

## Supplementary Data

1 CTAGTTAAAGATCTCTTCTCTCTCTAACTCTGGTTCTCTCTCTATCTCTCTCTA~~CC~~AAAAATCTACCCTCGAAAAATCTTCTCAATTTAACCAC  
 101 TCTGGAACTGCAGCTTCCTCT~~CA~~CAATACACGAATCAATTCTGTTGGAATCTCCATAGCTCTGAATCAAGGGCACTCTGTACACTCAGATTATTTCTGCTG  
 201 GCAAGCTTGCATGCTGTATATTTCAATCAATTTTGAAGCTTAACTTGGAGCTCTGTTGCATCCATAGTCAATGTTTGGACGAGCTATTACTTTTAATTT  
 301 GGATTGCTAAGATGCTGAATATAGTCTATTGATCCTGGATTGGTAGAAGCTTTTATGGAAATGTAGCTTTTACTCACTGCTTGACACCTTGCTTAATAG  
 401 ATAGGATTAT~~AG~~CCAAACATCAAAATGAGAAGTGGATCGACAATTTGCAGTTCTCTCATTATTTTGGCCGCTCCGCAAGATGTAGAACAAGAAAGGCTCA  
 M A T S N E K W I D N L Q F S S L F W P P P Q D V E Q R K A Q  
 501 AACCACTGCATATGTTGAATATTTGGGCGAGTTCCACCTCAGAACAAATTTGCTGATGATATAGCAGAGTTAATCCGTAGCCGATTCATCCAAGGAGAAC  
 T T A Y V E Y F G Q F T S E Q F A D D I A E L I R S R Y P S K E N  
 601 AGGCTTTTGGACGATGTTCTTGAACATTTTGTCTTCCATCCGAGCATGGCCATGCTGTTCTACTTCTATTATTTCATGATTATTCGATAGCACAC  
 R L F D D V L A T F V L H H P E H G H A V L L P I I S C I I D S T L  
 701 TTGAATATGACAGAAGTGGTCCACCATTTGCTTCTTTCATTTCCTGTTTGGCCAAACAGTGAGAATGAGTATTCGGGCGAGTGGGCCCTTGATGGCGG  
 E Y D R S G P P P F A S F I S L F C P N S E N E Y S G Q W A L A C G  
 801 GGAGATACTAAGAATTTTGACACACTATAACAGACCAATTTACAAGTTGGAACGCCACGAAAATTTATCCGATAGAAGTAGCAGCGGAGATCATGCTTCA  
 E I L R I L T H Y N R P I Y K L E R H E N L S D R S S S G H A S  
 901 ACTAGCAAGTCAATAGATGGAGAACCTTCTTCTCACCTCCCTCACTGGCTGAGAGGAAACCTCTGAGGCCACTGTCTCCCTGGATTACTGATATCTTGC  
 T S K S I D G E P S S P S S L A E R K P L R P L S P W I T D I L L  
 1001 TTGCCACGCTCCCTGGTATAAGGAGTGATTATTTTCGATGGTGGAGGTGTTATGGGAAAATATGCCGCTGGAGAGCTGAAACCGCTATAACTTGCTAC  
 A T S L G I R S D Y F R W C G G V M G K Y A A G E L K P P I T A T  
 1101 TTCTCGTGGATCTGGGAAGCATCCCAGCTTATGCCATCAACTCCGAGTGGGCTGTGCTAATGGTCTGGAGTTATACTTAGTGTCTGTGACGAGGAA  
 S R G S G K H P Q L M P S T P R W A V A N G A G V I L S D G E E  
 1201 GTTCTCGCTATGAACTGCTACCTTGACTGCAGCTGCTGTCTTCTGCACTTCTGCTTCCACCTCCAACACACCTACAGATGAACATCTAGTCGCTGGTC  
 V S R Y E T A T L T A A A V P A L L L P P P T T P T D E H L A G L  
 1301 TGCCAGCGCTGAGCCCTATGCAGCTCTTTTTCACAGATATTATGCAATTGCTAGTCCAAGTGCCACCCAAAGGCTGCTTCTTGGACTTCTAGAGGCTCC  
 P A L E P Y A R L F H R Y Y A I A S P S A T Q R L L L L G L L E  
 1401 GCCTCTGGCTCCAGATGCTCTTGACGCTGCTGACAACCTAGTGGAACTTCTTGAGCCGCGAGAAGACTATGCCACAGGAGTGAAGCTCCCTAGGAAT  
 P S W A P D A L D A A V Q L V E L L R A A E D Y A T G M R L P R N  
 1501 TGGATGCACTTGCAATTTCTGCGAGCAATTTGGGATTGCAATGCTATGCGAGCTGGAATTGCTGCAGATTGCTGCAGCTGCCTTACTTTTCGGAATCTTT  
 W M H L H F L R A I G I A M S M R A G I A A D S A A A L F R I L S  
 1601 CTCACACAGCTTTACTTTTCCCCCACTGAGACAAGTTGATGGTATTGAATTCGCCACCCAGCCTCTAGGTGGTTATATTTCACTGAAAGAAAGCAGCA  
 Q P A L L F P P L R Q V D G I E F P H Q P L G G Y I S S E R K Q Q  
 1701 ACTTGAATTGACTGCAGCAGAACAGTACAAGCTACTGCCAAGGGAATGCATCTATGCTTTGTGCACATGGTCCAGAGGTTGAATGGCGAATTTGT  
 L E L T A A E A T V E A T A Q G I A S M L C A H G P E V E W R I C  
 1801 ACGATATGGGAAGCTGCTTATGGTTGATACTTTAAGTTCTTCTGCACTGATCTTCCAGATATCATAGTCGCAACACCGCTGCAGCCTCCGATACTAT  
 T I W E A A Y G L I P L S S S A V D L P D I I V A T P L Q P P I L S  
 1901 CATGGAATTTGATACCTCTTCTTAAAGTACTCGAATATCTTCTCGTGGGAAGCCACTGAAACCTGCTTATGAAAATATTTGTGCTACTGTTGA  
 W N L Y I P L L K V L E Y L P R G S P S E T C L M K I F V A T V E  
 2001 AGCAATTCTACAGAGGACATCCCACCTGAGTCTCTAGAGAACAACTAGGAAACAGATATATATTTGGGCTGCATGCCAAGAATCTTCTGCTGGCT  
 A I L Q R T F P P E S S R E Q T R K T R Y I F G S A S K N L A V A  
 2101 GAGCTTCGCACAATGGTCCACTCATTATCTTAGAGTCGTGCTTCTGAGAGCTTGCCCTCCGCTTCTTTTGTGTTTAACTGCTGTGCTGAGCC  
 E L R T M V H S L F L E S C A S V E L A S R L L F V V L T V C V S H  
 2201 ATGAAGCGCATCCCAACAGTAGCAAGCGCCCAAGGTGAAAAATCTCATCCGAAACATGTGGTGAAGACGGATGGAAGGAGAGAGAAATCGAAAAATAA  
 E A H P N S S K R P K G E K S H P E H V G E D G W K E R E S E N N  
 2301 TAAACAGTGCAGAAACAGGGCCCTGAGTGCATTGATTCTTATGTCATTGCTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCT  
 K Q C K K Q G P V A A F D S Y V I A A V C A L S C E L Q F F P L I  
 2401 TCTAAAGGAAGTAGTCTTTTAGGTGCTAGAAATGGTCTGACGTTGCTAAGCCTGCAAACTGAATGACCTGTCCAGTCAGTTGCAGAGTGGTATTGATT  
 S K G S S L L G A R N G S D V A K P A K L N D L S S Q G I D S  
 2501 CAGCTGTTGATCACACTCGCAGAAATATAGCAATTTAGAGGCCCTTTTCTTGTAAACCGTCTTCGATTGGCACTTCATGGAGTTACAGTTCAATAGA  
 A V D H T R R I L A I L E A L F S L K P S S I G T S W S Y S S N E  
 2601 AATAGTGGCTGCTGCTATGGTTGACGCTCATGTTTCTGATTGTTTGTAGACGCTCAAAAGCATGCTGCTGCTTCTTCCACCTGGTGAGATGCAAGTGG  
 I V A A A M V A A H V S D L F R R S K A C M R A L S T L V R C K W  
 2701 GATAGTGAATCTACTAGGGCGTCTCCCTCTCAATTTGATTGATATTACAGCAAAAGTTGTCGCTATATAGTCAACAAGCAGAGCCATTAGAGG  
 D S E I H S R A S S L F N L I D I H S K V V A S I V N K A P L E A  
 2801 CTCACCTACTCCATGCGCCATCGCTGAAGGAAATATCTCATGTTTTCATGGAAAAAGCTTGCAAAATGCGGTAGCTGTGCGCACTCAGAACAGGACA  
 H L L H A P S L K E I S S C F H G K K L A N C G S C R H S E P G H  
 2901 TCCATCTTATCCCGGTGGAAGACTTGCCCGTCTGAAGATTGGTTAATCATAGGAAGGTTGATTCTGGCGAAGCTGGAAGATGTACAATGGGCAAA  
 P S S S P C E D L P R S E D L V N H R K V D S G E A G R C T M G K  
 3001 GGAATTGGCGATTTCCTATGGATGCTTCAGATTGGCCAACTCTTACTATGGACAGGCATATAGGATTTAATTGCAACGGCAATGTTCTTCTGAGGT  
 G I A D F P M D A S D L A N F L T M D R H I G F N C N G N V L L R S  
 3101 CTGTCCTTGTGAAAAGCAAGAACTGTGTTTTCAAGTTGTTTCAATTGCTTTGGCACAATGATTGTATACCCGAGACACAACCAAGTGTGAGAGCAC  
 V L A E K Q G E L C F S V V S L L W H K L I V S P E T Q Q P S A E S T  
 3201 GTCTGCCAGCAAGGATGGAGACAGGTGGTGTGCACTTGCATGTAGTATCGGCATCACGCAAAAGCTGCAACAGCCGTTGCTTCCAGCGGAA  
 S A Q Q G W R Q V V A A L C N V V S A S P A K A A T A V V L Q A E  
 3301 AGGAATTGAAACCATGGATCGCTAAGATGATGATCTTGGGCAAAAAATGTGGAGAGTCAACAGAGAATCGTGAAGGTGATTGTGCGATTATGAGGA  
 R E L K P W I A K D D D L G Q K M W R V N Q R I V K V I T E L M R N  
 3401 ACCATGAGGAGCGGAGTCGCTAGTAATTTCTGCTAGCGCATCAGACGCTCCTCCTCGTCCACAGACGGGATGCTTGTGATGGAGAAGCATGCACCTT  
 H E A A E S L V I L A S A S D V L L R A T D G M L V D G E A C T L  
 3501 ACCTCAGCTGGAGCTTCTGGAAGTAACGGCGAGAGCAGTTCAAGCCGTGCTTGAATGGGGAGAGTCAGGATTAGCAGTTGCGAGATGGCTCTCAAAAGCTG  
 P Q L E L L E V T A R A V Q A V L E W G E S G L A V A D G L S N L  
 3601 TTGAAGTGTGCGCTCTCAGCCACCTGCTGCTGCTCTCACCCGAGCGCTCATGTCCGGCGCTGAGCACCTCAGTTCTCCGTGCAATCTCACACAACG  
 L K C R L S A T V R C V S H P S A H V R A L S T S V L R A I L H N G  
 3701 GGTGCGTGAACCAAGAGCAAACTACCGGAGCTGAACATAATCCACAATCCGGGTATCAGCAGTACTGAACAGATAGGACCACTTGAAGTGGCAAGC  
 S V K P K S K L P D V N I I H N P R Y Q Q Y L N K I G T I D W Q A  
 3801 TGACATTGAGAAGTCTGACGTGGGAAGCGCAGCGCGCTTGCAACGGGGCTGCCATCGAGTTTGTGGGTGCGCTGCGAAGGAGTTAGGTTGTGCT  
 D I E K C L T W E A H S R L A T G L P I E F V G S A A K E L G C A  
 3901 ATATCTGTG~~TGA~~AAGTAGTTTTAAACATTTGATGCTGTTTGAGCTAAAGAAACACTATGACTGAGATTGA~~AATTAG~~ATTATTGCTATT~~CA~~CCCTCTGT~~T~~  
 I S V \*  
 4001 AATAATTTTTCTAGACTACGCCATGTATGTATAATATTTA~~TATTA~~TTGTAGTATTATTTCTAT~~CA~~AAAAAAAAAAAAAAAAAAAAAA  
*ShGt-1*

*ShGI-2*

The start codon ATG and the stop codon TGA/TAG are in underlined bold face. Transcription start sites and polyadenylation sites are in underlined and italic bold face. The possible polyadenylation signals are in italic bold face. The 5'-UTR intron in *ShGl-1* is labeled with grey background. Poly A tail is in lower case. The sites corresponding to the qRT-PCR primers are underlined.