2P301

原核生物 87 種の膜貫通タンパク質におけるループ上機能ドメインの定義と分類

平山 慎太郎¹、〇松代 欣也¹、吉田 秀一¹、新居 真吏²、水田 智史¹、清水 俊夫¹ (¹ 弘前大・理工・電子情報システム工、² 東北大院・生命科学・生命機能)

膜貫通(TM)タンパク質の機能ドメインを定義する場合、ループ領域とTMセグメント領域に分けて考える必要がある。しかし、Pfamドメインの定義は、配列全体をベースに行われており、PfamドメインをTMタンパク質にアサインした場合、ループ領域とセグメント領域にまたがってアサインされたり、ループ上だけにアサインされたりと必ずしも一貫していない。また、立体構造から定義されるSCOPドメインは、構造既知のTMタンパク質が少ないため、エントリー数が少なく、今のところ網羅的な解析には適していない。そこで、本研究ではTMタンパク質のTMトポロジー(TMヘリックス本数及びその位置、膜貫通方向)を考慮にいたれたドメイン定義を行い、それらをTMタンパク質ループ上ドメインのデータベースとして構築する。

原核 87 ゲノムからの全 TM タンパク質配列を対象とし、TM トポロジー類似性に基づくクラスタリングを行い、さらに、より良いアライメントを得るために、クラスター内の全てのペアワイズ配列一致度が 20% 以上になるように配列選択処理を行った。この処理で得られたクラスターに対してマルチプルアライメントを行い、このアライメントのセグメント端の平均を計算し、ループ長が 60 残基以上でギャップの割合が 50% 以下となるループ領域のマルチプルアライメントを取り出した。このようにして得られたデータセットから、profile HMM としてドメインを定義した。これにより、TM タンパク質のドメイン・アーキテクチャーの詳細な解析が可能となる。

現在、本研究で定義したドメインと Pfam ドメインとの比較、検討を行っている。また、原核生物種だけではなく、真核生物種を含む全ての TM タンパク質配列を対象とした、ループ上のドメイン定義を同様の方法で進めているところである。

S.Hirayama, K.Matsushiro, S.Yoshida, M.Arai, S.Mizuta, T.Shimizu : Definition and classification of functional domains on loop regions of prokaryotic transmembrane proteins

2P303

立体構造情報を利用したタンパク質機能予測法 FCANAL: 酵素・結合タンパク質への応用

〇鈴木 歩¹、安藤 格士¹、松村 綾子¹、坂尾 宗旦¹、山登 一郎¹、宮崎 智² (¹ 東理大・基礎工・生物工、² 東理大・薬学・薬学)

構造ゲノムプロジェクトの進展によってタンパク質立体構造情報の蓄積は膨大な量になり、機能未知の立体構造も増加してきている。タンパク質の機能を予測する手法として配列相同性検索が一般的であるが、機能既知の遺伝子との相同性が見られない配列が多数存在することから、この手法には限界があると考えられている。そこで配列よりも機能部位における保存性が高く、機能と密接に関わっているタンパク質立体構造情報を利用した機能予測法が注目されてきている。私達の研究室では、立体構造情報を利用した高速・高精度な機能予測法 FCANAL (Fast Calculable protein function ANALyzer) の開発を行ってきた。本手法では機能部位とそれ以外の部位のアミノ酸残基の C α 原子間の距離と、その出現頻度を比較することによってスコア行列を作成し、機能予測に用いている。立体構造という 3 次元情報を、距離と出現頻度という 1 次元情報に置き換えているため計算速度が速く、また機能部位の構造変化に対しても柔軟に対応することができる。本研究では、FCANAL を網羅的なタンパク質機能予測法として一般化することを目指している。今回、機能について情報が充実している酵素と基質結合タンパク質を予測対象とした。これらタンパク質に対して機能部位の立体構造を反映したスコア行列を作成し、高い精度で機能部位を同定することが可能であった。本 FCANAL は配列相同性やフォールド類似性の低いタンパク質の機能予測にも有効であった。さらに、機能未知のタンパク質立体構造に対して FCANAL を実行したところ、機能に関する手がかりを得ることができた。今後、構造データベースの拡充とともに、本手法の有用性はさらに高まるものと考えられる。現在、FCANAL は <http://atgc002.ps.noda.tus.ac.jp/contents/fcanal/> で公開しており、今後さらに予測できる機能の種類を増やしていく。

A. Suzuki, T. Ando, A. Matsumura, H. Sakao, I. Yamato and S. Miyazaki : FCANAL: A structure based protein function prediction method. Application to enzymes and binding proteins

2P302

てんかん関連膜貫通タンパク質遺伝子の網羅的検出

〇井上 泰仁¹、金井 数明²、水田 智史¹、岡田 元宏³、兼子 直³、清水 俊夫¹ (¹ 弘前大・理工・電子情報システム工、² 千葉大・医・附属病院・神内、³ 弘前大・医・精神神経)

熱性けいれん (febrile seizure, FS) は、6 歳以下の全小児の数が罹患し、最も頻繁に見られるけいれん疾患である。また、全般てんかん熱性けいれんプラス (generalized epilepsy with febrile seizure plus, GEFS+) は、6 歳を越えてなお生じる熱性けいれんと、その後の高頻度での無熱性けいれんを特徴とするてんかん症候群である。GEFS+ は FS の亜型と考えられており、これまでに 5 つの責任遺伝子が報告されている。その多くはイオンチャネル機能のものである。現在のところ、FS の責任遺伝子は分かっているが、FS が GEFS+ の一連の疾患であることを考慮するならば、イオンチャネルが関与していると考えられる。

そこで、本研究では、イオンチャネルなどの膜貫通(TM)タンパク質に注目し、ヒトプロテオーム (Build 35, 28,104 配列) から、FS の責任遺伝子の存在が報告されている遺伝子座 (2q23-q24, 5q14-q21, 6q22-q24, 8q13-q21, 19p13.3) 周辺におけるてんかん関連遺伝子を検索した。これらの領域には、それぞれ 63、63、126、100、234 の遺伝子が存在し、27、11、40、13、59 個の TM タンパク質遺伝子が含まれている。これらの TM タンパク質遺伝子を SWISS-PROT に対してホモロジー検索を行うことで、それぞれ 8、1、4、1、1 配列の Na⁺、K⁺ チャネル、GABA_A 受容体、あるいは代謝型グルタミン酸受容体などのてんかん関連 TM タンパク質遺伝子候補を検出した。

また、我々が開発した BTP 法による各種イオンチャネルの検出、および、24 型 Na⁺ チャネルの TM ヘリックス空間配置予測モデルと分子病態との対応関係について解析した結果についても報告する。

Y.Inoue, K.Kanai, S.Mizuta, M.Okada, S.Kaneko and T.Shimizu : Comprehensive detection of transmembrane protein genes associated with epilepsy

2P304

立体構造にもとづくタンパク質-RNA 相互作用面のデータベース解析

〇Kim Oanh¹、由良 敬^{1,2}、郷 信広^{1,3,4} (¹ 原研・計算科学セ・量子生命、² CREST, JST、³ 原研・中性子セ・生体物質、⁴ 奈良先端大・情報)

Protein-RNA interactions play essential roles in a number of regulatory mechanisms of gene expression such as RNA splicing, transport and translation. Computational analyses of protein-DNA interface have been carried out a lot, while those of protein-RNA interface are rare. As the number of available protein-RNA complexes have increased, it is now possible to examine protein-RNA interactions based on 3D structures. We carried out computational analyses of 91 protein-RNA complexes retrieved from PDB. Interface residue propensity was calculated which give a measure for the relative importance of different amino acid residues in the RNA-binding site. The result showed that Arg, Lys, His, Phe, Tyr, Met and Gly were favored in RNA-binding site. The interface residue propensity was used as a parameter to predict protein-RNA interface. The accuracy of the prediction method was tested by jackknife method, which attained 60% on average. Although more analysis is required for the better prediction, the residue interface propensity was proved to be a good parameter for predicting protein-RNA interface. The prediction method was applied to RNA-binding proteins with known 3D structures: exon junction complex Mago-Y14 and mRNA export factors TAP and Mex67. These proteins have long been studied, but how they interact with RNAs is still unsolved. The prediction method enables us to point out the candidate interfaces, which are consistent with previous studies and can be tested by mutation experiments leading to understanding protein-RNA interactions.

O.T.P.Kim, K.Yura, N.Go : Structure-based bioinformatics analyses of protein-RNA interface toward developing a computational method to predict protein-RNA interface