Systems-level analysis of microbial community organization: Combinatorial labeling and spectral imaging (CLASI) of microbial diversity

Alex M. Valm\textsuperscript{a,b}, Jessica L. Mark Welch\textsuperscript{c}, Christopher W. Rieken\textsuperscript{a,1}, Yuko Hasegawa\textsuperscript{a,b}, Mitchell L. Sogin\textsuperscript{a,b}, Rudolf Oldenbourg\textsuperscript{a,b}, Floyd E. Dewhirst\textsuperscript{d}, and Gary G. Borisy\textsuperscript{a,2}

\textsuperscript{a}Marine Biological Laboratory, Woods Hole, MA 02543; \textsuperscript{b}Brown University, Providence, RI 02912; and \textsuperscript{1}The Forsyth Institute, Cambridge, MA 02142

Contributed by Gary G. Borisy, January 25, 2011 (sent for review December 22, 2010)

Microbes in nature frequently function as members of complex multitaxon communities, but the structural organization of these communities at the micrometer level is poorly understood because of limitations in labeling and imaging technology. We report here a combinatorial labeling strategy coupled with spectral image acquisition and analysis that greatly expands the number of fluorescent signatures distinguishable in a single image. As an imaging proof of principle, we first demonstrated visualization of Escherichia coli labeled by FISH with 28 different binary combinations of eight fluorophores. As a biological proof of principal, we then applied this Combinatorial Labeling and Spectral Imaging FISH (CLASI-FISH) strategy using genus- and family-specific probes to visualize simultaneously and differentiate 15 different phylotypes in an artificial mixture of laboratory-grown microbes. We then illustrated the utility of our method for the structural analysis of a natural microbial community, namely, human dental plaque, a microbial biofilm. We demonstrate that 15 taxa in the plaque community can be imaged simultaneously and analyzed and that this community was dominated by early colonizers, including species of Streptococcus, Prevotella, Actinomyces, and Veillonella. Proximity analysis was used to determine the frequency of inter- and intrataxon cell-to-cell associations which revealed statistically significant intertaxon pairings. Cells of the genera Prevotella and Actinomyces showed the most interspecies associations, suggesting a central role for these genera in establishing and maintaining biofilm complexity. The results provide an initial systems-level structural analysis of biofilm organization.

The first direct observation of bacteria is credited to Antony van Leeuwenhoek, who in the 17th century produced drawings with remarkable realism of several microbes that he had viewed in a sample of his own dental plaque (1). In hindsight, van Leeuwenhoek’s discovery clearly revealed microbes living not as isolated cells but rather as members of a complex community of morphologically distinct organisms. Modern methods of microbial imaging no longer rely solely on phenotypic characteristics, which provide only limited usefulness for taxonomic classification. An important advance was the introduction of in situ hybridization with oligonucleotide probes targeted to ribosomal RNA for the taxonomic identification and visualization of microbes in situ (2, 3). However, to date, limitations in current technologies have prevented a full systems-level study of microbial community organization.

In principle, FISH probes could be designed with DNA sequence specificity for nearly any microbial phylotype or taxon. In practice, the use of bandpass filters in fluorescence image acquisition and the excitation crosstalk and emission bleed-through of available organic fluorochromes limit the number of fluorophores that can be differentiated simultaneously (4). This limitation restricts routine FISH technology to the identification of only one or a few types of microbes in any single experiment. Thus, although it is possible in many cases to identify with high certainty a few microbial taxa within the context of an entire community, the simultaneous microscopic identification of more than a handful of taxa in a single sample has yet to be achieved.

Microbial communities can be considered as networks of interconnected organisms, often growing in biofilms attached to a substratum and surrounded by a pericellular matrix (5). The complexity of microbial biofilms varies from simple single-species populations to extraordinarily species-rich communities as in the case of soil particle biofilms or those associated with the epithelial surfaces of the human body (6, 7). Environmental metagenomics (8) and microbial coculture experiments (9) have suggested a number of specific molecular interactions between different phylotypes of microbes that make up communities. These interactions include communication via soluble factors (10), syntrophic metabolism in which the enzymatic breakdown of complex bio-polymers is carried out between two or more cooperating organisms (11), and specific structural associations mediated by cell-surface receptors and ligands (12). Molecular methods provide an exhaustive list of potential functions that are present in a community but provide little information on how these metabolic functions are segmented in space and how different organisms might cooperate. Analogous to protein complexes in cellular systems biology, the observation of physical cell–cell interactions between organisms has been suggested as strong indication of functional interaction in the community (13). Therefore, for a comprehensive understanding of microbial ecosystem function, a systems-level analysis of the cellular architecture of microbial communities is crucial.

Human dental plaque is a well-studied and highly diverse microbial community that serves as a model for biofilm biology in general (14). More than 600 species of microbes have been identified as resident in the human oral cavity by traditional molecular methods (15), and recent studies suggest far greater diversity (16). A small subset of oral microbes such as species of the genera Streptococcus and Actinomyces are able to colonize the salivary-coated surface of teeth after tooth brushing (17). Once the tooth surface is initially colonized, an ecological succession is thought to occur as other species including those of the genera Veillonella, Fusobacterium, and others bind to these founder organisms. Coadhesion and coaggregation of genetically distinct oral microbes, in addition to growth of already adherent cells, further contributes to biofilm maturation and ultimately may lead to a highly complex and species-rich microbial community (18). Biochemical mediators of species-specific cell-to-cell adhesion, the cell-surface adhesins and receptors, have been elucidated by means of in vitro biochemical and genetic studies, leading to a...
general hypothesis for the structure of dental plaque (e.g., 19–22). In addition, microscopic imaging studies using FISH or immuno-fluorescence to identify specific microbial taxa have confirmed some of the interspecies spatial relationships in dental plaque biofilms, but only for a few species or taxonomic groups in any single assay (e.g., 23–27). Thus, the system-level taxonomic spatial structure of dental plaque remains to be determined directly.

Recent advances in fluorescence spectral image acquisition and the application of linear unmixing algorithms to spectrally recorded image data allow the unambiguous identification of fluorophores with overlapping spectra, even within the same pixel in a digitally recorded image (28, 29). Here we report the development of a biological labeling and fluorescence image analysis strategy, which we call “Combinatorial Labeling and Spectral Imaging” (CLASI) (Fig. 1), and its application to the simultaneous identification of tens to potentially hundreds of microbial taxa in a single microscopic image. Our strategy entails the labeling by FISH of a given type of microorganism with two or more fluorophores selected from a library of fluorophores chosen to maximize spectral discrimination. With this combined labeling approach we greatly expand the number of different kinds of microbes distinguishable in a single field of view.

Results

Imaging Proof of Principle. We first sought to establish that microbe-sized objects labeled with binary combinations of fluorophores with highly overlapping emission spectra could be distinguished in spectrally acquired microscope images. *Escherichia coli* were used as test objects for CLASI-FISH. Eight versions of a FISH probe with the same oligonucleotide sequence, the Eub338 probe (30), targeting a conserved region of the bacterial 16S ribosomal RNA and confirmed to be present in *E. coli*, were custom synthesized (Materials and Methods). Each version had a different fluorophore covalently conjugated to its 5′ end. These eight probes were used to label separate aliquots of *E. coli* cells in FISH reactions so that 28 populations of microbes were created, each population labeled with a unique binary combination of probes, hereafter referred to as a “label type.” After FISH, the separately labeled *E. coli* populations were combined to create a mixture of the 28 different label types.

Specimens were imaged with a laser scanning confocal microscope equipped with a 32-anode spectral detector (Materials and Methods). For quantitative analysis, to maximize both fluorescence signal within pixels and number of cells in a single field of view, specimens were imaged with a low-magnification, high numerical aperture (NA) objective lens (20x/0.8 NA) (Fig. 2A and B). For ease of visualization, we also present a high-magnification image acquired with a 100x/1.4 NA objective (Fig. 2C and D). Separate spectral image acquisitions were made of each field of view of labeled *E. coli*, with each image acquired using a different laser line available on the microscope system for fluorophore excitation.

Linear unmixing was applied to the spectral images to generate multichannel images in which each channel consisted of measured intensities assigned to one of the eight fluorophores used in the experiment. To facilitate display of the raw data, we show the raw spectral image merge (Fig. 2A and C) in which the pixel-based intensity data from each of the fluorophore channels were pseudocolored in one of eight colors and combined. The data set was converted to a particle basis by segmentation from background in each channel using an intensity threshold and a size discriminator to exclude spurious fluorescence and aggregates of cells from the analysis. For each particle, the mean intensities of all eight fluorophores were tabulated to determine the label type, defined simply as the combination of the two fluorophores with the highest mean intensities in each particle. The taxon-assigned segmented image (Fig. 2B and D) discards the intensity information and assigns 1 of 28 different colors to each label type. For each field of view, the number of each of the 28 different label types was determined (Fig. 2E).

To assess the accuracy of the assay, the same acquisition and image-processing protocols used to characterize the microbial mixture were applied to preparations of the 28 pure populations of labeled *E. coli* (Fig. S1). In each case, more than 98% of cells were correctly identified as to their label type (Fig. S2). To confirm the cell density in each of the 28 *E. coli* tubes, a hemocytometer was used with semiautomatic cell counting. This input is plotted alongside the output as measured in the mixture with the spectral imaging assay (Fig. 2F). We conclude that input into the mixture correlates with output. Together, these results confirm the accuracy of CLASI-FISH and establish as proof of principle that a mixture of microbes labeled with binary combinations of eight different fluorophores can be distinguished correctly in a single spectrally acquired image after linear unmixing.

Biological Proof of Principle. In a natural community we expect to find tens to hundreds, if not more, of different species simultaneously present. In establishing biological proof of principle, we sought to demonstrate that multiple different species of microbes, mixed before labeling, could be labeled and distinguished with taxon-specific probes in a single FISH reaction. We identified previously published oligonucleotide probes or de novo designed genus- or family-level probes that targeted the 16S small subunit (SSU) rRNA of 15 of the most abundant taxa of microbes resident in the human oral cavity and tailored the probes to be

---

Fig. 1. Overview of the CLASI-FISH method. (A) Schematic diagram of the combinatorial labeling approach. To a sample of fixed microbes, fluorophore-conjugated oligonucleotide probes are added so that each kind of microbe is labeled with exactly two different fluorophores. (B) Image processing and analysis workflow. Linear unmixing is applied to raw spectrally acquired images, cells are segmented from background, and then each particle is analyzed for its binary fluorophore composition. The binary label restriction and averaging the unmixing result over all pixels within a segmented object result in less noisy data and a more robust analysis.
specific for their targets within the context of all microbes included in the Human Oral Microbiome Database (31). This curated database consists of the nearly full-length SSU rRNA sequences of 711 microbial taxa previously identified as resident in the human oral cavity and is assumed to include the majority of human oral microbial phyotypes identified by traditional sequencing methods. Specific details of probe design are given in Materials and Methods. Two versions of each probe were synthesized, each conjugated to one of six fluorophores to give 15 unique binary fluorophore combinations.

Laboratory stocks of representative species of these 15 taxa were procured from the American Type Culture Collection (ATCC), were grown as monocultures in media and under atmosphere ideal for each species’ growth, and then were fixed in 2% paraformaldehyde. We prepared a mixture of the 15 stock oral microbial cultures. To this cell mixture, all 30 of our probes (15 different taxon-specific oligonucleotide probes, each in two fluorophore versions) were added in a CLASI-FISH experiment. Spectral images were acquired sequentially with separate wavelength excitation; linear unmixing was applied to the spectral images; then the images were processed to segment cells from background and from each other. Particles were analyzed for their mean intensity over all the pixels within each segmented object in each fluorophore channel, giving each particle six fluorophore intensity values. As previously described for E. coli, the two highest-intensity fluorophores were identified in each particle, revealing the binary fluorophore label of each particle (Fig. 3 A and B). For all oral microbe experiments, an additional level of analysis was applied before taxon identity was declared. In addition to the two highest-intensity fluorophores in each particle, the third highest fluorophore was inspected. If the intensity of this fluorophore was $\geq 60\%$ of the second-highest fluorophore, the particle was declared ambiguous in its binary label composition and was assumed to result from inadequate image segmentation so that two or more microbes of different taxa were identified incorrectly as a single object. Ambiguous cells were declared unknown in their taxon identity, were removed from further analysis, and are colored gray in Fig. 3 A and B. In this experiment, we successfully distinguished 15 different taxon-specific binary spectral markers, thereby identifying all 15 different species simultaneously in one FISH experiment. The fraction of each taxon input into the artificial mixture correlated ($r = 0.91$) with the output result from our CLASI-FISH assay (Fig. 3C).

We conclude that input into the mixture of laboratory-grown microbes correlated with output after image analysis, thereby establishing biological proof of principle of CLASI-FISH.

Application of CLASI-FISH to a Natural Microbial Community. We next tested CLASI-FISH as a tool to characterize and quantify the spatial relationships of 15 different taxa of microbes in semidispersed human dental plaque. Volunteers provided samples of dental plaque removed using dental floss. These samples are assumed to consist mostly of supragingival plaque from between teeth as well as microbes from saliva, subgingival plaque, and plaque from soft tissue surfaces of the mouth. The plaque biofilm was dispersed partially in PBS with vortex mixing to break up large aggregates of cells while preserving smaller aggregates composed of a few to less than a hundred cells. We applied the same mixture of 15 human oral taxon-specific probes to the fixed plaque samples; then the FISH-labeled plaque was settled onto a glass slide and mounted for imaging (Fig. 4). For illustrative purposes, we present data from a single donor taken at a single time point. All 15 taxa were identified in the plaque sample, and the relative abundance of each taxon was calculated (Fig. S3). Relative abundance of each genus or family agreed with previous studies using molecular techniques to quantify the microbial taxa composition of plaque and saliva (12). We next performed a quantitative analysis of the intertaxon spatial relationships of cells within multicell aggregates. We developed an image analysis program that simply calculated the number of cells of each taxon that touched other cells of either the same or a different taxon. To assess the statistical significance of the observed associations, we constructed model images of randomly placed cells and compared the observed touching frequencies of taxa in the random model with the frequencies in natural plaque. This analysis identified 36 statistically significant ($P < 0.05$) associations in dental plaque cell aggregates (Fig. 5 and Tables S3–S6). Two genera, Leptotrichia and Treponema, were shown to
have no significant associations. Cells of the genera *Streptococcus* and *Prevotella* were the most abundant, whereas cells of the genera *Veillonella*, *Prevotella*, and *Actinomyces* showed the most interspecies associations including those with *Streptococcus*, *Rothia*, *Gemella*, *Pasteurellaceae*, *Porphyromonas* and *Neisseriaceae*. In addition, *Actinomyces* and *Prevotella* were found to have significant association frequencies with *Fusobacterium*.

**Discussion**

We have demonstrated a fluorescence imaging assay capable of distinguishing 28 differently labeled microbes in a single field of view, significantly improving the number of fluorescent labels previously distinguished in microbial FISH (32, 33). Our 15-taxon mixture of laboratory-grown oral microbes consisted of cells of varying complexity in their cell wall composition and represented both Gram-positive and Gram-negative organisms. Thus, the labeling strategy reported here is adequate to achieve specific labeling of multiple taxa in a single FISH reaction. Application of CLASI-FISH to human dental plaque, a microbial biofilm, enabled a quantitative analysis of the microscopic spatial relationships of microbes of 15 different taxa in a complex microbial community. We expect this method to provide a rapid, cost-effective characterization of microbial diversity and community composition. More importantly, we have demonstrated that this potential high-throughput imaging assay is well suited to test hypotheses regarding the spatial distribution of microbes in complex multispecies communities.

Proximity analysis of a partially disrupted human oral biofilm revealed 26 statistically significant intertaxon cell-to-cell associations. Cells of two genera, *Actinomyces* and *Prevotella*, each showed nine intertaxon associations, appreciably more than cells of any other taxon. This result suggests an important role for cells of these two genera in maintaining or establishing biofilm complexity. In addition, the relative abundance of all 15 probed taxa, and in particular that of *Prevotella*, is consistent with this sample being a mixture of plaque, predominantly supragingival plaque, and saliva, a result that is consistent with the sample extraction protocol used (Materials and Methods).

Successful spectral imaging requires effective handling of potential sources of error. For this initial test of the CLASI method, we restricted our approach to binary labeling, meaning that all cells of interest were labeled with exactly two different fluorophores or combinations. This restriction provided an a priori constraint that greatly increased the accuracy and robustness of our results. In our experiments, the linear unmixing of fluorophores in cell images typically resulted in high intensity assigned to two fluorophores, whereas a smaller amount of intensity was assigned to the remaining fluorophores. The binary constraint permitted us to reject the residual off-channel intensities as artifacts resulting from limitations of the unmixing algorithm, photon shot noise, and detector dark noise (Fig. S2). Another consideration is the accuracy of image segmentation. Cells of different taxa that overlap within the plane of focus or cells that otherwise touch and are not identified as separate objects with our segmentation procedure result in ambiguous binary label assignment, because more than two fluorophores contribute to the total fluorescence intensity within that segmented feature. In the future, we expect more fluorophores and more fluorophore combinations to be unmixed successfully. This advance will be made possible by more accurate fluorophore assignment using improved unmixing algorithms with more appropriate noise consideration, better treatment of uncertainty in the unmixing solution (34), and better image segmentation. For example, with 15 fluorophores, there exist 105 binary combinations and 455 ternary combinations. We believe this number of identifiable fluorophores or combinations is well within reach.
It has been reported that some species of oral microbes (e.g., species of the genera *Prevotella* and *Porphyromonas*, among others) exhibit endogenous autofluorescence when excited with short-wavelength light (35). Indeed, we observed long-wavelength autofluorescence from some bacteria in dental plaque when illuminated with 405-nm laser excitation but not with longer-wavelength excitation. To avoid the problems of autofluorescence, we eschewed coupling short-wavelength fluorophores with taxon-specific probes and only used the short-wavelength fluorophore, Pacific Blue, with the general bacterial probe Eub338, which labels most bacteria and was not used to discriminate taxa.

In the future, to increase the number of distinguishable targets, the use of short-wavelength excitation probes may be necessary. In this case, to reduce the influence of systematic errors from autofluorescence, one can measure the autofluorescence using an unlabeled sample and include these spectra in the library of reference spectra used for unmixing. Other unknown systematic biases can be detected through the analysis of the goodness of fit of unmixed fluorophores to recorded image data, for identifying particles or pixels with low confidence in fluorophore assignment.

It has been estimated that upwards of six to seven hundred species of microbes make up the human oral microbial community. In our analyses of human dental plaque, the number of nonlabeled cells outnumbered the number of taxon-labeled cells; therefore, we expect that further development of taxon-specific probes will allow a more complete analysis of this community. In addition, we expect the CLASI approach to be compatible with other forms of fluorescence labeling including immunohistochemistry with primary labeled antibodies, live-cell imaging with combinations of fluorescent proteins, and other molecular probes. This technique will expand the number and kinds of targets suitable for CLASI analysis. We expect future application of CLASI-FISH to intact oral biofilms or to model biofilms, as well as spatial analysis in three dimensions, to provide more detailed information on the spatial structure of these complex communities.

Finally, the CLASI approach is not limited to the analysis of the species composition of microbial aggregates. Many subdisciplines of biology could benefit from this method to expand the number of distinguishable fluorescent reporters within a single image. For example, the approach could be useful for the analysis of the spatiotemporal arrangements of complex biological structures such as cellular organelles and macromolecular assemblies within cells, allowing the precise microscopic analysis of many dynamic biological complexes. In general, to the extent that function is implicit in structure, the systems-level analysis of complex biological structures afforded by our combinatorial labeling and spectral imaging strategy offers an approach for deconstructing biological organization.

**Materials and Methods**

**Microbial Strains and Culture Conditions.** *E. coli* K12 (ATCC 10798) were grown in LB broth (Beckton Dickinson) at 37 °C under aerobic conditions and were harvested at an OD<sub>600</sub> of 0.5. A full list of oral microbial strains and culture conditions is provided in SI Materials and Methods.
Oligonucleotide Probe Design. For E. coli experiments, eight versions of the Eub338 probe were custom synthesized. The Eub338 probe targets a conserved region of the bacterial 16S rRNA which is predicted to be present in most bacteria (http://www.microbial-ecology.net/probebase) and was verified to be present in E. coli. Each version of the probe had a different fluorophore conjugated to its 5’ end.

Oligonucleotide probes used to label 15 of the most abundant genera or families of oral microbes were designed to target the 16S rRNA and are listed in Table 5. The probes used were identified as specific genus-level probes for oral organisms. Primers longer than 42 nucleotides were designed de novo using the probe design function of the ARB program (www.arb-home.de) with a database of 16S sequences from the Human Oral Microbiome Database (www.homed.org). Two versions of each of the 15 taxon-specific probes were synthesized. Both versions of each oligonucleotide probe were conjugated to one of two different fluorophores. Additional details on probe design are given in SI Materials and Methods.

Dental Plaque Sample Collection. The use of human subjects in this study was approved by the New England Institutional Review Board. Plaque was provided by donors who had given informed consent. Donors refrained from normal oral hygiene (tooth brushing, flossing, and mouthwash) for 24 h before plaque samples were taken. Donors used acetate dental floss (Proctor & Gamble) to corn-harvest plaque from above and below the gum line, mostly between teeth from all parts of the mouth. Additional details on plaque sample collection are given in SI Materials and Methods.

FISH. FISH was performed in Eppendorf tubes using a protocol modified from Pernthaler et al. (36). Details of the modified FISH protocols used in this study are given in SI Materials and Methods.

Image Acquisition. Spectral images were acquired with a laser scanning confocal microscope equipped with a 32-channel multianode photomultiplier tube. Multiple acquisitions were made of each field of view in series, using laser excitation appropriate for the fluorophores used in any experiment. Details on image acquisition and instrumentation are given in SI Materials and Methods.

Image Analysis. Spectrally acquired images were subjected to linear unmixing using the microscope manufacturer’s image acquisition and processing software. Unmixed images were imported into Image J (37) for image segmentation with a global-intensity threshold. Foreground objects were segmented further with a watershed operation. Computed fluorophore intensities were averaged over all the pixels within a segmented object, and the two highest-intensity fluorophores were identified to assign a binary fluorophore label to each cell. Specific details on image analysis are given in SI Materials and Methods.

Spatial Analysis. A computer program to analyze the spatial distribution of microbes in semidispersed dental plaque was developed on the Mathematica platform. The number of cells of each labeled type that touch each other or touch cells of a different label type was estimated as the difference between the number of objects in segmented images before and after two rounds of image dilution. Details on spatial analysis are given in SI Materials and Methods.

Image Modeling. Model images of randomly distributed dental plaque microbes were created in Mathematica with parameters of cell size, cell shape, and cell abundances derived from measurements made on images of plaque microbes. Details on model image construction are given in SI Materials and Methods.

Spatial Plot. A network diagram that summarizes all intertaxon cell-to-cell associations observed in dental plaque was constructed in Mathematica. Details on construction of this plot are given in SI Materials and Methods.

ACKNOWLEDGMENTS. We are grateful to Louis M. Kerr for invaluable assistance with microscopy. We thank Nikon Instruments and Carl Zeiss for gracious instrumentation support. This work was supported by Grant 2007-3-13 from the Alfred P. Sloan Foundation and National Institutes of Health Grant 1R1C-DE020630 from the National Institute of Dental and Craniofacial Research (NIDCR) (to G.G.B.) and by National Institutes of Health Fellowship F31-DE019576 from NIDCR (to A.M.V.).