

領域略称名：ゲノム情報
領域番号：316

文部省科学研究費補助金重点領域研究

ゲノム解析に伴う大量知識情報処理の研究

研究成果報告書

(平成3～7年度)

平成8年12月

領域代表者 金久 實(京都大学化学研究所)

まえがき

文部省科学研究費補助金重点領域研究「ゲノム解析に伴う大量知識情報処理の研究」（重点領域研究「ゲノム情報」）は平成3年4月より平成8年3月まで5カ年計画で行われました。この報告書は計画研究班員および公募研究班員による主要な発表論文をまとめたものです。また、最初に平成8年10月に行われた文部省での研究終了ヒアリングの資料も添付しています。なお、重点領域研究「ゲノム情報」による出版物はインターネットを通じて公開されています。<http://www.genome.ad.jp/documents/gi/> をご覧ください。

大量データをもたらすゲノム計画の波及効果により、基礎生物学から医学・薬学・農学まで、バイオサイエンスでは急速な情報化が進んでいます。重点領域研究「ゲノム情報」の5年間の研究目標は、ゲノム解析に伴うデータベースとデータ解釈の問題に、新しい情報処理技術を開発し実用化することでした。しかし、私どもはより長期的な展望から領域活動を行ってきました。それは、情報化が進むバイオサイエンスの急展開に対処するための情報インフラストラクチャーの整備と、若手研究者層の拡大でした。そして、バイオサイエンスとコンピュータサイエンスの融合の中に、新しい学問分野としての「ゲノム情報」を確立することでした。ヒアリング資料にまとめてあるように、またゲノムネットデータベースサービス（<http://www.genome.ad.jp/>）をご利用になればお分かりいただけるように、当初の研究目標は十分に達成され、新しいバイオサイエンス研究の流れを作ることができたと思います。5年間のご支援をどうもありがとうございました。

平成8年12月

重点領域研究「ゲノム情報」代表者
京都大学化学研究所・教授
金 久 實

〒611 宇治市五ヶ庄
京都大学化学研究所
TEL (0774)38-3270
FAX (0774)32-8235
kanehisa@kuicr.kyoto-u.ac.jp
www.kuicr.kyoto-u.ac.jp/kanehisa

目 次

研究経過等の報告書

1 . 領域全体の当初の研究目標と目標達成度、学界への貢献度	3
2 . 領域内における研究組織と研究班の連携状況	4
3 . これまでの主な研究成果	6
4 . 領域として研究を推進したうえでの問題点と対応措置	9
5 . 今後の研究成果の取りまとめ方策	10
6 . 研究を取りまとめるうえでの問題点	11
7 . 研究成果公表の状況（主な論文等一覧）	12
付録 A . 活動記録	17
付録 B . ゲノムネットデータベースサービス利用統計	22

主要発表論文

Akutsu, T., Onizuka, K. and Ishikawa, M.

New hashing techniques and their application to a protein structure database system.

Proc. 28th Hawaii International Conference on System Sciences, 5, 197-206 (1995)25

Akutsu, T. and Tashimo, H.

Protein structure comparison using representation by line segment sequences.

Proc. Pacific Symposium on Biocomputing '96, 25-40 (1996)35

Mukouchi, Y. and Arikawa, S.

Towards a mathematical theory of machine discovery from facts.

Theoret. Comput. Sci., 137, 53-84 (1995)51

Sugaya, K., Fukagawa, T., Matsumoto, K., Mita, K., Takahashi, E., Ando, A.,
Inoko, H. and Ikemura, T.

Three genes in the human MHC class III region near the junction with the class II:
gene for receptor of advanced glycosylation end products, PBX2 homeobox gene
and a Notch homolog, human counterpart of mouse mammary tumor gene int-3.

Genomics, 23, 408-419 (1994)83

Kanaya, S., Kudo, Y., Nakamura, Y. and Ikemura, T.

Detection of genes in Escherichia coli sequences determined by genome projects and
prediction of protein production levels, based on multivariate diversity in codan usage.

Comput. Appl. Biosci., 12, 213-225 (1996)95

Ishikawa, M., Toya, T., Hoshida, M., Nitta, K., Ogiwara, A. and Kanehisa, M. Multiple sequence alignment by parallel simulated annealing. Comput. Appl. Biosci., 9, 267-273 (1993)	108
Hirosawa, M., Totoki, Y., Hoshida, M. and Ishikawa, M. Comprehensive study on iterative algorithms of multiple sequence alignment. Comput. Appl. Biosci., 11, 13-18 (1995)	115
Hakata, K. and Imai, H. The longest common subsequence problem for small alphabet size between many strings. Lecture Notes in Computer Science, 650, 469-478 (1992)	121
Ogasawara, N., Nakai, S. and Yoshikawa, H. Systematic sequencing of the 180 kilobase region of the <i>Bacillus subtilis</i> chromosome containing the replication origin. DNA Res., 1, 1-14 (1994)	131
Ogiwara, A., Uchiyama, I., Seto, Y. and Kanehisa, M. Construction of a dictionary of sequence motifs that characterize groups of related proteins. Protein Eng., 5, 479-488 (1992)	145
Mizuno, M. and Kanehisa, M. Distribution profiles of GC content around the translation initiation site in different species. FEBS Lett., 352, 7-10 (1994)	155
Ogata, H., Akiyama, Y. and Kanehisa, M. A genetic algorithm based molecular modeling technique for RNA stem-loop structures. Nucleic Acids Res., 23, 419-426 (1995)	159
Tomii, K. and Kanehisa, M. Analysis of amino acid indices and mutation matrices for sequence comparison and structure prediction of proteins. Protein Eng., 9, 27-36 (1996)	167
Ogata, H., Fujibuchi, W. and Kanehisa, M. The size differences among mammalian introns are due to the accumulation of small deletions. FEBS Lett., 390, 99-103 (1996)	177

Ogiwara, A., Uchiyama, I., Takagi, T. and Kanehisa, M. Construction and analysis of a profile library characterizing groups of structurally known proteins. Prot. Sci., 5, 1991-1999 (1996)	182
Satou, K., Furuichi, E., Takiguchi, K., Takagi, T. and Kuhara, S. A deductive database system PACADE for analyzing 3-D and secondary structures of protein. Comput. Appl. Biosci., 9, 259-265 (1993)	191
Satou, K., Furuichi, E., Takiguchi, K., Kuhara, S. and Takagi, T. Application of a deductive database system PACADE toward discovery of cluster of similar structures in proteins. Proc. 27th Hawaii International Conference on System Sciences, 5, 160-169 (1994)	198
Nishikawa, A., Satou, E., Furuichi, E., Kuhara, S. and Ushijima, K. Data classification component in a deductive database system and its application to protein structural analysis. IEICE Trans. Information and Systems E78-D, 1377-1386 (1995)	208
Gotoh, O. Optimal alignment between groups of sequences and its application to multiple sequence alignment. Comput. Appl. Biosci., 9, 361-370 (1993)	219
Gotoh, O. Further improvement in methods of group-to-group sequence alignment with generalized profile operations. Comput. Appl. Biosci., 10, 379-387 (1994)	229
Gotoh, O. A weighting system and algorithm for aligning many phylogenetically related sequences. Comput. Appl. Biosci., 11, 543-551 (1995)	238
Sakamoto, N., Takagi, T. and Sakaki, Y. Development of overlapping oligonucleotide database and its application to signal sequence search of the human genome. Comput. Appl. Biosci., 9, 427-434 (1993)	247

Goto, S., Sakamoto, N. and Takagi, T. Object-oriented database with rule-based query interface for genomic computation. Proc. 3rd International Symposium on Database Systems for Advanced Applications, 65-72 (1993)	255
Sakamoto, N., Goto, S. and Takagi, T. A deductive database system for analyzing human nucleotide sequence data. Int. J. Bio-Medical Computing, 171-179 (1994)	263
Murakami, K., Tsukuni, S., Takagi, T. and Hattori, M. Evaluation of exon prediction tools using a long DNA sequence data. Proc. Gene-Finding and Gene Structure Prediction Workshop (1995)	272
Takaoka, T. Approximate pattern matching with samples. Lecture Notes in Computer Science, 834, 234-242 (1994)	279
Kamimura, M. and Takahashi, Y. Phi-psi conformational pattern clustering of protein amino acid residues using the potential function method. Comput. Appl. Biosci., 10, 163-169 (1994)	288
Nakai, K. and Sakamoto, H. Construction of a novel database containing aberrant splicing mutations of mammalian genes. Gene, 141, 171-177 (1994)	295
Horton, P. and Nakai, K. A probabilistic classification system for predicting the cellular location sites of proteins. ISMB, 4, 109-115 (1996)	302
Nakata, K. Prediction of zinc finger DNA-binding protein. Comput. Appl. Biosci., 11, 125-131 (1994)	309
Suyama, M., Ogiwara, A., Nishioka, T. and Oda, J. Searching for amino acid sequence motifs among enzymes: the Enzyme-Reaction Database. Comput. Appl. Biosci., 9, 9-15 (1993)	316

Suyama, M., Nishioka, T. and Oda, J. Searching for common sequence patterns among distantly related proteins. Protein Eng., 8, 1075-1080 (1995)	323
Shimada, T., Hagiya, M., Arita, M., Nishizaki, S. and Tan, C.L. Knowledge-based simulation of regulatory action in lambda phage. Int. J. Artificial Intelligence Tools, 4, 511-523 (1995)	329
Husimi, Y., Mori, Y., Ushijima, S. and Aita, T. Modes of population changeovers in continuous culture of bacterial and viral mutants spectrum. J. Theoret. Biol., 165, 577-591 (1993)	342
Aita, T. and Husimi, Y. Period-dependent selection in continuous culture of viruses in periodic environment. J. Theoret. Biol., 168, 281-289 (1994)	357
Nemoto, N. and Husimi, Y. A model of the virus-type strategy in the early stage of encoded molecular evolutions. J. Theoret. Biol., 176, 67-77 (1995)	366
Aita, T. and Husimi, Y. Fitness spectrum among random mutants on Mt. Fuji-type fitness landscape. J. Theoret. Biol., 182, 469-485 (1996)	377
Matsuda, H. and Kaneda, Y. An application of an OR-parallel prolog system to phylogenetic analysis. Lisp and Symbolic Computation, 7, 211-229 (1994)	394
Matsuda, H. Protein phylogenetic inference using maximum likelihood with a genetic algorithm. Proc. Pacific Symposium on Biocomputing '96, 512-523 (1996)	413
Suwa, M., Mitaku, S. and Kuroda, Y. Theoretical analysis of amino acid sequence of human dystrophin. Biochem. Biophys. Res. Commun., 191, 782-789 (1993)	425

Suwa, M., Hirokawa, T. and Mitaku, S. A continuum theory for the prediction of lateral and rotational positioning of α -helices in membrane proteins: bacteriorhodopsin. Proteins, 22, 363-377 (1995)	433
Dohi, H., Ishizuka, M., Minoshima, S. and Shimizu, N. GeneView: multi-language human gene mapping library with a graphical user interface. Comput. Appl. Biosci., 9, 459-464 (1993)	448
Shimozono, S., Shinohara, A., Shinohara, T., Miyano, S., Kuhara, S. and Arikawa, S. Finding alphabet indexing for decision trees over regular patterns: an approach to bioinformatical knowledge acquisition. Proc. 26th Hawaii International Conference on System Sciences, 1, 763-772 (1993)	454
Shimozono, S., Shinohara, A., Shinohara, T., Miyano, S., Kuhara, S. and Arikawa, S. Knowledge acquisition from amino acid sequences by machine learning system BONSAI. Trans. Information Processing Society of Japan, 35, 2009-2018 (1994)	464
Shoudai, T., Lappe, M., Miyano, S., Shinohara, A., Okazaki, T., Arikawa, S., Uchida, T., Shimozono, S., Shinohara, T. and Kuhara, S. BONSAI Garden: parallel knowledge discovery system for amino acid sequences. ISMB, 3, 359-366 (1995)	474
Miyano, S. Learning theory toward genome informatics. IEICE Trans. Information and Systems E78-D, 560-567 (1995)	482
Tateishi, E. and Miyano, S. A greedy strategy for finding motifs from yes-no examples. Proc. Pacific Symposium on Biocomputing '96, 599-613 (1996)	490
Karasawa, T., Tabuchi, K., Fumoto, M. and Yasukawa, T. Development of simulation models for protein folding in thermal annealing process - I: a simulation of BPTI folding by pearl necklace model. Comput. Appl. Biosci., 9, 243-251 (1993)	505
Yokomori, T., Ishida, N. and Kobayashi, S. Learning local languages and its application to protein alpha-chain identification. Proc. 27th Hawaii International Conference on System Sciences, 113-122 (1994)	514

Nakaya, A., Yamamoto, K. and Yonezawa., A.
RNA secondary structure prediction using highly parallel computers.
Comput. Appl. Biosci., 11, 685-692 (1995)524

Nakaya, A., Taura, K., Yamamoto, K. and Yonezawa, A.
Visualization of RNA secondary structures using highly parallel computers.
Comput. Appl. Biosci., 12, 205-211 (1996)532

領域略称名：ゲノム情報
領域番号：316

科学研究費補助金重点領域研究
に係る研究経過等の報告書

「ゲノム解析に伴う大量知識情報処理の研究」

(平成3～7年度)

領域代表者 京都大学・化学研究所・教授・金久 實

連絡先 (0774)38-3270

1. 領域全体の当初の研究目標と目標達成度、学界への貢献度

分子生物学の実験技術の進歩に裏付けされて発足したヒト・ゲノム解析計画は、さらに新たな技術革新をもたらし、加速度的に大量のシーケンスデータを産み出している。それに伴い分子・細胞レベルでの生命現象に関する基礎データから、病気の診断・治療への可能性を示す応用データまで、様々な生物学的なデータが急速に蓄積されている。生物学はかつて直面したことのない大量情報の時代に入り、新しい情報処理技術の開発がゲノム解析の進展のため、また生物学の発展のために不可欠の要素となった。新しい情報処理技術とは大量データの中から生物学の知識を体系化し、知識を利用する技術であり、これは知識処理や並列処理といった計算機科学の最先端の研究と深く関連している。重点領域研究「ゲノム情報」は平成3年度から7年度までこのような背景の下に行われ、生物科学と計算機科学の融合の中から、生物情報科学ともいべき新しい学問分野の確立を目指してきた。

具体的な研究目標は、(A) ゲノム解析がもたらす大量かつ多様なデータのデータベースの問題と、(B) 実験によるゲノム解析の最終産物であるシーケンスデータの生物学的意味解釈の問題に新しい情報処理技術を開発し、その実用化を行ってゲノム研究を推進することであった。前者については目標は十分に達成された。数々の研究成果は論文等として発表しただけでなく、実用的なデータベースやソフトウェアとして、とくにゲノムネットデータベースサービスとして取りまとめ、インターネットを通じて国際的に提供している（付録B参照）。後者については基礎的な研究に重点を置いたため、実用化については未だ不十分である。ただし国際的なゲノム研究の流れとしても、第1期がマッピングとシーケンシングの技術開発段階、第2期がその結果としての大量シーケンシングデータの生産段階と位置づけられるので、データ解釈（機能予測）の実用化は第2期（平成8年度以降）での継続事項とした。

本領域研究は情報処理技術の研究開発と実用化以外に、わが国におけるバイオサイエンス研究に対して2つの重要な貢献を行った。1つは情報インフラストラクチャーの整備、もう1つは研究者層の拡大である。大量データの出現によるバイオサイエンスの急激な展開に対応するためには、情報インフラストラクチャーの整備が不可欠であるとの認識から、総括班の活動の一部として1991年にゲノム研究ネットワーク「ゲノムネット」を設立し、これはその後東京大学医科学研究所ヒトゲノム解析センターの事業として運用されている（付録A参照）。この5年間にわが国のインターネット環境は大幅に向上したが、ゲノムネットはゲノム関連データベースを国際的に提供するためのバックボーンとなっている。一方、上記の研究目標を達成するには従来からの生物科学の研究者層だけでなく、計算機科学を中心とした幅広い研究者層の参加が不可欠であった。これも総括班の活動として、毎年夏に「ゲノム情報チュートリアル」を、毎年冬に「ゲノム情報ワークショップ」を開催し、生物系研究者と計算機系研究者の交流と共同研究を推進した（付録A参照）。その結果、数百名のメンバーからなるゲノム情報コミュニティとも呼べる研究者層ができています。

既存の学会に対する活動としては、分子生物学会と生物物理学会でシンポジウムやコ

ンピュータデモンストレーションを何度も開催し、生物物理学会に「ゲノム情報」セッションを新設し、さらに情報処理学会と物理学会で学会誌に「ゲノム情報」の特集を行った（付録A参照）。それなりのインパクトはあったと考えられるが、複数の分野の境界に位置する新しい研究領域の宿命として、既存の学会で市民権を得るまでにはまだ時間がかかるだろう。今後は新しい学会の設立も視野に入れて、生物情報科学の学問分野の確立を目指すべきである。

2. 領域内における研究組織と研究班の連携状況

重点領域研究「ゲノム情報」は、総括班と全体で1つの計画研究班、それに公募研究を加えて組織した。以下に示す通り5年間を通じて総括班員は20名、研究班員（計画研究および公募研究）は49名であった。研究班員49名のうち20名以上は計算機科学専門の研究者であり、本領域研究が真に生物系と計算機系の共同研究プロジェクトであったことを裏付けている。これはもちろんわが国では初めてのことであり、世界的にも類のない先駆的なプロジェクトであった。

総括班

石浜 明	(平成3～7年度)	国立遺伝学研究所
磯野 克己	(平成3～7年度)	神戸大学理学部
伊藤 彬	(平成4～7年度)	(財)癌研究会癌研究所
牛島 和夫	(平成5～7年度)	九州大学工学部
内田 俊一	(平成3～6年度)	(財)新世代コンピュータ技術開発機構
大須賀節雄	(平成3～7年度)	東京大学先端科学技術研究センター
金久 實	(平成3～7年度)	京都大学化学研究所
榊 佳之	(平成3～7年度)	東京大学医科学研究所
佐々木慎一	(平成3～5年度)	豊橋技術科学大学
清水 信義	(平成3～7年度)	慶應義塾大学医学部
鈴木 義昭	(平成3～7年度)	基礎生物学研究所
関口 睦夫	(平成3～7年度)	九州大学生体防御医学研究所
高木 利久	(平成5～7年度)	東京大学医科学研究所
高浪 満	(平成3～7年度)	京都大学化学研究所
田中 穂積	(平成3～5年度)	東京工業大学工学部
西岡 孝明	(平成3～4年度)	京都大学化学研究所
松原 謙一	(平成3～7年度)	大阪大学細胞生体工学センター
美宅 成樹	(平成3～7年度)	東京農工大学工学部
吉川 寛	(平成3～7年度)	大阪大学医学部、奈良先端科学技術大学院大学
米澤 明憲	(平成3～7年度)	東京大学理学部

研究班（計画研究および公募研究）

秋山 泰	(平成4～7年度)	京都大学化学研究所
阿久津達也	(平成4～7年度)	機械技術研究所、群馬大学工学部
鯨坂 恒夫	(平成4年度)	京都大学工学部

有川 節夫	(平成6年度)	九州大学理学部
池村 淑道	(平成4～6年度)	国立遺伝学研究所
石川 幹人	(平成6年度)	(財)新世代コンピュータ技術開発機構
石田 好輝	(平成5年度)	京都大学工学部
磯山 正治	(平成7年度)	(財)蛋白質研究奨励会
伊藤 彬	(平成3年度)	(財)癌研究会癌研究所
今井 浩	(平成3～7年度)	九州大学理学部、東京大学理学部
小笠原直毅	(平成4～7年度)	大阪大学医学部、奈良先端科学技術大学院大学
金久 實	(平成3～7年度)	京都大学化学研究所
久原 哲	(平成3～7年度)	九州大学大学院農学研究科
小谷 善行	(平成3年度)	東京農工大学工学部
後藤 修	(平成3～7年度)	埼玉県立がんセンター研究所
小原 雄治	(平成6～7年度)	国立遺伝学研究所
小山 照夫	(平成3～5年度)	学術情報センター
坂本 博	(平成3年度)	京都大学理学部
品川日出夫	(平成3年度)	大阪大学微生物病研究所
篠原 武	(平成3～4年度)	九州工業大学情報工学部
陶山 明	(平成3～7年度)	長岡技術科学大学、東京大学教養学部
瀬戸 保彦	(平成3～6年度)	(財)蛋白質研究奨励会
添田 栄一	(平成3年度)	理化学研究所ジーンバンク
高岡 忠雄	(平成6年度)	茨城大学工学部
高木 利久	(平成3～7年度)	九州大学情報処理教育センター、東京大学医科学研究所
高橋 由雅	(平成3～7年度)	豊橋技術科学大学工学部
寺野 隆雄	(平成3～4年度)	筑波大学社会工学系
中井 謙太	(平成4～7年度)	基礎生物学研究所、大阪大学細胞生体工学センター
中田 琴子	(平成5年度)	国立衛生試験所
那須 正夫	(平成5年度)	大阪大学薬学部
新島 耕一	(平成5年度)	九州工業大学情報工学部
西岡 孝明	(平成3～7年度)	京都大学化学研究所
新田 克己	(平成3～5年度)	(財)新世代コンピュータ技術開発機構
野口 俊之	(平成3年度)	名古屋大学理学部
萩谷 昌己	(平成3～7年度)	京都大学数理解析研究所、東京大学理学部
伏見 譲	(平成3～7年度)	埼玉大学工学部
藤山秋佐夫	(平成3～7年度)	国立遺伝学研究所
古川 哲也	(平成5年度)	九州大学大型計算機センター
古谷 博史	(平成5年度)	京都教育大学教育学部
堀 寛	(平成3年度)	広島大学原爆放射能医学研究所
松田 秀雄	(平成5～7年度)	神戸大学工学部
美宅 成樹	(平成3～7年度)	東京農工大学工学部
蓑島 伸生	(平成3～7年度)	慶應義塾大学医学部
宮野 悟	(平成3～7年度)	九州大学理学部
森 浩禎	(平成4～7年度)	京都大学ウイルス研究所、奈良先端科学技術大学院大学
安川 民男	(平成3～5年度)	東京農工大学工学部
安永 照雄	(平成6～7年度)	大阪大学遺伝情報実験施設
横森 貴	(平成6～7年度)	電気通信大学
米澤 明憲	(平成3～7年度)	東京大学理学部

(1) 総括班

総括班は領域全体の運営と研究活動の円滑化を図ることを目的とし、年度初めと年度終りに毎年2回ずつ、全部で10回開催した総括班会議において、研究計画の取りまとめと調整、研究の進捗状況に対する助言、研究成果の公表と評価などを行った（付録A参照）。また対外的には総括班は、開かれたゲノム情報コミュニティ作りとゲノム研究の現場との密接な連携を目的とした活動を行った。具体的には生物系研究者と計算機系研究者の研究交流と研究者層拡大のためのワークショップとチュートリアル開催、コンピュータネットワークの整備を軸としたゲノム研究のインフラストラクチャー作り、東大医科研ヒトゲノム解析センターと京大化研スーパーコンピュータラボラトリーを中心とした全国的な研究支援体制作りである。

(2) 研究班

本領域は計画研究および公募研究あわせても班員は比較的少数であり、チュートリアルとワークショップ、ヒトゲノム解析センター、その他の場所で緊密に交流する機会が多く、その結果としての共同研究も盛んであった。そこで、領域全体の方針としては、班員相互の緊密な連携をコアとしたゲノム情報コミュニティ全体との連携を重視した。例えば、班員による研究成果発表会も毎年チュートリアル場で公開して行い、班員であるかどうかの区別なく研究交流を行ってきた。一方、ゲノム情報コミュニティと実験系研究者の交流にも力を入れ、本領域の班員はヒトゲノム解析センターのデータベース研究会を通じて、創成的基礎研究「ヒト・ゲノム解析研究」の班員を始めとしたゲノム研究現場の実験系研究者との連携を行い、実用的なデータベースシステム等の共同開発を行った。すでにその成果はゲノムネットを通じて世界中に提供され、またヒトゲノム解析センター主催の講習会、学会年会や国際シンポジウムなどでのデモンストレーションにより普及にも努めている。

(3) ゲノム情報コミュニティ

毎年夏に開催するゲノム情報チュートリアルと毎年冬に開催するゲノム情報ワークショップは、本領域の最も主要な公開行事であった（付録A参照）。この2つの定期行事は重点領域研究「ゲノムサイエンス」により継続して開催されることになり、ゲノム情報コミュニティの定期行事として定着している。チュートリアルが泊り込みで行うインフォーマルな勉強会であるのに対して、ワークショップはフォーマルな研究発表の場である。チュートリアルは米国のゴードン会議のような、ワークショップは学会の年会のような性格をもっているとも言えるだろう。また、ゲノム情報コミュニティの機関誌ともいえる「ゲノム情報ニュース」を年4回発行し、約1500名に配布してきた。

3. これまでの主な研究成果

本領域研究は5ヶ年計画を前期3年と後期2年に分け、前期は主に技術開発の段階、

後期は主に実用化の段階と位置づけ、(A) ゲノム解析がもたらす大量かつ多様なデータのデータベースの問題と、(B) 実験によるゲノム解析の最終産物である配列データの生物学的意味解釈の2つの問題に取り組んできた。ただし(B)については後期も技術開発が中心となったことは前述の通りである。

研究項目(A)における技術開発には3つの要素が必要であった。第1はオブジェクト指向データベースや演繹データベースといった新しいデータベース技術、第2はクライアント・サーバーを中心としたネットワーク技術、第3はグラフィカルなユーザーインターフェース技術である。一方、実用化に関してはゲノム研究の現場で行われるマッピングやシーケンシングに伴うデータベースの問題と、公表されたデータを蓄積した公共データベースの統合化の問題に区別して考えた。データベースの問題に関する主な研究成果は以下の通りである。

(A1) 大腸菌、枯草菌、酵母ゲノムのシーケンシングをサポートするために、オブジェクト指向的な考え方とグラフィカルユーザーインターフェース、それにクライアント・サーバー機能を備えたシステムを設計し、ヒトゲノム解析センターとの協力で Genomatica を開発した(秋山)。これは実際に枯草菌ゲノムデータベース構築(小笠原)や大腸菌ゲノムデータベース構築(森)に利用された。

(A2) ヒトゲノムに対してはグラフィカルユーザーインターフェースにより、マッピングデータ入力と統合マップ作成のデータベースシステム Locus-in(蓑島)と、コンティグ地図作成のためのソフトウェアツール ContigMaker(陶山、萩谷)をヒトゲノム解析センターとの協力で開発し、一般公開している。

(A3) ヒトに関する既存の公共データを統合して利用できるようにするため、ヒトの遺伝的地図(GDB)、塩基配列(GenBank)、アミノ酸配列(PIR)、遺伝病(OMIM)を、グラフィカルな環境で統合し、クライアント・サーバーで利用できるシステム HyperGenome を開発した(久原、高木、金久)。

(A4) 塩基配列(GenBank/DDBJ, EMBL)、アミノ酸配列(SWISS-PROT, PIR, PRF)、立体構造(PDB, PDBSTR)、配列モチーフ(EPD, TRANSFAC, PROSITE)、酵素反応(LIGAND)、代謝パスウェイ(PATHWAY)、変異タンパク質(PMD)、アミノ酸指標(AAindex)、遺伝病(OMIM)、文献(Medline, LITDB)といった既存の分子生物学データベースの関連を2項関係で表現した簡単な演繹データベース LinkDB を開発し、全体を統合して利用できる DBGET システムを公開した(金久)。

(A5) 上記のようなデータベースの統合だけでなく、さらに様々なデータベース検索・解析システムを緩やかに統合し、自由に組み合わせて利用できる環境 Gidre を構築し提供した(高木)。

(A6) 蛋白質研究奨励会の文献データベース LITDB と配列データベース PRF を一般公開するとともに、米国 NCBI との協力体制を作り、PRF は NCBI の Entrez システムに定期的に取り込まれるようになった（瀬戸）。

研究項目 (B) は配列データの生物学的意味を解釈する問題、すなわち機能予測の問題である。その基本方針は既知のデータで配列と機能の関連を明らかにし、その知識を利用することである。そのためには大量データから知識を獲得し利用する技術を開発しなければならない。また、実用化に関しては情報処理技術に頼るだけでなく、手作業により個々の専門分野で実際に知識ベースを構築し、生物学の広範な専門的知識をゲノム解析で利用できる形に体系化しなければならない。これはもちろん5年間で達成できるような内容ではない。そこで、本研究では知識獲得の技術をとくに学習アルゴリズムという観点から研究し、実際の知識ベース構築に基づく知識の利用についてはいくつかのプロトタイプを開発するにとどめた。データ解釈の問題に関する主な研究成果は以下の通りである。

(B1) 配列グループに共通なパターンを見いだすための複数配列アライメントに、シミュレーテッドアニーリング法（石川、金久）、ならびに反復改善法（後藤）という新しいアルゴリズムを開発した。これらは並列処理の観点からも興味深いアルゴリズムである。

(B2) タンパク質立体構造での共通パターンを抽出するために、部分構造検索およびアライメントを行うアルゴリズムを開発した（阿久津）。構造予測の問題に関しては、膜タンパク質の構造予測のためのアルゴリズムを開発し、実用的なシステムとして提供した（美宅）。また、RNA二次構造予測の並列アルゴリズムを開発し、並列計算機に実装した（米澤）。

(B3) 正の例と負の例の配列データグループが与えられたときに、両者を区別する配列の特徴を自動的に知識として獲得する機械学習アルゴリズムを開発し、実際のシステム BONSAI を開発した（宮野）。また、知識利用システムの1つとしてタンパク質の配列モチーフ検索・解析システム MOTIF およびそのクライアント・サーバー版 SMART を開発し提供した（金久）。

(B4) 手作業による知識の体系化もいくつかのテーマで行われているが、とくに酵素反応式、化合物の化学構造、酵素のアミノ酸配列と立体構造などを統合した、酵素反応データベース LIGAND を構築し提供した（西岡）。また、核酸の知識ベースとして、異常スプライシングデータを蓄積し、RNAスプライス部位選択に関するルールベースの構築を行った（中井）。

(B5) その他いくつかの問題で配列データの機能的、進化的意味の解釈が行われた（7

の論文一覧参照)。また、ゲノム情報の今後の発展につながる研究として、個々の配列の解析ではなく、発生などにおける遺伝子発現制御ネットワークのモデリング（萩谷）や、生物種内突然変異体ポピュレーションの解析（伏見）を行った。

4．領域として研究を推進したうえでの問題点と対応措置

重点領域研究「ゲノム解析に伴う大量知識情報処理の研究」が発足したときの最大の問題点は、わが国ではバイオサイエンスの一部として情報科学的な研究を行うインフラストラクチャーも人材も存在しないことであった。しかしながら、ゲノムプロジェクトの発足とともに必然的に起こるであろう大量データの出現と、それに伴うバイオサイエンスの急激な展開に対応するためには、新しい情報処理技術の研究開発と実用化を行わなければならないことは明白であり、本領域はまさにそれを期待されていた。研究を行う以前の問題として、インフラストラクチャー整備と若手研究者を中心とした人材発掘が急務だったのである。

情報インフラストラクチャーの整備に関しては、1990年秋より総合研究Bの中でコンピュータネットワークの設立を準備し、1991年9月に東京大学医科学研究所ヒトゲノム解析センターと京都大学化学研究所が開通して、ゲノムネット（GenomeNet）が発足した。ゲノムネットは当初は重点領域の総括班の活動として、現在はヒトゲノム解析センターの事業として運用されている。また1995年まではTISNの一部として京都から各地へ接続がなされていたが、TISNの解散に伴い現在は京都から先は福岡への接続のみとなっている。ゲノムネットはもともと回線接続を提供することを目的としたものではなく、国内のインターネット整備に国が積極的に支援するようになった現在、回線接続サービスを行う意味はないと考えられる。わが国のゲノム関連データベースを国内外に提供するためのバックボーンという本来の目的に戻ったわけである。

一方の人材発掘に関しては、通産省の第5世代コンピュータプロジェクト（ICOT）の遺伝子情報処理ワーキンググループが出発点であった。ここで、ゲノム情報の学問に対する興味が高まり、領域発足後は計算機科学者としてもエース級の若手研究者が続々と参加することとなったのである。その際の重要なポイントは、並列処理や知識処理に代表される新しい計算機科学の分野と、分子生物学の技術革新により大量情報の時代を迎えた生物科学の分野の間には、学問としての自然な共通性と発展性があり、基礎的な研究面での融合を強調したことである。ゲノム情報（genome informatics）という言葉は、とくに米国では、ゲノム計画を支援する実用的なデータベース作りや解析ツール作りを意味する。わが国の「ゲノム情報」はゲノムを出発点とした生命系の情報に関する学問領域であり、結果的に研究支援まで含めた活動を行ったのである。また、個々の研究者を本領域の都合だけで強く束縛せず自由な研究を推進し、これはゲノム情報コミュニティという形で、わが国全体の研究水準の向上につながったと考える。各省庁がビッグプロジェクトを開始した近年の状況を考えても、これからの重点領域の1つの形態として、閉じた研究プロジェクトではなく、開かれた研究領域の組織作りを目指すことがあって

もよいのではないだろうか。

ところで、若手の人材発掘を他の分野に頼ることは短期的には有効であったが、長期的にはバイオサイエンスの中で人材を育成していくことが必要である。ここに教育体制の問題が係わってくる。これは1つの重点領域で解決できる問題ではないが、いくつかの試みを行った。大学の学部レベルの教育については、分子生物学会年会の中で「生物系学科における情報科学教育」と題したパネルディスカッションを行い、問題提起と少なくとも問題に対する共通の認識を目指した。すなわち、データベースや解析システムを上手に使える人が増えれば問題が解決するのではなく、そもそもデータベースや解析システムを開発できる技術、さらにはその背景となる情報科学全体に対する理解をもった人材が育たなければ、バイオサイエンスの新しい発展はあり得ない点である。また、高校レベルでの教育については、領域発足直後に「ヒトゲノム計画とは - 新しい生物学の出発点 - 」と題したパンフレットを作成し、全国の高校に配布した。領域終了後にはとりまとめ研究の中で、首都圏の高校の先生を対象とし、「ヒトゲノム計画とインターネット」と題した講習会を開催した。そこで出された強い要望は、ゲノムネットWWWサーバーの日本語化であった。現時点では国際対応のために英語のみでサービスしているゲノムネットも、わが国の将来という観点からは日本語の重要性が非常に高いことを強く認識した次第である。

5 . 今後の研究成果の取りまとめ方策

重点領域研究「ゲノム情報」は終了したが、国内外のゲノム研究はますます活発化しつつある。幸い、平成8年度より文部省ヒトゲノムプログラムの第2期として重点領域研究「ゲノムサイエンス」が発足し、その1つの研究項目として「ゲノムの生物知識情報」に関する研究が始まった。また、平成3年度に設置された東京大学医科学研究所ヒトゲノム解析センターのゲノムデータベース分野と平成5年度に設置されたDNA情報解析分野の存在により、ゲノム情報に関する長期的な研究活動と研究支援活動が可能となっている。従って、重点領域研究「ゲノム情報」の研究成果取りまとめは、ゲノムを出発点とした生命系の情報に関する研究の1つの区切りとして位置づけ、これまでの研究成果を基盤とし、今後の長期的な展望を与えることに重点を置きたい。

実はすでに「ゲノムサイエンス」の「ゲノムの生物知識情報」はそのような観点から研究計画が立案されている。重点領域研究「ゲノム情報」は、(1) 情報インフラクチャーの整備と、(2) 研究者層の拡大をベースに、(3) 新しい情報処理技術の開発を達成した。既存のデータベースを統合する問題では、技術開発とシステム開発を行い、ゲノムネットを通じて提供するまで実用化されたが、配列データの生物学的意味解釈の問題は、技術開発までであった。これは解釈の前提となる生物機能に関するデータあるいは知識がほとんどコンピュータ化されていないためであり、「ゲノムサイエンス」では「技術開発から知識の体系化へ」をキーワードに、(4) 実用的な知識ベースの構築を目標としている。しかし、そこでの知識ベースは「ゲノム情報」で考えていたような配列

と機能の関連を示す辞書だけではない。それは配列を通して個々の分子の働きが分かっていても、細胞あるいは生物個体の働きが分かるわけではないからである。個々の分子とはいわば生命系を構成する部品であり、部品間のつながり、つまり結線図のようなものが解明されなければ、生命系の働きを理解することはできない。実は遺伝子や分子の相互作用に関するデータは分子・細胞生物学の知識として広く存在しており、これをコンピュータ化して、ゲノム解析がもたらす遺伝子カタログとの対応をつけていくことが、「ゲノムサイエンス」における生物知識情報の体系化なのである。

このように重点領域研究「ゲノム情報」の研究成果の取りまとめの結果として、ゲノム研究の新しい展開が始まっている。なお、本領域研究の成果は各年度ごとの報告書と取りまとめ研究による5年間の報告書として出版し、発表論文リストとゲノム情報ワークショップの論文集はゲノムネットのWWWサーバーでも公開している。

6. 研究を取りまとめるうえでの問題点

情報インフラストラクチャー作りと技術開発・システム開発は若手の計算機科学者の参入により急速に進んだ。しかし今後の生物知識情報の体系化には生物学の知識とセンスを備えた若手研究者が参入しなければ達成できないだろう。とくに幅広い生物学の専門領域をカバーするには多数の生物学者の協力が必要である。重点領域研究「ゲノム情報」発足時には、まず指導的立場にある計算機科学者の理解を求め、それを軸に若手計算機科学者の発掘を行った。同様の方策は今後の若手生物学者の発掘についても必要である。重点領域研究「ゲノム情報」の取りまとめに対し、指導的立場にある生物学者の理解を求め、新しいゲノム情報研究の展開が可能になることを期待したい。

研究の取りまとめは学界に対してだけでなく、一般社会に対しても行う必要がある。それは本重点領域としてのゲノム情報だけでなく、広くゲノム計画がもたらす情報の観点から行うべきであろう。そして科学の最先端を一方的に伝えるのではなく、将来の科学者を育成する教育的見地が必要であろう。インターネットは一般社会と直結しているので、例えば現状では専門家のみを対象とした英語のゲノムネットWWWサーバーに日本語のページを加え、初歩から段階的に学んでいけるようにすることも考えられる。もっとも技術的には可能でも、我々研究者がこのような教育活動を行う意義が学界で認められなければ、実践することは難しいかもしれない。しかし、本領域は5年間にいくつの発表論文をだせるかといった短期的な研究成果をだけを目指したのではない。本領域が力を入れた情報インフラストラクチャー作りやゲノム情報コミュニティ作りは、長期的な研究推進のための基盤整備だったのである。長期的な展望に立てば教育は避けて通れない重要な問題である。その意味から、重点領域研究「ゲノム情報」の研究成果が一般社会の中に、とくに次の世代を担う子供たちの教育の中に取り込まれるよう努力をしていきたいと考えている。

7 . 研究成果公表の状況 (主な論文等一覧)

1) 主な図書

- Takagi, T., Imai, H., Miyano, S., Mitaku, S., and Kanehisa, M. (eds.); "Proceedings of Genome Informatics Workshop IV", Universal Academy Press, Tokyo, 1993.
- Miyano, S., Akutsu, T., Imai, H., Gotoh, O., and Takagi, T. (eds.); "Proceedings of Genome Informatics Workshop 1994", Universal Academy Press, Tokyo, 1994.
- Hagiya, M., Miyano, S., Nakai, K., Suyama, A., Yokomori, T., Takagi, T. (eds.); "Proceedings of Genome Informatics Workshop 1995", Universal Academy Press, Tokyo, 1995.
- Akutsu, T., Asai, K., Hagiya, M., Kuhara, S., and Miyano, S. (eds.); "Proceedings of the Seventh Workshop on Genome Informatics", Universal Academy Press, Tokyo, 1996.
- 美宅成樹、金久 實 (編) 「ヒトゲノム計画と知識情報処理」培風館、東京 (1995).
- 高木利久、金久 實 (編) 「ゲノムネットのデータベース利用法」共立出版、東京 (1996).
- 金久 實 「ゲノム情報への招待」共立出版、東京 (1996).

2) 主な発表論文

- Akutsu, T., Kobayashi, S., Hori, K., and Ohsuga, S.; Algorithms for finding the largest subtree whose copies cover all the leaves. IEICE Trans. Information and Systems E76-D, 707-710 (1993).
- Akutsu, T.; A linear time pattern matching algorithm between a string and a tree. Proc. 4th Symposium on Combinatorial Pattern Matching, pp. 1-10 (1993).
- Akutsu, T.; Efficient and robust three-dimensional pattern matching algorithms using hashing and dynamic programming techniques. Proc. 27th Hawaii International Conference on System Sciences, pp. 225-234 (1994).
- Akutsu, T.; A linear time pattern matching algorithm between a string and a tree. IEICE Trans. Information and Systems E77-D, 281-287 (1994).
- Akutsu, T.; Approximate string matching with don't care characters. Lecture Notes in Computer Science 807, 240-249 (1994).
- Akutsu, T., Onizuka, K. and Ishikawa, M.; New hashing techniques and their application to a protein structure database system. Proc. 28th Hawaii International Conference on System Science 5, 197-206 (1995).
- Akutsu, T.; Approximate string matching with don't care characters. Information Processing Letters 55, 235-239 (1995).
- Akutsu, T. and Tashimo, H.; Protein structure comparison using representation by line segment sequences. Proc. Pacific Symposium on Biocomputing '96, pp. 25-40 (1996).
- Mukouchi, Y. and Arikawa, S.; Towards a mathematical theory of machine discovery from facts. Thoret. Comput. Sci. 137, 53-84 (1995).
- Ikemura, T.; Codon usage. In "Plant Molecular Biology LABFAX" (Brown, T.A., ed.), pp. 38-48, Blackwell Sci. Pub. (1993).
- Sugaya, K., Fukagawa, T., Matsumoto, K., Mita, K., Takahashi, E., Ando, A., Inoko, H., and Ikemura, T.; Three genes in the human MHC class III region near the junction with the class II: gene for receptor of advanced glycosylation end products, PBX2 homeobox gene and a Notch homolog, human counterpart of mouse mammary tumor gene int-3. Genomics 23, 408-419 (1994).
- Fukagawa, T., Sugaya, K., Matsumoto, K., Okumura, K., Ando, A., Inoko, H., and Ikemura, T.; A boundary of long-range G+C% mosaic domains in the human MHC locus: pseudoautosomal boundary-like sequence exists near the boundary. Genomics 25, 184-191 (1995).

- Fukagawa, T., Nakamura, Y., Okumura, K., Ando, K., Inoko, H., Saitou, N., and Ikemura, T.; Human pseudoautosomal boundary-like sequences (PABLs): core and consensus sequence, expression, and involvement in formation of the present day pseudoautosomal boundary of short arm of human sex chromosomes. *Hum. Mol. Genet.* 5, 23-32 (1996).
- Ishikawa, M., Toya, T., Hoshida, M., Nitta, K., Ogiwara, A., and Kanehisa, M.; Multiple sequence alignment by parallel simulated annealing. *Comput. Appl. Biosci.* 9, 267-273 (1993).
- Hirosawa, M., Totoki, Y., Hoshida, M., and Ishikawa, M.; Comprehensive study on iterative algorithms of multiple sequence alignment. *Comput. Appl. Biosci.* 11, 13-18 (1995).
- Hakata, K. and Imai, H.; The longest common subsequence problem for small alphabet size between many strings. *Lecture Notes in Computer Science* 650, 469-478 (1992).
- Ogasawara, N., Nakai, S., and Yoshikawa, H.; Systematic sequencing of the 180 kilobase region of the *Bacillus subtilis* chromosome containing the replication origin. *DNA Res.* 1, 1-14 (1994).
- Yoshida, K., Sano, H., Miwa, Y., Ogasawara, N., and Fujita, Y.; Cloning an nucleotide sequence of a 15kb region of the *Bacillus subtilis* genome containing the *iol* operon. *Microbiology* 140, 2289-2298 (1994).
- Ogasawara, N., Fujita, Y., Kobayashi, Y., Sadaie, Y., Tanaka, T., Takahashi, H., Yamane, K., and Yoshikawa, H.; Systematic sequencing of the *Bacillus subtilis* genome: progress report of the Japanese group. *Microbiology* 141, 257-259 (1995).
- Akagawa, E., Kurita, K., Sugawara, T., Nakamura, K., Kasahara, Y., Ogasawara, N., and Yamane, K.; Determination of 17, 484 bp nucleotide sequence around the 390 region of the *Bacillus subtilis* chromosome and similarity analysis of the products of putative ORFs. *Microbiology* 141, 3241-3245 (1995).
- Ogiwara, A., Uchiyama, I., Seto, Y., and Kanehisa, M.; Construction of a dictionary of sequence motifs that characterize groups of related proteins. *Protein Eng.* 5, 479-488 (1992).
- Fujibuchi, W. and Kanehisa, M.; A method to extract functional motifs for transcriptional regulation in eukaryotic sequences. *Bull. Inst. Chem. Res., Kyoto Univ.* 71, 317-326 (1993).
- Ohkubo, Z. and Kanehisa, M.; Characterization of spatially close peptide segments in proteins. *Bull. Inst. Chem. Res., Kyoto Univ.* 71, 327-337 (1993).
- Uchiyama, I., Ogiwara, A., and Kanehisa, M.; A library of signature pentapeptides for the Protein Data Bank. In "Protein Structures by Distance Analysis" (Bohr, H. and Brunak, S., eds.), 237-246, IOS Press (1994).
- Mizuno, M. and Kanehisa, M.; Distribution profiles of GC content around the translation initiation site in different species. *FEBS Lett.* 352, 7-10 (1994).
- Ogata, H., Akiyama, Y., and Kanehisa, M.; A genetic algorithm based molecular modeling technique for RNA stem-loop structures. *Nucleic Acids Res.* 23, 419-426 (1995).
- Tomii, K. and Kanehisa, M.; Analysis of amino acid indices and mutation matrices for sequence comparison and structure prediction of proteins. *Protein Eng.* 9, 27-36 (1996).
- Ogata, H., Fujibuchi, W., and Kanehisa, M.; The size differences among mammalian introns are due to the accumulation of small deletions. *FEBS Lett.* 390, 99-103 (1996).
- Ogiwara, A., Uchiyama, I., Takagi, T., and Kanehisa, M.; Construction and analysis of a profile library characterizing groups of structurally known proteins. *Prot. Sci.* 5, 1991-1999 (1996).
- Kuhara, S., Kanehisa, M., and Yoshida, Y.; A gene mapping database HyperGenome System. *Proc. International Symposium on Advanced Computing for Life-Science*, pp. 250-252 (1992).
- Satou, K., Furuichi, E., Takiguchi, K., Takagi, T., and Kuhara, S.; A deductive database system PACADE for analyzing 3-D and secondary structures of protein. *Comput. Appl. Biosci.* 9, 259-265 (1993).
- Satou, K., Furuichi, E., Takagi, T., Kuhara, S., and Ushijima, K.; Similar structure search in deductive database. *Proc. International Symposium on Next Generation Database Systems and Their Applications*, pp. 130-137 (1993).
- Furuichi, E., Kamikawaji, N., Sasazuki, T., and Kuhara, S.; Estimation of the mechanism of peptide binding to HLA class III molecules. In "Modern Trends in Biothermokinetics" (Schuster, S. et al., eds.), Plenum, pp.

- 377-381 (1993).
- Satou, K., Furuichi, E., Takiguchi, K., Kuhara, S., and Takagi, T.; Application of a deductive database system PACADE toward discovery of cluster of similar structures in proteins. Proc. 27th Hawaii International Conference on System Sciences, Vol. 5, pp. 160-169 (1994).
- Nishikawa, A., Satou, E., Furuichi, E., Kuhara, S., and Takagi, T.; Equipments for statistical analysis in a deductive database system. Proc. International Symposium on Advanced Database Technologies and Their Integration, pp. 237-244 (1994).
- Nishikawa, A., Satou, E., Furuichi, E., Kuhara, S., and Ushijima, K.; Data classification component in a deductive database system and its application to protein structural analysis. IEICE Trans. Information and Systems E78-D, 1377-1386 (1995).
- Gotoh, O.; Optimal alignment between groups of sequences and its application to multiple sequence alignment. Comput. Appl. Biosci. 9, 361-370 (1993).
- Gotoh, O.; Further improvement in methods of group-to-group sequence alignment with generalized profile operations. Comput. Appl. Biosci. 10, 379-387 (1994).
- Gotoh, O.; A weighting system and algorithm for aligning many phylogenetically related sequences. Comput. Appl. Biosci. 11, 543-551 (1995).
- Niwa-Kawakita, M., Miyoshi, H., Gotoh, O., Matsushima, Y., Nishimura, M., Shisa, H., and Ohki, M.; Cloning and gene mapping of the mouse homologue of the CBFA2T1 gene associated with human acute myeloid leukemia. Genomics 29, 755-759 (1995).
- Barnes, T.M., Kohara, Y., Coulson, A., and Hekimi, S.; Meiotic recombination, noncoding DNA and genomic organization in *Caenorhabditis elegans*. Genetics 141, 159-179 (1995).
- Takagi, T.; ODS: Overlapping Oligonucleotide Database with deductive engine for signal sequence search. Proc. 2nd International Conference on Bioinformatics, Supercomputing & Complex Genome Analysis, pp. 263-272 (1992).
- Sakamoto, N., Takagi, T., and Sakaki, Y.; Development of Overlapping Oligonucleotide Database and its application to signal sequence search of the human genome. Comput. Appl. Biosci. 9, 427-434 (1993).
- Goto, S., Sakamoto, N., and Takagi, T.; Object-oriented database with rule-based query interface for genomic computation. Proc. 3rd International Symposium on Database Systems for Advanced Applications, pp. 65-72 (1993).
- Goto, S., Sakamoto, N., Takagi, T., and Ushijima, K.; A deductive language in object-oriented database for genome analysis. Proc. International Symposium on Next Generation Database Systems and Their Applications, pp. 123-129 (1993).
- Suzuki, T., Takagi, T., Kuhara, S., and Kanehisa, M.; Development of an integrated database for genome mapping and nucleotide sequences. Proc. 27th Hawaii International Conference on System Sciences, Vol. 5, pp. 68-76 (1994).
- Sakamoto, N., Goto, S., and Takagi, T.; A deductive database system for analyzing human nucleotide sequence data. Int. J. Bio-Medical Computing, 171-179 (1994).
- Murakami, K., Tsukuni, S., Takagi, T., and Hattori, M.; Evaluation of exon prediction tools using a long DNA sequence data. Proc. Gene-Finding and Gene Structure Prediction Workshop (1995).
- Takaoka, T.; Approximate pattern matching with samples. Lecture Notes in Computer Science 834, 234-242 (1994).
- Kamimura, M., Takahashi, Y., and Sasaki, S.; Phi-psi conformational pattern clustering of protein amino acid residues using the potential function method. J. Protein. Chem. 11, 393-394 (1992).
- Kamimura, M. and Takahashi, Y.; Phi-psi conformational pattern clustering of protein amino acid residues using the potential function method. Comput. Appl. Biosci. 10, 163-169 (1994).
- Nakai, K. and Sakamoto, H.; Construction of a novel database containing aberrant splicing mutations of mammalian genes. Gene 141, 171-177 (1994).

- Nakai, K., Tokimori, T., Ogiwara, A., Uchiyama, U., and Niiyama, T.; Gnome: an Internet-based sequence analysis tool. *Comput. Appl. Biosci.* 10, 547-550 (1994).
- Horton, P. and Nakai, K.; A probabilistic classification system for predicting the cellular location sites of proteins. *ISMB 4*, 109-115 (1996).
- Nakata, K.; Prediction of zinc finger DNA-binding protein. *Comput. Appl. Biosci.* 11, 125-131 (1994).
- Suyama, M., Ogiwara, A., Nishioka, T., and Oda, J.; Searching for amino acid sequence motifs among enzymes: the Enzyme-Reaction Database. *Comput. Appl. Biosci.* 9, 9-15 (1993).
- Suyama, M., Nishioka, T., and Oda, J.; Searching for common sequence patterns among distantly related proteins. *Protein Eng.* 8, 1075-1080 (1995).
- Hagiya, M.; A Typed lambda-Calculus for proving-by-example and bottom-up generalization procedure, algorithmic learning theory. *Lecture Notes in Artificial Intelligence 744*, 73-86, (1993).
- Hagiya, M. and Toda, Y.; On implicit arguments, logic, language and computation - Festschrift in honor of Satoru Takasu. *Lecture Notes in Computer Science 792*, 10-30 (1994).
- Shimada, T., Hagiya, M., Arita, M., Nishizaki, S., and Tan, C.L.; Knowledge-based simulation of regulatory action in lambda phage. *Int. J. Artificial Intelligence Tools 4*, 511-523 (1995).
- Yoshikawa, H., Nagai, H., Matsubara, K., and Fujiyama, A.; Two-dimensional gel electrophoretograms of human chromosome specific restriction DNA fragments. *Biochem. Biophys. Res. Com.* 196, 1566-1572 (1993).
- Husimi, Y., Murata, G., Yoshino, K., et al; Continuous culture of virus as a model of evolving systems and its application to evolutionary molecular engineering. In "The Origin and Evolution of the Cell" (Hartman, H. and Matsuno, K., eds.), pp.375-386, World Sci. (1992).
- Husimi, Y., Mori, Y., Ushijima, S., and Aita, T.; Modes of population changeovers in continuous culture of bacterial and viral mutants spectrum. *J. Theoret. Biol.* 165, 577-591 (1993).
- Aita, T. and Husimi, Y.; Period-dependent selection in continuous culture of viruses in periodic environment. *J. Theoret. Biol.* 168, 281-289 (1994).
- Nemoto, N. and Husimi, Y.; A model of the virus-type strategy in the early stage of encoded molecular evolutions. *J. Theoret. Biol.* 176, 67-77 (1995).
- Aita, T. and Husimi, Y.; Fitness spectrum among random mutants on Mt.Fuji-type fitness landscape. *J. Theoret. Biol.* 182, 469-485 (1996).
- Kaneda, Y. and Matsuda, H.; A priority control system for OR-parallel Prolog and its performance evaluation. *Lecture Notes in Computer Science 748*, 42-53 (1993).
- Matsuda, H. and Kaneda, Y.; An application of an OR-parallel prolog system to phylogenetic analysis. *Lisp and Symbolic Computation 7*, 211-229 (1994).
- Matsuda, H.; Querying genomic database by using a parallel logic programming system on distributed computing environment. *Proc. IEEE Pacific Rim Conference on Communications, Computers, and Signal Processing*, pp. 333-336 (1995).
- Matsuda, H.; Protein phylogenetic inference using maximum likelihood with a genetic algorithm. *Proc. Pacific Symposium on Biocomputing '96*, pp. 512-523 (1996).
- Mitaku, S., Suwa, M., and Saito, N.; A computer system for predicting membrane protein structure. *J. Protein Chem.* 11, 390-391 (1992).
- Suwa, M., Mitaku, S., Shimazaki, K., and Chuman, T.; Characterization of Transmembrane Helices by a Probe Helix of Molecular Energy Calculation. *Jpn. J. Appl. Phys.* 31, 951-956 (1992).
- Suwa, M., Mitaku, S., and Kuroda, Y.; Theoretical analysis of amino acid sequence of human dystrophin. *Biochem. Biophys. Res. Commun.* 191, 782-789 (1993).
- Suwa, M., Hirokawa, T., and Mitaku, S.; A continuum theory for the prediction of lateral and rotational positioning of a-helices in membrane proteins: bacteriorhodopsin. *Proteins* 22, 363-377 (1995).
- Dohi, H., Ishizuka, M., Minoshima, S., and Shimizu, N.; GeneView: multi-language human gene mapping library with a graphical user interface. *Comput. Appl. Biosci.* 9, 459-464 (1993).

- Arikawa, S., Kuhara, S., Miyano, S., Shinohara, A., and Shinohara, T.; A learning algorithm for elementary formal systems and its experiments on identification of transmembrane domains. Proc. 25th Hawaii International Conference on System Sciences, Vol. 1, pp. 675-684 (1992).
- Arikawa, S., Kuhara, S., Miyano, S., Mukouchi, Y., Shinohara, A., and Shinohara, T.; A machine discovery from amino acid sequences by decision trees over regular patterns. New Generation Computing 11, 361-375 (1993).
- Shimozono, S., Shinohara, A., Shinohara, T., Miyano, S., Kuhara, S., and Arikawa, S.; Finding alphabet indexing for decision trees over regular patterns: An approach to bioinformatical knowledge acquisition. Proc. 26th Hawaii International Conference on System Sciences, Vol. 1, pp. 763-772 (1993).
- Miyano, S.; Learning theory toward genome informatics. Proc. 4th International Workshop on Algorithmic Learning Theory, pp. 19-36 (1993).
- Shimozono, S., Shinohara, A., Shinohara, T., Miyano, S., Kuhara, S., and Arikawa, S.; Knowledge acquisition from amino acid sequences by machine learning system BONSAI. Trans. Information Processing Society of Japan 35, 2009-2018 (1994).
- Shoudai, T., Lappe, M., Miyano, S., Shinohara, A., Okazaki, T., Arikawa, S., Uchida, T., Shimozono, S., Shinohara, T., and Kuhara, S.; BONSAI Garden: parallel knowledge discovery system for amino acid sequences. ISMB 3, 359-366 (1995).
- Shimozono, S. and Miyano, S.; Complexity of finding alphabet indexing. IEICE Trans. Information and Systems E78-D, 13-18 (1995).
- Miyano, S.; Learning theory toward genome informatics. IEICE Trans. Information and Systems E78-D, 560-567 (1995).
- Tateishi, E. and Miyano, S.; A greedy strategy for finding motifs from yes-no examples. Proc. Pacific Symposium on Biocomputing '96, pp. 599-613 (1996).
- Karasawa, T., Tabuchi, K., Fumoto, M., and Yasukawa, T.; Development of simulation models for protein folding in thermal annealing process - I: a simulation of BPTI folding by pearl necklace model. Comput. Appl. Biosci. 9, 243-251 (1993).
- Izumisawa, J., Higuti, M., Yoshizawa, T., and Yasukawa, T.; Prediction of protein structure by Hopfield algorithm II: analysis of inter-residue interaction and application to small proteins. Repts. Progr. Polym. Phys. Japan 36, 621-624 (1993).
- Fumoto, M. and Yasukawa, T.; Effects of short range force on the folding processes of proteins. Repts. Progr. Polym. Phys. Japan 37, 719-712 (1994).
- Yoshizawa, T., Fumoto, M., and Yasukawa, T.; Prediction of protein conformations by a spin glass model (II). Repts. Progr. Polym. Phys. Japan 38, 593-596 (1995).
- Yokomori, T., Ishida, N., and Kobayashi, S.; Learning local languages and its application to protein alpha-chain identification. Proc. 27th Hawaii International Conference on System Sciences, pp. 113-122 (1994).
- Nakaya, A., Yamamoto, K., and Yonezawa, A.; RNA secondary structure prediction using highly parallel computers. Comput. Appl. Biosci. 11, 685-692 (1995).
- Nakaya, A., Taura, K., Yamamoto, K. and Yonezawa, A.; Visualization of RNA secondary structures using highly parallel computers. Comput. Appl. Biosci. 12, 205-211 (1996).

付録A . 活動記録

重点領域研究「ゲノム情報」公式行事

90.12. 3-4	第1回ワークショップ(東京、機械振興会館)	参加者	210名
91. 1. 7	総合B班会議(東京ガーデンパレス)		
91. 6.17	第1回総括班会議(東京ガーデンパレス)		
91. 7.11-12	第1回チュートリアル(京都平安会館)	参加者	130名
91.12. 9-10	第2回ワークショップ(東京、星陵会館)	参加者	300名
92. 1. 6-7	研究発表会(東京ガーデンパレス)		
92. 1. 7	第2回総括班会議(東京ガーデンパレス)		
92. 6.11	第3回総括班会議(東京ガーデンパレス)		
92. 7. 1-4	第2回チュートリアル(賢島、大阪府セミナーハウス)	参加者	120名
92.12.14-15	第3回ワークショップ(横浜パシフィコ)	参加者	370名
93. 2.26	第4回総括班会議(アルカディア市ヶ谷)		
93. 6.24	第5回総括班会議(アルカディア市ヶ谷)		
93. 7. 6-9	第3回チュートリアル(福岡、ホテル海の中道)	参加者	120名
93.12.13-15	第4回ワークショップ(横浜パシフィコ)	参加者	280名
94. 2.15	第6回総括班会議(東京ガーデンパレス)		
94. 6.23	第7回総括班会議(ヒトゲノム解析センター)		
94. 7. 6-9	第4回チュートリアル(富山、インテック研修センター)	参加者	110名
94.12.19-20	第5回ワークショップ(横浜パシフィコ)	参加者	294名
95. 3. 9-10	国際シンポジウム「ゲノム計画とコンピュータサイエンス」	参加者	121名
95. 3. 3	第8回総括班会議(アルカディア市ヶ谷)		
95. 6.22	第9回総括班会議(東京ガーデンパレス)		
95. 7. 5-8	第5回チュートリアル(神戸ゼロックス総合教育研究所)	参加者	111名
95.12.11-12	第6回ワークショップ(横浜パシフィコ)	参加者	275名
96. 3.12	第10回総括班会議(国立教育会館)		
96. 8.22-23	ヒトゲノム計画とインターネット講習会	参加者	31名

学会シンポジウム等

92.11. 6	生物物理学会(大阪)シンポジウム「ヒトゲノム計画で得られる大量ゲノム情報をどうするか？」		
92.12. 9	分子生物学会(京都)ミニシンポジウム「ゲノムデータベース構築の現状」		
93. 5	物理学会誌小特集「生命の設計図に迫る！ - ヒトゲノム計画と知識情報処理 - 」		
93.10.12	生物物理学会(名古屋)ゲノム情報セッション		
93.11.15-17	国際ヒトゲノム会議(神戸)オンサイト・インフォマティックス・セッション		
94. 9.29	生物物理学会(東京)ゲノム情報セッション		
94.12.13	分子生物学会(神戸)シンポジウム「分子生物学データベースの新しい流れ」		
95. 1. 3-6	HICCS Biotechnology Computing Track (Hawaii)		
95. 3. 9-10	国際シンポジウム「ゲノム計画とコンピュータサイエンス」		
95. 9.24	生物物理学会(札幌)ゲノム情報セッション		
95.12. 6	分子生物学会(名古屋)シンポジウム「情報生物学の展開」		
96.10	情報処理学会誌特集「ゲノム情報」		

ヒトゲノム解析センター主催の講習会等

93. 2. 1	稼働開始	
93. 2.22	ヒトゲノム解析センター計算機システム利用説明会	約 1 7 0 名
93. 3.17	CM-5、nCUBE2 講習会	1 3 名
93. 4.22-23	データベース利用講習会	約 1 0 0 名
93. 6.10-11	医科研内利用者講習会	約 4 0 名
93. 7.23	旧 VAX 利用者講習会	約 4 0 名
93. 8.17-19	CM-5 トレーニングコース	1 5 名
93. 9.27	QUANTA 講習会	1 5 名
93.10.21	QUANTA 講習会 (第 2 回)	6 名
93.12.21	nCUBE2 講習会	6 名
94. 1.19	新プロ合同班会議 (千葉) での講習会	約 4 0 名
94. 2.23-24	ヒトゲノム解析センター利用法入門講習会	1 2 8 名
94. 3. 5	九州大学での講習会	7 5 名
94. 5.16-17	ヒトゲノム解析センター計算機システム利用者講習会	3 0 名
94. 7. 4	QUANTA講習会	1 1 名
94. 7. 7-8	チュートリアルでの計算機講習会	約 5 0 名
94. 7. 26-27	GCG講習会	5 1 名
94.10.19-20	CM-5講習会	9 名
94.12.5	データベース検索と配列解析の入門	1 6 名
94.12.13-16	分子生物学会年会での講習会	約 1 3 0 名
95.2.2	ネットワーク管理者入門コース	1 8 名
95.4.21	利用法入門講習会	4 8 名
95.5.26	データベース利用法講習会	3 8 名
95.6.23	開発ソフト利用法講習会	1 9 名
95.7.6	チュートリアルでの計算機講習会	5 0 名
95.7.21	G C G 利用法講習会	2 5 名
95.8.23	nCUBE講習会	6 名
95.8.24-25	CM5利用法講習会	7 名
95.8.28	ネットワーク管理者入門コース	3 5 名
95.9.29	データベース利用法講習会	2 0 名
95.10.25	QUANTA利用法講習会 (基礎編)	5 名
95.11.17	QUANTA利用法講習会 (応用編)	5 名
96.2.17	九州大学での利用者入門講習会	4 5 名
96.4.26	利用者講習会	6 0 名
96.5.24	データベース利用法講習会	5 5 名
96.6.21	開発ソフト利用法講習会	2 2 名
96.7.26	G C G 利用法講習会	2 4 名
96.8.22	並列計算機 (CRAY) 利用法講習会	4 名
96.9.6	ネットワーク管理者入門コース	2 6 名
96.9.27	QUANTA講習会 (1)	4 名
96.10.25	データベース利用法講習会	1 9 名
96.11.15	QUANTA講習会 (2)	2 名

スーパーコンピューターラボラトリー主催の講習会等

92. 1. 4	稼働開始	
92. 1.14	スーパーコンピューター利用説明会（宇治地区）	約70名
92. 1.28-29	QUANTA 講習会	15名
92. 2.19	スーパーコンピューターラボラトリー竣工披露式	
92. 2.27	スーパーコンピューター利用説明会（京大、ゲノム）	約50名
92. 2.27-28	QUANTA 講習会	17名
92. 6. 2	CHARMm 講習会	48名
92.11.18	UniChem 講習会	16名
92.11.24	BIOSYM 講習会	8名
92.12. 1- 2	FORTTRAN 講習会	25名
93. 4.19	スーパーコンピューターラボラトリー利用説明会	約50名
93. 4.20	データベース利用講習会	32名
93. 5.18-20	UNICOS 講習会	18名
93. 6.15-17	C 講習会	8名
93. 7.13	QUANTA/CHARMm 講習会	14名
93. 7.28	Insight/Discover 講習会	9名
93.10.19	UniChem 2.0 講習会	4名
93.10.21	FORTTRAN/C 最適化講習会	8名
93.11.24	CHARMm 応用編講習会	11名
94. 1.28	ドキュメントツール講習会	7名
94. 4.19	スーパーコンピューターラボラトリー利用説明会	56名
94. 5.27	AVS5 講習会	4名
94. 6.24	Fortran 最適化講習会	7名
94. 7.14-15	UNIX 講習会	16名
94. 7.29	UniChem 講習会	4名
94. 9.21	ドキュメントツール講習会	10名
94. 9.27	QUANTA/CHARMm 講習会	14名
94.10.21	FORTTRAN90 講習会	5名
94.11.18	UniChem 2.3 講習会	1名
95. 2.17	ドキュメントツール講習会	10名
95. 4.18	スーパーコンピューターラボラトリー利用説明会（宇治地区）	31名
95. 4.19	ゲノムネットデータベースサービス利用説明会（宇治地区）	21名
95. 4.25	スーパーコンピューターラボラトリー利用説明会（吉田地区）	29名
95. 4.25	ゲノムネットデータベースサービス利用説明会（吉田地区）	78名
95. 5.19	ネットワーク入門講習会	16名
95. 5.19	UNIX 入門講習会	15名
95. 5.25	QUANTA/CHARMm 講習会	5名
95. 6. 2	UNIX 応用講習会	7名
95. 7.28	UniChem 講習会	5名
95. 8.11	C 言語入門	5名
95. 8.24	FORTTRAN 最適化	5名
95. 8.25	Fortran90 講習会	4名
95. 9. 1	ドキュメントツール講習会	2名
95. 9.14	ネットワーク入門	4名
95.10.20	WWW 講習会	13名

95.11.15	ゲノムネットデータベース利用説明会（病院地区）	5 1 名
95.11.17	CRAY ツール講習会	3 名
95.12.13	QUANTA NMR オプション説明会	3 名
96. 3.14	ドキュメントツール講習会	2 名
96. 4.19	スーパーコンピュータラボラトリー利用説明会（宇治地区）	8 名
96. 5.16	UNIX 入門講習会	9 名
96. 5.17	ネットワーク入門	1 2 名
96. 5.22	ゲノムネットデータベース利用説明会（吉田地区）	3 名
96. 6.21	UniChem 講習会	1 1 名
96. 6.28	QUANTA/CHARMm 講習会	8 名
96. 7.16	ゲノムネットデータベース利用説明会（宇治地区）	3 6 名
96. 7.25	Fortran90 講習会	2 名
96. 9.26	Fortran 最適化講習会	2 名
96.10. 2	ドキュメントツール講習会	5 名

ネットワーク接続（GenomeNet / TISN）

91. 9.20	東大理 - 京大化研、京大化研 - 阪大細胞工学が 64KBPS で開通
91.11. 7	東大理 - 東大医科研が 64KBPS で開通
92. 8.20	東大理 - 癌研が 64KBPS で開通
92. 9.30	東大理 - 京大化研が 192KBPS にアップグレード 京大化研 - WIDE京都NOC - 九大工が 192KBPS で開通
92.10. 2	東大理 - がんセンターが 64KBPS で開通
93.10.27	京大化研 - 蛋白奨励会が ISDN で開通
92.12.17	東大理 - 農水省筑波が 64KBPS で開通
93. 1.11	東大理 - 岡崎が 64KBPS で開通
93. 3.15	東大理 - 東大医科研が 1.5MBPS にアップグレード
93. 4.12	阪大細胞生体工学 - 阪大遺伝情報実験施設が LAN接続で開通
93. 5.26	東大理 - 米国西海岸が 512KBPS にアップグレード
93. 6. 2	東大医科研 - 国立衛生試験所が ISDN で開通
93. 7.29	京大化研 - 高輝度光科学研究センターが 64KBPS で開通
93. 8. 1	京大化研 - 蛋白工学研究所が 64KBPS で開通
93. 9. 1	東大理 - 京大化研が 512KBPS にアップグレード
93. 9. 1	京大化研 - けいはんなが 64KBPS で開通
93. 9. 1	HSP研究所 - WIDE京都NOCが LAN接続で開通
93. 9.11	京大化研 - 国際高等研究所が 64KBPS で開通
93.12.20	東大医科研 - 埼玉県立がんセンターが ISDN で開通
94. 1. 6	京大化研 - 藤田保健衛生大学が ISDN で開通
94. 1.14	東大理 - 米国西海岸が 768KBPS にアップグレード（WIDE との共同運用）
94. 5. 1	京大化研 - WIDE京都NOC - 九大工が 512KBPS にアップグレード
94. 6.10	東大医科研 - 京大化研間に 150MBPS の NTT 実験線が開通
95. 1.31	京大化研 - 高輝度光科学研究センターが 512KBPS にアップグレード
96. 4. 1	京大化研 - WIDE京都NOC - 九大工が 768KBPS にアップグレード
96. 5.17	海外線を TISN から IMnet へ移行
96. 5.29	東大医科研 - 京大化研間の 150MBPS NTT 実験線が STM から ATM へ移行

ゲノムネットデータベースサービス

サービス名	内容	開発者	利用形態	開始時期	備考
IRx	データベース検索	NLM	On-line	92. 9	
Entrez	統合データベース検索	NCBI	On-line	92. 9	95.9 廃止
DBGET	データベース検索	SCL	E-mail	92. 9	
			Gopher	94. 2	
			WWW	94. 7	
			Client	94. 8	
BLAST	ホモロジー検索	NCBI	E-mail	92. 9	
			WWW	94.12	
FASTA	ホモロジー検索	W.Pearson	E-mail	92. 9	
MOTIF	モチーフ検索	SCL	E-mail	93. 2	
			WWW	94.12	
CLUSTALW	マルチプルアライメント	SCL	WWW	96. 9	
Locus-in	マッピング支援	HGC/Keio	FTP	94. 4	
ContigMaker	マッピング支援	HGC/MKI	FTP	94. 4	
Genomatica	シーケンシング支援	HGC/SCL	FTP	94. 4	
HyperGenome	統合データベース検索	HGC/Kyushu	FTP	94. 4	
Gnome	配列解析支援	HGC/Senri	FTP	94. 4	
SMART	モチーフ解析	HGC/SCL	Client	94.10	
PSORT	局在部位予測	K.Nakai	WWW	94.12	
Gidre	データベース統合化環境	HGC	Client	95. 6	
98Entrez	Entrez の PC98 版	NCBI/HGC	FTP	95. 7	
ACEDB	ゲノムデータベース検索	R.Durbin	On-line	95.10	
KEGG	遺伝子・ゲノム百科辞典	SCL	WWW	95.12	

新規開発システム一覧

NetDBget: 分子生物学データベース統合検索システム (荻原 淳、内山郁夫、金久 實)

WebDBget: ハイパーリンクによる分子生物学データベース統合検索システム (秋山 泰、五斗 進、内山郁夫、金久 實)

Gnome: ネットワークを利用した配列解析支援システム (新山利幸、時森健夫、荻原 淳、内山郁夫、中井 謙太)

SMART: 配列モチーフ検索・解析ツール (荻原 淳、内山郁夫、高木利久、金久 實)

Gidre: データベース統合化環境構築システム (北野景彦、荻原 淳、高木利久)

HyperGenome: 染色体地図と配列情報を統合したデータベース (久原 哲、高木利久、五斗 進、金久 實)

Genomatica: 染色体上の物理位置に基づきDNA配列を整理したデータベース (秋山 泰、森 浩禎、小笠原直毅、久原 哲、村上康文、宮嶋伸行)

Locus-in: マッピングデータ入力と統合マップ作成のためのデータベースシステム (菟島伸生、清水信義)

ContigMaker: コンティグ地図作製のためのソフトウエアツール (陶山 明、萩谷昌己、伊藤隆司、藤山秋佐夫、大山 彰、高木利久)

SAND: Nested Deletion 法による配列結合編集システム (陶山 明、大山 彰、古畑芳晶、萩谷昌己、伊藤隆司、藤山秋佐夫、服部正平、榊 佳之、高木利久)

付録B．ゲノムネットデータベースサービス利用統計

