Table S1. Peak areas for xanthine alkaloids obtained from LC-MS analyses. A combination of retention times and unique MS/MS fragmentation patterns allowed for positive identification of each product. Ionization efficiencies differ between the various xanthine alkaloids thereby limiting direct quantitative comparisons of preferred product formation.

	X[a]	7X	3X	1X	TB	PX	TP
PcAncCS1 + 7X		2220			61		
PcAncCS1 + X	761		57				
PcAncCS1 + 3X			538				
PcAncCS1 + NaOH	27						
PcAncCS2 + X (total protein)	610		56				
PcAncCS2+ 7X (total protein)[b]		422			28	200	
PcAncCS2 + 3X (total protein)	112		483		71		
PcAncCS2 + NaOH (total protein)	160						
PcAncCS2 N314Y+ X	309		164				
PcAncCS2 N314Y+ NaOH							
PcAncCS2 T25S + 3X	111		526		107		
PcAncCS2 T25S+ NaOH	93						
PcCS2 + 3X (total protein)	311				242		
PcCS2 + NaOH (total protein)	286						
PcCS1 + X	554		129				
PcCS1 + NaOH							
CsAncCS + X	1000		25	44			
CsAncCS + 7X[b]		165			47	68	
CsAncCS + NaOH							
TcAncCS1 + 3X			870				88
TcAncCS1 + X	436		41	297			
TcAncCS1 + 7X		2590			119	406	
TcAncCS1 + NaOH							
TcAncCS1 GHRC +3X (total protein)[b]			436	183	85		
TcAncCS1 GHRC +X (total protein)[b]			115	551			
TcAncCS1 GHRC +7X (total protein)[b]		283	48.7	230	109	326	
TcAncCS1 GHRC +NaOH (total protein)[b]			54.3	286			
TcAncCS2 + X	354		686				
TcAncCS2 + 3X			470		146		
TcAncCS2 + NaOH	266						
TcAncCS2AEA + 3X (total protein)	2010	76	869		204		
TcAncCS2AEA + X (total protein)	989	382	219				
TcAncCS2AEA + NaOH (total protein)	1570	80	30				
TcCS1 + X	779		71				
TcCS1 + NaOH							
TcCS2 + X	898	115					
TcCS2 + 3X			498		150		
TcCS2 + NaOH							
pET15B + X (total protein)	274						
pET15B + NaOH (total protein)	193						

[a]Xanthine co-extracts with total cellular protein and it may therefore be detected in samples in which it was not added. In some cases, levels are sufficiently high to detect its methylated products for those enzymes that have high relative activity with it.

[b]Sample not scanned for xanthine fragmentation

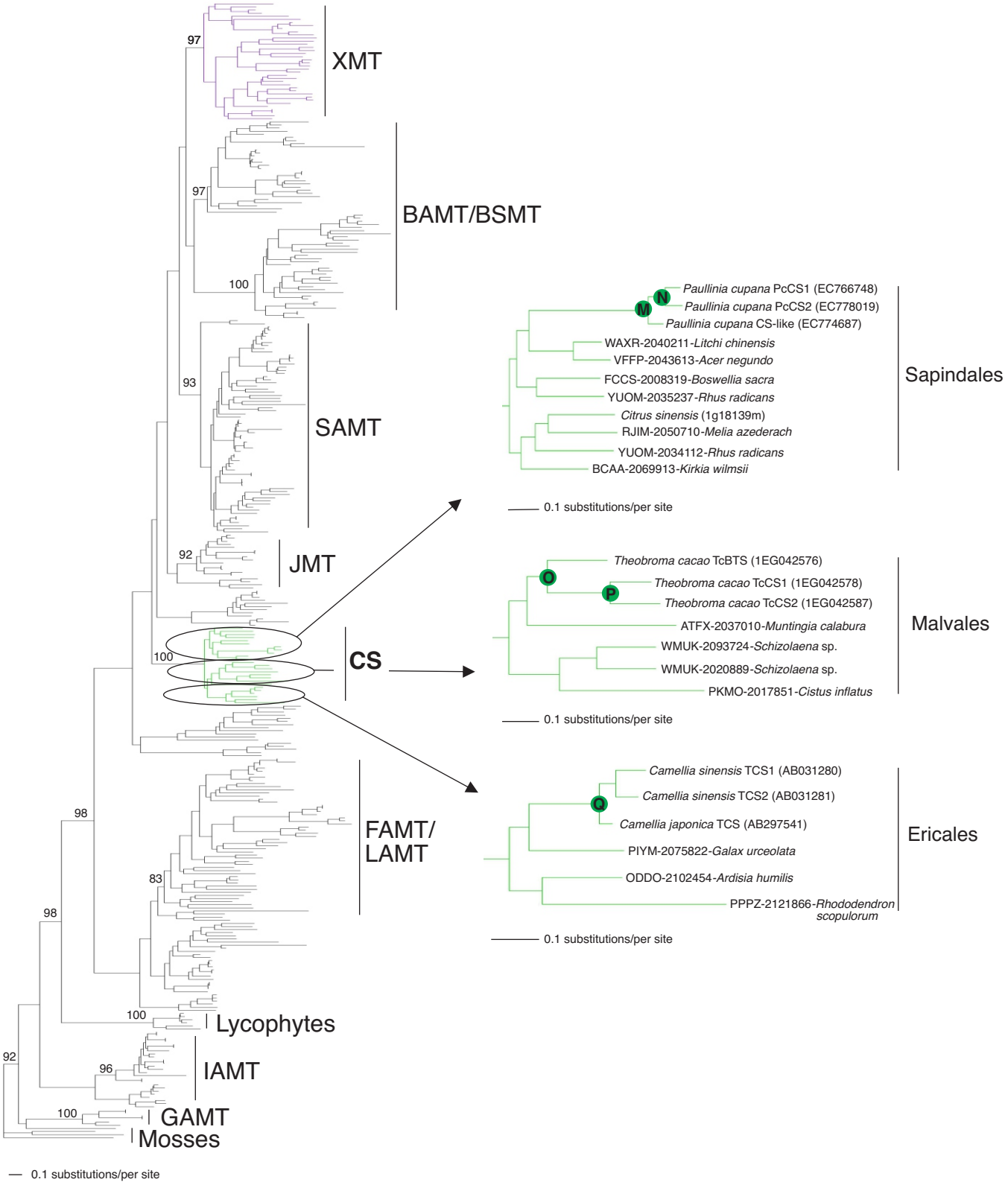


Figure S1. Phylogenetic relationships amongst 366 SABATH protein sequences (lnL = -128066.67176). Sequences were extracted from 11 complete genomes of land plants in addition to selected CS and XMT transcriptome sequences from the OneKP database. Lineages with functionally characterized sequences are labeled by enzyme name. Bootstrap support values are shown for selected nodes which define major enzyme lineages. Enzymes from *Camellia* (CS) or *Coffea* (XMT) known to be involved in caffeine biosynthesis are shown in green and purple, respectively. Sequences expressed in *Paullinia* and *Theobroma* are orthologous to CS sequences from *Camellia*. Arrows point to expanded views of the CS lineages to show recent duplication events within *Paullinia*, *Theobroma* and *Camellia*. Nodes for which ancestral resurrected proteins were studied are labeled M-Q. Accession numbers for OneKP and GenBank databases are shown before and after relevant sequences, respectively.

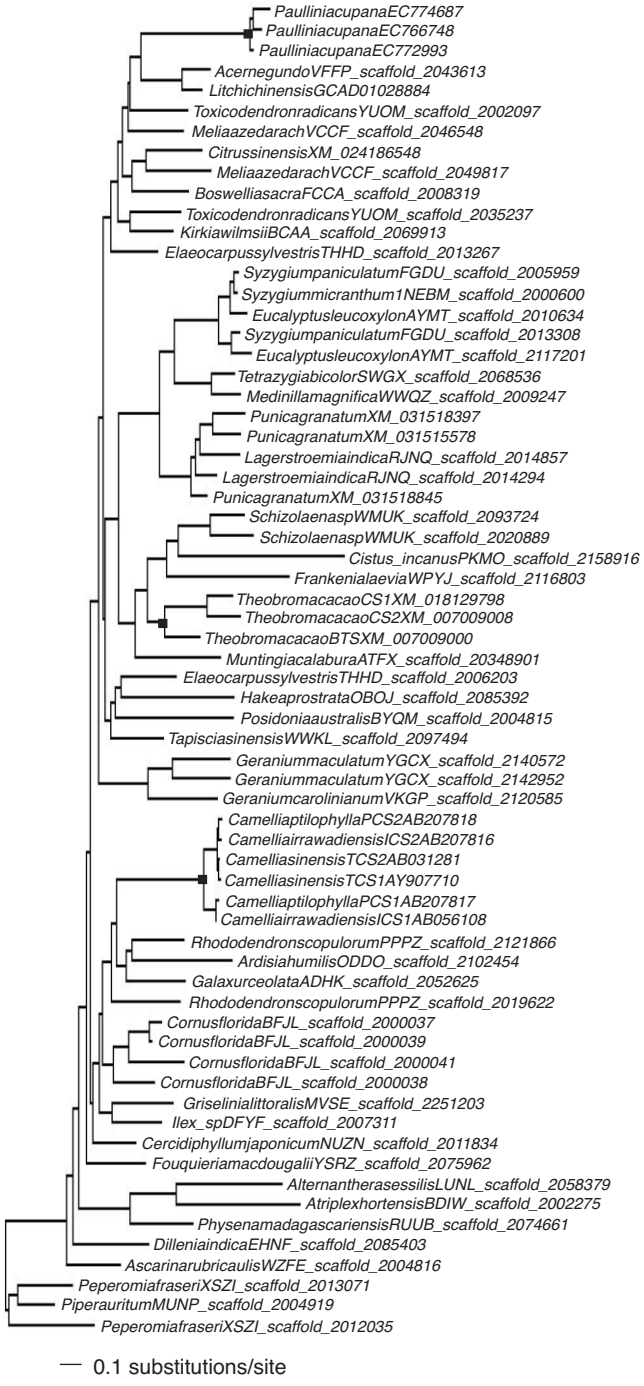
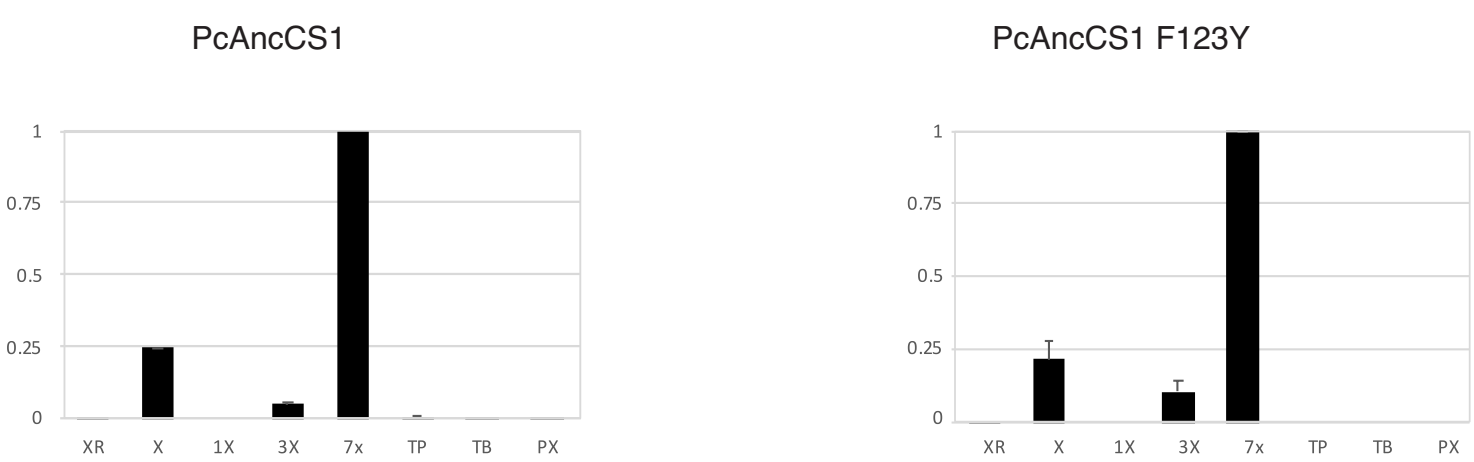


Figure S2. Detailed CS phylogeny (lnL = -19220.31407) used for ASR with accession numbers for OneKP and GenBank sequences shown before or after species names, respectively. Black squares represent branch support > 0.9 to show that the three clades used for ASR were confidently estimated.

A. PcAncCS1 (average site posterior probability = 0.976)

MDVKEVLCMNKGEGESSYLLNSKFTKITAIKSIPTLKRAIESLFKEESPPFEHLLNVADLGCASGSTSNTIMSTIVQTVVNKCRELNHKIPEFQF
YLNDLPSNDFNTLFKGLSGFMGSGGEEFENTSCFVMGAPGSFHGRLFPLNTIHLVYSNYSVHWLSKVPDLRDEKGNPINKGKFYISKTSPSA
VREAYLAQFQKDFTLFLKSRAEEMVSNGRVVLVLHGRLSQDFSCEKELRLPWLILSQAISRLVSKGLIDEEKLDSFEVPYYTPSKQEVKEVVE
REGSYAVELMETFAIEIGDKDEGIWSDARGFVNNLRSFTETMISHHFGPQILDELYDEIHDLLLQDFATQCSIVVGLKRN

B.



C.

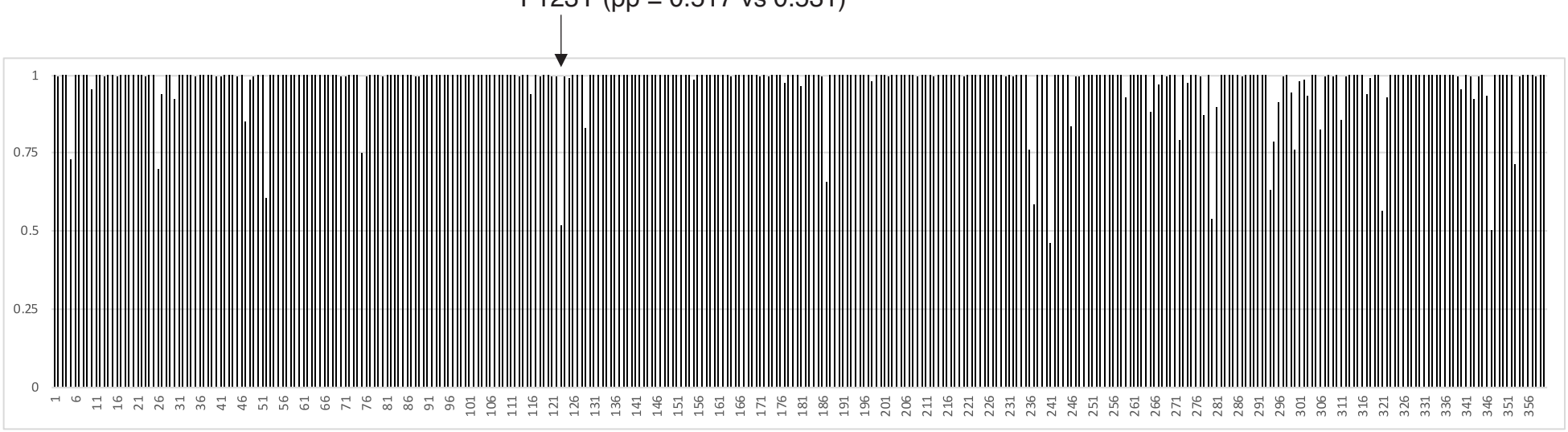
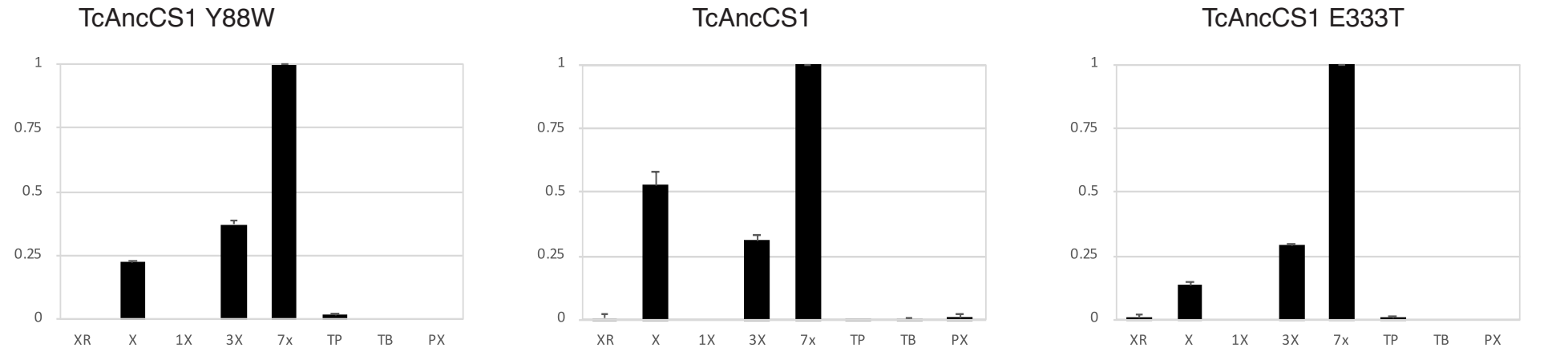


Figure S3. Analyses of two ancestral enzyme variants for PcAncCS1 (A) show similar relative substrate preferences (B). Average site posterior probability is 0.976 for this estimated ancestral sequence (C). The profile shows that F123 had relatively low probability but mutation to the alternative state did not change the enzymatic properties in terms of relative substrate preference.

A. TcAncCS1 (average site posterior probability = 0.935)

MEVKEVLFMNKGDGENSYVKTSGFTQKVAAMTQPVVYRAAQSLFTERNSLSYQVLNVADLGCASGPNTFTVMSTVIESIVDKCSELNY
QMPEIQFYLNDLVGNDFNTLFKGLSVIQEKYKNVSCFAMGAPGSFHGRLFPRNSMHLVHSSYSVHWLSKVPKITNEEGLPLNKGKIYISK
TSPPAVREAYLSQFQEDFSSFLRSRSPELVPDGRMVLILHGRKSADPTTKESCYTWELLAEAISYLVSQGLIDEEKLDSFNVPYYTPSQEE
VRELVDKEGSFTIEFIDTIEMEIGGINIWSSPENRVKNLRSFTEPMISHQFGEEVMDKLYDKVEEILVEDCKQGKESTKTISIVVALKKKES

B.

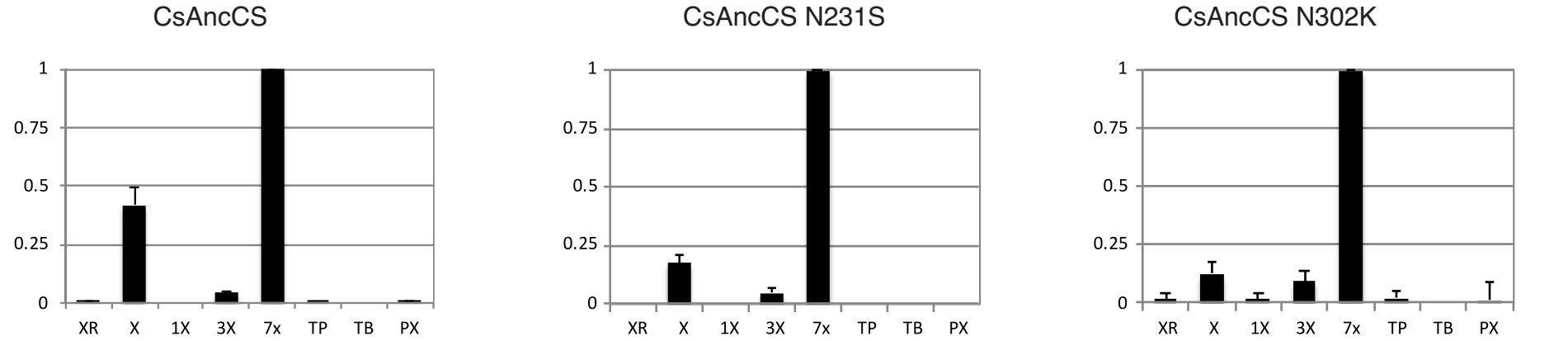


C.

Figure S4. Analyses of three ancestral enzyme variants for TcAncCS1(A) show similar relative substrate preferences (B). Average site posterior probability is 0.935 for this estimated ancestral sequence (C). The profile shows that Y88 and E333 had relatively low probabilities but mutation to alternative states did not change the enzymatic properties in terms of substrate preference.

A. CsAncCS (average site posterior probability = 0.997)
MEEVKEALFMNRGEGESSYAQNSSFTQKVASMTMPVLENAVETLFSKDFHLLQALNAADLGCAAGPNTFTVISTIKRMMEKKCRELNC
QTLELQVYLNDLPGNDFNTLFKGLSSKVVVGNKCEEVSCYVMGVPGSFHGRLFPRNSLHLVHSSYSVHWLSQAPKGLTSREGLALNK
GKIYISKTSPPVVREAYLSQFHEDFTMFLNARSQEVVPNGCMVLILHGRQSSDPSNMESCFTWELLAIAIAELVSQGLIDEDKLDTFNVP
YYTPSLEEVKDIVEREGSFTIDHMEGFELDSPQMQENDKWVRGEKLAKAVRAFTEPIISNQFGHEIMDKLYDKFTHIVVSDLEAKIPKTT
SIILVLSKIVG

B.



C.

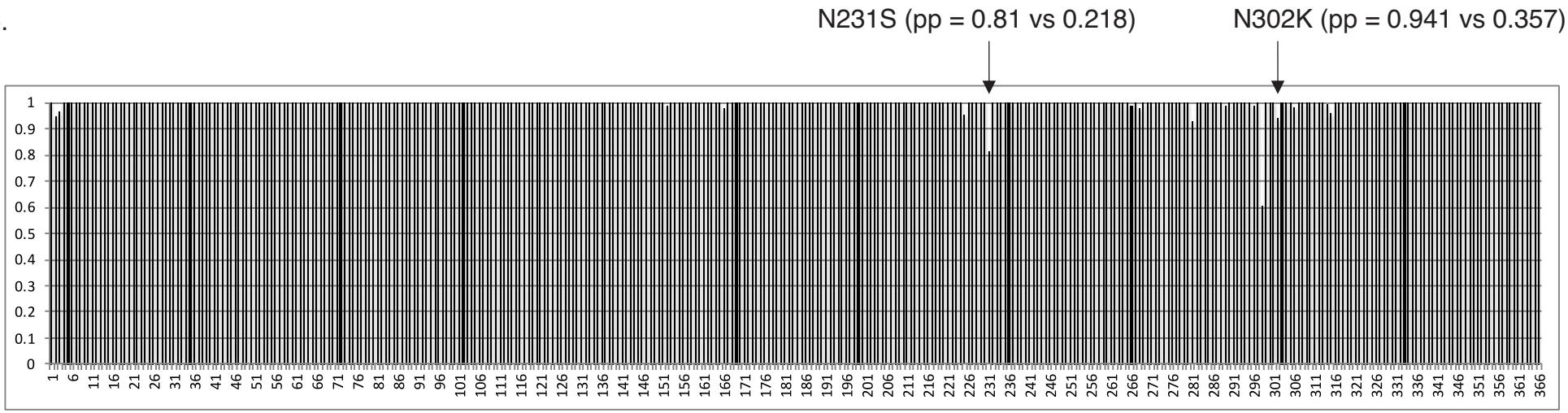
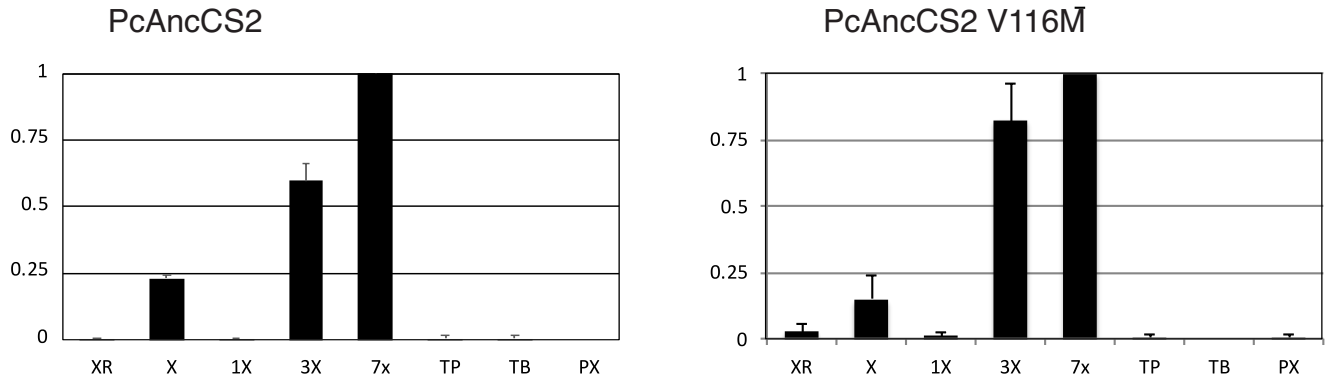


Figure S5. Analyses of three ancestral enzyme variants for CsAncCS (A) show similar relative substrate preferences (B). Average site posterior probability is 0.997 for this estimated ancestral sequence (C). The profile shows that N231 and N302 had relatively low probabilities but mutation to alternative states did not change the enzymatic properties in terms of substrate preference.

A.

PcAncCS2 (average site posterior probability = 0.986)
MDVKDVLCMNKGEGESSYLLNSKFTNITAVKSIPTLKRAIESLFKEESPPFEHLLNVADLGCASGSTSNTIMSTVVQTVVNKCRELN
HKIPEFQFYLNDLPSNDFNTLFKGLSGFVGSGGEEFENTSCLVMGAPGSFHGRLFPLNTIHLVYSNYSVHWLSKVPDLRDEKGNP
INKGTFYISKTSPSAVREAYLAQFQKDFTLFLKSRAEEMVSNGRVVLVLHGRLSQDFScEKELQLPWLILSQAISRLVSKGLIDEEKL
DSFEVPYYTPSAQEVKELVEGEGSYAVELMETFTIRIGARNEGIWSDARGFGNNLRSITETMISHHFGPQILDELYDEIQDLPLQDF
ATQCSFVVGLKRN

B.



C.

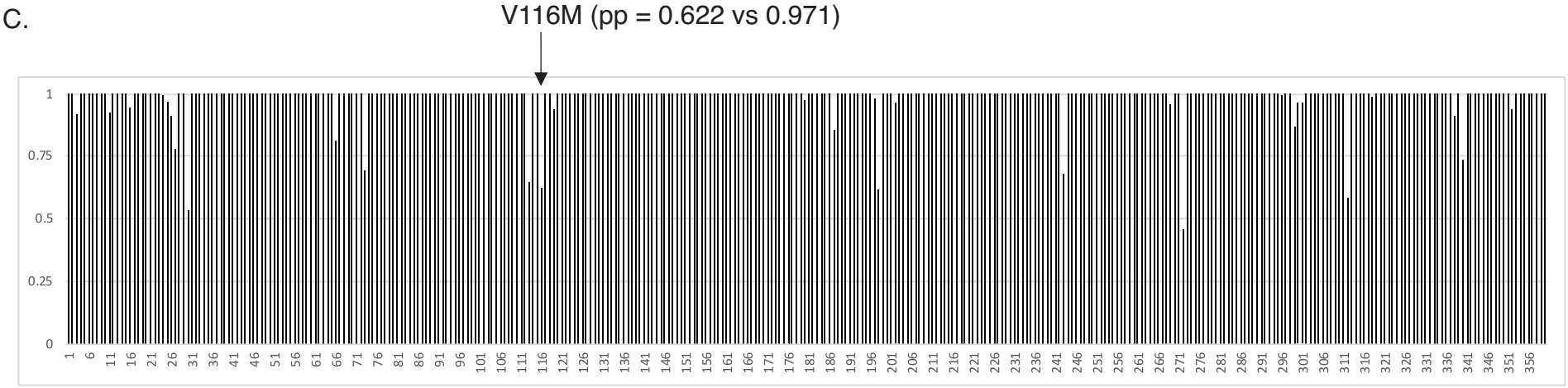
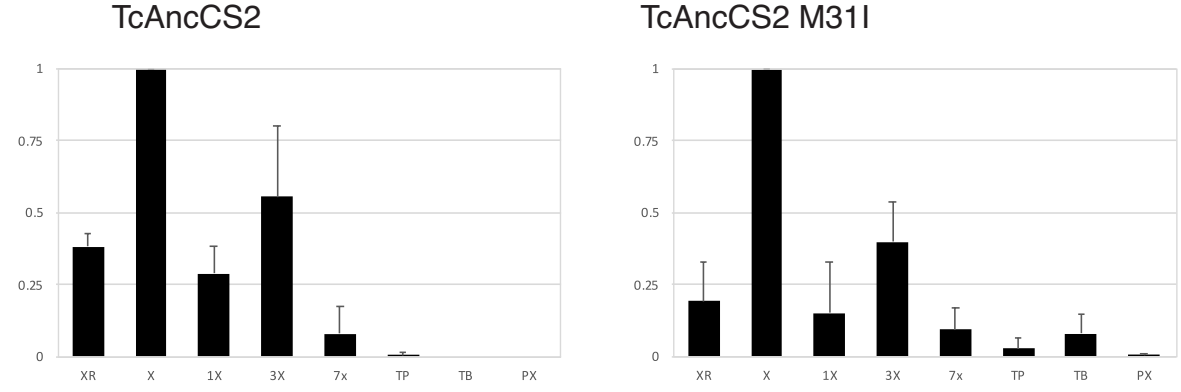


Figure S6. Analyses of two ancestral enzyme variants for PcAncCS2 (A) show similar relative substrate preferences (B). Average site posterior probability is 0.986 for this estimated ancestral sequence (C). The profile shows that V116 has relatively low posterior probability but mutation to an alternative state did not change the enzymatic properties in terms of substrate preference.

A.

TcAncCS2 (average site posterior probability = 0.954)
MEVKDVLFMNKGDGENSYVKSAGLTLKVIAMTQPIVQKAVQSLFTETHSIPLQVVNVADLGCALGPQPLEFMSTVIESIVEKCGELGCE
MPEIQFYLNDLVGNDFNTLFKGLSVVQEKYKNVSWFAMGAPGSFHGRLFPRNSMHLVHSCYSVHWLSKAPKITNEAGLPLNKGKIYMS
KTSPPAVREAYLSQFQEDFSSLLRFRSPELAPDGRMVLILNGRQSADPTNKDTCYTWDLLAEALSYLVSQGLIDEEKLDSFNVPYYNPS
QEEVKRLVDKEGSFTIEFIDTIELEIGGKNIWSSPESRIKGHRCFTEPILSHQFGEEVMDKLYDKAEEILVEDYKQGKEATKNISIVVVLKKK
KN

B.



C.

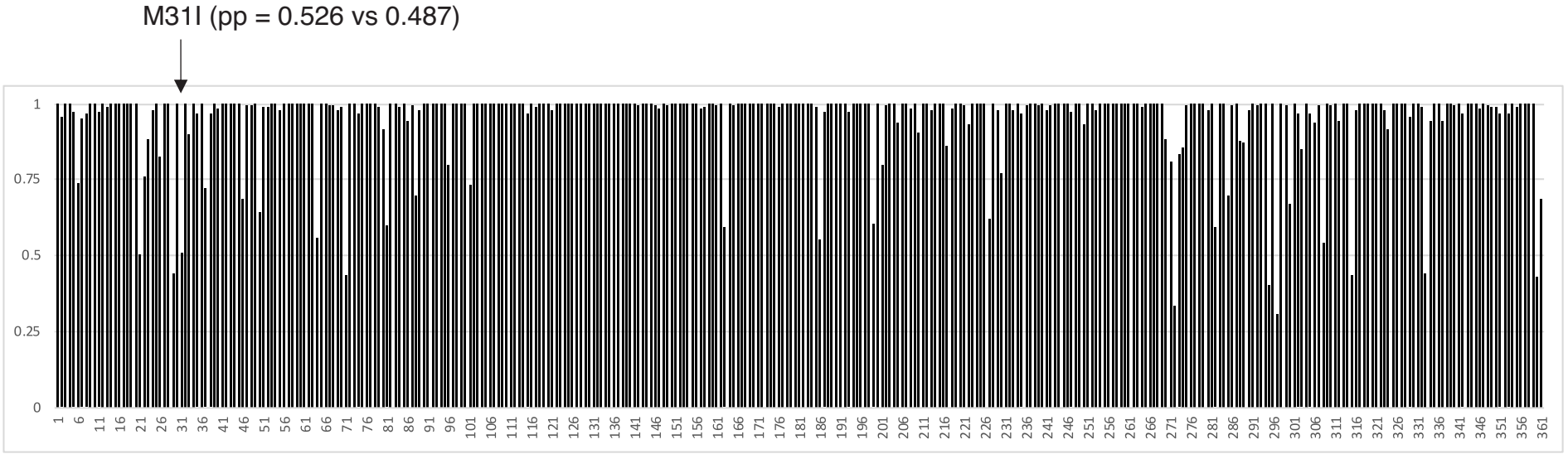


Figure S7. Analyses of two ancestral enzyme variants for TcAncCS2 (A) show similar relative substrate preferences (B). Average site posterior probability is 0.954 for this estimated ancestral sequence (C). The profile shows that M31 has relatively low posterior probability but mutation to an alernative state did not change the enzymatic properties in terms of substrate preference.

```
TcAncCS1  ME------VKEVLFMNKGDGENSYVKTSGFTQKVAAMTQPVVYRAAQSLFTER-NSLSYQ-VLNVADLGC 62
TcAncCS2  ME------VKDVLFMNKGDGENSYVKSAGLTLKVIAMTQPIVQKAVQSLFTET-HSIPLQ-VVNVADLGC 62
TcCS1     MAMK----VKDIVFMNKGDGENSYVKSAGLTLKVIAKTQPIVQKAVQSLFTGT-HSTPLQ-VVNVADLGC 64
TcCS2     MEA-----VKDVLCMNNGVGENSYVKAEALTIKVMAITKPIVPKAVQSLFTETDHSIPLQ-VVNVADLGC 64
PcAncCS1  MD------VKEVLCMNKGEGESSYLLNSKFTKITAIKSIPTLKRAIESLFKEE-SP-PFEHLLNVADLGC 62
PcAncCS2  MD------VKDVLCMNKGEGESSYLLNSKFTNITAVKSIPTLKRAIESLFKEE-SP-PFEHLLNVADLGC 62
PcCS1     MD------LKDVLCMNTGEGESSYLLNSKFTNITAIKSIPTLKRAIESLFKEE-SP-PFEHLLNVADLGC 62
PcCS2     MD------VKDVLCMNKGEGEGSYLLNSKVSIITAVKSIPTLKRAIESLFKEE-SP-PFEHLLNVADLGC 62
CsAncCS   MELATAGEVKEALFMNRGEGESSYAQNSSFTQKVASMTMPVLENAVETLFSKD-FH-LLQ-ALNAADLGC 67
TCS1      MELATAGKVNEVLFMNRGEGESSYAQNSSFTQQVASMAQPALENAVETLFSRD-FH--LQ-ALNAADLGC 66
TCS2      MKE-----VKEALFMNKGEGESSYAQNSSFTQTVTSMTMPVLENAVETLFSKD-FH-LLQ-ALNAVDLGC 62

TcAncCS1  ASGPNTFTVMSTVIESIVDKCSELNY-QMPEIQFYLNDLVGNDFNTLFKGLS-----VIQEKYKN----V 122
TcAncCS2  ALGPQPLEFMSTVIESIVEKCGELGC-EMPEIQFYLNDLVGNDFNTLFKGLS-----VVQEKYKN----V 122
TcCS1     ALGPQPLESMSIVIESIVEKCGELGC-EMPEIQFHLNDLAGNDFNTLFKGLS-----VVQEKYKN----V 124
TcCS2     AVGPQPLEFMSTVIESILKKCGEMGR-EMPEIQFFLNDLVGNDFNTLFKGLS-----VVQEKYKK----V 124
PcAncCS1  ASGSTSNTIMSTIVQTVVNKCRELNH-KIPEFQFYLNDLPSNDFNTLFKGLSGFMG-SGGEEFEN----T 126
PcAncCS2  ASGSTSNTIMSTVVQTVVNKCRELNH-KIPEFQFYLNDLPSNDFNTLFKGLSGFVG-SGGEEFEN----T 126
PcCS1     ASGSTSNTIMPTIVQTVVNRCRELNH-KIPEFQFYLNDLPSNDFNTLFKGLNGLVG-SGGEEFEN----T 126
PcCS2     ASGLTSNTIMSTVVQTVVNKCRELNH-KIPEFQFYLNDLPSNDFNTLFKGLSGFMG-S-GEEFEN----T 125
CsAncCS   AAGPNTFTVISTIKRMMEKKCRELNCAQTLELQVYLNDLPGNDFNTLFKGLSSKVDQVVGNKCEEGGGRV 137
TCS1      AAGPNTFAVISTIKRMMEKKCRELNC-QTLELQVYLNDLFGNDFNTLFKGLSSE---VIGNKCEE----V 128
TCS2      AAGPTTFTVISTIKRMMEKKCRELNC-QTLELQVYLNDLPGNDFNTLFKGLPSK---VVGNKCEE----V 124

TcAncCS1  SCFAMGAPGSFHGRLFPRNSMHLVHSSYSVHWLSKVPK-ITNEEGLPLNKGKIYISKTSPPAVREAYLSQ 191
TcAncCS2  SWFAMGAPGSFHGRLFPRNSMHLVHSCYSVHWLSKAPK-ITNEAGLPLNKGKIYMSKTSPPAVREAYLSQ 191
TcCS1     SWFAMGAPGSFHGRLFPRNSMHLVHSCYSVHWLSKAPK-ITSEAGLPLNKGKIYMSKTSPPAVREGYLSQ 193
TcCS2     SWFAMGAPGSFHGRLFPRNSMHLVYSCYSVHWLSEAPK-ITNEAGLPLNKGKIYMSKTSPPAVTKAYLSQ 193
PcAncCS1  SCFVMGAPGSFHGRLFPLNTIHLVYSNYSVHWLSKVPD-LRDEKGNPINKGKFYISKTSPSAVREAYLAQ 195
PcAncCS2  SCLVMGAPGSFHGRLFPLNTIHLVYSNYSVHWLSKVPD-LRDEKGNPINKGTFYISKTSPSAVREAYLAQ 195
PcCS1     SCLVMGAPGSFHGRLFPLNTIHLVYSNYSVHWLSKVPD-LKDEKGNPINKGTFYISKTSPSGVREAYLAQ 195
PcCS2     SCLVMGAPGSFHGRLFPLNTIHLVYSNYSVHWLSKVPD-LRDEKGNPINKGTFYLSKTSPSAVREAYLAQ 194
CsAncCS   SCYVMGVPGSFHGRLFPRNSLHLVHSCYSVHWLTQAPKGLTSREGLALNKGKIYISKTSPPVVREAYLSQ 207
TCS1      PCYVMGVPGSFHGRLFPRNSLHLVHSSYSVHWLTQAPKGLTSREGLALNKGKIYISKTSPPVVREAYLSQ 198
TCS2      SCYVVGVPGSFHGRLFPRNSLHLVHSCYSVHWLTQAPKGLTSKEGLALNKGKIYISKTSPPVVREAYLSQ 194

TcAncCS1  FQEDFSSFLRSRSPELVPDGRMVLILHGRKSADPT-TKESCYTWELLAEAISYLVSQGLIDEEKLDSFNV 260
TcAncCS2  FQEDFSSLLRFRSPELAPDGRMVLILNGRQSADPT-NKDTCYTWDLLAEALSYLVSQGLIDEEKLDSFNV 260
TcCS1     FEEDFSSVLRFRSPELAPDGRMVLILNGRQSADPT-EKDICYLWDLLAEALSYLVSEGLIDEEKLDSFNV 262
TcCS2     FQEDFSSLLKFRSQELAPNGRVVLIFNGRQTADPT-NKDTCYTWDLLAEALSYLVSQGLVDEGKLDSFNV 262
PcAncCS1  FQKDFTLFLKSRAEEMVSNGRVVLVLHGRLSQDFSCEKELRLPWLILSQAISRLVSKGLIDEEKLDSFEV 265
PcAncCS2  FQKDFTLFLKSRAEEMVSNGRVVLVLHGRLSQDFSCEKELQLPWLILSQAISRLVSKGLIDEEKLDSFEV 265
PcCS1     FQKDFTLFLKSRAEEMVSNGRVVLVLHGRLSQDFSCEKELQLPWLILSKAISRLVSKGLIDEEKLDSFEV 265
PcCS2     FRQDFTFFLKSRAEEMVSNGRVVLVLHGRLSQDFSCEKELQLPWLILTQAISRLVSKGLIDEEKLDSFEV 264
CsAncCS   FHEDFTMFLNARSQEVVPNGCMVLILHGRQSSDPS-DMESCFTWELLAIAIAELVSQGLIDEDKLDTFNV 276
TCS1      FHEDFTMFLNARSQEVVPNGCMVLILRGRQCSDPS-DMQSCFTWELLAMAIAELVSQGLIDEDKLDTFNI 267
TCS2      FHEDFTMFLNSRSQEVVPNGCMVLILRGRLSSDPS-DMGSCFTWELLAVAIAELVSQGLIDEDKLDTFNV 263

TcAncCS1  PYYTPSQEEVRELVDKEGSFTIEFIDTIEMEIG----GIN-IWSSPENRVKNLRSFTEPMISHQFGEEVM 325
TcAncCS2  PYYNPSQEEVKRLVDKEGSFTIEFIDTIELEIG----GKN-IWSSPESRIKGHRCFTEPILSHQFGEEVM 325
TcCS1     PYYNPSQEEVERVIDKEGSFTTEFSDTVVLEIG----GKN-AWSDPGLRIKGYRCFSEPILSHQFGEEVM 327
TcCS2     PYYNPSQEEIKYLVDKEGSLTIEFIDTIELEIG----GPNGYWSSPESRIRGHRCFTEPLLSHQFGERLM 328
PcAncCS1  PYYTPSKQEVKEVVEREGSYAVELMETFAIEIG--DKDEG-IWSDARGFVNNLRSFTETMISHHFGPQIL 332
PcAncCS2  PYYTPSAQEVKELVEGEGSYAVELMETFTIRIG--ARNEG-IWSDARGFGNNLRSITETMISHHFGPQIL 332
PcCS1     PYYTPSEQEVKELVEGEGSYAVELMETFTLKDG--ARNEG-IWSDARGFGNYLRSFTETMISHHFGPQIL 332
PcCS2     PYYAPSAQEVKELVEGEGSYAVELMETFTIRIG--VGDEG-IWSDARGFVNNLRSITETMISHHFGPQIL 331
CsAncCS   PSYFPSLEEVKDIVERDGSFTIDHMEGFELDSPLQMQEND-KWVRGEKFAKVVRAFTEPIISNQFGHEIM 345
TCS1      PSYFASLEEVKDIVERDGSFTIDHIEGFDLDSV-EMQEND-KWVRGEKFTKVVRAFTEPIISNQFGPEIM 335
TCS2      PSYFPSLEEVKDIVERNGSFTIDHMEGFELDSP-EMQEND-KWVRGEKFATVARAFTEPIISNQFGHEIM 331

TcAncCS1  DKLYDKVEEILVEDCKQGKESTKTI--SIVVALKKKES---- 361
TcAncCS2  DKLYDKAEEILVEDYKQGKEATKNI--SIVVVLKKKKN---- 361
TcCS1     DKLFDKAEEILAEDYKQGKEATKNI--SIVVVLKKKTNQTWT 367
TcCS2     DKLYDKATQILVEDYKHGKEATKNI--GIAVVLKKKKL---- 364
PcAncCS1  DELYDEIHDLLLQDFA--TQC------SIVVGLKRN------ 360
PcAncCS2  DELYDEIQDLPLQDFA--TQC------SFVVGLKRN------ 360
PcCS1     DELYDEIHNLPLQDFA--TQC------SFVVGLKRN------ 360
PcCS2     DELYDGILDLPLQDFA--TQC------NFVVGLKRN------ 359
CsAncCS   DKLYDKFTHIVVSDLE--AKIPKTTYVSIILVLSKIVG---- 381
TCS1      DKLYDKFTHIVVSDLE--AKLPKTT--SIILVLSKIDG---- 369
TCS2      DKLYEKFTHIVVSDFE--AKIPKIT--SIILVLSKIVG---- 365
```

Figure S8. Alignments for resurrected ancestral proteins with their modern-day descendant enzymes from each of the clades shown in Fig. 2C-E.