オーミクス研究者から見たデータベース
データ生産者の目線

伊藤隆司
（東大・理・生物情報）

© 2010 伊藤　隆司　(東京大学) licensed by CC表示2.1日本

Omics オーミクス

- Genome, Epigenome, Transcriptome, Proteome, Metabolome, Metagenome etc.

- Interactome

- Grantome
  - グラントの網羅的取得？
  - DBが必要？公開できるか‥

© 2010 伊藤　隆司　(東京大学) licensed by CC表示2.1日本

データ生産者と呼ばれて

- データ生産者とは
  - とっさの連想：生産者米価、消費者米価
  - お百姓さんか・・
  - 良心的オーミクス農家とは・・
    - 他人による再利用を意識してデータを出荷する人

- データベース構築者とは
  - 流通業者？

© 2010 伊藤　隆司　(東京大学) licensed by CC表示2.1日本

10年前‥

- インタラクトームデータの収穫期
  - 酵母のタンパク質間相互作用データ
    - 2ハイブリッド総当たり戦
  - IST: Interaction Sequence Tag（30K reads）

- データをどう見せるか？
  （作物をどう出荷するか？）
  - 消費者のニーズとは‥
    - 生産者＝最初の消費者　⇒　流通業者との違い

© 2010 伊藤　隆司　(東京大学) licensed by CC表示2.1日本

Yeast Interacting Proteins Database

INTRODUCTION

We performed a comprehensive two-hybrid analysis of the budding yeast proteome. We cloned almost all of the full-length yeast ORFs individually as a DNA-binding domain fusion (bait) in a MATa strain and an activation domain fusion in a MATa strain. These clones were subdivided into pools, each containing up to 96 clones. Following the systematic mating of these pools, survivors of two-hybrid selection were subjected to sequence-tagging of bait and prey to obtain Interaction Sequence Tags (ISTs).

While "full data" contain all of the 4,549 interactions that we detected, "core data" include 841 interactions with more than 3 IST hits. Both tables contain gene name, ORF name, brief description of each bait and prey, information on their co-occurrence in literature, and other lines of evidence for their biological relevance, from which you can learn whether the interaction has been reported by others, whether synthetic lethality is reported between the two genes, whether the two genes are co-expressed, whether the two proteins have alternative interaction paths, whether similar interactions are identified in other species and how many IST hits the interaction has. All of these feature would help you evaluate the biological relevance of each interaction to build your own model for interaction networks. A viewer software for model construction is developed by INTEC Web & Genome Informatics Corporation.

We also attached a table containing all of the available interaction data on yeast poteins.

REFERENCES

Ito, T., Tashiro, K., Muta, S., Ozawa, R., Chiba, T., Nishizawa, M., Yamamoto, K., Kuhara, S. & 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.

Ito, T., Chiba, T., Ozawa, R., Yoshida, M., Hattori, M. & Sakaki, Y.
"A comprehensive two-hybrid analysis to explore the yeast protein interactome"
Proc. Natl. Acad. Sci. USA 98, 4569-4574, 2001.

DATASETS

Full data (4,549 interactions)

Excel format file

吉田美寸夫さん
(INTEC W &G)

© 2010 伊藤　隆司　(東京大学) licensed by CC表示2.1日本

Proc. Natl. Acad. Sci. USA 97, 1143-1147, 2000.

Ito, T., Chiba, T., Ozawa, R., Yoshida, M., Hattori, M. & Sakaki, Y.
"A comprehensive two-hybrid analysis to explore the yeast protein interactome"
Proc. Natl. Acad. Sci. USA 98, 4569-4574, 2001.

DATASETS

Full data (4,549 interactions)

Excel format file

Compressed File (gzip format. 300K byte - Recomended)
Uncompressed file (1.3M byte)

HTML file

Huge file. (11.2M byte)

Text file

Core data (841 interactions with more than 3 IST hits)

Excel file

Compressed File (gzip format. 1K byte - Recomended)
Uncompressed file (274K byte)

HTML file

Huge file. (2.0M byte)

Text file

You can easily see the results of our comprehensive two-hybrid analysis and previously reported protein-protein interactions using:
Workbench System for Support of Gene Regulatory Network Construction

© 2010 伊藤　隆司　(東京大学) licensed by CC表示2.1日本

	A	B	C	D	E	F	G	H	I	J	K	
1	Bait ORF	Bait gene name	Bait description	Prey ORF	Prey gene name	Prey description	Literature on bait	Literature on prey	Literature shared by bait and prey	Literature sharing score	CuraGen	S.
2	YAR033W	YAR033W	Protein with si	YAR033W	YAR033W	Protein with si	6	6			0	
3	YBL006C	YBL006C	Protein of unkr	YDR489W	YDR489W	Protein of unkr	3	3	2	2	0	
4	YBL016W	FUS3	Serine/threoni	YDR480W	DIG2	MAP kinase-a	134	11	8	47	0	
5	YBL016W	FUS3	Serine/threoni	YPL049C	DIG1	MAP kinase-a	134	12	10	55	0	
6	YBL026W	LSM2	U6 snRNA-as	YCR077C	PAT1	Protein requir	18	12	4	17	0	
7	YBL033C	RIB1	GTP cyclohyc	YNL105W	YNL105W	Protein of unkr	10	2	2	2	0	
8	YBL033C	RIB1	GTP cyclohyc	YPR172W	YPR172W	Protein of unkr	10	3	3	4	0	
9	YBL039C	URA7	CTP synthase	YJR103W	URA8	CTP synthase	24	11	10	58	0	
10	YBL045C	COR1	Ubiquinol cyto	YPR191W	QCR2	Ubiquinol cyto	27	23	14	89	0	
11	YBL056W	PTC3	Protein serine	YDR071C	YDR071C	Protein of unkr						
12	YBL068W	PRS4	Phosphoribos	YOL061W	PRS5	Phosphoribos						
13	YBR059C	AKL1	Serine/threoni	YBR059C	AKL1	Serine/threoni						
14	YBR077C	YBR077C	Protein of unkr	YKR007W	YKR007W	Protein of unkr						
15	YBR094W	YBR094W	Protein of unkr	YOL149W	DCP1	mRNA decap						
16	YBR103W	SIF2	Protein involve	YJR141W	YJR141W	Essential prot						
17	YBR108W	YBR108W	Protein with cl	YDR388W	RVS167	Protein that af						
18	YBR108W	YBR108W	Protein with cl	YGR136W	YGR136W	Protein with we						
19	YBR135W	CKS1	Protein assoc	YBR160W	CDC28	Cyclin-depen						
20	YBR135W	CKS1	Protein assoc	YBR252W	DUT1	dUTP pyroph						
21	YBR135W	CKS1	Protein assoc	YGR108W	CLB1	G2/M-phase-s						
22	YBR137W	YBR137W	Protein of unkr	YBR137W	YBR137W	Protein of unkr						
23	YBR170C	NPL4	Nuclear pore	YDL126C	CDC48	Protein of the A						
24	YBR170C	NPL4	Nuclear pore	YGR048W	UFD1	Ubiquitin fusio						
25	YBR187W	YBR187W	Protein with si	YNR032W	PPG1	Protein serine						
26	YBR188C	NTC20	Splicing factor	YDR416W	SYF1	Pre-mRNA sp						
27	YBR216C	YBR216C	Protein of unkr	YML007W	YAP1	Transcriptiona						
28	YBR217W	APG12	Protein conjug	YLR423C	APG17	Protein involv						
29	YBR217W	APG12	Protein conjug	YMR159C	APG16	Coiled-coil pr						
30	YBR223C	YBR223C	Tyrosyl-DNA	YDR510W	SMT3	Ubiquitin-relat						
31	YBR228W	YBR228W	Protein of unkr	YLR135W	YLR135W	Protein of unkr						
32	YBR239C	YBR239C	Protein with si	YPL133C	YPL133C	Protein with si						
33	YBR246W	YBR246W	Protein of unkr	YDR520C	YDR520C	Protein with si						
34	YBR254C	TRS20	Protein of the T	YDR472W	TRS31	Protein of the T	9	6	6	26	1	
35	YBR254C	TRS20	Protein of the T	YKR068C	BET3	Low molecula	9	9	5	20	0	
36	YBR260C	RGD1	GTPase-activ	YMR032W	HOF1	Protein involv	10	12	3	4	0	

semifinal table 20001206

Systematic Nameで分かるユーザはいない

全遺伝子を知ってるユーザはいない

© 2010 伊藤　隆司　(東京大学) licensed by CC表示2.1日本

当時の相互作用DBは
ショボ過ぎ

文献情報しか頼れない

Yeast Proteome
Databaseの利用
(マニュアルキュレー
ションによる商用DB)

	A	B	C	D	E	F	G	H	I	J	K	
1	Bait ORF	Bait gene name	Bait description	Prey ORF	Prey gene name	Prey description	Literature on bait	Literature on prey	Literature shared by bait and prey	Literature sharing score	CuraGen	S.
2	YAR033W	YAR033W	Protein with sir	YAR033W	YAR033W	Protein with sir	6	6			0	
3	YBL006C					f unkr	3	3	2	2	0	
4	YBL016W					ase-a	134	11	8	47	0	
5	YBL016W					ase-a	134	12	10	55	0	
6	YBL026W					equire	18	12	4	17	0	
7	YBL033C					f unkr	10	2	2	2	0	
8	YBL033C					f unkr	10	3	3	4	0	
9	YBL039C					hase	24	11	10	58	0	
10	YBL045C					l cyto	27	23	14	89	0	
11	YBL056W					f unkr	13	3	0	0	0	
12	YBL068W					ribos	9	11	8	39	0	
13	YBR059C					reoni	8	8			0	
14	YBR077C					f unkr	3	6	1	3	1	
15	YBR094W					ecap	5	31	3	5	0	
16	YBR103W					prote	8	3	3	7	0	
17	YBR108W					at aff	7	34	4	11	0	
18	YBR108W					ith we	7	6	3	5	0	
19	YBR135W					pend	21	334	16	112	0	
20	YBR135W					roph	21	10	3	5	0	
21	YBR135W					ase-s	21	60	7	33	1	
22	YBR137W					f unkr	5	5			0	
23	YBR170C					f the A	15	46	3	5	0	
24	YBR170C					fusio	15	10	5	18	1	
25	YBR187W					erine,	6	8	2	2	0	
26	YBR188C					lA sp	7	7	4	7	1	
27	YBR216C					tione	3	85	3	4	0	
28	YBR217W					volve	16	6	4	7	0	
29	YBR217W					il pro	16	11	7	31	1	
30	YBR223C					relat	5	19	1	1	0	
31	YBR228W					f unkr	5	5	4	7	1	
32	YBR239C						4	5	3	4	0	
33	YBR246W	YBR246W	Protein of unkr	YDR520C	YDR520C	Protein with sir	3	2	2	2	0	
34	YBR254C	TRS20	Protein of the T	YDR472W	TRS31	Protein of the T	9	6	6	26	1	
35	YBR254C	TRS20	Protein of the T	YKR068C	BET3	Low molecula	9	9	5	20	0	
36	YBR260C	RGD1	GTPase-activ	YMR032W	HOF1	Protein involve	10	12	3	4	0	

semifinal table 20001206

© 2010 伊藤　隆司　(東京大学) licensed by CC表示2.1日本

This Protein Report of the BioKnowledge Library™ is produced by BIOBASE Corporation for employees of organizations that use the BioKnowledge Library in whole or in part in the ordinary course of performing pharmaceutical research and development for the subscribing organization. The BioKnowledge Library and this report are proprietary products of BIOBASE Corporation and are protected under U.S. copyright law and international treaty provisions. Employees or affiliates of any subscribing organization shall not at any time furnish copies of all or part of this Protein Report to any third party without the express prior written consent of BIOBASE Corporation.
Copyright BIOBASE Corporation

Quick Search: Name/ID

BIOBASE

Proteome BioKnowledge® Library
YPD ™
Saccharomyces cerevisiae

PDH1

Protein possibly involved in propionate utilization

Gene Symbol/Synonyms PDH1; YP9723.02; LPZ2W; YPR002W
Gene Families

Tools

Gene Ontology
Molecular Function Molecular function unknown [details]
Biological Process Propionate catabolism [E] [details]
Cellular Component Mitochondrial outer membrane [E]; Mitochondrion [E]; Cytoplasm [E] [details]

BIOKNOWLEDGE RETRIEVER

Mutant Phenotype Viable [E]; Altered sensitivity to drugs or other compounds [E]... [details]

Sequence
Full mflaknl...mdlfteg (1..516; 516 aa)
pI: 9.26 **MW:** 57694 **TM:** 0 [P]
Mature mflaknl...mdlfteg [P] (1..516; 516 aa) [details]
pI: 9.26 **MW:** 57694 **TM:** 0 [P]
Codon Usage **Codon Bias:** 0.018 **CAI:** 0.114
Gene **Chromosome:** XVI **Introns:** 0 [E]
3D Structure (PDB) 1SZQ_A (58%) [details]
Domain MmgE/PrpD family [details]

PROTEIN SEQUENCE

Related Proteins
H. sapiens
M. musculus IRG1 (24%) [details]
R. norvegicus

Mutant phenotype

- null mutant exhibits increased sensitivity to 50 mM propionate when grown on minimal c grown on 0.01% to 0.1% raffinose medium (11179416)
- null mutant shows comparable growth with wild-type cells on acetate medium (1117941

Localization

- localizes to the cytoplasm and mitochondria, as determined for a C-terminally tagged C
- detected in highly purified mitochondria by large-scale tandem mass spectrometry anal
- one of 112 proteins or precursor proteins localized to the mitochondrial outer membrane/surface fraction, based on experiments using subcellular fractionation and mass spectroscopy (16407407)

Regulation

- transcription in response to mitochondrial dysfunction is independent of Pdr3p function (11943188)
- transcription may be regulated by Rap1p, Rtg1p, and Ste12p, as suggested by computational analyses using support vector machines (SVMs) (14555624)
- mRNA and protein abundance is increased in response to diauxic shift, as determined by microarray and 2D-gel analyses (14597615)
- one of 1004 cell-cycle regulated genes, as determined by microarray and computational analyses (14960459)
- one of 552 SAGA-dominated genes whose transcription is regulated by the SAGA complex and is largely independent of Taf1p, as determined by microarray analysis (14992726)
- one of 312 class VI genes that are repressed in response to glucose, Ras2p overproduction, or Gpa2p overproduction, as determined by microarray analysis (15138498)
- one of 84 mRNAs whose abundance increases in cth2 null mutant cells in response to iron deprivation (15652485)
- mRNA contains at least one AU-rich element in the 3' untranslated region and may be regulated by Cth2p-mediated mRNA degradation in response to iron deprivation (15652485)

Abundance

- present in approximately 2.73E+03 molecules per log-phase cell, as determined by (14562106)

Gel mobility

- has an experimental isoelectric point of approximately 8.85, and runs at 53.3 kDa o

Functional Genomics

- null mutant, constructed in a systematic high-throughput deletion study, is viable (10436161)
- one of 43 genes coregulated in cells lacking total mitochondrial genome independently of cell density [FZF1 YJL213W ADH2 SPG2 FRE4 TNA1 PYC1 YNL203C YLR089C PXA1 VID24 BAT2 PEX11 YER187W YLR004C SPO1 ERR2 DLD1 YOL119C CTF19 YAR068W YMR191W PRM1 GPD2 AGP2] (11179416)
- one of 43 genes coregulated in cells lacking total mitochondrial genome independently of cell density [YJL037W YDR384C PDH1 CIT3 CIT2 HXT10 DIC1 ACS1 YIL057C DIP5 NCE103 CDA1 ALD4 PHO89 CRC1 YLR349W JEN1 SPS19] (11179416)
- one of 19 genes coinduced by treatment with 1mg/ml antimycin and in mutan

要旨収集
ではなくて
文献代読

オーミクスデータの
解析代行

© 2010 伊藤　隆司　(東京大学) licensed by CC表示2.1日本

文献情報から探る相互作用データの尤もらしさ

	A	B	C	D	E	F	G	H	I	J	K	L
1	Bait ORF	Bait gene name	Bait description	Prey ORF	Prey gene name	Prey description	Literature on bait	Literature on prey	Literature shared by bait and prey	Literature sharing score	CuraGen	S.
2	YAR033W	YAR033W	Protein with sir	YAR033W	YAR033W	Protein with sir	6	6			0	
3	YBL006C					f unkr	3	3	2	2	0	
4	YBL016W					se-a	134	11	8	47	0	
5	YBL016W					se-a	134	12	10	55	0	
6	YBL026W					equire	18	12	4	17	0	
7	YBL033C					f unkr	10	2	2	2	0	
8	YBL033C					f unkr	10	3	3	4	0	
9	YBL039C					hase	24	11	10	58	0	
10	YBL045C	COR1	Ubiquinol cyto	YPR191W	QCR2	Ubiquinol cyto	27	23	14	89	0	
11	YBL056W	PTC3	Protein serine/	YDR071C	YDR071C	Protein of unkr	13	3	0	0	0	
12	YBL068W	PRS4	Phosphoribos	YOL061W	PRS5	Phosphoribos	9	11	8	39	0	
13	YBR059C	AKL1	Serine/threoni	YBR059C	AKL1	Serine/threoni	8	8			0	
14	YBR077C	YBR077C	Protein of unkr	YKR007W	YKR007W	Protein of unkr	3	6	1	3	1	
15	YBR094W	YBR094W	Protein of unkr	YOL149W	DCP1	mRNA decap	5	31	3	5	0	
16	YBR103W	SIF2	Protein involve	YJR141W	YJR141W	Essential prot	8	3	3	7	0	
17	YBR108W	YBR108W	Protein with clu	YDR388W	RVS167	Protein that aff	7	34	4	11	0	
18	YBR108W	YBR108W	Protein with clu	YGR136W	YGR136W	Protein with we	7	6	3	5	0	
19	YBR135W	CKS1	Protein assoc	YBR160W	CDC28	Cyclin-depend	21	334	16	112	0	
20	YBR135W	CKS1	Protein assoc	YBR252W	DUT1	dUTP pyroph	21	10	3	5	0	
21	YBR135W	CKS1	Protein assoc	YGR108W	CLB1	G2/M-phase-s	21	60	7	33	1	
22	YBR137W	YBR137W	Protein of unkr	YBR137W	YBR137W	Protein of unkr	5	5			0	
23	YBR170C	NPL4	Nuclear pore	YDL126C	CDC48	Protein of the A	15	46	3	5	0	
24	YBR170C	NPL4	Nuclear pore	YGR048W	UFD1	Ubiquitin fusio	15	10	5	18	1	
25	YBR187W	YBR187W	Protein with sir	YNR032W	PPG1	Protein serine/	6	8	2	2	0	
26	YBR188C	NTC20	Splicing factor	YDR416W	SYF1	Pre-mRNA sp	7	7	4	7	1	
27	YBR216C	YBR216C	Protein of unkr	YML007W	YAP1	Transcriptione	3	85	3	4	0	
28	YBR217W	APG12	Protein conjug	YLR423C	APG17	Protein involve	16	6	4	7	0	
29	YBR217W	APG12	Protein conjug	YMR159C	APG16	Coiled-coil pro	16	11	7	31	1	
30	YBR223C	YBR223C	Tyrosyl-DNA	YDR510W	SMT3	Ubiquitin-relat	5	19	1	1	0	
31	YBR228W	YBR228W	Protein of unkr	YLR135W	YLR135W	Protein of unkr	5	5	4	7	1	
32	YBR239C	YBR239C	Protein with sir	YPL133C	YPL133C	Protein with sir	4	5	3	4	0	
33	YBR246W	YBR246W	Protein of unkr	YDR520C	YDR520C	Protein with sir	3	2	2	2	0	
34	YBR254C	TRS20	Protein of the T	YDR472W	TRS31	Protein of the T	9	6	6	26	1	
35	YBR254C	TRS20	Protein of the T	YKR068C	BET3	Low molecula	9	9	5	20	0	
36	YBR260C	RGD1	GTPase-activ	YMR032W	HOF1	Protein involve	10	12	3	4	0	

semifinal table 20001206

© 2010 伊藤　隆司　(東京大学) licensed by CC表示2.1日本

競合データセットとの重複

	K	L	M	N	O	P	Q	R	S	T	U
1	CuraGen	S. Fields	Association (YPD)	Complex (YPD)	Synthetic lethality	Co-induced by	Co-repressed by	Not affected by	Interologs	Expression similarity (BRITE)	Alternative path with 1 intervening protein
2	0	0	0	0	0						0
3	0	0									0
4	0	0									0
5	0	0									0
6	0	0									1
7	0	0									0
8	0	0									0
9	0	0	0	0	1		stationary phase				0
10	0	0	1	1	0	heme			2		7
11	0	0	0	0	0						0
12	0	0	1	1	0						0
13	0	0	0	0	0				11		0
14	1	0	0	0	0						0
15	0	0	0	0	0						1
16	0	0	0	0	0						0
17	0	0	0	0	0						1
18	0	0	0	0	0						0
19	0	0	1	1	0				5		2
20	0	0	0	0	0						0
21	1	0	0	1	0				1		0
22	0	0	0	0	0						0
23	0	0	0	0	0						0
24	1	0	0	0	0						0
25	0	0	0	0	0						0
26	1	0	0	0	0						0
27	0	0	0	0	0						0
28	0	0	0	0	0						0
29	1	0	1	0	0						1
30	0	0	0	0	0						0
31	1	0	0	0	0						0
32	0	0	0	0	0						0
33	0	0	0	0	0						0
34	1	0	0	1	0						0
35	0	0	1	1	0						1
36	0	0	0	0	0						0

semifinal table 20001206

© 2010 伊藤　隆司　(東京大学) licensed by CC表示2.1日本

	K	L	M	N	O	P	Q	R	S	T	U	
1	CuraGen	S. Fields	Association (YPD)	Complex (YPD)	Synthetic lethality	Co-induced by	Co-repressed by	Not affected by	Interologs	Expression similarity (BRITE)	Alternative path with 1 intervening protein	Alt inte p
2	0	0	0	0	0						0	
3	0	0	0								0	
4	0	0	1								0	
5	0	0	1								0	
6	0	0	0								1	
7	0	0	0								0	
8	0	0	0								0	
9	0	0	0								0	
10	0	0	1						2		7	
11	0	0	0								0	
12	0	0	1	1	0						0	
13	0	0	0	0	0				11		0	
14	1	0	0	0	0						0	
15	0	0	0	0	0						1	
16	0	0	0	0	0						0	
17	0	0	0	0	0						1	
18	0	0	0	0	0						0	
19	0	0	1	1	0				5		2	
20	0	0	0	0	0						0	
21	1	0	0	1	0				1		0	
22	0	0	0	0	0						0	
23	0	0	0	0	0						0	
24	1	0	0	0	0						0	
25	0	0	0	0	0						0	
26	1	0	0	0	0						0	
27	0	0	0	0	0						0	
28	0	0	0	0	0						0	
29	1	0	1	0	0						1	
30	0	0	0	0	0						0	
31	1	0	0	0	0						0	
32	0	0	0	0	0						0	
33	0	0	0	0	0						0	
34	1	0	0	1	0						0	
35	0	0	1	1	0						1	
36	0	0	0	0	0						0	

複合体情報・遺伝的相互作用・共発現情報
etc

semifinal table 20001206

© 2010 伊藤　隆司　(東京大学) licensed by CC表示2.1日本

ネットワークの構造

	M	N	O	P	Q	R	S	T	U	V	W
1	Association (YPD)	Complex (YPD)	Synthetic lethality	Co-induced by	Co-repressed by	Not affected by	Interologs	Expression similarity (BRITE)	Alternative path with 1 intervening protein	Alternative path with 2 intervening proteins	IST hit
2	0	0	n						0	0	4
3	0	0							0	0	7
4	1	0							0	7	8
5	1	0							0	7	10
6	0	1	0						1	6	3
7	0	0	0						0	0	4
8	0	0	0						0	0	4
9	0	0	1		stationary phase				0	0	4
10	1	1	0	heme			2		7	47	11
11	0	0	0						0	0	4
12	1	1	0						0	0	4
13	0	0	0				11		0	0	4
14	0	0	0						0	0	12
15	0	0	0						1	4	9
16	0	0	0						0	0	3
17	0	0	0						1	5	8
18	0	0	0						0	0	7
19	1	1	0				5		2	2	6
20	0	0	0						0	0	3
21	0	1	0				1		0	2	3
22	0	0	0						0	0	6
23	0	0	0						0	0	4
24	0	0	0						0	0	20
25	0	0	0						0	0	3
26	0	0	0						0	0	16
27	0	0	0						0	0	7
28	0	0	0						0	0	3
29	1	0	0						1	1	4
30	0	0	0						0	0	3
31	0	0	0						0	0	7
32	0	0	0						0	0	8
33	0	0	0						0	0	6
34	0	1	0						0	0	23
35	1	1	0						1	1	4
36	0	0	0						0	2	3

semifinal table 20001206

© 2010 伊藤　隆司　(東京大学) licensed by CC表示2.1日本

データの再現性（信頼性）

	M	N	O	P	Q	R	S	T	U	V	W
1	Association (YPD)	Complex (YPD)	Synthetic lethality	Co-induced by	Co-repressed by	Not affected by	Interologs	Expression similarity (BRITE)	Alternative path with 1 intervening protein	Alternative path with 2 intervening proteins	IST hit
2	0	0	n						0	0	4
3	0	0							0	0	7
4	1	0							0	7	8
5	1	0							0	7	10
6	0	1							1	6	3
7	0	0							0	0	4
8	0	0							0	0	4
9	0	0	1		stationary phase				0	0	4
10	1	1	0	heme			2		7	47	11
11	0	0	0						0	0	4
12	1	1	0						0	0	4
13	0	0	0				11		0	0	4
14	0	0	0						0	0	12
15	0	0	0						1	4	9
16	0	0	0						0	0	3
17	0	0	0						1	5	8
18	0	0	0						0	0	7
19	1	1	0				5		2	2	6
20	0	0	0						0	0	3
21	0	1	0				1		0	2	3
22	0	0	0						0	0	6
23	0	0	0						0	0	4
24	0	0	0						0	0	20
25	0	0	0						0	0	3
26	0	0	0						0	0	16
27	0	0	0						0	0	7
28	0	0	0						0	0	3
29	1	0	0						1	1	4
30	0	0	0						0	0	3
31	0	0	0						0	0	7
32	0	0	0						0	0	8
33	0	0	0						0	0	6
34	0	1	0						0	0	23
35	1	1	0						1	1	4
36	0	0	0						0	2	3

semifinal table 20001206

© 2010 伊藤　隆司　(東京大学) licensed by CC表示2.1日本

Workbench System for Support of Gene Regulatory Network Construction:

Genetic Network Visualization System

About this system

Using this system, you can easily see the results of our comprehensive two-hybrid analysis and previously reported protein-protein interactions. We expect that the system will help you built your own model for interaction networks. Both the system and our data can be freely downloaded from this site. The system was just developed by a Japanese company, INTEC Web and Genome Informatics Corporation. If you have some troubles with this system, please contact them by e-mail at the following address: mailto:webgen@gic.intec.co.jp.

Environment

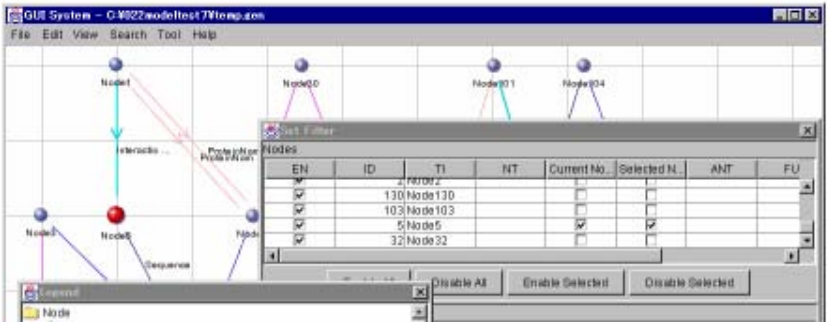
This system operates as an application of Java 2. To run this system, Java 2, version 1.2 or 1.3 (1.3 is recommended), is required. Java 2 Runtime can be obtained from the Web site of Sun Microsystems, Inc. We have already confirmed that it can run on Windows 95/98/2000/NT4.0, Linux 2.2.12 or higher and Solaris 2.6 or higher.

For Mac users:
We are afraid that this software cannot run on Mac OS because Java 2 does not support the current Mac OS.

Introduction

1. **Overview of the system**

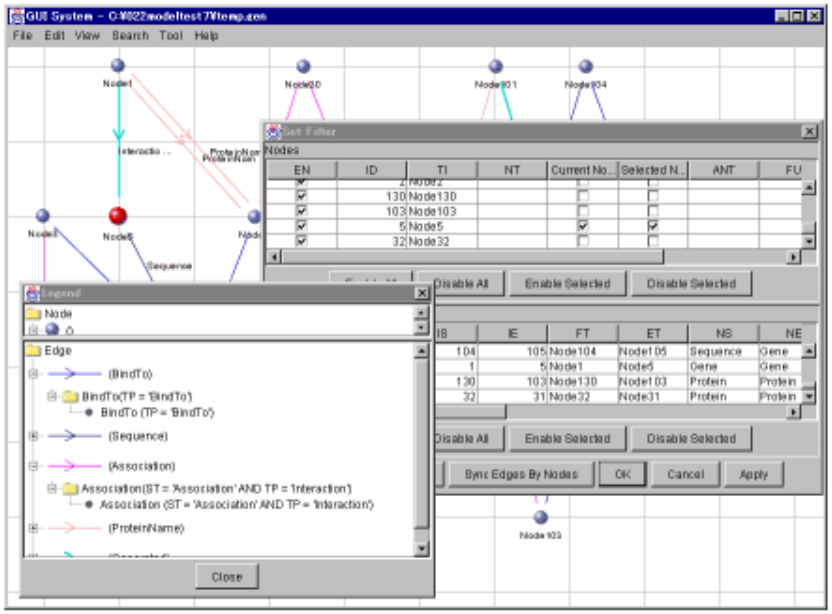
   This system supports the construction of a gene regulatory network. In this software, a genetic network is represented as a graph. You can see **protein-protein interactions**, **connected components** and **alternative paths**.



© 2010 伊藤　隆司　(東京大学) licensed by CC表示2.1日本

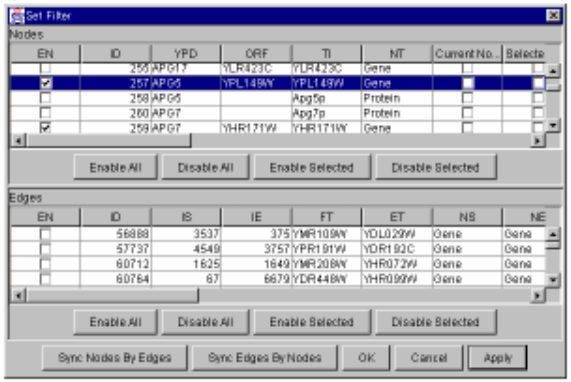
1. Overview of the system

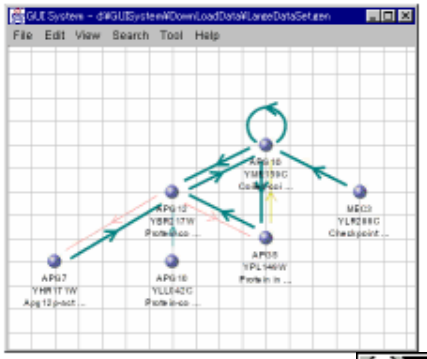
This system supports the construction of a gene regulatory network. In this software, a genetic network is represented as a graph. You can see **protein-protein interactions**, **connected components** and **alternative paths**.



2. Protein-protein interactions

You can search protein-protein interactions and extend it to a network, or you can confirm whether an interaction has been previously screened.





© 2010 伊藤　隆司　(東京大学) licensed by CC表示2.1日本

生産者の歓び

- これまでにない種類の作物を出荷したい
- 当然、流通業者も未経験
- 流通のさせ方にも工夫したい

© 2010 伊藤　隆司　(東京大学) licensed by CC表示2.1日本

生産者の悲哀

- 生産者の気遣いを無視する“心ない”流通業者が横行
  - 消費者のための品質表示の無視
    （core dataとnon-core data）
- 生産者直売店は閉鎖の憂き目
  - 前任地のwebsiteはリンク切れ
- そんなとき‥

© 2010 伊藤　隆司　(東京大学) licensed by CC表示2.1日本

Life Science Database Archive LSDB Archive

文部科学省委託研究開発事業 統合データベースプロジェクト
生命科学系データベース アーカイブ

Yeast Interacting Proteins Database

- データベースの説明
- ダウンロード
- 利用許諾
- このDBの更新履歴

ダウンロード

はじめに利用許諾をお読みください。
なお、ここに表示したデータ名およびその説明は、DBCLSのサイトからダウンロード可能なデータについてのものです。オリジナルのデータベースの内容とは一致しない場合があります。

「データ名」のリンクをクリックすると、データ項目の説明を参照できます。

#	データ名	データファイル	検索＆ダウンロード
1	README	README.html	-
2	Yeast Interacting Proteins DatabaseのCoreデータ(アノテーション更新版)	core_updated.zip (77KB)	検索＆ダウンロード
3	Yeast Interacting Proteins DatabaseのFullデータ(アノテーション更新版)	full_updated.zip (372KB)	検索＆ダウンロード
4	Yeast Interacting Proteins DatabaseのCoreデータ(オリジナル版)	core_original.zip (37KB)	検索＆ダウンロード
5	Yeast Interacting Proteins DatabaseのFullデータ(オリジナル版)	full_original.zip (180KB)	検索＆ダウンロード

アノテーション更新版

DBCLS Database Center for Life Science

＜利用者の方へ＞

- ダウンロードボタン([Download])を押す前に利用許諾を注意深くお読み下さい。ダウンロードボタンを押すことによって、本利用許諾を承諾したものと見なします。

Show Advanced Search | Search | Reset | Download

841 Found

15 | 1 2 3 .. 57 | Next

Bait ORF	Bait gene name	Bait description	Rows with this bait as bait	Rows with this bait as prey	Prey ORF	Prey gene name	Prey description	Rows with this prey as prey	Rows with this prey as bait	Literature on bait (YPD)	Literature on prey (YPD)
YAR033W	YAR033W	Protein with similarity to subtelomerically-encoded proteins such as Cos5p, Ybr302p, Cos3p, Cos1p, Cos4p, Cos8p, Cos6p, Cos9p, and Ycr007p	Rows with this bait as bait (1)	Rows with this bait as prey (1)	YAR033W	YAR033W	Protein with similarity to subtelomerically-encoded proteins such as Cos5p, Ybr302p, Cos3p, Cos1p, Cos4p, Cos8p, Cos6p, Cos9p, and Ycr007p	Rows with this prey as prey (1)	Rows with this prey as bait (1)	6	6
YBL006C	YBL006C	Protein of unknown function	Rows with this bait as bait (1)	Rows with this bait as prey (0)	YDR489W	YDR489W	Protein of unknown function	Rows with this prey as prey (5)	Rows with this prey as bait (0)	3	3
		Serine/threonine protein kinase required for cell cycle	Rows with this	Rows with this			MAP kinase-associated	Rows with this	Rows with this		

YBL068W	PRS4	pyrophosphokinase), enxyme that synthesizes phosphoribosylpyrophosphate (PRPP) from ribose-5' phosphate and ATP	this **bait** as **bait** (1)	this **bait** as **prey** (0)	**YOL061W**	PRS5	Phosphoribosylpyrophosphate synthetase (ribose-phosphate pyrophosphokinase)	this **prey** as **prey** (2)	this **prey** as **bait** (1)	9	11
YBR059C	AKL1	Serine/threonine protein kinase of unknown function	Rows with this **bait** as **bait** (1)	Rows with this **bait** as **prey** (3)	**YBR059C**	AKL1	Serine/threonine protein kinase of unknown function	Rows with this **prey** as **prey** (3)	Rows with this **prey** as **bait** (1)	8	8
YBR077C	YBR077C	Protein of unknown function	Rows with this **bait** as **bait** (1)	Rows with this **bait** as **prey** (1)	**YKR007W**	YKR007W	Protein of unknown function	Rows with this **prey** as **prey** (1)	Rows with this **prey** as **bait** (2)	3	6
YBR094W	YBR094W	Protein of unknown function	Rows with this **bait** as **bait** (1)	Rows with this **bait** as **prey** (0)	**YOL149W**	DCP1	mRNA decapping enzyme involved in mRNA turnover	Rows with this **prey** as **prey** (6)	Rows with this **prey** as **bait** (1)	5	31
YBR103W	SIF2	Protein involved in telomere silencing, interacts with Sir4p and targets Sir4p from telomeres to other sites, has WD (WD-40) repeats	Rows with this **bait** as **bait** (1)	Rows with this **bait** as **prey** (1)	**YJR141W**	YJR141W	Essential protein of unknown function	Rows with this **prey** as **prey** (3)	Rows with this **prey** as **bait** (1)	8	3

841 Found

15 1 2 3 .. 57 | Next

YBL068W	PRS4	for nucleotide, histidine, and tryptophan biosynthesis; one of five related enzymes, which are active as heteromultimeric complexes	this **bait** as **bait** (1)	this **bait** as **prey** (0)	**YOL061W**	PRS5	for nucleotide, histidine, and tryptophan biosynthesis; one of five related enzymes, which are active as heteromultimeric complexes	this **prey** as **prey** (2)	this **prey** as **bait** (1)	9	11	8	39
YBR059C	AKL1	Ser-Thr protein kinase, member (with Ark1p and Prk1p) of the Ark kinase family; involved in endocytosis and actin cytoskeleton organization	Rows with this **bait** as **bait** (1)	Rows with this **bait** as **prey** (3)	**YBR059C**	AKL1	Ser-Thr protein kinase, member (with Ark1p and Prk1p) of the Ark kinase family; involved in endocytosis and actin cytoskeleton organization	Rows with this **prey** as **prey** (3)	Rows with this **prey** as **bait** (1)	8	8	-	-
YBR077C	SLM4	Component of the EGO complex, which is involved in the regulation of microautophagy, and of the GSE complex, which is required for proper sorting of amino acid permease Gap1p; gene exhibits synthetic genetic interaction with MSS4	Rows with this **bait** as **bait** (1)	Rows with this **bait** as **prey** (1)	**YKR007W**	MEH1	Component of the EGO complex, which is involved in the regulation of microautophagy, and of the GSE complex, which is required for proper sorting of amino acid permease Gap1p; loss results in a defect in vacuolar acidification	Rows with this **prey** as **prey** (1)	Rows with this **prey** as **bait** (2)	3	6	1	3
YBR094W	PBY1	Putative tubulin tyrosine ligase associated with P-bodies	Rows with this **bait** as **bait** (1)	Rows with this **bait** as **prey** (0)	**YOL149W**	DCP1	Subunit of the Dcp1p-Dcp2p decapping enzyme complex, which removes the 5' cap structure from mRNAs prior to their degradation; enhances the activity of catalytic subunit Dcp2p;	Rows with this **prey** as **prey** (6)	Rows with this **prey** as **bait** (1)	5	31	3	5

データアーカイブ

- データ公開の代行
  - 維持困難DBの救済
- アノテーションの更新
  - 消費者にとっては便利
  - 生産者にとっては負担

© 2010 伊藤　隆司　(東京大学) licensed by CC表示2.1日本

データの永代供養

- データ公開の代行
  - 維持困難DBの救済
- アノテーションの更新
  - 消費者にとっては便利
  - 生産者にとっては負担
- 論文からのリンク先更新

© 2010 伊藤　隆司　(東京大学) licensed by CC表示2.1日本

Mirzayan Policy Fellowships
The National Academies
Now Accepting Applications

Sign up for PNAS Online eTOC alerts
Get notified by e-mail when new content is posted online

Listen Discover Experience
Science Sessions
PNAS Podcasts

Institution: UNIVERSITY OF TOKYO | Sign In as Member / Individual

Info for Authors | Editorial Board | About | Subscribe | Advertise | Contact | Feedback | Site Map

PNAS
Proceedings of the National Academy of Sciences of the United States of America

A comprehensive two-hybrid analysis to explore the yeast protein interactome

Takashi Ito [*], [†], Tomoko Chiba [*], Ritsuko Ozawa [‡], Mikio Yoshida [§], Masahira Hattori [‡], and Yoshiyuki Sakaki [‡], [¶]

+ Author Affiliations

Communicated by Satoshi Omura, The Kitasato Institute, Tokyo, Japan (received for review December 4, 2000)

Abstract

Protein–protein interactions play crucial roles in the execution of various biological functions. Accordingly, their comprehensive description would contribute considerably to the functi
genomes, which are flooded with nov
previously developed a system to exami
combinations between the ≈6,000 prot
cerevisiae. Here we have completed the
to identify 4,549 two-hybrid interactio
these data do not largely overlap with t
P., *et al.* (2000) *Nature (London)* 403,
expanded our knowledge on the protein interaction space or interactome of the yeast. Cumulative connection of these binary interactions generates a single huge network linking the vast majority of the proteins. Bioinformatics-aided selection of biologically relevant interactions highlights various intriguing subnetworks. They include, for instance, the one that had successfully foreseen the involvement of a novel protein in spindle pole body function as well as the one that may uncover a hitherto unidentified multiprotein complex potentially

ここから

« Previous | Next Article »
Table of Contents

This Article

Published online before print March 13, 2001, doi: 10.1073/pnas.061034498
PNAS April 10, 2001 vol. 98 no. 8 4569-4574

Abstract *Free*
Figures Only
» Full Text
Full Text (PDF)

– Classifications
Biological Sciences
Genetics

– Services

Request copyright permission

– Citing Articles
View citing article information
Citing articles via CrossRef
Citing Articles via Web of Science (1526)
Articles citing this article

+ Google Scholar

Search PNAS GO
advanced search >>

This Week's Issue
September 28, 2010, 107 (39)

From the Cover
- Sustaining Chile's marine resources
- Liquid silicon phase transitions
- ATP synthase stoichiometry
- Reducing corneal inflammation
- Prenatal famine and adult cognition

Alert me to new issues of PNAS

Early Edition
Archives
Online Submission
Feature Articles
Commentaries

MEXT Integrated Database Project

Life Science Database Archive

Yeast Interacting Proteins Database

- Database Description
- Download
- License
- Update History of This Database

Download

First of all, please read the license of this database.
Data names and data descriptions are about the downloadable data in this page. They might not correspond to the contents of the original database.

Click the links on "Data Name" for descriptions of the data.

#	Data name	File	Simple search and download
1	README	README_e.html	-
2	Core Data of Yeast Interacting Proteins Database (Annotation Updated Version)	core_updated.zip (77KB)	Simple search and download
3	Full Data of Yeast Interacting Proteins Database (Annotation Updated Version)	full_updated.zip (372KB)	Simple search and download
4	Core Data of Yeast Interacting Proteins Database (Original Version)	core_original.zip (37KB)	Simple search and download
5	Full Data of Yeast Interacting Proteins Database (Original Version)	full_original.zip (760KB)	Simple search and download

- Download via FTP

ここへ

飛べるとよい

データ(生産者)の成仏

- データ公開の代行
  - 維持困難DBの救済
- アノテーションの更新
  - 消費者にとっては便利
  - 生産者にとっては負担
- 論文からのリンク先更新

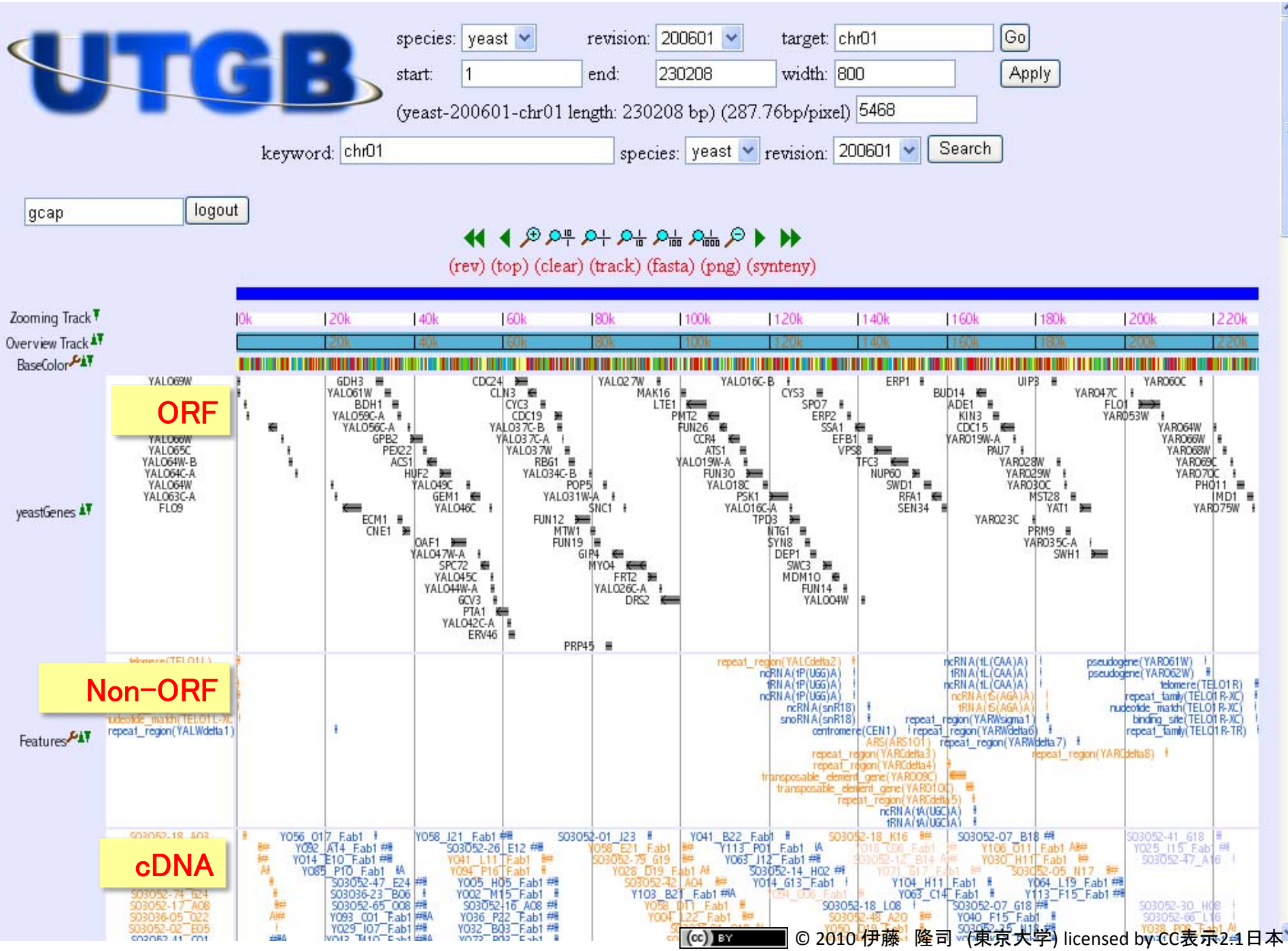
© 2010 伊藤　隆司　(東京大学) licensed by CC表示2.1日本

５年前‥

- 完全長cDNAデータの収穫期
  - 出芽酵母の転写開始点
    - ベクターキャップ法cDNAライブラリ
  - 5’シングルパス（80K reads）

- UTGB(UT Genome Browser)によるデータ公開

© 2010 伊藤　隆司　(東京大学) licensed by CC表示2.1日本



© 2010 伊藤　隆司（東京大学) licensed by CC表示2.1日本

species: yeast revision: 200601 target: chr03 Go

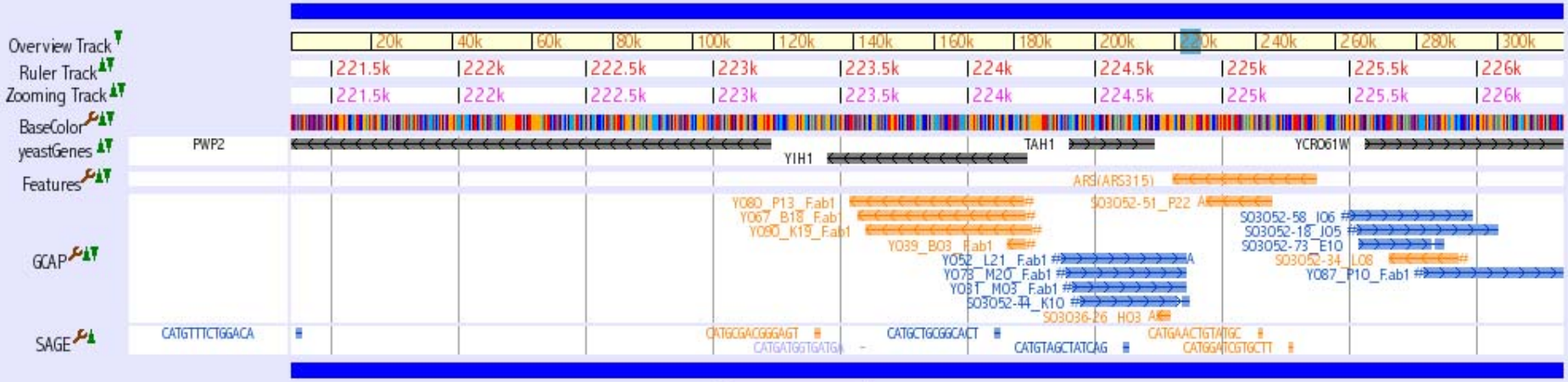
start: 221338 end: 226337 width: 800 Apply

(yeast-200601-chr03 length: 316617 bp) (6.25bp/pixel)

keyword: YIH1 species: yeast revision: 200601 Search

Guest login

(rev) (top) (clear) (track) (fasta) (png) (synteny)



save this view

© 2010 伊藤　隆司　(東京大学) licensed by CC表示2.1日本

Saccharomyces Genome Database

Search

Site Map | Search Options | Help | Home |

Community Info | Submit Data | BLAST | Primers | PatMatch | Gene/Seq Resources | Advanced Search | Community Wiki

Search Options
Advanced Search, Full-text Search (Textpresso), Search SGD web pages, Global Gene Hunter, Search Literature, and more.

Help Resources
Getting Started, Sitemap, FAQ, and more.

Analysis & Tools
BLAST, Gene/Seq Resources, Maps, and more.

Homology & Comparisons
PDB Homologs, Protein Domains/Motifs, Homologs, and more.

Function & Expression
Protein Info, Pathways, Expression Connection, and more.

GO Resources
GO Tutorial, What is GO?
GO Slim Mapper,
GO Term Finder, and mor

Community Inf
Search SGD colleagues,
Conferences & Seminars, Community wiki, and more.

SGD™ is a scientific database of the molecular biology and genetics of the yeast *Saccharomyces cerevisiae*, which is commonly known as baker's or budding yeast.

New and Noteworthy

- **SGD on Facebook and Twitter** - *August 26, 2010*

  SGD is now on Facebook and Twitter! 'Like' us on Facebook or follow @yeastgenome on Twitter to get updates about new features, tips on using SGD, and interesting tidbits about yeast.

- **SGD Curation News**

  - Genome Snapshot
  - New papers added to SGD this week.
  - View Genome-wide Analysis papers in SGD.
  - Do a literature search.

酵母の統合DB

GD. An
in the future

- **New resource links** - *July 14, 2010*

YIH1/YCR059C Summary

Help

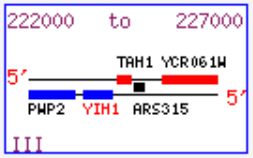
YIH1 BASIC INFORMATION

Standard Name	*YIH1* (*see Nomenclature conflict Note*)
Systematic Name	YCR059C
Feature Type	ORF, Verified
Description	Protein that inhibits activation of Gcn2p, an eIF2 alpha subunit protein kinase, by competing for Gcn1p binding, thus impacting gene expression in response to starvation; has sequence and functional similarity to the mouse IMPACT gene (1, 2, 3)
Name Description	Yeast Impact Homolog [1]
GO Annotations	*All **YIH1** GO evidence and references* *View Computational GO annotations for **YIH1***
Molecular Function Manually curated	• molecular_function unknown (ND)
Biological Process Manually curated	• regulation of cellular amino acid metabolic process (IMP, IPI)
Cellular Component High-throughput	• cytoplasm (IDA) • nucleus (IDA)
Mutant Phenotype	*All **YIH1** Phenotype details and references*

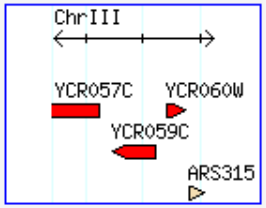
YIH1 RESOURCES

Click on map for expanded view

SGD ORF map



GBrowse



- **Literature**
  Literature Guide — View
- **Retrieve Sequences**
  Genomic DNA — View
- **Sequence Analysis Tools**
  BLASTP — View
- **Protein Info & Structure**
  Protein Info — View
- **Localization Resources**
  YeastRC Localization (Seattle) — View
- **Interactions**
  BioGRID (Toronto) — View

5 kbp の範囲を ChrIII から表示、塩基番号 221,338 から 226,337

説明
配列の名前、遺伝子名、ゲノム上の位置、その他のランドマークなどを使って検索します。ワイルドカード文字として * を使うことができます。ルーラー上でクリックした位置が中心になります。スクロールとズームボタンを使って拡大率と位置を変更します。

例: chrIII, chrV:80,000..120,000, SGS1, RPL28, YCR065W, flocculation, CAATGATTACGGCATT.

[バナーを隠す] [この表示をブックマーク] [この画像へのリンク] [高品質SVG画像] [ヘルプ] [リセット]

検索

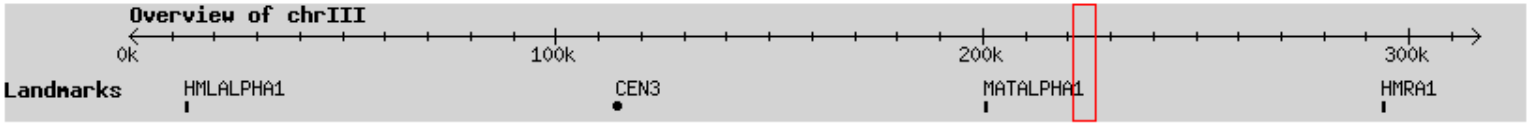
ランドマークまたは領域:
ChrIII:221338..226337 検索

データソース
S. cerevisiae from Saccharomyces Genome Database (SGD)

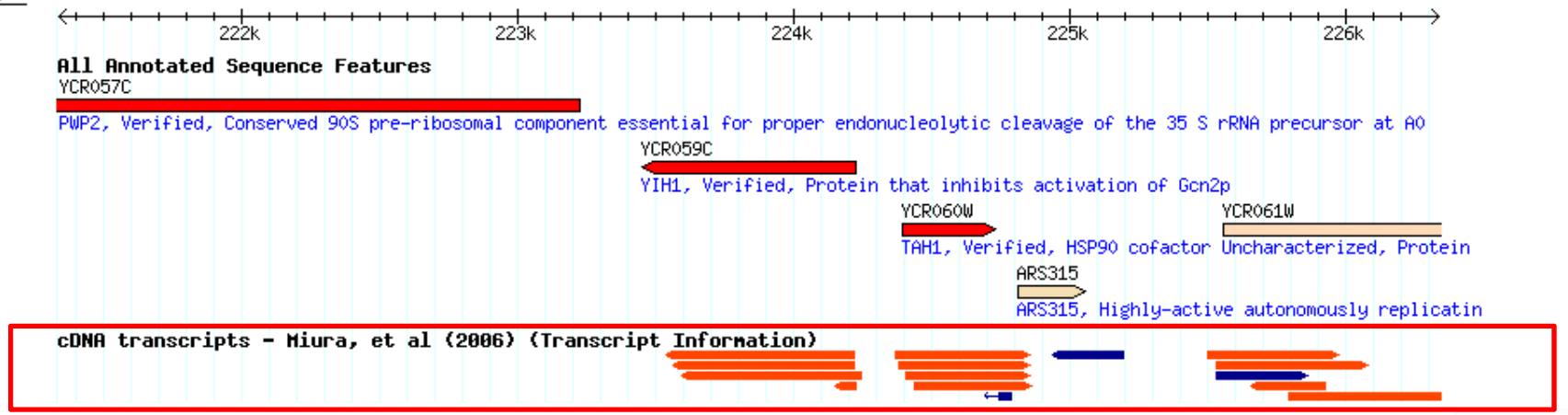
出力や解析などの操作:
解析 Restriction Sites 設定... 実行

スクロール/ズーム: 表示 5 kbp 反転

オーバービュー



詳細ビュー



強調表示を解除

画像を更新

Name:	**Y090_K19_F.ab1**
Type:	cDNA_clone
Source:	PMID:17101987
Position:	chrIII:223597..224245 (- strand)
Length:	649
Assignment:	single feature
Associated_feature:	YCR059C (5'-portion of the feature)
Average_quality_3'-end_sequence:	0
Average_quality_G-cap_sequence:	52
G-cap_judgement:	Perfect G-cap sequence
G-cap_sequence:	G
genomic_hit:	single genomic locus

>Y090_K19_F.ab1 class=Sequence position=chrIII:223597..224245 (- strand)

GAGAGGAAAG AAAAGCTCAA TGGATGACGA TCACGAACAG TTGGTCGAAG AACTGGAGGC CGTCGAGGCC ATCTATCCGG
ATCTTCTCTC CAAGAAGCAG GAAGACGGAA GCATCATCGT TGTGAAAGTG CCGCAGCATG AATACATGAC ACTGCAGATC
TCCTTCCCGA CACACTACCC CTCCGAGGAG GCTCCTAATG TCATCGAAGT TGGTGTCTGC ACTTCTTTGG CTAAGCGCGA
TCTCTACGAT ACCAAGTACC TTCAGCATTT GTTCCAGGAA GTGATGGACT CTGTTTTCCA CCGCGGATCT GTCTGTCTAT
TTGACTTCCT CACAGAACTC GACGGTGTCT TGTACGTTGA ACCAGAGGAG GAGACAGAAC CGGTCCAGCA GAGTGACATT
CCCACAGACC CCTTCGAGGG CTGGACCGCG TCGGACCCCA TTACTGATAG AGGCTCGACT TTCATGGCCT TTGCAGCACA
TGTTACCTCC GAGGAACAAG CGTTTGCCAT GCTAGACCTA CTGAAGACCG ACTCCAAGAT GCGTAAGGCA AACCATGTCA
TGAGTGCATG GCGAATCAAG CAGGATGGCT CTGCGGCAAC ATATCAAGAT TCCGATGATG ACGGTGAAAC GGCCGCCGGC
TCCAGAATG

SGD™ Database Copyright © 1997-2010 The Board of Trustees of Leland Stanford Junior University. Permission to use the information contained in this database was given by the researchers/institutes who contributed or published the information. Users of the database are solely responsible for compliance with any copyright restrictions, including those applying to the author abstracts. Documents from this server are provided "AS-IS" without any warranty, expressed or implied.

統合DBに求めたいもの

- 目につくこと
- 相乗効果を生む品揃え
- お墨付きが得られること

⇒　作物を置かせて貰いたくなるお店

- それを支える人材

⇒　優秀で親切な店員さん

© 2010 伊藤　隆司　(東京大学) licensed by CC表示2.1日本

生産者の懸念

- オーミクス論文で最も役に立つのはサプリ
  - DBにsubmitした生データよりも、解析済みデータの方が一般ユーザにはありがたい
- 消費者はあちこち廻ってサプリを落としてくる

© 2010 伊藤　隆司　(東京大学) licensed by CC表示2.1日本

Recommended
F1000 Factor 3.0
EndNote
Download citation
Send page by email

A large-scale full-length cDNA analysis to explore the budding yeast transcriptome.

Miura F, Kawaguchi N, Sese J, Toyoda A, Hattori M, Morishita S, Ito T

Proc Natl Acad Sci U S A 2006 Nov 21 **103**(47):17846-51 [abstract on PubMed] [citations on Google Scholar] [related articles] [FREE full text]

Selected by | Jonathan R Warner
Evaluated 5 Dec 2006
Relevant Sections

Faculty Comments & Author Responses

Faculty Member	Comments
Jonathan R Warner Albert Einstein College of Medicine, United States of America Cell Biology Tech Advance	**This article provides a substantial advance in our understanding of the transcriptome of Saccharomyces cerevisiae. This extensive analysis of full-length cDNAs derived fom yeast mRNAs provides a complete catalog of 5' untranslated regions (UTRs), identifies 45 new introns, rectifies several errors in open reading frame (ORF) assignments, and reveals a number of anti-sense transcripts.** The Supplemental Data files should be in the computer of any yeast researcher interested in mRNA and its regulation. **Competing interests:** None declared Evaluated 5 Dec 2006 How to cite this evaluation

Faculty Comments & Author Responses

How to cite the Faculty of 1000 Biology evaluation(s) for this paper

1) To cite all the evaluations for this article:

Faculty of 1000 Biology: evaluations for Miura F et al *Proc Natl Acad Sci U S A* 2006 Nov 21 103 (47) :17846-51 http://f1000biology.com/article/id/1051605/evaluation

2) To cite an evaluation by a specific Faculty member:

Jonathan R Warner: Faculty of 1000 Biology, 5 Dec 2006 http://f1000biology.com/article/id/1051605/evaluation

© 1999-2010 Faculty of 1000 Ltd unless otherwise stated < info@f1000.com > Terms and conditions Legal info

TEVA Discover the Teva Partners advantage for developing your pre-clinical compound
Now seeking pre-clinical compounds in Neurology, Oncology, and Auto-Immune / Inflammatory Diseases

Sign up for PNAS Online eTOC alerts
Get notified by e-mail when new content is posted online

Listen Discover Experience Science Sessions PNAS Podcasts

PNAS
Proceedings of the National Academy of Sciences of the United States of America

A large-scale full-length cDNA analysis to explore the budding yeast transcriptome

Fumihito Miura [*, †], Noriko Kawaguchi [*, †], Jun Sese [*], Atsushi Toyoda [‡], Masahira Hattori [*, ‡, §], Shinichi Morishita [*, †], and Takashi Ito [*, †, ¶]

+ Author Affiliations

Edited by David Botstein, Princeton University, Princeton, NJ, and approved October 3, 2006 (received for review July 7, 2006)

Abstract

We performed a large-scale cDNA analysis to explore the transcriptome of the budding yeast *Saccharomyces cerevisiae*. We sequenced two cDNA libraries, one from the cells exponentially growing in a minimal medium and the other from meiotic cells. Both libraries were generated by using a vector-capping method that allows the accurate mapping of transcription start sites (TSSs). Consequently, we identified 11,575 TSSs associated with 3,638 annotated genomic features, including 3,599 ORFs, to suggest that most yeast genes have two or more TSSs. In addition, we identified 45 previously undescribed introns, including those affecting current ORF annotations and those spliced alternatively. Furthermore, the analysis revealed 667 transcription units in the intergenic regions and transcripts derived from antisense strands of 367 known features. We also found that 348 ORFs carry TSSs in their 3′-halves to generate sense transcripts starting from inside the ORFs. These results indicate that the budding yeast transcriptome is considerably more complex than previously thought, and it shares many recently revealed characteristics with the transcriptomes of mammals and other higher eukaryotes. Thus, the genome-wide active transcription that generates

« Previous | Next Article »
Table of Contents

This Article

Published online before print November 13, 2006, doi: 10.1073/pnas.0605645103
PNAS November 21, 2006 vol. 103 no. 47 17846-17851

Abstract Free
Figures Only
» Full Text
Full Text (PDF)
Supporting Information

- Classifications

Biological Sciences
Genetics

- Services

Email this article to a colleague
Alert me when this article is cited
Alert me if a correction is posted
Similar articles in this journal
Similar Articles in Web of Science
Similar articles in PubMed
Add to My File Cabinet
Download to citation manager
Request copyright permission

- Citing Articles

View citing article information
Citing articles via CrossRef
Citing Articles via Web of Science (65)
Articles citing this article

Search PNAS GO
advanced search >>

This Week's Issue
September 28, 2010, 107 (39)

From the Cover

- Sustaining Chile's marine resources
- Liquid silicon phase transitions
- ATP synthase stoichiometry
- Reducing corneal inflammation
- Prenatal famine and adult cognition

Alert me to new issues of PNAS

Early Edition
Archives
Online Submission
Feature Articles
Commentaries

There's a big difference between visible

Anatrace

Sign up for PNAS Online eTOC alerts
Get notified by e-mail when new content is posted online

Listen Discover Experience
Science Sessions
PNAS Podcasts

PNAS
Proceedings of the National Academy of Sciences of the United States of America

A large-scale full-length cDNA analysis to explore the budding yeast transcriptome

Miura *et al.* 10.1073/pnas.0605645103.

Supporting Information

Files in this Data Supplement:

Supporting Figure 6
Supporting Figure 7
Supporting Figure 8
Supporting Figure 9
Supporting Table 1
Supporting Table 2
Supporting Table 3
Supporting Table 4
Supporting Table 5
Supporting Table 6
Supporting Table 7
Supporting Table 8
Supporting Table 9
Supporting Table 10

This Article

Published online before print November 13, 2006, doi: 10.1073/pnas.0605645103
PNAS November 21, 2006 vol. 103 no. 47 17846-17851

Abstract Free
Figures Only
Full Text
Full Text (PDF)
» Supporting Information

Search PNAS GO
advanced search >>

This Week's Issue
September 28, 2010, 107 (39)

From the Cover
■ Sustaining Chile's marine

Early Edition
Archives
Online Submission
Feature Articles
Commentaries

出版社がつぶれたらどうなるんだろう？

Supporting Figure 6

Fig. 6. Structure of 5′-end of *CDC19* cDNA clones. Genomic sequence of *CDC19* (chr1, 71702 to 71789) is shown on the top. The two most frequently-used TSSs and the initiation codon are underlined. Thirty different patterns of the 5′-ends for *CDC19* cDNA clones are shown with clone counts. Red and blue letters indicate 'typical' and 'atypical' G-cap sequences, respectively. Clones with typical

PubMed Central
Search | Journal List
This Article | Info for Authors | Subscribe | About
PNAS
Proceedings of the National Academy of Sciences of the United States of America

Proc Natl Acad Sci U S A. 2006 November 21; 103(47): 17846–17851. PMCID: PMC1693835
Published online 2006 November 13. doi: 10.1073/pnas.0605645103.

Copyright © 2006 by The National Academy of Sciences of the USA

Genetics

A large-scale full-length cDNA analysis to explore the budding yeast transcriptome

Fumihito Miura,[*†] Noriko Kawaguchi,[*†] Jun Sese,[*] Atsushi Toyoda,[‡] Masahira Hattori,[*‡§] Shinichi Morishita,[*†] and Takashi Ito[*†¶]

*Department of Computational Biology, Graduate School of Frontier Sciences, University of Tokyo, Kashiwa 277-8561, Japan;
†Institute for Bioinformatics Research and Development, Japan Science and Technology Agency, Tokyo 102-0081, Japan;
‡The Institute of Physical and Chemical Research (RIKEN) Genomic Sciences Center, RIKEN Yokohama Institute, Yokohama 230-0045, Japan; and
§Kitasato Institute for Life Sciences, Kitasato University, Tokyo 108-8641, Japan
¶To whom correspondence should be addressed. E-mail: ito@k.u-tokyo.ac.jp
Edited by David Botstein, Princeton University, Princeton, NJ, and approved October 3, 2006
Contributed by
Author contributions: F.M. and T.I. designed research; F.M. and N.K. performed research; J.S., A.T., M.H., and S.M. contributed new reagents/analytic tools; F.M., J.S., and T.I. analyzed data; and F.M. and T.I. wrote the paper.

Received July 7, 2006

This article has been cited by other articles in PMC.

ABSTRACT

We performed a large-scale cDNA analysis to explore the transcriptome of the budding yeast *Saccharomyces cerevisiae*. We sequenced two cDNA libraries, one from the cells exponentially growing in a minimal medium and the other from meiotic cells. Both libraries were generated by using a vector-capping method that allows the accurate mapping of transcription start sites (TSSs). Consequently, we identified

Other Sections

Formats: Abstract | Full Text | PDF (486K)

PubMed articles by these authors
- Miura, F.
- Kawaguchi, N.
- Sese, J.
- Ito, T.

PubMed related articles
- Review Unexpected complexity of the budding yeast transcriptome. [IUBMB Life. 2008]
- Reinvestigation of the Saccharomyces cerevisiae genome annotation by comparison to the genome [Genome Biol. 2003]
- The transcriptional landscape of the yeast genome defined by RNA sequencing. [Science. 2008]
- Genome-wide analysis of mRNA lengths in Saccharomyces cerevisiae. [Genome Biol. 2003]
- Review Transcriptome analyses of human genes and applications for proteome analys [Curr Protein Pept Sci. 2006]

» See reviews... | » See all...

Recent Activity
Turn Off Clear
A large-scale full-length cDNA analysis to explore the budding yeast transcriptome
A large-scale full-length cDNA analysis to explore the budding yeast transcriptome.
Miura F AND Ito T (8) PubMed
A comprehensive two-hybrid analysis to explore the yeast protein interactome
A comprehensive two-hybrid analysis to explore the

PNAS
PNAS

Data Availability.

All of our cDNA data can be browsed at http://yeast.utgenome.org and downloaded from http://itolab.cb.k.u-tokyo.ac.jp/GCap.

SUPPLEMENTARY MATERIAL

Supporting Information

Click here to view.

ACKNOWLEDGMENTS

We thank Seishi Kato and Satoshi Okada for advice on vector-capping clones and SK1 cells, respectively. This work was partly supported by a grant-in-aid for Scientific Research on Priority Areas from the Ministry of Education, Culture, Sports, Science, and Technology of Japan and a grant-in-aid for Scientific Research (B) from the Japanese Society of Promotion of the Science.

ABBREVIATIONS

TSS	transcription start site
SAGE	serial analysis of gene expression
uORF	upstream ORF
CUT	cryptic unstable transcript.

FOOTNOTES

The authors declare no conflict of interest.

This article is a PNAS direct submission.

Data deposition: The 5'-end sequence data for the 31,847 full-length cDNA clones have been deposited in the DNA Data Bank of Japan, www.ddbj.nig.ac.jp (DDBJ ID codes DB636784–DB668630).

森下研、伊藤研

雑誌のWebsite

DDBJ

© 2010 伊藤　隆司　(東京大学) licensed by CC表示2.1日本

生産者の懸念

- オーミクス論文で最も役に立つのはサプリ
  - DBにsubmitした生データよりも解析済みデータの方が一般ユーザにはありがたい
- 消費者はあちこち廻ってサプリを落としてくる
- サプリの散逸を防ぐ手立て
- サプリを使いやすくする手立て

© 2010 伊藤　隆司　(東京大学) licensed by CC表示2.1日本

将来のデータのために

- 論文概要
  - 主に普通の論文
  - アブストラクト以上、全文以下
- データ
  - 主にオーミクス系論文
  - サプリ
    - DBにsubmitしないタイプのデータ
    - 解析済みデータ

⇒　協力を得るにはストレスを感じさせない規制

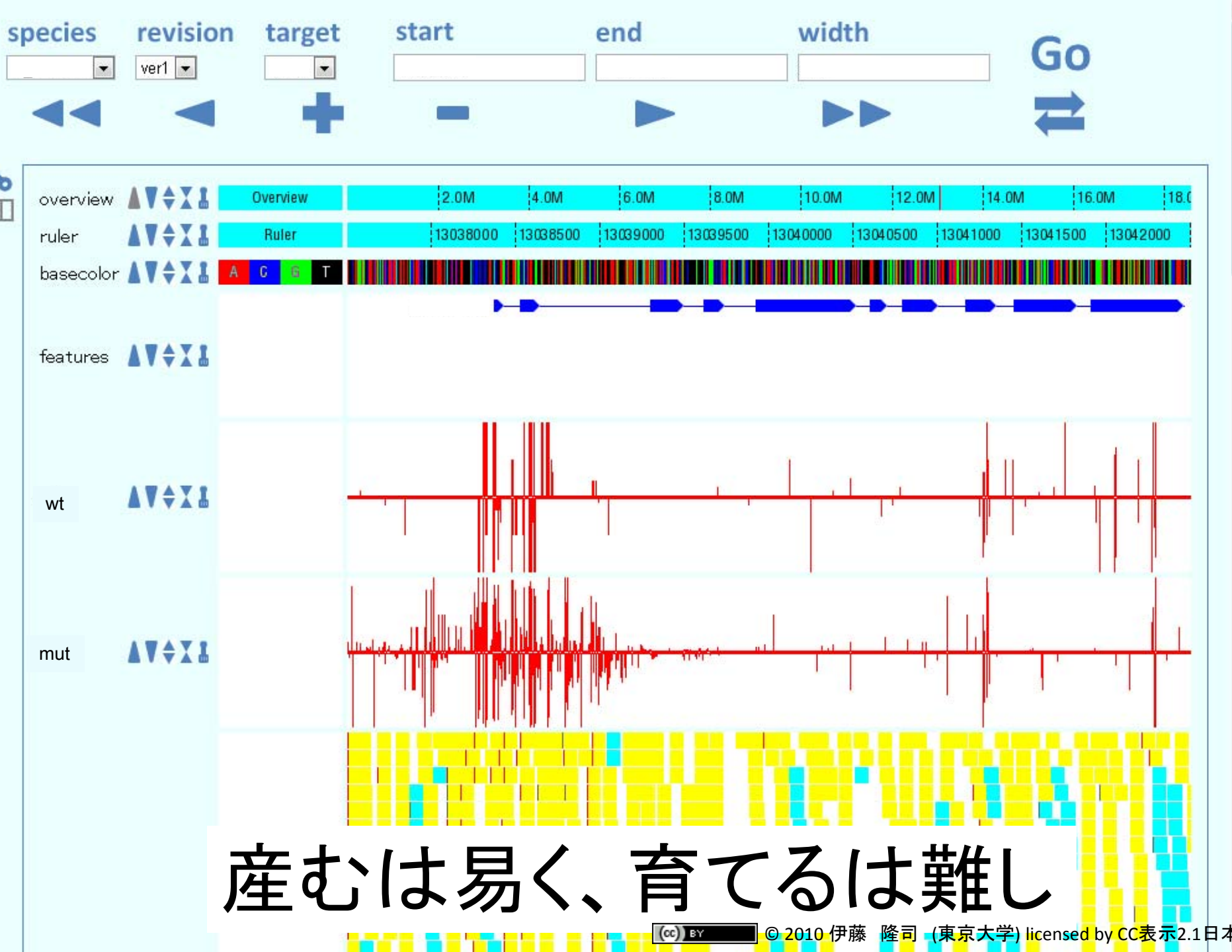
© 2010 伊藤　隆司　(東京大学) licensed by CC表示2.1日本

最近‥

- メチロームデータの成長期
  - 全ゲノムバイサルファイトシーケンス
  - >1G reads

- データの解析が大変‥
  - どうやって世に出そうか
    - 言い訳としてのデータ公開（＝　最も再利用しにくい形でのデータ公開）にはしたくない

© 2010 伊藤　隆司　(東京大学) licensed by CC表示2.1日本



産むは易く、育てるは難し

© 2010 伊藤　隆司　(東京大学) licensed by CC表示2.1日本

生産者から見た統合DBへの期待

- ユーザの目にとまりやすい環境
- データの相乗効果が期待できる環境
- よい生産者の証となるブランド力

- データの永代供養
- データの子育て支援

- 必要な人材の育成

© 2010 伊藤　隆司　(東京大学) licensed by CC表示2.1日本