

主論文

全ゲノム配列に基づくABC輸送系の比較解析

富井 健太郎

目次

概要	2
略語	3
1. 序論	4
1.1. ABC輸送系の生物学的的重要性	4
1.2. ゲノム配列解析の進歩と本研究における意味	4
2. 方法	7
3. 結果	10
3.1. ATP結合蛋白質の分類 (ortholog テーブル)	10
3.2. グループ1	20
3.3. グループ2	26
3.4. グループ3	26
3.5. グループ4	31
3.6. グループ5	31
3.7. グループ6	33
3.8. グループ7	34
3.9. グループ8	34
3.10. その他のグループ	36
4. 考察	37
4.1. オペロンの機能推定	37
4.2. 7生物種におけるオペロンの保存性	38
4.3. <i>Synechocystis</i> におけるABC輸送系遺伝子の特徴	41
4.4. 既存の分類との比較	41
4.5. ABC輸送系オペロンの進化	47
5. 結論	49
文献	50
謝辞	56

概要

本論文は、近年その全ゲノム配列が明らかにされた7種類の微生物（真正細菌から6種(*E. coli*, *H. influenzae*, *H. pylori*, *M. genitalium*, *M. pneumoniae*, *Synechocystis*)、古細菌から1種(*M. jannaschii*)）ゲノム上での、ABC輸送系オペロンの保存性を比較解析した研究成果である。

本研究ではまず、7生物種の持つABC輸送系のATP結合蛋白質、およびこれに類似した蛋白質全てを、orthologous関係とアミノ酸配列類似性に基づいて分類を行い、ATP結合蛋白質のorthologテーブルを構築した。そして次に各生物種のゲノム上で、それらの蛋白質をコードしている遺伝子の周辺遺伝子の保存性を、分類されたグループ内で解析した。これにより、ATP結合蛋白質の分類が、その含まれるであろうオペロンの構造に非常に良い結びつきがあることを示した。つまり、真正細菌の進化を通して、orthologous遺伝子のゲノム上での配置は専らシャッフルされているという報告[Tatusov *et al.*, 1996; Watanabe *et al.*, 1997]とは対照的に、一般にABC輸送系の遺伝子群は、生物種を越えて、さらには生物界のドメインを越えて非常に強固に集積し、オペロンを形成しているであろうことを明らかにした。今回比較を行った7生物種の中で、*Synechocystis*ゲノムでのみ、ABC輸送系オペロンの保存性に低下がみられる。これらの結果から、複数のABC輸送系オペロンの起源は古細菌と真正細菌の分岐以前にさかのぼること、また、各々のオペロンが一つの単位として進化してきたことが示唆される。

またこの保存性の解析を通して、次の幾つかの機能推定を行うことができた。これまでに報告されていなかった、新規のオペロン構成要素を4種類発見した。このうち一つは、*E. coli*におけるシグナル伝達に関与すると考えられる蛋白質である。さらに、接着および細胞壁の構築への関与が示唆されるABC輸送系オペロンの一群を同定することができた。*E. coli*の細胞分裂への関与が知られているオペロンの構成要素にも共有される、新規のアミノ酸配列モチーフを持つ蛋白質の一群を同定した。また、オペロンを形成せず、輸送以外の機能を担うであろうATP結合蛋白質の一群を同定することができた。*E. coli*でのABC輸送系オペロンとべん毛遺伝子の発現制御との関連についても示すことができた。

そしてさらに、既に発表されているABC輸送系オペロンのATP結合蛋白質以外の要素（膜蛋白質と基質結合蛋白質）の分類と、本研究の分類結果との比較から、7つのABC輸送系オペロンの祖先型を提案した。

本研究で得られたオペロンの機能に関する新たな知見は、今後、我々の研究室で構築しているKEGGデータベース(<http://www.genome.ad.jp/kegg/>)を通して公開される予定である。

略語

Ec: *Escherichia coli*

Hi: *Haemophilus influenzae*

Hp: *Helicobacter pylori*

Mg: *Mycoplasma genitalium*

Mp: *Mycoplasma pneumoniae*

Ss: *Synechocystis* sp.

Mj: *Methanococcus jannaschii*

ABC: ATP-binding cassette

1. 序論

1.1. ABC輸送系の生物学的的重要性

ABC輸送系は、その発見当初、真正細菌の栄養分摂取システムとして同定された。これらは、グラム陰性菌では、ペリプラズム基質結合蛋白質依存性輸送システムとして、またグラム陽性菌では、基質結合リポ蛋白質依存性輸送システムとして、現在までに多数の系が発見されている[Higgins, 1992]。

ABC輸送系は、幾つかの共通した特徴を有している。これらの系はいずれも、アミノ酸配列上、比較的良く保存されたATP結合ドメインを持ち、これが名前の由来 (ATP-binding cassette) ともなっている。このATP結合ドメインで行われるATP加水分解が、その詳細な機構は未だ不明であるものの、能動輸送のエネルギー供給源となっていることが示唆されている。現在までに明らかにされている典型的なABC輸送系は、4つのドメインから成る。うち2つは既に述べたATP結合ドメインであり、これら以外の2つは、各々6本の膜貫通領域を持つとされる膜内在性のドメインである。最初に述べた、ペリプラズム結合蛋白 (あるいは、リポ蛋白) の有無は、系に依ることが知られている。一般にこれらのドメインは、真正細菌では、別々の蛋白質として発現し、細胞中で互いに結合して機能単位を構成していると考えられている (図1)。真核生物では、これらのドメインが融合し、一つの蛋白質として発現している。

またABC輸送系は、真核高等生物においても、生物学的、医学的見地から見て非常に重要な現象、例えば、のう胞性繊維症、免疫系における抗原提示、がん細胞の薬剤耐性等への関連[Higgins, 1992]、さらには、神経伝達物質受容体との配列類似性が指摘されている[杉山等, 1997]。このようにABC輸送系は、真正細菌から真核高等生物に至るまで幅広く存在し、細胞にとって重要な系であると考えられ、非常に興味深い研究対象である。

1.2. ゲノム配列解析の進歩と本研究における意味

近年、急速な勢いで、様々な生物種の全ゲノム配列が明らかになりつつある。これにより、例えば化合物の代謝、合成、輸送経路やシグナル伝達経路といった各生物種のもつ分子間相互作用のネットワーク全体を、計算機上で自動的に比較することが可能になる[富井・金久, 1997]。これには、異なる生物種で同じ機能を果たすと考えられる orthologous 遺伝子の同定が必須である。

本研究では、近年その全ゲノム配列が明らかにされた7種類の微生物 (グラム陰性細菌から3種、グラム陽性細菌から2種、シアノバクテリアから1種の以上、真正細菌

から6生物種。そして古細菌から1種)の輸送、特に上述したABC輸送系に着目し、比較解析を行った。この輸送系に注目した理由は、既に述べた生物学的な機能の重要性以外に、ゲノム上での多量のABC輸送系遺伝子の存在[Tatusov *et al.*, 1996; Blattner *et al.*, 1997]が挙げられる。これまでに、orthologous遺伝子の配置は、真正細菌ゲノム上で専らシャッフルされているという報告[Tatusov *et al.*, 1996; Watanabe *et al.*, 1997]がある。これに対してオペロンの構造が、生物種、さらには生物界のドメインを越えて保存されているか否かを調べることは、非常に興味深い。このような意味から、ゲノム上に多量に存在し、かつオペロンを形成することが知られているABC輸送系遺伝子は、恰好の対象であると言える。

またこれまで、ABC輸送系オペロンの構成要素である基質結合蛋白質や膜蛋白質に基づく分類[Tam and Saier, 1993; Saurin and Dassa, 1994]は発表されているが、ATP結合蛋白質の全体を通じた系統的な分類は、本研究が最初である。

本研究では、まずABC輸送系のATP結合蛋白質、あるいはこれに類似した蛋白質を、アミノ酸配列類似性およびorthologous関係に基づき分類を行い、次に各生物種のゲノム上で、それらの蛋白質をコードしている遺伝子の周辺の遺伝子の保存性について解析を行った。さらに、ABC輸送系オペロンの進化についても考察を行った。

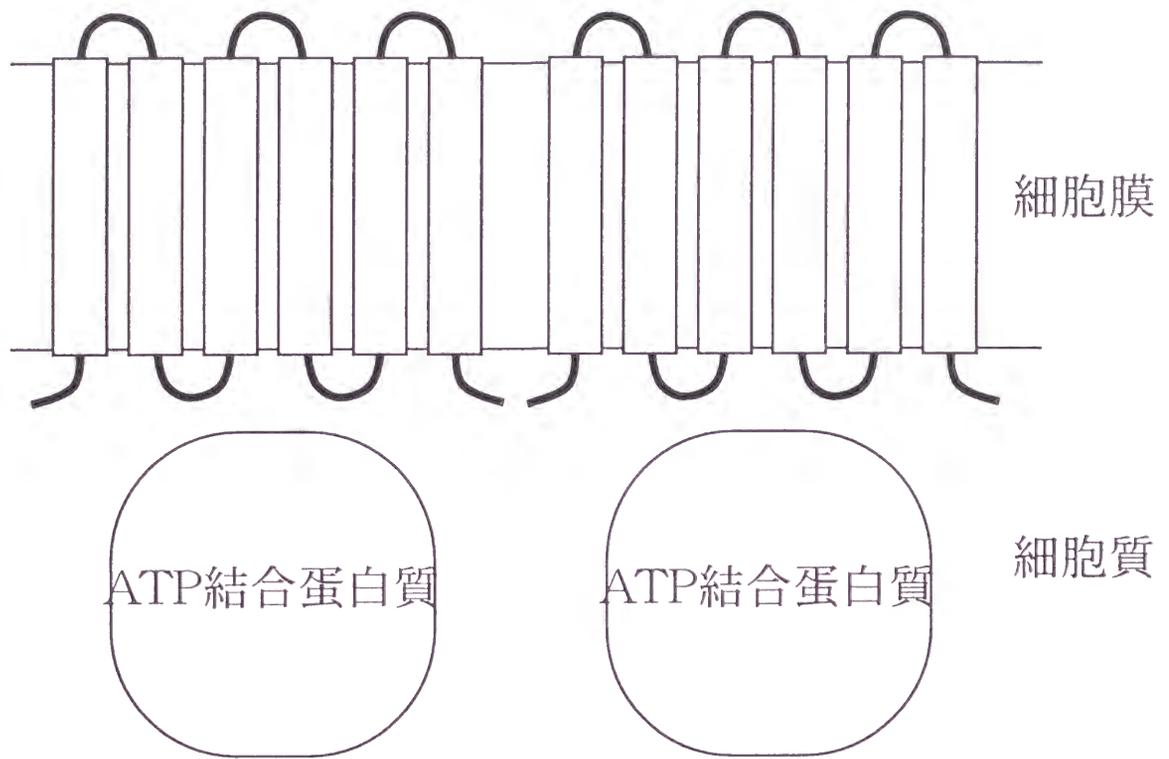


図1 典型的なABC輸送系のサブユニット構造の模式図。膜貫通領域は矩形で表した。

2. 方法

配列データ

本研究の解析に用いた生物種は、次の7種である。即ち、グラム陰性菌から次の3種、*E. coli* [Blattner *et al.*, 1997], *H. influenzae* [Fleischmann *et al.*, 1995], *H. pylori* [Tomb *et al.*, 1997], グラム陽性菌から次の2種、*M. genitalium* [Fraser *et al.*, 1995], *M. pneumoniae* [Himmelreich *et al.*, 1996], シアノバクテリアから1種、*Synechocystis* sp. PCC6803 [Kaneko *et al.*, 1996], の以上真正細菌から6種、そして古細菌から1種、*M. jannashii* [Bult *et al.*, 1996], である。

これら7生物種の持つ蛋白質のアミノ酸配列とそれらの機能に関する注釈は、我々の研究室で構築しているKEGGデータベース(<http://www.genome.ad.jp/kegg/>)から取得した。遺伝子の位置情報およびDNA塩基配列については、GenBankのcomplete genomes section (<ftp://ncbi.nlm.nih.gov/genbank/genomes/bacteria/>)の1997年9月10日版のデータを利用した。

配列類似性の基準およびorthologの定義

アミノ酸配列の比較には、SSEARCH [Pearson, 1991]を用いた。使用したパラメータは、ギャップの開始および伸長のコストが、それぞれ-12と、-2、アミノ酸置換行列がBLOSUM50 [Henikoff and Henikoff, 1992]である。このアミノ酸置換行列は、相同性解析に有効な置換行列群 [Tomii and Kanehisa, 1996]の一つと同じ手法で作られている。

配列類似性の基準には、データベース検索（後述）を行う時にSSEARCHプログラムで計算される期待値（E value）を使用した。期待値が0.001未満である時、配列類似性があると定義する。この設定は、立体構造既知の蛋白質分類の際、偽陽性がなかった [Brenner *et al.*, 1995] ことによる。この期待値には、厳密な対称性がないので、大きな期待値の方を各配列ペアの値とした。

2生物種間でデータベース検索を行って、ある蛋白質ペアが上記の配列類似性の基準を満たして、かつお互いに最高の期待値とアライメントスコアを示す場合に限り、その蛋白質ペアはorthologであるとした。

ATP結合蛋白質の分類（ortholog テーブルの構築）

ABC輸送系のATP結合蛋白質の、アミノ酸配列類似性およびorthologous関係に基

づく分類（ortholog テーブルの構築）は、図2にも示すように次の4段階を経て行われた。

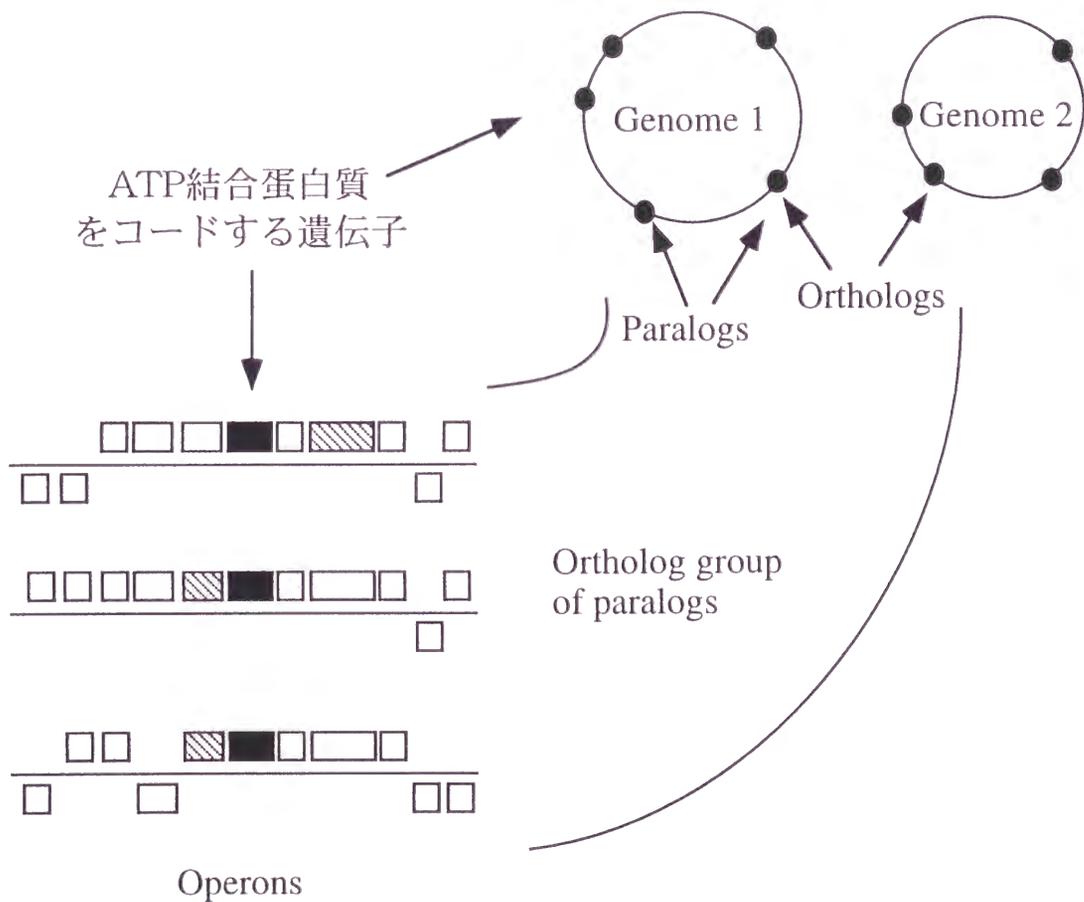
まず最短距離法を用いて、各生物種内でABC輸送系のATP結合蛋白質のparalogous遺伝子（産物）の同定を行った。この時の各配列間の距離には、E valueを用い、閾値は、 $E < 0.001$ に設定した。このために、あらかじめ各生物種内で、全配列を問い合わせ配列としたデータベース検索を行った。次に、各生物種内での明らかな重複、これは例えば、Ecにおけるペプチド輸送系や糖輸送系等、配列類似性からみても機能的にも非常に良く似たグループ、を同定するために、各生物種内でATP結合蛋白質のクラスター解析を行った。これには、1000を各ATP結合蛋白質間のアライメントスコアで割った値を、その蛋白質ペアの距離とし、最短距離法と最長距離法による階層的クラスター解析を行い、両手法で同じ構成要素を持つクラスターとして同定した。この時点の分類が、ortholog テーブルにおける分類の基礎となる。

3番目に、異なる生物種の持つATP結合蛋白質間のorthologous関係を知るために、各生物種の持つATP結合蛋白質の配列を問い合わせ配列とし、他生物種へのデータベース検索を行った。今回のorthologの定義は、上に述べた通りである。最後に、2段階目で得られたクラスターと、3段階目で得られたorthologous関係を基にortholog テーブルを構築した。テーブル構築には、orthologous関係を最優先した。また2段階目で得られたクラスターは同一グループに属するよう分類した。2段階目で複数の構成要素を持つクラスターに属さなかった配列は、類似性（アライメントスコア）に従って各グループに分類した。

こうして、ortholog テーブル構築後、各グループ内でATP結合蛋白質をコードしている遺伝子近傍（前後各々5）の遺伝子の保存性を、その産物のアミノ酸類似性から探索した。このためにまた、ATP結合蛋白質周辺の配列を問い合わせ配列とし、他生物種へのデータベース検索を行った。これらの結果は、3.2.以降に示す。

解析手順

- 1) 各生物種からABC輸送系ATP結合蛋白質を同定
- 2) 各生物種内で、ABC輸送系ATP結合蛋白質をクラスタリング。
- 3) 生物種間で、orthologを同定。
- 4) ABC輸送系ATP結合蛋白質のorthologテーブルを構築。
- 5) 分類したグループ内で、ATP結合蛋白質周辺の遺伝子を比較。遺伝子の配置の保存性を解析。



3. 結果

この章ではまず、ATP結合蛋白質の分類結果について述べる。そして次に、分類された各グループ内での、ATP結合蛋白質をコードしている遺伝子の周辺遺伝子の保存性、つまりオペロン構造について、各グループ毎に順次述べて行く。

3.1. ATP結合蛋白質の分類 (ortholog テーブル)

前章2で述べたように、まず最短距離法を用いて、各生物種内でABC輸送系のATP結合蛋白質のparalogous遺伝子(産物)の同定を行った。Ecでは、83配列が同定されたが、このうちDNA修復に関与するUvrA(b4058)と、Ec内にあるATP結合蛋白質と膜蛋白質との融合蛋白質の影響で混入された膜蛋白質と推定される4配列を除いた78配列を、今後の解析に用いた。Ssでも、UvrA(slr1844)と、同様の理由から14膜蛋白質、さらに非常に短い2つのATP結合蛋白質(ssr3570, ssr2802)を除いた54配列を今後の解析に用いた。Hi, Hp, Mpでも、UvrAに相当する配列を除いた。これらの結果は表1にまとめた。

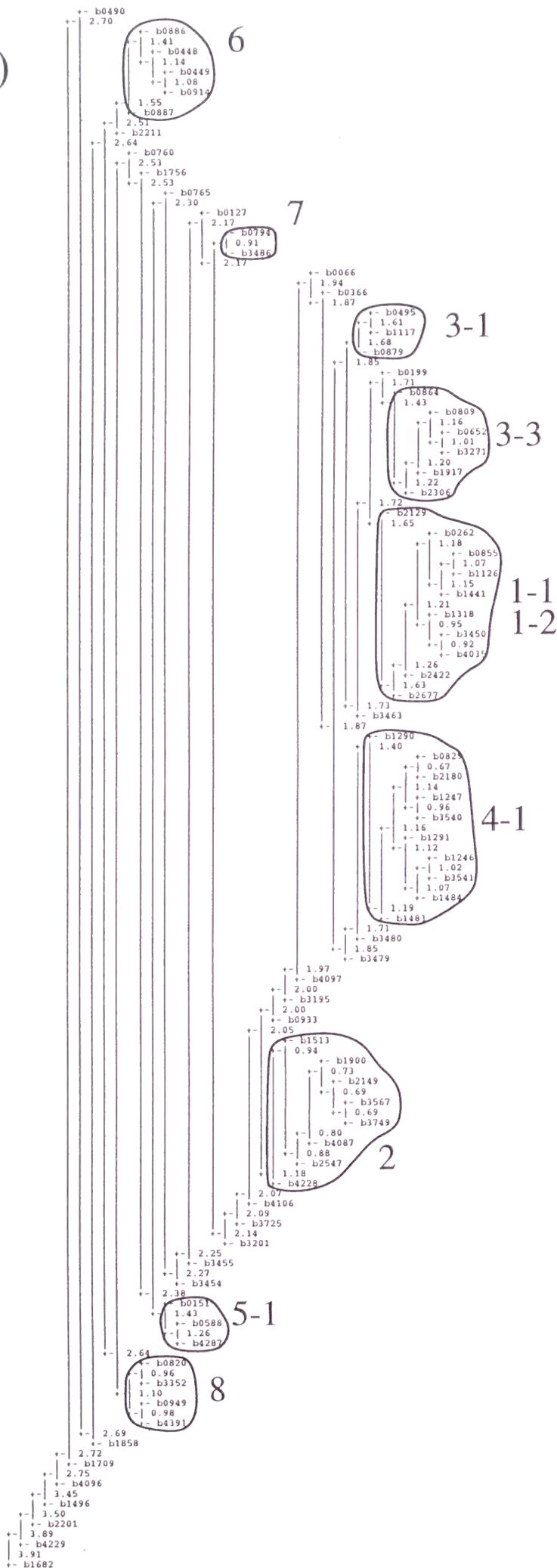
次に、各生物種内での重複を推定するために、各生物種内でATP結合蛋白質のクラスター解析を行った。ここ形成されたクラスターが、この後のATP結合蛋白質全体のグループ分けの基礎となる。ここでは、例としてEcでのクラスター解析の結果を図3に示す。ここで得られるクラスターは、例えばEcにおけるペプチド輸送系や糖輸送系等、配列も機能も良い類似を示すATP結合蛋白質が形成している、非常に分離の良い集合であることがわかる。7生物種における結果は、表1にまとめた。これらのクラスターは、基本的には、最短距離法と最長距離法の2つの手法で、同一要素を含むクラスターである。これをより明確に同定するために、Ecの場合、最短距離法では590点、最長距離法では470点の閾値を導入した。他の生物種については、およそこれと同じ値だが、閾値を調整して導入した。

そしてこの結果と既に得ているorthologous関係から、ortholog テーブルを構築した。これを、表2に示す。表2では、orthologous関係にあるものが同じ行に示されている。グループ分類は、先程のクラスターが基礎となっている。3生物種以上の間で比較を行うと、orthologous関係のペアに矛盾が生じる場合がある。今回、そうしたペアについては、アライメントスコアの高い順に関係を優先させ、テーブルを構築した。

表 1. 7 生物種におけるABC輸送系ATP結合蛋白質数およびそのクラスター解析の結果

生物種	Gram-negative bacteria			Gram-positive bacteria		Cyano-bacteria	Archaea
	Ec	Hi	Hp	Mg	Mp	Ss	Mj
ATP結合蛋白質数 (A)	78	41	18	16	17	54	17
全蛋白質数(B)	4289	1717	1566	467	677	3166	1680
(A)/(B) (%)	1.8	2.4	1.1	3.4	2.5	1.7	1.0
クラスター数	9	8	2	5	4	13	3
クラスター中のATP結合蛋白質の数	51	30	5	10	9	39	6
単独クラスターのATP結合蛋白質数	27	11	13	6	8	15	11

a)



b)

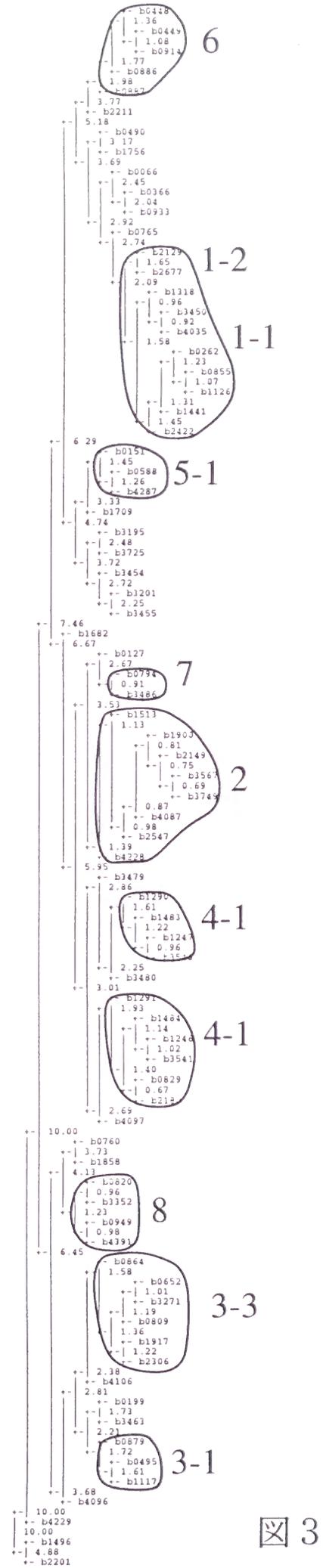


図3 EcでのATP結合蛋白質のクラスター解析の結果

a) 最短距離法 b) 最長距離法の結果を示す。クラスターに付された数字は表2のグループ名を示す。

表2. ABC輸送系ATP結合蛋白質のorthologテーブル

グループ	Gram-negative bacteria			Gram-positive bacteria		Cyano-bacteria	Archaea	4) 機能	
	Ec	Hi	Hp	Mg	Mp	Ss	Mj		
1-1						slr0354		?	
	b0760							molybdate	
	b0765	HI1691							
	b0066	HI1021						a) thiamin	
	b1756							?	
						slr0240		?	
			HP0475			sll0739		molybdate(
	b1318			MG187	E07_orf586	slr0747		?	
	b3450					slr1224		sn-glycerol-phosphate maltose/ maltodextrin	
	b4035							?	
	b0262	HI0126						?	
	b0855			MG042	D09_orf560L			putrescine spermidine putrescine	
	b1126	HI1347						?	
	b1441							?	
	b2422					slr1455	MJ1367	sulfate / thiosu	
1-2	b2677	HI0099				sll1878		b) iron glycine betar L-proline	
	b2129							?	
		HI0354						?	
		HI1474						?	
	b0366					sll1082 slr0043 slr0044 sll1452 sll1453	MJ0412	taurine nitrate nitrate	
	1-3	b3195	HI1087	HP1465			sll1001 sll1041		?
									?
	1-2	b0933							?
				HP0819					?
	1) 1-4	b0490						?	
2	b1513			MG119	A65_orf572			?	
	b1900							L-arabinose galactose	
	b2149	HI0823						?	
	b2547							D-xylose	
	b3567	HI1110						ribose	
	b3749	HI0502						?	
	b4087							?	
	b4228							?	
b4229							?		

表2 (続き)

グループ	Gram-negative bacteria			Gram-positive bacteria		Cyano-bacteria	Archaea	機能
	Ec	Hi	Hp	Mg	Mp	Ss	Mj	
3-1	b0495 b0879* b1117*		HP0179	MG065	R02_orf465		MJ1508*	cell division(
		HI1549*		MG467	K05_orf284 K05_orf339	slr0544 sll0484	MJ0796*	
			HP0748					cell division(
	b3463*	HI0769						cell division(
3-2	b0199	HI0621	HP1576					?
						sll1870		?
						sll0778		?
3-3	b0652 b0809 b0864 b1917* b2306 b3271	HI1180 HI1078	HP1171			slr1735		glutamate/aspa glutamine arginine ? histidine ?
3-4	b4106							phospho-nat
3-5	b3454					slr1881 sll0374	MJ1268	branched- chain
	b3455					slr0467 sll0764	MJ1267	amino acid
3-6	b3201	HI1148	HP0715			slr0251		?
3-7	b3725	2)		MG410	C12_orf329	sll0683 sll0684 slr1250	MJ1012	phosphate
4-1	b0829* b1246 b1291 b1484* b2180* b3541	HI1121 HI1641		MG079	G07_orf423	sll1927	MJ1662	c) ? oligo-peptid peptide ?
	b1247 b1290 b1483* b3540	HI1120 HI1642	HP0250 HP0301				MJ1242	? / c) dipeptide oligo-peptid peptide ?
		HI1184	HP0302	MG080	G07_orf851			dipeptide oligo-peptide nickel
	b3479							
	b3480							
	4-2	b4097						
b4096								

表2 (続き)

グループ	Gram-negative bacteria			Gram-positive bacteria		Cyano-bacteria	Archaea	機能
	Ec	Ili	Hp	Mg	Mp	Ss	Mj	
5-1	b1709							vitamin B ₁₂
		HI1470 HI1272	HP0888	3)MG290	3)A05_orf244		MJ0089	? ?
	b0151 b4287 b0588					slr1318	MJ0873	ferri-chrom iron(III) dicitr ferric entero-ba
5-2	b1858*	HI0361* HI0408*				sll1599* slr2044*		manganese
5-3				MG304*	A05_orf270L*			cobalt(?)
							MJ1572*	
		HI1618*		MG179* MG180*	GT9_orf274* GT9_orf303*	sll0385*	MJ1088*	
				MG303*	A05_orf382*			
6						sll1623		drug resistance(?)
	b0914 b0448* b0449*	HI0060	HP1082	MG014	D12_orf623	sll1725 sll1276 slr0615		
	b0886* b0887*	HI1156* HI1157* HI1051 HI0663 HI0664	HP1206	MG015	D12_orf634	slr1149 slr2019 sll1180 slr1488 slr1651		
						slr1494		
			HP0600					
				MG390				
	b1496	HI0036 HI1467				sll0182		
	b2211							
	7	b0127						
b0794 b3486						sll0759		
					F10_orf326	sll0415 sll0489 slr1113	MJ1023	
						slr1901		
			IIP1220					

表2 (続き)

グループ	Gram-negative bacteria			Gram-positive bacteria		Cyano-bacteria	Archaea	機能
	Ec	Hi	Hp	Mg	Mp	Ss	Mj	
8	b0820* b3352* b0949* b4391*	HI0658* HI1300* HI1342* HI1252*	HP0853*			slr0864* slr0912*	MJ0719*	? d)
9	b1682*					slr0075*	MJ0035*	?
10	b2201	HI1089						heme export
11						slr0575* slr0982* slr2108*		ABC-2(?)
12			HP0613					?
1) 13							MJ0121	?

オペロンを形成する遺伝子には、影をつけて示した。各グループのオペロン構造については、本文を参照。太線は、グループ分類を、中線、細線は、各々アライメントスコアが400点以上、500点以上であることを示す。点線は、クラスター解析の結果による。同じ行にある遺伝子はorthologである。

- 1) 本文参照
 - 2) ATP結合蛋白質は、suspected coding regionにコードされている。
 - 3) MG290 と A05_orf244 の近傍の膜蛋白質は、他のグループの膜蛋白質と類似 (本文参照)
 - 4) 機能情報に関しては、Higgins [1992] and Saurin & Dassa [1994] を参照。
- a) gb:U09984 (文献なし)。
 - b) gb:S72674 [Sanders *et al.*, 1994]。
 - c) MJ1242 and MJ1662 は、methylcoenzyme M 還元酵素 (本文参照)。
 - d) MJ0719 はヒトの RNase L inhibitor に類似 (本文参照)。

本研究では、orthologous遺伝子は異なる生物種内で同じ機能を担う、という作業仮説に基づき、ATP結合蛋白質全体の分類にorthologous関係を考慮した点が最大の特徴である。従って本研究の分類結果は、必ずしも既存の分類方法による結果と同一ではない。例えば、今回解析に用いた7生物種全てのATP結合蛋白質を、前述と同様の距離の定義と群平均法を用いて分類した場合、おおよそ今回の分類と良い一致を見るが、本研究でのグラム陽性菌のグループ1-1（後述）に属するATP結合蛋白質は、この手法では、明らかに異なるグループに分類される（図4）。しかしこの場合、次節3.2.でも述べるように、周囲にコードされている膜蛋白質が保存されているという事実が、本分類の妥当性を支持する。即ち、少なくともこのグループに属するオペロンが、共通の祖先から進化した系統であろうことが示唆されている。

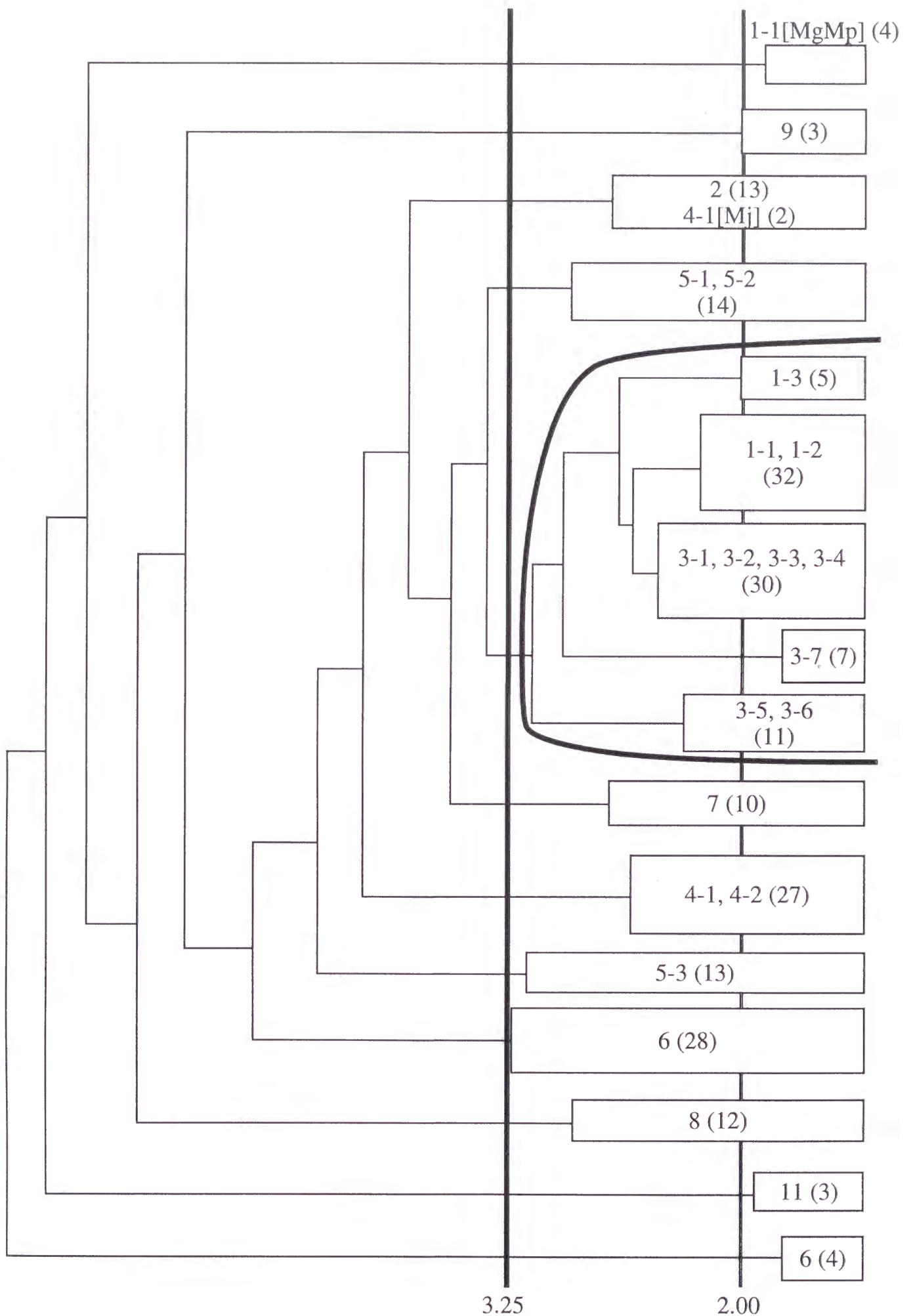


図4 全ATP結合蛋白質の群平均法によるクラスタリングの結果

3.2. グループ1

本グループに属するATP結合蛋白質を含むオペロンは、その周囲にコードされている蛋白質（主に膜蛋白質）の配列類似性により、さらに4つのサブグループに分類される。この分類は、ATP結合蛋白質のクラスタリングの結果とも、非常に良く一致する（図3）。モリブデン輸送に関与するATP結合蛋白質のうちb0760だけは、類似性からの判断ではなく、Ec内でのみ存在し、かつ、このグループのオペロンを構成している事実から、特別にこのグループに区分した。

グループ1-1に属するATP結合蛋白質は、今回解析に用いた7生物種全てが有しており、また多くの場合オペロン構造も保持されていた（図5）。しかし、各オペロンの基質特異性は、金属から糖に至るまで、非常に多様性を示している（表2）。その最も顕著な例は、Ecのb1318, b3450, b4035の3つのATP結合蛋白質と、それを含むオペロンである。これら3つのATP結合蛋白質は、互いに約50%の同一残基率を示し、アライメントスコアも1000を超える。また機能単位の構成要素である、膜蛋白質や基質結合蛋白質も統計的に有意な配列類似性を示す。にもかかわらず、これらは異なる基質の輸送に用いられている。（UgpC(b3450)はsn-glycerol-3-リン酸輸送 [Overduin *et al.*, 1988]、MalK(b4035)はmaltose/maltodextrin輸送 [Froshauer and Beckwith, 1984]、b1318は機能未知）

グループ1-1全体として見ると、オペロンにコードされている遺伝子の順序は、必ずしも保存されているとは言えない（図5）。特にparalogousオペロン間では、あまり遺伝子順序は保存されていない。しかし、複数のorthologousオペロン、例えばモリブデン輸送系 [Walkenhorst *et al.*, 1995] や、スペルミジン/プトレシン輸送系 [Furuichi *et al.*, 1991]、硫酸塩/チオ硫酸塩輸送系 [Hryniewicz *et al.*, 1989; Sirko *et al.*, 1990] 等、にコードされている遺伝子の順序は、種を越えて保存されている。また、その機能は未知であるものの、b0066(Ec)とHI1021(Hi)を含むorthologousオペロンのペア、HP0475(Hp)とsll0739(Ss)を含むorthologousオペロンのペアも、各々同じ遺伝子の順序をもつ。これらのオペロンにコードされているほとんどの膜蛋白質や基質結合蛋白質遺伝子についても各々orthologous関係にある。

このグループのオペロンの半数以上が、5'端から順に、基質結合蛋白質、膜蛋白質（多くの場合2蛋白質、あるいはドメイン）、そしてATP結合蛋白質という順序を保存している。paralogousオペロンでも、順序は保存されていないものの、一つの機能単位を構成するであろう遺伝子群は集中してコードされており、一つの転写単位となっている可能性が示唆される。

このように機能単位を構成するであろう遺伝子群は、ゲノム上で高度に集積していること、および、グラム陽性菌の基質結合蛋白質は、リポ蛋白質であることから、配列

の類似性こそないものの、MG186(Mg)とE07_orf301(Mp)の2リポ蛋白質は、隣接するオペロンの基質結合蛋白質である可能性を示唆したい(図5)。

グループ1-2のオペロンの特徴は、膜蛋白質を一つしか有しないことである(図6)。類似性のある基質結合蛋白質も、3生物種にわたり、6オペロンで保存されている。しかし、オペロン内の遺伝子順序は必ずしも保存されていない。また基質特異性も様々である(表2)。

Ec中のグループ1-2の次に述べる2つのオペロンは、硫酸塩飢餓反応に関与することが知られている。タウリン活用システムである、tauABCDオペロン[van der Ploeg *et al.*, 1996]のTauA (b0365) および、TauD (b0368) は、SSI (sulfate starvation induced proteins) 1 および 3 である[Quadroni *et al.*, 1996]。また、b0937とb0935は、各々SSI4および6である。この2蛋白質を含むオペロンの構成(図6)は、枯草菌(*B. subtilis*)のオペロン(GenBank accession no. L16808)[Quirk *et al.*, 1994]と類似している。このことは、Ecと同様の飢餓反応系が枯草菌にも存在することを示すのかもしれない。

グループ1-3のオペロンは、グラム陰性菌で非常に良く保存されている。EcとHiの間では7遺伝子が、順序も含めて保存されている(b3189からb3195と、HI1081からHI1087)。アミノ酸配列の保存度も、ほとんどのペアが同一残基率40%を超える。Hpにおいても7遺伝子のうち3遺伝子(HP1464(b3193に相当), HP1465(ATP結合蛋白質), HP1466(b3194に相当))が、保存されており、隣接してコードされている。Ssでは、この3遺伝子に相当する遺伝子を2組分持っているものの、ゲノム上で隣接してコードされているのは、sll1001(ATP結合蛋白質)とsll1002(b3193に相当)の2つだけである。7遺伝子のうち残りの、b3191とb3192は、7生物種のうちEcとHiのみに存在するようである。また、murA(b3189)は、グラム陰性菌3種とSsに、b3190はEc, Hi, Ssの3種に存在するようであるが、EcやHiの様に集積してコードされていない。グループ1-3のオペロン構成遺伝子は、他のABC輸送系オペロンの膜蛋白質や基質結合蛋白質とは、配列類似性を示さない。

グループ1-4のATP結合蛋白質、b0490の周囲には、今回比較を行なった7生物種のグループ1のオペロン構成遺伝子は保存されていない。しかし、Ec中のb0490とb0491の遺伝子の並びが、枯草菌[Kunst *et al.*, 1997]でも保存されている(yjkBA)。

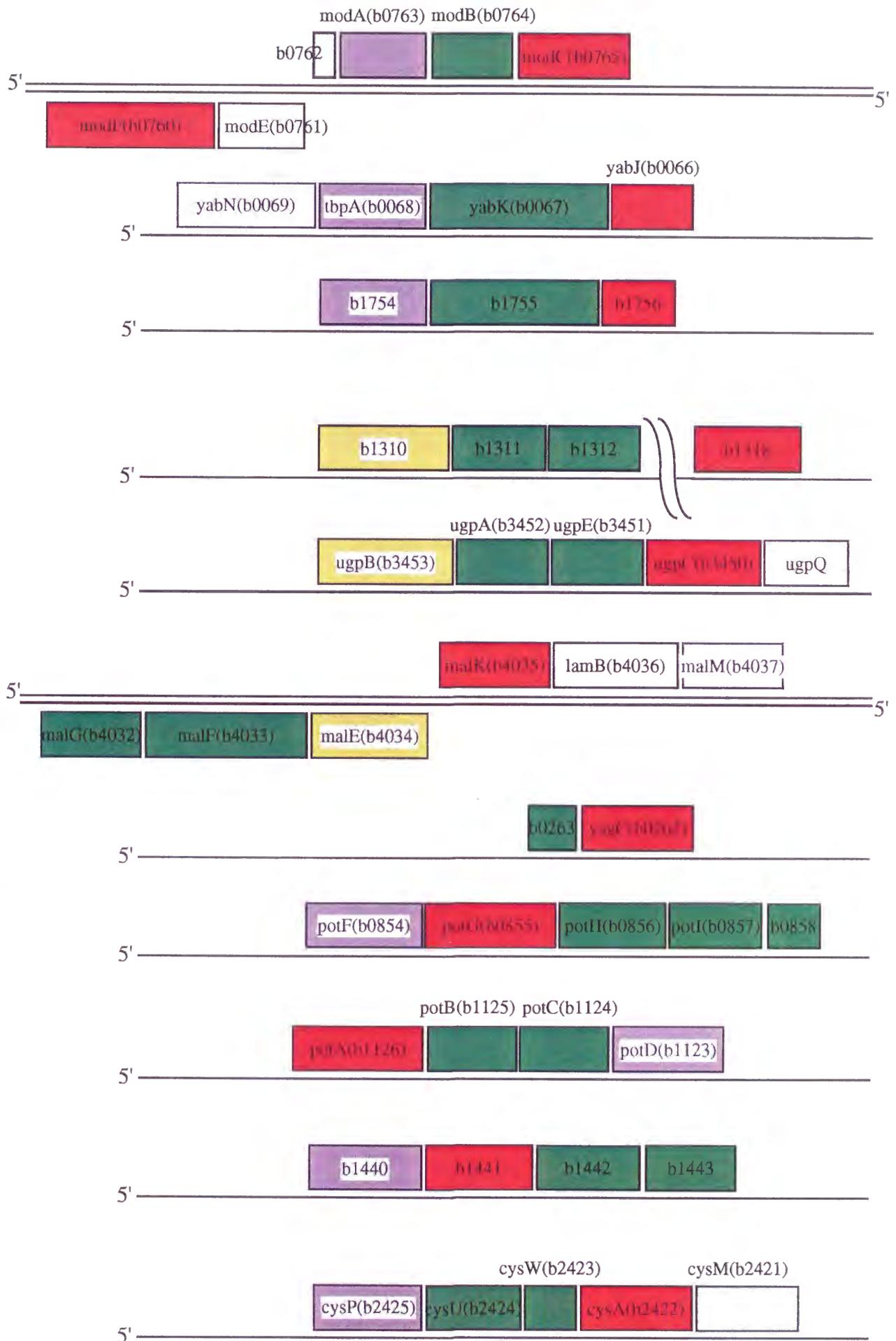


图 5 a)

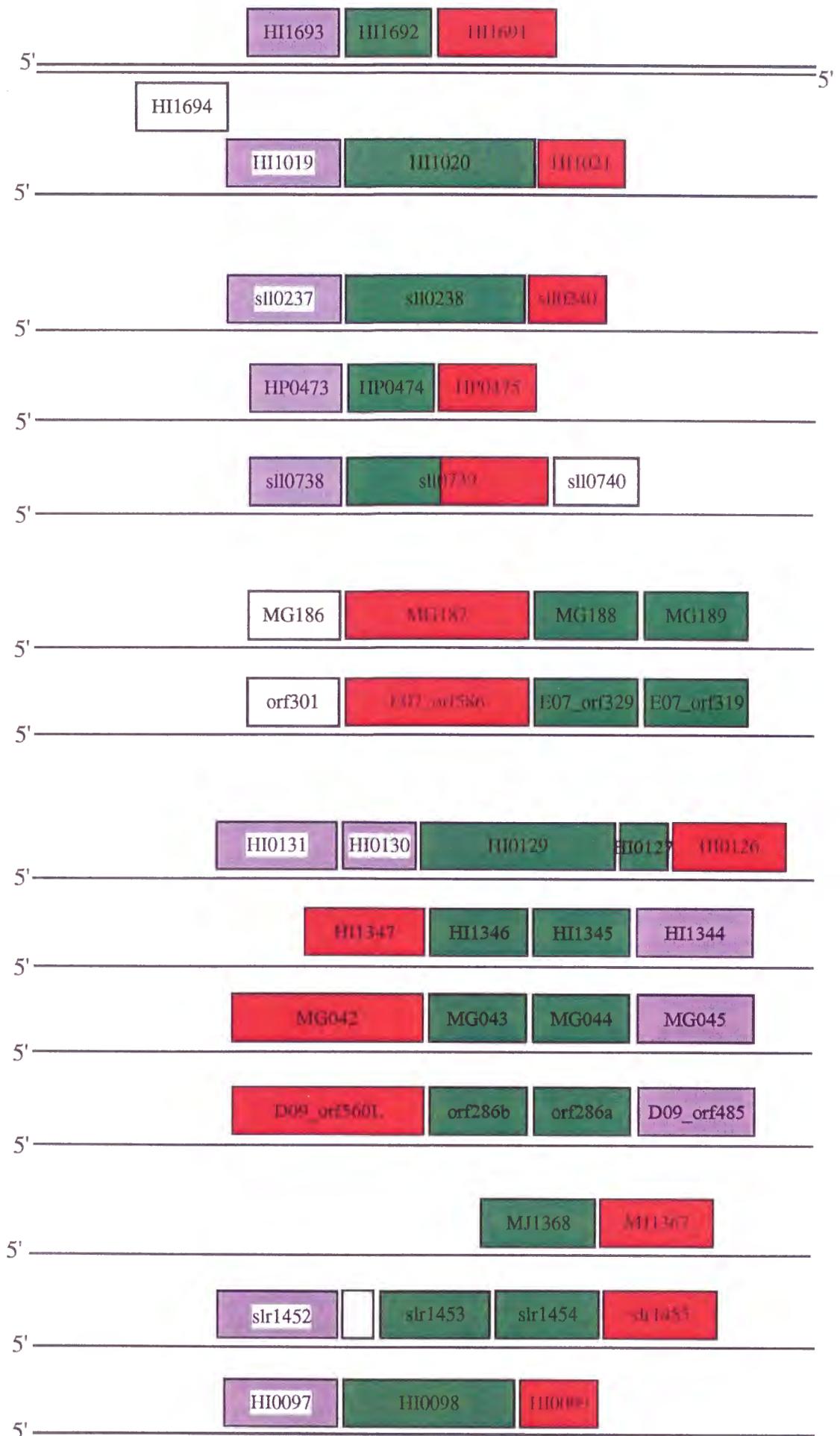


図5 (説明) グループ1-1のa) Ecでのb)それ以外の6生物種でのオペロン構造を、それぞれ示す。赤色はATP結合遺伝子、緑色は膜蛋白質を表す。(これはこの後の図も同様) 灰色をかけた部分が基質結合蛋白質にあたる。配列類似性により紫と黄に色分けした。

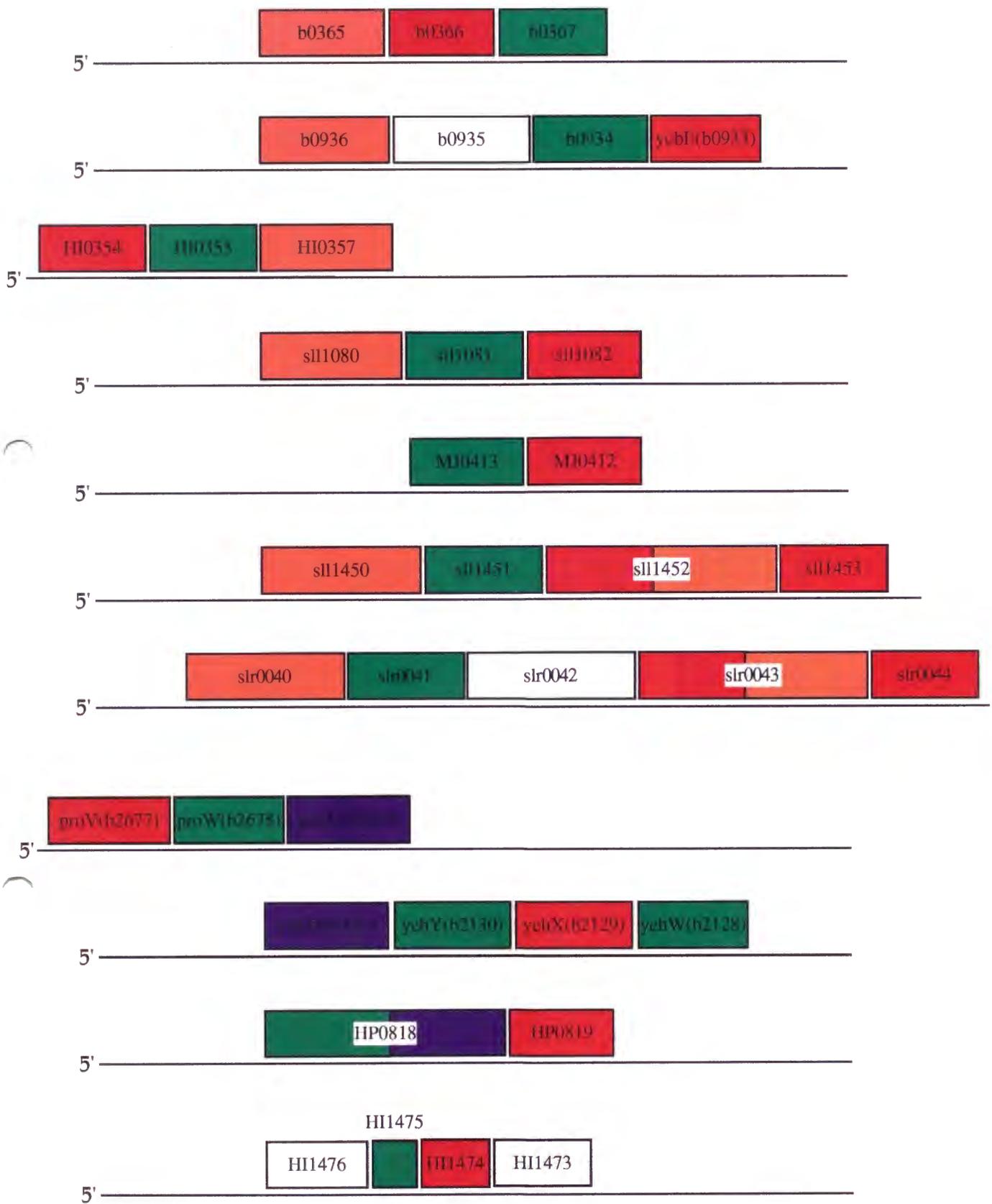


図 6

グループ 1-2 のオペロンの構成
 黄色と青色で示した遺伝子は、基質結合蛋白質を示す。
 白色の遺伝子は実験事実等から、(ATP結合蛋白質、
 膜蛋白質、基質結合蛋白質以外の) オペロン構成要素
 であると考えられるもの。

3.3. グループ2

グループ2のオペロンのうち、Ecの4つのオペロンは既に実験的に同定されたものであり、これらはいずれも糖輸送に関与している。このグループに属するATP結合蛋白質をコードする遺伝子の周囲の保存性から、Ecはこれら以外にもさらに4オペロンをもつと考えられる。Hiにも、このグループに属する3つのオペロンが存在する。MgとMpには、各々1つのこのグループのオペロンが存在する。Ecのb4229は非常に短い(90残基)蛋白質で、今回の分類の閾値を満たしていないが、配列類似性の面からみてこのグループに区分した。

オペロン内の遺伝子配置で最も多いのは、5'端から順に、基質結合蛋白質、ATP結合蛋白質、膜蛋白質であるが、例外も存在する(例えば、rbsオペロン(b3749)やb1513)。このグループに属するATP結合蛋白質は、2つのATP結合ドメインを持つ。既知のオペロンとの構成要素の類似から考えて、このグループの機能未知のオペロンも、糖輸送に関与していると推定される。

EcとHiの間で、orthologousオペロンは非常に良く保存されている。galactose輸送系(b2149, HI0823)では4遺伝子が、xylose輸送系(b3567, HI1110)では、2遺伝子(うち一つはsuspected coding region)の挿入(Hi)を挟んで6遺伝子が、ribose輸送系(b3749, HI0502)では6遺伝子が、それぞれ順序も含めて保存されている。従って、これらのオペロンの起源は、EcとHiの分岐以前にあることが示唆される。こうした良く保存されている遺伝子の中に、オペロンの発現を調節するリプレッサー遺伝子が含まれている。このタイプのリプレッサーは、EcとHiのみに存在した。リプレッサー蛋白質の認識するDNA配列も、EcとHiの間で保存されている。

3.4. グループ3

本グループに属するATP結合蛋白質を含むオペロンは、その隣にコードされている膜蛋白質の配列類似性により、7つのサブグループに分類される。

グループ3-1のATP結合蛋白質遺伝子に、ゲノム上で隣接している遺伝子の多くには、図7に示す保存領域を持つ蛋白質がコードされている。保存領域のアミノ末端の部位は、最初の膜貫通領域に相当すると考えられる。この他にも、2番目から3番目の膜貫通領域を繋ぐループ部分と、カルボニル末端のアミノ酸配列が保存されている。特にループ領域のG-x[9]-F-x[10]-Gというモチーフは、非常に良く保存されている。これは、新規のアミノ酸配列モチーフである。

さらに、機能が既にある程度知られているオペロンとの類似性を指摘しておく。Ec中で、ftsYEX(b3464, b3463, b3462)というオペロンが細胞分裂に関与することが知

られている[Gill *et al.*, 1986]。このうちftsEは、ABC輸送系のATP結合蛋白質と配列類似性を有している。そして前記の保存領域は、ftsXも有す(図7)。このことから、このグループのオペロンの機能が、類似している可能性が示唆される。また、こうした保存領域が、ftsXも含めた膜蛋白質、ひいてはオペロン全体の機能に重要な関りがある部位であると考えられる。

グループ3-2のオペロンでは、EcとHiの間で4つの蛋白質が、両ストランドにまたがって保存されている。それらは、配列類似性から推定されるヒスチジノールホスファターゼと、ATP結合蛋白質、膜蛋白質、リポ蛋白質の4つである。EcとHiは、共にグラム陰性菌であるにもかかわらず、このグループのオペロンの基質結合蛋白質は、リポ蛋白質であるのかもしれない。Hpでは、ATP結合蛋白質と膜蛋白質はゲノム上で隣接している。しかし、これら以外の2つの蛋白質は、EcとHiの様に集積してコードされていない。グループ3-2のオペロンのもう一つの特徴は、EcとHi両生物種で、このオペロンがrrnオペロンの隣にコードされていることである。これは、EcとHiの分岐以前の名残であろう。また、グループ3-3のうちの一つのオペロンとの関連も示唆される(後述)。このグループの機能は未知である。

グループ3-3のオペロンのうち、Ecの4オペロンの構成と機能が、実験的に確かめられている。これらは、いずれも極性あるいは荷電アミノ酸の輸送を行うことが知られている。その他の機能未知のオペロンのうちEcのb1917については、考察の項で述べるように、べん毛の形成に関与していることが予想される。

グループ3-3のオペロンでは、基質結合蛋白質、膜蛋白質、ATP結合蛋白質の3種類の蛋白質が保存されている。グラム陰性菌3生物種中に、全部で9つのオペロンが存在する。そのうち、6オペロンで、膜蛋白質の重複が見られる。Ssでは、このタイプのオペロンは形成されていない。

グラム陰性菌3生物種中で保存されている、このグループの膜蛋白質のアミノ酸残基の長さを比較すると、Ecのb3269とb3270(ともに368残基)だけが、長いことがわかる(他はいずれも210から250残基)。また、この2つの膜蛋白質は、グループ3-2のオペロンに保存されている膜蛋白質と、配列類似性を示す。さらにこの2つの蛋白質とATP結合蛋白質(b3271)がコードされている領域の隣には、グループ3-2のEcとHiのオペロンと同様、rrnオペロンがコードされている。このことは、Ec内部での2つのオペロンの遺伝子重複(基質結合蛋白質(b3268)をコードする遺伝子を加えた、4遺伝子(b3268-71)からなるオペロンとそれに隣接するrrnDオペロンと、グループ3-2で述べたオペロン(b0197-b0200)とrrnHオペロン)の跡を示すものかもしれない。

グループ3-4に属するPhnC(b4106)は、Ecに特有のオペロンのATP結合蛋白質である。このオペロンは、alkylphosphonateの輸送に関与することが明らかになっており、構成要素も、典型的なABC輸送系オペロンよりも多い[Makino *et al.*, 1991]。この

オペロンには、後述するグループ4-2に属するATP結合蛋白質も含まれている。

グループ3-5に属するオペロンは、Ecでは、分枝した側鎖を持つアミノ酸の輸送に関与している。EcとMjで、基質結合蛋白質、膜蛋白質、ATP結合蛋白質の3種類の蛋白質がオペロンを構成している。Ssでは、こうしたオペロンの形成は見られない。このグループの膜蛋白質は、糖輸送に関与するグループ2の膜蛋白質と配列類似性を示す。

グループ3-6に属するATP結合蛋白質をコードする遺伝子の隣には、EcとHpでrpoN (RNAポリメラーゼシグマ(54)因子) がコードされている。このシグマ因子と配列類似性の基準を満たす蛋白質は、今回解析に用いた7生物種の内、EcとHpの2生物種以外には存在しなかった。また、EcとHiでは5遺伝子が保存されている(図8)。この5遺伝子の中には、ABC輸送系以外の輸送システムの蛋白質であるptsN(b3204) [Powell *et al.*, 1995]が存在する。Ssでは、これら5遺伝子の多くは存在しない。

グループ3-7のオペロンは、7生物種中Hpを除く6生物種で、3種類の蛋白質が保存されている。遺伝子の順序も、重複等多少の変化はあるものの、5'端から基質結合蛋白質、膜蛋白質、ATP結合蛋白質の順で、生物種を越えて保存されている。このグループのオペロンは、Ecでリン酸輸送に関与することが知られている[Surin *et al.*, 1985]。Ssは、このグループのオペロンを二組持つ。

Ecでリン酸輸送に関与する蛋白質と共にコードされている蛋白質PhoU(b3724)は、輸送には直接関係しないが、膜周辺に存在し、細胞内のリン酸代謝を調節する[Surin *et al.*, 1985]。この蛋白質はEc以外では、Mg (MG409)、Mp (C12_orf225)、Mj (MJ1009)の3生物種で、やはり輸送系蛋白質と共にコードされている。Mjには、これ以外にももう一つ別の個所に、PhoUと配列類似性を示す蛋白質(MJ1641)が存在する。Ssにも、PhoUと配列類似性を示す蛋白質(slr0741)が存在するが、輸送系蛋白質の近傍にコードされている訳ではない。

このグループのオペロンを持たないHpは、リン酸の摂取をどのように行っているのだろうか？ Hpには、別のタイプのリン酸輸送蛋白質(HP1491)が存在する。この蛋白質と配列類似性を示す蛋白質は、Ecが2つ(b2987, b3493)、Hiが1つ(HI1604)、そしてMjも1つ(MJ0630)持っている。

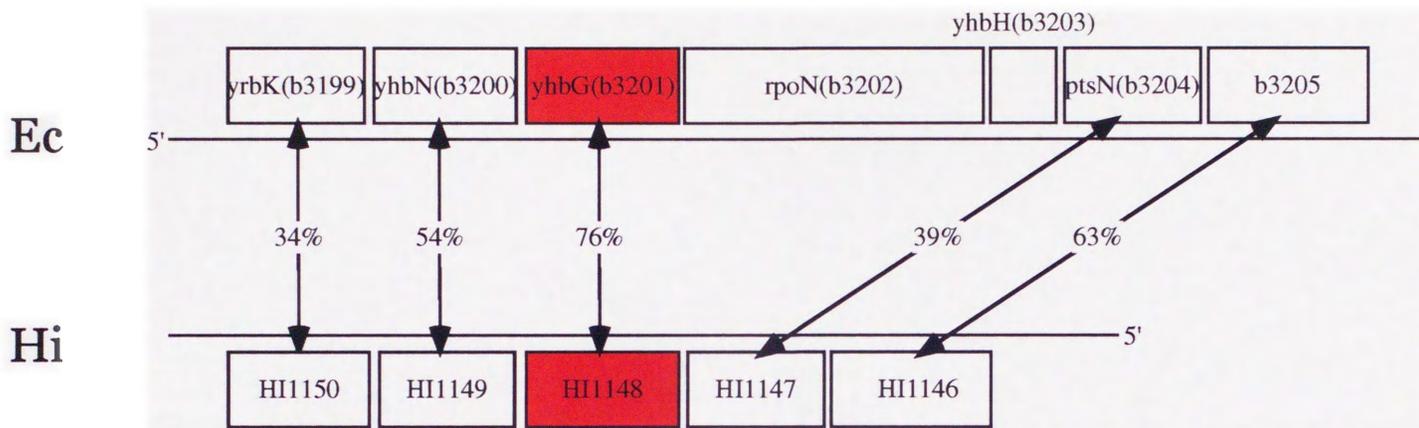


図 8 EcとHiで保存されている5遺伝子のゲノム上での配置
ATP結合蛋白質を赤で示す。矢印は対応する遺伝子を、矢印に添えた数字は、各遺伝子のコードするアミノ酸配列の同一残基率を示す。

3.5. グループ4

本グループに属するATP結合蛋白質を含むオペロンは、その構成により、2つのサブグループに分類される。一つ（グループ4-1）は、グループ内のオペロン構成が、Mjの例外を除き、生物種を越えて良く保存されている。もう一つ（グループ4-2）は、前述したグループ3-4に属するalkylphosphonateの輸送に関与するオペロンであり、Ecのみに存在するものである。

グループ4-1のオペロンは、ATP結合蛋白質を2つ、あるいは2つのATP結合ドメインが融合した蛋白質を持つ。このグループのオペロン内の典型的な遺伝子配置は、5'端から順に、基質結合蛋白質、2つの膜蛋白質、そしてATP結合蛋白質である。このグループのオペロンは、これまで明らかになってきているペプチド輸送のオペロンに類似している。Ecのb3479, b3480のみは、ニッケルの摂取に関与している[Navarro *et al.*, 1993]。このグループに属するMjの2つのATP結合蛋白質(MJ1242, MJ1662)は、オペロンを構成していないことおよび、methylcoenzyme M還元酵素[Kuhner *et al.*, 1993]との高い配列類似性から、輸送蛋白質としては機能していないと考えられる。

Ecのみに存在する3つのオペロン(b0829, b1483-84, b2180)には、上に述べた3種類の蛋白質以外にも保存されている蛋白質が2種類存在する。一種類は、b0833, b1489, b2176の3蛋白質であり、もう一種類は、b0834, b1490の2蛋白質である。これら5つの蛋白質は、共通したドメインを有し、それは*Caulobacter crescentus*（グラム陰性菌 α subdivision）のもつGGDEFドメインに類似している。このドメインは、細菌のシグナル伝達への関与が示唆されている[Hecht and Newton, 1995]。

Ssのゲノム中には、このグループのオペロンにコードされている基質結合蛋白質と配列類似性を示すものが2つ(sll1699, slr1740)、膜蛋白質と示すものが4つ(sll0312, sll1768, slr0324, sll0833)、ATP結合蛋白質が1つ(sll1927)、各々存在している。しかし、それらは別々にコードされており、上に述べたようなオペロンは形成されていない。MgとMpには、このグループの基質結合蛋白質と配列類似性を示す蛋白質は存在しない。

グループ4-2のATP結合蛋白質が含まれてオペロンは、14遺伝子からなる[Makino *et al.*, 1991]。このオペロンの構成遺伝子のうち、phnGHIJMはEcに特有で、他の6生物種中には、配列類似性を示す蛋白質は見出せない。

3.6. グループ5

本グループに属するATP結合蛋白質を含むオペロンは、膜蛋白質の配列類似性により、3つのサブグループに分類される。

グループ5-1のオペロンは、今回比較した7生物種全てにわたって、少なくともATP結合蛋白質と膜蛋白質を保存している。しかし、オペロン内での遺伝子の順序は、一様ではない。このグループのオペロンは、Ecで、主に鉄の複合体の輸送を担っていることが知られている。Ecではこれらのオペロンは、これまで述べてきたように、基質結合蛋白質、膜蛋白質とATP結合蛋白質だけではなく、外膜に存在するレセプター蛋白質等多くの他の要素を含む複雑な構成になっている [Ozenberger *et al.*, 1987; Shea and McIntosh, 1991; Hunt *et al.*, 1994]。Ecのb1709は、ビタミンB₁₂を輸送するオペロンbtuCEDを構成するATP結合蛋白質である。Ssには、2つの膜蛋白質 (slr1316, slr1317)、ATP結合蛋白質 (slr1318)、そして基質結合蛋白質 (slr1319) からなると推定されるオペロンが存在する。これに隣接して、Ecの外膜レセプターや、基質結合蛋白質、ATP結合蛋白質 (slr1488, slr1494 (いずれもグループ6)) 等を複数個ずつ含む、非常に長い領域がある。この領域も、一つのオペロンを形成するのかもしれない。

今回の分類で、このグループに分類されたMgおよびMpのオペロンは、*Mycoplasma hyorhinis*で決定されたオペロン [Dudler *et al.*, 1988] に非常に良く類似している。これらのオペロン中の膜蛋白質は、このグループの膜蛋白質と統計的に有意な類似を示さない。しかし、Mpの膜蛋白質 (A05_orf542) は、グループ1-1の膜蛋白質HP0474 (Hp) と、200アミノ酸残基以上にわたって約25%の同一残基率を示す。

グループ5-2のオペロンでは、ATP結合蛋白質をコードする遺伝子の周囲に、基準は満たさないものの、グループ5-1の膜蛋白質と弱い配列類似性を示す膜蛋白質、およびadhesion蛋白質が保存されている。このオペロン構成は、グラム陽性菌である *Streptococcus parasanguis* の fimA locus (GenBank accession no. M26130) に非常に良く似ている。ただしEcでは、adhesion蛋白質 (b1857) は逆ストランドにコードされている。Hiでは、このオペロンが2つ存在し、各々の隣に細胞壁構築に関与する遺伝子が配置されている。Ssでもこのオペロンが2つ存在し、そのうち1つはマンガン輸送系であることが知られている [Bartsevich and Pakrasi, 1995]。これは、マンガンとadhesion (接着) との関連が知られている [Bauer *et al.*, 1993] ので、妥当な帰結であろう。

グループ5-3のオペロンでは、*Salmonella typhimurium* のコバルト輸送遺伝子に類似した遺伝子が保存されている。*Salmonella* のコバラミン合成オペロンには、5'端から順にcbiMNQOというコバルト輸送遺伝子が存在する [Roth *et al.*, 1993]。このうち、cbiQは膜蛋白質をコードしており、cbiOにATP結合蛋白質がコードされている。残りの2つの遺伝子の機能等は不明である。CbiQに相当する蛋白質は5生物種、8オペロンで保存されている。CbiNと有意な配列類似性を示す蛋白質はMJ1090一つであるが、これがコードされているのは別の2つのオペロンで、CbiMに相当する蛋白質

(HI1621, MJ1569)とCbiQに相当する蛋白質(HI1619, MJ1571)との間に膜蛋白質がコードされている。MgとMpでは、このグループのオペロンが重複しているが、これらの非常に小さな遺伝子セットしか持たない生物種が全く同じ機能を有する輸送系を保存しているとは考えずらいので、各々微妙に異なる役割を担っているのかもしれない。

このように配列類似性から、コバルト輸送系であると推定されるオペロンが、5生物種に保存されている訳であるが、他の2生物種(EcとHp)に存在しない理由を考慮する。まず、コバルトを細胞内に摂取する理由として、コバラミンの合成が考えられる[Roth *et al.*, 1993]。コバラミン(ビタミンB₁₂)は、コバルトを含む化合物で、ある種の細菌により産出される。また、細菌の発育因子でもある。この化合物を合成できないEc[Raux *et al.*, 1996]が、このグループのオペロンを持たず、ビタミンB₁₂を輸送するオペロンを有しているのは妥当である。しかしこの理由だけでは、不都合な場合も出てくる。なぜなら、このグループのオペロンを持つ5生物種のうち、Hi, Mg, Mpの3生物種は、現在知られているコバラミン生合成遺伝子と類似性を示す遺伝子が乏しいことから、コバラミンを合成能を欠く可能性が大きいためである。これらの持つオペロンの基質はコバルト以外の物質なのであろうか？ただこの点に関しては、生物種により、コバラミン合成経路は大きく異なる場合もあるという指摘[Raux *et al.*, 1996]もあることを付け加えておく。もし他のなんらかの理由で、コバルトの摂取を必要とする場合、EcとHpはどうしているのか？答えは、他の輸送系を使用することである。実際、ABC輸送系とは異なるファミリーに属する輸送蛋白質CorA(と配列類似性を示す蛋白質)が、MgとMpを除く5生物種に存在する。CorAは、コバルトとマンガンの輸送を司る[Smith *et al.*, 1993]。

3.7. グループ6

このグループに属するATP結合蛋白質は、Ecの細菌毒素(溶血素)分泌蛋白質であるhemolysin Bや、哺乳類の薬剤耐性蛋白質と配列類似性を示す。また、このグループに属する蛋白質は、ATP結合ドメインのアミノ端に、膜貫通ドメインが融合した構造を持っている。このため各蛋白質の長さは、550アミノ酸残基を越える。このグループの膜貫通ドメインは、他のグループのオペロンの構成要素である膜蛋白質とは、配列類似性を示さない。Mjには、このグループのATP結合蛋白質は存在しない。

このグループのATP結合蛋白質が2つペアでコードされている場合が、4生物種(Ec, Hi, Mg, Mp)にわたり、6つみられる。これらは、おそらくオペロンを形成し、ヘテロダイマーとして機能単位を構成するのであろう。これらの構造に加えて、b0886, b0887(Ec)とHI1156, HI1157(Hi)のオペロンでは、チオレドキシン還元酵素trxB(b0888およびHI1158)が保存されている。また、ともにEcのb0886, b0887と

b0448, b0449の間では、転写因子(b0889およびb0447)が保存されている。

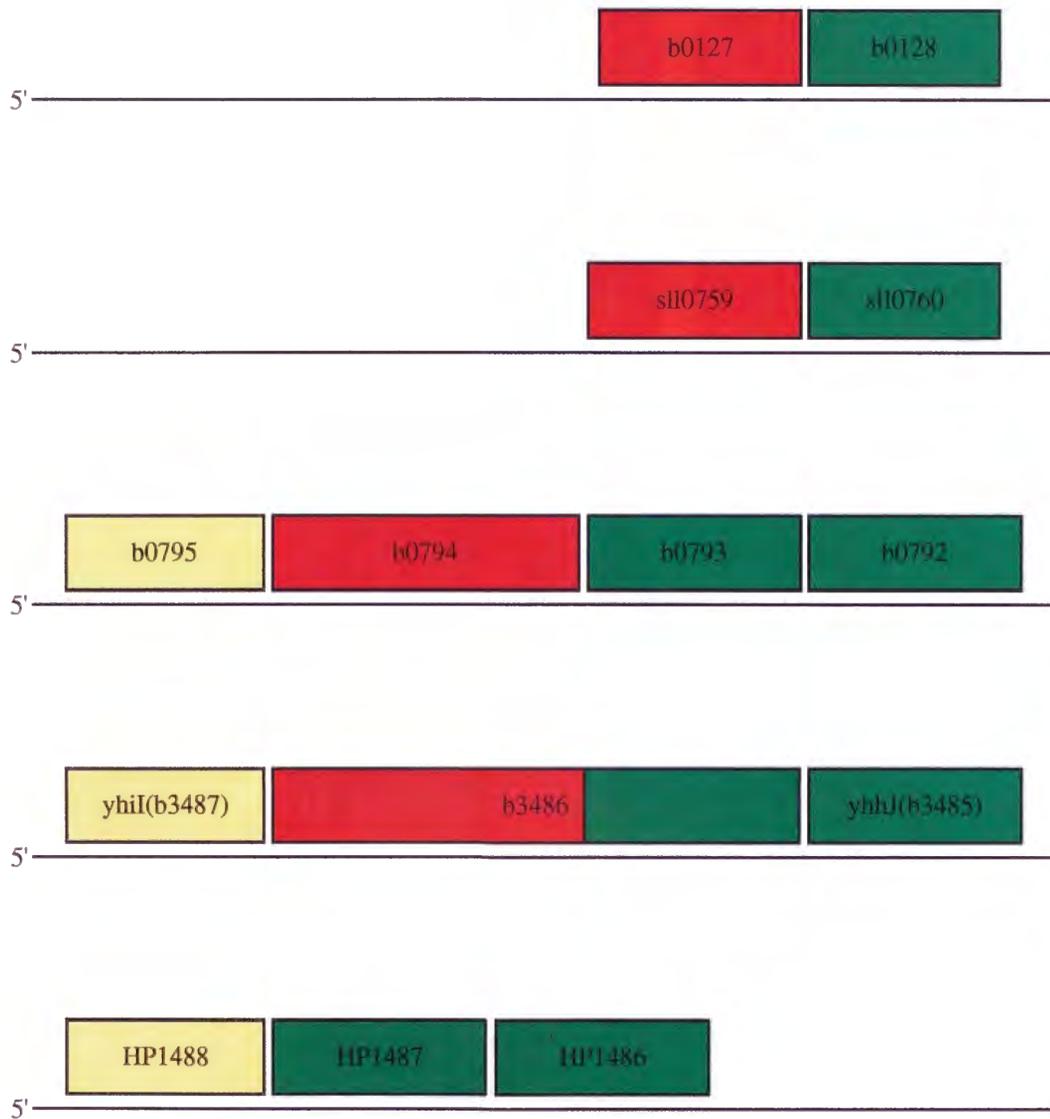
EcのmsbA(b0914)とその両隣の遺伝子(b0913, b0915)は、Hiでも保存されている。これら3遺伝子のうち2遺伝子(b0914, b0915)は、細菌(Ec)に必須であることが知られている[Karow and Georgopoulos, 1993]。このグループのうち、4蛋白質(b1496(Ec), HI0036, HI1467(Hi), sll0182(Ss))がEcの抗菌ポリペプチドの摂取に関与すると考えられているSbmA(b0377)蛋白質[Lavina *et al.*, 1986]に類似したドメインを有している。

3.8. グループ7

このグループのオペロンに保存されている膜蛋白質は、ABC-2型[Raizer *et al.*, 1992]の膜蛋白質である。このグループに属するATP結合蛋白質をコードしている遺伝子の近傍で、保存されている遺伝子を図9に示す。ATP結合蛋白質と膜蛋白質以外にも保存されている要素がある(図中の黄色で示した遺伝子)。これらの遺伝子がコードしている蛋白質は、細菌毒素(溶血素)分泌蛋白質であるHlyDと配列類似性を持つ。Hpでは、ATP結合蛋白質(HP1220)とは離れたところに、膜蛋白質とHlyD類似蛋白質がまとまってコードされている(図9)。Mjでは、ATP結合蛋白質(MJ1023)の近傍には、今回比較した生物種の持つ膜蛋白質やHlyD類似蛋白質と配列類似性を持つ蛋白質は、コードされていない。しかし、MJ1023の隣にあるMJ1024がコードしている蛋白質は、枯草菌のNatB[Cheng *et al.*, 1997](これはEcの膜蛋白質YadH(b0128)と配列類似性を持つ)と配列類似性を有することを、MJ1024の遺伝子産物がこのグループの膜蛋白質である間接的な証拠として挙げておく。

3.9. グループ8

このグループに属するATP結合蛋白質をコードしている遺伝子の周辺の遺伝子には、生物種内でも、生物種を越えても、顕著な保存性はみられない。これは、今まで見てきたグループとは、明らかに傾向を異にするものである。このグループのATP結合蛋白質は、HI1342(Hi)を除いて、全て530アミノ酸残基長を越え、2つのATP結合ドメインを有する。また、このグループに分類したMJ0719(Mj)は、ヒトのリボヌクレアーゼLの阻害蛋白質[Bisbal *et al.*, 1995]と約50%もの同一残基率を示す。従ってMJ0719は、輸送蛋白質というよりも、このヒトの蛋白質と同じであることが予想される。これらの事実を総合すると、少なくともこのグループに属するATP結合蛋白質は、輸送蛋白質ではない可能性が考えられる。ただし、真正細菌の持つこのグループのATP結合蛋白質は、MJ0719程ヒトのリボヌクレアーゼL阻害蛋白質と高い配列類似性を示さ



☒ 9

グループ7 (ABC-2型) のオペロン構造。
黄色は、HlyD類似遺伝子を示す。

ない（同一残基率は約25%）。

3.10. その他のグループ

グループ9から13までに属するATP結合蛋白質が、自分の属するグループ以外のATP結合蛋白質と300点以上のアライメントスコアを示すこと、また、同一残基率が30%を超えることはほとんどない。まれに300点を越える場合があるが、その中でも例外的に高い得点を示すのは、グループ9のMJ0035(Mj)と、グループ3-6のb3201(Ec)の395点（同一残基率約34%）、一例のみである。

グループ9のオペロンに保存されている遺伝子には、他のグループにない特徴的な遺伝子が2種類含まれている。そのうちの1つである窒素固定遺伝子nifSは、Ec(b1680)とSs(slr0077)で保存されている。もう1つは、Ec(b1681, 1683), Ss(slr0074, slr0076), Mj(MJ0034)で保存されている。この遺伝子がコードする蛋白質は、グラム陽性菌である*Mycobacterium leprae*のもつintein, XheA[Smith et al., 1997a]に類似している。また、このXheAをコードする遺伝子付近の構成もグループ9のオペロンに酷似している。*M. leprae*ゲノム中でxheA近傍にあるATP結合蛋白質とグループ9のそれとの同一残基率はそれぞれ、47%(b1682), 49%(slr0075), 35%(MJ0035)である。このオペロンは真正細菌、古細菌にまたがり、広く分布している。

グループ10から13のATP結合蛋白質は、比較的限られた生物種にのみ存在している。グループ10はチトクロムcの成熟に関与するオペロン[Thony-Meyer et al., 1995]で、その順序も含めて全部で8遺伝子が、EcとHiの間で保存されている。この8遺伝子は、Ecでは、より大きなオペロンの一部である[Thony-Meyer et al., 1995]。グループ11のオペロンは、それに含まれている膜蛋白質の配列類似性から、ABC-2型の輸送系オペロンの仲間であると考えられる。このグループのオペロンは、Ssのみに存在する。グループ12のATP結合蛋白質(HP0613)は、今回比較した7生物種のもつ、自分自身以外のどのATP結合蛋白質とも230点を越えるアライメントスコアを示さない。この蛋白質の機能やオペロン構造の有無については、不明である。グループ13のATP結合蛋白質(MJ0121)をコードする遺伝子の近傍の構造は、今回比較した生物種の中では保存されている訳ではない。しかし、この遺伝子とその下流2遺伝子の並びが、同じく古細菌である*Methanobacterium thermoautotrophicum*[Smith et al., 1997b]でも保存されている。

4. 考察

4.1. オペロンの機能推定

ここでは、本研究の分類を通して得られた各グループの機能に関する新しい知見（表2の該当箇所にアスタリスクを付記した）について主なものを纏める。さらに、グループ3-3のEcの機能未知のオペロンとべん毛形成との関りについて述べる。

本研究で得られたABC輸送系オペロンの機能に関する知見は、大きく分けて次の8点である。

- ・グループ3-1のオペロンの構成要素に共通する配列上の保存領域および、モチーフの発見。またその保存領域および、モチーフが、Ecの細胞分裂に関与するオペロンの構成要素にも存在することを発見。このグループのオペロンの細胞分裂への関与が示唆される。

- ・Ecのグループ4-1のオペロンの構成要素に、シグナル伝達への関与が示唆されているGGDEFドメインが保存されていることを発見。

- ・グループ5-2のオペロンのマンガン輸送系に対する類似を発見。またこれらのオペロンの、細胞壁（再）構築／接着への関与に対する示唆。

- ・グループ5-3のオペロンの、コバルト輸送系に対する類似を示唆。

- ・EcとHiのグループ6のオペロンに、新規の構成要素を発見。

- ・グループ8のATP結合蛋白質が、輸送以外の機能を担うであろうことを示唆。

- ・グループ9のオペロンが、非常に独特な構成要素(intein)を持つこと、また、この種のオペロンが、古細菌と真正細菌にわたって、さらに真正細菌の間でも広く存在することを明らかにした。

- ・グループ11のオペロンが、ABC-2型のオペロンであることを明らかにした。（ただし、このグループについては、同様の発見が同時期にReizerによりなされたことを付記しておく。PROSITE [Bairoch and Bucher, 1994] (PDOC00692)参照。）

Ecのべん毛形成とABC輸送系オペロンとの関り

Ecのべん毛遺伝子の発現には、3つの制御構造の段階がある。このうち3番目の段階に当たる、class IIIのべん毛遺伝子の転写に関与するシグマ因子*fliA* (b1922)を含むオペロン*fliAZY*の存在が報告されている[Mytelka and Chamberlin, 1996]。このうち、*fliY* (b1920)には、ABC輸送系の基質結合蛋白質に類似した蛋白質がコードされている。

fliY (b1920)の遺伝子産物は、グループ3-3オペロンの基質結合蛋白質と明らかな配列類似性を示す。さらに、*fliY*近傍の遺伝子の並びが、グラム陰性菌3種(Ec, Hi, Hp)で保存されている。つまり、Ecでは、*fliY*、その下流1つにおいて膜蛋白質(b1918)、そしてATP結合蛋白質(b1917)、Hiでも5'端から順に、基質結合蛋白質(HI1080)、膜蛋白質(HI1079)、そしてATP結合蛋白質(HI1078)、Hpでは、5'端から順に、基質結合蛋白質(HP0940)、膜蛋白質(HP0939)、の並びが保存されている。そして、これらの蛋白質は、該当するもの同士、互いにorthologous関係にある。

また、これまで3章で見てきたように、あるいは次の4.2.節でも述べるように、ABC輸送系の遺伝子群は、各生物種のゲノム上で強固に集積し、オペロンを形成し、その構成要素が一つの機能単位となる場合が非常に多く見られる。

以上の事実から、この*fliY* (b1920)の遺伝子産物が、その近傍にコードされている膜蛋白質(b1918)、およびATP結合蛋白質(b1917)と機能単位を形成し、ABCタイプの輸送に関与するであろうことが示唆される。また、*fliAZY*のオペロンも実は、5'端の*fliA* (b1922)から、3'端のATP結合蛋白質(b1917)までで、構成されるのであろう。

4.2.7 生物種におけるオペロンの保存性

表3に纏めたように、今回比較を行った7生物種のうち、Ssを除く6生物種のゲノム上では、ABC輸送系の遺伝子群は非常に強固に集積し、オペロンを形成している。これは、真正細菌ではortholog遺伝子のゲノム上での配置は専らシャッフルされているというTatusovら[1996]あるいは、Watanabeら[1997]の結果とは非常に対照的である。このことは、MushegianとKoonin[1996]が示唆したように、機能単位としてサブユニット構造をもつ遺伝子群には、その遺伝子がゲノム上で離散しないような淘汰圧の存在を示すのかもしれない。本研究では、古細菌(Mj)ゲノムでも、ABC輸送系の遺伝子群はオペロン構造を持つであろうことが示された。

これに対してSsでは、他の6生物種程ABC輸送系のオペロン構造の保存が見られなかった。表3に示したように、ATP結合蛋白質と膜蛋白質が融合しているグループ6と、オペロンを形成しないグループ8に属するATP結合蛋白質を除いても、21個のABC輸送系オペロンを形成しないであろうATP結合蛋白質が、Ssには存在する。これ

は、個数の面から見ても、各生物種の持つATP結合蛋白質数との割合で考えても、明らかに他の6生物種とは異なる傾向を示す数字である。Ssのこうした傾向に関する考察は、次節4.3.で詳述する。

Ss以外の生物種に目を向けると、グループ6と8に属するATP結合蛋白質を除いてもABC輸送系オペロンを形成しないであろうATP結合蛋白質が、4生物種(Hp, Mg, Mp, Mj)に存在する。こういった蛋白質のうち、Hpの1個(HP1220)は、グループ7の節(3.8.)で述べた膜蛋白質とHlyD類似蛋白質と、レギュロンを構成するのかもしれない。Mpの4個のうちの一蛋白質、F10_orf326(グループ7)は、他の生物種との比較から、ABC輸送系様のオペロンを形成するであろうことが示唆されている[Himmelreich *et al.*, 1996]。グループ3-1に、2個と3個の「オペロンを形成しない」ATP結合蛋白質が、各々MgとMpに存在する。この蛋白質をコードしている遺伝子の隣接遺伝子産物は、ごくごく弱い配列類似性を示すので、これらがオペロンを形成するか否かは、さらなる解析を待ちたい。Mjの2個は、グループ4の節(3.5.)で述べたmethylcoenzyme M還元酵素と推定される蛋白質である。

本研究では、ABC輸送系(および類似)オペロンを25のグループ(サブグループも含めて)に分類した。このうち、7生物種全体に存在しているのは、グループ1-1と5-1の二グループのみである。しかし、グラム陰性菌、グラム陽性菌、シアノバクテリア、古細菌という分類で見ると、これら4種にまたがって存在しているオペロンは、前記の2つとグループ3-7、グループ5-3の四グループである。また、グラム陰性菌、陽性菌のうちいずれかと、シアノバクテリア、古細菌にまたがって存在しているオペロンは、六グループ(1-1、1-2、3-7、5-1、5-3、9)ある。さらに、真正細菌と古細菌という分類で見ると、これら生物界の2ドメインにまたがって存在しているオペロンは、前記の6グループとグループ3-1、3-5、7の九つである。これらのオペロンは、細胞を維持していく上で非常に重要な系なのであろう。

最近その全ゲノム配列が明らかにされた枯草菌[Kunst *et al.*, 1997]とも比較を行うと、生物種を越えて保存されているオペロン数は増えるかもしれない。枯草菌のゲノムサイズは、同じグラム陽性菌に属するMgとMpよりも大きく、遺伝子数もEcに匹敵する。

表3. 各生物種のもつABC輸送系オペロンの数

生物種	Gram-negative bacteria			Gram-positive bacteria		Cyano-bacteria	Archaea
	Ec	Hi	Hp	Mg	Mp	Ss	Mj
オペロンを形成している ATP結合蛋白質の数	72	32	11	13	13	21	14
オペロン数	61	30	10	9	9	18	14
オペロンを形成していな いATP結合蛋白質の数	6	7	7	3	4	33	3
オペロンを形成していな いATP結合蛋白質の数(グ ループ6と8を除く)	0	0	3	2	4	21	2

4.3. *Synechocystis*におけるABC輸送系遺伝子の特徴

前節4.2.で述べたように、今回比較した7生物種の中では、SsのゲノムでのみABC輸送系オペロンの強固な保持が見られない。この点に関していくつかの考察を述べる。

Ssのゲノム上で、ABC輸送系オペロンを形成しないATP結合蛋白質をコードする遺伝子の周辺に、新規に保存されている遺伝子がないか探索した結果、sll0488とsll0913が、各々グループ7(sll0489)とグループ8(sll0912)の遺伝子の隣にコードされていることが判明した。これらの遺伝子産物は、約200アミノ酸残基長で、今回比較した7生物種の中では、Ssのみにみられる。これらの遺伝子産物に配列類似性蛋白質は、前述の2つ以外に18個Ss中に存在し、しかも、そのうち2個(sll0740, slr0981)が、各々グループ1-1(sll0739)とグループ1 1(slr0982)のABC輸送系オペロン近傍にコードされている。これらの蛋白質は、輸送に関与しているのだろうか？あるいはSsに特有の新規な機能を担うのであろうか？

上記の例以外に、Ssに特有の構成のオペロンがなければ、SsはレギュロンによりABC輸送系蛋白質の発現を制御している可能性が示唆される。なぜなら、結果の章でも触れているように、ABC輸送系オペロンの構成要素、つまりATP結合蛋白質以外にも、同じグループの膜蛋白質や基質結合蛋白質を、そのゲノム上に保持しているからである。ただしこれをさらに確かめるには、そうした膜蛋白質や基質結合蛋白質をコードする遺伝子周辺についても、新規なオペロンの有無を調査する必要がある。

もう一つSsのABC輸送系蛋白質の発現に関して特徴的なこととして、ABC輸送系蛋白質の融合が挙げられる。TamとSaierのレビュー[Tam and Saier, 1993]によれば、それまで基質結合蛋白質とABC輸送オペロンの他の要素との融合蛋白質は発見されていなかった。しかし、Ssでは、基質結合蛋白質と膜蛋白質との融合蛋白質(sll1270)や基質結合蛋白質とATP結合蛋白質との融合蛋白質(sll1452, slr0043)が存在する。他にも膜蛋白質とATP結合蛋白質との融合蛋白質が幾つか存在する。

4.4. 既存の分類との比較

これまで、ABC輸送系のATP結合蛋白質、あるいはこれに類似した蛋白質の、アミノ酸配列類似性に基づく分類が、ABC輸送系オペロンの分類にも非常に良く結びつくことを示してきた。典型的なABC輸送系オペロンの構成要素は、ATP結合蛋白質の他に、膜蛋白質と基質結合蛋白質がある。ここでは、本研究の分類結果を、既に発表されている膜蛋白質と基質結合蛋白質の分類と比較し、本研究の結果が妥当であることを示す。さらに、この輸送系オペロンの祖先型についても考察する。

比較に用いた分類は、膜蛋白質については1件[Saurin and Dassa, 1994]、基質結合

蛋白質についてが2件[Tam and Saier, 1993; Murzin *et al.*, 1995]である。Murzin等は、立体構造既知の蛋白質について、その構造類似性に基づく分類を、SCOPというデータベースとして公開している(<http://scop.mrc-lmb.cam.ac.uk/scop/>)。

図10に本研究の分類結果と他の3件の分類とを併せて示す。この図から明らかな様に、本研究の分類結果は、これら3件の分類とほとんどの場合良い一致を示す。図10中の、A, B, Fのクラスターは、膜蛋白質を閾値 $E < 0.001$ で、最短距離法を用いて分類した時に、新しく形成されるクラスターである。この新たな分類も、3.1.で述べた7生物種のATP結合蛋白質全体でのクラスター分析の結果と、概して良い一致を示す。本研究での分類のうち、グループ3のオペロンのみが、やや不一致な、モザイク的オペロン構成を示す。しかしこれは、ATP結合蛋白質全体でのクラスター分析の結果から考えても、本研究の分類が特に誤っているという訳ではない。グループ3のATP結合蛋白質は、相当早い時期にA, B, Cの3種類のオペロンを形成したのかもしれない(次節4.5.参照)。

SCOPでは、ABC輸送系の基質結合蛋白質の立体構造は、IとIIという2つのタイプのfoldに大別されている。この分類もまた、非常に良い一致を示す。図10中のクラスターAに属するオペロンのうち、グループ1-1、3-3、3-7、4-1のオペロンの基質結合蛋白質の立体構造が明らかになっており、それらはいずれもIIタイプのfoldに属している。また、クラスターBに属するオペロンの、グループ2と3-5、両方の基質結合蛋白質の立体構造が明らかになっており、それらはいずれもIタイプのfoldに属していることが知られている。

以上の事実を踏まえて整理統合された、図10中のクラスターAからGまでの7タイプが、ABC輸送系、もしくはそれに類似したオペロンの祖先型として考えられる。ただし、この7クラスター以外のグループについては、存在する生物種が限られていることから、祖先型か否かの判断は保留してある。

話を再び基質結合蛋白質に戻す。クラスターAのグループ1-2、および3-2のオペロンの基質結合蛋白質の立体構造は、まだ明らかにされてはいないが、上述した事実から、IIタイプのfoldを持つ可能性が示唆される。ただし、グループ3-2のオペロンに保存されているリポ蛋白質が、基質結合蛋白質であるかどうかはまだ検討の余地がある。

図10中のクラスターで、AとB以外に基質結合蛋白質を持つクラスターに、D(グループ5-1)があるが、このグループの基質結合蛋白質はIタイプあるいはIIタイプ以外の立体構造をとるのだろうか? 手掛かりは、一つある。このグループのEcのオペロンの基質結合蛋白質BtuE(b1710)は、その明らかな配列類似性から、既に立体構造が判明している牛のグルタチオンペルオキシダーゼに似た立体構造をとると考えられる。グルタチオンペルオキシダーゼは、前述のIタイプあるいはIIタイプのfoldとは異なる

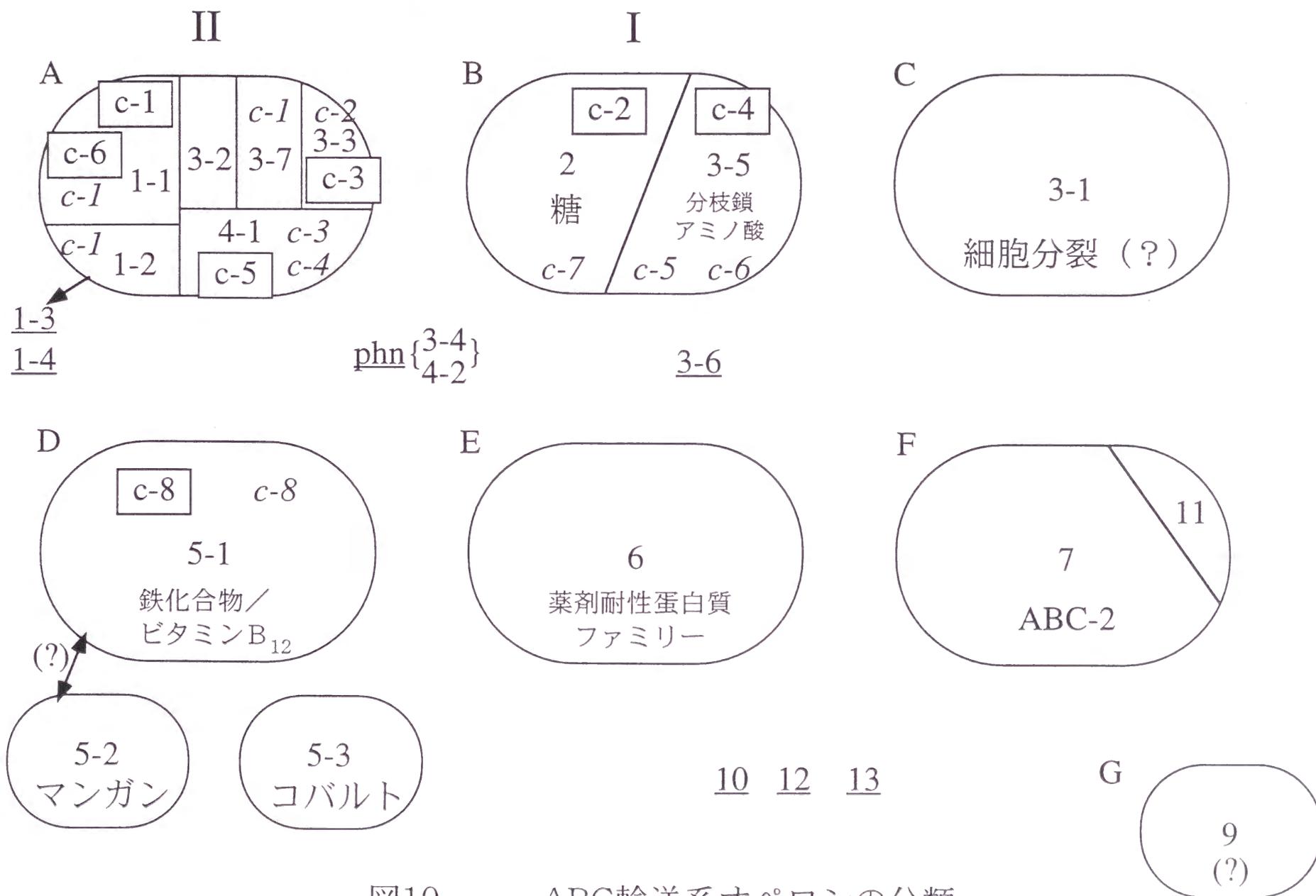


図10 ABC輸送系オペロンの分類

図10 四角囲みの数字はTamとSaier [1993]による基質結合蛋白質の分類、斜字体の数字はSaurinとDassa [1994]による膜蛋白質の分類、それ以外の数字は本研究による分類を示す。IIとIは、SCOPでの基質結合蛋白質の立体構造の分類を示す。

立体構造である。この構造こそが、クラスターDの基質結合蛋白質の立体構造である可能性もある。しかし、これには次の2点から、検討の余地があることを付記しておく。まず、この蛋白質BtuEは、輸送には必須ではない[Rioux and Kadner, 1989]。次に、この蛋白質は、クラスターDの他の基質結合蛋白質と十分な配列類似性を示さない。

Tam and Saierによる分類のうち、cluster 7の基質結合蛋白質[Shaw *et al.*, 1991]は、ABC輸送系オペロンを構成していない。しかし、この種の基質結合蛋白質をコードする遺伝子の周辺には、ABC輸送系とは全く別の種類の遺伝子がEcとHiで保存されている(図11)。基質結合蛋白質は、化学受容器として機能することも知られている[Tam and Saier, 1993]ので、これらの遺伝子はそういった機能に関係しているのかもしれない。

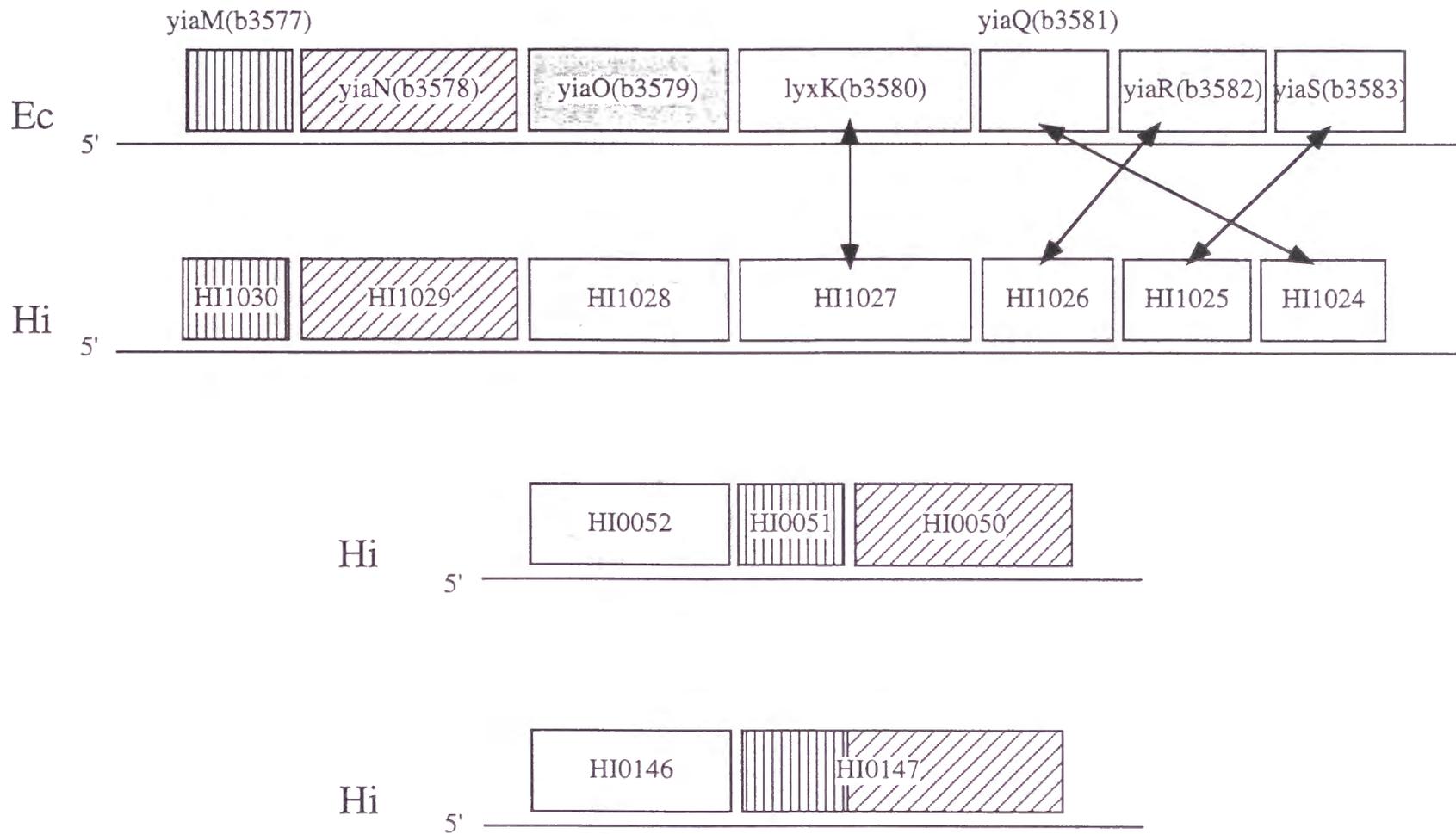


図11 Tam と Saier [1993] の分類のクラスター7に属する基質結合蛋白質の遺伝子周辺の保存影をつけた箱が、クラスター7の基質結合蛋白質類似遺伝子である。周辺に保存されている遺伝子の対応関係は、網かけと矢印で示した。

4.5. ABC輸送系オペロンの進化

大局的に眺めると、真正細菌ゲノムでは、異なる生物種間で、orthologous遺伝子の配置は保存されていない[Tatusov *et al.*, 1996; Watanabe *et al.*, 1997]。しかしこれとは対照的に、局所的にみると、4.2.節でも述べたように、今回比較を行った7生物種のうちSsを除く6生物種のゲノム上では、ABC輸送系の遺伝子群は非常に強固に集積し、しばしばortholog遺伝子の配置もそのまま保存され、オペロンを形成している。

また基本的に、ABC輸送系のATP結合蛋白質、あるいはこれに類似した蛋白質の、アミノ酸配列類似性に基づく本研究での分類が、ABC輸送系オペロンの分類にも非常に良く結びつくことを示してきた。つまり、オペロンを形成している要素（典型的には、ATP結合蛋白質、膜蛋白質、基質結合蛋白質の3つ）の分類に、相関が見られるということである。

以上の事実は、ABC輸送系オペロンが、一つの遺伝単位として進化してきたであろうことを、強く示唆する。

4.2.節で述べたように、異なる9種類のABC輸送系オペロンが、細菌と古細菌という生物界の2ドメインにまたがって、現存する生物のゲノム上に保存されている。保存されている理由には、もちろんこれらのオペロンが、細胞を維持していく上で非常に重要な系であるという理由もあろう。それに加えて、こうした種類の異なるオペロンが、進化上非常に離れた細菌と古細菌にわたり存在するという事は、この生物界の2ドメイン分岐以前に、既にこの9種類のABC輸送系オペロンが形成されていた可能性がきわめて高いことを示すと考えられる。こう考えると、これらが細胞を維持していく上で必須の存在であったからこそ、細菌と古細菌の分岐以前の共通祖先の生物が、既にこの9種類のオペロンを持っていたと考えられる。

これら9種類のオペロンを、図10と比較するとさらに興味深い可能性が示される。図10では、オペロン構成要素の配列・立体構造類似性から、ABC輸送系オペロンの祖先型をAからGの7種類に分類した訳であるが、これによると、真正細菌と古細菌の分岐以前の共通祖先の生物の時点で、クラスターAは他のグループに比べて良いアミノ酸配列保存度を有しているにもかかわらず、既に3種（1-1、1-2、3-7）に分岐していたことになる。これは、現存する生物でのこのクラスターの驚く程の機能の多様化から考えても、妥当な結果と思える。このクラスターは、配列を保存しながら、様々な機能を獲得して行く遺伝単位の好例であろう。クラスターAのオペロンの起源は相当古いのかもしれない。

その他の祖先型オペロンは、クラスターBCDFGから各々1グループずつと、グループ5-3オペロンが保持されていることになる。（ただしグループ5-3オペロンは、本研究の結果（ATPに結合蛋白質による分類に）従えば、クラスターDの一員であるが、

オペロンの他の構成要素が、クラスターDのそれと配列類似性を示さないので祖先型Dに含まれるか否かの判断は保留してある。)これは、図10の分類が妥当なものであることを裏付ける。クラスターEのオペロンは、古細菌(Mj)には存在しないが、哺乳類に明らかな類似蛋白質が存在する(3.7.参照)ので、おそらく、真正細菌と古細菌の分岐以前の共通祖先にも存在したであろう。また、グループ8の蛋白質は、輸送以外の機能を持ちオペロンを形成しないと推定されるが、真正細菌にも古細菌(および真核生物(3.9.参照))にも存在する。従って、この蛋白質も共通祖先には既に存在したと考えられる。

以上をまとめると、真正細菌と古細菌(さらには真核生物)の分岐以前の共通祖先は、3つに分岐したクラスターAのオペロンとBからGのオペロンを1つずつ、既に有していたと考えられる。

最後に、種内でのオペロンの重複について述べる。本研究の結果に従えば、Ecにおける多量の輸送系遺伝子の存在[Tatusov *et al.*, 1996; Blattner *et al.*, 1997]は、ABC輸送系に関する限り、冗長性というよりはむしろ、多様な基質に対する特異性の獲得を反映していると考えられる。これは、表2より明らかである。

5. 結論

全ゲノム配列が明らかにされた7生物種中に存在するABC輸送系のATP結合蛋白質および類似蛋白質の分類を、そのアミノ酸配列の類似性およびorthologous関係に基づいて行った。そしてこの分類が、その蛋白質が含まれるであろうオペロンの構造に非常に良い結びつきがあることを示した。またこの分類は、オペロンの機能についても、ある程度の良い分類となっている。これは、orthologous関係を考慮した本研究の分類法が、非常に有効なものであることを示す。

今回、分類されたABC輸送系のATP結合蛋白質をコードしている遺伝子の近傍の遺伝子の保存性を、orthologous関係だけでなくparalogous関係についても解析することにより、新規のオペロン構成要素を発見した。また今回の解析を通して、4.1.1で纏めた8点について、各オペロンの機能を推定することができた。さらに、EcでのABC輸送系オペロンとべん毛遺伝子の発現制御との関連について示すことができた。こうした本研究のオペロン探索手法は、これから益々さかんになっていくであろうゲノム配列の解析に、一つの有効な方途を示したと言えるであろう。

真正細菌の進化を通して、orthologous遺伝子のゲノム上での配置は専らシャッフルされているという報告[Tatusov *et al.*, 1996; Watanabe *et al.*, 1997]とは対照的に、ABC輸送系の遺伝子群は一般に非常に強固に集積し、オペロンを形成している。また上にも述べたように、オペロン要素（典型的には、ATP結合蛋白質、膜蛋白質、基質結合蛋白質の3つ）の分類の間に、非常に良い相関が見られることから、ABC輸送系の遺伝子群は、オペロンを遺伝単位として進化してきたであろうことが示唆される。

さらに、今回の解析および他のグループの膜蛋白質や基質結合蛋白質の分類結果との比較から、真正細菌と古細菌の分岐以前の共通祖先は、3つに分岐したクラスターAのオペロンとBからGのオペロンを1つずつ、既に有していたと考えられる。また、ABC輸送系オペロンの祖先型、7種類を提案した。

現在、続々と異なる生物種のゲノム配列が明らかになりつつある。本研究の成果は、こうした配列を用いて比較解析を行う生物種を増すことにより、さらに確かめることが重要であるし、また、より確かなものになると考えられる。

本研究で得られたオペロンの機能に関する新たな知見は、今後、我々の研究室で構築しているKEGGデータベース(<http://www.genome.ad.jp/kegg/>)を通して公開される予定である。

文献

- Bairoch, A., Bucher, P. 1994. PROSITE: recent developments. *Nucleic Acids Res.* 22:3583-3589.
- Bartsevich, V. V., Pakrasi, H. B. 1995. Molecular identification of an ABC transporter complex for manganese: analysis of a cyanobacterial mutant strain impaired in the photosynthetic oxygen evolution process. *EMBO J.* 14:1845-1853.
- Bauer, P. D., Trapp, C., Drake, D., Taylor, K. G., Doyle, R. J. 1993. Acquisition of manganous ions by mutants group streptococci. *J. Bacteriol.* 175:819-825.
- Bisbal, C. Martinand, C., Silhol, M., Lebleu, B., Salehzada, T. 1995. Cloning and characterization of a RNase L inhibitor. A new component of the interferon-regulated 2-5A pathway. *J. Biol. Chem.* 270:13308-13317.
- Blattner, F. R., Plunkett III, G., Bloch, C. A., Perna, N. T., Burland, V., Riley, M., Collado-Vides, J., Glasner, J. D., Rode, C. K., Mayhew, G. F., Gregor, J., Davis, N. W., Kirkpatrick, H. A., Goeden, M. A., Rose, D. J., Mau, B., Shao, Y. 1997. The complete genome sequence of *Escherichia coli* K-12. *Science* 277:1453-1462.
- Brenner, S. E., Hubbard, T., Murzin, A., Chothia, C. 1995. Gene duplications in *H. influenzae*. *Nature* 378:140.
- Bult, C. J., White, O., Olsen, G. J., Zhou, L., Fleischmann, R. D., Sutton, G. G., Blake, J. A., FitzGerald, L. M., Clayton, R. A., Gocayne, J. D., Kerlavage, A. R., Dougherty, B. A., Tomb, J. F., Adams, M. D., Reich, C. I., Overbeek, R., Kirkness, E. F., Weinstock, K. G., Merrick, J. M., Glodek, A., Scott, J. L., Geoghagen, N. S. M. and Venter, J. C. 1996) Complete genome sequence of the methanogenic archaeon, *Methanococcus jannaschii*. *Science* 273:1058-1073.
- Cheng, J., Guffanti, A. A., Krulwich, T. A. 1997. A two-gene ABC-type transport system that extrudes Na⁺ in *Bacillus subtilis* is induced by ethanol or protonophore. *Mol. Microbiol.* 23:1107-1120.
- Dudler, R., Schmidhauser, C., Parish, R. W., Wettenhall, R. E., Schmidt, T. 1988. A mycoplasma high-affinity transport system and the in vitro invasiveness of mouse sarcoma cells. *EMBO J.* 7:3963-3970.
- Fleischmann, R. D., Adams, M. D., White, O., Clayton, R. A., Kirkness, E. F., Kerlavage, A. R., Bult, C. J., Tomb, J.-F., Dougherty, B. A., Merrick, J. M., McKenney, K., Sutton, G., FitzHugh, W., Fields, C. A., Gocayne, J. D., Scott, J. D., Shirley, R., Liu, L.-I., Glodek, A., Kelley, J. M., Weidman, J. F., Phillips, C. A., Spriggs, T., Hedblom, E., Cotton, M. D., Utterback, T. R., Hanna, M. C., Nguyen, D. T., Saudek, D. M., Brandon, R. C., Fine, L. D.,

- Fritchman, J. L., Fuhrmann, J. L., Geoghagen, N. S. M., Gnehm, C. L., McDonald, L. A., Small, K. V., Fraser, C. M., Smith, H. O. and Venter, J. C. 1995. Whole-genome random sequencing and assembly of *Haemophilus influenzae* Rd. *Science* 269:496–512.
- Fraser, C. M., Gocayne, J. D., White, O., Adams, M. D., Clayton, R. A., Fleischmann, R. D., Bult, C. J., Kerlavage, A. R., Sutton, G., Kelley, J. M., Fritchman, J. L., Weidman, J. F., Small, K. V., Sandusky, M., Fuhrmann, J. L., Nguyen, D. T., Utterback, T. R., Saudek, D. M., Phillips, C. A., Merrick, J. M., Tomb, J.-F., Dougherty, B. A., Bott, K. F., Hu, P.-C., Lucier, T. S., Peterson, S. N., Smith, H. O., Hutchison, C. A. III and Venter, J. C. 1995. The minimal gene complement of *Mycoplasma genitalium*. *Science* 270:397–403.
- Froshauer, S., Beckwith, J. 1984. The nucleotide sequence of the gene for malF protein, an inner membrane component of the maltose transport system of *Escherichia coli*. Repeated DNA sequences are found in the malE-malF intercistronic region. *J. Biol. Chem.* 259:10896–10903.
- Furuchi, T., Kashiwagi, K., Kobayashi, H., Igarashi, K. 1991. Characteristics of the gene for a spermidine and putrescine transport system that maps at 15min on the *Escherichia coli* chromosome. *J. Biol. Chem.* 266:20928–20933.
- Gill, D.R., Hatfull, G. F., Salmond, G. P. 1986. A new cell division operon in *Escherichia coli*. *Mol. Gen. Genet.* 205:134–145.
- Hecht, G. B., Newton, A. 1995. Identification of a novel response regulator required for the swarmer-to-stalked-cell transition in *Caulobacter crescentus*. *J. Bacteriol.* 177:6223–6229.
- Henikoff, S., Henikoff, J. G. 1992. Amino acid substitution matrices from protein blocks. *Proc. Natl. Acad. Sci. U. S. A.* 89:10915–10919.
- Higgins, C. F. 1992. ABC transporters: from microorganisms to man. *Annu. Rev. Cell Biol.* 8:67–113.
- Himmelreich, R., Hilbert, H., Plagens, H., Pirkl, E., Li, B.-C. and Herrmann, R. 1996. Complete sequence analysis of the genome of the bacterium *Mycoplasma pneumoniae*. *Nucleic Acids Res.* 24:4420–4449.
- Hryniewicz, M., Sirko, A., Hulanicka, D. 1989. Identification and mapping of the sulphate permease promoter region in *Escherichia coli*. *Acta Biochim. Pol.* 36:353–363.
- Hunt, M. D., Pettis, G. S., McIntosh, M. A. 1994. Promoter and operator determinants for fur-mediated iron regulation in the bidirectional fepA-fes control region of the *Escherichia coli* enterobactin gene system. *J. Bacteriol.* 176:3944–3955.

- Kaneko, T., Sato, S., Kotani, H., Tanaka, A., Asamizu, E., Nakamura, Y., Miyajima, N., Hirose, M., Sugiura, M., Sasamoto, S., Kimura, T., Hosouchi, T., Matsuno, A., Muraki, A., Nakazaki, N., Naruo, K., Okumura, S., Shimpo, S., Takeuchi, C., Wada, T., Watanabe, A., Yamada, M., Yasuda, M., Tabata, S. 1996. Sequence analysis of the genome of the unicellular cyanobacterium *Synechocystis* sp. strain PCC6803. II. Sequence determination of the entire genome and assignment of potential protein-coding regions. *DNA Res.* 3:109-136.
- Karow, M., Georgopoulos, C. 1993. The essential *Escherichia coli* *msbA* gene, a multicopy suppressor of null mutations in the *htrB* gene, is related to the universally conserved family of ATP-dependent translocators. *Mol. Microbiol.* 7:69-79.
- Kuhner, C. H., Lindenbach, B. D., Wolfe, R. S. 1993. Component A2 of methylcoenzyme M reductase system from *Methanobacterium thermoautotrophicum* delta H: nucleotide sequence and functional expression by *Escherichia coli*. *J. Bacteriol.* 175:3195-3203.
- Kunst, F., Ogasawara, N., Moszer, I., Albertini, A. M., Alloni, G., Azevedo, V., Bertero, M. G., Bessieres, P., Bolotin, A., Borchert, S., Yoshida, K., Yoshikawa, H. F., Zumstein, E., Yoshikawa, H. and Danchin, A. et al. 1997. The complete genome sequence of the Gram-positive bacterium *Bacillus subtilis* *Nature* 390:249-256.
- Lavina, M., Pugsley, A.P., Moreno, F. 1986. Identification, mapping, cloning and characterization of a gene (*sbmA*) required for microcin B17 action on *Escherichia coli* K12. *J. Gen. Microbiol.* 132:1685-1693.
- Makino, K., Kim, S. K., Shinagawa, H., Amemura, M., Nakata, A. 1991. Molecular analysis of the cryptic and functional *phn* operons for phosphonate use in *Escherichia coli* K-12. *J. Bacteriol.* 173:2665-2612.
- Murzin, A. G., Brenner, S. E., Hubbard, T., Chothia, C. 1995. SCOP: a structural classification of proteins database for the investigation of sequences and structures. *J. Mol. Biol.* 247:536-540.
- Mushegian, A.R., Koonin, E.V. 1996. Gene order is not conserved in bacterial evolution. *Trends Genet.* 12:289-290.
- Mytelka, D. S., Chamberlin, M. J. 1996. *Escherichia coli* *fliAZY* operon. *J. Bacteriol.* 178:24-34.
- Navarro, C., Wu, L. F., Mandrand-Berthelot, M. A. 1993. The *nik* operon of *Escherichia coli* encodes a periplasmic binding protein-dependent transport system for nickel. *Mol. Microbiol.* 9:1181-1191.
- Overduin, P., Boos, W., Tommassen, J. 1988. Nucleotide sequence of the *ugp* genes of *Escherichia coli* K-12: Homology to the maltose system. *Mol.*

Microbiol. 2:767-775.

- Ozenberger, B. A., Nahlik, M. S., McIntosh, M. A. 1987. Genetic organization of multiple *fep* genes encoding ferric enterobactin transport functions in *Escherichia coli*. *J. Bacteriol.* 169:3638-3646.
- Pearson, W. R. 1991. Searching protein sequence libraries: comparison of the sensitivity and selectivity of the Smith-Waterman and FASTA algorithms. *Genomics* 11:635-650.
- Powell, B. S., Court, D. L., Inada, T., Nakamura, Y., Michotey, V., Cui, X., Reizer, A., Saier, M. H. Jr., Reizer, J. 1995. Novel proteins of the phosphotransferase system encoded within the *rpoN* operon of *Escherichia coli*. Enzyme IIANtr affects growth on organic nitrogen and the conditional lethality of an *erats* mutant. *J. Biol. Chem.* 270:4822-4839.
- Quadroni, M., Staudenmann, W., Kertesz, M., James, P. 1996. Analysis of global responses by protein and peptide fingerprinting of proteins isolated by two-dimensional gel electrophoresis. Application to the sulfate-starvation response of *Escherichia coli*. *Eur. J. Biochem.* 239:773-781.
- Quirk, P. G., Guffanti, A. A., Clejan, S., Cheng, J. and Krulwich, T. A. 1994. Isolation of Tn917 insertional mutants of *Bacillus subtilis* that are resistant to the protonophore carbonyl cyanide *m*-chlorophenylhydrazone. *Biochim. Biophys. Acta* 1186:27-34.
- Raux, E., Lanois, A., Levillayer, F., Warren, M. J., Brody, E., Rambach, A., Thermes, C. 1996. *Salmonella typhimurium* cobalamin (vitamin B12) biosynthetic genes: functional studies in *S. typhimurium* and *Escherichia coli*. *J. Bacteriol.* 178:753-767.
- Reizer, J., Reizer, A., Saier, M. H. Jr. 1992. A new subfamily of bacterial ABC-type transport systems catalyzing export of drugs and carbohydrates. *Protein Sci.* 1:1326-1332.
- Rioux, C.R., Kadner, R.J. 1989. Vitamin B12 transport in *Escherichia coli* K12 does not require the *btuE* gene of the *btuCED* operon. *Mol. Gen. Genet.* 1989 217:301-308.
- Roth, J. R., Lawrence, J. G., Rubenfield, M., Kieffer-Higgins, S., Church, G. M. 1993. Characterization of the cobalamin (vitamin B12) biosynthetic genes of *Salmonella typhimurium*. *J. Bacteriol.* 175:3303-3316.
- Sanders, J. D., Cope, L. D. and Hansen, E. J. 1994. Identification of a locus involved in the utilization of iron by *Haemophilus influenzae*. *Infect. Immun.* 62:4515-4525.
- Saurin, W., Dassa, E. 1994. Sequence relationships between integral inner membrane proteins of binding protein-dependent transport systems: evolution by recurrent gene duplications. *Protein Sci.* 3:325-344.

- Shaw, J. G., Hamblin, M. J., Kelly, D. J. 1991. Purification, characterization and nucleotide sequence of the periplasmic C4-dicarboxylate-binding protein (DctP) from *Rhodobacter capsulatus*. *Mol. Microbiol.* 5:3055-3062.
- Shea, C. M., McIntosh, M. A. 1991. Nucleotide sequence and genetic organization of the ferric enterobactin transport system: homology to other periplasmic binding protein-dependent systems in *Escherichia coli*. *Mol. Microbiol.* 5:1415-1428.
- Sirko, A., Hryniewicz, M., Hulanicka, D., and Bock, A. 1990. Sulfate and thiosulfate transport in *Escherichia coli* K-12: nucleotide sequence and expression of the *cysTWAM* gene cluster. *J. Bacteriol.* 172:3351-3357.
- Smith, D. R., Richterich, P., Rubenfield, M., Rice, P.W., Butler, C., Lee, H.-M., Kirst, S., Gundersen, K., Abendschan, K., Xu, Q., Chung, M., Deloughery, C., Aldredge, T., Maher, J., Lundstrom, R., Tulig, C., Falls, K., Imrich, J., Torrey, D., Engelstein, M., Breton, G., Madan, D., Nietupski, R., Seitz, B., Connelly, S., McDougall, S., Safer, H., Gibson, R., Doucette-Stamm, L., Eiglmeier, K., Bergh, S., Cole, S. T., Robison, K., Richterich, L., Johnson, J., Church, G. M., Mao, J.-I., 1997a. Multiplex sequencing of 1.5 Mb of the *Mycobacterium leprae* genome. *Genome Res.* 7:802-819.
- Smith, D. R., Doucette-Stamm, L. A., Deloughery, C., Lee, H.-M., Dubois, J., Aldredge, T. *et al.*, 1997b. Complete genome sequence of *Methanobacterium thermoautotrophicum* delta H: functional analysis and comparative genomics. *J. Bacteriol.* 179:7135-7155.
- Smith, R. L., Banks, J. L., Snavely, M. D., Maguire, M. E. 1993. Sequence and topology of the CorA magnesium transport systems of *Salmonella typhimurium* and *Escherichia coli*. Identification of a new class of transport protein. *J. Biol. Chem.* 268:14071-14080.
- 杉山 成、五十嵐一衛、森川耿右 1997. ポリアミン結合タンパク質(PotD)の構造と機能. *生化学.* 69:113-119.
- Surin, B., Rosenberg, H., Cox, G.B. 1985. Phosphate-specific transport system of *Escherichia coli*: Nucleotide sequence and gene-polypeptide relationships. *J. Bacteriol.* 161:189-198.
- Tam, R., Saier, M. H. Jr. 1993. Structural, functional, and evolutionary relationships among extracellular solute-binding receptors of bacteria. *Microbiol. Rev.* 57:320-346.
- Tatusov, R.L., Mushegian, A.R., Bork, P., Brown, N.P., Hayes, W.S., Borodovsky, M., Rudd, K.E., Koonin, E.V. 1996. Metabolism and evolution of *Haemophilus influenzae* deduced from a whole-genome comparison with *Escherichia coli*. *Curr. Biol.* 6:279-291.

- Thompson, J.D., Higgins, D.G., Gibson, T.J. 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res.* 22:4673-4680.
- Thony-Meyer, L., Fischer, F., Kunzler, P., Ritz, D., Hennecke, H. 1995. *Escherichia coli* genes required for cytochrome c maturation. *J. Bacteriol.* 177:4321-4326.
- Tomb, J. F., White, O., Kerlavage, A. R., Clayton, R. A., Sutton, G. G., Fleischmann, R. D., Ketchum, K. A., Klenk, H. P., Gill, S., Dougherty, B. A., Nelson, K., Quackenbush, J., Zhou, L., Kirkness, E. F., Peterson, S., Loftus, B., Richardson, D., Dodson, R., Khalak, H. G., Glodek, A., McKenney, K., Fitzgerald, L. M., Lee, N., Adams, M. D., Venter, J. C. et al. 1997. The complete genome sequence of the gastric pathogen *Helicobacter pylori*. *Nature* 388:539-547.
- Tomii, K., Kanehisa, M. Analysis of amino acid indices and mutation matrices for sequence comparison and structure prediction of proteins. *Protein Eng.* 9:27-36.
- 富井健太郎、金久實 1997. パスウェイ解析の現状. 蛋白質・核酸・酵素 3026-3032.
- van der Ploeg, J.R., Weiss, M.A., Saller, E., Nashimoto, H., Saito, N., Kertesz, M.A., Leisinger, T. 1996. Identification of sulfate starvation-regulated genes in *Escherichia coli*: a gene cluster involved in the utilization of taurine as a sulfur source. *J. Bacteriol.* 178:5438-5446.
- Walkenhorst, H. M., Hemschemeier, S.K., Eichenlaub, R. 1995. Molecular analysis of the molybdate uptake operon, modABCD, of *Escherichia coli* and modR, a regulatory gene. *Microbiol. Res.* 150:347-361.
- Watanabe, H., Mori, H., Itoh, T., Gojobori, T. 1997. Genome plasticity as a paradigm of eubacteria evolution. *J. Mol. Evol.* 44:57-64.

謝辞

本研究を進めるにあたり、御指導頂いた金久實教授に深く感謝致します。また、五斗進、藤渕航、緒方博之の各先生には、貴重な御助言を頂きました。深く感謝致します。研究室の学生、秘書の方々には、数々の御協力に対し、深い感謝の念を抱いております。本当にありがとうございました。私をたえず励まし、支えて下さった、家族、友人達に感謝の言葉を送ります。本当にありがとうございました。本研究は日本学術振興会の助成を受け、行われました。