

# Systematic Engineering of *Saccharomyces cerevisiae* for the De Novo Biosynthesis of Genistein and Glycosylation Derivatives

Yongtong Wang <sup>1,2,3</sup>, Zhiqiang Xiao <sup>1,2,3</sup>, Siqi Zhang <sup>1,2,3</sup>, Xinjia Tan <sup>1,2,3</sup>, Yifei Zhao <sup>1,2,3</sup>,  
Juan Liu <sup>1,2,3,4</sup>, Ning Jiang <sup>2,3</sup> and Yang Shan <sup>1,2,3,\*</sup>

<sup>1</sup> Longping Branch, College of Biology, Hunan University, Changsha 410125, China; quarknorthstar@163.com (Y.W.); xiaozhiqiang@hnu.edu.cn (Z.X.); 18861822835@163.com (S.Z.); tanxinjia0209@163.com (X.T.); zhaoyifei@hnu.edu.cn (Y.Z.); liujmax2019@163.com (J.L.)

<sup>2</sup> Agriculture Product Processing Institute, Hunan Academy of Agricultural Sciences, Changsha 410125, China; jiangning0806@outlook.com

<sup>3</sup> Hunan Key Lab of Fruits & Vegetables Storage, Processing, Quality and Safety, Hunan Agricultural Products Processing Institute, Changsha 410125, China

<sup>4</sup> Department of Life Sciences, Chalmers University of Technology, SE 412 96 Gothenburg, Sweden

\* Correspondence: sy6302@sohu.com; Tel.: +86-0731-84691289; Fax: +86-0731-82873368

**Table S1 Strains used in this study.**

Strains	Parent strain	Description	Sources
<i>S. cerevisiae</i>			
CENPK2-1D	CENPK2-1D	<i>MAT<math>\alpha</math></i> ; <i>ura3-52</i> ; <i>trp1-289</i> ; <i>leu2-3,112</i> ; <i>his3<math>\Delta</math>1</i> ; <i>MAL2-8C</i> ; <i>SUC2</i>	Laboratory
<i>S. cerevisiae</i>	CENPK2-1D	<i>gal80::G418</i> , <i>CENPK2-1D</i>	Laboratory
C800			
C04	C800	<i>XI-5::</i> ( <i>GAL10p-AtC4H-CYC1t</i> )+( <i>GAL10p-AtPAL2-CIT2t</i> )+( <i>THD3p-AtATR2-SDH1t</i> )+( <i>SED1p-CYB5-PDB1t</i> )	This study
C05	C800	<i>XI-5::</i> ( <i>GAL10p-AtC4H-CYC1t</i> )+( <i>THD3p-AtATR2-SDH1t</i> )+( <i>SED1p-CYB5-PDB1t</i> ); <i>XI-2::</i> ( <i>GAL7p-FjTAL-DIT1t</i> )	This study
C06	C800	<i>XI-5::</i> ( <i>GAL10p-AtC4H-CYC1t</i> )+( <i>GAL10p-AtPAL2-CIT2t</i> )+( <i>THD3p-AtATR2-SDH1t</i> )+( <i>SED1p-CYB5-PDB1t</i> ); <i>XI-2::</i> ( <i>GAL7p-FjTAL-DIT1t</i> )	This study
N001	C06	<i>Ty3::GAL10p-PhCHS-ALD5p-MsCHI-ARO7p-Pc4CL-LEU2p-LEU2</i> , C800	This study
N002	C06	<i>Ty3::GAL10p-PhCHS-ALD5p-MsCHI-ARO7p-Pc4CL-LEU2p-LEU2</i> , C800	This study

N003	C06	<i>Ty3::GAL10p-PhCHS-ALD5p-MsCHI-ARO7p-P</i> <i>c4CL-LEU2p-LEU2, C800</i>	This study
N004	C06	<i>Ty3::GAL10p-PhCHS-ALD5p-MsCHI-ARO7p-P</i> <i>c4CL-LEU2p-LEU2, C800</i>	This study
N005	C06	<i>Ty3::GAL10p-PhCHS-ALD5p-MsCHI-ARO7p-P</i> <i>c4CL-LEU2p-LEU2, C800</i>	This study
N006	C06	<i>Ty3::GAL10p-PhCHS-ALD5p-MsCHI-ARO7p-P</i> <i>c4CL-LEU2p-LEU2, C800</i>	This study
G01	C800	<i>pY26-SED1p-GeHIS-Ter22-GmCPR-TDH1p-IN</i> <i>O1p-GmHID-CYC1t</i>	This study
G02	C800	<i>pY26-SED1p-GmIFS-Ter22-GmCPR-TDH1p-IN</i> <i>O1p-GmHID-CYC1t</i>	This study
G03	C800	<i>pY26-SED1p-GmIFS2-Ter22-GmCPR-TDH1p-I</i> <i>NO1p-GmHID-CYC1t</i>	This study
G04	C800	<i>pY26-SED1p-GuIFS-Ter22-GmCPR-TDH1p-IN</i> <i>O1p-GmHID-CYC1t</i>	This study
G05	C800	<i>pY26-SED1p-MtIFS-Ter22-GmCPR-TDH1p-IN</i> <i>O1p-GmHID-CYC1t</i>	This study
G06	C800	<i>pY26-SED1p-PlIFS-Ter22-GmCPR-TDH1p-INO</i> <i>1p-GmHID-CYC1t</i>	This study
G07	C800	<i>pY26-SED1p-TpIFS-Ter22-GmCPR-TDH1p-INO</i> <i>1p-GmHID-CYC1t</i>	This study

G08	C800	<i>pY26-SED1p-GeHIS-Ter22-GmCPR-TDH1p-IN</i> <i>O1p-GeHID-CYC1t</i>	This study
G09	C800	<i>pY26-SED1p-PlIFS-Ter22-GmCPR-TDH1p-INO</i> <i>1p-PlHID-CYC1t</i>	This study
G11	C800	<i>pY26-SED1p-GmIFS-Ter22-CrCPR2-TDH1p-IN</i> <i>O1p-GmHID-CYC1t</i>	This study
G12	C800	<i>pY26-SED1p-GmIFS-Ter22-AtCPR-TDH1p-INO</i> <i>1p-GmHID-CYC1t</i>	This study
G13	C800	<i>pY26-SED1p-GmIFS-Ter22-LjCPR-TDH1p-INO</i> <i>1p-GmHID-CYC1t</i>	This study
G16	N006	<i>pY26-SED1p-GmIFS-Ter22-AtCPR-TDH1p-INO</i> <i>1p-GmHID-CYC1t</i>	This study
G17	N006	<i>EXG1::SED1p-GmIFS-Ter22-AtCPR-TDH1p-IN</i> <i>O1p-GmHID-CYC1t;</i>	This study
G18	N006	<i>pY26-SED1p-GmIFS-Ter22-AtCPR-TDH1p-INO</i> <i>1p-GmHID-CYC1t; PRS424-SED1p-STB5-Ter22</i>	This study
G19	N006	<i>pY26-SED1p-GmIFS-Ter22-AtCPR-TDH1p-INO</i> <i>1p-GmHID-CYC1t; pRS424-SED1p-YEF1-Ter22</i>	This study
G20	N006	<i>pY26-SED1p-GmIFS-Ter22-AtCPR-TDH1p-INO</i> <i>1p-GmHID-CYC1t;</i> <i>pRS424-SED1p-EcPntAB-Ter22</i>	This study
G21	N006	<i>pY26-SED1p-GmIFS-Ter22-AtCPR-TDH1p-INO</i>	This

		<i>1p-GmHID-CYC1t;</i>	study
		<i>pRS424-SED1p-ALD6-TEr22</i>	
		<i>pY26-SED1p-GmIFS-Ter22-AtCPR-TDH1p-INO</i>	This
G22	N006	<i>1p-GmHID-CYC1t;</i>	study
		<i>pRS424-SED1p-HEM2-TEr22-HEM3-TDH3p</i>	
<hr/>			
		<i>oca5Δ;</i>	This
G23	N006	<i>pY26-SED1p-GmIFS-Ter22-AtCPR-TDH1p-INO</i>	study
		<i>1p-GmHID-CYC1t</i>	
NHG 12	GO06	<i>TY1::SED1p-GmIFS-Ter22-AtCPR-TDH1p-INO</i>	This
		<i>1p-GmHID-CYC1t</i>	study
NHG 34	GO06	<i>TY1::SED1p-GmIFS-Ter22-AtCPR-TDH1p-INO</i>	This
		<i>1p-GmHID-CYC1t</i>	study
NHG 58	GO06	<i>TY1::SED1p-GmIFS-Ter22-AtCPR-TDH1p-INO</i>	This
		<i>1p-GmHID-CYC1t</i>	study
NHGO01	NHG12	<i>exg1::SED1p-HEM2-TEr22-HEM3-TDH3p</i>	This study
NHGO02	NHGO01	<i>oca5Δ,</i>	This study
NHGO03	NHGO01	<i>fas1::(ACC1p-FAS1)</i>	This study
NHGO04	NHGO01	<i>fas1::(ZWF1p-FAS1)</i>	This study

NHGO05	NHGO01	<i>fas1::(RPL3p-FAS1)</i>	This study
NHGO06	NHGO01	<i>fas1::(TEF1p-FAS1)</i>	This study
NHGO07	NHGO01	<i>fas1::(ITR1p-FAS1)</i>	This study
NHGO08	NHGO01	<i>fas1::(CCW12p-FAS1)</i>	This study
NHGO10	NHGO08	<i>pY26-SED1p-GmUGT4-DIT1t</i>	This study
NHGO11	NHGO08	<i>pY26-SED1p-PIUGT43-DIT1t</i>	This study
NHGO12	NHGO08	<i>pY26-SED1p-GmUGT4-DIT1t;</i> <i>pRS424-TDH1p-PGM1-ADH1t</i>	This study
NHGO13	NHGO08	<i>pY26-SED1p-GmUGT4-DIT1t;</i> <i>pRS424-TDH1p-PGM2-ADH1t</i>	This study
NHGO14	NHGO08	<i>pRS424-TDH1p-PGM1-ADH1t-INO1p-UGP1-C</i> <i>YC1t</i>	This study
NHGO15	NHGO08	<i>pY26-SED1p-GmUGT4-DIT1t;</i> <i>pRS424-TDH1p-PGM2-ADH1t-INO1p-UGP1-C</i> <i>YC1t</i>	This study

NHGO16	NHGO08	<i>pY26-SED1p-PIUGT43-DIT1t;</i> <i>pRS424-TDH1p-PGM1-ADH1t</i>	This study
NHGO17	NHGO08	<i>pY26-SED1p-PIUGT43-DIT1t;</i> <i>pRS424-TDH1p-PGM2-ADH1t</i>	This study
NHGO18	NHGO08	<i>pY26-SED1p-PIUGT43-DIT1t;</i> <i>pRS424-TDH1p-PGM1-ADH1t-INO1p-UGP1-C</i> <i>YC1t</i>	This study
NHGO19	NHGO08	<i>pY26-SED1p-PIUGT43-DIT1t;</i> <i>pRS424-TDH1p-PGM2-ADH1t-INO1p-UGP1-C</i> <i>YC1t</i>	This study

---

**Table S2 Genes used in the present study**

<b>Gene</b>	<b>Original organisms</b>	<b>EC numbers</b>
<i>AtC4H</i>	<i>Arabidopsis thaliana</i>	1.14.14.91
<i>AtPAL2</i>	<i>Arabidopsis thaliana</i>	4.3.1.24
<i>AtATR2</i>	<i>Arabidopsis thaliana</i>	1.6.2.4
<i>FjTAL</i>	<i>Flavobacterium johnsoniae</i>	2.2.1.2
<i>PhCHS</i>	<i>Petunia hybrida</i>	2.3.1.74
<i>MsCHI</i>	<i>Medicago sativa</i>	5.5.1.6
<i>Pc4CL</i>	<i>Petroselinum crispum</i>	6.2.1.12
<i>GeHIS</i>	<i>Glycyrrhiza echinata</i>	1.14.14.87
<i>GmIFS</i>	<i>Glycine max</i>	1.14.14.87
<i>GmIFS2</i>	<i>Glycine max</i>	1.14.14.87
<i>GuIFS</i>	<i>Glycyrrhiza uralensis</i>	1.14.14.87
<i>MtIFS</i>	<i>Medicago truncatula</i>	1.14.14.87
<i>PlIFS</i>	<i>Pueraria lobata</i>	1.14.14.87
<i>TpIFS</i>	<i>Trifolium patense</i>	1.14.14.87
<i>GmHID</i>	<i>Glycine max</i>	4.2.1.105
<i>GeHID</i>	<i>Glycyrrhiza echinata</i>	4.2.1.105
<i>PlHID</i>	<i>Pueraria lobata</i>	4.2.1.105
<i>GmCPR</i>	<i>Glycine max</i>	1.6.2.4
<i>CrCPR2</i>	<i>Catharanthus roseus</i>	1.6.2.4



---

<i>AtCPR</i>	<i>Arabidopsis thaliana</i>	1.6.2.4
<i>LjCPR</i>	<i>Lotus japonicus</i>	1.6.2.4
<i>EcPntAB</i>	<i>Escherichia coli</i>	7.1.1.1
<i>GmUGT4</i>	<i>Glycine max</i>	2.4.1.170
<i>PlUGT43</i>	<i>Pueraria lobata</i>	2.4.1.17
<i>STB5</i>	<i>Saccharomyces cerevisiae</i>	7.3.2.1
<i>YEF1</i>	<i>Saccharomyces cerevisiae</i>	2.7.1.23
<i>ALD6</i>	<i>Saccharomyces cerevisiae</i>	1.2.1.3
<i>HEM2</i>	<i>Saccharomyces cerevisiae</i>	4.2.1.24
<i>HEM3</i>	<i>Saccharomyces cerevisiae</i>	2.5.1.61
<i>PGM1</i>	<i>Saccharomyces cerevisiae</i>	5.4.2.2
<i>PGM2</i>	<i>Saccharomyces cerevisiae</i>	5.4.2.7
<i>UGP1</i>	<i>Saccharomyces cerevisiae</i>	2.7.7.9
<i>CYB5</i>	<i>Saccharomyces cerevisiae</i>	1.6.2.2

---

**Table S3 Primers used in the present study**

Primers	Nucleotide sequence (5' - 3')	Application
<b>Construction of the synthesis pathway for p-coumaric acid</b>		
<i>GAL10p-1-F</i>	ACAACATGGCTTCAATTTGATCCATT	Amplification of <i>GAL10p</i> promoter
	ATAGTTTTTTCTCCTTGACGTTAAAGT	
	ATAG	
<i>GAL10p-1-R</i>	AGACTTTTCCAACAACAACAAATCC	Amplification of <i>GAL10p</i> promoter
	ATTTATATTGAATTTTCAAAAATTCTT	
	ACTTTTT	
<i>AtC4H-F</i>	ATCGACAAAGGAAAAGGGGCCTGTT	Amplification of <i>AtC4H</i> fragment
	TAACAGTTTCTTGGCTTCATAACAAT	
	AATAGAG	
<i>AtC4H-R</i>	ATGGATTTGTTGTTGTTGGAAAAGTC	Amplification of <i>AtC4H</i> fragment
	TTTG	
<i>CYC1t-F</i>	ACAGGCCCTTTTCCTTTGTCGATA	Amplification of <i>CYC1t</i> terminator
<i>CYC1t-R</i>	GCAAATTAAAGCCTTCGAGCGTCC	Amplification of <i>CYC1t</i> terminator
<i>AtPAL2-F</i>	ATGGATCAAATTGAAGCCATGTTGTG	Amplification of <i>AtPAL2</i> fragment
	TG	

---

<i>AtPAL2-R</i>	GTAAAAGTAGGATGTAATCCAAATTT	Amplification of
	AACAAATTGGAATTGGAGCACCG	<i>AtPAL2</i> fragment
<i>CIT2t-F</i>	TTGGATTACATCCTACTTTTACACCC	Amplification of <i>CIC2t</i>
		terminator
<i>CIT2t-R</i>	GCCGGAATAGTGCAAATTGTATG	Amplification of <i>CIC2t</i>
		terminator
<i>THD3p-F</i>	AGCAATCAGTGGTGTATGGAAGGTA	Amplification of <i>TDH3p</i>
	GAATAAAAAACACGCTTTTTCAGTTC	promoter
<i>THD3p-R</i>	GAG	
	CAACTTGTAGTTCATCATGTTCTTCAT	Amplification of <i>TDH3p</i>
<i>THD3p-R</i>	TTTGTTTGTTTATGTGTGTTTATTCTGA	promoter
	AAC	
<i>AtATR2-F</i>	TTACCAAACATCTCTCAAGTATCTAC	Amplification of
	CG	<i>AtATR2</i> fragment
<i>AtATR2-R</i>	ATGAAGAACATGATGAACTACAAGT	Amplification of
	TG	<i>AtATR2</i> fragment
<i>SDH1t-F</i>	GGTAGATACTTGAGAGATGTTTGGTA	Amplification of <i>SDH1t</i>
	ATTGGAACCTCATTGTATTTTACGG	terminator
<i>SDH1t-R</i>	GATTCATACAATTTGCACTATTCCGG	Amplification of <i>SDH1t</i>
	CGACAGCACCTTGTACAGCAATTC	terminator
<i>SED1p-F</i>	TCTACCTTCCATACACCACTGATTG	Amplification of <i>SED1p</i>

---

---

		promoter
<i>SED1p-R</i>	CATCTTAATAGAGCGAACGTATTTTA TTTTG	Amplification of <i>SED1p</i> promoter
<i>CYB5-F</i>	AATAAAATACGTTCGCTCTATTAAGA TGATGCCTAAAGTTTACAGTTACCAA G	Amplification of <i>CYB5</i> fragment
<i>CYB5-R</i>	CTCTCCTTCCTATTGGATTGAAGTTTT ATTCGTTCAACAAATAATAAGCAAC AC	Amplification of <i>CYB5</i> fragment
<i>PDB1t-F</i>	AACTTCAATCCAATAGGAAGGAGAG	Amplification of <i>PDB1t</i> terminator
<i>PDB1t-R</i>	ATGAAGCAGCCAATACCTGTTTTACA ACTTTCGCTTAGGTTCCATGCC	Amplification of <i>PDB1t</i> terminator
<i>FjTAL-F</i>	ATGAACACCATTAATGAATACTTGAG	Amplification of <i>FjTAL</i> fragment
<i>FjTAL-R</i>	TTAATTGTTAATCAAATGATCCTTAA CC	Amplification of <i>FjTAL</i> fragment
<i>GAL7p-F</i>	TTTGCCAGCTTACTATCCTTCTTG	Amplification of <i>GAL7p</i> promoter
<i>GAL7p-R</i>	CTCAAGTATTCATTAATGGTGTTTCATT TTTGAGGGAATATTCAACTGTTTTTTT	Amplification of <i>GAL7p</i> promoter

---

---

<i>DIT1t-F</i>	TTAAGGATCATTGATTAACAATTAA	Amplification of <i>DIT1t</i>
	ATAAAGTAAGAGCGCTACATTGG	terminator
<i>DIT1t-R</i>	ACTTGTTACTCCGCAACGC	Amplification of <i>DIT1t</i>
		terminator
<b>Construction of the synthesis pathway for naringenin</b>		
<i>GAL10p-2-F</i>	TTATATTGAATTTTCAAAAATTCTTACT	Amplification of
	TTTTTTTTGGATGGACGR	<i>GAL10p</i> promoter
<i>GAL10p-2-R</i>	CATTGTACTTCAATCTATAGTTTTTCT	Amplification of
	CCTTGACGTTAAAGTATAGAG	<i>GAL10p</i> promoter
<i>PhCHS-F</i>	CAAACCTCTGGCGAAGAATTGTTAATT	Amplification of <i>PhCHS</i>
	AAAGATCTCCGCGGTCCG	fragment
<i>PhCHS-R</i>	GAATTTTTGAAAATTCAATATAAATGG	Amplification of <i>PhCHS</i>
	TTACGGTGGAAGAATACCGC	fragment
<i>ALD5p-F</i>	CTATAGATTGAAGTACAATGACGCTAA	Amplification of <i>ALD5p</i>
	CACTAAGTTATG	promoter
<i>ALD5p-R</i>	TTCTTCTTCTTTGGCTTTTGTTAATGTT	Amplification of <i>ALD6p</i>
	GTG	promoter
<i>MsCHI-F</i>	CAAAAAGCCAAAGAAGAAGAAATGGC	Amplification of <i>MsCHI</i>
	AGCAAGCATTACGGC	fragment
<i>MsCHI-R</i>	GACTGAATCAAATGTAATCCACCTAGG	Amplification of <i>MsCHI</i>
	TCAGTTACCGATTTTAAAGGC	fragment

---

---

<i>ARO7p-F</i>	TGGATTACATTTGATTCAGTCATACAC	Amplification of <i>ARO7p</i>
	GAATTATG	promoter
<i>ARO7p-R</i>	ATCTTATACCAATTTTATGCAGGATGCT	Amplification of <i>ARO7p</i>
	GAGTG	promoter
<i>Pc4CL-F</i>	GCATAAAATTGGTATAAGATATGGGTG	Amplification of <i>Pc4CL</i>
	ACTGCGTTGCCCC	fragment
<i>Pc4CL-R</i>	TCGACGGTATCGATAAGCTTTTACTTC	Amplification of <i>Pc4CL</i>
	GGCAGGTCGCCGC	fragment
<i>LEU2-F</i>	TTCGCCCTATAGTGAGTCGTATTACA	Amplification of <i>LUE2</i>
	ACTGTGGGAATACTCAGGTATCGTAA	fragment
	G	
<i>LEU2-R</i>	TTAAGCAAGGATTTTCTTAACCTTCTTC	Amplification of <i>LUE2</i>
	GGC	fragment
<b>Construction of the synthesis pathway for genistein</b>		
<i>RPY26-F</i>	TCTAGAACTAGTGGATCCCCCGGG	Amplification of pY26
		vector
<i>RPY26-R</i>	GAGCTCCAGCTTTTGTTCCTTTTAGT	Amplification of pY26
		vector
<i>SED1p-F</i>	CTAAAGGGAACAAAAGCTGGAGCTC	Amplification of <i>SED1p</i>
	AATTATCTCCCAGACGGCACCG	promoter
<i>SED1p-R</i>	CTTAATAGAGCGAACGTATTTTATTT	Amplification of <i>SED1p</i>
	TGCTTGTC	promoter

---

---

<i>Ter22-F</i>	AATGTCAAAAGCCTCAAGGTGCC	Amplification of <i>Ter22</i> terminator
<i>Ter22-R</i>	ATAGGTTGGCTTCCATGTTGGCTAT	Amplification of <i>Ter22</i> terminator
<i>TDH1p-F</i>	TTTGTTTTGTGTGTAAATTTAGTGAAG TACTG	Amplification of <i>TDH1p</i> promoter
<i>TDH1p-R</i>	GCACCGGCCTCATCGTCTTCATGTCC CACCAGCCAACACTTG	Amplification of <i>TDH1p</i> promoter
<i>INO1p-F</i>	GAAGACGATGAGGCCGGTGC	Amplification of <i>INO1p</i> promoter
<i>INO1p(Gm)-R</i>	CAATTCCTTAACGATTTCCTTAGCCA TTGTTACTTCTTTTTCACTGGAAAAAA AAGGG	Amplification of <i>INO1p</i> promoter
<i>INO2p(Ge)-R</i>	GTGGTAGTAGAAGTAGAAGAAGCCA TTGTTACTTCTTTTTCACTGGAAAAAA AAGGG	Amplification of <i>INO1p</i> promoter
<i>INO1p(Pl)-R</i>	TTCCTTATTAGAGTTTTCGTTAGCCAT TGTTACTTCTTTTTCACTGGAAAAAA AAGGG	Amplification of <i>INO1p</i> promoter
<i>GeHIS-F</i>	CAAAATAAAATACGTTCGCTCTATTA AGATGTTGGTTGAATTGGCTATCACT TTGTTAG	Amplification of <i>GeHIS</i> fragment

---

---

	GCACCTTGAGGCTTTTGACATTTTAT	
<i>GeHIS-R</i>	GAAGAAAACAATTTTGGAAGTCTGCTG AAGA	Amplification of <i>GeHIS</i> fragment
	GCAAAATAAAATACGTTTCGCTCTATT	
<i>GmIFS-F</i>	AAGATGTTGTTAGAATTAGCTTTGGG TTTATTC	Amplification of <i>GmIFS</i> fragment
	GGCACCTTGAGGCTTTTGACATTTTA	
<i>GmIFS-R</i>	GGACAACAATTTAGAAGCAACACCA AT	Amplification of <i>GmIFS</i> fragment
	GCAAAATAAAATACGTTTCGCTCTATT	
<i>GmIFS2-F</i>	AAGATGTTGTTGGAATTGGCTTTGGG T	Amplification of <i>Gm2IFS</i> fragment
	GCACCTTGAGGCTTTTGACATTTTAA	
<i>GmIFS2-R</i>	GACAATAACTTAGAAGCAACACCAA TTCTAGC	Amplification of <i>Gm2IFS</i> fragment
	CAAAATAAAATACGTTTCGCTCTATTA	
<i>GuIFS-F</i>	AGATGTTGGTTGAATTGGCTATCACC TTG	Amplification of <i>GuIFS</i> fragment
	GCACCTTGAGGCTTTTGACATTTTAT	
<i>GuIFS-R</i>	GAAGAGAACAACCTTTGGGACAGC	Amplification of <i>GuIFS</i> fragment



<i>MtIFS-F</i>	GCAAAATAAAATACGTTCGCTCTATT	Amplification of <i>MtIFS</i> fragment
	AAGATGTTGGTTGAATTAGCCGTTAC	
	TTTGTT	
<hr/>		
<i>MtIFS-R</i>	GCACCTTGAGGCTTTTGACATTTTAG	Amplification of <i>MtIFS</i> fragment
	GAAGATAACAACCTTATCAGCAACAC	
	C	
<i>PlIFS-F</i>	CAAAATAAAATACGTTCGCTCTATTA	Amplification of <i>PlIFS</i> fragment
	AGATGTTATTAGAATTAGCTTTAGGT	
	TTGTTAG	
<i>PlIFS-R</i>	GCACCTTGAGGCTTTTGACATTTTAA	Amplification of <i>PlIFS</i> fragment
	GACAAAACCTTAGCAGCGACG	
<i>TpIFS-F</i>	GCAAAATAAAATACGTTCGCTCTATT	Amplification of <i>TpIFS</i> fragment
	AAGATGTTGTTGGAATTGGCCTTGGG	
<i>TpIFS-R</i>	GCACCTTGAGGCTTTTGACATTTTAA	Amplification of <i>TpIFS</i> fragment
	GACAACAACCTTAGAAGCAACACCAA	
	TTC	
<i>GmCPR-F</i>	ATAGCCAACATGGAAGCCAACCTAT	Amplification of <i>GmCPR</i> fragment
	TTACCAAACATCTCTCAAGTATCTAC	
	CAGTAG	
<i>GmCPR-R</i>	CAGTACTTCACTAAATTTACACACAA	Amplification of <i>GmCPR</i> fragment
	AACAAAATGCAAGATTCTGGTTCCAT	

GAAGATTTC

<i>AtCPR-F</i>	ATAGCCAACATGGAAGCCAACCTAT	Amplification of <i>AtCPR</i>
	TTACCAAACGTCACGCAGGTAAC	fragment
<hr/>		
<i>AtCPR-R</i>	CAGTACTTCACTAAATTTACACACAA	Amplification of <i>AtCPR</i>
	AACAAAATGAGCTCTAGCTCTTCTTC	fragment
<i>CrCPR-F</i>	TTCTACC	
	ATAGCCAACATGGAAGCCAACCTAT	Amplification of <i>CrCPR</i>
<i>CrCPR-R</i>	TTACCAAACATCTCTTAAGTATCTAC	fragment
	CTGTC	
<i>LjCPR-F</i>	GTACTTCACTAAATTTACACACAAAA	Amplification of <i>LjCPR</i>
	CAAAATGGATTCTTCATCTGAAAAGT	fragment
<i>LjCPR-R</i>	TGTCACC	
	ATAGCCAACATGGAAGCCAACCTAT	Amplification of <i>LjCPR</i>
<i>GmHID-F</i>	TTACCAGACATCCCTCAAGTACCTAC	fragment
	C	
<i>GmHID-R</i>	GTACTTCACTAAATTTACACACAAAA	Amplification of <i>GmHID</i>
	CAAAATGGAGGAATCCTCTTCTATGA	fragment
<i>GmHID-F</i>	AGATCAG	
	ATGGCTAAGGAAATCGTTAAGGAAT	Amplification of
<i>GmHID-R</i>	TGTTG	<i>GmHID</i> fragment

<i>GmHID-R</i>	CCGGGGGATCCACTAGTTCTAGATTA AACCAAAAATGAAGCTAATCTCTTG ATCATTG	Amplification of <i>GmHID</i> fragment
<i>GeHID-F</i>	ATGGCTTCTTCTACTTCTACTACCACT TC	Amplification of <i>GeHID</i> fragment
<i>GeHID-R</i>	CCGGGGGATCCACTAGTTCTAGATTA AACCAAGAAAGAAGCAATTCTACCA ATCAAA	Amplification of <i>GeHID</i> fragment
<i>PIHID-F</i>	ATGGCTAACGAAAACCTAATAAGG AAATTGTTAAG	Amplification of <i>PIHID</i> fragment
<i>PIHID-R</i>	CGGGGGATCCACTAGTTCTAGATTAA ACCAAGAAGGAAGCTAATCTCTTGA TCAT	Amplification of <i>PIHID</i> fragment
<b>Modification of the genistein pathway</b>		
<i>RPRS424-F</i>	CAGCTTTTGTTCCCTTTAGTGAGGG	Amplification of pRS424 vector
<i>RPRS424-R</i>	CGCGCGCTCACTGGCCGTCG	Amplification of pRS424 vector
<i>SED1p-(PRS42 4)-F</i>	CGACGGCCAGTGAGCGCGCAATTA TCTCCCAGACGGCACCG	Amplification of <i>SED1p</i> promoter
<i>SED1p-R</i>	CTTAATAGAGCGAACGTATTTTATTT TGCTTGTC	Amplification of <i>SED1p</i> promoter

<i>Ter22-F</i>	AATGTCAAAAGCCTCAAGGTGCC	Amplification of <i>Ter22</i> terminator
<i>Ter22-(PRS424)</i>	CCCTCACTAAAGGGAACAAAAGCTG	Amplification of <i>Ter22</i>
<i>-R</i>	ATAGGTTGGCTTCCATGTTGGCTAT	terminator
<hr/>		
<i>STB5-F</i>	GACAAGCAAAATAAAATACGTTCGC TCTATTAAGATGGATGGTCCCAATTT TGCACATC	Amplification of <i>STB5</i> fragment
<i>STB5-R</i>	GGCACCTTGAGGCTTTTGACATTTCA TACAAGTTTATCAACCCAAGAGACG TC	Amplification of <i>STB5</i> fragment
<i>YEF1-F</i>	GCAAAATAAAATACGTTCGCTCTATT AAGATGAAAAGTATAGATTACTGA TTAACGCT	Amplification of <i>YEF1</i> fragment
<i>YEF1-R</i>	GGCACCTTGAGGCTTTTGACATTTTA GATTGCAAAATGAGCCTGACGAG	Amplification of <i>YEF1</i> fragment
<i>EcPntAB-F</i>	GACAAGCAAAATAAAATACGTTCGC TCTATTAAGTGAAGGGAATATCATG CGAATTGG	Amplification of <i>EcPntAB</i> fragment
<i>EcPntAB-R</i>	GGCACCTTGAGGCTTTTGACATTTTA CAGAGCTTTCAGGATTGCATCCAC	Amplification of <i>EcPntAB</i> fragment

ALD6-F	CAAGCAAAATAAAATACGTTTCGCTC	Amplification of <i>ALD6</i> fragment
	TATTAAGATGACTAAGCTACACTTTG	
	ACACTGCT	
ALD6-R	GGCACCTTGAGGCTTTTGACATTTTA	Amplification of <i>ALD6</i> fragment
	CAACTTAATTCTGACAGCTTTTACTTC	
	AGTG	
<hr/>		
HEM2-F	GCAAAATAAAATACGTTTCGCTCTATT	Amplification of <i>HEM2</i> fragment
	AAGATGCATACAGCTGAATTTTTGGA	
	AACAG	
HEM2-R	GGCACCTTGAGGCTTTTGACATTTTA	Amplification of <i>HEM2</i> fragment
	GTTTTCTTCATCTAACCAGTCTAGGA	
	ACTC	
HEM3-F	GCCAACATGGAAGCCAACCTATTCA	Amplification of <i>HEM3</i> fragment
	TTTGATTCTGTCTAAATTAATTTCATC	
	CAG	
HEM3-R	TTAGTTTCGAATAAACACACATAAAC	Amplification of <i>HEM3</i> fragment
	AAACAAAATGGGCCCTGAAACTCTA	
	CATATTGG	
TDH3p-F	TTTGTTTGTTTATGTGTGTTTATTCGA	Amplification of <i>TDH3p</i> promoter
	AACTAAGTTC	
TDH3p-R	CCCTCACTAAAGGGAACAAAAGCTG	Amplification of <i>TDH3p</i> promoter
	TCATTATCAATACTGCCATTTCAAAG	

	AATACG	
	CAATAATCAGAGATTACAGTCGGCA	
<i>ACC1p-F</i>	TCAATTTTCCTTTTCCTTATTCTACTCTT	Amplification of <i>ACC1p</i>
	TTTATC	promoter
	GGTCTTGTGGAGTAAGCGTCCATGGT	
<i>ACC1p-R</i>	AGAAACTTGATTTTTTCTAATTTTCTG	Amplification of <i>ACC1p</i>
	CGC	promoter
<hr/>		
	CAATAATCAGAGATTACAGTCGGCA	
<i>ZWF1p-F</i>	TCAAGTACATAGTGACATTTAAATAA	Amplification of <i>ZWF1p</i>
	TAGCAAG	promoter
	GGTCTTGTGGAGTAAGCGTCCATCTT	
<i>ZWF1p-R</i>	GCCTTATGTGGTTTTCTATTCTATTGG	Amplification of <i>ZWF1p</i>
	ATTTAC	promoter
	CGACAATAATCAGAGATTACAGTCG	
<i>RPL3p-F</i>	GCATGCGAACAAACCAGCTGCTAAG	Amplification of <i>RPL3p</i>
		promoter
	GGTCTTGTGGAGTAAGCGTCCATGAT	
<i>RPL3p-R</i>	TGATTGTTGTAGTAACTGTGTTGTTCT	Amplification of <i>RPL3p</i>
	AGAG	promoter

TEF1p-F	CAATAATCAGAGATTACAGTCGGCA	Amplification of <i>TEF1p</i> promoter
	GGTGTCGTTAATTACCCGTAATAAG	
	G	
TEF1p-R	GTCTTGTGGAGTAAGCGTCCATTTTG	Amplification of <i>TEF1p</i> promoter
	TAATTAAACTTAGATTAGATTGCTA	
	TGCTTTC	
ITR1p-F	CAATAATCAGAGATTACAGTCGGCA	Amplification of <i>ITR1p</i> promoter
	TAACATAATAAGCGCATATACTTAGT	
	TCTCTCC	
<hr/>		
ITR1p-R	GGTCTTGTGGAGTAAGCGTCCATTTT	Amplification of <i>ITR1p</i> promoter
	AACACCCACTGCAGAAACAAAGAA	
	AATG	
CCW12p-F	CAATAATCAGAGATTACAGTCGGCA	Amplification of <i>CCW12p</i> promoter
	AACGCCACCAAAAAAAAAAAAAATA	
	AAAGCCAATC	
CCW12p-R	GGTCTTGTGGAGTAAGCGTCCATTAT	Amplification of <i>CCW12p</i> promoter
	TGATATAGTGTTTAAGCGAATGACAG	
	AAG	
Glycosylation of genistein		
GmUGT4-F	GCAAAATAAAATACGTTCGCTCTATT	Amplification of <i>GmUGT4</i> fragment
	AAGATGACAATGAAGGATTCTATCG	
	TTTTGTAC	

	GACCAATGTAGCGCTCTTACTTTATTT	
<i>GmUGT4-R</i>	AATGTTCTCTCCACAATTCAACCAAC TTATTC	Amplification of <i>GmUGT4</i> fragment
	GCAAAATAAAATACGTTTCGCTCTATT	
<i>PlUGT43-F</i>	AAGATGACTAGATACGAAGTTGTTTT TATTGC	Amplification of <i>PlUGT43</i> fragment
	AATGTAGCGCTCTTACTTTATTTAAG	
<i>PlUGT43-R</i>	ATGTTAATTCTTGAATTAATGAAACC AAATTAG	Amplification of <i>PlUGT43</i> fragment
<hr/>		
<i>DIT1t(PRS424)</i>	CGACGGCCAGTGAGCGCGCGACTTG	Amplification of
-F	TTACTCCGCAACGCTTTTC	<i>GmUGT4/PlUGT43</i> fragment
<i>SED1p(ADH1t)</i>	CCAATTCAGCTGGCGTAATAGCAATT	Amplification of
-R	ATCTCCCAGACGGCACCG	<i>GmUGT4/PlUGT43</i> fragment
<i>SED1p(CYC1t)-</i>	CCAATTCGCCCTATAGTGAGTCGTAT	Amplification of
R	TACAATTATCTCCCAGACGGCACCG	<i>GmUGT4/PlUGT43</i> fragment
<i>DIT1t-F</i>	ATAAAGTAAGAGCGCTACATTGGTCT ACC	Amplification of <i>DIT1t</i> terminator
<i>DIT1t-R</i>	CACGACGTTGTAAAACGACGGCACT TGTTACTCCGCAACGCTTTTC	Amplification of <i>DIT1t</i> terminator



# Homology arms of integration sites

<i>XI-5-up-F</i>	GCGGAGAAGTCGTTGATAGC	Amplification of <i>XI-5-up</i> fragment
<i>XI-5-up-R</i>	TGGTGCACGGAGTTTATGG	Amplification of <i>XI-5-up</i> fragment
<i>XI-5-down-F</i>	TTGGATTACATCCTACTTTT	Amplification of <i>XI-5-down</i> fragment
<i>XI-5-down-R</i>	GCCGGAATAGTGCAAATTGT	Amplification of <i>XI-5-down</i> fragment
<hr/>		
<i>XI-2-up-F</i>	TAACTCTTCGTATGAGGATTTTCGAT G	Amplification of <i>XI-2-up</i> fragment
<i>XI-2-up-R</i>	AAGGATAGTAAGCTGGCAAATTCTA TGGCACATTTTTCTGTTGAG	Amplification of <i>XI-2-up</i> fragment
<i>XI-2-down-F</i>	CGTTGCGGAGTAACAAGTCCACAAG TAAAGCTCGTTGAC	Amplification of <i>XI-2-down</i> fragment
<i>XI-2-down-R</i>	ATGGTTGAAAAGGTTACAGAGGATC	Amplification of <i>XI-2-down</i> fragment
<i>Ty3-up-F</i>	TGTTGTATCTCAAAATGAGATATGTC AGTATGACAATACG	Amplification of <i>Ty3-up</i> fragment
<i>Ty3-up-R</i>	GCTCCAATTCAGCTGGCGTAATAGCT CGGATCTAAACTAATTGTTTCAGGCAT TTATACT	Amplification of <i>Ty3-up</i> fragment

<i>Ty3-down-F</i>	AGAAGTTAAGAAAATCCTTGCTTAA GATTCCGCGCTTCCACCACT	Amplification of <i>Ty3-down</i> fragment
<i>Ty3-down-R</i>	TGTTGTATTACGGGCTCGAGTAATAC C	Amplification of <i>Ty3-down</i> fragment
<i>EXG1-up-F</i>	ACGAGCCTGAGACAAGCCCG	Amplification of <i>EXG1-up</i> fragment
<i>EXG1-up-R</i>	TTTAGTTGGTAATTAAGTAGAAAAAG AAAGTAAACAAAAAATCAAAGG	Amplification of <i>EXG1-up</i> fragment
<hr/>		
<i>EXG1-down-F</i>	GTCGCCCTCAGTCCGCTCATTTTAG	Amplification of <i>EXG1-down</i> fragment
<i>EXG1-down-R</i>	GTTGTTTAAGTTCTTTATCCTTCCTT AGATAACACC	Amplification of <i>EXG1-down</i> fragment
<i>Ty1-up-F</i>	TGTTGGAATAAAAAATCAACTATCATC TACTAACTAGTATTTACGTTAC	Amplification of <i>Ty1-up</i> fragment
<i>Ty1-up-R</i>	TATGTTAATATTCATTGATCCTATTAC ATTATCAATCCTTGCG	Amplification of <i>Ty1-up</i> fragment
<i>Ty1-down-F</i>	CTTTTTTTTTCTCTTGAACCTCGATAAA ACGGAATGATGAATAATATTATAG AATTGTG	Amplification of <i>Ty1-down</i> fragment
<i>Ty1-down-R</i>	TGAGAAATATGTGAATTTTGAGATAA TTGTTGGGATTC	Amplification of <i>Ty1-down</i> fragment

<i>Ty2-up-F</i>	TGTTGGAATAAAAATCAACTATCATC TACTAACTAGTATTTACGTTAC	Amplification of <i>Ty2-up</i> fragment
<i>Ty2-up-R</i>	GTTAATATTCATTGATCCTATTACATT ATCAATCCTTGCG	Amplification of <i>Ty2-up</i> fragment
<i>Ty2-down-F</i>	CGAATATAAAAATGATGATAATAATA TTTATAGAATTGTGTAGAATTGCAG	Amplification of <i>Ty2-down</i> fragment
<i>Ty2-down-R</i>	GAGAATGTGGATTTTGATGTAATTGT TGGG	Amplification of <i>Ty2-down</i> fragment
<i>FAS1p-up-F</i>	TACTATGCGGTCTCGTCCTCTACG	Amplification of <i>FAS1p-up</i> fragment
<i>FAS1p-up-R</i>	TGCCGACTGTAATCTCTGATTATTGTC G	Amplification of <i>FAS1p-up</i> fragment
<i>FAS1p-down-F</i>	ATGGACGCTTACTCCACAAGACC	Amplification of <i>FAS1p-down</i> fragment
<i>FAS1p-down-R</i>	TCGGTGTTACCTTGACCACCG	Amplification of <i>FAS1p-down</i> fragment
<i>OCA5-up-F</i>	GGCGCAATATGCAGGTACCG	Amplification of <i>OCA5-up</i> fragment
<i>OCA5-up-R</i>	GTAGTATATAATTCTAAAGGAAATA ATTCAACTTTATTTGCTAGCGAGCC CTTTACC	Amplification of <i>OCA5-up</i> fragment

<i>OCA5-down-F</i>	AAAGTTGAAATTATTTCTTTAGAAT	Amplification of
	TATATACTACTTCTAC	<i>OCA5-down</i> fragment
<i>OCA5-down-R</i>	TATCTCGATAGTTGGAAGATTTAGTA	Amplification of
	CCATAAGTAGC	<i>OCA5-down</i> fragment

---

**Table S4 Codon optimized exogenous gene sequences**

Synthetic gene	Nucleotide Sequence (5' - 3')
<i>AtC4H</i>	TTAACAGTTTCTTGGCTTCATAACAATAATAGAGTGGTTCAAA ATGTGCAAAGAGAATTGACCACCCTTTTCAGAGGTGTCGACCT TAGATTGACCTGGTGGTGGCAACAATTCAAAATTTTGGACCAT TCTACCAATGGTAATACCTAAAATTGGCAAAGCCAAAATAAT ACCAGGACAAGATCTTCTACCAACACCAAATGGGACGTATCT AAAATCATTACCATTAGCTTCAACGTGAGATTCTTCTTCGAAA AATCTTTCTGGTCTGAATTCTTCTGGTTTCTTCCATGAATTTGGG TTGTTAGCCAACCACCAAGCGTTAACCAAAATCTTACTTTCAG CAGGGATATCGTAACCAGCTAACTTAGCATCATGCAAATTCAT ATGTGGGACTAACAAAGGAATAGCCATTCTCAAACGCAAAGT TTCCTTAACAACTGCTTGCAAATATGGCAATTTATGCAAGTCA GGTTCAGTAACTTGGACACCAGGACCTAAAACAGTATCTAATT CATTTCTCAACTTAGATTGAATTTTCAGGATGATTAATAATTCA GCAATACCCCATTCATGGACCATAAAGTAGTTTCAATGGCAG CAACGTTAATATTCTCAACAATATACAAAACATTATCTTCATTA ATTTACCCCTTTTGCTCGGCTTCCAAAATGTGGTCGATAGCACA CTTAAGACCTTCAGAACCAGTTGGCTTTGAAGAGGCAATTTGC TTACGTTTCATCAACAAAGTACTTCTTAAACAAAGCAATTCTTCT ATCCTTAACATCTTGACAAATCTTTAGGTAACCTCTTAAAAATG GTCTTAAATTTGGAATGAAGTCACCGTAGTTGTATTGGAATGA TTGGGCCAATCTGGATCTTTACCGTTCAAAGCCTTCAAACGC AAGAATAAAGGATCATCTTCTGATTCAAATCTACGGTTCGAACA TAATTCTAAACATGTTATTGTACATCATCAATTGTAATCTCTTTC TCAAACGATACCCTTGGTAGCAGAATCAGGATTCTTCTTAAC ATCTTCGACAACGGAAGCAGCTTCAAATTCCCAACCTTCTCTA TTTTGTTGAACAACCTTTGTTAGTGAAGAAAGGGACGGTCATTAT TCTTCTCATCTTACGCCAGTGTTACCGTAAACTGTGAAAACCA TGCTTGACCTTTACCAGTGAAAATATCAAAAACAACGTTTCT AGTTCTAGAACCGAATTCAACACCTTGAGTTAACAAAACCTCC TTAGTCAAATCTGGAGAAGAAACAACAACCAAATTTCTTTGGC CCATTCTCAATAAGAACAAGTCACCGAACTTTTTAGCGTAATC AACCAAGTTTCTATGGTTCAAATCGTCACCAACTTGTAACCAG TTACCGAAAATTGGAATTGGGATTGGACCTGGTGGCAACTTCA ATTTTTTACCTCTCAACTTGGAATGACAGTGGCCAAAATAAC AGCGACGAAAACGGCAATCAAAGACTTTTCCAACAACAACAA ATCCAT
<i>AtPAL2</i>	ATGGATCAAATTGAAGCCATGTTGTGTGGTGGTGGTGAAAAAA CTAAAGTCGCAGTCACTACCAAGACCTTGGCTGATCCTTTGAA CTGGGGTTTAGCTGCTGATCAAATGAAGGGTTCTCATTTGGATG

---

AAGTTAAGAAGATGGTTGAAGAATATCGTAGACCAGTTGTTAA  
CCTAGGTGGCGAAACATTGACCATTGGTCAAGTTGCTGCTATC  
TCTACTGTTGGTGGTTCCGTTAAAGTTGAATTAGCTGAACTTC  
AAGAGCTGGTGTAAAGCTTCTTCTGACTGGGTATGGAATCT  
ATGAACAAAGGTACTGATTCTTATGGTGTACTACTGGTTTTGG  
TGCTACTAGTCATAGAAGAACTAAGAATGGTACAGCTTTGCAA  
ACTGAATTGATCAGATTTTTGAACGCTGGTATTTTCGGTAACAC  
TAAGGAACTTGTGCATACCCTACCACAATCTGCCACTCGTGCT  
GCTATGTTGGTTAGAGTAAACACCTTATTGCAAGGTTATTCTGG  
TATTAGATTGCAAATCTTGGAAGCCATTACATCTTTGTTAAATC  
ATAACATCTCTCCTTCCTTGCCATTGAGAGGTACTATTACTGCT  
TCTGGTGACTTAGTTCCATTGTCATACATTGCCGTTTTGTTGAC  
CGGTAGACCAAACCTAAGGCTACTGGTCCAGATGGTGAATCA  
TTGACCGCTAAGGAAGCTTTCGAAAAGGCTGGTATTAGTACTG  
GTTTCTTTGACTTGCAACCAAAAAGAAGGGTTAGCTTTAGTCAA  
CGGTACTGCTGTCGGTTCCGGTATGGCCTCAATGGTTTTGTTTG  
AAGCTAACGTTCAAGCCGTTTTGGCTGAAGTCTTATCTGCTATT  
TTCGCTGAAGTTATGTCCGGCAAACCAGAATTCAGTATCATT  
GACTCATAGATTGAAACACCATCCAGGTCAGATTGAAGCTGCT  
GCTATTATGGAACATATTTTGGATGGTTCTTCTTACATGAAGCT  
AGCTCAAAAGGTACATGAAATGGATCCATTACAAAAGCCAAA  
GCAAGATAGATACGCACTTAGAACTTCTCCACAATGGTTGGGT  
CCACAAATTGAAGTTATTAGACAAGCTACTAAATCTATTGAAA  
GAGAAATTAAGTCTGTCAATGATAATCCATTGATTGATGTTTCC  
AGAAACAAAGCTATTCACGGTGGTAATTTCCAAGGTACTCCAA  
TTGGTGTTTCTATGGATAATACTAGATTAGCTATAGCTGCCATT  
GGTAAATTAATGTTTGCCCAATTTTCTGAATTGGTTAATGATTT  
CTACAACAATGGTTTGCCATCAAACCTAACAGCTTCTTCTAACCC  
CATCTTTGGATTACGGTTTTAAGGGTGCTGAAATTGCCATGGCT  
TCTTACTGTTCTGAATTACAATACTTGGCTAACCCAGTTACCTC  
CCATGTTCAATCAGCTGAACAACACAACCAAGATGTTAACTCT  
TTGGGCCTAATTTCTCTAGAAAGACTTCTGAAGCTGTTGATAT  
CTTAAAGTTAATGTCTACAACCTTCTTAGTTGGTATCTGTCAAG  
CTGTTGACTTGAGACATTTGGAAGAAAATTTGAGACAAACAGT  
TAAGAATACTGTTAGTCAAGTTGCTAAAAAAGTTTTGACTACT  
GGTATTAACGGTGAATTGCATCCTAGTAGATTCTGTGAAAAGG  
ATTTGTTAAAAGTCGTTGATAGAGAACAAGTTTTACATACGTT  
GATGATCCATGTTTACGCTACCTACCTTTAATGCAAAGATTAA  
GACAAGTTATTGTTGATCATGCCTTATCTAATGGTGAACTGA  
AAAGAACGCTGTCACCTTCTATTTTCAAAGATTGGTGCCTTTG  
AAGAAGAATTGAAGGCTGTTTTGCCAAAGGAAGTTGAAGCAG  
CTAGAGCTGCATATGGTAATGGTACCGCTCCTATTCCTAATAG  
AATTAAGGAATGTAGAAGTTATCCATTGTATAGATTGTTAGA  
GAAGAAGTAAAGTAAATGTTGACTGGTGAAAAAGTCGTTT

---

	CTCCAGGTGAAGAATTTGATAAGGTTTTTACAGCTATGTGTGA AGGAAAGTTAATTGATCCATTAATGGATTGCTTAAAGGAATGG AACGGTGCTCCAATTCCAATTTGTTAAAT
<i>AiATR2</i>	TTACCAAACATCTCTCAAGTATCTACCGGAGGTTTGCAAGTTCT TAACGAAACCTTCAGCTTTAGTTGAGTCCATAGAACCTTGTTCT TGAGCAATAGTGTGCAAAGATCTATGAACATCTCTAGCCATAC CTTTAGCATCGCCACAAACATACAAGTAAGCACCTTGAGAAAT CATGTTCCAGATGTCAGAAGCTTTATCCATCATCTTATGTTGAA CGTATTCCTTAGTTGGACCTTCTCTTGAGAAAGCAACTGACAAT TCAGCTAAAGCACCAGATTCAACAAATCTTTGCAATTCTTCTTC ATAGATAAAATCCATTCTTCTATTTCTGCAACCGAAGAACAAA ACAGATGGACCAAGTTCAACACCAGATTCAACTAAAGCCAAT CTTTCTTGCAAGAAACCTCTGAATGGTGCTAAACCAGTACCTG GACCAATCATGATGATTGGGACCTTAGAATCAGATGGTAACTT AAAATTAGATTGTCTAACAAAAATTGGAGCAGAAGAACAATT TTCAGACTTTTCATATGGAACAGCATTCTTCATCCAAGTAGAAC AGACACCCTTATGGATTCTACCGGTTGGCATTTTTTTCATAGACT AAAGCGCAGGTAACATGAATTCTTGTTTCAGCGATTTTTGGGG AGCTAGAGATAGAATAGAATCTTGGCTGTAATCTTGGGGCGAC ACCGGCAAAGAACACACCCAATGGTGGTTTAGCAGATGGGAA TTCAGCCATAACTTCTAGTAATGATCTTTGAGATTCAACAACCC ACTTGGAGTATTCATCCTTTCCAGCTGGGGAAGCCAAATGTTTC AATCTTTTCGGCTTCAGTTGGATCACTAGCGTGTGCGGCTAGGG CAACTAGGGCTGACTTTTTAGGAGAGGATAGCAAACAAGCGT ATCTAGTTAAAGCGGTTCTAAGATTACATGGAGGGAATGGAGG TGGCAAAGAAGAGGAGATCGGAGTACCATCTTCCTTTTCAGCA TGCAAAGAAAAATAGGTGTCAGGAGACATATCCAATAATCTC AAGGCTTCATCAACAGTTTCGGACAAGTTGTCACACAAAACAC CAACATGGTCACCAGTTTCGTAAGTTAAACCAGAACCGGCTAT ATCGAATTCTAGGTGGATACAAGAACGATCAGATTCTGGAGTA TGTAATTCTCTTTTGACTGCTACATTAGCCTTATAAGGATGTTG AGCATCAAAAACGGTGTAACCATTGCCGTTAGCCATGTTTATA TCGTAAACTTAGCATCTTCAGAATCATGGATACTAACTCTATA TTCTAAGACAGCTGCAGTATATGGTGTAGCAACAGCGGTGTCA CCTTCTTCTCTTAGAATGGTATCTAGTTCTGGCCACAAGGCTTC ACGCCATGCGGTGAAGTCATCTTCAATACATTGATCATCATCC CCCAAACCAACTTGAACCAATCTTTGAGCACCTTGTTGACCA AAATGTCATCAACAACCTTAGCGACCTTATTAATGTTTCGTA TTGTCTATTACCCAAACCAAAAACCCCATATTTTAAGTTCTTCA ACCATTACCTCTATCGTTACCTTCGGTGAACCACTTATAAAAT CTGGCAGCGTTATCAGTAGGTTACCGTCACCGTAAGTAGCTA AAAAGAAAAAGGCAACATCTTCTTTCTTTAGCTTTTCCTCATAT TCATCGTCGTCAGCGGCGTAATCATCAAATCAACGATCTTGA ATCTAGTCTTTTCATATCTAGCCTTAGCTTCTTCACCCAAAGCC

---

TTAGCGAAACCTTCAGCAGTACCAGTTTGAGTACCGAAGAAA  
 ATAGTAACTTTCTTTCTACCGTCATCAATTTCTTCTTCTTGGC  
 TTTATAACCAAAGGTTTTAAAGGTTCAACTCTCTTAGAGTTACC  
 CGAACCTGAACGCCTCCAACTAGCATAACAATACAACCAAT  
 CAAAACAGCGATGGAAGTAGTGACAATCATAGCAAATTGTCT  
 ATTTTCAATCAACATAGAACTAAGTTCAGCGGCGACAGATTCTG  
 TATGCAGAAGCATTAGCGGGATCGGAAACAATAACTGGTTCA  
 CCCTTTATAATGGCAGCCATCAAGTCGATCATGGAAGTAGAGG  
 AGGAAGAACTAGAGGACATTGCATTGTAGATGTTATAAGAAG  
 AAGAAGATAATAATAAGAAAAAGAAAGAAGAAGATGGCAAC  
 AACAAGTCTCTTTCAGTATGAATAGTTGGTGGCTTAGTCAAGT  
 GTGGAGTTGGAGACAAGGAAACACCCTTACTATTCTTAGAAAC  
 ACTACACAATTTCAACTTGTAGTTCATCATGTTCTTCAT

---

*FjTAL*

ATGAACACCATTAATGAATACTTGAGTTTAGAAGAATTCGAAG  
 CAATAATCTTCGGTAACCAAAAAGTAACTATCTCTGATGTTGT  
 CGTAAACAGAGTTAACGAAAGTTTTAACTTCTTAAAGGAATTT  
 TCTGGTAATAAGGTTATATATGGTGTAAACACTGGTTTCGGTCC  
 AATGGCTCAATACAGAATCAAGGAATCTGATCAAATCCAATTG  
 CAATACAATTTGATAAGAAGTCATTCTTCAGGTACTGGTAAAC  
 CATTATCTCCTGTTTGTGCTAAGGCTGCAATCTTGGCAAGATTG  
 AACACATTGTCCTTAGGCAACTCAGGTGTTACCCATCTGTTAT  
 TAATTTGATGTCTGAATTGATAAAACAAAGACATCACTCCTTG  
 ATATTCGAACATGGTGGTGTGGTGCATCTGGTGACTTGGTCCA  
 ATTGTCCCACTTGGCCTTAGTATTGATAGGTGAAGGTGAAGTTT  
 TCTATAAAGGTGAAAGAAGACCAACACCTGAAGTCTTCGAAA  
 TCGAAGGTTTTAAAGCCTATACAAGTAGAAATCAGAGAAGGTTT  
 AGCTTTGATTAATGGTACTTCTGTCATGACAGGTATAGGTGTTG  
 TCAACGTATACCATGCTAAGAAATTGTTGGATTGGTCATTGAA  
 GTCCAGTTGTGCCATTAATGAATTGGTTCAAGCATATGATGAC  
 CATTTCTCTGCAGAATTGAACCAAACCAAGAGACACAAGGGT  
 CAACAAGAAATCGCATTGAAGATGAGACAAAATTTGTCCGAT  
 AGTACATTGATCAGAAAGAGAGAAGACCACTTATACTCAGGT  
 GAAACACCGAAGAAATTTTCAAAGAAAAGGTTCAAGAATAC  
 TACTCCTTGAGATGCGTCCCACAAATCTTGGGTCCTGTATTGGA  
 AACTATTAATAACGTTGCCTCAATCTTGGGAAGATGAATTCAAT  
 TCCGCTAACGATAACCCAATCATCGACGTTAAAAATCAACATG  
 TTTATCACGGTGGTAACTTCCATGGTGACTIONTTCTTTAGAA  
 ATGGACAAATTGAAGATAGTTATCACAAAATTGACCATGTTGG  
 CTGAAAGACAATTGAACACTTGTGAACTCAAAGATTAAACGA  
 ATTGTTGCCACCTTTCGTTAATTTGGGTACATTGGGTTTTAACTT  
 CGGTATGCAAGGTGTTCAATTCACCGCCACTTCAACTACAGCT  
 GAATCCCAAATGTTGAGTAACCCAATGTACGTTCAATCCATCC  
 CTAACAACAACGATAACCAAGACATCGTCTCTATGGGTACCA  
 ACTCAGCCGTCATTACTTCCAAAGTAATAGAAAACGCATTCCA

---



	AGTTTTGGCCATCGAAATGATCACAATTGTCCAAGCTATCGAT TACTTGGGTCAAAAGGACAAGATCTCTTCTGTTTCTAAGAAAT GGTACGATGAAATAAGAAACATAATCCCAACCTTTAAGGAAG ACCAAGTTATGTACCCTTTCGTACAAAAGGTAAAGGATCATTT GATTAACAATTAA
<i>PhCHS</i>	TTAGGTAGCCACACTATGCAGAACCACCGTTTCCACCGTCAGA CCCGGACCAAAGCCGAACAGCACGCCCCATTCCAGGCCTTCA CCCGTGGTGCCCAGACCTTCTTTAGCCGATGCTTTGCGCATTTC ATCCAGAATAAACAGGACGCAGGCAGAGCTCATATTGCCGTA GTTTGACAGAACATTACGGGTCGCTTTCAGTTTTTCCGGTTTCA GACCCAGTTTAATTTTCGACCTGGTCCAGGATTGCCGGGCCGCC CGGATGTGCAATCCAGAACAGGCTGTTCCAATCGGAGATACTC AGCGGGCGAAACGCTTCTTCCAGGCTTTTTTCAATGTTTTTTGA GATCAGGCCCCGGCACATCTTTCAGCAGATGGAACGTCAGACC GACTTCACGCAGGTGACCGTCAATTGCGCCGTGGCTATCCGGC AGCAGGGTTTGAGCTGCGCTGACCAGTTCAAACAGCGGACGTT CAACGCCCCGAATCGGGTCGCTACCAATGATAATGGCACCTG CACCATCACCGAACAGGGCCTGGCCAACCAGGGAGTCCAGGT GCGTATCGTTCGGACCACGAAACGTAACAGCGGTGATTTCACT ACACACCACCAGAACACGCGCACCTTTATTGTTTTTCGGCCAGA TCTTTCGCCAGACGCAGAACGGTACCGCCTGCGAAGCAGCCTT GCTGGTACATCATCAGGCGTTTCACGCTCGGACGCAGACCCAG CAGTTTTCGTCAGTTGATAGTCACAACCCGGCATATCCACACCC GACGTGGTGCAAAAGACCAGATGGGTAATTTTTGATTTCCGGTT GGCCCCATTCTTTGATCGCTTCTGGGCCGCTTCTTTACCCAGTT TCGGCACTTCGACAACCACAATGTCCTGGCGAGCATCCAGAG ACGGTGCCATGTATTCACACATGCTCGGATTTTCTTTCAGAATT TCTTCGGTCAGGTGCATATAGCGTTTTTTGATCATGCTTTTTTCG CACATACGTTTGAATTTTTCTTTCAGATCCGTTTTATGTTTCAGAG TTGGTGATACGAAAGTAATAGTCCGGATAGGTGCTCTGATCAA CACAGTTCGTCGGGGTTGCCGTGCCAATCGCCATACCGTCGC CGGGCCTTCAGCGCGTTGAGCTTTGCGGTATTCTTCCACCGTAA CCAT
<i>MsCHI</i>	ATGGCAGCAAGCATTACGGCAATCACGGTTGAAAATCTGGAA TATCCGGCGGTGCGTTACCTCTCCGGTCACGGGCAAATCATACTT TCTGGGCGGTGCCGGTGAACGTGGTCTGACCATGAGGGTAAC TTTATCAAATTCACGGCAATTGGCGTTTATCTGGAAGATATCGC GGTCGCCTCACTGGCGGCCAAATGGAAAGGTAAAAGCTCTGA AGAACTGCTGGAAACCCTGGATTTTTACCGTGACATTATCTCA GGCCCGTTCGAAAAACTGATCCGTGGTTCGAAAATTCGCGAAC TGAGCGGCCCCGGAATATTCTCGCAAAGTCATGGAAAACCTGCGT GGCTCATCTGAAATCCGTCCGCACGTACGGTGACGCAGAAAGCT GAAGCGATGCAGAAATTTGCCGAAGCATTCAAACCGGTGAAT TTTCCGCCGGGTGCCAGTGTTTTCTATCGTCAATCCCCGGATGG

---

CATCCTGGGTCTGTCATTTTCGCCGGACACCAGCATCCCGGAA  
AAAGAAGCAGCTCTGATTGAAAATAAAGCTGTGAGTTCCGCG  
GTTCTGGAAACGATGATTGGCGAACACGCGGTTTCTCCGGATC  
TGAAACGCTGTCTGGCTGCTCGCCTGCCGGCTCTGCTGAATGA  
AGGTGCCTTTAAAATCGGTAACCTGACCTAGG

---

*Pc4CL*

ATGGGTGACTGCGTTGCCCCGAAAGAGGATCTGATCTTCCGCA  
GCAAACCTGCCGGACATTTACATTCCAAAGCATCTGCCGCTGCA  
TACCTATTGTTTTGAGAACATCAGCAAGGTTGGCGACAAGAGC  
TGTCTGATCAACGGCGCAACCGGCGAAACCTTTACCTACAGCC  
AGGTTGAGCTGCTGTCCCGTAAAGTTGCCAGCGGCCTGAACAA  
GCTGGGCATTCAACAAGGTGATACCATTATGCTGCTGCTGCCG  
AACTCCCCGGAGTACTTTTTCGCTTTCCTGGGTGCGAGCTATCG  
CGGTGCAATCAGCACTATGGCGAAACCCATTCTTTACCAGCGCA  
GAAGTGATCAAGCAACTGAAAGCGAGCCAAGCGAAGCTGATT  
ATCACCCAGGCATGCTATGTTGACAAGGTTAAGGACTACGCAG  
CGGAGAAAAACATCCAGATCATTTGTATTGACGATGCACCGCA  
GGATTGCCTGCACTTTAGCAAGCTGATGGAAGCGGATGAGAG  
CGAAATGCCGGAAGTGGTTATTAACAGCGATGATGTGGTGGC  
ACTGCCGTACAGCTCTGGCACCACCGGCCTGCCGAAAGGCGTT  
ATGCTGACCCACAAGGGTCTGGTTACCAGCGTTGCACAACAGG  
TGGATGGTGATAACCCGAACCTGTATATGCACTCCGAGGATGT  
TATGATCTGCATCCTGCCACTGTTCCATATCTATAGCCTGAACG  
CTGTTCTGTGTTGTGGTCTGCGTGCGGGCGTTACCATTCTGATC  
ATGCAAAAGTTGACATTGTGCCGTTTCTGGAGCTGATTGAGA  
AGTATAAGGTTACCATTGGTCCGTTTGTTCGCCCGATCGTGCTG  
GCCATCGCGAAAAGCCCCGTTGTTGACAAGTACGACCTGTCTA  
GCGTGCGCACCGTTATGAGCGGTGCAGCGCCGCTGGGTAAAG  
AGCTGGAGGACGCTGTTTCGTGCGAAATTCGCGAACGCGAAGC  
TGGGTCAAGGCTATGGCATGACCGAAGCCGGTCCGGTTCTGGC  
GATGTGTCTGGCGTTCGCCAAAGAGCCGTATGAGATTAAGTCT  
GGCGCATGCGGTACCGTTGTGCGTAACGCCGAGATGAAAATC  
GTTGACCCAGAAACCAACGCGTCTCTGCCGCGTAACCAGCGTG  
GTGAGATTTGCATCCGTGGTGATCAGATTATGAAAGGTTACCT  
GAACGACCCGGAAGCACCCGCACCACCATCGACGAAGAGG  
GTTGGCTGCACACCGGTGACATTGGTTTCATCGACGATGACGA  
TGAACCTGTTCAATTGTTGATCGTCTGAAAGAAATCATTAAGTAC  
AAAGGTTTTCAAGTTGCTCCGGCGGAGCTGGAAGCACTGCTGC  
TGACCCACCCGACCATCAGCGATGCCGCGGTGGTTCCGATGAT  
TGACGAGAAAGCGGGTGAAGTGCCAGTGGCGTTTGTGTGCGT  
ACCAACGGTTTTACCACCACCGAAGAAGAAATCAAACAATTT  
GTGAGCAAACAGGTTGTGTTCTACAAACGTATCTTCCGCGTTTT  
CTTCGTTGACGCTATTCCGAAATCCCCGAGCGGCAAGATTCTG  
CGTAAGGATCTGCGCGCTCGTATTGCGAGCGGCGACCTGCCGA  
AGTAA

---

---

*GeHIS* ATGTTGGTTGAATTGGCTATCACTTTGTTAGTTATTGCATTGTTT  
 ATTCACTTAAGACCAACATTGTCTGCAAAGTCTAAGTCATTGA  
 GACATTTGCCAAATCCACCATCACCAAAACCAAGATTACCATT  
 CGTTGGTCATTTGCATTTGTTGGATAAGCCATTGTTGCATTACT  
 CTTTGATTGATTTGTCAAAGAGATACGGTCCATTGTACTCTTTG  
 TACTTCGGTTCAATGCCAACTGTTGTTGCTTCTACACCAGAATT  
 GTTTAAATTGTTTTTACAACTCATGAAGCTTCTTCTTTTAATAC  
 TAGATTCCAAACATCTGCAATCAGAAGATTGACATACGATAAC  
 TCAGTTGCTATGGTTCCATTTGGTCCATACTGGAAGTTTATTAG  
 AAAGTTGATCATGAACGATTTGTTAAATGCTACTACAGTTAAT  
 AAGTTAAGACCATTGAGATCTCAAGAAATCAGAAAGGTTTTA  
 AGAGTTATGGCTCAATCTGCAGAATCACAAGTTCCATTGAACG  
 TTAAGTTAAGTTGTTGAAGTGGACTAACTCTACAATCTCAAG  
 AATGATGTTGGGTGAAGCTGAAGAAATTAGAGATATTGCAAG  
 AGATGTTTTGAAGATCTTCGGTGAATACTCTTTGACAGATTTC  
 TCTGGCCATTGAAGAAATTGAAGGTTGGTCAATACGAAAAGA  
 GAATCGATGATATTTTCAATAGATTTGATCCAGTTATTGAAAG  
 AGTTATTAAGAAAAGACAAGAAATCAGAAAGAAAAGAAAGG  
 AAAGAAATGGTGAAATTGAAGAAGGTGAACAATCAGTTGTTT  
 TCTTGGATACTTTGTTGGATTTTCGCTGAAGATGAACTATGGAA  
 ATTAATAATCACAAAGGAACAAATTAAAGGTTTGGTTGTTGATT  
 TCTTTTCTGCAGGTACTGATTCAACAGCTGTTGCAACAGATTGG  
 GCTTTGTCTGAATTGATTAATAACCCAAGAGTTTTCCAAAAGG  
 CTAGAGAAGAAATCGATGCAGTTGTTGGTAAAGATAGATTGGT  
 TGATGAAGCTGATGTTCAAAATTTGCCATACATCAGATCAATC  
 GTTAAGGAACTTTTAGAATGCATCCACCATTGCCAGTTGTTA  
 AGAGAAAGTGTGTTCAAGAATGTGAAGTTGATGGTTATGTTAT  
 TCCAGAAGGTGCTTTGATTTTGTTTAATGTTTGGGCAGTTGGTA  
 GAGATCCAAAATACTGGGATAGACCAACAGAATTCAGACCAG  
 AAAGATTTTTAGAAAATGTTGGTGAAGGTGACCAAGCAGTTGA  
 TTTGAGAGGTCAACATTTTCAATTGTTACCATTGTTGTTCTGGTA  
 GAAGAATGTGTCCAGGTGTTAATTTGGCTACTGCTGGTATGGCT  
 ACATTGTTGGCATCTGTTATTCAATGTTTCGATTTGTCAGTTGTT  
 GGTCCACAAGGTAAAATCTTGAAGGGTAACGATGCTAAAGTTT  
 CAATGGAAGAAAGAGCTGGTTTAACTGTTCCAAGAGCACATA  
 ATTTGATTTGTGTTCCAGTTGCTAGATCTTCAGCAGTTCCAAAA  
 TTGTTTTCTTCATAA

---

*GmIFS* ATGTTGTTAGAATTAGCTTTGGGTTTATTCGTTTTGGCTTTGTT  
 TTGCACTTGAGACCAACTCCATCTGCTAAGTCTAAGGCTTTGA  
 GCACTTGCCAAACCCTCCTTCTCCAAAGCCAAGATTACCATT  
 CATCGGTCAATTTGCATTTGTTAAAGGATAAGTTGTTGCACTATG  
 CTTTAATTGATTTGTCTAAGAAGCATGGTCCATTGTTCTCATTGT  
 CTTTCGGTCCATGCCAACTGTTGTTGCTTCCACTCCAGAATTG  
 TTCAAGTTGTTCTTGCAAACTCATGAAGCTACCTCTTTTAACAC

---

---

TAGATTCCAAACCTCTGCTATTAGAAGATTAACCTACGATAAT  
TCAGTTGCTATGGTTCCATTTGGTCCATACTGGAAGTTTGTTAG  
AAAGTTAATCATGAATGATTTGCTAAACGCTACTACTGTTAAC  
AAGTTGAGACCATTAAGAACCCAACAAATTAGAAAATTCTTG  
AGAGTTATGGCTCAATCAGCTGAAGCTCAAAAGCCATTGGATG  
TCACTGAAGAACTTTTGAAGTGGACTAATTCTACCATTTCATG  
ATGATGTTGGGTGAAGCTGAGGAAATTCGTGATATCGCTAGAG  
AAGTTTTAAAAATTTTGGTGAATACTCCTTGACCGATTTTATTT  
GGCCATTGAAATACCTGAAGGTTGGTAAATACGAAAAAAGAA  
TTGATGACATTTTGAACAAGTTCGATCCTGTTGTTGAAAGAGTT  
ATCAAGAAGCGTAGAGAAATTGTTAGGCGTCGTAAGAACGGC  
GAAGTTGTTGAAGGTGAAGCTTCAGGTGTCTTTTTGGACACATT  
ATTAGAATTTGCTGAAGATGAAACCATGGAAATCAAGATTACT  
AAGGAACAAATCAAGGGTCTAGTCGTCGATTTCTTCTCTGCTG  
GTACTGATTCCACTGCTGTTGCAACTGAATGGGCTCTAGCTGA  
ACTAATTAACAACCCTAGAGTGTTACAAAAGGCTAGAGAGGA  
AGTTTACTCTGTCGTTGGGAAGGATCGTTTAGTCGATGAAGTTG  
ACACTCAAACTTGCCATATATTAGAGCTATTGTTAAGGAAAC  
TTCCGTATGCATCCACCTTTGCCAGTCGTTAAGAGAAAAATGTA  
CAGAAGAATGTGAAATTAATGGTTACGTTATTCCAGAAGGTGC  
CTTAGTTTTGTTCAATGTTTGGCAAGTAGGTAGAGATCCAAAGT  
ATTGGGATCGTCCATCTGAATTCAGACCTGAAAGATTTTLAGA  
AACTGGTGCTGAAGGTGAAGCAGGTCCTTTAGATTTGAGAGGT  
CAACATTTTCAATTGTTACCATTCCGGTCCGGTAGAAGAATGTG  
TCCAGGTGTTAACTTAGCTACATCAGGTATGGCTACTTTATTGG  
CTTCCTTAATTCAATGTTTCGATCTTCAAGTCTTAGGTCCCTCAA  
GGTCAAATCTTGAAAGGCGATGATGCTAAAGTTTCCATGGAAG  
AAAGAGCCGGTTTGACTGTTCCAAGAGCTCATAGTTTAGTTTGT  
GTTCCATTAGCTAGAATTGGTGTTGCTTCTAAATTGTTGTCCTA  
A

---

*GmIFS2*

ATGTTGTTGGAATTGGCTTTGGGTTTGTTGGTTTTAGCTTTGTTT  
TTGCATTTGAGACCAACTCCAAGTCTAAGGCTTTGA  
GACACTTACCAAATCCACCATCTCCTAAGCCAAGATTACCTTT  
CATCGGTCACTTGCAATTTGTTGAAGGACAAGTTATTGCACTATG  
CTTTGATTGATTTGTCCAAGAAACATGGTCCATTGTTTCTTTAT  
ACTTCGGTTCTATGCCAACCGTTGTGCTTCTACTCCTGAATTGT  
TTAAATTATTCTTGCAAACACATGAAGCTACCTCTTCAACACT  
AGATTCCAAACATCTGCTATTAGAAGATTAACCTACGATTCAT  
CCGTTGCTATGGTTCCCTTCGGTCCATATTGGAAATTCGTTAGA  
AAGTTAATTATGAATGACTTGTTAAACGCTACTACCGTCAACA  
AATTGAGACCATTGAGAACTCAACAAATTCGTAAGTTTTTAAG  
AGTCATGGCTCAAGGTGCTGAAGCTCAAAAACCATTTGGATCTA  
ACTGAAGAATTGCTAAAATGGACTAATAGTACTATTCTATGA  
TGATGTTAGGTGAAGCTGAAGAAATTAGAGACATTGCTAGAG

---

---

AAGTTTTAAAGATTTTCGGTGAATACTCTTTGACCGACTTCATT  
TGGCCATTGAAACACTTAAAAGTTGGTAAATACGAAAAGAGA  
ATTGATGACATCTTAAATAAGTTTGATCCTGTTGTTGAAAGAGT  
TATTAAGAAAAGAAGAGAAATTGTAAGACGTCGTAAAAACGG  
TGAAGTTGTTGAAGGTGAAGTTAGTGGTGTTTTTCTTGATACTT  
TGTTAGAATTCGCTGAAGATGAAACTATGGAAATTAAGATAAC  
TAAGGACCATATTAAGGGTTTAGTTGTTGATTTCTTTCTGCCG  
GTACAGATTCTACTGCTGTTGCCACAGAATGGGCATTAGCAGA  
ATTGATTAATAACCCAAAGGTTTTGGAAAAGGCAAGAGAAGA  
AGTTTATTCTGTTGTTGGTAAAGATAGATTGGTCGATGAAGTTG  
ATACCCAAAATTTGCCTTACATTAGAGCTATTGTAAAGAAAC  
TTTTAGAATGCACCCTCCATTACCAGTTGTAAAGAGAAAATGT  
ACTGAAGAATGTGAAATCAATGGTTACGTGATTCCAGAAGGTG  
CTTTGATCTTGTTTAATGTTTGGCAAGTTGGTAGAGACCCTAAA  
TATTGGGATCGTCCTTCTGAATTTTCGTCCAGAACGTTTCTTAGA  
AACTGGTGCTGAAGGTGAAGCAGGTCCATTAGATCTAAGAGG  
TCAACATTTTCAATTATTACCATTTCGGTTCTGGTAGAAGAATGT  
GTCCAGGTGTAACTTAGCAACCTCTGGTATGGCTACTTTATTG  
GCTTCTTTGATTCAATGTTTTGACCTACAAGTCTTAGGTCCACA  
AGGTCAAATTTGAAGGGTGGTGACGCTAAGGTCTCCATGGAA  
GAAAGAGCTGGTTTGACTGTTCCACGTGCTCATTCAATTGGTTTG  
TGTCCATTAGCTAGAATTGGTGTTGCTTCTAAGTTATTGTCTTA  
A

---

*GuIFS*

ATGTTGGTTGAATTGGCTATCACCTTGTTGGTCATTGCTTTGTTC  
ATTCATTGAGACCAACCCCAAGTGCTAAGTCTAAGTCCTTGA  
GACATTTGCCAAACCCACCATCTCCAAAGCCAAGATTGCCATT  
CGTCGGTCACTTGCAATTTGTTGGACAAACCATTATTACATAACT  
CTTTGATCGATTTGTCCAAAAGATACGGACCATTGTATTCTCTT  
TACTTTGGTTCTATGCCAACTGTTGTCGTTTCTACTCCTGAATTG  
TTCAAGTTGTTTTTGCAAACCTCACGAAGCCTCTTCCTTTAATAC  
CAGATTTCAAACCTCCTGCTATTAGAAGATTGACCTACGATAAC  
TCCGTTGCTATGGTCCCATTGTTGTCCTTATTGGAAGTTCATTAG  
AAAGTTGATTATGAACGATCTATTGAACGCCACTACCGTTAAC  
AAATTACGTCCATTGAGATCTCAAGAAATTAGAAAGGTTTTGA  
GAGTTATGGCCTTGTCTGCTGAATCTCAAGTTCCTTTGAACGTT  
ACTGAAGAATTATTGAAGTGACTAATTCTACTATTTCCCGTAT  
GATGTTAGGTGAAGCTGAAGAAATTAGAGATATTGCCAGAGA  
TGTTTTGAAGATTTTTGGTGAATATTCTTTAACTGATTTCATCTG  
GCCATTGAAAAAGTTAAAAGTTGGTCAATACGAAAAGAGAAT  
AGATGATATTTTTAATAGATTTGATCCAGTAATTGAACGTGTAA  
TTAAGAAGAGACAAGAAATCAGAAAAAAGCGTAAGGAAAGA  
AACGGTGAAGTTGAAGAAGGTGAGCAATCCGTCGTTTTCTTG  
ATACCCTTTTGGATTTTCGCTGAAGATGAAACTATGGAAATTAA  
GATTACTAAGGAACAAATCAAAGGTTTGGTTGTTGACTTCTTTT

---

---

CTGCCGGTACTGATTCTACTGCTGTTGCAACAGAATGGGCTTTG  
TCTGAATTGATTAACAACCCTAGAGTTTTGCAAAAAGCAAGAG  
AAGAAGTTGACGCTGTTGTTGGTAAAGATCGTTTGGTTGATGA  
AGCTGATGTTCAAAACTTGCCTTACATTAGATCTATTGTTAAGG  
AAACCTTCAGAATGCATCCTCCATTGCCAGTTGTTAAGAGAAA  
ATGTGTTCAAGAATGTGAAATTGATGGTTACGCTATTCCAGAA  
GGTGCTCTAATTTTGTTCATGTTTGGGCTGTAGGTAGAGATCC  
AAAATATTGGGATAGACCTACTGAATTTTCGTCCAGAAAGATTT  
CTAGAAAATGTAGGTGAAGGTGACCAAGCCGTTGATTTGAGA  
GGTCAACACTTTCAATTATTACCATTTGGTTCTGGTAGAAGAAT  
GTGTCCAGGTGTTAATTTGGCTACAGCTGGTATGGCTACCTTAT  
TAGCTTCTGTTATTCAATGTTTTGATTTGTCTGTTGTTGGTCCAC  
AAGGAAAGATTTTGAAAGGTAATGATGCTAAGGTTTCTATGGA  
AGAATCCGCCGGTCTAACTGTGCCAAGAGCTCATAACTTAGTT  
TGTGTTCCAGTCGCTCGTTCTTCTGCTGTCCCAAAGTTGTTCTCT  
TCATAA

---

*MtHFS*

ATGTTGGTTGAATTAGCCGTTACTTTGTTGTTGATTGCTTTGTTT  
TTACACTTGAGACCAACTCCAAGTCTAAATCTAAAGCTTTAA  
GACATTTGCCTAACCCTCCATCTCCAAAGCCAAGATTGCCTTTT  
ATCGGTCACTTGCATTGTTAGATAATCCATTATTGCATCACAC  
GCTTATCAAGTTAGGCAAGAGATACGGTCCATTGTACACTTTA  
TACTTCGGTTCCATGCCAACTGTTGTTGCTTCTACTCCAGATTTG  
TTTAAGTTGTTCTTACAACTCACGAAGCTACTTCTTTCAACAC  
TAGATTCCAACTAGTGCTATTTCCCGTTTGACTTATGACAATT  
CTGTTGCTATGGTTCCATTGCTCCATACTGGAAGTTCATTAGA  
AAGTTGATTATGAATGATTTGTTAAATGCTACAAGTGTAAACA  
AGTTGAGACCATTGAGATCCAGAGAAATTTTAAAGGTGTTAAA  
GGTTATGGCTAACTCTGCTGAAACTCAACAACCATTGGATGTT  
ACTGAAGAATTACTTAAATGGACAAATTCTACTATTTCTACTAT  
GATGTTGGGTGAAGCTGAAGAAGTTAGAGATATCGCTAGAGA  
TGTCTTAAAAATCTTTGGTGAATACTCCGTTACTAATTTTATTG  
GCCACTAAATAAGTTCAAATTTGGTAACTATGATAAAAGAACT  
GAAGAAATTTTCAACAAGTATGATCCAATTATTGAAAAGGTTA  
TCAAGAAGAGACAAGAAATCGTTAACAAAAGAAAGAACGGT  
GAAATTGTGCAAGGTGAACAAAACGTTGTTTTTCTAGATACTTT  
GTTAGAATTCGCTCAAGATGAAACCATGGAAATTAAGATTACT  
AAGGAACAAATCAAGGGTTTGGTCGTGGATTTCTTTTCCGCCG  
GTAAGTATAGCACAGCTGTTTCTACCGAATGGACCTTGTCCGA  
ATTGATCAACAACCCTAGAGTCTTGAAGAAGGCTAGAGAAGA  
AATCGACTCAGTCGTAGGTAAAGATAGATTAGTTGACGAATCT  
GATGTTCAAACTTACCTTATATTAAGGCTATTGTCAAGGAAG  
CCTTCAGATTGCATCCTCCATTACCAGTTGTCAAAAGAAAGTG  
TACCCAAGAATGTGAGATTGACGGTTACGTTGTGCCAGAAGGT  
GCTTTGATTTTGTTCATGTCTGGGCTGTTGGACGTGATCCAAA

---

---

GTACTGGGTAAACCATTGGAATTCAGACCTGAACGTTTCATT  
GAAAATGTTGGTGAAGGTGAAGCTGCCTCTATTGATTTGAGAG  
GTCAACATTTACATTATTACCATTTGGTTCTGGTAGAAGAATG  
TGTCCAGGTGTTAACTTGGCCACTGCTGGTATGGCCACAATGA  
TTGCTTCTATCATCCAATGTTTTGATTTGCAAGTTCCAGGTCAA  
CATGGTGAAATTTTGAACGGTGACTACGCTAAGGTTTCTATGG  
AAGAAAGACCAGGTTTGACTGTTCCAAGAGCTCATAACTTGAT  
GTGTGTTCCATTGGCTAGAGCTGGTGTTGCTGATAAGTTGTTAT  
CTTCCTAA

---

*PIIFS*

ATGTTATTAGAATTAGCTTTAGGTTTGTTAGTTTTAGCTTTGTTT  
TTACACTTGAGACCAACCCCAACTGCTAAATCTAAAGCCTTGA  
GACATTTGTTGAATCCACCATCTCCAAAGCCAAGACTACCATT  
CATCGGTCATTTGCACTTATTGAAGGATCAATTGTTGCACCATG  
CTTTGATCGATTTGTCTAAGAAGTATGGTCCATTGTACTCTTTGT  
ATTTCCGTTCTATGCCAACTGTTGTTGCCTCTACTCCAGAACTA  
TTCAAGCTTTTCTTGCAAACCTCATGAAGCAACCTCCTTTAATAC  
TAGATTTCAAACCTCAGCTATTAGAAGATTAACCTTACGATAAC  
TCTGTCGCTATGGTCCCATTCGGTCCATACTGGAAGTTCATTAG  
AAAGTTGATTATGAACGATTTGCTTAATGCTACTACCGTTAATA  
AGTTACGTCCATTGCGTACTGAACAAATTAGAAAGTTTCTAAG  
AGTTATGGCCCAATCCGCTGAAGCTCAACAACCATTGAACGTT  
ACTGAAGAATTGTTAAAATGGACTAATTCTACTATCAGTATGA  
TGATGTTAGGTGAAGCTGAAGAAATTAGAGATATAGCTAGAG  
AAGTCTTGAAAATCTTCGGTGAATATTCTGTTACAACTTTATT  
TGGCCATTGATGTATTTGAAGGTTGGTAAATACGAAAAGAGAA  
TCGATGATATTTTAAACAAGTTTCGATCCAGTTATTGAAAGAGTT  
ATCAAAAAGAGAAGAGAAATTGTTAGACGTAGAAAGAACGGT  
GAAGTCGTTGAAGGTGAAGCTTCCGGTGTTTTCTTAGATACTTT  
ATTGGAATTGCTGAAGATGAAACTATGGAAATCAAGATAACT  
AAGGAACAAATTAAGGGTTTAGTTGTTGATTTCTTTTCTGCAGG  
TACTGATTCTACTGCTGTTGCTACTGAATGGGCCTTAGCTGAAT  
TGATTAATAACCCAAGAGTTCTTCAAAAAGCTAGAGATGAAGT  
TACTCCGTTGTCGGTAAAGATAGATTAGTTGATGAAGTTGAC  
ACTCAAACTTGCCATACATTAGAGCTATTGTTAAGGAAACAT  
TCAGAATGCACCCACCATTGCCTGTTGTTAAGAGAAAGTGTGT  
TGAAGAATGCGAAGTTAATGGTTTTGTCATCCCTGAAGGTGCT  
TTAATCTTGTTTAAACGTCTGGGCTGTTGGTAGAGATCCTAAATA  
CTGGGATAGACCACTAGAATTCAGGCCAGAAAGATTCTTAGA  
ATCCGGTGCTGAGGGTGAAGTCGGTCCAATTGACTTGAGAGGT  
CAACATTTCCAATTGTTACCATTCCGTTCTGGTAGACGTATGTG  
TCCAGGTGTTAATTTGGCCACATCTGGTATGGCTACATTGTTGG  
CTTCCTTGATTCAATGTTTTGATTTGCAAGTAGTTGGTCCAAAA  
GGTGAAATTTTGAAGGGAAAAGATGCCAAGGTCTCCATGGAA  
GAATCCGCTGGTTTGACTGTTCTAGAGCTCACAACCTTGAGAT

---

	GTGTTCCATTGGCTAGAACCTCCGTCGCTGCTAAGGTTTTGTCT TAA
<i>TrpFS</i>	ATGTTGTTGGAATTGGCCTTGGGTTTGTGTTGGTCTTGGCTTTGTTCTGCACTTGAGACCAACTCCAAGCTAAATCTAAAGCTTTGAGACATTTGCCAAACCCACCATCTCCAAAACCTAGATTGCCATT CATTGGTCATTTGCATTTGTTGAAGGATAAGTTGTTGCATTACGCTTTGATTGATTTGTCTAAGAAACATGGTCCATTATTTTCATTAT ATTTCCGTTCTATGCCAACTGTCGTTGCTTCTACTCCTGAATTGT TTAAATTATTCTTACAAACCCATGAAGCTACTAGTTTAAACACT AGATTTCAAACCTTCAGCTATTAGAAGATTAACTTATGATTCTTC TGTTGCCATGGTTCCAATTGGTCCATATTGGAAATTCGTTAGAA AGTTAATTATGAATGACTTGTGAACGCTACTACTGTTAACAA GTTAAGACCATTAAAGAACGCAACAAATTAGAAAGTTTTTAAG AGTTATGGCACAAGGTGCTGAAGCTCAAAAACCACTAGATTTG ACTGAAGAATTGTTAAAATGGACCAATTCTACCATTTCCATGA TGATGTTAGGTGAAGCTGAAGAAATTAGAGATATTGCTAGAGA AGTTTTAAAGATCTTCGGTGAATATTCCTTGACTGATTTCAATT GGCCTCTTAAACATTTGAAGGTTGGCAAGTACGAAAAGAGAA TTGACGATATTTTAAACAAGTTTGATCCAGTTGTTGAAAGAGTT ATTAAGAAGCGTAGAGAAATTGTTAGACGTCGTAAGAACGGT GAAGTTGATGAAGGTGAAGTTTCTGGTGTTCCTTGGATACTTT GTTAGAATTTGCTGAAGATGAACTACTGAAATCAAAATCACA AAGGATCATATTAAGGGTTTAGTTGTTGACTTTTTCTCTGCTGG TACTGATTCTACCGCTGTTGCTACTGAATGGGCTTTAGCTGAAT TAATTAATAACCCTAAAGTTTTAGAAAAAGCTAGAGAGGAAG TTTATTCTGTTGTTGGTAAAGATAGATTAGTTGATGAAGTTGAT ACTCAAACTTGCCATACATTAGAGCCATCGTCAAGGAACTT TTAGAATGCATCCACCATTGCCTGTCGTTAAGAGAAAATGTAC TGAAGAATGTGAAATCAATGGTTACGTCATTCCAGAAGGTGCA TTGATTTTGTTCACGTTTGGCAAGTTGGTAGAGATCCTAAATA CTGGGATAGACCAAGTGAATTCGTCCTGAAAGATTTTGTAGAA ACTGGCGCTGAAGGTGAAGCTCGTCCATTGGACTTAAGAGGTC AACATTTTCAATTGTTACCTTTTGGTTCTGGTAGAAGAATGTGC CCAGGTGTTAACTTGGCTACTTCTGGAATGGCTACTTTGTTGGC TTCTTTAATTCAATGTTTTGATTTGCAAGTTTTAGGTCCACAAG GTCAAATTTTGAAGGTGGTGACGCTAAAGTTAGTATGGAAGA AAGAGCAGGTTTGACTGTTCCAAGAGCTCATTCTTTGGTTTGTG TTCCATTGGCTAGAATTGGTGTGCTTCTAAGTTGTTGTCTTAA
<i>GmHID</i>	ATGGCTAAGGAAATCGTTAAGGAATTGTTGCCATTGATCAGAG TTTACAAGGATGGTTCTGTTGAAAGATTGTTGTCTTCAGAAAAC GTTGCTGCATCACCAGAAGATCCACAACTGGTGTTCCTTCAA AGGATATCGTTATCGCTGATAACCCATACGTTTCTGCAAGAAT TTTCTTGCCAAAATCACATCATACAAACAATAAGTTGCCAATT TTCTTGTAACCTTCATGGTGGTGCTTTCTGTGTTGAATCTGCTTTT



	<p>CTTCTTTGTTTCATAGATACTTAAACATCTTGGCTTCTGAAGCA  AACATCATCGCTATCTCAGTTGATTTTCAGATTGTTGCCACATCA  TCCAATTCCAGCTGCATATGAAGATGGTTGGACTACATTGAAG  TGGATCGCTTCTCATGCAAACAACACTAACACTACAAATCCAG  AACCATGGTTGTTGAACCATGCTGATTTCACTAAGGTTTACGTT  GGTGGTGAAACATCTGGTGCTAACATCGCACATAAATTTGTTGTT  GAGAGCTGGTAACGAATCATTACCTGGTGACTTGAAAATTTTG  GGTGGTTTATTGTGTTGTCCATTTTTCTGGGGTTCTAAACCAATT  GGTTCAGAAGCTGTTGAAGGTCATGAACAATCTTTGGCAATGA  AAGTTTGGAATTTTGCTTGTCCAGATGCACCAGGTGGTATTGAT  AATCCATGGATTAATCCATGTGTTCCAGGTGCTCCATCTTTGGC  TACTTTGGCATGTTCAAAGTTGTTGGTTACTATCACTGGTAAAG  ATGAATTCAGAGATAGAGATATCTTGTACCATCATACAGTTGA  ACAATCTGGTTGGCAGGGTGAATTACAATTGTTTGATGCTGGT  GACGAAGAACATGCATTCCAATTGTTTAAACCAGAAACACATT  TGGCTAAGGCAATGATCAAGAGATTAGCTTCATTTTTGGTTTAA</p>
<i>GeHID</i>	<p>ATGGCTTCTTCTACTTCTACTACCACTTCTAAAGAAATTGATAG  AGAATTACCACCATTGTTGAGAGTTTACAAGGATGGTACCGTT  GAAAGATTCTTGGGTTCCTCTTTTGTTCCTCCTTCTCCAGAAGA  TCCAGAAACCGGTGTTTCTACTAAGGATATCGTCATTTCTGAA  AACCCAACTATCTCTGCTAGAGTTTACTTGCCAAAGTTAAACA  ACACTACTGAAAAGTTGCCAATTTTAGTTTATTACCACGGTGGT  GCTTTTTGTTTAGAATCTGCTTTTTCTTTCCTTCATCAAAGATAC  TTAAACATTGTTGCTTCTAAAGCTAACGTTTTGGTAGTCTCTATT  GAATATAGATTAGCTCCAGAACATCCATTGCCAGCAGCTTACG  AAGATGGTTGGTACGCTTTGAAGTGGGTACCTCTCACTCTACC  AACAAACAACAAGCCTACTAACGCTGATCCATGGTTGATTAAGC  ACGGCGATTTC AACAGATTTTACATTGGTGGCGACACCTCTGG  TGCTAATATTGCCCATAAATGCCGCCTTGAGAGTTGGTGCTGAA  GCTTTGCCAGGTGGTTTAAAGAATTGCTGGTGTTTTATCTGCTTTC  CCATTGTTTTGGGGTTCCAAGCCTGTTTTGTCCGAACCAGTTGA  AGGTCATGAAAAGTCCTCCCCAATGCAAGTTTGGAATTTTCGTT  TATCCAGATGCTCCAGGTGGTATTGATAATCCATTGATTAACCC  ATTGGCTCCAGGTGCTCCAACTTGGCTACATTGGGTTGTCCA  AAGATGTTGGTTTTTCGTCGCTGGTAAAGATGACTTGAGAGATA  GAGGTATTTGGTACTACGAAGCTGTAAAGGAATCTGGTTGGAA  GGGTGACGTTGAATTGGCTCAATATGAAGGTGAAGAACATTGT  TTCCAAATCTACCATCCAGAAACTGAAAACCTCTAAGGATTGA  TTGGTAGAATTGCTTCTTTCTTGGTTTAA</p>
<i>PIHID</i>	<p>ATGGCTAACGAAAACCTCTAATAAGGAAATTGTAAAGGAAGTTT  TGCCATTGATTAGAGTTTACAAGGATGGTACTGTTGAAAGATT  ATTGTCTCTCCAAACGTTGCTGCTTCTCCAGAAGATCCAGAA  ACTGGTGTTAGTTCTAAGGATATTGTTATTGCTCATAACCCATA  TGTTTCTGCTAGAATCTTCTTGCCTAACATCAATAAGTCTCATA</p>

---

ACAAGTTGCCAATCTTCGTTTACTTTTCATGGTGGTGGCTTTCTGTG  
TTGAATCTGCTTTCTCTTTCTTTGTTTCATAGATACTTGAATATTTT  
AGCTTCTCAAGCTAATATTATCGCCGTTTCTGTTGATTTTCAGATT  
GTTGCCACATCATCCTTTGCCAGCCGCTTATGAAGATGGTTGG  
ACTACTTTGCAATGGATCGCCTCACATGCTAACAACACTGCTA  
CTAATCCAGAACCATGGCTATTGAACCATGCTGATTTTAACAA  
GTTGTACGTTGGTGGTGAACTTCTGGTGCTAATTTGGCTCATA  
ACTTGCTATTGCGTGCTGGTAACGGTAATCAATCCTTGCCAGGT  
GACTTAAAGATTTTGGGTGGTTTGTGTGTGTGCCATTCTTCTGG  
GGTTCTAAGCCAATCGGTTCTGAACCAGTTGATGAACATGAAC  
AATCCTTAGCTATGAAGGTTTGGAACTTAGCTTGTCCAGATGCT  
CCAGGTGGTATTGATAATCCATGGATTAATCCTTGTGTTGCTGG  
TGCTCCTTCTTTAGCCACCTTGGGTTGTTCTAAGTTGTTAGTTAC  
TATTACTGGTAGAGATGAGTTTAGAGATAGAGATATTTTGTATC  
ATGATACCGTTAAAAAATCAGGTTGGGAAGGTCAATTGGAATT  
GTTTCGACGCCGGGGATGAAGAACATGCCTTCCAATTGTTCAAA  
CCAGAACTGATACTGCTAAGGCTATGATCAAGAGATTAGCTT  
CCTTCTTGGTTTAA

---

*GmCPR*

TTACCAAACATCTCTCAAGTATCTACCAGTAGTTTGCAAGTTCT  
TAACCATAGATTTCAGCTTTACTACTATCTAGTGAACCTTGTCT  
TGCAAGATAGTATGTAGGGCTCTATGAACATCACGTGCCATAC  
CCTTAGCATCTCCACAAACGTAGATGTAAGCACCTTGAGAAAT  
CATGGACCAAATTCAGAAGCCTTTTCCATCATCTTGTGTTGGA  
CGTATTCTTTGTTGGACCTTCTCTGCTGAAAGCTAAGATCAAT  
TCAGATAAAGCACCTGTATTGACGAAATGGTTCAATTCATCTT  
CATAGATATAATCCATTTGTCTATTTCTACAACCGAAGAATAA  
AACAGAAGGACCCAATTCAGCACCACCACCCTTTAAGGCCAA  
TCTTTCTTGTAAGAAACCCCTAAAAGGGGCCAAACCGGTACCT  
GGACCAATCATAATGATTGGAACCTTATTATCTGCTGGTAATCT  
GAAATTGCTAGTACGGACGAAAATTGGAGCCCAAGAGCAGTC  
TTGGGATTTTTCTAATGGGACAGAGTTCTTCATCCAAGTAGAAC  
AAACACCCTTATGAATTCTACCAGTTGGCATCTTATCATGAAC  
AAAGCGCAGGTAACATGAATTCTGTTAGGAACCATCTTGGGG  
AGGAAGAGATAGAATAAAATCTAGGTTGTAACCTAGGTGCGA  
CGGCGGCGAAAAAGACTCCAATTGGTGGCTTAGCTGATGGGA  
ACTCAGCCATAACTTCCAATAATGATCTTTGGGAAGTGATGAC  
CCATTCGGAGTATTCATCTTTACCAGCCGGGGAAGCCAAATGT  
CTCAATCTATCAGCTTCAGATGGATCAGAAGCATGAGCAGCCA  
AGGCTAATAAAGCAGACTTTTTAGGAGAAGATAAAACGTCAG  
CGTATTGAGCTAAAGCTTTACGTAAAGTACATGGTGGGAATGT  
TGGTGGTAAGGAAGAACCACCTAATGGCTTGCCATCTTCATCA  
TCAGTATGAATAGAGAAGTACGTGTCTGGGGATAAACCAATC  
AATCTAATAGCTTCTTCAACGGTTTCACTTAAGTTTTACAGTA  
AACACCGACATGGTCACCGGTTTCGTAAGTAACACCGGTACCG

---

---

GAAATGTCAAATTCTAAGTGAGTACATGATCTATCTGAAGCGG  
GGGTGTGCAATTCTTTTCTAACAGCAACGTTAGCACGAACTGG  
ATGTTGAGCGTCAACAATGGCGTGACCGTTAACATTATGCCAT  
TTCTTTTCATCAACTGAAGCTTCCAATGGATCATGAATAACGAC  
TCTGTATTCTAAGACAGCAGCGGTGTATGGAGTGGAACAGTA  
GTATCATCTTCATCTCTTAGCAAATCGTCCAAGGCTGGCCACA  
ATTCTTCCTTCCAAGCGGTAAAATCATCTTCAATACATTGGTCA  
TCATCTCCTAAACCAACTGGAATAATCTTTTACCACCTTGTTT  
AGCCAACATATCATCTACGACCTTGGCAACCTTGTTAAAGTGT  
TCATATTGACGATTACCTAGACCGAAAACACCATAGTGCAAGT  
TTTCTAACCAACCTTCACCTTTTACCACCTTCAGTGAACCAC  
TTGTAGAAACGAGCAGCATTATCAGTGGGTTACCGTCACCGT  
AGGTGGCTAAGAAAAATAAAGCGTGAGTTTCTTTTTTGAAACG  
TTCTTCATATTCATCATCATCAGCAGCGTAATCATCCATATCAA  
CCACTCTAAAAGTGGCTTTATCATATCTAGCCTTAGCTTCTTCA  
GCTATAGCTTTAGCGAAACCTTCAGCTGTTCCAGTTTGAGTACC  
GAAGAAAATAGTAACTTTTTTAGTACCATCATCAACCTCAATTT  
CAGGTAATTTTTCAATAACTCTCTTTGGTGGTTCTAAAGGTTTA  
GCTTTTGGAGAGGAAGATCTTCTCCAAATAAAAACAACGAAA  
CAACCAACCAAAACGGCTACAGAAGTAGTCAATAACATTAAG  
AATTGTCTATTTTCTAACAAAACGGCAGCAGCTGCATTAGAAG  
AAGAAGATGAAACGTTAGATGGATCCAATTTACCCTTAATAAT  
AGCAGACATCAAATCTAATGGGGAAATCTTCATGGAACCAGA  
ATCTTGCAT

---

*CrCPR2*

TTACCAAACATCTCTTAAGTATCTACCTGTCATTTGCAAGTTCT  
TAACGAAACCTTCAGCTTGAGTTGAATCCATAGAACCTTGTTCT  
TGTGCAATTGTATGCAAAGTTCTATGAACATCTCTTGCCATACC  
CTTAGCATCACCACAAACGTAAACATAAGCACCATCAGAGAT  
CATTCTCCAGATATCTGAAGCCTTTTCTGCCATCTTATGTTGAA  
CGTATTGCTTAGTTGGACCTTCTCTAGAGAATGCAACCAACAA  
TTCTGACAAAGCACCGATTTCCAAGAAATGGTTCAATTCATCTT  
CGTAGATGTAATCCATCTTTCTGTTTCTACAACCGAAAAAGAA  
AACTGCTGTACCCAATTACGACCTTCTTCTTTTAAAGCCAATC  
TTTCTTGCAAGAAACCTCTAAATGGTGCCAAACCAGTACCTGG  
ACCAATCATAATAACTGGAACCTTTGGATCAGCTGGCAACTTG  
AAGTTAGATTGTCTAACGAAGATTGGAGCCCATGAACAATCTC  
TAGATTCTTCCAATGGAATTGCATTTTTTCATCCAAGTTGAACAA  
ACACCTTTATGAATTCTACCACCTGGTGTTTTTTCGTAAACCAA  
AGCACAAGTAACATGAATTCTAGATGGTGCCATTCTTGGTGAA  
GATGAGATTGAGTAGAATCTTGTTGCAATCTTGAGCAATTG  
CAGCAAAGAAAACACCCAATGGTGGTTTTGCAGATGGAAATT  
CAGCCATAACTTCCAACAAAGATCTTTGGTTTGCAACCAATGA  
TTGAGCGTATTCATCTTTACCAGCTGGTGATGCTAAGTACTTCA  
ATCTATCTGCTTCGTTTGGATCAGAAGCGTATGCAGCCAAAGC

---

---

TAACAATGCAGATTTTTTAGGTGTGTTCAACAAATCTGCGTATC  
TTGTCAAAGCAGTTCTCAATGTACATGGTGGAAATGGTGGTGG  
TAAAGATGAACCAGCCAATGGTGTACCATCTTCTTTATCTGCAT  
GCAAAGAGAAGTAAGTTTCTGGTGGCAAGTTCAACAATCTTTC  
AGCTTCTTCAACAGTTTCAGACAAGTTATCACAGTAAACACCA  
ACATGATCACCTGTACCATATGATAAACCAGTACCAGCAATAT  
CGAAATCCAAATGTGTACATGATCTATCAGAAGCTGGAGTATG  
CAATTCCTTTCTAACTGCAACGTTAGATCTACATGGATGTTGAG  
CATCGTAAACTGTGTTACCGTTTTCGTAACCGTTAGCATGACC  
ATTTGCTTCAGAGATCAATGAATCAGACTTATCAGGGAAAACA  
ACTCTGTATTCTGGAATTGCAGCAGTGTATGTAGTTGAACTGT  
AGTATCATCTTCATCTCTCAACAAGTTATCCAATTCTGGCCAAA  
CATTTTCTCTCCATGCAGCAAAATCATCTTCGATACATTGATCA  
TCATCACCTAAAACCAATGGAACAATTCTTTTACCACCTTGTTT  
TGCAACTTTTTTCATCAACAACCTTTAGCAATCTTGTTGAAATGTT  
CGTATTGTCTGTTGCCTAAACCGAAAACACCGTATTGCAAGTT  
CTTCAACCAATCACCACGATCGTTACCCTCAACGAACCACTTG  
TAGAATCTTGCAGCATTATCTGTTGGTTCACCATCACCATAAGT  
TGCCAAGATAAAGAAAGCCAATGTTTCCTTTCTGAACTTTTCTT  
CGTATTCTTCATCATCTGCAGCATAATCATCGATATCGATAACC  
TTGATAACAGCTTTTTTCATATCTAGCTTTTGCTTCTTCAGCCAAT  
GCTTTAGCAAAACCTTCTGCAGTACCTGTTTGAGTACCAAAGA  
AAATAGTGAATTTTTTCTTACCTTCATCGATTTCTTCTGGTTCAA  
CAACAGACTTTGGAACGATCAACTTTGGTGGTTCAACAACCTTT  
TTTACCTGAACCAGATGATCTTCTCCAGATCAAAACAACAACA  
CAACCGATCAAAACAGCAACAGATGTAGTCAAGATCATAACC  
AATTCCTTGTTTTCCAACAACATAGCCATAACTGCTGGAGAAA  
CAGCAACACCTGAATCAGATGAGTTGGAACCATCCAACCTTTC  
ACCTTTCAAGATAGCAGACATCAATTCGAATGGTGACAACCTT  
TCAGATGAAGAATCCAT

---

*AtCPR*

TTACCAAACGTCACGCAGGTAACGGCCAGAGGTCTGCAGGTTT  
TTCACGAAACCTTCCGCTTTGGTAGAATCCATAGAACCTGTTC  
TTGCGCGATGGTGTGCAGAGAACGGTGAACATCACGCGCCAT  
ACCTTTTGCATCGCCACAAACGTACAGGTACGCGCCCTGAGAA  
ATCATGTTCCAGATATCGCTTGCTTTATCCATCATTTTGTGCTGA  
ACATATTCTTTGGTCCGACCTTCACGGCTGAACGCGACGGACA  
GTTTCGGCCAGCGCACCGCTTTCAACAAAGCGCTGCAGTTCTTC  
TTCGTAGATGAAATCCATACGGCGGTTACGGCAGCCGAAGAA  
CAGAACGGACGGACCCAGTTCAACGCCAGATTCAACCAGCGC  
CAGACGTTCTGCAGGAAGCCACGGAACGGAGCCAGACCGGT  
ACCCGGACCGATCATAATAATCGGAACCTTAGAATCAGACGG  
CAGTTTAAAGTTGCTCTGGCGCACGAAGATCGGCGCGGAGGA  
GCAGTTTTCAGATTTTTCGTATGGAACCGCGTTTTTCATCCAGG  
TGCTACATACGCCTTTGTGGATACGACCAGTCGGCATTTTTTCA

---

---

TACACCAGCGCGCAAGTAACGTGAATGCGGGTTTCAGCGATTT  
TCGGGCTGCTAGAGATGGAATAGAAACGCGGCTGAAGGCGCG  
GAGCCACGCCAGCGAAAAACACGCCAGCGGCGGTTTGGCGC  
TCGGGAATTCCGCCATAACTTCCAGCAGAGAGCGCTGGCTTTC  
CACCACCCATTTAGAGTATTCGTCTTTGCCTGCCGGAGAGGCC  
AGGTGTTTCAGACGTTCCGGCTTCGGTTGGGTCGGATGCGTGCG  
CCGCCAGCGCTACCAGCGCGGATTTTTTCGGGGAGCTCAGCAG  
GCATGCGTAGCGGGTCAGCGCAGTGCGCAGGTTGCACGGCGG  
GAACGGTGGCGGGAGGGAGGATGAGATCGGGGTACCATCTTC  
TTTTTCAGCATGCAGACTGAAATAGGTATCCGGAGACATGTCC  
AGCAAACGCAGAGCTTCATCAACGGTTTCAGACAGGTTGTAC  
ACAGAACGCCAACGTGGTCACCGGTTTCATAGGTCAGACCGCT  
GCCCCGAATGTGCAATTCAGGTGGATACAGGAGCGGTCAGA  
TTCTGGGGTATGCAGTTCACGTTTAACGGCCACGTTGCTTTGT  
ACGGATGCTGTGCGTCGAACACAGTGTAGCCGTTGCCGTTGCG  
CATGTTAATGTCGTTGAATTTAGCATCTTCGCTATCATGGATGC  
TAACACGGTATTCCAGAACCGCAGCGGTATACGGGGTTGCCAC  
GGCGGTATCACCTTCTTCACGCAGGATCGTATCCAGTTCCGGC  
CACAGCGCTTCGCGCCAAGCGGTGAAGTCATCTTCGATGCACT  
GGTCATCATCACCCAGGCCTACCTGAACCAGACGCTGTGCGCC  
CTGTTCAACCAGGATATCGTCTACCACTTTCGCAACTTTGTTGA  
AGTGTTCACTACTGGCGGTTACCCAGGCCAAACACACCGTATTT  
CAGGTTTTTTCAGCCATTACCCAGATCGTTACCTTCAGTGAACC  
ATTTATAGAAACGAGCCGCGTTATCAGTCGGTTCGCCGTCACC  
GTAGGTAGCCAGGAAGAAAAACGCAACATCTTCTTTTTTTCAGT  
TTTTCTTCATATTATCATCATCTGCCGCGTAATCGTCCAAGTC  
AACGATTTTGAAGCGGGTTTTTTTCATAACGAGCTTTGGCTTCTT  
CACCCAGCGCTTTCGCGAAACCTTCCGCCGTGCCGGTCTGAGT  
GCCGAAGAAGATGGTAACTTTTTTACGGCCATCATCGATTTCTT  
CTTCACGCGGTTTGATCACCCAGCGGTTCAACGGTTCAACACG  
TTTGAGTTACCGCTACCGCTACGGCGCCAAACCAGCATAACG  
ATGCAACCGATCAGAACCGCGATGCTGGTGGTCACGATCATA  
GCGAACTGACGGTTTTTCGATCAGCATAGAGCTCAGTTCAGCCG  
CTACAGATTTCGTATGCGCTCGCGTTAGCCGGGTCAGACACGAT  
CACCGGTTACCTTTGATGATCGCCGCCATCAGATCAATCATG  
CTGGTAGAAGAAGAAGAGCTAGAGCTCAT

---

*LjCPR*

TTACCAGACATCCCTCAAGTACCTACCGTTCAATTGCAGGTTCT  
TGACCATACCTTCAGCCTTAGAAGAGTCCAAAGAACCCTGTTC  
TTGCAGGATAGTGTGCAAGGTTCTGTGAACATCTCTAGCCATA  
CCCTTAGCATCACACAGACGTAGATGTAAGCACCTTGGGAGA  
TCATGTTCCAGATGTCAGAAGCCTTTTCCATCATCTTGTGCTGG  
ACGTATTCTTTAGTTGGACCTTCCCTGGAGAAGGCAACGATCA  
ATTCAGACAAAGCACCAGAGTTAACGAAATGGTTCAATTCATC  
TTCGTAGATGTAATCCATTTGTCTGTTTCTACAGCCGAAGAACA

---

---

AAACAGATGGACCCAATTACAGCACCATCCTCCTTCAAAGCCA  
GTCTCTCTTGCAGAAAACCTCTGAATGGAGCCAAACCAGTGCC  
TGGACCGATCATGATAATTGGAACCTTGTTGTCGGCTGGCAAC  
TTGAAGTTAGATTGTCTGACGAAGATTGGAGCCCAAGAACAAT  
CTTGGGACTTCTCCAATGGGACAGAGTTCTTCATCCATGTGGA  
ACAGACACCTCTATGGATTCTACCTGTTGGCATTTTGTCTGTTGA  
CCAGAGCACAGGTAACATGAATCCTAGATGGGGCCATTCTTG  
AGAAGAAGAGATGGAGTAAAATCTTGGTTGCAGTCTTGGAGC  
GATAGCAGCGAAAAAACGCCAATTGGTGGCTTAGCAGATGG  
AAATTCGGCCATAACCTCCAGCAATGATCTTTGAGAGGCGATA  
ACCCACTCAGAGTACTCATCTTTACCAGCTGGAGAAGCCAAAT  
GCCTCAACCTATCAGCTTCAGATGGGTTAGAAGCATGAGCAGC  
CAAAGCCAACAAGACAGACTTCTTTGGGGAGGACAAAACATC  
AGCGTATCTAGCGATAGCGGTTCTCAAAGTACATGGTGGGAAA  
GTTGGTGGCAAAGAAGAGCCAGACAATGGCTTACCATCTTCAT  
CATCAGTATGAACAGAGAAGTAAGTATCTGGGGATAAACCCA  
ACAGTCTGACAGCTTCTTCGACTGTTTCAGACAGGTTCTCGCAG  
TAAACACCAACATGATCACCAGTCTCGTAAGCAACACCAGTA  
CCAGAGATATCGAACTCCAGATGGGTGCATGATCTATCAGAA  
ACTGGGGTATGCAGCTCCTTTCTGACAGCAACATTGGATCTGA  
CTGGATGTTGAGCGTCAACAATAGCATGGCCGTTAACGTTGTG  
CCACTTCTTTTCATCAACAGAAGCGTCTAATGGGTCGTGGATA  
ACGACTCTGTACTCCAAAACAGCAGCAGTGTATGGTGTAGAG  
ACTGTAGTGTCTCATCACCTCTCAACAACCTCATCCAAAGCTG  
GCCACAATTCTTCCTTCCAAGCAGTGAAATCATCTTCGATACAT  
TGATCATCGTCACCCAAGCCAACCTTAACCAGTCTTTTACCAC  
CGAAATCGGCCAACTTGTCATCAACTTCGATAGCAACCTTGTT  
GAAATGTTTCGTATTGTCTATTACCCAAGCCGAAAACGGCGTAT  
TCCAAGTTTCTCAACCAACCTTCTTCCTTTTCGTCACCTTCCAGG  
AACCCTTGTAGAATCTAGCAGCGTTATCAGTTGGTTTACCAT  
CGCCGTAAGTAGCCAAGAAGAACAAGCCAAAGTTTCTTCTT  
CAACTTTTCTTCGTATTCATCATCATCTTGAGCGTAATCATCCAT  
ATCAACGATCTTGAACCTTAGCTTTCTCGTACCTAACCTTGGCTT  
CTTCAGCAATAGCCTTAGCAAAAACCTTCGGCAGTACCAGTTTG  
AGTACCGAAGAAGATAGTAACCTTTCTAGTACCATCATCAACT  
TCAGCTTCATCAGACAACCTTTCAACAACCTCTTTTGGTGGTTC  
GATGGACTTGGCCTTGTTACCAGTTGATCTTCTCCAGATGAAAA  
CAACAACACAACCGATCAAAAACAGCGATAGAAGTAGTCAAA  
ACCATAACAAATTCTCTGTTTTCCAAAAAACAGAACAGCAC  
CGGAGGTAGAAGAAACGTTAGATGGGTCCAAGGTGCCCTTAA  
TCATAGCAGACATCAGATCCAGGGGACTGATCTTCATAGAAG  
AGGATTCCTCCAT

---

*EcPntAB*

TGGAAGGGAATATCATGCGAATTGGCATAACCAAGAGAACGGT  
TAACCAATGAAACCCGTGTTGCAGCAACGCCAAAAACAGTGG

---

---

AACAGCTGCTGAAACTGGGTTTTACCGTCGCGGTAGAGAGCGG  
CGCGGGTCAACTGGCAAGTTTTGACGATAAAGCGTTTGTGCAA  
GCGGGCGCTGAAATTGTAGAAGGGAATAGCGTCTGGCAGTCA  
GAGATCATTCTGAAGGTCAATGCGCCGTTAGATGATGAAATTG  
CGTTACTGAATCCTGGGACAACGCTGGTGAGTTTTATCTGGCCT  
GCGCAGAATCCGGAATTAATGCAAAAACCTGCGGAACGTAAC  
GTGACCGTGATGGCGATGGACTCTGTGCCGCGTATCTCACGCG  
CACAATCGCTGGACGCACTAAGCTCGATGGCGAACATCGCCG  
GTTATCGCGCCATTGTTGAAGCGGCACATGAATTTGGGCGCTT  
CTTTACCGGGCAAATTACTGCGGCCGGGAAAGTGCCACCGGC  
AAAAGTGATGGTGATTGGTGCGGGTGTTGCAGGTCTGGCCGCC  
ATTGGCGCAGCAAACAGTCTCGGCGCGATTGTGCGTGCAATTCG  
ACACCCGCCCCGGAAGTGAAAGAACAAGTTCAAAGTATGGGCG  
CGGAATTCCTCGAGCTGGATTTTAAAGAGGAAGCTGGCAGCG  
GCGATGGCTATGCCAAAGTGATGTCGGACGCGTTCATCAAAGC  
GGAAATGGAACCTTTGCCGCCAGGCAAAGAGGTCGATAT  
CATTGTCACCACCGCGCTTATTCCAGGCAAACCAGCGCCGAAG  
CTAATTACCCGTGAAATGGTTGACTCCATGAAGGCGGGCAGTG  
TGATTGTCGACCTGGCAGCCCCAAAACGGCGGCAACTGTGAAT  
ACACCGTGCCGGGTGAAATCTTCACTACGGAAAATGGTGTCAA  
AGTGATTGGTTATACCGATCTTCCGGGCCGTCTGCCGACGCAA  
TCCTCACAGCTTTACGGCACAAACCTCGTTAATCTGCTGAAAC  
TGTTGTGCAAAGAGAAAGACGGCAATATCACTGTTGATTTTGA  
TGATGTGGTGATTGCGGGCGTGACCGTGATCCGTGCGGGCGAA  
ATTACCTGGCCGGCACCGCCGATTCAGGTATCAGCTCAGCCGC  
AGGCGGCACAAAAAGCGGCACCGGAAGTGAAAACCTGAGGAA  
AAATGTACCTGCTACCGTGCGGTAAATACGCGTTGATGGCGC  
TGGCAATCATTCTTTTTGGCTGGATGGCAAGCGTTGCGCCGAA  
AGAATTCCTTGGGCACTTCACCGTTTTCGCGCTGGCCTGCGTTG  
TCGGTTATTACGTGGTGTGGAATGTATCGCACGCGCTGCATAC  
ACCGTTGATGTCGGTCACCAACGCGATTCAGGGATTATTGTTG  
TCGGAGCACTGTTGCAGATTGGCCAGGGCGGCTGGGTTAGCTT  
CCTAGTTTTATCGCGGTGCTTATAGCCAGCATTAAATATTTTCG  
GTGGCTTCACCGTGACTCAGCGCATGCTGAAAATGTTCCGCAA  
AAATTAAGGGGTAAACATATGTCTGGAGGATTAGTTACAGCTGC  
ATACATTGTTGCCGCGATCCTGTTTATCTTCAGTCTGGCCGGTC  
TTTCGAAACATGAAACGTCTCGCCAGGGTAACAACTTCGGTAT  
CGCCGGGATGGCGATTGCGTTAATCGCAACCATTTTTGGACCG  
GATACGGGTAATGTTGGCTGGATCTTGCTGGCGATGGTCATTG  
GTGGGGCAATTGGTATCCGTCTGGCGAAGAAAGTTGAAATGA  
CCGAAATGCCAGAACTGGTGGCGATCCTGCATAGCTTCGTGGG  
TCTGGCGGCAGTGCTGGTTGGCTTTAACAGCTATCTGCATCATG  
ACGCGGGAATGGCACCGATTCTGGTCAATATTCACCTGACGGA  
AGTGTTCTCGGTATCTTCATCGGGGCGGTAAACGTTACGGGTT

---

---

CGGTGGTGGCGTTCGGCAAACCTGTGTGGCAAGATTTTCGTCTAA  
ACCATTGATGCTGCCAAACCGTCACAAAATGAACCTGGCGGCT  
CTGGTCGTTTCCTTCCTGCTGCTGATTGTATTTGTTTCGCACGGAC  
AGCGTCGGCCTGCAAGTGCTGGCATTGCTGATAATGACCGCAA  
TTGCGCTGGTATTCGGCTGGCATTAGTCGCCTCCATCGGTGGT  
GCAGATATGCCAGTGGTGGTGTGCTGCTGAACTCGTACTCCG  
GCTGGGCGGCTGCGGCTGCGGGCTTTATGCTCAGCAACGACCT  
GCTGATTGTGACCGGTGCGCTGGTTCGGTTCTTCGGGGGCTATCC  
TTTCTTACATTATGTGTAAGGCGATGAACCGTTCCTTTATCAGC  
GTTATTGCGGGTGGTTTCGGCACCGACGGCTCTTCTACTGGCGA  
TGATCAGGAAGTGGGTGAGCACCGCGAAATCACCGCAGAAGA  
GACAGCGGAACTGCTGAAAACTCCCATTCACTGATCATTACT  
CCGGGGTACGGCATGGCAGTCGCGCAGGCGCAATATCCTGTG  
GCTGAAATTACTGAGAAATTGCGCGCTCGTGGTATTAATGTGC  
GTTTCGGTATCCACCCGGTCGCGGGGCGTTTGCCTGGACATAT  
GAACGTATTGCTGGCTGAAGCAAAAGTACCGTATGACATCGTG  
CTGGAAATGGACGAGATCAATGATGACTTTGCTGATACCGATA  
CCGTAATGGTATTGGTGCTAACGATACGGTTAACCCGGCGGC  
GCAGGATGATCCGAAGAGTCCGATTGCTGGTATGCCTGTGCTG  
GAAGTGTGGAAGCGCAGAACGTGATTGTCTTTAAACGTTCTGA  
TGAACACTGGCTATGCTGGTGTGCAAAACCCGCTGTTCTTCAA  
GGAAAACACCCACATGCTGTTTGGTGACGCCAAAGCCAGCGT  
GGATGCAATCCTGAAAGCTCTGTAA

---

*GmUGT4*

TTAATGTTCTCTCCACAATTCAACCAACTTATTCAATGCCATGA  
TTGAAGAACCACCTTTTGCCATAGCTTCAGTAGCTGAAATTTTC  
ATCTTAAAAATTCTTTGTCTGATTTCTTTACCCTTATCAGAAATCC  
ATCAATTCCATAACTCTGTCACCCAATTCTGTTGAAGAACTA  
AACCATCCTTATTTTGCTTAACAGCCAAACCAACCTTCATTCT  
TCAACCAAAATAACCTTATTCAACTTTTGTTCGCTACAATGG  
CCAAGCAACCATTTGGAACACCTTCACAACTGCTTCTAAAACA  
GAATTCCAACCACAATGAGTAACAAAACCACTGAATCA  
TGAGATAAAATTGCAGCTTGTGGAGCCCAATCTCTAACAACCA  
AACCTTTTCCTTTGTTCTTTCCAAGAAACCTTCTGGTAACAATT  
CATCCAATGATGGTGGTTCACCAGAGTCACCCTCTTCGAATTCT  
GATCTAACAACCCATAAAAAATCTTTGTTTCTGATTTTCCAAACC  
GATAGCGATTTCTCTCAATTGTGTTCTAGAAAATCTACCCATTG  
AACCAAAAGACAAGAAAACAACCTGAATGAGATGGTTGTGAAT  
CTAACCAAGACAAACAACCATATCATCTTTTCTACATGGAGC  
TGAAGAAATAACTGGACCAATACAGAAAACCTTTGGTGTAGTA  
CCTTCCATTAAACCTTCTGAAAATGCTTCAACAACCTCTACCTTC  
CATAGCTTCACAAGTATTAACAATAACACCATCAGAATCTCTC  
ATACATGTTGCAATATCAATAAAAACTTGGTAACCTTCGTTAG  
CTCTATCTTGCAATTGTTCTGGCAAATCATCAGTATGGATCTTTG  
GCAAACCTGGAATAACCAAAATGCATATTCAAATCTTTAAAAG

---



---

ACTTTGTAGAGTTTTTCATGGATGATGATTTGTTGCAAGAAAATT  
GCTAAAGTTGAAGCACCAGATGTGTAGTAGAAGTAAGTTGGG  
ATTTGCAATGCGTTAGTAACTCTTGCAGCTGAGTAGTTCATGAA  
ATCCAAAACAATAGCTTTCAAATTAGATGTTTGTGAGATAGAG  
TTCAAGATTCTTCTCAAATGATGACCAGTTGCTCTACACAATTC  
AAATGTTAAAGCCATTGGTGGTAAAACAGTTGGAACAGAAAT  
TTGTGGGATTCTATGGAATGTGATTGATGGTGTAGATGCAGTA  
ACTGCAGCAATGTATTTTGCAGTAGCATCACATGTAAAAGCAG  
TTGGAGATGTTGGAGTATCTTGATTGGTGGTGGAGTCAAGAA  
AATGATAGTGATTGACAAAGATGGATGATGTGACAAGATCAA  
TTACCCAATTCAACCATAGAAACCAAATGACCTCTACCCAAA  
GCTGAGTACAAAACGATAGAAATCCTTCATTGTCAT

---

*PIUGT43*

TTAAGATGTTAATTCTTGAATTAATGAAACCAAATTAGAATAT  
GAAGAACCATTTTCTAACAAAGCACCTCTACAAGTATCTGACA  
TTTCCTTAACCTTTCTTCTGATTTTCGTCACCACCTTTCATCAAAG  
ATCTCAAACCGTTTTCAACTTCTTCTGCTCTAACCAAATCAGCA  
CCAACTCTGTAATCAACTGAAATTTCAACTGCTAAACCCAAAT  
CTCTAACCATTGGAAGCGTTCATTTGTTGTTTCAAGTACAAT  
GGCCATGTTGCAACTGGAACACCATGCCATAATGATTCCAAAA  
TAGAATTCCAACCACAATGAGAAACAAAACCACCAACTGCTT  
TATGAGCTAAAACAACAGCTTGTGGAACCCAACCACAAACCA  
AACCGATGTTGTTAGTTCTTTCCAAGAAACCATCTGGCAAAAC  
ATCCTTATGGTTTTCGTAATCTCTTGGGTATTCCAATTGTGCTTT  
TGGAGATTCTCTTAAAGCCCACAAAAATCTAACACCTGCTCTT  
TCCAAACCGATAGCGATTTCTTCAACTTGGTTAGCTTCCAATGA  
ACCCATAGAACCGAAACACAACAAAACAACCTGAAGACAATG  
GTTGTTGATCCAACCATTCATAATTCTTTTGTATTGTGCTGGAT  
TTGGATCCCATTGAGCAGAACCAACCAAATCCAAAATTGGAC  
CAATTGGATAAACTCTTTGCAATTGTGAATCGTTATGCAAAGA  
TTGCAAAGCATGTGTTTCTAATTCTTGCAAAGTATTAATAACGA  
TACCCTTTGTTTCCTTGTATCTTCTTGCAATGGTATGACAACCA  
AAAGTACCATCTTTAGCATCCAAAACCAAATTTGGGAAAACA  
GATCTTGGCAACAAATTCTTGAAACATGGAACAGTGAATTCTG  
ACTTAGATTCTGCCAAATCGAATCTTGGCAAATCCAAAGTAAA  
ACCCAAACATGATGCTGGAGAAGCGAAAAACAAATAACATG  
GAACAGCCAATTCTGCAGCAACATCAATTAATGTAGTTGAGAA  
CATATCAACGAAGATTGCAGCCAATCTAACTGAATTTGAAGAA  
TTAGATTCTGTAGTCTTCATCAAATTCACAATGCATGCTTAAC  
ATGATGTTTATGGTTTTGCATATGCAAAGACAAGAAACCAACC  
AAAGTTTGGTATTGTTCTGGAGCTGGTGGATCAACGATTGGCA  
ATTGCAACAACCTTGATGTTTGCAGCTGAAGATGCTCTAGCTTG  
AACGTATGTGTTTCATCAATGGTCTTTGTGGCATTGAACTGTCA  
AGATAGTTGCAGAAAATCTTGGATCATGCTTAGTCAACAAATT  
AGCGAATTCAACTTGTGGAACCAAATTACCCAATGTTGGGATA

---

---

GCAATAAAAACAACTTCGTATCTAGTCAT

---

**Table S5 Guide RNA sequences used in this study.**

Site	Guide RNA sequences	Role
<i>XI-5</i>	TGTCCCTGTATACTGCAAAG	Common integration site
<i>XI-2</i>	GGTCCCTAAGGATATCTCGG	Common integration site
<i>EXG1</i>	GATCATGGAAGCTATTCCAG	Glycoside hydrolase site
<i>OCA5</i>	CTTCAGCAGAAATTTACGA	Inositol pyrophosphatase
<i>FAS1p</i>	TTGATAAGGAAAAAGAACCA	Promoter site of the <i>FAS1</i> gene

**Table S6 Plasmids used in this study.**

<b>Name</b>	<b>Relevant characteristics</b>
pY26-1	<i>pY26-SED1p-GeHIS-Ter22-GmCPR-TDH1p-INO1p-GmHID-CYC1t</i>
pY26-2	<i>pY26-SED1p-GmIFS-Ter22-GmCPR-TDH1p-INO1p-GmHID-CYC1t</i>
pY26-3	<i>pY26-SED1p-GmIFS2-Ter22-GmCPR-TDH1p-INO1p-GmHID-CYC1t</i>
pY26-4	<i>pY26-SED1p-GuIFS-Ter22-GmCPR-TDH1p-INO1p-GmHID-CYC1t</i>
pY26-5	<i>pY26-SED1p-MtIFS-Ter22-GmCPR-TDH1p-INO1p-GmHID-CYC1t</i>
pY26-6	<i>pY26-SED1p-PlIFS-Ter22-GmCPR-TDH1p-INO1p-GmHID-CYC1t</i>
pY26-7	<i>pY26-SED1p-TpIFS-Ter22-GmCPR-TDH1p-INO1p-GmHID-CYC1t</i>
pY26-8	<i>pY26-SED1p-GeHIS-Ter22-GmCPR-TDH1p-INO1p-GeHID-CYC1t</i>
pY26-9	<i>pY26-SED1p-PlIFS-Ter22-GmCPR-TDH1p-INO1p-PlHID-CYC1t</i>
pY26-10	<i>pY26-SED1p-GmIFS-Ter22-CrCPR2-TDH1p-INO1p-GmHID-CYC1t</i>
pY26-11	<i>pY26-SED1p-GmIFS-Ter22-AtCPR-TDH1p-INO1p-GmHID-CYC1t</i>
pY26-12	<i>pY26-SED1p-GmIFS-Ter22-LjCPR-TDH1p-INO1p-GmHID-CYC1t</i>

pY26-13	<i>pY26-SED1p-GmUGT4-DIT1t</i>
pY26-14	<i>pY26-SED1p-PlUGT43-DIT1t</i>
pRS424-1	<i>pRS424-SED1p-STB5-TEr22</i>
pRS424-2	<i>pRS424-SED1p-YEF1-TEr22</i>
pRS424-3	<i>pRS424-SED1p-EcPntAB-TEr22</i>
pRS424-4	<i>pRS424-SED1p-ALD6-TEr22</i>
pRS424-5	<i>pRS424-SED1p-HEM2-TEr22-HEM3-TDH3p</i>
pRS424-6	<i>pRS424-TDH1p-PGM1-ADH1t</i>
pRS424-7	<i>pRS424-TDH1p-PGM2-ADH1t</i>
pRS424-8	<i>pRS424-TDH1p-PGM1-ADH1t-INO1p-UGP1-CYC1t</i>
pRS424-9	<i>pRS424-TDH1p-PGM2-ADH1t-INO1p-UGP1-CYC1t</i>
Cas-1	<i>Ca9 gRNA-XI5</i>
Cas-2	<i>Ca9 gRNA-XI2</i>
Cas-3	<i>Ca9 gRNA-EXG1</i>
Cas-4	<i>Ca9 gRNA-OCA5</i>
Cas-5	<i>Ca9 gRNA-FAS1</i>

---