

ノ ー ト

高速液体クロマトグラフ-四重極飛行時間型質量分析計を用いた清酒の品質予測

牧野 正知*・大渡 康夫**・秋吉 渚月**・田畑 光正*

1. 目 的

清酒には原材料ならびに発酵に関与する酵母・麹菌など微生物由来の代謝産物（以下、メタボライトと略す）が280種類以上含まれており、それらが清酒の香味を特徴付けている¹⁾。メタボライトの種類やその濃度は、使用する米の種類やその精米歩合、醸造に用いる酵母や麹菌の株といった原材料の特徴のほか、米の吸水歩合、麴の破精込み状態や、醸造時の温度や時間といった醸造パラメータによって影響を受ける。そのためメタボライトの全体像と醸造パラメータの関係に対する理解とその利用は、清酒の香味特性や酒質を科学的にコントロールする新しい醸造技術として期待されている。そのような背景から島根県産業技術センターでは、ガスクロマトグラフ-質量分析計（以下、GCMSと略す）を用いメタボライトに注目した清酒の客観的評価法の開発に取り組んできた²⁾。

GCMSは、比較的高分解能で保持時間（以下、RTと略す）再現性の高い測定が可能で、気体を移動相とする仕組みから揮発性メタボライトの測定に適している。またアミノ酸、糖や核酸塩基など不揮発な親水性メタボライトのうちいくつかは、誘導体化することにより測定可能である。しかしメタボライトの中には誘導体化できない物質あるいは誘導体化しても揮発しないもの、また熱分解するものもあり、GCMSによって試料に存在する全てのメタボライトの測定をカバーすることは難しい。そのため分析対象の化学的特徴に応じて別の測定系である高速液体クロマトグラフ-質量分析計（以下、LCMSと略す）、キャピラリー電気泳動-質量分析計や核磁気共鳴装置なども用いられてきた。

LCMSは、保持時間の揺らぎや汎用同定MSスペクトルライブラリが利用できないなどの課題があるものの、幅広い化学的特性に対して適応可能で、また誘導体化など前処理が必要ないことから食品分野での測定に活用されてきた。特に清酒などの液体試料の場合、希釈してそのまま測定に供することができるという簡便性のため、近年は親水性相互作用クロマトグラフィーカラムを用いたLCMSに

よって清酒オリゴ糖が³⁾、またLC部に低流量液体クロマトグラフを用いたLCMSによって山廃仕込み清酒のメタボライトが分析された⁴⁾。岩下らの研究グループは、逆相クロマトグラフィーカラムを用いたLCMSによって198種の清酒メタボライトを同定する方法を構築し、複数の醸造条件から得られた清酒を分析することで醸造パラメータとメタボライトの関係を明らかにした⁵⁾。

このような背景から、筆者らはこれまでに取り組んできたGCMSによる清酒の評価方法に、新たにLCMS測定を組み合わせることによって、より多角的に清酒の特徴を機器分析で捉えることができると考えた。そこで本研究は平成28酒造年度島根県新酒技術研究会に出品された清酒を試料としてLCMS測定し、官能評価で分類された品質、香気成分量、吟醸酒と純米酒の特徴などが清酒LCMSデータから説明できるか検証したので報告する。

2. 試料および方法

2.1 分析試料

2.1.1 清酒試料

本研究では、平成28酒造年度島根県新酒技術研究会（以下、研究会と略す）において島根県内の酒造メーカーより「吟醸酒の部」と「純米酒の部」に出品された合計134点の清酒を試料とした（以下、「吟醸酒の部」と「純米酒の部」にそれぞれ出品された清酒を吟醸酒と純米酒と略す）。酒造メーカーによっては、同一ロット（醪）に由来する清酒が複数出品された（例：斗瓶違い）。これらの品質差については本研究では取り扱わないこととし、研究会における官能評価の結果、同一メーカー酒のうち指摘事項が重複し後述する品質クラスの近いものについては、本研究の研究対象から除外することとした。

2.1.2 清酒モデル試料

データを統合するプロセスの一つであるデータアライメントでは、RTと質量分析計の質量電荷比（以下、 m/z と略す）のそれぞれの許容範囲の設定が必要である。それらを求めるため清酒に含まれる成分¹⁾の一部を含む以下の溶液（以下、清酒モデル試料と略す）を調製し、これを繰り返し測定して評価を行った。15% (v/v) エタノール、1.4% (w/v) グルコース、110 mg/L L-アルギニン、110 mg/L

*食品技術科, **生物応用科

表1 質量分析計の測定条件

GCMS	
装置	GCMS 7000B (Agilent) TurboMatrixHS40 (PerkinElmer)
ヘッドスペースサンプラー	
抽出法	60°C, 24 min
注入口温度, イオン源温度	250°C, 200°C
トランスファーライン温度	230°C
MRMトランジション	
アセトアルデヒド	44 <i>m/z</i> > 43 <i>m/z</i>
酢酸エチル	70 <i>m/z</i> > 55 <i>m/z</i>
イソバレルアルデヒド	86 <i>m/z</i> > 58 <i>m/z</i>
n-プロピルアルコール	60 <i>m/z</i> > 59 <i>m/z</i>
イソブチルアルコール	74 <i>m/z</i> > 41 <i>m/z</i>
酢酸イソアミル	70 <i>m/z</i> > 55 <i>m/z</i>
イソアミルアルコール	70 <i>m/z</i> > 55 <i>m/z</i>
カプロン酸エチル	115 <i>m/z</i> > 87 <i>m/z</i>
フェネチルアルコール	122 <i>m/z</i> > 91 <i>m/z</i>
LCMS	
装置	Nexera (Shimadzu) micrOTOF-QII (Bruker Daltonics)
ドライガス流量, 温度	8.0 L/min, 200°C
イオン源, 極性, 電圧	ESI, ポジティブ, 4500 V
ネブライザー	0.16 MPa
サンプリングレード, 範囲	1.0 s ⁻¹ , 50–1000 <i>m/z</i>

L-チロシン, 130 mg/L L-プロリン, 120 mg/L L-アスパラギン酸, 120 mg/L L-メチオニン, 100 mg/L ロイシン, 120 mg/L DL-ピログルタミン酸, 110 mg/L パントテン酸カリウム, 120 mg/L バニリン, 110 mg/L L-アスコルビン酸, 130 mg/L チアミン, 7.0 mg/L リボフラビン.

2.2 官能評価

官能評価は研究会にて6名の審査員が審査カード⁶⁾に従って実施した。本研究における官能評価データはここで記録された官能評価プロファイルとした。また本研究における清酒の品質クラスは、研究会の審査では総合評価の悪いものを脱落させながら複数回(純米酒の部は3回, 吟醸酒の部は4回)審査されたことを反映させ、審査された回数とした。すなわち1回目の審査で脱落した清酒の品質クラスは1, 4回目の審査まで残ったものは4とした。

2.3 GCMS分析

試料0.9 mLは内部標準溶液(アセトン-d₆, 1-ペンタノール, カプロン酸メチルの混合液)0.1 mLとガラスバイアル中で混合し, ヘッドスペースサンプラーを接続したトリプル四重極型GCMSにて分析した。質量分析計の条件は表1に示した。使用したカラムはDB-WAX (20 m,

0.18 mm, 0.18 μm)で, 溶出条件としてオープンの温度を0.00–1.00 min : 35°C, 1.00–6.25 min : 35–140°C, 6.25–9.25 min : 140–230°C, 9.25–11.3 min : 230°Cとした。定量した成分は以下のとおり: アセトアルデヒド, 酢酸エチル, イソバレルアルデヒド, n-プロピルアルコール, イソブチルアルコール, 酢酸イソアミル, イソアミルアルコール, カプロン酸エチル, フェネチルアルコール。

2.4 グルコース濃度測定

清酒に含まれるグルコース濃度は, グルコースCIIテストワコー(富士フィルム和光純薬)で測定した。値は重量%として求めた。

2.5 LCMS分析

試料は2.3 mg/L カフェイン(和光純薬)含有メタノールにて10倍希釈し, 遠心分離して得られた上澄み液とした。用いた質量分析計は四重極飛行時間型で, その分析条件を表1に示した。分離に用いたカラムはACQUITY UPLC HSS T3 (1.8 μm, 2.1 mm × 150 mm, Waters)で, 0.1%ギ酸含有水と0.1%ギ酸含有アセトニトリルによるグラジエント溶出とした。溶出条件は0.0–5.0 min: 0%, 5.0–15 min: 0–100%とし, オープンは40°C, 流速は0.3 mL/minとした。繰り返し測定は, 全ての試料を一通り測定してからそれを更に2回繰り返す方法とした(測定セット1, 2, 3)。このとき清酒モデル試料は10回の清酒試料測定後に1回測定されることとし, このシーケンスを測定セット内のサブセットと設定した。各測定セットにおけるサブセットの順序は, 質量分析計の感度振れや日間誤差の影響を最小化するためランダム化した。なおいずれの測定においても質量電荷比補正用に測定開始後0.0–0.5 minにわたり5 mmol/Lのギ酸ナトリウムを質量分析計に導入した。

2.6 LCMSデータの統合

2.6.1 アライメントパラメーターの決定

はじめに全てのLCMSデータの質量電荷比を, DataAnalysis 4.0 (Bruker Daltonics)により同一データ内のギ酸ナトリウムマスペクトルを基準に補正した。次に清酒モデル試料を測定したLCMSデータをMS-DIAL⁷⁾で解析し, 標準成分ごとにRTとプレカーサイオンの*m/z*を抽出した。この操作を全ての清酒モデル試料(n=22)で行い, 標準成分ごとにRTと*m/z*を整理してそれぞれの標準偏差を求めた。本研究では標準成分間で標準偏差を比較してから最大値を求め, それを5倍した値をピークアライメントにおけるRTと*m/z*のそれぞれの揺らぎの許容範囲とした。

2.6.2 データセットの作成

次に清酒試料を測定したLCMSデータのピーク検出をMS-DIALによって行った。先に求めた許容範囲にてデータ間のアライメントを行い, RTと*m/z*の組み合わせたデータポイント(以下, マーカーと略す)を定めた。次に試料ごとにマーカー×ピークエリア面積で整理し, 3次元構造

表2 清酒試料の品質クラスとグルコース濃度および香気成分の統計量

清酒の属性	吟醸酒				純米酒			
	品質クラス	1	2	3	4	1	2	3
試料数 (測定データ数)		12(36)	11(33)	9(27)	8(24)	11(33)	9(27)	6(18)
項目	平均値*	最小値	最小値	最大値	平均値*	最小値	最大値	
グルコース(%(w/w))	2.4±0.6	1.5		4.0	2.0±1.1	0.50		4.9
アセトアルデヒド(mg/L)	75±18	48		120	76±34	30		160
酢酸エチル(mg/L)	49±23	25		130	87±43	26		180
イソバレルアルデヒド (mg/L)	0.1±0.1	0.0		0.3	0.1±0.1	0.0		0.7
n-プロピルアルコール(mg/L)	56±13	42		92	65±15	32		110
イソブチルアルコール(mg/L)	39.0±8.0	29		72	52±16	30		86
酢酸イソアミル(mg/L)	2.0±1.0	0.90		6.1	3.5±2.2	1.1		9.0
イソアミルアルコール(mg/L)	110±16	88		160	140±25	98		180
カブロン酸エチル(mg/L)	13.0±4.6	1.2		24	6.6±6.2	1.2		25
フェネチルアルコール(mg/L)	170±45	110		270	160±43	61		260

* 平均値±標準偏差

(試料名×マーカー×ピークエリア面積)を有するデータマトリックスがLCMSデータセットとして出力された。

2.6.3 データリダクション

測定再現性の乏しいマーカーを解析から除外するため、試料内 (n=3) のマーカーごとにピークエリア面積値の変動係数 (以下, CVと略す) を計算した。次に試料間で各マーカーに対応する最大CV (CV_{max}) を求めた。本研究ではCV_{max}>20を示すマーカーを除外することとした。

2.7 統計解析

統計解析のためLCMSデータセット、清酒の属性 (吟醸酒あるいは純米酒)、官能評価データ、品質クラス、香気成分とグルコース濃度のそれぞれをマイクロソフトExcelにより統合した。以降の統計解析とグラフ作成はプログラミング言語R⁸⁾を用いた。はじめに吟醸酒属性と純米酒属性のそれぞれのデータセット (以下、吟醸酒データセットと純米酒データセットと略す) に対し、LCMSデータを説明変数とした主成分分析 (Principal Component Analysis, 以下PCAと略す) とUMAP (Uniform Manifold Approximation and Projection, version 0.2.7.0, 教師なしモデル) で解析し、次元縮約した結果と品質クラス及び香気成分を照らし合わせ、それらの関係を考察した。次にLCMSデータセットから清酒の属性を判別できるかを試みた。判別分析は直交部分最小二乗判別分析 (Orthogonal Partial Least Squares-Discriminant Analysis, 以下, OPLS-DAと略す) を用いた。このとき目的変数と説明変数をそれぞれ清酒の属性とLCMSデータとし、7セグメントによる交差検証とした。次にグルコース濃度と香気成分がLCMSデータから予測できるか部分的最小二乗回帰 (Partial Least Squares Regression, 以下, PLSと略す) で検証した。本検討ではLCMSデータセットをランダムに二分割して学習と検証

データセットとした。ただし繰り返し測定 (n=3) における同一試料は同じ学習あるいは検証区分になるようにした。まず学習データセットから10の潜在変数で予測モデルを構築し、それに学習と検証データセットをそれぞれ入力して、得られた目的変数 (予測値) と実測値のピアソンの積率相関係数の値 (以下, R値と略す) を求めた。最後に品質クラスがPLSとサポートベクトルマシン (以下, SVMと略す) の組み合わせで予測できるかを検証した。吟醸酒と純米酒データセットのそれぞれのLCMSデータを上記と同様に学習と検証データセットに分け、PLSによって二乗平均平方根誤差 (Root Mean Square Error of Estimation, 以下, RMSEEと略す) が最小となる潜在変数の数で次元縮約した。次に次元縮約した各試料のPLSスコアに対してSVM (kernlab, version 0.9-29) で学習・分類させ、検証データセットにおける品質クラスの実際の値と予測値の一致度を予測精度とした。

3. 結果および考察

3.1 官能評価, グルコース濃度と香気成分

6人の審査員によって審査された134点の出品酒のうち、官能評価の結果を参考にして分析に供する試料を選択した結果、吟醸酒40点、純米酒26点の合計66点を研究試料とした。このとき品質クラス、グルコース濃度及び香気成分の平均値と標準偏差は表2のとおりであった。

3.2 データ統合

清酒モデル試料を繰り返し測定 (n=22) して得られた標準物質の [M+H]⁺ イオンの m/z と RT を集計した。標準物質ごとにそれぞれ標準偏差を求め比較したところ、最大のそれらはそれぞれ 0.0020 m/z (リボフラビン) と 0.0080 min (チアミン) であった。これらの5倍値を許容範囲として清

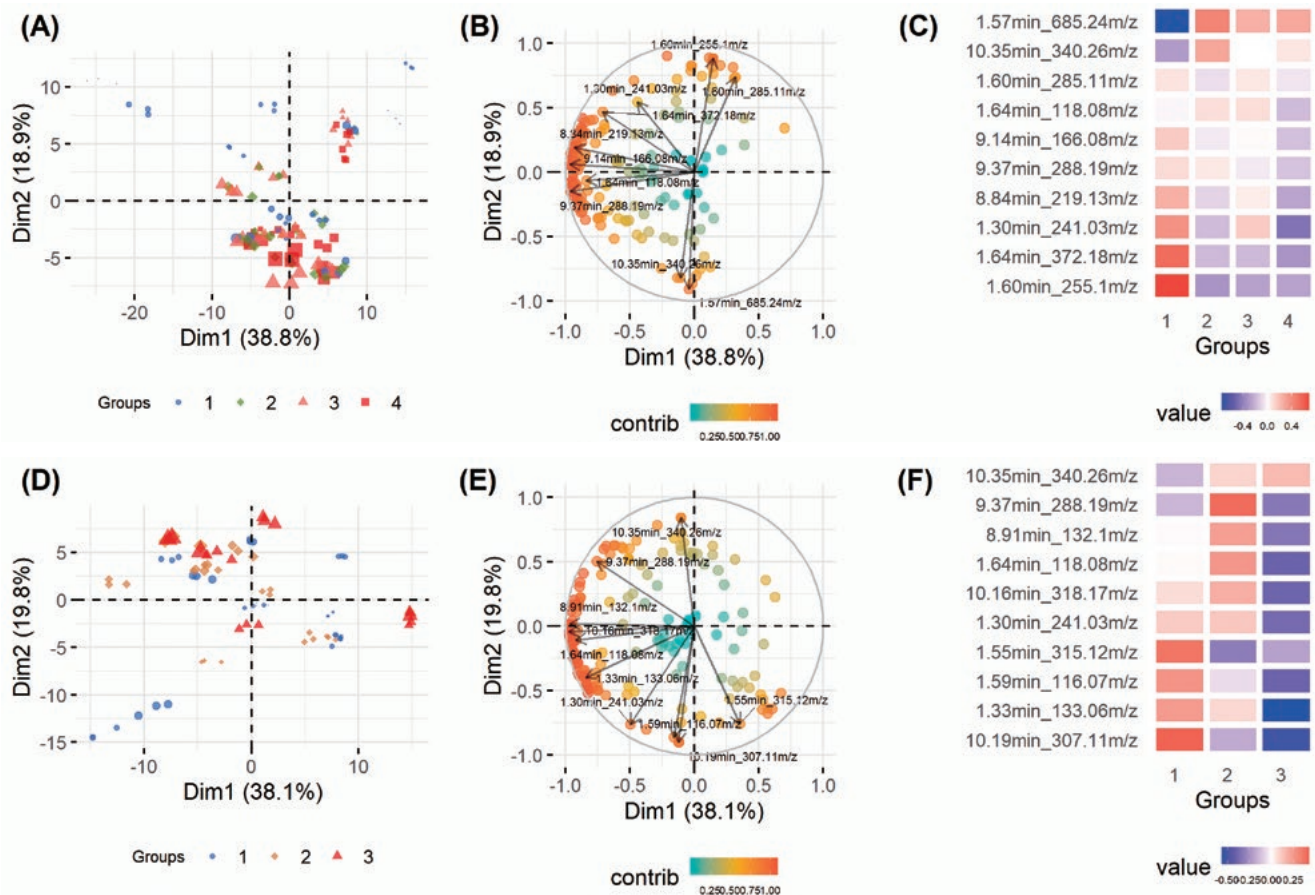


図1 主成分分析 (第1, 2主成分, A, B, C: 吟醸酒データセット, D, E, F: 純米酒データセット)

(A, D) スコアプロット: ポイントの色彩とポイントのサイズはそれぞれ品質クラスとグルコース濃度に対応している. (B, E) ローディングプロット: 合計 149 マーカーを相対寄与率で色付けし, そのうち寄与率の大きなものについてはベクトル表示とともにマーカー名を示した. (C, F) ローディングプロットでベクトル表示したマーカーにおける正規化面積値.

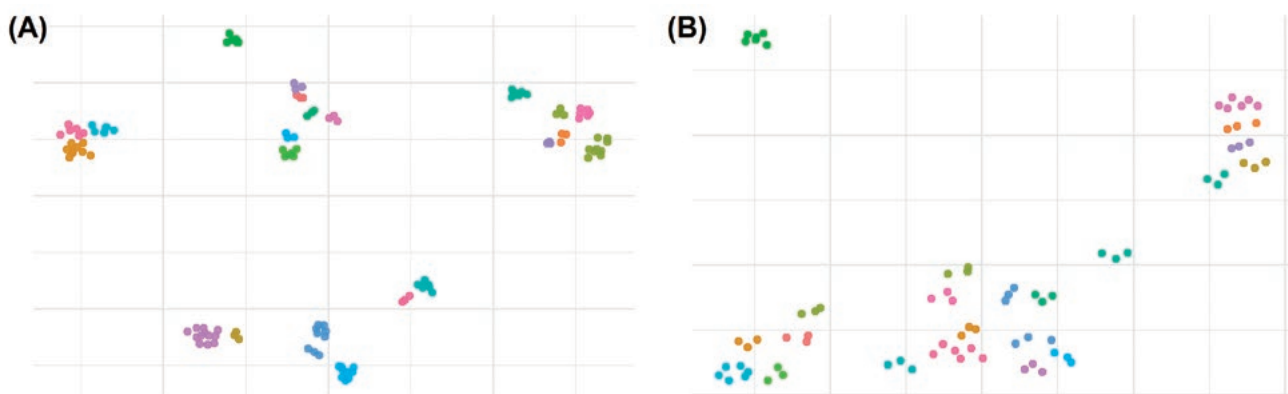


図2 UMAP を用いた次元縮約

(A) 吟醸酒データセット, (B) 純米酒データセット, プロットの色は酒造メーカーを表している.

酒試料から得られた LCMS データに対してアライメントした結果, 5,519 マーカー (RT_m/z) × 66 (試料数) × 3 (繰り返し数) のデータセットが得られた. CV 値を基準に解析に用いるマーカーを選択した結果, 最終的なマーカー数は 149 となった.

3.3 次元削減法を用いた測定データの可視化

はじめに品質クラスが LCMS データに反映されているかを PCA によって調べた. PCA は教師なし機械学習に分類される古典的な線形次元縮約の解析方法である. 今回の解析では PCA によって清酒試料の分布状況を求めると

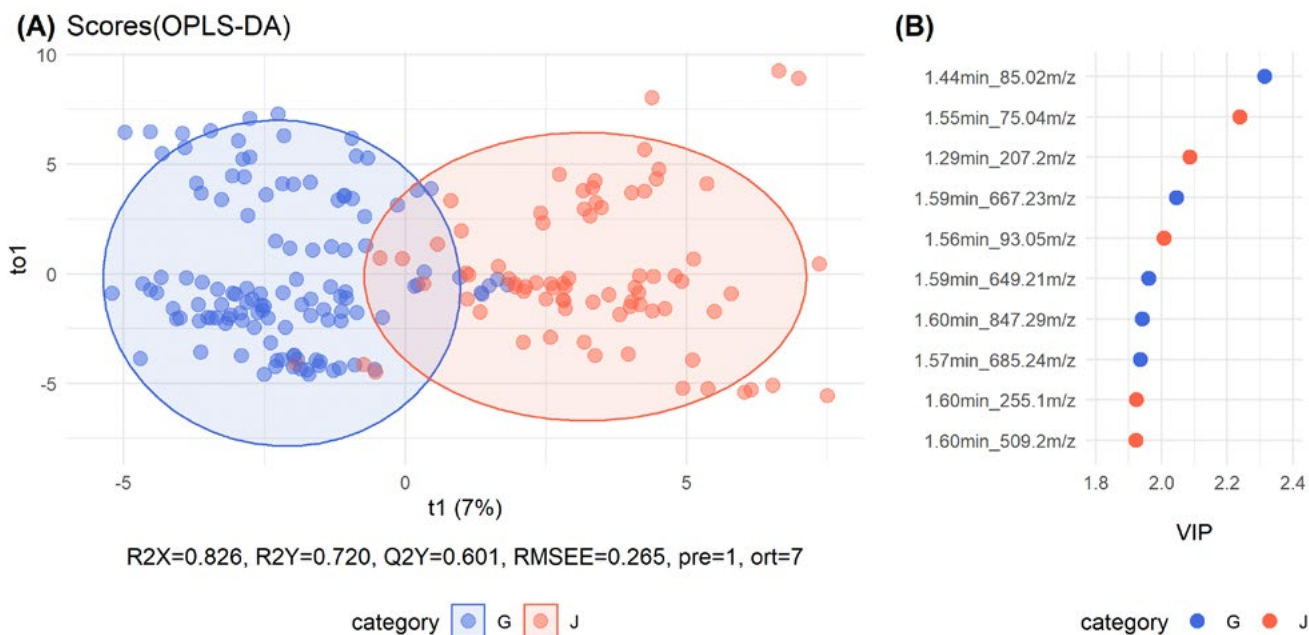


図3 判別分析

OPLS-DA を用いた吟醸酒と純米酒データセットの判別分析結果。(A) スコアプロット：各試料の第1, 第2成分のスコアをそれぞれGとJラベルしてプロット。(B) VIP：上位10位までのVIPとそのスコアを示した。色は面積値が大きい方の属性を示している。

ともに、ばらつきやグループ間の傾向、ひいてはデータの背後に存在する本質的な構造の抽出を試みた。得られたスコアプロットとローディングプロットはそれぞれ図1(A,D)と図1(B,E)となった。なおスコアプロットとローディングプロットのそれぞれの軸の方向は対応関係にあり、またローディングプロットは各マーカーの主成分軸に対する重みを原点からベクトルで表している。第1主成分と第2主成分に対する寄与率は吟醸酒と純米酒データセットとでそれぞれ38.8%, 18.9%と、38.1%, 19.8%となり、共に約6割の情報が縮約されたといえる。また吟醸酒データセットにおける第2主成分は、グルコース濃度と負の相関があり ($p < 0.01$), その相関係数は -0.651 であった。清酒のグルコース濃度はアミラーゼ系酵素の供給源である麹菌の影響を受けるため、第2主成分軸方向には麹菌に関する情報が縮約されていると考えられる。なお、図1(A)において原点付近のほか第1象限に一部の試料が集中しているが、これらはアミラーゼ系酵素低活性タイプの麹菌を使用した清酒である可能性が高い。品質クラスに着目すると、品質クラス2以上の試料はそれぞれのスコアプロットにおいて原点付近に集中していた。一方、品質クラス1の試料は第1主成分、第2主成分とも原点から離れた位置にプロットされた。次にローディングプロットにおいて相対寄与率の高いマーカー(図1(B,E)におけるベクトル表示したマーカー)について注目し、品質クラスごとにそれらの面積値(平均値)を正規化して比較した結果を図1(C,F)に示した。品質クラスに対応して面積値(平均値)は単純増加あるいは単純減少、または凹凸であった。

これらのマーカーと面積値は、LCMSデータを用いて品質クラスを判断するうえでの指標の一つと示唆される。なお品質クラス2以上については、スコアプロットの分布においてこれ以上注目すべき特徴を見出すことは難しく、第5主成分(累積寄与率79.4%)まで確認したが同様であった。

次に最新の機械学習による非線形次元削減分類方法の一つで、リーマン幾何学と代数的位相幾何学に基づいた理論的フレームワークにより次元を圧縮するUMAP⁹⁾を採用して、吟醸酒と純米酒データセットがそれぞれどのように次元縮約して分類されるかを調べることにした。UMAPで次元縮約された結果は図2となり、複数の酒造メーカー試料で構成される集団をいくつか形成した。このとき品質クラスや醸造に使用した麹菌と酵母の種類を反映するグルコース濃度とカプロン酸エチル濃度をそれぞれ指標に対応関係を調べたが、分布上目立った特徴は見出せなかった。しかし本結果は、集団内の試料において共通する未知の要素の存在を示唆している。

3.4 判別分析

清酒の属性がLCMSデータから判別できるかをOPLS-DAによって検証した。得られた予測モデルにおける予測性能の指標である Q^2 値と予測残差を示すRMSEE値はそれぞれ0.601と0.265になり、このとき潜在変数の数は7であった。一般的にモデルの評価は $Q^2 > 0.5$ でgood, > 0.9 でexcellentと見なされ¹⁰⁾、本モデルはある程度の予測性能があるものと考えられる。スコアプロットは図3(A)のようになり、清酒の属性に応じて第1成分方向に対してそれぞれ離れるようにプロットされた。しかし原点付近にプ

ロットされた試料もあり、これらは今回構築したモデルによる判別は難しいことを示している。研究会には純米酒であっても高精白で吟醸造りしたものが「純米酒の部」として出品されたケースもあるようで、判別が難しい試料については、そのような造りをした清酒が該当する可能性が高い。

また清酒の属性を判別する指標として、OPLS-DA から得られた説明変数（マーカー）の重要度（Variable Importance in Projection, 以下、VIP と略す）に注目した。10 位までの VIP と、それが清酒の属性のどちらにおいて強く観測されたかについて図 3 (B) で示した。例えば吟醸酒属性と純米酒属性の特徴的なマーカーとして、それぞれ 1.44min_85.02m/z, 1.55min_75.04m/z が見出された。今回の結果は LCMS データから清酒の属性を判別するうえでの指標候補を示唆している。

3.5 グルコース濃度・香気成分と品質クラスの予測

LCMS データから PLS によってグルコース濃度と香気成分量を予測できるかを調べた。学習データセットを用いた PLS は表 3 (学習) の相関を示す予測モデルを構築した。これらの予測モデルのうち検証データセットにおいて最も精度良く予測できた成分はイソアミルアルコールで、その R 値 (以下、 R 値 (検証) と略す) は 0.931 ($p < 0.01$) であった (表 3, 図 4 (A, B))。そのほかグルコースや酢酸エチル, n-プロピルアルコール, イソブチルアルコール, 酢酸イソアミル, カブロン酸エチルでは R 値 (検証) が 0.70 を超えた (表 3)。一方、予測の難しい成分がフェネチルアルコールで、 R 値 (検証) は 0.180 となり (表 3, 図 4 (C, D))、検定の結果、相関はなかった ($p > 0.05$)。清酒の香気成分は醸造に使用する酵母の影響を強く受けることが知られており¹¹⁾、酵母に関する情報が清酒に含まれるメタボライトを介して LCMS で捉えられることは注目に値する。

品質クラスの予測は PLS-SVM 法を用いた。PLS は今回の研究対象のようなサンプルサイズよりも圧倒的に説明変数が多い場合や説明変数間の共線性が高い場合に有効とされる手法である。それを拡張した PLS-SVM 法は、PLS によって高次元データを目的変数と関連の強い低次元データ (PLS スコア) に変換し、それらを SVM により超平面で分類する方法である。最近では本手法を用いた食肉動物種の FTIR スペクトルによる識別が報告されている¹²⁾。なお本研究においては PLS-SVM 法で清酒の属性を分類させると、検証データセットにおける正解率は 80.0% となることを予め確認している。品質クラスの分類に際し、学習データセットを用いて RMSEE が最小となる PLS モデルを作成したところ、その潜在変数の数は 10 となった。次に学習データの PLS スコアセットで SVM 分類器を学習させたのち、検証データセットのそれを同様に分類器に入力して品質クラスを分類させたところ、吟醸酒と純米酒データセットにおける分類結果の単純な正解率は、それぞれ 40%

表 3 ピアソンの積率相関係数一覧

データセットの種類	学習	検証
グルコース	0.985*	0.764*
酢酸エチル	0.996*	0.870*
アセトアルデヒド	0.948*	0.330*
イソバレルアルデヒド	0.989*	0.651*
n-プロピルアルコール	0.986*	0.817*
イソブチルアルコール	0.990*	0.910*
酢酸イソアミル	0.994*	0.794*
イソアミルアルコール	0.991*	0.931*
カブロン酸エチル	0.985*	0.904*
フェネチルアルコール	0.969*	0.180

* $p < 0.05$

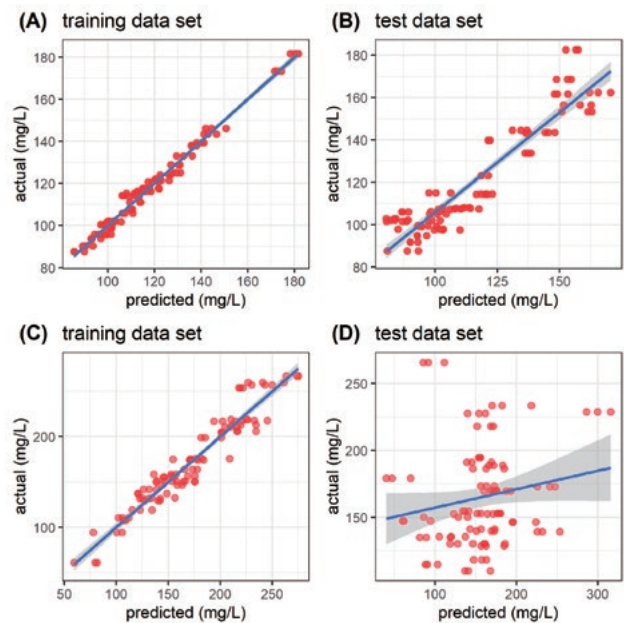


図 4 PLS による香気成分量の予測

(A,B) イソアミルアルコールと (C,D) フェネチルアルコールについて、学習と検証データセットのそれぞれの予測値と実測値をプロットした。回帰直線を青線で、また回帰直線の 95% 信頼区間を灰色で示した。

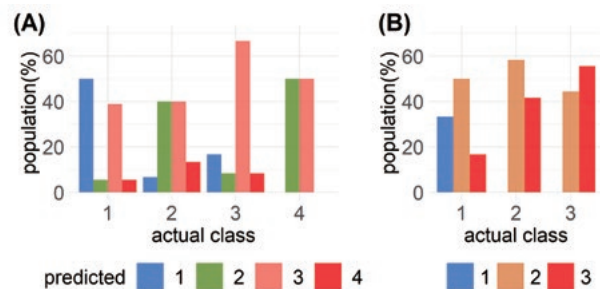


図 5 PLS-SVM 法による分類

(A) 吟醸酒データセット, (B) 純米酒データセット。

と46%であった。またそれぞれの品質クラスにおける分類予想は図5のようになった。吟醸酒データセットでは品質クラス3試料における分類一致率は66.7%で比較的高い結果となったものの、品質クラス4試料では一致率は0%となった(図5(A))。また純米酒データセットにおける品質クラスが2と3の試料について注目すると、1と予測されたものはなく、一致率はそれぞれ58%と55%となった(図5(B))。これらのことから、本研究で用いたLCMSデータから精度よく清酒の品質クラスを予測するには至らなかったが、一方で示唆的な結果も得られていると考えられた。

4. まとめ

本研究ではLCMS分析によって清酒の香味特徴を多角的にとらえることができるか、ひいては品質が予測できるか、それらの可能性について検証した。その結果、品質クラスや吟醸酒・純米酒の判別に関与するマーカー候補を見出した。またLCMSデータがイソアミルアルコールなどいくつかの香り成分の定量的な予測に適応できることを示した。PLS-SVM法による品質クラスの分類では、その可能性が示唆されたものの、精度よく予測するには至らなかった。試料の情報は確かに測定できているのか、またそれはデータリダクションや次元圧縮・機械学習時に欠損していないかなど検討すべき課題は残ったままである。データ分析の観点からはサンプル数を増やすことや別の機械学習の手法を検討することも重要で、そのためLCMSデータによる清酒品質評価・予測に関する可能性は今後も検討されなければならない。なお本研究においてマーカーとメタボライトの関係については踏み込まなかった。もちろんメタボライトの同定が進めば、それが何に由来するのか、どのようなプロセスで生産されるのかといった科学的理解に繋がり、ひいては高度な醸造技術の開発へ貢献すると考えられる。一方でLCMSデータをあえてそのまま使用することで清酒の品質予測や品質管理が可能になるのであれば、その利用は新しい価値創出に対する近道になり得るのかもしれない。

5. 謝 辞

LCMSを用いた清酒分析は酒類総合研究所の岩下和裕成分解析研究部門長、小林拓嗣研究員にご指導いただいた。清酒の官能評価には酒類総合研究所 品質評価部門 伊豆英恵主任研究員、広島国税局 鑑定官室 宮脇達也 鑑定官に、また試料調製の一部は島根県産業技術センターの太田

ゆかり研究助手、二瀬充幸研究助手に協力していただいた。ここに謝意を示します。

6. 文 献

- 1) 日本醸造協会 編. 醸造物の成分. 日本醸造協会. 1999, p.1-108.
- 2) 大渡康夫, 牧野正知, 松林和彦, 田畑光正. メタボローム解析を用いた食品成分の評価技術の確立(第1報) - メタボローム解析による日本酒の品質評価技術について -. 島根県産業技術センター研究報告. 2020, vol.56, p.7-12.
- 3) Tokuoka, M.; Honda, C.; Totsuka, A.; Shindo, H.; Hosaka, M. Analysis of the oligosaccharides in Japanese rice wine, sake, by hydrophilic interaction liquid chromatography-time-of-flight/mass spectrometry. *Journal of Bioscience and Bioengineering*. 2017, vol.124, p.171-177.
- 4) Tatsukami, Y.; Morisaka, H.; Aburaya, S.; Aoki, W.; Kohsaka, C.; Tani, M.; Hirooka, K.; Yamamoto, Y.; Kitaoka, A.; Fujiwara, H.; Wakai, Y.; Ueda, M. Metabolite profiling of the fermentation process of "yamahai-ginjo-shikomi" Japanese sake. *PLOS ONE*. 2018, vol.13, p.e0190040.
- 5) Yazawa, H.; Tokuoka, M.; Kozato, H.; Mori, Y.; Umeo, M.; Toyoura, R.; Oda, K.; Fukuda, H.; Iwashita, K. Investigation of relationship between sake-making parameters and sake metabolites using a newly developed sake metabolome analysis method. *Journal of Bioscience and Bioengineering*. 2019, vol.128, p.183-190.
- 6) 大渡康夫, 永田善明, 土佐典照, 田畑光正, 秋吉渚月. 味覚センサーによる清酒の評価とその活用方法の検証. 島根県産業技術センター研究報告. 2018, vol.54, p.27-32.
- 7) 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*. 2015, vol.12, p.523-526.
- 8) R Core Team; R: A language and environment for statistical computing. <https://www.R-project.org/>. 2019.
- 9) McInnes, L.; Healy, J.; Melville, J. UMAP: Uniform Manifold Approximation and Projection for Dimension Reduction. *ArXiv e-prints* 1802.03426, 2018.
- 10) 玉田佳大, 樺島文恵, 櫻井昌文, 徳井美里, 山下伸雄, 窪寺隆文, 明石貴裕. ノンターゲット GC/MS による清酒の「押し味」のモデル化. *生物工学会誌*. 2018, vol.96, p.234-239.
- 11) Tsutsumi, H. Mechanism of aroma production in Japanese sake. *Journal of Japan Association on Odor Environment*. 2015, vol.46, p.346-349.
- 12) Tsakanikas, P.; Karnavas, A.; Panagou, E. Z.; Nychas, G.-J. A machine learning workflow for raw food spectroscopic classification in a future industry. *Scientific reports*. 2020, vol.10, p.11212.