

プロテオームプロファイリングによる

ヒトがん細胞株の腫瘍モデルとしての評価に関する研究

2020

高橋 知里

目次

略語表.....	1
緒言.....	4
第1章 ヒト乳がん細胞株のプロテオームおよびリン酸化プロテオームプロファイリングと腫瘍モデルとしての評価	7
第1節 はじめに	7
第2節 乳がん細胞株のプロテオームおよびリン酸化プロテオームプロファイリング ..	9
第3節 乳がんの臨床的サブタイプとの相関	18
第4節 乳がんの内因性サブタイプとの相関	32
第5節 乳がん細胞株の腫瘍モデルとしての評価.....	42
第6節 おわりに	50
第2章 細胞内キナーゼ活性計測のためのリン酸化ペプチドの選択反応モニタリングの開発	51
第1節 はじめに	51
第2節 標的リン酸化ペプチドの選択	54
第3節 SRM 測定系の構築.....	57
第4節 測定方法のバリデーション	60
第5節 過バナジン酸処理 HeLa 細胞のキナーゼ活性測定	67
第6節 おわりに	70
結論.....	72
実験の部	73
発表論文	79
謝辞.....	80
引用文献.....	81

略語表

CCLE	がん細胞株エンサイクロペディア (cancer cell line encyclopedia)
CNV	コピー数多型 (copy number variation)
CPTAC	臨床プロテオミクス腫瘍分析コンソーシアム (clinical proteomic tumor analysis consortium)
Depmap	がん依存性マップ (the cancer dependency map)
DMEM	ダルベッコ改変イーグル培地 (Dulbecco's modified eagle's medium)
DTT	ジチオトレイトール (dithiothreitol)
ER	エストロゲン受容体 (estrogen receptor)
ESI	エレクトロスプレーイオン化法 (electrospray ionization)
FBS	ウシ胎児血清 (fetal bovine serum)
FDA	アメリカ食品医薬品局 (food and drug administration)
GO	遺伝子オントロジー (gene ontology)
HAMMOC	ヒロドキシ酸修飾酸化金属クロマトグラフィー (hydroxy acid-modified metal oxide chromatography)
Her2	ヒト上皮増殖因子受容体 2 (human epidermal growth factor receptor2)
IAA	ヨードアセトアミド (iodoacetamide)

ICGC	国際がんゲノムコンソーシアム (international cancer genome consortium)
IMBR	Isobaric match between runs (isobaric match between runs)
LC	液体クロマトグラフィー (liquid chromatography)
MS	質量分析法/質量分析計 (mass spectrometry/mass spectrometer)
NCCN	全米総合がん情報ネットワーク (national comprehensive cancer network)
NCI	米国国立がん研究所 (national cancer institute)
NGS	次世代シーケンシング (next-generation sequencing)
ODS	オクタデシルシリル (octadecylsilyl)
PAM50	PAM50 (prediction analysis of microarray 50)
PBS	リン酸塩緩衝生理食塩水 (phosphate buffer saline)
PR	プロゲステロン受容体 (progesterone receptor)
PVDF	ポリフッ化ビニリデン (poly(vinylidene difluoride))
RNA	リボ核酸 (ribonucleic acid)
RPPA	逆相タンパク質アレイ (reverse phase protein array)
RSD	相対標準偏差 (relative standard deviation)

SDB	ポリスチレンジビニルベンゼン (poly(styrene-co-divinylbenzene))
SRM	選択反応モニタリング (selected reaction monitoring)
StageTip	ステージ・チップ (stop-and-go-extraction tip)
TCGA	がんゲノムアトラス (the cancer genome atlas)
TEAB	重炭酸トリエチルアンモニウムバッファー (triethylammonium bicarbonate buffer)
TMT	タンデムマスタグ (tandem mass tag)
Tris	トリス (ヒドロキシメチル) アミノメタン (tris(hydroxymethyl)aminomethane)

緒言

近年、がん治療においてプレシジョン・メディシン (Precision Medicine) が着目されている。2015 年 1 月、米国のオバマ大統領 (当時) が一般教書演説で初めて Precision Medicine Initiative という言葉を使用した。高精度の遺伝子解析が可能になったことや治療薬の開発が進んだことを背景に、従来の個別化医療 (Personalized Medicine) と比べて、プレシジョン・メディシンには”患者の特徴ごと”に分類して適切に治療を選択するシステムを目指す意味合いが込められている。全米では、100 万人規模のコホートを創設し、ゲノムやその他のオミックスデータを用いて疾患の詳細なサブグループ分類を行い、各グループの治療や疾患の予防法の確立を目指す取り組みとされている。

がんのゲノム解析は、がん遺伝子やがん抑制遺伝子の同定から、第一世代の標的治療法の開発に至るまで、がんの理解と治療において大きな進歩を遂げてきた。大量並列シーケンシング (次世代シーケンシング: next-generation sequencing, NGS) の開発により、がんにおける体細胞変異の包括的なプロファイル、すなわち、がんゲノムを安価かつ迅速に解読することが可能になり、これまで分子生物学が中心であったがんの研究分野において、急速な変化が起こっている。さらに、がんゲノムアトラス (The Cancer Genome Atlas, TCGA)^{1,2} や国際がんゲノムコンソーシアム (International Cancer Genome Consortium, ICGC)³ における多数の患者から得た試料の塩基配列を解読するゲノム研究により、その変異ががんの発生や進展の直接的な原因になるドライバー遺伝子が数多く同定され、種々のがんにおいて変異のカatalogが作成された。また、これらの研究により、同じ種類のがん患者の間でもがんゲノムにいちじるしい不均一性 (腫瘍間不均一性) のあることが明らかにされ、個別化医療ならびにプレシジョン・メディシンの必要性が高まっている^{4,5}。

しかし、これらのゲノムの異常は必ずしもタンパク質の異常に反映されてはいない。なぜなら、mRNA の量とそれから翻訳されるタンパク質の量には密接な相関が無く、場合によっては 20 倍以上も異なることが報告されているからである⁶。さらに、タンパク質は、リン酸化、糖化、アセチル化などの翻訳後修飾や分解により、機能の全く異なる状態に変化する。ゲノムからの転写・翻訳の産物であるプロテオームは、生命活動をつかさどり、機能を有する分子として発現するタンパク質の総体であり、ほとんどの抗がん剤や治療薬の標的はタンパク質である。がん研究において、ゲノムに蓄えられた情報の機能的な解釈のためには、プロテオーム情報が必須である⁷。このように、がんにおける

病態生理の解明にプロテオーム解析は欠くことのできない解析手段であるとの認識が広く受け入れられ、急速にプロテオーム解析に対する関心が高まっている。その背景には、近年の質量分析計の高性能化および種々の定量技術の開発によるタンパク質解析技術の飛躍的進歩がある。それに伴い、大規模ながん細胞株のプロテオーム解析が様々ながん種で行われている^{8,9,10,11,12}。

昨今、ゲノムとプロテオームの研究を統合した“プロテオゲノミクス”という新たな研究分野が作られた^{13,14}。プロテオゲノミクスという言葉は、2004年にJaffeら¹⁵によって初めて科学文献に登場した。プロテオゲノミクスは、次世代DNAおよびRNAシーケンシングと質量分析ベースのプロテオミクスを組み合わせた腫瘍プロファイリングのアプローチであり、タンパク質やリン酸化などの翻訳後修飾の定量を可能にする。これらのデータを用いて分子メカニズムの理解を深めることで、疾患や障害の予防、早期発見、診断、治療の改善を図るとともに、分子知見の臨床応用を加速させ、分子特性に基づいた治療法を選択するプレシジョン・メディシンの実現に貢献できると考えられている。この新しい分野の重要性を理解した米国国立がん研究所 (National Cancer Institute, NCI) は、がんに関連するプロテオーム解析の実施を標準化するために、臨床プロテオミクス腫瘍分析コンソーシアム (Clinical Proteomic Tumor Analysis Consortium, CPTAC) プログラムを設立した。CPTAC は、がんの分子基盤を理解し、それらの知見の臨床への応用を加速するために、がんゲノム (および、それに関連する生物学的プロセス) の変化に関連するタンパク質を体系的に同定し、複数のがん種にわたる深層プロテオゲノミクスプロファイリングの実施を目指している。

がん細胞株は、腫瘍モデルとして、がんの原理の研究と新薬開発に関して必須のツールとなっている。同時に、プレシジョン・メディシンの実現には、がんについて膨大な情報を理解する必要があり、そのためにがん細胞株は汎用されている。しかし、がん細胞株を用いた従来の前臨床がんモデルは、腫瘍内の不均一性や *in vivo* の腫瘍活性を再現できないなど、いくつかの限界があることが知られている。Cancer Cell Line Encyclopedia (CCLE) プロジェクトは、米国ブロード研究所とNovartisが提供するヒトがん細胞株のデータポータルである^{16,17}。CCLEでは現在、約1,500種に及ぶ様々ながん細胞株における遺伝子発現、変異、RNAスプライシング、DNAメチル化、ヒストンH3修飾、マイクロRNA発現、逆相タンパク質アレイ (Reverse Phase Protein Array, RPPA) などのデータがまとめられている。CCLEで提供されているデータに加え、米国ブロード研究所と英国サンガー研究所が中心となって提供しているがん依存性マップ (the Cancer Dependency Map, depmap) による薬剤感受性などの薬理学データとの関

連付けを行うことで、抗がん剤および関連するバイオマーカー探索などを効率的にすすめることができるようになってきている¹⁸。Depmap では、500 以上の細胞株について、遺伝子依存やオミクスデータ、薬剤感受性のデータがまとめられている。がん細胞株を腫瘍モデルとして用いるにあたり、ゲノム発現や変異がどのように細胞の性質に反映されているかを知るためにもプロテオーム解析は重要である。

タンパク質の受ける翻訳後修飾の中で、リン酸化酵素（キナーゼ）によって触媒されるリン酸化は非常に重要性が高く、細胞周期、増殖、アポトーシス、シグナル伝達経路といった様々な細胞プロセスの調節において重要な役割を果たしている。さらに、リン酸化シグナルはがん細胞のさまざまな機能を制御し、リン酸化シグナルを標的にした抗がん剤も多数開発されている。そのため、リン酸化タンパク質の総体であるリン酸化プロテオームによる知見は、新薬の開発や診断などにおいて新たな展開をもたらすと期待されている。LC/MS/MS を用いたリン酸化プロテオミクスの発展により、一度の分析で数万種のリン酸化部位が定量できるようになった。しかし、20 万種程度知られているヒトタンパク質リン酸化部位のうち、責任キナーゼが分かっているものは少ない（~5%）¹⁹。その要因の一つとして、キナーゼと基質の相互作用は一過的であることが挙げられる。そのため、得られた大規模データにはその基質へのリン酸化反応を直接触媒した責任キナーゼを示す情報が含まれない。そのため、リン酸化修飾を担ったキナーゼに情報を収斂させるアプローチが望ましいと考えられる。

第 1 章では、乳がん細胞株と腫瘍組織のプロテオームおよびリン酸化プロテオームプロファイルを比較し、がん細胞株がプロテオームレベルで腫瘍モデルとして妥当かどうかを検討した。第 2 章では、EGFR-MAPK シグナル伝達経路に参与する 10 種類のキナーゼを標的としたキャピラリー LC/MS/MS による SRM アッセイプラットフォームを開発し、HeLa 細胞内の 10 種類のキナーゼの活性を測定した。

第1章

ヒト乳がん細胞株のプロテオームおよびリン酸化プロテオーム プロファイリングと腫瘍モデルとしての評価

第1節 はじめに

乳がんは日本人女性において最も罹患率の高いがんである。国立がん研究センターがん情報サービスがん登録・統計サイトによる「最新がん統計」によると、日本人女性が生涯で乳がん罹患する確率は、10.6 % (9人に1人) となっている。一口に乳がんといっても様々な性質があり、その予後や治療効果などの臨床経過は患者個人によって著しく異なる。従来、臨床経過を予測するための乳がんの分類方法として、病理学的組織分類に加えて、estrogen receptor (ER)、progesterone receptor (PR) などのホルモン受容体や human epidermal growth factor receptor2 (Her2) による分類が行われてきた。加えて、2000年に Perou らによって、cDNA マイクロアレイを用いた遺伝子発現の網羅的解により、乳がんの内因性サブタイプ分類 (intrinsic subtype) の概念が提唱された^{20,21}。PAM50 (Prediction Analysis for Microarrays 50)²² は、乳がんの内因性サブタイプ分類のための最も一般的な遺伝子検査の一つで、2015年に National Comprehensive Cancer Network (NCCN) (<http://www.nccn.org/>) によって予後遺伝子シグネチャーアッセイとして認められている。乳がん試料から mRNA を抽出後、マイクロアレイによる遺伝子発現のクラスター解析を行ったところ、大きく①luminal A、②luminal B、③basal-like、④Her2amp、⑤normal breast-like に分類されることが明らかになった^{21,22}。これらのサブタイプごとに全生存率や無再発生存率について解析を行うと、サブタイプごとに予後と有意な相関がみられた。その後、化学療法や内分泌療法の治療効果とサブタイプ分類との間に相関がみられ、本サブタイプ分類は薬剤感受性を示す効果予測因子としての側面も重要視されている^{23,24}。この内因性サブタイプ分類と、従来臨床現場で使用されてきた ER、PR、Her2 発現との関係を見てみると、luminal A は大部分が ER 陽性、luminal B もまた大部分 ER 陽性だが、ER や関連遺伝子の発現レベルが低い、basal-like は ER、PR、Her2 がいずれも陰性のいわゆる「トリプルネガティブ」に大部分が含まれ、Her2amp は ERBB2 (Her2) 遺伝子の高発現と対応していることが多い。

初代組織を用いたがんの機構論的研究が困難であることから、*in vitro* および *in vivo* 実験のための細胞モデルの必要性は明らかである。しかし、これらの細胞モデル

が疾患を反映していなければ、得られた実験データに臨床的有用性がほとんどないため、可能な限り腫瘍を代表するものでなければならない。1958年に最初の乳がん細胞株、BT-20が確立され²⁵、ほとんどの細胞株は1970年代後半に確立された。多くの細胞株が確立されてから時間が経過しているため、長期間増殖させた場合の突然変異の蓄積や交差汚染のリスクと相まって、多くの細胞株の組織起源の不正確な割り当てにつながっている。最近では、TCGA^{1,2}が確立されたことにより、ゲノム情報やトランスクリプトーム情報に基づいて最も代表的な細胞株を選択することができるようになってきている。TCGAは、2006年から米国で開始された大型がんゲノムプロジェクトであり、20種類以上のがん種について、ゲノム・メチル化異常、遺伝子・タンパク質発現異常について網羅的な解析を行っている。乳がんについては、1098の腫瘍サンプルのmRNA発現プロファイル、DNAエクソーム解析、コピー数多型 (copy number variation, CNV)、DNAメチル化、およびタンパク質発現情報を公開している²⁶。本データにより、多くの重要な乳がん遺伝子や経路が系統的に検出されている^{27,28,29}。Liuらは、遺伝子発現の特徴を比較することで、広く用いられている乳がん細胞株を乳がんモデルとして評価した³⁰。さらに、臨床で見られるさまざまなサブタイプの転移性乳がんにより密接に類似した細胞株を同定した。Jiangらは、TCGAコンソーシアムを通じて入手可能な1375の原発性乳がんとCCLEの68の乳がん細胞株との間で、mRNA発現、CNV、タンパク質発現等の包括的な比較を行ったところ、PAM50遺伝子の発現に基づいて、細胞株と腫瘍の両方においてluminal AとB、Her2amp、Basal-likeの4つの主要な乳癌サブタイプに分類された³¹。タンパク質発現では、ER、PR、Her2の発現は細胞と腫瘍の間で一貫したパターンを示していたが、逆相タンパク質アレイによるその他のタンパク質については低い相関が見られた。

これらの研究は、乳がんのゲノムプロファイルや逆相タンパク質アレイによる限られた数のタンパク質発現などの観点から細胞株と腫瘍の比較を行っており、プロテオームおよびリン酸化プロテオームベースで腫瘍と細胞株の間の相関を比較した報告はない。ゲノムレベルおよびトランスクリプトームレベルの解析は、必ずしも表現型を定義するプロテオームプロファイルを反映しているとは限らない。本章では、これらの研究で報告された分子特性が、どの程度プロテオームレベルで表現されているか乳がん細胞株を用いて検証した。また、従来の臨床サブタイプおよびPAM50による内因性サブタイプに対するプロテオームおよびリン酸化プロテオームマーカーを同定した。最後に、プロテオーム、リン酸化プロテオームプロファイル、および、乳がん関連分子の発現の観点から、腫瘍モデルとしての乳がん細胞株を評価した。

第2節 乳がん細胞株のプロテオームおよびリン酸化プロテオームプロファイリング

ゲノム解析によって分類された 24 の乳がん細胞株のプロテオームおよびリン酸化プロテオーム解析を行った (表 1.1、図 1.1)。これらの細胞株は、PAM50 で定義された 4 種類のサブタイプを含んでおり²²、3 つの luminal A 細胞株、4 つの luminal B 細胞株、12 の Basal-like 細胞株、そして 5 つの Her2 (ERBB2) amp 細胞株が含まれる³¹。乳がん細胞株の PAM50 分類は Jiang らの論文に従った³¹。本章では、Tandem mass tag (TMT) ベースのアイソバリック標識によりタンパク質グループ ($n = 2$) およびリン酸化部位 ($n = 4$) の同定、定量を行った。24 種類の乳がん細胞株からタンパク質を抽出し、酵素消化して得られたペプチド試料を TMT 標識後、図 1.1A のバッチ J~P に示すように混合し、LC/MS/MS 測定を行った。イオン化された各細胞株由来の TMT 標識ペプチドは、フラグメンテーション後、MS/MS スペクトルにおいて低 m/z 領域に固有のレポーターイオンを生成する。データベース検索で同定されたタンパク質に対し、それぞれのレポーターイオンの強度を定量した。

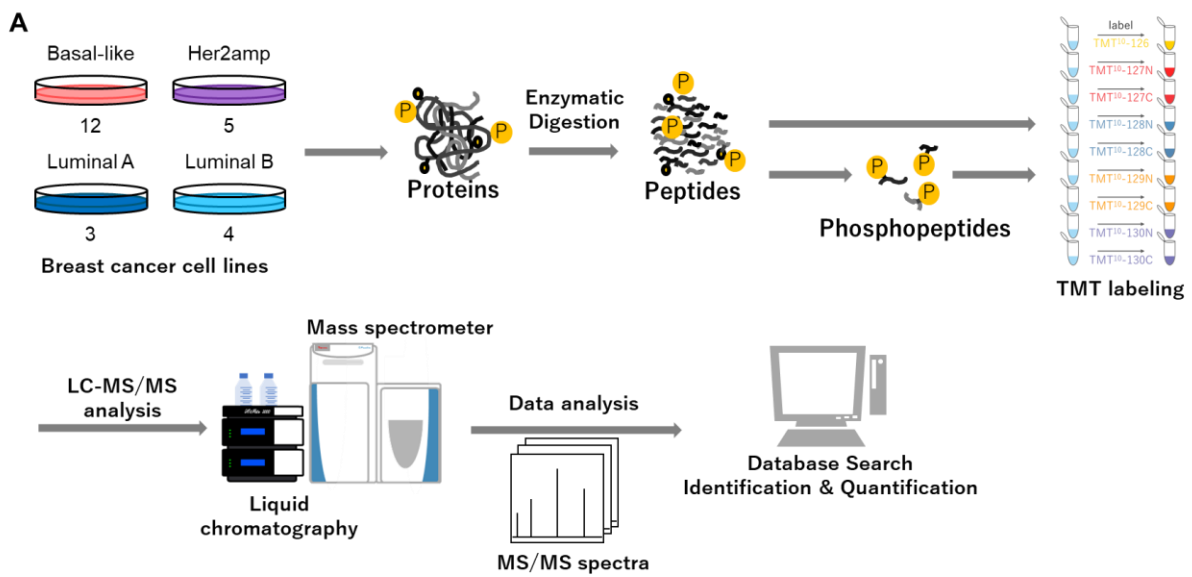
アイソバリック標識におけるバッチ間での欠損値を軽減するために、Isobaric match between runs (IMBR) を行った³²。IMBR では、同定に必要な y および b 系列のイオンの検出率が低いことにより、ペプチドの同定には至らなかった MS/MS スペクトルを保持時間およびプリカーサーイオンの同位体パターンから検索し、MS/MS スペクトルのレポーターイオン強度に基づいて定量する。図 1.2 では、プロテオームおよびリン酸化プロテオームデータセットについて、一定数の TMT バッチで検出されたタンパク質グループとリン酸化サイトの分布を IMBR の有無で示している。IMBR を適用することで、すべてのサンプルで一貫して検出され、定量されたタンパク質グループの数は、IMBR を適用しない場合に比べてわずかに増加した。

マルチプレックス TMT は定量的なプロテオミクスのための有用な方法であるが、複数の TMT バッチ間で大規模なプロテオーム解析を行う際には、各バッチの定量値の偏りがデータ品質に影響を与える可能性があることが知られている。バッチ間の影響を軽減するために、タンパク質発現推論³³ から標準基準線³⁴ まで、いくつかの方法が提案されている。本章では、バッチ間の影響を制御するための内部標準試料として 64 種類のがん細胞株から抽出したペプチドの等量混合物を使用した (図 1.1)。さらに、Cox らの論文³² にしたがって正規化を行った後、バッチ間の影響を評価するためにクラスタリングを行った (図 1.3)。その結果、プロテオームデータセット ($n = 2$) とリン酸化

プロテオームデータセット ($n = 4$) の樹形図で、同じ細胞株がクラスターを形成し、バッチ効果の軽減を確認した。

表 1.1 使用した乳がん細胞株

Cell Name	PAM50 subtype	ER status	PR status	Her2 status
MDA-MB-468		Negative	Negative	Negative
DU4475		Negative	Negative	Negative
HCC-1806		Negative	Negative	Negative
HCC-38		Negative	Negative	Negative
HCC-1143		Negative	Negative	Negative
HCC1937	Basal-like	Negative	Negative	Negative
HCC70		Negative	Negative	Negative
BT-20		Negative	Negative	Negative
Hs578T		Negative	Negative	Negative
BT-549		Negative	Negative	Negative
MDA-MB-231		Negative	Negative	Negative
MDA-MB-436		Negative	Negative	Negative
HCC1419		Positive	Negative	Positive
UACC-812		Positive	Negative	Positive
HCC1954	Her2amp	Negative	Negative	Positive
SKBR3		Negative	Negative	Positive
AU565		Negative	Negative	Positive
MCF-7		Positive	Positive	Negative
HCC1500	Luminal A	Positive	Positive	Negative
HCC2218		Negative	Negative	Positive
T47D		Positive	Positive	Negative
HCC1428	Luminal B	Positive	Positive	Negative
BT-474		Positive	Positive	Positive
MDA-MB-453		Negative	Negative	Positive



B

Batch	J	K	L	M	N	O	P
TMT10-126	Ctrl	Ctrl	Ctrl	Ctrl	Ctrl	Ctrl	Ctrl
TMT10-127N	MDA-MB-468	MDA-MB-468	Hs578T	Hs578T	BT-474	BT-474	
TMT10-127C	DU4475	DU4475	BT-549	BT-549	HCC1419	HCC1419	
TMT10-128N	HCC-1806	HCC-1806	MDA-MB-231	MDA-MB-231	UACC-812	UACC-812	
TMT10-128C	HCC-38	HCC-38	MDA-MB-436	MDA-MB-436	HCC1954	HCC1954	
TMT10-129N		HCC-1143	HCC-1143	MCF-7	MCF-7	SKBR3	SKBR3
TMT10-129C		HCC1937	HCC1937	T47D	T47D	HCC2218	HCC2218
TMT10-130N		HCC70	HCC70	HCC1500	HCC1500	AU565	AU565
TMT10-130C		BT-20	BT-20	HCC1428	HCC1428	MDA-MB-453	MDA-MB-453

図 1.1 実験ワークフロー(A)と TMT ラベル(B)

A) 乳がん細胞からタンパク質を抽出し、消化酵素によりペプチドに断片化して得られたペプチドおよび濃縮したリン酸化ペプチドを TMT 標識し、LC/MS/MS 測定を行った。B) J~P のバッチにおける TMT ラベルとコントロール試料 (TMT10-126) および乳がん細胞株 (TMT10-127N~130C) の組み合わせを示す。

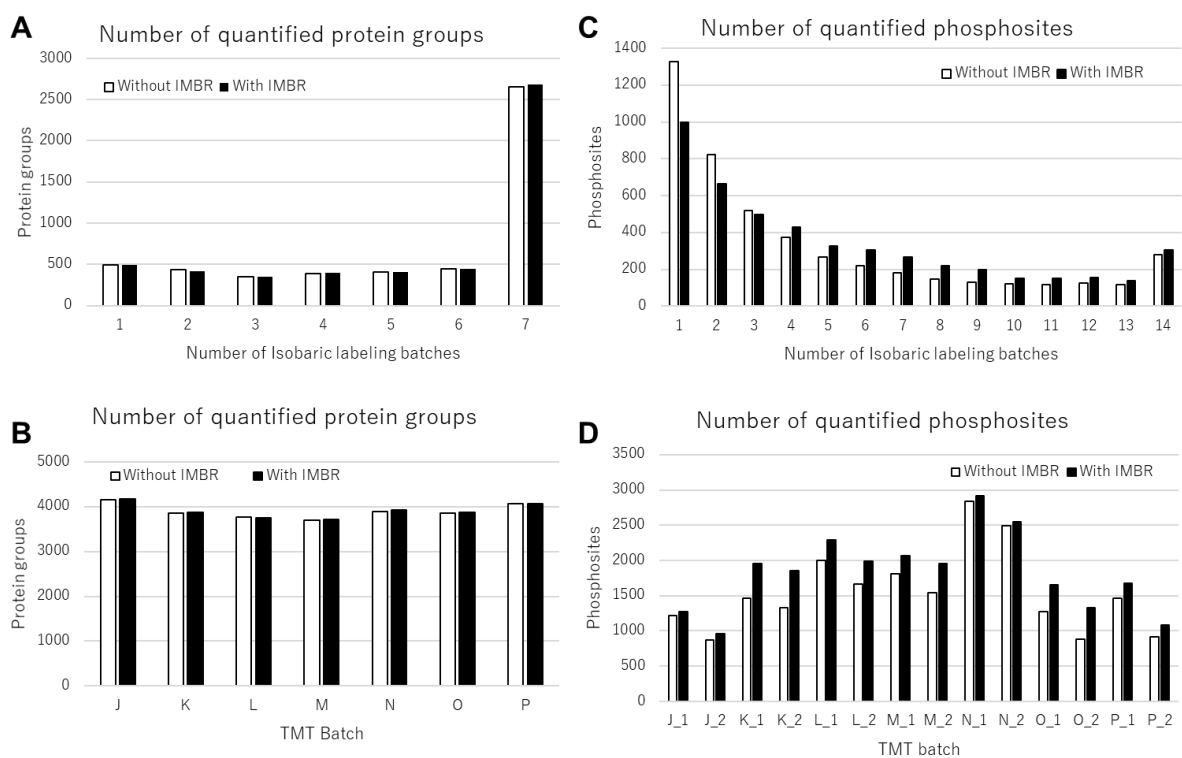


図 1.2 IMBR の有無による定量されたタンパク質グループおよびリン酸化サイト数

A、C) n 種類のタンパク質グループおよびリン酸化サイトが何個の TMT バッチで定量されたかを示す。白色のバーと黒色のバーは、それぞれ IMBR を用いない場合と用いた場合の結果を示す。B、D) 各 TMT バッチで定量されたタンパク質グループおよびリン酸化サイト数。

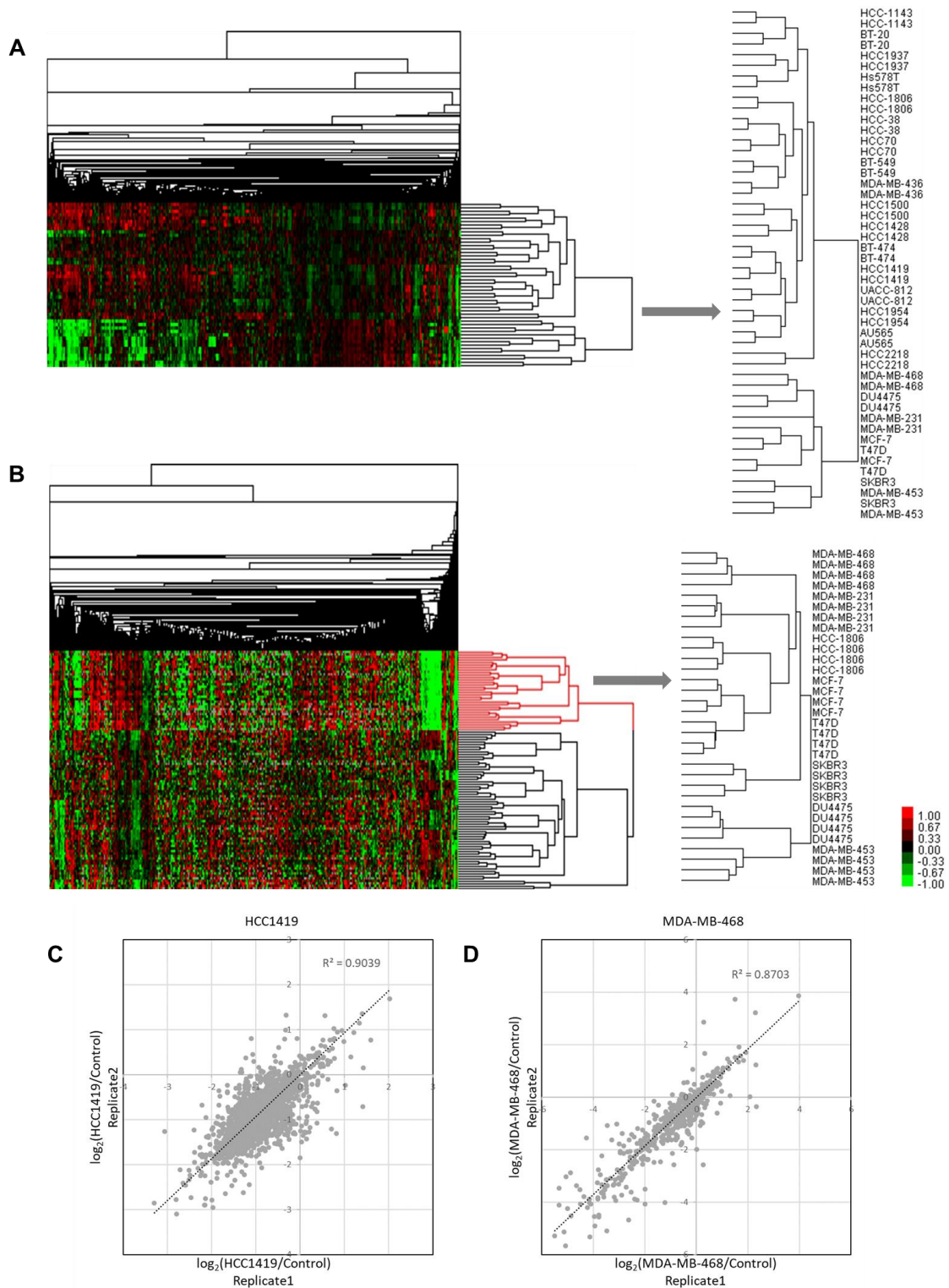


図 1.3 階層的クラスタリングと繰り返し測定間の再現性

A) 5,177 タンパク質グループの発現プロファイルに基づくクラスタリング。B) 3,508 リン酸化サイトの発現プロファイルに基づくクラスタリング。C) プロテオームデータセットでの再現性の一例。D) リン酸化プロテオームデータセットでの再現性の一例。

プロテオームデータセットでは、細胞株あたり平均 4,400 タンパク質グループ、2,800 ペプチドが同定、定量され、合計で 5,177 タンパク質グループの定量値 ($\log_2(\text{sample}/\text{control})$) が得られた (図 1.4)。これらのタンパク質は、150 種類のキナーゼや 160 種類の転写因子を含んでいた。リン酸化プロテオームデータセットでは、細胞株あたり平均 3,500 リン酸化ペプチドが同定、定量された。定量された合計 4,893 リン酸化ペプチドのうち、1 リン酸化ペプチド 4153、2 リン酸化ペプチド 655、3 リン酸化ペプチド 82、4 リン酸化ペプチドが 3 であった。今後の解析には、1 リン酸化ペプチドのみを用いた。その結果、細胞株あたり平均 2,400 リン酸化サイト、合計 3,508 リン酸化部位の定量値 ($\log_2(\text{sample}/\text{control})$) が得られた。また、リン酸化サイトの分布は、 $\text{pS}:\text{pT}:\text{pY} = 88.1:10.7:1.2$ であった (図 1.5)。

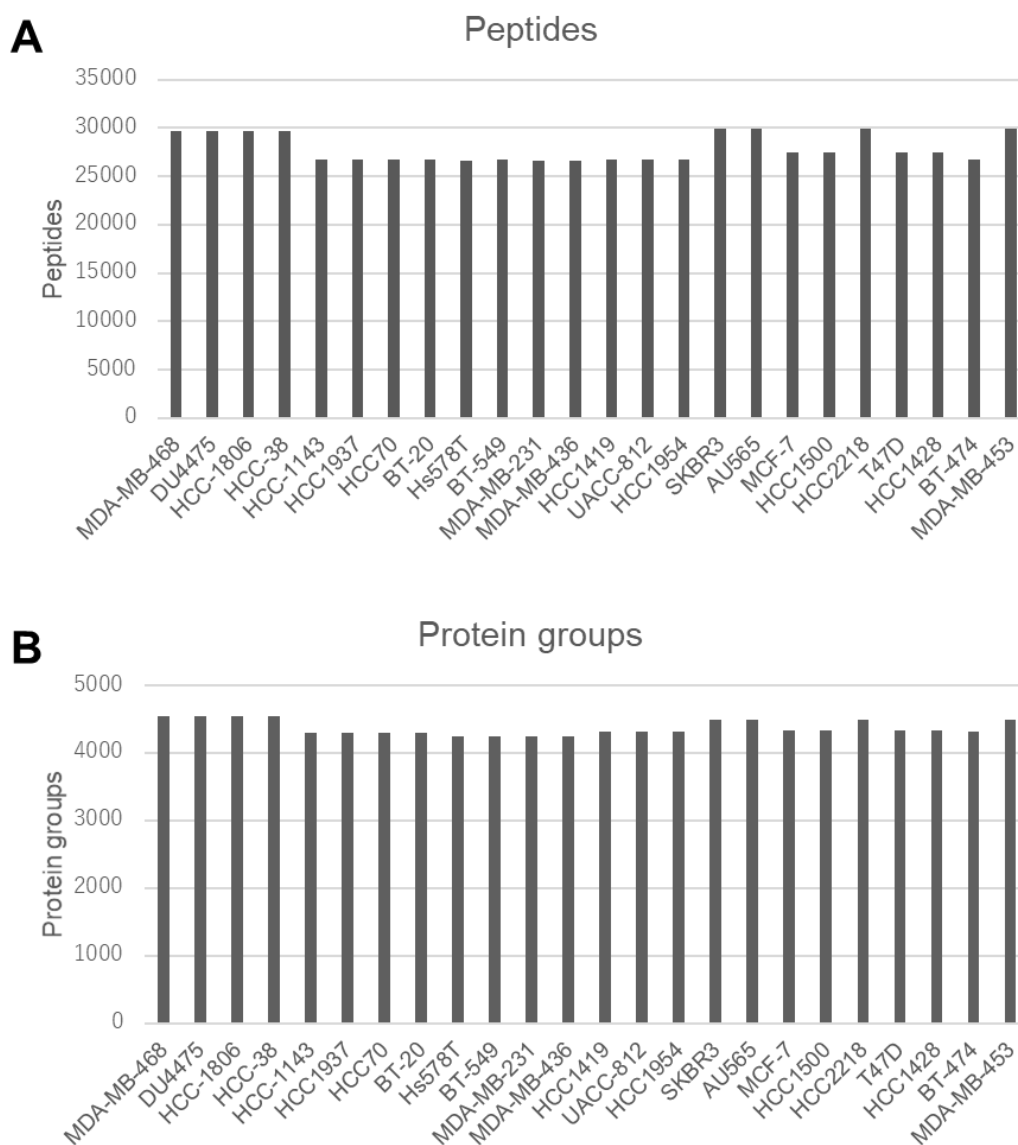


図 1.4 細胞株ごとの定量されたペプチド数 (A) およびタンパク質グループ数 (B)

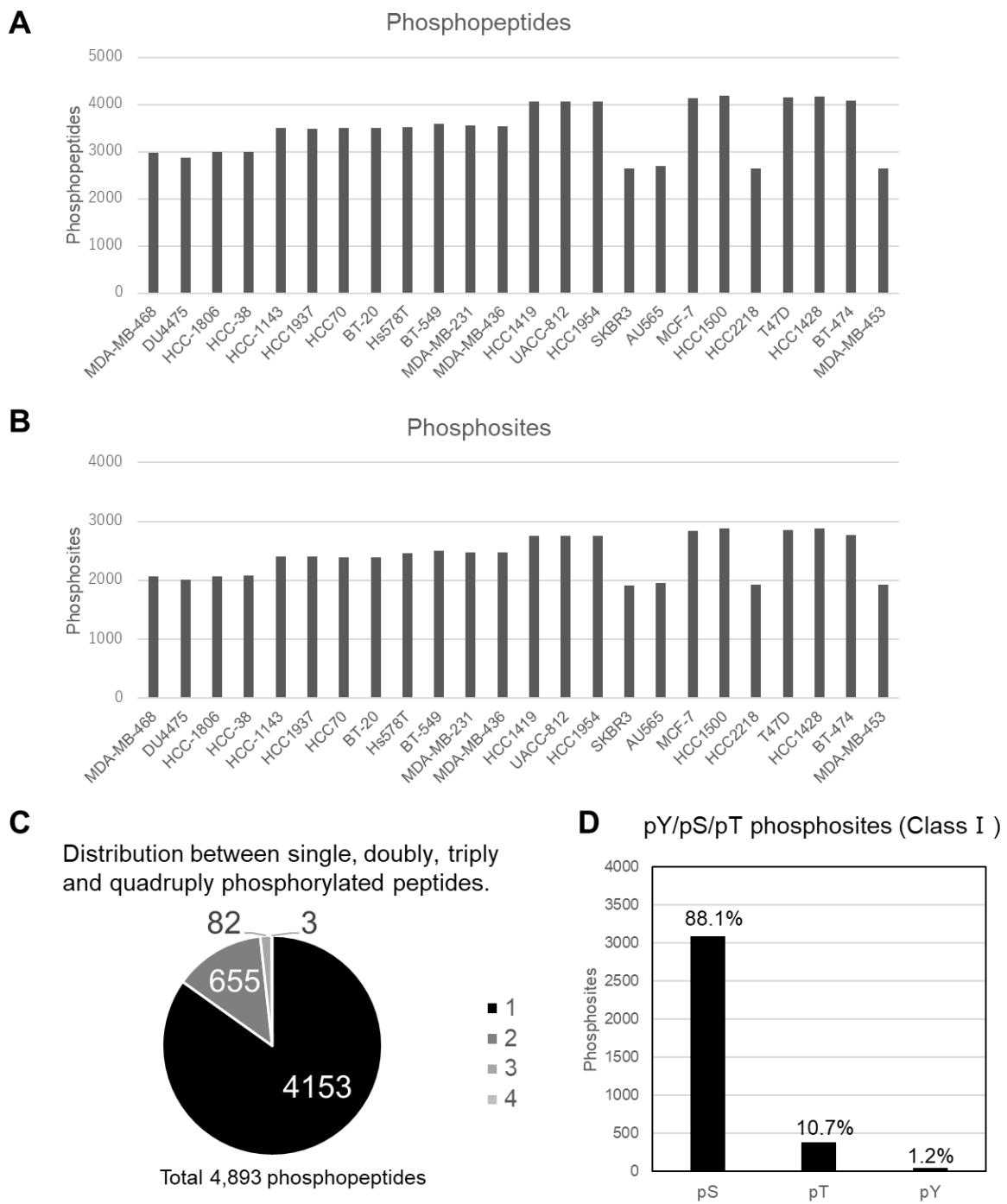


図 1.5 リン酸化プロテオームプロファイル

A) 細胞株ごとの定量されたリン酸化ペプチド数。B) 細胞株ごとの定量されたリン酸化サイト数。
 C) 1 リン酸化ペプチド、2 リン酸化ペプチド、3 リン酸化ペプチド、4 リン酸化ペプチドの割合。D)
 pS/pT/pY リン酸化サイト数。

第3節 乳がんの臨床的サブタイプとの相関

乳がん診療では長い間、腫瘍の浸潤径、腋窩リンパ節転移状況、組織型、病理学的悪性度、エストロゲン受容体 (ER)、プロゲステロン受容体 (PR)、Her2 発現状況等により、予後予測や治療方針決定がなされてきた。24 細胞株の 5,177 のタンパク質グループの発現に基づく階層型クラスタリングにより、2 つの主要な細胞株グループが得られた (図 1.6A)。これらのクラスターの組み合わせと ER/PR/Her2 の発現プロファイルと比較すると、乳がん細胞のプロテオームは、ER/PR/Her2 の発現のみには依存しないプロテオームプロファイルを示した。また、同様に 24 細胞株の 3,508 のリン酸化サイトに基づく階層型クラスタリングにより、乳がん細胞株のリン酸化プロテオームは、ER/PR/Her2 の発現のみには依存しないリン酸化プロテオームプロファイルを示唆した (図 1.6B)。また、プロテオームプロファイルとリン酸化プロファイルのクラスターを比較すると、一部の細胞株は類似したクラスターを形成していたが、異なる細胞株同士によるクラスターも形成されており、乳がん細胞株は異なるプロテオームおよびリン酸化プロテオームプロファイルを持つことが示された (図 1.6C)。

次に、ER/PR/Her2 の発現と相関するタンパク質の小グループを決定することを目的として、解析を行った。5,177 種類のタンパク質グループについて、ER 陽性細胞株と陰性細胞株を比較したところ、36 種類のタンパク質グループが異なる発現を示していた (Welch の t 検定; FDR < 0.05; 図 1.7A)。発現変動しているタンパク質グループの中には、ER 陽性乳がんで高発現している既知のマーカー、TFF1 (Trefoil factor 1)³⁵ が同定された。同定された 36 種類のタンパク質グループについてクラスタリングを行ったところ、ER 陽性細胞株からなるクラスターが形成された (図 1.7B)。さらに、抽出されたタンパク質グループの妥当性を確認するために ChIP-Atlas³⁶ の Enrichment analysis を用いて発現変動しているタンパク質グループの遺伝子に結合しうる、あるいは上流でそれらの遺伝子の発現を制御する可能性がある転写因子を検索した。ChIP-Atlas は、論文などで報告された ChIP-seq データを閲覧、活用するためのウェブサービスで、データソースは、公開 NGS データレポジトリ (NCBI、EMBL-EBI、DDBJ) に登録されたほぼ全ての ChIP-seq データである。その結果、ER 陽性細胞株で発現が上昇していたタンパク質グループの上流には、ESR1 や ER 関連遺伝子として知られる FOXA1 および GATA3 が抽出され、本アプローチの妥当性を支持した^{37,38}。同様の解析を PR 陽性株/陰性株、Her2 陽性株/陰性株についても行い、それぞれ、25 種類、45 種類の

発現変動しているタンパク質グループを同定した (図 1.8、1.9)。合計、ER/PR/Her2 の発現と相関する合計 91 種類のタンパク質グループが同定された (表 1.2)。

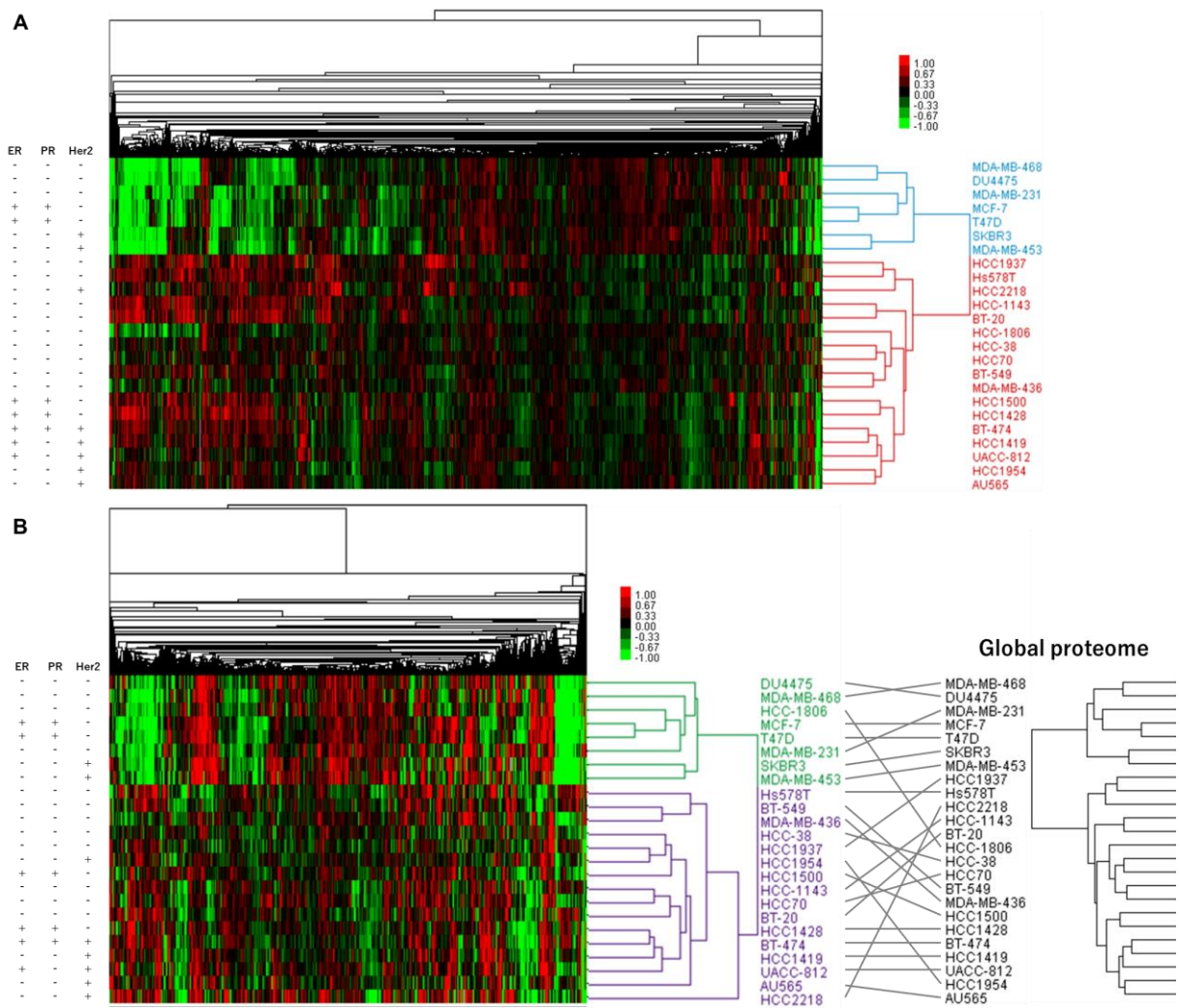


図 1.6 プロテオームクラスタリングと ER、PR、Her2 発現との関連

A) 24 細胞株と 5,177 のタンパク質グループのクラスタリングと ER、PR、Her2 発現との関係、B)

24 細胞株と 3,508 のリン酸化サイトのクラスタリングと ER、PR、Her2 発現との関係

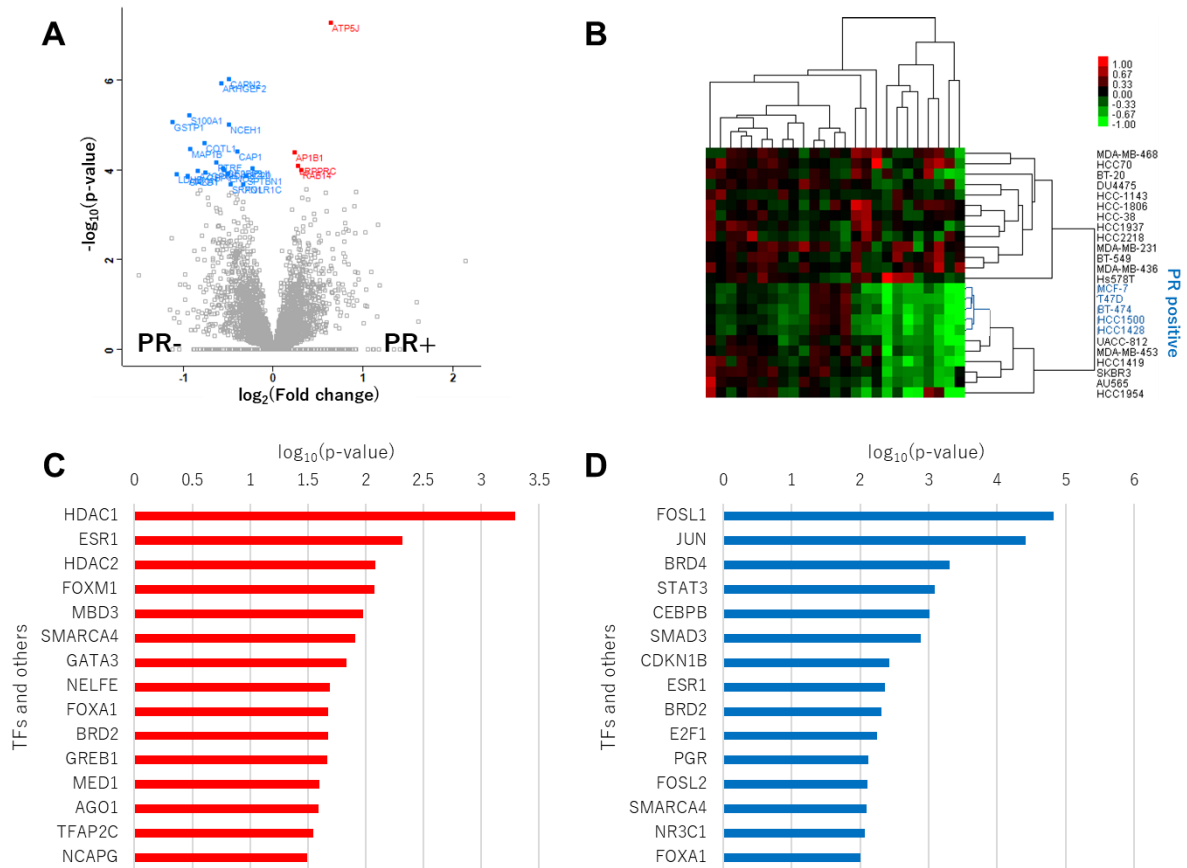


図 1.8 PR と発現が相関するタンパク質グループの同定

A) PR 陽性株と陰性株の間で発現が異なるすべてのタンパク質について、Welch の t 検定由来の $-\log_{10}(\text{p-value})$ (y 軸) に対する $\log_2(\text{Fold change})$ (x 軸) を表示したボルケーノプロット (Benjamini-Hochberg (BH) $\text{FDR} < 0.05$)。有意に減少したタンパク質は青、有意に増加したタンパク質は赤で示す。B) データセット中の 25 種類の有意に変化したタンパク質を用いたクラスター分析。C、D) ChIP-Atlas による PR 陽性株で有意に増加したタンパク質 (C) と減少したタンパク質 (D) の遺伝子に結合しうる、あるいは上流でそれらの遺伝子の発現を制御する可能性がある転写因子

表 1.2 ER/PR/Her2 発現と相関して有意に変化したタンパク質

Gene names	Protein names	Significant correlation with ER/PR/Her2	
MYO1C	Unconventional myosin-Ic	ER	
BIN1	Myc box-dependent-interacting protein 1	ER	
CLDN3	Claudin-3	ER	
MCM3AP	Germinal-center associated nuclear protein	ER	
TACC1	Transforming acidic coiled-coil-containing protein 1	ER	
TFF1	Trefoil factor 1	ER	
LCP1	Plastin-2	ER	
PEBP1	Phosphatidylethanolamine-binding protein 1	ER	
GPC1	Glypican-1;Secreted glypican-1	ER	
SERPINB5	Serpin B5	ER	
PLOD1	Procollagen-lysine,2-oxoglutarate 5-dioxygenase 1	ER	
MTX1	Metaxin-1	ER	
NAB2	NGFI-A-binding protein 2	ER	
ANXA8L2;ANXA8	Annexin A8-like protein 2;Annexin A8	ER	
PPP1R18	Phostensin	ER	
TMEM205	Transmembrane protein 205	ER	
MINK1	Misshapen-like kinase 1	ER	
CLPTM1L	Cleft lip and palate transmembrane protein 1-like protein	ER	
PPIL3	Peptidyl-prolyl cis-trans isomerase-like 3	ER	
ARMT1	Protein-glutamate O-methyltransferase	ER	
TM9SF3	Transmembrane 9 superfamily member 3	ER	
IGF2BP3	Insulin-like growth factor 2 mRNA-binding protein 3	ER	PR
ANXA1	Annexin A1	ER	PR
LDHB	L-lactate dehydrogenase B chain	ER	PR
GSTP1	Glutathione S-transferase P	ER	PR
CAPN2	Calpain-2 catalytic subunit	ER	PR
ATP5PF	ATP synthase-coupling factor 6, mitochondrial	ER	PR
GBP1	Interferon-induced guanylate-binding protein 1	ER	PR
MAP1B	Microtubule-associated protein 1B	ER	PR
SPTBN1	Spectrin beta chain, non-erythrocytic 1	ER	PR
CALD1	Caldesmon	ER	PR
AP1B1	AP-1 complex subunit beta-1	ER	PR
COTL1	Coactosin-like protein	ER	PR
CAVIN1	Caveolae-associated protein 1	ER	PR
NCEH1	Neutral cholesterol ester hydrolase 1	ER	PR
AGPS	Alkyldihydroxyacetonephosphate synthase, peroxisomal	ER	Her2
POLR1C	DNA-directed RNA polymerases I and III subunit RPAC1		PR
ENO2	Gamma-enolase		PR
S100A1	Protein S100-A1		PR
AZGP1	Zinc-alpha-2-glycoprotein		PR
LRPPRC	Leucine-rich PPR motif-containing protein, mitochondrial		PR
RAB14	Ras-related protein Rab-14		PR
CAP1	Adenylyl cyclase-associated protein 1		PR
FLII	Protein flightless-1 homolog		PR
ARHGEF2	Rho guanine nucleotide exchange factor 2		PR
SRXN1	Sulfiredoxin-1		PR
KIDINS220	Kinase D-interacting substrate of 220 kDa		PR

(続く)

(続き)

Gene names	Protein names	Significant correlation with ER/PR/Her2
MIEF1	MIEF1 upstream open reading frame protein	Her2
CCN1	CCN family member 1	Her2
PMM2	Phosphomannomutase 2	Her2
EPB41L2	Band 4.1-like protein 2	Her2
ORC5	Origin recognition complex subunit 5	Her2
GLS	Glutaminase kidney isoform, mitochondrial	Her2
HLA-A	HLA class I histocompatibility antigen, A alpha chain	Her2
ERBB2	Receptor tyrosine-protein kinase erbB-2	Her2
ITGB1	Integrin beta-1	Her2
ITGAV	Integrin alpha-V	Her2
LGALS1	Galectin-1	Her2
AKR1B1	Aldose reductase	Her2
ANPEP	Aminopeptidase N	Her2
CD44	CD44 antigen	Her2
ITGB6	Integrin beta-6	Her2
MX1	Interferon-induced GTP-binding protein Mx1	Her2
ACSL1	Long-chain-fatty-acid--CoA ligase 1	Her2
SERPINH1	Serpin H1	Her2
MVD	Diphosphomevalonate decarboxylase	Her2
RAD23A	UV excision repair protein RAD23 homolog A	Her2
RAP1B	Ras-related protein Rap-1b	Her2
FABP5	Fatty acid-binding protein, epidermal	Her2
CAV1	Caveolin-1	Her2
CD47	Leukocyte surface antigen CD47	Her2
THOC5	THO complex subunit 5 homolog	Her2
FSCN1	Fascin	Her2
MICOS13	MICOS complex subunit MIC13	Her2
MRPL54	39S ribosomal protein L54, mitochondrial	Her2
HS2ST1	Heparan sulfate 2-O-sulfotransferase 1	Her2
CCDC50	Coiled-coil domain-containing protein 50	Her2
ASCC3	Activating signal cointegrator 1 complex subunit 3	Her2
GPX8	Probable glutathione peroxidase 8	Her2
PSMF1	Proteasome inhibitor PI31 subunit	Her2
TFG	Protein TFG	Her2
COG8	Conserved oligomeric Golgi complex subunit 8	Her2
VAT1	Synaptic vesicle membrane protein VAT-1 homolog	Her2
MIEN1	Migration and invasion enhancer 1	Her2
PSMG3	Proteasome assembly chaperone 3	Her2
HINT2	Histidine triad nucleotide-binding protein 2, mitochondrial	Her2
ELP3	Elongator complex protein 3	Her2
RTN4	Reticulon-4	Her2
PDLIM7	PDZ and LIM domain protein 7	Her2
CORO1C	Coronin-1C	Her2
ARFGEF1	Brefeldin A-inhibited guanine nucleotide-exchange protein 1	Her2

同様の解析をリン酸化プロテオームデータについても行った (Welch の t 検定; FDR < 0.05; 図 1.10)。ER 陽性株/陰性株、PR 陽性株/陰性株、Her2 陽性株/陰性株について、それぞれ、15、8、10 の発現変動しているリン酸化サイトを同定した。これにより、ER/PR/Her2 の発現と相関する合計 30 のリン酸化サイトが同定された (表 1.3)。プロテオームおよびリン酸化プロテオームから、同定されたこれらのタンパク質グループおよびリン酸化サイトの定量値に基づくクラスタリングの結果、図 1.6 のすべてのタンパク質グループおよびリン酸化サイトによるクラスタリングと比べて、ER/PR/Her2 の発現に相関したクラスターが得られ、ER/PR/Her2 の発現と相関するマーカー候補タンパク質が同定できたと考えられる (図 1.11)。

表 1.2 の 91 のタンパク質グループの遺伝子オントロジー (GO) ターム解析の結果から、ER/PR/Her2 の発現と相関するタンパク質は、細胞組成では extracellular exosome や focal adhesion、分子機能では、actin binding や actin filament binding、生物学的プロセスでは、negative regulation of phospholipase A2 activity や extracellular matrix organization に濃縮されていた (表 1.4)。同定された 30 のリン酸化サイトの内、表 1.2 のタンパク質上のリン酸化サイトは 5 サイトであった。この 5 つのリン酸化サイト (EPB41L2; S614、ERBB2; S1054、ERBB2; T1166、CAVIN1; S167、MINK1; S763) とタンパク質の発現を図 1.12 に示す。5 つのリン酸化サイトの存在する 4 タンパク質には、キナーゼである ERBB2 と MINK1 が含まれていた。また、Cavin1 としても知られるポリメラーゼ I および転写物放出因子 (PTRF) は、カベオラ形成の重要な因子として以前に報告されており、乳がんを含め、さまざまな悪性腫瘍で異常な PTRF 発現が報告されている³⁹。タンパク質の発現とそのリン酸化は、多くの細胞株で正の相関を示したが、いくつかの細胞株で負の相関を示した。

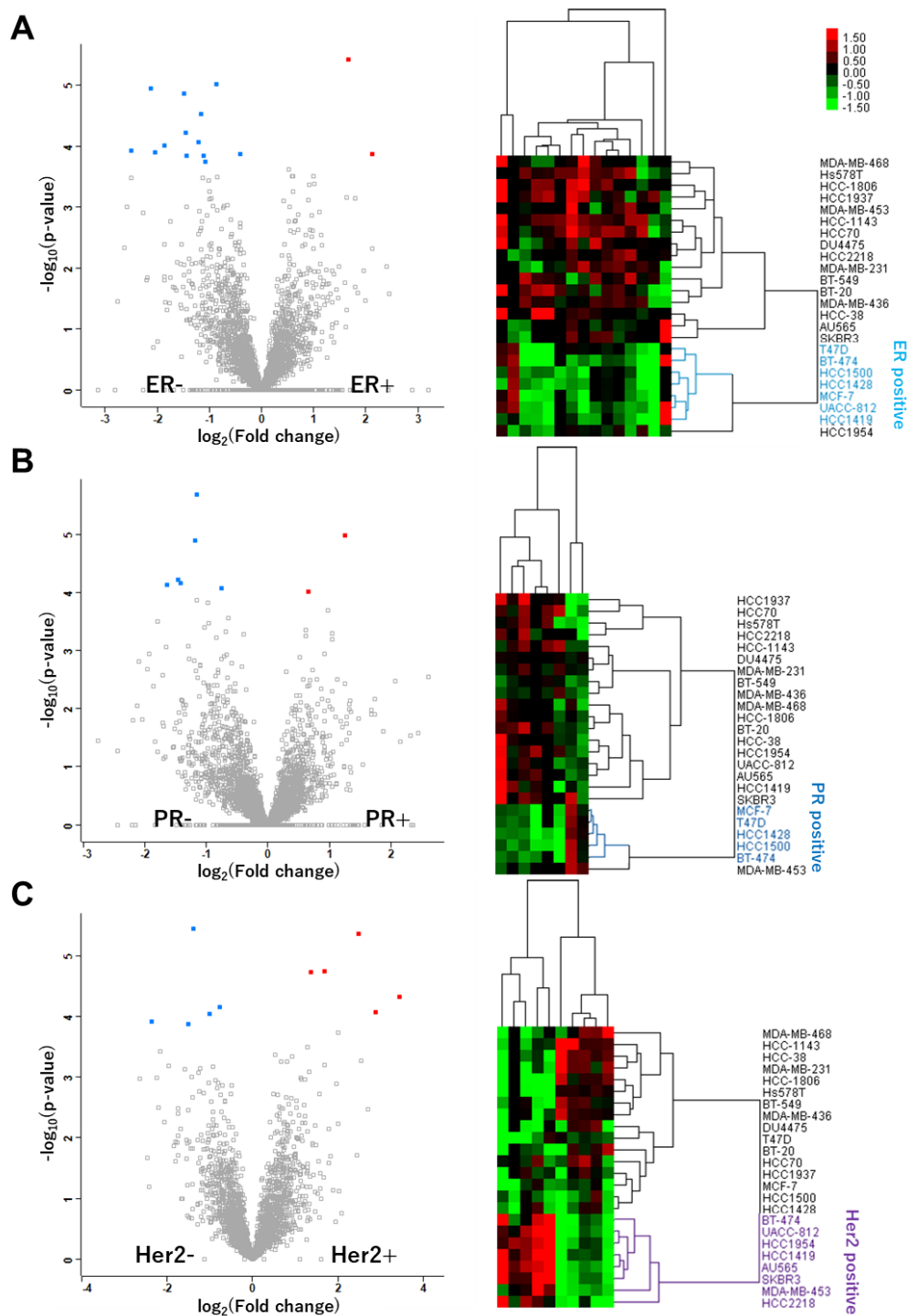


図 1.10 ER/PR/Her2 発現と相関するリン酸化サイト

A-C)ER (A)、PR (B)、Her2 (C) 陽性株と陰性株の間で発現が異なるすべてのタンパク質について、Welch の t 検定由来の $-\log_{10}(\text{p-value})$ (y 軸) に対する $\log_2(\text{Fold change})$ (x 軸) を表示したボルケーノプロット (Benjamini-Hochberg (BH) FDR < 0.05)。データセット中で ER/PR/Her2 発現と相関して有意に変化したリン酸化サイトを用いたクラスター分析。

表 1.3 ER/PR/Her2 発現と相関して有意に変化したリン酸化サイト

Gene names	Protein names	Phosphosites	Significant correlation with ER/PR/Her2	
AKAP2	A-kinase anchor protein 2	S393	ER	
BCAS1	Breast carcinoma-amplified sequence 1	S552	ER	
CAVIN1	Polymerase I and transcript release factor	S167	ER	
CCNYL1	Cyclin-Y-like protein 1	S344	ER	
FOSL2	Fos-related antigen 2	S200	ER	
FOKK2	Forkhead box protein K2	S398	ER	
GREB1	Protein GREB1	S1146	ER	
LUZP1	Leucine zipper protein 1	S659	ER	
MICAL3	Protein-methionine sulfoxide oxidase MICAL3	S1704	ER	
MINK1	Misshapen-like kinase 1	S763	ER	
SLC16A1	Monocarboxylate transporter 1	S467	ER	
SLC16A1	Monocarboxylate transporter 1	T466	ER	
TMEM230	Transmembrane protein 230	S24	ER	
LYN	Tyrosine-protein kinase Lyn	S13	ER	PR
SP100	Nuclear autoantigen Sp-100	S18	ER	PR
ARHGAP29	Rho GTPase-activating protein 29	S1019		PR
MYO18A	Unconventional myosin-XVIIIa	S140		PR
SPECC1L	Cytospin-A	S384		PR
TMPO	Lamina-associated polypeptide 2, isoforms beta/gamma	S184		PR
TOP2B	DNA topoisomerase 2-beta	S1466		PR
SERBP1	Plasminogen activator inhibitor 1 RNA-binding protein	S203	PR	Her2
AHNAK2	Protein AHNAK2	S260		Her2
CIAPIN1	Anamorsin	S185		Her2
EPB41L2	Band 4.1-like protein 2	S614		Her2
ERBB2	Receptor tyrosine-protein kinase erbB-2	S1054		Her2
ERBB2	Receptor tyrosine-protein kinase erbB-2	T1166		Her2
LMNA	Prelamin-A/C;Lamin-A/C	S652		Her2
PML	Protein PML	S403		Her2
RANBP2	E3 SUMO-protein ligase RanBP2	S2454		Her2
S100A16	Protein S100-A16	S2		Her2

表 1.4 ER/PR/Her2 発現と相関して有意に変化したタンパク質群に濃縮されている GO ターム
上位 20 種

Category	GO Term	Count	%	P-Value
GOTERM_CC_DIRECT	extracellular exosome	39	41.9	1.40E-09
GOTERM_CC_DIRECT	focal adhesion	14	15.1	7.10E-08
GOTERM_MF_DIRECT	actin binding	12	12.9	1.60E-07
GOTERM_CC_DIRECT	cytosol	37	39.8	1.40E-06
GOTERM_CC_DIRECT	membrane raft	9	9.7	8.90E-06
GOTERM_MF_DIRECT	actin filament binding	7	7.5	5.70E-05
GOTERM_BP_DIRECT	negative regulation of phospholipase A2 activity	3	3.2	1.60E-04
GOTERM_MF_DIRECT	protein binding	62	66.7	2.70E-04
GOTERM_BP_DIRECT	extracellular matrix organization	7	7.5	5.70E-04
GOTERM_CC_DIRECT	ruffle membrane	5	5.4	7.60E-04
GOTERM_BP_DIRECT	mRNA transport	4	4.3	1.90E-03
GOTERM_CC_DIRECT	lateral plasma membrane	4	4.3	2.40E-03
GOTERM_CC_DIRECT	stress fiber	4	4.3	2.50E-03
GOTERM_BP_DIRECT	cell adhesion mediated by integrin	3	3.2	2.70E-03
GOTERM_BP_DIRECT	cell proliferation	8	8.6	3.10E-03
GOTERM_BP_DIRECT	leukocyte migration	5	5.4	3.90E-03
GOTERM_MF_DIRECT	identical protein binding	11	11.8	4.70E-03
GOTERM_CC_DIRECT	actin cytoskeleton	6	6.5	4.70E-03
GOTERM_BP_DIRECT	actin cytoskeleton organization	5	5.4	4.80E-03
GOTERM_CC_DIRECT	cytoplasm	39	41.9	5.00E-03

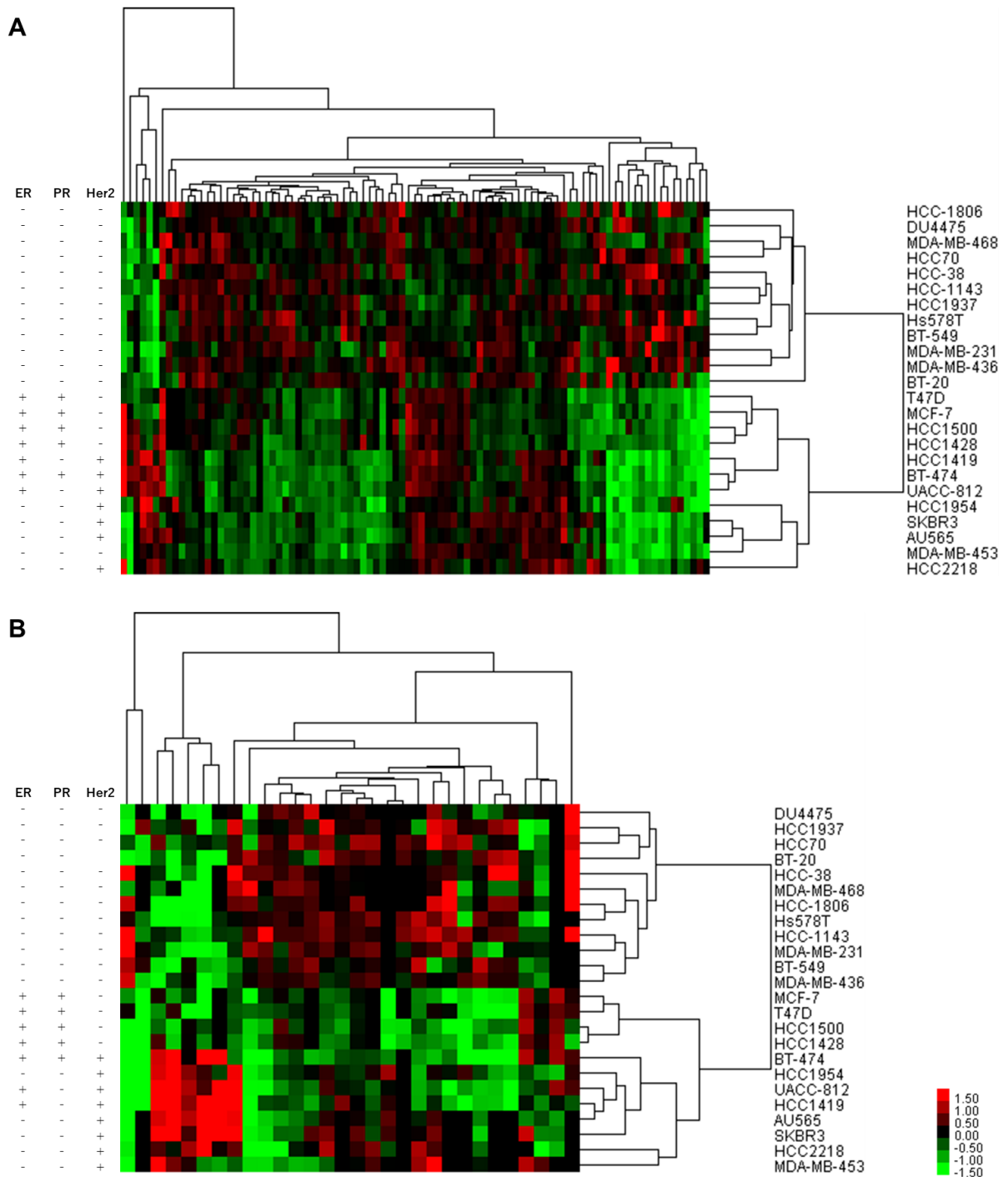


図 1.11 ER/PR/Her2 発現と相関するタンパク質グループおよびリン酸化サイトによるクラスタリング

A) プロテオームデータで ER/PR/Her2 発現と相関して有意に変化したタンパク質グループを用いたクラスタリング。B) リン酸化プロテオームデータセット中で ER/PR/Her2 発現と相関して有意に変化したリン酸化サイトを用いたクラスタリング。

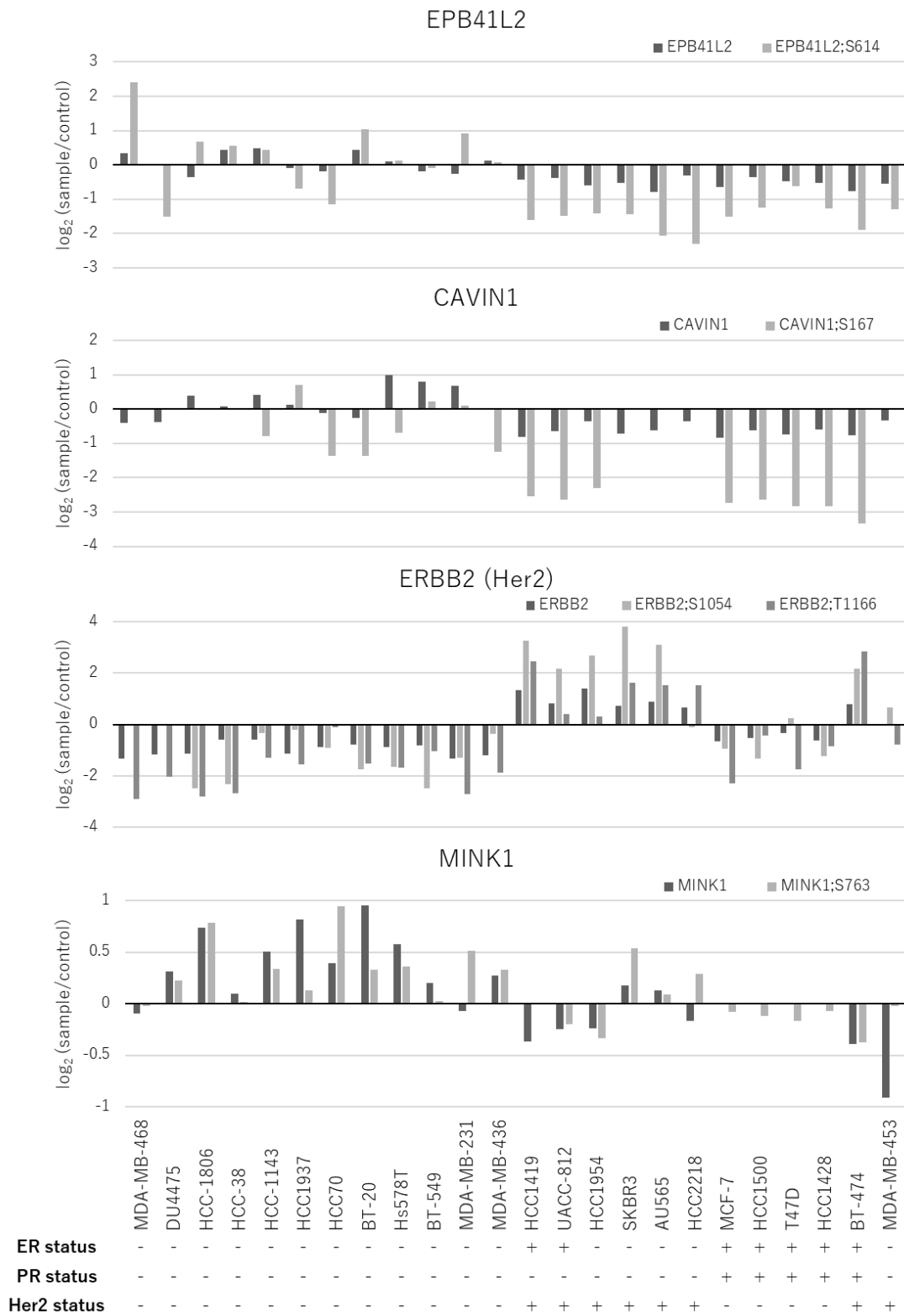


図 1.12 ER/PR/Her2 発現と相関して有意に変化したタンパク質上のリン酸化サイト
 データセット中で ER/PR/Her2 発現と相関して有意に変化したリン酸化サイトのうち、ER/PR/Her2 発現と相関して有意に変化したタンパク質上のリン酸化サイト

第4節 乳がんの内因性サブタイプとの相関

PAM50 によるサブタイプ (basal-like、Her2amp、luminal A、luminal B) 分類は非常に堅牢であることが証明されており、現在でも予後予測や治療レジメンの設計に広く用いられている⁴⁰。しかし、日常臨床においてマイクロアレイによる遺伝子発現情報を得ることは困難であることから、より簡易的なサブタイプの分類が必要とされ、Cheang らの報告⁴¹ などから、免疫組織化学 (IHC) マーカー (ER、PR、HER2、Ki67) の発現情報を組み合わせて、luminal A-like、luminal B-like、Her2 type、Triple negative type といった代替的な分類により、治療法を選択するようになった。

しかし、この方法では患者個人の予後や治療効果を予測するには十分とは言えず、3人に1人の患者が誤分類されている可能性が示唆されている^{42,43}。近年、個々の患者の再発リスクや治療効果を適切に判断するために種々の多遺伝子アッセイが開発・実用化されている (例: MammaPrint⁴⁴、Oncotype DX⁴⁵、Curebest 95GC Breast⁴⁶) が、いずれの検査も国内では保険適応されておらず、すべての患者が容易に利用できるわけではない。

近年、質量分析計を用いたハイスループットなタンパク質定量技術の進歩により、乳がんのタンパク質発現に基づいた分子特性評価が可能となった。プロテオームによる分類は、乳がんの腫瘍間の不均一性をより正確に反映したものであると考えられ、PAM50分類をタンパク質レベルで代替する新規の免疫組織化学的バイオマーカー候補が同定できると考えられる。本節では、プロテオームベースで、治療法の選択のための正確な内因性サブタイプ分類が困難な患者をより簡便に識別できる分子の同定を試みた。

プロテオームプロファイルの階層的クラスタリングは、細胞株の一部を PAM50 サブタイプと一致してクラスター化した (図 1.13A)。Basal-like サブタイプ、Her2amp サブタイプは一部でクラスターを形成した。しかしながら、luminal A サブタイプと luminal B サブタイプが混在しており、分子表現型の類似性を示していた。次に、定量プロテオームデータ中に含まれる PAM50 遺伝子 21 種のみを考慮したタンパク質発現プロファイルについて階層的クラスタリングを行ったが、PAM50 サブタイプと完全には一致しなかった (図 1.13B)。しかし、5,177 のタンパク質グループの定量値に基づくクラスターと PAM50 メンバーのみのクラスターは、類似しており、細胞株のプロテオーム全体のプロファイルがより小さいサブセットに由来することを示した (図 1.13C)。

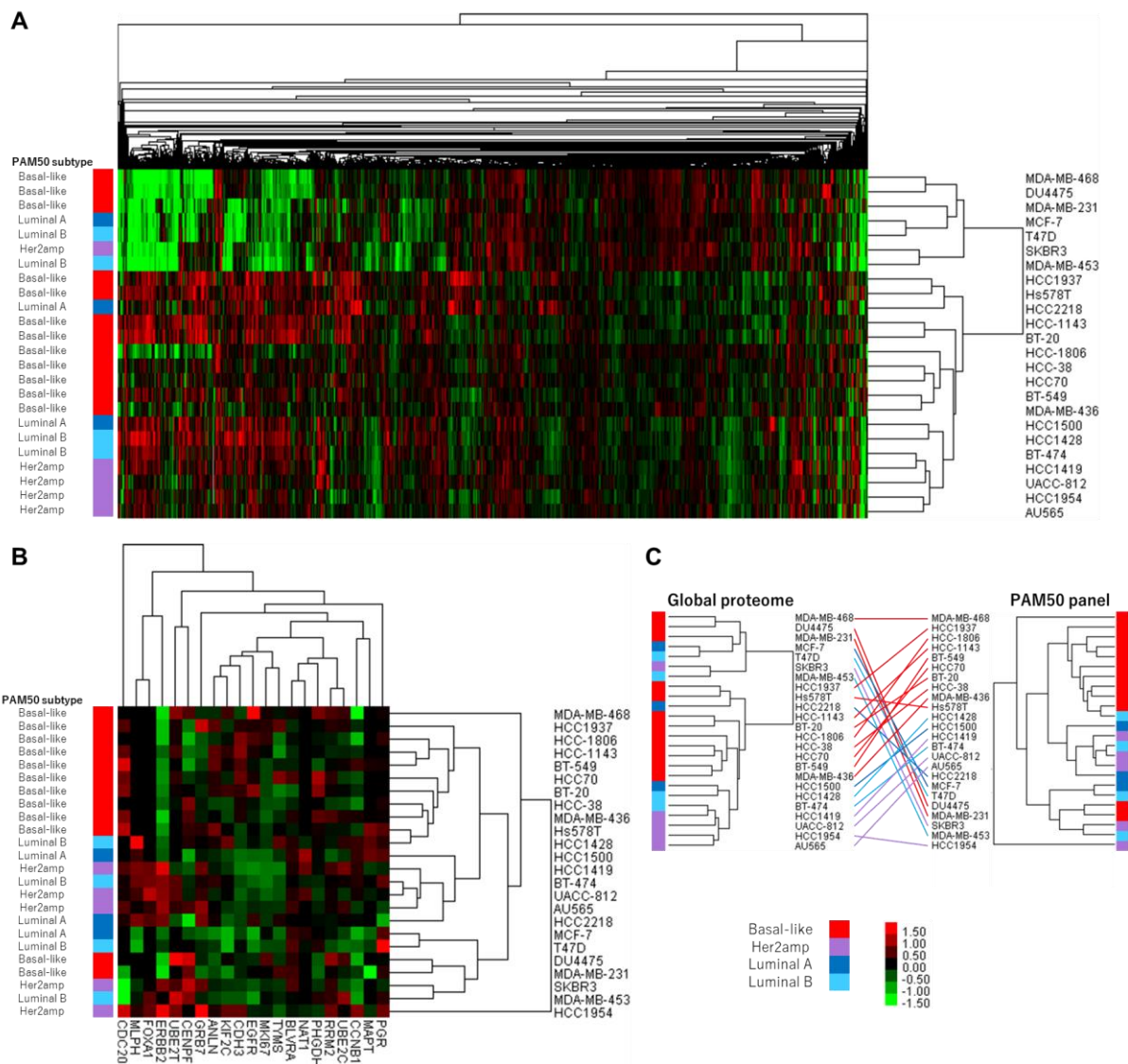


図 1.13 乳がん細胞株のプロテオームプロファイルと PAM50 サブタイプの相関

A) グローバルプロテオームと PAM50 サブタイプ。B) PAM50 パネルから同定され、定量されたタンパク質のクラスタリング (n = 21)。C) グローバルプロテオームと PAM50 パネル分子によるクラスターの比較

次に、分散分析 (ANOVA 検定、FDR < 0.05) を行い、4 つのサブタイプのいずれかの間で有意に変化する 59 種類のタンパク質グループを同定した。これらのタンパク質グループの階層的クラスタリングは、2 つの主要なクラスターを示し、そのうちの1つのクラスターは、1 つのサブタイプ (Basal-like) にのみ特異的な (より高発現の) タンパク質を含み、他のクラスターは 3 つのサブタイプ間で共有されていた (図 1.14A)。Basal-like クラスターには、CD44、CAV1、GPC1、TGB1、MSN などの腫瘍の増殖と浸潤につながる可能性がある KEGG の Proteoglycans in cancer パスウェイ上の分子が含まれていた。リン酸化プロテオームについても、同様の分散分析を行い、4 つのサブタイプのいずれかの間で有意に変化する 12 のリン酸化サイトを同定した。これらのリン酸化サイトの階層的クラスタリングは、3 つの主要なクラスターを示し、そのうちの1つのクラスターは、1 つのサブタイプ (Basal-like) にのみ特異的な (より高発現の) リン酸化サイトを含んでいた (図 1.14B)。

臨床的サブタイプと同様に、PAM50 サブタイプと相関する分子群を同定するためにプロテオームとリン酸化プロテオームについて Welch の t 検定を行い、特定のサブタイプとその他のサブタイプ間で有意に変化する 650 タンパク質グループおよび 596 リン酸化サイトを同定した (図 1.15、1.16)。これらのタンパク質グループおよびリン酸化サイトの階層的クラスタリングは、図 1.13 のすべてのタンパク質グループおよびリン酸化サイトによるクラスタリングと比べて、PAM50 サブタイプに相関したクラスターが得られ、PAM50 サブタイプと相関するマーカー候補タンパク質が同定できたと考えられる (図 1.17)。

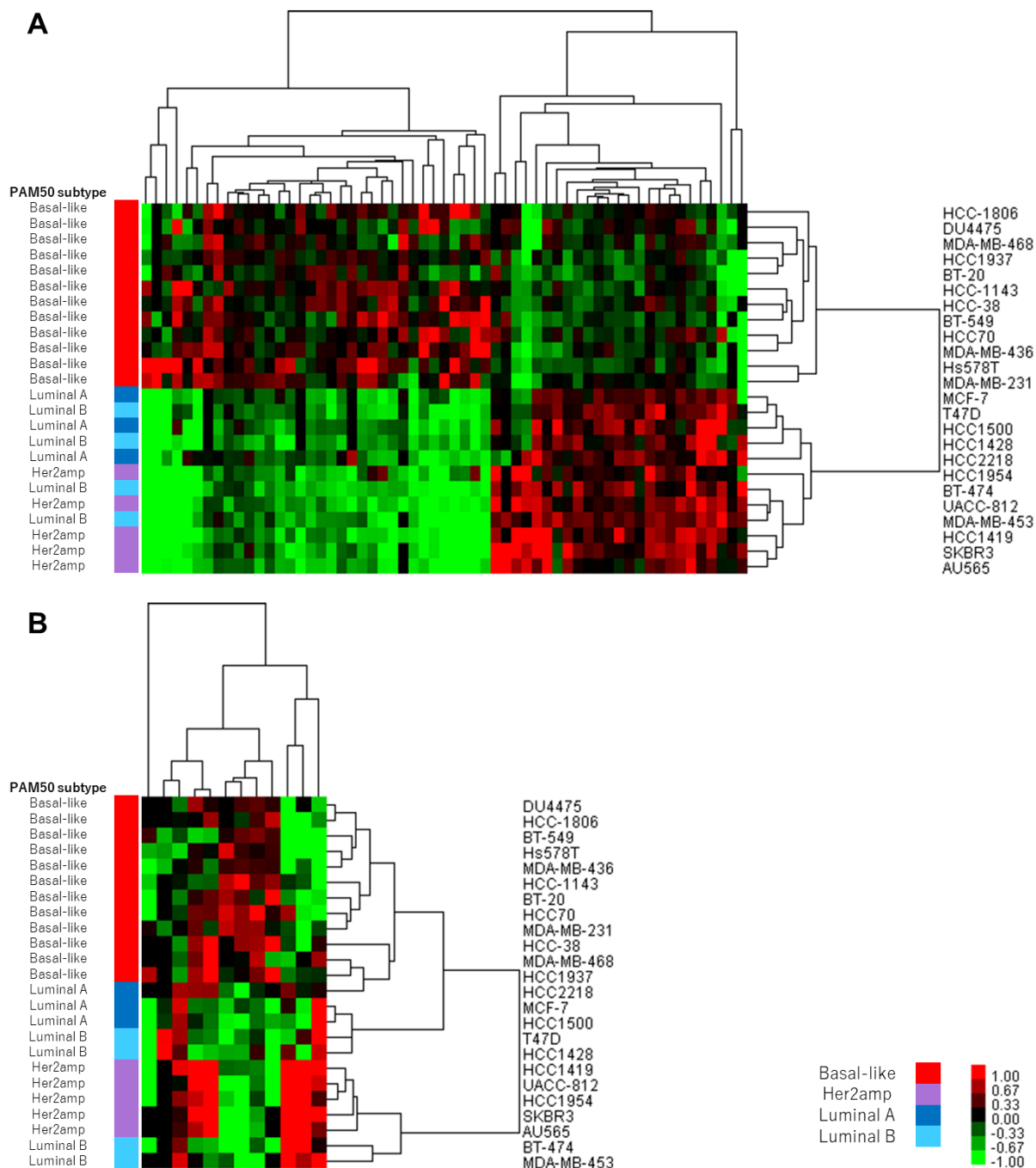


図 1.14 サブタイプ間で発現が異なるタンパク質およびリン酸化サイトの階層的クラスタリング

4 つのがんサブタイプ間で発現が異なるタンパク質を、ANOVA 分析 (FDR < 0.05) を用いて同定した。タンパク質はその発現レベルに応じて色分けされている。サブタイプ特異的なクラスターは、樹形図上で区別することができる。A) プロテオームプロファイルから同定されたサブタイプ間で発現が異なるタンパク質のクラスタリング。B) リン酸化プロテオームプロファイルから同定されたサブタイプ間で発現が異なるリン酸化サイトのクラスタリング。

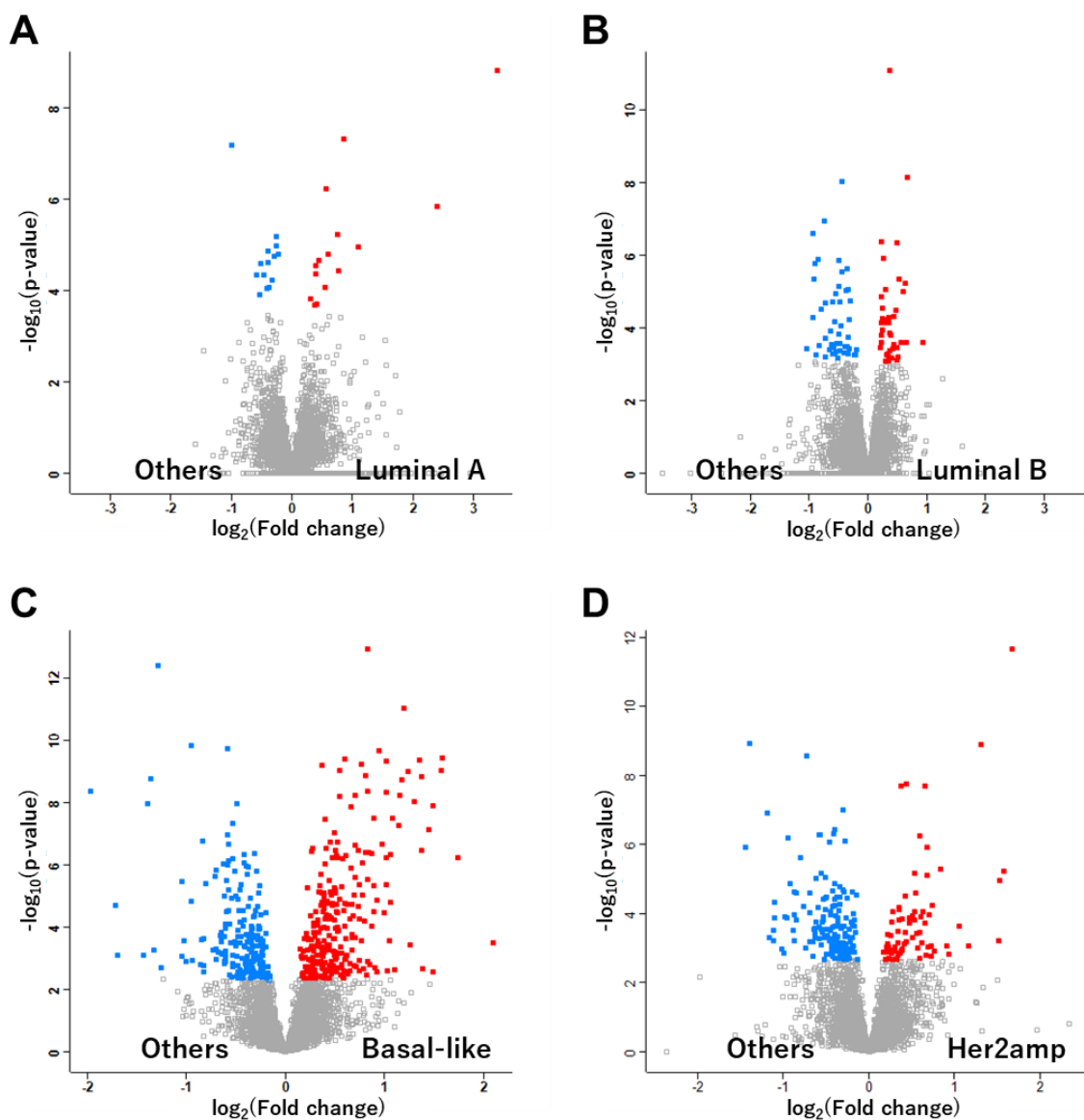


図 1.15 サブタイプ間で発現が異なるタンパク質グループの同定

1 種類のサブタイプとその他のサブタイプの間で発現が異なるタンパク質について、Welch の t 検定由来の $-\log_{10}(\text{p-value})$ (y 軸) に対する $\log_2(\text{Fold change})$ (x 軸) を表示したボルケーノプロット (Benjamini-Hochberg (BH) FDR < 0.05)。有意に減少したタンパク質は青、有意に増加したタンパク質は赤で示す。A) Luminal A、B) Luminal B、C) Basal-like、D) Her2amp

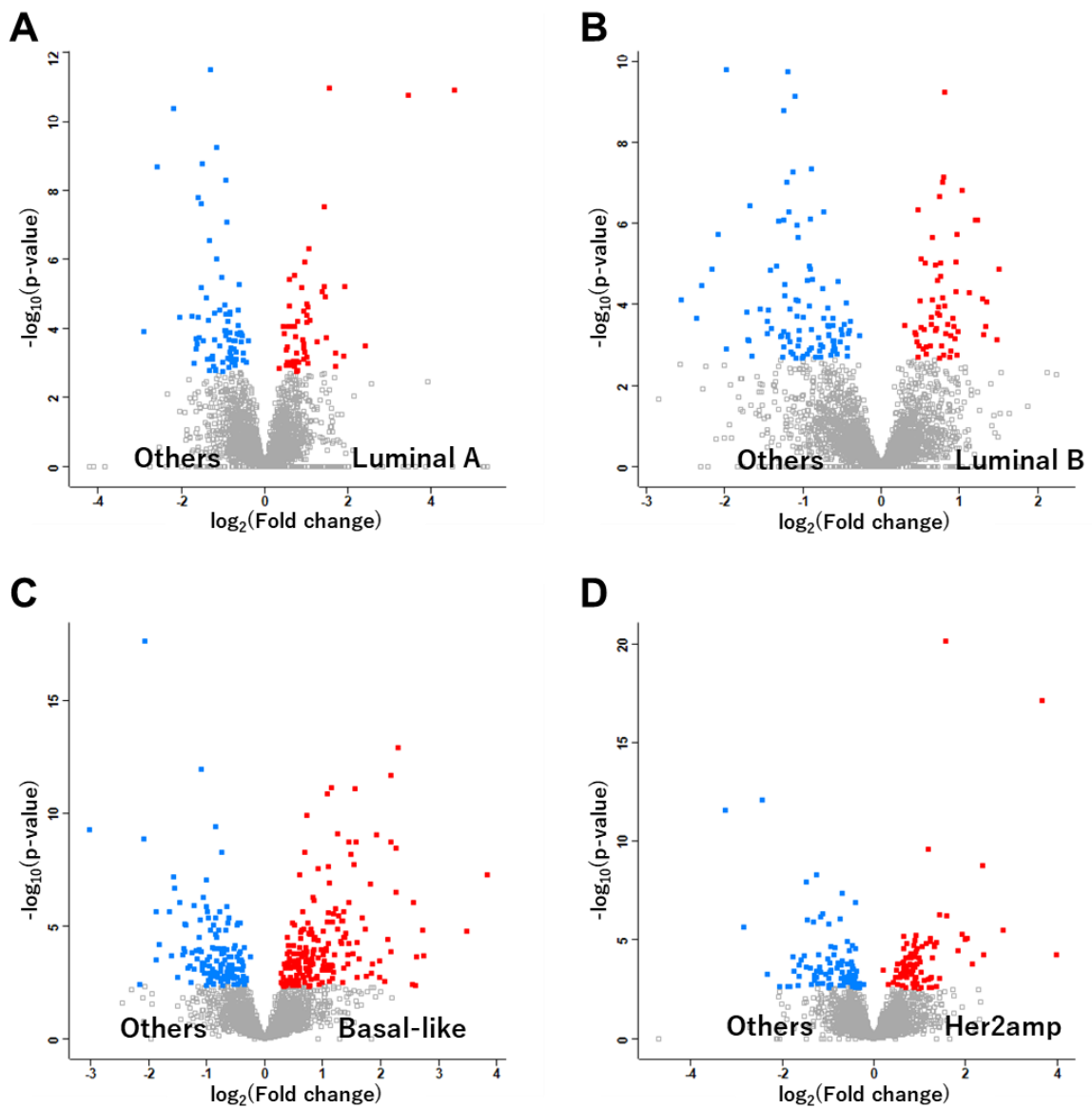


図 1.16 サブタイプ間で発現が異なるリン酸化サイトの同定

1 種類のサブタイプとその他のサブタイプの間で発現が異なるリン酸化サイトについて、Welch の t 検定由来の $-\log_{10}(\text{p-value})$ (y 軸) に対する $\log_2(\text{Fold change})$ (x 軸) を表示したボルケーノプロット (Benjamini-Hochberg (BH) FDR < 0.05)。有意に減少したタンパク質は青、有意に増加したタンパク質は赤で示す。A) Luminal A、B) Luminal B、C) Basal-like、D) Her2amp

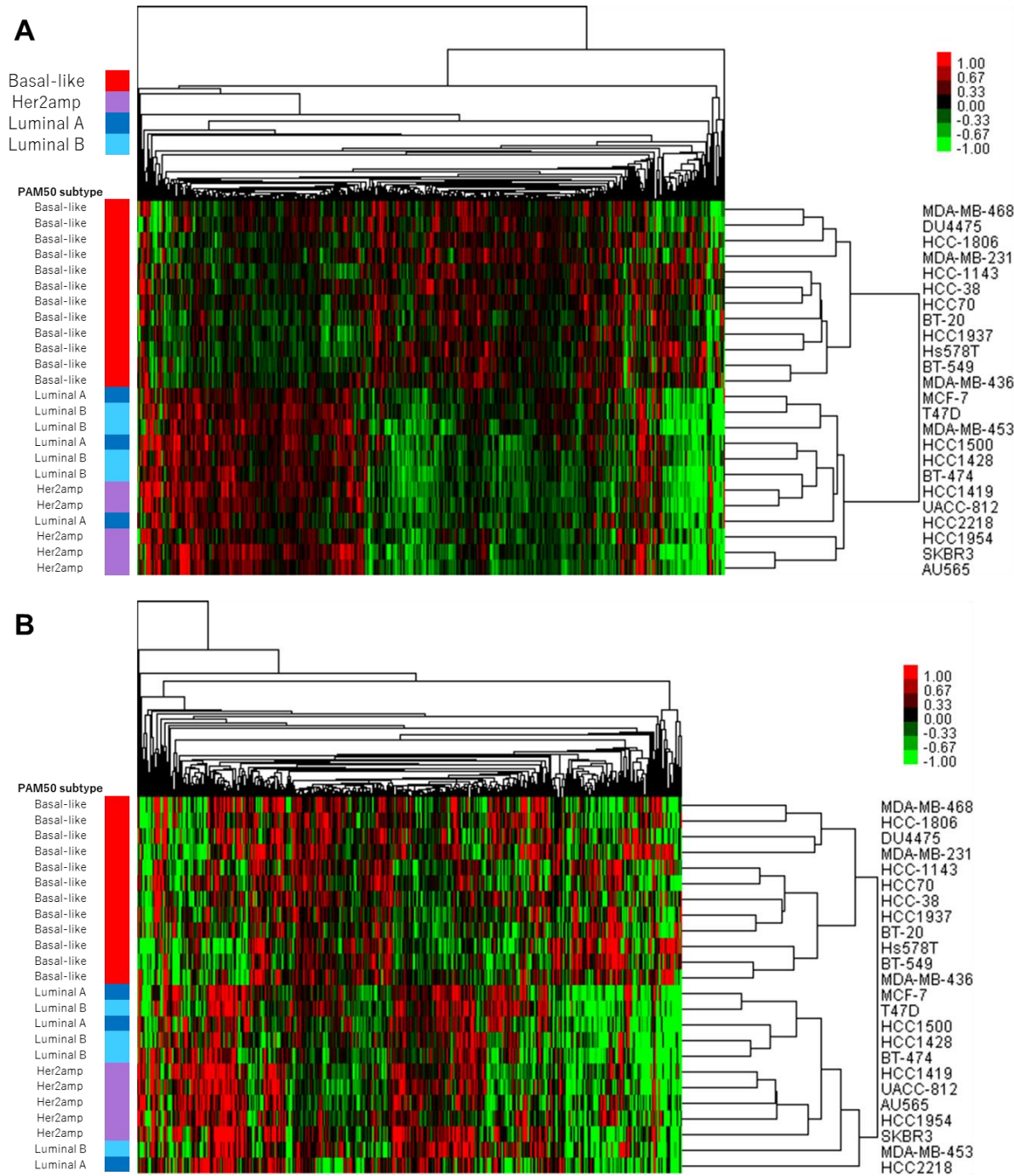


図 1.17 サブタイプ間で発現が異なるタンパク質およびリン酸化サイトの階層的クラスタリング

1 種類のサブタイプとその他のサブタイプの間で発現が異なるタンパク質グループおよびリン酸化サイトを Welch の t 検定 (Benjamini-Hochberg (BH) FDR < 0.05) を用いて同定した。A) プロテオームプロファイルから同定された特定のサブタイプで発現が異なるタンパク質のクラスタリング。B) リン酸化プロテオームプロファイルから同定された特定のサブタイプで発現が異なるリン酸化サイトのクラスタリング。

第3節および本節で同定した ER/PR/Her2 の発現と関連するタンパク質グループおよびリン酸化サイトと、PAM50 サブタイプと関連するタンパク質グループおよびリン酸化サイトには、30 タンパク質、7 リン酸化サイトの重複があった (表 1.5、1.6)。この 30 タンパク質、7 リン酸化サイトの内、PAM50 パネルに含まれているのは ERBB2 のみであった。また、アメリカ食品医薬品局 (Food and Drug Administration, FDA) が承認している薬剤のターゲットタンパク質として、ANXA1、ERBB2、CD44、ITGT1 が含まれていた。他にも、UniProt データベースで、がん、神経疾患、全身疾患、心血管疾患など、様々な疾患状態に関与していることが実験的に証明されている遺伝子のうち、既知の薬物標的タンパク質クラス (酵素、輸送体、受容体、イオンチャネル) に属しており、ドラッグバンクのデータベース⁴⁷では FDA 承認薬や実験薬の標的にはなっておらず、薬物標的の可能性があると考えられているタンパク質として、PLOC1、CAV1、MVD、AGPS、SLC16A1 が含まれていた。

表 1.5 ER/PR/Her2 の発現および PAM50 サブタイプと関連するタンパク質グループ

*は、FDA が承認している薬剤のターゲットタンパク質。**は、UniProt データベースで、がん、神経疾患、全身疾患、心血管疾患など、様々な疾患状態に関与していることが実験的に証明されている遺伝子のうち、既知の薬物標的タンパク質クラス(酵素、輸送体、受容体、イオンチャネル)に属しており、ドラッグバンクのデータベースでは FDA 承認薬や実験薬の標的にはなっていないタンパク質。

Gene Names	Protein names
GPC1	Glypican-1;Secreted glypican-1
SPTBN1	Spectrin beta chain, non-erythrocytic 1
EPB41L2	Band 4.1-like protein 2
PDLIM7	PDZ and LIM domain protein 7
TMEM205	Transmembrane protein 205
ANPEP	Aminopeptidase N
PLOD1**	Procollagen-lysine,2-oxoglutarate 5-dioxygenase 1
ANXA1*	Annexin A1
GBP1	Interferon-induced guanylate-binding protein 1
ARFGEF1	Brefeldin A-inhibited guanine nucleotide-exchange protein 1
FSCN1	Fascin
CAVIN1	Caveolae-associated protein 1
MAP1B	Microtubule-associated protein 1B;MAP1B heavy chain;MAP1 light chain LC1
ERBB2*	Receptor tyrosine-protein kinase erbB-2
CAV1**	Caveolin-1
HLA-A	HLA class I histocompatibility antigen, A-3 alpha chain
IGF2BP3	Insulin-like growth factor 2 mRNA-binding protein 3
FABP5	Fatty acid-binding protein, epidermal
MRPL54	39S ribosomal protein L54, mitochondrial
GLS	Glutaminase kidney isoform, mitochondrial
MVD**	Diphosphomevalonate decarboxylase
PPP1R18	Phostensin
CD44*	CD44 antigen
MIEN1	Migration and invasion enhancer 1
PSMG3	Proteasome assembly chaperone 3
CALD1	Caldesmon
ORC5	Origin recognition complex subunit 5
ITGB1*	Integrin beta-1
LGALS1	Galectin-1
AGPS**	Alkyldihydroxyacetonephosphate synthase, peroxisomal

表 1.6 ER/PR/Her2 の発現および PAM50 サブタイプと関連するリン酸化サイト

*は、FDA が承認している薬剤のターゲットタンパク質。**は、UniProt データベースで、がん、神経疾患、全身疾患、心血管疾患など、様々な疾患状態に関与していることが実験的に証明されている遺伝子のうち、既知の薬物標的タンパク質クラス(酵素、輸送体、受容体、イオンチャネル)に属しており、ドラッグバンクのデータベースでは FDA 承認薬や実験薬の標的にはなっていないタンパク質。

Gene Names	Protein names	Phosphosites
SLC16A1**	Monocarboxylate transporter 1	S467
CAVIN1	Caveolae-associated protein 1	S167
MYO18A	Unconventional myosin-XVIIIa	S140
PML	Protein PML	S403
LMNA	Prelamin-A/C;Lamin-A/C	S652
S100A16	Protein S100-A16	S2
ERBB2*	Receptor tyrosine-protein kinase erbB-2	S1054

第5節 乳がん細胞株の腫瘍モデルとしての評価

ヒトがん由来の細胞株は、がんを理解するためのモデルとして、抗悪性腫瘍薬の開発の諸段階において広く利用されている。細胞株と腫瘍の間には多くの類似点が見出され、がん細胞株を使用して、疾患における制御が緩和された遺伝子やシグナル伝達経路に関する情報が多く報告されている。しかし、最初のがん細胞株が開発されて以来、これらのモデルの臨床的妥当性は、絶えず議論されてきた。元の腫瘍とそこから派生したがん細胞株との間には、完全ではないがゲノム類似性があるとされ、元の腫瘍と比較した場合、がん細胞株は抗がん剤に対して同様の反応を示すことが報告されている。これまで、これらの細胞株の臨床的妥当性の検討は遺伝子発現の比較によって行われることが多く、タンパク質発現を比較した報告は少ない。本節では、2016年にCPTACによって報告されたヒト乳がん組織 77 試料のプロテオームプロファイルおよび105 試料のリン酸化プロテオームプロファイル¹³と、本研究で取得した 24 乳がん細胞株のプロテオームおよびリン酸化プロテオームプロファイルの比較を行い、乳がん細胞株の腫瘍モデルとしての妥当性を評価した。

乳がん組織試料と細胞株の間で 5,050 タンパク質、2,467 リン酸化サイトが共通して同定、定量された。これらの定量値を Quantile 正規化したのち、乳がん組織試料-細胞株間のプロテオーム、リン酸化プロテオームプロファイルのピアソン相関係数を算出した (図 1.18A、B)。相関係数はおおよそプロテオームで-0.1 から 0.2、リン酸化プロテオームで-0.2 から 0.4 までの範囲に分布した。プロテオームとリン酸化プロテオームの発現相関解析は、4 つのサブタイプの細胞株すべてが腫瘍サンプルと異なる相関分布を有することを示唆している。リン酸化プロテオームの相関において、MCF7 と T47D は、乳がん腫瘍と比較的高い相関を示した。Jiang らによって報告された逆相タンパク質アレイによるがんに関連する 50 のタンパク質およびリン酸化タンパク質の発現における細胞株と腫瘍の相関係数は、-0.61 から 0.84 までの範囲であり、本研究のプロテオームおよびリン酸化プロテオームより高い相関係数を示していた³¹。その中でも、T47D、AU565 は、タンパク質レベルで腫瘍に近い相関を有する上位 3 つの細胞株であったが、プロテオームの相関においてはその傾向は見られなかった。これは、プロテオームおよびリン酸化プロテオームでは、がんに関連する以外にも多くの分子が発現しているために低い相関係数を示したと考えられる。一方で、乳がん腫瘍と乳がん細胞株との間の遺伝子発現プロファイルの類似性を比較するために、Gene Expression Omnibus (GEO) GSE41998⁴⁸の乳がん腫瘍 (279 腫瘍) と本研究で使用した 24 細胞

株を含む GEO GSE50811⁴⁹ を使用した。図 1.18C は、乳がん細胞株と原発腫瘍との間の遺伝子発現の相関係数の分布を示し、細胞株と腫瘍の間で 0.7~0.8 程度であった。これらの結果から、細胞株は乳がんのサブタイプが異なっても、遺伝子発現プロファイルにおいて腫瘍と高い類似性を維持していることがわかった。

プロテオーム、リン酸化プロテオームプロファイルの階層的クラスタリング分析を行った結果、細胞株と組織は混じりあうことなくクラスタリングされた (図 1.19)。さらに、乳がん腫瘍のプロテオームおよびリン酸化プロテオームプロファイルから PAM50 分類のためのプロテオームおよびリン酸化プロテオームマーカーとして、2,534 タンパク質、および 2,086 リン酸化サイトを同定した (図 1.20、1.21)。これらのマーカーと第 4 節で乳がん細胞株における PAM50 分類のためのプロテオームおよびリン酸化プロテオームマーカーとして同定された、650 タンパク質グループおよび 596 リン酸化サイトとの重複を確認したところ、185 タンパク質と 45 リン酸化サイトが重複していた。これらの直接乳がんに関係していると考えられる分子のみを用いた階層的クラスタリングを行った (図 1.22)。その結果、図 1.19 と比較して、PAM50 サブタイプによるクラスターが形成され、その中に細胞株試料と腫瘍組織試料がそれぞれクラスターを形成していた。このようにがん関連分子にフォーカスすることで、細胞株と腫瘍組織の類似した特徴を一部抽出することは可能だが、細胞株と腫瘍組織がまとまって分類されることはなく、乳がん細胞株は腫瘍組織とは異なるプロテオームおよびリン酸化プロテオームプロファイルを有していることがわかった。

一方で、乳がんの臨床バイオマーカー (ERPR、ERBB2、MKI67) および既知の乳がんに関連する分子⁵⁰の腫瘍と細胞株の間でのタンパク質とそのリン酸化体の発現変動を示す (図 1.23)。腫瘍と細胞株の両方で定量された PGR、ERBB2、MKI67、CD44、KRT18、GATA3、FOXA1、TP53 は、腫瘍と細胞株の間でタンパク質とそのリン酸化体の発現変動に高い類似性を示した。これらの結果から、乳がん細胞株と乳がん腫瘍組織は、異なるプロテオームプロファイル、リン酸化プロテオームプロファイルを示すが、内因性サブタイプに関連する分子群や、増殖・細胞周期関連因子、ER シグナル関連因子などの乳がんの特徴的な分子の発現は保存されていることがわかった。

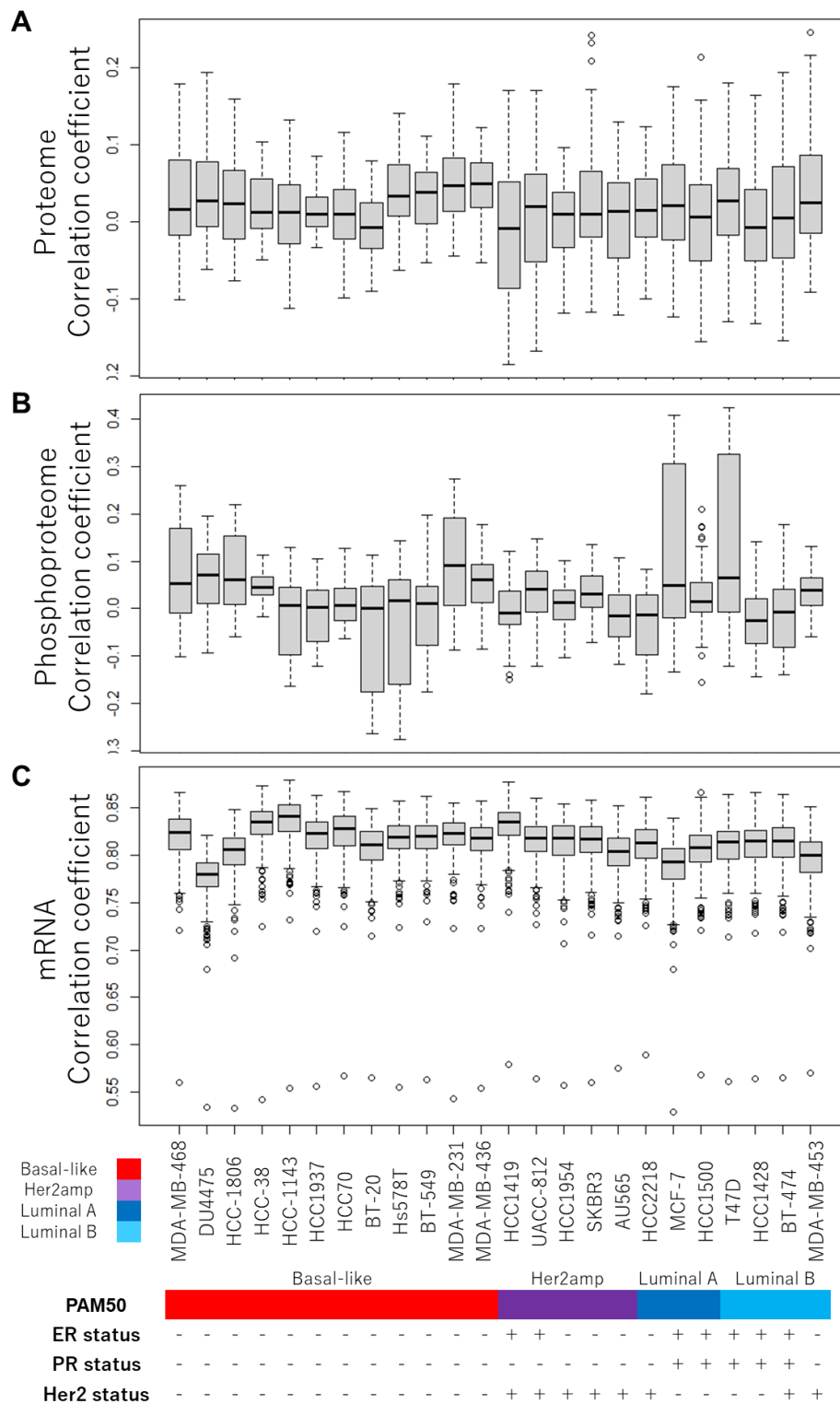


図 1.18 乳がん細胞株と腫瘍のプロテオーム、リン酸化プロテオーム、mRNA 発現の比較

A-C) 乳がん細胞株と腫瘍との間の発現相関解析。X 軸は 24 の細胞株、Y 軸は腫瘍と細胞とのピアソン相関係数を示す。A) プロテオーム発現の相関係数、B) リン酸化プロテオーム発現の相関係数、C) mRNA 発現の相関係数

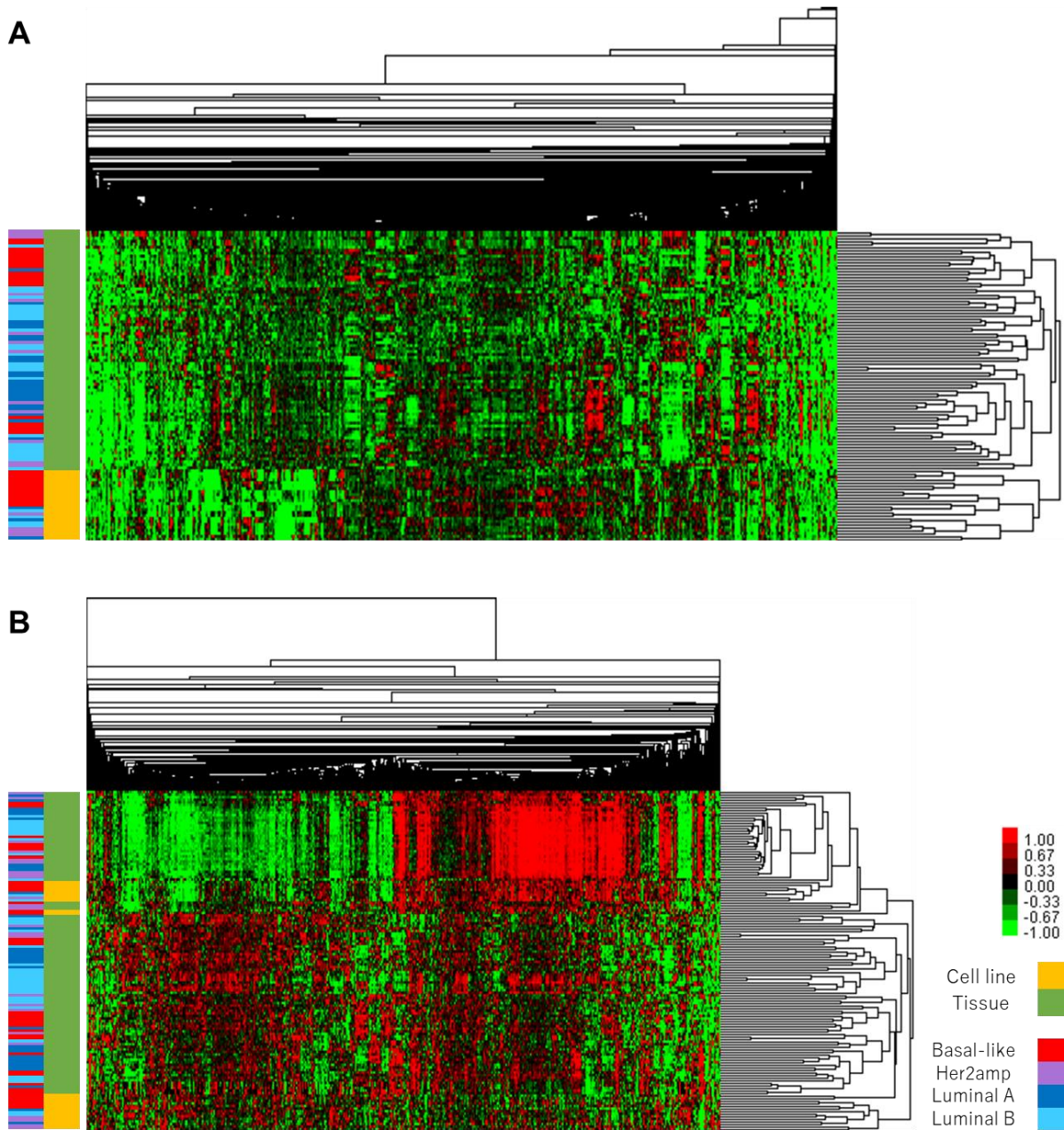


図 1.19 乳がん細胞株と腫瘍のクラスタリング

A) 乳がん細胞株と腫瘍のプロテオームに基づくクラスタリング。B) 乳がん細胞株と腫瘍のリン酸化プロテオームに基づくクラスタリング。

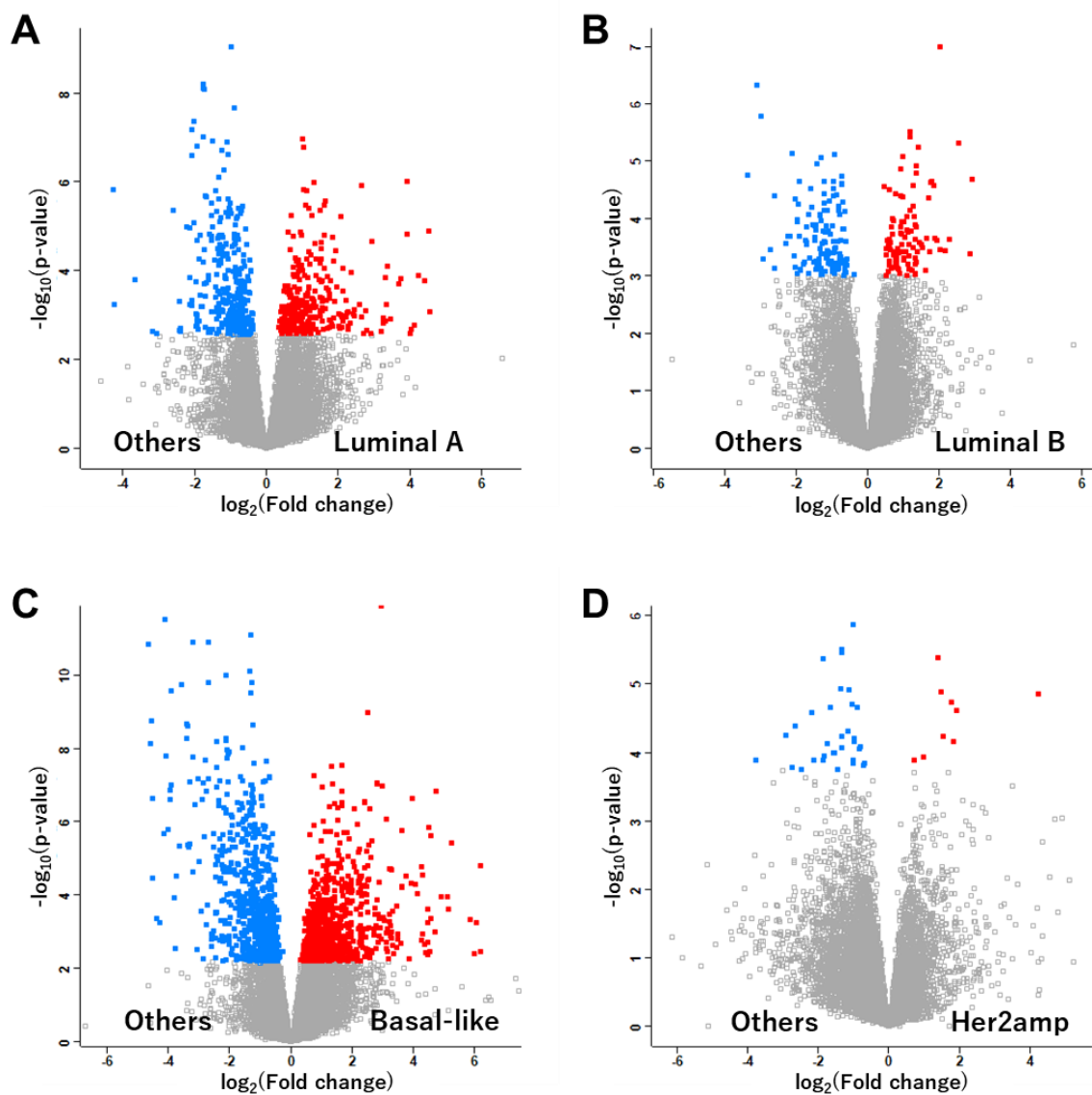


図 1.20 サブタイプ間で発現が異なるタンパク質の同定 (乳がん腫瘍)

1 種類のサブタイプとその他のサブタイプの間で発現が異なるタンパク質について、Welch の t 検定由来の $-\log_{10}(\text{p-value})$ (y 軸) に対する $\log_2(\text{Fold change})$ (x 軸) を表示したボルケーノプロット (Benjamini-Hochberg (BH) FDR < 0.05)。有意に減少したタンパク質は青、有意に増加したタンパク質は赤で示す。A) Luminal A、B) Luminal B、C) Basal-like、D) Her2amp

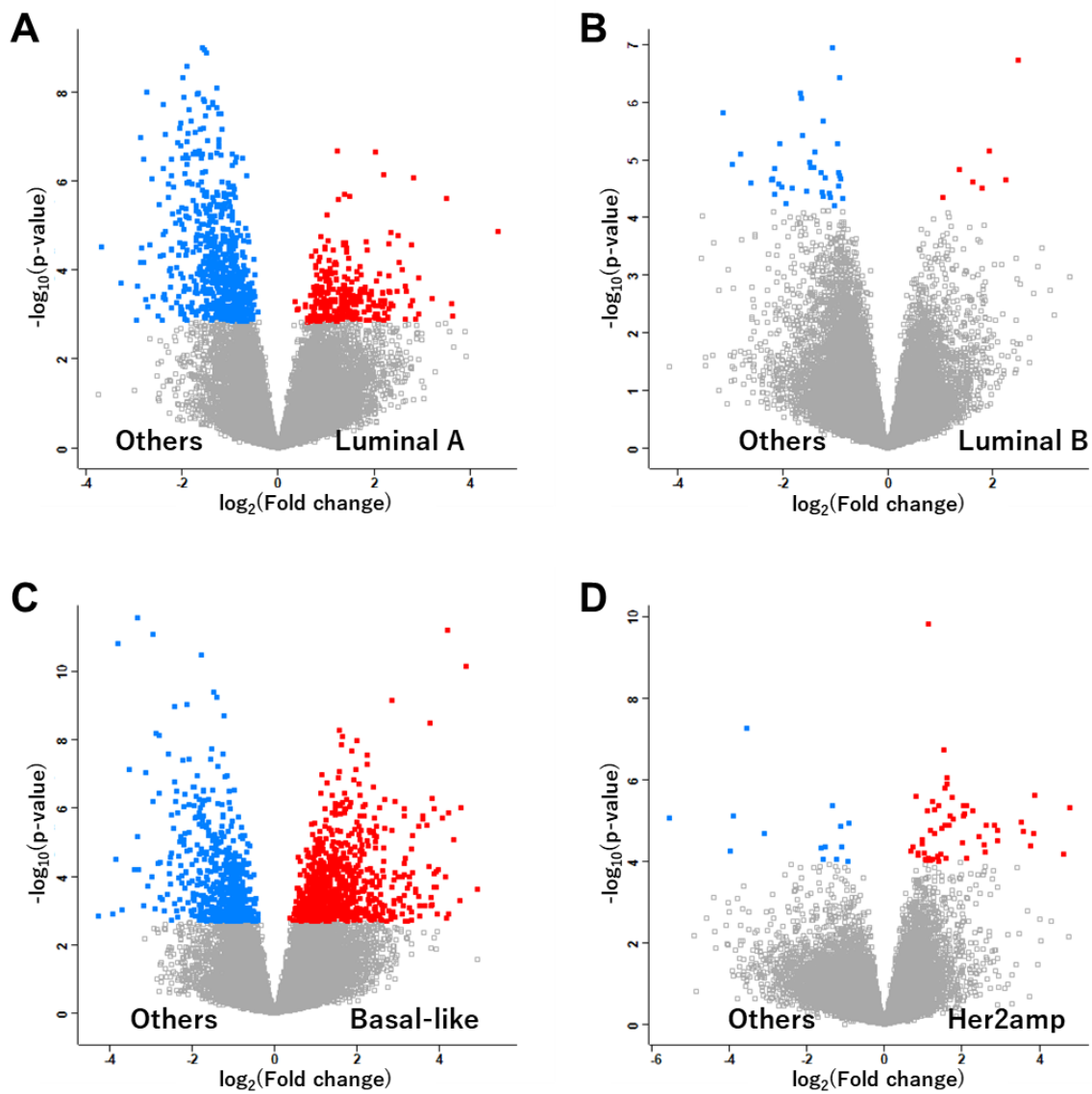


図 1.21 サブタイプ間で発現が異なるリン酸化サイトの同定 (乳がん腫瘍)

1 種類のサブタイプとその他のサブタイプの間で発現が異なるリン酸化サイトについて、Welch の t 検定由来の $-\log_{10}(\text{p-value})$ (y 軸)に対する $\log_2(\text{Fold change})$ (x 軸)を表示したボルケーノプロット (Benjamini-Hochberg (BH) $\text{FDR} < 0.05$)。有意に減少したタンパク質は青、有意に増加したタンパク質は赤で示す。A) Luminal A、B) Luminal A、C) Basal-like、D) Her2amp

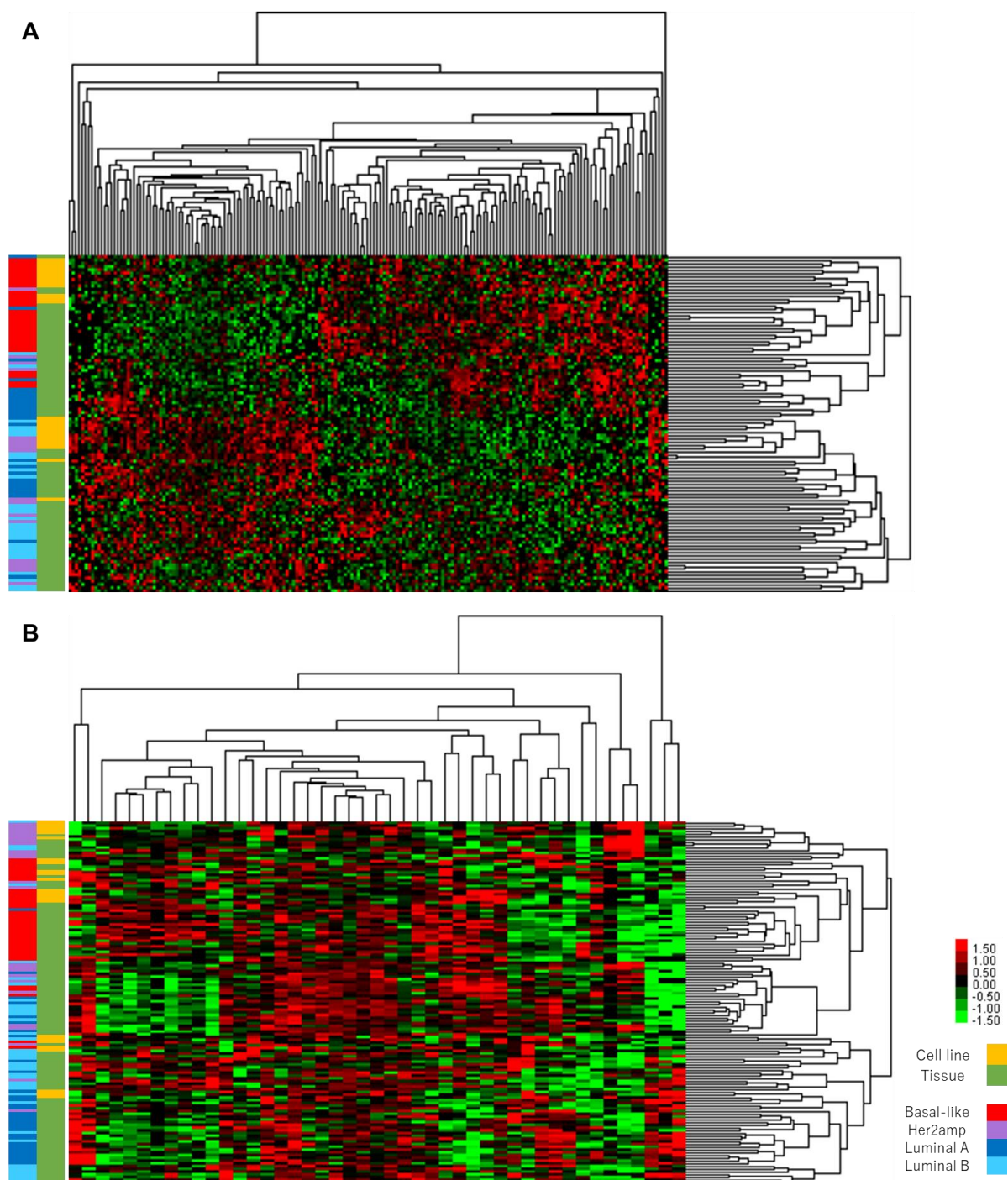


図 1.22 PAM50 分類のためのマーカーに基づく乳がん細胞株と腫瘍のクラスタリング

A) 乳がん細胞株と腫瘍で PAM50 分類のためのマーカーとして共通同定されたの 185 タンパク質に基づくクラスタリング。B) 乳がん細胞株と腫瘍で PAM50 分類のためのマーカーとして共通同定された 45 リン酸化サイトにに基づくクラスタリング。

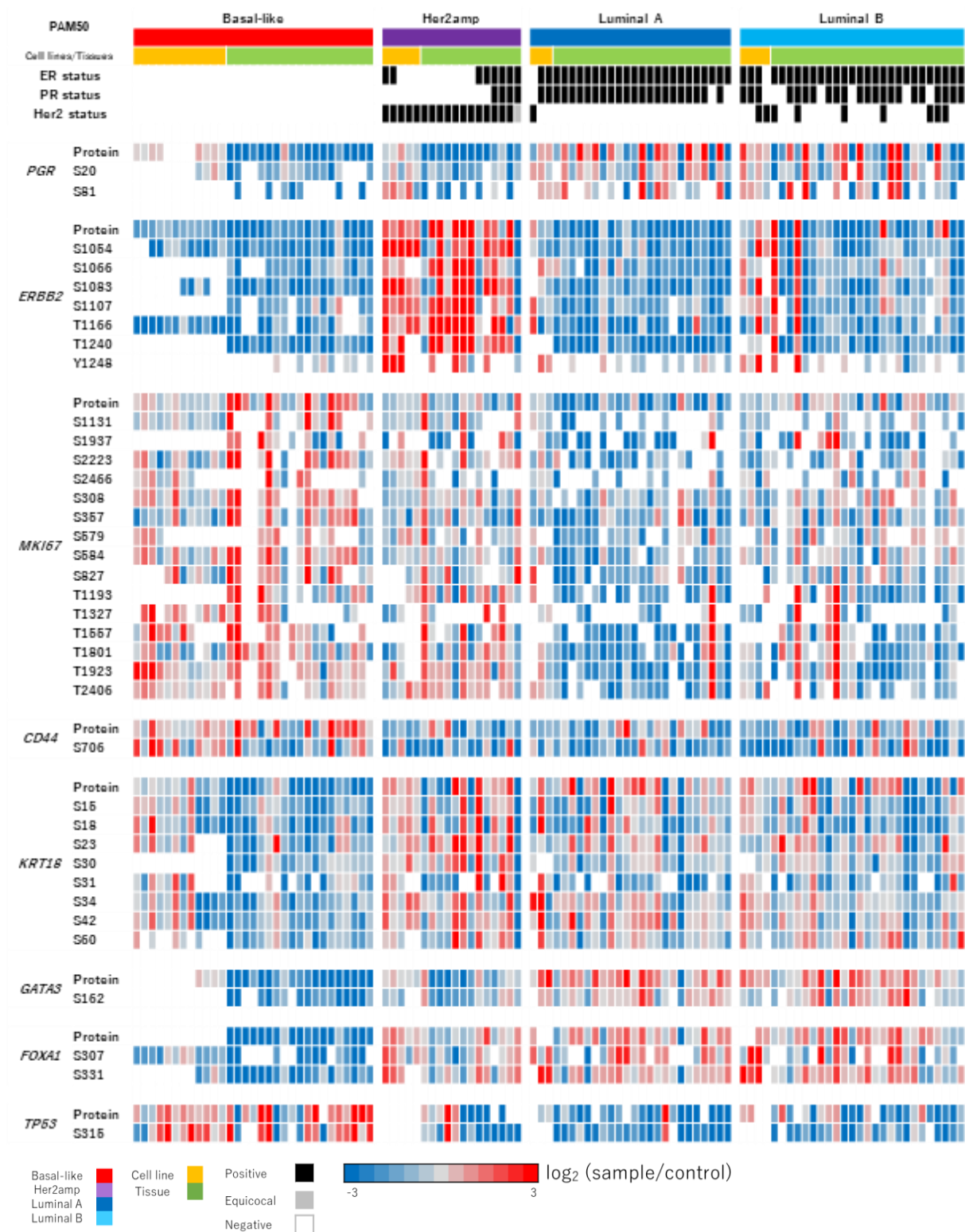


図 1.23 乳がん細胞株および腫瘍にわたる乳がん関連遺伝子のタンパク質発現のヒートマップ
PAM50 サブタイプ、ER、PR、Her2 の状態がアノテーションされている。各遺伝子について、タンパク質およびリン酸化サイトの発現を示す。

第6節 おわりに

本章では、24 種の乳がん細胞株のプロテオームおよびリン酸化プロテオームプロファイルを取得した。取得した 5,177 のタンパク質グループと 3,508 のリン酸化サイトの定量値を用いて、従来の臨床モデルや PAM50 分類のためのプロテオームおよびリン酸化プロテオームマーカーを同定した。これらのマーカー分子には既知の PAM50 パネル分子以外の新規のマーカー分子が多く含まれていた。

ヒト乳がん組織 77 試料のプロテオームプロファイルおよび 105 試料のリン酸化プロテオームプロファイルを既報から収集して、細胞株のプロファイルと発現の相関解析を行ったところ、mRNA の発現プロファイルは細胞株と腫瘍組織の間で高い相関を示した。一方で、プロテオームプロファイルおよびリン酸化プロテオームプロファイルは、低い相関を示した。また、階層的クラスタリングを行ったところ、細胞株と組織は混じりあうことなくクラスタリングされた。したがって、乳がん細胞株は腫瘍組織とは異なるプロテオームおよびリン酸化プロテオームプロファイルを有していることがわかった。さらに、PAM50 分類のためのプロテオームおよびリン酸化プロテオームマーカーを同定し、これらの直接乳がんに関係していると考えられる分子のみを用いた階層的クラスタリングを行った。その結果、細胞株試料と腫瘍組織試料は、PAM50 分類によるサブタイプからなるクラスターの中で分かれて分類された。また、細胞株と腫瘍において、PGR、ERBB2、MKI67、CD44、KRT18、GATA3、FOXA1、TP53 などの既知の乳がんの臨床マーカーや関連分子は、タンパク質とそのリン酸化体の発現様式に高い類似性を示した。この結果、乳がん腫瘍と細胞株の間には分子的特徴の類似性と非類似性が存在することがわかった。細胞株は、原発腫瘍の分子特性の一部を反映しているにすぎず、すべての分子特性を反映しているわけではないため、細胞株は腫瘍モデルとして最適であるとは言い難く、細胞株を用いた実験で得られた結果については、腫瘍組織を用いた検証がより重要であると考えられる。

第2章

細胞内キナーゼ活性計測のためのリン酸化ペプチドの 選択反応モニタリングの開発

第1節 はじめに

タンパク質の可逆的なリン酸化は、プロテインキナーゼによって媒介される最も普遍的な翻訳後修飾の一つであり、様々な生物学的機能を制御している。がんや動脈硬化、免疫疾患などの多くの疾患は、リン酸化に伴うシグナル異常が原因であると考えられている^{51,52}。そのため、キナーゼ全体の活性プロファイリング、すなわちキノームプロファイリングは、疾患診断や分子標的薬を用いた個別化治療ならびにプレシジョン・メディシンに重要な役割を果たしている。キナーゼ活性を測定するための最も直接的な手法は、リン酸化された基質を定量することである。現在の液体クロマトグラフィー/タンデム質量分析 (LC/MS/MS) 技術では、リン酸化ペプチドを特異的に濃縮することで、複雑な生物サンプルから数千個のリン酸化ペプチドを同定することが可能である^{53,54}。しかし、多くの場合、キナーゼと基質の関係についての情報は限られており、従来のリン酸化プロテオミクスではキナーゼ活性を測定することは困難である。

キナーゼ活性は触媒ドメイン上のリン酸化によって制御されていることが知られており^{55,56}、リン酸化部位特異的抗体を用いたウェスタンブロットによりキナーゼ活性の変化が調べられている。しかし、この方法では、結果が抗体の品質に依存することや、リン酸化されたタンパク質の絶対量を測定することが困難であることから、キノームワイドな定量には適していなかった。一方、ヒトプロテオームのプロファイリングには、LC/MS/MS を用いた定量法が用いられている^{57,58,59,60,61}。特に、安定同位体を用いた選択反応モニタリング (SRM) 法は、複雑な試料中のタンパク質を正確に定量するターゲットプロテオミクスアプローチとして活用されている^{62,63,64}。具体的には、GRB2アダプターシグナル伝達⁶⁵、インスリンシグナリング⁶⁶、キナーゼ発現^{67,68}、キナーゼ基質^{69,70,71}、そして、ERK⁷²、EGFR⁷³、Akt⁷⁴などの単一標的キナーゼ上のリン酸化サイトの測定に用いられている。しかし、これまでのところ、キナーゼの活性をパスウェイやキノームスケールでプロファイリングした報告はなかった。

本研究では、EGFR-MAPK シグナル伝達経路に関与する 10 種類のキナーゼを標的としたキャピラリーLC/MS/MS による SRM アッセイプラットフォームを開発した。ターゲットとするリン酸化ペプチドは、これまで得られた実験データや公開されているデータ

ベースをもとに、リン酸化サイトの位置と質量分析計での検出性を考慮して選択した。次に、SRM アッセイ法の精度、真度、直線性、検出限界 (LOD)、定量限界 (LOQ) を検証した。最後に、本アッセイ法を過バナジン酸処理した HeLa 細胞に適用し、HeLa 細胞内の 10 種類のキナーゼの活性を測定した (図 2.1)。

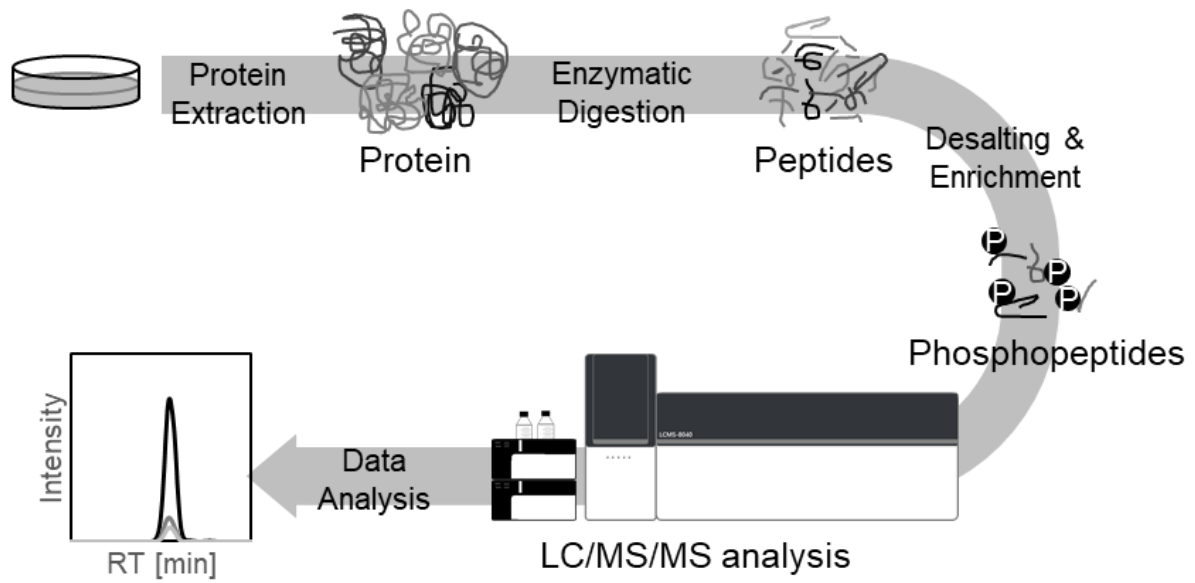


図 2.1 SRM 法によるキナーゼ活性測定の仕事フロー

HeLa 細胞からタンパク質を抽出し、タンパク質を消化酵素によりペプチドに断片化した後に、濃縮したリン酸化ペプチドをキャピラリーLC/MS/MS を用いて SRM モードで分析する。

第2節 標的リン酸化ペプチドの選択

図 2.2 に示す EGFR-MAPK パスウェイに含まれる 10 種類のキナーゼを標的とした。10 種類のキナーゼに対応するリン酸化ペプチド配列を研究室内で有するリン酸化プロテオームライブラリー (64 種類のがん細胞株由来の 46,000 リン酸化ペプチドを含む) と *in vivo* でのキナーゼと基質の関係やキナーゼの活性調節部位情報が利用可能な PhosphoSite-Plus⁷⁵ のリン酸化サイト情報を基に、質量分析計での検出性を考慮して選定した (表 2.1)。検出性については、(1) 標的ペプチドの長さが 7 残基以上 20 残基以下であること、(2) 消化酵素による未切断部位がないこと、(3) 定量対象以外のリン酸化修飾部位が存在しないこと、を基準にして選択した。ただし、基準 (2) については、研究室内の実験データでは、消化酵素による未切断部位を持たない BRAF ペプチドは同定されていなかったため、未切断部位を 1 つ含む BRAF ペプチドを選択した。基準 (3) については、GSK3A、ERK2、PAK2、EGFR、BRAF は、PhosphoSite-Plus では配列上に他のリン酸基が報告されているが、表 2.1 のモノリン酸化ペプチドを選択した。これは、これらの研究室内のリン酸化プロテオームライブラリーにおいて、モノリン酸化ペプチド以外がほとんど同定されなかったためである。プロテインキナーゼの活性は、一般に活性化ループ (t-loop) のリン酸化によって制御されていることが知られているが⁵⁶、他の領域に位置するリン酸化物がキナーゼ活性を制御している例もある。本研究では、4 つのペプチドを t-loop 上の活性調節部位から、5 つのペプチドを t-loop 外の活性調節部位から選択した (表 2.1)。また、JNK1 の活性を測定するための基質ペプチドを 1 種類選んだ。

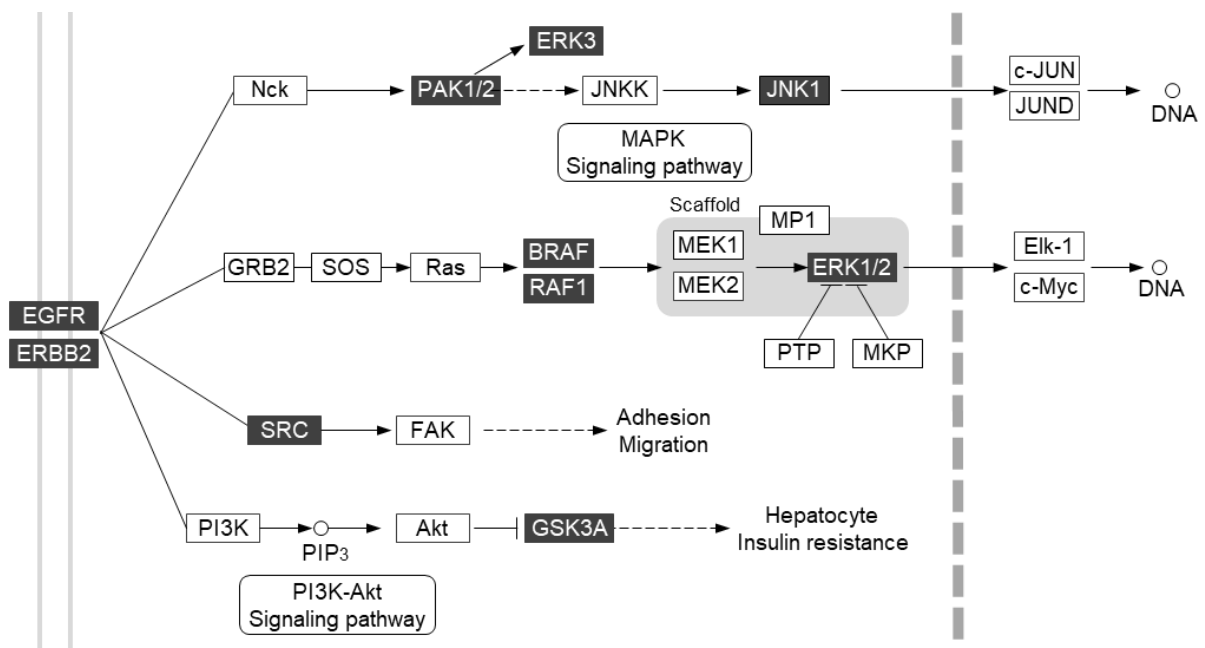


図 2.2 EGFR-MAPK シグナル伝達経路

KEGG ERBB シグナル伝達経路 (hsa04012) と MAPK シグナル伝達経路 (hsa04010) より作成。
 ERK3 は文献情報に応じて追加している。本研究で標的としたキナーゼを黒地で示している。

表 2.1 本研究における標的キナーゼとリン酸化ペプチド

No.	Target kinase	Target phosphopeptide	Phosphosite position in protein	Location	Isotope label	Q1		Q3		RT	Collision Energy
1	SRC	LIEDNEpYTAR	Y419 (SRC)	t-loop	light	666.30	1190.47	1077.39	833.32	13.7	-24.0
					heavy	669.31	1190.47	1077.39	833.32	13.7	-24.1
2	ERK3	GHLpSEGLVTK	S189 (ERK3)	t-loop	light	588.80	993.57	856.51	743.43	15.7	-21.7
					heavy	594.84	999.60	862.55	749.46	15.7	-21.8
3	JNK1	LApSPELER	S100 (JUND)		light	511.75	881.38	783.40	643.34	16.0	-19.4
					heavy	514.77	881.38	783.40	643.34	16.0	-19.4
4	RAF1	AAHTEDINACILTTpSPR	S642 (RAF1)		light	655.96	653.29	440.71	641.27	16.4	-23.7
					heavy	657.97	659.32	443.73	641.27	16.4	-23.7
5	GSK3A	GEPNVSpYICSR	Y279 (GSK3A)	t-loop	light	695.30	1175.49	964.40	865.33	17.3	-24.9
					heavy	698.31	1175.49	964.40	865.33	17.3	-24.9
6	ERK2	VADPDHDHTGFLpTEpYVATR	T185/Y187 (ERK2)	t-loop	light	778.00	901.40	960.90	1053.90	20.9	-27.3
					heavy	789.00	901.40	960.90	1053.95	20.9	-27.7
7	PAK2	YLpSFTPPEK	S141 (PAK2)		light	609.30	928.51	815.43	599.34	21.3	-22.3
					heavy	615.33	934.55	821.46	605.37	21.3	-22.5
8	EGFR	GSHQISLDNPDpYQQDFFPK	Y1172 (EGFR)		light	791.36	980.48	809.42	681.36	22.7	-27.7
					heavy	795.38	986.51	815.45	687.39	22.7	-27.9
9	BRAF	RDpSSDDWEIPDGQITVGQR	S446 (BRAF)		light	761.34	784.25	1099.37	1070.56	22.8	-26.8
					heavy	763.35	790.28	1105.40	1070.56	22.8	-26.9
10	ERBB2	GTPTAENPePYLGLDVPV	Y1248 (ERBB2)		light	940.44	314.21	215.14	699.33	28.5	-32.2
					heavy	943.46	314.21	215.14	705.36	28.5	-32.3

第3節 SRM 測定系の構築

10 種類のキナーゼの活性を SRM で測定するために、キャピラリー LC システムを使用し、流速 2 $\mu\text{L}/\text{min}$ 、1 サンプルあたりの LC/MS 測定時間を 55 分に設定した。本条件下で、標的ペプチドの SRM パラメータを最適化した(表 2.1)。2+および 3+の両方のプレカーサーイオンをフラグメント化に使用し、プロダクトイオンスキャンの結果に基づいて、b-イオンおよび y-イオンからピーク強度上位 3 つのプロダクトイオンを選択した。衝突エネルギー (CE) は、次式に従って SRM トランジションごとに最適化した⁷⁶。

$$\text{CE} = 0.03 \times (\text{プレカーサーイオンの } m/z) + 4.0$$

本研究の LC 条件では、最大で 4 つのペプチドを同時に定量する必要があり、感度を最大化するために各 SRM トランジションの dwell time (測定対象物質のデータ取り込み時間) を 80 msec に設定した。その結果、最大ループ時間は 1.146 秒と、ピークを定量するためのデータポイントを取得するのに十分な間隔だった (図 2.3)。続いて、合成リン酸化ペプチドを、選択したトランジションを用いて分析し、SRM ピークを確認した(図 2.4)。1 pmol 注入の場合、得られたペプチドあたりのピーク面積は $8.6 \times 10^4 \sim 2.7 \times 10^6$ の範囲であり、感度に大きなばらつきがあることを示しているが、ピーク形状の多様性も感度差の原因となっていると考えられる。

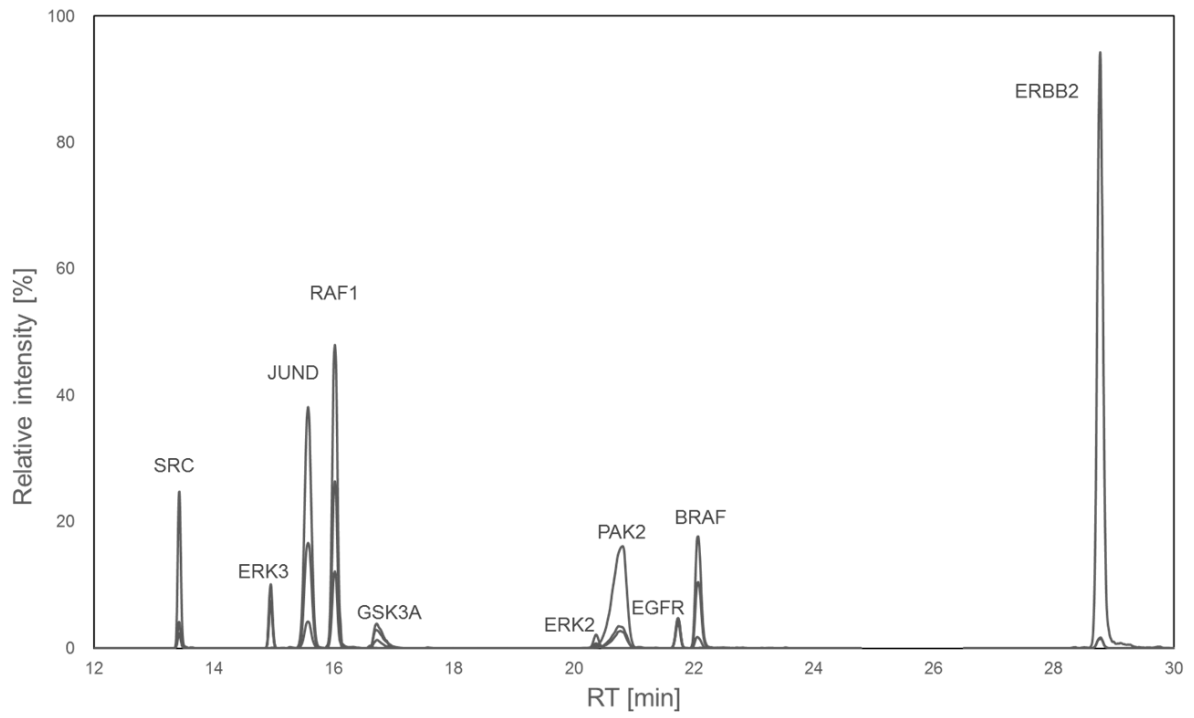


図 2.3 10 種類の標的リン酸化ペプチドの代表的な SRM クロマトグラム

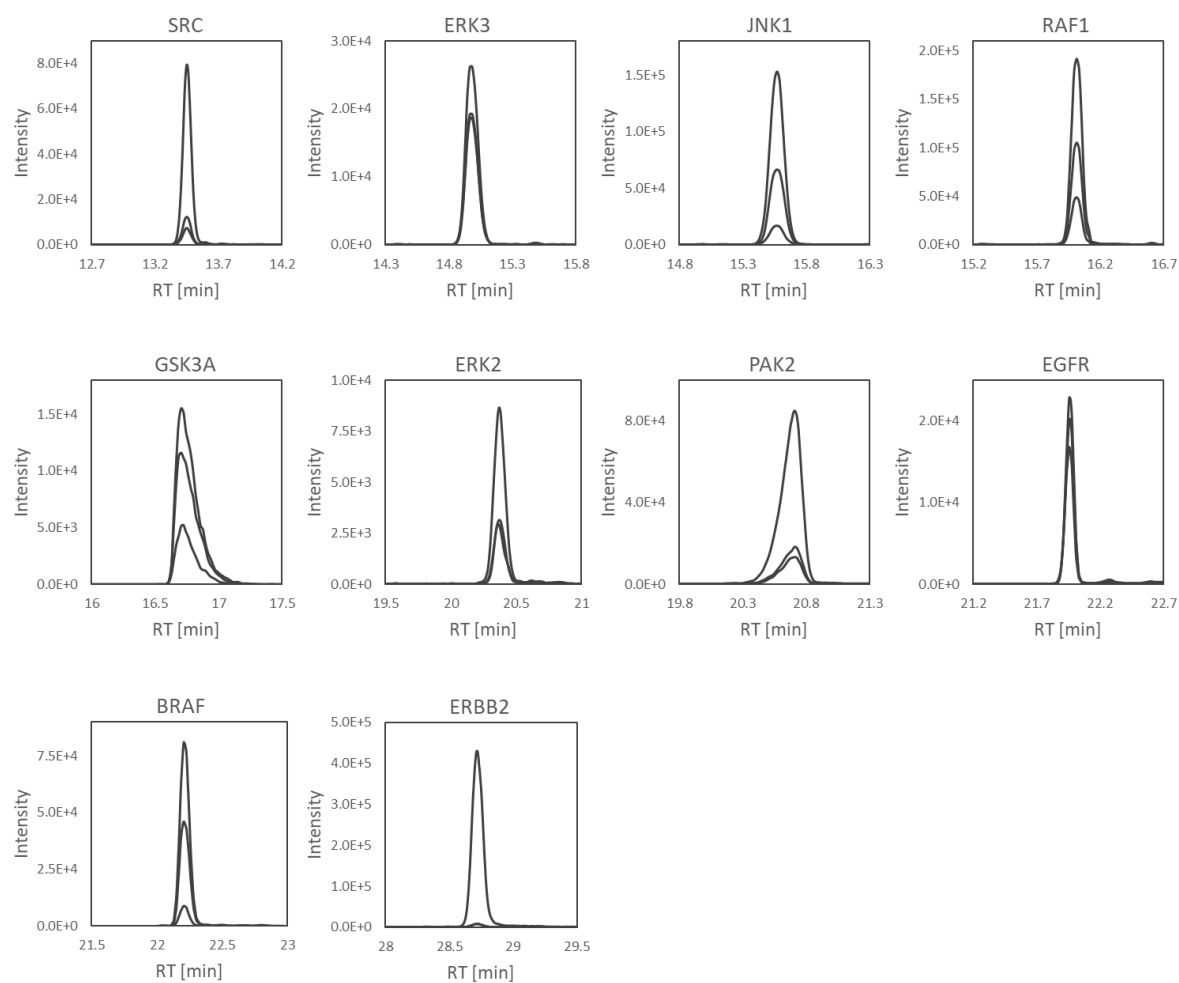


図 2.4 選択したトランジションによる標的リン酸化ペプチドの標準的な SRM ピーク

標的リン酸化ペプチドの詳細を表 2.1 に示す。合成標準リン酸化ペプチド (各 1 pmol) の混合物を、選択したトランジションによる SRM モードでキャピラリー LC/MS/MS により測定した。

第4節 測定方法のバリデーション

構築した SRM アッセイ法を直線性、精度、真度、検出限界 (LOD) および定量限界 (LOQ) に関して検証した。100 μg の HeLa 細胞溶解液から濃縮したリン酸化ペプチドをサンプルマトリックスとして添加した場合と添加しない場合のリン酸化ペプチドの日間解析 ($n = 3$) を行った。7 濃度 (2、10、25、50、100、150、200 pmol/mL) にサンプルを調製し、各濃度を 3 回/日で測定した。合成リン酸化ペプチドは Heavy label、HeLa 細胞由来のマトリックスペプチドは light label した。サンプル調製については、リン酸化ペプチドがバイアルに吸着する傾向があるため、4 つの市販品を用いて吸着性を評価した (図 2.5)。4 %ACN、0.5 %酢酸中に溶解した合成リン酸化ペプチドを新しいバイアルに繰り返し移し、SRM モードで LC/MS 分析した。その結果、ポリマーコーティングを施したポリプロピレン製の TORAST-H Bio バイアルは、他のプラスチック製バイアルと比較して最も高い回収率を示した。

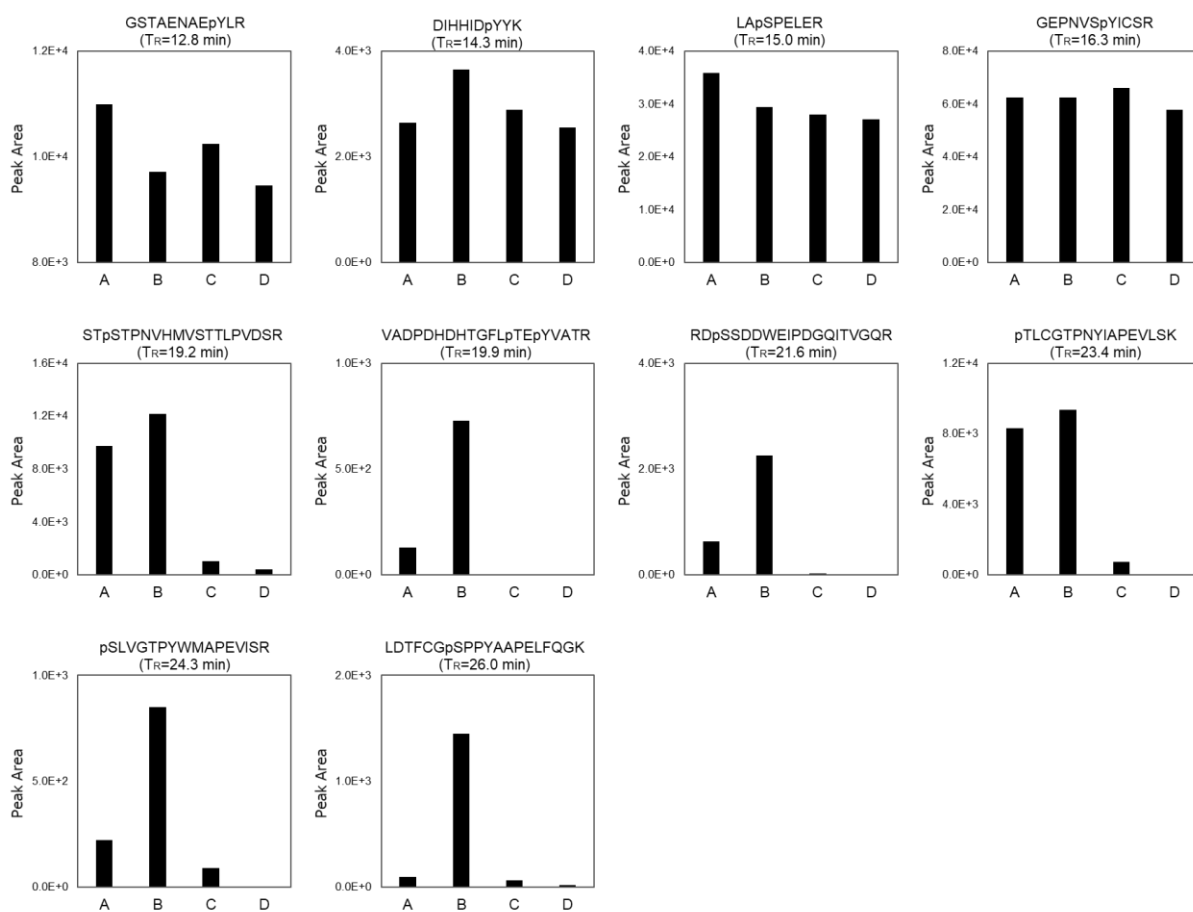


図 2.5 異なるプラスチックバイアルからの 10 種類のリン酸化ペプチドの回収

バイアル A: ポリプロピレン製エッペンドルフチューブ (容量 1.5 mL)、バイアル B: ポリプロピレン製ポリマーコーティング済みバイオバイアル TORAST-H (容量 300 μ L)、バイアル C: TPX[®] (polymethylpentene) 製 Sn バイアル (容量 300 μ L)、バイアル D: ポリプロピレン製 96 ウェル PCR プレート (容量 200 μ L)。バイアル A (溶液 A) を用いて、4 %ACN、0.5 %酢酸溶液に溶解した合成ペプチドを調製した。このバイアル A の溶液を、バイアル B、C、D を用いてそれぞれ 5 回ずつ新しいバイアルに移した (溶液 B、C、D)。次いで、溶液 A~D を SRM モードでキャピラリー LC/MS/MS により分析した。

マトリックスを添加した場合と添加しなかった場合の精度、真度および直線性を図 2.6 および表 2.2、2.3 に示す。すべてのリン酸化ペプチドで、マトリックス効果が見られたが、マトリックスの有無にかかわらず、2 pmol/mL -200 pmol/mL の範囲で良好な直線性 ($R^2 > 0.99$) を示した。精度については、マトリックスの有無にかかわらず、相対標準偏差 (RSD%) はそれぞれ 1.9 % から 30.7 % および 1.0 % から 34.6 % の範囲であった (表 2.2、2.3)。RSD が 20 % 以上で最も高かったのは、2 pmol/mL の EGFR および ERK2 リン酸化ペプチドであった。これは、図 2.3 に示しているように、低いシグナル応答に起因するものと考えられる。真度については、マトリックスを添加した場合は、全濃度範囲で一定の回収率が得られた (82.0 % ~ 115.0 %)。一方で、低濃度のマトリックス添加しなかった場合では偏りのある回収率が得られた (75.2 % ~ 154.55 %)。

LOD および LOQ は、第十七改正日本薬局方に準拠し、ノイズレベル (σ) および検量線の傾き (傾き) に基づいて、以下の式を用いて求めた。

$$\text{LOD} = 3.3\sigma/\text{傾き}$$

$$\text{LOQ} = 10\sigma/\text{傾き}$$

LOD の値は、マトリックスの有無により、それぞれ 0.04 pmol/mL-0.29 pmol/mL、0.03 pmol/mL-0.18 pmol/mL、LOQ の値は、マトリックスの有無により、それぞれ 0.11 pmol/mL -0.87 pmol/mL、0.10 pmol/mL-0.54 pmol/mL であった (表 2.4)。ERK2 ペプチドが最も高い LOD 値と LOQ 値を示したのは、ERK2 キナーゼを完全に活性化するために必要な ERK2 の二重リン酸化によるイオン化効率の低下のためであると考えられる (図 2.3)。測定間のキャリーオーバーを調べるため、200 pmol/mL の試料を注入した後、ブランク試料 (5 μ L) を 3 回続けて分析した。キャリーオーバーは、1 回目で 0.5 % 未満、3 回目では 0.03 % 未満であり、キャリーオーバーがこの SRM アッセイに干渉しないことを示した。3 日間 (合計 $n = 63$) における保持時間の再現性は、0.1~0.8 % (マトリックスあり) と 0.3~1.6 % (マトリックスなし) であり、LC 分離の高い頑強性を示している (表 2.5)。この特徴は、標的ペプチドの数を最大化するための scheduled SRM アッセイにとって、重要である。

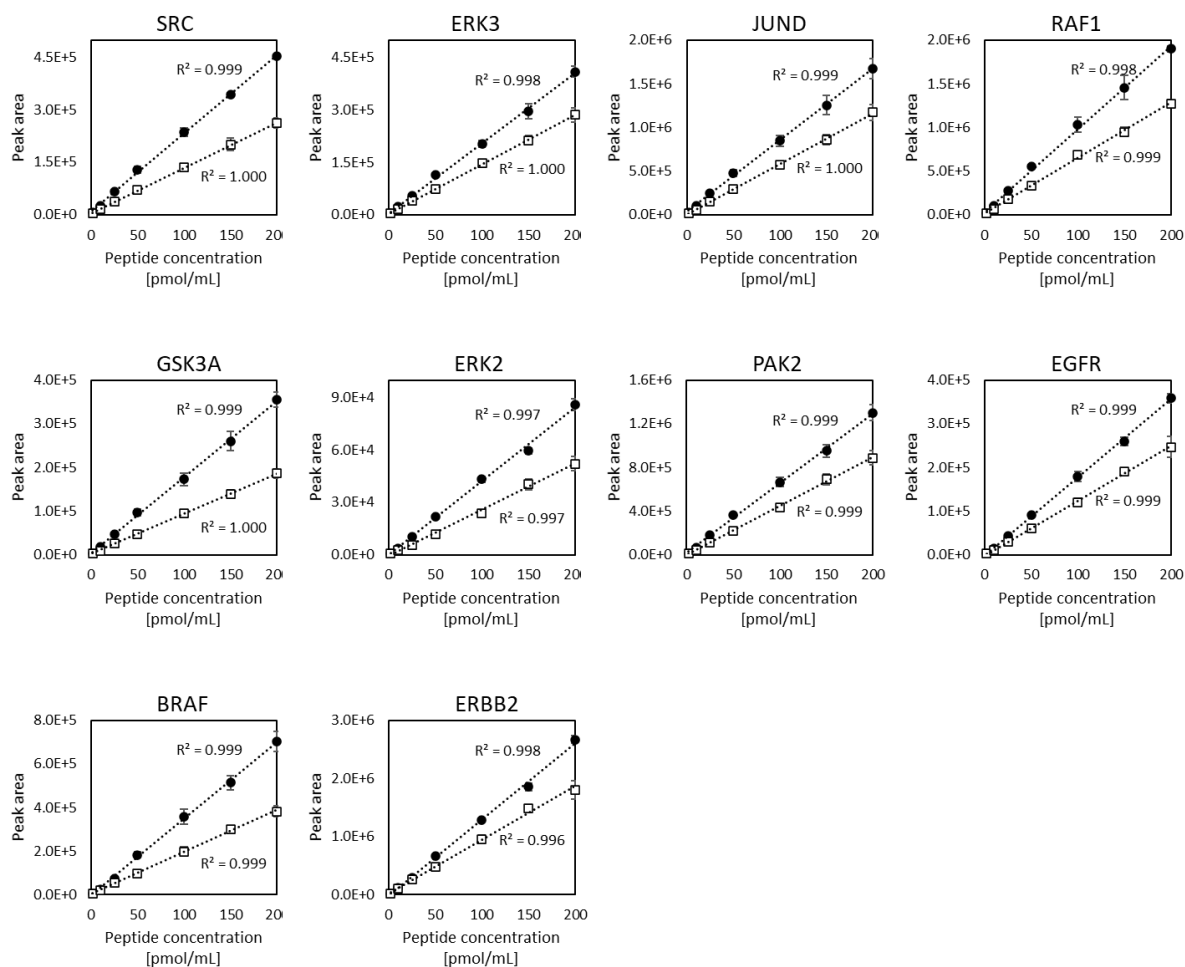


図 2.6 マトリックス添加(□)とマトリックス非添加(●)の Heavy label ペプチドの検量線

7 濃度 (2, 10, 25, 50, 100, 150, 200 pmol/mL) の分析を濃度ごとに 1 日に 3 回行い、HeLa 細胞溶解液から濃縮したリン酸ペプチドをマトリックスとして添加した場合と添加しなかった場合の日間分析 ($n = 3$) を行った。x 軸は Heavy label ペプチドの濃度を示し、y 軸は平均ピーク面積 \pm SD を示す。各検量線の R^2 を示す。

表 2.2 サンプルマトリックスを添加した場合の精度と真度

	Precision [%]						
	2 pmol/mL	10 pmol/mL	25 pmol/mL	50 pmol/mL	100 pmol/mL	150 pmol/mL	200 pmol/mL
SRC	8.4	4.9	1.9	5.2	3.6	8.9	5.7
ERK3	9.6	8.4	5.6	5.0	4.8	6.6	7.0
JUND	5.4	7.9	2.9	5.8	2.1	6.9	7.6
RAF1	10.5	8.1	5.5	6.3	5.7	5.5	3.8
GSK3A	9.7	8.3	5.4	6.5	5.1	7.0	5.3
ERK2	30.7	7.0	7.9	4.8	5.5	7.4	7.7
PAK2	6.8	3.4	4.9	4.9	4.6	7.3	7.5
EGFR	9.3	11.4	5.9	7.3	7.1	2.0	9.7
BRAF	11.5	8.9	8.5	5.0	9.8	4.3	6.2
ERBB2	7.3	6.3	8.3	5.8	7.6	3.4	8.5

	Accuracy [%]						
	2 pmol/mL	10 pmol/mL	25 pmol/mL	50 pmol/mL	100 pmol/mL	150 pmol/mL	200 pmol/mL
SRC	87.2	87.9	88.4	93.7	98.9	99.1	101.5
ERK3	92.8	94.6	92.7	97.2	98.0	100.9	100.3
JUND	94.3	101.1	95.7	99.7	101.9	101.4	98.8
RAF1	87.2	95.1	92.0	96.1	93.9	101.7	101.1
GSK3A	87.8	98.2	91.6	97.8	98.7	100.7	100.2
ERK2	82.0	105.3	115.0	109.8	109.0	96.3	99.3
PAK2	89.8	101.2	98.3	101.3	103.0	97.0	101.0
EGFR	93.7	108.6	107.7	104.0	102.7	97.4	100.4
BRAF	83.3	99.2	88.2	99.7	98.5	97.1	102.3
ERBB2	91.4	95.6	90.5	96.8	98.7	94.6	104.1

表 2.3 サンプルマトリックスを添加しなかった場合の精度と真度

	Precision [%]						
	2 pmol/mL	10 pmol/mL	25 pmol/mL	50 pmol/mL	100 pmol/mL	150 pmol/mL	200 pmol/mL
SRC	5.8	6.2	1.7	7.6	5.0	1.3	1.1
ERK3	6.9	6.9	4.1	3.8	4.7	7.1	3.8
JUND	8.8	9.5	8.7	8.8	7.4	8.7	7.1
RAF1	7.1	4.0	6.9	3.5	8.4	9.7	1.6
GSK3A	7.6	5.8	7.5	7.8	8.0	8.4	5.0
ERK2	17.0	8.1	6.9	1.6	4.0	2.7	3.9
PAK2	4.1	1.8	1.0	4.2	5.8	6.1	5.5
EGFR	34.6	8.4	3.8	1.1	6.7	3.9	3.0
BRAF	8.2	6.5	9.2	7.4	9.3	6.5	6.8
ERBB2	13.9	7.9	4.1	2.3	2.7	3.8	2.2

	Accuracy [%]						
	2 pmol/mL	10 pmol/mL	25 pmol/mL	50 pmol/mL	100 pmol/mL	150 pmol/mL	200 pmol/mL
SRC	85.2	87.4	87.0	90.1	97.9	100.6	101.7
ERK3	86.0	91.2	92.3	89.0	99.8	103.1	99.3
JUND	75.2	84.7	85.6	88.5	99.4	100.9	100.8
RAF1	95.5	92.0	89.4	88.5	94.3	100.2	102.4
GSK3A	99.9	101.0	93.3	91.5	102.1	101.6	99.3
ERK2	133.6	112.6	99.9	95.8	97.2	106.1	97.8
PAK2	92.2	94.6	89.1	89.1	97.7	102.6	100.1
EGFR	154.5	119.9	104.7	96.8	98.9	102.5	99.0
BRAF	134.2	134.2	118.2	95.8	97.6	102.0	99.5
ERBB2	122.4	114.8	108.7	97.4	101.0	105.0	97.2

表 2.4 標的リン酸化ペプチドの LOD と LOQ

	With matrix		Without matrix	
	LOD	LOQ	LOD	LOQ
	[pmol/mL]	[pmol/mL]	[pmol/mL]	[pmol/mL]
SRC	0.12	0.35	0.08	0.23
ERK3	0.12	0.37	0.06	0.18
JUND	0.07	0.21	0.06	0.17
RAF1	0.25	0.77	0.09	0.27
GSK3A	0.11	0.34	0.08	0.23
ERK2	0.29	0.87	0.18	0.54
PAK2	0.04	0.11	0.03	0.10
EGFR	0.11	0.33	0.16	0.48
BRAF	0.10	0.30	0.14	0.41
ERBB2	0.07	0.20	0.10	0.29

表 2.5 標的リン酸化ペプチドの保持時間の再現性 (各 n = 63)

	With matrix		Without matrix	
	RT [min]	RSD [%]	RT [min]	RSD [%]
SRC	13.6	0.3	13.4	0.7
ERK3	15.3	0.8	15.0	1.6
JUND	15.9	0.5	15.6	0.8
RAF1	16.2	0.5	15.9	0.9
GSK3A	17.1	0.3	16.9	0.6
ERK2	20.7	0.1	20.5	0.4
PAK2	21.0	0.3	20.8	0.6
EGFR	22.3	0.3	22.1	0.8
BRAF	22.6	0.2	22.4	0.4
ERBB2	28.9	0.2	28.7	0.3

第5節 過バナジン酸処理 HeLa 細胞のキナーゼ活性測定

広く用いられているチロシンホスファターゼ阻害剤の一つである過バナジン酸処理した HeLa 細胞におけるキナーゼ活性の定量分析に上述の SRM アッセイを適用した。Heavy label された内因性ペプチドの量は、添加した light label 標準合成ペプチドに基づいて定量した (図 2.7)。その結果、コントロール試料、処理試料ともに内在性の ERK3 リン酸化ペプチドの量は LOD 以下であったが、10 個のリン酸化ペプチドの定量に成功した。HeLa 細胞の過バナジン酸誘導リン酸化が活性化されると 6 種類の標的リン酸化ペプチドが増加し、EGFR、ERBB、ERK2、JNK1、SRC、RAF1 の 6 種類のキナーゼが活性化されたのに対し、BRAF の標的リン酸化ペプチドは EGFR-MAPK 経路に属しているにもかかわらず変化がなかった(図 2.7)。これは、これまでに報告されているように、HeLa 細胞では RAF1 が BRAF よりも優勢であるためだと考えられる^{77, 78, 79}。

本結果を確認するために、リン酸化部位特異的抗体を用いて、ウェスタンブロットを行った (図 2.8)。得られた結果は、SRM の結果 (図 2.7) と一致していたが、RAF1 や BRAF ではバンドは検出されず、SRM 解析がウェスタンブロットよりも高感度であることが示唆された。

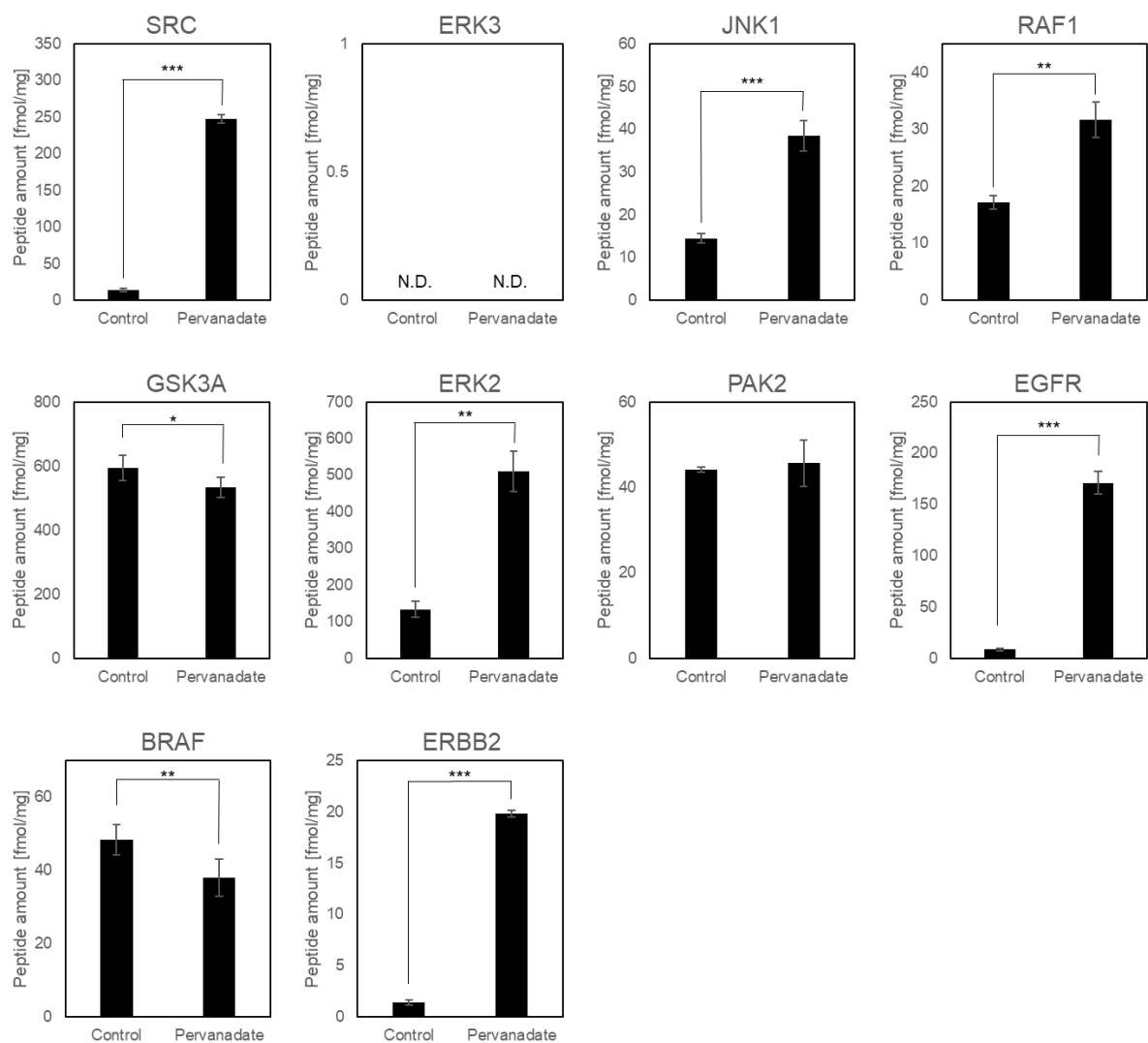


図 2.7 キナーゼ由来のリン酸化ペプチド 9 種類と基質由来のリン酸化ペプチド 1 種類の SRM アッセイによる 10 種類のキナーゼ活性の定量

過バナジン酸処理した HeLa 細胞から得られたリン酸ペプチドは、Heavy label 後、light label 合成リン酸ペプチドと混合した。Heavy label したリン酸ペプチドを定量した (n = 4)。片側 t 検定による p 値を示す。*** p < 0.001, ** p < 0.01, * p < 0.05。

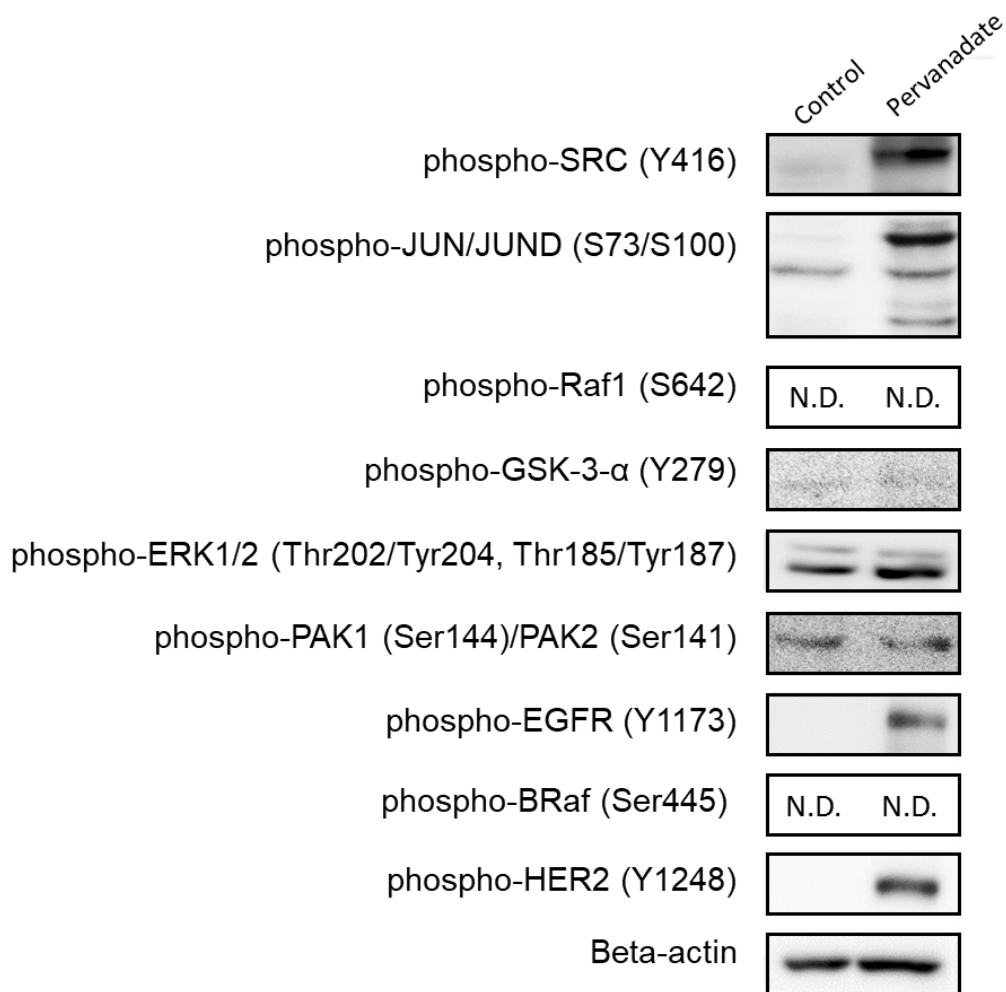


図 2.8. 過バナジジン酸処理した HeLa 細胞のウェスタンブロット

ウェスタンブロットには、8 種類のキナーゼ上のリン酸化部位特異的抗体 (phospho-SRC (Y416)、phospho-RAF1 (S642)、phospho-GSK-3-alpha (Y279)、phospho-ERK1/2 (Thr202/Tyr204、Thr185/Tyr187)、phospho-PAK1 (Ser144)/PAK2 (Ser141)、phospho-EGFR (Y1173)、phospho-BRaf (Ser445)、phospho-HER2 (Y1248))、および基質上のリン酸化部位特異的抗体 (phospho-JUN/JUND (S73/S100)) を使用した。ローディングコントロールとして β -アクチンを使用した。

第6節 おわりに

本章では、細胞内キナーゼ活性をモニターするために、キャピラリーLC/MS/MSをベースとしたアプローチを開発した。キナーゼによるリン酸化を受けたキナーゼ基質消化ペプチド、またはキナーゼ活性を調節するリン酸化部位を含むキナーゼ消化ペプチドを定量するために、選択反応モニタリング (SRM) モードを採用した。EGFR-MAPK シグナル伝達経路に存在する 10 種類のキナーゼを対象とし、標的リン酸化ペプチド配列は、研究室内で有するリン酸化プロテオームライブラリーとキナーゼ活性を調節するリン酸化修飾部位に関する文献情報を用いて選択した。SRM トランジション、LC パラメータ、サンプル前処理ステップなどの実験条件を最適化した。また、精度、真度、直線性、検出限界、定量限界についてのバリデーション試験を実施し、生体試料中のリン酸ペプチド量からキナーゼ活性を測定できることを確認した。最後に、本手法をチロシンホスファターゼ阻害剤である過バナジン酸処理した HeLa 細胞のキナーゼプロファイリングに SRM アッセイを適用した。その結果、10 種類中 6 種類のキナーゼが活性化されており、従来のリン酸化部位特異的抗体を用いたウェスタンブロット法の結果と一致した。

本アプローチでは、SRM 標的ペプチドをあらかじめ設計することができれば、iMPAQT 法 (in vitro proteome-assisted MRM for Protein Absolute Quantification)⁸⁰ のように 1 回の実験で複雑な生体試料のキノームワイドな解析が可能となる。そのための標的リン酸化ペプチドの選択は、適切な SRM アッセイを設計するための重要なポイントの一つである。本章においては、ERK2 の場合、キナーゼ活性は Thr185-X-Tyr187 での二重リン酸化に依存することがよく知られているが、リン酸化プロテオミクスでは pT-または pY-モノリン酸化ペプチドがよく同定される。ERK2 の活性化機構は十分に研究されているので、これらのモノリン酸化ペプチドをモニターしてキナーゼ活性を測定する必要はないが、標的キナーゼの活性化機構が不明な場合には、SRM アッセイのためにあらゆるリン酸化サイトの組み合わせを検討しなければならない。もう一つのポイントは、選択されたペプチド配列の冗長性である。本章で JNK1 の活性を測定するために使用した標的ペプチドは JUND 98-105^{81, 82} で、AP-1 71-78 と同一である。また、これらの部位は JNK1 だけでなく、場合によっては PLK3 によってもリン酸化される⁸³。このような場合には、得られた結果を慎重に解釈する必要がある。このようにキナーゼ活性計測のための SRM アッセイ開発における大きな課題は、キナーゼ活性を代表するリン酸化物の選択である。しかし、この方法は、リン酸基特異的な抗体の入手可能性や品質を考慮する必要がない

ため、現在の抗体法と比較して大きな利点がある。また、限られた試料しか入手できない場合でも、キナーゼ由来または基質由来のリン酸化ペプチドを用いることで、短時間で容易にキナーゼ活性を定量することが可能となる。

結論

プロテオームプロファイリングによるヒトがん細胞株の腫瘍モデルとしての評価と細胞内キナーゼ活性計測のためのリン酸化ペプチドの選択反応モニタリングの開発に向けて研究を行い、以下の知見を得た。

第 1 章では、プロテオームおよびリン酸化プロテオームプロファイリングに基づくヒト乳がん細胞株の腫瘍モデルとしての評価を行うために、24 の乳がん細胞株のプロテオームおよびリン酸化プロテオームプロファイルを取得した。取得したプロファイルと既報から収集したヒト乳がん腫瘍組織 77 試料のプロテオームプロファイルおよび 105 試料のリン酸化プロテオームプロファイルは、相関が低く、PAM50 サブタイプ分類のためのプロテオームおよびリン酸化プロテオームマーカーに絞っても、階層的クラスタリングにおいて、細胞株と腫瘍で異なるクラスターを形成した。しかし、乳がんの臨床バイオマーカーや既知の乳がんに関連するタンパク質およびそのリン酸化サイトの発現パターンは、細胞株と腫瘍間で相関しており、細胞株と腫瘍組織は異なるプロテオームおよびリン酸化プロテオームプロファイルを有していることものの、特定の分子の発現プロファイルには相関があることが示唆された。以上の通り、乳がん腫瘍と細胞株の間には分子的特徴の類似性と非類似性が存在することがわかった。細胞株は、原発腫瘍の分子特性の一部を反映しているにすぎず、すべての分子特性を反映しているわけではないため、乳がん研究のための細胞株モデルを選択する際には、より多くの根拠が必要であると考えられる。

第 2 章では、細胞内キナーゼ活性をモニターするために、キャピラリー LC/MS/MS をベースとしたアプローチを開発した。選択反応モニタリング (SRM) モードを採用し、EGFR-MAPK シグナル伝達経路に存在する 10 種類のキナーゼの活性を測定した。また、精度、真度、直線性、検出限界、定量限界についてのバリデーション試験を実施し、生体試料中のリン酸ペプチド量からキナーゼ活性を測定できることを確認した。本手法を用いて、チロシンホスファターゼ阻害剤である過バナジン酸処理した HeLa 細胞のキナーゼプロファイリングに SRM アッセイを適用した。その結果、10 種類中 6 種類のキナーゼが活性化されており、従来のリン酸化部位特異的抗体を用いたウェスタンブロット法の結果と一致する結果を得た。本手法をキノームワイドに適応することにより、短時間で容易にキノーム活性を定量することが可能となると考えられる。

実験の部

1. 材料

ATCC (American Type Culture Collection) のヒト乳がん細胞株 24 株の細胞ペレットは、木村剛之博士、佐藤義明博士、小田吉哉博士 (エーザイ株式会社) より提供を受けた。HeLa 細胞は、ヒューマンサイエンス研究資源バンク (大阪) より購入した。ダルベッコ改変イーグル培地(DMEM)、カナマイシン、リン酸緩衝生理食塩水 (PBS)、ジチオスレイトール (DTT)、ヨードアセトアミド (IAA)、ペペリジン、重炭酸アンモニウム、デオキシコール酸ナトリウム (SDC)、N-ラウロイルサルコシン酸ナトリウム (SLS)、Mエンドプロテイナーゼ lys-C、酢酸エチル、アセトニトリル、酢酸、メタノール、トリフルオロ酢酸 (TFA) は、富士フイルム和光純薬 (大阪) から購入した。プロテアーゼ阻害薬、重炭酸トリエチルアンモニウムバッファー (TEAB) は、シグマアルドリッチジャパン (東京) より購入した。ウシ胎児血清(FBS)および BCA タンパク質定量キットは、サーモフィッシャーサイエンティフィック株式会社 (横浜) より購入した。PVDF 膜、およびウェスタンブロットリング検出試薬 (ECL Prime western blotting detection reagent) は GE Healthcare (東京) より購入した。ブロッキング試薬はブロッキングワン P をナカライテスク (京都) より購入した。ウェスタンブロットリングの一次抗体には、phospho-SRC (Y416)(Cell Signaling Technology, CST, Danvers, MA, USA)、phospho-JUN/JUND (S73/S100、CST)、phospho-Raf1 (S642、Abnova)、phospho-GSK-3- α (Y279、Santa Cruz Biotechnology Inc.)、phospho-ERK1/2(Thr202/Tyr204、Thr185/Tyr187、CST)、phospho-PAK1 (Ser144)/PAK2 (Ser141、CST)、phospho-EGFR (Y1173、CST)、phospho-BRaf (Ser445、CST)、phospho-HER2 (Y1248、CST)および β -アクチン (CST) を用いた。検出用抗体は全てウサギ由来のものを CST より購入した。トリプシンはプロメガ株式会社 (東京) より購入した。Empore disk は 3M (St. Paul, MN, USA) から入手した。水は Milli-Q システム (Merck, Darmstadt, Germany) で精製した。バイアル評価には、TORAST-H bio バイアル (CatNo. GLCTV-H-BIO) を島津 GLC (東京)、エッペンドルフチューブ (CatNo.0030120086) をエッペンドルフ (Hamburg, Germany)、TPX[®] (polymethylpentene) 製の Sn バイアル (CatNo.1030-51056) をジーエルサイエンス (東京)、96 ウェル PCR プレート (CatNo.PCR-96-FS-C) をコーニング (NY, USA) より入手した。本研究で使用した合成ペプチドは、サーモフィッシャーサイエンティフィック株式会社 (横浜) から購入した。システイン残基を含むペプチドは、使用前にカルバミドメチル化した。

2. HeLa 細胞培養と過バナジン酸処理

HeLa 細胞の培養には 10 % FBS と 100 $\mu\text{g}/\text{mL}$ カナマイシンを添加した DMEM 培地を用いた。過バナジン酸は、100 mM オルトバナジン酸ナトリウムおよび過酸化水素を用いて調製し、DMEM 培地中で適切な濃度に希釈した。HeLa 細胞 (~80 %コンフルエント) を 30 分間、過バナジン酸で処理した。過バナジン酸溶液は、調製後 5 分以内に使用した。過バナジン酸処理した細胞を氷冷リン酸緩衝生理食塩水 (PBS) で洗浄後、氷冷 PBS を用いて氷上で収穫し、使用するまで $-80\text{ }^{\circ}\text{C}$ で凍結した。

3. タンパク質の抽出と消化

試料の前処理は、相間移動可溶化 (PTS) 法⁸⁴に変更を加えて行った。12 mM デオキシコール酸ナトリウム (SDC)、12 mM N-ラウロイルサルコシン酸ナトリウム (SLS)、100 mM トリス-HCl (pH 8.5)、プロテインホスファターゼ阻害剤カクテル 2 および 3 (Sigma-Aldrich) とプロテアーゼ阻害剤 (Sigma-Aldrich) を用いて、ペレットからタンパク質を抽出した。95 $^{\circ}\text{C}$ にて 5 分間静置した後に 10 分間超音波処理した。タンパク質量を BCA タンパク質アッセイキット (サーモフィッシャーサイエンティフィック株式会社) を用いて確認した。10 mM DTT を加えて、30 分間室温で還元し、55 mM IAA を加え室温で 30 分間暗所でアルキル化した。タンパク質混合物を 50 mM 重炭酸アンモニウムバッファーで 5 倍に希釈し、lys-C/trypsin 消化を以前に記載されているように行った⁸⁵。同量の酢酸エチルを加え、0.5 % TFA で酸性化した。混合物を 2 分間振盪し、15,700 g で 2 分間遠心し、水相を回収した。トリプシン消化ペプチドは StageTip^{86,87} により脱塩した。

4. リン酸化ペプチド試料の調製

リン酸化プロテオーム解析用の試料は、 TiO_2 を用いたヒドロキシ酸修飾酸化金属クロマトグラフィーを応用した HAMMOC カラム^{85,88} によるリン酸化ペプチド濃縮法により、ペプチド消化物からリン酸化ペプチドの濃縮を行った。HAMMOC カラムからの溶出には 0.5 % ピペリジンを用いた。その後、SDB-XC を充填した StageTip^{86,87} を用いてペプチドの脱塩・濃縮を行い、LC/MS/MS に導入する試料とした。

5. 同位体ラベル

5.1. TMT ラベル

乳がん細胞のプロテオミクスおよびリン酸化プロテオミクス分析は、TMT9-plex 実験として構成された。分析されたすべてのサンプル間の定量的な比較は、各 9-plex に存在する個々のサンプルと共通の参照サンプル（64 種類のがん細胞由来のトリプシンペプチドの等量混合物）間の TMT レポーターイオン比を使用することによって行った。ペプチドおよびリン酸化ペプチドは、pH 8.5 で 200 mM HEPES 中に懸濁させた。TMT10-plex 試薬（サーモフィッシャーサイエンティフィック）は ACN に溶解した。試料は、図 1.1B に示すように標識した。15 °C で 1 時間ペプチドと TMT 試薬を混合してインキュベートした後、5 %ヒドロキシルアミン溶液を用いて反応をクエンチした。更に15分間のインキュベーションの後、ペプチド溶液を 1 %TFA で酸性化した。標識ペプチドは、逆相-StageTip を使用して脱塩し、その後の LC/MS/MS 分析のためのローディングバッファー（0.5 %TFA、4 %ACN）に懸濁させた。

5.2. ジメチルラベル

安定同位体ジメチルラベルは、Boersema らの報告に変更を加えて行った⁸⁹。手順を簡潔に述べると、脱塩後の溶液中のペプチドを乾固した後に、HeLa 細胞由来のトリプシン消化ペプチドまたは合成ペプチドを 100 μ L の 100 mM TEAB に再溶解した。次いで、ペプチドを 4 %ホルムアルデヒド同位体(¹³CD₂O (Isotec) または CH₂O (Sigma-Aldrich)) 4 μ L と混合し、その後、0.6 M シアノ水素化ホウ素ナトリウム (NaBH₃CN) をそれぞれの試料に 4 μ L ずつ添加し、室温にて一時間静置した。各試料に 16 μ L の 1 %水酸化アンモニウムを添加して反応を止めた。次いで、20 μ L の 10 %TFA をジメチル標識ペプチドをと混合した。その後、二つの試料を混合し、StageTip にてペプチドを精製した。上述の通り、リン酸化ペプチド濃縮を行い、LC/MS/MS を用いて分析した。

6. LC/MS/MS

6.1. NanoLC/MS/MS システム

乳がん細胞株のプロテオームおよびリン酸化プロテオーム測定は、次に示す条件で行った。LC/MS/MS 装置は HTC-PAL (CTC Analytics, Zwingen, Switzerland) および Ultimate3000 pump (DIONEX, Germering, Germany) を装備した、Q Exactive (Thermo Fisher Scientific, San Jose, CA, USA) を使用した。ペプチドの分離には、長さ 200 cm のモノリス型シリカ-C18 キャピラリーカラムを使用した。リン酸化ペプチドの分離には、粒子径 3 μ m の ReproSil-PurC18-AQ (Dr. Maisch, Ammerbuch, Germany) を窒素ガスにて自家充填したスプレーチップ一体型の nanoLC カラム (250 mm length

x 100 μm i.d.) を使用した。試料注入量は 5 μL 、移動相流速は 500 nL/min に設定した。移動相には (A) 0.5 % 酢酸水溶液、および (B) 0.5 % 酢酸含有 80 % アセトニトリル水溶液を用いた。グラジエント条件はペプチド分離には、5-40 %B (8 時間) \rightarrow 40-100 %B (5 分間) \rightarrow 99 %B (10 分間)、またリン酸化ペプチド分離には 5-10 %B (5 分間) \rightarrow 10-40 %B (180 分間) \rightarrow 40-99 %B (5 分間) \rightarrow 99 %B (10 分間) を適用した。スプレー電圧は 2,300 V (ポジティブイオンモード) とし、スキャン範囲を m/z 300-1,500 に設定した。MS スキャンは Orbitrap で 70,000 の分解能で、MS/MS スキャンは Orbitrap で 35,000 の分解能で行った。MS のオートゲインコントロールは 3×10^6 、MS/MS のオートゲインコントロールは 1×10^5 とした。MS スペクトル中最も強度の強い 10 個のプレカーサーイオンに対して、MS₂ スペクトルを取得した。同一イオンに対する MS₂ スキャンを防ぐため、すでにスキャンされたイオンは 30 秒間 MS₂ スキャンの対象から除外するように設定した。また、質量分析計による測定精度を保つため lock mass 機能を採用した。

6.2. キャピラリーLC/MS/MS システム

SRM モードでのペプチドの分析は、島津製作所 LC-20AD ナノポンプとオートサンプラー SIL-30AC を装備した島津製作所 LCMS-8040 トリプル四重極液体クロマトグラフ質量分析計 (島津製作所、京都) を用いて行った。分析カラムは、粒子径 3 μm の ReproSil-PurC18-AQ (Dr. Maisch, Ammerbuch, Germany) を窒素ガスにて自家充填したスプレーチップ一体型の nanoLC カラム (150 mm length x 200 μm i.d.) を用いた⁹⁾。移動相には (A) 0.5 % 酢酸水溶液、および (B) 0.5 % 酢酸含有 80 % アセトニトリル水溶液を用いた。ペプチド混合物を 0.5 % TFA、4 % ACN で再構成した。試料注入量は 5 μL 、移動相流速は 2 $\mu\text{L}/\text{min}$ に設定した。グラジエント条件は、5-10 %B (5 分間)、10-40 %B (20 分間)、40-99 %B (5 分間)、99 %B (5 分間)、5 %B (20 分間) を適用した。LC/MS/MS データは、時間的にスケジュールされた SRM モード (保持時間 \pm 2 分) で取得した。装置のパラメータは以下のように設定した。全ての SRM トランジションに 80 msec の dwell time を使用した。スプレー電圧は 2,300 V (ポジティブイオンモード) とした。HeLa 細胞のキナーゼ活性は、400 μg の HeLa 細胞からリン酸化ペプチド濃縮後に 100 fmol の合成ペプチドをスパイクして測定した。内因性リン酸化ペプチドとスパイクされた内部標準ペプチドのピーク面積値を測定した。得られたデータは、島津 Lab solution ソフトウェアを使用して定量した。

7. データベースサーチ

すべての得られた解析データは、MaxQuant v1.6.14.0 を用いてデータベースサーチを行った⁹¹。ペプチドおよびタンパク質は、Human UniprotKB/Swiss-Prot データベース (2020-06-29 ダウンロード) に対する Andromeda 検索エンジンを用いた自動データベース検索を行った。さらに、デフォルトの contaminant タンパク質データベースが含まれていた。計算質量との許容誤差範囲をプレカーサーイオンで 20 ppm、プロダクトイオンで 4.5 ppm とした。システインのカルバミドメチル化を、すべてのグループの固定修飾として設定した。可変的な修飾としては、メチオニンの酸化、タンパク質の N 末端アセチル化、およびリン酸化プロテオームデータのためのセリン、スレオニンおよびチロシン残基のリン酸化が考慮された。制限酵素にトリプシン/P を選択し、最大 2 つの未切断部位を許容した。ペプチドスペクトラムマッチ (PSM)、PTM サイト、Protein level で False discovery rate (FDR) フィルタリングを 1 % に設定した^{91,92}。IMBR とバッチ効果を除去するための正規化を有効にした³²。その他の MaxQuant のパラメータの値は、デフォルト値から変更しなかった。

8. 情報解析

基本的な配列処理、および情報解析は、Perseus バージョン 1.6.14.0 と RStudio v1.3.3.1093 上の R64 ビットバージョン 4.0.3 の R スクリプトを使用して行った。既知の汚染物質、サイトまたはリバーズによってのみ同定されたタンパク質グループまたはリン酸化サイトは削除した。データは、さらなる分析のために対数化され、各サンプルの中央値の強度を用いて正規化した。

リン酸来基の局在化については、MaxQuant で計算された局在化確率が最も低いものを用いて、0.75 以上のしきい値で局在化されたリン酸化部位をフィルタリングした。階層型クラスタリングは、フリーウェアの Cluster 3.0 を用いて行った⁹³。クラスターは、JAVA TREEVIEW で可視化した。遺伝子発現または、PAM50 分類に伴って、有意にアップまたはダウンレギュレートされたタンパク質グループおよびリン酸化サイトは、Welch の t 検定 (FDR < 0.05) によって決定した。PAM50 のサブタイプ間で有意にアップまたはダウンレギュレーションされたタンパク質は、ANOVA (FDR < 0.05) によっても決定した。DAVID (v.6.8)⁹⁴ による機能解析には、同定された SwissProt ID を解析対象データ (foreground) として使用した。バックグラウンド (background) データとして、web 上で用意されている全てのヒトタンパク質リストの 'Homo sapiens' を選択した。結果として得られた機能同定チャート (functional annotation chart) から、GO タ

ームを p-value (<0.01) が昇順になるようにリストした。FDA が承認した創薬標的の遺伝子リストは、Protein Atlas (<http://www.proteinatlas.org/humanproteomedruggable>) からダウンロードした。細胞株と腫瘍間の mRNA の発現相関解析のためのマイクロアレイによる遺伝子発現データは、GEO から GSE50811⁴⁹ および GSE41998⁴⁸ をダウンロードして用いた。これらのデータは、Affymetrix Microarray Suite5.0 (MAS5.0) アルゴリズムで正規化されている。

9. ウェスタンブロットティング

HeLa 細胞を、125 mM Tris-HCl、10 % (v/v) スクロース、10 % (v/v) 2-メルカプトエタノール、および 4 % (w/v) SDS (pH 6.8) からなる SDS サンプルバッファーで溶解し、等量の細胞溶解液を SDS ゲル上にロードし、ランニングバッファー (25 mM Tris、192 mM グリシン、および 0.1 % SDS) 中で 1 ゲルあたり 20 mA にて泳動した。ウェスタンブロットティング用の PVDF 膜は、メタノールに浸した後、蒸留水に浸し、最後に転写バッファー (25 mM Tris、195 mM グリシン、および 20 % (v/v) メタノール) に浸して親水化した。タンパク質を転写した後、ブロッキング試薬に浸漬して膜表面をブロッキングした。一次抗体と PVDF 膜をインキュベートした後に、膜を洗浄バッファー (20 mM Tris-HCl および 137 mM NaCl、pH 7.6) で洗浄した。続いて、HRP (horseradish peroxidase) が結合した抗ウサギ IgG 二次抗体とインキュベートを行い、膜を洗浄した。その後、ルミノールを含む検出液を用いて、抗体結合部位を検出し、CCD (Charged Coupled Device) カメラ (Image Quant LAS4000) にて撮影した。

発表論文

第 2 章

Selected Reaction Monitoring of Kinase Activity-Targeted Phosphopeptides

Chisato TAKAHASHI, Tatsuya YAZAKI, Naoyuki SUGIYAMA, Yasushi
ISHIHAMA

2019 年 6 月発行 *Chromatography* 第 40 巻第 2 号 39-47 頁

謝辞

この学位論文の執筆に際し、多くの方々のご指導、ご支援、ご厚意をいただきました。心より感謝申し上げます。

本研究を行うに当たり、終始ご指導を賜りました、京都大学大学院薬学研究科 石濱泰教授に深謝申し上げます。

また、本研究の遂行を基礎から指導して頂きました、京都大学大学院薬学研究科 杉山直幸准教授ならびに、今見考志特任講師、吉沢明康特定助教、国立循環器病研究センター 創薬オミックス解析センター 若林真樹室長に深く感謝申し上げます。

勤務先である同志社女子大学ならびに、同志社女子大学薬学部 前川京子教授には、勤務中の学位論文のための作業にご快諾いただき、勤務と学位論文の両立にご理解をいただきました。深く感謝申し上げます。

本研究を進めるにあたり、ATCC (American Type Culture Collection) のヒト乳がん細胞株 24 株の細胞ペレットをご提供していただいた木村剛之博士、佐藤義明博士、小田吉哉博士 (エーザイ株式会社) に感謝申し上げます。

また、本研究を進めるにあたり、有益なご助言を賜りました田畑剛氏、京都大学 iPS 細胞研究所 岩崎未央特定助教、システム・バイオロジー研究機構 金子 (今村) 春菜博士、シオノギ製薬 矢崎達也氏、京都大学大学院薬学研究科 小形公亮特定助教、京都大学大学院薬学研究科 製剤機能解析学分野の卒業生ならびに在学生の皆様に感謝致します。

最後に、これまでの研究生活を支えてくれた家族に感謝致します。

引用文献

1. Integrated genomic analyses of ovarian carcinoma. *Nature* **474**, 609–615 (2011).
2. Tomczak, K., Czerwińska, P. & Wiznerowicz, M. Review The Cancer Genome Atlas (TCGA): an immeasurable source of knowledge. *Współczesna Onkol.* **1A**, 68–77 (2015).
3. Hudson, T. J. *et al.* International network of cancer genome projects. *Nature* **464**, 993–998 (2010).
4. Vogelstein, B. *et al.* Cancer genome landscapes. *Science* **339**, 1546–1558 (2013).
5. Garraway, L. A. & Lander, E. S. Lessons from the cancer genome. *Cell* **153**, 17–37 (2013).
6. Gygi, S. P., Rochon, Y., Franza, B. R. & Aebersold, R. Correlation between Protein and mRNA Abundance in Yeast. *Mol. Cell. Biol.* **19**, 1720 LP – 1730 (1999).
7. Anderson, N. L. & Anderson, N. G. Proteome and proteomics: new technologies, new concepts, and new words. *Electrophoresis* **19**, 1853–1861 (1998).
8. Gholami, A. M. *et al.* Global proteome analysis of the NCI-60 cell line panel. *Cell Rep.* (2013) doi:10.1016/j.celrep.2013.07.018.
9. Coscia, F. *et al.* Integrative proteomic profiling of ovarian cancer cell lines reveals precursor cell associated proteins and functional status. *Nat. Commun.* **7**, (2016).
10. Frejno, M. *et al.* Proteome activity landscapes of tumor cell lines determine drug responses. *Nat. Commun.* **11**, 1–12 (2020).
11. Tyanova, S. *et al.* Proteomic maps of breast cancer subtypes. *Nat. Commun.* **7**, 1–11 (2016).
12. Nusinow, D. P. *et al.* Quantitative Proteomics of the Cancer Cell Line Encyclopedia. *Cell* **180**, 387–402.e16 (2020).
13. Mertins, P. *et al.* Proteogenomics connects somatic mutations to signalling in breast cancer. *Nature* **534**, 55–62 (2016).
14. Zhang, H. *et al.* Integrated Proteogenomic Characterization of Human High-Grade Serous Ovarian Cancer. *Cell* (2016) doi:10.1016/j.cell.2016.05.069.
15. Jaffe, J. D., Berg, H. C. & Church, G. M. Proteogenomic mapping as a

- complementary method to perform genome annotation. *Proteomics* **4**, 59–77 (2004).
16. Barretina, J. *et al.* The Cancer Cell Line Encyclopedia enables predictive modelling of anticancer drug sensitivity. *Nature* **483**, 603–607 (2012).
 17. Ghandi, M. *et al.* Next-generation characterization of the Cancer Cell Line Encyclopedia. *Nature* **569**, 503–508 (2019).
 18. Tsherniak, A. *et al.* Defining a Cancer Dependency Map. *Cell* **170**, 564–576.e16 (2017).
 19. Needham, E. J., Parker, B. L., Burykin, T., James, D. E. & Humphrey, S. J. Illuminating the dark phosphoproteome. *Sci. Signal.* **12**, (2019).
 20. Perou, C. M. *et al.* Molecular portraits of human breast tumours. *Nature* **406**, 747–752 (2000).
 21. Sørliie, T. *et al.* Repeated observation of breast tumor subtypes in independent gene expression data sets. *Proc. Natl. Acad. Sci. U. S. A.* **100**, 8418–8423 (2003).
 22. Bernard, P. S. *et al.* Supervised risk predictor of breast cancer based on intrinsic subtypes. *J. Clin. Oncol.* **27**, 1160–1167 (2009).
 23. Ellis, M. J. *et al.* Randomized phase II neoadjuvant comparison between letrozole, anastrozole, and exemestane for postmenopausal women with estrogen receptor-rich stage 2 to 3 breast cancer: clinical and biomarker outcomes and predictive value of the baseline PAM50-based in. *J. Clin. Oncol. Off. J. Am. Soc. Clin. Oncol.* **29**, 2342–2349 (2011).
 24. Chia, S. K. *et al.* A 50-gene intrinsic subtype classifier for prognosis and prediction of benefit from adjuvant tamoxifen. *Clin. cancer Res. an Off. J. Am. Assoc. Cancer Res.* **18**, 4465–4472 (2012).
 25. LASFARGUES, E. Y. & OZZELLO, L. Cultivation of human breast carcinomas. *J. Natl. Cancer Inst.* **21**, 1131–1147 (1958).
 26. Koboldt, D. C. *et al.* Comprehensive molecular portraits of human breast tumours. *Nature* **490**, 61–70 (2012).
 27. Kandoth, C. *et al.* Mutational landscape and significance across 12 major cancer types. *Nature* **502**, 333–339 (2013).
 28. Curtis, C. *et al.* The genomic and transcriptomic architecture of 2,000 breast tumours reveals novel subgroups. *Nature* **486**, 346–352 (2012).

29. Nik-Zainal, S. *et al.* Mutational processes molding the genomes of 21 breast cancers. *Cell* **149**, 979–993 (2012).
30. Liu, K. *et al.* Evaluating cell lines as models for metastatic breast cancer through integrative analysis of genomic data. *Nat. Commun.* **10**, (2019).
31. Jiang, G. *et al.* Comprehensive comparison of molecular portraits between cell lines and tumors in breast cancer. *BMC Genomics* **17**, (2016).
32. Cox, J., Yu, S. H. & Kyriakidou, P. Isobaric matching between runs and novel PSM-level normalization in maxquant strongly improve reporter ion-based quantification. *J. Proteome Res.* **19**, 3945–3954 (2020).
33. Schwacke, J. H., Hill, E. G., Krug, E. L., Comte-Walters, S. & Schey, K. L. iQuantitator: A tool for protein expression inference using iTRAQ. *BMC Bioinformatics* **10**, 342 (2009).
34. Herbrich, S. M. *et al.* Statistical inference from multiple iTRAQ experiments without using common reference standards. *J. Proteome Res.* **12**, 594–604 (2013).
35. Yi, J. *et al.* Trefoil factor 1 (TFF1) is a potential prognostic biomarker with functional significance in breast cancers. *Biomed. Pharmacother.* **124**, 109827 (2020).
36. Oki, S. *et al.* ChIP-Atlas: a data-mining suite powered by full integration of public ChIP-seq data. *EMBO Rep.* **19**, 1–10 (2018).
37. Carroll, J. S. *et al.* Chromosome-wide mapping of estrogen receptor binding reveals long-range regulation requiring the forkhead protein FoxA1. *Cell* **122**, 33–43 (2005).
38. Voduc, D., Cheang, M. & Nielsen, T. GATA-3 expression in breast cancer has a strong association with estrogen receptor but lacks independent prognostic value. *Cancer Epidemiol. Biomarkers Prev.* **17**, 365–373 (2008).
39. Qian, X. L. *et al.* Caveolin-1: A multifaceted driver of breast cancer progression and its application in clinical treatment. *Onco. Targets. Ther.* **12**, 1539–1552 (2019).
40. Coates, A. S. *et al.* Tailoring therapies--improving the management of early breast cancer: St Gallen International Expert Consensus on the Primary Therapy of Early Breast Cancer 2015. *Ann. Oncol. Off. J. Eur. Soc. Med. Oncol.* **26**, 1533–

- 1546 (2015).
41. Cheang, M. C. U. *et al.* Ki67 index, HER2 status, and prognosis of patients with luminal B breast cancer. *J. Natl. Cancer Inst.* **101**, 736–750 (2009).
 42. Senkus, E. *et al.* Primary breast cancer: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up. *Ann. Oncol. Off. J. Eur. Soc. Med. Oncol.* **26 Suppl 5**, v8-30 (2015).
 43. Kim, H. K. *et al.* Discordance of the PAM50 Intrinsic Subtypes Compared with Immunohistochemistry-Based Surrogate in Breast Cancer Patients: Potential Implication of Genomic Alterations of Discordance. *Cancer Res. Treat.* **51**, 737–747 (2019).
 44. van 't Veer, L. J. *et al.* Gene expression profiling predicts clinical outcome of breast cancer. *Nature* **415**, 530–536 (2002).
 45. Cobleigh, M. A. *et al.* Tumor gene expression and prognosis in breast cancer patients with 10 or more positive lymph nodes. *Clin. cancer Res. an Off. J. Am. Assoc. Cancer Res.* **11**, 8623–8631 (2005).
 46. Naoi, Y. & Noguchi, S. Multi-gene classifiers for prediction of recurrence in breast cancer patients. *Breast Cancer* **23**, 12–18 (2016).
 47. Wishart, D. S. *et al.* DrugBank: a comprehensive resource for in silico drug discovery and exploration. *Nucleic Acids Res.* **34**, D668-72 (2006).
 48. Horak, C. E. *et al.* Biomarker analysis of neoadjuvant doxorubicin/cyclophosphamide followed by ixabepilone or paclitaxel in early-stage breast cancer. *Clin. Cancer Res.* **19**, 1587–1595 (2013).
 49. Dezso, Z. *et al.* Gene expression profiling reveals Epithelial Mesenchymal Transition (EMT) genes can selectively differentiate eribulin sensitive breast cancer cells. *PLoS One* (2014) doi:10.1371/journal.pone.0106131.
 50. Dai, X., Xiang, L., Li, T. & Bai, Z. Cancer Hallmarks, Biomarkers and Breast Cancer Molecular Subtypes. *J. Cancer* **7**, 1281–1294 (2016).
 51. Choudhary, C. & Mann, M. Decoding signalling networks by mass spectrometry-based proteomics. *Nat. Rev. Mol. Cell Biol.* (2010) doi:10.1038/nrm2900.
 52. Beltrao, P., Bork, P., Krogan, N. J. & Van Noort, V. Evolution and functional cross-talk of protein post-translational modifications. *Mol. Syst. Biol.* **9**, 1–13 (2013).

53. Doll, S. & Burlingame, A. L. Mass spectrometry-based detection and assignment of protein posttranslational modifications. *ACS Chem. Biol.* **10**, 63–71 (2015).
54. Riley, N. M. & Coon, J. J. Phosphoproteomics in the Age of Rapid and Deep Proteome Profiling. *Anal. Chem.* **88**, 74–94 (2016).
55. Taylor, S. S. & Radzio-Andzelm, E. Three protein kinase structures define a common motif. *Structure* **2**, 345–355 (1994).
56. Nolen, B., Taylor, S. & Ghosh, G. Regulation of protein kinases: Controlling activity through activation segment conformation. *Mol. Cell* **15**, 661–675 (2004).
57. IWASAKI, M. & ISHIHAMA, Y. Challenges Facing Complete Human Proteome Analysis. *CHROMATOGRAPHY* **35**, 73–80 (2014).
58. Breitkreutz, A. *et al.* A global protein kinase and phosphatase interaction network in yeast. *Science* **328**, 1043–1046 (2010).
59. Gingras, A.-C., Gstaiger, M., Raught, B. & Aebersold, R. Analysis of protein complexes using mass spectrometry. *Nat. Rev. Mol. Cell Biol.* **8**, 645–654 (2007).
60. Kaake, R. M., Wang, X. & Huang, L. Profiling of protein interaction networks of protein complexes using affinity purification and quantitative mass spectrometry. *Mol. Cell. Proteomics* **9**, 1650–1665 (2010).
61. Rinner, O. *et al.* An integrated mass spectrometric and computational framework for the analysis of protein interaction networks. *Nat. Biotechnol.* **25**, 345–352 (2007).
62. Gerber, S. A., Rush, J., Stemman, O., Kirschner, M. W. & Gygi, S. P. Absolute quantification of proteins and phosphoproteins from cell lysates by tandem MS. *Proc. Natl. Acad. Sci.* (2003) doi:10.1073/pnas.0832254100.
63. Lange, V., Picotti, P., Domon, B. & Aebersold, R. Selected reaction monitoring for quantitative proteomics: a tutorial. *Mol. Syst. Biol.* (2008) doi:10.1038/msb.2008.61.
64. Picotti, P., Bodenmiller, B., Mueller, L. N., Domon, B. & Aebersold, R. Full Dynamic Range Proteome Analysis of *S. cerevisiae* by Targeted Proteomics. *Cell* (2009) doi:10.1016/j.cell.2009.05.051.
65. Bisson, N. *et al.* Selected reaction monitoring mass spectrometry reveals the dynamics of signaling through the GRB2 adaptor. *Nat. Biotechnol.* (2011) doi:10.1038/nbt.1905.

66. Dong, M.-Q. *et al.* Quantitative Mass Spectrometry Identifies Insulin Signaling Targets in *C. elegans*; *Science* (80-.). **317**, 660 LP – 663 (2007).
67. Xiao, Y., Guo, L. & Wang, Y. A Targeted Quantitative Proteomics Strategy for Global Kinome Profiling of Cancer Cells and Tissues. *Mol. Cell. Proteomics* (2014) doi:10.1074/mcp.M113.036905.
68. Urisman, A. *et al.* An optimized chromatographic strategy for multiplexing in parallel reaction monitoring mass spectrometry: Insights from quantitation of activated kinases. *Mol. Cell. Proteomics* **16**, 265–277 (2017).
69. Wolf-Yadlin, A., Hautaniemi, S., Lauffenburger, D. A. & White, F. M. Multiple reaction monitoring for robust quantitative proteomic analysis of cellular signaling networks. *Proc. Natl. Acad. Sci.* (2007) doi:10.1073/pnas.0608638104.
70. Kennedy, J. J. *et al.* Immobilized Metal Affinity Chromatography Coupled to Multiple Reaction Monitoring Enables Reproducible Quantification of Phospho-signaling. *Mol. Cell. Proteomics* (2016) doi:10.1074/mcp.O115.054940.
71. Yang, T. Y., Eissler, C. L., Hall, M. C. & Parker, L. L. A Multiple Reaction Monitoring (MRM) Method to Detect Bcr-Abl Kinase Activity in CML Using a Peptide Biosensor. *PLoS One* **8**, 1–10 (2013).
72. Shi, T. *et al.* Sensitive targeted quantification of ERK phosphorylation dynamics and stoichiometry in human cells without affinity enrichment. *Anal. Chem.* (2015) doi:10.1021/ac503797x.
73. Zhang, G. *et al.* Mass Spectrometry Mapping of Epidermal Growth Factor Receptor Phosphorylation Related to Oncogenic Mutations and Tyrosine Kinase Inhibitor Sensitivity. *J. Proteome Res.* **10**, 305–319 (2011).
74. Atrih, A. *et al.* Stoichiometric quantification of Akt phosphorylation using LC-MS/MS. *J. Proteome Res.* (2010) doi:10.1021/pr900572h.
75. Hornbeck, P. V *et al.* PhosphoSitePlus: a comprehensive resource for investigating the structure and function of experimentally determined post-translational modifications in man and mouse. *Nucleic Acids Res.* **40**, D261-70 (2012).
76. Fumio Matsuda. ターゲットプロテオーム解析をはじめのノウハウ A Practical Introduction to Targeted Proteomics. *Proteome Lett.* **1**, 89–94 (2016).

77. McPhillips, F. *et al.* Raf-1 is the predominant Raf isoform that mediates growth factor-stimulated growth in ovarian cancer cells. *Carcinogenesis* **27**, 729–739 (2006).
78. Yoshiki, S. *et al.* Ras and calcium signaling pathways converge at Raf1 via the Shoc2 scaffold protein. *Mol. Biol. Cell* **21**, 1088–1096 (2010).
79. Zhao, Z., Tan, Z., Diltz, C. D., You, M. & Fischer, E. H. Activation of mitogen-activated protein (MAP) kinase pathway by pervanadate, a potent inhibitor of tyrosine phosphatases. *J. Biol. Chem.* **271**, 22251–22255 (1996).
80. Matsumoto, M. *et al.* A large-scale targeted proteomics assay resource based on an in vitro human proteome. *Nat. Methods* **14**, 251–258 (2017).
81. Hibi, M., Lin, A., Smeal, T., Minden, A. & Karin, M. Identification of an oncoprotein- and UV-responsive protein kinase that binds and potentiates the c-Jun activation domain. *Genes Dev.* **7**, 2135–2148 (1993).
82. Huang, J. *et al.* The same pocket in menin binds both MLL and JUND but has opposite effects on transcription. *Nature* **482**, 542–546 (2012).
83. Wang, L., Dai, W. & Lu, L. Stress-induced c-Jun activation mediated by Polo-like kinase 3 in corneal epithelial cells. *J. Biol. Chem.* **282**, 32121–32127 (2007).
84. Masuda, T., Tomita, M. & Ishihama, Y. Phase transfer surfactant-aided trypsin digestion for membrane proteome analysis. *J. Proteome Res.* (2008) doi:10.1021/pr700658q.
85. Sugiyama, N. *et al.* Phosphopeptide Enrichment by Aliphatic Hydroxy Acid-modified Metal Oxide Chromatography for Nano-LC-MS/MS in Proteomics Applications. *Mol. Cell. Proteomics* (2007) doi:10.1074/mcp.T600060-MCP200.
86. Rappsilber, J., Mann, M. & Ishihama, Y. Protocol for micro-purification, enrichment, pre-fractionation and storage of peptides for proteomics using StageTips. *Nat. Protoc.* (2007) doi:10.1038/nprot.2007.261.
87. Rappsilber, J., Ishihama, Y. & Mann, M. Stop And Go Extraction tips for matrix-assisted laser desorption/ionization, nanoelectrospray, and LC/MS sample pretreatment in proteomics. *Anal. Chem.* (2003) doi:10.1021/ac026117i.
88. Kyono, Y., Sugiyama, N., Imami, K., Tomita, M. & Ishihama, Y. Successive and selective release of phosphorylated peptides captured by hydroxy acid-modified metal oxide chromatography. *J. Proteome Res.* (2008) doi:10.1021/pr800305y.

89. Boersema, P. J., Raijmakers, R., Lemeer, S., Mohammed, S. & Heck, A. J. R. Multiplex peptide stable isotope dimethyl labeling for quantitative proteomics. *Nat. Protoc.* (2009) doi:10.1038/nprot.2009.21.
90. Ishihama, Y., Rappsilber, J., Andersen, J. S. & Mann, M. Microcolumns with self-assembled particle frits for proteomics. in *Journal of Chromatography A* (2002). doi:10.1016/S0021-9673(02)01402-4.
91. Cox, J. & Mann, M. MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. *Nat. Biotechnol.* (2008) doi:10.1038/nbt.1511.
92. Elias, J. E. & Gygi, S. P. Target-decoy search strategy for increased confidence in large-scale protein identifications by mass spectrometry. *Nat. Methods* **4**, 207–214 (2007).
93. Eisen, M. B., Spellman, P. T., Brown, P. O. & Botstein, D. Cluster analysis and display of genome-wide expression patterns. *Proc. Natl. Acad. Sci. U. S. A.* **95**, 14863–14868 (1998).
94. Huang, D. W., Sherman, B. T. & Lempicki, R. A. Bioinformatics enrichment tools: Paths toward the comprehensive functional analysis of large gene lists. *Nucleic Acids Res.* **37**, 1–13 (2009).