

Table S1: Soils characteristics and archaeal and bacterial *amoA* copy numbers

	Location	Sampling date	Usage	Soil type	Sand/silt/clay (%)	pH ¹	WEON ^{1,2}	WEOC ^{1,3}	AOA <i>amoA</i> copies g ⁻¹ soil (S.D. ⁴)	AOB <i>amoA</i> copies g ⁻¹ soil (S.D.)	isoprenoidal GDGT [µg g ⁻¹ dry soil]	Crenarchaeol [µg g ⁻¹ dry soil]
sandy ecosystem (RUD)	Darmstadt, Germany	Sept. 2005	Natural grasslands	Sandy soil	n.d. ⁴	7.06	0.9	7.6	5.5x10 ⁷ (5.4x10 ⁵)	1.0x10 ⁶ (1.5x10 ⁵)	0.44	0.21
limestone soil (GSF)	Munich, Germany	Oct. 2005		Sandy soil	n.d.	6.9	1.08	10.2	3.5x10 ⁷ (1.2x10 ⁷)	2.5x10 ⁶ (1.8x10 ⁶)	n.d.	n.d.
pasture land (KRO)	Bergen, Norway	Oct. 2005		Sandy loam	n.d.	6.1 (H ₂ O)	3.76	25.85	4.7x10 ⁷ (3.9x10 ⁶)	3.2x10 ⁷ (5.9x10 ⁶)	3.24	0.33
grassland (STO)	Bergen, Norway	Oct. 2005		Sandy loam	n.d.	5.5 (H ₂ O)	1.16	5.5	1.3x10 ⁸ (1.4x10 ⁷)	5.2x10 ⁷ (4.2x10 ⁷)	1.11	0.05
tomato field (B77T)	Santorini, Greece	Oct. 2005		Sandy soil	16-20/70-75/9-15	6.0	0.76	5.4	2.6x10 ⁷ (3.5x10 ⁵)	1.1x10 ⁷ (3.2x10 ⁶)	0.24	0.06
vineyard (B77V)	Santorini, Greece	Oct. 2005		Sandy soil		6.2	0.87	5.6	3.5x10 ⁷ (1.1x10 ⁷)	6.5x10 ⁶ (5.3x10 ⁵)	0.18	0.06
barley field (E16)	Santorini, Greece	Oct. 2005		Sandy soil		6.0	0.56	3	7.3x10 ⁶ (1.3x10 ⁵)	2.6x10 ⁶ (6.9x10 ⁴)	0.12	0.05
non-fertilized site (L-I)	Bad Lauchstädt, Germany	July 2002	Agricultural ecosystems	Sandy loam	6-10/70-73/19-21	6.4	1.22	35.2	1.5x10 ⁷ (8.0x10 ⁵)	6.5x10 ⁶ (2.8x10 ⁵)	n.d.	n.d.
mineral fertilized site (L-n)	Bad Lauchstädt, Germany	July 2002		Sandy loam		6.3	2.34	57.1	5.6x10 ⁷ (9.6x10 ⁵)	7.2x10 ⁶ (5.9x10 ⁵)	n.d.	n.d.
mineral & organic fertilized site (L-h)	Bad Lauchstädt, Germany	July 2002		Sandy loam		6.7	4.76	85.1	7.0x10 ⁷ (7.5x10 ⁵)	4.7x10 ⁶ (2.0x10 ⁵)	n.d.	n.d.
non-tillaged site (R-nt)	Rommersheim, Germany	July 2002		Silty clay loam	2-5/69-75/33-35	7.28	2.3	16.3	7.9x10 ⁷ (3.0x10 ⁶)	1.1x10 ⁶ (4.4x10 ⁵)	n.d.	n.d.
ploughed site (R-p)	Rommersheim, Germany	July 2002		Silty clay loam		7.33	2.4	7.7	6.1x10 ⁷ (1.6x10 ⁶)	6.6x10 ⁶ (1.7x10 ⁶)	0.54	0.18
ploughed site (R-p 30)	Rommersheim, Germany	July 2002		Silty clay loam		7.3	2.05	9.3	7.7x10 ⁷ (2.2x10 ⁷)	7.8x10 ⁶ (5.9 x10 ⁵)	0.71	0.22

¹ measured using the CaCl₂ extraction if not indicated otherwise

²WEON: water extractable organic nitrogen [mg/kg dw soil]

³WEOC: water extractable organic carbon [mg/kg dw soil]

⁴n.d.: not determined

⁵ S.D.: standard deviation

Table S2: Relative abundance of AOA and AOB in investigated depth profiles

	depth (cm)	WEON ¹	WEOC ²	AOA <i>amoA</i> copies g ⁻¹ soil (S.D. ⁵)	AOB <i>amoA</i> copies g ⁻¹ soil (S.D.)	AOA <i>amoA</i> copies µg ⁻¹ DNA (S.D.)	AOB <i>amoA</i> copies µg ⁻¹ DNA (S.D.)	%AOA / microbiota ³	%AOB / microbiota ⁴	isoprenoidal GDGT [µg g ⁻¹ dry soil]	Crenarchaeol [µg g ⁻¹ dry soil]
sandy ecosystem (RUD)	0-10	0.9	7.6	5.5x10 ⁷ (5.4x10 ⁶)	1.0x10 ⁶ (1.5x10 ⁵)	2.3x10 ⁸ (2.3x10 ⁵)	3.5x10 ⁴ (5.0x10 ³)	1.0	0.006	0.44	0.21
	10-20	0.9	7.6	7.2x10 ⁷ (1.3x10 ⁷)	4.3x10 ⁵ (1.1x10 ⁵)	4.8x10 ⁸ (9.1x10 ⁵)	2.9x10 ⁴ (7.5x10 ³)	2.1	0.005	n.d. ⁵	n.d.
	20-30	0.7	6.3	3.6x10 ⁷ (5.8x10 ⁶)	2.1x10 ⁵ (1.7x10 ⁴)	4.2x10 ⁸ (6.8x10 ⁵)	2.5x10 ⁴ (2.0x10 ³)	1.9	0.004	n.d.	n.d.
	30-40	0.6	5.9	1.4x10 ⁷ (4.9x10 ⁶)	5.9x10 ⁴ (1.8x10 ⁴)	2.0x10 ⁸ (7.1x10 ⁵)	8.5x10 ³ (2.7x10 ³)	0.9	0.001	0.10	0.04
	40-50	0.4	5.3	1.8x10 ⁷ (3.1x10 ⁶)	1.6x10 ⁴ (1.8x10 ³)	7.2x10 ⁸ (1.2x10 ⁶)	6.2x10 ³ (7.0x10 ²)	3.2	0.001	n.d.	n.d.
	60-70	0.4	4.5	3.2x10 ⁶ (3.9x10 ⁵)	3.8x10 ³ (1.7x10 ³)	2.5x10 ⁸ (3.0x10 ⁵)	2.9x10 ³ (1.3x10 ³)	1.1	0.001	0.04	0.02
non-fertilized acre (L-l)	0-15	1.22	35.2	1.5x10 ⁷ (8.0x10 ⁵)	6.5x10 ⁴ (2.8x10 ⁴)	4.8x10 ⁸ (2.6x10 ⁵)	3.1x10 ⁴ (9.0x10 ³)	2.1	0.005	n.d.	n.d.
	15-30	1.09	28.4	1.8x10 ⁷ (1.2x10 ⁶)	2.7x10 ⁴ (3.6x10 ³)	7.7x10 ⁸ (5.3x10 ⁵)	1.2x10 ⁴ (1.6x10 ³)	3.4	0.002	n.d.	n.d.
	30-40	0.99	27.3	1.4x10 ⁷ (6.7x10 ⁵)	4.6x10 ³ (6.1x10 ²)	6.8x10 ⁸ (3.3x10 ⁶)	2.3x10 ³ (3.0x10 ²)	3.0	0.0004	n.d.	n.d.
mineral fertilized acre (L-n)	0-15	2.34	57.1	5.6x10 ⁷ (9.6x10 ⁶)	7.2x10 ⁵ (5.9x10 ⁵)	1.3x10 ⁷ (2.3x10 ⁶)	2.2x10 ⁵ (1.4x10 ⁵)	5.8	0.038	n.d.	n.d.
	15-30	1.44	53.3	4.2x10 ⁷ (1.5x10 ⁶)	9.7x10 ⁴ (1.1x10 ⁴)	1.1x10 ⁷ (4.0x10 ⁵)	2.6x10 ⁴ (3.1x10 ³)	4.9	0.005	n.d.	n.d.
	30-40	3.37	57.5	1.5x10 ⁷ (2.8x10 ⁶)	2.2x10 ⁴ (1.3x10 ⁴)	5.7x10 ⁸ (1.1x10 ⁶)	8.5x10 ³ (5.1x10 ³)	2.5	0.001	n.d.	n.d.
mineral & organic fertilized acre (L-h)	0-15	4.76	85.1	7.0x10 ⁷ (7.5x10 ⁶)	4.7x10 ⁵ (2.0x10 ⁵)	9.0x10 ⁸ (9.7x10 ⁵)	1.6x10 ⁵ (2.6x10 ⁴)	3.9	0.028	n.d.	n.d.
	15-30	5.24	85.6	9.3x10 ⁷ (1.0x10 ⁷)	7.3x10 ⁵ (1.3x10 ⁵)	1.1x10 ⁷ (1.3x10 ⁶)	9.0x10 ⁴ (1.6x10 ⁴)	5.0	0.016	n.d.	n.d.
	30-40	7.87	78.2	5.2x10 ⁷ (8.4x10 ⁶)	2.1x10 ⁵ (4.3x10 ⁴)	1.1x10 ⁷ (1.9x10 ⁶)	4.7x10 ⁴ (9.6x10 ³)	5.0	0.008	n.d.	n.d.

¹WEON: water extractable organic nitrogen [mg/kg dw soil]

²WEOC: water extractable organic carbon [mg/kg dw soil]

³assuming an average genome size for all microorganisms of 4Mbp and 1 *amoA* copy per AOA genome

⁴assuming an average genome size for all microorganisms of 4Mbp and an average of 2.5 *amoA* copies per genome in AOB

⁵ S.D.: standard deviation

Table S3: Results of MPN-PCR experiments with DNA extracted from two different top soils (RUD and L-h)

L-h	series	10 ng	1 ng	100 pg	10 pg	1 pg	100 fg	10 fg
Arch 16S rRNA gene (Arch20F/958R)	I	+	+	+	+	+	-*	-
	II	+	+	+	+	+	-	-
	III	+	+	+	+	+	-	-
Bact 16S rRNA gene (CTO189F/654R)	I	+	+	+	-	-	-	-
	II	+	+	+	-	-	-	-
	III	+	+	+	-	-	-	-
Arch amoA gene (amoA19F/634R)	I	+	+	+	+	+	-	-
	II	+	+	+	+	-	-	-
	III	+	+	+	+	+	+	-
Bact amoA gene (A189/ amoA-2R)	I	+	+	+	+	-	-	-
	II	+	+	+	-	-	-	-
	III	+	+	+	+	+	-	-
Bact amoA gene (amoA-1F*/2R)	I	+	+	+	+	-	-	-
	II	+	+	+	+	-	-	-
	III	+	+	+	-	-	-	-
RUD	series	10 ng	1 ng	100 pg	10 pg	1 pg	100 fg	10 fg
Arch 16S rRNA gene (Arch20F/958R)	I	+	+	+	+	+	-	-
	II	+	+	+	+	-	-	-
	III	+	+	+	+	-	-	-
Bact 16S rRNA gene (CTO189F/654R)	I	+	+	+	-	-	-	-
	II	+	+	+	-	-	-	-
	III	+	+	+	-	-	-	-
Arch amoA gene (amoA19F/634R)	I	+	+	+	+	+	-	-
	II	+	+	+	+	+	-	-
	III	+	+	+	+	+	-	-
Bact amoA gene (A189/amoA-2R)	I	+	+	+	+	-	-	-
	II	+	+	+	+	-	-	-
	III	+	+	+	+	-	-	-
Bact amoA gene (amoA-1F*/2R)	I	+	+	+	+	-	-	-
	II	+	+	+	-	-	-	-
	III	+	+	+	+	+	-	-

* +/- indicates whether a PCR product was obtained or not at the given amount of DNA (first row) with the respective primer set (shown in the first column). Note that characterised AOB have 2 to 3 copies of the *amoA* gene per chromosome¹, and only 1 copy of the 16S rRNA gene per chromosome⁵¹, which accounts for the differences in sensitivity in this assay.

Table S4: Statistics from large-scale transcriptional study

No. of runs on a Roche GS20 sequencer	1
No. of cDNA sequence reads	314,041
Average read length (bp)	96.4
Total sequences (Mb)	30.27
Similarity in non-redundant (nt) DB ¹	210,523 (Archaea ² 2,878 = 1.37%)
Similarity in ribosomal DB (16S rRNA) ³	106,643

¹Search performed using MEGABLAST against nucleotide database release of February 8th, 2006.

²Search performed by keyword search on the BLAST output file using the classification names of the Taxonomic Outline of Prokaryote Genera of Bergey's Manual.

³Search performed using MEGABLAST against ribosomal database release of February 2th, 2006 (RDP 9.36).

Table S5: Sequence reads from large-scale cDNA study *

Header	Sequence
Arch958	Number = 19
002422_0743_1016	ATTTGGTTCAATTGGAGCCAACGCCGGAATCTTACCGGGGCGACAGC AGAATGAAGGTCAAGCTGAAGACTTTACCAGACAAGCTGAGC
127285_0249_0851	GACTGAAACTTAAAGGAATTGGCGGGGGAGCACCACAAGGGTGAAGC CTGCGGTTCAATTGGAGTCAACGCCGGC
164769_0880_3174	CTCAGCTTGTCTGAGTAAAGTCTTCAGCTTGACCTTCACACTGCTGTCGC CCCGGTAAGATTTCCGGCGTTGGCTCCAATTGAACCGCAGGCTTCACC CCTTGTTGGTG
177800_0587_3036	CCCGGCGTTGACTCCAATTGAACCGCAGGCTTCACCCTTGTGGTGCTC CCCGCCAATTCCTTTAAGTTTCAGTCTTGCACCGTACTACCAGGCGG CAGACTTAACGGCTTCCCTGCGG
193756_0814_3802	CTGAAACTTAAAGGAATTGGCGGGGGAGCACCACAAGGGTGAAGCCT GCGGTTCAATTGGAGTCAACGCCGGAATCTTACCGGGGCGACAGCA GAATGAAGGTCAAGC
228651_1046_3166	CTCAGCTTGTCTGGTAAAGTCTTCAGCTTGACCTTCACTCTGCTGTCGC CCCGGTAAGATTTCCGGCGTTGGCTCCAATTGAACCGCAGGCTTCACCC TTGTGGTG
235050_1827_1019	CCACAAGGGTGAAGCCTGCGGTTCAATTGGAGCCAACGCCGGAATC TTACCGGGGGCGACAGCAGTGTGAAGGTCAAGCTGAAGACTTTACTC GACAAGC
284324_1031_1738	CACAAGGGGTGAAGCCTGCGGTTCAATTGGAGTCAACGCCGGGAATC TTACCGGGGGCGACAGCAGAGTGAAGGTCAAGCTGAAGACTTTACCAG ACAAGC
300882_0679_1699	GCTTGTCTGGTAAAGTCTTCAGCTTGACCTTCATTCTGCTGTCGCCCCG GTAAGAGTTTCCGGCGTTGGCTCCAATTGAACCGCAGGCTTCACCCTT GTGGTGCTCCCCGC
048555_2861_0356	CCCCGGTAAGATTTCCGGCGTTGGCTCCAATTGAACCGCAGGCTTCAC CCCTTGTGGTGCTCCCCGCCAATTCCTTTAAGTTTCAGTCTTGCACCG TACTCCCCAGGCGGCAGACTTAACGGC
069159_3343_3764	CCCCCGGTAAGATTTCCGGCGTTGGCTCCAATTGAACCGCAGGCTTCA CCCCTTGTGGTGNTCCCCGCCAATTCCTTTAAGTTTCAGTCTTGCACCG GTATCCCCAGGCGGCAGCTTAACGGC
097478_2639_3510	GTACGGTCGCAAGACTGAACTTAAGGAATTGGCGGGGGAGCACCACA AGGGTGAAGCCTGCGGTTCAATTGGAGTCAACGCCGGGAATCTTACC GGGGGCGACAGCAGAATGAAGGTC
121130_2652_2871	GGGAGCACCACAAGGGGTGAAGCCTGCGGTTCAATTGGAGCCAACGC CGGAAATCTTACCGGGGGCGACAGCAGTGTGAAGGTCAAGCTGAAGA CTTTACTCGACAAGC
140270_2678_3074	GCTGTTGCCCCGGTAAGATTTCCGGCGTTGACTCCAATTGAACCGCA GGCTTCACCCCTTGTGGTGCTCCC
172828_3497_3796	CTCTCTCAGTCTTGTCTGGTAAAGTCTTCAGCTTGACCTTCACTTCTCGC TGTCGCCCCGGTAAGATTCCCGGCGTTGGCTCCAATTGAACCGCAGGC TTCACCC
205931_2772_0327	CTGCTGTCGCCCCGGTAAGATTCCCGGCGTTGACTCCAATTGAACCGC AGGCTTCACCCCTTGTGGTGCTCCC
226287_2487_1761	CGCCTGGGGAGTACGGTCGCAAGACTGAACTTAAAGGATTGGCGGG GGAGCACCACAAGGGGTGAAGCCTGCGGTTCAATTGGAGCCAACGCC GGAATCTTACGGGGGCACAGCAG
230978_2898_1783	CTCAGCTTGTCTGGTAAAGTCTTCAGCTTGACCTTCACTCTGCTGTCGC CCCGGTAAGATTTCCGGCGTTGGCTCCAATTGAACCGCAGGCTTCACC CTTGTTGGTG
275823_3382_0855	GCCGCAGGGAAGCCGTTAAGTCTGCCGCCTGGGTAGTACGGTTCGCAA GACTGAACTTAAAGGAATTGGCGGGGGAGCACCACAAGGGGTGAAGC CTGCGGTTCAATTGGAGCCAACGCCGGG
Cren1.lb_537	Number = 25
042121_1208_3490	TGCCCGCTAAGGTATCTTTTGGCACCTCTAAAATGGTGCAGAATAAGG

	GTGGGCAAGTCTGGTGTGACCCGCCGCGGTAATACCAGCACCCCGAG TTGTCGGGACGATTATTGGG
089221_0616_0138	CTGACCACTCGGGGTGCTGGTATTACCGCGGCGGCTGACACCAGACTT GC
129308_0091_3855	GCCCGCTAAGGGTATCTTTTGGCACCTCTAAAAATGGTGCAGAATAAG GGGTGGGCAAGTCTGGTGTGACCCGCCGCGGTAATACCAGCACCCCG AGTTGTCTGGGACC
159367_0522_1513	CCCGACCACTCGGGGTGCTGGTATTACCGCGGCGGCTGACACCAGACT TGCCCACCCCTTATTCTGCACCATTTTTAGAGGTGCCAAAAGACTTCCT TAGCGGAAATCAC
176111_0806_2417	CCAGCACCCCGAGTGGTCTGGGACGATTATTGGGCCTAAAGCATCCGTA GCCGGTTCTACAAGTCTTCCGTTAAATCCACCTGCTTAACAGATGGGC TGCGGAGGATACTATGG
197661_0745_3327	GGAGGACTTGACAGAACAGGCTACGGATGCTTTAGGCCAATAATCGTC CCGACAACCTCGGGGTGCTGGTATTACCGCGGCGGCTGACACCAGACTT GCCACCCTTATTC
218120_0548_2711	CTTTTGGCACCTCTAAAAATGGTGCAGAATAAGGGGTGGGCAAGTCTG GTGTGACCCGCCGCGGTAATACCAGCACCCCGAGTTGTCTGGGACGATT ATTGGGCCTAAAGCATCCG
218729_1652_2811	CCCGCTAAGGGATCTTTTGGCACCTCTAAAAATGGTGCAGAATAAGGG TGGGCAAGTCTGGTGTGACCCGCCGCGGTAATACCAGCACCCCGAGT GGTCGGGACGATTATTGGGCC
219814_0622_1537	GCCCAATAATCGTCCCGACCACTCGGGGTGCTGGTATTACCGCGGCGG CTGACACCAGACTTGCCCACCCCTTATTCTGCACCATTTTTAGAGGTGCC AAAGATCCCCTTAGCGGGAA
254939_1915_0787	AGAAAGTCTTTTGACATCTCTAAAATGGTGTAGAATAAGGGTGGGCA AGTCTGGTGTGACCCGCCGCGGTAATACCAGCACCCCGAGTGGTCGG GACGATTATTGAGCC
290065_0506_0181	CCCGGCTACGGATGCTTTAGGCCCAATAATCGTCCCGACCACTCGGGG TGCTGGTATTACCGCGGCGGCTGACACCAGACTTGCCCACCCCTTATT CTACACCATTTTTAGAGA
296022_0559_1392	GGGTGGGCAAGTCTGGTGTGACCCGCCGCGGTAATACCAGCACCCCG AGTGGTCGGGACGATTATTGGGCCTAAAGCATCCGTAGCCGGTTTTAC AAGTCCTCCGTTAAATCCAACCTGCTTAAC
311638_0531_3974	TGCAGAATAAGGGGTGGGCAAGTCTGGTGTGACCCGCCGCGGTAATA CCAGCACCCCGAGTTGTCTGGGACGATTATTGGGCCTAAAGCATCCGTA GCCTGTTCTGCAAG
314145_0111_2664	CCAATAATCGTCCCGACCACTCGGGGTGCTGGTATTACCGCGGCGGCT GACACCAGACTTGCCCACCCCTTATTCTACACCATTTTTAGAGATGTCA AAAGACTTTCCTAGCGAAAG
321901_1209_0270	CCAATAATCGTCCGACCACTCGGGGTGCTGGTATTACCGCGGCGGCTG ACACCAGACTTGCCCACCCCTTATTCTGCACCATTTTTAGAGTGCCAAA AGATCCCTTAGCGGAA
063423_3228_3658	CGCCGCGGTAATACCAGCACCCCGAGTGGTCTGGGACGATTATTGGGC CTAAAGCATCCGTAGCCGGTTCTACAAGTCTTCCGTTAAATCCACCTG CTTAACAGATGGGCTGCGGAGG
085313_2324_2665	ATGCTTTAGGCCCAATAATCGTCCCGACCACTCGGGGTGCTGGTATTA CCGCGGCGGCTGACACCAGACTTGCCCACCCCTTATTC
097244_2805_0323	AGAATAAGGGGTGGGCAAGTCTGGTGTGACCCGCCGCGGTAATACCA GCACCCCGAGTGGTCTGGGACGATTATTGGGCC
105649_2794_0282	TTTGACATCTCTAAAATGGTGTAGAATAAGGGGTGGGCAAGTCTGGTG TCAGCCGCCGCGGTAATACCAGCACCCCGAGTGGTCTGGGACGATTATT GGCCTAAGCC
152648_2663_2303	CCCAATATCGTCCGACCACTCGGGGTGCTGGTATTACCGCGGCGGCTG ACACCAGTACTTAGCCACCCTTATTCTG
214162_3474_1143	CCAGCACCCCGAGTGGTCTGGGACGATTATGGCCTAAGCATCCGTAGC CGGTTCTGCAAGTCCCCGTTAATCCAACCTGCTCAACAGTGGCTGCGGA GGATACTACAG
224401_2385_1232	ATGCTTTAGGCCAATAATCGTCCCGACCACTCGGGGTGCTGGTATTAC

	CGCGGCGGCTGACACCAGACTTGCCCACCCTTATTC
291669_2827_0674	GCCCAACTAATCGTCCCGACCACTCGGGGTGCTGGTATTACCGCGGCG GCTGACACCAGACTTGCCCACCCTTATTCTGCACCATTTTTAGAGGT GCCAAAAGATCCCCTTAGCGGGAA
299880_2124_0294	CCCGACCACTCGGGGTGCTGGTATTACCGCGGCGGCTGACACCAGGCT TGCCCACCCTTATTCTGCACCATTTTTAGAGGTGCCAAAAGATTTTCCTT AGCGGAAATCAC
307308_2991_4050	GCAAGGGGTGTCAGCCGCCGCGGTAATACCAGCACCCCGAGTGGTCG GGACGTTTATTGGGCCCTAAAGCATCCGTAGCCGGTTCTACAAGTCTTC CGTTAAATCCACCTGCTTAACAGATGGG
CTO189	Number = 1
204487_2125_2018	GGAACGTGTCTTCAAGTGGGGGATAACGCACCGAAAGGTGTGCTAAT ACCGCATAATCTCTACGGAGAAAAGCAGGGGATCGCAAGACCTTGCG CTTTTGGAGCGGCCGA
CTO654	Number = 1
128302_3419_0725	CCCTCTGCTACACTCTAGCCTTGTAGTTTCAAACGCAGTTCCAGGTAA GCCCCGGGATTTACATCTGACTTACAAAACCGCCTGCGCACGCTC
Nso190	Number = 1
204487_2125_2018	GGAACGTGTCTTCAAGTGGGGGATAACGCACCGAAAGGTGTGCTAAT ACCGCATAATCTCTACGGAGAAAAGCAGGGGATCGCAAGACCTTGCG CTTTTGGAGCGGCCGA
Nso1225	Number = 1
081434_0164_2330	GCGCCATTGTATTACGTGTGAAGCCTACCCATAAGGGCCATGAGGACT TGACGC
Nsp436	Number = 3
280392_3122_3636	GAAGGCCTTCGGGTTGTAAAGCTCTTTCAGCCGGAACGAAACGGTAC ACGGCTAATACCGTGACAACCTGACGGTACCGGAAGAAGAAGCACCGG CTAACTACGTGCCAGCAGCCGCGGTAATACG
308604_0914_1313	GCTCTTTCGGCCGGAACGAAACGGCCGTGCCAATACCACGGGCTAAT GACGGTACCGGAGGAAGAAGCACCGGCTAACTACGTGCCAGCAGCCG CGGTAATACGTAGGGTGCAAGCGTTAA
224541_2557_3737	AAGCTCTTTCAGCCGGAACGAACGGTCACGGCTAATACCGTGACAAC GACGGTACGGAAGAAGAAGCACCGGCTAACTACGTGCCAGCAGCCGC GGTAATACGTAGGGTNCAG
AOA amoA genes	Number = 2
244261_0531_0298	GGGTTAGAGATTTAGTCCCATTACTCCAAGAGGCCATCCATCTGTAT GTCAAGTGTTAAGCTTCGCGCCTAATGCAGCTATAGGAACGCTTAGTA CTGCTCC
012110_0815_2177	GTGTCATGTATGGGGTAGTGTGGCCTAGGATACTTGAATGCCATTT TAGTGGATCACGTATGGTCAATAGGTTGATCATGTAAAC

* sequences that were identified through group-specific oligonucleotides or *amoA* genes, respectively (see Supplementary Methods). All were verified by individual BLASTn searches.