

해양유래 초고온성 고세균 *Thermococcus* sp. NA1 유전체 해독과 활용

이 정 현(한국해양연구원, 해양생물자원연구본부, 해양바이오신소재연구단)

인간유전체 해독의 선행 연구로 병원성 미생물 해모필러스(*Haemophilus influenza*)의 유전체 해독이 1995년에 완성된 이래로, 완성되었거나 혹은 진행되는 미생물(원핵세포로 이루어진 단세포 생물을 지칭하며 세균(Bacteria)과 고세균(Archaea)을 포함함) 유전체 해독은 세계적으로 공식 확인된 것만 2006년 10월 현재, 1400 여건에 이른다(http://www.genomesonline.org/gold_statistics.htm). 국내의 경우, 2002년부터 10년간 계획으로 시작된 프론티어사업인 미생물유전체활용기술개발사업과 미생물연구관련 기관 및 연구자들이 *Vibrio*, *Zymomonas*, *Helicobacter*, *Corynebacterium*, *Manheimia*, *Phaenibacillus*, *Pedococcus*, *Streptomyces* 등 십 여종 이상의 균주를 대상으로 유전체 해독이 이루어졌다.

미생물유전체의 경우 다른 동·식물에 비해 유전체 크기가 작아 비교적 적은 비용으로 해독작업이 가능하며, 유전체 정보를 용이하게 분석하고 활용가치가 우수한 특징이 있다. 유전체 해독의 대상이 되는 미생물의 특성을 보면, 의학용 연구 목적인 병원성 미생물이 가장 많으며, 다음으로 산업미생물, 환경미생물, 농업용 미생물 그리고 진화연구 목적의 미생물 순이다. 최근에는 해양생물의 응용에 관한 관심이 높아지면서, 해양미생물의 유전체 해독 연구가 국내외에서 두드러지고 있다. 우리나라의 경우, 수산양식업에 큰 피해를 주는 적조 원인생물의 방제에 이용될 수 있는 해양미생물인 *Hahella chejuensis*의 유전체 해독이 최근에 완료되었다. 또한 2003년부터 시작된 해양수산부의 마린바

이오21사업의 해양극한생물분자유전체연구단이 수종의 해양미생물을 대상으로 유전체 해독 작업이 진행 중에 있다. 특히, 미국의 무어 재단(Moore Foundation; <http://www.moore.org/microgenome/>)은 2004년부터 해양 원핵미생물 150 여종의 유전체에 대해 염기서열을 결정하는 거대한 해양미생물 유전체 프로젝트를 시작하였다. 해양미생물 중에 대한 종합적인 유전체 해독을 통해 바다의 건강과 생산성을 유지하는데 핵심적인 역할을 담당하는 해양미생물의 기능을 체계적으로 밝히고자 하는 것이다. 즉, 자동화된 장비와 분자적 분석, 생명정보해석 기술의 발달로 해양생태계의 생물학적, 화학적 과정을 좀 더 근원적으로 이해하고자 하는 것이다.

해양환경의 특징 중 하나로 다양한 극한조건을 들 수 있다. 즉, 사해와 같은 호염성, 심해저의 높은 압력과 저온환경, 해저화산지역의 초고온 조건은 다양한 극한 미생물이 서식하기 좋은 환경을 제공한다. 생물분류 체계상 계(Kindom)보다 한단계상위의 도메인(domain) 수준에서 고세균류('Archaeobacteria'에서 'Archaea'로 수정) 도메인은 호염성, 메탄생성, 초고온성, 호산성 등의 극한환경 조건에 서식하는 미생물들을 포함하고 있다. 대부분의 고세균류는 혐기적 조건, 초고온성, 호산성 등의 극한환경 조건에서 서식하기 때문에 분리와 배양이 비교적 어렵다. 현재까지 29 종의 고세균류에 대한 유전체 해독이 완성되었다(표1). 세균류 383종의 유전체 해독이 완성된 것에 비하면 그 수는 적지만, 고세균류 미생물은 극한 조건에서의 생명현상 유지, 진화적 관계성 규명, 생명공학적 응용 측면에서 각광을 받고

Table 1. 유전체 해독이 완성된 고세균류와 그 특성

고 세 균	특 성
<i>Aeropyrum pemix</i>	The first aerobic member of the Archaea to be sequenced
<i>Archaeoglobus fulgidus</i>	Metabolises sulfur
<i>Haloarcula marismortui</i>	A Halophilic microorganism that thrives in extreme saline environments such as the Dead Sea
<i>Halobacterium salinarum</i>	Responsible for the bright pink or red appearance of the Dead Sea and other bodies of salt water
<i>Methanobacterium thermoautotrophicum</i>	Produces methane from CO ₂
<i>Methanocaldococcus jannaschii</i>	Cannot tolerate oxygen and is one of the methane-producing microbes
<i>Methanococcus maripaludis</i>	Converts organic waste material into methane and carbon dioxide
<i>Methanopyrus kandleri</i>	Produces methane and grows optimally at temperatures near and above the boiling point of water
<i>Methanosarcina acetivorans</i>	One of the most versatile methane-producing microbes
<i>Methanosarcina barkeri</i>	Causes methane gas production in cattle and contains a "new" amino acid
<i>Methanosarcina mazei</i>	Ferment acetate, methylamines and methanol to methane, carbon dioxide and ammonia
<i>Methanosphaera stadtmanae</i>	Generates methane only by reduction of methanol with H ₂
<i>Nanoarchaeum equitans</i>	Hhas one of the smallest genomes of any sequenced microbe
<i>Natronomonas pharaonis</i>	An extremely haloalkaliphilic archaeon that is studied for Nitrogen metabolism
<i>Picrophilus torridus</i>	One of the most thermoacidophilic organisms known
<i>Pyrobaculum aerophilum</i>	Useful for investigating the molecular basis of heat resistance
<i>Pyrococcus abyssi</i>	A hyperthermophile
<i>Pyrococcus furiosus</i>	Highly resistant to radiation and may possess an efficient system for repairing DNA
<i>Pyrococcus horikoshii</i>	A hyperthermophile that grows optimally at 98°C
<i>Sulfolobus acidocaldarius</i>	A model for research on mechanisms of DNA replication
<i>Sulfolobus solfataricus</i>	A model species for investigating DNA replication, the cell cycle, and RNA processing
<i>Sulfolobus tokodaii</i>	Convert hydrogen sulfide to sulfate and may have industrial applications, such as wastewater treatment
<i>Thermococcus kodakaraensis</i>	A hyperthermophilic archaeon with an optimum growth temperature of between 65 and 100°C
<i>Thermoplasma acidophilum</i>	Lives in a particularly harsh environment — hot and acidic — without the protection of a rigid cell wall
<i>Thermoplasma volcanium</i>	Able to live in the presence of oxygen as well as in environments lacking oxygen
<i>Methanosaeta thermophila</i> PT	Metabolize acetate into methane and it may be useful as a producer of biofuels
<i>Haloquadratum walsby</i> HBSQ001	Hyperhalophile of specific interest because of its unique square shape and extraordinarily large size
<i>Methanococcoides burtonii</i> DSM6242	Ability to produce methane, a potential source of fuel
<i>Methanosprillum hungatei</i> JF-1	Methanogenic archaeon that has been isolated from bovine rumens, anaerobic sludge digestors and mud

(<http://www.ebi.ac.uk/2can/genomes/archaea.html>)

있다. 2002년 한국해양연구원 해양바이오연구팀은 서태평양 심해열수 환경지역에서 고세균 *Thermococcus* sp. NA1종을 분리하여 분류작업을 통해 신종으로 밝혔고, 유전체 해독을 완성하여 분석 작업을 진행 중이다.

1. 미생물 분리 및 분류

서태평양 PACMANUS 열수구 지역의 퇴적물 시료

로부터 여러 번의 계대배양을 통해 순수 분리된 균주를 확보하였으며 NA1이라고 명명하였다. 80°C에서 2 일 간 배양된 NA1의 콜로니 형태는 원형이며, 크림색깔을 띄었다. NA1은 polar flagella를 이용하여 운동성을 가지며, 세포크기는 대략적으로 0.5 - 1.0 μm로 구형이다(그림 1). NA1 균주의 최적생육온도는 80°C(63-90°C, 생육 온도 범위), 최적생육 pH는 8.5(5.0-9.0, 생육 pH 범위)이며, 최적생육 NaCl 농도는 3.5%(1.0-5.0%, 생육 NaCl

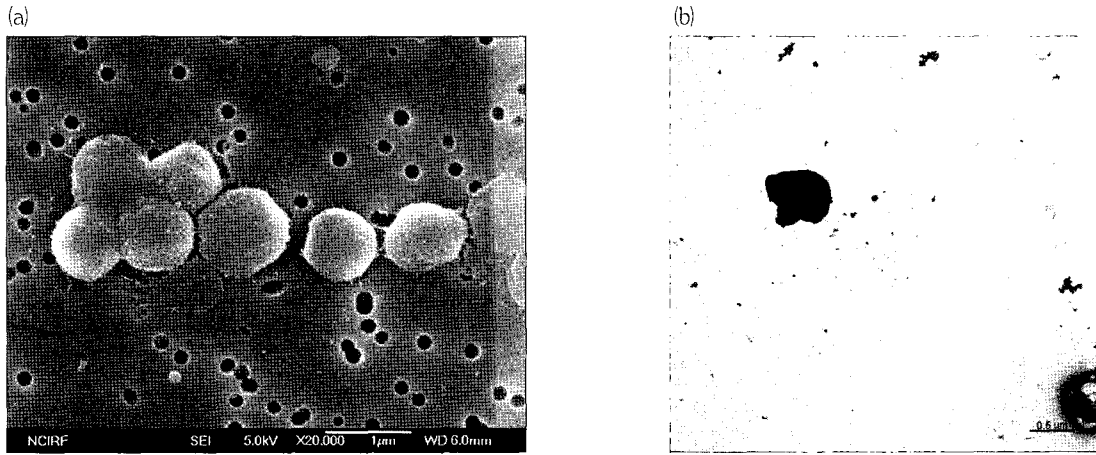
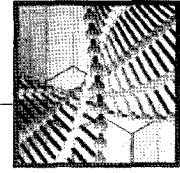


그림 1. 해양 초고온성 고세균 NA1의 전자현미경사진: (a), scanning electron micrograph; (b), transmission electron micrograph. Scale bar represents 1 μm in (a) and 0.5 μm in (b).

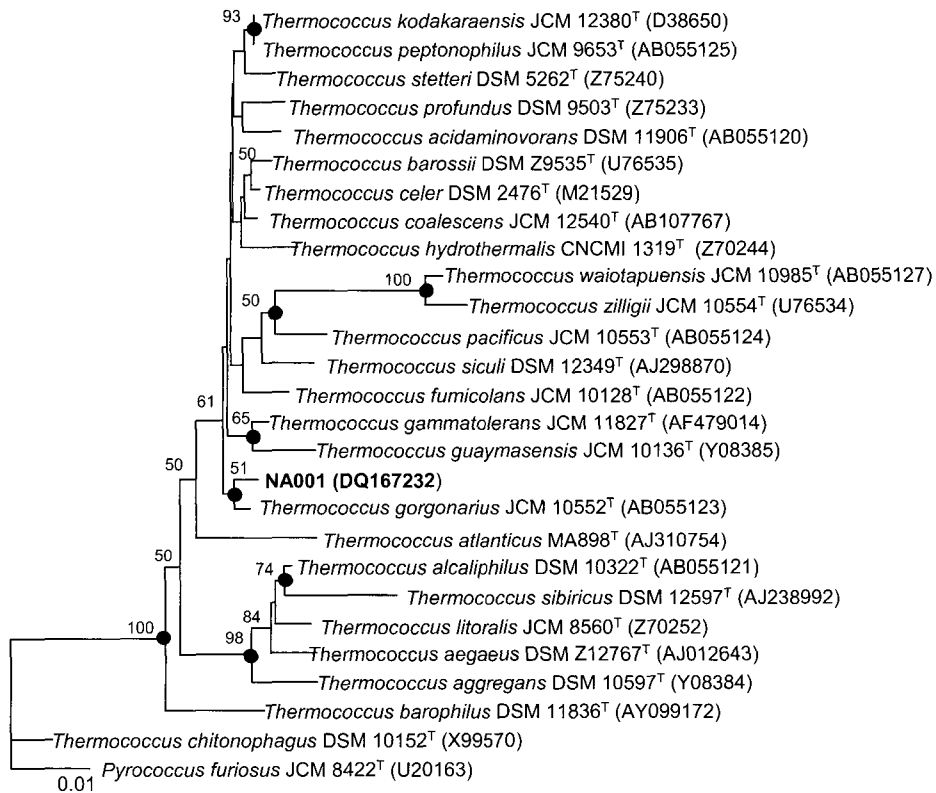


그림 2. NA1의 16S rRNA 유전자 염기서열에 기초한 *Thermococcus* 속의 계통도(Bootstrap percentage values higher than 50 % based on neighbor-joining analyses of 1000 replications were included at nodes, and solid circles indicate that corresponding nodes (groupings) were also recovered in Fitch-Margoliash, maximum-parsimony, and maximum-likelihood trees. Bar, 0.01 nucleotide substitution per position)

범위)이다. 이 균주는 종속영양세균으로 유일 탄소원으로 $H_2/CO_2(80:20, 200\text{ kPa})$ 혼합가스를 넣어준 무기영양배지에서는 성장하지 않는다. 균주가 성장하기 위해서는 elemental sulfur를 요구하며 전자수용체로 작용하여 황화수소로 환원된다. 분리된 NA1균주의 다른 전자수용체 사용유무를 확인하기 위해 elemental sulfur 대신 cystine(10%), polysulfide(10 mM), sodium thiosulfate(10 mM), sodium sulfate(20 mM), and sodium sulfite(3 mM)를 첨가한 배지에서는 성장하지 않았다. 또한, 20개의 아미노산을 0.2 mM의 농도로 첨가한 YPS 최소배지에서는 성장하지 않는다. NA1 균주는 chloramphenicol, ampicillin, kanamycin, vancomycin, streptomycin 그리고 rifampicin과 같은 항생제에 저항성이 있다. 기질요구도 조사에서 NA1은 beef extract, casein, peptone, tryptone, yeast extract, 그리고 starch와 같은 기질에서는 성장하지만 maltose, lactose, sucrose, cellobiose, xylose, gelatin, glycogen, casamino acids, acetate, succinate, propionate, 그리고 pyruvate에서는 성장하지 않는다. NA1의 16S rDNA 유전자 서열에서 총 1,457 bp 염기서열이 결정하였고 상동성 분석 결과 분리된 NA1 균주는 Thermococcus 속에 속하는 것으로 나타났다(그림 2). 16S rDNA 염기서열 계통분석을 한 결과 NA1은 T. gorgonarius (99.7%), T. kodakaraensis (99.5%), T. peptonophilus (99.5%) 그리고 T. celer (99.4%)와 높은 상동성을 나타내었지만 DNA-DNA hybridization 결과 genomic DNA의 상동성이 20%로 나타나 새로운 종임을 확인하였다.

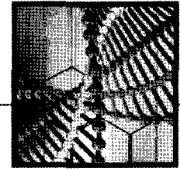
2. 초고온성 고세균 NA1의 유전체 해독

2kb shotgun library의 경우는 Nebulizer를 이용한 방법으로 진행하였다. 대규모 염기서열 분석 후 assembly 과정에서 repeat sequence, chimera sequence 등 문제점 해결을 위해 정확한 크기의 insert의 준비가 매우 중요하므로 최소한 2번 이상의 연속적인 gel purification

과정을 수행하였다. 이렇게 sizing되어진 insert들은 2kb의 경우 pCR4-Blunt TOPO Vector (Invitrogen사)에 cloning하였다. 두 개의 shotgun library를 제작하였으며 insert의 길이를 조사한 결과 2kb shotgun library는 각각 1.973kb (stdev=439)와 1.692kb (stdev=570) 이었다. 전기영동시 high molecular weight DNA를 40kb 정도로 분포하도록 random shearing하고 Pulse-Field Gel Electrophoresis를 이용하여 40kb insert를 분리하고 이를 이용하여 cosmid library를 제작하였다. 벡터로는 pWEB:TNC (Epicentre)를 사용하였다. Assembly 상의 cosmid의 길이는 37.691kb (stdev=3.5)이었다. DNA 추출은 반자동화 로봇 시스템을 이용하여 수행함으로써 수작업에 의한 샘플 quality 차이를 최소화하였다.

마크로젠이 보유한 3730 DNA analyzer를 이용하여 염기서열분석을 진행하였으며 base-caller로는 KB basecaller를 사용하였다. Phred score 20 이상을 기준으로 trim할 때 평균 838bp의 read length를 보이는 결과를 얻을 수 있었다. 전체적으로 22,000 여개의 분석된 염기서열 정보는 KB base-caller를 통하여 base calling을 수행하였으며, Phrap 프로그램을 이용하여 contig assembly를 수행하였다. Assembly 수행결과 총 146개의 contig가 생성이 되었으며 contig의 길이의 합은 1.857 kb이었다. 최종 finishing 단계에서는 새로운 시퀀싱 방법인 454 방법을 이용하여 20x 수준의 염기서열을 더 확보하였다.

생성된 contig 및 supercontig들의 양끝에서 Consed의 auto finish의 기능을 이용하여 design되어진 primer를 이용하여 genomic DNA를 template로 한 direct genome sequencing과 assembly상의 정보를 이용하여 gap을 가지고 있는 clone들을 선별하여 수행하는 clone walking 또는 gap size가 5kb 이상인 경우는 시간을 단축시키기 해당 clone에 대해 shotgun library를 제작하였으며, 다음 방법으로 scaffold map을 바탕으로 각각의 contig 양 말단에서 합성된 PCR primer를 디자인하여 합성한 뒤 PCR 과정을 거쳐 이웃한 두개



의 contig 순서를 확인하고 이들 PCR product의 염기서열 분석을 통하여 염기서열 분석이 이루어지지 않은 부분을 sequencing하였다. 한편 scaffold가 형성되지 않는 contig들에 대해서는 combination PCR에 의하여 ordering과 gap filling을 실시하였다. 이렇게 하여 얻어진 sequencing파일을 이용하여 reassembly하여 오류가 없음을 확인하였다. 이렇게 하여 형성되어진 circular single contig의 길이는 1,847,622 bases이며 51.85%의 GC content를 나타내었다. Quality가 Phred score 40이하인(99.99%)인 부분에 대하여 보완을 하기 위해 해당부분을 manual로 확인하였다.

3. 초고온성 고세균 NA1의 유전체 정보 분석

NA1 게놈(1,847,622 bp)로부터 Glimmer2.0 프로그램을 이용하여 ORF 2,039개를 예측하였다. 이 중에서

1-100개의 아미노산을 가진 단백질은 226개, 100-199 사이는 556개, 200-299는 512개, 300-399는 348개, 400-499는 215개, 500-599는 71개, 600-699는 54개, 700-799는 22개, 800-899는 13개, 900개 이상의 아미노산으로 이루어진 OF는 22개이었다. Uniprot과 NCBI 상동성 분석결과로부터 NA1 단백질의 상동성 검색 결과, 높은 상동성(E-value가 10^{-50} 이하, identity 50%이상, coverage 70%이상)을 나타내는 단백질 중, putative, predictive, hypothetical protein으로 annotation된 것은 putative 혹은 hypothetical 그룹으로 처리했다. 그리고 낮은 상동성(E-value가 10^{-10} 이하, identity 30-50%, coverage 70% 미만)을 나타내는 ORF들은 putative로 처리했다. 이러한 기준으로 known protein은 1,042개(51%), putative protein은 318개(16%), hypothetical protein은 679개(33%)이었다.

NA1_genome / NA1_genome

[View map](#) [View gene table](#) [View operon table](#)

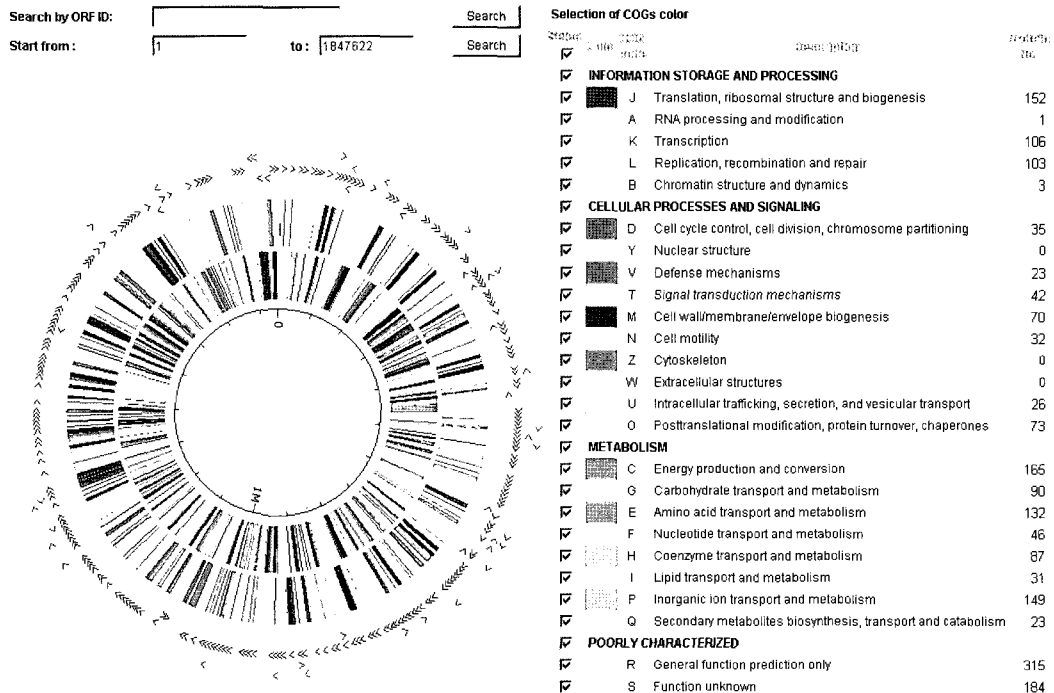


그림 3. 해양유래 초고온성 고세균 Thermococcus sp. NA1의 유전체 web browser 화면.



● 인적사항

성명 : 이정주(李廷珠)

소속기관 : 가천의과대학교 생명과학부 교수연구소 /
서울대학교 생명과학부

직위 : 겸임교수 / 명예교수

연락처 : 전화 02-856-3345

전자메일 : chung-c-lee@hanmail.net

전공 : 유전학(인류유전학, 초파리유전학)

● 교육배경

- 1959 서울대학교 문리과대학 생물학과 이학사
- 1962 서울대학교 대학원 생물학과 이학석사
- 1975 서울대학교 대학원 동물학과 이학박사

● 주요경력

- 1967 한국원자력청 방사선농학연구소 연구원
- 1970 서울대학교 전임강사, 조교수, 부교수, 교수
- 1979 미국 Temple대학원 의대 교환교수, 연구교수
- 1989~현재 국립과학수사연구소 유전자감식 자문위원
- 2001~현재 가천의과대학교 생명과학부 겸임교수