

# ゲノム比較と系統的相互作用解析に基づく遺伝子・分子ネットワークの解明

(H10年度～H14年度) 研究代表者：金久 實(京都大学) 研究体制：宝バイオ(株)(中核機関)他7機関

## < 研究の概要・目標 >

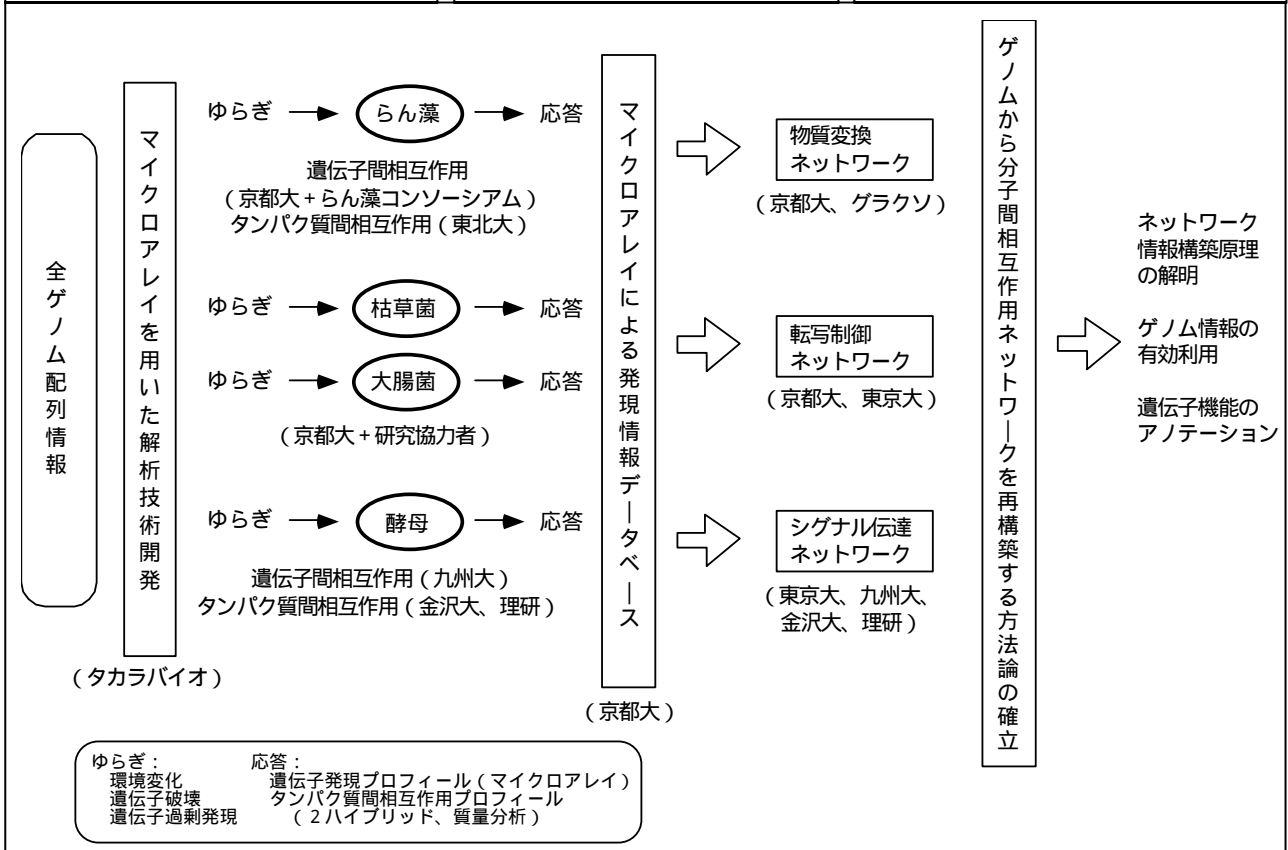
- 1 何を目標しているのか  
ゲノムの全塩基配列情報とマイクロアレイによる遺伝子発現情報、及びタンパク質間相互作用情報から遺伝子のネットワークを解明し、生命システムの機能を解読する。
- 2 何を研究しているのか  
らん藻を中心に、ゲノム全塩基配列がすでに分かっている生物種を対象として、遺伝子・分子ネットワークの階層性、共通性と多様性、遺伝情報と物質情報の関連について研究。また、これに必要な基盤技術を整備。
- 3 何が新しいのか  
様々な生物種や機能に関する遺伝子ネットワークに共通する情報構築原理を確立しようとしているアイデアが新しい。また、情報技術の観点から遺伝子発現やタンパク質間相互作用実験を行う点が新しい。

## < 諸外国の現状 >

- 1 現状及び我が国の水準  
他国においても、ある生物種や特定の機能に着目して遺伝子ネットワークを研究することは行われているが、ネットワークの全体像をカバーする構築原理の確立を目指しているものはない。我が国は原理の解明に必須な遺伝子・分子ネットワークの情報を世界最大のパスウェイデータベース KEGG として蓄積しており、我が国が優位に立っている。

## < 研究推進によるメリット >

- 1 世界の水準との関係  
ゲノム塩基配列情報から生命活動を担う遺伝子のネットワークを推定することが可能となる。これにより、塩基配列は分かっているもののその機能が分かっていない多数の遺伝子について世界に先駆けその機能を推定することができる。
- 2 波及効果  
本研究により解明された原理に基づいて生命活動をコンピュータ上で再現することができ、生物そのものを使わなくても薬物の効果の予測が可能となるなど、生物学、医学、薬学分野への幅広い波及効果が期待される。



## 所要経費（ 期 ）

（単位：千円）

研究項目	担当機関等	研究担当者	H10 年度	H11 年度	H12 年度	所用 経費
1. ネットワーク階層性の原理解明：転写制御ネットワークに関する研究						
(1) 転写因子予測とプロモーター予測に関する研究	東京大学医科学研究所	中井 謙太	5,256	5,667	8,135	19,058
(2) 枯草菌の遺伝子間相互作用に関する実験研究	福山大学工学部	藤田 泰太郎	10,899	51,393	23,080	85,372
(3) 線虫の遺伝子間相互作用に関する実験研究	国立遺伝学研究所	小原 雄治	14,289	14,289	12,973	41,551
2. ネットワーク共通性と多様性の原理解明：シグナル伝達ネットワークに関する研究						
(1) 最適なネットワークを演繹的に再構築する情報処理技術に関する研究	東京大学医科学研究所	高木 利久	8,645	9,246	10,806	28,697
(2) 酵母のネットワーク関連遺伝群に関する研究	東京大学医科学研究所	宮野 悟	9,463	11,356	12,337	33,156
(3) 酵母の遺伝子間相互作用に関する実験研究	九州大学大学院生物資源環境科学研究科	久原 哲	33,638	23,407	20,934	77,979
(4) 酵母のタンパク質間相互作用に関する実験研究	金沢大学がん研究所	伊藤 隆司	20,266	13,432	16,522	50,220
(5) タンパク質のリン酸化ネットワークに関する実験研究	理化学研究所播磨研究所	谷口 寿章			35,867	35,867
3. 遺伝子情報と物質情報との関連に係る原理の解明：物質変換ネットワークに関する研究						
(1) ネットワーク再構築に関する情報科学研究および実験研究	京都大学化学研究所	金久 實	33,988	35,526	44,435	113,949
(2) 発現プロファイル解析の代謝・排出への応用研究	グラクソ・ウエルカム（株）筑波研究所	荻原 淳	5,548	15,590	12,900	34,038
4. 遺伝子・分子ネットワーク解明のための基盤技術開発						
(1) マイクロアレイの解析技術開発に関する研究	宝酒造株式会社バイオ研究所	浅田 起代蔵	52,523	60,578	45,825	158,926
(2) 発現パターン情報のデータベース化	日本 SGI（株）	福島 信宏	8,878	18,524	22,186	49,588
5. 研究管理	宝酒造株式会社バイオ研究所	浅田 起代蔵	1,014	1,346	18,669	21,029
合計			204,407	260,354	284,669	749,430

## 所要経費（ 期 ）

（単位：千円）

研究項目	担当機関等	研究担当者	H13 年度	H14 年度		所用 経費
1．ネットワーク原理の 解明						
(1) ネットワーク再構 築に関する情報科学 研究及び実験研究	京都大学化学研究所	金久 實	66,900	61,015		127,915
(2) 最適なネットワー クを演繹的に再構築 する情報処理技術に 関する研究	東京大学医科学研究所	高木 利久	13,788	4,916		18,704
(3) らん藻のネットワ ーク関連遺伝子群に 関する研究	東京大学医科学研究所	宮野 悟	13,924	14,130		28,054
(4) 遺伝子ネットワー ク予測に関する情報 科学的研究及び実験 研究	九州大学大学院生物資源 科学研究府	久原 哲	19,898	12,707		32,605
(5) 発現プロフィール 解析の代謝・排出への 応用研究	グラクソ・スミスクライ ン(株)筑波研究所	荻原 淳	16,311	17,557		33,868
2．情報技術と実験技術 の融合						
(1) 2 ハイブリッド法 による系統的タンパ ク質間相互作用解析 に関する研究	金沢大学がん研究所	伊藤 隆司	14,249	12,032		26,281
(2) 質量分析法による タンパク質間相互作 用解析に関する研究	理化学研究所播磨研究所	谷口 寿章	41,295	4,601		45,896
(3) らん藻のタンパク 質間相互作用解析に 関する実験研究	京都大学大学院生命科学 研究科	福澤 秀哉	21,885	20,377		42,262
(4) らん藻のタンパク 質複合体のプロテオ ーム解析に関する研 究	東北大学大学院生命科学 研究科	東谷 篤志	25,051	10,773		35,824
(5) マイクロアレイの 解析技術開発に関す る研究	タカラバイオ株式会社	浅田 起代蔵	49,069	49,238		98,307
3．研究推進	タカラバイオ株式会社	浅田 起代蔵	1,354	13,823		15,177
合計			283,72 4	221,16 9		504,893

# 研究成果の概要

## 総括

本研究では、ゲノムシーケンシングが明らかにする全遺伝子のデータ(生命システムを構成する部品のカatalog情報)と、トランスクリプトームやプロテオームに関する系統的实验による遺伝子間または分子間の相互作用データ(部品間のつながりの情報)を、新しい情報技術で融合することにより、ゲノムから生命システムを再構築する方法論を確立し、その応用の可能性を探ることを目的とした。とくに、らん藻を中心とし、枯草菌、大腸菌、酵母との比較解析も含め、マイクロアレイによる発現プロファイル解析、ならびに2ハイブリッドシステムと質量分析によるタンパク質間相互作用解析を行い、システム再構築の方法論を実践した。バクテリアのマイクロアレイ解析は、本研究が我が国において最初に本格的に取り組み、実験技術・情報処理技術の開発・改良とともに、その有用性を明らかにした。同時に我が国のらん藻コミュニティおよび枯草菌コミュニティとの共同研究を推進し、マイクロアレイデータはKEGG/EXPRESSIONデータベースとして解析システムとともに公開し、また遺伝子機能アノテーションの情報はCYORFデータベースおよびBSORFデータベースとして集大成した。当初計画にある実験と情報を融合した技術開発の部分は十分に達成され、ゲノムからネットワーク原理の解明を目指した研究も個々の成果としては達成することができた。らん藻ゲノムに多数残っている未知遺伝子の系統的機能予測を行うことについては、我が国のらん藻コミュニティと連携し、新しい実験データや文献情報をもとに約1/3の遺伝子の再アノテーションを行った。本研究には原著論文等の成果以外に、KEGG/EXPRESSION、CYORF、BSORFといった研究基盤情報を整備した成果があり、とくにCYORFはらん藻遺伝子に関する全世界の研究者の知識を集約した「コミュニティデータベース」と位置づけ、国際的ならん藻コミュニティとの連携研究にまで発展した。

## サブテーマ毎、個別課題毎の概要

本研究は「ネットワーク原理の解明」と「実験技術と情報技術の融合」の2つの研究項目に分け、それぞれ5つの研究課題を設定した。「ネットワーク原理の解明」では、転写制御ネットワーク、シグナル伝達ネットワーク、物質変換(代謝)ネットワークにおいて、階層性、共通性と多様性、遺伝情報と化学情報の二面性を理解するため以下の研究を行った。マイクロアレイデータを他のゲノム関連データと統合して分子間相互作用ネットワークを予測する情報技術を開発し、らん藻が外界の環境変化(化学情報の変化)や特定遺伝子の破壊(遺伝情報の変化)に対してどのような応答をするかのマイクロアレイ実験を行い、またらん藻遺伝子機能の再アノテーションを行った(金久)。シグナル伝達パスウェイに関するデータや知識を計算機上に表現する方式を考案し、入力、表示、検索、推論を行うシステムを開発した(高木)。マイクロアレイデータから個々の遺伝子間の関係をネットワークとして抽出するデータマイニング技術を開発し、その情報管理・可視化解析システムを作成して、酵母とらん藻で遺伝子制御関係の解析を行った(宮野)。酵母の270種の転写制御遺伝子の破壊株を中心に500を超えるマイクロアレイプロファイル解析を行い、グラフィカルモデリング法等により発現制御ネットワークの解析を行った(久原)。薬物代謝等の二次代謝系を対象とし、基質物質の添加で惹起される発現プロファイルの変化から、関与する酵素群を探索する方法の開発と予備的な実験を行った(荻原)。「実験技術と情報技術の融合」ではマイクロアレイ技術だけでなく2ハイブリッド法や質量分析法を加え、情報技術と実験技術を融合した以下の研究を行った。酵母の系統的2ハイブリッド解析によりタンパク質間相互作用を網羅的に解析した後、相互作用モチーフの発見につながる保証付き逆2ハイブリッド法および関連の周辺技術を確立した(伊藤)。質量分析法を基盤とした微量タンパク質同定法、およびそこから得られる部分配列情報をゲノムに対応させる技術を開発し、枯草菌等において新規遺伝子、遺伝子予測のエラーを見いだすことができた(谷口)。らん藻でタンパク質間相互作用を網羅的に解析するため、大腸菌を用いた2ハイブリッド法と遺伝子破壊株を使った機能解析システムを構築し、光合成関連遺伝子およびストレス関連遺伝子の解析を行った(福澤)。らん藻のタンパク質複合体を質量分析で同定するため、細胞抽出液に化学架橋剤を添加する方法、および組み換えタンパク質を利用する方法を確立し、強光ストレスを与えて発現変動するタンパク質群の解析を行った(東

谷)。高品質なマイクロアレイ作製のための標準的な操作法を開発し、同時に解析データの評価法、品質チェック法等の最適化を行って、本研究の基盤となるらん藻マイクロアレイを作製し、また新世代の DNA チップやプロモータマイクロアレイ等の作製も行った(浅田)。本研究の最終年度には公開シンポジウム「バイオインフォマティクスに基づくポストゲノム研究のフロンティア」(平成 14 年 10 月 11~12 日、プラザ平成)を開催し、外国人招待講演 4 件、日本人招待講演 9 件(うち本研究担当者 7 件)、それにポスター発表により、本研究の成果発表を行った。

#### 波及効果、発展方向、改善点等

実験生命科学と情報科学の融合は、ゲノムやプロテオームに代表される生命科学の大量データの処理・解析のために情報科学の実用的なツールが不可欠であることから始まった。すなわち、情報科学は実験プロジェクトを単にサポートする役目だったわけである。これに対して本研究では、情報科学の概念や方法論を実験プロジェクトを立案する段階から取り入れること、あるいは情報科学の予測を実験で検証することなど、真の意味での実験生命科学と情報科学の融合を目指し、これはある程度達成できたと考えている。情報系研究者が代表者の実験研究プロジェクトとして、本研究は恐らく我が国で最初の試みであり、またプロジェクト内だけでなく、実験系の研究コミュニティとの連携も成功させたことは、今後のゲノム関連の研究プロジェクトの編成に影響を与えるだろう。本研究はマイクロアレイを中心とした技術を開発し、それを普及される役割を果たした。ゲノム研究はもともと基盤的な研究である。情報の点でも技術の点でも、ゲノム研究の成果は基盤情報あるいは基盤技術として研究コミュニティに普及させべきものである。実際、らん藻研究は本プロジェクトが提供したマイクロアレイ技術により、日本が米国に 2 年の差をつけたと言われている。また、らん藻コミュニティの知識を集約したコミュニティデータベースをいち早く立ち上げ、米国が日本の基盤情報データベースに依存する形になった。今後の発展方向として、個々の分野でこのような基盤技術や基盤情報を開拓していくことが、我が国の国際競争力の優位性につながると期待される。本研究には遺伝子・分子ネットワークの解明というタイトルがついているが、実際には解明のための技術開発が中心であった。らん藻コミュニティとの共同研究には、生物学的に面白い結果がたくさん得られているが、プロジェクト内部での生物学的な成果は必ずしも期待通りではなかった。これは、研究担当者の多くが他のプロジェクトにも関与しているため、成果の切り分けが難しかったことが一つの要因として考えられる。複数の異なるタイプのプロジェクトに関与することの相乗効果を積極的に評価することが望まれる。

# 研究成果公表等の状況

## (1) 研究発表件数

	原著論文による発表	左記以外の誌上発表	口頭発表	合計
国内	第 期 2 件 第 期 2 件	第 期 10 件 第 期 12 件	第 期 10 件 第 期 33 件	第 期 22 件 第 期 47 件
国際	第 期 18 件 第 期 35 件	第 期 0 件 第 期 9 件	第 期 14 件 第 期 26 件	第 期 32 件 第 期 70 件
合計	第 期 20 件 第 期 37 件	第 期 10 件 第 期 21 件	第 期 24 件 第 期 59 件	第 期 54 件 第 期 117 件

## (2) 特許等出願件数

第 期 0 件 (うち国内 0 件、国外 0 件)

第 期 0 件 (うち国内 0 件、国外 0 件)

合計 0 件 (うち国内 0 件、国外 0 件)

## (3) 受賞等

第 期 0 件 (うち国内 0 件、国外 0 件)

第 期 1 件 (うち国内 1 件、国外 0 件)

1. 金久 實:「大川出版賞」, 2001.11.29

## (4) 主な原著論文による発表の内訳

### 国外誌

- Ogata, H., Fujibuchi, W., Goto, S., and Kanehisa, M.: A heuristic graph comparison algorithm and its application to detect functionally related enzyme clusters. *Nucleic Acids Res.* 28, 4021-4028 (2000)
- Fujibuchi, W., Ogata, H., Matsuda, H., and Kanehisa, M.: Automatic detection of conserved gene clusters in multiple genomes by graph comparison and P-quasi grouping. *Nucleic Acids Res.* 28, 4029-4036 (2000)
- Kanehisa, M., Goto, S., Kawashima, S., and Nakaya, A.: The KEGG databases at GenomeNet. *Nucleic Acids Res.* 30, 42-46 (2002)
- Kanehisa, M. and Bork, P.: Bioinformatics in the post-sequence era. *Nat. Genet.* 33, 305-310 (2003)
- Fukuda, K. and Takagi, T.: Knowledge representation of signal transduction pathways. *Bioinformatics* 17, 829-837 (2001)
- De Hoon, M.J., Imoto, S., and Miyano, S.: Statistical analysis of a small set of time-ordered gene expression data using linear splines. *Bioinformatics* 18, 1477-1485 (2002)
- Akutsu, T., Miyano, S., and Kuhara, S.: Inferring qualitative relations in genetic networks and metabolic pathways. *Bioinformatics*, 16, 727-734 (2000)
- Aburatani, S., Tashiro, K., Savoie, C.J., Nishizawa, M., Hayashi, K., Ito, Y., Muta, S., Yamamoto, K., Enomoto, A., Masaki, M., Watanabe, S., Maki, Y., Takahashi, Y., Eguchi, Y., Sakaki, Y., and Kuhara, S.: Discovery of novel transcription control relationships with gene regulatory networks generated from multiple-disruption full genome expression libraries. *DNA Res.* 10, 1-8 (2003)

9. Ito, T., Tashiro, K., Muta, S., Ozawa, R., Chiba, T., Nishizawa, M., Yamamoto, K., Kuhara, S. and Sakaki, Y: Toward a protein-protein interaction map of the budding yeast: a comprehensive system to examine two-hybrid interactions in all possible combinations between the yeast proteins. Proc. Natl. Acad. Sci. USA 97, 1143-1147 (2000)
10. Ito, T., Chiba, T., Ozawa, R., Yoshida, M., Hattori, M., and Sakaki, Y.: A comprehensive two-hybrid analysis to explore the budding yeast protein interactome. Proc. Natl. Acad. Sci. USA 98, 4569-4574 (2001)
11. Oyama, T., Kitano, K., Satou, K., and Ito, T.: Extraction of knowledge on protein-protein interactions by association rule discovery. Bioinformatics 18, 705-714 (2002)
12. Yamauchi, E., Nakatsu, T., Matsubara1, M., Kato1, H., and Taniguchi, H.: Crystal structure of a MARCKS peptide containing the calmodulin-binding domain in complex with Ca<sup>2+</sup>-calmodulin. Nature Struc. Biol. 10, 226-231 (2003)

(5)主要雑誌への研究成果発表

Journal	Impact Factor	サブテーマ1	サブテーマ2	サブテーマ3	合計
P NATL ACAD SCI USA	54.48	1	4		5
NAT GENET	29.6	1	0		1
J BIOL CHEM	29.032	0	4		4
NUCLEIC ACIDS RES	25.492	4	0		4
EMBO J	24.918	0	2		2
PLANT CELL	22.162	2	0		2
MOL CELL	16.611	1	0		1
BIOINFORMATICS	13.684	4	0		4
MOL MICROBIOL	12.796	2	0		2
NAT STRUCT BIOL	11.707	0	1		1
J NEUROSCI	8.178	0	1		1
TRENDS BIOTECHNOL	10.12	0	2		2
J BACTERIOL	3.984	1	0		1
BIOCHEM BIOPH RES CO	2.946	1	0		1
YEAST	2.54	1	0		1
ARCH BIOCHEM BIOPHYS	2.476	0	1		1
BIOPOLYMERS	1.845	1	0		1
J COMPUT BIOL	1.728	1	0		1
SIGNAL PROCESSING	0.488	1	0		1
LECT NOTES COMPUT SC	0.415	1	0		1

impact factorは全てISI/JCRの2001年の値による