

研究講演会レポート

第288回CBI学会研究講演会

バイオインフォマティクスの最近の話題：

次世代シーケンサーデータ解析とプロテオミクスによるバイオマーカー探索



会場風景

第288回CBI研究講演会は、2008年8月22日午後、東大医科研講堂で行われた。今回は東大医科研の中井謙太とエーザイ(株)の河合隆利(現CBI学会長)が世話人を務め、日本バイオインフォマティクス学会の後援を得て、バイオインフォマティクスにおける最近の話題を取り上げた。具体的には、いわゆる次世代シーケンサーの登場がバイオインフォマティクスに与えるインパクトと、実用化に向けて新たな展開を見せ始めたプロテオーム解析におけるバイオマーカー探索である。幸い会員?名、非会員?名の参加があり、この分野への関心の高さが裏付けられた。当日の講演の詳しい内容については、以下の東京農工大学大学院工学府博士前期課程2年増田莉恵氏によるレポートを参照されたい。

(東京大学医科学研究所 中井謙太)

開催趣旨より

従来法と比べ桁違いに高速にDNA配列を決定できる次世代シーケンサーの実用化の時代がやってきた。個人のゲノム配列を現実的な時間とコストで読むことも夢物語ではなくなった。SNPや点変異も感度よく同定でき、定量性を加味することでマイクロアレイの守備範囲であった遺伝子発現解析やChIP-chip実験をも飲み込んでしまう勢いである。しかし、そのデータ生成量は生半かなものではなく、膨大な情報をどう料理するかがバイオインフォマティクスの新たな課題になっている。一方、プロテオミクス研究分野も着実に進展し、最近には特に臨床応用に直結したバイオマーカー探索が世界的にもよく研究されている。そこでも質量分析を用いたプロテオミクスデータからいかにマーカータンパク質を絞り込むか、マーカータンパク質を検出する抗体をいかに効率的に作成するかなど、インフォマティクスの適用分野がますます拡大している。そこで、本講演会ではこれら二つの分野の第一線で活躍しておられる研究者を講師にお招きして、現状をざっくばらんにお話しいただくことにした。

プログラム

1. はじめに 河合隆利(会長・世話人)
2. 次世代シーケンサーの原理と最近の開発動向
北川正成(タカラバイオ株式会社)
3. 大規模シーケンサーを用いたトランスクリプトーム解析
山下理宇(東京大学医科学研究所)
4. 次世代シーケンサーによるメタゲノム解析の可能性
黒川 顕(東京工業大学)
5. 臨床応用へ向けた蛋白質バイオマーカーの探索と検証：プロテオミクスからのアプローチ
川上隆雄(株式会社メディカルプロテオスコープ・東京医科大学)
6. ランダム免疫法による効率的な血清腫瘍マーカーの開発
前田忠計(北里大学)
榎本友美(株式会社ダイナコム)
7. 総合討論

1. 次世代シーケンサーの原理と最近の開発動向

2003年に提唱された「\$1,000 ゲノム」の実現に向け、近年、多種多様なシーケンサーが開発されている。タカラバイオ株式会社の北川氏は、これまでのゲノム配列決定研究を支えてきたサンガー法とは異なる原理に基づく「次世代シーケンサー」の主要な3種の装置について、その特徴を概説するとともに、Next-next と呼ばれる次々世代のシーケンサーの開発動向について紹介した。

現在、次世代シーケンサーは3種の装置が商業ベースに乗っている。

「Genome Sequencer FLX system」は、微小なウェルをもつプレートを用いて、ピロシーケンス法により塩基配列を解析する。本手法は他の2種の解析装置に比べ、一度に解析できる配列長が200から300 bases と圧倒的に長い点の特徴である。また、得られるデータ量は、1ランで約100 Mbと他の2機種に比べると劣るが、サンガー法を用いた従来のシーケンサーに比べ、大幅に向上している。さらに、ウェルの数を増やすことでデータ量の向上を目指している。本手法は塩基の取り込みに伴い生成するピロリン酸を用いて化学発光により検出を行うため1種の塩基が長く続く場合に精度が落ちるといった欠点があるが、これは冗長度を上げることでカバーしている。

「GenomeAnalyzer」は、蛍光標識したdNTPの取り込みを蛍光顕微鏡により検出することで塩基配列を解析する。本手法は、一度に解析できる配列長が35 baseと非常に短い。しかし、1ランにつきシングルエンドリードでは1.5 Gb以上と、大量の解析データが得られる。また、両端からの解析により解析の精度を上げることが可能である。現在は、読み取り配列の正確性向上が課題となっている。

最後に紹介された「SOLiD system」では、ライゲーションによりシーケンス解析を行っている。本手法も、一度に解析できる配列長が35 baseと非常に短い一方で、両端からの解析が可能であり、1ランで得られる解析データは最大6 Gbと非常に大きい。また、サンガー法には劣るが、読み取り配列の正確性は3種の次世代シーケンサーの中で最も高い。

一方、Next-next シーケンサーとしては、一分子シーケンサーが開発の中心である。Helicos Biosciences社が開発を手がけ、既に販売に至っているものは、配列長は18 baseと短い、1ランにつき50 Gbの解析データが得られる。また、より迅速かつ長い配列長取得を目的に、現在3社が開発を進めている。Pacific Biosciences社は、DNA合成反応をリアルタイムで解析する装置の開発を進めている。1kbaseの配列長を1時間で10 Gb検出することを目標にしているが、実用例は未だ報告されていない。一方で、ZS genetics社は電子顕微鏡を用いた装置の計画を発表しており、電子顕微鏡により塩基配列を確認できたという報告があるが、具体的なデータは発表されていない。他にも、蛍光標識ポリメ



北川正成氏

ラーゼと蛍光修飾ヌクレオチドとの相互作用(FRET)を利用した一分子計測によるリアルタイム検出技術がVisiGen Biotechnologies社から提案されているが、こちらもデータの報告はない。一方で、マウスやヒトといった真核生物のシーケンス解析には未だ数千万のコストがかかるという問題がある。この問題には、現在、Roche社とAgilent社が取り組んでいる。

以上のように、シーケンス解析技術は、現在、飛躍的な進歩を遂げている。この技術の進歩は塩基配列取得による構造解析やTag配列取得による機能解析など幅広い解析を支えるものとなるだろう。一方で、得られた膨大なデータの取捨選択、バックアップの形などの検討が今後の新たな課題となる、と北川氏は述べられ、講演を終えられた。

2. 大規模シーケンサーを用いたトランスクリプトーム解析

東京大学の山下特任助教は、トランスクリプトーム解析における次世代シーケンサーの有用性について解説した。具体的には、トランスクリプトーム解析に必要な転写開始点のデータベースを公開しているDBTSSの紹介し、トランスクリプトーム解析の課題であるmRNAの絶対定量およびプロモーターの組織特異性について述べられた。

細胞内でDNAから転写されたmRNAを一度に解析するトランスクリプトーム解析では、発現量の定量や転写開始点の保証されたデータが求められる。しかし、現在、多量に登録されているmRNAに対するcDNA配列は、転写開始点が保証されていないという問題がある。そこで山下氏らの研究グループは、転写開始点が保証されたデータベースDBTSS(<http://abtss.hgc.jp>)の構築を進めている。DBTSSは、mRNAからcDNAを合成する際、5'端に目印をつけることで転写開始点の決定が可能なオリゴキャッピング法を用いて、cDNA配列を解析し、ゲノムへマッピングすることにより、転写開始点を特定したものである。また、mRNAの発現量の定量は、従来までは、マイクロアレイやRT-PCRといったハイブリダイゼーションを基にした技術により、組織別のmRNA量相対値の比較が行われてきた。しかし、これらの手法ではmRNA量の絶

対定量は困難であり、どの遺伝子が最も多く発現しているなどといった遺伝子間の発現量の比較は困難である。一方、次世代シーケンズ技術は多量なcDNAの同定が可能であり、配列数の比較により遺伝子間での発現量の比較が可能となる。そこで山下氏は、次世代シーケンサー・GenomeAnalyzer を用いて5' 端配列の同定を進めている。5' 端配列は、1つのmRNAに対し1つしか存在しないため、より正確にmRNA量を推定できる。また、5' 端配列を同定すれば、転写開始点の同定も同時に可能となる。これまでにHEK293細胞とMCF7細胞の解析を行い、



山下理宇氏

cDNA配列ゲノムへのマッピングが行われた。その結果、約1000万のシーケンズ配列に対し、既存のmRNAと800万がマッピングでき、転写開始点はHEK293細胞では130万箇所、MCF7細胞では68万箇所がそれぞれ同定できた。さらに、現在、8種の細胞種由来の15種類の検体で実験を進めており、合計で約3億以上の配列データが得られている。さらに得られたデータを用いて解析を行った結果、細胞あたりの発現遺伝子量の絶対定量、ある細胞に特異的に発現している遺伝子の同定、細胞間での遺伝子発現量の揺らぎなど、発現プロファイルの有効な結果が示された。

さらに山下氏は、転写開始点も定義できる本手法の特徴を活かし、転写開始を制御するプロモーター配列の細胞特異性について様々な解析を進められている。また、プロモーターの特異性を決定する要因について、プロモーター領域のCpG Islandに着目して検討された。その結果、CpG Islandが存在すれば、非常に強く発現し、遺伝子発現量の変動も小さいことが明らかとなり、次世代シーケンサーを用いた解析の有用性を示された。

3. 次世代シーケンサーによるメタゲノム解析の可能性

東京工業大学の黒川教授は、メタゲノム解析におけるシーケンサーの可能性について話をされた。具体的には、まず、従来のサンガー法に基づくシーケンズ解析によるヒト腸内フローラのメタゲノム解析について述べられ、続いて、次世代シーケンサーを用いたメタゲノ

ム解析の可能性について報告された。

ヒトの健康に密接に関わる細菌は、ヒトの体内から、水中、土壌など、あらゆる場所に存在する。環境中では、多種多様な菌が相互作用しながら生命活動を維持している。そのため、近年、細菌全体としての機能を理解する目的で、群集をまるごとゲノム解析するメタゲノム解析が関心を集めている。黒川氏は、従来型のシーケンサーを用いて、ヒトの腸内細菌のメタゲノム解析を行っている。

まず、生後3ヶ月から45歳の男女、うち2家族を含む13人により、シーケンズ解析を行い、得られた塩基配列のアセンブルを行った結果、約65%のアセンブルに成功した。そこから、人の腸内に存在する細菌の種類はそれほど多様ではないことが示唆された。さらに、得られたゲノム配列の遺伝子予測を行った結果、1検体あたり2万から6万、合計66万の遺伝子予測に成功した。続いて、検体間の相同性解析を行った結果、離乳前の乳児は個人間で多様性が高い一方、大人と同じように食事をしている離乳後の子供は、ある程度、腸内環境が類似していることが明らかとなった。なお、大人においても子供と同様の結果が得られた。また、家族や性別などで腸内環境が類似は見られなかった。さらに、得られたゲノム情報から、アミノ酸配列の相同性解析を行った結果、約6割のみが相同性を示した。その上で、腸内に存在す



黒川 顕氏

る細菌の構成を解析した結果、細菌の構成においても、上記の遺伝子による相同性解析と同様の結果が得られた。一方で、腸内に存在する細菌の種は、人によって大きく異なり、また、6から7割は、未知のものであることが明らかとなった。

続いて黒川氏は、次世代シーケンサーを用いたメタゲノム解析の可能性を検討する為、腸内フローラのメタゲノムデータを用いてシミュレーションを行った。まず、ゲノム配列をランダムに50 baseで区切り、6フレームでORFの抽出を行い、元のデータベースで相同性検索を行った。その結果、次世代シーケンサーによるメタゲノム解析が可能であることが窺えた。しかし、本解析には、新しい遺伝子検索や機能に迫る解析は非常に困難であり、黒川氏が行っているような従来のシーケン

サーによるメタゲノム解析や、個々の細菌のデータベースなどの参照データの充実が不可欠である。そのため、本解析は、食後の体内環境の変化や薬の効果などを経時的に調べる上で有効な手段になると考えられる。

なお、メタゲノム解析には、細菌のゲノム情報を得るためのシーケンス解析が、「いつ」、「どのような環境で」行われたのかといった情報が十分でない、その先の解析は非常に困難を要する。その為、解析データの周りのデータである「メタデータ」が非常に重要となる。なお、必要なメタデータの内容は、現在、検討が進められている。

これまで、シーケンサー開発の現状および応用例について3名の方がご講演下さった。さらに、シーケンス解析のような種類の解析技術から得られる様々なバイオインフォマティクスを、より、臨床現場に応用するための技術について、川上氏、前田氏、榎本氏をご講演下さった。

4. 臨床応用へ向けた蛋白質バイオマーカーの探索と検証：プロテオミクスからのアプローチ



川上隆雄氏

株式会社メディカルプロテオスコープの川上氏は、プロテオミクスを用いたバイオマーカーの探索・検証と、実験により得られた情報の処理の双方における研究開発の現況を講演された。

近年、病態進展の把握や治療効果の予測における定量的な指標、特に、個人に対する最適な医療を実現し得るものとして、バイオマーカーが注目を集めている。その際、バイオマーカー探索のプールの一つとしてプロテオームが注目されており、これまでに様々な探索手法が試みられてきた。しかし、ここ10年で当局認可にまで到達しているバイオマーカーは多くても年4例であり、発見から市場化までは非常に遠いことが現状である。そこで、川上氏らのグループは、プロテオミクスにおいてすでに確立されているショットガン解析法を改良し、LC-MS/MSによって解析を行うことで、バイオマーカーの探索を進めている。

まず、タンパク質の混合物を加水分解によりペプチド

の混合物にする。続いて、LC-MS/MS 解析によりペプチドプロファイルの情報を得る。得られた情報の内、ペプチド溶出時間、イオン化したペプチド分子の m/z 値および検出強度から間接的にタンパク質の量比を解析する。また、生成イオンスペクトル群、生成イオンの m/z 値および検出強度からペプチド/タンパク質の同定を行う。さらに、川上氏らは動的プログラミング法を応用したプロファイル重ね合わせアルゴリズム「internal standard-guided Optimal ALignment, i-OPAL」を開発することで、個別に分析して得られた各試料のペプチドプロファイルを比較解析することを可能にした。

本手法を用いて、川上氏は2つの症例において実際に解析が行われた。1つは初期原発性肺腺癌組織のプロテオーム解析であり、術後の補助化学療法の効果の予測と検証が行われた。もう1つは、膵臓癌の神経浸潤に関するバイオマーカーの探索が行われた。とくに後者では、プロテオミクスによるアライメントとcDNAマイクロアレイによりバイオマーカーの探索を行い、214のタンパク質と49の遺伝子を候補に挙げた。さらに、両データで共通に発現量の増加が観察された1つのタンパク質を詳細に解析したところ、そのタンパク質が実際に術後のリスク因子となり得るという結果が得られた。

さらに今後の課題として、探索においては解析ラインのさらなる精密化を、検証においてはタンパク質を直接検出可能な解析系への効率的な移行などを挙げ、臨床使用に耐えられる手法の確立における問題点を述べられ、講演を終えられた。

5. ランダム免疫法による効率的な血清腫瘍マーカーの開発

最後に、ランダム免疫法による血清腫瘍マーカーの探索について、北里大学前田教授が講演し、株式会社ダイナコム榎本氏がバイオインフォマティクスによる支援ツールを紹介した。

臨床診断に用いられる検体は、主に血液と尿に限られる。その中でDNAを含むものは白血球くらいしかないので、臨床診断はタンパク質解析が主となる。その際、対象となるタンパク質は、主に、「疾病特異的な修飾を受



前田忠計氏



榎本友美氏

けたタンパク質」と「疾患特異的なタンパク質」が挙げられる。しかし、これらは、単離・精製が難しく、また、微量タンパク質も多く、疾病によるタンパク量の変化の解析は困難である。また、仮にタンパク質の探索に成功しても、疾患特異的な変異部位の構造解析、変異部位を保持する抗体の大量生産などが非常に困難となる。そこで前田氏は、抗原ターゲットを絞ることなくガン診断用の抗体を迅速、系統的に探索するランダム免疫法に着目し、研究を進めている。

ランダム免疫法では、まず、腫瘍細胞の抽出液を精製することなく直接、マウスに免疫することで抗体を作製する。マウスから得られた多種多様な抗体産生細胞は、希釈して培養することでモノクローナル化することができる。次に、培養上清を腫瘍細胞に加え、免疫染色を行い、反応の検討を行った結果、様々な染色パターンが観察された。しかし、免疫染色は非常に手順が煩雑であるため、前田氏は新たにドットプロット法により抗体の検

出を試みた。具体的には、細胞抽出液を希釈した後、抗体を加え、結合した抗体のみを検出した。その結果、免疫染色による染色パターンとほぼ一致する結果が得られた。そこで、ガン患者7人と健常者4人の血清を用いて実験を行った結果、両者の血清で抗体反応の違いが観察された。続いて、得られた抗体をウェスタンブロットで検討した結果、反応性が残るものは3分の1程度であることが明らかとなった。ウェスタンブロットでは抗原が変性しているためだと考えられる。しかし、反応性が失われた3分の2に診断に使用可能な抗体がある可能性があり、スクリーニング法の検討が必要であることが示された。続いて、有用抗体における抗原の同定を行った。なお、この抗原同定は、免疫沈降法、細胞抽出液の分画、二次元電気泳動、LC-MSなどの全てを試みる必要がある。しかし、抗原の同定が出来ない場合も多々ある。なお、この抗原同定の速度は、本手法の中でも課題の1つであり、現在、完全長タンパク質の発現系技術を擁する期間と協力することで、改善を目指している。

最後に前田氏は、ガンの早期診断抗体の作製における課題と、患者数の少ないガン疾患における研究機関の役割について述べられた。前者は、早期診断用ガン抗体を作製しても、それを評価するための、早期ガン患者の一定プールが得られないという問題がある。そのため、診断用抗体を現場で多数、使用し、その結果と、患者の方の予後データの蓄積が重要な鍵になる。また、後者は、患者数の少ない疾患は営利企業では開発に着手されないという現状がある。そのため、大学の研究機関の役割は非常に大きく、低コスト、低労力で製造が可能な診断薬の開発が、今後、この問題の解決策となると述べられ、講演を終えられた。

絶
賛
発
売
中

薬づくりの真実 臨床から投資まで

神沼二真 訳

多田幸雄 堀内正監修

(CBI 学会出版)

販売中！ CBI 学会 HP よりご注文いただけます。

http://cbi-society.org/cbi/DDpress/book_DrugDiscovery.html

定価 3,000 円

【読者の声】

- 以前から気になっていた本の和訳があることを、「ファルマシア」の新刊紹介で知りました。翻訳していただき感謝します。
- この本を発見されたこと、紹介の重要性を強く認識されたことに敬意を表します。バイオ・ビジネス、ベンチャー・ビジネスに1989年ころからグローバルに関わっている身には、認識していることばかりですが、アカデミア、製薬企業の研究者、経営者までが、ほとんど認識していないのが現状ですので、本書の価値は非常に大きく、1000部は即売り切れになり、増刷になるものと思います。

「出版経緯」の記事

http://www.scripps.edu/newsandviews/e_20060313/bartfai.html

今後の講演会予定

第292回CBI研究講演会「神経変性疾患の標的と創薬-I」

開催趣旨：現在CBI学会では、“Pathway/Network to Disease and Drug Discovery”を重要な活動指針としている。そしてまず、“Nuclear Receptors and Metabolic Syndrome”に焦点を当てて10月22-24日の国際シンポジウムの開催にこぎつけたが、これに続く領域としては、アルツハイマー疾患、パーキンソン病などを含む神経変性症をとりあげたいと考えている。その理由は、これらの疾患が社会的に大きな問題になっていることと、ER(endoplasmic reticulum、小胞体)ストレスを介してメタボリック症候群 Metabolic Syndrome と深く関係していることを示唆する知見がえられつつあるからである。その可能性を探るために、この講演会では、アルツハイマー疾患、ERストレスとタンパク質の miss-folding, Metabolic Syndrome と Inflammation と神経変性症との関連、治療薬の論理的な開発などの話題を取り上げ、研究の第一線にある気鋭の研究者を講師としてお招きして、解説していただくこととした。これらの幅広い話題に関係した研究者が多数参加されることを期待する。

日時：2008年12月16日(火) 13:10-17:30

場所：東京大学医科学研究所講堂

世話人：神沼二眞(広島大学、東京医科歯科大学)、多田幸雄(大鵬薬品)

プログラム

1. 「ER(小胞体)におけるタンパク質の品質管理ならびにERストレスに対する細胞応答」森和俊(京大)
2. 「神経変性の治療標的としてのストレスシグナル」一條秀憲(東大院薬)
3. 「ERストレス応答経路による代謝制御機構」親泊政一(徳島大ゲノム機能研究センター)
4. 「構造生物学的アプローチによる抗プリオン化合物の開発」桑田一夫(岐阜大医)

講演会参加費：法人賛助会員：無料、個人会員(非営利)：無料、個人会員(一般企業)：¥5,000、ビジター(非営利)：¥1,000、ビジター(一般企業)：¥10,000

出席を希望される方は事前に必ず事務局セミナー受付 seminar@cbi-society.org に連絡してください。

第293回CBI研究講演会「アカデミアにおける創薬と計算化学」

開催趣旨：現在、「薬をつくる」ことは専ら製薬企業で行われています。それは、薬となる候補化合物を、実際に規制の対象である医薬品に仕上げる過程に巨額の資金がいるからです。この30年ほどの間に、多くのいわゆるバイオベンチャーが創薬に挑んできましたが、ほとんどが失敗してきました。その理由は、科学としての創薬には、さまざまな分野を統合した知が必要なことと、開発過程の後半に巨額の資金がいることが挙げられています。一方、ヒトゲノム解読計画は、多くの画期的な薬を世に出すことに寄与すると期待されていましたが、現実はまだ、そうなっていません。そこで注目されているのは、アカデミアにおける創薬への取り組みですが、そこには新しい戦略思考が要求されます。CBI学会は、1981年の活動開始以来、産官学の研究者が連携して、論理的な創薬方法論の普及を基本目標としてきました。いまやそこで蓄積してきたさまざまな経験と知恵を、新しい戦略思考に転換すべき時が来たと感じております。Workshopの形式をとったこの講演会は、このことを目的として開催されます。この課題に関心のある幅広い関係者の参加を期待しています。なお、CBI学会が刊行している「薬づくりの真実」と、G. P. ピサノ著、「サイエンス・ビジネスの挑戦」、日経BP、2008が、このWorkshopのよい背景資料になっています。

日時：2009年1月16日(金) 13:00-17:00

場所：東京大学山上会館大会議室

世話人：岡部隆義(東京大学)、多田幸雄(大鵬薬品)、神沼二眞(広島大学、東京医科歯科大学)

プログラム

1. 趣旨説明、多田幸雄(大鵬薬品)
2. 「アカデミアにおける創薬基盤」岡部隆義(東京大学)
3. 「大学が有する化合物利用の仕組みづくり」奥山彬(NPO 法人化合物活用センター)
4. 「大学化合物ライブラリーを用いたパーキンソン病治療薬の *in silico* 開発」有賀寛芳(北大大学院薬学研究科)
5. 総合討論：創薬に関する産学連係の戦略
大学の視点から(安西偕二郎、日本大学薬学部)、医薬品開発の産学連係とは(多田幸雄、大鵬薬品)、
アカデミアへの計算創薬の開放(神沼二眞)、他。

講演会参加費：法人賛助会員：無料、個人会員(非営利)：無料、個人会員(一般企業)：¥5,000、ビジター(非営利)：¥1,000、ビジター(一般企業)：¥10,000

出席を希望される方は事前に必ず事務局セミナー受付 seminar@cbi-society.org に連絡してください。

発行：情報計算化学生物学会(CBI学会)事務局 〒158-0097 東京都世田谷区用賀 4-3-16 イイダビル 301

Tel.03-5491-5423 Fax.03-5491-5462 cbi-staff@cbi.or.jp http://www.cbi.or.jp/
