

Supporting Information

Diterpene Cyclases and the Nature of the Isoprene Fold

Rong Cao¹, Yonghui Zhang², Francis M. Mann³, Cancan Huang², Dushyant Mukkamala¹,
Michael P. Hudock¹, Matthew Mead³, Sladjana Prusic³, Ke Wang², Fu-Yiang Lin¹, Ting-Kai
Chang², Reuben Peters^{3,*}, and Eric Oldfield^{1,2,*}

¹Center for Biophysics & Computational Biology,

University of Illinois at Urbana-Champaign, Urbana, IL 61801, USA;

²Department of Chemistry, University of Illinois at Urbana-Champaign, 600 South Mathews Avenue,

Urbana, IL 61801, USA; ³Department of Biochemistry, Biophysics and Molecular Biology,

Iowa State University, Ames, IA 50011, USA

*To whom correspondence may be addressed. E-mail: rjpeters@iastate.edu or
eo@chad.scs.uiuc.edu

Table of Contents

Supporting Tables	S3-S9
The predicted and experimental numbers of helices and turns	S3
QSAR table	S4-S5
Distance Matrix of 20 terpene synthases by SSM program.....	S6-S7
Distance Matrix of 20 terpene synthases by CE program.....	S8-S9
Supporting Figures.....	S10-S57
Sequence alignment between bacterial diterpene cyclases and triterpene cyclases ...	S10-S11
Predicted and experimental number of helices and turns	S12
Full alignment of 228 terpene synthases obtained by using the PSI-BLAST.....	S13-S37
Full Structures of <i>Abies grandis</i> abietadiene synthase inhibitors investigated	S38

Structure Dendrogram	S39
PSI-BLAST output using Blr2149 as query	S40-S54
FASTA search program output alignment by using <i>Bradyrhizobium japonicum</i> γ -domain only as the query	S55
Supporting Experimental Section	S56-S62
Supporting references	S62-S63

Table S1: The number of helices (N_h) and turns (N_t) calculated using the COUDES¹ program versus the number determined experimentally for 10 terpene synthases. The numbers of helices and turns predicted for two bacterial diterpene cyclases (*B. japonicum* CPS, *M. tuberculosis* Rv3377c) and one plant diterpene cyclase (AgABS) are also shown.

		N_h (calc)	N_t (cacl)	N_h (expt)	N_t (expt)
3sqc	SHC	24	31	27	33
1w6k	OSC	26	35	28	34
1kzo	Farnesyl transferase	31	29	31	28
3dra	Geranylgeranyl transferase I	28	20	30	21
3dss	Geranylgeranyl transferase II	27	19	26	22
2a73	complement C3 β subunit	12	7	13	9
2f8z	FPPS	14	9	13	11
2q80	GGPPS	14	10	12	7
1ezf	SQS	13	8	14	11
2zcg	<i>S. aureus</i> CrtM	11	5	12	6
	<i>B. japonicum</i> CPS	23	19		
	<i>M. tuberculosis</i> Rv3377c	19	26		
	AgABS	34	32		

Table S2: CoMSIA training and test set predictions for AgABS Inhibition using GGPP as the substrate

Compound ^a	Experimental		Training	Test Set pIC ₅₀ (M) ^{b, c}						Predicted	Residual
	IC ₅₀ (μM)	pIC ₅₀ (M) ^b		1	2	3	4	5	6		
Aza-GGPP	0.01	7.9	7.8	7.9	8	7.9	7.9	7.9	7.9	7.9	0
BPH-916	0.15	6.8	6.7	6.8	6.8	6.7	6.7	6.8	6.8	6.7	0.1
BPH-882	0.19	6.7	6.7	6.7	6.7	6.8	6.7	5.6	6.7	5.6	1.1
BPH-890	0.24	6.6	6.6	6.7	6.7	6.7	6.7	6	6.6	6	0.6
BPH-922	0.25	6.6	6.6	6.6	6.5	6.7	6.6	6.6	6.5	6.5	0.1
BPH-901	0.29	6.5	6.5	6.5	6.2	6.5	6.6	6.7	6.5	6.2	0.3
BPH-899	0.35	6.5	6.5	6.5	6.6	6.5	6.4	6.5	6.4	6.4	0.1
BPH-921	0.48	6.3	6.2	6.1	6.1	6.1	6.1	5.9	6.2	5.9	0.4
BPH-892	0.53	6.3	6.3	6.4	6.2	6.3	6.3	6.3	6.3	6.4	-0.1
BPH-917	0.77	6.1	6.1	6.1	6.1	6	6.2	6.1	6.1	6	0.1
BPH-907	0.94	6	6	6	6	5.9	6	6	6	5.9	0.1
BPH-909	1.02	6	6.1	6.3	6	6.1	6	6	6.1	6.3	-0.3
BPH-908	1.08	6	6	6	5.9	6	6	5.9	5.9	5.9	0.1
BPH-924	1.18	5.9	6.2	6.1	6.1	6.2	6.1	5.9	6.2	5.9	0
BPH-896	1.45	5.8	5.7	5.8	5.6	5.7	5.8	5.5	5.4	5.4	0.4
BPH-863	1.72	5.7	5.7	5.8	5.6	5.7	5.7	5.8	5.5	5.5	0.2
BPH-923	1.94	5.7	5.8	6.2	5.7	5.7	5.7	5.7	5.7	6.2	-0.5
BPH-905	2.47	5.6	5.5	5.6	5.6	5.5	5.3	5.5	5.5	5.3	0.3
BPH-902	3	5.5	5.6	5.5	5	5.4	5.5	5.6	5.6	5	0.5
BPH-661	10.68	5	5	5	5	5	5.2	5.1	5	5.2	-0.2
BPH-904	12.58	4.9	4.9	4.9	4.9	4.9	4.9	4.9	4.9	4.9	0
BPH-913	15	4.8	4.8	4.9	4.6	4.8	4.8	4.7	4.8	4.6	0.2
BPH-918	30.34	4.5	4.5	4.5	4.4	4.5	5.3	4.5	4.5	5.3	-0.8
BPH-876	41.63	4.4	4.4	4.4	4.6	4.4	4.4	4.5	5	5	-0.6
BPH-891	41.72	4.4	4.3	4.4	4.3	5.6	4.3	4.4	4.3	5.6	-1.2
BPH-883	67.7	4.2	4.2	4.1	4.2	4.2	5.6	4.2	4.2	5.6	-1.4
BPH-673	175.68	3.8	3.7	3.7	5.2	3.8	3.7	3.8	3.7	5.2	-1.4
BPH-914	274	3.6	3.6	4.4	3.6	3.6	3.5	3.6	3.6	4.4	-0.8
		q ²	0.67	0.56	0.59	0.59	0.64	0.59	0.60		

r ²	0.99	0.99	0.99	0.99	0.99	0.99	0.99
N	6	6	5	6	5	6	6
F	734	474	503	503	602	603	916
n	28	23	23	23	23	24	24
%Steric	0.17	0.18	0.16	0.17	0.17	0.15	0.14
%Electrostatic	0.18	0.19	0.20	0.22	0.20	0.25	0.18
%Donor	0.33	0.31	0.33	0.31	0.37	0.37	0.34
%Acceptor	0.31	0.32	0.29	0.32	0.25	0.29	0.33

N = number of components; n = Number of training set compounds

^a the structures of all inhibitors shown in Figure S4.

^b $pIC_{50} = -\log_{10}IC_{50}[M]$

^c bold values indicated compounds not included in training set.

Table S3. Distance matrix for 20 terpene synthase structures obtained by using the Protein Structure Alignment by SSM program². The values shown are the normalized C^α rmsd values (1000* (C^α rmsd (Å)/N), where N is the number of aligned residues in structures *i* and *j*). The letters shown are the PDB file designations.^a

	1JFG	1DI1	5EAS	1N1B	2ONG	1PS1	2F7M	2HER	2IHI	1YHK	2FOR	1UBV	2EWG	1RQI	1RTR	2E8T	2J1O	2Q80	2ZCP	1EZF
1JFG ^b	-	17.4	13.1	14.7	14.9	15.6	17.4	19.0	19.9	17.7	21.1	17.4	18.2	19.9	20.2	20.0	21.4	21.1	18	19.9
1DI1	17.4	-	24.9	15.3	13.7	11.3	21.2	20.2	20.9	19.8	22.4	21.1	19.7	22.1	25.0	25.7	27.8	25.5	20.5	22.1
5EAS	13.1	24.9	-	4.3	4.0	12.0	15.6	16.2	16.6	15.1	19.7	15.0	15.5	17.8	19.9	14.5	19.9	14.5	17.2	16.9
1N1B	14.7	15.3	4.3	-	2.5	11.9	15.5	16.7	16.6	16.0	20.9	16.5	16.2	20.8	20.3	15.5	20.0	16.2	18.6	18.2
2ONG	14.9	13.7	4.0	2.5	-	11.2	14.8	16.4	14.9	14.9	19.9	15.6	18.7	17.8	20.2	15.1	20.7	15.6	17	17.3
1PS1	15.6	11.3	12.0	11.9	11.2	-	16.9	17.5	18.3	16.5	22.5	17.1	16.2	21.3	21.8	17.5	23.0	16.7	17.6	18.1
2F7M	17.4	21.2	15.6	15.5	14.8	16.9	-	7.0	5.1	5.9	10.8	2.7	5.8	10.8	10.5	9.5	12.1	10.3	19.1	18
2HER	19.0	20.2	16.2	16.7	16.4	17.5	7.0	-	5.1	5.7	11.8	7.5	5.4	11.1	11.9	10.3	13.9	10.0	20.1	17.2
2IHI	19.9	20.9	16.6	16.6	14.9	18.3	5.1	5.1	-	4.0	10.4	5.7	4.7	9.1	10.7	10.0	12.3	9.6	19.7	18.2
1YHK	17.7	19.8	15.1	16.0	14.9	16.5	5.9	5.7	4.0	-	10.4	5.8	2.7	9.6	10.9	10.2	12.4	10.1	19.6	17.4
2FOR	21.1	22.4	19.7	20.9	19.9	22.5	10.8	11.8	10.4	10.4	-	11.1	10.3	3.8	4.8	18.5	5.8	10.8	19.6	18.7
1UBV	17.4	21.1	15.0	16.5	15.6	17.1	2.7	7.5	5.7	5.8	11.1	-	6.0	11.0	10.5	9.1	12.3	10.5	19.9	18.8
2EWG	18.2	19.7	15.5	16.2	18.7	16.2	5.8	5.4	4.7	2.7	10.3	6.0	-	8.3	11.3	9.3	12.7	8.9	19	16.3
1RQI	19.9	22.1	17.8	20.8	17.8	21.3	10.8	11.1	9.1	9.6	3.8	11.0	8.3	-	6.2	9.7	7.4	8.8	18.6	17.5
1RTR	20.2	25.0	19.9	20.3	20.2	21.8	10.5	11.9	10.7	10.9	4.8	10.5	11.3	6.2	-	11.8	6.3	11.4	18.1	19
2E8T	20.0	25.7	14.5	15.5	15.1	17.5	9.5	10.3	10.0	10.2	18.5	9.1	9.3	9.7	11.8	-	19.2	4.8	20.3	18.6
2J1O	21.4	27.8	19.9	20.0	20.7	23.0	12.1	13.9	12.3	12.4	5.8	12.3	12.7	7.4	6.3	19.2	-	16.4	21.3	23
2Q80	21.1	25.5	14.5	16.2	15.6	16.7	10.3	10.0	9.6	10.1	10.8	10.5	8.9	8.8	11.4	4.8	16.4	-	17.8	15.7
2ZCP	18	20.5	17.2	18.6	17	17.6	19.1	20.1	19.7	19.6	19.6	19.9	19	18.6	18.1	20.3	21.3	17.8	-	10.8
1EZF	19.9	22.1	16.9	18.2	17.3	18.1	18	17.2	18.2	17.4	18.7	18.8	16.3	17.5	19	18.6	23	15.7	10.8	-

^aThese normalized C^α rmsd values were used to construct the phylogenetic tree shown in Fig 6a.

^b1JFG: *Fusarium sporotrichioides* trichodiene synthase
1DI1: *Penicillium roqueforti* aristolochene synthase
5EAS: *Nicotiana tabacum* 5-epi-aristolochene synthase
1N1B: *Salvia officinalis* (+)-bornyl diphosphate Synthase
2ONG: *Mentha spicata* 4S-limonene synthase
1PS1: *Streptomyces sp.* pentalenene synthase
2F7M: *Homo sapiens* farnesyl diphosphate synthase (FPPS)
2HER: *Cryptosporidium parvum* FPPS
2IHI: *Plasmodium vivax* FPPS
1YHK: *Trypanosoma cruzi* FPPS
2FOR: *Shigella flexneri* FPPS
1UBV: *Gallus gallus* FPPS
2EWG: *Trypanosoma brucei* FPPS
1RQI: *Escherichia coli* FPPS
1RTR: *Staphylococcus aureus* FPPS
2E8T: *Saccharomyces cerevisiae* geranylgeranyl
diphosphate synthase (GGPPS)
2J1O: *Sinapis alba* GGPPS
2Q80: *Homo sapiens* GGPPS
2ZCP: *S. aureus* dehydrosqualene synthase (crtM)
1EZF: *Homo sapiens* squalene synthase

Table S4: Distance matrix for 20 terpene synthase structures obtained by using the Protein Structure Alignment by Incremental Combinatorial Extension (CE) program³. The values shown are the normalized C^α rmsd values (1000* (C^α rmsd (Å)/N), where N is the number of aligned residues in structures *i* and *j*). The letters shown are the PDB file designations.^a

	1JFG	1DI1	5EAS	1N1B	2ONG	1PS1	2F7M	2HER	2IHI	1YHK	2FOR	1UBV	2EWG	1RQI	1RTR	2E8T	2J1O	2Q80	1EZF	2ZCP
1JFG ^b	-	20.0	13.8	14.8	15.4	13.1	16.9	16.0	33.1	15.9	17.4	16.8	16.5	20.8	16.6	18.7	18.4	23.4	20.4	19.7
1DI1	20.0	-	13.8	15.7	14.3	10.3	18.3	16.6	32.8	20.6	35.0	22.6	22.7	18.8	23.2	20.7	23.5	22.5	20.4	17.9
5EAS	13.8	13.8	-	4.2	4.0	13.4	14.5	18.6	14.2	13.8	16.8	13.9	14.8	16.5	15.7	16.7	18.1	17.3	16.4	15.1
1N1B	14.8	15.7	4.2	-	2.6	15.4	14.5	20.0	17.0	16.7	20.0	16.6	16.7	19.3	20.4	16.0	19.4	17.0	17.3	17.4
2ONG	15.4	14.3	4.0	2.6	-	14.2	16.3	19.5	18.0	16.0	19.5	15.9	16.0	18.6	17.8	16.8	23.1	18.0	17.1	16.9
1PS1	13.1	10.3	13.4	15.4	14.2	-	18.8	17.3	19.0	16.8	22.4	18.2	16.9	20.8	20.1	21.5	22.7	18.2	17.4	17.1
2F7M	16.9	18.3	14.5	14.5	16.3	18.8	-	6.3	4.5	5.0	9.9	5.1	5.3	9.0	9.7	9.3	11.5	9.0	16.9	17.8
2HER	16.0	16.6	18.6	20.0	19.5	17.3	6.3	-	5.9	6.0	9.9	8.9	5.9	9.2	9.8	9.7	11.1	8.8	17.6	19.2
2IHI	33.1	32.8	14.2	17.0	18.0	19.0	4.5	5.9	-	4.9	8.9	5.8	4.7	8.5	8.7	8.8	10.5	9.3	17	17.6
1YHK	15.9	20.6	13.8	16.7	16.0	16.8	5.0	6.0	4.9	-	9.2	6.4	2.5	7.6	9.8	9.1	10.9	8.9	17.1	18.3
2FOR	17.4	35.0	16.8	20.0	19.5	22.4	9.9	9.9	8.9	9.2	-	9.8	9.2	3.4	5.6	10.4	9.3	9.7	21.1	19.6
1UBV	16.8	22.6	13.9	16.6	15.9	18.2	5.1	8.9	5.8	6.4	9.8	-	6.4	9.8	8.7	9.5	10.6	9.8	17.1	17.8
2EWG	16.5	22.7	14.8	16.7	16.0	16.9	5.3	5.9	4.7	2.5	9.2	6.4	-	7.9	9.2	9.1	10.9	8.5	19.3	19.2
1RQI	20.8	18.8	16.5	19.3	18.6	20.8	9.0	9.2	8.5	7.6	3.4	9.8	7.9	-	6.7	10.7	12.9	8.5	21.2	20.6
1RTR	16.6	23.2	15.7	20.4	17.8	20.1	9.7	9.8	8.7	9.8	5.6	8.7	9.2	6.7	-	10.9	9.6	10.4	20.3	20
2E8T	18.7	20.7	16.7	16.0	16.8	21.5	9.3	9.7	8.8	9.1	10.4	9.5	9.1	10.7	10.9	-	12.7	4.5	19.6	21.8
2J1O	18.4	23.5	18.1	19.4	23.1	22.7	11.5	11.1	10.5	10.9	9.3	10.6	10.9	12.9	9.6	12.7	-	16.5	23.8	21
2Q80	23.4	22.5	17.3	17.0	18.0	18.2	9.0	8.8	9.3	8.9	9.7	9.8	8.5	8.5	10.4	4.5	16.5	-	16.5	17.8
1EZF	20.4	20.4	16.4	17.3	17.1	17.4	16.9	17.6	17	17.1	21.1	17.1	19.3	21.2	20.3	19.6	23.8	16.5	-	11.4
2ZCP	19.7	17.9	15.1	17.4	16.9	17.1	17.8	19.2	17.6	18.3	19.6	17.8	19.2	20.6	20	21.8	21	17.8	11.4	-

^aThese normalized C^α rmsd values were used to construct the phylogenetic tree shown in Fig S5.

^b1JFG: *Fusarium sporotrichioides* trichodiene synthase
1DI1: *Penicillium roqueforti* aristolochene synthase
5EAS: *Nicotiana tabacum* 5-epi-aristolochene synthase
1N1B: *Salvia officinalis* (+)-bornyl diphosphate Synthase
2ONG: *Mentha spicata* 4S-limonene synthase
1PS1: *Streptomyces sp.* pentalenene synthase
2F7M: *Homo sapiens* farnesyl diphosphate synthase (FPPS)
2HER: *Cryptosporidium parvum* FPPS
2IHI: *Plasmodium vivax* FPPS
1YHK: *Trypanosoma cruzi* FPPS
2FOR: *Shigella flexneri* FPPS
1UBV: *Gallus gallus* FPPS
2EWG: *Trypanosoma brucei* FPPS
1RQI: *Escherichia coli* FPPS
1RTR: *Staphylococcus aureus* FPPS
2E8T: *Saccharomyces cerevisiae* geranylgeranyl
diphosphate synthase (GGPPS)
2J1O: *Sinapis alba* GGPPS
2Q80: *Homo sapiens* GGPPS
2ZCP: *S. aureus* dehydrosqualene synthase (crtM)
1EZF: *Homo sapiens* squalene synthase

B.japonicum CPS 232 -----GIEGVFPNVWPINVFEPWSLYTLHLGLFAHPALAEAVRVI^{VAQLDA}
X.oryzae cyclase 233 -----GIEGIVPNVWPINVFEPAWSLYTLHLGLFAHPALAEAVRATVAPLAA
B.multivorans cyclase 242 -----AVPGVFPYVWPLDVFECAVYL^{YSLGGLGGLAAHPALTEPIARTATFLGA}
S.cellulosum cyclase 239 -----GIPGVVPTVWPIGRFEQAVVLYALRAF^{DLDDHPRLGDVARPQLDLDLGA}
M.tuberculosis cyclase 231 -----GAP-----AFYQAEIFEIVWSLWNLSRTD^{IDLSDPEIVR^{TY}LPYLDHVE}
S.clavuligerus cyclase 231 -----RYGGLFPETARITVFERLWLVLTLLHRA^{GLLAT--FEPLARRWVSALAA}
K.griseola cyclase 226 -----RYGGAI^{PMGSSMPYFEVLWVLN----}-LVLKYFPDVP^{IPREIEEIAA}
A.acidocaldarius SHC 286 GWEGLELYGVLDYGGW^{MFQASISPVWDTGLAVLALRAAGLPA}DDHRLVKAGEW^{LDDRQI}
Human OSC 361 HVSRI^{PDYLWMLDGMKMQGTNGSQIWDTAFAIQALLEAGGHRP}EFSSCKQKAHEFLRL
Consensus 361 -----g-gvvp--w-i-vfe-aw-ly-l--agl-a-p-----r--v--laa

B.japonicum CPS 280 RL^G-----VHGLGP-----ALHFAAD^{ADDDTAVALCVLHLA}GRDPAVDALRH
X.oryzae cyclase 281 RMS-----ARGLGP-----ALHFAAD^{ADDDTAVVLSVHLA}GRAPTADALRQ
B.multivorans cyclase 290 RIT-----ERGASF-----AAGFNP^{DGDTTSVAIAALS}MAGAATSCHSLAH
S.cellulosum cyclase 287 ALR-----PEGIGM-----SDAFTQ^{DGDI^{TS}TVIATLGELAA}PRAI^{EA}LRR
M.tuberculosis cyclase 275 QHW-----VRRGVG--WTGNSTLE^{CDTTSVAYDVL^{SK}FGRSPDIAVGL}
S.clavuligerus cyclase 277 PGG-----VPVPG-----FEP^{ADDDTAVTLHLA}TEL^{GVPYRPEVLD}
K.griseola cyclase 269 GFS-----DSGIGG-----GGLPFP^{DGDDTAYANLAGDKL}GAPTHPEILMK
A.acidocaldarius SHC 346 TVPGD---WAVKRPNLKPGGF^{AF--QFDNVYYPD}VDDTAVVVWALN^{TLRLRPDERRRDA}
Human OSC 421 SQVPDNP^{PDYQKYRQMRKGGFS}FTLDCGWI^{VS}^{DCTAEAL}KAVLL^{LQEK}CPHVTEH^{IPR}
Consensus 421 -l-----glg-----a--f--Daddtav-1-vl--g-p---d-l--

B.japonicum CPS 321 FEIGELFVTFPGERN-----ASVSTNIHALHA
X.oryzae cyclase 322 FERGALFVTFPGERN-----ASVSTNIHALHA
B.multivorans cyclase 331 FASDGVVATCPGERN-----PSLSTTVHAMHA
S.cellulosum cyclase 328 FERGGMFITYANELQ-----PSLTTNAHAMHA
M.tuberculosis cyclase 319 FEDADWFRTYFHEVG-----PSISTNVHVLGA
S.clavuligerus cyclase 315 FRTGDHFACYLGEDT-----GSVSTNAHVLLA
K.griseola cyclase 310 FWAEDHFVSY^{PEQ}T-----PSVTNAHALEY
A.acidocaldarius SHC 400 MTKGFRWIVGMOSS^{NGGWGA}YDVNTSDLPNHIP^{FCDFGEV---}TDPPSE^{DVT}AHVLE^C
Human OSC 481 ERLCDAVAVLLNMRN^{PDGGFATYETKRGHLELLNPS}EVFGDIMIDY^{TYVECTSAVMQA}
Consensus 481 f--ge-fvtf-ge-n-----psvstn-hal-a

B.japonicum CPS 348 LRLLGK--PAAGASA-----YVEANRNPHGLWDNEK^{WHVS}WLYPT^{TAHAVA}LAQ^G
X.oryzae cyclase 349 LRLLGE--PAVATRT-----YVETNRNPEGVWGNEK^{WHVS}WLYPT^{TAHAIA}LAQ^G
B.multivorans cyclase 358 LRLQRE--PRTVASA-----HVLAKRDADGLWRGDK^{WHVS}PF^{FLT}CHAVAA^LDHD
S.cellulosum cyclase 355 LASRGE--KASEPAR-----YLLERQQFDGRWVGDK^{WHSS}WLYTT^SQVILALSHE
M.tuberculosis cyclase 346 LKQAGY--DKCHPRV---KVLEFIRSSKEP^GRFCWRDK^{WHRS}AYYT^{TAHLI}CAASN-
S.clavuligerus cyclase 342 LGTWTR--HHPDTADHGNTIRLLGRWL^{VER}QHG^{DGHWD}-DKWHAS^{PPY}YATA^{AKVT}ALSRH
K.griseola cyclase 337 LNHLRM--RRGITEFG-AVEDACA^{EWVISQ}TEDGCWY-DKWNV^{SPY}YSTA^{ACV}EALLDA
A.acidocaldarius SHC 456 FGSFGY--DDAWK^{VIR---}-RAVEY^{LKRE}QK^{PDGS}WFG-RWGVN^{YLY}YGT^{GAVV}SALKAV
Human OSC 541 LK^{YFH}KRFP^{EHRAE}IRE^{TLT}QGLE^{FCRRQ}Q^{RADGS}WEGS-^{WGVCF}TYG^{TW}FGLE^AFACM
Consensus 541 lr--g-----yv---q-pdg-w--dkWhvsw-y-Ta--v-Al--

B.japonicum CPS 396 K-----PQWRDERALAA^{LLQAQR}DDGGW^{AGR-GST}FEETAYALFALH^{VMDG-SEEAT}
X.oryzae cyclase 397 H-----PRWRDECALAA^{LLQAQQ}ADGGW^{AGS-APT}FEETAYALFALH^{AVDS-GETPE}
B.multivorans cyclase 406 E-----VDRHGAQ^{TLEGLLQAQR}ADGGW^{AGTG-TPN}FEETAYAVLTI^{DRLCARGPMP}
S.cellulosum cyclase 403 G-----QLPAVKRGL^{EALLRQR}GDGGW^{AGG-APT}AETAYAVLALRA^{ARRNRELES}
M.tuberculosis cyclase 398 Y-----DDALCSDA^{IGWILNTQR}PDGS^{WGF}FDGQATA^{EETAYC}IQA^{LAH}WQ^{RESS}TSL
S.clavuligerus cyclase 399 GGPE---AADALRRA^{ARWVRETR}TDSG^{WGI}W-GTAE^{EYAA}QIL^LLD^{PAPEPT---}
K.griseola cyclase 393 RKQDEP--QLDSLRR^{AREWLLRHQT}DSG^{WMAE--}PSPE^{EYAA}QIL^LDLFAS^{RGGEG-}
A.acidocaldarius SHC 508 G--IDT--REPYIQ^{KALDWVEQHNP}PDGGW^{EDC--RS}YED^{PAYAG}KGAS^TSQTAWALM
Human OSC 600 GQTYRDGTCAEVS^{RACDFLLSRQ}MADGGW^{EDF--ES}CE^{RRY}VQS^{AQS}QI^{HNT}CWAMM
Consensus 601 -----v-ralewll--qr-dGgWg-----t-eetaYal-al-----

B.japonicum CPS 447 GRRR---IAQVVARALEWMLAR^{AAHGLPQTP}LWIGKELY^{CPTRVVR}VAELAGL^{LWL}ALR
X.oryzae cyclase 448 ARRL---IAQAVAQARAWMLAR^{YVPHALPQTT}LWIGKELY^{CPLLVVR}VT^{ELAG}LWLALQ
B.multivorans cyclase 458 AAR-----TSL^{LKA}FRLMRDCY^{SPRAQAS}MLWINKELSR^{PTRVVR}AVELAGL^{WVAHR}
S.cellulosum cyclase 455 MAR-----GAWRR^{GRD}WLSNHD^{IGAGQ}EDLRWIGKELY^{CPARVDR}AWVL^{GALL}ASQP
M.tuberculosis cyclase 451 S-----AQISRA^{GGWLSQ}CEPPYAP---LWI^{AK}TLYCSAT^{VV}KAA^{ILS}ALRL^{VD}
S.clavuligerus cyclase 449 ------TDVLC^{GAH}LTARADDD^{GPPAL}WHD^{KT}LFA^{PD}IVRAEVL^{STL}RL^{LD}
K.griseola cyclase 448 ------AE^{ECAA}ISRAKE^FTDESREN^{PL}WMG^{KD}LY^{TF}FRI^VDVTVM^{CGRA}AV^{VR}
A.acidocaldarius SHC 562 ALIAGRAE^{EAAR}RGV^{YLVE}TQR^{PDGGW}DEPY^{YT}GTG^{FP}DFL^{GY}TM^{YR}HV^{PT}LAL
Human OSC 658 GLMAVRHPDIEAQER^{GVRC}LLER^{QL}PNGD^{WP}QENIAG-VFNKSA^{IS}YS^{YR}NI^{FP}I^WAL
Consensus 661 g-----v-ra--wml-----ga---lwigkely-p--vvr--l-g-l-1--

B.japonicum CPS 503 WGRRLVLA--EGAGAAP-----
X.oryzae cyclase 504 WEQR^{AAD}GV^{THT}GAAP-----
B.multivorans cyclase 511 RAEAL^EAA^{CC}-----
S.cellulosum cyclase 508 SVE^S-----
M.tuberculosis cyclase 499 S^{NQ}-----
S.clavuligerus cyclase 499 RLPAPAPVPPGFDAART^{GPAD}
K.griseola cyclase 499 Y-----
A.acidocaldarius SHC 622 GR^{YK}QAIERR-----
Human OSC 717 GR^FSQLYPERALAGHP-----
Consensus 721 -----

Figure S2: Graph showing correlation between the number of predicted helices (red) (COUDES program)¹ and turns (blue) versus number seen experimentally in SHC, OSC, FTase I, GGTase I, GGTase II, FPPS, GGPPS, SQS and CrtM (≥ 5 helical residues define a helix; ≥ 4 turn residues define a turn).

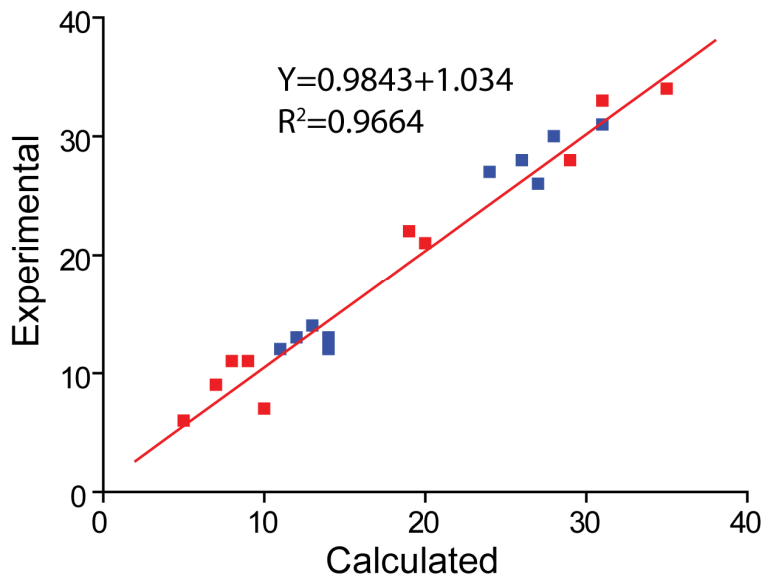


Figure S3: Full alignment of 228 terpene synthases obtained by using the JPRED3⁴ program. The target sequence (at the top of the output) is abietadiene synthase from *Abies grandis*, minus the plastid targeting segment⁵.





Figure S3 (continued).

QUERY
 UniRef90_Q947C4
 UniRef90_Q40SN4
 UniRef90_Q40SN5
 UniRef90_Q39Y43
 UniRef90_Q2PHF3
 UniRef90_Q00G37
 UniRef90_Q39548
 UniRef90_Q69D55
 UniRef90_Q00G36
 UniRef90_Q60HB5
 UniRef90_Q9XEH9
 UniRef90_Q32TN7
 UniRef90_Q1SDQ3
 UniRef90_Q673F9
 UniRef90_Q9FXV8
 UniRef90_Q8FXV9
 UniRef90_Q2R4G8
 UniRef90_Q38802
 UniRef90_Q95T35
 UniRef90_Q85AK2
 UniRef90_Q41771
 UniRef90_Q04408
 UniRef90_Q673G0
 UniRef90_Q68BJ4
 UniRef90_Q9LRB1
 UniRef90_Q672R4
 UniRef90_Q69D58
 UniRef90_Q64K29
 UniRef90_Q9MAX2
 UniRef90_Q6C3H3
 UniRef90_Q5GIA9
 UniRef90_Q43714
 UniRef90_Q6C3H2
 UniRef90_Q32W37
 UniRef90_Q67SL1
 UniRef90_Q6PWU1
 UniRef90_Q675K9
 UniRef90_Q40SN3
 UniRef90_Q673K6
 UniRef90_Q6E7D7
 UniRef90_Q5UBY0
 UniRef90_Q84KL3
 UniRef90_Q67SL2
 UniRef90_A0RZ12
 UniRef90_P93665
 UniRef90_Q22340
 UniRef90_Q84SM8
 UniRef90_Q4VM12
 UniRef90_Q94EN2
 UniRef90_Q84LK9
 UniRef90_Q71MJ3
 UniRef90_Q5UB05
 UniRef90_Q9FYU6
 UniRef90_Q7XLD9
 UniRef90_A2XXF1
 UniRef90_A2XXF1
 UniRef90_Q81634
 UniRef90_A1UH12
 UniRef90_Q85A63
 UniRef90_Q84UJ3
 UniRef90_Q927N4
 UniRef90_A0T2T4
 UniRef90_Q9FQ26
 UniRef90_Q66PX8
 UniRef90_Q9A7D0
 UniRef90_Q40577
 UniRef90_Q9LLR9
 UniRef90_Q49SP5
 UniRef90_Q8LSJ7
 UniRef90_Q1XBU4
 UniRef90_Q9ZT06
 UniRef90_Q84KL4
 UniRef90_Q5UB07
 UniRef90_Q40322
 UniRef90_Q672F7
 UniRef90_Q70E26
 UniRef90_Q9AXF5
 UniRef90_Q948Z0
 UniRef90_A2Q3S6
 UniRef90_Q5D1Y6
 UniRef90_Q5CD82
 UniRef90_Q5CD81
 UniRef90_Q6F5H3
 UniRef90_Q8FV75
 UniRef90_A0SNS9
 UniRef90_Q6F5H1
 UniRef90_Q95DN9
 UniRef90_Q1XBU5
 UniRef90_Q8L6K3
 UniRef90_Q9LDF1

F E T A R I M Q L F V M Q G D G L L S H D M E K E H V K N C F P P V I A
 F N T A R V M O L E Y M Y R D G F G L I S K K E M K D H V S R T L L E D P P V I A
 F E E T R V I T M V I F K D G D R L S N S K L E M K D H V K E C L L E D P P L
 M E E T R V I T M V I F K E G D G F G I S K T K I K D V I K D C L L E D P P L
 F N L R L Q V Q I F Y K F G D G F G I A N E E T K D V I R K V I L D P I
 W K M C K I L H L F Y Y Q T D G F S S P K E M V I S A V N A V I L D P P L
 W K M C K I L H L F Y Y Q T D G F S S P K E M V I S A V D A V I L D P P L
 W K L M S I V N L F Y M E D D G Y T S N R G L M N T V K A M F L E O P M
 W F M S K V L H F Y M E K D G Y F Y A D G V M A S A N A V I L D P P L
 W E M C K I L V F Y F Y S E D D A Y R T P K E T M S S A R A V I L D P P L
 W N L Y K T S H V F Y S Q A D G F S S P K E M M G A M N G V I L D P P L
 W N Q H V L N F Y A N D D G F T G N T L L D T V K D I L Y N R V
 L T V A K S V Y Y A A Y F D A W T I N Y H I A K V L F E R V
 W K M M K A S Y L F Y D E N D K F T L N E M Y S T V N A V I L D P P L
 W K M C K I L H L F Y F R T D G F S S P K E M A S A V N A V I L D P P L
 W N M C N V L N L F Y A T D D G F T G N A L L D V V K E L Y E P V S
 W H T G K V A S V F Y Q E G D Y A R K A M R S M A N A V I L E P L
 L D V A K A F Y Y F A L C G D H L Q T H L S K V L F O K V
 L K M S K V L N L F Y R K D D G F T S N D L M S L V K S V I Y E P V S
 W D V L R S L Y A T H S P Q H M I D R H V S R V I F E P V
 F T L A K S F Y A A F C D S R T L N F H L A K V L F E K V
 W K M C K V Q Y F F Y S R G D G F S S P
 L N L A R V I D V V Y K D E D G Y T N A G T V L K D L V S A L L L D P V
 F T V V K S I Y Y T T V C D L G T T N Y H M S K V L F E R V
 L N L A R V I D V T Y K N E D G Y T H F G A V L K D F V T S M L D P V
 W T L T N A A Y S L H T H G F W I S L P R G L K D P Y Q R Q N L O
 L N L A R V M D V I Y R E G D G Y T H V G K A A K G G I T S L L D P I
 L N F T R T M D V I Y K E O D S Y T H P V G K V M R D N I A S V I N A V
 L N L A R V L V I Y Q D E D G Y T H P G T K V E N E V S V L I D S M
 F D I T R A L H L Y K Y R D G Y S V I A T K E T K S V L S M V I E P V
 M N L A R M A Q M Y Q H G D G H G I E Y G E T E D R V L S L L V E P I
 F D I T R A L H L Y K Y R D G Y S V I A N K E T K N L V E K T L L E P M
 Y D V L R A F H V G Y K Y R D G Y S V I A N I E T K N E V A S V I E P V
 F D I L R A F H L Y K Y R D G Y S V I A K I E T K N L V M R T L E P V
 L N L S R V T D V I Y K D G D Y T N S Q G S T K N D I T S L L L N P
 F D I Q R V F H V G Y K Y R D G F S V I A S L I E T K N L V T R T V E T V
 F D I T R A F H L Y K Y R D G Y S V I A R E I R N L V M T T V I H V
 L N L S R F M D V I Y K N K D N Y T D S D G L M K T Q I K E V L V D P P I
 L N L A R V M D V I Y R E G D G Y T Y V G K A A K G G I T S L L E P I A
 F D I T R A F H L Y K Y R D G Y T W I S A N K E T K N L V M K T V I E P V
 F D I N R G L H H Y N Y R D G Y T V I A S N E T K N L V I K T V I E P V
 L N L A R M C E V I Y S V N D G F T H A E G D M K S Y M K S F F H P M
 L N F S R M D V I Y K E D D G Y T N P K S K V K D L T A S L V E S V
 L N L S R V M D V I Y K E D G Y T Y V I R K A M K D V I T S L L E P V
 L N L A R A L D F I Y K E D D G Y T H S Y L K K D O I A S V L G D P I
 L N F S R F M D V I Y T K D N Y T H V I E G L M K T Y I K D V L D P I
 F N I A R V T E F L N K G D W Y T H S H A L K D Q L A A V L R D P V
 W S M C K V C H F F Y S G D G F S S P
 L N L P R M A E T V Y K K C D S Y T H S S T M K D V I M S L L V E P I S
 L N L A R M A Q M Y Q H G D G H G H G N H E T K D R I S S L L I O P I
 L N L A Q V M D V I Y K K D G F T N V I G E E L K D H I K S L L V H P I
 T S C Q K V F E M F F N K K N R Y D S N T E M L H I D I K K A L Y D P I
 C L K I F N M F F N S N L Y D S H D D T H L F D I O K A F I L P
 F N M L Q V I H C L Y Q O K D A Y T F S N N S T K E M I T S V L V D P V
 L N F T R V I D T L Y Q E E D E Y T N A K G K L K N M I H S I L I E S V
 L N L T R V M A L L Y S H E D G Y T N S K G R T K L L Q S L L D P L
 F D I L K G S L H G Y K Y R D G F S V I A N K E T K N W V R T V E S V
 L N L A R L V E V I Y H N D G Y T H P E K V L K P H I L L V D S I
 L N F S R S I D V I Y K N K D H F T H V G V E V I N I K S L V D A I
 R N F S R M D V I Y K N E D G Y S K P E K I L K P L T A A F V D Q
 M S M G R M G Q Y M Y H Y G D G H G I O H S L I H Q O M S T L F H P
 K N I A T T A H C M Y L K G D S H G I O N T D V K N S I S N L H P
 L N L A R I D V I Y K H N D G Y T H P E K V L K P H I L L V D S I
 F D I T R A F H Y G Y K Y R D G Y S V I S S V E T K S L V M R T L L E P V
 T N L A R I S H G T Y Q Y G D G H G A P D T L A K N R L K A L I E P I
 V D L G R M A Q L M Y H N G D G H G T O H P T L H Q O M T R T L E P F A
 T N L A R I A O C Y Q T G D G H G A P D S I A K N R V K S L I E P I S
 L N L A R V I D M L Y K N D N L K N V G O E I Q D Y I K S L E I N A I
 L S L A R V S D I L Y K T N N G Y N H A G E V I L M K V I L V H P L
 F D I S R V W H G Y R Y R D G Y S F A N V E T K S L V M R T V I E P V
 E D L G R M A Q F M Y H Y G D G H G T O H P K I H Q O I A A C L F O P F A
 F T V A K S F Y A A F C D S R T I N F H I A K V L F E D K V I
 L N L E R L H V A Y K D S D G Y T Q P E K V L K P H I L L E P I A
 M N L A R M S Q C M Y Q F G D G H G H G V Q E L T K A R V L S L L V D P I
 M N L A R M A Q C Y Q N G D R H T V I Q D N S K N R V L P L L I H P I
 A D M G R M A Q F M Y H N G D G H G I O N P O I H Q O M T D I L E P
 A N L G R M A Q C Y Q H R D G F G M O H S L V N K M L R G L L L E D
 M N L V R M S O C M Y L Y G D G H G V I Q E N T K D R V L S L E P I
 L N V V R V S H F M Y L H G D H G A Q N O E T M O D V F T L E P
 K N I T K A S H F M Y L H S D E K V L K P H I V A L L V D S I
 L N F N R I M D V I Y K N K D G Y T Q P E K V L K P H I L L E P D O I
 M N L A R M S O C M Y Q Y G D H G C P E K A K T I D R V Q S L L V D P I

Figure S3 (continued).

	1	10	20	30	40	50	60	70	80
--	---	----	----	----	----	----	----	----	----

UniRef90_Q94G53
UniRef90_Q84LU6
UniRef90_Q84KL2
UniRef90_Q84404
UniRef90_Q8ZUH4
UniRef90_A2Q848
UniRef90_Q8W119
UniRef90_Q2XS05
UniRef90_Q9FV12
UniRef90_Q23516
UniRef90_Q09HU5
UniRef90_A2Q852
UniRef90_Q2IHU6
UniRef90_Q675L0
UniRef90_Q8LH31
UniRef90_Q81192
UniRef90_Q86376
UniRef90_Q86CL9
UniRef90_Q84U04
UniRef90_Q8EJ97
UniRef90_Q84405
UniRef90_Q2TJK5
UniRef90_P58287
UniRef90_Q5CL36
UniRef90_A0RZ13
UniRef90_Q0PC13
UniRef90_Q8GU54
UniRef90_Q9LQ27
UniRef90_Q9C6W6
UniRef90_A2YDV8
UniRef90_Q485P6
UniRef90_UP1000004870C
UniRef90_A2X444
UniRef90_Q01L21
UniRef90_Q5SBP0
UniRef90_A2YR19
UniRef90_Q5SBP6
UniRef90_A0ERE2
UniRef90_Q485P4
UniRef90_Q8H2B4
UniRef90_Q5SBP4
UniRef90_Q84NE5
UniRef90_Q8SPN1
UniRef90_Q8YN71
UniRef90_Q5SBP1
UniRef90_Q66NG3
UniRef90_Q485P3
UniRef90_A0F037
UniRef90_A2XGN2
UniRef90_Q2XSC4
UniRef90_Q84LU9
UniRef90_Q8AR04
UniRef90_A0ERE1
UniRef90_Q84LV1
UniRef90_Q2M4M3
UniRef90_Q01KN7
UniRef90_Q0VHD8
UniRef90_A2ZSN9
UniRef90_UP10000162AD9
UniRef90_Q5G4S4
UniRef90_Q0JE23
UniRef90_Q29VN3
UniRef90_Q84LS8
UniRef90_Q0J834
UniRef90_Q86PX9
UniRef90_Q81193
UniRef90_Q0JE19
UniRef90_Q2NM14
UniRef90_Q01LN7
UniRef90_Q75WN1
UniRef90_Q8LA1
UniRef90_Q8LVP7
UniRef90_Q01HA9
UniRef90_Q1PD02
UniRef90_Q6JD69
UniRef90_Q5GJ60
UniRef90_Q60WJ1
UniRef90_Q1T1D0
UniRef90_Q7570
UniRef90_Q22980
UniRef90_Q5GJ59
UniRef90_Q5SBP3
UniRef90_Q84LV0
UniRef90_Q65437
UniRef90_Q65435
UniRef90_Q5GJ58
UniRef90_Q10L24
UniRef90_Q10L17
UniRef90_Q5SBP7
UniRef90_Q8ZVM4
UniRef90_Q4U3F6
UniRef90_Q7XNL0

SSSSEQLQFLVLDKVKRESLSSSSNTONTLPYDTAWLALTPHPHHGRPFMEKQLQWTLHNOTPQGFWAAAGDDVITLDQLLS
ASTSELPLPYMILESLKSKLRAAASLGETTVISAYDTAWLALVNRQLDGRSPQFPEALDQWLAARNQLPDGSGWDAGMFLVQDRLLIN

Figure S3 (continued).

UniRef90_Q94G53
UniRef90_Q84UJ6
UniRef90_Q84L2
UniRef90_Q64404
UniRef90_Q9ZUJ4
UniRef90_AZC549
UniRef90_Q8W1J9
UniRef90_Q2XSC5
UniRef90_Q9FYL2
UniRef90_Q23516
UniRef90_Q08HU5
UniRef90_A20K52
UniRef90_Q20HU6
UniRef90_Q67SL0
UniRef90_Q8LH31
UniRef90_Q81192
UniRef90_Q96376
UniRef90_Q88CL9
UniRef90_Q84UJ4
UniRef90_Q8E197
UniRef90_Q64405
UniRef90_Q2TJK5
UniRef90_Q59287
UniRef90_Q50L36
UniRef90_A0RZ13
UniRef90_Q0PC13
UniRef90_Q8GUE4
UniRef90_Q8LQ27
UniRef90_Q8C3W6
UniRef90_A2YDV8
UniRef90_Q49SP6
UniRef90_UPI00004870C
UniRef90_A2X414
UniRef90_Q0L121
UniRef90_Q5S8P0
UniRef90_A2YRL9
UniRef90_Q5S8P6
UniRef90_A0ERE2
UniRef90_Q49SP4
UniRef90_Q8H2B4
UniRef90_Q5S8P4
UniRef90_Q84NE5
UniRef90_Q6SPN1
UniRef90_Q6YN71
UniRef90_Q5S8P1
UniRef90_Q8R0K3
UniRef90_Q49SP3
UniRef90_A0F0E7
UniRef90_A2XGN2
UniRef90_Q2XSC4
UniRef90_Q84UJ9
UniRef90_Q8AR04
UniRef90_A0ERE1
UniRef90_Q84UJ1
UniRef90_Q2M4M3
UniRef90_Q01KN7
UniRef90_Q0V1D6
UniRef90_A2ZSN9
UniRef90_UPI000162AD9
UniRef90_Q5GAS4
UniRef90_Q0JE23
UniRef90_Q29VN3
UniRef90_Q84J58
UniRef90_Q0J834
UniRef90_Q66FX9
UniRef90_Q81193
UniRef90_Q0JE19
UniRef90_Q2NM14
UniRef90_Q01LN7
UniRef90_Q75WN1
UniRef90_Q8LIA1
UniRef90_Q8LVP7
UniRef90_Q01HA6
UniRef90_Q1PDD2
UniRef90_Q6JD69
UniRef90_Q5GJ60
UniRef90_Q5QWJ1
UniRef90_Q11D0
UniRef90_Q75I70
UniRef90_Q226R0
UniRef90_Q5GJ58
UniRef90_Q5S8P3
UniRef90_Q84UJ0
UniRef90_Q65437
UniRef90_Q85435
UniRef90_Q5GJ58
UniRef90_Q10L24
UniRef90_Q10L17
UniRef90_Q5S8P7
UniRef90_Q6ZVM4
UniRef90_Q4U3F6
UniRef90_Q7XNL0

90 100 110 120 130 140 150 160

T L A G L V A L K R W Q L A P D M I H K G L E F V N R N T E R V M K Q K P S D V P R W F T I M F P A M L E L A G A S S L R V D F S E R I L V E L S Q N R D D L L
T L G C V Y A L J A T W G V R E E Q R A R G L J A Y L Q D N L W R L G E D D E E W M M V G F E I L T F P V L L E K A R N L G L D I N Y D D P A L Q G L Y A K R Q L K L

Figure S3 (continued).

UniRef90_Q94G53
UniRef90_Q84U06
UniRef90_Q84K12
UniRef90_Q64404
UniRef90_Q9ZUH4
UniRef90_A22649
UniRef90_Q8W1J9
UniRef90_Q2XSC5
UniRef90_Q8FYL2
UniRef90_Q23516
UniRef90_Q08HU5
UniRef90_A22652
UniRef90_Q20HU6
UniRef90_Q67510
UniRef90_Q9LH31
UniRef90_Q81192
UniRef90_Q96376
UniRef90_Q86C19
UniRef90_Q84U04
UniRef90_Q6EJ87
UniRef90_Q64405
UniRef90_Q2TJK5
UniRef90_P59287
UniRef90_Q59136
UniRef90_A0RZ13
UniRef90_Q0PC13
UniRef90_Q86UE4
UniRef90_Q9LQ27
UniRef90_Q9C6W6
UniRef90_A22VY9
UniRef90_Q49SP6
UniRef90_UP1000004870C
UniRef90_A2X444
UniRef90_Q01L21
UniRef90_Q5SBP0
UniRef90_A22YR19
UniRef90_Q5SBP6
UniRef90_A0ERE2
UniRef90_Q49SP4
UniRef90_Q8H2B4
UniRef90_Q5SBP4
UniRef90_Q8AN65
UniRef90_Q8SPN1
UniRef90_Q6YN71
UniRef90_Q5SBP1
UniRef90_Q66NG3
UniRef90_Q49SP3
UniRef90_A0R057
UniRef90_A2XGN2
UniRef90_Q2XSC4
UniRef90_Q84UJ9
UniRef90_Q9AR04
UniRef90_A0ERE1
UniRef90_Q84UJ1
UniRef90_Q2M4M3
UniRef90_Q01KN7
UniRef90_Q0VH06
UniRef90_A2ZSN9
UniRef90_UP10000162AD9
UniRef90_Q5GAS4
UniRef90_Q0JE23
UniRef90_Q29VN3
UniRef90_Q84IS9
UniRef90_Q0J834
UniRef90_Q66PX9
UniRef90_Q81193
UniRef90_Q0JE19
UniRef90_Q2NM14
UniRef90_Q01LN7
UniRef90_Q75WN1
UniRef90_Q9LIA1
UniRef90_Q9LV57
UniRef90_Q01HA9
UniRef90_Q1PD02
UniRef90_Q6LD99
UniRef90_Q5GJ60
UniRef90_Q6QWJ1
UniRef90_Q1T1D0
UniRef90_Q75I70
UniRef90_Q22680
UniRef90_Q5GJ59
UniRef90_Q5SBP3
UniRef90_Q84UV0
UniRef90_Q65437
UniRef90_Q65435
UniRef90_Q5GJ58
UniRef90_Q10L24
UniRef90_Q10L17
UniRef90_Q5SBP7
UniRef90_Q6ZVM4
UniRef90_Q4U3F6
UniRef90_Q7XNL0

170 180 190 200 210 220 230 240

S T Y F S P S L W G D H F L S V S D R G E F D E E R T M K P L V K D M C M S S Q
H S L N S P Y G A P A Y Y E L L Q K L T O E L K H L L L T E M E
F S P S L W G D V F L Q Y D S Q P M E L N D N Q M Q K E E V R K L F Q S S S
L H V D E T R R S G N Y R P S A W D S N Y L Q S K E K K L T R L E G L L E Q V K E L G T K
D V K P V Q E G R R S G N Y Q P S L W D F N Y V Q S L N Y K E E R Y L T R H A E L L V Q V K P
S L F S P S L W G D Y F L S V F D D A R E L E S V M K P Y V R D R L T S S H
T R E E V D E K K Q Y S P L L L F L E A L P A Q S Y D N D V L K N L S N D G S L L Q S P S A T A R A Y M I T G N T R O L S Y L H S L T N S G N G G V P S F Y P
A K L P R E A L H A R P T T L L L H S L E G M E N L D W E R L L Q P K P A G S L H S S P A A S A Y A L S E T G D K E L L E Y L E T A T T N F D G G A P D T Y P
V R G M L N R V D
A Y H L L G A N E W E E K E H G D L K E E V R K M L L M T S
K P K V R N M C M S S H
F L Y H S S T T A A S Y Q Q Q E W L V Q L E R Q K E E V G N M I T S S V T
D L A W M L R L E K L K E E V A T M I T S S L K
P T T R R S G N Y E P S A W D F N Y L Q N Y H H K E E R Y L R R D L L E K V K M C L K E E K
L L Q E V S K M L N E T E
F P P S L W A N R F L S F S N S E L E A Y A N A H E K E S V R S L T D T T
Y O P S A W D F N Y L Q S L N N H S K E E R H L E R K A K L T M L L E Q E
F H P A L F G D F L N N S D E W E E R V D Q L V E E V G R M L E V C K
M M M D E A
P
H D P
D P V L K T S K E D D Q S T K F H P S L W G D Y F L Q N S S L S A E E S T O R M I K R V E E L K V Q V K S
D D G S Q A L E A L A L A H S W T Q A C M K G R A E V L R E E V R T L L K G L K
F E P S V W G D F F L N Y D P K P L Q R S E D M M E K A D K L L F E T H
H A A L K Q E V R R M I T T T A
C P R S W A D L L M K R K A H V S E E Q V R R M L E L E E K M E L V S I T T
L L K E E V R K T L K A A
E R A E K L K D D V R T M F E T
E L L K E E V R K T L K A A
F S P S Y W G D H F L S V S L D E A L E K I E T V F K P K V R D M C M S P H
R K L T V D P T
E V L R E R V R K V L K G S T
V L L K E E V R N M I L K G S H
L L S K Y D G D E Y R K L S E K L L E E V R K L Y L S A E T
F H P T V W G D Y F L Q V S D E T M A E R L R H L R E E V S G M F Q A C
Q F H N K T S L E Y S H E L N L K K K N C L S A N V
O A M K E R A E V R E D V R K L L M G P
A T D R H R S E E W R E R V D R L K M O V G K L L K T O N
T P
S S L K E T V R G L L M A S K
K E N V R R M L N P A

Figure S3 (continued).

UniRef90_Q04653 V N G N G F K F 330 G D I T S M L I L I Y E A S Y H G V E P S E S E D I L M D A S E A E A F F T T O N Y L L K E W I E N T E 370 D O N I A L C F I H A L V I F P H C 400
UniRef90_Q84U06 E D G R F F K E S L A K D V R G M L Q L L Y E A S Y H G V E P S E S E D I L M D A S E A E A F F T T O N Y L L K E W I E N T E 370 D O N I A L C F I H A L V I F P H C 400
UniRef90_Q84KL2 H N G G F F T C G F T G E D K Q V R S M L S L Y E A S Y H G V E P S E S E D I L M D A S E A E A F F T T O N Y L L K E W I E N T E 370 D O N I A L C F I H A L V I F P H C 400
UniRef90_Q04604 F N G G F F T C G F T G E D K Q V R S M L S L Y E A S Y H G V E P S E S E D I L M D A S E A E A F F T T O N Y L L K E W I E N T E 370 D O N I A L C F I H A L V I F P H C 400
UniRef90_Q9ZU44 N O G K F L N E N T S D D T K D V G G M W S L L Y E A S Y H G V E P S E S E D I L M D A S E A E A F F T T O N Y L L K E W I E N T E 370 D O N I A L C F I H A L V I F P H C 400
UniRef90_A20649 E G S F K A R L L D D D T K D V G G M W S L L Y E A S Y H G V E P S E S E D I L M D A S E A E A F F T T O N Y L L K E W I E N T E 370 D O N I A L C F I H A L V I F P H C 400
UniRef90_Q8W1J9 N K G F L N E N T S D D T K D V G G M W S L L Y E A S Y H G V E P S E S E D I L M D A S E A E A F F T T O N Y L L K E W I E N T E 370 D O N I A L C F I H A L V I F P H C 400
UniRef90_Q2XSC5 E N G Y F F K H H D D T K D V G G M W S L L Y E A S Y H G V E P S E S E D I L M D A S E A E A F F T T O N Y L L K E W I E N T E 370 D O N I A L C F I H A L V I F P H C 400
UniRef90_Q9FVL2 A N G K F K E S L C O D D T K D V G G M W S L L Y E A S Y H G V E P S E S E D I L M D A S E A E A F F T T O N Y L L K E W I E N T E 370 D O N I A L C F I H A L V I F P H C 400
UniRef90_Q23516 E T G F F R K T L H S D L K G L L S L Y E A S Y H G V E P S E S E D I L M D A S E A E A F F T T O N Y L L K E W I E N T E 370 D O N I A L C F I H A L V I F P H C 400
UniRef90_Q09HU5 D S G S F K A S L S D N T K G L L S L Y E A S Y H G V E P S E S E D I L M D A S E A E A F F T T O N Y L L K E W I E N T E 370 D O N I A L C F I H A L V I F P H C 400
UniRef90_A20652 P K G G F F N E K I S I G D K D V G G M W S L L Y E A S Y H G V E P S E S E D I L M D A S E A E A F F T T O N Y L L K E W I E N T E 370 D O N I A L C F I H A L V I F P H C 400
UniRef90_Q20HU6 E N G K F F O S L T G D K R V R S M L S L Y E A S Y H G V E P S E S E D I L M D A S E A E A F F T T O N Y L L K E W I E N T E 370 D O N I A L C F I H A L V I F P H C 400
UniRef90_Q675L0 E N G O F F O S S T G D G M L T L F R A S E I S F H G G E K V M O E A K A F F T T O N Y L L K E W I E N T E 370 D O N I A L C F I H A L V I F P H C 400
UniRef90_Q9LH31 E D G R L K E S L T V G D V R G M L Q L L Y E A S Y H G V E P S E S E D I L M D A S E A E A F F T T O N Y L L K E W I E N T E 370 D O N I A L C F I H A L V I F P H C 400
UniRef90_Q81192 E K G D F K A S L A Q D T K G M L Q L L Y E A S Y H G V E P S E S E D I L M D A S E A E A F F T T O N Y L L K E W I E N T E 370 D O N I A L C F I H A L V I F P H C 400
UniRef90_Q84U44 D D E L R D H I E T N Y E E F A A V L L N V Y R A T D L M V S G E V I L D E A R S F A T R K L L O R K L D E P H E S S K T C R E D D N I L L S O S P L H W K
UniRef90_Q68CL9 K D G G Y F C A G O S S I T A M Y N S Y R A S O L V G T H G G E D I L D E A L A F A T R S R L L E I E S S A S H P H L R R R K L L D E I L D
UniRef90_Q84U44 E K G G Y F C A G O S S I T A M Y N S Y R A S O L V G T H G G E D I L D E A L A F A T R S R L L E I E S S A S H P H L R R R K L L D E I L D
UniRef90_Q8EJ97 K E G G F S G E L K G D V G G M W S L L Y E A S Y H G V E P S E S E D I L M D A S E A E A F F T T O N Y L L K E W I E N T E 370 D O N I A L C F I H A L V I F P H C 400
UniRef90_Q64405 D N G O F F G S T A Y N K H V R G L L S L Y E A S Y H G V E P S E S E D I L M D A S E A E A F F T T O N Y L L K E W I E N T E 370 D O N I A L C F I H A L V I F P H C 400
UniRef90_Q21LK9 G O G Y N K A L Y S N H G L L S L Y E A S Y H G V E P S E S E D I L M D A S E A E A F F T T O N Y L L K E W I E N T E 370 D O N I A L C F I H A L V I F P H C 400
UniRef90_P58287 S D G K F K E S L R G D A K G M L S L Y E A S Y H G V E P S E S E D I L M D A S E A E A F F T T O N Y L L K E W I E N T E 370 D O N I A L C F I H A L V I F P H C 400
UniRef90_Q50L38 S O N G F L E N L K E T K A I K M L S L Y E A S Y H G V E P S E S E D I L M D A S E A E A F F T T O N Y L L K E W I E N T E 370 D O N I A L C F I H A L V I F P H C 400
UniRef90_A0R233 K S G N F N E L T S D K D V G G M W S L L Y E A S Y H G V E P S E S E D I L M D A S E A E A F F T T O N Y L L K E W I E N T E 370 D O N I A L C F I H A L V I F P H C 400
UniRef90_Q0PC13 E S G S F V K L L G D V G G M W S L L Y E A S Y H G V E P S E S E D I L M D A S E A E A F F T T O N Y L L K E W I E N T E 370 D O N I A L C F I H A L V I F P H C 400
UniRef90_Q86UE4 E K G F K P S L S M D I K G M L S L Y E A S Y H G V E P S E S E D I L M D A S E A E A F F T T O N Y L L K E W I E N T E 370 D O N I A L C F I H A L V I F P H C 400
UniRef90_Q81027 N D G R F K E S L V R D F R G M L Q L L Y E A S Y H G V E P S E S E D I L M D A S E A E A F F T T O N Y L L K E W I E N T E 370 D O N I A L C F I H A L V I F P H C 400
UniRef90_Q9C6W6 N D G R F K E S L V R D F R G M L Q L L Y E A S Y H G V E P S E S E D I L M D A S E A E A F F T T O N Y L L K E W I E N T E 370 D O N I A L C F I H A L V I F P H C 400
UniRef90_A2YDV8 E K G G F F K P S S A K N F M D L L Y E A S Y H G V E P S E S E D I L M D A S E A E A F F T T O N Y L L K E W I E N T E 370 D O N I A L C F I H A L V I F P H C 400
UniRef90_Q49SP3 E D G R F K E S L A E D E A A H L G T P O G E V I D E A L A F A T R S R L L E I E S S A S H P H L R R R K L L D E I L D
UniRef90_UPI00004870C E D G R F K E S L A E D E A A H L G T P O G E V I D E A L A F A T R S R L L E I E S S A S H P H L R R R K L L D E I L D
UniRef90_A2X444 E G G F A F A N S P R D I T L Y E A S Y H G V E P S E S E D I L M D A S E A E A F F T T O N Y L L K E W I E N T E 370 D O N I A L C F I H A L V I F P H C 400
UniRef90_Q01L21 E K G G F R P D L V D D T Q G L L S L Y E A S Y H G V E P S E S E D I L M D A S E A E A F F T T O N Y L L K E W I E N T E 370 D O N I A L C F I H A L V I F P H C 400
UniRef90_Q5SBP0 E K G G F R P D L V D D T Q G L L S L Y E A S Y H G V E P S E S E D I L M D A S E A E A F F T T O N Y L L K E W I E N T E 370 D O N I A L C F I H A L V I F P H C 400
UniRef90_A2YRL9 N O G G F D D E S V M N N V G M L S L Y E A S Y H G V E P S E S E D I L M D A S E A E A F F T T O N Y L L K E W I E N T E 370 D O N I A L C F I H A L V I F P H C 400
UniRef90_Q5SBP9 E K G G F R P D L V D D T Q G L L S L Y E A S Y H G V E P S E S E D I L M D A S E A E A F F T T O N Y L L K E W I E N T E 370 D O N I A L C F I H A L V I F P H C 400
UniRef90_A0ERE2 E K G G F R P D L V D D T Q G L L S L Y E A S Y H G V E P S E S E D I L M D A S E A E A F F T T O N Y L L K E W I E N T E 370 D O N I A L C F I H A L V I F P H C 400
UniRef90_Q49SP4 H D G F F E E S L S G D K D V G G M W S L L Y E A S Y H G V E P S E S E D I L M D A S E A E A F F T T O N Y L L K E W I E N T E 370 D O N I A L C F I H A L V I F P H C 400
UniRef90_Q84NE5 K A G D F E E S L S G D K D V G G M W S L L Y E A S Y H G V E P S E S E D I L M D A S E A E A F F T T O N Y L L K E W I E N T E 370 D O N I A L C F I H A L V I F P H C 400
UniRef90_Q8SPN1 K D G G F F K A L Y S N H G L L S L Y E A S Y H G V E P S E S E D I L M D A S E A E A F F T T O N Y L L K E W I E N T E 370 D O N I A L C F I H A L V I F P H C 400
UniRef90_Q6YV71 T T T G T F D E V T K D V G G M W S L L Y E A S Y H G V E P S E S E D I L M D A S E A E A F F T T O N Y L L K E W I E N T E 370 D O N I A L C F I H A L V I F P H C 400
UniRef90_Q5SBP1 E G G S D F S A N L A D D K G M L S L Y E A S Y H G V E P S E S E D I L M D A S E A E A F F T T O N Y L L K E W I E N T E 370 D O N I A L C F I H A L V I F P H C 400
UniRef90_Q86NG3 S D G F K E S L T S D V K G M L S L Y E A S Y H G V E P S E S E D I L M D A S E A E A F F T T O N Y L L K E W I E N T E 370 D O N I A L C F I H A L V I F P H C 400
UniRef90_Q49SP3 G K G S F K E S L V R D F R G M L Q L L Y E A S Y H G V E P S E S E D I L M D A S E A E A F F T T O N Y L L K E W I E N T E 370 D O N I A L C F I H A L V I F P H C 400
UniRef90_Q2XSC4 D D G S F F A W V V I D V S V L S L Y E A S Y H G V E P S E S E D I L M D A S E A E A F F T T O N Y L L K E W I E N T E 370 D O N I A L C F I H A L V I F P H C 400
UniRef90_Q84U09 D D G S F F A W V V I D V S V L S L Y E A S Y H G V E P S E S E D I L M D A S E A E A F F T T O N Y L L K E W I E N T E 370 D O N I A L C F I H A L V I F P H C 400
UniRef90_Q9AR04 D D G S F F A W V V I D V S V L S L Y E A S Y H G V E P S E S E D I L M D A S E A E A F F T T O N Y L L K E W I E N T E 370 D O N I A L C F I H A L V I F P H C 400
UniRef90_A0ERE1 D E G E F F P L N D N T K G L L S L Y E A S Y H G V E P S E S E D I L M D A S E A E A F F T T O N Y L L K E W I E N T E 370 D O N I A L C F I H A L V I F P H C 400
UniRef90_Q84U01 K K E G F N K N D V O E L L E L L Y E A S Y H G V E P S E S E D I L M D A S E A E A F F T T O N Y L L K E W I E N T E 370 D O N I A L C F I H A L V I F P H C 400
UniRef90_Q2M4M3 E P R Q L I K N T V R G M L S L Y E A S Y H G V E P S E S E D I L M D A S E A E A F F T T O N Y L L K E W I E N T E 370 D O N I A L C F I H A L V I F P H C 400
UniRef90_Q01KN7 D G G F F M N G I A K E P R Q L I K N T V R G M L S L Y E A S Y H G V E P S E S E D I L M D A S E A E A F F T T O N Y L L K E W I E N T E 370 D O N I A L C F I H A L V I F P H C 400
UniRef90_Q0VHD6 D N E G F F S I S L N G D K K G L L S L Y E A S Y H G V E P S E S E D I L M D A S E A E A F F T T O N Y L L K E W I E N T E 370 D O N I A L C F I H A L V I F P H C 400
UniRef90_AZ2S9 E K G G F F S I S L N G D K K G L L S L Y E A S Y H G V E P S E S E D I L M D A S E A E A F F T T O N Y L L K E W I E N T E 370 D O N I A L C F I H A L V I F P H C 400
UniRef90_UPI0000162AD9 S E N G F F K E S L T R D A R G M L S L Y E A S Y H G V E P S E S E D I L M D A S E A E A F F T T O N Y L L K E W I E N T E 370 D O N I A L C F I H A L V I F P H C 400
UniRef90_Q5G4S4 E O G G F G F T A N S A R A D V E S L Y E A S Y H G V E P S E S E D I L M D A S E A E A F F T T O N Y L L K E W I E N T E 370 D O N I A L C F I H A L V I F P H C 400
UniRef90_Q01E23 S E N G F F K E S L T R D A R G M L S L Y E A S Y H G V E P S E S E D I L M D A S E A E A F F T T O N Y L L K E W I E N T E 370 D O N I A L C F I H A L V I F P H C 400
UniRef90_Q29VN3 E D G G F F S S D I A D D P K G L L S L Y E A S Y H G V E P S E S E D I L M D A S E A E A F F T T O N Y L L K E W I E N T E 370 D O N I A L C F I H A L V I F P H C 400
UniRef90_Q94J38 D E G K F K E S L I N D P R G M L S L Y E A S Y H G V E P S E S E D I L M D A S E A E A F F T T O N Y L L K E W I E N T E 370 D O N I A L C F I H A L V I F P H C 400
UniRef90_Q0J834 E N G F N T R D D N N G T R S E L L S L Y E A S Y H G V E P S E S E D I L M D A S E A E A F F T T O N Y L L K E W I E N T E 370 D O N I A L C F I H A L V I F P H C 400
UniRef90_Q66PX9 K K S O S F N A N L A Q D T K G M L S L Y E A S Y H G V E P S E S E D I L M D A S E A E A F F T T O N Y L L K E W I E N T E 370 D O N I A L C F I H A L V I F P H C 400
UniRef90_Q81193 D D G R F V A N L A Q D T K G M L S L Y E A S Y H G V E P S E S E D I L M D A S E A E A F F T T O N Y L L K E W I E N T E 370 D O N I A L C F I H A L V I F P H C 400
UniRef90_Q0JE19 E G G F F V A N L A Q D T K G M L S L Y E A S Y H G V E P S E S E D I L M D A S E A E A F F T T O N Y L L K E W I E N T E 370 D O N I A L C F I H A L V I F P H C 400
UniRef90_Q2NM14 E G G F F V A N L A Q D T K G M L S L Y E A S Y H G V E P S E S E D I L M D A S E A E A F F T T O N Y L L K E W I E N T E 370 D O N I A L C F I H A L V I F P H C 400
UniRef90_Q01LN7 D H G T F N E T S L A D D P R G M L S L Y E A S Y H G V E P S E S E D I L M D A S E A E A F F T T O N Y L L K E W I E N T E 370 D O N I A L C F I H A L V I F P H C 400
UniRef90_Q75WN1 K D G G F L C C T Q L E E E E I K S I N Q L F R A S L H L F H G P E L V E A I S F A R H H L K O V F O K S H L P L A G T E T S S N L S T H I R N A L S I S Q H W
UniRef90_Q9L1A1 E D G K F K E S L T E D A V R G M L S L Y E A S Y H G V E P S E S E D I L M D A S E A E A F F T T O N Y L L K E W I E N T E 370 D O N I A L C F I H A L V I F P H C 400
UniRef90_Q81VP7 N G K F K E S L T E D A V R G M L S L Y E A S Y H G V E P S E S E D I L M D A S E A E A F F T T O N Y L L K E W I E N T E 370 D O N I A L C F I H A L V I F P H C 400
UniRef90_Q01HA9 D D G K F F K E S L T E D A V R G M L S L Y E A S Y H G V E P S E S E D I L M D A S E A E A F F T T O N Y L L K E W I E N T E 370 D O N I A L C F I H A L V I F P H C 400
UniRef90_Q1PDD2 S S G F F K A S I L T S D V G V G L S L Y E A S Y H G V E P S E S E D I L M D A S E A E A F F T T O N Y L L K E W I E N T E 370 D O N I A L C F I H A L V I F P H C 400
UniRef90_Q6J069 E G G F A G A A A D T R S L S L Y E A S Y H G V E P S E S E D I L M D A S E A E A F F T T O N Y L L K E W I E N T E 370 D O N I A L C F I H A L V I F P H C 400
UniRef90_Q5GJ60 E L I G N F V E A D T R S L S L Y E A S Y H G V E P S E S E D I L M D A S E A E A F F T T O N Y L L K E W I E N T E 370 D O N I A L C F I H A L V I F P H C 400
UniRef90_Q60WJ1 E K G L E N R H F A H L K G M L S L Y E A S Y H G V E P S E S E D I L M D A S E A E A F F T T O N Y L L K E W I E N T E 370 D O N I A L C F I H A L V I F P H C 400
UniRef90_Q11TD0 V M D N K F S S N A K D F E L E A S H L G E D E I L D A K I F A K W L K D A F S Y S N I K I V V E R V V H A L E E L P S H W
UniRef90_Q7570 K G N L N G I C S D T N A L L A L Y E A A H L G T P N E E I L R E A Q V E T T T N Q L K R I V D O L E R A G A S S P H I L M R N R A L O M P P H Y
UniRef90_Q22680 E N E K F M E S I T S D V K G M L S L Y E A A H L G T P N E E I L R E A Q V E T T T N Q L K R I V D O L E R A G A S S P H I L M R N R A L O M P P H Y
UniRef90_Q5GJ59 E D G S F L I D G T N D P K G L S L Y E A A H L V T H D E T L E D A I A F A R O H L E A A R R O S L K S P L A E O V G R A L G T P H K
UniRef90_Q5SBP3 E R G G F D E S D T L G L S L Y E A S N L G V T G E E L L E A M E F F A E R L R R S S E R E S H O K G T P P H L A R H R N A L G L S H K
UniRef90_Q84UV0 K K G G K D V K N D V K G M L S L Y E A S N L G V T G E E L L E A M E F F A E R L R R S S E R E S H O K G T P P H L A R H R N A L G L S H K
UniRef90_Q65437 S N G S F K E S L I G D A K G M L S L Y E A A H L A T K O Y I L D E A L I F T S S H E L V A T I V V E R V V H A L E E L P S H W
UniRef90_Q65435 N D G N F K G G I S D A K G L L A L Y E A A O L R T T T E Y I M E A I S F F T S S N L E L A D G T P P H L A R H R N A L G L S H K
UniRef90_Q5GJ58 E G G S F A S T S T A D D T A L S L Y E A A Y L R K H G E A L D D A I A F A R R S Q S L E A G O G A T R S P M A K O V S R A L D T P H K
UniRef90_Q10L24 S T G F R E S T S T A D D T A L S L Y E A A Y L R K H G E A L D D A I A F A R R S Q S L E A G O G A T R S P M A K O V S R A L D T P H K
UniRef90_Q10L17 N R G S F R A L S S D T R G L S L Y E A A H M A M P G E A L D D A I A F A R R S Q S L E A G O G A T R S P M A K O V S R A L D T P H K
UniRef90_Q5SBP7 D D G G F F K E S L A E D A K D V G G M W S L L Y E A S Y H G V E P S E S E D I L M D A S E A E A F F T T O N Y L L K E W I E N T E 370 D O N I A L C F I H A L V I F P H C 400
UniRef90_Q32VM4 K D G K F K E S L A E D A K D V G G M W S L L Y E A S Y H G V E P S E S E D I L M D A S E A E A F F T T O N Y L L K E W I E N T E 370 D O N I A L C F I H A L V I F P H C 400
UniRef90_Q4U3F6 S S G F F E E D M T R D V R G M L S L Y E A S A O L R I V H G E I L D E A G A F A E S R L K K T I E K T L D G T L A O V K H R E R P N R
UniRef90_Q7XNLO D G D A R L S S V L A D R D V R G M L S L Y E A S A O L R I V H G E I L D E A G A F A E S R L K K T I E K T L D G T L A O V K H R E R P N R

Figure S3 (continued).

```

UniRef90_Q94G53      MVIPLRVIETSWFIIIVYPKKVG 420
UniRef90_Q84UJ6      NIIEVLVIAREYISYVEEQQEY
UniRef90_Q84KL2      NIPLRLVIAREYISYVEEQQEY
UniRef90_Q64U44      SVIPDWEARSFIIIVYGNHNSWLN
UniRef90_Q82UJ4      OMRRRLISTRWYIDVYGRKQNYK
UniRef90_A20E49      GVPRLRIETIYIISYVEEPPS
UniRef90_Q8W1J9      SIIPRLIEARWFLIDRYATR
UniRef90_Q2XSC5      RAORLIEARWFLIDAYARRP
UniRef90_Q8FVL2      SIIPRLIEARWFLIDRYATR
UniRef90_Q23516      SIIPRLIEARWFLIDRYATR
UniRef90_Q09HU5      RIORLIEARWFLIDAYATR
UniRef90_A20E52      RVIPRLIEARWFLIDRYATR
UniRef90_Q20HJ6      SVIPDWEARSFIIIVYGNHNSWLN
UniRef90_Q67SLO      SLPLRWEARSFIIIVYGNHNSWLN
UniRef90_Q8LH31      NMIEPLVIAREYISYVEEQQEY
UniRef90_Q81192      RIQSVIEARWFLIDAYARRP
UniRef90_Q96376      RMDRVIEENRIWLEIASSALWFGKSSYR
UniRef90_Q68CL8      AASRDIEVRYISYVEEQQEY
UniRef90_Q84UJ4      GITSRIETROYISYVEEQQEY
UniRef90_Q8EJ97      RLHLRIEARWFLIDRYATR
UniRef90_Q64405      SVIPDWEARSFIIIVYGNHNSWLN
UniRef90_Q27JK5      TIIVRVIAARKYISYVEEQQEY
UniRef90_P59287      GVPRLRIETIYIISYVEEPPS
UniRef90_Q50L36      RTQRRIEARSFIIIVYGNHNSWLN
UniRef90_A0R273      RMARLIEARWFLIDRYATR
UniRef90_Q0PC13      RMPRLIEARSFIIIVYGNHNSWLN
UniRef90_Q8GUE4      RLRIEALVIAREYISYVEEQQEY
UniRef90_Q8L027      TIIVRVIAARKYISYVEEQQEY
UniRef90_Q8C8W6      NIIEVLVIAREYISYVEEQQEY
UniRef90_A2YD8      RVIGLIEARWFLIDRYATR
UniRef90_Q49SP4      GMDQVIEARWFLIDRYATR
UniRef90_UPI00004870C  NLIEPLVIAREYISYVEEQQEY
UniRef90_A2X44      RVIRIIEARWFLIDRYATR
UniRef90_Q01121      TIIVRVIAARKYISYVEEQQEY
UniRef90_Q5SBP0      RQRLIEARWFLIDRYATR
UniRef90_A2YRL9      RVIRIIEARWFLIDRYATR
UniRef90_Q5SBP6      KMLRIEARWFLIDRYATR
UniRef90_A0ERE2      RVIORLIEARWFLIDRYATR
UniRef90_Q49SP4      GMDQVIEARWFLIDRYATR
UniRef90_Q892B4      RVPQRLIEARWFLIDRYATR
UniRef90_Q5SBP4      RVIGLIEARWFLIDRYATR
UniRef90_Q84NE5      KMLRIEARWFLIDRYATR
UniRef90_Q5SPN1      GMPVIEARWFLIDRYATR
UniRef90_Q6YN71      RIORLIEARWFLIDRYATR
UniRef90_Q5SBP1      TLVRLIEARWFLIDRYATR
UniRef90_Q49SP3      GLPRLIEARWFLIDRYATR
UniRef90_A0FB7      TMKRRIEARWFLIDRYATR
UniRef90_A2X2V2      NKRRIEARWFLIDRYATR
UniRef90_Q2XSC4      RVIGLIEARWFLIDRYATR
UniRef90_Q84UJ9      NLIEPLVIAREYISYVEEQQEY
UniRef90_Q5A9P4      SVIORPNSRWLIDAYARRP
UniRef90_Q84UJ1      RMRRRIEARWFLIDRYATR
UniRef90_Q2MM43      RPLRIEALVIAREYISYVEEQQEY
UniRef90_Q01KN7      TSKRTETIHYLSEYGOEIGHMS
UniRef90_Q0VHD6      RVIKLIEARWFLIDRYATR
UniRef90_A2Z5N9      RARLIEARWFLIDRYATR
UniRef90_UPI0000162AD9  NMEMLVPIVFIPIVYVEEATR
UniRef90_Q5GAS4      RVIRIIEARWFLIDRYATR
UniRef90_Q0J223      TYKRVIEMLHYMIEYGOEAAE
UniRef90_Q29VN3      TLKRLIEARWFLIDRYATR
UniRef90_Q84J58      AAARLIEARWFLIDRYATR
UniRef90_Q03834      RPLRIEALVIAREYISYVEEQQEY
UniRef90_Q66PX9      NPLRLKARHHIIGLSEKPTR
UniRef90_Q81193      RVIGLIEARWFLIDRYATR
UniRef90_Q0E19      TYRRIEALVIAREYISYVEEQQEY
UniRef90_Q2NM14      RLRIEALVIAREYISYVEEQQEY
UniRef90_Q01LN7      THRRRIEARWFLIDRYATR
UniRef90_Q75W11      NPLRIEALVIAREYISYVEEQQEY
UniRef90_Q8L1A1      SLIEPLVIAREYISYVEEQQEY
UniRef90_Q8LV77      NMIEPLVIAREYISYVEEQQEY
UniRef90_Q01H49      TYKRVIEMLHYMIEYGOEAAE
UniRef90_Q1PDD2      GMPRLGTRQYISYVEEQQEY
UniRef90_Q6J069      RVIGLIEARWFLIDRYATR
UniRef90_Q5GJ60      RVIGLIEARWFLIDRYATR
UniRef90_Q60WJ1      RVIOVDFIKWQIINAYEKDLCRV
UniRef90_Q1T1D0      RVIPWFDLKHVYKQYQANK
UniRef90_Q75170      NMKRRIEARWFLIDRYATR
UniRef90_Q2980      NAEMLVPIVFIPIVYVEEATR
UniRef90_Q5GJ59      TLKRLIEARWFLIDRYATR
UniRef90_Q5SSP3      RMARLIEARWFLIDRYATR
UniRef90_Q84UJ0      TVIRGLTSSKRFIIIVYGNHNSWLN
UniRef90_Q65437      NFEEQVISAHPIVYGNHNSWLN
UniRef90_Q65435      QMEVLVIAREYISYVEEQQEY
UniRef90_Q5GJ58      RVIGLIEARWFLIDRYATR
UniRef90_Q10L24      APKRLIETIHYIISYVEEPPS
UniRef90_Q10L17      APKRLIETIHYIISYVEEPPS
UniRef90_Q5SBP7      AAILRLGAEARFIIIVYGNHNSWLN
UniRef90_Q8ZVM4      GITSRIETROYISYVEEPPS
UniRef90_Q4U3FE      GORVIEARWFLIDRYATR
UniRef90_Q7XNL0      TCKRIEALVIAREYISYVEEPPS

```

Figure S3 (continued).

UniRef90_Q94G53
UniRef90_Q84UU6
UniRef90_Q84KL2
UniRef90_Q64404
UniRef90_Q9ZUH4
UniRef90_A2Q649
UniRef90_Q8W1J9
UniRef90_Q2XSC5
UniRef90_Q9FVL2
UniRef90_Q23516
UniRef90_Q09HU5
UniRef90_A2Q652
UniRef90_Q20HU6
UniRef90_Q675L0
UniRef90_Q9LH31
UniRef90_Q81192
UniRef90_Q96376
UniRef90_Q68CL9
UniRef90_Q84UU4
UniRef90_Q6EJ97
UniRef90_Q64405
UniRef90_Q2TJK5
UniRef90_P59287
UniRef90_Q50L36
UniRef90_A0RZ13
UniRef90_Q0PC13
UniRef90_Q8GUE4
UniRef90_Q9LQ27
UniRef90_Q9C6W6
UniRef90_A2YDV8
UniRef90_Q49SP6
UniRef90_LP100004870C
UniRef90_A2X414
UniRef90_Q01L21
UniRef90_Q5SBP0
UniRef90_A2YRL9
UniRef90_Q5SBP6
UniRef90_A0ERE2
UniRef90_Q49SP4
UniRef90_Q8H2B4
UniRef90_Q5SBP4
UniRef90_Q84NE5
UniRef90_Q9SPN1
UniRef90_Q6YN71
UniRef90_Q5SBP1
UniRef90_Q66NG3
UniRef90_Q49SP3
UniRef90_A0F0B7
UniRef90_A2XGN2
UniRef90_Q2XSC4
UniRef90_Q84UU9
UniRef90_Q9AR04
UniRef90_A0ERE1
UniRef90_Q84UV1
UniRef90_Q2M4M3
UniRef90_Q01KN7
UniRef90_Q0VHD6
UniRef90_A2ZSN9
UniRef90_LP10000162AD9
UniRef90_Q5GAS4
UniRef90_Q0JE23
UniRef90_Q29VN3
UniRef90_Q94JS8
UniRef90_Q0J834
UniRef90_Q66PX9
UniRef90_Q81193
UniRef90_Q0JE19
UniRef90_Q2NM14
UniRef90_Q01LN7
UniRef90_Q75WN1
UniRef90_Q9LIA1
UniRef90_Q9LVP7
UniRef90_Q01HA9
UniRef90_Q1PDD2
UniRef90_Q6JD69
UniRef90_Q5GJ60
UniRef90_Q6QWJ1
UniRef90_Q1T1D0
UniRef90_Q75170
UniRef90_Q22680
UniRef90_Q5GJ59
UniRef90_Q5SBP3
UniRef90_Q84UV0
UniRef90_Q65437
UniRef90_Q65435
UniRef90_Q5GJ58
UniRef90_Q10L24
UniRef90_Q10L17
UniRef90_Q5SBP7
UniRef90_Q9ZVM4
UniRef90_Q4U3F6
UniRef90_Q7XNL0

730 740 750

```

T N L P R M G H F M Y T D E G D G K Y G K E D P M F K P Y V F S L F V N P I
L N F A R L I F D V M Y K Y R D D G Y G S D A P K T E V K N L V K R T L V H E P V
F D I C R I I F H H G Y Q E K Y R D D G Y G S D A P K T E V K N L V K R T L V H E P V
T N L A R M S Q C M Y Q E K Y R D D G Y G S D A P K T E V K N L V K R T L V H E P V
M N V A R V I T E F A Y E N E D K Y T K P E E L L K Q Q W I V S M L F E P P Y A
UniRef90_Q8W1J9 R N L G R A A Q F M Y F D D G D G N H S Q L H Q R I A C L L F E P P Y A
UniRef90_Q2XSC5 A N L G R A A Q F M Y F D D G D G N Q S S L L Q Q W I V S M L F E P P Y A
UniRef90_Q9FVL2 L N L A R I T I D V T Y K H N D D G Y T H P P E K T K T V D Y I Q S V L F N P V
UniRef90_Q23516 A N L G R A A Q F M Y F D D G D G H G V Q D H S K I T H Q Q M G G L M F P V
UniRef90_Q09HU5 L N F A R G T H F I Y E D E D K Y T P D L L K D D Y I A A L L V D D Q P S
UniRef90_A2Q652 F N I V R G L Q F V Y K Y R D D G Y G S D A P K T E V K N L V K R T L V H E P V
UniRef90_Q20HU6 F N V A R G L R L I Y K Y R D D G Y G S D A P K T E V K N L V K R T L V H E P V
UniRef90_Q675L0 L N L A R I T I D V T Y K H N D D G Y T H P P E K T K T V D Y I Q S V L F N P V
UniRef90_Q9LH31 A N L G R V A Q F I Y K H G D G F G V Q H S K T Y E H L I A G L L F E P Y A
UniRef90_Q81192 F S C L K V F N M F F N
UniRef90_Q96376
UniRef90_Q68CL9
UniRef90_Q84UU4
UniRef90_Q6EJ97
UniRef90_Q64405
UniRef90_Q2TJK5
UniRef90_P59287
UniRef90_Q50L36
UniRef90_A0RZ13
UniRef90_Q0PC13
UniRef90_Q8GUE4
UniRef90_Q9LQ27
UniRef90_Q9C6W6
UniRef90_A2YDV8
UniRef90_Q49SP6
UniRef90_LP100004870C
UniRef90_A2X414
UniRef90_Q01L21
UniRef90_Q5SBP0
UniRef90_A2YRL9
UniRef90_Q5SBP6
UniRef90_A0ERE2
UniRef90_Q49SP4
UniRef90_Q8H2B4
UniRef90_Q5SBP4
UniRef90_Q84NE5
UniRef90_Q9SPN1
UniRef90_Q6YN71
UniRef90_Q5SBP1
UniRef90_Q66NG3
UniRef90_Q49SP3
UniRef90_A0F0B7
UniRef90_A2XGN2
UniRef90_Q2XSC4
UniRef90_Q84UU9
UniRef90_Q9AR04
UniRef90_A0ERE1
UniRef90_Q84UV1
UniRef90_Q2M4M3
UniRef90_Q01KN7
UniRef90_Q0VHD6
UniRef90_A2ZSN9
UniRef90_LP10000162AD9
UniRef90_Q5GAS4
UniRef90_Q0JE23
UniRef90_Q29VN3
UniRef90_Q94JS8
UniRef90_Q0J834
UniRef90_Q66PX9
UniRef90_Q81193
UniRef90_Q0JE19
UniRef90_Q2NM14
UniRef90_Q01LN7
UniRef90_Q75WN1
UniRef90_Q9LIA1
UniRef90_Q9LVP7
UniRef90_Q01HA9
UniRef90_Q1PDD2
UniRef90_Q6JD69
UniRef90_Q5GJ60
UniRef90_Q6QWJ1
UniRef90_Q1T1D0
UniRef90_Q75170
UniRef90_Q22680
UniRef90_Q5GJ59
UniRef90_Q5SBP3
UniRef90_Q84UV0
UniRef90_Q65437
UniRef90_Q65435
UniRef90_Q5GJ58
UniRef90_Q10L24
UniRef90_Q10L17
UniRef90_Q5SBP7
UniRef90_Q9ZVM4
UniRef90_Q4U3F6
UniRef90_Q7XNL0

```

Figure S3 (continued).

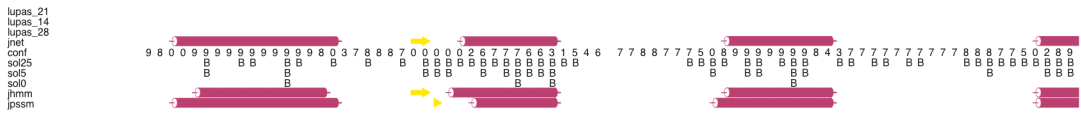
UniRef90_Q9LS76
 UniRef90_Q01KU6
 UniRef90_A3AHW2
 UniRef90_Q9C8E3
 UniRef90_Q10L25
 UniRef90_Q6ZCP0
 UniRef90_Q10L27
 UniRef90_A2Y215
 UniRef90_Q5UB06
 UniRef90_Q49SP7
 UniRef90_A2ZV57
 UniRef90_Q9C748
 UniRef90_A2XGY4
 UniRef90_Q84NC9
 UniRef90_Q5W283
 UniRef90_Q9F127
 UniRef90_Q0U750
 UniRef90_Q5SBP5
 UniRef90_A3A876
 UniRef90_A3AS12
 UniRef90_Q6ZH94
 UniRef90_Q65Y14
 UniRef90_Q29V2
 UniRef90_Q1SIE9
 UniRef90_Q7XNT3
 UniRef90_Q4KSH9
 UniRef90_Q9TJ9
 UniRef90_Q1SMG8
 UniRef90_A3AS12
 UniRef90_Q2184
 UniRef90_Q84ZV8
 UniRef90_A2X696
 UniRef90_Q9C749
 UniRef90_Q0IWT3
 UniRef90_A3AHW3
 UniRef90_Q9LR09
 UniRef90_A1D1K6
 UniRef90_Q7XLE2
 UniRef90_Q68V14
 UniRef90_Q13284
 UniRef90_Q5BCV6
 UniRef90_Q8F5E5
 UniRef90_UP1000021ADD4
 UniRef90_A2RA08
 UniRef90_Q4WWB3
 UniRef90_UP10000219AC9
 UniRef90_Q7S1N5
 UniRef90_Q2GRK4
 UniRef90_Q2JIF6
 UniRef90_Q5B878
 UniRef90_A2QWS7
 UniRef90_A0UKP0

1 10 20 30 40 50 60 70 80

WVLAQVLDK D I N G S E T P Q F P S S L K W I A N N Q I A D G S W G D S S I F L V Y D R V I N

Q A C Q L V I O Q I R T K D G E I G F M S P A V Y D T A W V S M V K K T T P E G Q W L L P E Q F E Y L R T Q I E D G G W E T Y A Q D D G I L L N

E E E A R A L V R K V G N A V D P I G F S T T S Q I I V D T A W A A M I S K E F G D K V W L F P S S F H Y L L K T O A A D G S W G Y P T T O T A G I L L D
 Q A S Q L V Q I R T K D G E I G F M S A A V I V D T A W V S M V Q K T T P E G R W L L P K Q F F E Y I L R T Q I E D G S W E T Y A S D V I D G I L L N
 Q A R S L L T R L V H N S D D V I G A F G L S S Q I I V D T A W L S L V K K V D G Q O R W L F P S S F F O V Y L L R N O Q D D G S W E A Y Q S E I D G I L L N
 R A Q H L V T G I A K G Q O N G F G G F S S I L V D T A W L S M V I T K K D G V Q H W L F P S S F F O V Y L R Q Q G D G W G P Q V T L L D I T L L Q
 A K G L T A R L S S I N D P K A F S T A S S I I V D T A W I A M V I T K P T G S V W A F P A A F O A I L O E Q S P G S W G S T S E L L D V S
 Q S Q I D S L I S F I A S H Q S K G L G T S S A I V D T A W L S M I Q K P A N S S H F L P Q G F T Y L L R Q L P S G G W T S Y A T P V D G I L L N
 T Y I S E I L S Q S P D K L K R Q G F T F S A S I I V D T A W L S M I N R R E N Q I L W L F P Q G F D V L L S Q L L D G A W S P S P A S T V D G I L L N
 A D A L L V S E L R R L V T G L P D R T A A S V Y E T A L L V L O V P A I G D R R R R V Q S F F E Y L L Q T I D G W G Y P E N I W I Q D G V I H



UniRef90_Q9LS76
 UniRef90_Q01KU6
 UniRef90_A3AHW2
 UniRef90_Q9C8E3
 UniRef90_Q10L25
 UniRef90_Q6ZCP0
 UniRef90_Q10L27
 UniRef90_A2Y215
 UniRef90_Q5UB06
 UniRef90_Q49SP7
 UniRef90_A2ZV57
 UniRef90_Q9C748
 UniRef90_A2XGY4
 UniRef90_Q84NC9
 UniRef90_Q5W283
 UniRef90_Q9F127
 UniRef90_Q0U750
 UniRef90_Q5SBP5
 UniRef90_A3A876
 UniRef90_A3AS12
 UniRef90_Q6ZH94
 UniRef90_Q65Y14
 UniRef90_Q29V2
 UniRef90_Q1SIE9
 UniRef90_Q7XNT3
 UniRef90_Q4KSH9
 UniRef90_Q9TJ9
 UniRef90_Q1SMG8
 UniRef90_A3AS12
 UniRef90_Q2184
 UniRef90_Q84ZV8
 UniRef90_A2X696
 UniRef90_Q9C749
 UniRef90_Q0IWT3
 UniRef90_A3AHW3
 UniRef90_Q9LR09
 UniRef90_A1D1K6
 UniRef90_Q7XLE2
 UniRef90_Q68V14
 UniRef90_Q13284
 UniRef90_Q5BCV6
 UniRef90_Q8F5E5
 UniRef90_UP1000021ADD4
 UniRef90_A2RA08
 UniRef90_Q4WWB3
 UniRef90_UP10000219AC9
 UniRef90_Q7S1N5
 UniRef90_Q2GRK4
 UniRef90_Q2JIF6
 UniRef90_Q5B878
 UniRef90_A2QWS7
 UniRef90_A0UKP0

90 100 110 120 130 140 150 160

T I A C V A I K S W N V D S H K C E I G M S F R E N E N S I G D N A E H M P I G F E V A F T S I E E G K K I G D P F Y D S P I A A N Y A R R T K I K I

T A A S L L A L E I H R T D P E R I G R A R A A S R E I Q A W D V K D T V H V G F E M I P A L L R L R K K G H E F E F D G R A E I E R N R R K I K I

T A S A L L A L C H V Q E P Q I E H G V A S L K R Q L A V W N D V D D T N H I G V E L T I P A L L S M I E K E I D V P F F E F P C R D V I E R M H E E K I K I
 T A A A L L A L R H V K N P L Q I O R G R A S L E E O L L A W D V D L D T N H I G V E M I P A L L R L L R E K G H E F E F E G R A E L D R I W R I K I
 T A A S L L A E T H V I E M K E F I G R A R A A L S R Q L A W S V K D T V H V G F E I I L P A L L R L L R E K G H E F E F D G R A E L D R I W R I K I
 T A A G L L V I K R Y R R D P I Q L D S H R L A T A S A L H S L E E W D V I T H V G F E I I V P T M L D L L V O E D P S I S F E K G R A E L T E I R E A K M
 T A A S L L A L Q R H S V N P I Q L D V G E L C R V E L A R A A W N V S E T T H V G F E I I V P T L L R L L D E E G I S I S O F E P G K A T I M E I N A A K M
 T M A G L L S L L A H R S P L S D N G A K L D S R I S R A E W D V A S T I Q V G S E V L V P L L Q L S L O D A K S E D P A G A G I R E M N A O K M
 T L A A L L S L O K H E D L H S R I L K A K O F L O T A L K G L D R L L A C T L S V G E V R L P E L L D L L E L E G H T F N F D R T Y I T E I Q S K K L
 T A A A L L A I K K H L K E I Q I R K G S S S R T O L Q N W D D V I S T N H I G V E I V A P A L L D L K O E D P A L D F F E P A G A G I R E M N A O K M
 T A A G L I S I D N W L H S Q A S A L O S M L L H W V A T T D Q V G F E I L V I T S L O D A K S E D P A G A G I R E M N A O K M
 T A A L L A I K K H L K E I Q I R K G S S S R T O L Q N W D D V I S T N H I G V E I V V S L L T L L E K E A G V S V D F P P G T I Q A L R D A K I
 T L A A L L A L O R H A E D S R A A R O D D N S R I K A K A D L D G T D Q V G F E I V P G L L R O I S H G I T F K F C Q C R L E A L R A A K I
 S A L C R H V R R S S H L P D D I L F R I F S A K S L R L D A K I Q K W P D D G I H F E L I V P V L L H L L R E E G V I T F Q F P A K R D D L S K Y T A A
 S L A A L L A H R A G D R I D A G A F L R Q D P Y W S A T L D D I P V G L E L L W P R L Q A D A Q R A G L R F S L S P A L N A L G Q A R I A K I

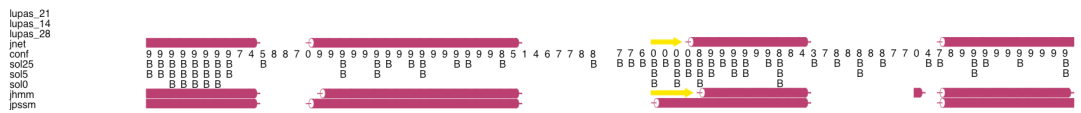


Figure S3 (continued).

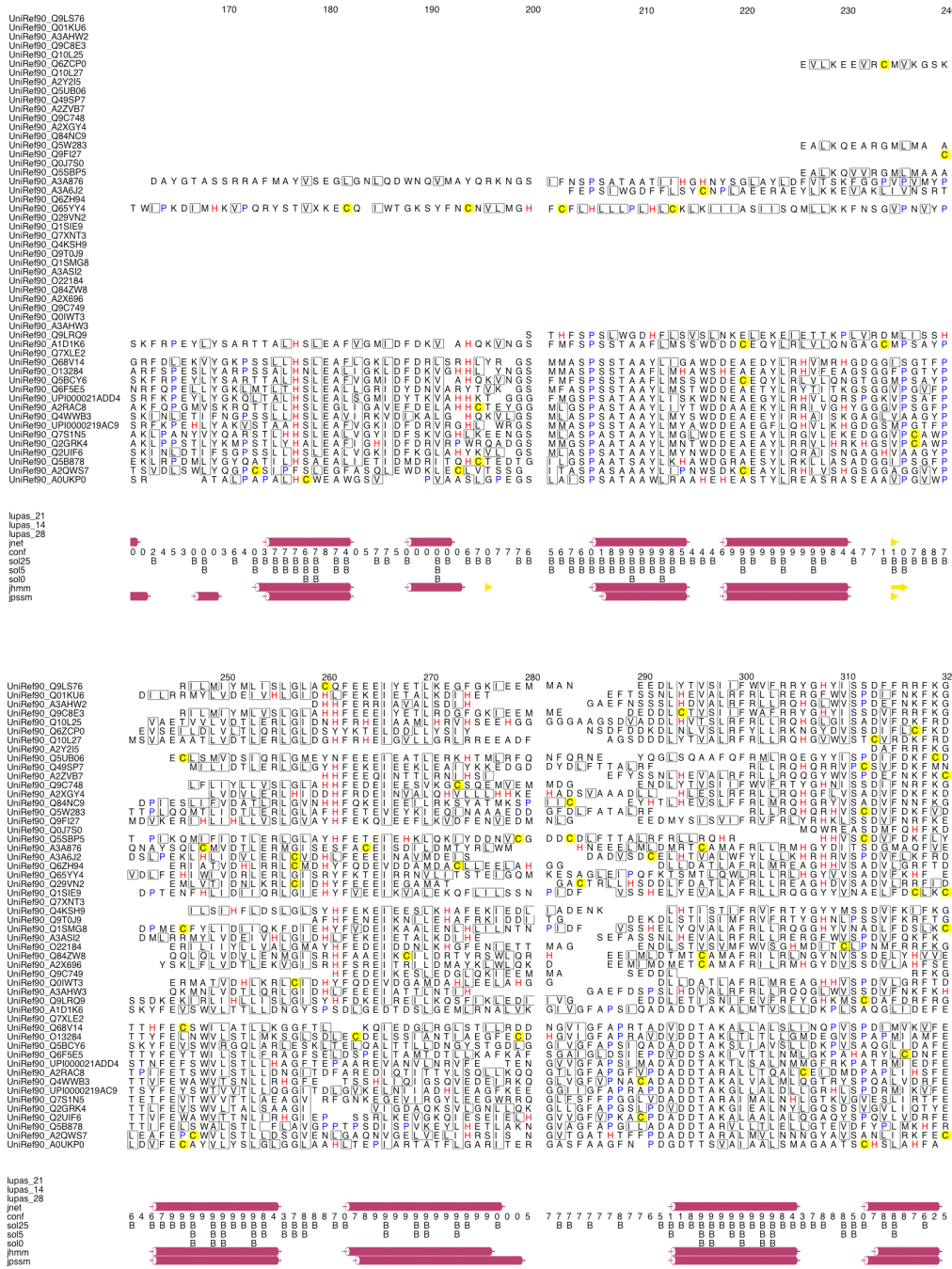


Figure S3 (continued).



Figure S3 (continued).



Figure S3 (continued).



Figure S3 (continued).

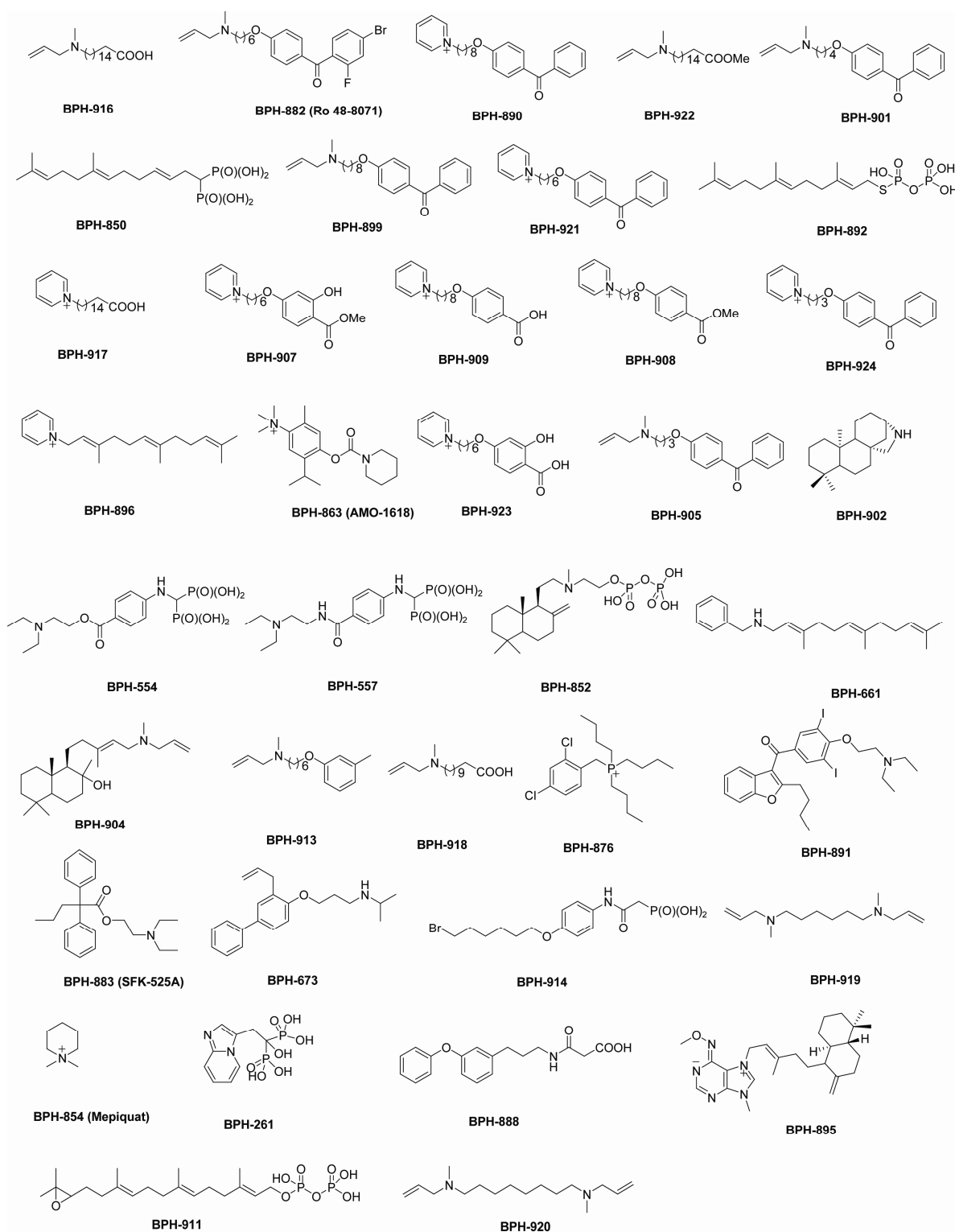


Figure S4: Structures of *Abies grandis* abietadiene synthase inhibitors investigated using GGPP as substrate. Structures are shown in order of decreasing potency (top left is the most potent).

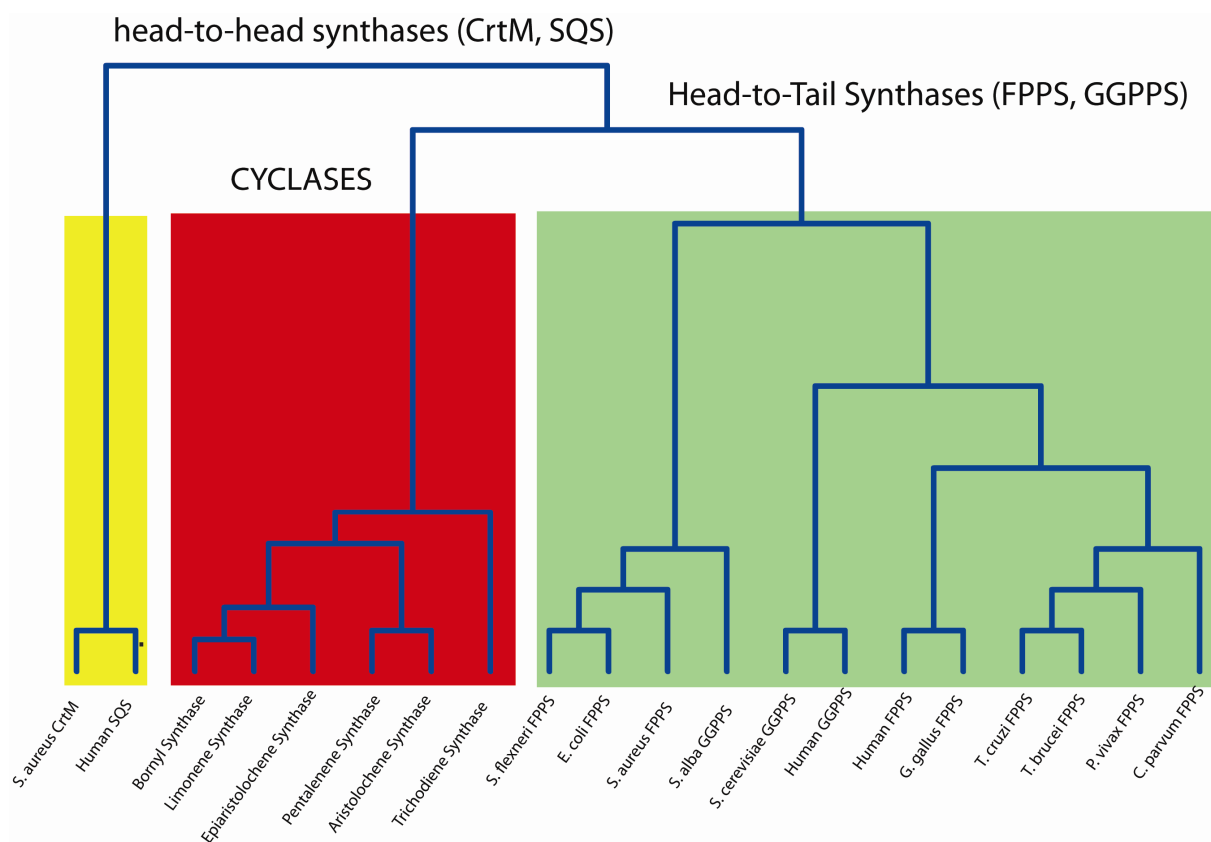
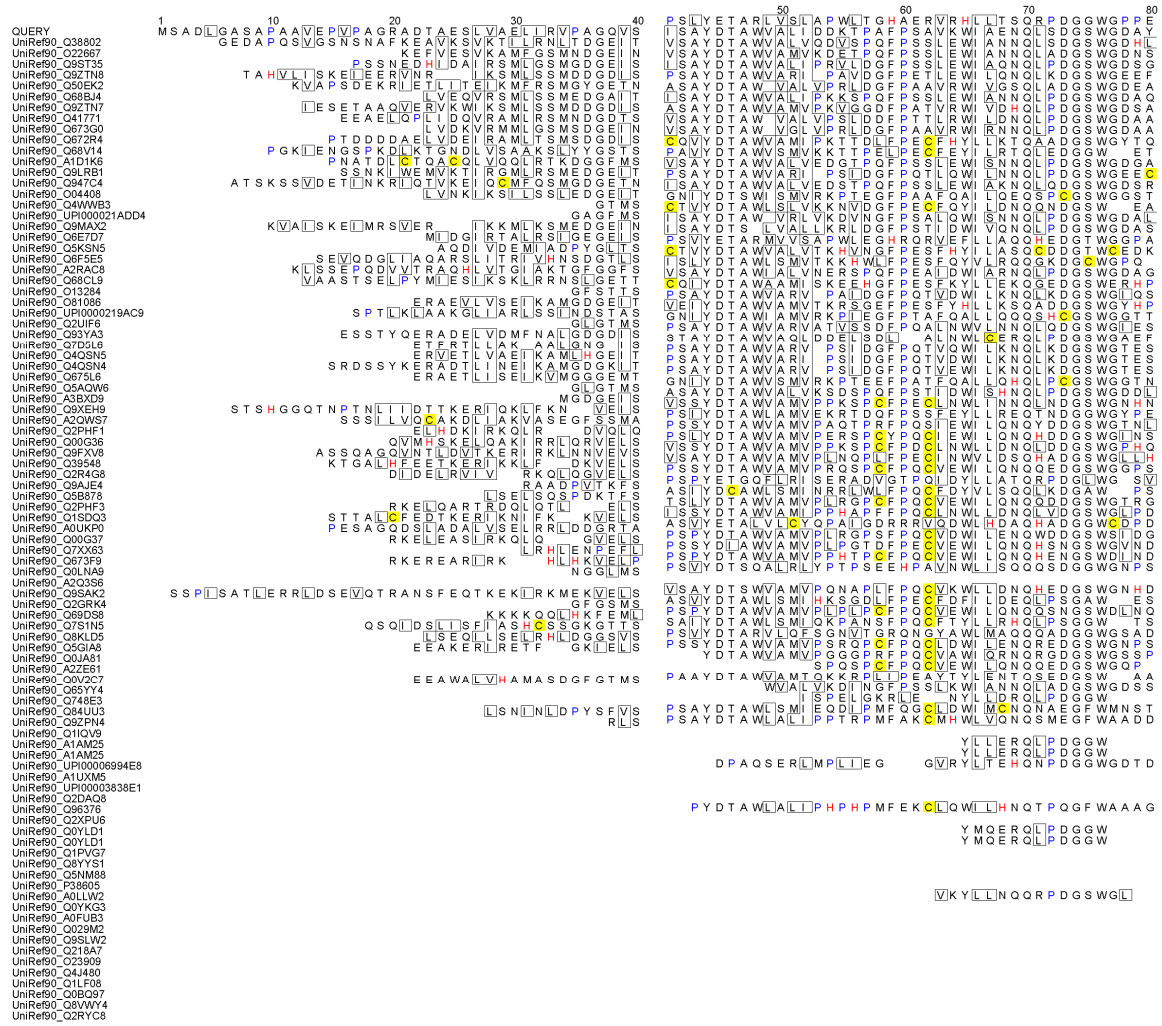


Figure S5: Structure dendrogram for α and $\alpha\beta$ proteins showing strong clustering of GGPPS/FPPS, separate from that seen with the monoterpene and sesquiterpene cyclases. This dendrogram was generated from the structural alignments obtained by using the Incremental Combinatorial Extension (CE) Program³.

Figure S6: JPRED3⁴ output using *Bradyrhizobium japonicum* cyclase blr2149 as the query.




```

QUERY
UniRef90_Q38802
UniRef90_Q22867
UniRef90_Q8ST35
UniRef90_Q82TN8
UniRef90_Q29K42
UniRef90_Q88B44
UniRef90_Q82TN7
UniRef90_Q41771
UniRef90_Q673G0
UniRef90_Q672R4
UniRef90_Q68V14
UniRef90_A1D1K6
UniRef90_Q9LR81
UniRef90_Q947C4
UniRef90_Q04408
UniRef90_Q44WB3
UniRef90_UPI000021ADD4
UniRef90_Q9MA22
UniRef90_Q867D7
UniRef90_Q8SKN5
UniRef90_Q6F5E5
UniRef90_Q29AC8
UniRef90_Q68CL9
UniRef90_Q13284
UniRef90_Q91086
UniRef90_UPI0000219AC9
UniRef90_Q2UJF6
UniRef90_Q93YA3
UniRef90_Q7D5L6
UniRef90_Q4QSN5
UniRef90_Q4QSN4
UniRef90_Q675L6
UniRef90_Q5A0W6
UniRef90_A38XD9
UniRef90_Q9XEH9
UniRef90_Q2QW57
UniRef90_Q2PFP1
UniRef90_Q0Q3G8
UniRef90_Q673V8
UniRef90_Q39548
UniRef90_Q2R4G8
UniRef90_Q94JE7
UniRef90_Q58876
UniRef90_Q2PFP3
UniRef90_Q15DQ3
UniRef90_A0UKP0
UniRef90_Q0Q537
UniRef90_Q7XG63
UniRef90_Q673F9
UniRef90_Q0LN49
UniRef90_Q2Q3G6
UniRef90_Q98AK2
UniRef90_Q2GR44
UniRef90_Q68D58
UniRef90_Q791N5
UniRef90_Q84LD5
UniRef90_Q5GIA8
UniRef90_Q0JA81
UniRef90_Q2Z5E1
UniRef90_Q0V2C7
UniRef90_Q65Y74
UniRef90_Q748E3
UniRef90_Q84UJ3
UniRef90_Q32PN4
UniRef90_Q1QV9
UniRef90_A1AM25
UniRef90_A1AM25
UniRef90_UPI00006994E8
UniRef90_A1UXM5
UniRef90_UPI00003383E1
UniRef90_Q2D4O8
UniRef90_Q98376
UniRef90_Q2XPJ6
UniRef90_Q0YLD1
UniRef90_Q0YLD1
UniRef90_Q1PYG7
UniRef90_Q8YYS1
UniRef90_Q5NM88
UniRef90_Q386D5
UniRef90_A0LW2
UniRef90_Q0YKG3
UniRef90_A0YU3
UniRef90_Q029M2
UniRef90_Q9SLW2
UniRef90_Q218A7
UniRef90_Q23909
UniRef90_Q4480
UniRef90_Q11F08
UniRef90_Q0BC97
UniRef90_Q8WY4
UniRef90_Q2RYC8

```

D D V I T D C L L S T L A C V L A L K R W Q L A P D M I K G L E F V N R N R L V M K Q K P S D V W F T M F P A M L E L A G A S S L R V D F S E N L N
P L Y A V D G N V N I S A S V K A L G H D K N A P H M V I R A R Q I L S V F T R I T L A P W H T A P A M P I E I V P L L Y A T Q P I G R L Q V N E G L T E
P L Y A V D G N V N I S A S V K A L G H D K N A P H M V I R A R Q I L S V F T R I T L A P W H T A P A M P I E I V P L L Y A T Q P I G R L Q V N E G L T E

Y Y G D G G N L S T T E A Y F A L K A G E H E P M R R A R E F I L S F T K I W L A D W K K V P S M P V E L V L L P S L S W A R G T V V P L S I R C P L P A
V P L M V L A A P R A V N R D D G L E D E L F

Figure S6 (continued).

QUERY
UniRef90_Q38802
UniRef90_Q22667
UniRef90_Q38T35
UniRef90_Q32T88
UniRef90_Q39EK2
UniRef90_Q68B14
UniRef90_Q32T87
UniRef90_Q41771
UniRef90_Q673G0
UniRef90_Q672F4
UniRef90_Q63V14
UniRef90_A1D1K6
UniRef90_Q3LR81
UniRef90_Q547C4
UniRef90_Q04408
UniRef90_Q4W8E3
UniRef90_UPI000021ADD4
UniRef90_Q6EMX2
UniRef90_Q6ETD7
UniRef90_Q5KSN5
UniRef90_Q6F5E5
UniRef90_Q2RAC8
UniRef90_Q68CL9
UniRef90_Q13284
UniRef90_Q31086
UniRef90_UPI000021AC9
UniRef90_Q2UIF6
UniRef90_Q39Y43
UniRef90_Q7D5L6
UniRef90_Q4Q3N5
UniRef90_Q4Q3N4
UniRef90_Q675L6
UniRef90_Q5ACW6
UniRef90_Q39X09
UniRef90_Q9XE19
UniRef90_Q2QW57
UniRef90_Q00G36
UniRef90_Q39XV8
UniRef90_Q39549
UniRef90_Q2R4G8
UniRef90_Q8AJE4
UniRef90_Q59B78
UniRef90_Q2PHF3
UniRef90_Q1SD02
UniRef90_A0UKP0
UniRef90_Q00G37
UniRef90_Q7XG63
UniRef90_Q673F9
UniRef90_Q0LN49
UniRef90_Q2Q3S6
UniRef90_Q9SAK2
UniRef90_Q2GRK4
UniRef90_Q69D38
UniRef90_Q7S1N5
UniRef90_Q8KLDS
UniRef90_Q5G1A8
UniRef90_Q0UJ481
UniRef90_A2ZE81
UniRef90_Q0Y2C7
UniRef90_Q6S9Y4
UniRef90_Q748E3
UniRef90_Q84UJ3
UniRef90_Q3ZPN4
UniRef90_Q1QV9
UniRef90_A1AM25
UniRef90_A1AM25
UniRef90_UPI00006994E8
UniRef90_A1UXM5
UniRef90_UPI00003838E1
UniRef90_Q2Q4C8
UniRef90_Q63V16
UniRef90_Q2XPU6
UniRef90_Q0YLD1
UniRef90_Q0YLD1
UniRef90_Q1PVG7
UniRef90_Q8Y51
UniRef90_Q5NM88
UniRef90_P38605
UniRef90_A0LLW2
UniRef90_Q0YK63
UniRef90_A0FUB3
UniRef90_Q029M2
UniRef90_Q9SLW2
UniRef90_Q21847
UniRef90_Q23909
UniRef90_Q4J480
UniRef90_Q1LF08
UniRef90_Q09C07
UniRef90_Q3WVY4
UniRef90_Q2RYC8

WR A A A R L R I P A G M D R R L A V H G L G A G A A L P E K L L H A L L E
L K D I Y A K K E L K L T R I P K E M V L M H K V P T T L L H S L E
L K E I Y A R R D I K L T K I I P M E V L M H K V P T T L L H S L E
L Q D I Y A S R O L K L T R I P K D I L M H K V P T T L L H S L E
R S Q I K L E T R I P K E L M H K V P T T L L H S L E
L Q O I L E K R E A K L R I P K D I L M H K V P T T L L H S L E
L K G I Y A N R E L K L K R I P K P D M M H K V P T T L L H S L E
L Q O I N A R R R L K L A R I P K P D M M H K V P T T L L H S L E
L O S I Y A N R E V K L R I P K P D M M H K V P T T L L H S L E
L Q R I Y A N R E V K L R I P K P M E M M H K I P T S I L H S L E
L O G I Y S R E L K M K R I P K E V M H T V P S I L H S L E
R D V I Y R M E E K K R I P K E V Y G K P S S L L H S L E
F E F G K D R A R E L N R I P K L S K F R P E Y L V S A R T I T A L L H S L E
F E K L R K R K R F K L K I P R D I L M H K V P T T L L H S L E
K R D E K L T R I P K E I L N V L H K V P T T L L H S L E
T F N F D R T Y L T M E I N A A K M S R F K P E Y L V G K T O L I T A L L H S L E
R N T K K K K I P K D I L M H K V P T T L L H S L E
L K A T Y A E R E R K L K P R D V L M H K P T S L L H S L E
R I L D L O A D T G T S G I R E L L R O N T G I L P E K T W S L I E
L I K G R E A L T E R E A K M N R P O P E L L Y O G K L I M I L L H S L E
L O D I Y A K R O L K L A K R E A H A R P T T L L H S L E
H E F E S H L L M Q M Y K E M A R F S P E S L Y R A P S S A L L H S L E
F P A O A G L R E M N A Q M S R F K P C H L Y E A K V S T A A H S L I E
F D F D R T Y L N K S I S T A F E R I T D I V A A D N P A N L M A V I S
D I L D I S K Y T A A S T S P D L S W L Y G P S S I L L H S L E
E E S I L I G E L V I G R E O Q L R K L G S K I N K H I L L S V L
S T P P K R O E R I A N L E R D A Y G A S S R R A F M A Y V I S
I R L L D T R K R O E R I A N L E S M D K I H F G G I L L S V L
I P D F E R R L T L K L S K K V S K I L N F D T I E S G P O S S I L L H S L E
O T D F S I L M K R E L E O K R C H S N E I D G I Y A I L S E
M O T D I D G I L I R K I E L E R D A Y G A S S R R A F M A Y V I S
T D I D E L F H L R K I E L E R D S G G T I S A R K A F M A Y V I S
V S I M V H E R E L L E L R R C H S N G R E A Y L A Y I A V I S
V D A I L W K K E L L E L R R C H S N G R E A Y L A Y I A V I S
M E T D I D S L F R L R E V E F E R D A G G T A S A R K A F M A Y V I S
A S Q E D F P P G A N A E L W O L S D R I A R G O A K P T A W H T L L E
E A L R A A K R P P G A N A E L W O L S D R I A R G O A K P T A W H T L L E
O T D I V D R L R E L R E L E R E A G H S Y G R K A Y A M A V I T
A I T O K R D R E L K R S F E A G S E G W K A Y L A Y I S
I G G A R I L A K I S R A T A P A G H P A L A Y I S
Q N D I D S L N H I R E M K T O R E A G H S Y G R K A Y A M A V I T
I D Q I T H L R D M E L K R I A G D E S I G K E A Y F A Y I A E G L I E
O T D I D G I L H O W E M L K R O A E E K S Y G R E A Y M A V I A E
L R L N I P L P R E L I A L I G H K R K I V I Y S W E
L K N F A A K R D E K K R P R E L M H K V P T T L L H S L E
D M I R K R D L K O D S E K F S K L M H K V P T T L L H S L E
R A L L R D A K L A K P P S T I Y K M P S T I Y H A L L E
R V C H Q E T E L A K P P S T I Y K M P S T I Y H A L L E
Q O A L H D A K L A K I L P A N Y I V Q P A R S T L L H S L I E
D P D P N L V N M M L G E R L T R E A L K W E F E G N M A N V E F A
D R D A G V I C H G G R T I T A Y I S
Y R E F R S R E A L M S N T A K I L A R I C S A L S S K K T T L H S L I E
K F E L L A A N Y A R R T K L I T W P K D I M H K
L F S T D K R H L D G F Q G Y W R K N R L K R F N E F I S A L L H K
S M K I Q O E L R E K I L V D G N Y P I L A Y L E V L P S K I Y
V L D E L F E N R N T L L R E E L S D K T Q Y A P L M F L E
L F S T D M L V H H D F R Y A W R K N R L K R T M H L V P G R I K R R
L F S T D M L V H H D F R Y A W R K N R L K R T M H L V P G R I K R R
O K F Y K L V K I P V V I Y A P A L V A L G O A R Y F H R P P N L M R G L
R I L V I E L S Q N R D D I L T R E E V D E K K Q O Y S P L L L F L E
R H L D K F R H H A W R K N V I A L D R V L K R T M H L V P G R I K O H
L F H L D K F R H H A W R K N V I A L D R V L K R T M H L V P G R I K O H
K S K E K E Y V P G S K H T H K I F F D R I A A F E R R P V P S I R N K
L Q D P S K S K M P A K A P H Q S W A L F K F L A G O D A V L R T P K R L R D H
A V Q A A E T W V D H Q E D S G D W G G Q P S V L A Y Y I G V P L D H E V T
A I K A A V I S F V D E R I N G V D G L G A I F P A M A N S I M M Y P A D H P N R
A V E K A K T W M I E R L H S D G L G A I Y P S V M A I L V I G A K D D P L P R
A K L A V D F E E R L N G E D G L G A I Y P P M A N T I M M Y P E D H P P R
A I R K A E A W F L P R M N G E D G L G A I F P P M V N Y L G Y P K D H P A R
A L L O R A V I E W S E A R M N G E D G G I F P P M V I S Y L D Y P E D H P L R
A V D F V R E R I N G E D G L G A I Y P A M A N S I M M Y S P D H P E A

Figure S6 (continued).

QUERY UniRef90_Q38802 UniRef90_Q22667 UniRef90_Q95T35 UniRef90_Q92TN8 UniRef90_Q50EK2 UniRef90_Q68BJ4 UniRef90_Q92TN7 UniRef90_Q41771 UniRef90_Q673C0 UniRef90_Q672R4 UniRef90_Q66V14 UniRef90_A1D1K6 UniRef90_Q9LRE1 UniRef90_Q947C4 UniRef90_Q04408 UniRef90_Q4WWB3 UniRef90_UPI000021ADD4 UniRef90_Q9NAK2 UniRef90_Q6E7D7 UniRef90_Q5KSN5 UniRef90_Q6F5E5 UniRef90_Q2RAC3 UniRef90_Q68CL9 UniRef90_Q13284 UniRef90_Q81086 UniRef90_UPI0000219AC9 UniRef90_Q2LHF5 UniRef90_Q93YA3 UniRef90_Q7D5L6 UniRef90_Q4G5N5 UniRef90_Q4G5N4 UniRef90_Q675L6 UniRef90_Q5AQW6 UniRef90_Q3BXD9 UniRef90_Q9XEH9 UniRef90_Q2QWS7 UniRef90_Q2PHF1 UniRef90_Q00G36 UniRef90_Q9FXV8 UniRef90_Q39548 UniRef90_Q2R4G8 UniRef90_Q9AJE4 UniRef90_Q5B678 UniRef90_Q2PHF3 UniRef90_Q1SDQ3 UniRef90_A0UKP0 UniRef90_Q00G37 UniRef90_Q7X263 UniRef90_Q673F9 UniRef90_Q0LNA9 UniRef90_Q2Q3S6 UniRef90_Q9S52 UniRef90_Q2GRK4 UniRef90_Q69DS8 UniRef90_Q7S1N5 UniRef90_Q9KLD5 UniRef90_Q5GIA8 UniRef90_Q0JA81 UniRef90_Q2ZE61 UniRef90_Q0V2C7 UniRef90_Q85YY4 UniRef90_Q748E3 UniRef90_Q84UU3 UniRef90_Q92FN4 UniRef90_Q1IQV9 UniRef90_A1AM25 UniRef90_A1AM25 UniRef90_UPI00008994E8 UniRef90_A1UXM5 UniRef90_UPI00003838E1 UniRef90_Q2DAQ8 UniRef90_Q96376 UniRef90_Q2XPL6 UniRef90_Q0YLD1 UniRef90_Q0YLD1 UniRef90_Q1PVG7 UniRef90_Q8YYS1 UniRef90_Q5NM88 UniRef90_P38605 UniRef90_A0LLW2 UniRef90_Q0YK63 UniRef90_A0FUB3 UniRef90_Q029M2 UniRef90_Q95LW2 UniRef90_Q218A7 UniRef90_Q23309 UniRef90_Q4J480 UniRef90_Q1LF08 UniRef90_Q0BQ87 UniRef90_Q8VWY4 UniRef90_Q2RYC8

RGRAYLLETGTTAHDHPGPPWLWHDKDLYRPA MIVRAAVVAAR HLAGAAGPATA
 490 500 510 520 530
 RGVSWLKSCL H SQDLTWTSKTAYE VGFVAEAYKLAAL QASLEVPAA
 RGRKFLQNS AGEKPEYLWVIEKVTYGS LLLSKSYVLAAL KVISERS
 KGLEFIFLARSQRSPTDDOQLWLDKTLAYAIPTVSDSYIMAAV QA
 RGEYLLQPKQSPSG DYLLWLEKVTYGS PVLISQVYRLLAAL KLSSSAPT S
 AGRFISGHVIEDDRRNPPLWHDKDLYTPVVRVIRAEITLATO RAAEAEKEARA
 RGITFL TTTIDKPPMEYLWSDKVISYGSAYLAEYVLAAR RAAESTS
 KGRDYLAHTDDWTGDYLLWLEKVTYKLPITLAEYVLAAM KI
 KAADVMQRWSCSDHDCD WTSKTAYRVDVIAAAYRLLAAM KASS
 RGAHYLKTTRKVGHH DASWTSKSAYRVAIVIAEAYELAAAL NQKAT
 RGVQFLTLTKGQRSQTDQDLWLDKTLAYALPTVSDSYIMAAAL QA
 RAGGWL SQHCEPPYAPLWMAKTLAY SATVVKAAALLSAL RLVI
 KGLQFLLSKQSRSQTDQDLWLDKTLFAPITVSDSYIVAAAV QA
 R GKQFLLETWTMGQK PDRITWTGKVLHGHSYVHDAHVLAAL KINRA
 RAKEFFTDESREN PPLWMGKDLYTFPRITVDVTVMCGR AVVCR
 RGGYLLTDHQHP RREDYVWLEKVISYVSSFLRKVITYVLAAL HASRKQSPCS
 KAFRLMRDCRSPRAQASMLWNKELSRPTRVVIRAVLELAAL WVVAHRRAEA
 RGYNWLCIHYRPRMKEYQWLNKEIYCPQRITDRAYELLSAM LAIV
 LGKAFLLNSNRDAWAKGRHLWLEKVTYSSPVLISE
 SGKAFLEAHKHHDWKGOYLWIEKVTYASPVLLCEAYELAA WLAGRRIPPT
 HALIEWMLARHEAHEPQTPLLWLGKELYCPTRVVIRVLELAAL
 R GKAYLEENHRKEWREAEERLWLEKTVYGS PNLISQAYQLAAM KVV VPTT
 RGVRYLLD HQNQWGTWE EKHFTGTGF
 RGITAWLLENQKPNGS WWEKWLITGT
 RGI EFLREEQEPFGGWFGRWGVIYIYGTWSV
 RGI EFLREEQEPFGGWFGRWGVIYIYGTWSV
 PAVGWLQRVIEGSSYFSKLLWYIYKLY PRVMTVTSLAAL QANASVPPA
 RGV DYLLGTQREHGLWD
 RGVN YLMRTQAGDGLW
 RGITAWLQSIVIEDDGSW
 RAARYLITNAQOMENGDFPQEEIMGVFNRRNMIYAAAYR
 RGIAYLKKKEOEPPGGWFGRWGVIYIYGTWSV LLSGL HQAGE
 RGIAYLKKKEOEPPGGWFGRWGVIYIYGTWSV LLSGL HQAGE
 RGI DFLISTQKEDGSWD
 RGINHLVSTQQPDGSW
 KALLDVRSEOEAEAGWFGWGVIYIYGTGAVLPA LAAL
 RAARYLITNAQOMENGDFPQEEIMGVFNRRNMIYAAAYR
 RGIQYLRLRTQKR DGTWEEL
 RGVQYL LATAQKSDGTWD
 RGI DYLQREORDGLWD
 HGLEYLLETQRSDGSWD
 RA AKLLINSOENGDFFPQEEIMGVFNKNMIYAAAYR
 RGVIEYLKSTQKGLWD
 RA AVGLINSQLENGDFPQEEIMGVFNKNMIYAAAYR
 RGIAYLLLAQOQTDGLW
 RGAAYLLDAQNE DGFW
 RGIAYLLQATQKPDGEWD
 RA AKVLLINMQMENGEFPQEEIMGVFNKNMIIS
 RGI RYLLLEAPRDGKWE EEELY

Figure S6 (continued).

UniRef90_C6WQZ4
UniRef90_Q39RT6
UniRef90_Q0AFY0
UniRef90_Q029U6
UniRef90_P84466
UniRef90_A0W5P5
UniRef90_Q3X7Y9
UniRef90_Q3JBH8
UniRef90_P55348
UniRef90_Q8DQK6
UniRef90_Q6IWA7
UniRef90_Q59080
UniRef90_Q59080
UniRef90_LPI000038D45A
UniRef90_LPI000068B70A6
UniRef90_Q2Y5X6
UniRef90_Q07I43
UniRef90_Q07I43
UniRef90_A1AN91
UniRef90_Q8BLN5
UniRef90_A0JND6
UniRef90_A0L1S0
UniRef90_P33990
UniRef90_Q7NE23
UniRef90_P73914
UniRef90_Q4CA42
UniRef90_A1D8P1
UniRef90_A1D8P1
UniRef90_P48449
UniRef90_Q2U3C4
UniRef90_Q2RH04
UniRef90_Q74FC1
UniRef90_Q759C3
UniRef90_A1DR10
UniRef90_Q86WJ0
UniRef90_Q3A7M9
UniRef90_Q62TY2
UniRef90_Q10231
UniRef90_A0RCV5
UniRef90_F54924
UniRef90_A0KC81
UniRef90_C6N3F4
UniRef90_A0LVK9
UniRef90_Q35C98
UniRef90_Q116B5
UniRef90_Q2E5Y5
UniRef90_Q7UR25
UniRef90_Q69446
UniRef90_A1DRH9
UniRef90_A1HVC3
UniRef90_A1CP22
UniRef90_Q4WDN6
UniRef90_Q2JET5
UniRef90_A1DRH6
UniRef90_Q57ZG6
UniRef90_A1DRH2
UniRef90_LPI0000699273
UniRef90_Q2B3H0
UniRef90_A1D8P8

1 10 20 30 40 50 60 70 80

S N S **C** M E A E W **V** L A M H L G I D D D P K Y I R A I L G E O R A D G S W G **V** F H

S N S **C** I E A E W I L A M **V** K D D P K Y D K **V** V Q A I L N E Q R E D G S W A **V** Y Y

R W I A D Q Q N D D G G F G D T D

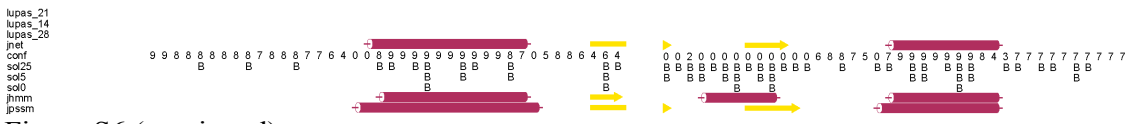


Figure S6 (continued).

UniRef90_Q8W3Z4
UniRef90_Q39RT6
UniRef90_Q0AFY0
UniRef90_Q029U6
UniRef90_P84466
UniRef90_A0V5P5
UniRef90_Q9X7V9
UniRef90_Q32BH8
UniRef90_P52346
UniRef90_Q8DGK8
UniRef90_Q6VWA7
UniRef90_Q59060
UniRef90_Q59060
UniRef90_UPI000039D45A
UniRef90_UPI00006B70A6
UniRef90_Q2Y5X6
UniRef90_Q07143
UniRef90_Q07143
UniRef90_A1AN91
UniRef90_Q881NS
UniRef90_A9JND6
UniRef90_A0LUS0
UniRef90_P33990
UniRef90_Q7NE23
UniRef90_P73914
UniRef90_Q4CA42
UniRef90_A1D8P1
UniRef90_A1D8P1
UniRef90_P48449
UniRef90_Q2JJC4
UniRef90_A2RH34
UniRef90_Q74FC1
UniRef90_Q796C3
UniRef90_A1DR10
UniRef90_Q96WJ0
UniRef90_Q347M9
UniRef90_Q62TY2
UniRef90_Q10Z31
UniRef90_A09GV5
UniRef90_P54924
UniRef90_A0KCC1
UniRef90_Q6H3F4
UniRef90_A0LVK9
UniRef90_Q35C98
UniRef90_Q118B5
UniRef90_Q2E5Y5
UniRef90_Q7UP25
UniRef90_Q68446
UniRef90_A1DRH9
UniRef90_A1HVC3
UniRef90_A1CPZ2
UniRef90_Q4WDM6
UniRef90_Q2ET5
UniRef90_A1DRH6
UniRef90_Q57ZG6
UniRef90_A1DRH2
UniRef90_UPI0000699273
UniRef90_Q283H0
UniRef90_A1D8P8

90 100 110 120

K A P N G D [] N T T [] E Y A A [] R A S G [] A P E S A P [] S S A R E

D A P A G D [] N A T [] E A Y A A [] R T A G F G A G D E R [] H G G [] K F T R Y

R S H S N [] A T S Y [] V [] A A W T [] S D Q [] T C E T T D A N A [] R R N



130 140 150 160

P E Q E R H

W L A P W E G T P T P P E L I F P R M A S W A R S T L V P L S L S A R R P

W L A P W D E T P A L A P E I L L P A W C P L A R A T L V P L S V L S V R P V

W L Q L A G E L D G L R D K T F V V P I L T N M A L G L V P W K K V S A L P F



Figure S6 (continued).

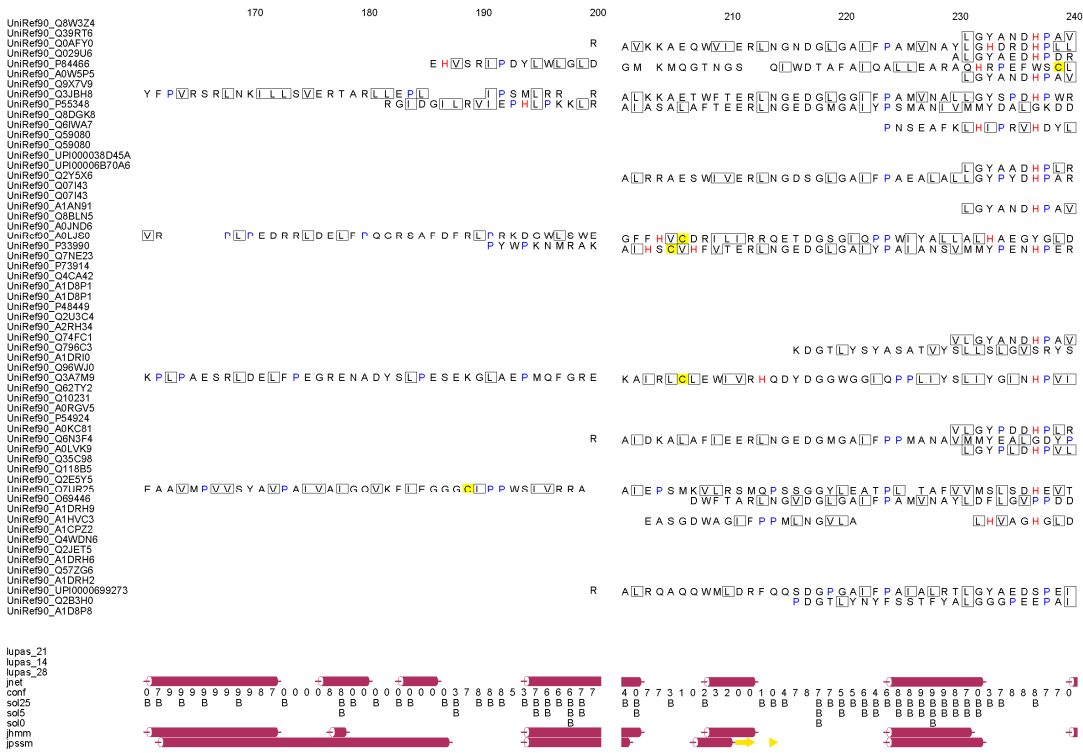


Figure S6 (continued).

```

UniRef90_Q8W3Z4
UniRef90_Q39T6
UniRef90_Q0AFY0
UniRef90_Q029Y6
UniRef90_P84466
UniRef90_A0W5P5
UniRef90_Q8X7V9
UniRef90_Q3JBH6
UniRef90_P55348
UniRef90_Q8DC89
UniRef90_Q8W47
UniRef90_Q59080
UniRef90_Q59080
UniRef90_UP1000038D45A
UniRef90_UP100008870A6
UniRef90_Q2Y5X6
UniRef90_Q07143
UniRef90_Q07143
UniRef90_A1AN91
UniRef90_Q8BLN5
UniRef90_A0JND6
UniRef90_A0LJ50
UniRef90_P33990
UniRef90_Q7NE23
UniRef90_P73914
UniRef90_Q4CA42
UniRef90_A1D8P1
UniRef90_A1D8P1
UniRef90_P48449
UniRef90_Q2J0C4
UniRef90_Q2RH34
UniRef90_Q74FC1
UniRef90_Q796C3
UniRef90_A1DR10
UniRef90_Q96WJ0
UniRef90_Q3A7M9
UniRef90_Q82T72
UniRef90_Q10231
UniRef90_A0RGV5
UniRef90_P54924
UniRef90_A0KC81
UniRef90_Q8N3F4
UniRef90_A0L1V9
UniRef90_Q35C98
UniRef90_Q118B5
UniRef90_Q2E5Y5
UniRef90_Q7UR25
UniRef90_Q89446
UniRef90_A1DRH9
UniRef90_A1HYC3
UniRef90_A1CP22
UniRef90_Q4WDN6
UniRef90_Q2JET5
UniRef90_A1DRH6
UniRef90_Q57ZG6
UniRef90_A1DRH2
UniRef90_UP10000699273
UniRef90_Q2B3H0
UniRef90_A1D8P8

```

250 260 270 280 290 300 310 320

A K G I Q A L N F C I E S D D E L V S C I T S P V W D T A T L A L M A M V D S E
Q O O R Q S L R E L L V D E G E O P C V S P V W D T V L A T L L A L Q E A G
S E A L I R H F E S L L L E T D D F O P C V S P V W D T A L C A F A L Q E A G
R K A H E Y I R I S Q I V P D R F P Q K Y Y R H M E S F S T L D G W
A K G L E A L N F T I E D S D S L V S C I T S P V W D T A L V L L Q A M O E A S
V I Q A K K A L Q N L V K N S A S O P C L S P V W D T L A T T I A L A D A G G
A I A R R A L D K L L V T I G E E O P C L S P V W D T A L T C H A L L Q E I V G
W V A E D G M K M Q G Y N G S G V I Q P C V S P V W D T A L V M R A M V D S G
L W D T A F A V I Q A L T A T G

Q O V R K G L R D L V R A Y V O P C V S P V W D T A V M R A L V I E S S
E O R K A L R L L L V E E G E O P C V S P V W D T V L T C L A F E D T E

A K G L D A L A N F C I E D D D C I T S C V S P V W D T A L A L V A L Q E A D
L R I L S

H P I R A G L R A F D S H W S Q A S E S P V W D T L S I R A L A D G
A I A R R A V E K L M D K E V I Y O P C L S P V W D T A L T M R A L V D S G
Q A C V S P V W D T A W V V R A L A E A D
Q A C V S P V W D T A W V V R A L T D S G
P T A W R S D P F A L T L S I L H R L D
P T A W R S D P F A L L F T L S I L H R L D

K G K R V O P C V S P V W D T T L M S I A L S D A A
S T V W D T A L L N Y A L M K S G

A K G I Q A L N F C I E G E D E L V S C I T S P V W D T A L L G Y A L Q E A N
P R A L T G L R S L V T K C N G L P Y L S T V W D T A L I S Y A L Q K N G

S K G L D A F N P A Y E K N G G V Y L Q O S E S P V W D T S L V I V A M O E A G
S T V W D T A L L S Y A M O E S G
V W E T S F T L Q A L V I E S S
V W D T A L L S Y A L Q E A D
V W D T T L T A H A L L E A G
V W D T A L A S H A V L E A D
V W D T A L A V T A L R D A G

E A D R F A I E N E D S Y Q A C V S P V W D T A V W T R S L V D S G
S P S T V W D T A L L S Y A L Q E A N
Q N G R F R L D S M L P D G W P I D T N I A N W A T S I A I T A
E R A R E S L D R L L V F Q G O P C V S P V W D T A L L S I T L L Q E A G

A G L E A E R F A V S D R E G F R I E A Q S P V W D T T L A L T G L L D S G
C I T S P V W D T L L M L R A L O D A S

S P V W D T A L A V V I L S D A G
S P V W D T P W M L E A L L E T G

Q Y C L D H L E R L V L E D G E L O P C K S P V W D T S I T L R A L Q E A G
R N A V A G L R M T V K A D G R T I Q Y T A V W N T A L T S H A L A E A G
T T L R D S S R M L T G L E D A G

V L I E D O P G D L N F W Y R H I T S K G P F F S T A D H G W P I L S D C T A E G L L K
W P T D P A L V K A Q W L L S R E G F F Q N D W Y P D V D D S G I V M M
O D S E P I R R A L D W I T O I L D E P F F S N H Y P D O Y A N H Y P D I D D T A A V I A W
N T D D R T L A A D W L I S D W S I K R F E F A N E F Y P D I D D T A M V L L
V I A D P A E A L K S I L L L Q E A F F E F C N D W Y P D V D D S T V V M I
V I L D H P S I L K A S O W L D R A F F E F C N D W Y P D V D D S T V V M I
V P E D H P O L V K A S D W D W S V K R F E F H N D Y P D I D D T A E V V L
G G H T A P V I G D W O V O H P N L K G G W A N S Y P D I D D T A A V I A W
N A W A K A K G S D W L R G W L D V K G F O Y N N A H Y P D I D D T A A V V M
V I A P D L S K Q L L D Y G D W H I K N K K F E F E N R F Y P D I D D T A A V V M
L V L D O P G D L S Y W Y R H I S K G P F S T A D H G W P I L S D C T A E G L L K
E K A L E W I L A H O E A D G G W G P W F Y A L L A L K
E K A L E W I L A H O E A D G G W G P W F Y A L L A L K
S S A Q V R H A L D W L Q A D W S E H F O F R N D H Y P D I D D T A A V I A W
V I D Q K P I R K A L D W L V A D W O E D H F O Y A N H Y P D I D D T A A V I A W
I D H V L Q W I I R H O D Y D G S W G G F W I Y G L M
V P A D E V R R K G D W O V K C P E L E G G G W A N D W Y P D V D D S G F V M I
O V R P E P W S R K A E Y L R I S Q V P D N F P D X O K Y Y R H M K
E E R K S I A S A L E W L N D W A V R A F O R A N S F Y P D I D D T A V I A L
G D E A E K S A I G D W A W R R D L R P G G W A N D Y P D V D D T A V I T M
L P P D L K Q L L D Y G W A V K R R A F E F N R F Y P D I D D T A V V M
L S K D H F A L Y K A T Y G D W O L K R F E F N D W Y P D I D D T A V V M
L E K K O V M E Y G D W A V K N K A G K P G G W A N R F Y P D I D D S A T V V M
G L R S F N K G V I D W L E H Q E D M G D I G D I M P L H G A M L
G L R S F N K G V I D W L E H Q E D M G D I G D I M P L H G A M L
Q V P D N P D Y O K Y Y R O M R K G S F S D G W I V S D C T A E A L L K
T R N O I P R G D W R V Y R R L A P G G F F E N S Y P D I D D S A A I L
I S D N O I T S A A R F L E R Q K K G F N I N N P D O D D T A V L L K
V D H P S L S K A A Q W L D R E V R P F F M N D W Y P D V D D S G I V M M
V I T E T D V T K A A D F L L A D W S V K G F N I N N P D O D D T A V L L K
N P V I D P A L V K A Q W L L D R E A F F L N D W Y P D V D D S G F V M I
L D P E K K A L D F L D K O I L O K N F S T R O O G Y T V S D C T A E A L L K
S F D O T P P A L D W I L S D W G V K A F E R A N T A Y P D V K D S A A V D A
V P E S S T S A A D Y L L K R Q H K G F N I N N P D I D D T A A A L K
L Y E K F K D I A K A L Y N S L G A W P F N I T O G Y T V S D T S E A L R
V S K D N I Q N A T A Y L L K Q H T K K G F S D V N T P D I D D T A V L R
I P K Q E L E V K G D W A V K R P D V R P G G W A N A Y Y P D I D D T A V V M
V A A P E T G D T A S G A L R D L D K R A F O Y A N Y P D I D D T A V I A R
G G A K V R P A L D W L R Q V L D L K G F O Y N N A H Y P D I D D T A V V M
L P A O H P R V Q A A R W L V S E A F E F A N D Y P D I D D T A E V V L
V I E W M I R H O D P D G A W G P W I Y G L M
I S P S O I L D Y G D W V F K N K F G K P G G W A N R F Y P D I D D T A V V M
N E S T L O T A E Y I L Q R H E K G F S D V N T T P D I D D T A V I N W
U T M D P D D S G Y Q E R H T G A D P G S W G T S V D S O D T E G A I L
V A R H T A D L R I D A A L S D W R E G G W F G N D Y P D V D D S G I V A H
V P T D X I S K Q I L T G V H G D W A V K N F E F E N D Y P D I D D T A E V V L
E S T D P R L V A A R R W E G M F E Y A N S W Y P D I D D T A A V I V I
L G I K S D P R L A D A A W D W R Y I Q F E Y S N T W Y P D I D D T A A V L
V P A D H P A M V R A G R W L V D E E A F E Y N T W Y P D I D D T A A V L
V P P G D P A K Q I L T G V R G D W S P K G F E F E N D Y P D I D D T Q V L T
P A N D P X I L L T E V R G D W A V K N F E F E N D Y P D I D D T A E V L R
L G L A E P T O V E W L S D W T N N V F E Y A N E F Y P D I D D T S M G I M
V P P K E H R R F G D W I V H N T K A E P G G W G N T I N P D I D D T A A L R
L D A P W F R K S O W L O R C L L P D V I N N D A S K T F E V I D D M A S A L H

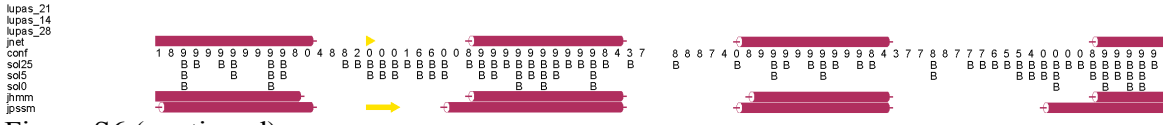


Figure S6 (continued).

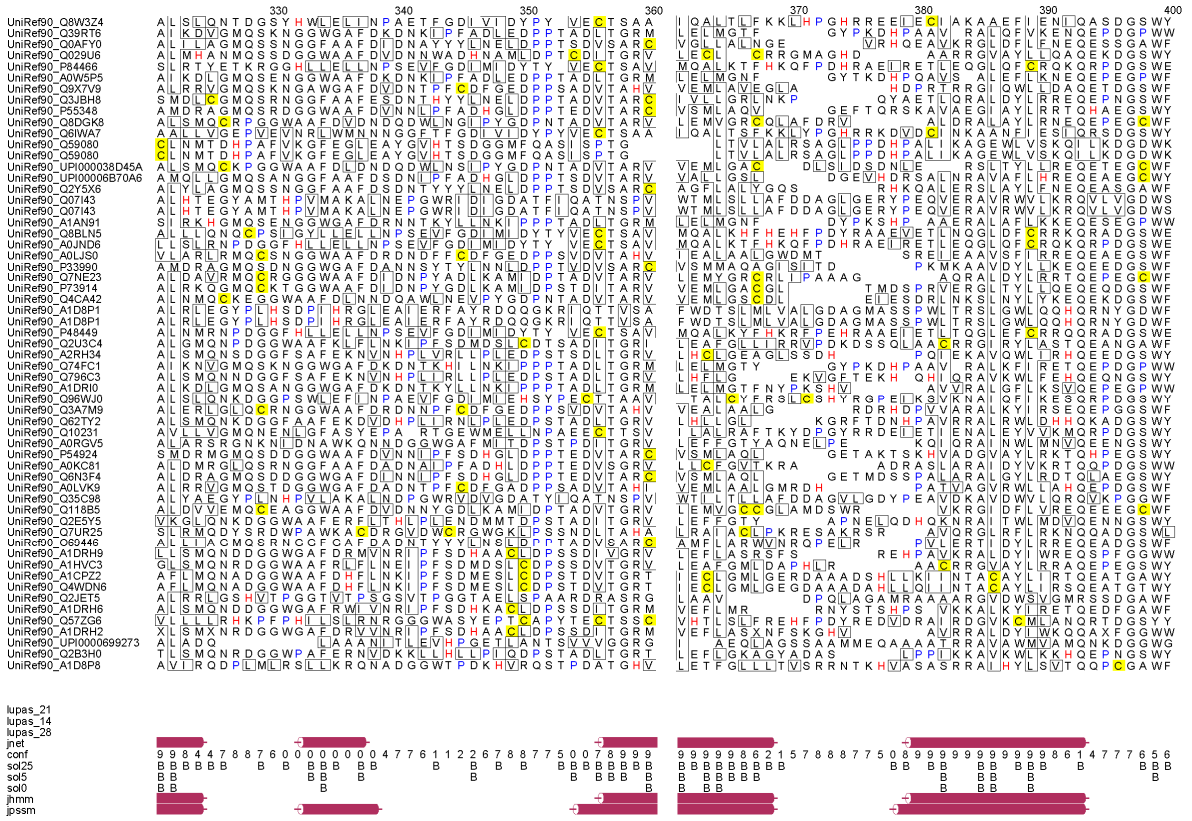


Figure S6 (continued).

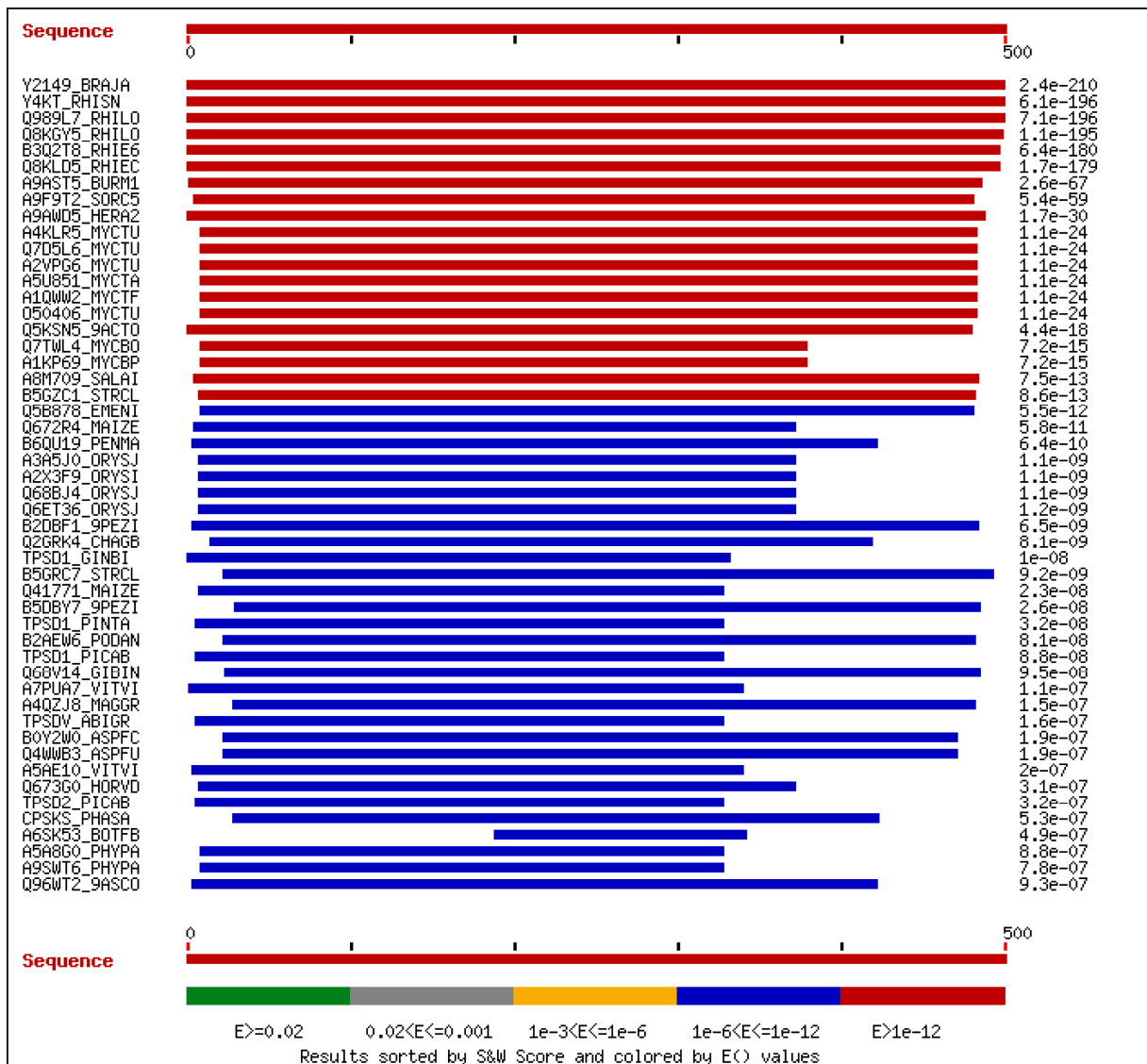


Figure S7: FASTA⁶ search program output alignment by using *Bradyrhizobium japonicum* γ -domain only as the query. Note the high similarity to *M. tuberculosis* Rv3377c, rice and maize CPS.

Supporting Experimental Section:

Synthetic Aspects. All reagents used for synthesis were purchased from Aldrich (Milwaukee, WI). The purities of all compounds were routinely monitored by using ^1H and ^{31}P NMR spectroscopy at 400 or 500 MHz on Varian (Palo Alto, CA) Unity spectrometers using, in some instances, absolute spin-count quantitative analyses. **BPH-882** was purchased from Biomol international (Plymouth Meeting, PA); **BPH-863** was purchased from EMD Biosciences, Inc. (San Diego, CA). **BPH-891**, **BPH-854** were purchased from Sigma-Aldrich. The syntheses of **BPH-892**, **BPH-904**, **BPH-852**, **BPH-911**, **BPH-902** were reported previously⁷⁻¹⁰. **BPH-895** was a gift from Dr. L.L. Gundersen. **BPH-850** was prepared using the procedure reported¹¹⁻¹³. **BPH-261 (Minodronate)**, **BPH-661**, **BPH-673** were available from our previous work^{12, 13}.

General procedure for BPH-554 and BPH-557. An amine (3 mmol), triethyl orthoformate (3.6 mmol), and diethyl phosphite (12 mmol) were heated to 140 °C under N_2 overnight. The mixture was subjected to column chromatography to give the tetraethyl ester, which was then treated with TMSBr (8 equiv) in acetonitrile for 8 h at room temperature. Upon removal of the solvent, the residue was dissolved in ethanol or acetone to precipitate the substituted aminomethylene bisphosphonate.

{[4-(2-Diethylamino-ethylcarbamoyl)-phenylamino]-phosphono-methyl}-phosphonic acid. (BPH-557).

Anal. ($\text{C}_{14}\text{H}_{27}\text{N}_3\text{Na}_2\text{O}_9\text{P}_2$) C, H, N. ^1H NMR (400 MHz, D_2O): δ 7.62 (d, 2H, $J = 8.0\text{Hz}$), 6.60 (d, 2H, $J = 8.0\text{Hz}$), 4.15-4.38 (m, 2H), 3.70 (t, 1H, $J = 22\text{Hz}$), 3.25-3.38 (m, 2H), 3.13 (t, 4H, $J = 7.2\text{Hz}$), 1.11 (t, 4H, $J = 7.2\text{Hz}$). ^{31}P NMR (162 MHz, D_2O): δ 16.1.

4-[(Bis-phosphono-methyl)-amino]-benzoic acid 2-diethylamino-ethyl ester. (BPH-554).

Anal. (C₁₄H₂₇N₃Na₂O₉P₂) C, H, N. ¹H NMR (400 MHz, D₂O): δ 7.47 (d, 2H, *J* = 8.0Hz), 6.64 (d, 2H, *J* = 8.0Hz), 3.89 (t, 1H, *J* = 21Hz), 3.45-3.60 (m, 2H), 3.00-3.19 (m, 4H), 1.15(t, 4H, *J* = 7.2Hz). ³¹P NMR (162 MHz, D₂O): δ 16.4.

BPH-876 was prepared by refluxing of 3, 4-dichlorobenzyl chloride (1 mmol) and tributyl phosphine (1 mmol) in acetone (5 mL). Removal of solvent gave BPH-876 as a white solid. Yield 89%. Anal. (C₁₉H₃₂Cl₃P) C, H. ¹H NMR (400 MHz, CDCl₃): δ 8.22 (d, 1H, *J* = 8.4Hz), 7.43 (s, 1H), 7.33 (d, 1H, *J* = 8.4 Hz), 4.56 (d, 2H, *J* = 12 Hz), 2.44-2.53 (m, 6H), 1.38-1.42 (m, 12H), 0.94 (t, 3H, *J* = 6.8 Hz).

BPH-916, BPH-922, BPH-901, BPH-899, BPH-905, BPH-904, BPH-913, BPH-918, BPH-919, BPH-920 were prepared using a similar procedure to that previously reported, using alkylbromides and methylallylamine as reactants¹⁴.

16-(N-allyl-N-methylamino)hexadecanoic acid•hydrobromide (BPH-916).

Anal. (C₂₁H₄₂BrNO₂) C, H, N. ¹H NMR (400 MHz, CDCl₃): δ 6.12-6.21 (m, 1H), 5.50-5.60 (m, 2H), 3.6-3.75 (m, 2H), 2.90-3.05 (m, 2H), 2.73 (s, 3H), 2.35 (t, 2H, *J* = 7.2 Hz), 1.80-1.98 (m, 2H), 1.61-1.69 (m, 2H), 1.20-1.40 (m, 22 H).

Methyl 16-(N-allyl-N-methylamino)hexadecanoate•hydrobromide (BPH-922).

Anal. (C₂₁H₄₁NO₂Br•0.3H₂O) C, H, N. ¹H NMR (400 MHz, CDCl₃): δ 11.5 (s, 1H), 6.14-6.19 (m, 1H), 5.51-5.60 (m, 4H), 3.66 (s, 3H), 3.55-3.70 (m, 3H), 2.81-3.09 (m, 2H), 2.30 (s, 3H), 2.30 (t, 2H, *J* = 6.5Hz), 1.79-2.00 (m, 2H), 1.59 (t, 2H, *J* = 6.5Hz), 1.24-1.32 (m, 24H).

(4-(4-(N-allyl-N-methylamino)butoxy)phenyl)(phenyl)methanone•hydrobromide (BPH-901).

Anal. (C₂₁H₂₆BrNO). C, H, N. ¹H NMR (400 MHz, CDCl₃): δ 7.85 (d, 2H, *J* = 8 Hz), 7.69 (d, 2H, *J* = 7.2Hz), 7.47 (t, 1H, *J* = 7.2Hz), 7.45 (t, 3H, *J* = 7.2Hz), 6.10-6.21 (m, 2H), 5.52-5.64 (m, 2H), 3.90 (t, 2H, *J* = 6.4 Hz), 3.60-3.67 (m, 2H), 2.80-3.18 (m, 2H), 2.77 (s, 3H), 1.68-2.00 (m, 4H).

(4-(8-(N-allyl-N-methylamino)octyloxy)phenyl)(phenyl)methanone (BPH-899).

Anal. (C₂₅H₃₄BrNO•0.3H₂O) C, H, N. ¹H NMR (400 MHz, CDCl₃): δ 7.83 (d, 2H, *J* = 8.8Hz), 7.74 (d, 2H, *J* = 7.2Hz), 7.57 (t, 1H, *J* = 7.2Hz), 7.46 (t, 3H, *J* = 7.2Hz), 6.14-6.21 (m, 2H), 5.51-5.61 (m, 2H), 3.62-3.68 (m, 2H), 2.80-3.18 (m, 2H), 2.73 (s, 3H), 1.78-2.00 (m, 4H), 1.32-1.49 (m, 8H).

(4-(3-(methylallylamino)propoxy)phenyl)(phenyl)methanone•hydrobromide (BPH-905).

Anal. (C₂₀H₂₄BrNO₂) C, H, N. ¹H NMR (400 MHz, CDCl₃): δ 11.8 (s, 1H), 8.04 (d, 2H, *J* = 9.0 Hz), 7.7 (d, 2H, *J* = 9.0Hz), 7.59 (t, 1H, *J* = 7.5Hz), 7.56 (t, 2H, *J* = 7.5Hz), 6.92 (t, 2H, *J* = 9.0Hz), 6.19-6.23 (m, 1H), 5.56-5.65 (m, 2H), 4.18 (t, 3H, *J* = 4.8 Hz), 3.60-3.82 (m, 2H), 3.10-3.42 (m, 2H), 2.81 (s, 3H), 2.4-2.60 (m, 2H).

(E)-N-allyl-5-((4aR,5R)-decahydro-1,1,4a-trimethylnaphthalen-5-yl)-N,3-dimethylpent-2-

en-1-amine•hydrobromide (BPH-904). Anal. (C₂₄H₄₄BrNO) C, H, N. ¹H NMR (400 MHz, CDCl₃): δ 6.12-6.20 (m, 1H), 5.42-5.80 (m, 4H), 3.4-3.75 (m, 4H), 2.65 (s, 3H), 0.6-2.4

6-(m-tolyloxy)-N-allyl-N-methylhexan-1-amine (BPH-913). Anal. (C₁₇H₂₇NO) C, H, N. ¹H NMR (400 MHz, CDCl₃): δ 7.18 (t, 2H, *J* = Hz), 6.82-6.87 (m, 3H), 5.8-6.0 (m, 1H), 5.11-5.21 (m, 2H), 3.96 (t, 2H, *J* = 6.8Hz), 3.00 (d, 2H, *J* = 6.4Hz), 2.31-2.35 (m, 5H), 2.21 (s, 3H), 1.76-1.80 (m, 2H), 1.39-4.58 (m, 6H).

10-(N-allyl-N-methylamino)decanoic acid•hydrobromide (BPH-918). Anal. (C₁₄H₂₈BrNO₂)

C, H, N. ¹H NMR (400 MHz, CDCl₃): δ 6.12-6.20 (m, 1H), 5.52-5.60 (m, 2H), 3.6-3.75 (m, 2H),

2.90-3.05 (m, 2H), 2.75 (s, 3H), 2.35 (t, 2H, $J = 7.2$ Hz), 1.80-1.98 (m, 2H), 1.61-1.69 (m, 2H), 1.20-1.40 (m, 10H).

N1, N6-diallyl-N1,N6-dimethylhexane-1,6-diamine•Hydrobromide (BPH-919).

Anal.(C₁₄H₃₀Br₂N₂) C, H, N. ¹H NMR (400 MHz, CDCl₃): δ 10.81 (s, 2H), 6.08-6.17 (m, 2H), 5.54-5.59 (m, 4H), 3.43-3.78 (m, 4H), 3.16-3.18 (m, 4H), 3.02-3.04 (m, 4H), 2.7 (s, 6H), 1.87-1.98 (m, 2H), 1.50-1.53 (m, 4H).

N1, N8-diallyl-N1, N8-dimethyloctane-1, 8-diamine•hydrobromide (1:2) (BPH-920).

Anal.(C₁₆H₃₄Br₂N₂) C, H, N. ¹H NMR (400 MHz, CDCl₃): δ 10.96 (s, 2H), 6.09-6.18 (m, 2H), 5.53-5.70 (m, 4H), 3.45-3.79 (m, 4H), 3.16-3.18 (m, 4H), 3.02-3.04 (m, 4H), 2.7 (s, 6H), 1.83-1.98 (m, 2H), 1.34-1.53 (m, 8H).

BPH-890, BPH-921, BPH-917, BPH-907, BPH-909, BPH-908, BPH-924, BPH-896, BPH-923 were prepared by refluxing different alkyl bromides in pyridine followed by removal of solvent and recrystallization from acetone.

[(4-phenylcarbonyl)-phenoxyoctyl]pyridinium bromide (BPH-890). Anal. (C₂₆H₃₀BrNO₂) C, H, N. ¹H NMR (400 MHz, CDCl₃): δ 9.54 (d, 2H, $J = 5.6$ Hz), 8.50 (t, 1H, $J = 8$ Hz), 8.10 (t, 2H, $J = 6.8$ Hz), 7.80 (d, 2H, $J = 6.8$ Hz), 7.60 (d, 2H, $J = 6.8$ Hz), 7.55 (t, 1H, $J = 7.6$ Hz), 7.50 (t, 2H, $J = 7.6$ Hz), 6.95 (d, 2H, $J = 8$ Hz), 5.05 (t, 2H, $J = 7.6$ Hz), 4.02 (t, 2H, $J = 6.4$ Hz), 2.04-2.08 (m, 2H), 1.74-1.82 (m, 2H), 1.36-1.46 (m, 8H).

[6-(4-(formylphenyl)(phenoxy))hexyl]pyridinium bromide (BPH-921). Anal. (C₂₄H₂₆BrNO₂) C, H, N. ¹H NMR (400 MHz, CDCl₃): δ 9.60 (d, 2H, $J = 5.6$ Hz), 8.46 (t, 1H, $J = 7.6$ Hz), 8.07 (t, 2H, $J = 7.2$ Hz), 7.64-7.17 (m, 4H), 7.50 (t, 1H, $J = 7.8$ Hz), 7.42 (t, 2H, $J = 7.8$ Hz), 6.90 (t, 2H, $J = 8.8$ Hz), 5.00 (t, 2H, $J = 7.2$ Hz), 3.95 (t, 2H, $J = 6.0$ Hz), 2.00- 2.07 (m, 2H), 1.73-1.78 (m, 2H), 1.43-1.47 (m, 4H).

[1-carbamoylhexadecanyl]pyridinium chloride (BPH-917). Anal. ($C_{21}H_{36}ClNO_2 \bullet 0.3H_2O$) C, H, N. 1H NMR (400 MHz, CD_3OD): δ 9.56 (d, 2H, $J = 5.6$ Hz), 8.42 (t, 1H, $J = 7.6$ Hz), 8.01 (t, 2H, $J = 6.8$ Hz), 4.9 (t, 2H, $J = 7.6$ Hz), 2.60 (t, 2H, $J = 7.2$ Hz), 2.2-2.4 (m, 2H), 1.30-2.4 (m, 24H).

[6-(4-(methoxycarbonyl)-3-hydroxyphenoxy)hexyl]-pyridinium bromide (BPH-907). Anal. ($C_{19}H_{24}BrNO_4 \bullet 1.4H_2O$) C, H, N. 1H NMR (400 MHz, D_2O): δ 8.74 (d, 2H, $J = 6$ Hz), 8.35 (t, 1H, $J = 7$ Hz), 7.82 (t, 2H, $J = 7$ Hz), 7.54 (d, 1H, $J = 9$ Hz), 6.23 (d, 1H, $J = 9$ Hz), 6.20 (s, 1H), 4.31 (t, 2H, $J = 7.5$ Hz), 3.80 (t, 2H, $J = 6$ Hz), 3.70 (s, 3H), 1.82-1.85 (m, 2H), 1.50-1.80 (m, 2H), 1.29-1.51 (m, 4H).

[8-(4-carbonyl)phenoxy]octyl] pyridinium bromide (BPH-909). Anal. ($C_{20}H_{26}BrNO_3$) C, H, N. 1H NMR (400 MHz, D_2O): δ 8.63 (d, 2H, $J = 5.5$ Hz), 8.33 (t, 1H, $J = 8$ Hz), 7.82 (t, 2H, $J = 6.5$ Hz), 7.78 (d, 1H, $J = 9$ Hz), 6.84 (d, 1H, $J = 8.5$ Hz), 4.37 (t, 2H, $J = 7.5$ Hz), 3.92 (t, 2H, $J = 6.5$ Hz), 1.75-1.82 (m, 2H), 1.55-1.80 (m, 2H), 1.29-1.51 (m, 8H).

1-[8-(4-(methoxycarbonyl)phenoxy)octyl]-pyridinium bromide (BPH-908). Anal. ($C_{21}H_{28}BrNO_3 \bullet H_2O$) C, H, N. 1H NMR (400 MHz, D_2O): δ 8.66 (d, 2H, $J = 5.5$ Hz), 8.33 (t, 1H, $J = 8$ Hz), 7.86 (t, 2H, $J = 6.5$ Hz), 7.52 (d, 1H, $J = 8.5$ Hz), 6.52 (d, 1H, $J = 8.5$ Hz), 4.37 (t, 2H, $J = 7.5$ Hz), 3.66 (s, 3H), 3.51 (t, 2H, $J = 6.5$ Hz), 1.71-1.75 (m, 2H), 1.30-1.55 (m, 2H), 0.98-1.20 (m, 8H).

[6-(4-(phenylcarbonyl)(phenoxy))propyl]pyridinium bromide (BPH-924).

Anal. ($C_{21}H_{20}NO_2Br$) C, H, N. 1H NMR (400 MHz, $CDCl_3$): δ 9.62 (d, 2H, $J = 6$ Hz), 8.48 (t, 1H, $J = 8.0$ Hz), 8.04 (t, 2H, $J = 8.0$ Hz), 7.60-7.63 (m, 4H), 7.48 (t, 1H, $J = 7.5$ Hz), 7.36 (t, 2H, $J = 7.5$ Hz), 6.73 (d, 2H, $J = 7.5$ Hz), 5.27 (t, 2H, $J = 7.0$ Hz), 4.23 (t, 2H, $J = 5.5$ Hz), 2.57-2.62 (m, 2H).

[6-(4-(carboxy)-3-hydroxyphenoxy)hexyl]-pyridinium chloride (BPH-923A). Anal. ($C_{18}H_{22}ClNO_4$) C, H, N. 1H NMR (400 MHz, D_2O): δ 8.64 (d, 2H, $J = 6$ Hz), 8.32 (t, 1H, $J = 8$ Hz), 7.82 (t, 2H, $J = 7$ Hz), 7.50 (d, 1H, $J = 9$ Hz), 6.26(d, 1H, $J = 9$ Hz), 6.18 (s, 1H), 4.41 (t, 2H, $J = 7.5$ Hz), 3.80 (t, 2H, $J = 6$ Hz), 1.80-1.84 (m, 2H), 1.49-1.80 (m, 2H), 1.19-1.51 (m, 4H).

General procedure for BPH-888 and BPH-914. To a solution of malonic acid, monoethyl ester (1 mmol) were added N-ethyl-N-(3-dimethylaminopropyl)-carbodiimide (EDC) (1.5 mmol) and 1-hydroxybenzotriazole (1 mmol). After stirring for 2h at room temperature, 50 mL of ethyl acetate was added and the reaction mixture washed successively with 1N HCl (5 mL), water (5 mL) and saturated $NaHCO_3$ (5 mL), dried, and evaporated. The amide was purified using flash chromatography.

N-[3-(3-Phenoxy-phenyl)-propyl]-melonic acid monoamide (BPH-888). 3-(3-phenoxy-phenyl)-propylamine was coupled with monoethyl ester of malonic acid to give the ethyl ester of BPH-888, which was then hydrolyzed with 3 equiv. of KOH in MeOH/ H_2O for 1h. The reaction mixture was acidified, extracted with ethyl acetate, and the organic layer evaporated. The oily residue was dissolved in methanol, neutralized with KOH and evaporated to give **BPH-888** as a white powder (250 mg, 66% overall yield). Anal. ($C_{18}H_{18}KNO_4 \cdot 0.5H_2O \cdot 0.25KCl$) C, H, N. 1H NMR (400 MHz, D_2O): δ 6.70-7.40 (m, 9H), 3.33-3.40 (m, 2H), 3.32 (s, 2H), 2.62 (t, 2H, $J = 7.6$ Hz), 1.80-1.90 (m, 2H).

(4-(6-bromohexyloxy)phenylcarbamoyl)methylphosphonic acid (BPH-914). 4-(6-bromohexyl)aniline was coupled with dibenzyl phosphonoacetic acid, according to the general method, to give the dibenzyl ester of **BPH-914**. The benzyl groups were removed by hydrogenation for 1hr, catalyzed with 5% Pd/C in methanol, followed by neutralization with NaOH to give **BPH-914** as a white powder (289 mg, 60%). Anal. ($C_{14}H_{19}NaNO_5 \cdot 2.5H_2O$) C, H,

N. ¹H NMR (400 MHz, D₂O): δ 7.22 (d, 2H, *J* = 9Hz), 3.90 (t, 3H, *J* = 7Hz), 3.45 (t, 3H, *J* = 7Hz), 2.51 (d, 2H, *J* = 18.5Hz), 1.56-1.62 (m, 4H), 1.28-1.32 (m, 4H).

References:

1. Fuchs, P. F.; Alix, A. J., High accuracy prediction of beta-turns and their types using propensities and multiple alignments. *Proteins* **2005**, 59, (4), 828-39.
2. Krissinel, E.; Henrick, K., Secondary-structure matching (SSM), a new tool for fast protein structure alignment in three dimensions. *Acta Crystallogr D Biol Crystallogr* **2004**, 60, (Pt 12 Pt 1), 2256-68.
3. Shindyalov, I. N.; Bourne, P. E., Protein structure alignment by incremental combinatorial extension (CE) of the optimal path. *Protein Eng* **1998**, 11, (9), 739-747.
4. Cuff, J. A.; Clamp, M. E.; Siddiqui, A. S.; Finlay, M.; Barton, G. J., JPred: a consensus secondary structure prediction server. *Bioinformatics* **1998**, 14, (10), 892-893.
5. Vogel, B. S.; Wildung, M. R.; Vogel, G.; Croteau, R., Abietadiene synthase from grand fir (*Abies grandis*). cDNA isolation, characterization, and bacterial expression of a bifunctional diterpene cyclase involved in resin acid biosynthesis. *J Biol Chem* **1996**, 271, (38), 23262-8.
6. Pearson, W. R.; Lipman, D. J., Improved tools for biological sequence comparison. *Proc Natl Acad Sci U S A* **1988**, 85, (8), 2444-8.
7. Roy, A.; Roberts, F. G.; Wilderman, P. R.; Zhou, K.; Peters, R. J.; Coates, R. M., 16-Aza-ent-beyerane and 16-Aza-ent-trachylobane: potent mechanism-based inhibitors of recombinant ent-kaurene synthase from *Arabidopsis thaliana*. *J Am Chem Soc* **2007**, 129, (41), 12453-60.
8. Ravn, M. M.; Jin, Q.; Coates, R. M., Synthesis of allylic isoprenoid diphosphates by S_N2 displacement of diethyl phosphate. *Eur. J. Org. Chem.* **2000**, 2000, (8), 1401-10.

9. Peters, R. J.; Ravn, M. M.; Coates, R. M.; Croteau, R. B., Bifunctional abietadiene synthase: free diffusive transfer of the (+)-copalyl diphosphate intermediate between two distinct active sites. *J Am Chem Soc* **2001**, 123, (37), 8974-8.
10. Prusic, S.; Xu, J.; Coates, R. M.; Peters, R. J., Probing the role of the DXDD motif in Class II diterpene cyclases. *Chembiochem* **2007**, 8, (8), 869-874.
11. Wiemer, A. J.; Yu, J. S.; Shull, L. W.; Barney, R. J.; Wasko, B. M.; Lamb, K. M.; Hohl, R. J.; Wiemer, D. F., Pivaloyloxymethyl-modified isoprenoid bisphosphonates display enhanced inhibition of cellular geranylgeranylation. *Bioorg Med Chem* **2008**, 16, (7), 3652-60.
12. Szabo, C. M.; Martin, M. B.; Oldfield, E., An investigation of bone resorption and *Dictyostelium discoideum* growth inhibition by bisphosphonate drugs. *J Med Chem* **2002**, 45, (14), 2894-903.
13. Liu, C. I.; Liu, G. Y.; Song, Y.; Yin, F.; Hensler, M. E.; Jeng, W. Y.; Nizet, V.; Wang, A. H.; Oldfield, E., A cholesterol biosynthesis inhibitor blocks *Staphylococcus aureus* virulence. *Science* **2008**, 319, (5868), 1391-1394.
14. Dehmlow, H.; Aebi, J. D.; Jolidon, S.; Ji, Y. H.; von der Mark, E. M.; Himber, J.; Morand, O. H., Synthesis and structure-activity studies of novel orally active non-terpenoid 2,3-oxidosqualene cyclase inhibitors. *J Med Chem* **2003**, 46, (15), 3354-70.