LETTERS

Deciphering the evolution and metabolism of an anammox bacterium from a community genome

Marc Strous[1], Eric Pelletier[2], Sophie Mangenot[2], Thomas Rattei[3], Angelika Lehner[4,5], Michael W. Taylor[4], Matthias Horn[4], Holger Daims[4], Delphine Bartol-Mavel[2], Patrick Wincker[2], Valérie Barbe[2], Nuria Fonknechten[2], David Vallenet[2], Béatrice Segurens[2], Chantal Schenowitz-Truong[2], Claudine Médigue[2], Astrid Collingro[4], Berend Snel[6], Bas E. Dutilh[6], Huub J. M. Op den Camp[1], Chris van der Drift[1], Irina Cirpus[7], Katinka T. van de Pas-Schoonen[1], Harry R. Harhangi[1], Laura van Niftrik[1,7], Markus Schmid[1], Jan Keltjens[1], Jack van de Vossenberg[1], Boran Kartal[1], Harald Meier[8], Dmitrij Frishman[3], Martijn A. Huynen[6], Hans-Werner Mewes[3,9], Jean Weissenbach[2], Mike S. M. Jetten[1,7], Michael Wagner[4] & Denis Le Paslier[2]

Anaerobic ammonium oxidation (anammox) has become a main focus in oceanography and wastewater treatment[1,2]. It is also the nitrogen cycle's major remaining biochemical enigma. Among its features, the occurrence of hydrazine as a free intermediate of catabolism[3,4], the biosynthesis of ladderane lipids[5,6] and the role of cytoplasm differentiation[7] are unique in biology. Here we use environmental genomics[8,9]—the reconstruction of genomic data directly from the environment—to assemble the genome of the uncultured anammox bacterium *Kuenenia stuttgartiensis*[10] from a complex bioreactor community. The genome data illuminate the evolutionary history of the Planctomycetes and allow us to expose the genetic blueprint of the organism's special properties. Most significantly, we identified candidate genes responsible for ladderane biosynthesis and biological hydrazine metabolism, and discovered unexpected metabolic versatility.

Since the biological nitrogen cycle was drafted at the end of the nineteenth century, the possibility of anaerobic ammonium oxidation (anammox) was generally overlooked[2]. A chance discovery just ten years ago led to the identification of the responsible chemolithoautotrophic bacteria and the appreciation of their applied and ecological significance[2,10]. Currently it is estimated that anammox contributes up to 50% to the removal of fixed nitrogen from the oceans globally[1]. Unfortunately, anammox bacteria divide only once in two weeks at maximum speed and they are not available in pure culture[2]. Environmental genomics[8,9] thus offers a unique possibility to gain insight into the metabolism and evolution of these important bacteria. A laboratory bioreactor, anoxic and fed with synthetic wastewater containing the substrates ammonium, nitrite and bicarbonate, was the starting point for the present study. The bioreactor had been operated in the laboratory for one year, corresponding to 10 to 15 generations of anammox bacteria. Fluorescence *in situ* hybridization with a suite of ribosomal-RNA-targeted probes showed that this reactor was dominated (73 ± 5%) by the anammox bacterium *K. stuttgartiensis*. No other known anammox bacterium was detectable[10]. Comparative analysis of all 16S rRNA gene sequences retrieved from the different libraries constructed for genome sequencing (see below) showed that the microbial community as a whole contained at least six recognized bacterial phyla and two lineages made up exclusively of uncultured bacteria (Supplementary Fig. 1). In total, 29 different operational taxonomic units were detected.

Genomic DNA was extracted from the bioreactor, and shotgun, fosmid and bacterial artificial chromosome (BAC) libraries were generated. Scaffolds were constructed from the shotgun library after binning, and gaps were closed with BAC and fosmid clones as well as polymerase chain reaction. Ultimately, five contigs were assembled (Supplementary Materials and Methods, Supplementary Fig. 2 and Supplementary Table 1). The five remaining gaps could not be closed and their size remains unknown. The robustness of the assembly was thoroughly validated with more than 100 confirmation points (Supplementary Fig. 2). The near completeness and correct assembly of the *K. stuttgartiensis* genome was further confirmed by the lack of suspicious redundancy or missing essential genes in major biosynthetic pathways. Apart from a gene for leucyl-transfer-RNA synthetase, the 64 clusters of orthologous groups of proteins (COGs) present in all currently sequenced bacterial genomes represented in the STRING database[11] were also present here. We therefore estimate the genome to be more than 98% (1 − 1/64) complete. Single nucleotide polymorphisms were virtually absent from the assembled scaffolds (less than 1 per 15,000 bases), strongly suggesting that this enrichment culture contained a single anammox strain.

The almost complete, 4.2-megabase genome of *K. stuttgartiensis* could then be used to deduce the biochemical pathway of anaerobic ammonium oxidation. Anammox is the one-to-one combination of ammonium and nitrite into dinitrogen gas. It was already known that hydrazine (N_2H_4) is an intermediate, and that the bacteria oxidize hydrazine to dinitrogen gas with a hydroxylamine-oxidoreductase-like protein (HAO)[2–4]. On the basis of this experimental evidence, a pathway for anammox catabolism was previously postulated and involves hydrazine and hydroxylamine (NH_2OH)[3]. In the *K. stuttgartiensis* genome we detected more than 200 genes relevant to anammox catabolism and respiration (Supplementary Table 2).

Two hundred genes directly involved in catabolism and respiration exceeds the number possessed by most other bacteria (Supplementary Fig. 3). So far, a comparable level of redundancy has only been observed for versatile heterotrophic bacteria such as *Geobacter*

[1]Department of Microbiology, IWWR, Radboud University Nijmegen, Toernooiveld 1, 6525 ED Nijmegen, the Netherlands. [2]CNRS UMR 8030 and Genoscope, 2 rue Gaston Crémieux CP 5706, 91057 Evry, France. [3]Department of Genome Oriented Bioinformatics, Technische Universität München, Am Forum 1, 85354 Freising, Germany. [4]Department of Microbial Ecology, University of Vienna, Althanstr. 14, 1090 Vienna, Austria. [5]Institut für Lebensmittelsicherheit und Hygiene, Universität Zürich, CH-8057 Zürich, Switzerland. [6]Center for Molecular and Biomolecular Informatics, Nijmegen Center for Molecular Life Sciences, Radboud University Nijmegen Medical Centre, Toernooiveld 1, 6525 ED Nijmegen, The Netherlands. [7]Kluyver Laboratory for Biotechnology, Delft University of Technology, 2628 BC Delft, The Netherlands. [8]Lehrstuhl für Rechnertechnik und Rechnerorganisation, Parallelrechnerarchitektur, Institut für Informatik, Technische Universität München, 85748 Garching, Germany. [9]Institute for Bioinformatics, GSF-National Research Center for Environment and Health, Ingolstädter Landstrasse 1, 85764 Neuherberg, Germany.

sulfurreducens and *Shewanella oneidensis*[12,13], while the aerobic ammonia oxidizer *Nitrosomonas europaea* has only 50 such genes[14]. Multiple divergent paralogues of respiratory complexes indicate the presence of a branched respiratory chain[15] which would enable the respiration of different energy sources with different electron acceptors. Triggered by the large number of encoded *c*-type cytochromes and the previous detection of many *c*-type cytochromes in the metal-respiring organisms *G. sulfurreducens* and *S. oneidensis*, we experimentally investigated iron and manganese respiration by *K. stuttgartiensis*. Indeed, both iron and manganese oxides were respired with formate as electron donor (Supplementary Table 3). Iron was also oxidized with nitrate as electron acceptor. Thus the present study confirms and extends the recently discovered versatility[16] of anammox bacteria.

Among the proteins encoded by the above-mentioned 200 genes were nine HAO-like proteins and two enzymes active in denitrification, nitrite::nitrate oxidoreductase (NarGH) and cd_1 nitrite::nitric oxide oxidoreductase (NirS). The presence of NirS was not compatible with the previously postulated model for anammox catabolism[3], but the new genomic data allowed the deduction of a new metabolic pathway for anammox involving NO as intermediate (Fig. 1a). This new pathway is the only possible one consistent with the available experimental data[2–4], is thermodynamically and biochemically feasible, and conforms to 'Ockham's razor'—that is, it invokes minimum biochemical novelty and requires the fewest number of biochemical reactions.

Anammox bacteria use the energy from their catabolism for autotrophic growth. Based on the isotopic composition of cell carbon, it was proposed that anammox bacteria make use of the acetyl-coenzyme A (CoA) pathway for carbon fixation[17]. In the *K. stuttgartiensis* genome we detected the complete acetyl-CoA pathway (Supplementary Table 4), while all other known carbon fixation pathways remain either missing or incomplete. The acetyl-CoA pathway depends on electrons at very low redox potential (−0.41 V), usually derived from the oxidation of molecular hydrogen[18]. Interestingly, anammox bacteria derive their electrons from the anaerobic oxidation of nitrite to nitrate (+0.43 V) and their use of the acetyl-CoA pathway is not straightforward. To show that the acetyl-CoA pathway is really operational in anammox bacteria, we experimentally demonstrated the activity of two key enzymes (formate dehydrogenase and carbon monoxide dehydrogenase) in cell-free extracts of *K. stuttgartiensis* (Supplementary Table 5). With the genomic data we subsequently deduced a biochemical pathway that explains how the acetyl-CoA pathway can be reconciled with nitrite as electron donor for carbon fixation (Fig. 1b). In this pathway, high energy electrons from hydrazine are transferred via ferredoxin to the acetyl-CoA synthetase/CO dehydrogenase and the replenishment of the hydrazine pool to compensate for the hydrazine invested in carbon fixation requires no additional enzymes except reverse electron transport.

In the pathways of Fig. 1 there are two enzymes unique to anammox bacteria (dark blue circles and square). Hydrazine hydrolase (hh) is the enzyme which produces hydrazine from nitric

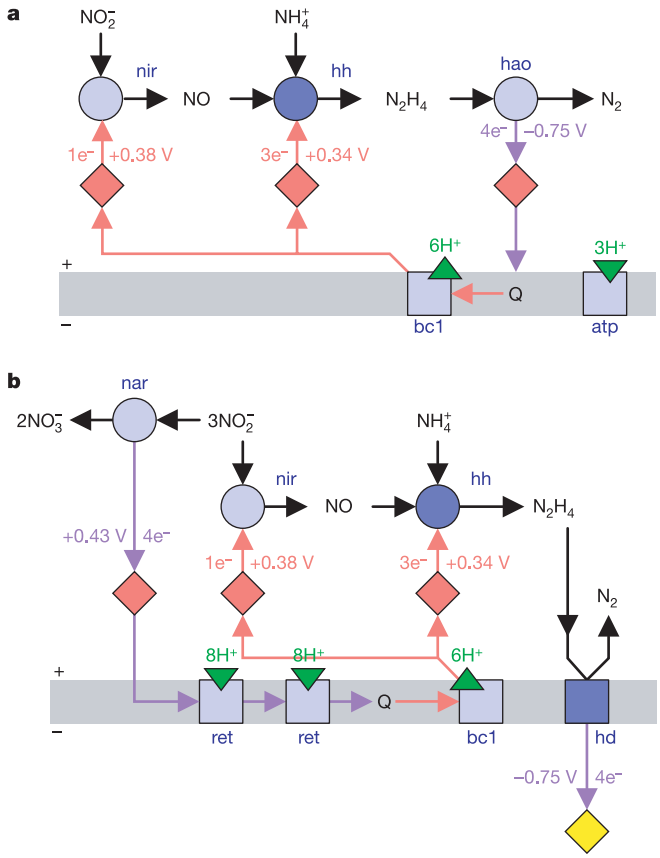

Figure 1 | Metabolic pathways of *K. stuttgartiensis* inferred from the present genomic data, previous experimental evidence and thermodynamic calculations. **a**, Anammox central catabolism with nitric oxide as intermediate, electron transport and energy conservation; **b**, combination of central catabolism with nitrate reductase to generate high potential electrons for the acetyl-CoA pathway. Three-letter abbreviations below enzymes refer to the genes and operons in Supplementary Tables 1 and 2. The dark blue 'hh' and 'hd' gene clusters are shown in Fig. 2. Red diamonds, cytochromes; yellow diamond, ferredoxin; red arrows, reductions; purple arrows, oxidations.

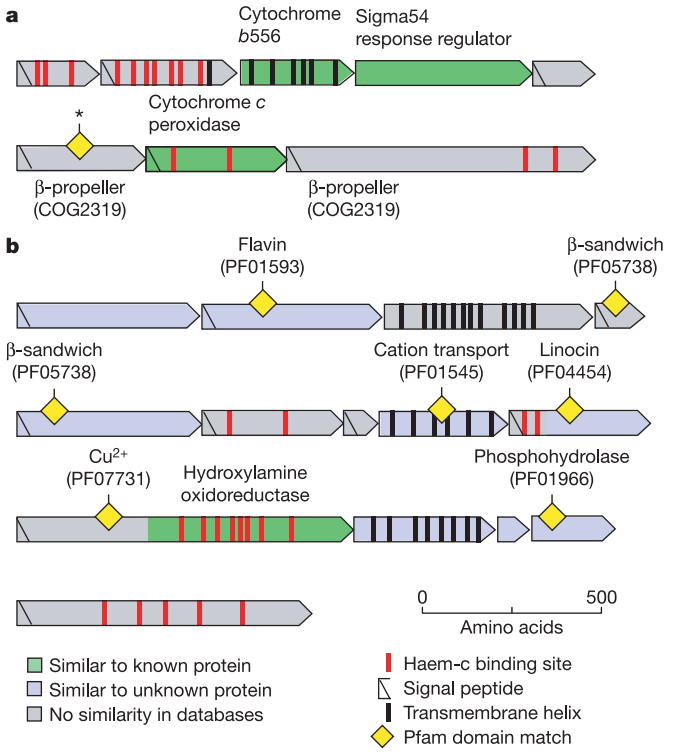

Figure 2 | The operons representing the best candidates to encode hydrazine hydrolase (hh) and hydrazine dehydrogenase (hd). **a**, Open reading frames (ORFs) *kuste2854–kuste2861* encoding a β-propeller complex (such as nitrous oxide reductase) consisting of several cytochromes. The asterisk indicates a novel cofactor binding site with weak homology to cytochrome *d* and quinoheme binding sites. **b**, ORFs *kuste2469–kuste2483* encode a new multicopper oxidase (such as the nitrite reductase NirK), a flavin containing amine oxidase and several integral membrane proteins.

oxide and ammonium, and hydrazine dehydrogenase (hd) is the enzyme which transfers the electrons from hydrazine to ferredoxin. Interestingly, several new genes containing domains involved in electron transfer and catalysis were detected. The latter genes were mainly present in two gene clusters, which apparently code for two completely novel enzyme complexes. These are the most likely candidates to catalyse these novel steps (Fig. 2).

With a doubling time of two weeks, anammox metabolism according to Fig. 1 proceeds very slowly, and passive diffusion of protons and reactive intermediates (that is, hydrazine) over membranes is more of a problem at low enzymatic turnover. Therefore, it was proposed that the anammox bacteria have evolved ladderane lipids as the major component of their biomembranes[5]. These lipids impart unusual density and impermeability to the membrane because of their structural rigidity and size. The mode of biosynthesis is totally unknown and unprecedented because of the structural novelty and high ring strain[6]. Interestingly, in the *K. stuttgartiensis* genome fatty acid biosynthesis is represented by four gene clusters, two of which consist of a conspicuous combination of homologues of known fatty acid biosynthesis and *S*-adenosylmethionine (SAM) radical enzyme genes. The largest of these clusters is shown in Fig. 3. Radical SAM enzyme-encoding genes are rare in bacterial genomes and these enzymes are frequently reserved for the most difficult chemical reactions[19]. Although the ladderane biosynthetic pathway could not be immediately inferred, the above-mentioned gene clusters are the most promising candidates to encode this unique pathway. The encoded proteins suggest the involvement of methylation, cyclization via oxidative radical chemistry, and addition of an aromatic residue, combined with regular fatty acid elongation.

Considering the many unique adaptations, it was no surprise that previous studies based on rRNA genes indicated that anammox bacteria evolved from a common ancestor. This ancestor arose deep within the phylum Planctomycetes[2,20]. Recently, the evolutionary position of this phylum within the tree of life has been the subject of intensive debate. The Planctomycetes are perceived as either the deepest branching bacterial phylum which forms a missing link between eukaryotes and prokaryotes[20,21] or as a 'normal' bacterial phylum related to the intracellular parasites of the Chlamydiae[22]. Both possibilities are supported by several genomic and cell-biological features[21–25]. With the genome of *K. stuttgartiensis* and the genomic data of the planctomycete *Gemmata obscuriglobus* (http://www.tigr.org) it became possible to tackle this issue. We assembled a dataset of 49 concatenated amino acid sequences (Supplementary Table 6) from 90 representative bacterial genomes comprising 7,568 slowly evolving positions. A concatenated 5S-16S-23S rRNA dataset with 3,846 conserved nucleotide positions was also constructed. Phylogenetic analysis strongly supported the evolutionary grouping of Planctomycetes and Chlamydiae (Fig. 4), whereas an early divergence of the Planctomycetes from the other bacteria was not apparent, even when the trees were rooted with archaeal sequences (Supplementary Figs 4 and 5).

Next, we searched the planctomycete genomes for orthologous groups shared specifically with eukarya and other bacterial phyla. Two deduced *K. stuttgartiensis* proteins showed low similarity with eukaryal signature proteins, much less than the 17 and 10 proteins observed for *Gemmata* sp. Wa-1 and *Prosthecobacter dejongeii*, respectively[25]. This finding suggests that gene transfer occurred between planctomycetes and eukaryotes after the divergence of the anammox bacteria. The planctomycetes, however, shared 17 orthologous groups with Chlamydiae, more than with any other bacterial phylum (Supplementary Table 7). Notably, all planctomycetes possess a gene encoding a protein highly similar to the well-known 60-kDa cysteine-rich outer membrane protein of Chlamydiae[23], which does not occur in any other prokaryote for which a genome sequence is available. Thus it represents a unique characteristic of the Planctomycetes-Chlamydiae superphylum. Members of this superphylum are also characterized by peptidoglycan-less cell envelopes[21,23] even though the Chlamydiae are sensitive to penicillin and their genomes encode the complete peptidoglycan biosynthetic pathway except for penicillin-binding proteins 1 and 2 (Supplementary

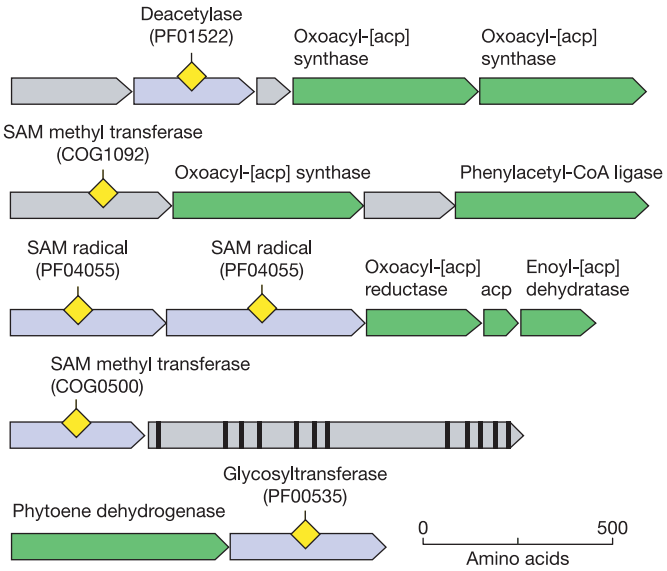

Figure 3 | The largest of four operons encoding fatty acid biosynthesis in *K. stuttgartiensis* (ORFs *kuste3352–kuste3335*). The presence of several hypothetical *S*-adenosylmethionine (SAM) radical and methylase proteins, highly unusual in the context of fatty acid biosynthesis, make this operon the best candidate to encode the ladderane lipid biosynthetic pathway. acp, acyl carrier protein.

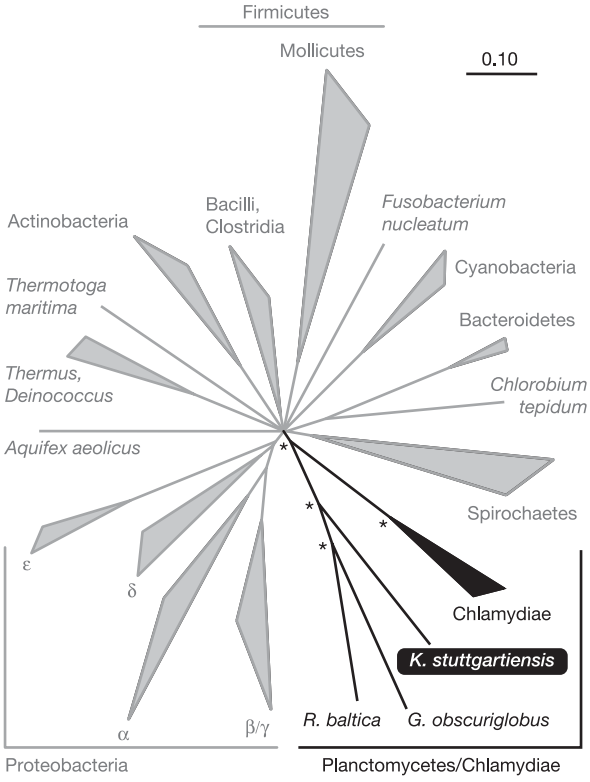

Figure 4 | Unrooted phylogenetic consensus tree based on 49 concatenated protein sequences, showing the phylogenetic positioning of *K. stuttgartiensis* within the Bacteria. Asterisks at nodes represent 100% distance- and maximum parsimony-based bootstrap support. Scale bar represents 10% estimated sequence divergence.

Table 8). In contrast to the planctomycete *Rhodopirellula baltica*, which lacks almost all peptidoglycan synthesis genes[24], *K. stuttgartiensis* encodes 19 out of the 21 genes required for peptidoglycan biosynthesis. This situation resembles that of the Chlamydiae, and this suggests that the common ancestor of the Planctomycetes-Chlamydiae superphylum evolved from a bacterium that was still able to synthesize a 'normal' Gram-negative cell wall.

Not so long ago, anammox bacteria were considered to be slowly growing, highly specialized chemolithoautotrophs with minor ecological importance. It has now become clear that although slow and specialized, they are global players in the biological nitrogen cycle[1,2]. With the genome of *K. stuttgartiensis*, reconstructed from the microbial community of a laboratory bioreactor, we begin to realize that anammox bacteria could actually be generalists, not specialists; even though anammox metabolism requires distinct adaptations, it appears to be compatible with a versatile lifestyle—a lifestyle that may link the biological nitrogen cycle to the carbon and metal cycles in new ways.

METHODS

Laboratory bioreactor. The anammox bioreactor consisted of an anoxic 10-litre gas-lift reactor fed with synthetic wastewater, and was inoculated with sludge from the nitrification stage of the Dokhaven wastewater treatment plant (Rotterdam, The Netherlands). At the time of sampling (one year after inoculation), *K. stuttgartiensis* made up 73% of the microbial community[2,26].

DNA preparation. Bioreactor samples were immobilized in agarose plugs and DNA was extracted after incubation (3 h, 37 °C) in EDTA (100 mM), NaCl (50 mM), Tris-HCl (10 mM), Na-lauroyl sarcosine (0.5%), lysozyme (1 mg ml^{-1}), mutanolysine (10 U ml^{-1}), lipase (1 mg ml^{-1}), peptidase (1 mg ml^{-1}) and β-glucuronidase (1 mg ml^{-1}) pH 8. The plugs were then incubated twice for 24 h at 50 °C in EDTA (0.5 M), Na-lauroyl sarcosine (0.5%) pH 8 containing protease (500 ng ml^{-1}) and proteinase K (2 mg ml^{-1}).

Library construction. A BAC library (8,448 clones) was constructed from partial *Hind*III digests[27]. Fosmid (6,432 clones) and shotgun-randomly-sheared-DNA-plasmid libraries were constructed with pEpiFOS5 (Epicentre) and pCDNA2 vectors respectively. Clones from all libraries were picked and bidirectionally sequenced using standard protocols.

Assembly. In a preliminary global assembly (192,713 shotgun sequence reads, performed with Phrap) a great number of small contigs were produced. After the removal of all sequences harbouring repeated motifs, a second global assembly (127,557 reads) resulted in 9,787 contigs. By BAC and fosmid bridging, 493 of these contigs were assembled into 68 supercontigs. Three genes indicative of anammox organisms (16S rDNA, a cytochrome *c* and hydroxylamine oxidoreductase) were detected on three of these supercontigs. The subsequent assignment of 17 other supercontigs to *K. stuttgartiensis* was guided by the GC content (41%) and a high percentage of BAC-end sequences (>5%, see Supplementary Methods) characteristic for these 3 supercontigs. After a second round of BAC and fosmid bridging, 5 final contigs (totalling 4,218,325 nucleotides, EMBL accession numbers CT030148, CT573071–CT573074) were finally obtained (Supplementary Fig. 2). The mean read coverage over the 5 contigs was 22×.

Confirmation of assembly. The robustness of the assembly was supported by the phylogenetic analysis of 33 conserved marker genes. In neighbour-joining or maximum parsimony trees, these markers clustered with the planctomycetes *R. baltica* or *G. obscuriglobus*. Further, 73 out of 88 shotgun sequences obtained from a 99.5% pure cell preparation of the related anammox bacterium *Brocadia anammoxidans* were most similar to sequences in the *K. stuttgartiensis* assembly when compared against a database of 180 complete bacterial genomes (including *K. stuttgartiensis*) using the program Blast.

Genome annotation. Prediction of coding sequences was performed with the programs dps/orpheus, AMIGene, glimmer/rbsfinder and genemarks/genemark.hmm. Annotation was performed in PEDANT[28]. Genes with sequence homology (>20% identical) to genes with established functions (that is, by knockout/complementation experiments, structural analysis) were annotated as '(strongly) similar to'. Genes with sequence homology to uncharacterized genes were annotated as '(conserved) hypothetical protein'. Genes without significant homology (<20% identical) were annotated as 'unknown'. tRNA genes were identified with the tRNAscan-SE program and rRNA genes were located by homology. Genome synteny against other complete genomes was computed with LAGAN.

Phylogenetic analysis. Phylogenetic analysis of *K. stuttgartiensis* was based on concatenated data sets of protein and rRNA sequences. The data set was based on 98 representative (<97% similarity of 16S rDNA genes) genomes plus *K. stuttgartiensis* and *G. obscuriglobus*, a planctomycete currently being sequenced at The Institute for Genomic Research (TIGR). Sequences were extracted for 44 ribosomal proteins, 3 DNA-directed RNA polymerase subunits and 3 other proteins (GyrB, RecA and EF-TU) commonly used for bacterial phylogeny (Supplementary Table 6). For proteins, distance (FITCH), maximum parsimony (PROTPARS), and maximum likelihood (MOLPHY) analyses were used to construct consensus trees from concatenated alignments (CLUSTAL, default settings) with a positional conservation filter of 30%. For ribosomal RNA (5S, 16S and 23S), neighbour-joining, maximum parsimony and maximum likelihood analysis was performed in ARB after alignment with ARB-FastAligner followed by manual refinement, concatenation and application of a 50% conservation filter. Bootstrapping (100 resamplings) was performed with PHYLIP 3.61. Rooting of ribosomal protein trees with members of another domain is often not attempted owing to alignment difficulties and the risk of biases due to long-branch attraction. Here, to test whether the current data set suggests deep-branching of the Planctomycetes within the Bacteria tree, we included ribosomal protein sequences from five archaea according to the alignments of Vishwanath *et al.*[29]. After concatenation and application of conservation filters, phylogenetic analyses were performed as described above and with PHYML[30] for maximum-likelihood-based phylogeny with the following evolutionary models: Dayhoff, JTT, MtRev and WAG. For each PHYML inference topology optimization was applied on a BIONJ starting tree, the proportion of invariable sites and the gamma distribution parameter were estimated, and 16 substitution rate categories were applied. These calculations were performed on Quad AMD Opteron 2.4 GHz computing nodes, equipped with 8 GB shared main memory. Methods are described and referenced in detail in the Supplementary Information.

Received 22 December 2005; accepted 15 February 2006.

1. Arrigo, R. A. Marine microorganisms and global nutrient cycles. *Nature* **437,** 349–355 (2005).
2. Strous, M. & Jetten, M. S. M. Anaerobic oxidation of methane and ammonium. *Annu. Rev. Microbiol.* **58,** 99–117 (2004).
3. Van de Graaf, A. A., De Bruijn, P., Robertson, L. A., Jetten, M. S. M. & Kuenen, J. G. Metabolic pathway of anaerobic ammonium oxidation on the basis of N-15 studies in a fluidized bed reactor. *Microbiology (UK)* **143,** 2415–2421 (1997).
4. Schalk, J., de Vries, S., Kuenen, J. G. & Jetten, M. S. M. A novel hydroxylamine oxidoreductase involved in the anammox process. *Biochemistry* **39,** 5405–5412 (2000).
5. Sinninghe Damsté, J. S. *et al.* Linearly concatenated cyclobutane lipids form a dense bacterial membrane. *Nature* **419,** 708–712 (2002).
6. Macitti, V. & Corey, E. J. Total synthesis of pentacyclic anammoxic acid. *J. Am. Chem. Soc.* **126,** 15664–15665 (2004).
7. Lindsay, M. R. *et al.* Cell compartmentalization in planctomycetes: novel types of structural organization for the bacterial cell. *Arch. Microbiol.* **175,** 413–429 (2001).
8. DeLong, E. F. Microbial community genomics in the ocean. *Nature Rev. Microbiol.* **3,** 459–469 (2005).
9. Tringe, S. G. *et al.* Comparative metagenomics of microbial communities. *Science* **308,** 554–557 (2005).
10. Schmid, M. C. *et al.* Molecular evidence for genus level diversity of bacteria capable of catalyzing anaerobic ammonium oxidation. *Syst. Appl. Microbiol.* **23,** 93–106 (2000).
11. Von Mering, C. *et al.* STRING: known and predicted protein–protein associations, integrated and transferred across organisms. *Nucleic Acids Res.* **33,** D433–D437 (2005).
12. Methe, B. A. *et al.* Genome of *Geobacter sulfurreducens:* Metal reduction in subsurface environments. *Science* **302,** 1967–1969 (2003).
13. Heidelberg, J. F. *et al.* Genome sequence of the dissimilatory metal ion-reducing bacterium *Shewanella oneidensis*. *Nature Biotechnol.* **20,** 1118–1123 (2002).
14. Chain, P. *et al.* Complete genome sequence of the ammonia-oxidizing bacterium and obligate chemolithoautotroph *Nitrosomonas europaea*. *J. Bacteriol.* **185,** 2759–2773 (2003).
15. Brasseura, G. *et al.* Apparent redundancy of electron transfer pathways via bc1 complexes and terminal oxidases in the extremophilic chemolithoautotrophic *Acidithiobacillus ferrooxidans*. *Biochim. Biophys. Acta* **1656,** 114–126 (2004).
16. Güven, D. *et al.* Propionate oxidation by and methanol inhibition of anaerobic ammonium-oxidizing bacteria. *Appl. Environ. Microbiol.* **71,** 1066–1071 (2005).
17. Schouten, S. *et al.* Stable carbon isotopic fractionations associated with inorganic carbon fixation by anaerobic ammonium oxidizing bacteria. *Appl. Environ. Microbiol.* **40,** 3785–3788 (2004).
18. Drake, H. L. & Daniel, S. L. Physiology of the thermophilic acetogen *Moorella thermoacetica*. *Res. Microbiol.* **155,** 422–436 (2004).
19. Layer, G., Heinz, D. W., Jahn, D. & Schubert, W. D. Structure and function of radical SAM enzymes. *Curr. Opin. Chem. Biol.* **8,** 468–476 (2004).

20. Brochier, C. & Philippe, H. Phylogeny—A non hyperthermophilic ancestor for bacteria. *Nature* **417,** 244 (2002).
21. Fuerst, J. A. Intracellular compartmentation in planctomycetes. *Annu. Rev. Microbiol.* **59,** 299–328 (2005).
22. Teeling, H., Lombardot, T., Bauer, M., Ludwig, W. & Glöckner, F. O. Evaluation of the phylogenetic position of the planctomycete '*Rhodopirellula baltica*' SH 1 by means of concatenated ribosomal protein sequences, DNA-directed RNA polymerase subunit sequences and whole genome trees. *Int. J. Syst. Evol. Microbiol.* **54,** 791–801 (2004).
23. Chopra, I., Storey, C., Falla, T. J. & Pearce, J. H. Antibiotics, peptidoglycan synthesis and genomics: the chlamydial anomaly revisited. *Microbiology (UK)* **144,** 2673–2678 (1998).
24. Glöckner, F. O. *et al.* Complete genome sequence of the marine planctomycete *Pirellula* sp. Strain 1. *Proc. Natl Acad. Sci. USA* **100,** 8298–8303 (2003).
25. Staley, J. T., Bouzek, H. & Jenkins, C. Eukaryotic signature proteins of *Prosthecobacter dejongeii* and *Gemmata* sp Wa-1 as revealed by *in silico* analysis. *FEMS Microbiol. Lett.* **243,** 9–14 (2005).
26. Daims, H., Lücker, S. & Wagner, M. DAIME, a novel image analysis program for microbial ecology and biofilm research. *Environ. Microbiol.* **8,** 200–213 (2006).
27. Le Paslier, M. C. *et al.* Construction and characterization of a *Schistosoma mansoni* bacterial artificial chromosome library. *Genomics* **65,** 87–94 (2000).
28. Frishman, D. *et al.* Functional and structural genomics using PEDANT. *Bioinformatics* **17,** 44–57 (2001).
29. Vishwanath, P., Favaretto, P., Hartman, H., Mohr, S. C. & Smith, T. F. Ribosomal protein-sequence block structure suggests complex prokaryotic evolution with implications for the origin of eukaryotes. *Mol. Phylogenet. Evol.* **33,** 615–625 (2004).
30. Guindon, S. & Gascuel, O. A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. *Syst. Biol.* **52,** 696–704 (2003).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements Preliminary sequence data from *G. obscuriglobus* was obtained from The Institute for Genomic Research through the website at http://www.tigr.org. Sequencing of *G. obscuriglobus* was accomplished with support from the United States Department of Energy. A. Cabezas is acknowledged for technical assistance. M.H., A.C. and M.W. were supported by the Austrian Ministry for Science and Education. M.T. was funded by the German Federal Ministry for Education and Research (Biolog II). H.D. was supported by the Wiener Wissenschafts-, Forschungs- und Technologiefonds (WWTF). We thank B. Anneser, F. Maixner, M. Leitner, A. Pol and B. Meijerink for help in the first round annotation and M. Ott for his effort in high performance phylogenetic treeing. The Dutch anammox research was supported by the Netherlands Science Foundation for Earth and Life Science (ALW), the Foundation for Applied Research (STW) and the European Union. The Radboud University and S. Wendelaar Bonga financially supported the initial phase of the project.

Author Information Reprints and permissions information is available at npg.nature.com/reprintsandpermissions. The authors declare no competing financial interests. Correspondence and requests for materials should be addressed to M.S.M.J. (m.jetten@science.ru.nl) or M.W. (wagner@microbial-ecology.net).