



HOKKAIDO UNIVERSITY

Title	光合成生物のオミックスデータベース利用法
Author(s)	成川, 礼; 岡本, 忍
Description	6章 バイオインフォマティクス 4
Relation	光合成研究法. 北海道大学低温科学研究所, 日本光合成研究会共編
Citation	低温科学, 67, 673-678
Issue Date	2009-03-31
Doc URL	https://hdl.handle.net/2115/39210
Type	departmental bulletin paper
File Information	67-094.pdf



4. 光合成生物のオミックスデータベース利用法

成川 礼¹⁾, 岡本 忍²⁾

近年のゲノム解析の進展と、それにともなうゲノムレベルの様々な網羅的解析手法の開発により、膨大な量の計測データがデータベースに蓄積され続けている。生物学者は研究を進める上で、これらの情報を有効に活用する必要性に迫られている。本章では、光合成生物に関連するオミックスデータベースの内容と所在について紹介する。

Utilization of omics databases

Rei Narikawa, Shinobu Okamoto

Over the last decade, genomic and post-genomic data are accumulating at an ever-accelerating rate. Experimental biologists are on the verge of effectively utilizing these data to design experimental procedures. In this chapter, we introduce location and context of omics databases of phototrophic organisms.

4.1 ゲノム配列データベース

ゲノムの塩基配列データは、実際に配列決定を行った研究機関が国際塩基配列データベース (International Nucleotide Sequence Databases) に登録することにより、世界中に公開され利用可能になる。国際塩基配列データベースとは、日本の DDBJ (DNA Data Bank of Japan, 表 1-1), 米国 NCBI (National Center for Biotechnology Information, 表 1-2), および欧州 EBI (European Bioinformatics Institute, 表 1-3) からなる組織であり、世界中の塩基配列データ登録の窓口になっている。登録された塩基配列データは、国際規約に従いアクセッション番号 (Accession number) が付与される。ひとつの機関に登録された配列データは、公開と同時に同一のアクセッション番号で DDBJ, GenBank, EMBL の三者間で共有される。配列決定が完了した生物種に関しては、DDBJ の Gene Trek in Prokaryote Space (GTPS, 表 1-4) や GenBank の RefSeq¹⁾ (表 1-5) のように独自の判断基準と自動処理を用いて、均一な遺伝子構造抽出を行ったゲノム情報なども公開されている。また、Joint Genome Institute (JGI, 表 1-6) や J. Craig Venter Institute (JCVI, 表 1-7) など、配列機関やプロジェクトによっては、ゲノム配列決定完了前のドラフト状態での公開もおこなわれている。

配列決定のコストが下がり、膨大な数のゲノムプロジェクトが精力的に行われた結果、研究者は自分に関係のある生物種のゲノムプロジェクトの把握すら困難な状況にある。このような状況を受け、プロジェクト単位、生物種単位で配列決定プロジェクトの進捗状況やプロジェクトの推進機関などメタデータの重要性が高まっている。Genome Online Database²⁾ (GOLD, 表 1-8) は、プロジェクト毎の進捗状況に応じて、解読が完全に終了した生物のゲノムプロジェクト、真核生物、古細菌、真正細菌の現在進行形のゲノムプロジェクト、メタゲノムプロジェクトという視点で分類、整理され閲覧可能である。このデータベースは、JGI の微生物ゲノムプロジェクトのチームリーダーである Nikos Kyrpides を中心としてボランティアベースで提供されている。また、NCBI Genome Project (表 1-9) は、NCBI の豊富なゲノム情報リソースをゲノムプロジェクト単位で整理し直すことで、ゲノム配列情報、生物種情報やプロジェクトに関連する主要機関などの情報を参照可能である。

4.2 個別生物種の統合データベース

網羅的解析時代以前は、個別の研究プロジェクトデータベースが、それぞれの関係者により開発、維持されてきた。網羅解析時代になると、そのようなデータベースが乱立し、利用者が必要な情報にたどり着くことが困難になった。また、技術や資金の問題から、プロジェクト毎でのデータベースの維持管理も難しくなっている。このような背景から、近年、データベースを統合して利便

1) 東京大学大学院総合文化研究科広域科学専攻

2) かずさ DNA 研究所

表1：ゲノム配列・ゲノムプロジェクトデータベース

No	データベース名	URL
1	DDBJ	http://www.ddbj.nig.ac.jp/index-e.html
2	NCBI	http://www.ncbi.nlm.nih.gov/
3	EBI	http://www.ebi.ac.uk/
4	GTPS	http://gtps.ddbj.nig.ac.jp/
5	RefSeq	http://www.ncbi.nlm.nih.gov/RefSeq/
6	JGI	http://www.jgi.doe.gov/
7	JCVI	http://www.jcvi.org/
8	GOLD	http://www.genomesonline.org/
9	NCBI Genome Project	http://www.ncbi.nlm.nih.gov/sites/entrez?db=genomeprj
10	ELIXIR	http://www.elixir-europe.org/

性を上げようという試みが加速している。各研究機関や研究者コミュニティは、分散している大量なゲノム情報を収集整理して検索機能、比較機能などをつけることにより利用者の便宜をはかり、情報の付加価値を高めようとしている。このような試みは、米国 NCBI や英国 ELIXER (表 1-10) など、ライフサイエンスの包括的なデータベース統合だけでなく、光合成のモデル生物毎にも精力的に行われている。ここでは光合成に関連する生物の統合データベースの一部を紹介する。

4.3 シアノバクテリア、 光合成細菌データベース

シアノバクテリア *Synechocystis* sp. PCC 6803 は、酸素発生型光合成生物として 1996 年に世界で初めてゲノム配列が決定された³⁾。CyanoBase⁴⁾ (表 2-1) は、当初 *Synechocystis* のゲノムデータベースとして開発され、その後、現在 30 種以上のシアノバクテリアと光合成細菌のゲノム情報が収集公開されている。

CyanoBase では、遺伝子構造、遺伝子産物情報、ゲノムマップなどの他に、プロテオーム解析の結果⁵⁾、Yeast two-hybrid 法によるタンパク質相互作用実験情報⁶⁾ や、外部データベースへの豊富なリンクも提供している。ま

た、1998 年には変異株の情報を共有したいという利用者の要望にこたえて CyanoBase 内に変異株データベース CyanoMutants が、さらに、その機能を拡張するかたちで研究者自身の遺伝子注釈を自然言語で記入することが出来る CyanoGenes が提供された。2002 年には京都大学のバイオインフォマティクスセンターからコミュニティ型アノテーションの先駆けとなる CYORF (表 2-2) が提供され、専門家によるジャンボリー形式の遺伝子注釈が試みられ、文献証拠に基づいた遺伝子注釈が提供されている。また、アメリカの Elhai らによって、CyanoBIKE (表 2-4) と呼ばれるコミュニティデータベースも近年開発され、Web 上から様々なコマンドを使用したバイオインフォマティクス解析が可能となっている。2007 年には、Web ページに注釈付けが行えるシステム、KazusaAnnotation (表 2-3) を用い、S.PCC 6803 の約 1200 報の原著論文全文中に記述された遺伝子・タンパク質名を抽出する試みがなされている。それ以外にも、東京大学の園池らは網羅的に遺伝子破壊株のクロロフィル蛍光を測定し、そのプロファイルによる機能分類を行ったデータベース、Fluorome⁷⁾ (表 2-5) を公開している。

4.4 シロイヌナズナデータベース

世界における主なシロイヌナズナデータベースを表 3

表2：シアノバクテリア、光合成細菌データベース

No	データベース名	URL
1	CyanoBase	http://bacteria.kazusa.or.jp/cyanobase
4	CYORF	http://cyano.genome.jp/
3	KazusaAnnotation	http://a.kazusa.or.jp/
4	CyanoBIKE	http://cyanobike-community.csbc.vcu.edu/welcome-frame.php
5	Fluorome	http://sunlight.k.u-tokyo.ac.jp/fluorome/about.php

表3：シロイヌナズナ関連データベース

No	データベース名	URL
1	TAIR	http://www.arabidopsis.org
2	AGNS	http://www.mgs.bionet.nsc.ru/agns
3	AGRIS	http://arabidopsis.med.ohio-state.edu
4	AMPdb	http://www.plantenergy.uwa.edu.au/applications/ampdb/index.html
5	MPSS	http://mpss.udel.edu/at/
6	AtNoPDB	http://bioinfo.scri.sari.ac.uk/cgi-bin/atnopdb/proteome_comparison
7	ARAMEMNON	http://aramemnon.botanik.uni-koeln.de
8	ARTDEdb	http://omicspace.riken.jp/ARTADE
9	ASRP	http://cgrb.orst.edu/smallRNA/db
10	AtGDB	http://www.plantgdb.org/AtGDB
11	AthaMap	http://www.athamap.de
12	ATTED-II	http://www.atted.bio.titech.ac.jp
13	CATMA	http://www.catma.org
14	DATF	http://datf.cbi.pku.edu.cn
15	GABI-Ka	http://www.GABI-Kat.de
16	MAtdB	http://mips.gsf.de/proj/thal/db
17	NASCarrays	http://affymetrix.arabidopsis.info
18	PlantQTL-GE	http://www.scbt.org/qtl2gene/new
19	Plprot	http://www.pb.ipw.biol.ethz.ch/proteomics
20	RARGE	http://rarge.gsc.riken.jp
21	SeedGenes	http://www.seedgenes.org
22	SUBA	http://www.suba.bcs.uwa.edu.au
23	WAtDB	http://www.watdb.nl
24	GeneFarm	http://urgi.versailles.inra.fr/Genefarm
25	TIGR Arabidopsis annotation db	http://www.tigr.org/tdb/e2k1/ath1/ath1.shtml

にまとめた。なかでも圧倒的に整備されているのが、The Arabidopsis Information Resource⁹⁾ (TAIR, 表3-1)である。TAIRは、シロイヌナズナ (*Arabidopsis thaliana*) の包括的なデータベースであり、30以上の研究機関による遺伝子構造、遺伝子産物情報、代謝、遺伝子発現の他に、ゲノムマップ、遺伝子マーカー、文献情報とシロイヌナズナ研究コミュニティに関する情報などを統合して提供している。2008年時点で、月間3万人以上の利用者がある。また、TAIRはThe Gene Ontology Consortium⁹⁾でも中心的な役割を果たしており、GO termの整備やGOに基づいた遺伝子注釈が精力的に行われている。遺伝子構造予測と遺伝子注釈は、年に1~2回の頻度で更新されており、最新のTAIR8(2008年4月リリース版)では、27,235の推定タンパク質を含む38,963の遺伝子モデルを提供している。さらにTAIRは、そのWebページから他のシロイヌナズナ資源まで広範囲なリンクを提供し、ポータルサイトの役目も果たしている。

The TIGR Arabidopsis thaliana Genome Annotation Database¹⁰⁾ (表3-25)に含まれる配列は、TIGRを含む全てのArabidopsis Genome Initiative (AGI)に所属する研究機関に由来するものである。このデータベースはこれらのデータの標準化のため、解析とアノテーションを行ったものである。本データベースは、TAIR、MIPS (表6-1)、GO、INTERPRO (表6-2)、PRINTS (表6-3)、PRODOM (表6-4)、PROSITE (表6-5)、Pfam (表6-6)などの情報が付加されている。また検索ツールが3種類、解析ツールが2種類、利用できる。

MAtdB¹¹⁾ (MIPS Arabidopsis thaliana Database, 表3-16)は、AGIの一部としてEuropean Scientists Sequencing Arabidopsis (ESSA)計画のゲノム配列データのリポジトリとして開始された。現在は、分散しているデータの統合と各種解析・可視化ツールの提供により、生物学的知識の統合リソースとなっている。また、各プロジェクトによる全長cDNAのID統合データベースを持ち、AGI Locusコードを軸にTIGR、TAIRと相互

参照が可能になっている。他の植物ゲノムとの比較ゲノム解析ツールも充実しており、シンテニーブラウザー SynBrowse¹²⁾ により、ミヤコグサ、タルウマゴヤシ、トウモロコシ、イネなどとゲノムレベルで比較することができる。

ATTED-II¹³⁾ (表 3-12) は、公開されているシロイヌナズナのマイクロアレイデータ、58 実験、1,388 スライドと予測 cis-element から推測された遺伝子の共発現解析に基づいた相互制御関係のデータを提供している。また、

遺伝子発現プロファイルを可視化するツール、LabAT-TED や、検索ツールなども提供している。データベースの更新状況や、データ数などの統計情報など、解析を支援する周辺情報も充実している。

4.5 イネゲノム解析プロジェクトとイネデータベース

世界的に主要なイネ関連データベースを表 4 にまとめた。IRGSP (The International Rice Genome Sequence

表 4：イネ科植物データベース

No	データベース名	URL
1	RAP-DB	http://rapdb.lab.nig.ac.jp/
2	TIGR Rice Genome Annotation	http://www.tigr.org/tdb/e2k1/osa1/
3	INE	http://rgp.dna.affrc.go.jp/giot/INE.html
4	Oryzabase	http://www.shigen.nig.ac.jp/rice/oryzabase/top/top.jsp
5	Gramene	http://www.gramene.org/Oryza_sativa/index.html
6	Rice MPSS	http://mpss.udel.edu/rice/
7	RiceGE (Rice Functional Genomics Database)	http://signal.salk.edu/cgi-bin/RiceGE
8	RAD (Rice Annotation Database)	http://rad.dna.affrc.go.jp/index_j.html
9	KOME (Knowledge-based Oryza Molecular biological Encyclopedia)	http://cdna01.dna.affrc.go.jp/cDNA/
10	RiceGAAS	http://ricegaas.dna.affrc.go.jp/
11	MOsDB (MIPS Oryza sativa Database)	http://mips.gsf.de/proj/plant/jsf/rice/index.jsp
12	MGOS (Magnaporthe Grisea Oryza Sativa Database)	http://www.mgosdb.org/
13	OsGDB	http://www.plantgdb.org/OsGDB/index.php
14	RetrOryza	http://www.retroryza.org/
15	RED (The Rice Expression Database)	http://red.dna.affrc.go.jp/RED/
16	Rice Proteome Database	http://gene64.dna.affrc.go.jp/RPD/main_en.html
17	DRTF (The Database of Rice Transcription Factor)	http://drtf.cbi.pku.edu.cn/
18	TIGR Oryza Repeat Database	http://www.tigr.org/tdb/e2k1/osa1/blastsearch.shtml
19	GTOP (Genomes TO Protein structures and functions)	http://structure.rice.dna.affrc.go.jp/gtop/gtop.html
20	RKD (Rice Kinase Database)	http://rkd.ucdavis.edu/index.htm
21	The Genome Cluster Database (GCD)	http://bioweb.ucr.edu/databaseWeb/index.jsp
22	CSRDB	http://sundarlab.ucdavis.edu/smrnas/
23	RiceTFDB (Rice Transcription Factor DataBase)	http://ricetfdb.bio.uni-potsdam.de/v2.1/
24	BarleyBase	http://www.plexdb.org/plex.php?database=Rice
25	ASIP (Alternative Splicing In plant)	http://www.plantgdb.org/ASIP/
26	IRGSP (The International Rice Genome Sequencing Project)	http://rgp.dna.affrc.go.jp/J/IRGSP/index.html
27	RMG (Rice Mitochondrial Genome Information)	http://rmg.rice.dna.affrc.go.jp/
28	RPSD (Rice Protein Structure Database)	http://structure.rice.dna.affrc.go.jp/
29	ricemarker	http://shenghuan.shnu.edu.cn/genefunction/ricemarker.htm
30	Rice Pipeline	http://cdna01.dna.affrc.go.jp/PIPE/index_ja.html
31	PLACE	http://www.dna.affrc.go.jp/PLACE/
32	Rice SNPs Database	http://www.pgcdna.co.jp/snps/index.html
33	Rice Research Database (RRDB)	http://www.pgcdna.co.jp/igs_system/index.html
34	Wild Rice Database (WLD)	http://www.pgcdna.co.jp/cgi-bin/wrdb/content.cgi
35	Oryza Tag Line	http://urgi.versailles.inra.fr/OryzaTagLine/
36	IRIS - International Rice Information System	http://www.iris.irri.org
37	POGs/PlantRBP	http://plantrbp.uoregon.edu/

Project, 表 4-26) は, 日本を含む 10 カ国からなる国際イネゲノム塩基配列解析プロジェクトで決定したイネ (*Oryza sativa* L. ssp. Japonica) ゲノムデータを公開している。イネの 12 本の染色体, 合計約 4 億塩基対に対して, 階層型ショットガン方式により, PAC, BAC クローンをシーケンスし, DNA マーカーに基づき再配置が行われている。現在の最新版は Build 5.0 (2008 年 12 月リリース版) である。

RAP-DB¹⁴⁾ (Rice Annotation Project Database, 表 4-1) は, 独立行政法人農業生物資源研究所 (NIAS), 独立行政法人産業技術総合研究所 (CBRC), 大学共同利用機関法人 情報・システム研究機構 国立遺伝学研究所 (NIG) がイネゲノム全塩基配列上に存在する 29,550 の予測遺伝子の位置を決定し, Gene Ontology, EST 配列, タンパク質機能ドメイン (InterPro), TIGR Annotation Pseudomolecules などの情報を付与したアノテーションデータベースである。特に専門家による遺伝子のマニュアルキュレーションに注力しており, ジャンボリー形式の遺伝子アノテーションが過去 4 回行われ, 成果が RAP4 (2007 年 10 月) として反映されている。

TIGR Rice Genome Annotation Resource (表 4-2) は, The Institute for Genomic Research (TIGR, 現 JCVI) が, 予測遺伝子 42,653 個の位置を決定し, 機能アノテーションの付与, 発現データ, 完全長 cDNA データ, MPSS データ, SAGE データ, Genbank EST をマッピングしたプローブ, Agilent 社や Affymetrix 社の Rice Genome Array プローブ情報などを統合したアノテーションデータベースである。このプロジェクトでは, コンピュータ解析と人手によるキュレーション作業により, 高精度のゲノム注釈を行うことを目的としている。リリース 4 の段階で, 42,653 の遺伝子 (トランスポゾン関連遺伝子を除く) を同定した。

Gramene¹⁵⁾ (表 4-5) データベースは, イネ科植物に関

する包括的なゲノム情報を収集したデータベースである。ヒトやマウスゲノムプロジェクトデータベースに携わってきた Cold Spring Harbor 研究所の Lincoln Stein が開発している。イネ科植物の基本的なゲノム情報以外にも, 文献情報によるタンパク質の機能構造注釈, 代謝マップおよび EST 配列や QTL などの情報が検索可能である。また, 比較ゲノムツール CMap を用いて, イネ, ムギ, トウモロコシ, アワ, シロイヌナズナなどの種間でシンテニー領域を比較することができる。ゲノムブラウザーには, 多くの真核生物ゲノムプロジェクトで標準的な Ensembl Genome Browser を使用しており, EST 配列, 発現データ, SNPs 情報などの膨大なデータがゲノム上にマッピングされている。また Web ベースのデータマネジメントシステムである BioMart¹⁶⁾ (表 6-7) を用いて, 共通プロモーター領域や共通の機能ドメインをもつ任意のゲノム配列といったデータを一括取得することが可能となっている。

4.6 その他のオミックスデータベース

EST データベース: かずさ DNA 研究所では, クラミドモナス, ミヤコグサ, シロイヌナズナ, スサビノリ, トマトの EST データ (表 5-1) が公開されている。様々な真核生物の EST を集めている TBestDB (表 5-2) というデータベースも存在する。真核藻類のデータ取得等に便利である。

トランスクリプトームデータベース: 広範囲の生物のトランスクリプトームデータのポータルサイトとしては, ArrayExpress (表 5-3) や GEO DataSets (表 5-4) などが存在する。シロイヌナズナの発現データはかなり蓄積されている。前者のデータベースにはシアノバクテリアのデータは少ないが, 後者のデータベースには 24 つのデータセットが蓄積されている。GEO DataSets には近年プロテオームデータも蓄積している。シアノバクテリアに関しては, KEGG EXPRESSION (表 5-5) にもデータが蓄積されている。

表 5: その他のオミックスデータベース

No	データベース名	URL
1	Kazusa Plant cDNA project	http://www.kazusa.or.jp/e/resources/database.html
2	TBestDB	http://tbestdb.bcm.umontreal.ca/searches/welcome.php
3	ArrayExpress	http://www.ebi.ac.uk/microarray-as/ae/
4	GEO DataSets	http://www.ncbi.nlm.nih.gov/sites/entrez?db=gds&cmd=search&term
5	KEGG EXPRESSION	http://www.genome.jp/kegg-bin/get_htext?htext=Exp_DB&hier=1
6	NSF2010 Metabolomics	http://lab.bcb.iastate.edu/projects/plantmetabolomics/

メタボロームデータベース：メタボローム解析に関しては近年データが蓄積し始めているので、データベースとして公開されている例が少ない。NSF2010 Metabolomics (表5-6) にシロイヌナズナのメタボローム解析データが集められている。

参考文献

- 1) K. D. Pruitt, T. Tatusova, & D. R. Maglott, *Nucleic Acids Res.* **35** (2007) D.61.
- 2) K. Liolios, K. Mavromatis, N. Tavernarakis, & N. C. Kyrpides, *Nucleic Acids Res.* **36** (2008) D.475.
- 3) T. Kaneko, S. Sato, H. Kotani, A. Tanaka, E. Asamizu, Y. Nakamura, N. Miyajima, M. Hirosawa, M. Sugiura, S. Sasamoto, T. Kimura, T. Hosouchi, A. Matsuno, A. Muraki, N. Nakazaki, K. Naruo, S. Okumura, S. Shimpou, C. Takeuchi, T. Wada, A. Watanabe, M. Yamada, M. Yasuda, & S. Tabata, *DNA Res.* **3** (1996) P. 109.
- 4) Y. Nakamura, T. Kaneko, & S. Tabata, *Nucleic Acids Res.* **28** (2000) P.72.
- 5) T. Sazuka, M. Yamaguchi, & O. Ohara, *Electrophoresis.* **20** (1999) P.2160.
- 6) S. Sato, Y. Shimoda, A. Muraki, M. Kohara, Y. Nakamura, & S. Tabata, *DNA Res.* **14** (2007) P.207.
- 7) H. Ozaki, M. Ikeuchi, T. Ogawa, H. Fukuzawa, & K. Sonoike, *Plant Cell Physiol.* **48** (2007) P.451.
- 8) D. Swarbreck, C. Wilks, P. Lamesch, T. Z. Berardini, M. Garcia-Hernandez, H. Foerster, D. Li, T. Meyer, R. Muller, L. Ploetz, A. Radenbaugh, S. Singh, V. Swing, C. Tissier, P. Zhang, & E. Huala, *Nucleic Acids Res.* **36** (2008) D.1009.
- 9) *Nucleic Acids Res.* **36** (2008) D.440.
- 10) A. P. Chan, P. D. Rabinowicz, J. Quackenbush, C. R. Buell, & C. D. Town, *Methods Mol. Biol.* **406** (2007) P. 113.
- 11) H. Schoof, R. Ernst, V. Nazarov, L. Pfeifer, H. W. Mewes, & K. F. Mayer, *Nucleic Acids Res.* **32** (2004) D. 373.
- 12) X. Pan, L. Stein, & V. Brendel, *Bioinformatics.* **21** (2005) P.3461.
- 13) T. Obayashi, K. Kinoshita, K. Nakai, M. Shibaoka, S. Hayashi, M. Saeki, D. Shibata, K. Saito, H. Ohta, *Nucleic Acids Res.* **35** (2007) D.863.
- 14) Rice Annotation Project, T. Tanaka, B. A. Antonio, S. Kikuchi, T. Matsumoto, Y. Nagamura, H. Numa, H. Sakai, J. Wu, T. Itoh, T. Sasaki, R. Aono, Y. Fujii, T. Habara, E. Harada, M. Kanno, Y. Kawahara, H. Kawashima, H. Kubooka, A. Matsuya, H. Nakaoka, N. Saichi, R. Sanbonmatsu, Y. Sato, Y. Shinso, M. Suzuki, J. Takeda, M. Tanino, F. Todokoro, K. Yamaguchi, N. Yamamoto, C. Yamasaki, T. Imanishi, T. Okido, M. Tada, K. Ikeo, Y. Tateno, T. Gojobori, Y. C. Lin, F. J. Wei, Y. I. Hsing, Q. Zhao, B. Han, M. R. Kramer, R. W. McCombie, D. Lonsdale, C. C. O'Donovan, E. J. Whitfield, R. Apweiler, K. O. Koyanagi, J. P. Khurana, S. Raghuvanshi, N. K. Singh, A. K. Tyagi, G. Haberer, M. Fujisawa, S. Hosokawa, Y. Ito, H. Ikawa, M. Shibata, M. Yamamoto, R. M. Bruskiwicz, D. R. Hoen, T. E. Bureau, N. Namiki, H. Ohyanagi, Y. Sakai, S. Nobushima, K. Sakata, R. A. Barrero, Y. Sato, A. Souvorov, B. Smith-White, T. Tatusova, S. An, G. An, S. Oota, G. Fuks, G. Fuks, J. Messing, K. R. Christie, D. Lieberherr, H. Kim, A. Zuccolo, R. A. Wing, K. Nobuta, P. J. Green, C. Lu, B. C. Meyers, C. Chaparro, B. Piegu, O. Panaud, & M. Echeverria, *Nucleic Acids Res.* **36** (2008) D.1028.
- 15) P. Jaiswal, J. Ni, I. Yap, D. Ware, W. Spooner, K. Youens-Clark, L. Ren, C. Liang, W. Zhao, K. Ratnapu, B. Faga, P. Canaran, M. Fogleman, C. Hebbard, S. Avraham, S. Schmidt, T. M. Casstevens, E. S. Buckler, L. Stein, & S. McCouche, *Nucleic Acids Res.* **34** (2006) D. 717.
- 16) S. Durinck, Y. Moreau, A. Kasprzyk, S. Davis, B. De Moor, A. Brazma, & W. Huber, *Bioinformatics.* **21** (2005) P.3439.

表6：Web アプリケーション

No	アプリケーション名	URL
1	MIPS	http://mips.gsf.de/
2	INTERPRO	http://www.ebi.ac.uk/interpro/
3	PRINTS	http://www.bioinf.manchester.ac.uk/dbbrowser/PRINTS/index.php
4	PRODOM	http://prodom.prabi.fr/prodom/current/html/home.php
5	PROSITE	http://ca.expasy.org/prosite/
6	Pfam	http://pfam.sanger.ac.uk/
7	BioMart	http://www.biomart.org/