Quantitating pathogenic biofilm architecture in biopsied tissue

Shareef M. Dabdoub*
The Ohio State University Biophysics Program

Brian A. VanderBrink
Nationwide Children's

Sheryl S. Justice
Center for Microbial Pathogenesis at
Nationwide Children's Research Institute

William C. Ray†
Battelle Center for Mathematical Medicine at
Nationwide Children's Research Institute

ABSTRACT

Accurate diagnosis and treatment of biofilm infections require identification of the pathogenic organism(s) as well as determining the progress of the disease. Current tools in clinical use, including culturing and PCR tests, are extremely useful for identifying organisms, but are destructive in nature – resulting in the loss of important information regarding biofilm architecture and state. Improving clinical understanding of these, often treatment-resistant, infections is of great importance, and new non-destructive imaging-based tools must be developed in order to gather crucial information regarding disease.

Here we present new software, *ProkaryMetrics*, designed to take advantage of available microscopy imaging modalities, providing a unique platform for 3D imaging and analysis of biofilm samples. We demonstrate the software capabilities by analysis of murine tissue biopsy samples containing uropathogenic *Escherichia coli* biofilms: wild type UTI89 and UTI89Δ*kpsF* strains. Using *ProkaryMetrics*, we establish significant architectural differences with qualitative 3D visualizations as well as quantitative measurements including volumetric biofilm size, bacterial counts, community density, orientations, and lengths.

Index Terms: J.3 [Computer Applications]: Life and Medical Science—Biology and genetics; D.2.11 [Software]: Software Architectures—Domain-specific architectures

1 INTRODUCTION

Proper treatment of an infectious disease requires the identification of the causative agent and state of the disease, as well as the susceptibility of the organism to standard treatment. While culture-based techniques remain the gold standard for identification of bacterial and fungal pathogens, diagnostically significant features of an infection also include aspects of the current activity and state of the pathogen, in addition to its simple identity as available through culturing [1]. There is increasing recognition that the state information lost through the culturing process can be critical for properly identifying causative agents and appropriate treatment. However, there is a dearth of quantitative approaches to acquiring state-related measures such as pathogen morphology, biofilm/community organization, and cellular localization from pathology specimens. The current approaches rely either on automated applications of computer-vision, or on manual applications of expert-user visual assessments from (typically) serial microscopy/histology sections. Unfortunately, there are significant impediments to both of these approaches, as, at the diagnostic endpoint there is insufficient homogeneity across either samples or image-acquisition systems for any automated system to be universally, or even widely successful, and simultaneously there is sufficient variation in clinical-user expertise that evaluations from different experts are not quantitatively comparable. Until significantly more sophisticated imaging capabilities are routinely available to endpoint clinical caregivers, any successful approach to integrating quantitative assessments of pathogen state information into treatment decisions, will require systems that can extract quantitatively comparable data from numerous disparate imaging systems and imaging modalities, without requiring more than a lay expertise in applying or adapting the computational approach. Explicitly, we propose that enabling a rural physician with a white-light microscope to, with human intervention, make quantitatively comparable measurements of clinically relevant variables, to those produced by a research laboratory with a scanning confocal instrument, is far more clinically useful than developing an automated approach for the confocal data alone.

To this end, we introduce *ProkaryMetrics*, a Visual Analytics tool for extracting quantitative measures of pathogen community morphology, density and architecture from microscopy images. *ProkaryMetrics* leverages straightforward computer vision and volume segmentation/visualization approaches that can be applied on commodity hardware, to provide a guided interface through which a human expert can rapidly annotate salient pathogen/community features for quantitative analysis. By applying algorithmic volume segmentation/visualization as a guide, rather than as a direct producer of quantitative results, *ProkaryMetrics* can be applied to input data across a wide range of imaging modalities, resolutions, histological approaches, and ultimately absolute quality, without requiring modification of the algorithm, or adaptation of numerous parameters. By guiding the user to make specific quantitative measures, rather than relying on subjective expert assessments, *ProkaryMetrics* can be applied by users with widely varying expertise levels, and still produce quantitatively comparable results.

In this manuscript we validate *ProkaryMetrics* for quantitating two pathogen state variables of known clinical importance. The first is the morphology of the organism. The ability of the microorganism to alter its size, by regulation of cell division, provides advantages during disease. "Morphological plasticity" is a well-known survival strategy for fungal pathogens. Its utility for bacterial pathogens is becoming evident through studies of persistence of uropathogenic *Escherichia coli* [6] (UPEC), and Mycobacterium tuberculosis [4]. In addition to being resistant to the host immune response, filamentous morphotypes of organisms are typically resistant to antibiotics even when their non-filamentous progeny are sensitive. This inherent resistance to killing underscores the importance of determining the prevalence of the filamentous morphotypes in infected samples, as successful treatment regimens must be adapted to eliminate these tenacious survivors. The second is the morphology and architecture (in terms of coherent organization) of the pathogenic community. While only gaining widespread acceptance in the past decade, pathogen community architecture, and internal and external organization, have become well understood as directly modulating the effect and effectiveness of treatments for

*e-mail: dabdoub.2@osu.edu
†e-mail:ray.29@osu.edu

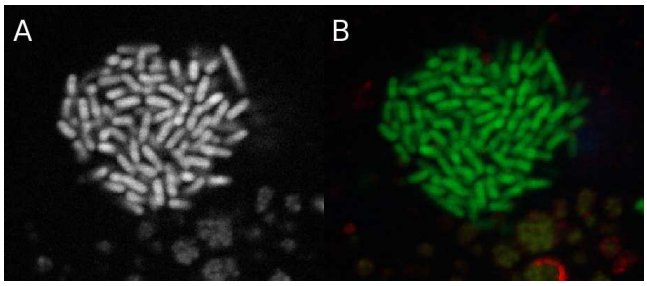

Figure 1: Optical sections of UPEC infected murine bladder biopsy. (A) The green channel of a UTI89 wild type IBC. (B) The same IBC reimaged with a lower-resolution, 3-color modality.

infection.

By leading the user to recognize the general structure and organization of a pathogenic community – a task that can be accomplished with a generalized visualization algorithm that is adequate for descriptive rather than quantitative purposes – and enabling the user to then make specific quantifiable assessments and measures within this generalized presentation, *ProkaryMetrics* provides a mechanism for quantitation of these population-state traits that can be consistently applied with existing technology in typical clinical settings. Our results demonstrate that this approach is adequate to differentiate uropathogenic *E. coli* morphotypes and strains, and to extract quantitatively comparable assessments of community state variables, using both state-of-the-art research, and typical clinical microscopy samples.

2 Software Design

The guiding requirement for *ProkaryMetrics* has been, and continues to be, that the system enable end-user practitioners, without programming experience, and with varying imaging modalities at their disposal, to make quantitatively similar assessments of pathogen/community state traits. As such, the system presupposes nothing more than that the user can acquire one or more digital images of the infected tissue, at sufficient resolution and quality that the user can differentiate individual members of the community by visual inspection. Simple volume segmentation algorithms, and volume visualization approaches enable the user to interactively explore and annotate the pathogen population within the images, with as much dimensional detail as is available in the images themselves. These algorithms and approaches are required only to provide guidance and support for the user's identification of pathogen features, and to quantitatively report the identifications. They are at no point required to automatically determine quantitative features without user guidance.

While this assisted-manual-analysis methodology requires user interaction for every diagnostic analysis, we propose that this is entirely appropriate for clinical applications. Not only does this facilitate timely analysis of samples that are not amenable to automated approaches, it is a practical requirement that any automated clinical diagnostic based on data with quality as variable as microscopy imaging, must be confirmed by inspection by a human expert. Since such inspection is necessarily visual, it is no impediment that our *ProkaryMetrics* approach starts with this process.

3 Software Implementation

ProkaryMetrics is written entirely in Python, relying on the Visualization Toolkit (VTK) [8] to enable visualization of and interaction with volumetric data. Users begin by loading volumetric image data, typically as a series of single-channel 2-dimensional image slices of the sample to be studied. This data is first smoothed using a Gaussian filter, and then isosurfaced with a user-modifiable target

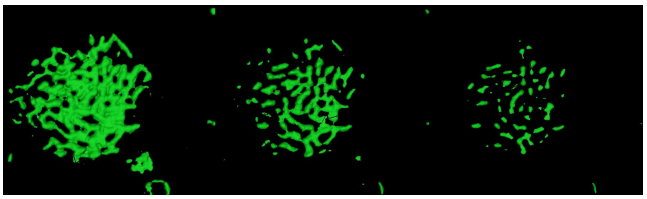

Figure 2: No single isosurface value is appropriate for automated analysis. However, a human assisted analysis of a range of isosurfaces, as shown here, produces nearly identical results for both data sets.

pixel intensity value. This volumetric surface rendering is displayed in the visualization window allowing the user to manipulate and explore the data in 3D. The system provides a cursor controlled by the mouse that attaches itself to the nearest rendered surface through ray tracing. In this manner, the user can simply click on a surface outlining a bacterium to place a spherical marker object. One or more markers can then be recorded as representing a bacterium.

4 Analysis

In order to allow comparison between samples, *ProkaryMetrics* supplies a suite of numeric and statistical tools for investigating the mathematical properties of the biofilms. Using the Khachiyan method [7] for calculating a bounding ellipsoid $E \subseteq \mathbb{R}^n$ for a set of m points, users are enabled to estimate the volume of space occupied by the biofilm under investigation. Furthermore, assuming a standard width and depth for the bacteria, and the length provided by the user, we can estimate the total volume occupied by the sum of the individuals. Combined with the ratio of the two volumes we can present a quantitative picture of the size, shape, and relative packing density of the bacterial community (see Figure 3).

Using the main axis vector of each bacterium, we calculate its scalar projection in the direction of each of the orthonormal basis vectors in $\mathbb{R}^3$. Gathering this information for all of the recorded bacteria, we can compile statistics on the overall layout and orientations of the bacteria within a community. We use these three projections to set the RGB components of the diffuse color of the corresponding bacterium. The resulting visual representation (Figure 4) provides a clear indication of the general orientation trends of the community organization, as well as a means to compare visually between different samples.

Finally, using the midpoint of each bacterium, we borrow a technique from the field of data mining to calculate the average interpoint Euclidean distance (Equation 1), giving another measure of bacterial community packing and another means to compare between communities.

$$d(x_1, x_2) = \sqrt{\sum_{j=1}^{m} (x_{ij} - x_{2j})^2} \tag{1}$$

5 Results

We have applied *ProkaryMetrics* to the visualization and analysis of UPEC, the major causative agent of urinary tract infections (UTIs). UPEC causes both acute and recurring (mainly in women) UTIs, and results in billions of dollars in medical costs and lost productivity annually [5]. These infections are particularly difficult to treat because UPEC has evolved highly effective means for evading host defenses, as well as medical treatment. The major components of their evasion strategy centers on intracellular invasion of the superficial epithelial cells of the host bladder and morphological change by filamentation. During filamentation, bacteria continue to

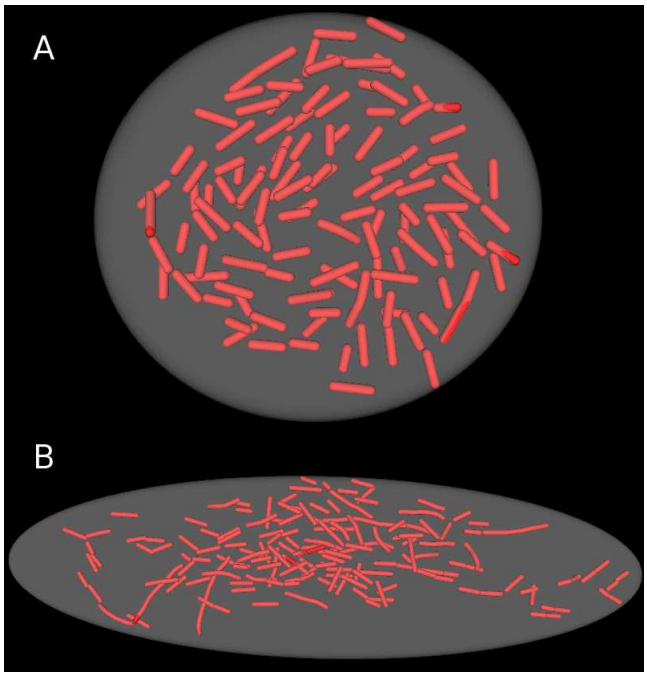

Figure 3: Minimum volume bounding ellipsoids calculated for the user-marked bacteria in (A) a wild type UTI89 IBC and (B) a capsule mutant (UTI89Δ*kpsF*) IBC. The wild type fills a volume of space approximately $1.95x10^3\mu m^3$ and is nearly perfectly circular in the x and y dimensions, making it an oblate spheroid. The kpsF mutant is much larger, occupying a volume of $3.87x10^4\mu m^3$, with a clear bias in one dimension.

grow but are unable to complete the process of division, producing a long strand of conjoined bacteria up to approximately $70\mu m$ [5]. Establishment of intracellular bacterial communities (IBCs) allows UPEC to avoid the hostile environment of the bladder until the late stages of infection when the host cells begin to porate and apoptose[5]. At this point, the bacteria are exposed and filaments are resistant to neutrophil and macrophage killing, as well as antibiotic treatment [5]. Additionally, the organism traverses distinct stages of development during the infection cycle, three of which are specific to biofilm establishment and growth, and during which the changes in morphology occur [5].

In order to highlight the capabilities of *ProkaryMetrics*, we focus on two UPEC data sets, each containing a single IBC and at least 100 individual bacteria. The first data set is of the wild type UTI89 (a clinical isolate), and the second is a mutant of UTI89 with a defect in the production of capsular polysaccharides, specifically the *kpsF* gene. This capsule mutant is known to produce visually distinct IBCs in size, shape, cohesion, and apparent early onset of morphological change. As with other mutants, viewers could visually comprehend the differences as compared to the wild type, but were limited to vague qualitative descriptions.

In Figure 3 we have used *ProkaryMetrics* to estimate the volume of space occupied by the mass of each IBC, wild type and capsule mutant. By visual inspection, the Δ*kpsF* mutant IBC is clearly much larger and less regular in diameter. Fitting an ellipsoid to the data, as is seen in the Figure, we calculate the wild type fills a volume of $1.95x10^3\mu m^3$ and, within tolerance, fits the category of an oblate spheroid (two of the radii are equal). The Δ*kpsF* mutant fits with the visual inspection and has occupies a much larger volume of space of $3.87x10^4\mu m^3$.

The question of orientation is an interesting one, and certainly important when regarding biofilms. In fact, it is their structure as a community and the spatial heterogeneity of the individuals that contributes greatly to their role as a common cause of persistent infection and their ability to resist treatment [2]. However, obtaining such important information is impossible with destructive techniques such as PCR. As we described in Section 4, the orientation of the main axis vector running along the length of each bacterium is compared to the three orthonormal basis vectors in $\mathbb{R}^3$. In Figure 4, we have used the three orientation calculations to fill the RGB components of the color for each bacterium. *ProkaryMetrics* currently provides three different, user selectable, coloring schemes, and Figure 4 displays the orientation to color mapping: x→blue, y→green, z→red (only bacilli are colored by orientation). In both data sets the bacteria are nearly perfectly aligned with the plane of the image, indicated by the general lack of red in the bacilli. Additionally, with this visualization, it is immediately obvious that the Δ*kpsF* mutant is largely dominated by orientation relative to the y-axis (Figure 4B is rotated 90°).

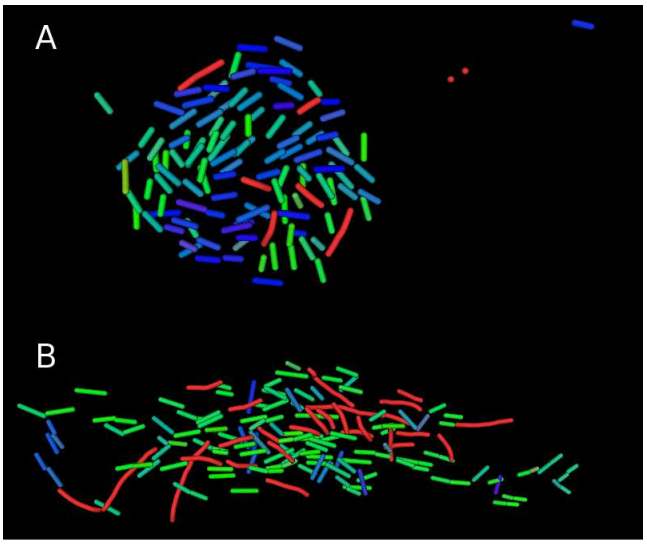

Figure 4: Visual representation of the orientation of the main axis of each bacterium with respect to the three orthonormal basis vectors in $\mathbb{R}^3$. The xyz components of the orientation are represented in the RGB channels of the image: x→blue, y→green, z→red. Filaments and coccoid bacteria are left in their original color. (A) The UTI89 w.t. IBC is clearly split into two fairly distinct populations with one having the main component of its orientation in the y (green) axis, the other with the main component in the x (blue) axis. (B) The IBC formed by the Δ*kpsF* mutant is dominated by bacteria oriented along the y axis, indicated by the predominance of green. The image is rotated such that the y-axis is presented horizontally.

While Figures 3 and 4 provide important semi-quantitative analysis regarding overall biofilm organization and architecture, mathematical and statistical comparisons of architectural characteristics are necessary to establish quantitative descriptors that can be used to prove key differences. Table 1 gives a summary comparison of such data, as well as additional information that is not provided by the previous visual representations. Case-in-point, Figure 5 establishes statistical proof of the obvious qualitative differences visible in the orientation visualization in Figure 4. As we would expect, the bacteria in both samples exhibit almost no orientation change in the plane of the image (z-axis). However, both the x-axis and y-axis orientation data show significant differences, with the wild type strain exhibiting similar orientation distribution in both, while the Δ*kpsF* strain is dominated by orientation difference in the y-axis (as we expected from Figure 4B by the majority of green-hued bacteria).

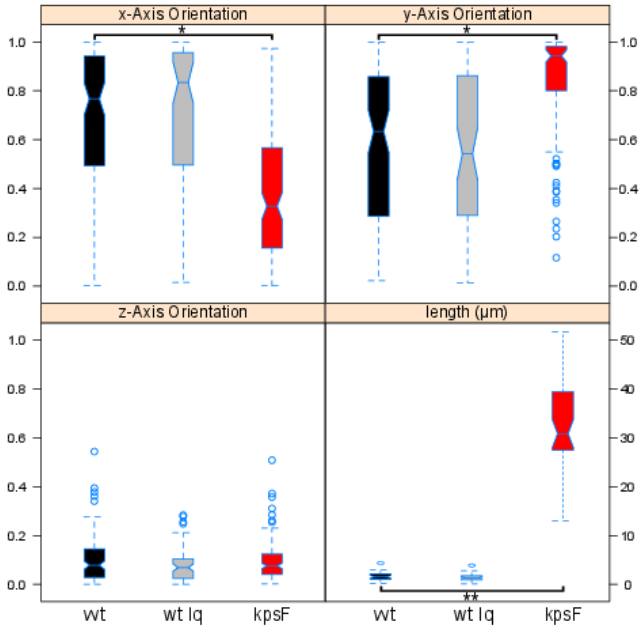

Figure 5: Quantitative representation of the orientations and lengths of the bacteria within the w.t. and Δ*kpsF* UTI89 IBCs. The relative orientations are calculated as the projection of each dimensional component on its respective axis. In both, as noted in Figure 4, the relative orientation of the bacteria with respect to the z-axis is nearly flat. Both the x and y-axis projections, however, show significant differences, as do the overall bacterial lengths between the two samples. (*) 2-tailed Mann-Whitney U test, $p < 0.0001$ (**) unpaired 2-tailed Student's t-test, $p < 0.0001$. Finally, the gray barplots in each graph represent the data gathered from the lower quality data set seen in Figures 1B and 2. Despite the loss of information, we were able to achieve nearly identical quantitative results with *ProkaryMetrics*.

Quantifier	UTI89 *w.t.*	UTI89Δ*kpsF*
Count	120	161
Length (μm)		
IBC Volume (μm^3)	1949	38660
IBC Diameter (μm)	21.6 19.6 3.31	32.0 93.8 12.9
Volume Ratio	0.58	0.32
Orientation (0.0-1.0)		
IB Distance (μm):	15.7	62.5

Table 1: This table summarizes the quantitative descriptors we have developed for architectural comparisons of biofilm infections. The volume ratio is the total volume of bacteria to the IBC volume. The orientation histograms are colored green, blue, pink for the x,y,z orientations respectively.

6 Conclusion

It is becoming clear that in many cases, current methods for clinical investigation of biofilm infections are either insufficient or inaccurate due to the amount of time required or their destructive nature. An integral aspect of the nature of biofilms is their overall architecture as well as the arrangement of the individual pathogens. Indeed, in UPEC distinct architectural and morphological changes occur through the three stages of its intracellular growth cycle [5]. Understanding and calculating these various properties in a quantitative manner can be important for identification of disease state and potential susceptibility of an infectious organism. A non destructive 3D microscopy-based tool is ideal for meeting these needs. Here we present new software, *ProkaryMetrics*, as a tool to fit these requirements, providing 3D visualization and qualitative, as well as quantitative, analyses for user-assisted identification of bacteria from volumetric microscopy data.

In order to demonstrate the utility of the software, we have applied *ProkaryMetrics* to the visualization and analysis of a model organism that is recognized as the causative agent of most urinary tract infections: UPEC. As an infectious organism, it prefers intracellular existence during the majority of its lifecycle. While intracellular, it forms biofilm-like structures (IBCs) that are necessary for pathogenesis [5]. Applying the software to an IBC of wild type UTI89 and the UTI89Δ*kpsF* mutant, we have established significant qualitative and quantitative differences between them in overall architecture and individual characteristics including IBC volume and shape, as well as aggregate and specific orientation and length parameters. While the software was developed with UPEC in mind, it is generalizable and easily modified to handle any sort of organism. Currently, in addition to UPEC analysis, we are generating analyses in collaboration with researchers studying non-typeable *Haemophilus influenzae*. Furthermore, we are investigating computational image analysis techniques for partially automated processing of microscopy data to aid the user in identification [3].

Developing such algorithms, descriptors, and objective analyses is necessary for accurate identification and comparison of clinical biofilm samples in all stages of infection. Such non-destructive imaging analyses will provide rapid and important guidance for clinicians and improve the suite of tools available for disease assessment.

References

[1] J. W. Costerton, J. C. Post, G. D. Ehrlich, F. Z. Hu, R. Kreft, L. Nistico, S. Kathju, P. Stoodley, L. HallStoodley, G. Maale, G. James, N. Sotereanos, and P. DeMeo. New methods for the detection of orthopedic and other biofilm infections. *FEMS Immunology & Medical Microbiology*, 61(2):133–140, Mar. 2011.

[2] J. W. Costerton, P. S. Stewart, and E. P. Greenberg. Bacterial biofilms: A common cause of persistent infections. *Science*, 284(5418):1318–1322, 1999.

[3] S. M. Dabdoub, S. S. Justice, and W. C. Ray. A dynamically masked gaussian can efficiently approximate a distance calculation for image segmentation. *Advances in Experimental Medicine and Biology*, 696:425–432, 2011.

[4] K. England, R. Crew, and R. Slayden. Mycobacterium tuberculosis septum site determining protein, ssd encoded by rv3660c, promotes filamentation and elicits an alternative metabolic and dormancy stress response. *BMC Microbiology*, 11(1):79, 2011.

[5] D. A. Hunstad and S. S. Justice. Intracellular lifestyles and immune evasion strategies of uropathogenic escherichia coli. *Annu Rev Microbiol*, 64:203–21, Oct 2010.

[6] S. S. Justice, D. A. Hunstad, L. Cegelski, and S. J. Hultgren. Morphological plasticity as a bacterial survival strategy. *Nat Rev Micro*, 6(2):162–168, Feb. 2008.

[7] L. G. Khachiyan. Rounding of polytopes in the real number model of computation. *Mathematics of Operations Research*, 21(2):307–320, May 1996.

[8] W. Schroeder, K. Martin, and B. Lorensen. *Visualization Toolkit: An Object-Oriented Approach to 3D Graphics, 4th Edition*. Kitware, 4th edition, Dec. 2006.