Comparability of 16S rRNA gene sequencing in fresh frozen and formalin fixed paraffin embedded colorectal cancer samples.

Serna G. [1]*, Borgognone A. [2]*, Alonso L. [1], Noguera-Julian M. [2], Parera M. [2], Fasani R.[1], Paredes R[2#]. Nuciforo P. [1#]

*(lead presenters): gserna@vhio.net; aborgognone@irsicaixa.es. [#]co-senior authors.

[1] Molecular Oncology Group, Vall d'Hebron Institute of Oncology, Barcelona, Spain

[2] Infectious Diseases Service & Institut de Recerca de la SIDA - irsiCaixa, Barcelona, Spain

Keywords: microbiota, colorectal cancer, 16S gene sequencing, in situ hybridization.

Introduction

Advancements in microbial sequencing technologies have resulted in the discovery of several previously overlooked microorganisms that assume important roles in human diseases, including cancer. However, the impact of sample preservation on the microbiome composition has not been fully elucidated.

Methods

We analysed the microbiota of 10 paired fresh-frozen (FF) and formalin-fixed, paraffin-embedded (FFPE) colorectal cancer (CRC) biopsies using 16S rRNA gene sequencing. A negative control for PCR (purified water) was included in the dataset. After amplification of V3-V4 16S rRNA gene region, PCR products were indexed and sequenced on an Illumina MiSeq platform (Illumina, San Diego, CA). DADA2 software was used for data processing[1] and resulting outputs were analyzed by R and Bioconductor packages for metagenomics profiling (vegan[2] and phyloseq[3]) and data visualization (ggplot2[4]). Before performing downstream microbiome analysis, data were filtered for low-quality, eukaryotic, contaminant and low-abundant (singletons) features to include only study-specific sequences. Normalization was obtained by computing relative abundance (raw abundances divided by total counts per sample) for each sample. Sequencing data processing and downstream analysis were performed in the programming language R[5]. RNA in situ hybridization (ISH) was used to determine the spatial distribution and quantification of bacteria in FFPE samples. B-*Fusobacterium* 23S RNA, B-*propionibacterium* 16SrRNA and EB-16S-rRNA probes were obtained by ACD. Bacteria abundance was analysed with Visiopharm® image analysis software.

Results

FF samples displayed higher species richness and diversity when compared to the FFPE group (α-diversity Chao1 p=0.019 and Shannon p=0.064). Archea sequences were detected exclusively among FF samples, covering up to 0.2% of the total microbial composition. Orders from phyla *Actinobacteria* and *Proteobacteria* were more enriched among FFPE samples (p=0.002), while *Firmicutes* and *Fusobacteria* were significantly more abundant in FF group (p=0.052 and p=0.036, respectively). Beta diversity based on Jaccard (presence/absence) and Bray-Curtis (abundance) indices also showed significant differences between FF and FFPE groups (PCoA, adonis test $p=0.007$ and $p=0.001$, respectively). *Bacteroides* and *Propionibacterium* genera were identified across all FF and FFPE samples. Both genera, in addition to *Paracoccus, Streptococcus, Faecalibacterium, Geobacillus, Parabacteroides, Clostridium, Parvimonas* and *Fusobacterium* represented the top ten most abundant genera, cosidering all samples in dataset. RNA-ISH confirmed the presence of tumor-associated *Propionibacterium* in FFPE samples. *F. nucleatum* was found in two out of 10 samples by ISH, in line with the low abundance detected by 16S rRNA in FFPE samples.

Conclusion

Sample preservation significantly influences the microbiome characterization of CRC biopsies. This implies that microbiome results from FF cannot be directly extrapolated to FFPE, or vice-versa. However, despite these limitatons, 16S sequencing of CRC biopsies bears the potential to identify relevant microbes in CRC pathogenesis,