Skylineデータ非依存性取得

# データ非依存性取得 (DIA)1,2 とは大規模ターゲットのプロテオミクス実験を実施するための高度なテクニックです。選択反応モニタリング (SRM) および並列反応モニタリング (PRM) などのターゲット取得アプローチでは、取得スケジューリングなしの場合はごく少量のペプチド、スケジューリングしたマススペクトルランでは数十～数百のペプチドに制限されます。DIAでは、SRMと比べて感度、選択性、および再現性をわずかに犠牲にするだけで、より多数のペプチド（数千またはプロテオーム全体）の測定が可能です。またDIAは、測定するペプチドを事前に指定またはスケジュールする必要がなく、むしろDIAランから取得後に広範囲のプレカーサーm/z内で任意のペプチドのプロダクトイオンクロマトグラムを抽出することが可能であるという利点があります。

DIAデータからのクロマトグラム抽出をサポートする機能が2010年10月に初めてSkylineに実装されて以来、このサポートは定期的に改善され、バージョン21.1のSkylineではDIAデータ解析のための一般的な複数のストラテジーおよびワークフローがサポートされています。また、SkylineではSCIEX、Agilent、Bruker、およびWaters各社のQ-TOFや、Thermo社のQ-Orbitrap装置など、すべてのDIA対応装置がサポートされています。

特定のタイプの装置およびクロマトグラフィー設定では、1つの有効なDIAワークフローで任意の数の初期データ依存性取得 (DDA) ランを開始します。DIAを実行するのと同じ装置でこれらの初期ショットガン測定を実行することは有益です。しかし、同様のフラグメンテーションテクニックとクロマトグラフィーが使用されている場合、装置プラットフォームの間でターゲットアッセイが移動する可能性もあります。これらの初期DDA測定では、試料を分割するか、そうでなければ簡素化して高いプロテオームカバレッジを実現できます。DDAランはペプチド-スペクトルマッチングパイプラインで処理され、その結果のペプチドID、スペクトルおよび保持時間が使用されてスペクトルライブラリおよび保持時間 (iRT) ライブラリ（Skylineの場合）、または（フラグメントイオンのサブセットについて同様の情報を伴う）「アッセイライブラリ」と呼ばれる拡張トランジションリスト（他のツールの場合）が作成されます。これらのフラグメントイオン相対存在量および正規化保持時間 (iRT) のライブラリは、保持時間のアラインメントのために同じ装置と標準ペプチドを用いる後続のDIAランで何度も使用することができます。

DDA検索結果を、このライブラリDIA分析アプローチに適したライブラリに変換するメソッドは数多く存在し（[Skylineチュートリアルウェビナー#2](https://skyline.gs.washington.edu/labkey/project/home/software/Skyline/events/2014%20Webinars/Webinar%202/begin.view)の「以前のナレッジワークフロー」で説明され、以降はウェビナー[#14](https://skyline.ms/webinar14.url)、[#15](https://skyline.ms/webinar15.url)、[#18](https://skyline.ms/webinar18.url)およびチュートリアル[DIA/SWATHデータの分析](https://skyline.ms/tutorial_dia_swath.url)でより詳しく文書化）、DIA分析を開始する最も直接的でアクセスしやすい方法は、同じ装置にDIAランとDDAランを組み込み、DDA結果のスペクトルおよび保持時間をDIAランの予測として使用することです。

このチュートリアルでは、Skylineで簡単な散在DDA/DIAアプローチを使用してDIAランを設定、インポート、およびプロセスする方法を学習します。研究でDIAを使用する場合には、上記のウェビナーやチュートリアルで説明されているより強力で複雑な方法に進むことを強くお勧めします。

# はじめに

チュートリアルを始める前に、以下のzipファイルをダウンロードしてください。

<https://skyline.ms/tutorials/DIA-20_2.zip>

ファイルのサイズが膨大であることにご注意ください（ダウンロードファイルは4.5 GB、解凍後は6.0 GB）。これは一般的に非常に大きなファイルであるDIAランを含んでいるためです（時にSRMランの100～200倍のサイズとなります）。このダウンロードに時間がかかりすぎるか十分なディスク容量がない場合、代わりに以下からより小さなチュートリアルバージョンのファイルをダウンロードできます（ダウンロードファイルは660 MB、解凍後は918 MB）。

<https://skyline.ms/tutorials/DIALibrary-20_2.zip>

小さなバージョンを選択した場合、マススペクトルの生データファイルが含まれていないため、チュートリアルのいくつかの手順を省略する必要があります。下のテキストが表示されます。どちらのZIPファイルを選択しても、次の手順は以下のようなコンピュータのフォルダにファイルを解凍することです。

C:\Users\damodei\Documents

これにより以下の新しいフォルダが作成されます。

C:\Users\damodei\Documents\DIA-20\_2

本チュートリアルを始める前にSkylineを使用していた場合には、Skylineをデフォルト設定に戻すことをお勧めします。デフォルト設定に戻すには、以下の操作を行います。

* Skylineを起動します。
* **開始ページ**で、以下のような**空のドキュメント**をクリックします。

Graphical user interface, application, Word

Description automatically generated

* [ **設定** ] メニューで、[ **デフォルト** ] をクリックします。
* 現在の設定を保存するかどうかを尋ねるフォームで [ **いいえ** ] をクリックします。

Skylineのこのインスタンスのドキュメント設定がデフォルトにリセットされました。

本チュートリアルはプロテオミクスに関するものであるため、以下の操作を行うとプロテオミクス用インターフェイスを選択できます。

* Skylineウィンドウの右上隅にあるユーザーインターフェイス管理をクリックし、[ **プロテオミクス用インターフェイス** ] をクリックします。

Text

Description automatically generated

Skylineは、ウィンドウの右上隅のタンパク質アイコン で表示される分子モードで動作しています。

DIA分析の場合は、Skylineを強制して抽出された定量トランジションすべてのピーク面積を積分します。このためには、以下のように操作します。

* [ **設定** ] メニューで [ **すべて積分** ] をクリックします。

# DIAの単離スキームの設定

デフォルト設定のみがある新しい空のドキュメントでSkylineが表示されます。SkylineでDIAデータを分析する完全なプロセスを説明するために、 必要な設定、トランジション、スペクトルライブラリ、および保持時間の情報を入力して最初からSkylineドキュメントを構築します。

散在するDDAワークフローを伴うDIAを使用して実験を行う場合、カバーされるプリカーサーm/z範囲にターゲットが含まれているかを確認するなど、関心があるペプチドターゲットをごく一般的に考慮するだけで、最初にDIAおよびDDAランの両方向けに装置を設定できます。DDAメソッドも残されていますが、Skylineはユーザーが「単離スキーム」を定義できるようにすることで、DIAメソッドの設定を支援します（MS/MSのフラグメントに対するプリカーサー単離ウィンドウのパターン）。たとえDIAデータを収集済みであっても、SkylineがユーザーのDIAランをプロセスできるようにするには、使用した単離スキームを定義します。このチュートリアル実験の単離スキームを定義するには、以下の手順に従います。

* [ **設定** ] メニューで [ **トランジション設定** ] をクリックします。
* [ **フルスキャン** ] タブをクリックします。
* [ **MS1フィルタ** ] セクションの [ **含まれる同位体ピーク** ] ドロップダウンリストで、「数」を選択します。
* [ **ピーク** ] フィールドに「3」と入力します。
* [ **MS1フィルタ** ] セクションの [ **取得メソッド** ] ドロップダウンリストで「DIA」を選択し、DIAデータをインポートすることをSkylineに伝えます。

多くのトリプシンペプチドでは、最初の3つの同位体ピークが最も強くなります。ベース（最も強い）同位体ピークのパーセンテージに基づき強度閾値を使用することもできますが、これらの設定はトリプシンペプチドにとって合理的なデフォルトです。

* [ **プリカーサー質量分析** ] および [ **プロダクト質量分析** ] を「Centroided」に設定します。

|  |
| --- |
| **注：**このデータはQ-Exactiveで収集され、Orbitrapを使用してMS1およびMS2スキャンが実行されていますが、質量分析計を「Orbitrap」として指定した場合には、セントロイド化されたスペクトルからクロマトグラムを抽出すると、ネイティブプロファイルスペクトルを使用するよりも優れた結果が得られることがわかりました。 |

* [ **MS1フィルタ** ] および [ **MS/MSフィルタ** ] 両者の [ **質量精度** ] フィールドに「20」ppmと入力します。

[ **トランジションの設定** ] フォームは以下のようになります。

Graphical user interface, application

Description automatically generated

|  |
| --- |
| **注：**[ **保持時間のフィルタ** ] オプションは、[ **MS/MS IDのスキャンの中で** [5] **分前後のスキャンのもののみを使用** ] のままにしておいてください。これによって、Skylineはライブラリ内で見つかったペプチドスペクトル一致から5分以内で取得されたDIAスペクトルからの強度のみを抽出します。単一のペプチドスペクトル一致の場合、総抽出ウィンドウは10分です。所与のペプチドに対して1つ以上のIDがある場合、Skylineは最小ID時間マイナス5分と最大ID時間プラス5分の範囲で取得されたスペクトルから抽出します。（後で説明しますが、すべてのID保持時間は構築した非冗長ライブラリ内にも保存されます。) 2つ目のオプションである [ **予測RTの中で** [5] **分前後のスキャンのもののみを使用** ] では、RT予測（たとえばSSRCalc、または一般的にはiRTライブラリ）を使用して抽出時間範囲を決定します。本チュートリアルではRT予測は使用しませんが、これは前記の他のチュートリアルやウェビナー、そして[アッセイのインポート](https://skyline.gs.washington.edu/labkey/_webdav/home/software/Skyline/%40files/tutorials/ImportingAssayLibraries-2_6.pdf)のヒント（Skylineのウェブサイトの [ [ヒント](https://skyline.gs.washington.edu/labkey/wiki/home/software/Skyline/page.view?name=tips) ] > [ [他の定量ツールでの作業](https://skyline.gs.washington.edu/labkey/wiki/home/software/Skyline/page.view?name=other_tools) ]）でも説明しています**。** |

これで、基本的なフルスキャン装置パラメータの設定が完了しました。次に、DIA単離スキーム、またはDIAを実行する際に装置がサイクルするプリカーサー*m/z*範囲のパターンを指定します。例えば、このチュートリアルのデータセットでは、Q Exactive装置は、500～520*m/z*からプリカーサー *m/z* 範囲を開始し、次に520～540 *m/z*、それから880～900 *m/z*まで増加し（あるいは500から900 *m/z*まで連続する20個の20 m/z ウィンドウ）、その後このサイクルを繰り返します。Skylineでこの単離スキームを指定するには、以下の手順に従います。

* [ **単離スキーム** ] ドロップダウンメニューで、**<追加...>**を選択します。

次のような [ **単離スキームを編集** ] フォームが表示されます。

Graphical user interface, application

Description automatically generated

* [ **事前に指定した単離ウィンドウ** ] を選択します。

ここで、単離ウィンドウを指定できるグリッドが有効になります。手動でグリッドにウィンドウ境界を入力することもできますが、この場合、ウィンドウ境界のサイクルは規則性が高いため（20 *m/z*の増分で500～900 *m/z*）となります。次のように素早く境界を指定する方法があります。

* [ **計算** ] ボタンをクリックして [ **単離スキームを計算** ] フォームを表示します。
* [ **m/zの開始** ] フィールドに「500」と入力します。
* [ **m/zの終了** ] フィールドに「900」と入力します。
* [ **ウィンドウ幅** ] フィールドに「20」と入力します。
* [ **ウィンドウ位置の最適化** ] チェックボックスをオンにします。

これでウィンドウ境界が1価および2価のペプチドプリカーサーが発生するところから約0.25 m/zのところに最適化されます。3価および4価のプリカーサーの中には、これらの境界に近いものもある可能性がありますが、20 m/zのウィンドウであればこのようなことが頻繁に起きる可能性は低く、Q1単離が境界で合理的に効率的であればSWATH2に関するオリジナルの論文で提案されているようなウィンドウのマージンをオーバーラップさせる（各側に0.5 m/z）必要が少なくなります3。



**図：**観測されるペプチドの共通ライブラリで見られる1価から4価のペプチドプリカーサーから計算したヒストグラム。一番高いピークは1価から4価のプリカーサーを表し、2番目に高いピークは2価と4価のプリカーサー、そしてさらに小さなピークが3価のみのプリカーサーを表し、4価のみのプリカーサーは3価のみのピークの小さな肩程度にしか現れません。「最適」境界と記されている縦の線は、この範囲での最適化ウィンドウ配置境界を示しています。（Brian Searleのご厚意により提供）



Gillet, et al. MCP 2012よりの補足図S1：c) は定義された範囲の完璧な長方形単離を示し、d) は定義された四重極単離範囲のエッジでの0.5 m/zにわたるあまり理想的ではない信号ドロップオフを示し、e) とf) はマージンがオーバーラップしない場合とする場合とで完全な同位体分布がどのように分割されるかを示しています。同位体分布全体はターゲットとなっているモノアイソトピックm/zのみを通じて単離されているため、モノアイソトピックプリカーサーm/zが単離ウィンドウのエッジに非常に近いように見える場合でも強いフラグメントイオン信号が得られることを覚えておくことが重要です。

作業中のDIAデータが既に取得済みである場合、現段階で定義された単離スキームが取得に使用される装置の設定を反映することが重要です。この場合、通常は最も安全な状態であるため、[ **インポート** ] ボタンをクリックしてSkylineを既存のデータファイルの1つにポイントします。

[ **単離スキームを計算** ] フォームは以下のようになります。

Graphical user interface, application

Description automatically generated

* [ **OK** ] ボタンをクリックします。

Skylineは、500～900 *m/z*の範囲を20 *m/z*間隔でカバーするために必要な20個のウィンドウの境界を自動入力します。

[ **単離スキームを編集** ] フォームは以下のようになります。

Table

Description automatically generated

また、Skylineでは、入力した内容が正しいか確認するために、プリカーサー*m/z*範囲の単離を経時的に視覚化できます。

* 先ほどクリックした [ **計算** ] ボタンの横にある [ **グラフ** ] ボタンをクリックします。

Y軸をサイクル/時間、X軸を*m/z*にした、経時的な単離ウィンドウのサイクルのグラフが表示されるはずです。

Chart, line chart

Description automatically generated

* [ **閉じる** ] ボタンをクリックします。
* [ **名前** ] フィールドに、この単離スキームの名前として「500 to 900 by 20」と入力します。
* [ **単離スキームを編集** ] フォームで [ **OK** ] ボタンをクリックします。
* [ **トランジションの設定** ] フォームで [ **OK** ] ボタンをクリックします。

DIAでは、装置は測定するために特定のトランジションセット (SRM) やプリカーサー (PRM) を必要と*しません*。そのため、空のターゲットであるにもかかわらずユーザーのドキュメントはここでDIAランを設定するために必要な情報すべてを保持しています。DIA単離スキームは以下のようにして装置にエクスポートできます。

* [ **ファイル** ] メニューで、[ **エクスポート** ] を選択し [ **単離リスト** ] をクリックします。

[ **単離リストをエクスポート** ] フォームが表示され、単離リストのエクスポート形式が選択できます。

* [ **装置タイプ** ] ドロップダウンリストから**Thermo Q Exactive**を選択します。

フォームは以下のようになります。

Graphical user interface, application

Description automatically generated

* [ **OK** ] ボタンをクリックします。
* 表示された保存フォームで、このチュートリアル用に作成したフォルダに移動します。
* [ **ファイル名** ] フィールドに、「DIA\_tutorial\_isolation\_list.csv」と入力します。
* [ **保存** ] ボタンをクリックします。

保存したファイルを開きます。これは以下のようになります。



この単離スキームは、（このチュートリアルのデータが取得された）Thermo Q Exactive用にフォーマットされていますが、Skylineでは他の複数の装置タイプにエクスポートすることも可能です。

これが実際のDIA実験であるならば、この単離リストファイルを使用してユーザーの装置でDIA取得を実行することが妥当といえます。代わりに、装置ソフトウェアで単に単離スキームを手動で指定することもできます。データ取得の他のメソッドパラメータ（MS/MS単離幅および分解能など）は、メソッドファイルに手動で設定する必要があります。

# ペプチド検索のインポートウィザードの使用

本チュートリアルでは、実際に実験のすべてのDDAおよびDIAランが完了していますので、そのように仮定します。Peptide ProphetおよびTrans Proteomic Pipeline (TPP) を実行した場合は、X! Tandemのような検索エンジンを使用してDDAランからのMS/MSスペクトルをペプチドに一致させてデータ分析ワークフローを開始すると、一連の.xtan.xmlファイルまたは .pep.xmlファイルが生成されます。本チュートリアルでは、DIA-20\_2.zip（DIALibrary-20\_2.zipではありません）をダウンロードした場合はDDAランの単一の.pep.xmlファイルと、オリジナルの生DDAデータファイルの.mzXMLファイル変換 (804 MB) が提供されます。Skylineドキュメントを関連DIAラン分析用に設定する最初のステップは、この検索結果をSkylineにインポートし、MS/MSスペクトルと取得時の保持時間を含むスペクトルライブラリを作成することです。DDA検索結果をインポートするには、以下の操作を実行します。

* [ **ファイル** ] メニューで、[ **インポート** ] を選択して [ **ペプチド検索** ] をクリックします。

コンピュータディスクドライブのどこで分析を実施するのかがわかるよう、現在のドキュメントを保存する必要があるというメッセージが表示されます。

* [ **OK** ] ボタンをクリックします。
* 本チュートリアル用に作成したフォルダに移動します。
* [ **ファイル名**] フィールドに、「DIA-tutorial」と入力します。
* [ **保存** ] ボタンをクリックします。

以下のような**ペプチド検索のインポート**ウィザードが表示されます。

Graphical user interface, text, application

Description automatically generated

このデータセットでは、[ **カットオフスコア** ] フィールドが「0.95」に設定されており、PeptideProphetスコアが0.95以上のペプチドスペクトル一致が含まれることを意味します。これはDDAデータがTPPで処理されたためです（SEQUESTによるペプチドスペクトル一致の後）。q値またはゼロが最高で1が最低である期待スコアを使用する他のスペクトル一致パイプラインの場合は、1 – スコアのカットオフスコアを使用します。したがって、0.95は≤0.05となります。再利用可能なライブラリの場合は、0.99などのもっと厳密なカットオフを使用します（≤0.01またはq値の偽発見率1%）。

* [ **ファイルを追加** ] ボタンをクリックします。
* 表示される [ **入力ファイルを追加** ] フォームで、チュートリアルが入っているディレクトリから以下のファイルを選択します。
  + interact-20130311\_DDA\_Pit01.pep.xml

このファイルには、1回のDDAランからのペプチドスペクトル一致結果が含まれています。実際の実験では、質量分析計で実際にDDA取得を行い、それから検索エンジンで出力ファイルを実行し、このようなファイルを1つ以上生成します（通常はTPPが生成するpepXMLに対して1つ）。ここでは、ファイルが提供されています。元のDDAランデータファイル（mzXMLに変換）であるinteract-20130311\_DDA\_Pit01.mzXMLも同じフォルダにありますので注意してください。このファイルをクロマトグラム抽出用にインポートする必要はありませんが、ライブラリビルダーが.pep.xmlファイルには存在しない、そのライブラリのMS/MSスペクトルを検出するためにこのファイルが存在している必要があります。Mascot DATファイル、Proteome Discoverer pdResultファイル、X! TandemネイティブXMLファイルなどのその他のスペクトル一致パイプラインの出力では、1つの出力ファイルに必要な情報すべてが含まれています。Skylineウェブサイトで[対応するツールおよびファイル形式の全リスト](https://skyline.ms/build-blib.url)をご覧ください。

* [ **開く** ] をクリックしてこのファイルを構築しているライブラリに追加します。
* [ **ワークフロー** ] ボックスで [ **DIA** ] オプションをクリックします。

ウィザードの [ **スペクトルライブラリ** ] ページは以下のようになります。

Graphical user interface, text, application

Description automatically generated

* [ **次へ** ] をクリックしてこれらの検索結果からスペクトルライブラリを構築します。

Skylineがスペクトルライブラリの構築を終了すると、DDAランからのMS/MSスペクトルに一致するすべてのペプチドとその保持時間をSkylineドキュメントに含められるようになります。まずは、[ **クロマトグラムを抽出** ] ページが表示されます。このページでは、クロマトグラムの抽出、ピーク検出、ピーク面積計算に使用するDIAデータファイルがどこにあるかを指定できます。

* [ **参照** ] ボタンをクリックします。
* DIAフォルダで、このチュートリアルと共に含まれている.rawファイル両方を選択します。

[ **結果ファイルを参照** ] フォームは以下のようになります。

Graphical user interface, text, application

Description automatically generated

* [ **開く** ] ボタンをクリックします。

**ペプチド検索のインポート**フォームは以下のようになります。

Graphical user interface, text, application

Description automatically generated

* [ **次へ** ] ボタンをクリックします。

Skylineは、Skylineインターフェイスで表示される繰り返し測定名を短くし、一般的に作業しやすくなるよう、共通プリフィックスである「20130311\_DIA\_Pit0」を削除するかどうかをユーザーに尋ねます。ただし、この場合、名前は「1」や「2」まで短縮されます。

* [ **共通プリフィックス** ] フィールドで、サフィックス「Pit0」を削除します。

[ **結果をインポート** ] フォームは以下のようになります。

Graphical user interface, text, application

Description automatically generated

* [ **OK** ] ボタンをクリックします。

[ **修飾を追加** ] ページが表示され、スペクトルライブラリには、Unimod修飾である「酸化(M)」に一致したM[+16]があるペプチドシークエンスが含まれていると表示されます。この種の修飾をターゲットリストに許可するには修飾名の横にあるチェックボックスをオンにするだけですが、このチュートリアルでは行いません。

* [ **次へ** ] ボタンをクリックします。

以下のような [ **トランジション設定を行う** ] ページに移動します。

Graphical user interface, application, Word

Description automatically generated

このチュートリアルのデフォルト設定を使用することもできますが、今後の自分のデータに合った選択肢を考えるのもいいでしょう。

* プリカーサーイオン2価と3価は、通常最も一般的で安全です。
* イオン電荷とタイプは、yイオンとbイオン、1価と2価、そしてMS1スペクトルから抽出されたプリカーサーイオンです。
* プロダクトイオンの範囲は、選択性が低いために一般的には除外されるy1、y2、b1、b2以外のすべてです。
* 最小および最大m/z範囲は、MS/MSスペクトルのみに言及します。これは許容されるプリカーサーm/z範囲がDIA単離スキームによって定義されるためです。
* DIAプリカーサーウィンドウ範囲を除外するオプションは、SkylineにDIA単離ウィンドウに入るトランジションを含まないように指示します（たとえばプリカーサーが500から520 *m/z*のDIAウィンドウに入るのであれば513 *m/z*にあるプロダクトイオンは含まれません）。フラグメンテーションをすり抜けてしまうプリカーサーイオンは、元の単離範囲にあるMS/MSスペクトル中でノイズや場合によっては干渉を増加させる可能性があり、この範囲にあるフラグメントイオンに基づいた定量化の信頼性が低くなるため、これによる除外が好ましいこともあります。
* イオン許容誤差0.05は、DDAライブラリスペクトルをフラグメントイオンに一致させるときに使用される許容誤差であり、ライブラリ強度を割り当てます。これは大雑把に言って50 ppm（1000 *m/z*にて）であり、おそらくOrbitrap MS/MSスペクトルの場合にはもう少し広くなります。
* 6個のフラグメントイオンを選択し、6個のフラグメントイオンを必要とするオプションは、すべてのターゲットで、6個以上のイオンが一致しないスペクトルがあるちょうど6個のトランジションとペプチドが除外されることを意味します。

次のページに移動するには、以下の操作を行います。

* [ **次へ** ] ボタンをクリックします。

[ **フルスキャン設定を行う** ] ページに移動します。このページには、先ほど行った設定と同じものが表示されますが、以下のようになっていることを確認してください。

Graphical user interface, application

Description automatically generated

* [ **次へ** ] ボタンをクリックします。

[ **FASTAをインポート** ] ページに移動します。このページでは、ペプチドへの消化に使用したプロテアーゼ酵素と共にターゲットにしたいタンパク質をFASTA形式で指定するように要求されます。この酵素は、最大未切断数0で、デフォルトのTrypsinのままにしておいてかまいません。

* [ **参照** ] ボタンをクリックします。
* [ **FASTAを開く** ] フォームで、ここまで作業をしてきたチュートリアルフォルダに移動し、「pituitary\_database.fasta」という名前のファイルをダブルクリックします。

フォームは以下のようになります。

Text

Description automatically generated

今FASTAテキストに6個のターゲットタンパク質だけを含めたこの簡単な導入では、デコイやmProphetアルゴリズムで生成された動的ピーク選択モデルは使用しません。これらのオプションについての詳細は、[DIA/SWATHデータの分析](https://skyline.ms/tutorial_dia_swath.url)のチュートリアルを参照してください。

* [ **完了** ] ボタンをクリックします。

Skylineがタンパク質を6個、ペプチドを21個追加するというフォームを表示します。これによって25個のプリカーサーと225個のトランジションが対象となります。どのタンパク質でも、スペクトルライブラリに存在していたペプチドのみがこのリストに含まれます。

Graphical user interface, text

Description automatically generated

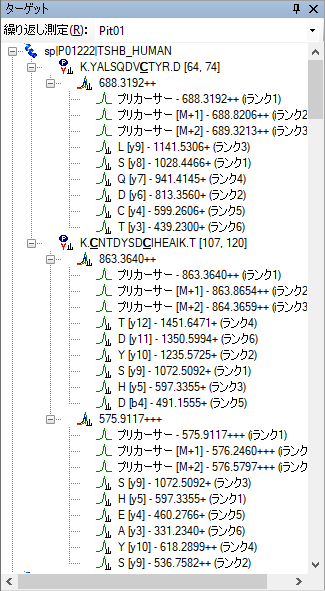
* [ **OK** ] ボタンをクリックします。

タンパク質とペプチドが [ **ターゲット** ] ペインに表示され、選択した生データファイルからクロマトグラム抽出が自動的に開始します。以下に進む前に、進行状況ウィンドウを移動させたり閉じたりできます。

* [ **編集** ] メニューで [ **すべて展開** ] を選択して、[ **プリカーサー** ] をクリックします。

これによってSkylineが含めたトランジションセットすべてが表示されます。すべてのペプチドには500と900 m/zの間にある対象プリカーサーすべてが含まれ、すべてのプリカーサーには6個のbトランジションとyトランジションの他、3個のプリカーサーイオン (M、M+1、M+2) が含まれていることに注意してください。また、すべてのプリカーサーには該当するライブラリスペクトルがあることにも注意してください。各トランジションのテキストは、「rank」または「irank」番号で終わっています。これは、フラグメントイオンのライブラリスペクトルまたはプリカーサー同位体イオンの同位体分布内での強度ランキングを示します。

[ **ターゲット** ] ビューは以下のようになります。



DIAでは、装置が500-900 *m/z*の範囲にある考えられるすべてのプリカーサーとプロダクトの組み合わせを網羅しているため、あらゆるセットのプロダクトイオンを抽出できます（少なくともMS/MSスペクトルの測定範囲内にあるもの）。しかしながら、実際には考えられるすべてのプロダクトイオンを抽出する必要はないことが多く、事実、ノイズが多く、必要とするペプチドの信号が非常に低いクロマトグラムを追加してペプチドの検出を妨げる場合があります。最も強度が高い6個のトランジション(y1、y2、b1、b2を含まない)を必要とするように設定するプロダクトイオンの選択規則は、プロテオミクスコミュニティにおいて広く受け入れられるようになりました。

大規模なDIA実験では、高度なピーク選択（ペプチド検出）モデルをトレーニングするデコイペプチド6があると有用な場合があります。これによってSkyline自動ピーク選択が向上し、Skylineが検出確率スコア（q値）を生成できるようになります。デコイペプチドは厳密にはDIAワークフローには必要ではなく、カスタムピークスコアメソッドを適用する場合にのみ必要となります。多くの場合、Skylineデフォルトピーク選択が十分にうまく機能するので、カスタムピークスコアは必要ありません。カスタムピークスコアを使用する場合でも、デコイに代わるものはあります（[高度なピーク選択モデル](https://skyline.gs.washington.edu/labkey/wiki/home/software/Skyline/page.view?name=tutorial_peak_picking)のチュートリアルを参照）。本チュートリアルでは、ドキュメントにデコイを追加しません。

# DIA結果の調査

DIAランがすべてインポートされたら、結果を見ることができます。

* クロマトグラムペインで「Pit01」と言うタブが選択されていることを確認します。
* [ **編集** ] メニューで、[ **すべて折り畳む** ] を選択して [ **ペプチド** ] をクリックします。
* [ **ビュー** ] メニューで [ **トランジション** ] を選択し、[ **グラフの分割** ] をクリックします。
* [ **ビュー** ] メニューで、 [ **自動ズーム** ] を選択し [ **最適ピーク** ] (F11) をクリックします。
* [ **ビュー** ] メニューで、[ **グラフを配置** ] を選択して [ **行** ] をクリックします。
* [ **ビュー** ] メニューで [ **ピーク面積** ] を選択し、[ **繰り返し測定の比較** ] ( F7) をクリックします。
* クロマトグラムグラフの1つを右クリックし、[ **凡例** ] をクリックしてチェックをオフにします。
* [ **ピーク面積－繰り返し測定の比較** ] グラフを [ **ライブラリ一致** ] ビューの上にあるメインウィンドウの右端にドッキングします。
* [ **ターゲット** ] ビュー内の最初のペプチドK.YALSQDV**C**TYR.D [64, 74]を選択します。

Skylineウィンドウは以下のようになります。

Graphical user interface, application, histogram

Description automatically generated

これはここから始めるのに適した例ではありません。実際よりもずっと悪いように見えてしまいます。人間の観察者として、ユーザーはピーク積分境界に調整が必要なことがすぐにわかります。これはおそらくSkylineがフラグメントのクロマトグラムに基づいて積分境界を選択するためだと思われます。よく見てみると、信号がX軸まで落ちているところがあるのがわかります。セントロイド化されたデータでは、これは通常、先ほど+/- 20 ppmに設定した抽出ウィンドウ外へのピークのシフトによるものです。

このような信号ドロップアウトのエビデンスが見られる場合は、抽出設定の幅が狭すぎるのではないかと考えて当然です。Skylineでそれを確認する最良の方法は、以下の操作を行うことです。

* [ **ビュー** ] メニューで [ **質量誤差** ] を選択し、[ **ヒストグラム** ] をクリックします。
* 表示されるグラフ内を右クリックし、[ **トランジション** ] を選択して [ **すべて** ] をクリックします（チェックマークをオンにします）。

もっと多くのターゲットがあるドキュメントでは、各ペプチドプリカーサーに対して強度が最も高いトランジションのみを使用して質量誤差分布を把握します。この2番目のステップは不要かも知れませんが、この場合は分布がわかりやすくなり、以下のようになります。



経験則として、以下を確認します。

1. 平均 (2.7) がゼロから遠くない。
2. 設定した許容誤差 (20 ppm) が平均+3\*標準偏差 (12.6) 以上（標準偏差は、正規分布値の99.8%が収まると予想される値）。

この規則によって、インポートされるファイルに入っている質量スペクトルにとって、15 ppmと言う選択はおそらく悪くないでしょう。

* グラフ内を右クリックして [ **トランジション** ] を選択し、[ **プリカーサー** ] をクリックします。

平均値 (2.9) と標準偏差 (2.1) により、9.2 ppmが十分であることがわかります。抽出されたクロマトグラムで見られる信号ドロップアウトは、20 ppmの質量許容誤差が狭すぎるためではないと考えるのにはこれで十分です。この異常を生じている原因については、後ほどまた説明します。

ここでは、以下の操作を行います。

* [ **質量誤差－ヒストグラム** ] ウィンドウの右上角にあるXをクリックしてウィンドウを閉じます。
* プリカーサーやフラグメントの場合は、いずれかのX軸のすぐ下をピークの片側から反対側までクリックアンドドラッグして両ファイル内の積分範囲を調整し、ピーク全体が収まるようにします（積分境界を調整する前に、マウスホイールを使用するとズームを調整できます）。
* クロマトグラムを右クリックし、[ **ペプチドが同定された回数** ] をクリックして [ **他の実行から** ] を選択します。

この最後のステップで、スペクトルライブラリのペプチドスペクトル一致の保持時間のところで、クロマトグラムグラフに青緑色の縦線が追加されます。これはクロマトグラムの+/- 5分を抽出することにしたおよその回数です。DIAおよびDDAランがよく似たクロマトグラフィーを示す場合は、選択されたピークの近くにこのような線が表示されるはずです。

これでSkylineウィンドウは以下のようになります。

Graphical user interface, histogram

Description automatically generated

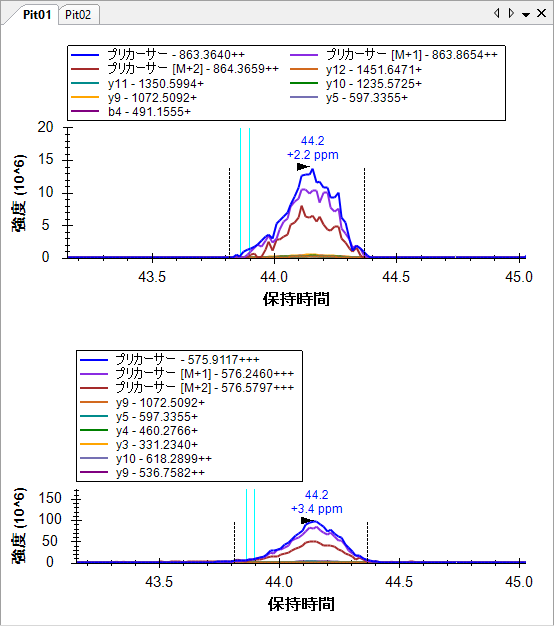
ここでは以下に注意します。

1. クロマトグラムグラフは、まだ元の積分範囲を紫色で強調表示しています。
2. クロマトグラムグラフは、45.7分で見られる4つのピークすべての頂点に注釈を付けます。
3. ピークには、2.3から6.4 ppmの質量誤差値で注釈が付いています。フラグメントイオンの平均質量誤差値は2.7、プリカーサーイオンは2.9であることがわかっているため、すべてがその平均から4 ppm未満であると言えます。
4. プリカーサーの**ピーク面積**プロットでは、「idotp 0.97」で注釈が付いていることがわかります。これは一番左の棒で示される予想される同位体分布に対する類似度であり、0.97はかなりいい値です。
5. フラグメントの**ピーク面積**プロットでは、「dotp 0.88」と「dotp 0.95」で注釈が付いています。これは一番左の棒で示されるライブラリスペクトルに対する類似度であり、0.88は受け入れられる値、0.95はかなりいい値です。

今度は次のペプチドに進みます。

* K.**C**NTDYSD**C**IHEAIK.T [107, 120]をクリックするか、下向き矢印キーを押します （[ **ターゲット** ] ビューが青の場合）。

荷電状態が複数ある場合には、トランジション分割表示が電荷状態で分割され、イオンタイプでは分割されないことがわかります。また、MS1スキャンから抽出されたプリカーサーイオン強度は、通常MS/MSスキャンから抽出されたプロダクトイオン強度を小さく見せることもわかります。



|  |
| --- |
| **注**：明確にするため、凡例は上記画像に合わせて復元され、プロットはサイズ変更されています。 |

* ペプチドの左側にある [ + ] をクリックするか、右矢印キーを押します。
* 863.3640++プリカーサーをクリックするか、下向き矢印キーを押します。

これで2つ目のペプチドの2価のプリカーサーが選択され、Skylineウィンドウは以下のようになります。

Graphical user interface

Description automatically generated with medium confidence

最後のペプチド、質量誤差、idotp、dotpに関して導入されたすべての測定は、このプリカーサーでは向上しているように見えます。特に「Pit02」内のピークに対しては、別の積分境界を選択したいと思うかも知れません。別の積分境界を選択し、その後Skylineの [ 元に戻す/やり直し] ボタンを使用することで、それがピーク面積をあまり変えないことを確認できます。最初のペプチドのように積分境界がピークの大部分を逃していない限り、通常はこのようなわずかな違いは許容する方がいいでしょう。

* 575.9117+++プリカーサーをクリックするか、下向き矢印キーを押します。

これは約7倍強度の高いプリカーサーです。また、かなりいいように見えますが、y5クロマトグラムに干渉があることに気付かれたかもしれません。これをもう少し明確にするために、以下の操作を行います。

* プリカーサーの左側にある [ + ] をクリックするか、右矢印キーを押します。
* プリカーサーの下にあるH [y5] - 597.3355+トランジションをクリックします。

クロマトグラムは以下のようになります。

y5クロマトグラムが赤で強調表示され、積分境界間にある点がX軸の上にある赤の点線で強調表示されています。その右側には、カウント（17と20）が注記されています。また、右側にはPit01で45分、Pit02で44.6分のところに2つ目のピークがあります。どちらも44.1分のところにあるピークにはあまり寄与していないように見えますが、大量の測定を予定していた厳密な定量分析では、やはりこのトランジションを削除し、対象としているピークとおそらくもっと高強度で共溶出しないようにすることも可能です。

次のペプチドに進みます。

* K.ELVYETVR.V [72, 79]ペプチドをクリックするか、そのペプチドが選択されるまで下向き矢印を押します。

一見すると、このペプチドは共溶出、積分境界、質量誤差、ドット積値と言う面では非常に優れているように見えます。もっとよく見てみると、プリカーサークロマトグラムの右端がフラグメントクロマトグラムよりも低下するのが速いことに気付かれるかもしれません。これもやはり信号ドロップアウトであり、以前のものと原因は同じです。これについては後でまた説明します。

4つ目のペプチドは3つ目と比べてあまり魅力はありませんが、53.5分辺りで常にピークを積分しているように見えます。

抽出されたクロマトグラム全範囲を表示するには、以下の操作を行います。

* [ **ビュー** ] メニューで [ **自動ズーム** ] を選択し、[ **なし** ] (Shift+F11) をクリックします。

これによってSkylineがDDA ID時間の周囲+/-5分枠でクロマトグラフを抽出しただけであることが示されます。IDは53.5分近辺で発生していますので、Skylineは48.5～58.5分を抽出します。

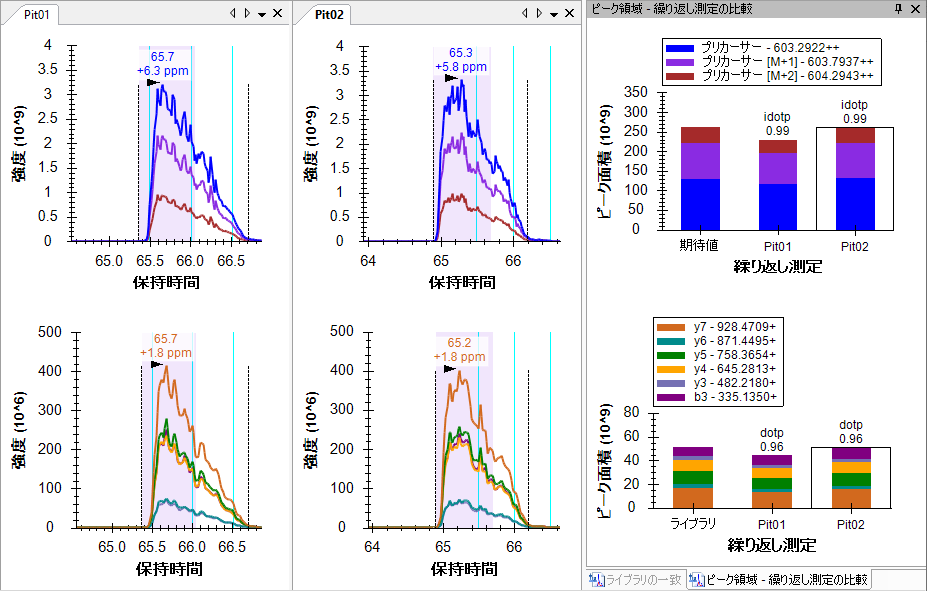
DIAでは、プリカーサー単離ウィンドウが非常に幅広い（例：10～25 m/z）ため、多数の干渉が考えられます。しかしながら、Skylineによるピーク選択の自動化は、非常に多数の干渉がある場合でも、多くの場合正しいピークを選択できます。この場合、より多くのピークがあることがわかりますが、Skylineはそれでも利用可能な最良のピークを選んでいます。以下の操作を行ってピークの拡大縮小を復元したら、これはこのままにしておいてそのまま下に進みます。

* [ **ビュー** ] メニューで、 [ **自動ズーム** ] を選択し [ **最適ピーク** ] (F11) をクリックします。

次の5個のプリカーサーは、ペプチドK.NYGLLY**C**FR.K [184, 192]までは優れたものに見えます。このペプチドのピークは、幅が1分にわたり、あまりガウス形状とは言えません。Skylineは、両ピークに誤って正しい積分境界を与えているため、ピークが2/3しかカバーされていません。これを修正するには、以下の操作を行います。

* 正しい積分境界を示す黒の点線の上にマウスカーソルを合わせると、カーソルが左右の矢印のあるスプリッターカーソル () に変わります。
* ピーク全体がカバーされるように境界を好きな場所にクリックアンドドラッグします。

このペプチドのプロットは、以下のようになります。



青緑色のIDライン、idotp 0.99、そしてdotp 096の組み合わせにより、このピークの割り当てにはかなりの信頼性があります。

これらのランにおけるクロマトグラフィーにより、このリストにある後半のペプチドにはこの種の修正を複数回繰り返す必要があります。複数の公開されたソフトウェア比較論文に基づくと、Skylineはピーク積分において他のツールと比較しても劣るものではありません。難しいピーク形状では、他のツールにも同じような問題があります。

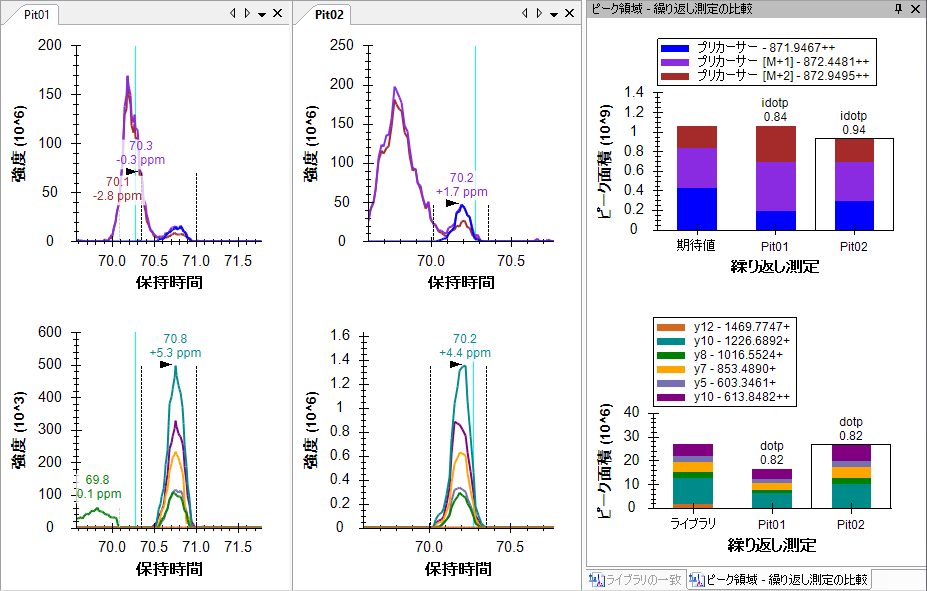
# 抽出されたクロマトグラムの理解

次の2つのペプチドであるAVVLSHYIHNLSSEMFSEFDKとAINS**C**HTSSLATPEDKには、DDAからペプチドが同定された回数に関して説得力のある信号が欠けています。Shift-F11を使用して拡大したり、マウススクロールホイールを使ってマウスを中心とした拡大縮小をしたり、ctrl-click-and-dragでパンしたりすることができますが、どちらのペプチドに対しても、クロマトグラム中には何も見つからないでしょう。これらは削除するか、または単にペプチドSWNEPLYHLVTEVRに進みます。

このペプチドには2つのプリカーサーがあります。871.9467にある2価と、581.6335にある3価です。

* 最初のプリカーサー871.9467++ (idotp 0.94、dotp 0.82) を選びます。

どちらのプリカーサーにも、うまく共溶出しているフラグメントイオンのクロマトグラムがあり、明白な干渉はありません。また、どちらもMS1クロマトグラムの左側で明白な干渉を示しています。



**ピーク面積**プロットを見ると、強度の規模（イオン個数/秒）が把握できます。ここでプリカーサー (MS1) は総数100万近く、プロダクト (MS/MS) は総数2万から2万5千または40から50倍の差です。先ほど見たように、MS1信号はMS/MSを小さく見せますが、同時にノイズや干渉も起こしやすくなっています。

|  |
| --- |
| **注**：以下の**フルスキャン**ビューの使用はすべて、フルチュートリアルデータセットをダウンロードした場合にのみうまくいきます。そうでなければ、Skylineは生データファイルがないというメッセージを表示します。 |

このクロマトグラムが抽出されたMS1スキャンで何が起きたのかをよりよく把握するため、以下の操作を行います。

* 積分されたピークの左側にある「プリカーサー[M+1]」ピークの頂点にマウスを合わせます。紫色の丸が表示され、マウスカーソルが指差ししている手に変わります。
* 丸の中をクリックします。

Skylineは、プリカーサークロマトグラムの点が抽出された幅広いMS1スキャンを表示する**フルスキャン**ビューを表示します。虫眼鏡のボタンをクリックし、ビューの右上角にあるプラス () をクリックするとMS1スキャン全体が見えるようになります。代わりに、以下の操作を行います。

* 70.19分（タイトルに表示）で取得されたスペクトルが表示されていることを確認します。または、右上角にある矢印ボタンを使ってこのスペクトルまで移動します。

**フルスキャン**グラフは以下のようになります。



セントロイドモードピークの強度は棒のようになります。クロマトグラフで点を形成していた強度は、クロマトグラムの色で強調表示され、抽出範囲は影付き領域として表示されます。抽出されていなかったM+3およびM+4ピークが影付きピークの右側にあるのが見えます。また、干渉しているピークの頂点を形成していたスペクトルが、このプリカーサーの積分されたピークの左側に表示されています。ターゲットペプチドよりも1ダルトン重い2価のプリカーサーの同位体分布全体が見られます。モノアイソトピックピークは、ターゲットのM+1に非常に近くなっています。

* 今度は矢印ボタン（、**フルスキャン**ビューのツールバー内）を使って、これと積分されたピークの頂点を形成していたものとの間にあるスキャンを確認します。

70.56分には、ターゲットモノアイソトピックフィルタ範囲にあるピークを含む最初のMS1スペクトルが青色で強調表示されます。70.67分までは、モノアイソトピックピークが最も高い最初のMS1スペクトルは見えません。

プロダクトイオンMS/MSスキャンに同様の分析を実施するには、以下の操作を行います。

* 「y10+」ピークの頂点にマウスを合わせます。水色の丸が表示され、マウスカーソルが指差ししている手に変わります。
* 丸の中をクリックします。

**フルスキャン**ビュー内のグラフが以下のように表示されます。



今クリックしたクロマトグラムの点となったMS/MSピークが水色で強調表示され、y10と注釈が付くことがわかります。同位体ピークは1 m/zずつ離れていることからも、これらが1価のイオンからによるものであるという確信が持てます。

* 紫色のy10++クロマトグラムにも上記のステップを繰り返します。

以下のようなスペクトルプロットが表示されます。



ここでも、一連の中で一番高いピークが強調表示されて注釈が付きます。そのうち、以下のピークは0.5 *m/z*ずつ離れており、2価のイオンの同位体分布であることを示します。これにより、プロダクトイオンクロマトグラムで抽出された信号がターゲットとしているものと一致するという、より確かな確信が持てます。

## 干渉がどのように信号ドロップアウトを生成するかの理解

上記で確認したMS1スペクトルでは、ニュートラル質量が1ダルトン異なるにも関わらず、干渉イオンと対象としているターゲットイオンの質量電荷比はかなり類似しています。あまり類似していないイオン*m/z*値でも十分に近く、質量分析計では単一の未解決プロファイルピークとなることがあります。この場合、この単一のピークがターゲット*m/z*値よりも20 ppm以上高くても*m/z*値にセントロイド化されることがあります。本チュートリアルで既に見てきた信号ドロップアウトの最も一般的な原因はこれです。

以下の操作を行い、**フルスキャン**ビューを使用してこの様子を見ます。

* K.ELVYETVR.V ペプチドをクリックします。（[ **ターゲット** ] ビューの3番目）
* 「Pit02」繰り返し測定の「プリカーサー」ピークの頂点にマウスを合わせます。青い丸が表示され、マウスカーソルが指差ししている手に変わります。
* 丸の中をクリックします。
* 右矢印ボタン（、**フルスキャン**ビューのツールバー内）を使って、41.7分のところまでMS1スペクトルを確認します。

時間が経つほど、M-1（モノアイソトピックプリカーサーの左）にあるピークが強くなっていくのがわかります。

41.67分には、最大のピークとなります。



41.68分で、抽出されたピークのm/zが左にシフトするのがわかります。



そして41.70分では、+/- 20 ppmの抽出範囲を超えてシフトしています。



DIAのセントロイド化されたMS/MSスペクトルからの抽出は、特にThermo Orbitrap装置では、幅広い統計検定でより信頼性があることが実証されていますが、データを見るときには、これがやはり理解に有用な概念です。この設定での干渉は、通常は信号が増加するであろうと期待されるような状況で総信号を減少させることがあります。

# まとめ

本チュートリアルでは、DIA装置メソッドの生成またはこのチュートリアル同様に単なるデータ分析に使用するDIA取得スキームを定義しました。DDA検索結果からスペクトルライブラリを構築し、測定されたRTに基づいてライブラリ内で保持時間制限を設定し、対象とするタンパク質とDDAランからの一致するスペクトルに基づいて測定するトランジションセットを定義し、DIAランをインポートし、その結果生じるデータ品質を分析しました。最後に、すべてのSkylineドキュメント同様、対象とするペプチドのピーク面積と統計情報が生成できます。本チュートリアルでは、さらに高度な調査のためにその他に複数のチュートリアル（[ターゲットメソッドの編集](https://skyline.ms/tutorial_method_edit.url)、[既存の定量的実験](https://skyline.ms/tutorial_existing_quant.url)、[iRT保持時間の予測](https://skyline.ms/tutorial_irt.url)、[高度なピーク選択モデル](https://skyline.ms/tutorial_peak_picking.url)、[アッセイライブラリのインポート](https://skyline.gs.washington.edu/labkey/_webdav/home/software/Skyline/%40files/tutorials/ImportingAssayLibraries-2_6.pdf)および[パノラマクロマトグラムライブラリ](https://panoramaweb.org/labkey/wiki/home/page.view?name=chromatogram_libraries)）に言及しています。

本ワークフローでは、DIAランの前、後、またはその間に取得したDDAランがあれば、SkylineであらゆるDIAデータセットを分析できるようになります。これはDIAデータ分析を始める簡単な方法の1つですが、[DIA/SWATHデータの分析](https://skyline.ms/tutorial_dia_swath.url)チュートリアルで勉強を続けることが強く奨励されます。そこでは、較正されたiRT値を使用してスペクトルライブラリを構築し、クロマトグラムピークのスコアと選択にこの正規化された保持時間とmProphet統計モデルを使用します。また、Skylineでのグループ比較も使用して、調整されたp値で各ペプチドまたはタンパク質の条件による倍率変化を推定します。

# 参考文献

1. Venable, J. D., Dong, M.-Q., Wohlschlegel, J., Dillin, A. & Yates, J. R. Automated approach for quantitative analysis of complex peptide mixtures from tandem mass spectra. *Nat. Methods* **1,** 39–45 (2004).

2. Gillet, L. C. *et al.* Targeted data extraction of the MS/MS spectra generated by data-independent acquisition: a new concept for consistent and accurate proteome analysis. *Mol. Cell. Proteomics MCP* **11,** O111.016717 (2012).

3. Egertson, J. D. *et al.* Multiplexed MS/MS for improved data-independent acquisition. *Nat. Methods* **10,** 744–746 (2013).

4. Krokhin, O. V. *et al.* An improved model for prediction of retention times of tryptic peptides in ion pair reversed-phase HPLC: its application to protein peptide mapping by off-line HPLC-MALDI MS. *Mol. Cell. Proteomics MCP* **3,** 908–919 (2004).

5. Escher, C. *et al.* Using iRT, a normalized retention time for more targeted measurement of peptides. *Proteomics Accept.* (2012).

6. Reiter, L. *et al.* mProphet: automated data processing and statistical validation for large-scale SRM experiments. *Nat. Methods* **8,** 430–435 (2011).