

Supplementary Materials

for

**Comparative phylogenomic analysis reveals evolutionary genomic changes and novel toxin families in endophytic *Liberibacter* pathogens**

Yongjun Tan<sup>1</sup>, Cindy Wang<sup>1</sup>, Theresa Schneider<sup>1</sup>, Huan Li<sup>1</sup>, Robson Francisco de Souza<sup>2</sup>,  
Xueming Tang<sup>3</sup>, Kylie D. Swisher Grimm<sup>4</sup>, Tzung-Fu Hsieh<sup>5,6</sup>, Xu Wang<sup>7,8,9</sup>, Xu Li<sup>5,6</sup>, and  
Dapeng Zhang<sup>1,10,\*</sup>

<sup>1</sup>Department of Biology, College of Arts & Sciences, Saint Louis University, St. Louis, MO 63103

<sup>2</sup>Departamento de Microbiologia, Instituto de Ciências Biomédicas, Universidade de São Paulo, São Paulo 05508-900, Brazil

<sup>3</sup>School of Agriculture and Biology, Shanghai Jiao Tong University, Shanghai, China

<sup>4</sup>United States Department of Agriculture—Agricultural Research Service, Temperate Tree Fruit and Vegetable Research Unit, Prosser, WA

<sup>5</sup>Department of Plant and Microbial Biology, North Carolina State University, Raleigh, NC 27695

<sup>6</sup>Plants for Human Health Institute, North Carolina State University, Kannapolis, NC 28081

<sup>7</sup>Department of Pathobiology, College of Veterinary Medicine, Auburn University, Auburn, AL 36849

<sup>8</sup>Alabama Agricultural Experiment Station, Auburn University, Auburn, AL 36849

<sup>9</sup>HudsonAlpha Institute for Biotechnology, Huntsville, AL 35806

<sup>10</sup>Bioinformatics and Computational Biology Program, College of Arts & Sciences, Saint Louis University, St. Louis, MO 63103

Correspondence and requests for materials should be addressed to D.Z. (email: [dapeng.zhang@slu.edu](mailto:dapeng.zhang@slu.edu))

## Content list

1. Supplementary Table 1. Detailed genome information of Huanglongbing-associated bacteria and other related bacteria that were used in this study (Fig.1 & Fig. 2).
2. Supplementary Table 2. Genomic positions of the identified *Liberibacter* prophage loci.
3. Supplementary Table 3. A list of orthogroups that were lost in *Liberibacter* pathogens (Appended).
4. Supplementary Table 4. A list of representative genes that were gained in the ancestor of pathogenic *Candidatus* *Liberibacter* species.
5. Supplementary Table 5. The KEGG pathways whose components (enzymes) are present in ancestral *L. crescens* but lost in *Liberibacter* pathogens (Appended).
6. Supplementary Figure 1. Gene ontology analysis (in the category of biological processes) of genes that were lost in pathogenic *Liberibacter* bacteria.
7. Supplementary Figure 2. The maps of KEGG pathways whose components were lost in pathogenic *Liberibacter* bacteria using the Blast2GO program.
8. Supplementary Figure 3. Multiple sequence alignment (MSA) of the PD domain.
9. Supplementary Figure 4. MSA of the Ntox52 domain.
10. Supplementary Figure 5. MSA of the beta-sandwich domain associated with the YdjM domain.
11. Supplementary Figure 6. MSA of the Exonuclease-Endonuclease-Phosphatase (EEP) family.
12. Supplementary Figure 7. Phylogenetic relationships of the *Liberibacter* EEP related sequences and other known EEP enzymes.
13. Supplementary Figure 8. MSA of the potential transporter (Peptidase\_M75).
14. Supplementary Figure 9. MSA of the Phage\_lysozyme family.
15. Supplementary Figure 10. MSA of a new protein family which contains four conserved Cysteines.
16. Supplementary Data 1. Detailed genomic information of the identified *Liberibacter* prophage loci (Appended).
17. Supplementary Data 2. HMM profiles of the newly-identified domains in this study.

**Supplementary Table 1.** Detailed genome information of Huanglongbing-associated bacteria and other related bacteria that were used in this study (Fig.1 & Fig. 2).

Species	Strain	Source	GenBank Accession Number	Length (bp)	Year	Reference
<i>Candidatus Liberibacter asiaticus</i>	gxpsy	Guangxi, China	NC_020549.1	1,268,237	2012	1
<i>Candidatus Liberibacter asiaticus</i>	psy62	Florida, USA	NC_012985.3	1,227,328	2010	2
<i>Candidatus Liberibacter asiaticus</i>	A4	Guangdong, China	NZ_CP010804.2	1,230,251	2019	3
<i>Candidatus Liberibacter asiaticus</i>	JXGC	Jiangxi, China	NZ_CP019958.1	1,225,162	2017	4
<i>Candidatus Liberibacter asiaticus</i>	Ishi1	Okinawa, Japan	NZ_AP014595.1	1,190,853	2014	5
<i>Candidatus Liberibacter asiaticus</i>	AHCA1	California, USA	NZ_CP029348.1	1,233,755	2018	6
<i>Candidatus Liberibacter africanus</i>	PTSAPSY	Pretoria, South Africa	NZ_CP004021.1	1,192,232	2012	7
<i>Candidatus Liberibacter solanacearum</i>	ZC1	Texas, USA	NC_014774.1	1,258,278	2010	8
<i>Candidatus Liberibacter americanus</i>	Sao Paulo	Sao Paulo, Brazil	NC_022793.1	1,195,201	2013	9
<i>Candidatus Liberibacter europaeus</i>	NZ1	Canterbury, New Zealand	PSQJ00000000.1	1,331,154	2018	10
<i>Liberibacter crescens</i>	BT0	Puerto Rico, USA	NZ_CP010522.1	1,522,119	2015	Not available
<i>Liberibacter crescens</i>	BT1	Puerto Rico, USA	NC_019907.1	1,504,659	2012	11

**Supplementary Table 2.** Genomic positions of the identified *Liberibacter* prophage loci.<sup>1</sup>

Species	Strain	Accession Number	LC1		LC2		SC		UT	
			start	stop	start	stop	start	stop	start	stop
	gxpsy	NC_020549.1			1118314	1137878	1223784	1264363	1173699	1185042
<i>Candidatus</i>	psy62	NC_012985.3			1132029	1149043	1186049	1223266		
<i>Liberibacter</i>	A4	NZ_CP010804.2			1129881	1146896	1196268	9590	1184862	1195729
<i>asiaticus</i>	JXGC	NZ_CP019958.1			1129886	1149448	1194901	3341	1182715	1194055
	Ishi1	NZ_AP014595.1			1127556	1144572	1196849	3320	1185268	1196611
	AHCA1	NZ_CP029348.1			1129943	1149509	1200706	3333	1180544	1788
<i>Candidatus</i>	PTSA-PSY	NZ_CP004021.1			26536	48147	449353	488428		
<i>Liberibacter</i>							176477	216722		
<i>Liberibacter</i>	ZC1	NC_014774.1			814348	832703	1214438	1257724		
<i>solanacearum</i>							749038	787835		
<i>Candidatus</i>	Sao Paulo	NC_022793.1			555495	576335	788985	806744		
<i>Liberibacter</i>		PSQJ01000004.1			44516	62901				
		PSQJ01000015.1					1	16745		
<i>Candidatus</i>	NZ1	PSQJ01000003.1					1	13653		
<i>Liberibacter</i>		PSQJ01000006.1					173	39658		
<i>europaeus</i>		PSQJ01000005.1					20427	41563		
		PSQJ01000013.1					195	3703		
<i>Liberibacter</i>	BT0	NZ_CP010522.1	588495	601703	898955	916810				
<i>crecens</i>	BT1	NC_019907.1	544655	557864	850693	868545				

<sup>1</sup> The shading blocks indicate that bacteria do not harbor the corresponding prophage loci.

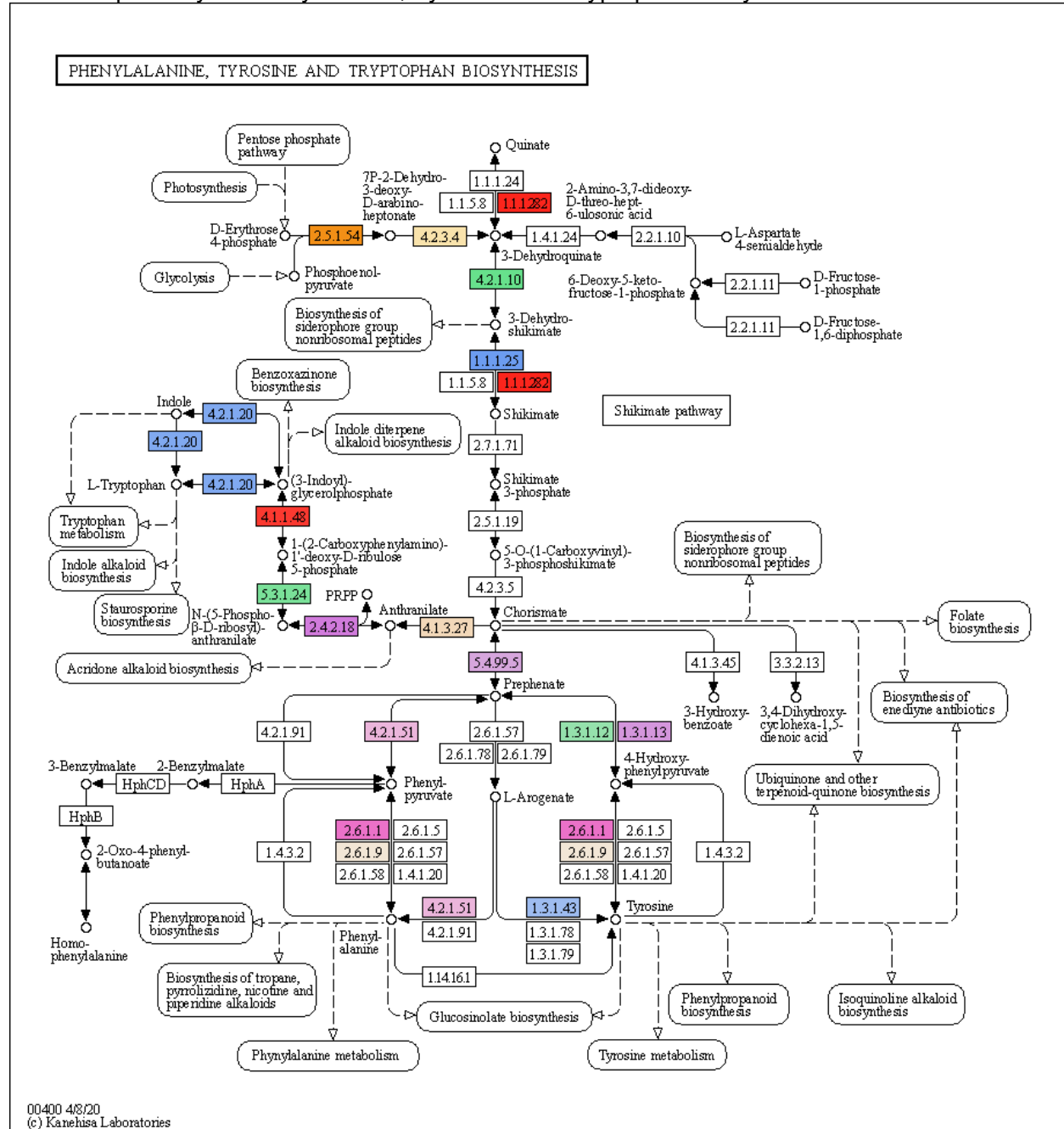
**Supplementary Table 4.** A list of representative genes that were gained in the ancestor of pathogenic *Candidatus* Liberibacter species.

Ortholog Groups Representative Accession	Length	Definition	Domain Architecture (Pfam + novel domains)	Function Annotation
Toxins/effectors				
WP_012778667.1	311	hypothetical protein	PD+Ntox52	Ntox52 is distantly homologous to the EspG/VirA effectors
WP_012778668.1	459	hypothetical protein	PD+Ntox52	
WP_015452837.1	304	hypothetical protein	SP+EEP	EEP family <sup>12</sup> (see Supplementary Figure 6)
WP_015452856.1	163	metal-dependent hydrolase	YdjM	Homologous to alpha-toxin and the Het-C toxin domain <sup>13</sup>
Transporters				
WP_012778500.1	201	SCO family protein	SP+SCO1-SenC	periplasmic copper binding protein <sup>14</sup>
WP_012778517.1	118	EamA family transporter	SP+EamA	small multidrug resistance protein
WP_012778542.1	491	NTP/NDP exchange transporter	SP+TLC	NTP/NDP transporter
WP_015452442.1	260	metal ABC transporter permease	ABC-3	Ion ABC transporter
WP_015452443.1	240	metal ABC transporter ATP-binding protein	ABC_tran	Metal ABC transporter
WP_015452444.1	294	zinc ABC transporter solute-binding protein	SP+ZnuA	Zinc ABC transporter
WP_102134465.1	418	peptidase	SP+Peptidase_M75 (wrong annotation)	Iron transporter <sup>15</sup> (see Supplementary Figure 8)
WP_015452664.1	314	cation transporter	Cation_efflux	Cation transporter
WP_015452813.1	202	LysE family translocator	LysE	Lysine translocator
WP_015452909.1	242	outer membrane beta-barrel protein	SP+beta-barral	A beta barrel membrane protein
WP_015452959.1	225	hypothetical protein	SP+ beta-barral	
WP_015452999.1	420	dicarboxylate/amino acid:cation symporter	SDF	Sodium:dicarboxylate transporter
WP_015452595.1	195	ABC transporter substrate-binding protein	SP+SBP_bac_6	ABC transporter
WP_015824913.1	107	hypothetical protein	SBP_bac_8	periplasmic iron-binding protein
Enzymes for DNA/RNA synthesis and regulation				
WP_109298658.1	1781	DUF4011 domain-containing protein	DUF4011+AAA_11 +RNA-Helicase_SF1 +REase_AbiJ-DUF2726	Replication/transcription
WP_015452495.1	445	RsmB/NOP family class I SAM-dependent RNA methyltransferase	NusB+Methyltr_RsmB-F	16S rRNA (cytosine(967)-C(5))-methyltransferase <sup>16</sup>
WP_015452499.1	471	exodeoxyribonuclease I	35exo_RnaseT.1 +Exonuc_X-T_C	exodeoxyribonuclease I involved in DNA recombination and repair <sup>17</sup>
WP_015452977.1	83	helix-turn-helix transcriptional regulator	HTH	transcriptional factor
WP_015452998.1	483	deoxyribodipyrimidine photo-lyase	DNA_photolyase +FAD_binding_7	DNA photolyase <sup>18</sup>
Enzymes for small molecule metabolism				
WP_015452887.1	171	lysozyme	Phage_lysozyme	glycoside hydrolase <sup>19</sup> (see Supplementary Figure 9)
WP_015452479.1	250	5'/3'-nucleotidase SurE	SurE	Survival protein; metal-dependent phosphatase <sup>20</sup>
WP_015452406.1	216	MBL fold metallo-hydrolase	Mbetalac	beta-lactamase; antibiotic resistance <sup>21</sup>
WP_015452826.1	152	acyl-CoA thioesterase	4HBT	thioesterase <sup>22</sup>
WP_015452942.1	298	polyprenyl synthetase family protein	polyprenyl_synt	polyprenyl synthetases
Transmembrane proteins				
WP_012778543.1	61	hypothetical protein	TM	Transmembrane protein
WP_015452822.1	151	hypothetical protein	TM+TM	Transmembrane protein
WP_015452847.1	302	DUF4065 domain-containing protein	DUF4065+TM+TM	Transmembrane protein
WP_047263992.1	119	hypothetical protein	TM+Unknwon region	Transmembrane protein
Function unknown proteins				
WP_015452797.1	229	hypothetical protein	SP+4Cys	May be a novel toxin (see Supplementary Figure 10)
WP_012778451.1	98	hypothetical protein	DUF167	
WP_047263804.1	189	hypothetical protein	DUF6456	

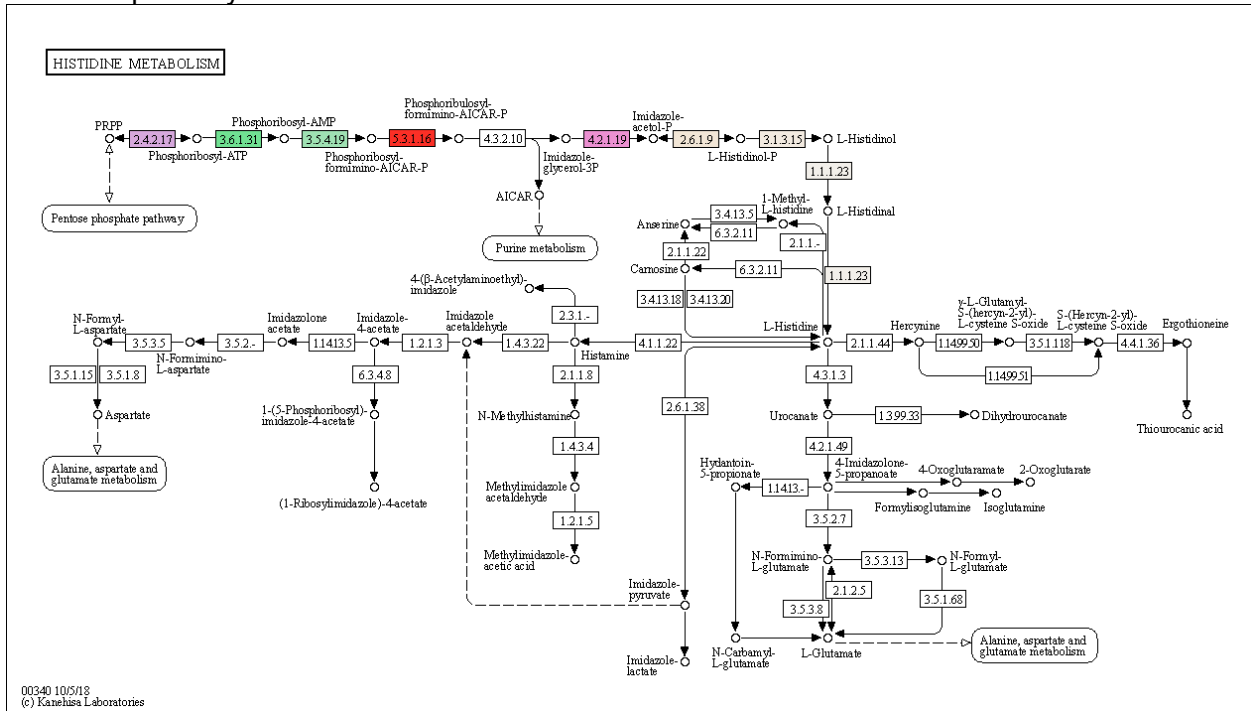


**Supplementary Figure 2. The maps of KEGG pathways whose components were lost in pathogenic *Liberibacter* bacteria.**

**a. KEGG pathway of Phenylalanine, Tyrosine and Tryptophan biosynthesis.**

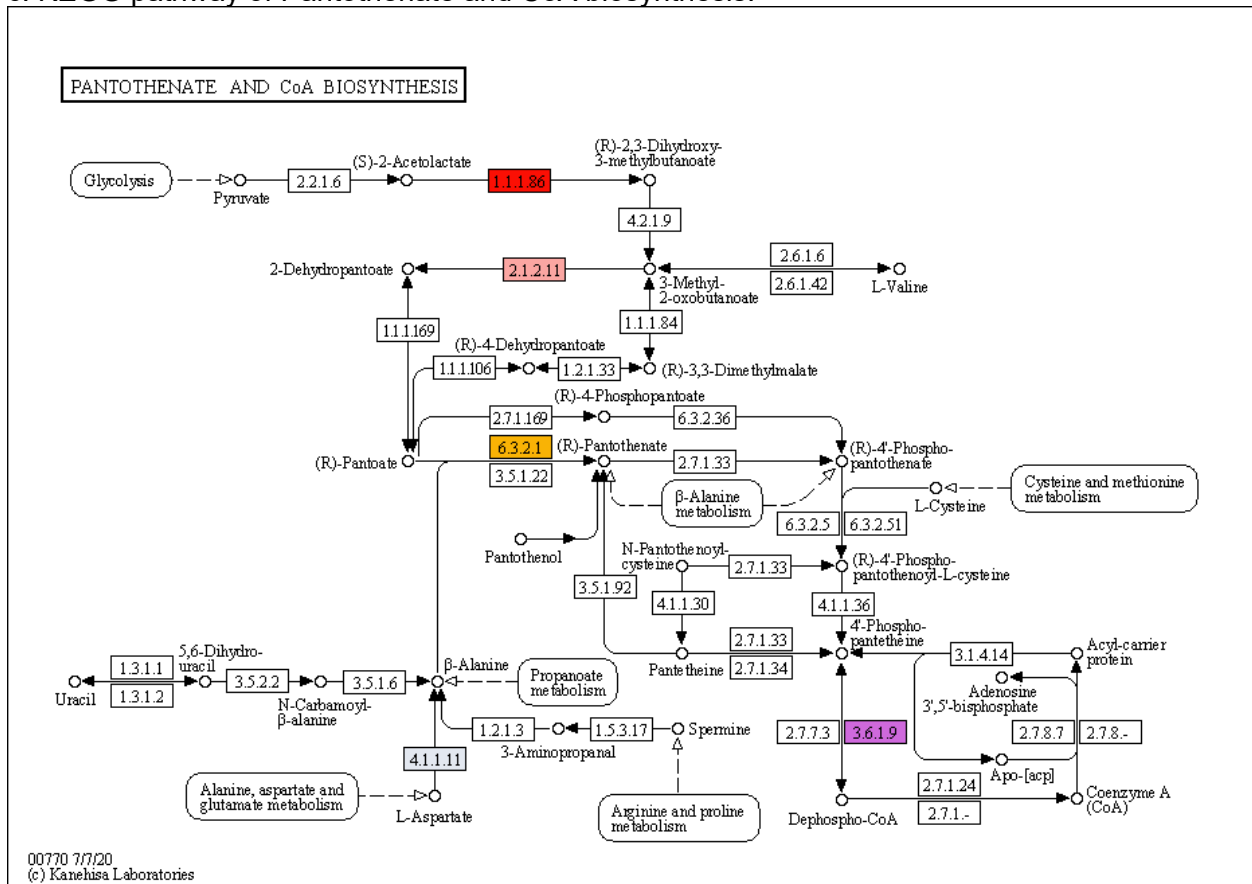


b. KEGG pathway of Histidine metabolism.

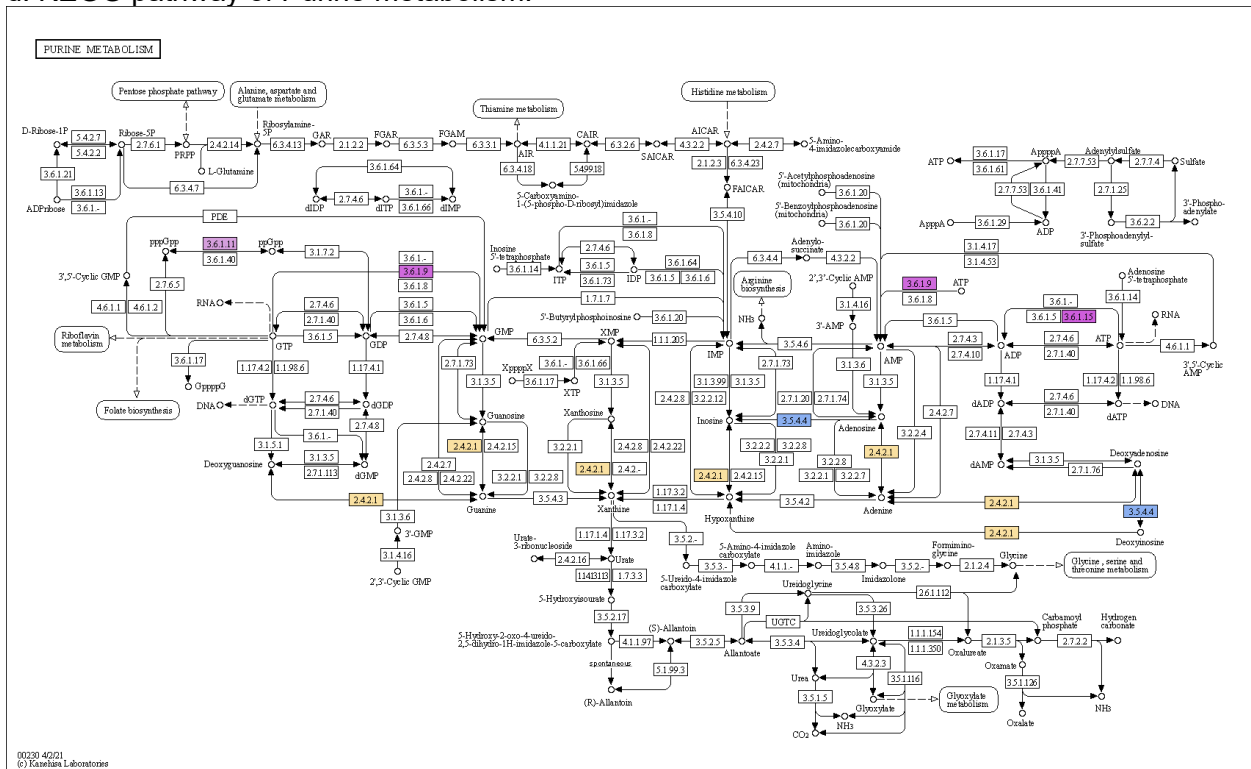




c. KEGG pathway of Pantothenate and CoA biosynthesis.



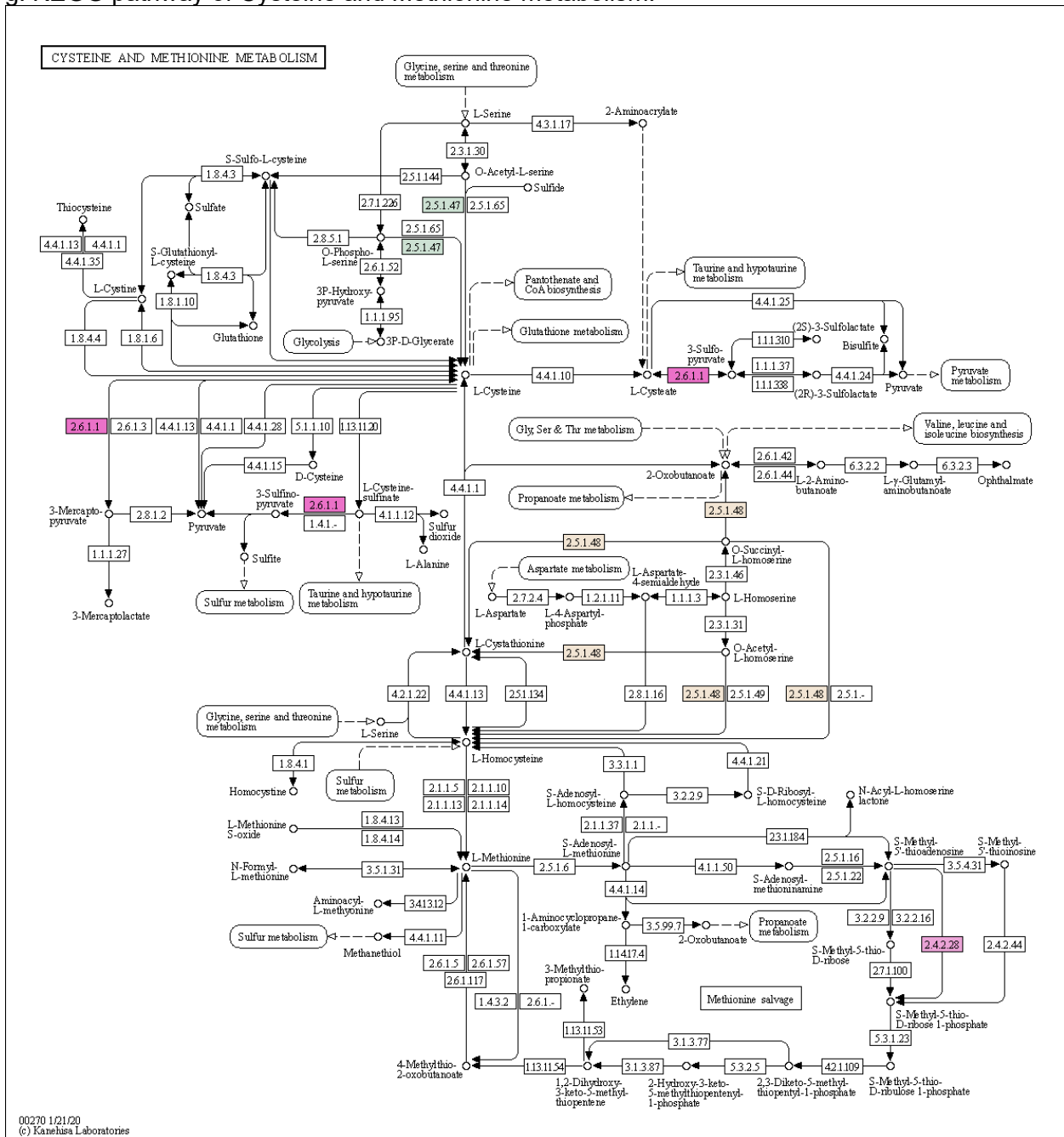
d. KEGG pathway of Purine metabolism.







g. KEGG pathway of Cysteine and Methionine metabolism.





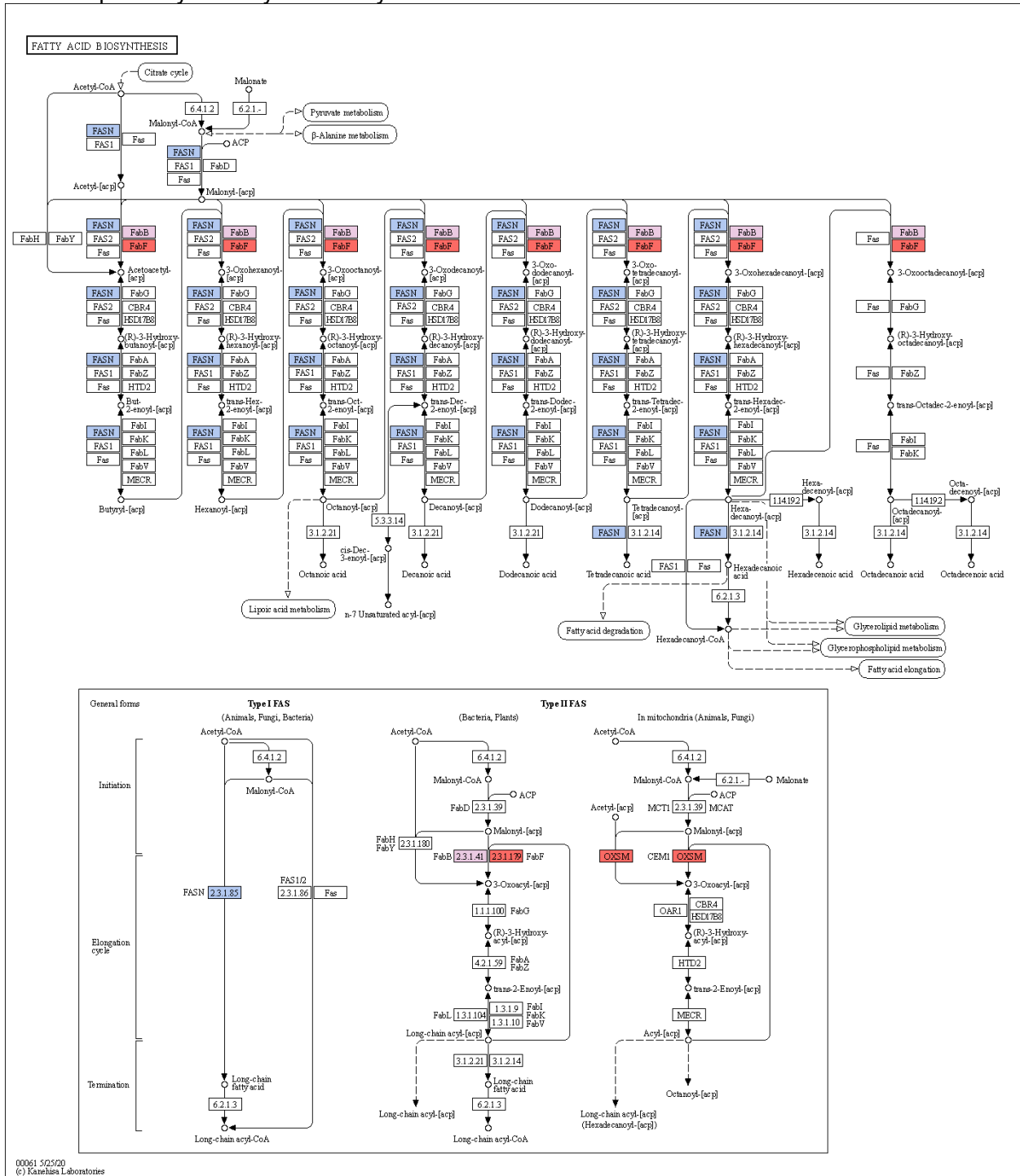








# I. KEGG pathway of Fatty acid biosynthesis.



**Supplementary Figure 3. Multiple sequence alignment (MSA) of the PD domain.** The sequences in MSA are labelled by their species abbreviation followed by NCBI accession number. The secondary structure predicted by JPRED is shown above the MSA, while the consensus result calculated by a custom perl script is displayed below the MSA. The MSA was colored according to the consensus in 80% of the sequences based on 7 amino acid classes (see Methods section for details). Coloring theme: the amino acid involved in either catalytic activities or specific function is highlighted in red background with white font color and bold font style; any amino acid present above the threshold is highlighted in black background with white font color and bold font style; hydrophobic (h) in yellow background with bold font style; aromatic groups (a) in orange background with bold font style; big amino acids (b) in light grey background with bold font style; small amino acids (s) in green background with bold font style; polar (p) group in blue background with bold font style; charged (c) amino acids in pink background with bold font style. The sequence, CLaf\_WP\_148407341.1, consist of two protein sequences, WP\_052775023.1 and WP\_148407341.1, which originated from the same gene in *C. Liberibacter africanus*.

```

Secondary Structure  HHHHHHHHHHHHHHHH  HHHHHHHHH  HH  HHHHHHHHHHHHH  HHH  HHHHHH  HHHHHHHHHHHHHHHH
Clas_WP_012778668.1  STTKVVDVAVDVAVSM-IPIYGTYREFKCNYG---WG-IVGAISDAALLI-PVVGYGARAA---INLVRGGSI---ALKGATAGTIAAEKEACT
Clas_WP_012778667.1  -----M-IEVYGTIQEFKCNYG---WG-ALGIVDVALLA-IPAAYLGKVL---FLVRGSSI---ATKIATGIATVVQEATV
Claf_WP_148407341.1  LEKAKECKEAAISM-IEVYGTIQEFK/-----MG-ANALPSEFKLL---PSKNSAKIS---VAKS-----ESKLDNIYLSGKTINK-
Clso_WP_013461761.1  ISLESIAKKSLIAA-IPIYGTIQAFKEKESG---WG-ILGITDVLTLI--GIGYGIKGA---AALIRGSS---AATAMAAGTAIEGG
Clso_WP_044054101.1  KNEIYHLAKEVALSS-IPIYGTIQAFKEGEIG---WG-IFGIVTDALLIV-PAIGVGAKLA---VNFI-----ARRSKIASTMRSTATIA
Clso_WP_013461760.1  SSKIGETAKEALLSL-IPIYGTIQSFKRGEIG---WG-IFGAITDVLTLV-PVGYGAKMV---GALARGGNA---AIKKSKAGAIASASATAST
SSDFFSNAKDVMVNL-IPIYGTINEFKRCNTG---WG-IVGVITDALLIV-PVFGYGIKVV---ATLARGGAKV---AENAIKAKTIINDFKVLQ
Clam_WP_023466137.1  MDRVSQVAKDVLMSS-IPIYGTIQSFKRGEIG---WG-IFGAITDLLTLV-PVLGSSAKLA---PAASIATRE---ARGLLATAKAVKATSDVV
Clam_WP_007556842.1  IKNILIEIAKNAAINS-IPIYGTIQAFKEKCETK---WG-IFGIVTDVLTLV-PIIGIGSKLI---GSVIKAGSI---TTKFRTMKGGKASIELK
Clam_WP_023466135.1  IDNILNIAKEAAISS-IPIYGTIQSFKRGEIG---WG-IFGAITDFLTLV-PALGFSVKFA---GSLIRGGGV---AMKVETKVVAETMKIANV
CLeu_PTL86945.1  SSNIFYIAKEVFTSM-IPIYGSIQAFKKNTL---YG-LFGIADLLTLI-HIMALGDKLE---GFLIQG-----CKAIVSSIVLV-----
CLeu_PTL86946.1  MSTFHVAKEVLSS-IPIYTIRSFKRGEIG---WG-IFGIVTDVLTIV-PVLGVSAKLA---GSLIRGGGA---AIKAGATAEVIAKTGAA
CLeu_PTL86771.1  ANQIHAVKEALIST-IPIYGTIQEFKRGEIR---WG-IFGAVTDVLLV-PVLGCSAKLI---GSLARGAKVAKNAIEIGKTGVASKVIHG
Lsp_MBA5723779.1  ESPYLEVAKEVGISM-IPIYGTIHEFRCNIG---WG-IFSAVTDALSLI-PVIGAGCKAT---GALIRGGET---AIKGEACVVTESINT
Lsp_MBA5723780.1  ESPYLEAKEIGISM-IPIYGTIHEFRCNIG---WG-IFSAVTDALSLI-PVIGAGCKAT---GALIRGGET---AIKGEACVVTESINAT
Ssym_NIH12428.1  QVSAGEISKEVALSM-IPIYGTVREFRCNIG---WG-IFGAITDALTLI-PVIGAGAKAI---GTAIRGGSAAVKALRVGLTAEVTSKIT-
Xalb_CBA15982.1  AKTYEKVIDEFALSL-IPFYECCKNIYCKIE---EA-VIECTVDVLGLI-PLAVPAGRVV---MKFGE---ALSFALRALRQSSA---
Vcam_WP_103412121.1  VTTDRDILFMVLQAF-IEVGTVESFRCKM---GS-VLGLTGDFMFV--GVGTGIKSA(4)QLLLKRS---KHLFGRGVSLAVKIGI---
Vaqu_WP_126575537.1  REKARREFSKSIGKAI-LLWGTIDDFKGNNS---AG-VAGVGDIMFFV-PIFKASASA---VGRDVTL---SLTLKALKNSSKSFARGI
Srub_WP_163833448.1  QHTARQSMEMLDMLVPFYGFAQAIKKCNVG---EA-IISFADPLSLI-PVAGAAGKLG---GTALKTGF---KVANIARAQALKAGLKIG
Fwei_KMN12671.1  EHPVTRQLLNLAVNS-IPLVSSIKSFARGDIA---GG-VTDLTDIFGFIVPGVKA-GKAG---GLAKGAST---TSRYITLASKVAKALIVT
Pvir_WP_029242134.1  WAKEREYADTAARFL-IPFYGCIQDLRAGEHS---AGVIVGVLDAAFALVPLGQFAGATA---RIVMRAGEL---SVLSITRLTGKALGR---
IDGEADKWIGSALAF-IPFYTAIKNFIAGNTA---MG-LFYASLDVFGLM---LGGTGNRLV---SELVRSAAG---KGTLKSLAALAQAPS---
Psp_WP_116628907.1  KLKTDKSLKDLTIRM-IPLVAAAQDAKACNYG---EA-AVDLAFDAFGL-APELKLIEAGG---SALAKIG---SREVESLASSAGAKEV---
Prho_MBB5004468.1  SREARTRAHNFLLGF-VEVGAYQEFKACNIG---SG-LANLAVDIVGVA-IGAAGPAR---GLIRSAK---ALAFKPLKGIARRLTATV
Pkor_WP_077573129.1  ERTKEKIATYFVDLAVEFKKIEDISSCEHNRVDG-IYGLMDAIALG-GAFYGAGV---KALSI---SAKTISAASKARVTKLV
Pfre_KAF2390978.1  IKETFDSRFDTVKSF-IPFWGSIEDITSCDTDRMFG-FLGLMDVVSFALPIGKFVSGSV---RLVKIASASVAQLSTKATIPSFSTL-
Pflu_WP_044273868.1  EKEKQEFRNLVLGL-IFGSALWNLARCNYK---DA-AADIFDIVLYA--TTAGLKAGS---CAARAG---SSRFKYLTAGKSTAGVG
Pcos_SEE43809.1  EVTVEQIQAVTRAL-IPFASAIDSFRCKIG---QG-LIFLAFDIFGFVVGVGALNRIS---KVVKVGSVGGKAGRLARGLSAANFFAG
Psil_RKP54526.1  IQKRNALVDFLKSL-VEFHDFIQHVRACQSS---EA-VGALGRDVIALI-PLARAIGKVTGLTRLARIGGA---TARAGSANAAGAIT---
CRin_EFL91365.1  EATIIEKIADEFLKHL-IPFYDAIKNLQCKKL---EA-AGYFVLDLMLFL-PLIGEVGKLS---VGM---SEVLISAANISR---
Balb_TKI06771.1  HHKVLHAARDEFLLSL-IPFYSCVDGVKTDPLG---ET-LFDCGMDIVSLL-PVLGPLSKMT(4)TVFIRGLQK---TGREALMMAARKSLGNI
Bfla_ANN79506.1  GQTISEKLAEAERGL-IPFYATYRAIRCDVR---GA-ILSGLMDVATFL-PAIGAGLKLGMTASAIKA---GLEAGLQG-----
Consensus/80% .pp..p.h.phhsh.bpbassbp.bp.Gph.....s.hshhhDhhhh..hhhh.schh.....hhcss.....s..h..h.....

```

HHH → alpha helix

EEE → beta strand



**Supplementary Figure 5. Multiple sequence alignment (MSA) of the beta-sandwich domain associated with the YdjM domain. For coloring themes refer to the Supplementary Figure 3 legend.**

```

Secondary Structure  EEEEEEE  EEEEEEE  EEEEEEE  EEEEEEE  HHHHHHHHH  HHHHHH  EEEEEEE  EEEEE
Tfra_WP_051965836.1  IDALSLVPT--LNFSKMPFIVEKK---EKYVTRAS-LP--STLQVEGVYKYN---SVPPGLLNIALQAE-PIQAQDFARHPMITYKTCG--EGHEVVKF
Ss1l_WP_094604719.1  LTRLLVMESL-KSVFSQDFVVEK---DQYHVQOVGLLQ--QFLTIHKKFKTQ---TLSFATQEAHKTIV-LGERSSEFTPFYSYFTEYQDNHACKVHI
Opfe_WP_054876990.1  IRKIHLLSM-IAFYKFOYILLEG---MSKIVGDINFLY--HRVVFDFIKGQ----DDGIEEHVLRKSR-TAKVYKKEFTPIFHIDFEKIN--DGYKALL
Nbac_MBI5406191.1  SDSVAALQOP-FSPFKWVLLIDTG-DTLYQSFDIDIKG-NDRHDKINMFWKK---WPDSFWQKARKLP-GVGFVLYFARFPVAMVEETD-DGHHLVEF
Nbac_MBI4823699.1  VKLCLPLP---NDFLRWVFAKSN---DGIQVCFADLFT--KRICIQETYPGQ---ADEFAIERSKDDK-VVRSGLYFAKYPYAEVKKQN--GKTVVTV
Ntru_OWZ84931.1  INKMVVLAS-KSFFHDFIVEVD---YIVVYQAPIFR--RAVRIKELSKRIEIEIENSLINKALNSK-IGKLEQFETPHYLVKHKHTS-DNKHILITF
Maby_GCE26688.1  PGHYTLVPT--IRWVMSLVLEEK---DRVMSQEW--LG--GQIVWTVQLDKR---ALEHFAVKASRKDS-KIQALVLESYAYPQVKETE--TCYEVVRW
Fbac_PKM87113.1  DSKLVILP---LRFVMDFIASLP---DKKVVGVDSIK--RKVRILGCFDN----IRNHLQAVLASV-LGKVEVFTFPYHYISCEASE--GKYVARL
Fbac_NFV27839.1  LKRVSVLPV-VGISNMFIVETE---NSMFLGNVDLNR---GVITREVLKPK---QVDFVVTAMRAP-AAEVVLGFAHHPYLYEIKLG--DRYLVKVI
Dlac_WP_073029145.1  KDQVKVIFAM-YHPTWVNFLLFQE---EHVVFQVIRN---QIPSTCRALPKL---TEDNPSVTNLEGN-VAEISGCFPHYHVIAHYLD-NEBCRVVF
Dalk_WP_031517982.1  AGAYSVVEP--LGINWVHCVVEKR(5)KKILVGHIS-TA--GRVVVEREFIVN---PPRYIMDSLKKE-LAQCEISFARHLAVWINKKA--DGYEIVL
Dbac_MBI5328482.1  SSNIEALPL-LSPFRWVVEVDD---ERFYQVNTDILK--YNLINSFEKPK---QANNIIEKWONLE-IVKYLWFAFPVAVVVEKQD--SGYLVVEY
Cend_RKP58239.1  SSQVLIPT--FSLFNWSIVLVGH---DAVQAQVWRY---GNIIWQSPINLG---EAGDRYKCEQTD-AVKLLKNSPYLAVQKITHP--DGIKIRW
Chac_PLX31380.1  IISIDVLPAL-MAFHKNVFLSTN---SHNIAQVNLN--KRITRKKLRRE---DKDLVKVFKGSA-AGQVFTNLSPNYHVTMKSDE--GRITLEG
Chac_NLI60421.1  NQRIVVMSP-LRFRKDFILYQE---KRIITRVSILT--KDVVEVDYLER---SHSKLENVMDTK-IGRPEKESPHYHVEWTKDG--DNHKAIV
Chac_HHZ17469.1  PFRIVVLSGRMSFLRMDPIVKT---LKDIVSIDLGS--GRWRVRELDHP---PEEVREIQILPLPLPGQPSSEFTPLHYIKCEKLG--DKFICHP
Chac_NLT94498.1  DGRITLLEPL-LGLRKNFIVERC---GTYYLQVDSFS--GEITIKEEELFPE---EECPVLSKEDP-EVKIHLDFARFPWYSIEKQD--QDITVTKW
Cfer_WP_028051626.1  EAKITVYDPP-LDFFSWRYIAQEG---KTIHQCAKFLS--GKIITETKLESS---EIPESRVLSTP-EAKASLNFARFVYVVKVDTP--EGRKYYL
Bace_AYO29893.1  LHRVLMESM-FGIFSWIVIVETK---KSLIVSINSFPKLFKVLKELLC---KDHFMVQALNSK-PGEHSEKFTPFYIDLIRKG--EYFYVQF
Pant_WP_138192560.1  GDEYTALETV--NWSVNVVRKNRE--GNVYLVWRSG---RLKWTDEYVCE---SSPLIEKSKRHP-AVASLSESSHPCPEERRG--SETIVRW
Consensus/80%      ..ph.hbPs..bsb.pwbbhhppp...p.bhGhph.h....h.h.p.b.p.....p.hhp.sbps...sphF..Fs.a.hhphppps...bbhpb
  
```

```

Secondary Structure  EEEEE  EEEEEEE  EEEEEEE
Tfra_WP_051965836.1  TDLRYRFRD-YHPPFATVTLNQNLOITGACLGK
Ss1l_WP_094604719.1  YDLRYEFNQ-DFLHSATIIIFDCNQHPCDSYIKT
Opfe_WP_054876990.1  TDLRYELKG-EFLHATIIYDEKMNILEEKENP
Nbac_MBI5406191.1  IDLRFILRGRIPFTHRVEFNEGGVIMEKETT
Nbac_MBI4823699.1  KETSEFSFIY-GEHFVAKVFMDENGKVLSSYKFK
Ntru_OWZ84931.1  IDLRYHIGE-RFLHATVVKCHNQGMIEEELQFP
Maby_GCE26688.1  VDLRYRTRK-HFPFQAIALLDHCVRVIESHVGN
Fbac_PKM87113.1  MDLRYRVKD-RFLHNATLVMSDLKVEEATLQFP
Fbac_NFV27839.1  VDLRYQVRN-RSPPFVVVQLDRNLNVIDSGLGS
Dlac_WP_073029145.1  LDRLYVWVG-DFLYTGNVYLLNNGEISHEIFHP
Dalk_WP_031517982.1  GDLRYRHRQ-QYPPFAAYIKLSNNLEVLDEDLRY
Dbac_MBI5328482.1  EDLRFNALPPRKPFLLSLFLDKNGSLNHAELMF
Cend_RKP58239.1  LDRLYGERA-NYCLMACVLIDQSYTIQRSYIGW
Chac_PLX31380.1  IDLRYEMRD-NFMHATVVIDENLNIVESLQFP
Chac_NLI60421.1  TDLRYVVKG-KYLHRATALFDKLPVHAARQFP
Chac_HHZ17469.1  MDLRYRVEN-RFMHNGVLVLENREVEKAFELP
Chac_NLT94498.1  SDLRYKLR-KEHFTLITAKSKBEIILWKAN----
Cfer_WP_028051626.1  LDRLYHFRG-HYVFGVLIELDRQNKVVKTYLQ
Bace_AYO29893.1  YDLRYLLKD-RFLHSATVIVSEGDLEGLFHP
Pant_WP_138192560.1  LDRLYQHRK-QYPPFVAVVKYNAKPEVFSYVGV
Consensus/80%      hDLRYbhp..pb.a.shhbspsphph.p..b..
  
```

HHH → alpha helix

EEE → beta strand

**Supplementary Figure 6. Multiple sequence alignment (MSA) of the Exonuclease-Endonuclease-Phosphatase (EEP) (PF03372) family.** For coloring themes refer to the Supplementary Figure 3 legend. The sequence, CLas\_WP\_015452829.1, consist of two protein sequences, WP\_015452829.1 and WP\_015452830.1, which originated from the same gene in *C. Liberibacter asiaticus*.

```

Secondary Structure
Clas_WP_015452837.1 -----VAQKEEEEEEVRLVSNINNTSEEQEG-----VSLWKNVSKVHHHHHHHHHHHHHHHRTSDEEEEEEYLLRQYAKHHHHHHHNLD--ADEEEEEEIVFLEEEEEEQEMGSYNEEEEEEAVAKVFPKNTEEEEEEWCIFEEEEEEYSTEEEEEEERLINHSKR--
Clas_WP_015452829.1 -----LAQEEEEEERIEEEEEERIASEEEEEEWNINNTSEKEEEEEEESG-----VALFKNSVIREDDEEEEEEYALLQKYAEEEEEEEQLD--ADEEEEEEIVCLEEEEEEQEEEEEEIGSYEEEEEEELIKREEEEEEVPNDKEEEEEEWDLEEEEEELYSG-----
Claf_WP_083965963.1 -----FAQEEEEEEKEEEEEEMRVEEEEEEASEEEEEEWNINNTSEKEEEEEEESG-----VNIWRNSVIREKSDEEEEEEYDLLRQYAMEEEEEENLN--ADEEEEEEIVSLEEEEEEQEMGSYEEEEEEGVAKEEEEEEKEEEEEEIFEEEEEEPDEEEEEEWEEEEEEQEEEEEEIFEEEEEEYSGEEEEEEDSIWSSIL--
Clso_WP_013461676.1 -----FAQEEEEEERIEEEEEERIASEEEEEEWNINNTSEKEEEEEEESG-----MPLLNKNSVIREDDEEEEEEYDLLRQYAEEEEEEEREEEEEELN--ADEEEEEEIVCLEEEEEEQEMGSYEEEEEEAVAKREEEEEEVPEEEEEEPDEEEEEEWEEEEEELYSG-----
Clam_WP_007556761.1 -----SAQEEEEEEVRLASEEEEEEWNINNTSEISEEEEEESG-----KSILKNSVIREDSEEEEEEYDLLKQYAEEEEEEEIN--ADEEEEEEIVSLEEEEEEQEMGSYNEEEEEEAKREEEEEEIFEEEEEEPDEEEEEEWEEEEEELYAE-----
Cleu_PTL87010.1 -----RAQEEEEEEKEEEEEEVREEEEEEIASEEEEEEWNINNTSEKEEEEEEESG-----HSLKNSIEEEEEEIREDDEEEEEEYDLLRQYAEEEEEEEIG--ADEEEEEEIVSLEEEEEEQEMGSYNEEEEEEAVAKREEEEEEIFEEEEEEPDEEEEEEWEEEEEELYSGSSWENLLL--
Rsp_WP_155133432.1 TGIEEEEEEDGEEEEEETREEEEEEVRLASEEEEEEWNIANEEEEEEHWKEEEEEEEG-----EHLPGRPSAPAREEEEEEASEEEEEEYKEEEEEERIANEEEEEEYEEEEEESAID--ADEEEEEEIVFLEEEEEEQEMEEEEEENGEEEEEEPDEEEEEEAVAKEEEEEEVPEEEEEEPPEEEEEEQEEEEEEYEEEEEEHIEEEEEEVSEEEEEEQREEEEEEFEDDR-----
Esp_WP_071023116.1 -----SAEEEEEEETREEEEEEIAEEEEEETEEEEEEWNINNTSEKEEEEEEESG-----VALRDRAPEEEEEEIREDDEEEEEEYVLLQKYAEEEEEEELD--ADEEEEEEIVEEEEEEALEEEEEEQEMGNEEEEEEPAEEEEEELREEEEEEVPEEEEEEPEEEEEEESEEEEEEEWEEEEEEDMEEEEEEVEEEEEEESEEEEEERYK-----
Hsp_WP_152010586.1 PTANAEEEEEEEELEEEEEEKEEEEEEIGEEEEEESSEEEEEEWNIANEEEEEEAEEEEEEEGPEEEEEE-----IALREEEEEEGHVEEEEEETEEEEEEEEDEEEEEEYEEEEEERIEEEEEELDEEEEEETIEEEEEEKEEEEEEALD--VDEEEEEEILEEEEEEALEEEEEEQEEEEEEELEEEEEEGSEEEEEEEAEEEEEEAEEEEEEEILEEEEEEPD--GYEEEEEEKEEEEEEVEEEEEEFEEEEEEEDEEEEEERCEEEEEETEEEEEESNEEEEEEAENEEEEEECK--
Odia_WP_126539690.1 GVAEEEEEESEEEEEEAFDEEEEEELEEEEEEKEEEEEEVATEEEEEEWNLEEEEEEGWEEEEEEHVEEEEEESKEEEEEEAE(55)GSEEEEEEFDEEEEEETVEEEEEEAVEEEEEETVEEEEEEAAEEEEEEYEEEEEEKREEEEEEAVEEEEEEQEEEEEEIASEEEEEEIKEEEEEEKGEEEEEEID--EEEEEEEVEEEEEEIAEEEEEEFEEEEEEQEEEEEEESEEEEEEGEEEEEEAGEEEEEEATEEEEEEREEEEEEEVEEEEEELPEEEEEEGDEEEEEEGDEEEEEEWEEEEEENVEEEEEECSEEEEEEFG--
uncM_WP_161722589.1 ASAAEEEEEEEQEEEEEEVREEEEEEVATEEEEEEWNIEEEEEEANEEEEEEFEEEEEEWEEEEEEVEEEEEEEGE--EYLREEEEEEPEEEEEEEPGEEEEEEGGEEEEEEQEEEEEEGIEEEEEEFTEEEEEEADEEEEEEYEEEEEERAEEEEEEADEEEEEELAEEEEEEREEEEEEELN--ADEEEEEEIVEEEEEEGLEEEEEEQEEEEEEEMEEEEEEGSEEEEEETEEEEEEEAREEEEEERAEEEEEEPEEEEEEEATEEEEEEYEEEEEEDLEEEEEEIFEEEEEESREEEEEEEEEEEEEEDEEEEEELEEEEEERAEEEEEEDPEEEEEEPS
Hsp_WP_154738790.1 TAQEEEEEETYEEEEEETREEEEEELEEEEEERLEEEEEEATEEEEEEWEEEEEELEEEEEEEHEEEEEEADEEEEEETEEEEEENG-----EGCEEEEEERREEEEEEASEEEEEEDEEEEEEADEEEEEEYEEEEEELEEEEEELEEEEEEKEEEEEEYEEEEEEAEEEEEEQEEEEEELK--ADEEEEEEIVEEEEEEALEEEEEEQEEEEEEEVEEEEEEENEEEEEEEPEEEEEEWEEEEEECREEEEEEIFEEEEEEPDEEEEEEQEEEEEEEWEEEEEEIEEEEEEEISEEEEEEWREEEEEEHDEEEEEEQEEEEEENPEEEEEEPEEEEEEETEEEEEECK--
Blit_WP_090881175.1 -----YAEEEEEEETREEEEEEIAEEEEEETEEEEEEWEEEEEEIEEEEEEEHEEEEEERAEEEEEEEN-----IGEEEEEESVEEEEEEPLEEEEEEAAEEEEEEDEEEEEEEVEEEEEERLEEEEEEGMEEEEEEYEEEEEEADEEEEEERLEEEEEEGE--ADEEEEEEIVEEEEEEALEEEEEEQEEEEEEEVEEEEEEDNEEEEEEGAEEEEEEAREEEEEEVPEEEEEEPEEEEEEEREEEEEEEYEEEEEEDFEEEEEEFFEEEEEESGR-----
Rbac_WP_162765173.1 TGAEEEEEEELEEEEEETREEEEEEVRLASEEEEEEWNLENEEEEEETSEEEEEEEAG-----VSLEEEEEEFPEEEEEEQEEEEEEAEEEEEEEAREEEEEEPDEEEEEELEEEEEEKEEEEEEYEEEEEEQEEEEEEYEEEEEELEEEEEEFEEEEEEQEEEEEEMD--ADEEEEEEIVEEEEEEALEEEEEEQEEEEEEEVEEEEEEDNEEEEEEGAEEEEEEAREEEEEEVPEEEEEEPEEEEEEEREEEEEEEYEEEEEEDLEEEEEEFSEEEEEEVREEEEEEQEEEEEEAFEEEEEEDWEEEEEENR--
Rbac_WP_115039404.1 PAREEEEEEGEEEEEEEATEEEEEEISEEEEEESNEEEEEEIANEEEEEEHHEEEEEEVPEEEEEE-----YEEEEEEAYEEEEEERGEEEEEEEEEEEEEIEEEEEEQEEEEEEWEEEEEETFEEEEEEADEEEEEELEEEEEEAIEEEEEEKKEEEEEEILEEEEEEVDEEEEEEQEEEEEEQ--EEEEEEEVEEEEEEILEEEEEEQEEEEEEELEEEEEEASEEEEEEGEEEEEEELEEEEEEAVEEEEEERDEEEEEEVLEEEEEEGD--YEEEEEEAIEEEEEEIGEEEEEETREEEEEEEEEEEEEYEEEEEEDAEEEEEERREEEEEEDNEEEEEEDN--
Alph_WP_054169531.1 VAALEEEEEEAKEEEEEEDEEEEEEVEEEEEEKEEEEEEVATEEEEEEWNLEEEEEEGWEEEEEETLEEEEEERAEEEEEE-----EEEEEEEINGEEEEEEPNEEEEEEADEEEEEEPKEEEEEERAEEEEEEIEEEEEEYEEEEEEQEEEEEEWEEEEEETFEEEEEEADEEEEEEYEEEEEEFAEEEEEERLEEEEEERYEEEEEEAAEEEEEEKLEEEEEELD--ADEEEEEEIVEEEEEEALEEEEEEQEEEEEEEVEEEEEEDNEEEEEEGAEEEEEEAREEEEEEVPEEEEEEPEEEEEEEREEEEEEEYEEEEEEDLEEEEEEFSEEEEEEVREEEEEEQEEEEEEAFEEEEEEDWEEEEEENR--
Nsp_MXX10824.1 GIEEEEEEANEEEEEEDREEEEEEILEEEEEETLEEEEEEATEEEEEEWEEEEEELEEEEEEEHEEEEEEADEEEEEEDD-----KEEEEEEGCEEEEEERPEEEEEENSEEEEEEQEEEEEEYEEEEEEEDEEEEEERLEEEEEEKEEEEEEYEEEEEELREEEEEETR--ADEEEEEEIVEEEEEEALEEEEEEQEEEEEEEVEEEEEEDNEEEEEEGAEEEEEEAREEEEEEVPEEEEEEPEEEEEEEREEEEEEEYEEEEEEDLEEEEEEFSEEEEEEVREEEEEEQEEEEEEAFEEEEEEDWEEEEEENR--
Aheg_WP_106457830.1 SNEEEEEETELEEEEEEQEEEEEETLEEEEEEATEEEEEEWEEEEEELEEEEEEEHEEEEEEAEEEEEEEKEEEEEENG-----EEEEEEEGCEEEEEERPEEEEEEQEEEEEENAEEEEEEDEEEEEEYEEEEEEQEEEEEEYEEEEEEANEEEEEERL(4)VDEEEEEEITEEEEEEQEEEEEEEVEEEEEEENEEEEEEEAEEEEEEAREEEEEEVPEEEEEEPAEEEEEESKEEEEEENIEEEEEEVEEEEEELSEEEEEEAREEEEEEPSEEEEEESSEEEEEEQEEEEEEPREEEEEECK--
Chac_MYD94332.1 --AAEEEEEEVEEEEEEAGEEEEEEELEEEEEERIEEEEEEANEEEEEELEEEEEEEHEEEEEEEDEEEEEETEEEEEEEGE-----QEEEEEECCEEEEEELDEEEEEEDREEEEEEDDEEEEEEYEEEEEEDEEEEEEALEEEEEEAREEEEEEIVEEEEEEELEEEEEEDE--ADEEEEEEIVEEEEEEALEEEEEEQEEEEEEEVEEEEEEDNEEEEEEGAEEEEEEAREEEEEEVPEEEEEEPEEEEEEEREEEEEEEYEEEEEEDLEEEEEEFSEEEEEEVREEEEEEQEEEEEEAFEEEEEEDWEEEEEENR--
Fdor_WP_026210086.1 EEEEEEEIIEEEEEEHEEEEEEENREEEEEELEEEEEERLEEEEEEATEEEEEEWEEEEEELEEEEEEEHEEEEEEESEEEEEEEDEEEEEEGE-----QEEEEEESTEEEEEEYEEEEEETEEEEEEGSEEEEEERPEEEEEESIEEEEEEKAEEEEEEBDEEEEEEYEEEEEEEREEEEEELEEEEEEKEEEEEEYEEEEEEIREEEEEEMFEEEEEEEN--EEEEEEEDEEEEEEILEEEEEEAVEEEEEEQEEEEEEEVEEEEEEDNEEEEEEGAEEEEEEAREEEEEEVPEEEEEEPEEEEEEEREEEEEEEYEEEEEEDLEEEEEEFSEEEEEEVREEEEEEQEEEEEEAFEEEEEEDWEEEEEENR--
Abac_MXZ35846.1 VCAGEEEEEEAAEEEEEEEEEEEEEKEEEEEEIASEEEEEEWNIEEEEEEYEEEEEETSEEEEEEEDEEEEEE-----DEEEEEEHHEEEEEEGREEEEEEANEEEEEEDEEEEEEEDLEEEEEEKEEEEEEYEEEEEELAEEEEEEAYEEEEEEANEEEEEELD--ADEEEEEEIVEEEEEEALEEEEEEQEEEEEEEVEEEEEEDNEEEEEEGAEEEEEEAREEEEEEVPEEEEEEPEEEEEEEREEEEEEEYEEEEEEDLEEEEEEFSEEEEEEVREEEEEEQEEEEEEAFEEEEEEDWEEEEEENR--
Abac_MYB31970.1 QEEEEEESREEEEEEPNEEEEEENREEEEEEITEEEEEEVATEEEEEEWNIEEEEEEAWEEEEEEGGEEEEEEGIEEEEEER--DEEEEEEGVEEEEEELREEEEEEGNEEEEEEQREEEEEEGREEEEEEFLEEEEEEQEEEEEEEAEEEEEEDLEEEEEERLEEEEEERGEEEEEEVEEEEEESREEEEEEILEEEEEEDLEEEEEEDEEEEEEEVEEEEEEVEEEEEEGLEEEEEEQEEEEEEEVEEEEEEENEEEEEEAAEEEEEEVREEEEEERIEEEEEEFPEEEEEEEDEEEEEEGDEEEEEEWEEEEEEDLEEEEEEFSEEEEEETREEEEEEED--
Ldzo_WP_135756106.1 AKEEEEEEGPEEEEEEIKEEEEEELEEEEEEKEEEEEEVEEEEEENSEEEEEEYEEEEEEFFEEEEEEYEEEEEEDEEEEEEEIGEEEEEEDD-----SEEEEEEHKEEEEEEFPEEEEEEKEEEEEEDREEEEEEHREEEEEEEDEEEEEEFTEEEEEEKKEEEEEEIIEEEEEELSEEEEEENH--EEEEEEEDEEEEEEIIEEEEEEIGEEEEEEQEEEEEEEVEEEEEEENEEEEEETAEEEEEEHLEEEEEEILEEEEEEIN--EEEEEEEVEEEEEEECAEEEEEETEEEEEEATEEEEEEVTEEEEEEP-----
Ladl_PJZ53041.1 ---QEEEEEEDLEEEEEELEEEEEEKEEEEEELEEEEEEASEEEEEENEEEEEEAMPEEEEEEYEEEEEEDEEEEEEEVGEEEEEEDD-----GEEEEEEKKEEEEEEFPEEEEEEKEEEEEERIEEEEEEPKEEEEEESDEEEEEEEEEEEEEFKEEEEEEIREEEEEETHEEEEEELSEEEEEEKLEEEEEEDD--EEEEEEEDEEEEEEIIEEEEEEALEEEEEEQEEEEEEEVEEEEEEENEEEEEETAEEEEEEHLEEEEEEILEEEEEEIN--EEEEEEEVEEEEEEECAEEEEEETEEEEEEATEEEEEEVTEEEEEEP-----
Vtub_WP_004744376.1 GLEEEEEEVSEEEEEEAAEEEEEEPLEEEEEETLEEEEEEITEEEEEEWEEEEEETEEEEEEEHEEEEEEMSEEEEEEEEEEEEEER(36)TYEEEEEECEEEEEENEEEEEEALEEEEEESWEEEEEEPKEEEEEEVPEEEEEEWEEEEEEINDEEEEEEHEEEEEEREEEEEEEKEEEEEEVEEEEEESSEEEEEELEEEEEEQEEEEEESLA(4)SDEEEEEEIVEEEEEEFFEEEEEEQEEEEEEELEEEEEESDEEEEEEAKEEEEEETEEEEEEEELEEEEEELSEEEEEESDEEEEEEKAEEEEEETSEEEEEEISEEEEEESFEEEEEEELEEEEEELK-----
Saby_WP_106608755.1 YDAEEEEEEIDEEEEEEGEEEEEEELEEEEEETLEEEEEEATEEEEEEWEEEEEELEEEEEEEHEEEEEEQEEEEEESG-----EEEEEEEPLEEEEEERMEEEEEERAEEEEEEAREEEEEEEDEEEEEEYEEEEEETREEEEEELREEEEEEDEEEEEEEFAEEEEEETSEEEEEELG--ADEEEEEEIVEEEEEEALEEEEEEQEEEEEEEVEEEEEEDNEEEEEEGAEEEEEEAREEEEEEVPEEEEEEPEEEEEEEREEEEEEEYEEEEEEDLEEEEEEFSEEEEEEVREEEEEEQEEEEEEAFEEEEEEDWEEEEEENR--
Consensus/80% .....hphssWb..L.p..s.....rp..Da.bb..bh.pb.....sDhshbQeh.s..Ah.chbs...aphhs.p.....

```

```

Secondary Structure
Clas_WP_015452837.1 -----DSNEEEEEENDEEEEEEIEEEEEEHTEEEEEEAIEEEEEEVREEEEEEKKEEEEEENVEEEEEERVEEEEEELEEEEEEQEEEEEE-----SYEEEEEEPLEEEEEELEEEEEEGAEEEEEEKEEEEEEADEEEEEE-----SEEEEEEFSEEEEEEREEEEEEAGEEEEEENREEEEEERAEEEEEEVEEEEEEELLEEEEEEVEEEEEEEING--KEEEEEEKEEEEEEITEEEEEEWEEEEEELDEEEEEEIEEEEEEHLEEEEEEKEEEEEESEEEEEEECFEEEEEELEEEEEEDSEEEEEELE-----NTEEEEEEYSEEEEEEFSEEEEEECS
Clas_WP_015452829.1 (14)IDEEEEEETEEEEEEDESEEEEEEINEEEEEETAEEEEEEAIEEEEEETREEEEEEKEEEEEEDEEEEEEVAREEEEEEVLEEEEEEQEEEEEEV-----SYEEEEEELPEEEEEEPAEEEEEEQEEEEEEEEEEEEEI-----TEEEEEEPREEEEEEMGEEEEEENREEEEEEKEEEEEERAEEEEEEVEEEEEEELLEEEEEEIEEEEEEEIDD--QEEEEEEQEEEEEEVEEEEEEWEEEEEELDEEEEEEINEEEEEEVEEEEEEHLEEEEEEKEEEEEESEEEEEEECVEEEEEEVEEEEEEKKEEEEEEIQ-----
Clas_WP_012778745.1 -----SEEEEEENTEEEEEEDKEEEEEEHAEEEEEEMEEEEEEHTEEEEEEAIEEEEEEVREEEEEEKKEEEEEEGAEEEEEEIHEEEEEELEEEEEELEEEEEEQEEEEEEK-----SYEEEEEELPEEEEEEMDEEEEEETEEEEEEEGE-----LEEEEEEDSEEEEEEKEEEEEEAGEEEEEERREEEEEERAEEEEEEVEEEEEEELLEEEEEEFEEEEEEEVDG--REEEEEEKEEEEEEITEEEEEEWEEEEEELDEEEEEEIEEEEEEHLEEEEEEKEEEEEESEEEEEEECFEEEEEELEEEEEEDSEEEEEEIE-----DSEEEEEEYEEEEEEISEEEEEESEEEEEEECY
Claf_WP_083965963.1 (5)KYEEEEEENSEEEEEERSEEEEEENSEEEEEEINEEEEEETAEEEEEEAVEEEEEEVREEEEEEKKEEEEEESMEEEEEEHEEEEEEVLEEEEEEQEEEEEEV-----SYEEEEEEFPEEEEEEVGEEEEEEIEEEEEEEEN-----SEEEEEEHSEEEEEEREEEEEEAGEEEEEERREEEEEERAEEEEEEVEEEEEEELLEEEEEEIEEEEEEEING--QEEEEEEQEEEEEEVEEEEEEWEEEEEELDEEEEEEIEEEEEEHLEEEEEEKEEEEEESEEEEEEECFEEEEEELEEEEEENKEEEEEELE-----NEEEEEETREEEEEENSEEEEEECS
Clso_WP_013461676.1 -----NEEEEEEDSEEEEEEDEEEEEEEHTEEEEEEVEEEEEEHTEEEEEEAIEEEEEEVREEEEEEAKEEEEEETEEEEEEGVEEEEEEHEEEEEELEEEEEEEK-----SYEEEEEELSEEEEEEMDEEEEEETEEEEEENEEEEEEK-----LEEEEEEDSEEEEEEKEEEEEEAGEEEEEERREEEEEERSEEEEEEVEEEEEEELLEEEEEEFEEEEEEEING--IEEEEEEKEEEEEEITEEEEEEWEEEEEELDEEEEEEIEEEEEEHLEEEEEEKEEEEEESEEEEEEECFEEEEEEVEEEEEEDSEEEEEELK-----DEEEEEEAYEEEEEETLEEEEEESCY
Clam_WP_007556761.1 -----NEEEEEEDSEEEEEEGDEEEEEE--SEEEEEEHTEEEEEEAIEEEEEEVREEEEEEKEEEEEEEKEEEEEEIKEEEEEEIEEEEEEEK-----NYEEEEEEISEEEEEETEEEEEEINEEEEEEEN-----DEEEEEEIPEEEEEEVMEEEEEEGSEEEEEERAEEEEEEVEEEEEEELLEEEEEEFKEEEEEEVEEEEEENG--KEEEEEEKEEEEEEVEEEEEEWEEEEEELLEEEEEECVEEEEEEHLEEEEEEKEEEEEESEEEEEEECHEEEEEELEEEEEEDEEEEEEHLA-----SEEEEEEYNEEEEEEKSEEEEEECA
Cleu_PTL87010.1 -----KEEEEEENGEEEEEENEEEEEEGEEEEEEQEEEEEENIEEEEEEHTEEEEEEAIEEEEEEVREEEEEEKKEEEEEEDTEEEEEEIKEEEEEELEEEEEEKT-----SYEEEEEELSEEEEEEMEEEEEENTEEEEEEEG-----MEEEEEEYEEEEEEQEEEEEENREEEEEEGSEEEEEERAEEEEEEEMEEEEEELEEEEEEEFEEEEEEEVEEEEEENG--KEEEEEEEEEEEEEFEEEEEEWEEEEEELLEEEEEEGVEEEEEEHLEEEEEEKEEEEEESEEEEEEECFEEEEEELEEEEEEDREEEEEELE-----NEEEEEEQEEEEEEYEEEEEENSEEEEEECS
Rsp_WP_155133432.1 -----DEEEEEEQEEEEEEEREEEEEEDTEEEEEEDHEEEEEEITEEEEEEAVEEEEEEVEEEEEEKKEEEEEESGEEEEEESVEEEEEESIEEEEEEAGEEEEEEQ-----TYEEEEEENEEEEEEALEEEEEESIEEEEEEIEEEEEEHEEEEEEDP-----DEEEEEEGDEEEEEEAREEEEEEPTEEEEEERAEEEEEEGTEEEEEEEVEEEEEELEEEEEEELPEEEEEENG--EEEEEEEPLEEEEEEQEEEEEEVMEEEEEESVEEEEEEHLEEEEEEKEEEEEESEEEEEEECHEEEEEEAGEEEEEESLEEEEEEDD-----NEEEEEEPREEEEEENEEEEEEEDEEEEEECV
Esp_WP_071023116.1 -----PEEEEEEGTEEEEEEDDEEEEEEPDEEEEEEIFEEEEEETAEEEEEEAVEEEEEEVEEEEEEKKEEEEEEGREEEEEEATEEEEEEIVEEEEEEEGEEEEEEK-----DEEEEEEYEEEEEEEEEEEEEPLEEEEEEGVEEEEEEIN-----PEEEEEEGDEEEEEEGREEEEEEPTEEEEEEVREEEEEERGEEEEEEIQEEEEEEVLEEEEEEEAEEEEEEDE--KEEEEEEREEEEEEFEEEEEEWEEEEEELLEEEEEEGVEEEEEEHLEEEEEEKEEEEEESEEEEEEECFEEEEEEPAEEEEEEKSEEEEEELT-----QEEEEEEPEEEEEEVEEEEEESDEEEEEECK
Hsp_WP_152010586.1 -----KEEEEEEDEEEEEEVEEEEEEDDEEEEEEIFEEEEEETAEEEEEEAYEEEEEERGEEEEEEDKEEEEEEFPEEEEEEDDEEEEEEKEEEEEEIEEEEEEEQEEEEEEI-----DEEEEEEDLEEEEEEDIEEEEEEHEEEEEEQEEEEEESEEEEEEECT-----TEEEEEEEEEEEEEEAREEEEEEKEEEEEEVREEEEEEGAEEEEEEPAEEEEEEVAEEEEEEKTEEEEEEWEEEEEEDG--QEEEEEEKEEEEEEYEEEEEEAVEEEEEEPSEEEEEEHLEEEEEEKEEEEEESEEEEEEECFEEEEEEGREEEEEENSEEEEEEASEEEEEESN--TEEEEEEPDEEEEEELEEEEEEADEEEEEECV
Odia_WP_126539690.1 -----YEEEEEEKKEEEEEEVEEEEEEQEEEEEERLEEEEEEGEEEEEEEFAEEEEEEWREEEEEEKSEEEEEESLEEEEEEGDEEEEEEAVEEEEEEET-----CEEEEEEEPEEEEEEDDEEEEEEALEEEEEESLEEEEEEPEEEEEEEV-----PEEEEEEDKEEEEEEDREEEEEEVREEEEEEPGEEEEEESLEEEEEEGLEEEEEEKGEEEEEEVDG--KEEEEEELEEEEEEVEEEEEEREEEEEEFLEEEEEESVEEEEEEHLEEEEEEKEEEEEESEEEEEEECFEEEEEEVEEEEEESLEEEEEEGAEEEEEEE(8)AGEEEEEEQEEEEEEDPEEEEEEACEEEEEEQ
uncM_WP_161722589.1 KLAN-----YEEEEEEPKEEEEEEKREEEEEEDIEEEEEEYEEEEEETEEEEEEALEEEEEELREEEEEEKKEEEEEEDKEEEEEEVEEEEEEKLEEEEEEEREEEEEEIE-----PEEEEEEIAEEEEEEEEEEEEELEEEEEEAIEEEEEEIT-----DEEEEEEREEEEEEEGHEEEEEEPTEEEEEERDEEEEEEGTEEEEEEAAEEEEEELEEEEEEEVEEEEEEEAG--REEEEEEREEEEEEFEEEEEEAVEEEEEESLEEEEEEHLEEEEEENSEEEEEECAEEEEEERIEEEEEEEDEEEEEEPA(7)PNEEEEEELEEEEEEEEAEEEEEECE
Hsp_WP_154738790.1 (3)---AEEEEEEPMEEEEEEITEEEEEEQEEEEEERTEEEEEEGEEEEEEEFAEEEEEEIREEEEEEGIEEEEEEPYEEEEEETREEEEEENP-----DEEEEEEVEEEEEETEEEEEEALEEEEEEDV-----GEEEEEEGTEEEEEENREEEEEEHREEEEEEGDEEEEEEVDEEEEEEITEEEEEELEEEEEEEAG--VEEEEEEPEEEEEEIREEEEEEMLEEEEEESVEEEEEEHLEEEEEEKEEEEEESEEEEEEECFEEEEEEADEEEE
```



```

Secondary Structure
Clas_WP_015452837.1 HHHHHHHHHHHHHHHHHHHHHH EEEE HHHHHHH EEEEE HHH
LLSQCAQWLKDWLTKKESLVPFVLAQDFNKKINYLGN---DFWKTIDPNDL-----IRFPKEDSRCN---ANKNLRNKPIIDYFVMDQNAV
Clas_WP_015452829.1 MLLSQGEMLKKWADQKIKTGIPEVLAQDFNKKINSIGDT---DFWQKMDPDGLL-----IRFPQEKESTCN---VKRNKSSLDYFVIDRDNK
Clas_WP_012778745.1 MNLQCATWLLKQWVQKNNLNPFIILAQDFNKKINSHSG---IKDELWQKIQNDNTL-----MRLPHKKN-----
Claf_WP_083965963.1 LLNKQSKWLKLDQDKKKTGIPFIIAQDFNKKINYLEDK---DSFWQKINPDNSL-----IRFPQOKETS CN---ANKNLLNKPTIDYFIMDKNAV
CLso_WP_013461676.1 TLNLQVNLNLRWHLQKRSNIPFIIAQDFNKKINHFQDN---DELWGKISKDTIL-----IRVFNKKRSCN---AHKSIRKRPEIDYFVMDQNAV
Clam_WP_007556761.1 LLNKQAIMLNKRNKRNKNEENPFIIISGDFNKKINRYGED---DEFWQKISNNTNL-----IRFPYKRSKCN---ANHYKNSEPIDYFIMDKNAV
Cleu_PTL87010.1 LLKLDLWLLKRWLNKKNKASNIPFIIAQDFNKKMNNSGN---DEAWLEINQNNL-----IRYPSYKRSKCR---TASKYKNSPIDYFVMDQNAV
Rsp_WP_155133432.1 TLANQRAELEAWDELTEEDNTPFIIAQDFNKKINRDRFAQN---DHLWGEIDDNSPAP---LDLARFPEGITSECL---AGTSGHFPPIDYFVLDQVAL
Esp_WP_071023116.1 TLRAMQKPLPEWIDEKEATGLPVVLAQDFNKKIPDVHGQN---DHLWGEIDDADPASLDL---IRTFREASRCP---TTRESDRPEPIDYFIVTNKLAT
Hsp_WP_152010586.1 TQAKCIDRLIEHWLCEPEQAGYRFFILICDFNKKINRLEGTSS---DTIRKRLEAVDANV-----RFEFQERVYCW(28)WAYTPDSGNDIDYFVVLGVPS
Odia_WP_126539690.1 VLQOQVALEAWTERKAAGAARIVVLGDFNRRNLPHEIAA(21)ASRIGEVDGEPAA(6)EAHCPVTPDAEAA(17)ENLGCNRPGLDHYLHVLGSGVL
uncm_WP_161722589.1 MLVRQPLLEFRMIDVRERAGEDFVVLGDFNRRVDRAN-----EFLWRDLDLDRDPQT---LDLFMVPYRQALVCS---GHSPEPDRSHAYVILNKRRLW
Hsp_WP_154738790.1 PLRDQSKVLSWTEARRKDGLEFVVLGDFNRRRLQNEE---EVVGLLGVRSGLT-----LSVREAVSRCH-----AWTKDFIIDYFVLDQKSK
Blit_WP_090881175.1 RLKQCPVLEWIDWADARIEGVFAVILGDFNRRFEIDENLPSE---DSFWQDIDDGKPKGLDL---SRVTAGRTSECW---GGRHPDFIIDYFVLDQIAY
Rbac_WP_162765173.1 LLRAQAEIINNMLEKNIKAGASTILGDFNRRNIFRSEDPFT---SELRAVRSSGTYK---RIVPSSGVTCPC---TRKFRIIDYFVIAVGGIV
Rbac_WP_115039404.1 LLYKQSLIGDLKKKAGQDDAVLVVGDFFNRRLLNPR---LLEWKRKYAPWAAS-----LIAPPACI-----LRPSLOPIDYFVILSKLPT
Alph_WP_054169531.1 VVIAQMPILSKWIDARQAEGLSLLVAGDFNRRLRDE---DPLWKVLDNGPSPFL-----KLAITAGKANHCW-----AGEVPEPIDYFVILGGPVV
Nsp_MXX10824.1 LLNKQINVLSSMNRDRSCQNFIIAQDFNRRQLDQLS---DAMWRLLETGCKPG---TYVDLEKALHGVKHP---KPNYKRYVPIIDYFVYMQALD
Aheg_WP_106457830.1 ELATQPLLEKRWIDRVRKGERFAIAGDFNRRNLNIPG---DEFWQEIDDSEPTD---GADLETTIETRGLSSCN-----RYKDYIIDYFVLDKGVY
Cbac_MYD94332.1 TLRAGVNLKAWADARREEGAFFVILGDFNRRNIALPG---DVAWVLSPPAAPL---RLITQGVAFRCN---PRYPAFIIDYFVIAVGGGAT
Fdor_WP_026210086.1 TLMDCQVLEWIDWADARIEGVFAVILGDFNRRNRLDLS---DRVWSEIDDAEPAN---ADLTTVTLNMPFLSCR-----DNRYTEIIDYFVLDKKSF
Abac_MXZ35846.1 ELKSOVRLEWIDRERGRSEFVILGDFNRRNRRDRG---DPLWRFLADGNPKP---VYRVNAGKRPKCW---GSRFPFVIDYFVIVGPIAG
Abac_MYB31970.1 ELRMOFAVLEWIDRERMAEIEPTVVAQDFNRRFLSRPA---DEAAAGLEGEAL-----VLPFAGSAPACW-----DGFDFHYVIDYFVIVFAPRG
Ldzo_WP_135756106.1 LRKQCLALPKDILPSL---GKFFVILGDFNRRVLERR---VDLLKILQRLK---KVANYKQKSCW---QHNDFIIDYFVILNMDWQ
Ladl_PJZ53041.1 LRSRQVLEWIDRERKILKIK---KNYFVILGDFNRRVPLGKD---NRSWNLLSKGLDL---KNFGRYTKQICW---GHESLFIIDYFVILDLAIS
Vtub_WP_004744376.1 TLYAKVLEWIDRERKMTEDGIDFVILGDFNRRVFKVDW---SYDARSDDGSDAQ---SSPTKDSKTRSF(23)KFSDWNCYKIDRFFLSKAFS
Saby_WP_106608755.1 TLAQVLEWIDRERKMDRGRGNVVLGDFNRRNLDGTGASADVWDLNLDKGGAP---LVKGEISDDPDCW---PNHGHRVSHIDYFVILGDRLV
Consensus/80% hb..Qh..Lp.Wbp.p...s.sbhbgDFNRpb.....s.hb..bs.s.....s.....C.....IDabhhs....

```

```

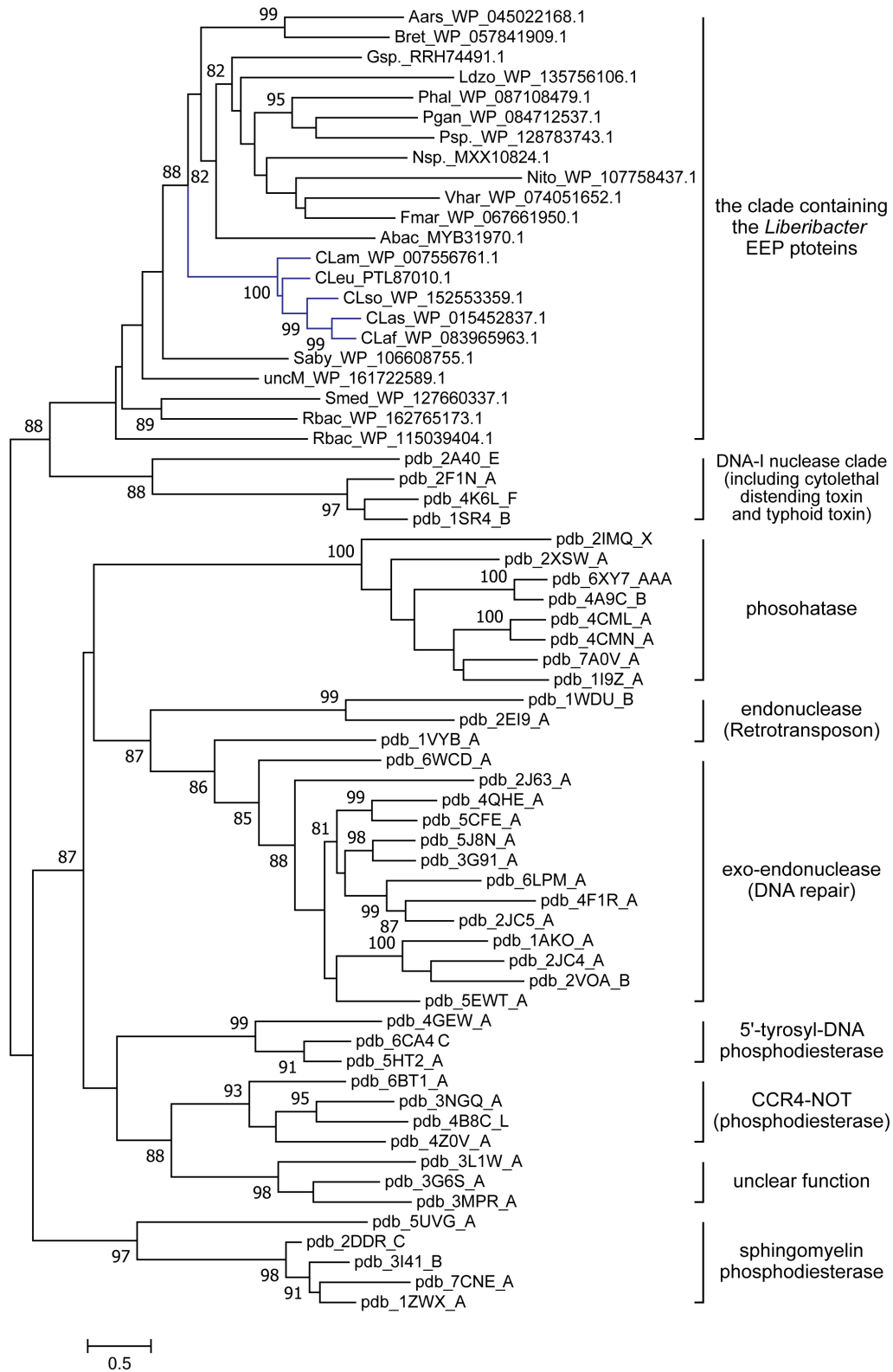
Secondary Structure
Clas_WP_015452837.1 HHHH EEEEE EEEEE
KFLIQES---FSEILYNEDDIKS---RGKRLSDHCPHSIDYDF
NELLDNS---FSIVSYDQSDLDT---RRSKLSHCPHTIIEYDF
Clas_WP_012778745.1
KLLIKNS---FSEVLYDSNDIQI---KGNRLSDHCPHTIDYDV
KYLIENTS---YSEVSYIEDDIK---RGYRLSDHCPHTIDYDF
QYHVEGS---FYFESYRNQDIKE---RRYNLSDHCPHTIIEYDF
RYPFKGS---FSEFLYRDQDIKE---IGRRLSDHCPHTIDYDF
RWADDRDS---FQIVDFKPNDRP---HRKKLSHCHALYIDLVI
EALIDQS---FVELLYDETEAAE---LGNRLSDHCPKSIDKFK
DIQISAE(32)NKVLVFSF-----ANPESDHCPHRLMELGS
SDGANKV(5)-GRSLPADQNHLD---PLLAVSDHCPHRAVVRV
AWAEPSSD---SPKLDV-----TGKNVSDHCPVFLDLRL
AFAEFTH---FAELKFAEP-----EAKYPSDHCPVSVDTV
GWVSSGS---FEQLVYDEAD-EQ---FDSVLSHCHPSTVLDI
DAIIESS---VREYIPTYERI---FGAKVSDHCEMSVDLFL
GGTIAGS---STLSNLLPTLD---HPARISDHCPHSLRVDL
RRAKEGS---EGVLVYPTETSRD---DKKRVSDHCPVSTIIR-
GLIV-----EAKTSPDA-----KGGKYSDHCPHPLFAVFDL
RAVSERS---FQQLVYKEA-----ESEHPSDHCPHGVITDM
VVPGS-----FREMPP-----RRGFHDHCPVYADFLR
EWDVLS---FRQVNYRQADKA---DWDKLSHCHPSTVILTV
KAVKPNs---FEELYSERNYFG---WRKRLSDHCPHSTIRLVF
REIAPRA---FEEVRYRA-PRS---ARDRLSDHCPHVALTGF
QGS-----FVQTKFESDDGNF(5)AEKGLSDHCPVSAIDL
NAD-----FRQFPFEDDGEF(5)SERGLSDHCPVLELAL
ETLIWKT(5)EAHGKNYPN-FSK---DKVKPSDHCPHTLPL
SEAQMST---ARKIGYDAYDDPT(8)GRYRLSDHCPVILTLGG
Consensus/80% .hh.....h.h.h.....hSDHCP.hpb.h

```

HHH → alpha helix

EEE → beta strand

**Supplementary Figure 7. Phylogenetic relationships of the *Liberibacter* EEP related sequences and other known EEP enzymes.**









**Supplementary Figure 9. Multiple sequence alignment (MSA) of the Phage\_lysozyme (PF00959) family. For coloring themes refer to the Supplementary Figure 3 legend.**

```

Secondary Structure  HHHHHHHHHH  EEEEE  EEEE  HHHHHHHHHHHHHHHHHHHHH  HHHHHHHHH
Clas_WP_015452887.1  IPVPNALIKMLKEFEG---LRLTAVRDIIGG---SVWTIGYGHGTS---DVTEGMITTEKBEEDFLKQASKSLNLLLESSPAL---KSTSENRLVAVADF
Claf_WP_047264530.1  VLIPLGLLVDVVERYEG---VRVTAMKDPGR---SVWTIGYGHGTP---DVHEGMITTEQEAELKQGLRKRHVLDLLSASPIM---ASVSENRLAVGDF
Clso_WP_013461579.1  --MPHLLDLVLKGFEG---LRLKAVRCSA---SVWTIGYGHGAGN---DVFNLAITEKQANDLAKWVSKKLSQVFTVSPIL---INAGENRISAVGDF
Clso_WP_013462155.1  IIPPLLNLLIKRFEG---LRLKAVQCPA---SVWTIGYGHGAGN---DVFKDLVITEQKAEESLQKQVLFKLTQVFKISPSL---IDAGENRISAVGDF
Clso_WP_044054126.1  -----MKRFAG---LRLTAVRCSA---SVWTIGYGHGAGN---DVFNLAITEQCADDLKQWVSKKLSQVFTVSPIL---IHAGENRISAVGDF
Clam_WP_007554787.1  ITIPQLLISLLIKQFEG---LRLKAVRCSA---SVWTIGYGHGAGN---DVVDGLVITEQCAENLNLYDIRRHLNVTIKLSPSL---LSDNDCLSAVGDF
Clau_PTL86076.1     -----MLLDLLEFEG---LRLKAVQCPA---SVWTIGYGHGAGN---DVVEGMVITQPOVYEFLLQDASQCIRQTFEVSPII---KECGANRISAVGDF
Cleu_PTL86791.1     TTIPSLLDLVLKNEFEG---LRLKAVRCSA---SVWTIGYGHGAGN---DVCEDMNISQEADELLELILQVSKCLNQVLVNVSPII---INANENRISAVGDF
Abau_6ET6_A        KTTSDDDGIDLITSEFEG---TRFNAMQDGV---SVWTIGYGHGTPVYNGVKVKGGDTCTAEQAKTYFKHDLAKFEKTVNESVT---APLTCNQFDALVSL
Khan_WP_048859119.1  MNLPTYLIQLVSESEFEG---LRLKAVQDCC---SVWTIGYGHGAGN---IKPGQAITPEQAEWILSLLEAIWQAITARSTIQ---LETYQTALADF
Atro_WP_039904868.1  MDIPQTLVELVVEFEG---FKAEAYQDIG---SVWTIGYGHGTPD---LHKGTTFTKEBALQRLSETLQSIWNQIEQKTSIA---LTSKQMBALSSF
Dpig_WP_204673942.1  AFDGPDLLITTKHYEG---FKAHFVLCPPA---SVWTIGYGHGTEGV---SRNPPVFNKEBAEQLIKDLVATHAVLRHAGPA(6)PEEAYRRLVALASW
Dsp_WP_009303990.1  DFNHMEATVAVKHYEG---LRLVPKKCPA---SVWTIGYGHGCLCP---PDORPISKEBAERLRLQWMLDAFGAMLRQWPEG---ADAPAHQLALTSF
Pshi_WP_010864476.1  LKVSDEGLKLIADFEG---CRLTPYRCSA---SVWTIGYGHGTEGV---VTSKSVTEROVAENLVAVAR-TEKALAH-CMP---VTMPOFVYDVAVSW
Pzuc_PZQ57970.1     -AIIAAAVKLAKFEG---WRARFVLCPPA---SVWTIGYGHGTEGV---VKFTNPFVREHCEAVISEMDT-FAGALLAYSPN-LKQHPDVGCGALDF
uCau_CAB4164981.1   -MIPDLAIALIARFEG---FSAKPYLCPA---SVWTIGYGHGCLCP---PDHPPVTEEBQALLIGYLER-VAPHILVQAPA---LENSEARLAVLSW
Maer_NCR22735.1     --IPQRAIADLKEFEG(7)GRVOAMPDPN(5)KLPTIGYGHGTRCP(4)-VRKQIITRDEAEYLMYLDKIRRPALCKIPTW---KDMNENORALYSF
Lsp_NEQ54273.1     --VPQAAADLKEFEG(9)DRVKAMPDPG(4)SLPTIGYGHGTYYP(4)-VQQNDIITRESEERFLTHEVEEKCRLALETIPTW---GGMNFORALYSF
uCau_CAB4127944.1   ---MKDLILLIKRFEG---CRLKAVFCPPA---SVWTIGYGHGTEGV---ISIGDIWTEQCDKRLQDQTYK-FFVGVSKLIPN---ASDNKFIACSDF
Bbac_MBE6304922.1   ---MENASILLIKRFEG---LRLSAYRCPA---SVWTIGYGHGAGN---VRLPMSVSHABADELLRNQVSE-IFTKLQLSAE-ANVTLSNQLSALTSF
Mbac_GFI06907.1     ---MIYNLIRRFEG---LRLHAYRCEG---SVWTIGYGHGTEGV---VNAGTIVTEDEAMRLIREQVAE-VRGRIRHMCQAQ-DGVTLSAVQADALTSF
Bbac_MBD5239742.1   ---MKDLIKRFEG---LRLSAYRCPA---SVWTIGYGHGTEGV---VTAGMITTEBAERLLEDELRP---VVAALRPG---LSDNQRALASL
Bbac_MBE6293835.1   -KISDVAEFKIKDFEG---LRLKAVLCSPG---SVWTIGYGHGTEGV---VHEGMVITRQBAERLLEDELLY-FENFLSKRYA---EDITQOWDALVSF
Bbac_NLA94747.1     MKATNYATLLIKSEFEG---LRLNSVLCPPA---SVWTIGYGHGTEGV---VNOGMILTEKPADAFRQDTHH-AEAEVNRIDAD---LTCQFDALVSF
uCau_OJV19870.1     MKVSQCGVOLLKESEFEG---LRLQAVRCPA---SVWTIGYGHGTEGV---ARAGMVTADKAEQLQEDIRD-TEGLTNSGLL---LTCNQFDALVSF
uCau_CAB4127944.1   MEISKNGLDLKEFEG---CRLKAVLCPPA---SVWTIGYGHGTYYP(4)-VKAGDVITEERAVGLKLVAN-FYHEIKG---INCNQFDALTSF
Bsp_WP_002562728.1  MKTSESAQAIGTFEG---LRLKAVLCPPA---SVWTIGYGHGTEGV---VHKGMQITKEQALAFSLSSLDAD-VERNLTFRFSP---ISCNKFDALISL
Bbac_MBE6322682.1   VRANDEAKALIKSEFEG---LRLKAVKCPA---SVWTIGYGHGTEGV---VTQGMELISQVAEQYFEOQLVYNAEYFISDIFYK-AKVTLTCNQFDALVSF
Consensus/80%      .....hhpbppaEG...hchpsYbss...GHWITIGYGPt.....h..s..hspppa..bL..Dh.p.h...bp..s.....hs.spbsAbssa
    
```

```

Secondary Structure  HHH  HHHH  HHHHH  H  HH  HHHH  H  EE  E  E  HHHHHHHHHHHHH
Clas_WP_015452887.1  VFNLCIGNYR---KSTFKQRV---DAQDWE-KAAE-ECKKWTKA---GGK---VLEGLVRRRDAEAKLLLE--
Claf_WP_047264530.1  VFNLCIGNYR---KSTFKKCI---DAEDWE-NAAT-ECKKWVHG---GGK---VLEGLVRRRNEBAALLISG--
Clso_WP_013461579.1  VFNLCIGRYR---NSTLRRKV---DREDWI-NASH-EICKWVFA---GGK---KLGGLVRRRIBADALLLKSE-
Clso_WP_013462155.1  VFNLCIGRYR---KSTLRRKV---DVGDWK-SASD-ECKKWVFA---GQK---KLGGLVRRRKBVADALLKKG-
Clso_WP_044054126.1  VFNLCIGRYR---ASLRRKCV---DAEDWV-TASH-EIRKWVFA---GGK---KLGGLVRRRCEVBAELLKKN-
Clam_WP_007554787.1  VFNLCITRYK---NSTLRRCV---NSQOME-EASQ-ECRRWVYV---GGK---RLKGLVRRRBAIBAEALLGG-
Cleu_PTL86076.1     VFNLCIGNYR---SSTLRRKV---DVKDWD-GAVD-EIKKWVHV---GKK---TLKGLVRRRNSBACALLMG-
Cleu_PTL86791.1     VFNLCIGRYK---ASTFRKRV---DEEDWD-EASK-ECTRWVFA---GGR---KLGGLVRRRIBADALLLQC-
Abau_6ET6_A        TVNLCGSAFV---NSTLLKLI---NKGDIQ-GAAD-QFLVWVKA---GGK---VMKGLVRRRBAERLFLK--
Khan_WP_048859119.1  AYNLGCSALF---KSTLWQVL---QANQPI-NASL-EFSKWVYA---GKH---IEFGLVRRRKBALFLG--
Atro_WP_039904868.1  AYNLGCSALF---KSTLWKYL---EGEBIE-QAAQ-EFPKWVSHV---NGK---VTIGLKRVRVKEQLFLS--
Dpig_WP_204673942.1  TENLCEGNLS---SSTMLKRI---RVLDWD-AAAN-EMLRWVKA-TVGGR-TVALLGLVRRRKBTAHYFRIT--
Dsp_WP_009303990.1  VFNLCIGNEA---SSTLQRLI---RARKFS-DVPA-QIRRWKNTVDGVK---KRLKGLVRRRQCBALCFEAGE-
Pshi_WP_010864476.1  GFNVCPSAAC---RSTLRAHFV---NQRNWS-QACQ-QLSRWVYV---GGV---KNAKGLVRRRQRELAHCLR--
Pzuc_PZQ57970.1     VFNLCITAYK---ASTLRRKV---AAEDWE-EVEY-QLQRWVFG---GGR---KLFGLVRRRKBABEQDQGA--
uCau_CAB4164981.1   VFNLCIGNFT---ASTFRRCI---AAQDWP-GAAF-QCRWVRA---GGK---RLKGLVRRRREVBARLLCG--
Maer_NCR22735.1     SFNLCGSGFYV(9)VCDSQRW---NDQEWV---KE-QFVKYRNP---GTN---VESGLVRRRREBAAKLFCT--
Lsp_NEQ54273.1     AYNLGAGFYG(9)VCDSQRW---NDQQWV---AE-QFGKYVRS---NNE---VLEGLVRRRBAEAQLFCK--
uCau_CAB4240787.1   AYNLGAGFAK---SSTLRRKI---LAGDIE-GAKM-QLARWVNSK---GGK---ILKGLVRRRQAEIAYI---
Bbac_MBE6304922.1   IENVCESAER---RSTLWKLII---RENPDNPDIAE-QFAKWVYA---AGK---VMSGLVRRRREBATALYFK--
Mbac_GFI06907.1     AFNVCTEARL---RSTLWLSLV---KRDPGDPDIAE-EFSRWVYV---GGR---RLKGLVRRRREBAAALYFS--
Bbac_MBD5239742.1   CFNLCISAFV---RSTLRRKV---CANPADPSIAA-EFLRWVYK---RGR---ELKGLVRRRBAEAVYMS--
Bbac_MBE6293835.1   IENLCIGNFI---SSTLRNKL---INIDEP-TIPD-EFRWVFA---KQK---KLGGLVRRRREBAAQMYEW--
Bbac_NLA94747.1     VFNLCIQAFR---VSTLRKLV---ERNPSDFKID-EFRWVVA---GDK---KLFGLVRRRREBTKLYYS--
uCau_OJV19870.1     IENFGAKPFL---SSTFYRYAKA-SSSDPR---IPA-ELRRWVHS---TDD(5)VQGLVRRRREBATALYK--
uCau_CAB4127944.1   CFNLCIGNFN---ASTLKKKI---VANPNDFSTEA-EFLRWVKA---GGK---VLEGLVRRRREBAAALYFSK--
Bsp_WP_002562728.1  SENLHCQAFN---SSTLYRKAKA-NLNDPS---IRI-EFMKWVHS---RGK---VLEGLVRRRREBAAALYFS--
Bbac_MBE6322682.1   VYNCNCGNEK---KSTLKKKAL-ANFND---KSIYCFEMWVKA---GGK---EVLKGLVRRRREBAAALYFK--
Consensus/80%      sanBg.s.a.....STbbpbh...p.p...h...ph.cwSbs...ssc...b.GLSbRRp.Eh.bbh...
    
```

HHH → alpha helix

EEE → beta strand



## Reference

- 1 Lin, H. *et al.* Complete genome sequence of a Chinese strain of “Candidatus Liberibacter asiaticus”. *Genome announcements* **1** (2013).
- 2 Duan, Y. *et al.* Complete genome sequence of citrus huanglongbing bacterium, ‘Candidatus Liberibacter asiaticus’ obtained through metagenomics. *Molecular Plant-Microbe Interactions* **22**, 1011-1020 (2009).
- 3 Zheng, Z., Deng, X. & Chen, J. Whole-genome sequence of “Candidatus Liberibacter asiaticus” from Guangdong, China. *Genome Announcements* **2** (2014).
- 4 Zheng, Z. *et al.* A type 3 prophage of ‘Candidatus Liberibacter asiaticus’ carrying a restriction-modification system. *Phytopathology* **108**, 454-461 (2018).
- 5 Katoh, H., Miyata, S.-i., Inoue, H. & Iwanami, T. Unique features of a Japanese ‘Candidatus Liberibacter asiaticus’ strain revealed by whole genome sequencing. *PLoS One* **9**, e106109 (2014).
- 6 Dai, Z. *et al.* Prophage diversity of ‘Candidatus Liberibacter asiaticus’ strains in California. *Phytopathology* **109**, 551-559 (2019).
- 7 Lin, H. *et al.* Complete genome sequence of “Candidatus Liberibacter africanus,” a bacterium associated with citrus huanglongbing. *Genome announcements* **3** (2015).
- 8 Lin, H. *et al.* The complete genome sequence of ‘Candidatus Liberibacter solanacearum’, the bacterium associated with potato zebra chip disease. *PLoS One* **6**, e19135 (2011).
- 9 Wulff, N. A. *et al.* The complete genome sequence of ‘Candidatus Liberibacter americanus’, associated with citrus huanglongbing. *Molecular Plant-Microbe Interactions* **27**, 163-176 (2014).
- 10 Frampton, R. A. *et al.* Draft genome sequence of a “Candidatus Liberibacter europaeus” strain assembled from broom psyllids (*Arytainilla spartiophila*) from New Zealand. *Genome announcements* **6** (2018).
- 11 Leonard, M. T., Fagen, J. R., Davis-Richardson, A. G., Davis, M. J. & Triplett, E. W. Complete genome sequence of *Liberibacter crescens* BT-1. *Standards in genomic sciences* **7**, 271-283 (2012).
- 12 Mol, C. D., Kuo, C.-F., Thayer, M. M., Cunningham, R. P. & Tainer, J. A. Structure and function of the multifunctional DNA-repair enzyme exonuclease III. *Nature* **374**, 381-386 (1995).
- 13 Romier, C., Dominguez, R., Lahm, A., Dahl, O. & Suck, D. Recognition of single-stranded DNA by nuclease P1: High resolution crystal structures of complexes with substrate analogs. *Proteins: Structure, Function, and Bioinformatics* **32**, 414-424 (1998).
- 14 Buggy, J. & Bauer, C. E. Cloning and characterization of senC, a gene involved in both aerobic respiration and photosynthesis gene expression in *Rhodobacter capsulatus*. *Journal of bacteriology* **177**, 6958-6965 (1995).
- 15 Xu, Q. *et al.* Structural and sequence analysis of imelysin-like proteins implicated in bacterial iron uptake. *PLoS One* **6**, e21875 (2011).
- 16 Foster, P. G., Nunes, C. R., Greene, P., Moustakas, D. & Stroud, R. M. The first structure of an RNA m5C methyltransferase, Fmu, provides insight into catalytic mechanism and specific binding of RNA substrate. *Structure* **11**, 1609-1620 (2003).
- 17 Busam, R. D. Structure of *Escherichia coli* exonuclease I in complex with thymidine 5'-monophosphate. *Acta Crystallographica Section D: Biological Crystallography* **64**, 206-210 (2008).
- 18 Scheerer, P. *et al.* The class III cyclobutane pyrimidine dimer photolyase structure reveals a new antenna chromophore binding site and alternative photoreduction pathways. *Journal of Biological Chemistry* **290**, 11504-11514 (2015).

- 19 Morita, M. *et al.* Antibacterial activity of *Bacillus amyloliquefaciens* phage endolysin without holin conjugation. *Journal of bioscience and bioengineering* **91**, 469-473 (2001).
- 20 Antonyuk, S. V. *et al.* Structure of SurE protein from *Aquifex aeolicus* VF5 at 1.5 Å resolution. *Acta Crystallographica Section F: Structural Biology and Crystallization Communications* **65**, 1204-1208 (2009).
- 21 Carfi, A. *et al.* The 3-D structure of a zinc metallo-beta-lactamase from *Bacillus cereus* reveals a new type of protein fold. *The EMBO journal* **14**, 4914-4921 (1995).
- 22 Leesong, M., Henderson, B. S., Gillig, J. R., Schwab, J. M. & Smith, J. L. Structure of a dehydratase–isomerase from the bacterial pathway for biosynthesis of unsaturated fatty acids: two catalytic activities in one active site. *Structure* **4**, 253-264 (1996).