

Agilent LC/Q-TOF を用いたタンパク質 消化物の信頼性の高いペプチドマッピング

著者

Stephen Madden, Crystal Cody,
and Jungkap Park
Agilent Technologies, Inc.
Santa Clara, California, USA

はじめに

バイオ医薬品業界では、パイプラインに加わるタンパク質治療薬の数が増大しています。5 年間の年平均成長率 (CAGR) は 7.3 % で、販売額は 2020 年までに 2,487 億ドルに達するものと予測されています¹。5 年間の CAGR が 16.6 % と、最も急速に成長している分野はモノクローナル抗体 (mAb) で、がんや自己免疫疾患などの一定の疾患で使用が認可されています。タンパク質治療薬の開発中には、タンパク質配列や翻訳後修飾 (PTM) などの確認を含め、広範囲の重要な品質特性 (CQA) のモニタリングが必要です。

通常、タンパク質が酵素で消化されるボトムアップ手法を用いて、5 ～ 60 個のアミノ酸で構成されたペプチドが作られます。ペプチド混合物は、液体クロマトグラフィーで分離された後、質量分析計 (MS) で検出され、各混合物についてマススペクトルが得られます。大量のマスデータの中から、ペプチド由来のスペクトルを見つけて、配列および PTM を正確かつ効率的に抽出することは困難であり、検査による手動アプローチは多大な時間と専門知識が必要です。

Agilent MassHunter BioConfirm B.09.00 により、ペプチドの発見と結果の視覚化のためのペプチドマッピングワークフローが進歩しました。MS/MS データを使用するソフトウェアアルゴリズムによって可能性のあるペプチド範囲を絞ることで、手動検査と比べて時間を短縮できます。結果は配列カバーマップに表示され、ユーザーは複数のサンプル (例: 複数の消化酵素) を同時に確認することができます。反復 MS/MS と呼ばれる Agilent 6545XT AdvanceBio LC/Q-TOF での Agilent MassHunter 採取ソフトウェアの新しいモードは、低濃度の PTM に焦点を当て、これを配列カバーマップで視覚化することができます。

プロセスの概要

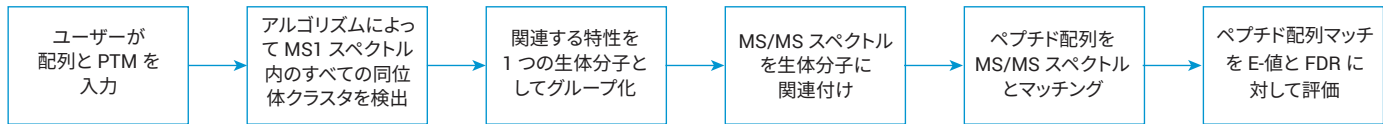


図 1. ペプチドマッピングプロセスの概要

タンパク質配列および PTM の設定

ユーザーは、MassHunter BioConfirm に付属するシーケンスマネージャーを使用し、タンパク質配列を手動で入力するか、あるいは .TXT、.FASTA、.PSQ ファイルをインポートして入力します (図 2)。配列は、複数の鎖、グローバルに明示されている PTM あるいは配列の特定の位置で指定される PTM に合わせて設定できます。

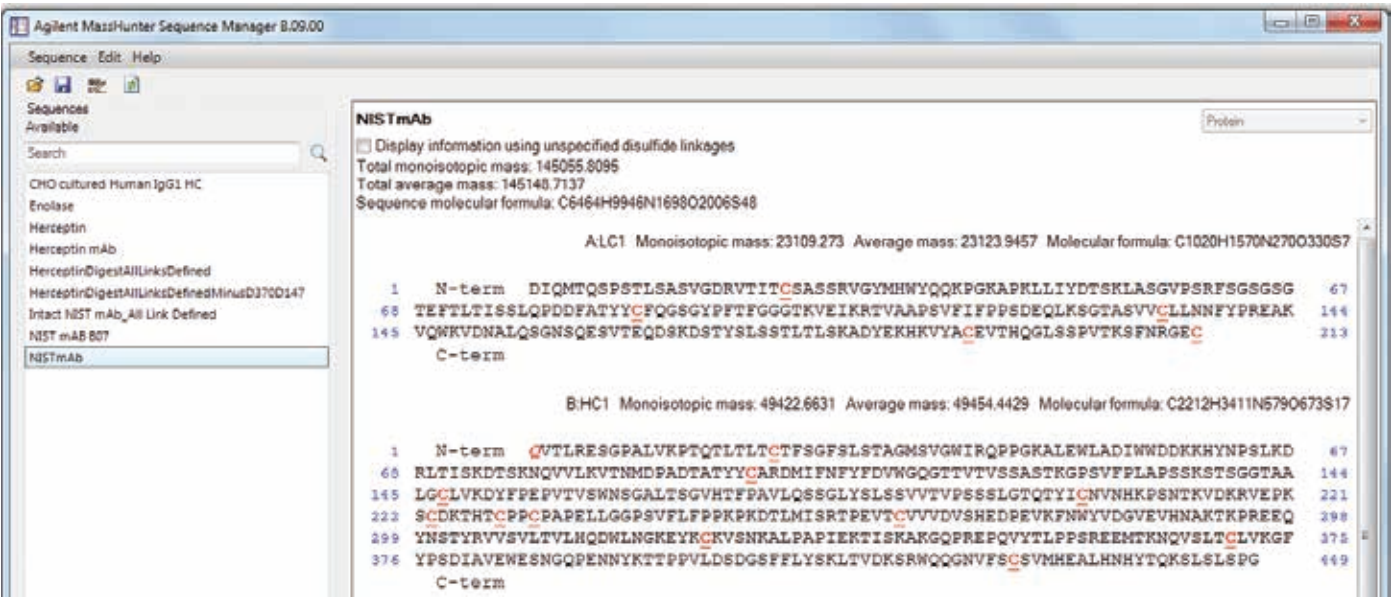


図 2. タンパク質情報を入力するためのシーケンスマネージャー

ユーザーは、ジスルフィド結合の位置を指定できます (図 3)。例えば、型 IgG1 の mAb には鎖間および鎖内に 16 のジスルフィド結合があります。

最後に、許容される切断ミスを含め、トリプシンや LysC などの消化酵素を指定することができます。この情報を用いて、BioConfirm ではタンパク消化で作成される可能性があるペプチド (その PTM および結合リンクを含め) のコンピュータによる完全なモデルを生成できます。ただし、メインの BioConfirm プログラム内でデータファイルを処理しているときに消化酵素が選択されることもあります。

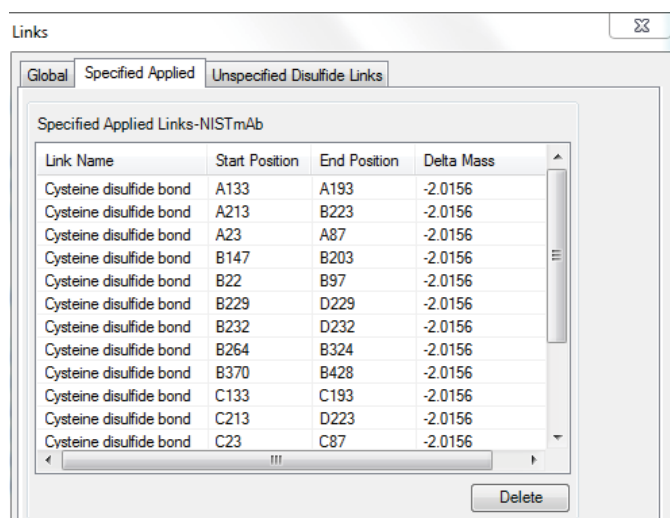


図 3. シークエンスマネージャーで指定されたシステインジスルフィド結合

生体分子のリストの生成

ペプチドマッピングのプロセスは、ハイスルーブット、自動、MassHunter Walkup によるオープンアクセスなど、複数の方法で開始できます。サンプルごとに、ユーザーは [Reprocess Sample](#) ダイアログボックスを用いてサンプルを処理できます (図 4)。ユーザーは、Agilent LC/Q-TOF で採取した MS/MS データファイルで開始して、最初にタンパク分解ワークフローを選択し、次に条件 (ジスルフィド結合の還元・非還元)、配列 (シークエンスマネージャーから入力)、適用する任意の修飾またはプロファイル (例: 糖鎖の型、アルキル化)、消化酵素を選択します。[Reprocess](#) をクリックすると、分析が開始されます。

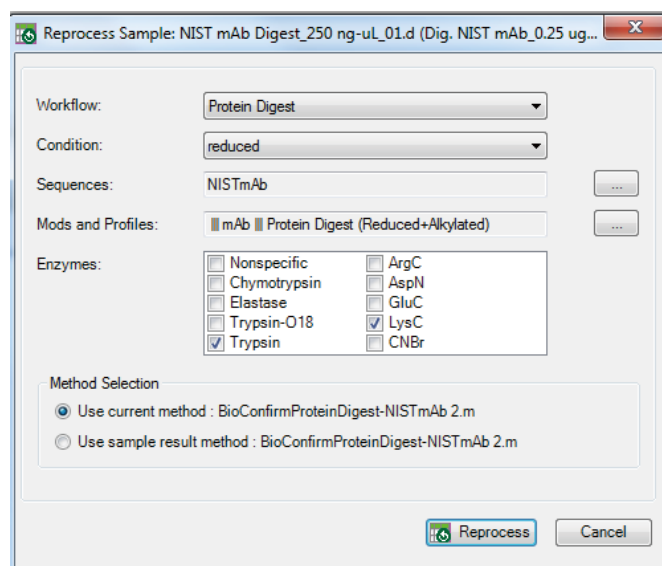


図 4. Reprocess Sample ダイアログでは、カスタムパラメータを使用して 1 つのデータファイルの分析を設定できます。

次に、ペプチドマッピングアルゴリズムがデータファイル全体に適用され、ペプチドに属する可能性のある特性を持つ MS1 データ中の同位体クラスが検出されます。荷電状態、リテンションタイム全体に対するスキャン範囲、付加体 (例えば図 5 の BioConfirm メソッドで指定されているプロトン) を基に、関係するペプチド特性が生体分子にクラスタ化されます。

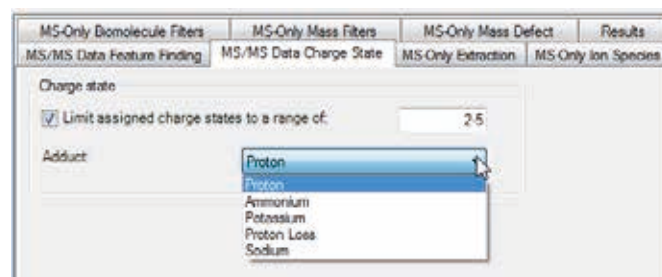


図 5. MS/MS Data Charge State タブで、ペプチド生体分子の発見に使用する荷電状態と付加体を指定できます。

生体分子と MS/MS スペクトルとの関連付け

ペプチドマッピングアルゴリズムは、MS/MS アイソレーション幅とリテンションタイムを使用し、MS1 スペクトルを基に、MS/MS スペクトルと生体分子を関連付けます (図 6)。キメラ (つまり、マルチペプチド) スペクトルの場合、1 つの MS/MS スペクトルが複数の生体分子と関連付けられる場合があります。MS/MS スペクトルは、フラグメンテーションの特異性のために MS1 よりも情報量が豊富で、より高い信頼性のペプチドスペクトルマッチ (PSM) が得られる可能性があります。生体分子が存在する場合、MS/MS スペクトルは生体分子と関連付けられますが、生体分子が存在しない場合、ソフトウェアアルゴリズムはプリカーサ質量を使用して生体分子を作成します。この方法で、すべての MS/MS スキャンが、可能性のある PSM について評価されます。

MS/MS スペクトルに対するペプチド配列のマッチングと検証

ペプチドマッピングのアルゴリズムは、MS/MS スペクトルに対してすべての可能性のある配列を照合し、また、PTM を考慮します。検索時間を短縮するために、コンパクトなデータ構造にいくつかの可能性のある PTM をコード化できるグラフベースのアルゴリズムを使用します²。PSM 候補のリストが作成されると、そのような結果を偶然観測できる確率を示すスペクトルの E-値を使用して、PSM がランク付けされます³。ランク付けは、PSM の生スコアと、可能性のあるすべてのペプチドのスペクトル固有のヒストグラムを比較することによって完了します (図 7)。スペクトル E-値が最高のペプチドが MS/MS スペクトルに割り当てられます。

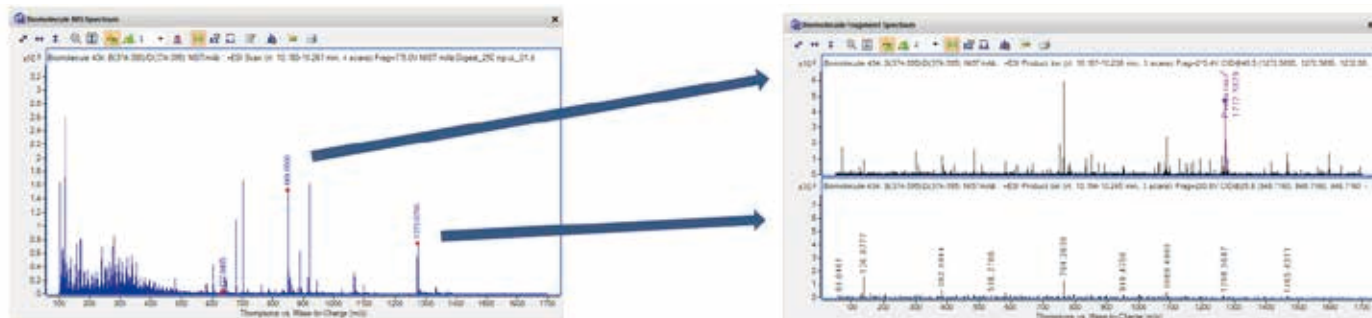


図 6. プリカーサイオンとリテンションタイム範囲が異なる MS/MS スペクトルが生体分子 MS1 スペクトルと関連付けられます。

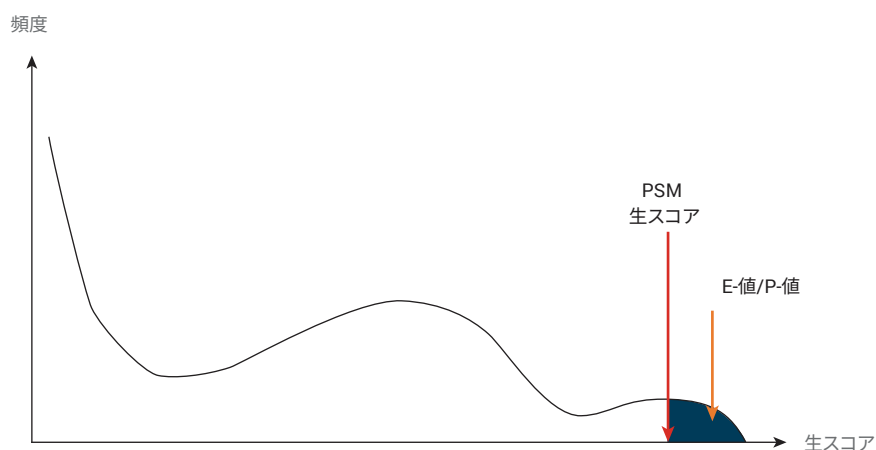


図 7. 生スコアをベースとした配列マッチ候補の頻度のスコアヒストグラムにより、E-値を計算できます。

偽発見率 (FDR) を推定するために、アルゴリズムはターゲットデコイアプローチを使用します⁴。ターゲット配列の逆となるデコイ配列について検索を繰り返します。デコイ配列に対するペプチドスペクトルマッチのランク付け (つまり、アルゴリズムが偽ペプチド配列をヒットするまでの順序) により、FDR が決定します。例えば、ある基準より上に 100 個のターゲットスペクトルマッチと 10 個のデコイスペクトルマッチがある場合、マッチの結果 (つまり、ターゲットスペクトルマッチ) の FDR は 10 % です (10 マッチのうち 1 個が偽陽性)。FDR の閾値は Method Editor の FDR タブで設定され適用され、一部のペプチドスペクトルマッチが捨てられます (図 8)。

配列マッチの結果の視覚化

MS/MS スペクトルには、b、y、インモニウムイオンなどのフラグメント情報が付けられています。下付きの数字 (例: y_6 の 6) は、フラグメントペプチド配列のアミノ酸の数を表します (図 9)。

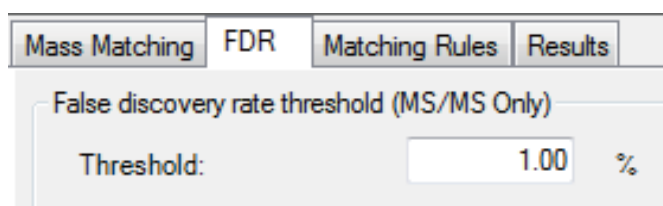


図 8. FDR タブで、偽発見率について閾値を設定できます。

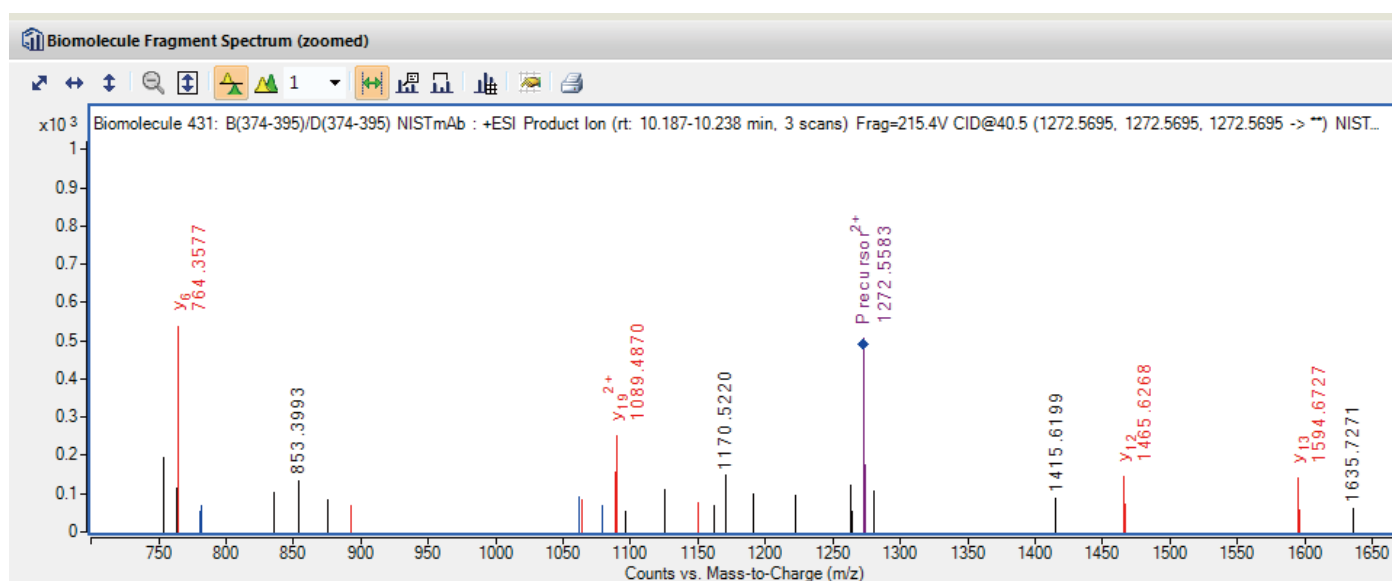


図 9. b (青色) と y (赤色) のイオンが付けられた生体分子の MS/MS 平均スペクトル

ペプチドマッピングの結果を容易に視覚化するために、MS/MS 配列マップはタンパク質配列に対する実線として BioConfirm 配列カバレッジに表示されます (図 10)。緑色で表示されたアミノ酸残基には、ペプチドマッピングからの MS/MS ベースの証拠が含まれています。ユーザーが処理の前に MS のみを基にした特性を表示するように Method Editor で選択した場合、特性は配列上に破線で示されます。Biomolecule テー

ブルの行をクリックすると (図 11) 生体分子と関連付けられたスペクトルが表示され (図 9)、配列カバレッジでは緑色でペプチド配列が強調表示されます (図 10)。



図 10. 配列カバレッジで、MS/MS-ベースのタンパク質配列を確認できます。

Biomolecules: 493 found											
General					Sequence Match						
Label	RT	Height	MS/MS Count		Seq Name	Seq Loc	Tgt Seq Mass	Diff (Bio, ppm)	Sequence	FDR	Score (Bio)
Biomolecule 430	6.284	177507	7								
Biomolecule 431	10.209	250644	6		NISTmAb	B(374-395)/D(374	2543.1241	1.12	GFYPSDIAVEWES	0	74.67
Biomolecule 433	9.856	38831	4								
Biomolecule 434	10.232	9854	2								

図 11. 生体分子テーブルは、各生体分子の情報を表示し、スペクトル表示および配列カバレッジとリンクされます。

Biomolecule テーブルのペプチドマッピングの結果は、0 ～ 100 の範囲の FDR とスコア (Bio) の両方の列を表示でき、フラグメントイオンの数を使用して計算され、連続するイオン、補完的なイオンが考慮されます。これは存在量で重み付けされるため、大きな存在量の質量ピークからのイオンはより大きなスコアに貢献します。

複数のデータファイル、例えば異なる酵素消化物または反復 MS/MS 分析からのデータファイルを同時に検査することが可能です。図 12 は、ロードされた 2 つのデータファイルを用いて、CHO 細胞培養のヒト IgG1 mAb の重鎖 (HC) を配列カバーマップで示しています。最初の採取 (黒色) で 96.5 % の配列が確認され、2 つめの反復 MS/MS でさらに詳細に解析されて、A(294-306) で追加のペプチドが発見され、配列確認が 99.1 % に高まりました。

結論

タンパク質治療薬を正確に分析するという要件を考慮すると、速度と結果の信頼性についてのニーズはきわめて重要なものになります。Agilent MassHunter BioConfirm ペプチドマッピングアルゴリズムでは、抽出できる情報が豊富な MS/MS スペクトルをペプチドの推定に使用しています。PTM およびリンクを含め可能性のあるペプチド配列を効率的にモデル化し、ペプチドの一致を評価して偽陽性の数を最小に抑えるとともに、真の陽性の数を最大にします。アルゴリズムにより、ターゲットデコイメソッドを使用して追加の検証が実行されます。また、偽陽性を示す閾値をユーザーが設定できるほか、ペプチドの一致についての品質を示すスコアが算定されます。最終的には、配列カバーの証拠をスペクトル、表、グラフィック形式で同時にすばやく検査できるように、結果が視覚化されます。

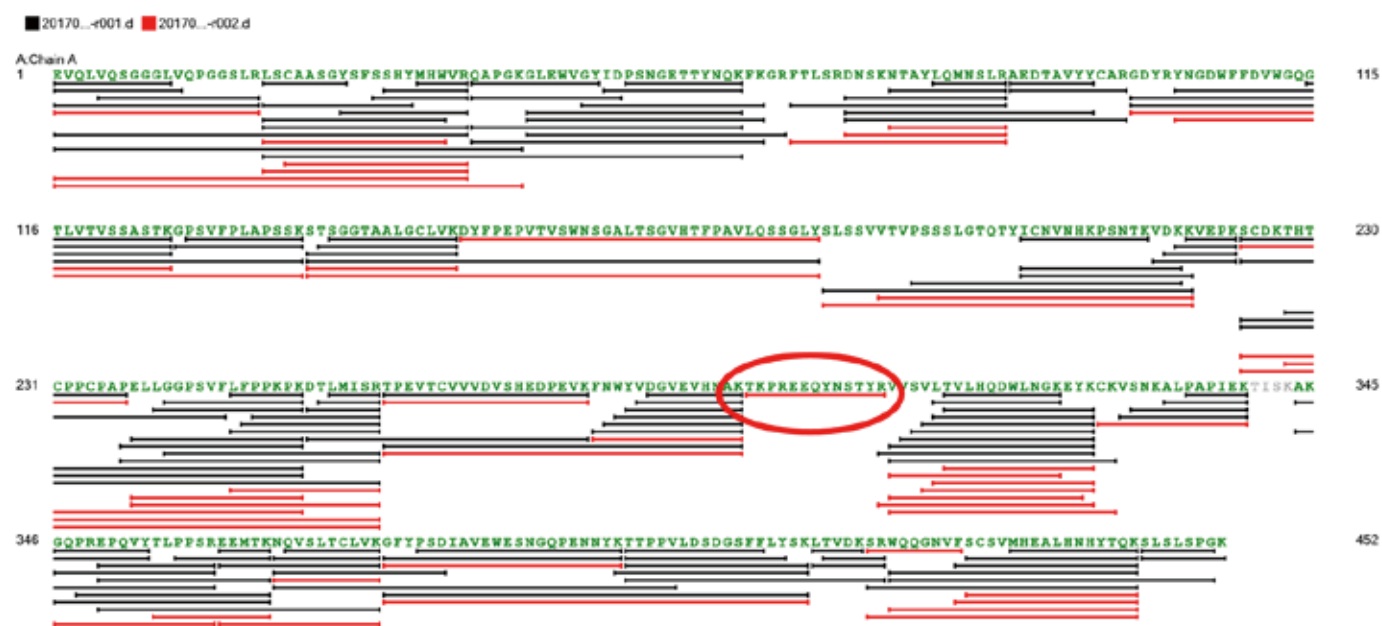


図 12. 複数の反復 MS/MS データファイルがロードされた配列カバーマップ

参考文献

1. Sullivan, et al Drug Development & Delivery November/December 2016 Vol 16 No 9: 44-47.
2. Park, et al.Nat. Methods (2017): doi:10.1038/nmeth.4388
3. Kim, et al.Nat Commun.(2014) 5:5277
4. Elias, et al.Methods Mol Biol.(2010) 604: 55–71.

詳細情報

アジレント製品とサービスの詳細については、アジレントのウェブサイト www.agilent.com/chem/jp をご覧ください。

ホームページ

www.agilent.com/chem/jp

カスタマコンタクトセンタ

0120-477-111

email_japan@agilent.com

本製品は一般的な実験用途での使用を想定しており、医薬品医療機器等法に基づく登録を行っておりません。本文書に記載の情報、説明、製品仕様等は予告なしに変更されることがあります。

アジレント・テクノロジー株式会社

© Agilent Technologies, Inc. 2017

Printed in Japan, September 29, 2017

5991-8552JAJP