

**Manuscript title:** Disruption of the *Snf1* gene enhances cell growth and reduces the metabolic burden in cellulase-expressing and lipid-accumulating *Yarrowia lipolytica*

**Additional file 1:**

**File type:** .pdf

**List of contents:** **Page no.**

**Supplementary Materials and Methods**

|   |          |
|---|----------|
| <b>Outlines for plasmid construction and pathway engineering</b>  | <b>2</b> |
| <b>Plasmids for <i>ACL</i> and <i>DGA1</i> gene expression in <i>Yarrowia</i></b>   | <b>2</b> |
| <b>Plasmids for HisG-<i>URA3</i>-His flanked by <i>Snf1</i>up and <i>Snf1</i>down, and fused with <i>cbl1-cbh2-eg2</i></b>  | <b>3</b> |
| <b>Genomic DNA extraction, primer design, and PCR</b>   | <b>4</b> |
| <br>  |          |
| <b>Supplementary Table S1.</b> Sequences of forward (F) and reverse (R) primers   | <b>5</b> |
| <b>Supplementary Figure S1.</b> A three-step process for constructing plasmids pYLEX1- <i>ACL</i> , pMT015-TEFin- <i>DGA1</i> and pMT015-TEFin- <i>DGA1-ACL</i>                 | <b>6</b> |
| <b>Supplementary Figure S2.</b> Main steps in building the constructs for <i>Snf1</i> depletion, recyclable <i>URA3</i> marker and cellulase expression in <i>Y. lipolytica</i> | <b>7</b> |
| <br>  |          |
| <b>Nucleotide sequences SEQ NO 1 to 11</b>  | <b>8</b> |

## Outlines for plasmid construction and pathway engineering

All plasmids used in this study are described in Table 1. For the deletion of *Snf1* and the expression of *cbh1*, *cbh2*, and *eg2* genes in *Yarrowia*, the overall approach for plasmid construction in this study can be divided into three parts, as described in below sections: (1) for *ACL1* and *DGA1* gene expression (**Supplementary Figure S1**), (2) for recyclable marker HisG-*URA3*-His flanked by *Snf1* gene's upstream (i.e. *Snf1* up) and downstream (i.e. *Snf1* down) sequences (**Supplementary Figure S2A**), and (3) for fusion with *cbh1-cbh2-eg2* cassette (**Supplementary Figure S2B**).

### Plasmids for *ACL* and *DGA1* gene expression in *Yarrowia*

The constructs for expressing singular *ACL* and *DGA1* genes; as well as *ACL-DGA1*, were built in the backbone of vector pYLEX1 which contains an EXP1 promoter and EXP1 terminator for *ACL* gene, TEFin promoter and XPR2 terminator for *DGA1* gene, respectively, as outlined in Supplementary Figure S1.

**Step 1.** Plasmid pMT015-YTEFin-DGA1 (*i.e.*, construct 203) was constructed on the backbone of plasmid pMT015-YTEFin, which is a derivative of plasmid pYLEX1 with its hybrid promoter (hp4d) being replaced with TEFin (*i.e.*, TEF promoter with intron). The map and sequence information were retrieved from the previously published supplemental materials in literature (Tai and Stephanopoulos, 2013), listed as SEQ NO 4 for the sequence record in this additional file 1. To insert *Y. lipolytica* *DGA1* gene (YALI0E32769g) into plasmid pMT015-YTEFin, we removed the ATG and added TAACCGCAG in the 5' end of the sequence of *Y. lipolytica* *DGA1* gene to complete the intron after digestion with SnaBI. The synthesized *DGA1* sequence (1560 bp; this additional file 1, SEQ NO 5) had a blunt end on 5' and a NsiI site (ATGCAT) on the 3' end, and was cloned into the SnabI/NsiI double cut vector of pMT015-YTEFin. The resultant plasmid was named as pMT015-YTEFin-DGA1 (Supplementary Figure S1).

**Step 2.** Plasmid pYLEX1-*ACL* was constructed by a sequential cloning procedure described below. The related gene, promoter and terminator sequences were synthesized by GenScript with appropriate restriction sites at the 5' and 3' ends for assembling into destiny vector (Supplementary Figure S1), and described as below:

- (1) Clone the synthesized promoter sequence of SalI-pEXP1-PmlII (*i.e.*, construct 205; 1014 bp, this additional file 1, SEQ NO 1) into vector pYLEX1 at the restriction cut sites of SalI-PmlII;
- (2) Clone the synthesized terminator sequence of KpnI-tEXP1-SalI-Clal (*i.e.* construct 206; 517 bp, this additional file 1, SEQ NO 2) into vector pYLEX1 at KpnI-Clal cut sites;
- (3) Clone the synthesized blunt end – *ACL*– stop codon KpnI (*i.e.*, construct 207; 3312 bp, this additional file 1, SEQ NO 3) into vector pYLEX1 at restriction cut sites of PmlII – KpnI. The resultant plasmid was named as pYLEX1-*ACL* (*i.e.*, construct 208, Supplementary Figure S1).

**Step 3.** To generate a single plasmid for co-expressing both ACL and DGA1, a promoter-gene-terminator cassette of pEXP1-*ACL-tEXPI* was cut from the above pYLEX1-*ACL* plasmid using restriction enzyme SalI and inserted into the SalI-cut pMT015-DGA1 with the same orientation. This resultant plasmid was named as pMT015-*TEFin-ACL-DGA1* (*i.e.*, construct 209, Supplementary Figure S1).

### **Plasmids for HisG-*URA3*-His flanked by *Snf1up* and *Snf1down*, and fused with *cbh1-cbh2-eg2***

The scheme for a six-step construct building process is illustrated in **Supplementary Figure S2**. The purpose of this process for building HisG-*URA3*-HisG cassette flanked by *Snf1* upstream (*i.e.* *Snf1up*) and downstream (*i.e.* *Snf1down*) sequences (Supplementary Figure S2A; steps 1-3), followed by fusing with *cbh1-cbh2-eg2* cassette (Supplementary Figure S2B; steps 4-6). The related genes and sequences were synthesized by GenScript with appropriate restriction sites at the 5' and 3' ends for cloning into destiny vectors as described in Table 1.

#### Steps 1 and 2. Synthesis of *Snf1up* and *Snf1down*

*Y. lipolytica Snf1* gene's genomic sequence is 1.74kb region from base 236133 to 237872 ([http://www.ncbi.nlm.nih.gov/nuccore/NC\\_006070.1?from=236133&to=237872&report=fasta](http://www.ncbi.nlm.nih.gov/nuccore/NC_006070.1?from=236133&to=237872&report=fasta)). *Yarrowia Snf1* upstream 1.2 kb (base 234933 to 236132 of *Y. lipolytica* CLIB122 chromosome D complete sequence) and downstream 1.2 kb (base 237873 to 239073) sequences were retrieved from GenBank, and are also listed as SEQ NO 6 and NO 7, respectively, in this additional file 1. These sequences were synthesized with the appropriate restriction sites at the 5' and 3' end (**Supplementary Figure S2**).

#### Step 3. Synthesis of HisG-*URA3*-HisG

In HisG-*URA3*-HisG (*i.e.*, construct 150, SEQ NO 8, this additional file 1), the *URA3* gene of *Y. lipolytica* flanked by direct repeats of *Salmonella typhimurium hisG* DNA (Supplementary Figure S2, Step 3). The 1157 bp fragment *HisG* gene of *S. typhimurium* was obtained from literature (Alani et al., 1987; Voth et al., 2001) and retrieved from GenBank (accession no. AF324729). The promoter, CDS and the terminator of *Y. lipolytica URA3* gene was retrieved from the genomic DNA sequence of U40564 at GenBank. The full length of U40564 is 1710 bp, among which 1-384 nt is promoter region, 385-1245 nt is the CDS of *URA3* gene (from start to stop codons), and 1246-1710 nt is the terminator region. We used 1-1545 nt (in which 1246-1545 nt was the core part of terminator) as the *URA3* cassette to be inserted between two HisG repeats.

#### Step 4. Assembly of *Snf1up*-HisG-*URA3*-HisG-*Snf1down*

For knocking out *Snf1* gene, the 1.2 kb DNA fragments of upstream and downstream nucleotide sequences of *Y. lipolytica Snf1* gene (as synthesized in Supplementary Figure S2, Steps 1 and 2) were added to the 5' and 3' end of HisG repeats, respectively (Supplementary Figure S2, Step 4). The assembled fragment was named as *Snf1up*-HisG-*URA3*-HisG-*Snf1down* (*i.e.*, construct 161, SEQ NO 9, this additional file 1), cloned in vector pUC57-Simple by EcoRV.

### Step 5. *cbh1-cbh2-eg2* cassettes

*cbh1-cbh2-eg2* cassettes (*i.e.*, construct 162, SEQ NO 10, this additional file 1) is a row of three cassettes of *cbh1-cbh2-eg2*, with SalI-PamII site on its 5' end, and KpnI site on its 3' end. This construct was synthesized by our group in a recent study (Wei et al., 2019), in which the construct was built in the backbone of the vector pYLSC1 with the TEFin promoter and XPR2 terminator for Te-Tr (*Talaromyces emersonii-Trichoderma reesei*) chimeric *cbh1* gene, the GPD promoter and the Lip2 terminator for Tr *cbh2* gene, EXP1 promoter and EXP1 terminator for Tr *eg2* gene, respectively, as illustrated in Supplementary Figure S2.

### Step 6. Construct of *Snf1* up-*cbh1-cbh2-eg2*-HisG-*URA3*-HisG-*Snf1* down

The construct of *Snf1* up-*cbh1-cbh2-eg2*-HisG-*URA3*-HisG-*Snf1* down (*i.e.*, construct 163, SEQ NO 11, 14064 bp, this additional file 1) was built by inserting the fragment of PamII-*cbh1-cbh2-eg2*-KpnI (from construct 162) into the site of construct 161 between its *Snf1* up and first HisG, as illustrated in (Supplementary Figure S2).

## Genomic DNA extraction, primer design, and PCR

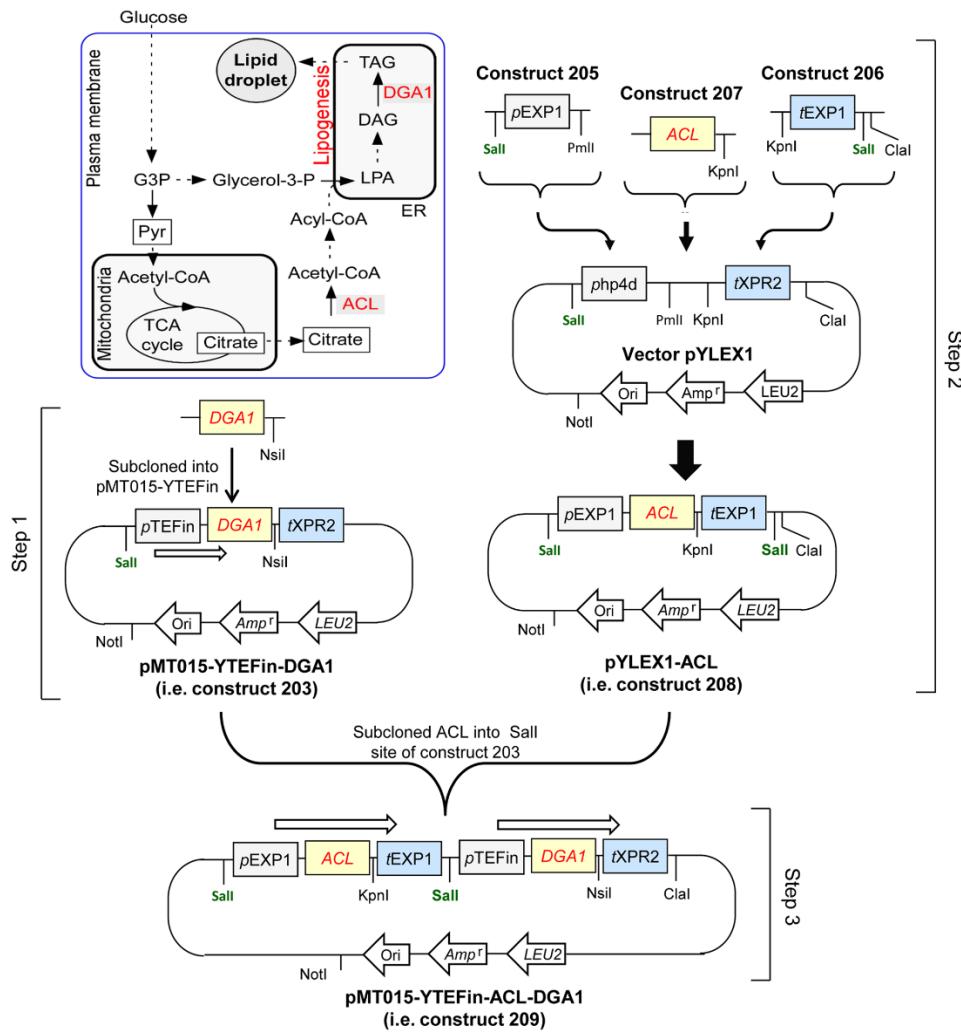
Genomic DNA was isolated from *Y. lipolytica* cell pellets or colony patch by using the ZR Fungal/Bacterial DNA Miniprep kit (cat.# D6005; Zymo Research, Irvine, CA), and by following the procedure described in literature (Xu et al., 2017). The concentration of extracted genomic DNA was determined using Nanodrop, and adjusted to 20 ng  $\mu$ L<sup>-1</sup> and stored at -80 °C until use.

Primers were designed to characterize the modes for the insertion of construct 163 into the host cell genome. The initial screening primers were primers CBHI-F/CBHI-R that designed to align to *cbh1* cassette region of construct 163 and were used to confirm the insertion of construct 163 into the genome without distinguishing random insertion and target insertion. Furthermore, primers 163F8/163R8 and 163F10/163R10 were designed to flank the 5' end and 3' end of construct 163 presumably inserted into the genome at the site of disrupted *Snf1* gene in the desired mutants (see **Supplementary Table S1**, and as illustrated in the corresponding primer design and PCR result figure in the Results and Methods section).

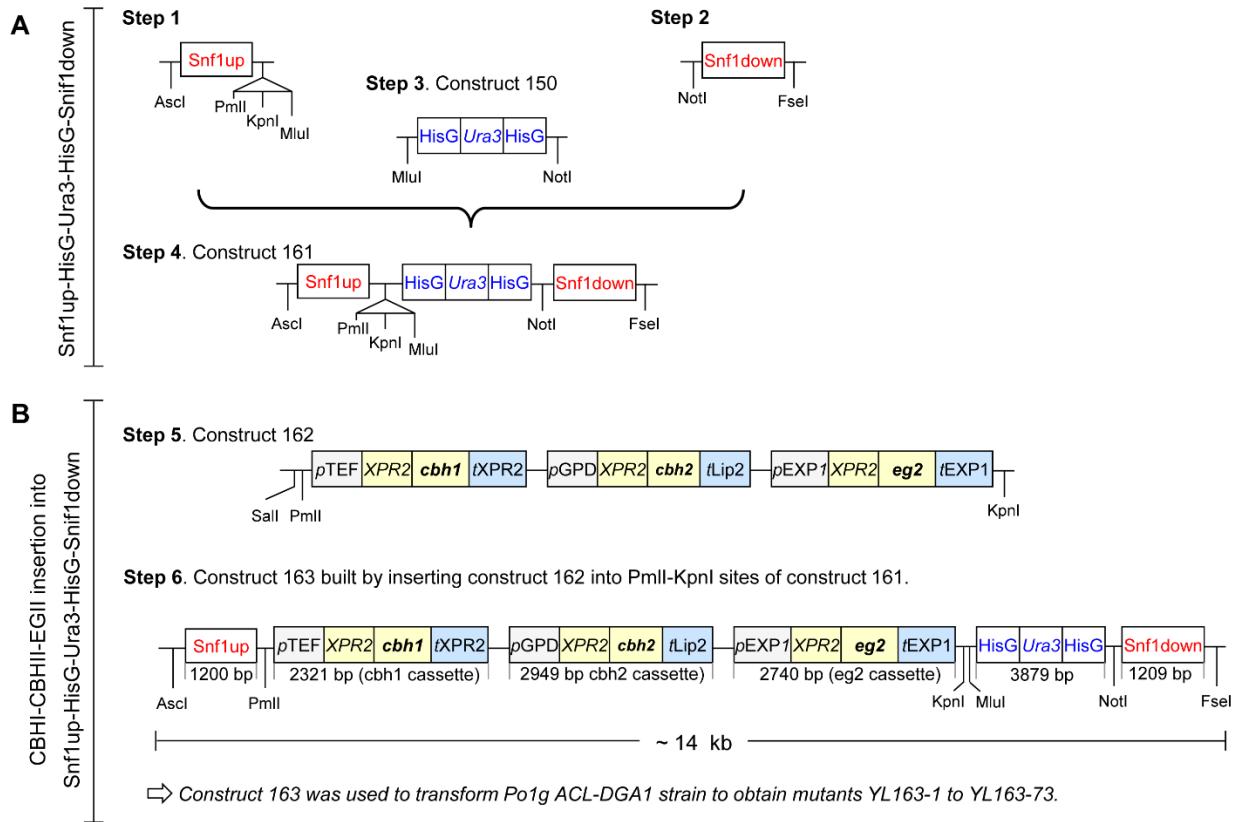
To confirm the knocking-out of *snf1* gene and knocking-in cellulase cassette in mutants. Their genomic DNA samples were PCR-amplified using the designed primers and the Q5 Hot start high-fidelity 2x master mix (New England BioLabs M0494). Based on the manual instruction for Q5 Hot start high-fidelity 2x master mix, thermal cycling conditions consisted of 30 s at 98 oC for initial denaturation, 32 cycles of amplification (10 s at 98 °C, 20 s at 60 °C, 30 s/kb at 72°C), 5 min at 72 °C for final extension, and hold at 4 °C.

**Supplementary Table S1. Sequences of forward (F) and reverse (R) primers.** Primers were used to characterize the random verse target insertions of construct into the *Snfl* gene site in the genome of *Y. lipolytica* strains. An illustration of primer alignment regions can be found in Figure for PCR results in the Results and Discussion section.

| Primer name      | Sequences and primer alignment region   | Use of PCR and amplicon size  |
|------------------|---|---|
| CBHI-F<br>CBHI-R | F: TTCACCGCATCTAACCCACC;<br>align to the <i>cbl</i> cassette region.<br>R: TCCACACCCCCACAAAAAGAC;<br>align to the <i>cbl</i> cassette region. | For initial confirming of construct 163 insertion into genome via either random or target insertion.<br><br>Product size: 517 bp.                       |
| 163F8<br>163R8   | F: TCGTCACCATGTCCTTCAGA;<br>flank 5' end of Snfl up.<br>R: CCCGCTACTGGGTCAATT; align to <i>cbl</i> cassette.                                  | For confirming construct 163's target insertion into genome at the site of disrupted <i>Snfl</i> gene in desired mutants.<br><br>Product size: 1493 bp. |
| 163F10<br>163R10 | F: TATCCGCATGATCTGTCAA;<br>align to <i>URA3</i> nearing the second HisG.<br>R: GAGATCAAAGCCGAAAAATGC;<br>flank 3' end of Snfl down.           | For confirming construct 163's target insertion into genome at the site of disrupted <i>Snfl</i> gene in desired mutants.<br><br>Product size: 3354 bp. |



**Supplementary Figure S1. A three-step process for constructing plasmids pYLEX1-ACL, pMT015-TEFin-DGA1 and pMT015-TEFin-DGA1-ACL.** The roles of ACL and DGA1 in lipogenesis are shown in the left-top insert diagram and highlighted in red color. The generated plasmids were followed by transforming into Po1g URA3<sup>-</sup> cells to obtain transformants of Po1g ACL, Po1g DGA1, and Po1g ACL-DGA1, respectively. See abbreviation section for the acronyms of genes and their components.



**Supplementary Figure S2. Main steps in building the constructs for *Snf1* depletion, recyclable *URA3* marker and cellulase expression in *Y. lipolytica*.** The details are described in the Materials and Methods section with construct sequences being listed in this additional file 1. The vector for each plasmid construct is also briefly described in Table 1. In steps 5 and 6, TEFin promoter and XPR2 terminator for Tr *cbh1* gene; GPD promoter and Lip2 terminator for Tr *cbh2*; EXP1 promoter and EXP1 terminator for Tr *eg2*, respectively. Other symbols and abbreviations: p, promoter; Snf1down, 1.2 kb upstream downstream nucleotide sequences of *Snf1* gene; Snf1up, 1.2 kb upstream nucleotide sequences of *Snf1* gene; t, terminator; XPR2, alkaline extracellular protease 2 pre-region (signal sequence).

## Nucleotide sequences SEQ NO 1 to 11

### SEQ ID NO 1

Promoter sequence of SalI-pEXP1-PmlI (1014 bp).

Note: SalI site (GTCGAC) at the 5' end; PmlI site (CACGTG) at the 3' end; start codon (ATG) for the downstream gene prior to the pmlI site. Cloning vector was pUC57.

GTCGACGAGTTGGCGCCGTTTCGAGCCCCACACGTTCGGTGAGTATGAGCGGCCGC  
AGATTGAGCGTTCCGGTTCCCGCGCTGGACGAGAGCCCATGATGGGGCTCCCACCACC  
AGCAATCAGGGCCCTGATTACACACCCACCTGTAATGTCATGCTGTTCATCGTGGTTAATGC  
TGCTGTGCTGTGTGTGTTGGCGCTATTGTTGCGTTATGCAGCGTACACCCACA  
ATATTGGAAGCTTATTAGCCTTCTATTTTCTGCAAGGCTAACAAACATTGCTGTGGA  
GAGGGATGGGGATGGAGGCCGCTGGAGGGAGTCGGAGAGGCCGTTGGAGCGGCTTGGC  
CTGGCGCCCAGCTCGCGAAACGCACCTAGGACCCCTTGGCACGCCGAAATGTGCCACTTTTC  
AGTCTAGTAACGCCTTACCTACGTCAATTCCATGCATGCATGTTGCGCCTTTTCCCTGCC  
TTGATGCCACACAGTACAGTCAGTACAGTGGAGGTTGGGGGGCTTAGATGGGAG  
CTAAAAGCGGCCTAGCGGTACACTAGTGGATTGTATGGAGTGGCATGGAGCCTAGGTGGA  
GCCTGACAGGACGCACGACCGGCTAGCCCGTGACAGACGATGGGTGGCTCCTGTTGTCCACC  
GCGTACAAATGTTGGCCAAAGTCTTGTCAAGCCTGCTGGAACCTAACCTCCAAATTG  
CACTTCGACCCCCATTGATCGAGCCCTAACCCCTGCCATCAGGCAATCCAATTAAGCTCG  
CATTGTCTGCCCTGTTAGTTGGCTCCTGCCGTTCGCGTCCACTGCACAAACACAAAC  
AAGCATTATATATAAGGCTCGTCTCCCTCCCAACCACACTCACTTTTGCCCGTCTCCC  
TTGCTAACACAAAAGTCAAGAACACAAACCACCCCAACCCCTACACACAAGACATA  
TCTACAGCAATGCACGTG,

### SEQ NO 2

### Terminator sequence of KpnI-tEXP1-SalI-ClaI (517 bp)

Note: KpnI site (GGTACC) at 5' end, and SalI and ClaI sites (GTCGACATCGAT) at the 3' end. Cloning vector was pUC57.

GGTACCAAGCTTGGCGAAACTCGATTCTCACCCCTGATAACTCGACTCACCCCCCTTAAC  
TAAAATTCACTTACGACAAACAAACGCTCTGATACCGACTACCCCTCGACTTCTCGCAATCT  
CGACTTCAATCAGAGGACCTCAAACAACCAACTTTTCTTACGATTCTAATTATTTACCCAT  
TCATTAATTCCCGTGCCTCGTCCAGCAATGTCCGAGAGCATTGCTGCTCTGGGCC  
CATCCATTAAATTGGGTCCATCCTCCGGCAGATCCACAGTCCAGTTGTCGCCGACTGGA  
TGGTTAGTAGATCCGCCTTTAGTTGAACATGTTGCGAGTTGATACTGAACATCAGAGT  
TTAGCTTCTGTACTAATATACATTCCTTGTAGTGGAGGTCTGTACACTGTACACTGTATTAT  
ATTGTAGCGTTGTATTTTGTACCGTGTACAGGTGTACAAGTATGTACGAT  
CGTCGACATCGAT,

SEQ NO 3

**Construct 207 for gene ACL (3312 bp)**

Note: It is mouse ACL gene with codon optimization for expression in *Yarrowia*. It had blunt end at 5' end, and stop codon and KpnI at 3' end; the start codon is at the 3' end of above "SalI-pEXP1-PmlI" sequence prior to the pmlI site Cloning vector PUC57

TCCGCTAAGGCTATCTCGAGCAGACTGGAAAGGAGCTGCTCTACAAGTACATTGCACTAC  
CTCCGCCATTCAAACGATTCAAGTACGCCGAGTGACCCCCGACACTGATTGGGCTCAC  
TGCTCCAGGACCATCCCTGGCTGCTCTCCAGTCGCTGGTGGCAAGCCTGATCAGCTCATC  
AAGCGACGAGGCAAGCTGGACTCGTCGGTGTAAACCTGTCGCTCGACGGCGTCAAGTCTTG  
GCTGAAGCCTCGACTCGGACACGAGGCTACTGTGGCAAGGCTAAGGGCTTCCGAAGAAC  
TTTCTCATCGAGCCCTCGTTCTCATTCCCAGGCCAGGAGTTTACGTGTATCTACGCT  
ACCCGAGAGGGCGACTACGTCCTGTTCCACCATGAGGGCGGAGTGGATGTCGGCGACGTTG  
ATGCTAAGGCCAGAACGCTGCTCGTTGGAGTGGACGAGAACGCTGAACACCCGAGGATATCAA  
GCGACACCTGCTCGTTCATGCCCGAGGACAAGAAGGAGGTGCTGGCTTCGTTCATTTCTG  
GCCTCTCAACTTTACGAGGATCTACTTACTTACCTGGAGATCAACCCCTCGTTGTGA  
CCAAGGACGGAGTCTACATTCTGGATCTCGCCGCTAACGGTTGACGCTACCGCCGATTACATC  
TGTAAGGTGAAGTGGGTGACATTGAGTTCCCTCCTTGGCCGAGAGGCTTACCGCTGA  
GGAGGCTTACATCGCTGACCTGGATGCCAAGTCCGGTGTTCGCTGAAGCTCACTGCTCA  
ACCTAAGGCCGAATCTGGACCATGGTCGCTGGTGGAGCTCGGTGTTACTCTGAC  
ACCATTGCGATCTGGTGGCGTGAACGAGACTCGCCAACACTACGGAGAGTACTCCGGTGC  
CTCGGAGCAGCAGACCTACGACTACGCCAAGACTATCCTGTCCTCATGACCCGAGAGAAC  
CACCTGAGGGCAAGATCCTGATCATTGGAGGTTGATTGCTAACCTCACCAACGTGGCGC  
TACTTTAAGGGTATCGTCCGAGCCATTGAGACTACCAGGGCCCCCTGAAGGAGCATGAGG  
TGACCATTTCTCGACGAGGCGGACCTAACGTTTCGGAACCGAGACTCATATGACCGCCA  
TCGTGGCATGGCTCTGGACACCGACCCATTCTAACCGCTCCACCGCCGCTCATACT  
GCCAACTCCTGCTCAACGCTCTGGCTCCACCTCGACTCCTGCTCCTCTCGAACCGCTTCT  
TTTCCGAGTCTCGAGCTGACGAGGTCGCTCTGCCAACGAGGCTAACGCTGCTATGCTCA  
GGATTCTGTGCCCTCCCTCGATCGCTGCAGGGAAAGTCGGCCACTCTCTCTCGACACAC  
CAAGGCTATCGTTGGGCATGCAGACTCGAGCCGTGCAGGGTATGCTGGACTTGATTACG  
TCTGCTCCGAGACGAGCCTCGGCGTGCATGGTTACCCCTCACCGAGATCACAAG  
CAGAAGTTTACTGGGTATAAGGAGATCCTGATTCCGTGTTCAAGAACATGGCTGACGC  
CATGAAGAACCCCTGAGGTCGATGTTCTGATTAACCTTGCTCTCCGATCCGCTTACGA  
CTCCACCATGGAGACTATGAACACTACGCCAGATCCGAACTATTGCCATATTGCTGAGGGCA  
TCCCCGAGGCTCTGACCCGAAAGCTCATTAAGAACGGCCGACAGAACGGGTGTC  
ACTATCATTGGCCCCGCCACCGTTGGTGGCATCAAGCCTGGATGTTCAAGATTGGTAACACC  
GGCGCGTACATCCTCGCTTCAAGCTGTACCGACCCGGCTCGGTTGCCTACGTGTCTCGAT  
CCGGCGGAATGTCTAACGAGCTGAACAAACATCATTCCGAACCAACTGACGGTGTACGAG  
GGCGTCGCCATTGGTGGCGACCGATACCCGGATCTACTTCTATGGACCAACGTGCTGCGATA  
CCAGGATAACCCCTGGCGTCAAGATGATCGTGGCCTCGGAGAGATTGGAGGTACTGAGGAG  
TACAAGATCTGCCGAGGAATTAAAGGAGGGCCACTGACCAAGCCCCTGTTGTTGGTCAT  
CGGAACCTGTGCTACTATGTTCTCTCGAGGTCCAGTTGGTCACGCTGGTGCCTGCGCTAA

CCAGGCTTCCGAGACCGCTGTGGCTAAGAACCAAGGCCCTGAAGGAGGCTGGAGTGTCGTC  
CCTCGATCGTTGACGAGCTGGGAGAGATCATTCACTGTTACGAGGATCTCGTGGCCAA  
GGCGCTATTGTCCCTGCTCAGGAGGTTCCCTCCTACCGTGCCTATGGACTACTCTTGGGC  
TCGAGAGCTGGGACTCATCGAAAGCCGCCTTTCATGACTTCCATTGTGACGAGCGGG  
GTCAGGAGCTGATCTACGCCGGATGCCTATTACCGAGGTCTTAAGGAGGAGATGGGAATC  
GGCGGAGTTCTGGGCTGCTCTGGTCCAGCGACACTCCCCAAGTACTCTTGTCAAGTTATT  
GAGATGTGCCTGATGGTACTGCTGACCACGGCTGCCCTCCGGAGCTCATAACACCAT  
CATTGCGCCCAGCTGGAAAGGACCTCGTGTGCTCTGACCTCCGGTGTGCTCACTATCG  
GCGACCGATTGGTGGCGCTCTGGATGCCGTCCAAGATGTTCTCTAAGGCCTTGACTCC  
GGAATCATTCCCATTGGAGTTCGTCAACAAGATGAAGAAGGAGGGCAAGCTGATCATGGAA  
TTGGTCACCGAGTGAAGTCCATCAACAACCCCTGACATGCGAGTCCAGATTCTGAAGGATTTC  
GTTAAGCAGCATTTCGCCACCCCTGCTCGACTACGCTCTCGAGGTCGAGAAGATCAC  
CACTTCTAAGAACGCCCACCTGATCCTCAACGTGGACGGATTGATTGGTGTGCCCTTGTG  
ATATGCTGCAAACGTGGCTCGTTACCCGAGAGGGAGGCCACGAGTACGTGGATATCGG  
CGCTCTGAACGGAATTTCGTCCTCGGACGATCTATGGCTTATTGGACACTACCTGGACC  
AGAAGCGACTCAAGCAGGGCTCTACCGACACCCCTGGATGATATTCTACGTCCCTGCCT  
GAGCACATGTCTATGTAATAGGGTACC,

#### SEQ NO 4

Full sequence of vector pMT015-YTEFin was described in literature (Tai and Stephanopoulos, 2013).

#### SEQ NO 5

DGA1 gene synthesized (1560 bp), with a **5' blunt end and a NsiI site on the 3' end**

TAACCGCAGACTATCGACTCACAAACTACAAGTCGCGAGACAAAAACGACACGGCACCCA  
AAATCGCGGAATCCGATATGCCCGCTATCGACACCATTACTCAACCGATGTGAGACCTTC  
TCTCTGGTCTGGCACATTTCAGCATTCCACTTCCCTACAATTTCATGCTATGCTGCGCA  
ATTCCACTGCTCTGGCCATTGTGATTGCGTATGTAGTGTACGCTGTTAAAGACGACTCCCG  
TCCAACGGAGGAGTGGTCAAGCGATACTCGCCTATTCAAGAAACTCTTCTCATCTGGAAAGCT  
CTTGGCCGCTACTTCCCCATAACTCTGCACAAGACGGTGGATCTGGAGGCCACGCACACAT  
ACTACCCCTGGACGTCCAGGAGTATCACCTGATTGCTGAGAGATACTGGCCGAGAACAAAG  
TACCTCCGAGCAATCATCTCCACCATCGAGTACTTCTGCCCGCTCATGAAACGGTCTCTT  
TCTATCAACGAGCAGGAGCAGCCTGCCAGCGAGATCCTCTCCTGTCTCCGTTCTCCAG  
CTCTCCGGTTCTAACCTGACAAGTGGATAACCACGACAGCAGATATAGCCGTGGAGAAT  
CATCTGGCTCCAACGCCACGCCCTGGCTCCGAACCTAACGGCAACGGCAACAATGGCAC  
CACTAACCGACGACCTTGTGTCGCTCCACTGCATCTGATTCCACGCTTCT  
TAACGGGTCCCTCAACTCCTACGCCAACAGATCATGGCAAACGACCCACAGCTGTCGC  
CCACAAAACCTCAAGCCCACGGCAGAAAATACATCTTCCGCTACCACCCACGGCATTATC  
GGCATGGGAGCCTTGGTGAATTGCCACCGAGGGAGCTGGATGGTCAAGCTTCCGG  
CATCCCTGTTCTCTTATGACTCTCACCAACAACCTCCGAGTGCCTCTACAGAGAGTACCT  
CATGAGTCTGGAGTCGCTCTGTCCTCAAGAACGACTCTGCAAGGCCCTCTCAAGCGAAACC  
AGTCTATCTGCATTGCGTTGGAGCACAGGAAAGTCTTCTGGCAGACCCGGTGTGTCATG  
GACCTGGTACTCAAGCGAAAGGGTTTGTGACTTGGTATGGAGGTCGGAAATGTCGC

CCTTGTCCCACATGGCCTTGGTGAGAACGACCTCTATGACCAGGTTAGCAACGACAAGT  
CGTCCAAGCTGTACCGATTCCAGCAGTTGTCAAGAACCTTCCTGGATTCAACCTCCTTGA  
TGCATGCCGAGGCCTTCAACTACGATGTCGGTCTGTCCCCTACAGGCGACCCGTCAAC  
ATTGTGGTTGGTCCCCATTGACTTGCCTATCTCCCACACCCCACCGACGAAGAAGTGTCC  
GAATACCACGACCGATAACATGCCGAGCTGCAGCGAATCTACAACGAGCACAAGGATGAAT  
ATTCATCGATTGGACCGAGGGAAAGGAGCCCCAGAGTCCGAATGATTGAGTAATA  
GATGCAT

SEQ NO 6

SNFup sequence: (length: 1238 bp). Cloning vector name: pUC57. It contains:

- (1) AscI (GGCGCGCC)
- (2) The *Yarrowia Snf1* upstream (base 234933 to 236132 of *Yarrowia lipolytica* CLIB122 chromosome D complete sequence; 1200 bp)
- (3) PmlII (CACGTG), KpnI (GGTACC), and MluI (ACGCGT)

GGCGCGCCTGATTACACCCTGGAACCTGTTGGGTGGAGAACAGCCAGATATTGGGGAGATCC  
AGCTGAATCTGAACGGGTTCTGGAGGAGAACACAGCCAAGTCTGCAAGGAGCTGTGGGA  
GCTGCTGGTGGCTGCCAGAAAGACAAGGACGGCATTCCGCCGCAGTTGATTGCCATCAAG  
AAGGAGCAGATGGAGCAGGAGCGGGTGAAGCGGGAAATCAAGATTGAGAGTTGGAGTGGT  
GGGGCGGTGGTGAAGAGACAGAGGTGATAGAAGGGAAAGAACGACAGAAAGAGACGC  
AAAAGGAAGAACAGAAGAGATGGAGATAGGAGGAATGGCGATCGAAGAGACAGAGACA  
GAGATAGAGATAGAGACGACAGGTCAAGTCGTTCAAGTCGGTACAACCGATCAAGATCAAG  
ATCTCCGACTGTCAAAAGGAAACAGATGAATATGGCCGGGACCGGAAGGACTAACTGCAT  
TATATATAAAATAGATTCTGATTAGTGATCAAAGCATGAGATACTGTCAAGGCAAGCAAG  
CCTTACAATATCTGTTGGCTTAGTTAAATCTTTTGAGATGTAATAATTGGGAAACCTCTGT  
TAGATGAACAGTTGCGGCTTACTCATCTTAGACATGGTCAGTTAGGCGTAATTAAATTATC  
AGTTTAATGCGCCTCTACGTAAATAAGCGTTGCCACAAGTTACAGTAGAGATGATACCATTG  
CTACAGGACCGTATAGCGTATATAATGACGAGGCTAATTGATGATGATCGACAGCGAAGAC  
ATATCTACGTCAATCTGACGCTTCAATTGCTCTGTTGGATTGTGTTCTCCAGTGC  
ACCCAGTTTTGCTGACCATTGGTCTTACAGTTGCTCGCGCTTCTCGATTGAAC  
TCTGCTGCGAGGGTGGTGAAGACACACTCCCTGTTAGAACATGCATTGAGGTGTGACATT  
TCCATCTCTGTTAAATCAAACGAGTGACACAAACGGAATAATCTACATATACGTCGAACCC  
TCTTCT  
CACCATCAAAGTTAGGGTGTGCTCCAACACCCACCTCCACACCCACACTCTATTATCACCAC  
TCCACGACTACATACCACTCCCTCCACCCACCTCACAAACACGTGAAGCTCTAGAGGTACCAC  
GCGTA,

SEQ NO 7

Snf1down Sequence (Length: 1217 bp); Vector name: pUC57.

GGGGCCGCGCACTTGTAGAGCACACTAGGGATTAGAGGGATTATGGCACGTACAATATA  
GATAATTAAGCAGTAGCTGAGTCAGTTGAATGATCAGAGGTGAAACATGAGTGTGGATGGA  
TTGTGTAGAGTCGTTAAAATAATGAGTTAAGAATAATTATACGACTACAGGATACGAT  
GTACTTGTATTGTATCGATACAGTACATACAGTACATACGTGAAACATACACTCCTAACTGTTG  
CATCACCTACAACCTCAACTAGTCGGTCATAATTCTTACGTTCCCTGGTAGTCTAG  
GCCAAACACTCCTGGTAACTCTTGAGAAAGATTGCTCTAGACTCCTGAATTCTCTA  
GGCTGGATTCGGCCAAAAAGAACCCACAAACGGCGCAATACTCTTGTAGCCGGAATA  
GTTGTCTCGAATGACAATATCCGGTTGGCAGACCGCTGCTGATGCTCCTGACGATGATAC  
TCTCAAAGTGGTCCACGGTAGACATGACAATAAGAGCCGACGAGGTTTGTCTGTAGTTG  
GCGATGAGCGTTAAGTCCCCAGAGTTCTCCACCACCTTGTGGATCAGCTGATGCTTTGT  
TCCTCGCTGAGTCCTCCGAGGCCAGGGGATAGCCAAGCAGGTGAGAATAGCGCCTGTGTC  
TCGTTAGAAGAGGTGCTGTTGTATCATCCGCATGCTTGTGATGAACACTGAACAAATTCA  
TAGCGTGCTCCACAGCTCATTCCAGTGGTAATTGAGGGAGTGTCCCTGTAACCATGAGTG  
TTGCTTCAGTGACGACCAAGCGAGATCATAATGCTGGAGAGGTTCTCATATTGTAT  
CTGAGACAGCAAAGAAGTGCCTCCAAAAGGCCGGTCCAAAGGTGAGGCCTCTGTTCTG  
GAAGGGGGGTTCCGTTGGCTGTTGGCGTTGATTTAATAGATGCTTGTGTTGAACGACTGG  
GCAATTCCGGCAGAAGAACGCTGCAGAATGATTAGAGCCACCTACTGTACATTCAGCTCGT  
GGGTTTTGTTGGTGGAAAAGAGGTACGAACAGAGAGAGAAAGGCTGCCAGGGCAG  
GCTCGGTGTCATAGTTCTGGTACGATGTCCTTATCGTAATGAGTGTCTTGCAAACAACT  
TGTTCTCTGCACGAATTCAAAGTGGGGCCGGCC,

#### SEQ NO 8

Construct pNREL150, *i.e.*, HisG-*URA3*-HisG (length: 3885 bp). Cloning vector: pUC57. Cloning site: BamHI-HindIII.

BamHI site GGATCC (5' end, position 1 to 6); MluI site ACGCGT + A (7 to 13); HisG of *S. typhimurium* (position 14 to 1169); *URA3* gene cassette (1170 to 2715); HisG of *S. typhimurium* (2716 to 3871); HindIII site AAGCTT (3' end, 3880 to 3885); NotI site GCGGCCGC (3872 to 3879).

GGATCCACCGTAGATCTCCAGTGGTCATGAACGCATGAGAAAGCCCCCGGAAGATCAT  
CTTCCGGGGCTTTTTGGCGCGATACAGACCGGTTAGACAGGATAAAGAGGAACG  
CAGAATTTAGACAACACCCGCTACGCATAGCTATTAGAAATCAGGCCGTTAAGCGATG  
ATTACGAGAATTGCTGGCCGCTCGGCATAAAATTAAATTACACACTCAGCGCTGATT  
GCGATGGCGAAAACATGCCATTGATATCCTGCGCGTGCCTGATGACATTCCGGGTCT  
GGTAATGGATGGCGTGGTCATCTGGTATTATCGCGAAAACGTGCTGGAAGAAGAGCTA  
CTCAACCGCCGCGCACAGGGCGAAGATCCACGCTATTAAACCTGCGCCGCTTGAACCG  
CGGCTGCCGTTATCGCTGGCAACACCGGTTGACGAAGCCTGGACGGCCGCCGCGCTGG  
ACGGTAAACGTATCGCTACCTCATATCCGACCTCCTCAAACGCTACCTCGACCAGAAAGGC  
GTCTCTTAAATCGTGTGTTAAATGGTCTGCGAAGTGCACGCGCAGCAGCTGAGCTAACGG  
GACGCTATCTGCATTGGTCTACCGCGCAGCAGCTGAGCTAACGGCCTGCGTGAAGT  
CGAAGTTATCTACCGCTAAAGCCTGTGATTAGCGACGGTGAGATGGCACAGAGCA

AGCAAGAGCTGATCGATAAATTGCTGACCCGTATTCAGGGCGTATTGAGGCAGCGAATC  
GAAATACATCATGATGCACGCCAACAGTGAACGCCCTGGAAGAGGTTATGCCCTGCTGCCA  
GGCGCCGAAAGGCCGACAATTCTGCCGCTGGCAGGCAGAACAGCGCGTGGCGATGCACA  
TGGTCAGCAGCAGAACGTTGTTCTGGAAACCATGGAGAAACTGAAAGCGCTGGCGCCAG  
CTCGATTCTGGTACTGCCGATCGAGAAGATGATGGAGTATGACGCCCTGATGGCGCTGCG  
CTTATCAGGCCTACGTAATGCGTTGATATTGGGTTCTGTAGGCCGATAAGGCCAACCC  
TGTGATGGAGTAAAGACCATGAGCTCAATACCCTGATTGACTGGAACAGCGGATCTGGTCG  
ACGAGTATCTGACTCGTCATTGCCCTTGAGTACGACTCCAACATGAGTGTGCTT  
GGATCACTTGACGATACATTCTCGTTGGAGGCTGAGGCTGACAGCTGCGTTTCATGATCACATTG  
GGTTGGCCGACAAACAATATCAGCTGCAACGTCATTGCTGGCTTCATGATCACATTG  
TCGGCAAAGGCGACGCCAGAGAGCCATTGACGTTCTTCTAATTGGACCGATAGCCGTAT  
AGTCCAGTCTATCTATAAGTTCAACTACTCGTAACTATTACCATACATACCTCACTGCC  
CCAGATAAGGTTCCGATAAAAAGTCTGAGACTAAATTATTAGTCTCAGTCTCCTTCAACCACC  
AAAATGCCCTCCTACGAAGCTCGAGCTAACGTCACAAGTCCGCTTGCCGCTCGAGTGCT  
CAAGCTCGTGGCAGCCAAGAAAACCAACCTGTGTGCTCTGGATGTTACCAACCAAGG  
AGCTCATTGAGCTGCCGATAAGGTCGGACCTTATGTGTGATGATCAAGACCCATATCGAC  
ATCATTGACGACTTCACCTACGCCGGCACTGTGCTCCCCCTCAAGGAACCTGCTCTTAAAGCA  
CGGTTCTCCTGTTGAGGACAGAAAGTTCGAGATATTGGCAACACTGTCAAGCACCAGT  
ACAAGAACGGTGTCTACCGAATGCCGAGTGGTCCGATATCACCAACGCCACGGTGTACCC  
GGAACCGGAATCATTGCTGGCCTGCGAGCTGGTCCGAGGAAACTGCTCTGAACAGAAGA  
AGGAGGACGTCTGACTACGAGAACTCCCAGTACAAGGAGTTCTGGTCCCTCTCCAAAC  
GAGAAGCTGCCAGAGGCTGCTCATGCTGCCAGCTGTCTGCAAGGGCTCTGGCAC  
TGGCGAGTACTCCAAGCAGACCATGAGCTTGGCCGATCCGACCCCGAGTTGTGGTGGCT  
TCATTGCCAGAACGACCTAACGGCGACTCTGAGGACTGGCTATTGACCCCCGGGGTG  
GGTCTTGACGACAAGGGAGACGCTCTGGACAGCAGTACCGAACTGTTGAGGATGT  
CTACCGGAACGGATATCATAATTGTCGGCCAGGCTGTACGGCCAGAACCGAGATCCTATT  
GAGGAGGCCAAGCGATACCGAGAACGGCTGGCTGGAGGCTTACCGAGAACGATTA  
GGTAGACTATGGATATGTCATTAACTGTGTATATAGAGAGCGTGCAAGTATGGAGCGCTT  
GTTCAAGCTGTATGAGGTCAGACGACCTGCTGATCGAGTATGTATGATACTGCACAACCT  
GTGTATCCGATGATCTGCAATGGGATGTTGTTGTTCTCGATACGGAGATGCTGG  
GTACAAGTAGCTAATACGATTGAACTACTTAACTGTGTATATGAGGCTTGAAAGAAAGCTGACTT  
GTGTATGACTTATTCTCAACTACATCCCCAGTCACAATACCACCACTGCACGGATCTCCAGT  
GGTGCATGAACGCATGAGAAAGCCCCCGGAAGATCATCTCCGGGGCTTTTTGGCGC  
GCGATACAGACCGGTTCAGACAGGATAAAAGAGGAACGCAGAACGAGCT  
ACGCATAGCTATTCAAGAACATCAGGCCGTTAAGCGATGATTGACGAGAACG  
GCGGCATAAAATTAAATTACACACTCAGCGCTGATTGCGATGGCGAAAACATGCCGATT  
GATATCCTGCGCGTGTGATGATGACATTCCGGCTGGTAATGGATGGCGTGGCGATCT  
CGGTATTATCGCGAAAACGTGCTGGAAGAAGAGACTACTCAACCGCCGCGACAGGGCGAA  
GATCCACGCTATTAAACCTCGCCGCTTGAACGACTTCCGGCGCTGCCGTTATCGCTGGCAACA  
CCGGTTGACGAAGCCTGGACGGCCGGCGCTGGACGGTAAACGTATCGCTACCTCAT  
ATCCGCACCTCCTCAAACGCTACCTCGACCAGAAAGGCGTCTTTAAATCGTGTCTGTTAA  
ATGGTTCTGTCGAAGTCGCGCCGCGCGGGCTGGCGACGCTATCTGCGATTGGTCTCT  
ACCGGCGCGACGCTTGAAGCTAACGGCCTGCGTGAAGTCGAAGTTATCTACCGCTCTAAAGC  
CTGTCTGATTAGCGCGACGGTGAGATGGCACAGAGCAAGAGCAAGAGCTGATCGATAATTG

CTGACCCGTATTCAAGGCGTGATTCAAGGCAGCGAATCGAAATACATCATGATGCACGCGCC  
AAAGTGAACGCCTGGAAGAGGTTATCGCCCTGCTGCCAGGCGCCAAAGGCCACAATTCTG  
CCGCTGGCAGGCGAGCAACAGCGCTGGCGATGCACATGGTCAGCAGCGAAACGTTGTTCT  
GGGAAACCAGGAGAAACTGAAAGCGCTGGCGCCAGCTGATTCTGGTACTGCCGATCGA  
GAAGATGATGGAGTGTACTGACGCCTGATGGCGCTGCCTTATCAGGCCTACGTAATGCGTT  
GATATTGGGTTCTGTAGGCCGATAAGGCCGAAACCGTGTGATGGAGTAAAGACCATGAG  
CTTCAATACCGTATTGACTGGAACAGCGGATCTGGCGGCCAAGCTT,

SEQ NO 9

Construct 161 (length: 6313 bp) for Snf1up-HisG-*URA3*-HisG-Snf1down. Cloning vector: pUC57-Simple.

AscI site (GG^CGCG\_CC) at the 5' end; FseI site (GG\_CCGG^CC) at the 3' end.

GGCGCGCCTGATTACACCTGGAACCTGTTGGGAGACAAGCCAGATATTGGGAGATCC  
AGCTGAATCTGAACGGGTTCTGGAGGAGAACACAGCCAAGTCTGCAAGGAGCTGGGA  
GCTGCTGGTGGCTGCCAGAAAGACAAGGACGGCATTCCGCCGAGTTGATTGCCATCAAG  
AAGGAGCAGATGGAGCAGGAGCGGGTGAAGCGGAAATCAAGATTGAGAGTTGGAGTGGT  
GGGGCGGTGGTGAAGAGACAGAGGTGATAGAAGGGAAAGAACGACAGAAGAGACGC  
AAAAGGAAGAACAGAAGAGATGGAGATAGGAGGAATGGCGATCGAAGAGACAGAGACA  
GAGATAGAGATAGAGACAGCTCAAGTCGTTCAAGTCGGTACAACCGATCAAGATCAAG  
ATCTCCGACTGTCAAAAGGAAACAGATGAATATGCCGGGACCGGAAGGACTAATGCAT  
TATATATAATATAGATTCTGATTAGTGTCAAAGCATGAGATACTGTCAAGGCAAGCAAG  
CCTTACAATATCTGGCTTAGTTAAATCTTTGAGATGTAATAATTGGGAAACCTCTGT  
TAGATGAACAGTTGCGGCTACTCATCTTAGACATGGTCCAGTTAGCGTAATTATTATC  
AGTTTAATGCGCCTCTACGTAAAGCGTTGCCACAAGTTACAGTAGAGATGATACCGATTG  
CTACAGGACCGTATAGCGTATATAATGACGAGGCTAATTGATGATGATCGACAGCGAAGAC  
ATATCTACGTCAATCTGACGCTCATTGTGCTTCTGGATTGTGTTCTCCAGTGC  
ACCCAGTTTGCTGACCAATTGGCTTACCAAGTTTGCTCGCGCTTCTCGATTGAAC  
TCTGCTGCGAGGGTGGTGAAGACACACTCCCTGTTAGAACATGCATTGAGGTGTGACATT  
TCCATCTCTGTTAAATCAAACGAGTGACACAAACGGAATAATCTACATATACGTCGAACCC  
TCTTCTCTCTCTCTCTCTTAATAGACACCCTCGTATCGACTCTGCTCCTCTTATCA  
CACCATCAAAGTTAGGGTGTCCAACACCACCTCCACACCACACTCTATTATCACCACC  
TCCACGACTACATACCACTCCTCCACCACCTCACAAACACGTGAAGCTCTAGAGGTACAC  
GCGTAGATCTCCAGTGGCATGAACGATGAGAAAGCCCCCGGAAGATCATCTGGGG  
GGCTTTTTGGCGCGCATACAGACCGGTTAGACAGGATAAAGAGGAACCGAGAATGT  
TAGACAACACCCGCTACGCATAGCTATTAGAAATCAGGCCGTTAAGCGATGATTACGA  
GAATTGCTGGCCGCTGCGGATAAAAATTAAATTACACACTCAGCGCCTGATTGCGATGGC  
GGAAAACATGCCGATTGATATCCTGCGCGTGTGATGATGACATTCCGGTCTGGTAATGG  
ATGGCGTGGCGATCTCGGTATTATCGGCAGAACCGTGTGGAAGAAGAGCTACTCAACCG  
CCGCGCACAGGGCGAAGATCCACGCTATTAAACCTGCGCCGTTGACTCAGGCCGCTGCC  
GTTTATCGCTGGCAACACCGGTTGACGAAGCCTGGACGGGCCGGCGCTGGACGGTAA  
ACGTATCGCTACCTCATATCCGACCTCCTCAAACGCTACCTGACCAGAAAGCGTCTCTT

TAAATCGTGTCTGTTAAATGGTTCTGTCGAAGTCGC CGCC GCG C GCG GGG CTGCC GAC GCTA  
TCTCGGATTGGTCTCTACCGCGCGACGCTGAAGCTAACGGCCTGCGTGAAGTCGAAGTT  
ATCTACCGCTCTAAAGCCTGCTGATT CAGCGCAGCGT GAGATGGCACAGAGCAAGCAAG  
AGCTGATCGATAAATTGCTGACCCGATT CAGGGCGTGATT CAGGC GCG CAATCGAAATAC  
ATCATGATGCACCGCCAAGTGAACGCCCTGGAAGAGGTTATGCCCTGCTGCCAGGC GCG  
AAAGGCCACAATTCTGCCGCTGGCAGGCGAGCAACAGCGCGTGGCGATGCACATGGTCAG  
CAGCGAAACGTTCTGGGAAACCATGGAGAAACTGAAAGCGCTGGCGCCAGCTCGATT  
CTGGTACTGCCGATCGAGAAGATGATGGAGTGATCTGACGCC TGATGGCGCTGCGCTTATCA  
GCCCTACGTAATGCGTTGATATTGGGTTCTGTAGGCCGATAAGGCCGAACCCCTGTGATG  
GAGTAAAGACCATGAGCTCAATACCCGATTGACTGGAACAGCGGATCTGGTCAGCAG  
ATCTGTCTGACTCGT CATTGCCGCC TTGGAGTGACTCCAACTATGAGTGTGCTTGGATCA  
CTTGACGATACATTCTCGTTGGAGGCTGTTGACAGCTGCCGTTTCATCATGATCACATTGTCGGCA  
CCGACAACAATATCAGCTGCAACGTCATTGCTGGCTTCAATTTGGACCGATAGCCGTATAGTCCA  
GTCTATCTATAAGTCACACTCGTAACTATTACCATACATACACTGCCCCAGAT  
AAGGTTCCGATAAAAAGTCTGCAGACTAAATTATTCAGTCTCCTCTCACCAACCAAAAT  
GCCCTCCTACGAAAGCTGAGCTAACGTCACAAGTCCGCC TTGCCGCTCGAGTGCTCAAGC  
TCGTGGCAGCCAAGAAAACCAACCTGTGTGCTTCTCTGGATGTTACCAACCAAGGAGCTC  
ATTGAGCTTGGCGATAAGGTCGGACCTTATGTGTCATGATCAAGACCCATATGACATCAT  
TGACGACTTCACCTACGCCGGCACTGTGCTCCCCCTCAAGGAACTTGCTCTTAAGCACGGTT  
CTTCCTGTTGAGGACAGAAAGTTCGAGATATTGGCACACTGTCAAGCACCAGTACAAGA  
ACGGTGTCTACCGAATGCCGAGTGGTCCGATATCACCAACGCCACGGTGTACCCGGAACC  
GGAATCATTGCTGGCCTGCGAGCTGGTGCCGAGGAAACTGTCTCTGAACAGAAGAAGGAGG  
ACGTCTGACTACGAGAACTCCCAGTACAAGGAGTTCTGGTCCCCTCCCAACGAGAAG  
CTGGCCAGGGTCTGCTCATGCTGCCGAGCTGTCTTGCAAGGGCTCTGGCCACTGGCGA  
GTACTCCAAGCAGACCATTGAGCTGCCGATCCGACCCCGAGTTGTGGTGGCTTCTGATT  
CCCAGAACCGACCTAAGGGCGACTCTGAGGACTGGCTTATTCTGACCCCCGGGTGGCTT  
GACGACAAGGGAGACGCTCTGGACAGCAGTACCGAACTGTTGAGGATGTCATGTCTACCG  
GAACGGATATCATATTGTCGGCCGAGGTCTGTACGGCCAGAACCGAGATCCTATTGAGGA  
GGCCAAGCGATACCAGAAGGCTGGCTGGAGGCTTACCAAGAAGATTAACGTGTTAGAGGTTA  
GAECTATGGATATGTCATTAACTGTGTATATAGAGAGCGTGAAGTATGGAGCGCTTGTCA  
GCTTGTATGATGGTCAGACGACCTGTCTGAGTATGTATGATACTGCACAAACCTGTGTA  
TCCGCATGATCTGCCAATGGGCATGTTGTTGTTCTCGATACGGAGATGCTGGGTACA  
AGTAGCTAATACGATTGAACTACTTAACTTATGAGGCTTGAAGAAAGCTGACTTGTGTA  
TGACTTATTCTCAACTACATCCCCAGTCACAATACCACCACTGCACGGATCTCCAGTGGTC  
ATGAACGCATGAGAAAGCCCCGGAAGATCATCTCCGGGGCTTTTTGGCGCGCGAT  
ACAGACGGTTCAGACAGGATAAAGAGGAACGCAGAATGTTAGACAACACCCGCTACGCA  
TAGCTATTCAAGAACATCAGGCCGTTAACGCGATGATTCACGAGAATTGCTGGCCGCTGCC  
ATAAAAAATTAAATTACACACTCAGCGCCTGATTGCGATGGCGAAAACATGCCGATTGATAT  
CCTCGCGTGTGATGATGACATTCCGGCTGGTAATGGATGGCGTGGTCATCTCGTA  
TTATCGCGAAAACGTGCTGGAAGAAGAGCTACTCAACCGCCGCGCACAGGGCGAAGATCC  
ACGCTATTAAACCTCGGCCGCTTGACTTCGGCGGCTGCCGTTATCGCTGGCAACACCGGT  
TGACGAAGCCTGGACGGCCGGCGCTGGACGGTAAACGTATCGCTACCTCATATCCGC  
ACCTCCTCAAACGCTACCTCGACCAGAAAGCGTCTTTAAATCGTGTGTTAAATGGTT

CTGTCGAAGTCGCCCGCGCGGGCTGCCGACGCTATCGGATTGGTCTTACCGGC  
GCGACGCTGAAGCTAACGGCCTCGTGAGTCGAAGTTATCTACCGCTCTAAAGCCTGTCT  
GATTCAAGCGACGGTGAGATGGCACAGAGCAAGCAAGAGCTGATCGATAAATTGCTGACC  
CGTATTCAAGGGCGTATTCAAGCGCGCGAACATCGAAATACATCATGATGCACCGCCAAGTG  
AACGCCTGGAAGAGGTTATGCCCTGCTGCCAGGCCGAAAGGCCACAATTCTGCCGCT  
GGCAGGGCGAGCAACAGCGCTGGCGATGCACATGGTCAGCAGCGAACAGTTGTTCTGGGAA  
ACCATGGAGAAACTGAAAGCGCTGGCGCAGCTGATTCTGGTACTGCCGATCGAGAAGA  
TGATGGAGTGATCTGACGCCATGGCGCTCGCTTATCAGGCCTACGTAATCGTTGATAT  
TTTGGGTTCTGTAGGCCGATAAGCGGAACCCCTGTGATGGAGTAAAGACCATGAGCTCAA  
TACCCCTGATTGACTGGAACAGCGGATCTGGCGCCGCACTTGTAGAGCACACTAGGGATT  
TAGAGGGGATTATGGCACGTACAATATAGATAATTAAGCAGTAGCTGAGTCAGTTGAATGA  
TCAGAGGTGTAACATGAGTGTGGATTGTAGAGTCGTTAAAATAATGAGTTAA  
GAATAATTATACGACTACAGGATACGATGTACTTGTATTGTATCGATACTACAGTA  
CATACGTGTAACATACTCCTAAACTGTTGCATCACCTACAACCTCCAACTAGTCGGTCATAATT  
CATTAATACGTTCCCTGGTAGTCTAGGCCAAACACTCCTGGTGAACCTCTGAGAAAAG  
ATTCGCTCTCTAGACTCCTGAATTCTCTAGGCTGGATTTCGGCCAAAAAGAACCCACAAA  
CGCGCAAAACTCTTGTAGCCGAATAGTTGTCTCGAATGACAATATCCGGGTTGGCAG  
ACCGCTGCTGATGCTCCTGACGATGATACTCTCAAAGTGGTCCACGGTAGACATGACAATA  
AGAGCCGACGAGGTTTGCTCTGTAGTTGGCGATGAGCGTTAAGTGCCCCAGAGTCTC  
CACCACCTGTGGATCAGCTGATGCTTGTCTCGCTGAGTCCTCGAGGCCAGGGGAT  
AGCCAAGCAGGTGAGAATAGCGCCTGTCTCGTTAGAAGAGGTGCTGTTGATCATCCG  
CATGCTTGTGATGAACTGAACTAAATTCACTAGCTGCTCCACAGCTCATTCCAGTGGTAA  
TTTGAGGGAGTGTCCCTGTAACCATGAGTGTGCTTCAGTGACGACCAGAGCGAGATCATA  
AATGTCTGGAGAGGTTCTCATATTGTATCTGAGACAGCAAAGAAGTGCCTCAAAGGC  
CGGTTCCAAAGGTGAGGCTCTGTTCTGGAAGGGGGTTCCGTTGCTGGCGTTGATT  
TAATAGATGCTTGTGTTGACGACTGGTCAATTGGCAGAAGAAGCTGCAGAATGATT  
AGAGCCACCTTACTGTACATTCACTCGTGGTTTTGGTTGGAAAAGAGGTACGA  
ACAGAGAGAGAGAAAAGGCTGCCAGGGCAGGCTCGTGTAGTTCTGGTACGATGCCT  
TTATCGGTAATGAGTGTCTTGCAAACAACTTGTCTGCACGAATTCCAAAGTGGGGC  
CGGCC,

SEQ NO 10

Construct 162. *cbh1-cbh2-eg2* (Length: 7780 bp). Cloning vector: pUC57.

Sequence:

Sall site (GTCGAC) and PmlI site (CACGTG) at the 5' end. KpnI site (GGTACC) at the 3' end.

GTCGACCACGTGAGAGACCGGGTTGGCGCGCATTGTGTCCTAAACAGCCCCAATTG  
CCCCAATTGACCCAAATTGACCCAGTAGCGGGCCAACCCCGCGAGAGCCCCCTCTCCC  
CACATATCAAACCTCCCCGGTCCCACACTTGCCTTAAGGGCGTAGGGTACTGCAGTCTG

GAATCTACGCTTGTTCAGACTTGTACTAGTTCTTGCTGCCATCCGGTAACCCATGCC  
GGACGCAAAATAGACTACTGAAAATTGTTGCTTGTGGTTGGACTTAGCCAAGGGTAT  
AAAAGACCACCGTCCCCGAATTACCTTCTCTCTCTCTCCTGTCAACTCACACC  
CGAAATCGTTAAGCATTCTCTGAGTATAAGAACATCATTCAAATGGTGAGTTCAGAGGC  
AGCAGCAATTGCCACGGGCTTGAGCACACGGCCGGGTGGTCCCATCCATCGACACAA  
GACGCCACGTATCCGACCAGCAGTACTAACCAGAAGCTCGTACCGCCCT  
TACTATTCTCACGCCCTGGCCCAGCAGGCAGGGACCCTACCGCCGAGAACCATCCTC  
CCCTCACCTGGCAGGAGTGCACCGCACCGGATCCTGACACAGCAGAACGGCTGTCGTG  
CTGGACGCAAATTGGCGATGGGTTACGTGAAACGGTACACTAAGTGTACAG  
ACACATGGGATCCTACCTACTGCCCGACGACGAAACATGTGCCAGAACTGTGCCCTGAC  
GGAGCTGACTACGAGGGAACCTACGGGTCACCTCGTCCGGAGCAGTCTCAAGCTCAACT  
TTGTGACCGGTTCAAACGTCGGTACGGCTATCTGCTCCAGGACGGACTCGACCTACCAG  
ATCTCAAACGTGAACCGAGAGTTCTCGTTGACGTTGATGTTCGAACTTGCCTGCGGA  
CTTAATGGTGCCTGTACTTGTCAATGGATGCTGACGGCGGAGTCTCTAAGTACCCAA  
CAACAAGGCTGGTGCACGGTACGGCTATTGTGACAGTCAGTGCCTAGAGATCTA  
AAATTCAATTGATGGCGAGGCCAACGTCGAGGGCTGGCAACCGAGCAGCAATAACGCCAATA  
CTGGAATCGGCACCGCCTGTCAGCGAGATGGACGTGTGGGAAGCGAACCTCCAT  
TAGTAATGCTGTAACACCCCCATCCGTGCGACACTCCGGACAGACGATGTGTTCCGGTACG  
ATTGTGGCGGCACCTACTCCAACGATCGATACGCGGGTACATGCGATCCGACGGCTGCGAT  
TTCAACCCATACGGATGGTAATACATCCTCTATGGACCAGGTAAGATTATCGACACTAC  
GAAGCCTTCACCGTGGTACGCAGTTTGACAGATGACGGACTGATAACCGAACCTT  
CGGAAATCAAGCGCTTACATCCAGAACTCTAATGTTATTCCCAACCCACAGTGACATT  
TCCGGAGTTACTGGAAACTCTATCACTACAGAGTTGTACAGCCAAAAGCAGGCCCTCGG  
GGACACCGACGACTTCTCAGCACGGAGGCCTGGCAAAATGGAGCTGCCATGCAACAG  
GGAATGGTGCCTGTCATGCTCTGTGGATGACTACGCCGCTCAAATGCTGTGGCTGGACTC  
TGACTACCCACTGATGCCGATCCAACACTACCCCTGGCATGCCAGAGGTACTTGTCCCACCG  
ACAGCGCGTCCGACGTAGAGTCCCAGTCGCCAACAGCTATGTGACGTACTCGAAC  
ATTAAGTTCGGCCCTATCAACTCTACGTTACCGCATCTAACCCACCTGGTGGTAACCGAGG  
TACTACGACCACTCGACGTCCTGCTACTACCACTGGTGTACCGACCCACCCAGTCTC  
ACTACGGACAGTGTGGCGGAATTGGATACTCTGGTCCCACCGTGTGCGTCTGGAACTACC  
TGTCAGGTCCCTCAACCCCTACTACTCCCAATGCCGTAAATAGGCAATTAAACAGATAGTTGCC  
GGTGATAATTCTCTAACCTCCCACACTCCTTGACATAACGATTATGTAACGAAACTGAA  
ATTGACCAAGATATTGTTGAAATAGAAAATCTGGCTGTAGGTGGCAAAATGCGGCGTCTT  
TGTTCATCAATTCCCTCTGTGACTACTCGTCATCCCTTATGTTGACTGTCGTATTCTTATT  
TTCCATACATATGCAAGTGAGATGCCGTGCGAATTGACCGAGTAGGATGCGTCTGCCACG  
GGCTTTGTGGGGTGTGGAGAAAGGGGTGTGGAGATGGAAGCCGGTAGAACCGGCT  
GCTGTGCTGGAGATGGAAGCCGGTAGAACCGGCTGCTGGGGGATTGGGCCGCTG  
GGCTCCAAAGAGGGTAGGCATTGTTGGGTACGTAATTGCGGATTGGGGCTGCG  
GCATGTCCCATTGGTCAGAATTAGTCCGGATAGGAGACTTACGCCAATCACAGCGCCGGA  
TCCACCTGTAGGTTGGGTGGGAGCACCCCTCACAGAGTAGAGTCAAACAGCAGCA  
GCAACATGATAGTGGGGTGTGCGTGTAAAGGAAAAAGAAGCTGGTTATATTCC  
CGCTCTATTAGAGGTTGCCGGATAGACGCCGACGGAGGGCAATGGCGCCATGGAACCTG  
CGGATATCGATACGCCGCCGGACTGCGTCCGAACCAGCTCCAGCAGCGTTTCCGGC  
CATTGAGCCGACTGCGACCCGCCAACGTGTCTGGCCCACGCACATGTCATGTTGGTGT

TGGGAGGCCACTTTAAGTAGCACAAGGCACCTAGCTCGCAGCAAGGTGTCCGAACCAAA  
GAAGCGGCTGCAGTGGTGCACCGGGCGAACCGGGAAAAGCCACGGGGCACGA  
ATTGAGGCACGCCCTCGAATTGAGACGAGTCACGGCCCATTGCCGCAATGGCTCGC  
CAACGCCGGTCTTGACCACATCAGGTTACCCAAAGCCAACCTTGTGTTAAAAGCT  
TAACATATTATACCGAACGTAGGTTGGGCGGGCTGCTCCGCTGTCCAAGGCAACATT  
TATAAGGGTCTGCATGCCGGCTCAATTGAATCTTTCTCTCTCTATATT  
TTGAATTAAACACACATCAACATGAAGCTCGTACCGCCTTACTATTCTCACGGCGTTCTG  
GCCAGGGCGTGCAGTAGTGGTGGGCCAGTCGGTGGTCAAAACTGGTCTGGCCGACTG  
CTGTGCTCTGGTGCACCTGTGTTACTCGAACGATTACTACTCTCAATGCCTGCCGGAGC  
TGCCAGCTCCAGTCATCCACGAGAGCCCGTCTACAACATCACGAGTGTCCCCACGACCT  
CTCGCTCATCGAGCGCAACTCCTCCCCCTGGTCTACCACAACCCGGTCCCACCTGTGG  
TCGGGAACGCCACGTACTCCGGCAATCCTTGTGGCGTGACACCTGGCAAACGCC  
CTACGCTAGTGAAGTGTCGTCGCTGCCATTCCCTCTCACTGGAGCTATGGAACGGCTG  
CCGCTGCCCGTCAAAGGCTCCCTCTCATGTGGCTGGACACTCTTGATAAGACCCC  
ATGGAGCAAACCTGGCAGACATTGAACACTCGAACAAAAACGGTGGTAAC  
AGTCGGTCTATGACCTCCGACAGAGACTCGCCTGCCCTGGCCTTAACGGCGAATAT  
TCTATTGCGGATGGAGGGTCGCAAGTACAAAAACTATATTGACACGATCCGGCAGATCGT  
CGTTGAGTACTCTGACATCCGTACCCCTCTCGTCATCGAGCCGATTCCCTGCCAACTGGT  
TACCAACCTGGTACACCTAACGTGCAAACGCCAGTCTGCTTACCTGGAGTGTATT  
ATGCTGTTACCAACTGAACCTCCCCAATGCGCATGTACCTCGATGCTGGACACGCC  
TGGCTGGGTGGCCTGCTAATCAGGACCCGCTGCTCAGCTTCGCTAATGTATA  
CGCTTCCTCCCCAGGGCCCTCGAGGACTGCCACTAACGTTGCCAACTACAACGGTGG  
ACATTACCTCTCCCCCATCCTACACCCAGGGTAACCGAGTGTATAACGAGAAGCT  
CATGCTATCGGACCCCTGCTGCCAATCATGGATGGAGCAACGCC  
GGGCCATCGGGCAAGCAGCCCACCGGCCAGCAGCAGTGGGCGATTGGTGAATGTGATC  
GGAACCGGTTCGGTATCGCATCTGCCAATACAGGCAGTCCCTGCTGACTATTG  
TGGGTCAAGCCTGGCGAGAGTGTGACGGCACCTCGGACAGCTCCGCTCCG  
GCCACTGCGCACTGCCGGACGCTTGCAGCCGCTCTCAGGCCGGAGCGTGGTCAAAG  
TACTTGTGAGCTATTGACTAACGCCAACCCCTGTTCTATGATAAGCTATT  
TTACAACCTTACCTCAACTATCTACTTAAATAATGAATATCGTTATTCT  
TATATGCGTTCTCTAACGACAATCGAAACCCAGCATGCGATCGAATGG  
TTCCGAAGTTGATCAATGCTGATAGTCAGGCAGCTGAGAAGATTGACACAG  
CGTAGGGAACCGATCAACCTGTCTACCAGCGTTACGAATGCC  
TTGAATCCTGCAATGGTGCCTGGATACTGATGTCACAAACTTAAGAAG  
CTCTTCCTCGAAACTGAGTTGGCGCCGTTTCGCGCTGGACAGAG  
GGCCAGATTGAGCGTTCCGGTTCCGCGCTGGAGCGAGAG  
CCACCAACAGCAATCAGGCCCTGATTACACACCCACCTG  
GTTAATGCTGCTGTGCTGTGTTGGCGCTCATGTTGCGTTATG  
ACACCACAATATTGGAAGCTTATTAGCCTTCTATT  
GCTGTGGAGAGGGATGGGATATGGAGGCC  
GGCTTGGCCTGGCGCCAGCTCGCAAACGCC  
CACTTTCACTAGTAACGCC  
CCTGCC  
GATGCC  
GAGCT  
GATGGGAGCT  
AAAGCGGC  
TAGCGGT  
ACACTAG  
TGGGATT  
GTGAGT  
GGCATGG  
GAGCC

TAGGTGGAGCCTGACAGGACGCACGACC GGCTAGCCCGTAGCCCGTACAGACGATGGGTGGCTCCTG  
 TTGTCCACCGCGTACAAATGTTGGGCCAAGTCTTGTCAAGCCTTGCTGCGAACCTAATTCC  
 CAATTTCGACTTCGACCCCCATTGATCGAGCCCTAACCCCTGCCCATCAGGCAATCCAAT  
 TAAGCTCGATTGTCTGCCTTGTAGTTGGCTCCTGCCGTTCGGCGTCCACTGCACAA  
 ACACAAACAAGCATTATAAGGCTCGTCTCCCTCCAACCACACTCACTTTTGCCC  
 GTCTTCCCTGCTAACACAAAAGTCAAGAACACAAACCACCCCAACCCCTACACACA  
 AGACATATCTACAGCAATGAAGCTCGTACCGCCTTACTATTCTCACGGCCGTTCTGGCC  
 AACAGACGGTGTGGGCCAATGTGGCGCATTGGCTGGTGGTCCGACCAACTGTGCTCCT  
 GGCTCGGCTGTTGACTCTGAACCCCTACTACGCTCAGTGTATCCCGGAGCGACCCAT  
 CACCACGTCTACTCGACCCCCCTCAGGACCCACTACCACCACGCGAGCGACCTCCACCTCCA  
 GCTCTACTCCCCTACATCCTCTGGTGTTCGATTGCTGGCGTGAACATGCCGGTTCGACT  
 TCGGATGCACGACTGACGGTACATGCGTACATCCAAGGTCTATCCCCGTTGAAAAACTTC  
 ACCGGGTCTAACAAATTATCCTGACGGCATTGGTCAGATGCAACACTTGTCAACGAGGATGG  
 CATGACCATCTCCGACTCCCTGTCGGCTGGCAGTATCTAGTGAACAATAACCTGGCGGAA  
 ACCTTGACAGCACAAGCATTCCAAGTACGATCAATTGGTGCAGGGATGCCTCAGTTGGGT  
 GCTTAUTGCATCGTTGACATTACAACACTACGCTAGGTGGAACGGGGAAATCATCGGACAGG  
 GTGGCCCCACTAACGCCAGTTACAAGTCTGTGGTCCCAGCTCGCCTTAAGTACGCATCT  
 CAGTCCGGGTGTGGTTGGAATCATGAATGAGCCCCATGACGTCAACATTAAACACTTGGG  
 TGCCACCGTACAGGAAGTCGTTACCGCATTGTAACGCCAGTGGCAGTGGCAGTTCATCT  
 CGCTGCCTGGAAATGACTGGCAGTCTGCTGGCTTCAATTCCGATGGATCAGCAGCCGCT  
 CTCAGCCAAGTCACGAACCCCTGATGGATCTACAACCAACCTGATTTGACGTGCATAAGTA  
 CCTGGACTCGGATAACAGCGGAACCCACGCCAGTGGCAGTGTACCAACTAACACATTGACGGTGC  
 TTTTCTCCTCTCGCAACGTGGCTCAGACAGAACACAGACAGGCCATTCTGACTGAAACTGG  
 TGGCGGCAACGTTCAATCATGTATCCAGGACATGTGCCAGCAGATTCACTGAATCAGA  
 ATAGTGACGTGTACCTGGCTACGTGGTTGGGAGCAGGGCCTTGATTCTACCTACGTC  
 CTGACCGAGACTCCGACATCCTCCGGCAACTCGTGGACTGATAACATCTCTGTTCGTCATGC  
 CTTGCCCGCAAGTGATAAAAGCTTGGCGAAACTCGATTCTCACCCCTCGATAACTCGACTC  
 ACCCCCCCTTAACTTAAACTTACGACAAACAAACGTCTGATACCGACTACCCCTCGA  
 CTTCTCGCAATCTGACTTCAATCAGAGGACCTAAACAAACCACTTTTCTACGATTCTA  
 ATTATTTACCCATTCAATTCCCGTGCCTCGTCCAGCAATGTCCGAGAGCATTCTGCCT  
 GTGTCTCTGGGCCATCCATTAAATTGGTCCATCCTCCGGCAGATTCCACAGTCCAGTT  
 GTCGCCGACTGGATGGTTAGTAGATCCGCCTTTAGTTGAACATGTTGCGAGTTGATACC  
 TGAACATCAGAGTTTAGCTTGTACTAATATACATTCCCTTGTACACT  
 GTACACTGTATTATATTGTAGCGTTGTATTTGATCCCTGGTTGTTCGTTACGGTGTAC  
 AAGTATGTACGATCGGTACC,

#### SEQ NO 11

Construct 163 (length: 14064 bp) for Snflup-*cbh1-cbh2-eg2-HisG-URA3-HisG-Snfl*down; Cloning vector: pUC57-Simple.

AscI site (GGCGCGCC) at the 5' end; FseI site (GGCCGGCC) at the 3' end.

GGCGCGCCTGATTACACCCTGGAACTGTTGGGTGGAGACAAGCCAGATATTGGGAGATCC  
AGCTGAATCTGAACGGGTTCTGGAGGAGAACACAGCCAAGTCTGCAAGGAGCTGTGGGA  
GCTGCTGGTGGCTGCCAGAAAGACAAGGACGGCATTCCGCCGCAGTGATTGCCATCAAG  
AAGGAGCAGATGGAGCAGGAGCAGGTGAAGCGGGAAATCAAGATTGAGAGTTGGAGTGGT  
GGGGCGGTGGTGGAAAGAGACAGAGGTGATAGAAGGGAAAGAACGACAGAAGAGACGC  
AAAAGGAAGAAACAGAAGAGATGGAGATAGGAGGAATGGCGATCGAAGAGACAGAGACA  
GAGATAGAGATAGAGACACAGGTCAAGTCGTCAAGTCGGTACAACCGATCAAGATCAAG  
ATCTCCGACTGTCAAAAGGAAACAGATGAATATGCCGGGACCGGAAGGACTAACTGCAT  
TATATATAAATATAGATTCTGATTAGTGATCAAAGCATGAGATACTGTCAAGGCAAGCAAG  
CCTTACAATATCTGTTGGCTTAGTTAAATCTTTTGAGATGTAATAATTGGGAAACCTCTGT  
TAGATGAACAGTTGCGGCTTACTCATCTTAGACATGGTCCAGTTAGGCGTAATTAAATTATC  
AGTTTAATGCGCCTCTACGTAAAGCGTTGCCACAAGTTACAGTAGAGATGATACATTG  
CTACAGGACCGTATAGCGTATATAATGACGAGGCTAATTGATGATGATCGACAGCGAAGAC  
ATATCTACGTCAATCTGACGCTTCAATTGCTTCTGCTTCTGCGATTGTGTTCTCCCAGTGC  
ACCCAGTTTTGCTGACCATTGGTCTTACAGTTGCTCGCGCTTCTCGATTGAAC  
TCTGCTGCGAGGGTGGTGAAGACACACTCCGTAGAACATGCATTGAGGGTGTGACATT  
TCCATCTCTGTTAAATCAAACGAGTGACACAAACGGAATAATCTACATATACGTCGAACCC  
TCTTCTCTCTCTCTCTCTCTCTCTTAATAGACACCCCTCGCTATCGACTCTGCTCCTCCTTATCA  
CACCATCAAAGTTAGGGTGTGCCCCAACACCACCTCCACACCACACTCTATTATCACCACC  
TCCACGACTACATACCACTCCTCCACCACCTCACAAACACGTGAGAGACCGGGTTGGCGCG  
CATTGTGCTCCAAAAAACAGCCCCATTGCCCAATTGCCCAATTGACCCAAATTGACCCAGTAGCGG  
GCCCAACCCGGCGAGAGCCCCCTCTCCCCACATATCAAACCTCCCCGGTCCCACACTT  
GCCGTTAAGGGCGTAGGGTACTGCACTGCTGGAAATCTACGCTGTTAGACTTTGACTAGTTT  
CTTGTCTGCCATCGGGTAACCATGCCGACGCAAAATAGACTACTGAAAATTGGT  
CTTGTGGTGGGACTTAGCCAAGGGTATAAAAGACCACCGTCCCCAATTACCTTCT  
TCTTTCT  
AATCATTCAAATGGTGAGTTCAGAGGACGAGCAATTGCCACGGCTTGAGCACACGG  
CCGGGTGGTCCCATTCCATCGACACAAGACGCCACGTACCGACAGCACCTTTGCA  
GTACTAACCGCAGAAGCTCGTACCGCTTACTATTCTACGGCCGTTCTGGCCAGCAG  
CAGGGACCGCTACCGCCGAGAACCATCCTCCCCCACCTGGCAGGAGTGCACCGCACCGG  
ATCCTGCACCACCGAGAACGGCGCTGCTGCTGGACGCAAATTGGCGATGGGTCACGATG  
TGAACGGGTACACTAACTGCTATACAGGCAACACATGGGATCCTACCTACTGCCCGACGAC  
GAAACATGTGCCAGAACACTGTGCCCTGACGGAGCTGACTACGAGGGAACCTACGGGTCA  
CCTCGTCCGGAGCAGTCTCAAGCTCAACTTGTGACCGGTTCAAACGTCGGTCACTGGCT  
TATCTGCTCCAGGACGACTCGACCTACAGATCTTCAAACACTGTTGAACCGAGAGTTCTGTT  
GACGTTGATGTTCGAACTTGCCTGCGGACTTAATGGTGCTCTGTACTTGTGCAATGGAT  
GCTGACGGCGGAGTCTCAAGTACCCAAACAACAGGCTGGCCAAGTACGGTACGGCT  
ATTGTGACAGTCAGTGCCTAGAGATCTAAATTGATGGCGAGGCCAACGTCGAGGGC  
TGGCAACCGAGCAGCAATAACGCCAATACTGGAATCGGCACCGCCTGTTGCGCG  
AGATGGACGTGTGGGAAGCGAACCTCATTAGTAATGCTGTAACACCCATCCGTGCGACACT  
CCGGGACAGACGATGTGTTCCGGTGACGATTGTGGCGGACCTACTCCAACGATCGATACGC  
GGGTACATGCGATCCGACGGCTGCGATTCAACCCATACGGGATGGTAATACATCCTTCT  
ATGGACCAGGTAAGATTATCGACACTACGAAGCCTTCACCGTGGTCACGCAGTTTGACA  
GATGACGGCACTGATACCGAACCCCTTCGGAAATCAAGCGCTTACATCCAGAACTCTAA

TGTTATTCCCCAACCAACAGTGACATTCCGGAGTTACTGGAAACTCTATCACTACAGAGT  
TTTGTACAGCCCCAAAAGCAGGCCCTCGGGGACACCGACGACTTCTCAGCACGGAGGCCTG  
GCCAAAATGGGAGCTGCCATGCAACAGGGATGGTCTGCTATGTCTGTGGATGACTA  
CGCCGCTCAAATGCTGTGGCTGGACTCTGACTACCCCACGTGATGCCATCCAACACTACCCCTG  
GCATGCCAGAGGTACTTGTCCCACCGACAGCGGCGTCCGACGTAGAGTCCCAGTCG  
CCCAACAGCTATGTGACGTACTCGAACATTAAGTTCGGCCCTATCAACTCTACGTTCACCGC  
ATCTAACCCACCTGGTGGTAACCGAGGTACTACGACCACTCGACGTCTGCTACTACCAACTG  
GTTCGTCACCCGGACCCACCCAGTCTCACTACGGACAGTGTGGCGGAATTGGATACTCTGGT  
CCCACCGTGTGCGTCTGGAACCTACCTGTCAAGGTCTCAACCCCTACTACTCCAATGCCTG  
TAATAGGCAATTAAACAGATAGTTGCCGGTGATAATTCTCTAACCTCCCACACTCCTTGAC  
ATAACGATTATGTAACGAAACTGAAATTGACCAGATATTGTTGAAATAGAAAATCTGGC  
TTGTAGGTGGCAAAATGCCCGTCTTGTTCATCAATTCCCTGTGACTACTCGTCATCCCT  
TTATGTTGACTGTCGTATTCTTATTTCCATACATATGCAAGTGAGATGCCGTGTCGAA  
TTCGACGCAGTAGGATGTCCTGCACGGTCTTTGTGGGGTGTGGAGAAAGGGGTGCTTGG  
AGATGGAAGCCGGTAGAACCGGGCTGCTGTGCTGGAGATGGAAGCCGGTAGAACCGGGC  
TGCTTGGGGGATTGGGCCGCTGGCTCAAAGAGGGTAGGCATTGTTGGGTTACG  
TAATTGCCGATTGGCTCTGCCATGCGCATGCCCATTGGTCAGAATTAGTCCGGATAGGAGA  
CTTATCAGCCAATCACAGCGCCGGATCCACCTGTAGGTTGGGTTGGGAGCACCCCTCC  
ACAGAGTAGAGTCAAACACAGCAGCAGCAACATGATAGTTGGGTTGTGCGTGTAAAGGAAA  
AAAAAGAAGCTGGTTATTCCGCTCTATTAGGTTGGGATAGACGCCGACGGAG  
GGCAATGGCGCCATGGAACCTTGCCTGCGGATATCGATACGCCGCGGACTGCGTCCGAACCA  
GCTCCAGCAGCGTTTCCGGCATTGAGCCACTGCGACCCCGCCAACGTGTCTTGGCC  
CACGCACTCATGTCATGTTGGTGTGGAGGCCACTTTAAAGTAGCACAAGGCACCTAGCT  
CGCAGCAAGGTGTCGAACCAAAGAAGCGGCTGCAGTGGTCAAACGGGCGGAAACGGC  
GGGAAAAGCCACGGGGCACGAATTGAGGCACGCCCTCGAATTGAGACGAGTCACGGCC  
CCATTGCCCGCGCAATGGCTCGCCAACGCCGGTCTTGCACCACATCAGGTTACCCAA  
GCCAAACCTTGTTAAAAAGCTTAACATATTACCGAACGTAGGTTGGGCGGCTTGC  
TCCGTCTGCCAAGGCAACATTATATAAGGGTCTGCATGCCGGCTCAATTGAATCTTTT  
CTTCTCTCTCTATATTCAATTCTGAATTAAACACACATCAACATGAAGCTCGTACCGC  
CTTACTATTCTACGGCCGTTCTGCCAGGCGTGCAGTAGTGTGTTGGGCGAGTGCCTG  
GTCAAAACTGGCTGGCCGACTGCTGTGCTCTGGTCACCTGTGTTACTCGAACGATT  
ACTACTCTCAATGCCGCCGGAGCTGCCAGCTCCAGTTCATCCACGAGAGCCCGTCTACA  
ACATCACGAGGTGCCCCACGACCTCTCGCTCATCGAGCGCAACTCCTCCCCCTGGGTCTACC  
ACAACCCGGTCCCACCTGTGGGACTGGGAACTGCCACGTACTCCGGCAATCCTTTGTGGG  
CGTGACACCTTGGCAAACGCCACTACGCTAGTGAAGTGTGCGCTGCCATTCCCTCTC  
TCACTGGAGCTATGGCAACGGCTGCCGTGCCCAAAGGTCCCCCTCATGTGGCTG  
GACACTCTGATAAGACCCACTCATGGAGCAAACCTGGCAGACATTGAACACTCGAAC  
AAAACGGTGGTAACACGCCGACAGTTGTTGAGTACTCTGACATCCGTACCCCTCGTAC  
GCCCTGGCCTCTAACGGCGAATATTCTATTGCGGATGGAGGGTGCCTAACGTT  
TATTGACACGATCCGGCAGATCGCTGTTGAGTACTCTGACATCCGTACCCCTCGTAC  
GCCGATTCCCTGCCAACCTGGTACCAACCTGGTACACCTAACGTTGCAACACGCCAGT  
CTGCTTACCTGGAGGTGTTAAACTATGCTGTTACCCAACTGAACCTCCCCAATGTC  
ACCTCGATGCTGGACACGCCGGTTGGCTGGGCTGCTAACGAGCCGCTGCTCAG  
CTTTGCTAATGTATAACAAAACGCTCCTCCCCAGGGCCCTCGAGGACTCGCCACTAA

CGTTGCCAACTACAACGGTTGGAACATTACCTCCCCATCCTACACCCAGGGTAACGCAG  
TGTATAACGAGAAGCTGTACATCCATGCTATCGGACCCCTGCTGCCAATCATGGATGGAGC  
AACGCCTTTCATCACCGACCAGGGCCGATCGGGCAAGCAGCCCACCGGCCAGCAGCAGT  
GGGCGATTGGGTAAATGTGATCGGAACCGGTTGGTATTGTCATCTGCCAATACAGGC  
GATTCCCTGCTCGACTCATTGTGTTGGTCAAGCCTGGCGAGAGTGTGACGGCACCTCGGA  
CAGCTCCGCTCCCGATTGATAGCCACTGCGCACTGCCGGACGCTTGAGCCCCGCTCCTC  
AGGCCGGAGCGTGGTCCAAGCCTACTTGAGCTATTGACTAACGCCAACCCCTCGTT  
CTATGATAAGCTATTATCACTCTTACAACCTACCTCAACTATCTACTTTAATAATGAA  
TATCGTTATTCTATGATTACTGTATATGCGTTCTCTAAGACAAATCGAAACCAGCATGC  
GATCGAATGGCATACAAAAGTTCTTCCGAAGTTGATCAATGTCCTGATAGTCAGGCAGCT  
GAGAAGATTGACACAGGTGGAGGCCGTAGGGAACCGATCAACCTGTCCTACCAGCGTACGA  
ATGGCAAATGACGGGTCAAAGCCTGAATCCTGCAATGGTGCCTGGATACTGATGTCAC  
AAACTTAAGAAGCAGCCGCTTGTCTCTCGAAACTGAGTTGGCGCCGTTTCGAG  
CCCCACACGTTCCGGTAGTATGAGCGCGGCAGATTGAGCGTTCCGGTTCCGCGCTG  
GACGAGAGCCCATGATGGGGCTCCCACCACAGCAATCAGGGCCCTGATTACACACCCAC  
CTGTAATGTCATGCTGTTCATCGTGGTAATGCTGCTGTGCTGTGTTGGTTGG  
CGCTCATTGTTGCGTTATGCGCTACACCACAATTGGAAGCTTATTAGCCTTCTATT  
TCGTTGCAAGGCTTAACAACATTGCTGAGAGGGATGGGATATGGAGGCCGCTGGA  
GGGAGTCGGAGAGGCCGTTGGAGCGGCTGGCCTGGCCTGGGCCAGCTCGCAAACGCC  
GGACCCCTTGGCACGCCAAATGTGCCACTTTCACTGAGCTAGTAACGCCAACAGTACAGT  
CATGCATGCATGTTGCCCTTTCCCTGCCCTGATGCCAACACAGTACAGTGC  
ACAGTGGAGGTTGGGGGGCTTAGATGGAGCTAAAAGCGGCCAGCGGTACACTAGT  
GGGATTGTATGGAGTGGCATGGAGCCTAGGTGGAGCCGTACAGGACGCACGCC  
CCGTGACAGACGATGGTGGCTCCTGTTGTCACCGCGTACAAATGTTGGCCAAAGT  
GTCAGCCTGCTGCAACCTAATCCCATTGTCAGTCGCACCCCCATTGATCGAGGCC  
TAACCCCTGCCATCAGGCAATCCAATTAAAGCTCGCATTGTCAGTCCTGTTAGTTGG  
TGCCCGTTCGCGTCCACTGACAAACACAAAGCATTATATAAGGCTCGTCTC  
CCTCCAACCACACTCACTTTGCCGCTTCCTGCTAACACAAAAGTCAAGAACACAA  
ACAACCACCCAAACCCCTACACACAAAGACATATCTACAGCAATGAAGCTCG  
TACTATTCTACGGCCGTTCTGGCCAACAGACGGTGTGGGCCAATGTGGCG  
TGGTCGGGCTCGACCAACTGTGCTCTGGCTCGGCCGTTGACTCTGAACCC  
CAGTGTATCCCCGGAGCGACCACATCACCGTCACTCGACCCCTCAGGACCC  
CACCACCGAGCGACCTCCACCTCAGCTACTCCCCCTACATCCTGGTGT  
TGGCGTGAACATGCCGGTTGACTTCGGATGCACGACTGACGGTACATCG  
AGGTCTATCCCCGGTTGAAAAACTTCACCGGGCTAACAAATTATCCTG  
AGCAGTCAGTGGCATTGGTCAG  
ATGCAACACTTGTCAACGAGGATGGCATGACCATCTCCGACTCC  
TCTAGTGAACAATAACCTGGCGAACCTGACAGCACAAGCATCTCAAGTAC  
TTGGTGCAGGGATGCCTCAGTTGGGTGCTACTGCATCGTGA  
GTGGAACGGGGGAATCATCGGACAGGGTGGCCCCACTAAC  
TCCCAGCTGCCCTAACGCTACAGCAGTCAGTGGCTGGTTG  
CCATGACGTCAACATTAAACACTTGGGCTGCCACCGTACAGGAAG  
ACGCCGGAGCCACCTCGCAGTCATCTCGCTGCC  
TTCATTTCCGATGGATCAGCAGCCCTCTCAG  
AACCTGATTTGACGTGCATAAGTAC  
CTGGACTCGGATAACAGCGGAACCCACCGCGGAGT

GTACCACTAACAAACATTGACGGTGCCTTCCTCGCAACGTGGCTCAGACAGAACAAAC  
AGACAGGCCATTCTGACTGAAACTGGTGGCGGCAACGTTCAATCATGTATCCAGGACATGTG  
CCAGCAGATTCACTGAATCAGAATAGTGACGTGTACCTGGTACGTGGTGGGAG  
CAGGGTCTTGTATTCTACCTACGTCCTGACCAGACTCCGACATCCTCCGGCAACTCGTGG  
ACTGATACTCTTGTTCGTATGCCTGCCGCAAGTGATAAAGCTTGCGCGAAACTC  
GATTCTCACCCCTCGATAACTCGACTCACCCCCCTTAACCTAAAATTCACTTACGACAAACAA  
ACGTCTGATACCGACTACCCCCCTCGACTTCTCGCAATCTGACTTCAATCAGAGGACCTCA  
AACAAACACTTTTCTTACGATTCTAATTATTACCCATTCAATTATCCCCTGCGCTCGTC  
CAGCAATGTCCGAGAGCATCCTGCCGTGTCTCTGGGCCATCCATTAAATTGGGTCCATCC  
TCCCGGGCAGATTCCACAGTCCAGTTGCGACTGGATGGTAGATCCGCTTTTTA  
GTTTGAACATGTTGCGAGTTGATACTGAACATCAGAGTTAGCTTGTACTAATATACA  
TTCCTTTAGTGAGGTCTGTACACTGTACACTGTATTATATTGAGCTTGTATTGGTAC  
CTTGGTTGTTGTTACCGGTACAAGTATGTACGATCGTACCGTAGATCTCCAG  
TGGTGCATGAACGCATGAGAAAGCCCCCGAAGATCATCTCCGGGGCTTTTTGGCG  
CGCGATACAGACCGGTTCAGACAGGATAAAGAGGAACGCAGAATGTTAGACAAACACCGCT  
TACGCATAGCTATTCAAGAAATCAGGCCGTTAACCGATGATTACGAGAATTGCTGGCCGC  
TGCAGCATAAAATTACACACTCAGCGCCTGATTGCGATGGCGAAAACATGCCGAT  
TGATATCCTGCGCGTGTGATGATGACATTCCGGGTCTGGTAATGGATGGCGTGGTCGATC  
TCGGTATTATCGCGAAAACGTGCTGGAAGAAGAGACTACTCAACCGCCGCGCACAGGGCA  
AGATCCACGCTATTAACCCCTGCCGTCTGACTTCGGCGCTGCCATTATCGCTGGCAAC  
ACCGGTTGACGAAGCCTGGGACGGCCGGCGCTGGACGGTAAACGTATCGCTACCTCA  
TATCCGCACCTCCTCAAACGCTACCTCGACCAAGAAAGCGCTCTTTAAATCGTGTCTGTTA  
AATGGTTCTGTCGAAGTCGCGCCGCGCGGGCTGGCGACGCTATCGGATTTGGTCTC  
TACCGCGCGACGCTTGAAGCTAACGGCCTGCGTGAAGTCAAGTTATCTACCGCTAAAG  
CCTGCTGATTCAAGCGCAGCGTGAAGATGGCACAGAGCAAGCAAGAGCTGATCGATAAATT  
GCTGACCGTATTCAAGGGCGTGAATTCAAGCGCGCGAATCGAAATACATCATGATGCACGCGC  
CAAGTGAACGCCTGGAAGAGGTTATGCCCTGCTGCCAGGCAGCGAACAGTTGTT  
GCCGCTGGCAGGCAGCAACAGCGCTGGCGATGCACATGGTCAGCAGCGAACAGTTGTT  
TGGGAAACCATGGAGAAACTGAAAGCGCTGGCGCCAGCTGATTCTGGTACTGCCGATCG  
AGAAGATGATGGAGTGACTIONGACGCGCTGATGGCGCTGCGTTATCAGGCCTACGTAATCGT  
TGATATTGGGTTCTGTAGGCCGATAAGGCCGAAACCTGTGATGGAGTAAAGACCATGAG  
CTTCAATACCCCTGATTGACTGGAACAGCGGATCTGGCGACGAGTATCTGCTGACTCGTCA  
TTGCCGCTTGGAGTACGACTCCAACATGAGTGTGCTGGATCACTTGACGATACTTCT  
TCGTTGGAGGCTGGGTCTGACAGCTGCCTTCGGCGCGTTGGCGACAAACAATATCAG  
CTGCAACGTCATTGCTGGCTTCACTCATGATCACATTGTCGGCAAAGGCAGGCCAGA  
GAGCCATTGACGTTCTTCAATTGGACCGATAGCGTATAGTCCAGTCTATCTATAAGTC  
AACTAACTCGTAACTATTACCATACATACCTCACTGCCAGATAAGGTTCCGATAAAA  
AGTTCTGAGACTAAATTATTCAGTCTCCTCTTCAACCACCAAAATGCCCTCTACGAAGCT  
CGAGCTAACGTCACAAGTCCGCTTGGCGCTCGAGTGCTCAAGCTCGTGGCAGCCAAGAA  
ACCAACCTGTGCTCTGGATGTTACCAACCAAGGAGCTATTGAGCTTGGCGATA  
AGGTCGGACCTTATGTGTGCATGATCAAGACCCATATCGACATCATTGACGACTTCACCTAC  
GCCGGCACTGTGCTCCCCCTCAAGGAACCTGCTCTTAAAGCACGGTTCTCCTGTTGAGGAC  
AGAAAGTCGAGATATTGGCAACACTGTCAAGCACCAGTACAAGAACGGTGTCTACCGAA  
TCGCCGAGTGGTCCGATATCACCAACGCCACGGTGTACCCGGAACCGGAATCATTGCTGGC

CTGCGAGCTGGTGCCGAGGAAACTGTCTCTGAACAGAAGAAGGAGGACGTCTGACTACG  
AGAACTCCCAGTACAAGGAGTCCTGGTCCCTCTCCCAACGAGAAGCTGGCCAGAGGTCTG  
CTCATGCTGGCCGAGCTGTCTGCAAGGGCTCTGGCCACTGGCGAGTACTCCAAGCAGAC  
CATTGAGCTTCCCCATCCGACCCCGAGTTGTGGTCTCATTGCCAGAACCGACCTA  
AGGGCGACTCTGAGGACTGGCTTATTCTGACCCCCGGGTGGGCTTGACGACAAGGGAGA  
CGCTCTCGGACAGCAGTACCGAAGTGTGAGGATGTCATGTCTACCGGAACGGATATCATAA  
TTGTGGCCGAGGTCTGTACGCCAGAACCGAGATCCTATTGAGGAGGCCAACGATACCA  
GAAGGCTGGCTGGAGGCTTACAGAAGATTAACGTAGAGGTTAGACTATGGATATGTC  
ATTTAAGTGTATATAGAGAGCGTGCAAGTATGGAGCGCTGTTAGCTTGTATGATGGTC  
AGACGACCTGTCTGATCGAGTATGTATGATACTGCACAACCTGTATCCGATGATCTGTC  
CAATGGGCATGTTGTTCTCGATACGGAGATGCTGGGTACAAGTAGCTAACGAT  
TGAACACTTAACTTATATGAGGCTGAAAGAAAGCTGACTTGTATGACTTATTCTCAACT  
ACATCCCCAGTCACAATACCACACTGCACGGATCTCCAGTGGCATGAACGCATGAGAA  
AGCCCCCGGAAGATCATCTCCGGGGCTTTTTGGCGCGCGATACAGACCGGTTCGA  
CAGGATAAAGAGGAACGCAGAATGTTAGACAACACCCGCTACGCATAGCTATTAGAAAT  
CAGGCCGTTAACCGATGATTACGAGAATTGCTGGCCGCTGCCATAAAAAATTAAATTAA  
CACACTCAGCGCCTGATTGCGATGGCGAAAACATGCCGATTGATATCCTGCGCGTGC  
TGATGACATTCCGGGCTGGTATGGATGGCGTGGTCGATCTCGGTATTATCGCGAAAACG  
TGCTGGAAGAAGAGCTACTCAACCGCCGCACAGGGGAAGATCCACGCTATTAAACCC  
GCGCGTCTGACTTCGGCGCTGCCGTTATCGCTGGCAACACCGGTTGACGAAGCCTGG  
ACGGCCCGCGCGCTGGACGGTAAACGTATCGTACCTCATATCCGACCTCCTCAAACGC  
TACCTCGACCAGAAAGCGCTCTTTAAATCGTGTGTTAAATGGTCTGCGAAGTCGCG  
CCGCGCGGGGCTGGCCGACGCTATCTCGGATTGGTCTCTACCGCGCGACGCTGAAAGC  
TAACGGCCTGCGTGAAGTCGAAGTTATCTACCGCTCTAAAGCCTGCTGATTAGCGCG  
GTGAGATGGCACAGAGCAAGCAAGAGCTGATCGATAAATTGCTGACCCGTATTAGGG  
GATTAGCGCGCGAATCGAAATACATCATGATGCGACGCGCCAAGTGAACGCC  
GTTATGCCCTGCTGCCAGGCGCCAAAGGCCACAATTCTGCCGCTGGCAGGCGAGCAAC  
AGCGCGTGGCGATGCGACATGGTCAGCGAGCGAAACGTTCTGGAAACCATGGAGAA  
GAAAGCGCTGGCGCCAGCTGATTGACTGCCGATCGAGAAGATGATGGAGTGATCTG  
ACGCCTGATGGCGCTGCGCTTATCAGGCCCTACGTAATCGTTGATATTGGGTTCTG  
CGGATAAGGCAGGAAACCGCTGTGATGGAGTAAAGACCATGAGCTCAATACCG  
GAACAGCGGATCTGGCGCCGCACTGTAGAGCACACTAGGGATTAGAGGGATTATG  
GCACGTACAATATAGATAATTAGCAGTAGCTGAGTCAGTTGAATGATCAGAGGT  
TGAGTGTGGATGGATTGTAGAGTCGTTAAAATAATGAGTTAAGAATAATTACGA  
CTACAGGATACGATGTACTTGTATTGATCGATACAGTACATACAGTACATACGT  
ACTCCTAAACTGTGATCACCTACAACCTCAACTAGTCGGTCATAATTCA  
CCTGGTAGTCTAGGCCAAACACTCCTGGTAACCTTGTGAGAAAGATTGCTCT  
CTCCTGAATTCTCTAGGCTGGATTGCGCCAAAAGAACCCACAAACGGCG  
TCTTGACTCTAGGCTGGATTGCGATGAGCGTTAAAGTGGCCAGAGTCT  
TCCTGACGATGATACTCTCAAAGTGGCACGGTAGACATGACAATAAGAG  
TTTGCTCTGTAGTGGCGATGAGCGTTAAAGTGGCCAGAGTCT  
TCAGCTGATGCTTTGTTCTCGCTGAGTCCTCCGAGGCCAGGGG  
AGAATAGCGCCTGTCGTTAGAAGAGGTGCT  
AACTGAACAAATTCA  
GACTGAGGTGTC

CCTGTAAACCATGAGTGGTGCCTCAGTGACGACCAGAGCGAGATCATAATGTCTGGGAGA  
GGTTTCTTCATATTGTATCTGAGACAGCAAAGAAGTGCCTCAAAAGGCCGGTCCAAAGGT  
GTAGGCCTCTGTTCTGGAAGGGGGGTTCCGTTGCTGGCGTTGATTTAATAGATGCTT  
TGTGTTGAACGACTGGTGCAATTGGGCAGAAGAACGCTGCAGAATGATTAGAGCCACCTTA  
CTGTACATTCAGCTCGTGGGTTTTGTTGGTGGAAAAGAGGTACGAACAGAGAGAGAG  
AAAGGCTGCCAGGGCAGGCTCGGTGTCATAGTTCTGGTACGATGTCCTTATCGGTAATGA  
GTGTCTTGCAAACAACTTGTTCTGCACGAATTCCAAAGTGGGGGCCGGCC

## References

- Alani, E., Cao, L., and Kleckner, N. (1987). A method for gene disruption that allows repeated use of URA3 selection in the construction of multiply disrupted yeast strains. *Genetics* 116, 541-545.
- Tai, M., and Stephanopoulos, G. (2013). Engineering the push and pull of lipid biosynthesis in oleaginous yeast *Yarrowia lipolytica* for biofuel production. *Metabolic Engineering* 15, 1-9.
- Voth, W.P., Richards, J.D., Shaw, J.M., and Stillman, D.J. (2001). Yeast vectors for integration at the HO locus. *Nucleic Acids Research* 29, e59-e59.
- Wei, H., Wang, W., Alahuhta, M., Vander Wall, T., Baker, J.O., Taylor, L.E., Decker, S.R., Himmel, M.E., and Zhang, M. (2014). Engineering towards a complete heterologous cellulase secretome in *Yarrowia lipolytica* reveals its potential for consolidated bioprocessing. *Biotechnology for Biofuels* 7, 148.
- Wei, H., Wang, W., Alper, H.S., Xu, Q., Knoshaug, E.P., Van Wychen, S., Lin, C.-Y., Luo, Y., Decker, S.R., and Himmel, M.E. (2019). Ameliorating the metabolic burden of the co-expression of secreted fungal cellulases in a high lipid-accumulating *Yarrowia lipolytica* strain by medium C/N ratio and a chemical chaperone. *Frontiers in microbiology* 9, 3276.
- Xu, Q., Knoshaug, E.P., Wang, W., Alahuhta, M., Baker, J.O., Yang, S., Wall, T., Decker, S.R., Himmel, M.E., Zhang, M., and Wei, H. (2017). Expression and secretion of fungal endoglucanase II and chimeric cellobiohydrolase I in the oleaginous yeast *Lipomyces starkeyi*. *Microbial cell factories* 16, 126.