ORIGINAL PAPER

Sorel Fitz-Gibbon · Anthony J. Choi Jeffrey H. Miller · Karl O. Stetter · Melvin I. Simon Ronald Swanson · Ung-Jin Kim

A fosmid-based genomic map and identification of 474 genes of the hyperthermophilic archaeon *Pyrobaculum aerophilum*

Received: October 18, 1996 / Accepted: October 30, 1996

Abstract We have constructed a physical map of the approximately 1.7-Mb genome of the hyperthermophilic archaeon Pyrobaculum aerophilum. Derived from a 12× coverage genomic fosmid library with an average insert size of 36Kb, the map consists of a single circular contig of 96 overlapping fosmid clones with 211 markers ordered along them. One hundred of the sequence markers have strong similarities to known genes. Many overlaps were also checked using restriction fingerprint analysis. This map is an important step in the elucidation of the sequence of the entire genome of Pyrobaculum aerophilum. To this end we have determined more than 95% of the genome with 15000 random sequences. Each sequence has been screened against the public sequence databases to identify similarities to known genes. We report here a list of the 474 putative genes we have identified.

Key words *Pyrobaculum aerophilum* · Genome · Hyperthermophile · Archaea · Eocyte

Communicated by J. Wiegel

S. Fitz-Gibbon · A.J. Choi · J.H. Miller (⊠) Department of Microbiology and Molecular Genetics and The Molecular Biology Institute, University of California at Los Angeles, 1602 Molecular Sciences Building, 405 Hilgard Avenue, Los Angeles, CA 90095-1489, USA Tel. +1-310-825-8460; Fax +1-310-206-3088

e-mail: jhmiller@ewald.mbi.ucla.edu

K.O. Stetter Archaeenzentrum, Regensburg University, 93053 Regensburg, Germany

M.I. Simon \cdot R. Swanson¹ \cdot U.-J. Kim Division of Biology, 147-75, California Institute of Technology, Pasadena, CA 91125, USA

Present address: ¹Recombinant Biocatalysis Inc., Philadelphia, PA, USA

Introduction

Pyrobaculum aerophilum is a rod-shaped hyperthermophilic archaeon isolated from a boiling marine water hole (Völkl et al. 1993). Due to its distinctive characteristics, *P. aerophilum* is both an interesting and a suitable organism for evolutionary and comparative studies. It has an extremely high optimal growth temperature of 98°C, and yet, in contrast to other hyperthermophiles, it is facultatively aerobic. Furthermore, P. aerophilum can be plated to form colonies within four days with up to 100% efficiency (Völkl et al. 1993). Since strictly anaerobic metabolism predominates at the upper temperature limits of life (Robb and Place 1995), cultivation of hyperthermophiles has been difficult and consequently they are only poorly described. Based on 16S rRNA analysis, hyperthermophiles are the living organisms with the greatest similarity to ancient organisms (Stetter 1992; Woese 1987). Thus hyperthermophiles are of great interest and, due to its ease of cultivation, P. aerophilum is an ideal candidate for development as a model hyperthermophilic organism. We plan the construction of a genetic system which will allow for mutagenesis and reverse genetic studies, and for the analysis of gene function. Determining the entire genome sequence will simplify the development of a genetic system for P. aerophilum and will provide valuable data not only for research on hyperthermophily but also for phylogeny research and the elucidation of the least studied domain of life, Archaea.

Our strategy for sequencing the entire genome of *P. aerophilum* has been primarily a random "shotgun" approach. We have sequenced random fragments from the genome, used computer algorithms to compare the sequences, and then assembled the overlapping sequences into contigs. While the "shotgun" sequencing approach (Fleischmann et al. 1995) is an efficient way to obtain most of the sequence, it becomes impractical to complete the sequence by this method. To close the final gaps in the sequence, a directed approach is desirable. We have developed a genome map in order to facilitate a directed approach to sequence

contigs, and to aid in resolving repetitive and otherwise difficult regions.

Results

We have constructed a genomic map consisting of 211 markers on 96 overlapping large clones. The large inserts (~36 Kb) are carried in the fosmid vector pFOS1. This vector uses the *E. coli* F-factor origin of replication and thus is kept at a low copy number. This vector has been shown to maintain large insert DNA fragments with relatively high stability (Kim et al. 1992). The entire genomic fosmid library consists of 768 clones, which corresponds to a 12× genomic coverage based on an estimated genome size of 1.7 Mb. This level of redundancy was sufficient to cover the whole genome without gaps. By hybridizing the library with various probes we were able to place markers on the map and determine the overlaps of the fosmids.

Construction and screening of fosmid library

Genomic DNA was partially digested with *Sau*3AI before ligation to the *Bam*HI cloning site of the fosmid arms. The ligation reaction contained a large molar excess of fosmid arms over genomic DNA in order to minimize the chance of chimera formation. Packaging was done using the Stratagene GigaPack XL kit which preferentially packages large inserts. The average insert size for this library was 36Kb as determined by *Not*I digestion and pulsed field gel electrophoresis (not shown). The entire library was gridded onto single small ($8 \text{ cm} \times 12 \text{ cm}$) nylon membranes using the Biomek 1000 laboratory workstation. The library was screened by a variety of probes that are described in detail later.

Fosmid contig assembly and verification

Three methods were used to identify overlapping fosmids. First, small inserts from the pUC18 clones were amplified by PCR and hybridized to high-density colony membranes of the large insert fosmid library. As part of our large-scale random sequencing project, both ends of these pUC18 inserts had previously been sequenced and compared to the public sequence databases. Of the 146 markers on the map, 112 were putative genes. These include ribosomal genes, DNA and RNA polymerase genes, tRNA genes, various repair genes, recA homologs, and numerous metabolic genes (Fig. 1). Analysis of these data revealed several ambiguous regions, possibly due to repetitive sequences in the markers. These regions were clarified by a fingerprinting method consisting of computer assisted analysis of the banding patterns of fosmids digested with BanI and MspI and labeled at the BanI ends (Coulson et al. 1986; Sulston et al. 1988). Twenty-four overlaps were confirmed by this method (Fig. 1), yielding 10 clear contigs. The final 10 gaps were closed using RNA probes transcribed from the T7 and Sp6 promoters flanking the fosmid vector cloning site. These RNA probes were radiochemically labeled and hybridized back to the whole fosmid library. Prior to generating the probe, the fosmid DNA was digested with HincII, which yields DNA fragments averaging approximately 1000 bp. HincII was chosen over a less frequent cutter, to keep the length of the labeled transcript short. This was important for minimizing the inclusion of repetitive sequences in the riboprobes. We found larger probes often gave ambiguous results. Using this method we successfully identified fosmids spanning the 10 remaining gaps and confirmed 38 other overlaps (Fig. 1). In three cases, a second round of probing was necessary to cross the gaps, using probes from the ends of the fosmids identified in the first round. An overlapping set of 96 fosmids was identified, giving a redundancy of approximately 2×. From the average size of the 96 clones in the map and the overall redundancy that they represent, we estimate the genome size of this microbe to be 1.7 Mb.

Genome sequence

The map reported here will be used as an aid to determining the entire genome sequence of P. aerophilum. The sequencing effort has so far yielded more than 15000 random singlepass sequences which were determined using fluorescent multiplex automated sequencers (ABI 373, Applied Biotechnology). Base calls were made from the ABI trace data using the Phred software package (Brent Ewing). This software not only assigns bases from the trace data but also assigns a quality value for each base call. This factor becomes extremely important when the sequences are assembled, obviating the need for extensive manual editing of contigs. The Phrap software package (Phil Green) makes use of the quality information provided by the Phred software, while other assembly programs allow regions of low-quality data to disrupt the consensus sequence. The Phrap software then uses a novel method of assembly to create contigs which are a mosaic of the highest quality parts of reads (Table 1).

After having completed the random sequencing phase of the project, we have slightly over 800 contigs and thus approximately 800 gaps in our sequence. Poisson analysis suggests an average gap size of less than 200 bases long. Several methods are being used in parallel in order to move quickly to a finished sequence. Where possible, new sequences are being derived from the adjacent regions of partially sequenced pUC18 clones. In most cases an oligo primer has to be designed and synthesized for this purpose. On a small number of clones, a deletion subcloning method (unpublished) is being used to reach the unsequenced region at the center of the insert while still using the standard priming sites on the vector.

Of the 800 gaps in our sequence there are approximately 200 gaps which are not represented in the pUC18 library. The fosmid map and templates will be used to extend the sequences across these gaps. The fosmid map includes 146 sequence markers which have been assembled into the se-



Fig. 1. The *horizontal lines* represent the inserts of fosmid clones covering the entire circular genome of *Pyrobaculum aerophilum*. Fosmid designations are indicated *immediately below the horizontal lines*. *Vertical dashed lines* represent sequence markers placed along the

38

genome via hybridization experiments. Where the sequence tag has similarity to a known gene, the gene name is written in *bold*. All markers were tentatively identified as the listed genes by their similarity to RNA or amino acid sequences in the public databases.





the horizontal lines represent the regions from which the riboprobes were transcribed. *Vertical solid lines* represent the 24 overlaps confirmed by fingerprinting

1	Ω
+	υ

Table 1a. Similarity data for 100 putative genes which were placed on the fosmid map (Fig. 1). All matches have a spurious match probability of less than 10^{-5} except for eukaryotic IF-4A (90). The numbers (*Map* #) correspond to the markers' position on the fosmid map (Fig. 1)

Map #	Hit identification	Probability (-E)	Accession	Organism
1	S9, 30S ribosomal	30	sp P39468	Sulfolobus acidocaldarius
4	16S ribosomal	54	gb M35966	Thermoproteus tenax
8	glutamine synthetase	35	sp P36205	Thermotoga maritima
12	chorismate synthase	37	pir \$13070	Synechocystis sp.
14	3-dehydroquinate synthase	62	sp P07639	Éscherichia coli
15	transketolase	26	gi 1149710	Clostridium perfringens
18	serine protease	26	sp P29139	Bacillus polymyxa
20	O6-methyltransferase	9	sp P16455	Homo sapiens
23	antioxidant	55	gi 1045502	Sulfolobus sp.
25	DNA polymerase	150	gi 807830	Pvrodictium occultum
26	L18. 50S ribosomal	52	sp P14033	Methanococcus vannielii
27	TF55 heat shock protein	79	sp P28488	Sulfolobus shibatae
30	DNA polymerase	200	gi 807828	Pvrodictium occultum
33	RNase PH	15	sp P37939	Mycobacterium leprae
35	agmatinase	16	sp P37819	Streptomyces clavuligerus
38	succinvl-CoA synthetase alphaβ	86	sp P25126	Thermus flavus
39	synaptic vesicle protein	37	nir \$34961	Rattus norvegicus
41	DNA Helicase recO	22	sp P30015	Escherichia coli
43	NADH oxidase	20	sp P37061	Esterococcus faecalis
46	anthranilate synthase	91	sp[006128]	Sulfolobus solfataricus
40	anthranilate phosphoribosyltransferase	69	gi 149037	Bacillus numilus
49	nicotine dehydrogenase	44	pir \$37570	Arthrobacter nicotinovorans
50	acyl-CoA dehydrogenase	33	sp P46703	Mycobacterium lenrae
52	ABC transporter	11	sp P39326	Escherichia coli
54	adenylyl sulphate reductase alpha subunit	44	ai 1183005	Desulfovibrio vulgaris
55	sulfite reductase alpha-subult	56	pir \$27470	Archaeoglobus fulgidus
58	nitrate reductase commo subunit	24	sp[P42177]	Bacillus subtilis
50 67	2 ovoisovalorate debudrogenese, beta	24 82	sp F 42177	Pacillus staarotharmonhilus
65	2-0x0isovalerate deliverogenase, beta	82 21	sp[F3/941]	Enterococcus hirac
67	polyculfido roductoso choin per 4	20	sp[1 52115]	Wolinglla sugging and
70	putrescipe transport	10	sp[F31075]	Fscharichia coli
70	argininosuscinate synthese	100	sp[F31130]	Mathanooogus yannialii
71	argininosuccinate synthase	60	sp[F15250]	Lantospira interrogans
74	NA DH ubiquinono oridoroduotoso	40	sp[141394]	Lepiospiru interroguns Bog taumug
74 75	NADH-ubiquilione oxidoreduciase	40	sp[F42020]	BOS laurus Sulfolobus solfatarious
75	aiconor denydrogenase	28	spjr 59402	Suijoiobus soijuuricus
70	SIF2	30	gi 043000	Basillus subtilis
11	DNA tongiamerasa III	62	gi 1140247	Baculus sublitis
10	DNA topoisoinerase III	41	gi 1292915	The server of th
02 00	uKNA	12	ail1072122	Homo aming
90	eukaryoue IF-4A	3	gi 10/2122	Mothers a base of the second s
92	mur Y = adenine giycosylase	37 14	sp P29388 ail1165245	Meinanobacierium inermo.
95	cytosine specific DNA metnyitransierase	14	g1 1103243	Nelsseria gonorrhoeae
94	raasi	24	pir 53/6/3	Homo sapiens
95	replication protein A	7	sp P03082	Agrobacierium rnizogenes
98	inosine monophosphate denydrogenase	1	pir B48868	Methanosarcina thermophila
99 102	pnospnogiycerate denydrogenase	55	g1 1146196	Baculus subtilis
105	ribose transport, nign amnity	23 21	g1/290599	Escherichia coli Mucch actarium 1
100	giucose-1-phosphate adenyiyitranterase	21	g1 099151	Mycobacterium leprae
10/	L13, JUS FIDOSOMAI	8	sp P14032	Methanococcus vannielu
110	235 ribosomal KNA	228	emb[X05480]	Desulfurococcus mobilis
111	molybaopterin biosynthesis	44	gi 1002858	Anabaena sp.
113	str2	21	g1 845686	Staphylococcus aureus
114	pnospno- <i>N</i> -acetylmuramoyl-pentapeptide-transferase		sp Q03521	Bacillus subtilis
117	citrate synthase	64	sp P39120	Bacillus subtilis

Table 1a. Continued

Map #	Hit identification	Probability (-E)	Accession	Organism
118	DNA excision repair helicase, XP-bc/ERCC-3	14	gi 902048	Arabidopsis thaliana
119	nitrate reductase beta	157	sp P42176	Bacillus subtilis
121	5S ribosomal RNA	40	gb M16530	Sulfolobus sp.
124	subtilisin	61	sp P29139	Bacillus polymyxa
131	diadenosine tetraphosphatase (mutT-like)	18	gi 1054947	Sus scrofa
134	superoxide dismutase (Mn)	19	sp Q08713	Sulfolobus acidocaldarius
136	glutamine transport protein glnQ	46	sp P10346	Escherichia coli
138	precorrin methylase	38	sp Q05630	Salmonella typhimurium
139	precorrin decarbocylase	18	sp Q05632	Salmonella typhimurium
144	indole-3-glycerol-phosphate synthase	19	sp Q06121	Sulfolobus solfataricus
145	porphobilinogen deaminase	52	sp P16616	Bacillus subtilis
147	dihydroxyacid dehydrase	43	sp P31959	Clostridium pasteurianum
148	X-PRO aminopeptidase	41	gi 1303901	Bacillus subtilis
149	rad51	91	gi 1054624	Xenopus laevis
150	ATPase, alpha, membrane associated	126	sp P09639	Sulfolobus acidocaldarius
151	RNA polymerase B	94	sp P11513	Sulfolobus acidocaldarius
152	S9, 30S ribosomal	30	sp P39468	Sulfolobus acidocaldarius
153	rad2	24	sp P26793	Saccharomyces cerevisiae
156	S7, 30S ribosomal	55	sp P41206	Desulfurococcus mobilis
159	pyruvate kinase	30	gi 1041097	Bacillus psychrophilus
162	DNA ligase	32	sp Q02093	Desulfurolobus ambivalens
163	DNA ligase	122	sp Q02093	Caenorhabditis elegans
164	DNA replication licensing factor (MCM2, cdc46)	98	gi 852053	Drosophila melanogaster
165	cdc21	5	sp P33991	Homo sapiens
168	pyruvate synthase	43	pir \$22396	Halobacterium halobium
170	16S ribosomal	47	gb M38637	Thermoplasma acidophilum
171	L-asparaginase	34	gi 496102	Lupinus albus
172	ATP phosphoribosyltransferase	21	gi 1117923	Yarrowia lipolytica
173	hisF (cvclase) (histidine biosynthesis)	70	sp O02133	Lactococcus lactis
176	23S ribosomal RNA	26	gb M86626	Pvrobaculum occultum
177	recF	8	sp P24900	Salmonella thyphimurium
178	ATP-binding protein	23	gi 1184189	Escherichia coli
179	phosphoglycerate dehydrogenase	37	sp P37666	Escherichia coli
180	pyruvate ferredoxin oxidoreductase	55	gi 1197358	Pvrococcus furiosus
184	Cu transporting ATP-binding <i>nosF</i>	11	sp P19844	Pseudomonas stutzeri
185	3-hydroxyisobutyrate DH	47	gi 1001605	Synechocystis sp.
188	RNAse HII	9	pir \$53\$08	Saccharomyces cerevisiae
189	nucleotide-binding protein <i>ugpC</i>	80	sp P10907	Escherichia coli
192	polysulfide reductase chain C. psrC	28	sp P31076	Wolinella succinogenes
196	cvtochrome <i>b</i>	19	gi 927525	Sulfolobus acidocaldarius
197	eIF-2 beta subunit	6	sp[P41375]	Drosophila melanogaster
201	asp-tRNA	9	gb[L07321]	Thermococcus celer
202	vagX	11	gi 1303871	Bacillus subtilis
206	NADH-quinone oxidoreductase	25	sp P29925	Paracoccus denitrificans
207	DNA excision repair helicase XP-bc/ERCC-3	16	oi 902048	Arabidonsis thaliana
210	adenvlate kinase	29	gi 1086550	Methanococcus jannaschii
211	phosphoglycerate kinase	69	sp P20971	Methanothermus fervidus

Approximately 15 000 sequences were screened individually against the nonredundant set of public databases using BLAST (Altschul et al. 1990; States et al. 1991) with default settings. The spurious match probabilities (as determined by the BLAST algorithm) are reported as orders of magnitude, e.g., 3.6E-24 becomes 24. All of the matches reported have spurious match probabilities of less than 10^{-5} ; most of the matches reported have spurious match probabilities of less than 10^{-10} . Detailed analyses of matches have not been performed. Duplicate copies of genes are not known or reported. The identification and categorization of the matches should be considered tentative. Assignments are based upon Monica Riley's categories (Riley 1993; see also Bult et al. 1996). The cited names of the organisms are as they appear in the data entry of the gene bank and thus do not necessarily reflect the present correct nomenclature of the listed organisms. Upon completion of the sequence, extensive analysis will be performed and reported. We are developing a web site on which our data will be publicly available. E-mail to sorel@ewald.mbi.ucla.edu for web site information or for help with specific data.

Table 1b. Similarity data for 474 Pyrobaculum aerophilum random sequence tags

ID #	Hit identification	Probability (-E)	Accession	Organism
Amino ac	id biosynthesis		1	
1	glutamate-1-semialdehyde 2,1 aminomutase	24	gp X53695	Synechococcus sp.
2	NMHase	27	gp A17961	Arthrobacter sp.
3	ornithine aminotransferase	12	gp X81802	Bacillus subtilis
4	S-adenosyl-L-homocysteine hydrolase	81	gp Z50174	Sulfolobus solfataricus
5	2-isopropylmalate synthase	42	sp P05342	Lactococcus lactis
6	2-oxoisovalerate dehydrogenase, beta	43	sp P3/941	Bacillus subtilis
7	3-phosphoshikimate 1-carboxyvinyltrans.	13	sp P24497	Klebsiella pneumoniae
8	anthranilate phosphoribosyltransferase	45	gp M65060	Methanobacterium thermo.
9	anthranilate synthase I	68	sp Q08653	I nermotoga maritima
10	anthrannate synthase II	52	gp[M98048]	Suijolodus soljalaricus
11	chorismate mutase	9	sp P2/005	Pseudomonas siuizeri
12	dibudrowy goid debudrotogo	54 21	gp[A0/310]	Synechocysus sp.
15	hisE	51	sp r 39322	A compinitive brasilence
14	nist hist	45	sp[r20/21]	Azospirilium Drasilense
15	histidinal phosphata transaminasa	11	sp[Q02132]	Euclococcus lucits Escherichia coli
10	imidazolegiyeerol phosphate debydratase	20	ph [A 30270]	Saccharomycas caravisiaa
18	indole 3 alveerol phosphate synthese	20	sp[100035]	Sulfolobus solfataricus
10	ATP phosphoribosyltransferase	15	sp[Q00121]	Lactococcus lactis
20	ketol-acid reductoisomerase	51	$p_{1} = \frac{3p_{1} + 20212}{2}$	Corvnehacterium alutam
20	kynurenine/alpha-aminoadinate aminotrans	22	gn Z 50144	Rattus norvegicus
21	<i>N</i> -acetylornithine aminotransferase	47	pir \$44189	Anahaena sp
23	phosphoribosylanthranilate isomerase $(trnF)$	7	on[X17149]	Vibrio parahaemolyticus
23	phosphoribosylawinamate isomerase (<i>iipi</i>)	15	gp M76553	Leishmania donovani
25	tryptophan synthase, alpha	5	gp M65060	Methanobacterium thermo
26	tryptophan synthase, beta	6	sp P19868	Bacillus stearothermophilus
27	acetylglutamate kinase	8	sp P36840	Bacillus subtilis
28	argininosuccinate synthetase (Ass)	51	sp P13256	Methanococcus vannielii
29	aspartate semialdehyde dehydrogenase	35	pir A44846	Leptospira interrogans
30	aspartate aminotransferase	12	gp D50624	Streptomyces virginiae
31	aspartate beta-semialdehvde dehvdrogenase	27	gp M77500	Leptospira borgpetersenii
32	aspartate kinase-homoserine dehydrogenase	15	gp[L33912]	Zea mays
33	ATP phosphoribosyltransferase	20	gp[U07830]	Schizosaccharomyces pombe
34	cystathionine gamma-synthase	17	sp P00935	Escherichia coli
35	L-isoaspartyl protein carboxyl methyltrans.	32	gp M63493	Escherichia coli
36	lysine-sensitive aspartokinase III	20	sp P08660	Escherichia coli
37	dihydrodipicolinate synthase (mosA)	15	sp Q07607	Rhizobium meliloti
38	succinyl-diaminopimelate desuccinylase	8	sp P24176	Escherichia coli
39	homoserine kinase (<i>thrB</i>)	16	gp Y00522	Calothrix sp.
40	3-isopropylmalate dehydrogenase	37	gp U07980	Bos taurus
41	acetolactate synthase, large subunit	41	sp P37251	Bacillus subtilis
42	branched-chain-amino-acid-transaminase	39	pir S30668	Escherichia coli
43	branched-chain amino acid transport BRAF	17	sp P21629	Pseudomonas aeruginosa
44	branched-chain amino acid transport BRAG	23	sp P21630	Pseudomonas aeruginosa
45	ketol-acid reductoisomerase	56	gp L03181	Bacillus subtilis
46	argininosuccinate lyase	14	sp P11447	Escherichia coli
47	argininosuccinate synthetase	16	gp M21315	Methanococcus vannielii
48	ferredoxin-dependent glutamate synthase	32	gp U03006	Spinacia oleracea
49	glutamate-1-semialdehyde 2,1 aminomutase	25	pir A35789	Hordeum vulgare
50	glutamate dehydrogenase	34	gp L19995	Thermococcus litoralis
51	glutamine synthetase	46	gp[X60160]	Thermotoga maritima
52	L-giu-D-fructose-o-phosphate amidotrans.	31	gp U21932	Bacillus subtilis
55 54	NADP-specific glutamate denydrogenase	52	sp P39475	Sulfolodus snidatae
54	aspartate carbamoyitransierase	14	sp P19930	Serralia marcescens
33 56	nydroxymetnyllfans.,5-metnyl-2-0x00ut.	46	sp[P31057]	Escherichia coli Muccollagua conitalium
50	serine O costultronsforoso	12	gp[002131]	Mycopiusmu genuulum Eacharichia coli
59	thiorodoxin roductoso	12	sp[F03790]	Escherichiu coli Eubacterium acidamin ophilum
50	tinoredoxin reductase	19	pii 356966	Еибистенит испитторниит
Biosynthe	esis of cofactors, prosthetic groups, and carriers			
59	3-oxoacyl [acyl-carrier protein] reductase	21	sp P27582	
60	adgA	18	gp X59399	Rhodobacter capsulatus
61	cobyrinic acid a,c-diamide synthase	19	sp P21632	Pseudomonas denitrificans
62	dihydroflavonol 4-reductase	24	pir S38474	Lycopersicon esculentum
63	mRNA for dihydroflavonol-4-reductase	12	gp X15536	Antirrhinum majus
64	pqqF/A/B/C	9	gp X87299	Pseudomonas fluorescens
65	precorrin 3 methylase	43	sp Q05590	Salmonella typhimurium
66	precorrin-8W decarboxylase	10	sp Q05632	Salmonella typhimurium
67	GTP cyclohydrolase I	27	gp X85954	Campylobacter jejuni

Table 1b. Continued

68 GTP cyclohydnionen dominase 23 gpP17621 Bacillus subdits 70 uroporphysini-III C-methytransferase 14 spP12929 Bacillus subdits 71 lipotet hosynthesis A 19 gpP125028 Bacillus subdits 72 thiamine biosynthesis C 7 gpD11750 Zeer mays 73 mobiberum collator biosynthesis B 12 gpP121021 Katerochia coll 74 mobiberum collator biosynthesis B 12 gpD121071 Max moredus 76 robolistum collator biosynthesis B 12 gp1210271 Max moredus 77 and station synthase, bota 12 gp1210275 Max moredus 78 crythromycin bosynth sensory transductin 19 sp1299023 Hacillus subdits 80 taminin receptor honolog 25 gp135366 Hamon suptase 81 transmendynae 24 gp10009 Exclametha actillus subdits 82 ankyrin gp123001 Drocophila indigenate gp123558 84 transmendynae 13 sp121756 Bacillus subdits 85 prolinebethaine	ID #	Hit identification	Probability (-E)	Accession	Organism
60porphobilingen deaminase23gpMX7078Beatilus megarium71uroporphytric III C-methylamaderase14spP2928Bacilus megarium72inpost biosynthesis A17gpU2088Iterarophilus influencese73molybdopterin biosynth. (mocA)25gpV22151Excheribia coli74molybdopterin biosynth. (mocA)15gpV22181Excheribia coli75molybdopterin biosynth. sensory transduct19gpV22181Excheribia coli76erylitomycin biosynth. sensory transduct19gpV2321Bacilus sublifis77erylitomycin biosynth. sensory transduct19gpV2321Bacilus sublifis78erylitomycin biosynth. sensory transduct17gpU210561Bacilus sublifis79gatama-aminobutyrate permease9gpU211561Bacilus sublifis81integra inchronare phosphoprotein band 7.2b41gpU2301Drosophia melanogaster82rolitorie biosing interior7spP13031Salenacing and	68	GTP cyclohydrolase II (<i>ribA</i>)	14	sp P17620	Bacillus subtilis
70incopriprint II C-methyltransferase14 \mathfrak{sp} P2928]Boolibs influenzae71lipote biosynthesis7gpU12508Zen mays72chiE/M7gpU1750Zen mays73chiE/Mripote biosynthesis7gpU1750Zen mays74chiE/Mripote biosynthesis82gpU1758Exchericha coli75molyblemum colicitor biosynthesisB32gpU32510Mass musculus76ribotemin synthase, heta14gpU23510Mass musculus78erythromycin biosynth. sensory transductin19spU32568Boolibus solibits80laminin receptor homolog25gpX55061Mass musculus81transmenbrane24gpU00059Excherichia coli82ankyrinmar ghorphoprotein hand 7.2b14gpX55061Drosophia micropaster83prolinichestane transporter37spPX10758Drosophia micropaster84oplopridic permese7gpX10501Drosophia micropaster85prolinichestane transporter37spPX10758Drosophia micropaster86otiopridic permese11spP10758Boolibus tobilits90earthonic anhylyrase11spP17581Excherichia coli87membrane spanning25gj(007301Brucella micropaster88prolonic anhylyrase11spP17581Excherichia coli91EXC topridic permese26gpX7064Boolibus tobilis<	69	porphobilinogen deaminase	23	gp M57676	Bacillus subtilis
71lipoute biosynthe isogene19gpU12388I Incomphilus influenzace72thillingExchericha coli22gpU113151Exchericha coli73andybbarni obiesynth (morch)12gpU113161Exchericha coli76ribollovin synthase, beta14gpU123810Hacmophilus influenzace77Arythomycin biosynth, censory transduct9gpU1231576Bacillus subtilis78arythomycin biosynth, censory transduct9gpU1231561Bacillus subtilis79gamma-aninobutyrate permasse9gpU137561Bacillus subtilis79gamma-aninobutyrate permasse9gpU137561Bacillus subtilis81transmembrane24gpU107397Mus musculas82ankyrin77gpU17397Mus musculas83integral membrane glutamyl-phosphoprotein band 7.2b41gpU172971Mus musculas84nikyrin7gpU17397Mus musculas85membrane synaming25gp607906Bacillus subtilis86nikygperbide permase21gp077666Bacillus subtilis87membrane synaming25gp1607301Braculus subtilis88n-acetyl-gamma glutamyl-phosphate reduct12sp07976691ERV operon: (r.X15sp172331Exchericha coli92subdisshiftgg12331Exchericha coli93subdisshiftgg123321Exchericha coli94gg123331Exchericha	70	uroporphyrin-III C-methyltransferase	14	sp P29928	Bacillus megaterium
72thamine biosynthetic7gpU17151Zea mays73 $obleynth_{10}$ 22gpW11151Excheribla coli74malybioquerin biosynthesis B32pp12511Excheribla coli75molybioquerin colicion biosynthesis B32pp125211Excheribla coli76molybioquerin biosynth, samory transductin19gp125201Haemophilas coli77A10gp125201Mas musculusgp125201Haemophilas coli78expliring philosynth, samory transductin29gp125301Haemophilas80laminin receptor homolog25gp1253061Homo sapiers81transmembrane24gp112021Mas musculus82ankyrin17gp135611Drosophila melanogaster84polysulfide reductuse A7gp1217551Wolzella accinogenes85prolino-bitain transporter37gp124041Salamonella colicia scanagenes86enligopeptide permease7gp127753Baeillas stahlis87explorino-fitain colicia scanagenes7gp127821Excheribia coli88n-acctj-gamma ghtamj-phospharterduct12gp007906Baeillas stahlis89earbon dioxide binding25gj1607301Brucella stahlis89earbon dioxide binding25gj1607301Brucella stahlis89earbon dioxide binding25gj1607301Brucella stahlis80earbon dioxide binding25gj1607301Brucella stahlis </td <td>71</td> <td>lipoate biosynthesis A</td> <td>19</td> <td>gp U32688 </td> <td>Haemophilus influenzae</td>	71	lipoate biosynthesis A	19	gp U32688	Haemophilus influenzae
73chlE/N22gp/M2151Excherioba coli75molybidopterin biosynth. (mocA)15spP12281Excherioba coli75molybidopterin biosynth. sensory transduct14gp/U22810Homophilus influenzae76A 1026gp/E3061Mas musculus78erythromycin biosynth. sensory transduct19spP129621Bacillus subtilis80laminin teceptor homolog20gp/S0089Homo sphilus81integral membrane phosphoprotein band 7.2b41gp/U127971Mas musculus82integral membrane phosphoprotein band 7.2b41gp/U127971Mas musculus83integral membrane phosphoprotein band 7.2b41gp/U127971Mas musculus84op/S1016 reductase A7spP/S10751Woinella succinogenes85prolinchetaine transporter37spP/S10581Sadomoella typhinurium86membrane spanning25gj(4017341Bacillus subtilis87membrane spanning25gj(4017341Brucella melitansis90carhon dioxida binding25gj(4017341Brucella melitansis91ERV operosi./rvX15spP125131Excheriobia coli92stad33gp/Z5802Saccharomphilus93fa/EX (cli division control)30gp/X34594Excheriobia coli94fa/EX (cli division control)30gp/X34594Excheriobia coli95stadgp/Z5802Saccharomphranephilus9	72	thiamine biosynthetic	7	gp U17350	Zea mays
74 molybaptern hosynth: (moc4) 15 app[128] Excherible coli 75 molybaptern hosyntheses 14 gpU22010 Electricities coli 76 molybaptern hosynth: sensory transducti 14 gpU22101 Mus nusculus 77 cryphronycli biosynth: sensory transducti 19 gpP2023. Reclins sublific 77 cryphronycli biosynth: sensory transducti 19 gp1201754 Bucillis sublific 78 cryphronycli biosynth: sensory transducti 19 gp1201754 Bucillis sublific 78 gryphronycli biosynth: sensory transducti 17 gp1201754 Mus nusculus 81 transmembrane 23 gp1201754 Mus nusculus 82 anityrin 17 gp125601 Drosophila melanogator 84 polysufide reductase A 7 gp1201754 Wolnella succhasengene 85 proline-betaine transporter 37 gp120451 Sucharoniterophilas 85 polysufide reductase A 7 gp12052 Excheribia coli subanolis succharoniterophilas 86 polysufide reductase A 7 gp12051	73	chlE/N	22	gp M21151	Escherichia coli
$\begin{array}{llllllllllllllllllllllllllllllllllll$	74	molybdopterin biosynth. (moeA)	15	sp P12281	Escherichia coli
nonlowin synthese, beta14gp(12.52.10)Intermediats influences77A1026gp(12.1027)Max musculus78erythronycin hiosynth, sensory transductin19sp(12.1027)Max musculus79gamma-aminobulyrate permease9gp(13.1756)Bucillos subilits81transmembrane24gp(100039)Escherichia coli81antryrinmrane phosphoprotein band 7.2b41gp(13.1751)Wolmelangester84polysulfide roductase A7gp(13.1751)Wolmelangester85proline/betaine transporter37sp(20.9796)Bucillos succinogenes86oligopeptide permease7gp(27.1756)Bucillos succinogenes87membrane spanning26gp(27.1766)Bucillos succinogenes88m-acetyl-gamma-glutamyl-phosphate reduct12sp(07.9706)Bucillos succinogenes90earbon dioxide binding25gp(17.1862)Excherichia coli91earbon dioxide binding25gp(12.182)Excherichia coli92tradit division control)30gp(12.182)Excherichia coli93fu3*EX (cell division control)30gp(12.182)Excherichia coli94fer.37gcho06emNY.54.201Plannodium fide/parum95heat-shock54gp(12.182)Excherichia coli96spant recognition particle42sp(17.182)Excherichia coli972-hydytashptin2-4-dieme-17-dioat isom. <td< td=""><td>75</td><td>molybdenum cofactor biosynthesis B</td><td>32</td><td>pir \$31880 </td><td>Escherichia coli</td></td<>	75	molybdenum cofactor biosynthesis B	32	pir \$31880	Escherichia coli
17.26gpl[2302]Max mean due78crythronycin hisopry transduct1999F306Hadilus subilitis79gamma-aminohutyrate permesse9gpl[2306]Horn sapiens81transmembrane24gpl[200039]Excherichia coli82ankyrin17gpl[21501]Drosophila melanogaster83integral membrane phosphoprotein band 7.2b41gpl[2117297]Max meaching84polysaulide reductase A7spl[73107]Max meaching85proline/betaine transporter37spl[73107]Max meaching86ontoperide permesse7spl[73107]Max meaching87resolvig ammeglutannyl phosphate reduct12spl[007906]Bacelline skenothermophilax88resolvig ammeglutannyl phosphate reduct12spl[007906]Bacelline skenothermophilax90carbonic anhydrase11spl[731752]Excherichia coli91ERV operon: (rr.X13spl[73138]Spl[74108]92carbon dioxide binding25gpl[60790]Bacelline skenothermophilax93trike(rd.2) cpnf06gpl[73138]Excherichia coli94trike(rd.2) cpnf06gpl[73138]Excherichia coli95heatsbock54gpl[73138]Excherichia coli96signal recognition particle42spl[73138]Excherichia coli972-hydroxyhepts2-4-dione-1,7-dioate ison.19gpl[73138]Excheric	Coll onv	alana	14	gp[032810]	Huemophius injiuenzue
78crythromycin biosynth. sensory transductn19spl [P3175] <i>Bacillus subilitis</i> 79gama-animoburyns permease9ppl [3175] <i>Bacillus subilitis</i> 80laminin receptor homolog25gpl [33506] <i>Homo supins</i> 81transmembrane24gpl [00039] <i>Bacillus subilitis</i> 82ankyrin17gpl [3501] <i>Drosophila melinogaster</i> 83integral membrane phosphoprotein band 7.2b41gpl [011297] <i>Homo supins</i> 84polysulidis roductiss A7spl [3051] <i>Homo supins</i> 85prolino/betain transporter37spl [20044] <i>Bacillus subilitis</i> 86oligopeptide permease7gpl X7686 <i>Bacillus subilitis</i> 87membrane spanninggpl 23 <i>Bacillus subilitis</i> gpl X768688n=actil-jagmm-gluamyl-phosphate reduct12spl [70768] <i>Bacillus subilitis</i> 91carbon canbydrase11spl [73602] <i>Bacillus subilitis</i> 92subilitis43gpl [23007] <i>Bracella principiis</i> 93fits? LX (cell division control)30gpl [23002] <i>Bacillus cubilitis</i> 94fits? gpn606embX13502] <i>Plannolium falciparum</i> 95keat-shock54gpl [2302] <i>Bacillus cubilitis</i> 96sighal-anylase16gpl [2312] <i>Escherichia coli</i> 972-hydroxyheptin2-4.4tidne-1, d-foate isom.19spl [23132] <i>Escherichia coli</i> 98sighal-anylase16 <td>77</td> <td>A10</td> <td>26</td> <td>gp L 21027</td> <td>Mus musculus</td>	77	A10	26	gp L 21027	Mus musculus
79gimma-minohutyrate permession9gip [13:756] <i>Buillies subditis</i> 80lamitin receptor homolog25gp [35:80] <i>Homo suplers</i> 81transmembrane24gp [10:0039] <i>Excherichia coli</i> 82ankyringp [13:756] <i>Buscherichia coli</i> 83integral membrane phosphoprotein band 7.2b41gp [17:277] <i>Muscherichia coli</i> 84polysubide roductse A7sp [17:376] <i>Buscherichia coli</i> 85proline/betaine transporter37sp [17:376] <i>Buscherichia coli</i> 86elisoperpide permesse7gp [X7:756] <i>Buscherichia coli</i> 87membrane spanning26gp [X7:756] <i>Buscherichia coli</i> 89carboni anhydrase11sp [17:582] <i>Excherichia coli</i> 90carboni anhydrase15sp [27:380] <i>Buscherichia coli</i> 91EKV operon: $fr X$ 15sp [12:380] <i>Buscherichia coli</i> 92stado6cm [X7:5420]Plasmodum falciparum93fabrYEX (cell division control)30gp [12:3782] <i>Excherichia coli</i> 94(for3) cpn006cm [10:372] <i>Pyrcoccus sp.</i> 95signal recognition particle24sp [12:374] <i>Subodialis avalatica</i> 96signal recognition particle26gp [X7:852] <i>Excherichia coli</i> 972.hydroxybatyri_2.4-dinent.7.dioate isom.19sp [12:372] <i>Excherichia coli</i> 98signal recognition particle26gp [X7:852] <td>78</td> <td>ervthromycin biosynth, sensory transductn</td> <td>19</td> <td>sp[P39623]</td> <td>Bacillus subtilis</td>	78	ervthromycin biosynth, sensory transductn	19	sp[P39623]	Bacillus subtilis
80Iminin recept fomolog25ppl/S3900How suplex81transmembrane24ppl/S0030Excharichta coli82ankyrin17ppl/S1075Water methods84polysulfide reductase A7spl/S1075Water methods85polysulfide reductase A7spl/S1075Water methods86oligopeptide permease7gpl/X1756Bacillus static coli87membrane spanning26gpl/X1766Bacillus static coli88n=acetyl-gamma-glutamyl-phosphate reduct12spl/O07906Bacillus static coli89carbon dioxide binding25gpl/X1705Katellus state coli90carbon dioxide binding25gpl/X1708Excharichta coli91ERV operon fn/X15spl/S2103Excharichta coli92static colispl/X1430Baccharonyces crevisiae93fst/EX (cell division control)30gpl/X1430Baccharonyces crevisiae94fcd-3 (pnd)6gpl/X14500Plasmodini faciparam95heat-shock54gpl/Z3271Excharichta coli96signal recognition particle29gpl/X3212Excharichta coli972.hydroxyheta 2.4-diene-1.7-dioate isom.19gpl/Z3212Excharichta coli98abat-shock54gpl/Z3212Excharichta coli99eaby/CACA hydratise26gp/X3989Clostridiani difficile91abat/forxyhutyryh-CA dobe, rodione ductase <t< td=""><td>79</td><td>gamma-aminobutyrate permease</td><td>9</td><td>gp U31756 </td><td>Bacillus subtilis</td></t<>	79	gamma-aminobutyrate permease	9	gp U31756	Bacillus subtilis
81transmembrane24gpl[U0039]Excherichia coli82ankyrin17gpl[3501]Drosophia melanogoster83integral membrane phosphoprotein band 7.2b41gpl[3775]Woltnella succinogenes84polysulfale reductse A7spl[3775]Woltnella succinogenes85proline/betaine transporter37spl[37786]Bacillus subdilis86nigoeptide permease7gpl[X7785]Bacillus subdilis87membrane spanning26gpl[X7785]Bacillus subdilis88n activ/jegnima-glutamyl-phosphate reduct12spl[X7785]Bacillus subdilis90earbonic anhydrase11spl[27313]Excherichia coli91ENV operon: fr.X15spl[27313]Excherichia coli92subdision control)30gpl[X5002]Saccharomyces cerevisiae93fuk7FX (cell division control)30gpl[Z3403]Excherichia coli94(fc7) cpn606embl[X75420]Pharondium falciparum95signal recognition particle26gpl[X3722]Excherichia coli99control opharosha and saccum19spl[27414]Saalfolobus acidocaldarius91altyricobian28gpl[X3728]Nicoinat abaccun92contralium difficile10gpl[X372]Excherichia coli93indyrotybapta24-diene-1.7-dioate isom.19spl[X372]Excherichia coli94altyricosha and acidocaldarius16gpl[X373] </td <td>80</td> <td>laminin receptor homolog</td> <td>25</td> <td>gp S35960 </td> <td>Homo sapiens</td>	80	laminin receptor homolog	25	gp S35960	Homo sapiens
82ankyrin17gp [J13501]Drosphila melanogaster83integral membrane phosphoprotein band 7.2b41gp [J1797]Mus musculus84polysulidic reductase A7sp [P31075]Willenela succinogenes85polino-Detaine transporter37sp [P31075]Willenela succinogenes86oligopeptide permease7gp [X7566]Bacillus subilis87membrane spanning26gp [X7566]Bacillus subilis88 n -acetyl-gamma-glutamyl-phosphate reduct12sp [V7566]Bacillus subilis89carbon dioxide binding25gi [900730]Bracella meltensis90carbon dioxide binding25gp [2X153]Excherichia coli91ERV operon fr/X15sp [2X1540]Palsoncian diciparam92stad5pp [ZX540]Plasoncian diciparam93for FZX (cell division control)30gp [ZX540]Plasoncian diciparam95heat-shock54gp [J2372]Prococcus sp.96signal recognition particle42sp [Z714]Sulfolobus acidocaldaritis972-hydrosyheptha2-4-fine-1.7-diotati som.19sp [ZY359]Carachian adficial98alpha-anylase (mr/A)12gp [ZY352]Carachian adficial99endyrid-avdosin oxidoreductase29gp [YY359]Carachian adficial91alytosyheptha2-4-dimentary29gp [YY359]Carachian adficial91antify antify adficial colina adficia	81	transmembrane	24	gp U00039	Escherichia coli
83integral membrane phosphoprotein band 7.2b41gp[U17297]Wolsmetculas84polysulfide reductase A7sp[P2105]Wolsmetculas85proline/betine transporter37sp[P20848]Sachorizhia coli86oligopptide permease7gp[X77636Bacillus subtilis87membrane spanning26gp[X77636Bacillus subtilis88n-accvtf-gamma-glutamyl-phosphate reduct12sp[007906Bacillus subtilis90carboni doxide binding25gi[000730]Brucella melinentis91ERV operon: fr/X15sp[P2135]Escherichia coli92stad43gp[Z38002]Bacillus starohtemosphile93fs/FK (cell division control)30gp[X3498]Escherichia coli94(fc7) cpn606cmb[X75400]Parsonolium falciparum95heat-shock54gp[D2072]Pyrozencus pr96signal recognition particle42sp[P2735]Escherichia coli972-hydroxyhpta-2-4-diene-1,7-dioate isom.19sp[P3352]Escherichia coli98alpha-amylase (amyA)12gp[X3728]Nicotiani afficile100gyoxysomal cirtate synthase16gp[X3728]Nicotiani afficile101inorganic pyrophosphatse28gp[X5746]Bradyhticobiani ignoriaum102malate oxidoreductase39sp[P16468]Bradyhticobiani ignoriaum103nodulation ATP-binding 18gp[D504]Decul	82	ankyrin	17	gp L35601	Drosophila melanogaster
84polysillide reductase A/spl P310/5Wolanella succuragenes85proline/betaine transporter37spl P2084Escherichia coli86oligopeptide permease7gplX05491Salmonella typhinurium87membrane spanning26gplX7766Bacillus starohermophilus88 n -acetyl-gamma-glutamyl-phosphate reduct12splO07906Bacillus starohermophilus89carbon dioxide binding25gi[600730]Bracellu melliennis90earbon dioxide binding25gi[600730]Bracellu melliennis91EKN operon: <i>In-X</i> 15splP21331Escherichia coli92statz43gplZ2802Saccharonyces crevisiae93fp12 X (cell division control)30gplX2802Escherichia coli94(fc7) cpn606embX75420Phamodium fakiparum95hast shock24gplZ2132Pyrococcus sp.96signal recognition particle26gpX79899Clostridiau dificite972-hydroxyhepta-2/-dione isom.19splE37521Escherichia coli98apha-amylase (amy A)12gplL31327Escherichia coli99endyritase26gpX79899Clostridium dificite90gyvysomal cirate synthase16gpDX3833Klebsichia coli910gyvysomal cirate synthase16gpDX3834Klebsichia coli910gyvysomal cirate synthase16gpDX3833Klebsichia coli9	83	integral membrane phosphoprotein band 7.2b	41	gp U17297	Mus musculus
85profile befane transporter37splp PMS48Sathmedia tophimutua86oligoppild e permease7gp X77636Bacillus subdifs87membrane spanning26gp X77636Bacillus subdifs88n-acetty spanning glutamyl-phosphate reduct12splQ07960Bacillus subdifs90carboni davide binding25gl (600730)Brucella melitencis91ERV operon: $fn X$ 15splP17582]Excherichia coli92stads43gp/Z38002Saccharomyces cerevisiae93fstYEK (cell division control)30gp Xa4398Excherichia coli94(fcr3) cpn606emb/X75420Plasmoduim falciparum95beat-shock54gp J27352Excherichia coli96signal recognition particle42splP27414Sulfolobus acidocaldarius972-hydroxyhepta-2-4-diene-1,7-dioate isom.19splP27352Excherichia coli98alpha-amylase (amyA)12gp JL3279Clostridium difficile100gyoxysomal citrate synthase16gp D28312Cucurbita sp.101inorganic pyrophophatase28gp X2728Nicotaan tabacum102malate oxidoreductase39splP16468Bradilyticobium inponicum103nodulation ATP-binding I8gp JD0504Excherichia coli103nodulation ATP-binding I8gp JD2333Excherichia coli104pyrotacharodoxin oxidoreductase14sp JP05345	84	polysulfide reductase A	7	sp P31075	Wolinella succinogenes
8000g0pcptu0 permease1gplx03931Salinoneid synumizution87membrane spanning-iphosphate reduct12splC07306Bacillus subility88 n -acetyl-gamma-glutamyl-phosphate reduct12splC07306Bacillus subility89carbon dioxide binding25gjl600730Brucella melicensis90carbonic anhydrase11splP17582Erscherichia coli91ERV operon: fn/X15splP23153Erscherichia coli92siad30gglX04398Erscherichia coli93firyF2X (cell division control)30gglX04398Erscherichia coli94(fcr3) cpu606embX754201Plazomodium fidiciparum95hcat-shock54gplD296721Proceccus sp.96signal recognition particle22splP27414Sulfolobus acidocaldarius972-hydroxyhepta-2,4-dione-1,7-dioate isom.19splP37352Erscherichia coli98alpha-amylase (amyA)12gplX3929Clostridium difficile100glyoxyonal citrat synthase16gplD38132Cucarbia sp.101inorganic pyrophosphatase28gplX3939SplP16468102malate coxidoreductase (nifi)10splP33333Klebicalla pneumoniae103nodulation ATP-binding I8splP23532Erscherichia coli104cyndroxyobutyrut-CoA dehydrogenase21splP23532Erscherichia coli105ribodinase11splP23532Erscher	85	proline/betaine transporter	3/	sp P30848	Escherichia coli
$ \begin{array}{c c c c c c c c c c c c c c c c c c c $	80 87	membrane spanning	26	gp[X03491]	Saimonella lypnimurium Bacillus subtilis
Cellodar processes server the theory gamma gamma gamma program trends in the processes generation of the processes in the process in the proces	88	<i>n</i> -acetyl-gamma-glutamyl-phosphate reduct	12	sp 007906	Bacillus stearothermophilus
Cumum processesgi [600730]Brucella melitensis90carbonic anhydrase11splP17512Escherichia coli91ERV operon; rr_X 15splP132133Escherichia coli92stad543gpl238021Saccharomyces cerevisiae93ftsYEX (cell division control)30gplX43988Escherichia coli94(fcr3) cptd06emb(X75420)Plasmodium falciparum95heat-shock54gplD296721Pyrococcus sp.96signal recognition particle42splP27414Sulfolobus acidocaldariusCentral intermediary metabolism972-hydroxyhepta-2,4-diene-1,7-dioate isom.19splP373521Escherichia coli98alpha-amylase (amyA)12gplX37291Escherichia coli99enoyl-CoA hydratase26gplX37291Escherichia coli100glyoxysomal citrate synthase16gplX37218Nicotiana tabacum101inorganic pyrophosphatase28gplX37218Nicotiana tabacum102malate oxidoreductase14splP050544Escherichia coli103nodulation ATP-binding I8ppl7527218Escherichia coli104pyrotyta-Edhydrogenase11splP050544Escherichia coli105ribkinase10splP050544Escherichia coli106desulfoviridin, gamma18gpl24321Clostridium acetoburylicum1073-hydroxybutyrd-CA dehydrogenase25splP236431 <t< td=""><td>Collular</td><td></td><td></td><td>spicessool</td><td></td></t<>	Collular			spicessool	
0carbonic anhydrase1 gp </td <td>89 80</td> <td>carbon dioxide binding</td> <td>25</td> <td>gil600730</td> <td>Brucella melitensis</td>	89 80	carbon dioxide binding	25	gil600730	Brucella melitensis
91ERV operon: fvX 15sp[P3213] <i>Excherichia coli</i> 92sta543gp[Z38302]Saccharmyces cerevisiae93 fx^rLX (cell division control)30gp[X04398] <i>Escherichia coli</i> 94 $(fcr3)$ $cpnd0$ 6emb[X7320] <i>Plasmodium falciparum</i> 95heat-shock54gp[D29572] <i>Pyrococcus sp.</i> 96signal recognition particle42sp[P27414] <i>Sulfolobus acidocaldarius</i> Central intermediary metabolism972-hydroxyhepta-2,4-diene-1,7-dioate isom.19sp[P37352] <i>Escherichia coli</i> 98alpha-amylase (any,4)12gp[X1379] <i>Escherichia coli</i> 99enoyl-CoA hydratase26gp[X7399] <i>Clostridium difficile</i> 100glyoxysomal citrate synthase16gp[D38312] <i>Cucurbia sp.</i> 101inorganic pyrophosphatase28gp[X83728] <i>Nicotiana tabacum</i> 102malate oxidoreductase39sp[10468] <i>Bacillus stearothermophil.</i> 103nodulation ATP-binding I8pir[S27496] <i>Bradyrhizobium ipaonicum</i> 104pruravate-flavodoxin oxidoreductase (nifj)10sp[P2352] <i>Clostridium acetoburylicum</i> 105ribokinase11pir[A43723] <i>Clostridium acetoburylicum</i> 106desalloviridin, gamma18gp[L05610] <i>Desulfovibrio valgaris</i> 1073-hydroxybutyrl-CoA dehydrogenase25sp[P2352] <i>Clostridium acetoburylicum</i> 1083-hydroxybutyrl-CoA dehydrog	90	carbonic anhydrase	11	sp[P17582]	Escherichia coli
92such43eight 28002Succharomyces cerevisiae93fix EX (cell division control)30gplX04398Succharomyces cerevisiae94(fcr3) cpn606emblX75420Plasmodium falciparum95heat-shock54gpl2971Pyrococcus sp.96signal recognition particle42splP27414Sulfolobus acidocaldariusCentral intermediary metabolism972-hydroxyhepta-2.4-diene-1.7-dioate isom.19splP37552Escherichia coli98alpha-amylase (amyA)12gplX3529Escherichia coli99encyl-CoA hydratase26gplX9899Clostridium difficile100glyoxysomal citrate synthase16gplX983728Nicotiana tabacum101inorganic pyrophosphatase39splP16468Bradyhritzobium japonicum103nodulation ATP-binding I8pirJ873728Klebsiella neumoniae104pyruxete-flavodosin oxidoreductase (nifi)10splP05610Desulfovibrio vulgaris105ribokinase11pirJ43723Clostridium acetobutylicum106desulfoviridi, gamma18gpl23523Clostridium acetobutylicum1073-hydroxybuttyrb-CoA dehydrogenase11pirJ43723Clostridium acetobutylicum1083-hydroxybuttyrb-CoA dehydrogenase25spl23528Escherichia coli1093-oxoacyl-lecyl-carrier-protein reductase26gplX51028Escherichia coli1104-ocoumarate-CoA ligase26 </td <td>91</td> <td>ERV operon: frvX</td> <td>15</td> <td>sp P32153 </td> <td>Escherichia coli</td>	91	ERV operon: frvX	15	sp P32153	Escherichia coli
93 $fxYEX$ (cell division control)30 gp [X04398] $Excherichia coli94(fcr3) cpn606cmb X75420]Plasmodium falciparum95heat-shock54gp[D29672]Pyrococcus sp.96signal recognition particle42sp[2714]Salolobus acidocaldarius972-hydroxyhepta-2.4-diene-1,7-dioate isom.19sp[237352]Escherichia coli98alpha-amylase (am)A)12gp[L13279]Escherichia coli99encyl-CoA hydratase26gp[X7989]Clostridium diffcile100glyoxysomal citrate synthase16gp[S3132]Cuarrbia sp.101inorganic pyrophosphatase28gg[X3138]Nicotiana tabcum103nodulation ATP-binding I8pir[S27496]Bradyhticobium iaponicum104pyruvate-flavodoxin oxidoreductase (nifj)10sp!P03054 Escherichia coli105ribokinase14sp!P0504 Desulfovibriaganicum106desulfoviridin, gamma18gp[L0510]Desulfovibriaganicum1073-hydroxyisotyrate dehydrogenase25sp!P2352]Escherichia coli1083-hydroxyisotyrate dehydrogenase20gp[M3508]Methanothix soelnegenii1115-oxopent-3-enel.2,5-tricarbox. decarbox.20gp[M3508]Methanothix soelnegenii112acetylcLox A synthetase36gp[M3598]Methanothix soelnegenii113adorysultar reductase28pp[293139]Bacillas polymyxa$	92	sua5	43	gp[Z38002]	Saccharomyces cerevisiae
94 $(fcr3)$ cpn606cmb X75420 Plasmodium falciparum95heat-shock54gp D29672 Pyrococcus sp.96signal recognition particle42sp P27414 Sulfolobus acidocaldarius972-bydroxyhepta-2.4-diene-1.7-dioate isom.19sp P37352 Escherichia coli98alpha-amylase $(anyA)$ 12gp L13279 Escherichia coli99encyl-CoA hydratase26gp/X9899 Clostridium difficile100glyoxysomal citrate synthase16gp/D38132Nicotiana tabacum101inorganic pyrophosphatase28gp/X3728 Nicotiana tabacum102malate oxidoreductase39sp P16468 Bradyrhitzobium igponicum103nodulation ATP-binding I8pir S27496 Bradyrhitzobium igponicum104pyruvate-flavodoxin oxidoreductase (nifj)10sp P05054 Escherichia coli105ribokinase14sp P05054 Escherichia coli106desulfoviridin, gamma18gp L0510 Desulfovirbio vulgaris1073-hydroxybutyrl-CoA dehydrogenase25sp P25323 Escherichia coli1083-hydroxybutyrl-CoA derbydrogenase20gp X1502 Pinus taeda1115-toxopent-3-enet_1-2,-stricarbox. decarbox.20gp X1502 Pinus taeda112acetyl-CoA synthetase36gp M63968 Methanothrix soelingenii113adenylsulfate reductase28pip18438 Archaeoglobus fulgidus114 <td>93</td> <td>ftsYEX (cell division control)</td> <td>30</td> <td>gp[X04398]</td> <td>Escherichia coli</td>	93	ftsYEX (cell division control)	30	gp[X04398]	Escherichia coli
95heat-shock54gp]P27414 <i>Sulfolobus acidocaldarius</i> 96signal recognition particle42sp]P27414 <i>Sulfolobus acidocaldarius</i> 972-hydroxyhepta-2,4-diene-1,7-dioate isom.19sp]P37352 <i>Escherichia coli</i> 98alpha-amylase (amyA)12gp]L13279 <i>Escherichia coli</i> 99enoyl-CoA hydratase26gp]X79899 <i>Clostrialum difficile</i> 100glyoxysomal citrate synthase16gp]X38132 <i>Clostrialum difficile</i> 101inorganic pyrophosphatase28gp]X38728 <i>Nicotianu tabacum</i> 102malate oxidoreductase39sp]P16468 <i>Bacillus stearothermophil.</i> 103nodulation ATP-binding I8pp183233 <i>Klebsiella pneumoniae</i> 104pyruvate-flavodxin oxidoreductase (nifj)10sp]P05054 <i>Escherichia coli</i> 105ribokinase14sp]P050514 <i>Escherichia coli</i> 106desulfoviridin, gamma18gp]L05610 <i>Desulfovibrio vulgaris</i> 1073-hydroxybutryl-CoA dehydrogenase25sp]P23523 <i>Escherichia coli</i> 1083-hydroxybutryl-CoA dehydrogenase20gp]U12012 <i>Pinus taeda</i> 1115-oxopent-3-enel.2,5-tricarbox. decarbox.20gp]N53086 <i>Methanostrix solingenii</i> 112acetyl-CoA synthetase36gp]M63968Methanostrix solingenii113adenysulfate reductase28pif1822986 <i>Archaeoglobus fulgidus</i> 114subditisin61sp]P291391 <i>Bacillus st</i>	94	(fcr3) cpn60	6	emb X75420	Plasmodium falciparum
96signal recognition particle42splP27414Sulfolobus acidocaldariusCentral intermediary metabolism972-hydroxyhepta-2.4-diene-1.7-dioate isom.19splP37352Excherichia coli98alpha-amylase (amy.4)12gplX79899Clostrichia coli99enoyl-CoA hydratase26gplX79899Clostrichian difficile100glyoxysomal citrate synthase16gplD38132Cucurbita sp.101inorganic pyrophosphatase28gplX79899Clostrichian tabacum102malate oxidoreductase39splP164648Bacillus stearothermophil.103nodulation ATP-binding I8piP138333Klebsilla pneumoniae104pyruate-flavodoxin oxidoreductase (nifj)10splP038333Klebsilla pneumoniae105ribokinase14splP050544Excherichia coli106desulfoviridin gamma18gpl205614Excherichia coli1073-hydroxybutyryl-CoA dehydrogenase11pir/A437231Clostridium acetobutylicum1083-hydroxybutyryl-CoA ligase20gplX75028Excherichia coli1104-coumarte-CoA ligase20gplX75028Excherichia coli111subtilisin61splP291391Bacillus polymyta112acetyl-CoA synthetase36gplX75028Excherichia coli113achyloughesh esplates75splP1391Bacillus polymyta114subtilisin61splP291391Bacillus polymyta	95	heat-shock	54	gp D29672	Pyrococcus sp.
Central intermediary metabolism972-hydroxyhepta-2,4-diene-1,7-dioate isom.19sp[P37352]Escherichia coli98alpha-amylase (amyA)12gp[L13279]Escherichia coli99enoyl-CoA hydratase26gp DX9899]Clostridium difficile100glyoxysomal citrate synthase16gp D3812Cucurbia sp.101inorganic pyrophosphatase28gp DK3728]Nicoitana tabacum102malate oxidoreductase39splP16468]Bacillus stearothermophil.103nodulation ATP-binding I8pirS27496]Bradyrhizobium japonicum104pyruvate-flavodoxin oxidoreductase (nifj)10splP05833]Klebsiella pneumoniae105ribokinase14splP0584Escherichia coli106desulfoviridin, gamma18gplL05610]Desulfovibrio vulgarisEnergy metabolism1073-hydroxyiobutyrut-CoA dehydrogenase25splP23523]Escherichia coli1093-oxoacyl-lacyl-carrier-protein] reductase14splP26643]Cuphea lanceolata1104-coumarate-CoA ligase20gp [M15028]Archaeoglobus fulgidus1115-oxopent-3-ene1,2,5-tricarbox. decarbox.20gp [M5028]Archaeoglobus fulgidus112acetyl-CoA synthetase28pirS18928]Archaeoglobus fulgidus113adenylsulfate reductase28pirS18928]Archaeoglobus fulgidus114subtilisin61spl20139]Bacil	96	signal recognition particle	42	sp P27414	Sulfolobus acidocaldarius
972-hydroxyshepta-2,4-diene-1,7-dioate isom.19splsplState98alpha-amylase (any A)12gpL13279Escherichia coli99enoyl-CoA hydratase26gpX79899Clostridium difficile100glyoxysomal citrate synthase16gpD38132Cucurbia sp.101inorganic pyrophosphatase28gpX83728Nicoiana tabacum102malate oxidoreductase39spP16468Bacillus stearothermophil.103nodulation ATP-binding I8pirS27496Bradynhicobium iaponicum104pyruvate-flavodoxin oxidoreductase (nifj)10spP03833Klebsiella pneumoniae105ribokinase14spSplScherichia coli10106desulfoviridin, gamma18gpL05610Desulfovibri vulgarisEnergy metabolism1073-hydroxysbutyryl-CoA dehydrogenase11pirA43723Clostridium acetoburylicum1083-hydroxysbutyrate dehydrogenase20gpL1212Pinus taeda1104-coumarate-CoA ligase36gpMchanodirix soehgenii111soxoacyl-facyl-carrier-protein] reductase36gpMchanodirix soehgenii112acetyl-CoA synthetase36gpMchanodirix soehgenii113adenylsulfate reductase28pirS1828Archaeoglobus fulgidus114subtilsin61spsp115aninopeptidase P25 <td>Central</td> <td>intermediary metabolism</td> <td></td> <td></td> <td></td>	Central	intermediary metabolism			
98alpha-amylase $(any.4)$ 12gp[L13279] <i>Excherichia coli</i> 99enoyl-CoA hydratase26gp[X83728](Costridium difficile100glyoxysomal citrate synthase16gp[D38132] <i>Cucurbita sp.</i> 101inorganic pyrophosphatase28gp[X83728]Nicotiana tabacum102malate oxidoreductase39sp[Pl6468]Bacillus stearothermophil.103nodulation ATP-binding I8pir[S27496]Bradythizobium iponicum104pyruvate-flavodoxin oxidoreductase (nifj)10sp[P05054] <i>Escherichia coli</i> 105ribokinase14sp[P05054] <i>Escherichia coli</i> 106desulfoviridin, gamma18gp[L15210] <i>Desulfovibrio vulgaris</i> Energy metabolismEnergy metabolismEnergy metabolismEnergy metabolismEnergy metabolismExcherichia coli ony pir[A43723] <i>Clostridium acetobutylicum</i> 1083-hydroxylutyrl-CoA dehydrogenase25sp[P23523] <i>Escherichia coli</i> 1093-oxoacyl-[acyl-carrier-protein] reductase10gp [X5028] <i>Escherichia coli</i> 11soxoacyl-[acyl-carrier-protein] reductase26gp [X75028] <i>Escherichia coli</i> 11Soxoacyl-[acyl-carrier-protein] reductase36gp [M26396] <i>Mechanothrix soehngenii</i> 111Soxoacyl-[acyl-carrier-protein] reductase28 <t< td=""><td>97</td><td>2-hydroxyhepta-2,4-diene-1,7-dioate isom.</td><td>19</td><td>sp P37352 </td><td>Escherichia coli</td></t<>	97	2-hydroxyhepta-2,4-diene-1,7-dioate isom.	19	sp P37352	Escherichia coli
99enoyl-CoA hydratase26gp[X7989]Clostridium difficile100glyoxysomal citrate synthase16gp[38132]Cucurbia sp.101inorganic pyrophosphatase28gp[X83728]Nicotiana tabacum102malate oxidoreductase39sp[P16468]Bacillus stearothermophil.103nodulation ATP-binding I8pir[S27496]Bradyrhizobium japonicum104pyruvate-flavodoxin oxidoreductase (nifj)10sp[P03833]Klebsiella pneumoniae105ribokinase14sp[P05054]Escherichia coli106desulfoviridin, gamma18gp[L05610]Desulfovibrio vulgarisEnergy metabolism1073-hydroxybutyryl-CoA dehydrogenase11pir[A43723]Clostridium acetobutylicum1083-hydroxybutyrate dehydrogenase25sp[P23523]Escherichia coli1104-coumarate-CoA ligase20gp[W12012]Pinus taeda1114-coumarate-CoA synthetase36gp[M63968]Methanothrix soehngenii113adenylsulfate reductase28pir[S18928]Archaeoglobus fulgidus114subtilisin61sp[P29139]Bacillus searothermophil.115aminopeptidase P25gp[0436]Methanostrix soehngenii116N-acetylglutamate-g-semialdehyde DH (argC)15sp[0436]Methanostric acli118arP13gp[X86160]Escherichia coli119beta-lactamase9gp[D436]Methanostric acli<	98	alpha-amylase (<i>amyA</i>)	12	gp L13279	Escherichia coli
100glyoxysomal citrate synthase16gp[D38152]Cucurbita sp.101inorganic pyrophosphatse28gp[X83728]Nicotiana tabacum102malate oxidoreductase39sp[P16468]Bacillus stearothermophil.103nodulation ATP-binding I8pir[S27496]Bradyrhizobium japonicum104pyruvate-flavodoxin oxidoreductase (nifj)10sp[P05054]Escherichia coli105ribokinase14sp[P05054]Escherichia coli106desulfoviridin, gamma18gp[L05610]Desulfovirio vulgaris1073-hydroxybutyryl-CoA dehydrogenase25sp[P23523]Escherichia coli1083-hydroxyisobutyrate dehydrogenase20gp[U12012]Pinus taeda1104-coumarate-CoA ligase20gp[U12012]Pinus taeda1115-oxopent-3-enel,2,5-tricarbox. decarbox.20gp[M53968]Methanothrix solengenii113adenylsulfate reductase28pir[S18928]Archaeoglobus fulgidus114subtilisin61sp[P2319]Bacillus stearothermophil.115aminopeptidase P25gi[1046027]Mycoplasma genitalium116N-acetylglutamate-g-semialdehyde DH (argC)15sp[D0906]Bacillus stearothermophil.117ATPase31gj[1304392]Sulfolobus solfataricus118atrP13gp[X86160]Escherichia coli119beta-lactamase9sp[P00811]Escherichia coli120carbamoylphospha	99	enoyl-CoA hydratase	26	gp X79899	Clostridium difficile
101morganic pyrophosphatase28gp[X85/28]Nicotiana tabacum102malate oxidoreductase39sp[P16466]Bacillus staarothermophil.103nodulation ATP-binding I8pir[S27496]Bradyrhizobium japonicum104pyruvate-flavodxin oxidoreductase (nifj)10sp[P0383]Klebsiella pneumoniae105ribokinase14sp[P05054]Excherichia coli106desulfoviridin, gamma18gp[L05610]Desulfovibrio vulgarisEnergy metabolism1073-hydroxybutyryl-CoA dehydrogenase11pir[A43723]Clostridium acetobutylicum1083-hydroxybutyrate dehydrogenase25sp[P23523]Excherichia coli1093-oxoacyl-facyl-carrier-protein] reductase14sp[P28643]Cuphea lanceolata1104-coumarate-CoA ligase20gp[N5028]Escherichia coli112acetyl-CoA synthetase36gp[M3968]Methanothrix sochngenii113adenylsulfate reductase28pir[S18928]Archacoglobus fulgidus114subtilisin61sp[P29139]Bacillus polymyxa115aminopeptidase P25gj[N46027]Mycoplasma genitalium116A-acetylgutamate-g-semialdehyde DH (argC)15sp[Q07906]Bacillus stearothermophil.117ATPase9gp[N48160]Escherichia coli120carbamoylphosphate synthetase, large31gj[N34392]Sulfolobus solfataricus121cellulase <t< td=""><td>100</td><td>glyoxysomal citrate synthase</td><td>16</td><td>gp D38132 </td><td>Cucurbita sp.</td></t<>	100	glyoxysomal citrate synthase	16	gp D38132	Cucurbita sp.
102Indiate Oxtobeuteuse39splp(1946)Database intermonial103nodulation ATP-binding I8splp(1946)Bradyrhizobium inponicum104pytruate-flavodoxin oxidoreductase (nifj)10splp03833Klebsiella pneumoniae105ribokinase14splP05054 Escherichia coli106desulfoviridin, gamma18gplL05610 Desulfovibrio vulgarisEnergy metabolism1073-hydroxyisobutyrate dehydrogenase25splP236343Cuphea lanceolata1083-hydroxyisobutyrate dehydrogenase20gplU12012 Pinus taeda1093-oxoacyl-facyl-carrier-protein] reductase20gplX57028 Excherichia coli1104-coumarate-CoA ligase20gplX57028 Excherichia coli121acetyl-coA synthetase36gplM63968 Methanothrix soehngenii113adenylsulfate reductase28pirS18928 Archaeoglobus fulgidus114subtilisin61splP2139 Bacillus polymyxa115aminopeptidase P25gpl1046027 Mcycoplasma geniallium116N-acetylglutamate-g-semialdehyde DH (argC)15spl2007906 Bacillus stearothermophil.117ATPase9spl208610 Excherichia coli128carbamoylphosphate synthetase, large31gj1304392 Sulfolobus sulfaricus129carbamoylphosphate synthetase, large9spl20801 Excherichia coli120carbamoylphosphate synthetase, large<	101	inorganic pyrophosphatase	28	gp X83/28	Nicotiana tabacum Pacillus stagnothermonkil
1031031031031041041041041041041041041041041041041061061071061071071071083107310831073108310731083108	102	nodulation ATP hinding I	59 8	sp[F10400]	Bradyrhizobium japonicum
105ribokinase14sp[P05054]Escherichia coli106desulfoviridin, gamma18gp[L05610]Desulfovibrio vulgarisEnergy metabolism1073-hydroxybutyryl-CoA dehydrogenase11pir[A43723]Clostridium acetobutylicum1083-hydroxybotyryl-CoA dehydrogenase25sp[P23523]Escherichia coli1093-oxoacyl-[acyl-carrier-protein] reductase14sp[P28643]Cuphea lanceolata1104-coumarate-CoA ligase20gp[V12012]Pinus taeda1115-oxopent-3-enel.2.5-tricarbox. decarbox.20gp[X5028]Escherichia coli112acetyl-CoA synthetase36gp[M63968]Methanothrix soehngenii113adenylsulfate reductase28pir[S18928]Archaeoglobus fulgidus114subilisin61sp[P2039]Bacillus polymyxa115aminopeptidase P25gi[1046027]Mycoplasma genitalium116N-acetylglutamate-g-semialdehyde DH (argC)15sp[Q07906]Bacillus stearothermophil.117ATPase87gp[X86160]Escherichia coli120carbamoylphosphate synthetase, large31gi[X3742]Caldocellum saccharolyticum121cellulase26gp[X5172]Sulfolobus solfataricus122cephalosporin C acylase9gp[X0132]Sulfolobus solfataricus123chloroplast ATPase26gp[X60752]chloroplast Gammateria124dihydrolipoamide dehydrogenase27	103	pyruvate-flavodoxin oxidoreductase (nifi)	10	sp[P03833]	Klebsiella nneumoniae
106107108108108108108106desulfoviridin, gamma18gp[L05610]Desulfovibrio vulgaris1073-hydroxybutyryl-CoA dehydrogenase11pir[A43723]Clostridium acetobutylicum1083-hydroxyisobutyrate dehydrogenase25sp[P23523]Escherichia coli1093-oxoacyl-[acyl-carrier-protein] reductase14sp[P28643]Cuphea lanceolata1104-coumarate-CoA ligase20gp[U12012]Pinus taeda1115-oxopent-3-ene1,2,5-tricarbox, decarbox.20gp[M53968]Methanothrix soehngenii113adenylsulfate reductase28pir[S18928]Archaeoglobus fulgidus114subilisin61sp[P29139]Bacillus polymyxa115aminopeptidase P25gi [1046027]Mycoplasma genitalium116N-acctylglutamate-g-semialdehyde DH (argC)15sp[Q07906]Bacillus stearothermophil.117ATPase9gp[X5160]Escherichia coli120carbamoylphosphate synthetase, large31gi [1304392]Sulfolobus solfatricus121celluase26gp [X5075]Caldocellum saccharolyticum122cephalosporin C acylase9gp [X60752]chloroplast Odontella sine.123chloroplast ATPase26gp [X60752]chloroplast Odontella sine.124dihydrolipoamide dehydrogenase27sp [P1959]Bacillus stearothermophil.125DR-nm2316gp [U29656]Homo sapiens	105	ribokinase	10	sp P05054	Escherichia coli
Energy metabolism1073-hydroxyisobutyryl-CoA dehydrogenase11pir[A43723]Clostridium acetobutylicum1083-hydroxyisobutyrate dehydrogenase25sp[P23523]Escherichia coli1093-oxoacyl-[acyl-carrier-protein] reductase14sp[P28643]Cuphea lanceolata1104-coumarate-CoA ligase20gp[U12012]Pinus taeda1115-oxopent-3-enel,2,5-tricarbox. decarbox.20gp[M63968]Methanothrix soelngenii113acetyl-CoA synthetase28pir[S18928]Archaeolobus fulgidus114subtilisin61sp[P29139]Bacillus polymyxa115aminopeptidase P25gi[1046027]Mycoplasma genialium116N-acetylglutamate-g-semialdehyde DH (argC)15sp[007906]Bacillus stearothermophil.117ATPase87gp]N886160 Escherichia coli119beta-lactamase9sp[P00811]Escherichia coli120carbamo/plphosphate synthetase, large31gi[304392]Sulfolobus solfataricus121celhalosporin C acylase26gp]X50752 Chlooroplast Odonella sincu122cephalosporin C acylase26gp]X20752 chlooroplast Odonella sincu123dihydrolipoamide dehydrogenase27sp]P11959 Bacillus stearothermophil.124dihydrolipoamide dehydrogenase17sp]P29201 Haloarcula125DR-nm2316gp[U292056]Homo sapiens126enolase17<	106	desulfoviridin, gamma	18	gp L05610	Desulfovibrio vulgaris
1073-hydroxybutyryl-CoA dehydrogenase11pirpirA43723Clostridium acetobutylicum1083-hydroxyisobutyrate dehydrogenase25spP23523Escherichia coli1093-oxoacyl-[acyl-carrier-protein] reductase14spP28643Cuphea lanceolata1104-coumarate-CoA ligase20gpU12012Pinus taeda1115-oxopent-3-ene1,2,5-tricarbox. decarbox.20gpX75028Escherichia coli112acetyl-CoA synthetase36gpMethanothrix soehngenii113adenylsufate reductase28pirS18928Archaeoglobus fulgidus114subtilisin61spP29139Bacillus polymyxa115aminopeptidase P25gi1046027Mycoplasma genitalium116N-acetylglutamate-g-semialdehyde DH (argC)15spQ07906Bacillus stearothermophil.117ATPase87gpJ04836Methanosarcina barkeri118atrP13gpX86160Escherichia coli120carbamoylphosphate synthetase, large31gi134392Sulfolobus solfataricus121cellulase26gpL32742Caldocellum saccharolyticum122cephalosporin C acylase9gpJ104392Sulfolobus solfataricus123chloroplast ATPase26gpK0752chloroplast Odontella sine.124dihydrolipoamide dehydrogenase27spP1959Bacillus stear	Energy	netabolism			
1083-hydroxyisobutyrate dehydrogenase25sp[P23523]Escherichia coli1093-oxoacyl-[acyl-carrier-protein] reductase14sp[P28643]Cuphea lanceolata1104-coumarate-CoA ligase20gp[U12012]Pinus taeda1115-oxopent-3-enel,2,5-tricarbox. decarbox.20gp[X75028]Escherichia coli112acetyl-CoA synthetase36gp[M63968]Methanothrix soehngenii113adenylsulfate reductase28pir[S18928]Archaeoglobus fulgidus114subtilisin61sp[P29139]Bacillus polymyxa115aminopeptidase P25gi[1046027]Mycoplasma genitalium116N-acetylglutamate-g-semialdehyde DH (argC)15sp[Q07906]Bacillus stearothermophil.117ATPase87gp[J04836]Methanosarcina barkeri118atrP13gp[X86160]Escherichia coli120carbamoylphosphate synthetase, large31gi[I304392]Sulfolobus solfataricus121cellulase26gp[X2742]Caldocellum saccharolyticum122cephalosporin C acylase9gp[X17015]Pseudomonas diminuta123chloroplast ATPase26gp[X2656]Homo sapiens124dihydrolipoamide dehydrogenase27sp[P1959]Bacillus stearothermophil.125DR-nm2316gp[V29656]Homo sapiens126enolase17sp[P29201]Haloarcula125fadBA operon (fatty acid oxidizing)8 <td>107</td> <td>3-hydroxybutyryl-CoA dehydrogenase</td> <td>11</td> <td>pir A43723 </td> <td>Clostridium acetobutylicum</td>	107	3-hydroxybutyryl-CoA dehydrogenase	11	pir A43723	Clostridium acetobutylicum
1093-oxoacyl-[acyl-carrier-protein] reductase14sp [P28643]Cuphea lanceolata1104-coumarate-CoA ligase20gp [U12012]Pinus taeda1115-oxopent-3-enel,2,5-tricarbox. decarbox.20gp [X75028]Escherichia coli112acctyl-CoA synthetase36gp [M63968]Methanothrix soelngenii113adenylsulfate reductase28pir [S18928]Archaeoglobus fulgidus114subtilisin61sp [P29139]Bacillus polymyxa115aminopeptidase P25gi [1046027]Mycoplasma genitalium116N-acetylglutamate-g-semialdehyde DH (argC)15sp [Q07906]Bacillus stearothermophil.117ATPase87gp [J04836]Methanosarcina barkeri118atrP13gp [X86160]Escherichia coli120carbamoylphosphate synthetase, large31gj [1304392]Sulfolobus solfataricus121cellulase26gp [X3772]Caldocellum saccharolyticum122cephalosporin C acylase9gp [A17015]Pseudomonas diminuta123chloroplast ATPase26gp [X60752]chloroplast Odontella sine.124dihydrolipoamide dehydrogenase17sp [P29201]Haloarcula125DR-m2316gp [V2956]Homo sapiens126enolase17sp [P29201]Haloarcula127fadBA operon (fatty acid oxidizing)8gb [M74164]Escherichia coli128formate dehydrogenase9	108	3-hydroxyisobutyrate dehydrogenase	25	sp P23523	Escherichia coli
1104-coumarate-CoA ligase20gp[U12012]Pinus taeda1115-oxopent-3-enel,2,5-tricarbox, decarbox.20gp[X75028]Escherichia coli112acetyl-CoA synthetase36gp[M63968]Methanothrix soehngenii113adenylsulfate reductase28pir[S18928]Archaeoglobus fulgidus114subtilisin61sp[P29139]Bacillus polymyxa115aminopeptidase P25gi[1046027]Mycoplasma genitalium116N-acetylglutamate-g-semialdehyde DH (argC)15sp[Q07906]Bacillus stearothermophil.117ATPase87gp J04836]Methanostrica barkeri118atrP13gp[X86160]Escherichia coli120carbamoylphosphate synthetase, large31gi[1304392]Sulfolobus solfataricus121cellulase26gp [X6752]Chloroplast Odontella sine.122cephalosporin C acylase9gp [X6752]chloroplast Odontella sine.123chloroplast ATPase26gp [X6752]chloroplast Odontella sine.124dihydrolipoamide dehydrogenase27sp[P11959]Bacillus stearothermophil.125DR-nm2316gp [U29656]Homo sapiens126enolase17sp [P29201]Haloarcula127fadBA operon (fatty acid oxidizing)8gb [M74164]Escherichic acoli128formate dehydrogenase25pir[S34619]Thermoplasma acidophilum129fumarter reductase25 <td>109</td> <td>3-oxoacyl-[acyl-carrier-protein] reductase</td> <td>14</td> <td>sp P28643 </td> <td>Cuphea lanceolata</td>	109	3-oxoacyl-[acyl-carrier-protein] reductase	14	sp P28643	Cuphea lanceolata
1115-oxopent-3-enel,2,5-tricarbox. decarbox.20gp X75028 Escherichia coli112acetyl-CoA synthetase36gp M63968 Methanothrix soehngenii113adenylsulfate reductase28pir S18928 Archaeoglobus fulgidus114subtilisin61sp P29139 Bacillus polymyxa115aminopeptidase P25gi 1046027 Mycoplasma genitalium116N-acetylglutamate-g-semialdehyde DH (argC)15sp Q07906 Bacillus stearothermophil.117ATPase87gp X86160 Escherichia coli118atrP13gp X86160 Escherichia coli120carbamoylphosphate synthetase, large31gi 1304392 Sulfolobus solfataricus121cellulase26gp X7015 Pseudomonas diminuta122cephalosporin C acylase9gp X71015 Pseudomonas diminuta123chloroplast ATPase26gp X71015 Pseudomonas diminuta124dihydrolipoamide dehydrogenase27sp P11959 Bacillus stearothermophil.125DR-nm2316gp U29656 Homo sapiens126enolase17sp P29201 Haloarcula127fadBA operon (fatty acid oxidizing)8gb X71164 Escherichia coli128formate dehydrogenase25pir S34619 Thermoplasma acidophilum139glucosamine-fruc6-phos, aminotransferase24sp P39754 Bacillus subtilis	110	4-coumarate-CoA ligase	20	gp U12012	Pinus taeda
112acetyl-CoA synthetase36gp[M63968]Methanothrix soehngenii113adenylsulfate reductase28pir[S18928]Archaeoglobus fulgidus114subtilisin61sp[P29139]Bacillus polymyxa115aminopeptidase P25gi[1046027]Mycoplasma genitalium116N-acetylglutamate-g-semialdehyde DH (argC)15sp[Q07906]Bacillus stearothermophil.117ATPase87gp[J04836]Methanosarcina barkeri118atrP13gp[X86160]Escherichia coli119beta-lactamase9sp[P00811]Escherichia coli120carbamoylphosphate synthetase, large31gi[1304392]Sulfolobus solfataricus121cellulase26gp[X2742]Caldocellum saccharolyticum122cephalosporin C acylase9gp[X60752]chloroplast Odontella sine.123chloroplast ATPase26gp[X29656]Homo sapiens124dihydrolipoamide dehydrogenase27sp[P11959]Bacillus stearothermophil.125DR-nm2316gp[V29656]Homo sapiens126enolase17sp[P29201]Haloarcula127fadBA operon (fatty acid oxidizing)8gb[M74164]Escherichia coli128formate dehydrogenase25pir[S34619]Thermoplasma acidophilum130glucosamine-fruc6-phos. aminotransferase24sp[P39754]Bacillus subtilis	111	5-oxopent-3-ene1,2,5-tricarbox. decarbox.	20	gp X75028	Escherichia coli
113adenylsultate reductase28 $pr[S18928]$ Archaeoglobus fulgidus114subtilisin61 $sp[P29139]$ Bacillus polymyxa115aminopeptidase P25 $gi[1046027]$ Mycoplasma genitalium116N-acetylglutamate-g-semialdehyde DH (argC)15 $sp[Q07906]$ Bacillus stearothermophil.117ATPase87 $gp[J04836]$ Methanosarcina barkeri118 $atrP$ 13 $gp[X86160]$ Escherichia coli119beta-lactamase9 $sp[P00811]$ Escherichia coli120carbamoylphosphate synthetase, large31 $gi[1304392]$ Sulfolobus solfataricus121cellulase26 $gp[X80752]$ Caldocellum saccharolyticum122cephalosporin C acylase9 $gp[X60752]$ chloroplast Odontella sine.123chloroplast ATPase26 $gp[U29656]$ Homo sapiens124dihydrolipoamide dehydrogenase27 $sp[P11959]$ Bacillus stearothermophil.125DR-nm2316 $gp[U29656]$ Homo sapiens126enolase17 $sp[P29201]$ Haloarcula127fadBA operon (fatty acid oxidizing)8 $gb[N74164]$ Escherichia coli128formate dehydrogenase25 $pi[S34619]$ Thermoplasma acidophilum130glucosamine-fruc6-phos. aminotransferase24 $sp[P39754]$ Bacillus subtilis	112	acetyl-CoA synthetase	36	gp M63968	Methanothrix soehngenii
114Sublish61sp[P29139]Baculus polymyxa115aminopeptidase P25gi[1046027]Mycoplasma genitalium116N-acetylglutamate-g-semialdehyde DH (argC)15sp[Q07906]Bacillus stearothermophil.117ATPase87gp[J04836]Methanosarcina barkeri118atrP13gp[X86160]Escherichia coli119beta-lactamase9sp[P00811]Escherichia coli120carbamoylphosphate synthetase, large31gi[1304392]Sulfolobus solfataricus121cellulase26gp[X2742]Caldocellum saccharolyticum122cephalosporin C acylase9gp[A17015]Pseudomonas diminuta123chloroplast ATPase26gp[X260752]chloroplast Odontella sine.124dihydrolipoamide dehydrogenase27sp[P11959]Bacillus stearothermophil.125DR-nm2316gp[U29656]Homo sapiens126enolase17sp[P201]Haloarcula127fadBA operon (fatty acid oxidizing)8gb[M74164]Escherichia coli128formate dehydrogenase9gp [X54057]Wolinella succinogenes129fumarate reductase25pir[S34619]Thermoplasma acidophilum130glucosamine-fruc6-phos, aminotransferase24sp[P39754]Bacillus subtilis	113	adenylsultate reductase	28	pir \$18928	Archaeoglobus fulgidus
115animopeptidase P25gl104027Mycoplasma gentalium116N-acetylglutamate-g-semialdehyde DH (argC)15sp Q07906 Bacillus stearothermophil.117ATPase87gp J04836 Methanosarcina barkeri118atrP13gp X86160 Escherichia coli119beta-lactamase9sp P00811 Escherichia coli120carbamoylphosphate synthetase, large31gi 1304392 Sulfolobus solfataricus121cellulase26gp X50752 Caldocellum saccharolyticum122cephalosporin C acylase9gp X60752 chloroplast Odontella sine.123chloroplast ATPase26gp U29656 Homo sapiens124dihydrolipoamide dehydrogenase27sp P11959 Bacillus stearothermophil.125DR-nm2316gp U29656 Homo sapiens126enolase17sp P29201 Haloarcula127fadBA operon (fatty acid oxidizing)8gb M74164 Escherichia coli128formate dehydrogenase9gp X54057 Wolinella succinogenes129fumarate reductase25pir S34619 Thermoplasma acidophilum130glucosamine-fruc6-phos. aminotransferase24sp P39754 Bacillus subtilis	114	subulisin aminopontidaga B	01	sp P29139 ail1046027	Bacillus polymyxa Muoonlaama oonitalium
110Arectylgitalinate geschnadedryde D11 (arge)13sp[20706]Datamas stationitrinophil.117ATPase87gp[J04836]Methanosarcina barkeri118 $atrP$ 13gp[X86160]Escherichia coli119beta-lactamase9sp[P00811]Escherichia coli120carbamoylphosphate synthetase, large31gi[1304392]Sulfolobus solfataricus121cellulase26gp[X30742]Caldocellum saccharolyticum122cephalosporin C acylase9gp[X60752]chloroplast Odontella sine.123chloroplast ATPase26gp[X60752]chloroplast Odontella sine.124dihydrolipoamide dehydrogenase27sp[P11959]Bacillus stearothermophil.125DR-nm2316gp[U29656]Homo sapiens126enolase17sp[P29201]Haloarcula127fadBA operon (fatty acid oxidizing)8gb[M74164]Escherichia coli128formate dehydrogenase9gp[X54057]Wolinella succinogenes129fumarate reductase25pir[S34619]Thermoplasma acidophilum130glucosamine-fruc6-phos. aminotransferase24sp[P39754]Bacillus subtilis	115	$N_{-acetylglutamate_g_semialdehyde DH(argC)$	23 15	gi 1040027 sp 007906	Mycopiusma genianum Bacillus stearothermonhil
111	117	ATPase	87	gn 104836	Methanosarcina barkeri
119beta-lactamase9sp[P00811]Escherichia coli120carbamoylphosphate synthetase, large31gi[1304392]Sulfolobus solfataricus121cellulase26gp[L32742]Caldocellum saccharolyticum122cephalosporin C acylase9gp[A17015]Pseudomonas diminuta123chloroplast ATPase26gp[X60752]chloroplast Odontella sine.124dihydrolipoamide dehydrogenase27sp[P11959]Bacillus stearothermophil.125DR-nm2316gp[U29656]Homo sapiens126enolase17sp[P29201]Haloarcula127fadBA operon (fatty acid oxidizing)8gb[M74164]Escherichia coli128formate dehydrogenase9gp[X54057]Wolinella succinogenes129fumarate reductase25pir[S34619]Thermoplasma acidophilum130glucosamine-fruc6-phos. aminotransferase24sp[P39754]Bacillus subtilis	118	atrP	13	gp X86160	Escherichia coli
120carbamoylphosphate synthetase, large31gi 1304392 Sulfolobus solfataricus121cellulase26gp L32742 Caldocellum saccharolyticum122cephalosporin C acylase9gp A17015 Pseudomonas diminuta123chloroplast ATPase26gp X60752 chloroplast Odontella sine.124dihydrolipoamide dehydrogenase27sp P11959 Bacillus stearothermophil.125DR-nm2316gp U29656 Homo sapiens126enolase17sp P29201 Haloarcula127fadBA operon (fatty acid oxidizing)8gb M74164 Escherichia coli128formate dehydrogenase9gp X54057 Wolinella succinogenes129fumarate reductase25pir[S34619 Thermoplasma acidophilum130glucosamine-fruc6-phos. aminotransferase24sp P39754 Bacillus subtilis	119	beta-lactamase	9	sp P00811	Escherichia coli
121cellulase26gp[L32742]Caldocellum saccharolyticum122cephalosporin C acylase9gp[A17015]Pseudomonas diminuta123chloroplast ATPase26gp[X60752]chloroplast Odontella sine.124dihydrolipoamide dehydrogenase27sp[P11959]Bacillus stearothermophil.125DR-nm2316gp[U29656]Homo sapiens126enolase17sp[P29201]Haloarcula127fadBA operon (fatty acid oxidizing)8gb[M74164]Escherichia coli128formate dehydrogenase9gp[X54057]Wolinella succinogenes129fumarate reductase25pir[S34619]Thermoplasma acidophilum130glucosamine-fruc6-phos. aminotransferase24sp[P39754]Bacillus subtilis	120	carbamoylphosphate synthetase, large	31	gi 1304392	Sulfolobus solfataricus
122cephalosporin C acylase9gp A17015 Pseudomonas diminuta123chloroplast ATPase26gp X60752 chloroplast Odontella sine.124dihydrolipoamide dehydrogenase27sp P11959 Bacillus stearothermophil.125DR-nm2316gp U29656 Homo sapiens126enolase17sp P29201 Haloarcula127fadBA operon (fatty acid oxidizing)8gb M74164 Escherichia coli128formate dehydrogenase9gp X54057 Wolinella succinogenes129fumarate reductase25pir[S34619 Thermoplasma acidophilum130glucosamine-fruc6-phos. aminotransferase24sp P39754 Bacillus subtilis	121	cellulase	26	gp L32742	Caldocellum saccharolyticum
123chloroplast ATPase26gp X60752 chloroplast Odontella sine.124dihydrolipoamide dehydrogenase27sp P11959 Bacillus stearothermophil.125DR-nm2316gp U29656 Homo sapiens126enolase17sp P29201 Haloarcula127fadBA operon (fatty acid oxidizing)8gb M74164 Escherichia coli128formate dehydrogenase9gp X54057 Wolinella succinogenes129fumarate reductase25pir S34619 Thermoplasma acidophilum130glucosamine-fruc6-phos. aminotransferase24sp P39754 Bacillus subtilis	122	cephalosporin C acylase	9	gp A17015	Pseudomonas diminuta
124dihydrolipoamide dehydrogenase27sp[P11959]Bacillus stearothermophil.125DR-nm2316gp[U29656]Homo sapiens126enolase17sp[P29201]Haloarcula127fadBA operon (fatty acid oxidizing)8gb[M74164]Escherichia coli128formate dehydrogenase9gp[X54057]Wolinella succinogenes129fumarate reductase25pir[S34619]Thermoplasma acidophilum130glucosamine-fruc6-phos. aminotransferase24sp[P39754]Bacillus subtilis	123	chloroplast ATPase	26	gp X60752	chloroplast Odontella sine.
125DR-nm2316gp U29656 Homo sapiens126enolase17sp P29201 Haloarcula127fadBA operon (fatty acid oxidizing)8gb M74164 Escherichia coli128formate dehydrogenase9gp X54057 Wolinella succinogenes129fumarate reductase25pir S34619 Thermoplasma acidophilum130glucosamine-fruc6-phos. aminotransferase24sp P39754 Bacillus subtilis	124	dihydrolipoamide dehydrogenase	27	sp P11959	Bacillus stearothermophil.
120enotase17sp[P29201]Haloarcula127fadBA operon (fatty acid oxidizing)8gb[M74164]Escherichia coli128formate dehydrogenase9gp[X54057]Wolinella succinogenes129fumarate reductase25pir[S34619]Thermoplasma acidophilum130glucosamine-fruc6-phos. aminotransferase24sp[P39754]Bacillus subtilis	125	DR-nm23	16	gp U29656	Homo sapiens
127Juarda operion (raity actio oxidizing)8gb[M/4104]Escherichia coli128formate dehydrogenase9gp[X54057]Wolinella succinogenes129fumarate reductase25pir[S34619]Thermoplasma acidophilum130glucosamine-fruc6-phos. aminotransferase24sp[P39754]Bacillus subtilis	120	enolase fad R A oppropr (fotty poid oridining)	1/	sp P29201 ab M74164	Haloarcula Escharichia coli
120formate denytrogenase5gp[X54057]wound a succoogenes129fumarate reductase25pir[S34619]Thermoplasma acidophilum130glucosamine-fruc6-phos. aminotransferase24sp[P39754]Bacillus subtilis	127 128	formate dehydrogenase	0 0	g0[1v1/4104] gp[X5/057]	Escherichia con Wolinella succinogenes
130glucosamine-fruc6-phos. aminotransferase24sp[P39754]Bacillus subtilis	129	fumarate reductase	25	pir 834619	Thermoplasma acidonhilum
	130	glucosamine-fruc6-phos. aminotransferase	24	sp P39754	Bacillus subtilis

Λ	Λ
+	÷

Table 1b. Continued

ID #	Hit identification	Probability (-E)	Accession	Organism
131	glyceraldehyde 3-phosphate dehydrogenase	49	sp P19315	Methanobacterium formic.
132	indolepyruvate decarboxylase	19	sp P2323	Enterobacter cloacae
133	L-isoaspartyl protein carboxyl methyltrans.	24	sp P24206	Escherichia coli
134	lactaldehyde dehydrogenase	11	gp M64541	Escherichia coli
135	lipoic acid metabolism (<i>lipB</i>)	37	sp P30976	Escherichia coli
136	membrane-associated ATPase alpha	49	gp J03218	Sulfolobus acidocaldarius
137	membrane-associated ATPase beta	58	sp P13052	Sulfolobus acidocaldarius
138	mmgC	15	gp[U29084]	Bacillus subtilis
139	NADH oxidase	25	sp P3/061	Enterococcus faecalis
140	NADH-plastoquinone oxidoreductase	20	sp P2/24	Synechocysus sp. Battus normagicus
141	nups/ nitrate reductase, alpha (narZ)	55	sp[Z .34922]	Kuuus norvegicus Escharichia coli
142	nitrate reductase, apria (<i>nurZ</i>)	58	sp[P42176]	Bacillus subtilis
143	nitrate reductase, seta (<i>nul 1</i>)	9	gn 749884	Bacillus subtilis
145	nitrate transporter	9	sp P38044	Svnechococcus sp
146	periplasmic divalent cation tolerance (<i>cutA</i>)	17	sp P36654	Escherichia coli
147	phospho-2-dehydro-3-deoxyheptonate aldol.	29	pir S21418	Bacillus subtilis
148	porphobilinogen deaminase	25	gp M95623	Homo sapiens
149	porphobilinogen synthase (hem B)	48	pir \$42531	Synechococcus sp.
150	precorrin decarbocylase	9	sp Q05632	Salmonella typhimurium
151	precorrin methylase	22	s p Q05630	Salmonella typhimurium
152	precorrin-3 methylase	17	sp P21922	Pseudomonas denitrificans
153	glucose-1-phosphate cytidylyltrans. (<i>Psa1p</i>)	11	gp U19608	Saccharomyces cerevisiae
154	thiosulfate sulfurtransferase	21	sp P16385	Saccharopolyspora erythr.
155	pyruvate phosphate dikinase	41	gp U02529	Entamoeba histolytica
150	pyruvate synthase	5/	pir 822397	Halobacterium nalobium
157	ribere 5 reperiete incorrect A	44	sp P11155	Zea mays
150	sucrase isomaltase	10	gp[U10436]	Rattus norvegicus
160	succinate dehydrogenase (ubig) iron-sulfur	15	sp P21914	Drosophila melanogaster
161	sulfite reductase	18	pir[\$27479]	Archaeoglobus fulgidus
162	enovl-CoA hydratase	19	gp[Z27079]	Caenorhabditis elegans
163	tungsten formylmethanofuran dehydrogenase	7	gp X87970	Methanobacterium thermo.
164	tartrate dehydratase	9	sp[P05847]	Escherichia coli
165	ttuD	16	gp U32375	Agrobacterium vitis
166	UDP-glucose 4-epimerase	29	sp P45602	Klebsiella pneumoniae
167	uroporphyrinogen III methyltransferase	11	gp U05002	Bacillus megaterium
168	x-aconitate hydratase	41	sp P09339	Bacillus subtilis
169	anaerobic dimethyl sulfoxide reductase A	13	gp U32785	Haemophilus influenzae
170	NADH dehydrogenase	45	sp P42026	Bovine
1/1	Carbamate kinase	31 44	sp P13982	Pseudomonas aeruginosa
172		44	sp[332028]	Lupinus arboraus
174	4-hydroxybutyrate dehydrogenase	14	gp I 36817	Alcaligenes eutrophus
175	formate DH nitrate-inducible alpha	12	pir S18213	Wolinella succinogenes
176	iron sulfur	26	gp[M96826]	Methanothermus fervidus
177	L-lactate dehvdrogenase	33	gp M93720	Plasmodium falciparum
178	adenosine triphosphatase	15	gp L33259	Helicobacter pylori
179	ATP sulfurylase	7	gp U07353	Penicillium chrysogenum
180	aldehyde:ferredoxin oxidoreductase	8	gp X79777	Pyrococcus furiosus
181	cytochrome b	14	sp P39480	Sulfolobus acidocaldarius
182	cytochrome c biogenesis	24	gp Z22517	Bradyrhizobium japonicum
183	cytochrome c oxidase assembly factor (cyoE)	30	gp U00013	Mycobacterium leprae
184	cytochrome c oxidase polypeptide I	35	sp P08681	Mitochondria Chlamydomonas
185	cytochrome oxidase I	57	gp U08900	Mitochondria Trichoniscus pusil.
180	cytochrome oxidase II	15	g1 155085	<i>Thermus thermophilus</i>
10/	b-lactate denydrogenase (cytochrome)	14	gp[Z07730]	Saccharomyces cerevisiae
180	electron transfer flavoprotein beta	40	sp 01/110	Paracoccus danitrificans
190	ferredoxin oxidoreductase	20	gp[X64521]	Halobacterium halobium
191	ferredoxin-nitrite reductase	14	pir \$30920	Nicotiana tomentosiformis
192	formaldehyde:ferredoxin oxidoreductase	57	gp[X83963]	Thermococcus litoralis
193	molybdenum- iron-sulfur flavoprotein ($codH$)	12	gp[X82447]	Oligotropha carboxidovora.
194	polysulfide reductase B	43	sp P31076	Wolinella succinogenes
195	polysulfide reductase C	5	sp P31077	Wolinella succinogenes
196	quinol oxidase polypeptide I/II	22	sp P39481	Sulfolobus acidocaldarius
197	sulfide dehydrogenase	7	sp Q06530	Chromatium vinosum
198	thiosulfate reductase	40	gp L32188	Salmonella typhimurium
199	alcohol dehydrogenase	26	sp P12311	Bacillus stearothermophil.
200	alcohol dehydrogenase family member Ke 6	19	pir A48154	Mus musculus

Table 1b. Continued

ID #	Hit identification	Probability (-E)	Accession	Organism
201	aldehyde dehydrogenase, cytosolic	34	sp P13601	Rattus norvegicus
202	(salicyl)aldehyde dehydrogenase	31	gp U19817	Arthrobacter globiformis
203	3-phosphoglycerate kinase	21	gp M55529	Methanothermus fervidus
204	4-aminobutyrate aminotransferase	13	sp P40829	Mycobacterium leprae
205	fructose 1,6-biphosphate aldolase (class II)	21	gp X14436	Escherichia coli
206	otsA	27	gp U15187	Mycobacterium leprae
207	pgk	49	gp X80178	Sulfolobus solfataricus
208	phosphoenolpyruvate-utilizing	27	gp S74619	Staphylothermus marinus
209	phosphoglycerate kinase	17	sp P20971	Methanothermus fervidus
210	pyruvate kinase	12	gp U12980	Saccharomyces cerevisiae
211	pyruvate, water dikinase	59	gp U08376	Pyrococcus furiosus
212	fumarate dehydrogenase	46	gp X/5402	Sulfolobus solfataricus
213	rumarate nydratase class II	41	sp[P39461]	Sulfolobus solfataricus Bacillus magatarium
214	transkatolasa	20	gp[D90044]	Homo sapians
215	lipoamide dehydrogenase	15	sp P35484	Acholenlasma laidlawii
210	npoannue denyurogenase	15	ap D37809	Streptomyces hygroscopic
217	pyruvate debydrobenase F1	24	sp P35488	Acholenlasma laidlawii
210	galactokinase	7	sp 001415	Homo sapiens
220	aconitate hydratase	50	sp[P37032]	Legionella pneumophila
221	carboxyphosphonoenolpyruvate mutase	36	gp X67953	Streptomyces hygroscopic.
222	citrate synthase	34	gp X55282	Thermoplasma acidophilum
223	citrate synthase II	42	gp U05257	Bacillus subtilis
224	hydroxymethylglutaryl-CoA reductase	54	gp X54658	Hevea brasiliensis
225	malic acid	24	gb M19485	Bacillus stearothermophil.
226	phosphorylated isocitrate dehydrogenase	57	pdb 4ICD	Escherichia coli
227	succinyl-CoA synthetase	49	sp P25126	Thermus aquaticus
Fatty acid	and phospholipid metabolism			
228	acetyl-CoA synthetase	71	sp P27095	Methanothrix soehngenii
229	acyl-CoA dehydrogenase	19	sp P15650	Rattus norvegicus
230	acyl-coA ligase(luciferase)	49	sp P29212	Escherichia coli
231	beta-ketothiolase	33	gp L37/61	Acinetobacter sp.
232	enoyl-coA nydratase, mitochondrial	14	gp X15958	Rattus norvegicus
233	gronylgoronyl pyrophosphate synthese	20	gp[A95015]	Sulfolobus acidocaldarius
234	geranyltranstransferase	12	sp[0.08291]	Bacillus stearothermonhil
235	hexaprenyl pyrophosphate synthetase	12	gn 726494	Saccharomyces cerevisiae
237	HMG-CoA reductase	25	sp[P34135]	Haloferax volcanii
238	isoprenyl diphosphate synthase	12	gp \$75695	Methanobacterium thermo.
239	long chain fatty acid coA ligase	12	gp U32686	Haemophilus influenzae
240	medium chain fatty acid coA ligase	31	sp Q00594	Pseudomonas oleovorans
241	thiolase	28	pir JC4032	Clostridium acetobutylicum
Purines, p	yrimidines, nucleosides, and nucleotides			
242	diadenosine tetraphosphatase (<i>apaH</i>)	12	gp X04711	Escherichia coli
243	mannosyltransferase B	6	gp D43637	Escherichia coli
244	deoxycytidine triphosphate deaminase	41	sp Q02103	Desulfurolobus ambivalens
245	CTD surplus and (surplus)	36	gp L04500	Eubacterium acidaminophi.
240	rDNA (adapaging N6 N6) dimethyltranef	02	gp[000021]	Mycobacierium ieprae Eachariakia aali
247	adenulate kinase	11	sp[F00992]	Escherichia coli Mathanococcus jannaschii
240	amidonhosphoribosyltransferase	11	gp[039862]	Arabidonsis thaliana
250	glu phosphoribosylpyrophosphate amidotrans	7	sp P00497	Racillus subtilis
250	GMP synthase (glutamine-hydrolyzing)	44	sp P29727	Bacillus subtilis
252	guaA	33	gp[U00015]	Mycobacterium leprae
253	phosphoribosylamine-glycine ligase	17	sp P12039	Bacillus subtilis
254	phosphoribosylformylglycinamidine synthase I	21	sp P12041	Bacillus subtilis
255	phosphoribosylformylglycinamidine synthase II	29	gp M85265	Lactobacillus casei
256	phosphoribosylglycinamide formyltransf.	11	pir \$37105	Arabidopsis thaliana
257	pur operon encoding purine biosynthesis	20	gp J02732	Bacillus subtilis
258	purL	12	gp U15182	Mycobacterium leprae
259	carbamoyl-phosphate synthetase	47	gp J05503	Mesocricetus auratus
260	dihydroorotate synthetase	19	gp U09990	Methanobacterium thermo.
261	dihydroorotate dehydrogenase	6	sp P32747	Schizosaccharomyces pom.
262	orotate phosphoribosyltransferase	19	pir A30492	Bacillus subtilis
263	phosphoribosyl-amp 1,6 cyclohydrolase	25	gp X82010	Rhodobacter sphaeroides
264	carbamyl phosphate synthetase	31	gp X87371	Saccharomyces cerevisiae
265	thiA	17	gp U26178	Bacillus subtilis
260	UDP galactose 4-epimerase	21	gp M94964	Klebsiella pneumoniae
207	UMP synthase	24	gp 022260	Caenornadaiatis elegans

Table 1b. Continued

ID #	Hit identification	Probability (-E)	Accession	Organism
268	uracil phosphoribosyl transferase	17	gp U10246	Toxoplasma gondii
269	cytidine deaminase	28	sp P19079	Bacillus subtilis
270	HPRT	8	gb M88110	Plasmodium falciparum
271	uridine phosphorylase	45	sp P12758	Escherichia coli
272	galactose-1-phosphate uridylyltransferase	21	sp P31764	Haemophilus influenzae
Regulate	ory functions	24	Droz col	
273	adenosylhomocysteinase	36	sp P10760	rat
274	ATP synthase	11	gp M22402 sp P15078	Sulfolodus aclaocalaarius Escharichia coli
275	GTP binding $(hfl X)$	14	ap[1100019]	Mycobacterium leprae
277	hydrobenase expression	9	sp[P31905]	Alcaligenes eutrophus
278	indoleacetamide hydrolase	12	sp P06618	Pseudomonas syringae
279	GTP-binding	10	gp Z49068	Caenorhabditis elegans
280	lactose operon repressor	6	sp P03023	Escherichia coli
281	mg11	14	gp U15635	Mus musculus
282	pleiotropic regulatory	44	gp M29002	Bacillus stearothermophilus
283	signal recognition particle (SPP54)	/ 11	gp M81646	Agrobacterium tumefacie.
204 D. U. A	·	11	sp r 57100	Arubiaopsis inaliana
Replicat	ION DMC1/rad51/rec 4	43	sp P25453	veast
285	recF	9	sp P24900	s typhimurium
287	recO	25	sp P30015	E coli
288	DNA repair helicase (rad25)	7	sp P19447	yeast
289	DNA repair (<i>rad2</i>)	8	sp P39750	Schizosaccharomyces pom.
290	DNA repair, ionizing radiation (XRCC1)	5	sp P18887	human
291	mutT	8	gp U00021	Mycobacterium leprae
292	O6-methyltransferase	9	sp P16455	human
293	methyltransferase (uvrC)	20 20	gp L29642	Pseudomonas fluorescens
294	possible G-T mismatches repair	20 16	sp 500115	Methanobacterium thermo
296	umuD/lerA-type UV SOS operon	9	sp P04153	Escherichia coli
297	SIR2 (silent information regulator)	7	sp P06700	veast
298	superoxide dismutase (Mn)	19	sp Q08713	Sulfolobus acidocaldarius
299	reverse gyrase	41	gp L10651	Sulfolobus acidocaldarius
300	DNA replication licensing factor	19	gp L41762	Drosophila melanogaster
301	DNA topoisomerase I	11	gp L27797	Bacillus subtilis
302	DNA topoisomerase (ATP hydrolysing)	21	sp P13099	Saccharomyces cerevisiae
303	modification methylase FNUDI	42	sp[P34906]	Erwinia carolovora Eusobacterium nucleatum
305	heat shock	25	gi 473965	Pyrococcus sp.
306	TF55 heat shock	80	sp P28488	Sulfolobus shibatae
307	TF56	57	gp L34691	Sulfolobus shibatae
308a	DNA polymerase	34	gp M74198	Thermococcus litoralis
308b	DNA polymerase	59	gi 807828	Pyrodictium occultum
308c	DNA polymerase	63	gp D38573	Pyrodictium occultum
308a	DNA polymerase	60 22	gp D385/4	Pyroaictium occuitium Sulfalabus salfatariaus
309	replication factor C small subunit	22	sp P25249	Homo saniens
310	replication factor C, large subunit	12	sp P35601	Mus musculus
311	adenylylsulfate reductase	36	gp X63435	Archaeoglobus fulgidus
312	alkyl hydroperoxide reductase	49	gp U36479	Sulfolobus sp.
313	diadenosine tetraphosphatase	19	gi 1054947	Sus scrofa
314	DNA-ligase	122	sp Q02093	Desulfurolobus
315	endonuclease III Hallidau investige DNA haliaaa	19	gp U11289	Bacillus subtilis
310	minichromosome maintenance	5	sp[032/10]	Saccharomycas caravisiaa
318	protein-L-isoaspartate (D-asp) O-methyltrans	13	sp[129409] sp[U09669]	Caenorhabditis elegans
319	structure specific endonuclease	10	sp P26793	Saccharomyces cerevisiae
320	XPBara	9	gp U29168	Arabidopsis thaliana
321	YSA1	7	sp Q01976	Saccharomyces cerevisiae
Transcri	ption			
322	RNA polymerase sigma factor (<i>ntrA</i>)r	11	gp X69959	Azorhizobium caulinodans
323	TATA-binding	57	gp U23419	Sulfolobus shibatae
324 325c	KNASE PH (tRNA nucleotidyltransferase)	15	gp L10328	Escherichia coli Mathanobastarium thormeo
325h	DNA-directed RNA polymerase II PpR10	41 16	sn P20100	Haloarcula marismortui
325c	DNA-directed RNA polymerase II, RpB10	17	pir D44126	Halobacterium marismort
326	DNA-directed RNA polymerase A	81	sp P31813	Thermococcus celer

46

Table 1b. Continued

ID #	Hit identification	Probability (-E)	Accession	Organism
327	DNA-directed RNA polymerase B	65	sp P11513	Sulfolobus acidocaldarius
328	DNA-directed RNA polymerase C	38	gp X14818	Sulfolobus acidocaldarius
329	DNA-directed RNA polymerase E	29	sp P39466	Sulfolobus acidocaldarius
330	RNA helicase	20	gp L01622	Escherichia coli
331	transcription activator	11	gp M73546	Bacillus subtilis
332	transcription elongation factor S-II	26	sp Q07271	Sulfolobus acidocaldarius
333	transcription factor IIB	18	gp U20899	Sulfolobus shibatae
334 T	transcriptional activator (<i>tenA</i>)	11	sp P25052	Baculus subtilis
Translation 335	on 5S rRNA	41	gb[M16530]	Sulfolopus sp
336	16S rRNA	43	gb M38637	Thermoplasma acidophilum
337	23S rRNA	65	emb[X05480]	Desulfurococcus mobilis
338	tRNA/5S rRNA gene cluster	12	emb X00916	Methanococcus vannielii
339	ala-tRNA	16	emb X05069	Thermoproteus tenax
340	ala-tRNA synthetase	64	sp P35029	Sulfolobus solfataricus
341	asn-tRNA synthetase	29	sp P10/23	Brugia malayi (nematode)
342	asp-IKNA	20	emb[X0/692]	Methanobacterium inermo.
343 344	asp-iKNA synthetase	50 50	gp[U52610] gp[X56234]	Escherichia coli
345	olu-tRNA	7	gb[L36898]	Saccharomyces cerevisiae
346	glu-tRNA synthetase	17	gp X07466	Homo sapiens
347	gly-tRNA	96	emb X14835	Thermofilum pendens
348	gly-tRNA synthetase	27	gp X78993	Saccharomyces cerevisiae
349	his-tRNA synthetase	19	gp L36863	Sulfolobus shibatae
350	ile-tRNA synthase	43	pir S21569	Methanobacterium thermo.
351	Iso-tRNA synthetase	27	sp P41368	Staphylococcus aureus
352	leu tPNA synthetese	25	emb[X050/1]	I hermoproteus tenax mitochondrial waast
354	met-tRNA synthetase	9	sp F11525 gb M26978	Mathanothermus fervidus
355	met-tRNA synthetase	13	pir A 25424	Saccharomyces cerevisiae
356	pro-tRNA	8	gb J01365	Saccharomyces cerevisiae
357	pro/glu-tRNA synthetase	28	gp M74104	Drosophila melanogaster
358	ser-tRNA	10	gb M97644	Methanopyrus kandleri
359	ser-tRNA synthetase	49	sp P26636	Cricetulus griseus
360	thr-tRNA synthetase	22	sp P18256	Bacillus subtilis
361	trp-tRNA	11	gb K02528	Halobacterium volcanii
302 363	val tPNA synthetase	13	gp L12221 sp P36420	Lactobacillus casai
364	elongation factor. EF 1 alpha	72	sp P35021	Sulfolobus solfataricus
365	elongation factor, EF 2	75	gp[X69297]	Sulfolobus solfataricus
366	Initiation factor, IF 4A	5	sp P10081	Saccharomyces cerevisiae
367	initiation factor, IF 5A	19	sp P28461	Sulfolobus acidocaldarius
368	L2, 50S ribosomal	16	sp P21479	Methanococcus vannielii
369	L3, 60/50S ribosomal	14	sp P35684	Oryza sativa
370	L4, 50S ribosomal	12	sp Q06845	Halobacterium halobium
371	L0,505 filosomal	21	sp P14030 sp P35023	Meinanococcus vaneili Sulfolobus solfataricus
372	L11 50S ribosomal	21	sp P35025	Sulfolobus solfataricus
374	L12, ribosomal	24	gp[X59038]	Sulfolobus solfataricus
375	L13, 50S ribosomal	22	sp P29198	Haloarcula marismortui
376	L15, 50S ribosomal	28	gp U16148	Thermoplasma acidophilum
377	L18, 50S ribosomal	50	sp P14033	Methanococcus vannielii
378	L23, ribosomal	19	gp Y00772	Methanococcus vannielii
379	L24, SUS ribosomal	26	sp P14034	Methanococcus vannielii
381	L29, 50S ribosomal	10	sp P04457 sp P14035	Mathanococcus vanalli
382	L 32e 50S ribosomal	17	sp P14549	Methanococcus vannielii
383	L37a, 60S ribosomal	16	pir S24170	Gallus gallus
384	S3, 30S ribosomal	15	sp P20281	Haloarcula marismortui
385	S5, 30S ribosomal	40	sp P14036	Methanococcus vannielii
386	S6, ribosomal	10	sp P17116	Escherichia coli
387	S7, 30S ribosomal	43	sp P35026	Sulfolobus solfataricus
388	S8, 30S ribosomal	35	sp P14038	Methanococcus vannielii
389 300	S10, 505 ribosomal	12	sp P20460	Huloarcula marismortul Sulfolobus acidocaldarius
390	S11, 505 HOSSOIIIai S12, 30S ribosomal	62	sp[r 39409] sp[P20161]	Suijoioous aciaocaiaarius Thermococcus celer
392	S12, 505 ribosomal	46	sp[005761]	Zea mays
393	S15, ribosomal	21	pir A35908	Homo sapiens
394	S17, ribosomal	14	pir A24028	Rattus norvegicus

1	0
4	0

Table 1b. Continued

ID #	Hit identification	Probability (-E)	Accession	Organism
395	S18, 40S ribosomal	20	gp Z46260	Saccharomyces cerevisiae
396	S26e, 40S ribosomal	13	pir \$47942	Saccharomyces cerevisiae
397	ribosomal alanine acetyltransferase	10	sp P09453	Escherichia coli
398	trmlp	26	gp Z48758	Saccharomyces cerevisiae
399	tRNA-splicing endonuclease ß	6	sp P16658	Saccharomyces cerevisiae
400	ARD1 N-acetyl transferase	12	gp X77588	Homo sapiens
401	multifunctional aminoacyl-tRNA synthetase	26	sp P07814	Homo sapiens
402	ATP-dependent protease	10	gp[L19301]	Myxococcus xanthus
403	beta-type proteasome	25	gp U22157	Methanosarcina thermophi.
404	dipeptidase	17	gp D13142	Sus scrofa
405	leuA	11	gp U18657	Haemophilus influenzae
406	lysosomal alpha glucosidase	28	gp Y00839	Homo sapiens
407	prolidase	11	gp Z34896	Lactobacillus delbrueckii
408	proteasome, beta	18	sp P28061	Thermoplasma acidophilum
409	protein serine/threonine phosphatase 2A	14	gp[X56261]	Saccharomyces cerevisiae
410	TVG	29	gp D13178	Thermoactinomyces vulga.
411	sulfite reductase	10	pir \$27478	Archaeoglobus fulgidus
Transport	and hinding proteins	10	PH[027.170]	
412	$2'_{-5'}$ oligoadenvlate binding	25	nir 852166	Homo sapiens
413	annR	11	gn[U20909]	Bacillus subtilis
414	ATP-binding	25	gp 725798	Bacillus subtilis
415	cvsA	20	gp[Z25750] gp[M32101]	Escherichia coli
416	maltose hinding periplasmic	11	sp[P10576]	Salmonalla typhimurium
410	nat Δ/B	31	ap[130873]	Bacillus subtilis
418	n87	16	gp[030075]	Bos sp
410	por	16	gp[347919]	Bos sp. Escharichia coli
419	selenium binding: ap56	26	sp[D 17563]	Mus musculus
420	transport social alpha	15	sp[117505]	Duranomonas salina
421	argining permassa substrate hinding	13	spjr 56579	I yrenomonus suinu Listaria monoautoganas
422	a ginne permease, substrate-officing	17	sp[0008005]	Saacharomycas coravisiae
423	dei A D	18	$\frac{\text{sp}[1,52691]}{\text{sp}[\mathbf{V}56678]}$	Pacillus subtilis
424	dinantida transport quatam normaaga dan P	24	gp[A30076]	Each anishing and
425	dipeptide transport system permease dppB	24	sp[F5/510]	Escherichia coli Haamophilus influenzae
420	aluterine normana	32	gp[U1/293]	Each anishing a shi
427	glutamine permease	2/	gp[A14160]	Escherichia coli
420	grutamme-omonig periplasinic	19	sp[F10544]	Escherichia coli
429	oranched-chain amino acid transport	40	sp P30294	Saimoneila typnimurium
430	spook operon	20	gp[NI57089]	Bacillus sublills
431	oligopeptide perm. (<i>oppC</i> , sporulation initiation)	20	sp P24139	Baculus subtilis
432	oligopeptide transport permease <i>appB</i>	10	sp P42062	Bacillus sbutilis
433	oligopeptide transport permease <i>app</i> C	8	sp P42063	Baculus subtilis
434	oligopeptide transport ATP binding appF	32	sp P42065	Bacillus subtilis
435	oligopeptide transport ATP-binding oppF	51	sp P08007	Salmonella typhimurium
436	phosphonates transport ATP-binding phnC	21	sp P16677	Escherichia coli
437	phosphonates transp. sys. permease (phnE)	9	sp P16683	Escherichia coli
438	translocase SecY	23	gp X85020	Sulfolobus acidocaldarius
439	proline/betaine transporter	24	gp U14003	Escherichia coli
440	putrescine transport system permease (<i>potl</i>)	10	sp P31136	Escherichia coli
441	spermidine/putrescine-binding periplasmic	10	gp U32731	Haemophilus influenzae
442	phosphate transport ATP-binding <i>pstB</i>	45	sp P07655	Escherichia coli
443	lactose transport ATP-binding lack	13	sp Q01937	Agrobacterium radiobacter
444	sterol carrier -2	24	gp M75884	Homo sapiens
445	sugar-binding	7	gp M77351	Streptococcus mutans
446	cation efflux system CZCD	11	sp P13512	Alcaligenes eutrophus
447	copper transport atp-binding nosf	17	sp P19844	Pseudomonas stutzeri
448	mercuric reductase	14	gp X73112	Pseudomonas fluorescens
449	<i>n</i> -ferritin repressor (FRP)	48	gb M95815	Oryctolagus cuniculus
450	Na ⁽⁺⁾ -ATPase B	17	sp Q08637	Enterococcus hirae
451	periplasmic divalent cation tolerance	14	gp Z36905	Escherichia coli
452	potassium/copper transporting ATPase	32	sp P32113	Enterococcus hirae
453	AP56	39	gp S56599	Mus sp.
454	rhodanese-like (cysA)	41	gp M29612	Saccharopholyspora erythr.
455	SN-glycerol-3-phos. transport ATP-binding	56	sp P10907	Escherichia coli
456	sulfate permease (<i>cysA</i>)	19	sp P14788	Anacystis nidulans
Other cat	egories			
457	7-alpha-hydroxysteroid hydrogenase	8	pir A42468	Eubacterium sp.
458	acylphosphatase	13	sp P07032	Gallus gallus
459	cheA/W/Y	22	gp U30501	Thermotoga maritima

Table 1b. Continued

ID #	Hit identification	Probability (-E)	Accession	Organism
460	FixA (nitrogen fixation)	8	sp P31573	Escherichia coli
461	FixB (nitrogen fixation)	41	gp M91817	Clostridium acetobutylic.
462	FixC (nitrogen fixation)	30	sp[P31575]	Escherichia coli
463	FixX (nitrogen fixation)	33	pir \$49190	Azotobacter vinelandii
464	homocitrate synthase (<i>nifV</i> , nitrogen fixation)	39	sp P05342	Azotobacter vinelandii
465	phosphomannomutase	15	gp M60873	Pseudomonas aeruginosa
466	ABC transporter (<i>pstC</i> -1)	14	gp Z47982	Mycobacterium tuberculos.
467	sec61	12	gp X77805	Pyrenomonas salina
468	threonine dehydratase biosynthetic	6	sp Q02145	Lactococcus lactis
469	vacuolar H ⁺ phosphatase	49	gp M81892	Arabidopsis thaliana
470	multidrug resistance	18	gp Z49126	Caenorhabditis elegans
471	phenylacrylic acid decarboxylase	38	gp L09263	Saccharomyces cerevisiae
472	daunorubicin-doxorubicin polyketide synth.	16	gp L35560	Streptomyces peucetius
473	pyrroline-5-carboxylate reductase	6	pir JC2078	Thermus aquaticus
474	amylomaltase	43	gp J01796	Streptococcus pneumoniae

quence contigs. The distance between these sequence markers is 11.5 Kb on average, while the average distance between the 200 sequence gaps will be 8.5 Kb, or in other words the average contig size will be 8.5 Kb. Thus, a randomly placed contig is likely to fall across a sequence marker which consequently maps that contig to particular fosmid clones. The few contigs which do not cross a sequence marker will be mapped by hybridization to the 96 fosmid clones.

Discussion

The mapping and sequencing of the *Pyrobaculum aerophilum* genome will provide valuable resources to researchers in many disciplines. We report here a list of 474 putative genes and a genome map. Access to these gene sequences now will provide researchers with a valuable tool for a multitude of studies.

A few whole genome sequencing projects have already been completed (Bult et al. 1996; Fleischmann et al. 1995; Fraser et al. 1995) and several others are in progress (Burland et al. 1995; Charlebois et al. 1996; Coulson 1996; Levy 1994). These efforts are concentrated on key organisms dispersed throughout the domains of life and will be important in elucidating the evolutionary path of modern organisms.

Pyrobaculum aerophilum is a member of the crenarchaea, which were thought to be a branch of the archaea but whose position in the phylogenetic tree of life has been a subject of controversy. It is uncertain whether the archaea are truly monophyletic (sharing a common ancestor distinct from other groups) or whether the crenarchaea in fact only share a distinct common ancestor with the eukaryotes and not with the rest of the archaea (Rivera and Lake 1992). Tracing the evolution of a few single genes is not sufficient to resolve these questions since

many factors complicate the analyses including gene duplications, exon shuffling, and horizontal transfers (i.e., transfer of a gene from one organism to another). Accumulating large amounts of sequence data from representative organisms is the most efficient way to unscramble these evolutionary puzzles.

Complete genome sequences from representative organisms also provide researchers with the information needed to quickly isolate their gene of interest from virtually any organism. Furthermore, researchers will frequently be able to trace the evolutionary history of their favorite gene directly from the sequence databases. Quickly determining the absence of a particular gene in an organism is in itself a powerful tool. Currently, this takes weeks of hybridization experiments and still leaves the researcher uncertain of the gene's absence.

Although a detailed analysis comparing this sequence to other completed sequences will wait for the final edited sequence, there are already some surprises. One example is in the area of repair systems. We have identified tags to members of many of the major repair systems found in E.coli and the human, including those involved in excision repair, repair of oxidative damage, and methylated bases. Pyrobaculum aerophilum is a particularly interesting organism for the study of repair systems because of its remarkable similarity, at the amino acid sequence level, to eukaryotes. For instance, we have found homologs to genes involved in human repair disorders, such as xeroderma pigmentosum. Interestingly, despite the fact that we have sequenced more than 95% of the Pyrobaculum genome, we have not detected any genes involved in either the mismatch repair systems or deamination of cytosines. The latter is expected to be enhanced at high temperatures. The reported complete genome sequence of Methanococcus jannaschii (Bult et al. 1996) also fails to list genes involved in either of these repair pathways. Therefore, either the Archaea have a different way of repairing replication mismatches, or we must entertain the fascinating possibility that they have high mutation rates in vivo.

Materials and methods

Laboratory procedures were performed as described in the laboratory manual (Sambrook et al. 1989) unless otherwise mentioned.

Genomic DNA extraction

Pyrobaculum aerophilum (type strain IM2; DSM 7523) was grown anaerobically and pelleted at Karl Stetter's laboratory, University of Regensburg, Germany, under the previously published conditions (Völkl et al. 1993). Approximately 0.5 g (wet weight) of the cell pellet was chilled with liquid nitrogen, ground with a mortar and pestle, and resuspended in 500 ml TE buffer. Lysozyme (5 ml, to give 50 mg/ml) was added to the cell suspension, followed by the addition of 30 ml 20% w/v SDS, and the lysis mixture was incubated at 37°C for 1 h. Lysed cells were phenol extracted once, ethanol precipitated, then resuspended in TE at a final concentration of 0.5 mg/ml.

Fosmid library construction

For the construction of the genomic fosmid library, 10µg of genomic DNA was resuspended in 500µl Sau3AI buffer by gently swirling on a rotating platform for 3 h at 4°C followed by 10min at 37°C. The DNA was restriction digested by adding 0.1 unit Sau3AI enzyme (New England Biolabs, Beverly, MA, USA) and incubating at 37°C for 15 minutes, followed by addition of another 0.1 U Sau3AI enzyme with additional incubation at 37°C for 15 minutes. Then, the DNA was extracted by chloroform, ethanol precipitated, and resuspended in 20µl TE. Analysis on an agarose gel showed that most fragments were within 20-40 Kb. From this DNA stock a 3-µl aliquot was taken and dephosphorylated in a 25-µl reaction with 2 units HK-Phosphatase (Epicentre Technologies, Madison, WI, USA) at 30°C for 1h, followed by heat inactivation of the phosphatase at 65°C for 15 min. Fosmid arms were prepared as described previously (Kim et al. 1992). Ligations were done with a large molar excess of vector DNA (50-100×) to minimize the probability of chimera formation. BamHI digested fosmid arms $(1 \mu g)$ were ligated with approximately 15 ng of the genomic DNA in a 15-µl reaction with 1 unit of T4 DNA ligase (Epicentre Technologies). The ligation product thus generated (3µl) was packaged into phage particles using GigaPack XL in-vitro packaging extract (Stratagene, La Jolla, CA, USA) as instructed by the vendor. The packaged particles were suspended in 500µl SM buffer. E coli DH10B cells were grown in LB + 0.2% maltose + 10 mM MgSO₄ to an optical density at 600 nm (OD₆₀₀) of 0.5–0.9, centrifuged at 4°C, and resuspended in 10 mM MgSO₄ to a final concentration of $OD_{600} = 1.0$. Undiluted packaging lysate (25 µl) was mixed with 25 µl resuspended cells and kept at room temperature for 30 min. LB (200 µl at 37°C) was then added and kept at 37°C for 1h with gentle shaking every 15 min. Cells were centrifuged and resuspended in 50µl LB, which was subsequently plated on LB plates containing 10μ g/ml chloramphenicol. Emerging colonies from the plates were inoculated into eight 96-well microtiter plates containing LB + 10% glycerol, which were then grown at 37°C overnight and kept frozen at -70° C until use.

Construction of pUC18 library

Genomic DNA (10µg) was diluted in 300µl TE and sheared by sonicating with a Branson cell disrupter 200 sonicator at setting 3.5 for 2s. This sheared DNA (180µl) was mixed with $180\mu l \simeq Bal31$ buffer and kept at $30^{\circ}C$ for 15 min. Then, 0.015 units of *Bal*31 enzyme (New England Biolabs) was added and the mixture incubated at 30°C for 10min. The mixture was then diluted to 660µl with TE, extracted with phenol: chloroform and chloroform, and precipitated with ethanol. One fifth of the remaining DNA was run on a preparative 1.2% SeaKem GTG (FMC Bioproducts, Rockland, ME, USA) agarose gel. After brief staining with ethidium bromide, fragments of 1.5-2 Kb were excised and recovered by QiaexII (Qiagen, Chatsworth, CA, USA) binding following manufacturers instructions. Elution was in 20µl 10mM Tris pH 8.5. This entire volume was used for subcloning into the SmaI site of pUC18 (Ready-To-Go subcloning kit, Pharmacia Biotech, Piscataway, NJ, USA). These subclones were randomly sequenced from both ends (unpublished) using fluorescent multiplex automated sequencers (ABI 373, Applied Biotechnology, Foster City, CA, USA), and the public databases were searched with these sequences by BLAST (Altschul et al. 1990; States et al. 1991) to find similarities to previously known sequences.

Generation of sequence marker probes by PCR amplification

Typically, 0.005 pmoles of template DNA, 20 pmoles of each primer, and 1 unit of TaqI polymerase (Gibco BRL, Gaithersburg, MD, USA) were combined in a 25-µl reaction. Polymerase chain reaction (PCR) products were isolated by running the PCR reactions on 1% SeaPlaque low melting point agarose gel (FMC Bioproducts, Rockland, ME, USA) in 1× TAE and excising DNA bands. The DNA fragments, in low melting point agarose, were radiochemically labeled using the DecaPrime kit (Ambion, Austin, TX, USA) as instructed by the vendor except that each reaction was scaled down to half the instructed values.

Generation of riboprobes

Fosmid clone templates were prepared from 3 ml LB cultures using an AutoGen 740 automated miniprep machine (Integrated Separation Systems, Natick, MA, USA). Half of the DNA isolated from a 3-ml culture was used in a 100- μ l digestion reaction with 5 units *Hinc*II restriction enzyme. Digested products were ethanol precipitated and resuspended in 5 μ l TE or water. This concentrated template (1.5 μ l) was used in the standard transcription reaction described in the RiboScribe kit protocol (Epicentre Technologies), except that all volumes were scaled down to half, giving a final reaction size of 10μ l. After the 2-h labeling period the entire contents of the reactions were used for hybridization.

Hybridization

Fosmid library colony membranes were prepared by gridding the entire library, represented by eight 96-well microtiter plates, at high density onto single 8 cm × 12 cm nylon membranes using a Biomek 1000 laboratory workstation (Beckman Instruments, Fullerton, CA, USA). Labeled probes (PCR products or riboprobes, as described in the previous section) were transferred directly into hybridization chambers without purification. Colony hybridizations were incubated overnight in 1 M NaCl, 50 mM Tris-HCl pH 8.0, 5 mM EDTA pH 8.0, 10% PEG 8000, at 65°C. The membranes were washed in 2× standard saline citrate (SSC), 0.1% sodium dodecyl sulfate (SDS) at 65°C for 30 min, followed by another wash in 0.2× SSC, 0.1% SDS at 65°C for 30 min. The membranes were exposed to X-ray film from 3h to overnight.

Fingerprinting

Restriction fingerprint analysis of fosmid clones was done essentially according to the previously described modified protocol that was adapted from the original cosmid fingerprinting procedure (Coulson et al. 1986; Sulston et al. 1988). The fingerprint gels were digitized using a Phosphorimager (Molecular Dynamics, Foster City, CA, USA) and analyzed using the contigc program (http:// www.sanger.ac.uk).

Assembly of fosmid contig

Ordering and alignment of contigs was performed manually and based primarily on the results of hybridization of the markers to the fosmid clones. Two clones are considered overlapping if they are hit by one or more common probes. Ambiguities in the overlapping relations were resolved by fingerprint analysis or additional hybridization with riboprobes that are specific to clone ends.

Acknowledgments We thank Dr. Phil Green and Dr. David Gordon for use of their Phred and Phrap programs, Dr. David Mathog for helping us use the GCG DNA sequence analysis programs and Jack Gerson for critical reading of the manuscript. This work was supported by funding from the United States Department of Energy to M.I.S. and from the Office of Naval Research to J.H.M. (N00014-95-1-0938).

References

- Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990) Basic local alignment search tool. J Mol Biol 215:403–410
- Bult CJ, White O, Olsen GJ, Zhou L, Fleischmann RD, Sutton GG, Blake JA, FitzGerald LM, Clayton RA, Gocayne JD, Kerlavage AR, Dougherty BA, Tomb JF, Adams MD, Reich CI, Overbeek R, Kirkness EF, Weinstock KG, Merrick JM, Glodek A, Scott JL, Geoghagen NSM, Weidman JF, Fuhrmann JL, Venter JC et al. (1996) Complete genome sequence of the methanogenic archaeon, Methanococcus jannaschii Science 273:1058–1073
- Burland V, Plunkett GR, Sofia HJ, Daniels DL, Blattner FR (1995) Analysis of the *Escherichia coli* genome VI: DNA sequence of the region from 92.8 through 100 minutes. Nucleic Acids Res 23:2105– 2119
- Charlebois RL, Gaasterland T, Ragan MA, Doolittle WF, Sensen CW (1996) The *Sulfolobus solfataricus* P2 genome project. FEBS Lett 389:88–91
- Coulson A (1996) The *Caenorhabditis elegans* genome project. *C. elegans* Genome Consortium. Biochem Soc Trans 24:289–291
- Coulson A, Sulston J, Brenner S, Karn J (1986) Toward a physical map of the genome of the nematode *Caenorhapditis elegans*. Proc Natl Acad Sci USA 83:7821–7825
- Fleischmann RD, Adams MD, White O, Clayton RA, Kirkness EF, Kerlavage AR, Bult CJ, Tomb JF, Dougherty BA, Merrick JM et al. (1995) Whole-genome random sequencing and assembly of *Haemophilus influenzae* Rd Science 269:496–512
- Fraser CM, Gocayne JD, White O, Adams MD, Clayton RA, Fleischmann RD, Bult CJ, Kerlavage AR, Sutton G, Kelley JM et al. (1995) The minimal gene complement of *Mycoplasma genitalium* Science 270:397–403
- Kim UJ, Shizuya H, de Jong PJ, Birren B, Simon MI (1992) Stable propagation of cosmid sized human DNA inserts in an F factor based vector. Nucleic Acids Res 20:1083–1085
- Levy J (1994) Sequencing the yeast genome: an international achievement. Yeast 10:1689–1706
- Riley M (1993) Microbiol Rev 57:862
- Rivera MC, Lake JA (1992) Evidence that eukaryotes and eocyte prokaryotes are immediate relatives. Science 257:74–76
- Robb FT, Place AR (eds) (1995) Archaea : A laboratory manual. Thermophiles. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, 3 vol
- Sambrook J, Fritsch EF, Maniatis T (1989) Molecular cloning : a laboratory manual, 2nd edn. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 3 vol
- States DJ, Gish W, Altschul SF (1991) Improved sensitivity of nucleic acid databases searches using application-specific scoring matrices. Methods 3:66–70
- Stetter KO (1992) Life at the upper temperature border. In: Tran Than Van JK, Mounolou JC, Schneider J, McKay C (eds) Colloque Interdisciplinaire du Comite National de la Recherche Scientifique, Frontiers of Life
- Sulston J, Mallett F, Staden R, Durbin R, Horsnell T, Coulson A (1988) Software for genome mapping by fingerprinting techniques. Comput Appl Biosci 4:125–32
- Völkl P, Huber R, Drobner E, Rachel R, Burggraf S, Trincone A, Stetter KO (1993) *Pyrobaculum aerophilum* sp. nov., a novel nitratereducing hyperthermophilic archaeum. Appl Environ Microbiol 59:2918–2926
- Woese CR (1987) Bacterial evolution. Microbiol Rev 51:221-271