

遺伝子位置情報による巨視的ゲノム比較：グラフィカル・ガウシアン・モデリングによる遺伝子制御ネットワーク推定

●堀本 勝久¹⁾ ◆藤 博幸²⁾

1) 東京大学医科学研究所 2) 九州大学生体防御研究所

〈研究の目的と進め方〉

2000～2003年度の3年間は、ゲノム構造の巨視的な比較の開発と適用を行った。

ゲノム配列の決定に伴って、多数の遺伝子が同定され、さらに配列の類似性に基づきゲノム間での遺伝子の類縁関係も調べられ、これらは数種のデータベースに整理・収集されている。このように個々の遺伝子の配列類似性に関する観点からの研究は盛んに行われているが、これらの知識に基づいてゲノムの全体構造を比較する観点からのゲノム研究は極めて少ない。現在、遺伝子の配置に関する数理的研究は、カナダのSankoffらのグループによる遺伝子の並びの順序に着目した研究があるが、この方法は計算時間に莫大な時間がかかることや正確な遺伝子配置の代わりに遺伝子の順序を変数とすることなど、実際のゲノム解析に未だ十分に開発されていない。

ゲノムの全体構造の相異を巨視的に捉える一つの試みとして、遺伝子配置データに基づいてゲノム構造を比較する数理的方法を提案する。ゲノム構造を忠実に反映した形で比較するために、原核生物を中心とする環状ゲノムと真核生物を中心とする鎖状ゲノムそれぞれについて方法を開発し、原核生物の環状ゲノムについてはミトコンドリアゲノムによって、真核生物の鎖状ゲノムについては酵母ゲノムによって、それぞれの方法の有効性を検証する。さらにこれらの方法の開発・改良によって、膨大な遺伝子情報データについて、ゲノムの全体構造における遺伝子配置の相異という、ひとつの新しい巨視的な観点を提供することができる。このように本研究は、ゲノム上にコードされる遺伝子の配置情報に基づいて、ゲノムの全体構造の比較を行う数理的方法の開発と全構造が決定されたゲノムへの適用を目的とする。具体的には、

1. 細菌ゲノムにおいて、コードされる相対的な位置が保たれている遺伝子群が存在するのかを調べる
2. また、存在すればどのような機能を持つ遺伝子群であるのかを調べる。
3. 細菌ゲノム全体の重複が起こった可能性があるかどうかを調べる。
4. 一生物種のもつ鎖状ゲノムの複数のクロモゾーム間の相異を、遺伝子配置の観点から比較する方法の開発を行う。
5. 4.の方法を用いて、定性的に示唆されている真核生物ゲノムの重複について、遺伝子配置の観点から、具体的にどのクロモゾームが重複したペアであるのかを明らかにする。
6. 真核生物間で複数の鎖状クロモゾーム上にコードされる遺伝子の配置を比較する方法を開発する。
7. 6.の方法によって、真核生物ゲノム間で遺伝子配置の比較をおこない、遺伝子配置からみたクロモゾーム間の類似性を数値化する。
8. 比較法をWeb上で公開するため、解析ツールを開発し、一般に利用できるようにする。

上記の研究により、ゲノムのマクロな構造のダイナミ

クスを、機能及び進化の観点から、数量データに基づいて解明できることが予測される。上記目的の1, 2について、急速な細菌ゲノム情報の蓄積を鑑みて、多数の細菌ゲノムについて解析を行う。目的3についても、遺伝子配置の観点から定量的な結果を導くことを試みる。また、真核生物についても、目的4, 5, 6, 7の問題意識の基でゲノムの全体構造の比較法を確立する。目的8によって、その解析法を公表し我々の解析結果との照応によって、より一般的な知見の確立を目指す。

さらに、現在の計画では、取り扱うデータを類縁関係のある遺伝子の配置情報に限定しているが、その他の何らかの関連性が示される遺伝子配置情報データへの適用は容易である。例えば、分子間相互作用や遺伝子発現に関するデータについて、それらの遺伝子の配置情報を解析することで、相互作用や発現と遺伝子配置との関連を数量的に解明することも試みる。

2003年度と2004年度の2年間は、グラフィカル・ガウシアン・モデルの遺伝子発現プロファイルへの適用による制御ネットワーク推定を行った。

この研究は、多変量解析法の一手法であるグラフィカル・モデリングに基づいて、直接的な相互作用を間接的な相互作用から判別し相互作用の全体像を推定する方法を開発し、遺伝子発現プロファイル、蛋白質配列・立体構造情報などの生物情報に適用することで、生命を構成する物質間の関係性の理解を深めることを目的とする。

具体的には、

1. 階層的クラスタリング法とグラフィカル・ガウシアン・モデリングの組み合わせによる遺伝子間の制御関係の枠組みを推定する方法を開発し、量的データからの相関係数行列を入力するだけで、相互作用の枠組みを自動的に推定するシステムを構築する。
2. 1.のシステムを遺伝子発現プロファイルデータと蛋白質配列・立体構造データに適用し、遺伝子ネットワーク推定及び立体構造上相互作用に関与する領域の推定をそれぞれ行う。
3. 1.によって推定されたネットワークを構成する遺伝子について、ネットワーク形成と遺伝子配置との関連について解明する。
4. 1.のシステムをWeb上に公開する。
5. 質的データ解析のために、グラフィカル対数線形モデルに基づく推定法を開発し、定性的な関係性データによる相互作用推定を試みる。

本研究で採用する統計的手法は、解析の前提となる条件が極めて緩く、正確な測定が困難な場合が多い生物情報の解析に適した手法の一つである。また、第一次近似としての知見が比較的速やかに得られる。さらに、解析結果について微分方程式やアルゴリズムに基づく因果律的なアプローチとの連携によって、さらに詳細な知見が得られる可能性が高い。

〈研究開始時の研究計画〉

2000年度

1. ゲノム比較法の開発
環状ゲノム比較法を開発する。特に統計的有意性を見積もるためにシミュレーション法の点について重点をおく。そして、開発された方法を、実際のデータ解析のためにワークステーション上でプログラム化する。
2. 遺伝子配置情報の整備
各種データベースより、細菌ゲノム及び酵母ゲノムの遺伝子配置情報を収集し、計算プログラムに取り込めるように整備する。
3. ゲノム比較の実行
ワークステーション上で、収集した配置情報データについて、解析プログラムを実行する。
- 2001年度
4. ゲノム比較法の開発
鎖状ゲノム比較法については、ひとつの生物種のゲノム上で異なるクロモゾーム間の遺伝子配置の比較法を開発し、さらにこの方法を異なる生物種ゲノム上のクロモゾーム間の遺伝子配置を網羅的に比較する方法に拡張する。
5. 遺伝子配置情報の整備
各種データベースより、鎖状ゲノムの遺伝子配置情報を収集し、計算プログラムに取り込めるように整備する。この際、プログラムとデータとの互換性をもたせるため、ワークステーションにおいて作業をおこなう。
6. ゲノム比較の実行
開発した方法を、酵母ゲノムの遺伝子配置データに適用する。
7. 比較法のweb上への公開
比較法の解析ツールを開発し、web上で一般に公開する。この際、大量なデータの通信が見込まれるので専用のwebサーバを設置する。
- 2002年度
8. 環状ゲノム比較法の確立
現在の環状ゲノム配置比較法について、類似度の統計的評価と計算速度の点に関して改良を行う。
9. 遺伝子配置情報の整備
各種データベースより、46の細菌ゲノムのorthologous遺伝子及び13の細菌ゲノムと酵母ゲノムのparalogous遺伝子に関して既存のデータベースから遺伝子位置情報を収集し、解析プログラムに取り込めるフォーマットに整備する。この際、整備データのプログラムへの取り込みを迅速にするために、ワークステーションにおいてすべての作業を行う。
10. ゲノム比較の実行
ワークステーション上で、細菌ゲノム及び酵母ゲノムの配置情報データについて、解析プログラムを実行する。
11. 鎖状クロモゾーム比較法の開発
異なる生物種間の複数鎖状クロモゾームの遺伝子配置比較法開発の予備的な調査として、様々な既存の統計解析法を適用する。この適用は、統計ソフトウェアパッケージをインストールしたパソコンで行う。
12. 解析ソフトウェアの公開
web上で、解析ソフトウェアを利用できるwebサイトの作成を行う。2002年度は、利用者の入力した遺伝子位置データから、比較解析が行えるように利用環境を整える。
- 2003年度
13. データの収集と整備
遺伝子発現プロファイルについて、web上に公開されている遺伝子発現データを収集し、プログラムに取り

- 込めるように整備する。同時に立体構造既知の蛋白質配列についても、各種データベースより収集し、配列アライメントを行う。
14. 自動グループ化法と相互作用推定法を統合したシステムの構築
クラスター法を利用し自動的にグループ化されたデータから、グラフィカル・ガウシアン・モデリングによって直接相互作用を推定する方法を、一つのシステムとして実行できるように統合する。
 15. 統合システムの遺伝子発現プロファイルへの適用
統合システムによって、遺伝子の相互作用を推定し、推定結果を転写因子による制御関係などの既知の知識と照応を行う。この解析は主に研究代表者が行う。
 16. 統合システムの配列・立体構造データへの適用
統合システムによって、配列の相互作用領域の推定を行い、その結果を、立体構造データを照応し、評価を行う。この解析は主に研究分担者が行う。
 - 2004年度
 17. 遺伝情報データの収集と整備
2003年度開発したシステムを用いて解析可能にするため、最新のデータについて、収集・整備する。遺伝子発現プロファイルについては、web上に公開されている遺伝子発現データを収集し、プログラムに取り込めるように整備する。同時に立体構造既知の蛋白質配列についても、各種データベースより収集し、配列アライメントを行う。研究代表者と分担者との共同で行う
 18. システムの遺伝子発現プロファイルへの適用
システムによって、遺伝子の相互作用関係を推定し、推定結果を転写因子による制御関係などの既知の知識と照応を行う。この解析は主に研究代表者が行う。
 19. システムの配列・立体構造データへの適用
システムによって、配列の相互作用領域の推定を行い、その結果を立体構造データと照応し、評価を行う。この解析は主に研究分担者が行う。
 20. システムの確立
上記18と19の適用に際して生じる現在のシステムの問題点を洗い出し、それらを解決することでより汎用的な相互作用推定システムの確立を試みる。研究代表者と分担者との共同で行う。
 21. Web siteの結果表示法の改良
2003年度システムをのweb siteを公開したが、解析結果の表示法に関して不十分な点が多い。利用者からの様々な意見を取り入れながら表示法を改良し、利用の促進を図る。主に研究代表者が行う。
 22. 遺伝子相互作用と遺伝子配置との関連の解明
相互作用が推定された遺伝子間の配置を数量的に解析し、相互作用と遺伝子配置の関連を解明する。主に研究代表者が行う。
 23. グラフィカル対数線形モデルに基づく相互作用推定法の開発
生物情報における質的データを利用した相互作用推定のために、グラフィカル対数線形モデルに基づいた相互作用推定法を開発する。主に研究代表者が行う。

〈研究期間の成果〉

2000年度

ゲノムの全体構造の相異を巨視的に捉える一つの試みとして、遺伝子配置データに基づいてゲノム構造を比較する数理的方法を提案した。ゲノム構造を忠実に反映した形で比較するために、原核生物を中心とする環状ゲノムと真核生物を中心とする鎖状ゲノムそれぞれにつ

いて方法を開発し、原核生物の環状ゲノムについてはミトコンドリアゲノムによって（研究業績 25）、真核生物の鎖状ゲノムについては酵母ゲノムによって（研究業績 26）、それぞれ方法の有効性を検証した。これらの方法の開発・改良によって、膨大な遺伝子情報データについて、ゲノムの全体構造における遺伝子配置の相異という、ひとつの新しい巨視的な観点を提供することができる。また、ゲノムのマクロな構造のダイナミクスを、機能及び進化の観点から、数量データに基づいて解明できることが予測される。さらに、現解析では、取り扱うデータを類縁関係のある遺伝子の配置情報に限っているが、その他の何らかの関連性が示される遺伝子配置情報データへの適用は容易である。例えば、分子間相互作用や遺伝子発現に関するデータについて、それらの遺伝子の配置情報を解析することで、相互作用や発現と遺伝子配置との関連を数量的に解明することが可能になる。

2001年度

既存のデータベースから収集・整備した19生物種の細菌ゲノムのorthologous遺伝子配置情報について適用した。その結果、細菌ゲノムにおいて、相対的な位置が類似な遺伝子群が存在すること、最も類似な配置は複製点を一致させる方向にゲノムを回転した際に実現されること、相対的位置が類似な遺伝子群には遺伝子機能の観点から分類した場合統計的に有意に偏りがあること、が判明した(研究業績 24)。

急速な細菌ゲノム構造情報の蓄積を鑑みて、19の細菌ゲノムについて得られた結果を踏まえ、更に多数の細菌ゲノム間のorthologous遺伝子の配置情報について収集・整備を行い、遺伝子群の存在とその機能についてさらに研究を行う準備し、同時に、細菌ゲノム全体の重複が起こった可能性について調べるため、paralogous遺伝子の配置情報も収集・整備した。また、原核生物を中心とする環状ゲノムの比較法の開発に目途がついたので、真核生物に特徴的な鎖状ゲノムの比較法開発とその予備的な調査を開始した。orthologous遺伝子の配置情報については、各種データベースよりデータを収集し、paralogous遺伝子の配置情報については、web上で公開されている9つの細菌ゲノムの情報について解析のためのデータ整備をしたが、収集後の解析において、比較するゲノムの組み合わせが指数関数的に増加し、パソコンのパフォーマンス（特にメモリ）の問題が起こった。また、鎖状ゲノムの比較法開発の予備的な調査のため、遺伝子の配置情報について様々な統計解析法を試みた。

2002年度

環状ゲノム比較法においてbootstrap法を導入することで環状ゲノムの配置類似性に関して統計検定が行えるように改良した(研究業績 5,10)。その方法によって、既存のデータベースから収集・整備した19生物種の細菌ゲノムのorthologous遺伝子配置情報について、相対的位置が類似な遺伝子群には遺伝子機能の観点から分類した場合統計的に有意に偏りがあること、が判明した。

鎖状クロモソーム比較法については、同一生物種の複数クロモソーム間においてparalogous遺伝子の配置を指標にした方法を開発し、酵母ゲノムに適用することでゲノム重複の可能性を探った。ただし、この比較法では、類似度評価に関して不十分な点があり改良の余地がある。また、異なった生物種の複数クロモソームの比較については、定式化は終了しているが、プログラム化及び実際のデータの実行は行っていない。

将来において、遺伝子の配置と発現・制御との関連を探るために、遺伝子発現プロファイルから類似発現パタ

ーンを示す遺伝子の自動グループ化法(研究業績 22, 23)とそれらグループ間の相互作用推定法(研究業績 18, 19)をグラフィカル・ガウシアン・モデル (GGM) に基づいて開発した。

各種データベースより、46の細菌ゲノムのorthologous遺伝子及び13の細菌ゲノムと酵母ゲノムのparalogous遺伝子に関して既存のデータベースから遺伝子位置情報を収集し、解析プログラムに取り込めるフォーマットに整備した。この際、整備データのプログラムへの取り込みを迅速にするために、ワークステーションにおいてすべての作業を行った。

環状ゲノム比較については、急速な細菌ゲノム構造情報の蓄積を鑑みて、46の細菌ゲノム間のorthologous遺伝子配置情報について収集・整備を行い、遺伝子群の存在とその機能についてさらに研究を行う準備をした。同時に、遺伝子配置の観点から細菌ゲノム全体の重複が起こった可能性について調べるため、web上で公開されている13の細菌ゲノムのparalogous遺伝子配置情報を収集・整備を完了し、解析を実行中した。

2003年度

開発済みであった二つの遺伝子発現プロファイル解析法を、発現プロファイルデータに限らず相関係数を入力することで自動グループ化と相互作用推定が可能のようにそれぞれ改良し、それらを統合することで一つのシステムを構築した(研究業績 7)。システムの有効性を検証するために、公開されている酵母遺伝子の発現プロファイルについて適用し、既知の遺伝子制御関係データと照応した結果、遺伝子制御関係に関する高い再現率を得た(研究業績 15)。この解析から、本システムで推定されるネットワークのグラフにおける辺の有無、統計学的には条件付き独立・従属の関係が、分子生物学における転写開始因子とそれが直接制御する遺伝子との関係の強弱を示していることが判明した。このことは、本システムによる理論的な相互作用推定が分子生物学的な現象を探索する上において有効であることを示している。

さらに、本システムをweb上に公開し一般の利用を可能にした (<http://eureka.ims.u-tokyo.ac.jp/asian/>) (研究業績 4, 7)。公開に併せプログラムの改良を行い、従来の計算速度に比べ約15分の1の時間で計算が終了するようになった。さらにweb siteの改良、特に解析結果の表示について改良を図ると共に、システムを利用して、遺伝子分類に止まっている発現データについてさらに解析を進めて相互作用推定を行い、新たな生物学的知見の発見を試みた。また、蛋白質配列と立体構造との関連について、特に蛋白質相互作用領域推定のためにもシステムを利用するため、構造既知の蛋白質について、配列アライメントとその相互作用領域の照応を行った。

2004年度

2003年度構築したウェブサイト公開し (<http://eureka.ims.u-tokyo.ac.jp/asian/>) について、画像データを出力するなど改良を加え、利便性を増した(研究業績 1)。システムの有効性を検証するために、公開されている大腸菌トリプトファン代謝関連遺伝子の及びSOS応答関連遺伝子の発現プロファイルについて適用し、既知の遺伝子制御関係データと照応した結果、遺伝子制御関係に関する高い再現率を得たのに加え、生物学的に整合性の取れる新しい制御関係の候補を推定した(研究業績 2, 6)。このことは、本システムで推定されるネットワークのグラフにおける辺の有無、統計学的には条件付き独立・従属の関係が、分子生物学における転写開始因子とそれが直接制御する遺伝子との関係の強弱を示している

ことを再確認し、本システムによる理論的な相互作用推定が分子生物学的な現象を探求する上において有効であることを示している。さらに4例の発現データについて解析を行った。また、既存の知識を活用することで有向独立グラフ及び連鎖独立グラフを出力し因果関係の推定にまで拡張する試みのために、酵母の制御ネットワークの知見の整備とGGMのモデリング過程の改良を行っている。

蛋白質相互作用推定については、構造既知の蛋白質について配列アライメントとその相互作用領域の照応を行い、そのデータを本システムで解析し予備的な結果を得た。ただし、構造に関する定性的データと解析によって出力される数値データとの照応が、データ間の精度が原因となって難しいことがあった。そこで、解析データ数を増やすと共に、GGMの混合法を実装することで蛋白質構造に関する定性データも解析に組み込む準備をしている。一方、蛋白質相互作用データに関して、実験法に依存した相互作用蛋白質ペアの相異などが指摘されていることを鑑みて、相互作用蛋白質を構造から直接予測する方法の開発を試みた。その第一歩として、仮想的な化学架橋剤によってアミノ酸残基間距離の情報、アミノ酸残基の埋もれ度、二次構造から、立体構造の予測を試みた。その結果、予測された埋もれ度、二次構造に加え、蛋白質の総残基数Nの時、 $N/4 \sim N/10$ の化学架橋剤によってアミノ酸残基間距離の情報で、構造予測が可能であることが判明した。このことは、質量分析との組み合わせにより化学架橋によるアミノ酸残基間距離の情報が得られれば、我々の方法でかなりの精度で立体構造が予測できることを意味し、さらに複合体構造の詳細な予測やそれに基づく蛋白質相互作用の予測の基礎となると考えられる。

遺伝子発現と遺伝子のゲノム上の位置の関連については、相関係数を類似度として網羅的に関連を見積もることを行い、予備的な結果を得た。現在、最終的な結果を算出し報告の準備をしている。ゲノムワイドでの発現量と位置との関連情報は多数の生物種にわたるため、相互作用推定システムの直接的な適用は煩雑な解析結果が得られる可能性があり、まず簡潔な方法で全体像の把握を試みている。

質的データによる相互作用解析に関しては、解析プログラムを作成中である。統計学で想定される通常のデータ量に比べ、ゲノム関連情報が大量であるため、解析実行上の問題点が浮かび上がってきた。

〈達成できなかったこと、予想外の困難、その理由〉

2000～2002年度の3年間のゲノム構造の巨視的な比較法の開発と適用については、遺伝子発現解析法の開発を同時に試みたため、環状ゲノム比較の解析結果をまとめること及び鎖状クロモゾーム解析法の開発が遅れた。また、遺伝子発現データの解析法は開発したが、遺伝子の位置との関連を自動的に照応するシステムの構築はできなかった。

2003年度と2004年度の2年間のグラフィカル・ガウシアン・モデルの遺伝子発現プロファイルの適用による制御ネットワークの研究については、成果発表が予備的・部分的なものに留まり、最終的な成果の発表が遅れた。特に、蛋白質相互作用に関する解析について、予備的な解析にとどまった。

〈今後の課題〉

ゲノムの巨視的比較についての課題は以下の通りである。

1. 環状ゲノム比較法を細菌ゲノムに適用することで、細菌ゲノムの遺伝子配置を遺伝子機能ごとに解析し、遺伝子の位置と機能の関連を詳細に調べる。
 2. 大腸菌と光合成細菌について、ゲノム重複の可能性を、配置情報に加えて遺伝子機能や分子系統樹など多面的な解析を行うことによって探る。
 3. 鎖状クロモゾームの巨視的な比較法の開発及び適用を行う。
 4. 遺伝子発現解析法について、遺伝子の位置との関連付けを行えるようにすると共に、その方法によって細菌及び酵母のデータに適用し、遺伝子配置に関する比較結果と合わせて遺伝子の位置と機能を詳細に解析する。制御ネットワーク推定に関しては、発現プロファイルに加え、様々な生命情報を利用できるようにネットワーク推定システムの適用範囲を拡張し、様々な生命現象について物質に基づいた関係性の解明を試みる。同時にシステムの適用可能性をさらに広げるために、それぞれの生物情報データに特化した方法への改良を試みる。具体的な課題は以下の通りである。
1. 近年開発したGGMによるシステム間の関連構造推定法に基づいた、グラフィカル連鎖モデル (GCM) による因果推定法を確立する。
 2. グラフィカル・モデルの一つであるpath consistency (PC) アルゴリズムの改良によって、分子間関連及び因果推定法を開発する。
 3. 知識ベースに推定された関連・因果モデルを測定データに基づいて評価する方法を開発する。
 4. GGMによる関連構造推定法を、カテゴリカル・データが解析可能な混合法に拡張する。
1. 2. 3. の推定法を統合することによって、遺伝子制御ネットワークの推定精度の向上を図る。
 2. 4. の統合により蛋白質間相互作用領域の推定を行う。
- 2004年以降、ゲノム比較については、以上課題の内2.について成果をみることができた。ネットワーク推定については、1.2.について成果を得られ、3.についても報告の準備段階である。その他の課題については、進行の速度は遅く成果発表に至ってはいないが研究継続中である。

〈研究期間の全成果公表リスト〉

1) 論文/プロシーディング

1. Aburatani, S., Goto, K., Saito, S., Toh, H. and Horimoto, K. ASIAN: A Web Server for Inferring a Regulatory Network Framework from Gene Expression Profiles. *Nucleic Acids Res.*, 33, W659-W664, 2005.
2. Aburatani, S., Sakai, H., Murakami, H. and Horimoto, K. Elucidation of the relationships among the LexA-regulated genes in SOS response. *Genome Informatics*, 16, 95-105, 2005.
3. Murakami, H., Sakai, H., Aburatani, S. and Horimoto, K. Relationship between Segmental Duplications and Repeat Sequences in Human Chromosome 7. *Genome Informatics*, 16, 13-21, 2005.
4. Aburatani, S., Goto, K., Saito, S., Fumoto, M., Imaizumi, A., Sugaya, N., Murakami, H., Sato, M., Toh, H. and Horimoto, K. ASIAN: A Web Site for Network Inference. *Bioinformatics*, 20, 2853-2856, 2004.
5. Sugaya, N., Sato, M., Murakami, H., Imaizumi, A., Aburatani, S. and Horimoto, K. Genome-Size Increase in *Anabaena* sp. PCC7120. *Proceedings of the 8th World Multi-Conference on Systemics, Cybernetics and Informatics*, 7, 35-40, 2004.

6. Imaizumi, A., Sugaya, N., Murakami, H., Sato, M., Aburatani, S. and Horimoto, K. Network Reconstruction of Genetic Relationship in Tryptophan Metabolism from Gene Expression Profiles. Proceedings of the 8th World Multi-Conference on Systemics, Cybernetics and Informatics. 7, 4-11, 2004.
 7. Aburatani, S., Sugaya, N., Murakami, H., Sato, M., Imaizumi, A. and Horimoto, K. Internet Tool for Network Inference: ASIAN (Automatic System for Inferring A Network). Proceedings of the 8th World Multi-Conference on Systemics, Cybernetics and Informatics. 7, 1-3, 2004.
 8. Akutsu, T., Hayashida, M., Tomita, E., Suzuki, J. and Horimoto, K. Protein Threading with Profiles and Constraints. Fourth IEEE Symposium on Bioinformatics and Bioengineering (BIBE2004), 537-544, 2004.
 9. Murakami, H., Sugaya, N., Sato, M., Imaizumi, A., Aburatani, S. and Horimoto, K. Detection of Inter-Spread Repeat Sequence in Genomic DNA Sequence. Genome Informatics, 15, 170-179, 2004.
 10. Sugaya, N., Sato, M., Murakami, H., Imaizumi, A., Aburatani, S. and Horimoto, K. Causes for the Large Genome Size in a Cyanobacterium *Anabaena* sp. PCC7120. Genome Informatics, 15, 229-238, 2004.
 11. Ichihara, H., Daiyasu, H., Toh, H. How does a topological inversion change the evolutionary constraints on membrane proteins? Protein Eng Des Sel. 17, 235-244, 2004.
 12. Daiyasu, H., Ishikawa, T., Kuma, K., Iwai, S., Todo, T., Toh, H. Identification of cryptochrome DASH from vertebrates. Genes to Cells 9, 479-495, 2004.
 13. Stamdley, D., Toh, H., Nakamura, H. Detecting Local Structural Similarity in Proteins by Maximizing the Number of Equivalent Residues. PROTEINS, 57, 381-391, 2004
 14. Toh, H., Osaka, K., Takei, T. Molecular phylogenetic analysis of pulmonary surfactant protein C and the relatives. J. Jpn. Med. Soc. Biol. Interface 35, 24-31, 2004.
 15. Aburatani, S., Kuhara, S., Toh, H. and Horimoto, K. Deduction of a Gene Regulatory Relationship Framework from Gene Expression Data by the Application of Graphical Gaussian Modeling. Signal Processing, 83, 777-788, 2003.
 16. Suzuki T, Yamasaki K, Fujita S, Oda K, Iseki M, Yoshida K, Watanabe M, Daiyasu H, Toh H, Asamizu E, Tabata S, Miura K, Fukuzawa H, Nakamura S, Takahashi T. Archaeal-type rhodopsins in *Chlamydomonas*: model structure and intracellular localization. Biochem Biophys Res Com. 301, 711-717, 2003.
 17. Brudler, R. Hitomi, K. Daiyasu, H., Toh, H., Kucho, K., Ishiura, M., Kanehisa, M., Roberts, V.A., Todo, T., Tainer, J.A., Getzoff, E. G. Identification of a New Cryptochrome Class: Structure, Function, and Evolution. Molecular Cell 11, 59-67, 2003.
 18. Toh, H. and Horimoto, K. System for Automatically Inferring a Genetic Network from Expression Profiles. J. Biol. Phys., 28, 449-464, 2002.
 19. Toh, H. and Horimoto, K. Inference of a Genetic Network by a Combined Approach of Cluster Analysis and Graphical Gaussian Modeling. Bioinformatics, 18, 287-297, 2002.
 20. Daiyasu, H., Hiroike, T., Koga, Y., Toh, H. An analysis of membrane stereochemistry with homology modeling of sn-glycerol-1-phosphate dehydrogenase. Protein Engineering 15, 987-995, 2002.
 21. Akutsu, T. and Horimoto, K. Local Multiple Alignment of Real Numerical Sequences: Detection of Subtle Motifs from Protein Sequences and Structure. Genome Informatics, 12, 83-92, 2001.
 22. Horimoto, K., Aburatani, S., Kuhara, S. and Toh, H. ASIAN - Automatic System for Inferring a Network from Gene Expression Profiles. Res. Commun. Biochem. Cell Mol. Biol., 5, 192-207, 2001
 23. Horimoto, K. and Toh, H. Statistical Estimation of Cluster Boundaries in Gene Expression Profile Data. Bioinformatics, 17, 1143-1151, 2001.
 24. Horimoto, K., Fukuchi, S. and Mori, K. Comprehensive Comparison between Locations of Orthologous Genes on Archaeal and Bacterial Genomes. Bioinformatics, 17, 791-802, 2001.
 25. Horimoto, K. and Fukuchi, S. Feasibility of Gene-Location Distance in Microbial Genome Analysis. INFORMATION, 3, 263-269, 2000.
 26. Fukuchi, S. and Horimoto, K. Classification of Linear Chromosomes. In: Begehr, H.G.W., Gilbert, R.P. and Kajiwara, J. (eds) Proceedings of Second International Congress ISAAC vol.2, Kluwer Academic Publisher, New York, p.1265-1274, 2000.
- 2) データベース/ソフトウェア
ウェブサイトの公開
ASIAN (<http://eureka.ims.u-tokyo.ac.jp/asian/>)
ソフトウェアの商用化
"Auto Net Finder" (インフォコム株式会社)
Option in "Pathway Assist" (Ariadne Genomics Corp, 米国)
- 3) 特許など
なし
- 4) その他顕著なもの
受賞
2001年 Oxford University Press Bioinformatics Prize
from The Twelfth International Conference on Genome Informatics
プレスリリース
日経産業新聞2004年10月21日 p.4
日経産業新聞2004年12月9日 p.4
化学工業日報2004年12月9日 p.6
化学工業日報2004年12月14日 p.9