3H1400

全ゲノム立体構造予測データベース”GTOP”

○川端 猛¹、福地 佐斗志¹、石井 崇洋¹、太田 元規¹、伊藤 武彦²、落合 孝正²、市吉 伸行²、西川 建¹ (¹ 遺伝研、² 三菱総研)

ゲノム由来の膨大なタンパク質配列群を解析・整理することの重要性は論を待たないという時代である。ついに我々も構造ゲノム科学(配列→構造→機能)の考えに基づき、全ゲノム規模の立体構造予測データベース”GTOP”を構築し公開するに至った。立体構造予測は、構造データベース(PDB)に対する配列相同性検索で行う。配列検索のツールとして、BLAST, FASTAに加え、プロファイル型検索ソフト”PSI-BLAST”を双方向で使用している。FASTAだけではORF数の20-30%の予測が可能であったのが、PSI-BLASTを使うことで35-50%まで向上する。これら以外に、機能既知の配列(Swissprot)に対する配列検索、SOSUIによる膜貫通ヘリックス予測、コイルドコイル予測、シグナル配列予測、ProSite、Pfamの結果も取り込み、総合的にタンパク質の特徴付けができる環境を目指している。現在までに、大腸菌、枯草菌、酵母、線虫など10種の解析を終えた。実はこういうPSI-BLASTによる全ゲノムの立体構造予測は世界的には決して新規なことではない。むしろ当然行うべき「常識」レベルの解析だ。オリジナリティはこれらの膨大な結果から何をどうやって読み取るかにある。我々は立体構造と機能の橋渡しをするのは間違いないリガンドであると考え、各予測構造にリガンド結合サイトが表示される仕組みを作り、機能の保存性の検討が容易にできるようにした。また、ゲノム全体として眺めるときに便利ようにファミリー(SCOPスーパーファミリー・ProSite・Pfam)ごとのタンパク質の一覧も作成し、生物種ごとの特徴付けも可能になっている。しかし、とにかく解析結果は膨大である。多くの人に閲覧してもらい、結果を検討していただければと思う。

アドレスは <http://spock.genes.nig.ac.jp/~genome/index.html>

T.Kawabata, S.Fukuchi, T.Ishii, M.Ota, T.Ito, T.Ochiai, N.Ishiyoshi, K.Nishikawa : GTOP : database for "Genome TO Protein structure and function"

3H1430

受容体内部の極性環境とリガンド特性の相関性：7本膜貫通ヘリックス型受容体の機能分類

○諏訪 牧子¹、Yudate Henrik. T¹、増保 安彦¹、美宅 成樹² (¹ヘリックス研究所、²東京農工大学・工)

7本膜貫通ヘリックス型受容体は、実に多彩なリガンドに結合して機能を発現しており、様々な分野で機能予測が急がれている。予測には、機能既知の相同配列を検索する方法が一般的だが、これでは機能発現のメカニズムを得られず、他と相関性が無い未知配列には情報を与えられない。リガンド分子は、受容体と特異的に結合して複合体を作るため、分子量、疎水性値などの物理化学的な特性は、受容体の構造特徴と相関がある。つまり、受容体の配列が他と相関性が無くとも、配列から構造特徴を得れば、リガンド特性を予測できると考えた。そこで機能に関わる構造(細胞外側、細胞質側ループ、膜貫通ヘリックス)に注目し、リガンド特性との相関性を調べた。SWISSPROTからリガンド既知の受容体配列を全て選択し、無冗長化した。リガンドの分子量と親水性値を計算すると、分子量>750のペプチド性リガンドでは分子量の増大に伴い顕著に親水性が高くなった。この様な分子に対し、まず受容体の膜貫通ヘリックス部内部の環境との相関性を調べた。前回の年会で、ヘリックスのN、C末端領域ごとに周囲の極性表面の形を定量的に表現してタイプ分けしたが、今回は、ヘリックス部分での各タイプの領域の存在割合から内部の環境の度合を定義した(強、中間、弱極性環境)。3つの極性環境のタイプに対応して、細胞外側、細胞質側の総ループ長を軸に各リガンドをプロットすると、物理化学的特性が近いリガンドがクラスターとなり判別できた。(強極性環境では、細胞外側の総ループ長が150残基以上、細胞質側の総ループ長が200残基未満では、分子量が3000以上の親水性ペプチドに限られたなど。)以上の結果は、膜外ループ部分、膜内ヘリックス部分の内部環境とリガンドの物理化学的な特性の相関性を示唆し、受容体側の構造特徴のみから、リガンドの特性が予測できる可能性を示している。

M. Suwa, H. T. Yudate, Y. Masuho, S. Mitaku : Functional Classification of 7 Transmembrane Helix Receptors, based on Inner Polar Environment and Ligand Property.

3H1415

シグナルペプチド判別・領域予測システム SOSUI Signalの開発

○五味 雅裕、美宅 成樹(農工大・工・生命工)

シグナルペプチドは、分泌タンパク質および一部の膜タンパク質前駆体のN末端に存在し、タンパク質の生体膜透過と膜への組み込みにおける中心的な役割を担っている機能性セグメントである。膜タンパク質におけるシグナルペプチドは、トポロジーを決定する大きな要因の一つであり、膜タンパク質の立体構造形成を議論するのに欠かせない。またシグナルペプチドはタンパク質の輸送経路と最終的な行き先を決定する”””荷札”””的な役割を持っており、細胞内におけるタンパク質局在を支配する主たる要素である。このような背景から、任意のアミノ酸配列が与えられた場合にシグナルペプチドを判別し、その領域を予測することは、タンパク質の機能予測上、また膜タンパク質の立体構造を予測する上でも有意義なことである。本研究ではシグナルペプチド判別・領域予測を行うアルゴリズムとして主流となっている機械学習的アルゴリズムに代わる物理化学的パラメータによるアルゴリズムを提案し、これに基づくシグナルペプチド判別・領域予測システムSOSUI Signalを開発する。SOSUI Signalはアミノ酸配列の入力に対してシグナルペプチドの有無を判別し、分泌タンパク質、非分泌型水溶性タンパク質、膜タンパク質を分類する。疎水性セグメントの位置、N-端側の正電荷、C-端側(正電荷から18ないし27残基の領域の配列)に見られる特徴などを用いる。SOSUI Signalは、膜タンパク質判別・膜貫通領域システムSOSUIと連携を取ることによってシグナルペプチドを有する膜タンパク質について膜貫通領域を高精度に予測し、限定的ながらトポロジーを予測するなど膜タンパク質解析ツールとして有用である。

M Gomi, S Mitaku : Development of Signal Peptide Discrimination System: SOSUI Signal

3H1445

膜タンパク質予測システムSOSUIによる生物ゲノム比較解析の展開

○美宅 成樹、高江洲 宏智、赤澤 史嗣、林部 敏之、園山 正史(東京農工大・工)

ヒトゲノム計画が最終段階を迎え、得られたすべての配列データからできるだけ多くの構造・機能情報を得られることが求められている。そうした網羅的情報処理システムとして私たちは膜タンパク質予測システムSOSUIを開発してきた。このシステムでは、任意のアミノ酸配列に対して、それが膜タンパク質か水溶性タンパク質かを判別することができる。このシステムの特徴はホモロジー解析とは異なり、アミノ酸の配列によらず、同じ質の予測結果が得られるということ、生物ゲノムの特徴をつかむのに適したシステムとなっている。しかし、オリジナルSOSUIはいくつかの問題があった。(1)膜貫通領域予測で96%と少し悪く膜貫通ヘリックス本数予測の精度が良くない、(2)ゲノムレベルの解析にはバッチ処理が必要である、(3)予測された膜貫通領域のグラフィカルな表現が車輪図だけである、(4)UNIXしか対応していない。そこで本研究では、それらの問題を解決すべく新しいシステムの開発を行った結果を報告する。具体的に検討していることは、以下の4つのことである。(1)膜貫通ヘリックスの本数予測の精度を95%以上とする。(2)バッチ処理をWEB上でもできるようにすると同時に、結果を統計的に表現できるようにする。(3)予測結果における膜貫通ヘリックスを横からも表現できるようにする。(4)Windows版を作り、より多くのユーザーが使えるようにする。このシステムを用いて、ショウジョウバエなど新しく得られた生物ゲノムの比較解析結果の例を示す予定である。

S. Mitaku, H. Takaesu, F. Akazawa, T. Hayashibe and M. Sonoyama : Comparison of Total Genome Sequences by Improved Membrane Protein Prediction System SOSUIver2