

Table S1. Splicing scheme and SNPs of *agNAT8* (ENSANGG00000015692) gene

#	Exon / Intron Chr 3	Start	End	Length	Sequence
1	5' upstream sequence ENSANGE00000062606	11089528	11090052	525 bpctgccgaaaagacctggaataaaactttactactctgtctttttccagata aacaacttcttcgcgcaacttcccaactctccaccacacctcagaccATGGAAGGACG GGATAATAAATGGGTTCATCGGGGACAACAGCCCCATCGATAG CGGGCCAGTACCGTTGGACTACACCGGCAGCACCGAATGGA GTTACACGTCACCCACGGAGGACTGGCCCTACTGATGTTGAG CTGGCTGTAGGAGGTAAAAC(a/t)GTCCCCAGACCAACTCCAG CTACAGG(g/a)GAAGATGCGGGACGAGTCCCTACC(c/t)C(c/t)AT AATACCCGAGCCCAATGCCGATCGGGACCAGTGGGGCAAGG GCGTCGAGTTCCTG(a/t)TGTCCTGCATTGCCATGTCGGTCCGG TTGGGCAACGTGTGGCGGTTCCGTTTCGTAGCGCTCGAGAAC GGAGGCGGCGC(a/g)TTCGTGATACCGTACAT(c/t)ATAGTGCT GCTGCTTGTTCGGCAAACCCGCTACTACATGGAGATGA(t/c)C ATCGGACAGTTCTCGAGCCGGGGCAGCGT(a/g)AGGTGTACG ATATGGCACCGATCATGCGAG
2	Intron 1-2 ENSANGE00000062603	11090053 11090650	11090649 11090878	597 bp 229 bp	gtaagtggaggcgtccttggactt.....cccttcttctctgacgcgccaag GTGT(t/c)GGTTACGGGCAGCTGTTCTCGGTGACCGCTCTAAT CACGTACTACTCCTCCCTAATGGCAC(t/c)AATTGCACGGTAC ATGATCGAC(t/g)CGTTCATG(a/t)ACCCGCTCCCGGGCCCA CTGTCTGCTCGGAATGGCAACCGAACTGCATCGATTCCGTGG CCGGCAGTGCAAGCCCTAATGACAGCTCATCCAATCGCACA CTGACCAGTAGTT(c/g)GGAGCTTTACTTTAC
3	Intron 2-3 ENSANGE00000062605	11090879 11091034	11091033 11091149	155 bp 116 bp	gtaagtttggcgggttcggttta.....ctcactagtctgttacttacag CAAGGT(a/g)GTCCCTGAAGGAGCTGGATGGCATTTGACGATGG AATCGGACTTCCAGATTTGAGACTAACGCTATTCCTGGTCC TATCTTGGTTCGCTCGTGTTCCTCACCTCAT(t/a)AAAG
4	Intron 3-4 ENSANGE00000062608	11091150 11091221	11091220 11091846	71 bp 626 bp	gtaacgaaaactattcttctca.....taaattcgtttcacacctttgcag GTGTTAAAAGCTCCGGCAAAGCCTCTACTTCTTAGCGCTC TTCCCGTACGTCGTCATG(a/g)CGGTGCTGCT(c/t)GTCCGCGC TTG(t/c)ACACT(c/g)CCGGG(c/t)GCAGTGGA(t/c)GGTATCGTGT ACTTCTCAAGCC(g/c)CAGTGGGACAAAATCTACGACCC(a/g)AAGGTGTGGT(a/g)CGCCGCCGTCACGCAGTGCTTCTTCG CTGTCGATCTGCTTCGGCAACATCATCATGTACTCGTCGTAC AACAA(a/g)TTCCGGCACAAATGTGTACCGCGATGCGACGATC GTGACGTTCGATCGATACGTTACGTCGCTGCTGGCCGGGTG TACAATCTTCGGCATACTGGGACATCTGGCACACGTGACGG GGAAGACGGACGTGGGAAATGTGGTGAAATCCGATGCGGG ACTTGCATTTATTCGTACCCGGAAGCGATCGCAAAGTTCG AGGTGCTGCCGCAAGCATTTTCGGTGTCTGTTCTTCTGATGC TGTTTGTGCTCGGCATCGGCAGCAATGTGGCGATGACGTCCG TGCGTGATGACGGTGATAAAGGATCAGTTCGCCGCGGTGCG CAACTGGCAGGC(c/g)GCCACGATATTGCCATCTGTGGCGT GCTGCTCGGTAGCATCTACGTGACGCCG
5	Intron 4-5 ENSANGE00000062609	11091847 11091919	11091918 11092195	72 bp 277 bp	gtaggtgtttcttagcttctta.....atcctttctcctgctcatttcag GGCGGTCAAGT(a/g)TGTGTTGAAG(C/t)TGGTGGATTATTACGG CGCCTCATCGATCGCGCTCGTGCTTGCCATTGCTGAGCTGAT CGCGATCGGGTGGGTTTA(t/c)GGCGTGGATCGACTGTGCAA GGACACGGAGTTTATGCTTGGCCATCGGCCCAACTGTACT GGCGTCTGTGCTGGCGCTGGATCACGCCCTGCTAATGTTT GTCATTCTCATCTACAACCTTGTACGCTCGAGCCGCTCATG TACA(a/g)GCAGTACGTTTATCCTACGGTCGCTTACG
6	Intron 5-6 ENSANGE00000062604	11092196 11092280	11092279 11092374	84 bp 95 bp	gtgagtatgaccattaagagctaa.....actgctgctgttcatccccgag GCATTGGTTGGTGCATTTTTGCCTTTGGGCT(g/t)CTGCAGCT (c/t)CCGATTTGGGC(g/t)GCCACGCCGCTACAAGCAGAGC GGCAAATCGCT(c/g)AACGAG
7	Intron 6-7 ENSANGE00000062602	11092375 11092438	11092437 11092530	63 bp 93 bp	gtacggcaccgaaactagctcgc.....ttaactgtttcattattcttttag AAAATTA(a/g)AACGCCTTCAAACCGACCGCTGCCTGGGGT CCGATTGATCCTGCCACGATTACGAGTACAAAAAGTTCAT CGACGAGGACTGA
	3' downstream sequence				ctttagagggaaaccttaaccgggacggtcagcaggttaattctta.....

SNPs are shown in parentheses (EMBL prediction/vs./cloned *agNAT8*).

Table S2. Putative post-translational modification sites in the agNAT8 protein along with predicted transmembrane domain (TMD) boundaries

Putative motif definition	Start	Stop	Length
NH₂			
1. Proline dependent kinase phosphorylation site	13	16	4
2. Myosin 1 heavy chain kinase phosphorylation site	21	25	5
3. Proline dependent kinase phosphorylation site	24	27	4
4. Casein kinase II phosphorylation site (S or T PO4'D)	41	46	6
5. Myosin 1 heavy chain kinase phosphorylation site	54	58	5
6. Proline dependent kinase phosphorylation site	56	59	4
7. Casein kinase II phosphorylation site (S or T PO4'D)	59	64	6
TMD1	84	111	27_TMD1
TMD2	115	144	29_TMD2
8. Protein kinase C phosphorylation site	142	146	5
9. Glycogen synthase kinase 3 phosphorylation site (last S must be PO4 S)	142	148	7
10. cGMP dependent protein kinase phosphorylation site	143	146	4
TMD3	176	212	36_TMD3
11. Glycogen synthase kinase 3 phosphorylation site (last S must be PO4 S)	209	215	7
12. Proline dependent kinase phosphorylation site	215	218	4
13. Casein kinase II phosphorylation site (S or T PO4'D)	215	220	6
14. Glycogen synthase kinase 3 phosphorylation site (last S must be PO4 S)	215	221	7
1g. Tentative N-glycosylation site	218	221	NDSS
15. Protein kinase C phosphorylation site	221	225	5
2g. Tentative N-glycosylation site	223	226	NRTL
16. Casein kinase II phosphorylation site (S or T PO4'D)	227	232	6
TMD4	252	272	20_TMD4
17. Myosin 1 heavy chain kinase phosphorylation site	254	258	5
18 Protein kinase C phosphorylation site	277	281	5
TMD5	280	300	20_TMD5
19. Glycogen synthase kinase 3 phosphorylation site (last S must be PO4 S)	277	283	7
1c. Tentative chaperonin binding epitope	293	300	8
20. Tyrosine kinase phosphorylation site	314	323	10
TMD6	324	349	25_TMD6
21. Consensus sequence for phosphorylation of mannose	347	352	6
22. Consensus sequence for phosphorylation of mannose	350	355	6
23. Casein kinase II phosphorylation site (S or T PO4'D)	365	370	6
TMD7	358	390	32_TMD7
24. Protein kinase C phosphorylation site	390	394	5
25. Casein kinase II phosphorylation site (S or T PO4'D)	409	414	6
TMD8	422	454	32_TMD8
2c. Tentative chaperonin binding epitope	428	435	8
TMD9	460	481	21_TMD9
26. Proline dependent kinase phosphorylation site	480	483	4
TMD10	486	512	26_TMD10
1n. NADP binding consensus sequence	494	513	20
Tmd11	535	569	34_TMD11
27. Proline dependent kinase phosphorylation site	543	546	4
28. Consensus sequence for phosphorylation of mannose	563	568	6
TMD12	575	604	29_TMD12
29. Consensus sequence for phosphorylation of mannose	593	598	6
30. Protein kinase C phosphorylation site	599	603	5
31. Casein kinase II phosphorylation site (S or T PO4'D)	602	607	6
32. Casein kinase II phosphorylation site (S or T PO4'D)	624	629	6
33. Consensus sequence for phosphorylation of mannose	628	633	6
COOH			