White Paper



Microbial identification using Axiom[®] Microbiome Array versus 16S rRNA gene sequencing technology

Abstract

Over the past decade, massively parallel high-throughput, short-read 16S ribosomal RNA (rRNA) gene sequencing has replaced traditional long-read Sanger sequencing for the identification of bacteria within a community.¹ The transition to this next-generation, targeted amplicon approach has provided a more efficient and cost-effective method to analyze microbial samples; however, it has come at the expense of taxonomic resolution and is limited to detection of bacteria through the use of conserved 16S rRNA gene primers, which do not share homology with other members of microbial populations such as fungi, viruses, and protists. While alternative approaches, such as whole-genome metagenomic sequencing, are also being used to profile microbial communities, these methods can require deep read depths to accurately infer taxonomic origins and can be cost-prohibitive on a per sample basis. Here we present a novel method of microbial analysis based on the Axiom[®] genotyping platform from Affymetrix, that enables detection of all organisms in a sample including bacteria, virus, protozoa, fungi, and archaea down to species- and often strain-level resolution within a single reaction. We show that Axiom Microbiome Array provides increased accuracy and resolution while leveraging a streamlined protocol with simple analysis software.

Results

Axiom Microbiome Array was designed to interrogate microbial entities from all kingdoms or domains of life. Axiom Microbiome Array leverages Axiom[®] 2.0 assay biochemistry, which includes an isothermal whole-genome amplification (WGA) step. DNA extracted from microbes can be used as the direct input for the initial step of the workflow. To access the subset of viral genomes comprised of RNA, extracted RNA can be reverse transcribed, and the resulting cDNA can be used as a template for the Axiom[®] assay workflow. Axiom Microbiome Array data is analyzed through a software interface called Axiom[™] <u>Mi</u>crobial <u>D</u>etection <u>A</u>nalysis <u>S</u>oftware (MiDAS). This software utilizes the Composite Likelihood Maximization (CLiMax) algorithm developed at Lawrence Livermore National Laboratories.^{2,3} Axiom MiDAS converts probe-level hybridization intensities to a two-state model (on/off) by thresholding the probe intensities at the 99th percentile of background anti-genomic, non-targeting probes also present on the array. This algorithm then uses this information to build a predictive model of the sample that best explains the probe-level data. This iterative 'greedy' process adds targets to the description of the sample and culminates in a list of detected organisms. The initial list of detected targets is further refined by implementing an empirically determined filtering parameter based on the ratio of the conditional log likelihood divided by the initial log likelihood. A threshold value of 0.2 for this quotient is recommended in order to remove targets with minimal contribution to the predictive model. Axiom MiDAS generated results in the studies described below all use this final threshold value.

The WGA step does not specifically target microbial DNA, in contrast to the bacterial or fungal specificity of 16S or 18S rRNA PCR-based approaches, respectively. Thus, host genomic DNA (gDNA) present in a sample will be amplified along with microbial DNA since there is no microbial-specific enrichment *per se*. To assess the limit of detection (LOD) of Axiom Microbiome Array, three bacterial genomes were selected to be used as input for Axiom 2.0 assay target preparation at input doses ranging from 1 to 1 million genome equivalents (Figure 1). Bacterial genomes were selected to represent distinct families and span a range of both genome size and GC content (*Bacillus cereus:* 5.4 Mb, 35% GC; *Thermotoga maritima*: 1.9 Mb, 46% GC; *Burkholderia thailandensis*: 6.7 Mb, 67% GC) to control for any potential effect these factors may have on the Axiom 2.0 assay biochemistry. The sensitivity of the assay was assessed across a range of background human gDNA concentrations as microbial detection applications can include the enumeration of organisms derived from different body sites which can have a wide range of host background DNA.⁴

Figure 1: Limit of Detection (LOD) on Axiom[®] Microbiome Array. For each of the indicated bacterial genomic DNA (gDNA), 1 to 1 million copies were used as input into Axiom[®] 2.0 assay target preparation, either in the presence or absence of the indicated amounts of HapMap human gDNA (NA18501). "+" indicates input bacterial gDNA was detected using Axiom[™] MiDAS. "-" indicates input bacterial gDNA was not detected using Axiom[™] MiDAS.

			No ba	ckground	DNA		
	1,000,000	100,000	10,000	1,000	100	10	1
Bacillus cereus	+	+	+	+	+*	-	-
Burkholderia thailandensis	+	+	+	+**	-	-	-
Thermotoga maritima	+	+	+	+	+***	_	_
				ng NA850			
	1,000,000	100,000	10,000	1,000	100	10	1
Bacillus cereus	+	+	+	+	+*	-	-
Burkholderia thailandensis	+	+	+	+	-	-	-
Thermotoga maritima	+	+	+	+	+***	-	-
) ng NA850			
	1,000,000	100,000	10,000	1,000	100	10	1
Bacillus cereus	+	+	+	+	-	-	-
Burkholderia thailandensis	+	+	+	+	-	-	-
Thermotoga maritima	+	+	+	+	_	-	_
	100 ng NA8501						
	1,000,000	100,000	10,000	1,000	100	10	1
Bacillus cereus	+	+	+	-	-	-	-
Burkholderia thailandensis	+	+	+	-	_	-	-
Thermotoga maritima	+	+	+	-	-	-	-

*Only plasmid pBClin15 was detected.

**Correct species, alternate strain (other than E264) was detected.

***Thermotoga sp. strain RQ2 was detected.

The microbial gDNA for these LOD experiments was procured from ATCC. Lyophilized DNA was resuspended in low-EDTA TE buffer (10 mM Tris Ultrapure, 0.1 mM EDTA, pH 8.0) and concentrations were determined using Quant-iT[™] PicoGreen[®] dsDNA Assay Kit (Thermo Fisher Scientific). To calculate the mass necessary for each genome equivalent, the mass of chromosomal DNA was added to the weight of one plasmid copy. The human gDNA sample used as background was HapMap NA18501 (Coriell Institute for Medical Research).

The results indicate that Axiom Microbiome Array was able to identify both the correct species and strain of interest when 10,000 genomic equivalents (~20–70 pg DNA depending on the strain tested) were present along with the highest (100 ng) amount of background human DNA in the sample (Figure 1), equating to the detection of an organism present at less than 0.1% of the starting material (<100 pg/100 ng). When 1,000 (~2–7 pg DNA depending on the strain tested) bacterial genome copies were spiked into 10 ng or less of NA18501 background DNA, the correct species for all genomes were identified. When 100 genome copies were in the background of either 1 ng or no human DNA, a prediction of a target within the same genus for *Thermatoga maritima* (*Thermotoga sp. RQ2*) and the pBClin15 plasmid was made, consistent with the *Bacillus cereus* strain tested. The most likely explanation for the lack of identification of the chromosomal target for *Bacillus cereus* is that plasmids can exist in multi-copy numbers for some genomes of interest such that the plasmid copy number may be higher than the indicated 100 genome copies.

Following LOD studies, Axiom Microbiome Array was challenged with defined samples of increasing biological complexity, and results were compared to the output from an alternative technology often used for bacterial community profiling, namely 16S rRNA gene sequencing. While the description of complex biological-derived communities from *bona fide* diverse sample types, such as gut and oral cavities, is of intense interest, we first sought to benchmark the comparative performance of Axiom Microbiome Array to 16S sequencing by analyzing artificial, mock-defined samples, where *a priori* knowledge of ground-level truth is known, in the presence of NA18501 background DNA.

16S rRNA gene sequencing is a powerful technique that has provided the foundation for the understanding of many aspects of microbial communities such as taxonomic structure and functional characteristics. The ubiquitous use of this technique and its variations along with the concomitant wealth of data generated underscore its importance and usefulness in many microbiome applications. Briefly, this technique utilizes the conserved nature of the 16S rRNA gene (*rrnA*), which is composed of both conserved and variable regions. PCR amplification using primers targeted towards the conserved regions generates amplicons spanning hypervariable regions, enabling discrimination of bacteria in a complex mixture after sequence alignment algorithms are used. Despite the power of this technique, one key limitation of 16S rRNA gene sequencing is the inability to describe all viral, protozoan, and fungal components of a sample of interest in a single assay unlike Axiom Microbiome Solution.

As a test case we chose to interrogate sample HM-276D (Microbial Mock Community B (Even, High Concentration), v5.1H, for Whole Genome Shotgun Sequencing; BEI Resources), which has been used in previous studies to assess 16S rRNA and whole-genome shotgun sequencing technologies.^{5,6} HM-276D is a mixture of gDNA from 20 bacterial strains containing equimolar rRNA operon counts. For this sample, 16S rRNA sequencing was done using 16S MetaVx[™] Mammalian Sequencing of V3 and V4 hypervariable regions carried out at GeneWiz[®] Services (2 x 250 bp sequencing on Illumina[®] MiSeq[®] System).^{7,8} This sample yielded 1.3 million reads for data analysis. Data processing and taxonomic assignments were performed using the Illumina 16S Metagenomics BaseSpace App. In the absence of any PCR bias, 16S rRNA sequence reads should be evenly distributed among the community members since known variability in 16S gene copy number between different species has been normalized.⁹

HM-276D was tested in duplicate on Axiom Microbiome Array using 100 ng input and analyzed with Axiom MiDAS. The results were compared to 16S sequencing results (Figure 2). Family-level read assignments show that all families in the sample were represented; however, deviations from the expected counts were observed. All families except *Streptococcaceae* and *Staphylococcaceae* should comprise 5% of the total reads. *Streptococcaceae* should be represented by 15%, as it has three species in the sample, and *Staphylococcaceae* should be represented by 10%, with two species present. The relative abundance at the family-level assignment deviated more than 10-fold for only one interrogated family (*Listeriaceae* at 0.4%).

Inpu	Axiom™ MiDAS Target description				16S V3 and V4 rRNA sequencing				
						% Re	% Reads – Taxon-level		
Family	Species	Family	Genus	Species	Strain	Family	Genus	Species	
Moraxellaceae	Acinetobacter baumannii	+	+	+	+	10.7	10.8	9.7	
Actinomycetaceae	Actinomyces odontolyticus	+	+	+		2.3	2.3	2.0	
Bacillaceae	Bacillus cereus	+	+	+	+	6.2	6.2	*	
Bacteroidaceae	Bacteroides vulgatus	+	+	+	+	9.5	9.5	8.1	
Clostridiaceae	Clostridium beijerinckii	+	+			12.4	12.4	0.1	
Deinococcaceae	Deinococcus radiodurans	+	+	+	+	1.3	1.3	1.3	
Enterococcaceae	Enterococcus faecalis	+	+	+	+	3.9	3.9	2.4	
Enterobacteriaceae	Escherichia coli	+	+	+	+	7.1	7.0	0.02*	
Helicobacteraceae	Helicobacter pylori	+	+	+	+	6.5	6.5	6.3	
Lactobacillaceae	Lactobacillus gasseri	+	+	+	+	5.2	5.2	4.3	
Listeriaceae	Listeria monocytogenes	+	+	+	+	0.4	0.4		
Neisseriaceae	Neisseria meningitidis	+	+	+		7.3	7.3	6.6	
Propionibacteriaceae	Propionibacterium acnes	+	+	+	+	1.1	1.1	0.8	
Pseudomonadaceae	Pseudomonas aeruginosa	+	+	+	+	0.7	0.7	1.0	
Rhodobacteraceae	Rhodobacter sphaeroides	+	+	+	+	2.3	2.3	2.2	
Staphylococcaceae	Staphylococcus aureus	+	+	+		10.3	10.3	4.0*	
Staphylococcaceae	Staphylococcus epidermidis	+	+	+		10.3	10.3	0.1*	
Streptococcaceae	Streptococcus agalactiae	+	+	+	+	12.6	12.6	6.1	
Streptococcaceae	Streptococcus mutans	+	+	+	+	12.6	12.6	4.6	
Streptococcaceae	Streptococcus pneumoniae	+	+	+	+	12.6	12.6	1.30	

Figure 2: Comparison of Axiom[®] Microbiome Array results to 16S rRNA sequencing read assignments for

HM-276D. A complex sample of 20 bacterial species normalized to 16S rRNA copy number was interrogated by both Axiom[®] Microbiome Array and 16S rRNA sequencing (MetaVX[™] Mammalian, V3-V4; GeneWiz[®] Services). Results were compared with respect to the most discrete taxonomic discrimination that could be assigned.

*Indicates that the most prevalent call for this species was a genus-level assignment by Illumina 16S Metagenomics App

While 16S rRNA sequencing data is more commonly mined at higher taxonomic levels, such as family-level calls, in order to profile within and between community diversity and assess general microbial community trends, we investigated the ability to make species-level determinations and compared the results to those from Axiom MiDAS. If both a chromosomal- and plasmid-level determination was made for the same species, it was condensed into a single species-level assignment.

Axiom MiDAS made the correct species-level determination for 19 out of the 20 input species when challenged with the HM-276D set. For 16 of these, it was able to make a target-level prediction consistent with the correct strain used as the actual input. The species for which the software was unable to make the correct determination was *Clostridium beijerinckii*. For this strain, the species determination made by the software was either *Clostridium diolis* or *Clostridium pasteurianum*.

While 16S rRNA sequencing is not able to make strain-level descriptions, it was able to make species-level assignments for 18 of the 20 species in the sample. However, for a number of these, the most prevalent species detected from the input was not the species of interest. For example, while *Clostridium beijerinckii* was the species interrogated, the most prevalent species-level assignment in this genus was *Clostridium saccharoperbutylacetonicum*, which accounted for approximately 148,000 reads in the sample and represents 800-fold more reads than seen for *Clostridium beijerinckii*.

Since the actual composition of sample HM-276D is known, we assessed the true positive rate (TPR) as a measure of sensitivity and the positive predictive value (PPV) as a measure of the assay specificity, and compared these metrics to 16S rRNA sequencing species-level performance (Figure 3). The TPR and PPV are defined as follows:

 $TPR = \frac{True Positives}{(True Positives + False Negatives)}$

 $PPV = \frac{True Positives}{(True Positives + False Positives)}$

Figure 3: Comparison of true positive rate (TPR) and positive predictive value (PPV) from Axiom[™] MiDAS and 16S rRNA sequencing. TPR and PPV were calculated using both Axiom[™] MiDAS assignments of Axiom[®] Microbiome Array data using standard default filtering of data and unfiltered and filtered 16S rRNA sequencing reads at different levels of taxonomic resolution.

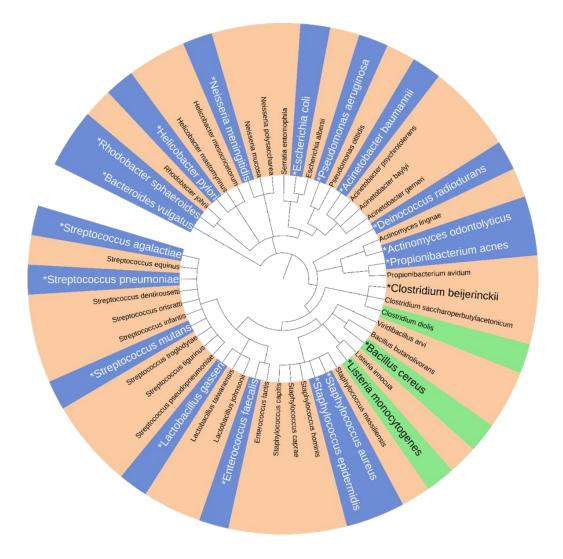
Axiom™ MiDAS TPR and PPV									
Family-level									
Condition True False False TPR PP									
HM-276D – Replicate 1	17	0	0	1	1				
HM-276D – Replicate 2	17	0	0	1	1				
Genus-level									
Condition	TPR	PPV							
HM-276D – Replicate 1	17	1	0	1	1				
HM-276D – Replicate 2	17 0 0		0	1	1				
	Spe	cies-level							
Condition	True positives	False positives	False negatives	TPR	PPV				
HM-276D – Replicate 1	19	1	1	0.95	0.95				
HM-276D – Replicate 2	19	1	1	0.95	0.95				
Strain-level									
Condition	Condition True False False TPR PPV								
HM-276D – Replicate 1	16	7	4	0.8	0.696				
HM-276D – Replicate 2	16	5	4	0.8	0.762				

16S rRNA sequencing V3 and V4 TPR and PPV									
	Family-le	evel							
Condition	True positives	False positives	False negatives	TPR	PPV				
HM-276D, unfiltered	17	115	0	1	0.129				
HM-276D, filtered to assignements with ≥ 5836 reads (99.9% of reads)	17	0	0	1	1				
Genus-level									
Condition True False False TPR positives positives negatives									
HM-276D, unfiltered	17	217	1	1	0.073				
HM-276D, filtered to assignements with ≥ 5829 reads (99.8% of reads)	17	0	1	1	1				
	Species-	level							
Condition	True positives	False positives	False negatives	TPR	PPV				
HM-276D, unfiltered	18	374	2	0.9	0.05				

Axiom MiDAS and 16S rRNA-derived species-level assignments were analyzed with these two metrics. The 16S rRNA sequencing results were parsed as a function of read depth using two different conditions. In the first condition, all of the 16S rRNA reads were considered in the calculations. The second condition maximized the PPV calculation for 16S rRNA sequencing by filtering the data to only species-level assignments that were predicted by \geq 186 reads. This threshold was the lowest read count that yielded a true positive call when compared to known truth. The comparison of species calls between Axiom MiDAS and 16S rRNA sequencing results filtered for \geq 186 reads is rendered in Figure 4 with Interactive Tree of Life (iTOL)¹⁰ tools using taxonomic trees generated with PhyloT (http://phylot.biobyte.de/).

It is important to note that these comparisons are not static and can be affected by the use of different reference databases and analysis pipelines. In this analysis, both technologies performed well at capturing the expected species, and generally reflect similar proportions of species calls within genera. For example, the majority of distinct species calls fall within the *Streptococcus* and *Staphylococcus* genera for both Axiom MiDAS and 16S rRNA sequencing results, indicating both technologies give similar relative genus-level landscapes. In contrast, the number of unique species calls between Axiom MiDAS and 16S rRNA sequencing diverges to 3 and 32, respectively. These values, in addition to correct taxonomic assignments, were used in the TPR and PPV species-level assessments.

Figure 4: Species taxonomic tree comparing Axiom™ MiDAS and 16S rRNA sequencing results. Species calls associated with the sample input are denoted with an asterisk (*). Blue represents Axiom™ MiDAS and 16S rRNA sequencing matched calls, green represents Axiom™ MiDAS only calls, and tan represents 16S rRNA sequencing only calls.



Another feature distinguishing Axiom Microbiome Array from 16S rRNA sequencing is the ability to make plasmidlevel target identification. For example, one component of the HM-276D panel is the *Bacillus cereus* strain NRS 248 (which corresponds to ATCC 10987). This ability to identify extra-chromosomal DNA is a distinct advantage as plasmids are often the DNA moiety harboring virulence elements or toxins, as is the case with the pXO1 and pXO2 plasmids in pathogenic *Bacillus anthracis* strains.¹¹ Axiom MiDAS was able to make target-level assessments consistent with this strain input. Alternatively, the 16S rRNA sequencing methods and analysis used were able to make family- and genus-level determinations consistent with this input, but were not able to identify the correct species (Figure 5).

Figure 5: Comparison of 16S rRNA sequencing and Axiom[™] MiDAS descriptions from *Bacillus cereus* strain included in HM-276D. For *Bacillus cereus* strain NRS 248 (ATCC 10987) 16S sequencing makes the correct family-level and genus-level determination, while Axiom[®] Microbiome Array provides target-level information at both the chromosome and plasmid level consistent with the input strain.

Family	Species	% Reads Family-level	% Reads Genus-level	% Reads Species-level	Axiom™ MiDAS Target-level descriptions
Bacilliaceae	Bacillus cereus	6.2	6.2	None	<i>Bacillus cereus</i> ATCC 10987
					<i>Bacillus cereus</i> ATCC 10987 plasmid pBc10987

A key goal of Axiom Microbiome Array was to develop a platform agnostic of sample type and composition, meaning that the composition of microbes in any sample could be interrogated using the same platform with no modifications to the sample preparation workflow. To this end, single genomes from different kingdoms represented on the array were tested, including that of a negative sense single-stranded (ss) RNA virus. Because the isothermal WGA step in the Axiom 2.0 assay workflow uses a DNA template, a reverse transcription (RT) reaction is required to generate cDNA as input to interrogate RNA genomes. The only modification from the standard RT assay is a two-hour rather than one-hour incubation at 42°C using Superscript[®] VILO (Thermo Fisher Scientific). Next, 17.5 µL of a standard 20 µL negative single-stranded RNA virus reaction was added to 2.5 µL of low-EDTA TE as an alternative input to gDNA for this specific application.

Figure 6 summarizes the Axiom MiDAS results of other single genomes tested with Axiom Microbiome Array. For these, 50 ng input was used for the following samples: *Methanobrevibacter smithii* (ATCC, 35061D-5), *Candida albicans* (ATCC, MYA2876D-5), and *Toxoplasma gondii* (BEI Resources, NR-33510). In addition, cDNA derived from 500 pg of RNA from *Respiratory syncytial virus A* (BEI Resources, NIAID, NIH: Genomic RNA from *Human Respiratory Syncytial Virus*, A1997/12-35, NR-43976) was also tested as input to the Axiom 2.0 assay workflow.

Figure 6: Summary of Axiom[™] MiDAS output from single genome samples of non-bacterial kingdoms represented on Axiom[®] Microbiome Array. Summary data from Axiom[™] MiDAS output from samples derived from the indicated species and source. Indicated are the number of Axiom[™] MiDAS algorithm iterations run; the initial and conditional log composite likelihood scores used to determine the species identified; and the expected and detected probes. The family, species, and target descriptions generated by Axiom[™] MiDAS are also displayed.

Archaebacteria – <i>Methanobrevibacter smithii</i> , ATCC 35061D-5										
Axiom [™] MiDAS iterations	Initial score	Conditional score	Expected probes	Detected probes	Conditional/ Initial	Family	Species	Target description		
1	434.36	434.36	168	123	1	Methanobacteriaceae	Methanobrevibacter smithii	Methanobrevibacter smithii ATCC 35061 chromosome		
	Fungi – Candida albicans, ATCC MYA2876D-5									
Axiom™ MiDAS iterations	Initial score	Conditional score	Expected probes	Detected probes	Conditional/ Initial	Family	Species	Target description		
1	455.82	455.82	159	124	1	Debaryomycetaceae	Candida albicans	Candida albicans A20, combined chrs		
2	420.21	8.67	154	116	0.02	Debaryomycetaceae	Candida albicans	Candida albicans A155, combined chrs		
	Protozoa – <i>Toxoplasma gondii,</i> BEI NR-33510									
Axiom™ MiDAS iterations	Initial score	Conditional score	Expected probes	Detected probes	Conditional/ Initial	Family	Species	Target description		
1	403.53	403.53	176	116	1	Sarcocystidae	Toxoplasma gondii	Toxoplasma gondii p89, combined chrs		
	Viral – Respiratory Syncitial Virus A, BEI NR-43976									
Axiom™ MiDAS iterations	Initial score	Conditional score	Expected probes	Detected probes	Conditional/ Initial	Family	Species	Target description		
1	344.7	344.7	202	110	1	Paramyxoviridae	Human respiratory syncytial virus	Human respiratory syncytial virus strain RSVA/Homo sapiens/USA/629- 4360/1998		
2	338.64	68.13	141	97	0.2	Paramyxoviridae	Human respiratory syncytial virus	Human respiratory syncytial virus strain RSVA/human/USA/A1997-12- 35/1997		

For these single genome samples, Axiom MiDAS provides target identification in the primary iteration that is consistent with all of the species tested. For two of the samples (*Candida albicans* and *Human respiratory syncytial virus A*), a subsequent iteration of the software identifies a target sequence that disagrees with the primary iteration at the strain level. When considering the results from later iterations of the algorithm, it is important to note that they have low conditional scores and therefore are contributing far less in terms of log likelihood for the overall sum of the predictions as compared to those of previous iterations.

Conclusion

Axiom Microbiome Array enables enumeration of organisms from all kingdoms or domains of life within a single assay by leveraging proven Axiom[®] technology. Analysis with user-friendly Axiom MiDAS follows a streamlined workflow that results in higher accuracy at the species level than 16S rRNA gene sequencing technology and provides additional information such as the presence of plasmids. While 16S rRNA sequencing can offer quantitative results, the method often suffers from a low positive identification value caused by a large number of false positives as well as missed calls. We suggest that for many microbiome-related applications the increased accuracy and robust workflow of Axiom Microbiome Solution represent a viable alternative to current sequencing-based methodologies.

References

- 1. Singer E., *et al*. High-resolution phylogenetic microbial community profiling. *The ISME Journal* (2016). doi:10.1038/ismej.2015.249.
- Gardner S. N., *et al.* A microbial detection array (MDA) for viral and bacterial detection. *BMC Genomics* **11**(668) (2010).
- 3. McLoughlin K. S., Microarrays for pathogen detection and analysis. *Briefings in Functional Genomics* **10**(6):342–353 (2011).
- 4. The Human Microbiome Project Consortium, A framework for human microbiome research. *Nature* **486**(7402):215–221 (2012).
- 5. Nelson M. C., *et al.* Analysis, optimization and verification of Illumina-generated 16S rRNA gene amplicon surveys. *PLoS One* **9**(4):e94249 (2014).
- Jones M. B., *et al.* Library preparation methodology can influence genomic and functional predictions in human microbiome research. *Proceedings of the National Academy of Sciences of the United States of America*. **112**(45):14024–14029 (2015).
- 7. Fu S. F., *et al*. The chemical properties and microbial community characterization of the thermophilic microaerobic pretreatment process. *Bioresource Technology* 198:497–502 (2015).
- You J., et al. Microbial community dynamics in Baolige oilfield during MEOR treatment, revealed by Illumina MiSeq sequencing. Applied Microbiology and Biotechnology 100(3):1469–1478 (2015).
- 9. Větrovský T., Baldrian P. The variability of the 16S rRNA gene in bacterial genomes and its consequences for bacterial community analyses. *PLoS One* **8**(2):e57923 (2013).
- 10. Letunic I., Bork P. Interactive tree of life (iTOL) v3: an online tool for the display and annotation of phylogenetic and other trees. *Nucleic Acids Research* (2016). doi:10.1093/nar/gkw290
- 11. Pilo P., Frey J. *Bacillus anthracis*: molecular taxonomy, population genetics, phylogeny and patho-evolution. *Infection, Genetics and Evolution* **11**(6):1218–1224 (2011).

Affymetrix, Inc: (US) +1-888-362-2447, +1-408-731-5000 ■ (EU) +44-(0)1628-552550 ■ (JP) +81-(0)3-6430-4020 ■ (CN) +86-21-63915511 eBioscience Products: (US) +1-888-999-1371, +1-858-642-2058 ■ (EU) +43 1 796 40 40 305 ■ (JP) +81-(0)3-6430-4020 USB Products: (US) +1-800-321-9322, +1-216-765-5000 ■ (EU) +44-(0)1628-552600

www.affymetrix.com Please visit our website for international distributor contact information.

For Research Use Only. Not for use in diagnostic procedures.

P/N GCL05561 Rev. 2

© 2016 Affymetrix, Inc. All rights reserved. Affymetrix[®], Axiom[®], GeneChip[®], Command Console[®], CytoScan[®], DMET[™], Eureka [™], Eureka Genomics[®], Eureka Genotyping[™], Expression Console[™] GeneAtlas[®], GeneChip-compatible[™], GeneTitan[®], Genotyping Console[™], myDesign[™], MyGeneChip[™], NetAffx[®], OncoScan[®], Powered by Affymetrix[™], PrimeView[®], and ViewRNA[®] are trademarks or registered trademarks of Affymetrix, Inc. Please see affymetrix.com/trademarks for a complete list of Affymetrix trademarks. All other trademarks are the property of their respective owners.