

科学研究費助成事業 研究成果報告書

平成 28 年 6 月 10 日現在

機関番号：33916

研究種目：基盤研究(C) (一般)

研究期間：2012～2015

課題番号：24510271

研究課題名(和文)イントロン長分布の多角的解析によるヒト・スプライシング機構の進化的解明

研究課題名(英文)Bioinformatic and molecular biological study on ultra-short introns

研究代表者

嶋田 誠 (SHIMADA, Makoto)

藤田保健衛生大学・総合医科学研究所・講師

研究者番号：00528044

交付決定額(研究期間全体)：(直接経費) 4,300,000円

研究成果の概要(和文)：ヒト遺伝子のイントロンの中に、従来考えられていなかった65塩基長以下のものを発見し、微小イントロンとして発表した。

実験的な裏付けをもつ転写物配列調査の結果、ヒトで見つかった微小イントロンと起源を共にするイントロンがヒト以外の生物で見つかったのは7微小イントロンであった。それ以外の微小イントロンはヒト以外の転写物配列情報が少ないのみならず、ゲノム配列の面から検討してもイントロン両端を示す配列が近縁種で保存されていない場合が多かった。そのため、微小イントロンの多くは起源が新しいか発現が限定的なものであると考えられる。

研究成果の概要(英文)：We found introns that are shorter than previously expected and published them as "ultra-short introns."

Because exon/intron boundary is dynamically defined, we evaluated the ultra-short introns using transcript and genome sequences of vertebrate species as well as human. We found seven orthologous ultra-short introns in experimentally validated transcript sequences. In case of other ultra-short introns, few transcript sequences are available in non-human species. Moreover, splice di-nucleotides sequences of these ultra-short introns scarcely conserved in genome sequences. These results suggest that at least some of ultra-short introns have been appeared within human or closely related species and/or are specifically expressed depending on cell types or development stages.

研究分野：ゲノム進化学

キーワード：イントロン スプライシング 進化

1. 研究開始当初の背景

我々は、ヒト遺伝子のイントロンの中に従来考えられていなかった 65 塩基長以下のものを発見し、22 の微小イントロンとして発表した。

ところで、一般にイントロン/エクソンの定義はダイナミックなものであり、多くの遺伝子が選択的スプライシングと呼ばれる機構によって、エクソンとイントロンの組合せを変化させ複数種類の遺伝子産物を産生する。そのため、一口にイントロン長で区切って定義した微小イントロンそれぞれについて、それらがいつどのような細胞で発現されるのか実態を把握する必要がある。それにはこれらのセットを別の方面から再検討することが重要である。

2. 研究の目的

- (1) 微小イントロンの発現の実態を解明する。
- (2) ヒト遺伝子のイントロン長と人類進化上の特徴との関係を解明する。

3. 研究の方法

(1) ヒト遺伝子データベース H-InvDB (<http://hinv.jp>)のアノテーション情報に則り、イントロンとしてゲノム上にマップされた転写物の配列やその信頼性を様々な方面から検討した。

(2) 公共データベース (NCBI,EMBL,DDBJ)上の転写物配列やゲノム配列を用いて、我々が定義した ultra-short intron (微小イントロン)の相長的配列を探索し、配列比較を行った。

4. 研究成果

(1) イントロン配列特性からの分類
スプライシングに必要な中心的配列要素の解析：

ヒトのイントロン配列では短いイントロンほど GC 含量が高くなることが知られているが、我々の微小イントロンに関しても、同様の傾向を認めた。特に 22 微小イントロンの約半数の 12 イントロンにおいて G 塩基が最も高頻度であった。これらはすべて G 塩基の 3 塩基以上の反復配列を含んでいた。さらに我々の先行研究で G に富む 11 塩基の配列が微小イントロンのスプライシング効率を上げるイントロン内スプライシング・エンハンサー(ISE)を同定したが、これと同じか高度に似た配列を持つイントロンが計 4 つ確認できた。

上記と異なり、G 塩基に富まないタイプのイントロンを認めた。それらの中には一般的にはスプライシングのシグナル配列要素であるランチサイトとピリミジン配列を含むものが多かった。

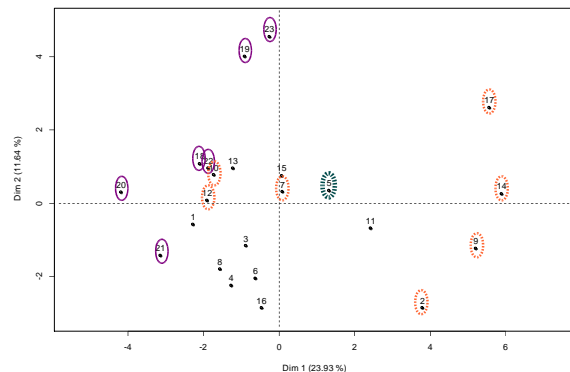
また、スプライシングのシグナル配列要素の強度を SROOGLE web tool によってスコア化して、比較した。その結果、我々が微小イントロンを選抜した際にイントロン両

端 (i.e., 5' サイトと 3' サイト)の GT-AG 配列を含むことを条件に選抜したにもかかわらず、微小イントロンすべてにおいて、3 つのスプライシングに必要な中心的配列要素 (5' サイト、3' サイト、ランチサイト)を全て持ち合わせているイントロンは無かった。さらに 3 要素それぞれにおいて、スコアが 0 もしくは 0 に近い、すなわちスプライシング・シグナルの配列には全く類似していない、イントロンが複数認められた。そのような配列であっても、ヒト由来の細胞に対して行った RT-PCR 実験により、微小イントロンの位置でスプライシングが行われた内在性 mRNA の存在を示すイントロンが認められた。

(2) 諸特徴による多変量解析による分類

我々が定義した 22 の微小イントロンと微小イントロンの可能性のある一つのイントロンを加えて計 23 イントロンの間で、諸特徴の共有具合からグループ化できるか、またどのような特徴がグループを特徴づけているかを調べるために主成分分析を行った(図 1)。

図 1. イントロンの特徴による主成分分析



評価した特徴には塩基配列構成上の特徴、スプライシングに必要な配列要素の有無や強さ、実験的証拠の得やすさ、微小イントロンが発見された転写物配列における相対的位置情報やその他のイントロンに関する情報、微小イントロンを含んだ遺伝子の機能、進化的保存性などの情報を総合的に組み込んで解析を行ったところ、次のような傾向が示唆された。

A 塩基の頻度が高いイントロンは 5' ESE の密度が高い傾向がある (dim1)。これは、我々の先行研究 (Sasaki-Haraguchi, Shimada, & Mayeda 2012) でのスプライシングに周囲のエクソン配列が必要であるとする実験報告を裏付けるものである。

G 塩基の頻度が高いイントロンはそれとは別の機構により、短いながらもスプライシングが起こった (dim1) ことを示唆する。

Splice Signal 強度と逆相関している splicing regulator 密度がある (たとえば、3. SC35, dim2)。これはスプライシング強度が強い ultra-short intron は SC35 のような ESE に依存

しないスプライシングによって除かれたことを示唆する。

(3) 微小イントロンの存在意義や役割の検討：

もし微小イントロンの領域がイントロンとして認識されずエクソンの一部となった場合、塩基配列より読み枠のずれ等により、停止コード (PTC) が出現し、NMD により分解されることが予想されるのは、22 微小イントロンのうち 15 個であった。しかもそのうち、11 個はイントロンとして認識される実験的証拠をもつ 15 微小イントロンのカテゴリに入っている。したがって、これらのことはイントロンとして認識される生体機能的意義があると考えられる。

また、実際に同じ遺伝子座より転写される別の既知スプライシング変異体配列との間で翻訳されるアミノ酸配列の違いを比較検討したところ、つぎのような 3 種類の興味深い例が観察された。

両者の間で翻訳開始点とコドン読み枠フレームが異なるものの、片方の読み枠では翻訳開始点から 15 アミノ酸目 (45 塩基目) で 55 塩基長の微小イントロンの位置をまたぐことにより読み枠が変更され、それ以降の下流ではもう片方のスプライス変異体と同じ読み枠になる例。

全長 1257 塩基長のイントロン領域中に、84 塩基長のスキップ・エキソンと 50 塩基長の微小イントロンが同伴することによって前後の読み枠を他のスプライス変異体と同様にそろえる役割をしている例。

終止コドンの上流 13 アミノ酸の位置に微小イントロンが認識され、アミノ酸配列の C 末端部に他のスプライス変異体産物とは異なるアミノ酸配列をもつが、機能ドメインに影響しない位置であると考えられる例。

(4) 進化的保存性

ゲノム配列と転写物配列との間の違いとその意味：

公共データベース中のヒト以外の転写物には推定配列が多く、実際に実験的に得られたことが確実な転写物が少ない。そのために転写物配列で種間配列比較ができない場合も認められた。その場合は、ゲノム配列の種間比較によって、イントロン両端の GT-AG 配列の保存性・共有性を確認して、イントロンとして認識している分類群を推定した。この解析によって、ヒトの転写物配列およびゲノム配列データでは GT-AG 型のイントロンであることが推定できても、そこをイントロンとして使っていない転写物配列が一定数存在する事が分かった。ヒトで見つかった微小イントロンと起源を共にするイントロンがヒト以外の生物で見つかったのは 22 微小イントロンのうち 7 微小イントロンであった。また、ゲノム配列比較により、ヒトの微小イントロン全てで共有している両端の

GT-AG 配列は単純に進化のある時期に一度獲得した GT-AG 配列がすべての系統で維持されているというシナリオではなく、ヒト以外のいくつかの系統では失われていることが観察された。このことは微小イントロンが進化的に一般的なイントロンとは今の段階では認められないことを示している。

パターンの概説とその意味：

上記の解析の結果は次のような 4 パターンに分類できた。

(i) ゲノム配列で GT-AG がそろっていない。
(ii) ゲノム配列で GT-AG が揃っているが、エクソンとして認識されている。

(iii) ゲノム配列で GT-AG が揃っているが、もっと長いイントロン領域の一部となっている。

(iv) ゲノム配列で GT-AG が揃い、イントロン両端として認識されている。

上記(iv) が出現することを示せた微小イントロンはほとんど霊長類の種分化以降であった。

これらのことにより、ヒトで微小イントロンとされているものは、多くの生物種で必要なイントロンというわけではない実態を明らかにしたと言える。

あるいは、人類の特徴であることを示しているのかもしれない。

また、ゲノム配列・転写物配列の種間比較情報、ヒトの RNA-seq 情報、様々な情報を組み合わせることにより、実験系に適したイントロンが少なかった実態を示している。

5. 主な発表論文等

(研究代表者、研究分担者及び連携研究者には下線)

[雑誌論文](計 6 件)

Shimada MK, Sanbonmatsu R, Yamaguchi-Kabata Y, Yamasaki C, Suzuki Y, Chakraborty R, Gojobori T, Imanishi T: Selection pressure on human STR loci and its relevance in repeat expansion disease, *Molecular Genetics and Genomics* (in press), doi: 10.1007/s00438-016-1219-7, 査読有.

Shimada MK, Sasaki-Haraguchi N, Mayeda A: Identification and Validation of Evolutionarily Conserved Unusually Short Pre-mRNA Introns in the Human Genome, *International Journal of Molecular Sciences* 16(5): 10376-10388, 2015, doi:10.3390/ijms160510376, 査読有.

Shimada MK: An attempt to predict the functional haplotype differences between archaic and modern humans, *Research Bulletin (Saito Ho-on Kai Museum)* 78: 77-85, 2014, 査読無.

Gough CA, Homma K, Yamaguchi-Kabata Y, Shimada MK, Chakraborty R, Fujii Y, Iwama H, Minoshima S, Sakamoto S, Sato Y, Suzuki Y, Tada-Umezaki M, Nishikawa K, Imanishi T, Gojobori T: Prediction of protein-destabilizing polymorphisms by manual curation with protein structure, *Plos One*, 7(11): e50445, 2012,

doi: 10.1371/journal.pone.0050445, 査読有。
Shimada MK: Survival for All: Let's Share Benefits and Hardships, *International Journal of Evolution* 1:1. 2012, doi:10.4172/2324-8548.1000e103, 査読無,招待。

Sasaki-Haraguchi N, Shimada MK, Taniguchi I, Ohno M, Mayeda A: Mechanistic insights into human pre-mRNA splicing of human ultra-short introns: Potential unusual mechanism identifies G-rich introns, *Biochemical and Biophysical Research Communications* 423(2): 289-294, 2012, doi:10.1016/j.bbrc.2012.05.112, 査読有。

〔学会発表〕(計 10 件)

嶋田 誠, 三本松 良子, 山口 由美, 山崎千里, 鈴木 善幸, Chakraborty, Ranajit, 五條堀孝, 今西 規: ポリグルタミン病発症機序には共通して長いグルタミン反復が要求される過程を含んでいる、日本遺伝学会 第 87 回大会、2015 年 9 月 24-26 日、東北大学、仙台。

Shimada MK, Sasaki-Haraguchi N, Mayeda A: Ultra-short G-rich introns with completely inefficient 5' and 3' splice sites are spliced in vivo, but how? 2015 Cold Spring Harbor Laboratory Meeting "Eukaryotic mRNA processing", Aug. 18-22, 2015, Cold Spring Harbor, NY, USA.

嶋田 誠, 佐々木(原口)典子, 前田 明: ヒト微小イントロンは 4 種類の進化様式で出現した、第 16 回日本 RNA 学会年会、2014 年 7 月 23-25 日、ウインクあいち、名古屋。

嶋田 誠, 三本松良子, 山崎千里, 山口由美, 五條堀 孝, 今西 規: ヒト転写物と多型のデータベース統合による、グルタミン反復配列長大化への進化的解析、日本遺伝学会 第 85 回大会、2013 年 9 月 19-21 日、慶応大学、日吉。

嶋田 誠: 1000 Genomes データを用いた古代人からの流入配列候補の比較、NGS 現場の会第三回研究大会、2013 年 9 月 4-5 日、神戸国際会議場神戸。

嶋田 誠, 佐々木(原口)典子, 前田 明: Analysis of the intron distribution of the human genes implicates the existence of ultra-short introns (ヒト遺伝子のイントロン長分析から明らかになった微小イントロンの存在)、第 35 回日本分子生物学会年会、2012 年 12 月 11-14 日、福岡国際会議場・マリンメッセ福岡、福岡。

Shimada MK: An attempt to estimate the functional difference in haplotypes between archaic and modern humans. *International Conference on Replacement of Neanderthals by Modern Humans: Testing Evolutionary Models of Learning (RNMH2012)*, Nov. 18-24, 2012. National Institute of Informatics, Tokyo, Japan.

嶋田 誠, 佐々木(原口)典子, 前田 明: ヒトにおける規格外に短いイントロンの発見とその特徴、第 14 回日本 RNA 学会年会、2012 年 7 月 18-20 日、東北大学、仙台。

嶋田 誠, 佐々木(原口)典子, 前田 明: 外れ値にあたる短さをもつヒト・イントロンの進化的保存性、日本霊長類学会第 28 回学術大会、

2012 年 7 月 6-8 日、椋山女学院大学、名古屋。

嶋田 誠: ゲノムデータと学習仮説 (Genome data and the learning hypothesis for the replacement of Neanderthals by modern humans)、国立遺伝学研究所研究会「人類集団の進化的起源と文化的分化要因～学習戦略による旧人と新人の交替劇に関連して～」、2012 年 4 月 13 日、国立遺伝学研究所、三島。

〔その他〕

ホームページ等

- (1) Labarchives[転写物配列データへの相同性検索結果]
<http://dx.doi.org/10.6070/H4V985ZZ>
- (2) Labarchives[ゲノム配列データへの相同性検索結果]
<http://dx.doi.org/10.6070/H400001W>
- (3) researchmap [研究者情報、資料公開]
<http://researchmap.jp/shimada-mk/>

6. 研究組織

(1) 研究代表者

嶋田 誠 (SHIMADA, Makoto)
藤田保健衛生大学・総合医科学研究所・講師
研究者番号: 0 0 5 2 8 0 4 4

(2) 研究分担者

()

(3) 連携研究者

前田 明 (MAYEDA, Akila)
藤田保健衛生大学・総合医科学研究所・教授
研究者番号: 5 0 2 1 2 2 0 4

(4) 研究協力者

佐々木(原口)典子 (SASAKI-HARAGUCHI Noriko)
藤田保健衛生大学・総合医科学研究所・博士研究員(当時)
研究者番号: 9 0 5 4 6 6 4 9