

・研究種目：基盤研究 (B)
 研究期間：2006～2009
 課題番号：18370061
 研究課題名(和文) 選択的スプライシングを受けたタンパク質の立体構造モデリングによる機能解析
 研究課題名(英文) Studying Function of Alternative Splicing Products Based on Protein Structure Modeling
 研究代表者
 郷 通子 (MITIKO GO)
 長浜バイオ大学・バイオサイエンス学部・特別客員教授
 研究者番号：70037290

研究成果の概要(和文): ゲノムにコードされている遺伝子の数は予想よりも少なく、遺伝子の転写後に起きる選択的スプライシングによって、ひとつの遺伝子から複数個の転写産物ができることが明らかになっている。ひとつの遺伝子由来の複数のタンパク質にどのような構造の違いがあり、どのように機能が変化しているのかを実験的に調べることは困難である。そこでタンパク質の構造と機能の変化を計算生物学の手法で推定できるようにした。選択的スプライシングを起こす遺伝子を実験データから明らかにする方法を構築し、それらのタンパク質の構造を推定する方法を開発し、タンパク質の構造を推定することができた。立体構造の違いから、機能がどのように変化するのか、いくつかのタンパク質の場合について推定することができた。

研究成果の概要(英文): The number of genes in the genome was turned out to be unexpectedly small and alternative splicing (AS), a mechanism to produce more than one product out of a single gene, is expected to fill the gap in the numbers. It is difficult to experimentally elucidate the mechanisms of functional diversity of proteins produced by alternative splicing of the single genes. We built methods to identify AS products from the measured data, to predict structures of AS products (proteins), and to estimate difference of biochemical function of the products. The methods elucidated the differences in a number of cases.

交付決定額

(金額単位：円)

	直接経費	間接経費	合計
2006年度	5,600,000	1,680,000	7,280,000
2007年度	3,700,000	1,110,000	4,810,000
2008年度	3,000,000	900,000	3,900,000
2009年度	3,000,000	900,000	3,900,000
総計	15,300,000	4,590,000	19,890,000

研究分野： 生命情報学・生物物理学・進化学

科研費の分科・細目： 生物科学・生物物理学

キーワード： タンパク質立体構造・選択的スプライシング・ホモロジーモデリング・データベース・RNA編集・機能部位・モジュール・蛋白質間相互作用

1. 研究開始当初の背景

ヒトゲノムの決定により、2万5千程度の遺伝子がゲノム塩基配列中にコードされていることが明らかになった。これは当初予想されていた遺伝子数から極端に少ない。ESTとゲノム上の遺伝子座の対応関係の解析により、ヒトゲノムにコードされている遺伝子のうち40%ぐらいが、2つ以上の異なる成熟mRNAを生み出していることがわかってきた。mRNAの成熟過程において、状況に応じてあるエクソンをイントロンと見なして切り出し、同一のmRNAから複数の転写産物を得られることが、古くから知られている(選択的スプライシング)。エクソン境界のマイクロアレイにより、ヒトゲノムで選択的スプライシングが70%程度起こっていることが明らかになった(Johnson et al. Science, 302, 2141-2144, 2003)。成熟mRNAの多様化により、ひとつの遺伝子からアミノ酸配列が部分的に異なるタンパク質が翻訳され、その結果として多様な機能をもつタンパク質が生み出されることが期待される。

タンパク質の機能は通常、立体構造を形成することで実現される。選択的スプライシングにより生じるアイソフォームの立体構造解析は本研究開始までに数例なされており、ほとんどの場合は選択的スプライシングがループ部分の構造変化をもたらしていた。選択的スプライシングによるタンパク質立体構造の変化はループの挿入欠失程度と思われる。しかし、Piccolo C2A ドメインでは、選択的スプライシングによって、予想外の構造変化が起こっていることが判明した(Garcia et al. Nat. Struct. Mol. Biol., 11, 45-53, 2004)。短いアイソフォームではシートの一部を形成している部分が、長いアイソフォームではヘリックス構造を形成し、長いアイソフォームに挿入された9残基がシートの一部を形成するようになっている。選択的スプライシングによるアイソフォームの立体構造は、選択的に用いられるエクソン部分の単なる挿入欠失ではなく、タンパク質の全体構造に影響を及ぼす場合がある。

選択的スプライシングによって生産されるアイソフォームに、立体構造上どのような変化をもたらされ、選択的スプライシングがタンパク質の分子性および細胞性機能にどのような変化をもたらすのかを知ることは、ゲノムにコードされる遺伝子の機能の理解や、高等真核生物における選択的スプライシングの意義を理解するために必要である。しかし、選択的スプライシングがヒト遺伝子の40%から70%で起こっていることより、すべての選択的スプライシング産物であるタンパク質立体構造を実験的に決定することは、短い時間では不可能である。そのために、選択的スプライシングにより生じるアイソフォ

ームの立体構造を推定する手法の開発が重要である。

2. 研究の目的

ヒトゲノムにコードされている大部分の遺伝子では、mRNAの成熟過程において選択的スプライシング(以下AS)により、状況に応じてひとつの遺伝子から2つ以上の異なる成熟mRNAを生み出していることがわかってきた。そこで本申請研究では、ASによってできるタンパク質がどのような立体構造をとっているのかを計算生物学の手法で推定し、分子性機能および細胞性機能にどのような変化が起きるのかを推定することを目的とする。

3. 研究の方法

(1) ASデータベースの作成

ゲノムが決定されている生物種の全塩基配列に全長mRNAのデータを対応させることで、ASが起こることにより変化するタンパク質上の領域(AS領域)を同定するアルゴリズムを開発した。ゲノムが決定していない生物種については、タンパク質アミノ酸配列のデータベースであるUniProtKBにもとづいてAS領域の情報を収集するアルゴリズムを作成した。また、得られたAS領域とタンパク質の立体構造を対応させ、タンパク質立体構造安定性に関わる疎水性コア、および、立体構造情報から求めた他分子との相互作用部位とAS領域との関係を求めるアルゴリズムを開発した。さらに、これらのアルゴリズムにより得られた結果を検索し、AS領域と立体構造の関係を簡単に観察することのできる機能を持つWebデータベースを作成した。このデータベースに格納されたデータの中から、他分子との相互作用部位を変化させ、かつ、疎水性コアには影響を与えず安定な立体構造を形成できると考えられるAS領域のデータを収集し、ASアイソフォームのモデリングを行った。AS同様にmRNA上でゲノム情報が変化し遺伝情報の多様化に関係している可能性があるRNA編集についても、国際塩基配列データベースから関連情報を取得するアルゴリズムを作成し、植物オルガネラ由来のRNA編集の情報を収集した。RNA編集により変化するアミノ酸残基の部位とタンパク質の立体構造との関係を簡単に検索することができるWebデータベース(RESOPS)を作成した。

(2) AS産物のホモロジーモデリング法の開発

ホモロジーモデリング法は、テンプレート

(鋳型タンパク質) の検索法、アラインメント (アミノ酸残基の対応関係導出) 法、原子座標の導出法に分解することができる。テンプレートとターゲット (立体構造を推定したいタンパク質) のアミノ酸配列一致度は一般的には低いいため、ホモロジーモデリングにおいて、低いアミノ酸配列一致度であっても正しくアラインメントを構築することが問題になっている。そこで A S 産物のホモロジーモデリング用に、低いアミノ酸配列の一致度であっても精度よいアラインメントを得られる方法を構築した。類縁タンパク質の立体構造を比較することで、タンパク質の構造変化は立体構造のどのような部位で発生するのかを導出し、この経験則を定式化し自動アラインメント内に組み込むことで、高精度アラインメントを実現した。

(3) A S による機能部位変化の検証

A S データベースから A S が機能部位に改変をもたらしている可能性がある場合を検索し、本研究で開発したアラインメント法を利用してホモロジーモデリングを行い、アイソフォーム間の立体構造変化を明らかにし、変化にもとづく機能変化を推定した。さらに推定された変化が実際に起こっているかを文献から調べた。

4 . 研究成果

(1) A S データベースの作成

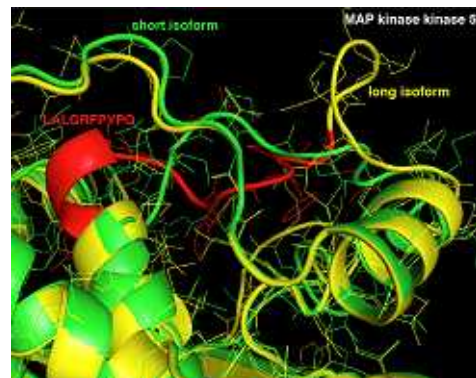
ゲノム全塩基配列が決定されている生物において、A S により影響を受けるアミノ酸配列上の領域をゲノム塩基配列データと遺伝子発現データから同定する計算手法を確立した (Shionyu, Go, et al., 投稿準備中。発表後にホームページから詳細を公開)。開発した方法を、高等真核生物 (ただしヒトとマウスを除く) に適用し、4300 個以上の A S データを集めることができた。得られたデータを解析し、C 末端側のアミノ酸配列が変化する A S が多く見られることが明らかになった。転写産物をゲノム全塩基配列に対応させる過程で、一塩基の違いが植物オルガネラで頻繁に起こっていることが明らかになった。これらは A S とは異なる機構で mRNA が改変された結果であり、これらの変化がタンパク質の立体構造形成に影響をおよぼす変化であることを明らかにした (発表論文 3)。植物オルガネラで起こっている RNA 編集に関しては、独立のデータベースを構築し公開するとともに、その起源が植物の上陸と関係があることを推定した (発表論文 1)。

(2) A S 産物のホモロジーモデリング法の開発

A S 産物のホモロジーモデリングにおいては、テンプレートとターゲットのアミノ酸配列の一致度が非常に低い場合と、A S を起こしているエクソン部分のテンプレートとのアミノ酸配列一致度が非常に低い場合が想定される。アミノ酸配列の一致度が低いタンパク質では、アミノ酸残基の置換のみならず、アミノ酸残基の挿入欠失が非常に多く起こっていることが、共通祖先由来のタンパク質の立体構造を比較することで明らかにできた。それらの挿入欠失部位がタンパク質立体構造のどのような部位で起こっているのかを調べたところ、タンパク質の内部よりも表面の方で非常に頻繁に起こっていることが明らかになった。アミノ酸残基の溶媒接触率と挿入欠失の頻度との間には相関関係があり、溶媒接触率から挿入欠失の頻度を推定する経験式を導出することができた。その式をアラインメントの手続きに導入することで、ホモロジーモデリングをするために必要な精度よいアラインメントを構築することができるようになった。同方法を ALAdeGAP と命名した (Yura, Go, et al. 投稿中。発表後にホームページから詳細を公開)。

(3) A S による機能部位変化の検証

A S データベースを用いて、機能部位 (基質結合部位、サブユニット相互作用部位など) に A S がおよぶ場合を、見いだすことができるようになった。データベースからそのような変異を起こしている A S 産物の組みを抽出し、ALAdeGAP を用いたアラインメントにもとづきホモロジーモデリングを行った。



MAP kinase kinase 5 の場合は、サブユニットとの相互作用を考えると考えられている部位で、A S により 10 残基の欠失が発生する。この欠失によりどのような立体構造変化が発生するかをホモロジーモデリングで明らかにすることができた (上図)。ひとつのエクソンにコードされている 10 アミノ酸残基 (上図赤部分) を失うことで、タンパク質の表面につきだしているヘリックスが短くなるとともに、10 残基の直後にコードされているアミノ酸残基で構築される、表面

に突き出したループも失ってしまうことがわかった。このループの部分が失われた10アミノ酸残基部分を補っていることがわかった。これらの変化がMAPキナーゼカスケードにおけるタンパク質の物理的相互作用に影響を及ぼすことが考えられる。

アシルタンパク質チオエステラーゼ1では、ASにより2量体形成部位を失ってしまい、2量体が形成できなくなる可能性がホモロジーモデリングにより示唆された(発表論文5)。エクソンにコードされている部位を失うことで、2量体間の接触表面積が1300平方から900平方程度になり、その結果2量体としての構造は維持できないと考えられる。このタンパク質は2量体を形成することで活性を調整(抑制)していると考えられている。2量体形成が不可能になることで、活性を調整できなくなり、活性を常に持っている構造になると考えられる。

ミトコンドリアにあるリン脂質アシル基転移酵素には、第5番目のエクソンにコードされる部分が存在しない産物があり、この酵素の立体構造が植物にある遠縁のタンパク質をテンプレートとしてホモロジーモデリングをすることができた。ホモロジーモデリングの結果、エクソン5にコードされる部分は活性部位の側にあり、かつ不安定なループ構造を形成する可能性が高いことが明らかになった。このループの有無によって基質特異性に影響が出ている可能性があることがわかった(Hijikata, Yura, Go, et al., 投稿準備中。発表後にホームページから詳細を公開)。

以上のように、ASによるタンパク質の構造変化をホモロジーモデリングをすることができるようになり、構造変化から機能にどのような変化がおよぶかを推定できるようになった。

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7件)

1. Yura, K., Sulaiman, S., Hatta, Y., Shionyu, M., Go, M., RESOPS: A database for analyzing the correspondence of RNA editing sites to protein three-dimensional structures. *Plant and Cell Physiology*, 査読有, **50**, 2009, 1865-1873
2. 由良敬、郷通子, 陸上植物オルガネラのRNA編集の役割, *生物物理*, 査読有, **49**, 2009, 244-245

3. Yura, K., Go, M. Correlation between amino acid residues converted by RNA editing and functional residues in protein three-dimensional structures in plant organelles. *BMC Plant Biology*, 査読有, **8**, 2008, 79

4. Yamasaki, C., Murakami, K., Fujii, Y., Sato, Y., Harada, E., ..., Yura, K. (66th), ..., Shionyu, M. (101th), ..., Go, M. (107th), ..., Gojobori, T. (138th) The H-Invitational Database (H-InvDB), a comprehensive annotation resource for human genes and transcripts. *Nucleic Acids Research*, 査読有, **36**, 2008, D793-D799.

5. Go, M., Yura, K., Shionyu, M. Contribution of computational biology and structural genomics to understand genome and transcriptome. *Proceedings of the International Symposium on Frontiers of Computational Science 2005 (ISFCS2005)*, (eds. Kaneda, Y., Kawamura, H., Sasai, M.), Springer, 査読有, 2007, 75-80.

6. Yura, K., Shionyu, M., Hagino, K., Hijikata, A., Hirashima, Y., Nakahara, T., Eguchi, T., Shinoda, K., Yamaguchi, A., Takahashi, K., Itoh, T., Imanishi, T., Gojobori, T., Go, M. Alternative splicing in human transcriptome: functional and structural influence on proteins. *Gene*, 査読有, **380**, 2006, 63-71.

7. Iida, K., Go, M., Survey of Conserved Alternative Splicing Events of mRNAs Encoding SR Proteins in Land Plants. *Mol. Biol. Evol.* 査読有 **23**, 2006, 1085-1094.

[学会発表](計25件)

1. Go, M., Information flow from genes to proteins, International Symposium: Fifty Years of Biophysics Research at Nagoya University, 2010年3月12日, 名古屋大学(名古屋市)
2. Shionyu, M., Takahashi, K., Go, M., AS-ALPS: a database for inferring effects of alternative splicing based on the protein 3D structure, Asian Young Researchers Conference on

- Computational and Omics Biology 2010, 2010年3月10日, National Cheng Kung University (Taiwan)
3. 郷通子, 遺伝情報の流れの複雑性からみる生物進化, 「生物進化の持続性と転移」国際高等研究所研究会, 2010年3月2日, 国際高等研究所(木津川市)
 4. 郷通子, 選択的スプライシングの曖昧さとその意味, 大沢文夫研究会, 2010年2月20日, 名古屋大学(名古屋市)
 5. 由良敬, RNA編集とその生物学的意味, 大沢文夫研究会, 2010年2月20日, 名古屋大学(名古屋市)
 6. Shionyu, M., Takahashi, K., Go, M., A web service for functional annotations of protein isoforms produced by alternative splicing, 第32回日本分子生物学会年会, 2009年12月11日, パシフィコ横浜(横浜市)
 7. Yura, K., Sulaiman, S., Hatta, Y., Shionyu, M., Go, M., RESOPS: A database of RNA editing sites in plant organelle mapped onto protein three-dimensional structures, 第32回日本分子生物学会年会, 2009年12月10日, パシフィコ横浜(横浜市)
 8. Go, M., Events during information flow from genomic sequence to protein structure/function in biological system, The 4th Global COE International Symposium 2009 joint with the 19th Hot Spring Harbor Symposium Molecular Evolution and Bioinformatics, 2009年11月1日, 九州大学病院キャンパス(福岡市)
 9. 郷通子, 筋蛋白質のモジュールと筋収縮について, 生物物理学会サテライトミーティング「筋収縮のメカニズムを考える」, 2009年10月29日, 徳島文理大学(徳島市)
 10. 郷通子, たくさんの師との巡り会いと海外で学んだ経験, 第9回日本蛋白質科学会年会, 2009年5月22日, 熊本全日空ホテル(熊本市)
 11. Yura, K., Diversification of protein repertoire by alternative splicing, The 66th Korean Society for Biochemistry and Molecular Biology Annual Meeting, 2009年5月13日, COEX (Korea)
 12. Yura, K., Go, M., Effect of residue substitution by RNA editing on protein three-dimensional structures in plant organelles, BMB2008(第31回日本分子生物学会年会、第81回日本生化学会年会 合同大会), 2008年12月10日, 神戸国際会議場(神戸市)
 13. Go, M., Yura, K., Correlation between RNA editing sites on amino acid residues and protein three-dimensional structures in plant organelles, 日本生物物理学会第46回年会, 2008年12月5日, 福岡国際会議場(福岡市)
 14. 郷通子, タンパク質の立体構造と遺伝子の構造の進化的起源に関する研究、日本進化学会賞/木村資生記念学術賞受賞講演、第10回日本進化学会大会, 2008年8月23日, 東京大学駒場キャンパス(東京)
 15. 郷通子, 由良敬, 植物オルガネラにおけるRNAエディティングの部位とタンパク質立体構造の関係, 第10回日本進化学会大会, 2008年8月22日, 東京大学駒場キャンパス(東京)
 16. Yura, K., Go, M., Information Flow from DNA Sequence to Protein Function, Science in Japan Forum 2008 "Interaction of Physics and Biology", 2008年6月20日, JSPS, 米国ワシントンDC
 17. 郷通子, 研究は最初から国際舞台で: 男性も女性も輝くために、生物物理若手夏の学校、2008年7月22日、八王子セミナーハウス(八王子市)
 18. 郷通子, シミュレーションが切り開く新しい生物学・夢と発想の展開へ、次世代生命体統合シミュレーション研究開発プロジェクトシンポジウム2007、2007年12月25日、MY PLAZA ホール(東京)
 19. 郷通子, 由良敬, 高橋健一, 塩生真史, 選択的スプライシングがもたらすタンパク質構造への影響、シンポジウム「構造プロテオミクスを支援するバイ

オインフォマティクス」、生物物理学会
年会、2007年12月22日、横浜パ
シフィコ(横浜)

20. 由良敬, 塩生真史, 郷通子, 真核生物の
遺伝子構造とタンパク質立体構造・機能
の関係, 日本進化学会第9回京都大会
シンポジウム, 2007年9月2日, 京都大
学(京都市)
21. Shionyu, M., Yura, K., Go, M., A
possible effect of alternative
splicing on multiple regions of a
single protein, 5th EABS & 44th Annual
Meeting of Biophysical Society of
Japan, 2006年11月13~14日, 沖縄コ
ンベンションセンター(宜野湾市)
22. Go, M., Computational Biology to
Understand Genome and Transcriptome,
5th EABS & 44th Annual Meeting of
Biophysical Society of Japan, 2006
年11月12日, 沖縄コンベンションセ
ンター(宜野湾市)
23. 由良敬, 郷通子, 真核生物遺伝子産物の
多様性を生み出している選択的スプラ
イニングのタンパク質立体構造への影
響, 日本進化学会第8回大会シンポジウ
ム, 2006年8月30日, 国立オリンピッ
ク記念青少年総合センター(東京都渋谷
区)
24. 郷通子, タンパク質の予期せぬ部品モ
ジュール数理と生物の橋渡しをめざし
て, 第27回数理の翼夏季セミナー,
2006年8月10日, 広島大学西条研修セ
ンター(東広島市)
25. Shionyu, M., Yura, K., Hijikata, A.,
Nakahara, T., Shinoda, K., Yamaguchi,
A., Takahashi, K., Go, M., Systematic
detection of protein regions affected
by alternative splicing, 20th
International Congress of
Biochemistry and Molecular Biology,
2006年6月22日, 国立京都国際会館
(京都市)

[その他]

ホームページ等

<http://as-alps.nagahama-i-bio.ac.jp/>

<http://cib.cf.ocha.ac.jp/RNAEDITING/>

<http://teapot.lib.ocha.ac.jp/ocha/>
<http://cib.cf.ocha.ac.jp/yura/present0.html>

<http://cib.cf.ocha.ac.jp/yura/present0.html>

6. 研究組織

(1) 研究代表者

郷通子 (MITIKO GO)

長浜バイオ大学・バイオサイエンス学
部・特別客員教授

研究者番号: 70037290

(2) 研究分担者

塩生真史 (MASAFUMI SHIONYU)

長浜バイオ大学・バイオサイエンス学部・
講師

研究者番号: 30345847

(3) 連携研究者

由良敬 (KEI YURA)

お茶の水女子大学・大学院人間文化創成
科学研究科研究院・教授

研究者番号: 50252226