Table S3. Estimation of genome completeness based on 104 single-copy genes present in 5,449 complete bacterial genomes using CheckM (v0.9.7).

	Marker sets occurrence as: [a]							Est. genome size (Mbp)
Genome/SAGs	Singletons	Doubletons	Tripletons	% Completeness	% Contamination [b]	% Heterogeneity [c]	Assembly size (bp)	
Nitrospina gracilis	101	1	0	96.6	1.72	0	3,067,213	3.17
***Ca.* Nitromaritima sp. A02**	**18**	**0**	**0**	**19.8**	**0**	**0**	**1,404,272**	**2.53**
***Ca.* Nitromaritima sp. C22**	**26**	**0**	**0**	**39.7**	**0**	**0**	**1,678,166**	**2.69**
Ca. Nitromaritima sp. B06 (Current)	41	0	0	29.9	0	0	708,914	1.21
Ca. Nitromaritima sp. B18 (Current)	85	1	0	73.3	1.72	0	2,139,866	2.71
Ca. Nitromaritima sp. L16	83	0	0	75.2	0	0	2,077,614	2.59
Nitrobacter winogradskyi	104	0	0	100.0	0	0		
Nitrobacter hamburgensis	104	0	0	100.0	0	0		
Nitrococcus mobilis	103	0	0	98.3	0	0		
Ca. Nitrospira defluvii	102	0	0	97.4	0	0		
Ca. Kuenenia stuttgartensis	101	0	0	94.8	0	0		
Nitrolancetus hollandicus	95	1	0	91.9	0.86	0		

[a] More details on marker sets can be found in Parks et al. (2015).

[b] Determined by the number of single-copy genes that occur more than once.

[c] Determined from the number of multi-copy genes, which exceed a specified amino acids threshold (see methods in CheckM publication; Parks et al., 2015).