

平成 30 年 6 月 9 日現在

機関番号：17102

研究種目：基盤研究(C) (特設分野研究)

研究期間：2015～2017

課題番号：15KT0108

研究課題名(和文) ネットワークダイナミクスの数理的解析によるインスリン情報伝達機構の解明

研究課題名(英文) Elucidation of insulin signal transduction mechanism with mathematical analysis of network dynamics

研究代表者

松崎 芙美子 (Matsuzaki, Fumiko)

九州大学・生体防御医学研究所・助教

研究者番号：10631773

交付決定額(研究期間全体)：(直接経費) 3,700,000円

研究成果の概要(和文)：インスリン受容によりマウス肝臓で惹起されるタンパク質リン酸化量の経時変化を網羅的に計測し、リン酸化ダイナミクスの統計的・情報学的・数理的な解析を行った。その結果、10,000にも上るノード(タンパク質リン酸化部位)を含めたインスリン情報伝達ネットワークのダイナミクスと、様々な生物学的機能との関連性が明らかになった。さらに、インスリン-インスリン情報伝達分子-その他のタンパク質リン酸化の常微分方程式モデルが大規模に作成され、巨大なインスリン情報伝達ネットワークの全体像に加えて、その動的特性が定量的に明らかになった。

研究成果の概要(英文)：In this project, we focused on insulin signal transduction to see this phenomenon and its dynamic characteristics as completely as we could. Phosphoproteomics of time series liver samples taken during a four-hour period after insulin treatment was performed and the data were analyzed using statistical, informatics and mathematical approaches. As a result, the dynamics of an insulin signaling transduction network with around 10,000 nodes representing protein phosphorylation, and its relation to various biological function were revealed. Furthermore, ordinary differential equation models among insulin, insulin signaling molecules and other protein phosphorylation were built on a large scale and the dynamics of the entire insulin signaling transduction network were characterized quantitatively.

研究分野：定量生物学

キーワード：インスリン情報伝達機構 リン酸化プロテオミクス

1. 研究開始当初の背景

糖尿病は血糖の調節不全による全身性の疾患であり、生活の質を著しく低下させるにとどまらず、命を奪うことも珍しくない。今や日本人の1/5が糖尿病かその予備群であると言われており、その病態の解明や治療法の開発が強く望まれている。糖尿病の発症は、食事により血中に分泌されたインスリンに対して様々な組織の中の細胞が適切な応答をできなくなることから開始するため、この段階で予防策や治療を施すことが最も有効である。しかし、糖尿病において各細胞に惹起されるインスリン応答異常の詳細は未だ解明されていないのが現状である。

インスリンを受容した細胞の中では、タンパク質にリン酸基が共有結合する(タンパク質のリン酸化)という現象が連鎖的に起こることで、「タンパク質リン酸化のネットワーク」が形成される。その結果、インスリンを受容したという情報が細胞の中に散在している機能タンパク質に伝達され、最終的に栄養源の利用や貯蔵が行われる。インスリン受容によるタンパク質のリン酸化は既に数万カ所において検出されていることから、「インスリンの情報伝達」は、巨大なネットワークの形成により実現されていることが容易に予想される。さらに、我々は近年、インスリンの情報伝達において、リン酸化の時間変化のパターンも重要であることを明らかにしていることから、インスリンの情報伝達機構を真に解明するためには、リン酸化のネットワーク構造の時間変化(ネットワークダイナミクス)を考慮する必要がある。しかし現段階では、巨大なリン酸化ネットワークの構造はごく一部しか解明されていない上に、そのネットワークダイナミクスはほとんど不明といっても過言ではない。インスリンによって惹起されるタンパク質リン酸化を網羅的に検出し、そのネットワーク構造とネットワークダイナミクスを解析することで、インスリン情報伝達機構の全体像が初めて解明され、インスリン応答異常の改善方法を提示できる可能性がある。

2. 研究の目的

本研究では、インスリン受容により惹起されるタンパク質リン酸化量の時系列データを網羅的に取得し、そのデータを用いてタンパク質リン酸化ネットワークの構造やダイナミクスを統計的・情報学的・数理的に解析することで、インスリン情報伝達機構の全体像とその動的特性を明らかにすることを目的とする。

3. 研究の方法

(1) インスリン投与後のタンパク質リン酸化量時系列データの取得

インスリン(コントロール: PBS)を投与したマウスの肝臓を用いて、タンパク質のリン酸化の時間変化を定量的に計測した。具体的には、C57BL/6J野生型マウスにインスリンを腹腔内投与(0.7 U/kg)し、尾静脈血および肝臓を4時間にわたり経時的に採取した(それぞれ計13もしくは8時点)。採取した尾静脈血を用いて、血糖値および血中インスリン濃度を定法に従い計測した。既知のインスリンシグナル伝達分子のうちpIR(Y1150/Y1151)、pAKT(T308)、pGSK3B(S9)については、各肝臓サンプルより抽出したタンパク質を用いてウェスタンブロットティングによる定量を行った。抽出したタンパク質を酵素消化することでペプチドとした後に、リン酸化されたペプチドのみをIMAC法で濃縮して、リン酸化ペプチドサンプルとした。調製したサンプルについてラベルフリーでのLC-MS/MS解析(Orbitrap Velos Pro)を行った。個体サンプルを用いていることとラベルフリーでの定量であることから、生物学的なばらつきと技術的誤差が共に大きいことが予想されるため、計測時点につき生物学的・技術的にそれぞれ3回ずつ(計9回)反復計測を行った。MaxQuant・Perseusソフトウェアを用いて、得られたリン酸化ペプチドのマスペクトルをタンパク質配列情報(UniProt)と照合することでリン酸化ペプチドの同定を行い、このうちペプチド上のリン酸化部位の信頼性が高いもののみ(Localization probability > 0.75)を抽出した。計測毎にシグナル強度データの正規化(width adjustment)を行い、各リン酸化ペプチドの補正シグナル強度を得た。同条件のreplicateの補正シグナル強度において、平均±3SDを超えるデータを外れ値として除外して平均し、さらにリン酸化部位毎にリン酸化ペプチド補正シグナル強度平均の和をとることで、各部位のリン酸化の定量値とした。

(2) タンパク質リン酸化量時系列データのクリーニング・変動評価

タンパク質リン酸化量時系列データのうち、全時点(0-4時間の8点)中4時点以上の定量値が得られているリン酸化部位のみを抽出後、大きなノイズを含むもの(各時点の両側での定量値の傾きの符号が異なる場合の差の絶対値が、全体の上位5%であるデータ点を2点以上持つ時系列データ)を除外した。インスリン/PBS投与依存的な量的変動の有無を、時刻0の値に対して1.5倍以上もしくは0.67以下の値が2時点以上あるか否かで判定した。インスリンとPBS投与間での時間変動の差の有無を、Area Under Curve (AUC)の比がPBSのAUCのMean Absolute Deviation (MAD)以上か否かで判定した。各処理段階での時点毎および時系列を通したリン酸化部位数等を集計した。これらの処理はRソフトウェアを使用して行った。

(3) インスリンにより惹起されるタンパク質リン酸化の統計的・情報学的・数理解析
 クリーニング済みのタンパク質リン酸化量時系列データを入力として、主成分分析・階層的クラスタリングを適切に正規化した上で実施し、インスリン投与後の時系列変動の傾向を解析した。リン酸化が見られる各タンパク質に既知の生物学的情報 (KEGG・GO) を紐付けし、上位主成分における因子負荷量の大きなリン酸化タンパク質群や各クラスターに対して、特定の生物学的プロセス・分子機能・細胞内構成要素等が統計的に有意に濃縮されているかを確認した (Enrichment 解析)。

既知のリン酸化酵素の標的アミノ酸配列情報を基に、各リン酸化反応がどのリン酸化酵素により触媒されているかを **networKIN** で推定した。推定されたリン酸化酵素のうち、**insulin signaling (KEGG)** に含まれるものについて抽出し、これらによるリン酸化反応を既知のインスリン情報伝達経路の下流の推定部分ネットワークのエッジとし、リン酸化部位をノードとした。

定量された全タンパク質リン酸化について、血中インスリンを入力とした常微分方程式を作成した。既知のインスリンシグナル伝達分子のうち **AKT** もしくは **GSK3B** の標的候補のリン酸化部位については、**pAKT(T308)** もしくは **pGSK3B(S9)** を入力とする常微分方程式も同様に作成した。各リン酸化量は 2 パラメータを含む非線形関数で記述し、二乗誤差最小となるパラメータをグリッドサーチにより探索した。フィッティングの判定は χ^2 検定 (有意水準: 0.01) で行い、有意にフィッティングできているリン酸化について、得られた最適パラメータの関連性を解析した。

これらの解析は **R・Python・MATLAB** ソフトウェアを使用して行った。

4. 研究成果

(1) インスリン投与後のタンパク質リン酸化量時系列データの取得

インスリンのマウスへの腹腔内投与後、血中のインスリン濃度は、約 15 分後にピークを持ち 120 分後には元のレベルに戻る一過性の変動を示した。肝臓中インスリンシグナル分子のリン酸化量にも、同様に 15 分後程度にピークを持つ一過的な増加が見られ、投与したインスリンのシグナルが肝臓組織中にも伝達されていることが確認できた [図 1]。

インスリン投与後経時的に採取した肝臓組織より調製したリン酸化ペプチドの質量分析を行い、網羅的にタンパク質リン酸化を解析した結果、条件毎の平均で約 13,000、最大 16,500 のリン酸化部位について定量された。平均 $\pm 3SD$ 以上を外れ値として除外して [図 2]、条件毎の平均で約 10,000、最大 13,000 のリン酸化部位について定量値が得られた。

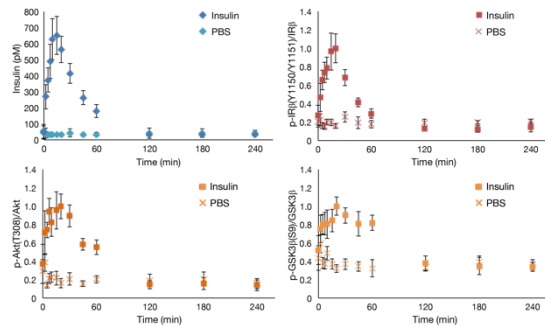


図 1. インスリン投与後の血中インスリン濃度と肝臓組織中インスリンシグナル分子のリン酸化レベル

(2) タンパク質リン酸化量時系列データのクリーニング・変動評価

得られたタンパク質リン酸化量時系列データの処理スキームを図 2 のように確立し、実行した。

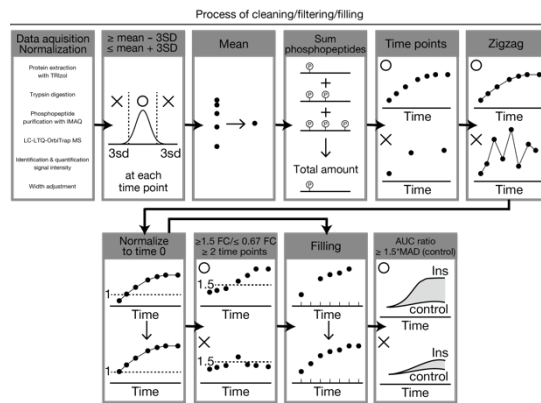


図 2. タンパク質リン酸化量時系列データの処理スキーム

各クリーニング段階でのリン酸化部位数は表 1 のようになっており、最終的に、条件毎の平均で約 10,000 のリン酸化部位について、以後の解析に使用可能な良好なデータであると判断した。

Condition	Phosphosites (Mean)	Phosphosites (Time points)	Phosphosites (Zigzag)
Insulin	12,154	10,080	9,901
PBS	12,275	10,295	10,067

表 1. 条件毎の各クリーニング段階におけるリン酸化部位数

インスリンと PBS 投与で変動差が見られたリン酸化部位数は約 3,400 であり、インスリン依存的な時間変動が見られたのは約 1,700 であった。

(3) インスリンにより惹起されるタンパク質リン酸化の統計的・情報学的・数理解析

クリーニング済みのリン酸化量時系列データについて主成分分析を行ったところ、明らかにタンパク質リン酸化量はインスリン投与後に変化していく傾向が見られた [図 3]。

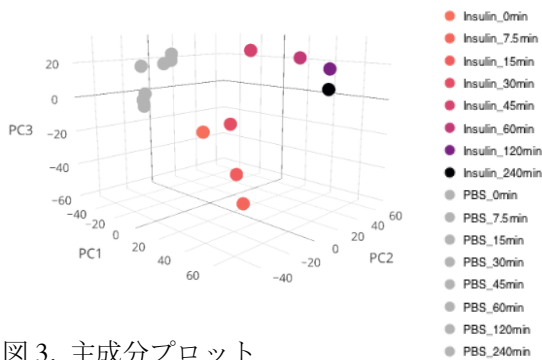


図 3. 主成分プロット

上位の主成分に大きく寄与している生物学的プロセス (GOBP) を探索したところ、インスリン投与後 45 分以内の短い時間スケールでサンプル間の分散の大きな第 3 主成分 (PC3) には、インスリン関連経路 (Insulin resistance, FoxO signaling pathway, AMPK signaling pathway, Insulin signaling pathway, Type II diabetes mellitus 等) に関連するタンパク質が有意に寄与していた [図 4]。このことから、インスリン受容の情報が肝臓内タンパク質リン酸化量の時間変動として順当に反映されていることが期待される。一方、インスリン投与後 30-240 分の時間スケールで大きな分散を示す第 2 主成分 (PC2) には、Citrate acid cycle, Nitrogen metabolism, Carbon metabolism, Glyoxylate and dicarboxylate metabolism 等の代謝経路が有意に寄与していた [図 4]。インスリン作用の最終機能分子である代謝関連タンパク質には、既知のインスリン関連経路と比較して遅い時間スケールで情報が伝達されていることが示唆される。

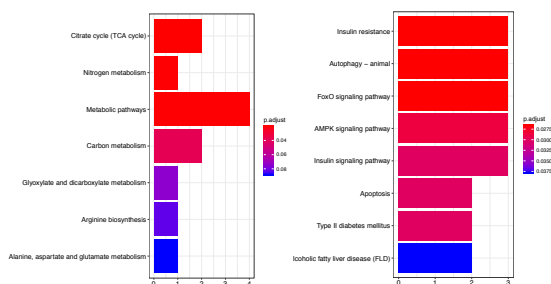


図 4. 第 2, 第 3 主成分に有意に寄与する生物学的プロセスの棒グラフと p 値

次に、時間変動が見られたリン酸化の階層的クラスタリングを行ったところ、インスリン投与依存的な変動が同様に確認でき、さらに時間変動パターンが異なる多数のクラスターに分割された [図 5-6]。

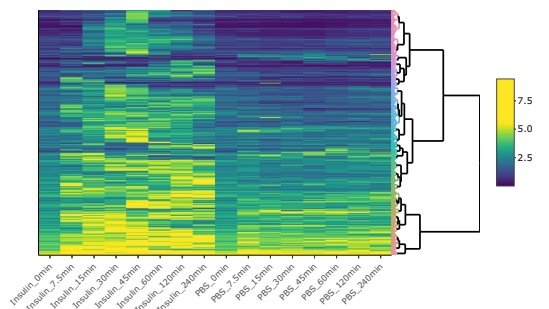


図 5. 階層的クラスタリング結果のヒートマップとデンドログラム

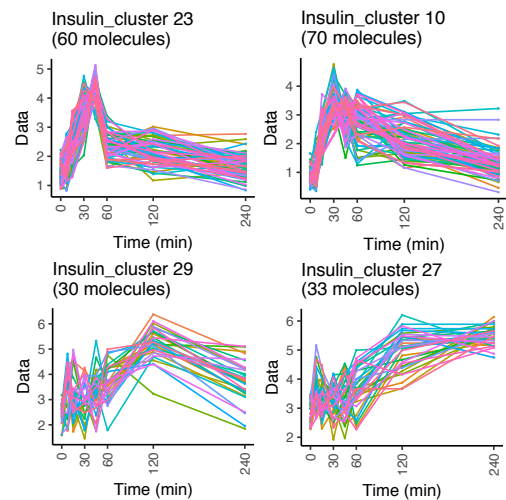


図 6. クラスター毎のタンパク質リン酸化量の時間変動の例

これらのクラスターの中には、各種シグナル伝達経路や代謝関連経路の他、特定の分子機能や細胞内構成要素のタンパク質が有意に含まれるものもあり、生物グループによっては、異なるダイナミクスでインスリンからの情報が伝達されていることが考えられる。一方、特定の生物グループを有意に含まないクラスターや、どのクラスターにも有意に含まれない生物グループも多く存在することから、既存の分類に依らない機序で経時的なインスリン情報伝達ネットワークの変遷 (ネットワークダイナミクス) が起きていることが考えられる。

次に、インスリン依存的にリン酸化量の時間変動を引き起こすタンパク質の探索を既知の標的アミノ酸配列情報を元に行ったところ、約 1,000 種類のタンパク質の 1,600 リン酸化部位について、関与が考えられるリン酸化酵素等が推定され、インスリンにより惹起されるリン酸化ネットワーク中のエッジが 4,700 得られた。このうち約 130 タンパク質の 160 リン酸化部位については、insulin signaling (KEGG) 中の 12 のリン酸化酵素によりリン酸化されると推定された [図 7]。

推定されたネットワーク構造に加えて、定量されている約 10,000 のリン酸化部位へのインスリンからのエッジも仮定し、血中インスリン濃度や insulin signaling 中の分子のリン酸化量変動をそれぞれ入力とする常微分方

程式モデルを作成した [図 8]。

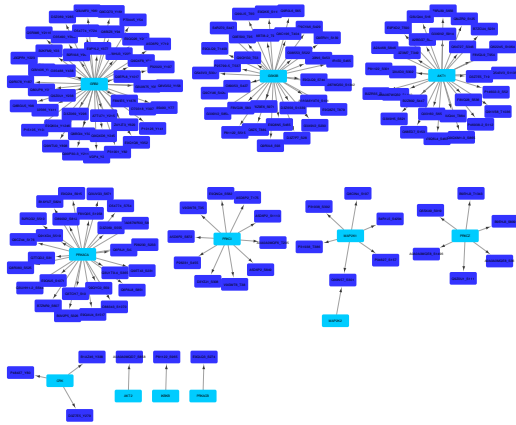


図 7. Insulin signaling (KEGG) 中の分子 (水色) を親ノードとする部分ネットワーク

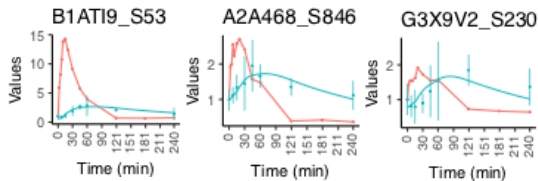


図 8. インスリン・pAKTT(T308)・pGSK3B(S9)量 (赤色) をそれぞれ入力とする常微分方程式モデルによるタンパク質リン酸化量 (水色) 変化のシミュレーション結果の例

血中インスリン濃度を入力とした場合で、フィッティングも良好であった約 2,400 のリン酸化について 2 つのパラメータの最適値の関係性を解析したところ、強い負の相関が見られた [図 9]。この 2 つのパラメータは、入力の線形もしくは非線形な効果をそれぞれ示すものであるため、親ノードのリン酸化の両効果の影響は、子ノードのリン酸化に対してある程度排他的に現れていると言える。

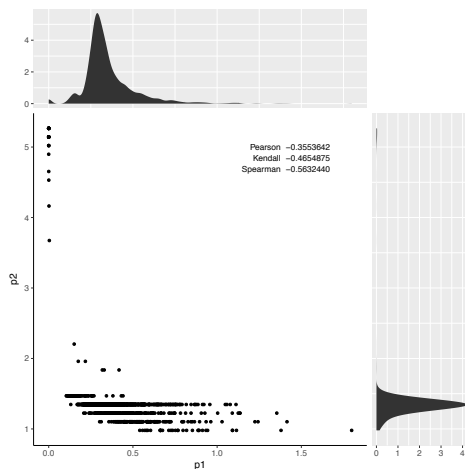


図 9. 推定された 2 パラメータの散布図とヒストグラム

強い非線形性制御 (3 次以上) を受けるタンパク質は、gamma-tubulin complex、lateral plasma membrane (GOCC) 等、細胞内に多岐にわたる影響を与える複合体や、シグナル伝達上重要な細胞内構成要素に有意に含まれており、インスリン受容に迅速に急激に反応する分子機構の存在が考えられる [図 10]。一方、線形制御が強いもの (線形係数 1.2 以上) は mitochondrial matrix、large ribosomal subunit (GOCC) といったオルガネラ等に含まれるものが多く、これらは最終機能分子近傍でインスリン受容に強めに応答していることが考えられる [図 10]。

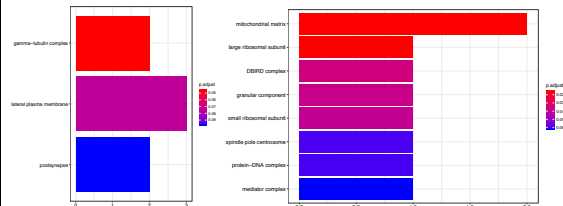


図 10. 非線形もしくは線形制御を強く受けているリン酸化タンパク質が有意に含まれる細胞内構成要素の棒グラフと p 値

本研究により、10,000 にも上るノード (タンパク質リン酸化部位) を含めたインスリン情報伝達ネットワークのダイナミクスと、様々な生物学的情報との関連性が明らかになった。さらに、インスリン-インスリン情報伝達分子-その他のタンパク質リン酸化の常微分方程式モデルが大規模に作成され、巨大なインスリン情報伝達ネットワークの全体像に加えて、その動的特性が定量的に明らかになった。本研究により明らかになったインスリン情報伝達ネットワークの構造や動的特性の観点から、糖尿病におけるインスリンへの応答異常を解析することによって、インスリン情報伝達機構のさらなる理解に加えて、将来的には糖尿病病態のネットワークレベルでの解明や病態改善方法を提示できる可能性がある。

5. 主な発表論文等

(研究代表者、研究分担者及び連携研究者には下線)

[学会発表] (計 11 件)

- (1) Tetsuo Imai, Takayasu Kawaguchi: Analysis of a coauthor network of nursing researchers. *2018 International Symposium on Nonlinear Theory and Its Applications*, Tarragona, Spain (2018)
- (2) Fumiko Matsuzaki: Metabolic systems analysis with multi-omics data. *9th International Conference and Exhibition on Metabolomics and Systems Biology*,

Prague, The Czech Republic (2017)

- (3) 松崎芙美子: トランスオミクスデータを用いた代謝システム解析の試み. *JSBi九州地域部会セミナー*, 下関 (2017)
- (4) 松崎芙美子, 宇田新介, 久保田浩行: 定量的トランスオミクス解析の試み. *定量生物学の会 第八回年会*, 岡崎 (2017)
- (5) 今井哲郎, 川口孝泰: 日本の看護学研究における研究者ネットワークの分析. *第14回情報処理学会ネットワーク生態学シンポジウム*, 能美 (2017)
- (6) 松崎芙美子: マウス肝臓におけるインスリン作用のトランスオミクス解析. *第1回代謝モデリング勉強会*, 横浜 (2016)
- (7) 松崎芙美子, 松本雅記, 押川清孝, 中山敬一: ヒト全代謝酵素の絶対定量と代謝システム解析. *第6回数値デザイン道場*, 三島 (2016)
- (8) 永田裕史, 今井哲郎, 吉澤康介, 三宅修平: ウェブページのランク付け指標の妥当性の検証. *第49回日本大学生産工学部学術講演会*, 習志野 (2016)
- (9) 永田裕史, 今井哲郎, 吉澤康介, 三宅修平: Webページの評価指標の妥当性の評価. *第13回ネットワーク生態学シンポジウム*, 木更津 (2016)
- (10) Fumiko Matsuzaki, Masaki Matsumoto, Kiyotaka Oshikawa, Fumihide Shiraishi, Keiichi I. Nakayama: Absolute quantification of all human metabolic enzymes and metabolic systems analysis. *International Symposium on Synthetic Systems Biology*, Fukuoka, Japan (2015)
- (11) 松崎芙美子, 松本雅記, 押川清孝, 中山敬一: ヒト全代謝酵素の絶対定量と代謝システム解析. *日本プロテオーム学会第13回大会*, 熊本 (2015)

[その他]

ホームページ等

<http://www.bioreg.kyushu-u.ac.jp/labo/omics/>

6. 研究組織

(1) 研究代表者

松崎 芙美子 (MATSUZAKI, Fumiko)

九州大学・生体防御医学研究所・統合オミクス分野・助教

研究者番号 : 10631773

(2) 研究分担者

今井 哲郎 (IMAI, Tetsuo)

東京情報大学・看護学部・遠隔看護実践研究センター・博士研究員

研究者番号 : 10436173

宇田 新介 (UDA, Shinsuke)

九州大学・生体防御医学研究所・統合オミクス分野・准教授

研究者番号 : 20599609

久保田 浩行 (KUBOTA, Hiroyuki)

九州大学・生体防御医学研究所・統合オミクス分野・教授

研究者番号 : 40376603

(3) 連携研究者

松本 雅記 (MATSUMOTO, Masaki)

九州大学・生体防御医学研究所・プロテオミクス分野・准教授

研究者番号 : 60380531