Periodontology 2000, Vol. 55, 2011, 70–86
Printed in Singapore. All rights reserved

High-throughput methods for analysis of the human oral microbiome

Alex E. Pozhitkov, Thomas Beikler, Thomas Flemmig & Peter A. Noble

The vast majority (ca. 90%) of cells in the human body are not human at all – but rather microbial (64). Many of these microbial cells play important roles in normal human physiology because they are involved in nutrient absorption (5, 6, 10), synthesis of vitamins and protection of human cells from infection (39, 134). The human body contains numerous ecological niches (such as the oral cavity) that are suitable for acquiring and retaining microbial cells. The oral cavity is a warm, wet and nutrient-rich environment that is ideal for supporting microbial growth. It also shelters microbes by providing specific ecological niches (e.g. the tongue, and supra- and subgingival areas of the tooth), which protect microbes from mechanical and environmental stress caused by brushing, chewing, pH and temperature (58). These benefits are not without consequences: some microbes are a burden on the host because they continually challenge oral tissues for resources. The host deals with these microbes by regulating microbial growth through enzymes and by influencing the composition of oral microbes through defence mechanisms (e.g. immune effectors and defensins). Hence, the host and the microbial cells co-exist in the oral cavity by maintaining a dynamic balance that arises out of numerous host–microbe interactions (79). In healthy individuals, this balance is regarded as 'dynamic' because it changes in response to endogenous (e.g. stress) and exogenous (e.g. smoking) factors. A substantial disruption of this balance, however, can lead to a variety of chronic oral infectious diseases, including periodontitis, peri-implant diseases and dental decay (47).

Thousands of diverse species of microorganisms colonize the human oral cavity. Most of these microorganisms are found in complex biofilms that are attached to soft and hard tissues in the oral cavity (57, 89). These biofilms are assembled in an ordered process that begins with the adherence of microbial colonizers to a surface and is followed by the colonization of different microbial species that attach to the initial colonizers (65, 100, 101, 119). The initial colonizers are predominantly gram-positive aerobes, such as streptococci (93). These microbes normally exist in commensal harmony with host cells, while secondary colonizers, which are predominately gram-negative bacteria, may or may not exist in harmony with host cells. Host cells provide specific receptor ligands that coordinate the initial microbial colonization of a surface, while subsequent biofilm assembly depends upon the physiology and metabolic capabilities of the interacting microbial cells (52). Hence, the physiological processes of both the host and microbial cells are responsible for biofilm assembly (61). Differences in the composition of oral microbial biofilms within the same individual on different tissue surfaces, such as mucosa vs. tooth, or within the same habitat (e.g. one vs. another periodontal pocket) (7–9, 55, 77, 89, 97) are presumably caused by variations in endogenous and exogenous factors.

The perception of oral microbial communities, biofilm formation and host–microbe interactions by the scientific community is inherently linked to the availability and capability of technologies that can be used to describe biological events. These technologies have enabled researchers to study known (previously described) microbes, and have helped to elucidate the role of microbes in oral disease. However, most conventional technologies are based on detecting only one nucleic acid sequence (e.g. by

using the quantitative polymerase chain reaction [PCR]) of a known gene/microbe rather than multiple nucleic acid sequences of several genes/microbes (both known and unknown). These technologies are apparently not high-throughput. As a consequence, much information exists about the physiology and interactions of specific microbes, for example *Porphyromonas gingivalis* and *Streptococcus mutans* (62, 93). Yet, little information exists about entire microbial communities and their emergent properties (e.g. population and metabolic dynamics in response to environmental conditions). The scientific community is keen to increase the availability of high-throughput technologies because such methods will enable researchers to study the physiology and interactions of microbial communities and human hosts in a cost-effective manner.

This review is organized into three distinct sections. The first section reflects on the past and current molecular technologies that are used to routinely identify and quantify microbes in clinical practice and research. The second section critically evaluates these technologies, highlighting the major problems that prevent the correct interpretation of the data. The last section deals with emerging technologies that will probably substantially improve our understanding of oral cavity biofilms and disease because they are based on holistic approaches.

Identification and quantification of microbial species: past and present

Culture-dependent approaches

Culture-dependent approaches to identify microbes involve growing the microorganisms on defined media, followed by identification based on phenotypic and biochemical criteria, differential staining methods, metabolic end-product analysis and cell-membrane composition (98). Culture-dependent approaches are extremely useful for determining the antibiotic susceptibility of oral microbes and for assessing the pathogenicity of individual species (37, 130). It is well established that not all commonly used antibiotics are effective against periodontal pathogens, and patients respond differently to the same antibiotics (8, 43). Hence, the best clinical outcome achieved is by first determining the baseline response of microbial communities (before treating the patient) (42) and then administering the most effective course of antibiotic treatment based on antibiotic susceptibility testing (138). It should be noted that both of these tasks (identifying microbes and determining antibiotic susceptibility) require experienced laboratories and are relatively costly and time-consuming.

Currently, the culture-dependent approach may involve extracting nucleic acid from a single colony, cloning the sequence into a plasmid vector, sequencing the ribosomal RNA genes and identifying the sequence using a ribosomal RNA database. The extracted ribosomal RNA sequences can be also used to identify microorganisms through terminal restriction fragment length polymorphisms, denaturing gradient gel electrophoresis, hybridization to micro- and macro-arrays (e.g. checkerboard hybridizations) and quantification by real-time polymerase chain reaction (PCR). The main problem with conventional culture and culture-based analytical technologies, however, is that less than half of bacterial species in biological samples seem to be culturable (117), thus rendering these approaches unsuitable for holistic studies.

Culture-independent approaches

Counting clones and sequencing

Phylogenetic approaches using 16S ribosomal RNA gene clone libraries have been applied to investigate the diversity of culturable and nonculturable species in the human oral cavity (1, 50, 60, 97, 124). The method is labor- and time-intensive, nonquantitative and there is a significant cloning bias inherent with the method (35).

Checkerboard hybridization

The checkerboard hybridization method involves extracting DNA from oral samples and hybridizing the sample against labeled probes representing whole genomes or 16S ribosomal RNA genes of known microbes. The reason it is called 'checkerboard' is that the genome or ribosomal RNA probes are hybridized at right-angles to the DNA of multiple oral samples, and processed images of the hybridizations look like a checkerboard. The checkerboard DNA–DNA hybridization technique has been widely used to comprehensively examine the types and numbers of bacteria in supragingival and subgingival plaque and saliva in healthy subjects and in patients with periodontitis (25, 146, 147). This approach has also been used to study the influence of cigarette smoking on the composition of the subgingival microbiota (13, 42), the differences between the subgingival microbiota in subjects from different geographic locations (41), the relationship of ethnic/racial

groups, occupational and periodontal disease status (24) and the effects of different periodontal therapies (32, 125). Although this technique is rapid, sensitive and relatively inexpensive, nonspecific target binding is still a major shortcoming (131).

Fingerprinting of amplified PCR products

Terminal restriction fragment length polymorphism is a molecular approach that allows the assessment of complex bacterial communities (22, 56, 69, 123) by profiling based on the position of a restriction site close to a labeled end of an amplified gene. The method is based on digesting a mixture of PCR-amplified variants of a single gene using one or more restriction enzymes and detecting the size of each of the individual resulting terminal fragments using a DNA sequencer. However, taxa with different sequences can have the same size of terminal restriction fragments, making analysis of parallel clone libraries necessary for identification (149). Related techniques include denaturing gradient gel electrophoresis (36, 152). However, it is difficult to create a database from the band profiles obtained by denaturing gradient gel electrophoresis and compare them to fragment profiles obtained by terminal restriction fragment length polymorphism analysis. Moreover, DNA with different sequences can migrate to the same point in a denaturing gradient gel electrophoresis gel, making sequencing of individual bands necessary for identification (149).

Quantitative PCR

Quantitative reverse-transcription (RT) PCR allows the detection and quantification of genes / bacteria in microbiological samples. Quantitative RT-PCR using the TaqMan system was first used to quantify *Tannerella forsythia* in subgingival plaque samples (127) and to measure the density of *P. gingivalis* and the total number of bacterial cells in plaque samples (74). In addition, quantitative RT-PCR using SYBR Green dye and the LightCycler system (Roche Diagnostics, Mannheim, Germany) was first used to detect and quantify periodontal pathogens. Currently, detection and quantification of bacteria by quantitative RT-PCR is generally accepted and many studies have reported the usefulness of this system for the identification of oral bacteria (3, 14–16, 54, 76, 91, 94, 95, 133, 148).

The quantitative RT-PCR technique is a highly selective method that may be considered as a 'Gold standard' when quantitative analysis of specific bacteria is needed. However, this method is only able to detect and to quantify known species, and thus is inadequate for the analysis of complex microbial communities containing unidentified species.

Pyrosequencing

The core technology of pyrosequencing relies upon a sequence of enzyme-triggered reactions, which ultimately results in the production of a luminescence signal. The enzyme cascade is activated by the generation of a pyrophosphate, which is released after successful incorporation of a nucleotide into the DNA sequence (Fig. 1). The pyrophosphate is converted into ATP by an ATP sulfurylase and subsequently used by the luciferase to convert luciferin to oxyluciferin, yielding a luminescence signal. To perform pyrosequencing, a mix of a nucleic acid template and primers is prepared and deoxyribonucleotide triphosphates are added in a known order (e.g. A, T, G, C, A…) to a reaction vessel. Because the order in which the deoxyribonucleotide triphosphates are added is known and a charge-coupled device camera records the intensity of the luminescence bursts, the template sequence and its quantity can be reconstructed (105).

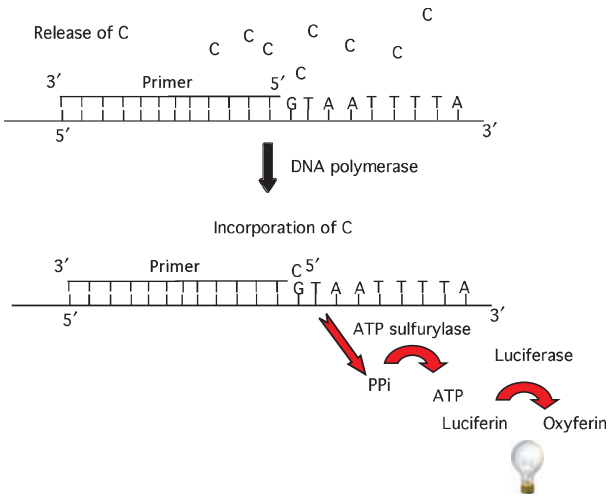

Fig. 1. Cartoon showing pyrosequencing reactions. The oligonucleotide primer hybridizes to the template. There is an ordered release of deoxyribonucleotide triphosphates and the addition of polymerase. If a deoxyribonucleotide triphosphate is incorporated, an enzyme cascade is triggered. The pyrophosphate (PPi) released from the incorporated nucleotide (in this case, dCTP) is converted to ATP by sulfurylase, and the ATP is used by the luciferase to oxidize luciferin to oxyferin, which results in the liberation of luminescence. In the next round, all the excess dCTP is washed away and another deoxyribonucleotide triphosphate is released. Because the order of the deoxyribonucleotide triphosphates is known, the bursts of light can be used to decode the order of nucleotides in a sequence.

Pyrosequencing has been used to identify mycobacteria (135), clinically relevant yeasts (88) and pathogenic *Candida* species (17).

In summary, the techniques described above (see Fig. 2) have helped to improve the understanding of the oral microbiota by analyzing single pathogens / commensals or clusters / profiles of pathogenic / commensal microorganisms and their association with health and oral disease. In addition, some of the techniques are crucial in daily routine to guide clinical decision-making. However, these techniques are of very limited use when trying to analyze the complete oral microbiome and transcriptome. To successfully master this task, high-throughput technologies are needed.

High-throughput approaches

Microarrays

Oligonucleotide microarrays have been widely used to identify microorganisms and determine the expression of genes. Microarrays contain oligonucleotide probes that target RNA and DNA genes in biological samples. All microarray platforms share the common attribute that a sensor detects a signal from target sequences which hybridize to immobilized oligonucleotide probes. The intensity of this signal provides a measure of the amount of bound nucleic acid in a sample. Microarray platforms vary in the number of oligonucleotide probes they contain and in the way the probes are attached to the microarray. High-density microarrays (e.g. Affymetrix, NimbleGen, Agilent) contain thousands to millions of probes that are directly synthesized on the microarray surface. Low-density microarrays (e.g. custom-design) contain hundreds to thousands of probes that are spotted on the microarray surface.

Both high- and low-density DNA microarrays have been used to identify microbes in the oral cavity (23, 51, 115). Huyghe et al. (51) was able to classify PCR-amplified ribosomal RNA genes to various nodes on a phylogenetic tree using high-density microarrays. Each node on the tree was represented by a hierarchically nested probe. Therefore, a target hybridized to a microarray could be classified to a specific taxonomic level (e.g. phyla, family, genera and species). A potential advantage of this approach is that probe hybridization signals of an 'uncharacterized' ribosomal RNA gene could be fit onto the tree. Although Huyghe et al. (51) showed that the hierarchically nested probes approach could detect a single bacterial species spiked into a complex oral sample, the utility of the approach to detect microorganisms in a complex sample has yet to be fully demonstrated or properly evaluated.

Low-density microarrays have been designed by Presza et al. (115) and Columbo et al. (23) to identify ribosomal RNA genes of more than 300 oral microbes. Presza et al. (115) compared their microarray results with those obtained in a previously published sequencing study (114). Although they found that most of the genes identified in the microarray study were also identified in a sequencing study, several ribosomal RNA genes were not found in the microarray results (see the Supplementary Table S1 in Ref. 115). The authors concluded that these results could be caused by stochastic effects that arise from low copy number. However, it could be argued that microarray problems caused the discrepancy. The challenge of interpreting the results

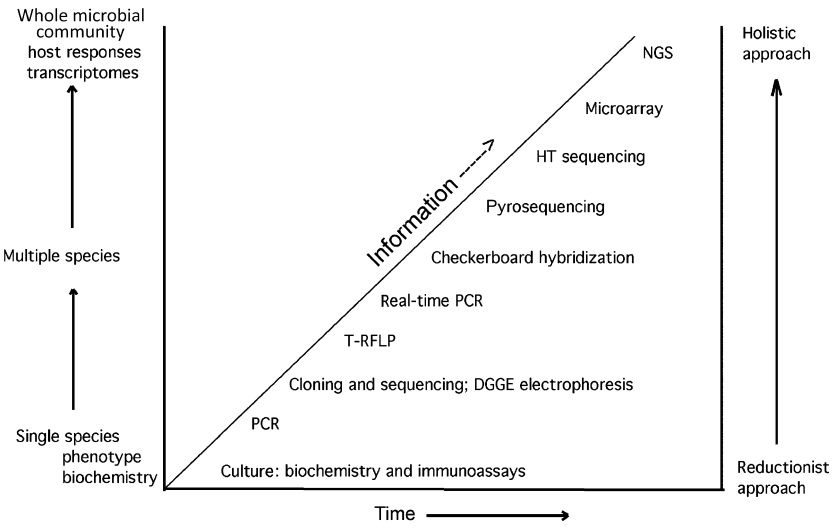

Fig. 2. Chronology of analytic method availability and the capacity of these methods to characterize microbes and microbial communities. DGGE, denaturing gradient gel electrophoresis; HT, high throughput; NGS, next-generation sequencing; PCR, polymerase chain reaction; T-RFLP, terminal restriction fragment length polymorphism.

presented by Columbo et al. (23) is that extensive data manipulation was carried out.

High-density microarrays have also been used to investigate host gene expression of human cells infected with *P. gingivalis* (83). Meka et al. (83) revealed that *P. gingivalis* down-regulates cytokines involved in bone resorption in soft tissue but had no effect on cytokines in hard tissue. Interestingly, the authors report no correspondence between the expression levels obtained using the microarray and those obtained using quantitative PCR. Therefore, it is difficult to interpret the biological significance of their study. One could argue that the observed differences in cytokine regulation in soft and hard tissues could be caused by the methodological artifacts.

In summary, microarrays have been used to examine the gene-expression profiles of host cell–microbial interactions in the oral cavity and to identify oral microbial communities. A common problem of these studies is that the microarrays used have not been properly validated and, as a consequence, the interpretation of the results may not accurately reflect the biology. The next section will discuss the physicochemical aspects of microarrays and provide a perspective on the microbiological interpretation of microarray results and possible solutions.

Evaluation of high-throughput microarrays

Oligonucleotide microarrays offer a significant potential for the accurate quantification and identification of multiple gene targets in a biological sample because they typically contain hundreds of thousands of different immobilized probes, with each probe acting as an individual sensor with its own specificity and sensitivity to different gene targets (12). Typically, a set of unique probes is used to identify specific gene sequences (i.e. targets) (71). Here, the assumption is that the intensity of a probe signal provides a measure of the amount of target in a sample. However, interpreting the signal in terms of bound targets is difficult because of the interactions between multiple targets at different concentrations (113), which is further complicated by electrostatic factors arising from the microarray surface (107). The observed signal intensity on a microarray spot (i.e. the probe) is a composite of all specific and nonspecific targets bound to that probe. Therefore, it is extremely difficult to determine how much of the probe signal is represented by the specific (i.e. 'true') target and how much is represented by the nonspecific target in a heterogeneous target mixture. Three main strategies have been devised to determine the amount of specific target bound to a probe.

The first strategy is to subtract nonspecific signal from the signal in question. This is achieved by subtracting the signal intensity of a probe containing a central mismatch from that of a perfect match probe. The idea is that the mismatch probes report nonspecific binding because they do not bind to the same target as the perfect match probe. This idea evolved from early molecular studies, which showed that some probes containing a mismatch had lower binding affinities than complementary (perfect match) probes (68). A mismatch probe and its corresponding perfect match probe are known as a 'probe-pair'. Many microarray studies have subtracted the intensity of mismatch probes from the intensity of perfect match probes before downstream analyses because it was thought that mismatch intensities account for nonspecific signal (e.g. 19, 20, 26, 143). Yet, several studies have shown that mismatch probes are not adequate measures of nonspecific target binding because the sources of cross-hybridization signals are different for perfect match and mismatch probes (113, 144). In addition, it has been shown that mismatch probe signal intensity is correlated to perfect match probe signal intensity (106) (Fig. 3); therefore, mismatch probes partially contain a specific perfect match signal, which suggests that mismatch probes are probably binding the same target as perfect match probes, as well as other

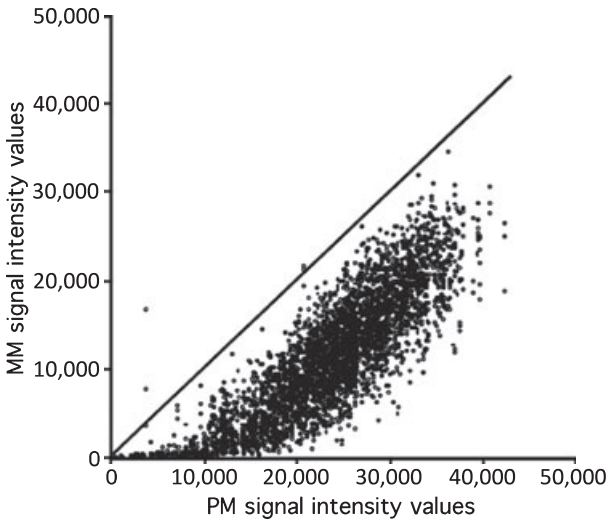

Fig. 3. Distribution of average signal intensity values for perfect match (PM) and designed mismatched (MM) duplexes using target ribosomal RNA from *Porphyromonas gingivalis* hybridized to one microarray. The line represents SI(PM) = SI(MM). SI, signal intensity. Adapted from Pozhitkov et al. (106).

targets. As a consequence, mismatch probes are now rarely used in microarray research.

The second strategy is to design oligonucleotide probes that have high specificity and sensitivity to the nucleic acid targets of interest (11, 46, 67, 68, 81, 82, 104, 121). The assumption here is that a properly designed probe should hybridize only to perfectly complementary targets and not to targets having high sequence similarity to the probe. The principles of designing probes for oligonucleotide arrays have been published in numerous articles and they have been widely implemented in microbiology studies. These studies fall into two categories, namely (i) 'common sense' probe design (72, 85, 99, 136, 137) and (ii) thermodynamic probe design (11, 46, 67, 68, 81, 82, 104, 121) – which are discussed below.

The 'common sense' approach of probe design is based on *in silico* predictions (e.g. GC content, probe length, secondary structure and mismatches to nontarget sequences), followed by experiments carried out to optimize the specificity and sensitivity of the probes by adjusting the conditions of hybridization and stringent washing. For the *in silico* predictions, each probe on the microarray is individually evaluated using software tools to determine its uniqueness in terms of number, types and positions of mismatches to nontarget ribosomal RNA sequences. Next, the thermodynamic properties (described in more detail in the next paragraph) of the probes are evaluated using *mfold* (153). A recent study using 7,693 probes and 13 different targets found no relationship between *in silico* parameters and probe specificity and sensitivity, indicating the limitations of the common sense approach (104). Currently, no algorithm exists that can satisfactorily predict probe behavior on the surface of a microarray.

The thermodynamic probe-design approach determines the specificity and sensitivity of immobilized probes based on their thermodynamic properties. As proposed by Matveeva et al. (81), thermodynamic properties include all four free energy-binding terms (i.e. ΔG^0_b, ΔG^0_p, ΔG^0_t and ΔG^0_d) shown in Fig. 4, as well as the effects of secondary structure of the probe and target (153). At the same time, Naef & Magnasco (92) and Mei et al. (82) described an *ad hoc* model that examined the affinity of a probe to a target based on the sum of position-dependent base-specific contributions. Zhang et al. (151) described another *ad hoc* model that considered position-dependent nearest-neighbor effects. Held et al. (48) examined, using methods similar to those described earlier in this paragraph, the effects of free energies of RNA/DNA duplex formation. Wu & Irizarry (145) developed a model that considered both stochastic and deterministic aspects of probe–target hybridizations. Li et al. (66) developed a hybridization model that explains probe signal intensities through known nearest-neighbor parameters, but without the excessive parameter fitting used by the previous authors. Unfortunately, none of the studies satisfactorily predict signal-intensity values on oligonucleotide microarrays because there are significant disagreements between actual and predicted values (107). Therefore, the basic assumption that probes can be designed to be specific to one target and not to other targets is flawed.

A third strategy is to physically remove nonspecific targets bound to probes by performing a stringent wash (i.e. rinse the array with low-salt buffer at a constant temperature [isothermal wash]). It is widely believed that once the isothermal wash has been performed, nonspecific duplexes are washed away and the observed signal is the 'true' signal of the specific (perfect match) duplexes. However, stringent washing actually distorts the 'true' signal of specific probe–target duplexes. This distortion was revealed when the signal of duplexes at equilibrium was recorded (using isopropanol) and compared with the

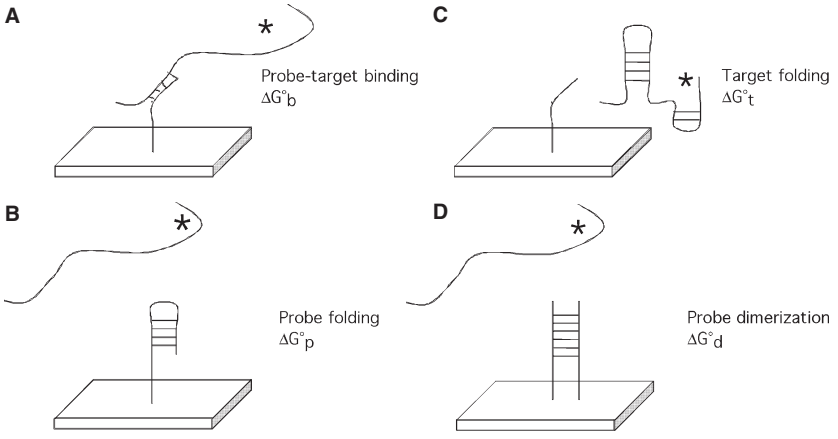

Fig. 4. Presumptive hybridization processes and ΔG° terms that influence signal intensities. Each panel depicts a labeled (*) target and an immobilized oligonucleotide probe on an array. Adapted from Pozhitkov et al. (104).

signal after a stringent wash (nonequilibrium). Apparently, in experiments that used pure targets only (to prevent cross-hybridization effects), the degree of distortion varied by hybridization time, probe concentration and probe sequence. Extended hybridization times resulted in fewer targets being removed by the stringent wash (107, 129). In general, the higher the concentration of the microarray probe, the lower the amount of bound target removed by stringent washing. Although some probe sequences were affected more by stringent washing than others, no pattern could be found, which is consistent with previous studies (111, 112). Attempts to model the effects of stringent washing on probe–target duplexes have not been successful because stringent washing affects probes in many different ways (e.g. hybridization time, probe sequence, probe concentration), which are still poorly understood.

In summary, none of the above strategies are sufficient for minimizing or eliminating the effects of nonspecific target binding to immobilized probes. As a consequence, determining the portion of the total signal represented by the binding of specific targets vs. the portion represented by the binding of nonspecific targets cannot be resolved. These findings have a substantial effect on the interpretation of microarray results because potentially all signals from a microarray could be the result of nonspecific targets hybridizing to 'sticky' probes. Hence, a microarray researcher cannot confidently assess if a gene was expressed without verifying the results using an alternative approach, such as quantitative PCR. It is important to emphasize that the problem of nonspecific target binding to probes does not improve for poorly expressed genes (or genes at low concentrations) because more nonspecific targets are available to bind to probes than specific targets. Hence, noise in the probe signal is inversely correlated with the concentration of specific targets. Poor correspondence between microarray and quantitative PCR results, as reported by Meka et al. (83), are consistent with the results obtained in other biological disciplines (40, 90, 141). This lack of correspondence is a reason to question the validity of some microarray studies. Lastly, many microarray studies involve a tremendous amount of data manipulation (e.g. *ad hoc* criteria) that is not based on physicochemical principles, as shown previously (103, 104, 106, 109–113, 122).

Alternative microarray applications

As discussed, nonspecific target binding appears to be an inherent problem of all microarray experiments. To circumvent this, Pozhitkov et al. (106, 108) devised an analytic approach that uses signal intensities of all oligonucleotide probes on a microarray as a source of information. The approach permits direct analysis of hybridization patterns of targets in mixed target samples and avoids the problems associated with nonspecific target binding that affect all the microarray approaches mentioned above. The basic principle of the approach is that a fluorescence pattern of a mixed sample is a superposition of individual fluorescence patterns (i.e. fingerprints) of each target (Fig. 5). The superposition can be deconvoluted in terms of individual target mass or concentration, and the quality of the quantification can be assessed by linear regression analysis.

Rather than use probe-design strategies, this approach involves (i) making a tiling microarray from the sequences of targets under investigation (25-mers with a 1 base-pair offset), removing redundant probes and then synthesizing the probes *in situ* on the microarray, (ii) generating a fingerprint library by experimentally recording the fingerprints of individual pure targets, (iii) filtering out probes that have the same signal intensity for all targets and (iv) quantifying the sample. Two studies (106, 108) have demonstrated the utility of the approach (Fig. 6). The disadvantage of the approach is that it requires additional microarrays to record the fingerprints of each target in a mixture.

In contrast to other microarray studies (e.g. 26, 96, 143), the analytic approach does not attempt to minimize, reduce, or ignore nonspecific target

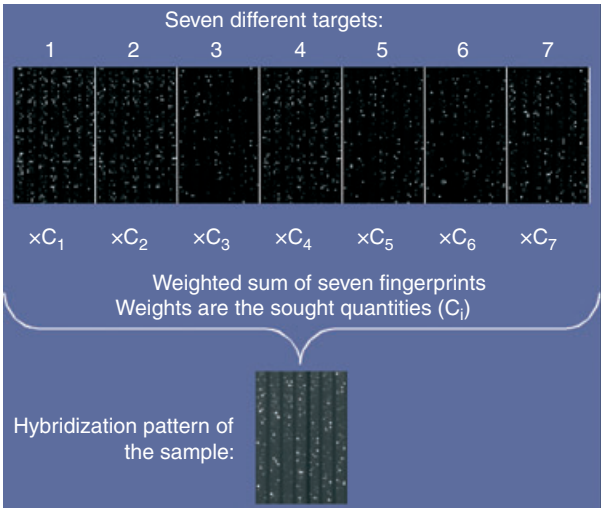

Fig. 5. Schematic representation of the fingerprint approach. Seven different ribosomal RNA target fingerprints and the hybridization pattern of a mixture is shown. The hybridization pattern of a sample is considered as a weighted sum of the patterns of pure targets (see the text for details). C_1–C_7, concentrations determined of targets 1–7.

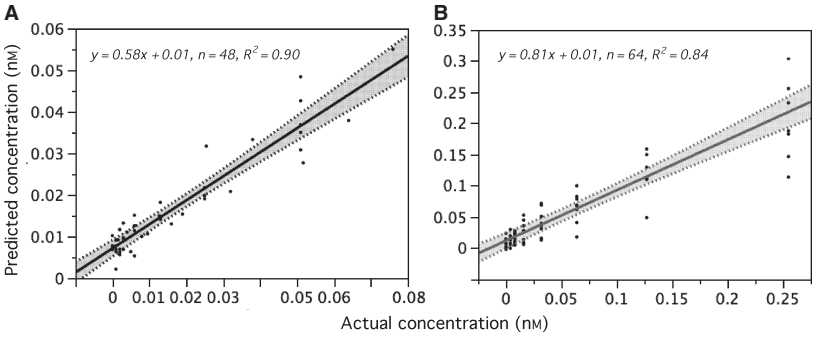

Fig. 6. Relationship between actual and predicted concentrations of the two independent Latin-square experiments using the fingerprint approach. (A) Results from the first Latin-square experiment (six targets × eight mixtures). (B) Results from the second Latin-square experiment (eight targets ×eight mixtures). The grey lines represent the 99% confidence limits of the regression line. Adapted from Pozhitkov et al. (108).

binding; rather, all probes are used as a source of information for quantifying ribosomal RNA targets. The three major advantages of the fingerprint approach over the other array-based approaches (all of which involve PCR amplification) are: (i) the fingerprint approach does not rely on the design of specific probes; (ii) *ad hoc* criteria and normalization for interpreting microarray results is not required; and (iii) the quality of the quantification can be determined from the R^2 of the linear regression.

In summary, the fingerprinting approach offers significant potential to accurately quantify microbes in complex mixtures with statistical confidence (106, 108).

Next-generation technologies and clinical implications

After the extensive review on microarray-based methodologies, we now discuss next-generation sequencing. Since 2006, there has been an explosion of studies that have used next-generation sequencing devices to sequence whole genomes of different organisms, transcriptomes of single organisms and even communities of organisms (Fig. 7, right panel). The projected number of peer-reviewed published papers dealing with next-generation sequencing is expected to reach more than 800 articles in 2010 and surpass 2,500 articles in 2011 (Fig. 7, left panel).

There are several advantages of characterizing genes by sequencing, rather than by using microarray hybridization. Specifically, sequencing does not require knowledge of the targets to be hybridized, gene sequencing provides a 'digital' rather than an 'analog' signal (i.e. a gene sequence is either present or absent from a sample), and once a gene has been sequenced, putative identification of the gene can be rendered by comparing the sequence to reference genes in DNA databases.

For the purpose of this review, there are three distinctly different generations of 'next-generation

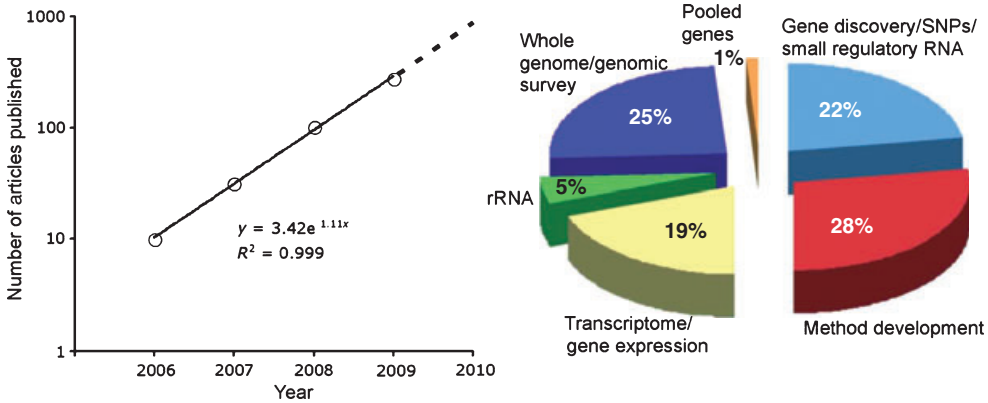

Fig. 7. Cumulative number of next-generation sequencing articles published in the Web of Science (left panel) and a breakdown of their primary content (n = 265 articles) (right panel). The analysis was based on the following search words: [454 pyrosequencing] OR [454 sequencing] OR [Illumina Solexa sequencing]. The regression (left panel) predicts 875 articles in 2010, 2,655 in 2011 and 8,100 in 2012. rRNA, ribosomal RNA; SNPs, single nucleotide polymorphisms.

sequencing' systems, which are all capable of high-throughput sequencing. The first-generation systems rely on DNA amplification and therefore do not permit single molecule detection. Second- and third-generation next-generation sequencing systems can detect single molecules. What distinguishes the third-generation systems from the second-generation systems is that the third-generation systems are based on nanopore technology, they do not require polymerase and they are not yet commercially available. The next section provides an overview of leading next-generation sequencing systems by generation, and discusses their strengths and weaknesses.

First-generation next-generation sequencing

Two consistent themes of first-generation systems are: the ligation of oligonucleotide adaptors to DNA fragments and the immobilization of the fragments to a solid surface, such as a bead. The purpose of the adaptors is twofold: to anchor the fragments to a solid surface; and to serve as primers for amplification and/or sequencing. The following three next-generation sequencing devices are discussed: Roche/454 (Roche 454 Life Sciences, Branford, CT), Illumina/Solexa (Illumina, San Diego, CA) and Applied Biosystems/SOLiD (Life Technologies, Carlsbad, CA).

The first instrument of this kind is the Roche/454 (http://www.454.com/). In addition to the above-mentioned common theme, it is based on the same core pyrosequencing technology mentioned in the first section. A simplified version of the protocol is the following: (i) clonal amplification of templates on beads; (ii) deposition of the beads onto picotiterplate wells; (iii) controlled delivery of deoxyribonucleotide triphosphates by laminar fluidics, and (iv) a high-resolution charge-coupled device camera that detects the luminescent burst upon deoxyribonucleotide triphosphate incorporation. In detail, clonal amplification of a template is accomplished by mechanically shearing the DNA and ligating the end of each fragment to different adaptor sequences and amplifying the bead in a thermostable water-oil emulsion droplet by emulsion PCR (78). Each droplet acts as a micro-reactor, whereby only one bead is enclosed and only one type of PCR fragment is produced, resulting in the production of millions of immobilized, clonally amplified targets. Next, the emulsion is broken, the DNA strands are denatured and beads containing the single stranded (ss) DNA are deposited in the picotiterplate wells. Smaller beads containing immobilized enzymes (sulfurylase and luciferase) are added to each well and the core pyrosequencing reactions (previously discussed) are commenced. Different deoxyribonucleotide triphosphates are sequentially diffused into each well via microfluidic laminar flow. A luminescence signal indicates the successful incorporation of a deoxyribonucleotide triphosphate.

The strengths of 454 sequencing lie in its long read lengths (400–500 nucleotides) and the amount of sequence generated (0.5 Gb) (75, 140). The long reads can handle repetitive regions better than other next-generation sequencing systems. A major weakness of the 454 sequencing system is that sometimes more than one nucleotide is incorporated in the DNA template during a cycle, making it difficult to resolve homopolymeric stretches of sequence (e.g. CCCCC or AAAAA). In later sections we will provide a comparison with other systems that have solved this problem. Another problem is that the processing of samples before they are put into the sequencing machine is extremely labor intense. In retrospect, 454 protocols are changing monthly and it is anticipated that the protocols will become more finely tuned with time.

The Illumina/Solexa GA system (http://www.illumina.com/) is very different from 454 sequencing. Specifically, in the Solexa system, the targets are amplified on a solid surface. After amplification, only one of the strands is sequenced with all four deoxyribonucleotide triphosphates present during sequencing (not one at a time, as in the case of the 454 system). Each deoxyribonucleotide triphosphate has a unique fluorophore. Reversible terminator nucleotides (also called 'cyclic reversible termination' (86) are used to prevent the insertion of multiple nucleotide bases in the same cycle. In detail, the DNA is fragmented and adaptor sequences are added to each end of the fragments. The fragments (ca. DNA templates) are then sent to a lawn of immobilized oligonucleotides that are grafted to the surface of a microfluidic chamber. The DNA templates are hybridized to the immobilized oligonucleotides by the adaptors. Once attached, the DNA templates are copied using 'bridge amplification' (2). Bridge amplification involves the tethered DNA template arching over and hybridizing to an adjacent anchored oligonucleotide, forming a bridge. Amplification of a single DNA molecule results in a cluster of molecules composed of the same sequence. Following amplification, the reverse strands of the DNA are denatured and washed away, resulting in clusters of unique immobilized ssDNA. DNA sequencing begins with

the addition of polymerase, fluorescently labeled deoxyribonucleotide triphosphates and a primer that hybridizes to one of the adaptors. The incorporation of a complementary base results in a burst of light that is recorded by a charge-coupled device camera. Unlike the 454 sequencing system, the fluorophore is removed from the incorporated base, washed away and the cycle is repeated. This prevents the addition of more than one base per cycle.

The strength of the Solexa system is that it can generate 1.5 Gb of sequence per run. However, read lengths range from 35 to 100 bases and each run requires 3–5 days to complete (120, 140). To deal with short read length, the confidence of the sequence reads is improved by using 'pair-end sequencing', which means that both ends are sequenced. A weakness of the Solexa system is that it tends to produce biased sequence coverage that occurs in AT-rich repetitive sequences, presumably because of short read lengths (44). The system can also be affected by 'dephasing noise' that occurs when a complementary nucleotide is not incorporated or when the fluorophore is not properly cleaved at the end of the cycle – blocking the incorporation of the next nucleotide base. As a consequence, the sequence is out-of-phase for the remainder of the template (29). It is also important to recognize that only 43% of the raw reads are useable compared with 95% for 454 sequencing.

The SOLiD ('support oligonucleotide ligation detection') system is similar to the previously mentioned systems in that it depends on the ligation of adaptors to fragmented templates and the attachment of the adaptors to solid supports (beads in the case of SOLiD), which is followed by emulsion PCR amplification. Beads with extended templates undergo modification so that they can covalently attach to the surface of a slide (Fig. 8). DNA sequencing involves five different primers that differ in their position on the adaptor template by one nucleotide. In the first round of ligation, the first primer, P1, hybridizes to the template, and a set of 16 fluorescently labeled dibase probes competes for ligation (e.g. dibase probe: CANNNNNN or GGNNNNNN). The specificity of the probe is achieved by interrogating every first and second base in each ligation reaction. A charge-coupled device camera records the dibase signal and then the fluorophore is cleaved off, allowing other probes to interrogate the template. Multiple cycles of ligation, detection and cleavage are performed with the number of cycles determining the read length. Following these ligation cycles, the extension product is removed and the template is reset with the next primer complementary to the n-1 position for the next round of ligation cycles. It should be noted that each base in the template is interrogated twice

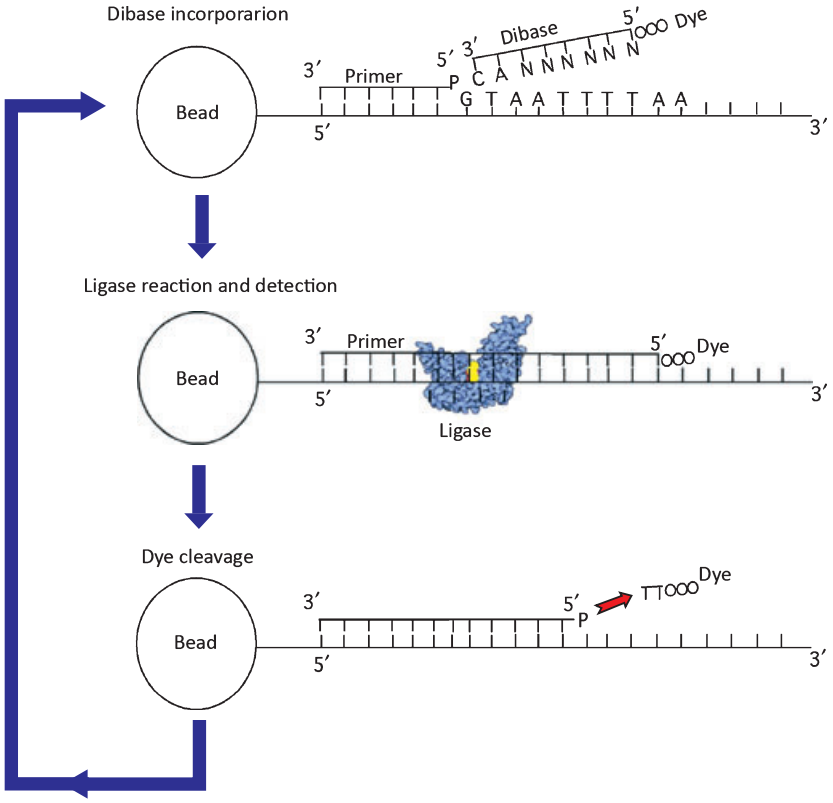

Fig. 8. Cartoon of the SOLiD (support oligonucleotide ligation detection) system, showing the hybridization of a fluorescently labeled dibase. First, the P1 primer hybridizes to the adaptor, then one of the 16 dibases competes for hybridization to the template. The dibase is then ligated to the 5′ phosphate (P) of the primer, the excess dibases are washed away and the fluorescence of the dye is recorded. Next, the dye is cleaved off and another cycle begins with the addition of all the dibases. This cycle is repeated a further nine times. Then, the P1 primer and a new sequence strand are melted away, a new primer (P2) is added, and the cycle is repeated for a total of 10 times. Altogether the primer is reset five times, resulting in a total of 50 ligation cycles.

because the primers are shifted to the left (upstream). Differences in the fluorescent dyes and bioinformatics are used to identify the dibases that have hybridized to the DNA template. Bioinfomatics are used to assemble the code of the template.

The strength of the SOLiD system is that it can generate 4 Gb of sequence but the reads are only 35 nucleotides (140). The two-base encoding system provides better sequence fidelity than the one-base next-generation sequencing systems. The weakness of the SOLiD system is, similarly to the Solexa system, that it yields biased sequence coverage in AT-rich repetitive sequences (44) and only 35% of the raw reads are useable, compared with 95% for the 454 system. Another weakness is it requires long run times (ca. 6 days).

In summary, 454 sequencing seems to be the most accurate system, although sample preparation involves an extensive amount of work. Other systems may also be highly labor-intensive; however, we have no experience with them. Regarding physicochemical considerations, the next-generation sequencing systems may be affected by surface effects in the same way as microarrays. It remains to be tested if these systems could provide quantitative measurements of the targets.

Second- and third-generation next-generation sequencing

Second-generation next-generation sequencing

Second-generation next-generation sequencing involves sequencing single molecules in real time, does not involve DNA amplification and thus eliminates error rates associated with PCR and intensity biases. The Pacific Biosystem system (http://www.pacificbiosciences.com/) allows detection of single molecules in real time using phospho-linked nucleotides for incorporation with each nucleotide having a different fluorescent label. The advantage of these deoxyribonucleotide triphosphates is that the phospholinked fluorophores are released from the base when it is incorporated in the DNA sequence by the polymerase. Pacific Biosystems uses a modified phi29 DNA polymerase that has high kinetic properties and can effectively handle the phospho-linked fluorophores (140). Unlike the previously mentioned technologies, the phi29 polymerase is immobilized to the bottom of a 70-nm-wide metallic well, which is illuminated through a glass support. As the DNA polymerase incorporates a complementary nucleotide, the enzyme cleaves the fluorescently labeled phosphate from the base, resulting in a burst of light (31). Using a nanophototonic visualization chamber and zero-mode waveguide technology, the burst of light is recorded in real time (59). Key to the visualization chamber is the fact that the zero-mode waveguide only records light in close proximity to the surface and therefore disregards other fluorescent events, such as those of fluorophores diffusing away into the solution. It was recently demonstrated that the approach resulted in long read lengths ($>$1,000 nucleotides) at a speed of 10 bp$/$s (33).

The strength of the Pacific Biosystem system is that it allows very long reads ($>$1,000 nucleotides). A weakness lies in their highest error rate (ca. 17%) (31) of all next-generation sequencing systems. For example, of the 27 errors in 158 bases, 12 were deletions, eight were insertions and seven were mismatches. These problems might be addressed in future versions of the system. For example, Eid et al. suggests that re-engineering the polymerase enzyme and$/$or the zero-mode waveguide technology might improve the system's performance. Improvements might also be addressed by performing repeated sequence analyses. Repeating the same sequence 15 times improved the consensus accuracy by $>$99% (31).

Third-generation next-generation sequencing

At the time this review was written, third-generation next-generation sequencing systems were not commercially available. However, the development of these systems is being revealed in the scientific literature. These systems seem to be label-free, do not require the analytes to be immobilized on a surface, do not depend on DNA polymerase, allow single-molecule analysis with a high signal-to-noise ratio and have the potential to sequence DNA strands at high speed and low cost (18). These new systems are based on nanopore technology. It is important to know that nanopores vary in size (≥2–5 nm), shape and composition (solid-state membranes, protein channels in lipid membranes, or polymer-based fluidic channels) (49). The idea behind nanopore technology is that a strand of nucleic acid can be electrophoretically driven through a nanopore, with each nucleotide base modulating the ionic current as it passes through the nanopore (18). The modulation of ionic current is caused by the partial blockage of the 'translocating' molecule, which varies for different molecules (4). Although initial experiments have shown that modulation of ionic current occurs for polyA and polyC within the same strand, these modulations were later found to be more a reflection of base stacking and other secondary structural

effects than single nucleotide detection (18). In other words, it was not possible to measure differences in individual nucleotide bases because the ion-current differences are too small. Furthermore, the ion-current blockages were found to be a consequence of 10–15 nucleotides positioned at the opening of the nanopore. Hence, it was not possible to detect a single molecule. The issue is further complicated by pore size and the speed at which the DNA strand moves through the nanopore, which are interdependent. The pore has to be big enough to let a nucleotide strand through and, at the same time, allow the electrical reading from one side of pore to the other. Apparently, DNA sequences move too fast through the nanopore to resolve the differences in charge of the bases (84). According to Branton et al. (18), one approach to resolve this issue is to slow the movement of the strand through the pore so that each base occupies the detector for greater than 1 ms. The movement of the DNA strand can be slowed by controlling electrolyte temperature, salt concentration, viscosity and the electrical bias voltage across the nanopore (34), by using enzymes that affect DNA processivity, or by using nanopores in conjunction with target hybridization (i.e. hybridize small oligonucleotide probes to the target and then distinguish double-stranded DNA from single-stranded DNA as it passes through a nanopore).

Oxford Nanopore Technology (http://www.nanoporetech.com) and NABsys (http://www.nabsys.com/) have developed solutions to the above-mentioned problems. A review by Branton et al. (18) concludes that future nanopore systems will probably be based on nanopores that are hybrids between solid-state (e.g. NABsys) and alpha-hemolysins (e.g. Oxford Nanopore Technology), and that improvements in electronic sensing using tunneling probes or capacitors will probably improve the detection of single bases through nanopores. Ultimately, nanopore technology promises to result in faster and inexpensive DNA sequencing.

Outlook

Implementation of next-generation sequencing systems to understand microbial communities and microbial dynamics is changing our view of the human oral microbiome, in the same way that Sanger sequencing changed our view of the diversity and metabolism of the microbial world. Despite the short time that next-generation sequencing systems have been commercially available to researchers (ca. 4 years), there has been an explosion in the number of papers publishing research that has used this technology. Next-generation sequencing systems have been used to examine the microbial diversity and structure of 16S ribosomal RNA genes or other specific genes in different biological and environmental settings (21, 30, 38, 63, 73, 80, 87, 102, 118, 126, 132, 142, 150), to determine the total metabolic capabilities of entire microbial communities (27, 28, 30, 53, 116, 128), to examine the natural sequence heterogeneity in microbial functional genes (139) and to conduct whole-genome analysis of a single microbial species (45, 70). In the near future, next-generation sequencing will not only tell us who is there in the oral cavity, but also what they are metabolically doing. We will also be able to study the emergent properties of microbial communities. This is very important in a clinical sense because it will enable us to see how entire communities respond to therapies and to specific host conditions. Furthermore, we will understand how the host responds to changes in microbial communities.

As we have shown in this review, microarrays were once thought to be the big hope for high-throughput microbial identification and quantification. In reality, many problems were discovered and a great number of publications have been specifically devoted to solving those problems. Ways to deal with these problems do exist, and microarrays might have some utility as diagnostic tools in the future.

Currently, we see a blossoming of new next-generation sequencing technologies, such as Pacific Biosystems and Oxford Nanopore Technology. Soon, a flood of data will settle into various databases for public use. However, two questions persist. How significant will this large body of information be in terms of advancing science? Will these data be translated into biologically- and, even more challenging, clinically meaningful knowledge? Like all technologies, next-generation sequencing has its problems – some of which have not yet been revealed. It is very tempting to rush on with applications and to generate gigabases of sequences. However, we must pause to consider the lessons learned from microarrays. Hopefully, with next-generation sequencing, the physicists and chemists will work together with the biologists to provide reality checks for the data obtained.

Acknowledgments

Funding was provided by the National Science Foundation CREST (HRD#0734232) and NSF DBI-1038671.

References

1. Aas JA, Paster BJ, Stokes LN, Olsen I, Dewhirst FE. Defining the normal bacterial flora of the oral cavity. *J Clin Microbiol* 2005: **43**: 5721–5732.
2. Adessi C, Matton G, Ayala G, Turcatti G, Mermod JJ, Mayer P, Kawashima E. Solid phase DNA amplification: characterisation of primer attachment and amplification mechanisms. *Nucleic Acids Res* 2000: **28**: e87.
3. Asai Y, Jinno T, Igarashi H, Ohyama Y, Ogawa T. Detection and quantification of oral treponemes in subgingival plaque by real-time PCR. *J Clin Microbiol* 2002: **40**: 3334–3340.
4. Ashkenasy N, Sánchez-Quesada J, Bayley H, Ghadiri MR. Recognizing a single base in an individual DNA strand: a step toward DNA sequencing in nanopores. *Angew Chem Int Ed Engl* 2005: **44**: 1401–1404.
5. Bäckhed F, Ding H, Wang T, Hooper LV, Koh GY, Nagy A, Semenkovich CF, Gordon JI. The gut microbiota as an environmental factor that regulates fat storage. *Proc Natl Acad Sci USA* 2004: **101**: 15718–15723.
6. Bäckhed F, Manchester JK, Semenkovich CF, Gordon JI. Mechanisms underlying the resistance to diet-induced obesity in germ-free mice. *Proc Natl Acad Sci USA* 2007: **104**: 979–984.
7. Beikler T, Abdeen G, Schnitzer S, Salzer S, Ehmke B, Heinecke A, Flemmig TF. Microbiological shifts in intra- and extraoral habitats following mechanical periodontal therapy. *J Clin Periodontol* 2004: **31**: 777–783.
8. Beikler T, Prior K, Ehmke B, Flemmig TF. Specific antibiotics in the treatment of periodontitis – a proposed strategy. *J Periodontol* 2004: **75**: 169–175.
9. Beikler T, Schnitzer S, Abdeen G, Ehmke B, Eisenacher M, Flemmig TF. Sampling strategy for intraoral detection of periodontal pathogens before and following periodontal therapy. *J Periodontol* 2006: **77**: 1323–1332.
10. Berg RD. The indigenous gastrointestinal microflora. *Trends Microbiol* 1996: **4**: 431–435.
11. Binder H, Preibisch S. Specific and nonspecific hybridization of oligonucleotide probes on microarrays. *Biophys J* 2005: **89**: 337–352.
12. Bishop J, Chagovetz AM, Blair S. Kinetics of multiplex hybridization: mechanisms and implications. *Biophys J* 2008: **94**: 1726–1734.
13. Bostrom L, Bergstrom J, Dahlén G, Linder LE. Smoking and subgingival microflora in periodontal disease. *J Clin Periodontol* 2001: **28**: 212–219.
14. Boutaga K, Savelkoul PH, Winkel EG, van Winkelhoff AJ. Comparison of subgingival bacterial sampling with oral lavage for detection and quantification of periodontal pathogens by real-time polymerase chain reaction. *J Periodontol* 2007: **78**: 79–86.
15. Boutaga K, van Winkelhoff AJ, Vandenbroucke-Grauls CM, Savelkoul PH. The additional value of real-time PCR in the quantitative detection of periodontal pathogens. *J Clin Periodontol* 2006: **33**: 427–433.
16. Boutaga K, van Winkelhoff AJ, Vandenbroucke-Grauls CMJE, Savelkoul PHM. Comparison of real-time PCR and culture for detection of *Porphyromonas gingivalis* in subgingival plaque samples. *J Clin Microbiol* 2003: **41**: 4950–4954.
17. Boyanton BL Jr, Luna RA, Fasciano LR, Menne KG, Versalovic J. DNA pyrosequencing-based identification of pathogenic *Candida* species by using the internal transcribed spacer 2 region. *Arch Pathol Lab Med* 2008: **132**: 667–674.
18. Branton D, Deamer DW, Marziali A, Bayley H, Benner SA, Butler T, Di Ventra M, Garaj S, Hibbs A, Huang X, Jovanovich SB, Krstic PS, Lindsay S, Ling XS, Mastrangelo CH, Meller A, Oliver JS, Pershin YV, Ramsey JM, Riehn R, Soni GV, Tabard-Cossa V, Wanunu M, Wiggin M, Schloss JA. The potential and challenges of nanopore sequencing. *Nat Biotechnol* 2008: **26**: 1146–1153.
19. Brodie EL, DeSantis TZ, Joyner DC, Baek SM, Larsen JT, Andersen GL, Hazen TC, Richardson PM, Herman DJ, Tokunaga TK, Wan JM, Firestone MK. Application of a high-density oligonucleotide microarray approach to study bacterial population dynamics during uranium reduction and reoxidation. *Appl Environ Microbiol* 2006: **72**: 6288–6298.
20. Brodie EL, DeSantis TZ, Parker JP, Zubietta IX, Piceno YM, Andersen GL. Urban aerosols harbor diverse and dynamic bacterial populations. *Proc Natl Acad Sci USA* 2007: **104**: 299–304.
21. Brown MV, Philip GK, Bunge JA, Smith MC, Bissett A, Lauro FM, Fuhrman JA, Donachie SP. Microbial community structure in the North Pacific ocean. *ISME J* 2009: **3**: 1374–1386.
22. Clement BG, Kehl LE, DeBord KL, Kitts CL. Terminal restriction fragment patterns (TRFPs), a rapid, PCR-based method for the comparison of complex bacterial communities. *J Microbiol Methods* 1998: **31**: 135–142.
23. Colombo AP, Boches SK, Cotton SL, Goodson JM, Kent R, Haffajee AD, Socransky SS, Hasturk H, Van Dyke TE, Dewhirst F, Paster BJ. Comparisons of subgingival microbial profiles of refractory periodontitis, severe periodontitis, and periodontal health using the human oral microbe identification microarray. *J Periodontol* 2009: **80**: 1421–1432.
24. Craig RG, Boylan R, Yip J, Bamgboye P, Koutsoukos J, Mijares D, Ferrer J, Imam M, Socransky SS, Haffajee AD. Prevalence and risk indicators for destructive periodontal diseases in 3 urban American minority populations. *J Clin Periodontol* 2001: **28**: 524–535.
25. Darout IA, Albandar JM, Skaug N, Ali RW. Salivary microbiota levels in relation to periodontal status, experience of caries and miswak use in Sudanese adults. *J Clin Periodontol* 2002: **29**: 411–420.
26. DeSantis TZ, Stone CE, Murray SR, Moberg JP, Andersen GL. Rapid quantification and taxonomic classification of environmental DNA from both prokaryotic and eukaryotic origins using a microarray. *FEMS Microbiol Lett* 2005: **245**: 271–278.
27. Dinsdale EA, Edwards RA, Hall D, Angly F, Breitbart M, Brulc JM, Furlan M, Desnues C, Haynes M, Li L, McDaniel L, Moran MA, Nelson KE, Nilsson C, Olson R, Paul J, Brito BR, Ruan Y, Swan BK, Stevens R, Valentine DL, Thurber RV, Wegley L, White BA, Rohwer F. Functional metagenomic profiling of nine biomes. *Nature* 2008: **452**: 629–632.
28. Dinsdale EA, Pantos O, Smriga S, Edwards RA, Angly F, Wegley L, Hatay M, Hall D, Brown E, Haynes M, Krause L, Sala E, Sandin SA, Thurber RV, Willis BL, Azam F, Knowlton N, Rohwer F. Microbial ecology of four coral

atolls in the Northern Line Islands. *PLoS ONE* 2008: **3**: e1584.
29. Dohm JC, Lottaz C, Borodina T, Himmelbauer H. Substantial biases in ultra-short read data sets from high-throughput DNA sequencing. *Nucleic Acids Res* 2008: **36**: e105.
30. Edwards RA, Rodriguez-Brito B, Wegley L, Haynes M, Breitbart M, Peterson DM, Saar MO, Alexander S, Alexander EC Jr, Rohwer F. Using pyrosequencing to shed light on deep mine microbial ecology. *BMC Genomics* 2006: **7**: 57.
31. Eid J, Fehr A, Gray J, Luong K, Lyle J, Otto G, Peluso P, Rank D, Baybayan P, Bettman B, Bibillo A, Bjornson K, Chaudhuri B, Christians F, Cicero R, Clark S, Dalal R, Dewinter A, Dixon J, Foquet M, Gaertner A, Hardenbol P, Heiner C, Hester K, Holden D, Kearns G, Kong X, Kuse R, Lacroix Y, Lin S, Lundquist P, Ma C, Marks P, Maxham M, Murphy D, Park I, Pham T, Phillips M, Roy J, Sebra R, Shen G, Sorenson J, Tomaney A, Travers K, Trulson M, Vieceli J, Wegener J, Wu D, Yang A, Zaccarin D, Zhao P, Zhong F, Korlach J, Turner S. Real-time DNA sequencing from single polymerase molecules. *Science* 2009: **323**: 133–138.
32. Feres M, Haffajee AD, Allard K, Som S, Socransky SS. Change in subgingival microbial profiles in adult periodontitis subjects receiving either systemically-administered amoxicillin or metronidazole. *J Clin Periodontol* 2001: **28**: 597–609.
33. Flusberg BA, Webster DR, Lee JH, Travers KJ, Olivares EC, Clark TA, Korlach J, Turner SW. Direct detection of DNA methylation during single-molecule, real-time sequencing. *Nat Methods* 2010: **7**: 461–465.
34. Fologea D, Uplinger J, Thomas B, McNabb DS, Li JL. Slowing DNA translocation in a solid-state nanopore. *Nano Lett* 2005: **5**: 1734–1737.
35. Forns X, Bukh J, Purcell RH, Emerson SU. How *Escherichia coli* can bias the results of molecular cloning: preferential selection of defective genomes of hepatitis C virus during the cloning procedure. *Proc Natl Acad Sci USA* 1997: **94**: 13909–13914.
36. Fujimoto C, Maeda H, Kokeguchi S, Takashiba S, Nishimura F, Arai H, Fukui K, Murayama Y. Application of denaturing gradient gel electrophoresis (DGGE) to the analysis of microbial communities of subgingival plaque. *J Periodontal Res* 2003: **38**: 440–445.
37. Genco RJ. Using antimicrobial agents to manage periodontal diseases. *J Am Dent Assoc* 1991: **122**: 30–38.
38. Gilbert JA, Field D, Swift P, Newbold L, Oliver A, Smyth T, Somerfield PJ, Huse S, Joint I. The seasonal structure of microbial communities in the Western English Channel. *Environ Microbiol* 2009: **11**: 3132–3139.
39. Gill SR, Pop M, Deboy RT, Eckburg PB, Turnbaugh PJ, Samuel BS, Gordon JI, Relman DA, Fraser-Liggett CM, Nelson KE. Metagenomic analysis of the human distal gut microbiome. *Science* 2006: **312**: 1355–1359.
40. Git A, Dvinge H, Salmon-Divon M, Osborne M, Kutter C, Hadfield J, Bertone P, Caldas C. Systematic comparison of microarray profiling, real-time PCR, and next-generation sequencing technologies for measuring differential micrRNA expression. *RNA* 2010: **16**: 991–1006.
41. Haffajee AD, Bogren A, Hasturk H, Feres M, Lopez NJ, Socransky SS. Subgingival microbiota of chronic periodontitis subjects from different geographic locations. *J Clin Periodontol* 2004: **31**: 996–1002.
42. Haffajee AD, Socransky SS. Relationship of cigarette smoking to the subgingival microbiota. *J Clin Periodontol* 2001: **28**: 377–388.
43. Haffajee AD, Socransky SS, Gunsolley JC. Systemic antiinfective periodontal therapy. A systematic review. *Ann Periodontol* 2003: **8**: 115–181.
44. Harismendy O, Ng PC, Strausberg RL, Wang X, Stockwell TB, Beeson KY, Schork NJ, Murray SS, Topol EJ, Levy S, Frazer KA. Evaluation of next generation sequencing platforms for population targeted sequencing studies. *Genome Biol* 2009: **10**: R32.
45. Harris TD, Buzby PR, Babcock H, Beer E, Bowers J, Braslavsky I, Causey M, Colonell J, Dimeo J, Efcavitch JW, Giladi E, Gill J, Healy J, Jarosz M, Lapen D, Moulton K, Quake SR, Steinmann K, Thayer E, Tyurina A, Ward R, Weiss H, Xie Z. Single-molecule DNA sequencing of a viral genome. *Science* 2008: **320**: 106–109.
46. He Z, Wu L, Li X, Fields MW, Zhou J. Empirical establishment of oligonucleotide probe design criteria. *Appl Environ Microbiol* 2005: **71**: 3753–3760.
47. Heitz-Mayfield LJ, Lang NP. Comparative biology of chronic and aggressive periodontitis vs. peri-implantitis. *Periodontol 2000* 2010: **53**: 167–181.
48. Held GA, Grinstein G, Tu Y. Relationship between gene expression and observed intensities in DNA microarrays – a modeling study. *Nucleic Acids Res* 2006: **34**: e70.
49. Howorka S, Siwy Z. Nanopore analytics: sensing of single molecules. *Chem Soc Rev* 2009: **38**: 2360–2384.
50. Hutter G, Schlagenhauf U, Valenza G, Horn M, Burgemeister S, Claus H, Vogel U. Molecular analysis of bacteria in periodontitis: evaluation of clone libraries, novel phylotypes and putative pathogens. *Microbiology* 2003: **149**: 67–75.
51. Huyghe A, Francois P, Charbonnier Y, Tangomo-Bento M, Bonetti EJ, Paster BJ, Bolivar I, Baratti-Mayer D, Pittet D, Schrenzel J; Geneva Study Group on Noma (GESNOMA). Novel microarray design strategy to study complex bacterial communities. *Appl Environ Microbiol* 2008: **74**: 1876–1885.
52. Jenkinson HF, Lamont RJ. Oral microbial communities in sickness and in health. *Trends Microbiol* 2005: **13**: 589–595.
53. Kalyuzhnaya MG, Lapidus A, Ivanova N, Copeland AC, McHardy AC, Szeto E, Salamov A, Grigoriev IV, Suciu D, Levine SR, Markowitz VM, Rigoutsos I, Tringe SG, Bruce DC, Richardson PM, Lidstrom ME, Chistoserdova L. High-resolution metagenomics targets specific functional types in complex microbial communities. *Nat Biotechnol* 2008: **26**: 1029–1034.
54. Kawada M, Yoshida A, Suzuki N, Nakano Y, Saito T, Oho T, Koga T. Prevalence of *Porphyromonas gingivalis* in relation to periodontal status assessed by real-time PCR. *Oral Microbiol Immunol* 2004: **19**: 289–292.
55. Kilian M, Frandsen EV, Haubek D, Poulsen K. The etiology of periodontal disease revisited by population genetic analysis. *Periodontol 2000* 2006: **42**: 158–179.
56. Kitts CL. Terminal restriction fragment patterns: a tool for comparing microbial communities and assessing community dynamics. *Curr Issues Intest Microbiol* 2001: **2**: 17–25.
57. Kolenbrander PE. Oral microbial communities: biofilms, interactions, and genetic systems. *Annu Rev Microbiol* 2000: **54**: 413–437.

58. Kolenbrander PE, Palmer RJ Jr, Periasamy S, Jakubovics NS. Oral multispecies biofilm development and the key role of cell-cell distance. *Nat Rev Microbiol* 2010: **8**: 471–480.
59. Korlach J, Marks PJ, Cicero RL, Gray JJ, Murphy DL, Roitman DB, Pham TT, Otto GA, Foquet M, Turner SW. Selective aluminum passivation for targeted immobilization of single DNA polymerase molecules in zero-mode waveguide nanostructures. *Proc Natl Acad Sci USA* 2008: **105**: 1176–1181.
60. Kroes I, Lepp PW, Relman DA. Bacterial diversity within the human subgingival crevice. *Proc Natl Acad Sci USA* 1999: **96**: 14547–14552.
61. Kuboniwa M, Lamont RJ. Subgingival biofilm formation. *Periodontol 2000* 2010: **52**: 38–52.
62. Kuramitsu HK, He XS, Lux R, Anderson MH, Shi WY. Interspecies interactions within oral microbial communities. *Microbiol Mol Biol Rev* 2007: **71**: 653–670.
63. Lazarevic V, Whiteson K, Huse S, Hernandez D, Farinelli L, Osterås M, Schrenzel J, François P. Metagenomic study of the oral microbiota by Illumina high-throughput sequencing. *J Microbiol Methods* 2009: **79**: 266–271.
64. Ley RE, Peterson DA, Gordon JI. Ecological and evolutionary forces shaping microbial diversity in the human intestine. *Cell* 2006: **124**: 837–848.
65. Li J, Helmerhorst EJ, Leone CW, Troxler RF, Yaskell T, Haffajee AD, Socransky SS, Oppenheim FG. Identification of early microbial colonizers in human dental biofilm. *J Appl Microbiol* 2004: **97**: 1311–1318.
66. Li S, Pozhitkov A, Brouwer M. A competitive hybridization model predicts signal intensity on high density DNA microarrays. *Nucleic Acids Res* 2008: **20**: 6585–6591.
67. Li X, He Z, Zhou J. Selection of optimal oligonucleotide probes for microarrays using multiple criteria, global alignment and parameter estimation. *Nucleic Acids Res* 2005: **33**: 6114–6123.
68. Lipshutz RJ, Fodor SP, Gingeras TR, Lockhart DJ. High-density synthetic oligonucleotide arrays. *Nat Genet* 1999: **21**: 20–24.
69. Liu WT, Marsh TL, Cheng H, Forney LJ. Characterization of microbial diversity by determining terminal restriction fragment length polymorphisms of genes encoding 16S rRNA. *Appl Environ Microbiol* 1997: **63**: 4516–4522.
70. Lowder BV, Guinane CM, Ben Zakour NL, Weinert LA, Conway-Morris A, Cartwright RA, Simpson AJ, Rambaut A, Nübel U, Fitzgerald JR. Recent human-to-poultry host jump, adaptation, and pandemic spread of *Staphylococcus aureus*. *Proc Natl Acad Sci USA* 2009: **106**: 19545–19550.
71. Loy A, Bodrossy L. Highly parallel microbial diagnostics using oligonucleotide microarrays. *Clin Chim Acta* 2006: **363**: 106–119.
72. Loy A, Lehner A, Lee N, Adamczyk J, Meier H, Ernst J, Schleifer KH, Wagner M. Oligonucleotide microarray for 16S rRNA gene-based detection of all recognized lineages of sulfate-reducing prokayotes in the environment. *Appl Environ Microbiol* 2002: **68**: 5064–5081.
73. Lumini E, Orgiazzi A, Borriello R, Bonfante P, Bianciotto V. Disclosing arbuscular mycorrhizal fungal biodiversity in soil through a land-use gradient using a pyrosequencing approach. *Environ Microbiol* 2010: **12**: 2165–2179.
74. Lyons SR, Griffen AL, Leys EJ. Quantitative real-time PCR for *Porphyromonas gingivalis* and total bacteria. *J Clin Microbiol* 2000: **38**: 2362–2365.
75. MacLean D, Jones JD, Studholme DJ. Application of 'next-generation' sequencing technologies to microbial genetics. *Nat Rev Microbiol* 2009: **7**: 287–296.
76. Maeda H, Fujimoto C, Haruki Y, Maeda T, Kokeguchi S, Petelin M, Arai H, Tanimoto I, Nishimura F, Takashiba S. Quantitative real-time PCR using Taq-Man and SYBR Green for *Actinobacillus actinomycetemcomitans, Porphyromonas gingivalis, Prevotella intermedia, tetQ* gene and total bacteria. *FEMS Immunol Med Microbiol* 2003: **39**: 81–86.
77. Mager DL, Ximenez-Fyvie LA, Haffajee AD, Socransky SS. Distribution of selected bacterial species on intraoral surfaces. *J Clin Periodontol* 2003: **30**: 644–654.
78. Margulies M, Egholm M, Altman WE, Attiya S, Bader JS, Bemben LA, Berka J, Braverman MS, Chen YJ, Chen Z, Dewell SB, Du L, Fierro JM, Gomes XV, Godwin BC, He W, Helgesen S, Ho CH, Irzyk GP, Jando SC, Alenquer ML, Jarvie TP, Jirage KB, Kim JB, Knight JR, Lanza JR, Leamon JH, Lefkowitz SM, Lei M, Li J, Lohman KL, Lu H, Makhijani VB, McDade KE, McKenna MP, Myers EW, Nickerson E, Nobile JR, Plant R, Puc BP, Ronan MT, Roth GT, Sarkis GJ, Simons JF, Simpson JW, Srinivasan M, Tartaro KR, Tomasz A, Vogt KA, Volkmer GA, Wang SH, Wang Y, Weiner MP, Yu P, Begley RF, Rothberg JM. Genome sequencing in microfabricated high-density picolitre reactors. *Nature* 2005: **437**: 376–380.
79. Marsh PD. Dental plaque as a biofilm and a microbial community – implications for health and disease. *BMC Oral Health* 2005: **6**: S14.
80. Martínez I, Wallace G, Zhang C, Legge R, Benson AK, Carr TP, Moriyama EN, Walter J. Diet-induced metabolic improvements in a hamster model of hypercholesterolemia are strongly linked to alterations of the gut microbiota. *Appl Environ Microbiol* 2009: **75**: 4175–4184.
81. Matveeva OV, Shabalina SA, Nemtsov VA, Tsodikov AD, Gesteland RF, Atkins JF. Thermodynamic calculations and statistical correlations for oligo-probes design. *Nucleic Acids Res* 2003: **31**: 4211–4217.
82. Mei R, Hubbell E, Bekiranov S, Mittmann M, Christians FC, Shen MM, Lu G, Fang J, Liu WM, Ryder T, Kaplan P, Kulp D, Webster TA. Probe selection for high-density oligonucleotide arrays. *Proc Natl Acad Sci USA* 2003: **100**: 11237–11242.
83. Meka A, Bakthavatchalu V, Sathishkumar S, Lopez MC, Verma RK, Wallet SM, Bhattacharyya I, Boyce BF, Handfield M, Lamont RJ, Baker HV, Ebersole JL, Kesavalu L. *Porphyromonas gingivalis* infection-induced tissue and bone transcriptional profiles. *Mol Oral Microbiol* 2010: **25**: 61–74.
84. Meller A, Nivon L, Branton D. Voltage-driven DNA translocations through a nanopore. *Phys Rev Lett* 2001: **86**: 3435–3438.
85. Metfies K, Medlin LK. Feasibility of transferring fluorescent in situ hybridization probes to an 18S rRNA gene phylochip and mapping of signal intensities. *Appl Environ Microbiol* 2008: **74**: 2814–2821.
86. Metzker ML. Sequencing technologies – the next generation. *Nat Rev Genet* 2010: **11**: 31–46.
87. Miller SR, Strong AL, Jones KL, Ungerer MC. Bar-coded pyrosequencing reveals shared bacterial community properties along the temperature gradients of two alkaline hot springs in Yellowstone National Park. *Appl Environ Microbiol* 2009: **75**: 4565–4572.

88. Montero CI, Shea YR, Jones PA, Harrington SM, Tooke NE, Witebsky FG, Murray PR. Evaluation of pyrosequencing technology for the identification of clinically relevant non-dematiaceous yeasts and related species. *Eur J Clin Microbiol Infect Dis* 2008: **27**: 821–830.
89. Moore WE, Moore LV. The bacteria of periodontal diseases. *Periodontol 2000* 1994: **5**: 66–77.
90. Morey JS, Ryan JC, Van Dolah FM. Microarray validation: factors influencing correlation between oligonucleotide microarrays and real-time PCR. *Biol Proced Online* 2006: **8**: 175–193.
91. Morillo JM, Lau L, Sanz M, Herrera D, Silva A. Quantitative real-time PCR based on single copy gene sequence for detection of *Actinobacillus actinomycetemcomitans* and *Porphyromonas gingivalis*. *J Periodontal Res* 2003: **38**: 518–524.
92. Naef F, Lim DA, Patil N, Magnasco M. DNA hybridization to mismatched templates: a chip study. *Phys Rev E Stat Nonlin Soft Matter Phys* 2002: **65**: 040902.
93. Nobbs AH, Lamont RJ, Jenkinson HF. *Streptococcus* adherence and colonization. *Microbiol Mol Biol Rev* 2009: **73**: 407–450.
94. Nonnenmacher C, Dalpke A, Mutters R, Heeg K. Quantitative detection of periodontopathogens by real-time PCR. *J Microbiol Methods* 2004: **59**: 117–125.
95. Nonnenmacher C, Dalpke A, Rochon J, Flores-de-Jacoby L, Mutters R, Heeg K. Real-time polymerase chain reaction for detection and quantification of bacteria in periodontal patients. *J Periodontol* 2005: **76**: 1542–1549.
96. Palmer C, Bik EM, Eisen MB, Eckburg PB, Sana TR, Wolber PK, Relman DA, Brown PO. Rapid quantitative profiling of complex microbial populations. *Nucleic Acids Res* 2006: **34**: e5.
97. Paster BJ, Boches SK, Galvin JL, Ericson RE, Lau CN, Levanos VA, Sahasrabudhe A, Dewhirst FE. Bacterial diversity in human subgingival plaque. *J Bacteriol* 2001: **183**: 3770–3783.
98. Paster BJ, Dewhirst FE. Molecular microbial diagnosis. *Periodontol 2000* 2009: **51**: 38–44.
99. Peplies J, Lachmund C, Glöckner FO, Manz W. A DNA microarray platform based on direct detection of rRNA for characterization of freshwater sediment-related prokaryotic communities. *Appl Environ Microbiol* 2006: **72**: 4829–4838.
100. Periasamy S, Kolenbrander PE. Central role of the early colonizer *Veillonella* sp. in establishing multispecies biofilm communities with initial, middle, and late colonizers of enamel. *J Bacteriol* 2010: **192**: 2965–2972.
101. Periasamy S, Kolenbrander PE. Mutualistic biofilm communities develop with *Porphyromonas gingivalis* and initial, early, and late colonizers of enamel. *J Bacteriol* 2010: **191**: 6804–6811.
102. Poretsky RS, Hewson I, Sun S, Allen AE, Zehr JP, Moran MA. Comparative day/night metatranscriptomic analysis of microbial communities in the North Pacific subtropical gyre. *Environ Microbiol* 2009: **11**: 1358–1375.
103. Pozhitkov A, Chernov B, Yershov G, Noble PA. Evaluation of gel-pad oligonucleotide microarray technology using artificial neural networks. *Appl Environ Microbiol* 2005: **71**: 8663–8676.
104. Pozhitkov A, Noble PA, Domazet-Loso T, Nolte AW, Sonnenberg R, Staehler P, Beier M, Tautz D. Tests of rRNA hybridization to microarrays suggest that hybridization characteristics of oligonucleotide probes for species discrimination cannot be predicted. *Nucleic Acids Res* 2006: **34**: e66.
105. Pozhitkov A, Stemshorn K, Tautz D. An algorithm for the determination and quantification of components of nucleic acid mixtures based on single sequencing reactions. *BMC Bioinformatics* 2005: **6**: 281.
106. Pozhitkov AE, Bailey KD, Noble PA. Development of a statistically robust quantification method for microorganisms in mixtures using oligonucleotide microarrays. *J Microbiol Methods* 2007: **70**: 292–300.
107. Pozhitkov AE, Boube I, Brouwer MH, Noble PA. Beyond Affymetrix arrays: expanding the set of known hybridization isotherms and observing pre-wash signal intensities. *Nucleic Acids Res* 2010: **8**: e2.
108. Pozhitkov AE, Nies G, Kleinhenz B, Tautz D, Noble PA. Simultaneous quantification of multiple nucleic acid targets in complex rRNA mixtures using high density microarrays and nonspecific hybridization as a source of information. *J Microbiol Methods* 2008: **75**: 92–102.
109. Pozhitkov AE, Noble PA. Comment on ''Discrimination of shifts in a soil microbial community associated with TNT-contamination using a functional ANOVA of 16S rRNA hybridized to oligonucleotide microarrays''. *Environ Sci Technol* 2007: **41**: 1797–1798.
110. Pozhitkov AE, Noble PA. High variability in melting profiles from gel pad arrays. *Environ Microbiol* 2007: **9**: 1865.
111. Pozhitkov AE, Rule RA, Stedtfeld RD, Hashsham SA, Noble PA. Concentration dependency of nonequilibrium thermal dissociation curves in complex target samples. *J Microbiol Methods* 2008: **74**: 82–88.
112. Pozhitkov AE, Stedtfeld RD, Hashsham SA, Noble PA. Revision of the nonequilibrium thermal dissociation and stringent washing approaches for identification of mixed nucleic acid targets by microarrays. *Nucleic Acids Res* 2007: **35**: e70.
113. Pozhitkov AE, Tautz D, Noble PA. Oligonucleotide microarrays: widely applied – poorly understood. *Brief Funct Genomic Proteomic* 2007: **6**: 141–148.
114. Preza D, Olsen I, Aas JA, Willumsen T, Grinde B, Paster BJ. Bacterial profiles of root caries in elderly patients. *J Clin Microbiol* 2008: **46**: 2015–2021.
115. Preza D, Olsen I, Willumsen T, Boches SK, Cotton SL, Grinde B, Paster BJ. Microarray analysis of the microflora of root caries in elderly. *Eur J Clin Microbiol Infect Dis* 2009: **28**: 509–517.
116. Qu A, Brulc JM, Wilson MK, Law BF, Theoret JR, Joens LA, Konkel ME, Angly F, Dinsdale EA, Edwards RA, Nelson KE, White BA. Comparative metagenomics reveals host specific metavirulomes and horizontal gene transfer elements in the chicken cecum microbiome. *PLoS ONE* 2008: **3**: e2945.
117. Rappé MS, Giovannoni SJ. The uncultured microbial majority. *Annu Rev Microbiol* 2003: **57**: 369–394.
118. Roesch LF, Lorca GL, Casella G, Giongo A, Naranjo A, Pionzio AM, Li N, Mai V, Wasserfall CH, Schatz D, Atkinson MA, Neu J, Triplett EW. Culture-independent identification of gut bacteria correlated with the onset of diabetes in a rat model. *ISME J* 2009: **3**: 536–548.
119. Rosan B, Lamont RJ. Dental plaque formation. *Microbes Infect* 2000: **2**: 1599–1607.
120. Rothberg JM, Leamon JH. The development and impact of 454 sequencing. *Nat Biotechnol* 2008: **26**: 1117–1124.

121. Rouillard JM, Herbert CJ, Zuker M. Oligoarray: genome-scale oligonucleotide design for microarrays. *Bioinformatics* 2002: **18**: 486–487.
122. Rule RA, Pozhitkov AE, Noble PA. Use of hidden correlations in short oligonucleotide array data are insufficient for accurate quantification of nucleic acid targets in complex target mixtures. *J Microbiol Methods* 2009: **76**: 188–195.
123. Sakamoto M, Takeuchi Y, Umeda M, Ishikawa I, Benno Y. Application of terminal RFLP analysis to characterize oral bacterial flora in saliva of healthy subjects and patients with periodontitis. *J Med Microbiol* 2003: **52**: 79–89.
124. Sakamoto M, Umeda M, Ishikawa I, Benno Y. Comparison of the oral bacterial flora in saliva from a healthy subject and two periodontitis patients by sequence analysis of 16S rDNA libraries. *Microbiol Immunol* 2000: **44**: 643–652.
125. Sakellari D, Belibasakis G, Chadjipadelis T, Arapostathis K, Konstantinidis A. Supragingival and subgingival microbiota of adult patients with Down's syndrome. Changes after periodontal treatment. *Oral Microbiol Immunol* 2001: **16**: 376–382.
126. Sanapareddy N, Hamp TJ, Gonzalez LC, Hilger HA, Fodor AA, Clinton SM. Molecular diversity of a North Carolina wastewater treatment plant as revealed by pyrosequencing. *Appl Environ Microbiol* 2009: **75**: 1688–1696.
127. Shelburne CE, Prabhu A, Gleason RM, Mullally BH, Coulter WA. Quantitation of *Bacteroides forsythus* in subgingival plaque. Comparison of immunoassay and quantitative polymerase chain reaction. *J Microbiol Methods* 2000: **39**: 97–107.
128. Simon C, Wiezer A, Strittmatter AW, Daniel R. Phylogenetic diversity and metabolic potential revealed in a glacier ice metagenome. *Appl Environ Microbiol* 2009: **75**: 7519–7526.
129. Skvortsov D, Abdueva D, Curtis C, Schaub B, Tavaré S. Explaining differences in saturation levels for Affymetrix GeneChip arrays. *Nucleic Acids Res* 2007: **35**: 4154–4163.
130. Slots J, Jorgensen MG. Efficient antimicrobial treatment in periodontal maintenance care. *J Am Dent Assoc* 2000: **131**: 1293–1304.
131. Socransky SS, Haffajee AD. Periodontal microbial ecology. *Periodontol 2000* 2005: **38**: 135–187.
132. Suchodolski JS, Dowd SE, Westermarck E, Steiner JM, Wolcott RD, Spillmann T, Harmoinen JA. The effect of the macrolide antibiotic tylosin on microbial diversity in the canine small intestine as demonstrated by massive parallel 16S rRNA gene sequencing. *BMC Microbiol* 2009: **9**: 210.
133. Suzuki N, Yoshida A, Saito T, Kawada M, Nakano Y. Quantitative microbiological study of subgingival plaque by real-time PCR shows correlation between levels of *Tannerella forsythensis* and *Fusobacterium* spp. *J Clin Microbiol* 2004: **42**: 2255–2257.
134. Tannock GW. *Normal microflora*. New York: Chapman and Hall, 1995.
135. Tuohy MJ, Hall GS, Sholtis M, Procop GW. Pyrosequencing as a tool for the identification of common isolates of *Mycobacterium* sp. *Diagn Microbiol Infect Dis* 2005: **51**: 245–250.
136. Urakawa H, El Fantroussi S, Smidt H, Smoot JC, Tribou EH, Kelly JJ, Noble PA, Stahl DA. Optimization of single-base-pair mismatch discrimination in oligonucleotide microarrays. *Appl Environ Microbiol* 2003: **69**: 2848–2856.
137. Urakawa H, Noble PA, El Fantroussi S, Kelly JJ, Stahl DA. Single-base-pair discrimination of terminal mismatches by using oligonucleotide microarrays and neural network analyses. *Appl Environ Microbiol* 2002: **68**: 235–244.
138. van Winkelhoff AJ, Winkel EG. Antibiotics in periodontics: right or wrong? *J Periodontol* 2009: **80**: 1555–1558.
139. Varaljay VA, Howard EC, Sun S, Moran MA. Deep sequencing of a dimethylsulfoniopropionate-degrading gene (dmdA) by using PCR primer pairs designed on the basis of marine metagenomic data. *Appl Environ Microbiol* 2010: **76**: 609–617.
140. Voelkerding KV, Dames SA, Durtschi JD. Next-generation sequencing: from basic research to diagnostics. *Clin Chem* 2009: **55**: 641–658.
141. Wang Y, Barbacioru C, Hyland F, Xiao W, Hunkapiller KL, Blake J, Chan F, Gonzalez C, Zhang L, Samaha RR. Large scale real-time PCR validation on gene expression measurements from two commercial long-oligonucleotide microarrays. *BMC Genomics* 2006: **7**: 59.
142. Webster NS, Taylor MW, Behnam F, Lücker S, Rattei T, Whalan S, Horn M, Wagner M. Deep sequencing reveals exceptional diversity and modes of transmission for bacterial sponge symbionts. *Environ Microbiol* 2010: **12**: 2070–2082.
143. Wilson KH, Wilson WJ, Radosevich JL, DeSantis TZ, Viswanathan VS, Kuczmarski TA, Andersen GL. High-density microarray of small-subunit ribosomal DNA probes. *Appl Environ Microbiol* 2002: **68**: 2535–2541.
144. Wu C, Carta R, Zhang L. Sequence dependence of cross-hybridization on short oligo microarrays. *Nucleic Acids Res* 2005: **33**: e84.
145. Wu Z, Irizarry RA. Stochastic models inspired by hybridization theory for short oligonucleotide arrays. *J Comput Biol* 2005: **12**: 882–893.
146. Ximenez-Fyvie LA, Haffajee AD, Socransky SS. Comparison of the microbiota of supra- and subgingival plaque in subjects in health and periodontits. *J Clin Periodontol* 2000: **27**: 648–657.
147. Ximenez-Fyvie LA, Haffajee AD, Socransky SS. Microbial composition of supra- and subgingival plaque in subjects with adult periodontitis. *J Clin Periodontol* 2000: **27**: 722–732.
148. Yoshida A, Suzuki N, Nakano Y, Oho T, Kawada M, Koga T. Development of a 5′ fluorogenic nuclease-based real-time PCR assay for quantitative detection of *Actinobacillus actinomycetemcomitans* and *Porphyromonas gingivalis*. *J Clin Microbiol* 2003: **41**: 863–866.
149. Zak DR, Blackwood CB, Waldrop MP. A molecular dawn for biogeochemistry. *Trends Ecol Evol* 2006: **21**: 288–295.
150. Zaura E, Keijser BJ, Huse SM, Crielaard W. Defining the healthy ''core microbiome'' of oral microbial communities. *BMC Microbiol* 2009: **9**: 259.
151. Zhang L, Miles MF, Aldape KD. A model of molecular interactions on short oligonucleotide microarrays. *Nat Biotechnol* 2003: **7**: 818–821.
152. Zijnge V, Harmsen HJM, Kleinfelder JW, van der Rest ME, Degener JE, Welling GW. Denaturing gradient gel electrophoresis analysis to study bacterial community structure in pockets of periodontitis patients. *Oral Microbiol Immunol* 2003: **18**: 59–65.
153. Zuker M. Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic Acids Res* 2003: **31**: 3406–3415.