

Complete Genome Sequence of an Aerobic Hyper-thermophilic Crenarchaeon, *Aeropyrum pernix* K1

Yutaka KAWARABAYASI,^{1,2,*} Yumi HINO,¹ Hiroshi HORIKAWA,¹ Syuji YAMAZAKI,¹ Yuji HAIKAWA,¹ Koji JIN-NO,¹ Mikio TAKAHASHI,¹ Mitsuo SEKINE,¹ Sin-ichi BABA,¹ Akiho ANKAI,¹ Hiroki KOSUGI,¹ Akira HOSOYAMA,¹ Shigehiro FUKUI,¹ Yoshimi NAGAI,¹ Keiko NISHIJIMA,¹ Hidekazu NAKAZAWA,¹ Minako TAKAMIYA,¹ Sayaka MASUDA,¹ Tomomichi FUNAHASHI,¹ Toshihiro TANAKA,¹ Yutaka KUDOH,¹ Jun YAMAZAKI,¹ Norihiro KUSHIDA,¹ Akio OGUCHI,¹ Ken-ichi AOKI,¹ Kenji KUBOTA,¹ Yoshinobu NAKAMURA,¹ Norimichi NOMURA,³ Yoshihiko SAKO,³ and Hisasi KIKUCHI¹

National Institute of Technology and Evaluation, 2-49-10 Nishihara, Shibuya, Tokyo 151-0066, Japan,¹ National Institute of Bioscience and Human-Technology, 1-1 Higashi, Tsukuba, Ibaraki 305-0046, Japan,² and Kyoto University, Sakyo-ku, Kyoto 606-8502, Japan³

(Received 9 March 1999)

Abstract

The complete sequence of the genome of an aerobic hyper-thermophilic crenarchaeon, *Aeropyrum pernix* K1, which optimally grows at 95°C, has been determined by the whole genome shotgun method with some modifications. The entire length of the genome was 1,669,695 bp. The authenticity of the entire sequence was supported by restriction analysis of long PCR products, which were directly amplified from the genomic DNA. As the potential protein-coding regions, a total of 2,694 open reading frames (ORFs) were assigned. By similarity search against public databases, 633 (23.5%) of the ORFs were related to genes with putative function and 523 (19.4%) to the sequences registered but with unknown function. All the genes in the TCA cycle except for that of alpha-ketoglutarate dehydrogenase were included, and instead of the alpha-ketoglutarate dehydrogenase gene, the genes coding for the two subunits of 2-oxoacid:ferredoxin oxidoreductase were identified. The remaining 1,538 ORFs (57.1%) did not show any significant similarity to the sequences in the databases. Sequence comparison among the assigned ORFs suggested that a considerable member of ORFs were generated by sequence duplication. The RNA genes identified were a single 16S–23S rRNA operon, two 5S rRNA genes and 47 tRNA genes including 14 genes with intron structures. All the assigned ORFs and RNA coding regions occupied 89.12% of the whole genome. The data presented in this paper are available on the internet homepage (<http://www.mild.nite.go.jp>).

Key words: aerobic hyper-thermophile; crenarchaeota; *Aeropyrum pernix* K1; genome sequencing; whole genome shotgun method

1. Introduction

The complete genome sequences of four kinds of thermophilic and anaerobic archaea have been reported,^{1–4} including the one sequenced by our group (*Pyrococcus horikoshii* OT3).¹ Among these archaea, *Pyrococcus horikoshii* OT3 is a hyper-thermophilic strain with an optimal growth temperature of 98°C, whereas the other three strains grow in the range of 60 to 85°C. To accumulate the information on the genetic system of or-

ganisms living at hyper-thermophilic conditions, we now analysed the genome of another archaeon, *Aeropyrum pernix* K1. This strain has been isolated in 1993 from coastal solfatric thermal vent in Kodakara-jima Island in Kyusyu, Japan.⁵ Like *P. horikoshii* OT3, it grows in the range of 90 to 98°C with an optimal temperature of 95°C. However, compared with *P. horikoshii* OT3 that belongs to an anaerobic euryarchaeota, *A. pernix* K1 is an aerobic strain classified as crenarchaeota. Therefore, the sequence data on its genome should also provide information on the origin of the oxygenic respiratory system.

To sequence the entire genome, we first constructed the restriction maps of the genome, using three kinds of 8-base cutters, *Pac* I, *Sgf* I and *Swa* I. Then the shotgun libraries with short and long inserts were constructed from both the entire genomic DNA and restriction fragments

* Communicated by Mituru Takanami

† To whom correspondence should be addressed. Tel. +81-3-3481-8951, Fax. +81-3-3481-8424, E-mail: kyutaka@kazusa.or.jp

‡ The entire genome sequence has been deposited in DDBJ/Genbank/EMBL databases under the accession numbers AP000058-AP000064.

generated by *Sgf*I and *Swa*I. The single-run sequence data of the shotgun clones were accumulated, and the entire genomic sequence was directly deduced by assembling the sequence data with the software PhredPhrap,^{6,7} followed by gap-filling by walking with long-insert shotgun clones.

2. Materials and Methods

2.1. Bacterial strains and genomic DNA

The *A. pernix* K1 strain deposited in the Japan Collection of Microorganisms (JCM number 9820) was used for genome sequencing in this study. The cells were inoculated into 100 ml of the JXT medium⁵ which was prepared in 500-ml Erlenmeyer flasks, and cultured at 90°C with vigorous shaking. The genomic DNA was isolated principally based on the method of Sako et al.⁵ *Escherichia coli* DH10B was used for the preparation of plasmid clones.

2.2. Construction of shotgun clones

The *A. pernix* K1 genomic DNA was sonicated to three different extents (5, 10 and 20 sec at the L position of Biorupter; Cosmo Bio, Tokyo, Japan). Followed by size-fractionation with agarose gel electrophoresis, the fractions 0.8 to 1.2 kb and 2.0 to 2.5 kb were independently cloned into the *Hinc*II site of pUC118 (short and long fragment-shotgun libraries).

For the construction of the shotgun libraries from restriction fragments, the genomic DNA was prepared in the agarose-plug as described by Smith et al.,⁸ and after complete digestion of DNA in the plugs with *Sgf*I or *Swa*I, the digests were resolved by 0.4% agarose gel electrophoresis. The band regions were cut out, and agarose gel was dissolved by incubation overnight at 50°C with Agarase (FMC, Rockland, ME, USA). The resulting solution containing DNA fragments was sonicated, and the short fragment-shotgun libraries were prepared as above, except for *Sgf*I fragment A (1015 kb) from which both the short and long fragment-shotgun libraries were constructed.

2.3. DNA sequencing

Plasmid DNA in each shotgun library was prepared by the Autogen 740 automatic DNA preparation system (Autogen, Framington, MA, USA). The sequencing reaction was performed using two kinds of cycle sequencing kits, dye-primer cycle sequencing kit and dye-terminator cycle sequencing kit, and ABI-DNA sequencers (373XL and 377XL; Perkin-Elmer ABI, Foster City, CA, USA). For the sequence determination at the gaps, the clones bridging two contigs were screened from the long fragment-shotgun libraries, and sequenced by the dye-terminator method.

2.4. Data assembly and computational analysis

The accumulated files of single-run sequence data were first treated using the software Phred⁶ to eliminate any contaminated sequences derived from *E. coli* or vector DNAs, and the treated data were assembled into contigs using the software Phrap.⁷ The assembled sequences were split into 30-kbp segments, and the sequence in each segment was re-assembled and edited by Sequencher (Gene Codes, Ann Arbor, MI, USA).

The criteria used for the assignment of potential coding regions on the genomic sequence are as follows. All the ORFs of larger than 100 sense codons starting with ATG or GTG were first assigned, and then ORFs of 50 to 99 sense codons were taken, if the sequences showed any similarity to the coding sequences in the databases or to the protein motif sequences. Similarity search of assigned ORFs was performed by the Smith-Waterman algithrm.⁹ The databases used for similarity search were Genbank release 109, EMBL relaease 56.0, Swiss-Prot release 36.0, PIR release 57.0 and Owl release 31.0.

3. Results and Discussion

3.1. Construction of physical map

The genomic DNA of *A. pernix* K1 was digested with three kinds of 8-base cutters, *Sgf*I, *Pac*I and *Swa*I, which generated 5, 8 and 13 fragments, respectively. The resulting fragments were resolved and purified by the CHEF Mapper gel electrophoresis system (BioRad, Richmond, CA, USA). Then the fragments were ordered by hybridization analysis among the fragments and their double-digestion products.

Figure 1 shows the physical maps which were finally confirmed by the determination of the entire genomic sequence. Three of the *Swa*I sites which were present in the close vicinity of other *Swa*I sites were assigned on the basis of the deduced sequence. Fragments A to E generated by *Sgf*I and fragments F to H generated by *Swa*I fragments were used for the construction of restriction fragment-specific shotgun libraries.

3.2. Determination of entire genome sequence

For deduction of the entire genome sequence, two kinds of shotgun libraries, one with approximately 1-kb inserts and the other with approximately 2.2-kb inserts, were prepared from both the genomic DNA and restriction fragments as mentioned above. The clones from the short fragment-shotgun libraries were single-run sequenced from one end and those from the long fragment-shotgun libraries from both the ends. A total of 29,000 single-run sequence data, corresponding to approximately 8 times the entire genomic length, were accumulated and used for assembling by PhredPhrap. The contigs primarily assembled by PhredPhrap constituted 14 islands. For the gap-filling, the clones carrying

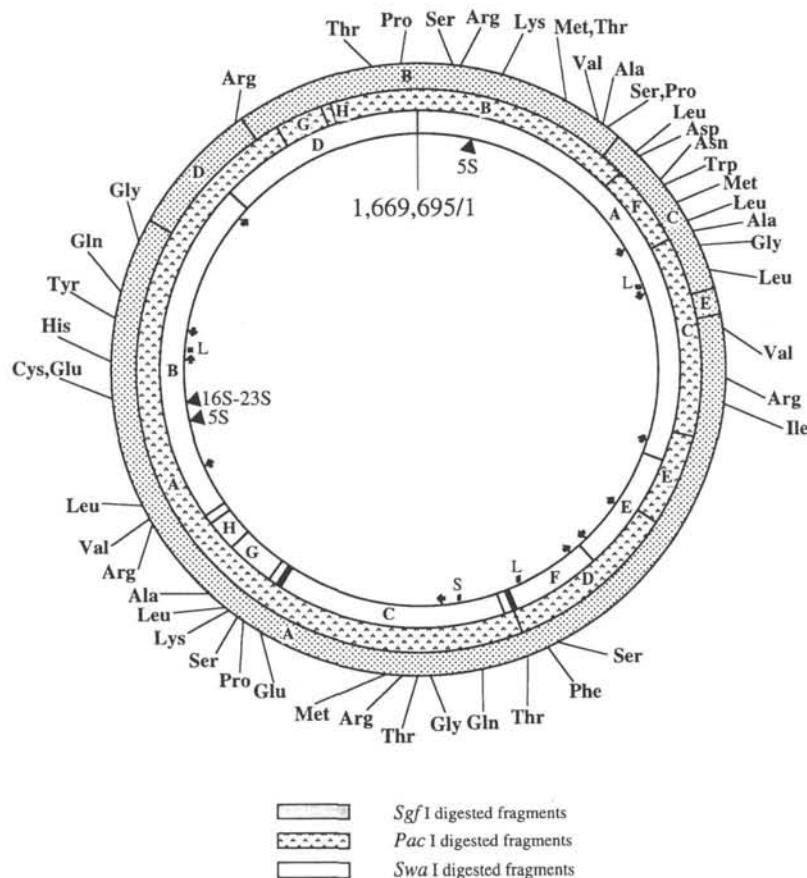


Figure 1. The physical map of the *A. pernix* K1 genome and location of RNA genes and characteristic repeating units. The restriction sites for *Sgf* I, *Pac* I and *Swa* I are indicated in the circular maps of the genome. The nucleotide positions are numbered in the clockwise direction from the junction of the *Swa* I A and D fragments (the position is given as 1,669,695/1). The positions and species of the tRNA genes are indicated by bars with the names of chargeable amino acids. The position of the 5S and 16S-23S rRNA genes are shown by triangles labeled with 5S and 16-23S, and those of type A and type B repeating units by boxes labeled L and S, respectively. The arrows along the inner circle indicate the position and direction of the type C repeating units.

the end sequences of contigs were screened from the long fragment-shotgun libraries and sequenced. All the gaps were closed by sequencing two clones per gap. The gap lengths finally confirmed were 59 to 553 bp. The criterion taken for the sequence confirmation was coincidence of the sequences deduced from different shotgun clones.

To confirm the authenticity of the genomic sequence constructed, every 15-kb region of the genome was directly amplified from the genomic DNA by long PCR, and the occurrence of restriction sites were compared with those deduced from the sequence data.

In the course of editing, it was noted that different nucleotide signals appear at the same position and that the results were reproducible. Such positions were identified at least 35 different positions along the entire genome: the mixture of G and A signals at 15 positions, that of C and T signals at 9 positions and that of other base signals at 11 positions. The occurrence of different nucleotides at the same position could be resolved by sequencing of the PCR products directly amplified from the genomic DNA. We assume that this is due to the presence of multiple

copies of the genome in a single cell. In the sequence registered in the databases, the nucleotides giving the major signals are indicated.

The total length of the genome finally confirmed was 1,669,695 bp. The nucleotide position was numbered from the one end of the *Swa* I restriction site located on the *Sgf* I B and *Pac* I B fragments, as shown in the physical map in Fig. 1. The distribution of ACGT along the strand of the entire genome was 21.6% A, 28.4% C, 28.0% G and 22.1% T, resulting in a G + C content of 56.3%. This is the highest G + C content among the genomes of four archaea of which the entire genomic sequence were reported. Through the processes of physical mapping and genomic sequencing, there was no evidence of extra-chromosomal units.

3.3. Assignment and similarity search of potential protein-coding regions

The assignment of potential protein-coding regions was done as in the previous paper.¹ The ORFs which consist

Table 1. Codon-usages of the predicted proteins coded for by the 2,618 APE-class ORFs.

UUU (Phe)	6.01	3823	UCU (Ser)	9.07	5767	UAU (Tyr)	12.65	8042	UGU (Cys)	3.20	2037
UUC (Phe)	21.35	13573	UCC (Ser)	15.55	9886	UAC (Tyr)	20.80	13225	UGC (Cys)	6.07	3859
UUA (Leu)	4.74	3015	UCA (Ser)	7.00	4448	UAA (Stop)	0.84	531	UGA (Stop)	0.90	573
UUG (Leu)	7.16	4554	UCG (Ser)	10.82	6880	UAG (Stop)	2.38	1514	UGG (Trp)	13.09	8323
CUU (Leu)	16.85	10712	CCU (Pro)	17.96	11420	CAU (His)	6.11	3883	CGU (Arg)	3.75	2385
CUC (Leu)	38.48	24467	CCC (Pro)	24.18	15373	CAC (His)	13.06	8306	CGC (Arg)	6.75	4291
CUA (Leu)	18.00	11443	CCA (Pro)	9.08	5776	CAA (Gln)	3.04	1934	CGA (Arg)	2.63	1671
CUG (Leu)	28.12	17879	CCG (Pro)	13.09	8323	CAG (Gln)	15.92	10123	CGG (Arg)	6.99	4446
AUU (Ile)	7.92	5036	ACU (Thr)	8.94	5682	AAU (Asn)	4.49	2852	AGU (Ser)	6.72	4276
AUC (Ile)	12.44	7911	ACC (Thr)	16.16	10278	AAC (Asn)	15.81	10052	AGC (Ser)	25.82	16418
AUA (Ile)	31.36	19942	ACA (Thr)	10.57	6720	AAA (Lys)	7.71	4900	AGA (Arg)	11.16	7099
AUG (Met)	19.62	12475	ACG (Thr)	11.03	7012	AAG (Lys)	27.40	17426	AGG (Arg)	45.48	28922
GUU (Val)	20.98	13343	GCU (Ala)	24.76	15743	GAU (Asp)	12.62	8022	GGU (Gly)	14.87	9457
GUC (Val)	22.62	14383	GCC (Ala)	37.01	23533	GAC (Asp)	26.00	16536	GGC (Gly)	32.66	20766
GUU (Val)	14.62	9296	GCA (Ala)	13.51	8592	GAA (Glu)	11.53	7331	GGA (Gly)	11.98	7616
GUG (Val)	28.93	18395	GCG (Ala)	19.67	12510	GAG (Glu)	54.34	34556	GGG (Gly)	25.66	16318

Numerals in the second column indicate the frequency of occurrence per thousand, and those in the third column the sum of codons occurred.

of longer than 100 sense codons starting with ATG or GTG were first searched, and are represented by a three-letter code (APE) plus a four-digit number indicating the ORF position. There was a total of 2,618 of this class of ORF including the overlapped ORFs in the same or either strand. In the next, shorter ORFs consisting of 50 to 99 codons were extracted from the regions where no ORF of longer than 100 codons was assigned, and the ORFs which possess significant similarity to either the registered sequences or protein motifs were taken as the probable protein-coding regions. Those were named with a four-letter code (APES) plus a two-digit number indicating the ORF position.

The ORFs thus assigned along the entire genomic sequence were totalled to 2,694, (APE-class ORFs = 2,618, APES-class ORFs = 76). An average size of the ORFs was 237 amino acid residues, and the longest one was consisted of 1,933 residues (APE0620). The assigned 2,694 ORFs occupy 88.83% of the genome.

For the anticipation of gene function, the products with a Zscore of ≥ 25 or with over 30% amino acid sequence identity along the entire coding region were taken into account for the APE-class ORFs, and the ZScore of ≥ 10 was used only for the APES-class ORFs. Among the ORFs assigned, the number of ORFs related to the genes with functions were 633 (23.5%), ORFs showed significant similarity to the registered sequences with unknown functions were 523 (19.4%) and those just containing some motifs were 202 (7.5%). The remaining 1,336 ORFs showed no significant similarity to the sequences in public databases.

The ORFs showing similarities to the registered genes with a known function and these containing protein-motifs are listed in the last two figures (Tables 5 and

6). The codon-usages of the hypothetical proteins coded by the total APE-class ORFs are summarized in Table 1. The codons with G or C at the third position appeared to be more frequently used, probably reflecting the relatively high G + C content of this strain.

3.4. RNA coding genes

The entire genomic sequence was subjected to a similarity search against the rRNA sequences registered in the databases, and a single 16S–23S rRNA cluster and two 5S rRNA genes were assigned, as shown in Fig. 1. Such an organization is similar to that of *A. fulgidus*⁴ and *P. horikoshii* OT3.¹ It has been reported that both the 16S and 23S rRNA genes of *A. pernix* K1 contain introns¹⁰ and that the intron in the 16S rRNA gene encodes a homing DNA endonuclease.¹⁰ We noted that an ORF similar to the DNA endonuclease gene is present in one intron of the 23S rRNA gene (APE1920 in Table 5).

By searching with tRNAscan,¹¹ 47 tRNA genes were identified, 14 of which were found to contain 18- to 121-bp introns. The species of tRNA genes identified are shown in Table 2, and those containing introns and the organization of the introns within each gene are summarized in Table 3. In all of the intron-containing tRNA genes except for the Thr- and Trp-tRNA genes, the introns were located one base from the anticodon region. Forty-one of them were discretely mapped, while the remaining six tRNA genes were mapped as clusters of two tRNA genes. As noted in other microorganisms,^{1,12} no tRNA genes containing A at the first position of the anticodon were present. As in all the archaea genomes sequenced, the Met-tRNA gene was present in triplicate. However, the sequences had no similarity each other, suggesting that the different Met-tRNA species may be used for the

Table 2. Summary of tRNA gene assignment.

UUU (Phe)		UCU (Ser)		UAU (Tyr)		UGU (Cys)	
UUC (Phe)	○	UCC (Ser)	○	UAC (Tyr)	○	UGC (Cys)	○
UUA (Leu)	○	UCA (Ser)	○	UAA (Stop)		UGA (Stop)	
UUG (Leu)	○	UCG (Ser)	○	UAG (Stop)		UGG (Trp)	○
CUU (Leu)		CCU (Pro)		CAU (His)		CGU (Arg)	
CUC (Leu)	○	CCC (Pro)	○	CAC (His)	○	CGC (Arg)	○
CUA (Leu)	○	CCA (Pro)	○	CAA (Gln)	○	CGA (Arg)	○
CUG (Leu)	○	CCG (Pro)	○	CAG (Gln)	○	CGG (Arg)	○
AUU (Ile)		ACU (Thr)		AAU (Asn)		AGU (Ser)	
AUC (Ile)	○	ACC (Thr)	○	AAC (Asn)	○	AGC (Ser)	○
AUA (Ile)		ACA (Thr)	○**	AAA (Lys)	○	AGA (Arg)	○
AUG (Met)	○*	ACG (Thr)	○	AAG (Lys)	○	AGG (Arg)	○
GUU (Val)		GCU (Ala)		GAU (Asp)		GGU (Gly)	
GUC (Val)	○	GCC (Ala)	○	GAC (Asp)	○	GGC (Gly)	○
GUA (Val)	○	GCA (Ala)	○	GAA (Glu)	○	GGA (Gly)	○
GUG (Val)	○	GCG (Ala)	○	GAG (Glu)	○	GGG (Gly)	○

The tRNA genes identified are indicated by circles and those containing introns by double-circles in the corresponding triple boxes.

*: Two of the three Met-tRNA genes possess the introns.

**: One of the two Thr-tRNA genes possesses the intron.

Table 3. The size and position of introns in the intron-containing tRNA genes.

tRNA species	anticodon	gene size (nt)	start	intron position* stop	intron size (nt)	mature tRNA (nt)
Arg-tRNA	TCT	122	40	83	44	78
Asp-tRNA	GTC	199	41	161	121	78
Cys-tRNA	GCA	96	39	56	18	78
Lys-tRNA	CTT	113	39	72	34	79
Lys-tRNA	TTT	113	39	72	34	79
Met-tRNA	CAT	115	39	76	38	77
Met-tRNA	CAT	125	39	86	48	77
Pro-tRNA	GGG	115	40	76	37	78
Pro-tRNA	CGG	122	40	83	44	78
Ser-tRNA	CGA	125	39	74	36	79
Thr-tRNA	CGT	124	37	85	49	75
Thr-tRNA	TGT	95	22	40	19	74
Tyr-tRNA	GTA	114	37	75	39	75
Trp-tRNA	CCA	114	32	68	37	77

*: Nucleotide positions from the 5' end of tRNA genes.

translation initiation of different classes of genes.

The genes for all the corresponding tRNA synthetases have been identified in the assigned ORFs, except for those of Gln-tRNA and Asn-tRNA. Instead, the three subunits of Glu-tRNA amidotransferases, which are necessary for Gln- and Asn-tRNA synthesis, were identified.

3.5. Other features noted

3.5.1. ORF containing intein

The presence of inteins, defined as the self-splicing portions of a polypeptide sequence,^{13,14} has been reported

in 36 archaea genes.^{1-3,15-19} To investigate the possible occurrence of ORFs containing inteins in the genome of *A. pernix* K1, all the ORF sequences were subjected to a similarity search against the intein-specific motif sequences described by Pietrokovski²⁰ and Perler,²¹ and also the amino acid sequences in the reported inteins using the Smith-Waterman algorithm. According to the results, one ORF (APE0745) was found to contain a putative intein element. The ORF length was 726 amino acids, of which the intein element occupied 468 amino acids. However, this ORF belonged to the ORF class coding for hypothetical proteins.

Table 4. Homologous ORFs identified by sequence comparison among the 2618 APE-class ORFs.

number of homologous ORFs*	number of groups	total ORFs
10	2	20
9	2	18
8	3	24
7	6	42
6	6	36
5	7	35
4	7	28
3	25	75
2	127	254
	total	532

*: ORFs with ZScore higher than 8 in SW search and amino acid identity higher than 30% in 70% of the entire region were taken.

3.5.2. Duplication of ORFs

By sequence comparison among the ORFs assigned to the genome, 532 ORFs could be grouped into 185, the ORFs in each group coding for proteins with considerable similarity (Table 4). The result can be interpreted that the ORFs in each group were generated by duplication of an ancestral sequence. Such homologous ORFs were present in the genome as either tandem repeats of a single ORF or the repeats of a single or a cluster of ORFs at different locations.

3.5.3. Unique repeating sequences

Three types of repeating units were found on the *A. pernix* K1 genome by similarity search, while no IS- or Tn-like elements were identified. The repeating unit composed of a well-conserved 24 bp sequence (GAATCTTC-GAGATAGAATTGCAAG) followed by a variable 37–52 bp sequence and 366- or 365-bp repeating units was classified as a type A repeating unit. This type of the repeating unit was present at three different positions, as indicated in Fig. 1 (positions 325,586–325,975, 716,882–718,997, and 1,277,298–1,278,859). Type B repeating unit was composed of 42 times-repeating well-conserved 23-bp sequence (ATATCCCTAAAGGGAATAGAAAG) followed by a variable 38–49 bp sequence. This was present at a single position (786,659–789,355). The type C repeating unit was merely dispersion of approximately 300-bp repeats along the entire genome (see Fig. 1). The biological significance of these repeats is not known.

3.5.4. Genes relating to respiration

Since *A. pernix* K1 is strictly aerobic, the genes involved in the TCA cycle were first searched, and all of the genes in this cycle except for alpha-ketoglutarate dehydrogenase were found. Instead of the genes for alpha-ketoglutarate dehydrogenase, the genes coding for the two subunits of 2-oxoacid:ferredoxin oxidoreductase,

which possesses the same function in archaea as alpha-ketoglutarate dehydrogenase,²² was identified on the genome. This suggests that *A. pernix* K1 had already acquired a respiratory system similar to that seen in mitochondria of higher eukaryotes.

As other genes relating to the respiratory chain, those for NADH dehydrogenase, NADH-quinone oxidoreductase, succinate dehydrogenase, iron-sulfur protein, flavoprotein, cytochrome *b* and cytochrome *c* oxidase were identified, and one ORF was found to have weak homology with the Rieske protein. However, none of the genes coding for quinol-cytochrome *c* oxidoreductase, cytochrome *c* reductase, cytochrome *c*₁, cytochrome *bc*₁, cytochrome *b*₆*f*, cytochrome *c* and quinol oxidase, were identified by similarity search.

Like all other aerobic organisms the gene for superoxide dismutase was also present in this organism.

Analysis of gene organization as well as gene structure of the *A. pernix* K1 genome in comparison with those of the *P. horikoshii* OT3 genome is under investigation to find any clue for understanding the thermostability of the gene products.

Acknowledgments: We thank Dr. S. Tabata and H. Kotani and their staffs at Kazusa DNA Research Institute for their valuable suggestions and help. Thanks are also due to Dr. D. Söll for his kind help with the assignment of tRNA genes and to Dr. P. Green for donation of the assembling software PhredPhrap. This work was supported by the Ministry of International Trade and Industry.

References

- Kawarabayasi, Y., Sawada, M., Horikawa, H. et al. 1998, Complete Sequence and Gene Organization of the Genome of a Hyper-thermophilic Archaeabacterium, *Pyrococcus horikoshii* OT3, *DNA Res.*, **5**, 55–76 & 147–155.
- Bult, C. J., White, O., Olsen, G. J. et al. 1996, Complete Genome Sequence of the Methanogenic Archaon, *Methanococcus jannaschii*, *Science*, **273**, 1058–1073.
- Smith, D. R., Doucette-Stamm, L. A., Deloughery, C. et al. 1997, Complete Genome Sequence of *Methanobacterium thermoautotrophicum* ΔH: Functional Analysis and Comparative Genomics, *J. Bacteriol.*, **179**, 7135–7155.
- Klenk, H. P., Clayton, R. A., Tomb, J. F. et al. 1997, The complete genome sequence of the hyperthermophilic, sulphate-reducing archaeon *Archaeoglobus fulgidus*, *Nature*, **390**, 364–370.
- Sako, Y., Nomura, N., Uchida, A. et al. 1996, *Aeropyrum pernix* gen. nov., sp. nov., a Novel Aerobic Hyper-thermophilic Archæon Growing at Temperatures up to 100°C, *International Journal of Systematic Bacteriology*, **46**, 1070–1077.
- Ewing, B., Hillier, L., Wendl, M. C., and Green, P. 1998, Base-calling of Automated Sequencer Traces Using phred. I. Accuracy Assessment, *Genome Res.*, **8**, 175–185.

7. Ewing, B. and Green, P. 1998, Base-calling of Automated Sequencer Traces Using phred. II. Error Probabilities, *Genome Res.*, **8**, 186–194.
8. Smith, C. L., Economou, J. G., Schutt, A., Klco, A., and Cantor, C. R. 1987, A Physical Map of the *Escherichia coli* K12 Genome, *Science*, **236**, 1448–1453.
9. Smith, T. F. and Waterman, M. S. 1981, Identification of Common Molecular Subsequences, *J. Mol. Biol.*, **147**, 195–197.
10. Nomura, N., Sako, Y., and Uchida, A. 1998, Molecular Characterization and Postslicing Fate of Three Introns within the Single rRNA Operon of the Hyperthermophilic Archaeon *Aeropyrum pernix* K1, *J. Bacteriol.*, **180**, 3635–3643.
11. Lowe, T. M. and Eddy, S. R. 1997, tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence, *Nuc. Acids Res.*, **25**, 955–964.
12. Nakamura, Y. and Tabata, S. 1997, Codon-Anticodon Assignment and Detection of Codon Usage Trends in Seven Microbial Genomes, *Microbial & Comparative Genomics*, **2**, 299–312.
13. Hirata, R., Ohsumi, Y., Nakano, A., Kawasaki, H., Suzuki, K., and Anraku, Y. 1990, Molecular Structure of a Gene, VMA1, Encoding the Catalytic Subunit of H⁺-Translocating Adenosine Triphosphatase from Vacuolar Membranes of *Saccharomyces cerevisiae*, *J. Biol. Chem.*, **265**, 6726–6733.
14. Kane, P. M., Yamashiro, C. T., Wolczyk, D. F., Neff, N., Goebel, M., and Stevens, T. H. 1990, Protein Splicing Converts the Yeast TFP1 Gene Product to the 69-kD Subunit of the Vacuolar H⁺-Adenosine Triphosphatase, *Science*, **250**, 651–657.
15. Riera, J., Robb, F. T., Weiss, R., and Fontecave, M. 1990, Ribonucleotide reductase in the archaeon *Pyrococcus furiosus*: a critical enzyme in the evolution of DNA genomes?, *Proc. Natl. Acad. Sci. USA*, **94**, 475–478.
16. Xu, M., Southworth, M. W., Mersha, F. B., Hornstra, L. J., and Peler, F. B. 1993, In vitro protein splicing of purified precursor and the identification of a branched intermediate, *Cell.*, **75**, 1371–1377.
17. Takagi, M., Nishioka, M., Kakihara, H. et al. 1997, Characterization of DNA polymerase from *Pyrococcus* sp. strain KOD1 and its application to PCR, *Appl. Environ. Microbiol.*, **63**, 4504–4510.
18. Peler, F. B., Comb, D. G., Jack, W. E. et al. 1992, Intervening sequences in an Archaea DNA polymerase gene, *Proc. Natl. Acad. Sci. USA*, **89**, 5577–5581.
19. Niehaus, F., Frey, B., and Antranikian, G. 1997, Cloning and characterisation of a thermostable alpha-DNA polymerase from the hyperthermophilic archaeon *Thermococcus* sp. TY, *Gene*, **204**, 153–158.
20. Pietrokovski, S. 1994, Conserved sequence features of inteins (protein introns) and their use in identifying new inteins and related proteins, *Protein Science*, **3**, 2340–2350.
21. Perler, F. B., Olsen, G. J., and Adams, E. 1997, Compilation and analysis of intein sequences, *Nuc. Acids Res.*, **25**, 1087–1093.
22. Zhang, Q., Iwasaki, T., Wakagi, T., and Oshima, T. 1996, 2-Oxoacid:Ferredoxin Oxidoreductase from the Thermoacidophilic Archaeon, *Sulfolobus* sp. Strain 7, *J. Biochem.*, **120**, 587–599

Table 5. List of ORFs showing similarities to registered genes with known function.

ORF ID	Position Initiation	Position Termination	Direction	Length (aa)	Predicted Product	Locus	Zscore	% Identity
ORFs with known function								
APE0009	6,091	5,774	<	105	transport protein	OWL:PSMERTRP2	23.96	45.46
APE0012	8,908	9,840	>	310	sugar kinase	PIR:A69073	53.30	32.27
APE0013	9,925	11,652	>	575	dihydroxy-acid dehydratase	PIR:S76895	126.31	49.09
APE0017	13,928	12,387	<	513	xylulose kinase	PIR:C69415	158.76	42.55
APE0026	16,932	18,800	>	622	phosphoenolpyruvate synthase	PIR:G69015	65.16	40.53
APE0028	18,728	19,384	>	218	phosphoenolpyruvate synthase	PIR:A70043	42.28	59.00
APE0031	19,387	20,850	>	487	antibiotic resistance protein	PIR:A70022	34.30	24.27
APE0038	24,026	24,823	>	265	multiple sugar-binding transport system permease protein	OWL:AP000004171	117.76	61.50
APE0040	24,790	25,698	>	302	multiple sugar-binding transport system permease protein	OWL:AP000004169	132.29	58.80
APE0043	25,667	26,812	>	381	multiple sugar-binding transport ATP-binding protein	OWL:AB00947118	56.01	51.50
APE0045	26,994	28,115	>	373	phosphate-binding periplasmic protein	PIR:F64426	89.04	37.99
APE0047	28,329	29,060	>	243	phosphate transport system permease protein pstC	Swiss_Prot:Q58420	97.33	47.11
APE0048	29,053	29,907	>	284	phosphate transport system permease protein pstA	Swiss_Prot:Q58419	90.53	41.52
APE0050	29,907	30,668	>	253	phosphate transport ATP-binding protein pstB	Swiss_Prot:Q58418	53.44	50.20
APE0056	35,796	35,035	<	253	enoyl-CoA hydratase	PIR:A69390	49.22	39.44
APE0063	42,608	44,086	>	492	ABC transporter ATP-binding protein	OWL:AB00952437	59.00	30.74
APE0069	46,617	47,108	>	163	deoxycytidine triphosphate deaminase	PIR:C69388	32.74	41.05
APE0072	49,294	50,331	>	345	heme exporter protein C	PIR:C69497	31.28	33.33
APE0077	53,276	53,791	>	171	molybdopterin biosynthesis mog protein	PIR:H69658	35.08	37.41
APE0084	58,602	57,823	<	259	NADH dehydrogenase	Swiss_Prot:P94424	43.79	37.02
APE0086	59,910	60,401	>	163	bacterioferritin comigratory protein	OWL:MTV0464	44.91	39.07
APE0092	66,333	65,371	<	320	acetylpolyamine aminohydrolase	OWL:AP000005176	77.93	44.84
APE0098	68,571	68,972	>	133	eukaryotic Ig lambda-1 chain V region	OWL:MUSIGLAS	13.41	35.71
APE0099	72,118	69,239	<	959	DNA-directed DNA polymerase (pfu polymerase)	PIR:A56277	181.58	51.14
APE0105	76,470	77,261	>	263	GTP-binding protein	OWL:AP00000342	59.38	34.73
APE0115	86,209	85,004	<	401	flap endonuclease-1	OWL:AP000006120	138.05	55.06
APE0117	88,755	87,535	<	406	threonyl-tRNA synthetase	Swiss_Prot:O29703	31.22	45.22
APE0119	88,998	89,957	>	319	radA protein	Swiss_Prot:Q55075	118.93	70.76
APE0122	90,706	91,260	>	184	hit-like protein	OWL:AB00946623	97.32	55.06
APE0134	99,329	100,084	>	251	electron transfer flavoprotein beta subunit	Swiss_Prot:P97089	40.31	35.79
APE0135	100,104	101,030	>	308	electron transfer flavoprotein alpha subunit	Swiss_Prot:P77378	63.14	35.77
APE0137	102,380	103,879	>	499	gamma-glutamyltranspeptidase	PIR:F70068	61.23	30.44
APE0142	106,321	106,863	>	180	diphtheria toxin repressor	PIR:D69225	22.12	34.34
APE0146	109,644	108,844	<	266	pyrroline-5-carboxylate reductase	Swiss_Prot:P74572	62.96	40.28
APE0150	110,740	111,534	>	264	DNA-(apurinic or apyrimidinic site)lyase	PIR:S75373	55.24	34.70
APE0159	120,096	119,248	<	282	bacteriochlorophyll synthase	PIR:D69300	44.04	32.19
APE0161	120,170	121,858	>	562	lysyl-tRNA Synthetase	Swiss_Prot:Q57959	71.05	29.39
APE0162	122,618	121,827	<	263	Proliferating cell nuclear antigen	OWL:AP00000362	64.24	35.83
APE0166	124,978	124,412	<	188	1,3-propanediol dehydrogenase	PIR:D70046	33.44	38.82
APE0168	125,988	125,632	<	118	Class II heat shock protein	PIR:S75382	39.64	50.00
APE0169	127,303	126,080	<	407	aminotransferase	PIR:S75385	132.22	44.64
APE0171	128,428	127,388	<	346	glyceraldehyde 3-phosphate dehydrogenase	PIR:S63529	116.67	52.07
APE0173	129,664	128,432	<	410	phosphoglycerate kinase	Swiss_Prot:P50317	100.92	45.46
APE0177	131,800	130,640	<	386	molybdopterin biosynthesis protein	OWL:MKFWUGDB4	42.40	32.06
APE0178	132,218	131,790	<	142	molybdopterin-guanine dinucleotide biosynthesis protein	Swiss_Prot:P44902	22.37	34.65
APE0181	134,653	133,391	<	420	cleavage and polyadenylation specificity factor subunit	PIR:D69316	75.61	33.84
APE0188	137,967	140,066	>	699	DNA replication licensing factor mcm	PIR:D69103	81.95	42.42
APE0191	140,209	142,317	>	702	helicase	OWL:AP000005189	132.28	38.41
APE0209	153,634	152,483	<	383	S2P metalloprotease	OWL:AB00947313	39.00	45.95
APE0212	154,578	155,789	>	403	processing protease	Swiss_Prot:Q04805	37.11	26.33
APE0218	158,482	157,766	<	238	50S ribosomal protein L2	OWL:AB0095255	87.77	58.75
APE0222	162,588	161,632	<	318	carboxyphosphoenolpyruvate phosphomutase	Swiss_Prot:P54528	117.70	49.46
APE0227	164,048	165,082	>	344	50S ribosomal protein L3	Swiss_Prot:P54014	108.05	48.67
APE0228	165,110	165,931	>	273	50S ribosomal protein L4	Swiss_Prot:P54015	85.73	50.60
APE0230	167,231	166,380	<	283	nucleotide-binding protein	Swiss_Prot:Q57731	92.31	50.19
APE0234	170,182	169,346	<	278	acetoin utilization acub protein	PIR:F69030	44.84	34.32
APE0236	172,094	171,336	<	252	uroporphyrinogen III C-methyltransferase	Swiss_Prot:P29928	61.72	41.10
APE0246	177,978	176,965	<	337	ethylene-responsive protein 1	Swiss_Prot:Q26762	108.21	60.75
APE0255	183,088	182,321	<	255	mannose-1-phosphate guanyltransferase	OWL:AB0095234	33.92	34.47
APE0257	184,505	183,945	<	186	DNA-directed RNA polymerase subunit E'	Swiss_Prot:Q57840	61.55	40.26
APE0260	186,772	185,843	<	309	dihydroorotate dehydrogenase	OWL:AP000006223	85.47	42.95
APE0261	188,024	186,735	<	429	dihydroorotate	OWL:SSPYRCDE	50.17	33.99
APE0263	190,121	188,799	<	440	alkaline protease	PIR:A57690	51.40	44.48
APE0266	192,941	195,373	>	810	aldehyde ferredoxin oxidoreductase	PIR:A69535	84.00	32.47
APE0270	197,378	196,677	<	233	transport system permease protein	PIR:F70308	46.40	29.76
APE0272	198,262	197,456	<	268	molybdate-binding protein	PIR:G69223	30.60	29.87
APE0275	201,164	200,817	<	115	DNA-3-methyladenine glycosidase I	Swiss_Prot:Q92383	22.94	29.17
APE0288	210,667	211,257	>	196	transcriptional regulator	OWL:AB01076	18.31	28.46
APE0291	212,566	212,234	<	110	bacterioferritin comigratory protein	Swiss_Prot:P23480	39.24	40.00
APE0293	213,313	213,008	<	101	transcription regulator	PIR:D69058	42.08	42.70
APE0300	217,618	216,356	<	420	oligopeptide transport ATP-binding protein oppF	OWL:AP000006118	51.72	54.04
APE0301	218,470	217,622	<	282	oligopeptide transport ATP-binding protein oppD	OWL:AP000006117	66.52	50.88
APE0302	219,513	218,587	<	308	oligopeptide transport system permease protein oppC	OWL:BFU645142	68.21	41.52
APE0303	220,534	219,506	<	342	oligopeptide transport system permease protein oppB	PIR:C64557	80.35	37.35
APE0304	222,959	220,551	<	802	oligopeptide-binding protein oppA	OWL:HEAHBPA	43.49	28.76
APE0306	223,382	224,818	>	478	Glycerol kinase	PIR:B69358	33.27	28.00
APE0308	224,872	226,164	>	430	alkyldihydroxyacetonephosphate synthase	PIR:D69358	65.71	33.75
APE0311	227,383	228,213	>	276	citrate lyase beta chain	OWL:MTCY07A73	52.03	36.55

Table 5. Continued.

ORF ID	Position Initiation	Position Termination	Direction	Length (aa)	Predicted Product	Locus	Zscore	% Identity
APE0314	228,849	229,709	>	286	dipeptide transport protein	OWL:AP000006293	83.37	38.43
APE0316	230,366	229,719	<	215	agmatinase	Swiss_Prot:Q57757	52.21	37.57
APE0317	232,020	230,611	<	469	MRSA protein	PIR:B69307	95.80	40.71
APE0319	232,765	232,223	<	180	proteinase I	Swiss_Prot:Q51732	105.81	61.21
APE0320	233,260	232,913	<	115	ferredoxin	Swiss_Prot:P00211	26.12	41.51
APE0326	237,935	236,514	<	473	TATA-binding protein-interracing protein 49	OWL:AB00952622	146.13	54.86
APE0333	239,958	240,503	>	181	deoxycytidine triphosphate deaminase	Swiss_Prot:Q02103	97.51	53.80
APE0336	243,168	240,826	<	780	penicillin acylase	PIR:G69310	64.39	43.40
APE0337	244,424	243,483	<	313	microsomal dipeptidase	OWL:AP000004158	45.92	31.82
APE0338	244,452	245,948	>	498	beta-lactamase	OWL:AP000001145	93.70	44.21
APE0342	250,548	249,451	<	365	acyl-CoA hydrolase	Swiss_Prot:P49851	25.45	25.69
APE0343	251,092	250,616	<	158	50S ribosomal protein L15	Swiss_Prot:P54047	50.77	45.52
APE0345	251,594	251,106	<	162	50S ribosomal protein L30	OWL:AB00952524	66.97	47.80
APE0346	252,287	251,631	<	218	50S ribosomal protein S5	OWL:SSSPCOPER9	106.49	62.56
APE0347	252,960	252,316	<	214	50S ribosomal protein L18	OWL:AB00952522	99.64	55.33
APE0348	253,448	252,981	<	155	50S ribosomal protein L19	OWL:AB00952521	67.84	47.95
APE0349	253,906	253,469	<	145	50S ribosomal protein L32	OWL:SSSPCOPER6	56.87	49.07
APE0350	254,454	253,906	<	182	50S ribosomal protein L6	OWL:SSSPCOPERS	90.39	51.96
APE0352	254,982	254,575	<	135	30S ribosomal protein S8	Swiss_Prot:P05636	78.44	54.14
APE0354	255,762	255,178	<	194	50S ribosomal protein L5	OWL:AB00952515	88.26	58.38
APE0356	256,526	255,753	<	257	30S ribosomal protein S4	OWL:AB00952514	98.14	52.42
APE0358	256,956	256,558	<	132	50S ribosomal protein L24	OWL:AB00952513	91.92	74.14
APE0359	257,398	256,976	<	140	50S ribosomal protein L14	Swiss_Prot:P54037	70.45	62.81
APE0360	257,773	257,411	<	120	30S ribosomal protein S17	OWL:AB00952510	56.84	49.00
APE0363	259,013	258,273	<	246	30S ribosomal protein S3	Swiss_Prot:P54034	76.49	46.08
APE0365	259,488	259,018	<	156	50S ribosomal protein L22	PIR:G69489	67.28	48.03
APE0367	260,006	259,515	<	163	30S ribosomal protein S19	OWL:AB00952526	78.05	67.41
APE0374	266,894	263,700	<	1,064	isoleucyl-tRNA synthetase	Swiss_Prot:O27428	145.88	47.15
APE0385	272,063	270,465	<	532	acyl-CoA dehydrogenase, short-chain specific	OWL:MTV026	50.43	38.44
APE0389	272,759	272,277	<	160	acetyltransferase	Swiss_Prot:P029729	20.54	28.57
APE0392	274,609	273,269	<	446	malate oxidoreductase	PIR:F69465	155.72	55.45
APE0396	276,343	274,664	<	559	medium-chain acyl-CoA synthetase	PIR:D69274	126.49	45.04
APE0402	278,630	278,016	<	204	membrane-associated ATPase gamma chain	Swiss_Prot:P22721	70.90	39.90
APE0404	280,111	278,711	<	466	membrane-associated ATPase beta chain	OWL:AB00946425	81.69	71.27
APE0405	281,940	280,144	<	598	membrane-associated ATPase alpha chain	Swiss_Prot:P09639	94.81	63.85
APE0412	283,068	284,156	>	362	histidine decarboxylase	PIR:E69015	39.66	28.85
APE0426	299,481	300,602	>	373	dolichol-phosphate-mannosyltransferase	OWL:AB00946629	79.85	40.67
APE0427	300,649	301,296	>	215	DNA polymerase homologous protein	OWL:AP000006178	83.93	51.46
APE0436	305,941	305,192	<	249	eukaryotic translation initiation factor eIF-2 alpha chain	Swiss_Prot:Q57581	72.95	37.10
APE0441	308,859	308,158	<	233	proliferating cell nuclear antigen	OWL:AP00000362	63.17	30.77
APE0442	309,515	309,201	<	104	DNA-directed RNA polymerase subunit M	OWL:AP00000361	47.62	42.86
APE0443	310,223	309,915	<	102	DNA-directed RNA polymerase subunit L	Swiss_Prot:P46217	20.43	30.00
APE0449	313,902	313,378	<	174	50S ribosomal protein L10	Swiss_Prot:Q57963	81.98	55.49
APE0450	313,995	315,728	>	577	molybdopterin biosynthesis moeA protein	Swiss_Prot:P12281	40.27	31.65
APE0457	317,638	318,996	>	452	4-aminobutyrate aminotransferase	OWL:AP000006128	96.66	44.72
APE0465	324,526	323,288	<	412	thiazole biosynthesis gene thi I	Swiss_Prot:P77718	70.40	32.95
APE0471	328,152	327,082	<	356	transport ATP-binding protein	OWL:MTCI36427	47.35	34.71
APE0475	331,501	332,634	>	377	DNA polymerase II	OWL:D84670	130.21	50.00
APE0478	334,672	333,605	<	355	molybdopterin cofactor biosynthesis protein A	OWL:AB00946813	71.41	37.76
APE0487	339,842	338,421	<	473	D-lactate dehydrogenase	PIR:H69350	162.66	50.86
APE0489	340,772	342,148	>	458	pyruvate kinase	Swiss_Prot:P94939	60.33	35.83
APE0496	344,881	345,564	>	227	ribonuclease HII	OWL:AB00952244	59.96	36.32
APE0498	345,570	346,643	>	357	carbamoylphosphate synthetase	OWL:AB016521	25.73	32.90
APE0506	350,041	349,385	<	218	endonuclease V	Swiss_Prot:P32679	46.40	39.33
APE0507	350,868	350,149	<	239	proteasome, beta subunit	PIR:A69310	71.41	41.76
APE0511	353,676	352,615	<	353	RNA 3'-terminal phosphate cyclase	PIR:G69122	93.79	47.19
APE0514	354,317	355,375	>	352	sun protein	OWL:AP00000679	80.22	35.29
APE0517	355,875	356,681	>	268	fkbp-type peptidyl-prolyl cis-trans isomerase	PIR:D69498	50.40	34.51
APE0519	357,799	356,687	<	370	glycerol dehydrogenase	OWL:AP000006181	128.89	49.28
APE0521	357,894	358,571	>	225	proteasome, beta subunit	OWL:AP000006107	64.31	45.70
APE0522	358,793	360,823	>	676	cleavage and polyadenylation factor subunit	PIR:B69310	162.15	45.67
APE0526	362,154	363,203	>	349	Xaa-Pro dipeptidase	OWL:AP00000555	71.02	42.42
APE0528	364,431	363,298	<	377	acetaminidase	PIR:S74213	84.99	37.14
APE0532	365,457	366,251	>	264	ABC transporter ATP-binding protein	PIR:E69497	38.58	34.65
APE0545	374,615	373,908	<	235	uracil phosphoribosyltransferase	Swiss_Prot:P39149	52.68	35.64
APE0553	377,718	378,551	>	277	dimethyladenosine transferase	PIR:C69043	40.48	38.29
APE0564	386,001	384,859	<	380	chorismate synthase	PIR:F69333	112.59	52.54
APE0569	387,274	385,991	<	427	3-phosphoshikimate 1-carboxyvinyltransferase	PIR:D69202	80.56	39.00
APE0574	388,781	388,104	<	225	shikimate dehydrogenase	PIR:C69130	38.95	34.15
APE0575	388,186	388,533	>	115	surface protein-I	OWL:AF000412	9.39	27.16
APE0579	390,673	389,597	<	358	3-dehydroquinate synthase	Swiss_Prot:P36919	63.15	35.33
APE0581	391,485	390,655	<	276	phospho-2-dehydro-3-deoxyheptonate aldolase	Swiss_Prot:P39912	117.66	54.09
APE0583	392,455	391,487	<	322	transketolase	Swiss_Prot:Q58092	72.42	38.69
APE0586	393,598	392,495	<	367	transketolase	Swiss_Prot:Q58094	43.09	40.00
APE0591	397,515	396,676	<	279	3-hydroxybutyryl-CoA dehydratase	PIR:C69370	37.70	36.62
APE0592	397,109	397,465	>	118	arabinogalactan-protein	OWL:NAU13066	12.19	36.67
APE0603	404,690	405,439	>	249	molybdopterin (mpt) converting factor, subunit 2	OWL:AB00947033	30.42	31.85
APE0607	411,387	407,392	<	1,331	surface layer-associated STABLE protease	OWL:SMU57968	93.49	39.62
APE0624	427,232	428,431	>	399	adenosylhomocysteinase	Swiss_Prot:P50252	113.59	63.66

Table 5. Continued.

ORF ID	Position Initiation	Position Termination	Direction	Length (aa)	Predicted Product	Locus	Zscore	% Identity
APE0650	438,846	436,309	<	845	phosphoenolpyruvate synthase	Swiss_Prot:P46893	230.03	62.41
APE0662	442,453	443,769	>	438	histidyl-tRNA synthetase	OWL:AB00947347	75.65	36.05
APE0665	444,837	444,124	<	237	ribose 5-phosphate isomerase	PIR:G69180	52.49	41.71
APE0669	447,540	446,854	<	228	transcriptional activator TENA	OWL:AP00000567	90.78	42.52
APE0672	449,885	448,944	<	313	malate dehydrogenase	Swiss_Prot:P49814	69.89	44.08
APE0674	452,071	453,288	>	405	aspartate aminotransferase	PIR:JC4537	65.28	34.85
APE0675	453,671	453,318	<	117	methionyl-tRNA synthetase	PIR:F70338	61.44	59.78
APE0676	454,639	453,815	<	274	3-methyl-1-oxobutanoate hydroxymethyl transferase	OWL:AP00000483	121.08	51.55
APE0680	455,882	456,316	>	144	ORF6	PIR:H70429	10.43	25.24
APE0683	457,930	458,598	>	222	phosphoserine phosphatase	OWL:AB00952815	24.95	27.89
APE0686	459,143	460,255	>	370	translation initiation factor eIF-2B alpha subunit	Swiss_Prot:Q57896	96.92	45.75
APE0689	461,506	462,813	>	435	isocitrate dehydrogenase	OWL:D86855	136.42	65.96
APE0695	465,770	464,346	<	474	Vir B11	PIR:B69292	69.46	39.67
APE0698	466,369	467,247	>	292	deoxyhypusine synthase	OWL:AP000006102	64.33	34.45
APE0703	469,708	468,551	<	385	type II DNA topoisomerase VI subunit a	Swiss_Prot:O05208	219.46	69.35
APE0706	471,436	469,739	<	565	type II DNA topoisomerase VI subunit b	Swiss_Prot:O05207	188.41	50.76
APE0708	471,778	474,039	>	753	aldehyde oxidoreductase	PIR:S73093	146.82	39.56
APE0712	476,314	475,118	<	398	N-Acylamino acid racemase	PIR:F69991	131.66	44.11
APE0715	476,453	478,012	>	519	TRK system potassium uptake protein	OWL:AB00952631	69.88	31.12
APE0739	492,098	493,135	>	345	acyl-CoA dehydrogenase	Swiss_Prot:P45857	31.92	37.95
APE0741	493,787	493,143	<	214	superoxide dismutase	Swiss_Prot:Q08713	56.35	51.69
APE0754	499,144	499,716	>	190	mitochondrial-type HSP70	PIR:B69001	66.31	41.14
APE0756	499,786	500,649	>	287	3-hydroxyisobutyrate dehydrogenase	PIR:C70303	34.50	25.89
APE0767	505,469	506,389	>	306	spermidine synthase	PIR:F70305	61.64	40.63
APE0772	508,748	509,086	>	112	signal recognition particle	PIR:E69087	28.38	42.42
APE0775	511,570	510,839	<	243	glutaredoxin-like protein	PIR:S54843	49.09	35.87
APE0777	513,100	512,300	<	266	serine/threonine protein phosphatase pp2a catalytic subunit	PIR:E69477	26.41	31.94
APE0782	515,085	516,314	>	409	N2,N2-dimethylguanosine tRNA methyltransferase	PIR:G69025	64.92	35.97
APE0789	521,170	521,580	>	136	30S ribosomal protein S8	OWL:AP0000058	60.28	51.91
APE0792	524,929	525,633	>	234	cytochrome C oxidase polypeptide II	PIR:A23711	24.20	39.83
APE0793	525,637	528,093	>	818	cytochrome C oxidase polypeptide I	OWL:MTV01256	32.29	45.91
APE0801	531,735	532,616	>	293	survival protein surE	PIR:F69367	31.74	38.28
APE0805	534,953	533,865	<	362	integrase/recombinase xerD	Swiss_Prot:P21891	41.34	39.29
APE0807	535,109	536,770	>	553	delta-1-pyrroline-5-carboxylate dehydrogenase	OWL:MTV00521	130.15	47.60
APE0809	536,898	538,352	>	484	threonyl-tRNA synthetase	PIR:F70444	65.43	42.58
APE0829	550,795	550,433	<	120	GAGE-2 protein	Swiss_Prot:Q13066	7.78	25.00
APE0872	571,101	572,450	>	449	cytosine-specific methylase	Swiss_Prot:P34905	39.49	59.34
APE0875	573,400	574,260	>	286	A/G-specific adenine glycosylase	Swiss_Prot:P29588	56.95	34.36
APE0903	593,926	594,750	>	274	alanyl-tRNA synthetase	Swiss_Prot:Q57984	40.49	35.95
APE0907	595,377	597,050	>	557	thermosome subunit	OWL:PO6549	94.49	70.38
APE0912	597,205	597,987	>	260	dehydrogenase	OWL:AP000007226	83.06	47.84
APE0914	598,062	598,373	>	103	DNA-directed RNA polymerase subunit	Swiss_Prot:P36595	31.93	48.61
APE0916	600,648	599,458	<	396	mu-crystallin	OWL:AF039391	35.29	31.92
APE0917	600,515	602,017	>	500	branched-chain amino acid binding protein	PIR:F69423	71.59	32.50
APE0919	602,220	602,984	>	254	high-affinity branched-chain amino acid transport	PIR:E69423	70.28	52.82
APE0921	602,987	603,700	>	237	high-affinity branched-chain amino acid transport	PIR:D69423	57.49	45.30
APE0922	603,714	604,778	>	354	high-affinity branched-chain amino acid transport	PIR:D64458	53.78	39.51
APE0924	604,782	605,954	>	390	high-affinity branched-chain amino acid transport	PIR:E64458	68.81	33.52
APE0929	606,778	607,980	>	400	nonspecific lipid-transfer protein	PIR:C69285	161.29	58.72
APE0931	609,422	608,538	<	294	diphthine synthase	PIR:A64459	33.54	35.93
APE0935	612,995	611,982	<	337	glutamate synthase small chain	PIR:A70477	63.02	38.37
APE0941	615,984	615,181	<	267	spermidine/putrescine transport system permease protein	PIR:E69450	60.65	40.00
APE0943	616,758	615,985	<	257	spermidine/putrescine transport system permease protein	PIR:F69450	54.47	32.38
APE0944	617,719	616,742	<	325	spermidine/putrescine transport ATP-binding protein	OWL:AP00000655	43.44	55.56
APE0945	619,088	617,865	<	407	spermidine/putrescine-binding periplasmic protein	Swiss_Prot:P23861	56.10	31.40
APE0946	620,204	619,287	<	305	fumarate reductase iron-sulfur protein	PIR:B69335	80.34	51.97
APE0950	622,861	621,140	<	573	fumarate reductase flavoprotein subunit	PIR:A69335	122.06	57.14
APE0951	623,425	622,982	<	147	glycine cleavage system H protein	PIR:E70395	60.97	50.00
APE0952	623,518	624,243	>	241	hexulose-6-phosphate synthase	OWL:AB00952937	39.57	32.38
APE0955	626,217	625,330	<	295	potassium channel	PIR:S09045	29.98	49.06
APE0959	627,918	627,391	<	175	methylated-DNA--protein-cysteine methyltransferase	PIR:B69182	33.65	49.35
APE0960	629,115	627,961	<	384	p60 katanin	Swiss_Prot:Q09803	45.23	46.22
APE0976	642,023	641,706	<	105	50S ribosomal protein L14	Swiss_Prot:P54054	51.11	62.82
APE0979	644,226	642,889	<	445	amidase	OWL:AE00126911	22.39	44.19
APE0981	645,352	644,738	<	204	adenylate kinase	Swiss_Prot:P35028	77.40	44.10
APE0983	646,833	645,349	<	494	preprotein translocate secY subunit	Swiss_Prot:P49978	117.72	42.31
APE0989	649,439	649,762	>	107	nucleolar protein	OWL:AB00952250	39.76	63.16
APE0990	649,660	650,541	>	293	nucleolar protein	PIR:A69141	98.92	47.89
APE0993	650,508	651,299	>	263	purine nucleoside phosphorylase	Swiss_Prot:P50389	72.81	44.96
APE1011	659,078	659,848	>	256	protein-L-isospartate O-methyltransferase	PIR:B69540	86.17	55.35
APE1015	664,422	661,543	<	959	leucyl-tRNA synthetase	OWL:AP00000497	178.61	43.46
APE1023	664,590	665,474	>	294	N-acetylneuraminate lyase	Swiss_Prot:P39359	39.07	30.47
APE1026	665,554	666,162	>	202	heme exporter protein A	OWL:SC1B56	25.73	39.49
APE1038	669,592	669,963	>	123	cytidine deaminase	Swiss_Prot:P19079	60.65	54.55
APE1044	673,174	673,923	>	249	reductase	Swiss_Prot:P39605	56.59	35.11

Table 5. Continued.

ORF ID	Position Initiation	Position Termination	Direction	Length (aa)	Predicted Product	Locus	Zscore	% Identity
APE1045	675,039	673,939	<	366	heat shock protein htpX	PIR:C69279	59.76	35.02
APE1047	675,333	676,019	>	228	translation initiation factor eIF-2B	OWL:AP000002134	23.15	29.17
APE1053	677,563	678,489	>	308	iron (III) dicitrate transport system permease protein	PIR:S74438	53.59	36.30
APE1055	678,473	679,225	>	250	iron (III) dicitrate transport ATP-binding protein	PIR:H69424	32.00	36.82
APE1061	682,347	681,316	<	343	thioredoxin reductase	OWL:AP000006131	75.05	43.71
APE1065	683,529	684,746	>	405	succinyl-CoA synthetase beta chain	Swiss_Prot:P25126	101.53	44.59
APE1072	684,753	685,646	>	297	succinyl-CoA synthetase alpha chain	PIR:A69523	98.05	62.76
APE1081	689,806	688,868	<	312	signal recognition particle protein	Swiss_Prot:P27414	88.31	51.56
APE1085	691,314	690,613	<	233	eukaryotic translation initiation factor 6	OWL:AP000002225	76.53	44.69
APE1087	691,724	691,407	<	105	50S ribosomal protein L31	Swiss_Prot:P54009	33.83	55.10
APE1090	692,832	692,371	<	153	30S ribosomal protein S19	OWL:AP00000630	86.19	58.55
APE1094	695,386	693,527	<	619	DNA ligase	Swiss_Prot:Q02093	202.86	58.79
APE1098	695,697	696,863	>	388	fmn protein	OWL:AB0094648	94.21	41.55
APE1114	702,053	701,208	<	281	methionine aminopeptidase	OWL:AP00000323	84.98	43.02
APE1119	704,785	702,971	<	604	RNase L inhibitor	OWL:AP000003259	127.34	51.08
APE1129	706,474	708,192	>	572	methionyl-tRNA synthetase	OWL:AP000004125	122.87	42.22
APE1132	708,253	708,612	>	119	30S ribosomal protein S24	PIR:A69389	30.99	38.54
APE1135	708,868	709,917	>	349	O-sialoglycoprotein endopeptidase	OWL:AB00946412	109.50	47.73
APE1139	711,198	711,650	>	150	30S ribosomal protein S13	Swiss_Prot:P54012	86.08	60.42
APE1140	711,805	713,226	>	473	aspartate kinase	Swiss_Prot:Q57991	63.12	39.02
APE1143	713,223	714,272	>	349	aspartate-semialdehyde dehydrogenase	PIR:E69206	108.95	47.09
APE1144	714,263	715,267	>	334	homoserine dehydrogenase	PIR:C69154	30.80	33.59
APE1154	719,058	719,723	>	221	30S ribosomal protein S3	OWL:AB00946620	91.91	50.79
APE1178	730,712	730,149	<	187	dTDP-4-dehydrorhamnose 3,5-epimerase	OWL:LIU61226	66.16	47.28
APE1179	731,633	730,716	<	305	dTDP-4-dehydrorhamnose reductase	OWL:AP000002111	62.86	41.25
APE1180	732,619	731,627	<	330	dTDP-glucose 4,6-dehydratase	PIR:D69290	77.27	50.75
APE1181	733,699	732,632	<	355	glucose-1-phosphate thymidylyltransferase	PIR:E69290	78.50	39.39
APE1191	746,786	745,695	<	363	capM protein	PIR:C69098	43.66	30.73
APE1192	747,663	746,791	<	290	dolichol monophosphate mannose synthase	PIR:E69322	19.29	28.49
APE1193	748,343	747,756	<	195	capsular synthesis regulator component B	OWL:AF0712151	11.39	22.68
APE1195	749,434	748,874	<	186	adenylylsulfate kinase	PIR:S74917	62.65	45.24
APE1197	750,622	751,791	>	389	sulfate adenylyltransferase	PIR:B69839	118.71	48.54
APE1202	753,560	754,642	>	360	sugar phosphate transferase	PIR:D64437	76.76	37.91
APE1205	754,793	756,670	>	625	glucosamine-fructose-6-phosphate aminotransferase	OWL:AB009473	103.64	42.42
APE1208	756,722	757,804	>	360	glucose-1-phosphate thymidylyltransferase	PIR:D64437	22.09	31.85
APE1211	763,287	763,670	>	127	gastrin precursor	Swiss_Prot:P48757	9.35	25.97
APE1221	772,391	773,461	>	356	tropomyosin	OWL:U9387282	29.36	42.50
APE1223	773,599	774,780	>	393	cystathione beta-synthase	PIR:C42790	61.64	41.40
APE1226	774,777	775,931	>	384	cystathione gamma-lyase	Swiss_Prot:P18757	89.33	52.37
APE1245	792,718	791,720	<	332	alcohol dehydrogenase	Swiss_Prot:Q57517	44.29	29.32
APE1253	795,518	796,438	>	306	ABC-transporter ATP-binding protein	PIR:B69377	43.82	46.67
APE1255	796,450	797,211	>	253	ABC transporter ATP-binding protein	OWL:AP000006241	46.27	41.18
APE1261	800,410	800,093	<	105	formylmethanofuran dehydrogenase	OWL:MKFWUGDB6	8.17	33.33
APE1265	801,213	802,235	>	340	dehydrogenase	OWL:AP000002294	36.50	31.48
APE1275	806,006	807,409	>	467	tryptophanase	Swiss_Prot:P31014	122.59	49.02
APE1279	810,828	810,304	<	174	riboflavin biosynthesis enzyme	Swiss_Prot:Q58085	37.56	44.36
APE1285	813,976	815,097	>	373	acetylpolyamine aminohydrolase	OWL:AP000005176	56.34	36.69
APE1288	816,917	820,699	>	1,260	nitrate reductase alpha chain	Swiss_Prot:P42175	176.97	51.32
APE1294	820,733	822,133	>	466	nitrate reductase beta chain	Swiss_Prot:P42176	128.61	50.32
APE1300	824,070	824,618	>	182	nitrate reductase delta chain	Swiss_Prot:P42178	20.90	32.30
APE1301	825,361	824,615	<	248	high-affinity branched-chain amino acid transport ATP-binding protein	Swiss_Prot:P30294	48.66	48.00
APE1304	827,831	826,782	<	349	high-affinity branched-chain amino acid transport permease protein	PIR:H69352	42.39	32.66
APE1305	828,720	827,836	<	294	high-affinity branched-chain amino acid transport permease protein	PIR:A69353	43.40	29.97
APE1307	830,682	828,727	<	651	long-chain-fatty-acid-CoA ligase	Swiss_Prot:P44446	75.18	27.30
APE1308	831,438	830,686	<	250	high-affinity branched-chain amino acid transport ATP-binding protein	PIR:G69352	56.40	44.00
APE1310	831,654	833,243	>	529	luciferase	PIR:A69831	49.31	28.60
APE1317	834,911	835,831	>	306	methanol dehydrogenase regulator	PIR:C69791	104.86	50.17
APE1324	839,616	839,092	<	174	heterocyst differentiation protein	PIR:S58288	20.15	27.70
APE1328	840,722	843,079	>	785	hyuA	PIR:D64420	208.87	49.63
APE1331	843,086	844,774	>	562	hyuB	PIR:C64420	181.47	52.01
APE1335	845,237	846,256	>	339	deblocking aminopeptidase	OWL:AP000002215	107.81	45.81
APE1340	849,459	853,589	>	1,376	reverse gyrase	OWL:PFU66557	177.77	41.88
APE1346	857,515	855,533	<	660	acetyl-CoA synthetase	PIR:G69371	161.31	54.10
APE1350	858,472	858,005	<	155	molybdenum cofactor biosynthesis protein	PIR:F69518	72.97	53.90
APE1353	858,753	862,574	>	1,273	ATP-dependent DNA helicase	PIR:F69901	25.34	36.81
APE1366	869,385	868,912	<	157	6,7-dimethyl-8-ribityllumazine synthase	PIR:H69515	82.32	57.04
APE1367	869,642	871,822	>	726	transitional endoplasmic reticulum ATPase	Swiss_Prot:Q07590	110.09	70.29
APE1376	874,765	878,433	>	1,222	reverse gyrase	OWL:PFU66557	164.08	36.83
APE1380	878,464	879,414	>	316	transport ATP-binding protein	OWL:AP00000671	32.74	36.04
APE1382	880,421	881,773	>	450	hyuE	OWL:SCI A65	20.17	23.87
APE1386	882,894	884,165	>	423	glutamate dehydrogenase	Swiss_Prot:P80053	97.17	56.79
APE1395	888,205	889,215	>	336	dipeptide transport system permease protein	OWL:AP000006115	132.26	52.71
APE1396	889,217	890,380	>	387	dipeptide transport system permease protein	OWL:AP000006116	152.98	56.28
APE1399	890,388	891,389	>	333	dipeptide transport system ATP-binding protein	OWL:AP000006117	103.01	66.98
APE1402	891,395	892,438	>	347	oligopeptide transport ATP-binding protein	OWL:AP000006118	78.84	65.20

Table 5. Continued.

ORF ID	Position Initiation	Position Termination	Direction	Length (aa)	Predicted Product	Locus	Zscore	% Identity
APE1404	893,748	892,465	<	427	3-phosphoshikimate 1-carboxyvinyltransferase	PIR:D69202	57.39	30.94
APE1410	900,652	899,246	<	468	NADH dehydrogenase (ubiquinone) chain 2	Swiss_Prot:P72714	33.28	34.80
APE1411	902,771	900,645	<	708	NADH dehydrogenase (ubiquinone) chain 5	PIR:A69478	32.68	33.84
APE1415	904,316	902,775	<	513	NADH dehydrogenase (ubiquinone) chain 13	PIR:H69477	46.86	39.22
APE1417	904,630	904,322	<	102	NADH-plastoquinone oxidoreductase chain 4L	Swiss_Prot:P26524	22.61	38.46
APE1418	905,159	904,638	<	173	NADH dehydrogenase (ubiquinone) chain 6	PIR:S44413	15.51	26.44
APE1419	905,706	905,110	<	198	NADH-plastoquinone oxidoreductase subunit I	OWL:AP000006152	24.89	41.24
APE1421	906,854	905,703	<	383	NADH dehydrogenase (ubiquinone) chain 1	PIR:F69478	52.63	33.13
APE1422	908,107	906,863	<	414	NADH dehydrogenase (ubiquinone) chain 7	PIR:E69478	94.59	44.85
APE1426	908,661	908,092	<	189	NADH-plastoquinone oxidoreductase subunit J	PIR:D69478	40.17	42.75
APE1428	909,282	908,674	<	202	NADH dehydrogenase (ubiquinone) chain 10	PIR:D69478	63.85	56.93
APE1432	912,063	909,853	<	736	elongation factor ef-2	Swiss_Prot:P23112	152.25	68.12
APE1440	914,604	914,233	<	123	prefoldin subunit	PIR:E69393	33.13	33.33
APE1445	918,316	917,486	<	276	autoantigen	PIR:H69190	103.14	45.02
APE1447	919,053	918,313	<	246	ribonuclease PH	OWL:AP000006260	120.60	57.90
APE1449	920,680	919,904	<	258	proteasome , alpha subunit	OWL:AP000006264	66.41	46.19
APE1452	922,735	922,058	<	225	50S ribosomal protein L15	Swiss_Prot:P54060	98.74	64.74
APE1454	923,075	925,582	>	835	cation-transporting ATPase	PIR:G69071	81.64	38.71
APE1458	927,922	926,537	<	461	mercuric reductase	Swiss_Prot:P16171	57.92	34.40
APE1461	929,369	928,500	<	289	acetylglutamate kinase	OWL:AB0066813	43.19	34.67
APE1462	930,400	929,333	<	355	N-acetyl-gamma-glutamyl-phosphate reductase	PIR:H69212	88.20	43.91
APE1463	930,732	931,592	>	286	ribosomal protein S6, modification protein	OWL:AB006681	73.71	38.91
APE1464	931,616	932,782	>	388	acetylornithine aminotransferase	PIR:G69044	71.13	38.46
APE1471	934,775	935,218	>	147	GTP cyclohydrolase I	Swiss_Prot:P51599	15.51	30.00
APE1472	936,176	935,247	<	309	2-oxoacid--ferredoxin oxidoreductase , beta chain	PIR:JC4920	134.52	55.59
APE1473	938,016	936,166	<	616	2-oxoacid--ferredoxin oxidoreductase , alpha chain	PIR:JC4919	154.27	40.26
APE1484	944,227	942,248	<	659	3-hydroxybutyryl-CoA hydratase	PIR:E69285	116.03	45.28
APE1488	944,833	945,870	>	345	threonine synthase	OWL:AB0095219	56.20	34.03
APE1491	945,876	946,547	>	223	siroheme synthase	PIR:G69448	35.62	32.67
APE1498	951,607	950,558	<	349	threonine dehydratase	Swiss_Prot:P05792	77.13	42.26
APE1507	953,816	955,150	>	444	IMP dehydrogenase	OWL:AP000001314	67.85	37.05
APE1517	958,932	960,110	>	392	myo-inositol-1-phosphate synthase	OWL:AP000006316	158.72	53.19
APE1522	961,033	962,073	>	346	replication factor C subunit	PIR:C69507	121.48	55.06
APE1524	962,273	963,517	>	414	replication factor C subunit	PIR:D64410	72.41	31.89
APE1526	964,410	963,985	<	141	CDP-diacylglycerol--glycerol-3-phosphatidyltransferase	PIR:B69093	32.29	40.88
APE1533	967,700	968,161	>	153	translation initiation factor 2 beta subunit	OWL:AP0000003	53.55	40.30
APE1535	969,228	970,439	>	403	pyruvate formate lyase activating protein	OWL:AP00000518	105.48	49.46
APE1547	977,626	975,878	<	582	acylamino-acid-releasing enzyme	PIR:S74053	118.88	35.01
APE1548	977,772	978,788	>	338	nodulation ATP-binding protein I	OWL:RSU533277	30.12	36.32
APE1557	982,204	983,073	>	289	alcohol dehydrogenase	PIR:F69905	22.13	26.43
APE1570	992,544	991,690	<	284	4-hydroxybenzoate octaprenyl transferase	PIR:H69521	65.45	44.40
APE1576	998,727	997,594	<	377	oligopeptide transport ATP-binding protein	OWL:AB0095304	76.25	55.08
APE1578	999,704	998,730	<	324	oligopeptide transport ATP-binding protein	OWL:AB0095303	77.69	59.43
APE1581	1,001,294	999,711	<	527	oligopeptide transport system permease protein	OWL:AB0095302	104.78	57.83
APE1582	1,002,404	1,001,313	<	363	oligopeptide transport permease protein	OWL:AB0095301	78.13	36.64
APE1586	1,006,080	1,007,249	>	389	cysteine synthase	Swiss_Prot:P16703	35.69	30.28
APE1591	1,008,887	1,009,201	>	104	acylphosphatase	Swiss_Prot:P75877	27.26	43.24
APE1592	1,009,242	1,010,669	>	475	cysteinyl-tRNA synthetase	OWL:AP00000332	130.30	47.68
APE1593	1,010,318	1,009,722	<	198	malonyl-CoA:acyl carrier protein acyltransferase	OWL:STMFABD	11.90	29.90
APE1594	1,010,654	1,011,457	>	267	HESA protein	OWL:AB00946624	66.25	42.48
APE1604	1,018,377	1,016,761	<	538	CTP synthetase	OWL:AB00952610	110.97	58.33
APE1605	1,018,511	1,019,278	>	255	4-nitrophenylphosphatase	OWL:AAM59612	62.27	37.45
APE1616	1,026,137	1,024,773	<	454	3-phosphonopyruvate decarboxylase	PIR:F69468	85.42	37.90
APE1618	1,029,520	1,026,908	<	870	aconitate hydratase	Swiss_Prot:P37032	110.33	54.49
APE1623	1,032,717	1,031,071	<	548	Cytochrome C oxidase polypeptide	OWL:AB0087572	110.83	42.45
APE1629	1,034,461	1,034,156	<	101	protein translation factor SU11	Swiss_Prot:O29348	55.09	57.14
APE1632	1,036,627	1,037,493	>	288	sugar transport system permease protein	OWL:AB00946515	134.84	55.68
APE1633	1,037,501	1,038,313	>	270	ABC transporter permease protein	OWL:AB00946516	112.42	54.31
APE1635	1,038,314	1,039,378	>	354	ABC transporter ATP-binding protein	OWL:AB00946517	60.03	57.90
APE1637	1,040,534	1,039,518	<	338	agmatinase	Swiss_Prot:P19268	61.64	34.07
APE1639	1,040,567	1,042,318	>	583	glycyl-tRNA synthetase	OWL:AP000006325	150.16	47.99
APE1646	1,043,361	1,043,924	>	187	tRNA intron endonuclease	PIR:D69131	39.87	35.66
APE1647	1,044,514	1,043,921	<	197	3-octaprenyl-4-hydroxybenzoate carboxy-lyase	PIR:G69062	69.35	44.39
APE1648	1,045,068	1,044,484	<	194	protein 2	PIR:E70385	60.40	45.57
APE1652	1,046,533	1,047,702	>	389	coenzyme PQQ synthesis protein	PIR:D69390	64.68	37.07
APE1655	1,047,692	1,048,903	>	403	coenzyme PQQ synthesis protein	PIR:F69551	122.47	44.53
APE1657	1,048,939	1,049,586	>	215	fuculose-1-phosphate aldolase	PIR:C69054	40.48	35.26
APE1659	1,049,589	1,050,596	>	335	K+ channel beta 1 subunit	OWL:AB0094658	41.53	37.87
APE1662	1,052,016	1,051,528	<	162	aspartate carbamoyltransferase regulatory chain	Swiss_Prot:P74766	58.30	44.08
APE1663	1,052,951	1,052,013	<	312	aspartate carbamoyltransferase catalytic chain	OWL:AP000003117	89.81	52.46
APE1668	1,054,112	1,053,792	<	106	periplasmic divalent cation tolerance protein	PIR:H70386	58.64	48.49
APE1669	1,055,599	1,054,196	<	467	lipoamide dehydrogenase component (E3) of pyruvate dehydrogenase complex	Swiss_Prot:P11959	60.74	39.01
APE1671	1,056,922	1,055,684	<	412	dihydrolipoamide acetyltransferase component (E2) of pyruvate dehydrogenase complex	Swiss_Prot:P11961	72.06	40.57
APE1674	1,057,915	1,056,938	<	325	pyruvate dehydrogenase E1 component, beta subunit	Swiss_Prot:P21874	82.09	50.46
APE1677	1,059,225	1,057,930	<	431	pyruvate dehydrogenase E1 component, alpha subunit	OWL:MTCY07A72	80.33	44.00
APE1681	1,059,437	1,060,618	>	393	acyl-CoA dehydrogenase	PIR:D69411	70.75	41.87
APE1686	1,062,093	1,061,602	<	163	methylmalonyl-CoA mutase alpha-subunit	OWL:AB00947332	72.49	58.92

Table 5. Continued.

ORF ID	Position Initiation	Position Termination	Direction	Length (aa)	Predicted Product	Locus	Zscore	% Identity
APE1687	1,063,762	1,062,065	<	565	methylmalonyl-CoA mutase alpha-subunit	OWL:AP0000069	162.50	59.40
APE1691	1,066,642	1,065,833	<	269	gufA protein	OWL:AP00000328	86.49	43.31
APE1692	1,066,818	1,067,441	>	207	inorganic pyrophosphatase	Swiss_Prot:P50308	101.45	61.99
APE1694	1,068,275	1,068,760	>	161	cytochrome AA3 controlling protein	Swiss_Prot:Q04443	24.81	27.91
APE1695	1,069,051	1,070,100	>	349	aminomethyltransferase	Swiss_Prot:P54378	63.53	35.07
APE1702	1,072,392	1,073,153	>	253	ATP-dependent transporter	Swiss_Prot:P80866	67.51	46.25
APE1713	1,079,588	1,080,772	>	394	citrate synthase	Swiss_Prot:Q53554	71.63	36.51
APE1717	1,082,098	1,083,564	>	488	Glu-tRNA amidotransferase, subunit A	PIR:F70322	103.42	48.49
APE1718	1,083,561	1,085,021	>	486	Glu-tRNA amidotransferase, subunit B	PIR:D69514	114.15	41.63
APE1720	1,085,946	1,086,536	>	196	cytochrome oxidase subunit II	OWL:NPCBAABCDI	40.71	33.99
APE1724	1,090,069	1,090,827	>	252	Rieske protein	PIR:S71348	37.81	45.46
APE1725	1,090,849	1,092,564	>	571	cytochrome b	Swiss_Prot:P39480	56.99	34.31
APE1730	1,094,856	1,096,523	>	555	molybdenum transport system permease protein	OWL:SHU753493	29.75	29.45
APE1732	1,096,530	1,097,606	>	358	transporter ATP-binding protein	OWL:AB00946517	53.09	65.11
APE1735	1,099,472	1,098,147	<	441	signal recognition protein	OWL:SAFFH	116.03	58.50
APE1737	1,099,840	1,100,175	>	111	30S ribosomal protein S13	Swiss_Prot:P95986	73.88	63.64
APE1739	1,100,190	1,100,711	>	173	30S ribosomal protein S4	Swiss_Prot:P95987	91.76	58.24
APE1742	1,100,734	1,101,129	>	131	30S ribosomal protein S11	Swiss_Prot:P95988	82.03	78.13
APE1744	1,101,197	1,102,069	>	290	DNA-directed RNA polymerase subunit D	Swiss_Prot:P95989	61.44	34.31
APE1746	1,102,066	1,102,428	>	120	50S ribosomal protein L18	OWL:AB00952118	59.87	55.26
APE1747	1,102,461	1,102,928	>	155	50S ribosomal protein L13	Swiss_Prot:P95991	61.10	41.14
APE1749	1,102,925	1,103,398	>	157	30S ribosomal protein S9	Swiss_Prot:P95992	62.16	51.35
APE1750	1,103,648	1,104,265	>	205	30S ribosomal protein S2	Swiss_Prot:P95993	92.53	57.87
APE1752	1,104,541	1,104,882	>	113	enterovirus capsid protein VPI	OWL:EV7M4276B	11.80	26.15
APE1756	1,108,694	1,106,760	<	644	arginyl-tRNA synthetase	Swiss_Prot:O29368	47.68	35.64
APE1764	1,112,675	1,111,662	<	337	geranyltransferase	Swiss_Prot:P39464	79.63	40.75
APE1775	1,119,301	1,117,685	<	538	thermostable carboxypeptidase	OWL:AP000002160	207.98	52.22
APE1786	1,125,447	1,123,990	<	485	aldehyde dehydrogenase	OWL:TTGAPN	121.81	55.79
APE1791	1,128,165	1,126,858	<	435	tRNA nucleotidyltransferase	OWL:SSU66004	59.91	36.00
APE1794	1,129,032	1,131,053	>	673	DNA topoisomerase	OWL:AP00000317	131.96	39.48
APE1798	1,132,365	1,133,171	>	268	myo-inositol-1(or 4)-monophosphatase	Swiss_Prot:P29218	16.61	27.01
APE1800	1,135,503	1,134,469	<	344	DOM34	PIR:S75405	70.20	32.25
APE1805	1,139,582	1,136,913	<	889	valyl-tRNA synthetase	Swiss_Prot:Q58413	122.70	38.36
APE1816	1,144,139	1,145,494	>	451	fumarate hydratase	Swiss_Prot:P39461	127.27	54.84
APE1818	1,146,015	1,146,398	>	127	30S ribosomal protein HS6	OWL:AP000006203	90.91	76.03
APE1829	1,154,829	1,153,792	>	345	ATP-binding transport protein	OWL:AB0095277	49.05	51.79
APE1831	1,155,133	1,156,176	>	347	D-3-phosphoglycerate dehydrogenase	OWL:AP000002294	117.35	56.21
APE1832	1,156,263	1,158,242	>	659	acylamino-acid-releasing enzyme	OWL:AP000003263	141.51	40.06
APE1834	1,158,657	1,159,625	>	322	ribose-phosphate pyrophosphokinase	OWL:AB00952922	62.63	40.36
APE1842	1,164,100	1,163,786	<	104	30S ribosomal protein S10	Swiss_Prot:P17199	77.94	69.00
APE1844	1,165,746	1,164,433	<	437	elongation factor l-alpha	Swiss_Prot:P41203	50.47	73.21
APE1846	1,166,218	1,165,820	<	132	30S ribosomal protein S7	Swiss_Prot:P41206	88.19	73.49
APE1849	1,167,539	1,167,096	<	147	30S ribosomal protein S12	Swiss_Prot:P11524	91.84	74.15
APE1851	1,168,373	1,168,065	<	102	50S ribosomal protein L30	Swiss_Prot:P29160	47.84	44.32
APE1852	1,169,635	1,168,403	<	410	DNA-directed RNA polymerase subunit A"	Swiss_Prot:P11514	93.89	45.23
APE1853	1,172,328	1,169,641	<	895	DNA-directed RNA polymerase subunit A'	Swiss_Prot:P11512	130.30	57.60
APE1856	1,175,868	1,172,359	<	1,169	DNA-directed RNA polymerase subunit B	Swiss_Prot:P11513	147.40	52.87
APE1862	1,178,219	1,177,551	<	222	TATA-box binding protein	OWL:SABAPG	89.44	63.59
APE1869	1,184,154	1,182,889	<	421	3-hydroxy-3-methylglutaryl-coenzyme A reductase	Swiss_Prot:O08424	101.72	58.87
APE1872	1,185,763	1,184,540	<	407	acetyl CoA synthase	OWL:PFPORVOR3	70.34	33.60
APE1873	1,186,811	1,185,714	<	365	acyl carrier protein synthase	OWL:PFPORVOR4	145.02	58.17
APE1877	1,189,640	1,188,870	<	256	membrane protein	PIR:G70414	58.20	46.99
APE1878	1,189,798	1,190,715	>	305	cytochrome C oxidase assembly factor	PIR:C69504	79.40	39.35
APE1885	1,194,313	1,195,140	>	275	5'-methylthioadenosine phosphorylase	OWL:AB00946824	125.28	57.92
APE1891	1,199,034	1,198,318	<	238	glutamine transport ATP-binding protein	PIR:H69334	40.31	50.66
APE1892	1,199,733	1,199,071	<	220	glutamine transport system permease protein	PIR:H69278	65.16	42.86
APE1893	1,200,715	1,199,837	<	292	ABC transporter binding protein	PIR:S68867	37.07	44.68
APE1905	1,208,330	1,207,719	<	203	flagellin	PIR:F69381	33.98	35.64
APE1907	1,209,092	1,208,352	<	246	flagellin	PIR:G69381	29.27	53.23
APE1911	1,211,547	1,210,693	<	284	ABC transporter ATP-binding protein	OWL:AF0474293	31.25	32.62
APE1914	1,212,745	1,211,627	<	372	drug resistance protein	PIR:D69415	23.94	28.40
APE1921	1,215,982	1,215,479	<	167	DNA endonuclease I in intron	OWL:AB0087451	147.36	99.23
APE1929	1,220,003	1,219,284	<	239	homing DNA endonuclease in intron	OWL:AB008745	229.46	100.00
APE1934	1,224,882	1,224,052	<	276	nodulation ATP-binding protein	OWL:AP000006241	44.06	42.34
APE1954	1,234,938	1,234,363	<	191	N-terminal acetyltransferase complex subunit	OWL:AB0094744	31.54	51.28
APE1958	1,238,444	1,237,704	<	246	NH(3)-dependent NAD(+) synthetase	PIR:E69068	80.00	43.83
APE1959	1,239,957	1,238,644	<	437	DNA/pantothenate metabolism flavoprotein	OWL:AP000006150	83.56	35.73
APE1961	1,240,727	1,239,909	<	272	stress protein	OWL:AB0094658	85.13	45.97
APE1962	1,240,800	1,242,128	>	442	serine hydroxymethyltransferase	OWL:AB00952240	133.20	53.03
APE1963	1,242,436	1,243,635	>	399	alcohol dehydrogenase	PIR:H69252	70.97	34.71
APE1968	1,246,865	1,245,915	<	316	carbamate kinase	OWL:AB0165212	104.52	47.28
APE1976	1,251,300	1,252,691	>	463	seryl-tRNA synthetase	Swiss_Prot:O28244	140.47	54.07
APE1988	1,257,636	1,256,611	<	341	eukaryotic peptide chain release factor subunit I	OWL:AP000006295	95.74	48.46
APE1992	1,259,437	1,258,229	<	402	ornithine carbamoyltransferase	OWL:AP000003123	84.82	56.17
APE2002	1,265,434	1,264,451	<	327	modification methylase	Swiss_Prot:Q58600	90.63	49.62
APE2007	1,267,678	1,266,560	<	372	GTP-binding protein	OWL:AP00000538	86.53	39.02
APE2012	1,268,321	1,269,550	>	409	26S protease regulatory subunit	Swiss_Prot:O28303	68.68	61.13
APE2021	1,272,972	1,273,400	>	142	ferric uptake regulation protein	PIR:H69528	26.24	30.77
APE2023	1,277,043	1,275,808	<	411	tRNA splicing protein SPL1	PIR:F70019	94.18	39.28

Table 5. Continued.

ORF ID	Position Initiation	Position Termination	Direction	Length (aa)	Predicted Product	Locus	Zscore	% Identity
APE2035	1,281,536	1,282,690	>	384	histidinol-phosphate aminotransferase	OWL:STU906251	30.28	29.76
APE2039	1,283,441	1,284,412	>	323	cobalamin biosynthesis protein	PIR:F69054	50.59	38.92
APE2044	1,286,947	1,286,021	<	308	hydrogenase expression/formation protein	OWL:AP000006284	67.01	37.50
APE2048	1,288,037	1,287,039	<	332	5-methyltetrahydropteroylglutamate-homocysteinemethyltransferase	OWL:AP000004223	89.00	38.13
APE2051	1,289,514	1,289,164	<	116	non specific lipid-transfer protein	Swiss_Prot:Q42999	7.37	31.51
APE2052	1,289,288	1,290,394	>	368	developmentally regulated GTP-binding protein	OWL:AB00952222	110.47	45.38
APE2060	1,295,463	1,294,297	>	388	multiple sugar-binding transport ATP-binding protein	OWL:AB00947118	62.77	55.77
APE2062	1,296,184	1,299,135	>	983	ribonucleotide reductase	OWL:MTV0396	48.12	44.91
APE2066	1,301,838	1,300,660	<	392	N-acetylglucosaminyl-phosphatidylinositol biosynthetic protein	OWL:AB00952712	71.32	33.68
APE2067	1,301,874	1,302,905	>	343	homoserine kinase	Swiss_Prot:Q58504	47.12	32.34
APE2068	1,302,912	1,304,081	>	389	cystathione gamma-synthase	OWL:AP000004227	87.49	44.67
APE2071	1,305,827	1,305,009	<	272	hypothetical proteinxanthine phosphoribosyltransferase	OWL:AB0094675	61.16	39.61
APE2072	1,306,352	1,308,019	>	555	thermosome, subunit	Swiss_Prot:Q427435	91.03	66.67
APE2074	1,308,204	1,309,298	>	364	tyrosyl-tRNA synthetase	Swiss_Prot:P59582	159.08	57.46
APE2081	1,313,720	1,314,868	>	382	aminopeptidase	Swiss_Prot:P42778	32.49	31.92
APE2085	1,316,082	1,316,528	>	148	initiation factor 5A	Swiss_Prot:P28461	93.75	59.56
APE2091	1,321,210	1,320,080	<	376	glucokinase	OWL:MTY20H1029	40.84	36.33
APE2098	1,327,105	1,324,751	<	784	DNA-directed DNA polymerase	PIR:B56277	234.66	58.83
APE2105	1,328,604	1,329,368	>	254	uridine phosphorylase	Swiss_Prot:P12758	70.49	42.11
APE2106	1,329,332	1,330,273	>	313	ribokinase	OWL:SSU521101	37.14	29.11
APE2108	1,331,478	1,330,270	<	402	acetyl-CoA acetyltransferase	PIR:S75020	76.96	45.29
APE2121	1,340,506	1,338,941	<	521	glycine dehydrogenase subunit 2	OWL:AB0094645	159.98	54.18
APE2124	1,341,924	1,340,509	<	471	glycine dehydrogenase subunit 1	OWL:AB0094644	120.41	43.68
APE2125	1,342,056	1,342,550	>	164	bacterioferritin comigratory protein	OWL:MTV0095	39.57	41.41
APE2126	1,342,967	1,344,901	>	644	2-oxoacid-ferredoxin oxidoreductase alpha chain	PIR:JC4919	148.57	41.60
APE2128	1,344,891	1,345,847	>	318	2-oxoacid-ferredoxin oxidoreductase beta chain	PIR:JC4920	154.13	64.91
APE2130	1,347,580	1,346,396	<	394	glutamate-1-semialdehyde 2,1-aminomutase	PIR:H69404	76.49	37.40
APE2136	1,350,296	1,351,270	>	324	immunogenic protein	PIR:C69329	105.06	44.98
APE2142	1,354,216	1,352,828	<	462	glutamine synthetase	Swiss_Prot:P23794	81.02	46.10
APE2149	1,360,726	1,361,550	>	274	thiazole biosynthetic enzyme	OWL:AP00000662	81.62	41.27
APE2153	1,363,899	1,363,063	<	278	erythrocyte band 7 integral membrane protein	PIR:C69427	90.38	57.71
APE2158	1,369,691	1,368,480	<	403	acyl-CoA dehydrogenase	Swiss_Prot:P45867	67.65	44.30
APE2162	1,371,866	1,373,680	>	604	carbon starvation protein A	PIR:E70409	78.71	35.98
APE2165	1,374,458	1,375,051	>	197	arsenite-translocating ATPase	PIR:E64442	26.63	27.54
APE2166	1,377,767	1,375,065	<	900	alanyl-tRNA synthetase	OWL:AB0094745	164.79	41.48
APE2170	1,378,331	1,377,996	<	111	50S ribosomal protein L12	PIR:C69436	36.72	49.04
APE2171	1,379,419	1,378,385	<	344	acidic ribosomal protein P0	OWL:U852623	72.42	36.94
APE2173	1,380,062	1,379,424	<	212	50S ribosomal protein L1	PIR:F69091	71.22	43.60
APE2174	1,380,449	1,380,123	<	108	50S ribosomal protein L11	Swiss_Prot:P35025	67.88	62.00
APE2176	1,380,990	1,380,532	<	152	transcription termination-antitermination factor	Swiss_Prot:P27341	60.94	41.26
APE2192	1,392,773	1,391,508	<	421	aspartyl-tRNA synthetase	OWL:AP000004152	122.92	45.50
APE2196	1,394,092	1,393,391	<	233	fibrillarin	OWL:AB00946628	100.07	55.61
APE2199	1,395,378	1,394,107	<	423	nucleolar protein NOP5	OWL:AB00946627	116.36	58.09
APE2200	1,395,938	1,397,221	>	427	L-asparaginase	OWL:AP000006169	155.59	52.94
APE2202	1,397,224	1,399,152	>	642	Glu-tRNA amidotransferase, subunit B	Swiss_Prot:Q60325	176.88	49.84
APE2213	1,404,723	1,405,214	>	163	isoquinoline 1-oxidoreductase alpha subunit	PIR:S49123	85.80	56.95
APE2216	1,405,322	1,407,574	>	750	nicotine dehydrogenase chain C	OWL:MTV0367	176.72	47.15
APE2219	1,407,574	1,408,452	>	292	nicotine dehydrogenase chain A	PIR:A56279	84.31	39.02
APE2229	1,415,093	1,413,180	<	637	DNA-directed DNA polymerase	Swiss_Prot:Q07635	58.95	43.24
APE2239	1,419,485	1,418,406	<	359	alcohol dehydrogenase	Swiss_Prot:P50381	70.64	39.94
APE2247	1,424,120	1,423,536	<	194	DNA-3-methyladenine glycosidase	PIR:D70082	20.58	34.95
APE2248	1,425,383	1,424,154	<	409	aspartate aminotransferase	Swiss_Prot:P14909	117.60	50.39
APE2252	1,427,240	1,426,329	<	303	maltose transport system permease protein	OWL:AB00947119	114.72	51.04
APE2253	1,428,067	1,427,237	<	276	lactose transport system permease protein	OWL:AP000001208	139.45	59.06
APE2257	1,431,509	1,433,056	>	515	transport system binding protein	PIR:F69470	139.91	57.49
APE2258	1,433,183	1,434,202	>	339	peptide transport system permease protein	PIR:C69470	136.78	56.42
APE2260	1,434,223	1,435,065	>	280	peptide transport system permease protein	PIR:H69470	98.72	50.56
APE2261	1,435,073	1,436,035	>	320	peptide transport ATP-binding protein	PIR:A69471	68.87	59.54
APE2263	1,436,036	1,437,079	>	347	peptide transport ATP-binding protein	PIR:B69471	66.14	52.92
APE2267	1,438,441	1,439,805	>	454	glutamate decarboxylase	PIR:C69536	135.29	49.10
APE2274	1,446,780	1,449,779	>	999	cytochrome C-type biogenesis protein	PIR:H69523	27.26	30.98
APE2277	1,452,278	1,453,141	>	287	nitrilase	Swiss_Prot:P77192	48.65	37.91
APE2278	1,453,439	1,454,191	>	250	thioredoxin peroxidase	Swiss_Prot:Q58146	100.73	55.81
APE2284	1,457,775	1,459,574	>	599	long-chain-fatty-acid-CoA ligase	PIR:C69471	100.87	50.80
APE2286	1,461,124	1,459,943	<	393	threonine synthase	PIR:G69131	110.88	50.68
APE2290	1,464,742	1,463,024	<	572	acylamino-acid-releasing enzyme	OWL:AP000003263	100.90	34.85
APE2291	1,465,851	1,465,132	<	239	glucose-1-phosphate thymidyltransferase	PIR:C69070	55.96	39.38
APE2296	1,470,180	1,468,930	<	416	glutamyl-tRNA reductase	Swiss_Prot:P42809	43.25	32.37
APE2298	1,471,097	1,470,180	<	305	porphobilinogen deaminase	PIR:A69405	66.04	43.83
APE2299	1,472,398	1,471,100	<	432	glutamate-1-semialdehyde 2,1-aminomutase	PIR:C70371	91.83	44.89
APE2300	1,473,432	1,472,434	<	332	delta-aminolevulinic acid dehydratase	Swiss_Prot:P43087	84.87	51.24
APE2302	1,473,768	1,475,189	>	473	phenylalanyl-tRNA synthetase alpha chain	Swiss_Prot:P95961	125.10	43.74
APE2305	1,475,191	1,476,837	>	548	phenylalanyl-tRNA synthetase beta chain	PIR:S75388	139.68	39.12
APE2316	1,482,920	1,481,571	<	449	tryptophan synthase beta chain	OWL:AP000006294	184.25	69.70
APE2317	1,484,624	1,482,990	<	544	glutamyl-tRNA synthetase	Swiss_Prot:P95968	148.51	49.37
APE2324	1,486,490	1,486,089	<	133	50S ribosomal protein L35	OWL:AP000007149	26.62	35.63
APE2326	1,487,116	1,487,511	>	131	vacuolar ATP synthase subunit	Swiss_Prot:P23040	35.98	46.74
APE2328	1,487,591	1,489,048	>	485	prolyl-tRNA synthetase	OWL:AP000004138	165.59	50.73

Table 5. Continued.

ORF ID	Position Initiation	Position Termination	Direction	Length (aa)	Predicted Product	Locus	Zscore	% Identity
APE2341	1,496,791	1,495,997	<	264	lipoate-protein ligase A	Swiss_Prot:P47512	38.74	37.66
APE2344	1,498,453	1,497,566	<	295	lipoic acid synthetase	PIR:E70417	93.71	52.65
APE2349	1,499,625	1,500,161	>	178	orotate phosphoribosyltransferase	OWL:SSPYREGEN	57.21	42.35
APE2354	1,502,075	1,503,292	>	405	acyl-CoA dehydrogenase	PIR:D69411	50.11	32.50
APE2366	1,509,658	1,508,396	<	420	eukaryotic translational initiation factor 2 gamma subunit	OWL:AB00952443	138.36	62.50
APE2371	1,510,453	1,509,716	<	245	30S ribosomal protein S6	Swiss_Prot:P54067	46.16	57.53
APE2374	1,510,539	1,512,392	>	617	transcription initiation factor IF-2	OWL:SAU43413	162.40	56.05
APE2377	1,512,835	1,512,410	<	141	nucleoside diphosphate kinase	PIR:G69345	69.51	64.03
APE2382	1,515,936	1,517,018	>	360	non specific lipid-transfer protein	PIR:H69370	107.37	39.61
APE2384	1,517,045	1,518,229	>	394	non specific lipid-transfer protein	PIR:G69370	199.18	66.16
APE2397	1,521,873	1,522,976	>	367	Sua5 protein	OWL:AP000002129	114.33	48.99
APE2400	1,524,394	1,523,000	<	464	phosphomethylpyrimidine kinase	OWL:AP00000561	156.73	53.20
APE2406	1,525,788	1,525,096	<	230	ABC transporter ATP-binding protein	OWL:PDAJ34606	22.31	32.35
APE2408	1,526,558	1,525,773	<	261	ABC transporter ATP-binding protein	PIR:G64435	26.41	31.25
APE2414	1,530,219	1,528,777	<	480	DNA-damage-inducible protein	OWL:AP000002274	74.13	30.60
APE2418	1,530,851	1,531,696	>	281	electron transfer flavoprotein beta-subunit	Swiss_Prot:P97089	77.02	41.22
APE2419	1,531,703	1,532,821	>	372	electron transfer flavoprotein alpha-subunit	Swiss_Prot:P71153	91.49	40.30
APE2421	1,532,842	1,534,152	>	436	electron transfer flavoprotein-quinone oxidoreductase	Swiss_Prot:P77337	92.70	37.07
APE2423	1,534,230	1,535,378	>	382	soluble hydrogenase subunit	PIR:G64419	95.62	39.78
APE2426	1,538,333	1,537,095	<	412	Na+/H+ antiporter	PIR:B64459	35.27	27.30
APE2430	1,539,118	1,538,627	<	163	thioredoxin	OWL:AE0012605	27.46	30.39
APE2433	1,539,715	1,541,193	>	492	phosphosugar mutase	OWL:AP000005116	78.66	34.51
APE2437	1,543,770	1,544,351	>	193	deoxyribose-phosphate aldolase	PIR:C69209	50.76	43.01
APE2439	1,545,900	1,546,874	>	324	mevalonate kinase	Swiss_Prot:Q50559	41.35	33.10
APE2441	1,546,931	1,548,706	>	591	acylamino-acid-releasing enzyme	Swiss_Prot:P34422	45.47	33.62
APE2443	1,548,791	1,549,768	>	325	transcription initiation factor IIB	Swiss_Prot:P50387	88.12	43.10
APE2450	1,556,022	1,554,448	<	524	cytosol aminopeptidase	OWL:AE0012792	68.37	42.57
APE2452	1,556,181	1,557,719	>	512	GMP synthase	PIR:E70321	115.50	49.90
APE2458	1,559,754	1,561,052	>	432	enolase	OWL:AB00952941	75.35	57.18
APE2460	1,561,523	1,561,825	>	100	ferredoxin like protein	Swiss_Prot:P53658	21.15	36.00
APE2461	1,561,889	1,563,013	>	374	tryptophanyl-tRNA synthetase	Swiss_Prot:O26352	136.07	48.77
APE2463	1,564,321	1,564,647	>	108	steroidogenic acute regulatory protein	OWL:SSHTOPR	58.56	49.49
APE2464	1,565,833	1,564,661	<	390	phoH like protein	Swiss_Prot:P46343	30.19	27.18
APE2465	1,566,014	1,567,276	>	420	fmn protein	OWL:AP000003251	104.49	37.65
APE2474	1,575,244	1,573,145	<	699	transitional endoplasmic reticulum ATPase	PIR:H69411	100.27	63.65
APE2480	1,576,023	1,575,652	<	123	translation elongation factor 1-beta	Swiss_Prot:Q64214	45.44	44.94
APE2486	1,578,488	1,579,273	>	261	beta-aspartate methyltransferase	OWL:MTCY26113	43.75	31.95
APE2495	1,584,068	1,583,607	<	153	histidine triad (HIT) protein	PIR:S74070	44.72	39.81
APE2496	1,584,007	1,584,609	>	200	thioredoxin	Swiss_Prot:P33791	30.60	33.02
APE2500	1,585,832	1,589,350	>	1,172	antibiotic transport-associated protein	PIR:C40046	26.48	21.85
APE2503	1,590,424	1,591,155	>	243	3-oxoacyl-[acyl-carrier protein] reductase	PIR:H70447	43.34	39.17
APE2507	1,591,862	1,592,842	>	326	formate dehydrogenase	PIR:B69472	114.32	55.45
APE2509	1,592,859	1,594,529	>	556	acetolactate synthase large subunit	PIR:C69472	123.49	43.66
APE2516	1,597,183	1,596,212	<	323	ATP-binding protein	OWL:AB0095277	57.56	50.85
APE2520	1,599,949	1,599,131	<	272	arylmalonte decarboxylase	OWL:AP000004117	123.82	55.79
APE2521	1,599,940	1,601,232	>	430	Leu/Ile/Val-binding protein	PIR:S75741	34.09	28.79
APE2522	1,601,401	1,602,198	>	265	high-affinity branched-chain amino acid transport permease protein	PIR:A69353	57.27	32.71
APE2523	1,602,195	1,603,193	>	332	high-affinity branched-chain amino acid transport permease protein	PIR:H69352	47.10	31.58
APE2524	1,603,181	1,603,909	>	242	high-affinity branched-chain amino acid transport ATP-binding protein	PIR:G69352	46.22	44.03
APE2526	1,603,909	1,604,631	>	240	high-affinity branched-chain amino acid transport ATP-binding protein	PIR:F69352	58.79	52.50
APE2528	1,604,628	1,606,691	>	687	hydantoinase	PIR:D64420	228.98	54.05
APE2530	1,606,685	1,608,352	>	555	hydantoinase	PIR:C64420	188.67	52.75
APE2546	1,620,144	1,619,359	<	261	indole-3-glycerol phosphate synthase	Swiss_Prot:Q06121	46.61	39.82
APE2548	1,622,065	1,620,782	<	427	tryptophan synthase beta chain	OWL:AP000006294	125.48	53.08
APE2550	1,622,829	1,622,071	<	252	tryptophan synthase alpha chain	Swiss_Prot:P50382	35.35	26.14
APE2551	1,622,991	1,624,018	>	345	anthranilate phosphoribosyltransferase	Swiss_Prot:Q57686	75.45	40.55
APE2553	1,624,385	1,625,338	>	317	anthranilate synthase component I	Swiss_Prot:Q06128	75.84	47.91
APE2555	1,625,331	1,625,921	>	196	anthranilate synthase component II	Swiss_Prot:Q06129	71.12	50.79
APE2571	1,632,892	1,634,757	>	621	heavy-metal transporting ATPase	Swiss_Prot:P20021	50.80	28.84
APE2575	1,637,467	1,636,274	<	397	tyrosine aminotransferase	OWL:LSAJ13303	37.34	28.62
APE2578	1,639,283	1,638,432	<	283	membrane protein	PIR:S74047	58.76	42.77
APE2579	1,640,286	1,639,285	<	333	ABC transporter ATP-binding protein	PIR:S74048	58.86	44.82
APE2583	1,643,937	1,642,672	<	421	proton/sodium-glutamate symport protein	OWL:AP000005205	138.79	60.62
APE2586	1,646,191	1,645,286	<	301	ribose ABC transport system permease protein	OWL:AB00952439	117.80	48.68
APE2588	1,648,776	1,647,280	<	498	ribose ABC transport system ATP-binding protein	OWL:AB00952437	89.79	49.90
APE2595	1,651,372	1,650,716	<	218	thiosulfate sulfurtransferase	Swiss_Prot:O05793	95.32	53.49
APE2601	1,654,183	1,655,007	>	274	N-G,N-G-dimethylarginine dimethylaminohydrolase	OWL:D86041	23.80	39.73
APE2604	1,657,628	1,656,087	<	513	acetyl-coenzyme A synthetase	PIR:F69295	102.50	52.14
APE2605	1,658,304	1,659,056	>	250	polysulfide reductase	PIR:E69269	98.54	59.32
APEs001	76,180	76,416	>	78	ferredoxin	Swiss_Prot:P29603	20.94	33.90
APEs002	84,782	84,934	<	50	50S ribosomal protein L40	Swiss_Prot:P54058	33.01	59.46
APEs003	133,178	133,381	>	67	50S ribosomal protein L38	Swiss_Prot:O22860	20.36	39.66
APEs004	165,982	166,188	>	68	50S ribosomal protein L23	PIR:E71187	52.61	65.15
APEs005	183,720	183,890	<	56	DNA-directed RNA polymerase,subunit E"	PIR:C69389	29.34	44.68
APEs006	184,574	184,744	<	56	ferredoxin	Swiss_Prot:P80168	16.22	45.16

Table 5. Continued.

ORF ID	Position Initiation	Position Termination	Direction	Length (aa)	Predicted Product	Locus	Zscore	% Identity
APES008	196,429	196,611	<	60	repressor protein	Swiss_Prot:P06965	8.71	32.65
APES015	306,121	306,342	<	73	30S ribosomal protein S27	Swiss_Prot:P54028	42.73	52.54
APES016	306,359	306,550	<	63	50S ribosomal protein L44	Swiss_Prot:P27074	25.68	49.12
APES019	315,792	315,986	<	64	ribonuclease p3	PIR:S69085	18.00	87.50
APES022	361,519	361,752	>	77	small nuclear ribonucleoprotein	PIR:C69186	36.33	46.67
APES023	361,782	361,970	>	62	50S ribosomal protein L37	PIR:C71028	53.31	62.90
APES024	370,582	370,857	>	91	50S ribosomal protein L13	OWL:RLAJ3198	18.41	49.06
APES041	658,128	658,400	>	90	DNA-directed RNA polymerase, subunit M	Swiss_Prot:O13896	8.62	33.96
APES045	760,865	761,038	>	57	aerolysin regulatory protein	Swiss_Prot:P09165	10.52	39.47
APES054	917,209	917,469	<	86	50S ribosomal protein L37	PIR:B71204	42.89	46.91
APES058	1,103,406	1,103,639	>	77	DNA-directed RNA polymerase, subunit N	PIR:A69391	41.43	45.76
APES061	1,175,870	1,176,142	<	90	DNA-directed RNA polymerase, subunit H	PIR:D69485	51.10	58.57
APES067	1,305,985	1,306,209	<	74	30S ribosomal protein S17	Swiss_Prot:P54026	44.93	58.93
APES068	1,395,524	1,395,676	>	50	30S ribosomal protein S30	Swiss_Prot:Q05472	19.30	46.51
APES070	1,513,640	1,513,831	<	63	50S ribosomal protein L24	PIR:F69345	48.89	60.38
APES071	1,513,891	1,514,166	>	91	30S ribosomal protein S28	PIR:C71025	68.57	76.81
APES074	1,657,629	1,657,865	<	78	Acetyl-CoA synthetase	Swiss_Prot:P27097	20.50	53.66

The ORF name, initiation and termination positions, direction of reading and length of amino acid residues of ORFs are shown in the left five columns. The right four columns represent the probable gene products, gene locus of genes which showed the highest similarity to corresponding ORFs, Zscore taken from Smith-Waterman similarity search and percent identity between the reported genes and assigned ORFs.

Table 6. List of ORFs with motifs.

ORF ID	Position Initiation	Position Termination	Direction	Length (aa)	Motifs assigned
APE004	2,174	1,260	<	304	gram-positive cocci surface proteins 'anchoring' hexapeptide
APE033	20,898	22,379	>	493	ATP/GTP-binding site motif A (P-loop)
APE034	22,033	21,668	<	121	prokaryotic membrane lipoprotein lipid attachment site
APE054	33,600	33,058	<	180	tonB-dependent receptor proteins signatures
APE0073	50,533	52,197	>	554	N-6 Adenine-specific DNA methylases signature
APE0080	56,596	54,512	<	694	ATP/GTP-binding site motif A (P-loop)
APE0090	64,466	63,975	<	163	prokaryotic membrane lipoprotein lipid attachment site
APE0111	81,414	81,947	>	177	prokaryotic membrane lipoprotein lipid attachment site
APE0112	83,240	83,749	>	169	prokaryotic membrane lipoprotein lipid attachment site
APE0130	97,001	97,999	>	332	prokaryotic membrane lipoprotein lipid attachment site
APE0138	103,068	102,703	<	121	ATP synthase alpha and beta subunit signature
APE0154	115,851	114,493	<	452	prokaryotic membrane lipoprotein lipid attachment site
APE0163	124,215	122,824	<	463	prokaryotic membrane lipoprotein lipid attachment site
APE0170	126,927	127,283	>	118	ATP/GTP-binding site motif A (P-loop) ; prokaryotic membrane lipoprotein lipid attachment site
APE0179	133,085	132,240	<	281	SRP54-type proteins GTP-binding domain signature
APE0186	137,641	137,309	<	110	crystallins beta and gamma 'Greek key' motif signature
APE0189	139,900	139,598	<	100	prokaryotic membrane lipoprotein lipid attachment site
APE0219	159,880	158,495	<	461	ATP/GTP-binding site motif A (P-loop)
APE0223	161,921	162,328	>	135	prokaryotic membrane lipoprotein lipid attachment site
APE0232	167,920	168,267	>	115	prokaryotic membrane lipoprotein lipid attachment site
APE0242	175,106	175,429	>	107	2Fe-2S ferredoxins, iron-sulfur binding region signature
APE0247	177,017	177,601	>	194	thiolases signatures
APE0256	183,717	183,160	<	185	prokaryotic membrane lipoprotein lipid attachment site
APE0283	206,831	208,840	>	669	ATP/GTP-binding site motif A (P-loop)
APE0292	213,021	212,701	<	106	gram-positive cocci surface proteins 'anchoring' hexapeptide
APE0294	213,536	214,009	>	157	multicopper oxidases signatures
APE0298	215,194	215,586	>	130	prokaryotic membrane lipoprotein lipid attachment site
APE0305	221,950	222,600	>	216	prokaryotic membrane lipoprotein lipid attachment site
APE0340	249,296	247,047	<	749	ATP/GTP-binding site motif A (P-loop)
APE0366	259,045	259,611	>	188	prokaryotic membrane lipoprotein lipid attachment site
APE0386	270,592	271,005	>	137	prokaryotic membrane lipoprotein lipid attachment site
APE0397	274,689	275,045	>	118	prokaryotic membrane lipoprotein lipid attachment site
APE0409	282,320	281,976	<	114	serine proteases, trypsin family, active sites
APE0415	291,639	288,073	<	1,188	regulator of chromosome condensation (RCC1) signatures
APE0424	298,856	299,269	>	137	prokaryotic membrane lipoprotein lipid attachment site
APE0428	303,153	301,243	<	636	prokaryotic membrane lipoprotein lipid attachment site
APE0431	303,833	304,396	>	187	ATP/GTP-binding site motif A (P-loop)
APE0434	305,005	305,367	>	120	prokaryotic membrane lipoprotein lipid attachment site
APE0472	329,651	329,133	<	172	tonB-dependent receptor proteins signatures
APE0481	335,324	335,016	<	102	4Fe-4S ferredoxins, iron-sulfur binding region signature
APE0486	338,138	337,794	<	114	prokaryotic membrane lipoprotein lipid attachment site
APE0540	372,037	370,874	<	387	ATP/GTP-binding site motif A (P-loop)
APE0548	376,327	375,746	<	193	prokaryotic membrane lipoprotein lipid attachment site
APE0554	378,497	379,144	>	215	N-6 Adenine-specific DNA methylases signature
APE0565	384,956	385,537	>	193	prokaryotic membrane lipoprotein lipid attachment site
APE0576	388,764	389,156	>	130	prokaryotic membrane lipoprotein lipid attachment site
APE0588	393,438	396,176	>	912	prokaryotic membrane lipoprotein lipid attachment site
APE0614	418,219	416,585	<	544	lipocalin signature
APE0617	418,227	418,640	>	137	prokaryotic membrane lipoprotein lipid attachment site
APE0620	425,424	419,623	<	1,933	mitochondrial energy transfer proteins signature
APE0625	427,791	427,396	<	131	prokaryotic membrane lipoprotein lipid attachment site
APE0626	428,224	427,715	<	169	prokaryotic membrane lipoprotein lipid attachment site
APE0627	429,153	428,758	<	131	zinc finger, C2H2 type, domain
APE0641	434,604	433,555	<	349	thioredoxin family active site
APE0643	434,144	434,674	>	176	prokaryotic membrane lipoprotein lipid attachment site
APE0653	438,163	438,798	>	211	gram-positive cocci surface proteins 'anchoring' hexapeptide
APE0705	469,437	469,916	>	159	gram-positive cocci surface proteins 'anchoring' hexapeptide; prokaryotic membrane lipoprotein lipid attachment site
APE0759	500,630	501,706	>	358	gram-positive cocci surface proteins 'anchoring' hexapeptide
APE0774	509,651	510,844	>	397	N-6 Adenine-specific DNA methylases signature
APE0787	519,490	520,416	>	308	transcription factor TFIIB repeat signature
APE0796	528,664	529,365	>	233	legume lectins signatures
APE0823	545,622	546,302	>	226	prenyl group binding site (CAAX box) ; prokaryotic membrane lipoprotein lipid attachment site
APE0824	546,649	545,948	<	233	neutral zinc metallopeptidases, zinc-binding region signature
APE0837	553,499	554,215	>	238	aminoacyl-transfer RNA synthetases class-II signatures
APE0851	559,911	559,507	<	134	prokaryotic membrane lipoprotein lipid attachment site
APE0867	567,267	568,112	>	281	neutral zinc metallopeptidases, zinc-binding region signature
APE0880	576,234	577,838	>	534	ATP/GTP-binding site motif A (P-loop)
APE0892	586,334	587,476	>	380	ATP/GTP-binding site motif A (P-loop)
APE0897	588,901	589,251	>	116	prokaryotic membrane lipoprotein lipid attachment site
APE0898	589,753	592,311	>	852	gram-positive cocci surface proteins 'anchoring' hexapeptide
APE0963	629,175	629,762	>	195	prokaryotic membrane lipoprotein lipid attachment site
APE0984	646,204	645,629	<	191	prokaryotic membrane lipoprotein lipid attachment site
APE0999	652,949	653,536	>	195	ATP/GTP-binding site motif A (P-loop)
APE1007	657,169	658,074	>	301	aminoacyl-transfer RNA synthetases class-II signatures
APE1076	686,760	686,443	<	105	aldehyde dehydrogenases active sites
APE1082	689,187	689,591	>	134	prokaryotic membrane lipoprotein lipid attachment site
APE1088	691,810	692,289	>	159	prokaryotic membrane lipoprotein lipid attachment site
APE1122	704,552	704,968	>	138	binding-protein-dependent transport systems inner membrane components signature

Table 6. Continued.

ORF ID	Position Initiation	Position Termination	Direction	Length (aa)	Motifs assigned
APE1151	718,419	717,868	<	183	prokaryotic membrane lipoprotein lipid attachment site
APE1175	728,323	727,871	<	150	prokaryotic membrane lipoprotein lipid attachment site
APE1213	764,158	768,282	>	1,374	serine proteases, subtilase family, active sites
APE1216	768,368	769,969	>	533	ribosomal protein S2 signatures
APE1218	770,849	770,142	<	235	prokaryotic membrane lipoprotein lipid attachment site
APE1234	781,349	780,435	<	304	prokaryotic membrane lipoprotein lipid attachment site
APE1244	791,512	789,800	<	570	lipases, serine active site ; prokaryotic membrane lipoprotein lipid attachment site
APE1274	804,971	805,525	>	184	prokaryotic membrane lipoprotein lipid attachment site
APE1277	809,886	809,320	<	188	phosphopantetheine attachment site
APE1303	826,697	825,405	<	430	prokaryotic membrane lipoprotein lipid attachment site
APE1309	830,771	831,136	>	121	prokaryotic membrane lipoprotein lipid attachment site
APE1311	832,234	831,854	<	126	bacterial regulatory proteins, lysR family signature
APE1315	833,830	834,936	>	368	crystallins beta and gamma 'Greek key' motif signature
APE1322	837,325	837,822	>	165	aminoacyl-transfer RNA synthetases class-II signatures
APE1326	840,888	839,725	<	387	aldehyde dehydrogenases active sites; nitrite and sulfite reductases iron-sulfur/siroheme-binding site; prokaryotic membrane lipoprotein lipid attachment site
APE1348	856,641	857,072	>	143	prokaryotic membrane lipoprotein lipid attachment site
APE1392	887,352	886,810	<	180	prokaryotic membrane lipoprotein lipid attachment site
APE1413	901,702	902,028	>	108	prokaryotic membrane lipoprotein lipid attachment site
APE1433	909,998	910,468	>	156	prokaryotic membrane lipoprotein lipid attachment site
APE1435	911,306	911,803	>	165	prokaryotic membrane lipoprotein lipid attachment site
APE1439	913,927	914,475	>	182	prokaryotic membrane lipoprotein lipid attachment site
APE1468	933,223	932,807	<	138	ATP synthase alpha and beta subunits signature
APE1482	941,320	940,946	<	124	prokaryotic membrane lipoprotein lipid attachment site
APE1493	946,525	947,247	>	240	prokaryotic membrane lipoprotein lipid attachment site
APE1499	950,583	951,041	>	152	prokaryotic membrane lipoprotein lipid attachment site
APE1504	953,346	953,038	<	102	prokaryotic membrane lipoprotein lipid attachment site
APE1519	960,138	960,536	>	132	prokaryotic membrane lipoprotein lipid attachment site
APE1534	969,058	968,192	<	288	prokaryotic membrane lipoprotein lipid attachment site
APE1544	973,098	974,015	>	305	prokaryotic membrane lipoprotein lipid attachment site
APE1553	980,168	980,536	>	122	prokaryotic membrane lipoprotein lipid attachment site
APE1561	987,536	987,952	<	138	neutral zinc metallopeptidases, zinc-binding region signature
APE1566	988,955	989,299	>	114	prokaryotic membrane lipoprotein lipid attachment site
APE1615	1,024,726	1,023,671	<	351	TNF family signature and profile
APE1619	1,027,311	1,027,715	>	134	prokaryotic membrane lipoprotein lipid attachment site
APE1709	1,076,164	1,076,556	>	130	prokaryotic membrane lipoprotein lipid attachment site
APE1711	1,078,508	1,079,467	>	319	aldo/keto reductase family signatures
APE1715	1,081,752	1,080,844	<	302	peroxidases signatures
APE1716	1,081,808	1,082,122	>	104	glycine radical signature
APE1719	1,085,123	1,085,707	>	194	cytochrome c family heme-binding site signature
APE1726	1,092,371	1,091,922	<	149	prokaryotic membrane lipoprotein lipid attachment site
APE1733	1,097,503	1,097,066	<	145	prokaryotic membrane lipoprotein lipid attachment site
APE1748	1,103,355	1,102,810	>	181	prokaryotic membrane lipoprotein lipid attachment site
APE1757	1,107,397	1,108,080	>	227	aldo/keto reductase family signatures
APE1787	1,124,639	1,124,995	>	118	prokaryotic membrane lipoprotein lipid attachment site
APE1799	1,133,202	1,134,035	>	277	cytochrome c family heme-binding site signature
APE1802	1,134,791	1,135,135	>	114	prokaryotic membrane lipoprotein lipid attachment site
APE1810	1,140,809	1,142,485	>	558	ATP/GTP-binding site motif A (P-loop)
APE1812	1,141,995	1,141,384	<	203	bacterial regulatory proteins, gntR family signature
APE1815	1,145,037	1,144,132	<	301	prokaryotic membrane lipoprotein lipid attachment site
APE1820	1,149,068	1,147,656	<	470	apoptosis regulator proteins, Bcl-2 family BH domains signatures/profiles
APE1855	1,172,045	1,172,368	>	107	prokaryotic membrane lipoprotein lipid attachment site
APE1860	1,174,574	1,175,329	>	251	prokaryotic membrane lipoprotein lipid attachment site
APE1879	1,191,036	1,190,704	<	110	prokaryotic membrane lipoprotein lipid attachment site
APE1883	1,191,742	1,192,152	>	136	prokaryotic membrane lipoprotein lipid attachment site
APE1920	1,214,934	1,215,437	>	167	prokaryotic membrane lipoprotein lipid attachment site
APE1923	1,216,957	1,216,640	>	105	prokaryotic membrane lipoprotein lipid attachment site
APE1924	1,217,925	1,218,350	>	141	prokaryotic membrane lipoprotein lipid attachment site
APE1956	1,236,638	1,235,436	<	400	lipocalin signature
APE1973	1,248,412	1,250,532	<	706	ribosomal protein S2 signatures
APE1978	1,252,953	1,253,483	>	176	prokaryotic membrane lipoprotein lipid attachment site
APE2001	1,264,439	1,263,975	<	154	bacterial regulatory proteins, gntR family signature
APE2005	1,265,616	1,266,005	>	129	prokaryotic membrane lipoprotein lipid attachment site
APE2008	1,266,696	1,267,109	>	137	prokaryotic membrane lipoprotein lipid attachment site
APE2028	1,278,621	1,278,151	<	156	trp-Asp (WD-40) repeats signature
APE2040	1,284,563	1,284,210	<	117	prokaryotic membrane lipoprotein lipid attachment site
APE2043	1,285,931	1,285,602	<	109	ribosomal protein S14 signature
APE2056	1,292,756	1,292,289	>	155	mutT domain signature
APE2079	1,313,114	1,312,668	<	148	prokaryotic membrane lipoprotein lipid attachment site
APE2082	1,314,444	1,313,980	<	154	prokaryotic membrane lipoprotein lipid attachment site
APE2083	1,314,454	1,314,062	<	130	immunoglobulins and major histocompatibility complex proteins signature
APE2093	1,321,205	1,322,749	>	514	ATP/GTP-binding site motif A (P-loop)
APE2119	1,338,235	1,336,766	<	489	prokaryotic membrane lipoprotein lipid attachment site
APE2123	1,339,800	1,340,186	>	128	prokaryotic membrane lipoprotein lipid attachment site
APE2129	1,345,980	1,346,351	>	123	gram-positive cocci surface proteins 'anchoring' hexapeptide
APE2132	1,346,786	1,347,175	>	129	prokaryotic membrane lipoprotein lipid attachment site
APE2138	1,351,675	1,351,307	<	122	prokaryotic membrane lipoprotein lipid attachment site
APE2148	1,359,442	1,359,897	>	151	prokaryotic membrane lipoprotein lipid attachment site
APE2163	1,373,730	1,374,053	>	107	peroxidases signatures

Table 6. Continued.

ORF ID	Position Initiation	Position Termination	Direction	Length (aa)	Motifs assigned
APE2164	1,374,060	1,374,467	>	135	ATP/GTP-binding site motif A (P-loop)
APE2169	1,377,928	1,378,284	>	118	prokaryotic membrane lipoprotein lipid attachment site
APE2187	1,386,739	1,385,831	<	302	prokaryotic membrane lipoprotein lipid attachment site
APE2190	1,389,893	1,390,957	>	354	EF-hand calcium-binding domain ; gram-positive cocci surface proteins 'anchoring' hexapeptide
APE2191	1,390,932	1,390,258	<	224	prokaryotic membrane lipoprotein lipid attachment site
APE2211	1,403,501	1,403,055	<	148	prokaryotic membrane lipoprotein lipid attachment site
APE2233	1,414,987	1,414,583	<	134	prokaryotic membrane lipoprotein lipid attachment site
APE2246	1,422,983	1,423,549	>	188	eukaryotic putative RNA-binding region RNP-1 signature
APE2292	1,467,563	1,466,016	<	515	lipases, serine active site ; tonB-dependent receptor proteins signatures
APE2310	1,477,855	1,478,952	>	365	prokaryotic membrane lipoprotein lipid attachment site
APE2325	1,486,623	1,487,096	<	157	phospholipase A2 active sites signatures
APE2329	1,489,289	1,488,822	<	155	prokaryotic membrane lipoprotein lipid attachment site
APE2358	1,504,387	1,503,749	<	212	prokaryotic membrane lipoprotein lipid attachment site
APE2368	1,508,742	1,509,155	>	137	pkB family of carbohydrate kinases signatures
APE2383	1,516,744	1,516,247	<	165	ATP synthase alpha and beta subunits signature
APE2385	1,517,802	1,517,449	<	117	prokaryotic membrane lipoprotein lipid attachment site
APE2401	1,523,649	1,524,407	>	252	prokaryotic membrane lipoprotein lipid attachment site
APE2405	1,525,109	1,524,783	<	108	gram-positive cocci surface proteins 'anchoring' hexapeptide
APE2409	1,526,992	1,526,483	>	169	ATP/GTP-binding site motif A (P-loop) ; prokaryotic membrane lipoprotein lipid attachment site
APE2411	1,527,536	1,526,943	<	197	G-protein coupled receptors signature
APE2429	1,538,592	1,539,062	>	156	prokaryotic membrane lipoprotein lipid attachment site
APE2434	1,540,709	1,540,353	<	118	prokaryotic membrane lipoprotein lipid attachment site
APE2454	1,558,082	1,557,762	<	106	prokaryotic membrane lipoprotein lipid attachment site
APE2478	1,575,425	1,575,826	>	133	prokaryotic membrane lipoprotein lipid attachment site
APE2487	1,579,230	1,578,730	<	166	prokaryotic membrane lipoprotein lipid attachment site
APE2508	1,593,077	1,592,763	<	104	prokaryotic membrane lipoprotein lipid attachment site
APE2534	1,610,413	1,610,099	<	104	prokaryotic membrane lipoprotein lipid attachment site
APE2552	1,624,302	1,623,898	<	134	ATP/GTP-binding site motif A (P-loop) ; prokaryotic membrane lipoprotein lipid attachment site
APE2560	1,628,248	1,628,625	>	125	prokaryotic membrane lipoprotein lipid attachment site
APE2567	1,631,425	1,630,772	<	217	prokaryotic membrane lipoprotein lipid attachment site
APE2568	1,632,097	1,631,432	<	221	prokaryotic membrane lipoprotein lipid attachment site
APE2593	1,649,757	1,650,080	>	107	prokaryotic membrane lipoprotein lipid attachment site
APE2600	1,652,475	1,652,804	>	109	prokaryotic membrane lipoprotein lipid attachment site
APE2614	1,664,255	1,663,881	<	124	prokaryotic membrane lipoprotein lipid attachment site
APES013	295,342	295,518	>	58	prokaryotic membrane lipoprotein lipid attachment site
APES018	309,642	309,842	<	66	lipocalin signature
APES034	503,403	503,633	>	76	prokaryotic membrane lipoprotein lipid attachment site
APES036	549,567	549,731	>	54	phospholipase A2 active sites signatures
APES037	568,627	568,878	>	83	prokaryotic membrane lipoprotein lipid attachment site
APES042	729,946	730,107	>	53	prenyl group binding site (CAAX box)
APES048	837,843	838,043	>	66	ATP/GTP-binding site motif A (P-loop)
APES051	869,419	869,589	<	56	prokaryotic membrane lipoprotein lipid attachment site
APES052	871,917	872,078	<	53	prenyl group binding site (CAAX box)
APES059	1,103,409	1,103,606	<	65	prokaryotic membrane lipoprotein lipid attachment site
APES072	1,563,366	1,563,638	>	90	prokaryotic membrane lipoprotein lipid attachment site
APES073	1,614,964	1,615,182	>	72	zinc carboxypeptidases, zinc-binding regions signatures

The ORF name, initiation and termination positions, direction of reading and length of amino acid residues of ORFs are shown in the left five columns, and the assigned protein-motifs in the right column.

