Oligonucleotide microarray methodology for taxonomic and functional monitoring of microbial community composition

M. Kyselková[1,2,3], J. Kopecký[4], M. Ságová-Marečková[4], G.L. Grundmann[1,2,3], Y. Moënne-Loccoz[1,2,3]

[1]*Université de Lyon, F-69622, Lyon, France*
[2]*Université Lyon 1, Villeurbanne, France*
[3]*CNRS, UMR5557, Ecologie Microbienne, Villeurbanne, France*
[4]*Crop Research Institute, Prague, Czech Republic*

ABSTRACT

Microarray analysis is a cultivation-independent, high-throughput technology that can be used for direct and simultaneous identification of microorganisms in complex environmental samples. This review summarizes current methodologies for oligonucleotide microarrays used in microbial ecology. It deals with probe design, microarray manufacturing, sample preparation and labeling, and data handling, as well as with the key features of microarray analysis such as specificity, sensitivity and quantification potential. Microarray analysis has been validated as an effective approach to describe the composition and dynamics of taxonomic and functional microbial communities, in environments including soil, compost, sediment, air or humans. It is now part of the technical arsenal available to address key issues in microbial community ecology, ranging from biogeography to ecosystem functioning.

Keywords: oligonucleotide microarray; taxonomic microarray; functional microarray; microbial community

The knowledge of microbial community composition is important to understand the role of microorganisms in various environmental processes, their response to changing conditions or their preference for certain types of environment. Microarray technology, originally developed for gene-expression studies (Schena et al. 1995), has been adapted in the last decade for a rapid assessment of the composition of complex microbial communities.

A microarray consists of a large number of DNA probes that are fixed on a solid support (Figure 1). The solid support is miniaturized; hence microarrays are also called microchips. The probes hybridize, based on DNA complementarity, with corresponding targets. Target nucleic acids are extracted directly from environmental samples (thereby avoiding cultivation), amplified and fluorescently labeled prior to hybridization. If a targeted microorganism is present in a sample, a fluorescent signal is detected on the corresponding probe(s) after hybridization. In addition, if probes targeting higher taxonomical levels are included, a microarray may also detect unknown microorganisms, i.e. those for which a low-taxonomic level probe is not available. Signal intensity is proportional to target quantity. Signals recorded from individual probes are numerized and analyzed statistically. Oligonucleotide microarrays comprise short (usually 20–70 nucleotides) DNA probes. Besides them, whole genome microarrays are used in environmental microbiology for detection and monitoring of selected strains (Zhou 2003, Wu et al. 2006). This approach, which is

Supported by the Ministry of Agriculture of the Czech Republic, Institutional Research Concept, Projects No. MZE 0002700604 and QH92151.

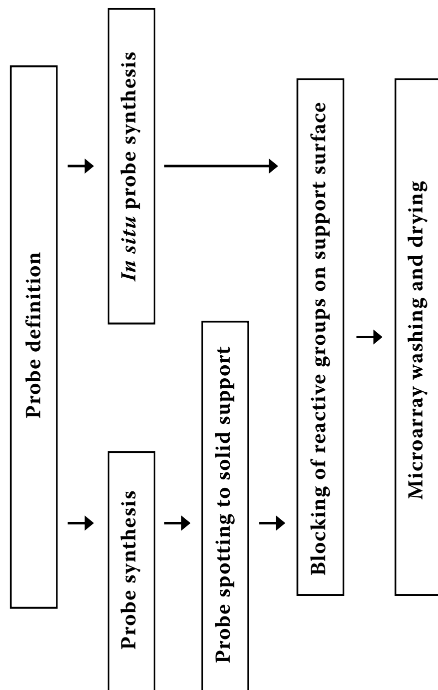

Figure 1. Microarray analysis of microbial community composition in environmental samples, from probe design and sample extraction to microarray hybridization and data analysis

limited by the fact that strains must be cultivated prior to microarray development, is not a subject of this review.

This review aims to summarize the recent progress in the methodology of oligonucleotide microarrays for microbial ecology. It deals with technical aspects typical for microarray technology, i.e. probe set design, microarray manufacturing, sample preparation and labeling, and data handling. In addition, we discuss key issues related to microarray methodology, i.e. specificity, sensitivity and quantification potential.

Taxonomic vs. functional microarrays

Taxonomic (phylogenetic) microarrays are an effective tool for the assessment of the taxonomic composition of a microbial community. They are based on molecular phylogeny markers, mostly on 16S rRNA gene *rrs* in the case of bacteria. The different level of conservation of regions within the *rrs* sequence allows to design probes with different taxonomic resolutions. For some groups of microorganisms, however, the taxonomic resolution of the *rrs* gene is poor and some related species (e.g. within the *Pseudomonas* genus, Sanguin et al. 2008) or even genera (e.g. within the *Enterobacteriaceae* family, Naum et al. 2008), cannot be distinguished from one another based on its sequence. To achieve species/strain-level resolution, the 16S rRNA gene may be replaced by the sequence of the internal transcribed spacer (Günther et al. 2006) or a housekeeping gene like the ribosomal protein S1-coding gene *rpsA* (Martens et al. 2008). The number of currently available rrs sequences nevertheless exceeds several times that for other genes (see e.g. Greengenes, http://greengenes.lbl.gov/cgi-bin/nph-index.cgi; Ribosome Database Project 10 (RDP 10), http://rdp.cme.msu.edu/; SILVA (http://www.arb-silva.de/), making rrs the number one choice for taxonomic probe definition in bacteria. Taxonomic microarrays available so far focus either on a particular taxonomic group of bacteria, e.g. *Bacillus* spp. (Liu et al. 2001), *Rhodocyclales* (Loy et al. 2005), *Alphaproteobacteria* (Sanguin et al. 2006a), or attempt to cover most bacterial phyla (Brodie et al. 2006, Kyselková et al. 2009).

Functional microarrays may be based either on one functional gene, e.g. *pmoA* coding for methane monooxygenase, and thus allowing to characterize the methanotrophic community (Stralis-Pavesse et al. 2004, Gebert et al. 2008), or several genes encompassing different functions. Rhee et al. (2004) for instance developed a functional microarray targeting genes involved in organic compound degradation and metal resistance, which is suitable for the assessment of polluted soils. The most comprehensive functional microarray developed so far is GeoChip of He et al. (2007), targeting genes involved in various biochemical and ecological processes such as carbon, nitrogen, phosphorus and sulfur cycling. Functional communities may be also assessed with *rrs* gene-based microarrays, provided that a function is carried out by distinctive bacterial taxonomic group(s), as for example in the case of nitrifiers (Kelly et al. 2005) or sulfate reducers (Loy et al. 2002).

Technical aspects

Probe definition. Definition of a probe for a desired taxonomic or functional group requires comparison of a large number of aligned sequences and is therefore rarely performed manually (Günther et al. 2006). Usually, probes are designed with a software, e.g. CommOligo (Li et al. 2005, He et al. 2007), PRIMEGENES (Xu et al. 2002, Rhee et al. 2004), ARB (Urakawa et al. 2002, Ludwig et al. 2004, Stralis-Pavese et al. 2004, Franke-Whittle et al. 2005) or CASCADE-P (DeSantis et al. 2003). For *rrs* gene-based taxonomic probes, a length around 20 nucleotides was shown a good compromise between probe specificity and sensitivity (Liu et al. 2001, Sanguin et al. 2006a). For functional gene-based probes, the length is usually 50–70 oligonucleotides (He et al. 2007), which is appropriate for transcriptional studies (Kane et al. 2000). When designing probes, one must be aware of some common pitfalls due to deficiencies in sequence databases (i.e. sequencing errors, wrong taxonomic affiliation of targets, under-representation of certain groups) or the targeted gene itself. In the case of *rrs*, this includes (in addition to its limited taxonomic resolution, as mentioned above) different copy numbers of ribosomal operons in bacterial genomes (up to 15, Klappenbach et al. 2001), *rrs* gene polymorphism within a single species (Marchandin et al. 2003, González-Escalona et al. 2005) and (rarely) horizontal transfer (Schouls et al. 2003), which may affect identification and abundance assessment based on *rrs* probes.

Microarray manufacturing. Microarrays developed so far vary in types of solid support and techniques for probe synthesis and attachment. Accordingly, they differ in (i) probe density, which may range from hundreds to hundreds of thousands probes per

square cm for 'high-density microarrays' (Lipshutz et al. 1999, Bodrossy and Sessitch 2004), as well as (ii) sensitivity and signal background (Lindroos et al. 2001, Hessner et al. 2004). The two main types of solid support are glass slides and gel pads.

The majority of microarrays used in microbial ecology are based on glass slides after surface modification with various functional groups allowing covalent binding of probes. Aldehyde, epoxysilan or 3D-Link (long hydrophilic chains with amine-reactive groups) surface groups bind covalently with amino-modified probes (Kane et al. 2000, Zammatteo et al. 2000, Lidroos et al. 2001). Good accessibility of the amino group for slide binding is ensured by adding a short carbon chain (usually C6) between the amino group and the 5' (or 3') end of the oligonucleotide (Franssen-val Hal et al. 2002). Some microarrays are based on peptidic bond formed between aminopropyl or aminophenyl group and succinate-modified probe (Dolan et al. 2001, Oh et al. 2002). Aminopropyl slides are suitable for long probes (about 50 nucleotides – usually on functional microarrays, He et al. 2007). Alternatively, probes may be bound non-covalently to positively charged surfaces by electrostatic adsorption (Belosludtsev et al. 2001). Lindroos et al. (2001) compared different probe immobilization techniques in terms of attachment efficiency and background signals. In their study, 3D-Link slides yielded high signals, but also high background. Aldehyde slides, when intact aldehyde groups were reduced with sodium borohydride after spotting, gave a good signal-to-noise ratio. Indeed, aldehyde slides are often used in microbial ecology studies (Loy et al. 2002, Stralis-Pavese et al. 2004, Franke-Whittle et al. 2005, Sanguin et al. 2006a).

The highest density of probes on glass microarrays is achieved when probes are synthesized *in situ* (i.e. directly on a slide). This is possible with (i) light-directed synthesis (photolithography, McGall et al. 1996), as used by Affymetrix (Lipshutz et al. 1999), or (ii) thermal inkjet printing, used by Agilent's Life Science (Fisher and Zhang 2007). Affymetrix technology was used to create the PhyloChip high density taxonomic microarray (Brodie et al. 2006).

Gel pad chips consist of thin polyacrylamide gel pads bound to the surface of glass slide. The gel contains aldehyde groups that bind covalently to the amino group of probe aminolinker (Fotin et al. 1998). This type of microarray was used mainly by the group of Kelly et al. (2005), for example for detection of nitrifying bacteria in wastewater.

Sample processing. Total DNA or RNA is extracted directly from environmental samples. This is usually followed by (RT) PCR amplification of the target(s). This step may already include sample labeling, when fluorescently-labeled primers (Franke-Whittle et al. 2005) or deoxyribonucleotides (Sanguin et al. 2006a) are used. Alternatively, the PCR step may be followed by incorporation of a fluorescent dye [most often Cy3 or Cy5 bound to one (deoxy) ribonucleotide], with the use of Klenow fragment (Loy et al. 2002) or in an additional *in vitro* transcription reaction (Bodrossy et al. 2003). The additional *in vitro* transcription brings several advantages, i.e. further amplification of target genes, efficient sample labeling and possibility of random chemical fragmentation of RNA, which leads to increased signal intensity (Bodrossy et al. 2003). PCR may affect relative abundances of targets and is difficult to implement when a high number of different functional genes are to be analyzed, but it can be avoided with the whole community RNA amplification technique developed by Gao et al. (2007). Whole community RNA amplification involves RNA extraction, its reverse transcription using random primers (adjusted with T7 RNA polymerase promoter sequence) and finally *in vitro* transcription of cDNA using labeled nucleotides. This method seems to preserve the original relative abundances of RNA targets.

Data processing. Even the data from carefully-performed microarray experiments may exhibit substantial variability, for example due to different level of fluorescent target hybridization, bad spot quality, or improper slides (Lee et al. 2000). To minimize the possibility of misinterpretation, microarray data should be filtered and normalized. Filtration is usually based on the comparison between probe signal and probe background intensity. Only probes whose signal significantly exceeds the background fluorescence level signal are considered positive (Peplies et al. 2004). Generally, the local background level is subtracted from probe signal (Sanguin et al. 2006a). Non-specific signal may also be estimated by using a mismatch probe in pair with each perfect-match probe. In that case, the resulting probe signal is computed by subtraction of mismatch probe signal from perfect-match probe signal (De Santis et al. 2005). Several repetitions of a probe per slide (or even better several hybridization repetitions) and statistical evaluation of repetitions are necessary to circumvent mistakes caused by local aberrancies (bad spots, impurities) on slides (Lee et al. 2000). To obtain normal distribution of signal intensities, microarray data are often transformed. The most common approaches for data transformation are square root and log2 transformations (DeSantis et al. 2005, Sanguin et al. 2006b).

Overall signal intensity on slide may vary among slides, e.g. due to different levels of fluorescent dye incorporation into target. This may be wrongly interpreted as difference in target abundance. Microarray

data normalization helps to overcome this problem. The simplest method of normalization is to divide probe signal by the signal of a positive control, which is expected to hybridize to all sequences. On *rrs*-based microarrays, the universal eubacterial probe EUB338 (Amann et al. 1990) is usually used as positive control (Loy et al. 2002, Bodrossy et al. 2003, Sanguin et al. 2006a). Sanguin et al. (2006b) showed that variability among technical repetitions was lower when probe signal was divided by the sum of all probe signals obtained per spotted motif rather than by the universal probe signal only. Alternatively, data can be normalized to the signal obtained with an internal standard added to each sample (Brodie et al. 2006).

Outputs from microarray analysis of microbial communities are large data sets often difficult to interpret. Output data matrices may contain tens to thousands variables (probes representing microbial groups or genes) and it is therefore tedious to employ classical univariate analysis like ANOVA (analysis of variance) for assessing treatment differences for each variable. In addition, biological data do not necessary fulfill all conditions required for formal application of such a parametric test.

Multivariate analyses are well adapted to facilitate interpretation of multivariate data and are used when assessing microbial communities with microarrays or fingerprinting methods (Ramette 2007). Based on data knowledge prior to analysis, these techniques are either exploratory (nothing is assumed about the samples, e.g. principal component analysis, correspondence analysis, non-metric multi-dimensional scaling, hierarchical clustering) or hypothesis-driven (some explanatory data are also considered in the analysis). The explanatory data may be either a factor (grouping of the samples into classes, e.g. between-group analysis, or analysis of similarities) or another table containing environmental data, e.g. chemical composition of samples (so-called 'two-table methods', e.g. redundancy analysis, canonical correspondence analysis). Hypotheses may be tested by permutation (Monte Carlo) test (Hope 1968).

Microarray specificity and sensitivity

Both probe specificity (distinguishing between targets and non-targets) and sensitivity (the smallest amount or relative abundance of sequence detected) depends on thermodynamics of probe-sequence binding. Ideally, target (matching perfectly with probe) is bound with much lower free energy than non-target (having some mismatches with probe), or in other words, melting temperature of matching duplex should be much higher than that of mismatch duplex. Algorithms available at present (e.g. nearest neighbor method, SantaLucia et al. 1996) allow to estimate melting temperature of duplexes in solution. However, they are not very reliable when designing microarray probes, since melting behavior of immobilized oligonucleotides differs from that in solution (Pozhitkov et al. 2006).

The effects of mismatch type and position (Pozhitkov et al. 2006, Zhang et al. 2006), type and length of probe spacer (Shchepinov et al. 1997, Peplies et al. 2004, Franke-Whittle et al. 2005), type and length of target (Kelly et al. 2005, Liu et al. 2007), and formamide concentration (Urakawa et al. 2002) on microarray duplex stability were assessed. For 20-mer oligonucleotides, Pozhitkov et al. (2006) compared the effect of all combinations of mismatch type × mismatch position on probe signal intensity for RNA/DNA hybridization. They reported that the best discrimination between target and non-target molecules was achieved with mismatches at the positions 5 to 13. Liu et al. (2007) found that optimal target length (for *rrs* genes) was between 20 and 100 nucleotides and longer targets should be therefore fragmented before hybridization. Long targets may form stable secondary structures within the molecule, which hinders probe binding (Southern et al. 1999). Secondary structures may result also in the opposite situation, i.e. false-positive result, when a hairpin formed within a target or probe molecule masks a mismatch, allowing the ends of molecules to form a stable duplex (Kyselková et al. 2008).

Optimal hybridization conditions (probe melting temperature, formamide concentration, etc.) may differ among probes spotted and this may pose a problem when designing a specific probe set (Loy et al. 2002). Several approaches were adopted to deal with this problem. First, probe-target hybridization behavior is carefully checked *in silico* and only probes meeting some narrow criteria are spotted (Bodrossy et al. 2003). Second, probe set specificity is tested a posteriori, by hybridization to pure strains/clones with known target gene sequence and non-specific probes are excluded (Loy et al. 2002, Sanguin et al. 2006a, b). Third, tetramethylammonium chloride (TMAC) or betaine added to hybridization reaction allows GC content-independent hybridization (Maskos and Southern 1992, Rees et al. 1993). Fourth, multiple probe subsets are designed for each taxon and a given taxon is considered present in a sample if the majority of probes in the subset are positive (Brodie et al. 2006).

Sensitivity of microarrays varies with microarray type and hybridization protocol. Different micro-

arrays are therefore difficult to compare, as their authors may have used different approaches for sensitivity assessment and result representation. Some spiked microorganisms to environmental sample (Franke-Whittle et al. 2005), others mixed known quantities of labeled PCR products before hybridization (Palmer et al. 2006, Sanguin et al. 2006a), and others used a complementary method for quantity estimation, e.g. Fluorescence *In Situ* Hybridization (FISH, Loy et al. 2005). When known quantities of targets are mixed before hybridization, targets with 0.03–5% relative abundance in a sample can usually be detected (Peplies et al. 2004, Sanguin et al. 2006b, Wagner et al. 2007). This is in line with the finding of Loy et al. (2005), who could detect bacteria whose relative abundance was less than 1% in activated sludge based on comparison with FISH. Franke-Whittle et al. (2005) could evidence 10^5 cells introduced into a complex community (compost). With a functional microarray (50-mer oligonucleotides), Rhee et al. (2004) could detect about 10^7 cells in the presence of background RNA.

Quantification with microarray

Changes in bacterial community composition in time occur typically because changes in environmental conditions have different consequences for the fate and abundance of different taxa. Quantitative community assessment may be therefore important to understand environmental processes. Different approaches were proposed for quantification with microarray, but only a restricted number of targets could be precisely quantified in these studies (Cho and Tiedje 2002, Palmer et al. 2006, Pozhitkov et al. 2007). Even if bias due to nucleic acid extraction and PCR amplification can be minimized, parallel microarray quantification of multiple taxa or functional groups in environmental samples is complicated by two facts:

First, a change of probe signal intensity may be due a change of target quantity as well as a change in number of weakly-binding mismatch targets (Zhou and Thompson 2002). The weak binding of mismatch sequences by probes was observed in several studies when pure strains were hybridized to microarray (Loy et al. 2002, Sanguin et al. 2006b, Kyselková et al. 2008). Whether this is also the case of complex environmental samples remains unknown. We suppose that, if perfectly matching targets are present in the sample, they probably outcompete mismatch targets from probe binding. In absence of perfect targets, weak signals may result in binding of mismatch sequences, but they should be much weaker than those obtained with pure cultures (as the mismatch target is diluted in the complex sample). To avoid this bias, the very weak non-specific hybridizations may be filtered with the use of negative controls, i.e. probes that theoretically should not hybridize with any *rrs* sequence, and whose possible weak signal is subtracted from each probe signal (Sanguin et al. 2006b).

Second, with the same quantity of matching targets, signal intensity is probe-dependent (Loy et al. 2002). This puts in doubt the possibility of absolute target quantification, regardless of the efficacy of an internal standard. For a given probe, however, the signal intensity was shown to be linearly dependent on target quantity (Rhee et al. 2004). In this context, the microarray is useful to assess the relative quantitative differences of a target group between samples (but not of different groups within a sample).

A promising alternative for parallel quantification of microorganisms in environmental samples is based on the use of padlock probes (van Doorn et al. 2007). With this approach, quantity is not inferred from hybridization signal but from TaqMan real-time amplification of ligated probes. A padlock probe is ligated (and may be therefore amplified) only when hybridized with a perfectly matching target from the sample. In addition to parallel quantification, this method has the advantage of distinguishing between perfect-match and one-mismatch targets.

Outlook

Because of the enormous diversity and complexity of microbial communities in the environment, research on microbial community composition has often been restricted to the description of a situation at one place, lacking a general theory (Prosser et al. 2007). When a theory is to be tested for its general applicability, high number of samples must be processed and this may be achieved with microarray analysis, allowing high-throughput assessment of taxonomic or functional diversity of microorganisms in the environment, as well as genotypic characterization of community members. Indeed, microarray analysis has proved effective to reach these goals with various types of microbial ecosystems, using microarrays focused on particular taxonomic groups or with a broad coverage of bacterial diversity (DeAngelis et al. 2009, Franke-Whittle et al. 2009, Sanguin et al. 2008, Zhou et al. 2008). Yet, bioinformatics and

biostatistics tools will have to be improved further to handle more effectively the huge amount of data produced, and to find significant features in noisy biological data.

In the context of microbial community analysis, current progress in sequencing technology (e.g. pyrosequencing, Roesch et al. 2007) is making this approach increasingly interesting to characterize community members. However, sequencing remains more useful to target dominant populations than to assess community members present in lower amounts. For example, to quantify a taxa with a relative abundance of 0.1% in the community (i.e. at a relative abundance where it can be commonly studied by microarray analysis), as many as 46 020 clones would need to be sequenced per sample replicate to expect (at $P = 0.01$) 10 sequences of this taxa (based on the equation of Taylor 2002). Indeed, microarray analysis was shown to reveal much more diversity than cloning/sequencing approach based on current sequencing efforts (DeSantis et al. 2007). However, sequencing is very useful in parallel to microarray approaches, both to guide design of new microarray probes and (when necessary) for a posteriori confirmation of particular microarray results.

A key challenge in microbial ecology is the relation between diversity and function. This may be resolved by coupling genotyping microarrays with other techniques, e.g. stable isotope probing. With this latter approach, microbial subcommunities that metabolized an isotope-labeled substrate may be revealed by comparative microarray analysis of heavy and light nucleic acids isolated from a sample (Cébron et al. 2007). This is one of several approaches (Loy et al. 2005) by which the high-throughput capacities of microarrays for analysis of microbial community composition can be combined with other methodologies for hypothesis testing and a better understanding of the ecological significance of the microbial world.

REFERENCES

Amann R.I., Binder B.J., Olson R.J., Chisholm S.W., Devereux R., Stahl D.A. (1990): Combination of 16S rRNA-targeted oligonucleotide probes with flow cytometry for analyzing mixed microbial populations. Applied and Environmental Microbiology, *56*: 1919–1925.

Belosludtsev Y., Iverson B., Lemeshko S., Eggers R., Wiese R., Lee S., Powdrill T., Hogan M. (2001): DNA microarrays based on noncovalent oligonucleotide attachment and hybridization in two dimensions. Analytical Biochemistry, *292*: 250–256.

Bodrossy L., Sessitsch A. (2004): Oligonucleotide microarrays in microbial diagnostics. Current Opinion in Microbiology, *7*: 245–254.

Bodrossy L., Stralis-Pavese N., Murrell J.C., Radajewski S., Weilharter A., Sessitsch A. (2003): Development and validation of a diagnostic microbial microarray for methanotrophs. Environmental Microbiology, *5*: 566–582.

Brodie E.L., DeSantis T.Z., Joyner D.C., Baek S.M., Larsen J.T., Andersen G.L., Hazen T.C., Richardson P.M., Herman D.J., Tokunaga T.K., Wan J.M.M., Firestone M.K. (2006): Application of a high-density oligonucleotide microarray approach to study bacterial population dynamics during uranium reduction and reoxidation. Applied and Environmental Microbiology, *72*: 6288–6298.

Cébron A., Bodrossy L., Chen Y., Singer A.C., Thompson I.P., Prosser J.I., Murrell J.C. (2007): Identity of active methanotrophs in landfill cover soil as revealed by DNA-stable isotope probing. FEMS Microbiology Ecology, *62*: 12–23.

Cho J.-C., Tiedje J.M. (2002): Quantitative detection of microbial genes by using DNA microarrays. Applied and Environmental Microbiology, *68*: 1425–1430.

DeAngelis K.M., Brodie E.L., DeSantis T.Z., Andersen G.L., Lindow S.E., Firestone M.K. (2009): Selective progressive response of soil microbial community to wild oat roots. ISME Journal, *3*: 168–178.

DeSantis T.Z., Brodie E.L., Moberg J.P., Zubieta I.X., Piceno Y.M., Andersen G.L. (2007): High-density universal 16S rRNA microarray analysis reveals broader diversity than typical clone library when sampling the environment. Microbial Ecology, *53*: 371–383.

DeSantis T.Z., Dubosarskiy I., Murray S.R., Andersen G.L. (2003): Comprehensive aligned sequence construction for automated design of effective probes (CASCADE-P) using 16S rDNA. Bioinformatics, *19*: 1461–1468.

DeSantis T.Z., Stone C.E., Murray S.R., Moberg J.P., Andersen G.L. (2005): Rapid quantification and taxonomic classification of environmental DNA from both prokaryotic and eukaryotic origins using a microarray. FEMS Microbiology Letters, *245*: 271–278.

Dolan P.L., Wu Y., Ista L.K., Metzenberg R.L., Nelson M.A., Lopez G.P. (2001): Robust and efficient synthetic method for forming DNA microarrays. Nucleic Acids Research, *29*: 107.

Fisher W., Zhang M.J. (2007): A biochip microarray fabrication system using inkjet technology. IEEE Transactions on Automation Science and Engineering, *4*: 488–500.

Flanagan J.L., Brodie E.L., Weng L., Lynch S.V., Garcia O., Brown R., Hugenholtz P., DeSantis T.Z., Andersen G.L., Wiener-Kronish J.P., Bristow J. (2007): Loss of bacterial diversity during antibiotic treatment of intubated patients colonized with *Pseudomonas aeruginosa*. Journal of Clinical Microbiology, *45*: 1954–1962.

Fotin A.V., Drobyshev A.L., Proudnikov D.Y., Perov A.N., Mirzabekov A.D. (1998): Parallel thermodynamic analysis of duplexes on oligodeoxyribonucleotide microchips. Nucleic Acids Research, *26*: 1515–1521.

Franke-Whittle I.H., Klammer S.H., Insam H. (2005): Design and application of an oligonucleotide microarray for the investigation of compost microbial communities. Journal of Microbiological Methods, *62*: 37–56.

Franke-Whittle I.H., Knapp B.A., Fuchs J., Kaufmann R., Insam H. (2009): Application of COMPOCHIP microarray to investigate the bacterial communities of different composts. Microbial Ecology, *57*: 510–521.

Franssen-val Hal N.L.W., Vorst O., Kramer E., Hall R.D., Keijer J. (2002): Factors influencing cDNA microarray hybridization on silylated glass slides. Analytical Biochemistry, *308*: 5–17.

Gao H., Yang Z.K., Gentry T.J., Wu L., Schadt C.W., Zhou J. (2007): Microarray-based analysis of microbial community RNAs by whole-community RNA amplification. Applied and Environmental Microbiology, *73*: 563–571.

Gebert J., Stralis-Pavese N., Alawi M., Bodrossy L. (2008): Analysis of methanotrophic communities in landfill biofilters using diagnostic microarray. Environmental Microbiology, *10*: 1175–1188.

González-Escalona N., Romero J., Espejo R.T. (2005): Polymorphism and gene conversion of the 16S rRNA genes in the multiple rRNA operons of *Vibrio parahaemolyticus*. FEMS Microbiology Letters, *246*: 213–219.

Günther S., Groth I., Grabley S., Munder T. (2006): Design and evaluation of an oligonucleotide-microarray for the detection of different species of the genus *Kitasatospora*. Journal of Microbiological Methods, *65*: 226–236.

He Z.L., Gentry T.J., Schadt C.W., Wu L.Y., Liebich J., Chong S.C., Huang Z.J., Wu W.M., Gu B.H., Jardine P., Criddle C., Zhou J. (2007): GeoChip: a comprehensive microarray for investigating biogeochemical, ecological and environmental processes. ISME Journal, *1*: 67–77.

Hessner M., Meyer L., Tacke, J., Muheisen S., Wang X. (2004): Immobilized probe and glass surface chemistry as variables in microarray fabrication. BMC Genomics, *5*: 53.

Hope A.C.A. (1968): A simplified Monte Carlo test procedure. Journal of the Royal Statistical Society Series B, *30*: 582–598.

Kane M.D., Jatkoe T.A., Stumpf C.R., Lu J., Thomas J.D., Madore S.J. (2000): Assessment of the sensitivity and specificity of oligonucleotide (50 mer) microarrays. Nucleic Acids Research, *28*: 4552–4557.

Kelly J.J., Siripong S., McCormack J., Janus L.R., Urakawa H., El Fantroussi S., Noble P.A., Sappelsa L., Rittmann B.E., Stahl D.A. (2005): DNA microarray detection of nitrifying bacterial 16S rRNA in wastewater treatment plant samples. Water Research, *39*: 3229–3238.

Klappenbach J.A., Saxman P.R., Cole J.R., Schmidt T.M. (2001): Rrndb: the ribosomal RNA operon copy number database. Nucleic Acids Research, *29*: 181–184.

Kyselková M., Kopecký J., Felföldi T., Čermák L., Omelka M., Grundmann G., Moënne-Loccoz Y., Ságová-Marečková M. (2008): Development of a 16S rRNA gene-based prototype microarray for the detection of selected actinomycetes genera. Antonie van Leeuwenhoek, *94*: 439–453.

Kyselková M., Kopecký J., Frapolli M., Défago G., Ságová-Marečková M., Grundmann G.L., Moënne-Loccoz Y. (2009): Comparison of rhizobacterial community composition in soil suppressive or conducive to tobacco black root rot disease. ISME Journal, doi:10.1038/ismej.2009.61

Lee M.-L.T., Kuo F.C., Whitmore G.A., Sklar J. (2000): Importance of replication in microarray gene expression studies: statistical methods and evidence from repetitive cDNA hybridizations. Proceedings of the National Academy of Sciences of the United States of America, *97*: 9834–9839.

Li X.Y., He Z.L., Zhou J.Z. (2005): Selection of optimal oligonucleotide probes for microarrays using multiple criteria, global alignment and parameter estimation. Nucleic Acids Research, *33*: 6114–6123.

Lindroos K., Liljedahl U., Raitio M., Syvanen A.-C. (2001): Minisequencing on oligonucleotide microarrays: comparison of immobilisation chemistries. Nucleic Acids Research, *29:* 69.

Lipshutz R.J., Fodor S.P.A., Gingeras T.R., Lockhart D.J. (1999): High density synthetic oligonucleotide arrays. Nature Genetics, *21*: 20–24.

Liu W.T., Guo H.L., Wu J.H. (2007): Effects of target length on the hybridization efficiency and specificity of rRNA-based oligonucleotide microarrays. Applied and Environmental Microbiology, *73*: 73–82.

Liu W.T., Mirzabekov A.D., Stahl D.A. (2001): Optimization of an oligonucleotide microchip for microbial identification studies: a non-equilibrium dissociation approach. Environmental Microbiology, *3*: 619–629.

Loy A., Lehner A., Lee N., Adamczyk J., Meier H., Ernst J., Schleifer K.H., Wagner M. (2002): Oligonucleotide microarray for 16S rRNA gene-based detection of all recognized lineages of sulfate-reducing prokaryotes in the environment. Applied and Environmental Microbiology, *68*: 5064–5081.

Loy A., Schulz C., Lucker S., Schopfer-Wendels A., Stoecker K., Baranyi C., Lehner A., Wagner M. (2005): 16S rRNA gene-based oligonucleotide microarray for environmental monitoring of the betaproteobacterial order *Rhodocyclales*. Applied and Environmental Microbiology, *71*: 1373–1386.

Ludwig W., Strunk O., Westram R., Richter L., Meier H., Yadhukumar, Buchner A., Lai T., Steppi S., Jobb G., Forster W., Brettske I., Gerber S., Ginhart A.W., Gross O., Grumann S., Hermann S., Jost R., Konig A., Liss T., Lussmann R., May M., Nonhoff B., Reichel B., Strehlow R., Stamatakis A., Stuckmann N., Vilbig A., Lenke M., Ludwig T., Bode A., Schleifer K.H. (2004): ARB: a software environment for sequence data. Nucleic Acids Research, *32*: 1363–1371.

Marchandin H., Teyssier C., Siméon de Buochberg M., Jean-Pierre H., Carriere C., Jumas-Bilak E. (2003): Intra-chromosomal heterogeneity between the four 16S rRNA gene copies in the genus *Veillonella*: implications for phylogeny and taxonomy. Microbiology, *149*: 1493–1501.

Martens M., Dawyndt P., Coopman R., Gillis M., De Vos P., Willems A. (2008): Advantages of multilocus sequence analysis for taxonomic studies: a case study using 10 housekeeping genes in the genus *Ensifer* (including former *Sinorhizobium*). International Journal of Systematic and Evolutionary Microbiology, *58*: 200–214.

Maskos U., Souther E.M. (1992): Parallel analysis of oligodeoxyribonucleotide (oligonucleotide) interactions. I. Analysis of factors influencing oligonucleotide duplex formation. Nucleic Acids Research, *20*: 1675–1678.

McGall G., Labadie J., Brock P., Wallraff G., Nguyen T., Hinsberg W. (1996): Light-directed synthesis of high-density oligonucleotide arrays using semiconductor photoresists. Proceedings of the National Academy of Sciences of the United States of America, *93*: 13555–13560.

Naum M., Brown E.W., Mason-Gamer R.J. (2008): Is 16S rDNA a reliable phylogenetic marker to characterize relationships below the family level in the Enterobacteriaceae? Journal of Molecular Evolution, *66*: 630–642.

Palmer C., Bik E.M., Eisen M.B., Eckburg P.B., Sana T.R., Wolber P.K., Relman D.A., Brown P.O. (2006): Rapid quantitative profiling of complex microbial populations. Nucleic Acids Research, *34*: 5.

Oh S.J., Cho S.J., Kim C.O., Park J.W. (2002): Characteristics of DNA microarrays fabricated on the various aminosilane layers. Langmuir, *18*: 1764–1769.

Peplies J., Lau S.C.K., Pernthaler J., Amann R., Glockner F.O. (2004): Application and validation of DNA microarrays for the 16S rRNA-based analysis of marine bacterioplankton. Environmental Microbiology, *6*: 638–645.

Pozhitkov A., Noble P.A., Domazet-Loso T., Nolte A.W., Sonnenberg R., Staehler P., Beier M., Tautz D. (2006): Tests of rRNA hybridization to microarrays suggest that hybridization characteristics of oligonucleotide probes for species discrimination cannot be predicted. Nucleic Acids Research, *34*: 66.

Pozhitkov A.E., Bailey K.D., Noble P.A. (2007): Development of a statistically robust quantification method for microorganisms in mixtures using oligonucleotide microarrays. Journal of Microbiological Methods, *70*: 292–300.

Prosser J.I., Bohannan B.J.M., Curtis T.P., Ellis R.J., Firestone M.K., Freckleton R.P., Green J.L., Green L.E., Killham K., Lennon J.J., Osborn A.M., Solan M., van der Gast C.J., Young J.P.W. (2007): The role of ecological theory in microbial ecology. Nature Reviews Microbiology, *5*: 384–392.

Ramette A. (2007): Multivariate analyses in microbial ecology. FEMS Microbiology Ecology, *62*: 142–160.

Rees W., Yager T., Korte J., Von Hippel P. (1993): Betaine can eliminate the base pair composition dependence of DNA melting. Biochemistry, *32*: 137–144.

Rhee S.-K., Liu X., Wu L., Chong S.C., Wan X., Zhou J. (2004): Detection of genes involved in biodegradation and biotransformation in microbial communities by using 50-mer oligonucleotide microarrays. Applied and Environmental Microbiology, *70*: 4303–4317.

Roesch L.F.W., Fulthorpe R.R., Riva A., Casella G., Hadwin A.K.M., Kent A.D., Daroub S.H., Camargo F.A.O., Farmerie W.G., Triplett E.W. (2007): Pyrosequencing enumerates and contrasts soil microbial diversity. ISME Journal, *1*: 283–290.

Sanguin H., Herrera A., Oger-Desfeux C., Dechesne A., Simonet P., Navarro E., Vogel T.M., Moënne-Loccoz Y., Nesme X., Grundmann G.L. (2006b): Development and validation of a prototype 16S rRNA-based taxonomic microarray for *Alphaproteobacteria*. Environmental Microbiology, *8*: 289–307.

Sanguin H., Kroneisen L., Gazengel K., Kyselková M., Remenant B., Prigent-Combaret C., Grundmann G.L., Sarniguet A., Moënne-Loccoz Y. (2008): Development of a 16S rRNA microarray approach for the monitoring of rhizosphere *Pseudomonas* populations associated with the decline of take-all disease of wheat. Soil Biology and Biochemistry, *40*: 1028–1039.

Sanguin H., Remenant B., Dechesne A., Thioulouse J., Vogel T.M., Nesme X., Moënne-Loccoz Y., Grundmann G.L. (2006a): Potential of a 16S rRNA-based taxonomic microarray for analyzing the rhizosphere effects of maize on *Agrobacterium* spp. and bacterial communities. Applied and Environmental Microbiology, *72*: 4302–4312.

SantaLucia J., Allawi H.T., Seneviratne A. (1996): Improved nearest-neighbor parameters for predicting DNA duplex stability. Biochemistry, *35*: 3555–3562.

Schena M., Shalon D., Davis R.W., Brown P.O. (1995): Quantitative monitoring of gene-expression patterns with a complementary-DNA microarray. Science, *270*: 467–470.

Schouls L.M., Schot C.S., Jacobs J.A. (2003): Horizontal transfer of segments of the 16S rRNA genes between species of the *Streptococcus anginosus* group. Journal of Bacteriology, *185*: 7241–7246.

Shchepinov M.S., Case-Green S.C., Southern E.M. (1997): Steric factors influencing hybridisation of nucleic acids to oligonucleotide arrays. Nucleic Acids Research, *25*: 1155–1161.

Southern E., Mir K., Shchepinov M. (1999): Molecular interactions on microarrays. Nature Genetics Supplement, *21*: 5–9.

Stralis-Pavese N., Sessitsch A., Weilharter A., Reichenauer T., Riesing J., Csontos J., Murrell, J.C., Bodrossy L. (2004): Optimization of diagnostic microarray for application in analysing landfill methanotroph communities under different plant covers. Environmental Microbiology, *6*: 347–363.

Taylor A.F.S. (2002): Fungal diversity in ectomycorrhizal communities: sampling effort and species detection. Plant and Soil, *244*: 19–28.

Urakawa H., Noble P.A., El Fantroussi S., Kelly J.J., Stahl D.A. (2002): Single-base-pair discrimination of terminal mismatches by using oligonucleotide microarrays and neural network analyses. Applied and Environmental Microbiology, *68*: 235–244.

van Doorn R., Szemes M., Bonants P., Kowalchuk G.A., Salles J.F., Ortenberg E., Schoen C.D. (2007): Quantitative multiplex detection of plant pathogens using a novel ligation probe-based system coupled with universal, high-throughput real-time PCR on OpenArrays™. BMC Genomics, *8:* 276.

Wagner M., Smidt H., Loy A., Zhou J.Z. (2007): Unravelling microbial communities with DNA-microarrays: challenges and future directions. Microbial Ecology, *53*: 498–506.

Wu L., Liu X., Schadt C.W., Zhou J. (2006): Microarray-based analysis of subnanogram quantities of microbial community DNAs by using whole-community genome amplification. Applied and Environmental Microbiology, *72*: 4931–4941.

Xu D., Li G., Wu L., Zhou J., Xu Y. (2002): PRIMEGENS: robust and efficient design of gene-specific probes for microarray analysis. Bioinformatics, *18*: 1432–1437.

Zammatteo N., Jeanmart L., Hamels S., Courtois S., Louette P., Hevesi L., Remacle J. (2000): Comparison between different strategies of covalent attachment of DNA to glass surfaces to build DNA microarrays. Analytical Biochemistry, *280*: 143–150.

Zhang L., Chunlei W., Carta R., Zhao H. (2006): Free energy of DNA duplex formation on short oligonucleotide microarrays. Nucleic Acids Research, *35*: 18.

Zhou J. (2003): Microarrays for bacterial detection and microbial community analysis. Current Opinion in Microbiology, *6*: 288–294.

Zhou J., Kang S., Schadt C.W., Garten C.T. (2008): Spatial scaling of functional gene diversity across various microbial taxa. Proceedings of the National Academy of Sciences of the United States of America, *105*: 7768–7773.

Zhou J., Thompson D.K. (2002): Challenges in applying microarrays to environmental studies. Current Opinion in Biotechnology, *13*: 204–207.

Received on July 1, 2009

Corresponding author:

Prof. Yvan Moënne-Loccoz, Université Lyon 1, UMR CNRS 5557 Ecologie microbienne,
43 Bd. du 11 Novembre 1918, 69622 Villeurbanne, France
e-mail: moenne@biomserv.univ-lyonl.fr