

第3回 日本ゲノム微生物学会年会 要旨集

2009年3月5日(木)～7日(土)

於：中央大学理工学部(後楽園キャンパス)



SGMJ

日本ゲノム微生物学会
Society of Genome Microbiology, Japan

セッション7-3 病原微生物

11:15-11:30 7301

次世代シーケンサを用いた感染症のメタゲノミック診断

○飯田 哲也¹

¹大阪大学微生物病研究所 感染症国際研究センター ゲノム病原細菌学

11:30-11:45 7302

タイリングアレイによるらい菌ゲノム全域の網羅的発現解析

○赤間剛¹、鈴木幸一¹、谷川和也¹、川島晃¹、Huhehasi Wu¹、林もゆる¹、石藤雄子¹、石井則久¹

¹国立感染症研究所ハンセン病研究センター生体防御部

Luncheon Seminar-2

12:00-12:45

司会：服部正平（東京大学大学院新領域創成科学研究科 生命システム観測分野）

次世代シーケンサの可能性

○大島健志朗¹

¹東京大学大学院新領域創成科学研究科 情報生命科学専攻

沖縄先端バイオ事業で見いだされてきたSOLIDの新たな可能性

○照屋 盛実^{1, 2}、塚原 正俊^{3, 2}、喜久里 育也^{4, 2}、藤森 一浩^{5, 2}、町田 雅之^{5, 2}、河原林 裕^{5, 2}、平野 隆^{5, 2}

¹沖縄県工業技術センター、²沖縄先端ゲノム、³(株) トロピカルテクノセンター、⁴OSTC、⁵産総研

共催 アプライドバイオシステムズ ジャパン株式会社

セッション7-4 バイオインフォマティクス

13:00-13:30

7401 (招待講演)

Structure evaluation using the consensus and quality assessment for protein modeling

寺師玄記¹、酒井博子¹、加納和彦¹、平田朋子¹、竹田一志鷹真由子¹、岩館満雄²、

○梅山秀明¹

¹北里大学薬学部、²中央大学理工学部

13:30-13:45

7402

新代謝経路探索のための化学量論的代謝ネットワーク拡張

○太田潤¹

¹岡山大学

13:45-14:00

7403

一括学習型自己組織化マップ法による全ゲノム・全タンパク質配列を対象にした大規模ボストゲノム解析

○阿部貴志¹、金谷重彦²、池村淑道¹

¹長浜バイオ大学、²奈良先端科学技術大学院大学

7404 ポスター発表

大規模なゲノムデータを活用した比較ゲノム解析のワークベンチ RECOG

○内山郁夫¹, 樋口俊夫², 北橋竜雄²

¹ 自然科学研究機構基礎生物学研究所, ² インテックシステム研究所

ゲノム配列解析は多様な生物種へと拡大を続けており、新型シーケンサの登場によりその流れはさらに加速する勢いである。特に微生物については千種近くのゲノム情報がすでに利用可能であり、近縁ゲノムの多重比較やメタゲノム解析など、これらのデータを活用した大規模な比較解析研究がますます重要になりつつある。そこで、我々は比較ゲノム解析を進めるためのワークベンチ RECOG を開発している。これは、我々がこれまで開発、公開してきた、オーソログ解析を中心とした微生物ゲノム比較データベース MBGD を発展させたもので、Java で開発されたインターフェイスを持ち、タキソノミーデータベースや遺伝子機能分類の情報などを活用しつつ、大規模なオーソログ解析結果を詳細に検討するための様々な機能を含んでいる。オーソログテーブル解析機能としては、系統パターン類似性に基づくクラスタリングや並べ替え、「シアノバクテリア門の半数以上にあるが、プロテオバクテリア門にはない」といったタキソノミーを用いたフィルタリングなどの機能を含んでいる。また、特に利用者が興味のあるゲノムを中心とした近縁ゲノム比較解析を支援する機能にも力を入れており、興味のあるゲノムを含む系統群を内群、その他を外群として指定して、タキソノミー情報を取り込んだオーソログ対応付けを行う機能や、近縁ゲノム間で遺伝子の並び順が保存性された領域を「ゲノムコア構造」として抽出し、アライメント風に表示する CoreAligner 機能など、ユニークな解析機能を含んでいる。RECOG は MBGD サーバに接続してそのまま利用できるほか、ローカルにサーバを立ち上げることにより独自のデータを用いた解析を行うこともできる。プログラムは <http://mbgd.genome.ad.jp/RECOG/> より取得できる。

7405 口頭発表

GenomeMatcher: 比較ゲノムソフトウェアとその付属機能

○大坪嘉行¹, 大坪和香子¹, 永田裕二¹, 津田雅孝¹

¹ 東北大学生命科学研究科

我々は、比較ゲノム用のソフトウェア GenomeMatcher* を公開した (プロジェクトホームページ: <http://www.ige.tohoku.ac.jp/joho/gmProject/gmhomeJP.html>)。GenomeMatcher は BLAST の各種プログラム (blastn, tblastx, tblastn, blastp) と MUMmer による 2 つのゲノム配列の比較結果をグラフィカルに表示する。本ソフトウェアは、複数のゲノムを相互比較する機能を備える一方で、数塩基レベルの一致を検出して表示する機能をも備えており、数 bp から数十 Mb までの広いレンジの解析が可能な仕様となっている。本ソフトウェアは 1) 任意部位の再解析が容易、2) 注釈情報の追加および参照が可能、3) 編集可能な図が描画可能、4) 相同性の高低が色調により表現、5) 点変異 (SNP) をグラフィックス上で表示、といった多くの特色を備えている。

また本ソフトウェアにはゲノム解析に限らない様々な用途に適した機能付属が付けられている。これらは、1) GenBank / DDBJ 形式の情報を表計算シートに適合する形式に変換する機能、2) 連想配列を取り扱う機能、3) 2 つの集合に共通する要素と片方にのみ存在する要素に分類する機能、4) BLAST の実行と、実行結果を表計算シートに適合する形式に変換する機能、5) 数値、文字などのテキストデータを基に描画する機能、6) 文字列処理機能 (複数の検索置換を同時に行う機能、表計算シートにある情報を FASTA 形式に変換する機能、改行コードの変換機能など)、等である。これはいずれもグラフィカルユーザーインターフェースから実行可能である。

*参考文献: GenomeMatcher: A graphical user interface for DNA sequence comparison. BMC Bioinformatics 2008, 9:376 (16 September 2008)

7406 口頭発表

大量オーソログデータを用いた原核生物の系統樹解析

○堀池徳祐¹, 官田大輔², 館野義男³

¹ 静岡大学 創造科学技術大学院, ² 千葉商科大学 商経学部,

³ 国立遺伝学研究所 生命情報・DBJ 研究センター

分子進化学において原核生物全体の系統関係を理解することは様々な進化的現象を理解する上で大変重要である。しかし、多数の遺伝子水平伝達やアウトパロログの欠失の影響により、正しい系統樹を作成するのは困難だった。以前の研究で我々はこの問題を克服すべく、独自の方法でオーソログ配列データセットを作成し、その連結系統樹から種の系統関係を推定した。また、系統樹の各結節の信頼性を評価するために結節の支持率を定義し計算した (Horiike et al. Gene 2009, 429:59-64)。この方法では、まず連結系統樹を各オーソログの系統樹と比較し、連結系統樹の各結節において分歧パターンが矛盾しない系統樹の数を集計する。そして、この集計値を比較可能な結節を持つ系統樹の数で割る事で計算される。しかしこの支持率計算法では遺伝子水平伝達などのイベントが起こった遺伝子がオーソログ配列データセットに含まれている場合、伝達元と伝達先までのすべての結節が支持されなくなるため、より根に近い結節の支持率が低くなる傾向が強かった。我々は遺伝子水平伝達が起こったとしても、それに直接関与しない結節は支持率計算時に支持される方法を新たに考案した。本発表ではこの方法を用いた解析結果を報告する。

7407 ポスター発表

CIRCLE: アミノ酸側鎖環境と二次構造類似度に基づく、タンパク質予測立体構造評価法の開発

○寺師玄記¹, 西井博子¹, 加納和彦¹, 平田朋子¹, 竹田一志 眞由子¹, 梅山秀明¹

¹ 北里大学薬学部

The accurate prediction of protein structure is one of the major challenges in the field of bioinformatics. The Model Quality (MQ) assessment technique for distinguishing the near native models (high quality models) from decoys which are inferior models is one of the most important factors to achieve the accurate protein structure prediction.

CIRCLE considers two terms for the model quality: (1) model quality calculated from the side-chain environment of each residue; and (2) similarity between the secondary structure propensities predicted for an amino acid sequence by PSI-PRED and the secondary structure of the three-dimensional model. The side-chain environment for each residue is determined from the fraction of the molecular surface area of the side-chain covered by the polar atoms, the fraction of the side-chain area buried by any other atoms, and the secondary structure.

7408 ポスター発表

FAMSD: Individual comparative modeling server using SP3 & SPARKS2, FAMS and CIRCLE in CASP8

加納和彦¹, 平田朋子¹, 〇寺師玄記¹, 酒井博子¹, 竹田一志鷹真由子¹, 梅山秀明¹

¹ 北里大学薬学部

We develop protein structure prediction method based on comparative modeling called FAMSD. FAMSD method consists of following four steps:

(1) making sequence alignments between target protein and template structures, (2) constructing three-dimensional structures based upon each alignment, (3) selecting the best structure model and (4) refinement of the selected model.

We have developed an automatic protein structure prediction server called FASMD. Programs such as SP3, FAMS (Full Automatic Modeling System), CIRCLE and Molecular dynamics were used at the each step (1) ~ (4), respectively.

We had participated in the recent Critical Assessment of Techniques for Protein Structure Prediction (CASP) experiments as an automatic predictor using FAMSD method for the purpose of assessing our method. We describe the algorithm of FAMSD method and our results for CASP8.

7410 ポスター発表

FAMS_multi: Automated homology modeling based upon multiple reference proteins in CASP8

加納和彦¹, 平田朋子¹, 〇寺師玄記¹, 酒井博子¹, 竹田一志鷹真由子¹, 梅山秀明¹

¹ 北里大学薬学部

We developed an automated method of protein structure prediction called FAMS (Full Automatic Modeling System). FAMS is a homology modeling program consisting of database search and simulated annealing, and can construct high accuracy model when appropriate reference protein was detected. For predicting more accurate model, especially of loop structure and side chain torsion angles, we developed a new version of FAMS, called FAMS-multi, which uses multiple reference proteins.

For the purpose of assessing this method, we participated in CASP8 (8th Critical Assessment of Techniques for Protein Structure Prediction) experiment as 'FAMS_multi' team. CASP is a world-wide experiment for protein structure prediction. We participated as a human predictor, but all processes were performed automatically. Models which were predicted by other automatic predictors were used to generate better alignments, and we rebuilt models by using FAMS-multi program which uses multiple reference proteins. In the following, we describe the scheme of this method and our results for CASP8.

7409 ポスター発表

FAMSD_OA: Assessing model quality using the side chain environment consensus score in CASP8

加納和彦¹, 平田朋子¹, 〇寺師玄記¹, 酒井博子¹, 竹田一志鷹真由子¹, 梅山秀明¹

¹ 北里大学薬学部

Selecting the best quality model from a set of predicted structures is one of the most important parts of protein structure prediction. In the CASP7 experiment, the new prediction category called "QA" (Quality Assessment) was implemented. In this category participants assess the quality of the predicted models by any servers. Each QA predictor gives a "reliability score" to these models. After the experimental structures of CASP7 targets became available, the GDT_TS score was calculated for all predicted models and was compared with the "reliability score". As a result of QA category in CASP7, we found that the most powerful method was consensus method like 3D-Jury. 3D-Jury score is the summation of the number of CA within 3.5Å from each predicted model. This method could select "good backbone" models but the quality of the side chain of selected models is not so good. Accordingly we developed a alternative consensus method which considers side chain environment for the purpose of selecting good side chain models. We describe the algorithm of this method and our results for CASP8.

7411 ポスター発表

fams-ace2: タンパク質予測立体構造評価法の開発

〇寺師玄記¹, 酒井博子¹, 加納和彦¹, 平田朋子¹, 竹田一志鷹真由子¹, 梅山秀明¹

¹ 北里大学薬学部

The CASP is carried out in order to survey the capabilities and limitations of current methods of modeling protein structure from sequence. The methods are assessed on the basis of the analysis of a large number of blind predictions of protein structures. In the eighth round (CASP8), our fams-ace2 group participated in the 3D coordinate prediction category as a human group. We applied two different scoring functions for the fams-ace2: (1) the local consensus total score (LocalCons_TS); and (2) the model quality score based on classification of the side-chain environment for each residue. The LocalCons_TS was used as a filter to select the models which have locally similar structures comparing with the set of models. The model quality score was then used for the final selection of the best model. This model quality score was calculated by our model quality assessment program CIRCLE.