

In silico 解析と in vitro 解析による RNA スプライシング機構の研究

●大野 欽司

名古屋大学大学院医学系研究科神経遺伝情報

<研究の目的と進め方>

ヒトゲノムプロジェクトが完了し、ヒトは22,000個という限られた数の遺伝子から組織特異的・発達段階特異的に splicing trans-factors を発現させ、alternative splicing を行い、10万種類以上と予想される多様なタンパク質を実現していることが明らかになり、splicing の生理・病態機構の解明は従前にも増して重要になってきている。

本研究の目的は、ヒトゲノム上に annotation をされた全 transcriptome の in silico 解析と培養細胞を用いた in vitro 解析により各 splicing cis-element 内、及び、各 cis-elements 相互の塩基配列の法則を解明することである。さらに、各 cis-element 内、及び、cis-elements 間の法則を明らかにすることにより degeneracy の統合的な代償機構を解明し、splicing 異常を起こす遺伝子変異を予測する精度の高い汎用アルゴリズムを構築することである。

1. Splice donor site の予測アルゴリズムの web service

U1 snRNA が結合をする splice donor sites における遺伝子変異が splicing 異常を起こすことが従来より数多く報告されてきているが、現在でも見落とされている変異が多く存在する (personal observations)。2007年度までの研究にて、splice donor sites における塩基間相互の関係を明らかにし、splicing 変異を効率よく予測するアルゴリズムの構築を行った (Nucleic Acids Res, 35, 5995-6003, 2007)。本年度は支援班のサポートにより本アルゴリズムのウェブサービスの開始を検討する。

2. ヒト branch point コンセンサス塩基配列の決定

ヒトを含む高等動物の branch point コンセンサス配列は degenerative であるが、degeneracy が全く存在しない yeast の branch point 配列に基づいて予測アルゴリズムが作られている。自験例の検討によるとこのアルゴリズムは不正確である。またひとつのイントロンに複数の branch points が使われることが多く、branch point sequence の degeneracy の tolerance の定量化と正確な branch point 予測アルゴリズムは branch point sequence を破断する遺伝子変異の研究に重要である。本研究では、全 transcriptome 解析・lariat RT-PCR 法・ribonuclease protection assay 法によりヒトにおける branch point コンセンサス配列の決定を試みる。

3. Aberrant splicing を惹起するエクソン 5' 末端遺伝子変異の予測アルゴリズムの構築

Exon +1 位変異による aberrant splicing は従来 3 例しか報告されておらず見逃されている例が多いと想定される。従来ほとんど研究が行われていない U2AF65、及び、U2AF35 が結合をする intron 3' 末端と exon 5' 末端の塩基間相互の関係を明らかにし、AG-dependent splice sites と AG-independent splice sites の判別アルゴリズムを構築する。これらの解析結果に基づき、exon +1 位を含む exon 5' 末端領域の遺伝子変異の中で、splicing 異常を

起こすものを検出するためのアルゴリズムを構築する。

これらの解析を組み合わせるにより疾患関連遺伝子変異が splicing 異常を起こす機構の解明と、splicing 異常を惹き起こす遺伝子変異を予測する汎用アルゴリズムの構築を試みる。

<2008 年度の研究の当初計画>

1. Splice donor site の予測アルゴリズムの web service

昨年度までの本特定領域研究で得られた成果にもとづく SD Score アルゴリズムの web service を開始する。

2. ヒト branch point コンセンサス塩基配列の決定

ヒトにおける branch point コンセンサス配列を実験結果に基づいて決定を行うために house keeping genes を対象に 100ヶ所の branch points を lariat RT-PCR により同定を目指す。House keeping genes は各種細胞において高発現であり、intron が短く、alternative splicing を受ける exon が少ないために本研究の対象遺伝子として最適であると考えられる。これらの結果を training dataset として、branch point 予測アルゴリズムを構築する。ヒト全ゲノムを対象にこの予測アルゴリズムによる branch points の予測を行い、予測結果の検定を lariat RT-PCR により複数の遺伝子の複数の introns を用いて行う。

3. Aberrant splicing を惹起するエクソン 5' 末端遺伝子変異の予測アルゴリズムの構築

Spliceosome の構築に intron 3' 末端の AG が必要な AG-dependent splice sites と不必要な AG-independent splice sites が存在する。AG-dependent splice sites においては、intron 3' 末端の AG に加えて exon 5' 末端の G が U2AF35 の結合に重要であり、exon 5' 末端の G の遺伝子変異が aberrant splicing の原因になるという仮説を立て実証を行う。FECH と GHI では exon 5' 末端の G の変異が exon skipping を起こすが、LPL と HEXA では起こさない。それぞれの遺伝子の mini gene を作成し、各種 mutagenesis を行い exon skipping を調べることにより exon 5' 末端遺伝子変異が aberrant splicing を惹起するのに必要な cis-elements の配列要素を決定する。

<2008 年度の成果>

1. Splice donor site の予測アルゴリズムの web service

支援班のサポートを認めていただき、現在業者と SD Score アルゴリズムの web service の仕様策定を進めている。

2. ヒト branch point コンセンサス塩基配列の決定

ヒト 20 種類の house keeping genes の 52 の introns を対象に 367 個の lariat clones を解析した。367 個の clones のうち branch point に misincorporated nucleotides を認めたものは 181 clones で、これらを用いて解析を行った。残りの 186 個のクローンは lariat RT-PCR において branch point がスキップをした可能性を

否定できなかったため解析対象としなかった。Branch pointsの塩基は、92.3% A, 3.3% C, 1.7% G, 2.8% Uであった。解析の結果、ヒトの branch point コンセンサス配列は yUnAy であることが判明した。Branch points の83%は -34位から -21位に存在した。Polypyrimidine tractは branch point の下流4から24塩基に存在した。ヒトの branch point コンセンサス配列は、予想されるよりも degenerative であり、branch point は他の splicing cis-element(s) と同時に認識をされている可能性が高いと思われた(論文1)。

3. Aberrant splicing を惹起するエクソン 5' 末端遺伝子変異の予測アルゴリズムの構築

exon 5' 末端の G の変異が exon skipping を起こす *FECH*, *GHI, EYA1* 遺伝子において polypyrimidine tract (PPT) の延長によりスプライシングの正常化を実現できた。一方、遺伝子変異が exon skipping を起こさない *LPL* 遺伝子においては PPT を短縮させることにより exon skipping が誘発された。同様に、遺伝子変異が exon skipping を起こさない *HEXA* 遺伝子においては PPT の短縮にても exon skipping が誘発されず、さらなる PPT の短縮を行うとともに、U2AF65-independent splicing 機能が存在する可能性を考慮して検討を行っている。

<国内外での成果の位置づけ>

この1年間に類似の研究は国内外で行われていない。

一方、splice donor site の予測アルゴリズムの web service に関する海外からの要望を受け取っており早急なサービスの開始を予定している。

<達成できなかったこと、予想外の困難、その理由>

ヒト branch point コンセンサス塩基配列は予想外に degenerative であり信頼性のあるアルゴリズムの構築が困難であった。

<今後の課題>

Aberrant splicing を惹起するエクソン 5' 末端遺伝子変異の研究をさらに進め、splicing cis-element 破断遺伝子変異の効率的な予測アルゴリズムの構築をめざす。

<成果公表リスト>

1) 論文/プロシーディング

1. 0805161026

Gao K, Masuda A, Matsuura T, Ohno K. Human branch point consensus sequence is yUnAy. *Nucleic Acids Res* 2008, 36: 2257-2267.

2. 0805161030

Kurosaki T, Matsuura T, Ohno K, Ueda S. Long-range PCR for the diagnosis of spinocerebellar ataxia type 10. *Neurogenetics* 2008, 9: 151-152.

3. 0901161125

Shen X-M, Fukuda T, Ohno K, Sine SM, Engel AG. Congenital myasthenia-related AChR delta subunit mutation interferes with intersubunit communication essential for channel gating. *J Clin Invest* 2008, 118: 1867-1876.

4. 0901161134

Masuda A, Shen XM, Ito M, Matsuura T, Engel AG, Ohno K.

hnRNP H enhances skipping of a nonfunctional exon P3A in *CHRNA1* and a mutation disrupting its binding causes congenital myasthenic syndrome. *Hum Mol Genet* 2008, 17: 4022-4035.

5. 0901161138

Bian Y, Masuda A, Matsuura T, Ito M, Okushin K, Engel AG, Ohno K. Tannic acid facilitates expression of the polypyrimidine tract binding protein and alleviates deleterious inclusion of *CHRNA1* exon P3A due to an hnRNP H-disrupting mutation in congenital myasthenic syndrome. *Hum Mol Genet* 2009, in press

2) データベース/ソフトウェア

なし