

Annex 4 Figure S1 The coding sequence of SiER1_X4 gene

GenBank accession number: OP492075

SiER1_X4 (1-2973)

Total amino acid number: 990, MW=107753

Max ORF starts at AA pos 1 (may be DNA pos 1) for 990 AA (2970 bases), MW=107753

```

10      20      30      40      50      60      70      80      90
1      ATGACCCGCCTCCTCCGGGCCTCGCCGCTTCCTCCTCCTCGCGCCGTCGCCGTCGCCGACGACGCGGGCAACACTGCTGGAGATCAAG
1      M T R L L R A L A A F L L L A A V A V A D D G A T L L E I K

100     110     120     130     140     150     160     170     180
91     AAATCCTTCAGCAACGGAGGCAACGCGCTGCACGATTGGTCCGGCGAGGGCGCATCGCCGACCTACTGCTCGTGGCGGGCGTGTCTATGC
31     K S F S N G G N A L H D W S G E G A S P T Y C S W R G V L C

190     200     210     220     230     240     250     260     270
181    GACAACGTCACCTTCGCTGTGCGGCGCTCAACCTTTTCAGGGCTCAATCTCGAGGGTGAAATCTCACCGGCATCGGGAGCCTGAAACGT
61    D N V T F A V A A L N L S G L N L E G E I S P A I G S L K R

280     290     300     310     320     330     340     350     360
271    GTTATCTCAATAGATTGAAAGTCTAATGGACTTTCGGGCGAGATCCCTGATGAGATTGGCGATTGTTGCTTGCTTAAACTTTGGACTTG
91    V I S I D L K S N G L S G Q I P D E I G D C S L L E T L D L

370     380     390     400     410     420     430     440     450
361    TCCTTAACAATCTAGAAGGAGACATACCATTCTCCATATCTAAGCTCAAGCACCTTGAGAACCTGATTTGAAGAACAACCACTGGTT
121    S S N N L E G D I P F S I S K L K H L E N L I L K N N Q L V

460     470     480     490     500     510     520     530     540
451    GGGGTGATACCATCAACACTATCTCAACTTCCATTTGAAGATATTGGACTTGGCTCAAAAACAGCTAAGTGGTGAAATCCAAACCTA
151    G V I P S T L S Q L P N L K I L D L A Q N K L S G E I P N L

550     560     570     580     590     600     610     620     630
541    ATATATTGGAATGAGGTTCTCAATACTTGGGATTGCGGAGCAATAATTTAGAAGGAAGCCTCTCTCCGGATATGTGCCAGTTAACTGGC
181    I Y W N E V L Q Y L G L R S N N L E G S L S P D M C Q L T G

640     650     660     670     680     690     700     710     720
631    CTGTGTACTTTGATGTGAAGAATAATAGCTTGTGGGTACGATACCAGAAACCATAGGGAAGTGTACCAGCTTTCAGGCTTGGATTG
211    L W Y F D V K N N S L M G T I P E T I G N C T S F Q V L D L

730     740     750     760     770     780     790     800     810
721    TCAAACAATCAGCTTACTGGAGAAATCCCATTCAATATTGGTTTCTGCAAGTGGCTACGCTGTCTTTGCAAGGGAACAAGTTCTCTGGC
241    S N N Q L T G E I P F N I G F L Q V A T L S L Q G N K F S G

820     830     840     850     860     870     880     890     900
811    CCAATTCATCAGTGATTGGCCTTATGCAGGCGCTTGCAGTGTCTGATCTGAGTTTCAATGAGCTATCTGGACCAATACCATCTATACTG
271    P I P S V I G L M Q A L A V L D L S F N E L S G P I P S I L

910     920     930     940     950     960     970     980     990
901    GGCAACTTGACATACACTGAGAAATTATACCTGCAAGGCAATAGGCTAACTGGATCGATACCTCCAGAGCTTGGTAATATGACGACACTG
301    G N L T Y T E K L Y L Q G N R L T G S I P P E L G N M T T L

1000    1010    1020    1030    1040    1050    1060    1070    1080
991    CATTACCTTGAAGTGAATGACAATCTATTGACTGGGTTTCTCTCTGATCTTGGGAAGCTCACAGAATTGTTTGACTTGAACCTTGCA
331    H Y L E L N D N L L T G F I P P D L G K L T E L F D L N L A

1090    1100    1110    1120    1130    1140    1150    1160    1170
1081   AACAAACCTTGGAGGACTATCCCTGATAATATAAGTTTCATGCATAAATCTCATTAGTTTGAATGCTTATGACAATAAATTAATGA
361   N N N L G G P I P D N I S S C I N L I S L N A Y D N K L N G

1180    1190    1200    1210    1220    1230    1240    1250    1260
1171   ACCATTCCACGTTTCAATTTCAAAGCTCGAGAGTCTGACTTATCTGAATCTGTCTCATCAAATCATCTCAGTGGAGCACTTCCGATTGAGGTA
391   T I P R S F Q K L E S L T Y L N L S S N H L S G A L P I E V

1270    1280    1290    1300    1310    1320    1330    1340    1350
1261   GCAAGATGAGAAATTTAGACACACTGGACTTATCCTGTAACATGATCACTGGTTCAATTCCTCAGCTATTGGAAGACTAGAGCATCTT
421   A R M R N L D T L D L S C N M I T G S I P S A I G R L E H L

1360    1370    1380    1390    1400    1410    1420    1430    1440
1351   CTGAGGCTCAATTTAAGCAAAAATGCTCTGGTTGGACACATTCTGCTGAGTTTGGGAAGCTTGGAGGATCATGGAGATCGATTGTC
451   L R L N L S K N A L V G H I P A E F G N L R S I M E I D L S

1450    1460    1470    1480    1490    1500    1510    1520    1530
1441   AGTAACTACCTCCGTGGCTGATTCCACAAGAGGTTGGGATGCTACAAAATCTGATACTGTAAAATAGAAAATAATAGTATTACTGGA
481   S N Y L R G L I P Q E V G M L Q N L I L L K L E N N S I T G

1540    1550    1560    1570    1580    1590    1600    1610    1620
1531   GATGTCTACCACTCACTAACTGTTTCAGTCTCAATAACTTAAATGATCATACAACGACCTTGCTGGTATTGTACCTACAGACAACAAC
511   D V S P L T N C F S L N N L N V S Y N D L A G I V P T D N N

1630    1640    1650    1660    1670    1680    1690    1700    1710
1621   TTCTCACGATTTTCACCTGACAGCTTCTTGGGTAACCTGGACTTTTGGGCTATTGGCGTGGTTCTCGTTCTCTGCTCTCTTTATCA
541   F S R F S P D S F L G N P G L C G Y W R G S R S S C S P L S

1720    1730    1740    1750    1760    1770    1780    1790    1800
1711   TCCAGTATTGAGCGAAAAGGAGATCATCTATATCAAAGGCTGCATTTCTTGGTATTGGTGTGGTGGGCTTGTATCTCTGCTCGTTATC
571   S S I E R K R R S S I S K A A F L G I G V G G L V I L L V I

1810    1820    1830    1840    1850    1860    1870    1880    1890
1801   CTAGCAGCTGCTTGTGGCGCACAACTCACCGTTCTCAAAGATGTCTCTGTAAGCAAACCAAGACAACCTTGTTCAGCATCAAGCAGT
601   L A A A C W P H N S P V L K D V S V S K P D N L V A A S S S

1900    1910    1920    1930    1940    1950    1960    1970    1980

```

