

科学研究費助成事業 研究成果報告書

平成 26 年 6 月 3 日現在

機関番号：20101

研究種目：基盤研究(C)

研究期間：2011～2013

課題番号：23500368

研究課題名(和文) バイオインフォマティクスによるロイシンリッチリピートの進化、構造、機能の研究

研究課題名(英文) Evolution, Structure, and Function of Leucine rich Repeats as Revealed by Bioinformatics

研究代表者

松嶋 範男 (MATSUSHIMA, NORIO)

札幌医科大学・医療人育成センター・教授

研究者番号：60137403

交付決定額(研究期間全体)：(直接経費) 3,600,000円、(間接経費) 1,080,000円

研究成果の概要(和文)：ロイシンリッチリピート(LRR)は、同じまたは似た配列を繰り返すタンデムリピートの一つであり、2万個以上の蛋白質に存在する。多くのLRRは蛋白質やリガンドと相互作用し、動物や植物の自然免疫などの機能に關与している。LRRの繰り返しの長さは20から30アミノ酸残基からなり、8つのクラスが存在する。本研究により、植物にのみ存在するLRRクラスがバクテリアにも存在することを見出し、このLRR遺伝子は植物からバクテリアに水平伝播したことを示した。また、単細胞真核生物とバクテリア由来蛋白質から、これまでLRRモチーフと同定されていない新規なLRRモチーフを見出した。

研究成果の概要(英文)：Leucine rich repeats (LRRs) are present in over 20,000 proteins from viruses to eukaryotes. Two to sixty-two LRRs occur in tandem. Most LRR proteins are involved in protein, ligand and in protein, protein interactions; these include the plant immune response and the mammalian innate immune response. Each repeat with 20-30 residues can be divided into a highly conserved segment and a variable segment. Eight classes have been recognized. Plant specific LRRs (class: PS-LRR) had previously been recognized in only plant proteins. However, we find that PS-LRRs are also present in twenty proteins from eleven bacterial species. We indicate that horizontal gene transfer of genes/gene fragments encoding PS-LRR domains occurred between bacteria and plants, as opposed to descent from a common ancestor. Also we find novel LRRs that were unrecognized until now. The novel LRR domains are present over three hundred proteins including fungal ECM33 protein from unicellular eukaryotes and bacteria.

研究分野：総合領域

科研費の分科・細目：情報学・生体生命情報学

キーワード：ロイシンリッチリピート タンデムリピート 水平進化

1. 研究開始当初の背景

ロイシンリッチリピート(LRR)は、同じまたは似た配列を繰り返すタンデムリピートの一つであり、ウイルスからヒトまで幅広い生物種の20,000個以上の蛋白質に存在する。多くのLRRは蛋白質と蛋白質間や、蛋白質とリガンド間の相互作用部位として用いられており、自然免疫や神経システムの発生などの重要な機能に関与している。

LRRの繰り返し単位は20から30アミノ酸残基である。この繰り返し単位は、前半部分の保存性がよい Highly conserved segment (HCS)と、後半部分の長さでコンセンサス配列が多様な Variable segment (VS)に分けられる。HCSは通常 LxxLxLxxNxL の11残基、または LxxLxLxxCxxL の12残基からなる。現在、VSの長さでコンセンサス配列の違いにより、"Typical", "SDS22-like", "IRREKO", "Bacterial", "plant-specific(PS)", "TpLRR", "RI-like", "Cysteine-containing" の8つのLRRクラスが認識されている。

LRRはドメイン全体で馬蹄形構造またはらせん体構造をとる。HCSの3-5番目のアミノ酸が α ストランド構造を取り、隣り合う次の繰り返し単位の同じ部分とスタッキングして、馬蹄形の内側を裏打ちしている。VS部分の二次構造はそのコンセンサス配列によって決まり、例えば SDS22-like の場合は3(10)-ヘリックスを、IRREKO の場合は extended 構造をとっている。多くのLRRはその疎水性のコアを覆うキャップ構造を持つ。LRRドメインのN末端側のキャップをLRRNT、C末端側のキャップをLRRCTと呼ぶ。

これら8つのクラスのLRRの進化的起源が一つであるのかどうかという問題は未だ解かれていない。また、"PS"クラスのLRR、PS-LRRは植物にのみ存在するとされているように、8つのクラスは広い生物種に様に分布していない。LRRがどのように広い生物種に行きわたっていったのかという問いへの答えは、未だない。さらに、この8つのクラスに分類できないLRRが多数存在することがわかってきている。これらがどのように出来てきたのかという問題も未だ解かれていない。これらは明らかにされるべき問題である。

2. 研究の目的

LRRは著しい多様性を示し8種類のクラスが存在する。これまで、LRRを含む蛋白質の分子進化の研究はなされているが、各クラス間の分子進化研究は何もなされて

いない。本研究の目的はLRRのバイオインフォテクス研究を行うことである。具体的には、(1)新しいクラスのLRRを同定し各クラスの進化のスキームの提案、(2)植物からバクテリアへの植物蛋白質に特有なLRRクラスの水平進化の有無、(3)巨大LRR長周期の同定と分子進化である。

3. 研究の方法

さまざまなバイオインフォマティクスの手法を駆使して、次の3つの課題を実行する。

(1) 新規なLRRクラスの同定と進化研究: 我々の開発した方法を用いて、バクテリアからのLRRを含む蛋白質(LRR蛋白質)に存在する新規なLRRクラスを検出し、データベースを構築する。このデータを用いてマルチプルアライメントなどを用いて詳細な配列解析を実行する。

(2) LRRの水平進化の実証: バクテリアLRR蛋白質に存在する"Plant-specific LRR"ドメインの塩基配列を用いてホモロジー検索を行い、検出された"Plant-specific LRR"ドメインの系統樹を作成し水平進化(HGT)の有無を検証する。

(3) 新規な巨大LRR超周期の検出と分子進化: LRR-RLKが重複した仮想的な蛋白質のアミノ酸配列をクエリーにしたホモロジー検索を行い、巨大LRR超周期を検出する。また、配列解析を実行し進化のスキームを提案する。

4. 研究成果

(1) Non-canonical LRRを含む蛋白質: 新たなLRRである non-canonical LRRを持つ、30種の真核単細胞生物と73種のバクテリア由来の324個の蛋白質を同定した。この non-canonical LRRは、これまでの配列解析プログラムやデータベースにおいてLRRと認識されないものである。我々は Non-canonical LRRを、VSにより以下の4つのクラスに分類した。

Non-canonical SDS22-like,

Non-canonical IRREKO

Non-canonical Vibrio-predominant

Non-canonical bacteria-specific

このうちとは39個の蛋白質で共存していた。これらのHCSは、VxGx(L/F)x(L/C)xxNxLの12アミノ酸、またはVxGxLxLxxNxxLの13アミノ酸であった。三番目にGlyが挿入されているのが特徴である。

配列解析の結果: とを含む多くの蛋白質にLRRNTとLRRCTが見られた。二

次構造予測は 4 つのクラス全てで HCS の 4-6 番目はβストランドを強く好んだ。それに加え と は HCS の最初のアミノ酸と VS の最後のアミノ酸、 は VS の N 末端側でαストランドを強く好んだ。LRR 以外のドメインを持つ蛋白質が多数存在し、中には既知の LRR 蛋白質と同じドメインの構成を持つものが存在した。

我々は 324 個の蛋白質において、新しい LRR である non-canonical LRR を発見した。また、non-canonical SDS22-like@LRR と non-canonical IRREKO@LRR が同一の祖先から進化してきたことを提案した。機能については未知であるが、本研究はこれからの実験的な研究を促進するだろう。

(2) PS-LRR の水平進化 : 11 種のバクテリアから 20 個の PS-LRR 蛋白質を同定した。これらと significant similarity を持つ 83 個の真核生物の蛋白質(植物 82 個、珪藻 1 個)を見出した。バクテリアの PS-LRR と植物の PS-LRR のコンセンサス配列はほとんど同一であった。5 つのバクテリア PS-LRR 蛋白質は、LRR 以外に他の機能的ドメインを持っていた。EGID は *Synechococcus sp.* の CYA_1022 が他の生物種から来たと予測した。

植物とバクテリアの間での HGT : 結果をまとめると、以下ようになる。PS-LRR 蛋白質は植物には広く分布しているが、バクテリアでは 11 種からしか見つからなかった。古細菌には PS-LRR 蛋白質は存在しなかった。バクテリアの PS-LRR のコンセンサス配列は明らかに PS-LRR であることを示した。バクテリアの PS-LRR が 83 個の真核生物由来蛋白質と非常に強い similarity を持つ (E-value < 10⁻¹⁰) ことを明らかにした。さらに、いくつかのバクテリアの PS-LRR は、他のバクテリアのそれよりもむしろ真核生物の蛋白質に似ていることがわかった。EGID プログラムは CYA_1022 が外から来たと予想した。ゲノム配列解読済の 31 種のシアノバクテリアのうち、2 種のみが PS-LRR を持っていた。バクテリアの PS-LRR 以外のドメインは真核生物の蛋白質との similarity を持たなかった。□これらの PS-LRR 蛋白質を持つバクテリアのほとんどは水環境に生息している。また、今回の 11 種の中に、*Beggiatoa* はコメの根圏に生息しているなど、植物と生息地の近いバクテリアが含まれていることを文献検索により示した。

~ の結果は、バクテリアと植物の間で HGT が起こった証拠を与える。さらに、を合わせて考えると、それに関与したバ

クテリアの一つは *Synechococcus sp.* というのがもっともあり得るシナリオである。さらに、 と も HGT 仮説を支持する。

ここで対立仮説として、植物とバクテリアの PS-LRR がその共通祖先において一つの祖先 gene を持ちそこから発散進化し、多くの生物種では gene loss が起こり、現在の分布になったという仮説を考える。この仮説においては、この祖先 gene は少なくとも真核生物と原核生物の分岐の前、三十数億年前には存在したことになる。ここで、Leu は 6 つのコドンを持っている。もしこれほど前に祖先 gene が存在したのであれば、そこから非常に多い回数の非同義置換が起こるはずである。しかしながら、これは□の結果と矛盾するため、この仮説は成り立たない。さらに考えられるのが収斂進化の仮説であるが、この仮説も □の結果を説明できない。

HGT の向き : HGT の起こった向きは二通り考えられる。一つ目は、バクテリアから植物への向きである。この向きでは PS-LRR の gene は全ての植物の祖先に入ったと考える。しかし、これが起きたのは少なくとも一億五千万年前と非常に古くなり、□の結果を説明することができない。二つ目は、最近植物からいくつかのバクテリアに HGT が起こったという仮説である。この仮説は□~□の結果とよく合致する。

バクテリア間の HGT : 属をまたいだバクテリア由来蛋白質間で、塩基配列が非常に近い組み合わせが存在した。これは 11 種のバクテリアの間でも HGT が起こったことを意味している。

進化のスキーム : ここまでの全てを踏まえた最もありえる進化のシナリオは、植物とバクテリアの間で少なくとも一回の PS-LRR gene または gene fragment の HGT が起き、それに引き続きバクテリア間の HGT が起こったというシナリオである。

まとめ : バクテリアに存在する PS-LRR 蛋白質が植物からの HGT に由来することを示した。LRR が多様な生物種に広がっていく一つの driving force として HGT が存在すると考えられる。

(3) 新規な巨大 LRR 超周期 : 我々が注目していた巨大 LRR 超周期をもつ蛋白質は、その後の研究により、蛋白質として発現している可能性が否定された。それで、この研究を進めることをあきらめた。

5. 主な発表論文等

[雑誌論文](計 6 件)

1. Miyashita H, Kuroki Y, Matsushima N.

- “Novel leucine rich repeat domains in proteins from unicellular eukaryotes and bacteria” *Protein and Peptide Letters*, (2014) 21(3):292-305. 査読有
Doi: 10.2174/09298665113206660112
2. Miyashita H, Kuroki Y, Kretsinger RH, Matsushima N. “Horizontal gene transfer of plant-specific leucine-rich repeats between plants and bacteria” *Natural Science*, (2013) 5(5): 580-598. 査読有
doi:10.4236/ns.2013.55074
 3. Nunomura W, Jinbo Y, Isozumi N, Ohki S, Izumi Y, Matsushima N, Takakuwa Y. “Novel Mechanism of Regulation of Protein 4.1G Binding Properties Through Ca(2+)/Calmodulin-Mediated Structural Changes” *Cell Biochem Biophys*.(2013) 66(3):545-558. 査読有
doi: 10.1007/s12013-012-9502-7
 4. Matsushima N, Miyashita H. “Leucine-rich repeat (LRR) domains containing intervening motifs in plants” *Biomolecules*, (2012) 2: 288-311. 査読有
doi: 10.3390/biom2020288
 5. Mikami T, Miyashita H, Takatsuka S, Kuroki Y, Matsushima N. “Molecular evolution of vertebrate Toll-like receptors: Evolutional rate difference between their leucine-rich repeats and TIR domains” *Gene* (2012) 503(2):235-243. 査読有
doi: 10.016/j.gene.04.007
 6. Enkhbayar P, Boldgiv B, Matsushima N. “ ω -Helices”, *Current Topics in Peptide & Protein Research*, (2011)12:17-28. 査読有
http://www.researchtrends.net/tia/title_issue.asp?id=26&in=0&vn=12&type=3
〔学会発表〕(16件)
 1. Miyashita H, Kuroki K, Matsushima N. “Novel leucine rich repeats”, 24th International Conference on Genome Informatics (GIW2013), 16–18 December, 2013, Matrix, Biopolis, Singapore.
 2. 宮下博樹, 黒木由夫, 松嶋範男 “新しいロイシンリッチリピート” 第86回日本生化学会大会, 2013年9月11-13日, パシフィコ横浜, 横浜
 3. Matsushima N. “Structures and Evolution of Tandem Leucine Rich Repeat” The seminar of biophysics in Saarland University, Saarbruecken (Prof. Ingolf Bernhardt) Germany, June 24, 2013 (招待講演)
 4. Miyashita H, Kuroki Y, Kretsinger, RH, Matsushima N. “Horizontal gene transfer of plant-specific leucine-rich repeats between plants and bacteria” ISMB/ECCB 2013, 19-23 July, 2013, the Messe Berlin (ICC Berlin), Germany.
 5. Matsushima N, “Prediction and Evolution of Tandem Leucine Rich Repeats” The seminar of structural biology in University of Pittsburgh (Dr. Rieko Ishima), April 8, 2013, Pittsburgh, USA (招待講演)
 6. 宮下博樹, 黒木由夫, 松嶋範男 “Geometric architecture of solenoid structures of tandem leucine-rich repeats in proteins”, 2012年度日本生物物理学会北海道支部例会, 2013年3月5日, 北海道大学
 7. 宮下博樹, 黒木由夫, 松嶋範男, “単細胞生物のロイシンリッチリピートの比較解析” 第85回日本生化学会大会, 2012年12月14-16日, 福岡国際会議場・マリンメッセ福岡
 8. Miyashita H, Matsushima N. “Leucine-Rich Repeat Domains Intersected by Non-LRR Island Regions”, Genome Informatics Workshop (GIW2012), 12-14 December, 2012, Tainan, Taiwan.
 9. Miyashita H, Matsushima N. “LRRpred: A New Method for the Identification of Leucine-Rich Repeats”, Genome Informatics Workshop (GIW2012), 12-14 December, 2012, Tainan, Taiwan.
 10. Matsushima N, Enkhbayar P. “Structural features of helical secondary structures and leucine-rich repeat superhelix in proteins as revealed by HELFIT analyses” ICNAAM 2012 (10th International Conference on Numerical Analysis and Applied Mathematics) September 19-25, 2012, Kos, Greece (招待講演)
 11. Enkhbayar P, Matsushima N. “HELFIT: Algorithm and applications” ICNAAM 2012 (10th International Conference on Numerical Analysis and Applied Mathematics) September 19-25, 2012, Kos, Greece (招待講演)
 12. Matsushima N, “Molecular Evolution of Toll-like receptors” 2nd International Conference on Computation for Science and Technology (ICCST-2), July 9-11, 2012, Nigde, Turkey (招待講演)
 13. 宮下博樹, 黒木由夫, Robert H. Kretsinger, 松嶋範男, “バクテリア蛋白質のロイシンリッチリピートの分子進化” 日本生物物理学会北海道支部例会, 2012年3月6日, 旭川市民文化会館
 14. 宮下博樹, 黒木由夫, 松嶋範男, “ロイシンリッチリピートの分子進化” 第84

回日本生化学会大会, 2011年9月22日,
京都国際会館

15. Matsushima N, Mikami T, Miyashita H,
Yamada K. “Plant Leucine-rich Repeat
(LRR) - containing Proteins with Non-LRR,
Islands Interrupting LRRs” , XXIV SPPS
Congress, 21-25 August, 2011, Stavanger,
Norway
16. 宮下博樹, 黒木由夫, 松嶋範男 “バク
テリア“Plant-specific”ロイシンリッチ
リピートの分子進化 , 日本生化学会北
海道支部例会, 2011年8月5日, 札幌医
科大学

〔図書〕(計 1 件)

Matsushima N, Miyashita H, Mikami T,
Yamada K. “A new method for the
identification of leucine-rich repeats by
incorporating protein secondary structure
prediction” in *Bioinformatics: Genome
Bioinformatics and Computational Biology*
(Renu Tuteja editor); NOVA Science Publishers:
Hauppauge, NY, USA, (2011) pp. 61-88

〔産業財産権〕

- 出願状況 (計 0 件)
- 取得状況 (計 0 件)

〔その他〕

1. ホームページ:
<http://www.sapmed.ac.jp/~matusima/>
 2. 出前講義: 登別中学校での全校授業 “生
命科学の最前線 生命の歴史 ”
(2012年) 10月3日、登別市
6. 研究組織
- (1) 研究代表者
松嶋 範男 (MATSUSHIMA, norio)
札幌医科大学・医療人育成センター・
教授
研究者番号 : 60137403