Metagenomic Analysis Reveals an Oral Microbiota Present Among the Oral Cancer Patients

Vaidehi Patel

Johns Hopkins University, Zanvyl Krieger School of Arts and Sciences

ABSTRACT

Motivation: In the recent years, Metagenomics has widened the scope of targeting microbes responsible for inducing various type of the cancers. In fact, about 16.1% of cancers are associated with microbial infection [1]. The Oral cavity is part of the digestive tract which has most bacteria rich parts in the human body. Therefore, microbiomes present in oral cavity leading to OSCC is important to study regarding development of various cancer in human body. Many cancers including the oral squamous cell carcinoma (OSCC) are associated with microbial infection. Therefore, the relationship between Oral squamous cell carcinoma and microbiome has been studied in this study. The abundance of Fusobacterium nucleatum, Haemophilus influenzae Rd and Bacillus subtilis subsp. Subtilis str168,Zymomanoas mobilis subsp ZM, and Pseudomonas aeruginosa, Escherichia coli k-12 were found in both cancer patient samples.

1   INTRODUCTION

Oral squamous cell carcinoma (OSCC) is commonly occur-ring head and neck cancer which affects human population all over the world. The OSCC cancer has been increasing with 5-year survival rate around 50% worldwide [2]. Many risk factors including tobacco use, and alcohol consumption could lead to the oral cancer. The Association between oral microbiome and oral squamous cell carcinoma could pro-vide insight and identify biomarkers with early diagnosis. Metagenomics is the study of microbes in their natural living environment. Metagenomics has been revolutionized to study and identify microbiomes in type of various cancers. In this study, identification of microbial diversity present in the specimens of Oral squamous cell carcinoma cancer patients were investigated. Many studies have been conducted that shows the relationship between bacteria and human diseases. Recently, research has shown that many microbes has been responsible for promoting carcinogen pathogens leading to development of the cancer (Smoot, 1997).

In this study, the publicly available data has been used which consist of swab samples from three individual with ethnicity of white and Asian. Two of the patient samples suffered from oral cavity squamous cell carcinoma and one normal healthy patient with no cancer. The samples collect-ed were part of the study of Oral microbiome for cancer development, progression, and recurrence

This study attempts to find which microbiomes are present in the OSCC patient samples. Also, exploring microbial communities in normal and cancer patient samples will give us insight on key pathogens which could be responsible for inducing the Oral squamous cell carcinoma cancer. However, many carcinogens pathogen such as H.Pylori, hepatis B and C virtues and human papillomavirus and Epstein virus are classified as pathogen could lead to oral cancer[3]. In this study, metagenomic analysis was performed in order to identify the bacterial and viruses diversity associated with Oral squamous cell carcinoma(OSCC).

2   METHODS

Study Population and Data

The publicly available data was collected from NCBI SRA database for BioProject (PRJEB4953). The study was to develop a framework for exploiting the oral microbiome for monitoring oral cancer development. Three human oral metagenome samples including SRR3586068 (Cancer), SRR3586069 (Normal), SRR3586070 (Cancer) were used for further metagenomic analysis Furthermore, many bioinformatics tools were used for the analysis of microbiome present in cancer and Normal patient's samples.

Bioinformatics Analysis

The FASTQ files were downloaded from SRA. The raw reads were trimmed using Trimmomatic to remove low quality reads (Read dropped less than 36 bp). The human GRCH38.p14 reference genome was used in the meta-genomic analysis. BWA-MEM software was used to map the human reference genome GRCH38.p14 with default settings. The de-novo assembler, Megahit was used to assemble the sequences under default settings. Prodigal (Pro-karyotic Dynamic programming Gene-finding Algorithm) was used for gene prediction and translation initiation site in metagenome mode (- p meta). Diamond software was used to map protein-coding genes with NCBI swissprot database using BLASTP with e-value of 1e-5.The output of the Diamond software will be useful for matching the Taxonomic profiling analysis was performed using MEGAN6 software. The MEGAN6 software provides taxonomic information including kingdom, phylum, class, or-der, family, genus, species. The workflow of metagenomic analysis consist of each step performed was created by using Lucid chart as shown in Figure 1. This proposed workflow for this study aims at discovering the microbiome composition present in oral cancer patients. The supple-mental document has been attached to this paper with codes and tools information used in this study.

3   RESULTS

Sequence Data statistics

After trimming and quality control of raw sequences of control, cancer patient samples, reads are mentioned in Table1. The data from Sequence Read Archive Database under Bio project Accession number PRJEB4953 is used.

Table 1: Sequence Read Archive (SRA) Database under BioProject Accession Number (PRJEB4953)

Sample	SRA Accession No.	Read Sequence Count
Cancer	SRR3586068	303367
Normal	SRR3586069	745635
Cancer	SRR3586070	428095

Figure 1: Workflow of Metagenomics Analysis

Oral microbiome composition of patients with oral cancer and normal healthy:

Megan6 software performed the taxonomic analysis using the diamond output daa format files. Meganizer was used to convert the daa format files for MEGAN6. The results revealed an abundance of microbial communities in both cancer samples as seen as in Figure 2 and 3. Taxonomic analysis shows that abundance of microbial composition consists of Fusobacterium nucleatum, Haemophilus influenzae, and Bacillus subtilis subsp as seen in Figure 2 for SRR3586068 cancer sample. Also, Figure 3 shows that microbial composition of Escherichia coli, Bacillus subtilis subsp. Subtilis str., Haemophilus.influezae Rd, Pseudomonasaeruginosa PAO1, Zymomonas mocilis subsp present in second oral cancer (SRR35866070) sample. The results shows that abundance of microbial communities found in cancer patients has been known oral

pathogens related to oral cancer according to many research study related to Oral Cancer[5].

Figure 2: Metagenomics Analysis of Bacteria Taxonomy in Cancer Sample (SRR3586068)

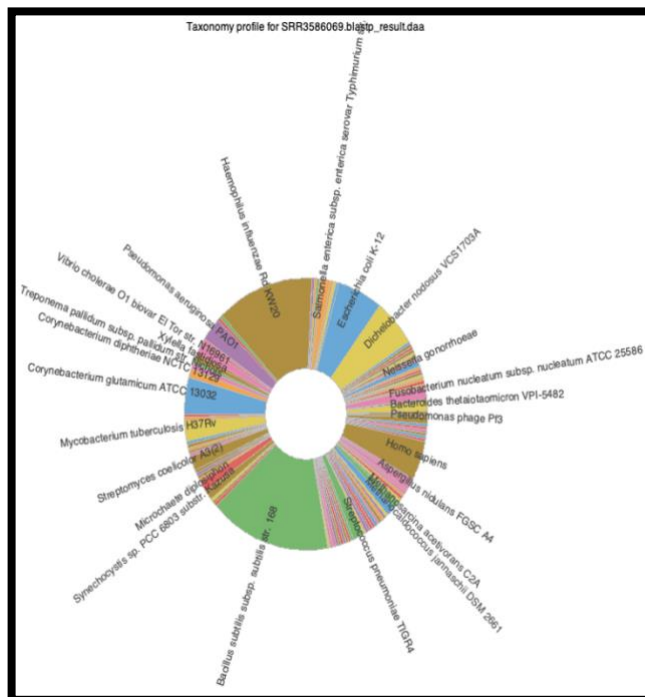*Figure 3: Metagenomics Analysis of Bacteria Taxonomy in Cancer Sample (SRR3586070)*

Figure 4: Metagenomics Analysis of Bacteria Taxonomy in Normal Healthy Sample

The Metagenomic analysis shows that abundance of Bacillus subtilis subsp, dichelobacter nodosus, and Haemophilus influenzae present in the normal healthy patient samples as seen in figure 4.

Some research suggested involvement of viruses in Oral Squamous cell carcinoma[4]. Therefore, discovery of viruses present in the cancer samples could provide insight of viruses involved in oral cancer. In the Figure 5 shows the viruses including Caudovirales, Muvirus, Egavirus, phikzvirus, Entrobacteria phage P21, Invertebrate iridescent virus six present in SRR3586070 sample.
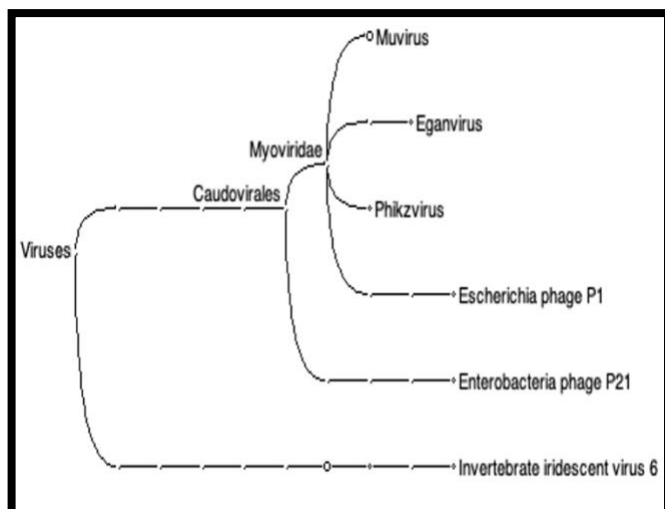

Figure 5: Metagenomics Analysis of Virus in Cancer Sample (SRR3586070)

4 CONCLUSION

The Research shows that microbiome of patients with Oral Cancer and normal healthy patient. In the conclusion, viruses present in the oral cancer samples are also shown in this study. The microbiome diversity of abundant bacteria present in the oral cancer patients in this study was also found in other research related to oral cancer. This research could potentially have scope of analysis of the data further.

5 REFERENCES

[1] Banerjee, J., Mishra, N., & Dhas, Y. (2015). Meta-genomics: A new horizon in cancer research. Meta gene, 5, 84-89.

[2] Peter, T. K., Withanage, M. H., Comnick, C. L., Pendleton, C., Dabdoub, S., Ganesan, S., ... & Zeng, E. squamous cell carcinoma associated oral microbiome. Frontiers in Microbiology, 13.

[3] Sharma, M. K., Srivastav, V. K., Joshi, C. K., Kumar, M., & Bhat, K. M. (2022). Metagenomic analysis of oral microbiota among oral cancer patients and to-bacco chewers in Rajasthan, In-dia. Bioinformation, 18(9), 757-763

[4] Gupta, K., & Metgud, R. (2013). Evidences suggesting involvement of viruses in oral squamous cell carcinoma. Pathology research international, 2013

[5] Liu, Y., Li, Z., Qi, Y., Wen, X., & Zhang, L. (2022). Metagenomic Analysis Reveals a Changing Microbiome Associated with the Depth of Invasion of Oral Squamous Cell Carcinoma. Frontiers in micro-biology, 13.Alexandrescu,A. (2001) *Modern C++ Design: Generic Programming and Design Patterens Applied.* Addision Wesley Professional, Boston.