Shotgun metagenomics of the human stomach

ASSIGN BUSTER

Materials and Methods

Biopsy samples

Four fresh biopsy specimen taken from stomach. Two from antrum and two from the body. One biopsy from each side was stored in liquid nitrogen and another piece was used for the culture. For culturing, the fresh biopsy tissues was homogenized and simultaneously inoculated onto selective plates [5% horse blood and in the case of selective media antibiotics included trimethoprim (5 µg/ml), vancomycin (10 µg/ml), nalidixic acid (20 µg/ml) and amphotericin B (5 µg/ml)] and non-selective plates and incubated at 37°C with 10% carbon dioxide. All the antibiotics were from Sigma- Aldrich Corporation (St. Louis, MO, USA). All data of the subjects were collected from patient's file.

Ethical Approval: This study is approved by the UMMC Medical Ethics Committee and biopsy samples for culturing were obtained with informed and written consent from patients.

DNA extraction

Reagents from the MasterPure™ Complete DNA and RNA Purification Kit (Epicentre Biotechnologies, Madison, WI, USA) were used. 300 µL Tissue & Cell Lysis solution containing the ProteinaseK was added to homogenize fresh tissue (1-5 mg) or grind frozen tissues. Mixtures were incubated (65°C; 15 min). Cool the samples to 37°C and add 1µl of 5µg/µl RNaseA to the sample. Mixtures were incubated (37°C; 30 min). One hundred and seventy five µL MPC solution was added and cell debris was pelleted by

centrifugation (13 000 rpm; 10 min). The supernatant was removed to clean tube, 500 μL of isopropanol added, and DNA pelleted by centrifugation (13 000 rpm; 10 min). The supernatant was removed and the pellet washed sequentially with 2×500 μL aliquots of ethanol. All ethanol was removed from the tube and the pellet re-suspended in 200 μL of 1× Tris-EDTA (TE) buffer which was then retained for PCR.

Illumina sequencing

For each strain, whole-genome sequencing was performed using an Illumina Hiseq 2000 by generating paired-end libraries (500 bp and 2 kb) following the manufacturer’s instructions. The read lengths were 90 bp and 50 bp for each library, from which more than 100 Mb of high-quality data was generated. The paired-end reads from the two libraries were de novo assembled into scaffolds using SOAPdenovo (http://soap. genomics. org. cnwebcite).

Data analysis

MG-RAST (Meta Genome Rapid Annotation using Subsystem Technology, v3. 1) server at the Argonne National Library (http://metagenomics. nmpdr. org) was conducted as combined taxonomic domain information analysis which supplies numerous methods to access the different data types, including phylogenetic and metabolic reconstructions, and has the ability to compare the metabolism and annotations of one or more metagenomes [13], [14]. MG-RAST also provides the protein similarities analysis for sequences pasted filtration, including both function classification and function annotation. The protein similarity was carried out using BLAT against M5NR protein database

(http://metagenomics. nmpdr. org), which is an integration of many sequence databases into a single and searchable database. A single similarity search at this server will allow retrieving similarities to several databases, including NCBI-nr, KEGG, SEED, and etc.

Project accession numbers

Registered as project “ Shotgun metagenomics of the human stomach” (ID 5712) with MG-RAST. Metagenome accession numbers are 4543593. 3 (U122), 4543594. 3 (U22), 4543595. 3 (U37) and 4543596. 3 (U67) for public access.

Results

Study population and samples

Gastric biopsy tissue were collected from four patients suffering from PUD, NUD and GC presenting for esophagogastroduodenoscopy (OGDS) between August 2011 and September 2012. First patient was the Indian lady, suffer from Gactric cancer (GC), 77 years old who passed away last year 2013. Second and third were Indian lady and male who suffer from Peptic ulcer disease (PUD) and the last was Indian lady suffer from None ulcer despesia (NUD). Population characteristics of these patients are presented in Table 1.

Human gut microbiome identification using shotgun metagenomics

In previous study, we have isolated and identified culturable bacterial species from gastric biopsies of 131 *Hp* -positive and 84 *Hp* -negative Malaysian residents suffering of gastric diseases and belonging to different

ethnic groups using MALDI-TOF-Biotyper (unpublished data) using MALDI-TOF-Biotyper. We identified 61 species of bacteria from 27 genera and 3 phyla. The predominant phylum was followed by Firmicutes, while Actinobacteria were rare. At the genus level, Streptococci was the most predominant (121 positive biopsies) followed by *Neisseria* (44 positive biopsies), *Klebsiella* (42 positive biopsies) and *Lactobacilli* (42 positive biopsies) (data not shown). The result reported here was based on culturable bacteria and underlying information of the patients were not provide in order to have a complete view of gut microbial community. The predominant genus in this study was similar to those previously recorded [15-17, 18-20].

Present study has completed our previous report using metagenomic analyzing. The number of studies focus on analyzing the composition of the gut microbiome is expanding rapidly. This is mainly due to continuous improvements on the next generation sequencing technologies such as " shotgun" Sanger sequencing or massively parallel pyrosequencing. The sequencing technology advancement ensuring that all the members of the sampled communities to get largely unbiased samples of all genes [21]. Because of the ability to expose the earlier hidden diversity of microscopic life, metagenomics provides a powerful platform for viewing the microbial world. Knowledge gained from metagenomic studies has the potential to transform our current understanding of the entire living world by providing more complete information on the genomic content of microorganisms, both live and dead; whereas other omic methods such as metatranscriptomics and metaproteomics, only able to provide information on the active part of the microbiome [22]. Our current study utilized a sensitive molecular method

to reveal previously uncharacterized features of the gastric bacterial biota among adults with gastric cancer and peptic disease. Interestingly, *H. pylori* was the most abundant phylotype in the stomach of the subjects who tested positive for this organism by using conventional clinical approaches. However, *H. pylori* DNA was also present in the biopsy sample of one subject (UM022) that were considered to be *H. pylori* -negative by conventional methods. It is possible that the UM022 biopsy was cross-contaminated by *H. pylori* DNA during tissue processing, although abundant extraction controls did not yield visible amounts of DNA on gel electrophoresis after broad-range PCR. Another alternative explanation is that the number of *H. pylori* bacteria in sample (UM022) was too low to be detected by conventional methods. Similar result was also reported by Bik et al (2006). They reported, detection of bacterial DNA does not necessarily indicate the presence of live, resident organisms. Bacterial DNA in stomach biopsy samples might reflect the existence of bacterial cell remnants or transient presence of specific bacteria. We illustrated that the human gastric environment contains 154 bacterial genera belonging to 26 phyla, with 19 genera, *Helicobacter* , *Streptococcus, Prevotella, Staphylococcus, Citrobacter, Enterobacter, Klebsiella, Neisseria, Campylobacter, Fusobacterium, Rothia* , *Bordetella, Burkholderia* , *Cupriavidus, Escherichia, Shigella, Acinetobacter, Lactobacilles, Pseudomonas* and *Xanthomonas, Bacteroides sp* comprising ~90% of the bacterial community.

At the species level, *Candidiate persephone* (9196 abundance) was predominant followed by *Beggitoa* . sp. (5141 abundance), *Burkholderia pseudomallei* (21550 abundance), *Helicobacter pylori* (49520 abundance),

Burkholderia mallei (1694 abundance), *Helcobacter acinonychis* (2103 abundance), *Fusobacterium ulcerans* (1374 abundance), *Burkholderia thailaudensis* (625 abundance), *Burkholderia cenocepacia* (653 abundance), *Burkholderia ambifaria* (358 abundance) and *Endoriftia persephone* (261 abundance) (Table 1).

Besides bacteria, Archea (Table 2), Eukayota (Table 3) and viruses (Table 4) were also identified to be present in the gastric samples. *Magnoporthe* (228 abundance) were the most abundance organism from the domain of eukaryote, followed by *Trichoplax adhaerense* (39 abundance). *Human endogenous retrovirus* (185 abundance) and *Human endogenous retrovirus K* (193 abundance) were the two most predominated viruses detected in subjects' gastric samples. This finding is not unexpected because that the composition of the gastric community is not only determined by niche-specific factors but also by stochastic colonization from upstream components of alimentary tract. Less than 5% of bacteria found in this study were previously uncharacterized. *Burkholderia* spp., especially *B. pseudomallei* (21550 abundance) was one of the uncharacterized genera detected. This bacterium is the etiological agent of melioidosis, an endemic infectious disease in Southeast Asia, northern Australia, Brazil, and other parts of the world [23, 24, 25]. To our knowledge, this study is the first to report of *B. pseudomallei* isolation from the stomach of asymptomatic human subjects.

The data presented in this study demonstrate a clear picture of the bacterial communities in gut which was consistent with previous studies [8]. However, we couldn't approach any aim through the microbial community between

gastric cancer and peptic ulcer disease because of the low number of subjects which taken in this study. In addition, the data from GC patient- *H. pylori* negative versus PUD patients- *H. pylori* positive suggests that *H. pylori* does not show any significantly affect on the diversity of human gastric microbiota. Consistent with our results, Bik *et al* . who used a metagenomic approach found no influence of *H. pylori* in the diversity gastric microbiota [8]. Likewise, Maldonado-Contreras *et al* . did not report a variation in the number of non- *H. pylori* present but found that *H. pylori* presence resulted in differences in the relative abundance of several phyla [19]. Interestingly our result showed the low percentage of *Lactobacteria* as the essential or beneficial flora in our microflora community. Beneficial flora are the housekeepers of the gut and without them the gut cannot be healthy [26]. There is recent evidence that supplementation with *Lactobacillus casei* may help reduce the recurrence of colorectal tumors in patients who have previously undergone surgery for colon cancer. Additional research is needed in this area. Similarly, the low percentage of probiotics bacteria in those patients suffer from GC and PUD previously has been reported [27]. Another group of bacteria which we reported was opportunistic flora that they are including *Staphylococci, Streptococci, Bacilli, Clostridia, Yeasts, Enterobacteria, Fuzobacteria, Eubacteria, Prevotella, Veillonella* and *Rothia* and many others. There are around 500 various species of microbes known to science so far, which can be found in the gut. In a healthy person their numbers are limited and are tightly controlled by the beneficial flora. Each of these microbes is capable of causing various health problems if they get out of control [15, 28].

The atrophic stomach caused to increase the pH of stomach and surviving and growing the other bacteria except *H. pylori* to in normal acidic of stomach. Furthermore, in atrophic stomach *Streptococcus* can be shift into stomach instead of *Prevotella* [29]. Francisco Aviles-Jimenez (2014) showed the present of *Prevotella, Streptococcus* and *Rothia* in the corpus predominant atrophy stomach. However in another study by Lee (2009) and her group using antrum biopsies, showed the opposite tendency. Abundance *Neisseria* was decreased in the gastritis stomach and increased in the atrophic stomach.

The predominance of *Proteobacteria* observed in the gut of two PUD, one NUD and one GC patients was unique in the present study. This is interesting because three types of gut microbiota, termed “ enterotypes,” were proposed [30] in which *Bacteroides, Beggiatoa sp, Neisseria gonorrhoeae , Klebsiella pneumoniae , Staphylococcus aureus , Acinetobacter baumannii , Burkholderia pseudomalle, Staphylococcus aureus , Burkholderia mallei , Rickettsia prowazekii , candidate division* TM7 were dominant respectively, and reported to be most common gut microbiota in our study and might look at them as the marker. In Furthermore, comparison of microbiota diversity between PUD/GC and NUD shows that conclusion, Illumina sequencing technology has yielded important information for microbiologic diagnosis. It has a great advantage for the analysis of microbial communities and the results are not biased by the limits of classic culture techniques. This may help to elucidate the role of the microbiota in their natural niche and provide interesting insights into the ecology of the gastrointestinal system.