

FIG S1 (A)

ScAAD10 (1) -----  
ScAAD10+1kb (695) GAATAGC-TTCATAACCCACCGCCGCAA<sup>AA</sup>-TGTGCTCTTCGAA<sup>AT</sup>GTCTGAGG-CTTTTGA<sup>CCT</sup>GCA<sup>CCTGAACCA</sup>CCTACCGAGTTAGGACGCTCTA  
ScAAD15 (1) -----  
ScAAD15+1kb (332) TTGAGAATGAGATTTTGACACCATTA<sup>CTAATAATCC</sup>TATTCACT<sup>AATACTTACT</sup>TTAAGC<sup>GGTA</sup>---GAAAA<sup>ACA</sup>CTG<sup>TA</sup>-<sup>CAAGAA</sup>-----  
ScAAD3+1kb (951) TTCA--ATTAC-TCTTTTCGATCATCTAG<sup>AGAAGCA</sup>TATCCA-TCGTG<sup>CTT</sup>AAAA<sup>TGATTC</sup>GGT<sup>CCGCGT</sup><sup>CCGACT</sup>CATCTAG<sup>CAAGTTAGGAC</sup><sup>GCTCC</sup>  
ScAAD3 (1) -----  
ScAAD16 (1) -----  
ScAAD16+1kb (291) CTGAATA-TCTCTTCTATATCATT<sup>TTCA</sup>AT-AGCGA<sup>TTTTAAAT</sup><sup>AT</sup>GGCTGAT<sup>T</sup>-TATTTGCT<sup>CCT</sup>GCT<sup>CCTGAACCA</sup>CTACAGAGTTAGGACGCTCTCA  
ScAAD6+1kb (959) CTGAATA-TCTCTTCTATATCATT<sup>TTCA</sup>AT-AGCGA<sup>TTTTAAAT</sup><sup>AT</sup>GGCTGAT<sup>T</sup>-TATTTGCT<sup>CCT</sup>GCT<sup>CCTGAACCA</sup>CTACAGAGTTAGGACGCTCTCA  
ScAAD6 (1) -----  
ScAAD14 (1) -----  
ScAAD14+11b (959) TACTTAA-CATATTGTCTATTGCATA<sup>AAA</sup>AA-ACCAC<sup>TATTCAGC</sup><sup>AT</sup>GACTGACT<sup>T</sup>-TGTTTAA<sup>ACCT</sup>CTA<sup>CCTGAACCA</sup>CCTACCGA<sup>ATT</sup>GGGACGCTCTCA  
ScAAD4 (1) -----  
ScAAD4+144kb (819) CACTAGAATACACTTCCCTATAGCTACAA<sup>AA</sup>TATTTAT<sup>CAATAAT</sup><sup>AT</sup>TAGTGA<sup>AT</sup>-TATTTAA<sup>ACCT</sup>CTA<sup>CCTGAACCA</sup>CCTACCGAGTTGGGACGCTCTCA  
  
ScAAD10 (1) -----  
ScAAD10+1kb (792) GGGTCTA<sup>TCT</sup>AAAA<sup>CAGC</sup>TGGTAT<sup>AAGAGTTTC</sup>TCCACT<sup>AATCC</sup>TGGGAGG<sup>TAT</sup>GTCTAT<sup>TGGT</sup>GAC<sup>GCCTGGTC</sup>TGGAT<sup>TCAT</sup>TGGAT<sup>CAAT</sup>GACAA  
ScAAD15 (1) -----  
ScAAD15+1kb (417) --T<sup>TGGCT</sup>GCT<sup>GAA</sup>-TGC<sup>TCCAT</sup>GGATTCAA<sup>CAC</sup>AGT<sup>GAAAAA</sup>GGT<sup>TAA</sup>GTGGAA-ACGT--GATAAACTG<sup>CGCCT</sup>GCT<sup>TGCTCG</sup>GATC-<sup>CAA</sup>  
ScAAD3+1kb (1047) GAT<sup>TTCTTTCT</sup>GAAACTGC<sup>CGCTAT</sup>TAAAGTAT<sup>CCCGT</sup>TAAATCC<sup>TAGGAGA</sup>AGTCTCATAC<sup>GAT</sup>GAGC<sup>ACGTT</sup>CGGAT<sup>TTCT</sup>CAAA<sup>TC</sup>AAATGAACAA  
ScAAD3 (47) GAT<sup>TTCTTTCT</sup>GAAACTGC<sup>CGCTAT</sup>TAAAGTAT<sup>CCCGT</sup>TAAATCC<sup>TAGGAGA</sup>AGTCTCATAC<sup>GAT</sup>GAGC<sup>ACGTT</sup>CGGAT<sup>TTCT</sup>CAAA<sup>TC</sup>AAATGAACAA  
ScAAD16 (1) -----  
ScAAD16+1kb (388) GAGTTC<sup>TTTCT</sup>AAAA<sup>TGCTGGTAT</sup>TAGAGTTTC<sup>CCCTCT</sup>CATTT<sup>TGGGAGGA</sup>ATGTC<sup>AAT</sup>TGG<sup>CGA</sup>CGCT<sup>TGGTC</sup>TGAA<sup>AT</sup>CCT<sup>TAGGAT</sup>CAATGAGCAA  
ScAAD6+1kb (1056) GAGTTC<sup>TTTCT</sup>AAAA<sup>TGCTGGTAT</sup>TAGAGTTTC<sup>CCCTCT</sup>CATTT<sup>TGGGAGGA</sup>ATGTC<sup>AAT</sup>TGG<sup>CGA</sup>CGCT<sup>TGGTC</sup>TGAA<sup>AT</sup>CCT<sup>TAGGAT</sup>CAATGAGCAA  
ScAAD6 (56) GAGTTC<sup>TTTCT</sup>AAAA<sup>TGCTGGTAT</sup>TAGAGTTTC<sup>CCCTCT</sup>CATTT<sup>TGGGAGGA</sup>ATGTC<sup>AAT</sup>TGG<sup>CGA</sup>CGCT<sup>TGGTC</sup>TGAA<sup>AT</sup>CCT<sup>TAGGAT</sup>CAATGAGCAA  
ScAAD14 (56) GGGTTC<sup>TTTCT</sup>AAAA<sup>TGCTGGTAT</sup>TAGAGTTTC<sup>CCCTCT</sup>CATTT<sup>TGGGAGGA</sup>GCTTCA<sup>AT</sup>CGG<sup>CGA</sup>CGCAT<sup>TGGTC</sup>AGC<sup>TT</sup>TAT<sup>TGGCT</sup>CTATGAATAA  
ScAAD14+11b (1056) GGGTTC<sup>TTTCT</sup>AAAA<sup>TGCTGGTAT</sup>TAGAGTTTC<sup>CCCTCT</sup>CATTT<sup>TGGGAGGA</sup>GCTTCA<sup>AT</sup>CGG<sup>CGA</sup>CGCAT<sup>TGGTC</sup>AGC<sup>TT</sup>TAT<sup>TGGCT</sup>CTATGAATAA  
ScAAD4 (1) -----  
ScAAD4+1kb (918) GGGTTC<sup>TTTCT</sup>AAAA<sup>TGCTGGTAT</sup>TAGAGTTTC<sup>CCCTCT</sup>CATTT<sup>TGGGAGGA</sup>GCTTCA<sup>AT</sup>CGG<sup>CGA</sup>CGCAT<sup>TGGTC</sup>AGC<sup>TT</sup>TAT<sup>TGGCT</sup>CTATGAATAA  
  
ScAAD10 (1) -----  
ScAAD10+1kb (892) AGAA<sup>CAA</sup>GCTTTTGA<sup>GCT</sup>ACTT<sup>GATGCTTTTTTA</sup>CCA<sup>AGCAGG</sup>CGGAAAT<sup>TCATTGATACTGCAA</sup>TAATTA<sup>TTA</sup>---G<sup>TAT</sup>GA---GCAGTCAGAGACT<sup>TT</sup>  
ScAAD15 (1) -----  
ScAAD15+1kb (510) CAA<sup>AAT</sup>CA<sup>TT</sup>ACG---GTG<sup>TGA</sup>ATA<sup>TGAAG</sup>AAGAA<sup>CAGATTT</sup>ATCT<sup>GCCA</sup>ATGTT<sup>GAGCAAA</sup>C<sup>ACT</sup>CTG<sup>ATGGA</sup>TTCA<sup>AT</sup>TTCT<sup>CTCA</sup>TAGC-T  
ScAAD3+1kb (1147) GAAT<sup>C</sup>GAGCTTTTGA<sup>AT</sup>TGCT<sup>TGAT</sup>ACTTTTT<sup>TACG</sup>AGCAGGT<sup>TGAAAT</sup>TCATT<sup>GATG</sup>CCGCA<sup>AA</sup>CA<sup>ACT</sup>GCCA---AAA<sup>CGA</sup>--GCA<sup>AT</sup>CAGAGAA<sup>AT</sup>  
ScAAD3 (147) GAAT<sup>C</sup>GAGCTTTTGA<sup>AT</sup>TGCT<sup>TGAT</sup>ACTTTTT<sup>TACG</sup>AGCAGGT<sup>TGAAAT</sup>TCATT<sup>GATG</sup>CCGCA<sup>AA</sup>CA<sup>ACT</sup>GCCA---AAA<sup>CGA</sup>--GCA<sup>AT</sup>CAGAGAA<sup>AT</sup>  
ScAAD16 (1) -----  
ScAAD16+1kb (488) GGAG<sup>C</sup>GAGCTTTTGA<sup>AT</sup>TGCT<sup>TGAT</sup>GCTTTTT<sup>TATGA</sup>GCAGGT<sup>TGAAAT</sup>TCATT<sup>GATACTGCAA</sup>TAATTA<sup>TCA</sup>---AAA<sup>CGA</sup>--ACAGTCAGAGCT<sup>TT</sup>

ScAAD6--+1kb (1156) GGAGCGAGCTTTTGAAGTTGCTCGATGCTTTTTATGAGCGAGGTGAAATTTTCATTGATACTGCAAAATAATTATCA---AAACGA--ACAGTCAGAAGCTTT  
ScAAD6 (156) GGAGCGAGCTTTTGAAGTTGCTCGATGCTTTTTATGAGCGAGGTGAAATTTTCATTGATACTGCAAAATAATTATCA---AAACGA--ACAGTCAGAAGCTTT  
ScAAD14 (156) GGAACAGGCTTTTGAAGTTCTTGATGCTTTTTATGAGCTGGAGGTAATTTGTATTGATACTGCAAAACAGTTACCA---AAATGA--AGAGTCAGAGATTT  
ScAAD14--+1lb (1156) GGAACAGGCTTTTGAAGTTCTTGATGCTTTTTATGAGCTGGAGGTAATTTGTATTGATACTGCAAAACAGTTACCA---AAATGA--AGAGTCAGAGATTT  
ScAAD4 (18) GGAACAGGCTTTTGAAGTTCTTGATGCTTTTTATGAGCTGGAGGTAATTTGTATTGATACTGCAAAACAGTTACCA---AAATGA--AGAGTCAGAGATTT  
ScAAD4--+1kb (1018) GGAACAGGCTTTTGAAGTTCTTGATGCTTTTTATGAGCTGGAGGTAATTTGTATTGATACTGCAAAACAGTTACCA---AAATGA--AGAGTCAGAGATTT

ScAAD10 (1) -----ATTCATCAA-GAAAACTGCGTGACCAGATTGTAATTGCCACTAAATTTACCA CGATTATAAGGGTATGATGTAGGCAAGG-GG  
ScAAD10--+1kb (987) GGATCGGTGAATGGATGGCATCAA-GAAAACTGCGTGACCAGATTGTAATTGCCACTAAATTTACCA CGATTATAAGGGTATGATGTAGGCAAGG-GG  
ScAAD15 (1) -----

ScAAD15--+1kb (605) CAATTTCGAACAGCAACCTGTACTTGAAAAGGACTGGCTTAGATATTCCTTACAAGT--GCTTGAAGGTGTCGTAA-----TACTACGCTCTCGATATGG  
ScAAD3--+1kb (1242) GGATTTGGTGAATGGATACAGTCCA-GAAGGTTACGTGATCAAATTTGTCATTGCAACCAAGTTTATAAAAGCGATAAAAGTATAAAGCAGGTGAAA-GT  
ScAAD3 (242) GGATTTGGTGAATGGATACAGTCCA-GAAGGTTACGTGATCAAATTTGTCATTGCAACCAAGTTTATAAAAGCGATAAAAGTATAAAGCAGGTGAAA-GT  
ScAAD16 (1) -----

ScAAD16--+1kb (583) GGATCGGTGAATGGATGGTTCAA-GAAAATTACCGGACCAGATTGTAATCGCCACCAAGTTTACCACTGACTATAAGAAATAACGACGTGGTGGTG-GT  
ScAAD6--+1kb (1251) GGATCGGTGAATGGATGGTTCAA-GAAAATTACCGGACCAGATTGTAATCGCCACCAAGTTTACCACTGACTATAAGAAATAACGACGTGGTGGTG-GT  
ScAAD6 (251) GGATCGGTGAATGGATGGTTCAA-GAAAATTACCGGACCAGATTGTAATCGCCACCAAGTTTACCACTGACTATAAGAAATAACGACGTGGTGGTG-GT  
ScAAD14 (251) GGATAGGTGAATGGATGGCATCAA-GAAAACTGCGTGACCAGATTGTAATTGCCACCAAGTTTACCGGAGATTATAAGAAGTATGAAGTAGGTGGTG-GT  
ScAAD14--+1lb (1251) GGATAGGTGAATGGATGGCATCAA-GAAAACTGCGTGACCAGATTGTAATTGCCACCAAGTTTACCGGAGATTATAAGAAGTATGAAGTAGGTGGTG-GT  
ScAAD4 (113) GGATAGGTGAATGGATGAAAICAA-GAAAATTGCGTGACCAAATTGTAATTGCCACCAAGTTTACCGGAGATTATAAGAAGTATGAAGTAGGTGGTG-GT  
ScAAD4--+1kb (1113) GGATAGGTGAATGGATGAAAICAA-GAAAATTGCGTGACCAAATTGTAATTGCCACCAAGTTTACCGGAGATTATAAGAAGTATGAAGTAGGTGGTG-GT

ScAAD10 (85) AAGAGTGCCAAATTTCTGTGGGAATCACAAGCCAGTTTGCATGTAAGTGTGAGAGATTCCCTTCGTAAAGTTG-----CAAACGTG--ATTGGATTGATAT  
ScAAD10--+1kb (1085) AAGAGTGCCAAATTTCTGTGGGAATCACAAGCCAGTTTGCATGTAAGTGTGAGAGATTCCCTTCGTAAAGTTG-----CAAACGTG--ATTGGATTGATAT  
ScAAD15 (1) -----

ScAAD15--+1kb (699) AACGCCGAAGA-TATGCTAGTTATAGTAGGATAGTTC-ACATTTCTGATTATCCTTTTGCTGCGCATAGTCTGGCGCAACATGACCACTATATCAGGGC  
ScAAD3--+1kb (1340) AACACTGCCAACTACTGTGGTAATCACAAGCGTAGTTTACATGTGAGTGTGAGGGATTCTCTCCGCAAA--TT--GC--AAACTG--ATTGGATTGATAT  
ScAAD3 (340) AACACTGCCAACTACTGTGGTAATCACAAGCGTAGTTTACATGTGAGTGTGAGGGATTCTCTCCGCAAA--TT--GC--AAACTG--ATTGGATTGATAT  
ScAAD16 (1) -----

ScAAD16--+1kb (681) AAAAGCGCAAACTACTGTGGCAATCACAAGCGTAGTTTGCATGTGAGTGTGAGGGATTCTCTCCGCAAAATG-----CAAACGTG--ATTGGATTGATAT  
ScAAD6--+1kb (1349) AAAAGCGCAAACTACTGTGGCAATCACAAGCGTAGTTTGCATGTGAGTGTGAGGGATTCTCTCCGCAAAATG-----CAAACGTG--ATTGGATTGATAT  
ScAAD6 (349) AAAAGCGCAAACTACTGTGGCAATCACAAGCGTAGTTTGCATGTGAGTGTGAGGGATTCTCTCCGCAAAATG-----CAAACGTG--ATTGGATTGATAT  
ScAAD14 (349) AAAAGTGCCAACACTACTGTGGTAATCACAAGCGTAGTTTACATGTGAGTGTGAGGGATTCTCTCCGCAAAATG-----CAAACGTG--ATTGGATTGATAT  
ScAAD14--+1lb (1349) AAAAGTGCCAACACTACTGTGGTAATCACAAGCGTAGTTTACATGTGAGTGTGAGGGATTCTCTCCGCAAAATG-----CAAACGTG--ATTGGATTGATAT  
ScAAD4 (211) AAAAGTGCCAACACTATTTGTGGTAATCACAAGCATAGTTTACATGTGAGTGTGAGGGATTCTCTCCGCAAAATG-----CAAACGTG--ATTGGATTGATAT  
ScAAD4--+1kb (1211) AAAAGTGCCAACACTATTTGTGGTAATCACAAGCATAGTTTACATGTGAGTGTGAGGGATTCTCTCCGCAAAATG-----CAAACGTG--ATTGGATTGATAT

ScAAD10 (177) TCTTTACGTTCACTGGT--GGGATTATATGAGCTCCATTGAGGAAGTTATGGATAGTTTGCACATTCTTGTGCAG-CAGGGCAAAGTACTCTATCTTAGGT  
ScAAD10--+1kb (1177) TCTTTACGTTCACTGGT--GGGATTATATGAGCTCCATTGAGGAAGTTATGGATAGTTTGCACATTCTTGTGCAG-CAGGGCAAAGTACTCTATCTTAGGT  
ScAAD15 (1) -----

ScAAD15--+1kb (797) ACTTTGTCCTTCGTTGACGGCATGGCGCTTGTGACTTTAGCAATT-ATGCTAA--TGATAT-TCTTGTGAGACA-CATCAAGAAG-TATACTAGGAC

ScAAD3--+1kb	(1432)	ACTTTACGTTCACTGGT--GGGATTATATGAGTTCAATCGAAGAA	TTTATGGATAGTTTGCATATTCT	GGTCCAG--CAGGGCAAGGTCCCTCTATTTGGGT
ScAAD3	(432)	ACTTTACGTTCACTGGT--GGGATTATATGAGTTCAATCGAAGAA	TTTATGGATAGTTTGCATATTCT	GGTCCAG--CAGGGCAAGGTCCCTCTATTTGGGT
ScAAD16	(1)	-----		
ScAAD16--+1kb	(773)	ACTTTACGTTCACTGGT--GGGATTATATGAGTTCTATCGAAGAAGTTATGGATAGTTT	ACACATTCTAGTGCAG--CAGG--CAAGGTCCCTCTATTTGGGT	
ScAAD6--+1kb	(1441)	ACTTTACGTTCACTGGT--GGGATTATATGAGTTCTATCGAAGAAGTTATGGATAGTTT	ACACATTCTAGTGCAG--CAGG--CAAGGTCCCTCTATTTGGGT	
ScAAD6	(441)	ACTTTACGTTCACTGGT--GGGATTATATGAGTTCTATCGAAGAAGTTATGGATAGTTT	ACACATTCTAGTGCAG--CAGG--CAAGGTCCCTCTATTTGGGT	
ScAAD14	(441)	ACTTTACATTCACTGGT--GGGATTATATGAGTTCAATCGAAGAAGTTATGGATAGTTT	GCATATTTTAGTTTCAG--CAGGGCAAGGTCCCTATAITTTAGGA	
ScAAD14--+1lb	(1441)	ACTTTACATTCACTGGT--GGGATTATATGAGTTCAATCGAAGAAGTTATGGATAGTTT	GCATATTTTAGTTTCAG--CAGGGCAAGGTCCCTATAITTTAGGA	
ScAAD4	(303)	ACTTTACGTTCACTGGT--GGGATTATATGAGTTCAATCGAAGAAGTTATGGATAGTTT	GCATATTTTAGTTTCAG--CAGGGCAAAGTCCCTCTATTTGGGT	
ScAAD4--+1kb	(1303)	ACTTTACGTTCACTGGT--GGGATTATATGAGTTCAATCGAAGAAGTTATGGATAGTTT	GCATATTTTAGTTTCAG--CAGGGCAAAGTCCCTCTATTTGGGT	
ScAAD10	(274)	GTGCTGATACCTCTGCCTGGGTTGTTTCTGCAGCAAATTA	CTACGCTACATCTCATGGTAAAACCTCCCTTTAGTATCTATCAAGGTA	AAATGGAACTGTAT
ScAAD10--+1kb	(1274)	GTGCTGATACCTCTGCCTGGGTTGTTTCTGCAGCAAATTA	CTACGCTACATCTCATGGTAAAACCTCCCTTTAGTATCTATCAAGGTA	AAATGGAACTGTAT
ScAAD15	(1)	-----		
ScAAD15--+1kb	(892)	CATTGTCATTCATAGAAAAGGCCATTTTTCACCGGATAGAAC	-----ATAGTAAAG---CAATTTCCAGTTCCAGCC-----TGGAACTGT	
ScAAD3--+1kb	(1529)	GTATCTGATACACCTGCTTGGGTTGTTTCTGCGGCAAACT	TACTACGCTACATCTTATGGTAAAACCTCCCTTTAGTATCTA	CAAGGTAATGGAACTGT
ScAAD3	(529)	GTATCTGATACACCTGCTTGGGTTGTTTCTGCGGCAAACT	TACTACGCTACATCTTATGGTAAAACCTCCCTTTAGTATCTA	CAAGGTAATGGAACTGT
ScAAD16	(1)	-----		
ScAAD16--+1kb	(869)	GTGCTGATACGCTCTGCCTGGGTTGTTTCTGCGGCAAAAT	TACTACGCTAAATCTCATGGTAAAACCTCTTTAGCATCTATCAAGGTA	AAATGGAACTGT
ScAAD6--+1kb	(1537)	GTGCTGATACGCTCTGCCTGGGTTGTTTCTGCGGCAAAAT	TACTACGCTAAATCTCATGGTAAAACCTCTTTAGCATCTATCAAGGTA	AAATGGAACTGT
ScAAD6	(537)	GTGCTGATACGCTCTGCCTGGGTTGTTTCTGCGGCAAAAT	TACTACGCTAAATCTCATGGTAAAACCTCTTTAGCATCTATCAAGGTA	AAATGGAACTGT
ScAAD14	(538)	GTATCTGATACACCTGCTTGGGTTGTTTCTGCGGCAAAAT	TACTACGCTACATCTCATGGTAAAACCTCTTTAGCATCTATCAAGGTA	AAATGGAACTGTAT
ScAAD14--+1lb	(1538)	GTATCTGATACACCTGCTTGGGTTGTTTCTGCGGCAAAAT	TACTACGCTACATCTCATGGTAAAACCTCTTTAGCATCTATCAAGGTA	AAATGGAACTGTAT
ScAAD4	(400)	GTGCTGATACACCTGCTTGGGTTGTTTCTGCGGCAAACT	TACTACGCTACATCTCATGGTAAAACCTCTTTAGTATCTATCAAGGTA	AAATGGAACTGT
ScAAD4--+1kb	(1400)	GTGCTGATACACCTGCTTGGGTTGTTTCTGCGGCAAACT	TACTACGCTACATCTCATGGTAAAACCTCTTTAGTATCTATCAAGGTA	AAATGGAACTGT
ScAAD10	(374)	TGAACAGGACTTTGAAACGTGATATCATTCCAATGGCTAGGCATTT	TGGTATGGCTCTTGCCTCATGGGATGTTATGGGAGGCGGGAGATTT	CAGAGTAA
ScAAD10--+1kb	(1374)	TGAACAGGACTTTGAAACGTGATATCATTCCAATGGCTAGGCATTT	TGGTATGGCTCTTGCCTCATGGGATGTTATGGGAGGCGGGAGATTT	CAGAGTAA
ScAAD15	(1)	-----		
ScAAD15--+1kb	(969)	TGAACAGAGATTTTGGAGCGTGATATATTCCAATGGCTAGGCATTT	CGGTATGGCCCTCGCCCCATGGGATGTCATGGGAGGTGGAAGATTT	CAGAGTAA
ScAAD3--+1kb	(1629)	TGAACAGAGATTTTGGAGCGTGATATATTCCAATGGCTAGGCATTT	CGGTATGGCCCTCGCCCCATGGGATGTCATGGGAGGTGGAAGATTT	CAGAGTAA
ScAAD3	(629)	TGAACAGAGATTTTGGAGCGTGATATATTCCAATGGCTAGGCATTT	CGGTATGGCCCTCGCCCCATGGGATGTCATGGGAGGTGGAAGATTT	CAGAGTAA
ScAAD16	(1)	-----		
ScAAD16--+1kb	(969)	TGAACAGAGACTTCGAACGTGATATCATTCCAATGGCTAGGCATTT	CGGTATGGCCCTCGCCCCATGGGATGTCATGGGAGGTGGAAGATTT	CAGAGTAA
ScAAD6--+1kb	(1637)	TGAACAGAGACTTCGAACGTGATATCATTCCAATGGCTAGGCATTT	CGGTATGGCCCTCGCCCCATGGGATGTCATGGGAGGTGGAAGATTT	CAGAGTAA
ScAAD6	(637)	TGA	-----	
ScAAD14	(638)	TGAACAGGACTTTGAGCGTGATATATTCCAATGGCTAGGCATTT	TGGTATGGCTCTAGCCCCATGGGATGTCATGGGAGGTGGAAGATTT	CAGAGTAA
ScAAD14--+1lb	(1638)	TGAACAGGACTTTGAGCGTGATATATTCCAATGGCTAGGCATTT	TGGTATGGCTCTAGCCCCATGGGATGTCATGGGAGGTGGAAGATTT	CAGAGTAA
ScAAD4	(500)	TGAACAGGACTTTGAGCGTGATATATTCCAATGGCTAGGCATTT	TGGTATGGCTCTAGCCCCATGGGATGTTATGGGAGGTGGAAGATTT	CAGAGTAA
ScAAD4--+1kb	(1500)	TGAACAGGACTTTGAGCGTGATATATTCCAATGGCTAGGCATTT	TGGTATGGCTCTAGCCCCATGGGATGTTATGGGAGGTGGAAGATTT	CAGAGTAA

ScAAD10 (474) AAAAGGCAGTGGAGAGCGGAAGAAGAAAGGAGAAGGCTTGGTACCTTTTITGGTACTTCGGAACAGACGGATATGGAAGTTAAAATCAGCGAAGCATTG  
ScAAD10--1kb (1474) AAAAGGCAGTGGAGAGCGGAAGAAGAAAGGAGAAGGCTTGGTACCTTTTITGGTACTTCGGAACAGACGGATATGGAAGTTAAAATCAGCGAAGCATTG  
ScAAD15 (69) AAAAGCAATGGAGGAACCGGAGGAAGAATGGAGAGTGTATTTCGTTCTTTCGTTGGCGCCTCCGAACAAACAGATGCAGAAATCAAGATTAGTGAAGCATTG  
ScAAD15--1kb (1069) AAAAGCAATGGAGGAACCGGAGGAAGAATGGAGAGTGTATTTCGTTCTTTCGTTGGCGCCTCCGAACAAACAGATGCAGAAATCAAGATTAGTGAAGCATTG  
ScAAD3--1kb (1729) AAAAGCAATGGAGGAACCGGAGGAAGAATGGAGAGGGTATTTCGTTCTTTCGTTGGCGCCTCCGAACAAACAGATGCAGAAATCAAGATTAGTGAAGCATTG  
ScAAD3 (729) AAAAGCAATGGAGGAACCGGAGGAAGAATGGAGAGGGTATTTCGTTCTTTCGTTGGCGCCTCCGAACAAACAGATGCAGAAATCAAGATTAGTGAAGCATTG  
ScAAD16 (69) AAAAGCAATGGAGGAACCGGAGGAAGAATGGAGAGGGTATTTCGTTCTTTCGTTGGCGCCTCCGAACAAACAGATGCAGAAATCAAGATTAGTGAAGCATTG  
ScAAD16--1kb (1069) AAAAGCAATGGAGGAACCGGAGGAAGAATGGAGAGGGTATTTCGTTCTTTCGTTGGCGCCTCCGAACAAACAGATGCAGAAATCAAGATTAGTGAAGCATTG  
ScAAD6--1kb (1737) AAAAGCAATGGAGGAACCGGAGGAAGAATGGAGAGGGTATTTCGTTCTTTCGTTGGCGCCTCCGAACAAACAGATGCAGAAATCAAGATTAGTGAAGCATTG  
ScAAD6 (640) -----  
ScAAD14 (738) AAAAGCAATGGAGGAACCGGAAGAAGAATGGAGAGGGTCTGGTACTTTTGTGGGTGGCCCGAACAAACAGAAATTGGAAGTTAAAATCAGCGAAGCATTG  
ScAAD14--1kb (1738) AAAAGCAATGGAGGAACCGGAAGAAGAATGGAGAGGGTCTGGTACTTTTGTGGGTGGCCCGAACAAACAGAAATTGGAAGTTAAAATCAGCGAAGCATTG  
ScAAD4 (600) AAAAGCAATGGAGGAACCGGAAGAAGAATGGAGAGGGTCTGGTACTGTTTCGGGTACTTCTAACAAGACGGATAAAGAGGTTAAGATCAGTGAAGCATTG  
ScAAD4--1kb (1600) AAAAGCAATGGAGGAACCGGAAGAAGAATGGAGAGGGTCTGGTACTGTTTCGGGTACTTCTAACAAGACGGATAAAGAGGTTAAGATCAGTGAAGCATTG

ScAAD10 (574) TTAAGAAGTTGCGGAAGAACATGGCACTGAGTCTGTCACTGCTATTGCGATAGCTTATGTTCCGCTCTAAAGCGAAAACATGTTTTCCATTAGTGGGAGGAA  
ScAAD10--1kb (1574) TTAAGAAGTTGCGGAAGAACATGGCACTGAGTCTGTCACTGCTATTGCGATAGCTTATGTTCCGCTCTAAAGCGAAAACATGTTTTCCATTAGTGGGAGGAA  
ScAAD15 (169) GCCAAGGTTGCTGAGGAACATGGCACTGAGTCTGTACTGCTATTGCTATTGCCTATGTTCCGCTCTAAGGCGAAAAATGTTTTCCGTCGGTTGAGGAG  
ScAAD15--1kb (1169) GCCAAGGTTGCTGAGGAACATGGCACTGAGTCTGTACTGCTATTGCTATTGCCTATGTTCCGCTCTAAGGCGAAAAATGTTTTCCGTCGGTTGAGGAG  
ScAAD3--1kb (1829) GCCAAGATTGCTGAGGAACATGGCACTGAGTCTGTACTGCTATTGCTATTGCCTATGTTCCGCTCTAAGGCGAAAAATTTTTTCCGTCGGTTGAGGAG  
ScAAD3 (829) GCCAAGATTGCTGAGGAACATGGCACTGAGTCTGTACTGCTATTGCTATTGCCTATGTTCCGCTCTAAGGCGAAAAATTTTTTCCGTCGGTTGAGGAG  
ScAAD16 (169) GCCAAGGTTGCTGAGGAACATGGCACTGAATCTGTTACTGCTATTGCTATTGCCTATGTTCCGCTCTAAGGCGAAAAATGTTTTCCATTGGTTGGAGGAA  
ScAAD16--1kb (1169) GCCAAGGTTGCTGAGGAACATGGCACTGAATCTGTTACTGCTATTGCTATTGCCTATGTTCCGCTCTAAGGCGAAAAATGTTTTCCATTGGTTGGAGGAA  
ScAAD6--1kb (1837) GCCAAGGTTGCTGAGGAACATGGCACTGAATCTGTTACTGCTATTGCTATTGCCTATGTTCCGCTCTAAGGCGAAAAATGTTTTCCATTGGTTGGAGGAA  
ScAAD6 (640) -----  
ScAAD14 (838) ACTAAAATTGCTGAGGAACATGGACAAGAGTCTGTTACTGCTATCGCTATTGCCTATGTTCCGCTCTAAAGCGAAAAATGTTTTCCATTGATTGGAGGAA  
ScAAD14--1kb (1838) ACTAAAATTGCTGAGGAACATGGACAAGAGTCTGTTACTGCTATCGCTATTGCCTATGTTCCGCTCTAAAGCGAAAAATGTTTTCCATTGATTGGAGGAA  
ScAAD4 (700) GCCAAGGTTGCTGAGGAACATGGCACTGAGTCTGTTACTGCTATTGCTATTGCCTATGTTCCGCTCTAAGGCGAAAAATGTTTTCCATTGGTTGGTGGAA  
ScAAD4--1kb (1700) GCCAAGGTTGCTGAGGAACATGGCACTGAGTCTGTTACTGCTATTGCTATTGCCTATGTTCCGCTCTAAGGCGAAAAATGTTTTCCATTGGTTGGTGGAA

ScAAD10 (674) GAAAGATCGAACATCTCAAACAGAACATTGAGGCTTTGAGCATTAAATTAACACCAGAACAATAAAGTACTTAGAAAGTATTGTTCCCTTTTGATGTCGG  
ScAAD10--1kb (1674) GAAAGATCGAACATCTCAAACAGAACATTGAGGCTTTGAGCATTAAATTAACACCAGAACAATAAAGTACTTAGAAAGTATTGTTCCCTTTTGATGTCGG  
ScAAD15 (269) GAAAAATTGAGGATCTCAAAGAGAACATTAAGGCTCTCAGTATCGATCTAACGCCGACAAATATAAAATACTTGAAAAATGTAAGTTCCTTTTGATGTCGG  
ScAAD15--1kb (1269) GAAAAATTGAGGATCTCAAAGAGAACATTAAGGCTCTCAGTATCGATCTAACGCCGACAAATATAAAATACTTAGAAAGTATTGTTCCCTTTTGATGTCGG  
ScAAD3--1kb (1929) GAAAAATTGAGGATCTCAAAGAGAACATTAAGGCTCTCAGTATCGATCTAACGCCGACAAATATAAAATACTTAGAAAGTATTGTTCCCTTTTGATGTCGG  
ScAAD3 (929) GAAAAATTGAGGATCTCAAAGAGAACATTAAGGCTCTCAGTATCGATCTAACGCCGACAAATATAAAATACTTAGAAAGTATTGTTCCCTTTTGATGTCGG  
ScAAD16 (269) GGAAAAATTGAACACCTCAAACAGAACATTGAGGCTTTGAGCATTAAATTAACACCAGAACAATAAAGTACTTAGAAAGTATTGTTCCCTTTTGATGTTGG  
ScAAD16--1kb (1269) GGAAAAATTGAACACCTCAAACAGAACATTGAGGCTTTGAGCATTAAATTAACACCAGAACAATAAAGTACTTAGAAAGTATTGTTCCCTTTTGATGTTGG  
ScAAD6--1kb (1937) GGAAAAATTGAACACCTCAAACAGAACATTGAGGCTTTGAGCATTAAATTAACACCAGAACAATAAAGTACTTAGAAAGTATTGTTCCCTTTTGATGTTGG  
ScAAD6 (640) -----  
ScAAD14 (938) GGAAAAATTGAACATCTCAAAGCAGAACATTGAGGCTTTGAGTATTAAATTAACACCAGAACAATAAAGTACTTAGAAAGTATTGTTCCCTTTTGATGTTGG

```

ScAAD14-+11b (1938) GGAAAATTGAACATCTCAAAGCAGAACATTGAGGCTTTGAGTATTAAATTAACACCGGAACAAATAGAATACCTGGAAAGTATTGTCCTTTTGAITGTTGG
ScAAD4 (800) GGAAAATTGAACACCTCAAACAGAACATTGAGGCTTTAAGTATCAAACTGACACCAGAACA GATA GAATACTTAGAAAGTATTATTCCTTTTGAITGTTGG
ScAAD4-+1kb (1800) GGAAAATTGAACACCTCAAACAGAACATTGAGGCTTTAAGTATCAAACTGACACCAGAACA GATA GAATACTTAGAAAGTATTATTCCTTTTGAITGTTGG

ScAAD10 (774) ATTTTCCACTAATTTTTATTG-GAGATGACCCAGCTGTTACCAAGAAACCTTCATTTTCTCACCGAAATGTCTGCAAGATTAGCTTTCGAAGATTTCG-----
ScAAD10-+1kb (1774) ATTTTCCACTAATTTTTATTG-GAGATGACCCAGCTGTTACCAAGAAACCTTCATTTTCTCACCGAAATGTCTGCAAGATTAGCTTTCGAAGATTTCAGTCAAAA
ScAAD15 (369) ATTTTCTAACACTTTTTATCGTGTAAATTCCTTGACTCAAAAATATGGTACGAATAATGTTTAC-----
ScAAD15-+1kb (1369) ATTTTCTAACACTTTTTATCGTGTAAATTCCTTGACTCAAAAATATGGTACGAATAATGTTTAGATAAATTTTTCAGTAAATCAACTACGCAAGTTAAGCAG
ScAAD3-+1kb (2029) ATTTTCTAATAATTTTTATCGTGTAAATTCCTTGACTCAAAAATATGGTACGAATAATGTTTAGATAAATTTTTCAGTAAATCAACTACGCAAGTTAAGCAG
ScAAD3 (1029) ATTTTCTAATAATTTTTATCGTGTAAATTCCTTGACTCAAAAATATGGTACGAATAATGTTTAC-----
ScAAD16 (369) TTTTCCTACTAATTTTTATCG-GTGATGATCCGGCTGTTACCAAGAAGGCTTCACCTTCTCACGGCAATGTCTGCGCAGATTTCCTTCGATTATA-----
ScAAD16-+1kb (1369) TTTTCCTACTAATTTTTATCG-GTGATGATCCGGCTGTTACCAAGAAGGCTTCACCTTCTCACGGCAATGTCTGCGCAGATTTCCTTCGATTATAAGCAGATTT
ScAAD6-+1kb (2037) TTTTCCTACTAATTTTTATCG-GTGATGATCCGGCTGTTACCAAGAAGGCTTCACCTTCTCACGGCAATGTCTGCGCAGATTTCCTTCGATTATAAGCAGATTT
ScAAD6 (640) -----

ScAAD14 (1038) CTTTTCCAAAAGTTTAAATAG-GAGATGACCCAGCGGTAAACCAAGAAGCTTTCACCCCTCACATCGATGTCTGCCAGGATAGCTTTTGACAATTCG-----
ScAAD14-+11b (2038) CTTTTCCAAAAGTTTAAATAG-GAGATGACCCAGCGGTAAACCAAGAAGCTTTCACCCCTCACATCGATGTCTGCCAGGATAGCTTTTGACAATTCAGTTAAA
ScAAD4 (900) TTTTCCTACTAATTTTTATCG-GTGATGATCCGGCTGTTACCAAGAAGGCTTCACCTTCTCACGGCAATGTCTGCGCAGATTTCCTTCGA----TCA-----
ScAAD4-+1kb (1900) TTTTCCTACTAATTTTTATCG-GTGATGATCCGGCTGTTACCAAGAAGGCTTCACCTTCTCACGGCAATGTCTGCGCAGATTTCCTTCGA----TCAATCAGA

```

**FIG S1 (A).** Nucleotide sequence alignment of the putative *AAD* ORFs in *S. cerevisiae* lab strain S288C. Alignment shows that several out-of-frame sequences share still high homology with other *AAD*s. SGD-defined start codons are shown in green and stop codons are in red. Hypothetical 'ancestral' start and stop codons are shown in black boxes. Consensus sequences are highlighted in blue.

FIG S1 (B)

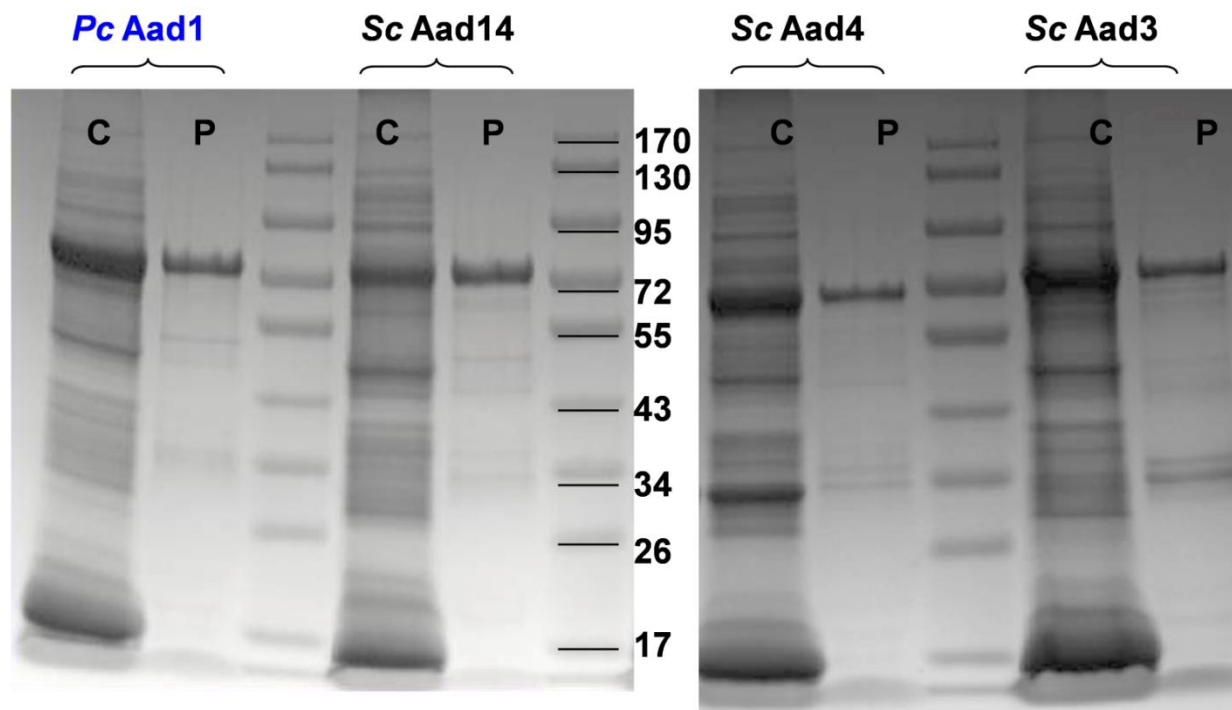
PcAad1p	(1)	-MNWAPAPEPPTKLGRRHQLAPGCGLHVSPIQLGAMSIGDKWHPYGMGTMDEKASFKLLDAFYNAGGNFIDTANVYQDE
ScAad15p	(1)	-----
ScAad6p	(1)	MADLFAPAPEPSTEELGRLRVLSKSAQIRVSPILLGMSIGDAWSE-ILGSMSEKRAFELLDIFYEAGGNFIDTANNYQNE
ScAad16p	(1)	-----
ScAad3p	(1)	---MIGSASDSSSKLGRRLRFLSETAATKVSPLILGEVSYDGARSD-FLKSMNKNRAFELLDIFYEAGGNFIDAANNQNE
ScAad10p	(1)	-----
ScAad14p	(1)	MTDLFKPLPEPSTEELGRLRVLSKTAGIRVSPILLGASIGDAWSG-FMGSMNKEQAFELLDIFYEAGGNCIDTANSYQNE
ScAad4p	(1)	-----
PcAad1p	(80)	TSEEFIGEWMEARGNRDQMVVATKYSLVYKRGASFEEIPQKTOYVGNLSLKSMMHSVHDSLRLKIRTSYIDIFYVHFWDYTC
ScAad15p	(1)	-----
ScAad6p	(80)	QSEAWIGEWVSRKLRDQIVIATKFTTDYKKYDVG--GKSANYCGNHKRSLSVSRDSLRLKQTDWIDILYVHWWDYMS
ScAad16p	(1)	-----
ScAad3p	(77)	QSEWIGEWIQSRRLRDQIVIATKFIKSDKKYKAGE--SNITANYCGNHKRSLSVSRDSLRLKQTDWIDILYVHWWDYMS
ScAad10p	(1)	-----
ScAad14p	(80)	ESEIWIGEWASRKLRDQIVIATKFTTDYKGYDVGK--GKSANFCGNHKRSLSVSRDSLRLKQTDWIDILYVHWWDYMS
ScAad4p	(34)	ESEIWIGEWKSRKLRDQIVIATKFTTDYKKYEVGG--GKSANYCGNHKHSLSVSRDSLRLKQTDWIDILYVHWWDYMS
PcAad1p	(160)	TIEEVMNGLHNLVAQGKVLYLGVSDTPAWVVSKANNYARMAAGKTPFVIYEGEWNTIMRDMERDIIPMCIHEGMAIAPWNV
ScAad15p	(1)	-----
ScAad6p	(158)	SIEEVMDSLHILVQQARSSIWVCLIRLPGILFLRQITTLNLMVQPLLASIKVNGTC-----
ScAad16p	(1)	-----
ScAad3p	(155)	SIEEFMDSLHILVQQGKVLYLGVSDTPAWVVSAAANYATSYGKTPFSIYQGKWNVLRNDFERDIIPMARHFGMALAPWDV
ScAad10p	(70)	SIEEVMDSLHILVQQGKVLYLGVSDTPAWVVSAAANYATSYGKTPFSIYQGKWNVLRNDFERDIIPMARHFGMALAPWDV
ScAad14p	(158)	SIEEVMDSLHILVQQGKVLYLGVSDTPAWVVSAAANYATSYGKTPFSIYQGKWNVLRNDFERDIIPMARHFGMALAPWDV
ScAad4p	(112)	SIEEVMDSLHILVQQGKVLYLGVSDTPAWVVSAAANYATSYGKTPFSIYQGKWNVLRNDFERDIIPMARHFGMALAPWDV
Consensus	(161)	SIEEVMDSLHILVQQGKVLYLGVSDTPAWVVSAAANYATSYGKTPFSIYQGKWNVLRNDFERDIIPMARHFGMALAPWDV
PcAad1p	(240)	LCAGKIRTDAAEEERRLKSGEGGRTILQFDGWLNRNETERKVSKALEKVAEEIGAKSITSVAIAYIMOKFPYVFPFVGGRRK
ScAad15p	(15)	MGGGRFQSKKAMEERRKNGECIRSFVG--ASEQTDAEIKISEALAKVAEEHGTESVTAIAIAYVRSKAKNVFPLVGGKI
ScAad6p	(213)	-----
ScAad16p	(15)	MGGGRFQSKKAMEERRKNGEGIRSFVG--ASEQTDAEIKISEALAKVAEEHGTESVTAIAIAYVRSKAKNVFPLVGGRKI
ScAad3p	(235)	MGGGRFQSKKAMEERRKNGEGIRSFVG--ASEQTDAEIKISEALAKVAEEHGTESVTAIAIAYVRSKAKNFVFPVGGKI
ScAad10p	(150)	MGGGRFQSKKAVERRKKGEGIRTFVG--TSEQTDMVKISEALKVAEEHGTESVTAIAIAYVRSKAKHVFPLVGGRKI
ScAad14p	(238)	MGGGRFQSKKAMEERRKNGEGIRTFVG--GPEQTELEVKISEALTKVAEEHGTESVTAIAIAYVRSKAKNVFPLVGGRKI
ScAad4p	(192)	MGGGRFQSKKAMEERRKNGEGIRTVSG--TSKQTDKEVKISEALAKVAEEHGTESVTAIAIAYVRSKAKNVFPLVGGRKI
Consensus	(241)	MGGGRFQSKKAMEERRKNGEGIRTFVG ASEQTD EIKISEALAKVAEEHGTESVTAIAIAYVRSKAKNVFPLVGGRKI
PcAad1p	(320)	EHLNANLEALDITLSLPEQMQLNDTVPFNKGFPLYLLFGDGSYDNLVHKAAGHYDKWPAQQAIRPQK

```

ScAad15p (93) EDLKENIKALSIDLTPDNIKYLENVVPFDIGFPNTFIVLNS--LTQKYGTNNV-----
ScAad6p (213) -----
ScAad16p (93) EHLKQNEALSIKLTPEQIKYLESIIPFDVGFPTNFIGDDP--AVTKKASL-LTAMSAQISFD--
ScAad3p (313) EDLKENIKALSIDLTPDNIKYLESIVPFDIGFPNNTFIVLNS--LTQKYGTNNV-----
ScAad10p (228) EHLKQNEALSIKLTPEQIKYLESIVPFDVGFPTNFIGDDP--AVTKKPSF-LTEMSAKISFED--
ScAad14p (316) EHLKQNEALSIKLTPEQIEYLESIVPFDVGFPTNFIGDDP--AVTKKLSL-LTSMASRIAFDN--
ScAad4p (270) EHLKQNEALSIKLTPEQIEYLESIIPFDVGFPTNFIGDDP--AVTKKASL-LTAMSAQISFD--

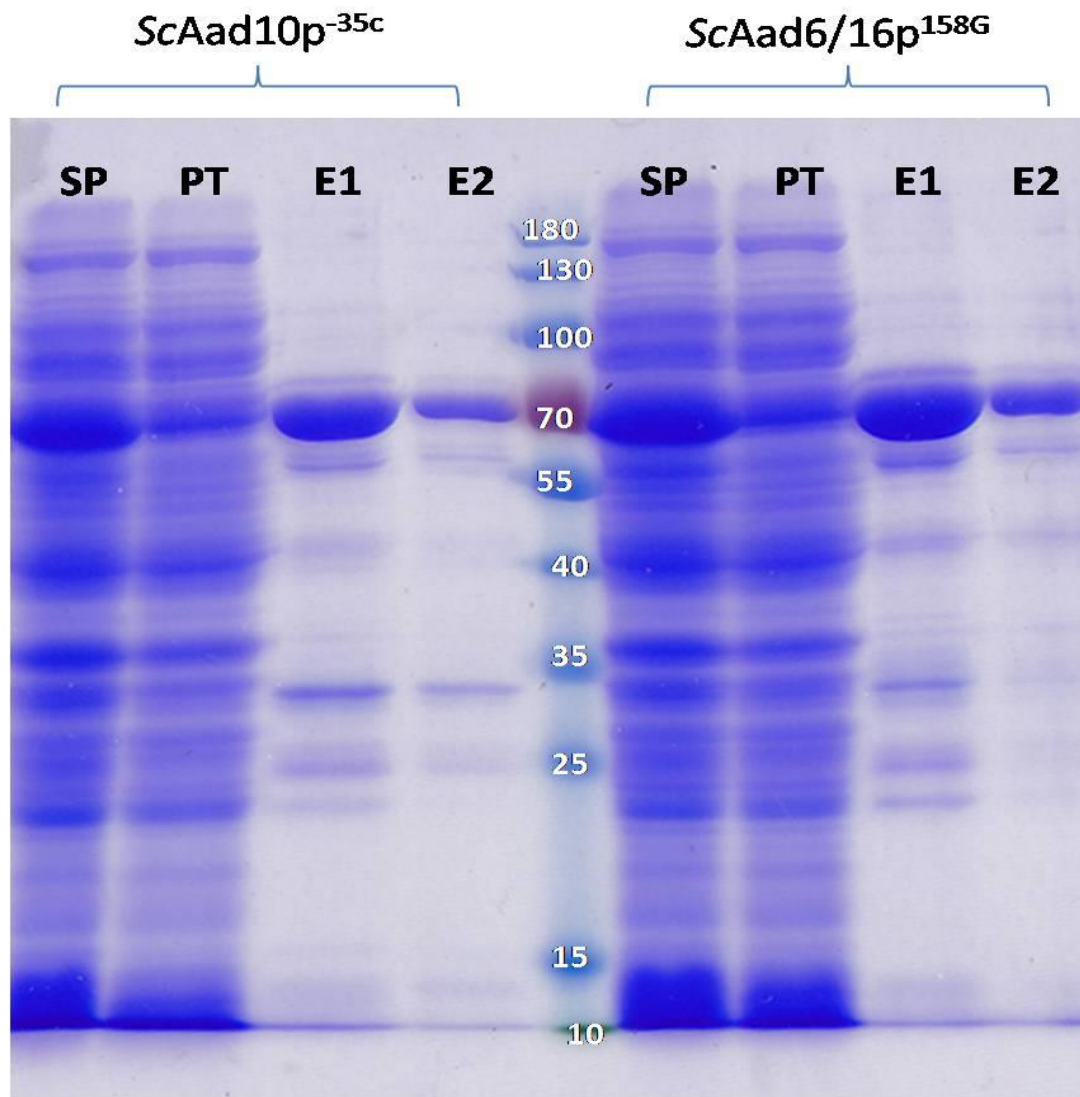
```

**FIG S1 (B).** Amino acid sequence alignment between *PcAad1p* (*Phanerochaete chrysosporium* Aad1p from strain BKM-F-1767) and putative *ScAadp* of S288C, as assigned in the current SGD. Identical residues are highlighted in blue; consensus sequences are highlighted in green.



**FIG S2.** Heterologous expression and GST-affinity batch purifications of ScAad3p, ScAad4p, ScAad14p, and PcAad1p. “C” denotes crude extract, “P” indicates for the first protein elution following purification. Proteins were separated by SDS-PAGE, and stained with Coomassie Brilliant Blue G-250.





**FIG S3.** Heterologous expression and GST-affinity batch purification of reconstructed *ScAad10*<sup>35C</sup>p and *ScAad6/16*<sup>158G</sup>p. “**SP**” indicates supernatant of crude extract, “**PT**” indicates the 1<sup>st</sup> bed volume (2 mL) pass-through elution, “**E1**” stands for 1<sup>st</sup> bed volume elution containing 10 mM reduced glutathione, “**E2**” stands for 2<sup>nd</sup> bed volume elution. Proteins were separated by SDS-PAGE and stained with Coomassie Brilliant Blue G-250.



**Pc Aad1p**

MNIWAPAPEPPTKLGRHRQLAPGCGLHVSPIQLGAMSIGDKWHPYGMGTMD  
β-sheet 1 coil  
 KEASFKLLDAFYNAAGGNFIDTANVYODETSEEFIGEWMEARGNRDOMVVAT  
α-helix 1 coil β-sheet 2 coil α-helix 2 coil β-sheet 3  
 KYSLVYKRGASFEEIPQKTOYVGNLSKSMHISVHDSLRLKLRTSYIDIFYVHFV  
coil α-helix 3 coil β-sheet 4  
 DYTCTIEEVMNGLHNLVAQGVLYLGVSDTPAWVVSKANNYARMAGKTPF  
coil α-helix 4 coil β-sheet 5 coil α-helix 5 coil  
 VIYEGEWNITMRDMERDIIPMCIHGMAIAPWNVLCAGKIRTAEEERRLKSG  
β-sheet 6 coil α-helix 6 coil β-sheet 7 coil  
 EGGRTLLQFDGWLRLNETERKVKALEKVAEEIGAKSITSVAIAYLMQKFPYV  
α-helix <sup>11</sup> coil α-helix 7 coil  
 FPIVGGRKVEHLYANLEALDISLSP EQMQLNDTVPFNKGFPYLLFGDGSYNI  
β-sheet 8 coil α-helix 8 coil α-helix <sup>14</sup>  
 VHKAAGHYDKWPAQQAIRPQK\*

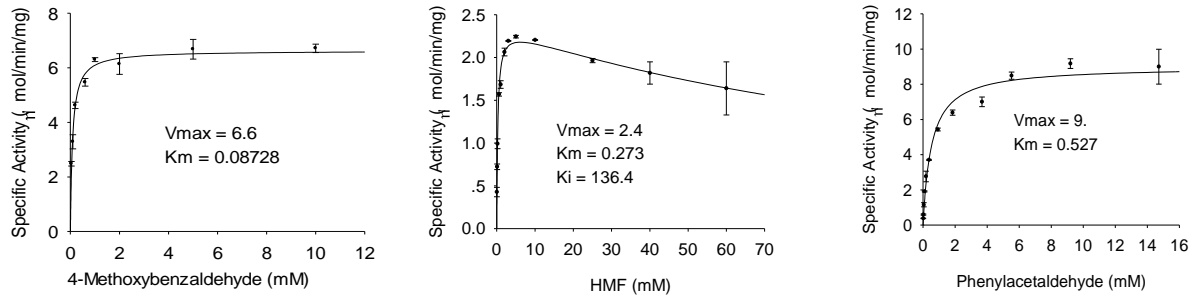
Strictly conserved a.a. among AKR:

**Aspartic acid<sup>71</sup>, Tyrosine<sup>76</sup>,**

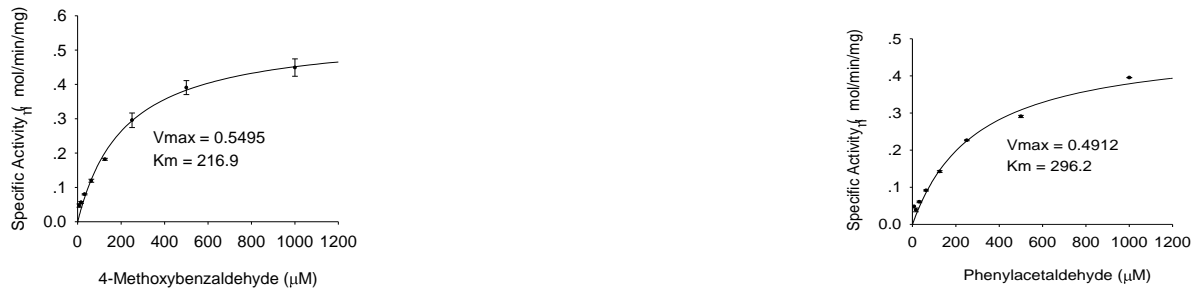
**Lysine<sup>103</sup> and Histidine<sup>152</sup>**

**FIG S4 Structural analysis of the reference PcAad1p.** *In silico* modeling reveals a (αβ)<sub>8</sub>-barrel motif typical of aldo-keto reductases. The α-helix is indicated in blue, β-sheet in purple. Two additional overhanging α-helices (at bottom of Fig. S4) and a sodium atom (in orange) are also noted. Within the catalytic tetrad, tyrosine acts as an acid in the reduction reaction, while the two additional amino acids complete a catalytic triad that activates the phenolic proton on tyrosine (Penning, TM, The aldo-keto reductase (AKRs): Overview. Chem Biol Interact. 234: 236-246, 2015; DOI : 10.1016/j.cbi.2014.09.024)

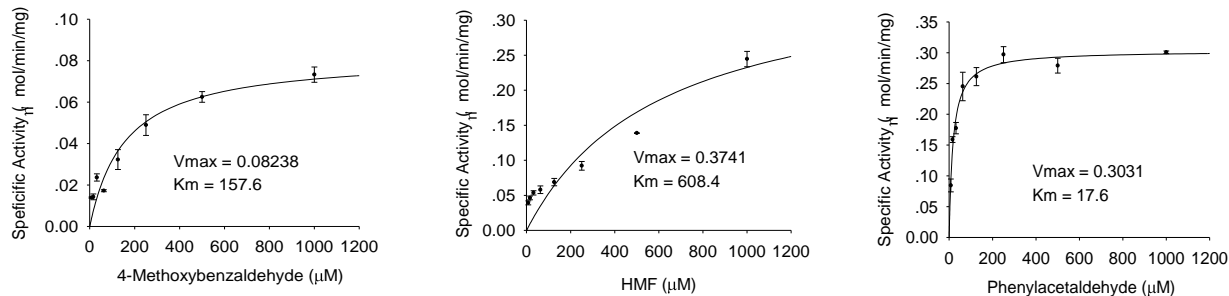
**(A) PcAad1p**



**(B) Reconstructed ScAad6<sup>518G</sup>p**



**(C) Reconstructed ScAad10<sup>-35</sup>p**



**FIG S5.** Recovery of enzyme activity from pseudogenized ScAad10 and ScAad6/16 following ancestral state reconstruction. Kinetic data were obtained by assaying purified enzymes in MES buffer (50 mM, pH 6.1) containing 0.3 mM NADPH. Each data point represents the mean  $\pm$  SEM of triplicate experiments.

**FIG S6 (A)**

ScAAD6<sup>518G</sup> (1) ATGGCTGATTTATTTGCTCCTGCTCCTGAACCACTACAGAGTTAGGACGTCTCAGAGTCTTTCTAAAAAGTGCTGGTAT  
T7 AAD6 (1) ATGGCTGATTTATTTGCTCCTGCACTGAACCACTACAGAGTTAGGACGTCTTAGGCTCTTCTAAAAAGTGCTGGTAT

ScAAD6<sup>518G</sup> (81) TAGAGTTTCCTCTCTATTTGGGAGGAATGTCAATTGGCGACGCTGGTCTGAAATCTTAGGATCAATGAGCAAGGAGC  
T7 AAD6 (81) TAGAGTTTCCTCACTAATCCTGGGAGGTATGTCTATTGGTGACGCTGGTCTGATTCATGGGATCAATGAGCAAGAAGC

ScAAD6<sup>518G</sup> (161) GAGCTTTTGAGTTGCTCGATGCTTTTATGAGGCAGGTGGAAATTTTCATTGATACTGCAAATAATTATCAAACGAACAG  
T7 AAD6 (161) GAGCTTTTGAGCTACTTATGATGCTTTTATCCAGGCAGGTGGAAATTTTCATTGATACTGCAAATAATTATCAAACGAACAG

ScAAD6<sup>518G</sup> (241) TCAGAAGCTTGGATCGGTGAATGGATGGTTTTCAAGAAAATTACGCGACCAGATTGTAATCGCCACCAAGTTTACCACTGA  
T7 AAD6 (241) TCAGAAGCTTGGATCGGTGAATGGATGGTTTTCAAGAAAATTACGCGACCAGATTGTAATCGCCACCAAGTTTACCACTGA

ScAAD6<sup>518G</sup> (321) CTATAAGAAATACGACGTTGGTGGTGGTAAAAGCGCAAACACTACTGTGGCAATCACAAGCGTAGTTTGCATGTGAGTGTGA  
T7 AAD6 (321) CTATAAGAAATACGACGTTGGTGGTGGTAAAAGCGCAAACACTACTGTGGCAATCACAAGCGTAGTTTGCATGTGAGTGTGA

ScAAD6<sup>518G</sup> (401) GGGATTCTCTCCGCAAATTGCAAACCTGATTGGATTGATATACTTTACGTTCACTGGTGGGATTATATGAGTTCTATCGAA  
T7 AAD6 (401) GGGATTCTCTCCGCAAATTGCAAACCTGATTGGATTGATATACTTTACGTTCACTGGTGGGATTATATGAGTTCTATCGAA

ScAAD6<sup>518G</sup> (481) GAAGTTATGGATAGTTTACACATTCTAGTGCAGCAGGGCAAAGTCTCTATTTGGGTGTGTCTGATACGCCTGCCTGGGT  
T7 AAD6 (481) GAAGTTATGGATAGTTTACACATTCTTGTGCAGCAGGGCAAAGTCTCTATCTAGGTGTGTCTGATACGCCTGCCTGGGT

ScAAD6<sup>518G</sup> (561) TGTTTCTGCGGCAAATTACTACGCTAAATCTCATGGTAAAACCCCTTTTAGCATCTATCAAGGTAAATGGAACCTGTTGA  
T7 AAD6 (561) TGTTTCTGCGGCAAATTACTACGCTAAATCTCATGGTAAAACCCCTTTTAGCATCTATCAAGGTAAATGGAACCTGTTGA

ScAAD6<sup>518G</sup> (641) ACAGAGACTTCGAACGTGATATCATTCCAATGGCTAGGCATTTTCGGTATGGCCCTCGCCCATGGGATGTCATGGGAGGT  
T7 AAD6 (641) ACAGAGACTTCGAACGTGATATCATTCCAATGGCTAGGCATTTTCGGTATGGCCCTCGCCCATGGGATGTCATGGGAGGT

ScAAD6518GG (721) GGAAGATTTTCAGAGTAAAAAGCAATGGAGGAACGGAGGAAGAATGGAGAGGGTATTCGTTCTTTCGTTGGCGCCTCTGA  
T7 AAD6 (721) GGAAGATTTTCAGAGTAAAAAGCAATGGAGGAACGGAGGAAGAATGGAGAGGGTATTCGTTCTTTCGTTGGCGCCTCTGA

ScAAD6<sup>518G</sup> (801) ACAAACAGATGCAGAAATCAAGATTAGTGAAGCATTGGCCAAGGTTGCTGAGGAACATGGCACTGAATCTGTTACTGCTA  
T7 AAD6 (801) ACAAACAGATGCAGAAATCAAGATTAGTGAAGCATTGGCCAAGGTTGCTGAGGAACATGGCACTGAATCTGTTACTGCTA

ScAAD6<sup>518G</sup> (881) TTGCTATTGCCTATGTTTCGCTCTAAGGCGAAAAATGTTTTTCCATTGGTTGGAGGAAGGAAAATTGAACACCTCAAACAG  
T7 AAD6 (881) TTGCTATTGCCTATGTTTCGCTCTAAGGCGAAAAATGTTTTTCCATTGGTTGGAGGAAGGAAAATTGAACACCTCAAACAG

ScAAD6<sup>518G</sup> (961) AACATTGAGGCTTTGAGCATTAATTAACACCAGAACAATAAAGTACTTAGAAAGTATTATTCCTTTTGATGTTGGTTT  
T7 AAD6 (961) AACATTGAGGCTTTAAGTATCAAACTGACACCAGAACAATAGAAATACTTAGAAAGTATTATTCCTTTTGATGTTGGTTT

ScAAD6<sup>518G</sup> (1041) TCCTACTAATTTTATCGGTGATGATCCGGCTGTTACCAAGAAGGCTTCACTTCTCACGGCAATGTCTGCGCAGATTTTCCT  
T7 AAD6 (1041) TCCTGCTAATTTTATCGGTGATGATCCGGCTGTTACCAAGAAGGCTTCACTTCTCACGGCAATGTCTGCGCAGATTTTCCT

ScAAD6<sup>518G</sup> (1121) TCGATTAA  
T7 AAD6 (1121) TCGATTAA

**FIG S6 (B)**

ScAad6p<sup>518G</sup> (1) MADLFAPAPEPS TELGRLRVLSK SAGIRVSPLILGGMSIGDAWS EILGSMSKER AFELLDAFYE AGGNFIDTANNYQNEQ  
T7 Aad6p (1) MADLFAPAPEP TELGRLRVLSK IAGIRVSPLILGGMSIGDAWS GFMGSMDKEQ AFELLDAFYQ AGGNFIDTANNYQNEQ

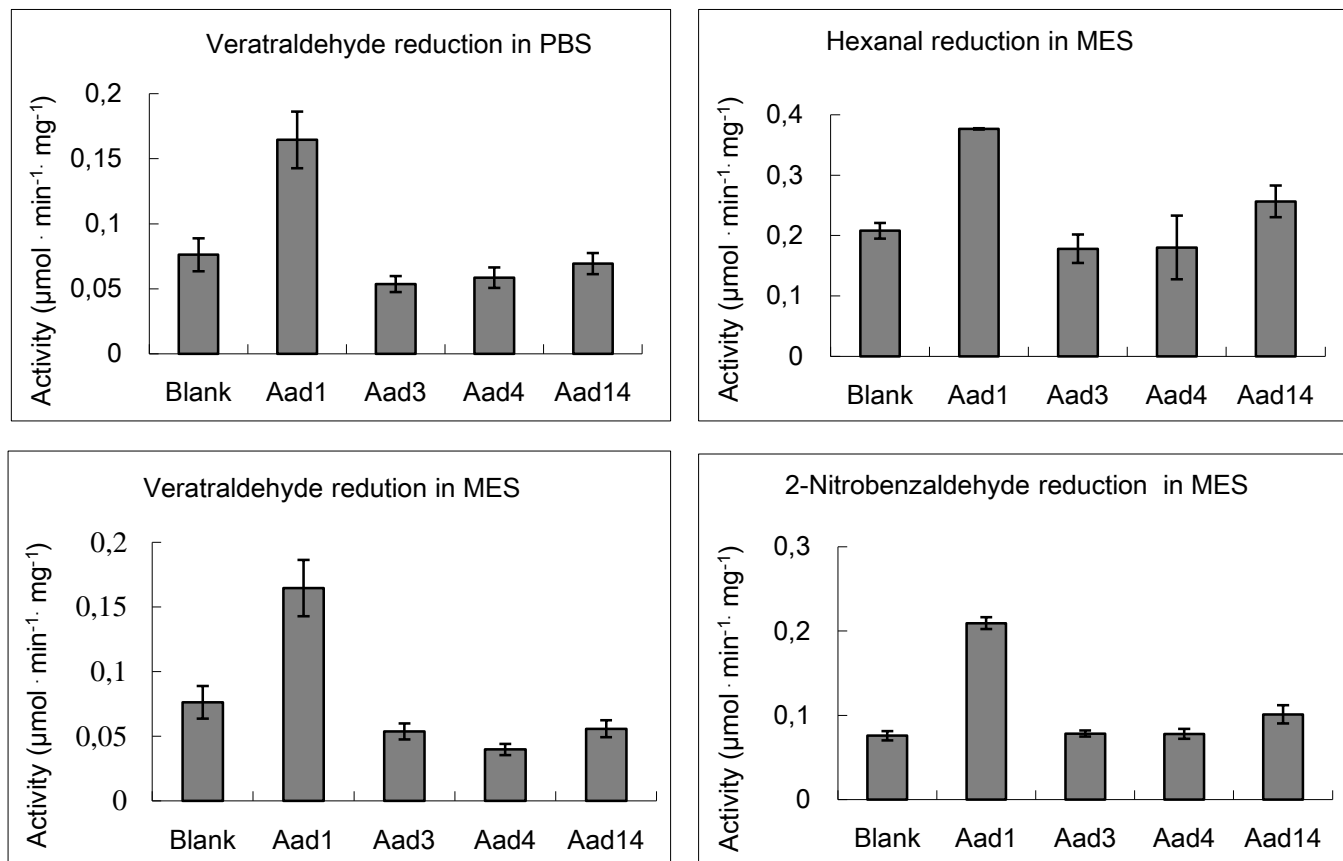
ScAad6p<sup>518G</sup> (81) SEAWIGEWVSRKLRDQIVIATKFT TDYKKYDVGGGKSANYCGNHKRSLHVSVRDSLRLKLQTDWIDILYVHWWDYMSSIE  
T7 Aad6p (81) SEAWIGEWVSRKLRDQIVIATKFT IDYKKYDVGGGKSANYCGNHKRSLHVSVRDSLRLKLQTDWIDILYVHWWDYMSSIE

ScAad6p<sup>518G</sup> (161) EVMDSLHILVQQGVLYLGVSDTPAWVVSAAANY AKSHGKTPFSIYQGKWNLLNRDFERDII PMARHFGMALAPWDVMGG  
T7 Aad6p (161) EVMDSLHILVQQGVLYLGVSDTPAWVVSAAANY T KSHGKTPFSIYQGKWNLLNRDFERDII PMARHFGMALAPWDVMGG

ScAad6p<sup>518G</sup> (241) GRFQSKKAMEERRKNGEGIRSFVGASEQTD AEIKISEALAKVAEEHGTESVTAIAIAYVRSKAKNVFPLVGGRKIEHLKQ  
T7 Aad6p (241) GRFQSKKAMEERRKNGEGIRSFVGASEQTD AEIKISEALAKVAEEHGTESVTAIAIAYVRSKAKNVFPLVGGRKIEHLKQ

ScAad6p<sup>518G</sup> (321) NIEALSIKLTPEQIK YLESIIPFDVGFP T NFIGDDPAVTKKAS L LTAMSAQISFD -  
T7 Aad6p (321) NIEALSIKLTPEQIE YLESIIPFDVGFP ANFIGDDPAVTKKAS F LTAMSAQISFD -

**FIG S6. A full-length AAD6 identified in *S. cerevisiae* T7 strain living in oak tree exudate.** This AAD6 exhibits >95% consensus positions with the reconstructed hypothetical AAD6. (A) and (B) show the alignment of nucleic acids and the amino acid sequences they encode.



**FIG S7.** Aldehyde reduction activity measured with *S. cerevisiae* crude enzyme extract overexpressing Aad. BY4741 strains harbor “blank” or Aad (“Aad1”, “Aad3”, “Aad4”, “Aad14”) YEPlac 195 vectors. Activity was assayed on crude cell extracts in 250  $\mu\text{L}$  MES buffer (50 mM, pH 6.1) containing 0.3 mM NADPH, 0.3 mM substrates and 20  $\mu\text{L}$  crude extract. Veratraldehyde (3,4-dimethoxybenzaldehyde), hexanal and 2-Nitrobenzaldehyde were used as they are the most reactive substrates (Fig 1) for purified *PcAad1p*, *ScAad4p* and *ScAad14p*.

**TABLE S1. High variability of AAD ORFs in sequenced *S. cerevisiae* genomes, as compared to other non-telomeric aldehyde reductases.**

Strains	BY4741	YJM789	EC1118	AWRI1631	AWRI796	Lalvin QA23	VL3	FostersB	FostersO	Kyokai no. 7	Sigma1278b
Query	Lab	Lab	Wine	Wine	Wine	Wine	Wine	Beer	Beer	Sake	Pathogenic
<b>AAD3</b>	100/100	96.4/93.7	99.8/99.5	99.9/99.5	≥99.4/82.1	99.7/67.6	99.8/85.2	N.R.	N.R.	N.R.	98.2/82.7
<b>AAD4<sup>-136G</sup></b>	99.9/83.5	93.2/93.9	94.3/84.6	94.4/84.3	94.3/84.3	94.2/71.3	94.2/84.3	95.9/84.3	94.3/84.3	93.4/94.7	N.R.
<b>AAD6<sup>518G</sup></b>	99.9/-	N.R.	N.R.	N.R.	N.R.	N.R.	N.R.	N.R.	N.R.	99.8/-	N.R.
<b>AAD10<sup>-35C</sup></b>	99.9/76.4	99.9/76.4	99.7/76.4	99.8/76.1	99.6/99.5	N.R.	92.6~93.2/72.9	N.R.	N.R.	*/*	87.3/55.2
<b>AAD14</b>	100/100	99.2/99.5	97.1/97.6	N.R.	97.1/97.6	96.9/52.0	96.9/65.3	97.3/97.1	97.5/97.3	100/100	100/100
<b>AAD15</b>	100/100	N.R.	97.7/95.1	N.R.	N.R.	N.R.	N.R.	N.R.	N.R.	N.R.	100/100
<b>YPL088W</b>	100/100	99.2/98.5	99.1/98.5	*/*	99.3/98.5	98.9/98.3	99.1/81.9	99.2/98.5	99.0/80.5	99.7/99.1	100/100
<b>ADH1</b>	100/100	98.6/98.6	99.0/98.6	98.6/98.6	98.6/98.3	98.9/98.6	98.6/98.6	98.7/98.6	98.4/98.0	98.9/98.6	100/100
<b>ADH2</b>	100/100	98.4/98.0	97.8/98.3	N.R.	97.8/98.3	97.7/43.0	97.8/98.3	97.2/97.7	94.1/95.4	99.5/99.1	N.R.
<b>ADH3</b>	100/100	100/100	100/100	100/100	100/100	100/100	100/100	99.7/99.5	99.7/99.5	99.9/100	100/100
<b>ADH4</b>	100/100	99.9/99.5	98.3/98.2	98.2/97.9	98.3/98.2	98.0/23.8	98.3/98.2	99.8/99.2	98.1/97.4	99.7/99.5	98.3/98.2
<b>ADH5</b>	100/100	99.6/100	99.6/100	99.4/99.4	99.6/100	99.6/100	99.6/100	99.5/99.4	99.7/99.4	99.9/100	100/100
<b>ADH6</b>	100/100	99.7/100	99.5/99.4	99.7/100	99.7/100	99.7/100	99.5/30.2	99.1/98.6	99.5/100	100/100	100/100
<b>ADH7</b>	100/100	99.8/99.4	100/100	100/100	100/100	100/100	100/100	N.R.	N.R.	N.R.	100/100
<b>GRE3</b>	100/100	100/100	99.9/99.4	100/100	100/10	99.9/99.4	100/100	99.9/99.4	99.8/87.8	99.7/100	100/100
<b>GCY1</b>	100/100	100/100	99.6/98.7	99.8/99.4	99.8/99.4	99.3/≤25.2	99.7/99.0	99.3/98.7	99.4/99.4	99.3/99.0	99.8/99.8
<b>YPR1</b>	100/100	99.3/12.8	100/100	100/100	99.8/93.6	100/100	99.6/99.4	99.7/99.4	99.6/67.4	99.8/99.4	100/100
<b>YDL124W</b>	100/100	99.4/98.4	99.4/98.4	99.0/97.4	99.0/97.4	99.4/98.4	99.3/98.4	99.8/99.0	99.4/98.7	98.0/70.0	98.4/97.1
<b>YJR096W</b>	100/100	99.4/99.3	100/100	99.9/48.4	100/100	99.9/86.2	100/100	99.9/100	99.9/100	99.6/98.9	100/100
<b>ARA1</b>	100/100	99.6/100	99.8/99.4	99.7/99.1	99.8/99.4	99.8/99.4	99.8/99.4	99.2/50.4	99.5/8.7	75.4/75.4	100/100

Nucleic acid/amino acid sequence percentage similarity relative to query sequences from lab strain S288C. N.R.: AAD homologs were not retrieved. Asterisks (\*): percentage similarity not calculated when only partial sequences were found at the end of a sequencing contig with missing 3' or 5' coding sequences.