

## Supplementary Materials

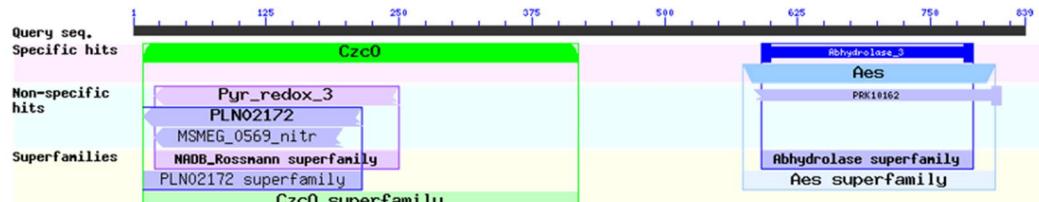
# Characterization of a Novel Family IV Esterase Containing a Predicted CzcO Domain and a Family V Esterase with Broad Substrate-Specificity from the Oil-Polluted Mud Flat Metagenomic Library

Jong Eun Park<sup>1</sup>, Geum Seok Jeong<sup>1</sup>, Hyun Woo Lee<sup>1</sup>, Sung Kyum Kim<sup>2</sup>, Jungho Kim<sup>2</sup>, Hoon Kim<sup>1,2,\*</sup>

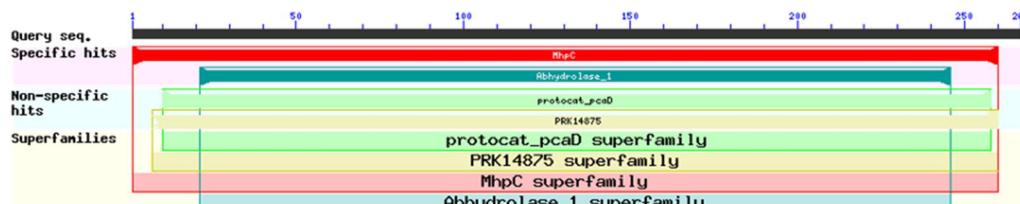
<sup>1</sup> Department of Pharmacy, and Research Institute of Life Pharmaceutical Sciences, Sunchon National University, Suncheon 57922, Republic of Korea; park140201@naver.com (J.E.P); fever41@naver.com (G.S.J); hyun\_325@naver.com (H.W.L)

<sup>2</sup> Department of Agricultural Chemistry, Sunchon National University, Suncheon 57922, Republic of Korea; rlatjdrua1234@nate.com (S.K.K); jungho@schnu.ac.kr (J.K)

\* Correspondence: hoon@sunchon.ac.kr; Tel.: +82617503751



(a)



(b)

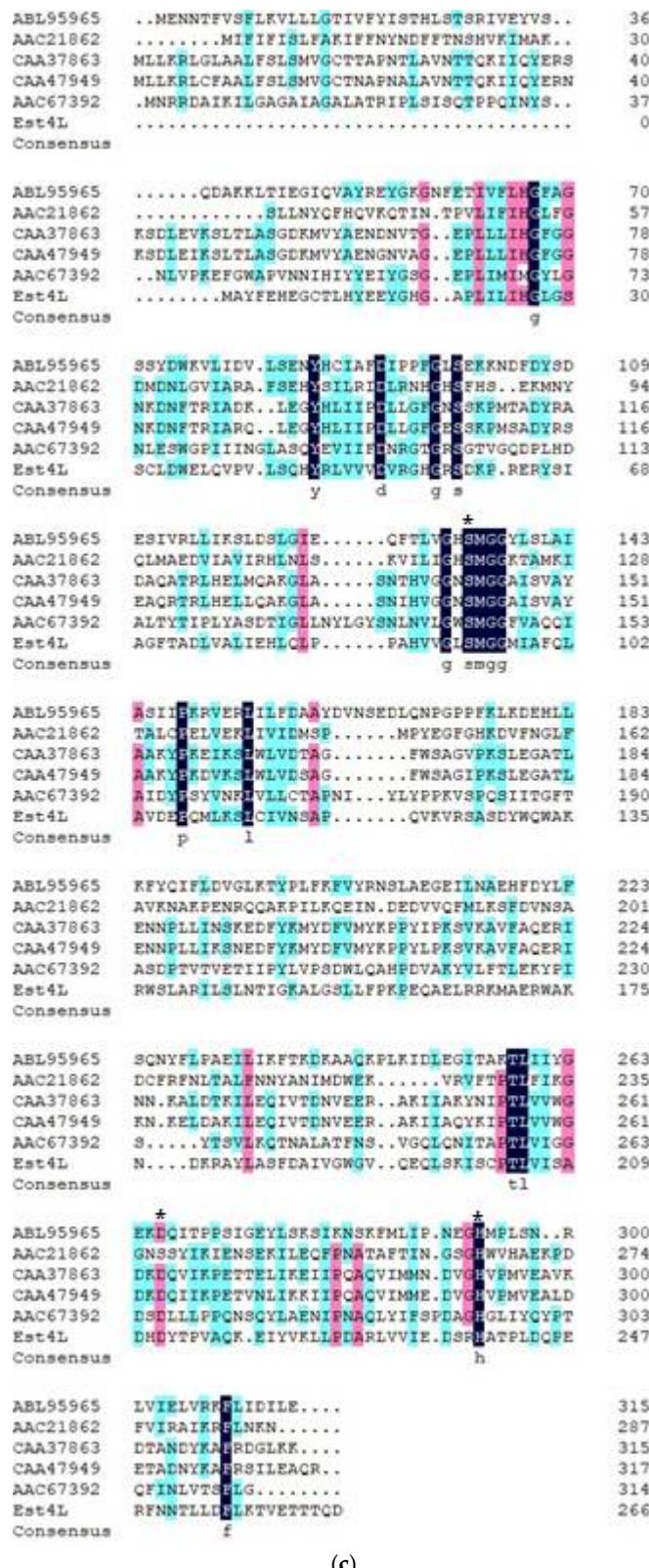
**Figure S1.** BLASTp of Est2L (a) and Est4L (b). Est2L was matched to two domains of a CzcO superfamily domain (10-419 amino acid residues) and an acetylesterase superfamily (573-811 amino acid residues).

Est2L Czco domain	THETEDCSATELEPIDIISGCGCLGMAIAHRAFGVHR...E	44	Est2L Czco domain	EEKQRLLRQFWFMKLVFPAHYLHFESFALGFTRIQLITK	269
007085	.....NYLTIVIEAGACAGACISIGYMKCSICK.....E	27	007085	.....IVYFIMIG.....KPSIEWW	208
NP_244544	.....MESIKWVIVCGCAIAMIGYVKEKEF.....E	29	NP_244544	.....ERELFLKL.....NKSMEWW	211
NP_355926	EPAAEIGYAEQFYVIIINGGGCIGALGARFCIGVE...T	196	NP_355926	GMITRPMALIASLFRIMHEFQVFIYIRIPECLAEFVAA	433
NP_390541	.....NYLTIVIEAGACAGACISIGYMKCSICK.....E	27	NP_390541	.....IVYFIMIG.....KPSIEWW	208
NP_601977	.....MSTNVEAIIIEAGACALDAAFEISRGCFTECKE	34	NP_601977	.....ETCREELAGWGIAVEFAVERFTNGKAFAAS	236
Consensus	y iiagacagia gyyl q gv f		Consensus	f pl ig s fw	
Est2L Czco domain	IILEPGQDVGVWPRENSVPAPACIVFSHLYSISSEFNFW	84	Est2L Czco domain	VVGGPFKKLLHTSVAS...FEIRPQNTIDYFIGCRRIILS	306
007085	IIILRSHEVGESWKR...YDSIVLFTSMYSSLPKMHLEGE	66	007085	FIKLGVIHASHTSIVG...KFIQKRCIFVFGHELKFAIKQ	245
NP_244544	MIIIPNEQVGDISWRNR...YDSIVLFTSMYSSLPKMHLEGE	68	NP_244544	IEWGGLIYAGITITPG...FWERKRCIFIGKRELSLIRK	248
NP_355926	IIIEENRERWKR...YKSICLICHFWYCHLFYIIFEFEN	235	NP_355926	IERAGFMILFGCDESGLMRYLFRGSGGYIIVGACILVID	473
NP_390541	IIILFSHEVGESWKR...YDSIVLFTSMYSSLPKMHLEGE	66	NP_390541	ERLGLVILHASHTSIVG...KFIQKRCIFVFGHELKFAIKQ	245
NP_601977	IIILESNIGHGAWRH...WISITIGAEGIALIICLMNNP	73	NP_601977	VVRTGIFQIYLIG...VWGLLIVSRGMEEFIFTGTVFG	275
Consensus	iildan vg srxrd ydslvltip yssllpg p g		Consensus	k g l a hds g k lk dp fg elk ik	
Est2L Czco domain	SRVAFPÇAEIHGYIÇHCAVYELKEHIFRGAEVFEAEFIT	124	Est2L Czco domain	NEYIKTFLCQFHVLITRGIPRITEHGIELIIGEEFAVIAI	346
007085	KHGFESRNEIVAYIKKVVKEFEIF...JCIIRTEVISVKLIK	104	007085	KEILMLK.....RVIAJKCNEIIIFKESSTIEVNIN	275
NP_244544	FNGEITRKLEMASVLCQIANEFNLF...MHHHTKVRVTRQQ	106	NP_244544	GQIHLRF.....RMNVNCQKEVEFAHHSRISSEFRI	278
NP_355926	WFVFTIKVKGKLEMTRVMELN...YWGTTICSAÇYEE	273	NP_355926	GSIKLHSG...LVLSHLTERAVVLRKGTEIIAIV	505
NP_390541	KHGFESRNEIVAYIKKVVKEFEIF...JCIIRTEVISVKLIK	104	NP_390541	KEILMLK.....RVIAJKCNEIIIFKESSTIEVNIN	275
NP_601977	DKFTIATSIAGYCAENESEA...VVRFVHSERFETS	111	NP_601977	EFFAEASEG...LGFSAGTIIAVEESWKFYIAGTHIIVENI	313
Consensus	gfp k ei yl y felp i tev sv		Consensus	i lk rv a enei fkds l vd i	
Est2L Czco domain	AN...ACWHVTCADDCQRHAPPRLIISMGQISPAHICLFG	161	Est2L Czco domain	IYIGIGIAPATEFLSF.MRITGRDGLIINQAWGCPAAAYIGI	385
007085	N.....YELIKINFEYCYIENLVIAGFFHTENIISISK	138	007085	IKVILCQEGRIIHHHRGVSVFVEGL	312
NP_244544	NG....RFBLLITNGWIEERKVVIAGCAQKEYLIFVFD	141	NP_244544	IKWIGSESFEYEWI...IIGVIMATNGWIFHNRCITNIKGL	315
NP_355926	TIGETVTVVFLAGKEWVILREKCLVLMIGMSGHANHKEFG	313	NP_355926	VYIIGNGSMNGWAPALISPIVAEKVGKCGWGLGSETTKCFG	545
NP_390541	N.....YELIKINFEYCYIENLVIAGFFHTENIISISK	138	NP_390541	IKVILCQEGRIIHHHRGVSVFVEGL	312
NP_601977	EELFS.SFLRVSSLGIREWIIIFMVNAIGTWINEYVHIFG	150	NP_601977	FWVIGCEFAIRHLAMKLRQCGRGIBILIEVSFASLAR.V	352
Consensus	n f kt e tk viatig f pn p ipg		Consensus	iwatgrf l wi i gd g iih rg s gl	
Est2L Czco domain	MASFEGRIESEAFLHIFYFLACKFVATIGTGASDICHVIE	201	Est2L Czco domain	SVIGCFNNEMLYGENTNLGHNSIIEMIESÇISHVÇARÇA	425
007085	ELISININGIASSCYNSKÇIAYGNVIVWGGNSCACIAME	178	007085	YFICL...WQHKGSGAIQCQVG...NIAEYIVKÇMN	343
NP_244544	SANNEMSCVİSSPARYRNFAÇIFGKSYIIVGGGASCACIAME	181	NP_244544	YFICL...WQYQRGSALVCGVG...NIAEYIVFFIK	346
NP_355926	QLVFEGCGCİSSCHEGFLAYAKRVVILIESNSHİHICAA	353	NP_355926	FMBEGCRNMİRETRÇIÅLWFHGGNİBÇSRHYSÇVİSLİLK	585
NP_390541	ELISININGIASSCYNSKÇIAYGNVIVWGGNSCACIAME	178	NP_390541	YFICL...WQHKGSGAIQCQVG...NIAEYIVKÇMN	343
NP_601977	IKFEGKGİIATVNYRPAEDEKGPWVWGGGİSVCİHİLE	190	NP_601977	IIVYG...STATIVGATFAGRVAG...FVAAKRİLVEVS	386
Consensus	sf g qlhscy n qlagk vlvvgggnsaaqlave		Consensus	yf glp wq rgsal gvg daeylv q	
Est2L Czco domain	VARÇIAELKVFCSEIAYIME.....FADREYSA	229	Est2L Czco domain	IYESLASHVVECER....	440
007085	ISKEFVITYIACSNK.....	192	007085	.GE.....	345
NP_244544	IAKEENVİMAİSF.....	195	NP_244544	HGCK.....	350
NP_355926	IWEACALVİMLFSSTHİVFSİSİMİGİGİLYSEFAVAN	393	NP_355926	ARÇAGLKTEVYGLÇKEHHBL	604
NP_390541	ISKEFVITYIACSNK.....	192	NP_390541	.GE.....	345
NP_601977	IEGLİETİTWİTRFFN.....	206	NP_601977	FHVN.....	390
Consensus	l ker t a sr		Consensus	g	

(a)

Est2L	HCRFCLGICRRLATSVWSGCKSWYVDERGHNSTNWPGFTW	480
PHS53692	LLTFSFAFFILVVSMTFSASTSGASTHNTDTASASEKQAD	42
WP_082823965	TLSFIFSSFILVVSMTFSSTSEASAHTDTAKT.....	38
AVP72281	.....	0
ARG92633	.....	0
Consensus	.....	
Est2L	SYRWLARYAGLAAYRLSSPLDTQSPACEGQRIAAPPDWLE	520
PHS53692	TVKISNESPLTVPAKTLPLPSASSDELKSAISCYPMPSVD	82
WP_082823965	....TEPLLTVPTRNVPLPSAASDELKSAISEYPMPSVD	74
AVP72281	.....MASPELD	7
ARG92633	.....MGRFN	5
Consensus	.....	
Est2L	RANAAFLRVFLRVSFRALIGPPRALASQRKIVDILLWLMP	560
PHS53692	EVINNTPQSIEQWRELIQIRNADCKKKIKKMRKGFDVVDVS	122
WP_082823965	EIINNTPQSVEQWRELIQCRNGDCKKKIKKNNRKGFDVVDVS	114
AVP72281	TVF....QMIKEWGENFGTTIEDNRLAYEKLVEPLFWVDD	43
ARG92633	QKLA...WMPFRIRRNRTPTPRVIQALLRSSCMVAGNELLKH	43
Consensus	.....	
Est2L	GCLGVNRRSMQLDGSLQIYEPGTRAPKGIVLYLHGGAF	600
PHS53692	...LEKINGVPVVRRLTPKTIAREFKN..KVVIDVHGGAYV	157
WP_082823965	...LEKINGVTVRRLTPKTIAREFKN..KVVIDVHGGAYV	149
AVP72281	...VKTERVGAGSAPAEMIIAGAEDG.PILLYLHGGGYV	79
ARG92633	...GLQAESRRVGSVPVRIIPKEKRAK.GVVIDIHGGGWV	79
Consensus	P hgg	
Est2L	LGSPRSHYSVTSRLARDSGCAVWVVEDYIRLAFEPHYFAALE	640
PHS53692	FFSGLPSIEESLLIAHHRVCGITVISIDYSPPIHAFFPAALN	197
WP_082823965	FFAGLPSIEESLLIANPVGITVISIDYSPPIHAFFPAALN	189
AVP72281	MGSMRTHRVMLAHISRAAGARVGLIDYIRLAEEFVFPACVE	119
ARG92633	IGNACMNDLNVAMVNACEVAVVSDYDPLAVNTEVEGILE	119
Consensus	v dy	
Est2L	GCCLACYDAIRACLC..PADKELIASISAGGNLVIALALAL	678
PHS53692	EVVSVVSSVVAEH...GAQNLFIGGTISAGACIVLAAVQTL	234
WP_082823965	EVVSVVSSVVAEH...DAHNLFIGGTISAGACIVLAAVQTL	226
AVP72281	DAVAAARWLLIANGS..DPKKIVISGSIISAGGGIMVATLVAL	157
ARG92633	ECLATARWLLADCEEFAGLPVIVVSEBAGGHLAAATTLLAL	159
Consensus	d a g sag l 1	
Est2L	KERGNGIARAGLMLLSPTDPLGGASMCSPCEADSPMIR..	716
PHS53692	IADKQPLPAAVYAGTPTWADLTKEGDTLYINEGVERILVTY	274
WP_082823965	IETRKCSLPLAVYAGTPTWADLTKEGDTLYTNEGIERILLY	266
AVP72281	PYLGEPPMPAAVGVLSTWWVIMEGTGEGTFITNAEVIPMVG..	195
ARG92633	KQSPELLARVSGAVLYYGVIYDLPCTPSVRTAGRETILLDG	199
Consensus	g	
Est2L	PRWLECALKAYA...APVQALSHRELEADLRGIPPMIO	752
PHS53692	CGFLEAAAANLYAGS..ESLTTHSSISIPLYGNFDGPPTEELI	312
WP_082823965	CGFLAAASAKLYAGS..ESLTTHPSISSLYGNFDGPPTEELV	304
AVP72281	KDLILQIAGVYLGG..KDPRPAPLASEINADLTGIPPELLO	233
ARG92633	PGMVERALRLITPGLSEDCRRQQPPLSLEYGDLACEPEALMF	239
Consensus	g pp	
Est2L	VGDQELLIAESTRLAERARQSGVACQCLEIHCGRMEVFCLQ	792
PHS53692	SGTREMELISITVVRVNPKLRDACVRTGLEVFEGLSADYVV	352
WP_082823965	SGTREMELISITVVRVNPKLRNANVRTQLEVFEGLSADYVV	344
AVP72281	VGSIETLIDCISNCLARLAKADGVEVVKVEVWDEMHPVFCDF	273
ARG92633	VGELEPLIIDLTLQMAERWAGSEVFLLPQRAHGFIEPTAM	279
Consensus	g l d v h	
Est2L	AFYLRSARNALLGLAPPARQCLETAGQTQQAADSPTTQP	832
PHS53692	AYETPESHSVYQCELKQCELLSVCTKSDD.....	379
WP_082823965	AYETPESLSVYQCELKQCELLSVSTNSFD.....	371
AVP72281	APILPEALCAIDGIGEFIKKHTG.....	296
ARG92633	S.....GRVLAYSREWITGRLRSVG.....	299
Consensus	.....	
Est2L	ATDPQP	838
PHS53692	.....	379
WP_082823965	.....	371
AVP72281	.....	296
ARG92633	.....	299
Consensus	.....	

(b)



(c)

**Figure S2.** Multiple sequence alignments of a CzcO domain of Est2L (a), an acetylesterase domain of Est2L (b), and Est4L (c) using the Clustal W method in DNA/MAN. Completely and highly conserved sequence(s) in the alignments are shown in black and pink boxes, respectively. Marks with star represent catalytic residues.

References for MSA sequences(Figure S2)

(Figure S2.a)

**O07085:** Sturr, M.; G.; Ablooglu, A.; J.; Krulwich, T.; A. A *Bacillus subtilis* locus encoding several gene products affecting transport of cations. *Gene*, **1997**, 188(1), 91–94.

**NP\_244544:** Nakasone K.; Masui N.; Takaki Y.; Sasaki R.; Maeno G.; Sakiyama T.; Hirama C.; Fuji F.; Takami H. Characterization and comparative study of the rrn operons of alkaliphilic *Bacillus halodurans* C-125. *Extremophiles*, **2000**, 4(4), 209-214.

**NP\_355926:** Wood D.W.; Setubal J.C.; Kaul R.; Monks D.E.; Kitajima J.P.; Okura V.K.; Zhou Y.; Chen L.; Wood G.E.; Almeida N.F.; Jr. Woo L.; Chen Y.; Paulsen I.T.; Eisen J.A.; Karp P.D.; Bovee D. Sr.; Chapman P.; Clendenning J.; Deatherage G.; Gillet W.; Grant C.; Kutyavin T.; Levy R.; Li M.J.; McClelland E.; Palmieri A.; Raymond C.; Rouse G.; Saenphimachak C.; Wu Z.; Romero P.; Gordon D.; Zhang S.; Yoo H.; Tao Y.; Biddle P.; Jung M.; Krespan W.; Perry M.; Gordon-Kamm B.; Liao L.; Kim S.; Hendrick C.; Zhao Z.Y.; Dolan M.; Chumley F.; Tingey S.V.; Tomb J.F.; Gordon M.P.; Olson M.V.; Nester E.W. The genome of the natural genetic engineer *Agrobacterium tumefaciens* C58. *Science*. **2001**, 14;294(5550), 2317-2323.

**NP\_390541:** Borriis R.; Danchin A.; Harwood C. R.; Médigue C.; Rocha E.; Sekowska A.; Vallenet D. *Bacillus subtilis*, the model Gram-positive bacterium: 20 years of annotation refinement. *Microb. Biotechnol.* **2018**, 11(1), 3–17.

**NP\_601977:** Silberbach M.; Schäfer M.; Hüser A. T.; Kalinowski J.; Pühler A.; Krämer R.; Burkovski A. Adaptation of *Corynebacterium glutamicum* to ammonium limitation: a global analysis using transcriptome and proteome techniques. *Appl. Environ. Microbiol.*, **2005**, 71(5), 2391–2402.

(Figure S2.b)

**PHS53692:** Won S. J.; Jeong H. B.; Kim H. K. Characterization of novel salt-tolerant esterase isolated from the marine bacterium *Alteromonas* sp. 39-G1. *J. Microbiol. Biotechnol.* **2020**, 30(2), 216–225

**WP\_082823965:** No available reference(article).

**AVP72281:** Borchert E.; Selvin J.; Kiran S.G.; Jackson S.A.; O'Gara F.; Dobson A.D.W. A novel cold active esterase from a deep sea sponge *Stelletta normani* metagenomic library. *Front. Mar. Sci.* **2017**, 4, 287.

**AKG92633:** Kim H.J.; Jeong Y.S.; Jung W.K.; Kim S.K.; Lee H.W.; Kahng H.Y.; Kim J.; Kim H. Characterization of novel family IV esterase and family I.3 lipase from an oil-polluted mud flat metagenome. *Mol Biotechnol.* **2015** 57(9),781-792.

(Figure S2.c)

**ABL95965:** Cai J.; Xie Y.; Song B.; Wang Y.; Zhang Z.; Feng Y. *Fervidobacterium changbaicum* Lip1: identification, cloning, and characterization of the thermophilic lipase as a new member of bacterial lipase family V. *Appl. Microbiol. Biotechnol.* **2011**, 89(5), 1463-1473.

**AAC21862:** Tatusov R.L.; Mushegian A.R.; Bork P.; Brown N.P.; Hayes W.S.; Borodovsky M.; Rudd K.E.; Koonin E.V. Metabolism and evolution of *Haemophilus influenzae* deduced from a whole-genome comparison with *Escherichia coli*. *Curr. Biol.* **1996**, 6(3), 279-291.

**CAA37863:** Feller G.; Thiry M.; Gerday C. Nucleotide sequence of the lipase gene lip3 from the antarctic psychrotroph *Moraxella* TA144. *Biochim. Biophys. Acta*. **1991**, 1088(2), 323-324.

**CAA47949:** Arpigny J.L.; Feller G.; Gerday C. Cloning, sequence and structural features of a lipase from the antarctic facultative psychrophile *Psychrobacter immobilis* B10. *Biochim Biophys Acta*. **1993**, 1171(3), 331-333.

**AAC67392:** Arpigny J.L.; Jendrossek D.; Jaeger K.E. A novel heat-stable lipolytic enzyme from *Sulfolobus acidocaldarius* DSM 639 displaying similarity to polyhydroxyalkanoate depolymerases. *FEMS. Microbiol. Lett.* **1998**, 167(1):69-73.