

User Manual Version 1.00

目次

1. 概要	2
1-1. SYGDとは	2
1-2. SYGDの特徴	2
2.SYGDチュートリアル	4
2-1. 解析画面共通の構成	4
2-2. 解析画面のトップページ(ゲノムリスト)の表示	5
2-3. ゲノム情報の閲覧	6
2-3-1. コンティグリストの表示	10
2-3-2. コンティグマップの表示	11
2-3-3. 遺伝子詳細情報の表示	14
2-3-3-1. Locus (Locus 情報の表示)	14
2-3-3-2. Transcript(転写産物情報の表示)	15
2-3-3-3. CDS(CDS情報の表示)	16
2-3-3-4. Otholog(オーソログ情報の表示)	18
2-3-4. シンテニーマップの表示	18
2-4. BLAST 検索	22
2-5. ドットプロット解析	23
2-6. 差分解析	24
2-7. 基本検索	25
2-7-1. コンティグ検索	25
2-7-2. 遺伝子、RNA、CDS 検索	27
2-7-3. オーソログ検索	33
2-7-4. COG、KOG、EC 検索	35

1. 概要

1-1. SYGD とは

SYGD(Sake Yeast Genome Database)は、ゲノム情報に関する膨大なデータや関連情報 を容易に閲覧、解析できる統合プラットフォームです。Web ブラウザを用いて SYGD にア クセスするだけで、迅速かつ簡便にデータを閲覧、解析することが可能です。

1-2. SYGD の特徴

SYGD が持つ機能には、以下のようなものがあります:

統計情報表示機能

総塩基数、コンティグ数、遺伝子数など、様々な統計データを表示します。この機能の詳 細は、以下で説明します:

⇒ 2-2 ゲノムリスト (解析画面のトップページ)の表示

⇒ 2-3 ゲノム情報の閲覧

ゲノム情報の表示機能(ゲノムビューワー、アノテーションビューワー)

コンティグや染色体上に遺伝子地図を表示します。またそれぞれの遺伝子に対して、 Locus情報、DNA配列、RNA配列、アミノ酸配列、エキソン・イントロン情報、EC number、オーソログ情報などの多彩な情報を表示します。 また配列情報はFASTA形式ファイルとして出力され、解析結果は Microsoft Excel 等の外 部ソフトウェアに取り込んで解析を進めることも可能となります。 この機能の詳細は、以下で説明します:

⇒ 2-3 ゲノム情報の閲覧

BLAST 検索機能

任意の配列と類似性のあるものが存在するかを調べるためにBLAST検索機能を用いることができます。この機能の詳細は、以下で説明します:

⇒ 2-4 BLAST 検索

ドットプロット解析機能

ゲノム、染色体、あるいはコンティグ間の類似性を全体的にみることができます。この機能の詳細は、以下で説明します:

⇒2-5 ドットプロット解析

差分解析機能

複数のゲノムグループ間で、共通に存在する遺伝子、あるいは特定のグループにのみに存 在する遺伝子を表示します。

2グループ間解析に加えて、3グループ間解析も実施することができます。この機能の詳細は、以下で説明します:

⇒ 2-6 差分解析

基本検索機能

コンティグ、遺伝子、そしてオーソログなどの様々な観点からデータを高速に検索するこ とができます。

基本検索機能では、大きく分けて以下の検索を行なうことができます:

- 1. コンティグ検索 (解説: 2-7-1 コンティグ検索)
- 2. 遺伝子, RNA, CDS検索 (解説: 2-7-2 遺伝子, RNA, CDS検索)
- 3. オーソログ検索 (解説: 2-7-3 オーソログ検索)
- 4. COG, KOG, EC 検索 (解説: 2-7-4 COG, KOG, EC 検索)

2. SYGD チュートリアル

2-1. 解析画面共通の構成

SYGD では各画面で共通のインターフェースを採用しています。解析メニューの選択や簡 易遺伝子検索が画面上部の共通インターフェースから行えます。

Comprehensive Sake Yeast Genome Database					F	Тор Мали	ual
SYGD	В	С	D	E	Search	Q Genome	\$
Search -	Genome -	BLAST	Comparative Genon	ne - Category	1- 1-		

図 1: 各解析画面上部の共通部分

A	基本検索	基本検索画面へのリンクになっています。 詳細は、2-7.節 基本検索 で解説します。
В	ゲノム選択	「2-3. ゲノム情報の閲覧」と同様に、ここからゲノムを選択するこ とができます。
С	BLAST 検索リンク	BLAST 検索画面へのリンクになっています。 詳細は、2-4.節 BLAST 検索 で解説します。
D	ドットプロット解析 差分解析リンク	ドットプロット解析画面・差分解析画面へのリンクになっています。 ドットプロット解析の詳細は、2-5.節 ドットプロット解析 で解説し ます。 差分解析の詳細は、2-6.節差分解析解析 で解説します
Е	Category Browser	各カテゴリーに分類されている遺伝子の一覧表を表示することがで きます
F	簡易 Full text サーチ	基本検索のgeneでのFull textサーチと同様の機能となっています
G	マニュアル参照	マニュアル(PDF形式)が表示されます。

2-2. 解析画面のトープページ(ゲノムリスト)の表示

所定のURLにブラウザでアクセスすると、以下の解析画面のトップページが表示されます (図2)。トップページでは、登録されているゲノムのリストが表示されます。

メインメニュー →	Search - Genome - BLAST Comp	arative Gen	ome-Cat	egory -						
	Genome List in Sake Yeast Gen	ome Dat	tabase							
	Organism	Acronym	Source	Taxon. id	Mb	Chrmosome	Contigs	CDSs	tRNAs	rRNAs
ゲノムリスト ―→	Saccharomyces cerevisiae S288c	S288c	SGD	559292	12.1	16	17	6008	299	14
	Saccharomyces cerevisiae Kyokai no. 7	K7	NRIB_SYGD	721032	11.8	16	17	5816	296	7
	All	-	-		-	32	34	11824	595	21

図 2: トップページ画面。ゲノムリストが表示されます。

図2のゲノムリストについて以下に説明します:

Organism	SYGD に登録されているゲノムの名称を示します。 この名称は、主にゲノムの分類学上の種 (species)・サンプル名・解析名などで示されま
Acronym	ッ。 ゲノムの略称を示します。 またゲノムトップ画面 (図 3) へのリンクにもなっています。
Source	ゲノム配列のデータソースを示しています。
Taxon. id	登録されているゲノムのEntrez Taxonomy database上でのtaxonomy IDを示します。 また NCBI taxonomy browser へのリンクにもなっています。
Mb	"Organism"で示された生物種のゲノムサイズを Mega base pair 単位で示します。
Chromosome	"Organism"で示された生物種のゲノムの染色体数を示します。
Contig	それぞれのゲノムに登録されているコンティグ数を示します。 またコンティグ検索結果表示画面へのリンクにもなっています (2-7-1 節 or 図 23 参 照)。
CDSs	それぞれのゲノムに登録されている CDS 数を示します。 また CDS 検索結果表示画面へのリンクにもなっています (2-7-2 節 or 図 27 参照)。
tRNAs	それぞれのぞれのゲノムに登録されている tRNA 数を示します。 また RNA 検索結果表示画面へのリンクにもなっています (2-7-2 節 or 図 26 参照)。
rRNAs	それぞれのゲノムに登録されている rRNA 数を示します。 また RNA 検索結果表示画面へのリンクにもなっています (2-7-2 節 or 図 26 参照)。

2-3. ゲノム情報の閲覧

トップ画面ゲノムリストの "Acronym" 欄 (図2) のリンクをクリックしますと、ゲノムト ップ画面が表示されます (図3)。

ゲノムトップ画面ではSYGDに登録しているゲノムのうち、クリックした "Acronym" に対応するゲノムについて、より詳細な情報が表示されます。

Chr. I	Text Search To	Genome size :12.1Mb Chromosomes:16
Chr. II Chr. II Chr. V Chr. V Chr. VI Chr. VI Chr. VI Chr. VI Chr. X	Full text	Chromosomes:16 Contigs :17 CDSs :6008 tRNAs :299 rRNAs :14 Summay statistics
Chr. XI Chr. XII Chr. XIV Chr. XIV Chr. XV Chr. XV Chr. XVI	C	

図 3: ゲノムトップ画面.

ゲノムトップ画面は以下の 4 つで構成されています。

- A: 選択したゲノムの名称などの情報を示します。
- B: 選択したゲノムまたはコンティグの情報を図で示します。
- C: 検索画面。選択したゲノムに対する様々な検索を行います。
- D: 選択したゲノムに関する統計情報 (ゲノムサイズ、遺伝子数など) を示します。

A. ゲノムの名称

選択したゲノムの名称、略称 (Acronym)、taxonomy IDなどの情報を示します。

"tax_id"はEntrez Taxonomy database上でのtaxonomy ID を示し、NCBI taxonomy browserへのリンクにもなっています。

B. 染色体情報

選択されたゲノムに染色体に関する情報がある場合には、染色体情報が示されます。 染色体毎に、2つのリンクがあります。Chr.番号のリンクをクリックすると、この染色体 に含まれるコンティグを示すコンティグリスト画面を見ることができます。Chr.番号の下 のバーをクリックすると、クリックした位置を中心としたコンティグマップ画面が現れま す。

染色体に関する情報が無い場合にはコンティグ情報が表示されます。この場合には、コン ティグ番号をクリックすると、コンティグ全体のコンティグマップが、下のバーをクリッ クするとクリックした位置を中心としたコンティグマップが現れます。

C. 検索画面

選択されたゲノムに対する検索インターフェースを提供いたします。

より厳密な検索や横断的な検索を実施したい場合には、BLAST検索(2-4参照)や基本検索 (2-7参照) をご利用してください。

検索インターフェースには Text Search, BLAST Search, Category Browser の3 つがあり ます:

• Text Search

選択されたゲノムに存在する遺伝子に対して、テキスト検索を行います。 検索結果の解説については、2-7-2 遺伝子、RNA、CDS検索 を参照してください。 検索フィールドとして、以下の 7 種類があります:

Full Text	登録されている遺伝子に対して全文検索を行います。
Gene ID	登録されている遺伝子に対して Gene ID で検索を行います。
Description	登録されている遺伝子に対してDescription情報で検索を行います。例えば、
	"enzyme" と入力すると、gene description 欄に "enzyme" という文字列が含まれ
	る遺伝子を検索します。
COG ID	登録されている遺伝子に対して COG (オーソログ) 情報の ID で検索を行います。
KOG ID	登録されている遺伝子に対して KOG (オーソログ) 情報の ID で検索を行います。
EC Number	登録されている遺伝子に対して EC Number 情報で検索を行います。
PRINTS	登録されている遺伝子に対して PRINTS (タンパク質モチーフ) 情報で検索を行いま
	す。

BLAST Search

FASTA形式の配列を用いて、選択されたゲノムに対してBLAST による相同性解析を行ないます。

検索結果の解説については、2-4 BLAST検索を参照してください。

"Select BLAST DB"のセレクトボックスから任意のデータセットを選択し、"FASTA Sequence"欄へFASTA形式の配列を入力して "BLAST" ボタンをクリックすることで BLAST検索が行なえます。 セレクトボックスから選択できるデータセットは、以下の通りです:

Genome	全コンティグ配列
All Genes	全遺伝子配列
All Transcripts	全転写産物 (RNA) 配列
All CDS	全 CDS 核酸配列
All Proteins	全アミノ酸配列
Intergenic Regions	遺伝子間領域
Upstream 2kb from all genes	全遺伝子の上流側 2kb の配列

Category Browser

遺伝子には、COG, PRINTS, EC number 等でアノテーションされているものがあります。

これらでは "RNA processing and modification", "Transcription" 等のカテゴリーで遺伝 子を分類しています。Category Browser では、これらカテゴリーの階層を表示すると ともに、各カテゴリーに分類されている遺伝子の一覧表を表示することができます (図 4)。

(a) Top / S288c / EC Category

Top / S288c / EC Category



(b)

	1 Oxidoreductases 288
	■1.1 to EC 1.3 108
	1.2 Acting on the aldehyde or oxo group of donors 47
	1.3 Acting on the CH-CH group of donors 59
	1.4 to EC 1.97 11
	1.5 Acting on the CH-NH group of donors 30
	1.6 Acting on NADH or NADPH 28
	1.7 Acting on other nitrogenous compounds as donors 8
	1.8 Acting on a sulfur group of donors 24
	1.9 Acting on a heme group of donors 12
	1.10 Acting on diphenols and related substances as donors 7
	1.11 Acting on a peroxide as acceptor 12
	1.12 Acting on hydrogen as donor 1
	1.13 Acting on single donors with incorporation of molecular oxygen (oxygenases)
	1.14 Acting on paired donors, with incorporation or reduction of molecular oxygen a
	1.16 Oxidizing metal ions 16
	1.17 Acting on CH or CH2 groups 9
	1.18 Acting on iron-sulfur proteins as donors 5
	1.20 Acting on phosphorus or arsenic in donors 2
	2 Transferases / 94
	4 Lyases 129
D	6 Ligases 98
	S Isomerases 75
	3 Hydrolases 644

Select T	arget: Co	ontig Gene RNA CD	S Ortholog COG KOG	EC						
Search										
CDS S	earch	Show Filter Option								
CDS list 1 - 50 of 108 Rows: 50 0 Page: 1 6 Top Bottom Page: 1, 2, 3										
No./108	Gene ID	Original description	Assigned description	Genome	Chromosome	Links	Protein sequence			
1	YLR0 70C	D-xylulose reductase X YL2	Sorbitol dehydrogenas e	Saccharomy ces cerevisi ae S288c (S 288c)	Chr. XII	SGD:S 00000 4060 NCBI_ GeneI D:850 759	SYLR070CHP 013171.1INC_001144.5:274141-275 MTDLTTQANLER07K1TNWSIPMSDPWEVTQLATG SGIVALISENWATLKVODIWLEDIDORSFENKSGIVN KLPDDSFEEGALIEN_SVAHANKLAKIKKGAKWKE TAQTGATHVISGDJNGTUSSVKALIGKKGAVVFE IQFHSIIPTKELTQGCFNCQQDVSDSIELVSSKLSLKP			
2	YLR1 00W	3-keto-steroid reducta se	3-keto sterol reductas e	Saccharomy ces cerevisi ae S288c (S 288c)	Chr. XII	SGD:S 00000 4090 NCBI_ GeneI D:850 790	SVLR100WINP_013201_LINC_001144_SI4480-042 WINK0AUTGINSULGUNPELITEDTINURTUVTSRT VLLV0STTIMVSVLNAYDDIKKYRAINYLFYNAGGIFSOL KSODMGLIPONYGWYSESLUCQTRCKAYTWISSI LUHATYKOLKKUGINQYWQPGIFTSHSFSYLNFFTYFOL VTRLAWNFEKQOVKYGSATSKOGMYKTQEIDPTGMST			
3	YLR1 74W	isocitrate dehydrogena se (NADP(+)) IDP2	NADP-dependent isoci trate dehydrogenase	Saccharomy ces cerevisi ae S288c (S 288c)	Chr. XII	SGD:S 00000 4164 NCBI_ GeneI D:850	□YLR174WINP_013275.1 INC_001144.5: 504591-505 MTKIKVANPIVEMOGDEQTRIIWHLIRDKLUPYLDVLLKV VKCATITPDEARVEEFHLKKWWKSNOTTIRNLIGGTVFREI ATDVITVPEGELRLVYKSKSGTHDVDLKVFDVPEHGGVAM TKNTILKKYDGKFDVFEAMYARSYKEKFESLGIWFEHRLI VAQEFOSLGLMTSVLITPOGKTFESEAAHGTVTRHFRQHC			

図 4: Category Browser. ここではEC categoryの例を示します。図3のEC categoryをクリックすると、EC number に関する最上層のカテゴリーリストが表示されます(a). EC number では、最初に6つのカテゴリー (OXidoreductases, Transferases, Hydolases, Lyases, Isomerases, Ligases)が示されています。各々のチェックボッ クスをチェックすると、より細かく分類されたカテゴリーが表示されます(b)。各々のカテゴリーの右側にあるリン ク付き数値は、そのカテゴリーに属するCDS数を示しています。リンクをクリックすると、そのカテゴリーの CDS の一覧表を示します(c)。

D. 統計情報

(C)

選択されたゲノムの統計情報が示されます。以下の情報が示されます:

- Genome size
- \circ Chromosomes
- Contigs
- CDSs
- ∘ tRNAs
- ∘ rRNAs

また Summary statistics をクリックすることで、このゲノムに対する染色体毎の統計情報 が示されます (図 5)。

Saccharomyces cerevisiae \$288c										
Sacci	Size[Mbo]	Inyces	Contine	rev	+RNIA	RNA	J200C			
Chr. I	0.2	230,218	1	94	4	-				
Chr. II	0.8	813,184	1	415	13	-				
Chr. III	0.3	316,620	1	168	10	-				
Chr. IV	1.5	1,531,933	1	767	28	-				
Chr. V	0.6	576,874	1	288	20	-				
Chr. VI	0.3	270,161	1	128	10	-	=			
Chr. VII	1.1	1,090,940	1	539	36	-	;			
Chr. VIII	0.6	562,643	1	290	11	-				
Chr. IX	0.4	439,888	1	215	10	-				
Chr. X	0.7	745,751	1	362	24	-	;			
Chr. XI	0.7	666,816	1	317	16	-				
Chr. XII	1.1	1,078,177	1	519	21	12				
Chr. XIII	0.9	924,431	1	469	21	-				
Chr. XIV	0.8	784,333	1	398	14	-				
Chr. XV	1.1	1,091,291	1	548	20	-				
Chr. XVI	0.9	948,066	1	472	17	-				
Mitochondria		85,779	1	19	24	2	0			
All	12.1	12,157,105	17	6008	299	14				

図 5: 統計情報。染色体毎のサイズ、コンティグ数、CDS 数などを表示。

2-3-1. コンディグリストの表示

図5のChr.列のリンクをクリックしますと、コンティグ配列のリストが表示されます。 真核生物で選択されたゲノムに染色体に関する情報がある場合、図3-BのChr番号のリンク をクリックしますと、1 つの染色体に属するコンティグ配列のリストが表示されます。表 示の例を図6 に示します。

Top / S288c / Contig list

Top / S288c / Statistics

Chr. I contig list

Contig	Strand	Position	Length[bp]	CDS	tRNA	rRNA
NC_001133.9	1	1-230218	230,218	94	4	-
All	-	-	230,218	94	4	-

図 6: コンティグリスト画面。"Contig"欄のリンクをクリックしますと、そのコンティグマップ (図7) が表示されます。

2-3-2. コンディグマップの表示

図6 Contigリンク部分、あるいは図3-BのChr.番号の下のバーをクリックしますと、図7の ようなコンティグマップ画面が表示されます。図は4つの領域: (1). 表示コントロールパネ ル (2) 染色体表示領域 (3). 表示範囲コントロール領域 (4) 遺伝子表示領域 分けることが できます。



図 7: コンティグマッブ画面. 一つのコンティグ配列について、染色体,遺伝子へのマップ情報を図にしたもの。

コンティグマップ画面は以下の 4 つで構成されています。

- A: 表示コントロールパネル
- B: 染色体表示領域
- C: 表示範囲コントロール領域
- D: 遺伝子表示領域

A: 表示コントロールパネル

このパネルのパラメータを変化させると、遺伝子表示領域のコンティグや遺伝子の表示を 変えることができます。

以下、パネル内のパラメータを解説します:

Focus area	右にある二つのボックスにコンティグ上始点位置、終点位置を入力すること
	で、表示範囲を指定できます。また、Focus area下にあるExport Sequenceをク
	リックすると、指定した範囲のコンティグ配列 (FASTA形式) がダウンロード
	できます。

Strand	矢印の向きを逆向きにすると、ゲノム配列を逆向き (相補的な配列) とした図 が再描画されます。
bp / pixel	1 ピクセル毎のbase pair数を変えることができます。例えばこの値を半分に すると、図は2倍に拡大されます。
Gene name	遺伝子表示領域に表示される遺伝子名の表記方法を変えます。
100%, 66%, 20%, 5%	現在のFocus areaを、遺伝子表示領域図全体の何%の割合で表示するかを決めます。 例えば66%とすると、遺伝子表示領域図全体の約2/3がFocus areaとなります。
Rows	遺伝子表示領域の段数を指定します。遺伝子表示領域では、存在する遺伝子を 正、負それぞれの方向に対して1行で表示していますが、遺伝子密度が高い場 合など見にくい場合があります。そのような場合にこの値を2以上にすること で、遺伝子を2行以上で表示することができます。
Width	画像の大きさ(幅)を指定します。この値を倍にすると、画像の幅が倍になります (表示する領域は変わらない)。
Gene Color	遺伝子表示領域に表示される遺伝子の色付け基準を切り替える。
New focus area	表示コントロール領域上にマウスを乗せると、コンティグ上の座標が表示されます。また、その右側にある > をクリックすると、Focus areaが 1/2 だけ右方向に移動すします。

B: 染色体表示領域

選択されているコンティグに対応する染色体が定義されている場合、染色体番号と共にオ レンジ色の線で示されます。

さらに、染色体に対する選択されたコンティグの部位がわかっていれば、その部位と方向 が赤い線で示されます。

現在選択されていないコンティグ配列は、赤ではなく灰色の線で示されます。

染色体、コンティグを示す線はリンクになっており、クリックすることでクリックした場 所周辺の表示へ移行します。

C: 表示範囲コントロール領域

この領域でマウスクリックを2回(始点,終点)を行うと、Focus areaが更新されます。指定 領域は上下二つあり、上の領域は大まかなFocus area指定に、下の領域はより細かな Focus area指定のために用いられます(図8)。

Focus area	28086	1804	90			02000)				_	_				
Export S	equence			Strand:	> 😳		bp/p	xel: 203.20	2 x2	/2		Gene	name:	Gene ID	
Zoom: 1	00%	66%	20% 5%	Rows: 1	0		Widt	1000	0			Gene	color:	KOG cat	egory ᅌ
New focus	area:900	64		Move:	~ •	: > >>									
			_			Chr. I									
				(Contig : N	4C_001133.9 (2	30218	op)							
									_						_
			50000	Ż	Ś	100000			15020	0				2000	00
					4										
)										
1		100911				**			2.5				: 28.82	8 8	
	-	2 2 2 2 2 2 2 2				22 E	11	222 222		11 1			1 11 11 1	11	1
1	*					ii f	╧╧	1 1111 111		H I	11		• • • • • •		=
¥ 11	11	1	r (1)1005	21 11	1 1 11 11		1	11 11 11	1 1	11	Πī) () ()	1	1	
			11101	11 11						11					
3 33	2.2	;	I AR SHALL	8.8 2.8	2 2 2 3			12 I I I		11	111	1111	8	1	



図8: 表示コントロール領域マウス操作による Focus area の変更。コントロール領域に対して、マウスを (a), (b) の順でクリックすると、その領域を新たに Focus area とした新しい図が描画されます (c)。「New focus area」 (赤枠)にはマウスのある位置が表示されます。

D: 遺伝子表示領域

この領域には、Focus area内にある遺伝子が示されています。遺伝子名、向きが表示され、表示コントロールパネル内で指定されるGene Colorに基づいた遺伝子機能による色分けがなされています。また、各々の遺伝子はリンクとなっており、ここをクリックすると遺伝子詳細画面(図9または 2-3-3参照)へ移動します。

2-3-3. 遺伝子詳細情報の表示

コンティグマップにおいて遺伝子表示領域にある遺伝子をクリックしますと、その遺伝子 についての詳細情報を見ることができます(図9)。遺伝子詳細情報はタブによって区切られ ています:

Locus(図9): 遺伝子名、塩基配列などのアノテーションの表示、コメントを編集
 Transcript(図10):転写産物 (RNA)の情報を表示
 CDS(図11):タンパク質コーディング領域 (CDS)の情報を表示
 Ortholog(図13):オーソログ (パラログ)情報を表示

2-3-3-1. Locus (Locus 情報の表示)

Locus タブ(図9)には遺伝子名、塩基配列などの各遺伝子のアノテーション情報が表示されます。Locus タブ内の情報は、A:基本情報表、B:コンティグマップ、C:遺伝子配列の3つに分かれています。



図9: 遺伝子詳細画面 (Locus). 選択された遺伝子に関する詳細情報。Aには基本情報が表で示されています。Bに は選択遺伝子を中心とした簡易的なコンティグマップ画面、Cには遺伝子配列が FASTA 形式で書き出されていま す。 A: 基本情報表

Gene ID	遺伝子 ID を示します。NCBI 由来のゲノムデータの場合は、locus tag 名を示します。
Original symbol	データソースに記載されていた遺伝子名を示します。
Assigned symbol	SYGD データ登録時に再アノテーションされた遺伝子名を示します。
Original description	データソースに記載されていた遺伝子に対する説明を示します。
Assigned description	SYGD データ登録時に再アノテーションされた遺伝子に対する説明を示し
	ます。
Length	遺伝子の塩基長(bp)を示します。
GC%	遺伝子の GC 割合を%単位で示します。
Organism	種名 (菌株名) を示します。
Genome	ゲノム名を示します。ゲノムトップ画面(図3)へのリンクにもなっています。
Chromosome	染色体情報のある場合に、染色体番号を示します(染色体情報が無い場合に
	は"Unmapped"と示されます)。
Contig Location	コンティグ名を示します。コンティグ名のリンクをクリックすると、この
-	遺伝子を中心としたコンティグマップ画面 へ移動します。
Links	外部データベースへのリンクです。

B: コンティグマップ (Contig Map)

選択された遺伝子を中心としたコンティグマップが表示されます。

C: 遺伝子配列(Gene Sequence)

選択された遺伝子の塩基配列が FASTA 形式で表示されます。 また遺伝子の上流と下流の塩基配列も表示できます。"with upstresm", "with downstream"のテキストボックスに整数(半角数字)を入力し"Apply"ボタンをクリッ クすると、入力した値の長さの配列が表示されます。

2-3-3-2. Transcript (Transcript 情報の表示)

Transcript タブ (図10) には転写産物 (RNA) の情報が表示されます。

Transcript タブ内の情報は、(1) 基本情報表、(2) RNA 配列情報の二つに分かれています。

	Pyruvate carboxylase	
	Locus Transcript CDS	Ortholog
ſ	Gene ID	YBR218C
	Туре	mRNA
	Assigned Description	Pyruvate carboxylase
	Length	3543 bp
-	Original Description	pyruvate carboxylase 2 Pyruvate carboxylase isoform; cytoplasmic enzymethat converts pyruvate to oxaloacetate; differentiallyregulated than isoform Pyc1p; mutations in the humanhomolog are associated with lactic acidosis; PYC2 has aparalog, PYC1, that arose from the whole genomeduplication
	Exon Start-End	1-3543
L	Links	NCBI_GeneID:852519
Г	RNA Sequence	
	РИК_001178566-11 Jength 1343 Аталаская такосамааматтооссоя Соорсосых такахото соорсосых такахото соорсосых соорсосых такахото соорсосых такахото самаясоватите соорсосых такахото самаясоватите соорсосых такосто обращее соорсосых такосто мого такахото соорсосых такосто такахото соорсосых такосто такахото соорсосых такосто соорсосых такосто такосто соорсосых такосто соорсосых такосто соорсосых такосто такосто соорсосых такосто соорсосых такосто такосто соорсосых такосто соорсосых такосто соорсосых такосто такосто соорсосых такосто соорсосых такосто соорсосых такосто соорсосых такосто такосто соорсосых такосто с	CTARGERIANTTEASTTRETCOGOGUNAMAGNIAN STRAIGNITTTRANSCOGCTOCOGUNAMAGNIAN TRAIGNITTTRANSCOGCTORGANOCOGCTAGUNA OCTINGTOGOLANGTAGUNAGOGUCAGUNAGUNA TECNICOGUNAGUNAGUNAGUNAGUNAGUNAGUNAGUNA TECNICOGUNAGUNAGUNAGUNAGUNAGUNAGUNAGUNAGUNAGUNA

図 10: 遺伝子詳細画面 (Transcript)

A:	基本情報表
----	-------

Gene ID	遺伝子 ID を示します。NCBI 由来のゲノムデータの場合は、locus tag 名を示します。
Assigned description	SYGD データ登録時に再アノテーションされた遺伝子に対する説明を示します。
Туре	RNA のタイプを示します。mRNA, tRNA,rRNA の3種類があります。
Length	遺伝子の塩基長 (bp) を示します。
Original description	データソースに記載されていた遺伝子に対する説明を示します。
Exon Start-End	遺伝子配列中、Exon の (一般には複数の) 始点-終点を示します。
Links	外部データベースへのリンクです。

B: RNA 配列情報(RNA Sequence)

選択された遺伝子の RNA 配列が FASTA 形式で表示されます。

2-3-3-3. CDS (CDS 情報の表示)

CDSタブ(図11)にはタンパク質コーディング領域(CDS)の情報が表示されます。 CDS タブ内の情報は、(1)基本情報表、(2)CDS 配列に分かれます。



図 11: 遺伝子詳細画面 (CDS).

A: 基本情報表

データソースに記載されていた遺伝子名を示します。 Original symbol

Assigned symbol	SYGD データ登録時に再アノテーションされた遺伝子名を示します。
Original description	データソースに記載されていた遺伝子に対する説明を示します。
Assigned description	SYGD データ登録時に再アノテーションされた遺伝子に対する説 明を示します。
Length	遺伝子の塩基長(bp)を示します。
Links	外部データベースへのリンクです。
similarity to KOG	KOG 配列との類似性があった場合、KOG ID とその説明, BLAST 検索によるアラインメントの e-value が示されます。e-value 値 はリンクとなっており、ここをクリックするとアラインメントの詳 細を見ることができます。
similarity to COG	上と同様。COG 配列との類似性を示します。
similarity to SwissProt	上と同様 SwissProt 配列との類似性を示します。
EC number	遺伝子に EC number 情報があった場合、EC number を表示します。
Pfam	Pfamドメインとの類似性があれば ID とその説明を示します。 "map" をクリックすると、類似性が認められる部分をグラフィカル に表示します
PRINTS	上と同様。PRINTS ドメインとの類似性を示します。 "map" をクリックすると、類似性が認められる部分をグラフィカ ルに表示します。
Prosite	上と同様。Prosite ドメインとの類似性を示します。 "map" をクリックすると、類似性が認められる部分をグラフィカ ルに表示します。
TMs	膜貫通領域予測計算結果を示します。 "map" をクリックすると、膜貫通領域と予測される部分をグラフ ィカルに表示します。
Signal peptides	signal peptide 予測計算結果を示します。
GPI-anchor	GPI アンカーとの結合が予測された部分を示します。

B: CDS 配列(Protein Sequence)

選択された遺伝子のアミノ酸配列が FASTA 形式で書かれます。

CDS タブ (図11) の最上部には、"Annotation Map" ボタンがあります。これをクリック すると、アノテーション情報の詳細が得られます (図12)。



図12: Annotation Map. CDS アミノ酸配列(上)とアノテーション情報(下).下部アノテーション情報では、Pfam, PRINTS等との類似性がグラフィカルに示されています。マップの対応する部分をマウスでクリックすると、上部 アミノ酸配列中で該当する部分の色が変化します。SwissProt等、アノテーション計算の際の配列類似性計算に BLAST検索を実行しているものに対しては、BLAST検索結果を別ウインドウに表示します。

2-3-3-4. Ortholog (Ortholog 情報の表示)

Ortholog タブ(図 13)では、選択された遺伝子についてのオーソログ(パラログ)情報を示します。

2 [Saccharomyces cerevisiae S	288c]			
ryruvate carboxylase	g			
Genome	Genes	Gene ID	Original Description	Links
	_	YGL062W	pyruvate carboxylase 1	NCBI_GeneID:852818 SGD:S000003030
Saccharomyces cerevisiae S288c	2	YBR218C	pyruvate carboxylase 2	NCBI_GeneID:852519 SGD:S000000422
		SYK7_005261	K7_Pyc2p	
Construction constitution Kitchel and R	-	_		

Orthologタブでは、ゲノムリスト(図2)のゲノムにある遺伝子配列との類似性を検索し、種を超えて相同性のある遺伝子についての一覧表が表示されます。

オーソログ計算には、類似性閾値などの違いなどいくつか異なる計算方法があります。計 算方法は、左上セレクトメニューで選択できます。

また左上のセレクトメニューの下にあるリンクをクリックすると、このオーソログを用いたシンテニー解析画面を表示します(図14)。

2-3-4. シンテニーマップの表示

SYGD ではオーソログ (パラログ) の関係のある遺伝子群について、シンテニー解析を行なった結果を表示する機能があります。

図 13 オーソログのリンクをクリックすると、シンテニーマップ画面を見ることができます (図 14)。

図 13: 遺伝子詳細表示(Ortholog)

1 [1288c] Saccharomyces convolutes 5286c TBR218C (PrC2) NC,001194.858707-662249 Deleter 2 [1288c] Saccharomyces convolutes 5286c TCL.6C2W (PrC1) NC,001139.9383196->388732 Deleter 3 [177] Saccharomyces convolutes 5286c TCL.6C2W (PrC1) NC,001139.9383196->388732 Deleter 4 [177] Saccharomyces convolutes 5286c TCL.6C2W (PrC1) D0000038.11636499-c 640001 Deleter 4 [177] Saccharomyces convolutes 5286c TSK7_0052561 (K7_PTC2) D000004.11638499-c 640001 Deleter Redraw Not x2 /2 /4 Gene color; K0C category (K) D000004.11389278->382814 Deleter Redraw TSK7_025561 (K7_PTC1) D000004.11389278->382814 Deleter Deleter Statistic dool (K) and (1 [528c] Saccharomyces cerevise \$286c YB218C (PrC2) NC_00134.86.58707 ~-662249 Detext 2 [528bc] Saccharomyces cerevise \$286c YG.06219 (PrC1) NC_00139.9.385196 ~-388712 Detext 3 [77] Saccharomyces cerevise \$286c YG.062261 (K7_PrC2) D0000038.1634594 ~-460001 Detext 4 [67] Saccharomyces cerevise \$296c SK7_052561 (K7_PrC1) D0000043.1389278 ~-598214 Detext Model 600 @ bp/pixet; 30 x4 x2 ?2 //4 Gene color; MCG category © Gene name; Gene D 9 9/117 />10 Mdt Sectory TYB218C Gene color; MCG category © Gene name; Gene D 9/117 />10 Mdt Gene D YB218C Gene color; MCG category © Gene name; Gene D 9/117 />10 Mdt Adgreed Description YM218C Gene color; MCG category © Gene name; Gene D 9/117 //10 Mdt Gene B YM218C YM220 YM2200 YM2200 YM2200 YM2200 YM2200 YM2200 YM2200 YM220
2 [15286] Saccharomyces cerevisiae S288c YGL062W (PrC1) NC_00113.9.9.385196388732 Detects 3 [17] Saccharomyces cerevisiae S288c YGL062W (PrC1) DC000008.1.636459640001 Detects 4 [17] Saccharomyces cerevisiae Kyolai no. 7 SYR7_005261 (K7_PrC2) DC000008.1.636459640001 Detects 4 [17] Saccharomyces cerevisiae Kyolai no. 7 SYR7_025361 (K7_PrC2) DC0000043.1.389278382814 Detects 8 [17] Saccharomyces cerevisiae Kyolai no. 7 SYR7_025361 (K7_PrC1) DC0000043.1.389278382814 Detects 1 [17] Saccharomyces cerevisiae Kyolai no. 7 SYR7_025361 (K7_PrC1) DC0000043.1.389278382814 Detects 1 [17] Saccharomyces cerevisiae State Systemy map Systemy map Systemy map 1 [17] Saccharomyces cerevisiae State Systemy map Systemy map System Systems YBR210C 1 [18] Saccharomyces cerevisiae State YBR210C YBR210C YBR210C YBR210C 1 [18] Saccharomyces cerevisiae State YBR210C YBR210C YBR210C YBR210C YBR210C YBR210C YBR210C	2 [1288c] Saccharomyces cervisiae S288c Y0L062W (PC1) NC_001139.9.385196388732 Detect 3 [1C7] Saccharomyces cervisiae S288c Y0L062W (PC1) NC_001139.9.385196388732 Detect 4 [1C7] Saccharomyces cervisiae Kyolai no. 7 SNC7_002561 (K7_PVC2) D000008.1:638459<-640001
3 [L7] Sacdaramyces cervisiae Kyolai no. 7 [VC7_005261 (K7_VPC2) DC000038.1:83449=-64001 Detec 4 [L7] Sacdaramyces cervisiae Kyolai no. 7 [VC7_052561 (K7_VPC1) DC000038.1:83449=-64001 Detec 4 [L7] Sacdaramyces cervisiae Kyolai no. 7 [VC7_022561 (K7_VPC1) DC000043.1:389278=-332814 Detec mdbt: 600 [VV7_VPC2] Gene color, K00 category Gene mane; Cerv ED V UVV IN Additional control of the color of the	3 [K7] Saccharomyces cerevisiae Kyokai no. 7 [SIX7_005261 (K7_yPC2) D6000058.1658459 Detect 4 [K7] Saccharomyces cerevisiae Kyokai no. 7 [SIX7_025561 (K7_yPC2) D6000043.1389278->392814 Detect A [K7] Saccharomyces cerevisiae Kyokai no. 7 [SIX7_025561 (K7_yPC1)] D6000043.1389278->392814 Detect Redraw Interview SIX7_025561 (K7_yPC1) D6000043.1389278->392814 Detect Redraw Six67_025561 (K7_yPC1) D6000043.1389278->392814 Detect Grees mino. Six67_025561 (K7_yPC1) D6000043.1389278->392814 Detect Grees mino. Six67_02562 Six67_000 Six67_000 Original Decorption Maxagend Description Six87_000 Six87_000 Grees mino. Six87_000 Six87_000 Six87_000 Six87_000 Grees mino. Six87_000 Six87_000 Six87_000 Six87_000 Six87_000 Grees mino. Six87_000 Six87_000 Six87_000 Six87_000 Six87_000 Grees mino. Six87_0000 Six87_0000 Six87_0000 Six87_0000 <t< td=""></t<>
4 [k7] Saccharomyces cerevisae Kyolan no. 7 [SXC_22561 (XC_PC1) DC000043.1:389278392814 Deiree Redrav Not x2 7.2 // 4 Gene color, KOC category () Gene name, Gene () 0.1 U U V H XXIII Redrav Not x2 7.2 // 4 Gene color, KOC category () Gene name, Gene () 0.1 U U V H XXIII Gene min0x. Systemy map Gene color, KOC category () Gene name, Gene () 0.1 U U V H XXIIII Gene min0x. Systemy map Gene color, KOC category () Gene name, Gene () 0.1 U U V H XXIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIIII	4 [k7] Saccharomyces cerevisiae Kyokai no. 7 [SK7_025961 (K7_P*C1) D6000043.1389278->392814 Deece Redraw Middle 600 2 paylaxet; 30 xk x2 /2 /4 Gene color; K0C category; Gene name; Gene ID U tr v h 3681 Middle 600 2 paylaxet; 30 xk x2 /2 /4 Gene color; K0C category; Gene name; Gene ID U tr v h 3681 Grees ID 1 Hit21 8C Systemy map Gene Color; K0C category; Gene Color; U tr v h 3681 Gregord 97C2 Gene Color; K0C category; Gene Color; Gene Color; Gene Color; Gene Color; K1220; YB22C; YB220; YB221;
Reform VIENDE 10 VIENDE 20 // 4 Serve coder KOC category Cene name: Cene ID // 11/2 / 14 Gene affor. Systemy map Systemy map Cene coder KOC category Cene name: Cene ID // 11/2 / 14 Cene coder KOC category Cene name: Cene ID // 11/2 / 14 Cene coder KOC category Cene name: Cene ID // 11/2 / 14 Cene coder KOC category VIENDE	Redaw Widzki 600 b byr/piedi 30 x4 x2 /2 /4 Gene eddari, KOC category Cene name: Gene D D U tr v h RKE Gene BD Y182/16C Greginal gregold Gregi
Syntaxy map Syntaxy map Original youndait PPC2 Assigned symbol PPC2 Opiginal youndait PPC2 Assigned symbol PPC2 Opiginal youndait Sprease Assigned Symbol PPC2 Statigned Symbol PPC1	Synterry map Cene: ID Y82118C Original synted IPrC2 Colspan="2">Colspan="2"Colspan="2">Colspan="2"Colspan=""2"Colspan=""2"Colspan=""2"Colspan=""2"Colspan=""2"Colspan=""2"Colspan=""2"Colspan=""2"Colspan=""2"Colspan=""2"Colspan=""2"Colspan=""2"Colspan=""2"Colspan=""2"Colspan=""2"Colspan=""2"Colspan=""2"Colspan=""2"Colspan=""2
Gene ID YRE218 C Original gymbol PPC2 Assigned bestryttering Provate carboxylase Staglard yymbol PPC2 Assigned bestryttering Provate carboxylase Langth 3343 pp (1180 a) Coh 43.05% Genome [1288:0] Saccharomyces cereviale S288c Original dymbol PrC1 Sasigned dymbol Sasigned dymbol Offsinal Discryttering Private carboxylase Sasigned Saryttering Sasigned Carboxylase Genome [5280:1] Saccharomyces cereviale S288c Domosome [5280:2] Saccharomyces cereviale S288c Demosome [5280:2] Saccharomyces cereviale S288c	Gene U YRR218 C Orginal symbol PYC2 Assigned symbol PYC2 Second Second Symbol Symbol Genome 1527401 (Stocharmyces cervisiae S288c VIR2201 VIR2201 VIR2201 VIR2201 VIR2201 VIR2201 VIR2201 VIR2201
Original grandidi mr.C.2 mr.C.2 Original Description mr.C.2 Original Description mr.C.2 Construct Mescription mr.C.2 Assigned Object (1800 as) mr.C.2 Construct Mescription mr.C.2 Gene Mescription mr.C.2 Massed Description mr.C.2 Mescription mr.C.2 Gene Mescription mr.C.2 Massed Symbol mr.C.2 Mescription mr.C.2	Original grandol Assigned Dysmich Assigned Dysmich Assigned Dysmich Assigned Description PrC2 60000 60000 Original Description PrC2 9822/2 9822/2 9822/2 9822/2 Assigned Operating Assigned Description 91/2 9822/2 9822/2 9822/2 9822/2 Cols 613.0% 1528/2 1528/2 9822/2 9822/2 9822/2 Cols 613.0% 1528/2 1528/2 9822/2 982/2
Assigned Symbol PrC2 60000 000000 00000 00000	Assigned Symbol PrC2 66000 60000 65000 Original Description Provide carbonylase 221 Y88221C Y88221C Y88221C Length 33-43 bp (1160 aa) 221 Y88221C Y88221C Y88221C Cols 43.30% Y88221C Y88221C Y88221C Y88221C Cols 152863 Sacharomyces cervisiae 5288C Y88221C Y88221C Y88221C
Original Description Image: Control of the control of th	Original Description Operating Section Assigned Description VIR.222 (VIR.222 (VIR
Assigned Description Provide control (162 dec) Provide	Assigned Description Pyruvate cantoolse Pyruv
Length 344 bp (1100 as) CON 43.00% Genome [12826] Saccharomyces cerevisas 5288c Otromosome Chr. 8 Genome Chr. 9 Genome Y16.20W Orginal Symbol Prc1 Sksigned Symbol Sksigned Symbol Script Prustic carbox/size Conse St280c) Saccharomyces cerevisae S286c Cenome [5280c] Saccharomyces cerevisae S286c Cenome St280c) Saccharomyces cerevisae S286c	Length 3341 bp (180 a) CCh 43.0% Cenome (16286) Stocharonyces cerevises 5288c Cenome (16286) Stocharonyces cerevises 5288c
CC6 43.30% YB223W A YB	COM 43.30% YBR22WA YBR22W
Genome [12280] Saccharomyces cereviae S288c Consign Nr. 1 Consign Nr. 200134.8 (558707 - 652249 bp.) BLAST Consign Nucl. 001394.8 (558707 - 652249 bp.) Consign symbol PrC1 Sees D YGL062W Original Symbol PrC1 Seegind Symbol Processee Consee [1228c] Saccharomyces cerevise S288c Consee [1228c] Saccharomyces cerevise S288c	Genome (S288c) Saccharomyces cerevisiae S288c
Okromasome Constag Ox. 8 BLAST Image: Constage BLAST Image: Constage Seaged symbol PrC1 Sassigned Symbol PrC2 Sassigned Symbol PrC2 Sassigned Symbol PrC2 <t< td=""><td>Channess Chan I</td></t<>	Channess Chan I
Contg VLC.001/8.8 (558707 - 652249 pp) BLAST	Chromosome Circ. II
BLAST	Contig NC_001134.8 [658707 - 662249 bp]
Cene ID VCL.062V Original grandb PVC1 Assigned symbol PVC1 Draynal Description 300,000 Assigned Symbol PVC1 Satigned Symbol 200,000 Satigned Description 200,000 COS 43.43% Comme C122801 Chromosome Chr. VB	BLAST
Original procisioned Incl Assigned symbol PrC1 Assigned symbol PrC1 Singland Description Processing Assigned Size (1178 a) Notew Cohe 43.4395 Cohese 12282(3) Saccharomyces cerevisiae 5288c Cohese Crivitic	Gene ID YCL062W
Assigned symbol PrC1 35000	Original symbol PYC1
Original Description 19000 300000 300000 30000	Assigned symbol PYC1
Assigned Description Provide control (see Contr	Original Description 38000 39000 390
Length 3537 bp (1178 ac) GC% 43,43% Genome [C2826] S5x7aromyces cerevisiae S288c Otromosome Chr. VB	Assigned Description Pyruvate carboxylase 2,066W YGL05W YGL05W YGL05W YGL05W YGL05W YGL05W
CCN 41.43% Genome [5285:] Sacharonyces cenvisas 5286. Genome [5285:] Charlanonyces cenvisas 5286.	Length 3537 bp (1178 aa)
Cenone (S228c; Saccharomyces cerevisue S288c Otranosone Otr. VI	GC% 43.43% YGL065C YGL064C YGL064C
Chromosome Chr. VII	Genome [S288c] Sacharomyces cerevisiae S288c
	Chromosome Chr. VII
Contig NC_001139.9 [385196 - 388732 bp]	Contig NC_001139.9 [385196 - 388732 bp]

図14:シンテニーマップ画面.オーソログ関係(同じ種であればパラログ関係)にある遺伝子群の情報を示します。 画面上部にオーソログとして選択された遺伝子の一覧表および、表示をコントロールするパネルがあります(図15)。 画面下部に、各遺伝子の基本情報(左, Genome info)と遺伝子マップ(右, Synteny map)を示します。左の遺伝子の 基本情報では、Gene IDやDescription などの情報、およびBLAST検索の結果が表示されます。右の遺伝子マップで はマップの中央に注目している遺伝子が配置され、その周辺の情報が表示されます。遺伝子マップは、"Width", "bp/pixel", "Gene color", Gane Name"のセレクトボックスで設定を変更することにより、表示を変更させることが できます。

シンテニーマップ (図14) の上には、オーソログ (パラログ) として選択された遺伝子の一 覧表があります (図15)。

これは、シンテニーマップの表示順序を変えたり、遺伝子を非表示にしたり (Delete) する コントロールパネルとなります。

Ortholog : ort1_c70_e100_142

No./4	Genome	Gene (Original symbol)	Location	Command	
1	[S288c] Saccharomyces cerevisiae S288c	YBR218C (PYC2)	NC_001134.8:658707<662249	Delete	
2	[S288c] Saccharomyces cerevisiae S288c	YGLO62W (PYC1)	NC_001139.9:385196>388732	Delete	
3	[K7] Saccharomyces cerevisiae Kyokai no. 7	SYK7_005261 (K7_PYC2)	DG000038.1:636459<640001	Delete	
4	[K7] Saccharomyces cerevisiae Kyokai no. 7	SYK7_025961 (K7_PYC1)	DG000043.1:389278>392814	Delete	

Redraw

図15: シンテニーマップのコントロールパネル。 Commandにあるボタンで表示を操作する。"Delete"ボタン を押 すと該当する遺伝子がシンテニーマップ表示から削除されます。また、矢印ボタン (↑↓) を押すと表示順序を変える ことができる。設定の後 Redraw ボタンを押すと、設定に基づき再描画される。

2-4. BLAST 検索

SYGDでは登録されている全ゲノムに対してBLAST検索を行なうことができます。 メインメニューの"BLAST"をクリックすると、BLAST 検索画面に移動します (図 16)。

LAST Searc	лт П	
Genome	<u>^</u>	
uery Sequence	<u>~</u>	
ile Upload 選択 ファイルが資 arget Genome	祝されていません。	submit リセット
ile Upload 選択 ファイルが刻 arget Genome ②[K7] Saccharo	訳だされていません。 myces cerevisiae Kyokai no. 7 マ[\$288	submit Uセット c] Saccharomyces cerevisiae S288c
ie Upload 選択 ファイルが弱 arget Genome ②[K7] Saccharo -value	訳 されていません。 myces cerevisiae Kyokai no. 7 ☑[\$288	submit リセット c] Saccharomyces cerevisiae S288c All Check All Clear
ile Upload 選択 ファイルが刻 arget Genome ②[K7] Saccharo -value 1e-10	駅されていません。 myces cerevisiae Kyokai no. 7 ☑[S288 1e-10 3	submit Uセット c] Saccharomyces cerevisiae S288c All Check All Clear

図 16: BLAST 検索画面。

BLAST 検索画面の説明を以下に記します:

Target DB	検索対象とする配列を選択します。 Genome: 全ゲノム配列 All Genes: 遺伝子配列 All Transcripts:RNA配列 All CDSs:CDS 配列 All Proteins:アミノ酸配列 Intergenic Regions:遺伝子間領域
	Upstream 2kb from all genes:遺伝子上流の 2kb の配列
Query Sequence/ File Upload	検索したい配列を Query Sequence 内に入力するか、あるいは File Upload か ら配列情報の入ったファイルをアップロードします。 対応フォーマット:配列のみ、FASTA 形式, GenBank 形式
Target Genome	ゲノムリストにあるゲノムのうち、どのゲノムを検索対象にするかを指定します。
E-value	BLAST 検索 e-value の閾値を決めます。 セレクトメニューまたは、テキストボックスにて閾値を入力してください。
Additional position	他のオプションを用いたい場合には、オプションと値をここに直接書き込むこと ができます。

BLAST 検索結果の例を図17 に示します。

<pre>>NC_001134.8 NC_001134.8 ([S288c] Saccharomyces cerevisiae S288c)</pre>								
Len								
Score = Identities = Strand = Plu	= 0/3532 (0%)							
Query 1	ATGAGCAGTAGCAAGAAATTGGCCGGTCTTAGGGACAATTTCAGTTTGCTCGGCGAAAAG	60						
Sbjct 662249	ATGAGCAGTAGCAAGAAATTGGCCGGTCTTAGGGACAATTTCAGTTTGCTCGGCGAAAAG	662190						
Query 61	AATAAGATCTTGGTCGCCAATAGAGGTGAAATTCCGATTAGAATTTTTAGATCTGCTCAT	120						
Sbjct 662189	AATAAGATCTTGGTCGCCAATAGAGGTGAAATTCCGATTAGAATTTTTAGATCTGCTCAT	662130						
Query 121	GAGCTGTCTATGAGAACCATCGCCATATACTCCCATGAGGACCGTCTTTCAATGCACAGG	180						
Sbjct 662129	GAGCTGTCTATGAGAACCATCGCCATATACTCCCATGAGGACCGTCTTTCAATGCACAGG	662070						
Query 181	TTGAAGGCGGACGAAGCGTATGTTATCGGGGAGGAGGGCCAGTATACACCTGTGGGTGCT	240						
Sbjct 662069	TTGAAGGCGGACGAAGCGTATGTTATCGGGGAGGAGGGCCAGTATACACCTGTGGGTGCT	662010						
Query 241	TACTTGGCAATGGACGAGATCATCGAAATTGCAAAGAAGCATAAGGTGGATTTCATCCAT	300						
Sbjct 662009	TACTTGGCAATGGACGAGATCATCGAAATTGCAAAGAAGCATAAGGTGGATTTCATCCAT	661950						
Query 301	CCAGGTTATGGGTTCTTGTCTGAAAATTCGGAATTTGCCGACAAAGTAGTGAAGGCCGGT	360						
Sbjct 661949	CCAGGTTATGGGTTCTTGTCTGAAAATTCGGAATTTGCCGACAAAGTAGTGAAGGCCGGT	661890						
Query 361	ATCACTTGGATCGGCCCTCCAGCTGAAGTTATTGACTCTGTGGGTGACAAAGTCTCTGCC	420						
Sbjct 661889	ATCACTTGGATCGGCCCTCCAGCTGAAGTTATTGACTCTGTGGGTGACAAAGTCTCTGCC	661830						
Query 421	AGACACTTGGCAGCAAGAGCTAACGTTCCTACCGTTCCCGGTACTCCAGGACCTATCGAA	480						
Sbjct 661829	AGACACTTGGCAGCAAGAGCTAACGTTCCTACCGTTCCCGGTACTCCAGGACCTATCGAA	661770						
Query 481	ACTGTGCAAGAGGCACTTGACTTCGTTAATGAATACGGCTACCCGGTGATCATTAAGGCC	540						
Sbjct 661769	ACTGTGCAAGAGGCACTTGACTTCGTTAATGAATACGGCTACCCGGTGATCATTAAGGCC	661710						
Query 541	GCCTTTGGTGGTGGTGGTAGAGGGTATGAGAGTCGTTAGAGAAGGTGACGACGTGGCAGAT	600						
Sbjct 661709	GCCTTTGGTGGTGGTGGTAGAGGTATGAGAGTCGTTAGAGAAGGTGACGACGTGGCAGAT	661650						

図 17: BLAST 検索結果の例

2-5.ドットプロット解析

Apply

ドットプロット解析を行なうことで2種類のゲノム、染色体、あるいはコンティグ間の類 似性を全体的に眺めることができます。

メインメニューの "Dot plot analysis" のリンクをクリックすることで、ドットプロット解 析画面に移動します(図 18)。

Search - Genome - BLAST Com	parative Genome -	Categ	gory -							
Genome List in Sake Yeast Do	t plot analysis ┥	-	clic	k						
Organism	ferential analysis	Ta	axon. id	Mb	Chrmosome	Contigs	CDSs	tRNAs	rRNAs	
Saccharomyces cerevisiae S288c	S288c SGD	5	59292	12.1	16	17	6008	299	14	
Saccharomyces cerevisiae Kyokai no. 7	K7 NRIB_S	rgd 7	721032	11.8	16	17	5816	296	7	
All			-	-	32	34	11824	595	21	
Top / Dotplot analysis										
Dotplot analysis [region based](count based)										
Genome vs Genome 3 50000 [bp/pixel] 3 1e-10										
QueryA (X axis) : select g	enome				٢					
QueryB (Y axis) : select genome										

図18: ドットプロット解析画面。メインメニューの "Dot plot analysis"をクリックすることで、ドットプロット解 析画面へ移動します。

左上のセレクトメニューで、ゲノム同士の比較、染色体同士の比較、コンティグ同士の比較を選択できます。中上のメニューでは、図の解像度(図19)を決めます。数値が小さくなるほど大きな図となります。右上メニューは検索(BLAST 検索)の e-value 閾値となります。一般に、ゲノム同士の比較を行うと非常に時間がかかる場合があります。region based はゲノム上の塩基単位で、count basedはCDS単位で図を作成します。

ゲノム同士のドットプロットから、見たい染色体、コンティグ、コンティグ内の領域にお ける具体的な遺伝子を見ることができます (図19)。



QueryA : [K7] Saccharomyces cerevisiae Kyokai no. 7 vs QueryB : [S288c] Saccharomyces cerevisiae S288c



図19: ドットプロット解析画面.ゲノム同士のドットプロット解析結果を示します。

2-6.差分解析

差分解析を行なうことにより、複数のゲノムグループ間で共通に存在する遺伝子、あるい は特定のグループにのみに存在する遺伝子を検索することができます。

メインメニューの "Differential analysis" のリンクをクリックすることで、差分解析画面に 行くことができます(図 20)。



図20: 差分解析画面. メインメニューの "Differential analysis"をクリックすることで、差分解析画面へ移動します。

用いる orthlog set を選択し, 作成するベン図 (No. of group. 1, or 2, or 3)を選択、さらに group に属する種 (図で はA,B に属する種) を選択する。

Search ボタンを押すと ortholog の検索画面 (図 32, 2-7-3 オーソログ検索参照)へ移動します。

2-7.基本検索

メインメニューの "Search" のリンクをクリックすることで、基本検索画面に行くことができます (図 21)。

Contig	- 12	ake Yeast Gen	ome Dat	tabase						
Gene <		СК	Acronym	Source	Taxon. id	Mb	Chrmosome	Contigs	CDSs	tRNA
RNA		visiae S288c	S288c	SGD	559292	12.1	16	17	6008	29
Ortholog		visiae Kyokai no. 7	K7	NRIB_SYGD	721032	11.8	16	17	5816	29
COG			-		-		32	34	11824	59
KOG										
EC										
/ Search (Gene)										
arch										
Select Tarpet: Co	🖦 Gen	E RNA CDS Ortholog	COG KOG EC							
under Filligen of		e not dos ordinoly								
search										
Sene Search	I I Hide F	ilter Option								
Full Text										
Gene ID										
Gene Symbol										
Description					1					
Genome	•		0		1					
Chromosome					1					
Contig										
Contig Location	From	То								
Comment										
Ortholog					1					
COG ID					1					
					1					
COG Functions	□3 □A	CK CL C8 C⊅ CY		M ON CZ OW						
COG Functions	□ □ □ A	CK CL CB CD CY CC CG CE CF CH		M ()N ()Z ()W Q ()R ()S						
COG Functions	0 U			M ()N ()Z ()W Q ()R ()S]					
COG Functions KOG ID				M N Z W						
COG Functions										
COG Functions KOG ID KOG Functions	· · · · · · · · · · · · · · · · · · ·			M N Z W						
KOG ID KOG Functions	· · · · · · · · · · · · · · · · · · ·			M N Z W Q R S M N Z W Q R S X						
COG Functions KOG ID KOG Functions PRINTS Category	· · · · · · · · · · · · · · · · · · ·			M N Z W Q R S M N Z W Q R S X						
COG Functions KOG ID KOG Functions PRINTS Category PRINTS				M N Z W Q R S M N Z W Q R S X						
COG Functions KOG ID KOG Functions PRINTS Category PRINTS EC Number	· · · · · · · · · · · · · · · · · · ·			M N Z W Q R 5 M N Z W Q R 5 X						
COG Functions KOG ID KOG Functions PRINTS Category PRINTS EC Number Pfam				M N Z W Q R S M N Z W Q R S X						
COG Functions KOG ID KOG Functions PRINTS Category PRINTS EC Number Pfam Prosite	· · · · · · · · · · · · · · · · · · ·			M N Z W Q R S M N Z W Q R S X						

図 21: 基本検索画面. メインメニューの "Search"をクリックすることで、基本検索画面へ移動します。

基本検索では、以下に示すものをキーワードや条件で検索することができます。

- (1). コンティグ
- (2). 遺伝子、RNA、CDS
- (3). オーソログ (パラログ)
- (4). COG、KOG、EC アノテーション

2-7-1. コンティグ検索

SYGD に登録されているゲノム情報の中から、コンティグを検索することができます。 コンティグ検索画面を図22、コンティグ検索結果表示画面を図23 に示します。 コンティグ検索画面は、基本検索画面 (図 21) の "Contig" リンクをクリックすることで表 示されます。

Search									
_	Select Target:	Contig	Gene	RNA	CDS	Ortholog	COG	KOG	EC
	Search								
	Contig Sea	arch : 🖽	de Filto	er Opti	on				
	Genome	-						0	
	Chromosome								
	Contig ID								
	Search	Reset							

図22: コンティグ検索画面. 基本検索画面の "Contig" をクリックする (図の矢印) ことで、コンティグ検索画面を 表示させることができます。基本検索画面の表示方法については、 2-7 基本けんさく参照。表中にある項 目に検索条件を入れて、"Search" ボタンを押すと、条件に合致したコンティグのリストが表示されます。

コンティグ検索画面では、検索したいゲノム、染色体をメニューから選択し、ID等を入力 することでコンティグの検索を行うことができます。コンティグ検索画面の詳細を以下に 記します:

検索したいゲノムを選択することができます。 セレクトメニューから適当なゲノム名を 選択してください。
"Genome"のセレクトメニューで染色体情報を持っているゲノムを選択した場合、検索 したい染色体を選択することができます。セレクトメニューから適当な染色体を選択 してください。
登録されているコンティグに対して、Contig ID で検索します。

コンティグ検索結果表示画面は以下の3つで構成されています(図23)。

また、以降説明する遺伝子、RNA、CDS、オーソログ、COG、KOG、EC アノテーションの検索結果表示画面も同様な構成になっています。

- (1). フィルターオプションの表示: フィルターオプションの表示/非表示
- (2). コンティグリスト: コンティグの一覧表が表示されます

(3). 表示コントロール部:一覧表の並び替え (ソート)を実行、一覧表のカラムの表示/非表示、データのダウンロードを行なうことができます。

earch Select Target: Contig Gene INA COS Ortholog COG KOG EC Search Contig Search : Hide Filter Option Genome [5286() Saccharomyces cerevisiae 5286] Dremnsome] Contig ID KC 201144.5 Search Reset Contig ID KC 201144.5 Search Reset Contig ID Organism Urromosome Length COC% No. of Genes 1 NC. 201144.5 Saccharomyces cerevisiae 5288 CVr. XII 1078177 bp 38.48% 572 Contig III - 1 of 1 Rows: 50 Page: 1 © Top Bottom Page: 1 No./I Contig ID Organism Urromosome Length COC% No. of Genes 5 Sect Columns First Key :: Select Key : Sacced © Second Key : Select Key : Ascend © Third Key : Select Key : Sacced © Ascend © Show / Hide Columns Contig ID Organism Genome Chromosome Length GC% No. of Genes _ Apply Download Data	Sete: Top Bottom Page: 1 Set: Set: Set: Contig Search : Hide Filter Option Genome: \$\$2886: Second Contig Search : Hide Filter Option Genome: \$\$2886: Second Contig Ib: 1.01144.5 Second Second: Rows: \$\$0 \$Page: 1 No: Contig Ib: Contig Ib: Page: 1 No: Contig Ib: Contig Ib: Contig Ib: No: Otopanism Chromosome: Ingth Ct/s No: Contig Ib: Contig Ib: Contig Ib: Page: 1 No: Contig Ib: Contig Ib: Contig Ib: Page: 1 Source Top Bottom Page: 1 Page: 1 Page: 1 Source Top Bottom Page: 1 Page: 1 Page: 1 Source Top Bottom Page: 1 Page: 1 Page: 1 Source Contig Ib:	/ Search (Contig)							
Seer target: Contig Gene NNA COS Onthelog COG KOG EC Search Contig Search : Hide Filter Option Genome [S286; Saccharomycos cerevisiae S288c] Corring ID NC.001144.5 Search Reset Contig III - 1 of 1 Rows: 50 Page: 1 D Top Bottom Page: 1 No./1 Contig ID Organism Chromosome Length CC% No. of Genes 1 NC.001144.5 Saccharomyces cerevisiae S288c (rr. XII 1078177 bp 38.48% 572 Contig III - 1 of 1 Rows: 50 Page: 1 D Top Bottom Page: 1 Search Organism Chromosome Length CC% No. of Genes First Key : Select Key D Ascend D Third Key : Select Key D Asc	Seet Target Contig Gene RNA COS Ontwood COS KOG EC Seet Target Contig Gene RNA COS Ontwood COS KOG EC Seet Target Contig Rear I Hide Filter Option Genome [5288c] SectoAromyces cerevisie S288c Contig ID K_001144.5 Search Reset Contig ID K_001144.5 Search Reset Contig ID K_001144.5 Search Reset Contig ID Organism Chromosome Length COS No. of Genes I NC.01514.5 [SectoAromyces cerevisies S288c Core. XII 1076177 bp 38.4% 572 Contig IBt - 1 of 1 Rows: S0 Page: 1 C Top Bottom Page: 1 Soft Columns Set Columns Contig ID Corganism Second Key : - Select Key - Ascend Third Key : - Select Key - Accend Paper Show / Hide Columns Contig ID Corganism Chromosome Length GC% No. of Genes Apply Show / Hide Columns	arch							
Contig Search: Hide Filter Option Contig Search: Hide Filter Option Corring ID Nc. 001144.5 Search Reset Contig II - 1 of 1 Rows: 50 @Page: 1 Top Bottom Page: 1 No./I Contig ID Organism Chromosome Length GCMs Mo. of Genes I NC.001144.5 Saccharomyces cerevisiae 520 @Page: 1 Top Bottom Page: 1 Rows: 50 @Page: 1 Top Bottom Page: 1 Sort Columns First Key : Select Key - @ Ascend @ Third Key : Select Key - @ Ascend @	Search igge Control go control in c	Select Tarnet: CO	ntia Gene RNA CI	is Ortholog COG I	KOG EC				
Search Contig Search : Hide Filter Option Genome [5288c] Saccharomyces cerevisiae 5288c] Contig ID Nc_001144.5 Search Reset Contig list 1 - 1 of 1 Rows: 50 Page: 1 Top Bottom Page: 1 No./1 Contig ID Organism Cerevisiae 528 (r. XII 1078177 bp 38.48% 572 Contig list 1 - 1 of 1 Rows: 50 Page: 1 Top Bottom Page: 1 Soft Columns First Key : Select Key @ Ascend @ Second Key : Select Key @ Ascend @ Third Key : Select Key @ Ascend @ First Key : Select Key @ Ascend @ Second Key : Select Key @ Ascend @ Third Key : Select Key @ Ascend @ First Key : Select Key @ Ascend @ Second Key : Select Key @ Ascend @ First Key : Select Key @ Ascend @	Sarch Contig Search : Hide Filter Option Genome [5288c] Saccharomyces cerevisiae 5288c © Orromosome © Contig ID N_001144.5 Search Reset Contig Ii 1 - 1 of 1 Rows: 50 Page: 1 Top Bottom Page: 1 No./1 Contig ID Organism Chromosome Length GC% No. of Genes 1 NCO1144.5 Saccharomyces cerevisiae 5288c Or. XII 1078177 bp 38.48% 572 Contig Ii 1 - 1 of 1 Rows: 50 Page: 1 Top Bottom Page: 1 Search Reset Search Reset S	Select Target.	indig dene wor de	is ortholog cod i	NOG EC				
Contig Ist 1 - 1 of 1 Rows: 50 Page: 1 Top Bottom Page: 1 No./L Cottig Ist 1 - 1 of 1 Rows: 50 Page: 1 Top Bottom Page: 1 No./L Cottig Ist 1 - 1 of 1 Rows: 50 Page: 1 Top Bottom Page: 1 No./L Cottig Ist 1 - 1 of 1 Rows: 50 Page: 1 Top Bottom Page: 1 No./L Cottig Ist 3 Saccharomyces cerevisiae 5288c Crv. XII 1078177 bp 38.48% 577 Contig Ist 1 - 1 of 1 Rows: 50 Page: 1 Top Bottom Page: 1 Soft Columns First Key : Select Key - 2 Ascend 2 Second Key : Select Key - 2 Ascend 3 Third Key : Select Key - 2 Ascend 3 Show / Hide Columns Contig 10 Organism Genome Chromosome Length GC% No. of Genes 2 Contig 10 Organism Genome Chromosome Length GC% No. of Genes 2 Contig 10 Organism Genome Chromosome Length GC% No. of Genes 2 Pagey	Contig Search : Hide Filter Option Genome [[5286] Seacharomyces cerevisiae 5288 Onromeant Contig ID NC_001144.5 Search Reset Contig list 1 - 1 of 1 Rows: 50 Page: 1 Top Bottom Page: 1 No./1 Contig ID Organism Chromosome Length GC% No. of Genes 1 NC_001144.5 Saccharomyces cerevisiae 5286 CW. XII 1078177 bp 38.48% 572 Contig list 1 - 1 of 1 Rows: 50 Page: 1 Top Bottom Page: 1 Sort Columns First Key : Collect Key - 2 Ascend 2 Second Key : Select Key - 2 Ascend 3 Third Key : Select Key - 3 Ascend 3 Apply Show / Hide Columns Contig ID Coganism Conomes Length: GC% No. of Genes Apply Download Data Nucleotide Sequences Tab Delimited File	Search							
Genome [\$2886] Saccharomyces cerevisiae \$288c Chrumasome Contig ID Nc. 001144.5 Search Reset Page: 1 Contig ID Organism Chromosome Page: 1 No./1 Contig ID Organism Chromosome Page: 1 No./1 Contig ID Organism Chromosome Page: 1 No./1 Contig ID Organism Chromosome Page: 1 Sort Columns First Key :	Genome [5288c] Saccharomyces cerevisiae 5288c Greenmoome • •	Contig Search	Hide Filter Option						
Chromesome • • •	Chramework	Genome	288c] Saccharomyces ce	revisiae S288c	3				
Contig ID NC_001144.5 Search Reset Contig Ist 1 - 1 of 1 Rows: 50 Page: 1 Top Bottom Page: 1 Rows: 50 Rows Rows Rows Rows Rows Rows Rows Rows	Contig ID NC_001144.5 Search Reset Contig ID Organism Dorganism Chromosome Ling ID Organism Contig ID Organism Contig ID Organism Chromosome Length GCMs No. of Genes I NC.00144.5 Saccharomyces cerevisiae S288c Orr. XII TOP Bottom Pege: 1 Sort Columns Top Bottom First Key : Sected Key :-> Ascend Sected Key :-> Ascend Third Key : Show / Hide Columns Contig ID Organism Contig ID Organism Contig ID Organism Contig ID Organism Contig ID Top Bottom	Chromosome	0						
Search Reset Contig list 1 - 1 of 1 Rows: 50 Page: 1 Top Bottom Page: 1 No./1 Ortganism Chromosome Length GC% No. of Genes 1 Nc00144.5 Saccharomyces cervisiae S280 (rr. XII 1078177 bp 38.48% 572 Contig list 1 - 1 of 1 Rows: 50 Page: 1 Top Bottom Page: 1 Sort Columns First Key :	Search Reset Contig lit 1 - 1 of 1 Rows: 50 Page: 1 Top Bottom Page: 1 No/1 Cotig lit Organism Chromosome Length GC% No. of Genes 1 No.01144.5 Saccharomyces cerevisiae 5288c Ore. XIII 1078177 bp 28.48% 572 Contig lit 1 - 1 of 1 Rows: 50 Page: 1 Top Bottom Page: 1 Sort Columns Top Bottom Page: 1 Top Bottom Page: 1 Show / Hide Columns Second Key : Select Key Selec	Contig ID NC	001144.5						
Contig list I - 1 of 1 Rows: 50 Page: 1 Top Bottom Page: 1 No./I Contig ID Organism Chromosome Length GC% No. of Genes 1 NC.00144.5 Saccharomyces cerevisiae 5288c Orr. XII 1078177 bp 38.48% 572 Contig list I - 1 of 1 Rows: 50 Page: 1 Top Bottom Page: 1 Sort Columns First Key : Select Key P Ascend P Second Key : Select Key P Ascend P Third Key : Select Key P Ascend P Ascend P Ascend P Show / Hide Columns Contig ID Organism Genome Chromosome Length GC% No. of Genes Apply Bownload Data	Sort Columns Page: 1 Top Bottom Page: 1 Sort Columns Image: 1 Rows: 50 Page: 1 Image: 1 Sort Columns Rows: 50 Page: 1 Image: 1 Image: 1 Sort Columns Rows: 50 Page: 1 Image: 1 Image: 1 Sort Columns Image: 1 Rows: 50 Page: 1 Image: 1 Sort Columns Image: 1 Image: 1 Image: 1 Image: 1 Sort Columns Image: 1 Image: 1 Image: 1 Image: 1 Show / Hide Columns Image: 1 Image: 1 Image: 1 Image: 1 Show / Hide Columns Image: 1 Image: 1 Image: 1 Image: 1 Show / Hide Columns Image: 1 Image: 1 Image: 1 Image: 1 Show / Hide Columns Image: 1 Image: 1 Image: 1 Image: 1 Show / Hide Columns Image: 1 Image: 1 Image: 1 Image: 1 Image: 1 Image: 1 Image: 1 Image: 1 Image: 1 Image: 1 Image:	Search Rese							
Contig list I - 1 of 1 Rows: 50 Page: 1 Top Bottom Page: 1 No./1 Cortig I D Organism Chromosome Length GC% No. of Genes 1 NC.001144.5 Saccharomyces cerevisiae 5288c (2rr, XII 1078177 bp 38.48% 572 Contig list I - 1 of 1 Rows: 50 Page: 1 Top Bottom Page: 1 Sort Columns Second Key : - Select Key S Ascend S Third Key : - Select Key S Ascend A Assend A Ascend A Ascend A Ascend A Ascend A Assend A Ascend A Assend A A	Contig liet 1 - 1 of 1 Rows: 50 Page: 1 Top Bottom Page: 1 No./1 Contig ID Organism Chromosome Length GC% No. of Genes 1 MC_001144.5 Saccharomyces cerevisiae S286c CM. XII 1078177 bp 38.48% 572 Contig liet 1 - 1 of 1 Rows: 50 Page: 1 Top Bottom Page: 1 Sort Columns								
Interference Image: Control Diagonal Control Diagonal Di	Source Chromosome Length Cords Page: 1 Source Chromosome Length CC/s No. of Genes 1 NC_001144.5 Saccharomyces cerevisiae S288c Chr. XII 1078177 by 38.48% S72 Contig list 1 - 1 of 1 Rows: So Page: 1 Page: 1 Page: 1 Source Ascend Page: 1 Top Bottom Page: 1 Source Ascend Second Key : Select Key P Ascend	Contia list 1 - 1 of	1	Rows: 50	ae: 1 🖸	Top Bo	ottom	Page: 1	
I NC.00144.5 Saccharomyces cerevisiae 5288 Cirr. XII 1078177 bp 38.45% 572 Contig list I - 1 of I Rows: 50 Page: I Top Bottom Page: I Sort Columns First Key : Select Key P Ascend Second Key : Select Key P Ascend Third Key : Select Key P Ascend Ascend Page: I Show // Hide Columns Contig ID Organism Genome Chromosome Length 6C% No. of Genes Apply Apply Download Data	I NC_001144.5 Saccharomyces cerevisiae S288c Chr. XII 1078177 bp 38.48% 572 Contig list I - 1 of I Rows: 50 Page: I Top Bottom Page: I Sort Columns First Key : - Select Key Ascend Third Key : - Select Key Ascend Ascend </th <th>No./1 Contig ID</th> <th>Organism</th> <th>Chrome</th> <th>some Length</th> <th>GC%</th> <th>No. of Genes</th> <th>- oger k</th> <th></th>	No./1 Contig ID	Organism	Chrome	some Length	GC%	No. of Genes	- oger k	
Contig list 1 - 1 of 1 Rows: 50 Page: 50 Pa	Contig list I - 1 of 1 Rows: 50 Page: 1 Top Bottom Page: 1 Sort Columns First Key : Select Key Ascend Third Key : Select Key Ascend Third Key : Select Key Ascend Apply Show / Hide Columns Contig 10 Organism Genome Chromosome Length CC% No. of Genes Apply Download Data Nucleotide Sequences Tab Delimited File	1 NC_001144	.5 Saccharomyces cerev	isiae S288c Chr. XI	I 1078177 b	p 38.48%	572		
Sort Columns First Key : Select Key Ascend Second Key : Select Key Ascend	Sort Columns First Key :Select Key Ascend Second Key :Select Key Ascend Third Key :Select Key Ascend Apply Show / Hide Columns Control D0 CorganismGenomeChromosome & Length & GC%No. of Genes Apply Download Data Nucleotide Sequences & Tab Delimited File	Contig list 1 - 1 of	1	Rows: 50 OPa	ge: 1 😳	Top Bo	ottom	Page: 1	
Sort Columns First Key : Select Key Ascend Select Key	Sort Columns First Key : Select Key Ascend Second Key : Select Key Ascend Third Key : Select Key Ascend Asc								
First Key : [Select Key] Ascend Second Key : [Select Key] Ascend Third Key : [Select Key] Ascend	I'rit Key (Select Key) Ascend Second Key (Select Key) Ascend Third Key (Select Key) Ascend Ascend Apply Show / Nide Columns Contig ID @rganism Genome Chromosome Length @GC% No. of Genes Apply Download Data Nucleotide Sequences Tab Delimited File	Sort Columns							
Apply Show / Hdc Columns Contig 1D _Organism _ Genome _ Chromosome _ Length _ GC% _ No. of Genes Apply	Appy Show / Nide Columns Contig ID @Organism Genome Chromosome Length GC% No. of Genes Appy Download Data Nucleotide Sequences Tab Delimited File	First Key : Sele	ct Key 🔋 Ascend	Second Key	Select Key	Asce	nd 🖸 Thir	rd Key : 🔄 Select Key 🔋 🕻 Ascend	
Show / Hide Columns Contig To Organism Genome Chromosome Length GC% No. of Genes Apply Download Data	Show / Hide Columns Contig DD @rganism @Genome @Chromosome @Length @GC% @No. of Genes Apply Download Data Nucleotide Sequences @ Tab Delimited File	Apply							
Contig ID Corpanism Genome Chromosome Length CGC% No. of Genes Apply Download Data	Contrig to Conganism Genome Chromosome Length GC% No. of Genes Apply Download Data Nucleotide Sequences © Tab Delimited File	Show / Hide Colu	imns						
Apply Download Data	Apply Download Data Nucleotide Sequences © Tab Delimited File	Contig ID Orga	inism 🗌 Genome 🛃 Chro	mosome <mark>Z</mark> Length	GC% No. of C	Genes			
Download Data	Download Data Nucleotide Sequences © Tab Delimited File	Apply							
	Nucleotide Sequences O Tab Delimited File	Download Data							

図 23: コンティグ検索結果表示画面.

コンティグ検索結果表示画面のカラム (図 23) の説明を以下に記します:

Contig ID	登録されているコンティグのContig ID を示します。
	またコンティグマップ画面へのリンクにもなっています。
Organism	分類学上の種 (species) を示します。
Genome	SYGDに登録されているゲノムの名称を示します。
	この名称は、主にゲノムの分類学上の種 (species)・サンプル名・解析名・略称など
	ていされより。 また蚊称けゲノムトップ両面 (M2) へのリンクにたたっています
	みた哈かはクノムトツノ回国(図3)へのワノクにもなりています。
	初期状態 (テノオルト) ぐは表示されません。表示コントロール C"Genome" ヘチェッ
	クすることで表示されます。
Chromosome	Chromosome情報がある場合には、属しているChromosome名が表示されます。
Chromosome	Chromosome情報がある場合には、属しているChromosome名が表示されます。 Chromosome情報がない場合には "Unmapped" と表示されます。
Chromosome	Chromosome情報がある場合には、属しているChromosome名が表示されます。 Chromosome情報がない場合には "Unmapped" と表示されます。 コンティグリスト画面へのリンクにもなっています。クリックすることで属している
Chromosome	Chromosome情報がある場合には、属しているChromosome名が表示されます。 Chromosome情報がない場合には "Unmapped" と表示されます。 コンティグリスト画面へのリンクにもなっています。クリックすることで属している Chromosome のコンティグリストー覧が表示されます ("Unmapped"である場合に
Chromosome	Chromosome情報がある場合には、属しているChromosome名が表示されます。 Chromosome情報がない場合には "Unmapped" と表示されます。 コンティグリスト画面へのリンクにもなっています。クリックすることで属している Chromosome のコンティグリストー覧が表示されます ("Unmapped"である場合に は、"Unmapped"のコンティグー覧が表示されます)。
Chromosome	Chromosome情報がある場合には、属しているChromosome名が表示されます。 Chromosome情報がない場合には "Unmapped" と表示されます。 コンティグリスト画面へのリンクにもなっています。クリックすることで属している Chromosome のコンティグリストー覧が表示されます ("Unmapped"である場合に は、"Unmapped"のコンティグー覧が表示されます)。 コンティグの長さを示します。
Chromosome Length GC%	Chromosome情報がある場合には、属しているChromosome名が表示されます。 Chromosome情報がない場合には "Unmapped" と表示されます。 コンティグリスト画面へのリンクにもなっています。クリックすることで属している Chromosome のコンティグリストー覧が表示されます ("Unmapped"である場合に は、"Unmapped"のコンティグー覧が表示されます)。 コンティグの長さを示します。 コンティグの GC% を示します。
Chromosome Length GC% No. of Genes	Chromosome情報がある場合には、属しているChromosome名が表示されます。 Chromosome情報がない場合には "Unmapped" と表示されます。 コンティグリスト画面へのリンクにもなっています。クリックすることで属している Chromosome のコンティグリストー覧が表示されます ("Unmapped"である場合に は、"Unmapped"のコンティグー覧が表示されます)。 コンティグの長さを示します。 コンティグの GC% を示します。 コンティグ上の遺伝子数を示します。
Chromosome Length GC% No. of Genes	Chromosome情報がある場合には、属しているChromosome名が表示されます。 Chromosome情報がない場合には "Unmapped" と表示されます。 コンティグリスト画面へのリンクにもなっています。クリックすることで属している Chromosome のコンティグリストー覧が表示されます ("Unmapped"である場合に は、"Unmapped"のコンティグー覧が表示されます)。 コンティグの長さを示します。 コンティグの GC% を示します。 コンティグ上の遺伝子数を示します。 また遺伝子検索結果表示画面 (図 25) へのリンクにもなっています。

2-7-2. 遺伝子、RNA、CDS 検索

SYGD に登録されているゲノム情報の中から、遺伝子・RNA・CDS を検索することができます。

遺伝子検索画面を図24 に示します (RNAやCDS検索画面も、ほぼ同じ仕様となっています)。また遺伝子検索結果表示画面を図25、RNA検索結果を図26、CDS検索結果を図27 に示します。

遺伝子検索画面は、基本検索画面 (図21) の "Gene" リンクをクリックすることで表示されます。同様に、RNA検索画面は "RNA"、CDS検索画面は "CDS" をクリックすることで表示されます。

Top / Search (Gene)									
Se	Search								
_	Select Target: Con	tig Gene RNA CDS Ortholog COG KOG EC							
1	Search								
	Gene Search	: Hide Filter Option							
	Full Text								
	Gene ID								
	Gene Symbol								
	Description								
	Genome	· · ·							
	Chromosome								
	Contig								
	Contig Location	From To							
	Comment								
	Ortholog								
	COG ID								
		1							

図24: 遺伝子検索画面. 基本検索画面の "Gene" をクリックする (図の矢印) ことで、遺伝子検索画面を表示させることができます。同様にRNA検索画面は "RNA" を、CDS検索画面は "CDS" をクリックすることで、それぞれの検索画面を表示させることができます。基本検索画面の表示方法については、2-7 基本検索参照。表中にある項目に検索条件を入れて、"Search" ボタンを押すと、条件に合致した遺伝子 (RNA, CDS) のリストが表示されます。

遺伝子検索画面の各項目の詳細を以下に記します:

Full Text	登録されている遺伝子に対して全文検索を行います。
Gene ID	登録されている遺伝子に対して Gene ID で検索を行います。
Gene Symbol	登録されている遺伝子に対して遺伝子の symbol 名で検索を行います。 スペースを空けて複数指定できます (or 検索) 。
Description	登録されている遺伝子に対してDescription情報で検索を行います。例えば、 "enzyme" と入力すると、gene description 欄に "enzyme" という文字列が含まれ る遺伝子を検索します。
Genome	検索したいゲノムを選択することができます。セレクトメニューから適当なゲノム名 を選択してください。
Chromosome	"Genome"のセレクトメニューで染色体情報を持っているゲノムを選択した場合、検索したい染色体を選択することができます。セレクトメニューから適当な染色体を 選択してください。
Contig	"Genome"のセレクトメニューでコンティグ情報を持っているゲノムを選択した場 合、検索したいコンティグを選択することができます。セレクトメニューから適当 なコンティグを選択してください。
Contig Location	コンティグ中の検索したい領域を指定できます。テキストボックスに検索したい領 域を数値 (単位bp) で入力してください。
Comment	登録されている遺伝子に対してComment情報で検索を行います。例えば、ある遺伝 子のComment欄に"enzyme"という単語を保存していた場合、"enzyme"と入力する ことでComment欄に"enzyme"と記載した遺伝子を検索します。
Ortholog	登録されている遺伝子に対して Ortholog ID で検索を行います。
COG ID	登録されている遺伝子に対して COG (オーソログ) 情報の ID で検索を行います。
COG Functions	登録されている遺伝子に対して COG (オーソログ) 情報の functional categories で検索を行います。
KOG ID	登録されている遺伝子に対して KOG (オーソログ) 情報の ID 情報で検索を行います。

KOD Functions	登録されている遺伝子に対して KOG (オーソログ) 情報の functional categories で検索を行います。
PRINTS	登録されている遺伝子に対して PRINTS (タンパク質モチーフ) 情報で検索を行いま
Category	す。
EC Number	登録されている遺伝子に対して EC Number 情報で検索を行います。
Pfam	登録されている遺伝子に対して Pfam (タンパク質モチーフ) 情報で検索を行います。
Prosite	登録されている遺伝子に対して Prosite (タンパク質モチーフ) 情報で検索を行います。
No. of TMs	登録されている遺伝子に対して膜貫通領域と予測されたドメインについてその数を もとに検索を行います。
	右のテキストボックスに適当なドメイン数を半角文字で入力し、左のセレクトメニ
	ューから、以上(>=), 同じ(=), 以下(<=) を選択してください。
GPI-anchor	GPI アンカーとの結合が予測された遺伝子について検索を行います。
Signal Peptide	シグナルペプチドと予測された遺伝子について検索を行います。

Gene list 1 - 2 of 2				Rows	: 50 <mark>0</mark> Page: (1 😳	Top Bottom	Page: 1		
	No./2	Gene ID	Original symbol	Assigned symbol	Original description	Assigned description	Links	Organism	Chromosome	Comment
	1	YBR218C	PYC2	PYC2	pyruvate carboxyl ase 2	Pyruvate carboxy lase	NCBI_GeneID:8 52519	Saccharomyces cerevisiae S 288c	Chr. II	-
	2	SYK7_005 261	K7_PYC2	PYC2	К7_Рус2р	Pyruvate carboxy lase		Saccharomyces cerevisiae K yokai no. 7	Chr. II	-
6	iene li	st 1 - 2 of 2	1	Rows	: 50 ᅌ Page: 🤇	1	Top Bottom	Page: 1		

図 25:遺伝子検索結果表示画面.

RNA list 1 - 50 of 11883					Rows: 50 📀	Page:	1	0		Тор	p Bottom		Pag	e:1, 2,	3, 4, 5, 6		
No./11883	Gene ID	Original symbol	Assigned symbol	Links	Description									RNA- type	Organism	Chromosome	Comment
1	YLLO6 7C		YRF1-1	NCBI_Ge neID:85 0659 NCBI_Rn a:NM_00 1181887 .1	Y' element ATP- Putative Y' elem	epeni ent AT	dent I 'P-dep	helicase pendent l	helicase					mRN A	Saccharomyc es cerevisiae S288c	Chr. XII	-
2	YLLO6 6W-B			NCBI_Ge neID:14 66404 NCBI_Rn a:NM_00 1184605 .1	hypothetical pro hypothetical pro	ein ein; c	overeo	xpressior	n causes a	a celicycle	delay or ar	rrest		mRN A	Saccharomyc es cerevisiae S288c	Chr. XII	-
3	YLLO6 6C		YRF1-4	NCBI_Ge neID:85 0660 NCBI_Rn a:NM_00 1181886 .1	Y' element ATP- Putative Y' elem	lepeni ent AT	dent I 'P-dep	helicase pendent l	helicase;	YLL066Cis	s not an ess	sential gene		mRN A	Saccharomyc es cerevisiae S288c	Chr. XII	-

図 26: RNA 検索結果表示画面.

CDS list 1 - 50 of 6008			6008	Rows: 50 OPage: 1	0	Top Bottom		Page: 1, 2, 3, 4, 5, 6
	No./6008	Gene ID	Original description	Assigned description	Genome	Chromosome	Links	Protein sequence
	1	YLL06 7C	Y' element ATP-dependent helicase	Y' element ATP-dependent helicas e YLL067C	Saccharomyces ce revisiae S288c (S 288c)	Chr. XII	SGD:S000 003990 NCBI_Gen eID:85065 9	>YLL067CINP_013033.1INC_001144.5:585-4202:-I120 MRTFTDFVSGAPIX-BLQKSTIRRYGYNLAPHMFLLLHVDELSI VRIEIX-YSGLOMMMXKCPRLGDVLATAGRITRRYNKNGSSE FSSTRNAEDVVAGEAASSDHDQKISRVTRKRPREPKSTNDILV DTPSVAVQAPPCYCKTELFHLDLIALASKGDVKYVSFLFVPYTV GCDGVTDLYVGIYDDLASTNFTDRIAAWENIVECTFRTNNVKL
	2	YLL06 6W-B	hypothetical protein	Uncharacterized protein YLL066W- B	Saccharomyces ce revisiae S288c (S 288c)	Chr. XII	SGD:S000 028672 NCBI_Gen eID:14664 04	>YLL066W-BINP_878115.1INC_001144.5:5605-5775:+156 MSLRPCLTPSSMQYSDIYIHHTHTPHPHHTHTHTHTHTPTHSLTSTF

図 27: CDS 検索結果表示画面.

遺伝子検索結果 (図 25) で初期状態 (デフォルト) で表示されるカラムの説明を以下にい

たします。

Gene ID	遺伝子の ID を示します。NCBI 由来のゲノムデータの場合は、locus tag 名を 示します。
Original symbol	データソースに記載されていた遺伝子名を示します。
Assigned symbol	SYGD データ登録時に再アノテーションされた遺伝子名を示します。
Original description	データソースに記載されていた遺伝子に対する説明を示します。
Assigned description	SYGD データ登録時に再アノテーションされた遺伝子に対する説明を示しま
	す。
Links	外部データベースへのリンクです。
Organism	種名 (菌株名) を示します。
Chromosome	染色体情報のある場合に、染色体番号を示します (染色体情報が無い場合には
	"Unmapped"と示されます)。
Comment	記録したコメントを示します。
Homologous genes	SYGDに登録された他のゲノムと相同性のある遺伝子を表示することができま
	す。
	セレクトボックスから任意のゲノムを選択し、"Apply"ボタンを押すことで相同
	性のある遺伝子が表示されます。

遺伝子検索結果 (図 25) で選択することにより表示されるカラムの説明を以下にいたしま す。

~	
u	

Description	
Genome	ゲノム名を示します。 リンクをクリックすると、 ゲノムトップ画面(図 3)へ移
	動します。
Chromosomal locateion	染色体情報のある場合に、染色体上の位置情報が示されます。
Contig	この遺伝子のあるコンティグ名が示されます。
Contig location	この遺伝子のあるコンティグの、コンティグ上の位置情報が示されます。
Мар	この遺伝子を中心としたコンティグマップが表示されます。
Gene Sequence	遺伝子配列が FASTA 形式で表示されます。
Upstream sequence	この遺伝子の上流の配列を表示します。 テキストボックスに数値 (半角数字)
	を入力し"Apply"ボタンを押してください。
Downstream sequence	この遺伝子の下流の配列を表示します。テキストボックスに数値(半角数字)
·	を入力し"Apply"ボタンを押してください。

Upstream variations	変異解析(M-option)を実施していた場合に利用できます。セレクトボックス からゲノムを選択し"Apply"ボタンを押すことで、遺伝子上流の変異の数が 表示されます。この数値は変異リストへのリンクにもなっています(詳細は 3-1節に記載)。
Ortholog	この遺伝子のオーソログ (パラログ) 情報がある場合、シンテニー解析画面 (図 14) へのリンクが表示されます。
COG	この遺伝子が COG (オーソログ) 情報を含む場合、COG ID を表示します。
COG Function	この遺伝子が COG (オーソログ) 情報を含む場合、COG functional categories を表示します。
KOG	この遺伝子が KOG (オーソログ) 情報を含む場合、KOG ID を表示します。
KOG Function	この遺伝子が KOG (オーソログ) 情報を含む場合、KOG functional categories を表示します。
PRINTS Category	この遺伝子が PRINTS ドメインと類似性がある場合に、PRINTS Category を表示します。
PRINTS	この遺伝子が PRINTS ドメインと類似性がある場合に、 PRINTS 情報を表示 します。
EC number	この遺伝子が EC number 情報がある場合に、EC number 情報を表示します。
Pfam	この遺伝子が Pfam ドメインと類似性がある場合に、PRINTS 情報を表示します。
Prosite	この遺伝子が Prosite ドメインと類似性がある場合に、PRINTS 情報を表示 します。
TMs	この遺伝子が膜貫通領域と予測されたドメインがある場合に、そのドメイン 数を示します。
GPI-anchor	この遺伝子が GPI アンカーとの結合が予測された配列を含む場合に、その 配列の長さを表示します。
Signal Peptide	この遺伝子がシグナルペプチドと予測された配列を含む場合に、その配列の 長さを表示します。

RNA 検索結果(図 26)で初期状態(デフォルト)で表示されるカラムの説明を以下にいたします。

Gene ID	遺伝子の ID を示します。NCBI 由来のゲノムデータの場合は、locus tag 名を示 します。
Original symbol	データソースに記載されていた遺伝子名を示します。
Assigned symbol	SYGD データ登録時に再アノテーションされた遺伝子名を示します。
Links	外部データベースへのリンクです。
Description	データソースに記載されていた遺伝子に対する説明を示します。
RNA-type	RNA のタイプを示します。mRNA, tRNA,rRNA の種類があります。
Chromosome	染色体情報のある場合に、染色体番号を示します (染色体情報が無い場合には
	"Unmapped"と示されます)。
Organism	種名 (菌株名) を示します。
Comment	記録したコメントを示します。
Homologous genes	SYGDに登録された他のゲノムと相同性のある遺伝子を表示することができま
	す。
	セレクトボックスから任意のゲノムを選択し、"Apply"ボタンを押すことで相同性
	のある遺伝子が表示されます。

RNA 検索結果 (図 26) で選択することにより表示されるカラムの説明を以下にいたします。 Genome ゲノム名を示します。リンクをクリックすると、ゲノムトップ画面 (図 3) へ 移動します。 8

RNA Sequence	RNA 配列が FASTA 形式で表示されます。
Upstream sequence	この遺伝子の上流の配列を表示します。テキストボックスに数値 (半角数字) を入力し"Apply"ボタンを押してください。
Downstream sequence	この遺伝子の下流の配列を表示します。テキストボックスに数値 (半角数字) を入力し"Apply"ボタンを押してください。
Ortholog	この遺伝子のオーソログ (パラログ) 情報がある場合、シンテニー解析画面 (図 14) へのリンクが表示されます。
COG	この遺伝子が COG (オーソログ) 情報を含む場合、COG ID を表示します。
COG Function	この遺伝子が COG (オーソログ) 情報を含む場合、COG functional categories を表示します。
KOG	この遺伝子が KOG (オーソログ) 情報を含む場合、KOG ID を表示します。
KOG Function	この遺伝子が KOG (オーソログ) 情報を含む場合、KOG functional categories を表示します。
PRINTS Category	この遺伝子が PRINTS ドメインと類似性がある場合に、PRINTS Category を 表示します。
PRINTS	この遺伝子が PRINTS ドメインと類似性がある場合に、PRINTS 情報を表示 します。
EC number	この遺伝子が EC number 情報がある場合に、EC number 情報を表示します。
Pfam	この遺伝子が Pfam ドメインと類似性がある場合に、 PRINTS 情報を表示しま す。
Prosite	この遺伝子が Prosite ドメインと類似性がある場合に、PRINTS 情報を表示します。
TMs	この遺伝子が膜貫通領域と予測されたドメインがある場合に、そのドメイン数 を示します。
GPI-anchor	この遺伝子が GPI アンカーとの結合が予測された配列を含む場合に、その配列の長さを表示します。
Signal Peptide	この遺伝子がシグナルペプチドと予測された配列を含む場合に、その配列の長 さを表示します。

CDS 検索結果 (図 27) で初期状態 (デフォルト) で表示されるカラムの説明を以下にいた

します。

Gene ID	遺伝子の ID を示します。 NCBI 由来のゲノムデータの場合は、 locus tag 名を示 します。
Original description	データソースに記載されていた遺伝子に対する説明を示します。
Assigned description	SYGD データ登録時に再アノテーションされた遺伝子に対する説明を示します。
Genome	ゲノム名を示します。リンクをクリックすると、ゲノムトップ画面 (図 3) へ移 動します。
Chromosome	染色体情報のある場合に、染色体番号を示します (染色体情報が無い場合には "Unmapped"と示されます)。
Links	外部データベースへのリンクです。
Protein Sequence	アミノ酸配列が FASTA 形式で表示されます。
Homologous genes	SYGDに登録された他のゲノムと相同性のある遺伝子を表示することができます。 セレクトボックスから任意のゲノムを選択し、"Apply"ボタンを押すことで相同 性のある遺伝子が表示されます。

CDS 検索結果 (図 24) で選択することにより表示されるカラムの説明を以下にいたします。

Description	
StartCodon	この遺伝子のスタートコドン、そのコドンから翻訳されるアミノ酸、遺伝子
	産物のアミノ酸配列長が表示されます。

CDS length	この遺伝子の CDS の配列長が表示されます。
CDS sequence	この遺伝子の塩基配列が FASTA 形式で表示されます。
Upstream sequence	この遺伝子の上流の配列を表示します。テキストボックスに数値 (半角数字) を入力し、"Apply"ボタンを押してください
	この遺伝子の下流の配列を表示します。テキフトボックフに数値(半角数字)
Downstream sequence	を入力し"Apply"ボタンを押してください。
Ortholog	この遺伝子のオーソログ (パラログ) 情報がある場合、シンテニー解析画面
-	(図 14) へのリンクが表示されます。
COG	この遺伝子が COG (オーソログ) 情報を含む場合、COG ID を表示します。
COG funtion	この遺伝子が COG (オーソログ) 情報を含む場合、COG functional categories
	を表示します。
KOG	この遺伝子が KOG (オーソログ) 情報を含む場合、KOG ID を表示します。
KOG Function	この遺伝子が KOG (オーソログ) 情報を含む場合、KOG functional categories
	を表示します。
PRINTS Category	この遺伝子が PRINTS ドメインと類似性がある場合に、PRINTS Category を表示します。
PRINTS	この遺伝子が PRINTS ドメインと類似性がある場合に、PRINTS 情報を表示します。
EC number	この遺伝子が EC number 情報がある場合に、EC number 情報を表示します。
Pfam	この遺伝子が Pfam ドメインと類似性がある場合に、PRINTS 情報を表示します。
Prosite	この遺伝子が Prosite ドメインと類似性がある場合に、PRINTS 情報を表示します。
TMs	この遺伝子が膜貫通領域と予測されたドメインがある場合に、そのドメイン 数を示します。
GPI-anchor	この遺伝子が GPI アンカーとの結合が予測された配列を含む場合に、その配列の長さを表示します。
Signal Peptide	この遺伝子がシグナルペプチドと予測された配列を含む場合に、その配列の 長さを表示します。

2-7-3. オーソログ検索

SYGD に登録されているゲノム情報の中から、オーソログ (パラログ) である遺伝子群 (オーソログクラスター) を検索することができます。

オーソログ検索画面を図28、コンティグ検索結果表示画面を図29 に示します。

オーソログ検索画面は、基本検索画面 (図 21) の "Ortholog" リンクをクリックすることで 表示されます。

Top / Search (Ortholog)	op / Search (Ortholog)					
Search	Search K					
Select Target: Contig Gene RNA CDS Ortholog COG KOG EC						
Search	Search					
Orthlog Sear	ch : Hide Filter Option					
ID						
Ortholog Set	select 🗘					
Genome All Positive All Negative All Ignore	- Genome P N - S288c O O K7 O O					
Differential	Num. of group select C Group Area Venn's diagram					
Search Rese	Search Reset					

図28: オーソログ検索画面。基本検索画面の "Ortholog" をクリックする (図の矢印) ことで、オーソログ検索画 面を表示させることができます。基本検索画面の表示方法については、 2-7 基本検索参照 。表中にある項目に検 索条件を入れて、"Search" ボタンを押すと、条件に合致したオーソログクラスターのリストが表示されます。

図25 のオーソログ検索画面には、4 つの項目 (ID, Ortholog Set, Genome, Differential) が あります。それぞれについて以下に説明します。これらは、組み合わせて検索することも 可能です。

ID	オーソログ内部ID を入力して検索します。 (図 13 における ort1_c70_e100_7960 等の ID を入力します)
Ortholog Set	以下の Genome, Differential で用いるオーソログ計算方法をメニューから選択します。
Genome	ゲノムリスト画面 (図2) にあるゲノムが表示されています。それぞれのゲノムにつ いて、P(positive)/N(negative)/ー(don't care) が選択できるようになっています。
Differential	差分解析を行います。No. of group のメニューを選択 (1, 2 or 3) すると、ベン図が 現れます (図32) ので、グループA, B (group=3 を選択したときは C も) に、同じグ ループに入れたいゲノムを選択します。Area では、右ベン図の領域を選択します。

Top / Search (Ortholog)

earch							
rthlog	Search :	Hide Filter Option					
ID							
Ortholog :	Set ort1	1_1					
Genome All Posi All Neg All Igne	itive gative ore S28	nome P N - Genome P 38c O O O K7 O	N -				
Differenti	ial Num	n. of group select ᅌ	Sroup Area Ven	n's diagram			
Search	Reset	513	Rows: 50	Page: 1	Top Bottom	Page: 1. 2. 3. 4. 5. 6	
Search Irthlog lis	Reset	513 Ortholog Set	Rows: 50 Genome count	Page: 1 Gene count S 2 K 8 7 L c	Top Bottom	Page: 1, 2, 3, 4, 5, 6	
Search Irthlog lis No./5513	Reset st 1 - 50 of 55 ID ort1_1_4725	Ortholog Set	Rows: 50 C Genome count	Page: 1 Gene Count S K 2 K 8 7 L L 1 1	C Top Bottom Annotation Repression factor of MSEs protein 1	Pege:1, 2, 3, 4, 5, 6	

図29: オーソログ検索画面.。ID にあるリンクをクリックするとシンテニーマップ (図14, 詳細は 2-3-4 シンテニ ーマップの表示を参照) が示されます。Gene countの数値のリンクをクリックすると、選択した種でこのオーソ ログID を持つ遺伝子を検索した結果が表示されます。

/ Search (Ortholog	9)					
earch						
Select Target: Carrig Gene INA COS Ortholog COS KOS EC						
Search						
Orthlog Sear	rch : 🕞	de Filter Option				
ID						
Ortholog Set	ort1_c70	_e100 🔁				
All Positive All Negative All Sphore	- Genome \$288c	P N - Genome O O O K7	P N -			
			Group		Area	Venn's diagram
Differential	Num. of g	ytagi 2 📑	A 47. Sacharemper convision 5288 : Sacharemyces convi 8 47. Sacharemyces convision 5288 : Sacharemyces convision 5288	Kyokai no. 7 siae 5288c Kyokai no. 7 siae 5288c	Area Courr AB 4442 AB 4442 AB 22 BB 24 All 4464	Ab AB aB
Search Res	et					
Orthiog list 1 - 7	of 7		Rows: 50 Page: 1	9	Тор Во	ttom Page: 1
No./7 ID		Ortholog Set		Genome count	Gene count S K A S 7 L C	saton
1 ort1_c70_	_e100_4458	ort1_c70_e100 (thresho	ids of cover=0.7,evalue=1e-100)	1	0 3 3 Pred	icted transporter (major facilitator superfamily)
2 ort1_c70_	_e100_4459	ort1_c70_e100 (thresho	ids of cover=0.7,evalue=1e-100)	1	0 2 2 Pred	icted transporter (major facilitator superfamily)

図 32: オーソログ検索画面 (Differential. 抜粋). Differential の項目で Num. of group を選択すると、上のようなベ ン図が現れます (この図ではgroup=2と選択したため、ベン図は図のようになる)。Group A, B としてゲノムを複数 選択 (Ctrl キーを押しながらマウスクリック, で複数選択できる) します。 Area 項目で Ab を選択し検索を実行しますと、図のように選択した領域 (ここでは Ab) が色つきで示されます。 Area, ベン図に記してある数値はそれぞれの条件に合致したオーソログクラスターの数です。

注: 各領域に表示されているラベル以下のような意味である: 大文字: そのグループに存在している。 小文字: そのグループに存在していない。 例) AB = (グループA に存在) かつ (グループ B に存在) Ab = (グループA に存在) かつ (グループ B に存在しない) aB = (グループA に存在しない) かつ (グループ B に存在) ab = (グループA に存在しない) かつ (グループ B に存在)

2-7-4. COG、KOG、EC 検索

SYGD に登録されているゲノム情報の中から、COG, KOG, EC アノテーションについての検索も可能です。

COG検索画面を図31に示します。KOGやEC検索画面も、ほぼ同じ仕様となっています。 COG検索画面は、基本検索画面 (図21) の "COG" リンクをクリックすることで表示され ます。同様に、KOG検索画面は"KOG", EC検索画面は "EC" をクリックすることで表示さ れます。

ip / Search (COG)													
Search	iearch												
Select Target: Con	Select Target: Contig Gene RNA CDS Ortholog COG KOG EC												
Search	Search												
COG Search :	Hi	de Filt	er Op	tion									
ID													
	-												
COG Functions	□1	□A	□ĸ	□L	⊡в	D	٦Y	٧	ਾ	⊡м		□z	□W
	U	0	C	G	E	OF	□н	I	P	□Q	□R	S	
Genome	-												
All Positive Genome P N - Genome P N -													
All Negative													
All Ignore	528	RSC		~ [*		/	10	\sim	•				

図31: COG検索画面. 基本検索画面の "COG" をクリックする (図の矢印) ことで、COG検索画面を表示させることが できます。同様にKOG検索画面は "KOG" を、EC検索画面は "EC" をクリックすることで、それぞれの検索画面を表 示させることができます。基本検索画面の表示方法については、 2-7 基本検索参照。表中にある項目に検索条件を入 れて、"Search" ボタンを押すと、条件に合致したオーソログクラスターのリストが表示されます。

図31 のCOG検索画面には、3つの項目(ID, COG Functions, Genome) があります。それぞ

れについて以下に説明します。これら	は、組み合わせ(快案9ることも可能で9。
-------------------	----------------------

ID	COG ID を入力して検索します。
COG Functions	COG functional categories で検索を行います。
Genome	ゲノムリスト画面(図2)にあるゲノムが表示されています。それぞれのゲノムについて、P(positive)/N(negative)/-(don't care)が選択できるようになっています。
	例えはクノムeco1, eco2には存在し、cgi1には存在しない、bsu1には存在していて も存在していなくとも良いといったオーソログクラスターを検索して表示します。

これらの検索画面は、オーソログ検索画面(図 29)と同様ですので、オーソログ検索(2-7-3 オーソログ検索)を参照してください。