

新規トランススプライシング機構を有する遺伝子の検索ツール開発と評価

杉崎 真 (筑波大学 生物学類) 指導教員: 橋本 哲男 (筑波大学 生命環境系)

《背景》

寄生性の真核生物 *Giardia intestinalis* では、特殊なトランススプライシングによって成熟 mRNA を生成する遺伝子が報告されている。このスプライシングでは、ゲノム上で非常に離れた領域から独立に転写されたエキソンが、付随する分割型イントロン (SPLit INTRON: splintron) を介して結合することで mRNA が成熟する (図 1)。これまでに他生物で splintron の報告はなく、*G. intestinalis* においても 3 遺伝子 (Dynein Heavy Chain, heat shock protein 90, p68 helicase) が報告されるに止まっている。Splintron が *G. intestinalis* 独自の機構かどうかは疑問の余地があるが、splintron をゲノムワイドに検出する方法はこれまでなかった。本研究の究極的な目的は、他生物での splintron の探索・解析を通じて、この機構の成立に至る進化的背景を解明し、真核生物のゲノム進化における新たな知見を得ることである。

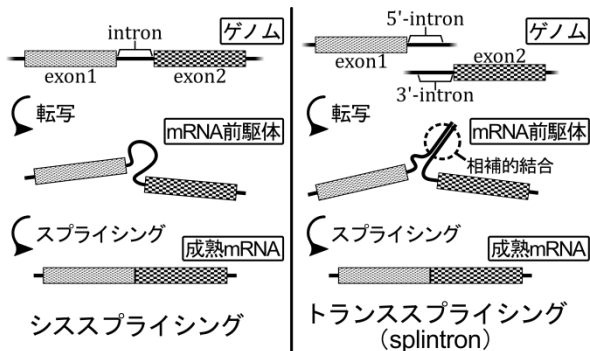


図 1

本研究では、*G. intestinalis* をテストデータとし、対象生物のゲノム及びトランスクリプトーム・データを元に splintron をゲノムワイドに検索するツールを開発し、既知の splintron を検出できるかどうか評価した。さらに、*G. intestinalis* 以外の生物でツールの実用性を検討するため、*G. intestinalis* の属する Fornicata 生物群で、系統的に *G. intestinalis* より早期に分岐する自由生活性生物 *Kipferlia bialata* のゲノム及びトランスクリプトームを用い、ツールの結果に対する実験的検証を行った。

《手法》

《ツールの開発》

目標とするツールの動作は、トランスクリプトーム上で一つの成熟 mRNA を構成する塩基配列がゲノム上の非常に離れた領域に由来することを確認した場合、その領域を splintron をもつ遺伝子の候補として検出し、さらに候補としての優位性をスコアリングすることである。これらの動作に必要な、ゲノム配列に対するトランスクリプトーム・データのマッピングには既存のツールである Burrows-Wheeler Aligner (BWA) を用いた。開発したツールは、*G. intestinalis* で splintron のもつ 3 つのタンパク質遺伝子に含まれる 8 領域を検出し、且つより高いスコアを付けるアルゴリズムを採用し、図 2 の式によってスコア (S) を算出した。*G. intestinalis* のデータは、公開されているゲノム・データ (GiardiaDB; <http://giardiadb.org/giardiadb/>)、及び所属

研究室で Illumina Hi-Seq の paired-end reads によって取得したトランスクリプトーム・データを用いた。

$$S = 1 - \frac{N_A + N_B - (n_A + n_B)}{n_A + n_B}$$

A: pair が非常に離れた領域にマップされる read が集中した領域
 B: A にマップされる read の pair が最も多くマップされた領域
 N_A : A にマップされた read の総数
 N_B : B にマップされた read の総数
 n_A : A にマップされた read のうち pair が B にマップされる read の総数
 n_B : B にマップされた read のうち pair が A にマップされる read の総数

図 2

《*Kipferlia bialata* を対象とした解析》

K. bialata のゲノム、トランスクリプトーム・データは所属研究室の先行研究によるデータを用いた。さらに、ツールの実行で高いスコアを得た、splintron の存在が予想される領域について RT-PCR 実験を行い、離れたゲノム領域に由来する転写産物が実際に一つの成熟 mRNA を構成するか、実験的に検証した。

《結果・考察》

《*Giardia intestinalis* のデータを用いたツール開発》

G. intestinalis の配列データセットからツールによって検出された 30 領域のうち、splintron をもつ 3 遺伝子 (Dynein Heavy Chain, heat shock protein 90, p68 helicase) に含まれる 8 領域すべてが確認された。各々の領域は 1, 2, 5, 6, 7, 8, 15, 16 番目に高いスコアとなった。また、残る 22 領域は variant-specific surface protein、或いは ankyrin repeat domain を含むキナーゼとして同定されており、ゲノム内で類似した配列が多数存在するため、塩基配列の類似性によるランダムなマッピングの結果、誤って推定されたと考えられる。一方で、splintron をもつ領域と同程度の高スコアで推定された領域については、新規の splintron の可能性があるため今後実験的検証が必要である。

《*Kipferlia bialata* を用いた評価・検証》

K. bialata の配列データセットに対し、3853 領域の splintron 候補が検出された。そのうちスコアが特に高い 280 領域の中から、偽陽性の可能性が高い領域を手動で除外し、残る 8 領域についてトランススプライシングの有無を確認する RT-PCR 実験の候補とした。8 領域それぞれのスプライシングによる結合を仮定し、得られた 4 種の塩基配列について BLAST による相同性検索を行った結果、cyclophilin-type peptidyl-prolyl cis-trans isomerase, protein KTI12, glutamine amidotransferase, S-adenosylmethionine-dependent methyltransferase との相同性が示唆された。分割されたエキソンの一方に forward プライマーを、もう一方に reverse プライマーを設計し RT-PCR 実験を行った結果、予想長の PCR 産物が得られ、塩基配列決定により目的の配列であることを確認した。以上から本研究で開発したツールが、splintron を *G. intestinalis* 以外の生物においても検出可能であると評価した。また、実際に他の生物でこの機構の存在を確認したことにより、生物種間で splintron の特性を比較する議論が可能となった。今後は、広範な真核生物のデータを用いた検出・検証を進め、さらなる知見を得ることを目指す。