SHIMADZU
Excellence in Science

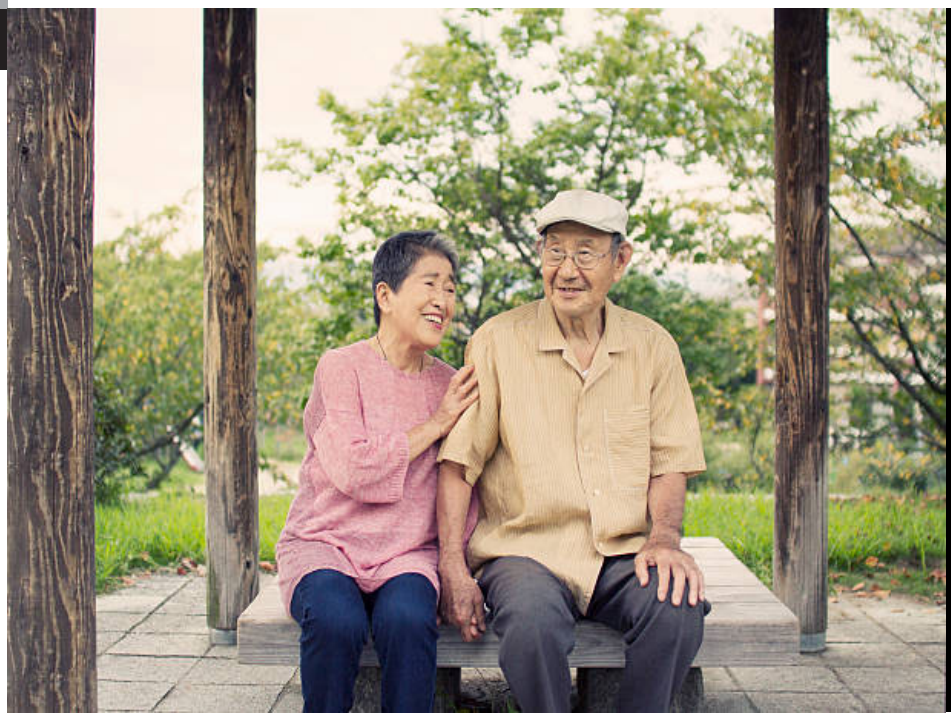
Application Note

No. 98

ライフサイエンス

長寿研究における次世代シーケンサーと質量分析計を用いたマルチオミクス解析

中川 裕貴[1]、飯伏 翼[2]、笠寺 浩介[3]、樫尾 宗志朗[4]



Life Science

■要旨

次世代シーケンサーと液体クロマトグラフ質量分析計システム（Nexera Mikros™を搭載したLCMS™-9050）を用いて長寿となる遺伝子変異を起こしたキイロショウジョウバエを野生型と比較しました。DNA解析では全ゲノム解析によりSNV/Indel83万箇所（構造変異は約4万箇所）、RNA解析では網羅的発現量解析によりp値が0.05以下の185遺伝子座を検出しました。タンパク質解析は網羅的解析により約1000個を同定し、代謝物はワイドターゲット分析で測定し約300成分を検出しました。これらの結果をマルチオミクス解析パッケージにより統合解析を行いました。本稿では、長寿となる遺伝子の探索結果とタンパク質や代謝物の関連性を調査した事例をご紹介します。

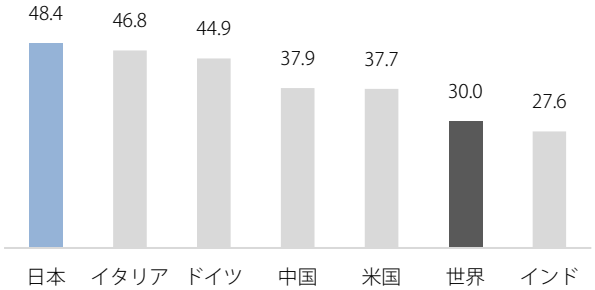


図1　世界各国の年齢中央値（2021年）[1)]

1 株式会社島津製作所
2 株式会社島津テクノリサーチ
3 インフォコム株式会社
4 国立大学法人東京大学

1. はじめに

現代の世界において、人口構成の変化は重要な課題です。世界の人口中央値は30.0歳である一方で、先進国では高齢化が進み労働力人口の減少や社会保障費の増加などの課題が生じています。特に日本では年齢中央値が48.4歳と世界有数の高さで、イタリアやドイツとともに長寿社会の在り方として注目を集めています（図1）。

日本の人口統計を詳しく解析すると、老衰が死因の3位に位置し、近年その割合が急上昇しています（図2）。老衰に至る高齢者の増加に伴い、介護やフレイル、認知症などの健康問題が増加しており、これらの課題に対処するためには、健康寿命を延ばすための研究が急務となっています。

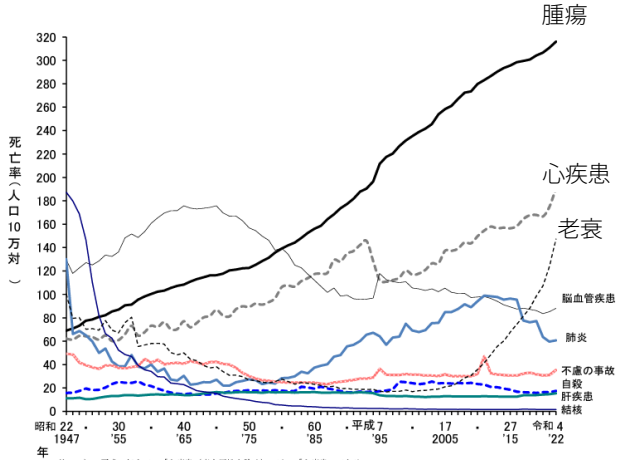


図2　主な死因別にみた死亡率（人口10万対）の年次推移（日本）[2)]

健康寿命を向上させる秘訣は一筋縄ではいかず、複雑な要素が絡み合っています。そのため人類の最高年齢である122歳（1997年ギネス記録）は20年以上更新されることなく、平均寿命のみが延びている状況です[3]。こうした課題に対し、基礎研究では人間を対象とするのではなく、ハエやマウス、シロイヌナズナなどのモデル生物を用いることが一般的です。キイロショウジョウバエは、短い世代周期とともに多様な遺伝子操作技術が確立されており、長寿についての研究に理想的なモデル生物のひとつです。

本アプリケーションでは、次世代シーケンサーと質量分析計の測定結果を代謝経路図に投影するマルチオミクス解析（DNA、RNA発現量解析、タンパク質解析、代謝物解析）により、長寿に関わる複数の生化学的プロセスや分子メカニズムを包括的に解析した例をご紹介します（図3）。

2. 測定方法

8世代バッククロス（戻し交配）したキイロショウジョウバエ（野生型n=2、長寿型n=2、各試料は5-10匹をまとめてすり潰した粉末）を多角的に評価するために、以下3つの装置を用いてDNA、RNA、タンパク質および代謝物を測定しました。

- **次世代シーケンサー**

DNA測定に関しては、QIAamp DNA Mini Kitを用いて抽出し、約20 µg/mLのDNAを含有する200 µLの抽出液（O.D.260/O.D.280 = 2.0弱）を測定に用いました。測定はLigation Sequencing Kit V14を用いて、R10.4.1フローセルとSUP精度（Guppy Ver.7）により5 khz（400 bps）ランモードでベースコールを実施しました。83万箇所のSNV/Indel変異箇所と、3万9千箇所の構造変異が検出されました。

RNA測定に関しては、RNeasy® Micro Kitを用いてTotal RNAを抽出し約20 µg/mLのRNAを含有する30 µLの抽出液（O.D.260/O.D.280 = 2.0弱）を測定に用いました。測定はPCR-cDNA Barcoding Kitを用いて、R9.4.1フローセルとHAC精度（Guppy Ver.7）で450 bpsランモードによりベースコールを実施しました。5920の遺伝子座発現量を検出し、そのなかでも統計的に有意（p値0.05以下）な遺伝子座である185個の遺伝子座を検出しました。

- **液体クロマトグラフ質量分析計 Nexera Mikros-LCMS9050**

各試料にリン酸緩衝食塩水（PBS）を入れてビーズ破砕した後に、等量の10 % SDS in 100 mM TEAB pH8.5を加えて可溶化しました。S-Trapのプロトコルに従ってTrypsin消化し、Iodoacetamide でアルキル化を実施しました[4]。Speed vacでバッファー液を乾固し、0.1 %ギ酸で再溶解してMikros-9050でノンターゲット分析により測定しました。Nexera Mikrosはミクロ流路対応液体クロマトグラフ質量分析計システムで、LCMS-9050は四重極飛行時間（Q-TOF）型質量分析計です。移動相は0.1%ギ酸水とアセトニトリルで、トラップカラムはYMC-Triart C18, 12 nm, 5 µm, 0.3 mm I.D. 50 mm、分析カラムはYMC-Triart C18, 12 nm, 1.9 µm, 0.3 mm I.D. 150 mmを用いて測定し、944個のタンパク質を同定しました。

- **ガスクロマトグラフ質量分析計 GCMS-TQ8040 NX**

GCMS-TQ8040 NXとSmart Metabolites Database™ Ver.2を用いて有機酸、糖、核酸、脂肪酸、アミノ酸などの一次代謝物488成分を23分で測定しました。面積値は内部標準物質である2-イソプロピルリンゴ酸で補正しました。各検体で300成分弱の代謝物が検出されました。

図4　装置写真
マルチオミクス解析パッケージ（上）、
Nexera Mikros™とTOFシステム（左下）、GCMS-TQ™8040 NX（右下）

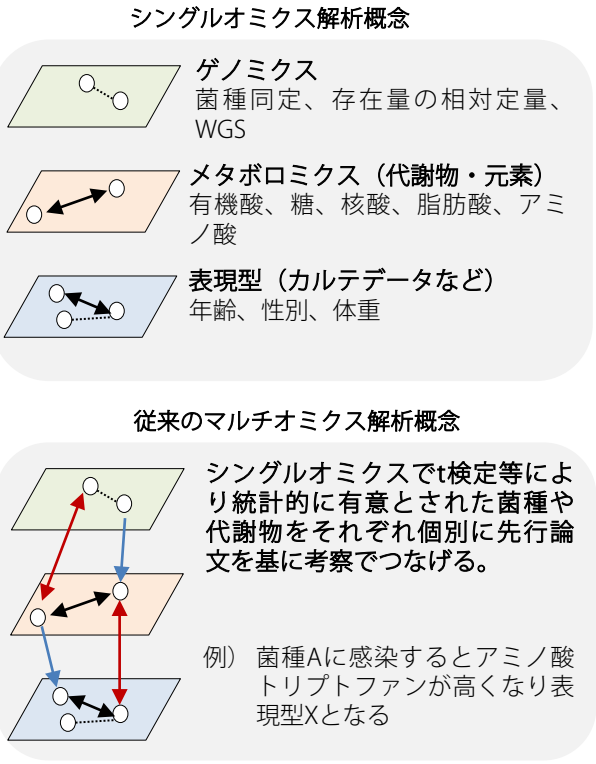


図3　シングルオミクス解析とマルチオミクス解析の概念

* ボルケーノプロットを代謝経路上に投影する手法や主成分分析を代謝経路上で解析する手法など弊社独自の解析手法を用います（特許出願中）。詳細はアプリケーションノートNo.94「次世代シーケンサーとGC/MSによる腸内細菌叢と代謝物と香気成分のマルチオミクス解析」をご参照ください。

3. ノンターゲットプロテオミクス解析結果

Mikros-9050（DDA*1）で得られたタンパク質のデータ解析（図5）にはPEAKS Studio XProソフトウェア（インフォコム株式会社）を用いました。まず*de novo*シーケンシングと呼ばれる機能により、MS/MS質量ピークリストから、ペプチドのアミノ酸配列（翻訳後修飾含）を算出しました。その結果、野生型のみで検出されたペプチド数は963個で、遺伝子変異型のみで検出されたペプチド数は1148個で、共通して検出されたペプチド数は3666個でした（図6）。

De novo sequencing結果（ペプチドのリスト）を利用して、UniprotのGene Count 13824データベースによりタンパク質名をFDR 1％（多重検定を考慮したp値：ペプチドもタンパク質も1％）で同定しました[5]。

*1 Data Dependent Acquisitionの略で、存在量の多いタンパク質から順に検出し同定する方法です。比較として、Data Independent Acquisition測定法は存在量に関係なく全てのMS/MSスペクトルを取得する代わりにデータ量が膨大となります）

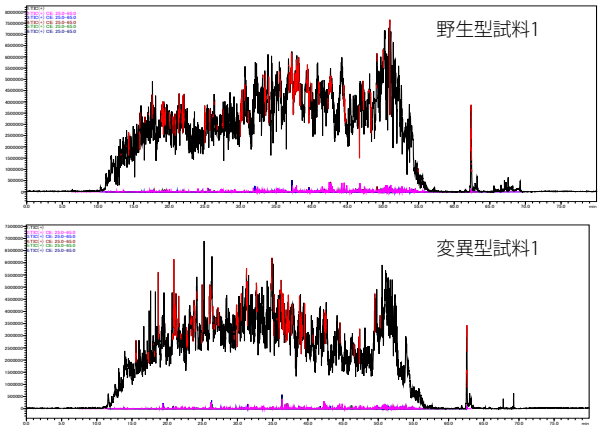


図5　網羅的プロテオミクス測定クロマトグラム（LabSolutions）

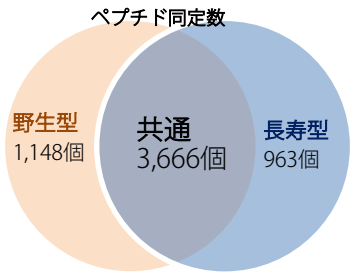


図6　PEAKS Studioによるペプチド同定数

「Fixed PTM:Carbamidomethylation」と「Variable PTM: Oxidation,Acetylation (N-term)」でデータベース検索を実施しました。該当タンパク質のペプチドとしてヒットしたペプチドのペプチドスコアを積算した値を分布図として図7に示します。野生型と長寿型でスコア分布に大きな差異が無く、分析と解析ともに良好であったことが伺えます。

そこで、それら5777個のペプチドを用いて944個のタンパク質を同定しました。944個のうち、野生型のみで検出されたタンパク質は162個、長寿型のみで検出されたタンパク質は126個、両群で検出されたタンパク質は656個でした（図8）。これらの同定されたタンパク質をANOVA解析（Significance > 10, Fold change > 1.3）に供し、統計的に有意な28個のタンパク質をボルケーノプロットとヒートマップで可視化しました（図9および10）。その結果、生殖細胞の形成に関与するタンパク質であるCG9075（ATP依存性RNAヘリカーゼeIF4A）など12個のタンパク質が野生型で特徴的に高く存在していることが分かりました。また、エネルギー生産（解糖経路）で重要なCG2171（トリオースリン酸イソメラーゼ）など16個のタンパク質は、逆に、長寿型ショウジョウバエで特異的に高く存在していました。

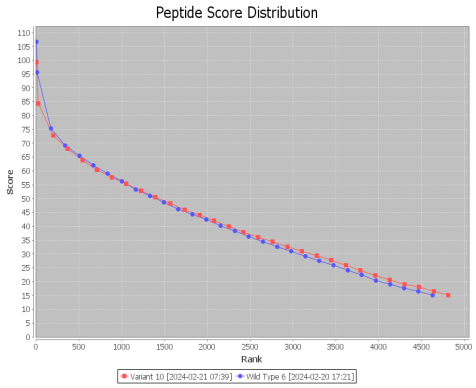


図7　ペプチド同定スコア分布

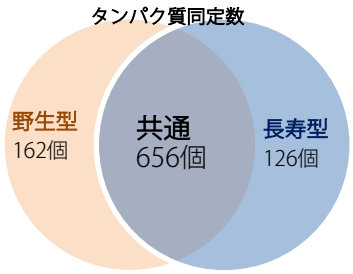


図8　PEAKS Studioによるタンパク質同定数

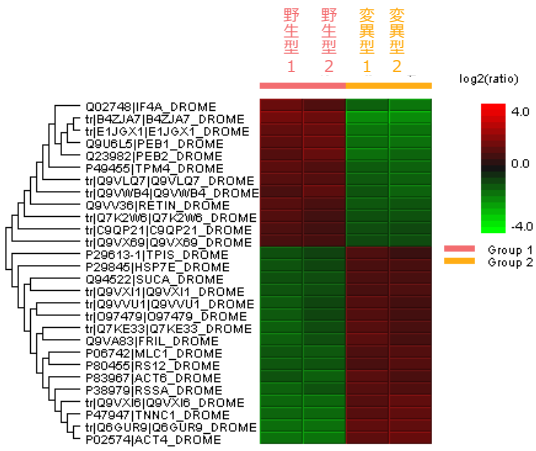


図9　PEAKS Studioによるヒートマップ解析
ANOVA Significance > 10, Fold change > 1.3で検出されたタンパク質

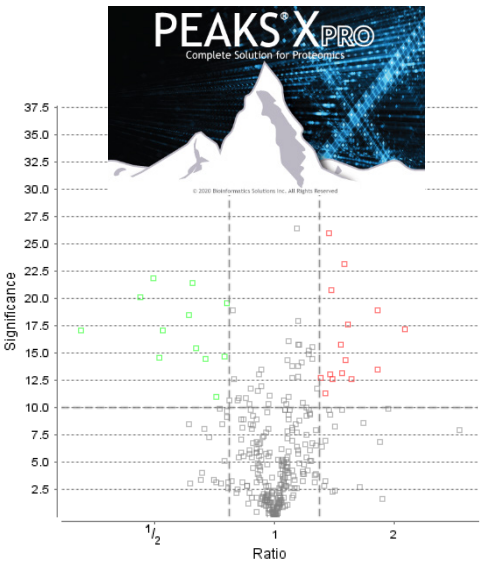


図10　PEAKS Studioによるボルケーノプロット解析

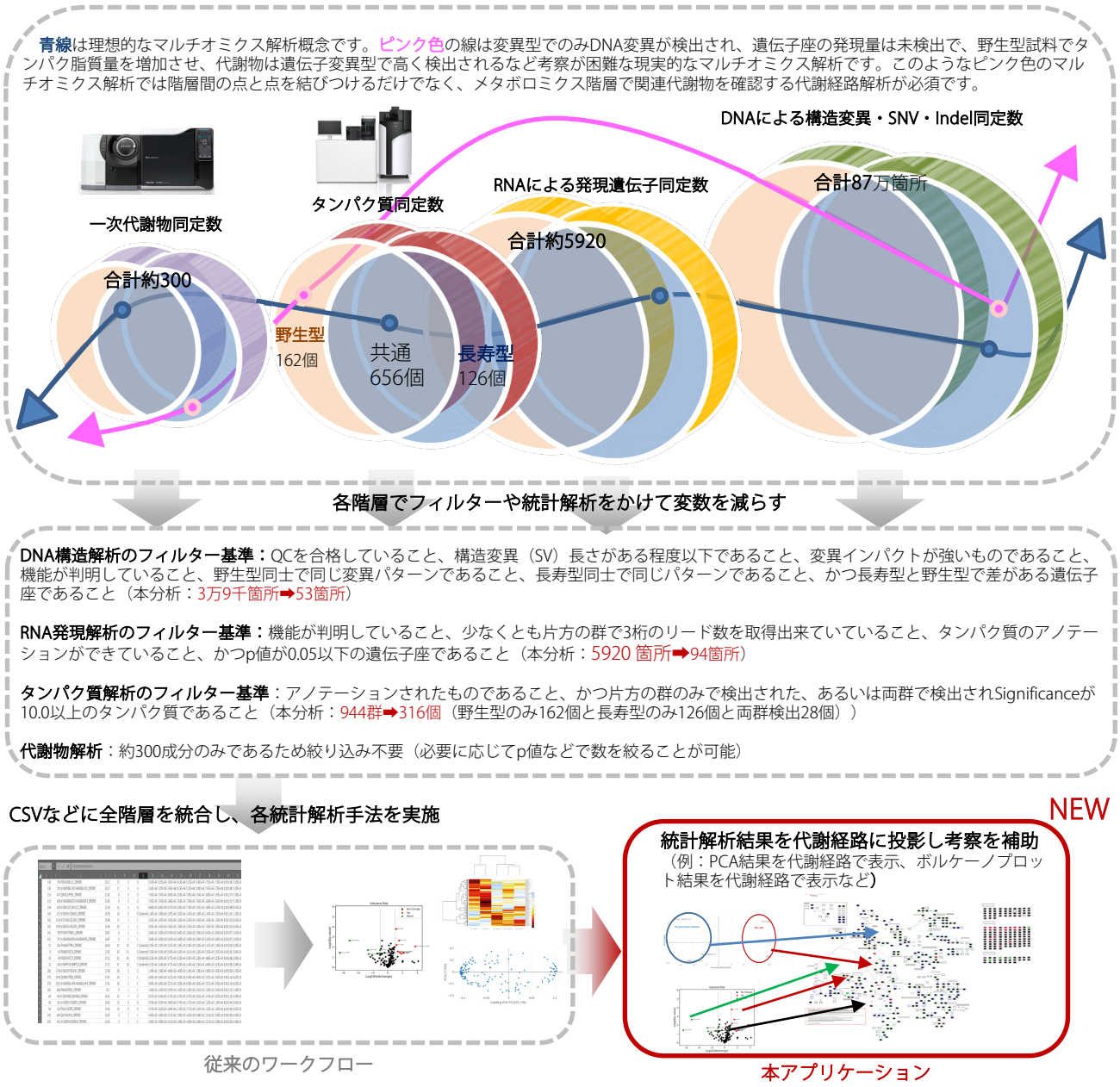


図11　マルチオミクス解析のワークフロー

前項の図9で長寿型で特異的に高く検出されたタンパク質のなかにCG14792（40Sリボソームタンパク SA）があります。このタンパク質は卵形成と成虫発生に必要となるもので、長寿型ショウジョウバエで特異的に高く検出されていることから、野生型と比較した際に、長寿型では成虫になる工程が遅延していることが推測されました。 CG14792 は特に40Sリボソームサブユニットの形成に必要で、長寿変異のメカニズムはこれらリボソームに関わりがある可能性が考えられます[6]。

このようにプロテオミクス階層でボルケーノプロットなどの統計解析を実施することでその階層の上流（DNA、RNA）や下流（代謝物、表現型）の群間差異を推測できます。しかし、実際にDNA、RNA、代謝物などを測定すると、プロテオミクスでの考察からは想定外の結果を得られることがあります。そのため、マルチオミクス解析パッケージを用いて、注目する代謝物だけでなく、その代謝経路周辺も同時に解析することや、ひとつの統計解析手法だけでなく複数の統計解析手法を組み合わせた手法（例：主成分分析結果を代謝経路図に投影）を用いることでマルチオミクス解析を進めることが重要です（図11）。

本報の実験でも、次世代シーケンサーでの全ゲノム解析測定結果は膨大で、モデル生物であるショウジョウバエでもDNA構造変異が39,000箇所（SNV/Indel83万箇所）やRNA発現変異が5,920箇所（遺伝子座）が確認されました。

これらの次世代シーケンサーの結果を単体で解析するよりも、より変数の少ないタンパク質や代謝物（有機酸、糖、核酸、脂肪酸、アミノ酸など）と組み合わせ解析することで次世代シーケンサーの結果解析時間が大幅に短縮されることが予測されます（図12）。

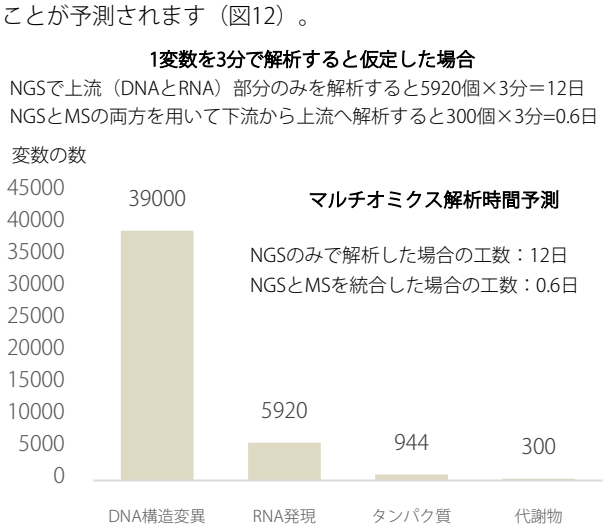


図12　各階層で検出された変数の数

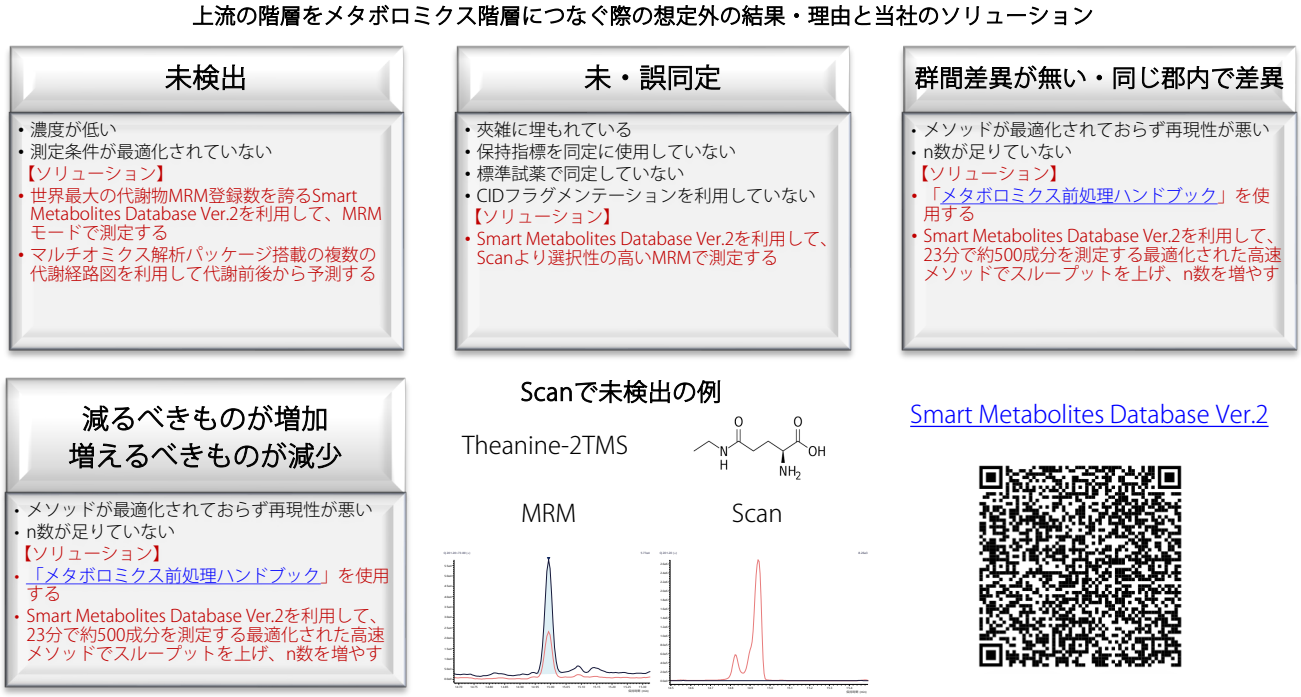


図13　代謝物を含むマルチオミクス解析に関する当社のソリューション一覧表

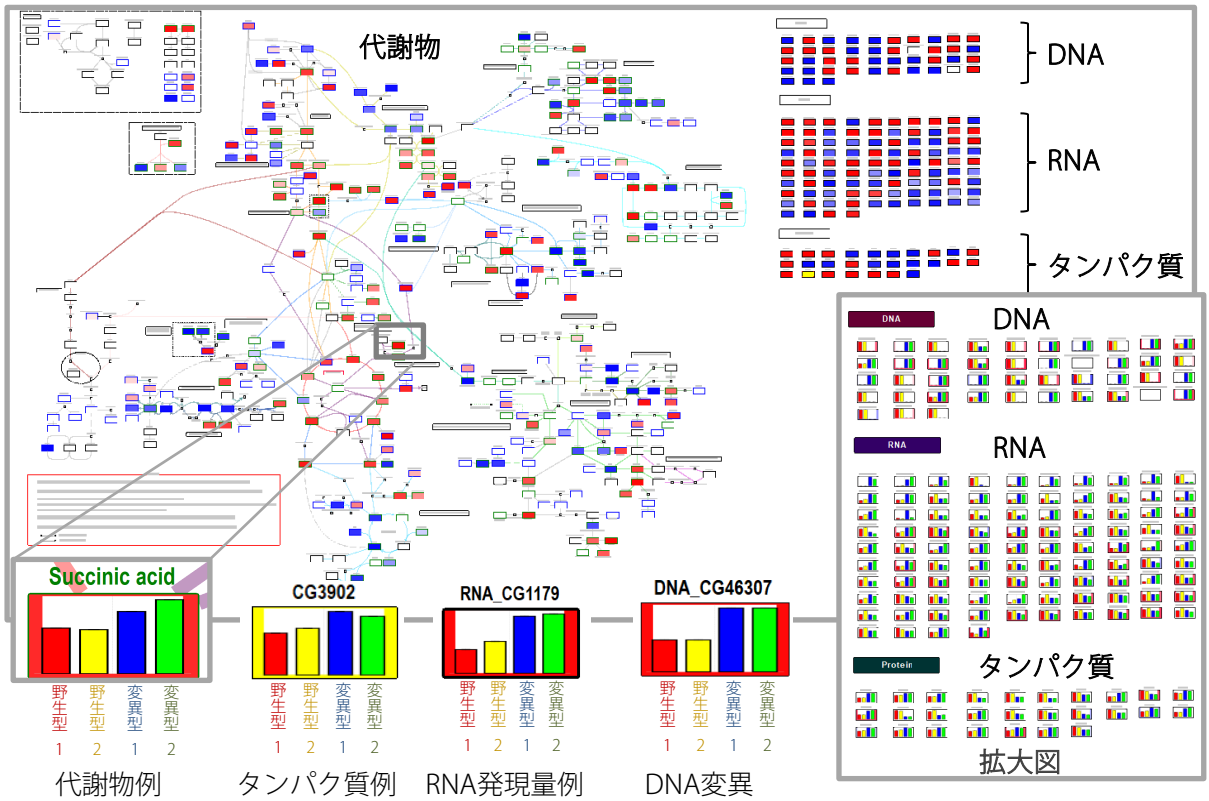


図14　マルチオミクス解析パッケージ（標準搭載の代謝物経路図にDNA、RNA、タンパク質のノードを追加しました）
*DNA、RNA、タンパク質は図9の通り各階層でフィルターにより絞ったもののみを投影しています。
RNAはリード数、DNA変異は0/0が0、1/0か0/1が1、1/1が2と入力しています。

しかし、上流（DNAとRNA）と下流（代謝物）を結びつける際に、注目する代謝物が未検出であったり、未同定や誤同定による結果の解釈のミスが起こる可能性があります。そのためSmart Metabolites Database Ver.2収載の最適化されたMRMメソッドにより、複雑な生体試料のなかでも注目する代謝物由来のピークを選択的に検出し、標準試薬により同定されたメソッドで確実にピークをアノテーション（同定）することがマルチオミクス解析において重要です（図13）。例えば、本報のショウジョウバエ試料に含有される代謝物テアニン（Theanine-2TMS）は、Scanでは未検出でMRMのみで検出されました。

前項の図11に示した方法でDNA、RNA、タンパク質の変数の数を各階層でのフィルターと統計解析で絞った結果、DNAは53箇所、RNA発現変異は94箇所、タンパク質は316個となりました。それらを代謝経路図に投影し、経路上でタンパク質CG3902を例にして相関解析を実施しました（図14）。ソフトウェアの機能により、タンパク質CG3902と正の相関関係があるDNA、RNA、代謝物は赤色、負の相関関係があるものは青色に自動で識別しました。

また654個の変数をボルケーノプロットにより統合解析し、野生型で特異的に検出される変数（左上：緑枠内7個）と長寿型で特異的に高く検出される変数（右上：赤丸15個）を検出しました（図15）。ボルケーノプロットで注目すべき変数を特定した後、代謝経路に投影しました。その結果、長寿型でアラニン代謝の最終生成物であるホモシステインが特異的に蓄積している一方で、長寿型でリシン分解代謝の中間物である2-アミノアジピン酸が減少し代謝経路が促進していることが分かりました（図16）。

ホモシステインは高濃度で蓄積すると脳や血管にダメージを与え、アルツハイマー型認知症を発症することが知られています。長寿に変異したショウジョウバエでホモシステインが高く検出されていることは、寿命が長くなるものの健康寿命は同様に伸びない可能性を示唆します[7]。タンパク質ではCG15616のような機能が判明していないものが同じく長寿型ハエで特異的に高く検出されており、同じ変動パターンで検出されている代謝物やRNAとの関連を調査することが次の研究の課題となることが分かります。

また長寿型で2-アミノアジピン酸はは老化したヒトの皮膚に存在し、皮膚のタンパク質がカルボニル酸化した際のマーカーとされる代謝物です[8]。しかし、ショウジョウバエでの形態については報告が無く、代謝経路上で同じく緑色で示されているRNAやタンパク質を注視する必要があります。

このように、どのDNAやRNAやタンパク質が特異的に片方の群で検出されているかボルケーノプロットと代謝経路図で色分けすることで、機能が判明していない遺伝子座や遺伝子変異やタンパク質と代謝経路の関係性を研究できます。

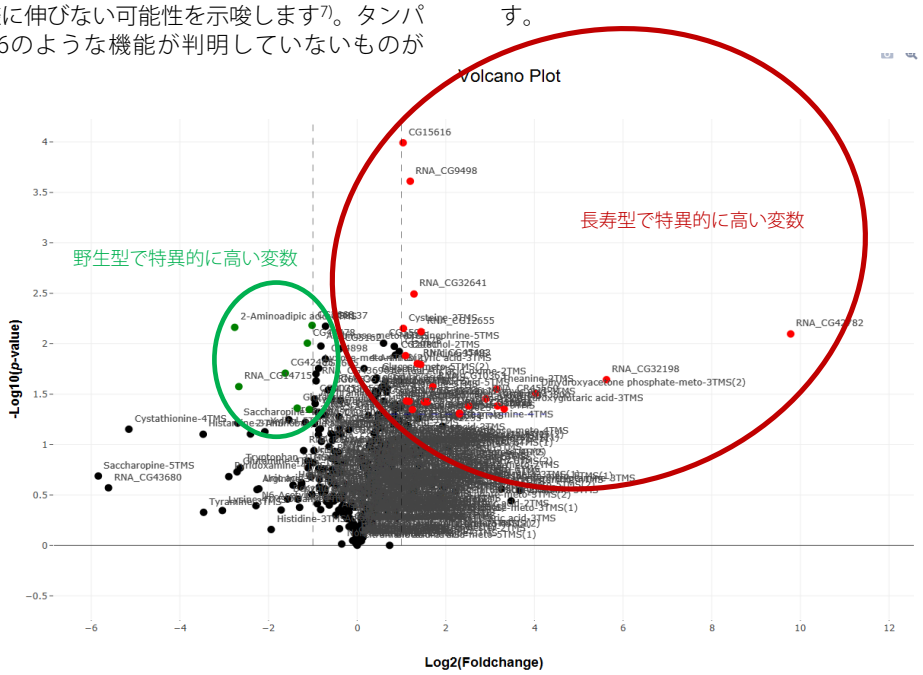
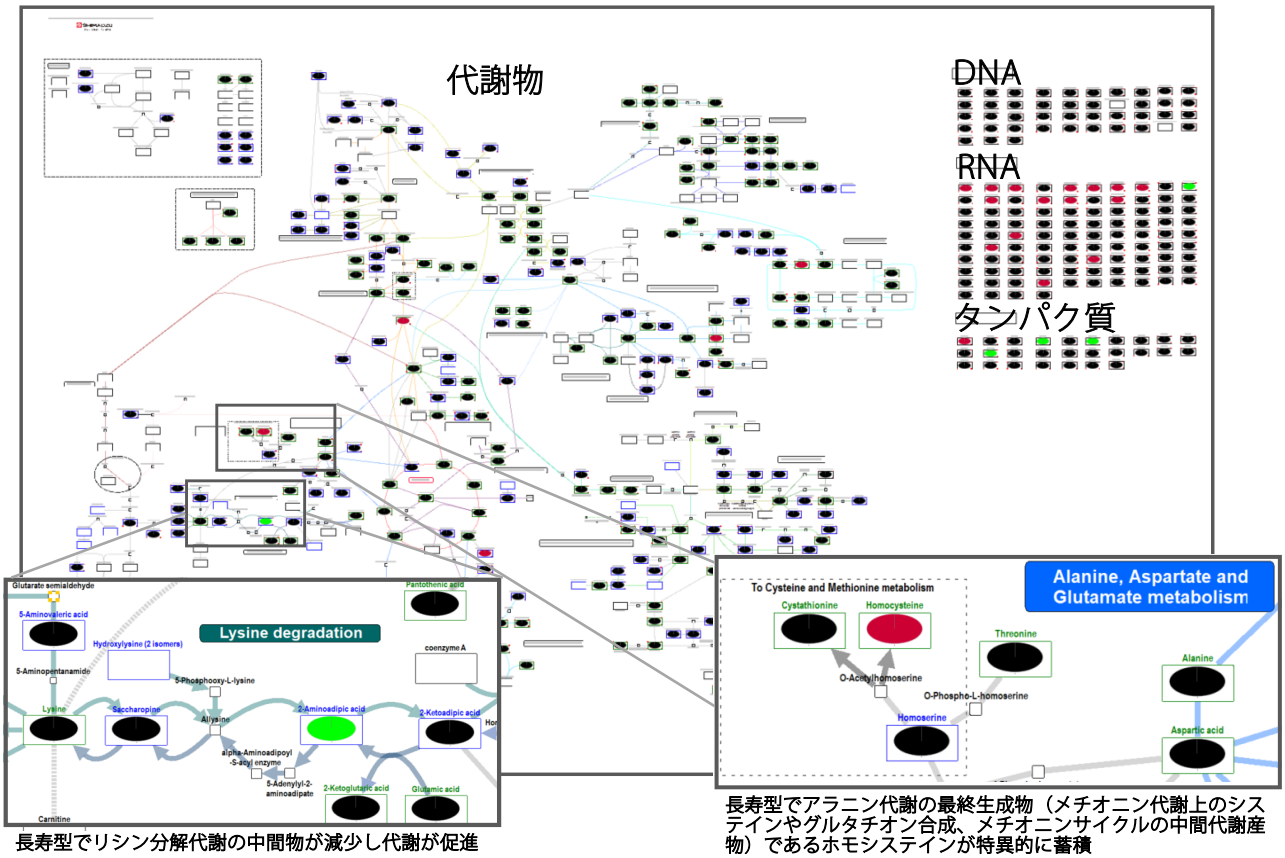


図15　野生型（n=2）と長寿型（n=2）のボルケーノプロット解析（DNA、RNA、タンパク質、代謝物データの654個の変数）



長寿型でリシン分解代謝の中間物が減少し代謝が促進

長寿型でアラニン代謝の最終生成物（メチオニン代謝上のシステインやグルタチオン合成、メチオニンサイクルの中間代謝産物）であるホモシステインが特異的に蓄積

図16　ボルケーノプロットの結果を代謝経路図に投影した解析

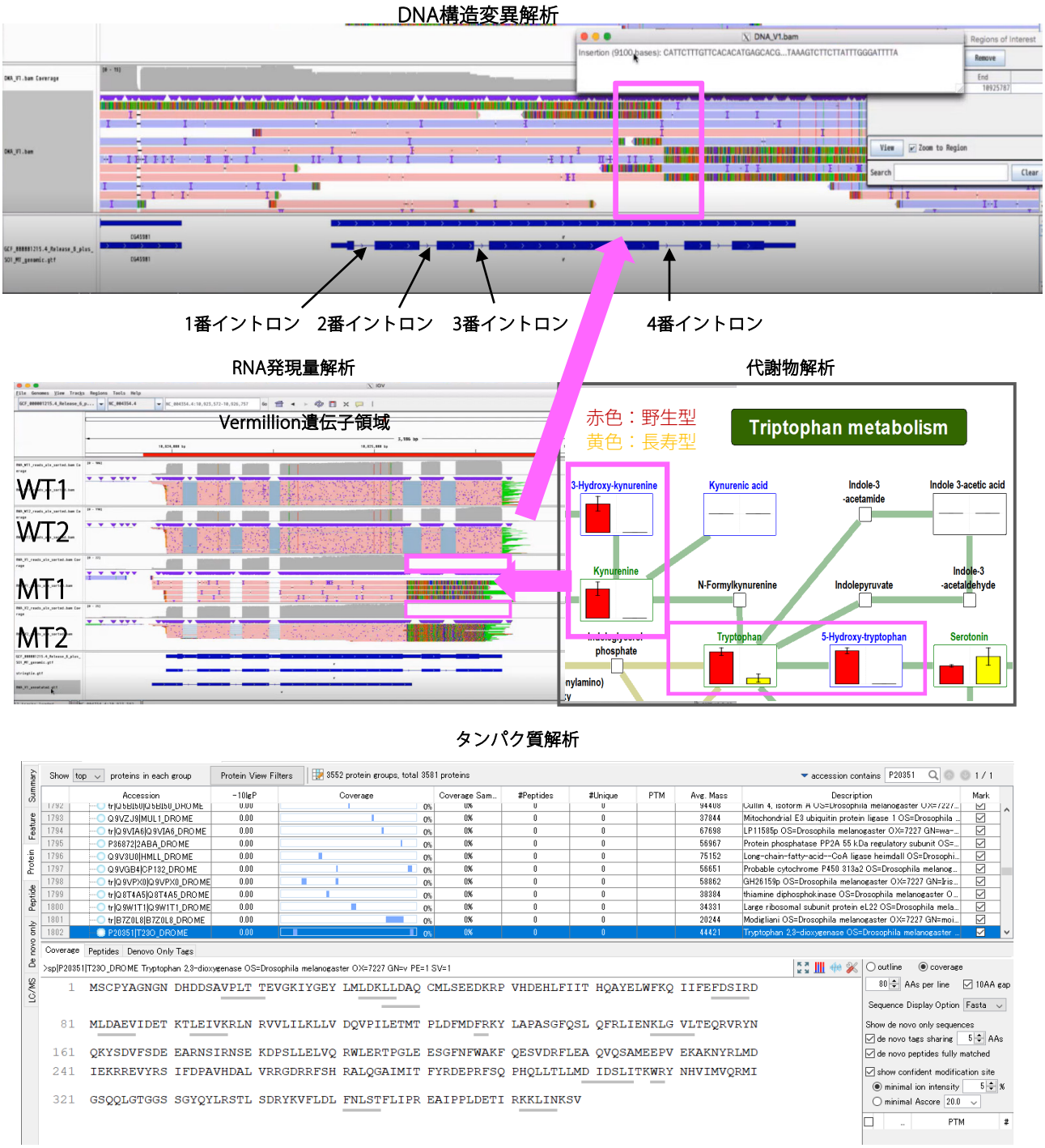


図17　長寿型ショウジョウバエにおけるキヌレニン代謝経路抑制とTDO2 RNA発現量変異確認と遺伝子座Vermillionの4番イントロンへの挿入確認
（DNAとRNAはIGV、タンパク質はPEAKS Studio、代謝物はマルチオミクス解析パッケージの画面キャプチャ）

また、代謝経路上でキヌレニン代謝経路が長寿型で抑制されていることから、律速酵素であるトリプトファン2,3-ジオキシゲナーゼ（TDO2）の濃度と発現量を確認しました（図17）。プロテオミクス解析では該当するタンパク質は低濃度で存在しているため信頼性の高い結果が得られなかったものの、RNA発現量解析では長寿型のVermillion遺伝子領域で構造の欠損が見られました。そのため、DNA解析でVermillion遺伝子座を確認すると、4番目のイントロン付近でクリッピングが確認され9100塩基ほどの長さで挿入があることが分かりました。

野生型と長寿型の遺伝子上差異である83万箇所のSNV/Indelや3万9千箇所の構造変異を単体で解析するよりも、代謝経路解析を活かすことでマルチオミクス解析の効率を上げることができました。

また、各代謝物の比を算出して、グリッド上に並べて、それらの比と注目するRNAやタンパク質との相関解析を実施することで基質や生成物の特定につなげられる可能性があります（次項図18）。

ストレス感受性に影響を及ぼすとされるタンパク質CG31508に正の相関関係があるものを赤色、負の相関関係があるものを青色で自動で色付けしました。その結果、2-デオキシ-D-リボースを生成物、6リン酸フルクトースを基質とする比と正の相関関係があることが分かりました。また、逆にAlloseを生成物、2-デオキシ-D-リボースを基質とする比と負の相関関係があることが分かりました。このことから、CG31508が2-デオキシ-D-リボースを蓄積する作用があることが示唆されました。

4. まとめ

次世代シーケンサーと液体クロマトグラフ質量分析計システムを用いて長寿となる遺伝子変異を起こしたキイロショウジョウバエとその野生型と比較しました。DNA解析では全ゲノム解析によりSNV/Indel83万箇所（構造変異は約4万箇所）、RNA解析では網羅的発現量解析によりp値 0.05以下の185遺伝子座を検出しました。タンパク質解析は網羅的解析により約1000個を同定し、代謝物では約500成分をワイドターゲット分析で測定し約250成分を検出しました。これらの結果をマルチオミクス解析パッケージにより統合解析を実施しました。次世代シーケンサーの結果（DNA、RNA）を単体で解析するよりも、より変数の少ないタンパク質や代謝物（有機酸、糖、核酸、脂肪酸、アミノ酸など）と組み合わせて解析することで次世代シーケンサーの結果解析時間が短縮され、マルチオミクス解析を促進させることができます。マルチオミクス解析パッケージの代謝経路解析を用いることで長寿型でのトリプトファン経路の抑制が見られ、Vermillion遺伝子座への挿入変異をDNAやRNAでも感知することができました。今後の展望としてエンリッチメント解析やDNAプロテオミクス階層でのDIA測定の実施などが考えられます。

<参考文献>
1) World Population Prospects 2022, United Nations, 2024年2月27日参照
2) 令和4年（2022年）人口動態統計月報年計（概数）の概況, 厚生労働省, 2024年2月27日参照
3) 120歳は長寿の上限？スーパーセンテナリアンの世界, 日経新聞, 2024年2月27日参照
4) Category：Protocols, PROTIFI, 2024年2月27日参照
5) Proteomes Drosophila melanogaster (Fruit fly), UniProt, 2024年2月27日参照
6) P38979 RSSA_DROME, UniProt, 2024年2月27日参照
7) アルツハイマー病や大病を引き起こす高ホモシステインとは？, DIAMOND online, 2024年2月27日参照
8) 2-aminoadipic acid is a marker of protein carbonyl oxidation in the aging human skin: effects of diabetes, renal failure and sepsis, David R Sell, 2024年2月27日参照

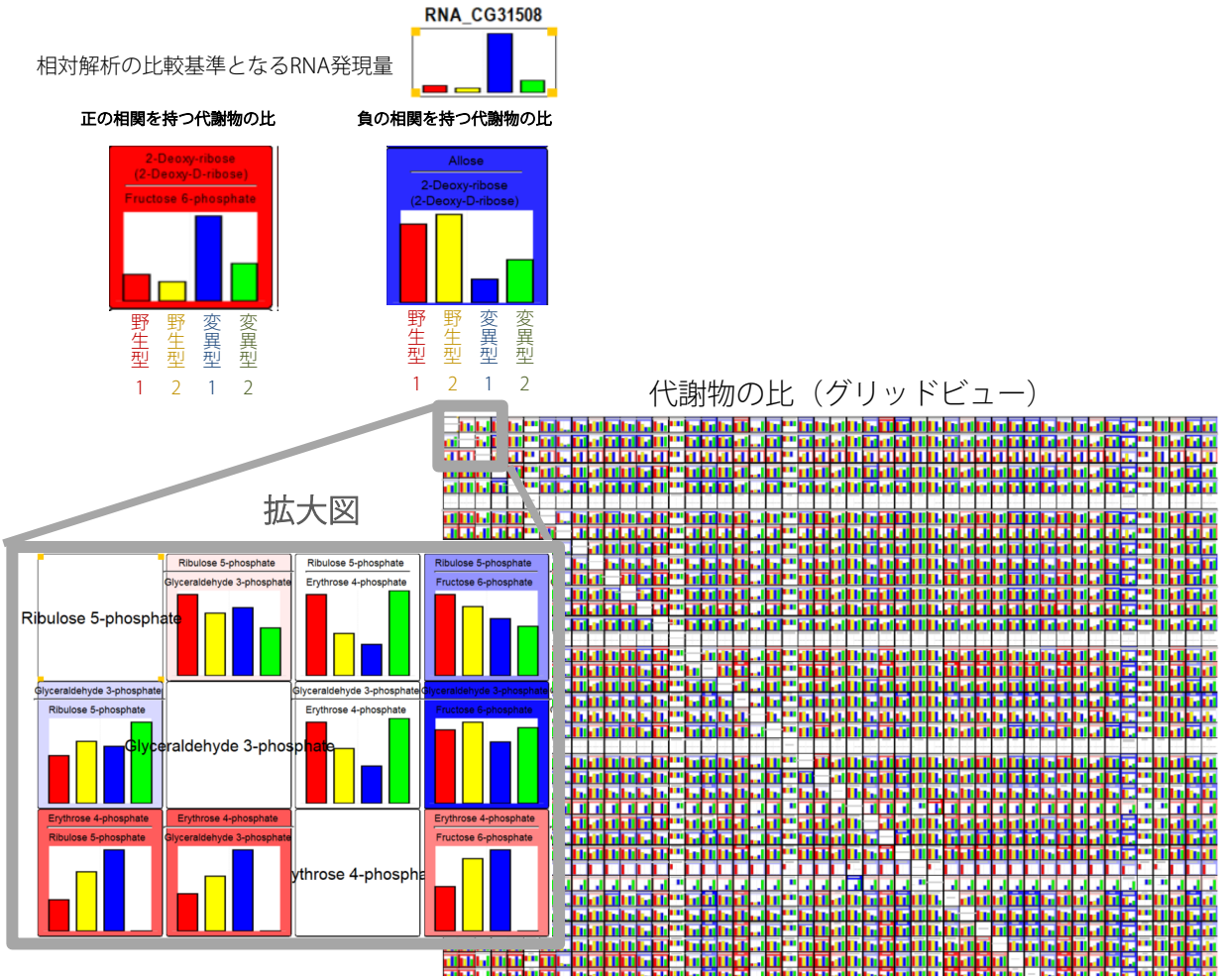


図18　NGSで検出されたRNA発現量（例：CG31508）と同じ群間差異パターンを有する基質/生成物ペアの探索
（赤がRNA発現量と正の相関関係、青が負の相関関係）

GCMS-TQ、LCMS、Mikros、Smart Metabolites Databaseは、株式会社島津製作所またはその関係会社の日本およびその他の国における商標です。

株式会社 島津製作所 分析計測事業部
https://www.an.shimadzu.co.jp/

初版発行：2024 年 4 月

本文書に記載されている製品は、医薬品医療機器法に基づく医療機器、体外診断用医薬品として承認・認証等を受けておりません。本文書に記載されている分析手法を治療診断目的およびその手続き上で使用することはできません。

本資料は発行時の情報に基づいて作成されており、予告なく改訂することがあります。
本文中に記載されている会社名、製品名、サービスマークおよびロゴは、各社の商標および登録商標です。
本文中では「TM」、「®」を明記していない場合があります。

本資料の掲載情報に関する著作権は当社または原著作者に帰属しており、権利者の事前の書面による許可なく、本資料を複製、転用、改ざん、販売等することはできません。
掲載情報については十分検討を行っていますが、当社はその正確性や完全性を保証するものではありません。また、本資料の使用により生じたいかなる損害に対しても当社は一切責任を負いません。

© Shimadzu Corporation, 2024