microLife, 2023, **4**, 1–10
DOI: 10.1093/femsml/uqad010
Advance access publication date: 23 March 2023
Research Article

Phylogenetic diversity of core rumen microbiota as described by cryo-ET

Benedikt H. Wimmer [1,‡], Sarah Moraïs [2,3,‡], Ran Zalk [4], Itzhak Mizrahi [2,3,*], Ohad Medalia [1,*]

[1]Department of Biochemistry, University of Zurich, Winterthurerstrasse 190, 8057 Zurich, Switzerland
[2]National Institute of Biotechnology in the Negev, Ben-Gurion University of the Negev, Beer-Sheva 84105, Israel
[3]Department of Life Sciences, Ben-Gurion-University of the Negev, Beer-Sheva 84105, Israel
[4]Ilse Katz Institute for Nanoscale Science and Technology, Ben-Gurion University of the Negev, Beer-Sheva 84105, Israel
*Corresponding author. Faculty of Natural Sciences, Life Sciences, Ben-Gurion University of the Negev, Beer-Sheva 84105, Israel. Tel: +972 8 647 9836; Fax: +972 8 647 9839; E-mail: imizrahi@bgu.ac.il; Department of Biochemistry, University of Zurich, Winterthurerstrasse 190, 8057 Zurich, Switzerland. Tel: +41 44 635 5522; E-mail: omedalia@bioc.uzh.ch
‡These authors contributed equally.

Abstract

Microbial taxonomy is critical for describing ecosystem composition, yet the link between taxonomy and properties of microbes, such as their cellular architecture, remains poorly defined. We hypothesized that the cellular architecture represents microbial niche adaptation. We used cryo-electron microscopy and tomography to analyze microbial morphology in order to associate cellular architecture with phylogeny and genomic contents. As a model system, we chose the core rumen microbiome and imaged a large isolate collection covering 90% of its richness at the order level. Based on quantifications of several morphological features, we found that the visual similarity of microbiota is significantly related to their phylogenetic distance. Up to the *Family* level, closely related microbes have similar cellular architectures, which are highly correlated with genome similarity. However, in more distantly related bacteria, the correlation both with taxonomy and genome similarity is lost. This is the first comprehensive study of microbial cellular architecture and our results highlight that structure remains an important parameter in classification of microorganisms, along with functional parameters such as metabolomics. Furthermore, the high-quality images presented in this study represent a reference database for the identification of bacteria in anaerobic ecosystems.

Keywords: rumen microbiome, cryo-electron microscopy, cryo-electron tomography, taxonomy, cellular architecture, gut microbiome, anaerobic microorganisms

Introduction

Bacterial species have historically been defined as clusters of similar organisms, combining both phenotype and common metabolic capabilities (American Society for Microbiology; Bergey and Breed 1957). Microbial systematics were later modified to combine phenotypic, genotypic, and sequence-based phylogenetic data within a framework of standards and guidelines for describing and identifying prokaryotes (Konstantinidis et al. 2006), but the definition of prokaryotic species remains controversial (Cohan 2002, Edgar 2018). The relationship between microbial taxonomy and function is sometimes vague and unclear (Taxis et al. 2015, Moraïs and Mizrahi 2019). Deepening our understanding of the similarity and dissimilarity of bacteria belonging to one taxon will help inform a better classification system, which is essential for the studies of microbial ecosystems and their unique compositions.

We hypothesize that the cellular architecture of a cell has evolved in response to a specific niche and environmental challenges. Deeper structural knowledge could thus serve as a proxy for function and ecological role. The conventional morphological classification used to describe microorganisms is still at the level of the one used by Antoni van Leeuwenhoek in 1676 using a very simplistic light microscope (Gest 2004). This description classifies all microorganisms into a few discrete shapes such as rods, spheres, and spirals. These visual phenotypes do not sufficiently reflect the uniqueness of phylogenetically diverse prokaryotes and are mostly limited by the low resolution of conventional brightfield light microscopy. Comprehensive taxonomy and genome databases cover many microbiome members, thus enabling omics-level discovery in complex biological samples. In contrast, no similar database of high-quality microbial images or structural parameters exists, making the identification of target species in mixed samples difficult.

Thus, we explored if the cellular architecture of microorganisms directly determined by cryogenic electron microscopy (cryo-EM) can expand our comprehension of microbial phylogeny and cell biology. In contrast to other techniques, cryo-EM allows the inspection of unstained biological materials in a near-native state at nanometer resolutions. Cryogenic electron tomography (cryo-ET) allows the retrieval of the 3D structure from a series of cryo-EM images acquired of the same object at multiple tilt angles. These projections are combined *in silico* to reconstruct its 3D architecture (Lucic et al. 2005, Irobalieva et al. 2016). Cryo-ET has opened a new window into the intra- and extracellular organization of prokaryotic cells, in some small bacteria even enabling a complete survey of the cells' contents (Oikonomou et al. 2016). It was used to uncover the architecture of chemoreceptor arrays (Briegel et al. 2012), contractile injections systems (Weiss et al. 2022), surface layers (von Kügelgen et al. 2020), and proteins supporting biofilm

Received: February 9, 2023. **Accepted:** March 22, 2023

formation (Melia et al. 2021). While cryo-ET has been instrumental in understanding key assemblies and features of prokaryotic cells, few studies have employed it to elucidate their emergence and diversity across taxonomies.

Comparative studies using cryo-ET highlight the potential of directly observing the outcomes of evolutionary processes. Chaban et al. (2018) compared subtomogram averages of flagellar motors in proteobacteria and traced the higher torque to an increased number of stator complexes in epsilonproteobacteria. Kaplan et al. (2021) surveyed tens of thousands of tomograms of 85 bacterial species to create a taxonomy of outer membrane (OM) projections in diderm bacteria. Recently, a team at Institut Pasteur combined genomics research with cryo-ET to reconstitute the emergence of monoderm bacteria through the loss of OM tethers (Witwinowski et al. 2022).

This study aims to provide a high-quality reference dataset of microbial cellular architecture to facilitate the design of future studies and the identification of species within biological samples. We use this dataset to investigate the link between phylogeny and cellular architecture, which we quantify using a set of five parameters. In line with previous studies, we illustrate the responsiveness of cellular architecture to environmental changes by growing a subset of three bacteria on several carbon sources. We show that the cellular architecture is relatively conserved within a genus, but not in higher taxonomic levels. In this context, we closely inspect members of the *Lachnospiraceae* family, which carry similar genes, but present different cellular architectures. We hypothesize that this divergence between phylogeny, genotype, and morphology reflects diverging ecological constraints, leading to specialized cellular architectures within a family or parallel emergence of similar features in unrelated phylogenies. Therefore, the cellular architecture remains an important factor that could deepen our understanding of microbial phylogeny, microbe–environment interactions in microbial ecology and microbial ecosystems.

Materials and methods

Bacterial culture

A collection of bacterial strains was assembled from commercially available culture collections (DSM and ATCC) or obtained from collaborations with USDA and the Rumen Microbial Genomics Network (Seshadri et al. 2018). The 69 bacterial strains were cultivated on their preferred media (Table S1, Supporting Information). Besides Brain Heart Infusion broth (Merck) and MRS (BD) that were purchased commercially, the various media were prepared as described previously, and included YCFA (Ze et al. 2012), M2 (Miyazaki et al. 1997), BY (Joblin 2005), and GS2 (Yoav et al. 2017). All media were supplemented with 1% (w/v) appropriate carbon sources as indicated in Table S1 (Supporting Information). The bacterial cells were grown anaerobically at 37°C for 24–48 h and the archaeal strains were incubated at 37°C until methane production was measurable. All strains reached stationary phase before cryo-EM sample preparation.

Whole-genome sequencing

The genomic DNA of *Agathobacter ruminis* DSM 29029, *Bacteroides cellulosilyticus* DSM 14838, *Bacteroides thetaiotaomicron* DSM 2079, *Bacteroides vulgatus* ATCC 8482, *Bacteroides caccae* JCM 9498, *Bacteroides uniformis* ATCC 8492, *Butyrivibrio fibrisolvens* CF3, *Blautia schinkii* DSM 10518, *Butyrivibrio sp.* DSM 10294, and *Bifidobacterium thermophilum* DSM 20209 were extracted using the phenol–chloroform method (Stevenson and Weimer 2007). The purified DNA was sequenced on an Illumina NovaSeq S1. The obtained sequences were assembled using SPAdes (Bankevich et al. 2012).

Cryo-EM sample preparation

For plunge-freezing, 1 ml of each culture was harvested, centrifuged for 5 min at 5000 *g*, washed briefly in TBS buffer, and resuspended in TBS prior to plunge-freezing. To minimize exposure to atmospheric oxygen, care was taken to plunge-freeze cells within 60 min of removing them from anaerobic culture. In parallel, 16S rRNA Sanger sequencing served to confirm the purity of the microbial cultures.

Cryo-EM samples were prepared by plunge-freezing into liquid ethane. A volume of 5 μl of bacteria diluted to OD = 0.15 in PBS were deposited on a glow-discharged Quantifoil R2/1 holey carbon grids (Quantifoil Micro Tools GmbH, Germany). A volume of 3 μl gold particles (10 nm BSA tracer, Aurion) concentrated 5-fold by centrifugation and resuspension in PBS were added. The grids were manually blotted for 4 s and vitrified by rapidly plunging into liquid ethane using a home-built plunging apparatus. The frozen samples were stored in liquid nitrogen until imaging.

Cryo-EM data acquisition

Grids carrying frozen-hydrated samples were clipped into AutoGrids (Thermo Fisher) and loaded into a Titan Krios microscope (FEI) equipped with an energy filter and a K2 direct electron detector (Gatan). Data acquisition was controlled by SerialEM 3.8 (Mastronarde 2003). All micrographs were collected in gain-corrected counting mode with a slit width of 20 eV. Overview montages at a magnification of 4800x (calibrated pixel size 3.4 nm) with defocus target −50 μm were used to measure diameter and shape of bacteria and guide further acquisition.

For suitably thin species, tilt series of dose-fractionated movies were acquired at 64 000x magnification (calibrated pixel size 0.221 nm) in a bidirectional scheme starting at −30° spanning tilts ±60° in 3° increments with −4–−6 μm defocus. For *Ruminococcus bromii*, tilt series were acquired in a dose-symmetric scheme (Hagen et al. 2017) from 0° to ±60° at 105 000x magnification (pixel size 0.137 nm). For species judged too dense for tilt series acquisition, single projections were acquired at a magnification of 11 500x (pixel size 1.30 nm) as dose-fractionated movies at a target defocus of −8 μm with a total dose of ∼ 20 e^-/A^2.

Data for *B. thetaiotaomicron* DSM 2079 on glucose, *B. uniformis* ATCC 8492, and *Corynebacterium vitaeruminis* Ga6A13 were collected using a Titan Krios G3i system (Thermo Fisher) equipped with an energy filter and a K3 direct electron detector (Gatan). Imaging parameters were chosen to parallel those on the original system: overview montages were acquired at 3600x magnification (pixel size 2.4 nm), tilt series were collected at 42 000x magnification (pixel size 0.216 nm). The projection dose was calibrated for each sample to yield a total dose of ∼140 e^-/A^2 for the tilt series (240 e^-/A^2 for *R. bromii*).

Cryo-EM data processing and visualization

Cryo-EM data was processed using the *imod* software package (Mastronarde and Held 2017). Montages were stitched using the tool *blendmont*, dose-fractionated movie frames were aligned using *alignframes* or the SerialEM SEMCCD plug-in. Tilt series were aligned based on fiducial markers, CTF-corrected by phase flipping, dose-weighted, and reconstructed via weighted back-projection in *etomo*. For figure preparation, tomograms were deconvolved using *tom_deconv.m* (Tegunov and Cramer 2019). Segmentations for Fig. 6 were created in Amira 2022.1 (Thermo

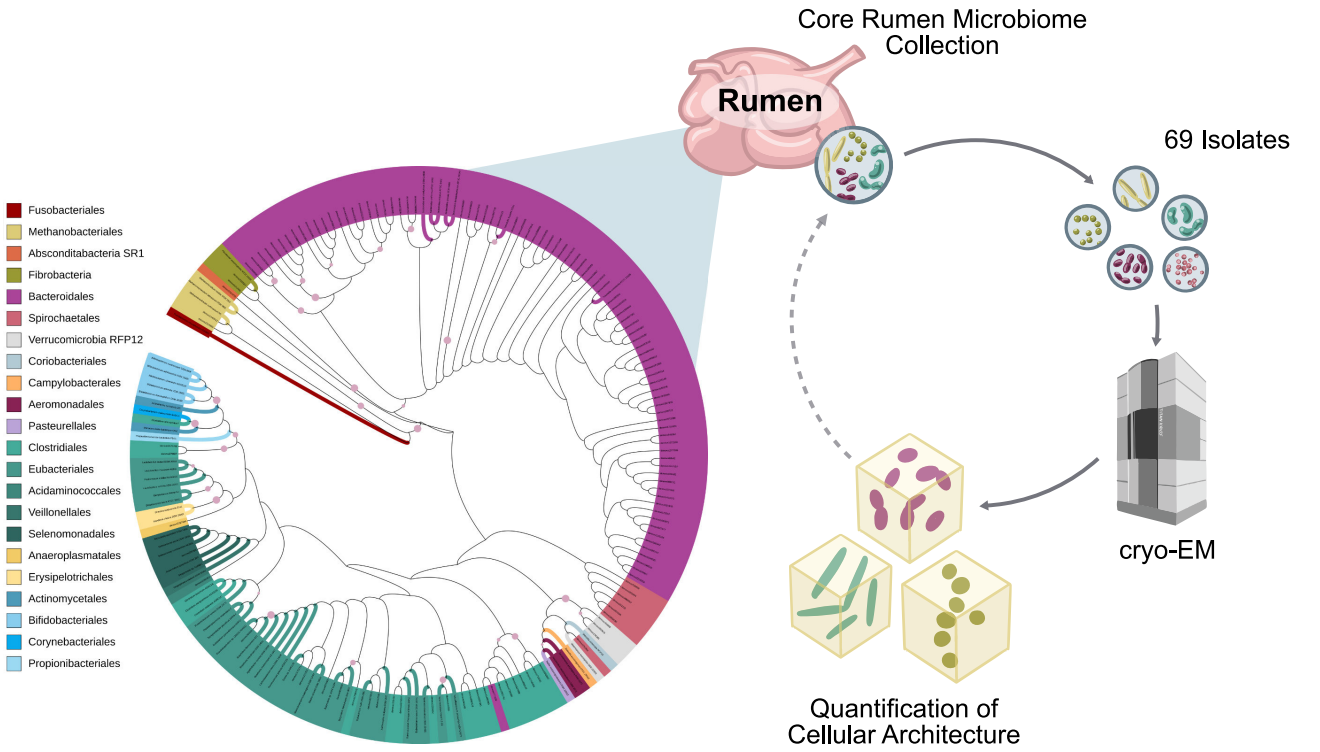

Figure 1. Visualizing the cellular architecture of core rumen microbiome members. To study the link between taxonomy and cellular architecture, the rumen core microbiome was selected as a model (phylogenetic tree, colored by order). A selection of 69 isolates (colored branches) was subjected to cryo-electron microscopy and subsequent analysis of cellular architecture. This analysis yields insights into the structural composition of the rumen ecosystem. The phylogenetic tree is based on the 16S rRNA gene of core microorganisms present in 80% of the individuals at the order level in a cohort of 1000 cows. Bootstraps with a confidence higher than 95% are displayed as pink circles. Phylogenetic tree generated using iTOL.

Fisher). Membranes and cell walls (CW) were presegmented using the *Membrane Enhancement Filter* and then manually refined. Flagella were segmented using the *Filament Tracer*. Other features were selected manually. Segmentation masks were then multiplied with tomogram densities to preserve the texture of features.

Imaging data analysis

Micrographs and tomograms were opened in FIJI (Schindelin et al. 2012) using the Bio-Formats Importer plugin. Length measurements were performed using the line measurement tool. Shape quantification was performed as follows: image thresholds were set so the CW or OM was clearly contrasted. Then, the boundary of the microbe was traced using the polygon selection tool. A spline was fit through the points. Then, the shape was measured and shape descriptors were extracted. Circularity, defined as $4\pi * \frac{area}{perimeter^2}$ was used as a shape descriptor. To quantify the number of cell bounding layers, a 10-nm slice was extracted from tomograms and a Gaussian filter with a radius of 1.75 px was applied. Then, a line profile was plotted and electron-dense layers were counted as local minima. Measurement values in Table S1 (Supporting Information) are reported as mean and standard deviation of 5–10 individual measurements for diameter and circularity or five measurements in three to five tomograms for CM—CW distance and CW thickness.

Statistical analysis

Statistical analysis was performed in R (version 4.2) (R Core Team 2018). Phylogenetic information was parsed from the NCBI taxonomy browser using the package taxize 0.9.99 (Chamberlain and Szöcs 2013). Distance matrices and ANOSIM were calculated using vegan 2.5–7 (Oksanen et al. 2020). For each bacterial strain, putative open reading frames (ORFs) were predicted using Prokka (Seemann 2014) and then Proteinortho (Lechner et al. 2011) was used to compare orthologous groups of proteins from all strains. Genomic distances between microbial genomes were calculated using Jaccard indices of their overall predicted orthologous proteins. Hierarchical clustering was performed using *hclust* using the ward.D2 algorithm and exported via ape 5.6–1 (Paradis and Schliep 2019) for visualization in itol (Letunic and Bork 2021).

Single-parameter comparisons for Figs 4 and 6 were calculated using GraphPad Prism 9.3.1 (GraphPad Software) as a Welch *t*-test. Multiple testing correction was performed with false discovery rate correction (Q = 1%) according to Benjamini, Krieger, and Yekutieli.

To compare cellular architecture to metabolomic profile, the metabolomics dataset published by the Sonneburg lab (Han et al. 2021) was downloaded from Zenodo (doi: 10.5281/zenodo.4890994). The four strains included in the preliminary analysis are *Marvinbryantia formatexigens* DSM 14469, *B. cellulosilyticus* DSM 14838, *B. caccae* ATCC 43185, and *Eubacterium siraeum* DSM 15702.

Results

The core rumen microbiome as a model for microbial structural characterization

The microbial population of the rumen exhibits a plethora of cellular morphologies. To make this diversity accessible to future studies of microbial ecosystems, we cataloged it by cryo-EM and -ET. We selected the core rumen microbiome as a model and gathered a large microbial isolate collection covering 90% of the core

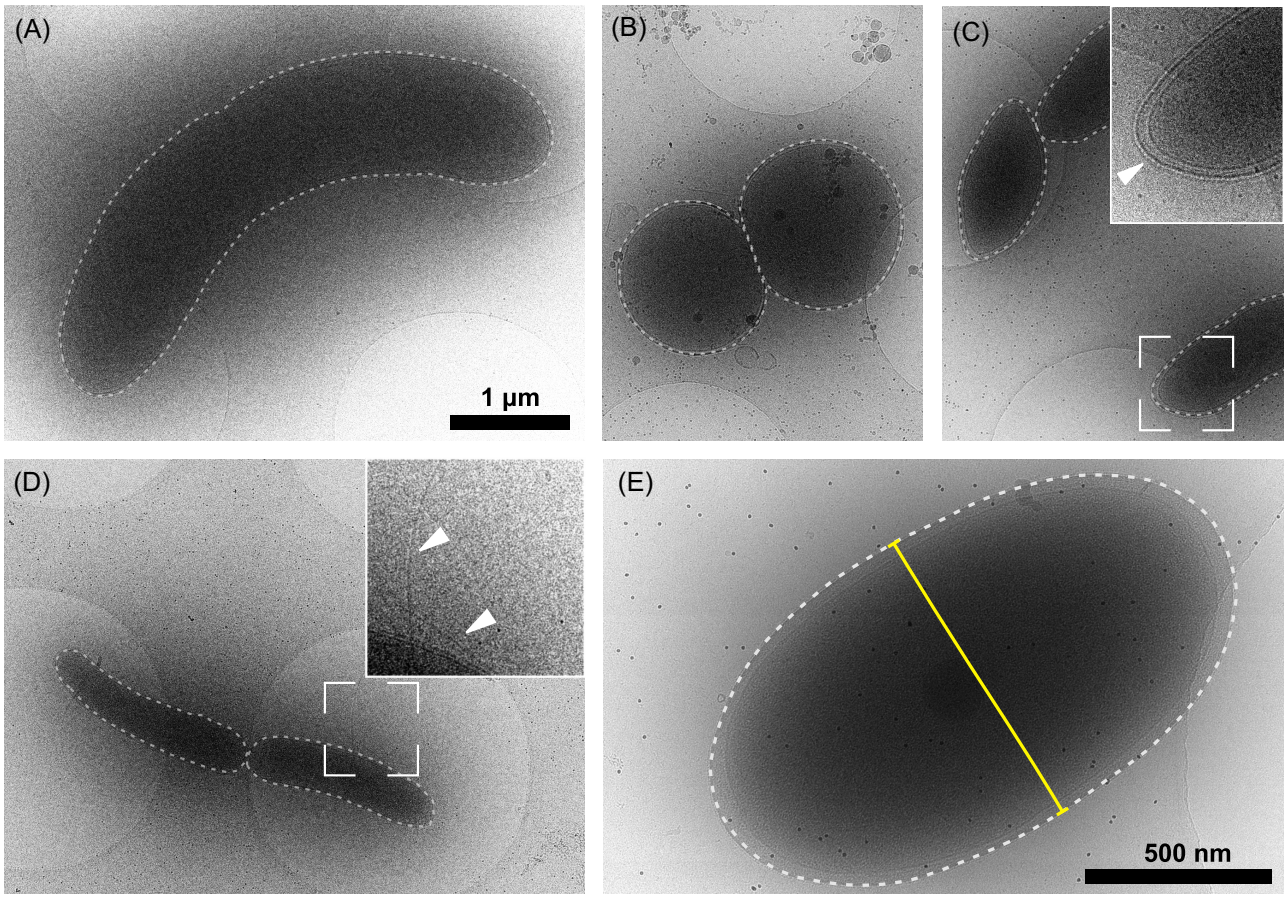

Figure 2. Bacteria of the core microbiome span a plethora of morphologies. Selection of representative micrographs of microbiome members. **(A)** *Selenomonas ruminantium* AB3002, a diderm rod-shaped bacterium isolated from the bovine rumen. **(B)** *Succinimonas amylolytica* DSM 2873, a diderm coccoid bacterium isolated from the bovine rumen. **(C)** *Prevotella ruminicola* ATCC 19189, a diderm ovoid bacterium isolated from the bovine rumen. Here, an S-layer is visible in the micrographs (insert, arrow). **(D)** *A. ruminis* DSM 29029, a monoderm vibrioid bacterium isolated from sheep rumen. Flagella are visible in the micrographs (insert, arrows). **(E)** High-exposure micrograph of *Akkermansia muciniphila* DSM 22959, an important player in the human intestinal microbiome linked to metabolic benefits in the early developing cow rumen (Furman et al. 2020, Ouyang et al. 2020). To compare shapes, the outline was manually traced (dotted line) and the circularity of the shape calculated in FIJI. Furthermore, the diameter was measured (yellow line). Panels (A)–(D) at the same magnification.

rumen microbiome diversity at the order level (microorganisms present in 80% of a 1000 cow cohort; Mizrahi et al. 2021), encompassing 45 genera in eight phyla (Fig. 1; Figure S1, Supporting Information). To probe the relationship between environment and cellular architecture, we diversified the carbon sources in the cultures of some microbes, resulting in 75 samples overall. In addition, we performed new whole-genome sequencing on 10 of the cultured strains where genomes were not available.

Sample thickness is a major limitation in cryo-EM. To maintain a signal-to-noise ratio (SNR) amenable to interpretation, we restricted ourselves to 2D cryo-EM for 29 strains and performed cryo-ET of 38 bacterial and two archaeal strains (seven phyla covering 55% of the core rumen microbiome diversity at order level). However, analysis levels did not cluster with taxonomy yielding a proper coverage of all phylogenies in the cryo-ET dataset (Figure S1, Supporting Information).

The core rumen microbiome spans a wide range of cell shapes and sizes

To gain an initial overview of the morphological diversity within the collection, we acquired ca. 4500 electron micrographs at magnifications between 3600x and 11 500x (examples in Fig. 2; Figure S2, Supporting Information). In these micrographs, the dimensions and shapes of each strain could be determined.

A total of 45 strains (65%) were classified as rod-shaped (e.g. *Selenomonas ruminantium* AB3002, Fig. 2A), eight strains (12%) as coccoid (e.g. *Succinimonas amylolytica* DSM 2873, Fig. 2B), 10 strains (14%) as ovoid (e.g. *Prevotella ruminicola* ATCC 19189, Fig. 2C), and six strains (9%) as vibrioid (e.g. *A. ruminis* DSM 29029, Fig. 2D). For some samples, further structural features were observed at this magnification, such as S-layers (Fig. 2C, insert) or flagella (Fig. 2D, insert).

To compare the cellular architecture across samples, the diameter and shape were quantified as outlined for *Akkermansia muciniphila* DSM 22959 in Fig. 2(E) (full results in Table S1, Supporting Information). The diameter was measured at the center of the individual cell, orthogonal to the long axis. The average diameter of the strains in the screening collection was 0.8 µm (range 0.38–1.81 µm). The shape was quantified by tracing the cell outline manually, fitting the points with a spline curve and then calculating the circularity, which provides a continuous description of shape that highly correlates with manual shape classification (ANOVA, $F = 58.39$, $P < 10^{-4}$).

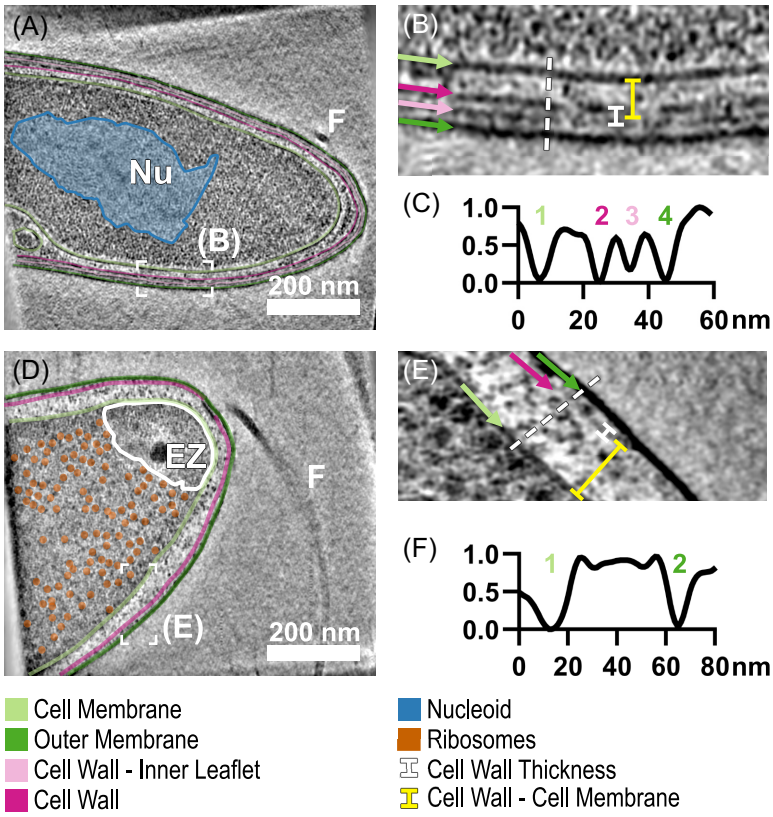

Figure 3. Cryo-ET elucidates the cellular architecture. Central cross-sections through tomograms of bacteria reveal the cellular architecture. **(A)** *Pseudobutyrivibrio sp.* LB 2011 contains a dense cytoplasm, the nucleoid DNA stands out by its stringy texture (blue overlay, Nu). Cell membrane (CM, light green), cell wall (CW, two leaflets in pink), and outer membrane (OM, dark green) are highlighted. A flagellum **(F)** is visible next to the cell. **(B)** Closer inspection of the cell surface reveals a diderm bacterium with a CW in two leaflets (pink), membranes are highlighted in light green (CM) and dark green (OM). The CW thickness (white line) and distance from the CM (yellow line) were measured as indicated. **(C)** A line profile through the cellular boundary clearly shows four electron-dense layers. **(D)** A central cross-section through a *Wolinella sp.* ATCC 33567 cell. Again, CM, CW, and OM are traced. Ribosomes are visible throughout the cytoplasm (orange), with the exception of a polar exclusion zone (white outline, EZ). A flagellum emerging from the tip is visible next to the cell **(F)**. **(E)** CM and OM are clearly visible bordering an electron-light periplasm. The CW is visible adjacent to the OM as a faint band. Measurements were performed as indicated. **(F)** The line profile plot shows that only CM and OM have significant electron density. Slice thickness 10 nm.

Phylogenies can be distinguished using five measures of cellular architecture

To analyze the cellular architecture in 3D, we applied cryo-ET to all samples with suitable thickness, i.e. < 600 nm in edge regions. Obvious features were manually annotated using reference images from the Cell Structure Atlas online textbook (Oikonomou and Jensen 2021). Tomograms from *Pseudobutyrivibrio sp.* LB 2011 and *Wolinella sp.* ATCC 33567 are shown in Fig. 3 (representative slices from all samples in Figure S2, Supporting Information).

In a central x–y section of the tomogram of *Pseudobutyrivibrio sp.* (Fig. 3A), a dense cytoplasm enclosed by the cell membrane (CM, light green) could easily be distinguished from the lighter periplasm framed by the OM (dark green). The cell surface organization was inspected carefully. In the case of the *Pseudobutyrivibrio sp.*, four electron-dense continuous layers were observed outside the CM (Fig. 3B, light green): a 13.4 ± 0.8 nm thick CW consisting of two parallel layers (pink) at a distance of 14.6 ± 0.6 nm and an OM (dark green) 28.8 ± 1.5 nm removed from the CM. The line profile of the electron density across the cell exterior confirms the presence of these four cell-enclosing layers (Fig. 3D).

For *Wolinella sp.*, cyto- and periplasm again are easily distinguishable (Fig. 3D). In the cytoplasm, ribosomes are ubiquitous (orange), with the exception of a polar exclusion zone (EZ, white outline), which contains the flagellar motor and a chemoreceptor array above the viewing plane. The periplasm itself is much less dense than that of *Pseudobutyrivibrio sp.*, and contains only one very light continuous CW layer close to the OM (Fig. 3D), best visible near the flagellar motor. A line profile in fact only picks up the cell and OMs. Even though macromolecular complexes such as flagella, chemoreceptor arrays, ribosomes, and nucleoids were frequently observed in the tomograms, we decided against describing species based on the presence or absence of specific complexes as only a part of the cell volume can be surveyed using tomography.

To compare cellular architectures across samples, we thus determined three parameters related to the surface structure of the cells: the number of cell-enclosing layers as detected by line profile, the CW thickness, and distance of the CW from the CM. Additionally, the circularity and diameter were included in a set of five cellular architecture measurements. Per sample, we included five replicate sets of measurements (data from five micrographs and three to five tomograms). This number was chosen to represent all samples equally in the dataset; fewer tomograms could be acquired on challenging samples. Where more measurements were available, we confirmed that the set included in the analysis did not significantly differ from the full set (data not shown).

To quantify the dissimilarity in microbial morphology based on our five cellular architecture parameters, we calculated the Eu-

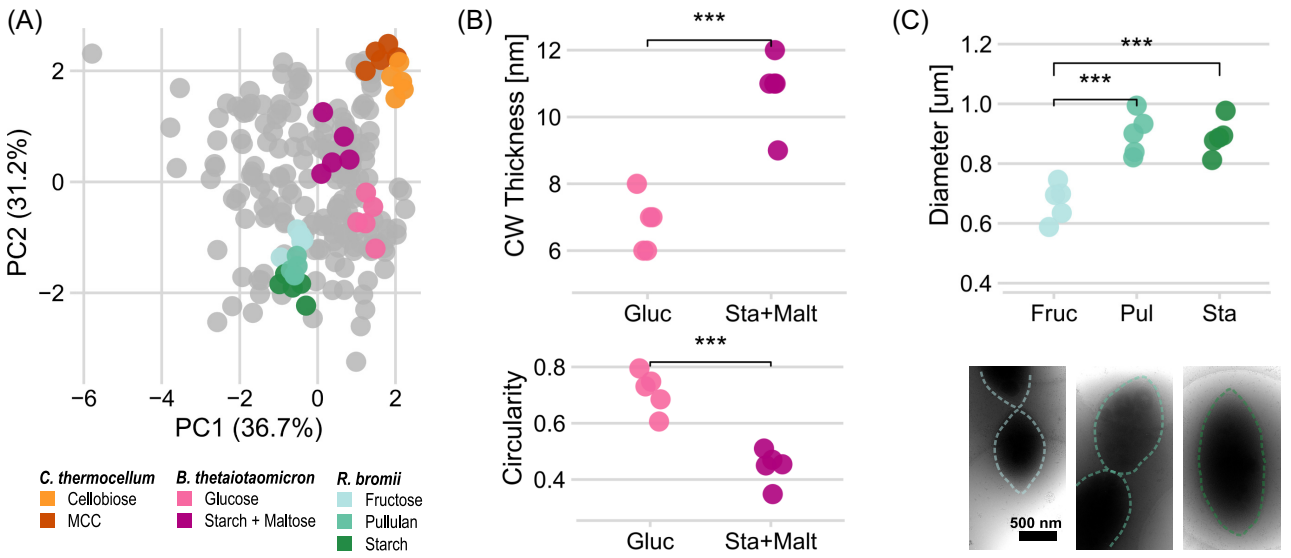

Figure 4. Cellular architecture reflects changing carbon sources. (A) PCA of the five parameters diameter, circularity, number of electron-dense bounding layers, CW thickness, and CW distance from the CM. Replicates for all 45 samples with associated 3D data are shown. Highlighted are measurements for strains *C. thermocellum* DSM 1313, *B. thetaiotaomicron* DSM 2079, and *R. bromii* L2-63 grown on simple or complex carbon sources. **(B)** Further analysis shows that the CW of *B. thetaiotaomicron* DSM 2079 is significantly thicker when grown on starch/maltose (unpaired *t*-test, $P < .001$), and that the cells are less circular ($P < .001$). (C) For *R. bromii* L2-63, the cell diameter is higher when grown on complex carbon sources starch or pullulan compared to cells grown on fructose (unpaired *t*-tests, both $P < .001$). Example micrographs shown below.

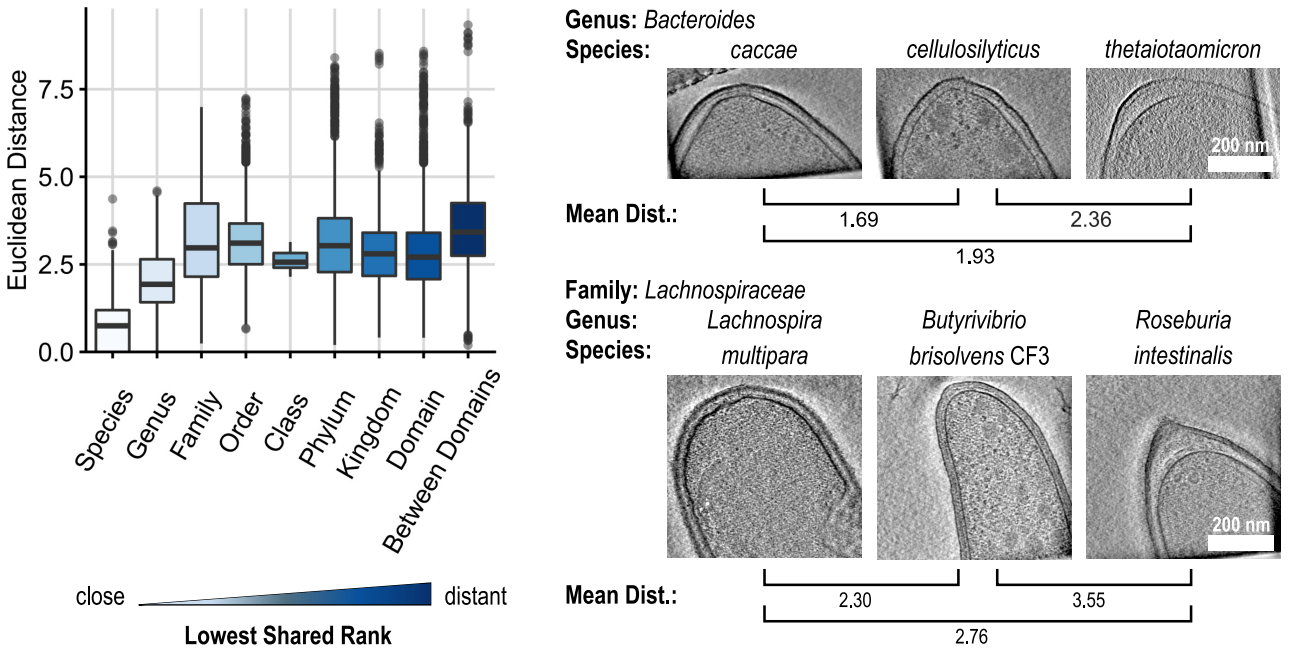

Figure 5. Visual diversity emerges in the phylogenetic branches. The Euclidean distance between each pair of replicates as a function of their lowest shared taxonomic rank. Samples within a species show a conserved cellular architecture, as do those within a genus. Starting at the family level, the cellular architectures do not become more diverse within the same domain. As an example, tomogram slices from three representative strains in the genus *Bacteroides* are shown on the top right: *B. caccae* DSM 19024, *B. cellulosilyticus* DSM 14838, and *B. thetaiotaomicron* DSM 2079. All three present as rod-shaped diderm bacteria with thin CWs. Correspondingly, the mean Euclidean distances between them are low. In contrast, three members of the *Lachnospiraceae* family show little resemblance and correspondingly higher Euclidean distances: *Lachnospira multipara* ATCC 19207, *B. fibrisolvens* CF3, and *Roseburia intestinalis* DSM 14610.

clidean distance between the 5D vectors defined by each replicate. To account for the different scales, ranges, and units of the parameters, measurements within each parameter were centered to a mean of 0 and scaled to a standard deviation of 1.

We assessed the robustness and specificity of the measurements by comparing the similarity between replicates within a sample to the similarity between samples as given by the analysis of similarities (ANOSIM) R. Replicate measurements exhibited

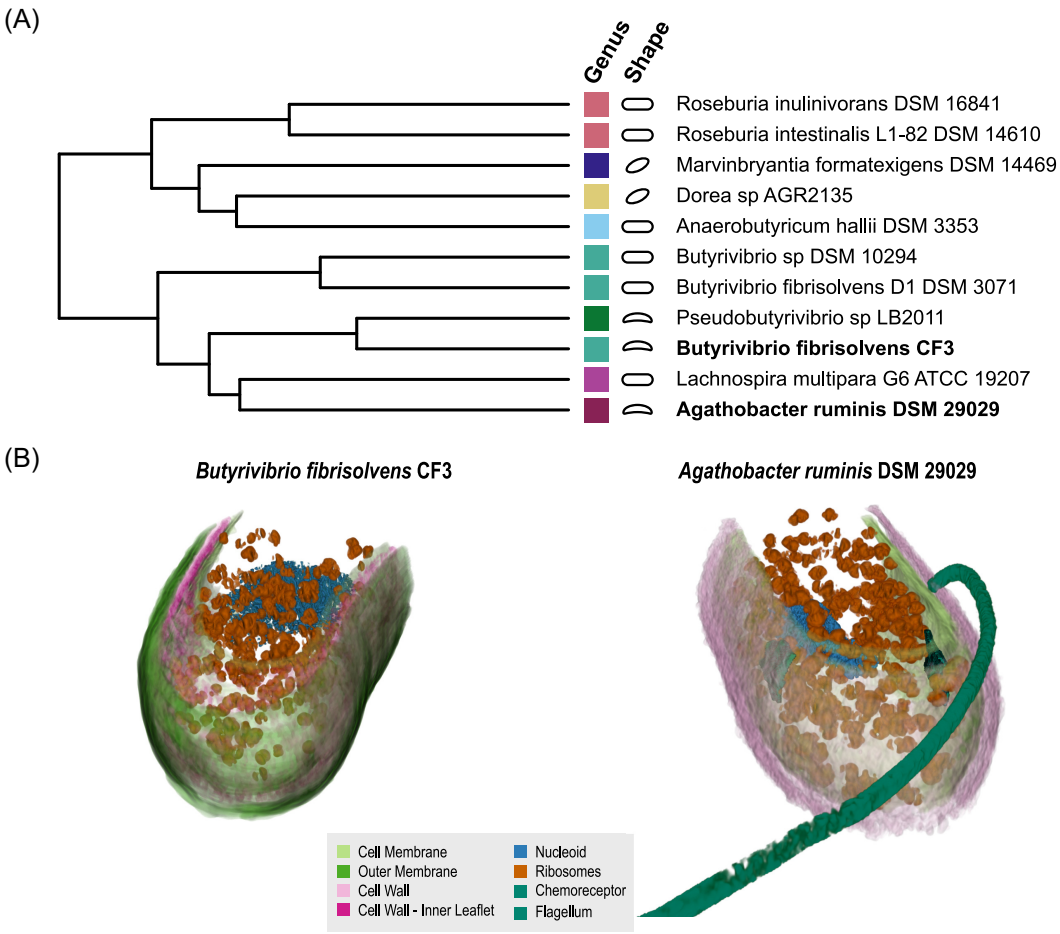

Figure 6. Hierarchical clustering and 3D rendering reveal divergence between genome and morphology. (A) All members of the *Lachnospiraceae* family were hierarchically clustered by genome contents. Members of the same genus frequently occupy neighboring branches, indicating a reliable link between taxonomy and genome contents at this level. Basic cell shapes are indicated by icon, either rod, ovoid, or vibrioid. **(B)** 3D renderings of *B. fibrisolvens* CF3 and *A. ruminis* cells. The divergent cell architectures are immediately visible: *B. fibrisolvens* is bound by a CM, a double-layered CW (pink), and an OM (dark green), while *A. ruminis* only carries one CW layer (pink) on top of the CM (light green). In both cells, the cytoplasm contains a high density of presumptive ribosomes (orange), as well as a nucleoid (blue). *A. ruminis* additionally contains two chemoreceptor arrays (teal), as well a flagellum in the extracellular space (teal).

a high degree of similarity ($R = 0.923$, $P < 10^{-4}$), indicating that our set of five parameters achieves a meaningful separation between samples.

The cellular architecture is responsive to environmental changes

In the rumen microenvironment, the presence of specific carbohydrates is a determinant for niche availability. It changes protein expression profiles of microbiota (Honan and Greenwood 2020) and affects community composition in the long term (Furman et al. 2020). If the cellular architecture of a cell is indeed determined by niche adaptation, it should thus reflect changes in environmental conditions such as the nutrient availability. Indeed, structural alterations in response to environmental changes have been shown for other bacteria, such as *Vibrio cholerae* in changing osmotic environments (Brenzinger et al. 2019).

We cultivated the human gut bacteria *B. thetaiotaomicron* DSM 2079 and *R. bromii* L2-63, as well as the environmental isolate *Clostridium thermocellum* DSM 1313 either on a simple carbon source, i.e. glucose, fructose, or cellobiose, or their preferred polysaccharide starch, pullulan, or cellulose. The cellular architecture was quantified using five parameters as described above, followed by principal component analysis (PCA) across all measurements. This analysis indicated that changes in carbon source correspond to overall structural variations for all the three strains that were analyzed, pointing to a connection between niche availability and cellular architecture (Fig. 4A). Interestingly, there was no unified structural response among all species. In the case of *B. thetaiotaomicron* DSM 2079, a change from glucose to starch/maltose was associated with a thicker CW (Fig. 4B, 6.8 ± 0.8 nm vs. 10.8 ± 1.1 nm, $P < .001$) and a more elongated, less circular shape (0.713 ± 0.072 vs. 0.446 ± 0.060, $P < .001$). On the other hand, cells of *R. bromii* L2-63 grown on fructose exhibited a significantly lower diameter than those grown on either pullulan or starch (Fig. 4C, 0.673 ± 0.06 um, 0.898 ± 0.07 um and 0.889 ± 0.06 um, respectively, $P < .001$), while the diameter was matched between pullulan and starch ($P = .846$).

The alterations in morphology of *C. thermocellum* DSM 1313 were nonsignificant for each individual parameter that was measured. The spread observed in the PCA might thus represent a combination of more subtle adaptations, such as the changes in the cellulose degradation machinery we described in an earlier publication (Tatli et al. 2022).

Structural diversity arises in the phylogenetic branches

After establishing, that the morphological similarity between prokaryotic cells can be expressed as the Euclidean distance between five parameters, we asked whether the cellular architecture of a strain is linked to its phylogeny. Through ANOSIM, we found that replicates within the same strain or species show high degrees of structural similarity ($R = 0.912$ and $R = 0.903$, respectively, both $P < 10^{-4}$), some of which seems to be conserved within a genus ($R = 0.731$, $P < 10^{-4}$). However, in-group similarity seems to be negligible at the family level ($R = 0.222$, $P < 10^{-4}$) and insignificant above.

To evaluate this on a per-sample level, we plotted the similarity score for each combination of two measurements over their respective shared taxonomic rank (Fig. 5, left). The resulting graph clearly shows that cells belonging to the same species usually have very similar cellular architectures, which significantly diverge for those sharing only the genus. Representative tomogram slices of three *Bacteroides* strains illustrate a conserved cellular architecture within a genus, in this case rounded cell ends, a diderm architecture and thin CWs. On the contrary, three members of different genera in the family *Lachnospiraceae* exhibit distinct morphologies (Fig. 5, right), with between two and four cell-bounding layers and drastically varying cell diameters. The link between shared taxonomic rank and structural similarity breaks above the family level; two strains belonging to the same order are on average as similar as those sharing only the domain.

From this, we conclude that the diverse cellular architectures of microbiome members, which we highlight in Fig. 2, do not represent ancient branching events, but rather recent specialization. To evaluate the genomic changes underlying these diverse cellular architectures, we additionally considered the genes present in each strain. Based on 59 publicly available and 10 genome sequences published with this study, we calculated the genome similarity between each two strains as the Jaccard index of their overall predicted orthologous proteins. For all pairs, which share the species, genus, or family, structural and genome similarity are highly correlated (Pearson's $r = 0.73$, $P < 10^{-15}$). For those more distantly related, no correlation could be found ($r = -0.04$, $P < 10^{-10}$).

In-family heterogeneity hints at diverse ecological roles

To understand, why the link between cellular architecture and taxonomy breaks around the family level, we closely inspected the most abundant family in our dataset, the *Lachnospiraceae*, to probe the relationship between genome contents and cellular architecture. For its members, the correlation between shared orthologous proteins and structural similarity closely matches that of the entire dataset ($r = 0.71$, $P < 10^{-15}$). We first hierarchically clustered all strains within the family by genome contents. As expected, species of the same genus share many genes and thus frequently occupy neighboring branches. Strains belonging to the same clade frequently share a basic cell shape (Fig. 6A). To investigate, how genome contents and taxonomy diverge, we closely investigated the clade containing the genera *Butyrivibrio*, *Pseudobutyrivibrio*, *Agathobacter*, and *Lachnospira*. While *Butyrivibrio sp.* and *Butyrivibrio fibrisolvens* D1 cluster together by genome contents, *B. fibrisolvens* CF3 diverges and is genetically more similar to *A. ruminis* than to another strain from its own species.

To understand the implications of this divergence between taxonomy and genome contents, we created 3D segmentations of *B. fibrisolvens* CF3 and *A. ruminis* DSM 29029 (Fig. 6B, Supplementary videos 1 and 2). The diameter (*A. ruminis*: 0.377 ± 0.02 μm, *B. fibrisolvens*: 0.398 ± 0.01 μm) and circularity (0.407 ± 0.062 and 0.429 ± 0.108) are closely matched between these strains. However, *A. ruminis* has only two cell-bounding layers (Fig. 6B, CM in light green and CW in pink), while *B. fibrisolvens* has four with a two-layered CW (similar to that of *Pseudobutyrivibrio sp.* in Fig. 3). Correspondingly, *A. ruminis* has a significantly thicker CW than *B. fibrisolvens* (18.8 ± 1.6 nm vs. 11.4 ± 0.8 nm, $P < .001$). In both *A. ruminis* and *B. fibrisolvens* cells, a high number of presumptive ribosomes (orange) could be observed around a nucleoid (blue). In *A. ruminis*, a flagellum was observed connected to the cell, along with two chemoreceptor arrays (both teal) adjacent and opposite to the observed flagellar attachment site. The unique features of both strains are well-known examples of structure–function relationships in microbiomes: flagella provide motility and access to distant nutrients (Ferreira et al. 2019). These divergent attributes in closely related strains could thus be the outcome of ecological constraints.

Discussion

In this study, we provide a dataset of high-resolution 2D and 3D cryo-EM data of cardinal rumen microbiome members, to facilitate the design of future studies and the identification of species in mixed samples. We used a combination of genomics and statistical analysis to answer the question whether the cellular architecture of a microbe is directly related to its phylogeny. For most bacteria in our collection, no high-quality electron micrographs were previously publicly available.

Cryo-EM provides a detailed 2D description of individual cells, enriched by cryo-ET with 3D information on the organization of the CW and membrane, as well as the crowded cell interior. Even though cryo-EM is clearly the method of choice for researchers to describe cellular architecture in unstained samples, the access to the required instruments is a limiting factor in its application. This study required ~60 microscope-days for data acquisition. Recently published advances, including faster data acquisition (Eisenstein et al. 2022) and automated tilt-series alignment (Zheng et al. 2022) will help alleviate time constraints, along with automated sample thinning methods (Klumpe et al. 2021) making more strains amenable to analysis. This study is unique in the application of cryo-EM and -ET to study a large collection of isolates representing most of the diversity of the core rumen microbiome, thus bringing electron microscopy on the way toward omics level analysis. Our data complements a collection of bacterial and archaeal tomograms that were published by Grant Jensen and coworkers (https://etdb.caltech.edu).

This study provides general insights into variations of the cellular architecture of gut microorganisms. We show that measuring five parameters (diameter, circularity, number of cell-bounding layers, CW thickness, and CW–CM distance) can define the cellular architecture of a strain with a high specificity among 40 tested strains. While taxonomy and morphology are correlated below the family level, they diverge afterwards. We argue that structure is likely linked to ecological function and niche, and thus selecting structurally diverse strains may be advantageous when reconstituting a representative ecosystem.

Along with electron micrographs, we provide a table of the five parameters we used to quantify the cellular architecture. These values can be cross-referenced when visualizing prokaryotes using more widely accessible methods, such as light microscopy and conventional transmission and scanning electron microscopy.

While our measurements appear to consistently describe the cellular architecture in our experiments, they should be regarded only as reference values to detect changes in response to the cellular environment and not as a universal constant. The cellular architecture of bacteria is frequently remodeled in response to a wide range of stimuli, such as changes in the osmolarity of the environment, growth density, biofilm formation, and biotic interactions (Ramijan et al. 2018, Brenzinger et al. 2019). In this study, we show that remodeling can also happen in response to a change in the carbon source of a bacterium under otherwise identical growth conditions. We, thus used single-isolate cultures in a similar growth phase for our comparisons between strains, to reduce sources of variation. In future studies, systematically probing the effects of growth density, osmolarity, and biotic interactions on the cellular architecture of a wider number of species could help identify common mechanisms of niche adaptation across the tree of bacteria. This would require a novel pipeline including more automated microscopy and image analysis, to include a higher number of individual cells.

When comparing microbes' structural similarity to their lowest common taxonomic rank, we observed a steady decrease in similarity for relationships up to the family level, above which no link is discernible. Remarkably, this pattern was also observed for metabolic similarities by Han et al. (2021). This indicates that the current taxonomy of bacteria accurately portraits local relationships, but not global ones. A preliminary analysis of the four strains included both in our study and the metabolomics dataset shows a high degree of correlation between the cellular architecture and metabolomic phenotype as indicated by a Mantel test, albeit not reaching statistical significance of 0.05 due to the limited number of possible permutations ($r = 0.860$, $P = .08$).

The microbes imaged for this study are also present in other gut ecosystems, such as the human gut microbiome (Atherly and Ziemer 2014, Mukhopadhya et al. 2018). Therefore, this dataset may serve as a basis to greater understanding of the structure–function relationships among microorganisms and enable future structural studies of reconstituted multispecies microbiomes. In conclusion, we observed a divergence between taxonomy and cellular architecture above the family level, which highlights possible challenges with phylogeny-based clustering of microorganisms. The preliminary link between metabolomics and cellular architecture identified here in a small subset studied warrants further exploration. If indeed morphology and metabolic function are correlated, this will represent a large step forward towards understanding the structure–function relationship of microbiomes. Ultimately, we imagine a classification system for microorganisms, which does not only take 16S-derived phylogenetic information into account but could also encompass metabolomic information and high-resolution structural information from cryo-EM/ET. Such a classification system will help advance our understanding of the relationships within important microbial ecosystems.

Data and code availability

The tomograms used for the preparation of main text (Figs. 3, 5, and 6) have been uploaded to the Electron Microscopy Data Bank (EMDB) with the following accession codes: Fig. 3 EMD-15448 (*Pseudobutyrivibrio sp.*) and EMD-15449 (*Wolinella sp.*). Figure 5 EMD-15447 (*B. fibrisolvens* CF3), EMD-16474 (*B. caccae*), EMD-16477 (*B. cellulosilyticus*), EMD-16478 (*L. multipara*), EMD-16479 (*R. intestinalis*), and EMD-16488 (*B. thetaiotaomicron*). Figure 6 EMD-15447 (*B. fibrisolvens* CF3) and EMD-15445 (*A. ruminis*). Genome assemblies were deposited in GenBank (BioProject PRJNA867738). Data and R code used for the statistical analysis are available on GitHub at https://github.com/bwmr/microbiome_morphology. Images of all samples in Figure S2 (Supporting Information) are available on Zenodo https://doi.org/10.5281/zenodo.6874893.

Acknowledgments

This work was funded by the DIP (2476/2 –1) to I.M. and O.M., ERC (866530) to I.M., ISF (1947/19) to I.M. and S.M., and the Swiss National Foundation (31003_207453) to O.M. The authors thank Dr Meltem Tatli for providing the micrographs and tomograms of *C. thermocellum* on MCC and cellobiose and Daphne Perlman for assisting with graphic design. We acknowledge bacterial samples kindly provided by the USDA, the Hungate Collection (Dr Graeme T Attwood and Kerri Reilly), the Rumen Microbial Genomics Network, Claudia Moresi (ETH Zurich, *Bacteroides thetaiotaomicron* DSM 2079 on BHI-Hemin), and Professor Magnus Øverlie Arntzen (NMBU Norway, *Fibrobacter succinogenes* S85 glycerol stocks). Cryo-electron microscopy imaging was performed with equipment maintained by the Center for Microscopy and Image Analysis (ZMB), University of Zurich. The authors thank Ariane Briegel and Uri Gophna for their constructive comments during the peer review process.

Supplementary data

Supplementary data is available at *FEMSML* online.

Conflict of interest statement. None declared.

References

American Society for Microbiology, Bergey DH, Breed RS. *Bergey's Manual of Determinative Bacteriology, by Robert S. Breed [and Others]*. Baltimore: Williams & Wilkins Co., 1957.

Atherly T, Ziemer CJ. Bacteroides isolated from four mammalian hosts lack host-specific 16S rRNA gene phylogeny and carbon and nitrogen utilization patterns. *Microbiologyopen* 2014;**3**:225–38.

Bankevich A, Nurk S, Antipov D *et al.* SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *J Comput Biol* 2012;**19**:455–77.

Brenzinger S, van der Aart LT, van Wezel GP *et al.* Structural and proteomic changes in viable but non-culturable *Vibrio cholerae*. *Front Microbiol* 2019;**10**:793.

Briegel A, Li X, Bilwes AM *et al.* Bacterial chemoreceptor arrays are hexagonally packed trimers of receptor dimers networked by rings of kinase and coupling proteins. *Proc Natl Acad Sci USA* 2012;**109**:3766–71.

Chaban B, Coleman I, Beeby M. Evolution of higher torque in *Campylobacter*-type bacterial flagellar motors. *Sci Rep* 2018;**8**:97.

Chamberlain SA, Szöcs E. taxize: taxonomic search and retrieval in R. *F1000Res* 2013;**2**:191.

Cohan FM. What are bacterial species?. *Annu Rev Microbiol* 2002;**56**:457–87.

Edgar RC. Updating the 97% identity threshold for 16S ribosomal RNA OTUs. *Bioinformatics* 2018;**34**:2371–5.

Eisenstein F, Yanagisawa H, Kashihara H *et al.* Parallel cryo electron tomography on in situ lamellae. *Nat Methods* 2023;**20**:131–8.

Ferreira JL, Gao FZ, Rossmann FM *et al.* γ-proteobacteria eject their polar flagella under nutrient depletion, retaining flagellar motor relic structures. *PLoS Biol* 2019;**17**:e3000165.

Furman O, Shenhav L, Sasson G *et al.* Stochasticity constrained by deterministic effects of diet and age drive rumen microbiome as-

sembly dynamics. *Nat Commun* 2020;**11**. doi:10.1038/s41467-020-15652-8.

Gest H. The discovery of microorganisms by Robert Hooke and Antoni Van Leeuwenhoek, fellows of the Royal Society. *Notes Rec R Soc Lond* 2004;**58**:187–201.

Hagen WJH, Wan W, Briggs JAG. Implementation of a cryo-electron tomography tilt-scheme optimized for high resolution subtomogram averaging. *J Struct Biol* 2017;**197**:191–8.

Han S, Van Treuren W, Fischer CR *et al.* A metabolomics pipeline for the mechanistic interrogation of the gut microbiome. *Nature* 2021;**595**:415–20.

Honan MC, Greenwood SL. Characterization of variations within the rumen metaproteome of Holstein dairy cattle relative to morning feed offering. *Sci Rep* 2020;**10**:3179.

Irobalieva RN, Martins B, Medalia O. Cellular structural biology as revealed by cryo-electron tomography. *J Cell Sci* 2016;**129**:469–76.

Joblin KN. Methanogenic archaea. In: Makkar HPS, McSweeney CS (eds), *Methods in Gut Microbial Ecology for Ruminants*. Dordrecht: Springer, 2005, 47–53.

Kaplan M, Chreifi G, Metskas LA *et al.* In situ imaging of bacterial outer membrane projections and associated protein complexes using electron cryo-tomography. *Elife* 2021;**10**. doi:10.7554/eLife.73099.

Klumpe S, Fung HK, Goetz SK *et al.* A modular platform for automated cryo-FIB workflows. *Elife* 2021;**10**. doi:10.7554/eLife.70506.

Konstantinidis KT, Ramette A, Tiedje JM. The bacterial species definition in the genomic era. *Philos Trans R Soc Lond B Biol Sci* 2006;**361**:1929–40.

von Kügelgen A, Tang H, Hardy GG *et al.* In situ structure of an intact lipopolysaccharide-bound bacterial surface layer. *Cell* 2020;**180**:348–58.e15.

Lechner M, Findeiß S, Steiner L *et al.* Proteinortho: detection of (co-)orthologs in large-scale analysis. *BMC Bioinf* 2011;**12**:124.

Letunic I, Bork P. Interactive Tree of Life (iTOL) v5: an online tool for phylogenetic tree display and annotation. *Nucleic Acids Res* 2021;**49**:W293–6.

Lucic V, Förster F, Baumeister W. Structural studies by electron tomography: from cells to molecules. *Annu Rev Biochem* 2005;**74**:833–65.

Mastronarde DN. SerialEM: a program for automated tilt series acquisition on Tecnai microscopes using prediction of specimen position. *Microsc Microanal* 2003;**9**:1182–3.

Mastronarde DN, Held SR Automated tilt series alignment and tomographic reconstruction in IMOD. *J Struct Biol* 2017;**197**:102–13.

Melia CE, Bolla JR, Katharios-Lanwermeyer S *et al.* Architecture of cell-cell junctions in situ reveals a mechanism for bacterial biofilm inhibition. *Proc Natl Acad Sci USA* 2021;**118**. doi:10.1073/pnas.2109940118.

Miyazaki K, Martin JC, Marinsek-Logar R *et al.* Degradation and utilization of xylans by the rumen anaerobe *Prevotella bryantii* (formerly *P. ruminicola subsp.brevis*) B14. *Anaerobe* 1997;**3**:373–81.

Mizrahi I, Wallace RJ, Moraïs S. The rumen microbiome: balancing food security and environmental impacts. *Nat Rev Microbiol* 2021;**19**:553–66.

Moraïs S, Mizrahi I. The road not taken: the rumen microbiome, functional groups, and community states. *Trends Microbiol* 2019;**27**:538–49.

Mukhopadhya I, Moraïs S, Laverde-Gomez J *et al.* Sporulation capability and amylosome conservation among diverse human colonic and rumen isolates of the keystone starch-degrader *Ruminococcus bromii*. *Environ Microbiol* 2018;**20**:324–36.

Oikonomou CM, Chang Y-W, Jensen GJ. A new view into prokaryotic cell biology from electron cryotomography. *Nat Rev Microbiol* 2016;**14**:205–20.

Oikonomou CM, Jensen GJ. The atlas of bacterial & archaeal cell structure: an interactive open-access microbiology textbook. *J Microbiol Biol Educ* 2021;**22**. doi:10.1128/jmbe.00128-21.

Oksanen J, Blanchet FG, Friendly M *et al.* vegan: community ecology package. R package version 2.5-7. 2020.

Ouyang J, Lin J, Isnard S *et al.* The bacterium *Akkermansia muciniphila*: a sentinel for gut permeability and its relevance to HIV-related inflammation. *Front Immunol* 2020;**11**. doi:10.3389/fimmu.2020.00645.

Paradis E, Schliep K. ape 5.0: an environment for modern phylogenetics and evolutionary analyses in R. *Bioinformatics* 2019;**35**:526–8.

Ramijan K, Ultee E, Willemse J *et al.* Stress-induced formation of cell wall-deficient cells in filamentous actinomycetes. *Nat Commun* 2018;**9**:5164.

R Core Team. *R: A Language and Environment for Statistical Computing*. Vienna: R Foundation for Statistical Computing, 2018.

Schindelin J, Arganda-Carreras I, Frise E *et al.* Fiji: an open-source platform for biological-image analysis. *Nat Methods* 2012;**9**:676–82.

Seemann T. Prokka: rapid prokaryotic genome annotation. *Bioinformatics* 2014;**30**:2068–9.

Seshadri R, Leahy SC, Attwood GT *et al.* Cultivation and sequencing of rumen microbiome members from the Hungate1000 Collection. *Nat Biotechnol* 2018;**36**:359–67.

Stevenson DM, Weimer PJ. Dominance of *Prevotella* and low abundance of classical ruminal bacterial species in the bovine rumen revealed by relative quantification real-time PCR. *Appl Microbiol Biotechnol* 2007;**75**:165–74.

Tatli M, Moraïs S, Tovar-Herrera OE *et al.* Nanoscale resolution of microbial fiber degradation in action. *Elife* 2022;**11**. doi:10.7554/eLife.76523.

Taxis TM, Wolff S, Gregg SJ *et al.* The players may change but the game remains: network analyses of ruminal microbiomes suggest taxonomic differences mask functional similarity. *Nucleic Acids Res* 2015;**43**:9600–12.

Tegunov D, Cramer P. Real-time cryo-electron microscopy data preprocessing with Warp. *Nat Methods* 2019;**16**:1146–52.

Weiss GL, Eisenstein F, Kieninger A-K *et al.* Structure of a thylakoid-anchored contractile injection system in multicellular cyanobacteria. *Nat Microbiol* 2022;**7**:386–96.

Witwinowski J, Sartori-Rupp A, Taib N *et al.* An ancient divide in outer membrane tethering systems in bacteria suggests a mechanism for the diderm-to-monoderm transition. *Nat Microbiol* 2022;**7**:411–22.

Yoav S, Barak Y, Shamshoum M *et al.* How does cellulosome composition influence deconstruction of lignocellulosic substrates in *Clostridium* (*Ruminiclostridium*) *thermocellum* DSM 1313?. *Biotechnol Biofuels* 2017;**10**:222.

Ze X, Duncan SH, Louis P *et al.* *Ruminococcus bromii* is a keystone species for the degradation of resistant starch in the human colon. *ISME J* 2012;**6**:1535–43.

Zheng S, Wolff G, Greenan G *et al.* AreTomo: an integrated software package for automated marker-free, motion-corrected cryo-electron tomographic alignment and reconstruction. *J Struct Biol X* 2022;**6**:100068.