

単一のオリゴプールを由来とする、 的を絞ったマルチプレックス化フォーカスト CRISPRライブラリ

要約

CRISPR/Cas9システムを用いた機能的ゲノムスクリーニングは、大規模に遺伝子機能を探索するための強力なツールである一方、当該技術の利用と実施については、所定の既存CRISPR sgRNAライブラリの制約を受けます。今回、バイオインフォマティクスおよびウェットラボの簡便なワークフローとしてCustom Library Multiplexed（マルチプレックス化カスタムライブラリ）クローニング（CLUE：[Custom Library Multiplexed Cloning]：crispr-clue.de）と、Twist Bioscienceの単一オリゴプールを組み合わせて用いることで、多数の高品質なカスタムsgRNAライブラリの作製を行いました。Twist Bioscienceのオリゴプールの優れた忠実性により、得られるCLUEライブラリにはsgRNAが均一に存在し、ライブラリ間でのクロスコンタミネーションも起こりません。Twist BioscienceのオリゴプールCRISPRスクリーニングアプリケーションを成功に導くための厳しい要求に応えるものです。

はじめに

CRISPR/Cas9ゲノムエンジニアリングは、遺伝子機能を調べるために、実質的にどのような規模においても用いることが可能です。sgRNAの合成、クローニング、導入およびスクリーニングをまとめて行うプール型スクリーニングアプリケーションにおいて、このことは特にあてはまります。ゲノムワイドなsgRNAライブラリは偏りのないヒット化合物同定を可能とするものの、その膨大な規模と大変な複雑さから、特定の実験系（すなわちin vivo）や経験の少ないCRISPRスクリーニング担当者にとっては障害となり得ます。

sgRNAライブラリ構築に用いるオリゴ鎖の合成からヒット化合物同定にいたるまで、CRISPRスクリーニングプロセスのあらゆる局面において、スクリーニング開始前の慎重な検討と最適化が求められます（Doench, 2018）。CRISPRスクリーニング実験の複雑さは、使用するsgRNAライブラリが複雑になるほど増大します。非常に複雑なsgRNAライブラリは、不均一なオリゴ合成、過剰なPCR増幅、および競合につながってしまう過剰な細菌増殖により、ライブラリを調整する間、スクリーニングを行う前に偏ってしまう恐れがあります（Joung et alら, 2017）。

焦点を絞ったカスタムsgRNAライブラリは実験としての忠実性が高いため好ましく、初代培養細胞、オルガノイド、動物の全身といった有限または複雑な系において実施されるスクリーニングにも適しています。そのような系は研究課題となっている生物学的現象をより良く再現し、結果として実際に生体で起きていることにより関係しており影響力のある洞察を与えてくれるため、CRISPRはそのような系においてますます用いられるようになっています（Doench, 2018）。ただし、ゲノムワイドなスクリーニングに必要な細胞数が多いため、CRISPRスクリーニングに対してそれらの系が広く用いられているわけではありません。

焦点を絞ったカスタムsgRNAライブラリが利用しやすくなると、生物学的により実際に近いモデルに対してCRISPRスクリーニングを適用することが容易になります。しかし、多くのラボでは、そのようなライブラリの作製に必要な技術やバイオインフォマティクスの専門知識が不足しています。そのため、ワークフローにおいて使いやすいsgRNAライブラリ設計ツールと質の高いオリゴヌクレオチドプール合成を組み合わせることで、あらゆるユーザーにとってCRISPRスクリーニングが容易になります。

Helmholtz Zentrum Münchenおよびthe Ludwig-Maximilians大学ミュンヘンの研究者たちは、複数のより大きな単一のライブラリにプールされたサブゲノム規模のsgRNAライブラリを効率的に構築するパイプラインであるCLUE（crispr-clue.de）を開発しました（Beckerら, 2020）。sgRNAオリゴ鎖にはバーコード配列が層状に隣接しており、的を絞ったフォーカストCRISPRスクリーニングアプリケーションのための、プール全体の増幅とライブラリ個別のクローニングが可能です。

このアプリケーションノートでは、高品質オリゴプールから始めて、どのようにしてCLUEのワークフローによってサブプールの品質が保たれ、ライブラリ間のクロスコンタミネーションが実質的に起きない均一なsgRNAライブラリが生み出されるのかについてご紹介します。Twist Bioscienceのオリゴプールの高い忠実性により、PCRで増幅されクローニングされたサブライブラリにおけるsgRNA発現が維持されます。結果として、多数の高品質なカスタムsgRNAライブラリを作製するためのend-to-endのパイプラインが提供されます。

方法

sgRNAライブラリ作製

CLUEパイプラインを用いて単一のTwist Bioscienceオリゴプールから10個のCLUEライブラリを下記の通り作製しました。

CLUEパイプライン用のマスターsgRNAライブラリを作製するために、カスタマイズしたPythonスクリプトを用いて、ゲノムワイドな既存sgRNAライブラリからsgRNAを選択しました（Wangら, 2015；Doenchら, 2016；Horlbeckら, 2016）。それらのリファレンスライブラリには、マウスとヒト両方のゲノムに及ぶsgRNAおよび複数のCRISPR摂動の戦略（すなわちCRISPRノックアウト、CRISPRi、CRISPRa）が含まれます。選択したsgRNA配列に、プロモーター（U6またはH1）、sgRNA鋳型、および複数のアダプター配列を付加して、オリゴプールおよびライブラリ個別の増幅を可能にしました。得られたオリゴ配列は、出力ファイルに書き込まれ、オリゴ合成のためにTwist Bioscience®に送られました。

凍結乾燥したTwist Bioscienceのオリゴプールを再懸濁し、Kapa[®] Hifi DNAポリメラーゼ (Roche) および初期増幅プライマーを用いて以下に述べる条件 (PCR 1) でPCR増幅を行いました：98°C 3分；98°C 30秒、62°C 15秒、72°C 10秒を15サイクル；72°C 2分。Zero Blunt[®] TOPO[®] PCRクローニングキット (Life Technologies Corp.) を用いて直接クローニングを行いました。TOPOプラスミドを精製し、Endura[®]エレクトロコンピテントセル (Lucigen Corp.) 中で形質転換し、単離しました。得られたTOPOライブラリについては、アガロースゲル電気泳動と次世代シーケンシング (NGS) を用いてクオリティチェックを行いました。

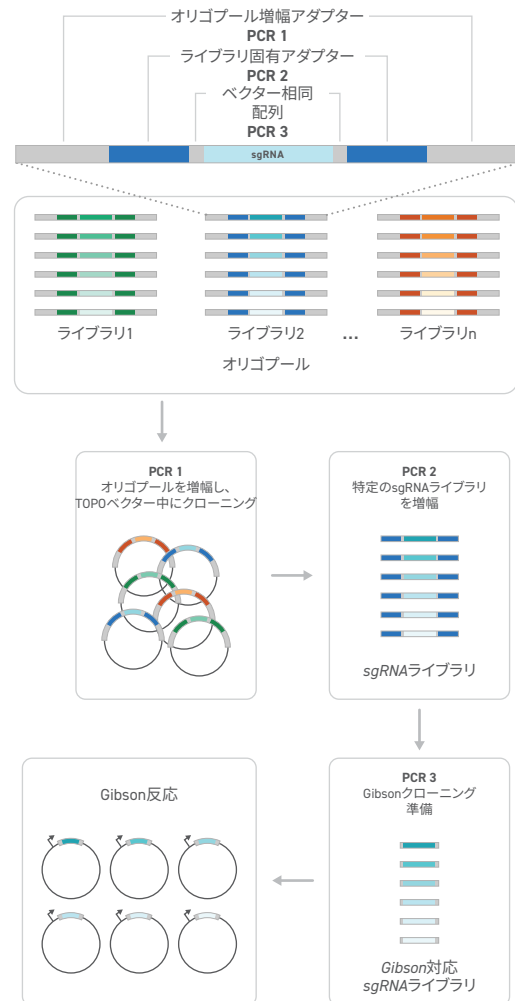
TOPOライブラリは、その後のクローニングで各CLUEライブラリを増幅するためのテンプレートとして機能します。CLUEライブラリについて、ライブラリに固有のプライマーと以下の温度サイクル条件 (PCR 2) を用いてPCR増幅を行いました：98°C 2分；98°C 20秒、57~62°C 15秒、72°C 1秒を30サイクル；72°C 1分。増幅したCLUEライブラリはカラム精製し、Gibson assembly[®] (PCR 3) (Synthetic Genomics, Inc.) 用のクローニング断片を増幅するために用いました。PCR 3の温度サイクル条件はPCR 2と同一であり、アニーリング温度は62°Cで一定としました。PCR 3によって生成された断片については、NEBuilder HiFi DNA Assembly[®] Master Mix (New England Biolabs, Inc.) を用いて線状化sgRNA発現ベクター中にクローニングしました。クローニング産物をEnduraエレクトロコンピテントセル中で形質転換し、NucleoBond[®] Xtra Maxiキット (Macherey, Nagel, GmbH) を用いて精製しました。

得られたCLUEライブラリは、NGSを用いてクオリティチェックしました。各ライブラリについて、100万本のsgRNAごとにPCRを実行し、アダプター配列とバーコード配列を追加しました。アンプリコンはIllumina HiSeq[®] 2000 (Illumina, Inc.) (50 bpのシングルエンドリード) でsgRNAあたり500リードを超える深さで精製および配列決定しました。

プール型CRISPRiスクリーニング

U-87 MG細胞 (ATCC[®] HTB-14[™]) に、高感度緑色蛍光タンパク質 (EGFP) でタグ付けされたヌクレアーゼ不活化Cas9 (dCas9) をコードするレンチウイルスを用いて形質導入を行いました。得られた細胞集団を蛍光活性化セルソーティング (FACS) によって濃縮してから、低い感染多重度 (MOI) でCLUEライブラリで形質導入し、各細胞中にsgRNAが一つずつ確実に取り込まれるようにしました。CLUEライブラリで形質導入されたU-87 MG dCas9-EGFP細胞は、FACSで濃縮し、スクリーニング開始前に培地中で短時間増殖させました。Wizard[®]ゲノムDNA精製キット (Promega) を用いて、0日後、16日後にゲノムDNAを採取しました。継代中、細胞数はsgRNAあたり250細胞となるよう維持しました。取り込まれたsgRNA配列を、Kapa Kapa Hifi DNAポリメラーゼ (Roche) およびP5-Hif1、P7-EF1aの2つのプライマーを用いて精製ゲノムDNAからPCRで増幅させました。精製後、得られたアンプリコンをHiSeq 2000 (50bpのシングルエンドリード) を用いて配列決定しました。カスタマイズしたPythonスクリプトでNGSのリードのマッピングとカウントを行い、MAGeCKアルゴリズムを用いて解析を行いました。

CLUEのワークフロー



ステップ1: スクリーニングの設計

- ・生物種の選択 (例：ホモサピエンス)
- ・遺伝子の摂動法の選択 (CRISPRi, CRISPRa, CRISPRko)
- ・標的遺伝子群または含めるべきsgRNA配列の選択
- ・sgRNAベクターの選択
- ・sgRNA/遺伝子の数の選択

ステップ2: 情報をCLUEにアップロード (crispr-clue.de)

ステップ3: マルチプレックス化オリゴプール配列情報をダウンロードし、Twist Bioscienceのオンライン注文プラットフォームに直接送信

ステップ4: オリゴプールを受け取る

ステップ5: オリゴプールを保存し、オリゴプールの品質を評価 (オプション):

- ・最も外側のプライマー結合部位でオリゴプールを増幅 (「PCR 1」)
- ・アンプリコンをTOPO[®]ベクター (Life Technologies Corp.) 中に一括クローニング
- ・TOPOベクターへの挿入配列をNGSで決定し、1 log以内に表れる90%のオリゴプールを評価

ステップ6: 個々のサブライブラリーのクローニング

- ・ライブラリ固有のPCRを実行 (「PCR 2」)
- ・sgRNAベクターに相同なプライマーでPCR 3を実行することにより、sgRNAベクターにクローニングするためのアンプリコンを準備 (「PCR 3」)
- ・線状化sgRNAベクターへのアンプリコンの一括クローニング
- ・クローニングしたライブラリをNGSで配列決定して品質を評価



ライブラリ数	意図したライブラリサイズ (sgRNA)	NGSで同定したsgRNA数	sgRNAクローニング効率 (%)	完全一致NGSリード (%)	その他のサブライブラリへのリードマッピング (%)	空のベクターリード (%)
1	265	265	100.00	92.53	0.100	0.02
2	705	705	100.00	92.64	0.150	0.01
3	445	444	99.78	92.71	0.020	0.01
4	170	170	100.00	93.19	0.010	0.00
5	520	520	100.00	94.40	0.005	0.00
6	530	530	100.00	89.68	0.010	0.00
7	504	504	99.80	78.90	0.040	8.64
8	449	449	99.78	77.32	0.370	10.39
9	1024	1024	99.90	63.93	0.002	14.95
10	970	970	100.00	62.89	0.030	15.95

表1. CLUEパイプラインで作製した10個のプール型sgRNAライブラリのNGSに基づいた品質評価。

結果

CLUEワークフローにおけるプールの均一性

TOPOを用いてクローニングした最初のライブラリは正規分布で、ライブラリの90%超が1 log以内に表れました (図1)。本研究でTOPOライブラリから生成されたCLUEライブラリの品質管理基準を表1にまとめました。CLUEライブラリはいずれも、元のオリゴヌクレオチドプールに高確率 (63~94%) でマッピングされました。sgRNAの分布は10個のCLUEライブラリ全てにおいて正規分布でした。各ライブラリについて、sgRNAの90%超が分布平均の1 log以内に収まり、それはsgRNAが全てのライブラリにおいて均一に発現したことを示しています (図2)。6個のCLUEライブラリが完全に発現し、残りの4個はそれぞれ1つずつのsgRNAが不足していました。全てのライブラリにおいて、他のライブラリにマップされたリードは0.5%未満でした。つまり、CLUEワークフローはライブラリ間のクロスコンタミネーションなく均一に発現するsgRNAライブラリを生成することが可能です。

CLUEライブラリによる神経膠腫の転写調節因子の同定

CLUEライブラリの機能性をテストするため、ある標的CRISPR干渉 (CRISPRi) スクリーニングの設計を行いました。この目的のために、CRISPRiを媒介させるため、U87-MG神経膠芽腫細胞株へのヌクレアーゼ不活化Cas9 (dCas9) の形質導入を行いました。続いてdCas9を発現するU87-MG神経膠芽腫細胞株に、88個の転写調節因子を標的とする505種類のsgRNAを含むCLUEライブラリを形質導入しました。スクリーニングの0日目と16日目に単離されたゲノムDNAにおけるsgRNAの発現を比較して、膠芽腫細胞の増殖に関わる転写調節因子を同定しました。ジンクフィンガータンパク質217 (ZNF217) およびE2F1を含む既知の腫瘍形成調節因子を標的とするsgRNAが、スクリーニング終了時に枯渇していました。これらのヒット化合物は、sgRNAのランキングアルゴリズムであるMAGeCKを用いて確認しました (次ページの図3)。この結果は、プール型CRISPRスクリーニングにおいてCLUEライブラリが生物学的に重要なヒット化合物を同定できることを明確に示しています。

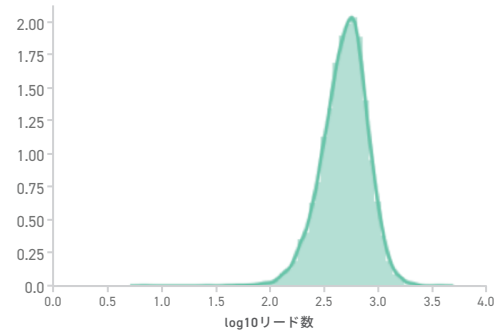


図1: TOPOプールにおいて発現するオリゴ分布を表すNGS解析。

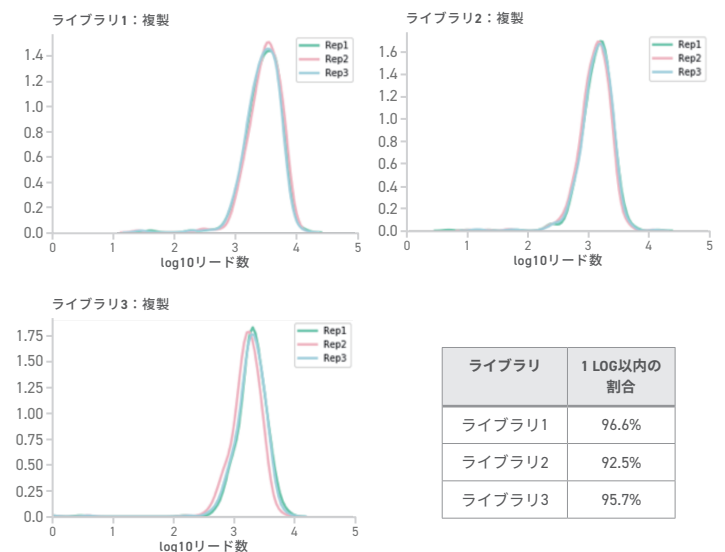


図2. TOPOライブラリによって厳密に均一性が保たれていることを示す表1のライブラリ1~3における分布。

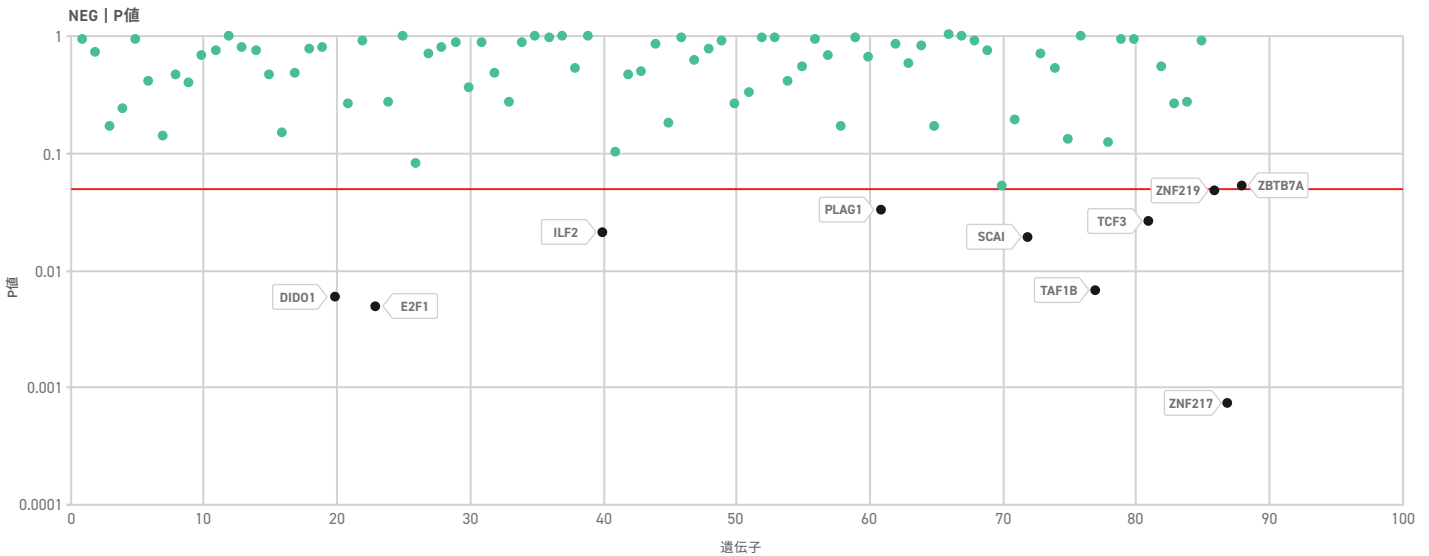


図3. 標的CRISPRiスクリーニングによって明らかとなる悪性神経膠腫細胞の増殖適応度に必要な転写因子。16日後についての枯渇の有意性を標的遺伝子に対してプロット。赤線：p=0.05。p<0.05を示す枯渇遺伝子は黒のデータポイントで示す。

結論

CRISPRスクリーニングは生物学的経路・プロセスの研究にとって拡張性に優れた効果的なツールです。しかし実際には、ゲノムワイドなCRISPRスクリーニングを実施することは困難である場合があります。的を絞ったフォーカストsgRNAライブラリは使いやすいものの、そのようなライブラリを作製するためのツールはほとんどなく、当該技術の実施に関心のあるユーザーにとってさらなる障害となっています。

的を絞ったマルチプレックス化フォーカストライブラリの利用は、単一のプールから複数の実験を実施できるようにすることで、コスト・効率の両面にメリットがあります。本稿に記載したパイプラインにより、複数標的のマルチプレックス化CRISPRライブラリの生成を可能にする使いやすいソリューションが実現します。CLUEを使用することで、ユーザーはライブラリのサイズ、標的生物種、摂動の戦略（すなわちCRISPRノックアウト、CRISPRi、またはCRISPRa）をカスタマイズすることで、実験における厳しい要求を満たすことが可能です。

CRISPRスクリーニングにおけるヒット化合物の同定は、選択プロセス前後のsgRNA発現の変化を比較して行います。そのため、配列の忠実性とsgRNAの均一性は、ライブラリ調製中に評価しなければならない重要なパラメータです。オリゴヌクレオチド合成の質が悪い場合、ライブラリ全体でsgRNA発現が偏り、意図しない、または遺伝子の摂動効果のないsgRNA配列が導入される恐れがあります。こうした影響は、細胞数とsgRNA発現のモニタリングがより困難な、細胞が限定されたスクリーニング（例：*in vivo*スクリーニング）において深刻になります。本研究で作製したCLUEライブラリは、均一なsgRNA分布と、それらが由来する最初のオリゴヌクレオチドプールへの高いマッピング率を示しました。CLUEをTwist Bioscienceの優れたオリゴ合成品質と組み合わせると、より大きなプールから非常に均一なサブライブラリを抽出できるため、研究者は忠実性の高い標的CRISPRスクリーニングですぐに使用できるsgRNAライブラリを簡便に設計・実現できます。

引用文献

- Becker M, Noll-Puchta H, Amend D, Nolte F, Fuchs C, Jeremias I, Braun CJ (2020) CLUE: a bioinformatic and wet-lab pipeline for multiplexed cloning of custom sgRNA libraries. *Nucleic Acids Res* 48:e78.
- Doench JG (2018) Am I ready for CRISPR? A user's guide to genetic screens. *Nat Rev Genet* 19:67–80.
- Doench JG, Fusi N, Sullender M, Hegde M, Vaimberg EW, Donovan KF, Smith I, Tothova Z, Wilen C, Orchard R, Virgin HW, Listgarten J, Root DE (2016) Optimized sgRNA design to maximize activity and minimize off-target effects of CRISPR-Cas9. *Nat Biotechnol* 34:184–191.
- Horlbeck MA, Gilbert LA, Villalta JE, Adamson B, Pak RA, Chen Y, Fields AP, Park CY, Corn JE, Kampmann M, Weissman JS (2016) Compact and highly active next-generation libraries for CRISPR-mediated gene repression and activation. *Elife* 5.
- Joung J, Konermann S, Gootenberg JS, Abudayyeh OO, Platt RJ, Brigham MD, Sanjana NE, Zhang F (2017) Genome-scale CRISPR-Cas9 knockout and transcriptional activation screening. *Nat Protoc* 12:828–863.
- Wang T, Birsoy K, Hughes NW, Krupczak KM, Post Y, Wei JJ, Lander ES, Sabatini DM (2015) Identification and characterization of essential genes in the human genome. *Science* 350:1096–1101.