

SCHEMA recombination uncovers stabilizing cellulase mutation

SUPPLEMENTAL TABLE 1

Two independent duplicate  $T_{50}$  values ( $^{\circ}\text{C}$ ) for parent CBH IIs, 23 original sample set CBH II chimeras and predicted thermostable CBH II chimeras. The 18 chimeras synthesized for this work are preceded by an asterisk.

Sample Set Chimeras & Parents				Predicted Thermostable Chimeras			
Sequence	$T_{50}(1)$	$T_{50}(2)$	Mean $T_{50}$	Sequence	$T_{50}(1)$	$T_{50}(2)$	Mean $T_{50}$
32333113	52	51	51.5	12332331	66.5	67	66.8
13111313	56	53.5	54.8	*13112332	67	67	67
11313121	55	55.5	55.3	22311331	68	68	68
21131311	57.5	57	57.3	*12111332	68	68	68
31212111	59	58	58.5	*12112332	68.5	67.5	68
<b>Parent 2</b>	<b>60</b>	<b>58</b>	<b>59</b>	12131331	68.5	69	68.8
23233133	61	61	61	*12131332	70	67.5	68.8
31311112	60	62	61	*12332332	69	69	69
22212231	63	61	62	12111131	70	68.5	69.3
13231111	63	6.5	63.3	12311332	70	69	69.5
12213111	63	63.5	63.3	13332331	70	69	69.5
<b>Parent 3</b>	<b>63.5</b>	<b>64.5</b>	<b>64</b>	12132331	70.5	69	69.8
12133333	64	64	64	*12132332	70.5	69	69.8
<b>Parent 1</b>	<b>64</b>	<b>65.5</b>	<b>64.8</b>	*13332332	69.5	70	69.8
33133132	65	66	65	12112132	71	68.5	69.8
11332333	64.5	66	65.3	13322332	71	68.5	69.8
23311333	65	66	65.5	*13131332	70	70	70
33213332	66	66	66	*12331332	71	69	70
13333232	67.5	67	67.3	*13312332	70	70	70
22232132	68	68	68	*11113332	69.5	70.5	70
11113132	71.5	71	71.3	*13113132	70.5	69.5	70
21333331	73.5	75.5	74.5	*11112132	70.5	70	70.3
21311131	75.5	75.5	75	*12113132	70.5	70.5	70.5
				*13132332	69.5	71.5	70.5
				*11111132	71	70.5	70.8
				13331332	72	70	71
				*13111132	72	69.5	71.3
				*12222132	72.5	70	71.3
				12222332	72	69.5	71.3
				13311332	71	71.5	71.7
				13311331	73.5	72.5	73

SCHEMA recombination uncovers stabilizing cellulase mutation

SUPPLEMENTAL TABLE 2

Amino acid sequences for CBH II parent and chimera catalytic domains shown in Supplemental Table 1. Table also includes catalytic domain for *P. chrysosporium* CBH II as described in text. All recombinant CBH IIs share the N-terminal CBM and linker from the native *H. jecorina* CBH II, CSSVWGQCGGNWSGPTCCASGSTCVYSNDYYSQCLPGAASSSSSTRAASTTSRVSPITTSRSSSA TPPPSTTTRVPPVGSATATYS.

Parent 1 (*H. insolens*)

GNPFEGVQLWANNYYRSEVHTLAIPQITDPALRAAASAVA E VPSFQWLDRNVTVDTL LVQTLSE  
IREANQAGANPQYAAQIVVYDLPDRDCAAAA SNGEWAIANNGVN NYKAYINRIREILISFSDVRT  
ILVIEPDSL ANMVTNMNV PKCSGAASTYRELT IYALKQLD LPHVAMYMDAGHAGWLGWPANI Q  
PAAELFAKIYEDAGK PRAVRGLATNV ANYNAWSVSSPPPYTSPNPNYDEKH YIEAFRPLLEARGF  
P\_AQFIVDQGRSGKQPTGQKEWGHWCNAIGTGF GMRPTANTGHQYVDAFVWVKPGGEC DGTS  
DTTAARYDYHCGLEDALKPAPEAGQWFNEYFIQLLRNANPPF\_

Parent 2 (*H. jecorina*)

GNPFVGVTPWANAYYASEVSSLAIPSLT\_GAMATAAAA VAKVPSFMWLD T\_LDKTPLMEQTLA  
DIRTANKNGGN\_YAGQFVVYDLPDRDCAALASNGEYSIADGGVAKYKNYIDTIRQIVVEYS DIR  
TLLVIEPDSL ANLVTNLGTPKCANAQSAYLE CINYAVTQLNLPNVAMYLDAGHAGWLGWPAN Q  
DPAAQLFANVYKNASSPRALRGLATNV ANYNGWNITSPPSY TQGNVYNEKLYIHAIGPLL ANH  
GWSNAFFITDQGRSGKQPTGQQQWGDWCNVIGTGF GIRPSANTGDSL LDFSFWVKPGGEC DG T  
SDSSAPRFDSHCALPDALQPAPQAGAWFQAYFVQLLTNANPSFL

Parent 3 (*C. thermophilum*)

GNPFSGVQLWANTYYSSEVHTLAIPSL S\_PELAAKAAKVAEVPSFQWLDRNVTVDTLFSGTLAEI  
RAANQRGANPPYAGIFVVYDLPDRDCAAAA SNGEWSIANNGANNYKRYIDRIRELLIQYSDIRTI  
LVIEPDSL ANMVTNMNVQKCSNAASTYKELTVYALKQLNLP HVAMYMDAGHAGWLGWPANI  
QPAAELFAQIYRDAGRPAAVRGLATNV ANYNAWSIASPPSYTSPNPNYDEKH YIEAFAPLLRNQ  
GFD\_AKFIVDTGRNGKQPTGQLEWGHWCNVKGTGF GVRPTANTGHELVD AFVWVKPGGESD G  
TSDTSAARYDYHCGLSDALTPAPEAGQWFQAYFEQLLINANPPF\_

*P. chrysosporium* CBH II

NNPWTGFQIFLSPYYANEVAAAA\_KQITDPTLSSKAASVAN IPTFTWLDSVAKIPDLGTYLASASA  
LGKSTGT\_KQLVQIVYDLPDRDCAAKASNGEFSIANNGQANYENYIDQIVAQIQQFPDVRVVA  
VIEPDSL ANLVTNLNVQKCANAKTTYLACVNYALTNLAKVGVYMYMDAGHAGWLGWPANLS  
PAAQLFTQVWQNAGKSPFIKGLATNV ANYNALQAASPD PITQGNPNYDEIHYINALAPLLQQAG  
WD\_ATFIVDQGRSGVQNI RQ\_QWGDWCNIKGAGFGTRPTTNTGSQFIDSIVWVKPGGEC DGTSN  
SSSPRYDSTCSLPDAAQPAPEAGTWFQAYFQTLVSAANPPL\_

32333113

GNPFSGVQLWANTYYSSEVHTLAIPSL S\_PELAAKAAKVAEVPSFMWLD T\_LDKTPLMEQTLADI  
RTANKNGANPPYAGIFVVYDLPDRDCAAAA SNGEWSIANNGANNYKRYIDRIRELLIQYSDIRTI  
LVIEPDSL ANMVTNMNVQKCSNAASTYKELTVYALKQLNLP HVAMYMDAGHAGWLGWPANI  
QPAAELFAQIYRDAGRPAAVRGLATNV ANYNAWSVSSPPPYTSPNPNYDEKH YIEAFRPLLEAR  
GFP\_AQFIVDQGRSGKQPTGQKEWGHWCNAIGTGF GMRPTANTGHQYVDAFVWVKPGGEC DG  
TSDTSAARYDYHCGLSDALTPAPEAGQWFQAYFEQLLINANPPF\_

13111313

GNPFEGVQLWANNYYRSEVHTLAIPQITDPALRAAASAVA E VPSFQWLDRNVTVDTLFSGTLAE

SCHEMA recombination uncovers stabilizing cellulase mutation

IRAAHQGANPQYAAQIVVYDLPDRDCAAAAASNGEWAIANNGVNKYKAYINRIREILISFSDVRT  
ILVIEPDSLANTMNMNPKCSGAASTYRELTIIYALKQLDLPHVAMYMDAGHAGWLGWPANIQ  
PAAELFAKIYEDAGKPRAVRGLATNVANYNAWSIASPPSYTSPNPNYDEKHIEAFAPLLRNQGF  
D\_AQFIVDQGRSGKQPTGQKEWGHWCNAIGTGFGMRPTANTGHQYVDAFVWVKPGGEDGTS  
DTSAARYDYHCGLSDALTPAPEAGQWFQAYFEQLLINANPPF\_

11313121

GNPFEGVQLWANNYYRSEVHTLAIPQITDPALRAAASAVAEPVSFQWLDRNVTVDTLVQTLSE  
IREANQAGANPPYAGIFVVYDLPDRDCAAAAASNGEWSIANNGANNYKRYIDRIREILISFSDVRTI  
LVIEPDSLANTMNMNPKCSGAASTYKELTVYALKQLNLPVAMYMDAGHAGWLGWPANI  
QPAAELFAQIYRDAGRPAAVRGLATNVANYNAWSVSSPPPYTSPNPNYDEKHIEAFRPLLEAR  
GFPNAFFITDQGRSGKQPTGQQWGDWCNVIGTGFGIRPSANTGDSLLDSFVWVKPGGEDGTS  
DTTAARYDYHCGLEDALKPAPEAGQWFNEYFIQLLRNANPPF\_

21131311

GNPFVGVTPWANAYYASEVSSLAIPSLT\_GAMATAAAAVAKVPSFQWLDRNVTVDTLVQTLSEI  
EIREANQAGANPQYAAQIVVYDLPDRDCAAAAASNGEWAIANNGVNKYKAYINRIRELLIQYSDI  
RTILVIEPDSLANTMNMNPKCSNAASTYRELTIIYALKQLDLPHVAMYMDAGHAGWLGWPA  
NIQPAAELFAKIYEDAGKPRAVRGLATNVANYNAWSIASPPSYTSPNPNYDEKHIEAFAPLLRN  
QGFD\_AQFIVDQGRSGKQPTGQKEWGHWCNAIGTGFGMRPTANTGHQYVDAFVWVKPGGEDC  
GTSDTTAARYDYHCGLEDALKPAPEAGQWFNEYFIQLLRNANPPF\_

31212111

GNPFSGVQLWANTYYSSEVHTLAIPSLT\_PELAAKAAKVAEPVSFQWLDRNVTVDTLVQTLSEI  
REANQAGGN\_YAGQFVVYDLPDRDCAALASNGEYSIADGGVAKYKNYIDRIREILISFSDVRTIL  
VIEPDSLANTMNMNPKCSGAASAYLECINYAVTQLNLPVAMYLDAGHAGWLGWPANQDP  
AAQLFANVYKNASSPRALRGLATNVANYNAWSVSSPPPYTSPNPNYDEKHIEAFRPLLEARGF  
P\_AQFIVDQGRSGKQPTGQKEWGHWCNAIGTGFGMRPTANTGHQYVDAFVWVKPGGEDGTS  
DTTAARYDYHCGLEDALKPAPEAGQWFNEYFIQLLRNANPPF\_

23233133

GNPFVGVTPWANAYYASEVSSLAIPSLT\_GAMATAAAAVAKVPSFQWLDRNVTVDTLFSGTLA  
EIRAAHQRRGN\_YAGQFVVYDLPDRDCAALASNGEYSIADGGVAKYKNYIDRIRELLIQYSDIR  
TILVIEPDSLANTMNMNPKCSNAASTYKELTVYALKQLNLPVAMYMDAGHAGWLGWPA  
NIQPAAELFAQIYRDAGRPAAVRGLATNVANYNAWSVSSPPPYTSPNPNYDEKHIEAFRPLLEA  
RGFP\_AKFIVDTGRNGKQPTGQLEWGHWCNVKGTGFGVRPTANTGHELVDVAFVWVKPGGESD  
GTSDTSAARYDYHCGLSDALTPAPEAGQWFQAYFEQLLINANPPF

31311112

GNPFSGVQLWANTYYSSEVHTLAIPSLT\_PELAAKAAKVAEPVSFQWLDRNVTVDTLVQTLSEI  
REANQAGANPPYAGIFVVYDLPDRDCAAAAASNGEWSIANNGANNYKRYIDRIREILISFSDVRTIL  
VIEPDSLANTMNMNPKCSGAASTYRELTIIYALKQLDLPHVAMYMDAGHAGWLGWPANIQP  
AAELFAKIYEDAGKPRAVRGLATNVANYNAWSVSSPPPYTSPNPNYDEKHIEAFRPLLEARGFP  
\_AQFIVDQGRSGKQPTGQKEWGHWCNAIGTGFGMRPTANTGHQYVDAFVWVKPGGEDGTS  
SSAPRFDSHCALPDALQPAPQAGAWFQAYFVQLLTNANPSFL

22212231

GNPFVGVTPWANAYYASEVSSLAIPSLT\_GAMATAAAAVAKVPSFMWLDL\_LDKTPLMEQTLA  
DIRTANKNGGN\_YAGQFVVYDLPDRDCAALASNGEYSIADGGVAKYKNYIDRIREILISFSDVR  
TILVIEPDSLANTMNMNPKCSGAASAYLECINYAVTQLNLPVAMYLDAGHAGWLGWPAN

SCHEMA recombination uncovers stabilizing cellulase mutation

QDPAAQLFANVYKNASSPRALRGLATNVANYNGWNITSPPSYTGQNAVYNEKLYIHAIGPLLAN  
HGWS\_AKFIVDTGRNGKQPTGQLEWGHWCNVKGTGFGVRPTANTGHELVDADFVWVKPGGESD  
GTSDDTAARYDYHCGLEDALKPAPEAGQWFNEYFIQLLRNANPPF\_

13231111

GPNFEGVQLWANNYYRSEVHTLAIPQITDPALRAAASAVAEPVPSFQWLDRNVTVDTLFSGTLAE  
IRAANQRGGN\_YAGQFVVYDLPDRDCAALASNGEYSIADGGVAKYKNYIDRIRELLIQYSDIRTI  
LVIEPDSLANMVTNMNVQKCSNAASTYRELTIVYALKQLDLPHVAMYMDAGHAGWLGWPANI  
PAAELFAKIYEDAGKPRAVRGLATNVANYNAWSVSSPPPYTSPNPNYDEKHIEAFRPLLEARGF  
P\_AQFIVDQGRSGKQPTGQKEWGHWCNAIGTGFGMRPTANTGHQYVDAFVWVKPGGECDGTS  
DTTAARYDYHCGLEDALKPAPEAGQWFNEYFIQLLRNANPPF\_

12213111

GPNFEGVQLWANNYYRSEVHTLAIPQITDPALRAAASAVAEPVPSFMWLDL\_LDKTPLMEQTLAD  
IRTANKNGGN\_YAGQFVVYDLPDRDCAALASNGEYSIADGGVAKYKNYIDRIREILISFSDVRTI  
LVIEPDSLANMVTNMNVKPCSGAASTYKELTVYALKQLNLPVAMYMDAGHAGWLGWPANI  
QPAAELFAQIYRDAGRPAAVRGLATNVANYNAWSVSSPPPYTSPNPNYDEKHIEAFRPLLEAR  
GFP\_AQFIVDQGRSGKQPTGQKEWGHWCNAIGTGFGMRPTANTGHQYVDAFVWVKPGGECDG  
TSDTTAARYDYHCGLEDALKPAPEAGQWFNEYFIQLLRNANPPF\_

12133333

GPNFEGVQLWANNYYRSEVHTLAIPQITDPALRAAASAVAEPVPSFMWLDL\_LDKTPLMEQTLAD  
IRTANKGANPQYAAQIVVYDLPDRDCAAAAASNGEWAIANNGVNNYKAYINRIRELLIQYSDIR  
TILVIEPDSLANMVTNMNVQKCSNAASTYKELTVYALKQLNLPVAMYMDAGHAGWLGWPA  
NIQPAAELFAQIYRDAGRPAAVRGLATNVANYNAWSIASPPSYTSPNPNYDEKHIEAFAPLLRN  
QGFD\_AKFIVDTGRNGKQPTGQLEWGHWCNVKGTGFGVRPTANTGHELVDADFVWVKPGGESD  
GTSDDTAARYDYHCGLSDALTPAPEAGQWFQAYFEQLLINANPPF\_

33133132

GPNFSGVQLWANTYYSSEVHTLAIPSLP\_PELAAKAAKVAEPVPSFQWLDRNVTVDTLFSGTLAEI  
RAANQRGANPQYAAQIVVYDLPDRDCAAAAASNGEWAIANNGVNNYKAYINRIRELLIQYSDIR  
ILVIEPDSLANMVTNMNVQKCSNAASTYKELTVYALKQLNLPVAMYMDAGHAGWLGWPANI  
QPAAELFAQIYRDAGRPAAVRGLATNVANYNAWSVSSPPPYTSPNPNYDEKHIEAFRPLLEAR  
GFP\_AKFIVDTGRNGKQPTGQLEWGHWCNVKGTGFGVRPTANTGHELVDADFVWVKPGGESD  
TSDSSAPRFDSHCALPDALQPAPQAGAWFQAYFVQLLTNANPSFL

11332333

GPNFEGVQLWANNYYRSEVHTLAIPQITDPALRAAASAVAEPVPSFQWLDRNVTVDTLVQTLSE  
IREANQAGANPPYAGIFVVYDLPDRDCAAAAASNGEWSIANNGANNYKRYIDRIRELLIQYSDIRTI  
LVIEPDSLANMVTNMNVQKCSNAASAYLECINYAVTQLNLPNVAMYLDAGHAGWLGWPANQ  
DPAAQLFANVYKNASSPRALRGLATNVANYNAWSIASPPSYTSPNPNYDEKHIEAFAPLLRNQ  
GFD\_AKFIVDTGRNGKQPTGQLEWGHWCNVKGTGFGVRPTANTGHELVDADFVWVKPGGESD  
TSDTSAARYDYHCGLSDALTPAPEAGQWFQAYFEQLLINANPPF\_

23311333

GPNFVGVTPWANAYYASEVSSLAIPSLT\_GAMATAAAAVAKVPSFQWLDRNVTVDTLFSGTLA  
EIRAANQRGANPPYAGIFVVYDLPDRDCAAAAASNGEWSIANNGANNYKRYIDRIREILISFSDVR  
TILVIEPDSLANMVTNMNVKPCSGAASTYRELTIVYALKQLDLPHVAMYMDAGHAGWLGWPANI  
QPAAELFAKIYEDAGKPRAVRGLATNVANYNAWSIASPPSYTSPNPNYDEKHIEAFAPLLRNQ

SCHEMA recombination uncovers stabilizing cellulase mutation

GFD\_AKFIVDTGRNGKQPTGQLEWGHWCNVKGTGFGVRPTANTGHELVDADFVWVKPGGESDG  
TSDTSAARYDYHCGLSDALTPAPEAGQWFQAYFEQLLINANPPF\_

33213332

GNPFSGVQLWANTYYSSEVHTLAIPSLS\_PELAAKAAKVAEVPSFQWLDRNVTVDTLFSGTLAEI  
RAANQRGGN\_\_YAGQFVVYDLPDRDCAALASNGEYSIADGGVAKYKNYIDRIREILISFSDVRTI  
LVIEPDSLANMVTNMNVPKCSGAASTYKELTVYALKQLNLPHVAMYMDAGHAGWLGWPANI  
QPAAELFAQIYRDAGRPAAVRGLATNVANYNAWSIASPPSYTSPNPNYDEKHIEAFAPLLRNQ  
GFD\_AKFIVDTGRNGKQPTGQLEWGHWCNVKGTGFGVRPTANTGHELVDADFVWVKPGGESDG  
TSDSSAPRFDSHCALPDALQPAPQAGAWFQAYFVQLLTNANPSFL

13333232

GNPFEGVQLWANNYYRSEVHTLAIPQITDPALRAAASAVA EVPSFQWLDRNVTVDTLFSGTLAE  
IRAANQRGANPPYAGIFVVYDLPDRDCAAAAASNGEWSIANNGANNYKRYIDRIRELLIQYSDIRTI  
LVIEPDSLANMVTNMNVQKCSNAASTYKELTVYALKQLNLPHVAMYMDAGHAGWLGWPANI  
QPAAELFAQIYRDAGRPAAVRGLATNVANYNGWNITSPPSYTOGNAVYNEKLYIHAIGPLLANH  
GWS\_AKFIVDTGRNGKQPTGQLEWGHWCNVKGTGFGVRPTANTGHELVDADFVWVKPGGESDG  
TSDSSAPRFDSHCALPDALQPAPQAGAWFQAYFVQLLTNANPSFL

22232132

GNPFVGVTPWANAYYASEVSSLAIPSLT\_GAMATAAAAVAKVPSFMWLDLTDKTPLEMTLA  
DIRTANKNGGN\_\_YAGQFVVYDLPDRDCAALASNGEYSIADGGVAKYKNYIDRIRELLIQYSDIR  
TILVIEPDSLANMVTNMNVQKCSNAASAYLECINYAVTQLNLPNVAMYLDAGHAGWLGWPAN  
QDPAAQLFANVYKNASSPRALRGLATNVANYNAWSVSSPPPYTSPNPNYDEKHIEAFRPLLEA  
RGFP\_AKFIVDTGRNGKQPTGQLEWGHWCNVKGTGFGVRPTANTGHELVDADFVWVKPGGESD  
GTSDDSSAPRFDSHCALPDALQPAPQAGAWFQAYFVQLLTNANPSFL

11113132

GNPFEGVQLWANNYYRSEVHTLAIPQITDPALRAAASAVA EVPSFQWLDRNVTVDTLVQTLSE  
IREANQAGANPQYAAQIVVYDLPDRDCAAAAASNGEWAIANNGVNNYKAYINRIREILISFSDVRT  
ILVIEPDSLANMVTNMNVQKCSGAASTYKELTVYALKQLNLPHVAMYMDAGHAGWLGWPANI  
QPAAELFAQIYRDAGRPAAVRGLATNVANYNAWSVSSPPPYTSPNPNYDEKHIEAFRPLLEAR  
GFP\_AKFIVDTGRNGKQPTGQLEWGHWCNVKGTGFGVRPTANTGHELVDADFVWVKPGGESDG  
TSDSSAPRFDSHCALPDALQPAPQAGAWFQAYFVQLLTNANPSFL

21333331

GNPFVGVTPWANAYYASEVSSLAIPSLT\_GAMATAAAAVAKVPSFQWLDRNVTVDTLVQTLSE  
EIREANQAGANPPYAGIFVVYDLPDRDCAAAAASNGEWSIANNGANNYKRYIDRIRELLIQYSDIR  
TILVIEPDSLANMVTNMNVQKCSNAASTYKELTVYALKQLNLPHVAMYMDAGHAGWLGWPA  
NIQPAAELFAQIYRDAGRPAAVRGLATNVANYNAWSIASPPSYTSPNPNYDEKHIEAFAPLLRN  
QGFD\_AKFIVDTGRNGKQPTGQLEWGHWCNVKGTGFGVRPTANTGHELVDADFVWVKPGGESD  
GTSDDTAAARYDYHCGLDALKPAPEAGQWFNEYFIQLLRNANPPF\_

21311131

GNPFVGVTPWANAYYASEVSSLAIPSLT\_GAMATAAAAVAKVPSFQWLDRNVTVDTLVQTLSE  
EIREANQAGANPPYAGIFVVYDLPDRDCAAAAASNGEWSIANNGANNYKRYIDRIREILISFSDVRT  
ILVIEPDSLANMVTNMNVQKCSGAASTYRELTIYALKQLDLPHVAMYMDAGHAGWLGWPANIQ  
PAAELFAKIYEDAGKPRAVRGLATNVANYNAWSVSSPPPYTSPNPNYDEKHIEAFRPLLEARGF  
P\_AKFIVDTGRNGKQPTGQLEWGHWCNVKGTGFGVRPTANTGHELVDADFVWVKPGGESDGTS  
DTTAARYDYHCGLDALKPAPEAGQWFNEYFIQLLRNANPPF\_

SCHEMA recombination uncovers stabilizing cellulase mutation

12332331

GNPFEGVQLWANNYYRSEVHTLAIPQITDPALRAAASAVAEVPSFMWLDT\_LDKTPLMEQTLAD  
IRTANKNGANPPYAGIFVVYDLPDRDCAAAAASNGEWSIANNGANNYKRYIDRIRELLIQYSDIRTI  
LVIEPDSLANMVTNMNVQKCSNAASAYLECINYAVTQLNLPNVAMYLDAGHAGWLGWPANQ  
DPAAQLFANVYKNASSPRALRGLATNVANYNAWSIASPPSYTSPNPNYDEKHIEAFAPLLRNQ  
GFD\_AKFIVDTGRNGKQPTGQLEWGHWCNVKGTGFGVRPTANTGHELVDADFVWVKPGGESDG  
TSDTTAARYDYHCGLEDALKPAPEAGQWFNEYFIQLLRNANPPF\_

13112332

GNPFEGVQLWANNYYRSEVHTLAIPQITDPALRAAASAVAEVPSFQWLDRNVTVDTLFSGTLAE  
IRAANQRGANPQYAAQIVVYDLPDRDCAAAAASNGEWAIANNGVNNYKAYINRIREILISFSDVRT  
ILVIEPDSLANMVTNMNVKCSGAASAYLECINYAVTQLNLPNVAMYLDAGHAGWLGWPANQ  
DPAAQLFANVYKNASSPRALRGLATNVANYNAWSIASPPSYTSPNPNYDEKHIEAFAPLLRNQ  
GFD\_AKFIVDTGRNGKQPTGQLEWGHWCNVKGTGFGVRPTANTGHELVDADFVWVKPGGESDG  
TSDSSAPRFDSHCALPDALQPAPQAGAWFQAYFVQLLTNANPSFL

22311331

GNPFVGVTPWANAYYASEVSSLAIPSLT\_GAMATAAAAVAKVPSFMWLDT\_LDKTPLMEQTLA  
DIRTANKNGANPPYAGIFVVYDLPDRDCAAAAASNGEWSIANNGANNYKRYIDRIREILISFSDVR  
TILVIEPDSLANMVTNMNVKCSGAASTYRELTIIYALKQLDLPHVAMYMDAGHAGWLGWPANI  
QPAAELFAKIYEDAGKPRAVRGLATNVANYNAWSIASPPSYTSPNPNYDEKHIEAFAPLLRNQ  
GFD\_AKFIVDTGRNGKQPTGQLEWGHWCNVKGTGFGVRPTANTGHELVDADFVWVKPGGESDG  
TSDTTAARYDYHCGLEDALKPAPEAGQWFNEYFIQLLRNANPPF\_

12111332

GNPFEGVQLWANNYYRSEVHTLAIPQITDPALRAAASAVAEVPSFMWLDT\_LDKTPLMEQTLAD  
IRTANKNGANPQYAAQIVVYDLPDRDCAAAAASNGEWAIANNGVNNYKAYINRIREILISFSDVRT  
ILVIEPDSLANMVTNMNVKCSGAASTYRELTIIYALKQLDLPHVAMYMDAGHAGWLGWPANI  
PAAELFAKIYEDAGKPRAVRGLATNVANYNAWSIASPPSYTSPNPNYDEKHIEAFAPLLRNQGF  
D\_AKFIVDTGRNGKQPTGQLEWGHWCNVKGTGFGVRPTANTGHELVDADFVWVKPGGESDGT  
DSSAPRFDSHCALPDALQPAPQAGAWFQAYFVQLLTNANPSFL

12112332

GNPFEGVQLWANNYYRSEVHTLAIPQITDPALRAAASAVAEVPSFMWLDT\_LDKTPLMEQTLAD  
IRTANKNGANPQYAAQIVVYDLPDRDCAAAAASNGEWAIANNGVNNYKAYINRIREILISFSDVRT  
ILVIEPDSLANMVTNMNVKCSGAASAYLECINYAVTQLNLPNVAMYLDAGHAGWLGWPANQ  
DPAAQLFANVYKNASSPRALRGLATNVANYNAWSIASPPSYTSPNPNYDEKHIEAFAPLLRNQ  
GFD\_AKFIVDTGRNGKQPTGQLEWGHWCNVKGTGFGVRPTANTGHELVDADFVWVKPGGESDG  
TSDSSAPRFDSHCALPDALQPAPQAGAWFQAYFVQLLTNANPSFL

12131331

GNPFEGVQLWANNYYRSEVHTLAIPQITDPALRAAASAVAEVPSFMWLDT\_LDKTPLMEQTLAD  
IRTANKNGANPQYAAQIVVYDLPDRDCAAAAASNGEWAIANNGVNNYKAYINRIRELLIQYSDIR  
TILVIEPDSLANMVTNMNVQKCSNAASTYRELTIIYALKQLDLPHVAMYMDAGHAGWLGWPANI  
QPAAELFAKIYEDAGKPRAVRGLATNVANYNAWSIASPPSYTSPNPNYDEKHIEAFAPLLRNQ  
GFD\_AKFIVDTGRNGKQPTGQLEWGHWCNVKGTGFGVRPTANTGHELVDADFVWVKPGGESDG  
TSDTTAARYDYHCGLEDALKPAPEAGQWFNEYFIQLLRNANPPF\_

SCHEMA recombination uncovers stabilizing cellulase mutation

12131332

GNPFEGVQLWANNYYRSEVHTLAIPQITDPALRAAASAVAEVPSFMWLDT\_LDKTPLMEQTLAD  
IRTANKNGANPQYAAQIVVYDLPDRDCAAAAASNGEWAIANNGVNKYAYINRIRELLIQYSDIR  
TILVIEPDSLANMVTNMNVQKCSNAASTYRELTIYALKQLDLPHVAMYMDAGHAGWLGWPANI  
QPAAELFAKIYEDAGKPRAVRGLATNVANYNAWSIASPPSYTSPNPNYDEKHIEAFAPLLRNQ  
GFD\_AKFIVDTGRNGKQPTGQLEWGHWCNVKGTGFGVRPTANTGHELVDADFVWVKPGGESDG  
TSDSSAPRFDSHCALPDALQPAPQAGAWFQAYFVQLLTNANPSFL

12332332

GNPFEGVQLWANNYYRSEVHTLAIPQITDPALRAAASAVAEVPSFMWLDT\_LDKTPLMEQTLAD  
IRTANKNGANPPYAGIFVVYDLPDRDCAAAAASNGEWSIANNGANNYKRYIDRIRELLIQYSDIRTI  
LVIEPDSLANMVTNMNVQKCSNAASAYLECINYAVTQLNLPNVAMYLDAGHAGWLGWPANQ  
DPAAQLFANVYKNASSPRALRGLATNVANYNAWSIASPPSYTSPNPNYDEKHIEAFAPLLRNQ  
GFD\_AKFIVDTGRNGKQPTGQLEWGHWCNVKGTGFGVRPTANTGHELVDADFVWVKPGGESDG  
TSDSSAPRFDSHCALPDALQPAPQAGAWFQAYFVQLLTNANPSFL

1211131

GNPFEGVQLWANNYYRSEVHTLAIPQITDPALRAAASAVAEVPSFMWLDT\_LDKTPLMEQTLAD  
IRTANKNGANPQYAAQIVVYDLPDRDCAAAAASNGEWAIANNGVNKYAYINRIREILISFSDVRT  
ILVIEPDSLANMVTNMNVKCSGAASTYRELTIYALKQLDLPHVAMYMDAGHAGWLGWPANIQ  
PAAELFAKIYEDAGKPRAVRGLATNVANYNAWSVSSPPPYTSPNPNYDEKHIEAFRPLLEARGF  
P\_AKFIVDTGRNGKQPTGQLEWGHWCNVKGTGFGVRPTANTGHELVDADFVWVKPGGESDGTS  
DTTAARYDYHCGLEDALKPAPEAGQWFNEYFIQLLRNANPPF\_

12311332

GNPFEGVQLWANNYYRSEVHTLAIPQITDPALRAAASAVAEVPSFMWLDT\_LDKTPLMEQTLAD  
IRTANKNGANPPYAGIFVVYDLPDRDCAAAAASNGEWSIANNGANNYKRYIDRIREILISFSDVRTI  
LVIEPDSLANMVTNMNVKCSGAASTYRELTIYALKQLDLPHVAMYMDAGHAGWLGWPANIQ  
PAAELFAKIYEDAGKPRAVRGLATNVANYNAWSIASPPSYTSPNPNYDEKHIEAFAPLLRNQGF  
D\_AKFIVDTGRNGKQPTGQLEWGHWCNVKGTGFGVRPTANTGHELVDADFVWVKPGGESDGTS  
DSSAPRFDSHCALPDALQPAPQAGAWFQAYFVQLLTNANPSFL

13332331

GNPFEGVQLWANNYYRSEVHTLAIPQITDPALRAAASAVAEVPSFQWLDRNVTVDTLFSGTLAE  
IRAANQRGANPPYAGIFVVYDLPDRDCAAAAASNGEWSIANNGANNYKRYIDRIRELLIQYSDIRTI  
LVIEPDSLANMVTNMNVQKCSNAASAYLECINYAVTQLNLPNVAMYLDAGHAGWLGWPANQ  
DPAAQLFANVYKNASSPRALRGLATNVANYNAWSIASPPSYTSPNPNYDEKHIEAFAPLLRNQ  
GFD\_AKFIVDTGRNGKQPTGQLEWGHWCNVKGTGFGVRPTANTGHELVDADFVWVKPGGESDG  
TSDTTAARYDYHCGLEDALKPAPEAGQWFNEYFIQLLRNANPPF\_

12132331

GNPFEGVQLWANNYYRSEVHTLAIPQITDPALRAAASAVAEVPSFMWLDT\_LDKTPLMEQTLAD  
IRTANKNGANPQYAAQIVVYDLPDRDCAAAAASNGEWAIANNGVNKYAYINRIRELLIQYSDIR  
TILVIEPDSLANMVTNMNVQKCSNAASAYLECINYAVTQLNLPNVAMYLDAGHAGWLGWPAN  
QDPAAQLFANVYKNASSPRALRGLATNVANYNAWSIASPPSYTSPNPNYDEKHIEAFAPLLRN  
QGFD\_AKFIVDTGRNGKQPTGQLEWGHWCNVKGTGFGVRPTANTGHELVDADFVWVKPGGESD  
GTSDDTTAARYDYHCGLEDALKPAPEAGQWFNEYFIQLLRNANPPF\_

12132332

GNPFEGVQLWANNYYRSEVHTLAIPQITDPALRAAASAVAEVPSFMWLDT\_LDKTPLMEQTLAD

SCHEMA recombination uncovers stabilizing cellulase mutation

IRTANKNGANPQYAAQIVVYDLPDRDCAAAAASNGEWAIANNGVNKYKAYINRIRELLIQYSDIR  
TILVIEPDSLANTMNTMNVQKCSNAASAYLECINYAVTQLNLPNVAMYLDAGHAGWLGWPANQ  
QDPAAQLFANVYKNASSPRALRGLATNVANYNAWSIASPPSYTSPNPNYDEKHYIEAFAPLLRN  
QGFD\_AKFIVDTGRNGKQPTGQLEWGHWCNVKGTGFGVRPTANTGHELVDADFVWVKPGGESD  
GTSDDSAPRFDSHCALPDALQPAPQAGAWFQAYFVQLLTNANPSFL

13332332

GNPFEGVQLWANNYYRSEVHTLAIPQITDPALRAAASAVAEPVPSFQWLDRNVTVDTLFSGTLAE  
IRAANQRGANPPYAGIFVVYDLPDRDCAAAAASNGEWSIANNGANNYKRYIDRIRELLIQYSDIRTI  
LVIEPDSLANTMNTMNVQKCSNAASAYLECINYAVTQLNLPNVAMYLDAGHAGWLGWPANQ  
DPAAQLFANVYKNASSPRALRGLATNVANYNAWSIASPPSYTSPNPNYDEKHYIEAFAPLLRNQ  
GFD\_AKFIVDTGRNGKQPTGQLEWGHWCNVKGTGFGVRPTANTGHELVDADFVWVKPGGESD  
TSDSSAPRFDSHCALPDALQPAPQAGAWFQAYFVQLLTNANPSFL

12112132

GNPFEGVQLWANNYYRSEVHTLAIPQITDPALRAAASAVAEPVPSFMWLDL\_LDKTPLMEQTLAD  
IRTANKNGANPQYAAQIVVYDLPDRDCAAAAASNGEWAIANNGVNKYKAYINRIREILISFSDVRT  
ILVIEPDSLANTMNTMNVKCSGAASAYLECINYAVTQLNLPNVAMYLDAGHAGWLGWPANQ  
DPAAQLFANVYKNASSPRALRGLATNVANYNAWSVSSPPPYTSPNPNYDEKHYIEAFRPLLEAR  
GFP\_AKFIVDTGRNGKQPTGQLEWGHWCNVKGTGFGVRPTANTGHELVDADFVWVKPGGESD  
TSDSSAPRFDSHCALPDALQPAPQAGAWFQAYFVQLLTNANPSFL

13322332

GNPFEGVQLWANNYYRSEVHTLAIPQITDPALRAAASAVAEPVPSFQWLDRNVTVDTLFSGTLAE  
IRAANQRGANPPYAGIFVVYDLPDRDCAAAAASNGEWSIANNGANNYKRYIDTIRQIVVEYSDIRT  
LLVIEPDSLANTLVTNLGTPKCANAQSAAYLECINYAVTQLNLPNVAMYLDAGHAGWLGWPANQ  
DPAAQLFANVYKNASSPRALRGLATNVANYNAWSIASPPSYTSPNPNYDEKHYIEAFAPLLRNQ  
GFD\_AKFIVDTGRNGKQPTGQLEWGHWCNVKGTGFGVRPTANTGHELVDADFVWVKPGGESD  
TSDSSAPRFDSHCALPDALQPAPQAGAWFQAYFVQLLTNANPSFL

13131332

GNPFEGVQLWANNYYRSEVHTLAIPQITDPALRAAASAVAEPVPSFQWLDRNVTVDTLFSGTLAE  
IRAANQRGANPQYAAQIVVYDLPDRDCAAAAASNGEWAIANNGVNKYKAYINRIRELLIQYSDIR  
TILVIEPDSLANTMNTMNVQKCSNAASTYRELTIIYALKQLDLPHVAMYMDAGHAGWLGWPANI  
QPAAELFAKIYEDAGKPRAVRGLATNVANYNAWSIASPPSYTSPNPNYDEKHYIEAFAPLLRNQ  
GFD\_AKFIVDTGRNGKQPTGQLEWGHWCNVKGTGFGVRPTANTGHELVDADFVWVKPGGESD  
TSDSSAPRFDSHCALPDALQPAPQAGAWFQAYFVQLLTNANPSFL

12331332

GNPFEGVQLWANNYYRSEVHTLAIPQITDPALRAAASAVAEPVPSFMWLDL\_LDKTPLMEQTLAD  
IRTANKNGANPPYAGIFVVYDLPDRDCAAAAASNGEWSIANNGANNYKRYIDRIRELLIQYSDIRTI  
LVIEPDSLANTMNTMNVQKCSNAASTYRELTIIYALKQLDLPHVAMYMDAGHAGWLGWPANIQ  
PAAELFAKIYEDAGKPRAVRGLATNVANYNAWSIASPPSYTSPNPNYDEKHYIEAFAPLLRNQGF  
D\_AKFIVDTGRNGKQPTGQLEWGHWCNVKGTGFGVRPTANTGHELVDADFVWVKPGGESD  
GTS DSSAPRFDSHCALPDALQPAPQAGAWFQAYFVQLLTNANPSFL

13312332

GNPFEGVQLWANNYYRSEVHTLAIPQITDPALRAAASAVAEPVPSFQWLDRNVTVDTLFSGTLAE  
IRAANQRGANPPYAGIFVVYDLPDRDCAAAAASNGEWSIANNGANNYKRYIDRIREILISFSDVRTI  
LVIEPDSLANTMNTMNVKCSGAASAYLECINYAVTQLNLPNVAMYLDAGHAGWLGWPANQD



SCHEMA recombination uncovers stabilizing cellulase mutation

PAAQLFANVYKNASSPRALRGLATNVANYNAWSIASPPSYTSPNPNYDEKHYIEAFAPLLRNQG  
FD\_AKFIVDTGRNGKQPTGQLEWGHWCNVKGTGFGVRPTANTGHELVDADFVWVKPGGESDGT  
SDSSAPRFDSHCALPDALQPAPQAGAWFQAYFVQLLTNANPSFL

11113332

GPNFEGVQLWANNYYRSEVHTLAIPQITDPALRAAASAVAEPVPSFQWLDRNVTVDTLVQTLSE  
IREANQAGANPQYAAQIVVYDLPDRDCAAASNGEWAIANNGVNNYKAYINRIREILISFSDVRT  
ILVIEPDSLANMVTNMNVPKCSGAASTYKELTVYALKQLNLPVAMYMDAGHAGWLGWPANI  
QPAAELFAQIYRDAGRPAAVRGLATNVANYNAWSIASPPSYTSPNPNYDEKHYIEAFAPLLRNQ  
GFD\_AKFIVDTGRNGKQPTGQLEWGHWCNVKGTGFGVRPTANTGHELVDADFVWVKPGGESDGT  
TSDSSAPRFDSHCALPDALQPAPQAGAWFQAYFVQLLTNANPSFL

13113132

GPNFEGVQLWANNYYRSEVHTLAIPQITDPALRAAASAVAEPVPSFQWLDRNVTVDTLFSGTLAE  
IRAANQRGANPQYAAQIVVYDLPDRDCAAASNGEWAIANNGVNNYKAYINRIREILISFSDVRT  
ILVIEPDSLANMVTNMNVPKCSGAASTYKELTVYALKQLNLPVAMYMDAGHAGWLGWPANI  
QPAAELFAQIYRDAGRPAAVRGLATNVANYNAWSVSSPPPYTSPNPNYDEKHYIEAFRPLLEAR  
GFP\_AKFIVDTGRNGKQPTGQLEWGHWCNVKGTGFGVRPTANTGHELVDADFVWVKPGGESDGT  
TSDSSAPRFDSHCALPDALQPAPQAGAWFQAYFVQLLTNANPSFL

11112132

GPNFEGVQLWANNYYRSEVHTLAIPQITDPALRAAASAVAEPVPSFQWLDRNVTVDTLVQTLSE  
IREANQAGANPQYAAQIVVYDLPDRDCAAASNGEWAIANNGVNNYKAYINRIREILISFSDVRT  
ILVIEPDSLANMVTNMNVPKCSGAASAYLECINYAVTQLNLPNVAMYLDAGHAGWLGWPANQ  
DPAAQLFANVYKNASSPRALRGLATNVANYNAWSVSSPPPYTSPNPNYDEKHYIEAFRPLLEAR  
GFP\_AKFIVDTGRNGKQPTGQLEWGHWCNVKGTGFGVRPTANTGHELVDADFVWVKPGGESDGT  
TSDSSAPRFDSHCALPDALQPAPQAGAWFQAYFVQLLTNANPSFL

12113132

GPNFEGVQLWANNYYRSEVHTLAIPQITDPALRAAASAVAEPVPSFMWLDL\_LDKTPLMEQTLAD  
IRTANKNGANPQYAAQIVVYDLPDRDCAAASNGEWAIANNGVNNYKAYINRIREILISFSDVRT  
ILVIEPDSLANMVTNMNVPKCSGAASTYKELTVYALKQLNLPVAMYMDAGHAGWLGWPANI  
QPAAELFAQIYRDAGRPAAVRGLATNVANYNAWSVSSPPPYTSPNPNYDEKHYIEAFRPLLEAR  
GFP\_AKFIVDTGRNGKQPTGQLEWGHWCNVKGTGFGVRPTANTGHELVDADFVWVKPGGESDGT  
TSDSSAPRFDSHCALPDALQPAPQAGAWFQAYFVQLLTNANPSFL

13132332

GPNFEGVQLWANNYYRSEVHTLAIPQITDPALRAAASAVAEPVPSFQWLDRNVTVDTLFSGTLAE  
IRAANQRGANPQYAAQIVVYDLPDRDCAAASNGEWAIANNGVNNYKAYINRIRELLIQYSDIR  
TILVIEPDSLANMVTNMNVQKCSNAASAYLECINYAVTQLNLPNVAMYLDAGHAGWLGWPAN  
QDPAAQLFANVYKNASSPRALRGLATNVANYNAWSIASPPSYTSPNPNYDEKHYIEAFAPLLRN  
QGFD\_AKFIVDTGRNGKQPTGQLEWGHWCNVKGTGFGVRPTANTGHELVDADFVWVKPGGESD  
GTSDSSAPRFDSHCALPDALQPAPQAGAWFQAYFVQLLTNANPSFL

11111132

GPNFEGVQLWANNYYRSEVHTLAIPQITDPALRAAASAVAEPVPSFQWLDRNVTVDTLVQTLSE  
IREANQAGANPQYAAQIVVYDLPDRDCAAASNGEWAIANNGVNNYKAYINRIREILISFSDVRT  
ILVIEPDSLANMVTNMNVPKCSGAASTYRELTIIYALKQLDLPHVAMYMDAGHAGWLGWPANIQ  
PAAELFAKIYEDAGKPRAVRGLATNVANYNAWSVSSPPPYTSPNPNYDEKHYIEAFRPLLEARGF

SCHEMA recombination uncovers stabilizing cellulase mutation

P\_AKFIVDTGRNGKQPTGQLEWGHWCNVKGTGFGVRPTANTGHELVDADFVWVKPGGESDGTS  
DSSAPRFDSHCALPDALQPAPQAGAWFQAYFVQLLTNANPSFL

13331332

GPNFEGVQLWANNYYRSEVHTLAIPQITDPALRAAASAVA EVPSFQWLDRNVTVDTLFSGTLAE  
IRAANQRGANPPYAGIFVVYDLPDRDCAAAAASNGEWSIANNGANNYKRYIDRIRELLIQYSDIRTI  
LVIEPDSL ANMVTNMNVKCSNAASTYRELTIYALKQLDLPHVAMYMDAGHAGWLGWPANIQ  
PAAELFAKIYEDAGKPR AVRGLATNVANYNAWSIASPPSYTSPNPNYDEKH YIEAFAPLLRNQGF  
D\_AKFIVDTGRNGKQPTGQLEWGHWCNVKGTGFGVRPTANTGHELVDADFVWVKPGGESDGTS  
DSSAPRFDSHCALPDALQPAPQAGAWFQAYFVQLLTNANPSFL

13111132

GPNFEGVQLWANNYYRSEVHTLAIPQITDPALRAAASAVA EVPSFQWLDRNVTVDTLFSGTLAE  
IRAANQRGANPQYAAQIVVYDLPDRDCAAAAASNGEWAIANNGVN NYKAYINRIREILISFSDVRT  
ILVIEPDSL ANMVTNMNVKCSGAASTYRELTIYALKQLDLPHVAMYMDAGHAGWLGWPANIQ  
PAAELFAKIYEDAGKPR AVRGLATNVANYNAWSVSSPPPYTSPNPNYDEKH YIEAFRPLLEARGF  
P\_AKFIVDTGRNGKQPTGQLEWGHWCNVKGTGFGVRPTANTGHELVDADFVWVKPGGESDGTS  
DSSAPRFDSHCALPDALQPAPQAGAWFQAYFVQLLTNANPSFL

1222132

GPNFEGVQLWANNYYRSEVHTLAIPQITDPALRAAASAVA EVPSFMWLDT\_LDKTPLMEQTLAD  
IRTANKNGGN\_\_YAGQFVVYDLPDRDCAALASNGEYSIADGGVAKYKNYIDTIRQIVVEYSDIRT  
LLVIEPDSL ANLVTNLGTPKCANAQSAYLECINYAVTQLNLPNVAMYLDAGHAGWLGWPANQ  
DPAAQLFANVYKNASSPRALRGLATNVANYNAWSVSSPPPYTSPNPNYDEKH YIEAFRPLLEAR  
GFP\_AKFIVDTGRNGKQPTGQLEWGHWCNVKGTGFGVRPTANTGHELVDADFVWVKPGGESDG  
TSDSSAPRFDSHCALPDALQPAPQAGAWFQAYFVQLLTNANPSFL

1222332

GPNFEGVQLWANNYYRSEVHTLAIPQITDPALRAAASAVA EVPSFMWLDT\_LDKTPLMEQTLAD  
IRTANKNGGN\_\_YAGQFVVYDLPDRDCAALASNGEYSIADGGVAKYKNYIDTIRQIVVEYSDIRT  
LLVIEPDSL ANLVTNLGTPKCANAQSAYLECINYAVTQLNLPNVAMYLDAGHAGWLGWPANQ  
DPAAQLFANVYKNASSPRALRGLATNVANYNAWSIASPPSYTSPNPNYDEKH YIEAFAPLLRNQ  
GFD\_AKFIVDTGRNGKQPTGQLEWGHWCNVKGTGFGVRPTANTGHELVDADFVWVKPGGESDG  
TSDSSAPRFDSHCALPDALQPAPQAGAWFQAYFVQLLTNANPSFL

13311332

GPNFEGVQLWANNYYRSEVHTLAIPQITDPALRAAASAVA EVPSFQWLDRNVTVDTLFSGTLAE  
IRAANQRGANPPYAGIFVVYDLPDRDCAAAAASNGEWSIANNGANNYKRYIDRIREILISFSDVRTI  
LVIEPDSL ANMVTNMNVKCSGAASTYRELTIYALKQLDLPHVAMYMDAGHAGWLGWPANIQ  
PAAELFAKIYEDAGKPR AVRGLATNVANYNAWSIASPPSYTSPNPNYDEKH YIEAFAPLLRNQGF  
D\_AKFIVDTGRNGKQPTGQLEWGHWCNVKGTGFGVRPTANTGHELVDADFVWVKPGGESDGTS  
DSSAPRFDSHCALPDALQPAPQAGAWFQAYFVQLLTNANPSFL

13311331

GPNFEGVQLWANNYYRSEVHTLAIPQITDPALRAAASAVA EVPSFQWLDRNVTVDTLFSGTLAE  
IRAANQRGANPPYAGIFVVYDLPDRDCAAAAASNGEWSIANNGANNYKRYIDRIREILISFSDVRTI  
LVIEPDSL ANMVTNMNVKCSGAASTYRELTIYALKQLDLPHVAMYMDAGHAGWLGWPANIQ  
PAAELFAKIYEDAGKPR AVRGLATNVANYNAWSIASPPSYTSPNPNYDEKH YIEAFAPLLRNQGF  
D\_AKFIVDTGRNGKQPTGQLEWGHWCNVKGTGFGVRPTANTGHELVDADFVWVKPGGESDGTS  
D\_TTAARYDYHCGLEDALKPAPEAGQWFNEYFIQLLRNANPPF\_

SCHEMA recombination uncovers stabilizing cellulase mutation

SUPPLEMENTAL TABLE 3

T<sub>50</sub> linear regression model parameters and p-values. Parameter values with  $p \leq 0.1$ , used to calculate the regression fit line of Figure 1, appear in bold. Block effects are reported relative to a parent 1 (*H. insolens* CBH II) reference state with 16 parameters representing substitution of each of the 8 blocks from parents 2 and 3.

Block	Parameter Value	p-value
<b>Parent1</b>	<b>62.8</b>	<b>0.00</b>
B12	-0.9	0.35
<b>B13</b>	<b>-3.5</b>	<b>0.00</b>
<b>B22</b>	<b>-1.7</b>	<b>0.06</b>
B23	-1.1	0.25
B32	0.5	0.68
<b>B33</b>	<b>1.2</b>	<b>0.10</b>
<b>B42</b>	<b>2.7</b>	<b>0.05</b>
B43	0.0	0.99
<b>B52</b>	<b>-1.3</b>	<b>0.10</b>
B53	-0.6	0.50
<b>B62</b>	<b>-3.5</b>	<b>0.02</b>
B63	-0.7	0.37
<b>B72</b>	<b>-3.8</b>	<b>0.05</b>
<b>B73</b>	<b>8.5</b>	<b>0.00</b>
B82	0.0	1.00
<b>B83</b>	<b>-5.6</b>	<b>0.00</b>

SCHEMA recombination uncovers stabilizing cellulase mutation

SUPPLEMENTAL TABLE 4

Specific activity values ( $\mu\text{g}$  glucose reducing sugar equivalent/ $(\mu\text{g}$  CBH II enzyme  $\times$  min  $\times 10^2$ ) for native, point mutant and selected thermostable chimeric CBH IIs. Error bars show standard errors, where standard error is defined as standard dev/sqrt (n), for three replicates. 2-hr reaction, 3 mg enzyme/g PASC, 50°C, 25 mM sodium acetate, pH 4.8.

CBH II Enzyme	Specific Activity $\mu\text{g}$ Reducing Sugar/ $(\mu\text{g}$ Enzyme $\times$ min) $\times 10^2$
<i>Humicola insolens</i> (Parent 1)	5.3 +/- 0.5
<i>Hypocrea jecorina</i> (Parent 2)	8.4 +/- 0.4
<i>Chaetomium thermophilum</i> (Parent 3)	4.8 +/- 0.3
<i>Phanerochaete chrysosporium</i>	7.7 +/- 0.3
<i>Humicola insolens</i> C314S	5.3 +/- 0.9
<i>Hypocrea jecorina</i> C311S	7.8 +/- 0.5
<i>Phanerochaete chrysosporium</i> C311S	8.5 +/- 0.1
HJPlus (Chimera 12222332)	9.6 +/- 0.8
Chimera 13111132	8.5 +/- 0.3
Chimera 22222232	7.7 +/- 0.3
Chimera 13311332	6.8 +/- 0.6
Chimera 13311331	6.2 +/- 0.3
Chimera 11111131	6.1 +/- 0.9
Chimera 13112332	5.6 +/- 0.4
Chimera 21311131	5.5 +/- 0.3
Chimera 11113132	5.3 +/- 0.5
Chimera 21333331	3.8 +/- 0.4

## SUPPLEMENTAL TABLE 5

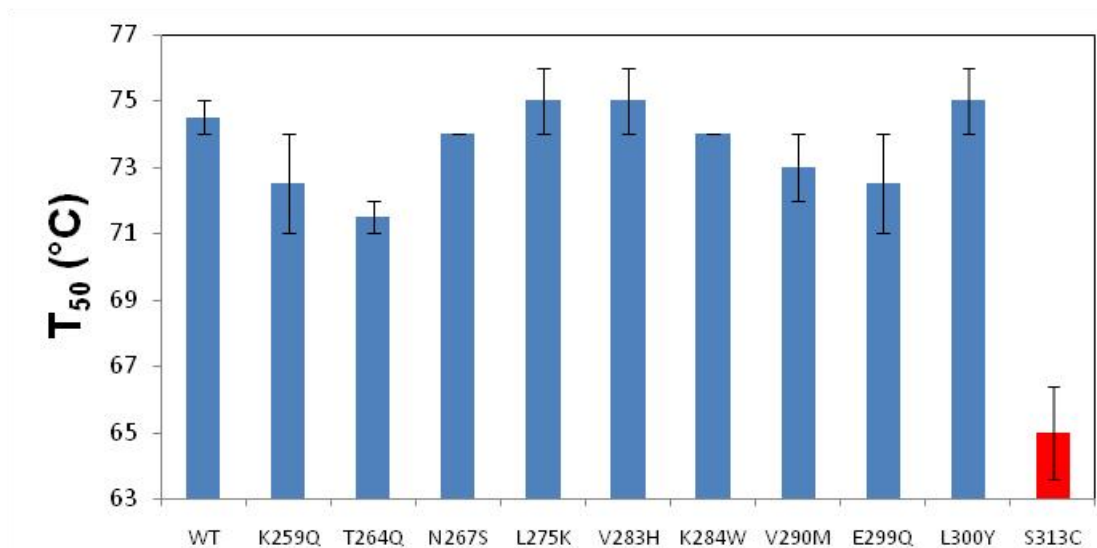
Total activity in synthetic (SDCAA) and rich (YPD) expression culture media supernatants for *H. jecorina* and *H. insolens* wild type, Ser point mutant and B7P3 block substitution CBH IIs. Values presented are  $\mu\text{g}$  glucose/mL cellulase activity assay per mL of expression culture supernatant CBH II equivalent added to cellulase activity assay. For SDCAA cultures, concentrated SDCAA culture supernatants were used and activity toward phosphoric acid swollen cellulose (1 mg/mL) at 50 °C for 100 minutes in 50 mM sodium acetate, pH 4.8, was measured. YPD supernatant CBH II was concentrated by binding to Avicel and activity toward Avicel (15 mg/mL) at 55 °C for 150 minutes in 50 mM sodium acetate, pH 4.8, was measured.

<b>CBH II</b>	<b>SDCAA (1)</b>	<b>SDCAA (2)</b>	<b>SDCAA Mean</b>	<b>YPD (1)</b>	<b>YPD (2)</b>	<b>YPD Mean</b>
<i>H. jecorina</i>	19	17	18	0.4	0.4	0.4
<i>H. jeco C311S</i>	50	43	47	6.1	5.6	5.9
<i>H. jeco B7P3</i>	35	33	34	3.9	3.6	3.8
<i>H. insolens</i>	73	83	78	6.2	6.0	6.1
<i>H. inso C314S</i>	100	97	98	8.8	8.0	8.4
<i>H. inso B7P3</i>	39	42	40	4.4	4.1	4.2



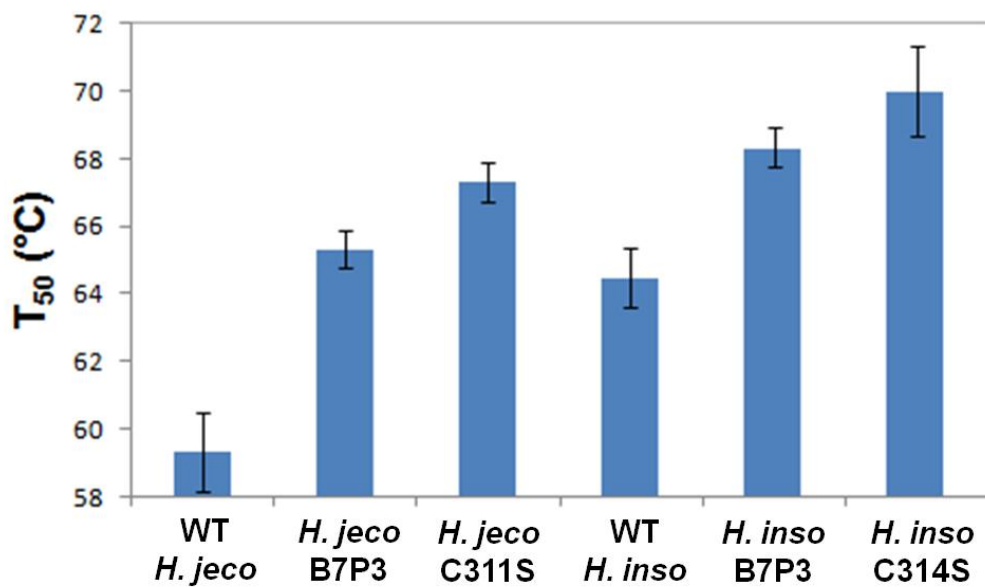
SUPPLEMENTAL FIGURE 2

$T_{50}$  values for 21111331 chimera point mutants. Values shown as average of two independent duplicates, error bars indicate duplicate  $T_{50}$  values for each point mutant. Inactivation was carried out for 10 minutes at the temperature being tested in 50 mM sodium acetate buffer, pH 4.8. Residual activity was determined by incubation with 1 g/L phosphoric acid swollen cellulose (PASC) in above buffer for 100 minutes at 50 °C.



SUPPLEMENTAL FIGURE 3

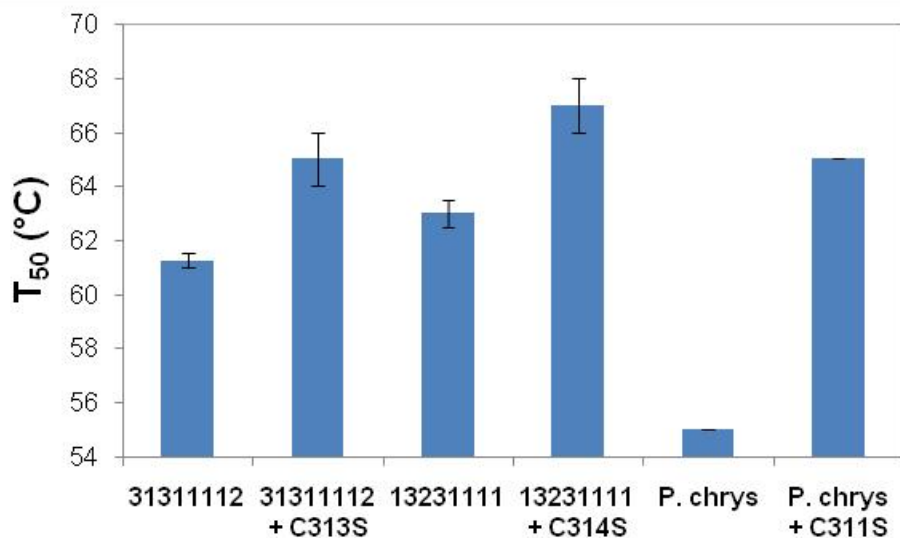
$T_{50}$  values for *H. insolens* and *H. jecorina* parent CBH IIs, Ser single point mutants and B7P3 block substitution chimeras. Values shown as average of three independent replicates, error bars indicate one standard deviation for each CBH II. Inactivation was carried out for 10 minutes at the temperature being tested in 50 mM sodium acetate buffer, pH 4.8. Residual activity was determined by incubation with 1 g/L phosphoric acid swollen cellulose (PASC) in above buffer for 100 minutes at 50 °C.





SUPPLEMENTAL FIGURE 4

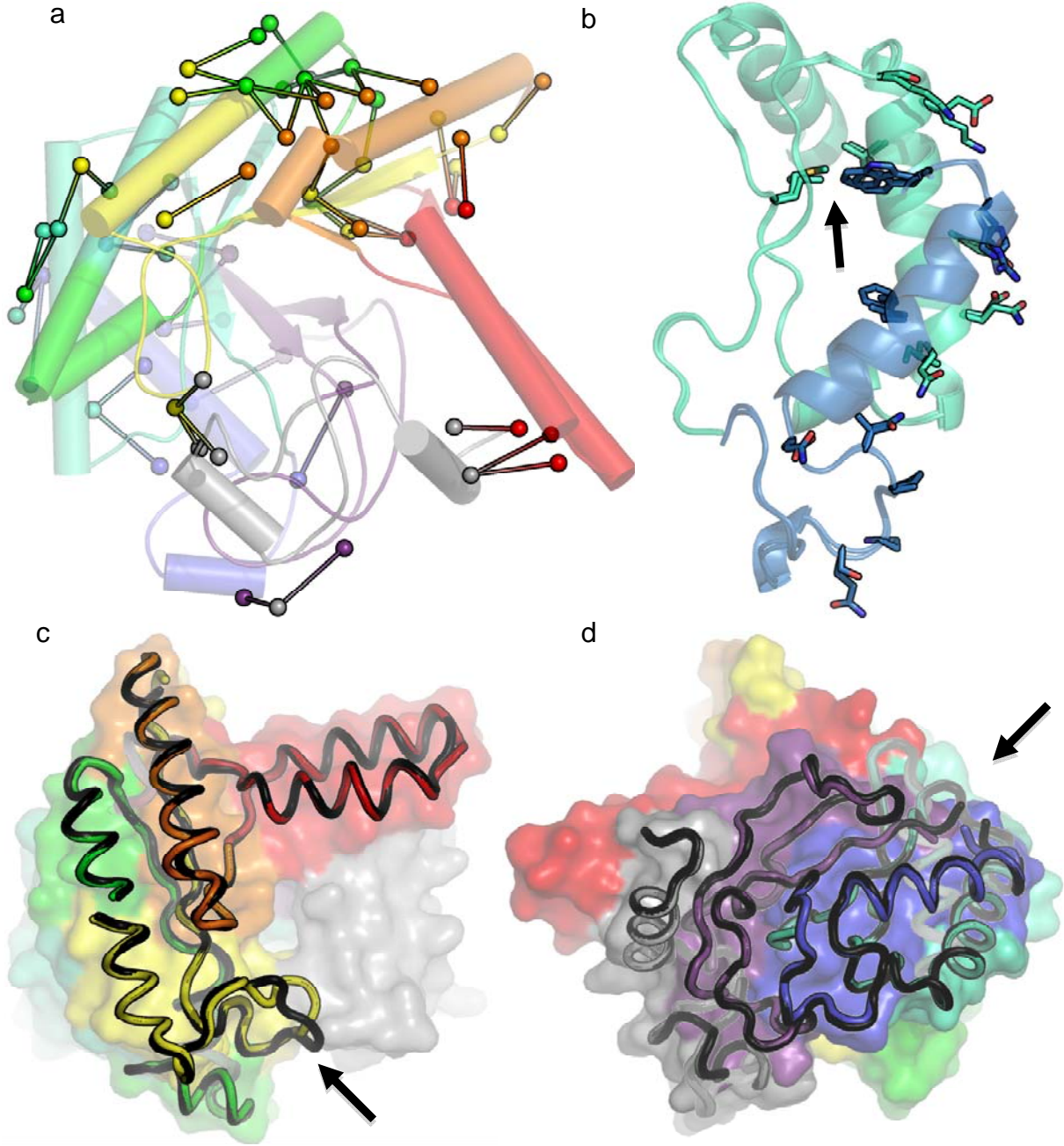
$T_{50}$  values for CBH II chimeras 31311112, 13231111 and the wild type CBH II catalytic domain from *P. chrysosporium* (fused to the *H. jecorina* CBM) and heterologously secreted from *S. cerevisiae*. Values shown as two independent replicates with error bars indicating values for each trial. Inactivation was carried out for 10 minutes at the temperature tested, in 50 mM sodium acetate buffer, pH 4.8. Residual activity was determined by incubation with 1 g/L phosphoric acid swollen cellulose (PASC) in above buffer for 100 minutes at 50 °C.



SUPPLEMENTAL FIGURE 5

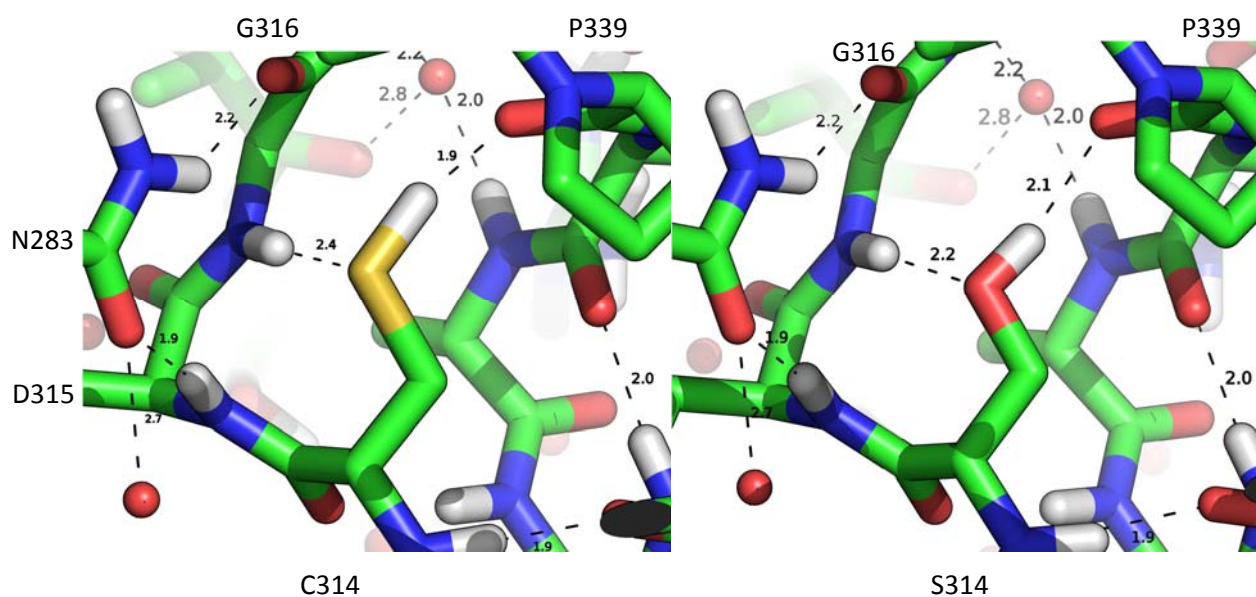
CBH II recombination block interfaces. (a) Inter-block sites where novel non-parental residue pairs are possible (connected spheres) are usually surface-exposed, potentially allowing solvent to screen the interactions. (b) An example interface (B5-B6) illustrates conservation of the backbone (cartoons for aligned *H. jecorina* and *H. insolens*), variable residues on the surface, and the comparatively rare possibility of a novel buried hydrophobic pair at residues 173 and 253 (arrow). (c) Blocks 1-4 from *H. jecorina* (black cartoon) match cognate *H. insolens* blocks (color-coded cartoon) without large deviations, though movement associated with substrate binding is observed (arrow) in part of B3 (yellow). (d) Cognate blocks 5-8 are also similar, though the indel at the B6,B7 junction (arrow) will require conformational change.

SCHEMA recombination uncovers stabilizing cellulase mutation



## SUPPLEMENTAL FIGURE 6

Structural analysis of C314S mutation and its stabilizing effect. (a) Hydrogen positions for high-resolution *H.insolens* structure (1ocn) were added with REDUCE.<sup>1</sup> (b) The reconfigured geometry of the analogous serine structure was modeled in PyMOL (<http://www.pymol.org>). Sidechain optimization in the SHARPEN<sup>2</sup> modeling platform (with an all-atom Rosetta energy function) also suggested that both the Cys314 and Ser314 would donate hydrogen bonds to the carbonyl of Pro339, and accept hydrogen bonds from the amide of Gly316. The superior hydrogen bonding capacity of serine may play a role in the greater stability of the serine containing variants. Another possible explanation is geometric complementarity. Specifically, the Cys position from 1ocn shows evidence of conformational strain in that the sidechain is noticeably bent (i.e. the improper dihedral angle from N-C-C $\alpha$ -C $\beta$  is 6° from the standard position), increasing the distance from the Pro carbonyl. Numbers in figure not preceded by letters denote hydrogen bond distances (Å).



SUPPLEMENTAL FIGURE 7

Alignment of 60 amino acid frame bracketing *H. jecorina* CBH II Cys311 for 196 protein sequences having highest identity to *H. jecorina* CBH II. Residues at 311 equivalent position denoted by shaded bar. Sequences for recombinant *C. thermophilum*, *H. insolens* and *P. chrys* CBH IIs studied in this work are denoted as *C. ther*, *H. inso* and *P. chrys*. Fifty-four of the 250 most identical sequences were excluded due to redundancy (i.e. point mutants for structural studies or >95% identical isoforms).

<i>H. jeco</i>	---	T---	G---	D---	S---	L---	LDSFVWVKPGGEC	SDG	--	T---	S-----	-----
XP_001395308	---	T---	G---	D---	A---	L---	VDAFVWVKPGGEC	SDG	--	T---	S-----	-----
CAP93233	---	T---	G---	D---	A---	L---	ADAFVWVKPGGEC	SDG	--	T---	S-----	-----
XP_662886	---	T---	G---	D---	S---	L---	LDAFVWVKPGGEC	SDG	--	T---	S-----	-----
XP_001552807	---	T---	G---	D---	A---	L---	EDAFVWVKPGGEC	SDG	--	T---	S-----	-----
XP_001216114	---	T---	G---	N---	S---	L---	VDAFVWVKPGGEC	SDG	--	T---	S-----	-----
XP_001258843	---	T---	G---	D---	P---	L---	QDAFVWVKPGGEC	SDG	--	T---	S-----	-----
XP_001392295	---	T---	G---	D---	E---	L---	ADAFVWVKPGGEC	SDG	--	T---	S-----	-----
XP_002380541	---	T---	G---	D---	E---	L---	VDAFVWVKPGGEC	SDG	--	T---	S-----	-----
AAQ72468	---	T---	G---	D---	A---	L---	VDAFVWVKPGGEC	SDG	--	T---	S-----	-----
P46236	---	T---	G---	D---	A---	L---	ADAFVWVKPGGEC	SDG	--	T---	S-----	-----
XP_001903170	---	T---	G---	L---	D---	I---	EDAFVWVKPGGEC	SDG	--	T---	S-----	-----
JC7931	---	T---	G---	N---	P---	L---	EDAFVWVKPGGEC	SDG	--	T---	S-----	-----
XP_001273717	---	T---	G---	D---	D---	L---	LDAFVWVKPGAES	SDG	--	T---	S-----	-----
XP_001210279	---	T---	G---	D---	E---	L---	VDAFVWVKPGGEC	SDG	--	T---	S-----	-----
XP_002484367	---	T---	G---	D---	P---	L---	EDAFVWVKPGGEC	SDG	--	T---	S-----	-----
XP_001264772	---	T---	G---	D---	K---	L---	VDAFVWVKPGGEC	SDG	--	T---	S-----	-----
XP_960770	---	T---	G---	S---	S---	L---	ADAFVWVKPGGEC	SDG	--	T---	S-----	-----
XP_658877	---	T---	G---	D---	E---	L---	VDAFVWVKPGGEC	SDG	--	T---	S-----	-----
ACH91035	---	T---	G---	D---	P---	L---	EDAFVWVKPGGEC	SDG	--	T---	S-----	-----
XP_001226029	---	T---	G---	H---	D---	L---	LDAFVWVKPGGEC	SDG	--	T---	S-----	-----
XP_360146	---	T---	G---	S---	E---	L---	ADAFVWVKPGGEC	SDG	--	V---	S-----	-----
<i>C. ther</i>	---	T---	G---	H---	E---	L---	VDAFVWVKPGGEC	SDG	--	T---	S-----	-----
XP_002149891	---	T---	G---	D---	P---	L---	EDAFVWVKPGGEC	SDG	--	T---	S-----	-----
XP_001269265	---	T---	G---	D---	E---	L---	VDAFVWVKPGGEC	SDG	--	T---	S-----	-----
<i>H. inso</i>	---	T---	G---	H---	Q---	Y---	VDAFVWVKPGGEC	SDG	--	T---	S-----	-----
XP_001598803	---	T---	G---	D---	A---	L---	EDAFVWVKPGGEC	ADG	--	T---	S-----	-----
AAM76664	---	T---	G---	N---	A---	L---	IDAVVWVKPGGEC	SDG	--	T---	S-----	-----
XP_001933777	---	T---	G---	N---	A---	N---	VDAFVWVKPGGEC	SDG	--	T---	S-----	-----
XP_001796781	---	T---	D---	D---	P---	L---	LDAFVWVKPGGEC	SDG	--	T---	S-----	-----
AA50608	---	T---	G---	S---	S---	L---	IDAVVWVKPGGEC	SDG	--	T---	S-----	-----
AAK28357	---	T---	G---	S---	S---	L---	IDSIVWVKPGGEC	SDG	--	T---	S-----	-----
BAF80327	---	T---	G---	S---	S---	L---	IDAVVWVKPGGEC	SDG	--	T---	S-----	-----
XP_001792324	---	T---	G---	N---	S---	L---	VDALVWVKPGGEC	SDG	--	T---	S-----	-----
BAG48183	---	T---	G---	S---	S---	L---	IDSIVWVKPGGEC	SDG	--	T---	S-----	-----
BAH59082	---	T---	G---	S---	P---	L---	IDSIVWVKPGGEC	SDG	--	T---	S-----	-----
AAT64008	---	T---	G---	S---	S---	L---	IDAVVWVKPGGEC	SDG	--	T---	S-----	-----
<i>P. chrys</i>	---	T---	G---	S---	Q---	F---	IDSIVWVKPGGEC	SDG	--	T---	S-----	-----
BAH59083	---	T---	P---	S---	S---	L---	IDSIVWVKPGGEC	ADG	--	T---	S-----	-----
XP_001833045	---	T---	P---	S---	S---	A---	IDAVVWVKPGGEC	ADG	--	T---	S-----	-----
AAF35251	---	T---	G---	N---	P---	L---	IDAVVWVKPGGEC	SDG	--	T---	S-----	-----
XP_001226566	---	T---	G---	A---	D---	L---	ADAFVWVKPGGEC	SDG	--	T---	S-----	-----
XP_001903893	---	T---	G---	L---	E---	L---	ADAFVWVKPGGEC	SDG	--	T---	S-----	-----
XP_002391276	---	T---	G---	S---	S---	L---	IDSIVWVKPGGEC	SDG	--	T---	S-----	-----
XP_956581	---	T---	G---	H---	D---	L---	ADAFVWVKPGGEC	SDG	--	T---	S-----	-----
XP_362054	---	T---	G---	S---	S---	L---	TDALLVWVKPGGEC	SDG	--	T---	S-----	-----
XP_001806560	---	T---	G---	N---	T---	L---	VDAFVWVKPGGEC	SDG	--	T---	S-----	-----
XP_001931623	---	W---	G---	D---	C---	L---	VDSFVWVKPGGEC	SDG	--	T---	S-----	-----
AA115038	---	T---	G---	S---	S---	L---	IDAVVWVKPGGEC	SDG	--	T---	S-----	-----
AA101211	---	P---	I---	S---	S---	MDYLD	DAFYWKPLGES	SDG	--	T---	S-----	-----
XP_001839612	---	QRILN	---	N---	T---	N---	VDAFVWVKPGGEC	SDG	--	T---	S-----	-----
AA192497	---	S---	G---	M---	D---	Y---	LDAFYWVKPLGES	SDG	--	Y---	S-----	-----
YP_001618727	---	P---	G---	G---	A---	TN	LDAFVWVKPPGDS	SDG	--	V---	G-----	-----
AAM94167	---	S---	G---	M---	D---	Y---	LDAFYWVKPYGES	SDG	--	T---	S-----	-----
XP_001836853	---	Q---	GVLQN	---	P---	N---	VDAFVWVKPGGEC	SDG	--	T---	S-----	-----
AA101212	---	S---	G---	M---	D---	Y---	LDAFYWKPLGES	SDG	--	T---	S-----	-----
XP_001841133	---	L---	D---	S---	E---	N---	VDAFVWVKPGGEC	SDG	--	T---	S-----	-----
XP_001841132	---	Q---	AIV	D---	S---	AN	VDAFVWVKPGGEC	SDG	--	T---	S-----	-----
AAF30749	---	A---	G---	M---	D---	Y---	LDAFYWVKPYGES	SDG	--	T---	S-----	-----
XP_001224490	---	T---	N---	N---	T---	V---	VDALVWVKPGGEC	SDG	--	T---	S-----	-----
XP_368004	---	T---	E---	N---	P---	H---	VDSIVWVKPGGEC	SDG	--	T---	S-----	-----
XP_001903209	---	T---	N---	N---	T---	V---	VDAFVWVKPGGEC	SDG	--	T---	S-----	-----
AA51055	---	P---	P---	D---	S---	SKP	L	LDAYMWIKTPGE	ADG	--	S-----	S-----
XP_001226518	---	T---	G---	N---	N---	N---	VDAFVWVKPGGEC	SDG	--	T---	S-----	-----
XP_957415	---	Q---	SVLQN	---	S---	N---	VDAFVWVKPGGEC	SDG	--	T---	S-----	-----
ABY52798	---	S---	G---	Y---	P---	L---	LDAFMWLKTPGE	ADG	--	S-----	A-----	-----
XP_001795501	---	T---	N---	N---	T---	H---	IDSIVWVKPGGEC	SDG	--	A-----	-----	-----

XP\_001903645 ----T---N---N---P---N--VDAILWVKPGGESDG--T---C-----  
 XP\_001934153 ----Q---N---N---T---N--VDSIVWIKPGGESDG--A-----  
 AAF34679 ----P---D---A---SMP-L-LDAYMWLKTTPGEGADG--S---A-----  
 ABY52797 ----P---S---K---P---L-LDAYMWIKTPGEGADG--S---S-----  
 AAC09228 ----S---G---M---P---L-LDAYMWLKTTPRESGDG--S---S-----  
 Q7SIG5 ----T---N---N---P---N--VDAIVVWVKPGGESDG-----  
 AAR08200 ----PNP-G---M---P---L-LDAYMWLKTTPGEGADG--S---S-----  
 AAD51054 ----V---N---M---P---L-LDAYMWLKPPEASDG--S---S-----  
 AAB92679 ----P---D---P---NMP-L-LDAYVWIKTPGEGADG-----  
 AAB92678 ----P---N---P---GSMPL-LDAYMWIKTPGEGADG--S---S-----  
 AAQ09256 ----A---G---M---P---L-LDAYMWLKT--PGES--D---G-----  
 AAD02028 ----P---A---DL---P---G-LDAYVWFKPPGESDG--S---D-----  
 AAC09066 ----V---N---M---P---L-LDAYMWLKPGGDSGDG-----  
 AAP33843 ----A---G---M---P---L-LDAYMWLKTTPGEGADG--S---S-----  
 AA047726 ----P---D---P---EKYPL-LDAYFWLKPPEASDG--S---S-----  
 AAQ09258 ----P---A---NY---P---L-LDAFVWIKTPGEGADG--N---A-----  
 ZP\_01907667 ----T---A---D---P---R-VDAYLWIKVPGESDG--E---C-----  
 NP\_627067 ----T---A---D---P---L-VDAYLWVKRPPGESDG-----  
 ABY52799 ----S---P---DPEKFP--L-LDAYFWLKPPEADG--S---D-----  
 YP\_879613 ----T---A---G---A---H-ADAYLWIKRPPGESDG--S-----  
 NP\_959214 ----T---A---G---A---H-ADAYLWIKRPPGESDG--S-----  
 YP\_002186959 ----T---A---D---P---L-ADAYLWVKRPPGESDG--T---C-----  
 YP\_001848433 ----T---A---G---A---H-ADAYLWIKRPPGESDG--S-----  
 YP\_001132274 ----T---G---N---P---M-VDAFLWVKRPPGESDG-----  
 YP\_890960 ----T---A---G---A---H-ADAYLWVKRPPGESDG--A---C-----  
 ZP\_04688022 ----T---A---D---P---L-VDAYLWVKRPPGESDG-----  
 YP\_642458 ----T---G---H---P---Q-VDAFLWVKRPPGESDG--S---C-----  
 YP\_001468205 ----D---D---D---S---H-QDARLWIKRVGESDG--T-----  
 ZP\_04366710 ----N---D---G---S---G-LDALLWIKLPGESDG-----  
 YP\_001506001 ----T---G---D---S---D-VDAFFWVWIKPGESDG--T-----  
 AAD41097 FIPHR---R---H---C---L-GQTRRRVNPAAKAG-----  
 ZP\_03911849 ----N---D---S---T---G-LDALLWIKLPGESDG--S---C-----  
 ZP\_04475891 ----T---G---H---D---R-VDAFLWIKRPPGESDG-----  
 ZP\_04368446 ----N---D---G---S---G-LDALLWIKRPPGESDG--T---C-----  
 ZP\_03888455 ----T---G---Q---P---R-VDAFLWVKRPPGESDG--T---C-----  
 ZP\_04704893 ----T---G---D---P---L-LDAYLWIKRPPGESDG--S-----  
 NP\_334477 ----T---A---G---A---H-ADAYLWIKRPPGESDG--T-----  
 YP\_924688 ----E---G---D---A---EH-LDAYVWVKPPGESDG--E---C-----  
 ZP\_03867905 ----A---S---G---A---H-RGNL-WIKLPGESDG--S---C-----  
 ZP\_04747451 ----T---A---G---A---H-ADAYLWVKRPPGESDG--S---C-----  
 YP\_002196808 ----T---G---D---P---L-VDAYLWVKRPPGESDG--D-----  
 NP\_826394 ----T---A---D---P---L-VDAYVWVKRPPGESDG--T-----  
 YP\_001511422 ----T---G---S---P---H-ADAFWVWIKTPGASDG-----  
 YP\_956608 ----T---G---N---P---M-VDAFLWVKRPPGESDG-----  
 NP\_778753 ----P---G---G---H---VDAFQWIKPPPGYSDG--S---SSLIP-----  
 AAC60491 ----T---G---D---A---N-IDAYLWVKPPGEGADG-----  
 NP\_298556 ----P---G---G---H---VDAFQWIKPPPGYSDG--S---SSLIP-----  
 YP\_715139 ----T---G---Y---P---H-ADAFWVWIKTPGASDG-----  
 NP\_522144 ----P---A---G---H---LDAFVWIKPAGEADG--S---SRLIP-----  
 YP\_002208806 ----T---A---D---P---L-VDAYLWVKRPPGESDG-----  
 ZP\_02961057 ----D---I---S---A-TVKTAWIKPPGESDG-----  
 YP\_001912004 ----P---G---G---H---IHAYVWVGKGGGESDG--S---SKYIP-----  
 NP\_638880 ----P---G---G---H---IHALVWVGKPPGESDG--A---S-----  
 AAC09068 ----V---K---M---P---L-LDAYMWLKTTPGEGADG--S---D-----  
 ZP\_04704563 ----T---G---D---D---L-VDAYLWIKRPPGESDG-----  
 YP\_001705499 ----T---A---S---P---N-ADAYLWVKRPPGESDG--S---C-----  
 P07984 ----N---D---G---S---G-LDALLWVKLPGESDG-----  
 NP\_821732 ----S---P---D---S---H-LDAFLWIKPPGESDG--A---SSDIP-----  
 ZP\_00651218 ----P---G---G---H---VDAFQWIKPPPGYSDG--S---SSMIP-----  
 YP\_872374 ----P---T---DFPN-A---H-LDAYVWVKPPGESDG--T---S-----  
 YP\_431432 ----PFG-A---G---A---P-VDAFVWIKPPGESDG--T---S-----  
 AAP42880 ----T---G---D---P---A-IDAYLWIKRPPGESDG--T---C-----  
 ZP\_04371095 ----T---G---D---A---A-VDAFLWIKPPGEGADG--C---A-----  
 YP\_001544904 ----A---P---V---S---G-IDAYVWVKPPGESDG--V---A-----  
 YP\_715138 ----T---G---N---P---R-ADAYLWIKVPGESDG--A---C-----  
 YP\_001826214 ----T---G---D---E---L-VDAYLWVKRPPGESDG--D-----  
 ZP\_04696280 ----T---G---D---E---R-LDAYLWIKRPPGDSGDG--T---C-----  
 P50401 ----P---S---GYAA-S---H-LDAFVWIKPPGESDG--A---STDIP-----  
 ZP\_03818362 ----T---S---D---S---Q-IDAFWVWIKVGEADG-----  
 NP\_625732 ----T---G---M---G---R-VDAYLWVKLPGESDG-----  
 ZP\_00944606 ----P---G---G---H---LDAVLWIKPPGESDG--S---SRLIP-----  
 ZP\_03817628 ----T---G---D---P---Q-IDAFWVWIKPGEADG-----  
 ZP\_04475504 ----P---A---A---G-LDAYVWVKPPGESDG--A---SKLIP-----  
 ZP\_04331392 ----T---G---N---P---L-IDAFIWTWIKLPGEADG-----  
 2BOE-X ----T---G---D---P---M-IDAFWVWIKLPGEADG-----  
 YP\_001828440 ----A---P---E---P---G-IDAYVWAKPPGESDG--S---S-----  
 ACR14000 ----P---A---P---G---VDAFVWVKPPGESDG--I---S-----  
 YP\_001823637 ----T---G---D---D---R-LDAYLWIKRPPGDSGDG--T---C-----  
 ZP\_01462143 ----P---A---GYTA-S---R-LDAFVWIKPPGESDG--A---SKEIP-----  
 ZP\_04367954 ----P---A---GAPSAS--H-LDAYVWVKPPGESDG--A---SKEIP-----  
 ZP\_04376648 ----S---G---D---H---V-I-ADVWIKPPGESDG--D---Y-----

YP_001902424	----P---G-----A---H--LDAYAWVKPPGESDGS--S---STLIP-----
ZP_03868469	----P---T---DFPN-S---H--LDAYVWVKPPGESDGS--S---SSEIP-----
ZP_04608509	----T---G---D---S---A--IAAYLWVKLPGEADG-----
YP_002208593	----P---A---T---G-----VDAYVWVKPPGESDGS--N---SAAVP-----
ZP_03912473	----S---G---Y---P---H--IDAFVWVKPPGESDGS--S---STEIP-----
ZP_04704085	----A---P---E---P---G--IDAYVWAKPPGESDGS--S---S-----
ZP_04707051	----T---G---E---A---R--IDAYLWVKLPGESDGS--C---S-----
ZP_04482497	----A---P---E---S---G--IDAYVWVKPPGESDGS--S---SKEIP-----
BAB83928	----P---A-----S---G--IDAYTWIKPPGESDGS--N---SAPVD-----
ZP_04684292	----G---G---A---E---M--L---LWIKVPGESDGS--NCGVGA-----
YP_002196127	----P---E---E---A---G--IDAYLWVKPPGESDGS--S---SSEIP-----
P33682	----G---G---A---E---M--L---LWIKTPGESDGS--NCGVGS-----
NP_823029	----P---A-----T---G--IDAYVWVKPPGESDGS--S---S-----
NP_828072	----T---G---E---A---G--IDAYLWVKLPGESDGS-----
P26414	----T---A---D---P---A--IDAFWLWIKPPGEADG-----
YP_925799	----V---F---D---G---A--FDGTLWVKHPGESDGS--T---K-----
ZP_04370960	----G---A---S---N---H--IIAYVWIKPPGESDGS--D---Y-----
ZP_04378286	----T---S---A---P---H--ADAYLWIKHPGLSDGS--T-----
ACR12723	----P---S---P---G-----IDAYVWVKPPGESDGS--V---S-----
ZP_04333277	----A---P---E---A---G--IDAYVWVKPPGESDGS--S---S-----
ZP_04483608	----V---G---E---G---A--QDARIWIKVPGESDGS-----
ZP_03390825	----T---T---G---A---H--ADAYLWINGGGGSDGS--E---C-----
ZP_04377655	----T---T---G---A---H--ADAYLWINGGGGSDGS--E---C-----
YP_001108158	----G---G---A---D---L--L---VWAKVPGESDGS-----
ZP_04377654	----T---T---G---A---H--ADAYLWINGGGGSDGS--E---C-----
ZP_04058051	----T---T---G---A---H--ADAYLWINGGGGSDGS--E---C-----
YP_002203140	----P---A-----A---G--IDAYVWIKPPGESDGS--A---SSAVP-----
YP_632998	----V---G---G---A---E--MT--LWIKVPGESDGS-----
ZP_04684732	----A---P---E---P---G--IDAYVWVKPPGESDGS--S---SKAID-----
ZP_04605447	----A---P---E---P---G--IDAYVWVKPPGESDGS--S---SKEIP-----
YP_288681	----P---A---P---G-----VDAYVWVKPPGESDGS--A---S-----
AAF75786	----P---A---P---G-----VDAYVWVKPPGESDGS--A---S-----
YP_001538764	----A---P---E---P---G--IDAYVWVKPPGESDGS--S---S-----
YP_527744	----A---P---E---P---G--VDAYVWVKPPGESDGS--I---SDPNFPIDPNPAK
NP_630629	----P---A-----A---G--IDAYVWVKPPGESDGS--S---SKLID-----
ZP_04032911	----T---G---H---R---L--VDAYLWVKRPPGESDGSQCT---R-----
YP_001160411	----A---P---E---P---G--IDAYVWVKPPGESDGS--S---S-----
YP_001982934	----P---A-----A---G--IDAYVWVKPPGESDGS--V---G-----
ZP_03817603	----P---K---P---N-----IDAYVWIKPPGESDGS--S---STQIP-----
XP_002396053	----T---G---S---S---L--IDAYVWAKPPGESDGS--T---S-----
ZP_04371165	----T---G---V---P---L--VDADLFIKTIGESDGSCT---R-----
AAD54679	----V---D---D---P---L--VAGYLWIKVPGESDGS--Q---CY-----
ZP_03911156	----T---G---D---P---L--VAAYLWIKVPGESDGS--L-----
YP_871896	----T---G---V---P---L--LDAYLWVKIPGESDGS--Q---C-----