Assessment of microbial population diversity in polymicrobial research samples by 16S single or multi-V region sequencing

E.Bolchakova, A. Schramm , A. Lackey, P. Balachandran

Thermo Fisher Scientific, Pharma Analytics group, 180 Oyster Point Blvd, South San Francisco, CA, 94080

Introduction

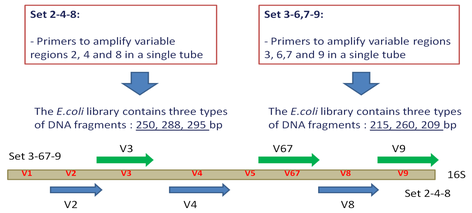
Analysis of 16S sequences in microbial populations using NGS gives a rapid overview of the community diversity, and is usually performed by sequencing one or two hypervariable regions (V-regions), out of the nine present in the 16S rRNA gene. In this study we compared the community structure of fecal, oral and water microbiomes by analyzing sequences from a single variable region, or from the seven V-regions (V2, V3, V4, V6-7, V8 and V9) included into Ion 16S™ Metagenomics Kit (multi-V analysis)

Ion 16S Metagenomics Kit

Ion PGM™ System

Ion Reporter™ Software

Fig 1. 16S metagenomics primer sets simultaneously amplify several hypervariable regions of bacterial 16S rDNA gene



Materials and Methods

Total DNA from fecal samples was prepared by suspending the material in PBS, centrifuging to remove crude debris, bead beating and finally extracting DNA using the PrepSEQ® sample preparation kit. Direct lysis method was applied to a water sample DNA , extracted from 1 ml of water sample or from the community cultured two days on R2A media on a 0.45 um PALL™ filter, was used as PCR template for analysis using Ion 16S Metagenomics workflow.

The microbial DNA content was estimated using a SYBR® Green based qPCR assay and the extracted DNA input amounts and number of PCR cycles for amplicon library preparation were adjusted appropriately. The sequencing was performed on an Ion PGM System, with Ion PGM Hi-Q™ or 400 bp sequencing chemistry. The reads were classified by 16S Metagenomics Analysis module in the Ion Reporter Software. The data were compared at the genus-only level, to avoid biases in species-level resolution capabilities by individual V-regions.

Results

Oral DNA sample contains mostly non-microbial DNA. The table below shows results of qPCR analysis of this sample for microbial 16S DNA content by SYBR® Green qPCR and suggests the number of PCR cycles.

Table 1. How many PCR cycles to use if sample contains mostly non-microbial DNA?

Sample	Total DNA content, pg/ul	16S qPCR content, pg/ul	Cycle number increase (compared to 1.5 ng input)	Required number of PCR cycles (2 ul DNA input)
#1	1.52E02	0.029	+15	33
#2	1.86E05	10.3	+7	25
#3	2.10E06	33	+6	24

For Ion 16S Metagenomics analysis, calculate number of required PCR cycles by using the equation:

PCR cycle number =18+ Log $_2$ (1500pg/ (qPCR quant (pg/ul stock)* desired sample input vol))

Table 2 . Oral microbiome study. If a single or two V regions are used to query the sample, true diversity of the sample will be missed.

Multi-V analysis reveals the more complex community structure. In the microbiome of an oral sample, from 107 total genera identified, 57% of diversity was observed in V3 analysis, with 39% of unique genera. The V2, V4, V6,7, V8 and V9 analysis uncovered 35%, 35%, 36,% ,25% and 10% of total diversity respectively.

	Genera Identified	% of Total	Unique genera ID only by this primer	% of unique
Full 16S kit	107	100%	62	100%
V2	37	35%	7	11%
V3	61	57%	24	39%
V4	37	35%	11	18%
V6,7	38	36%	9	15%
V8	27	25%	8	13%
V9	11	10%	2	3%

Fig 2. Monitoring growth-conditions-induced changes in water microbial community

Changes in community are easily seen: the entire group of microbes, amplified by V6,7 primer set is missing in cultured sample.

Original community

Culturable community (2 days of growth on R2A media on filter)

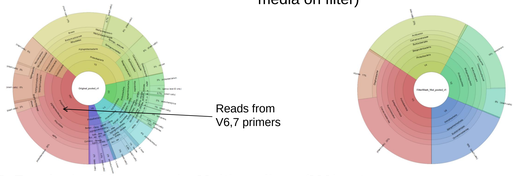
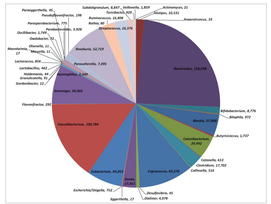


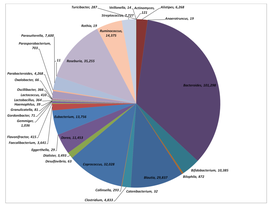
Fig 3. Fecal microbiome study. Multiamplicon (MA) set uncovers more quantitative and qualitative diversity in the sample

Substantial differences in the microbial profile were uncovered in the fecal sample. For example, with the multi-V, 14% and 13% of all reads were classified into genera *Coprococcus and Fecalibacterium*. For V4 only library, *Fecalibacterium* reads represented 0.52% of total reads while the proportion of *Coprococcus* reads remained nearly unchanged at 16%. When the data for a single V4 region from the multi-V- or V4-only sequencing libraries were compared, no distortion in the community structure was observed, indicating the absence of interference in the multiplex PCR reaction from other primers in the Ion16S Metagenomics kit.

MA set

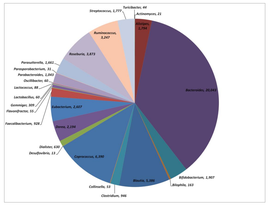


V4 only set

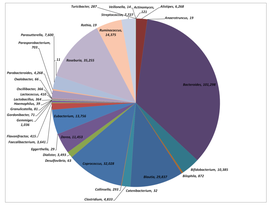


Unique Genera in MA Set (659856 reads)	Counts	V region	%of mapped reads	Unique Genera in V4 only set (291528 reads)	Counts	%of mapped reads
Butyricicoccus	1737	V8	0.19	Pediococcus	11	0
Catonella	613	V3	0.07			
Escherichia/Shigella	752	V2	0.08			
Holdemania	11	V2	0			
Mannheimia	17	V9	0			
Moryella	11	V3	0			
Olsenella	11	V3	0			
Paraeggerthella	45	V8	0			
Pseudoflavonifractor	136/61	V8/V3	0.01			
Subdoligranulum	8847	V2	0.96			

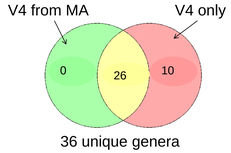
Fig 4. V4 community structure derived from V4 reads in the MA set and V4 only sequencing library are very similar



V4 from MA set (55223 mapped reads)



V4 only set (291528 mapped reads)



The minor differences in detection between V4 only set and V4 in MA set can be explained by sequencing depth. However, these “unique” genera are reliably detected with several other V regions in the multiamplicon set (Table 3)

Table 3. Detection of low abundant species by MA set
(reads number are shown in parenthesis)

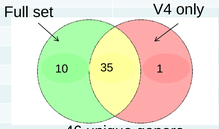
Unique genera in V4 only library	Counts	Detection by the MA set
Anaerotruncus	19	V2(10), V3(11), V8(12)
Catenibacterium	32	V3(13650) V6,7(13632), V2(3210)
Eggerthella	29	V8(17)
Gordonibacter	71	V8(12)
Granulicatella	81	V8(50), V3(41)
Haemophilus	39	V3(168), V6,7(491), V8(1162)
Oxalobacter	66	V2(22), V3(30), V8(20)
Pediococcus	11	The only unique genus in V4 only set
Rothia	19	V2(23), V3(57)
Veillonella	14	V2(277), V3(623), V8(959)

Conclusion

The use of seven variable regions detects more diverse bacterial populations while preserving the information from individual V-regions for a multi-dimensional analysis.

Full set

V4 only



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