Deep-learning deconvolution and segmentation of fluorescent membranes for high-precision bacterial cell-size profiling

Octavio Reyes-Matte[1,*], Carsten Fortmann-Grote[1], Beate Gericke[1], Nadja Hüttmann[1], Nikola Ojkic[2], Javier Lopez-Garrido[1,*]

[1] Max Planck Institute for Evolutionary Biology, Plön, Germany

[2] School of Biological and Behavioral Sciences, Queen Mary University of London, UK

Correspondence: lopezgarrido@evolbio.mpg.de; reyesmatte@evolbio.mpg.de

Keywords: image segmentation, image deconvolution, bacterial cell size, PBP1, *Priestia megaterium*, *Bacillus*

ABSTRACT

Evolutionary studies in bacteria have emphasized genetic and metabolic diversity, while cell-size variation has received less attention. Here we introduce MEDUSSA, a high-throughput method for precise bacterial cell-size profiling based on automatic segmentation of fluorescent membrane images, well suited to studying cell-size diversity. The approach couples deep-learning for membrane deconvolution and for cell segmentation with error-corrected cell measurement to extract accurate sizes from individual bacterial cells regardless of shape, size, chaining, or clustering. Our method overcomes limitations of phase-contrast segmentation, yielding reliable single-cell dimensions. We applied MEDUSSA to six strains of *Priestia megaterium* and found over twofold differences in cell volume across strains, largely driven by differences in cell width. We identified a partially-functional PBP1 allele that underlies the reduced width of one strain. Together, these results demonstrate the power of comparative analyses in bacterial cell biology and expand the toolkit to investigate the evolution of bacterial cell size.

INTRODUCTION

Bacterial cells are often perceived as uniformly small, which has led to a relative lack of attention to the evolution of bacterial cell size. Yet, careful microscopic observation has revealed an incredible diversity of bacterial shapes and sizes, with cell volumes spanning multiple orders of magnitude (Schavemaker & Munoz-Gomez, 2022; Schober et al., 2025). Even among closely related species, cell size can vary significantly. For instance, strains from two *Bacillus* species, *B. thuringiensis* and *B. pumilus*, can differ in volume by nearly an order of magnitude (Lopez-Garrido et al., 2018), suggesting that cell size is an evolvable trait capable of changing over relatively short evolutionary timescales.

Despite this diversity, our appreciation of bacterial cell size variation at finer evolutionary scales remains limited. Comparative studies analyzing the size of related bacterial strains are rare (Delaby et al., 2025; Gray et al., 2019), and most available data come from independent studies, often using different imaging methods and manual measurements of a few representative cells cultured in different conditions. While this highlights bacterial size diversity, methodological inconsistencies make cross-study comparisons difficult. As a result, we lack a clear picture of how bacterial cell size vary among closely related species and strains.

Recent advances in image analysis pose a unique opportunity for the standardization of methods to measure bacterial cell size and may help to address these challenges. In particular, the incorporation of machine-learning into image analysis pipelines have enabled the automatic identification and precise delineation of boundaries of individual bacterial cells in microscopy images. This process, known as image segmentation, generates masks of individual cells that can be used for accurate cell size measurements. The U-Net, a convolutional neural network architecture originally designed for the analysis of biomedical images (Falk et al., 2019; Ronneberger et al., 2015), is behind some of the most powerful image segmentation algorithms currently available, such as StarDist (Schmidt et al., 2018), DeLTA (Lugagne et al., 2020), Cellpose (Stringer et al., 2021), MiSiC (Panigrahi et al., 2021), Omnipose (Cutler et al., 2022), DeepBacs (Spahn et al., 2022)and DistNet (Ollion et al., 2024), to name a few. These tools have significantly improved the precision and automation of bacterial cell segmentation, paving the way for more consistent and reproducible cell size measurements.

Here, we have built upon these methods to develop a pipeline for extracting precise single-cell size information from fluorescent membrane images, which allows the characterization of individual cell sizes regardless of clustering or chaining patterns. We applied this pipeline, that we call ME-DUSSA, to investigate cell-size diversity across six strains of *Priestia megaterium*, a Gram-positive bacterium commonly used for physiological studies and biotechnological applications (Logan & Vos, 2015). We found significant differences in volume between the strains—spanning more than twofold—underpinned by consistent differences in cell length and width. We further link width differences between two of the strains to a specific mutation in *ponA*, which encodes penicillin binding protein 1 (PBP1), resulting in a hypomorphic PBP1 allele.

RESULTS

Limitations of phase contrast segmentation for single cell studies.

U-Net models are versatile and can segment different types of images, as long as they are trained with appropriately labeled datasets for the specific task. However, phase-contrast has become the default imaging modality for bacterial segmentation, as it provides excellent contrast between cells and background without the need for staining or genetic manipulation (Cutler et al., 2022; Ducret et al., 2016; O'Connor et al., 2022; Ollion et al., 2024; Paintdakhi et al., 2016; Stylianidou et al., 2016; Ursell et al., 2017; Van Valen et al., 2016). Despite its advantages, phase contrast images have limitations for cell size estimation. Firstly, they lack clear visual references to precisely define cell boundaries, which tends to result in an overestimation of cell size using currently available segmentation models (Hardo et al., 2024). Secondly, phase contrast does not allow the distinction of cells that remain attached after division, complicating the identification of individual cells in bacterial species that form cellular chains (Holt et al., 1994; Syvertsson et al., 2016). For instance, *Bacillus subtilis* and related species form heterogeneous populations in which single cells coexist with long chains consisting of dozens of cells (Kearns & Losick, 2005; Margulis et al., 1998). Because the separation of individual cells within a chain is not visible in phase-contrast images, state-of-the-art U-net models optimized for phase-contrast segmentation, such as the Omnipose "bact_phase_omni" model (Cutler et al., 2022), segment chains as very long individual cells, making it difficult to extract information at the single cell level (Fig. 1a,c).

This issue can be circumvented by staining the cell membranes with lipophylic fluorescent dyes such as FM 4-64 (Pogliano et al., 1999), which allows the visualization of the septal membranes that separate individual cells within chains (Fig. 1b,d). The fluorescence intensity profile of FM 4-64-stained membranes shows a sharp peak that can be traced around the cell (Fig 1e,f), providing an unambiguous guide to delimit the cell boundaries. However, existing fluorescence segmentation models fail to generate accurate masks from fluorescent-membrane images (Supplementary Fig. 1), hindering high-throughput image analysis.

To overcome these limitations, we trained a custom deep learning model to facilitate the segmentation of fluorescent-membrane images.

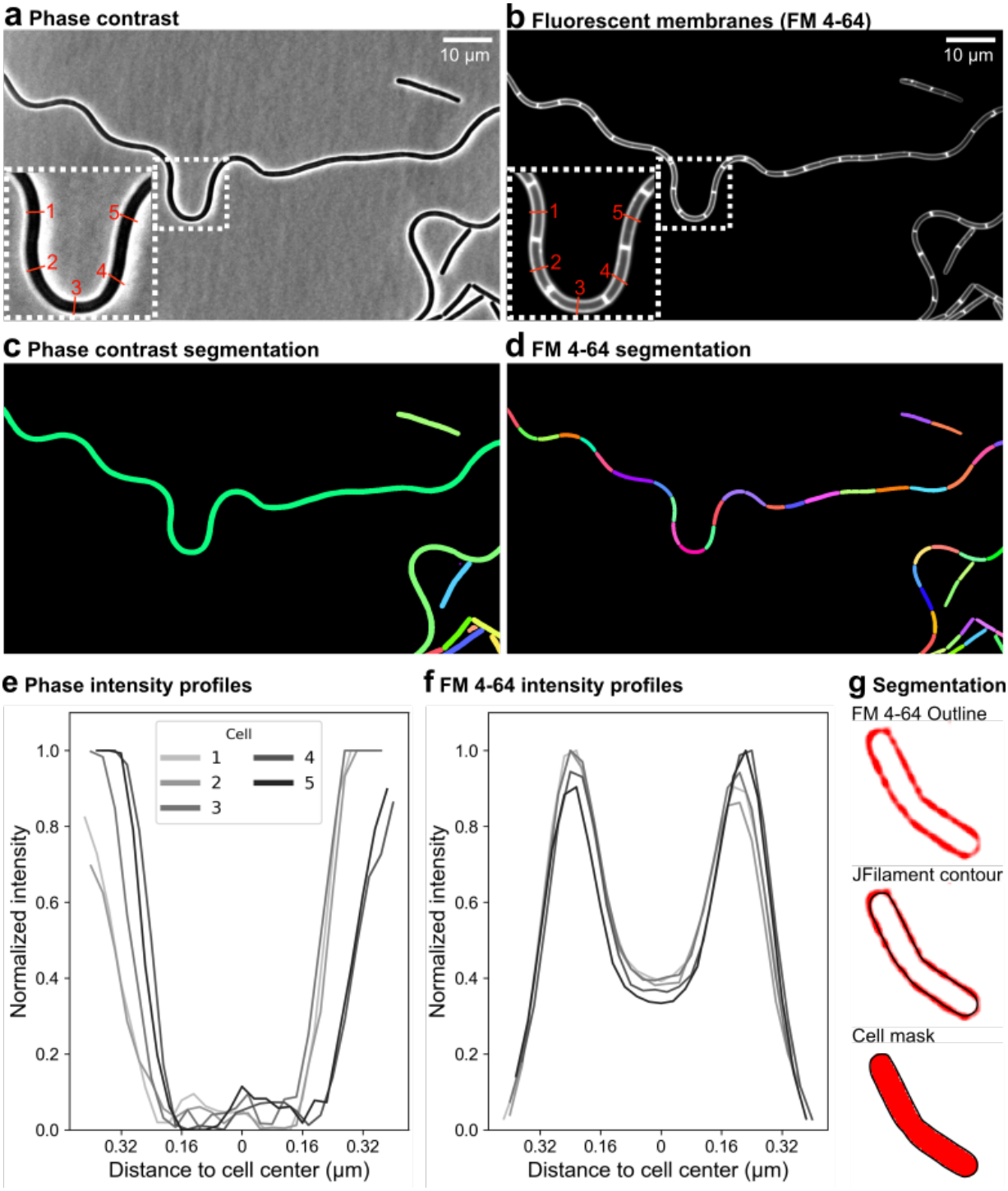

Figure 1. Phase-contrast versus fluorescent-membrane segmentation. a, **b**, Phase-contrast (**a**) and fluorescent membrane images (**b**) of *B. subtilis* growing vegetatively. Membranes were stained with FM 4-64. The insets show zoomed-in views of the boxed regions, with five cells (1 to 5) selected for intensity profiles across the red lines (shown in panels **e** and **f**). **c**, **d**, Segmentation masks for the phase-contrast (**c**) and fluorescent membrane (**d**) images shown in **a** and **b**. The phase-contrast image was automatically segmented with Omnipose (Cutler et al., 2022) and the fluorescent-membrane image was manually segmented with JFilament (Smith et al., 2010) (see also panel **g**). Every mask has a different color. **e**, **f**, Phase-contrast (**e**) and fluorescent membrane (**f**) intensity profiles of five selected cells measured along a 24-pixel line perpendicular to their long axis, centered at midcell (shown in insets in panels **a** and **b**). Pixel intensities were normalized to the maximum intensity observed across the five lines for each imaging modality. **g**, Example of manual fluorescent-membrane segmentation using the active contour software JFilament (Smith et al., 2010).

High-throughput segmentation of deconvolved fluorescent membranes.

U-net model training is performed in a supervised deep-learning setting using "ground truth" data consisting of manually or semi-manually annotated images, where a binary mask marking cell and non-cell pixels is generated for each image in the training dataset. The trained U-Net model will then be able to automatically identify cells in images that do not belong to the ground truth training dataset with a certain accuracy, which depends of different factors such as the quality, quantity and annotation accuracy of the training data set.

We employed the Omnipose algorithm (Cutler et al., 2022) due to its superior robustness with respect to different cell shapes—especially to extreme cellular elongation—compared to other segmentation algorithms. The new model, RawFMSeg, was trained using over seven-thousand FM 4-64-stained *B. subtilis* cells annotated with JFilament active-contour software (Fig. 1g) (Smith et al., 2010). The dataset included single cells, cells within chains and cells within clusters (see Methods for details). We tested the model with images of three bacilli of different sizes: *B. subtilis*, *B. thuringiensis*, and *Priestia megaterium* (Supplementary Fig. 2). The model produced segmentation masks lacking aberrations, truncations, and protrusions for isolated cells and individual cells within a chain, maintaining a proper rod-shaped morphology. However, it struggled with clustered cells, often producing indented or truncated masks (Supplementary Fig. S2).

We hypothesized that the difficulty in the segmentation of clustered-cells was related to the out-of-focus fluorescent signal, which increases noise, especially in regions of high cell density (Hardo et al., 2024). To address this issue, we retrained the model using deconvolved images. Deconvolution reassigns out of focus fluorescence back to its original focal plane, resulting in sharper membrane images with less noise (Agard, 1984; Scalettar et al., 1996). The retrained model, FMSeg, outperformed generalist models in fluorescent membrane segmentation (Fig. 2a). It successfully identified and segmented single cells within groups, isolated cells, chains, and even extremely elongated cells, regardless of the species or size (Fig. 2b,c). Additionally, FMSeg successfully segmented yeast and cocci of various sizes, both isolated and clustered, and worked with alternative membrane labels, including Mitotracker Green and GFP fused to a homogeneously distributed membrane protein (QoxB) (Fig. 2d), demonstrating its versatility.

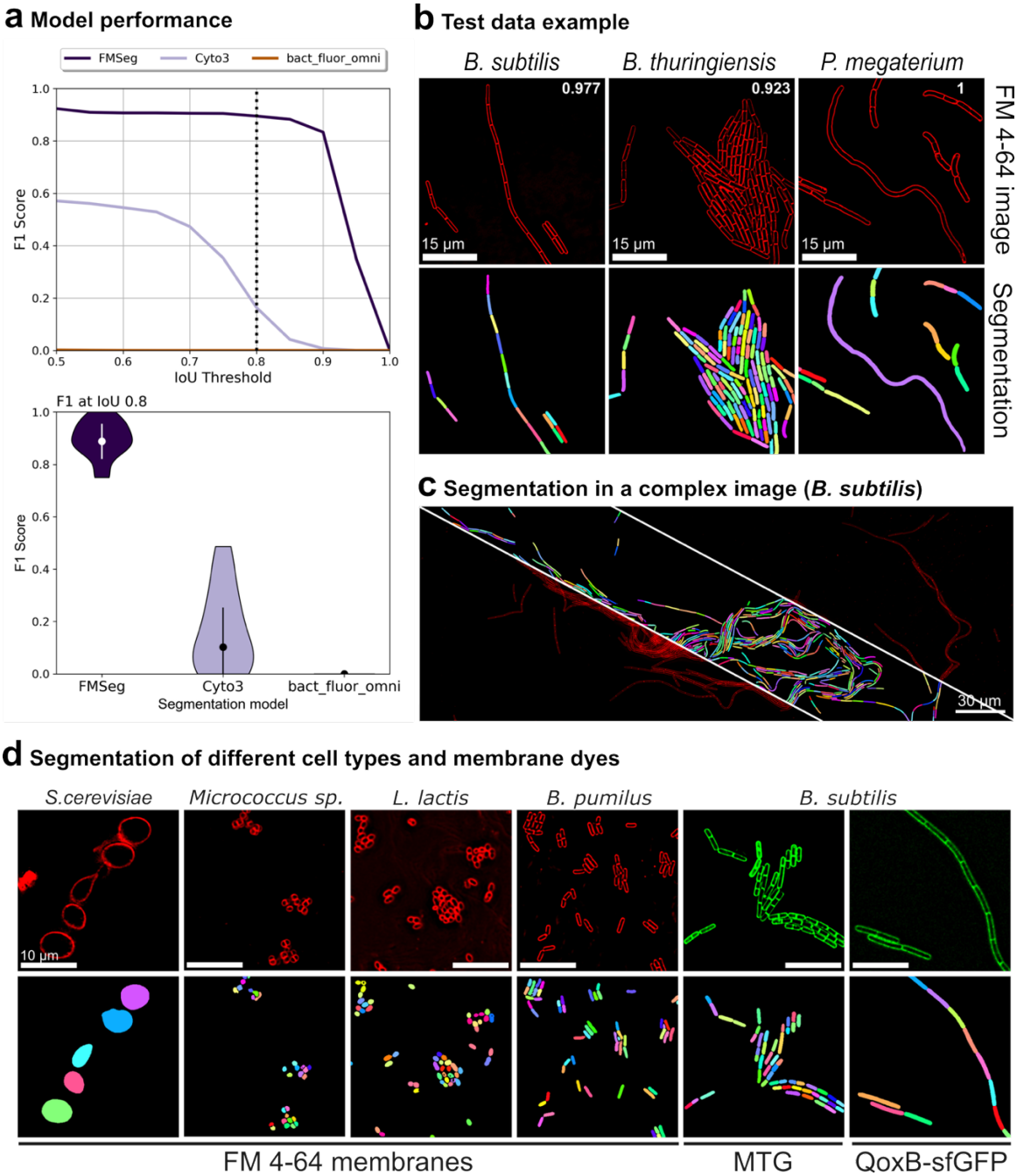

Figure 2. Qualitative and quantitative evaluation of a custom model for fluorescent-membrane segmentation. a, Quantitative benchmarking of FMSeg (dark purple) versus cyto3 (light purple) (Stringer & Pachitariu, 2025) and bact-fluore-omni (orange) (Cutler et al., 2022) on sixteen deconvolved fluorescent-membrane images of *B. subtilis*, *B. thuringiensis* and *P. megaterium*. Top, F1 score across Intersection over Union (IoU) thresholds. The dashed line marks the IoU value of 0.8, above which differences are visually indistinguishable (Cutler et al., 2022). Bottom, F1 at IoU of 0.8. The dots inside each violin plot represent the mean and the vertical lines the standard deviation. The mean FMSeg F1 score, 0.895. **b**, Examples of images showing chains, clusters, and heavily elongated cells. Top, deconvolved micrographs of *B. subtilis*, *B. thuringiensis* and *P. megaterium* cells stained with FM 4-64; bottom, FMSeg masks. F1 scores at IoU 0.8 are indicated at the top-right corner of each image. Scale bar, 15 µm. **c**, FMSeg segmentation of a *B. subtilis* microcolony. The image is split into three sectors. Segmentation results are shown in the middle sector, and FM 4-64 stained membranes in the flanking sectors. Scale bar, 30 µm. **d**, Segmentation of cells of different shapes and sizes, with membranes labeled with FM 4-64 or other fluorophores. Top, deconvolved fluorescent images; bottom, FMSeg segmentation. *Saccharomyces cerevisiae*, *Micrococcus sp.*, *Lactococcus lactis* and *B. pumilus* cells were stained with FM 4-64 (red). *B. subtilis* was stained with MitoTracker Green (MTG). We also imaged a *B. subtilis* strain producing a GFP fused to the membrane protein QoxB (QoxB-sfGFP, green), which is homogenously distributed in the membrane. Scale bars, 10 µm.

To facilitate the segmentation of fluorescent membrane images acquired with different microscopes, regardless of deconvolution capabilities, we trained an image restoration model to predict deconvolution from non-deconvolved images (FM2FM model; Fig. 3a). FM2FM was trained using the CARE framework (Weigert et al., 2018), in which a supervised U-Net predicts pixelwise fluorescence values while accounting for spatial context (i.e. neighboring pixel intensities). The fluorescence profiles of the predicted images closely matched those of deconvolved images (Fig. 3a,b), demonstrating the accuracy of the predicted deconvolution. Additionally, we trained a fluorescence translation model to infer deconvolved membranes from cytoplasmic fluorescence (FP2FM model; Fig. 3a). FP2FM is useful when membrane staining is not feasible (Fig. 3a,b).

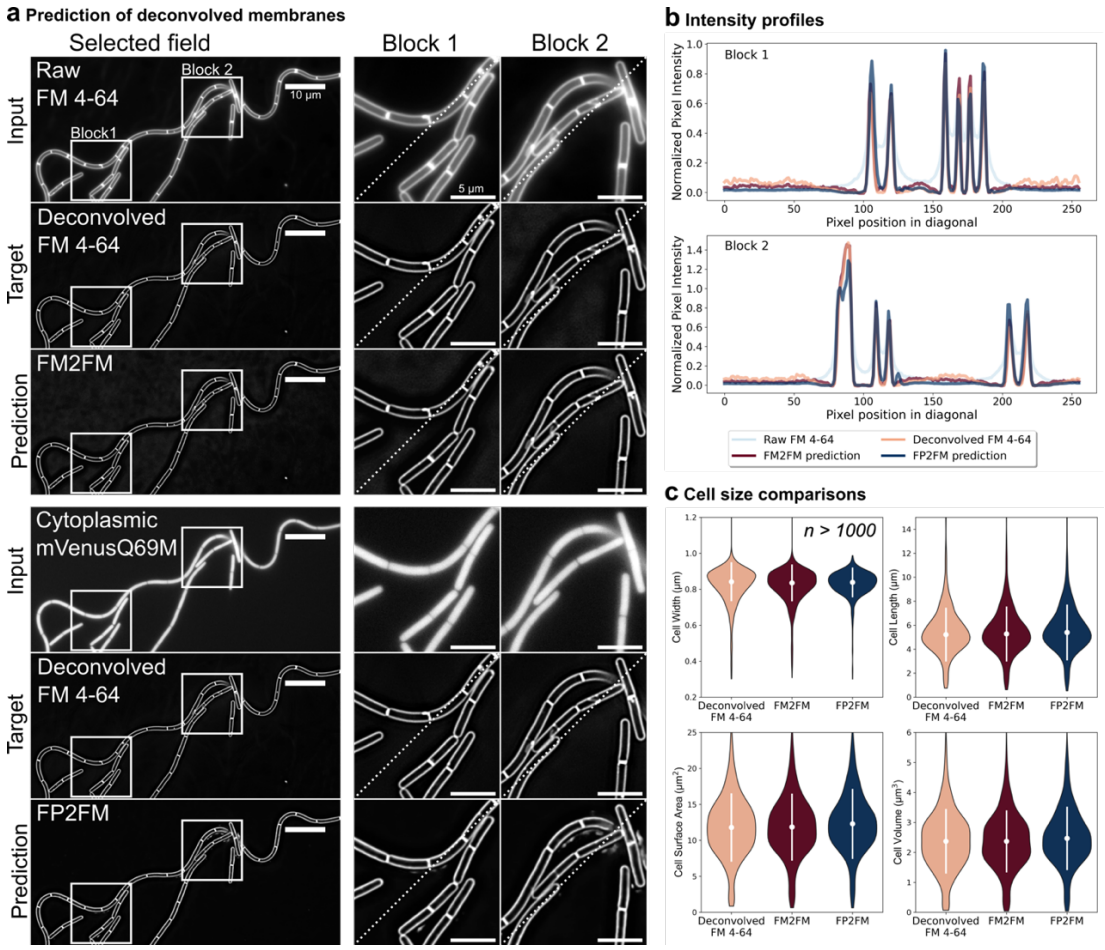

Figure 3. Deconvolution prediction. **a**, Two models, FM2FM (upper panel set) and FP2FM (lower panel set), were trained to predict deconvolved fluorescent images. FM2FM uses raw FM 4-64-stained membranes and FP2FM uses cytoplasmic fluorescence images (cytoplasmic mVenus-Q69M) as input (first row of each panel set), with true deconvolved fluorescent membrane images as target (second row of each panel set, deconvolved FM 4-64). The third row of each panel set shows the images predicted by each model. White squares mark regions zoomed in at right (Block 1 and Block 2). Diagonal dotted lines indicate profile traces in **b**. Scale bars: full images, 10 µm; zoomed-in blocks, 5 µm. **b**, Fluorescence profiles across the dotted lines in Block 1 (top) and Block 2 (bottom). Light blue, raw FM 4-64; orange, deconvolved FM 4-64; burgundy, FM2FM prediction; dark blue, FP2FM prediction. **c**, Violin plots of cell width, length, surface area and volume calculated from deconvolved FM 4-64 images (orange), FM2FM-predicted images (burgundy) and FP2FM-predicted images (dark blue). The white dots and vertical lines within the violin plots represent the medians and standard deviations. Twenty images (over 1000 cells) were processed and segmented with FMSeg. See the Methods for size calculation details.

To assess whether deconvolution prediction affects segmentation, we compared cell size parameters—width, length, surface area and volume (see next section and methods for details)—extracted from masks generated using deconvolved and predicted images. We used multiple size metrics to verify that information relevant to downstream analyses is preserved across processing scenarios, as single metrics often fail to predict how image-derived measurements propagate into downstream results (Gunawan et al., 2025). The size distributions were nearly identical (Fig. 3c), indicating that membrane deconvolution prediction has a minimal effect on segmentation and does not impact cell size calculations.

Altogether, our custom segmentation and deconvolution prediction models offer a robust approach for the high-throughput segmentation of individual bacterial cells with fluorescent membranes, regardless of their size, shape and chaining pattern.

MEDUSSA, an analytical pipeline for extracting cell size from fluorescent membrane images.

We designed an algorithm to extract precise 2D cell dimensions—width and length—and extrapolate 3D measurements such as membrane surface area and volume (see methods) (Fig. 4a). Integrating segmentation with size calculations revealed two key factors critical for reliable measurements (Fig. 4b-d; Supplementary Fig. 3; Supplementary Fig. 4).

The first is that the size of the segmentation masks is sensitive to the focal plane (Fig. 4b,c; Supplementary Fig. 3). For rod-shaped cells, width is maximal at the mid-cell focal plane but decreases above and below it (Fig. 4c). It is therefore critical to ensure that size measurements are taken from mid-cell focal planes. However, due to minor aberrations in microscopy slides, cells within a field of view often reside in slightly different focal planes (Fig. 4b; Supplementary Fig. 3). To address this, we acquired z-stacks and projected them onto a single plane before predicting deconvolution and segmentation (Fig. 4b). This approach yielded size estimations equivalent to those obtained from mid-cell focal planes (Fig. 4c), enabling accurate measurements across entire microscopy field, regardless of focal plane differences.

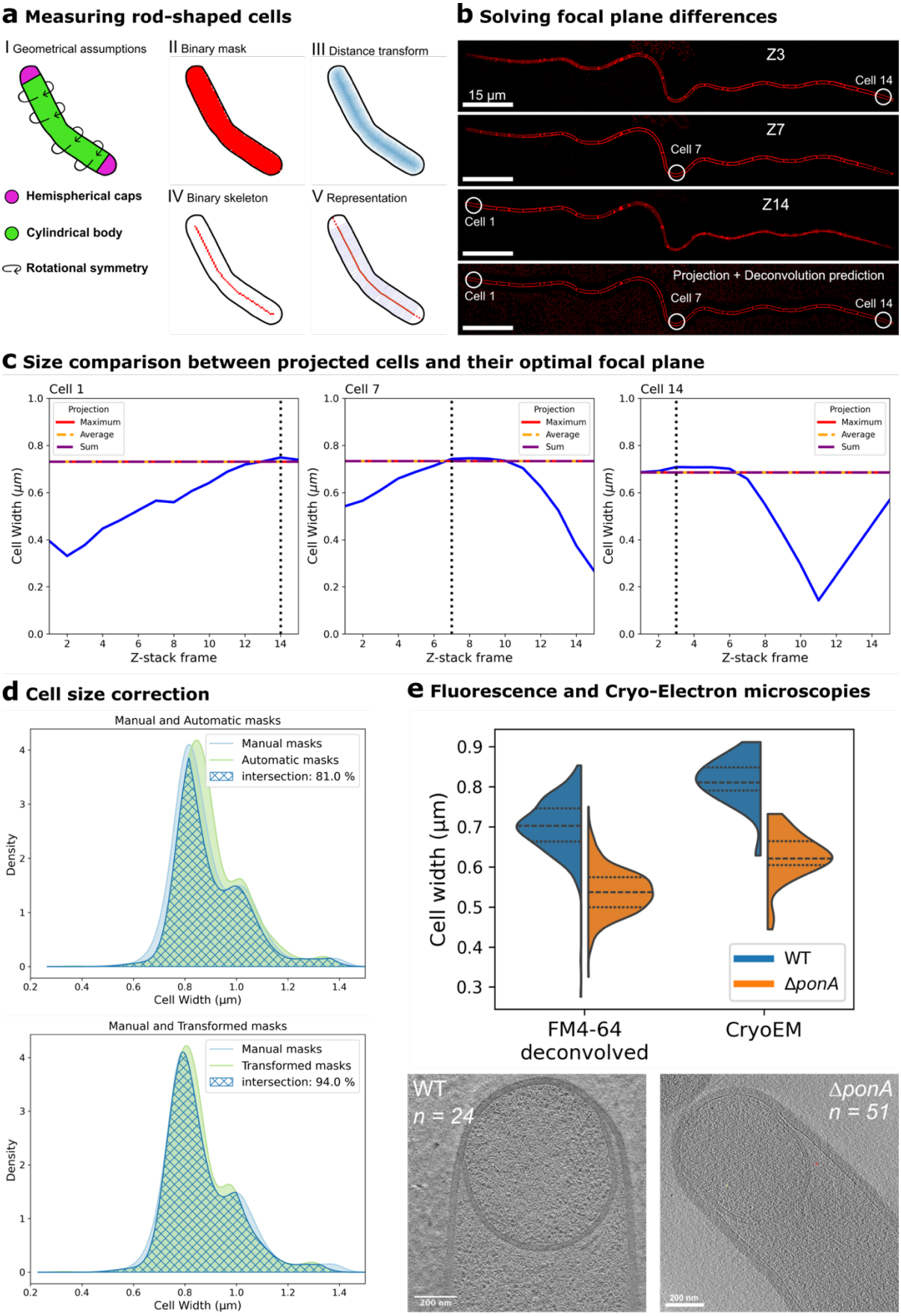

Figure 4. Cell size estimation from segmentation masks. a, Cell size calculation workflow. See the Methods for details. **b**, Representative z-slices from an optical-sectioning stack of a 14-cell *B. subtilis* chain stained with FM 4-64. Different cells are in focus in different slices. Circled cells are analyzed in **c**. The last image is a maximum-intensity

projection of whole stack, processed FM2FM, where all cells appear in focus. Scale bar, 15 μm. Full stack in Supplementary Fig. 3. **c**, Width profiles of the cells circled in **b** on different z-slices across the optical sectioning stack (blue lines): Cell 1 (right graph), Cell 7 (middle graph) and Cell 14 (right graph). Width is maximized on the slice in which the cell is in focus (vertical dotted lines). Cell width calculated from projected images is indicated by horizontal dashed lines. Results were equivalent regardless of the projection type (maximun intensity, red dashed line; average intensity, orange dashed line; sum, purple dashed line). Cells were segmented using JFilament in every z-slice and projections with FMSeg. **d**, FMSeg masks are slightly wider than manual JFilament masks. Density plots show the width distribution of 1100 bacilli cells. The top plot compares JFilament masks (blue) to FMSeg masks (green). The overlapping region between the two distributions is shown as a cross-hatched area, and the percentage of overlap (intersection) is indicated in the plot. After transforming the data with our sampling strategy (transformed masks, bottom plot), the distributions align almost perfectly. See Supplementary Fig. 4 and the Methods section for details. **e**, Widths of wild-type (WT, blue) and Δ*ponA* (orange) cells from sporulating *B. subtilis* cultures, measured from deconvolved fluorescent-membrane images segmented with FMSeg (FM 4-64 deconvolved; 326 and 617 cells for WT and Δ*ponA*, respectively), or from CryoEM images (24 and 51 cells for WT and Δ*ponA*, respectively). Dashed lines mark the median: WT 0.70 μm (FM) vs 0.81 μm (cryo-EM); Δ*ponA* 0.54 μm (FM) vs 0.63 μm (cryo-EM). FM widths were ~14% lower than cryo-EM, but the WT/ΔponA ratio was nearly identical (1.300 vs 1.304; <1% difference). Cryo-EM images courtesy of E. Tocheva (ΔponA) and E. Villa & K. Khanna (WT).

Second, we observed that the FMSeg model consistently generated masks slightly wider than ground truth masks (Fig. 4d; Supplementary Fig. 4a). This bias propagated into surface area and volume calculations (Supplementary Fig. 4). Fortunately, the overestimation was independent of cell size, allowing us to apply a systematic correction (Fig. 4d; Supplementary Fig. 4b). We developed a size correction pipeline that allowed us to predict the "true" cell size based on the dimensions of FMSeg-generated masks (Supplementary Fig. 4; see methods for details). We validated this approach by correcting the segmentation outputs for *B. subtilis*, *B. thuringiensis* and *P. megaterium* cells used in our FMSeg benchmarking (Fig. 2a). As shown in Fig. 4d, FMSeg-mask width distribution was slightly shifted compared to that obtained from ground-truth masks manually generated using JFilament. After applying the correction, both distributions became nearly identical (Fig. 4d).

To assess the accuracy of our cell size measurements, we compared estimates from our pipeline with those obtained from plunged-frozen *B. subtilis* imaged by cryo-electron microscopy (cryoEM), a technique that provides high-resolution images of cells preserved in their native physiological state. We focused on sporulating *B. subtilis* cells, as we could access sets of cryoEM images of these cells, kindly provided by Elitza Tocheva (Tocheva et al., 2013), and Elizabeth Villa and Kanika Khanna (Khanna et al., 2021; Khanna et al., 2019; Lopez-Garrido et al., 2018). To test robustness across different cell sizes, we measured both wild type cells and cells lacking PBP1 (*ponA*), which are ~25% thinner than wild type cells (Popham & Setlow, 1996). Our pipeline produced results similar to those from cryoEM images, for both the thicker wild-type and thinner *ponA* cells (Fig. 4e), indicating that our pipeline provides realistic cell size estimations and is robust to cell size differences.

Overall, our segmentation model, coupled with deconvolution prediction and size correction, enables high-throughput, accurate cell size measurements from fluorescent-membrane microscopy images. We refer to the full pipeline for bacterial cell-size measurement as MEDUSSA (MEmbrane DeconvolUtion and Segmentation for Size Analyses).

Insight into intraspecific cell-size diversity in *P. megaterium*.

We applied MEDUSSA to study variation in cell size across strains of *P. megaterium (de Bary, 1884)*, a member of the *Bacillaceae* generally considered to produce large cells but for which intraspecific morphological diversity has been reported (Logan & Vos, 2015; López & Alippi, 2009; Murumkar et al., 2012). We focused on six strains widely used in physiology and biotechnology: the type strain DSM 32 (ATCC 14581) (Arya et al., 2014) from the German Collection of Microorganisms and Cell Cultures (DSMZ); the classic multi-plasmid strain QM B1551 (Eppinger et al., 2011); the naturally plasmidless strain DSM 319 (Eppinger et al., 2011; Stahl & Esser, 1983); KM, used in early bacteriophage studies; 899, also used in early bacteriophage and bacteriocin studies; and WH320 (Rygus et al., 1991), a Lac^- derivative of DSM 319 commercially available for protein-expression. We sampled cultures of each strain during exponential and stationary phases in LB medium (Supplementary Fig. 5), imaged them in agarose pads supplemented with FM 4-64 for membrane staining, and processed the images through our complete pipeline to obtain estimations of the sizes of individual cells (Fig. 5; Supplementary Fig. 6). Although all strains tended to form chains (Fig. 5a), MEDUSSA could resolve single cells with high precision, providing robust single-cell dimensions that were similar between experimental runs (Supplementary Fig. 6).

We observed striking between-strain diversity in cell width: during exponential growth, median width ranged from 0.92 μm in WH320 to 1.39 μm in 899 (Fig. 5b; Supplementary Figure 6). Most strains were slightly thinner in stationary phase, except KM, whose width distribution broadened without a clear shift in the median (Fig. 5b). Although width showed a smaller coefficient of variation (CV) than length—consistent with tighter control—the width CVs of *P. megaterium* strains were generally higher than those reported for *B. subtilis* (Juillot et al., 2021; Nordholt et al., 2020; Sauls et al., 2019; Taheri-Araghi et al., 2015)and reproduced by us in Supplementary Fig. 6), with some strains reaching ~0.15 in exponential or stationary phase (Supplementary Fig. 6). Notably, cells of different widths were sometimes present within a single chain (arrowheads in Fig. 5a).

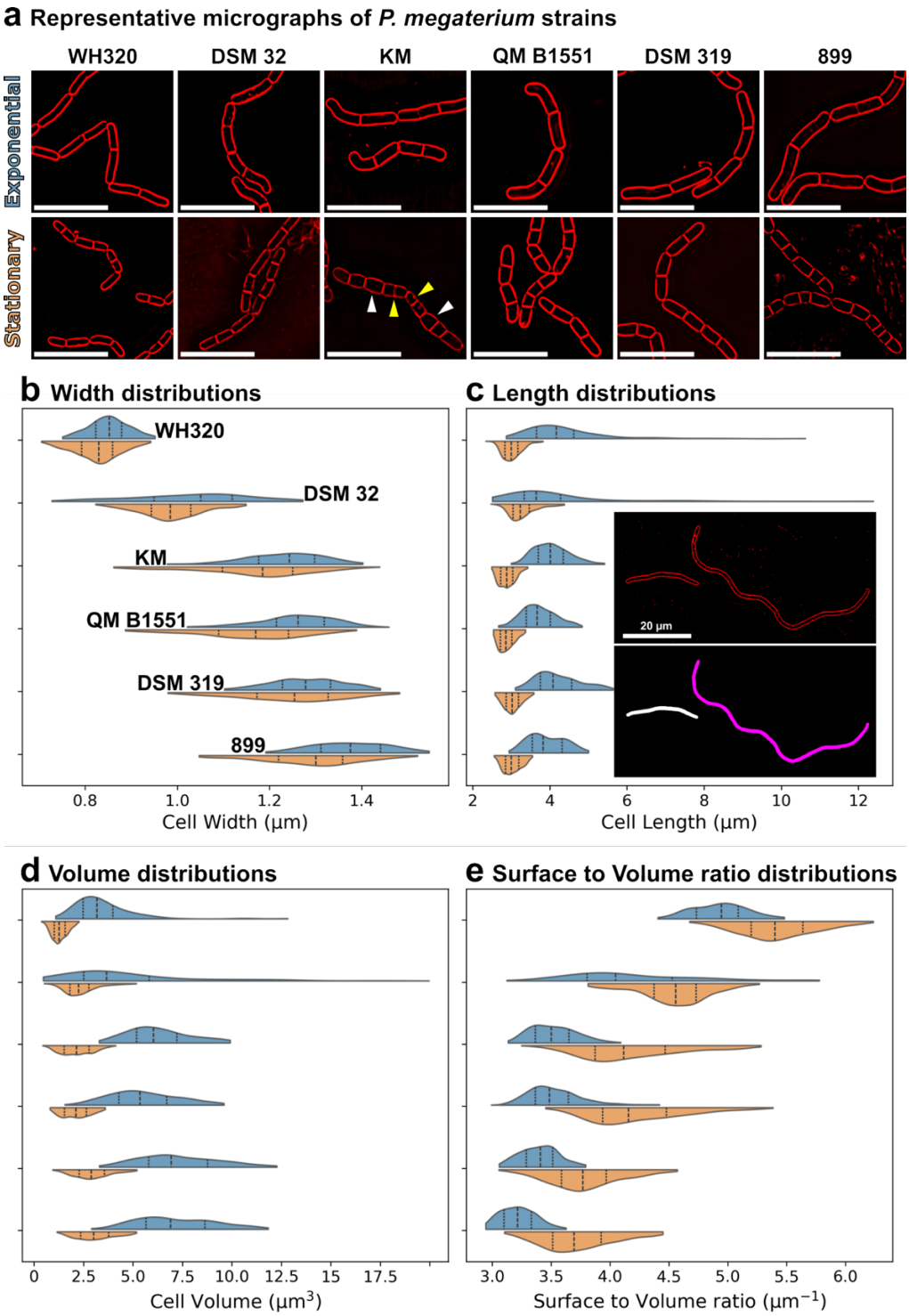

Figure 5. Morphological characterization of six *P. megaterium* strains. a, Fluorescence micrographs of FM 4-64-stained cells of different *P. megaterium* strains sampled form exponential and stationary phase LB cultures: WH320, DSM 32, KM, QM B1551, DSM 319 and 899. Arrowheads in the KM stationary image point to thicker (white) and thinner (yellow) cells within a chain. Scale bars, 10 μm. **b-e,** Distributions of cell with (**b**), length (**c**), volume (**d**) and surface to volume (**e**) for the indicated strains (strain names in panel **b** and same order maintained in the other panels) during exponential (blue) and stationary (orange) phase. The dashed line in the violin plots indicate the median, and the dotted lines the first and third quartiles. The inset in panel **c** shows a micrograph of a long cell produced by DSM 32 during exponential growth (membranes, red; scale bar, 20 μm). The following number of cells were analyzed for each strain in exponential and stationary phase: WH320, 153 and 961 cells; DSM 32, 192 and 904 cells; KM, 149 and 856 cells; QM B1551, 292 and 447 cells; DSM 319, 195 and 471 cells; 899, 758 and 548 cells. The central 95% of observed values are represented. See Supplementary Fig. 6 for comparison between experimental runs. Prior to measurement, segmentation output was manually curated to remove aberrant masks and truncated masks at image edges.

We also observed significant differences in cell length across strains: in exponential phase, the median ranged from 4.06 μm in DSM 32 to 5.76 μm in strain 899, and cells shortened in stationary phase, as expected (Fig. 5a,c). Within-strain variation in cell length was high in some cases (Figure 5c; Supplementary Fig. 6), with rare, extremely elongated cells occurring in several strains; the more dramatic cases were in DSM 32, with some cells over 10 μm long (Fig. 5c, inset). These extreme elongation events were not observed in stationary phase (Figure 5c), suggesting that it is a phenomenon associated with active growth and that these filaments septate into multiple cells or lyse upon entry into stationary phase.

These width and length differences underpinned differences in surface-to-volume ratios—thinner cells exhibited higher surface-to-volume ratios than thicker cells (Ojkic et al., 2019, 2022)—and contributed to the diversity of cell volumes observed between strains, spanning nearly threefold (Fig. 5d-e).

***P. megaterium* WH320 carries a partially functional version of PBP1.**

Most accounts of *P. megaterium* cell width report values of ~1.2-1.5 μm (Logan & Vos, 2015). However, two of our strains, DSM 32 and WH320, had widths below these values. We were particularly surprised by WH320, whose median width during exponential growth was ~0.9 μm, as this strain is a derivative of DSM 319, one of the thickest strains in our panel (median width in exponential phase ~1.4 μm).

WH320 was generated by ethyl-metanosulfonate (EMS) mutagenesis of DSM 319, selecting for Lac⁻ clones (Rygus et al., 1991). One potential explanation for the reduced cell width of WH320 is that an additional mutation introduced during EMS mutagenesis or at some point after the two strains diverged affected cell width. To explore this possibility, we sequenced the genome of WH320 and compared it to the parental DSM 319 reference genome (Eppinger et al., 2011). The two strains differed by 37 single nucleotide polymorphisms, with the majority of them being G·C→A·T transitions, consistent with EMS mutagenesis (Sega, 1984) (Supplementary Data in Zenodo repository). We also observed other mutations including transversions and indels, which might have occurred during transfer of the strains in the lab before we acquired them (Supplementary Data in Zenodo repository). One of the transitions caused a glycine-to-arginine amino acid substitution in the catalytic site of *P. megaterium* β-galactosidase, likely responsible for the Lac⁻ phenotype. We also found a C to A transversion in *ponA* that produced an alanine-to-aspartic acid

change at residue 214 in the transglycosylase domain of the classA penicillin binding protein PBP1 (Supplementary Data in Zenodo repository).

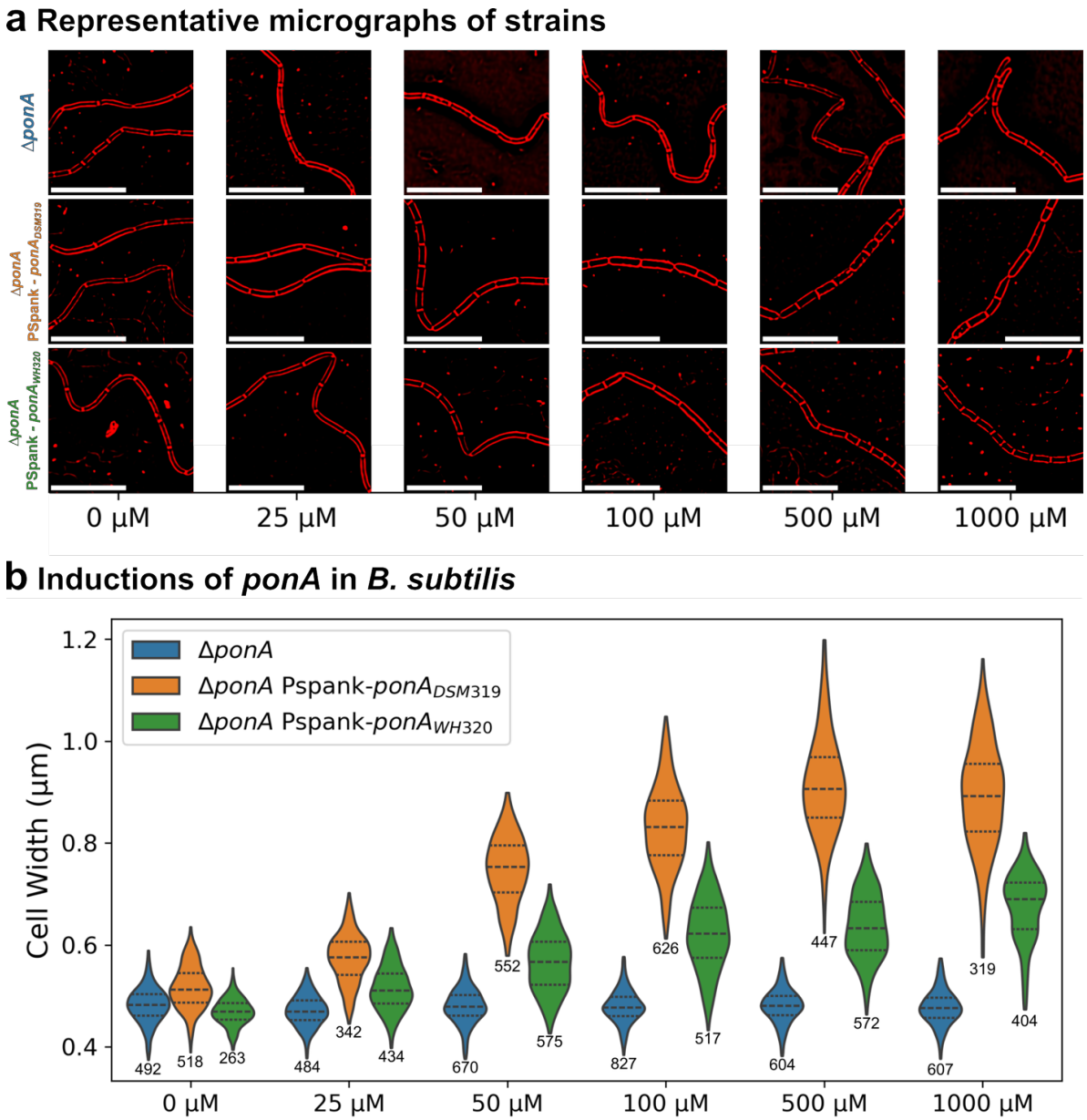

Figure 6. WH320 carries a hypomorphic PBP1 allele. **a,** Representative FM 4-64-stained cells corresponding to panel **b**. Scale bar, 10 µm. **b,** Cell width of a *B. subtilis* PY79 *ponA* mutant (Δ*ponA*, blue) and the same mutant expressing *P. megaterium* DSM 319 *ponA* (Δ*ponA* P_{spank}-*ponA*$_{DSM319}$, orange) or *P. megaterium* WH320 *ponA* (Δ*ponA* P_{spank}-*ponA*$_{WH320}$, green) from the IPTG-inducible P_{spank} promoter inserted at the *amyE* locus. Violin plots show width distributions at different IPTG concentrations. Central 99% of data points was sampled for each strain in each condition. The dashed line indicates the median; dotted lines mark the first and third quartiles. The number of cells for each violin is shown in the plot under each violin. Prior to measurement, segmentation output was manually curated to remove aberrant masks and truncated masks at image edges.

As noted above, PBP1 modulates cell width in *B. subtilis*, with *ponA* mutants being ~25% thinner that the wild type and *ponA* overexpression producing cells thicker than the wild type (Fig. 4e) (Dion et al., 2019; Popham & Setlow, 1996). We therefore hypothesized that the *ponA* mutation

reduces PBP1 activity and underlies the thin phenotype of WH320. To get insight into this possibility, we asked whether the *ponA* alleles from DSM 319 and WH320 could complement a *B. subtilis ponA* mutant when produced ectopically. We placed both alleles under the IPTG-inducible promoter P_{spank} at the ectopic *amyE* locus, and measured cell width after incubation with varying IPTG concentrations. As shown in Fig. 6, expression of DSM319 *ponA* produced a marked increase in cell width, even at low IPTG concentrations. In contrast, expression of WH320 *ponA* caused a milder increase in cell width and never reached the values observed with $ponA_{DSM319}$, even at high induction levels. These results indicate that the alanine-to-aspartate substitution in *ponA* reduces PBP1 function.

DISCUSSION

Here, we introduce MEDUSSA, an image-analysis pipeline that extracts accurate single-cell size measurements from fluorescent-membrane microscopy images, and demonstrate its utility for dissecting mechanisms of cell-size diversity among bacterial strains. The workflow combines deep-learning–based image restoration for deconvolution prediction with automated segmentation to delineate cell boundaries from membrane signal, followed by mask analysis to estimate cell size. Because each step is modular, the restoration (FM2FM), segmentation (FMseg), and measurement components can be used independently to suit specific study needs. Applied across strains and species, MEDUSSA yields cell-size measurements that can be further interrogated to identify the underlying mechanisms.

MEDUSSA is well suited for comparative studies because it addresses key shortcomings of previous methods. First, fluorescent membrane imaging enables precise delineation of cell boundaries by tracking the peak membrane signal around each cell (Fig. 1). This supports consistent training-set annotation and yields more reliable cell-size estimates than phase-contrast–based approaches (Hardo et al., 2024) (Fig. 1). Second, membrane dyes reveal septal membranes, allowing clear separation of daughter cells that remain attached after division and thus enabling reliable length measurements for both isolated cells and chains—something that is difficult or impossible with phase contrast (Fig. 1) (Syvertsson et al., 2016). Third, the pipeline corrects focal-plane differences among cells within the same image, which would otherwise bias measurements because focal-plane shifts alter mask size (Fig. 4b-c). Specifically, projecting fluorescence images from a z-stack and applying a deconvolution-prediction restoration algorithm (FM2FM) produces segmentation-ready images that yield mid-focal-plane masks. Thus, this standardized protocol of restoration, segmentation, and measurement converts membrane-fluorescence images into population-level size distributions.

A trade-off is the need for fluorescent labeling. Several membrane and cell-wall dyes are routinely used in bacterial cell biology (Greenspan & Fowler, 1985; Herman et al., 1994; Kuru et al., 2012; Lucidi et al., 2024; Monteiro et al., 2015; Pogliano et al., 1999; Sharp & Pogliano, 1999). Most experiments presented here use FM 4-64, a lipophilic styryl fluorophore that associates with the outer leaflet of the cytoplasmic membrane without disrupting membrane function (Pogliano et al., 1999). FM 4-64 is broadly applicable across bacterial species and compatible with live-cell imaging (Calvez et al., 2019; Dalebroux et al., 2015; Ma et al., 2009; Maloney et al., 2009; O'Hara et al., 2017; Ojkic et al., 2016; Shin et al., 2015; Specht et al., 2013; Studer et al., 2016; Sun & Margolin,

1998; Tsuge et al., 2008; Yeh et al., 2010). However, our segmentation model also works with other labels—including the membrane-permeable MitoTracker Green and with GFP fused to a membrane protein (Fig. 2d)—and, in principle, with any dye that provides uniform signal along the cell periphery. In addition, we provide a transformation model, FP2FM, that predicts membrane-like boundaries suitable for segmentation from cytoplasmic fluorescence, adding flexibility (Fig 3.). Altogether, the available dyes, combined with the transformation tools provided here, extend the applicability of the MEDUSSA method to virtually any bacterial species.

In this study, we applied MEDUSSA to quantify the size of six strains of *P. megaterium*. Our measurements reveal remarkable diversity in cell size both among strains and among individual cells within a strain (Fig. 5). Although all strains produce comparatively large cells by bacterial standards, some produce cells that are, on average, over twofold larger in volume than others. We also observed that some strains formed extremely elongated cells during exponential growth, which were not observed in stationary phase. To our knowledge, filamentation has not been reported previously in *P. megaterium*—likely because membrane dyes are not routinely used in microscopy. The mechanism underlying this filamentation, and whether such filaments occur in nature or are artifacts of laboratory culture, remains to be determined.

Cell width in bacteria is generally assumed to be tightly controlled within a species, with little cell-to-cell variation in steady state cultures (Campos et al., 2014; Chien et al., 2012; Ojkic & Banerjee, 2021; Ojkic et al., 2022; Sauls et al., 2019; Taheri-Araghi et al., 2015). Our dataset, however, reveals substantial width variation both within and between strains. Within strains, width values were approximately normally distributed, and some strains showed wide cell-to-cell variation; for example, in strain KM during stationary phase, the thinnest cells were up to twofold thinner than the thickest cells (Fig. 5). The coefficient of variation for width reached up to 0.15 (Supplementary Fig. 6)—more than twice that of *B. subtilis*—suggesting that mechanisms of width homeostasis in this bacterium are less precise than in *B. subtilis*.

Between strains, width distributions were shifted, with median widths ranging from ~0.9 to 1.4 μm (Fig. 5; Supplementary Fig. 6), indicating variation in width regulation across strains. We further examined differences between the thick strain DSM 319 and its Lac^- derivative WH320, which is ~0.5 μm thinner. Genome comparison revealed an alanine-to-aspartic acid substitution in the transglycosylase domain of PBP1. An alignment of PBP1 sequences from 854 strains indicates that alanine at position 214 is largely conserved (~75%); ~24% carry serine at this position, and fewer than 1% contain other residues (e.g., cysteine, threonine, valine)

(Supplementary Data in Zenodo repository). Notably, no sequence harbors an aspartic acid or any other charged residue at this position, suggesting that introducing charge there impairs PBP1 activity. Because PBP1 mutants yield thinner cells in *B. subtilis* (Popham & Setlow, 1996), it is plausible that the hypomorphic PBP1 allele in *P. megaterium* WH320 contributes—at least in part—to its reduced width relative to the DSM 319 parental strain.

Unlike the mutation in *lacZ*, the *ponA* mutation in *P. megaterium* WH320 is a G→T transversion. Because EMS typically induces G·C→A·T transitions (Sega, 1984), it is unlikely that the *ponA* mutation arose directly from EMS mutagenesis. Instead, we suspect that the *ponA* mutation in strain WH320 was either already present in the DSM 319 clone used for EMS treatment or was acquired during subsequent propagation of WH320 before we received the strain.

Overall, these results illustrate how MEDUSSA can be used to explore within- and between-strain diversity in bacterial cell size and to reveal differences that can be pursued to identify underlying mechanisms. Future work will expand our strain set and survey additional species and environmental isolates to advance our understanding of the evolution of bacterial cell size and contribute to the emerging field of evolutionary cell biology (Jiang et al., 2015; Lynch et al., 2014).

METHODS

Strains and plasmids

A list of strains is provided in Supplementary Table 1. Plasmid and primer lists are in Supplementary Tables 2 and 3, along with detailed descriptions of plasmid construction.

Culture media and growth conditions

Bacteria were routinely grown on LB agar plates (Lennox formulation, 5 g NaCl/L) overnight at 30 °C, with the exception of *Lactococcus lactis*, which was grown on MRS medium at 28 °C. *Saccharomyces cerevisiae* was grown on YPD at 37°C.

For size quantification of *B. subtilis* and *P. megaterium* strains, a single colony was inoculated into 5 mL fresh LB Lennox and grown overnight in a roller drum at 30 °C with constant rotation. The next morning, cultures were diluted 1:1000 into 30 mL of the appropriate medium in 250-mL baffled flasks and incubated in a water bath at 30 °C with shaking at 220 rpm. Samples were taken after five hours (exponential; Supplementary Fig. 5) and after ten hours (stationary; Supplementary Fig. 5) for microscopy (see next section).

For the induction experiments shown in Fig. 6, a single colony was inoculated into LB Lennox diluted 1/4, then grown at 25 °C in a shaking incubator at 220 rpm. The next morning, cultures were diluted to an OD600 of 0.02 in 1/4-diluted LB Lennox with the indicated IPTG concentration (0, 25, 50, 100, 500, or 1000 μM). Cultures were grown at 30 °C in a shaking incubator to an OD600 of 0.2–0.5 and then samples were taken for microscopy.

Sporulation (Fig. 4e) was induced by resuspension, as described elsewhere (Riley et al., 2018).

Microscopy

For imaging, 10 μL culture samples were imaged on 1.2% agarose pads in 1/4-diluted LB-Lennox supplemented with 0.5 μg/mL FM4-64. The pad was covered with a #1.5 glass coverslip. Ten-hour samples were diluted tree times to avoid overcrowded fields.

For cell-size estimation experiments for the different *B. subtilis* and *P. megaterium* strains, cells were imaged using a Leica Thunder Imager live cell with a Leica DFC9000 GTC sCMOS camera and a 100X HC PL APO objective (Numerical aperture 1.4, Refractive index 1.518 objective). FM 4-64 images were taken using a 510 nm excitation LED light source, with 70% intensity and 0.1 s exposure, and a custom filter cube produced by Chroma Technology GmbH (Olching, Germany)

with the following components: excitation filter, ET505/20x; dichroic mirror, T660lpxr; emission filter, ET700lp. Training images for the restoration (FM2FM) and translation (FP2FM) models were taken with a DeltaVision Ultra microscope with a PCO Edge sCMOS camera and an Olympus 100X UPLX APO objective (Numerical aperture 1.45, Refractive index 1.518 Ph3). FM 4-64 and mVenusQ69M were visualized with the following filters: excitation 542 nm and emission 679 nm for FM 4-64, and excitation 475 nm and emission 525 nm for mVenusQ69M, with light transmission at 100% and an exposure time of 0.2 s for both channels. Training images for the segmentation models (FMSeg, RawFMSeg) were taken using the same illumination conditions as the restoration model. For the membrane images of *E. coli*, the FM4-64 was captured with light transmission of 40% and an exposure time of 0.1 s.

For phase-contrast microscopy, the aforementioned DeltaVision microscope was used, light transmission was set to 100% and exposure time to 0.1 s.

Ground truth image annotation using active contours

We annotated cells in our ground-truth images using the ImageJ plugin JFilament (Smith et al., 2010), which allows the semi-automatic generation of sub-pixel resolution active contours (snakes) that follow the ridges of maximum fluorescence membrane intensity around the cell (Fig. 1f) (Lopez-Garrido et al., 2018; Ojkic et al., 2014, 2016). Default parameters were used except for Gamma, which we increased when the preset value caused excessive deformation of the snakes. For cell clusters, snakes were typically manually corrected because contour edges tended to deform toward neighboring cells. When membrane blobs distorted cell contours, we also corrected the snakes to follow an appropriate straight line. These steps ensured consistent single-cell segmentation for model training and improved robustness of the segmentation model to such confounding artifacts.

Training and evaluation of models for segmenting for fluorescent membranes (RawFM-seg and FMseg)

We annotated 143 raw and SoftWorx (Cytiva Life Sciences)-deconvolved FM 4-64 membrane images using JFilament (Smith et al., 2010), as described above, yielding over 7000 cells for training. Training images were maximum-intensity projections of FM4-64 z-stacks. Following the Omnipose documentation (Cutler et al., 2022), we trained the models from scratch, i.e., not using any existing model parameters as starting points. Models were trained with a batch size of 8 images for 4000 epochs using a learning rate of 0.1 and a Rectified Adam (RAdam) optimizer (Liu et al., 2019). Both models (raw and deconvolved) were trained using the same parameters. Training was

done in the "deepbio" computing cluster at the Max Planck Institute for Evolutionary Biology, using a single NVIDIA(R) GeForce GTX 2080 Ti GPU with 12GB memory.

For accuracy assessment, we annotated 16 images with JFilament. F1 score (Caicedo et al., 2019; Maier-Hein et al., 2024; Taha & Hanbury, 2015) was calculated at eleven Intersection over Union (IoU) thresholds spanning 0.5 to 1, in 0.05 intervals. True positives, false positives, and false negatives were calculated using Omnipose library function "average_precision". F1 was then computed using these values following Caicedo et al. 2019 (Caicedo et al., 2019).

Training and evaluation of models for deconvolution prediction (FM2FM) and fluorescence translation (FP2FM)

Deconvolution-prediction and fluorescence-translation models were trained using the Content Aware Restoration (CARE) framework (Weigert et al., 2018). For each model, 200 images (2048x2048 pixels) of vegetatively growing *B. subtilis* constitutively expressing mVenusQ69M (Cox et al., 2010) were acquired using a DeltaVision Ultra microscope with a PCO Edge sCMOS camera, capturing FM4-64 membrane fluorescence, and mVenusQ69M cytoplasmic fluorescence. Images were deconvolved in the DeltaVision software (SoftWorx) and then cropped to remove deconvolution edge artifacts, yielding 1914x1914 px image outputs. Non-deconvolved training inputs were cropped to the same size. Training input images were single-plane images of raw FM4-64 (for the FM2FM model) or mVenusQ69M (for the FP2FM model) z-stacks, and target images were the deconvolved FM4-64 images for both models. For training, 16 patches of 512x512 px were extracted from each image pair. Models were then trained using a batch size of 32 images for 100 epochs, with 100 training steps per epoch. Training was performed in the "deepbio" computing cluster at the Max Planck Institute for Evolutionary Biology, as above. Model parameters and training hardware was the same for both models.

To compare model performance against the membrane deconvolution target (SoftWorx), we conducted qualitative assessments of output images, compared intensity profiles, and evaluated cell-size estimates using membrane images deconvolved with SoftWorx or predicted with FM2FM or FP2FM (Fig. 3).

Extracting cell size information from masks of rod-shaped cells.

For rod-shaped bacteria, we assumed a sphero-cylindrical geometry with hemispherical caps (Ojkic et al., 2019). First, we computed the minimum Euclidean distance from each pixel inside the mask

to the closest pixel on the mask boundary (Borgefors, 1986), which is commonly referred to as the distance transform. Next, we generated a one-dimensional centerline (commonly known as "skeleton") (Zhang & Suen, 1984), overlaid it on the distance transform map, and determined the distance value at each skeleton pixel, thereby obtaining the local radius of the mask along the skeleton line (Fig. 4a). An alternative one-dimensional representation is the medial axis (Blum, 1967), also known as the "topological skeleton". We compared size estimates for an example cell using both representations and found minimal differences (Supplementary Fig. 7). We therefore proceeded with the skeleton representation.

To estimate the extend of the hemispherical caps, we took the distance between each skeleton endpoint and the corresponding mask tip as the cell radius at the skeleton endpoints. This yielded precise estimates of mask length and with profiles at every pixel along the skeleton. Assuming rotational symmetry and that the masks derive from medial focal-plane images, we estimated cell volume and surface area by adding the volumes and lateral surface areas of cylindrical segments generated by circularly projecting the radius of the mask at each pixel of the skeleton, plus the hemispherical caps (Fig. 4a). This approach accurately handles elongated masks, provides local morphological information of cell morphology, and is robust to cellular curvature. While tailored to rods, the framework can be adapted to other morphologies.

Dealing with branching skeletons

Due to the mask geometry, it is possible that the skeleton branches into multiple endpoints (Supplementary Fig. 7c). To handle both branched and non-branched cases, we developed the following workflow. First, we treat the skeleton as a graph, with each pixel considered a node. Assuming connectivity only between neighboring pixels (horizontal, vertical, or diagonal), we define edges for pairs of pixels whose Euclidean distance is less than 1.5 pixels, thereby retaining connected neighbors. Next, we implement a Depth-first search (DFS) algorithm to find the longest possible path connecting the opposite endpoints in the graph. If the skeleton is not branched, this path corresponds to the skeleton. For branched skeletons, multiple candidate paths exist. For each candidate path we compute the cell size metrics described in the next section and take the median of each metric as the cell size (Supplementary Fig. 7c).

Cell size measurements:

We combined the mask skeleton and distance transform to derive cell measurements under a spherocylindrical geometry with hemispherical caps and rotational symmetry. The measurements

are implemented in a custom code written in Python3, which can be found in the GitHub repository linked in the "Software and data availability section."

Cell length (L) is the sum of Euclidean distances d between two successive points in the mask skeleton, plus the distance from each skeleton endpoint and the corresponding mask tip, which corresponds to the local radius at the skeleton endpoint r_s. With that, L follows the equation:

$$L = r_{s1} + r_{s2} + \sum_{i=0}^{n-1} d(i,i+1). \qquad \text{Eq. (1)}$$

Cell width (w_i) is calculated as twice the value of the radius at each point in the skeleton (r_i). Mean w can then be calculated as:

$$\bar{w} = \frac{2}{n} \sum_{i=0}^{n} r_{c\,i}. \qquad \text{Eq. (2)}$$

Cell surface area (S) is calculated as the sum of lateral surface areas of the cylinders projected at each point in the skeleton, with radius r_c and height h of 1 pixel, plus the lateral surface area of the hemispherical caps with radius r_s. Thus, S is calculated as:

$$S = 2\pi \left(r_{s1}^2 + r_{s2}^2 + h \sum_{i=0}^{n} r_{c\,i}\right). \qquad \text{Eq. (3)}$$

Cell volume (V) is obtained similarly to S, as it adds the volume of each cylinder and the hemispherical caps (with same parameters of radius and height as S). V is then obtained from:

$$V = \pi\left(\frac{2}{3} r_{s1}^3 + \frac{2}{3} r_{s2}^3 + h \sum_{i=0}^{n} r_{c\,i}^2\right). \qquad \text{Eq. (4)}$$

Measurement transformation

We segmented 864 single cells of *Escherichia coli* stained with FM 4-64 using either JFilament (ground-truth) or FMseg. Cells were from early exponential phase (OD_{600} = 0.1-0.3) cultures in either LB Lennox, or MOPS minimal medium with 0.1% glucose, providing a range of cell widths from 0.53 to 1.04 μm, and a range of cell lengths from 0.87 to 9.61 μm (based on measurements from ground-truth masks) (Supplementary Figure 4a). Measurements from ground-truth masks were plotted against the values obtained from on the FMseg-generated masks, revealing a consistent size overestimation for FMseg masks (Supplementary Figure 4b). The data shows a linear relationship for the overestimation, which allows us to leverage statistical tools to compute the necessary parameters to transform the data between the automatic segmentation results to the ground truth results. A linear regression would be a direct way to do this, extracting the values for

the slope (m) and intercept (n) necessary to transform the data. While linear regression returns a coefficient of error that would allow us to simulate distributions to sample m and n, there is the chance that we may over or under-estimate these values, as well as sample pairs that do not make sense in the context of our data. Thus, we leveraged the PYMC library (Abril-Pla et al., 2023) and its use of Bayesian statistics for sampling the necessary parameters, which are m, the fit slope, n, the fit intercept, and σ, the standard deviation of our sampled observations. Since our data behaves linearly, we used a Bayesian Generalized Linear Model (GLM), assuming that our parameters (slope, intercept, size metric, and its standard deviation) follow certain types of distributions. For σ, we assume a Half-Cauchy distribution, which only returns positive, non-zero values, which is the only type of values a standard deviation can have. For m and n, we sample from normal distributions with mean 0 and standard deviation 20, giving us a wide range of possible parameters. Then, for computing the observed values, we also assume a normal distribution. PYMC performs a series of draws to fit the data, and thus after successive draws we end with the distributions of m, n and σ. In our case, we performed four chains of 3000 draws each, giving us a total of 12000 sampled values. We can then randomly sample pairs from these distributions in order to transform the data, as well as to include the uncertainty of the transformation in the analysis (Supplementary Figure 4b). For the examples shown in Figure 5, Figure 6, and Supplementary Figure 6, we sampled 250 pairs of m and n for each individual cell measurement, thus allowing us to compute new metrics.

CryoEM image analysis

Because the high magnification of the cryo-EM images typically did not capture entire cells, we estimated cell width as the straight-line distance between the mother-cell lateral membranes, measured perpendicular to the cell longitudinal axis. In cryo-FIB tomograms of wild-type cells, only a portion of the cellular volume was retained: focused ion-beam milling produced a thin lamella that sampled a section of the cell. We analyzed only tomograms in which this section included the mid-cell region. To verify this, we measured width across successive z-slices and selected tomograms showing a clear width peak at mid-cell that decreased in slices above and below.

Whole genome sequencing

For whole-genome sequencing, *Priestia megaterium* WH320 was streaked in an LB-Lennox agar plate and grown overnight at 30°C. Next morning, a single colony was picked and re-streaked in an LB-Lennox agar plate and grown overnight at 30°C. Cells were collected from the plate and genomic

DNA was extracted with the MasterPure Gram Positive DNA Purification Kit (Biosearch technologies). Extracted genome was then sequenced using an Illumina NextSeq 500 sequencer. Raw pair-end sequences were loaded into Geneious prime version 2025.1.2 for sequence trimming and pairing with an insert size of 500 base pairs. Paired sequence list was then exported in fastq format for analysis using breseq (Deatherage & Barrick, 2014).

Software and data availability

The segmentation (FMseg), deconvolution prediction (FM2FM) and fluorescence translation (FP2FM) models can be found in the Zenodo (https://zenodo.org/uploads/17397501). Segmentation model training and benchmarking data can be found at the BioImage Archive repository https://www.ebi.ac.uk/biostudies/studies/S-BIAD2350. Deconvolution model training and benchmarking data can be found at the BioImage Archive repository https://www.ebi.ac.uk/biostudies/studies/S-BIAD2353. Images of individual *E. coli* cells used for measurement transformation with their corresponding annotations and automatic segmentations be found at the BioImage Archive repository https://www.ebi.ac.uk/biostudies/studies/S-BIAD2352. Segmentation of *B. subtilis* and *P. megaterium* strains (wild-type and induction constructs) be found at the BioImage Archive repository https://www.ebi.ac.uk/biostudies/studies/S-BIAD2354.

The MEDUSSA GitHub repository contains all software-related material (https://github.com/OReyesMatte/MEDUSSA). The "MEDUSSA" folder contains the functions for size measurement (measure.py), Bayesian sampling, and measurement transformation (both in transform.py). The "CARE" folder contains the environment file which was used for training the deconvolution models, as well as the corresponding notebooks for preparing the data, train the model, and test it. The "Omnipose" folder contains the environment file which was used for training the segmentation models, the command that was used for training in the CLI, and an example notebook to run segmentation. The "Figures" folder contains notebooks to reproduce the plots shown in the manuscript. *P. megaterium* WH320 sequencing data, *ponA* protein sequence alignment, and results from breseq (Deatherage & Barrick, 2014) are available in Zenodo (https://zenodo.org/uploads/17392073).

ACKNOWLEDGMENTS

We are grateful to Elitza Tocheva, Elizabeth Villa and Kanika Khanna for kindly sharing CryoEM images of *B. subtilis* sporulating cells, Kaumudi H. Prabhakara for providing the *Saccharomyces cerevisiae* and *Lactococcus lactis* strains used in Fig. 2d, and Felipe Sinhorini for helping with the acquisition of the BGSC strains. We thank members of the Evolutionary Cell Biology group, the Ojkic lab, and the Microbial Population Biology department for continuous discussion and feedback. We also thank the members of the Scientific IT group at the Max Planck Institute for Evolutionary Biology for computational support. Work in JLG lab was supported by ERC starting grant 853323. NO was supported by BBSRC grant (BB/Y009002/1).

AUTHOR CONTRIBUTIONS

Conceptualization, ORM, NO, JLG; Methodology, ORM, JLG; Software, ORM, CFG, BG; Validation, ORM; Formal analysis, ORM, CFG, BG; Investigation, ORM, NH; Resources, NH; Data Curation, ORM; Writing - Original Draft, JLG; Writing - Review & Editing, All authors; Visualization, ORM, NO, JLG; Supervision, NO, JLG; Project administration, ORM, JLG; Funding acquisition, JLG.

COMPETING INTERESTS

The authors declare no competing interests.

REFERENCES

- Abril-Pla, O., Andreani, V., Carroll, C., Dong, L., Fonnesbeck, C. J., Kochurov, M., Kumar, R., Lao, J., Luhmann, C. C., Martin, O. A., Osthege, M., Vieira, R., Wiecki, T., & Zinkov, R. (2023). PyMC: a modern, and comprehensive probabilistic programming framework in Python. *PeerJ Computer Science*, *9*, e1516.

- Agard, D. A. (1984). Optical sectioning microscopy: cellular architecture in three dimensions. *Annual Review of Biophysics and Bioengineering*, *13*, 191-219.

- Arya, G., Petronella, N., Crosthwait, J., Carrillo, C. D., & Shwed, P. S. (2014). Draft genome sequence of *Bacillus megaterium* type strain ATCC 14581. *Genome Announcements*, *2*, e01124-01114.

- Blum, H. (1967). A transformation for extracting new descriptors of shape. In W. Walthen-Dunn (Ed.), *Models for the Perception of Speech and Visual Form* (pp. 362-380). MIT Press.

- Borgefors, G. (1986). Distance transformations in digital images. *Computer Vision, Graphics, and Image Processing*, *34*, 344-371.

- Caicedo, J. C., Roth, J., Goodman, A., Becker, T., Karhohs, K. W., Broisin, M., Molnar, C., McQuin, C., Singh, S., Theis, F. J., & Carpenter, A. E. (2019). Evaluation of deep learning strategies for nucleus segmentation in fluorescence images. *Cytometry*, *95*, 952-965.

- Calvez, P., Jouhet, J., Vie, V., Durmort, C., & Zapun, A. (2019). Lipid phases and cell geometry during the cell cycle of *Streptococcus pneumoniae*. *Frontiers in Microbiology*, *10*, 351.

- Campos, M., Surovtsev, Ivan V., Kato, S., Paintdakhi, A., Beltran, B., Ebmeier, Sarah E., & Jacobs-Wagner, C. (2014). A constant size extension drives bacterial cell size homeostasis. *Cell*, *159*, 1433-1446.

- Chien, A. C., Hill, N. S., & Levin, P. A. (2012). Cell size control in bacteria. *Current Biology*, *22*, R340-R349.

- Cox, R. S., 3rd, Dunlop, M. J., & Elowitz, M. B. (2010). A synthetic three-color scaffold for monitoring genetic regulation and noise. *Joural of Biological Engineering*, *4*, 10.

- Cutler, K. J., Stringer, C., Lo, T. W., Rappez, L., Stroustrup, N., Brook Peterson, S., Wiggins, P. A., & Mougous, J. D. (2022). Omnipose: a high-precision morphology-independent solution for bacterial cell segmentation. *Nature Methods*, *19*, 1438-1448.

- Dalebroux, Z. D., Edrozo, M. B., Pfuetzner, R. A., Ressl, S., Kulasekara, B. R., Blanc, M. P., & Miller, S. I. (2015). Delivery of cardiolipins to the *Salmonella* outer membrane is necessary for survival within host tissues and virulence. *Cell Host Microbe*, *17*, 441-451.

- de Bary, A. (1884). *Vergleichende Morphologie und Biologie der Pilze, Mycetozoen und Bacterien.* Engelmann.

- Deatherage, D. E., & Barrick, J. E. (2014). Identification of mutations in laboratory-evolved microbes from next-generation sequencing data using breseq. *Methods in Molecular Biology*, *1151*, 165-188.

- Delaby, M., Yang, L., Jacq, M., Gallagher, K. A., Kysela, D. T., Hughes, V., Pulido, F., Veyrier, F. J., VanNieuwenhze, M. S., & Brun, Y. V. (2025). Phenotypic plasticity in cell elongation among closely related bacterial species. *Nature Communications*, *16*, 5099.

- Dion, M. F., Kapoor, M., Sun, Y., Wilson, S., Ryan, J., Vigouroux, A., van Teeffelen, S., Oldenbourg, R., & Garner, E. C. (2019). *Bacillus subtilis* cell diameter is determined by the opposing actions of two distinct cell wall synthetic systems. *Nature Microbiology*, *4*, 1294-1305.

- Ducret, A., Quardokus, E. M., & Brun, Y. V. (2016). MicrobeJ, a tool for high throughput bacterial cell detection and quantitative analysis. *Nature Microbiology*, *1*, 16077.

- Eppinger, M., Bunk, B., Johns, M. A., Edirisinghe, J. N., Kutumbaka, K. K., Koenig, S. S., Creasy, H. H., Rosovitz, M. J., Riley, D. R., Daugherty, S., Martin, M., Elbourne, L. D., Paulsen, I., Biedendieck, R., Braun, C., Grayburn, S., Dhingra, S., Lukyanchuk, V., Ball, B.,…Vary, P. S. (2011). Genome sequences of the biotechnologically important *Bacillus megaterium* strains QM B1551 and DSM319. *Journal of Bacteriology*, *193*, 4199-4213.

- Falk, T., Mai, D., Bensch, R., Cicek, O., Abdulkadir, A., Marrakchi, Y., Bohm, A., Deubner, J., Jackel, Z., Seiwald, K., Dovzhenko, A., Tietz, O., Dal Bosco, C., Walsh, S., Saltukoglu, D., Tay, T. L., Prinz, M., Palme, K., Simons, M.,…Ronneberger, O. (2019). U-Net: deep learning for cell counting, detection, and morphometry. *Nature Methods*, *16*, 67-70.

- Gray, W. T., Govers, S. K., Xiang, Y., Parry, B. R., Campos, M., Kim, S., & Jacobs-Wagner, C. (2019). Nucleoid size scaling and intracellular organization of translation across bacteria. *Cell*, *177*, 1632-1648.

- Greenspan, P., & Fowler, S. D. (1985). Spectrofluorometric studies of the lipid probe, nile red. *Journal of lipid research*, *26*, 781-789.

- Gunawan, I., Marsh, R. J., Aggarwal, N., Meijering, E., Cox, S., Lock, J. G., & Culley, S. (2025). Image quality metrics fail to accurately represent biological information in fluorescence microscopy. *bioRxiv*, 2025.2008.2005.668508.

- Hardo, G., Li, R., & Bakshi, S. (2024). Quantitative microbiology with widefield microscopy: navigating optical artefacts for accurate interpretations. *Npj Imaging*, *2*, 26.

- Herman, P., Konopasek, I., Plasek, J., & Svobodova, J. (1994). Time-resolved polarized fluorescence studies of the temperature adaptation in *Bacillus subtilis* using DPH and TMA-DPH fluorescent probes. *Biochimica et biophysica acta*, *1190*, 1-8.

- Holt, J. H., Krieg, N. R., Sneath, P. H. a., Staley, J. T., & Williams, S. T. (1994). *Bergey's manual of determinative bacteriology ninth edition* (9th ed.). Lippincott Williams & Wilkins.

- Jiang, C., Caccamo, P. D., & Brun, Y. V. (2015). Mechanisms of bacterial morphogenesis: evolutionary cell biology approaches provide new insights. *BioEssays*, *37*, 413-425.

- Juillot, D., Cornilleau, C., Deboosere, N., Billaudeau, C., Evouna-Mengue, P., Lejard, V., Brodin, P., Carballido-Lopez, R., & Chastanet, A. (2021). A high-content microscopy screening identifies new genes involved in cell width control in *Bacillus subtilis*. *mSystems*, *6*, e0101721.

- Kearns, D. B., & Losick, R. (2005). Cell population heterogeneity during growth of *Bacillus subtilis*. *Genes and Development*, *19*, 3083-3094.

- Khanna, K., Lopez-Garrido, J., Sugie, J., Pogliano, K., & Villa, E. (2021). Asymmetric localization of the cell division machinery during *Bacillus subtilis* sporulation. *eLife*, *10*, e62204.

- Khanna, K., Lopez-Garrido, J., Zhao, Z., Watanabe, R., Yuan, Y., Sugie, J., Pogliano, K., & Villa, E. (2019). The molecular architecture of engulfment during *Bacillus subtilis* sporulation. *eLife*, *8*, e45257.

- Kuru, E., Hughes, H. V., Brown, P. J., Hall, E., Tekkam, S., Cava, F., de Pedro, M. a., Brun, Y. V., & VanNieuwenhze, M. S. (2012). In Situ probing of newly synthesized peptidoglycan in live bacteria with fluorescent D-amino acids. *Angewandte Chemie*, *51*, 12519-12523.

- Liu, L., Jiang, H., He, P., Chen, W., Liu, X., Gao, J., & Han, J. (2019). On the variance of the adaptive learning rate and beyond. *arXiv*, 1908.03265.

- Logan, N. A., & Vos, P. D. (2015). *Bacillus*. In *Bergey's Manual of Systematics of Archaea and Bacteria* (pp. 1-163).

- López, A. C., & Alippi, A. M. (2009). Diversity of *Bacillus megaterium* isolates cultured from honeys. *LWT - Food Science and Technology*, *42*, 212-219.

- Lopez-Garrido, J., Ojkic, N., Khanna, K., Wagner, F. R., Villa, E., Endres, R. G., & Pogliano, K. (2018). Chromosome translocation inflates *Bacillus* forespores and impacts cellular morphology. *Cell*, *172*, 758-770.

- Lucidi, M., Capecchi, G., Visaggio, D., Gasperi, T., Parisi, M., Cincotti, G., Rampioni, G., Visca, P., & Kolmakov, K. (2024). Expanding the microbiologist toolbox via new far-red-emitting dyes suitable for bacterial imaging. *Microbiology Spectrum*, *12*, e0369023.

- Lugagne, J. B., Lin, H., & Dunlop, M. J. (2020). DeLTA: Automated cell segmentation, tracking, and lineage reconstruction using deep learning. *PLoS Comput Biol*, *16*, e1007673.

- Lynch, M., Field, M. C., Goodson, H. V., Malik, H. S., Pereira-Leal, J. B., Roos, D. S., Turkewitz, A. P., & Sazer, S. (2014). Evolutionary cell biology: Two origins, one objective. *Proceedings of the National Academy of Sciences*, *111*, 16990-16994.

- Ma, L., Conover, M., Lu, H., Parsek, M. R., Bayles, K., & Wozniak, D. J. (2009). Assembly and development of the *Pseudomonas aeruginosa* biofilm matrix. *PLoS Pathogens*, *5*, e1000354.

- Maier-Hein, L., Reinke, A., Godau, P., Tizabi, M. D., Buettner, F., Christodoulou, E., Glocker, B., Isensee, F., Kleesiek, J., Kozubek, M., Reyes, M., Riegler, M. A., Wiesenfarth, M., Kavur, A. E., Sudre, C. H., Baumgartner, M., Eisenmann, M., Heckmann-Notzel, D., Radsch, T.,…Jager, P. F. (2024). Metrics reloaded: recommendations for image analysis validation. *Nature Methods*, *21*, 195-212.

- Maloney, E., Madiraju, M., & Rajagopalan, M. (2009). Overproduction and localization of *Mycobacterium tuberculosis* ParA and ParB proteins. *Tuberculosis (Edinb)*, *89 Suppl 1*(Suppl 1), S65-69.

- Margulis, L., Jorgensen, J. Z., Dolan, S., Kolchinsky, R., Rainey, F. A., & Lo, S.-C. (1998). The *Arthromitus* stage of *Bacillus cereus:* Intestinal symbionts of animals. *Proceedings of the National Academy of Sciences*, *95*, 1236-1241.

- Monteiro, J. M., Fernandes, P. B., Vaz, F., Pereira, A. R., Tavares, A. C., Ferreira, M. T., Pereira, P. M., Veiga, H., Kuru, E., Vannieuwenhze, M. S., Brun, Y. V., Filipe, S. R., & Pinho, M. G. (2015). Cell shape dynamics during the staphylococcal cell cycle. *Nature Communications*, *6*, 1-12.

- Murumkar, D. R., Borkar, S. G., & Chimote, V. P. (2012). Diversity for cell morphology, in-vitro phosphate solubilization and DNA profile of *Bacillus megaterium* isolates present in soils of Maharashtra. *World of Applied Sciences Journal*, *17*, 776-785.

- Nordholt, N., van Heerden, J. H., & Bruggeman, F. J. (2020). Biphasic cell-size and growth-rate homeostasis by single *Bacillus subtilis* cells. *Current Biology*, *30*, 2238-2247 e2235.

- O'Connor, O. M., Alnahhas, R. N., Lugagne, J. B., & Dunlop, M. J. (2022). DeLTA 2.0: A deep learning pipeline for quantifying single-cell spatial and temporal dynamics. *PLoS Computational Biology*, *18*, e1009797.

- O'Hara, B. J., Barth, Z. K., McKitterick, A. C., & Seed, K. D. (2017). A highly specific phage defense system is a conserved feature of the *Vibrio cholerae* mobilome. *PLoS Genetics*, *13*, e1006838.

- Ojkic, N., & Banerjee, S. (2021). Bacterial cell shape control by nutrient-dependent synthesis of cell division inhibitors. *Biophysical Journal*, *120*, 2079-2084.

- Ojkic, N., Lopez-Garrido, J., Pogliano, K., & Endres, R. G. (2014). Bistable forespore engulfment in *Bacillus subtilis* by a zipper mechanism in absence of the cell wall. *PLoS Computational Biology*, *10*, e1003912.

- Ojkic, N., Lopez-Garrido, J., Pogliano, K., & Endres, R. G. (2016). Cell-wall remodeling drives engulfment during *Bacillus subtilis* sporulation. *eLife*, *5*, e18657.

- Ojkic, N., Serbanescu, D., & Banerjee, S. (2019). Surface-to-volume scaling and aspect ratio preservation in rod-shaped bacteria. *eLife*, *8*, e47033.

- Ojkic, N., Serbanescu, D., & Banerjee, S. (2022). Antibiotic resistance via bacterial cell shape-shifting. *mBio*, *13*, e0065922.

- Ollion, J., Maliet, M., Giuglaris, C., Vacher, É., & Deforet, M. (2024). DiSTNet2D: Leveraging long-range temporal information for efficient segmentation and tracking. *PRX Life*, *2*, 023004.

- Paintdakhi, A., Parry, B., Campos, M., Irnov, I., Elf, J., Surovtsev, I., & Jacobs-Wagner, C. (2016). Oufti: an integrated software package for high-accuracy, high-throughput quantitative microscopy analysis. *Molecular Microbiology*, *99*, 767-777.

- Panigrahi, S., Murat, D., Le Gall, A., Martineau, E., Goldlust, K., Fiche, J. B., Rombouts, S., Nollmann, M., Espinosa, L., & Mignot, T. (2021). Misic, a general deep learning-based method for the high-throughput cell segmentation of complex bacterial communities. *eLife*, *10*, e65151.

- Pogliano, J., Osborne, N., Sharp, M. D., Abanes-De Mello, a., Perez, a., Sun, Y. L., & Pogliano, K. (1999). A vital stain for studying membrane dynamics in bacteria: a novel mechanism controlling septation during *Bacillus subtilis* sporulation. *Molecular Microbiology*, *31*, 1149-1159.

- Popham, D. L., & Setlow, P. (1996). Phenotypes of *Bacillus subtilis* mutants lacking multiple class A high-molecular-weight penicillin-binding proteins. *Journal of Bacteriology*, *178*, 2079-2085.

- Riley, E. P., Trinquier, A., Reilly, M. L., Durchon, M., Perera, V. R., Pogliano, K., & Lopez-Garrido, J. (2018). Spatiotemporally regulated proteolysis to dissect the role of vegetative proteins during *Bacillus subtilis* sporulation: cell-specific requirement of σ^H and σ^A. *Molecular Microbiology*, *108*, 45-62.

- Ronneberger, O., Fischer, P., & Brox, T. (2015). U-Net: Convolutional networks for biomedical image segmentation. In N. Navab, J. Hornegger, W. M. Wells, & A. F. Frangi, *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015* Cham.

- Rygus, T., Scheler, A., Allmansberger, R., & Hillen, W. (1991). Molecular cloning, structure, promoters and regulatory elements for transcription of the *Bacillus megaterium* encoded regulon for xylose utilization. *Archives of Microbiology*, *155*, 526-534.

- Sauls, J. T., Cox, S. E., Do, Q., Castillo, V., Ghulam-Jelani, Z., & Jun, S. (2019). Control of *Bacillus subtilis* replication initiation during physiological transitions and perturbations. *mBio*, *10*, 10.1128/mbio.02205-02219.

- Scalettar, B. A., Swedlow, J. R., Sedat, J. W., & Agard, D. A. (1996). Dispersion, aberration and deconvolution in multi-wavelength fluorescence images. *Journal of Microscopy*, *182*, 50-60.

- Schavemaker, P. E., & Munoz-Gomez, S. A. (2022). The role of mitochondrial energetics in the origin and diversification of eukaryotes. *Nature Ecology and Evolution*, *6*, 1307-1317.

- Schmidt, U., Weigert, M., Broaddus, C., & Myers, G. (2018). Cell detection with star-convex polygons. In A. F. Frangi, J. A. Schnabel, C. Davatzikos, C. Alberola-López, & G. Fichtinger, *Medical Image Computing and Computer Assisted Intervention – MICCAI 2018* Cham.

- Schober, I., Koblitz, J., Sarda Carbasse, J., Ebeling, C., Schmidt, M. L., Podstawka, A., Gupta, R., Ilangovan, V., Chamanara, J., Overmann, J., & Reimer, L. C. (2025). BacDive in 2025: the core database for prokaryotic strain data. *Nucleic Acids Research*, *53*(D1), D748-D756.

- Sega, G. A. (1984). A review of the genetic effects of ethyl methanesulfonate. *Mutation Research*, *134*, 113-142.

- Sharp, M. D., & Pogliano, K. (1999). An in vivo membrane fusion assay implicates SpoIIIE in the final stages of engulfment during *Bacillus subtilis* sporulation. *Proceedings of the National Academy of Sciences*, *96*(25), 14553-14558.

- Shin, J. Y., Lopez-Garrido, J., Lee, S.-H. H., Diaz-Celis, C., Fleming, T., Bustamante, C., & Pogliano, K. (2015). Visualization and functional dissection of coaxial paired SpoIIIE channels across the sporulation septum. *eLife*, *4*, e06474.

- Smith, M. B., Li, H., Shen, T., Huang, X., Yusuf, E., & Vavylonis, D. (2010). Segmentation and tracking of cytoskeletal filaments using open active contours. *Cytoskeleton*, *67*, 693-705.

- Spahn, C., Gomez-de-Mariscal, E., Laine, R. F., Pereira, P. M., von Chamier, L., Conduit, M., Pinho, M. G., Jacquemet, G., Holden, S., Heilemann, M., & Henriques, R. (2022). DeepBacs for multi-task bacterial image analysis using open-source deep learning approaches. *Communications Biology*, *5*, 688.

- Specht, M., Dempwolff, F., Schatzle, S., Thomann, R., & Waidner, B. (2013). Localization of FtsZ in *Helicobacter pylori* and consequences for cell division. *Journal of Bacteriology*, *195*, 1411-1420.

- Stahl, U., & Esser, K. (1983). Plasmid heterogeneity in various strains of *Bacillus megaterium*. *European journal of applied microbiology and biotechnology*, *17*, 248-251.

- Stringer, C., & Pachitariu, M. (2025). Cellpose3: one-click image restoration for improved cellular segmentation. *Nature Methods*, *22*, 592-599.

- Stringer, C., Wang, T., Michaelos, M., & Pachitariu, M. (2021). Cellpose: a generalist algorithm for cellular segmentation. *Nature Methods*, *18*, 100-106.

- Studer, P., Staubli, T., Wieser, N., Wolf, P., Schuppler, M., & Loessner, M. J. (2016). Proliferation of *Listeria monocytogenes* L-form cells by formation of internal and external vesicles. *Nature Communications*, *7*, 13631.

- Stylianidou, S., Brennan, C., Nissen, S. B., Kuwada, N. J., & Wiggins, P. A. (2016). SuperSegger: robust image segmentation, analysis and lineage tracking of bacterial cells. *Molecular Microbiology*, *102*, 690-700.

- Sun, Q., & Margolin, W. (1998). FtsZ dynamics during the division cycle of live *Escherichia coli* cells. *Journal of Bacteriology*, *180*, 2050-2056.

- Syvertsson, S., Vischer, N. O., Gao, Y., & Hamoen, L. W. (2016). When phase contrast fails: ChainTracer and NucTracer, two ImageJ methods for semi-automated single cell analysis using membrane or DNA staining. *PLoS ONE*, *11*, e0151267.

- Taha, A. A., & Hanbury, A. (2015). Metrics for evaluating 3D medical image segmentation: analysis, selection, and tool. *BMC Medical Imaging*, *15*, 29.

- Taheri-Araghi, S., Bradde, S., Sauls, J. T., Hill, N. S., Levin, P. A., Paulsson, J., Vergassola, M., & Jun, S. (2015). Cell-size control and homeostasis in bacteria. *Current Biology*, *25*, 385-391.

- Tocheva, E. I., Lopez-Garrido, J., Hughes, H. V., Fredlund, J., Kuru, E., Vannieuwenhze, M. S., Brun, Y. V., Pogliano, K., & Jensen, G. J. (2013). Peptidoglycan transformations during *Bacillus subtilis* sporulation. *Molecular Microbiology*, *88*, 673-686.

- Tsuge, Y., Ogino, H., Teramoto, H., Inui, M., & Yukawa, H. (2008). Deletion of cgR_1596 and cgR_2070, encoding NlpC/P60 proteins, causes a defect in cell separation in *Corynebacterium glutamicum* R. *Journal of Bacteriology*, *190*, 8204-8214.

- Ursell, T., Lee, T. K., Shiomi, D., Shi, H., Tropini, C., Monds, R. D., Colavin, A., Billings, G., Bhaya-Grossman, I., Broxton, M., Huang, B. E., Niki, H., & Huang, K. C. (2017). Rapid, precise quantification of bacterial cellular dimensions across a genomic-scale knockout library. *BMC Biology*, *15*, 17.

- Van Valen, D. A., Kudo, T., Lane, K. M., Macklin, D. N., Quach, N. T., DeFelice, M. M., Maayan, I., Tanouchi, Y., Ashley, E. A., & Covert, M. W. (2016). Deep learning automates the quantitative analysis of individual cells in live-cell imaging experiments. *PLoS Computational Biology*, *12*, e1005177.

- Weigert, M., Schmidt, U., Boothe, T., Muller, A., Dibrov, A., Jain, A., Wilhelm, B., Schmidt, D., Broaddus, C., Culley, S., Rocha-Martins, M., Segovia-Miranda, F., Norden, C., Henriques, R., Zerial, M., Solimena, M., Rink, J., Tomancak, P., Royer, L.,…Myers, E. W. (2018). Content-aware image restoration: pushing the limits of fluorescence microscopy. *Nature Methods*, *15*, 1090-1097.

- Yeh, Y. C., Comolli, L. R., Downing, K. H., Shapiro, L., & McAdams, H. H. (2010). The *Caulobacter* Tol-Pal complex is essential for outer membrane integrity and the positioning of a polar localization factor. *Journal of Bacteriology*, *192*, 4847-4858.

- Zhang, T. Y., & Suen, C. Y. (1984). A fast parallel algorithm for thinning digital patterns. *Communications of the ACM*, *27*, 236–239.