猫も杓子も

Geneious Prime でシーケンス解析

第 15 回 CRISPR サイトの検索



Geneious Prime の CRISPR ツールでは、配列中の CRISPR サイトを検索してアノテーション付けし、ゲノムに対してオフターゲット(非特異的)結合サイトを照合することができます。

CRISPR/Cas9 システムは、遺伝子編集のための RNA 誘導型エンドヌクラーゼ技術です。このシステムでは、Cas9 エンドヌクラーゼを切断部位に導くために、PAM(Protospacer Adjacent Motif)サイトに隣接する 20 bp のターゲット配列からなるガイド RNA(gRNA)を必要とします。

Find CRISPR sites ツールでは、選択した遺伝子の gRNA(「CRISPR」)サイトを検索し、さらにゲノム中のオフターゲット結合サイトを検索することができます。選択した配列の CRISPR サイトは、ターゲットサイトでの活性を予測する **Activity score** と、ゲノムの他の場所にあるオフターゲット結合サイトの存在に基づいた **Specificity score** が計算されます。

今回は、*Saccharomyces cerevisiae*(パン酵母)の LYP1 遺伝子の「GN(20)GG」gRNA (CRISPR)サイトを検索し、*Saccharomyces cerevisiae* ゲノムのオフターゲット結合サイトをチェックする手法をご紹介します。

チュートリアル用のデータは[こちら](#)からダウンロードすることができます。

Saccharomyces cerevisiae S288c の LYP1 CDS を使用し、S288c ゲノムをオフターゲットデータベースとして使用します。LYP1(Lysine permease)は、酵母ゲノムの 14 番染色体に存在するカチオン性アミノ酸の取り込みを担う 3 つのアミノ酸パーミターゼ(ALP1p, Can1p, Lyp1p)のうちの 1 つです。この遺伝子の詳細については[こちら](#)をご覧ください。

LYP1 CDS のドキュメントを選択し、**Annotate and Predict** → **Find CRISPR sites** へ進みます。

まず、オプションをデフォルトの設定にリセットするため、ウィンドウの左下にある歯車をクリックし、**Reset to defaults** を選択します(現在の設定がデフォルトの場合はグレーアウトしています)。

この LYP1 CDS 配列の全体から CRISPR サイトを探索するためには、**Find CRISPR Targets** の横の **Anywhere in sequence** を選択します。

Geneious Prime の CRISPR ツールは、3' Cas9 サイトと、5' Cpf1 サイトの検索に使用することができますので、**PAM Site Location** セレクトで、**3' (Cas9)**が選択されていることをご確認ください。

Target と **PAM Site** フィールドに、検索したい gRNA 配列を入力します。全ての CRISPR サイトを評価したい場合は **Target** に N(20)を入力します。オフサイト結合を確認したい特定の配列がある場合は **Target** に配列を入力することができます。例えば、"GN(20)GG"ガイド配列を検索したい場合は、**Target** に"GN(19)"、**PAM Site** に"NGG"と入力します。プレビューには、この **Target** と **PAM** のガイド配列、"GNNNNNNNNNNNNNNNNNNNNNNNGG"が表示されます。

デフォルトでは [Doench ら\(2016\)](#)の方法を用いて CRISPR サイトの活性をスコアリングします。活性・オンターゲットスコアリングでは、gRNA サイト自体の配列の特徴をモデル化し、活性を予測します。Doench ら(2016)のモデルを初めて実行する場合、Geneious は実行前に必要な環境 (python と R)をインストールするために、少し時間がかかることがあります。

オフターゲット結合に基づいてサイトをスコアリングするには、**Specificity Scoring** の **Score against an off-target database** にチェックを入れます。

Specificity Score は、gRNA が結合する可能性のあるオフターゲットサイトの数と、そのオフターゲットサイトが元の配列とどれだけ類似しているかを示す指標です。これは MIT の Zhang 研究室が開発した方法に従って計算されます(詳細は[こちら](#)をご参照ください)。各オフターゲットサイトは、元の CRISPR サイトとの類似度や、ミスマッチの発生箇所(PAM サイトに近いミスマッチは PAM サイトから遠いミスマッチよりも結合に影響を与える)に基づきスコアリングされます。オフターゲットサイトのスコアが高いほど、オリジナルの CRISPR サイトとの類似性が高いことを示します(CRISPR/Cas 複合体がオフターゲットに結合する可能性が高いことを意味します)。CRISPR サイトの総合的な Specificity Score は、100%からターゲットゲノム内のオフターゲットスコアの加重合計を差し引いたものです。そのため、Specificity Score が高いほど、オフターゲットの可能性が少なく、優れた CRISPR サイトであることを示しています。

次に、チュートリアル用データの場合、*Saccharomyces cerevisiae* S288c ゲノムは Yeast genome サブフォルダのなかに用意されていますので、オフターゲットデータベースとして使用するには、**Specificity Scoring** パネルにあるフォルダアイコンをクリックし、フォルダセレクトタから **Yeast genome** フォルダを選択します。

オフターゲットデータベースとしては、通常、対象生物の全ゲノムを用いますが、ターゲティングベクターなど、他の配列を含めることもできます。オフターゲットデータベースは、Geneious データベース内に新しく空のフォルダを作成し、使用したい配列をインポートすることで作成することができます。

研究者によって目的のゲノムが異なることや、ゲノム配列はデータサイズが非常に大きいこと、また、つねに新しいバージョンのゲノムアセンブリがリリースされる可能性があることから、Geneious にはデフォルトではフルゲノム配列のコピーが内蔵されていません。ゲノムはソースパネルの下にある **NCBI** フォルダを使用して NCBI から直接ダウンロードすることができます。一般的に研究されているゲノム(例えば、ヒト、ゼブラフィッシュ、ラットゲノムなど)は、Geneious Sample Documents の **Genomes** フォルダにあるリンクを使用して NCBI からダウンロードすることもできます。また、他のソースからダウンロードしたゲノムを一般的なファイルフォーマットで Geneious にインポートすることも可能です。

以下のスクリーンショットのように設定できたら、**OK** をクリックして解析を実行します。

Find CRISPR Sites

Find CRISPR Targets: Anywhere in sequence Selected region

PAM Site Location: 3' (Cas9) ▾

CRISPR Site ?

Target: GN(19) PAM Site: NGG

Preview: GNNNNNNNNNNNNNNNNNNNNNGG

Activity scoring: Doench et al. (2016) ?

Specificity Scoring ?

Score against an off-target database: Yeast genome

Maximum mismatches allowed against off-targets: 3 ▾

Maximum mismatches allowed to be indels: 0 ▾

Pair CRISPR Sites ?

Maximum overlap of paired sites: 6 ▾

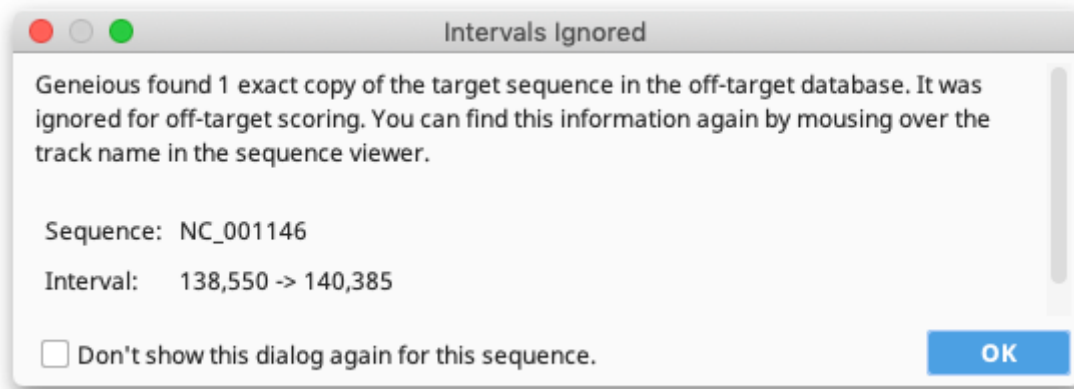
Maximum allowed space between paired sites: 16 ▾

Color CRISPR Sites by: Activity Score ▾

Cancel OK

大きな配列内から Selected region でサイトを探すことを選択し、Specificity Scoring でオフターゲットデータベースを選択しなかった場合は、自動的に配列の選択されていない領域に対してオフターゲット結合がテストされます。

解析を実行すると、以下のようなメッセージが表示されます。



このメッセージは、CRISPR サイトを検索している LYP1 CDS 配列が、酵母ゲノムの一部であるため、オフターゲットデータベースにも存在しているために表示されます。この領域内のマッチは、ガイドサイトそのものである可能性が高いため、オフターゲットのスコアリングでは無視されます。

Geneious 製品概要・フリートライアルリクエストについては[こちら](#)

『Geneious Prime でシーケンス解析』の過去の記事は[こちらでチェック！](#)

TDB News 12. 2022
トミーデジタルバイオロジー株式会社
Phone 03-6240-0843 Fax 03-6240-0461