

## 고온성과 초고온성 세균의 보존적 유전자 분석

이 동근 · 이재화 · 하배진 · 하종명 · <sup>1</sup>이정현 · <sup>1</sup>김상진 · † 이상현  
신라대학교 공과대학 생명공학과, <sup>1</sup>한국해양연구원 미생물연구실  
(접수 : 2005. 6. 9., 게재승인 : 2005. 10. 23.)

## Analysis of Conservative Genes in Thermophilic and Hyperthermophilic Bacteria

Dong-Geun Lee, Jae-Hwa Lee, Bae Jin Ha, Jong-Myung Ha, Jung-Hyun Lee<sup>1</sup>  
Sang-Jin Kim<sup>1</sup>, and Sang Hyeon Lee

Department of Bioscience and Biotechnology, College of Engineering, Silla University, Pusan 617-736, Korea  
<sup>1</sup>Microbiology Laboratory, Korea Ocean Research and Development Institute, Ansan PO Box 29, 425-600, Korea  
(Received : 2005. 6. 9., Accepted : 2005. 10. 23.)

Totally 16,299 conservative genes, commonly found in 13 thermophilic and hyperthermophilic bacteria, were analyzed. All genes were belong to 167 COGs (clusters of orthologous groups of proteins). COGs related to protein metabolism were 80 among 167 COGs. Conservative genes were not limited only thermophiles and hyperthermophiles, meaning thermal stability is independent of specific protein. However reverse gyrase was only found in all hyperthermophilic archaeobacteria and eubacteria, meaning DNA stability is important in hyperthermophiles. Hyperthermophilic eubacteria and thermophilic archaeobacteria had different position between phylogenetic tree of gene content and 16S rRNA gene. Thermophilic archaeobacteria, hyperthermophilic eubacteria and archaeobacteria had similar values by the statistical analysis of distance values with 167 COGs in each organism.

**Key Words :** Thermophiles, hyperthermophiles, 16S rRNA, gene content tree, COG (clusters of orthologous groups of proteins)

### 서 론

고온성 세균은 55°C 이상의 온도에서 성장하는 세균들로 고세균 (archaeobacteria)과 진정세균 (eubacteria)이 있으며, 고세균중 일부는 80°C 이상에서 최적성장을 하는데 이들을 초고온성 고세균이라고 한다(1). 이들은 고온에서 기능을 유지하는 효소들을 가지고 있어, 화학공정을 대처할 수 있으며 환경 문제와 경제적인 이유로 생물공학적인 응용과 산업적 가치도 높다(2, 3).

공통조상에서 유래하여 서로 다른 종에 있는 유전자들의 집합을 orthologs로 정의하며, COG (Clusters of Orthologous Groups of protein)는 ortholog들에서 유래된 단백질의 집합으로 유사한 구조와 기능을 갖는 것으로 알려져 있다(4). 각 COG는 적어도 3가지 이상의 종 (species)에서 유래된 단백질들로 구성되어 있어 하나의 공통조상 유전자에 해당하는 것으로 간주할 수 있다(5). 이 등은 유전체 염기서열 분석과

gene annotation을 통하여 확보된 COG를 이용한 보고들을 하였는데(6, 7), 신규자료의 확보 등으로 2005년 1월 현재 4873개의 COG 그룹으로 분류해 놓은 것을 인터넷을 통하여 접근할 수 있다(8).

유전체 염기서열과 gene annotation으로 얻어진 유전자는 기능을 예상만 할 수 있는 경우가 많지만, COG 알고리즘을 통해 예상 유전자에서 실제 기능을 확인한 보고들이 있어(9, 10) COG 알고리즘의 생물공학적 유용성을 알 수 있다.

고온성 효소들은 중온성 효소와 아미노산 서열에 차이가 있다는 보고는 있었지만(11), 고온성 세균들만이 가질 수 있는 공통 유전자에 대한 연구는 없었다. 본 연구에서는 고온성 및 초고온성 세균들만이 공통적으로 보유하는 유전자가 있는지 조사하고자 하였으며 유용한 단백질을 탐색하는 작업에 응용될 수 있는 COG를 이용한 분류법(6)을 시도하였다.

### 재료 및 방법

#### 재료

총 66종의 미생물 유전체를 NCBI의 공개 서버로부터 추출하였다(12). 이중 13종의 고온성 및 초고온성 원핵생물은

† Corresponding Author : Department of Bioscience and Biotechnology, Silla University, Pusan 617-736, Korea  
Tel : +82-51-999-5462, Fax : +82-51-999-5636  
E-mail : slee@silla.ac.kr

**Table 1.** Studied genomes derived from COGs database, number of COG and proteins for 13 species of thermophile and hyperthermophile

Phylogenetic Group	Organism	Abb. <sup>1</sup>	Number of COG <sup>2</sup>	Number of protein	NCBI # of 16S rRNA	Category
Eubacteria	<i>Aquifex aeolicus</i>	Aae	1088	1529		
	<i>Thermotoga maritima</i>	Tma	1167	1858		
Crenarchaeota	<i>Aeropyrum pernix</i>	Ape	947	1841		
	<i>Pyrobaculum aerophilum</i>	Pya	1016	2605		
	<i>Sulfolobus tokodaii</i>	Sso	1084	2825		Hyperthermophiles (80~113°C)
	<i>Archaeoglobus fulgidus</i>	Afu	1244	2420		
	<i>Methanococcus jannaschii</i>	Mja	1117	1729		
<i>Methanopyrus kandleri</i>	Mka	1022	1691			
Euryarchaeota	<i>Pyrococcus abyssi</i>	Pab	1065	1896	L19921	
	<i>Pyrococcus horikoshii</i>	Pho	1019	1956		
	<i>Methanobacterium thermoautotrophicum</i>	Mth	1123	1873		
	<i>Thermoplasma acidophilum</i>	Tac	890	1482	M38687	Thermophiles (55~80°C)
	<i>Thermoplasma volcanium</i>	Tvo	900	1499		

<sup>1</sup> Abbreviation<sup>2</sup> Number of COG was determined after compared 66 genomes of Archaea, Bacteria and microbial eucaryotes

Archaeobacteria가 11종, Eubacteria가 2종이었다. 미생물 유전자 유사성에 관한 자료는 COGs에서 정리된 자료를 이용하였는데(7) 이들은 2005년 1월 현재 66종의 미생물 유전체를 ortholog 그룹으로 분류하여, 총 144,324개의 유전자들을 4,873개의 COG 그룹으로 분류해 놓았다. Table 1은 본 연구에서 분석한 13종의 원핵생물의 분류 및 그들이 보유하고 있는 COG 자료를 나타내고 있다.

#### 계능 비교 및 유전자 보유 계통수 (gene content tree)

연구대상 각 생물종이 4,873개의 각 COG를 보유하고 있는지를 행렬로 작성하고 이를 Clustal 프로그램으로 neighbor-joining method와 bootstrap method (n = 1000)로 상관관계를 분석하였다(13).

#### 16S rRNA 유전자 염기서열 추출 및 분석

유전체 염기서열에서 rRNA 그룹을 조사하여 16S rRNA 유전자의 염기서열 중 1400 bp 이상인 것만을 추출하였고, 이것이 불가능한 *Pyrococcus abyssi*와 *Thermoplasma acidophilum*의 16S rRNA 유전자 염기서열은 NCBI 공개서버에서(14) 추출하였다. Table 1에 16S rRNA 유전자의 NCBI accession number가 표시되어 있다. Clustal 프로그램으로 다중염기배열 (multiple alignment)을 수행하고, 계능 비교와 동일하게 neighbor-joining method와 bootstrap method (n = 1000)로 분석하였다.

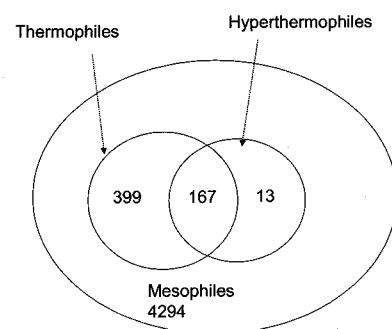
#### 유전자변이 비교

고온성 및 초고온성 세균 13종이 공통적으로 보유하는 167개의 COG에 속하는 총 16,299개의 단백질을 NCBI의 공개 서버에서 추출하고, Clustal 프로그램으로 정렬하였으며 각 COG당 그리고 각 생물당 distance value가 가장 적은 것을 추출하여 엑셀을 이용하여 통계처리하고 미생물 종별로 평균과 분산을 구하였다.

## 결과 및 고찰

### 계능 비교

Fig. 1은 고온성 및 초고온성 세균 전체가 중온성 세균 일부와 공통적으로 보유하는 COG의 수를 나타낸 것이다. 총 4,873개의 COG 중 12.2%인 599개의 COG를 고온성 세균 모두가 공통적으로 가지고 있었고, 180개의 COG를 초고온성 세균 모두가 가지고 있었다. 167개의 COG는 고온성과 초고온성 세균 전체가 공통적으로 가지고 있는 것으로 나타나 이들을 대상으로 분석을 계속하였다. 초고온성 고세균 8종 중 7종 이하에서 보이는 COG는 45개였고, 7개의 COG는 고온성 고세균, 초고온성 고세균 그리고 초고온성 진정세균의 각 일부에 존재하였지만, 이들은 전체에서 관찰되지 않아 분석대상에서 제외하였다. Table 2는 167개의 COG를 기능별로 분류한 것으로 각 번호 앞의 접두어 COG는 생략하였다. 해독 (translation) 관련 COG가 63개로 최대였고 아미노산 관련 COG가 17개 등 단백질 대사 관련 유전자가 80개로 전체 보존적 유전자의 47.9%에 해당하는 것을 알 수 있었다. 그의 일반적 기능만 알 수 있는 유전자, 핵산관련 유전자, 당관련 유전자, 복제관련 유전자 등의 순으로 보존적 유전자의 수가 많았다.



**Figure 1.** Number of common COG (Clusters of Orthologous Groups of protein) in all thermophiles and/or hyperthermophiles with mesophiles.

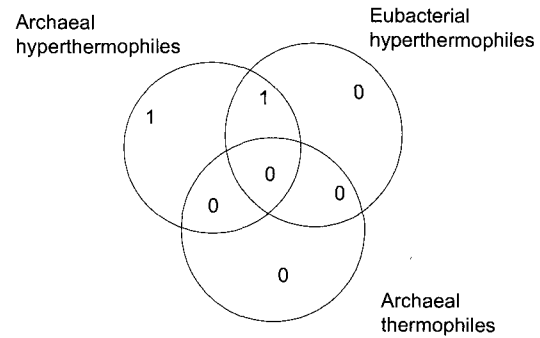
**Table 2.** Studied COGs and their functional category in thermophiles and hyperthermophiles. COG, prefix to each number, were omitted

Functional category	# of COG
Energy production and conservation	0644, 0674, 1013,1014
Cell cycle control, mitosis and meiosis	0037,0489
Amino acid transport and metabolism	0002, 0006, 0075, 0078, 0079, 0112, 0136, 0174, 0421, 0436, 0460, 0498, 0527, 0548, 0620, 0624, 4992
Nucleotide transport and metabolism	0005, 0044, 0125, 0127, 0167, 0284, 0461, 0504, 0518, 0519, 0528, 0540, 1618
Carbohydrate transport and metabolism	0057, 0061, 0063, 0126, 0148, 0149, 0483, 0574, 1109, 3635, 3839
Coenzyme metabolism	0142, 0171, 0237, 0294, 0351, 0452, 0499
Lipid transport and metabolism	0020, 0575
Translation	0008, 0009, 0012, 0013, 0016, 0018, 0024, 0030, 0048, 0049, 0051, 0052, 0060, 0072, 0080, 0081, 0087, 0088, 0089, 0090, 0091, 0092, 0093, 0094, 0096, 0097, 0098, 0099, 0100, 0102, 0103, 0124, 0130, 0143, 0162, 0172, 0180, 0182, 0184, 0185, 0186, 0197, 0198, 0199, 0200, 0231, 0244, 0255, 0256, 0343, 0361, 0441, 0442, 0480, 0495, 0522, 0525, 0532, 0621, 1514, 1841, 2519, 2890
	0085, 0086, 0195, 0202, 0250, 0640, 1758
	0164, 0177, 0258, 0350, 0358, 0419, 0420, 0468, 0470, 0550, 0592, 1110
	0438, 0463, 0668, 0750
	0071, 0330, 0459, 0492, 0533, 1180
	0704
	0073, 0312, 0433, 0446, 0456, 0491, 0517, 0535, 0705, 0714, 0730, 1078, 1355, 1708
	0062, 0432, 1371, 1690, 1912, 2078
	0467
	0201, 0541, 0552, 0681
0111, 0329, 0402, 0462, 0477, 0500	

**공통유전자**

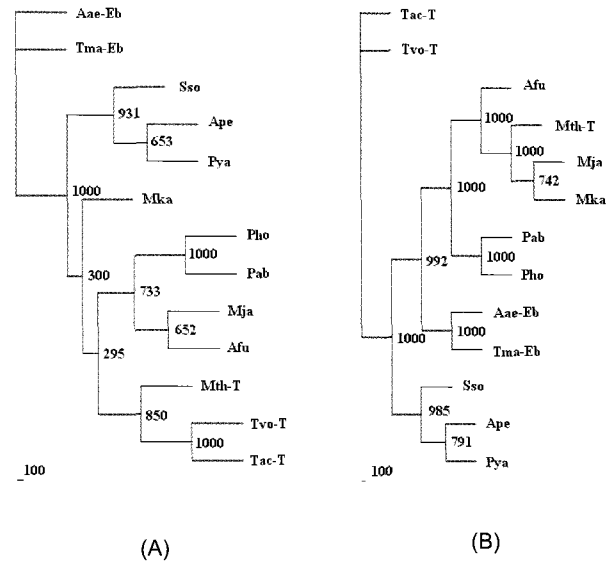
Fig. 2는 중온성 세균을 제외한 고온성 고세균, 초고온성 고세균 그리고 초고온성 진정세균 전체가 공통적으로 함유하는 COG의 개수를 나타내는 그림으로, 초고온성 세균과 고온성 세균 모두에 공통적인 유전자는 없었다. 공통적인 유전자가 없다는 것은 열안정성이 특정 단백질에 의해 주어지는 것은 아니라 단백질의 아미노산서열 수준 등에서 달성되는 것으로 생각되었다(11). 고온성 고세균 전체는 초고온성 고세균 및 진정세균 전체와 공통적인 유전자가 없었지만, 초고온성 고세균과 진정세균 전체는 COG1110 (reverse gyrase)를 공통유전자로 가지고 있었고, 초고온성 고세균 전체에서만 보이는 COG2248은 metallo-beta-lactamase superfamily에 속하는 hydrolase로 추정되는 공통

유전자였다.



**Figure 2.** Number of common COG (Clusters of Orthologous Groups of protein) in all thermophiles and/or hyperthermophiles excluding COGs showing in one more than mesophiles.

Reverse gyrase는 ATP를 사용하여 원형의 이중나선 DNA가 positive supercoiling을 이루게 하는데 고온에서 DNA의 안정성에 기여할 것으로 추정된다(15). 이는 80°C 이상의 고온에서는 고온성 세균이 성장하지 못하는 이유가 될 수도 있을 것이다. 그리고 고온성 고세균의 gene loss에 의해 초고온성을 상실했을 가능성과 reverse gyrase의 초고온성 진정세균으로의 horizontal gene transfer에 의한 가능성 등이 있는 것으로 사료되었다(16).



**Figure 3.** Comparison of the phylogenetic trees of thermophiles and hyperthermophiles obtained from neighbor-joining analysis of either 16S rRNA gene sequence (A) and gene content (B). Bootstrap values at each node are expressed as a number over 1000 trials. Abbreviation of organisms is in Table 1. Eubacteria and thermophiles were abbreviated as -Eb and -T after abbreviation of organism.

**16S rRNA 유전자와 유전자보유 계통수 비교**

Fig. 3A는 분석대상 세균들의 16S rRNA 유전자 계통수이다. 진정세균 (Tma, Aae)과 고세균이 따로 분리되고 (Fig. 3A), 고세균 내에서는 Crenarchaeota (Sso, Ape, Pya)와 Euryarchaeota가 따로 분리되었다. Euryarchaeota에서도 고온성 고세균들은 (Tac, Tvo, Mth) 내부적으로 응집하는 것을

볼 수 있었다.

COG 보유유무에 따른 유전자보유 계통수(Fig. 3B)와 16S rRNA 유전자 계통수를 비교한 후 도출된 공통점으로 는 첫째 진정세균 2종 (Tma, Aae), *Pyrococcus*속 2종 (Pho, Pab), *Thermoplasma*속 2종 (Tac, Tvo) 등이 모두 높은 bootstrap number로 같은 그룹에 위치하였다(6). 두번째 공 통점은 Crenarchaeota가 따로 분리되어 있는 점이다.

유전자보유 계통수와 16S rRNA 유전자 계통수의 차이점 은 첫째 초고온성 진정세균의 위치로 이들은 유전자보유 계 통수에서는 초고온성 고세균과 함께 분포하였지만 16S rRNA 유전자 계통수에서는 고세균들과 독립되는 것으로 나 타났다. 둘째 고온성 고세균인 Tac와 Tvo가 유전자보유 계 통수에서 다른 세균들과 독립적으로 존재하는 것이다. 셋째 메탄대사관련 세균들이 (Mja, Mka, Mth) 유전자보유 계통수 에서는 함께 분포하였지만 16S rRNA 유전자 계통수에서는 서로 떨어져 있는 점이다. 고온성 고세균이면서 메탄대사관 련 세균인 *Methanobacterium thermoautotrophicum* (Mth)를 유 전자보유 계통수로 보면, 초고온성 고세균이면서 메탄대사 관 련 세균인 Mja, Mka와 함께 존재하지만 다른 고온성 고세 균과 분리되어 있었다. 분석대상 메탄대사관련 고세균들이 공유하는 공통유전자수가, 고온성 고세균들 사이의 공통유 전자들보다 많다는 것을 생각할 수 있었다.

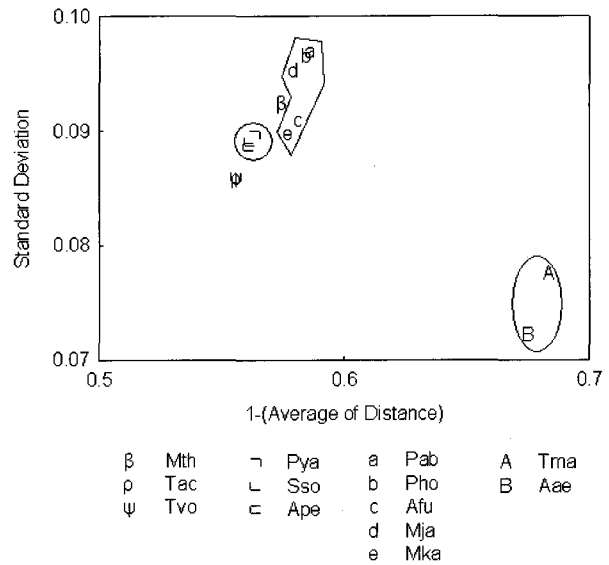
16S rRNA 유전자의 특정 부분은 진화 속도가 매우 느 리기 때문에, 세균들이 공통적으로 갖는 보존된 염기서열 과 이차구조를 보인다는 사실에 기초하여, 세균들을 중 수 준에서 구분할 수 있고 다양한 분류군의 상호 비교를 가 능하게 한다(17). 하지만 16S rRNA 유전자 비교는 단일 유 전자에 대한 염기서열의 유사성 (similarity) 비교라는 한계 가 있다(18). 이에 반해 COG 보유유무에 따른 계통수는 전체 계통을 고려한다는 점과 16S rRNA 유전자와 달리 보존성이 낮은 유전자까지 고려한다는 점에서 16S rRNA 의 단점을 보완할 수 있다(6). 각 계통에서 특정 COG의 보유유무에 따라서 binary sequence가 주어지므로 maximum parsimony와 bootstrap 같은 다양한 계통서열 분석법을 적용 할 수 있다(18). COG를 이용한 분류법은 세균들을 염기서 열 기반의 기능적 측면에서 분류하는 것으로, 전체 계통의 염기서열에 기초한 분류법과 직접 실험을 수행하는 생화 학적 분류법의 중간위치에 있는 분류법으로, 유용한 단백 질을 탐색하는 작업 등에 그 유용성이 높다고 할 수 있을 것이다(6).

Montague와 Hutchison(18)은 계통비교에 있어 COG 알고 리즘이 민감도 (sensitivity)와 설득력 (stringency) 면에서 우 수하다고 하였다. 수평적 유전자 전달 (lateral gene transfer) 에 의하여 진화적 계통수와 다른 계통수가 그려질 가능성 이 있지만, 수평적 유전자 전달은 전체 계통에서 차지하는 비율은 높지 않을 것이라는 것을 추정할 수 있으므로 이 를 이용하여 계통수를 작성하였고 neighbor-joining method 를 이용하였다(12).

**유전자의 변이**

하나의 유전자 정렬에서 얻어진 distance value로 유전체 의 특성을 표현하기는 어렵지만, 유전자 표본의 수가 많아

질수록 그 유전체의 특성을 가깝게 표현한다고 할 것이다 (19).



**Figure 4.** Distribution pattern of 13 genomes of thermophiles and hyperthermophiles by distance value. X-axis represents 1- (average of distance: sum of distance value for each COG was divided by 167) and Y-axis shows the standard deviation of distance value among 167 COGs of one genome. Species were grouped with their phylogenetic relatives excluding Mth. Alphabet in the figure represents thermophilic Euryarchaeota (β, ρ, ψ), hyperthermophilic Crenarchaeota (γ, λ, μ), hyperthermophilic Euryarchaeota (a-e), and hyperthermophilic Eubacteria(A, B).

Fig. 4는 분석대상 13종의 세균들이 공통적으로 함유하 는 167종의 공통유전자들이, 하나의 유전체내에서 보이는 distance value의 평균과 분산에 따른 분포이다. Distance value는 정규화된 유전자의 보존 정도로서 유전체 전체 수 준에서의 보존성의 정도를 나타내며, 특정 유전체의 distance value의 분산은 그 미생물이 가진 일부 유전자들의 유전자 변이(진화) 등을 설명하는 지표로 사용될 수 있다 (20). 고온성 진정세균 (Tma, Aae)은 고세균과는 완전히 독 립되어 있는 것으로 나타났으며, 고온성 Crenarchaeota (Pya, Sso, Ape)는 비슷한 평균값과 분산값을 갖는 것으로 나타났다. *Thermoplasma* 2종 (Tvo, Tac)의 경우 167 COG의 평균값이 높고 분산값은 다른 고세균에 비해 적은 것으로 나타났다. 높은 평균값은 각 유전자들의 변이가 심하다는 것으로 지구상에 출현한 지 오래되었다는 것이고, 분산값 이 낮다는 것은 167 COG들이 다른 고세균에 비해 고르게 변이를 했다고 할 수 있을 것이다(19). 고온성 고세균인 Mth는 다른 고온성 고세균인 Tvo, Tac보다는 Mka, Mja와 평균값이 유사한 것으로 나타나 유전자보유 계통수와 비 슷한 양상을 보였다. 이는 유전자 보유의 정성적 측면과 공통 COG의 정량적 측면 모두에서, 고온성 고세균보다는 메탄대사관련 초고온성 고세균에 더 가까운 것을 알 수 있었다.

## 요약

고온성 및 초고온성 세균과 고세균 13종 모두에서 관찰되는 167종류 총 16,299개의 보존적 유전자들에 대한 분석을 수행하였다. 단백질대사 관련 유전자들이 80개로 전체 보존적 유전자의 47.9%였으며, 중온성 세균을 제외하고 고온성과 초고온성 세균들에서만 관찰되는 공통유전자는 없어 열안정성은 특정 단백질의 유무에 따라 이루어지지 않는 것을 알 수 있었다. 하지만 초고온성 세균들은 reverse gyrase를 공통적으로 가지고 있어 고온에서의 DNA의 열안정성에 중요한 역할을 하는 것으로 생각되었다. 유전자보유 계통수와 16S rRNA 유전자 계통수의 비교결과 초고온성 진정세균과 고온성 고세균인 *Methanobacterium thermoautotrophicum*의 분포 양상이 서로 다르게 나타났다. 167개의 공통 유전자가 한 유전체에서 보이는 distance value들의 평균과 분산에서는 초고온성 진정세균, 초고온성 고세균, 고온성 고세균들끼리 유사한 값을 갖는 것으로 나타났다.

## 감사

본 연구는 부산대학교 김철민 선생님의 조언을 많이 받았으며, 해양수산부 마린바이오21사업 - 해양 극한생물 분자유전체연구단의 지원을 받아 수행되었습니다.

## REFERENCES

1. Radianingtyas, H. and C. Phillip (2003), Alcohol dehydrogenases from thermophilic and hyperthermophilic archaea and bacteria, *FEMS Microbiol. Rev.* **27**, 593-616.
2. Haki, G. D. and S. K. Rakshit (2003), Developments in industrially important thermostable enzymes: a review, *Biores. Technol.* **89**, 17-34.
3. Stetter, K. O. (1988), *Archaeoglobus fulgidus* gen. nov., sp. nov.: a new taxon of extremely thermophilic archaeobacteria. *Syst. Appl. Microbiol.* **10**, 172-173.
4. Tatusov, R. L., M. Y. Galperin, D. A. Natale, and E. V. Koonin (2000), The COG database, a tool for genome-scale analysis of protein functions and evolution, *Nucleic Acids Res.* **28**, 33-36.
5. Tatusov, R. L., E. V. Koonin, and D. L. Lipman (1997), A genomic perspective on protein families, *Science* **278**, 631-637.
6. Lee, D.-G., H.-Y. Kang, C.-M. Kim, S.-J. Kim, and J.-H. Lee (2002), Classification of Archaeobacteria and Bacteria using a gene content tree approach, *Korean J. Biotechnol. Bioeng.* **18**, 39-44.
7. Lee, D.-G., H.-Y. Kang, J.-H. Lee, and C.-M. Kim (2002), Detection of Conserved Genes in *Proteobacteria* by using a COG Algorithm, *Korean J. Biotechnol. Bioeng.* **17**, 560-565.
8. <http://www.ncbi.nlm.nih.gov/cog/new>
9. Ruiz, C., A. Blanco, F. I. J. Pastor, and P. Diaz (2002), Analysis of *Bacillus megaterium* lipolytic system and cloning of LipA, a novel subfamily I.4 bacterial lipase, *FEMS Microbiol. Lett.* **217**, 263-267.
10. Heath, C., A. C. Jeffries, D. W. Hough, and M. J. Danson (2004), Discovery of the catalytic function of a putative 2-oxoacid dehydrogenase multienzyme complex in the thermophilic archaeon *Thermoplasma acidophilum*, *FEBS Lett.* **577**, 523-527.
11. Fukuchi, S. and K. Nishikawa (2001), Protein surface amino acid compositions distinctively differ between thermophilic and mesophilic bacteria, *J. Mol. Biol.* **309**, 835-843.
12. <ftp://ftp.ncbi.nih.gov/genbank/genomes/Bacteria>
13. Amann, R., W. Ludwig, and K. H. Schleifer (1994), Identification of uncultured bacteria: a challenging task for molecular taxonomists. *ASM News* **60**, 360-365.
14. <http://www.ncbi.nlm.nih.gov>
15. Matoba, K., K. Mayanagi, S. Nakasu, A. Kikuchi, and K. Morikawa (2002), Three-dimensional electron microscopy of the reverse gyrase from *Sulfolobus tokodaii*, *Biochem. Biophys. Res. Co.* **297**, 749-755.
16. Jain, R., M. Rivera, and J. A. Lake (1999), Horizontal gene transfer among genomes : The complexity hypothesis, *Proc. Natl. Acad. Sci. USA.* **96**, 3801-3806.
17. Woese, C. R. (1987), Bacterial evolution, *Microbiol. Rev.* **51**, 221-271.
18. Montague, M. G. and C. A. Hutchison III (2000), Gene content phylogeny of herpesviruses, *Proc. Natl. Acad. Sci. USA.* **97**, 5334-5339.
19. Kang, H.-Y., C.-J. Shin, B.-C. Kang, J.-H. Park, D.-H. Shin, J.-H. Choi, H.-G. Cho, J.-H. Cha, D.-G. Lee, J.-H. Lee, H.-K. Park, and C.-M. Kim (2002), Investigation of Conserved Gene in Microbial Genomes using *in silico* Analysis, *Korean J. Life Sci.* **5**, 610-621.
20. Kimura, M. (1983), The neutral theory of molecular evolution, *Cambridge University Press*. London.