

令和 2 年 7 月 2 日現在

機関番号：34204

研究種目：基盤研究(C) (一般)

研究期間：2017～2019

課題番号：17K00419

研究課題名(和文)非モデル生物における条件依存型選択的スプライシングの網羅的発見手法の開発

研究課題名(英文) Exhaustive detection of condition-specific splicing variants in non-model organisms

研究代表者

米澤 弘毅 (Kouki, Yonezawa)

長浜バイオ大学・バイオサイエンス学部・助教

研究者番号：00374744

交付決定額(研究期間全体)：(直接経費) 3,500,000円

研究成果の概要(和文)：本研究では、非モデル生物における条件依存の選択的スプライシングの網羅的発見手法の評価および改良を行うとともに、様々な非モデル生物のRNA-seqデータを用いた条件依存の選択的スプライシングの抽出を行い、当該生物の特殊性との関連を探る手がかりを得ることが目的である。研究期間内において、(1)条件依存的な選択的スプライシング網羅的発見手法の開発および評価、(2)実際のヒトの様々な組織に関するRNA-seqデータに対して提案手法の適用を行い、脳特異的な選択的スプライシングを抽出した。また、シングルセルトランスクリプトーム解析データセットからの選択的スプライシング網羅的発見への本手法の適用を検討した。

研究成果の学術的意義や社会的意義

非モデル生物の条件依存型選択的スプライシングについては研究がそれほど進んでいない。これはモデル生物におけるリファレンスゲノムが無いことが一因であるが、これにより非モデル生物の特殊性についてのさらなる知見が得られることが予想される。またマリアナイトエラゴカイの温度耐性のように、ヒトの反応メカニズムを解明するための手掛かりを得ることも可能となる。また本手法では、従来の解析手法では困難だった巨大なRNA-seqに対して適用すること、さらに3つ以上の異なる条件下のRNA-seqデータをまとめて取り扱うことを可能にしている。

研究成果の概要(英文)：Alternative splicing is a mechanism to generate more than one mRNA isoforms from a single locus, and it increases the genetic diversity during post-transcriptional gene regulation. Furthermore, alternative splicing is often differentially regulated across tissues and during development. It suggests that each splicing isoform may have specific spatial and temporal roles in life system.

We have developed the differential alternative splicing variants estimation method. To assess applicability of our method using RNA-sequence data for estimation of conditional-specific alternative events, we used several RNA-seq datasets. We confirmed that the newest version of our method could extract many candidates of condition-specific alternative splicing variants.

研究分野：バイオインフォマティクス

キーワード：選択的スプライシング 条件依存的

様式 C-19、F-19-1、Z-19 (共通)

1. 研究開始当初の背景

各生物種によって遺伝子数に限りがある中で、その生物種が様々な環境下に対応していくための遺伝的多様性の保持には、選択的スプライシングが大きな役割を担っている。また選択的スプライシングは、組織や発生段階によって発現の様子が変化していることもよく知られている。選択的スプライシングのメカニズムは大腸菌からヒトまで広範囲の生物で確認されており、その影響は生命システムに深く関連している。例えばヒトの場合、95%以上の遺伝子が選択的スプライシングを受けて複数のアイソフォームを産んでおり、それらは器官や発生段階によって異なる機能を持っていることも知られている。

研究開始当初までに選択的スプライシングに関する研究は多数なされていたが、その大半はヒト、マウス、ハエなどモデル生物に関する研究であり、非モデル生物における選択的スプライシングの研究に関しては現在でもそれほど盛んではない。通常のスプライシングバリエーションを発見する手法はリファレンスとなるゲノム配列とアノテーション情報を必要としていることが理由として挙げられるが、選択的スプライシングの進化への影響や保存メカニズムを知るためには、より多くの生物種における選択的スプライシングの性質を明らかにすることが必要である。

2. 研究の目的

1. の問題に対して研究代表者らは、Trinity [1]による RNA-seq データを入力とし、リファレンスとなるゲノム配列がない非モデル生物に対する条件依存の選択的スプライシングの網羅的発見手法(Differential Alternative Splicing variants Estimation; 以降 DASE と略記) [2]を既に開発していたが、研究期間内においてその改良および評価指標の確立、さらに種々の RNA-seq データに対する DASE の適用を目指した。

3. 研究の方法

2. の目的を達成するため、研究期間内に以下の2点を達成する。

(1) 条件依存の選択的スプライシング網羅的発見手法の開発および評価

研究代表者らが設計した DASE と他の選択的スプライシング網羅的発見手法との比較実験を行う。また研究をしていく上で、転写産物の配列全体よりも同じ遺伝子由来の転写産物同士で異なる部分に着目すると、条件依存の選択的スプライシングの抽出性能を向上させることがわかったので、この点を考慮した提案手法の改善を目指した。

DASE の概要を図 1 に示す。DASE は異なる条件で得られたリードデータを混ぜて Trinity で de novo アセンブルを行い、異なる条件下の発現量(図 1 の 3)と MAFFT [3]によるマルチプルアライメントから得られる転写産物の塩基配列の一致度(図 1 の 4)を組み合わせて両者の 2 乗平均距離を計算し(図 1 の 5)、同じ遺伝子由来の他のコンティグとの距離が大きいコンティグを順に列挙する(図 1 の 6)。

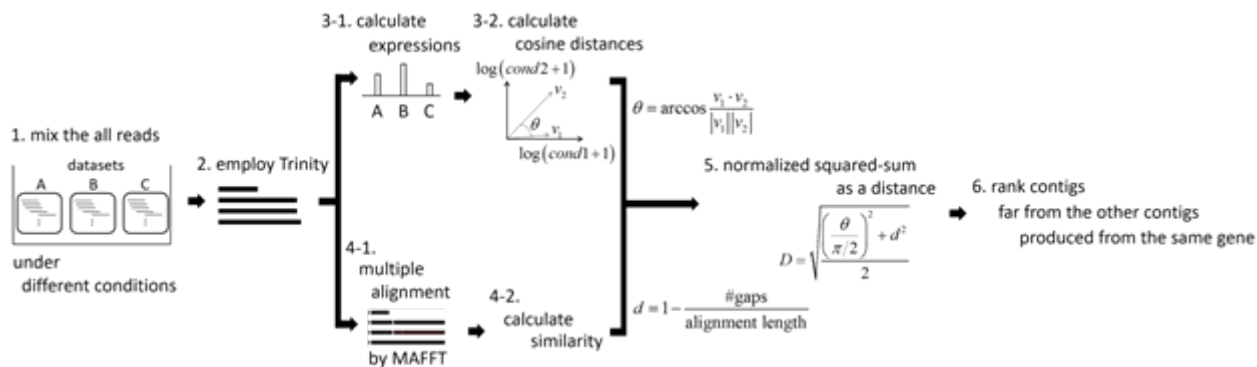


図 1: DASE メソッドの概略。

(2) 実際の RNA-seq データへの DASE の適用および生物学的知見のマイニング

クロレラやイカといった非モデル生物の RNA-seq データセットに提案手法である DASE を適用し、生物学的知見のマイニングを行った。

4. 研究成果

(1) 条件依存の選択的スプライシング網羅的発見手法の開発および評価

条件依存性を測る際の指標を変更することで提案手法の網羅的発見性能が良化することを示した。また、ヒトの様々な組織における RNA-seq データに対して提案手法を適用し、脳特異的な選択的スプライシングの抽出ができたことを確認した[4]。しかし[4]で提案した手法ではアーティファクトを疑われる転写産物や、5' 側や 3' 側が切れていて実験で確認しづらい転写産物が多く抽出されることが明らかになった。そこで、転写産物の配列全体よりも同じ遺伝子由来の転写産物同士で配列の 5' 末端や 3' 末端にない異なる配列構造に着目して提案手法の改良を行

った[5]. マウスの Rett 症候群に關係する RNA-seq データセット[6]で検証実験を行ったところ、条件依存的な選択的スプライシングの候補となる遺伝子数が 13%程度まで減少したことで実験担当者が検証しやすくなった。また、[6]で詳述されている条件依存的な遺伝子が、改良した手法ではより上位にランク付けされることも確認した。

(2) 実際の RNA-seq データへの DASE の適用および生物学的知見のマイニング

イカの筋肉と神経組織における RNA-seq データに対して提案手法を適用することで、イカの手から出される接着分子に関わる遺伝子の選択的スプライシングの候補を発見している。しかしながらイカを用いた実験が事前の想定よりも困難なことがわかり、現在まで実験による裏付けはなされていない。

また、シングルセルトランスクリプトーム解析データセットからの選択的スプライシング網羅的発見への本手法の適用を検討している。

引用文献

- [1] Grabherr et al., Nat. Biotech., 29(7), 644-652, 2011.
- [2] Yonezawa et al., Proc. 2016 IEEE Conference on Bioinformatics and Biomedicine (BIBM2016), 2016.
- [3] Katoh et al., Nucl. Acid. Res., 30(14), 3059-3066, 2002.
- [4] Yonezawa et al., Int. J. Data Mining and Bioinformatics, 18, 56-73, 2017.
- [5] Yonezawa et al., Bioinformatics: From Algorithms to Applications (BiATA 2019), 2019.
- [6] Osenberg et al., Proc. Natl. Acad. Sci. U. S. A., 115(23), E5363-E5372, 2018.

5. 主な発表論文等

〔雑誌論文〕 計1件（うち査読付論文 1件 / うち国際共著 0件 / うちオープンアクセス 0件）

1. 著者名 Kouki Yonezawa, Keisuke Nakata, Ryuhei Minei, Atsushi Ogura	4. 巻 18
2. 論文標題 DASE2: differential alternative splicing variants estimation method without reference genome, and comparison with mapping strategy	5. 発行年 2017年
3. 雑誌名 Int. J. Data Mining and Bioinformatics	6. 最初と最後の頁 56-73
掲載論文のDOI（デジタルオブジェクト識別子） 10.1504/IJDMB.2017.086101	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -

〔学会発表〕 計2件（うち招待講演 0件 / うち国際学会 0件）

1. 発表者名 米澤 弘毅, 嶺井 隆平, 小倉 淳
2. 発表標題 非モデル生物における条件依存的選択的スプライシングの網羅的発見手法の改良 -発現量に関する"外れ度"の定量化による精度向上-
3. 学会等名 第50回バイオ情報学研究会
4. 発表年 2017年

1. 発表者名 Kouki Yonezawa, Ryuhei Minei, Atsushi Ogura
2. 発表標題 DASE-AG: conditional-specific differential alternative splicing events estimation method for around-gap regions
3. 学会等名 Bioinformatics: From Algorithms to Applications (BiATA 2019)
4. 発表年 2019年

〔図書〕 計0件

〔産業財産権〕

〔その他〕

-

6. 研究組織

氏名 (ローマ字氏名) (研究者番号)	所属研究機関・部局・職 (機関番号)	備考
---------------------------	-----------------------	----