ISSN (Print) 1225-9918
ISSN (Online) 2287-3406
Journal of Life Science **2019** Vol. 29. No. 1. 123~128 DOI : https://doi.org/10.5352/JLS.2019.29.1.123

- Note -

Conserved Genes and Metabolic Pathways in Prokaryotes of the Same Genus

Dong-Geun Lee and Sang-Hyeon Lee*

Major in Pharmaceutical Engineering, Division of Bio-industry, College of Medical and Life Sciences, Silla University, Busan 617-736, Korea

Received August 14, 2018 / Revised September 28, 2018 / Accepted September 28, 2018

The use of 16S rDNA is commonplace in the determination of prokaryotic species. However, it has limitations, and there are few studies at the genus level. We investigated conserved genes and metabolic pathways at the genus level in 28 strains of 13 genera of prokaryotes using the COG database (conserved genes) and MetaCyc database (metabolic pathways). Conserved genes compared to total genes (core genome) at the genus level ranged from 27.62%(*Nostoc* genus) to 71.76%(*Spiribacter* genus), with an average of 46.72%. The lower ratio of core genome meant the higher ratio of peculiar genes of a prokaryote, namely specific biological activities or the habitat may be varied. The ratio of common metabolic pathways at the genus level was higher than the ratio of core genomes, from 58.79% (*Clostridium* genus) to 96.31%(*Mycoplasma* genus), with an average of 75.86%. When compared among other genera, members of the same genus were positioned in the closest nodes to each other. Interestingly, *Bacillus* and *Clostridium* genera were positioned in closer nodes than those of the other genera. Archaebacterial genera were grouped together in the ortholog and metabolic pathway nodes in a phylogenetic tree. The genera *Granulicella*, *Nostoc,* and *Bradyrhizobium* of the Acidobacteria, Cyanobacteria, and Proteobacteria phyla, respectively, were grouped in an ortholog content tree. The results of this study can be used for (i) the identification of common genes and metabolic pathways at each phylogenetic level and (ii) the improvement of strains through horizontal gene transfer or site-directed mutagenesis.

Key words : Conservative genes, conservative metabolic pathway, content phylogenetic tree, core-genome

서 론

종(species)은 분류의 기본 단위이며 새롭게 발견된 생물의 분류학적 위치 파악과 식의약 등의 분야에서 미생물을 GRAS (generally regarded as safe)로 인정할 때 사용된다. 원핵생물은 대개 무성생식을 하며 이분법으로 증식하므로, 종의 판단에 생식독립성이 아닌 다른 기준이 필요하다[13]. 전통적으로 원핵세포의 분류에는 세포모양과 크기, 편모 등의 형태학적 특성과 생화학적 시험이 사용되었다. 이를 위해 원핵생물은 순수분리 및 배양이 되어야 하는데 환경에 따라 최대 99% 정도의 원핵생물은 배양이 되지 않는다[8]. 또한 생화학적 시험에 영향을 미치는 유전자가 다른 종의 원핵생물로부터 전달되는 수평적 유전자 전달(horizontal gene transfer, HGT) [5, 12]에 의해 전통적 종의 개념에 혼란을 줄 수 있다.

염기서열분석 기술의 발달로 모든 원핵생물에 보존적인 16S rDNA 염기서열을 이용한 분류가 일반화되어 97% 이상의 유사도가 같은 종의 기준이 되었다[13]. 하지만 16S rDNA는 원핵생물의 게놈 하나에 1~15개가 분포하며 같은 게놈에서도 6.4%의 차이도 있고[14], 16S rRNA 유전자도 다른 종의 원핵생물로 수평전달되는 사례도 있어[1] 분류에서 16S rDNA의 절대성에 의문이 제기되었다. 이에 게놈에 1개 분포하면서 원핵생물들에 보존적인 GroEL chaperonin, rpoB, gyrB, dnaK 유전자 등이 대안으로 사용되고 있고, 하나의 유전자가 게놈 전체를 대표할 수 없다는 생각에 여러 보존적 유전자들을 활용하는 MLSA (multi-locus sequence analysis) 방법과 DNA-DNA hybridization 혹은 전체 게놈염기서열을 비교하여 70% 이상의 유사도를 나타내면 같은 종으로 인정하자는 제안이 있었다[13].

한편 동일한 공통조상 유전자에서 유래하여 생물종들에 분포하는 보존 유전자(conservative gene)들을 ortholog라 하고, 3가지 이상의 생물 종에 분포하는 ortholog에서 유래한 단백질들을 COG (Clusters of Orthologous Group of proteins)라고 하며, 현재 711개의 원핵생물의 COG들이 있다[4]. COG 방법으로 생물 종들 사이의 보존 유전자를 파악할 수 있고, 배양과 생화학적 실험 없이 원핵생물의 기능 유추가 가능하다[10, 11]. 생합성과 분해 등의 생물체의 물질대사는 개별 유전자가 아닌 유전자들에 기반한 효소들의 대사경로인데, COG 방법은 개별 보존유전자만 파악할 수 있다. MetaCyc 대사경로 database는 실험자료들을 토대로 구축되어, 2018년 현재 2,941

***Corresponding author**
Tel : +82-51-999-5624, Fax : +82-51-999-5628
E-mail : slee@silla.ac.kr

This is an Open-Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (http://creativecommons.org/licenses/by-nc/3.0) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

개의 생물에서 유래한 2,642개의 대사경로를 갖고 있다[3].

원핵생물을 속(genus) 수준에서 연구한 사례가 많지 않고 최근 *Enterococcus* [15], *Novosphingobium* [5], *Anaerobacillus* [2] 등에서 같은 속에 속하는 균주들의 게놈분석을 통해 분류 및 기능 측면에서 같은 속의 공통 유전자와 각 구성원의 개별 유전자를 비교하고 유전자의 수평전달[5] 등을 보고하였지만, 이들은 개별 속만을 분석하는 한계가 있었다. 이에 본 연구에서는 게놈분석이 완료된 13개 속의 28개 원핵생물의 보존적 유전자와 대사경로를 속 수준에서 비교하고자 하였다.

재료 및 방법

분석대상 원핵생물

본 연구에서 분석한 28개 원핵생물을 Table 1에 나타내었다. 보존유전자와 대사경로 파악을 위해 COG database [4]와 MetaCyc database [3]에 공통적인 원핵생물 중에서 속이 같고 종이 다른 28개 원핵생물을 선택하였다. 총 13개 속에서 10개 속은 진정세균이며, *Thermoplasma*, *Pyrococcus*, *Sulfolobus* 속은 고세균이다.

속 공통 및 특이 보존 유전자와 대사경로

'속 수준에서 공통 및 특이한 보존 유전자'를 파악하기 위해 MS사의 엑셀 프로그램(ver. 10)으로 분석대상 각 원핵생물이 전체 4,631개의 각각의 COG를 보유하고 있는 지를 행렬로 작성하였다[10]. 각 속의 원핵생물들에 공통인 COG를 통해 공통 보존 유전자를 탐색하였고, 각 속 보존 유전자들을 비교

Table 1. Studied prokaryotes and their numbers (#) and percentage (%) of total genes, conservative genes (orthologs) and metabolic pathways at each prokaryote and genus

Genus	Strains	Abbre-viation	Each strains						Each genus	
			Total genes		Orthologs		Metabolic pathways		Orthologs	Metabolic pathways
			# (A)	% (B/A)	# (B)	% (D/B)	# (C)	% (E/C)	Common (D) & Specific	Common (E) & Specific
Bacillus	*B. anthracis* Ames	Bacant	5,747	31.62	1,817	73.97	243	82.72		
	B. halodurans C-125	Bachal	4,058	43.99	1,785	75.29	226	88.94	1,344	201
	B. subtilis str. 168	Bacsub	4,381	41.34	1,811	74.21	246	81.71	& 53	& 10
	B. thuringiensis HD73	Bacthu	5,941	31.61	1,878	71.57	207	97.10		
Bradyrhizobium	*B. japonicum* USDA 6	Brajap	8,392	26.45	2,220	89.86	256	78.13	1,997	200
	B. oligotrophicum S58	Agroli	7,123	30.65	2,183	91.48	244	81.97	& 287	& 19
Clostridium	*C. acetobutylicum* ATCC 824	Cloace	3,885	41.57	1,615	76.97	238	49.16	1,243	117
	C. botulinum ATCC 3502	Clobut	3,702	41.79	1,547	80.35	171	68.42	& 66	& 17
Granulicella	*G. tundricola* MP5ACTX9	Acimp5	4,594	36.77	1,689	85.61	246	52.03	1,446	128
	G. mallensis MP5ACTX8	Gramal	4,833	34.41	1,663	86.95	184	69.57	& 91	& 15
Mycobacterium	*M. leprae* TN	Myclep	1,655	58.37	966	94.72	139	75.54	915	105
	M. tuberculosis H37Rv	Myctub	4,071	36.60	1,490	61.41	214	49.07	& 11	& 7
Mycoplasma	*M. genitalium* G37	Mycgen	557	65.89	367	98.37	41	95.12	361	39
	M. pneumoniae M129	Mycpne	1,046	38.15	399	90.48	40	97.50	& 1	& 2
Nostoc	*Nostoc* sp. PCC 7120	Nos712	6,212	29.93	1,859	87.79	251	58.57	1,632	147
	N. punctiforme PCC 73102	Nospun	7,074	25.32	1,791	91.12	249	59.04	& 144	& 37
Wolbachia	*W. endosymbiont*	Wolcul	1,310	52.37	686	95.19	102	69.61	653	71
	Wolbachia sp. wRi	Wolwri	1,287	53.61	690	94.64	73	97.26	& 21	& 2
Spiribacter	*S. salinus* M19-40	Ectm19	1,721	73.79	1,270	92.13	156	78.21	1,170	122
	Spiribacter sp. UAH-SP71	Spiuah	1,880	69.73	1,311	89.24	136	89.71	& 69	& 8
Streptococcus	*S. pneumoniae* TIGR4	Strpne	2,165	52.93	1,146	79.14	137	55.47	907	76
	S. pyogenes M1 GAS	Strpyo	1,801	58.97	1,062	85.40	110	69.09	& 46	& 3
Pyrococcus	*P. abyssi GE5*	Pyraby	1,933	56.39	1,090	92.94	73	94.52	1,013	69
	P. furiosus DSM 3638	Pyrfur	2,047	55.11	1,128	89.80	82	84.15	& 102	& 5
Sulfolobus	*S. acidocaldarius* DSM 639	Sulaci	2,296	46.60	1,070	92.90	85	75.29	994	64
	S. solfataricus P2	Sulsol	2,945	38.47	1,133	87.73	77	83.12	& 56	& 5
Thermoplasma	*T. acidophilum* DSM 1728	Theaci	1,571	58.31	916	94.98	65	84.62	870	55
	T. volcanium GSS1	Thevol	1,596	58.33	931	93.45	67	82.09	& 23	& 5

하여 속 특이 유전자의 파악을 시도하였다. MetaCyc database에서[3] 확보한 자료들로 분석대상 원핵생물들의 대사경로를 비교하였다. 전체 대사경로가 2,642개라는 것 이외에는 '속 수준에서 공통 및 특이한 보존 유전자'와 동일하게 처리하였다.

계통수

보존 유전자(ortholog) 및 대사경로(metabolic pathway) 보유 계통수는 행렬로 작성된 분석대상 원핵생물들의 COG 및 대사경로 자료들을 Mega 프로그램(ver. 7)의 phylogeny analysis를 이용하여 NJ (neighbor joining), ML (maximum likelihood), UPGMA (unweighted pair group method with arithmetic mean), ME (minimum evolution), MP (maximum parsimony) 방법으로 Presence-Absence phylogenetic tree를 작성하면서 bootstrap method (n=1,000)를 적용하였다[11]. 분석대상 원핵생물들의 16S rRNA 유전자 염기서열을 GenBank나 rdp II (https://rdp.cme.msu.edu/)에서 구한 후 Mega 프로그램으로 계통수를 구하였다(bootstrap method, n=1,000).

결과 및 고찰

속 공통 보존 유전자

각 원핵생물의 전체 및 보존 유전자 개수, 속에 공통(common) 및 특이한(specific) 보존 유전자와 대사경로의 개수 등을 Table 1에 나타내었다. 보존 유전자인 ortholog는 3개 이상의 생물에서 발견되는 COG를 기준으로 하였다. 각 속에 공통인 보존 유전자는 361~1,997개 범위였다. 원핵생물마다 유전자 개수가 달라 각 원핵생물 구성원의 보존 유전자와 속 공통 보존 유전자의 비율로 비교하면(Table 1의 D/B) 최소는 *Mycobacterium tuberculosis* H37Rv의 61.41%, 최대는 *Mycoplasma genitalium* G37의 98.37%로 평균 86.35%였다. Qin 등[12]은 전체 단백질의 50% 이상이 공통이면 같은 속의 원핵생물로 분류하였다. 이[11]는 같은 속 세균 4종이 전체 단백질의 29.38~46.93%가 보존적인 것을 밝혀 Qin 등[12]과 차이를 보여, 원핵생물의 분류체계가 완전하지 않은 것으로 해석할 수 있었다.

각 원핵생물과 속 수준의 공통 보존 유전자의 비율(Table 1의 D/B)을 평균한 속 수준의 범위는, *Bacillus* 속의 최저 73.76%에서 *Wolbachia* 속의 최대 94.91%였다. *Wolbachia* 속의 원핵생물들은 그람음성으로 곤충 등의 절지동물 세포내에 존재하는 절대 세포내 세균들이다[8]. 기타 보존적 유전자 비율이 94% 이상인 속은 세포내에서 기생하는 *Mycoplasma* (94.42%) 속과 통성혐기성 고세균으로 산성과 고온에서 생존하며 황과 유기산을 이용하는 *Thermoplasma* (94.21%) 속이었다. 속 수준에서 공통 보존유전자 비율이 높은 속들은 모두 서식환경이 숙주세포의 내부, 고온 및 산성 환경 등으로 다른 속에 비해 비교적 일정하다. 세포 내부 등의 일정한 환경에서 서식하면 적은 수의 유전자로도 생존이 가능하며[11], 공통조상과 서식환경이 유사하면 자손인 종 사이의 유전자 종류의 변화는 적었을 것이다[6].

각 원핵생물 특이 유전자 수

보존 유전자인 ortholog는 3종류 이상의 생물에 존재하는데, 본 연구에서는 *Bacillus* 속을 제외하고는 모든 속의 구성원이 2개라서 하나의 속에서만 특이한 보존 유전자를 구하기는 어려웠다. 각 원핵생물의 전체 유전자수 대비 보존 유전자의 비율(Table 1의 B/A)은 최저 *Nostoc punctiforme* PCC 73102의 25.32%에서 최대 *Spiribacter salinus* M19-40의 73.79% 범위였다. 보존 유전자는 3종류 이상의 생물에 분포하므로[4], 전체 유전자에서 보존 유전자의 비율이 낮다는 것은 1개 혹은 2개의 원핵생물에 분포하는 유전자의 비율이 높은 것이다. 공통이 아닌 유전자가 많은 원핵생물은 다양한 생합성 및 물질분해 효소, 저항 단백질 등의 생산이 가능하여 속이 여러 환경에서 생존할 수 있을 것이다[7]. 하지만 아직 모든 유전자 산물에 대한 기능이 밝혀지지 않아서 더 이상의 비교는 할 수 없었다.

같은 속인 구성원들의 공통 유전자들을 core-genome으로, 하나의 게놈에서 core-genome 이외의 유전자들을 shell-genome으로 나누었는데[5, 15], 본 연구에서는 속에 공통적인 orthologs를 core-genome으로 간주하였다. 각 원핵생물의 전체 유전자 개수와 core-genome 비율은 최저 *Mycobacterium tuberculosis* H37Rv의 22.48%에서 *Spiribacter salinus* M19-40의 67.98% 범위였다. 속 수준에서는 27.62%(*Nostoc* 속)~71.76%(*Spiribacter* 속) 범위였다. 게놈에서 core-genome 비율이 높은 원핵생물들은 각 원핵생물의 특이 유전자 비율이 낮다는 것으로, 종분화가 오래되지 않았거나 서식지가 공통조상과 비슷하다고 할 수 있었다[5, 15].

속 특이 보존 유전자

속 특이 보존 유전자가 많다면 다른 속과 차별되는 생명활동을 보일 것이며 생태학적으로 다른 속이 생존하기 힘든 서식지에서 생활할 수 있을 것이다[5, 6, 15]. COG는 3개 이상의 생물에서 발견되는 보존 유전자의 산물인데, 분석대상 속들의 구성원은 *Bacillus* 속의 4개를 제외하고는 모두 2개로 *Bacillus* 속을 제외하면 속 특이 보존 유전자를 파악하는 것은 불가능하였다. 13개 속에서 *Bacillus* 속 특이 보존 유전자로 확인된 53개도 COG database의 원핵생물 711개와 비교하니, 원핵생물 20~223개와 공통이어서 *Bacillus* 속 특이 보존 유전자는 없었다[자료 미제시].

연구대상인 13개 속 범위에서 특정 속에만 존재하는 속 특이 보존 유전자의 개수는 *Mycoplasma* 속의 최소 1개에서 *Bradyrhizobium* 속의 최대 287개로 평균 74.6개였다(Table 1). *Mycoplasma* 속 특이 보존 유전자인 COG1428은 deoxyadenosine/deoxycytidine kinase로 COG database의 원핵생물 142개에서 발견되었고, *Bradyrhizobium* 속의 보존 유전자 287

개 모두는 COG database 전체에서는 3개 이상의 속에 존재하였다. 따라서 *Bradyrhizobium* 속의 원핵생물이 추가되어도 *Bradyrhizobium* 속 특이 보존 유전자는 없을 것이다[자료 미제시].

속 공통 대사경로

Table 1을 보면 각 원핵생물의 대사경로의 개수는 *M. pneumoniae*의 최소 40개에서 *B. japonicum*의 최대 256개로 평균 156개였다. 각 속의 공통 대사경로 개수는 *Mycoplasma* 속의 최저 39개에서 *Bacillus* 속의 최대 201개로 평균 107개였다. 이[11]는 대사경로 개수가 작은 원핵생물은 합성과 분해할 수 있는 물질의 종류가 적고, 자유생활이 아닌 기생이나 공생생활과 연관이 있을 것으로 유추하였다.

각 속과 각 원핵생물의 공통 대사경로와 개별 대사경로의 개수로 비율(Table 1의 E/C)을 구하고 평균을 계산한, 속 수준의 공통 대사경로 비율은 *Clostridium* 속의 최저 58.79%에서 *Mycoplasma* 속의 최대 96.31%로 평균 75.86%였다. 공통 대사경로 비율이 높은 속은 순위대로 *Pyrococcus* (89.33%), *Bacillus* (87.62%), *Spiribacter* (83.96%), *Wolbachia* (83.43%), *Thermoplasma* (83.35%) 속이었다. 속 수준에서 전체 유전자 수 대비 core-genome의 비율은 27.62~71.76%(평균 46.72%) 범위였는데, 대사경로는 58.79~96.31%(평균 75.86%) 범위였다. 유전자의 수평전달[12], 중첩, 돌연변이 등으로 개별 게놈에 특이한 유전자의 수가 많아졌지만[6] 대사과정에 참여하지 못할 가능성 등이 원인으로 사료되었다.

속 특이 대사경로

각 속의 구성원들만 나타내는 속 특이 대사경로의 개수는 최소 2개(*Mycoplasma* 속과 *Wolbachia* 속)에서 최대 37개(*Nostoc* 속)였다(Table 1). 비교대상을 늘려 각 속 특이 대사경로를 가진 다른 속의 원핵생물을 확인하였다. *Pyrococcus* 속에만 발견될 것 같은 glycolysis V (*Pyrococcus*) 대사경로가 Crenarchaeota와 Euryarchaeota에서도 발견되었다. *Mycobacterium* 속 특이 대사경로로 추정되는 mycolate 생합성 경로가 *Mycobacterium* 속과 Actinobacteria 문(phylum)의 *Amycolicicoccus subflavus* DQS3-9A1에서 발견되고, dimycocerosyl phthiocerol 생합성 경로는 Mycobacteriaceae와 Nocardiaceae 과(family)에도 존재하는 등 속 특이적 대사경로는 없었다. *Novosphingobium* 속에 속하는 원핵생물 6개의 게놈 비교에서도 속 특이 유전자와 대사경로를 찾지 못했다[5].

보존 유전자 보유 계통수

Fig. 1에 분석대상 원핵생물들의 보존 유전자 보유 계통수를 나타내었다. ML, MP, ME, NJ, UPGMA 계통수 모두에서 발견되는 공통점은 다음과 같았다. 첫째, 각 속의 구성원은 모든 계통수에서 가장 인접하였는데 *Mycobacterium* (Fig. 1의 Myclep, Myctub) 속은 UPGMA, MP 계통수에서만 가장 인접

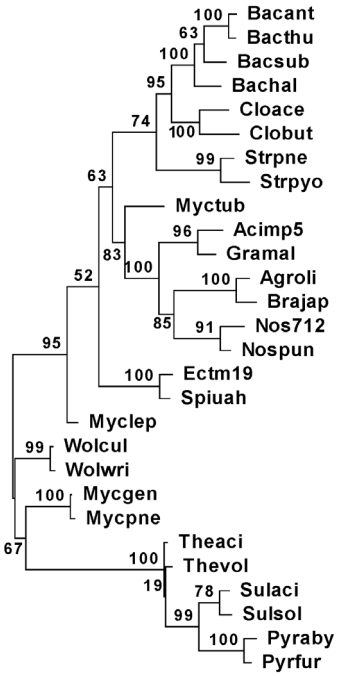

Fig. 1. ML(Maximum Likelihood) phylogenetic tree of 28 procaryotes in the point of presence or absence of the union of 4,631 COGs. Bootstrap values at each node are expressed as a percentage of 1,000 trials. See table 1 for abbreviations of prokaryotes.

하였다(자료 미제시). 둘째, 고세균인 *Thermoplasma* (Theaci, Thevol), *Sulfolobus* (Sulaci, Sulsol), *Pyrococcus* (Pyraby, Pyrfur) 속들은 하나의 그룹을 형성하였다. 셋째, *Bacillus* (Bacant, Bachal, Bacsub, Bacthu) 속과 *Clostridium* (Cloace, Clobut) 속이 가장 인접하였다. 이러한 점들은 16S rDNA를 이용한 계통수와 차이가 없거나 적다고 할 수 있다. 넷째, *Granulicella* (Acimp5, Gramal), *Nostoc* (Nos712, Nospun), *Bradyrhizobium* (Brajap, Agroli) 등 3개 속이 하나의 그룹이었는데 이들은 각각 Acidobacteria, Cyanobacteria, Proteobacteria 문으로 16S rDNA에서는 그룹이 되지 않는다. 본 연구의 *Bradyrhizobium*, *Spiribacter* (Ectm19, Spiuah), *Wolbachia* (Wolcul, Wolwri) 속이 Proteobacteria 문이다. 16S rDNA 분석은 미지의 원핵생물을 속 혹은 과로 분류하는데 사용하고, MLSA로 종이나 균주를 분류하자는 제안도 있었다[13].

대사경로 보유 계통수

Fig. 2에 분석대상 원핵생물들의 대사경로 보유 계통수를 나타내었다. ML, MP, ME, NJ, UPGMA 계통수 모두에서 발견되는 공통점은 다음과 같았다. 첫째, *Granulicella* (Fig. 2의 Acimp5, Gramal) 속과 *Nostoc* (Nos712, Nospun) 속을 제외한 모든 속의 구성원들은 가장 인접하였다. 둘째, *Granulicella* 속과 *Nostoc* 속의 구성원들은 항상 인접하지 않았고, *Granulicella* 속의 Acimp5와 *Nostoc* 속의 Nos712가 인접하였고 여기에

Nostoc 속의 Nospun이 결합하였다. 이들은 종이 달라서 core-genome은 다르더라도 다윈의 자연선택처럼[6] 서식지가 비슷하면 공통 유전자와[5, 15] 공통 대사경로를 보유할 수 있었을 것이다. 셋째, 고세균의 3개 속은 하나의 그룹을 형성하였다. 넷째, *Mycoplasma* (Mycgen, Mycpne)속과 *Wolbachia* (Wolcul, Wolwri) 속은 고세균에 인접하였다. 다섯째, *Bacillus* (Bacant, Bachal, Bacsub, Bacthu)와 *Clostridium* (Cloace, Clobut) 속은 서로 인접하였다. 대사경로 보유 계통수가 보존 유전자 보유 계통수와 대체로 일치하는 것은 첫째, 넷째, 다섯째이고 차이를 보이는 것은 둘째였고 넷째는 계통수의 종류에 따라 일치하거나 차이를 보였다.

16S rDNA 계통수와 비교

Fig. 3에 분석대상 원핵생물들의 16S rDNA의 염기서열로 작성한 계통수를 나타내었다. 보존 유전자 및 대사경로 보유 계통수(Fig. 1, Fig. 2) 그리고 16S rDNA 계통수(Fig. 3) 사이에 모두 일치하는 것은 고세균이 하나의 그룹을 형성한 것이다. 보존 유전자 및 대사경로 보유 계통수에서는 공통적인데 16S rDNA 계통수에서 차이를 보인 것은 첫째, *Mycoplasma* (Mycgen, Mycpne) 속과 *Wolbachia* (Wolcul, Wolwri) 속이 고세균과 떨어진 것이고, 둘째 *Bacillus* 속과 *Clostridium* 속이 인접하지 않은 것이다. 보존 유전자 보유 계통수와 16S rRNA 유전자 계통수에서 일치하는 것은 각 속의 구성원은 모든 계통수에서 가장 인접한 것이다. 그런데 보존 유전자 보유 계통수에서는 *Mycobacterium* (Myclep, Myctub) 속이 UPGMA, MP 계통수에서만 가장 인접하였는데, 16S rDNA 유전자 계통수에서는 ML, MP, ME, NJ, UPGMA 계통수 모두에서 가장 인접하였다. 전반적으로는 세 계통수에서 각 속의 구성원은 인접하는 경향이 있었고 세 계통수 사이에서는 대사경로 보유 계통수가 다른 계통수들과 차이를 많이 보였다고 할 수 있었다.

16S rDNA 계통수는 16S rRNA 유전자 하나로만 생물의 유연관계를 파악하고 원핵생물의 기능에 관여하는 다른 보존적 유전자들이나 대사경로의 공통점과 차이점은 알 수 없고 16S rRNA 유전자도 다른 종으로 수평전달되는 사례도 있다[1]. 따라서 본 연구처럼 보존 유전자 및 대사경로의 비교 그리고 이들을 이용한 계통수는 기능을 비교할 수 있어 16S rDNA 분석과 상호보완적일 수 있을 것이다.

연구결과 응용 가능성

원핵생물들의 생명현상은 학술적, 생태적, 경제적 측면 등에서 중요하다. 본 연구의 결과는 원핵생물의 환경에 대한 적응[5], 원핵생물 사이의 신호 및 유전자 전달[5], 오염된 환경의 정화, 수평적 유전자 전달[12]이나 부위특이적 돌연변이를 통

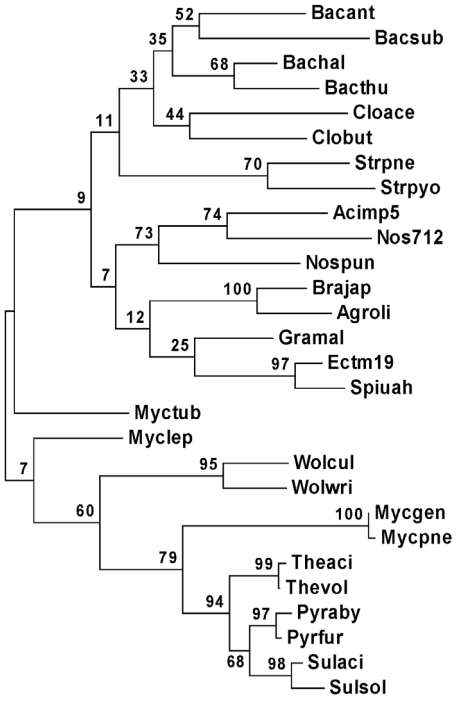

Fig. 2. ML (Maximum Likelihood) phylogenetic tree of 28 procaryotes in the point of presence or absence of the union of 2,526 metabolic pathways. Bootstrap values at each node are expressed as a percentage of 1,000 trials. See Table 1 for abbreviations of prokaryotes.

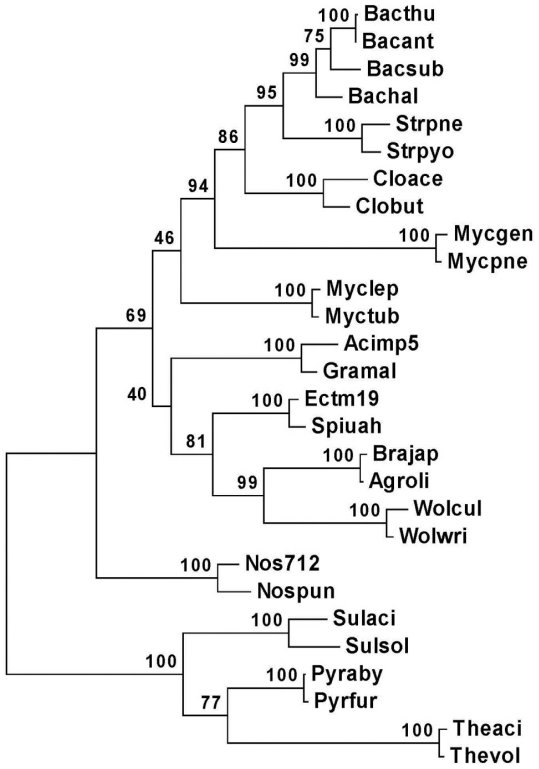

Fig. 3. ML (Maximum Likelihood) phylogenetic tree of 28 procaryotes in the point of 16S rRNA genes. Bootstrap values at each node are expressed as a percentage of 1,000 trials. See Table 1 for abbreviations of prokaryotes.

한 균주 개량[9]으로 생리활성 물질이나 항생물질을 대량생산하는 분야 등에서 기초자료로 활용될 수 있을 것이다.

References

1. Asai, T., Zaporojets, D., Squires, C. and Squires, C. L. 1999. An *Escherichia coli* strain with all chromosomal rRNA operons inactivated: complete exchange of rRNA genes between bacteria. *Proc. Natl. Acad. Sci. USA.* **96**, 1971-1976.
2. Bassil, N. M. and Lloyd, J. R. 2017. Draft genome sequences of four alkaliphilic bacteria belonging to the *Anaerobacillus* genus. *Genome Announc.* **5**, e01493-16.
3. Caspi, R., Altman, T., Dreher, K., Fulcher, C. A., Subhraveti, P., Keseler, I., Kothari, A., Krummenacker, M., Latendresse, M., Mueller, L. A., Ong, Q., Paley, S., Pujar, A., Shearer, A. G., Travers, M., Weerasinghe, D., Zhang, P. and Karp, P. D. 2012. The MetaCyc database of metabolic pathways and enzymes and the BioCyc collection of pathway/genome databases. *Nuc. Acids Res.* **40**, D742-D753.
4. Galperin, M. Y., Makarova, K. S., Wolf, Y. I. and Koonin, E. V. 2015. Expanded microbial genome coverage and improved protein family annotation in the COG database. *Nucleic Acids Res.* **43**, D261-D269.
5. Gan, H. M., Hudson, A. O., Rahman, A. Y., Chan, K. G. and Savka, M. A. 2013. Comparative genomic analysis of six bacteria belonging to the genus *Novosphingobium*: insights into marine adaptation, cell-cell signaling and bioremediation. *BMC Genomics* **28**, 431.
6. Gupta, R. S. 2016. Impact of genomics on the understanding of microbial evolution and classification: the importance of Darwin's views on classification. *FEMS Microbiol. Rev.* **40**, 520-553.
7. Han, K., Li, Z. F., Peng, R., Zhu, L. P., Zhou, T., Wang, L. G., Li, S. G., Zhang, X. B., Hu, W., Wu, Z. H., Qin, N. and Li, Y. Z. 2013. Extraordinary expansion of a *Sorangium cellulosum* genome from an alkaline milieu. *Sci. Rep.* **3**, 2101.
8. Johnson, K. N. 2015. Bacteria and antiviral immunity in insects. *Curr. Opin. Insect Sci.* **8**, 97-103.
9. Klein-Marcuschamer, D., Santos, C. N., Yu, H. and Stephanopoulos, G. 2009. Mutagenesis of the bacterial RNA polymerase alpha subunit for improvement of complex phenotypes. *Appl. Environ. Microbiol.* **75**, 2705-2711.
10. Lee, D. G. and Lee, S. H. 2015. Investigation of conservative genes in 711 prokaryotes. *J. Life Sci.* **25**, 1007-1013.
11. Lee, D. G. 2018. Comparison of metabolic pathways of less orthologous prokaryotes than *Mycoplasma genitalium*. *J. Life Sci.* **28**, 369-375.
12. Qin, Q. L., Xie, B. B., Zhang, X. Y., Chen, X. L., Zhou, B. C., Zhou, J., Oren, A. and Zhang, Y. Z. 2014. A proposed genus boundary for the prokaryotes based on genomic insights. *J. Bacteriol.* **196**, 2210-2215.
13. Rajendhran, J. and Gunasekaran, P. 2011. Microbial phylogeny and diversity : Small subunit ribosomal RNA sequence analysis and beyond. *Microbiol. Res.* **166**, 99-110.
14. Wang,Y., Zhang, Z. and Ramanan, N. 1997. Theactinomycete *Thermobispora bispora* contains two distinct types of transcriptionally active 16S rRNA genes. *J. Bacteriol.* **179**, 3270-3276.
15. Zhong, Z., Zhang, W., Song, Y., Liu, W., Xu, H., Xi, X., Menghe, B., Zhang, H. and Sun, Z. 2017. Comparative genomic analysis of the genus *Enterococcus*. *Microbiol. Res.* **166**, 95-105.

초록 : 동일한 속 원핵생물들의 보존 유전자와 대사경로

이동근 · 이상현*
(신라대학교 바이오산업학부 제약공학전공)

원핵생물 분류의 기본단위인 종(species)의 동정에 16S rDNA가 사용되지만 한계가 있고 원핵생물의 속(genus)에 대한 연구가 많지 않다. 본 연구에서는 보존 유전자를 확보한 COG database와 대사경로를 확보한 MetaCyc database에 공통적인 원핵생물 중 속이 같고 종이 다른 13개 속 28개의 원핵생물을 대상으로 속 수준에서 연구하였다. 전체 유전자에서 core-genome인 속 보존 유전자의 비율은 최저 27.62%(*Nostoc* 속)에서 71.76%(*Spiribacter* 속)의 범위로 평균 46.72%였다. 각 원핵생물에서 core-genome의 비율이 낮으면 특이한 생명현상을 보이거나 서식지가 다양할 수 있을 것이다. 속 수준의 공통 대사경로의 비율은 최저 58.79%(*Clostridium* 속)에서 최대 96.31%(*Mycoplasma* 속), 평균 75.86%로 core-genome의 비율보다 높았다. 비교대상을 확장하면 속 특이 보존 유전자와 대사경로는 확인할 수 없었다. 보존 유전자와 대사경로 보유 계통수에서는 대체로 같은 속의 구성원들이 가장 인접하였으며, *Bacillus*속과 *Clostridium* 속이 그룹을 형성하였고, 고세균끼리 그룹을 형성하였다. 보존 유전자 보유 계통수에서는 Acidobacteria, Cyanobacteria, Proteobacteria 문(phylum)의 *Granulicella*, *Nostoc*, *Bradyrhizobium* 의3개 속이 하나의 그룹을 형성하였다. 본 연구 결과는 (i) 각 계통 단계에서 보존유전자와 대사경로의 확인, (ii) 수평적 유전자 전달 또는 부위 지정 돌연변이를 통한 균주의 개선 등의 분야에 기초자료로 활용될 수 있을 것이다.