

バイオインフォマティクスによる生物機能開発



九州大学大学院農学研究院生物機能部門 久原 哲

1990年代に巻き起こったゲノム研究の急激な発展は、その塩基配列決定法の飛躍的な進歩が最も大きな要因である。1995年のインフルエンザ菌の全ゲノム配列の決定を皮切りに2002年のヒトゲノム配列の決定を経て多くの生物のゲノム配列が決定され、現在では、さらなる改良・開発が継続されヒトゲノムが1,000ドルで解析できる状況になり、より早く、安く、長く塩基配列が決定できる時代となっている。これらのゲノム解析技術の発展は従来からの解析手法では処理不可能な巨大なデータを生み出し、コンピュータを用いたデータの解析法の開発が喫緊の重要課題となってきている。また、ゲノム配列決定後の網羅的オミックス解析に至っては、コンピュータを用いた網羅的な解析手法の開発なくしては行えない状況となってきている。われわれは、この網羅的解析法、特に1)ゲノム配列解析手法、2)遺伝子発現解析手法、3)発現制御ネットワーク解析手法などをわが国において先駆的に生物学分野に導入し、バイオインフォマティクス分野の先導的開拓を行ってきた。

1. ゲノム配列解析手法の開発と応用

DNA データバンク GenBank が設立され DNA シークエンスデータの配布が始められた1980年代の始めに、われわれは世界に先駆けてDNA-データベースシステム GENAS (*Nucleic Acids Res.*, 1984) を九州大学大型計算機センターに構築しネットワーク上に公開した。このシステムはGenBankの塩基配列データベース、アミノ酸配列データベースを基盤とし、解析用のプログラムを付加した本格的データベースシステムであり、多くの研究者に利用された。このシステムを応用して、柿枝らの研究グループと共同で当時開発されたダイデオキシDNA塩

基配列決定法このデータベースとを基盤として、霊長類の反復配列であるL1ファミリーが1,294アミノ酸配列タンパク質をコードしていることを明らかにした。決定されたアミノ酸配列をアミノ酸配列データベースに対してホモロジー検索を行い、図1に示すように、この配列が出芽酵母のRNA依存DNA polymeraseなどに高い類似性をもつことを明らかにし1986年に*Nature*誌に発表した。この塩基配列解析技術を微生物のゲノム解析分野に応用し、バイオインフォマティクスの基盤を構築してきた。2000年代に入ると、ゲノムのショットガンシーケンシング法が開発されたことに伴い、いっそうのコンピュータを用いた解析が展開された。2001年にPNASに発表した*Clostridium perfringens* (ウエルシュ菌) はグラム陽性、嫌気性の桿菌であり、多数の毒素タンパク質を産生し、その協調作用でガス壊疽を起こす病原性細菌であり、全塩基数が3,031,430 bp、2,260個のタンパク質をコードすることを明らかにした。また、その病原性発症の機構は、一般の病原性細菌で見られるような、挿入された病原性のカセットに存在する原因毒素で引き起こされるものではなく、コラゲナーゼ、ヒアルロニダーゼなどの20個以上の病原性関連の酵素や毒素がゲノム上に散在し、これらの酵素や毒素がそれぞれの標的を効率よく破壊し、分解することによって成立することを明らかにした。

このようなゲノム解析手法は2000年以降飛躍的に拡張され、現在多くのゲノム解析支援システムの構築につながり、ゲノム配列決定の効率化の基盤となっている。

2. トランスクリプトーム解析手法の開発と応用

決定されたゲノム配列を基盤とし、マイクロアレイを用いた



図1 L1ファミリーのコンセンサス ORF のアミノ酸配列と RNA 依存性 DNA ポリメラーゼの類似性

トランスクリプトーム研究が2000年から増加してきた。この研究領域は全く新規に立ち上がったため、従前の機器が全く利用できず、基盤となる機器開発から始めなければならなかった。われわれのスタンスとしては、日本での技術開発を念頭におき、研究の基盤となるマイクロアレイの作製システム (SP-BIO) の開発を日立ソフトウェアエンジニアリング株式会社と測定機器 (FAL1000) の開発を富士フイルム株式会社と行い、現在のトランスクリプトーム解析分野の発展に大きく貢献し、研究コミュニティとの共同研究を行ってきた。その一例として、多くの遺伝子の協調作用として病原性を起こすことで知られているウエルシュ菌を対象として、病原遺伝子の発現機構の解明を行った。2成分制御系の遺伝子破壊および異なる培養条件を含めた100条件でのトランスクリプトーム解析を行い、図2に示すようにウエルシュ菌の2成分制御系センサー VirS が、外界

のシグナルをリン酸化を介して制御タンパク質 VirR に伝え、活性化 VirR は37の遺伝子の転写を直接制御することを明らかにした。加えて VR-RNA 遺伝子は転写調節RNAとして働き、下流にある複数の毒素遺伝子を含む94の遺伝子の転写を制御することを示した。このように VirR/VirS システムの包括的制御下にある遺伝子群はウエルシュ菌のガス壊疽の発症に強く関与している可能性を明らかにし、VirR/VirS システムによるグローバル制御系が本菌の病原性発現に関与するメカニズムであることを明らかにした。

これらのマイクロアレイを用いたトランスクリプトーム解析は現在多くの生物で適用され、発現制御解明に多大な貢献をしている。

3. 遺伝子発現ネットワーク解析手法の開発と応用

次に技術的に大きく発展したのが、トランスクリプトーム解

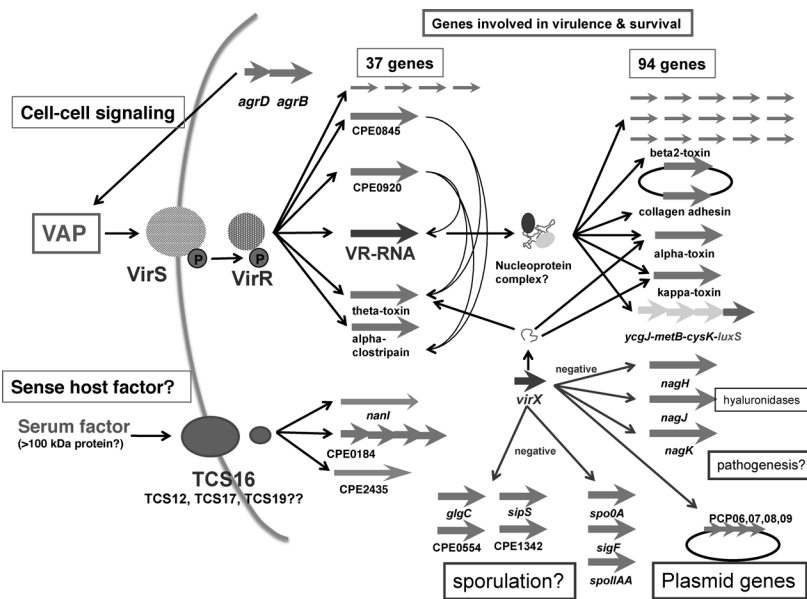


図2 ウエルシュ菌における VirR-VirS の2成分制御系を中心とする毒素関連遺伝子の発現制御システム

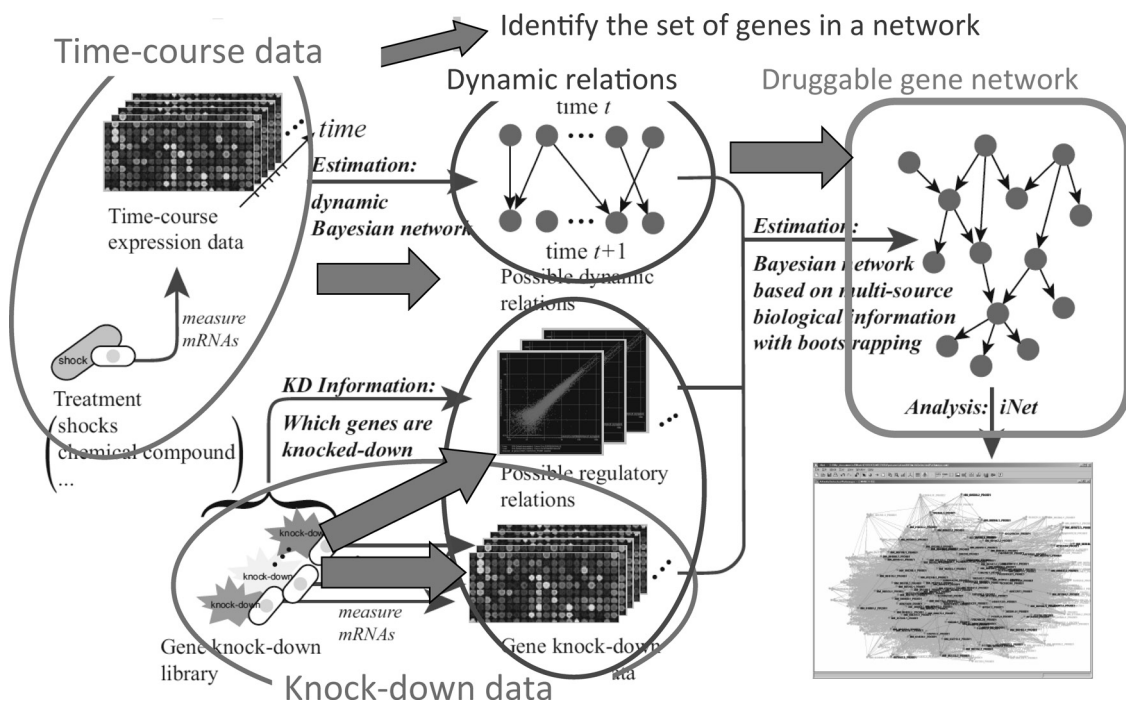


図3 遺伝子発現データからの遺伝子発現制御ネットワーク構築システムの概要

析である。特に、遺伝子発現の制御を網羅的に明らかにする試みが行われ、現在ではヒト全遺伝子産物を対象にした遺伝子ネットワーク推定が可能となり、創薬ターゲット遺伝子などの推定が可能となってきている。われわれは、遺伝子をグラフ上のノード、遺伝子と遺伝子の制御関係をノードとノードを結ぶエッジ、遺伝子の制御関係のネットワークは実験データを満足するグラフィカルモデルの構造推定であると考え、図3に示したベイジアンネットワークを基盤とした大規模遺伝子ネットワークを推定するためのアルゴリズムを開発した。このベイジアンネットワークはマイクロアレイデータに基づく遺伝子ネットワーク推定において極めて広く用いられている。

この方法論を用い、ヒト血管内皮細胞のプライマリーカルチャー細胞に対して270遺伝子のノックダウンのマイクロアレイ発現データと高脂血症薬（フェノフィブラート）を添加した

タイムコースのマイクロアレイ発現データとを統合して解析した。その結果、世界で初めて約1,000遺伝子から構成される遺伝子発現制御のネットワークの構築に成功した。このネットワークの中から、フェノフィブラートのターゲット遺伝子であるPPAR α が関与する制御パスウェイを明らかにした。また、これらのネットワークの構造を検討し、ハブ遺伝子の概念を導入し新たな創薬ターゲットの候補を発見している。

この解析プロセスはこれ以後多くの薬剤開発に適用されている。候補者らが2000年に遺伝子発現制御解析の手法を発表して以来、欧米でも遺伝子発現制御ネットワーク解析が盛んになり、バイオインフォマティクスの大きな分野を形成した。この波及効果は今や農芸化学分野を含む全生物学分野の研究に及んでいる。