

## Lipase 유전자의 보존적 영역 탐색

이동근 · 김철민<sup>1</sup> · 김상진<sup>2</sup> · 이상현 · 이재화\*

신라대학교 공과대학 생명공학과  
<sup>1</sup>부산대학교 의과대학 의학과  
<sup>2</sup>한국해양연구원 미생물연구실

## Investigation of Conserved Regions in Lipase Genes

Dong-Geun Lee, Cheol-Min Kim<sup>1</sup>, Sang-Jin Kim<sup>2</sup>, Sang-Hyeon Lee and Jae-Hwa Lee\*

*Department of Bioscience and Biotechnology, College of Engineering, Silla University, Busan 617-736, Korea*

*<sup>2</sup>School of medicine, College of Medicine, Pusan National University, Busan 609-735, Korea*

*<sup>3</sup>Microbiology Laboratory, Korea Ocean Research and Development Institute, Ansan PO Box 29, 425-600, Korea.*

### Abstract

For the investigation of conserved regions in lipase genes, 132 and 24 sequences were obtained from LED (Lipase Engineering Database) and COG (Clusters of Orthologous Groups of proteins), respectively. There was high diversity in lipase genes and peculiar amino acid sequences were conserved for each homologous family of LED. Similar conserved amino acid sequences were detected from COG0657 and *Moraxella* lipase 1 homologous group of LED. Although many studies have attempted to detect new lipase genes in procaryotes, they have been limited culturable bacteria. The importance of metagenome, including DNA from non-culturable bacteria, is known. Due to the high diversity, we assumed it might be possible to detect new lipase gene from metagenome. Due to the high diversity of nucleotide sequences in lipase genes, 10 primer sets were designed. Designed primer sets were inspected in BLAST of NCBI and they could amplify a part of the lipase gene from 222 to 713 bp. They can amplify 16.7%~60.0% of each lipase homologous group which was 3.6 fold higher than each sets. They might offer a high probability of detecting new lipase genes, owing to high efficiency and the diversity of lipase genes.

**Key words** — lipase, LED (lipase engineering database), metagenome

### 서 론

Lipase (E.C. 3.1.1.3)는 triacylglycerides의 ester 결합을 glycerol과 fatty acid로 가수분해하는 효소이다. Esterase는 lipase와 같이 ester 결합에 반응하지만 수용성 상태에서

반응이 주로 일어나고, 짧은 지방산을 갖는 ester 결합에 반응하는 세포내 효소 (intracellular enzyme)인데 비해 lipase는 세포외로 활발히 분비되는 효소로 esterase보다 다양한 스펙트럼의 기질특이성을 갖으며 뭉쳐진 기질 (aggregated substrate)에 활발한 활성을 갖는다[5].

Lipase는 생체내에서 지방산의 대사 및 glyceride와 phospholipid의 재합성에 관여하고 식품, 세제 산업 등에서 중요한 것으로 알려져 왔다. 그들은 비수용성(non-

\*To whom all correspondence should be addressed

Tel : 051-999-5748, Fax : 051-999-5636

E-mail : jhalee@silla.ac.kr

aqueous) 상태에서는 ester 합성과 transesterification에도 관여하며, 높은 특이성과 입체특이성 등이 있고 또한 온화한 반응 조건에서 특별한 조효소 (cofactor) 없이 반응이 진행되므로 산업적인 가치가 큰 효소이다[21].

미생물에서 유래한 lipase는 기질특이성이 넓고 유기용매에서 안정하며 다양한 pH와 온도에 적용할 수 있는 잇점이 있어 널리 이용되어 왔다[4,9,20]. 다양한 반응 조건에 적응하는 혹은 높은 활성을 갖는 새로운 lipase 유전자를 탐색하기 위하여 자연계에서 균주를 분리하거나[8], 분리된 균주에서 유전자를 확보하기 위한 cloning[18], 배양가능한 극한미생물에서 유전자를 얻기 위한 시도[17], lipase 생성균주의 genomic library를 이용한 lipase 유전자의 탐색[11,22,25] 등이 보고되고 있으며 lipase 활성 측정을 위한 여러 방법이 소개되고 있다[2].

한편 환경 문제에 대한 인식과 경제적인 이유로 인하여 화학공정을 대체할 수 있는 생물공학적인 응용에 대한 중요성은 새삼 언급할 필요가 없을 정도이다. 효소를 이용한 유기화학 반응의 범위는 아주 다양하며 생촉매 과정에 적용함에 있어 적절한 효소 후보군에 대한 선택은 매우 중요하다고 할 것이다[23]. Lipase 유전자 확보를 위하여 배양가능한 원핵생물에 PCR과 cloning을 적용한 사례가 많으나 현재의 배양기술로는 자연계에 존재하는 미생물의 1% 이상을 배양할 수 없는 것으로 알려져 있으며[26] 99% 이상의 미생물에 포함된 유전자원의 이용가능성은 무한하다고 할 수 있다. Schmidt 등에[23] 의해 배양불가능한 미생물의 유전자에 대한 분자생물학적 접근법이 시도된 이래로 자연환경에서 추출한 메타게놈은 유용 유전자 탐색을 위한 보고로 이용되어 왔다[6].

분자생물학의 발달에 따라 미국에서 생물정보학센터인 NCBI (National Center for Biotechnology Information)가 설립되었고 염기서열에 대한 방대한 자료들을 조직적으로 추가, 관리, 공유할 수 있는 데이터베이스인 GenBank, PIR (Protein Information Resource) 등이 생기게 되었다. 이들 데이터베이스를 이용하여 기존의 *Bacillus* lipase 유전자들에서 보존적 영역을 파악하고 이를 바탕으로 새로운 균주에 PCR을 적용하여 새로운 lipase 유전자를 찾는 시도[18]와 전체 게놈의 염기서열을 알고 있는 균주에서 lipase 유전자로 예상되는 DNA 부분을 증폭시켜 실제로 lipase를 획득한 시도[16]가 있었다. 이러한 보고들은 기존의 염기서

열을 이용하여 생물정보학적인 방법으로 새로운 유용유전자의 탐색이 가능하다는 것을 증명하는 사례들이라고 할 수 있을 것이다.

이 논문에서는 LED (Lipase Engineering Database)[19]와 COG[7]에서 알려진 lipase 유전자들의 염기서열 정보를 이용하여 lipase 유전자의 보존적 영역을 탐색하였다. 한편 이를 바탕으로 새로운 lipase 유전자를 메타게놈에서 탐색·검출하기 위해, PCR primer를 설계하고 새로운 lipase 유전자 검출 가능성에 대한 조사를 목적으로 하였다.

## 재료 및 방법

### 재료

Lipase 유전자중 원핵생물(prokaryote)이 보유하고 있는 lipase 유전자 염기서열만을 대상으로 분석하였다. 분석에 이용된 lipase 유전자의 염기서열은 NCBI 산하의 Genbank 등에서 추출(parsing)하였다. LED[19]를 경유하여 125개 그리고 COG[7]를 경유하여 24개의 염기서열을 추출하여 총 149개의(Table 1) 서열을 분석하였다.

LED (<http://www.led.uni-stuttgart.de/>)는 EMBL (European Molecular Biology Laboratory, <http://www1.embl-heidelberg.de/>), DDBJ (DNA DataBase of Japan, <http://www.ddbj.nig.ac.jp/>), GB (GenBank, <http://www.ncbi.nlm.nih.gov/>), SP (SWISS PROT, <http://www.expasy.org/sprot/>), PIR (Protein Information Resource, <http://pir.georgetown.edu/>), PDB (Protein Data Bank, <http://www.rcsb.org/pdb/>) 등에서 lipase 관련 자료를 취합한 공개용 데이터베이스이다. Lipase를 22개의 상동성그룹(homology family)와 11개의 superfamily로 분류해 놓은 LED 데이터베이스에서 염기서열 추출 대상을 선정하였다. 선정된 *Bacillus* sp., *Burkholderia* sp., *Moraxella* sp., *Mycobacterium* sp., *Pseudomonas* sp., 그리고 *Staphylococcus* sp. lipases와 non-heme peroxidases superfamily에 해당하는 염기서열을 NCBI의 공개용 데이터베이스에서 구하였다. 이 때 원핵생물과 진핵생물의 lipase들로 구성된 상동성그룹에서는 원핵생물의 lipase만 대상으로 하였다.

서열정렬(sequence alignment) 및 보존적 영역 탐색 LED와 COG를 매개로 획득한 lipase 염기서열과 아미노

Table 1. Homologous family of lipases and number of members and nucleotide accession number at NCBI. Nucleotide sequences were obtained by the way of LED (Lipase Engineering Database) and COG.

Homologous family	Number of Members	Accession Number at NCBI
<i>Bacillus</i> lipases	4	AF232707, AJ297356, AJ430831, M74010
<i>Burkholderia</i> lipases	24	AB8452A, AB8452B, AF031226, AF047691, AF050153, AF125523, AF216286, AF237723, AJ250176, AY091666, D10069, D10166, D50587, M14604, M58494, U07884, U33845, U75975, U88907, X14033, X63390, Y00557-A, Y00557-B, X70354
<i>Staphylococcus</i> lipase	24	AC005292, AC007354, AC009398, AF053006, AF086783, AF090142, AF096928, AF134840, AF141874, AF237623, AF208033, AF208229, AF429311, AP004831, AP003138, AY028918, AY095260, AY095262, M12715, M95577, X02844, X95309
<i>Moraxella</i> lipases 1	5	AB066349, AF009336, AL939106, U03114, X53053
LED <i>Haemophilus</i> lipase	7	AB070942, AE000172, AE005246, AE006071, AE008728, L35343, U32704
<i>Moraxella</i> lipases 3	12	AB004320, AB004321, AE012891, AB013096, AE013196, AB045874, AE007281, AF324946, AF395190, AP000994, X53869, X67712
<i>Mycoplasma</i> lipases	8	AE000035, AE000038, AY090779, L38402, U17036, U39712, U39716
Non-heme peroxidases	27	AB034986, AE006315, AE007285, AE007297, AE007766, AE007878, AE007908, AE008199, AE008208, AE008333, AE008336, AE008960, AE009406, AE010765, AF031153, AF031242, AF335493, AF361470, AL939105, D12484, D64005, M60743, M84990, U01096, U02635, U12537, U95170
<i>Pseudomonas</i> lipases	14	AB015053, AB025596, AB063391, AF069748, AF083061, AF144089, AF188366, AF202538, AF286062, AF307943, M74125, M86350, S77830, U11258
COG COG0657	24	NP_293859, NP_294545, NP_295208, NP_214731, NP_214734, NP_215592, NP_215915, NP_215916, NP_216901, NP_217486, NP_217600, NP_218004, NP_215942, NP_217001, NP_216800, NP_243114, NP_415009, NP_286217, NP_254071, NP_232882, NP_298542, NP_104647, NP_105354, NP_107583

산 서열들을 ClustalW의 윈도우 인터페이스인 ClustalX (ver. 1.64b)를 이용하여 정렬하였다. Gap opening penalty (GOP)와 Gap extension penalty (GEP)는 아미노산 (GOP=8, GEP=2)과 염기 (GOP=3, GEP=1)에 대하여 다르게 설정하였고 가중치 매트릭스로는 아미노산의 경우 BLOSUM30을 그리고 핵산의 경우에는 IUB를 이용하였다. 그 외의 parameter는 Clustal 프로그램의 기본값을 이용하였다. 이후 manual alignment로 Clustal 프로그램 정렬결과를 검증하고 보완하여 여러 lipase 사이에 존재하는 보존적인 염기서열과 아미노산 서열을 추적하였다.

#### Phylogram 작성

염기서열정렬 결과를 이용하여 neighbor joining method

로 각 염기서열간의 유사성을 조사하여 계통도(phylogram)를 그렸고 이때 bootstrap method (n=1000)를 이용하였다.

#### PCR primer 설계

새로운 lipase 유전자 후보군의 확인과 cloning에 필요한 PCR primer는 염기서열정렬을 통하여 확보된 자료를 바탕으로 하였다. 각 그룹별로 정렬된 lipase 유전자들의 3'-말단과 5'-말단 영역에서 정렬대상 lipase 유전자 전체 혹은 다수를 증폭시킬 수 있는 염기서열을 확보하고 PCR primer를 설계하였다. 정렬을 끝낸 각 lipase 그룹에서 공통적으로 보이는 염기서열들을 보존적 영역으로 간주하고, 17개 이상의 연속된 염기서열에서 3개 이하의 차이를 보이

는 경우 각 lipase 그룹을 증폭시킬 수 있는 primer 후보군으로 간주하였다. 이후 primer dimer 형성여부와 oligo-nucleotide melting temperature을 고려하여 선택하였다. 선택된 primer 들을 NCBI의 BLAST search (<http://www.ncbi.nlm.nih.gov/blast/>)를 통하여 lipase가 아닌 다른 염기서열까지 증폭시킬 수 있는 지 확인하였다.

### 결과 및 토의

#### 보존적 영역 탐색

LED에서는 2003년 1월 현재 lipase의 아미노산 염기서열의 유사도에 따라 11개의 superfamily와 22개의 상동성 그룹 (homologous family)로 나누고 있었으며 총 600개의 서열로 구성되어 있었다. 이 중 원핵생물이 위주가 되는 총 9 homologous families에 대한 각 구성원들의 염기서열을 총 132개 획득하였으며 100% 동일한 서열로 중복된 것들을 제거하고 최종적으로 125개의 서열에 대한 분석을 수행하였다 (Table 1). LED와 COG에서 구한 염기서열들에 대하여 clustal program을 이용한 alignment와 manual alignment를 수행하였으나 각 homologous family 내에서도 보존적인 염기서열을 구할 수는 없었다.

분석대상 lipase의 아미노산 서열을 분석하여 공통적으

로 발견되는 아미노산 조성을 보존적 영역이라 정의하고 조사하였다. 각 homologous family 수준에서의 보존적 아미노산 서열을 Table 2에 나타내었다. LED에서 추출한 서열들의 분석결과를 보면 임의의 아미노산 X를 포함하는 Ser-X-Gly의 아미노산 서열이 90% 이상의 lipase에서 발견되었으며, 80% 이상에서는 His-Gly 서열이 추가되었다. 이 두 서열은 70% 이상의 보존적 수준에서는 Gly-X<sub>1</sub>-Ser-X<sub>2</sub>-Gly-Gly와 Pro-X<sub>1</sub>-X<sub>2</sub>-X<sub>3</sub>-X<sub>4</sub>-His-Gly로 확대되었으며 60% 이상의 수준에서는 Leu-X<sub>1</sub>-X<sub>2</sub>-X<sub>3</sub>-Gly가 추가되어 총 3개의 영역이 보존적이었고 Gly-X<sub>1</sub>-Ser-X<sub>2</sub>-Gly-Gly는 Gly-His-Ser-X<sub>1</sub>-Gly-Gly로 바뀌는 것을 확인할 수 있었다. Gly-X<sub>1</sub>-Ser-X<sub>2</sub>-Gly-Gly 아미노산 서열은 Brumlik 등[3]과 Jaeger 등 [10]이 보고한 lipase의 active pentapeptide인 Gly(Ala)-X<sub>1</sub>-Ser-X<sub>2</sub>-Gly를 포함하는 것이었다. 각 homologous family 수준에서는 보존적 서열의 수가 증가하였으며 유사도 측면에서 보면 서열 전체 수준에서 유사성을 보이는 family와 일부에서만 유사성을 보이는 family로 나누어짐을 확인할 수 있었다. Active region은 *Bacillus* family에서 Ala-His-Ser-Met-Gly-Gly으로 나타났는데 *Deinococcus radiodurans*에서는 Ser-His-Ser-Met-Gly으로 나타나 특이적인 것으로 확인되었다. Gly(Ala)-His-Ser-X-Gly은 *Bacillus* sp., *Staphylococcus* sp., *Haemophilus* sp., *Pseudomonas* lipase

Table 2. Conservative amino acid sequences of lipase for each group. Conservative ratio was 100% except non-heme peroxidase (90%). Dashes represent arbitrary amino acids and amino acids in parenthesis could replace the very previous amino acid. Letters in bold font represent the active sites.

Homologous family		Conservative amino acid sequences
	<i>Bacillus</i> lipases	PVV(L)-VHG, TGA, VDIV(S) <b>A(S)</b> HSMG, TTIL, S-L-GA-N
	<i>Burkholderia</i> lipases	T-YP, Y---I---L---G--V, KVNL- <b>G-S-G</b> ,
	<i>Staphylococcus</i> lipase	GG-VDY(F)G, <b>G(A)</b> HS-GG(A), DLT--G, DH--V(I)
	<i>Moraxella</i> lipases 1	GT--YPT, G--GA, W-GPRLAS-GRVV, D-PDSR, <b>G-SMGGGG</b>
LED	<i>Haemophilus</i> lipase	D---G-S, <b>G(A)</b> HS-GG
	<i>Moraxella</i> lipases 3	<b>G(A)</b> - <b>S-GG(S)</b>
	<i>Mycoplasma</i> lipases	<b>G(A)</b> - <b>SM(IL)G</b>
	Non-heme peroxidases	D-G, F(L)-HG, GY(F)R-I(V), D-R--H-S, <b>G-S-G-G</b> , P---K---P--G, D---L(I)----P
	<i>Pseudomonas</i> lipases	O--I(L)-G, G-G-G(A), LGY-G(A), GL-G-D(E)V, <b>GHS</b> LGG(M), ASPTO, NIVS(N)F-DHYA(S)S, RR---WV-DLN-AE-H-G-T, TFLF-G-FG-D
COG	COG0657	Y--P, <b>G(A)</b> - <b>S-G</b> , P(A)---I(V,L,M)--G(A,S)

family에서 공히 관찰되었다.

COG는 같은 기원과 기능을 갖는 단백질을 그룹화한 것으로[12-14] lipase 관련 COG는 모두 8개였다. 이중 phospholipase 등을 제외한 esterase/lipase는 COG0657과 COG1647 이었고 각각 52개와 4개의 단백질을 포함하는 그룹이었다. 이중 COG1647은 모두 (carboxyl)esterase 이었고 COG0657에서 esterase와 추정 유전자를 제외한 lipase group은 24개로 나타났다. COG0657에 속하는 24개 lipase들은 모두 Tyr-X<sub>1</sub>-X<sub>2</sub>-X<sub>3</sub>-Pro, Gly(Ala)-X<sub>1</sub>-Ser-X<sub>2</sub>-Gly의 아미노산 서열을 보였다. 이것은 8개의 단백질이 속하는 LED의 *Moraxella* lipase 1 family와 유사한 것이었다. COG0657에 속하는 24개 구성원의 90% 이상이 Gly-Gly-Gly(Ala), Tyr-X<sub>1</sub>-X<sub>2</sub>-X<sub>3</sub>-Pro-X<sub>4</sub>-X<sub>5</sub>-X<sub>6</sub>-X<sub>7</sub>-Pro(Ala), Gly-X<sub>1</sub>-Ser-Ala-Gly-Gly(Ala)을 갖는 것으로 나타났다. Esterase를 포함한 COG0657 구성원 52개의 90% 이상이 Gly-X<sub>1</sub>-Ser-X<sub>2</sub>-Gly의 전형적인 pentapeptide와 Gly-Gly를 가지는 것으로 나타났다. His-Gly-Gly, Tyr-X<sub>1</sub>-X<sub>2</sub>-X<sub>3</sub>-Pro, Gly-X<sub>1</sub>-

Ser-Ala-Gly은 80% 이상의 COG0657 구성원들이 나타내는 것으로 확인되었다. 하나의 미생물이 여러 lipase 유전자를 가지는 경우도 있어 *Mycobacterium tuberculosis*가 12개, *Mesorhizobium loti*가 3개, *D. radiodurans*가 3개 등이었다.

Lipase 유전자의 다양성

아미노산 서열과 달리 LED의 각 homologous family 내에서도 보존적인 염기서열을 구할 수 없었고 COG0657의 24개 유전자에서도 마찬가지였다. 이는 amino acid와 codon간의 중복성 등에 의한 것으로 사료되었다. Fig. 1과 같이 각 homologous family 내에서도 유사성이 아주 높지는 않다는 것으로 확인할 수 있었다. 그리고 bootstrap 숫자가 900 이하가 많아 서로 간의 유사성에 관한 상관관계가 아주 높지는 않은 것을 알 수 있었다. 이것은 lipase 유전자의 다양성을 반영하는 것으로도 해석할 수 있었다. 실제로 lipase는 Ser, His, Asp의 catalytic triad를 내포하고 있으며 [3,10] active site의 Ser 잔기는 임의의 아미노산 X<sub>1</sub>과 X<sub>2</sub>를

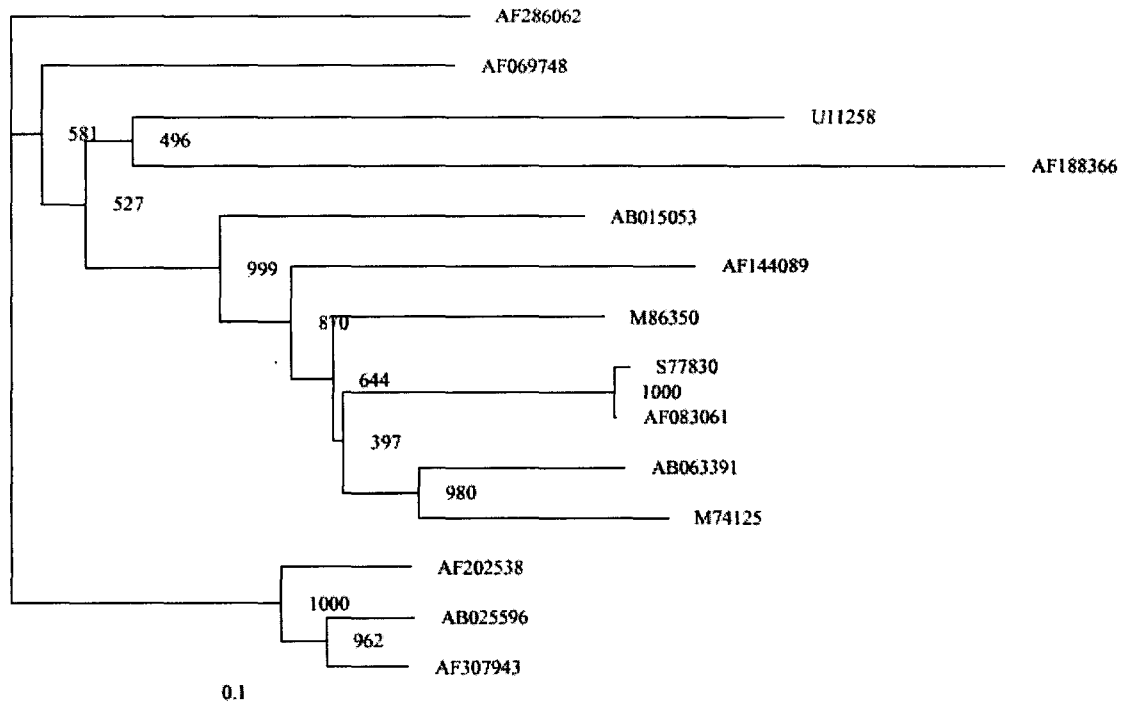


Fig. 1. Phylogram of *Pseudomonas* sp. homologous family from neighbor-joining analysis of lipase genes, which represents the diversity of gene sequence within same homologous family. Bootstrap values at each node are expressed as a number over 1000 trials. Alphabet and following numbers represent the accession numbers at Genbank of NCBI.

갖는 Gly(Ala)-X<sub>1</sub>-Ser-X<sub>2</sub>-Gly[10] 혹은 Gly-X<sub>1</sub>-Ser-X<sub>2</sub>-Gly[3]로 구성된 semiconservative pentapeptide 만이 보존적인 것으로 알려져 lipase 유전자의 염기서열의 보존도는 그다지 높지 않은 것을 알 수 있었다. 이는 많은 lipase 유전자를 증폭시킬 수 있는 universal primer set의 설계에 어려운 점을 제공하였지만 아직도 자연계나 메타게놈에서 새로운 lipase 유전자를 탐색할 가능성이 높다는 것도 동시에 내포하는 것으로 사료되었다.

#### PCR primer 설계와 유용성 검증

자연계에 존재하는 미생물의 1% 이상을 배양할 수 없는 현재의 기술로는 유용 유전자 탐색에 한계가 있으며[26] 이러한 한계는 분자생물학적 기술의 발달에 따른 메타게놈의 이용으로 극복할 수 있을 것이다[6]. 메타게놈에서 유용유전자 탐색을 위해 genome library를 구축하고 screening 하는 일은 많은 노동력과 시간을 필요로 한다. 한편 기존에 알려진 서열에 기초한 유용 유전자 탐색법은 이러한 수고를 덜 수 있는데 이는 PCR과 PCR-hybridization 등으로 가능하다[24]. 본 연구에서는 새로운 lipase 유전자를 검색하기 위해 메타게놈이나 메타게놈 library에 screening method로 적용 가능한 PCR primer set를 고안하였다 (Table 3). COG0657에 속하는 24개의 염기서열에서는 universal primer를 검색할 수 없었다. 기존의 lipase 유전자들에서 유래한 primer set이 알려진 lipase 유전자를 검출할 가능성도 있지만 lipase 유전자 염기서열의 다양성에 비추어 보면 새로운 lipase 유전자를 증폭할 가능성 또한 높다고 할 것이다. PCR에 기초한 방법은 메타게놈 library를 이용한 expression-based system에서 검출한계 이하로 발견되어 간과할 수 있는 유용 유전자까지 검출할 수 있는 장점까지 내포하고 있다[15].

Table 3과 같이 총 10개의 primer set은 222~713 bp 크기의 lipase 유전자 일부를 증폭시킬 수 있는 것으로 나타났다. Non-Heme peroxidase의 경우는 278 bp 혹은 286 bp의 증폭산물이 생성될 수 있으며 *Pseudomonas* homologous family의 primer set은 708 bp의 증폭산물도 생성될 수 있는 것으로 나타났다.

각 homologous group의 lipase 유전자 일부를 증폭시킬 수 있는 각 primer set의 효율은 16.7% (*Moraxella*3 lipase homologous group)와 60.0% (*Moraxella* lipase homolo-

gous group) 사이였고 전체적으로는 총 125개의 실험대상 lipase 유전자중 36개의 유전자를 증폭시킬 수 있어 분석대상 lipase 유전자의 28.8%를 증폭시킬 수 있는 것으로 판명되었다. 염기서열면에서 파악한 각 homologous lipase 그룹의 다양성을 보면 *Moraxella*3 lipase homologous group이 가장 높고 *Moraxella* lipase homologous group이 가장 낮은 것으로 판명되었다. 각 homologous group으로 구분하지 않고 전체 염기서열에 대한 분석을 수행하여도 각 그룹으로 구분한 것과 유사한 결과를 구할 수 있었다. 본 연구결과에서 도출된 primer set 들을 이용하면 개별적인 primer를 이용할 때 보다 3.6배 높은 효율로 lipase 유전자의 검출이 가능한 것으로 드러났다. 또한 lipase 유전자의 다양성을 고려하면 새로운 유전자를 메타게놈에서 탐색할 수 있는 가능성은 충분하다고 사료되었다.

Van Kampen 등은[27] *Staphylococcus*의 lipase 유전자를 cloning 하기 위하여 1kb 정도를 증폭시킬 수 있는 primer를 설계하였고 Amada 등은[1] *Pseudomonas* genome의 lipase 유전자를 cloning 하기 위하여 1.8 kb 정도의 산물을 갖는 PCR primer를 이용하였다. 그들은 배양가능한 균주에서 유래한 게놈을 이용하였지만 향후 메타게놈을 대상으로 연구를 수행하면 좁은 범위의 유전자를 증폭시키는 것이 보다 다양한 lipase 유전자를 탐색할 수 있을 것으로 사료된다. Ruiz[18] 등은 *Bacillus megaterium*에서 새로운 lipase 유전자를 탐색하기 위해 기존의 *Bacillus* lipase 일부를 증폭시킬 수 있는 primer set을 이용한 후 염기서열분석을 통해 새로운 primer를 설계하고 적용하는 기법을 이용하여 전체 lipase 유전자를 cloning 하였다. 본 연구에서 설계한 *Bacillus* homologous lipase의 일부를 증폭시키는 primer의 3' 말단은 Ruiz[18] 등이 이용한 primer set의 5' 말단과 겹치는 것으로 나타났다. 따라서 본 연구에서 설계된 primer 역시 새로운 lipase 유전자를 탐색하는데 충분한 효율가치가 있다고 할 것이다. 메타게놈 상태에서 PCR을 이용한 pre-screening으로 여러 primer set 중에서 일부만 증폭되면 genome library에서의 염기서열 추적에 높은 경제성과 신속함을 줄 수 있을 것이다. 그리고 배양을 통한 assay에서 간과할 수 있는 lipase 유전자까지도 검출할 수 있을 것으로 사료되었다.

설계된 primer set들은 degeneracy에 의해 목적 유전자 외에 다른 염기서열을 증폭시킬 가능성이 있었으므로,

Table 3. Applicable PCR primer sets to detect lipase gene candidate from metagenome and the lengths of PCR product at each homologous family. Ratio was expressed as the number of amplifiable sequences over that of analyzed sequences. Numbers after primer represent their melting temperature (T<sub>m</sub>, °C).

Homologous family	Length (bp)	Primer Set (T <sub>m</sub> )	Amplifiable Sequence	Ratio
<i>Bacillus</i>	342	F : AAA AAA GT(A/G) GAT ATT GTG GC (40) R : TTA ATT CGT ATT CTG (T/G)CC (40)	AJ297356 AF232707	2/4
<i>Burkholderia</i>	326	F : TAC CCC ATC GTG CTG GCC CA (60) R : TCG GCG GTG TCC GAA CCC TTG (61)	X63390 AB8452A D50587 AF237723 D10166	5/24
<i>Staphylococcus</i>	286	F : GTC GAT TAT GG(G/C) GCA GC (47) R : CGT CCC GTC ATG AGG AGT (52)	AF134840 AF237623 AY095260 AY095262 X95309	5/24
	283	F : GT(A/T) GA(T/C) TAT GGT GCA GC (39) R : TGT ACC ATT ATG TGG TGT (41)	AP003138 AP004831 M95577 M955771	4/24
<i>Moraxella</i>	559	F : ACT ACC CGA CCA GCA CC(A/G) G (54) R : AAG CGC TT(G/C) AGC CAG GAG A (53)	AF009336 AL939106 U03114	3/5
<i>Haemophilus</i>	222	F : ATC GTT CTT GTC CA(T/C) GGT C (50) R : ACC GCC CAT GGA GTG A(G/C)C (55)	AE000172 AE005246	7/14
<i>Moraxella 3</i>	237	F : GTC GT(A/G) CTG ATC CAC GGG (54) R : CCC GCC GAA GCT GTT GCC A (57)	AB004320 AB004321	2/12
<i>Mycoplasma</i>	401	F : TTG GT(A/C) A(T/G)G C(T/A)A TGG GTG (41) R : ACA CCA TCT TTT TCA CCA AG (46)	U17036 U170361	2/8
non-Heme peroxidase	280	F : AT(C/T) TAT T(T/A)C AAG GAC TGG GGC (50) R : TGT A(G/A)C G(G/C)G CCA C(A/T)T CG (49)	AF031153 AF0311531 D12484 D124841 U12537 U125371	6/27
<i>Pseudomonas</i>	713	F : TT(G/A) AGG TC(C/T) TG(C/G) ACC CAG (46) R : GTG CTG GGC AAG TAC GAT (49)	AB015053 AF083061 AB063391 AF144089 S77830	5/14

NCBI의 blast를 이용하여 조사하였다. Primer set 들은 원핵생물의 lipase외에 동물로는 초파리 (*Drosophila* sp.), 쥐, 인간의 DNA를 증폭시킬 수 있었으며 식물로는 애기장대 (*Arabidopsis thaliana*)의 DNA도 증폭시킬 수 있는 것으로 나타났다 (data not shown). 나머지 오염 가능성은 실험과정에서 조금만 주의하면 되지만, 인간 DNA에 의한 오염은 특히 주의하여야 할 것으로 사료되었다.

## 요 약

Lipase 유전자들의 보존적 영역을 탐색하기 위해 LED (Lipase Engineering Database)와 COG (Clusters of Orthologous Groups of proteins)를 통하여 각각 132개와 24개의 서열을 얻어 분석하였다. Lipase 유전자의 염기서열은 아주 다양하였고 LED의 각 상동성그룹 (homologous family) 별로 독특한 아미노산 서열이 보존적인 것을 확인할 수 있었다. COG0657에 속하는 lipase들은 LED의 *Moraxella* lipase 1 homologous group과 유사한 아미노산 보존적 영역이 있음을 확인하였다. 다양한 반응 조건에 적응하는 혹은 높은 활성을 갖는 새로운 lipase 유전자를 원핵생물에서 탐색하기 위하여 다양한 시도들이 수행되어 왔지만 그들은 모두 배양가능한 미생물에 국한되어 있었다. 배양할 수 없는 세균의 유전자원을 포함하는 메타게놈에 대한 유용성 역시 최근에 널리 인식되고 있다. Lipase 유전자의 다양성으로 보았을 때 메타게놈을 이용한 새로운 lipase 유전자를 찾는 작업도 가능할 것으로 사료되어 lipase 유전자 일부를 (222~713 bp) 증폭시키는 총 10개의 PCR primer를 설계하였으며 그 가능성을 NCBI의 BLAST를 통하여 검증하였다. Lipase 유전자의 염기서열은 아미노산 서열보다 아주 다양하여 비교적 많은 수의 primer set이 필요하였다. 각 primer set의 증폭효율은 각 LED group의 16.7%와 60.0% 사이였고 개별적인 primer set을 이용할 때 보다 3.6 배 효율이 높은 것으로 드러났다. Lipase 유전자의 다양성은 설계된 primer set 들을 이용하여 새로운 lipase 유전자를 검출할 가능성을 높이는 것으로 사료되었다.

## 참 고 문 헌

- Amada, K., M. Haruki, T. Imanaka, M. Morikawa and S. Kanaya. 2000. Overproduction in *Escherichia coli*, purification and characterization of a family I.3 lipase from *Pseudomonas* sp. MIS38. *Biochim. Biophys. Acta.* **1478**, 201-210.
- Beisson, F., T. Ali, R. Cluude and R. Verger. 2000. Methods for lipase detection and assay: a critical review. *Eur. J. Lipid Sci. Technol.* 133-153
- Brumlik, M. J. and J. T. Buckley. 1996. Identification of the catalytic triad of the lipase/acyltransferase from *Aeromonas hydrophila*. *J. Bacteriol.* **178**, 2060-2064.
- Demirjian, D. C., M. V. Francisco and C. S. Cassidy. 2001. Enzymes from extremophiles. *Curr. Opin. Biotechnol.* **5**, 144-151.
- Fojan, P., P. H. Jonson, M. T. Petersen and S. B. Petersen. 2000. What distinguishes an esterase from a lipase: A novel structural approach. *Biochimie.* **82**, 1033-1041.
- Handelsman, J., M. R. Rondon, S. F. Brady, J. Clardy and R. M. Goodman. 1998. Molecular biological access to the chemistry of unknown soil microbes: a new frontier for natural products. *Chemistry & Biology* **5**, R245-R249.
- <http://www.ncbi.nlm.nih.gov/COG/>
- Ishimoto, R., M. Sugimoto and F. Kawa. 2001. Screening and characterization of trehalose-oleate hydrolyzing lipase. *FEMS Microbiol. Lett.* **195**, 231-235.
- Jaeger, K. E. and T. Eggert. 2002. Lipases for biotechnology. *Curr. Opin. Biotechnol.* **13**, 390-397.
- Jaeger, K. E., S. Ransac, B. W. Dijkstra, C. Colson, M. van Heuvel and O. Misset. 1994. Bacterial lipases. *FEMS Microbiol. Rev.* **15**, 29-63.
- Litthauer, D., A. Ginster and E. van Eeden Skein. 2002. *Pseudomonas luteola* lipase: A new member of the 320-residue *Pseudomonas* lipase family *Enzyme Microb. Technol.* **30**, 209-215.
- Lee, D.-G., H.-Y. Kang, J.-H. Lee and C.-M. Kim. 2003. Detection of Conserved Genes in *Proteobacteria* by using a COG Algorithm. *Korean J. Biotechnol. Bioeng.* **17**, 560-565.
- Lee, D.-G., H.-Y. Kang, S. Kim, S.-H. Lee, C.-M. Kim, S.-J. Kim and J.-H. Lee. 2003. Classification of Archaeobacteria and Bacteria using a gene content tree approach. *Korean J. Biotechnol. Bioeng.* **18**, 39-44.
- Lee, D.-G., C.-M. Kim and J.-H. Lee. 2003. Genetic composition analysis of marine-origin Euryarchaeota by using a COG Algorithm. *Korean J. Life Sci.* **13**, 298-307.

1. Amada, K., M. Haruki, T. Imanaka, M. Morikawa



15. Lorenz, P., K. Liebeton, F. Niehaus and J. Eck. 2002. Screening for novel enzymes for biocatalytic processes: accessing the metagenome as a resource of novel functional sequence space. *Curr. Opin. Biotechnol.* **13**, 572-577.
16. Manco, G., E. Giosue, S. D'Auria, P. Herman, G. Carrea and M. Rossi. 2000. Cloning, Overexpression, and Properties of a New Thermophilic and Thermostable Esterase with Sequence Similarity to Hormone-Sensitive Lipase Subfamily from the Archaeon *Archaeoglobus fulgidus*. *Arch. Biochem. Biophys.* **373**, 182-192.
17. Rúa, M. R., C. Schmidt-Dannert, S. Wahl, A. Sprauer and R. D. Schmid. 1997. Thermoalkalophilic lipase of *Bacillus thermocatenulatus* large-scale production, purification and properties: aggregation behaviour and its effect on activity. *J. Biotechnol.* **56**, 89-102.
18. 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.
19. Pleiss, J., M. Fischer, M. Peiker, C. Thiele and R. D. Schmid. 2000. Lipase engineering database: Understanding and exploiting sequence - structure - function relationships, *Journal of Molecular Catalysis B: Enzymatic* **10**, 491-508.
20. Schmid, A., J. S. Dordick, B. Hauer, A. Kiener, M. Wubbolts and B. Witholt. 2001. Industrial biocatalysis today and tomorrow. *Nature* **409**, 258-268.
21. Schmidt-Dannert, C. 1999. Recombinant microbial lipases for biotechnological applications. *Bioorg. Med. Chem.* **7**, 2123 -2130.
22. Schmidt-Dannert, C., M. L. Rúa, H. Atomi and R. D. Schmid. 1996. Thermoalkalophilic lipase of *Bacillus thermocatenulatus*. I. Molecular cloning, nucleotide sequence, purification and some properties, *Biochim. Biophys. Acta.* **301**, 105-114.
23. Schmidt, T., E. DeLong and N. Pace. 1991. Analysis of a marine picoplankton community by 16S rRNA gene cloning and sequencing. *J. Bacteriol.* **173**, 4371-4378.
24. Seow, K. T., G. Meurer, M. Gerlitz, E. Wendt-Pienkowski, C. R. Hutchinson and J. Davies. 1997. A study of iterative type II polyketide synthases, using bacterial genes cloned from soil DNA: a means to access and use genes from uncultured microorganisms. *J. Bacteriol.* **179**, 7360-7368.
25. Sullivan, E. R., J. G. Leahy and R. R. Colwell. 1999. Cloning and sequence analysis of the lipase and lipase chaperone-encoding genes from *Acinetobacter calcoaceticus* RAG-1, and redefinition of a Proteobacterial lipase family and an analogous lipase chaperone family. *Gene* **230**, 277-285.
26. Torsvik, V., L. Ovreas and T. F. Thingstad. 2002. Prokaryotic diversity -magnitude, dynamics, and controlling factors. *Science* **296**, 1064-1066.
27. Van Kampen, M. D., R. Rosenstein, F. Götz and M. R. Egmond. 2001. Cloning, purification and characterisation of *Staphylococcus warneri* lipase 2. *BBA-Prot. Struc. Mole. Enzy.* **1544**, 229-241.

(Received June 23, 2003; Accepted October 20, 2003)