

学位論文要旨

Automated design and detailed profiling of MMEJ-assisted knock-in using a newly constructed computational pipeline

(新規に構築したコンピュータパイプラインを利用した MMEJ ノックインの自動設計及び詳細プロファイリング)

広島大学大学院理学研究科 数理分子生命理学専攻
氏名 中前 和恭

マイクロホモロジー媒介末端結合 (Microhomology-mediated end-joining: MMEJ) は、非同相末端結合 (Non-homologous end-joining: NHEJ) や相合同組換え (homologous recombination: HR) 修復と並んで、DNA 二本鎖切断 (double strand break: DSB) 時の主要な修復経路として知られている。とりわけ近年発展が著しい Clustered regularly interspaced short palindromic repeats (CRISPR)-Cas9 システムをはじめとしたゲノム編集技術においては、部位特異的な認識と切断によって起こる DSB と MMEJ を誘導するマイクロホモロジー (μ H) 配列を利用して多様なゲノム変化が展開されている。中でも当研究室が開発してきた MMEJ を介した遺伝子ノックイン法 (Precise integration into target chromosome (PITCh) 法) 及び Local accumulation of DSB repair molecules (LoAD) 法) は高効率かつドナー作製が簡便という長所を備えており、様々な培養細胞株やヒト、マウス及びゼブラフィッシュ等の生物種でその有効性が証明されている。しかしながら、PITCh 法及び LoAD 法においてどのような配列デザインが最適であるかという問題は、必要な実験的解析とあわせて議論されたことがなかった。そこで本研究では、MMEJ を介したノックインの効果に関わる傾向を明らかにするために、配列設計から解析までを画一的に実施するソフトウェアパイプラインを開発した。本研究では、数十の実験データを解析に供することで MMEJ ノックインの傾向を解明することを試みた。

第一章では、最適なノックインについて議論する前段階として、基本的な MMEJ ノックインの配列デザイン法の定義を行なった。その定義のためのアプローチとして MMEJ ノックインを画一的に自動設計するためのソフトウェア「PITCh designer」を JavaScript+HTML5 で開発した。PITCh designer は任意の標的配列の中からノックインのデザインが可能な領域を検索し、Single-strand guide RNA (sgRNA) や μ H、及びドナー構築に必要なプライマー情報を提供する。これによって基本的な MMEJ のノックインデザインをコンピュータ・アルゴリズムのレベルで定義することができた。それとともに MMEJ を利用したノックインのデザインの利便性の向上に貢献できたと考えられる。

第二章では、MMEJ ノックインの傾向を明らかにするための次世代シーケンス (NGS) 解析パイプライン「MMEJ-assisted chromosomal integration analysis tools (MaChIAto)」を開発した。MaChIAto は既成の変異解析ツール「CRISPResso」(Pinello *et al.* 2016; Lindsay *et al.*

2016)と連携してより包括的な変異解析及び数百種にわたる特徴解析を可能とする。また変異分類をより正確なものへと再分類する機能も搭載しており、その性能は従来の分類ツールの中でもトップクラスの正確性を示すことを確認した。この MaChIAto の機能実証を目して、「PITCh designer」で設計した 40 種類の遺伝子座に対する MMEJ ノックインデータを MaChIAto の解析に供した。その結果、40 遺伝子座の全てにおいて PITCh 法及び LoAD 法による正確なノックインシーケンスを検出することができた。また PITCh 法でのノックインの正確性は CRISPR-Cas9 によるインデル形成能力に依存せず、ドナーがもつ μ H 配列の熱力学的特性、二次構造因子、あるいは配列コンテキストや標的周辺へのアクセス性に依存していることが明らかとなった。また LoAD 法では類似した傾向を示す一方で、PITCh 法よりもこれら因子に対する依存性が緩和されていることが示された。最後に、インデルパターンとノックイン効率に着目したところ、高効率なノックインを示した遺伝子座グループでは三塩基以上の欠失が大半であるのに対し、低効率なノックインを示した遺伝子座グループでは一塩基挿入が高頻度に観察される傾向が見られた。これに関して配列的な特徴に注目してみると、高効率なノックインを示した遺伝子座グループは G リッチな μ H を介した MMEJ が比較的多く観察され、さらに三塩基周期 (6、9、12 bp 等) の欠失が顕著に確認された。一方で低効率なノックインを示した遺伝子座グループの中のいくつかの遺伝子座では CRISPR-Cas9 による切断面の T 塩基のような一塩基誘導性をもつ配列特徴を有していた。これらの結果から、MMEJ ノックインの効率は μ H 内の配列存在比のみならず配列の配置にも依存していることが示唆された。以上第二章の結果から MaChIAto の利用によりノックインの傾向の理解につながる様々な特徴を推測できることが示された。

以上、本研究では、MMEJ ノックインのデザインから解析に至るまでの一連のソフトウェア・パイプラインを開発した。これらのツールを利用することで PITCh 法と LoAD 法の傾向を解析できることを証明し、(エピ)ゲノムコンテキストとインデルパターンがノックインの効果に寄与していることを示唆した。本研究は MMEJ ノックインの傾向を初めて記述した基盤的な成果であるだけでなく、より数理的な視点でゲノム工学技術開発を進めるためのプラットフォーム形成にも寄与したと考えられる。