

ARTICLE

<https://doi.org/10.1038/s41467-019-13499-2>

OPEN

# Phagocytosis-like cell engulfment by a planctomycete bacterium

Takashi Shiratori <sup>1,2,4\*</sup>, Shigekatsu Suzuki <sup>3,4</sup>, Yukako Kakizawa<sup>1</sup> & Ken-ichiro Ishida<sup>1</sup>

Phagocytosis is a key eukaryotic feature, conserved from unicellular protists to animals, that enabled eukaryotes to feed on other organisms. It could also be a driving force behind endosymbiosis, a process by which  $\alpha$ -proteobacteria and cyanobacteria evolved into mitochondria and plastids, respectively. Here we describe a planctomycete bacterium, ‘*Candidatus Uab amorphum*’, which is able to engulf other bacteria and small eukaryotic cells through a phagocytosis-like mechanism. Observations via light and electron microscopy suggest that this bacterium digests prey cells in specific compartments. With the possible exception of a gene encoding an actin-like protein, analysis of the ‘*Ca. Uab amorphum*’ genomic sequence does not reveal any genes homologous to eukaryotic phagocytosis genes, suggesting that cell engulfment in this microorganism is probably not homologous to eukaryotic phagocytosis. The discovery of this “phagotrophic” bacterium expands our understanding of the cellular complexity of prokaryotes, and may be relevant to the origin of eukaryotic cells.

<sup>1</sup>Faculty of Life and Environmental Sciences, University of Tsukuba, Tsukuba, Ibaraki 305-0053, Japan. <sup>2</sup>Japan Agency for Marine-Earth Science and Technology (JAMSTEC), Yokosuka, Kanagawa 237-0061, Japan. <sup>3</sup>National Institute for Environmental Studies, Tsukuba, Ibaraki 305-0053, Japan. <sup>4</sup>These authors contributed equally: Takashi Shiratori, Shigekatsu Suzuki. \*email: [tshiratori@jamstec.go.jp](mailto:tshiratori@jamstec.go.jp)

There are significant differences between eukaryotes and prokaryotes in terms of cellular and genomic complexity<sup>1,2</sup>. During the emergence of ancestral eukaryotes (eukaryogenesis), cells increased in size and developed an endomembrane system, a nucleus and other membrane-bounded organelles. They also developed actin- and tubulin-based cytoskeletons, which facilitated uptake of large particles from the environment (phagocytosis)<sup>3</sup>. Phagocytosis not only supports efficient nutrient acquisition via consumption of other organisms (phagotrophy), but is also considered by some eukaryogenesis theories to be involved in the acquisition of the mitochondrion<sup>4–6</sup>. However, it remains unclear how the ancestral eukaryote acquired the phagocytosis ability since no known present organism exhibits a primitive phagocytosis.

Recent metagenomic studies have identified an archaeal group, Asgard, most closely related to Eukarya<sup>7,8</sup>. Asgard archaea possess genes that are homologous to eukaryotic genes involved in cytoskeleton formation, vesicle/membrane trafficking or remodelling and phagocytosis<sup>7,8</sup>. Recent cultivation and physiological studies of an Asgard archaeon, ‘*Candidatus Prometheoarchaeum syntrophicum*’, have shown that this organism is syntrophic with bacteria and occasionally produces long, branching protrusions<sup>9</sup>. However, it has not been reported that ‘*Ca. Prometheoarchaeum syntrophicum*’ displays any phagocytic-like behaviour.

Here, we describe a planctomycete bacterium, ‘*Candidatus Uab amorphum*’ (named after a hungry giant in Palauan mythology, called Uab) that exhibits several intriguing eukaryotic-like features, including the ability to engulf other microorganisms (bacteria and eukaryotes) through a phagocytosis-like process.

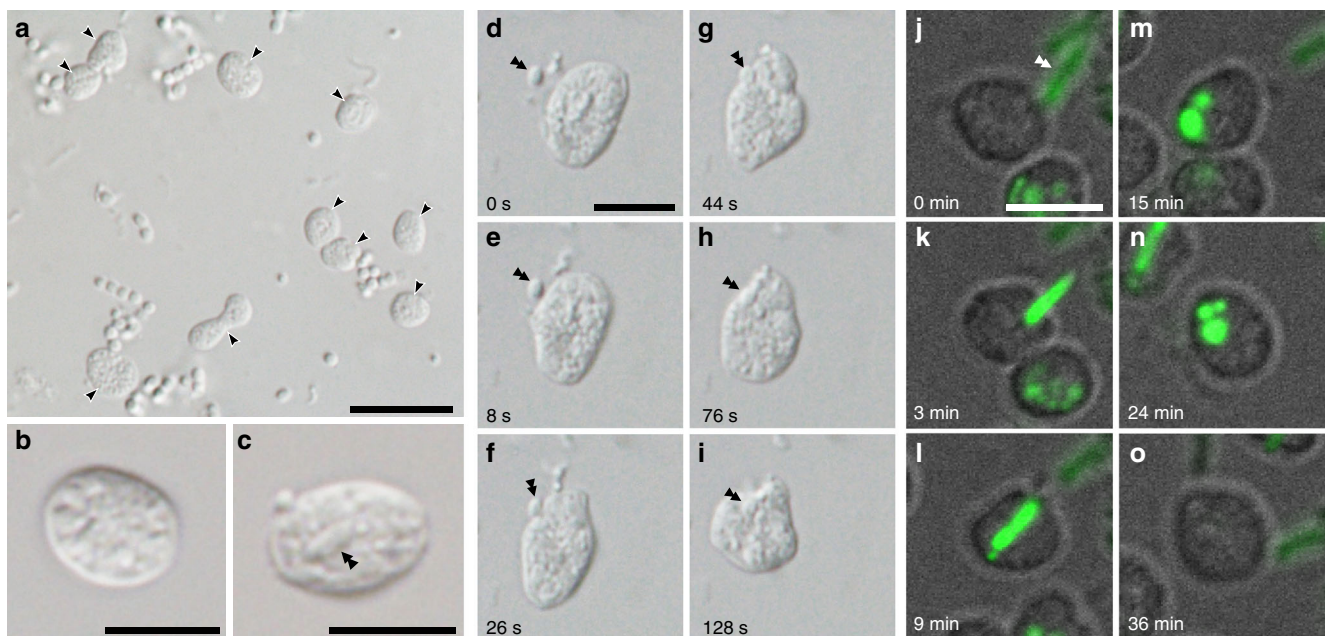
## Results

**Characteristics of ‘*Ca. Uab amorphum*’.** ‘*Ca. Uab amorphum*’ was discovered from a surface seawater sample collected at the Republic of Palau. It was maintained as an enriched or as a monoxenic culture with *Alteromonas macleodii* (NBRC 102226).

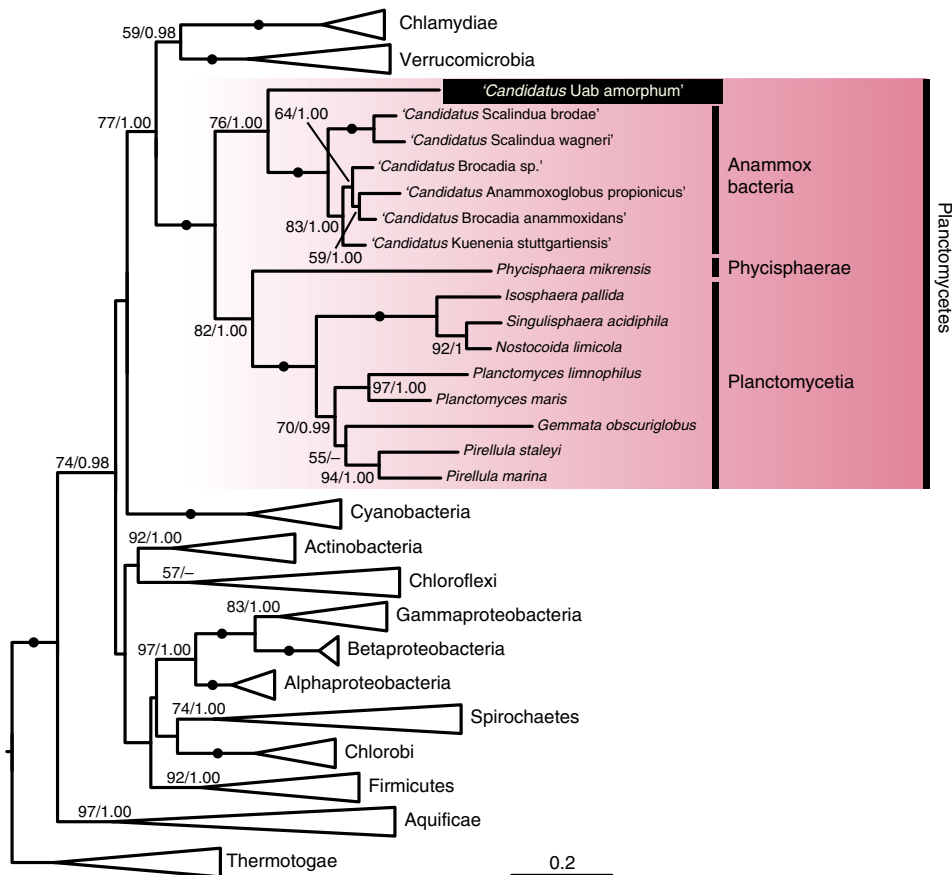
Our attempts to establish axenic cultures were unsuccessful. We found that ‘*Ca. Uab amorphum*’ has a flat, round or oval cell shape that is ~4–5 μm in diameter (Fig. 1a–c). However, cells occasionally reached 10 μm in diameter when cultures included many co-cultured bacterial cells (Supplementary Fig. 1a). ‘*Ca. Uab amorphum*’ moved on solid substrates at approximately 8–16 μm/min while frequently changing cell shape (Supplementary Movie 1). In addition, ‘*Ca. Uab amorphum*’ engulfed other bacteria in cultures (Fig. 1d, e, Supplementary Movie 2); the species description is included in the ‘Methods’ section.

Amplification of ribosomal RNA (rRNA) using PCR and fluorescent in situ hybridization (FISH) with specific oligonucleotide probes for ‘*Ca. Uab amorphum*’ revealed that this organism was not a eukaryotic amoeba, but a bacterium whose 16S rRNA gene sequence similarity with the closest known species was only 79% (Supplementary Fig. 2, Supplementary Table 1). Genome sequencing of the monoxenic culture produced a circular genome that encoded the 16S rRNA of ‘*Ca. Uab amorphum*’; the remaining sequences were small portions of the *A. macleodii* genomes (Supplementary Table 2). The ‘*Ca. Uab amorphum*’ genome was AT rich, 9.5 Mb in size and encoded 6660 protein-coding genes without plasmids (Supplementary Table 3).

To identify the phylogenetic position of ‘*Ca. Uab amorphum*’, we performed molecular phylogenetic analyses with two datasets, one using 16S rRNA gene sequences and the other using 171 proteins. Both analyses indicated that ‘*Ca. Uab amorphum*’ belongs to the bacterial phylum Planctomycetes (Fig. 2, Supplementary Fig. 3). ‘*Ca. Uab amorphum*’ was genetically separated from three subgroups of Planctomycetes: Planctomycetia, Phycisphaerae and anammox (anaerobic ammonium oxidation) bacteria<sup>10</sup>, with anammox bacteria being its closest relative. We found that various environmental rRNA gene sequences formed a clade with ‘*Ca. Uab amorphum*’ (Supplementary Fig. 4). This “Uab clade” included sequences derived from freshwater, hypersaline and marine environments, as well as the bovine rumen. The similarity between ‘*Ca. Uab amorphum*’ and these



**Fig. 1** Light and fluorescent micrographs of ‘*Candidatus Uab amorphum*’. **a** Cells of ‘*Ca. Uab amorphum*’ (arrowheads) in the xenic culture. **b, c** A cell of ‘*Ca. Uab amorphum*’. Double arrowheads indicate an engulfed bacterium. **d–i** Selected images of time-lapse video showing prey engulfment process of ‘*Ca. Uab amorphum*’. **j–o** Selected images of confocal fluorescent time-lapse video showing prey engulfment and digestion process of ‘*Ca. Uab amorphum*’. Green fluorescence indicates AcGFP1-labelled *Escherichia coli*. Scale bars, 10 μm (**a**) and 5 μm (**b–o**).



**Fig. 2** Maximum likelihood tree of ‘*Ca. Uab amorphum*’ and other bacterial 16S rRNA gene sequences. Left and right values on nodes indicate bootstrap value and Bayesian posterior probability, respectively. Scale bar indicates nucleotide substitution rate per site.

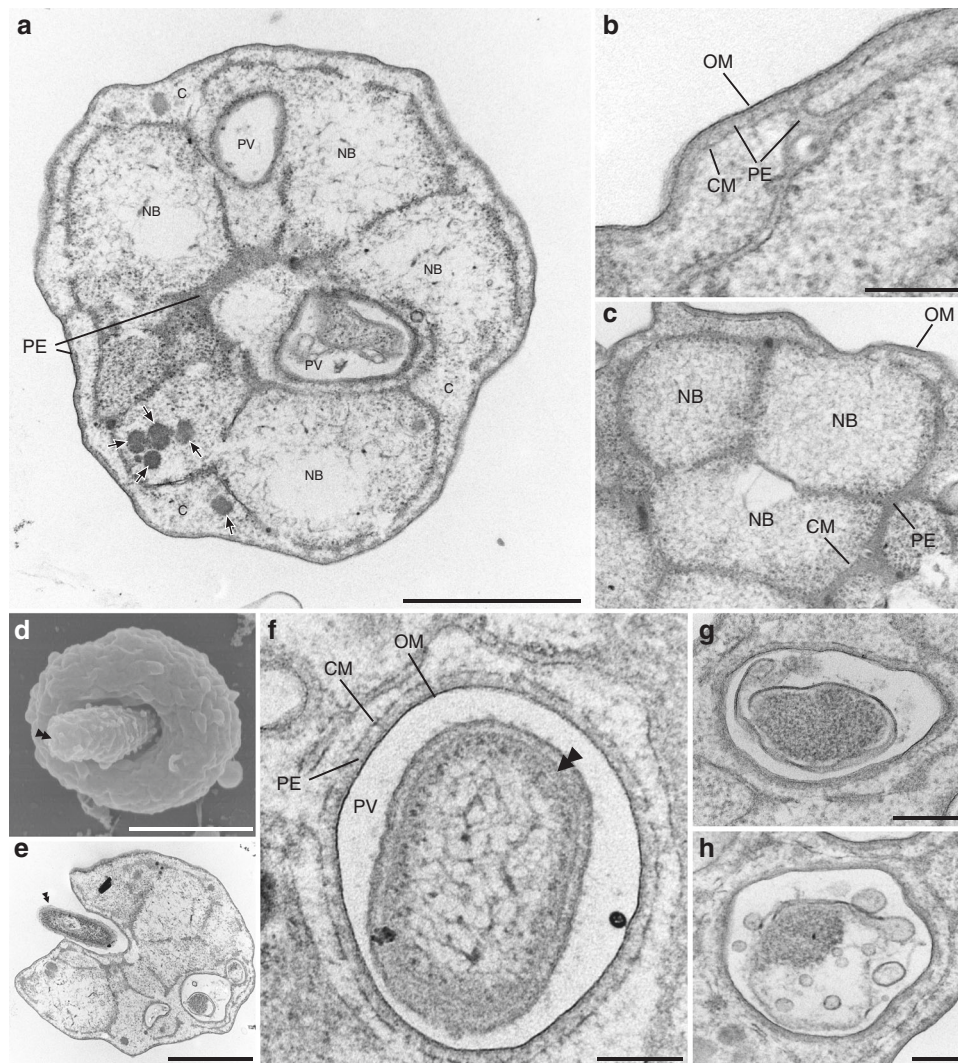
environmental sequences varied between 79.4% (KC604713) and 98.4% (KU631341). These data suggest that relatives of ‘*Ca. Uab amorphum*’ live in various environments and have high genetic diversity. We also screened 16S rRNA data from the Tara Oceans expedition<sup>11</sup> and found sequences related to the “Uab clade” in datasets derived from stations in the North and South Atlantic Oceans, the Indian Ocean, the North and South Pacific Ocean, the Red Sea and the Southern Ocean (Supplementary Fig. 5).

Transmission electron microscopy (TEM) demonstrated that ‘*Ca. Uab amorphum*’ has two membrane-bounded regions: an electron-dense ribosome-free region and a ribosome-containing region (Fig. 3a–c). The ribosome-free region located immediately beneath the outermost membrane and was highly invaginated into the ribosome-containing region, which showed a reticulate pattern in thin-section micrographs. The ribosome-containing region had multiple nuclear bodies that appeared to be separated by the invaginated ribosome-free region. This cellular structure consisting of ribosome-free and ribosome-containing regions with complex invaginations is typical of some planctomycete bacteria such as *Gemmata obscuriglobus*<sup>12</sup>. On the basis of electron microscopic observation and distribution of RNA<sup>13</sup>, it has been previously proposed that the outermost membrane surrounding the ribosome-free region is the cytoplasmic membrane, and the membrane surrounding the ribosome-containing region is an intracytoplasmic membrane. However, recent genomic and cell biological studies indicate that the cellular structure of planctomycetes resembles that of typical Gram-negative bacteria, with the outermost and inner membranes being equivalent to the outer membrane and cytoplasmic membrane, and the ribosome-free and ribosome-containing regions corresponding to the periplasm and cytoplasm,

respectively<sup>14,15</sup>. We found that the genome of ‘*Ca. Uab amorphum*’ includes genes encoding proteins similar to Gram-negative bacterial type II secretion systems (T2SS) (Supplementary Table 4), which would be consistent with a Gram-negative bacterial cell plan.

The genome of ‘*Ca. Uab amorphum*’ possesses most of the genes required for peptidoglycan biosynthesis (Supplementary Table 5), but the presence and properties of a peptidoglycan layer remain untested. Considering its cell flexibility, this organism is unlikely to have a rigid peptidoglycan layer. Synthesis and degradation of the peptidoglycan layer are known to be required for engulfment of the daughter cell during spore formation in other bacterial species<sup>16,17</sup>.

In comparison with other planctomycetes, ‘*Ca. Uab amorphum*’ exhibited a relatively large and gene-rich genome (Supplementary Table 3), with many genes involved in signal transduction (Supplementary Fig. 6). In addition, we identified 1515 genes of unknown function, and 1823 genes (27.6% of all genes) potentially derived from horizontal gene transfer (HGT) (Supplementary Table 6). Most of these HGT candidates (1778 genes) were apparently derived from bacteria. We detected only two genes potentially derived from eukaryotes, encoding proteins similar to actin (UABAM\_01738) and acyloxyacyl hydrolase (AOAH; UABAM\_05422), respectively. ‘*Ca. Uab amorphum*’ apparently lacks other eukaryotic-derived genes previously reported in Planctomycetes, such as genes associated with sterol biosynthesis, membrane coat protein, integrin and the inter-alpha-trypsin inhibitor<sup>18–20</sup>. We searched for homologues of eukaryote signature proteins (ESPs)<sup>21</sup> and found 27 ESP candidates (Supplementary Table 7). Most of these candidates were more similar to other bacterial proteins than to eukaryotic

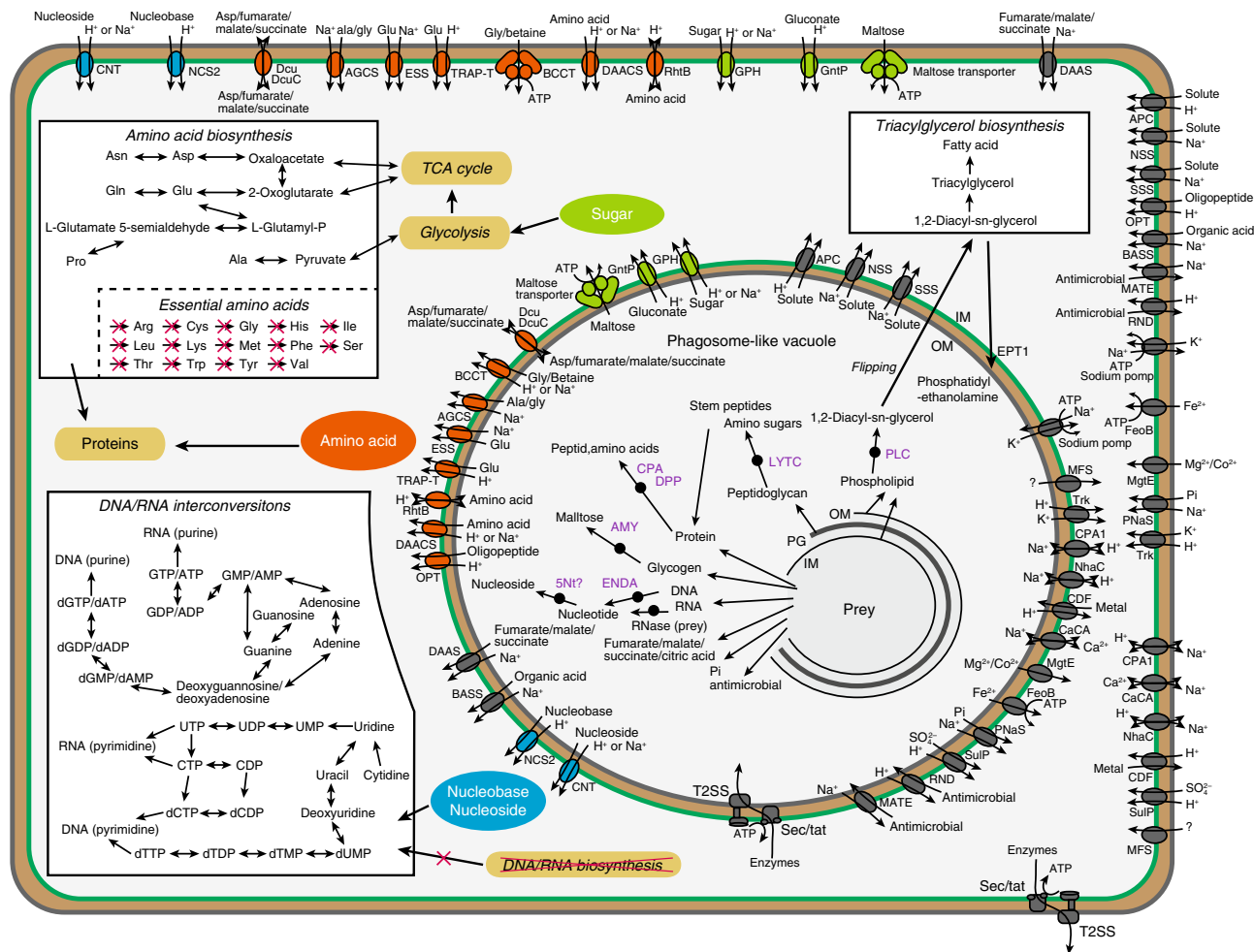


**Fig. 3 Ultrastructure of ‘*Candidatus Uab amorphum*’.** **a** Transmission electron micrograph of ‘*Ca. Uab amorphum*’ showing cytoplasm (C), cytoplasmic membrane (CM), nuclear body (NB), outer membrane (OM), periplasm (PE) and phagosome-like vacuole (PV). Arrows indicate dense granules that are probably glycogen granules. **b, c** High-magnification transmission electron micrograph showing the membrane structures. Cytoplasmic membrane (CM), nuclear body (NB), outer membrane (OM) and periplasm (PE). **d** Scanning electron micrograph of a ‘*Ca. Uab amorphum*’ that is about to engulf a prey bacterium (double arrowhead). **e** Transmission electron micrograph of a ‘*Ca. Uab amorphum*’ that is about to engulf a bacterium (double arrowhead). **f** Ultrastructure of PV that is surrounded by outer membrane (OM), periplasm (PE) and cytoplasmic membrane (CM). Double arrowhead indicates engulfed prey bacterium **g, h** that disrupted prey bacteria in PVs. Scale bars, 1  $\mu\text{m}$  (**a, d, e**), 500 nm (**b, h**) and 200 nm (**c, f, g**).

proteins, with the exception of an actin-like protein and four WD40 repeat-containing proteins (Supplementary Table 7). However, the actin-like protein appears to be more closely related to certain archaeal actins than to eukaryotic actins (see the section ‘Fibrous structures and an actin-like protein’).

**Phagocytosis-like cell engulfment.** The most notable feature of ‘*Ca. Uab amorphum*’ is its ability to engulf other organisms (prey), including Gram-negative bacteria (*Alteromonas macleodii* and *Escherichia coli*), Gram-positive bacteria (*Lactobacillus farciminis* and *Staphylococcus condimentii*) and a picoeukaryotic alga (*Bathycoccus prasinos*) (Supplementary Table 8, Supplementary Fig. 7). However, it was not able to engulf a marine yeast *Debaryomyces hansenii*, possibly due to its relatively large cell size (ca. 5  $\mu\text{m}$  in diameter, Supplementary Movie 3). Light and fluorescent microscopy showed that prey cells were taken up by surface invagination of the ‘*Ca. Uab amorphum*’ cell. In the presence of sodium azide (which inhibits the ATP hydrolase

activity of F-ATPase), cells did not show any locomotion or cell engulfment, which suggest that these may be energy-dependent processes (Supplementary Fig. 1b, c). Feeding experiments by using an AcGFP1-labelled *E. coli* showed that the rod-shaped signals were separated into several globules (Fig. 1j–o, Supplementary Video 4) that disappeared over time and did not diffuse into the cell. We detected few acidic and reactive oxygen species (ROS) signals in cells of ‘*Ca. Uab amorphum*’ (Supplementary Fig. 1d–i). Electron microscopy analyses showed invagination of the ‘*Ca. Uab amorphum*’ cell surface around prey cells (Fig. 3d, e). Consistent with the prediction that both outer and cytoplasmic membranes invaginate during prey engulf, the resulting compartment enclosing the prey cell (phagosome-like vacuole, PV) appeared to be surrounded by two membranes, i.e. the inner membrane of the PV would be the putative Gram-negative outer membrane, and the outer membrane of the PV would be the putative cytoplasmic membrane (Fig. 3f). Most prey cells in PVs were found to be disrupted to some extent (Fig. 3g, h), which suggests that digestion may occur in PVs. Some PVs exhibited



**Fig. 4 Putative transporters, enzymes and pathways of ‘*Candidatus Uab amorphum*’.** Arrows indicate pathways that are present. Pathways that are absent are crossed out. Purple enzymes are predicted to be secreted in the PV or localized at the outer membrane and involve to degradation of prey. Green line indicates the cytoplasmic membrane. Grey line indicates the outer membrane. Brown layers indicate periplasm. Blue circles indicate nucleoside and nucleotide transporters. Green circles indicate sugar transporters. Orange circles indicate peptide and amino acid transporters.

narrow ducts that appeared to be connected to the outside of the cell, or to another PV (Supplementary Fig. 8a, b). However, serial sectional observation showed that some PVs apparently lacked such ducts. It is therefore possible that ducts may disappear over time, and that PVs may eventually close (Supplementary Fig. 8c–k).

The ‘*Ca. Uab amorphum*’ genome includes genes encoding putative digestive enzymes, which might degrade prey components. We detected 179 putative peptidases (Supplementary Data 1), and some digestive proteins that may be secreted or localized at the outer membrane (Fig. 4, Supplementary Table 9). Genes encoding  $\alpha$ -amylase, carboxypeptidase and deoxyribonuclease I may have been acquired via HGT from Bdellovibrionales or Myxococcales, which are known to digest bacteria outside of the cells<sup>22,23</sup> (Supplementary Fig. 9a–c). In mammals, AOA (one of the proteins that might have been obtained by HGT from eukaryotes, Supplementary Fig. 9d) degrades and inactivates lipopolysaccharides of the bacterial outer membrane<sup>24</sup>. Therefore, it is possible that the putative AOA counterpart in ‘*Ca. Uab amorphum*’ could potentially be also involved in degradation of bacterial lipopolysaccharides.

In Gram-negative bacteria, the Sec/Tat pathway transports proteins from the cytoplasm to the periplasm, and the T2SS transports proteins from the periplasm to the outside of cells<sup>25</sup>. We found genes encoding most of the components of the Sec/Tat

pathway and T2SS in the ‘*Ca. Uab amorphum*’ genome (Supplementary Table 4). We speculate that these secretion systems may be used to secrete digestive enzymes into the PV. Other genes, encoding putative transporters of amino acids, nucleosides and sugars (Fig. 4), might be involved in the utilization of prey components.

De novo phospholipid synthesis may be important for formation and repair of PV membranes in ‘*Ca. Uab amorphum*’. Phosphatidylethanolamine (PE) is one of the main phospholipids in bacterial membranes and is synthesized from phosphatidylserine (PS) by phosphatidylserine decarboxylases (PSD) in bacteria<sup>26</sup>. In addition to a PSD gene candidate, we found that ‘*Ca. Uab amorphum*’ has two genes potentially encoding ethanolamine phosphotransferase (EPT1; UABAM\_02218 and UABAM\_03998), which is also found in a few myxobacteria and eukaryotes (Supplementary Fig. 10a). EPT1 synthesizes PE from 1,2-diacyl-sn-glycerol (DAG), which is the degradation product of PE and phosphatidylcholine<sup>27</sup>.

**Metabolic adaptation to phagotrophy.** The ‘*Ca. Uab amorphum*’ genome includes many genes similar to antibiotic-resistance genes (Supplementary Table 10). Planctomycetes are known to have  $\beta$ -lactamase,<sup>28</sup> however, we found genes encoding putative drug efflux systems, most of which are potentially derived from

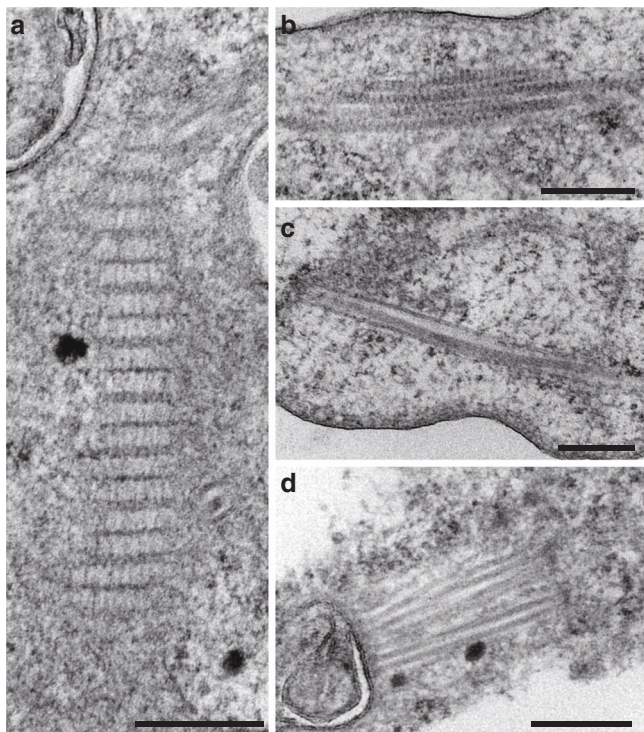
HGT. We speculate that these genes might confer protection against prey antibiotics that are released into the PV.

Dependence on phagotrophy can often result in starvation due to unstable food source. Therefore, certain predatory bacteria have developed specific resting forms to resist starvation<sup>23,29</sup>. However, we were unable to observe any ‘*Ca. Uab amorphum*’ resting forms via microscopy.

‘*Ca. Uab amorphum*’ possesses genes potentially encoding enzymes for the synthesis of triacylglycerol (TAG) from phospholipid via DAG (Supplementary Fig. 10b–d). Given that genes for this pathway were not detected in the genomes of non-phagotrophic relatives, we speculate that ‘*Ca. Uab amorphum*’ may use this metabolic pathway to store energy as TAG in order to counter starvation.

Phagotrophy can drive degenerative evolution in metabolism. Our genomic analysis suggested that ‘*Ca. Uab amorphum*’ may be able to synthesize only six amino acids de novo (Fig. 4, Supplementary Data 2). The inability to synthesize some amino acids is found in animals, phagotrophic and parasitic protists and predatory bacteria that acquire nutrients from preys or hosts<sup>30,31</sup>. Interestingly, ‘*Ca. Uab amorphum*’ also lacked candidate genes for de novo purine and pyrimidine biosynthesis, but had genes for salvage pathways (Fig. 4, Supplementary Figs. 11 and 12). It is therefore possible that ‘*Ca. Uab amorphum*’ must acquire some amino acids and nucleotides from prey. The absence of de novo purine and pyrimidine biosynthesis pathways in some obligate parasites has previously been reported<sup>32,33</sup>.

**Fibrous structures and an actin-like protein.** Actin is a eukaryote-specific protein that forms filaments by polymerization and plays crucial roles in cellular motility, cytoskeleton formation and phagocytosis<sup>34</sup>. TEM observation showed that ‘*Ca. Uab amorphum*’ seems to have at least four different types of fibrous structures. These include large striated fibres that occasionally reached 2  $\mu\text{m}$  in length, small striated fibres, short bundled fibres and cylindrical structures containing a linear fibre.



**Fig. 5** Fibrous structures in ‘*Candidatus Uab amorphum*’. **a** Large striated fibre. **b** Small striated fibre. **c** Short bundled fibre. **d** Cylindrical structures containing a linear fibre. Scale bars, 200 nm.

and cylindrical structures that contain linear fibres (Fig. 5a–d). Such conspicuous and large fibrous structures have not been reported in other bacteria as far as we know; we therefore propose that these structures may be involved in the specific locomotion and engulfment processes displayed by ‘*Ca. Uab amorphum*’.

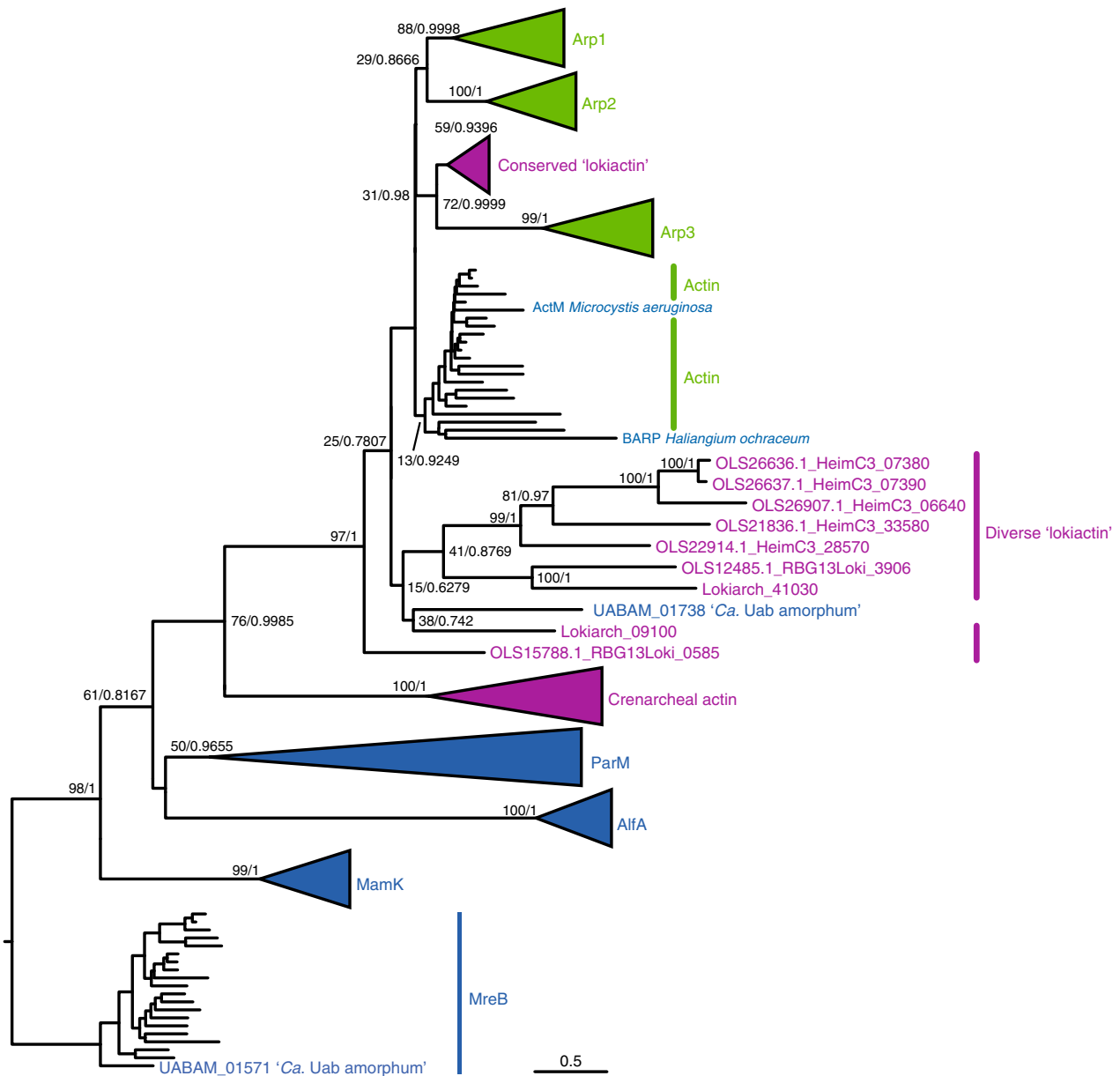
As we previously mentioned, the ‘*Ca. Uab amorphum*’ genome includes a gene encoding an actin-like protein. Molecular phylogenetic analysis showed that this protein was distant from the bacterial actin homolog MreB and the Crenarchaeal actin, and formed a clade with eukaryotic actins, actin-related proteins (ARPs) 1–3, bacterial actin-related protein (BARP) from *Haliangium ochraceum*<sup>35</sup>, ActM from *Microcystis aeruginosa*<sup>36</sup> and Asgard actins (Fig. 6). This analysis suggests that the Uab actin may not have been directly transferred from eukaryotes, but possibly derived from an Asgard or related archaeon. We investigated the presence of genes encoding actin-binding proteins (known to be involved in assembly, stabilisation and cross-link of actin filament<sup>37</sup>) and found two candidate proteins (UABAM\_01722 and UABAM\_04996) that are similar to Aip1 and Crn1, respectively (Supplementary Table 11). We used F-actin-binding phalloidin and an antibody specific for eukaryotic actin but failed to observe a signal on ‘*Ca. Uab amorphum*’ cells when examined by fluorescent microscopy. Therefore, it remains unclear whether the actin-like protein, or the putative actin-binding proteins, form fibrous structures and/or whether they are involved in the organism’s unique locomotion and engulfment abilities.

## Discussion

In this study, we describe the bacterium ‘*Candidatus Uab amorphum*’, which was discovered from a surface seawater sample collected at the Republic of Palau. Molecular phylogenetic analyses indicate that the organism belongs to the phylum Planctomycetes and is closely related to anammox bacteria. Planctomycetes are known to have unusual eukaryote-like features such as budding cell division, sterol synthesis and uptake of macromolecules via endocytosis-like behaviour. The endocytosis-like uptake was reported in *Gemmata obscuriglobus* as an ATP-dependent, membrane-involving process with similarities to eukaryotic endocytosis<sup>38</sup> however, a recent study reported that the macromolecules are accumulated in the periplasm<sup>13</sup>. Species of anammox bacteria have a unique membrane-bounded organelle called the anammoxosome that is involved in ammonium oxidation<sup>39</sup>. ‘*Ca. Uab amorphum*’ presents other eukaryotic-like features, such as a large and flexible cell, amoeba-like locomotion and uptake of other organisms by phagocytosis-like engulfment. In contrast to the endocytosis-like process previously reported for *G. obscuriglobus*, the phagocytosis-like engulfment observed in ‘*Ca. Uab amorphum*’ appears to involve invagination of both outer and inner membrane, and the resulting PV seems to be surrounded by a double membrane. In *G. obscuriglobus*, macromolecules are accumulated in the cytoplasm or periplasm<sup>13,38</sup>, while in ‘*Ca. Uab amorphum*’ the AcGFP1 signals of engulfed *E. coli* cells appear to be compartmentalized in PVs (Fig. 1j–o, Supplementary Movie 4).

Our microscopy studies show that ‘*Ca. Uab amorphum*’ can engulf various bacteria and picoeukaryotes, and appears to digest the engulfed organisms in specific compartments. The genomic analyses suggest that the organism may be able to utilize prey components such as sugar, amino acids and nucleotides from their prey. Therefore, we propose that engulfment by ‘*Ca. Uab amorphum*’ may serve a phagocytosis-like function.

Eukaryotic phagocytosis is regulated by many eukaryote-specific genes that are involved in recognition of prey, signal transduction, rearrangement of cytoskeleton, membrane



**Fig. 6 Maximum likelihood tree of actin-like proteins.** Left and right values on nodes indicate bootstrap value and Bayesian posterior probability, respectively. Bacterial actin-related protein (BARP) of *Haliangium ochraceum* and ActM of *Microcystis aeruginosa* form a clade with eukaryotic actin. Scale bar indicates amino acid substitution rate per site.

remodelling and phagosome maturation<sup>40</sup>. ‘*Ca. Uab amorphum*’ possesses an actin-like gene that might be derived from an Asgard archaeon or a relative. However, we did not identify any other genes with homology to eukaryotic genes involved in phagocytosis. In addition, the structure of the ‘*Ca. Uab amorphum*’ PV seems to be different from that of the eukaryotic phagosome. Therefore, the phagocytosis-like process observed in ‘*Ca. Uab amorphum*’ is probably not homologous to eukaryotic phagocytosis, despite apparent morphological and functional similarities.

The endokaryotic hypothesis proposes that the origin of an ancestral eukaryote involved a Gram-negative bacterial host engulfing an archaeon, accounting for origin of the nucleus as an engulfed cell<sup>41</sup>. ‘*Ca. Uab amorphum*’ is unlikely to be a descendant of such Gram-negative bacterial host, but it may be a good bacterial model to test this hypothesis. Although it has been suggested that phagocytosis without mitochondria is implausible due to the energetic cost of phagocytic functions<sup>42,43</sup>, our study indicates that prokaryotes can take up other microorganisms in

the absence of an energy-producing symbiont. Therefore, primitive phagocytosis may have originated in a mitochondrion-less eukaryote ancestor, with mitochondrion acquisition by phagocytosis being a later event.

Our analysis of environmental 16S rRNA gene data indicates that many uncultured lineages detected in various environments belong to the ‘*Ca. Uab amorphum*’ clade and have high genetic diversity. Therefore, it is tempting to speculate that some of these uncultured organisms may display other unusual phenotypes, may be acting as bacterial predators in marine, freshwater and terrestrial microbial ecosystems. Culture establishment and comparative analysis of Uab lineages will help to unveil their ecological importance and the evolution of eukaryotic-like features, and may provide insights about the origin of eukaryotes.

**Methods**

**Data reporting.** No statistical calculations were performed to pre-determine the sample size.

**Sample collection and culture establishment.** Surface seawater was collected in the Republic of Palau (longitude = 7.181386° N, latitude = 134.336947° E) on 27 October 2015. Approximately 500 mL of the seawater sample was filtered through 5 µm Isopore membrane filters (Millipore, Billerica, USA), and was subsequently concentrated to approximately 10 mL using 0.6 µm Isopore membrane filters (Millipore, Billerica, USA). The concentrated sample was serially diluted across 12 wells by using a 96-well culture plate filled with ESM medium<sup>44</sup>. Samples were incubated at 24 °C in the dark for 10 days. ‘*Candidatus* Uab amorphum’ and other bacterial cells were observed in several wells under light microscope following the incubation period. Single ‘*Ca.* Uab amorphum’ cells were isolated from the incubated sample with a micropipette, and were transferred into a new 96-well cell culture plate filled with the ESM medium. A small amount of the incubated sample was also added in the new 96-well cell culture plate together with isolated ‘*Ca.* Uab amorphum’ cell in order to inoculate bacteria as food sources. As a result, a xenic culture that consists of clonal ‘*Ca.* Uab amorphum’ and unidentified bacteria was established. The xenic culture was maintained in a culture flask or a culture plate in ESM medium at 20 °C in the dark, and was subcultured into a new medium every month.

To establish a monoxenic culture of ‘*Ca.* Uab amorphum’, *Alteromonas macleodii* (NBRC 102226), which was obtained from NITE Biological Resource Centre (NBRC), was inoculated into a 96-well culture plate filled with the ESM medium. A single cell of ‘*Ca.* Uab amorphum’ isolated by the micropipette was added into the culture plate. The monoxenic culture of ‘*Ca.* Uab amorphum’ was maintained in the same condition as the xenic culture. No antibodies were used during the culture establishment. The monoxenic culture was deposited at the Japan Collection of Microorganisms (JCM) as JCM 39082.

**Light and time-lapse microscopic observation.** Cells of ‘*Ca.* Uab amorphum’ and unidentified prey bacteria in the xenic culture were inoculated into glass-bottomed dishes filled with the ESM medium. Cells were incubated for 1–5 days prior to observation. Light micrographs and time-lapse videos were taken using an Olympus IX71 inverted microscope (Olympus, Tokyo, Japan) equipped with an Olympus DP73 CCD camera (Olympus, Tokyo, Japan). Sodium azide (76.9 mM) was added to the xenic culture 15 min before observation. For the feeding experiment, cells of *Debaryomyces hansenii* JCM 1439, *Lactobacillus farciminis* JCM 1097 and *Staphylococcus condimentii* JCM 6074 (obtained from JCM) were added to glass-bottomed dishes together with ‘*Ca.* Uab amorphum’ and observed immediately under light microscope. Cells of *Bathycoccus prasinos* NIES-2670 (obtained from the National Institute for Environmental Studies, NIES) were added to glass-bottomed dishes together with ‘*Ca.* Uab amorphum’, and were incubated for one day prior to light microscopic observation and specimen preparation for transmission electron microscopic (TEM) observation.

**DNA extraction, PCR and sequencing.** Cells of ‘*Ca.* Uab amorphum’ and unidentified prey bacteria in the xenic culture were collected by centrifugation. Total DNA was extracted using the DNeasy plant mini kit (Qiagen, Hilden, Germany) according to the manufacturer’s instructions. PCR was performed using primer pairs specific for the bacterial 16S rRNA gene (27F and 1492R)<sup>45,46</sup>. The PCR cycles (30 cycles) consisted of denaturation at 96 °C for 10 s, annealing at 55 °C for 30 min and extension at 68 °C for 2 min. The amplicon was purified with the QIAquick Gel Extraction kit (Qiagen, Hilden, Germany), and was then cloned into the p-GEM T-easy vector (Promega, Tokyo, Japan). Ten clones were completely sequenced with a 3130 Genetic Analyser (Applied Biosystems, Foster City, CA), using the BigDye Terminator v3.1 cycle sequencing kit (Applied Biosystems). Sequenced clones were assembled into six sequences by the CodonCode Aligner (CodonCode Co., Centerville, MA) with a sequence similarity threshold of 99%. The 16-S rRNA gene sequence of ‘*Ca.* Uab amorphum’ was deposited in the GenBank database with accession code LC496071.

**Fluorescent microscopy.** Fluorescein isothiocyanate (FITC)-labelled oligonucleotide probes EUB338 (ref. 47) and PLA886 (ref. 48) were purchased from FASMAC Co., Ltd. (Kanagawa, Japan). EUB338 completely matched all sequences derived from the xenic culture, while PLA886 matched only to the 16S rRNA gene sequence of ‘*Ca.* Uab amorphum’, and had at least three mismatches to other sequences (Supplementary Fig. 2). The xenic culture of ‘*Ca.* Uab amorphum’ was cultured on coverslips that were treated with 0.1% (w/v) poly L-lysine (Sigma Chemical Co., St. Louis, MO) for one week prior to observation. Fixation and hybridization were performed as described in a previous study<sup>47</sup>. Cells grown on coverslips were fixed with 4% (w/v) paraformaldehyde for 16 h at 4 °C, and were then treated with a series of 50%, 80% and 96% ethanol for 3 min each and air-dried. Hybridization buffer containing 0.9 M NaCl, 20 mM Tris/HCl (pH 8.0), 0.01% (w/v) SDS, 5 ng/µL probe and a specific amount of formamide (Supplementary Table 12) were mounted on coverslips and incubated at 46 °C for 1.5 h in humid chambers. In the case of PLA886, an equimolar amount of a competitor oligonucleotide (cPLA886) was added to the hybridization buffer to avoid binding to non-target bacteria. Coverslips were washed in pre-heated (48 °C) washing buffer containing 0.9 M NaCl, 20 mM Tris/HCl (pH 8.0) and 0.01% (w/v) SDS, and were then placed in the washing buffer for 20 min at 48 °C. They were then washed with distilled water and air-dried. Each coverslip was incubated with PBS

containing 0.1 µg/mL 4',6-diamidino-2-phenylindole (DAPI) for 10 min in the dark. Coverslips were washed with distilled water, air-dried, mounted with Slow-Fade Diamond (Invitrogen, Carlsbad, CA) and sealed with nail polish. Specimens were then observed using a Leica DMRD microscope (Leica, Wetzlar, Germany) equipped with an Olympus DP73 CCD camera (Olympus, Tokyo, Japan).

AcGFP1-labelled *Escherichia coli* was prepared by transformation using TOP10-competent cells (Thermo Fisher Scientific, MA, USA) and pAcGFP1 Vector (Takara, Tokyo, Japan). Cells of AcGFP1-labelled *E. coli* were added to glass-bottomed dishes together with ‘*Ca.* Uab amorphum’ and immediately observed under Nikon A1 confocal microscope (Nikon, Tokyo, Japan). For detection of acidic compartments in ‘*Ca.* Uab amorphum’ cells, ‘*Ca.* Uab amorphum’ and AcGFP1-labelled *E. coli* in glass-bottomed dishes were incubated with 1 µM L<sup>-1</sup> LysoTracker Red DND-99 (Thermo Fisher Scientific) in the dark for 1 h. Cells were washed three times by ESM medium and observed under Nikon A1 confocal microscope. For detection of reactive oxygen species, ‘*Ca.* Uab amorphum’ in glass-bottomed dishes was incubated with 1 µM L<sup>-1</sup> LysoTracker Red DND-99 and 1 mM L<sup>-1</sup> DCFH-DA (OxiSelect Intracellular ROS Assay Kit, Cell Biolabs, CA, USA) in the dark for 1 h. Cells were washed three times by ESM medium and observed under Nikon A1 confocal microscope.

**Electron microscopy.** For scanning electron microscopy, cells of ‘*Ca.* Uab amorphum’ were cultured on 8.5-mm-diameter glass SEM plates (Okenshoji Co., Tokyo, Japan) treated with 0.1% (w/v) poly L-lysine (Sigma Chemical Co., St. Louis, MO) for one week. Cells were pre-fixed with vapour of 4% (w/v) osmium tetroxide (OsO<sub>4</sub>) for 30 min, and were subsequently post-fixed with 1% (w/v) osmium tetroxide in culture medium for 2 h. Fixed cells were dehydrated in a series of 15–100% (v/v) ethanol. After dehydration, specimens were placed once in a 1:1 mixture of 100% ethanol and t-butyl alcohol, and then twice in 100% t-butyl alcohol, and chilled in the freezer. The specimens were freeze-dried using a VFD-21S (SHINKU-DEVICE, Ibaraki, Japan) freeze drier, and were then mounted on aluminium stubs using carbon paste. Specimens were sputter-coated with platinum-palladium using a Hitachi E-102 sputter-coating unit (Hitachi High-Technologies Corp., Tokyo, Japan), and were observed using a JSM-6360F field emission SEM (JEOL, Tokyo, Japan).

For TEM observation, cells were collected by centrifugation, and were pre-fixed with a mixture containing 1% glutaraldehyde, 0.1 M cacodylate buffer and 0.25 M sucrose for 1 h. Pelleted cells were washed twice with 0.2 M sodium cacodylate buffer (pH 7.2), and were post-fixed with a mixture containing 1% osmium tetroxide and 0.2 M cacodylate buffer. Fixed cells were dehydrated in a series of 30–100% (v/v) ethanol. After dehydration, cells were placed in a 1:1 mixture of 100% ethanol and acetone for 10 min, and 100% acetone for 10 min two times each. Resin replacement was performed with a 1:1 mixture of acetone and Agar Low Viscosity Resin R1078 (Agar Scientific Ltd., Stansted, England) for 30 min, and then with Low Viscosity Resin R1078 for 2 h. Resin was polymerized at 60 °C for 12 h. Ultra-thin sections were prepared on a Reichert Ultracut S ultramicrotome (Leica, Vienna, Austria). Cell sections were double-stained with 2% (w/v) uranyl acetate and lead citrate, and were observed using a Hitachi H-7650 electron microscope (Hitachi High-Technologies Corp., Tokyo, Japan) equipped with a Veleta TEM CCD camera (Olympus, Tokyo, Japan).

**Molecular phylogenetic analyses using 16S rRNA gene.** We prepared a dataset containing 16S rRNA gene sequences of major bacterial lineages and ‘*Ca.* Uab amorphum’. The sequences were first automatically aligned in MAFFT v7.273<sup>49</sup> with the G-INS-i algorithm at default settings, and were then manually edited with SeaView version 4.6<sup>50</sup>. Ambiguous regions in the alignment were trimmed with SeaView. The final alignment consisted of 71 OTUs and 1307 sites. The best substitution model was searched for using jModelTest 2.1.10<sup>51</sup>, and the GTR + Γ + I model was selected. A maximum likelihood (ML) tree was heuristically searched for using RAxML version 8.1.15<sup>52</sup> under the GTR + Γ + I model. Tree searches began with 20 randomized maximum-parsimony trees, and the highest log likelihood (lnL) was selected as the ML tree. A nonparametric bootstrap analysis with 1000 replicates was conducted using RAxML under the GTR + Γ + I model. A Bayesian analysis was run using MrBayes 3.2.2<sup>53</sup> with the GTR + Γ + I model for each dataset. One cold and three heated Markov chain Monte Carlo chains with default temperatures were run for 5 × 10<sup>6</sup> generations; lnL values and trees were sampled at 100-generation intervals. The first 2.2 × 10<sup>6</sup> generations with an average standard deviation of split frequencies (ASDSF) greater than 0.01 were discarded as “burnin”. Bayesian posterior probabilities (BPP) and branch lengths were calculated from the remaining trees.

In order to conduct molecular phylogenetic analysis of environmental sequences related to ‘*Ca.* Uab amorphum’, we prepared a 16S rRNA gene dataset including major subgroups of Planctomycetes, which were selected based on a previous phylogenetic study<sup>54</sup>. We added several Chlamydiae and Verrucomicrobia OTUs to the dataset as outgroups. Environmental sequences related to ‘*Ca.* Uab amorphum’ were screened using the BLASTn program, with the 16S rRNA gene sequence of ‘*Ca.* Uab amorphum’ as query. The top 300 BLASTn hit sequences were added to the dataset. Sequences in the dataset were automatically aligned in MAFFT with the G-INS-i algorithm at default settings, and were then manually aligned and trimmed using SeaView. Preliminary maximum likelihood tree was constructed by RAxML with rapid bootstrap analysis (100 replicates) under a



GTR +  $\Gamma$  model. We removed BLASTn-derived environmental sequences that did not form a clade with ‘*Ca. Uab amorphum*’ from the final dataset. Sequences in the final dataset were aligned and trimmed as shown above; the final alignment consisted of 100 OTUs and 1,351 sites. The best model was identified using jModelTest, and the GTR +  $\Gamma$  + I model was selected. The ML tree was heuristically searched for using RAxML under the GTR +  $\Gamma$  + I model. Tree searches began with 20 randomized maximum-parsimony trees, and the highest lnL was selected as the ML tree. A nonparametric bootstrap analysis with 1,000 replicates was conducted using RAxML under the GTR +  $\Gamma$  + I model. A Bayesian analysis was conducted using MrBayes with the GTR +  $\Gamma$  + I model for each dataset. One cold and three heated Markov chain Monte Carlo chains with default temperatures were run for  $5 \times 10^6$  generations; lnL values and trees were sampled at 100-generation intervals. The first  $1.0 \times 10^6$  generations with ASDSF greater than 0.01 were discarded as “burnin”. BPP and branch lengths were calculated from the remaining trees.

We screened for sequences related to ‘*Ca. Uab amorphum*’ from Tara Ocean 16S rRNA gene sequences<sup>10</sup> using BLASTn, with the ‘*Ca. Uab amorphum*’ 16S rRNA gene sequence and environmental Uab clade sequences as query. We added sequences of BLASTn hits with sequence similarities above 97% and sequence length over 100 bp to the dataset used in the environmental 16S rRNA gene sequence analysis. Sequences in the dataset were automatically aligned in MAFFT with the G-INS-i algorithm at default settings, and were then manually aligned and trimmed using SeaView. Preliminary maximum likelihood tree was constructed by RAxML with rapid bootstrap analysis (100 replicates) under a GTR +  $\Gamma$  model. We removed environmental sequences that did not form a clade with ‘*Ca. Uab amorphum*’ from the final dataset. Sequences in the final dataset were aligned and trimmed as shown above; the final alignment consisted of 189 OTUs and 1,509 sites. The ML tree was heuristically searched for using RAxML under the GTR +  $\Gamma$  + I model. Tree searches began with 20 randomized maximum-parsimony trees, and the highest lnL was selected as the ML tree. A nonparametric bootstrap analysis with 1,000 replicates was conducted using RAxML under the GTR +  $\Gamma$  + I model.

**Genome sequencing.** For genome sequencing, total DNA was extracted from the monoxenic culture of ‘*Ca. Uab amorphum*’ that was cultivated until most prey bacteria were consumed. Long reads (157,852 reads, 1.4 Gb) were sequenced using PacBio RS II (Pacific Biosciences, Melon Park, CA) with one cell (P6-C3 chemistry). Mean read length was 9.0 kbp. Sequencing was performed by Macrogen (Seoul, Korea).

**Genome assembly and annotation.** Raw long reads were error-corrected and assembled using Canu v1.4, and 49 contigs were acquired. To remove prey sequences, homology search was performed using BLASTn against the NCBI nr database; only the longest contig was derived from *Ca. Uab amorphum* (Supplementary Table 2), which was manually circularized. Any other sequences, e.g. plasmids, were not found in ‘*Ca. Uab amorphum*’. The contig was polished using raw long reads by Pbjelly 0.3.0<sup>55</sup> and Quiver 2.1.0 (Pacific Biosciences). Gene models were constructed using the DFAST web server<sup>56</sup> with the Refseq database. Functional annotation was performed using the EggNOG mapper web server<sup>57</sup>.

**Molecular phylogenetic analysis using 171 proteins.** Molecular phylogenetic analysis was performed using 171 highly conserved orthologues. The dataset was composed of 29 operation taxonomic units (OTU), which contained most available genomes from planctomycetes. Four species in Verrucomycrobia or Chlamydiae were used as an outgroup. The highly conserved orthologues were searched for by reciprocal blast best hit with the cut-off of coverage  $\geq 30\%$  and similarity  $\geq 40\%$ ; 171 proteins (52,272 amino acids) were conserved among  $\geq 90\%$  of the OTU, and were used for the phylogenetic analysis. The amino acids were aligned using Mafft v7.221, and highly diversified regions were manually trimmed on MEGA 7<sup>58</sup>. Model test was performed using IQ-TREE 1.4.3<sup>59</sup>, following the Bayesian information criterion (BIC). Phylogenetic analysis was performed as described above using RAxML version 8.2.9. A nonparametric bootstrap analysis with 200 replicates under the LG + gamma + I model was performed. A Bayesian analysis was run using MrBayes 3.2.6 with the same model for  $1.0 \times 10^6$  generations; lnL values and trees were sampled at 100-generation intervals. The initial 25% generations with ASDSF values greater than 0.01 were discarded as “burnin”. BPP and branch lengths were calculated from the remaining trees.

**Inference of digestive proteins in the FV.** Subcellular localization of proteins was predicted using PSORTb 3.0<sup>60</sup>, Cello v.2.5<sup>61</sup> and LocTree3<sup>62</sup> with an organism option: Gram-negative bacteria or bacteria. Proteins were categorized into “secreted” or “extracellular” by at least one of the three prediction programmes that were considered a candidate FV protein. Digestive proteins were identified using the KEGG database. Digestive proteins for DNA/RNA were searched for using GO terms (GO:0004518). Digestive proteins for peptidoglycan were searched for using the HyPe web server<sup>63</sup>.

**Detection of horizontal gene transfer.** Genes putatively derived by HGT were identified using HGTector v0.2.1<sup>64</sup> with the following cut-off values: e-value =  $1e-20$ ,

identity = 30 and coverage = 40. In this analysis, based on the phylogenetic relationships, “selfGroup” and “closeGroup” were set as ‘*Candidatus Brocadiaaceae*’ (taxonomic ID: 1127830) and the PVC (Planctomycetes–Verrucomicrobia–Chlamydiae) group (taxonomic ID: 1783257), respectively.

In addition, we searched Uab homologues of the 347 eukaryote-specific proteins (ESPs)<sup>21</sup> and actin-binding proteins<sup>37</sup> in the *Saccharomyces* Genome Database<sup>65</sup> by BLASTp with the cut-off: E-value <  $1E-5$ . The Uab proteins with any hits were re-checked by BLASTp against the NCBI nr database.

**Molecular phylogenomic analyses of single proteins.** For molecular phylogenetic analyses of single proteins ( $\alpha$ -amylase, actin, acylglycerol hydrolase (AOAH), phospholipase C (PLC), diacylglycerol acyltransferase (DGAT), carboxypeptidase, DNase I and EPT1), we screened each protein sequence by BLASTp, against the NCBI nr database and constructed the dataset. Datasets were aligned by MAFFT v7.273, and were then manually edited with SeaView version 4.6 or MEGA 7. The final alignments consisted of 331 amino acid positions and 17 OTUs for  $\alpha$ -amylase, 362 amino acid positions and 103 OTUs for actin, 534 amino acid positions and 28 OTUs for AOAH, 255 amino acid positions and 17 OTUs for PLC, 408 amino acid positions and 17 OTUs for DGAT, 206 amino acid positions and 6 OTUs for carboxypeptidase, 173 amino acid positions and 11 OTUs for DNase I and 154 amino acid positions and 10 OTUs for EPT1. ML trees were constructed using IQ-TREE 1.5.5 following the best-fit model, which was chosen in accordance with BIC (WAG + I + G4 for  $\alpha$ -amylase, LG + G4 for actin, DGAT and DNase I, LG + I + G4 for AOAH and PLC, WAG + G4 for carboxypeptidase and mtZOA + I + G4 for EPT1) with 200 replicates of nonparametric bootstrap. A Bayesian analysis was run using MrBayes 3.2.6 with the LG +  $\Gamma$  model for the actin dataset. One cold and three heated Markov chain Monte Carlo with default temperatures were run for  $1 \times 10^7$  generations; lnL values and trees were sampled at 100-generation intervals. The first  $7 \times 10^6$  generations with an average standard deviation of split frequencies (ASDSF) value greater than 0.03 were discarded as “burnin”. Bayesian posterior probabilities (BPP) and branch lengths were calculated from the remaining trees.

**Description of ‘*Candidatus Uab amorphum*’.** ‘*Candidatus Uab amorphum*’ (U.a.b masc. n. a giant of Palauan mythology. a.mor’phus. L. neut. adj. amorphum amorphous, deformed).

Marine free-living aerobic Gram-negative bacterium was collected from surface seawater in the Republic of Palau (7.181386° N, 134.336947° E). Cells are flattened and round or oval shape with granular cytoplasm (Fig. 1a–c). Flagellum is absent. Cells were 3.2–7.8  $\mu\text{m}$  ( $4.5 \pm 0.85 \mu\text{m}$ ) in the long axis and 2.8–5.5  $\mu\text{m}$  ( $4.0 \pm 0.57 \mu\text{m}$ ) in the short axis ( $n = 79$ ). When prey bacteria were abundant in the culture, cells occasionally reached 10  $\mu\text{m}$  in diameter (Supplementary Fig. 1a). Cells attach to substrate and show gliding motility with changing shape (Supplementary Movie 1). Cells reproduce by binary fission. Periplasm (or paryphoplasm) highly invaginates into the cytoplasm (or pirellulosome) and shows a reticulate pattern (Fig. 3a–c). The cytoplasm includes multiple nuclear bodies (Fig. 3a, c). The cell interior includes four types of fibrous structures: large striated fibres, small striated fibres, short bundle fibres and cylindrical structures that contain linear fibres (Fig. 5). Cells can engulf bacteria and picoeukaryotes (*Bathycoccus prasinos*) (Fig. 1d–o, Fig. 3d–h, Supplementary Fig. 7, Supplementary Movies 2 and 4). Engulfed bacteria and algae can be found in phagosome-like vacuoles (PVs) (Fig. 2f–h, Supplementary Fig. 7s–t). Some PVs connect to other phagosome-like structures or outside of the cell by narrow ducts (Supplementary Fig. 8a, b). Similarity of the 16S rRNA gene sequence (LC496071) to that of the closest species (the planctomycete ‘*Candidatus Kuenenia stuttgartiensis*’) is 79% (Supplementary Table 1). The genome of ‘*Ca. Uab amorphum*’ (AP019860) is circular and 9,503,110 bp in size. The G + C content of the genome is 39.4 mol%. A co-culture of ‘*Ca. Uab amorphum*’ has been deposited at the JCM as JCM 39082.

**Reporting summary.** Further information on research design is available in the Nature Research Reporting Summary linked to this article.

## Data availability

The 16S rRNA gene sequence of ‘*Ca. Uab amorphum*’ was deposited in DDBJ/GenBank/ENA with the accession number LC496071. Genome assembly was deposited in DDBJ/GenBank/ENA with accession number AP019860. Other relevant data are available in this article and its Supplementary Information files, or from the corresponding author upon request.

Received: 1 February 2019; Accepted: 8 November 2019;  
Published online: 11 December 2019

## References

- Lynch, M. & Conery, J. S. The origins of genome complexity. *Science* **302**, 1401–1404 (2003).

2. López-García, P. & Moreira, D. Open questions on the origin of eukaryotes. *Trends Ecol. Evol.* **30**, 697–708 (2015).
3. Yutin, N., Wolf, M. Y., Wolf, Y. I. & Koonin, E. V. The origins of phagocytosis and eukaryogenesis. *Biol. Direct* **4**, 9 (2009).
4. Poole, A. & Penny, D. Eukaryote evolution: engulfed by speculation. *Nature* **447**, 913 (2007).
5. Cavalier-Smith, T. Predation and eukaryote cell origins: a coevolutionary perspective. *Int. J. Biochem. Cell Biol.* **41**, 307–322 (2009).
6. Embley, T. M. & Williams, T. A. Evolution: steps on the road to eukaryotes. *Nature* **521**, 169–170 (2015).
7. Zaremba-Niedzwiedzka, K. et al. Asgard archaea illuminate the origin of eukaryotic cellular complexity. *Nature* **541**, 353–358 (2017).
8. Spang, A. et al. Complex archaea that bridge the gap between prokaryotes and eukaryotes. *Nature* **521**, 173–179 (2015).
9. Imachi, H. et al. Isolation of an archaeon at the prokaryote-eukaryote interface. Preprint at <https://www.biorxiv.org/content/10.1101/726976v1> (2019).
10. Kuenen, J. G. Anammox bacteria: from discovery to application. *Nat. Rev. Microbiol.* **6**, 320–326 (2008).
11. Sunagawa, S. et al. Structure and function of the global ocean microbiome. *Science* **348**, 1261359 (2015).
12. Sagulenko, E. et al. Structural studies of planctomycete *Gemmata obscuriglobus* support cell compartmentalisation in a bacterium. *PLoS ONE* **9**, e91344 (2014).
13. Lindsay, M. R. et al. Cell compartmentalisation in planctomycetes: Novel types of structural organisation for the bacterial cell. *Arch. Microbiol.* **175**, 413–429 (2001).
14. Boedeker, C. et al. Determining the bacterial cell biology of Planctomycetes. *Nat. Commun.* **8**, 14853 (2017).
15. Speth, D. R., van Teeseling, M. C. & Jetten, M. S. Genomic analysis indicates the presence of an asymmetric bilayer outer membrane in planctomycetes and verrucomicrobia. *Front. Microbiol.* **3**, 304 (2012).
16. Tocheva, E. I. et al. Peptidoglycan transformations during *Bacillus subtilis* sporulation. *Mol. Microbiol.* **88**, 673–686 (2013).
17. Tocheva, E. I. et al. Peptidoglycan remodeling and conversion of an inner membrane into an outer membrane during sporulation. *Cell* **146**, 799–812 (2011).
18. Santarella-Mellwig, R. et al. The compartmentalized bacteria of the Planctomycetes-Verrucomicrobia-Chlamydiae superphylum have membrane coat-like proteins. *PLoS Biol.* **8**, e1000281 (2010).
19. Pearson, A., Budin, M. & Brocks, J. J. Phylogenetic and biochemical evidence for sterol synthesis in the bacterium *Gemmata obscuriglobus*. *Proc. Natl Acad. Sci. USA* **100**, 15352–15357 (2003).
20. Jenkins, C., Kedar, V. & Fuerst, J. A. Gene discovery within the planctomycete division of the domain Bacteria using sequence tags from genomic DNA libraries. *Genome Biol.* **3**, research0031–research0031.11 (2002).
21. Hartman, H. & Fedorov, A. The origin of the eukaryotic cell: a genomic investigation. *Proc. Natl Acad. Sci. USA* **99**, 1420–1425 (2002).
22. Sockett, R. E. & Lambert, C. *Bdellovibrio* as therapeutic agents: a predatory renaissance? *Nat. Rev. Microbiol.* **2**, 669–675 (2004).
23. Kaiser, D. Myxococcus in *Encyclopedia of Microbiology* (eds. Schaechter, M.) 220–244 (Academic Press, Cambridge, 2009).
24. Munford, R., Lu, M. & Varley, A. Chapter 2: kill the bacteria...and also their messengers? *Adv. Immunol.* **103**, 29–48 (2009).
25. Green, E. R. & Mecsas, J. Bacterial secretion systems: an overview. *Microbiol. Spectr.* **4**, 1–32 (2016).
26. Moser, R., Aktas, M., Fritz, C. & Narberhaus, F. Discovery of a bifunctional cardiolipin/phosphatidylethanolamine synthase in bacteria. *Mol. Microbiol.* **92**, 959–972 (2014).
27. Hjelmstad, R. H. & Bell, R. M. The sn-1,2-diacylglycerol ethanolaminephosphotransferase activity of *Saccharomyces cerevisiae*. Isolation of mutants and cloning of the EPT1 gene. *J. Biol. Chem.* **263**, 19748–19757 (1988).
28. Jeske, O. et al. Planctomycetes do possess a peptidoglycan cell wall. *Nat. Commun.* **6**, 7116 (2015).
29. Sanchez-Amat, A. & Torrella, F. Formation of stable bdelloplasts as a starvation-survival strategy of marine bdellovibrions. *Appl. Environ. Microbiol.* **56**, 2717–2725 (1990).
30. Payne, S. H. & Loomis, W. F. Retention and loss of amino acid biosynthetic pathways based on analysis of whole-genome sequences. *Eukaryot. Cell* **5**, 272–276 (2006).
31. Crossman, L. C. et al. A small predatory core genome in the divergent marine *Bacteriovorax marinus* SJ and the terrestrial *Bdellovibrio bacteriovorus*. *ISME J.* **7**, 148–160 (2013).
32. Xu, P. et al. The genome of *Cryptosporidium hominis*. *Nature* **431**, 1107–1112 (2004).
33. Stephens, R. S. Genome sequence of an obligate intracellular pathogen of humans: *Chlamydia trachomatis*. *Science* **282**, 754–759 (1998).
34. Pollard, T. D. & Cooper, J. A. Actin, a central player in cell shape and movement. *Science* **326**, 1208–1212 (2009).
35. Wu, D. et al. A phylogeny-driven genomic encyclopaedia of Bacteria and Archaea. *Nature* **462**, 1056–1060 (2009).
36. Guljamov, A. et al. Horizontal gene transfer of two cytoskeletal elements from a eukaryote to a cyanobacterium. *Curr. Biol.* **17**, 757–759 (2007).
37. Moseley, J. B. & Goode, B. L. The yeast actin cytoskeleton: from cellular function to biochemical mechanism. *Microbiol. Mol. Biol. Rev.* **70**, 605–645 (2006).
38. Lonhienne, T. G. et al. Endocytosis-like protein uptake in the bacterium *Gemmata obscuriglobus*. *Proc. Natl Acad. Sci. USA* **107**, 12883–12888 (2010).
39. Yakushevskaya, A. et al. Combined structural and chemical analysis of the anammoxosome: A membrane-bounded intracytoplasmic compartment in anammox bacteria. *J. Struct. Biol.* **161**, 401–410 (2008).
40. Flannagan, R. S., Jaumouillé, V. & Grinstein, S. The cell biology of phagocytosis. *Annu. Rev. Pathol.* **7**, 61–98 (2012).
41. Lake, J. A. & Rivera, M. C. Was the nucleus the first endosymbiont? *Proc. Natl Acad. Sci. USA* **91**, 2880–2881 (1994).
42. Lane, N. & Martin, W. The energetics of genome complexity. *Nature* **467**, 929–934 (2010).
43. Martin, W. F., Garg, S. & Zimorski, V. Endosymbiotic theories for eukaryote origin. *Philos. Trans. R. Soc. B. Biol. Sci.* **370**, 20140330 (2015).
44. Kawachi M., et al. *MCC-NIES List of Strains* 9th edn Microalgae, endangered macroalgae and protists. (Microbial Culture Collection at the National Institute for Environmental Studies, Tsukuba, Japan, 2013).
45. Lane, D. J. 16S/23S rRNA Sequencing in *Nucleic acid techniques in bacterial systematic* (eds. Wtackebbrandt, E. & Goodfellow, M.) 115–175 (John Wiley and Sons, New York, 1991).
46. Turner, S., Pryer, K. M., Miao, V. P. & Palmer, J. D. Investigating deep phylogenetic relationships among cyanobacteria and plastids by small subunit rRNA sequence analysis. *J. Eukaryot. Microbiol.* **46**, 327–338 (1999).
47. Amann, R. I. et al. Combination of 16S rRNA-targeted oligonucleotide probes with flow cytometry for analyzing mixed microbial populations. *Appl. Environ. Microbiol.* **56**, 1919–1925 (1990).
48. Neef, A., Amann, R., Schlesner, H. & Schleifer, K. H. Monitoring a widespread bacterial group: in situ detection of planctomycetes with 16S rRNA-targeted probes. *Microbiology* **144**, 3257–3266 (1998).
49. Katoh, K. & Standley, D. M. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol. Biol. Evol.* **30**, 772–780 (2013).
50. Gouy, M., Guindon, S. & Gascuel, O. SeaView version 4: a multiplatform graphical user interface for sequence alignment and phylogenetic tree building. *Mol. Biol. Evol.* **27**, 221–224 (2010).
51. Hess, S., Suthaus, A. & Melkonian, M. “*CandidatusFinniella*” (Rickettsiales, Alphaproteobacteria), novel endosymbionts of Viridiraptorid Amoeboflagellates (Cercozoa, Rhizaria). *Appl. Environ. Microbiol.* **82**, 659–670 (2016).
52. Stamatakis, A. RAXML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* **30**, 1312–1313 (2014).
53. Ronquist, F. & Huelsenbeck, J. P. MrBayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics* **19**, 1572–1574 (2003).
54. Fukunaga, Y. et al. *Phycisphaera mikurensis* gen. nov., sp. nov., isolated from a marine alga, and proposal of Phycisphaeraeae fam. nov., Phycisphaerales ord. nov. and Phycisphaeraeae classis nov. in the phylum Planctomycetes. *J. Gen. Appl. Microbiol.* **55**, 267–275 (2009).
55. English, A. C. et al. Mind the gap: upgrading genomes with Pacific Biosciences RS long-read sequencing technology. *PLoS ONE* **7**, e47768 (2012).
56. Tanizawa, Y., Fujisawa, T. & Nakamura, Y. DFAST: a flexible prokaryotic genome annotation pipeline for faster genome publication. *Bioinformatics* **34**, 1037–1039 (2018).
57. Huerta-Cepas, J. et al. Fast genome-wide functional annotation through orthology assignment by eggNOG-Mapper. *Mol. Biol. Evol.* **34**, 2115–2122 (2017).
58. Kumar, S., Stecher, G. & Tamura, K. MEGA7: molecular evolutionary genetics analysis version 7.0 for bigger datasets. *Mol. Biol. Evol.* **33**, 1870–1874 (2016).
59. Nguyen, L. T., Schmidt, H. A., von Haeseler, A. & Minh, B. Q. IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Mol. Biol. Evol.* **32**, 268–274 (2015).
60. Yu, N. Y. et al. PSORTb 3.0: improved protein subcellular localization prediction with refined localization subcategories and predictive capabilities for all prokaryotes. *Bioinformatics* **26**, 1608–1615 (2010).
61. Yu, C. S., Chen, Y. C., Lu, C. H. & Hwang, J. K. Prediction of protein subcellular localization. *Proteins* **64**, 643–651 (2006).
62. Goldberg, T. et al. LocTree3 prediction of localization. *Nucleic Acids Res.* **42**, W350–W355 (2014).

63. Sharma, A. K., Kumar, S. K. H., Dhakan, D. B. & Sharma, V. K. Prediction of peptidoglycan hydrolases- a new class of antibacterial proteins. *BMC Genomics* **17**, 411 (2016).
64. Zhu, Q., Kosoy, M. & Dittmar, K. HGTECTOR: an automated method facilitating genome-wide discovery of putative horizontal gene transfers. *BMC Genomics* **15**, 717 (2014).
65. Cherry, J. M. et al. Saccharomyces Genome Database: the genomics resource of budding yeast. *Nucleic Acids Res.* **42**, D700–D705 (2012).

### Acknowledgements

We thank the Ministry of Natural Resources, Environment and Tourism, Palau for providing permits for collecting marine resources (permit number, RE-15-12). We thank Dr. Naoto Hanzawa at the Yamagata University for managing sampling trips. We thank Dr. Hirokazu Sakamoto for proposing the name “Uab”. This study was supported by the Japan Society for the Promotion of Science (JSPS) KAKENHI Grant Numbers: 13J00587, 18J02091 and 19K22445.

### Author contributions

T.S. collected samples, established a culture, performed light and electron microscopic observations, FISH experiments, genome sequencing and molecular phylogenetic analyses. S.S. performed molecular phylogenetic analyses, genome assembly and genome analyses. Y.K. performed light and electron microscopic observations. K.I. designed sampling trip and experiments. T.S. and S.S. wrote the paper. All authors edited and approved the paper.

### Competing interests

The authors declare no competing interests.

### Additional information

**Supplementary information** is available for this paper at <https://doi.org/10.1038/s41467-019-13499-2>.

**Correspondence** and requests for materials should be addressed to T.S.

**Peer review information** *Nature Communications* thanks John A Fuerst, Simon Johnston, Olga Lage and Tãm Mignot for their contribution to the peer review of this work.

**Reprints and permission information** is available at <http://www.nature.com/reprints>

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2019