,第二章の保持指標予測モデルの有用性について評価した.

3.2 実験材料および実験方法

3.2.1. 試薬およびカラム

サンプルは生薬センキュウを用いた. 全てのセンキュウは, 栃本天海堂 (Osaka, Japan) より入手した. センキュウは 6 種類あり, 種, 産地の有無, 加工法がそれぞれ異なっている. (表 3-1)

高速液体クロマトグラフ用メタノール,クロロホルムおよび,蒸留水 (Wako Pure Chemical Industries, Ltd., Osaka, Japan) を抽出溶媒として使用した. 誘導体化において, Methoxyamine hydrochloride (Sigma, WO, US), リビトール, ピリジン (Wako, Osaka, Japan), N-methyl-N- (trimethylsilyl). trifluoroacetamide (GL sciences Inc., Tokyo, Japan) を 使用した.

GC/MS に使用した高純度ヘリウムは NERIKI Co. Ltd. (Amagasaki, Japan) から購入した. カラムは CP-SIL 8 CB low bleed/MS (30 m × 0.25 mm i.d., Agilent Technologies Japan) を使用した.

代謝物同定に使用した標準品は表 3-2 に記載した. Butane-1,2,3-triol は Tokyo Chemical Industry Co., Ltd. (Tokyo, Japan) から, Diethylene glycol, 3-Deoxyglucosone, Lactitol, Melibiose は Wako Pure Chemical Industries, Ltd.から, 3,6-Anhydro-D-galactose, Palatintol は Funakoshi Co., Ltd. (Tokyo, Japan) から, Lactose は SIGMA からそれぞれ購入した.

-22-

表 3-1 本実験で用いたセンキュウ

種	原産国/県または州	湯通し	PCA (図 3-2)の記号		
Cnidium officinale Makino	日本/北海道	あり	•		
Cnidium officinale Makino	日本/北海道	なし	\bigcirc		
Cnidium officinale Makino	中国/四川彭州	あり			
Cnidium officinale Makino	中国/四川彭州	なし	\bigtriangleup		
Ligusticum chuanxiong Hort.	中国/四川彭州	あり	•		
Ligusticum chuanxiong Hort.	中国/四川彭州	なし			
※湯通しは収穫後に行われる加工の一種					

3.2.2. 試料調製方法

サンプルおよび QC サンプルは, 乾燥試料 15 mg を 2.0 mL チューブに秤量した. ジルコ ニアボール (直径 5 mm) を 1 つ, 2.0 mL チューブ内に入れ, 液体窒素で凍結させた. Ball Mill (MM 301, Verder Scientific Co. Ltd., Haan, German) を用いて, 20 Hz, 3 分の条件で粉 末状に破砕した. その後, メタノール-クロロホルム-蒸留水 (5:2:2, v/v/v) の混合溶液を 1 mL を 2 mL チューブに加えた. Ball Millを用いて, 20 Hz, 5 分の条件で疎水性および親水 性代謝物の分液抽出を行った. 遠心分離機器 (Eppendorf model 5415R, Eppendorf Co., Ltd., Tokyo, Japan) を用いて, 16000 g, 4°C, 3 分の条件で遠心分離を行った.

上澄み液 800 µLを1.5 mL チューブに回収した. 上澄みに 400 µL の蒸留水を加えて 10 秒間, ボルテックスで撹拌した. その後, 同条件で遠心分離を行い, 抽出液を上層と下層に 分離させた. 上層に親水性代謝物が抽出され, 下層には疎水性代謝物が抽出される. 本実 験では親水性代謝物を分析するため, 上層を分注した. 上層を 400 µL 分注し, 新しい 1.5 mL チューブに回収した. 回収した溶液を, 遠心濃縮機器 (Taitec model VC-96R, Taitec Co., Ltd., Saitama, Japan) を用いて 2 時間遠心濃縮した.

遠心濃縮後に、QCサンプルの作成ならびに希釈を行った.本論文におけるQCサンプル は全 6 種類のサンプルを等量ずつ混合したサンプルと定義している.QC サンプル用に、1 種類のサンプルあたり6 つチューブを用意し、合計 36 種を混合した.遠心濃縮後の抽出液 をすべて 15 mL のチューブに分注した. 回収した溶液の体積から QC サンプルの濃度を計算し, 10 mg, 20 mg, …60 mg 等量の溶質が含まれるように溶液を分注することで, 希釈系列を作成した. なお, 空のチューブを 0 mg とした.

その後, サンプルと QC サンプルの希釈系列を 18 時間凍結乾燥させた. (Taitec model VD-800F, Taitec Co., Ltd., Saitama, Japan)

乾燥後の抽出物に Methoxyamine hydroxychloride (20 mg/mL Pyridine) 100 µlを加えて 再溶解させた後,振盪機 (Thermomixer Comfort, Eppendorf Co., Ltd., Hambrug Germany)を用いて, 1200 rpm, 30°C, 90 min の条件で振盪することでのオキシム化を行っ た. 続いて N-methyl-N- (trimethylsilyl) trifluoroacetamide (MSTFA)を50 µL 加え, 1200 rpm, 37°C, 30 min の条件で振盪機を用いて振盪することで TMS 化を行なった. 反応後, 全量を GC-MS 用バイアル瓶に分注した.

3.2.3. 分析条件

GC-MS は GCMS-QP2010 Ultra (Shimadzu Co., Kyoto, Japan) を用いた. GCMS-QP2010 Ultra ではオートサンプラーAOC-20s (Shimadzu Co., Kyoto, Japan), オート インジェクターAOC-20i (Shimadzu Co., Kyoto, Japan) を用いた. GC 及び MS の制御, デ ータ取得にはソフトウェア GC MS solution Ver.2.7 (Shimadzu Co., Kyoto, Japan) を用いた.

GC の設定を以下に示す. キャリアガスには高純度ヘリウム (NERIKI Co. Ltd., Amagasaki, Japan) を用い, カラム流量が 1.12 mL/min となるようにした. インジェクターを 230°C, カラムオーブンを 80°C, GC と MS をつなぐトランスファーラインを 250°C に設定した. サンプルは1 µL をインジェクションし, スプリット比が 25:1 の条件で分析した. 分析開始 後はカラムオーブンを 80°C で 2 分間保持させ, その後 15°C /min で 330°C まで昇温させ, そして 330°C で 6 分間保持させた.

オートサンプラーの設定を以下に示す. QC サンプルの希釈系列は1つのバイアルから4回, 測定を行い. サンプルは1つのバイアルから1回だけ測定した.

MSの設定を以下に示す.イオンソースを200°C,フィラメント電圧を70 eV とし, EI 法を用 いイオン化を行った.スキャン速度を20 scan/sec,検出器の電圧を8.1 kV,質量範囲を*m/z* 85-500 に設定した.データ取得はインジェクションより3.5 分後から24 分後まで行った.た だし,15.6 分後から16.5 分後,19.7 分後から19.9 分は糖類のピークが過大に溶出するため, 機器への負担を軽減するためにフィラメントの電圧を切り,データを取得しなかった.

3.2.4.1 データ行列の作成

NetCDF 形式で測定した GC/MS のデータを GC MS solution Ver.2.7 から出力した. ABF file converter (http://www.reifycs.com/AbfConverter/index.html)で, ABF (analysis base framework) 形式のデータに変換した. MS-DIAL (ver2.48) のソフトウェアとMSPフォーマットの Osaka Univ. DB ライブラリを RIKEN PRIMe からダウンロードし, データ解析に使用した. (Tsugawa H., 2015, Lai Z., 2018) 以下, MS-DIAL のパラメータを列挙する. Smoothing level :3 minimum peak height: 2000, peak width 20, その他のパラメータは Tutorial に記載されたデフォルト値を使用した. 研究室内のライブラリを用いた代謝物同定では, total score of dot product, reverse dot product, existence percentage of fragment ions を 2:2:1 の重みづけ に設定した. その後, 目視で代謝物の同定結果を確認し, 127 個の代謝物を同定した.

3.2.4.2 ピークの選択

アライメントによって得られた 1975 個のピークに対し,保持指標順にピークの ID 番号を設定した.その後これらの RSD (relative standard deviation) をそれぞれ計算した. このうち,QC サンプル 20-50 mg,もしくはQC サンプル 30-60 mgの検量線の範囲で 直線性を示さないピークは除外した.また,RSD 値の平均値がQC サンプル 20-60 mg の間で 20 を超えたピークも除外した.この選択によって 330 個の未同定ピークを含 む,457 ピークが選択された.図 3-1 に選択されたピークと除外されたピークの例を 示し,付録-3 に選択されたピークの情報を記載した.



図 3-1. ピーク選択の例

左は代謝物由来として選択されたピーク,右 はノイズとして除外されたピーク

3.2.4.3 PCA

代謝物テーブルに UV 法による前処理を行った後, PCA を行った. PC1 (principle component1) と PC2 に対し有意に変動した変数を,山本らが提案した手法 (Yamamoto H., 2014) に基づいて選択した. (図 3-2) これらの統計処理は, RIKEN PRIMe で提供 されているオープンソースのエクセルマクロのプログラム (Statistics in Microsoft Excel, URL は引用に記載) によって行われた.

3.2.4.3 代謝物の同定とバリデーション

PCA の結果から統計的に有意であるとみなされた未同定代謝物の定性を行った. ま ず, MS-FINDER によって EI スペクトルによる代謝物同定を実施した. (Tsugawa H., 2016, Lai Z., 2018) MS-FINDER には MassBank と MoNA が既定のライブラリとして設定されてい る. これに加え, Wiley (第10版) を MSP 形式に変換し, "User-defined MSP database"として 設定した. このうち, SMILES (simplified molecular input line entry system) が登録されてい ない EI スペクトルは除外したため,この操作によって計 668,231 個のスペクトルが追加で登録された.

これらの EI スペクトルのライブラリから,同定したいピークの EI スペクトルと 80%以上の類似性を示す化合物を探索した.類似した EI スペクトルを持つ化合物 (以 下,候補化合物と呼ぶ)の保持指標を第二章の式に基づいて計算した.計算された予 測値と未同定ピークの保持指標値の実測値とを比較し,残差が100以上になったものを 除外した.以上の工程によって,未同定代謝物の候補化合物を絞り込んだ.(図 3-3)

候補化合物の純粋な標品による添加実験によって,同定結果のバリデーションを取った. まず,水を溶媒とした候補化合物の標準溶液(各 1 mmol)を調製した."3.2.2. 試料調製 方法"の手順に従って,センキュウの親水性代謝物の抽出サンプルを作成した.遠心濃縮 後のサンプルに標準溶液をそれぞれ 1,2,3 µl ずつ添加した.標準物質がスパイクされたサ ンプルに凍結乾燥,誘導体化の手順を行い,"3.2.3. 分析条件"の設定に従って GC/MS で 測定をした.標準品に由来するピークの保持指標を目視で確認した.

3.2.4.4 保持指標と誤同定率の計算

"3.2.4.1 データ行列の作成"で同定された 127 種のピークを用いて,代謝物同定の際 に設定される保持指標の許容値が,同定結果に与える影響を検証した. MassBank と MoNA に登録されている 13,570 個のスペクトルを搭載した MS-DIAL を用いて,127 種の代謝物を再同定した.保持指標の許容値を 10,20,30…,300 の範囲で変動させ, それぞれの結果における誤同定率を計算した.(図 3-4) 誤同定率は誤同定した代謝物 を全代謝物数 (127 種) で除算することで求めた.

3.3. 結果と考察

3.3.1. PCA の結果

PCA のスコアプロット, ローディングプロットを図 3-1 に記載した. 横軸である PC1 によって 日本種 (*Cnidium officinale Makino*) と中国種 (*Ligusticum chuanxiong* Hort) が分類され, 縦軸の PC2 によって中国で栽培された日本種の湯通しありがその他のサンプルから分離し た. (図 3-2 左) ローディング値に行った t-test の結果, データ行列にふくまれていた 457 個 のピーク (説明変数) のうち, 未同定ピーク 170 個を含む 245 個のピークが選択された. (図 3-2 右)





スコアプロットの各記号の意味は表 3-1 を参照.
ローディングプロットは説明変数(ピーク)がプロットされている.
黒;Bonferoni 補正 p-value < 0.05
白;Bonferoni 補正 p-value ≥ 0.05 となった説明変数

3.3.2. 未同定代謝物の定性

EIスペクトルの検索により, Alignment ID 301 のピークでは46 個の, Alignment ID 761 の ピークでは5 つの, そして Alignment ID 1538 のピークでは64 個の候補化合物が見つかっ た. その後の保持指標予測によって, 候補化合物の数はそれぞれ,5 つ,3 つ,7 つにまで 減少した.標準品によるバリデーションによって最終的にこれらは, butane-1,2,3-triol, 3-deoxyglucosone, palatinitol であると定性された. (図 3-3) Palatinitol にはヒトの腸内細菌の 増加に関わるという生理活性が報告されており, (van Weerden E. J., 1993) センキュウの薬 効に関与している可能性がある.

表 3-1 に測定した標準品の一覧を記載した. これらは保持指標予測によって絞り込まれた 候補化合物である. ただし, Alignment ID301の候補化合物である, phosphorochloridic acid dipropyl ester と 1-deoxyerythritol, 2-ethyl-3-Hydroxybutyrate, および Alignment ID761の 3,6-Anhydro-D-glucose 計 4 つの標準品は利用できなかったため, この表には記載していな い.

表 3-2 に記載された合計 13 個の候補化合物のうち, 予測値と実測値の残差が 50 以下と なった代謝物は 8 つとなった. 保持指標モデルはトレーニングセット以外の代謝物の予測へ 適用可能であることが示された. 3-deoxyglucosone や 3,6-anhydro-d-galactose, lactitol, lactose は,大きな誤差を示したが 95%の信頼領域には含まれている. 3-deoxyglucosone や 3,6-anhydro-d-galactose は類似した代謝物,つまりジカボニル化合物,または無水糖 は, トレーニングセットに含まれていなかったため,保持指標の予測精度が下がったと考えられ る.

Lactitol とlactoseは, *alpha*-1-4または*beta*-1-4結合をもつ二糖アルコール,または二糖である.同じく二糖であるが残差の値が小さくなった,lactoseとpalatinntolは*alpha*-1-6結合を, Gentiobioseはbeta-1-6結合をもつ.この結果から,立体構造も割合保持指標に影響を与えることが示唆された.

Compound	Actual RI	Predicted RI	Difference
Butane-1,2,3-triol	1283	1320	37
Diethylene glycol	1237	1225	12
3-Deoxyglucosone_1	1734	1669	65
3-Deoxyglucosone _2	1737	1669	68
3,6-Anhydro-D-galactose	1683	1736	53
Gentiobiose_1	2781	2811	30
Gentiobiose_2	2809	2811	2
Lactitol	2731	2848	117
Lactose_1	2656	2782	126
Lactose_2	2674	2782	108
Melibiose_1	2829	2839	10
Melibiose_2	2861	2839	22
Palatintol_1	2875	2894	19
Palatintol_2	2885	2894	9

表 3-2, 候補化合物の保持指標値の予測値と実測値

代謝物名_1, _2 となっているものは、オキシム化の違い

(E体,Z体)により二種類の誘導体化生成物をもつ代謝物である.






図 3-4 保持指標値の許容値と誤同定率のプロット

127 種類の生薬センキュウ中の既知代謝物が再同定された結果を示す. 赤のプロット:代謝物の立体構造が一致しなかった代謝物の割合 青のプロット:代謝物の骨格が一致しなかった代謝物の割合

3.3.3. 保持指標の許容値と誤同定率

"2.3.2. 保持指標の系統誤差"の結果では,異なる研究室間で保持指標値が示した系 統誤差は約50,保持指標予測モデルが示した残差は約100となった.そこで本節では 許容値を±50に設定した際の誤同定率を,保持指標ライブラリを使用する際に生じう る誤同定率とみなした.同じく±100の際の誤同定率を,予測モデルの利用によって 生じうる誤同定率とみなした.

図 3-4 では、立体異性体を区別した際の誤同定率(赤いプロット)と立体異性体を 区別しない場合の誤同定率(青いプロット)を示した.立体異性体を区別する場合は、 InChIKey が完全一致した代謝物のみを正しく同定された代謝物であるとみなした.保 持指標の許容値を±50 に設定すると 16.5%、±100 では 37.0%となった. 一方、立体 異性体を区別しない場合は、InChIKey の初めの 14 文字で定義される分子骨格のみの 一致に条件を緩和した.その結果、誤同定率はそれぞれ 9.4%、14.1%にまで低下した. さらに、保持指標を使わずに代謝物同定を行った場合は 19%となった. 127種のピー クを同定するために必要となった計算の回数は、保持指標値の許容値±50、±100 に 対し、それぞれ約 100000回、約 200000回となった.また、保持指標値を使用せずに 代謝物同定を行った場合は、約 1700000回となった.

以上の結果から,代謝物の立体構造まで同定するを行うためにはより精度の高い保 持指標の値がが必須となる.しかし,"3.3.2.未同定代謝物の定性"で示したように, 類似した EI スペクトルを持つ候補化合物の絞り込みは,分子骨格の識別によっても 十分に達成可能であり,さらに計算のコストを減らしスループットを上げる点におい ても,保持指標値の予測を行うメリットがあると結論付けた.

第四章 総括と展望

本研究では、本論文では誘導体化された代謝物を用いて、保持指標予測モデルの構築 することに成功した.構築したモデルは先行研究で報告されている保持指標予測モデルと 比べ、トレーニングセットに登録されている代謝物の数は小規模であるが、モデルに使用す る代謝物の誘導体化構造を精査することや分析条件の統一などによって、予測できる化合 物の種類の幅広さと正確さの両立に成功した.

類似した代謝物がトレーニングセットに含まれていない,もしくは少ない場合,保持指標予 測値は実測値との誤差が大きくなる傾向が見られた.記述因子と保持指標のプロットにより 予測したい代謝物が外れ値でないかを調査することで,この問題はある程度回避できると考 えられる.

これらに加え、本研究では全種類の測定サンプルを混合した QC サンプルの希釈系列を 基準として、サンプル中の代謝物に由来する信号の選択を実サンプルで立証した.この操 作によって GC/MS のデータに含まれる 1975 個のピークから、330 個の未同定代謝物由来 のピークを含む、457 個が選択された.これらのピークは QC サンプルによって信号強度 値の直線性が担保されているため、多変量解析に適した変数である.この手法には分析法 やサンプルの種類を問わずに適用できるという利点があり、汎用性が高い.さらに、多変量 解析の結果に基づいてピークの統計的な重要性を客観的に評価することで、同定すべきピ ークの数を 245 個にまで絞り込んだ.

構築したモデルを未同定代謝物の同定に利用して,先行研究では達成されなかった,異 性体の候補化合物の絞り込みにも成功した.

本手法はインハウスで構築した既存のデータベースから無償で利用可能なソフトのみで, 予測モデルを作成できるという利点も持ち, 今後の GC/ MS メタボロミクスの普及への貢献が 期待される. 本論文は,著者が大阪大学大学院工学研究科生命先端工学専攻において行った研究 をまとめたものです.

本研究を進めるにあたり,終始ご指導ご鞭撻を賜りました大阪大学大学院工学研究科生 命先端工学専攻 福崎英一郎 教授に謹んで感謝の意を表します.また,本研究の遂行な らびに本論文の執筆にあたり,懇切丁寧なご指導を賜りました国立研究開発法人理化学研 究所統合生命医科学研究センター 津川裕司 博士に厚く御礼申し上げます.また,本論 文を精読してくださり有用なご助言を賜りました大阪大学大学院工学研究科生命先端工学 専攻 大政 健史 教授ならびに大阪大学大学院情報科学研究科 松田 史夫 教授に心 より感謝申し上げます.さらに,日頃より研究室生活全般にわたり暖かいご援助を賜りました 大阪大学大学院工学研究科生命先端工学専攻 新間秀一 准教授,馬場健史 元准教授 (現 九州大学生体防御医学研究所 教授),小野比佐好 元助教に深く感謝いたします.

加えて,代謝物の誘導体化について議論をしてくださいましたジーエルサイエンス株式会社 宮川浩美 様に深く感謝の意を表します.

そして,研究活動のみならず,研究室生活全般において大変お世話になりました大阪大 学大学院工学研究科生命先端工学専攻生物資源工学領域の諸先輩方,同輩,後輩なら びにスタッフの皆様に心から感謝の意を表します.

最後に、博士後期課程に進学する機会を与え温かく見守ってくれた父 松尾二郎,母 律子,ならびに親族の方々,そして応援し続けてくれた多くの友人に深い感謝の意を表して 謝辞と致します.

引用文献

Aebersold R., Mann M.: Mass-spectrometric exploration of proteome structure and function. *Nature*, **537**, 347-355 (2016)

Allen F., Pon, A., Greiner R., Wishart. D.: Computational prediction of electron ionization mass spectra to assist in GC/MS compound identification. *Analytical Chemistry*, **88**, 7689-7697 (2016)

Beger R. D., Colatsky T.: Metabolomics data and the biomarker qualification process. *Metabolomics*, **8**, 2–7 (2012)

Cevallos-Cevallos J. M., Reyes-De-Corcuera J. I., Etxeberria E., Danyluk M. D., Rodrick G. E.: Metabolomic analysis in food science: a review. *Trends in Food Science & Technology*, **20**, 557-566 (2009)

Fatemi M. H., Elyasi M.: Prediction of gas chromatographic retention indices of some amino acids and carboxylic acids from their structural descriptors. *Journal of separation science*, **22**, 3216–3220 (2011)

Fiehn O., Kopka J., Dörmann P., Altmann T., Trethewey R. N., Willmitzer L.: Metabolite profiling for plant functional genomics. *Nature Biotechnology*, **18**, 1157–1161 (2000)

Fiehn O.: Metabolomics – the link between genotypes and phenotypes. *Plant Molecular Biology*, **48**, 155 – 171 (2002)

Fuhrer T., Zamboni N.: High-throughput discovery metabolomics. *Current Opinion in Biotechnology*, **31**, 73–78 (2015)

Godzien J., Alonso-Herranz V., Barbas C., Armitage E. G.: Controlling the quality of metabolomics data: new strategies to get the best out of the QC sample, *Metabolomics*, **11**, 518-528 (2015)

Greif P.A., Eck S.H., Konstandin N. P., Benet-Page's A., Ksienzyk B., Dufour A., Vetter A.T., Popp H.D., Lorenz-Depiereux B., Meitinger T., Bohlander S.K., Strom T. M.: Identification of recurring tumor-specific somatic mutations in acute myeloid leukemia by transcriptome sequencing. *Leukemia*, **25**, 821–827 (2011)

Hemmateenejad B., Javadnia K., Elyasi M.: Quantitative structure–retention relationship for the Kovats retention indices of a large set of terpenes: A combined data splitting-feature selection strategy. *Analytica Chimica Acta*, **592**, 72-81 (2007)

Hu R. J., Liu H. X., Zhang R. S., Xuea C. X, Yao X. J., Liu M. C., Hu Z. D., Fan B. T.: QSPR prediction of GC retention indices for nitrogen-containing polycyclic aromatic compounds from heuristically computed molecular descriptors. *Talanta*, **68**, 31–39 (2005)

Jones D. P., Park Y., Ziegler T. R.: Nutritional metabolomics: Progress in addressing complexity in diet and health. *Annual Review of Nutrition*, **32**, 183–202 (2012)

Kaliszan R.: Quantitative structure-retention relationships. *Analytical Chemistry*, **64**, 619–631 (1992)

Kanani H., Chrysanthopoulos P. K., Klapa M. I.: Standardizing GC–MS metabolomics. *Journal of Chromatography B*, 871, 191–201 (2008)

Kind T., Wohlgemuth G., Lee D. Y., Lu Y., Palazoglu M., Shahbaz S., Fiehn O.: FiehnLib: Mass spectral and retention index libraries for metabolomics based on quadrupole and time-of-flight gas chromatography/mass spectrometry. *Analytical Chemistry*, **81**, 10038– 10048 (2009)

Kobayashi S., Nagasawa S., Yamamoto Y., Donghyo K., Bamba T., Fukusaki E.: Metabolic profiling and identification of the genetic varieties and agricultural origin of *Cnidium officinale* and *Ligusticum chuanxiong*. *Journal of Bioscience and Bioengineering*, **114**, 86-91 (2012)

Koboldt D. C., Steinberg K. M., Larson D. E., Wilson R. K., Mardis E. R.: The Next-generation sequencing revolution and its impact on genomics. *Cell*, **155**, 27-38 (2013)

Kovats E.: Gas-chromatographische Charakterisierung organischer Verbindungen. Teil 1: Retentions indices aliphatischer Halogenide, Alkohole, Aldehyde und Ketone. *Helvetica Chimica Acta*, **41**, 1915–1932 (1958)

Kumari S., Stevens D., Kind T., Denkert C., Fiehn O.: Applying *in-Silico* retention index and mass spectra matching for identification of unknown metabolites in accurate mass GC-TOF mass spectrometry. *Analytical Chemistry*, **83**, 5895–5902 (2011) Lai Z., Tsugawa H., Wohlgemuth G., Mehta S., Mueller M., Zheng Y., Ogiwara A., Meissen J., Showalter M., Takeuchi K., Kind T., Beal P., Arita M., Fiehn O.: Identifying metabolites by integrating metabolome databases with mass spectrometry cheminformatics. *Nature methods*, **15**, 53-56 (2018)

Lee M. L., Vassllaros D. L.: Retention indices for programmed-temperature capillary-column gas chromatography of polycyclic aromatic hydrocarbons. *Analytical Chemistry*, **51**, 768–773 (1979)

Nicholson J. K., Lindon J. C., Holmes E.: "Metabonomics": understanding the metabolic responses of living systems to pathophysiological stimuli via multivariate statistical analysis of biological NMR spectroscopic data. *Xenobiotica*. **29**, 1181–1189 (1999)

Pongsuwan W., Fukusaki E., Bamba T, Yonetani T., Yamahara T., Kobayashi A.: Prediction of Japanese green tea ranking by gas chromatography/mass spectrometry-based hydrophilic metabolite fingerprinting. *Agricultural and food chemistry*, **55**, 231–236 (2007)

Righetti P. G., Sebastiano R., Citterio A.: Capillary electrophoresis and isoelectric focusing in peptide and protein analysis. *Proteomics*, **13**, 325–340 (2013)

Rivera C. M., Ren B.: Mapping human epigenomes. Cell, 155, 39-55 (2013)

Smart K. F., Aggio R. B. M., Houtte J. R. V., Villas-Bôas S. G.: Analytical platform for metabolome analysis of microbial cells using methyl chloroformate derivatization followed by gas chromatography–mass spectrometry. *Nature Protocols*, **5**, 1709-1729 (2010)

Smolinska A., Blancheta L., Buydens L. M. C., Wijmenga S. S.: NMR and pattern recognition methods in metabolomics: From data acquisition to biomarker discovery: A review. *Analytica Chimica Acta*, **750**, 82–97 (2012)

Spratlin J. L., Serkova N. J., Eckhardt S. G.: Clinical Applications of Metabolomics in Oncology: A Review. *Clinical Cancer Research*, **15**, 431-440 (2009)

Stein S. E., Babushok V. I., Brown R. L., Linstrom P. J.: Estimation of Kova'ts retention indices using group contributions. *Journal of Chemical Information and Modeling*, **47**, 975-980 (2007)

Strehmel N., Hummel J., Erban A., Strassburg K., Kopka J.: Retention index thresholds for compound matching in GC–MS metabolite profiling. *Journal of Chromatography B*, **871**, 182-190 (2008)

Styczynski M. P., Moxley J. F., Tong L. V., Walther J. L., Jensen K. L., Stephanopoulos G. N.: Systematic identification of conserved metabolites in GC/MS data for metabolomics and biomarker discovery. *Analytical Chemistry*, **79**, 966-973 (2007)

Sumner L. W., Amberg A., Barrett D., Beale M. H., Berger R., Daykin C. A., Fan T. W.
M., Fiehn O., Goodacre R., Griffin J. L., Hankemeier T., Hardy N., Harnly J., Higashi
R., Kopka J., Lane A. N., Lindon J. C., Marriott P., Nicholls A. W., Reily M. D., Thaden
J. J., Viant M. R.: Proposed minimum reporting standards for chemical analysis.

Tsugawa H., Tsujimoto Y., Arita M., Bamba T., Fukusaki E.: GC/MS based metabolomics:

development of a data mining system for metabolite identification by using soft independent modeling of class analogy (SIMCA). *BMC Bioinformatics*, **12**, 131–143 (2011)

Tsugawa H., Cajka T., Kind T., Ma Y., Higgins B., Ikeda K., Kanazawa M., VanderGheynst J., Fiehn O., Arita M.: MS-DIAL: data-independent MS/MS deconvolution for comprehensive metabolome analysis. *Nature methods*, **12**, 523-531 (2015)

Tsugawa H., Kind T., Nakabayashi R., Yukihira D., Tanaka W., Cajka T., Saito K., Fiehn
O., Arita M.: Hydrogen rearrangement rules: computational MS/MS fragmentation and structure elucidation using MS-FINDER software. *Analytical Chemistry*, 88, 7946–7958 (2016)

van Weerden E. J., Huisman J.: The digestion process of the sugar alcohol isomalt in the intestinal tract of the pig: 1. Studies with administration of isomalt in the feed. *British Journal of Nutrition*, **69**, 455-466 (1993)

Wagner C., Sefkow M., Kopka J.: Construction and application of a mass spectral and retention time index database generated from plant GC/EI-TOF-MS metabolite profiles. *Phytochemistry*, **62**, 887–900 (2003)

Wallace M., Cottell E., Gibney M. J., McAuliffe F. M., Wingfield M., Brennan L.: An investigation into the relationship between the metabolic profile of follicular fluid, oocyte developmental potential, and implantation outcome. *Fertility and Sterility*, **97**, 1078-1084 (2012)

Wishart D. S.: Current Progress in computational metabolomics. Briefings in bioinformatics,

Wrzesinska A. R., Bihan M. C. L., Andersen M. T., Roepstorff P.: 2D gels still have a niche in proteomics. *Journal of Proteomics*, 88, 4-13 (2013)

Xie G., Zhang S., Zheng X., Jia W.: Metabolomics approaches for characterizing metabolic interactions between host and its commensal microbes. *Electrophoresis*, **34**, 2787–2798, (2013)

Yamamoto H., Fujimori T, Sato H., Ishikawa G., Kami K., Ohashi Y.: Statistical hypothesis testing of factor loading in principal component analysis and its application to metabolite set enrichment analysis. *BMC Bioinformatics*, **15**, 51 (2014)

Yap C. W.: PaDEL-Descriptor: an open source software to calculate molecular descriptors and fingerprints. *Journal of Computational Chemistry*, **32**, 1466–1474 (2011)

Wang Z., Gerstein M., Snyder M.: RNA-Seq: a revolutionary tool for transcriptomics. *Nature Reviews Genetics*, **10**, 57-63 (2009)

第二章,第三章で参照した URL:

• Aloutput2

http://prime.psc.riken.jp/Metabolomics_Software/AIoutput/index.html 2018 年 1 月 9 日

• Pubchem download service

https://pubchem.ncbi.nlm.nih.gov/

2018年1月9日

• MetaboloDerivatizer

http://prime.psc.riken.jp/Metabolomics_Software/MetaboloDerivatizer/index.html, 2018 年 1 月 9 日

• Chemaxson Jchem Cxcalc http://www.chemaxon.com,

2018年1月9日

• Statistics in Microsoft Excel - RIKEN PRIMe

http://prime.psc.riken.jp/Metabolomics_Software/StatisticalAnalysisOnMicrosoftExcel/index.

<u>html</u>

2018年1月9日

本学位論文に関与する論文

 Teruko Matsuo, Hiroshi Tsugawa, Hiromi Miyagawa, Eiichiro Fukusaki: Integrated strategy for unknown EI–MS identification using quality control calibration curve, multivariate analysis, EI–MS spectral database, and retention index prediction. *Analytical Chemistry*, 89 (12), 6766–6773 (2017)

その他の原著論文

なし

総説等

なし

学会発表

国際会議

- O<u>Teruko Matsuo,</u> Yumiko Nagasawa, Takeshi Bamba, Eiichiro Fukusaki
 "A new integration method for metabolome analysis"
 10th Annual International Conference of Metabolomics society, Tsuruoka, June, 2014
- O<u>Teruko Matsuo</u>, Hiromi Miyagawa, Hiroshi Tsugawa, Eiichiro Fukusaki
 "Development of Retention Index Prediction Model for Annotation of Unknown Compounds for GC/MS-based Untargeted Metabolomics"
 Aachen-Osaka Joint Symposium "Biological and Chemical Methods for Selective Catalysis" (日独共同大学院プログラム), Aachen, September, 2016
- O<u>Teruko Matsuo</u>, Hiroshi Tsugawa, Hiromi Miyagawa, Eiichiro Fukusaki
 "Integrated strategy for unknown Electron impact Mass spectrometry annotation using Quality Control calibration curve, multivariate analysis, EI-MS spectra database, and retention index prediction"
 BIOTECHNOLOGY AND CHEMISTRY FOR GREEN GROWTH (日独共同大学院プ ログラム), Awaji-shima, March, 2017
- 4) OBenedikt Heyman, <u>Teruko Matsuo</u>, Lars Regestein, Eiichiro Fukusaki, Jochen Buechs
 "Biotechnological 2,3-Butanediol Production A step Towards Sustainable Fuel
 Production"

BIOTECHNOLOGY AND CHEMISTRY FOR GREEN GROWTH (日独共同大学院プ

ログラム), Awaji-shima, March, 2017

国内会議

- O<u>松尾 晃子</u>,長澤 由美子,馬場 健史,福崎 英一郎
 "定量的メタボローム解析に資する新規相対定量法"
 第 65 回日本生物工学会,広島,2013 年 9 月
- 2) <u>〇松尾 晃子</u>,宮川 浩美,津川 裕司,新間 秀一,馬場 健史,福崎 英一郎
 "ノンターゲットメタボローム解析に資する保持指標予測モデルの未同定化合物構造推定への応用"
 第9回メタボロームシンポジウム,三島,2015年10月
 本発表はポスター発表演題からロ頭発表演題に採択された.
- 3) ○宮川 浩美, <u>松尾 晃子</u>, 津川 裕司, 安藤 晶, 古野 正浩, 馬場 健史, 福崎 英一郎
 "メタボロミクスの定量的議論に資する GC/MS ライブラリの構築"

第9回メタボロームシンポジウム,三島,2015年10月

付録

1-選択された記述因子

保持指標と 0.8 以上の相関を示した計 80 種の記述因子を以下に示す. (Yap C. W., 2011)

AMR	ETA_Eta_B_RC	MLFER_L	SP-6
apol	ETA_Eta_F	MLogP	SP-7
ATSc1	ETA_Eta_L	MW	SPC-4
ATSm1	ETA_Eta_R	nAtom	SPC-5
ATSm2	ETA_Eta_R_L	nBonds	SPC-6
ATSm3	Exact mass	nBonds2	topoDiameter
ATSm4	fragC	nBondsS	topoRadius
ATSm5	GraphFP195	nBondsS3	VABC
ATSp1	GraphFP266	nC	VAdjMat
ATSp2	GraphFP31	nH	VP-0
ATSp3	GraphFP348	nHBa	VP-1
ATSp4	GraphFP44	nHBAcc_Lipinski	VP-2
ATSp5	GraphFP48	nHBAcc2	VP-5
bpol	GraphFP759	nHeavyAtom	VP-6
CrippenMR	Kier1	SP-0	WPATH
ECCEN	Kier2	SP-1	WPOL
ETA_Alpha	logP	SP-2	WTPT-1
ETA_Beta	McGowan_Volume	SP-3	WTPT-3
ETA_Beta_s	MDEC-12	SP-4	XLogP
ETA_Eta	MDEC-33	SP-5	Zagreb

2-モデル構築に使用された代謝物

Q _{ata}	Nama	Retention	n Set1>Set2		Set2>Set1	
Sets	Name	index	Predicted	Difference	Predicted	Difference
Set 1	Propyleneglycol	989.3	1064.9	75.5	1088.3	99.0
Set 2	n-Propylamine	1017.5	1038.6	21.2	1060.1	42.7
Set 1	2-Hydroxypyridine	1033.6	1291.3	257.7	1307.9	274.3
Set 2	Pyruvic acid	1041.9	1375.9	334.0	1380.2	338.3
Set 1	Lactic acid	1050.2	1095.6	45.4	1120.1	70.0
Set 2	Glycolic acid	1067.5	1044.8	22.7	1073.6	6.1
Set 1	n-Methylethanolamine	1073.6	1073.4	0.1	1101.6	28.0
Set 2	Valine_1	1085.6	1115.3	29.6	1125.2	39.6
Set 1	Isobutylamine	1090.3	1082.2	8.1	1099.7	9.3
Set 2	Alanine_1	1094.2	1117.0	22.8	1145.1	50.9
Set 1	n-Butylamine	1100.9	1134.9	34.0	1152.7	51.8
Set 2	Ketovaline_1	1101.6	1148.0	46.3	1172.3	70.6
Set 1	2-Hydroxybutyric acid	1118.3	1164.5	46.2	1181.1	62.7
Set 2	Oxalic acid	1124.6	1063.1	61.5	1093.4	31.2
Set 1	Ketovaline_2	1127.5	1148.0	20.5	1172.3	44.8
Set 2	Sarcosine	1130.4	1105.4	25.0	1134.6	4.2
Set 1	2-Aminoisobutyric acid	1136.4	1168.4	32.0	1191.6	55.2
Set 2	3-Hydroxybutyric acid	1152.8	1198.8	46.0	1219.5	66.7
Set 1	4-Hydroxypyridine	1155.2	1286.0	130.7	1302.8	147.5
Set 2	2-Aminobutyric acid	1164.6	1185.9	21.3	1205.9	41.4
Set 1	Ketoisoleucine_1	1172.7	1216.8	44.1	1233.1	60.4
Set 2	Ketoisoleucine_2	1192.3	1216.8	24.5	1233.1	40.8
Set 1	Glyceraldehyde_1	1195.5	1186.8	8.7	1212.1	16.6
Set 2	Malonic acid	1197.1	1180.2	17.0	1206.2	9.0
Set 1	Norleucine_1	1197.7	1264.1	66.5	1270.7	73.0
Set 2	1-Aminocyclopropane-1-c	1202.5	1221.4	18.9	1239.0	36.5
	arboxylic acid					
Set 1	3-Hydroxyisovaleric acid	1203.1	1248.9	45.8	1264.7	61.6
Set 2	Valine_2	1208.0	1229.5	21.5	1245.5	37.5
Set 1	Glyceraldehyde_2	1209.8	1186.8	23.0	1212.1	2.4

試薬のメーカーは Aloutput2 (Tsugawa H, 2011) に準ずる.

Set 2	Norvaline	1232.7	1254.7	22.0	1266.7	34.0
Set 1	Urea	1238.3	1168.2	70.0	1181.4	56.8
Set 2	Dihydroxyacetone	1243.5	1209.6	33.9	1231.3	12.2
Set 1	Serine_1	1251.9	1241.0	10.9	1252.7	0.8
Set 2	Oxamic acid	1254.2	1193.8	60.4	1210.2	44.0
Set 1	n-Acetyl alanine	1258.6	1257.5	1.1	1259.1	0.6
Set 2	2-Aminoethanol	1259.8	1206.4	53.4	1226.6	33.2
Set 1	Octanoic acid (C8)	1260.9	1371.4	110.5	1382.7	121.8
Set 2	Glycerol	1262.0	1270.7	8.7	1280.6	18.6
Set 1	Phosphoric acid	1262.8	1308.6	45.8	1301.1	38.3
Set 2	Leucine_2	1263.3	1298.2	34.9	1306.2	42.9
Set 1	Ethylmalonic acid	1272.9	1296.7	23.8	1310.5	37.6
Set 2	3-Hydroxypyruvic acid_1	1277.5	1238.2	39.3	1261.1	16.5
Set 1	Isoleucine	1284.7	1298.3	13.6	1306.3	21.6
Set 2	Threonine_1	1287.2	1288.0	0.8	1295.6	8.4
Set 1	Proline	1293.8	1275.2	18.7	1286.6	7.2
Set 2	Maleic acid	1297.5	1319.5	21.9	1344.8	47.2
Set 1	Nicotinic acid	1297.8	1403.7	105.9	1417.0	119.2
Set 2	Glycine	1299.5	1238.4	61.1	1259.7	39.8
Set 1	Succinic acid	1307.8	1288.3	19.5	1310.2	2.4
Set 2	Norleucine_2	1310.9	1351.0	40.1	1359.3	48.4
Set 1	Catechol	1315.4	1507.2	191.7	1512.8	197.4
Set 2	Glyceric acid	1318.9	1298.0	20.9	1309.1	9.8
Set 1	Methylsuccinic acid	1319.7	1333.2	13.6	1351.1	31.4
Set 2	3-Hydroxypyruvic acid_2	1324.3	1238.2	86.1	1261.1	63.2
Set 1	Oxamide	1325.5	1316.7	8.8	1320.3	5.2
Set 2	2-Methyl benzoic acid	1326.1	1479.3	153.3	1485.0	159.0
Set 1	Picolinic acid	1326.4	1399.4	73.0	1412.9	86.4
Set 2	Uracil	1332.9	1551.7	218.9	1530.9	198.0
Set 1	Fumaric acid	1344.9	1319.5	25.4	1344.8	0.1
Set 2	Serine_2	1349.5	1328.7	20.8	1342.1	7.4
Set 1	Alanine_2	1354.2	1285.5	68.8	1302.6	51.6
Set 2	Nonanoic acid (C9)	1357.5	1298.5	59.0	1302.9	54.7
Set 1	β -Cyano alanine	1367.3	1316.9	50.4	1330.8	36.5
Set 2	4-Methyl benzoic acid	1369.8	1506.7	137.0	1516.8	147.0
Set 1	n-Acetyl valine	1371.3	1370.0	1.3	1359.5	11.7
Set 2	n-Formyl glycine	1373.2	1243.1	130.1	1255.7	117.5

Set 1	Threonine_2	1373.3	1375.5	2.3	1384.8	11.6
Set 2	3-Pyridylacetic acid	1392.4	1501.3	108.9	1510.9	118.5
Set 1	Thymine	1394.6	1619.9	225.2	1591.3	196.7
Set 2	Glutaric acid	1398.6	1386.9	11.7	1405.0	6.5
Set 1	2-Methylglutaric acid	1408.8	1431.4	22.5	1445.5	36.6
Set 2	3-Methylglutaric acid	1418.7	1430.4	11.7	1444.6	25.9
Set 1	β -Alanine	1422.1	1340.4	81.8	1357.7	64.4
Set 2	Phenoxyacetic acid	1429.5	1518.3	88.8	1531.3	101.8
Set 1	Cinnamyl alcohol	1432.6	1599.0	166.4	1607.5	174.9
Set 2	Prolinamide	1435.0	1366.8	68.2	1378.3	56.7
Set 1	n-Acetyl-leucine	1435.4	1438.7	3.3	1420.3	15.2
Set 2	Homoserine	1438.5	1424.4	14.1	1435.0	3.5
Set 1	Citramalic acid	1461.7	1488.5	26.8	1491.8	30.0
Set 2	Mandelic acid	1470.8	1659.1	188.3	1654.9	184.1
Set 1	Malic acid	1476.8	1434.6	42.3	1442.8	34.1
Set 2	Threitol	1481.7	1523.5	41.8	1515.7	34.0
Set 1	Nicotinamide	1488.8	1521.8	33.1	1522.5	33.7
Set 2	meso-Erythritol	1489.6	1523.5	33.9	1515.7	26.1
Set 1	<i>p</i> -Hydroxybenzaldehyde	1498.2	1572.7	74.5	1596.5	98.3
Set 2	Adipic acid	1500.0	1483.3	16.7	1497.8	2.2
Set 1	4-Hydroxybenzyl alcohol	1501.3	1655.7	154.4	1664.1	162.8
Set 2	Homoserine lactone	1502.4	1068.4	434.0	1076.3	426.1
Set 1	Aspartic acid	1508.2	1457.8	50.4	1469.4	38.8
Set 2	Acetylsalicylic acid	1508.4	1737.7	229.3	1744.3	235.9
Set 1	Methionine	1515.2	1420.9	94.3	1416.3	98.9
Set 2	4-Hydroxyproline	1516.5	1487.8	28.7	1493.9	22.6
Set 1	α-Phenylglycine	1516.7	1680.4	163.7	1679.7	163.0
Set 2	Pyroglutamic acid	1522.3	1417.8	104.5	1413.1	109.2
Set 1	Cytosine	1522.4	1433.6	88.8	1435.0	87.4
Set 2	2-Thiouracil	1523.9	1524.1	0.2	1513.4	10.4
Set 1	4-Aminobutyric acid	1524.4	1438.9	85.5	1452.5	71.9
Set 2	Pyrogallol	1536.4	1715.6	179.2	1710.6	174.1
Set 1	Threonic acid	1543.0	1550.7	7.7	1544.1	1.1
Set 2	5-Methylcytosine	1548.6	1500.8	47.8	1494.5	54.0
Set 1	3-Hydroxybenzoic acid	1565.0	1665.5	100.5	1670.5	105.4
Set 2	2-Isopropylmalic acid	1566.0	1601.1	35.1	1592.3	26.3
Set 1	2-Hydroxyphenylacetic	1566.1	1733.4	167.4	1730.4	164.3

	acid					
Set 2	<i>α</i> -Ketoglutaric acid	1567.8	1475.9	91.8	1492.1	75.7
Set 1	4-Hydroxyphenethyl	1571.0	1751.8	180.9	1756.6	185.6
	alcohol					
Set 2	6-Hydroxynicotinic acid	1572.5	1611.2	38.7	1605.6	33.1
Set 1	Ethionine	1579.0	1517.6	61.3	1509.3	69.6
Set 2	1-Methyluracil	1581.9	1434.7	147.2	1410.4	171.5
Set 1	3-Phenyllactic acid	1582.5	1755.2	172.7	1747.3	164.8
Set 2	3-Amino-2,3-dihydrobenz	1582.7	1562.8	19.9	1569.9	12.8
	oic acid					
Set 1	Phosphoenolpyruvic acid	1587.9	1640.5	52.6	1625.3	37.4
Set 2	3-Hydroxy-3-methylglutari	1592.2	1590.7	1.6	1590.0	2.2
	c acid					
Set 1	Heptanedioic acid	1597.9	1579.6	18.3	1590.4	7.5
Set 2	Hypotaurine	1598.4	1592.7	5.7	1574.4	24.0
Set 1	β -Glutamic acid	1601.4	1553.5	47.9	1562.3	39.1
Set 2	3-Hydroxyphenylacetic	1607.0	1761.3	154.4	1762.6	155.6
	acid					
Set 1	Glutamic acid	1609.0	1558.5	50.5	1566.3	42.7
Set 2	Anthranilic acid	1613.2	1766.2	153.0	1752.6	139.4
Set 1	Tartaric acid	1620.9	1577.3	43.6	1571.9	49.0
Set 2	5-Aminovaleric acid	1623.9	1535.5	88.5	1545.4	78.6
Set 1	Phenylalanine	1626.7	1776.6	149.8	1772.2	145.5
Set 2	4-Hydroxybenzoic acid	1628.2	1693.4	65.2	1702.7	74.5
Set 1	Xylose_1	1630.3	1692.2	61.9	1682.0	51.8
Set 2	Lyxose_1	1630.5	1692.2	61.7	1682.0	51.6
Set 1	Xylose_2	1637.2	1692.2	55.0	1682.0	44.8
Set 2	4-Hydroxyphenylacetic	1638.8	1788.8	149.9	1794.4	155.6
	acid					
Set 1	Paeonol_1	1640.4	1581.9	58.6	1573.3	67.1
Set 2	Lyxose_2	1640.9	1692.2	51.3	1682.0	41.1
Set 1	threo-3-Hydroxy-aspartic	1641.8	1609.2	32.7	1606.1	35.8
	acid					
Set 2	Arabinose_1	1643.5	1692.2	48.6	1682.0	38.5
Set 1	Arabinose_2	1643.5	1692.2	48.6	1682.0	38.5
Set 2	Lauric acid(C12)	1650.9	1756.4	105.5	1752.9	102.0
Set 1	n-Acetyl-aspartic acid_1	1651.3	1598.0	53.3	1583.2	68.1

Set 2	Paeonol_2	1653.3	1668.9	15.7	1672.3	19.0
Set 1	n-Acetyl-aspartic acid_2	1655.7	1707.5	51.8	1691.5	35.9
Set 2	Homocysteine	1656.9	1442.6	214.3	1439.8	217.2
Set 1	Xylulose	1657.6	1715.0	57.5	1701.3	43.7
Set 2	Ribulose_1	1657.8	1715.0	57.2	1701.3	43.5
Set 1	Ribulose_2	1657.8	1715.0	57.2	1701.3	43.5
Set 2	Ribose_1	1659.1	1692.2	33.1	1682.0	23.0
Set 1	Ribose_2	1659.1	1692.2	33.1	1682.0	23.0
Set 2	5-Hydroxymethyluracil	1659.8	1855.6	195.9	1817.7	158.0
Set 1	Asparagine	1661.1	1576.3	84.7	1575.3	85.8
Set 2	Taurine	1665.1	1406.8	258.3	1421.9	243.2
Set 1	Cysteine sulfinic acid	1673.6	1709.3	35.7	1685.3	11.7
Set 2	Kojic acid	1678.0	1565.1	112.9	1567.0	111.0
Set 1	Bicine	1678.1	1601.7	76.4	1592.6	85.4
Set 2	Xylitol	1684.9	1776.4	91.5	1750.8	65.9
Set 1	n-Formyl methionine_1	1685.9	1483.5	202.4	1460.4	225.5
Set 2	n-Formyl methionine_2	1690.3	1594.3	96.0	1570.1	120.2
Set 1	Phthalic acid	1692.2	1755.3	63.1	1752.5	60.3
Set 2	Suberic acid	1692.5	1675.8	16.7	1682.9	9.6
Set 1	Methyl jasmonic acid_1	1693.6	1631.0	62.6	1621.0	72.6
Set 2	1,6-Anhydro-β-glucose	1695.6	1439.3	256.3	1426.1	269.5
Set 1	Arabitol	1699.1	1776.4	77.2	1750.8	51.6
Set 2	2-Aminoadipic acid	1703.7	1655.3	48.4	1659.4	44.3
Set 1	Ribitol	1704.0	1776.4	72.4	1750.8	46.8
Set 2	Glycerol-2-phosphonic	1707.3	1728.2	20.9	1693.8	13.5
	acid					
Set 1	Quinolinic acid	1728.0	1739.3	11.4	1736.7	8.7
Set 2	Ciliatine	1732.5	1724.2	8.3	1694.0	38.5
Set 1	Putrescine	1734.1	1587.6	146.5	1593.0	141.1
Set 2	Orotic acid	1734.5	1915.8	181.3	1883.3	148.8
Set 1	2,3-Dihydroxybenzoic acid	1735.4	1846.8	111.4	1836.3	100.9
Set 2	Methyl jasmonic acid_2	1735.5	1631.0	104.5	1621.0	114.5
Set 1	Aconitic acid	1737.2	1701.9	35.2	1698.7	38.4
Set 2	n-Acetyl-glutamic acid_1	1746.5	1808.1	61.6	1788.4	41.9
Set 1	Theanine_1	1754.0	1789.7	35.7	1772.6	18.6
Set 2	5-Amino levulinic acid_1	1755.5	1632.6	122.9	1640.3	115.3
Set 1	2,6-Pyridinedicarboxylic	1758.1	1763.9	5.8	1765.7	7.6

	agid					
G / 2		17(0.7	1704.5	22 0	1705.0	25.2
Set 2	1,3-Benzenedicarboxylic	1/60./	1784.5	23.8	1785.9	25.2
	acid	15(5.2	16565	00.7		00.6
Set I	Glutamine	1765.3	1676.5	88.7	16/1.7	93.6
Set 2	1-Methylhistamine	1767.5	1614.5	153.0	1614.1	153.3
Set 1	Gentisic acid	1768.1	1859.6	91.5	1847.6	79.5
Set 2	Dihydroorotic acid	1768.2	1568.3	199.9	1531.0	237.2
Set 1	4-Hydroxymandelic acid	1769.5	1939.5	170.0	1931.3	161.8
Set 2	3-Methoxy-4-Hydroxyphe	1769.8	1837.5	67.7	1835.7	65.9
	nylacetic acid					
Set 1	o-Phosphoethanolamine	1772.7	1692.1	80.6	1672.3	100.4
Set 2	n-Acetyl-glutamic acid_2	1777.5	1699.0	78.5	1680.4	97.2
Set 1	5-Amino levulinic acid_2	1778.8	1632.6	146.2	1640.3	138.5
Set 2	Theanine_2	1778.9	1675.4	103.5	1660.5	118.4
Set 1	Shikimic acid	1791.5	1859.6	68.1	1834.9	43.4
Set 2	1,4-Benzenedicarboxylic	1791.7	1812.1	20.4	1817.9	26.2
	acid					
Set 1	2-Aminopimelic acid	1800.0	1751.6	48.5	1752.0	48.0
Set 2	Glycylglycine_1	1801.6	1766.2	35.3	1758.7	42.9
Set 1	Isocitric acid	1803.3	1795.2	8.1	1778.5	24.8
Set 2	Citric acid	1803.3	1800.8	2.5	1783.3	20.0
Set 1	Hippuric acid_1	1805.5	1938.9	133.4	1923.5	118.0
Set 2	1-Methylhistidine_1	1805.6	1734.0	71.7	1727.8	77.8
Set 1	Indole-3-acetaldehyde_1	1806.3	1869.3	63.1	1843.1	36.8
Set 2	Ornithine	1806.4	1707.4	99.0	1707.0	99.4
Set 1	Hypoxanthine	1811.2	1718.0	93.2	1707.4	103.7
Set 2	3,4-Dihydroxybenzoic acid	1814.6	1877.8	63.2	1871.6	57.0
Set 1	Glycylglycine_2	1819.5	1597.5	222.0	1600.9	218.6
Set 2	o-Phospho serine	1821.0	1814.0	6.9	1787.4	33.5
Set 1	3,4-Dihydroxyphenylaceti	1822.5	1973.7	151.2	1963.9	141.3
	c acid					
Set 2	Homogentisic acid	1827.4	1958.7	131.2	1942.9	115.5
Set 1	Dimethylbenzimidazole	1831.4	1792.5	38.8	1772.7	58.6
Set 2	Cadaverine	1832.5	1683.9	148.7	1685.6	147.0
Set 1	2-Methyl hippuric acid_1	1833.9	2005.2	171.4	1982.1	148.3
Set 2	Phenanthrene	1836.9	2011.7	174.7	1984.0	147.1
Set 1	4-Aminobenzoic acid	1838.0	1804.5	33.5	1802.5	35.5

Set 2	Anthracene	1847.4	2011.7	164.3	1984.0	136.6
Set 1	Histidinol	1848.0	1833.7	14.3	1818.5	29.5
Set 2	2-Dehydro gluconic acid_1	1849.6	1996.4	146.8	1966.1	116.5
Set 1	Fructose_1	1849.7	1967.7	118.0	1936.3	86.6
Set 2	Sorbose_1	1850.9	1967.7	116.9	1936.3	85.4
Set 1	Homocysteic acid	1854.0	1626.1	227.9	1631.6	222.4
Set 2	Sorbose_2	1855.0	1967.7	112.7	1936.3	81.3
Set 1	Hippuric acid_2	1855.3	1829.3	26.1	1814.9	40.4
Set 2	Fructose_2	1859.9	1967.7	107.8	1936.3	76.3
Set 1	3-Hydroxyanthranilic acid	1861.5	1979.7	118.1	1954.8	93.3
Set 2	Mannose_1	1865.0	1944.9	79.9	1917.0	52.0
Set 1	Galactose_1	1870.8	1944.9	74.0	1917.0	46.2
Set 2	6-Amino-1-methyluracil	1874.6	1893.1	18.4	1856.5	18.1
Set 1	Adenine	1875.3	1887.0	11.6	1855.6	19.7
Set 2	Glucose_1	1877.1	1996.8	119.7	1964.0	86.9
Set 1	Histamine	1879.0	1775.0	104.0	1771.0	108.0
Set 2	Caffeine	1880.2	1820.3	60.0	1760.9	119.3
Set 1	Mannose_2	1882.3	1944.9	62.5	1917.0	34.7
Set 2	Sebacic acid	1888.4	1868.3	20.1	1868.0	20.4
Set 1	Pyridoxine	1890.8	1934.8	44.0	1905.5	14.7
Set 2	Galactose_2	1894.1	1944.9	50.8	1917.0	22.9
Set 1	2-Methylhippuric acid_2	1895.8	1895.5	0.3	1873.6	22.3
Set 2	Glucose_2	1897.8	1944.9	47.0	1917.0	19.2
Set 1	4-Hydroxyphenyl pyruvic	1901.9	1788.3	113.6	1804.9	97.0
	acid					
Set 2	Mannitol	1908.5	2029.2	120.7	1985.9	77.4
Set 1	Lysine	1911.8	1803.7	108.1	1799.6	112.2
Set 2	Tyramine	1913.4	1937.1	23.7	1934.6	21.1
Set 1	Galactosamine_1	1915.6	1965.8	50.1	1941.4	25.8
Set 2	Sorbitol	1915.9	2029.2	113.3	1985.9	70.0
Set 1	Histidine	1917.0	1866.8	50.3	1852.6	64.4
Set 2	5-Keto gluconic acid_1	1917.6	1995.4	77.8	1965.1	47.5
Set 1	Glucuronic acid_1	1918.1	1972.6	54.4	1945.9	27.8
Set 2	Galactitol	1920.7	2029.2	108.5	1985.9	65.2
Set 1	5-Keto gluconic acid_2	1924.9	1995.4	70.4	1965.1	40.2
Set 2	Galacturonic acid_1	1925.2	1972.6	47.4	1945.9	20.7
Set 1	Indole-3-acetaldehyde_2	1925.4	1959.9	34.5	1945.4	20.0

Set 2	Galactosamine_2	1926.8	1965.8	39.0	1941.4	14.6
Set 1	n-Carbamoyl aspartic	1926.9	1611.7	315.2	1612.2	314.7
	acid_2					
Set 2	Tyrosine	1933.0	2057.0	124.0	2048.6	115.6
Set 1	n-Acetyl glutamine	1934.2	1817.0	117.2	1785.7	148.5
Set 2	Ascorbic acid	1935.7	1885.8	49.9	1849.4	86.4
Set 1	Coniferyl alcohol	1940.2	1928.2	12.0	1925.2	15.0
Set 2	Cysteic acid	1942.6	1691.9	250.7	1690.3	252.3
Set 1	Gallic acid	1943.6	2058.2	114.5	2036.9	93.3
Set 2	<i>p</i> -Coumaric acid	1945.2	1916.5	28.7	1921.8	23.4
Set 1	Galacturonic acid_2	1946.2	1972.6	26.4	1945.9	0.3
Set 2	Epinephrine	1951.1	2015.2	64.1	1987.6	36.4
Set 1	Coniferyl aldehyde_1	1953.2	1845.4	107.9	1857.7	95.5
Set 2	1-Hexadecanol	1956.6	2104.5	147.9	2085.4	128.8
Set 1	Coniferyl aldehyde_2	1960.5	1845.4	115.1	1857.7	102.7
Set 2	n- α -Acetyl ornithine_1	1961.1	1957.0	4.1	1929.0	32.1
Set 1	<i>p</i> -Xanthine	1980.9	1857.7	123.3	1805.7	175.3
Set 2	Pantothenic acid	1984.2	2018.2	34.0	1984.8	0.6
Set 1	n-α-Acetyl ornithine_2	1992.8	1782.8	210.0	1766.8	226.0
Set 2	S-Benzyl cysteine	1998.3	2037.0	38.7	2012.4	14.1
Set 1	Urocanic acid	2006.7	1726.2	280.5	1725.8	280.9
Set 2	2-Deoxyribose	2006.8	2025.6	18.8	1989.3	17.5
	5-phosphatic acid					
Set 1	Xanthine	2015.7	2108.3	92.6	2051.7	36.0
Set 2	Palmitoleic acid	2025.0	2160.0	135.0	2144.4	119.4
Set 1	3-Methyladenine	2033.0	1833.4	199.6	1806.3	226.7
Set 2	o-Succinyl homoserine	2036.5	1803.9	232.6	1802.2	234.3
Set 1	1-Methylhistidine_2	2047.4	1902.5	144.9	1885.4	162.0
Set 2	n-Acetyl glucosamine_1	2055.3	2215.3	160.0	2163.4	108.1
Set 1	n- α -Acetyl lysine_1	2057.4	2053.3	4.1	2021.6	35.8
Set 2	n-Acetyl glucosamine_2	2064.7	2215.3	150.6	2163.4	98.7
Set 1	Dopamine	2073.6	2122.4	48.8	2104.3	30.7
Set 2	Kynurenic acid	2075.5	1993.1	82.4	1974.5	101.0
Set 1	Arabinose 5-phosphatic	2076.4	2229.7	153.3	2174.7	98.3
	acid_1					
Set 2	Ribulose 5-phosphatic acid_1	2080.2	2200.1	119.9	2146.3	66.1

Set 1	Arabinose 5-phosphatic	2083.4	2177.2	93.8	2127.1	43.7
	acid_2					
Set 2	Dopa	2085.9	2242.0	156.2	2218.1	132.2
Set 1	Uric acid	2087.8	2289.1	201.3	2218.1	130.3
Set 2	Ribulose 5-phosphatic	2094.8	2200.1	105.3	2146.3	51.5
	acid_2					
Set 1	trans-Ferulic acid	2096.5	1965.0	131.5	1962.9	133.6
Set 2	Pyridoxamine	2098.7	2120.8	22.1	2084.2	14.5
Set 1	n- α -Acetyl lysine_2	2103.8	1879.1	224.7	1859.5	244.3
Set 2	Dethiobiotin	2115.1	2153.2	38.0	2118.1	3.0
Set 1	Guanine	2127.0	1979.4	147.7	1964.1	163.0
Set 2	Methyldopa	2130.5	2293.4	163.0	2264.6	134.2
Set 1	Caffeic acid	2133.8	2101.3	32.6	2091.1	42.8
Set 2	Flavanone	2135.2	2097.9	37.2	2074.3	60.8
Set 1	Heptadecanoic acid (C17)	2142.9	2237.6	94.7	2215.7	72.8
Set 2	5-Sulfosalicylic acid_2	2171.4	2048.3	123.1	2030.4	141.0
Set 1	2,6-Diaminopurine	2186.7	2368.5	181.7	2308.3	121.6
Set 2	Kynurenine	2187.1	2293.6	106.5	2253.0	65.9
Set 1	Cystamine	2208.2	1933.6	274.6	1905.4	302.9
Set 2	5-Hydroxyindoleacetic	2212.9	2332.6	119.7	2295.3	82.4
	acid					
Set 1	Tryptophan	2217.1	2261.5	44.4	2216.8	0.3
Set 2	Elaidic acid	2223.7	2352.4	128.7	2329.5	105.7
Set 1	Stearic acid (C18)	2242.2	2333.8	91.6	2308.3	66.0
Set 2	Sinapic acid	2247.0	2009.4	237.5	2000.4	246.6
Set 1	Xanthurenic acid	2259.8	2240.0	19.8	2213.4	46.4
Set 2	Sinigrin	2289.9	2662.0	372.1	2582.1	292.2
Set 1	Sorbitol 6-phosphatic acid	2338.4	2513.5	175.1	2430.2	91.8
Set 2	3-Hydroxy kynurenine	2418.2	2507.2	89.0	2455.4	37.2
Set 1	5-Hydroxy tryptophan_1	2425.6	2601.2	175.5	2549.8	124.2
Set 2	Biopterin	2430.8	2487.2	56.4	2449.2	18.4
Set 1	2'-Deoxyuridine	2438.4	2201.2	237.1	2139.1	299.3
Set 2	Icosanoic acid (C20)	2439.6	2526.3	86.7	2493.4	53.7
Set 1	5-Hydroxy tryptophan_2	2448.9	2477.5	28.6	2430.2	18.7
Set 2	<i>p</i> -Aminohippuric acid	2454.7	2220.7	234.0	2191.0	263.7
Set 1	5-Hydroxy tryptamine_1	2460.5	2481.3	20.8	2435.8	24.7
Set 2	n-Acetyl-Glucosamine	2465.7	2591.3	125.7	2500.5	34.9

	6-phosphatic acid_1					
Set 1	Biotin	2479.3	2318.3	161.0	2260.8	218.5
Set 2	n-Acetyl-Glucosamine	2491.2	2591.3	100.1	2500.5	9.3
	6-phosphatic acid_2					
Set 1	Salicyl	2493.2	2579.6	86.4	2509.3	16.1
	alcohol-β-glucoside					
Set 2	Ribulose 1,5-bisphosphatic	2498.7	2685.0	186.3	2591.3	92.5
	acid_1					
Set 1	5-Hydroxy tryptamine_2	2501.8	2307.1	194.7	2273.6	228.2
Set 2	Ribulose 1,5-bisphosphatic	2510.6	2685.0	174.4	2591.3	80.7
	acid_2					
Set 1	Saccharopine	2515.6	2497.7	18.0	2460.8	54.9
Set 2	Homocystine	2517.7	2367.9	149.8	2320.6	197.1
Set 1	2'-Deoxyinosine	2549.2	2351.8	197.3	2311.4	237.8
Set 2	Inosine	2564.2	2622.8	58.6	2557.9	6.3
Set 1	2'-Deoxyadenosine	2629.2	2601.4	27.8	2536.4	92.8
Set 2	Adenosine	2629.7	2751.7	122.0	2672.8	43.1
Set 1	Behenic acid (C22)	2637.5	2718.8	81.2	2678.5	40.9
Set 2	Xanthosine	2656.8	2892.3	235.5	2791.9	135.1
Set 1	Trehalose	2714.2	2781.1	66.9	2673.5	40.7
Set 2	Spermine	2739.6	2558.3	181.3	2531.3	208.3
Set 1	Guanosine	2756.3	2708.1	48.2	2640.4	115.9
Set 2	2-Deoxyguanosine	2776.7	2558.2	218.5	2504.3	272.4
Set 1	5-Methylthioadenosine	2801.0	2825.0	24.0	2734.9	66.1
Set 2	Acetaminophen	2805.9	2848.2	42.3	2768.2	37.7
	glucuronide					
Set 1	Lignoceric acid (C24)	2836.3	2911.3	75.0	2863.6	27.3
Set 2	Epicatechin	2840.9	3061.6	220.7	2982.7	141.8
Set 1	Naringenin_1	2887.4	2781.2	106.2	2726.6	160.8
Set 2	Sakuranetin_1	2890.2	2645.1	245.1	2598.5	291.7
Set 1	Genistein	2961.4	2876.0	85.4	2825.0	136.4
Set 2	Daidzein	2964.8	2689.0	275.8	2654.0	310.8
Set 1	Chlorogenic acid_1	2965.3	3215.1	249.8	3128.9	163.5
Set 2	Naringenin_2	2967.1	2815.3	151.8	2763.8	203.3
Set 1	Sakuranetin_2	2971.5	2679.1	292.3	2635.7	335.8
Set 2	Kaempferol	3071.4	3036.6	34.9	2963.3	108.1
Set 1	Chlorogenic acid_2	3092.2	3215.1	122.9	3128.9	36.7

Set 2	Albiflorin	3570.9	3436.0	134.8	3294.0	276.8
Set 1	Podophyllotoxin	3428.1	2797.9	630.2	2702.0	726.0

3-生薬センキュウ中の代謝物

以下に記載した代謝物および未同定代謝物から成るデータ行列を PCA に供した.

RI: retention index, RT: retention time

Align	Average	Average	Quant	Metabolite name	InChIKey
-ment	RT (min)	RI	mass		
ID					
1	3.51	926.49	171	Unknown	No record
4	3.52	927.82	163	Unknown	No record
33	3.76	962.23	147	Unknown	No record
37	3.92	984.76	133	Unknown	No record
39	3.91	983.98	117	Propylene glycol;	DNIAPMSPPWPWGF-U
				GC-EI-TOF; MS; n	HFFFAOYNA-N
				TMS; RT	
49	4.01	997.91	162	Unknown	No record
50	4.04	1001.57	207	Unknown	No record
57	4.23	1017.86	86	Unknown	No record
59	4.23	1017.83	174	Unknown	No record
60	4.22	1017.49	175	Unknown	No record
61	4.23	1017.97	174	n-Propylamine;	WGYKZJWCGVVSQN-U
				GC-EI-TOF; MS; n	HFFFAOYSA-N
				TMS; RT	
66	4.28	1022.79	102	Unknown	No record
68	4.32	1025.93	98	Unknown	No record
74	4.39	1032.14	144	Unknown	No record
76	4.39	1032.14	117	Unknown	No record
77	4.4	1032.73	116	Unknown	No record
78	4.42	1034.71	152	2-Hydroxypyridine;	UBQKCCHYAOITMY-U
				GC-EI-TOF; MS; n	HFFFAOYSA-N
				TMS; RT	
79	4.42	1034.69	152	Unknown	No record
85	4.52	1043.16	174	Pyruvic acid;	LCTONWCANYUPML-U
				GC-EI-TOF; MS; n	HFFFAOYSA-N
				TMS; RT	

90	4.61	1051.48	117	L-(+)-Lactic acid;	JVTAAEKCZFNVCJ-UH
				GC-EI-TOF; MS; n	FFFAOYNA-N
				TMS; RT	
91	4.62	1052.43	194	Unknown	No record
99	4.78	1065.91	102	Unknown	No record
102	4.81	1068.74	147	Glycolic acid;	AEMRFAOFKBGASW-U
				GC-EI-TOF; MS; n	HFFFAOYSA-N
				TMS; RT	
109	4.95	1080.8	147	Unknown	No record
112	5.01	1086.24	146	L-Valine;	KZSNJWFQEVHDMF-B
				GC-EI-TOF; MS; 1	YPYZUCNSA-N
				TMS; n TMS; RT	
114	5.06	1090.59	174	Isobutylamine;	KDSNLYIMUZNERS-UH
				GC-EI-TOF; MS; n	FFFAOYSA-N
				TMS; RT	
116	5.12	1095.72	133	Unknown	No record
118	5.12	1096.21	149	Unknown	No record
119	5.11	1095.58	116	L-Alanine;	QNAYBMKLOCPYGJ-U
				GC-EI-TOF; MS; n	HFFFAOYNA-N
				TMS; RT	
122	5.13	1096.85	146	Unknown	No record
130	5.18	1101.63	174	n-Butylamine;	HQABUPZFAYXKJW-UH
				GC-EI-TOF; MS; n	FFFAOYSA-N
				TMS; RT	
131	5.25	1107.66	188	Unknown	No record
142	5.34	1116.16	102	Unknown	No record
146	5.37	1118.64	131	HydroxyButyric	AFENDNXGAFYKQO-U
				acid; GC-EI-TOF;	HFFFAOYNA-N
				MS; n TMS; RT	
149	5.43	1124.73	147	Oxalic acid;	MUBZPKHOEPUJKR-UH
				GC-EI-TOF; MS; n	FFFAOYSA-N
				TMS; RT	
165	5.59	1138.85	152	Unknown	No record
173	5.68	1147.34	141	Unknown	No record
174	5.76	1154.48	369	Unknown	No record
177	5.76	1154.28	86	Unknown	No record
178	5.76	1154.88	147	DL-beta-Hydroxyb	WHBMMWSBFZVSSR-U
				utyric acid;	HFFFAOYNA-N

				GC-EI-TOF; MS; n	
				TMS; RT	
183	5.79	1157.19	145	Unknown	No record
184	5.89	1166.01	130	L-2-Aminobutyric	QWCKQJZIFLGMSD-VK
				acid; GC-EI-TOF;	HMYHEASA-N
				MS; n TMS; RT	
188	5.94	1171.23	241	Unknown	No record
194	6	1176.57	86	Unknown	No record
198	6.05	1181.32	89	Unknown	No record
205	6.18	1193.27	170	Unknown	No record
211	6.23	1197.15	147	Malonic acid;	OFOBLEOULBTSOW-U
				GC-EI-TOF; MS; n	HFFFAOYSA-N
				TMS; RT	
215	6.24	1198.12	147	Unknown	No record
216	6.3	1204.26	131	3-Hydroxyisovaleri	AXFYFNCPONWUHW-U
				c acid;	HFFFAOYSA-N
				GC-EI-TOF; MS; n	
				TMS; RT	
218	6.34	1208.58	146	Unknown	No record
219	6.35	1208.76	128	Unknown	No record
221	6.35	1208.68	218	Unknown	No record
222	6.35	1208.92	144	L-Valine;	KZSNJWFQEVHDMF-B
				GC-EI-TOF; MS; 2	YPYZUCNSA-N
				TMS; n TMS; RT	
228	6.43	1217.36	211	Unknown	No record
239	6.57	1230.62	149	Unknown	No record
241	6.57	1230.72	147	Unknown	No record
253	6.66	1239.38	147	Urea; GC-EI-TOF;	XSQUKJJJFZCRTK-UHF
				MS; n TMS; RT	FFAOYSA-N
254	6.67	1241	147	Unknown	No record
263	6.78	1251.47	179	Benzoic acid;	WPYMKLBDIGXBTP-U
				GC-EI-TOF; MS; n	HFFFAOYSA-N
				TMS; RT	
265	6.79	1252.67	116	L-Serine;	MTCFGRXMJLQNBG-R
				GC-EI-TOF; MS; 2	EOHCLBHSA-N
				TMS; n TMS; RT	
266	6.82	1255.07	147	Oxamic acid;	SOWBFZRMHSNYGE-U
				GC-EI-TOF; MS; n	HFFFAOYSA-N

				TMS; RT	
269	6.86	1259.79	159	Unknown	No record
272	6.88	1260.92	174	Ethanolamine;	HZAXFHJVJLSVMW-UH
				GC-EI-TOF; MS; n	FFFAOYSA-N
				TMS; RT	
274	6.89	1262.28	298	Unknown	No record
275	6.9	1262.99	299	Pyrophosphoric	XPPKVPWEQAFLFU-UH
				acid; GC-EI-TOF;	FFFAOYSA-N
				MS; n TMS; RT	
278	6.9	1263.69	147	Glycerol;	PEDCQBHIVMGVHV-U
				GC-EI-TOF; MS; n	HFFFAOYSA-N
				TMS; RT	
279	6.92	1265.23	160	Unknown	No record
283	6.92	1264.87	158	L-Leucine;	ROHFNLRQFUQHCH-Y
				GC-EI-TOF; MS; n	FKPBYRVSA-N
				TMS; RT	
289	6.97	1269.73	103	Unknown	No record
297	7.06	1279.11	117	Unknown	No record
300	7.08	1280.98	217	Unknown	No record
301	7.1	1282.83	117	Unknown	No record
306	7.13	1286.42	158	L-Isoleucine;	AGPKZVBTJJNPAG-WH
				GC-EI-TOF; MS; n	FBIAKZSA-N
				TMS; RT	
310	7.16	1288.55	130	Unknown	No record
311	7.16	1288.85	117	Unknown	No record
315	7.22	1294.51	142	L-(-)-Proline;	ONIBWKKTOPOVIA-UH
				GC-EI-TOF; MS; n	FFFAOYNA-N
				TMS; RT	
321	7.25	1298.04	180	Unknown	No record
322	7.25	1297.78	180	Nicotinic acid;	PVNIIMVLHYAWGP-UH
				GC-EI-TOF; MS; n	FFFAOYSA-N
				TMS; RT	
326	7.28	1300.42	174	Glycine;	DHMQDGOQFOQNFH-U
				GC-EI-TOF; MS; n	HFFFAOYSA-N
				TMS; RT	
330	7.33	1306.5	175	Unknown	No record
331	7.35	1307.7	149	Unknown	No record
332	7.35	1308.24	247	Unknown	No record

334	7.35	1308.49	147	Succinic acid;	KDYFGRWQOYBRFD-U
				GC-EI-TOF; MS; n	HFFFAOYSA-N
				TMS; RT	
335	7.37	1309.92	146	Unknown	No record
336	7.38	1311.89	126	Unknown	No record
347	7.46	1319.49	147	Glyceric acid;	RBNPOMFGQQGHHO-U
				GC-EI-TOF; MS; n	WTATZPHSA-N
				TMS; RT	
352	7.59	1333.93	99	Uracil;	ISAKRJDGNUQOIC-UHF
				GC-EI-TOF; MS; n	FFAOYSA-N
				TMS; RT	
360	7.7	1345.15	245	Fumaric acid;	VZCYOOQTPOCHFL-U
				GC-EI-TOF; MS; n	HFFFAOYSA-N
				TMS; RT	
361	7.7	1345.79	148	Unknown	No record
364	7.72	1347.24	117	Unknown	No record
370	7.75	1350.39	204	L-Serine;	MTCFGRXMJLQNBG-R
				GC-EI-TOF; MS; 3	EOHCLBHSA-N
				TMS; n TMS; RT	
373	7.79	1354.77	189	Unknown	No record
375	7.79	1355.18	190	Unknown	No record
378	7.79	1355.65	188	Unknown	No record
386	7.85	1361.05	117	Unknown	No record
390	7.87	1363.44	156	DL-Pipecolic acid;	HXEACLLIILLPRG-UHF
				GC-EI-TOF; MS; n	FFAOYNA-N
				TMS; RT	
398	7.98	1375.57	117	L-(-)-Threonine;	AYFVYJQAPQTCCC-UH
				CC EL TOE. MO. a	EEEAOVALA N
				GC-EI-TOF, MS, h	FFFAUYNA-N
404				TMS; RT	FFFAOYNA-N
	8.07	1384.51	131	TMS; RT Unknown	No record
405	8.07 8.08	1384.51 1386.43	131 160	GC-EI-TOF, MS; n TMS; RT Unknown Unknown	No record No record
405 408	8.07 8.08 8.13	1384.51 1386.43 1391.85	131 160 145	GC-EI-TOF, MS, nTMS; RTUnknownUnknownUnknown	No record No record No record
405 408 409	8.07 8.08 8.13 8.16	1384.51 1386.43 1391.85 1394.07	131 160 145 116	GC-EI-TOF, MS, nTMS; RTUnknownUnknownUnknownUnknown	No record No record No record No record
405 408 409 412	8.07 8.08 8.13 8.16 8.17	1384.51 1386.43 1391.85 1394.07 1395.55	131 160 145 116 255	GC-EI-TOF, MS; nTMS; RTUnknownUnknownUnknownUnknownThymine;	No record No record No record No record RWQNBRDOKXIBIV-UH
405 408 409 412	8.07 8.08 8.13 8.16 8.17	1384.51 1386.43 1391.85 1394.07 1395.55	131 160 145 116 255	GC-EI-TOF, MS; n TMS; RT Unknown Unknown Unknown Unknown Thymine; GC-EI-TOF; MS; n	No record No record No record No record RWQNBRDOKXIBIV-UH FFFAOYSA-N
405 408 409 412	8.07 8.08 8.13 8.16 8.17	1384.51 1386.43 1391.85 1394.07 1395.55	131 160 145 116 255	GC-EI-TOF, MS, nTMS; RTUnknownUnknownUnknownUnknownThymine;GC-EI-TOF; MS; nTMS; RT	No record No record No record No record RWQNBRDOKXIBIV-UH FFFAOYSA-N
405 408 409 412 415	8.07 8.08 8.13 8.16 8.17 8.19	1384.51 1386.43 1391.85 1394.07 1395.55 1397.69	131 160 145 116 255 373	GC-EI-TOF, MS; nTMS; RTUnknownUnknownUnknownUnknownThymine;GC-EI-TOF; MS; nTMS; RTUnknown	No record No record No record No record RWQNBRDOKXIBIV-UH FFFAOYSA-N No record

419	8.23	1402.73	174	Unknown	No record
426	8.34	1414.81	104	Unknown	No record
434	8.39	1420.05	160	Unknown	No record
437	8.41	1422.82	174	beta-Alanine;	UCMIRNVEIXFBKS-UH
				GC-EI-TOF; MS; n	FFFAOYSA-N
				TMS; RT	
438	8.41	1422.76	174	Unknown	No record
453	8.56	1440.4	218	D-Homoserine;	UKAUYVFTDYCKQA-U
				GC-EI-TOF; MS; n	HFFFAOYNA-N
				TMS; RT	
465	8.68	1454.09	232	Unknown	No record
471	8.73	1458.89	142	Unknown	No record
480	8.83	1470.47	158	Unknown	No record
487	8.9	1478.49	211	Unknown	No record
489	8.89	1477.01	147	DL-Malic acid;	BJEPYKJPYRNKOW-UH
				GC-EI-TOF; MS; n	FFFAOYNA-N
				TMS; RT	
490	8.89	1477.96	147	Unknown	No record
492	8.94	1483.72	147	D-Threitol;	UNXHWFMMPAWVPI-U
				GC-EI-TOF; MS; n	HFFFAOYNA-N
				TMS; RT	
495	8.98	1488	179	Niacinamide;	DFPAKSUCGFBDDF-UH
				GC-EI-TOF; MS; n	FFFAOYSA-N
				TMS; RT	
501	9.01	1491.09	149	Unknown	No record
505	9.01	1491.62	147	meso-Erythritol;	UNXHWFMMPAWVPI-U
				GC-EI-TOF; MS; n	HFFFAOYNA-N
				TMS; RT	
508	9.05	1495.29	143	Unknown	No record
512	9.06	1497.35	230	Unknown	No record
519	9.11	1502.95	243	Unknown	No record
526	9.13	1505.59	157	Unknown	No record
529	9.16	1509.4	267	Unknown	No record
532	9.17	1510.37	232	L-Aspartic acid;	CKLJMWTZIZZHCS-UH
				GC-EI-TOF; MS; n	FFFAOYNA-N
				TMS; RT	
534	9.19	1512.13	87	Unknown	No record
539	9.22	1516.25	176	L-Methionine;	FFEARJCKVFRZRR-BYP
				GC-EI-TOF; MS; n	YZUCNSA-N
-----	------	---------	-----	------------------	--------------------
				TMS; RT	
542	9.25	1520.33	247	Unknown	No record
547	9.27	1521.73	156	DL-Pyroglutamic	ODHCTXKNWHHXJC-U
				acid; GC-EI-TOF;	HFFFAOYNA-N
				MS; n TMS; RT	
548	9.27	1522.43	122	Unknown	No record
554	9.3	1525.47	174	GABA;	BTCSSZJGUNDROE-UH
				GC-EI-TOF; MS; n	FFFAOYSA-N
				TMS; RT	
555	9.31	1527.15	292	Unknown	No record
559	9.32	1528.21	186	Unknown	No record
563	9.33	1529.82	191	Unknown	No record
565	9.36	1533.33	139	Unknown	No record
566	9.36	1533.3	155	Unknown	No record
567	9.37	1534.14	157	Unknown	No record
569	9.37	1534.18	129	Unknown	No record
570	9.37	1534.38	247	Unknown	No record
572	9.4	1538.54	219	Unknown	No record
578	9.45	1543.9	147	Unknown	No record
590	9.54	1555.17	120	Unknown	No record
595	9.62	1564.09	129	Unknown	No record
597	9.64	1567.36	275	2-Isopropylmalic	BITYXLXUCSKTJS-UHF
				acid; GC-EI-TOF;	FFAOYNA-N
				MS; n TMS; RT	
600	9.67	1570.09	117	Unknown	No record
601	9.68	1571.28	179	4-Hydroxypheneth	YCCILVSKPBXVIP-UHF
				yl alcohol;	FFAOYSA-N
				GC-EI-TOF; MS; n	
				TMS; RT	
604	9.75	1580.57	142	Unknown	No record
606	9.76	1581.98	193	D-3-Phenyllactic	VOXXWSYKYCBWHO-
				acid; GC-EI-TOF;	UHFFFAOYNA-N
				MS; n TMS; RT	
613	9.87	1594.58	129	Unknown	No record
615	9.87	1595.17	159	Unknown	No record
625	9.99	1610.52	133	Unknown	No record
626	9.99	1609.56	246	L-Glutamic acid;	WHUUTDBJXJRKMK-V

				GC-EI-TOF; MS; n	KHMYHEASA-N
				TMS; RT	
627	10	1611.59	159	Unknown	No record
631	10.08	1621.5	147	L-(+)-Tartaric acid;	FEWJPZIEWOKRBE-UH
				GC-EI-TOF; MS; n	FFFAOYNA-N
				TMS; RT	
640	10.11	1626.19	218	L-(-)-Phenylalanine	COLNVLDHVKWLRT-U
				; GC-EI-TOF; MS;	HFFFAOYNA-N
				n TMS; RT	
643	10.13	1628.62	200	Unknown	No record
644	10.15	1630.72	116	Unknown	No record
647	10.16	1631.94	245	Unknown	No record
651	10.2	1637.21	167	Unknown	No record
653	10.2	1637.66	103	D-(+)-Xylose;	SRBFZHDQGSBBOR-UH
				GC-EI-TOF; MS; n	FFFAOYNA-N
				TMS; RT	
659	10.26	1644.85	103	Unknown	No record
660	10.26	1644.73	147	DL-Arabinose;	PYMYPHUHKUWMLA-
				GC-EI-TOF; MS; n	UHFFFAOYNA-N
				TMS; RT	
665	10.31	1651.33	223	Unknown	No record
669	10.35	1656.68	245	Unknown	No record
672	10.37	1659.04	147	D-Ribulose;	LQXVFWRQNMEDEE-U
				GC-EI-TOF; MS; n	HFFFAOYNA-N
				TMS; RT	
673	10.37	1659.8	103	D-(-)-Ribose;	HMFHBZSHGGEWLO-S
				GC-EI-TOF; MS; n	OOFDHNKSA-N
				TMS; RT	
677	10.39	1661.81	132	Unknown	No record
679	10.39	1661.62	132	Unknown	No record
680	10.39	1661.92	116	L-Asparagine;	DCXYFEDJOCDNAF-U
				GC-EI-TOF; MS; n	WTATZPHSA-N
				TMS; RT	
682	10.42	1665.93	288	Unknown	No record
683	10.44	1667.84	160	Unknown	No record
684	10.44	1667.92	264	Unknown	No record
691	10.5	1676.07	105	Unknown	No record
693	10.5	1675.9	147	Unknown	No record

701	10.53	1679.85	103	Unknown	No record
702	10.58	1685.97	205	Unknown	No record
703	10.58	1686.6	307	Unknown	No record
705	10.58	1686.9	219	Unknown	No record
706	10.58	1686.76	217	Xylitol;	HEBKCHPVOIAQTA-UH
				GC-EI-TOF; MS; n	FFFAOYNA-N
				TMS; RT	
707	10.59	1687.74	231	Unknown	No record
710	10.62	1691.3	147	Phthalic acid;	XNGIFLGASWRNHJ-UH
				GC-EI-TOF; MS; n	FFFAOYSA-N
				TMS; RT	
712	10.64	1693.71	116	Unknown	No record
714	10.66	1696.32	204	1,6-Anhydro-beta-	TWNIBLMWSKIRAT-UH
				D-glucose;	FFFAOYNA-N
				GC-EI-TOF; MS; n	
				TMS; RT	
720	10.69	1700.96	144	Unknown	No record
721	10.69	1700.46	103	D-Arabitol;	HEBKCHPVOIAQTA-UH
				GC-EI-TOF; MS; n	FFFAOYNA-N
				TMS; RT	
722	10.71	1703.18	268	Unknown	No record
726	10.73	1705.59	333	Unknown	No record
727	10.73	1706.13	217	Ribitol;	HEBKCHPVOIAQTA-UH
				GC-EI-TOF; MS; n	FFFAOYNA-N
				TMS; RT	
729	10.76	1710.52	117	D-(+)-Fucose;	SHZGCJCMOBCMKK-S
				GC-EI-TOF; MS; n	VZMEOIVSA-N
				TMS; RT	
732	10.78	1712.73	157	Unknown	No record
736	10.83	1720.27	267	Unknown	No record
744	10.87	1725.59	117	D-(+)-Fucose;	SHZGCJCMOBCMKK-S
				GC-EI-TOF; MS; n	VZMEOIVSA-N
				TMS; RT	
746	10.93	1733.61	167	Unknown	No record
748	10.93	1733.69	174	Putrescine;	KIDHWZJUCRJVML-UH
				GC-EI-TOF; MS; n	FFFAOYSA-N
				TMS; RT	
749	10.94	1735.11	260	Unknown	No record

750	10.96	1737.01	211	Unknown	No record
752	10.95	1736.34	103	Unknown	No record
757	10.96	1737.45	147	Aconitic acid;	GTZCVFVGUGFEME-U
				GC-EI-TOF; MS; n	HFFFAOYSA-N
				TMS; RT	
758	10.97	1738.73	131	Unknown	No record
759	10.98	1740.1	142	Unknown	No record
761	10.98	1740.5	231	Unknown	No record
764	10.99	1741.46	184	Unknown	No record
767	11	1742.8	217	Unknown	No record
770	11.01	1744.59	159	Unknown	No record
771	11.02	1745.35	103	Unknown	No record
774	11.04	1748.3	107	Unknown	No record
779	11.06	1751.56	160	Unknown	No record
780	11.08	1753.5	217	Unknown	No record
799	11.17	1766.3	156	D-Glutamine;	ZDXPYRJPNDTMRX-UH
				GC-EI-TOF; MS; n	FFFAOYNA-N
				TMS; RT	
805	11.22	1773.5	172	Unknown	No record
806	11.22	1773.4	174	O-Phosphoethanola	SUHOOTKUPISOBE-UH
				mine; GC-EI-TOF;	FFFAOYSA-N
				MS; n TMS; RT	
809	11.24	1776.32	231	Unknown	No record
817	11.36	1792.23	117	Unknown	No record
819	11.36	1792.15	204	(-)-Shikimic acid;	JXOHGGNKMLTUBP-U
				GC-EI-TOF; MS; n	HFFFAOYNA-N
				TMS; RT	
830	11.45	1804.53	147	Citric acid;	KRKNYBCHXYNGOX-U
				GC-EI-TOF; MS; n	HFFFAOYSA-N
				TMS; RT	
831	11.47	1807.39	174	Unknown	No record
835	11.47	1807.32	142	L-Ornithine;	AHLPHDHHMVZTML-B
				GC-EI-TOF; MS; n	YPYZUCNSA-N
				TMS; RT	
839	11.49	1810.95	265	Unknown	No record
841	11.52	1814.54	193	Protocatechuic	YQUVCSBJEUQKSH-UH
				acid; GC-EI-TOF;	FFFAOYSA-N
				MS; n TMS; RT	

848	11.53	1816.12	157	L-Citrulline;	RHGKLRLOHDJJDR-BY
				GC-EI-TOF; MS; n	PYZUCNSA-N
				TMS; RT	
851	11.57	1821.26	297	Unknown	No record
852	11.57	1822.29	116	Unknown	No record
856	11.63	1830.43	103	D-Tagatose;	BJHIKXHVCXFQLS-UH
				GC-EI-TOF; MS; n	FFFAOYNA-N
				TMS; RT	
860	11.65	1833.8	205	Unknown	No record
861	11.66	1834.21	160	Unknown	No record
862	11.66	1834.34	146	Unknown	No record
863	11.66	1835.23	103	Psicose;	BJHIKXHVCXFQLS-UH
				GC-EI-TOF; MS; n	FFFAOYNA-N
				TMS; RT	
866	11.7	1840.35	267	Unknown	No record
867	11.71	1841.38	231	Unknown	No record
868	11.71	1841.83	146	Unknown	No record
869	11.71	1841.02	147	Quinic acid;	AAWZDTNXLSGCEK-F
				GC-EI-TOF; MS; n	NKGTGPASA-N
				TMS; RT	
870	11.73	1844.71	285	Unknown	No record
874	11.78	1851.36	349	Unknown	No record
875	11.78	1851.58	349	Unknown	No record
878	11.78	1851.97	349	Unknown	No record
879	11.78	1851.88	174	Unknown	No record
880	11.78	1851.79	103	D-(-)-Fructose;	BJHIKXHVCXFQLS-UY
				GC-EI-TOF; MS; n	FOZJQFSA-N
				TMS; RT	
881	11.79	1853.1	217	L-(-)-Sorbose;	BJHIKXHVCXFQLS-UH
				GC-EI-TOF; MS; n	FFFAOYNA-N
				TMS; RT	
882	11.8	1854.37	301	Unknown	No record
884	11.85	1861.03	145	Unknown	No record
885	11.85	1862.14	188	Unknown	No record
887	11.85	1861.21	278	Unknown	No record
888	11.85	1861.55	126	Unknown	No record
889	11.85	1861.32	103	D-(-)-Fructose;	BJHIKXHVCXFQLS-UY
				GC-EI-TOF; MS; n	FOZJQFSA-N

				TMS; RT	
890	11.88	1865.76	147	D-Mannose;	GZCGUPFRVQAUEE-UH
				GC-EI-TOF; MS; n	FFFAOYNA-N
				TMS; RT	
892	11.88	1865.78	96	Unknown	No record
895	11.89	1867.42	103	2-Dehydro-D-gluco	VBUYCZFBVCCYFD-U
				nic acid;	HFFFAOYNA-N
				GC-EI-TOF; MS; n	
				TMS; RT	
902	11.92	1871.33	147	D-(+)-Galactose;	WQZGKKKJIJFFOK-SVZ
				GC-EI-TOF; MS; n	MEOIVSA-N
				TMS; RT	
906	11.94	1875.05	264	Adenine;	MOSFIJXAXDLOML-UH
				GC-EI-TOF; MS; n	FFFAOYSA-N
				TMS; RT	
908	11.97	1878.82	154	Unknown	No record
910	11.97	1878.49	147	L-Glucose;	GZCGUPFRVQAUEE-UH
				GC-EI-TOF; MS; n	FFFAOYNA-N
				TMS; RT	
911	11.98	1880.34	319	D-Mannose;	GZCGUPFRVQAUEE-UH
				GC-EI-TOF; MS; n	FFFAOYNA-N
				TMS; RT	
915	12.01	1884.51	205	D-(+)-Allose;	WQZGKKKJIJFFOK-GA
				GC-EI-TOF; MS; n	SJEMHNSA-N
				TMS; RT	
922	12.05	1889.48	179	Unknown	No record
924	12.06	1891.24	147	Unknown	No record
927	12.06	1891.81	292	Unknown	No record
931	12.07	1892.5	273	Unknown	No record
932	12.07	1893.46	273	Unknown	No record
933	12.09	1896.3	319	D-(+)-Galactose;	WQZGKKKJIJFFOK-SVZ
				GC-EI-TOF; MS; n	MEOIVSA-N
				TMS; RT	
936	12.11	1899.32	203	Unknown	No record
937	12.11	1898.72	147	L-Glucose;	GZCGUPFRVQAUEE-UH
				GC-EI-TOF; MS; n	FFFAOYNA-N
				TMS; RT	
938	12.13	1901.53	103	Unknown	No record

939	12.14	1903.34	297	Unknown	No record
947	12.18	1909.45	319	Unknown	No record
952	12.19	1910.18	176	Unknown	No record
954	12.19	1909.96	147	D-(-)-Mannitol;	FBPFZTCFMRRESA-UH
				GC-EI-TOF; MS; n	FFFAOYNA-N
				TMS; RT	
956	12.19	1911.2	317	Unknown	No record
958	12.2	1911.63	230	Unknown	No record
960	12.2	1911.67	174	L-(+)-Lysine;	KDXKERNSBIXSRK-YF
				GC-EI-TOF; MS; n	KPBYRVSA-N
				TMS; RT	
961	12.22	1914.83	254	Unknown	No record
964	12.22	1915.31	154	L-Histidine;	MOSFIJXAXDLOML-UH
				GC-EI-TOF; MS; n	FFFAOYSA-N
				TMS; RT	
966	12.23	1916.53	147	D-(-)-Sorbitol;	FBPFZTCFMRRESA-UH
				GC-EI-TOF; MS; n	FFFAOYNA-N
				TMS; RT	
967	12.24	1918.58	338	Unknown	No record
971	12.25	1918.96	338	Unknown	No record
975	12.25	1919.75	338	Galactitol;	FBPFZTCFMRRESA-UH
				GC-EI-TOF; MS; n	FFFAOYNA-N
				TMS; RT	
982	12.29	1925.43	160	D-(+)-Galacturonic	AEMOLEFTQBMNLQ-Y
				acid; GC-EI-TOF;	MDCURPLSA-N
				MS; n TMS; RT	
984	12.29	1926.24	143	5-Keto-D-Gluconic	VBUYCZFBVCCYFD-U
				acid; GC-EI-TOF;	HFFFAOYNA-N
				MS; n TMS; RT	
987	12.33	1932.41	217	Unknown	No record
990	12.34	1933.02	218	L-Tyrosine;	OUYCCCASQSFEME-U
				GC-EI-TOF; MS; n	HFFFAOYNA-N
				TMS; RT	
997	12.38	1939.76	174	Unknown	No record
1002	12.4	1942.85	217	Unknown	No record
1004	12.41	1944.2	281	Gallic acid;	LNTHITQWFMADLM-U
				GC-EI-TOF; MS; n	HFFFAOYSA-N
				TMS; RT	

1000	10.40	1045 16	242	TT1	N
1006	12.42	1945.16	242		
1008	12.43	1947.06	160	D-(+)-Galacturonic	AEMOLEFTQBMNLQ-Y
				acid; GC-EI-TOF;	MDCURPLSA-N
				MS; n TMS; RT	
1015	12.47	1953.43	156	Unknown	No record
1017	12.48	1954.22	143	Unknown	No record
1022	12.49	1955.79	204	Unknown	No record
1030	12.62	1975.84	219	Unknown	No record
1031	12.63	1976.5	292	Unknown	No record
1032	12.63	1977.06	294	Unknown	No record
1037	12.66	1980.81	205	Unknown	No record
1039	12.66	1981.58	147	Gluconic acid;	RGHNJXZEOKUKBD-SQ
				GC-EI-TOF; MS; n	OUGZDYSA-N
				TMS; RT	
1040	12.67	1983.39	174	Unknown	No record
1041	12.71	1988.16	297	Unknown	No record
1043	12.75	1994.86	147	D-Glucaric acid;	DSLZVSRJTYRBFB-UHF
				GC-EI-TOF; MS; n	FFAOYNA-N
				TMS; RT	
1047	12.82	2006.2	203	Unknown	No record
1054	12.83	2006.96	204	Unknown	No record
1055	12.84	2008.76	318	Unknown	No record
1057	12.85	2010.28	305	Unknown	No record
1065	12.94	2025	147	Unknown	No record
1067	12.95	2026.75	218	Unknown	No record
1069	12.96	2028.25	198	Unknown	No record
1071	13.06	2043.16	132	Unknown	No record
1075	13.05	2042.68	306	Unknown	No record
1078	13.05	2042.03	143	Unknown	No record
1081	13.06	2043.86	117	Unknown	No record
1086	13.11	2050.99	252	Unknown	No record
1089	13.19	2064.55	147	N-Acetyl-D-glucos	OVRNDRQMDRJTHS-RT
				amine;	RLPJTCSA-N
1				,	
				GC-EI-TOF; MS; n	
				GC-EI-TOF; MS; n TMS; RT	
1099	13.27	2076	393	GC-EI-TOF; MS; n TMS; RT Unknown	No record
1099 1103	13.27 13.27	2076 2076.69	393 147	GC-EI-TOF; MS; n TMS; RT Unknown Inositol;	No record CDAISMWEOUEBRE-GP

				TMS; RT	
1105	13.33	2086.26	245	Unknown	No record
1106	13.34	2088.29	205	Unknown	No record
1115	13.39	2095.06	146	Unknown	No record
1116	13.39	2094.83	338	Ferulic acid;	KSEBMYQBYZTDHS-U
				GC-EI-TOF; MS; n	HFFFAOYSA-N
				TMS; RT	
1117	13.39	2095.45	323	Unknown	No record
1118	13.41	2098.53	252	Unknown	No record
1128	13.49	2112.18	157	Unknown	No record
1136	13.53	2118.91	103	Unknown	No record
1138	13.57	2125.11	352	Unknown	No record
1141	13.61	2131.97	87	Unknown	No record
1142	13.61	2132.46	204	Unknown	No record
1143	13.62	2132.68	217	Unknown	No record
1147	13.62	2133.83	219	Caffeic acid;	QAIPRVGONGVQAS-UH
				GC-EI-TOF; MS; n	FFFAOYSA-N
				TMS; RT	
1148	13.63	2135.32	397	Unknown	No record
1157	13.72	2150.33	290	Unknown	No record
1160	13.75	2154.45	144	Unknown	No record
1161	13.76	2156.55	117	Unknown	No record
1164	13.76	2156.63	204	Unknown	No record
1165	13.76	2156.23	129	Unknown	No record
1166	13.79	2161.01	103	Unknown	No record
1172	13.8	2163.92	156	Unknown	No record
1175	13.82	2166.31	252	Unknown	No record
1178	13.83	2168.77	204	Unknown	No record
1183	13.91	2181.49	156	Unknown	No record
1190	13.99	2193.9	147	Unknown	No record
1193	13.99	2193.98	320	Unknown	No record
1194	13.99	2194.62	235	Unknown	No record
1195	13.99	2194.79	131	Unknown	No record
1208	14.12	2217.36	202	L-Tryptophan;	UMQXPTSGLUXAQK-S
				GC-EI-TOF; MS; n	ECBINFHSA-N
				TMS; RT	
1209	14.14	2219.19	103	Unknown	No record

1216	14.2	2230.55	204	Unknown	No record
1225	14.27	2241.49	117	Stearic acid;	QIQXTHQIDYTFRH-UH
				GC-EI-TOF; MS; n	FFFAOYSA-N
				TMS; RT	
1226	14.27	2241.37	341	Unknown	No record
1232	14.33	2253.25	204	Unknown	No record
1248	14.44	2271.2	203	Unknown	No record
1249	14.44	2271.42	133	Unknown	No record
1250	14.45	2272.8	104	Unknown	No record
1251	14.44	2271.61	217	Unknown	No record
1252	14.45	2272.49	129	Unknown	No record
1255	14.45	2272.71	206	Unknown	No record
1257	14.47	2276.66	246	Unknown	No record
1258	14.49	2278.97	131	Unknown	No record
1259	14.49	2279.32	103	Unknown	No record
1264	14.49	2280.07	204	Unknown	No record
1265	14.5	2280.81	217	D-Fructose	GSXOAOHZAIYLCY-UH
				6-phosphate;	FFFAOYNA-N
				GC-EI-TOF; MS; n	
				TMS; RT	
1267	14.5	2281.66	316	Unknown	No record
1270	14.5	2281.48	315	Unknown	No record
1271	14.52	2285.12	98	Unknown	No record
1277	14.58	2294.87	387	D-Fructose	GSXOAOHZAIYLCY-UH
				6-phosphate;	FFFAOYNA-N
				GC-EI-TOF; MS; n	
				TMS; RT	
1284	14.68	2313.74	217	Unknown	No record
1289	14.7	2316.16	147	Unknown	No record
1294	14.76	2327.59	144	Unknown	No record
1313	14.92	2356.69	204	Unknown	No record
1317	14.98	2367.63	217	Unknown	No record
1321	15.01	2373.17	91	Unknown	No record
1322	15.03	2375.85	103	Unknown	No record
1340	15.26	2417.72	185	Unknown	No record
1342	15.3	2425.32	204	Unknown	No record
1343	15.3	2426.1	103	Unknown	No record
1345	15.3	2426.04	147	Unknown	No record

1246	15 21	2427.90	460	I Imlen aven	No magand
1346	15.31	2427.89	462	Unknown	No record
1354	15.37	2438.06	147	Uridine;	DRTQHJPVMGBUCF-XV
				GC-EI-TOF; MS; n	FCMESISA-N
				TMS; RT	
1357	15.45	2452.94	217	Uridine;	DRTQHJPVMGBUCF-XV
				GC-EI-TOF; MS; n	FCMESISA-N
				TMS; RT	
1362	15.48	2459.81	361	Unknown	No record
1370	15.53	2469.06	144	Unknown	No record
1378	15.6	2480.45	204	Unknown	No record
1379	15.6	2480.45	204	Unknown	No record
1382	16.51	2660.35	147	beta-Lactose;	GUBGYTABKSRVRQ-U
				GC-EI-TOF; MS; n	HFFFAOYNA-N
				TMS; RT	
1392	16.54	2665.51	218	Unknown	No record
1401	16.59	2675.03	217	beta-Lactose;	GUBGYTABKSRVRQ-U
				GC-EI-TOF; MS; n	HFFFAOYNA-N
				TMS; RT	
1415	16.69	2696.27	191	Unknown	No record
1425	16.76	2709.75	204	D-(+)-Maltose;	GUBGYTABKSRVRQ-PI
				GC-EI-TOF; MS; n	CCSMPSSA-N
				TMS; RT	
1426	16.78	2713.83	191	D-(+)-Trehalose;	HDTRYLNUVZCQOY-LI
				GC-EI-TOF; MS; n	ZSDCNHSA-N
				TMS; RT	
1432	16.82	2722.48	217	Unknown	No record
1439	16.86	2732.22	205	Lactitol;	VQHSOMBJVWLPSR-U
				GC-EI-TOF: MS: n	HFFFAOYNA-N
				TMS; RT	
1442	16.9	2738.89	345	Unknown	No record
1444	16.9	2738.74	204	D-(+)-Maltose:	GUBGYTABKSRVRO-PI
				GC-EI-TOF: MS: n	CCSMPSSA-N
				TMS· RT	
1445	16.9	2740.2	204	Unknown	No record
1446	16.9	2740 28	362	Unknown	No record
1448	16.91	2742.29	283	Unknown	No record
1455	16.93	2746.65	103	D-(+)-Turanose	SEWFWILIOVIHATO-UH
1100	10.75	2710.05	105	GC-FI-TOF MS n	FFFAOYNA-N
1	1	1	1		11110111111

				TMS; RT	
1463	16.98	2756.15	324	Guanosine;	NYHBQMYGNKIUIF-UU
				GC-EI-TOF; MS; n	OKFMHZSA-N
				TMS; RT	
1464	17.01	2761.57	319	Unknown	No record
1480	17.11	2784.02	147	Unknown	No record
1482	17.11	2784.2	191	Unknown	No record
1487	17.12	2786.32	103	Unknown	No record
1494	17.19	2800.11	236	5-Methylthioadeno	WUUGFSXJNOTRMR-IO
				sine; GC-EI-TOF;	SLPCCCSA-N
				MS; n TMS; RT	
1511	17.31	2826.22	204	Unknown	No record
1515	17.34	2831.61	204	Melibiose;	DLRVVLDZNNYCBX-A
				GC-EI-TOF; MS; n	BXHMFFYSA-N
				TMS; RT	
1516	17.33	2831.4	204	Unknown	No record
1521	17.37	2838.65	373	Unknown	No record
1522	17.37	2838.98	373	Unknown	No record
1523	17.37	2839.67	169	Unknown	No record
1527	17.38	2841.59	204	Unknown	No record
1537	17.49	2865.44	204	Melibiose;	DLRVVLDZNNYCBX-A
				GC-EI-TOF; MS; n	BXHMFFYSA-N
				TMS; RT	
1538	17.53	2874.85	204	Unknown	No record
1539	17.54	2876.48	205	Unknown	No record
1542	17.54	2875.72	103	Unknown	No record
1543	17.55	2877.93	204	Unknown	No record
1549	17.58	2884.43	315	Unknown	No record
1554	17.65	2899.65	297	Unknown	No record
1563	17.71	2912.35	204	Unknown	No record
1565	17.73	2918.5	204	Unknown	No record
1575	17.83	2939.63	204	Unknown	No record
1579	17.84	2943.41	345	Unknown	No record
1590	17.95	2967.55	255	Chlorogenic acid;	CWVRJTMFETXNAD-Z
				GC-EI-TOF; MS; n	NEHSRBWSA-N
				TMS; RT	
1591	17.96	2968.39	255	Unknown	No record
1592	17.96	2968.59	191	Unknown	No record

1595	17.96	2969.85	204	Unknown	No record
1599	17.99	2976.61	315	Unknown	No record
1601	18.01	2981.26	204	Unknown	No record
1605	18.03	2985.44	105	Unknown	No record
1611	18.04	2987.56	120	Unknown	No record
1612	18.04	2987.67	105	Unknown	No record
1620	18.09	2998.52	315	Unknown	No record
1633	18.34	3055.69	169	Unknown	No record
1640	18.37	3063.44	192	Unknown	No record
1644	18.37	3063.69	103	Unknown	No record
1646	18.43	3076.69	147	Unknown	No record
1647	18.43	3076.64	147	Unknown	No record
1648	18.43	3077.87	345	Unknown	No record
1650	18.44	3078.37	204	Unknown	No record
1656	18.49	3090.37	147	Unknown	No record
1657	18.49	3092.01	255	Unknown	No record
1662	18.5	3093.61	133	Unknown	No record
1663	18.5	3093.35	255	Chlorogenic acid;	CWVRJTMFETXNAD-Z
				GC-EI-TOF; MS; n	NEHSRBWSA-N
				TMS; RT	
1665	18.51	3096.75	398	Unknown	No record
1669	18.59	3112.9	91	Unknown	No record
1676	18.71	3141.91	249	Unknown	No record
1680	18.74	3148.61	307	Unknown	No record
1683	18.75	3150.92	307	Unknown	No record
1694	18.85	3174.73	307	Unknown	No record
1715	19.15	3239.84	259	Unknown	No record
1746	19.64	3338.97	451	Unknown	No record
1749	19.66	3342.58	147	Unknown	No record
1753	19.68	3346.85	207	Unknown	No record
1757	19.69	3348.58	217	Unknown	No record
1762	19.7	3350.33	204	D-(+)-Raffinose;	MUPFEKGTMRGPLJ-ZQ
				GC-EI-TOF; MS; n	SKZDJDSA-N
				TMS; RT	
1766	19.91	3392.04	147	Unknown	No record
1769	19.92	3393.97	204	Unknown	No record

1772	19.93	3395.91	362	Unknown	No record
1809	20.22	3445.98	217	D-(+)-Melezitose;	QWIZNVHXZXRPDR-W
				GC-EI-TOF; MS; n	SCXOGSTSA-N
				TMS; RT	
1827	20.25	3452.19	217	Unknown	No record
1836	20.29	3458.41	297	Unknown	No record
1881	20.68	3522.74	362	Maltotriose;	FYGDTMLNYKFZSV-U
				GC-EI-TOF; MS; n	HFFFAOYNA-N
				TMS; RT	