

論文 / 著書情報  
Article / Book Information

題目(和文)	古細菌のFタイプATPase遺伝子の発見
Title(English)	
著者(和文)	鷲見正人
Author(English)	
出典(和文)	学位:博士(理学), 学位授与機関:東京工業大学, 報告番号:甲第3890号, 授与年月日:1998年5月31日, 学位の種別:課程博士, 審査員:
Citation(English)	Degree:Doctor (Science), Conferring organization: Tokyo Institute of Technology, Report number:甲第3890号, Conferred date:1998/5/31, Degree Type:Course doctor, Examiner:
学位種別(和文)	博士論文
Type(English)	Doctoral Thesis

## 目次

第1章 序論 .....	1
第2章 PCR 法によるバクテリアのF/Vタイプ ATPase 遺伝子の確認 .....	6
第3章 古細菌からのFタイプ ATPase 遺伝子の単離と塩基配列決定 .....	15
第4章 古細菌におけるFタイプ ATPase 遺伝子の存在とバクテリア進化との関係 .....	26
引用文献 .....	33
付録A <i>M. barkeri</i> Fタイプ ATPase 遺伝子の全塩基配列と予想されるアミノ酸配列	
付録B <i>M. barkeri</i> Fタイプ ATPase 遺伝子産物と既知のFタイプ構成成分との相同性	
付録C VおよびFタイプ、触媒および非触媒サブユニットのアライメント	

## 第1章 序論

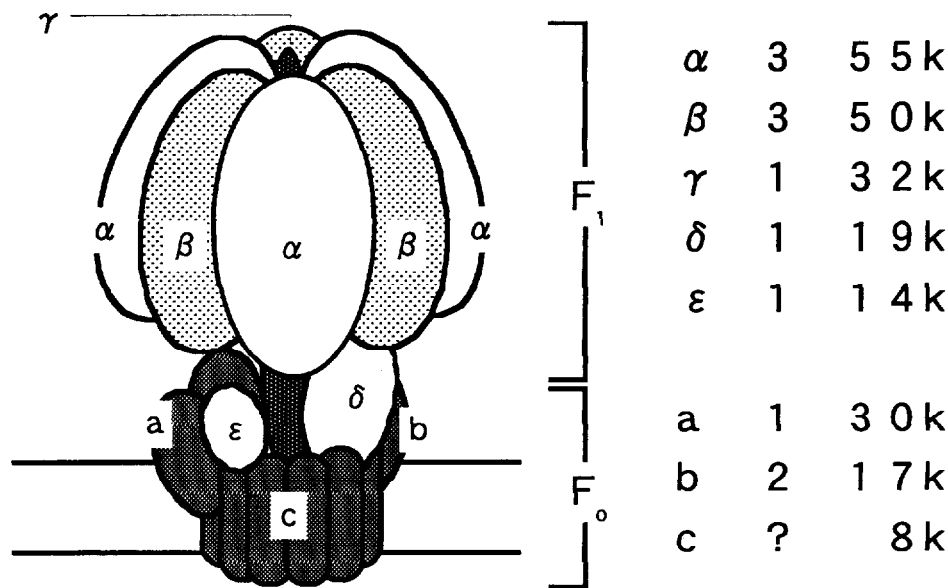
1977年に Woese らによりはじめて提唱され、以後、遺伝子解析技術の進歩とそれに伴う情報の蓄積により定着した古細菌という生物群には、当初、真正細菌（Fタイプ）とは異なる、真核生物の空胞系にあるものと同じ種類の ATPase（Vタイプ）のみが存在すると信じられていた。この状況から導き出された、古細菌と真正細菌がそれぞれ固有のタイプの ATPase を進化の過程で獲得したという説は魅力的であったが、真正細菌より Vタイプが見つかり修正の必要をせまられる。本研究は、F/Vタイプが真正細菌/古細菌の区別なく広範に分布する可能性を裏付ける情報を提供し、特にまだ報告されていない古細菌における Fタイプの存在を示唆する情報を得るために取り組まれた。

### 1.1 $F_0F_1$ -ATPase

生命体の能動的活動には ATP が必要である。この ATP のほとんどは、電子伝達系の酵素群などを用いて膜の内外に生じさせた電気化学的ポテンシャル差を利用し、ADP とリン酸より合成される。ATP 合成をつかさどる細胞内小器官としては、ミトコンドリアが有名だが、その内膜にあり ATP 合成の役割を演じているのが  $F_0F_1$ -ATP 合成酵素である。この酵素の反応は可逆で、ATP を水解することで膜を隔ててプロトンを輸送することが可能なこと、 $F_0$  から遊離した  $F_1$  部分のみで ATP 水解能を持ち、これが多くの研究者の標品となっていることから、本酵素は ATP 分解酵素を意味する ATPase とも呼ばれる。

$F_0F_1$ -ATPase は、膜に埋め込まれた疎水性ペプチドよりなるプロトンチャンネルの  $F_0$  部分と、ATP 合成/水解をになう触媒部位を含む大きな水溶性部分の  $F_1$  より構成される、比較的大きく複雑で特徴的な構造を持つ酵素である[Boyer, 1990; Weber & Senior, 1997]。ミトコンドリア[Walker *et al.*, 1985]以外にも、やや仕組みが単純化された同種の酵素が、同じ真核生物である植物の葉緑体[Graber, 1987]、原核生物でも大腸菌をはじめとする様々な真正細菌[Dunn and Futai, 1980]より見つかっている。

各種生物より単離された同酵素のサブユニット組成を比較することで、この酵素の基本的構成成分が分かっている[Foster & Fillingame, 1982; Yoshida *et al.*, 1979]。  $F_0$  部分については a、b、c の3種類、 $F_1$  部分については  $\alpha$ 、 $\beta$ 、 $\gamma$ 、 $\delta$ 、 $\epsilon$  の5種類のポリペプチドよりなる。 $F_1$  部分は、交互にそれぞれ3つずつ並ぶ 50kDa 程度の大きな  $\alpha$ 、 $\beta$  というサブユニットに基質結合部位があり、そのうち  $\beta$  が ATP 合成/水解活性を持つことが知られている[Boyer, 1987; Hisabori *et al.*, 1986; Dancna *et al.*, 1986] (図01)。また、この大きな2つの主要サブユニットはその一次構造において、お互いに 25% 程度の相同性を持ち、かつて遺伝子重複により分化したものであることが示唆されている[Gorgarten *et al.*, 1989]。さらに、 $\alpha$ 、 $\beta$  それぞれが種が変わっても 50% 程度の高い相同性を保つ。



〈図01〉  $F_0F_1$ -ATPase の構造を示す略図。薄いグレイが触媒サブユニットの  $\beta$  で、濃いグレイが  $\alpha, \beta, \gamma$  の中心を貫く  $\gamma$ 。中庸のグレイは  $F_0$  の構成成分を示す。右手の表は左からサブユニット名、酵素 1 分子に含まれる数、大腸菌におけるサブユニットの分子量である。c サブユニットの数はまだ決められていないが 10 から 12 とされている。

### 1.2 $F_0F_1$ -ATPase の普遍的分布と古細菌における例外

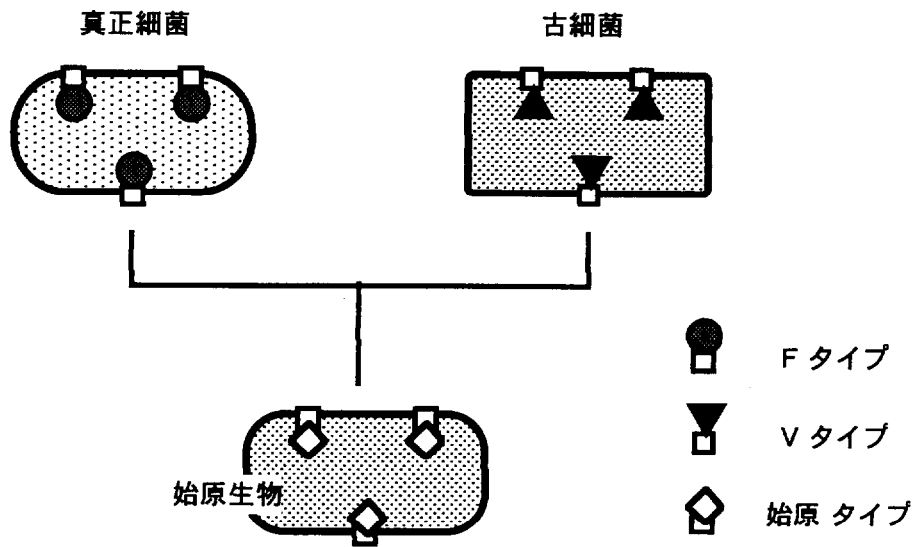
ATP の合成（もしくは逆反応による膜電位の生成）という機能の重要性や、バクテリアから高等生物までという幅広い分布と種間での保存性の高さより、当初は ATP 合成能を放棄した一部の寄生的細胞を除き、すべての生物が  $F_0F_1$ -ATPase を持っていると考えられていた。そのため、1977 年 Woese らが、16S rRNA の類似性にもとづき、真核生物と真正細菌とは別に第 3 の生命分類群として提唱した古細菌 [Woese *et al.*, 1987] にも、 $F_0F_1$ -ATPase が存在するものであると信じられていた。

しかし、1980 年代、古細菌に属する好塩菌 [Ihara *et al.*, 1991]、メタン産生菌 [Inatomi *et al.*, 1989]、好酸好熱性細菌 [Denda *et al.*, 1988a, 1988b, 1989, 1990] より相次いで単離された  $F_0F_1$ -ATPase 様酵素の  $F_1$  相当部分は、2 つ大きな  $\alpha$ 、 $\beta$  という構成成分を持つ点では  $F_0F_1$  との類似性を見ることができたが、 $\alpha$  が 60k と  $F_1$  の  $\alpha$  (55k) に比べ非常に大きいこと、 $F_1$  の特異的な阻害剤であるアザイドで阻害されないなどの決定的な違いを有していた。伝田らが好酸好熱性細菌、*Sulfolobus acidocaldarius* より単離した遺伝子の解析の結果、この ATPase の主要サブユニットは、 $F_1$  のそれと 25% 程度の相同性しか持たない一方で、ほぼ同時期に塩基配列が決定された真核生物の空胞系に発現している V (vacuole) タイプの ATPase と 50% 程度の高い相同性を持つことが分かった [Gorgarten *et al.*, 1989]。このことは、古細菌には F ( $F_0F_1$ ) タイプではなく、V タイプの ATPase のみが存在するという説を産むきっかけを作る。

### 1.3 古細菌と真正細菌の分化と V/F タイプ ATPase の分化との関係

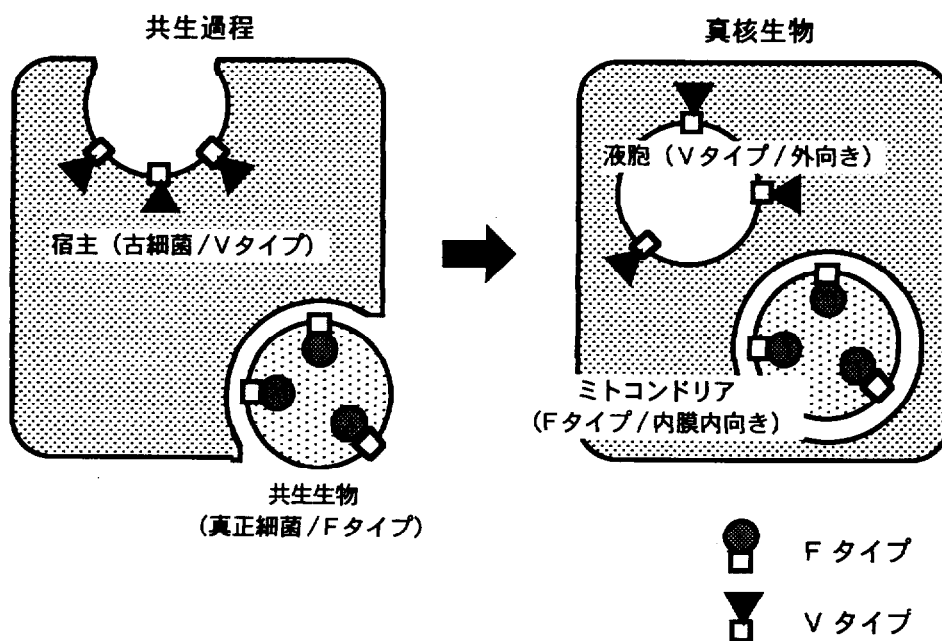
真正細菌には F タイプのみ、古細菌には V タイプのみが分布すると仮定すると、F、V 両タイプ ATPase に共通の始原タイプ ATPase を持つ始原生物が真正細菌と古細菌に分化、その後の進化を経て

真正細菌の ATPase は現在の F タイプに、古細菌の ATPase は現在の V タイプに変化したという仮説が成り立つ[Gogarten *et al.*, 1989] (図02)。



〈図02〉古細菌と真正細菌の分化と F/V タイプの分化時期が同じと考えたときの進化の仮説を模式か下もの。始原タイプの ATPase を継承した古細菌は進化の途中でそれを V タイプに変化させ、真正精動は F タイプに変化させたと考えられる。

この説は、Margulis らの細胞内共生による真核生物発生説[Margulis, 1981]をうまく説明するため適正なものと思われた。つまり、共生の際、宿主となる古細菌に、ミトコンドリアや葉緑体になる真正細菌が共生したと仮定すると、宿主のプラズマ膜の陥入により派生したと思われる空胞系の V タイプ、共生生物のプラズマ膜由来と思われるミトコンドリアなどの内膜に存在する F タイプの ATPase という具合に、分布もトポロジーも非常によく合致する (図03) [Nelson *et al.*, 1989]。

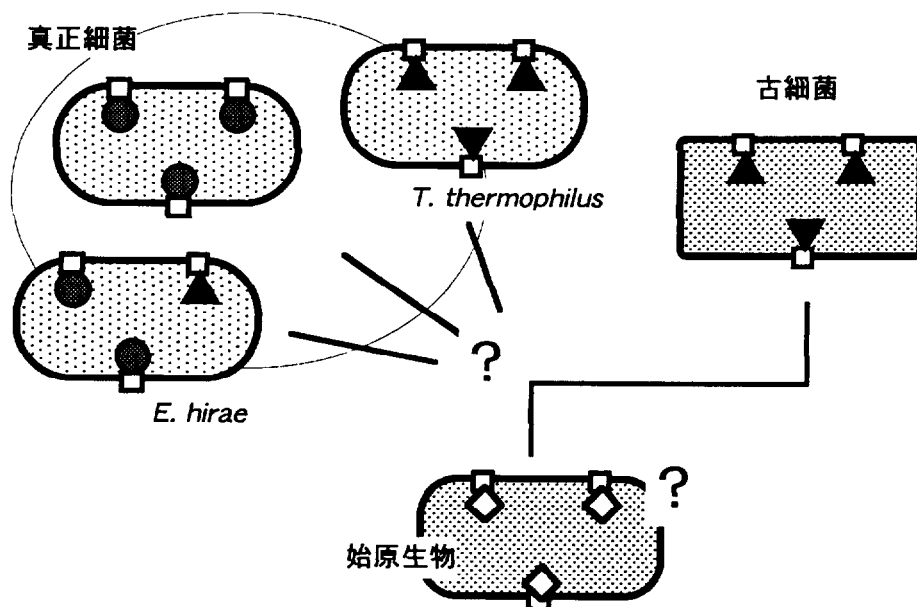


〈図03〉宿主の古細菌は V タイプ、共生生物の真正細菌は F タイプのみを保持するとし、真核生物が細胞内共生によって生

じたときの過程を模式化した図。空胞系は宿主のプラズマ膜が貫入して進化したもの、ミトコンドリアや葉緑体は共生生物がエンドサイトーシスなどで細胞内に取り込まれたものとする、現在の真核生物のVタイプおよびFタイプのトポロジーをよく説明できる。

#### 1.4 真正細菌におけるVタイプの発見と真正細菌におけるF/V両タイプの共存

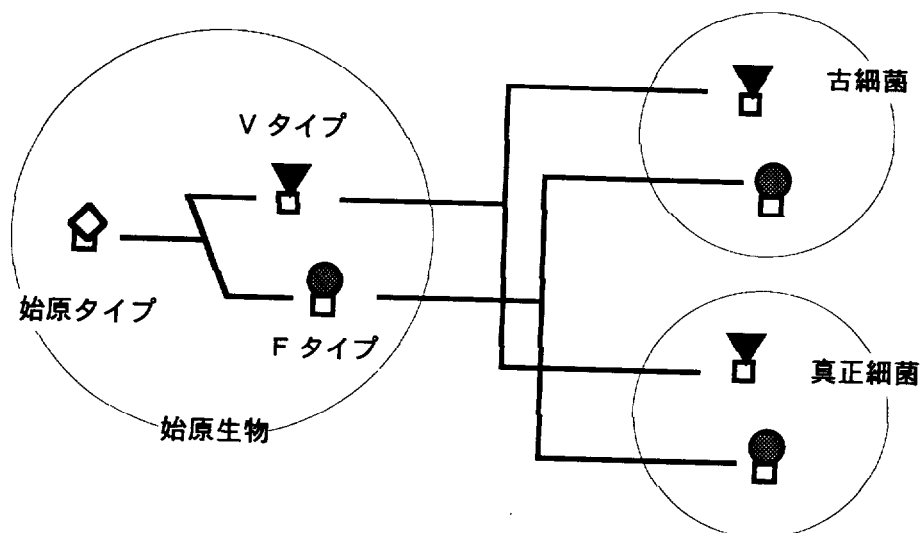
しかし、1990年、横山らにより、真正細菌に属する好熱菌、*Thermus thermophilus* より、明らかにVタイプに属すると思われるATPaseが単離される[Tsutsumi *et al.*, 1991; Yokoyama *et al.*, 1990, 1994]。この結果、前述のF/Vタイプの分化と真正細菌/古細菌の分化時期を同一とする仮説は修正を余儀なくされた(図04)。その後、柿沼らにより、すでにFタイプの存在が確認されている腸内球菌、*Enterococcus hirae* にもVタイプが存在することが報告された[Takase *et al.*, 1993, 1994; Kakinuma & Igarashi, 1994]。ここに至り、少なくともVタイプは古細菌に限らず真正細菌にも分布しうること、そして、同一バクテリア内に両タイプが共存することも可能であることが示唆された。



〈図04〉真正細菌にVタイプが見いだされたことにより不明確になった系統樹。真正細菌のVタイプが偶発的に古細菌のそれと同じようになったとは考えにくい。始原タイプの存在を仮定する必然性もなくなってしまう。

#### 1.5 本研究の目的

*T. thermophilus* におけるVタイプの発見をうけ、我々はF/Vタイプの分布に対する認識を改める必要を感じ、またATPase進化仮説の修正を行なうべく準備をすすめた。具体的には、始原タイプのATPaseがVおよびFタイプに分岐したのは、始原生物が古細菌と真正細菌に分岐するはるか以前で、すでに始原生物中にて分化は完了していたこと。その後、両タイプのATPaseは、真正細菌および古細菌に継承され、あるものは両タイプを維持し、あるものは片方を欠落させたという仮説である(図05)。



〈図05〉新しく考えられた仮説。始原生物内ですでに始原タイプは遺伝子重複などにより、FおよびVタイプに分化、その後、真正細菌/古細菌への分化を経て現在に至るとするもの。これらな、真正細菌におけるVタイプの存在も矛盾無く説明できる。

本研究はこの仮説に則り、1) まだF/Vの報告のないバクテリアより、いずれか、もしくは両方のタイプのATPaseの存在を確認すること、2) すでにF/VいずれかのATPaseが報告されているバクテリアについても、もう一方のタイプのATPaseが存在し得ないかを探索すること、を通じてより多くのATPaseのタイプ別分布に関する情報を集めることを当初の目的とした。また、F/Vの広範な分布を示すために不可欠であるがまだ報告されていない、古細菌のFタイプATPase遺伝子の発見を最終的な目標とした。その結果、メタン産生菌よりFタイプの遺伝子を発見することができたのでここに報告する [Sumi *et al.*, 1997]。

## 第2章 PCR法によるバクテリアのF/VタイプATPase遺伝子の確認

1990年初頭、より安価なサーマルサイクラーの普及により、PCR法を用いた遺伝子の単離が一般の研究室単位でも行なえるようになった。この系を応用し、F/Vタイプの遺伝子存在の有無を、DNA断片の増幅の有無を手がかりに手軽に判定できる系を構築した。その結果、いくつかの成果があがり有効性を示す一方で、PCR法一般にいわれるのと同じ問題もはらんでおり、この系の運用には注意が必要であることも判明した。

### 2.1 方法

#### 2.1.1 バクテリア

バクテリアは、生菌、集菌された凍結菌体、調製されたゲノムDNAのいずれかの形態で恵与されたものを使用した(表01)。*E. coli*、*P. denitrificans*、*T. aquaticus*は恵与された生菌を培養し、集菌後、ゲノムDNAを調製した。*P. modestum*、*M. barkeri*は凍結菌体よりゲノムDNAを調製した。

strain	gifted form	gifted from
<i>Escherichia coli</i> JM109	cell	
<i>Salmonella typhimurium</i> LT2	DNA	f
<i>Bacillus subtilis</i> NB22	DNA	e
<i>Staphylococcus aureus</i> 209P	DNA	f
<i>Paracoccus denitrificans</i>	cell	c
<i>Nitrosomonas europaea</i>	DNA	c
<i>Vibrio parahemolyticus</i> AQ3334	DNA	f
<i>Thiobacillus ferrooxidans</i> Fe1	DNA	c
<i>Propionigenium modestum</i>	DNA	b
<i>Enterococcus hirae</i> ATCC9790	DNA	g
<i>Thermus thermophilus</i> HB8	DNA	a
<i>Thermus thermophilus</i> HB27	DNA	a
<i>Thermusaquaticus</i>	cell	a
<i>Sulfolobus acidocaldarius</i> strain 7	DNA	
<i>Halobacterium salinarium</i> L-33	DNA	c
<i>Methanosarcina barkeri</i> DSM800	cell	d

a, 本学旧大島・有坂研究室; b, Lehrstuhl für Mikrobiologie, Technische Universität München, München, F.R.G.; c, 本学旧山中・副森研究室; d, 産業医科大学学生化学教室; e, 本学正田研究室; f, 岡山大学薬学部微生物薬品化学教室; g, 千葉大学薬学部臨床化学研究室。

(表01)

#### 2.1.2 培養

*E. coli* と *P. denitrificans* は、1.5% yeast extract、1.5% polypeptone、0.2% NaCl を含む培地を用い、37℃の好気的な条件下で培養した。*T. aquaticus* は、0.3% yeast extract、0.3% polypeptone、0.2% NaCl を含む培地で80℃で培養した。得られた菌体は、それぞれただちに2,000×g、4℃



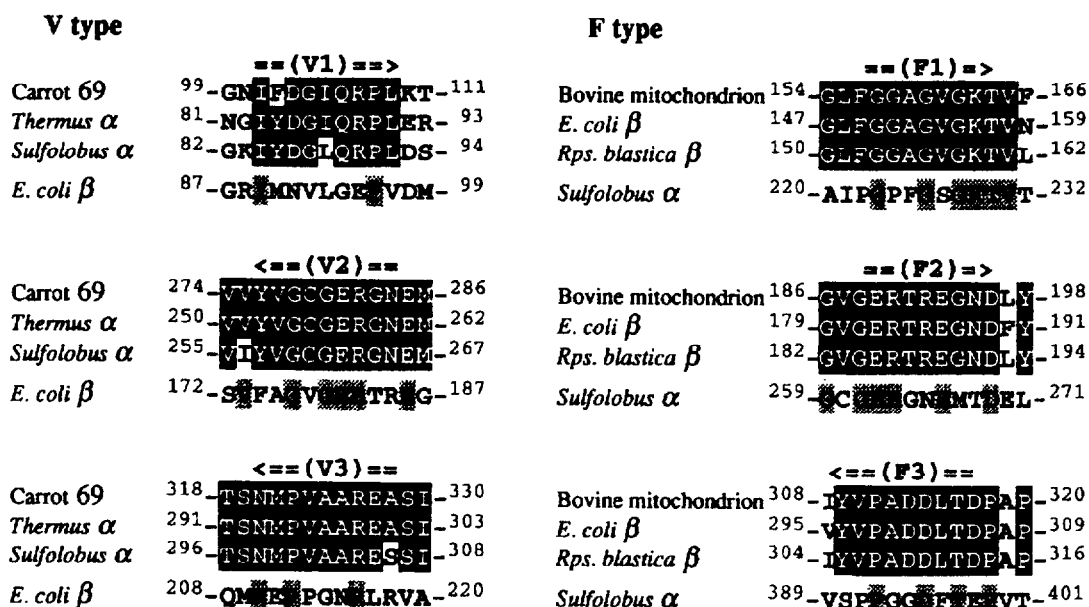
で集菌し、ゲノム DNA の調製に用いた。

### 2.1.3 ゲノム DNA 調製

*E. coli*, *P. denitrificans*, *T. aquaticus*, *M. barkeri* の菌体湿重量 10g を 25mL の TE バッファ (10 mM Tris-Cl pH 8.0, 0.1mM EDTA) によく懸濁し、10% SDS を 2.5mL 加えて菌体を溶解させた。溶解を確認した後、等量の TE 飽和フェノールを加え、30 分間穏やかに混和。6,000×g 10min 遠心したのち、上清水溶液部分を別の容器に移し、再び等量の TE 飽和フェノールを加え、15 分程度穏やかに混和。遠心分離までの同様操作を 3 度繰り返す、最後に得られた水溶液に無水エタノールを 3 倍量重層し穏やかに混和した際生じるゲノム DNA 沈殿をすくい上げ、70% エタノールで 2 度、無水エタノールで 2 度洗滌後、適量の TE バッファに溶解した。

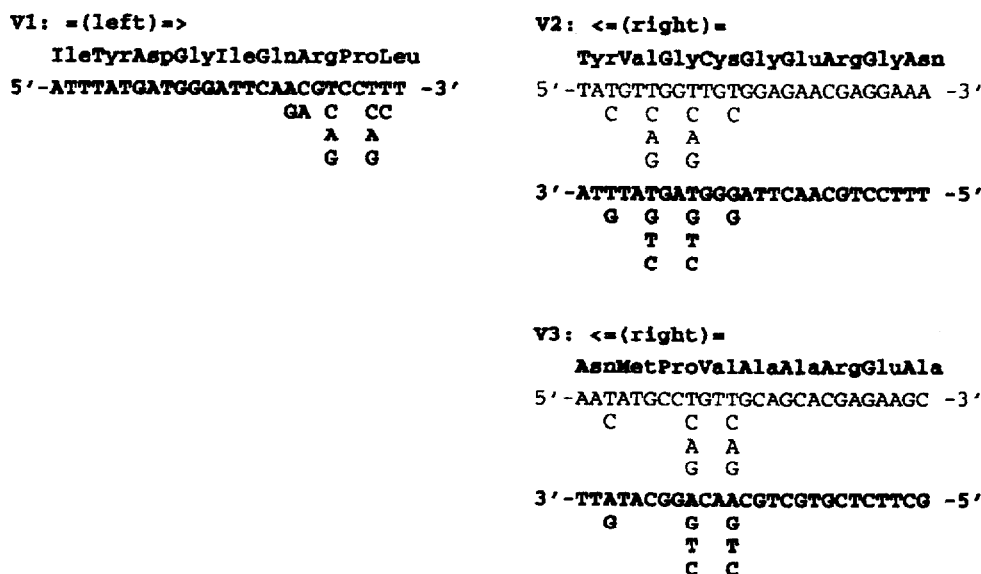
### 2.1.4 PCR 用プライマ合成

PCR 法で用いるオリゴヌクレオチド (プライマ) は、F/V それぞれのタイプの ATPase 触媒サブユニット (V タイプは  $\alpha$ 、F タイプは  $\beta$ ) のアミノ酸配列から非常に高い保存性を保つ領域を 3 カ所選びだし、これをもとに設計した (V タイプは V1、V2、V3、F タイプは F1、F2、F3 とそれぞれ名付ける; 図06)。鎖長は 30 塩基弱と、通常 PCR で使われるものより比較的長くし、特に伸長反応方向に位置する 3' 端の配列についてはすべてのコドンの組み合わせに対応できるように混合物として作成するなど、目標となる塩基配列が存在したときに取りこぼしのないよう配慮した (図07a、b)。



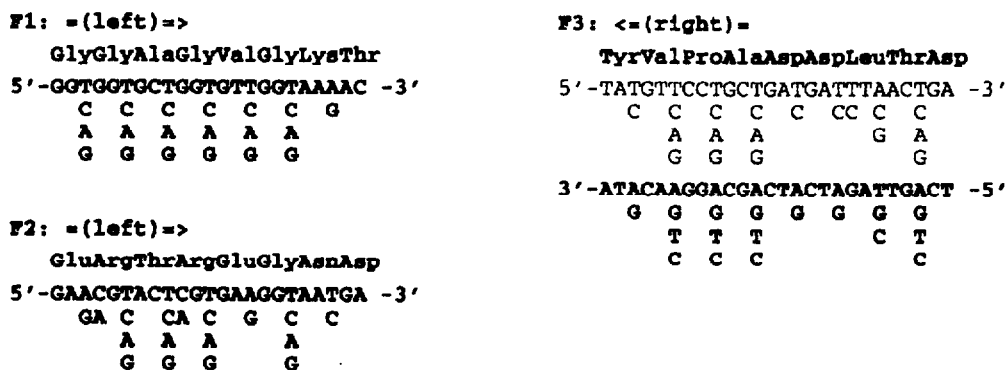
(図06) V1~3、F1~3 設計のもととなるアミノ酸配列

## V type



〈図07a〉 V1、V2、V3 の塩基配列

## F type



〈図07b〉 F1、F2、F3 の塩基配列

反応に用いるのは V1-V3、V1-V2、F1-F3、F2-F3 の組み合わせで、もし対応するタイプの遺伝子が存在すれば、それぞれ、669 bp、543 bp (*S. acidocaldairus* の場合)、492 bp、384 bp (*E. coli* の場合) 程度の長さをもつ DNA 断片の増幅が期待できる。

### 2.1.5 PCR

反応には、主に Taq ポリメラーゼ (BIOTECH) を用いた。反応液の総量は 100  $\mu$ L で、主要な成分は、0.2  $\mu$ g 程度のゲノム DNA (テンプレート)、各 100 pmol のプライマ (Vタイプ検査には V1-V3、Fタイプ検査には F1-F3)、10  $\mu$ L 10 $\times$ 反応溶液 (キット付属)、1.5 mM MgCl<sub>2</sub>、各 1.25mM デオキシヌクレオチド (dATP、dCTP、dGTP、dTTP)、2.5 単位の Taq ポリメラーゼである。混合反応液には、反応途中での水分蒸発を防ぐためにミネラルオイルを適量滴下した。増幅反応はアステッ

ク社製サーマルサイクラ (PC-700) を用いて行なった。温度調節は1サイクルにつき、94℃、55℃、74℃をおのおの1分ずつ、合計30サイクルのプログラムを利用した。増幅の確認は1%アガロースゲル電気泳動の後、エチジウムブロマイド染色を行ない紫外線下で写真撮影した。

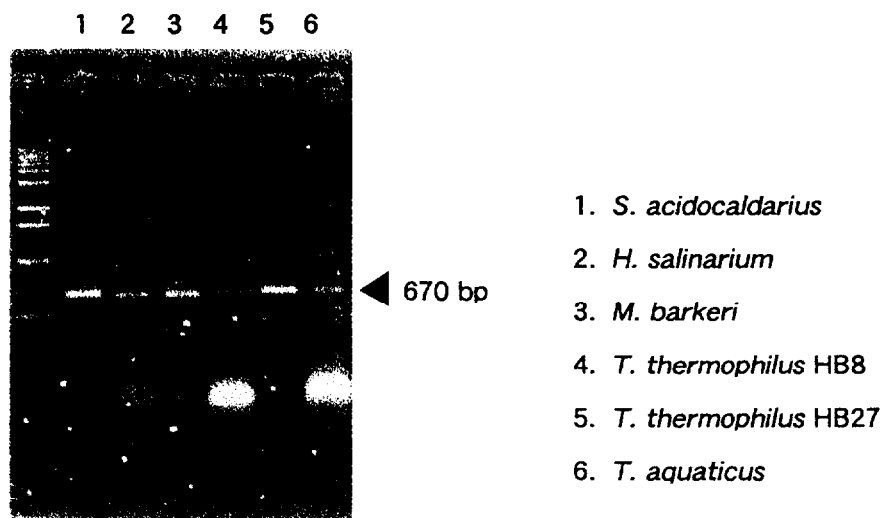
#### 2.1.6 確認のためのPCR

2回目の確認のための増幅に備え、1回目の増幅産物はゲルから切り出し精製の後、TEバッファに溶解させた。PCRは、テンプレートとして切り出し精製した増幅産物(約数ng)、プライマとしてV1-V2もしくはF2-F3を用いることを除いて、1回目の増幅と同じである。

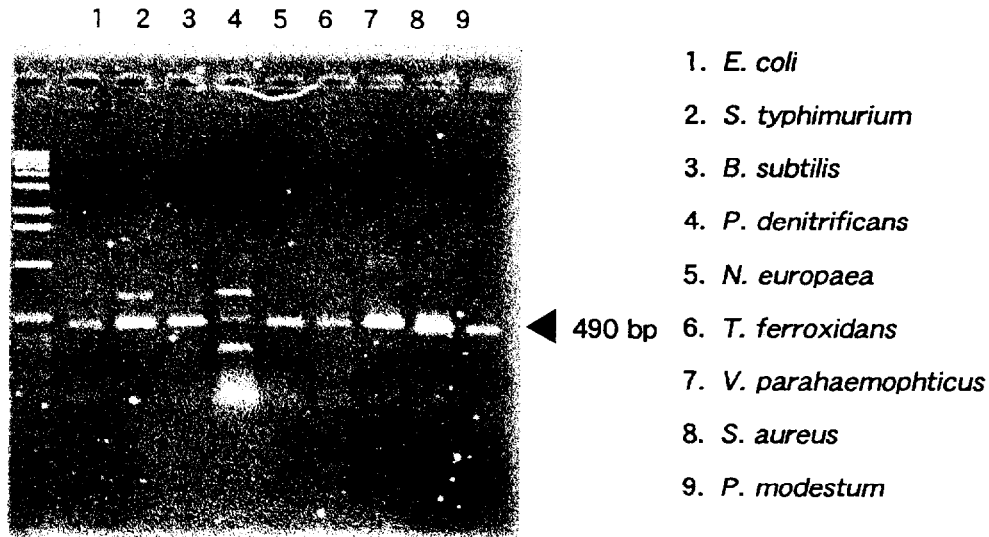
## 2.2 結果

### 2.2.1 既知のVタイプ、Fタイプ遺伝子の確認

従来の報告よりVタイプが存在することが分かっている古細菌および *Thermus* 属についてVタイプのテストを、Fタイプが存在することが予想される、もしくは知られている *Thermus* 属以外の真正細菌についてはFタイプのテストを、それぞれポジティブコントロール実験として行なったところ良好な結果が得られた(図8、9)。



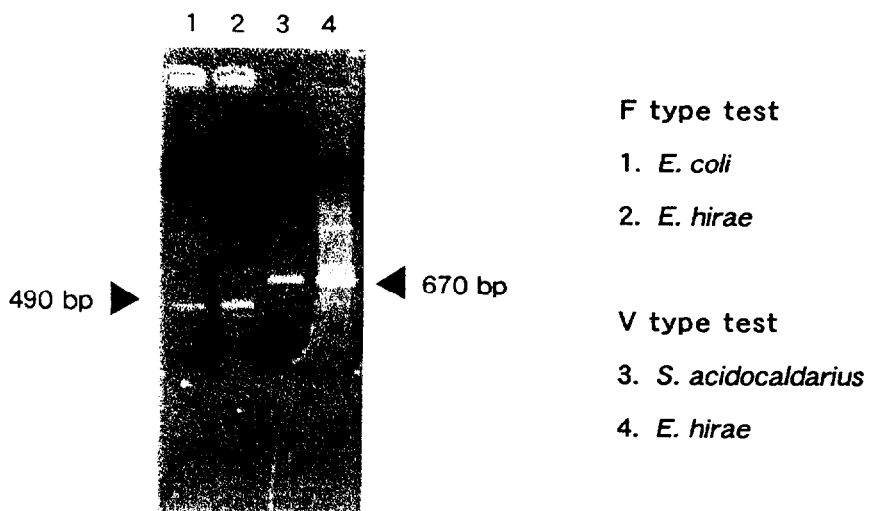
〈図08〉Vタイプテストのポジティブ実験。Vタイプ遺伝子が存在すれば、黒三角で示した670 bp付近にバンドが生じる。



〈図9〉 Fタイプテストのポジティブ実験。Fタイプ遺伝子が存在すれば、黒三角で示した 490 bp 付近にバンドが生じる。

### 2.2.2 *E. hirae* における V/Fタイプ遺伝子の確認

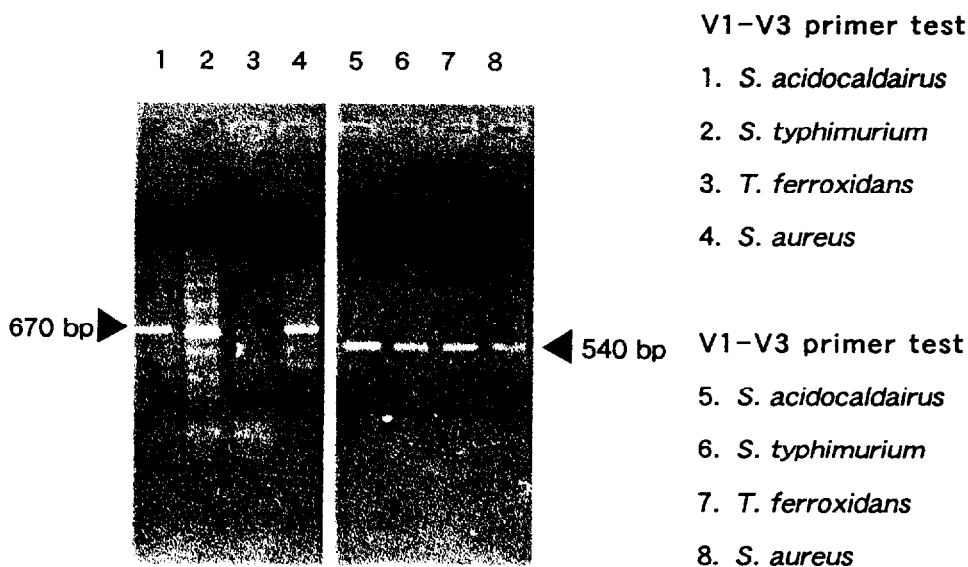
*E. hirae* は、呼吸鎖を持たず膜電位の維持には通常Fタイプの ATP 水解反応を用いている。しかしこの酵素の至適 pH は酸性よりであること、あるタンパク質を発現できない変異株はアルカリ性条件下では生育できないこと、などから、アルカリ条件下では別のナトリウムイオン輸送性 ATPase を発現して  $\Delta \mu \text{Na}^+$  により膜電位を維持している可能性が柿沼らにより示唆された[Kakinuma *et al.*, 1990]。このナトリウムイオン輸送性 ATPase は精製標品のサブユニット組成や阻害剤に対する挙動（アジドで阻害されず、硝酸イオンで阻害される）などからVタイプであると予想されたが当初は遺伝子による確認はされていなかった。柿沼らにより恵与された *E. hirae* のゲノム DNA を用いてFおよびVタイプのテストを行なったところ、両タイプにおいて増幅が見られた（図10）。



〈図10〉すでにタンパク質レベルで存在が確認されている *E. hirae* における、F/Vタイプのテスト。Vタイプのポジティブ標品には *S. acidocaldarius*、Fタイプには *E. coli* を用いた。Fタイプ、Vタイプ遺伝子の存在を示す、それぞれ 490、670 bp のバンドが確認できる。

### 2.2.3 *Thermus* 属以外の真正細菌におけるVタイプ遺伝子の確認

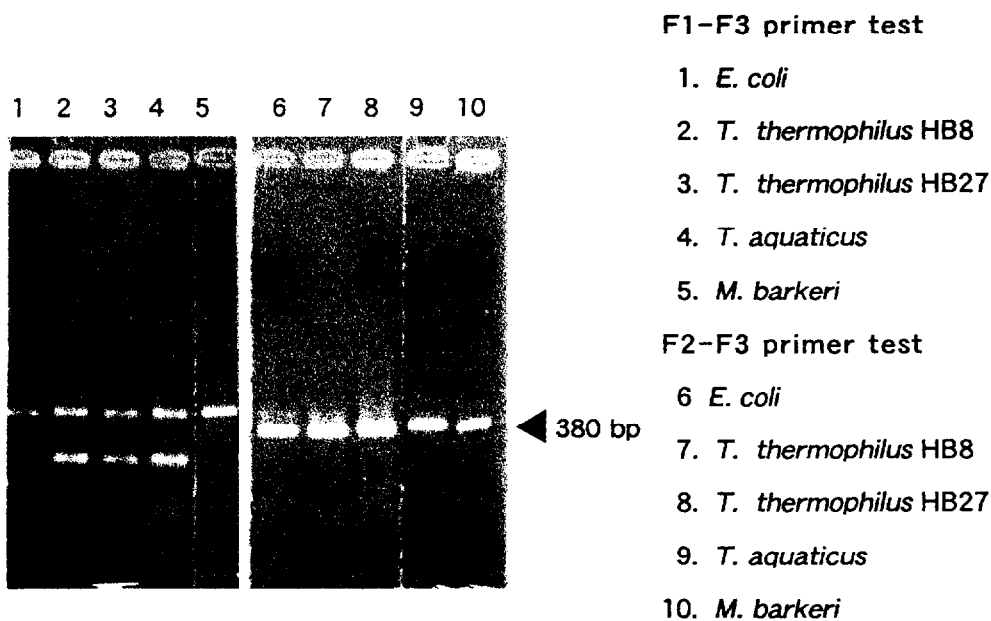
すでにVタイプの報告されている *T. thermophilus* HB8 とその同属バクテリアを除いた真正細菌においてVタイプ遺伝子が存在する可能性がないか検査を行なった。その結果、*S. aureus*、*S. typhimurium*、*T. ferrooxidans* において 690 bp 付近の増幅が確認された（図11、レーン2～4）。それぞれの増幅 DNA 断片を切り出し精製し、テンプレートに用いた2回目のPCR（V1-V2 プライマを使用）でも 540bp 付近に増幅が確認された（図11、レーン6～8）。



（図11）真正細菌のVタイプテスト。コントロール標品には *S. acidocaldarius* を使用した（レーン1、5）。1回目の増幅では非特異的バンドがあって不明瞭でも、2回目の増幅（レーン6～8）では単一のバンドで確認できる。

### 2.2.4 古細菌および *Thermus* 属におけるFタイプ遺伝子の確認

古細菌および *Thermus* 属ではVタイプの存在しか示されていない。まず、真正細菌である *Thermus* 属においては、*E. hirae* のようにV/F両方の遺伝子を保持する可能性が高い。このような予想の通り、*Thermus* 属からはFタイプ遺伝子由来と思われる増幅が確認できた。さらに古細菌からもメタン産生菌、*M. barkeri* と好塩菌、*H. salinarium* よりF由来の増幅が確認できた。いずれの増幅も切り出し精製し、それをテンプレートとして用いた2回目の増幅（F2-F3 プライマ使用）においても 380 bp 付近の単一のバンドが確認された（図12）。



〈図12〉 *Thermus* および古細菌におけるFタイプテスト。 *H. salinarium* の結果は示していない。

### 2.2.5 増幅産物の塩基配列確認

新規に確認された真正細菌のVタイプのうち、*S. aureus* のものを pUC119 に組み込んでサブクローニングし、その塩基配列を決定したところ、増幅が期待される領域のアミノ酸配列と相同性を持つコドンの並びを確認できた。新規に確認された古細菌のFタイプ、*M. barkeri* および *H. salinarium* より得られた増幅断片も同様に塩基配列決定を行なったところ、*M. barkeri* ではやはり期待されるアミノ酸配列と相同性を持つコドンの並びが確認されたが、*H. salinarium* では、偶然、同じ長さで2回目のPCRにおいても同じ挙動をする、まったく関係のないゲノムDNA断片が増幅されていることが判明した。

## 2.3 考察

### 2.3.1 検出系の長所と欠点、考えるその改善策

PCRで未知のDNA断片を増幅し単離する場合、増幅したいDNA断片の両端の塩基配列、20 bp程度が分かればよい。ある遺伝子が、たがいにそれほど離れていない（理想的には1~2 kbp程度の距離を持つ）2カ所以上の比較的高い保存性領域保持する場合、そこから予想されるDNA配列よりプライマを設計しPCRを行なうことで、その遺伝子の一部を単離することができる。逆に、対象となる遺伝子の保存領域において高度な保存性が期待される場合、PCRによるDNA断片増幅の有無により、その遺伝子の存在の有無を確認することができるはずである。FおよびVタイプATPaseはアミノ酸レベルで高い相同性を保持する領域をそれぞれ数カ所ずつ保持するため、この系によりその遺伝子の存在の有無を確認することが容易であると考えた。

間接的に酵素、もしくはその遺伝子の存在を確認するには、酵素独自の活性を測定する方法、プローブに抗体を用いるウエスタンブロット解析、プローブにDNA断片を用いるノザン/サザンブロット解

析などが考えられるが、いずれも状態のよい、比較的まとまった量の菌体や核酸などを必要とする上、標品調製および検出作業が煩雑で、多種に及ぶ標品の検定には向かない。それに対し、PCR法による検出システムでは、微量のゲノム DNA を手に入れることができれば、あとは必要な試薬を混合し反応を行なうだけで、簡便な検出が可能な点でアドバンテージを持つ。

一方で、PCR では、コンタミネーションやプライマの非特異的吸着によるアーチファクトの可能性に加え、テンプレートに用いるゲノム DNA の GC 含有量やそれに関連する高次構造による増幅障害などの可能性を考慮する必要がある。コンタミネーションについては、ゲノム DNA 標品への混入を除けば、反応系のネガティブコントロール実験（テンプレート DNA を加えずに反応を行なう）を定期的に行なうことで誤った判断をくだす危険性を排除できる。ゲノム DNA 標品への異物混入の検出は難しいが、小分けするなどして入手時のクオリティを保持する配慮は必要である。

V/Fタイプの触媒サブユニットには、25% 程度ではあるが相互に類似性がある。Fタイプの検出プライマでVタイプを、逆にVタイプの検出プライマでFタイプを増幅してしまわないようにプライマの設計においては可能な限り共通の保存領域は避けた。また、VにはあってFにはない挿入配列を挟んでプライマの設計することで、誤った増幅が行なわれても、その長さで判断が付くように工夫した。他方、非特異的増幅による判断ミスを避けるため、第3のプライマを用意し、1回目の増幅断片をテンプレートに2回目の増幅を確認し、都合3カ所の類似性を確認するシステムにした。

このチェック機構はおおかたうまく機能したが、残念ながら *H. salinarium* でのFタイプの増幅をサブクローン化し塩基配列を決定したところ、まったく関係のない領域が偶然、同じ長さで増幅していたことが分かった。この断片は、内部に第3のプライマとの相同領域も持つため、その塩基配列を決定するまで正常な増幅と区別することは難しい。このようなアーチファクトを産んだ背景には、高いGC含有のテンプレートと混合プライマを用いた際のPCRの難しさがある。

一般にGC含有量の高いDNA断片のPCR増幅は完全一致のプライマを用いても、その設計によっては増幅がうまくゆかず、対処に苦慮することがある。近年、超好熱菌より単離されたポリメラーゼを使用した様々なアプリケーションが提供されており、以前よりは状況はよくなったが、いずれも一般的な反応系において増幅が見られない際に試してみる価値はあるものの、決定打にはなり得ていない。従って、本検出系においても、増幅が無いことがすなわち遺伝子の存在を否定する結論には結びつかない。簡便な検出系という主旨からは外れるが、遺伝子の存在を否定する決定的な手段は、最近積極的に進められている全ゲノム塩基配列決定のみに限られる。

実際、本検出系においてもGC含量の問題は既知の遺伝子の増幅の際もたいへんな問題となった。GC含量が非常に高い（65%以上、コドン最終塩基では90%近く）*H. salinarium*、*Thermus* 属においては安定な増幅を得るのはいかなる状況においても難しい。先に触れた超好熱菌由来のポリメラーゼ

を用いたアプリケーションをはじめ、2価カチオン量、プライマ量のスクリーニング、テンプレートの断片化、アルカリ変性処理などを用いた一本鎖化など多くの手法を試したが、安定な増幅は期待できなかった。唯一、特異的な増幅が見られたのは、サーマルサイクラ不調で温度調節がプログラム通りにゆかなかったときであるが、残念ながらこれを正常なマシンで再現することはできなかった。

他方、やはり GC 含量の高い好塩菌の一種、*Haloferax volcanii* を用い、報告されている V タイプの塩基配列[Steinert et al., 1995]より、本検出系で設計した V タイプの検出用プライマとまったく同じ場所のプライマを作製し PCR 反応を行なったところ、安定な増幅を得ることができた。本検出系に用いた混合プライマの中には、このとき用いた完全一致のプライマと同じ配列のものも存在しているはずだが、安定した増幅は見られない。*H. volcanii* において増幅に十分な量の完全一致のプライマと、本検出系のプライマを混合して増幅したところ、混合比率によっては増幅が見られなくなることが分かった。これは反応条件によっては、多く存在する増幅には使用されないプライマが、極わずかに存在する適正なプライマによる増幅反応を阻害している可能性を示唆する。

本検出系は現在のままでも、中～低度の GC 含量のゲノム DNA を持つバクテリアには有効に機能する。高度に GC に富んだ配列を有する標品用に改良するとすれば、コドン第 3 塩基を GC に限った混合プライマを作製する、より適正な温度コントロールプログラムを考案するなどの工夫が必要であろう。

### 2.3.2 新奇 V/F タイプ遺伝子存在の可能性

前述の通り、問題は多く残されているが、現時点では本検出系により真正細菌で 3 種の V タイプ、古細菌および *Thermus* 属でつごう 4 種の新奇 F タイプ遺伝子の存在が示唆された。本研究で取り扱った真正細菌の数は現存しうるバクテリアの種類からすれば微々たるものであるにも関わらず、従来報告されている *T. thermophilus*、*E. hirae* 以外にも、塩基配列確認の済んだ *S. aureus*、そして他の 2 種の真正細菌に V タイプが存在する可能性が示唆されたことは、V タイプの真正細菌における分布を裏付ける上でも重要な情報を提供する。

また、古細菌の *M. barkeri* において F タイプ遺伝子の存在が確認されたことは、たいへんな成果である。後に分かったことだが、*M. barkeri* と大変近縁の *M. mazei* でのサザンプロット解析の結果[Müller, 私信]と、近年全ゲノム解析を終えた、*Methanococcus janaschii*、*Methanobacterium thermoautotrophicum*、*Archaeoglobus fulgidus* という 3 種の古細菌における類似性検索[<http://www.tigr.org/tdb/mdb/mdb.html>]の結果から、いずれも現時点では F タイプの存在を示唆する情報は得られていない。メタン産生菌として *M. barkeri* が入手可能であったこと、そのゲノム DNA の GC 含量が 40% 程度[Boone et al., 1993]と低かったため本系標品に向いていたことが幸いしたものの、古細菌において限られた種にしか存在しない F タイプの遺伝子を検出できたことは、本検出系の有効性を示すよい例となった。



### 第3章 古細菌からのFタイプATPase 遺伝子の単離と塩基配列決定

PCR法を用いた検出系により、古細菌に属するメタン産生菌、*M. barkeri* にFタイプ遺伝子が存在する可能性が示唆された。この増幅断片をプローブに古細菌のFタイプ遺伝子クラスターを単離し、その塩基配列を決定した。すでに、このメタン産生菌にはVタイプの遺伝子が存在することが報告されていたが、それとはまったく異なるFタイプの遺伝子であることが明らかにされた。この古細菌のFタイプ遺伝子には、 $\delta$ サブユニットをコードする遺伝子が存在しない、 $b$ サブユニットが異常に大きい、 $\gamma$ サブユニットがアミノ末端とカルボキシル末端とで別々の遺伝子にコードされている、などの特徴を持つ。残念ながらこの遺伝子の産物であるタンパク質や mRNA の検出は失敗に終わったため、本遺伝子が機能を持つかどうかは現時点では明らかでない。

#### 3.1 方法

##### 3.1.1 バクテリアとその培養

*M. barkeri* MS (基準株、DSM800) は、共同研究者の古賀らが Deutsche Sammlung für Mikroorganismen und Zellkulturen (DSM) より得、継代培養しているものを使用した。培養は絶対嫌気的条件下 (気層を 80% 窒素+20% 二酸化炭素にて置換) で行ない、0.5% メタノールを含む DSM120 (DSM が推奨するメタン産生菌用培地; pH 6.85、ナトリウムイオン濃度 約 40mM) を培地に用いた。pH の異なる培地はカリウムリン酸バッファ、ナトリウムイオン濃度の異なる培地は加える NaCl 量で調節し作製した。

##### 3.1.2 プローブのラベルと遺伝子の単離、塩基配列決定

PCR 増幅産物はアガロースゲル電気泳動後、切り出し精製を行ない、クローニングベクター pBlue-script II SK<sup>-</sup> (Stratagene 社) の *EcoRV* サイトに挿入し、サブクローン化した。このプラスミドを大腸菌 JM109 株に導入のち大量に得、*EcoRI* および *HindIII* により切断して切り出した DNA 断片をプローブとして使用した。プローブのラベルおよびサザン/コロニーハイブリダイゼーションは ECL ランダムプライマラベルシステム (Amersham 社) を使用した。

*Bam*HI で *M. barkeri* ゲノム DNA を切断し、8 kbp 付近の断片をアガロースゲル電気泳動と切り出し精製法を用いて回収した後、pBluescript II SK<sup>-</sup> ベクターの *Bam*HI サイトに組み込んだものをコロニーハイブリダイゼーションのライブラリに使用した。ポジティブクローンを選び、培養してプラスミドを得たのち、*Bam*HI で切り出されるベクター以外の挿入断片を切り出し精製法で回収した。挿入断片は超音波により適当な長さに破碎した後、600 bp 付近の断片を平滑化、クローニングベクター pUC18 の *Sma*I サイトへのライゲーション作業を経て大腸菌に導入した。無作為に選んだコロニーからプラスミドを精製し、シークエンス反応のテンプレートとして用いた。

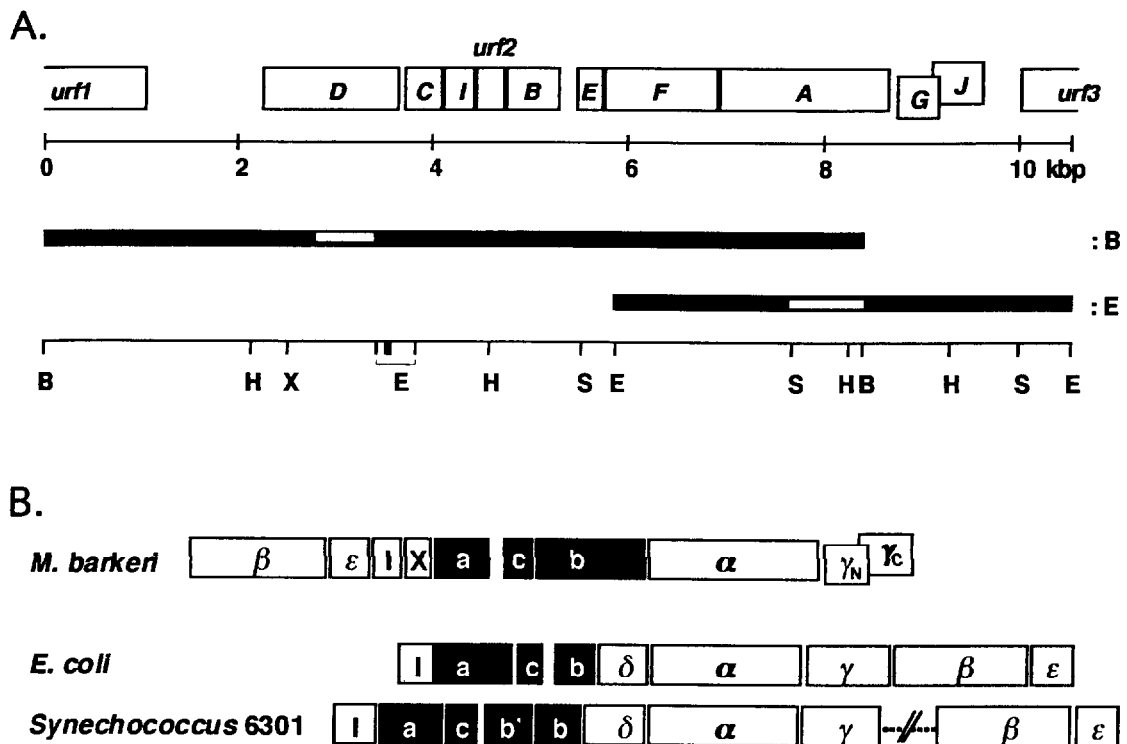
### 3.1.3 塩基配列の解析

シーケンス反応の結果は、随時 ATSQ ソフトウェア（自動塩基配列結合ソフト；Software Development 社）による処理を行ない、DNA 両方向からの配列決定を確認して作業を終えた。遺伝子の読み枠予測は GENETYX-Mac（Software Development 社）の機能を用いた。既知の遺伝子との類似性検索には、および、アライメントの作成は Genetics Computer Group（GCG；Wisconsin 大学）の提供するパッケージを使用した。

## 3.2 結果

### 3.2.1 遺伝子の単離

PCR で得られた 380 bp の DNA 断片（図13、黒い太線B上の白抜き部分）は、サザンブロット解析の結果、*M. barkeri* の *Bam*HI 切断産物の 8 kbp 付近で交叉した（図14）。約 100 クローンの 8 kbp 付近の *Bam*HI 切断産物で構成されるプラスミドライブラリーをスクリーニングしたところ、1 個のポジティブクローンを得ることができた（図13、黒い太線B）。塩基配列を決定したところ、このクローンには  $\alpha$  のカルボキシル末端と  $\gamma$ 、そして  $\delta$  サブユニットをコードする遺伝子を除き、F タイプ構成成分遺伝子のほとんどすべてが含まれていることが分かった。そこで、この *Bam*HI 断片の 3' 端に位置する約 800 bp の *Sac*I 断片（図13、E 上の白抜き）を切り出し精製し、新たなプローブとして染色体歩行を行なった。ハイブリダイゼーションの結果、*Eco*RI 切断産物で構成されるライブラリーに残りの遺伝子をすべて網羅するのに十分な長さ（約 4 kbp）の交叉が確認されたので、4 kbp 程度の長さを持つ *Eco*RI ライブラリーを作成、スクリーニングを行なった。約 400 クローン中より、4 つのポジティブクローンを単離した。これらの両端の塩基配列を決定したところ、同一かつ先に単離した *Bam*HI 断片と重複を持つ目指すクローンであると判断された（図13、黒い太線E）。このクローンについても *Bam*HI 断片と同様な方法でシーケンスを行なった。塩基配列決定の結果、残りの *Eco*RI 断片には予想通り、残りの F タイプ遺伝子および、転写終了領域と思われる配列を含んでいることが分かった。得られた両断片の塩基配列は、GenBank データバンクに AF028006 として登録した。



(図13) *M. barkeri* の *atp* オペロンと制限酵素サイト地図。A. *urf* は機能不明の読み枠、白枠内のアルファベットA、B、C、D、F、G、I、Jは遺伝子名を示す。太線B、Eはそれぞれ単離した8 kbp *Bam*HI 断片、4 kbp *Eco*RI 断片。太線内の白抜きはそれぞれの断片を単離するのに使用したプローブの位置を示す。B, *Bam*HI; E, *Eco*RI; H, *Hind*III; S, *Sac*I; X, *Xba*I。B. *M. barkeri*、*E. coli*、*Synechococcus* 6301 の *atp* (*unc*) オペロン内の遺伝子の並び。枠内のアルファベットおよびギリシャ文字はサブユニット名称を表わす。I は 14k タンパク質、X は *urf2* 産物、 $\gamma_N$  および  $\gamma_C$  はそれぞれ  $\gamma$  サブユニットのアミノ末端側、カルボキシル末端側との相同性を持つ遺伝子産物であることを示す。黒いボックスは  $F_0$  のサブユニット。

### 3.2.2 *M. barkeri* の F タイプ遺伝子の構成

単離した2つの DNA 断片の全塩基配列を決定した結果、その GC 含量は 43% と文献値[Boone *et al.* 1993]とほぼ一致した。これらのクローンには、10 個の遺伝子と思われる読み枠が存在し、それぞれ順に、*atpD*、*-C*、*-I*、*urf2*、*atpB*、*-E*、*-F*、*-A*、*-G*、*-J* と名付けた。*urf* は未知の読み枠の意で、完全ではないが得られたクローンの 5'端と 3'端にそれぞれ読み枠らしき配列 (仮に *urf1*、*urf3*) を見いだせたため番号は2となっている。その他の遺伝子の頭についている *atp* は初期に *unc* と名付けられそのまま定着した大腸菌を除く F タイプのオペロンの一般名称である。*atpJ* 以外の各サブユニットをコードすると予想される遺伝子の名前 (アルファベット) は慣例に従った (表02)。読み枠予想は、各フレームにおける終止コドン (TAA、TAG、TGA) の出現位置を視覚化し、開始コドン (ATG) との組み合わせで行なった (図14)





Gene	Sequence around RBS and start codon
<i>atpD</i>	CCAGCTTGGGGGAGCTGTTAAAATGGTAGATCGAAGTCG
<i>atpC</i>	CATGAATTCAGGACTCATGAATTCAGGACTCATGAATTC
<i>atpI</i>	GCCGATTAGCGGAGTTTCAGAATGTCCTGACAAACCATCA
<i>urf2</i>	AAGAGGAACAGGAGGATTACAATGAATGAAATATTGAAT
<i>atpB</i>	AATTGACAAAGGAGGCCAGTAATGCGACTTAGTCCTGAT
<i>atpE</i>	AACAACGGAAGGAAGCACTATGGCTTTGGACACCTATAT
<i>atpF</i>	GAAATTGAGAGGAAAGTAGACTATGCTGATTGATTGGTT
<i>atpA</i>	AGTCCTAATCTGTGGCCAAGATGAGATTAATCCCAAACC
<i>atpG</i>	CGGAGACAAAGGAAAAATCATGACCGAAACCACTCAGAG
<i>atpJ</i>	AGCTTTGCCAGGTAAAGCTCAGGTATGGGCTGTAGGGGA

(表03) 各 *atp* 遺伝子の開始部位とそれに対応するリボソーム結合部位。下線が開始コドン、太字が予想されるリボソーム結合部位。

### A.

```

GACGAAGCAATGGCCAAAAAATCAAGCCGGGAAAAATCATGAATTCAGGACTCATGAATT
                                RBS      M N S G L M N S
CAGGACTCATGAATTCAGGACTCATGAATCTTACGATTCCTTCTGCCTTTCCAGGTCCTTG
G L M N S G L M N S G L T I L L P F Q
(AtpC...)

```

### B.

```

(...AtpG)                                (AtpJ start)
TTATGCGGTCAGGGAACIGAAAGCTTIGCCAGGIAAAGCICAGGIATGGGCIGTAGGGGA
Y A V R E L K A L P G K A Q V W A V G E
                                M G C R G

GCTTGTIATTCACGCCTIGGAAGATGAAGGCTTIGCCGCTIATIGGACIIIIATAATIGGCCA
I V I H A W K M K A C R I L L D F I M C Q
A C Y S R L E D E G L P L I G L Y N V P

AAITCGGTTAAGGCAATTACCCCGCTTATAGCGCAAATICTTIGIGGAAAATGAAAAATTA
I R I R Q L P R L * (AtpGend)
N S V K A I T P L I A Q I L V E N E K L
(AtpJ...)

```

(図16) *M. barkeri atp* オペロンの特殊な構造。A. AtpC のアミノ末端は繰り返し構造を持つ。繰り返し単位を下線で、繰り返し単位中のリボソーム結合部位および開始コドンは太字で示した。B. γサブユニットをコードすると思われる遺伝子のオーバーラップ。AtpG 終了手前 110 bp から次の AtpJ のコードが始まっている。下線がオーバーラップ領域。太字は AtpJ 開始コドン。

### 3.2.3 他のFタイプ遺伝子との類似性比較

各遺伝子の予想されるアミノ酸配列は、*urf2*を除き、すでに報告されているFタイプのサブユニットのアミノ酸配列と何らかの類似性を持っていた。類似性検索の結果、AtpD (*atpD* から予想される翻訳産物; 469aa)、AtpC (144aa)、AtpB (190aa)、AtpE (91aa)、AtpA (588aa) はそれぞれ、す

で報告済みのFタイプの $\beta$  (459aa; *E. coli*の場合。以下同様)、 $\varepsilon$  (138aa)、 $a$  (271aa)、 $c$  (79aa)、 $\alpha$  (513aa) サブユニットと類似性を持っていた[GenBank M13704, L42023, MTU64318, X99078]。AtpF (386aa) は、その前半の配列(約120aa)でbサブユニット(156aa)との類似性を持つ[GenBank X99078]が、後半には他のどの遺伝子産物とも類似性を持たない余計な長いペプチド鎖が付加されていた。

ほとんどすべてのバクテリアの *atp* オペロンには、酵素の構成成分にならない約 14 k Da のタンパク質をコードする遺伝子がある。このタンパク質の機能はまだ分かっていない。*M. barkeri* の *atp* オペロンにも AtpI (112aa) をコードする遺伝子が見つかった。当初、類似性検索の結果では、AtpI が何に類似性を持つかが分からなかったが、既知の 14 kDa タンパク質に AtpI を加えてアライメントしたところ、*Rhodospirillum rubrum* の 14 kDa タンパク質[SWISS-PROT, P15011]と 92aa でオーバーラップしそのうち 24% 程度の相同性を持つことが分かった。

一見すると、*M. barkeri* のオペロン内の遺伝子の並びは、*E. coli* [Downie *et. al.*, 1979, 1981]などがそうで、一般に典型的と考えられている遺伝子の並びである  $I - a - c - b - \delta - \alpha - \gamma - \beta - \varepsilon$  とはかけ離れているものに見えるかも知れない。しかし、未知および  $F_0$  の構成成分である  $a$ 、 $c$ 、 $b$  サブユニットをコードする遺伝子クラスタ、 $\alpha - \gamma$  のクラスタ、 $\beta - \varepsilon$  のクラスタに分けて考えると、 $\delta$  がいないこと、 $\gamma$  が2つの遺伝子に分離している点などを除けば、 $\beta - \varepsilon$  クラスタがオペロンの先頭に移動しているだけで、さほど特異的な配列ではなくなる。 $\beta - \varepsilon$  クラスタがオペロンから分離し別の遺伝子座に位置したり、別の転写単位を形成することは、*Synechococcus* 6301 の *atp* 遺伝子などですでに報告されている(図13、B) [Cozens and Walker, 1987]。

### 3.2.4 主要サブユニットの保存性における特徴

Fタイプにおいて、 $\beta$  サブユニットの一次構造が種を越えて高度に保存されている点(60%以上が同一アミノ酸残基)はその特性のひとつと言っても過言ではない。ところが、*M. barkeri* の  $\beta$  サブユニットはどの種の  $\beta$  サブユニット対しても 50% 程度と比較的低い相同性しか持たない(図17)。 $\alpha$  についても同様のことが言える。しかし、酵素活性に重要なヌクレオチド結合モチーフ(別名 Walker 配列; モチーフ A、GXGXXGKT およびモチーフ B、ZZZZD; X は任意の、Z は疎水性のアミノ酸残基) [Higgins *et al.*, 1986; Walker *et al.*, 1982; Fry *et al.*, 1986] は  $\alpha$ 、 $\beta$  ともに保存されていたり、 $\beta$  の活性中心と見られる残基(ウシ心筋ミトコンドリアFタイプ $\beta$ では188番目のグルタミン酸)が保存されるなど、機能に必要なと思われる残基は変化していない。このことは、この遺伝子がATPaseの構成成分として機能している可能性を示唆する。 $\alpha$  サブユニットには、アミノ末端に 20aa、カルボキシル末端に 60aa 程度の他の  $\alpha$  には見られない余計なペプチド鎖が付加している(図17、パネルA)。

A: alpha subunits

			10	20	30	40	50	60	70																																																																
maize	chl		-----	-----	MATLRVDEINKILRE	REEDYNRKVGIE	NI	GRVVQVGGD	IARVIGLGE																																																																
maize	mit		-----	-----	MEFSPRAEELTTL	LES	MINFYTNL	KVD	EIGRVVSVCGDIARVYGLNE																																																																
bovine	mit		-----	-----	EKTGTAEVSSIL	EEFTLGADTSVD	LEETGRVLS	IGDGIARVHG	LRN																																																																
<i>E. coli</i>			-----	-----	MQLNS	TEISELIKQ	FA	GENYV	SEAHNEETIVSVSDGVTIRIHGLAD																																																																
<i>M. barkeri</i>			MRLIPNQSPHK	EPVNMETKSLKDV	FDKVFTEIHQV	---	ESVPTT	LT	PKLEVGRILTMSTGIANVSGLP																																																																
			80	90	100	110	120	130	140																																																																
maize	chl		IMSGELVEFAE	GTRGIALNLES	KNVGI	VL	MGDGL	ME	QEGSFVKATGRIA	OIPVSAYLGRVINALAKPTD																																																															
maize	mit		IQA	GEVVEFAS	GVKGI	LN	LEN	ENV	GIVVFGSD	ATKEGDUVKRTGSI	VDVPA	GRAL	LGRV	DAL	GV	PTD																																																									
bovine	mit		VQAE	EMVEFSS	GLK	GMS	LN	LE	PNV	GVV	FG	NDK	L	KE	GD	IV	KRTGAI	VDV	VG	EELL	GR	V	DAL	GN	ATD																																																
<i>E. coli</i>			CMO	GEM	IS	LP	GN	RY	ATA	LN	LR	DS	V	GA	V	MG	PP	Y	DL	AE	G	M	V	K	G	T	G	R	I	LE	V	P	V	G	R	G	L	L	GR	V	N	T	L	GA	P	T	D																										
<i>M. barkeri</i>			VG	FD	EL	IK	FP	GD	LF	GIA	FN	VD	EN	ET	GT	VL	LG	EYS	HL	HA	GD	Q	V	E	R	T	R	R	V	M	D	V	A	V	G	E	L	L	GR	V	I	D	P	L	GR	P	L	D																									
			150	160	170	180	190	200	210																																																																
maize	chl		QRGE	TV	ASE	R	L	IES	P	AP	G	T	I	S	R	S	V	Y	E	P	L	O	T	G	L	T	A	I	D	S	M	I	P	I	G	R	G	O	R	E	L	I	I	G	D	R	O	T	G	K	T	A	V	A	I	D	T	I	L	N	O												
maize	mit		GK	G	A	L	S	D	H	E	R	R	R	V	E	V	K	A	P	G	I	E	R	K	S	V	H	E	P	M	O	T	G	L	K	A	V	D	S	L	V	P	I	G	R	G	O	R	E	L	I	I	G	D	R	O	T	G	K	T	A	I	A	I	D	T	I	L	N	O			
bovine	mit		GK	GP	I	G	S	K	A	R	R	V	V	L	K	A	P	G	I	P	R	I	S	V	R	E	P	M	O	T	G	I	K	A	V	D	S	L	V	P	I	G	R	G	O	R	E	L	I	I	G	D	R	O	T	G	K	T	S	I	A	I	D	T	I	L	N	O					
<i>E. coli</i>			GK	GPL	D	H	D	C	F	S	A	V	E	A	I	A	P	G	V	I	E	R	C	S	V	D	O	P	V	O	T	G	K	A	V	D	S	L	V	P	I	G	R	G	O	R	E	L	I	I	G	D	R	O	T	G	K	T	A	I	A	I	D	T	I	L	N	O					
<i>M. barkeri</i>			NK	GP	I	V	S	K	R	L	P	I	E	R	P	S	P	A	I	M	D	R	A	P	V	T	P	L	O	T	G	V	K	V	I	D	A	L	I	P	I	G	R	G	O	R	E	L	I	I	G	D	R	O	T	G	K	T	A	I	A	I	D	T	I	L	N	O					
			220	230	240	250	260	270	280																																																																
maize	chl		R	---	---	G	O	D	V	I	C	V	V	A	I	G	O	R	A	S	S	V	A	O	V	V	T	T	F	H	E	E	G	A	M	E	Y	T	I	V	V	A	E	M	A	D	S	P	A	T	L	O	Y	L	A	P	Y	T	G	A	L	A	E										
maize	mit		R	O	M	N	S	R	G	T	N	E	S	E	T	L	Y	C	V	V	A	I	G	O	R	S	T	V	A	O	L	V	O	I	L	S	E	A	N	A	L	E	Y	S	M	L	V	A	A	T	A	S	D	P	A	P	L	O	F	L	A	P	Y	S	G	C	A	M	G	F			
bovine	mit		R	R	E	N	D	G	T	E	K	K	L	Y	C	I	Y	V	A	I	G	O	R	S	T	V	A	O	L	V	K	R	L	T	D	A	D	A	M	K	Y	T	I	V	S	A	T	A	S	A	A	P	L	O	Y	L	A	P	Y	S	G	C	M	G	E								
<i>E. coli</i>			R	---	---	D	S	G	I	K	C	I	Y	V	A	I	G	O	R	A	S	T	S	N	V	R	K	L	E	E	G	A	N	T	I	V	V	A	T	A	S	E	S	A	A	L	O	Y	L	A	R	M	P	V	A	L	M	G	E														
<i>M. barkeri</i>			R	N	F	---	---	N	V	L	C	V	Y	C	A	I	G	O	R	A	S	A	V	A	R	A	V	A	N	L	R	E	G	A	N	D	Y	T	I	V	V	V	T	E	G	N	D	P	S	G	L	I	Y	I	T	P	Y	A	A	T	S	I	A	E									
			290	300	310	320	330	340	350																																																																
maize	chl		Y	F	M	Y	R	E	R	H	L	I	I	Y	D	D	L	S	K	Q	A	C	A	Y	R	Q	M	S	L	L	L	R	P	P	G	R	E	A	Y	L	G	D	V	F	Y	L	H	S	R	L	L	E	R	A	A	K	L	N	S	L	L	G	E										
maize	mit		Y	F	R	D	N	G	M	H	A	L	I	I	Y	D	D	L	S	K	O	A	V	A	Y	R	Q	M	S	L	L	L	R	P	P	G	R	E	A	P	G	D	V	F	Y	L	H	S	R	L	L	E	R	A	A	K	M	N	D	A	F	G	G										
bovine	mit		Y	F	R	D	N	G	M	H	A	L	I	I	Y	D	D	L	S	K	O	A	V	A	Y	R	Q	M	S	L	L	L	R	P	P	G	R	E	A	P	G	D	V	F	Y	L	H	S	R	L	L	E	R	A	A	K	M	N	D	A	F	G	G										
<i>E. coli</i>			Y	F	K	D	K	G	E	D	A	L	I	I	Y	D	D	L	S	K	O	A	V	A	Y	R	Q	M	S	L	L	L	R	P	P	G	R	E	A	P	G	D	V	F	Y	L	H	S	K	L	L	E	R	A	R	V	N	A	B	Y	V	E	A	F	T	K	G	E	V				
<i>M. barkeri</i>			Y	F	M	E	S	G	R	D	V	L	I	Y	D	D	L	T	N	H	A	R	A	Y	R	E	L	S	L	L	L	R	P	P	G	R	E	A	P	G	D	T	F	Y	I	H	S	R	L	L	E	R	S	T	H	L	K	E	F	G	G												
			360	370	380	390	400	410	420																																																																
maize	chl		---	G	S	M	T	A	L	P	I	V	E	T	O	S	G	D	V	S	A	Y	I	P	T	N	V	I	S	I	T	D	G	O	I	F	L	S	A	D	L	F	N	A	G	I	R	P	A	I	N	V	G	I	S	V	S	R	V	G	S	A	A	O	I	T	R	A	N	H			
maize	mit		---	G	S	L	T	A	L	P	V	I	E	T	O	A	G	D	V	S	A	Y	I	P	T	N	V	I	S	I	T	D	G	O	I	C	L	E	T	E	L	F	Y	R	G	I	R	P	A	I	N	V	G	L	S	V	S	R	V	G	S	A	A	O	L	K	A	N	H				
bovine	mit		---	G	S	L	T	A	L	P	V	I	E	T	O	A	G	D	V	S	A	Y	I	P	T	N	V	I	S	I	T	D	G	O	I	F	L	E	T	E	L	F	Y	K	G	I	R	P	A	I	N	V	G	L	S	V	S	R	V	G	S	A	A	O	I	T	R	A	N	H			
<i>E. coli</i>			---	K	G	K	T	G	S	L	T	A	L	P	I	I	E	T	O	A	G	D	V	S	A	Y	I	P	T	N	V	I	S	I	T	D	G	O	I	F	L	E	T	N	L	F	N	A	G	I	R	P	A	V	N	G	I	S	V	S	R	V	G	G	A	A	O	I	T	R	A	N	H
<i>M. barkeri</i>			---	G	S	L	T	A	L	P	I	I	E	T	E	A	O	N	I	S	A	Y	I	P	T	N	L	I	S	I	T	D	G	O	I	Y	L	S	P	S	L	F	E	L	G	V	L	P	A	V	D	V	G	R	S	V	S	R	V	G	G	K	A	O	L	A	A	Y	R				
			430	440	450	460	470	480	490																																																																
maize	chl		Q	V	A	G	K	S	K	L	E	L	A	Q	F	A	E	L	Q	A	F	A	Q	F	A	S	A	L	D	K	T	S	O	N	L	A	R	G	R	L	R	E	L	L	Q	S	N	P	L	P	V	E	Q	V	A	T	I	Y	T	G	T	R	G	Y	L	D							
maize	mit		D	V	L	G	G	S	K	L	E	L	A	O	Y	R	E	V	A	A	F	A	O	F	G	S	D	L	D	A	A	T	O	A	L	L	N	R	G	A	R	L	T	E	V	P	K	O	P	O	Y	E	P	L	T	E	K	O	I	V	V	I	A	A	V	N	G	F	C	D			
bovine	mit		Q	V	A	G	T	M	K	L	E	L	A	O	Y	R	E	V	A	A	F	A	O	F	G	S	D	L	D	A	A	T	O	L	L	S	R	G	V	S	L	T	E	L	L	K	Q	O	Y	S	P	M	A	I	E	Q	V	A	V	I	A	G	V	R	G	Y	L	D					
<i>E. coli</i>			K	L	S	G	G	I	R	T	A	L	A	O	Y	R	E	L	A	A	F	S	O	F	A	S	D	L	D	A	T	R	N	Q	L	D	H	G	Q	K	V	T	E	L	L	K	O	K	O	Y	A	P	M	S	V	A	Q	S	L	V	L	F	A	A	E	R	G	Y	L	A			
<i>M. barkeri</i>			E	V	A	G																																																																			



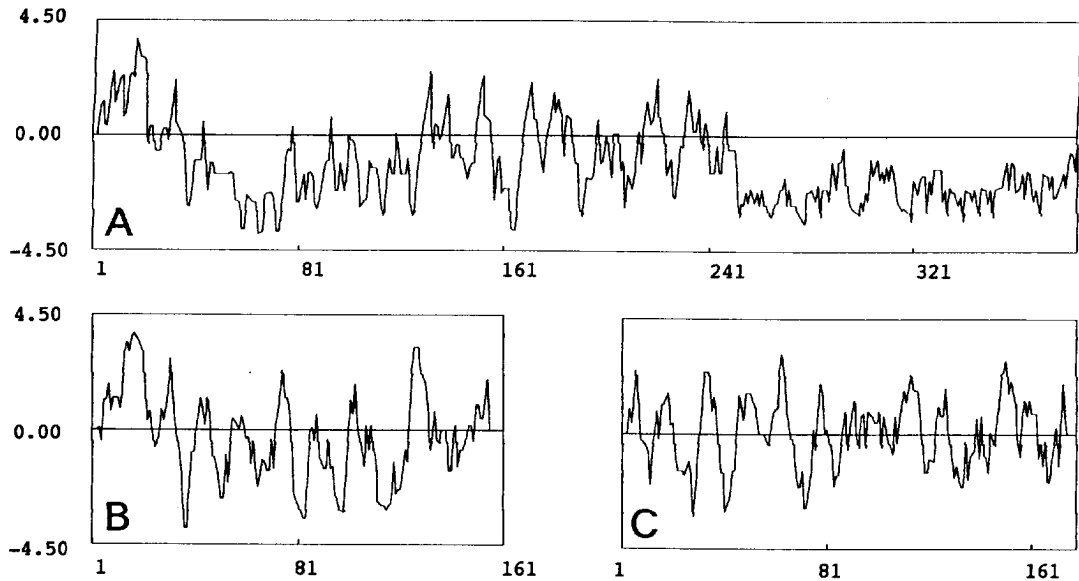
B: beta subunits

			10	20	30	40	50	60	70
maize	chl		MRTNPTTSRPGIS	IIEEKSVGR	DDIIGPVLGIT	FPPGKLPYIYNALIVKSRDTADKQ	INVTCEVQQLG		
maize	mit		---DEFTGAGAT---	GOVGOVIGAVVDVRF	DEGLPPILTALEVL---	DNNIRLVLEVAGHLG			
bovine	mit		---SPSPKAGATT---	GRIVAVIGAVVDVQF	DEGLPPILNALEVQ---	GRETRLVLEVAGHLG			
<i>E. coli</i>			---MAT---	GKIVQVIGAVVDVE	FPODAVPRVYDALEVO---	NGNEIRLVLEVQQLG			
<i>M. barkeri</i>			---MVDRSRLNL---	GKVISVRGSDVDVLE	EKHLPPVYTLIRA---	GRESQIAIEVLTQLD			
			80	90	100	110	120	130	140
maize	chl		NNRVRAVAVNSATEGLMRGHEVIDTGTPLS	VVPVGGATLGRIFENVL	GEPIDNLG	QVDTISATFP	PIHRSAPAFI		
maize	mit		ENMVRTIAMDGTEGLVRRGQKVLNLTGSPIT	VVVGRTATLGRITNVV	GEPIDEKGDIKT	NHFLPIHREAPAFV			
bovine	mit		ESTVRTIAMDGTEGLVRRGQKVLNLTGSPIT	VVVGRTATLGRITNVV	GEPIDEKGDIKT	NHFLPIHREAPAFV			
<i>E. coli</i>			GGIVRTIAGSSDGLRRGLDVKOLEHPI	VVVGKATLGRITNVV	GEPIVDMKGL	GEERWA	IHRAAPSYE		
<i>M. barkeri</i>			AHHVVGIALTPTEGLARGHAEVDTGGPLKAPVGREILSRM	EDVFGHTIDRRK	PPSDIOWRS	IHOTPPPLM			
			150	160	170	180	190	200	210
maize	chl		ELDTKLSIFETGIRKVDLLAPYR	RGGKIGLFGGAGVGKTVL	IMELINNI	AKAHGGVSVF	G	GVGERTREGN	
maize	mit		EQATEQDI LVTGIRKVDLLAPYR	RGGKIGLFGGAGVGKTVL	IMELINNVAKAHGG	F	SVFAGV	GVGERTREGN	
bovine	mit		ELSVIEQDI LVTGIRKVDLLAPYR	RGGKIGLFGGAGVGKTVL	IMELINNVAKAHGG	F	SVFAGV	GVGERTREGN	
<i>E. coli</i>			ELSNSQELLETGIRKVDLLAPYR	RGGKIGLFGGAGVGKTVL	IMELIRNIAIEHS	G	SVFAGV	GVGERTREGN	
<i>M. barkeri</i>			RRSTTSLET FETGIRKVDLLAPYR	RGGKIGLFGGAGVGKTVL	IMELINNVVNO	H	GVSTF	GVGERTREGN	
			220	230	240	250	260	270	280
maize	chl		DLYMEMKESGVIN	EKNIEESRVALVY	QGMNEPPGARMRVGLTA	LTMAEYFRD	VNKQD	VLLFIDNIFRFV	
maize	mit		DLYREMIESGVIN	LDKQSESKCAL	VYQGMNEPPGARARVGL	TGLTVAEH	FRDA	EGQDVL	
bovine	mit		DLYHEMIESGVIN	LKD--ATS	KVALVYQGMNEPPGARARVGL	TGLTVAEY	FRD	EGQDVL	
<i>E. coli</i>			DLYHEMIESGVIN	LDKQSESKCAL	VYQGMNEPPGARARVGL	TGLTVAEY	FRD	EGQDVL	
<i>M. barkeri</i>			ELYRDMNDAGVLE	-----PNT	VMVFGGMNEPPGARFRV	GHTAL	TMAEYFRD	ERQDVL	
			290	300	310	320	330	340	350
maize	chl		OAGSEVSALLGRMPSAVGYOPTL	STENGSL	LOERITSTRKGSITS	ICAVYVPADDL	TDPA	PATTF	
maize	mit		QANSFVSALLGRMPSAVGYOPTL	ATDMGT	LOERITSTRKGSITS	SVQAT	YVPADDL	DPAPATTF	
bovine	mit		OAGSEVSALLGRMPSAVGYOPTL	ATDMGT	LOERITSTRKGSITS	SVQAT	YVPADDL	DPAPATTF	
<i>E. coli</i>			LAGTEVSALLGRMPSAVGYOPTL	AEMGV	LOERITSTRKGSITS	VOAVYVPADDL	DPSPATTF	FAHLDAT	
<i>M. barkeri</i>			OAGSEVSGLMGQMP SRLIGYOPTL	GTLEL	SELEERIST	TDAGATMS	IQAVYVPADDL	FDPSAVHT	
			360	370	380	390	400	410	420
maize	chl		TVLSRGLASKGIYPAVDPLDSTS	IMLOR	RIVGNEHYETAORVKE	T	LORYKEL	ODIIAILGL	
maize	mit		TVLSRGLASKGIYPAVDPLDSTS	IMLOR	RIVGNEHYETAORVKE	T	LORYKEL	ODIIAILGL	
bovine	mit		TVLSRGLASKGIYPAVDPLDSTS	IMLOR	RIVGNEHYETAORVKE	T	LORYKEL	ODIIAILGL	
<i>E. coli</i>			TVLSRGLASKGIYPAVDPLDSTS	IMLOR	RIVGNEHYETAORVKE	T	LORYKEL	ODIIAILGL	
<i>M. barkeri</i>			IVLSRGLASKGIYPAVDPLDSTS	IMLOR	RIVGNEHYETAORVKE	T	LORYKEL	ODIIAILGL	
			430	440	450	460	470	480	490
maize	chl		TVARARKI	ERFLSOPFFVAEVFT	GSPGKYVGL	LAETIRGF	ILS	GELDGL	
maize	mit		TVARARKI	ORFLSOPFFVAEVFT	GAPGKYVEL	KS	FOGVL	DGKYDDL	
bovine	mit		TVARARKI	ORFLSOPFFVAEVFT	GAPGKYVEL	KS	FOGVL	DGKYDDL	
<i>E. coli</i>			VVARARKI	ORFLSOPFFVAEVFT	GSPGKYVLS	LDTIRGF	GLME	GEYDHL	
<i>M. barkeri</i>			VVARARLR	ERFLTOPFFITTE	FTIGKNS	ISLSDAL	D	CERTL	
			499						
maize	chl		LEEEESK	LKK					
maize	mit		LAEEHS	---					
bovine	mit		LAEEHS	---					
<i>E. coli</i>			L	---					
<i>M. barkeri</i>			REIS	---					

(図17) *M. barkeri* AlpA および AlpI の一般的なFタイプの  $\alpha$  (パネルA)、 $\beta$  (パネルB) サブユニットに対する相同性。chl は葉緑体Fタイプ、mit はミトコンドリアDタイプを示す。- は適正な類似性を持つために挿入されたギャップ。そのサイトで保存されているアミノ酸を黒く塗りつぶしてある。

3.2.5  $\delta$ サブユニット遺伝子の欠落

*M. barkeri* の *atp* オペロンには  $\delta$  サブユニットをコードすると思われる遺伝子が見つからない。 $\delta$  サブユニットをコードする遺伝子は、通常、b および  $\alpha$  サブユニットをコードする遺伝子の間隙に位置する。*M. barkeri* では、この領域に読み枠は見つからず、代わりに AtpF のカルボキシル末端に付加した長いペプチド鎖のコードが続いている。また、最近、立体構造上 b と  $\delta$  は相互作用をしている可能性が高いとの報告がある (Sawada *et al.*, 1997)。以上の状況から、AtpF カルボキシル末端の付加配列がなんらかの理由で b サブユニットに融合してしまった  $\delta$  サブユニットではないかという見方は、到達しやすい考え方ではある。しかしながら、アミノ酸配列の類似性検索の結果や疎水性/親水性プロットのいずれもそうした事実は示唆しない (図18)。



(図18) *M. barkeri* AtpF (パネルA) と *E. coli* b (パネルB) および  $\delta$  (パネルC) サブユニットの疎水性/親水性プロット。bサブユニットはアミノ末端に膜貫通部位と思われる特に強い疎水性領域を持つ。 $\delta$ は水可溶性タンパク質であるが全体に疎水性領域を有する。仮に AtpF のカルボキシル末端に  $\delta$ サブユニットが融合されてしまった場合、 $\delta$ の疎水性が見られるはずである。しかし、この領域は逆に親水性になっている。プロット、プラスが疎水性、マイナスが親水性。インデックスは Kyte と Doolittle のものを使用。

### 3.2.6 $\gamma$ サブユニット遺伝子の二分化

AtpG (149aa) は、*Arabidopsis thaliana* の葉緑体Fタイプの $\gamma$ と23%程度の相同性を持つ。しかし、一般的な $\gamma$ サブユニットの長さ(約300aa)に比べて、これは半分程度の長さでしかない。一方、AtpJ (186aa) は、当初、類似性検索には適当な候補が見つからなかったが、検索システムを変えると *Bacillus subtilis* の $\gamma$ のカルボキシル末端と31%の相同性を持つことが分かった。この情報を総合すると、AtpGは $\gamma$ のアミノ末端側の半分、AtpJはカルボキシル末端側半分をコードすることになる。

これら2つの遺伝子の読み枠のオーバーラップは110 bp と異常に長い(図16、B)。また、GC含有量の少ない *M. barkeri* には珍しくCの連続する配列も含まれるため、シーケンズ反応のミスという側面も考慮し、反応系を変えて再三配列決定を行なったが、読み間違いは発見できなかった。

Walker らの報告したウシ心筋ミトコンドリアF<sub>1</sub>部分の立体構造[Abrahams *et al.*, 1994]では、 $\gamma$ サブユニットは、アミノ末端とカルボキシル末端で構成されるコイルドコイル構造により、 $\alpha_3\beta_3$ の構成する球状構造の中心を貫く軸のような構造を形成することが分かっている。仮に、*M. barkeri* の $\gamma$ がアミノ末端とカルボキシル末端とで別にコードされたタンパク質であっても、こうした構造をとることは十分可能であり、機能上の障害にはならないと考えられる。

## 3.3 考察

### 3.3.1 *M. barkeri* のFタイプ遺伝子は機能しているのか?

*M. barkeri* の *atp* オペロンは、明らかに F タイプの特徴を持つ遺伝子ではあるが、個々の遺伝子産物の予想される性質は、従来報告されている F タイプのものと大きく異なる点が多い。もし、この遺伝子が発現し機能しているとすれば、F タイプの構造に関する研究に大きなインパクトを与えることは間違いないだろう。この意義をふまえて、*M. barkeri* における *atp* オペロン発現の証拠を得るべく、ノザンプロット、ウエスタンプロット解析などを行なった。しかし、残念ながらその痕跡を検出することはできなかった。

*M. barkeri* と同様に、V/F 両タイプの遺伝子を持つ *E. hirae* は、低い pH の際には F タイプの ATP 水解と供役したプロトン輸送で膜電位を作り出し、高い pH もしくは高いナトリウムイオン濃度の環境下では V タイプを用いたナトリウムイオン輸送機能を発現することが知られている [Kakinuma and Igarashi, 1989]。この際、V タイプの活性はその発現量により調節されている。*M. barkeri* の F タイプの発現量も、同じ様な仕組みで pH やナトリウムイオン濃度によって調節されている可能性があると考え、高低 pH (7.5~4.8)、高低ナトリウムイオン濃度 (400~11mM) の培地を特別に調製し、*M. barkeri* の培養を試みたが、得られた菌体を用いたノザン/ウエスタンプロット解析の結果には反映されなかった。

### 3.3.2 古細菌における F タイプの分布

その発現および機能の確認はできなかったが、*M. barkeri* に F タイプ遺伝子が存在することは明らかである。これは、遺伝子レベルでは少なくとも、F タイプの分布は真正細菌に限られてはいないことを示すのに十分な証拠となりうるが、古細菌において、どの程度 F タイプの遺伝子が分布しているかどうかは本研究では明らかにできない。同じ属の *M. mazei* ではサザンプロット解析の結果、F タイプ存在の可能性は低いという報告がある。また、全ゲノム解析の終わった、*M. jannaschii*、*M. thermoautotrophicum*、*A. fulgidus* を対象に提供されている検索システムでも、F タイプの類似性遺伝子は見つからないことから、*M. barkeri* の F タイプ遺伝子は古細菌の中でもかなり珍しい存在である可能性が高い。

## 第4章 古細菌におけるFタイプATPase 遺伝子の存在とバクテリア進化との関係

FタイプもしくはVタイプにおいて見られる主要2サブユニットの類似性の高さは、それらが共通の始原遺伝子から分化したものであることを示唆する。また、FタイプとVタイプ自身も互いに機能や構成成分などにおいて高い類似性を持ち、やはり共通の始原タイプから派生したものであることが予測可能だ。バクテリアから高等生物まで一貫して高度な保存性を保ちつつ、これほどまでに広範に分布するタンパク質は他にはないため、ATPaseはrRNAなどと並ぶ生物進化のマーカー遺伝子としてよく用いられてきた。新しく得られた *M. barkeri* Fタイプの情報も含め、改めてF/Vタイプの $\alpha/\beta$ サブユニットの一次配列を使って分子進化系統樹を作成した結果、Fタイプの系統樹に真正細菌と古細菌の分岐を示すと思われる新しい枝を継ぎ足すことができたので報告する。

### 4.1 方法

#### 4.1.1 塩基配列の収集

GenBank データベースを通じて、ATPase 関連の登録された遺伝子のうち、 $\alpha/\beta$ サブユニットの遺伝子を含む塩基配列を引用した。得られた塩基配列と付加された読み枠に関する情報をもとに、想定されるアミノ酸配列に翻訳し情報として用いた。真核生物については、ミトコンドリア、葉緑体（植物の場合）のFタイプ、空胞系のVタイプの遺伝子が揃って報告されているものを探した。使用した遺伝子を保持する生物の分類は表04。 [*Halobacterium salinarium*, GenBank HSATPABA; *Haloferax volcanii*, HVATPAB; *Methanococcus jannaschii*, MJU67477; *Methanosarcina barkeri*, MESATPAB; *Methanosarcina mazeii*, MMU47274; *Sulfolobus acidocaldarius*, SSOATPMA; *Enterococcus hirae*, ENENIP; *Thermus thermophilus*, D63799; *Bos taurus*, BOVA1PASEV, BOVATPF1A; *Drosophila melanogaster*, DMU19742; *Hordeum vulgare*, HVU36939; *Neurospora crassa*, NEUVMA1A, NEUATPB; *Saccharomyces cerevisiae*, YSCHATP, SCU46215; *Trypanosoma congolense*, TC-VACATPA; *Zea mays*, ZMU36436, CHZMXX, ZMATP2MT; *Cyanophora paradoxa*, CPU30821; *Marchantia polymorpha*, CHMPXX; *Nicotiana tabacum*, CHNTATPBG, NPATP21; *Oryza sativa*, RICCPATP, RICATPB; *Pinus thunbergiana*, PINCPTRPG; *Porphyra purpurea*, PPU38804; *Spinacia oleracea*, SOU23082; *Triticum aestivum*, WHTCPATPB; *Hevea brasiliensis*, HBATPB; *Rattus norvegicus*, RATMTATPSA]

## バクテリア

Hsa	<i>Halobacterium salinarium</i>
Hvo	<i>Haloferax volcanii</i>
Mja	<i>Methanococcus jannaschii</i>
Mba	<i>Methanosarcina barkeri</i>
Mma	<i>Methanosarcina mazei</i>
Sac	<i>Sulfolobus acidocaldarius</i>
Bsu	<i>Bacillus subtilis</i>
Ehi	<i>Enterococcus hirae</i>
Eco	<i>Escherichia coli</i>
Pmo	<i>Propionigenium modestum</i>
Rbl	<i>Rhodopseudomonas blastica</i>
Sli	<i>Streptomyces lividans</i>
Syn	<i>Synechococcus</i> sp.
Tth	<i>Thermus thermophilus</i>

## 真核生物

Cre	<i>Chlamydomonas reinhardtii</i>	コナミドリムシ
Dme	<i>Drosophila melanogaster</i>	ショウジョウバエ
Sce	<i>Saccharomyces cerevisiae</i>	酵母
Tco	<i>Trypanosoma congolense</i>	トリパノソーマ (寄生虫)
Bta	<i>Bos taurus</i>	ウシ
Rno	<i>Rattus norvegicus</i>	ラット
Spu	<i>Strongylocentrotus purpuratus</i>	ウニ
Ncr	<i>Neurospora crassa</i>	アカバンカビ
Cpa	<i>Cyanophora paradoxa</i>	藻類
Mpo	<i>Marchantia polymorpha</i>	コケ
Ppu	<i>Porphyra purpurea</i>	紅藻類
Hbr	<i>Hevea brasiliensis</i>	ゴム
Hvu	<i>Hordeum vulgare</i>	大麦
Nta	<i>Nicotiana tabacum</i>	タバコ
Osa	<i>Oryza sativa</i>	コメ
Pth	<i>Pinus thunbergiana</i>	マツ
Psa	<i>Pisum sativum</i>	エンドウ
Sol	<i>Spinacia oleracea</i>	ホウレンソウ
Tae	<i>Triticum aestivum</i>	コムギ
Zma	<i>Zea mays</i>	トウモロコシ

(表04)

本論文中では由来する生物の学名を3文字に省略したものを使用する。また、それぞれの一次配列を区別するために、1文字目 (m, c, v, e, a) で順にミトコンドリア、葉緑体、空胞系、真正細菌、古細菌の区別を、2文字目 (f, v) でF/Vタイプの区別を、3文字目 (c, r) で触媒サブユニット (Fタイプは $\beta$ 、Vタイプは $\alpha$ ) か非触媒サブユニット (Fタイプでは $\alpha$ 、Vタイプでは $\beta$ ) の区別を付ける記号をアンダーラインを隔てて先頭に付した。たとえば、古細菌に属する *M. barkeri* のFタイプ触媒サブユニットは afc\_Mba と表記する。

### 4.1.2 アライメント

GCG の提供するアライメントパッケージ (pileup) を利用し、より多くのアミノ酸残基サイトにおいて最適の類似性を示すように並べた。ギャップ作成に関するパラメータはデフォルト設定を使用 (gap creation penalty, 3.00; gap extension penalty, 0.10) 。

### 4.1.3 系統樹の作成

PHYLIP (Phylogeny Inference Package; Washington 大学, J. Felsenstein 提供) を使用した。アミノ酸差算出には Protdist、系統樹作成には Neighbor、系統樹の描画には Drawtree を使用した。アミノ酸置換行列は Dayhoff らの PAM インデックスを採用。ブーツストラップ用のアミノ酸配列は Seqboot を用い 100 セット作成した。アミノ酸差算出、系統樹描画の後、コンセンサス樹形の抽出には Consense を用いた。

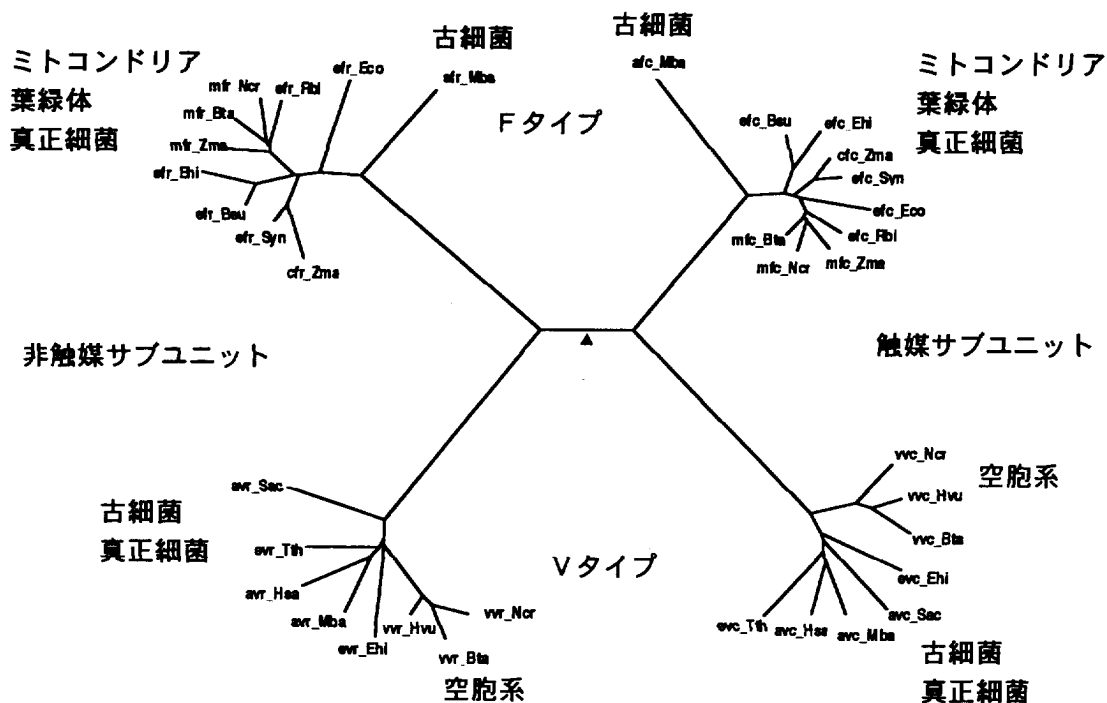
## 4.2 結果

### 4.2.1 アライメントと相同性マーク

入手した一次構造を一挙に並べ、GCG パッケージによりアライメントを取った（付録C）。計算には挿入配列や欠失の影響を受けていない保存配列を用いる。挿入配列や欠失を受けた領域は手作業で除いた。保存領域を見つけやすくするために、*B. subtilis* のFタイプ触媒サブユニットと相同のアミノ酸について黒く塗りつぶしマークした。こうして抽出されたサイトの数は約380であった。

### 4.2.2 系統樹の作成

真核生物では、植物、動物、菌類を、真正細菌では *Eco*、*Bsu* の他に、Vタイプの報告されている *Tth*、*Ehi*、ミトコンドリアおよび葉緑体のもとになった生物と比較的近縁と思われる *Rbl*、*Syn* を加えた。古細菌は *Sac*、*Hsa*、*Mba* を使用。*Mba* は本研究で明らかになったFタイプも加えた。以上の生物分類において、それぞれのFタイプ、Vタイプの触媒/非触媒サブユニットがどのような系統樹を形成するのかを確認した（図19）。



〈図19〉 触媒/非触媒サブユニット、F/Vタイプの系統樹。

### 4.2.3 系統樹の統計的検定

得られた系統樹の信頼性を調べるのには一般にブーツストラップ法が用いられる[Nei, 1987]。ブーツストラップ法というのは  $m$  個のアミノ酸サイトからなる複数の配列がある場合、この中から  $m$  個のサイトを重複を許して再抽出し、この新しいデータからまた系統樹を作成するものである。重複を許す再抽出により、サイト数を無限大と仮定した場合、その集団から  $m$  個のサイトを無作為に抽出したのと



成したり、アミノ酸配列の類似性に着目した過去の知見とは矛盾しない[Gorgarten *et al.*, 1989; Nelson, 1992]。

#### 4.3.2 新たに書き加えられた古細菌のFタイプの枝と真正細菌/古細菌の分化

本研究により、Fタイプの系統樹(図20の上半分)に新たに2本の枝を継ぎ足すことができた。触媒サブユニットと非触媒サブユニットのそれぞれ根に近い側にある *afc\_Mba*, *afr\_Mba* とラベルした枝がそれである。ブーツストラップ値もそれぞれ 97, 100 と十分大きく、第3章の他の種の $\alpha$ および $\beta$ サブユニットとの相同性比較からも直感的に推察されたように、古細菌のFタイプ遺伝子産物が、他のFタイプ遺伝子産物とは一線を画す存在であることが系統樹からも裏付けられた。

*M. barkeri*のFタイプ遺伝子が古細菌のFタイプを代表することが可能な存在かどうかは、1種しか発見できていない現段階では明言できない。この新たに書き加えられた *M. barkeri* Fタイプの枝に、今後、発見されるであろう他の古細菌のFタイプの遺伝子産物の枝が継ぎ足されるならば、このFタイプの系統樹を用いて、真正細菌と古細菌の分岐の時期を推察することが可能になるだろう。つまり、従来考えられていたように[Gorgarten *et al.*, 1989; Iwabe, N. *et al.*, 1989]、F/Vの分岐と真正細菌/古細菌の分岐時期は一致するのではなく、F/Vの分岐が起こってからかなり経過した後に真正細菌と古細菌の分化が起こったのであることを示すことが可能になる。

繰り返しになるが、このことを示すには、*M. barkeri*以外の古細菌、できれば高度好塩菌や好酸好熱古細菌からFタイプの遺伝子を発見する必要がある。

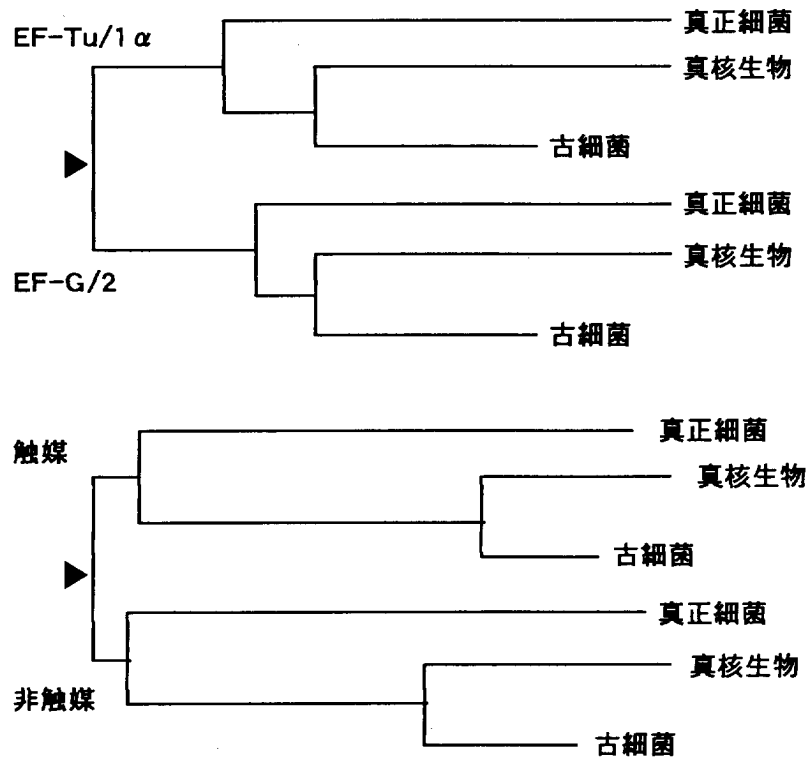
#### 4.3.2 系統樹の根

Woeseらが古細菌を提唱して以来、古細菌/真正細菌/真核生物を結ぶ系統樹のどの部分に根、つまり始原生物からの分岐点を設けるかについて議論が重ねられてきた。現存生物の分子の比較から、無根の系統樹を算出することは可能だが、この系統樹には時間の軸がないため、どの順序で各枝が分岐したのかを推察することはできない。無根の系統樹から有根の系統樹を作るには、比較した生物グループに比べ明らかに古い時期に分岐したことが分かっている別の生物の系統を比較に含める必要がある。しかし、全生物の有根系統樹を描く際に、現存する生物からこの目的に適したものを選ぶことはできない。

宮田らは、複合系統樹法によりその問題を解決した[Iwabe, N., *et al.*, 1989]。古細菌、真正細菌、真核生物の共通の祖先を根とするのではなく、さらに古い時代にさかのぼって根を設定しても同じ結果を導くことができる。つまり、遺伝子重複により生じた双子の遺伝子産物の系統樹、複合系統樹を利用することで派生的に有根系統樹を作成したのである。遺伝子重複は三者の分岐に先行するので、作成された系統樹の対称性より重複時期を決定することができ、しいては、この重複点を境に2つの有根系統樹が自動的に作成される。



複合系統樹から古細菌、真正細菌、真核生物の系統関係を知るために利用できる分子には制約がある。第1に超生物界に共通の遺伝子であること。第2に進化がたどれる程度に保存されていること。第3に、第2の条件を満たして、なお統計的計算に耐えられるよう十分な長さの配列であること。以上の3点を満たしている必要がある。このような条件を満たす遺伝子の数は少ない。宮田らは延長因子EF-Tu/1 $\alpha$ とEF-G/2の遺伝子対およびF/V-ATPaseの触媒/非触媒サブユニットの遺伝子対のそれぞれに対して複合系統樹を作成し、Woeseの系統樹に根をつけることに成功した(図21)。



(図21) Iwabe, N., et al., 1989 を改変。

この宮田らの研究成果はたいへんすばらしいものであったが、ただひとつ、F/Vタイプの分岐時期を真正細菌/古細菌の分岐時期と一致させる説(第1章、図02)に乗っ取っていること、つまりFタイプは真正細菌にのみ、Vタイプは古細菌にのみ分布するという前提に立ちATPaseに関する系統樹を作成している点で議論の余地が残されていると考えた。

本研究により、古細菌のFタイプの遺伝子を新しく利用できるようになった。しかし、系統樹からもうかがえるように細胞内小器官のFタイプはすべて真正細菌が細胞内共生により持ち込んだものである可能性が高い。そこでFタイプを用いて、系統樹の根、ひいては細胞内共生において宿主となった真核生物の始原生物を明らかにするためには、宿主由来のFタイプに情報を求めなければならない。しかし、真核生物の細胞膜や空砲系からFタイプは見つかっていない。これは先に述べた複合系統樹に用いる情報の第1条件「普遍的な分布」に反する。残念ながらFタイプで系統樹の根を決定する作業はできない。

一方、本研究における古細菌Fタイプ発見とは直接関係はないが、Vタイプには真正細菌のVタイプの2例を含め、真正細菌、古細菌、真核生物（宿主由来と想定される空砲系）のすべてから情報を得ることができる。比較的高度に保存され、十分な長さを持つこのVタイプの触媒/非触媒サブユニットをコードする遺伝子対は、複合系統樹の作成に向いていると思われる。しかし、今回作成した系統樹では特にVタイプの真正細菌の枝（図19のラベル ev\*；\* は c または r）はブートストラップ値が小さく、分岐位置に関する誤差が大きいため信頼性に欠け、根の位置を特定するといった議論には耐えられない。

本研究の他にも、Kibak らや Gogarten と Hilario がVタイプを用いて根の決定を行なおうとしているが、いずれの研究においても真正細菌のVタイプの枝は誤差を含んでおり、明確な結論には至っていない。また、両者は誤差を含みながらも本研究で作成した系統樹と同様に真正細菌のVタイプの枝が好塩菌およびメタン産生菌の形成する枝につがれる結果を受けて、*T. thermophilus* や *E. hirae* により提供される真正細菌のVタイプの遺伝子は、遺伝子の水平移動によって古細菌より獲得したものであるという仮説を提唱している[Kibak, H., et al, 1992; Hilario, E. & Gogarten, J. P., 1993]。

他方、宮田は、真正細菌に属し、Vタイプがすでに報告されている *T. thermophilus* と同属の *Thermus aquaticus* YT-1株のリンゴ酸デヒドロゲナーゼが、真正細菌よりむしろ真核生物のものによく似ていること、*T. thermophilus* が古細菌と等しいVタイプを保持していることを併せて、研究途上と前置きしながらも *T. aquaticus* が始原生物（細胞内共生説の宿主）であるという仮説を述べている[日本生化学会編、新生化学実験講座(16)分子進化実験法、第4章]（しかし、文献[Iwabe, N., et al., 1989]では、遺伝子の水平移動や、*Thermus* が古細菌であるという可能性にもふれている点は興味深い）。

生物進化のマーカとして、類希な好条件を備えていると思われた ATPase であったが、前述のような理由で、Fタイプ、Vタイプともにマーカとしてはそれほど有用ではないという結論を、少なくとも現時点においては導きざるをえないことは残念である。分子進化研究の立場では、今後の発展として、1) 複数の古細菌からFタイプの遺伝子が発見され、今回Fタイプの系統樹に継ぎ足された *M. barkeri* の枝が、Fタイプにおける古細菌の典型的な枝であることが証明される、2) 多くの真正細菌のVタイプの遺伝子が発見され、誤差の少ないVタイプの系統樹の作成や、現在報告されている真正細菌のVタイプが遺伝子の水平移動で獲得されたものであることを系統樹から証明できるようになる、ことが期待される。

## 引用文献

- Abrahams, J. P., Leslie, A. G. W., Lutter, R. and Walker, J. E. (1994) *Nature* **370**, 621-8.
- Allmansberger, R., Knaub S. and Klein, A. (1988) *Nucl. Acids Res.* **16**, 7419-36.
- Boone, D. R., Whiteman, W. B. and Rouvière, J. (1993) *Methanogenesis*, New York: Ferry, J. G. 62.
- Boyer, P. D. (1987) *Biochemistry* **26**, 8503-7.
- Boyer, P. D. (1997) *Annu. Rev. Biochem.* **66**, 717-49.
- Cozens, A. L. and Walker, J. E. (1987) *J. Mol. Biol.* **194**, 359-383.
- Dancan, T. M., Parsonage, D. and Senior, A. E. (1986) *FEBS. Lett.* **208**, 1-6.
- Downie, J. A., Gibson, F. and Cox, G. B. (1979) *Annu. Rev. Biochem.* **48**, 103-31.
- Downie, J. A., Cox, G. B., Langman, L., Ash, G., Becker, M. and Gibson, F. (1981) *J. Bacteriol.* **145**, 200-10
- Denda, K., Konishi, J., Oshima, T., Date, T. and Yoshida, M. (1988a) *J. Biol. Chem.* **263**, 6012-15.
- Denda, K., Konishi, J., Oshima, T., Date, T. and Yoshida, M. (1988b) *J. Biol. Chem.* **263**, 17251-17254.
- Denda, K., Konishi, J., Oshima, T., Date, T. and Yoshida, M. (1989) *J. Biol. Chem.* **264**, 7119-21.
- Denda, K., Konishi, J., Hajiro, K., Oshima, T., Date, T. and Yoshida, M. (1990) *J. Biol. Chem.* **265**, 21509-13.
- Dunn, S. D. and Futai, M. (1980) *J. Biol. Chem.* **255**, 113-8.
- Foster, D. L. and Fillngame, R. H. (1982) *J. Biol. Chem.*, **257**, 2009-15.
- Fry, D. C., Kuby, S. A. and Mildvan, A. S. (1986) *Proc. Natl. Acad. Sci. U.S.A.* **83**, 907-11.
- Gorgarten, J. P. Kibak, H., Dittrich, P., Taiz, L., Oshima, T., Konishi, J., Denda, K. and Yoshida, M. (1989) *Proc. Natul. Acad. Sci. U.S.A.* **86**, 6661-5.
- Graber, P. (1987) *Bioelectrochemistry II*, Plenum Publishing Corp., 379-429.
- Higgins, C. F., et al. (1986) *Nature* **323**, 448-450.
- Hilario, E. and Gogarten, J. P. (1993) *BioSystems* **31**, 111-9.
- Hisabori, T., Yoshida, M. and Sakurai, H. (1986) *J. Biochem.* **100**, 663-70.
- Iwabe, N., et at (1989) *Proc. Natl. Acad. Sci. USA* **86**, 9355-59.
- Kakinuma, Y. and Igarashi, K. (1989) *J. Bioenerg. Biomembr.* **21**, 679-92.
- Kakinuma, Y. and Igarashi, K. (1994) *J. Biochem. (Tokyo)* **116**, 1302-8.
- Kibak, H., Taiz, L., Starke, T., Bernasconi, P. and Gogarten, J. P. (1992), *J. Bioenerg. Biomembr.* **24**, 415-24.
- Margulis, I. *Symbiosis in Cell Evolution*, New York: Freeman, W. H.
- Nei, M. (1987) *Molecular Evolutionary Gnetics*, New York: Columbia Univ., Press.
- Nelson, N. (1992) *Biochem. Biophys. Acta.* **1100**, 109-124.
- Nelson, N. and Taiz, L. (1989) *Trends Biochem. Sci.* **14**, 113-6.
- Sawada, K., Kuroda, N., Watanabe, H., Moritani-Otsuka, C. and Kanazawa, H. (1997) *J. Biol. Chem.* **272**, in press.

- Steinert, K., Kroth-Pancic, P.G. and Bickel-Sandkotter, S. (1995) *Biochim. Biophys. Acta* **1249**, 137-144.
- Sumi, M., Yohda, M., Koga, Y. and Yoshida, M. (1997) *Biochem. Biophys. Res. Commun.* **241**, 427-433.
- Takase, K., Yamato, I. and Kakinuma, Y. (1993) *J. Biol. Chem.* **268**, 11610-6.
- Takase, K., Kakinuma, S., Yamato, I., Konishi, K., Igarashi, K. and Kakinuma, Y. (1994) *J. Biol. Chem.* **269**, 11037-44.
- Tsutsumi, S., Denda, K., Yokoyama, K., Oshima, T., Date, T. and Yoshida, M. (1991) *Biochim. Biophys. Acta* **1098**, 13-20.
- Walker, J. E., Saraste, M., Runswick, M. J. and Gay, N. J. (1982) *EMBO J.* **1**, 945-51.
- Walker, J. E., Fearnly, I. M., Gay, N. J., Gibson, B. W., Northrop, F. D., Powell, S. J., Runswick, M. J., Saraste, M. and Tybulewicz, V. L. J. (1985) *J. Mol. Biol.* **184**, 677-701.
- Weber, J. and Seinior, A. E. (1997) *Biochim. Biophys. Acta* **1319**, 19-58.
- Woese, C. R. (1987) *Microbiol. Rev.* **51**, 221-71.
- Yokoyama, K., Oshima, T. and Yoshida, M. (1990) *J. Biol. Chem.* **265**, 21946-50.
- Yokoyama, K., Akanabe, Y., Ishii, N. and Yoshida, M. (1994) *J. Biol. Chem.* **269**, 12248-53.
- Yoshida, M. and Allison, W. S. (1986) *J. Biol. Chem.* **261**, 5714-21.
- Yoshida, M., Sone, N., Hirata, H., Kagawa, Y. and Ui, N. (1979) *J. Biol. Chem.*, **254**, 9525-44.

付録A *M. barkeri* Fタイプ ATPase 遺伝子の全塩基配列と予想されるアミノ酸配列

10 20 30 40 50 60 70 80 90 100 110 120 130 140 150  
GATCCGAACGCATCGCAGGGCAGATTCGGTAACTTGTAAATTTTACACCCCTCCCGAGTGGATACAGACGGAGCAGTCCGTAGTGTATCTGGATATGGGGCTTTTGTGTGAATAACATCCGGAAAAATGTCAGCTCTGATCTTATAG

160 170 180 190 200 210 220 230 240 250 260 270 280 290 300  
GCCAAGAACAGATTCTTCAGGTATCGCTCATGGGCCCGCATATGATACCGCATATGGTGTGAACGCTTCAATGGAACACATCAACGAATGAATGGTCTTCAGTATGGGCAAGTGCAGATGTTTCGTGCAGAACGATGTTCCGATCT

310 320 330 340 350 360 370 380 390 400 410 420 430 440 450  
ACAGCCAATCGGTAATATGGCCCTCATCAATGGCTACAGGTTCCAGACGGCTGGAAAGTGTCTGGACTGACTGGGATACCGGTGAAATGTCATCAGACGATTTTTGGCGATGCAACTTGGAAACGGGGCATATGCAATGCTGG

460 470 480 490 500 510 520 530 540 550 560 570 580 590 600  
AGTACCTGGAAATGACGATCTGATTTTCAATTCCTCGTGGGACCAATTCGTATCAATTACAACACTTCAAACTGTCTCTGGAGCAGTTTCTATCGATAAACAACGACAAATATGTCTCTGGAGCAGATGCTGGCAGTTGGG

610 620 630 640 650 660 670 680 690 700 710 720 730 740 750  
AACTTATTTCTGAAATCCCGTTGGAGATGTAATAGAGATAAAAGGAACCGCTTTTGTGGAGATGCTGCCAAAGTGGCGCAACCTTTGAAAAAGATGTGATGTACAGGATGGAGAATCAATACCTGCTTGAAGATATAAATACC

760 770 780 790 800 810 820 830 840 850 860 870 880 890 900  
CTCAGGTTCCGACACAGTTTTACTGTGCAGGCAACAGTGGCCGATGACCTTAATGTGAGAGCAAAAGTTATCTGTGTTAAACGAAAGTGCAGAAAGGCACAAAACGGTACTGCTACTGTTCCAGAAAGGTTCTCCAGGAAACAT

910 920 930 940 950 960 970 980 990 1000 1010 1020 1030 1040 1050  
ATAAGATCAGGATCGATGGAAAGCAAGTGGCTCTAATGTCAAAATCAAGATCACAGCTGTCAAGAGGTGGAAGTAGATTGAGAAGAAATCTCAACTATGAATATAACTAAGTCTATACCGCAGGCAATTTTAAAGTAAAGTGC

1060 1070 1080 1090 1100 1110 1120 1130 1140 1150 1160 1170 1180 1190 1200  
GAGGCAATGAAAGCAGATTGAATATAAACTTCAGAGAGATTAATAAGCCTGCAGAAAGCTGCCTTTGGCAAAAGGCTCTTCTGTTATTTATGGATATCATATAATGTCTTCATAAAAAATCAATGCAGTCTCGAATGAAAA

1210 1220 1230 1240 1250 1260 1270 1280 1290 1300 1310 1320 1330 1340 1350  
TCATCTTATCCATAGAGAAATGGTGAAGGCCAAAATTAATTTGATAGCATTTCTTTACGCTTAAACGGATAACGTATGCAATCTTTTTAGCCCTCTTTTTTAAAGAAATAACACAGTATTGGACATCCCTCCGAAGACACAGTCCG

1360 1370 1380 1390 1400 1410 1420 1430 1440 1450 1460 1470 1480 1490 1500  
TATGTTTACAAAAAGAGTCTCGTGCAGAAATTTCTAGGATTGAATAAATAAAATTTATTATAACATTTAAACCCCTTGAATAATGATTTTTTTAAAGGATATTAACCTTATTTGCCAAATTAAGTGTAGTGTGCTCCTCTCC

1510 1520 1530 1540 1550 1560 1570 1580 1590 1600 1610 1620 1630 1640 1650  
CCGTAAGGTACATTATTAAAGAAATTTAATTTACTATTAAAGGAATTAAGTACTTTAATTTGGTAAATTAAGGTTAATTTTCTTACTTTTAAACATTAATTTTGTATGAATGAAATGATATCAGATGCCTTATTTCTAAATAACA

1660 1670 1680 1690 1700 1710 1720 1730 1740 1750 1760 1770 1780 1790 1800  
GAAATTTGAAAAAGAAATTAATTTCTCGTCTCTCTTCTGATTTAAAGACTATTATAGATTATTTTCTAAAAATGATTTTTAGAGTGTACCTTAATGGAAAAATCTAAATGCATACCTATTATCCAAAATAAGCATTGTATA

1810 1820 1830 1840 1850 1860 1870 1880 1890 1900 1910 1920 1930 1940 1950  
CCCTTAGTTTTAGTACTGATTTGGAAGGCCCTATATAACATTTGAATTAACCTGATTTTAGATTTTGCATAAATAAAAAATGTTAAAGTCAAATTAATATATTGAAATGATATTTCTTTATGGAAAGGGGAAACAAAGTA

1960 1970 1980 1990 2000 2010 2020 2030 2040 2050 2060 2070 2080 2090 2100  
ATAAAGAAATCCTGTTAAATGGGTTTATAAAGCACTGAAAGACACTATGAAGACACTATAATGAGTTCATTATAAAGACAATAATGAAATGATTTCCATTCATGTTTTCTTAGCGTTCTGTTCACTGTGGCTTTGTGAA

2110 2120 2130 2140 2150 2160 2170 2180 2190 2200 2210 2220 2230 2240 2250 (AtpD)  
TGAGGCAGATGTGAATCTTCATCATATTTAGTCTTACAAAAAGCTTTAGAAAAGGGGCACTCAAACCCGAATAATAATGCCAGTTTGGGGAGCTGTTAAATAGCAATATAATGCCAGCTTGGGGAGCTGTTAAATGTTAGA  
M V D

2260 2270 2280 2290 2300 2310 2320 2330 2340 2350 2360 2370 2380 2390 2400  
TCGAATCGGCTCTTGAATCTCGTAAAGTTCATCTCGGTGCGCGGCAGCATTGTAGATGATTTGTTGAGAAAGCACTTGCACCTGTATATACACTGCTACGTGCAGGAGAGAAAGCCAAATGCCATGAAAGTTTGACGCACTG  
R S R L L N L G K V I S V R G S I V D V L F E K H L P P V Y T L L R A G R E S Q I A I E V L T Q L D

2410 2420 2430 2440 2450 2460 2470 2480 2490 2500 2510 2520 2530 2540 2550  
CGCACATATGTCGCGGAAATGCTCTGACCCCACTGAAGGTCTTCTCGAGGTATGCGAGTGAAGATACAGGTGGACCTTTGAAAGCGCTTGGTGGGAAATCCTTTCTCGAATGTTGATGTTTGGAAATACCATAGACCG  
A H H V R G I A L T P T E G L A R G M A V E D T G G P L K A P V G R E I L S R M F D V F G N T I D R

2560 2570 2580 2590 2600 2610 2620 2630 2640 2650 2660 2670 2680 2690 2700  
CAGGAAGCTTCATCTGATATCCAGTGGCGCTCAATACATCAAAACCCGCCACCGCTAATGCGTGCATCACTACGCTGAAATATTCGAGACCGGTATCAAGCTATAGATGACTGGTACCGCTTGAAGCGGAGGTAAAGCAGGCT  
R K P P S D I Q W R S I H Q T P P L M R R S T T S E I F E T G I K A I D V L V P L E R G G K A G L

2710 2720 2730 2740 2750 2760 2770 2780 2790 2800 2810 2820 2830 2840 2850  
ATTCCGAGGAGCGGCGTTGGCAAGCAGTGTGCTAACCGAAATGATTATAACGTAGTCAAAACAACACCGAGGAGTGGCATCTTTGCGAAATAGGCGAGCGTGTGAGAAAGGGGAAAGAACTTTATCGTGATATGAAAGCAGCTGG  
F G G A G V G K T V L L T E M I H N V V K Q H Q G V S I F C G I G E R C R E G E E L Y R D M K D A G

2860 2870 2880 2890 2900 2910 2920 2930 2940 2950 2960 2970 2980 2990 3000  
TGTCTCCGAATACGTTATGGTATTCGGCCAGATGAACGAGCCAGGTCGCCCTTTCTGTAGTGTATACGCGCTGACAAATGGCAGAGTACTTCAGAGACGATGAACGCCGAGATGTTCTACTGCTTATAGATAACATTTTAC  
V L P N T V M V F G Q M N E P P G A R F R V G H T A L T M A E Y F R D D E R R D V L L L I D N I F R



6160 6170 6180 6190 6200 6210 6220 6230 6240 6250 6260 6270 6280 6290 6300  
TGACTGAAACAGTCTGGAAGAACCGGCAGTGTGATGTTTCAAGGCTCGCAATCTGCAAGTAATGAAAGAGCAGTTCAGCTCAGCGCTTAGTACATCACCAGGTCAGGTGCTCATCGTACTGATTTGATCTCCACAAA  
T E T S L E E R A V D V F V Q R L R M L Q D N E K K Q L A S A L S T S P G Q V L I R T A F D L P Q T

6310 6320 6330 6340 6350 6360 6370 6380 6390 6400 6410 6420 6430 6440 6450  
CGCAGCTGATTTATAAAAAGCGGATTAAGAAACTCTTGATATCGAACTCAACCAGGTTTGAGACTTTACCAGACTTCGTGAGTGAATTAATACAGACGGACAAAAGTTGCGTGAGGCAATTGATTTACAT  
Q R D S I K K A I K E T L D I E T Q P R F E T L P D L V S G I E L N T D G Q K V A W S I A D Y L T S

6460 6470 6480 6490 6500 6510 6520 6530 6540 6550 6560 6570 6580 6590 6600  
CGATGCAAAAAGTATCGACGAATTGCTGAACGAGCAGCTGAGAGCAATCCAAATCTGAGCCTCAAAACAAAATAAAGCTAAAACCGAACCTCAAAACAAAATAACAAAACAGGTTTAAAAACGAACGAGGCTGCAACAG  
M Q K S I D E L L N E Q P E S K S K S E P Q T K N N K A K T E P Q T K N N K P R F E N E P E A E Q E

6610 6620 6630 6640 6650 6660 6670 6680 6690 6700 6710 6720 6730 6740 6750  
AGCTGGAAGCTGAACTAATGAACCCAGACGAAACTGAGTCAGAAGTAAAACAGGCTGCACTGAGTATAATAATCCAGATCGAAAAAGAGCTAGAGACCGAAACAGAGTTTGAACAAAACAGGCTGAGGCTAGACAAG  
L E A E T N E P R P K T E S E V K T R P A T E Y N N P R S K K E L E T E Q E F D A K Q E S E A R Q E

(AtpA)  
6760 6770 6780 6790 6800 6810 6820 6830 6840 6850 6860 6870 6880 6890 6900  
AGCCTGATGCAAAACAGAGTCAGAGCTAAACAAAATCTGAAGCTGAACAAATGCTGATGCAAAACAAAGGTCAGAGCTAAAGCAAGAGTAAACAGCAATCTGGAATATGCTAATCTGTGCGCAAGTGAAGTAAAT  
P D A K Q E S E A K Q K S E A E Q M P D A K Q G S E A K Q E S E V K Q Q S G M S P N L W P R \*  
M R L I

6910 6920 6930 6940 6950 6960 6970 6980 6990 7000 7010 7020 7030 7040 7050  
CCAAACAGAGCCACATAAGAGGCAAGCAATGAAACTAAAAGCTTAAAGACTTTTCGATGAAAGTTTTCACGAAATACATCAGGTCAGGAATCAGTCACCCGACGCTTACCCCAAGGAGGTAGGCAAGTTCGACATG  
P N Q S P H K E P V N M E T K S L K D V F D K V F T E I H Q V R E S V T P T L T P K E V G R I L T V

7060 7070 7080 7090 7100 7110 7120 7130 7140 7150 7160 7170 7180 7190 7200  
TTCTACAGTATCGCAACGTATGGCTTCTACTGATGTTTGAAGCACTATAAAGTTTTCAGGGGATTTTGTGGATCGCATTTAATGTTGATGAAAAAGAAATAGTACCGTCTACTGGTGAATACTCACACCTGCATGC  
S T G I A N V S G L P T V G F D E L I K F P G D L F G I A F N V D E K E I G T V L L G E Y S H L H A

7210 7220 7230 7240 7250 7260 7270 7280 7290 7300 7310 7320 7330 7340 7350  
TGGAATCAGGTTGAACGCACAGAGCGGTTATGGATGTTGCCATGAGCAAGAAATGCTGGAGAGTATTAGCCACTAGGTCGCACTTGATAACAAAGGACCAATAGTTTCCAGTAACGCTTGCTATTGAACGCCCGAGTCC  
G D Q V E R T R R V M D V A V G E E L L G R V I D P L G R P L D N K G P I V S S K R L P I E R P S P

7360 7370 7380 7390 7400 7410 7420 7430 7440 7450 7460 7470 7480 7490 7500  
GGCAATATGATGATGAGTCTGTTACCGTGCCTCTCAGACCGGTGCAAGATTAATGATCGGTTAATACCAATAGGACGGGTGAGAGAGAAATGATTTTAGGGGACCGCAGACTGGCAAACTGCAATGCGATGACTATCTCT  
A I M D R A P V T V P L Q T G V K V I D A L I P I G R G Q R E L I L G D R Q T G K T A I A I D T I L

7510 7520 7530 7540 7550 7560 7570 7580 7590 7600 7610 7620 7630 7640 7650  
TAATCAGCGAAATTTAATGTCCTGTCGTTATTTGTCAGTTCAGGTCATCCGAGTTGCAAGAGCAGTGGCAAACTTACGGGAAAGAGGCGCAATGGATTAACAGATTGTTGTTGTAACGAAGCAATGACCATCAGGCGCT  
N Q R N F N V L C V Y C A I G Q R A S A V A R A V A N L R E R G A M D Y T I V V V T E G N D P S G L

7660 7670 7680 7690 7700 7710 7720 7730 7740 7750 7760 7770 7780 7790 7800  
AATCTATTAATCTTCTGTCGACAGCATTGCAAGATTTTATGGAATCAGTTCAGAGCTGCTAATGTTTACGATGACCTGACCAATATGCACTGCTTATCGAGAACTTCCCTTTTACTTGTGCGCTCTGGAAGGGA  
I Y I T P Y A A T S I A E Y F M E S G R D V L I V Y D D L T N H A R A Y R E L S L L L R R P P G R E

7810 7820 7830 7840 7850 7860 7870 7880 7890 7900 7910 7920 7930 7940 7950  
AGCCTATCCGGTGACATCTTTTATATCACTCAGCAGTACTGGAGCGGTCTACGACCTGCTCAAGAAATTTGTCGCGATCACTAACTGCTCTTCTTATATCGAAACCGAAGCGCAGAAATTTCCGCTTATATCCCACTAATCT  
A Y P G D I F Y I H S R L L E R S T H L L K E F G G G S L T A L P I I E T E A Q N I S A Y I P T N L

7960 7970 7980 7990 8000 8010 8020 8030 8040 8050 8060 8070 8080 8090 8100  
GATTTCAATTACGGAGCGGAGATTTACTCTTCAGCTTTCGCTTTTTCAGCTTTCGCTGCTTGTGTCGTCGAGTTCAGTTCGCGCAATCCGCTCTCTGTCAGGCGTAAAGCAGACTTGTGCTCATCGAGAGTGGCTGGAAGTCTCAAGCT  
I S I T D G Q I Y L S P S L F E L G V L P A V D V G K S V S R V G G K A Q L A A Y R E V V A G D L K L

8110 8120 8130 8140 8150 8160 8170 8180 8190 8200 8210 8220 8230 8240 8250  
TGCTATTTCACAATTTGAAGAGCTGGAACGTTTACCGGATTCGCGCAAGCTAGATGAAATACACGAAGACTCATAGAAATGCGGAGCGGATCCGCTCCTGAAAGCAGCTCATAAATTCCTGATCCGATCCTGATCAGAT  
A Y S Q F E E L E T F T R F G A K L D E N T R R L I E H G R R I R A L L K Q P H N S P V P V P D Q I

8260 8270 8280 8290 8300 8310 8320 8330 8340 8350 8360 8370 8380 8390 8400  
TGTTATCTTTGTCATTAACCGAAACTCTTCGATAATGTCGCTTGCAGAAATGTTAGAGGCTGAAACTCCCTAGCTAGGCAATGTCGAGATTTCTGAAGATGTGCGGAGAGGTTCAAGGCGATAAAGAAATGAGCGATAA  
V I L V A L N A K L F D N V P L D K M V E A E N S L R K A L P D I P E D V R E R F K G D K E L S D K

8410 8420 8430 8440 8450 8460 8470 8480 8490 8500 8510 8520 8530 8540 8550  
GGAGCTGAAACTCTCAATATCGCCCGCAAGCACTGGAACCTTACCAGCCAAAGCTGAATCGAATCAAGCCGAAAGCTAAAACGGAGGAAAGCGGAGTCAGAAACCAAACTGAGGCAAAATCTGAGTCAGAACTAAAAC  
D R E T I L N I A R K A L E P Y Q P K P E S E S K P E A K T E E N A E S E T Q T E D K S E S E A K T

8560 8570 8580 8590 8600 8610 8620 8630 8640 8650 8660 8670 8680 8690 8700  
TGAAGAAAAGTGGCGAAATCAGAGACTCAGACTAAGGGTAAAGCTGAGACAGAACTAAAATCAGGAAAACCTGAGTTGGAACCCGCAAGGAAAGAAAATAAGTAAATCAAAAGCCAGACTAAGGGCAGTCTGAGCCGAGACA  
E E K V A K S E T Q T K G K P E T E T K T E E K P E F G T R T E E K \*

(AtpG)  
8710 8720 8730 8740 8750 8760 8770 8780 8790 8800 8810 8820 8830 8840 8850  
AAGGAAAATCATGACCAGAACTCAGAGTTTGGCCACAAGATGATAGAGCAAAAGATCTTCACTGTCTGTCGCGCAGCATGAAAGCCCTGGCAGCTCGAACAATAGGGCAATGAAAATCAGTGAGTGGTGTCTGACTACT  
M T E T T Q S L R T K I D R A K D L Q S V V R T M K A L A A S H I G Q Y E K S V S A L S D Y Y

8860 8870 8880 8890 8900 8910 8920 8930 8940 8950 8960 8970 8980 8990 9000  
ATTATACAGTGAACCTGGGTTTGGCTTTATGTTTCAGGAAAATGCAATTCAGCCAGTCCACAGGAAAGGAAAGCGCAAAAGAAATACAGCTTATCGGAGCTGCTGTTTTCGGTTTCAGATCAGGGGCTTGGGTCAGTTAATGATA  
Y T V E L G L G L C F R K I A F M P A P Q E G K A Q K N T R L I G A V V F G S D Q G L V G Q F N D I

(AtpJ)  
9010 9020 9030 9040 9050 9060 9070 9080 9090 9100 9110 9120 9130 9140 9150  
TAATTACGATGAGCGGTCAGGAACTGAAAGCTTCCAGGTAAGGCTCAGTATGAGGCTGAGGAGGCTGTTATTCAGGCTGGAAGTGAAGGCTTCCCGCTTATGCACTTTAATAATGTGCCAATTCGTTAAGGCAATTC  
I T D Y A V R E L K A L P G K A Q V W A V G E L V I H A W M K M K A C R L L D F I M C Q I R L R Q L P  
M G C R G A C Y S R L E D E G L P L I G L Y N V P N S V K A I T

9160 9170 9180 9190 9200 9210 9220 9230 9240 9250 9260 9270 9280 9290 9300  
CCCCCTTATAGCCGAAATCTTGTGAAAATGAAAAATACCGCAGCGAGTGAAGATGCCGAACCTCATCTATTACAACCCCAAAAACAGAGTTACTATGAGCCAGTCAGCAGGACTGCTGCCATTTGATGAAACCTGGAG  
R L \*  
P L I A Q I L V E N E K L R S Q D E D A E L H L Y Y N R H K T R V T Y E P V S Q R L L P F D E T W R



9310 9320 9330 9340 9350 9360 9370 9380 9390 9400 9410 9420 9430 9440 9450  
AGACGACCTGACTAACTTTCCGCGCAACCAAAACTTGCCTGAGGTAATGGGCAATATCACAGAACTTACGGGCACTCATCCGCGAGTACCTTTTCGTTTCGTTTCAGGGCATGCGCTGAGTCCCTTGCAAGTAAAACTC  
D D L T K L S W P T K N L L P E V M G N I T E T L R A L I R E Y L F V S L F R A C A E S L A S E N S

9460 9470 9480 9490 9500 9510 9520 9530 9540 9550 9560 9570 9580 9590 9600  
TAGTCGACTGGCAGCAATGCGCGTCCGATAAAAACTTGAGGAGCTGTTAGTAATCTCAGTGGAGAACTACCGTGTGCCAGAGCGGCAATTGATGAGGAGTTGTTTGAAGTTGTTTCAGGTTTGAAGCTGTCCAGGGCCCG  
S R L A A M Q R A D K N I E E L L S N L S G E Y Y R V R Q S G I D E E L F E V V S G F E A L S R A R

9610 9620 9630 9640 9650 9660 9670 9680 9690 9700 9710 9720 9730 9740 9750  
TAGCTCTCGCCGGTAAAAAGCAGCGTGAGTCGGCAATGGTCCAGACAAAAATGATGAATCCGAAAAATGAGTTTATTGAGAAGTTTTGAGAGGCTGAAACGGGAAGAGATTTGTGAAATAAGAGAGACTTGTGAAATAAAAACT  
S S R R \*

9760 9770 9780 9790 9800 9810 9820 9830 9840 9850 9860 9870 9880 9890 9900  
TGTGAAATAAGAAAACTTGTGAAATAAGGGAGACTTGTGAAATAAGGAAAGCCGTTATATTAAAGAGAAGCATTGTGACATTCAGGAAATAATGTTGAATAAAGAGTTTTAGGGTCTGAAAAATCTCAAGTTTCTAAGATAT

9910 9920 9930 9940 9950 9960 9970 9980 9990 10000 10010 10020 10030 10040 10050  
CTAGGAGTCTGTGAAAAATTTCAAATTAACCTATTTAGATAATTTATTTTCTTAAAAAATATATATTAGTTGTTGACGCGATAGGTGCAGTAGACAGAGCTCTCAAGAAACAAGAGGCTCTTCTGAGCCTTTTGATTT

10060 10070 10080 10090 10100 10110 10120 10130 10140 10150 10160 10170 10180 10190 10200  
CTGGAATGGTAAAACTCGCAATCATTGTTTTCTGTTAGCGGGGTTTCGAGTTATGGAAATTCGAGCGGATACAATATAGGTTGAGAACTAAAAACAATTTTCTATTATGAGCCTGGCAAGATGCCTGATCTTCTTTTGA

10210 10220 10230 10240 10250 10260 10270 10280 10290 10300 10310 10320 10330 10340 10350  
CATGCATAATATTGCTTTCAGTTATATTATCAGGTGCCAAACCTGGCAACTGTAAACCGCAATAAGCTTCCCTTATTTCTCTAATTTTTTTCGTACATTTCCAGTATTATGGAGTTGTTTTGTGGAATCCCTGTACGAAA

10360 10370 10380 10390 10400 10410 10420 10430 10440 10450 10460 10470 10480 10490 10500  
CGAAGTAAAAATTCGGGCTTATTCAAGGAGTTTCTGGGAAAAGGATCCGCTCTTGTGCTGTGCGCTTAGCGATTGGTCTGGTACTTCTCGCGCTCTCGGAATATTCCTTCTCTTATACCCCTTATTCGTGAGGTATCCTCAGA

10510 10520  
CTTTCTTTGTCTGCAGTCTCGGAGG

付録B *M. barkeri* Fタイプ ATPase 遺伝子産物と既知のFタイプ構成成分との相同性

1st Amino Acid Sequence

File Name : atpA.aa  
Sequence Size : 588

2nd Amino Acid Sequence

File Name : atpA Moorella thermoacetica  
Sequence Size : 507

Unit Size to compare = 1

[51.3% / 86.6% / 464 aa]

```

1'  MRLIPNQSPHKEPVNMETKSLKDVFDKVFTEIHQVRESVTP TLT PKEVGRILT VSTGIAN
                               *** . * . ***
1"          MSIRPDEITSILKNQIEQYQLEVEMA EVGT V TQVGDGIAR

61'  VSGLPTVGFDELIKFP GD LFGIAFNVDEKEIGTVLLGEYSHLHAGDQV TRR VMDVAVG
   ** . . . . . * . . . . . * . . . . . * . . . . . * . . . . . * . . . . . * . . . . .
41"  IYGLDRAMAGELLEFP GD IYGMVLNLEEDNVGAVILGPYTHIKEGDQVKRTGRIVEVPVG

121' EELLGRVIDPLGRPLDNKGPIVSSKRLPIERPSPAIMDRAPVTVPLQTGVKVIDALIPIG
   * . . . . . * . . . . . * . . . . . * . . . . . * . . . . . * . . . . . * . . . . .
101" EALIGRVVNAMGQPIDGKGPIQTDKFRPVESPAGVVYRQP VNT PLQTGLKAIDSMVPIG

181' RGQRELILGDRQTGKTAIAIDTILNQRNFNVLCVYCAIGQRASAVARAVANLRERGAMDY
   * . . . . . * . . . . . * . . . . . * . . . . . * . . . . . * . . . . . * . . . . .
161" RGQRELIIGDRQTGKTAIAVDTIINQKQNVICIYVAIGQKASTVAGVVQRLEEAGAMEY

241' TIVVTEGNDP SGLIYITPYAATSIAEYFM-ESGRDVLIVYDDL TNHARAYRELSLLRR
   . *** . . . . . * . . . . . * . . . . . * . . . . . * . . . . . * . . . . . * . . . . .
221" IIVVMATASEPAPMLYIAPYAGCTMGEYFM YEQHRDVLCVYDDL SKHAAAYRELSLLRR

300' PPGREAYPGDIFYIHSRLLERSTHLLKEFGGSLTALPIIETEAQNI SAYIPTNLISITD
   * . . . . . * . . . . . * . . . . . * . . . . . * . . . . . * . . . . . * . . . . .
281" PPGREPYPGDV FYLHSGLLERPARL TDSLGGGSLTALPVIETQAGDVSAYIPTNVISITD

360' GQIYLSPSL FELGVLPAVDV GKS VSRVGGKAQLAAYREVAGDLK LAYSQFEELETFTFRFG
   * . . . . . * . . . . . * . . . . . * . . . . . * . . . . . * . . . . . * . . . . .
341" GQIFLESDFYAGQRPA INVALSVSRVGGAAQIKAMKQVAARLR L DLAQYRELA AFAQFG

420' AKLDENTRRLIEHRRIRALLKQPHNSPVPV PDQIVILVALNAK LFDNVPLDKMVEAENS
   .. * . . . . * . . . . . * . . . . . * . . . . . * . . . . . * . . . . . * . . . . .
401" SLDKATQARLARAERMMEILKQDQYQPMPVEEQVVVLYAAVNGFLDDL PVGRVRAFEKD

480' LRKALPDIPEDVRERFKGDKELSDKRETILNIARKALEPYQPKPESESKPEAKTEENAE
   . . . . . * . . . . . * . . . . . * . . . . .
461" FLRFLRNERPEVLAGIREKRQLDDNLQEQLKKSIEDFKGSFTAAGES

```

1st Amino Acid Sequence

File Name : atpB.aa  
Sequence Size : 190

2nd Amino Acid Sequence

File Name : atpB Moorella thermoacetica  
Sequence Size : 246

Unit Size to compare = 1

[36.0% / 79.1% / 172 aa]

```

1'                               MRLSPDELIFWQYDFIKLNATIVYTWGLMLVMV
                                * . . * . * * * * . * . . . .
1" MTGENDANKTRANSLATINMGLRALGEIMTHVRPVEIFHLGPIPIYSTVVNTWIIMILL
34' IGSKTITSKLSTGLERSRWQNILEIIVTGILEQIEDVGLDQPRKYLGF LGTLFLFIAPAN
   *   * . * * *   . * . * . * * * . . . . * . . . . * . * *   . . . * * * . * *   *
61" AGIFLATRKL SF-IPRGA-QHVLEMFFLEFFYGLLEEIIIGKEGRRYLPLVATL FIFILSLN
94' LCIIIPGYEPPTGSLSTTAALALCVFVAVPLFGIEEQGISNYLKTYTEPTIIMLPFNIIS
   * . * * *   . * * *   . * * * * * * * . * . . . * . * * *   . * . . . * . . . * . * * .
119" LSWFIPGMKPPTMDLSTTAFAVTTIILVQIFGIRKLG LRGYIRHFFQPAPFLFPLNVIE
154' EISRTLALAIRLFGNIMSGSMIVAILLTITPFIFLIL
    * . . . . * . * * * * . . . . * . * * *   . * * * . . * * . .
179" ELVKPVSLSLR LFGNLFGEEMVVTILFLMIPFLLPTPIMLLGVLMTIQAFVFTLLTITY

```

1st Amino Acid Sequence

File Name : atpC.aa  
Sequence Size : 144

2nd Amino Acid Sequence

File Name : atpC Haemophilus influenzae  
Sequence Size : 151

Unit Size to compare = 1

[26.9% / 67.7% / 130 aa]

```
1' MNSGLMNSGLMNSGLMNLTI LLPFQVFAEKKGVSRIVAEGREGSFGLLPHRLDCVATLGP
      **** . . * . . . * . . * . . * . . * . . * . . * . . * . . * . . * . . * . . *
1"   PSWISSPRTMATFNLTIVSAEQKIF-EGEVKQIQVTGVEGELGILPGHTPLLTAIKP

61' GILTYETESEGEVYVAVDEGVLIKNGQYVLVSVRDAIFGTDL--SQLHEAVEK--EFTL
      ** . . . . . * . . * . . * . . * . . * . . * . . * . . * . . * . . * . . *
57" GIVKFTLKDGN E E V I V S G G F L E V Q P N I V T V L A D I A I R G S E L D A D R I H E A K R K A E E N I V S

117' DETEQKIRSVMVKLETGLIRRLAEFQNV
      . . . . . . . * . . * . . * . . * . . * . . * . . * . . * . . * . . *
117" RGS D A D H D L L V A K L S K E L A - K L R A Y E L T E K L L K T R R
```



1st Amino Acid Sequence

File Name : atpE.aa  
Sequence Size : 91

2nd Amino Acid Sequence

File Name : atpE 0. neapolitana plastid  
Sequence Size : 82

Unit Size to compare = 1

[48.7% / 90.8% / 76 aa]

```
1' MALDTYITTTI AVASIATSGITIGIGVIGPAIGEGRAVATALSSLAQQPDASATITRTL FV
   .. **.....*..**.....***.**.* *.* *...**.***.*...* **..
1"  MNPIVSGASVVAAGLAIGLAAIGPGIGQGTAQAQAVEGLARQPEAEGKIRGTL L L
61' GLAMIESLAIYCFV VSMILIFANPFWRAL T
   .**..***.** **...*.*****
56" SLAFMESLTIYGLVVALCLLFANPFAG
```

1st Amino Acid Sequence

File Name : atpF.aa  
Sequence Size : 386

2nd Amino Acid Sequence

File Name : atpF Enterococcus hirae  
Sequence Size : 174

Unit Size to compare = 1

[26.0% / 76.0% /150 aa]

```
1'          MLIDWFTVIAQVINFLILVWLLKRFLYKPILNAVDARENKVANELKNADA
              . * . . . . * * * * . * * * * . * * . . . * * . * * * * . * .
1" MLNQLAIAEVGNPMLGNIIVVSGSFLILMFLFKHFAWGPISDILKKREDKIANDLDSA EK
51' KEAEAQKEKEEFKRKNEEFDQQRNDLLNKAKEEAQAERQRLFEDVREEASYLRAKQQEAL
      . . . . * . . * . . . . * . . . . * * * . . . * * . . . . * * . . . * . . . * . . .
61" SRINSAKMEQEREQQLLASRSDAADI IKNAKESGELSRQNILKETQEEVARLKSKAQTDI
111' RNEGQNL SQEIGRQTQQEVFSIARKILDTLTETSLEERAVDVVFVQRLRNLQDNEKKQLAS
      . * . . . . * * * * . . . . * . . . . *
121" MLERDTALNSVKDDVADLSLQIAAKILNKELSPEMHESLINQYIEGLGSSNETR
```



1st Amino Acid Sequence

File Name : atpG.aa  
Sequence Size : 149

2nd Amino Acid Sequence

File Name : atpG Arabidopsis thaliana chl  
Sequence Size : 373

Unit Size to compare = 1

[23.4% / 64.0% / 111 aa]

```

1'                                     MTETTQSLRTKI
1" MACSNLTTMWVSSKPSLSADSSLSFRSVLKCPNTSSPPSRASSVSPLQASLRÉLRDRI
13' DRAKDLQSVVRTMKALAASNIGQYEKSVSALSDDYYTVELGLGLCFRKIAFMPAPQEGKA
   *..*..*... ** ..**... ..**..*..*..*..*..*..*..*..*..*..*..*..*
61" DSVKNTQKITEAMKLVAAAKVRRRAQEAVVNGRPFSETLVEVLYNINEQLQTDVPLTK
73' QKNTRLIGAVVFGSDQGLVGQFNDIITDYAVRELKALPGKAQVWAVGELVIHAWKMKACR
   ..**..**..**..**..**..**..**..**..**..**..**..**..**..**..**..**
121" VRPVKKVALVVVTGDRGLCGGFNNFIIKKAEARIKELKGLGLETVISVGKKGNSYFLRR
133' LLDFIMCQIRLRQLPRL
181" PYIPVDKYLEAGTLPTAKEAQAVADDVFSLFISEEVDKVELLYTKFVSLVKSEPVHTLL

```



1st Amino Acid Sequence

File Name : atpJ.aa  
Sequence Size : 186

2nd Amino Acid Sequence

File Name : atpJ Bacillus subtilis  
Sequence Size : 290

Unit Size to compare = 1

[31.2% / 70.8% / 144 aa]

```

1'                                                                 MGC
61" RVSGNVKHPMLLSREVKKTAYLVITSDRGLAGAFNSSVLR SAYQAMQERHQS KDEYAVIA
 4' RGACYSRLEDEGLPLIGLYNVPNSVKAITPLIAQILVENEKLR SQDEDAELHLYNRHKT
    * . . . . * . . . . * . . . . * . . . . * . . . . * . . . . * . . . . *
121" IGRVGRDFFKKREIPII SELTGLGDEVTFTEIKDLARQTIQMFIDGAFDELHLVYNHFVS
64' RVTYEPVSQRLLPFDETWRDDLTKLSWPTKNLLPEVMGNITETLRALIREYLFVSLFRAC
  * * . . . . * . . . . * . . . . * . . . . * . . . . * . . . . * . . . . *
181" AITQEVTEKLLPL-----SDLGSGG-GKRTASYEFEPSEEEVLEVLLPQYAESLIFGAL
124' AESLASENSSRLAAMQRADKNIEELLSNLSGEYYRVRQSGIDEELFEVVS GF EALS RARS
  * * * * . . . * . * * * . * * * . * * * . * * * . * * * . * * * . * *
235" LDSKASEHAARMTAMKNATDNAKELIDSLSLSYNRARQAAITQEITEIVGGAAALE

```



付録C VおよびFタイプ、触媒および非触媒サブユニットのアライメント

```

+           +           +           +           +           +           +
vvr_Ncr -----MVLSDKE
vvr_Sce -----MALRAMRGIVNGAAPELVPVPTSGLA
vvr_Bta -----MNAQQ
vvr_Dme -----MAP
vvr_Hvu -----MLSKQE
vvr_Tco -----
avr_Mba -----
avr_Mma -----
avr_Hsa -----
avr_Hvo -----
avr_Mja -----
evr_Ehi -----
evr_Tth -----
avr_Sac -----
cfr_Osa -----MATLRVDEIHKI-----LR
cfr_Zma -----MATLRVDEINKI-----LR
cfr_Tae -----MATLRVDEIHKI-----VR
cfr_Nta -----MVTIRADEISNI-----IR
cfr_Sol -----MATIRADEISKI-----IR
cfr_Psa -----MVTSRADEISQI-----IR
cfr_Mpo -----MVNIRPDEISSI-----IR
cfr_Cpa -----MVSIRPDEISSI-----IR
cfr_Ppu -----MVNIRPDEISSI-----IR
efr_Syn -----MVSIRPDEISSI-----IR
efr_Bsu -----MSIKAEIISTL-----IK
efr_Ehi -----MAIKAEIISAL-----IK
efr_Pmo -----
mfr_Osa -----MEFSPRAAELTTL-----LE
mfr_Zma -----MEFSPRAAELTTL-----LE
mfr_Bta -----MLSVRYAAAVARALPRRAGLV-SKNALGSSFIAARNLHASNSRLQKTGTAEVSSI-----LE
mfr_Hsa -----MLSVRYAAAVRALPRRAGLV-SRNALGSSFIAARNFHASNTHLQKTGTAEVSSI-----LE
mfr_Dme -----MSIFSARLASSVARNLPKAAANQVACKAAYPAASLAARKLHVASTQ---RSAEISNI-----LE
mfr_Spu -----MAQVSVRLAAAAARSLSRQAPQVA-KRCIGAAVYSHRNVTTSAPR---PSTAEVSSI-----LE
mfr_Ncr -----MFRNALRQSTRAVGAFSATGRVAARNAAPVVSAYQARTYADA---KATPTEVSSI-----LE
efr_Rbl -----MGIQA-----AEISAI-----LK
mfr_Cre MRSQALSLARAGLLQLSSQTGASLEGGFALSKRAEQALIRASRAFASDAKALDELKPKFTSKYLINHVS
efr_Sli -----MAELTIRPEEIRD-----ALE
efr_Eco -----MQLNSTEISEL-----IK
afn_Mba -----MRLIPNQSPHKEPVMNETKSLKDVFDKVF
cfc_Osa -----
cfc_Zma -----
cfc_Tae -----
cfc_Nta -----
cfc_Sol -----
cfc_Psa -----
cfc_Mpo -----
cfc_Pth -----
cfc_Cpa -----
cfc_Ppu -----
efc_Syn -----
mfc_Osa MATRRALSSLVRAA-----SRLRGASP-APRPRGPLHRPSPSGYLFNRAAAAYATAAAAKERP--PAP
mfc_Zma MASRRVYSSLLRSA-----SRLRAASPAAPRPRAPPHRPSAGYLFNRAAAAYASSAAAQAAATPPP
mfc_Nta MASRRLASLLRQSAQRGGGLISRSLGNSIPKSASRASS-RASPKGFLNRAVQYATSAAPASQPSTPP
mfc_Hbr MASRRLSSLLRSSSRSSVSKSPISNINPKLSSSSPSSKSRASPYGYLLTRAAEYATSAAAAAPPQPPPA
mfc_Bta -----MLGLVGRVVAASASGALRGLSPSAPLPQAQLLLRAAPAA
mfc_Rno -----
mfc_Sce -----MVLPRLYTATSRAAFKAAKQSAPLLSTSWKRCMASAA
efc_Rbl -----
mfc_Ncr -----MFKSGISAFARTARPSFAAASRRRAVRPAALNLRAPALSR
efc_Eco -----
efc_Pmo -----
efc_Bsu -----
efc_Ehi -----
efc_Sli -----
afc_Mba -----
vvc_Ncr -----MAPQ
vvc_Sce -----MAGA
vvc_Hvu -----
vvc_Zma -----
vvc_Bta -----
vvc_Dme -----
vvc_Tco -----
avc_Mja -----
evc_Ehi -----
avc_Mba -----
avc_Mma -----
avc_Hsa -----
avc_Hvo -----
evc_Tth -----
avc_Sac -----

```

```

+ + + + +
vvr_Ncr ---MADPRVPSYVNTPRIRYNTVGGVNGPLVILDNVKFPRYN EIVTL-----TLPDGTQRSQGVVL
vvr_Sce LFAINKKAVEQGFNVKPRLNYNVTSGVNGPLVILEKVKFPRYN EIVNL-----TLPDGTVRQSGQVL
vvr_Bta GSREQALAVSRNYSQPRLTYKTVSGVNGPLVILDHYKFPRYA EIVHL-----TLPDGTKRSQGVVL
vvr_Dme AQREHVLAVSRDFISQPRLTYKTVSGVNGPLVILDEVKFPKFA EIVQL-----RLADGTVRSQGVVL
vvr_Hvu EMEEGTLEIG-----MEYRTVSGVAGPLVILDKVKGPKYQ EIVNI-----RLGDGTTTRRGQVL
vvr_Tco LLTEHINELRKDYVVKPRLEYTTIRAVNGPLVILGDVRLPTFA EIVNI-----ELSDGTMRRGQVL
avr_Mba -----MVKEYKTIITQIAGPLVVFVEKTEPVGYK EIVTI-----NLPDGTTRRGQVL
avr_Mma -----MVKEYKTIITQIAGPLVVFVEKTEPVGYN EIVNI-----KMGDGTVRRGQVL
avr_Hsa -----MKEYQTITEVSGPLVYVETDEPIGYD EIVQI-----ETPNGDVKRGQVL
avr_Hvo -----MKEYQTITEVSGPLVFAEVEDEPIGYD EIVEI-----ETPQGDKTRGQVL
avr_Mja -----MATAASAIEYSVKSIAAGPLLVIEGVEGAAYG EIVEV-----ICPDGGEKRMGQVL
evr_Ehi -----MIKEYRTIKEYVGPLMAVEKVSQVKYE E LIEV-----RMQNGEIRRQGVVL
evr_Tth -----MDLLKKEYTGITYISGPLLVFNENAKDLAYGATVDI-----KDGTRVRGGQVI
avr_Sac -----MSLLNVREYSNISMIGPLVIAVQGVSDAAYN E LVEI-----EMPDGSKRRGLVV
cfr_Osa ERIEQYNRKVG-----IENIGRVVQVGDGIARIIGLGEIMSG E LVEF-----AEGTRGIAL
cfr_Zma ERIEQYNRKVG-----IENIGRVVQVGDGIARIIGLGEIMSG E LVEF-----AEGTRGIAL
cfr_Tae ELIEQYNRKVG-----IENIGRVVQVGDGIARIIGLGEIMSG E LVEF-----AEGTRGIAL
cfr_Nta ERIEQYNREVK-----IVNTGTVLQVGDGIARIHGLDEVMAG E LVEF-----EEGTIGIAL
cfr_Sol ERIEQYNREVK-----VYNTGTVLQVGDGIARIHGLDEVMAG E LVEF-----EEGTIGIAL
cfr_Psa KRIEQYNTEVK-----IVNTGTVLQVGDGIARIYGLDVMAG E LVEF-----KEGTVGIAL
cfr_Mpo KQIEQYNQEVK-----IVNIGTVLQVGDGIARIYGLDKVMAG E LVEF-----EDGTVGIAL
cfr_Cpa QQIEQYDQEIQ-----VSNVGTVLQVGDGIARVYGLDKVMSG E LLEF-----EDGTIGIAL
cfr_Ppu QQIEKYDQDVE-----VANIGTVLQVGDGIARVYGLDEVMAG E LLEF-----EDKTIGVAL
efr_Syn QQIEQYEQSIN-----VDNVGTVLQVGDGIARVYGLDKVMAS E LVEF-----EDGTVGIAL
efr_Bsu QQIQNYQSDIE-----VQDVGTVLQVGDGIARVYHGLDNCMAG E LVEF-----SNGVLGMAQ
efr_Ehi EQIENYQNVLS-----VEEIGTVTVYVGDGIARAHGLENAMSG E LLEF-----SNGSYGMAQ
efr_Pmo -----MAL
mfr_Osa SRMNFYTNFQ-----VDEIGRVVSVGDGIARVYGLNEIQAG E MVEF-----ASGVKGIAL
mfr_Zma SRMNFYTNLK-----VDEIGRVVSVGDGIARVYGLNEIQAG E MVEF-----ASGVKGIAL
mfr_Bta ERILGADTSVD-----LEETGRVLSIGDGIARVHGLRNVQAE E MVEF-----SSGLKGMSL
mfr_Hsa ERILGADTSVD-----LEETGRVLSIGDGIARVHGLRNVQAE E MVEF-----SSGLKGMSL
mfr_Dme ERILGVAPKAD-----LEETGRVLSIGDGIARVYGLNNIQAE E MVEF-----SSGLKGMAL
mfr_Spu ERILGHTPKEN-----LEETGRVLSIGDGIARVYGLKNIQAE E MVEF-----SSGLKGMAL
mfr_Ncr QRIRGVQEEESN-----LAETGRVLSVGDGIARVHGMANVQAE E LVEF-----ASGVKGMCM
efr_Rbl EQIKNFQKDAE-----VAEYGRVLSVGDGIARVYGLDNVQAG E MVEF-----PGGIRGMAL
mfr_Cre EKLIPAVKWEKQYQPPVIHLGRVLSVGDGIARVYGLKSVQAG E LVCF-----DSGVKGMAL
efr_Sli NfVQSYKPDAAASR-----EEVGTVTLAGDGIAKVEGLPSAMANE L LKF-----EDGTLGLAL
efr_Eco QRIAQFNVVSE-----AHNETIVSVSDGVIIRIHGLDCMQG E MISL-----PGNRYAIAL
efr_Mba TEIHQVRESVPTLTP--KEVGRILTVSTGIANVSGLPTVGFDE L IKF-----PGDLFGIAF
cfc_Osa -MRTNPTTSRPGVSTIEEKSTGRIDQIIGPVLVDVTFP-RGKLPYIYNALVVKSRDTDGKQ--INVTCQEVQ
cfc_Zma -MRTNPTTSRPGISTIEEKSVGRIDQIIGPVLVDITFP-PGKLPYIYNALIVKSRDADKQ--INVTCQEVQ
cfc_Tae -MRTNPTTSPPGASTIEEKSTGRIDQIIGPVLVDVTFP-PGKLPYIYNALVVQSRDTDKQ--INVTCQEVQ
cfc_Nta -MRINPTTSKSGVSTLEKKNPGRVYQIIGPVLVDVAFP-PGKMPNIYNALVVQGRDSVGPQ--INVACQEVQ
cfc_Sol -MRINPTTSDPGVSTLEKKNLGRIAQIIGPVLVDVAFP-PGKMPNIYNALIVKGRDTAGQP--MNVTCQEVQ
cfc_Psa -MTITPPPSDTEVSVLENKKNLGRITQIIGPVLVDVFP-PGKMPYIYNALIVQGRDVTGKQ--INVTCQEVQ
cfc_Mpo -MKTNFLAF--GMSTLVAKNIGSIIQVIGPVLVDVAFS-PGKMPNIYNALIVKQNSAGEE--INVTCQEVQ
cfc_Pth -MRKNPLVL--GVSASVETNVGRIAQIIGPVLVDVFP-PGNMNIYNALIVKQGGTAGQE--IQVTCQEVQ
cfc_Cpa -----MATTTSKNTGVTQVIGPVLVDVFP-NGQLPKIYNALIVKQNEAGQD--ITVTCQEVQ
cfc_Ppu -----MVSSTKS--TGSVITQIIGPVLVDVAFP-NGQLPKVFNALVQST--D--GTITCQEVQ
efc_Syn -----MVTTAERTNVGFTQVIGPVLVDVFP-SGKMPATYNALRIVQGNAGLD--VAVTCQEVQ
mfc_Osa ATGKATGGGKITDEFTGAGAIQVQCVIGAVVDVRFD-EG-LPPIILTALEVL--DHN--IRLVLEVA
mfc_Zma ATGK-TGGGKITDEFTGAGAIQVQCVIGAVVDVRFD-EG-LPPIILTALEVL--DNN--IRLVLEVA
mfc_Nta KSG-SEPSGKITDEFTGAGSIIKVCQVIGAVVDVRFD-EG-LPPIILTALEVL--DNQ--IRLVLEVA
mfc_Hbr KPEGGKGGGKITDEFTGKGAIQVQCVIGAVVDVRFD-EG-LPPIILTSLEVL--DHS--IRLVLEVA
mfc_Bta LQPARDYAAQASPSKAGATTGRIVAVIGAVVDVQFD-EG-LPPIILNALEVQ--GRE--TRLVLEVA
mfc_Rno -----EFRTATQIVAVIGAVVDVQFD-EG-LPPIILNALEVQ--GRE--SRLVLEVA
mfc_Sce -----QSTPITGKVTAVIGAVVDVHFE-QSELPAIILNALETK--TPQ--GKLVLEVA
efc_Rbl -----MTMATTVSKGVTQVIGAVVDVQF-EGVLPAILNALETT--NNG--KKLILEVA
mfc_Ncr F-----ASSAGVGGKIYQVIGAVVDVDFD-TDKLPPILNALETO--NNG--QKLVLEVS
efc_Eco -----MATGKIVQVIGAVVDVQF-QDAVPRVYDALEVQ--NGN--ERLVLEVS
efc_Pmo -----MENGVITQIIGPVLVDVTFE-NE-LPRIYNALIKIDR--GNG--EYLVAEVQ
efc_Bsu -----MKGRVSVQVLPVVDVRFD-DGHLPEIYNALIKISQPAASENEVGIDLTLEVA
efc_Ehi -----MSSGKIVQVIGPVDVVEFSLDQSLPDINNALVVYK--NDEKKSXVVLETA
efc_Sli -----MTTIVETATATGRVARVIGPVDVVEFPVDA-MPEIYNALHVEVADPAKEGELKTLTLEVA
afc_Mba -----MVDRSRLNNGKVISVRSVGSIVDVLV--EKHLPPVYTLTRA--GRESQIAIEVL
vvc_Ncr QNGA-----EVDGIHTGKIYSVSGPVVVAEDMIGVAMYELVKV--GHDQLVGEVI
vvc_Sce IENARKEIKRISLEDHAESEYGAISVSGPVVIAENMIGCAMYELVKV--GHDNLVGEVI
vvc_Hvu -----ELVRV-----GHDSLIGETI
vvc_Zma -----MDFSKLPKIRDEDKESTFGYVHGVSGPVVTACDMAGAAMYELVRV-----GHSELVGEII
vvc_Bta -----MSNLRKFKDEERESEYGRVYAVSGPVVSAEAMSGSAMYELVRV-----GYVELVGEII
vvc_Dme -----MTSDKNPYKTEQRMGAVKAVSGPVVIAENMGSAMYELVQV-----GSFRLVGEII
avc_Mja -----MKREVENMPVVGKIICKIAGPVVVAEAGMKGAQMYEVVKV-----GEEKLTGEII
evc_Ehi -----MQIGKIKIVSGPLVMAENMSEASIQDMCLV-----GDLGVIGEII
avc_Mba -----MEVNGEYRVSQPVVTAIG-LQAKMYDLVKV-----GNEGLMGEVI
avc_Mma -----MEVNGEYRVAQPVVTAIG-LQAKMYDLVKV-----GNEGLMGEVI
avc_Hsa -----MSQA--EAITDTGEIESVSGPVVTAIG-LQAKMYDLVKV-----GDEGLMGEVI
avc_Hvo -----MSQATQDSVREDGVIASVSGPVVTAIG-LDARMNDVYVYV-----GDEGLMGEVI
evc_Tth -----MIQGVYQKIAGPVAIGKGLMARMYDICKV-----GEEGLVGEII
avc_Sac -----MVSEGRVYRVNGPLVIAAGMREAQMFVYVYV-----SDLKLVEIT

```

```

vvr_Ncr  EARGNRAVQVFEFGTSGIDVKKTKVEFTGESLKLGVSEDM LGRIFDGSGRATDKGPKVLAE-----
vvr_Sce  EIRGDRAIVQLFEFGTSGIDVKKTTVEFTGESLRIPVSEDM LGRIFDGSGRPIDNGPKVFAE-----
vvr_Bta  EVSGSKAVVQVFEFGTSGIDAKKTCCEFTGDILRTPVSEDM LGRVFNNGSGKPIDRGPVYLAE-----
vvr_Dme  EVSGSKAVVQVFEFGTSGIDAKNTLCEFTGDILRTPVSEDM LGRVFNNGSGKPIDKGPVYLAE-----
vvr_Hvu  EVDGEKAVVQVFEFGTSGIDNKYTTVQFTGEVLKTPVSLDM LGRIFNNGSGKPIDNGPPVLAE-----
vvr_Tco  EVDGDKAVVQVFEFGTSGIDVMRSKCEFTGKVMELGVSEDM LGRVFNNGSGKPIDDGPVYLAE-----
avr_Mba  DSSSDIVVIQIFEGTIGLD-KECGVVFGETLKLPAASIDL LGRILSGSGEPLDGGPRIVPD-----
avr_Mma  DSSLTWLRSRFSKVLVGLT-RTAVYVIFGETLKLPAASVDL LGRILFGSGEPRDGGPRIVPD-----
avr_Hsa  ESSDGFVAIQVFEFGTEGVG-KDASVRFLGETLKMPTEDL LGRVLDGSGNPTDGGPDIVPD-----
avr_Hvo  ESSEGLVAIQVFEFGTSGID-RNASVRFLGETLKMPTEDL LGRVLDGSGQPTDGGPEIVPD-----
avr_Mja  EAREGLAVVQVFEFGTIGLSTKOTRVRFTGRTAKIGVSMEM LGRIFNNGAGKPIDGGPEIVPE-----
evr_Ehi  EVQEDKAMVQVFEFGTSGINLKNSSVRFLGHPQLQLVSEDM IGRVFDGLGRPKDNGPEILPE-----
evr_Tth  EVSEEYAVIQVFEETIGLDLATTSSVSWVEDVARLGVSKEM LGRRFNGIGKPTDGLPITPE-----
avr_Sac  DSQMGVTFVQVFEFGTIGISPTGSKVRFLGRGLEVKISEEM LGRIFNPLGEPLDNGPPVIGG-----
cfr_Osa  NLESKNVGIIVLMDGLMIQ-EGSFVKATGRIAQIPVSEAY LGRVINALAKPIDGRGEIVAS-----
cfr_Zma  NLESKNVGIIVLMDGLMIQ-EGSFVKATGRIAQIPVSEAY LGRVINALAKPIDGRGEIVAS-----
cfr_Tae  NLESKNVGIIVLMDGLMIQ-EGSFVKATGRIQIPVSEAY LGRVINALAKPIDGKGEIAS-----
cfr_Nta  NLESNNVGVVLMGDGLLIQ-EGSSVKATGRIAQIPVSEAY LGRVINALAKPIDGRGEISAS-----
cfr_Sol  NLESNNVGVVLMGDGLMIQ-EGSSVKATGRIAQIPVSEAY LGRVINALAKPIDGRGEITAS-----
cfr_Psa  NLESKNVGVVLMGDGLMIQ-EGSSVKATGRIAQIPVSEAY LGRVINALAKPIDGKGEISTS-----
cfr_Mpo  NLESNNVGVVLMGDGLTIQ-EGSSVKATGKIAQIPVSDAY LGRVINALAQPIDGKQIPAS-----
cfr_Cpa  NLEADNVGVVLMGDGRNIL-EGSSVRATQKIAQIPVGDVAV IGRVVDALARPIDGKGDIAAT-----
cfr_Ppu  NLEADNVGVVLMGDGRDIL-EGSSVKGTKIAQIPVGDVAV IGRVVDALARPIDGKGEIPASN-----
efr_Syn  NLEEDNVGVVLMGAGLGLIE-EGSTVRATGKIASVPVGEAV IGRVVDALMRPIDGKGEIHAT-----
efr_Bsu  NLEESNVGIVTLGPFSEIR-EGDIVKRTGRIMEVVPVEEL IGRVINPLQPPVDGLGILTS-----
efr_Ehi  NLESNDVGIIILGDFESIR-EGDKVKTGKIMEVVPVGDAL IGRVINPLGQPTDGLGEIKTE-----
efr_Pmo  NLEENNVGAVILGDPITGVK-EGDEVKATQIAAVGAGEAL IGRVINPLGEPLDGGELKTE-----
mfr_Osa  NLENENVGIVVFGSDTAIK-EGDLVKRTGSIVDVPAGKAM LGRVVDALGVPIDGKGLSDH-----
mfr_Zma  NLENENVGIVVFGSDTAIK-EGDLVKRTGSIVDVPAGKAM LGRVVDALGVPIDGKGLSDH-----
mfr_Bta  NLEPDNVGVVVFVFGNDKLIK-EGDIVKRTGAIIVDVPVGEEL LGRVVDALGNAIDGKGPVIGSK-----
mfr_Hsa  NLEPDNVGVVVFVFGNDKLIK-EGDIVKRTGAIIVDVPVGEEL LGRVVDALGNAIDGKGPVIGSK-----
mfr_Dme  NLEPDNVGVVVFVFGNDKLIK-EGDIVKRTGAIIVDVPVGEEL LGRVVDALGNAIDGKGAINTK-----
mfr_Spu  NLEPDNVGVVVFVFGNDKLIK-EGDIVKRTGAIIVDVPVGEEL LGRVVDALGNAIDGKGPVIGTE-----
mfr_Ncr  NLEAGQVGVVLFVFGSDRLVK-EGETVKRTGEIVDVPVPEL IGRVIDALGNPIDGKGPVINC-----
efr_Rbl  NLEVDNVGIVTFGDDRDIL-EGDIVKRTGSIVDVPAGNGL LGRVVDALGNPIDGKGPVIVAS-----
mfr_Cre  NLQADHVGVVVFVFGNDSLIH-QGDLVYRTGIVNVVPGPTL IGRVVDALGQPIDGKGPVINTV-----
efr_Sli  NLEEREIGCVVLFVFGSIE-EGQPVSRTEGVLVAVAEGL IGRVVDALGNPIDGLGEIETS-----
efr_Eco  NLERDQSVGAVVMGPYADLA-EGMKVKTGRILEVVPVGRGL LGRVVDALGAPIDGKGPVLDH-----
avr_Mba  NVDEKEIGTVLLGEYSHLH-AGDQVETRRVMDVAVGEELL IGRVIDPLGRPLDNKGPVIVSS-----
cfr_Osa  QLLGNRRVRAVAMSAIDGLMRGMEVIDTGAPLSVPVGGATL IGRIFNVLGEPLDNLGVPVDT-----
cfr_Zma  QLLGNRRVRAVAMSAIDGLMRGMEVIDTGAPLSVPVGGATL IGRIFNVLGEPLDNLGVPVDT-----
cfr_Tae  QLLGNRRVRAVAMSAIDGLMRGMEVIDTGAPLSVPVGGATL IGRIFNVLGEPLDNLGVPVDT-----
cfr_Nta  QLLGNRRVRAVAMSAIDGLMRGMEVIDTGAPLSVPVGGATL IGRIFNVLGEPLDNLGVPVDT-----
cfr_Sol  QLLGNRRVRAVAMSAIDGLMRGMEVIDTGAPLSVPVGGATL IGRIFNVLGEPLDNLGVPVDT-----
cfr_Psa  QLLGNRRVRAVAMSAIDGLMRGMEVIDTGAPLSVPVGGATL IGRIFNVLGEPLDNLGVPVDT-----
cfr_Mpo  QLLGNRRVRAVAMSAIDGLMRGMEVIDTGAPLSVPVGGATL IGRIFNVLGEPLDNLGVPVDT-----
cfr_Pth  QLLGNRRVRAVAMSAIDGLMRGMEVIDTGAPLSVPVGGATL IGRIFNVLGEPLDNLGVPVDT-----
cfr_Cpa  QLLGDNQVRAVAMSAIDGLMRGMEVIDTGAPLSVPVGGATL IGRIFNVLGEPLDNLGVPVDT-----
cfr_Ppu  QLLGDNQVRAVAMSAIDGLMRGMEVIDTGAPLSVPVGGATL IGRIFNVLGEPLDNLGVPVDT-----
efc_Syn  QLLGDNQVRAVAMSAIDGLMRGMEVIDTGAPLSVPVGGATL IGRIFNVLGEPLDNLGVPVDT-----
mfc_Osa  QHLGENMVRTIAMDGTGLVGRQVRLNTGSPITVPVGRATL IGRIMNVIGEPLDEKGDITTN-----
mfc_Zma  QHLGENMVRTIAMDGTGLVGRQVRLNTGSPITVPVGRATL IGRIMNVIGEPLDEKGDITTN-----
mfc_Nta  QHLGENMVRTIAMDGTGLVGRQVRLNTGSPITVPVGRATL IGRIMNVIGEPLDEKGDITTN-----
mfc_Hbr  QHMGEGMVRTIAMDGTGLVGRQVRLNTGSPITVPVGRANP WTYHEVIGEPLDERGDIKTS-----
mfc_Bta  QHLGESTVRTIAMDGTGLVGRQVRLNTGSPITVPVGRANP WTYHEVIGEPLDERGDIKTS-----
mfc_Rno  QHLGESTVRTIAMDGTGLVGRQVRLNTGSPITVPVGRANP WTYHEVIGEPLDERGDIKTS-----
mfc_Sce  QHLGENTVRTIAMDGTGLVGRQVRLNTGSPITVPVGRANP WTYHEVIGEPLDERGDIKTS-----
efc_Rbl  QHLGENTVRTIAMDGTGLVGRQVRLNTGSPITVPVGRANP WTYHEVIGEPLDERGDIKTS-----
mfc_Ncr  QHLGENVVRTIAMDGTGLVGRQVRLNTGSPITVPVGRANP WTYHEVIGEPLDERGDIKTS-----
efc_Eco  QHLGENTVRTIAMDGTGLVGRQVRLNTGSPITVPVGRANP WTYHEVIGEPLDERGDIKTS-----
efc_Pmo  QHLGENTVRTIAMDGTGLVGRQVRLNTGSPITVPVGRANP WTYHEVIGEPLDERGDIKTS-----
efc_Bsu  QHLGENTVRTIAMDGTGLVGRQVRLNTGSPITVPVGRANP WTYHEVIGEPLDERGDIKTS-----
efc_Ehi  QHLGENTVRTIAMDGTGLVGRQVRLNTGSPITVPVGRANP WTYHEVIGEPLDERGDIKTS-----
efc_Sli  QHLGENTVRTIAMDGTGLVGRQVRLNTGSPITVPVGRANP WTYHEVIGEPLDERGDIKTS-----
avr_Mba  TQLDAHVRGIALTPTEGLARGMAVEDTGGPLKAPVGREILSR MFDVFGNTIDRRKP-PSDI-----
vvc_Ncr  RINGDQATIQQV-YEETAGVMYVGDVPLRTGKPLSVELGPG LNNIYDGIQRPLEKIAEASNSIYIPRGIAT-----
vvc_Sce  RIDGDKATIQQV-YEETAGLTVGDVPLRTGKPLSVELGPG LMETIYDGIQRPLKAIKEESNSIYIPRGIDT-----
vvc_Hvu  RLEGDQATIQQV-YEETAGLTVNDVPLRTGKPLSVELGPG LGGIILGNIYDGIQRPLKTIKSRDVIYIPRGVSV-----
vvc_Zma  ---ARATIQQV-YEETAGLMYNDVPLRTGKPLSVELGPG LGGIILGNIYDGIQRPLKTIKSGDVIYIPRGVSV-----
vvc_Bta  RLEGDQATIQQV-YEETAGLTVGDVPLRTGKPLSVELGPG LGGIILGNIYDGIQRPLKTIKSGDVIYIPRGVSV-----
vvc_Dme  RLEGDQATIQQV-YEETAGLTVGDVPLRTGKPLSVELGPG LGGIILGNIYDGIQRPLKTIKSGDVIYIPRGVSV-----
vvc_Tco  RLEGDQATIQQV-YEETAGLTVGDVPLRTGKPLSVELGPG LGGIILGNIYDGIQRPLKTIKSGDVIYIPRGVSV-----
avc_Mja  QLHDDKAVIQQV-YEETSGIKPGEVVTGAPLSVELGPGMLR AMYDGIQRPLTAIEKTVSIFIPRGVDV-----
evc_Ehi  EMRQDVASIQV-YEETSGIGPGEVVRTGEALSVELGPGIIS QMFDGIQRPLDTFMEVTQSNFLGRGVQL-----
avc_Mba  QILGPKTIIQQV-YEETAGIKPGEVVRTGSSLSVELGPG LSSIYDGVQRPLDYLEMDEMG-AFLDRGVDA-----
avc_Mmo  QIVGDKTIIQQV-YEETSGIKPGEVVTGAPLSVELGPGMLR AMYDGIQRPLTAIEKTVSIFIPRGVDV-----
avc_Hsa  EIEGQVTTIQQV-YEETSGIGPGEVVRTGSLVVELGPG LSSIYDGVQRPLDYLEMDEMG-AFLDRGVDA-----
avc_Hvo  EIEGQVTTIQQV-YEETSGIGPGEVVRTGSLVVELGPG LSSIYDGVQRPLDYLEMDEMG-AFLDRGVDA-----
evc_Tth  RLDGDTAFVQV-YEDTSGLVKGEVVTGSLVVELGPGMLN GIYDGIQRPLERIREKTGI-YITRGVVV-----
avr_Sac  RIEGDRAFIQQV-YESTDGLKPGDKVYRSVAPLSVELGPG LTIYDGIQRPLDSTAKVSNSPFVARGVSI-----

```



```

+ + + + + + +
vvr_Ncr -----
vvr_Sce -----
vvr_Bta -----
vvr_Dme -----
vvr_Hvu -----
vvr_Tco -----
avr_Mba -----
avr_Mma -----
avr_Hsa -----
avr_Hvo -----
avr_Mja -----
evr_Ehi -----
evr_Tth -----
avr_Sac -----
cfr_Osa -----
cfr_Zma -----
cfr_Tae -----
cfr_Nta -----
cfr_Sol -----
cfr_Psa -----
cfr_Mpo -----
cfr_Cpa -----
cfr_Ppu -----
efr_Syn -----
efr_Bsu -----
efr_Ehi -----
efr_Pmo -----
mfr_Osa -----
mfr_Zma -----
mfr_Bta -----
mfr_Hsa -----
mfr_Dme -----
mfr_Spu -----
mfr_Ncr -----
efr_Rbl -----
mfr_Cre -----
efr_Sli -----
efr_Eco -----
afr_Mba -----
cfc_Osa -----
cfc_Zma -----
cfc_Tae -----
cfc_Nta -----
cfc_Sol -----
cfc_Psa -----
cfc_Mpo -----
cfc_Pth -----
cfc_Cpa -----
cfc_Ppu -----
efc_Syn -----
mfc_Osa -----
mfc_Zma -----
mfc_Nta -----
mfc_Hbr -----
mfc_Bta -----
mfc_Rno -----
mfc_Sce -----
efc_Rbl -----
mfc_Ncr -----
efc_Eco -----
efc_Pmo -----
efc_Bsu -----
efc_Ehi -----
efc_Sli -----
afc_Mba -----
vvc_Ncr PALDRKKKWEFTPT--MKVGDHIAGGDYWGTVYENSFISVHKILLPPRARGTITRIAEKGEYTVEEKILE
vvc_Sce PALDRTIKWQFTPGK-FQVGDHISGGDIYGSVFENSLISSHKILLPPRSRGTITWIAPAGEYTLDEKILE
vvc_Hvu PALDKDQLWEFQPNK-LGVGDNITNGDLYATVFENTLM-KHHIALPPGAMGKISYIAPAGQYSLQDTVLE
vvc_Zma PALDKDVLWEFQPTK-LGVGDVITGGDLYATVFENTLM-QHHVALPPGSMGKISYIAPAGQYNLQDTVLE
vvc_Bta SALS RDVKWDFTPCKNLRVGS HITGGDIYGIVNENSLI-KHKIMLPPRNRGTVTYIAPPGNYDTSDVLE
vvc_Dme TALS RSEMWEFNP-LNVRVGS HITGGDLYGVVHENTLV-QQRMIVAPRAKGTVRYIAPAGNYNLEDIVLE
vvc_Tco KSLNDQKQWDFKPC--LKVGDLYSGGDIIGSVVENS LMYNHSIMIPPVNRGRVTSIVPSGNYTLQDDIIE
avc_Mja PALPRDIKWEFKPV--VNEG DYVEEGDIIGTVDET PSI-VHKILVPIGVK-GKIVEIKEGKFTVEETVAV
evc_Ehi PALDHEKQWWFEAT--IEEGTEVSAGDIIGYVDETKII-QHKIMVPNGIK-GTVQKIESGSFTIDDPICV
avc_Mba DGLDHKKLWDFKPI--VKKGDSVKGGDVI GVVQETVNI-EHKIMVPPDI-SGTISDIKSGNFTVVDTICT
avc_Mma DGLDHKKLWDFKPV--VKKGDGVI GGDVIGVVQETVNI-EHKIMLPPDI-SGTVSDIRSGSFTVVDTICT
avc_Hsa PGIDLDTDWEFEPT--VEAGDEVAAGDVVGTVDETVSI-EHKVLVPPRSDGGEVVAVESGFTVDDTVVE
avc_Hvo PGIDLDEKWEFEPT--VSEGDEVAPGDVVGTVPETVTI-EHKVMVPPDFGGGEVVAVEEGEFSVTEAVVE
evc_Tth HALDREKKWAWTPM--VKPGDEV RGGMVLGTVPE-FGF-THKILVPPDVRGRVKEVKPAGEYTVVEEPVVV
avc_Soc PALDRQTKWHFVPK--VKS GDKYVGGDIIGVVQETDLI-EHRILIPPVNHGTLKELAREGDYTVEDVVAV

```

```

+ + + + +
vvr_Ncr -----EYLDINGSPINPYAREYPOEMISTGSAIDTMNSIARGQKIPISAAAGLPHNEIAA
vvr_Sce -----DYLDINGSPINPYARIYPEMISTGVSAIDTMNSIARGQKIPISAAAGLPHNEIAA
vvr_Bta -----DFLDIMGQPINPHFRIYPEEMIQTGISAIDGMNSIARGQKIPISAAAGLPHNEIAA
vvr_Dme -----DFLDIQGQPINPWSRIYPEEMIQTGISAIDVMNSIARGQKIPISAAAGLPHNEIAA
vvr_Hvu -----AYLDISGSSINPSETYPEEMIQTGISTIDVMNSIARGQKIPISAAAGLPHNEIAA
vvr_Tco -----QYRDIQGIPINPRARVYPEEMIQTGISSIDVMTSISRQKIPISAAAGLPHNEIAA
avr_Mba -----QLLDINGAAMNPYARLPKDFIQTGISTIDGTNTLVRGQKLPISAAAGLPHNEIAL
avr_Mma -----QLLDINGAAMNPYARLPKDFIQTGISTIDGTNTLVRGQKLPISAAAGLPHNEIAL
avr_Hsa -----DRVDIVGEAINPHAREYPEFIQTGVSGIDGMNTLVRGQKLPISAAAGLPHSDLAL
avr_Hvo -----ERRDIVGAINPYSREYPEFIQTGVSAIDGMNTLVRGQKLPISAAAGLPHNDLAL
avr_Mja -----NELDINGVINPYSRKVPSDFIQTGISTIDGMNTLVRGQKLPISAAAGLPHNQLAA
evr_Ehi -----KYLIDINGEVINPIARDYDFIQTGISAIDHNLNTLVRGQKLPISAAAGLPHKELAA
evr_Tth -----NRLPITGLPLNPVARRKPEFIQTGISTIDVMNTLVRGQKLPISAAAGLPHNEIAA
avr_Sac -----EKRNINGDPINPATREYPEFIQTGISAIDGLNSLLRGSNTDLSAAAGLPANTLAA
cfr_Osa -----ESRLIESPAGIISRRSVYEPLOTGLIAIDSMIPIGRQRELIIQDRQTGKTAVAT
cfr_Zma -----ESRLIESPAGIISRRSVYEPLOTGLIAIDSMIPIGRQRELIIQDRQTGKTAVAT
cfr_Tae -----ESRLIESPAGIISRRSVYEPLOTGLIAIDSMIPIGRQRELIIQDRQTGKTAVAT
cfr_Nta -----EFRLIESAPAGIISRRSVYEPLOTGLIAIDSMIPIGRQRELIIQDRQTGKTAVAT
cfr_Sol -----ESRLIESPAGIISRRSVYEPLOTGLIAIDSMIPIGRQRELIIQDRQTGKTAVAT
cfr_Psa -----ESRLIESPAGIISRRSVYEPLOTGLIAIDSMIPIGRQRELIIQDRQTGKTAVAT
cfr_Mpo -----EFRLIESPAGIISRRSVYEPLOTGLIAIDSMIPIGRQRELIIQDRQTGKTAVAI
cfr_Cpa -----DTRLIESAPAGIISRRSVYEPLOTGITAIIDAMIPIGRQRELIIQDRQTGKTAVAI
cfr_Ppu -----GTRLIESMAPGIISRRSVYEPLOTGITAIIDSMIPIGRQRELIIQDRQTGKTAVAI
efr_Syn -----ATRLIESPAGIISRRSVYEPLOTGITAIIDAMIPIGRQRELIIQDRQTGKTAVAI
efr_Bsu -----NTRPIESAPGVMDRKSVHEPLOTGINKAIDALIPIGRQRELIIQDRQTGKTAVAI
efr_Ehi -----NVRPVEAMAPGVMQRKSVHEPMTGGLKAVDLSLVPIGRQRELIIQDRQTGKTAVAI
efr_Pmo -----NMMPLDRKAYGIISRRKPVHEPLOTGINSIDGMVPIGRQRELIIQDRQTGKTAVAI
mfr_Osa -----ERRRVEVKAPGIISRRKSVHEPMTGGLKAVDLSLVPIGRQRELIIQDRQTGKTAVAI
mfr_Zma -----ERRRVEVKAPGIISRRKSVHEPMTGGLKAVDLSLVPIGRQRELIIQDRQTGKTAVAI
mfr_Bta -----ARRRVGLKAPGIIPRISVREPMQTGINKAVDLSLVPIGRQRELIIQDRQTGKTAVAI
mfr_Hsa -----TRRRVGLKAPGIIPRISVREPMQTGINKAVDLSLVPIGRQRELIIQDRQTGKTAVAI
mfr_Dme -----DRFRVGIKAPGIIPRISVREPMQTGINKAVDLSLVPIGRQRELIIQDRQTGKTAVAI
mfr_Spu -----HRARVGTAPGIPRISVREPMQTGINKAVDLSLVPICRQRELIIQDRQTGKTAVAI
mfr_Ncr -----EKRRRAQLKAPGILPRQSVNQPVTGLKAVSDAMVPIGRQRELIIQDRQTGKTAVAI
efr_Rbl -----ERRIADVAPGIIPRQSVNQPVTGLKAVSDAMVPIGRQRELIIQDRQTGKTAVAI
mfr_Cre -----RSSLVEIKAPGIIPRQSVNQPVTGLKAVSDAMVPIGRQRELIIQDRQTGKTAVAI
efr_Sli -----GRRALELQAPTVMQRKSVHEPMTGYNKAVDAMTPIGRQRELIIQDRQTGKTAVAI
efr_Eco -----GFSAVEATAPGVIERQSVNQPVTGYNKAVDAMTPIGRQRELIIQDRQTGKTAVAI
avr_Mba -----NRLPIERPSAAMDRAVTVPLQTVKVVDAIDALIPIGRQRELIIQDRQTGKTAVAI
cfc_Osa -----ATFPIHRSAPAFIELDTKLSIFETGKVVLDLAPYRRGGKIGLFGGAGVGKTVLIM
cfc_Zma -----ATFPIHRSAPAFIELDTKLSIFETGKVVLDLAPYRRGGKIGLFGGAGVGKTVLIM
cfc_Tae -----ATFPIHRSAPAFIELDTKLSIFETGKVVLDLAPYRRGGKIGLFGGAGVGKTVLIM
cfc_Nta -----TTSPIHRSAPAFIELDTKLSIFETGKVVLDLAPYRRGGKIGLFGGAGVGKTVLIM
cfc_Sol -----TTSPIHRSAPAFIELDTKLSIFETGKVVLDLAPYRRGGKIGLFGGAGVGKTVLIM
cfc_Psa -----TTSPIHRSAPAFIELDTKLSIFETGKVVLDLAPYRRGGKIGLFGGAGVGKTVLIM
cfc_Mpo -----TTFPIHRAAPAFIELDTKLSIFETGKVVLDLAPYRRGGKIGLFGGAGVGKTVLIM
cfc_Pth -----ITSPIHRPAPAFIELDTKLSIFETGKVVLDLAPYRRGGKIGLFGGAGVGKTVLIM
cfc_Cpa -----STLPIHRPAPAFIELDTKLSIFETGKVVLDLAPYRRGGKIGLFGGAGVGKTVLIM
cfc_Ppu -----STLPIHRPAPAFIELDTKLSIFETGKVVLDLAPYRRGGKIGLFGGAGVGKTVLIM
efc_Syn -----ETLPIHRPAPAFIELDTKLSIFETGKVVLDLAPYRRGGKIGLFGGAGVGKTVLIM
mfc_Osa -----HFLPIHREAPAFVEQATEQQILVGTGKVVLDLAPYRRGGKIGLFGGAGVGKTVLIM
mfc_Zma -----HFLPIHREAPAFVEQATEQQILVGTGKVVLDLAPYRRGGKIGLFGGAGVGKTVLIM
mfc_Nta -----HFLPIHREAPAFVEQATEQQILVGTGKVVLDLAPYRRGGKIGLFGGAGVGKTVLIM
mfc_Hbr -----HFLPIHREAPAFVDAQATEQQILVGTGKVVLDLAPYRRGGKIGLFGGAGVGKTVLIM
mfc_Bta -----QFAAIIHAEAPAFVEQATEQQILVGTGKVVLDLAPYRRGGKIGLFGGAGVGKTVLIM
mfc_Rno -----QFAPIIHAEAPAFVEQATEQQILVGTGKVVLDLAPYRRGGKIGLFGGAGVGKTVLIM
mfc_Sce -----LRKPIHADPPSFAEQSTSAEILETGKVVLDLAPYRRGGKIGLFGGAGVGKTVLIM
efc_Rbl -----ENRAIHHQKAPDFQSQSTESQILVGTGKVVLDLAPYRRGGKIGLFGGAGVGKTVLIM
mfc_Ncr -----NFRPIHAEAPAFVEQATEQQILVGTGKVVLDLAPYRRGGKIGLFGGAGVGKTVLIM
efc_Eco -----ERWAIHRAAPSFEELSNSQELLETTGKVVLDLAPYRRGGKIGLFGGAGVGKTVLIM
efc_Pmo -----EYAPIHREAPAFEDQGTKEKVFETGKVVLDLAPYRRGGKIGLFGGAGVGKTVLIM
efc_Bsu -----KNDPIHROAPSFDLSTVEILETTGKVVLDLAPYRRGGKIGLFGGAGVGKTVLIM
efc_Ehi -----ERSEIHKKAPSFDELSTVEILETTGKVVLDLAPYRRGGKIGLFGGAGVGKTVLIM
efc_Sli -----ERWPIHREKAPNFDELSTVEILETTGKVVLDLAPYRRGGKIGLFGGAGVGKTVLIM
vvc_Mba -----QWRSIHQTTPPLMRRSSTSEIFETGINAIDVLPYRRGGKIGLFGGAGVGKTVLIM
vvc_Ncr -----VEFDG-KKTEYPMMQTWVVRVPRPAAEKHSANQPFL-VGQQRVLDALFPCVQGGTTAIPGAFGCGKTVISQ
vvc_Sce -----VEFDG-KKSDFTLYHTWVVRVPRPVTEKLSADYPLL-TGQQRVLDALFPCVQGGTTAIPGAFGCGKTVISQ
vvc_Hvu -----LEFQG-IKKEFTMLHTWVVRTPRPVASKLAADTPLL-TGQQRVLDALFPCVQGGTTAIPGAFGCGKTVISQ
vvc_Zma -----LEFQG-IKKEFTMLQTWVVRSPRPVASKLAADTPLL-TGQQRVLDALFPCVQGGTTAIPGAFGCGKTVISQ
vvc_Bta -----LEFQG-IKKEFTMLQTWVVRQVRPVTEKLANHPLL-TGQQRVLDALFPCVQGGTTAIPGAFGCGKTVISQ
vvc_Dme -----TEFDG-EITKHTMLQVWVVRHARPVTEKLANHPLL-TGQQRVLDALFPCVQGGTTAIPGAFGCGKTVISQ
vvc_Tco -----LEYNG-TVKSLKLMHRWVVRTPRPVASKLSGNHPLL-TGQQRVLDALFPCVQGGTTAIPGAFGCGKTVISQ
avc_Mja -----VETENGERKEITMMQWVVRKPRPYKEKIPLEIPLI-TGQRVEDTFFLAKGGTAAIPGPFSGKTVISQ
evc_Ehi -----IETEQGL-KELTMMQWVVRGRPIKQKLNPDVPMI-TGQRVIDTFFVTGGAAPVPPGAGKTVVQH
avc_Mba -----LTDGT----ELQMMQRWVVRPRPVKAKLTPTRPLV-TGMRILDGLFPVAKGGTAAIPGPFSGKTVTQH
avc_Hsa -----LTDGT----ELQMMQRWVVRKPRPYKRLTPKPLV-TGQRILDGLFPVAKGGTAAIPGPFSGKTVTQH
avc_Hvo -----LDSGE----EITMHQWVVRQARPAAEKKTTPREPLV-SGQRILDGLFPVAKGGTAAIPGPFSGKTVTQH
evc_Tth -----LEDGT----ELKMYHTWVVRARPVQRKLDPNTPL-TGMRILDGLFPVAKGGTAAIPGPFSGKTVTQH
avc_Sac -----VDMNGDEI-PVKMYQWVVRIPRPYKEKLEPVEPLL-TGIRVLDLTVFPAKGGTAAIPGPFSGKTVTQH

```

```

+ + + + +
vvr_Ncr QICRQAGLVQRQGITNKGVDHGHENFSIVFVGAMGVNLETARFFTRDFEE-----NGSLGRITTLF
vvr_Sce QICRQAGLVR-----PTKDVHDHGHENFSIVFVGAMGVNLETARFFKQDFEE-----NGSLERTSLF
vvr_Bta QICRQAGLYK-KSKD---YVDYSSEENFAIVFAAMGVNLETARFFKSDFEE-----NGSMDNVCLF
vvr_Dme QICRQAGLVKLPKGS---VLDDHTDNFAIVFAAMGVNLETARFFKQDFEE-----NGSMENVCLF
vvr_Hvu QICRQAGLVKRLEQSKHAAEGGEEDNFIVFAAMGVNLETARFFKQDFEE-----NGSMERVTLF
vvr_Tco QIVRQAGLYRK-----EGKQEDFCIVFAAMGVNLETARFFRTDFEE-----NGSMEKTVLF
avr_Mba QIARQAAMP-----GSESAFVAVFAAMGITNEEAQYFMSDFEK-----TGALERAVVF
avr_Mma QIARQASVP-----GSESAFVAVFAAMGITNEEAQYFMSDFEK-----TGALERAVVF
avr_Hsa QIARQASVPE---EEAETDDEGSEFVAVFGAMGITAEAEAEFMDDFER-----TGALERSVVF
avr_Hvo QIARQATVPE---DE---ESGESEFVAVFGAMGITQEEAEAEFMDDFER-----TGALERSVVF
avr_Mja QIARQAKVR-----GEGEKFAVFAAMGITSEEAFFMEFRK-----TGALERAVVF
evr_Ehi QIARQATVL-----GSDDFAVVFAAIGITFEAEFFMEDFRQ-----TGAIIDRSVVF
evr_Tth QIARQATVRP---D-LSGEKEEFPFAVFAAMGITQRELSYFIQFER-----TGALRSVLF
avr_Sac QIAKQATVR-----GSESNFAVFAAIGVRYDEALFFRKFFEE-----TGAINRVAMF
cfr_Osa DTILNQK-----GQDVIQVYVAIGQRASSVAQVVTTFHE-----EGAMEYTIIV
cfr_Zma DTILNQK-----GQDVIQVYVAIGQRASSVAQVVTTFHE-----EGAMEYTIIV
cfr_Tae DTILNQK-----GQDVIQVYVAIGQRASSVAQVVTTFHE-----EGAMEYTIIV
cfr_Nta DTILNQK-----GQDVIQVYVAIGQRASSVAQVVTTFHE-----EGAMEYTIIV
cfr_Sol DTILNQK-----GQDVIQVYVAIGQRASSVAQVVTTFHE-----EGAMEYTIIV
cfr_Psa DTILNQK-----GQDVIQVYVAIGQRASSVAQVVTTFHE-----EGAMEYTIIV
cfr_Mpo DTILNQK-----GQDVIQVYVAIGQRASSVAQVVTTFHE-----EGAMEYTIIV
cfr_Cpa DTILNQK-----GQDVIQVYVAIGQRASSVAQVVTTFHE-----EGAMEYTIIV
cfr_Ppu DTILNQK-----GQDVIQVYVAIGQRASSVAQVVTTFHE-----EGAMEYTIIV
efr_Syn DTILNQK-----GQDVIQVYVAIGQRASSVAQVVTTFHE-----EGAMEYTIIV
efr_Bsu DTILNQK-----GQDVIQVYVAIGQRASSVAQVVTTFHE-----EGAMEYTIIV
efr_Ehi DTILNQK-----GQDVIQVYVAIGQRASSVAQVVTTFHE-----EGAMEYTIIV
efr_Pma DTILNQK-----GQDVIQVYVAIGQRASSVAQVVTTFHE-----EGAMEYTIIV
mfr_Osa DTILNQK-----GQDVIQVYVAIGQRASSVAQVVTTFHE-----EGAMEYTIIV
mfr_Zma DTILNQK-----GQDVIQVYVAIGQRASSVAQVVTTFHE-----EGAMEYTIIV
mfr_Bta DTILNQK-----GQDVIQVYVAIGQRASSVAQVVTTFHE-----EGAMEYTIIV
mfr_Hsa DTILNQK-----GQDVIQVYVAIGQRASSVAQVVTTFHE-----EGAMEYTIIV
mfr_Dme DTILNQK-----GQDVIQVYVAIGQRASSVAQVVTTFHE-----EGAMEYTIIV
mfr_Spu DTILNQK-----GQDVIQVYVAIGQRASSVAQVVTTFHE-----EGAMEYTIIV
mfr_Ncr DTILNQK-----GQDVIQVYVAIGQRASSVAQVVTTFHE-----EGAMEYTIIV
efr_Rbl DTILNQK-----GQDVIQVYVAIGQRASSVAQVVTTFHE-----EGAMEYTIIV
mfr_Cre DTILNQK-----GQDVIQVYVAIGQRASSVAQVVTTFHE-----EGAMEYTIIV
efr_Sli DTILNQK-----GQDVIQVYVAIGQRASSVAQVVTTFHE-----EGAMEYTIIV
efr_Eco DTILNQK-----GQDVIQVYVAIGQRASSVAQVVTTFHE-----EGAMEYTIIV
afn_Mba DTILNQK-----GQDVIQVYVAIGQRASSVAQVVTTFHE-----EGAMEYTIIV
cfc_Osa ELINN-----IAKAHGGVSVFAGVGERREGNDLYMEMKE---SGVIN-EKNIEESKVALV
cfc_Zma ELINN-----IAKAHGGVSVFAGVGERREGNDLYMEMKE---SGVIN-EKNIEESKVALV
cfc_Tae ELINN-----IAKAHGGVSVFAGVGERREGNDLYMEMKE---SGVIN-EKNIEESKVALV
cfc_Nta ELINN-----IAKAHGGVSVFAGVGERREGNDLYMEMKE---SGVIN-EKNIEESKVALV
cfc_Sol ELINN-----IAKAHGGVSVFAGVGERREGNDLYMEMKE---SGVIN-EKNIEESKVALV
cfc_Psa ELINN-----IAKAHGGVSVFAGVGERREGNDLYMEMKE---SGVIN-EKNIEESKVALV
cfc_Mpo ELINN-----IAKAHGGVSVFAGVGERREGNDLYMEMKE---SGVIN-EKNIEESKVALV
cfc_Pth ELINN-----IAKAHGGVSVFAGVGERREGNDLYMEMKE---SGVIN-EKNIEESKVALV
cfc_Cpa ELINN-----IAKAHGGVSVFAGVGERREGNDLYMEMKE---SGVIN-EKNIEESKVALV
cfc_Ppu ELINN-----IAKAHGGVSVFAGVGERREGNDLYMEMKE---SGVIN-EKNIEESKVALV
efc_Syn ELINN-----IAKAHGGVSVFAGVGERREGNDLYMEMKE---SGVIN-EKNIEESKVALV
mfc_Osa ELINN-----IAKAHGGVSVFAGVGERREGNDLYMEMKE---SGVIN-EKNIEESKVALV
mfc_Zma ELINN-----IAKAHGGVSVFAGVGERREGNDLYMEMKE---SGVIN-EKNIEESKVALV
mfc_Nta ELINN-----IAKAHGGVSVFAGVGERREGNDLYMEMKE---SGVIN-EKNIEESKVALV
mfc_Hbr ELINN-----IAKAHGGVSVFAGVGERREGNDLYMEMKE---SGVIN-EKNIEESKVALV
mfc_Bta ELINN-----IAKAHGGVSVFAGVGERREGNDLYMEMKE---SGVIN-EKNIEESKVALV
mfc_Rno ELINN-----IAKAHGGVSVFAGVGERREGNDLYMEMKE---SGVIN-EKNIEESKVALV
mfc_Sce ELINN-----IAKAHGGVSVFAGVGERREGNDLYMEMKE---SGVIN-EKNIEESKVALV
efc_Rbl ELINN-----IAKAHGGVSVFAGVGERREGNDLYMEMKE---SGVIN-EKNIEESKVALV
mfc_Ncr ELINN-----IAKAHGGVSVFAGVGERREGNDLYMEMKE---SGVIN-EKNIEESKVALV
efc_Eco ELINN-----IAKAHGGVSVFAGVGERREGNDLYMEMKE---SGVIN-EKNIEESKVALV
efc_Pmo ELINN-----IAKAHGGVSVFAGVGERREGNDLYMEMKE---SGVIN-EKNIEESKVALV
efc_Bsu ELINN-----IAKAHGGVSVFAGVGERREGNDLYMEMKE---SGVIN-EKNIEESKVALV
efc_Ehi ELINN-----IAKAHGGVSVFAGVGERREGNDLYMEMKE---SGVIN-EKNIEESKVALV
efc_Sli ELINN-----IAKAHGGVSVFAGVGERREGNDLYMEMKE---SGVIN-EKNIEESKVALV
afn_Mba ELINN-----IAKAHGGVSVFAGVGERREGNDLYMEMKE---SGVIN-EKNIEESKVALV
vvc_Ncr SVSKF-----VANNHGGVSVFAGVGERREGNDLYMEMKE---SGVIN-EKNIEESKVALV
vvc_Sce SLASKY-----SNDVIVYVGGCGERGNEAEVLMDFPPELSEI---DGRKEPIMKRTTLI
vvc_Hvu ALSKY-----SNDVIVYVGGCGERGNEAEVLMDFPPELSEI---DGRKEPIMKRTTLI
vvc_Zma ALSKY-----SNDVIVYVGGCGERGNEAEVLMDFPPELSEI---DGRKEPIMKRTTLI
vvc_Bta SLASKY-----SNDVIVYVGGCGERGNEAEVLMDFPPELSEI---DGRKEPIMKRTTLI
vvc_Dme ALSKY-----SNDVIVYVGGCGERGNEAEVLMDFPPELSEI---DGRKEPIMKRTTLI
vvc_Tco ALSKF-----SNDVIVYVGGCGERGNEAEVLMDFPPELSEI---DGRKEPIMKRTTLI
avc_Mja QLAKW-----SDADVIVYVGGCGERGNEAEVLMDFPPELSEI---DGRKEPIMKRTTLI
evc_Ehi QLAKW-----SDADVIVYVGGCGERGNEAEVLMDFPPELSEI---DGRKEPIMKRTTLI
avc_Mba SLAKW-----SDTEIVYVGGCGERGNEAEVLMDFPPELSEI---DGRKEPIMKRTTLI
avc_Mma QLAKW-----SDTEIVYVGGCGERGNEAEVLMDFPPELSEI---DGRKEPIMKRTTLI
avc_Hsa SLAKF-----ADADIVYVGGCGERGNEAEVLMDFPPELSEI---DGRKEPIMKRTTLI
avc_Hvo QLAKW-----ADADIVYVGGCGERGNEAEVLMDFPPELSEI---DGRKEPIMKRTTLI
evc_Tth SLAKW-----SNADVIVYVGGCGERGNEAEVLMDFPPELSEI---DGRKEPIMKRTTLI
avc_Sac SLAKW-----SAAKVVIVYVGGCGERGNEAEVLMDFPPELSEI---DGRKEPIMKRTTLI

```

```

+ + + + +
vvr_Ncr LNLANDPTIERIITPRALTTAEYYAYQLEKHLVILTDLSSYCDALREVSAAAREEVPGRRGFPGYMYTD
vvr_Sce LNLANDPTIERIITPRALTTAEYFLAYQTERHVLILTDMSSYADALREVSAAAREEVPGRRGFPGYMYTD
vvr_Bta LNLANDPTIERIITPRALTTAEYFLAYQCEKHLVILTDMSYAEALREVSAAAREEVPGRRGFPGYMYTD
vvr_Dme LNLANDPTIERIITPRALTTAEYFLAYQCEKHLVILTDMSYAEALREVSAAAREEVPGRRGFPGYMYTD
vvr_Hvu LNLANDPTIERIITPRALTTAEYFLAYQCEKHLVILTDMSYADALREVSAAAREEVPGRRGFPGYMYTD
vvr_Tco MNLADDPPIERIVTPRMLTTAEYFLAYQCEKHLVILTDMSYADALREVSAAAREEVPGRRGFPGYMYTD
avr_Mba LNLADDPPIERIVTPRMLTTAEYFLAYQCEKHLVILTDMSYADALREVSAAAREEVPGRRGFPGYMYTD
avr_Mma LNLADDPPIERIVTPRMLTTAEYFLAYQCEKHLVILTDMSYADALREVSAAAREEVPGRRGFPGYMYTD
avr_Hsa LNLADDPPIERIVTPRMLTTAEYFLAYQCEKHLVILTDMSYADALREVSAAAREEVPGRRGFPGYMYTD
avr_Hvo TNLADDPPIERIVTPRMLTTAEYFLAYQCEKHLVILTDMSYADALREVSAAAREEVPGRRGFPGYMYTD
avr_Mja INLADDPPIERIVTPRMLTTAEYFLAYQCEKHLVILTDMSYADALREVSAAAREEVPGRRGFPGYMYTD
evr_Ehi MNLADDPPIERIVTPRMLTTAEYFLAYQCEKHLVILTDMSYADALREVSAAAREEVPGRRGFPGYMYTD
evr_Tth LNKADDPPIERIVTPRMLTTAEYFLAYQCEKHLVILTDMSYADALREVSAAAREEVPGRRGFPGYMYTD
avr_Sac VTLADDPPIERIVTPRMLTTAEYFLAYQCEKHLVILTDMSYADALREVSAAAREEVPGRRGFPGYMYTD
cfr_Osa AEMADSPATLQYLAPYTGAAALAEYFMYR-ERHTLIYDLSKQAQAYRQMSLLRRPPGREAVPGDVFYL
cfr_Zma AEMADSPATLQYLAPYTGAAALAEYFMYR-ERHTLIYDLSKQAQAYRQMSLLRRPPGREAVPGDVFYL
cfr_Tae AEMADSPATLQYLAPYTGAAALAEYFMYR-ERHTLIYDLSKQAQAYRQMSLLRRPPGREAVPGDVFYL
cfr_Nta AETADSPATLQYLAPYTGAAALAEYFMYR-ERHTLIYDPSKQAQAYRQMSLLRRPPGREAVPGDVFYL
cfr_Sol AETADSPATLQYLAPYTGAAALAEYFMYR-ERHTLIYDPSKQAQAYRQMSLLRRPPGREAVPGDVFYL
cfr_Psa AETADSPATLQYLAPYTGAAALAEYFMYR-ERHTLIYDPSKQAQAYRQMSLLRRPPGREAVPGDVFYL
cfr_Mpo AETADSPATLQYLAPYTGAAALAEYFMYR-ERHTLIYDPSKQAQAYRQMSLLRRPPGREAVPGDVFYL
cfr_Cpa AANADDPATLQYLAPYTGASIAEYFMYK-GQHTLIYDLSKQAQAYRQMSLLRRPPGREAVPGDVFYL
cfr_Ppu TANADSPATLQYLAPYTGAAALAEYFMYK-GKATLIYDLSKQAQAYRQMSLLRRPPGREAVPGDVFYL
efr_Syn AANADSPALQYLAPYTGASIAEYFMYK-GKHTLIYDLSKQAQAYRQMSLLRRPPGREAVPGDVFYL
efr_Bsu TASASQPAPLLYLAPYAGVTMAEYFMYN-GKHVLIYDLSKQAQAYRQMSLLRRPPGREAVPGDVFYL
efr_Ehi NAGASQPAPLLYLAPYAGVTMAEYFMYN-GKHVLIYDLSKQAQAYRQMSLLRRPPGREAVPGDVFYL
efr_Pmo AATASQPAPLQYMAPYTGVSMEYFMYK-GERVLIYDLSKQAVAYRQMSLLRRPPGREAVPGDVFYL
mfr_Osa AATASQPAPLQYMAPYTGVSMEYFMYK-GERVLIYDLSKQAVAYRQMSLLRRPPGREAVPGDVFYL
mfr_Zma AATASQPAPLQYMAPYTGVSMEYFMYK-GERVLIYDLSKQAVAYRQMSLLRRPPGREAVPGDVFYL
mfr_Bta SATASDAAPLQYLAPYSGCAMGEYFRDN-GKHALIYDLSKQAVAYRQMSLLRRPPGREAVPGDVFYL
mfr_Hsa SATASDAAPLQYLAPYSGCAMGEYFRDN-GKHALIYDLSKQAVAYRQMSLLRRPPGREAVPGDVFYL
mfr_Dme SATASDAAPLQYLAPYSGCAMGEYFRDN-GKHALIYDLSKQAVAYRQMSLLRRPPGREAVPGDVFYL
mfr_Spu SATASDAAPLQYLAPYSGCAMGEYFRDN-GKHALIYDLSKQAVAYRQMSLLRRPPGREAVPGDVFYL
mfr_Ncr AATASQPAPLQYLAPYTGACVGEYFRDN-GKHSVLIYDLSKQAVAYRQMSLLRRPPGREAVPGDVFYL
efr_Rbl AATASQPAPLQYLAPYTGACVGEYFRDN-GKHSVLIYDLSKQAVAYRQMSLLRRPPGREAVPGDVFYL
mfr_Cre SATASDAAPLQYLAPYTGACVGEYFRDN-GKHSVLIYDLSKQAVAYRQMSLLRRPPGREAVPGDVFYL
efr_Sli AAPASDPAPLQYLAPYTGSAIGQWYME-GKHVLIYDLSKQADAYRQMSLLRRPPGREAVPGDVFYL
efr_Eco VATASQPAPLQYLAPYTGSAIGQWYME-GKHVLIYDLSKQADAYRQMSLLRRPPGREAVPGDVFYL
afv_Mba YTEGNDPAPLQYLAPYTGSAIGQWYME-GKHVLIYDLSKQADAYRQMSLLRRPPGREAVPGDVFYL
cfc_Osa YGQMNEPPGARMRVGLTALTMAEYFRDVKQDVLFFIDNIFRFVQAGSEVSALLGRMPSAVGYQPTLSTE
cfc_Zma YGQMNEPPGARMRVGLTALTMAEYFRDVKQDVLFFIDNIFRFVQAGSEVSALLGRMPSAVGYQPTLSTE
cfc_Tae YGQMNEPPGARMRVGLTALTMAEYFRDVKQDVLFFIDNIFRFVQAGSEVSALLGRMPSAVGYQPTLSTE
cfc_Nta YGQMNEPPGARMRVGLTALTMAEYFRDVKQDVLFFIDNIFRFVQAGSEVSALLGRMPSAVGYQPTLSTE
cfc_Sol YGQMNEPPGARMRVGLTALTMAEYFRDVKQDVLFFIDNIFRFVQAGSEVSALLGRMPSAVGYQPTLSTE
cfc_Psa YGQMNEPPGARMRVGLTALTMAEYFRDVKQDVLFFIDNIFRFVQAGSEVSALLGRMPSAVGYQPTLSTE
cfc_Mpo YGQMNEPPGARMRVGLTALTMAEYFRDVKQDVLFFIDNIFRFVQAGSEVSALLGRMPSAVGYQPTLSTE
cfc_Pth YGQMNEPPGARMRVGLTALTMAEYFRDVKQDVLFFIDNIFRFVQAGSEVSALLGRMPSAVGYQPTLSTE
cfc_Cpa YGQMNEPPGARMRVGLTALTMAEYFRDVKQDVLFFIDNIFRFVQAGSEVSALLGRMPSAVGYQPTLSTE
cfc_Ppu YGQMNEPPGARMRVGLTALTMAEYFRDVKQDVLFFIDNIFRFVQAGSEVSALLGRMPSAVGYQPTLSTE
efc_Syn YGQMNEPPGARMRVGLTALTMAEYFRDVKQDVLFFIDNIFRFVQAGSEVSALLGRMPSAVGYQPTLSTE
mfc_Osa YGQMNEPPGARARVGLTGLTVAEHFRDAEGQDVLFFIDNIFRFTQANSEVSALLGRIPSAVGYQPTLATD
mfc_Zma YGQMNEPPGARARVGLTGLTVAEHFRDAEGQDVLFFIDNIFRFTQANSEVSALLGRIPSAVGYQPTLATD
mfc_Nta YGQMNEPPGARARVGLTGLTVAEHFRDAEGQDVLFFIDNIFRFTQANSEVSALLGRIPSAVGYQPTLATD
mfc_Hbr YGQMNEPPGARARVGLTGLTVAEHFRDAEGQDVLFFIDNIFRFTQANSEVSALLGRIPSAVGYQPTLATD
mfc_Bta YGQMNEPPGARARVGLTGLTVAEHFRDAEGQDVLFFIDNIFRFTQANSEVSALLGRIPSAVGYQPTLATD
mfc_Rno YGQMNEPPGARARVGLTGLTVAEHFRDAEGQDVLFFIDNIFRFTQANSEVSALLGRIPSAVGYQPTLATD
mfc_Sce YGQMNEPPGARARVGLTGLTVAEHFRDAEGQDVLFFIDNIFRFTQANSEVSALLGRIPSAVGYQPTLATD
efc_Rbl YGQMNEPPGARARVGLTGLTVAEHFRDAEGQDVLFFIDNIFRFTQANSEVSALLGRIPSAVGYQPTLATD
mfc_Ncr YGQMNEPPGARARVGLTGLTVAEHFRDAEGQDVLFFIDNIFRFTQANSEVSALLGRIPSAVGYQPTLATD
efc_Eco YGQMNEPPGNRLRVALTGLTMAEKFRD-EGQDVLFFIDNIFRFTQANSEVSALLGRIPSAVGYQPTLAE
efc_Pmo YGQMNEPPGARLRVGLTGLTMAEKFRD-EGQDVLFFIDNIFRFTQANSEVSALLGRIPSAVGYQPTLAE
efc_Bsu YGQMNEPPGARLRVGLTGLTMAEKFRD-EGQDVLFFIDNIFRFTQANSEVSALLGRIPSAVGYQPTLAE
efc_Ehi YGQMNEPPGARLRVGLTGLTMAEKFRD-EGQDVLFFIDNIFRFTQANSEVSALLGRIPSAVGYQPTLAE
efc_Sli YGQMNEPPGARLRVGLTGLTMAEKFRD-EGQDVLFFIDNIFRFTQANSEVSALLGRIPSAVGYQPTLAE
afv_Mba YGQMNEPPGARLRVGLTGLTMAEKFRD-EGQDVLFFIDNIFRFTQANSEVSALLGRIPSAVGYQPTLAE
vvc_Ncr ANTSMPPVAAREASIIYTGITIAEYFRDM-GKNSVMMADSSSRWAEALREISGRLEMPADQGFPAYLGA
vvc_Sce ANTSMPPVAAREASIIYTGITIAEYFRDM-GKNSVMMADSSSRWAEALREISGRLEMPADQGFPAYLGA
vvc_Hvu ANTSMPPVAAREASIIYTGITIAEYFRDM-GKNSVMMADSSSRWAEALREISGRLEMPADQGFPAYLGA
vvc_Zma ANTSMPPVAAREASIIYTGITIAEYFRDM-GKNSVMMADSSSRWAEALREISGRLEMPADQGFPAYLGA
vvc_Bta ANTSMPPVAAREASIIYTGITIAEYFRDM-GKNSVMMADSSSRWAEALREISGRLEMPADQGFPAYLGA
vvc_Dme ANTSMPPVAAREASIIYTGITIAEYFRDM-GKNSVMMADSSSRWAEALREISGRLEMPADQGFPAYLGA
vvc_Tco ANTSMPPVAAREASIIYTGITIAEYFRDM-GKNSVMMADSSSRWAEALREISGRLEMPADQGFPAYLGA
avc_Mja ANTSMPPVAAREASIIYTGITIAEYFRDM-GKNSVMMADSSSRWAEALREISGRLEMPADQGFPAYLGA
evc_Ehi ANTSMPPVAAREASIIYTGITIAEYFRDM-GKNSVMMADSSSRWAEALREISGRLEMPADQGFPAYLGA
avc_Mba ANTSMPPVAAREASIIYTGITIAEYFRDM-GKNSVMMADSSSRWAEALREISGRLEMPADQGFPAYLGA
avc_Mma ANTSMPPVAAREASIIYTGITIAEYFRDM-GKNSVMMADSSSRWAEALREISGRLEMPADQGFPAYLGA
avc_Hsa ANTSMPPVAAREASIIYTGITIAEYFRDM-GKNSVMMADSSSRWAEALREISGRLEMPADQGFPAYLGA
avc_Hvo ANTSMPPVAAREASIIYTGITIAEYFRDM-GKNSVMMADSSSRWAEALREISGRLEMPADQGFPAYLGA
evc_Tth ANTSMPPVAAREASIIYTGITIAEYFRDM-GKNSVMMADSSSRWAEALREISGRLEMPADQGFPAYLGA
avc_Sac ANTSMPPVAAREASIIYTGITIAEYFRDM-GKNSVMMADSSSRWAEALREISGRLEMPADQGFPAYLGA

```

```

+ + + + +
vvr_Ncr LSTIYERAGRVAGR-----NGSITQIPILTMPNDDITHPIFDLGYITEGQIFVDRQLHNRQ
vvr_Sce LSTIYERAGRVEGR-----NGSITQIPILTMPNDDITHPIFDLGYITEGQIFVDRQLHNRQ
vvr_Bta LATIYERAGRVEGR-----NGSITQIPILTMPNDDITHPIFDLGYITEGQIYVDRQLHNRQ
vvr_Dme LATIYERAGRVEGR-----NGSITQIPILTMPNDDITHPIFDLGYITEGQIYVDRQLHNRQ
vvr_Hvu LATIYERAGRIEGR-----KGSITQIPILTMPNDDITHPIFDLGYITEGQIYVDRQLHNRQ
vvr_Tco LACIYERAGRVLGR-----AGSITQIPILTMPNDDITHPIFDLGYITEGQIYVDRQLHNRQ
avr_Mba LATLYERAGIVKGA-----KGSVTQIPILSMPGDDITHPIFDLGYITEGQIYVDRQLHNRQ
avr_Mma LATLYERAGIVKGA-----KGSVTQIPILSMPGDDITHPIFDLGYITEGQIYVDRQLHNRQ
avr_Hsa LAQLYERAGRIEGK-----EGSVTQIPILTMPGDDITHPIFDLGYITEGQIMMNRDLNSQG
avr_Hvo LAQLYERAGRIQGR-----DGSVTQIPILTMPGDDITHPIFDLGYITEGQIYDRDLNSQG
avr_Mja LATIYERAGRVKGR-----TGTITQIPILTMPDODITHPIFDLGYITEGQIVLSRELHHRKG
evr_Ehi LATLFERAGRIRGL-----KGSVTQIPILTMPEDDKTHPIFDLGYITEGQIILTRRELHHRKG
evr_Tth LATIYERAGVVEGK-----KGSVTQIPILSMPDODRTHPIFDLGYITEGQIILTRRELHHRKG
avr_Sac LATIYERAGKVIGK-----KGSITQMPILTMPNDDMTHPIFDLGYITEGQIVLDRSLFNKKG
cfr_Osa HSRLLERAAKLNSLLGE-----GSMALPIVETQSGDVSAYIPTNVISITDGGIFLSADLFNAG
cfr_Zma HSRLLERAAKLNSLLGE-----GSMALPIVETQSGDVSAYIPTNVISITDGGIFLSADLFNAG
cfr_Tae HSRLLERAAKLNSLLGE-----GSMALPIVETQSGDVSAYIPTNVISITDGGIFLSADLFNAG
cfr_Nta HSRLLERAAKLSSLLGE-----GSMALPIVETQSGDVSAYIPTNVISITDGGIFLSADLFNAG
cfr_Sol HSRLLERAAKLSSLLGE-----GSMALPIVETQAGDVSAYIPTNVISITDGGIFLSADLFNAG
cfr_Psa HSRLLERAAKLSSLLGE-----GSMALPIVETQAGDVSAYIPTNVISITDGGIFLSADLFNAG
cfr_Mpo HSRLLERAAKLSSLLGE-----GSMALPIVETQAGDVSAYIPTNVISITDGGIFLSADLFNAG
cfr_Cpa HSRLLERAAKLSPQLGE-----GSMALPIVETQAGDVSAYIPTNVISITDGGIFLSADLFNAG
cfr_Ppu HSRLLERAAKLNSDLGG-----GSMALPIVETQAGDVSAYIPTNVISITDGGIFLSADLFNAG
efr_Syn HSRLLERAAKLNDALGG-----GSMALPIVETQAGDVSAYIPTNVISITDGGIFLSADLFNAG
efr_Bsu HSRLLERAAKLSDAKGA-----GSITALPFVETQAGDISAYIPTNVISITDGGIFLSADLFNAG
efr_Ehi HSRLLERAAKLSDDLGG-----GSMALPFVETQAGDISAYIPTNVISITDGGIFLSADLFNAG
efr_Pmo HSRLLERAAKLSDDEIGA-----GSITALPIVETQAGDVSAYIPTNVISITDGGIFLSADLFNAG
mfr_Osa HSRLLERAAKRSDQTGA-----GSLTALPVIETQAGDVSAYIPTNVISITDGGIFLSADLFNAG
mfr_Zma HSRLLERAAKRSDQTGA-----GSLTALPVIETQAGDVSAYIPTNVISITDGGIFLSADLFNAG
mfr_Bta HSRLLERAAKMNDAFGG-----GSLTALPVIETQAGDVSAYIPTNVISITDGGIFLSADLFNAG
mfr_Hsa HSRLLERAAKMNDAFGG-----GSLTALPVIETQAGDVSAYIPTNVISITDGGIFLSADLFNAG
mfr_Dme HSRLLERAAKMSPAMGG-----GSLTALPVIETQAGDVSAYIPTNVISITDGGIFLSADLFNAG
mfr_Spu HSRLLERAAKMNPDFGG-----GSLTALPVIETQAGDVSAYIPTNVISITDGGIFLSADLFNAG
mfr_Ncr HSRLLERAAKMNKTHGG-----GSMALPVIETQAGDVSAYIPTNVISITDGGIFLSADLFNAG
efr_Rbl HSRLLESAKMNGDFGA-----GSLTALPVIETQAGDISAYIPTNVISITDGGIFLSADLFNAG
mfr_Cre HSRLLERAAKLSAAMGG-----GSLTALPVIETQAGDISAYIPTNVISITDGGIFLSADLFNAG
efr_Sli HSRLLEPCAKLSDAEGA-----GSMGLPIVETKANDVSAYIPTNVISITDGGIFLSADLFNAG
efr_Eco HSRLLERAARVNAEYVEAFTKGEVKGKT-----GSLTALPVIETQAGDISAYIPTNVISITDGGIFLSADLFNAG
avr_Mba HSRLLESTHLLKEFGG-----GSLTALPVIETEAQNI SAYIPTNLISITDGGIYLSPSLFEFG
cfc_Osa MGS LQERIT-----S-----TKKGSITSIQAVYVPADDLTDPA PATTFAHLDATTVLSRGLASKG
cfc_Zma MGS LQERIT-----S-----TKKGSITSIQAVYVPADDLTDPA PATTFAHLDATTVLSRGLASKG
cfc_Tae MGS LQERIA-----S-----TKKGSITSIQAVYVPADDLTDPA PATTFAHLDATTVLSRGLASKG
cfc_Nta MGS LQERIT-----S-----TKKGSITSIQAVYVPADDLTDPA PATTFAHLDATTVLSRGLASKG
cfc_Sol MGS LQERIT-----S-----TKKGSITSIQAVYVPADDLTDPA PATTFAHLDATTVLSRGLASKG
cfc_Psa MGT LQERIT-----S-----TKKGSITSIQAVYVPADDLTDPA PATTFAHLDATTVLSRGLASKG
cfc_Mpo MGT LQERIT-----S-----TKKGSITSIQAVYVPADDLTDPA PATTFAHLDATTVLSRGLASKG
cfc_Pth MGS LQERIT-----S-----TKKGSITSIQAVYVPADDLTDPA PATTFAHLDATTVLSRGLASKG
cfc_Cpa MGS LQERIT-----S-----TKKGSITSIQAVYVPADDLTDPA PATTFAHLDATTVLSRGLASKG
cfc_Ppu MGS LQERIT-----S-----TKKGSITSIQAVYVPADDLTDPA PATTFAHLDATTVLSRGLASKG
efc_Syn VGS LQERIT-----S-----TKKGSITSIQAVYVPADDLTDPA PATTFAHLDATTVLSRGLASKG
mfc_Osa LGGLQERIT-----T-----TKKGSITSVQAIYVPADDLTDPA PATTFAHLDATTVLSRQISELG
mfc_Zma LGGLQERIT-----T-----TKKGSITSVQAIYVPADDLTDPA PATTFAHLDATTVLSRQISELG
mfc_Nta LGGLQERIT-----T-----TKKGSITSVQAIYVPADDLTDPA PATTFAHLDATTVLSRQISELG
mfc_Hbr LGGLQERIT-----T-----TKKGSITSVQAIYVPADDLTDPA PATTFAHLDATTVLSRQISELG
mfc_Bta MGMTQERIT-----T-----TKKGSITSVQAIYVPADDLTDPA PATTFAHLDATTVLSRAIAELG
mfc_Rno MGMTQERIT-----T-----TKKGSITSVQAIYVPADDLTDPA PATTFAHLDATTVLSRAIAELG
mfc_Sce MGLLQERIT-----T-----TKKGSITSVQAVYVPADDLTDPA PATTFAHLDATTVLSRAIAELG
efc_Rbl MGQLQERIT-----S-----TKKGSITSVQAVYVPADDLTDPA PATTFAHLDATTVLSRAIAELG
mfc_Ncr MGQMQERIT-----T-----TKKGSITSVQAVYVPADDLTDPA PATTFAHLDATTVLSRAIAELG
efc_Eco MGVLQERIT-----S-----TKKGSITSVQAVYVPADDLTDPA PATTFAHLDATTVLSRQIASLG
efc_Pmo MGALQERIT-----S-----TKIGSITSVQAVYVPADDLTDPA PATTFAHLDATTVLSRRIASLG
efc_Bsu MGQLQERIT-----S-----TNVGSVTSIQAVYVPADDLTDPA PATTFAHLDATTNLERKLTQEG
efc_Ehi MGQLQERIT-----S-----TKKGSITSIQAVYVPADDYTDPA PATTFAHLDATTNLERKLTQEG
efc_Sli MGLLQERIT-----S-----TRGHSITSMQAIYVPADDLTDPA PATTFAHLDATTNLERKLTQEG
afc_Mba LSELEERIS-----T-----TDAGAIMSIVQAVYVPADDLTDPA PATTFAHLDATTNLERKLTQEG
vvc_Ncr LASFYERAGKQVALGSP-----PREGGSIVGAVSPPGGDFSDPVTSA TLGIVQVFWGLDCKKLAQRK
vvc_Sce LASFYERAGKQVALGSP-----DRTGVSIVAVSPPAGGDFSDPVTSA TLGIVQVFWGLDCKKLAQRK
vvc_Hvu LASFYERAGKQVALGSP-----DRTGVSIVAVSPPAGGDFSDPVTSA TLGIVQVFWGLDCKKLAQRK
vvc_Zma LASFYERAGKQVALGSP-----DRTGVSIVAVSPPAGGDFSDPVTSA TLGIVQVFWGLDCKKLAQRK
vvc_Bta LASFYERAGKQVALGSP-----DRTGVSIVAVSPPAGGDFSDPVTSA TLGIVQVFWGLDCKKLAQRK
vvc_Dme LATFYERAGRVKCLGNP-----EREGSVIVGAVSPPGGDFSDPVTSA TLGIVQVFWGLDCKKLAQRK
vvc_Tco LASFYERAGRVKCLGNP-----EREGSVIVGAVSPPGGDFSDPVTSA TLGIVQVFWGLDCKKLAQRK
avc_Mja LAQFYERAGRVITLTKD-----NRQGVIVGAVSPPGGDFSDPVTSA TLGIVQVFWGLDCKKLAQRK
evc_Ehi LAEYYSRGRVIALGSD-----QREGSITAI SAVSPPGGDFSEPVTSNTLRIVKVFWALDANLARRR
avc_Mba LAEYFERAGVAESLCE-----TGTITVIGAVSPPGGDFSEPVTSNTLRIVKVFWALDANLARRR
avc_Mma LAEYFERAGVAETLCE-----KGSITAI SAVSPPGGDFSEPVTSNTLRIVKVFWALDANLARRR
avc_Hsa LSEFYERAGYFENFNGT-----EGSISVIGAVSPPGGDFSEPVTSNTLRIVKVFWALDANLARRR
avc_Hvo LSEFYERAGYFTVNGE-----EGSISVIGAVSPPGGDFSEPVTSNTLRIVKVFWALDANLARRR
evc_Tth LAAFYERAGKVITLGE-----EGAVTIVGAVSPPGGDFSEPVTSNTLRIVKVFWALDANLARRR
avc_Sac LAEYYPAGRVIALGNP-----ERYGSVITIASAVSPPGGDFSEPVTSNTLRIVKVFWALDANLARRR

```

```

+ + + + +
vvr_Ncr IYPPINVLPSLGLMKSAIG--EGMTRKDHGDVSNQLYAKYAIGRDAAMKAVVGEALSNEDKLSLEFL
vvr_Sce IYPPINVLPSLGLMKSAIG--EGMTRKDHGDVSNQLYAKYAIGRDAAMKAVVGEALTSIEDKLSLEFL
vvr_Bta IYPPINVLPSLGLMKSAIG--EGMTRKDHGDVSNQLYACYAIGRDVQAMKAVVGEALTSDDLLYLEFL
vvr_Dme IYPPINVLPSLGLMKSAIG--EGMTRKDHSDVSNQLYACYAIGRDVQAMKAVVGEALTPDDLYLEFL
vvr_Hvu IYPPINVLPSLGLMKSAIG--EGMTRRDHSDVSNQLYANYAIGRDVQAMKAVVGEALSSDDLYLEFL
vvr_Tco IYPPINVLPSLGLMKSAIG--EGMTRRDHSDVSNQMYARMRLD-DILANKAVVGEETLSSDDLHLEFL
avr_Mba IYPPINVLPSLGLMNSGIG--AGKTREDDHKAVSDQMYAGYAEGRDRLGLVAVVGEALSERDVKFLEFA
avr_Mma IYPPINVLPSLGLMNSGIG--AGKTREDDHKAVSDQMYAGYAEGRDRLGLVAVVGEALSERDTKFLEFA
avr_Hsa IYPPINVLPSLGLMNSGIG--AGKTREDDHGDVSDQLYAAYAEGEELRDLVNIIVGREALSERDNRYLDFFA
avr_Hvo IYPPINVLPSLGLMNSGIG--AGKTREDDHADVSDQMYAAYAEGEDLRLVNIIVGREALSERDNRYLDFFA
avr_Mja IYPPINVLPSLGLMNSGIG--AGKTREDDHKKVSNQYAYAAYAEGESLRLVAVVGEALTRDRAYLKFA
evr_Ehi IQPPIDVLPGLSRLKDKGTG--AGKTREDDHATMNLFAAYAQQKQAKELAVVYGESALSDIDKIIYAKFA
evr_Tth IYPPIDVLPGLSRLMNSGIG--AGKTREDDHKKVSDQLYAYANGVDIRKLVAVVGEALSDIDKIIYAKFA
avr_Sac IYPPINVLPSLGLMNSGIG--EGKTRDDHKKVSNQLFAAYARAQDIRGLAVVGEALSDIDKIIYAKFA
cfr_Osa IRPAINVGISVSRVGSAAQI--KAM----KQVAGKSKLELAQFAELQFAQFASALDKTSQNQLARG
cfr_Zma IRPAINVGISVSRVGSAAQI--KAM----KQVAGKSKLELAQFAELQFAQFASALDKTSQNQLARG
cfr_Tae IRPAINVGISVSRVGSAAQI--KAM----KQVAGKSKLELAQFAELQFAQFASALDKTSQNQLARG
cfr_Nta IRPAINVGISVSRVGSAAQI--KAM----KQVAGKSKLELAQFAELQFAQFASALDKTSQNQLARG
cfr_Sol IRPAINVGISVSRVGSAAQI--KAM----KQVAGKSKLELAQFAELQFAQFASALDKTSQNQLARG
cfr_Psa IRPAINVGISVSRVGSAAQI--KAM----KQVAGKSKLELAQFAELQFAQFASALDKTSQNQLARG
cfr_Mpo IRPAINVGISVSRVGSAAQI--KAM----KQVAGKSKLELAQFAELQFAQFASALDKTSQNQLARG
cfr_Cpa IRPAINVGISVSRVGSAAQI--KAM----KQVAGKSKLELAQFAELQFAQFASALDKTSQNQLARG
cfr_Ppu IRPAINVGISVSRVGSAAQI--KAM----KQVAGKSKLELAQFAELQFAQFASALDKTSQNQLARG
efr_Syn LRPAINAGISVSRVGSAAQI--KAM----KQVAGKSKLELAQFAELQFAQFASALDKTSQNQLARG
efr_Bsu VRPAINAGISVSRVGSAAQI--KAM----KQVAGKSKLELAQFAELQFAQFASALDKTSQNQLARG
efr_Ehi TRPAVDAGISVSRVGSAAQI--KAM----KQVAGKSKLELAQFAELQFAQFASALDKTSQNQLARG
efr_Pmo FRPAINAGISVSRVGSAAQI--KAM----KQVAGKSKLELAQFAELQFAQFASALDKTSQNQLARG
mfr_Osa IRPAINVGLSVSRVGSAAQL--KAM----KQVAGKSKLELAQFAELQFAQFASALDKTSQNQLARG
mfr_Zma IRPAINVGLSVSRVGSAAQL--KAM----KQVAGKSKLELAQFAELQFAQFASALDKTSQNQLARG
mfr_Bta IRPAINVGLSVSRVGSAAQT--RAM----KQVAGKSKLELAQFAELQFAQFASALDKTSQNQLARG
mfr_Hsa IRPAINVGLSVSRVGSAAQT--RAM----KQVAGKSKLELAQFAELQFAQFASALDKTSQNQLARG
mfr_Dme IRPAINVGLSVSRVGSAAQT--RAM----KQVAGKSKLELAQFAELQFAQFASALDKTSQNQLARG
mfr_Spu IYPAINVGLSVSRVGSAAQT--KAM----KQVAGKSKLELAQFAELQFAQFASALDKTSQNQLARG
mfr_Ncr YRPAINVGLSVSRVGSAAQL--KAM----KQVAGKSKLELAQFAELQFAQFASALDKTSQNQLARG
mfr_Rbl IRPAINVGLSVSRVGSAAQT--SAM----KQVAGKSKLELAQFAELQFAQFASALDKTSQNQLARG
mfr_Cre IRPALNVGLSVSRVGSAAQF--PGM----KQVAGKSKLELAQFAELQFAQFASALDKTSQNQLARG
efr_Sli QRPALNVGISVSRVGSAAQH--KAM----KQVAGKSKLELAQFAELQFAQFASALDKTSQNQLARG
efr_Eco IRPAINVGLSVSRVGSAAQT--KIM----KQVAGKSKLELAQFAELQFAQFASALDKTSQNQLARG
avr_Mba VLPADVYVYKSVSRVYVYKAVL--AAY----REVAAGDLKLAYSQFELETTFRFGAK--LDENTRRLIEHG
cfc_Osa IYPAVDPLDSTSTMLQP-----RIVGNEHYETAQRVKQTLQRYKELQDIIAILGLDELSEEDRLTVARA
cfc_Zma IYPAVDPLDSTSTMLQP-----RIVGNEHYETAQRVKQTLQRYKELQDIIAILGLDELSEEDRLTVARA
cfc_Tae IYPAVDPLDSTSTMLQP-----RIVGNEHYETAQRVKQTLQRYKELQDIIAILGLDELSEEDRLTVARA
cfc_Nta IYPAVDPLDSTSTMLQP-----RIVGNEHYETAQRVKQTLQRYKELQDIIAILGLDELSEEDRLTVARA
cfc_Sol IYPAVDPLDSTSTMLQP-----RIVGNEHYETAQRVKQTLQRYKELQDIIAILGLDELSEEDRLTVARA
cfc_Psa IYPAVDPLDSTSTMLQP-----RIVGNEHYETAQRVKQTLQRYKELQDIIAILGLDELSEEDRLTVARA
cfc_Mpo IYPAVDPLDSTSTMLQP-----WIVGEEHYETAQRVKQTLQRYKELQDIIAILGLDELSEEDRLTVARA
cfc_Pth IYPAVDPLDSTSTMLQP-----WIVGEEHYETAQRVKQTLQRYKELQDIIAILGLDELSEEDRLTVARA
cfc_Cpa IYPAVDPLDSTSTMLQP-----GIVGEEHYETAQRVKQTLQRYKELQDIIAILGLDELSEEDRLTVARA
cfc_Ppu IYPAVDPLDSTSTMLQP-----GIVGEEHYETAQRVKQTLQRYKELQDIIAILGLDELSEEDRLTVARA
efc_Syn IYPAVDPLGSTSNMLQP-----DIVGSEHYETAQRVKQTLQRYKELQDIIAILGLDELSEEDRLTVARA
mfc_Osa IYPAVDPLDSTSRMLSP-----HVLGEDHYNTARGVQKVLQNYKNLQDIIAILGMDELSEEDKLTIVARA
mfc_Zma IYPAVDPLDSTSRMLSP-----HVLGEDHYNTARGVQKVLQNYKNLQDIIAILGMDELSEEDKLTIVARA
mfc_Nta IYPAVDPLDSTSRMLSP-----HVLGEDHYNTARGVQKVLQNYKNLQDIIAILGMDELSEEDKLTIVARA
mfc_Hbr IYPAVDPLDSTSRMLSP-----HVLGEDHYNTARGVQKVLQNYKNLQDIIAILGMDELSEEDKLTIVARA
mfc_Bta IYPAVDPLDSTSRMLSP-----HVLGEDHYNTARGVQKVLQNYKNLQDIIAILGMDELSEEDKLTIVARA
mfc_Rno IYPAVDPLDSTSRMLSP-----HVLGEDHYNTARGVQKVLQNYKNLQDIIAILGMDELSEEDKLTIVARA
mfc_Sce IYPAVDPLDSTSRMLSP-----HVLGEDHYNTARGVQKVLQNYKNLQDIIAILGMDELSEEDKLTIVARA
efc_Rbl IYPAVDPLDSTSRMLSP-----HVLGEDHYNTARGVQKVLQNYKNLQDIIAILGMDELSEEDKLTIVARA
mfc_Ncr IYPAVDPLDSTSRMLSP-----HVLGEDHYNTARGVQKVLQNYKNLQDIIAILGMDELSEEDKLTIVARA
efc_Eco IYPAVDPLDSTSRMLSP-----HVLGEDHYNTARGVQKVLQNYKNLQDIIAILGMDELSEEDKLTIVARA
efc_Pmo IYPAVDPLDSTSRMLSP-----HVLGEDHYNTARGVQKVLQNYKNLQDIIAILGMDELSEEDKLTIVARA
efc_Bsu IYPAVDPLDSTSRMLSP-----HVLGEDHYNTARGVQKVLQNYKNLQDIIAILGMDELSEEDKLTIVARA
efc_Ehi IYPAVDPLDSTSRMLSP-----HVLGEDHYNTARGVQKVLQNYKNLQDIIAILGMDELSEEDKLTIVARA
efc_Sli IYPAVDPLDSTSRMLSP-----HVLGEDHYNTARGVQKVLQNYKNLQDIIAILGMDELSEEDKLTIVARA
afc_Mba IYPAVDPLDSTSRMLSP-----HVLGEDHYNTARGVQKVLQNYKNLQDIIAILGMDELSEEDKLTIVARA
LPAIDVPLDSTSRMLSP-----HVLGEDHYNTARGVQKVLQNYKNLQDIIAILGMDELSEEDKLTIVARA
HFPISINTSVSYSKYITLDKQWY--GREYDPFRLRDRIRQLSDSEELQVYVQLVQKASLSDSDKITLDTVA
vvc_Sce HFPISINTSVSYSKYTNVNLKQWY--DSNYPFPVLRDRMKEILSNAEELQVYVQLVQKASLSDSDKITLDTVA
vvc_Hvu HFPISINTSVSYSKYSTALEGYY--KFDPPGDFIDMRTKAREVLRQEDDLNETVQLVQKDALGESDKITLDTVA
vvc_Zma HFPISINTSVSYSKYSTALEGYY--KFDPPGDFIDMRTKAREVLRQEDDLNETVQLVQKDALGESDKITLDTVA
vvc_Bta HFPISINTSVSYSKYMRALDEYY--DKHFTFVPLRTRKAKEVLRQEDDLNETVQLVQKDALGESDKITLDTVA
vvc_Dme HFPISINTSVSYSKYMRALDEYY--DKHFTFVPLRTRKAKEVLRQEDDLNETVQLVQKDALGESDKITLDTVA
vvc_Tco HFPISINTSVSYSKYMRALDEYY--DKHFTFVPLRTRKAKEVLRQEDDLNETVQLVQKDALGESDKITLDTVA
avc_Mja HFPISINTSVSYSKYMRALDEYY--DKHFTFVPLRTRKAKEVLRQEDDLNETVQLVQKDALGESDKITLDTVA
evc_Ehi HFPISINTSVSYSKYMRALDEYY--DKHFTFVPLRTRKAKEVLRQEDDLNETVQLVQKDALGESDKITLDTVA
avc_Mba HFPISINTSVSYSKYMRALDEYY--DKHFTFVPLRTRKAKEVLRQEDDLNETVQLVQKDALGESDKITLDTVA
avc_Mma HFPISINTSVSYSKYMRALDEYY--DKHFTFVPLRTRKAKEVLRQEDDLNETVQLVQKDALGESDKITLDTVA
avc_Hsa HFPISINTSVSYSKYMRALDEYY--DKHFTFVPLRTRKAKEVLRQEDDLNETVQLVQKDALGESDKITLDTVA
avc_Hvo HFPISINTSVSYSKYMRALDEYY--DKHFTFVPLRTRKAKEVLRQEDDLNETVQLVQKDALGESDKITLDTVA
evc_Tth HFPISINTSVSYSKYMRALDEYY--DKHFTFVPLRTRKAKEVLRQEDDLNETVQLVQKDALGESDKITLDTVA
avc_Sac HFPISINTSVSYSKYMRALDEYY--DKHFTFVPLRTRKAKEVLRQEDDLNETVQLVQKDALGESDKITLDTVA
HYPAINWIGFSAAYVDLVAQWYHKNVDPNKEMRDTAMKVLITREDELRQIVRLVQVPEELAEKDKLVLEAA

```

Multiple sequence alignment of protein sequences from various species. The alignment shows conserved regions across different taxa, including mammals (e.g., vvr\_Ncr, vvr\_Sce), birds (e.g., vvr\_Bta, vvr\_Dme), and other vertebrates (e.g., vvr\_Hvu, vvr\_Hco). The sequences are aligned in columns, with gaps indicated by dashes. The alignment is flanked by '+' signs at the top, indicating conserved regions. The sequences are labeled on the left with their respective species and protein names.

```

+ + + + + + +
vvr_Ncr GKGKDKPTTKDTRDTAAPEEENLIDA-----
vvr_Sce DEDEEDPTQEAPVRRRTPAKKNL-----
vvr_Bta -----
vvr_Dme -----
vvr_Hvu -----
vvr_Tco -----
avr_Mba -----
avr_Mma -----
avr_Hsa TAD-----
avr_Hvo EAER-----
avr_Mja -----
evr_Ehi -----
evr_Tth GAPQALD-----
avr_Sac -----
cfr_Osa IISS-SKTFTEEAAILLKEAIEQLERFSLQEQT-----
cfr_Zma IISS-SKTFTEQAETLLKEAIEQLERFSLQEQT-----
cfr_Tae IISS-SKTFTEQAETLLKEAIEQLERFSLQ-----
cfr_Nta IISS-TKTFTEEAAILLKEAIEQMDFILQEQA-----
cfr_Sol IISS-TKTFTEEAAILLKEAIEQMERFLLQEQA-----
cfr_Psa IISS-TKTFTEEAAILLKEAIEQMELF-----
cfr_Mpo IIRS-TKVFTQAENLLKEAITEHIELFLFQEEK-----
cfr_Cpa IIRS-EKALTAEAENLLKDAISDCKQAFLSNI-----
cfr_Ppu SIRS-SKKLDTASEELLKKAIEDVKQGFVKA-----
efr_Syn AARS-SQKLDATAEAILLKEAIAEYKAGFTA-----
efr_Bsu GIAT-TGNLPADED--FKAIEGFKRTFAPSN-----
efr_Ehi TIRT-TKDLPEEER--LNQAIQEYKDIPLATKGNTSSTEDKLSIQNA-----
efr_Pmo ELKT-TQLFS-----
mfr_Osa EKGGLTNERKIEPDASLKQTAKEIN-----
mfr_Zma EKGGLTNERKMEPDASLKESALNL-----
mfr_Bta KIRTDG-KISEESDAKLKEIVTNFLAGFEA-----
mfr_Hsa TIRADG-KISEQSDAKLKEIVTNFLAGFEA-----
mfr_Dme TIAKDG-AISEASDAKLKDIVAKFMSTFQG-----
mfr_Spu TIRTEG-QLSPQTEATLKDVVLKFLDTFEG-----
mfr_Ncr TIDKEG-AISKDLEAKLRDVIQTFTKSLG-----
efr_Rbl DMTKNDRKVVAGDLEKAIKRAELDAFAKDA-----
mfr_Cre AKGKITPEINAHLAQQMSNLPVMTK-----
efr_Sli SIREGGK-MSDDTLQAVAEIAAFKKQFETSDGKLLGEDAPSAK-----
efr_Eco FI-NQTGGYNDIEGKLGILDSFKATQSW-----
avr_Mba RFKG-DKELSDKDREILNIARKALEPYQPKPESESKPEAKTEENAESQTEDEKSESEAKTEEKVAKSE
cfc_Osa ENKLLK-----
cfc_Zma ESKLLK-----
cfc_Tae ENKSQK-----
cfc_Nta ESNLKK-----
cfc_Sol ESKLLK-----
cfc_Psa -----
cfc_Mpo ES-----
cfc_Pth ES-----
cfc_Cpa EAK-----
cfc_Ppu -----
efc_Syn -----
mfc_Osa ESAS-----
mfc_Zma ESAS-----
mfc_Nta ESAA-----
mfc_Hbr ESAS-----
mfc_Bta EHS-----
mfc_Rno EHGS-----
mfc_Sce EAN-----
efc_Rbl AAA-----
mfc_Ncr ELEGQA-----
efc_Eco -----
efc_Pmo GAE-----
efc_Bsu EV-----
efc_Ehi -----
efc_Sli S-----
afc_Mba S-----
vvc_Ncr ATQDLQAQLKSLKFEVPS-EGQEKICKKYEAIQQQMLDKFASVIDE-----
vvc_Sce STGDVKHAVSSSKFFEPS-RGEKEVHGFEKLLSTMQERFAESTD-----
vvc_Hvu RMGDLFYRLVSQKFEDPA-EGEDVLVAKFKKLYDDLTAGFRNLEDEAR-----
vvc_Zma RLGDLYRLVSQKFEDPA-EGEEALVGKFKKLYDDLTTGFRNLEDEAR-----
vvc_Bta HMGEILYKLSMFKDPVKDGEAKIKADYAQLLEDQMNAFRSLED-----
vvc_Dme SMGGIMYQLSSMFKDPVKDGEQKIKADYDQLYEDLQAFRNLED-----
vvc_Tco MIPHITGLTENKFRDP-QEGEEANVEFYRKQNEEIVSAFASLLQ-----
avc_Mja DIARMKY-IPHDEFINVK-----KEIMEKIKNELGSLN-----
evc_Ehi RISRSKY-IPPEELAKISSINEE-----IKETIQILIVSEGGMTDD-----
avc_Mba VLAKV--KYEKF-----DES-----MNAVLAQMOKEFASLRGR-----
avc_Mma SLPKV--KYEENF-----EDS-----LNAVLAQMOKEFALGGK-----
avc_Hsa RLNRIGT-TPDDEH-----EAE-----VAEIKQQITEQLRELY-----
avc_Hvo RLNRIGV--QEDW-----EAY-----IEDLKAITEQLRELY-----
evc_Tth RIGRARY-VSEEF-----PAY-----FEEAMKEIQGAFKAGLKGERTWF-----
avc_Sac EIIRIKYTIKNDLKNIDEIENK-----LKFATFDSLLKEVS-----

```



```

+           +           +
vvr_Ncr -----
vvr_Sce -----
vvr_Bta -----
vvr_Dme -----
vvr_Hvu -----
vvr_Tco -----
avr_Mba -----
avr_Mma -----
avr_Hsa -----
avr_Hvo -----
avr_Mja -----
evr_Ehi -----
evr_Tth -----
avr_Sac -----
cfr_Osa -----
cfr_Zma -----
cfr_Tae -----
cfr_Nta -----
cfr_Sol -----
cfr_Psa -----
cfr_Mpo -----
cfr_Cpa -----
cfr_Ppu -----
efr_Syn -----
efr_Bsu -----
efr_Ehi -----
efr_Pmo -----
mfr_Osa -----
mfr_Zma -----
mfr_Bta -----
mfr_Hsa -----
mfr_Dme -----
mfr_Spu -----
mfr_Ncr -----
efr_Rbl -----
mfr_Cre -----
efr_Sli -----
efr_Eco -----
afr_Mba TQTKGKPETETKTEEKPEFGTRTEEK
cfc_Osa -----
cfc_Zma -----
cfc_Tae -----
cfc_Nta -----
cfc_Sol -----
cfc_Psa -----
cfc_Mpo -----
cfc_Pth -----
cfc_Cpa -----
cfc_Ppu -----
efc_Syn -----
mfc_Osa -----
mfc_Zma -----
mfc_Nta -----
mfc_Hbr -----
mfc_Bta -----
mfc_Rno -----
mfc_Sce -----
efc_Rbl -----
mfc_Ncr -----
efc_Eco -----
efc_Pmo -----
efc_Bsu -----
efc_Ehi -----
efc_Sli -----
afc_Mba -----
vvc_Ncr -----
vvc_Sce -----
vvc_Hvu -----
vvc_Zma -----
vvc_Bta -----
vvc_Dme -----
vvc_Tco -----
avc_Mja -----
evc_Ehi -----
avc_Mba -----
avc_Mma -----
avc_Hsa -----
avc_Hvo -----
evc_Tth -----
avc_Sac -----

```

## 謝 辞

本研究をまとめるにあたり、標品の供与や技術的指導、研究上の協力において多くの研究者の方々のお世話になりました。その中でも特に、取り扱いに特殊な機器や技術を要求する絶対嫌気性のメタン産生菌の培養に関しては、産業医科大学教授・古賀洋介先生の直接のご指導を、また、最終章の系統樹の作成については、統計数理研究所教授・長谷川政美先生のご助言をいただくことができました。本研究を通じて、このような分野ごとの権威の先生がたと関わりを持てたことをたいへんな誇りと思い、今後の研究活動に糧にしてゆきたいと考えております。そして長年にわたり、このような恵まれた機会と研究環境ををお与えくださった吉田賢右先生、大島泰郎先生に心からお礼をもうしあげます。