

**CARACTERIZAÇÃO DA MICROBIOTA GÁSTRICA RESIDENTE E
TRANSIENTE EM PACIENTES PORTADORES DE
ADENOCARCINOMA GÁSTRICO**

GABRIELA ESTRELA DE ALBUQUERQUE

**Dissertação apresentada à Fundação Antônio Prudente para
obtenção de Título de Mestre em Ciências**

Área de concentração: Oncologia

Orientador: Dr. Emmanuel Dias-Neto

Coorientadora: Dra. Thais Fernanda Bartelli

São Paulo

2020

FICHA CATALOGRÁFICA

Preparada pelo Ensino Apoio ao aluno da Fundação Antônio Prudente*

A345 Albuquerque, Gabriela Estrela de
Caracterização da microbiota gástrica residente e transiente em pacientes portadores de adenocarcinoma gástrico / Gabriela Estrela de Albuquerque – São Paulo, 2020.
108p.
Dissertação (Mestrado)-Fundação Antônio Prudente.
Curso de Pós-Graduação em Ciências - Área de concentração: Oncologia.
Orientador: Emmanuel Dias-Neto

Descritores: 1. Adenocarcinoma/Adenocarcinoma. 2. Neoplasias Gástricas/Stomach Neoplasms. 3. Metagenômica/Metagenomics. 4. Microbioma Gastrointestinal/Gastrointestinal Microbiome. 5. Microbiota/Microbiota.

Elaborado por Suely Francisco CRB 8/2207

*Todos os direitos reservados à FAP. A violação dos direitos autorais constitui crime, previsto no art. 184 do Código Penal, sem prejuízo de indenizações cabíveis, nos termos da Lei nº 9.610/08

SUPORTE À PESQUISA POR AGÊNCIA DE FOMENTO

Agradecemos o apoio financeiro recebido da Fundação de Amparo a Pesquisa do Estado de São Paulo (FAPESP). Este apoio se deu na forma de uma bolsa de mestrado (processo 2018/14267-2), além de um projeto temático, intitulado Epidemiologia e Genômica de Adenocarcinomas Gástricos no Brasil (processo 2014/26897-0), ao qual este projeto foi vinculado.

*“Nothing in Biology Makes Sense Except in
the Light of Evolution”*

(Theodosius Dobzhansky)

AGRADECIMENTOS

Gostaria de agradecer primeiramente a minha família (em especial a minha mãe, dona **Iris**) e meus amigos que acompanharam esta trajetória, foram pacientes ao ouvir dos meus problemas, e me incentivaram para que de alguma forma eu chegasse até aqui. **Bruna** (e **Bê**), **Sarah**, **Tiago**, **Wellington**, **Giulia**, **Talita**, **Ana**, **Lenin** e **Victor**, muito obrigada!

Gostaria de agradecer a todos do **Laboratório de Genômica Médica**, em especial meu orientador, **Dr. Emmanuel Dias-Neto** por ter me recebido no laboratório juntamente com a **Dra. Diana Nunes**, e por terem me mostrado o mundo da metagenômica e terem me dado a oportunidade de realizar este trabalho. Em segundo lugar, mas não menos importante, gostaria de agradecer a minha coorientadora, **Dra. Thais Bartelli**, por ter desenvolvido este projeto comigo desde o início e por estar sempre disposta a conversar e ajudar. Ao mesmo tempo gostaria de pedir desculpas pelas inúmeras vezes que te atrapalhei, pois não deve ser fácil ter sua aluna sentando bem ao seu lado. Gostaria de dizer que você foi fundamental para a realização deste projeto, e que eu não podia ter escolhido coorientadora melhor!

Gostaria também de agradecer aos alunos do laboratório, por terem compartilhado comigo esta jornada, seja discutindo experimentos ou apenas relaxando e tomando um café, conheci diversas pessoas do grupo, que jamais esquecerei, mas tem aquelas que vou levar pra sempre no coração: só **Luiza** me passou seu sotaque pro resto da vida, só **Isabella** consegue sempre ver o lado bom nas piores situações, e só **Marianna** pra fazer eu abraçar as pessoas.

Gostaria de agradecer também a **Dra. Maria Amorim**, por ter me ensinado o mundo da biologia molecular mais afundo, sempre com a sua vontade de fazer experimentos e comemorar bons resultados.

Gostaria de agradecer a todos do **Laboratório de Biologia Computacional**, pelo auxílio nas análises de sequenciamento, em especial gostaria de agradecer ao **Bruno**, pela paciência sem limites para qualquer problema, mas em especial quando o assunto é o R, e ao **Alexandre** pelas aproximadamente mil análises nesse tempo de mestrado.

Gostaria de agradecer a toda a equipe do **A.C.Camargo Cancer Center** e do **Centro Internacional de Pesquisa**, que me ajudaram de alguma forma, seja nas coletas utilizadas neste estudo, até as partes mais burocráticas do meu projeto que foram realizadas.

Por fim, gostaria de agradecer o apoio financeiro da **Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP)** que possibilitou o desenvolvimento deste projeto.

RESUMO

Albuquerque GE. **Caracterização da microbiota gástrica residente e transiente em pacientes portadores de adenocarcinoma gástrico.** [Dissertação]. São Paulo: Fundação Antônio Prudente; 2020.

O Adenocarcinoma Gástrico (AdG) está entre as neoplasias mais comuns no mundo, com alta incidência e mortalidade. A microbiota influencia ativamente o curso de diversos tumores, podendo contribuir para a tumorigênese, ativando/inativando quimioterápicos ou mesmo modulando a resposta imune que pode ajudar a combater a doença. A microbiota estomacal ainda é pouco estudada, mas certamente é composta por microrganismos transientes, derivados da cavidade oral, e também por verdadeiros residentes – que talvez tenham maior impacto direto na biologia tumoral. O estudo aqui apresentado teve como objetivo buscar evidências biológicas que ajudem a indicar bactérias com maior probabilidade de estarem vivas no estômago de pacientes com AdG. Tais evidências incluíram diferenças de microbiota entre saliva, biópsia e fluidos gástricos, análise da atividade transcricional buscando diferenças quantitativas entre leituras de DNA e RNA, além de crescimento de microrganismos de fluidos gástricos em meios de cultura diferentes na presença e ausência de oxigênio, seguido de sequenciamento para identificação taxonômica. Amostras de RNA e DNA foram extraídas e utilizados para abordagens de RNAseq, *shotgun* e sequenciamento de 16S rRNA. *Streptococcus pneumoniae* foi a bactéria encontrada em maior abundância tanto no DNA como no RNA de fluidos gástricos (11% e 28% respectivamente), indicando que ela está altamente ativa na região estomacal. Já as bactérias *Bacillus cereus*, *Elizabethkingia anophelis*, *Legionella pneumophila* e *Virgibacillus senegalensis* se destacaram como mais abundantes no RNA, enquanto que as espécies *Solobacterium moorei* e *Streptococcus pseudopneumoniae* se destacaram no DNA. Houve baixa associação positiva entre a abundância relativa da microbiota a nível de DNA versus RNA nos fluidos gástricos ($r=0,51$ e $p<0,05$). Tanto a diversidade alfa (índices *Observed* e *Simpson* $p<0,05$), quanto a diversidade beta (índices *Bray-Curtis* e *Unifrac* $p<0,05$) das amostras foram distintas. A partir da cultura de fluidos gástricos identificamos duas espécies de fungo (*Trichosporon asahii* e *Candida parapsilosis*), além de espécies bacterianas dos gêneros *Streptococcus*, *Staphylococcus*, *Bacillus* e *Rothia*. Já no

sequenciamento das biópsias (DNA x RNA) identificamos que no geral eles renderam composições similares, visto pela alta correlação entre a abundância dos gêneros bacterianos identificados pelas duas abordagens ($r=0,82$ e $p<0,05$), além de possuírem valores de diversidade similares. O único gênero bacteriano que se destacou nas biópsias a nível de DNA foi o *Prevotella*, também com elevada abundância nas amostras de saliva. Ao avaliarmos a microbiota da saliva, observamos que alguns gêneros bacterianos estão em proporções mais similares entre saliva e fluido gástrico, enquanto outros se assemelham na saliva e biópsia, indicando que algumas bactérias da cavidade oral podem se associar ao tecido, enquanto outras permanecem no lúmen estomacal. Por fim, concluímos que as abordagens de sequenciamento são complementares, visto que o RNA é altamente relevante para identificarmos a microbiota ativa transcricionalmente nos fluidos gástricos e que influenciam o microambiente mais ativamente. Já o cultivo dos microrganismos permite confirmarmos as bactérias que realmente estão ativas/vivas na região, e que possam ter papel relevante no AdG.

Descritores: Adenocarcinoma. Neoplasias Gástricas. Metagenômica. Microbioma Gastrointestinal. Microbiota

SUMMARY

Albuquerque GE. **[Characterization of the resident and transient gastric microbiota in patients with Gastric Adenocarcinoma]**. [Dissertação]. São Paulo: Fundação Antônio Prudente; 2020.

Gastric Adenocarcinoma (GAC) is one of the most common neoplasms in the world, with high incidence and mortality rates. The microbiota has an active role in cancer and may play relevant roles in carcinogenesis, activation/inactivating of chemotherapeutic agents, as well as in immune modulation that may determine health or disease. The gastric microbiota is still poorly studied, but certainly contains a population of transient microorganisms, derived from the oral cavity, as well as true residents - that maybe have a stronger impact on tumor biology. Our study evaluated biological evidence that would indicate bacteria more likely to be alive in the gastric environment of GAC patients. These evidences included differences in composition between saliva, gastric fluids and biopsies, analysis of transcriptional activity and quantitative differences between DNA and RNA reads, as well as the growth of microorganisms in distinct media and oxygen concentrations, followed by DNA sequencing for taxonomic identification. DNA and RNA samples were extracted and used for RNASeq, shotgun and 16S sequencing. *Streptococcus pneumoniae* was the most abundant bacterium for both DNA and RNA levels in gastric fluids (11% e 28% respectively), indicating its high activity and strong presence in the stomach. *Bacillus cereus*, *Elizabethkingia anophelis*, *Legionella pneumophila* and *Virgibacillus senegalensis*, on the other hand, were more abundant in the RNA, while the species *Solobacterium moorei* and *Streptococcus pseudopneumoniae* were differentially abundant in the DNA. We found low positive association between the relative abundance of the microbiota at DNA and RNA levels in gastric fluids ($r = 0.51$ and $p < 0.05$). Both the alpha (Observed and Simpson $p < 0.05$) and beta diversities (Bray-Curtis and Unifrac, $p < 0.05$) indicated distinct populations for DNA and RNA molecules reinforcing the complementarity of these data when considering the human microbiota. Two fungi (*Trichosporon asahii* and *Candida parapsilosis*) were identified in the gastric fluid cultures, in addition to bacterial species of the genera *Streptococcus*, *Staphylococcus*, *Bacillus*, and *Rothia* demonstrating these to be alive in this region. The 16S data on biopsies (both DNA x RNA), showed high correlation bacterial genera

identified by the two approaches ($r = 0.82$ and $p < 0.05$) and similar diversity values, indicating the identified species to be alive inside the gastric tissue. The only bacterial genus more abundant in the biopsies DNA was *Prevotella*, also highly abundant in the saliva samples. When evaluating the composition of the stomach versus the oral microbiota, we observed that some bacterial genera had similar proportions in saliva and gastric fluid, while others were similar in saliva and biopsy, indicating that some bacteria from the oral cavity may associate with the gastric tissue, while others remain in the stomach lumen. Finally, we concluded that the approaches used in this study are complementary; while DNA sequencing is important for the identification of all the microorganisms that pass through the stomach cavity, RNA is highly relevant to identify the transcriptionally active microbiota in the gastric fluids and the components that influence the microenvironment more actively. Meanwhile, culturomics not only allows us to confirm the bacteria that are actually active / alive in the gastric region, with a putative role in GAC but also enables us to isolate such species / strains for future functional studies.

Keywords: Adenocarcinoma. Stomach Neoplasms. Metagenomics. Gastrointestinal Microbiome. Microbiota.

LISTA DE FIGURAS

Figura 1	Abundância relativa de gêneros bacterianos presentes em fluidos gástricos de indivíduos controles e de pacientes com AdG.....	9
Figura 2	Ribossomo procarioto e gene da subunidade 16S rRNA	14
Figura 3	Razão do número de cópias de genomas bacterianos/humano em diferentes amostras biológicas	16
Figura 4	Desenho experimental proposto para este projeto	22
Figura 5	Cálculo utilizado para a identificação do número de cópias de DNA em 1ng de amostra	27
Figura 6	Visão geral da metodologia para cultivo <i>in vitro</i> de fluidos gástricos.....	29
Figura 7	Curva de saturação de <i>amplicons</i> derivados de biópsias (RNA e DNA), fluidos gástricos e saliva.....	33
Figura 8	Análise da integridade do RNA de amostras de fluido gástrico	40
Figura 9	Curvas padrão para DNA humano (beta-actina) e bacteriano (16S) relacionando o número de cópias de determinado genoma com seus respectivos CTs	42
Figura 10	Abundância relativa das espécies bacterianas com frequência superior a 1% nas amostras de fluido gástrico, estratificado por material sequenciado (DNA x RNA)	50

Figura 11	Análise comparativa da abundância das espécies bacterianas com frequência superior a 1% nas amostras de fluido gástrico (DNA x RNA)	54
Figura 12	Correlação da abundância relativa (log10) das espécies bacterianas nas amostras de fluidos gástricos (DNA x RNA)	56
Figura 13	Medidas de diversidade nas amostras de fluido gástrico (DNA total versus RNAseq)	57
Figura 14	Abundância relativa de <i>Helicobacter pylori</i> na microbiota bacteriana de fluidos gástricos (DNA x RNA).....	59
Figura 15	Crescimento microbiano de fluidos gástricos avaliado após 48h	61
Figura 16	Curvas de rarefação/saturação das amostras sequenciadas por 16S	69
Figura 17	Abundância relativa dos gêneros bacterianos com frequência superior a 1% nas amostras de biópsia, estratificado por material sequenciado (DNA x RNA)	73
Figura 18	Análise comparativa da abundância dos gêneros bacterianos com frequência superior a 1% nas amostras de biópsia (DNA x RNA)	74
Figura 19	Correlação da abundância relativa (log10) dos gêneros bacterianos presentes nas amostras de biópsias (DNA x RNA).....	75
Figura 20	Medidas de diversidade nas amostras de biópsia (DNA x RNA).	76
Figura 21	Comparação do número de gêneros encontrados	78
Figura 22	Abundância relativa dos gêneros bacterianos com frequência superior a 1% nas amostras de biópsia, fluido gástrico e saliva	79

Figura 23	Análise comparativa da abundância dos gêneros bacterianos com frequência superior a 1% nas amostras de biópsia, fluido gástrico e saliva	80
Figura 24	Correlação da abundância relativa (log10) dos gêneros bacterianos presentes nas amostras	80
Figura 25	Medidas de diversidade nas amostras de biópsia, fluido gástrico e saliva .	82

LISTA DE TABELAS

Tabela 1	Condições de conversão do RNA de fluidos gástricos para cDNA utilizados neste estudo	24
Tabela 2	Condições de amplificação dos genes 16S (região V1) e beta actina utilizadas neste estudo.	25
Tabela 3	Condições de termociclagem para amplificação dos genes 16S (região V1) e beta actina utilizadas neste estudo.	25
Tabela 4	Condições de amplificação do gene 16S (regiões V3-V4) utilizadas neste estudo	31
Tabela 5	Condições de termociclagem para amplificação do gene 16S (regiões V3-V4) utilizadas neste estudo	31
Tabela 6	Dados clínicos gerais dos pacientes deste estudo	36
Tabela 7	Quantificação de ácidos nucleicos (RNA e DNA) das amostras utilizadas neste estudo	38
Tabela 8	Média do CT de amplificação bacteriano e humano nas amostras de fluido gástrico (cDNA)	39
Tabela 9	Massas utilizadas, CT de amplificação e número de cópias dos genomas humano e bacteriano utilizados para construção de curvas padrão	41
Tabela 10	Média do CT, número de cópias e razão entre genomas bacterianos e genomas humanos nas amostras de fluido gástrico e biópsia.....	44

Tabela 11	Processamento das <i>reads</i> : amostras sequenciadas de fluidos gástricos (RNAseq e DNA total)	47
Tabela 12	Anotação das <i>reads</i> sequenciadas de fluido gástrico (DNA total) no genoma de <i>Mycobacterium leprae</i>	53
Tabela 13	Meios de cultura, condições de oxigênio e número de colônias observadas através do fluido gástrico de três pacientes do estudo	60
Tabela 14	Identificação dos microrganismos sequenciados, <i>reads</i> totais geradas e <i>reads</i> associados ao microrganismo potencialmente identificado	63
Tabela 15	Bactérias cultivadas e suas frequências relativas na microbiota sequenciada de fluidos gástricos (DNA x RNA)	65
Tabela 16	Processamento das <i>reads</i> : amostras sequenciadas (16S) a partir de fluido gástrico e saliva	68
Tabela 17	Processamento das <i>reads</i> : amostras sequenciadas (16S) obtidas a partir de biópsias (DNA e RNA)	70

LISTA DE SIGLAS E ABREVIATURAS

°C	Celsius
µg	Micrograma
µl	Microlitro
µM	Micromolar
ACRG	Do inglês, <i>Asian Cancer Research Group</i> , traduzido como Grupo Asiático de Pesquisa do Câncer
ACT1	Do inglês, <i>Actin-1</i> , traduzido como actina 1
AdG	Adenocarcinoma gástrico
CDH1	Do inglês, <i>cadherin 1</i> , traduzido como caderina 1
cDNA	Do inglês, <i>complementary DNA</i> , traduzido como DNA complementar
CEP	Comitê de Ética em Pesquisa
CG	Câncer gástrico
CT	Do inglês, <i>cycle threshold</i> , traduzido como limiar do ciclo
CTLA-4	Do inglês, <i>cytotoxic T-lymphocyte-associated protein 4</i> , traduzido como proteína T-linfócito-associada citotóxico 4
DNA	Do inglês, <i>deoxyribonucleic acid</i> , traduzido como ácido desoxirribonucleico
dNTP	Desoxirribonucleotídeos trifosfatos
EBD	Do inglês, <i>Epstein-Barr virus</i> , traduzido como vírus Epstein-Barr
ECF/ECX	Epirrubicina, cisplatina e 5-fluorouracil/ capecitabina
FLOT	5-Fluorouracil, leucovorina, oxaliplatina e docetaxel
Gb	Gibabyte
H	Hora
HER2	Do inglês <i>human epidermal growth factor receptor 2</i> , traduzido como receptor do fator de crescimento epidérmico humano 2
HPV	Do inglês, <i>human papillomavirus</i> , traduzido como papiloma vírus humano
M	Molar
Mb	Megabyte
MgCl ₂	Cloreto de magnésio
Min	Minuto
mM	Milimolar
Ng	Nanograma
OMS	Organização mundial de saúde

OUT	Do inglês, <i>Operational Taxonomic Units</i> , traduzido como unidades
Pb	Par de base
PCR	Do inglês, <i>Polymerase Chain Reaction</i> , traduzido como reação Reação em cadeia da polimerase
PD-1	Do inglês, <i>Programmed Death 1</i> , traduzido como proteína de morte celular programada 1
PD-L1	Do inglês, <i>Programmed Death-Ligand 1</i> , traduzido como ligante de morte programado 1
pH	Potencial Hidrogeniônico
qPCR	Do inglês, <i>quantitative Polymerase Chain Reaction</i> , traduzido como Reação em cadeia da polimerase quantitativa
QTneo	Quimioterapia neoadjuvante
RNA	Do inglês, <i>Ribonucleic acid</i> , traduzido como ácido ribonucleico
Rpm	Rotações por minuto
rRNA	Do inglês, <i>ribosomal RNA</i> , traduzido como RNA ribossomal
RT-PCR	Do inglês, <i>reverse-transcriptase polymerase chain reaction</i> , traduzido como Transcrição reversa seguida de reação em cadeia da polimerase
S	Segundo
SEER	Do inglês, <i>The surveillance, epidemiology, and end result</i> , traduzido como A vigilância, epidemiologia e resultado final
TCGA	Do inglês, <i>The Cancer Genome Atlas Research Network</i> , traduzido como Atlas do genoma do câncer
TCLE	Termo de Consentimento Livre e Esclarecido
TNM	Do inglês, <i>Tumor-Node-Metastasis</i> , traduzido como tumor-nódulo-metastase
TP53	Do inglês, <i>tumor protein P53</i> , traduzido como proteína do tumor P53
VEGFA	Do inglês, <i>vascular endothelial growth factor A</i> , traduzido como fator de crescimento endotelial vascular A
VEGFR-2	Do inglês, <i>receptor for vascular endothelial growth factor 2</i> , traduzido como receptor para o fator de crescimento endotelial vascular 2

SUMÁRIO

1	INTRODUÇÃO	1
1.1	Câncer Gástrico	1
1.2	Microbiota Gástrica.....	6
1.3	Microbiota e Câncer	10
1.4	Técnicas Utilizadas para Estudo da Microbiota	13
1.4.1	Estudos baseados na análise de DNA: 16S rRNA e shotgun	13
1.4.2	Identificando microrganismos transcricionalmente ativos: Metatranscriptômica ..	16
2	OBJETIVOS	19
2.1	Objetivo Geral	19
2.2	Objetivos Específicos.....	19
3	METODOLOGIA	20
3.1	Aspectos Éticos	20
3.2	População de Estudo.....	20
3.3	Coleta de Dados Clínicos	20
3.4	Coleta e Processamento de Amostras Biológicas	21
3.5	Extração de Ácidos Nucleicos	22
3.6	PCR em Tempo Real	23
3.6.1	cDNA de fluidos gástricos.....	23
3.6.2	DNA de fluidos gástricos e biópsias.....	26
3.6.2.2	Construção de curvas padrão.....	26
3.7	Sequenciamento de DNA total de Fluidos Gástricos (<i>shotgun</i>)	27
3.7.1	Processamento e filtragem das amostras sequenciadas	27
3.8	Sequenciamento de RNA total de Fluidos Gástricos (RNAseq).....	28
3.8.1	Processamento e filtragem das amostras sequenciadas	28
3.9	Cultivo, Sequenciamento e Identificação de Colônias Isoladas a Partir de Amostras de Fluidos Gástricos	28
3.9.1	Seleção de fluidos gástricos, meios de cultura e condições de cultivo	28
3.9.2	Seleção das colônias e extração de DNA	29
3.9.3	Sequenciamento e identificação de microrganismos.....	30
3.10	Sequenciamento de <i>Amplicons</i> do Gene 16S rRNA e de seu Transcrito.....	30
3.10.1	Processamento e filtragem das amostras sequenciadas	32
3.11	Análises Estatísticas.....	33

4	RESULTADOS E DISCUSSÃO	35
4.1	Características da População Estudada	35
4.2	Extração de Ácidos Nucleicos	37
4.3	Análise do RNA Extraído de Fluidos Gástrico: Integridade e PCR em Tempo Real..	38
4.4	Quantificação do Número de Cópias de Genomas Bacterianos x Genomas Humanos por PCR em Tempo Real.....	40
4.4.1	Construção de curvas padrão	40
4.4.2	Quantificando o número de cópias do genoma humano e bacteriano em fluidos gástricos e biópsias.....	43
4.5	Resultados Do Sequenciamento: DNA Total e RNAseq	45
4.5.1	Processamento das reads.....	45
4.5.2	Composição da microbiota bacteriana de fluidos gástricos avaliada por molécula sequenciada (DNA x RNA)	48
4.6	Cultivo Microbiano e espécies Identificadas nas Amostras de Fluido Gástrico	59
4.6	Resultados do Sequenciamento: rRNA 16S	66
4.6.1	Processamento das reads.....	66
4.6.2	Composição da microbiota bacteriana das biópsias avaliada por tipo de molécula sequenciada (DNA x RNA)	71
4.6.3	Composição da microbiota bacteriana: região estomacal x cavidade oral	77
5	CONCLUSÕES.....	83
6	REFERÊNCIAS.....	85

ANEXOS

Anexo 1 Carta de aprovação do Comitê de Ética em Pesquisa do projeto temático gástrico

Anexo 2 Carta de aprovação do Comitê de Ética em Pesquisa-CEP

APÊNDICES

Apêndice 1 Termo de Consentimento Livre e Esclarecido-TCLE

Apêndice 2 Manuscrito do paper

1 INTRODUÇÃO

1.1 CÂNCER GÁSTRICO

O câncer gástrico (CG) é a quinta neoplasia mais frequente no mundo, com mais de 1 milhão de novos casos por ano, o que contribui para que esta seja a terceira maior causa de morte por câncer, com aproximadamente 783 mil mortes anuais (Bray et al. 2018; *World Health Organization* 2020). A distribuição do CG varia de forma significativa ao redor do globo, sendo que a Ásia representa 73%, a Europa 15%, e as Américas Central e do Sul 7% dos casos mundiais (Carcas 2014). Entre os homens, o CG é o mais incidente e mortal em alguns países da Ásia Ocidental, tais como Irã, Turcomenistão e Quirguistão. Já na Ásia oriental, se destacam com elevados índices da doença países como Japão, Mongólia e Coreia do Sul, sendo o último o país com as mais altas taxas desta neoplasia, em ambos os sexos (Bray et al. 2018). No Brasil, por sua vez, só em 2017 ocorreram mais de 14 mil mortes, e estimativas para o triênio 2020-2022 apontam para mais de 21 mil novos casos por ano (Ministério da Saúde 2020).

O CG é uma doença multifatorial, e ambos aspectos ambientais e genéticos possuem papel em sua etiologia. Estes fatores de risco, por sua vez, podem ser classificados em nove grandes grupos: dieta, estilo de vida, predisposição genética, histórico familiar, tratamentos e condições médicas, infecções, características demográficas, exposição ocupacional e radiação ionizante (Yusefi et al. 2018). Os hábitos alimentares têm sido relatados como um dos principais fatores de risco para o CG, sendo o excessivo consumo de sal um dos mais ressaltados. Além disto, o baixo consumo de frutas e vegetais e o consumo de alimentos processados e ultraprocessados, como a carne vermelha e embutidos, também estão entre os fatores de risco para esta neoplasia (Bonequi et al. 2013; Fang et al. 2015). Já com relação ao estilo de vida, assim como em outros tipos de câncer, o álcool e o tabaco se destacam (González et al. 2003; Ma et al. 2017). Ainda, sabe-se que o risco da doença aumenta em idades mais avançadas, e é cerca de duas vezes mais elevado em homens do que em mulheres (Ajani et al. 2017). A infecção pela bactéria *Helicobacter pylori* também tem sido extensamente relatada na literatura como relevante na doença, principalmente como fator inflamatório inicial que contribui para o estabelecimento do CG, além de infecções pelo vírus

Epstein-Barr (EBV) e pelo papiloma vírus humano (HPV) (Nagini 2012; Zeng et al. 2016; Nishikawa et al. 2018). Com relação aos aspectos genéticos, sabe-se que mutações germinativas no gene CDH1 são as mais presentes em pacientes com CG hereditário (Petrovchich e Ford 2016), esta neoplasia também já foi associada a polimorfismos em genes codificadores de citocinas, além de tipos sanguíneos específicos, entre outros (Wang et al. 2012; Karimi et al. 2014; Ajani et al. 2017; Yusefi et al. 2018).

A variante histológica mais comum do CG é o Adenocarcinoma Gástrico (AdG), neoplasia proveniente de glândulas da camada mais superficial do estômago, a mucosa, representando cerca de 90-95% dos casos (Dicken et al. 2005). Outros tipos de CG menos comuns compreendem cerca de 5-10% destes tumores e incluem o linfoma, o tumor estromal gastrointestinal, tumores neuroendócrinos, entre outras variantes mais raras (Richman et al. 2017). A Organização Mundial de Saúde (OMS) reconhece quatro padrões histológicos principais do AdG: tubular, papilífero, mucinoso e células pobremente agregadas, além de variantes incomuns (Lauwers et al. 2010; Hu et al. 2012). A classificação histopatológica mais utilizada, porém, ainda é a de Lauren (1965), onde os AdGs são divididos em dois tipos principais: intestinal e difuso, além dos subtipos mistos e indeterminados. Clinicamente, porém, os diferentes subtipos são tratados de forma similar, visto que a classificação histopatológica não é determinante do tratamento a ser utilizado nestes pacientes (Cislo et al. 2018).

Na tentativa de melhor categorizar pacientes com AdG visando terapias mais direcionadas, classificações moleculares têm sido desenvolvidas. O TCGA (*The Cancer Genome Atlas Research Network*) sugeriu em 2014 uma classificação molecular que divide estes tumores em quatro subtipos: (1) tumores positivos para o vírus Epstein-Barr; (2) tumores que apresentam instabilidade de microssatélites; (3) tumores genomicamente estáveis e (4) tumores com instabilidade cromossômica. Já em 2015, o ACRG (*Asian Cancer Research Group*) também propôs sua classificação, que, no entanto, divide estes tumores em: (1) tumores com alta instabilidade de microssatélites; (2) tumores com estabilidade de microssatélites e maior expressão de genes relacionados à transição epitélio-mesênquima; (3) tumores com estabilidade de microssatélites e mutações/perda de função de TP53 e (4) tumores com estabilidade de microssatélites e TP53 funcional (Cristescu et al. 2015). Mais recentemente, um grupo do Chile fez uma classificação molecular baseada em pacientes deste país, sendo o único estudo deste tipo, apenas com pacientes latino-americanos, já publicado até o

momento. Este estudo utilizou técnicas mais acessíveis como imuno-histoquímica e hibridização *in situ*, visando aplicação desta estratégia na rotina clínica, classificando estes tumores em 5 subtipos: (1) tumores positivos para o vírus Epstein-Barr; (2) tumores com defeito na via de *mismatch repair*; (3) tumores de transição epitélio-mesênquima; (4) tumores com acúmulo de TP53 e (5) tumores sem detecção de TP53 (Pinto et al. 2020). Apesar de estas e outras classificações ainda não serem utilizadas na prática clínica, as mesmas possuem impactos importantes, estando relacionadas a informações clinicamente relevantes tais como idade ao diagnóstico, localização na cavidade gástrica, recorrência, potencial de metastatização e prognóstico (Wang et al. 2019).

Um dos subtipos histológicos mais comuns do AdG é o chamado subtipo intestinal, caracterizado pela formação de glândulas, sendo mais frequente em homens e em idades avançadas, representando cerca de 54% dos casos (Cisło et al. 2018). Sua etiologia é bem estabelecida, tendo sido descrita em 1975 uma cascata de eventos que culminam no desenvolvimento desta neoplasia, que se inicia com inflamação na região estomacal, culminando na gastrite crônica, posterior gastrite atrófica, metaplasia intestinal, displasia, e eventualmente, câncer (Correa et al. 1975). Sabe-se que a infecção por *H. pylori* pode iniciar esta cascata inflamatória e, devido em partes à redução da incidência desta infecção por erradicação desta bactéria na população, o AdG tem diminuído ao redor do mundo (Wroblewski et al. 2010; Crowe 2019). Já no subtipo difuso, também conhecido como pouco diferenciado, a formação de glândulas não é característica, mas a formação de fibrose é acentuada. Este subtipo representa cerca de 32% dos casos, é mais agressivo, ocorre em pacientes em geral mais jovens, de forma mais uniforme entre os sexos, além de ter um histórico familiar mais acentuado (Ansari et al. 2018; Cisło et al. 2018; Waldum e Fossmark 2018; Assumpção et al. 2020). Ao contrário do subtipo intestinal, há relatos na literatura do aumento em incidência do subtipo difuso, observado em ambos os sexos (Henson et al. 2004).

No A.C.Camargo Cancer Center, dentro do projeto temático coordenado por nosso grupo - “Genômica de Adenocarcinomas Gástricos do Brasil”, observamos que para os 342 pacientes recrutados e já laudados desde o início do projeto, em 2015, 46% foram classificados como do subtipo difuso, contra 42% do tipo intestinal e 12% do tipo misto. Visto que este subtipo tumoral vem crescendo ao longo dos anos, não tem uma clara relação com *H. pylori* e não tendo sua etiologia bem estabelecida até o momento, se fazem necessários mais estudos focados neste subtipo tumoral.

A taxa global de sobrevida em 5 anos para os pacientes com AdG, é de cerca de 32%, mas se estratificada por estágio da doença pode chegar a números tão díspares quanto 69,5% para pacientes com tumores localizados ou 5,5% em pacientes metastáticos ao diagnóstico. Infelizmente sabe-se que apenas 28% dos pacientes com AdG são diagnosticados em estágios iniciais da doença, contra 36% que são diagnosticados já em estágios avançados (National Cancer Institute 2016). A maioria dos pacientes com AdG em estágios iniciais são assintomáticos, fazendo com que esta neoplasia seja diagnosticada tardiamente, prejudicando seu tratamento e levando a uma menor sobrevida para estes pacientes (Cervantes et al. 2013). Estudo recente realizado no A.C.Camargo Cancer Center, com uma coorte de 112 pacientes com AdG, observou que a maioria foi diagnosticada já em estadiamento clínico III (42,9%), e que para estes pacientes a sobrevida livre de doença em 5 anos foi de 32% (contra 87% nos pacientes em estágios I e 62% em estágio II) (Freitas 2019). Diante disso, embora a incidência da doença esteja diminuindo ao redor do mundo (possivelmente devido à redução do uso do sal para a preservação dos alimentos, associada a um diagnóstico mais preciso e mais precoce) e o tratamento seja hoje muito mais eficiente, o prognóstico do AdG ainda é pouco favorável.

O AdG é comumente diagnosticado por exame endoscópico alto, seguido de análise da biópsia da lesão suspeita (Johnston e Beckman 2019). O método amplamente utilizado para estadiamento da doença e que dá suporte à abordagem terapêutica, é o sistema TNM, baseado em 3 critérios básicos: (T) tamanho do tumor, (N) número de linfonodos afetados e (M) ocorrência de metástase (Amin 2017). Dependendo do tamanho e da localização do tumor primário, os tratamentos clássicos com finalidade curativa para o AdG envolvem a cirurgia com gastrectomia subtotal ou total (Carcas 2014; Smyth et al. 2016). A quimioterapia tem sido recomendada como tratamento neoadjuvante para redução da massa tumoral antes da cirurgia, e a quimiorradioterapia é usada em um cenário adjuvante para destruir células tumorais remanescentes, sendo a cisplatina e o 5-fluorouracil (ou seus análogos oxaliplatina e capecitabina) os quimioterápicos atualmente mais utilizados mundialmente (Orditura et al. 2014).

O estudo do *British Medical Research Council* foi o primeiro a demonstrar a eficácia da quimioterapia neoadjuvante (QTneo). Neste trabalho, Cunningham et al. (2006) compararam os quimioterápicos epirrubicina, cisplatina e 5-fluorouracil como tratamento neoadjuvante versus a cirurgia isolada em pacientes com AdG, mostrando aumento de 13% na taxa de

sobrevida em 5 anos para os pacientes que foram tratados previamente à cirurgia (36% *versus* 23%). Mais recentemente, um estudo clínico comparou 2 regimes de QTneo: (1) tratamento ECF/ECX: epirrubicina, cisplatina e 5-fluorouracil/ capecitabina *versus* (2) tratamento FLOT: 5-fluorouracil, leucovorina, oxaliplatina e docetaxel. Neste trabalho foi observado o aumento da sobrevida global em 15 meses nos pacientes que foram tratados com o FLOT, além do aumento de 9% na sobrevida em 5 anos (Al-Batran et al. 2019). Apesar de estes e outros estudos evidenciarem o papel positivo da QTneo na sobrevida dos pacientes, apenas uma pequena porção destes parece se beneficiar do tratamento, sendo que os que não se beneficiam estão sendo expostos à toxicidade desnecessária, maior custo de tratamento, exposição a efeitos colaterais indesejáveis, além de terem o tratamento cirúrgico atrasado.

A terapia alvo, por sua vez, tem proporcionado melhora na sobrevida de alguns pacientes. Sabe-se que HER2 está amplificado em cerca de 20% dos AdGs, e a utilização do anticorpo monoclonal Trastuzumabe associada à quimioterapia padrão leva a um aumento na sobrevida nestes indivíduos (Bang et al. 2010; Oritura et al. 2014; Van Cutsem et al. 2015). A expressão aumentada do fator de crescimento endotelial vascular A (VEGFA), por sua vez, está relacionada a um pior prognóstico e já foi demonstrado que a utilização do anticorpo Ramucirumabe, que se liga ao receptor VEGFR-2, apresenta aumento de sobrevida global (7,4% para 9,6%) em comparação com pacientes tratados apenas com quimioterapia (Wilke et al. 2014; Park et al. 2015). Porém, mais uma vez, apenas uma pequena parcela de pacientes parece se beneficiar destes regimes.

Mais recentemente, com os excelentes resultados obtidos com uso de inibidores de *check-point* imunológicos, em diversos tipos tumorais tais como melanoma, carcinoma de células renais, câncer de pulmão de células não-pequenas, entre outros (Larkin et al. 2015; Motzer et al. 2015; Reck et al. 2016), estudos têm sido realizados com o intuito de avaliar a eficácia deste tratamento também no CG (Alsina et al. 2016; Zhao et al. 2019). Na classificação molecular proposta recentemente pelo TCGA, o gene *PD-L1* (*Programmed Death-Ligand 1*), que auxilia a célula tumoral induzindo a apoptose de células T, estava mais comumente amplificado em pacientes EBV positivo e com alta instabilidade de microssatélites (*The Cancer Genome Atlas Research Network* 2014; Derks et al. 2016), porém a resposta clínica ao tratamento anti PD-L1 também foi encontrada em pacientes PD-L1 e EBV-negativos, sendo que até o momento não está totalmente claro porque alguns pacientes respondem a este tratamento e outros não (Dolcetti et al. 2018; Brar et al. 2019; Morihito et al. 2019).

Atualmente no Brasil se tem aprovado como terceira linha de tratamento o Pembrolizumabe, anticorpo monoclonal que se liga a PD-1 (*Programmed Death 1*), porém mais estudos ainda precisam ser realizados para melhor entendimento da interação entre esses achados imunológicos neste tipo tumoral (Fuchs et al. 2018; Ribeiro 2020).

A microbiota intestinal, por sua vez, já foi associada com a resposta clínica a anti PD-1 e anti PD-L1 em pacientes com câncer gastrointestinal. Recentemente, Peng et al. (2020) observaram que pacientes que responderam a esta imunoterapia possuíam uma elevada razão entre *Prevotella*/Bacteroidetes em suas fezes. Além disto, foi observado que as bactérias capazes de produzir ácidos graxos de cadeia curta, tais como *Eubacterium*, *Lactobacillus* e *Streptococcus* foram positivamente associadas à resposta ao tratamento em diferentes tipos de neoplasias gastrointestinais. Já especificamente em pacientes com CG, neste estudo foram encontrados gêneros bacterianos associados tanto à resposta ao tratamento anti PD-1/PD-L1 (como *Prevotella* e *Bifidobacterium*), como associadas aos não respondedores (como *Megamonas*, *Butyricimonas* e *Agathobacter*), mostrando que esta nova modalidade de tratamento pode estar interligada com a microbiota do indivíduo, reforçando a relevância de estudos nesta área.

1.2 MICROBIOTA GÁSTRICA

Humanos e microrganismos coevoluíram ao longo do tempo desde o surgimento dos primeiros hominídeos, levando ao estabelecimento de uma complexa e equilibrada rede de interações (Moeller et al. 2016). Se por um lado o corpo humano fornece nutrientes e abrigo aos microrganismos ali presentes, os metabólitos por eles produzidos influenciam em funções fisiológicas importantes, tais como metabolismo, absorção de nutrientes e resposta imune, impactando na homeostase do hospedeiro (Shreiner et al. 2015). O conjunto de microrganismos que habitam o corpo humano é conhecido como microbiota humana, e contém não só organismos procariotos, como bactérias e archaeas, mas também eucariotos e vírus. O microbioma, por sua vez, é um termo mais amplo, que se refere não somente aos microrganismos presentes em determinado ambiente, mas também a seu material genético, e às condições do meio no qual eles se encontram (Marchesi et al. 2015). A microbiota humana é, portanto, dinâmica e sujeita a alterações ao decorrer da vida do hospedeiro em

resposta a diversos fatores, tais como dieta, ambiente, intervenções médicas e doenças (Barko et al. 2018).

Até recentemente acreditava-se que para cada célula humana, teríamos aproximadamente 10 células bacterianas (Savage 1977; Suau et al. 1999; Rosner 2014). No entanto, cálculos atuais estimam que o ser humano possui quantidades similares de células humanas e bacterianas (1:1,3, respectivamente) (Abbott, 2016; Sender et al. 2016). E apesar de uma proporção menor de bactérias, este número (que não considera outros componentes da microbiota) ainda é surpreendentemente alto, e certamente diz da importância da microbiota, em específico das bactérias, na composição do seu hospedeiro. O intestino, responsável pela absorção de nutrientes dos alimentos, é o órgão com maior número de bactérias, com aproximadamente 10^{14} somente no cólon (Sender et al. 2016). Outros locais com alto número de bactérias incluem saliva e pele (ambos com aproximadamente 10^{11} bactérias no total), além do estômago com um total estimado em 10^7 bactérias (Sender et al. 2016).

Por ser uma região de transição do trato gastrointestinal, ligando a cavidade oral e esôfago ao intestino, o estômago está sujeito à presença de bactérias provenientes de outras regiões do corpo. Anteriormente acreditava-se que este era um órgão estéril devido principalmente à sua acidez e ao peristaltismo gástrico que impossibilitariam sua colonização e a consequente existência de uma microbiota própria. Além disto, sua localização e a dificuldade de coletar amostras para análise, juntamente com a complexidade dos meios de cultura específicos necessários para crescimentos dessas bactérias, contribuíram para que a crença deste órgão ser estéril se prorrogasse por muitos anos (Wang e Yang 2013; Nardone e Compare 2015).

Em 1984 Barry Marshall e Robin Warren identificaram uma bactéria no estômago de pacientes com gastrite, inicialmente classificada como *Campylobacter pyloridis* e reclassificada posteriormente como *Helicobacter pylori* (Lancet 1989), que é até hoje é uma das bactérias mais relevantes e mais estudadas no ambiente estomacal (Salama et al. 2013; Burkitt et al. 2017; Brito et al. 2019). A *H. pylori*, é um bacilo gram-negativo presente em cerca de 50% da população mundial, a infecção por esta bactéria é usualmente adquirida na infância, e quando não tratada, persiste no hospedeiro ao decorrer da vida (Peek e Blaser 2002).

A *H. pylori* secreta a enzima urease, que converte ureia em amônia, reduzindo a acidez do estômago, facilitando a sobrevivência desta bactéria, permitindo que ela penetre na mucosa estomacal e colonize o epitélio. Uma vez no epitélio, ela pode provocar uma resposta inflamatória que danifica a mucosa gástrica, acarretando diversas doenças, que em uma fase inicial pode gerar uma gastrite crônica, mas, posteriormente parece contribuir para o estabelecimento do CG (Hatakeyama 2009; Carcas 2014). Embora a etiologia do CG seja multifatorial, 80% dos casos são atribuídos à infecção crônica por esta bactéria, sendo estabelecida, portanto, como importante fator de risco para o CG e agente carcinogênico definitivo, reconhecido pela Organização Mundial de Saúde desde 1994 (Nagini 2012).

A descoberta de microrganismos na cavidade estomacal incentivou diversos autores a buscarem desvendar a composição da microbiota desta região, visando identificar microrganismos viáveis que habitam esta região em simbiose com seu hospedeiro. Em um estudo realizado na Faculdade de Medicina da Universidade de São Paulo, vinte voluntários considerados saudáveis após procedimento endoscópico tiveram muco do seu estômago coletado e cultivado em diferentes meios de cultura, revelando os gêneros *Veillonella*, *Lactobacillus* e *Clostridium* como predominantes nesta região (Zilberstein et al. 2007). Em contrapartida, em um estudo mais recente realizado na Espanha, indivíduos saudáveis também tiveram suas biópsias e fluidos gástricos cultivados, evidenciando a presença majoritária dos gêneros *Propionibacterium*, *Lactobacillus*, *Streptococcus* e *Staphylococcus* (Delgado et al. 2013). É interessante destacarmos a divergência entre os dados obtidos nos dois estudos. Tal variabilidade pode ser decorrente da população de estudo, do material coletado, e dos meios de cultura utilizados. Uma vez que um meio específico pode favorecer o crescimento de determinadas bactérias em relação às demais, o não crescimento de uma bactéria específica não significa necessariamente a sua ausência, que pode ser simplesmente um reflexo da não utilização dos meios adequados ao seu cultivo. Diante disto, o desenvolvimento dos métodos de análise de material genético e a redução do custo das abordagens de sequenciamento, têm sido de grande importância para identificação da microbiota humana, incluindo-se aí a estomacal.

Dados recentes gerados em nosso laboratório por análise do gene 16S rRNA mostram dezenas de bactérias em amostras de fluído gástrico, tanto em indivíduos controle como em pacientes com AdG (**Figura 1**), evidenciando que de fato o estômago é uma região com alto número de bactérias. Enquanto diversas destas podem ter uma presença temporária,

bactérias transientes oriundas da cavidade oral, talvez uma parcela talvez seja de fato residente do estômago, tendo uma atividade metabólica própria e talvez um papel relevante no CG, que ainda necessita ser desvendado.



Figura 1 - Abundância relativa de gêneros bacterianos presentes em fluidos gástricos de indivíduos controles e de pacientes com AdG. Resultados preliminares do projeto temático: “Genômica de Adenocarcinomas Gástricos” em andamento no Laboratório de Genômica Médica do A.C.Camargo Cancer Center.

A literatura ainda é relativamente escassa sobre microbiota gástrica e apresenta resultados contrastantes. Em um dos primeiros estudos utilizando técnicas de sequenciamento para avaliar a composição de bactérias no estômago, foram identificados os filos Proteobacteria, Firmicutes, Actinobacteria, Bacteroidetes e Fusobacteria como os mais comuns em biópsia estomacais de 23 indivíduos. Interessantemente, eles também observaram que a presença de *H. pylori* não interferiu na composição bacteriana do estômago (Bik et al. 2006). Todavia, outros estudos, mais recentes, mostram que a presença de *H. pylori* influencia diretamente a composição da microbiota gástrica, até mesmo a longo prazo, após sua erradicação (Yap et al. 2016; Guo et al. 2020; Miao et al. 2020).

Outro fator importante a ser considerado no estudo da microbiota gástrica é o tipo de amostra estudada. Uma vez que biópsias podem favorecer a identificação de bactérias que

são mais proximamente associadas ao tecido gástrico e seu muco, o estudo dos fluidos gástricos pode permitir uma visão mais global da microbiota estomacal, já que este líquido está em contato com a parede estomacal de uma maneira generalizada e não restrita a somente um dado fragmento do estômago. Assim, ambas as abordagens são relevantes quando desejamos investigar a microbiota estomacal.

1.3 MICROBIOTA E CÂNCER

Microrganismos têm sido identificados como fatores de risco em cerca de 20% das neoplasias humanas, como por exemplo, os vírus de hepatite B e C no câncer de fígado, o papiloma vírus humano (HPV) no câncer de colo do útero, o parasita *Schistosoma haematobium* no câncer de bexiga, mais recentemente, a bactéria *Fusobacterium nucleatum* com o câncer colorretal, além da própria *H. pylori* no câncer gástrico (Martel et al. 2012; Shang e Liu 2018; Sun et al. 2019a). Apesar do importante papel da *H. pylori* na epidemiologia do AdG, parece improvável que apenas uma bactéria seja associada a este fenômeno. Além disso, apenas 1-3% das pessoas infectadas desenvolvem AdG (Wroblewski et al. 2010). Sendo assim, talvez não só esta bactéria, mas também outros microrganismos, podem estar associadas ao desenvolvimento de doenças gástricas relevantes.

O desequilíbrio da composição da microbiota, ou seja, a alteração dos microrganismos que habitam determinado nicho corporal, processo chamado disbiose, parece ter papel importante no surgimento e estabelecimento de várias doenças, tais como as inflamatórias intestinais, obesidade, diabetes e alergias, já tendo sido descrito até mesmo relação entre disbiose e autismo (Wang e Yang 2013; DeGruttola et al. 2016; Roy e Trinchieri 2017; Cong e Zhang 2018). No chamado *brain-gut axis*, sabe-se que a microbiota intestinal pode alterar o comportamento do hospedeiro, interferindo no seu desenvolvimento social, qualidade de vida, depressão, e até mesmo se relacionar com patologias do sistema nervoso central, como a doença de Parkinson (Hsiao et al. 2013; Sampson et al. 2016; Valles-Colomer et al. 2019). No desenvolvimento do câncer, por sua vez, este aspecto não é diferente, tendo sido amplamente estudado o papel da microbiota no desenvolvimento de diversas neoplasias tais como câncer colorretal, câncer de mama, câncer da cavidade oral, entre outros (Goodman e Gardner 2018; Helms et al. 2019; Thomas et al. 2019).

O intestino, por ser um dos locais do corpo com maior número de bactérias, tem sido uma das regiões mais estudadas e, em nosso grupo já foi observado que pacientes com câncer de reto possuem maior diversidade de bactérias que indivíduos controles, além de gêneros diferentes como os mais abundantes em cada população (Thomas et al. 2016). Já em pacientes com câncer colorretal, já foi relatado que a composição de bactérias do tecido tumoral é diferente da composição do tecido normal adjacente ao tumor do mesmo paciente (Gao et al. 2015). Além disto, como já citado, diversos estudos relataram que a bactéria *Fusobacterium nucleatum* possui importante papel no desenvolvimento e na progressão desta neoplasia (Kostic et al. 2013; Bullman et al. 2017). Ainda, em um estudo com modelo animal foi observado que quando fezes de pacientes com câncer colorretal foram transplantadas em camundongos desprovidos de microbiota (*germ-free*) estes animais tiveram a formação de pólipos, além de outros sinais de tumorigênese, em contraste com os controles (Wong et al. 2017). Em um estudo mais recente com uma coorte de 967 amostras de fezes, foi observado que pacientes com câncer colorretal possuem maior riqueza de bactérias que indivíduos controles, além de uma associação positiva entre o metabolismo de colina produzido pela microbiota intestinal, e bactérias associadas com o consumo de carne vermelha e este tipo de câncer (Thomas et al. 2019). Os estudos aqui destacados e tantos outros indicam que a microbiota intestinal quando em disbiose teria um papel importante no desenvolvimento desta neoplasia.

Se por um lado a microbiota pode estar envolvida no desenvolvimento de neoplasias, por outro lado os microrganismos também podem ter um papel potencialmente positivo, impedindo o crescimento tumoral ao ativar (direta ou indiretamente) o sistema imune que auxilia no combate às células tumorais (Wei et al. 2007; Leschner e Weiss 2010; Song et al. 2018). Além disto, estes microrganismos podem, ainda, estar relacionados com a eficácia dos quimioterápicos, influenciando positiva ou negativamente a resposta de seu hospedeiro a determinadas drogas. Iida et al. (2013) demonstraram que o uso de antibióticos reduz a eficácia de compostos à base de platina contra tumores subcutâneos em camundongos. Por outro lado, foi observado *in vitro* que dez agentes quimioterápicos tiveram eficácia inibida por bactérias, sendo que alguns tiveram sua estrutura química alterada (Lehouritis et al. 2015). Já em um estudo com modelos de câncer de colón, foi observado que gamaproteobactérias podem metabolizar o quimioterápico gencitabina no próprio tecido tumoral, inativando esta droga e reduzindo sua eficácia em poucas horas (Geller et al. 2017). Outro aspecto importante

é a capacidade de alguns quimioterápicos causarem danos à mucosa intestinal, permitindo a penetração de bactérias na lâmina própria, que por sua vez ativam células do sistema imune inato, iniciando uma inflamação local e sistêmica (Dzutsev et al. 2015; Roy e Trinchieri 2017). Deste modo, estes estudos podem contribuir com futuras estratégias que envolvam a alteração da microbiota visando melhor resposta terapêutica.

Considerando o papel da microbiota na modulação da resposta imune, tem sido de grande interesse avaliar a composição da microbiota intestinal e sua possível associação com a imunoterapia. Em 2015, Vétizou et al. mostraram que em melanoma, o tratamento com o anticorpo monoclonal anti-CTLA-4 não era efetivo em camundongos *germ-free*, nem em camundongos tratados com antibióticos, e que a ação antitumoral deste imunoterápico dependia da abundância de bactérias do gênero *Bacteroides* no intestino. Em outro estudo, foi comparado o crescimento tumoral de melanoma em camundongos da mesma linhagem, porém criados em dois laboratórios distintos, e tratados com monoclonais anti-PD-L1. Interessantemente, foi observado que havia resposta diferencial ao tratamento, e que esta resposta estava relacionada à microbiota intestinal dos camundongos nos diferentes laboratórios. Os camundongos que possuíam o gênero *Bifidobacterium* em sua microbiota tiveram melhor resposta ao tratamento, e quando a microbiota do grupo respondedor ao tratamento foi transferida aos animais do grupo não-respondedor, estes últimos passaram a responder ao tratamento (Sivan et al. 2015). Tais dados sugerem que a manipulação da microbiota intestinal pode modular também a resposta à imunoterapia.

Quando olhamos para a microbiota gástrica em AdG, no entanto, observa-se resultados conflitantes. Enquanto alguns estudos não identificaram mudanças significativas na composição taxonômica de microrganismos (Dicksved et al. 2009; Wang et al. 2016), outros reportaram aumento na abundância de certos gêneros como *Lactococcus*, *Veillonella*, *Fusobacterium*, *Leptotrichia* em pacientes com AdG em comparação com controles (Castaño-Rodríguez et al. 2017). Foi observada ainda uma diferente composição da microbiota em populações distintas na Colômbia com maior ou menor risco de desenvolvimento de AdG (Yang et al. 2016), evidenciando que o papel da microbiota no desenvolvimento desta neoplasia permanece pouco claro (Coker et al. 2018). Diante disto, é importante desvendarmos a composição da microbiota gástrica residente e tentarmos identificar bactérias que sejam relevantes nesse microambiente e possam desempenhar papéis na

patogênese e progressão do AdG. Tal conhecimento pode ajudar no desenvolvimento de estratégias potenciais de prevenção, detecção precoce e tratamento da doença no futuro.

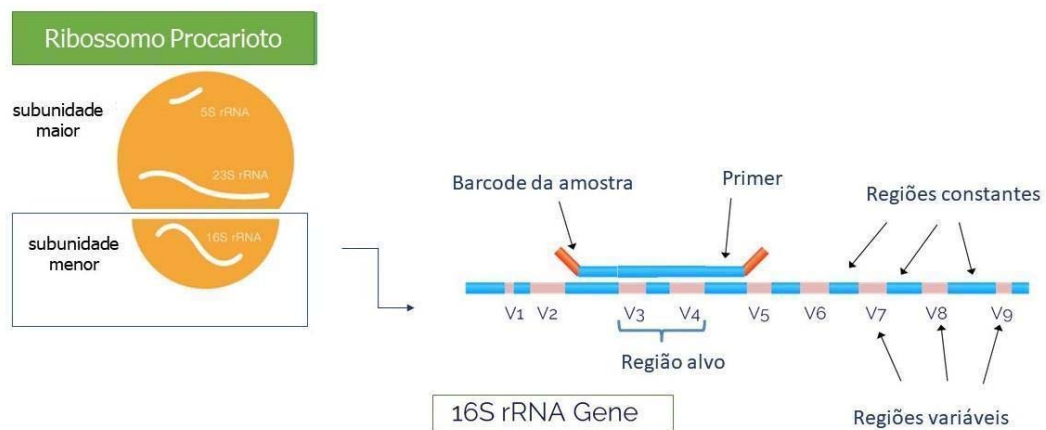
1.4 TÉCNICAS UTILIZADAS PARA ESTUDO DA MICROBIOTA

1.4.1 Estudos baseados na análise de DNA: 16S rRNA e *shotgun*

Estima-se que apenas 1% das bactérias são passíveis de cultivo in vitro (Staley e Konopka 1985; Amann et al. 1995; Torsvik e Øvreås 2002; Martiny 2019). Apesar de ser um consenso antigo, ainda se sabe que a maioria das bactérias não foi cultivada em laboratório, o que não necessariamente significa que elas não possam ser (Stewart 2012; Overmann 2017). Esta dificuldade pode estar relacionada a diversas razões, que vão desde presença das bactérias em baixa frequência em determinada amostra, ou ainda, a dificuldade de se estabelecer os nutrientes necessários para seu crescimento otimizado, principalmente para bactérias fastidiosas, que são mais restritas com relação aos nutrientes e condições do meio, como pH, temperatura, oxigênio, entre outras (Vartoukian et al. 2010). Diante disto, tem entrado em destaque a abordagem de *culturomics*, o uso simultâneo de condições de cultivo diversas e em larga escala. A cultura proveniente de diversos materiais, apesar de ser a prova mais sólida de que um determinado microrganismo está de fato vivo naquele ambiente, ainda é uma técnica relativamente lenta, que vai depender da seleção dos meios corretos, níveis de oxigênio e tempo de crescimento de cada organismo (Lagier et al. 2018; Diakite et al. 2020).

Desta forma, o isolamento do DNA presente em determinado ambiente, seguido de sequenciamento em larga escala permite identificar e caracterizar esta diversidade inexplorada de microrganismos. A área que permite esta caracterização é conhecida como metagenômica, que é o estudo do material genético de uma comunidade de organismos presentes em determinada amostra ambiental (Culligan et al. 2014; Aguiar-Pulido et al. 2016). A amplificação e posterior sequenciamento de marcadores gênicos específicos de microrganismos é um dos métodos mais utilizados em metagenômica, sendo que o gene utilizado vai depender do microrganismo que se quer estudar. Para bactérias utiliza-se como marcador o gene da subunidade 16S do RNA ribossomal (rRNA) (**Figura 2**). Este gene possui dez regiões altamente conservadas entre as diferentes bactérias, permitindo a identificação deste domínio, intercaladas com nove regiões variáveis (V1-V9), que por sua vez permitem a identificação de diferentes táxons bacterianos (Janda e Abbott 2007; Bikel et al. 2015). Este

método, apesar de ser relativamente barato e amplamente utilizado, possui ressalvas, como por exemplo, viés na identificação das bactérias presentes, podendo favorecer alguns táxons e subestimar a presença de outros dependendo da região do gene 16S que for amplificada (Guo et al. 2013; Ibarbalz et al. 2014). A avaliação via 16S em geral também não permite aprofundamento taxonômico em nível de espécie e cepa. Além disso, esta metodologia fornece apenas informações sobre a composição taxonômica das bactérias presentes, não sendo possível analisar diretamente as funções biológicas associadas aos microrganismos identificados.



Fonte: Adaptado de <https://help.ezbiocloud.net/16s-rna-and-16s-rna-gene/> e <https://www.lcsclences.com/discovery/applications/genomics/16s-rna-gene-sequencing-landing/16s-gene/>

Figura 2 - Ribossomo procarioto e gene da subunidade 16S rRNA. O ribossomo procarioto possui a subunidade maior, composta pelas regiões 5S e 23S, além da subunidade menor, a 16S. O gene da subunidade 16S, por sua vez, possui dez regiões constantes (em azul) intercaladas por nove regiões variáveis (em rosa), que servem para identificação do domínio bacteriano e dos diferentes táxons, respectivamente. A região em destaque (V3-V4) foi o alvo dos iniciadores usados neste trabalho para o sequenciamento de *amplicons*.

O estudo da microbiota pelo sequenciamento do DNA total (também conhecido como “*shotgun*”), por sua vez, é uma abordagem que permite que todo o DNA presente na amostra de interesse seja avaliado através de seu isolamento, fragmentação e posterior sequenciamento. Desta forma, é possível, em conjunto com ferramentas de bioinformática, identificar bactérias através do alinhamento das sequências geradas utilizando bancos de dados específicos. De forma semelhante, é possível identificar também archaeas, fungos e vírus que tenham genoma de DNA. Esta abordagem de estudo é bem mais abrangente,

permitindo uma visão da variabilidade funcional da microbiota, visto que todos os genes são sequenciados, e não somente genes específicos para inferir classificação taxonômica (Sharpton 2014).

Apesar das vantagens descritas, alguns desafios no uso de *shotgun* para análises de metagenômica incluem o custo mais elevado desta abordagem (10-100x mais caro, de acordo com a cobertura desejada/necessária), a grande quantidade de dados gerados além de maior complexidade de análise dos mesmos (Mulcahy-O'Grady e Workentine 2016). Outro desafio, ainda, é a proporção de DNA do hospedeiro em relação ao DNA microbiano presente na amostra de interesse. Neste sentido, o *Human Microbiome Project* relatou altos níveis de DNA humano em diversos tecidos e fluidos corporais, como vagina, saliva e garganta (Schmieder e Edwards 2011; *The Human Microbiome Project Consortium* 2012). Como resultado, o uso de estratégia *shotgun* pode gerar, em diversas situações, apenas uma ínfima porcentagem de sequências que sejam de fato correspondentes ao do microbioma, sendo que a maior parte das sequências geradas pode ser proveniente do genoma do hospedeiro.

Diante disto, em análises preliminares feitas em nosso laboratório (dados não publicados – manuscrito **Apêndice 2** - Albuquerque et al.) avaliamos por PCR em tempo real (qPCR) a proporção de DNA bacteriano em relação ao DNA humano presente em amostras biológicas de diversos tipos tumorais. Para isto, foi feita uma curva padrão com diferentes massas de DNA bacteriano e humano utilizando-se como alvos os genes 16S (região V1) e *ACT1* (beta-actina) humano. A partir da equação da reta gerada, foi possível relacionar o número de cópias de cada genoma com o ciclo de amplificação (CT, *cycle threshold*) da qPCR. Como é possível observar na **Figura 3**, apesar da variação, o fluido gástrico possui alta razão de cópias de genomas bacterianos/humano, indicando maior concentração de bactérias neste material, sugerindo ser viável o estudo da microbiota de fluido gástrico por abordagens de *shotgun*. As biópsias de estômago, por sua vez, possuem uma quantidade muito maior de DNA humano em relação ao bacteriano, inviabilizando a análise de metagenômica por sequenciamento de DNA total. Sendo assim, os dados aqui apresentados se baseiam em abordagem 16S para fluido gástrico, biópsia e saliva, além da abordagem de sequenciamento *shotgun* somente para as amostras de fluido gástrico.

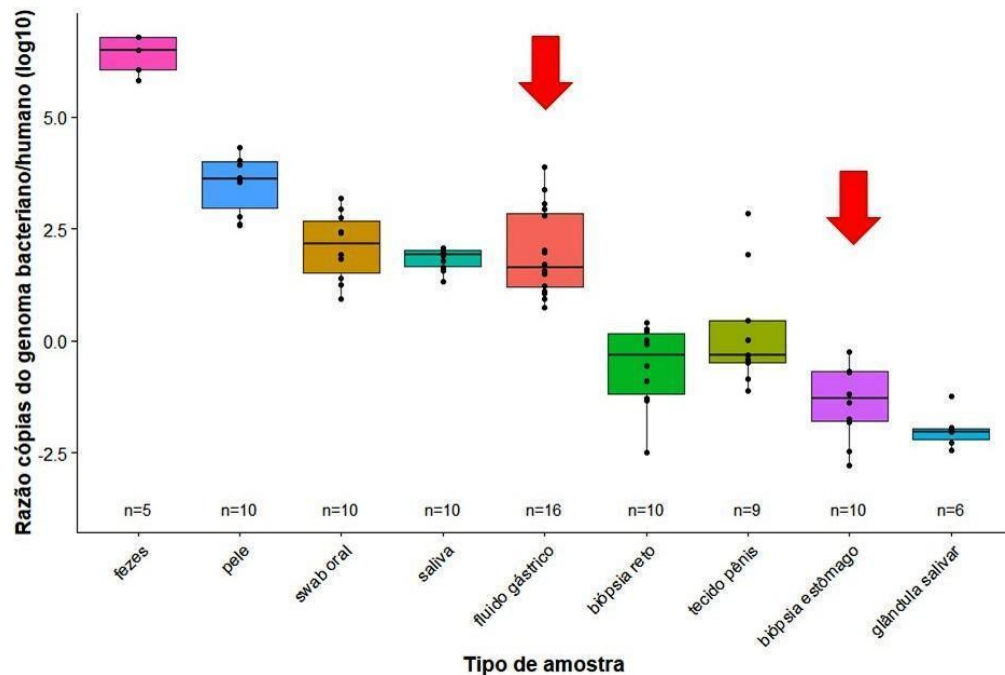


Figura 3 - Razão do número de cópias de genomas bacterianos/humano em diferentes amostras biológicas. Para este cálculo foram realizadas curvas padrão com diferentes massas de DNA bacteriano e humano. A partir da equação da reta gerada foi possível relacionar o CT (*cycle threshold*) de amplificação por PCR em tempo real com o número de cópias do genoma de determinado organismo em diferentes tipos de amostras humanas. As setas vermelhas destacam as amostras provenientes de estômago (fluido gástrico e biópsia).

Apesar das abordagens de sequenciamento por amplificação do gene 16S e por DNA total oferecerem dados altamente relevantes, estas técnicas não resolvem um grande desafio atual: determinar quais microrganismos da microbiota estão vivos e ativos em determinado nicho.

1.4.2 Identificando microrganismos transcricionalmente ativos: Metatranscriptômica

A determinação do conjunto de microrganismos vivos em uma comunidade, especialmente no contexto de populações com atividades metabólicas que possam influenciar de alguma forma o ambiente tumoral, é tema de grande importância. Se a diferenciação entre organismos vivos ou mortos é algo óbvio para seres macroscópicos, o tema é certamente mais desafiador para bactérias, onde em geral é necessário buscar evidências que indiquem a presença de células íntegras capazes de se reproduzir, e que sejam metabolicamente ativas (Emerson et al. 2017). Apesar da cultura de bactérias ser a técnica padrão para confirmar que um microrganismo está de fato vivo em determinado ambiente, esta técnica pode ser

relativamente lenta, pois dependerá do tempo de crescimento de cada organismo, e certamente será limitada e incapaz de indicar todos os microrganismos que estariam vivos em determinada região, já que diferentes microrganismos possuem diferentes requerimentos e distintas condições ótimas de crescimento.

Diante disto, alternativas para identificar microrganismos vivos têm sido elaboradas. Uma possibilidade explorada neste trabalho se baseia nos princípios de integridade de membrana e atividade metabólica. Sabe-se que a membrana celular protege os ácidos nucleicos de degradação, especialmente em ambientes desfavoráveis. Diante da labilidade do RNA, o estudo da microbiota baseado nestas moléculas deve permitir maior direcionamento das análises para microrganismos que estejam vivos na cavidade gástrica no momento da coleta, apontando com maior precisão aqueles que possam estar desempenhando funções relevantes no microambiente estomacal (Von Rosenvinge et al. 2013; Sharpton 2014). Obviamente, microrganismos transientes que caiam na cavidade gástrica de modo contínuo e em grande quantidade, talvez tenham seus RNAs em condição do tipo *steady-state*, fato que implica em comparações cuidadosas entre achados de DNA e RNA para tentar diferenciá-los de organismos de fato residentes. No entanto, é certo que dados das análises de metatranscriptômica (ou seja, o sequenciamento de RNA total, também denominado RNAseq) complementam os dados de metagenômica, elucidando com maior precisão quais bactérias encontradas estão transcricionalmente ativas, permitindo assim inferir também as funções e vias metabólicas ativas deste repertório de bactérias (Rosenvinge et al. 2013; Bashiardes et al. 2016).

É importante ressaltar, porém, que apesar destas abordagens permitirem a identificação de diferentes microrganismos, decidimos focar neste trabalho na análise de bactérias, por serem as mais abundantes da microbiota estomacal. Além disso, outros estudos em curso em nosso grupo avaliam a presença de fungos e vírus neste ambiente.

O RNAseq é uma abordagem mais recente nos estudos de metagenômica, e alguns trabalhos já demonstraram a importância desta. Franzosa et al. (2014) fizeram uso combinado da metatranscriptômica com metagenômica em amostras de fezes e observaram que 59% dos transcritos estavam regulados diferencialmente quando comparados com suas abundâncias nos dados genômicos. Estes autores também verificaram que várias famílias de genes menos abundantes no nível metagenômico podem ser muito ativas transcricionalmente, e vice-versa, mostrando a importância destas análises complementares no estudo do microbioma.

Recentemente, Thorell et al. (2017) analisaram a microbiota gástrica ativa em diferentes estágios da progressão do CG, através de análises de metatranscriptômica, observando a dominância de *H. pylori* na região estomacal, até mesmo em pacientes que eram considerados *H. pylori* negativos por metodologias padrão, além de relatarem uma associação positiva entre a presença de *H. pylori* e os gêneros *Campylobacter*, *Deinococcus* e *Sulfurospirillum*. Neste mesmo trabalho, foram observados altos níveis de expressão de genes envolvidos na regulação do pH e transporte de níquel, comprovando que a *H. pylori* induz genes de virulência relacionados à progressão da doença. Este estudo foi um dos primeiros a avaliar a microbiota gástrica a partir de seu RNA, mostrando que de fato esta abordagem é viável e que pode levar a novas descobertas sobre quais bactérias estão de fato ativas na região estomacal e como elas respondem in vivo a variações no ambiente. Apesar disto, este estudo analisou apenas amostras de biópsia de pacientes com CG, sendo sua grande maioria do sexo feminino (22/ 30 – 73,4%), e sendo todos os indivíduos identificados como *H. pylori* positivos. Diante disto, o trabalho aqui proposto é altamente relevante por se tratar de um estudo focado em uma população distinta de portadores desta neoplasia.

Uma vez que nem todos os microrganismos detectados ao nível de DNA estão abundantes na mesma proporção ao nível de RNA, e sabendo que estes microrganismos podem passar pelo estômago passivamente, pela ingestão de saliva, líquidos e nutrientes (Olsen e Yamazaki 2019), se torna interessante conhecermos também a microbiota oral dos pacientes com AdG e sua relação com a homeostasia estomacal. Trabalhos recentes têm destacado a similaridade desta microbiota com aquela proveniente da saliva (Tsuda et al. 2015), sugerindo ainda que a microbiota oral é mais complexa em pacientes com AdG (Coker et al. 2018; Sun et al. 2018). Uma vez que os microrganismos mais ativos na cavidade estomacal sejam identificados, caso estes tenham forte associação com a composição da microbiota da saliva, abre-se a possibilidade de estudos altamente relevantes e não invasivos, viabilizando estudos em escala ainda mais elevada.

2 OBJETIVOS

2.1 OBJETIVO GERAL

Contribuir para melhor caracterização da microbiota estomacal no câncer gástrico, buscando evidências de quais bactérias teriam maior probabilidade de estarem ativas neste órgão, em contraste com aquelas cujos ácidos nucleicos são encontrados livres, sem maiores comprovações de atividade metabólica local.

2.2 OBJETIVOS ESPECÍFICOS

- 1 Caracterizar a microbiota bacteriana a partir de biópsias tumorais e fluidos gástricos de pacientes com AdG usando a técnica de metagenômica por 16S rRNA;
- 2 Caracterizar a microbiota bacteriana a partir de biópsias tumorais de pacientes com AdG usando a técnica de metatranscriptômica por 16S rRNA;
- 3 Caracterizar a microbiota gástrica por metagenômica shotgun (DNA total) a partir de amostras obtidas do fluido gástrico de pacientes com AdG;
- 4 Caracterizar a microbiota gástrica por metatranscriptômica (RNAseq) a partir de amostras de fluido gástrico de pacientes com AdG;
- 5 Para uma parcela de pacientes, caracterizar a microbiota bacteriana presente na saliva usando a técnica de metagenômica – 16S rRNA;
- 6 Comparar as bactérias presentes na saliva com aquelas identificadas no estômago, buscando aquelas mais abundantes em algum destes locais;
- 7 Utilizar técnicas de cultivo para avaliar microrganismos presentes em fluidos gástricos, seguido de sequenciamento de colônias isoladas, de modo a caracterizar organismos vivos na cavidade gástrica;
- 8 Comparar os dados obtidos a partir das diferentes estratégias, buscando um conjunto de bactérias com maior probabilidade de estarem vivas e ativas na cavidade gástrica de pacientes com AdG.

3 METODOLOGIA

3.1 ASPECTOS ÉTICOS

Este projeto é afiliado ao projeto temático financiado pela Fapesp 14/26897-0 “Genômica de Adenocarcinomas Gástricos do Brasil” em andamento no Laboratório de Genômica Médica do A.C.Camargo Cancer Center e aprovado pelo Comitê de Ética em Pesquisa (CEP) da instituição sob o número 2134/15 (**Anexo 2**). Visto que as análises de metagenômica foram previamente previstas e aprovadas dentro do projeto citado, as amostras biológicas analisadas no presente estudo foram provenientes de pacientes que assinaram o Termo de Consentimento Livre e Esclarecido (TCLE) relacionado a este projeto temático (**Apêndice 1**). Todos os pacientes foram informados do objetivo da pesquisa e convidados a participar de modo voluntário. Este subprojeto também foi aprovado pelo CEP da instituição como projeto afiliado sob o número 2134/15E (**Anexo 2**).

3.2 POPULAÇÃO DE ESTUDO

A fim de reduzir a heterogeneidade do grupo amostral, selecionamos apenas pacientes do sexo masculino, com diagnóstico histopatológico de AdG do tipo difuso, avaliados *como H. pylori* negativos e com idade entre 40-70 anos. Devido à quantidade de abordagens selecionadas para este estudo, inicialmente estipulamos o número amostral de dez pacientes. Durante o projeto tivemos que excluir um paciente pois seu fluido gástrico não foi suficiente para a realização de todos os experimentos. Assim, o estudo incluiu 9 pacientes.

3.3 COLETA DE DADOS CLÍNICOS

Os pacientes que concordaram em participar do estudo responderam um questionário enquanto aguardavam a realização do procedimento endoscópico. Este questionário foi aplicado por entrevistadores treinados que fazem parte do projeto temático em andamento e, para este trabalho, foram selecionadas informações gerais coletadas tais como sexo, idade

e etnia autoatribuída, e informações sobre o uso de medicamentos como antibióticos e relacionados a problemas estomacais (inibidores de bombas de próton). Além disto, ao longo do tratamento foram coletadas informações clínicas, tais como tipo histológico tumoral, local da lesão, tipo de intervenção terapêutica, porcentagem de células tumorais viáveis após a cirurgia e resposta ao tratamento, quando possível.

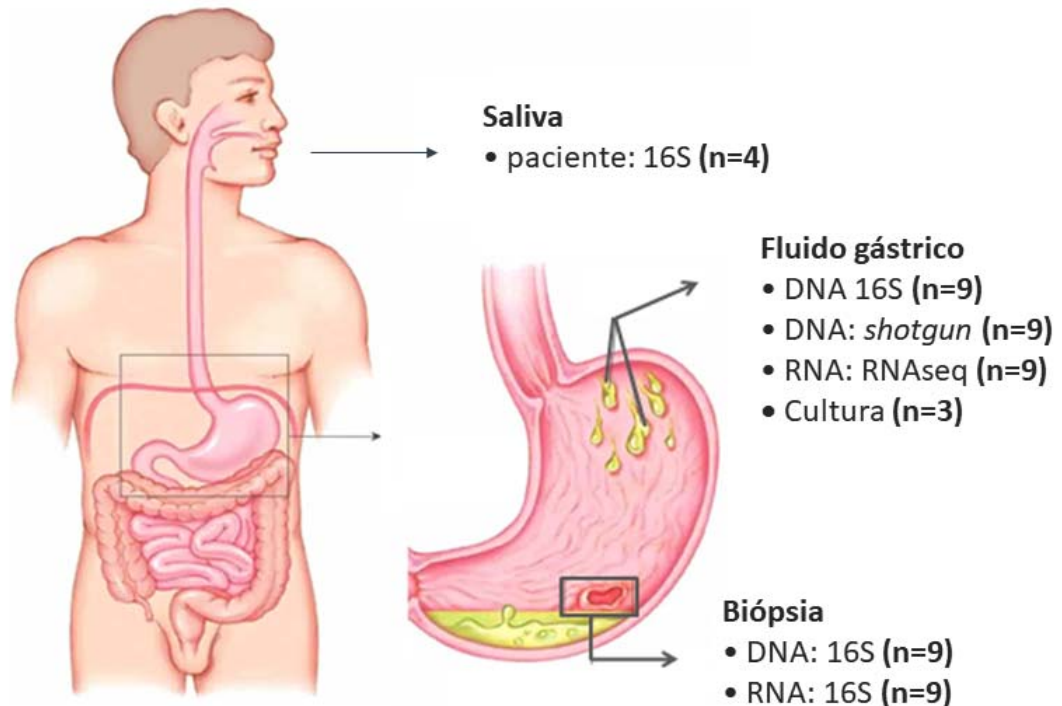
3.4 COLETA E PROCESSAMENTO DE AMOSTRAS BIOLÓGICAS

Amostras de biópsias teciduais do estômago foram coletadas exclusivamente para o projeto temático, em conjunto com amostras de fluido gástrico dos mesmos pacientes, obtidas simultaneamente, durante endoscopia de rotina realizadas durante o exame inicial de confirmação do diagnóstico, antes de qualquer tratamento oncológico. Todos os pacientes receberam instruções de preparo para o exame, que incluiu jejum de líquidos e sólidos por pelo menos seis horas.

Após coletadas, as biópsias foram congeladas e armazenadas no banco de tumores da instituição de acordo com os procedimentos operacionais padrão, até sua solicitação pelos pesquisadores. Já os fluidos gástricos foram imediatamente processados pelo nosso grupo, cujo processo incluiu medição e neutralização do pH, com alíquotas de 1,5ml armazenadas a -80°C com e sem glicerol 20% (quando possível), sendo o restante armazenado a -20°C até seu uso. As amostras de fluido gástrico são derivadas da lavagem do estômago do paciente com água mineral durante o exame endoscópico, a fim de facilitar a visualização da parede estomacal durante o exame pelo médico responsável. As amostras coletadas possuem volumes variáveis (geralmente de 3 até 50ml) e seriam descartadas durante o procedimento endoscópico.

Uma vez que a coleta de saliva foi incluída no estudo quando ele já estava em andamento, foi possível coletá-la para quatro dos nove pacientes. Nestes casos, amostras de saliva espontânea foram coletadas antes do procedimento endoscópico (1-2 ml de saliva não estimulada), e em seguida foram armazenadas a -80°C até o momento da extração de DNA. A **Figura 4** mostra, resumidamente, o planejamento experimental deste projeto.

► 9 pacientes : AdG difuso, sexo masculino, 40-70 anos, *H. pylori* negativo



Fonte: modificado de <<http://www.organsofthebody.com/stomach/stomach-ulcer-diet-and-precautions>>

Figura 4 - Desenho experimental proposto para este projeto. Foram selecionados nove pacientes do sexo masculino com AdG do tipo difuso, diagnosticados como *H. pylori* negativo, com idade entre 40 e 70 anos. Os fluidos gástricos foram coletados para análises de metagenômica por 16S e *shotgun* (DNA) e para análises de metatranscriptômica (RNAseq). As biópsias foram utilizadas para análises de sequenciamento por 16S (DNA e RNA) e a saliva foi utilizada, quando disponível, para análises de sequenciamento por 16S (DNA). Além disso, algumas amostras de fluido gástrico foram armazenadas com glicerol 20% para cultura microbiana.

3.5 EXTRAÇÃO DE ÁCIDOS NUCLEICOS

Para as amostras de fluido gástrico e saliva, 500µl do material coletado foram incubados por 3h em agitação (300 rpm) com 500µl do reagente de digestão *Cell Lysis Solution* (Qiagen) e 30µl de proteinase K a 20µg/µl. A partir deste ponto foi feita a extração de DNA utilizando fenol:clorofórmio:álcool isoamílico 25:24:1 (Merck), que consistiu em:

- 1 Adicionar 500µl do reagente. Homogeneizar por inversão 1min. Centrifugar 10min 13.200 rpm, temperatura ambiente;

- 2 Transferir a fase aquosa para um novo tubo. Adicionar 500µl do reagente. Homogeneizar por inversão 1min. Centrifugar 10min 13.200 rpm, temperatura ambiente;
- 3 Adicionar 500µl de clorofórmio (Merck). Homogeneizar por inversão 1min. Centrifugar 10min 13.200 rpm, temperatura ambiente;
- 4 Transferir a fase aquosa para um novo tubo. Adicionar 800µl de álcool 100% gelado + 80µl de acetato de sódio (3M). Incubar por 1h a -20°C;
- 5 Centrifugar 30min 14.000 rpm, 4°C. Remover sobrenadante;
- 6 Adicionar 1ml de álcool 70% gelado para lavagem do pellet;
- 7 Centrifugar por 5min 14.000 rpm, 4°C. Remover sobrenadante;
- 8 Adicionar 1ml de álcool 70% gelado;
- 9 Centrifugar por 5min 14.000 rpm, 4°C. Remover sobrenadante;
- 10 Secar o pellet a 42°C;
- 11 Ressuspender o pellet em 100ul de água ultrapura (Sigma);
- 12 Incubar a 55°C por 10min. Quantificar o DNA.

Para a extração de RNA de fluido gástrico, 2ml deste material foi concentrado no equipamento *CentriVap Concentration* (Labconco) em um volume final de 500µl, e a partir deste volume foi utilizado o kit de extração *RNeasy PowerMicrobiome* (Qiagen). Já para a extração de ácidos nucleicos das biópsias congeladas, foi utilizado o equipamento *Precellys* (*Bertin Technologies*) para homogeneizar o material, seguido de sua extração com *AllPrep DNA/RNA Mini Kit* (Qiagen), que permite o isolamento de DNA e RNA simultaneamente. O RNA das biópsias foi em seguida tratado com DNase (TURBO DNA-free Kit – ThermoFisher).

Após as extrações, todas as amostras foram quantificadas através do equipamento *Qubit* (ThermoFisher), e as amostras de RNA dos fluidos gástricos tiveram sua integridade avaliada através do equipamento *Bioanalyzer* (Agilent Technologies).

3.6 PCR EM TEMPO REAL

3.6.1 cDNA de fluidos gástricos

As amostras de RNA extraídas dos fluidos gástricos foram convertidas em cDNA para que fosse possível analisá-las por qPCR. Para isto, 8µl de RNA foram utilizados como molde

para conversão em cDNA com a enzima *SuperScript™ III First-Strand Synthesis System* (ThermoFisher) utilizando-se hexâmeros randômicos (N6). Cabe notar que aqui indicamos o volume de RNA usado, e não a massa, pois na maioria dos casos a quantidade presente neste volume estava abaixo do limite de detecção, o que impediu a quantificação. As condições e reagentes utilizados na RT-PCR (*reverse-transcriptase polymerase chain reaction*) se encontram na **Tabela 1**.

Tabela 1 - Condições de conversão do RNA de fluidos gástricos para cDNA utilizados neste estudo

Mix 1	volume (µl)
50ng/µl hexômeros randômicos	1
10 mM dNTP mix	1
até 5ug total RNA	8
Termociclador	
65°C 5 min	
4°C 2 min	
Mix 2	volume (µl)
10X RT Buffer	2
25 mM MgCl ₂	4
0.1 M DTT	2
RNaseOUT (40U/µl)	1
SuperScript III RT (200U/µl)	1
Termociclador	
25°C 10min	
50°C 50min	
85°C 5min	

Após a conversão do RNA total em cDNA, foi feita uma qPCR para avaliar a presença de DNA bacteriano e humano nas amostras de fluido gástrico. Para isto, 1µl de cDNA foram utilizados para cada reação, em conjunto com a enzima *Fast SYBR Green Master Mix* (ThermoFisher). Os iniciadores utilizados para a amplificação foram específicos para a região

V1 do gene 16S (F: 5'-AGAGTTTGATCMTGGCTCAG-3' e R: 5'-TTACTCACCCGTICGCCRCT-3'), e para um fragmento do gene humano beta actina (ACT1) (F: 5'-CCATCTACGAGGGGTATGC-3' e R: 5'-GGTGAGGATCTTCATGAGGTA-3'), ambos os iniciadores geram *amplicons* com cerca de 100bp. A reação final e condições estabelecidas são apresentadas nas **Tabelas 2 e 3**. Como controles foram utilizados DNA de *Escherichia coli* (controle positivo) e água ultrapura (Sigma) como controle negativo.

Tabela 2 - Condições de amplificação dos genes 16S (região V1) e beta-actina utilizadas neste estudo.

Reação	Volume (μl)
16S (V1)	
Iniciador V1F (5μM)	0,4
Iniciador V1R (5μM)	0,4
Fast SYBR Green 2x	5
Amostra (2ng) + H ₂ O	4,2
Beta-actina	
Iniciador ACTBF (5μM)	0,4
Iniciador ACTBR (5μM)	0,4
Fast SYBR Green 2x	5
Amostra (2ng) + H ₂ O	4,2
Volume final	10

Tabela 3 - Condições de termociclagem para amplificação dos genes 16S (região V1) e beta-actina utilizadas neste estudo.

Ciclagem		
Temperatura	Tempo	Ciclos
95°C	20s	1x
95°C	3s	40x
60°C	30s	

3.6.2 DNA de fluidos gástricos e biópsias

De forma similar às amostras de cDNA, 2ng de DNA foram utilizados para serem amplificados por qPCR para que fosse possível avaliar a proporção de DNA bacteriano e humano nas amostras.

3.6.2.2 Construção de curvas padrão

Com o intuito de estimar o número de genomas (humano e bacteriano) nas amostras, foi feita uma curva padrão para cada genoma. Para isto, utilizamos DNA humano comercial (*Human Genomic DNA* – Promega), e um mix composto de cinco espécies de bactérias e duas de fungo (*ZymoBIOMICS™ Microbial Community DNA Standard* - Zymo Research). Começando com 5ng foram preparadas diluições seriadas destes materiais, que foram amplificados por qPCR conforme condições descritas na seção anterior. Os CTs médios de amplificação correspondentes a cada massa foram utilizados para a realização da curva padrão.

O número de cópias por ng de DNA para cada organismo, por sua vez, foi calculado conforme demonstrado na **Figura 5**. Para os cálculos, o tamanho do genoma humano diploide foi considerado 6,4Gb, já para o genoma bacteriano, o tamanho estimado foi de 3,95Mb (valor médio obtido a partir de 16.031 genomas bacterianos depositados no *National Center for Biotechnology Information, National Institutes of Health, Genome Browse by Organism*. Assim, curvas padrão correlacionando a média dos CTs com o número total de cópias do genoma foram construídas. Cada curva possui duas linhas, representando o CT esperado *versus* o observado de cada genoma. O CT esperado foi construído a partir do menor valor de CT obtido, e os demais pontos foram calculados assumindo que a cada aumento de um CT, deveria haver uma redução em 50% na massa de DNA amplificado.

$\text{número de cópias} = \frac{\text{massa (ng)} * 6.022 \times 10^{23}}{\text{tamanho} * 1 \times 10^9 * 650}$
Onde: 6.022×10^{23} número de moléculas/mol = número de Avogadro 650 = peso médio de um par de base (bp) 1×10^9 = para converter g to ng tamanho = tamanho do genoma de cada organismo

Fonte: <http://cels.uri.edu/gsc/cndna.html>

Figura 5 - Cálculo utilizado para a identificação do número de cópias de DNA genômico em 1ng de amostra.

3.7 SEQUENCIAMENTO DE DNA TOTAL DE FLUIDOS GÁSTRICOS (*shotgun*)

As amostras de fluido gástrico foram sequenciadas utilizando-se a abordagem de DNA total (*shotgun*). Para isto, 500ng de DNA foram utilizados para a construção de bibliotecas através do kit *Nextera DNA Flex* (Illumina). As bibliotecas amplificadas foram avaliadas no *Bioanalyzer* (Agilent) e quantificadas por qPCR através do kit *Kapa Library Quantification* (Illumina). As amostras foram então sequenciadas utilizando-se o reagente *High Output Cartridge v2* (300 ciclos) pela plataforma NextSeq 500 (Illumina) que pode gerar até 400 milhões de *reads* por corrida.

3.7.1 Processamento e filtragem das amostras sequenciadas

As amostras sequenciadas foram processadas pelo grupo de Biologia Computacional do A.C.Camargo Cancer Center, cujo *pipeline* foi adaptado em conjunto com nosso grupo (Laboratório de Genômica Médica). Este processamento consistiu na utilização do *pipeline* *samsa2* com modificações (Westreich et al. 2018; Treiber et al. 2020), como o uso da ferramenta *KneadData*, para (1) unir as *reads* pareadas, (2) remoção de *reads* de baixa qualidade utilizando o programa *Trimmomatic* (Bolger et al. 2014) e (3) remoção de DNA contaminante do hospedeiro determinado após alinhamento com o genoma humano referência usando o *software* *Bowtie2* (Langmead e Salzberg 2012). A seguir as *reads* foram anotadas junto ao banco de dados do *RefSeq* (Tatusova et al. 2014), através do algoritmo DIAMOND (Buchfink et al. 2015) para posterior classificação taxonômica.

3.8 SEQUENCIAMENTO DE RNA TOTAL DE FLUIDOS GÁSTRICOS (RNAseq)

As amostras de fluido gástrico tiveram seu RNA total sequenciado. Para isto, todo o RNA extraído foi utilizado para construção de biblioteca utilizando-se o kit *Ion Total RNA-seq Kit v2* (ThermoFisher), e as bibliotecas amplificadas foram quantificadas no *Bioanalyzer* (Agilent). A PCR em emulsão foi realizada no equipamento *Ion OneTouch 2 System*, seguida do enriquecimento no *Ion OneTouch ES*, em ambas as etapas o kit utilizado foi o *Ion PI Hi-Q OT2 200* (ThermoFisher). O sequenciamento das amostras foi realizado no equipamento *Ion Proton* com o kit *Ion PI Hi-Q Sequencing 200*, com o *Ion PI Chip V3*, que gera entre 60 a 80 milhões de *reads* por corrida.

3.8.1 Processamento e filtragem das amostras sequenciadas

As amostras sequenciadas foram processadas pelo grupo de Biologia Computacional do A.C.Camargo Cancer Center, cujo *pipeline* foi adaptado em conjunto com o nosso grupo. De forma similar ao pipeline utilizado nas análises de metagenômica (DNA total), este processamento consistiu na utilização do *pipeline* *samsa2* (Westreich et al. 2018), que é específico para análises de metatranscriptômica, com algumas modificações, conforme descrito na seção 3.8.1. Adicionalmente para estas amostras, antes da anotação no banco de dados, foi realizada a filtragem de RNA ribossomal bacteriano através do *sortMeRNA* (Kopylova et al. 2012).

3.9 CULTIVO, SEQUENCIAMENTO E IDENTIFICAÇÃO DE COLÔNIAS ISOLADAS A PARTIR DE AMOSTRAS DE FLUIDOS GÁSTRICOS

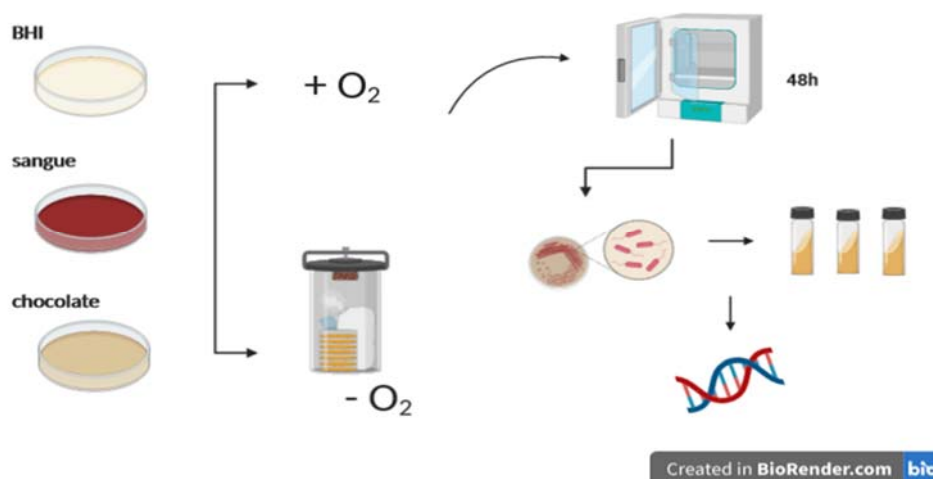
3.9.1 Seleção de fluidos gástricos, meios de cultura e condições de cultivo

Para o cultivo de fluidos gástricos, foram selecionadas as amostras que possuíam este material armazenado com glicerol 20% no freezer -80°C, sendo somente três dos nove pacientes. Foram utilizados os meios sólidos enriquecidos e não-seletivos ágar *Brain Heart Infusion* - BHI (Newprow), sangue (Newprow) e chocolate (Newprow), na presença e ausência de oxigênio. Para isto, 15µl de fluido gástrico foram adicionados em cada placa de petri e espalhados pela técnica de esgotamento. As placas cultivadas na presença de oxigênio foram

colocadas diretamente na estufa à 37°C após plaqueamento, enquanto aquelas cultivadas sem a presença de oxigênio foram colocadas em uma jarra específica para anaerobiose (Probac) junto com o reagente Anaerobac (Probac), um gerador de atmosfera com teor reduzido de oxigênio e maior concentração de CO₂. Para confirmarmos a ausência de oxigênio nas jarras, utilizamos fitas indicadoras que acompanham o produto, e que mudam de coloração de acordo com presença de oxigênio no microambiente. Cada experimento foi realizado em duplicata, e placas sem fluido gástrico foram também incubadas, como controles negativo.

3.9.2 Seleção das colônias e extração de DNA

O crescimento bacteriano foi avaliado nas placas após 24h e 48h. Para as amostras nas quais foi observado crescimento, colônias com diferentes características em cada placa foram isoladas e transferidas para 5ml de meio líquido Luria Bertani (LB). As amostras foram incubadas em agitação de 200 rpm a 37°C por 48h. Após incubação, as culturas foram centrifugadas por 14.000 rpm 2 min, e o pellet, com as bactérias, seguiu para extração de DNA com o reagente fenol:clorofórmio:álcool isoamílico 25:24:1 (Merck) como descrito na seção 3.5. A **Figura 6** mostra resumidamente o planejamento experimental desta etapa do projeto.



Fonte: Figura criada com BioRender.com

Figura 6 - Visão geral da metodologia para cultivo *in vitro* de microrganismos presentes nos fluidos gástricos. Os meios de cultura BHI, sangue e chocolate foram selecionados para cultivo microbiano de fluidos gástricos. Cada meio contendo 15µl do material foi armazenado diretamente na estufa (na presença de oxigênio) ou em uma jarra específica para geração de anaerobiose. Após 48h, as placas no qual foi observado crescimento tiveram diferentes colônias isoladas para crescimento em meio líquido LB. Após 48h foi realizada a extração de DNA destas colônias.

3.9.3 Sequenciamento e identificação de microrganismos

O DNA extraído das colônias foi fragmentado por 330 s utilizando o ultrasonicador S220 (Covaris). A construção das bibliotecas foi realizada seguindo o protocolo *Ion Xpress Plus gDNA Fragment Library Preparation* (ThermoFisher), com posterior quantificação destas bibliotecas por qPCR (*Ion Library Quantification kit* – ThermoFisher). A PCR em emulsão e o enriquecimento foram realizados no equipamento Ion Chef e sequenciadas no Ion S5 com o reagente Ion 540 Kit-Chef (ThermoFisher). O chip utilizado neste sequenciamento foi o Ion 540, com capacidade para gerar entre 60 a 80 milhões de *reads* por corrida.

Para identificação dos microrganismos encontrados, as *reads* foram anotadas através do classificador taxonômico kraken2, com as condições padrões do programa, visto que este banco identifica procariotos e eucariotos (Wood et al. 2019).

3.10 SEQUENCIAMENTO DE AMPLICONS DO GENE 16S rRNA E DE SEU TRANSCRITO

As amostras de DNA obtidas de fluido gástrico, biópsia e saliva tiveram as regiões V3-V4 do gene 16S do RNA ribossomal (rRNA) bacteriano amplificadas por PCR convencional. Para isto, foram padronizadas diferentes massas de DNA para cada material a ser amplificado, sendo 80ng para amostras de fluido gástrico, 50ng para biópsias e 20ng para saliva. Como já estabelecido no laboratório, quando uma amostra não era amplificada com a massa padrão, na PCR seguinte eram utilizados o dobro de massa e a metade, para a possibilidade de a não-amplificação ser devido a DNA bacteriano insuficiente na amostra ou alguma possível inibição da reação.

Os iniciadores utilizados para amplificação possuem adaptadores (sublinhados) adequados para sequenciamento na plataforma Illumina e geram *amplicons* com tamanho médio final de cerca de 500 pares de bases (U341F: 5' CACTCTTTCCCTACACGACGCTCTTCCGATCTCTACGGGSGCAGCAG 3' e 806R: 5' GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTGGACTACHVGGGTWTCTAAT 3') (Takahashi et al. 2014). A enzima *KAPA HiFi HotStart ReadyMix 2X* (Kapa Biosystems) foi selecionada para amplificação de DNA do fluido gástrico e saliva. Já a enzima *KAPA2G Robust HotStart ReadyMix PCR Kit* (Kapa Biosystems) foi utilizada para as biópsias, devido à dificuldade de amplificação do 16S neste material. A reação final e condições estabelecidas são apresentadas nas **Tabelas 4 e 5**. Como controles utilizamos DNA de *Escherichia coli* (controle positivo) e água ultrapura

(Sigma) como controle negativo. De forma similar, o RNA extraído das biópsias foi utilizado para síntese de cDNA utilizando iniciadores randômicos (*SuperScript III RT* – ThermoFisher) seguido de amplificação das regiões V3-V4 do gene 16S rRNA da mesma forma que os demais materiais.

Tabela 4 - Condições de amplificação do gene 16S (regiões V3-V4) utilizadas neste estudo. A enzima e a massa de DNA utilizada na reação dependeram do material a ser amplificado.

Reação	Volume (μl)
Iniciador U341F (5μM)	3,5
Iniciador 806R (5μM)	3,5
Mix da enzima e tampão	17,5
Amostra + H ₂ O	10,5
Volume final	35μl

Tabela 5 - Condições de termociclagem para amplificação do gene 16S (regiões V3-V4) utilizadas neste estudo. A quantidade de ciclos variou dependendo do tipo de amostra, para fluido gástrico e saliva foram utilizados 28 ciclos, já para as biópsias, 32 ciclos.

Etapas	Temperatura	Tempo	Ciclo(s)
Desnaturação inicial	95°C	2min	1x
Desnaturação	95°C	20s	28x (ou 32x)
Anelamento	54°C	15s	
Extensão final	72°C	5min	1x

Os *amplicons* gerados foram purificados com *beads* magnéticas (AMPure XP - Beckman Coulter), quantificados no *Qubit* (ThermoFisher) e sequenciados em serviço terceirizado (*Neoprosecta Microbiome Technologies* - Florianópolis, Santa Catarina, Brasil) que realizou a construção das bibliotecas e sequenciamento de acordo com protocolo estabelecido por eles. Basicamente, para a construção de cada biblioteca uma nova amplificação foi realizada em triplicata usando 5ng de *amplicon* purificado por reação, com a enzima *Taq Platinum* (Invitrogen) e *barcodes* específicos para cada amostra (condição de amplificação: 95°C por 5 min, seguindo de 10 ciclos de 95°C por 46s, 66°C por 30s e 72°C por 45s, e uma extensão final

de 72°C por 2min). As triplicatas foram juntadas e purificadas com o *Ampure XP Beads* (Beckman Coulter), e em seguida foi feito o *pool* das bibliotecas usando volume iguais de cada amostra (Christoff et al. 2017). A quantificação do *pool* foi feita por qPCR com o kit *KAPA Library Quantification* (KAPA Biosystems). Para o sequenciamento foi utilizado o kit *MiSeq Reagent v2* (500 ciclos) na plataforma MiSeq (Illumina), visando gerar ≥ 50.000 *reads* por amostra.

3.10.1 Processamento e filtragem das amostras sequenciadas

As amostras sequenciadas foram processadas pelo grupo de Biologia Computacional do A.C.Camargo Cancer Center, cujo pipeline foi estruturado em conjunto com nosso grupo. Este processamento consistiu em: (1) remoção dos adaptadores; (2) remoção das *reads* mapeadas no genoma humano; (3) importação das *reads* restantes para o programa *Qiime2* (Bolyen et al. 2019); (4) realização da junção das *reads* pareadas e filtro de qualidade e; (5) controle de qualidade para remoção de ruídos das *reads* através do *Deblur* (Amir et al. 2017). As sequências foram então clusterizadas como *Operational Taxonomic Units* (OTUs) quando apresentavam 97% de similaridade entre elas, seguida de busca de similaridade contra sequências presentes no banco de dados SILVA (Quast et al. 2013) visando a classificação taxonômica das OTUs.

A partir das análises das curvas de rarefação construídas para cada tipo de amostra (**Figura 7**), selecionamos aquelas com pelo menos 1 mil *reads* após aplicação dos filtros de qualidade para prosseguirmos as análises de caracterização da microbiota. Tal padronização foi realizada anteriormente por nosso grupo dentro do projeto temático.

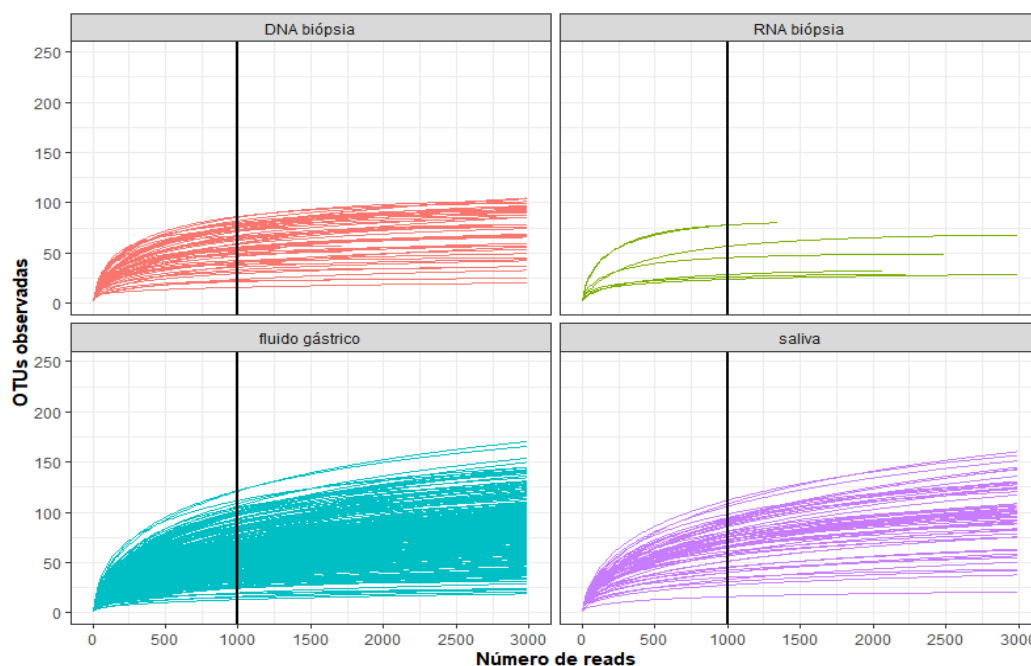


Figura 7 - Curva de rarefação de *amplicons* derivados de biópsias (RNA e DNA), fluidos gástricos e saliva. Os dados se referem a 439 amostras já sequenciadas dentro do contexto do projeto temático: “Genômica de Adenocarcinomas Gástricos” em andamento no Laboratório de Genômica Médica do A.C.Camargo Cancer Center. As curvas demonstram que com pelo menos 1.000 *reads* já é possível atingir saturação aceitável para cobrir a diversidade destas amostras.

Com o intuito de avaliar o nível taxonômico mais profundo possível, os resultados gerados foram baseados na classificação a nível de gênero. As *reads* que foram classificadas nível como “*Ambiguous_taxa*”, “*uncultured*”, “*uncultured bacterium*”, ou “*unidentified*” não foram analisadas neste momento, mas são de grande valia para estudos posteriores, notadamente pela possibilidade de identificação e análises de novos microrganismos.

3.11 ANÁLISES ESTATÍSTICAS

Para as análises de metagenômica e metatranscriptômica foi utilizado o pacote *phyloseq* (McMurdie e Holmes 2016) no programa *RStudio*. Assumindo a não-normalidade da amostra, para avaliar a existência de diferenças estatisticamente significativas na diversidade alfa (*Observed*, *Shannon* e *Simpson*) foi utilizado o teste *U de Mann-Whitney* para amostras não pareadas e o teste de *Wilcoxon* para amostras pareadas, e em relação à diversidade beta

(índice de *Bray-Curtis* e distâncias *Unifrac unweighted* e *weighted*) foi utilizado o teste multivariado *Analysis of Similarities* (ANOSIM). Para avaliar a diferença entre as abundâncias relativas das amostras, foi realizado o teste *Kruskal-Wallis* para comparação múltipla quando necessário, seguido do teste de *U de Mann-Whitney*, enquanto que para as amostras pareadas foi utilizado o teste de *Wilcoxon*. Por fim, para as análises de correlação, foi utilizado o teste de correlação de *Spearman*. Para todos os testes estatísticos, foi estabelecido um nível de significância de 5% ($p < 0,05$).

4 RESULTADOS E DISCUSSÃO

4.1 CARACTERÍSTICAS DA POPULAÇÃO ESTUDADA

Inicialmente selecionamos dez pacientes portadores de AdG, recrutados no contexto do projeto temático gástrico e que se enquadraram no perfil selecionado para este estudo. Escolhemos para o estudo apenas pacientes do sexo masculino, com idade entre 40 e 70 anos, diagnosticados com tumores gástricos do tipo difuso e negativos para infecção ativa por *H. pylori*, conforme coloração por *Giemsa*. Tais critérios de seleção visaram uma maior homogeneidade da população estudada e se encaixam em parâmetros mais comuns em nossa casuística. Após a seleção de casos e início dos estudos a amostra de um dos pacientes mostrou-se em quantidade insuficiente de fluido gástrico impedindo a realização dos os experimentos propostos. Diante disso, seguimos com o estudo de nove pacientes. Todos os pacientes possuíam biópsia coletada exclusivamente para o nosso projeto e além disto, para quatro destes pacientes também foi possível coletar saliva para análise.

Conforme observado na **Tabela 6** a maioria dos pacientes tem entre 61-70 anos (44,4%; n=4), seguido de pacientes entre 40-50 anos (33,3%; n=3) e de pacientes entre 51-60 anos (22,2%; n=2), além disto, a maioria se auto atribuiu como branco (77,8%; n=7). Com relação ao uso de medicamentos, 33,3% (n=3) relataram o último uso de antibióticos há mais de um ano ou não se lembravam (33,3%; n=3). Um paciente (paciente #2) informou ter feito o uso deste medicamento há menos de uma semana (11,1%) da coleta das amostras. Quanto ao uso de inibidores de bomba de próton, 55,6% dos pacientes relataram fazer uso desta medicação no momento do diagnóstico (n=5). Quando mensurado o pH do fluido gástrico destes pacientes, a maioria constava como pH neutro (55,6%; n=7), seguido de pH ácido (33,3%; n=3) e básico (11,1%; n=1).

Quando observamos dados específicos do tumor, 44,4% dos pacientes possuem lesão inicial no corpo do estômago (n=4), 33,3% seguido de lesão no esôfago distal (n=3) e 22% na região cardia/fundo ou antro (um paciente cada). Todos os pacientes fizeram terapia quimioterápica neoadjuvante, e 44,4% destes pacientes tiveram excelente resposta, restando entre 0-10% de células tumorais viáveis na peça cirúrgica, sendo classificados como respondedores de acordo com o critério de Becker 1 (In et al. 2018). Cabe notar que esta

proporção de respondedores não reflete a rotina observada na rotina de pacientes com câncer gástrico. Por exemplo, entre 159 pacientes de nosso projeto temático que realizaram QTneo entre 2015 a 2020, apenas 29 (18%) foram classificados como respondedores, números que são concordantes com a literatura que reporta que apenas cerca de 10-30% dos pacientes com CG respondem a QTneo (Becker et al. 2003; Mansour et al. 2007; Lorenzen et al. 2013; Achilli et al. 2017).

Tabela 6 - Dados clínicos gerais dos pacientes deste estudo.

Variáveis	N	%
Gênero		
Masculino	9	100,0
Idade ao diagnóstico		
40 - 50 anos	3	33,3
51 - 60 anos	2	22,2
61 - 70 anos	4	44,4
Etnia auto atribuída		
Asiático	1	11,1
Branco	7	77,8
Pardo	1	11,1
Classificação de Lauren		
Difuso	9	100,0
Local da lesão		
Antro	1	11,1
Cardia/fundo	1	11,1
Corpo	4	44,4
Esôfago distal	3	33,3
<i>H. pylori</i>		
Negativo	9	100,0

Cont/ Tabela 6

Variáveis	N	%
Uso de antibiótico pela última vez		
Uma semana	1	11,1
Entre 1-6 meses	2	22,2
Mais de um ano	3	33,3
Não se lembra	3	33,3
Uso de inibidores de bomba de próton		
Sim	5	55,6
Não	4	44,4
pH fluido gástrico		
ácido (1-6)	3	33,3
neutro (7)	5	55,6
básico (8-14)	1	11,1
Tratamento neoadjuvante		
Sim	9	100,0
% células tumorais na peça cirúrgica		
0-10%	4	44,4
11-60%	3	33,3
61-100%	2	22,2
Resposta a neoadjuvância		
Respondedor	4	44,4
Não respondedor ou respondedor parcial	5	55,6

4.2 EXTRAÇÃO DE ÁCIDOS NUCLEICOS

Após a extração de DNA e RNA das amostras deste estudo, elas foram quantificadas no *Qubit*, a **Tabela 7** mostra a quantificação total das mesmas. De modo geral as amostras de DNA de todos os materiais renderam boa massa total, sendo que para fluido gástrico a mediana foi de 1.000ng totais, para biópsia 9.056ng e para as amostras de saliva 2.150ng. Com relação às extrações de RNA, a mediana da massa total obtida nas extrações de biópsia foi de

136ng, porém, em duas das amostras a quantidade de RNA esteve abaixo do limite de quantificação. Já para o RNA de fluido gástrico, apenas 2/9 amostras deram quantificações detectáveis. Aqui é importante ressaltar que para a extração deste material já foi necessário utilizar um grande volume de fluido gástrico, e como para algumas amostras o volume total deste material era limitante, estabelecemos o máximo de 2ml para a extração (*versus* apenas 500µl para a extração de DNA do mesmo material).

Tabela 7 - Quantificação de ácidos nucleicos (RNA e DNA) das amostras utilizadas neste estudo.

Paciente	Fluido gástrico		Biópsia		Saliva
	DNA total	RNA total	DNA total	RNA total	DNA total
1	1.160	397	10.800	232	NA
2	2.540	0	23.700	768	NA
3	1.140	0	24.000	121,2	NA
4	1.010	0	712	0	NA
5	208	0	4.720	135,6	NA
6	978	0	7.440	111,2	2.260
7	1.330	570	3.820	248	2.020
8	101	0	13.600	0	4.560
9	640	0	7.260	644	2.040
média	1.012	107	10.672	251	2.720
mediana	1.010	0	9.056	136	2.150

4.3 ANÁLISE DO RNA EXTRAÍDO DE FLUIDOS GÁSTRICOS: INTEGRIDADE E PCR EM TEMPO REAL

Como a maioria das amostras de RNA não deram quantificação suficiente para serem detectadas no *Qubit*, foi feito o cDNA deste material e uma qPCR foi realizada para algumas amostras para avaliar a presença destes ácidos nucleicos no fluido gástrico. Através da **Tabela 8** é possível observar que apesar da baixa quantificação de RNA, todas as amostras amplificaram para o *primer* bacteriano (16-região V1), com mediana de CT de 20,32, enquanto

que para o *primer* humano (beta-actina), duas das seis amostras não amplificaram (identificados como *Undetermined*), indicando baixa quantidade de RNA humano (ou a baixa expressão do gene da beta-actina) neste material.

Tabela 8 - Média do CT de amplificação bacteriano e humano nas amostras de fluido gástrico (cDNA).

Paciente	CT 16S (bacteria)	CT actina (humano)
1	19,45	<i>Undetermined</i>
2	20,10	39,59
4	20,30	30,80
5	20,35	<i>Undetermined</i>
6	23,45	32,08
9	23,00	30,76
média	21,11	33,31
mediana	20,32	31,44

Após a confirmação da presença de RNA bacteriano nas amostras de fluido gástrico, sua integridade foi avaliada através do equipamento *Bioanalyzer* (Agilent Technologies), com o respectivo cálculo do seu número RIN (*RNA integrity number*), que permite inferir a qualidade do RNA eucariótico total, baseado em um *score* de 1 a 10, sendo 10 considerado a amostra intacta (Mueller et al. 2016). Para as amostras de RNA provenientes do fluido gástrico, o RIN ficou em torno de 2,5 (**Figura 8**), e apesar de estarmos interessados no RNA bacteriano, e não no eucariótico, a informação de qualidade deste é relevante, visto que ambos estão no mesmo ambiente e, a princípio, sujeitos ao mesmo processo de degradação. Uma limitação importante, porém, indicada pelos dados de amplificação do gene da beta-actina, é a praticamente ausência de RNA humano - incluindo seus RNAs ribossomais usados como referência para o cálculo, o que deixa claro uma certa inadequação desta abordagem.

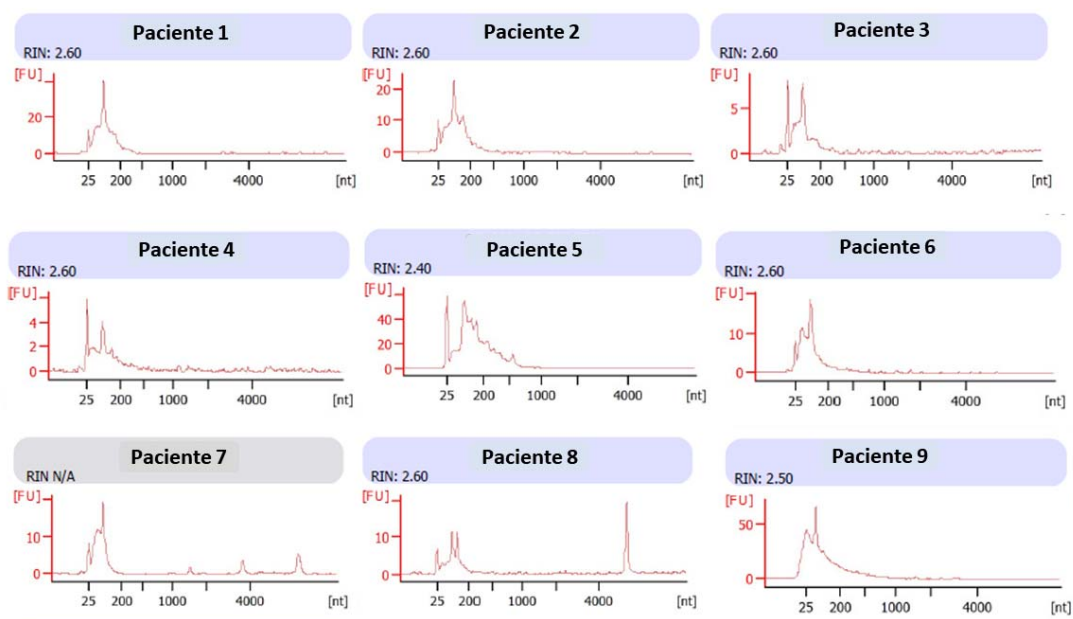


Figura 8 - Análise da integridade do RNA de amostras de fluido gástrico. O número de integridade de RNA (RIN) foi avaliado pelo *software* da *Agilent Technologies*, no equipamento *Bioanalyzer*. Esta abordagem tem como base a intensidade relativa dos picos de RNAs ribossomais humanos, que não são visíveis para nenhuma amostra. É possível observar que todas as amostras tiveram RINs calculados ao redor de 2,5 e pico entre 25 e 200 nt, sendo consideradas amostras altamente degradadas.

Apesar das limitações da nossa abordagem, decidimos seguir com o sequenciamento das amostras, visto que houve amplificação de cDNA bacteriano para estas, e um alto CT humano, o que poderia ser favorável para nosso propósito visto que o RNA deste material será sequenciado pela técnica de RNAseq. Além disto, apesar da possível degradação visualizada neste material, o protocolo utilizado para construção de bibliotecas requer a fragmentação dos RNAs, o que não impede o sequenciamento.

4.4 QUANTIFICAÇÃO DO NÚMERO DE CÓPIAS DE GENOMAS BACTERIANOS x GENOMAS HUMANOS POR PCR EM TEMPO REAL

4.4.1 Construção de curvas padrão

Após termos extraído DNA das amostras de interesse, partimos para a construção das curvas padrão correlacionando o número de cópias dos genomas humano e bacteriano, com

seus respectivos CTs de amplificação na qPCR, conforme trabalho de nosso grupo (**Apêndice 2**).

Para isto, iniciamos a curva com 5ng de DNA humano, seguido de cinco diluições seriadas, até que a beta-actina não fosse mais detectada na qPCR. De forma similar, iniciamos a curva para o DNA bacteriano com 5ng da *mock* utilizada (o que equivale a 4,8ng de DNA bacteriano, pois a mesma continha uma pequena massa de DNA fúngico), seguido de diluições seriadas até que o CT de amplificação ficasse em torno de 32, visto que para o DNA bacteriano o controle negativo começa a amplificar neste CT (já que a enzima utilizada na qPCR contém DNA bacteriano, esta amplificação tardia do controle negativo já é esperada). A **Tabela 9** exemplifica as massas utilizadas para a construção das curvas padrão, além do número de cópias calculado de cada genoma em 1ng de amostra conforme cálculo exemplificado anteriormente na **Figura 5**.

Tabela 9 - Massas utilizadas, CT de amplificação e número de cópias dos genomas humano e bacteriano utilizados para construção de curvas padrão.

Curva padrão DNA humano					
DNA humano (ng) usado na reação		CT médio de amplificação (actina)	Cópias do genoma em 1ng	Cópias do genoma em Xng	Log10
5,0000		25,39	144,77	723,85	2,86
0,8333		27,82	144,77	120,64	2,08
0,1389		31,08	144,77	20,11	1,30
0,0231		33,82	144,77	3,35	0,53
0,0039		36,90	144,77	0,56	-0,25
0,0006		Undetermined	144,77	0,09	-1,03
Curva padrão DNA bacteriano					
Massa DNA mock (ng)	Massa DNA bacteriano (ng)	CT médio de amplificação (16S-V1)	Cópias do genoma em 1ng	Cópias do genoma em Xng	Log10
5,000000	4,800000	12,06	234.275,40	1.124.521,94	6,05
0,450450	0,432432	15,78	234.275,40	101.308,28	5,01
0,040581	0,038958	19,81	234.275,40	9.126,87	3,96
0,003656	0,003510	23,84	234.275,40	822,24	2,91
0,000329	0,000316	27,14	234.275,40	74,08	1,87
0,000030	0,000028	30,01	234.275,40	6,67	0,82
0,000003	0,000003	32,44	234.275,40	0,60	-0,22

Com o CT de amplificação e o número de cópias do genoma calculado para cada massa utilizada na qPCR foi possível construir a curva de amplificação observada para o DNA bacteriano e humano. Já para o cálculo da curva esperada, utilizamos o maior valor de CT obtido a partir da massa na qual iniciamos a curva observada, e os demais pontos desta foram

calculados de forma que a cada metade da massa houvesse o aumento de um CT. Como podemos observar na **Figura 9**, ambas as curvas apresentaram alta correlação entre os valores esperados e calculados, com $R^2 = 1$ para as diluições de DNA humano e $R^2=0,99$ para diluições do DNA bacteriano, mostrando uma grande proximidade com a máxima eficiência da reação de amplificação. Diante disto, optamos por utilizar a equação da reta observada para os cálculos a seguir.

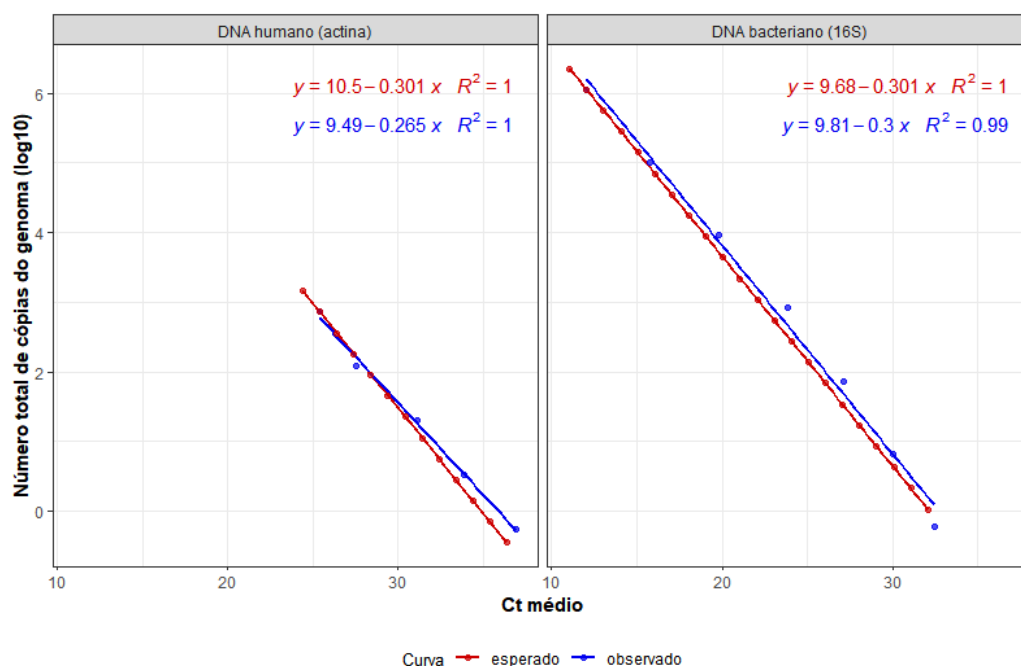


Figura 9 - Curvas padrão para DNA humano (beta-actina) e bacteriano (16S) relacionando o número de cópias de determinado genoma com seus respectivos CTs. Para o cálculo da curva padrão foram utilizados DNA humano comercial (*Human Genomic DNA* – Promega) e um mix composto de cinco espécies bacterianas (*ZymoBIOMICS™ Microbial Community DNA Standard*). Iniciando com 5ng, diversas diluições foram realizadas e amplificadas por PCR em tempo real. Foi calculado o número de cópias de determinado genoma para cada massa amplificada e, assim, curvas padrão correlacionando a média dos CTs com o número total de cópias do genoma foram construídas. Para cada genoma foram calculadas as curvas observadas (em azul) e esperadas (em vermelho), sendo que as curvas esperadas que foram calculadas a partir do menor valor de CT observado obtido, e os demais pontos foram adicionados assumindo que a cada aumento de um CT, deveria haver uma redução em 50% na massa de DNA amplificado.

4.4.2 Quantificando o número de cópias do genoma humano e bacteriano em fluidos gástricos e biópsias

Para quatro amostras de DNA de biópsia (pacientes #1, #2, #3 e #5) e seis de fluido gástrico (pacientes #1, #2, #3, #4, #5 e #9) amplificamos a região V1 do gene 16S bacteriano e o gene beta-actina humano por qPCR, de modo a termos uma relação quantitativa da microbiota. Como é possível observar na **Tabela 10**, todas as amostras de fluido gástrico possuem maior número de cópias de genomas bacterianos em relação ao número de cópias de genoma humano, com mediana de aproximadamente 10 mil genomas bacterianos *versus* 12 cópias do genoma humano por ng de amostra, sendo que os pacientes #1 e #3 possuem maior razão entre o número de cópias destes genomas, evidenciando como estas amostras possuem muito mais DNA bacteriano do que humano em sua composição (razões de cerca de 8 mil e 20 mil respectivamente). O paciente #9, por sua vez, possui um número de cópias de cada genoma mais similar do que os outros (razão de apenas 15,23). Considerando que cada genoma humano diploide tem uma massa cerca de 2000x maior que um genoma bacteriano, o sequenciamento *shotgun* de uma amostra nestas condições deve apresentar uma proporção de cerca de 260 *reads* de DNA humano para cada *read* originado de genomas DNA bacterianos. Curiosamente, o paciente #2, que alegou ter tomado antibiótico apenas uma semana antes da coleta dos materiais, ficou com razão entre os genomas (bacteriano:humano) em cerca de 400 (a terceira maior obtida). Isso sugere, ao menos para o momento do estudo, que não houve grande interferência quantitativa da antibioticoterapia na microbiota uma semana após o uso, assim, decidimos seguir com este paciente para o sequenciamento e para as análises subsequentes.

Tabela 10 - Média do CT, número de cópias e razão entre genomas bacterianos e genomas humanos nas amostras de fluido gástrico e biópsia.

Paciente	gene alvo (qPCR)	CT médio	genomas/ng de DNA total	razão cópias genoma bacteriano/humano
fluido gástrico				
1	16S	16,45	37.552,49	8.146,33
2	16S	20,05	3.109,46	412,62
3	16S	14,73	123.380,80	19.601,77
4	16S	18,90	6.924,17	45,34
5	16S	19,92	3.403,68	214,43
9	16S	17,98	13.052,68	15,23
1	Actina	32,17	4,61	
2	Actina	31,37	7,53	
3	Actina	31,66	6,29	
4	Actina	26,43	152,70	
5	Actina	30,14	15,87	
9	Actina	23,61	857,26	
Média				4.739,30
Mediana				313,56
Biópsia				
1	16S	30,10	3,01	0,10
2	16S	30,45	2,37	0,06
3	16S	26,60	33,84	1,17
5	16S	32,16	0,72	0,03
1	Actina	29,12	29,65	
2	Actina	28,70	38,25	
3	Actina	29,16	28,86	
5	Actina	29,25	27,35	
Média				0,34
Mediana				0,08

Quando avaliamos o DNA proveniente das biópsias, vemos que a quantidade de DNA humano entre as amostras é mais homogênea (mediana de 31 genomas por amostra) e a quantidade de genomas bacterianos se confirmando como muito baixa (mediana de 10 genomas), sendo que apenas o paciente #3 teve um número de cópias um pouco mais elevado (cerca de 34 genomas bacterianos por ng de DNA total), confirmando mais uma vez, que a melhor abordagem para este tipo de material é o sequenciamento de *amplicons*. Cabe notar que o paciente #3, que possui maior número de genomas bacterianos na biópsia, é o que também possui maior número de genomas bacterianos no fluido gástrico, com cerca de 123 mil genomas bacterianos por ng de DNA total. No geral, estes achados estão de acordo com o que é relatado na literatura, visto que biópsias contêm grande quantidade de DNA humano e uma baixa abundância de bactérias, quando comparado a superfícies mucosas (Zhang et al. 2015).

Diante do resultado da qPCR para fluidos gástricos, é esperado que as amostras dos pacientes #2, #4 e #9 tenham um alto percentual de *reads* humanos no sequenciamento *shotgun*. Mesmo assim, decidimos seguir com esta abordagem de sequenciamento para todos os casos buscando anotar mesmo os poucos *reads* bacterianos que eventualmente sejam sequenciados. Já para as biópsias, a técnica de sequenciamento de *amplicons* deve permitir um maior enriquecimento do DNA bacteriano presente na amostra.

4.5 RESULTADOS DO SEQUENCIAMENTO: DNA TOTAL e RNAseq

4.5.1 Processamento das *reads*

Com o intuito de identificarmos espécies bacterianas presentes na microbiota de fluidos gástricos, seu DNA e RNA total foram sequenciados e processados para remoção de *reads* de baixa qualidade, *reads* derivadas do genoma humano e rRNA residual (**Tabela 11**). A média de *reads* geradas no sequenciamento de RNA foi de 6 milhões por amostra, sendo que após todas as etapas de processamento e filtros esta média caiu para a metade, cerca de 3 milhões. De modo geral tivemos grande variação no número de *reads* de baixa qualidade, que chegou a 73% das *reads* para o paciente #4, contra apenas 15% para o paciente #3. Já com relação à remoção de DNA humano, os pacientes #6 e #9 tiveram mais de 65% das *reads* anotadas no genoma humano, sendo que nas demais amostras este percentual ficou abaixo de 6%. Ao observamos a qPCR realizada anteriormente com o cDNA de alguns pacientes, é

possível observar que aqueles onde praticamente não houve detecção de cDNA humano (pacientes #1, #2 e 5#), também demonstraram baixo número de *reads* humanos, <0,1%, o que mostra a importância da qPCR antes de seguirmos com o sequenciamento das amostras, visto que este dado nos ajuda a ter uma ideia da quantidade aproximada de genomas humanos ou bacterianos (além de fúngicos, conforme artigo no apêndice 2 - Albuquerque et al.) no material analisado, e a profundidade de leitura (milhões de *reads*) ideal. Nos dados de RNASeq, a quantidade de rRNAs de procariotos nas amostras variou de 0,1% no paciente #9 até cerca de 25% no paciente #3, sendo que as amostras com menor número de rRNA procarioto foram justamente as amostras com maior número de *reads* de origem humana.

Para o sequenciamento de DNA, obtivemos uma média de *reads* brutos de 44 milhões por amostra, valor bem superior à média obtida para as amostras de RNA. Tal diferença se deve ao uso de plataformas diferentes para o sequenciamento (Illumina para DNA, e Ion Proton para RNA). No geral, as *reads* de baixa qualidade ficaram abaixo dos 25% no DNA, e com relação ao número de *reads* contaminantes humanos, observamos dois grupos distintos: um contendo amostras que tiveram cerca de 80% de DNA humano em seu sequenciamento (pacientes #4, #6, #7 e #9), e outro grupo com porcentagem de *reads* humanos abaixo de 11% (pacientes #1, #2, #3, #5 e #8), ressaltando aqui mais uma vez que o paciente #2 relatou o uso recente de antibiótico, que no entanto parece não ter afetado de forma significativa o número de bactérias encontradas na região estomacal após uma semana de seu último uso. Comparando estes resultados com os dados de qPCR apresentados anteriormente, observamos que os pacientes #4 e #9 tinham baixa razão entre o número de cópias do genoma bacteriano e humano (razão de 45 e 15 respectivamente), corroborando o observado em seu sequenciamento, onde obtivemos um alto número de *reads* de DNA humano para estas amostras e, conseqüentemente tiveram um número de *reads* finais menor (~ 1 milhão de *reads* finais versus amostras com mais de 29 milhões de *reads* para aquelas com pouco DNA humano).

Tabela 11 – Processamento das *reads*: amostras sequenciadas de fluidos gástricos (RNAseq e DNA total).

Paciente	<i>Reads</i> brutas	<i>Reads</i> de baixa qualidade (%)	<i>Reads</i> humanas (%)	<i>Reads</i> rRNA humano (%)	<i>Reads</i> rRNA procarioto (%)	<i>Reads</i> finais	(%)
RNAseq							
1	7.570.545	49,48	0,08	0,24	18,98	3.818.758	50,44
2	8.294.206	34,56	0,01	0,14	21,81	5.427.332	65,44
3	4.845.468	15,97	2,60	1,66	24,83	3.945.371	81,42
4	6.837.008	73,61	2,53	0,00	6,42	1.631.614	23,86
5	8.993.790	32,21	0,02	0,57	20,06	6.095.038	67,77
6	3.683.773	20,52	67,92	0,05	0,36	425.794	11,56
7	6.177.041	49,74	5,07	0,13	10,48	2.791.266	45,19
8	6.465972	27,12	0,16	0,13	15,31	4.701.968	72,72
9	1.289.182	23,41	69,10	0,02	0,10	96.540	7,49
Média	6.017.443					3.214.853	
mediana	6.465.972					3.818.758	
Paciente	<i>Reads</i> brutas	<i>Reads</i> de baixa qualidade (%)	<i>Reads</i> humanas (%)	<i>Reads</i> finais		(%)	
DNA total							
1	57.422.506	25,22	10,95	36.649.732		63,82	
2	22.291.034	24,30	12,65	14.054.244		63,05	
3	44.735.162	21,67	9,36	30.851.982		68,97	
4	46.666.914	15,10	81,40	1.633.110		3,50	
5	44.714.260	23,21	10,48	29.649.164		66,31	
6	70.211.950	14,67	81,46	2.716.850		3,87	
7	23.632.558	15,64	79,80	1.078.226		4,56	
8	63.662.588	12,45	6,06	51.880.334		81,49	
9	27.824.426	17,03	78,18	1.333.646		4,79	
Média	44.573.489			18.871.920			
mediana	44,735,162			14.054.244			

Apenas duas amostras tiveram um alto número de *reads* humanos tanto a nível de DNA quanto de RNA (pacientes #6 e #9). A amostra #4 teve alto número de *reads* humanos somente a nível de DNA, mas a nível de RNA teve mais de 70% das *reads* consideradas de baixa qualidade, restando apenas cerca de 2,5% de DNA humano na amostra. De forma similar, o paciente #7 teve quase 50% das *reads* de baixa qualidade, sobrando apenas 5% de *reads* humanos para serem filtrados.

No geral, o fator que mais influenciou o número de *reads* bacterianos finais no RNAseq foi a qualidade da *read*, enquanto que no DNA total foi a presença de DNA humano na amostra. Uma vez que a degradação dos ácidos nucleicos ocorre no estômago, principalmente devido a enzima pepsina (Liu et al, 2015), é esperado que os fluidos gástricos possuam ácidos nucleicos parcialmente degradados, principalmente o RNA, que é mais lábil e possui usualmente uma meia-vida de algumas poucas horas (Tani e Akimitsu 2012), enquanto que o DNA é bem mais estável (Allentoft 2012). Além disso, como já citado, as amostras de RNA deram baixa quantificação no geral e estavam bem degradadas.

4.5.2 Composição da microbiota bacteriana de fluidos gástricos avaliada por molécula sequenciada (DNA x RNA)

No total, detectamos 7.088 espécies bacterianas no DNA e 1.492 espécies bacterianas no RNA, sendo que destas, 1.487 espécies estavam presentes em ambos os materiais. Assim, a maioria das espécies foi encontrada exclusivamente em *reads* de DNA (5.601 espécies), contra apenas 5 espécies exclusivas do RNA. Tal achado não é totalmente surpreendente diante de dois fatos: i) as moléculas de DNA são mais estáveis, e ii) a quantidade de *reads* geradas para as bibliotecas de DNA foi muito maior que a quantidade média de *reads* obtidas para as bibliotecas de RNA (respectivamente, médias brutas de 44 milhões x 6 milhões - um fator de 7,33X). Quando normalizamos o número de espécies encontradas por milhão de *read* sequenciados, para RNA encontramos cerca de 390 espécies, enquanto que para DNA foram 504 espécies por milhão de *reads*, ou seja, a diferença permanece (não pode ser atribuída apenas à profundidade de sequenciamento) e detectamos cerca de 114 espécies a mais ao nível DNA por milhão de *reads*, evidenciando que de fato temos mais espécies encontradas no sequenciamento de DNA total.

Ao avaliarmos a abundância das amostras exclusivas de cada material, ainda, observamos que as espécies exclusivamente encontradas no DNA representam apenas 1,5%

da abundância total de bactérias nos fluidos gástricos neste material, enquanto as exclusivas de RNA representam uma porcentagem ainda menor, apenas 0,001% do material sequenciado. Esta diferença entre o número de espécies encontradas entre DNA e RNA, mantida a proporção encontrada em um universo amostral distinto, reflete uma das ideias centrais deste projeto: ao estudarmos as bactérias presentes no fluido gástrico ao nível de DNA estamos possivelmente identificando todas que passam pela região estomacal, sendo elas residentes da cavidade gástrica ou transientes e oriundas da cavidade oral ou outras porções do trato gastrointestinal. Por outro lado, quando identificamos as bactérias presentes ao nível de RNA, identificamos aquelas com atividade transcricional detectável e que podem ter sua expressão gênica influenciando e sendo influenciada pelo microambiente estomacal. Se por um lado as bactérias mortas devem ter uma queda mais acentuada em seus *reads* quando DNA é comparado a RNA, as bactérias vivas, devem manter boa representatividade nas bibliotecas de RNA, pois a lise celular deve se dar apenas durante o processo de extração dos ácidos nucléicos.

Considerando ser menos provável que bactérias em baixa frequência sejam relevantes no contexto biológico, quando avaliamos a microbiota bacteriana com frequência maior que 1% nos fluidos gástricos, observamos várias espécies do gênero *Streptococcus*, sendo que a *Streptococcus pneumoniae* se destaca como a mais abundante tanto nos dados de DNA (11 %) quanto nos dados de RNA (28%), atividade transcricional que sugere esta ser uma bactéria residente na cavidade estomacal (**Figura 10**). Além disso, é possível observar que as demais bactérias com maior frequência no RNA (*Pseudoramibacter alactolyticus*, *Virgibacillus senegalensis* e *Bacillus cereus*, com cerca de 8% cada) não estão entre as mais abundantes no DNA, representando apenas 0,04%, 0,02% e 0,05% de sua composição respectivamente. A vasta diferença percentual entre os *reads* obtidos a partir de RNA, se comparados aos oriundos da biblioteca de DNA, pode sugerir que estas bactérias se encontram vivas, e possam ser residentes da cavidade gástrica.

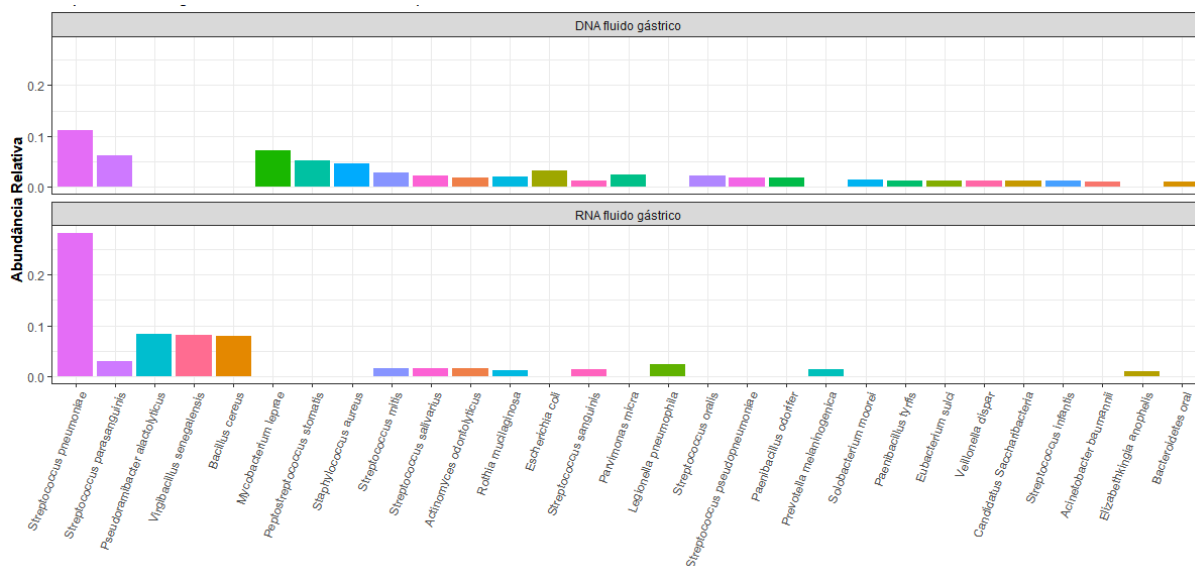


Figura 10 - Abundância relativa das espécies bacterianas com frequência superior a 1% nas amostras de fluido gástrico, estratificado por material sequenciado (DNA x RNA).

O gênero *Streptococcus* já foi identificado como abundante no fluido gástrico de indivíduos saudáveis (Bassis et al. 2015) e em biópsias de pacientes com CG pela amplificação do gene 16S (Rajilic-Stojanovic et al. 2020), quando avaliada a abundância bacteriana a nível de RNA, em um número menor de estudos, este gênero também já foi descrito como abundante na cavidade gástrica (Schulz et al. 2016). A espécie *S. pneumoniae*, por sua vez, é uma bactéria gram-positiva, patógeno oportunista que geralmente coloniza a mucosa do trato respiratório superior (Weiser et al. 2018). A sua disseminação está relacionada a diversas doenças inflamatórias, tais como otite, pneumonia, sepse e meningite (Kadioglu et al. 2008; Weiser et al. 2018), além disso, esta bactéria também já foi relatada como causadora de infecções pélvicas e intra-abdominais (Ronnachit et al. 1027). Em um dos poucos estudos que utilizaram sequenciamento de DNA total em amostras de fluido gástrico, Hu et al. em 2018 descrevem as espécies “*Streptococcus_mitis_oralis_pneumoniae*” entre os táxons mais abundantes no fluido gástrico de pacientes com CG, porém não individualizando estas espécies.

Curiosamente, a bactéria *Pseudoramibacter alactolyticus* foi detectada em alta frequência apenas ao nível de RNA no paciente #3, com cerca de 57% das *reads* associadas a esta bactéria, enquanto que nos demais essa proporção não passou de 0,06%, A *P. alactolyticus* já foi descrita como associada a infecções endodônticas (Siqueira e Rôças 2003;

Siqueira e Rôças 2009), e já foi encontrada no sistema nervoso central de um paciente após extração dentária (Liao et al. 2019). Além disto, estudos recentes têm mostrado a associação de doença periodontal e AdG, sendo possível que esta associação seja em parte determinada pela translocação da microbiota entre estes ambientes (Chou et al. 2018; Lo et al. 2020), algo que talvez tenha ocorrido no paciente #3.

Dentre as bactérias mais abundantes ao nível de DNA, a *Staphylococcus aureus* foi detectada com frequência de 4% neste material, porém em baixíssima quantidade no RNA dos fluidos gástricos (<0,004%). Quando quantificada por paciente, esta bactéria aparece com frequência acima de 4% em quatro pacientes ao nível de DNA (#4, #6, #7 e #9), chegando a até 10% da composição da amostra. *S. aureus* faz parte da microbiota nasal do hospedeiro, mas também pode colonizar o intestino e a região perineal (Acton et al. 2008; Tong et al. 2015; Balasubramanian et al. 2017). Esta bactéria também já foi descrita como principal agente etiológico causador de infecções, influenciando o prognóstico de pacientes com diversos tipos tumorais, incluindo câncer pancreático, de estômago e colorretal (Takeuchi et al. 2001; Bai et al. 2018), além de estar associada com infecções nosocomiais e intoxicação alimentar. Já a bactéria *Peptostreptococcus stomatis*, também em destaque na microbiota pela análise do DNA (5% contra apenas 0,07% no RNA), foi inicialmente encontrada na cavidade oral (Downes e Wade 2006) e alguns estudos já relacionaram a sua presença com câncer colorretal (Zeller et al. 2014; Yu et al. 2017b). Diante disso, é possível explicar a presença destas bactérias no estômago dos pacientes estudados, visto que as mesmas podem estar passando pelo estômago, mas são provenientes de outras regiões a ele conectadas.

Curiosamente, observamos o encontro de sequências de DNA aparentemente correspondentes a fragmentos genômicos de *Mycobacterium leprae*, agente causador da hanseníase (Scollard et al. 2006). Proporções mais elevadas destes fragmentos foram observadas em quatro pacientes (#4, #6, #7 e #9). *Reads* mapeadas no genoma de *M. leprae* foram encontradas nas bibliotecas de DNA - representando 7% da abundância total de espécies, e também nas bibliotecas de RNA, porém em frequência significativamente menor (0,008%). A frequência destas *reads* (DNA) variou entre 2,5% e 27% nestes quatro pacientes que, interessantemente, também foram os que apresentaram maior número de *reads* humanos nas mesmas bibliotecas. Nos demais pacientes sua frequência não passou de 0,01%.

O Brasil é, atualmente, o segundo país com maior incidência de hanseníase (WHO 2018), doença causada pela *M. leprae*, que produz uma infecção crônica que afeta

principalmente nervos periféricos e pele, mas também pode afetar outras regiões como olhos e membranas mucosas (White e Franco-Paredes 2015). Já foi relatado que por coleta de *swabs* nasais é possível a identificação desta bactéria (Job et al. 2008), sugerindo, portanto, que elas podem ser encontradas nos fluidos gástricos. Alguns estudos antigos já tentaram avaliar um possível aumento na incidência de câncer para pacientes com esta doença, porém sem sucesso (Purtilo e Pangi 1975; Kolonel e Hirohata 1977), e esta é a primeira vez que a *M. leprae* foi identificada no fluido gástrico de pacientes com AdG (que tenhamos conhecimento), e diante disso estudos adicionais são necessários para confirmar estes achados nestes indivíduos, e alguma possível associação entre estas doenças.

Com o intuito de validarmos o resultado obtido e descartarmos um possível artefato do sequenciamento *shotgun*, as amostras de DNA sequenciadas tiveram suas *reads* anotadas contra o genoma da bactéria *Mycobacterium leprae*. Para o mapeamento, foram consideradas as *reads* anotadas no genoma com alta similaridade (identidade $\geq 97\%$) e *E-value* $\leq 1e-50$. O valor de *E-value* representa quantos resultados podem ser encontrados apenas por acaso, dado o mesmo tamanho de um banco de dados aleatório, ou seja, quanto menor o *E-value*, melhor é a qualidade do alinhamento. Considerando o tamanho médio do genoma da *M. leprae* 3,3Mb (Cole et al. 2001), e que a mediana dos fragmentos gerados variou de 119 a 150pb entre os pacientes, podemos observar na **Tabela 12**, que os pacientes #4, #6, #7 e #9 tiveram suas *reads* cobrindo aproximadamente apenas 3% do genoma desta bactéria. Se por um lado esta cobertura (que variou entre 17Kb e 121Kb do genoma bacteriano) indica uma dispersão relativamente boa dos *reads* no genoma, não podemos descartar um possível artefato de mapeamento resultante de similaridade com elementos repetitivos do genoma de *M. leprae*.

Tabela 12 - Anotação das *reads* sequenciadas de fluido gástrico (DNA total) no genoma de *Mycobacterium leprae*.

Paciente	número de regiões anotadas no genoma ¹	reads mapeadas no genoma	Cobertura horizontal (nt) do genoma ²	% de cobertura do genoma ³	abundância relativa (<i>shotgun</i>)
1	347	4.187	42.681	1.29%	0.0057%
2	325	911	40.625	1.23%	0.0146%
3	148	368	17.612	0.53%	0.0003%
5	431	2.681	53.444	1.62%	0.0200%
4	646	16.419	96.900	2.94%	2.4739%
6	809	113.853	121.350	3.68%	27.3647%
7	700	32.535	105.000	3.18%	9.5461%
8	655	20.645	81.220	2.46%	0.0063%
9	731	32.598	109.650	3.32%	14.1938%

¹ Foram consideradas as *reads* anotadas no genoma de *M. leprae* com identidade $\geq 97\%$ e *E-value* $\leq 1e-50$. ² A cobertura horizontal foi calculada a partir da mediana dos fragmentos anotados no genoma de *M. leprae*, que variou de 119 até 150pb dependendo do paciente. ³ A porcentagem da cobertura do genoma foi calculada considerando o tamanho do genoma de *M. leprae* de 3,3Mb.

Ao compararmos a composição bacteriana das amostras (DNA x RNA) com abundância maior que 1% em pelo menos um material, foi possível observar que as bactérias *Bacillus Cereus*, *Elizabethkingia anophelis*, *Legionella pneumophila* e *Virgibacillus senegalensis* estão mais abundantes no RNA, enquanto apenas as espécies *Solobacterium moorei* e *Streptococcus pseudopneumoniae* foram mais abundantes no DNA dos fluidos gástricos de forma significativa (**Figura 11**).

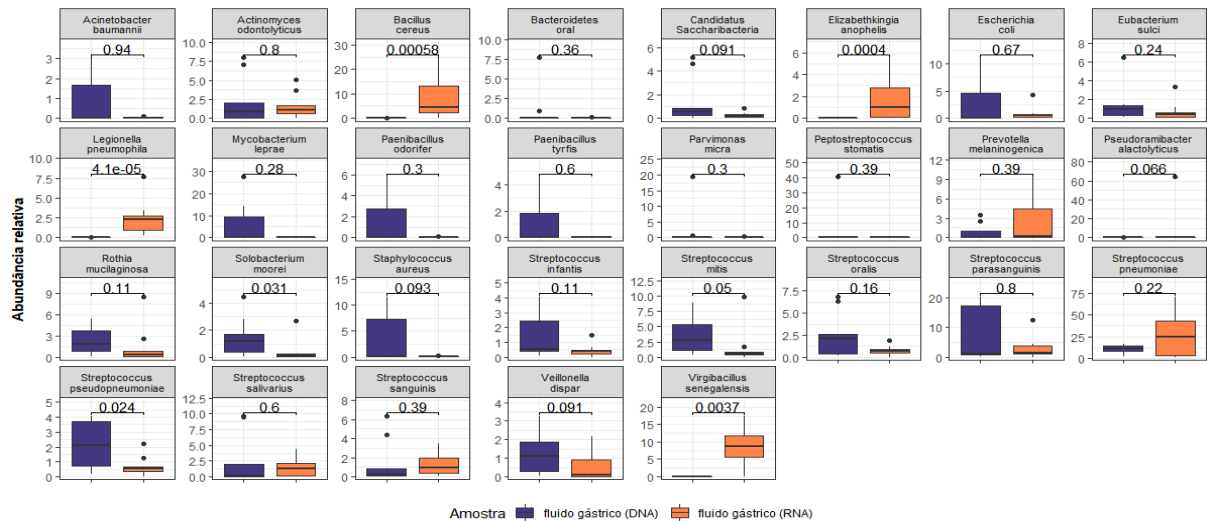


Figura 11 - Análise comparativa da abundância das espécies bacterianas com frequência superior a 1% nas amostras de fluido gástrico (DNA x RNA). Foram selecionadas as espécies que possuíam abundância relativa maior que 1% em pelo menos um material sequenciado de fluido gástrico (DNA ou RNA). Para análise estatística foi utilizado o teste não paramétrico *Wilcoxon*, para amostras pareadas. As espécies foram consideradas como diferencialmente abundantes quando $p \leq 0,05$.

Com relação às bactérias mais abundantes na microbiota ao nível de RNA, a espécie *Virgibacillus senegalensis*, é uma bactéria halofílica, descrita apenas em 2015, encontrada em fezes de um indivíduo senegalense saudável, o que gerou o nome da espécie (Seck et al. 2015). Esta bactéria não possui outros relatos na literatura até o momento, e a sua presença e atividade transcricional no estômago é de certa forma intrigante. Pela sua frequência relativamente alta, parece pouco provável que seja um artefato do sequenciamento, e quando observamos por paciente, 8 deles tiveram esta bactéria identificada, sendo que sua abundância ficou entre 6-14% em seis pacientes. Já ao nível de DNA esta bactéria se mostrou em frequência extremamente baixa ($< 0,0004\%$).

A bactéria *Bacillus cereus*, por sua vez, é uma bactéria gram-positiva, anaeróbia facultativa, conhecida principalmente por habitar o solo, podendo também ser encontrada na água e em vegetais. Enquanto que algumas cepas de *B. cereus* são utilizadas como probióticos (Zhu et al. 2016), outras estão relacionadas a infecções alimentares (Arnesen et al. 2008). Esta bactéria possui mecanismos de resistência ao pH ácido do estômago, visto que ela precisa passar por esta região para se alojar no intestino (Duport et al. 2016). Para 8/9 pacientes esta espécie foi identificada do nível de RNA, sendo que sua frequência variou de 19% no paciente #8 até 0,03% no paciente #6. Já no DNA, esta espécie foi identificada em frequência total de

apenas em 0,0005%. Curiosamente, as bactérias *V. senegalensis* e *B. cereus* só não foram detectadas no paciente #9.

Já a espécie *Elizabethkingia anophelis*, bactéria aeróbica gram-negativa, foi detectada com frequência entre 1-5% em três dos nove pacientes, e foi isolada pela primeira vez em 2012 no intestino de mosquitos da espécie *Anopheles gambiae*, o principal vetor da malária (Kämpfer et al. 2012; Hu et al. 2017). Diversas infecções já foram relacionadas a esta bactéria, como meningite e sepse (Lau et al. 2016; Janda e Lopez 2017). Até o momento, porém, esta bactéria não foi descrita como presente no estômago, ainda mais de forma ativa, sendo este achado passível de análises mais minuciosas.

Por fim, a bactéria *Legionella pneumophila* apareceu na microbiota (RNA) com frequência entre 1-8% em seis pacientes, enquanto que no DNA esta frequência ficou abaixo de 0,04%. A *L. pneumophila* é uma bactéria gram-negativa que pode causar pneumonia. A infecção por esta bactéria está relacionada, principalmente, com a ingestão de água contaminada, e já foi relatada até mesmo obstrução intestinal em pacientes devido a infecção por esta espécie (Prichard e Fick 2016). Mais uma vez nos deparamos com uma bactéria abundante na microbiota ao nível de RNA e não descrita, até o momento, como relevante na cavidade estomacal.

Quando avaliamos as bactérias encontradas majoritariamente na microbiota ao nível de DNA, a *Solobacterium moorei*, gram-positiva e anaeróbica obrigatória, está relacionada com halitose (Haraszthy et al. 2007; Barrak et al. 2020), sendo assim parece plausível que seu material genético esteja presente na microbiota do fluido gástrico, visto que bactérias da cavidade oral estão em contínua deglutição, caindo na cavidade gástrica. Já a *Streptococcus pseudopneumoniae*, possui características similares com a *S. pneumoniae* (Arbique et al., 2004) e já foi relatada como abundante em biópsias e mucosa estomacal de pacientes *H. pylori* negativos (Sohn et al. 2017). No geral, esta bactéria ficou com abundância em torno de 2% na microbiota ao nível de DNA versus 0,7% na microbiota RNA.

Ao avaliarmos se existe alguma associação entre as bactérias encontradas ao nível de DNA versus as encontradas ao nível de RNA, foi observado que existe uma pequena associação positiva entre a abundância das espécies em ambos os materiais (**Figura 12**, $r=0,51$ e $p<0,05$), porém é evidente que várias espécies em abundâncias variadas ao nível de DNA não aumentam em proporção ao nível de RNA (formando praticamente uma linha na borda do eixo x no gráfico). O que sugere que talvez estas bactérias estejam mortas ou tenham reduzida

atividade transcricional no estômago, ressaltando o valor do uso combinado do sequenciamento de DNA e de RNA para a avaliação da microbiota bacteriana estomacal.

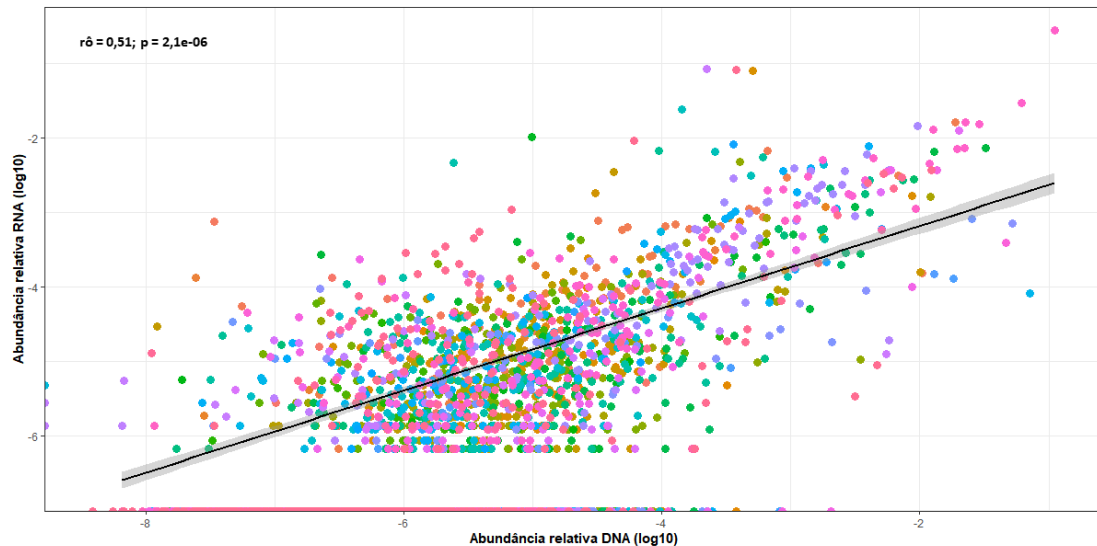


Figura 12 - Correlação da abundância relativa (log10) das espécies bacterianas nas amostras de fluidos gástricos (DNA x RNA). Cada círculo representa uma espécie bacteriana. A linha diagonal indica abundâncias similares nos conjuntos de dados de DNA e RNA, sendo que as espécies mais próximas da linha possuem maior correlação entre os materiais. Para análise estatística foi utilizada a correlação de *Spearman*, sendo que quanto $rô$ mais perto de 1, maior é a correlação entre as variáveis. A área sombreada representa o intervalo de confiança. A correlação foi considerada como estatisticamente significativa quando $p \leq 0,05$.

Para comparar a diversidade bacteriana das amostras, utilizamos as métricas de diversidade alfa e beta. Ao avaliarmos a diversidade alfa, verificamos se o número de bactérias diferentes (riqueza) entre o DNA e RNA é similar (Lande 1996). Como mostra a **Figura 13A**, é possível observar que a riqueza de espécies é maior na microbiota ao nível de DNA quando comparada com a encontrada no RNA (gráfico *Observed*). Quando levamos em consideração a abundância destas espécies na amostra, também houve diferença significativa pelo índice de *Simpson* ($p=0,04$), que dá um peso maior para as espécies mais frequentes.

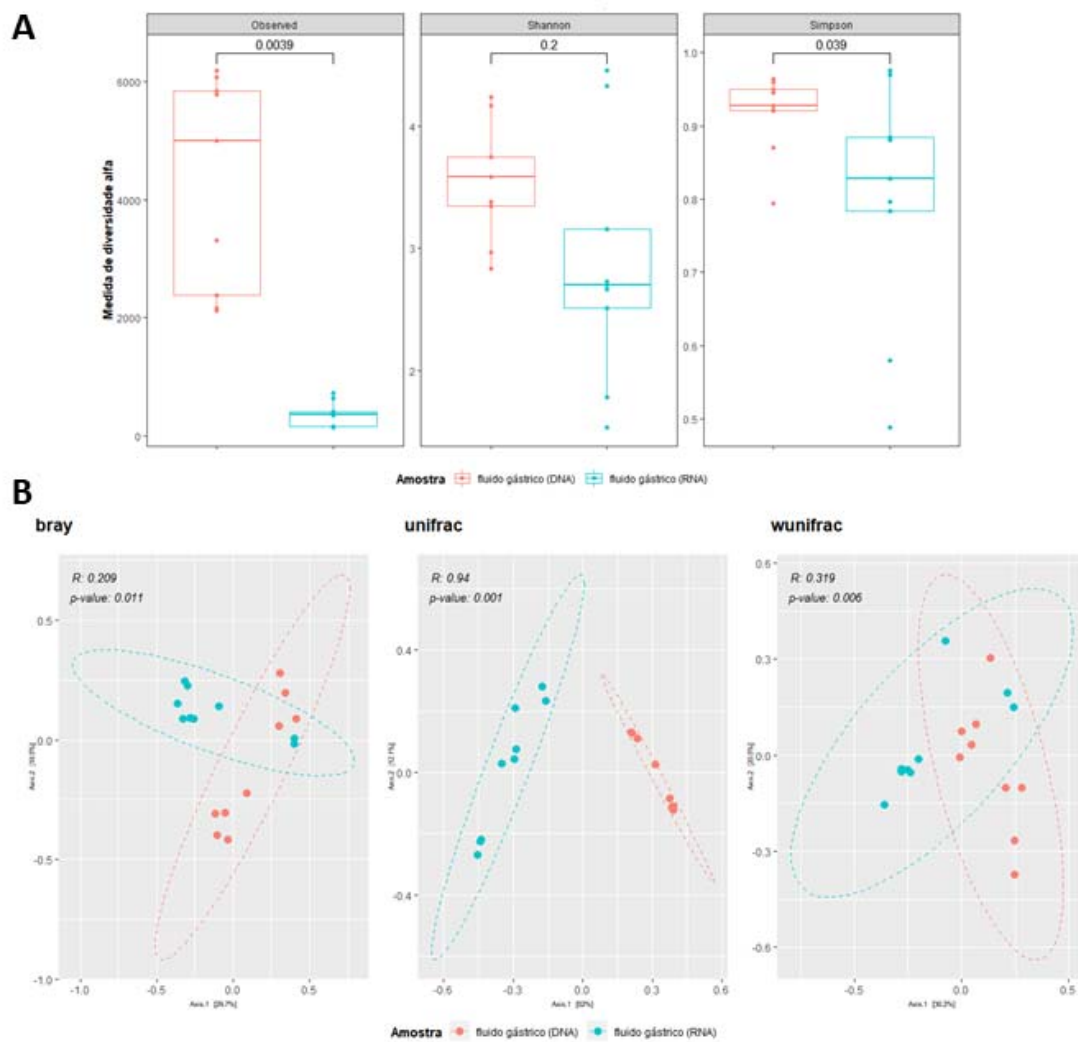


Figura 13 - Medidas de diversidade nas amostras de fluido gástrico (DNA x RNA). (A) A diversidade alfa foi avaliada pelos índices *Observed*, *Shannon* e *Simpson* respectivamente, e o teste não paramétrico U de Mann-Whitney foi utilizado para verificar a existência de diferença significativa entre as amostras. (B) A diversidade beta foi avaliada pelos índices de Bray-Curtis, e Unifrac (*unweighted* e *weighted*) respectivamente, e o teste não paramétrico ANOSIM foi utilizado para avaliar a similaridade entre as amostras, sendo o R o coeficiente de correlação. A diferença foi considerada estatisticamente significativa quando $p \leq 0,05$.

Outro parâmetro comumente utilizado em análises de diversidade é a diversidade beta, que avalia o quão similar é a composição bacteriana de diferentes amostras. Dentre os parâmetros utilizados, o índice *Bray-Curtis* avalia o quão similar são as amostras baseado na sua abundância, enquanto que índice *UniFrac* avalia esta similaridade a partir de quão próximas filogeneticamente as bactérias estão entre si e quanto mais distantes, menor é a similaridade das amostras (Lozupone et al. 2007). Diante disto, observamos que a composição bacteriana encontrada no RNA é diferente da vista no DNA, e isto é verdade tanto quando

consideramos apenas a abundância relativa das bactérias (índice *Bray-Curtis*), como quando levamos em consideração a distância das bactérias na árvore filogenética (índice *Unifrac unweighted* e *weighted*), com $p < 0,05$ em todas as métricas (**Figura 13B**), indicando, mais uma vez, que as diferentes abordagens de sequenciamento para este material geram resultados diferenciados, ou seja, a microbiota bacteriana ao nível de DNA parece representar muito mais que somente as bactérias presentes na região estomacal, precisando ser analisada em conjunto com o RNA deste material para inferir quais bactérias estão de fato transcricionalmente ativas na região estomacal.

Por fim, apesar de os pacientes selecionados para este estudo terem sido classificados como *H. pylori* negativos, podemos observar na **Figura 14** que na microbiota de todos os pacientes houve anotação para esta bactéria ao nível de DNA. Sua abundância, porém, é muito baixa, com no máximo 0,14% das *reads* associadas a esta bactéria para o paciente #6. Já a partir do RNA, apenas dois pacientes apresentaram *reads* mapeados no genoma de *H. pylori* (pacientes #4 e #7), e estes não chegaram a 0,006% da amostra, reforçando a baixa frequência e/ou baixa atividade transcricional desta bactéria nestes pacientes. Tais achados demonstram a capacidade de sequenciamento em larga-escala de revelar microrganismos de interesse médico, adicionando informação relevante e complementar a técnicas padrão. Neste sentido, cabe notar artigo recente que ao avaliar por *Deep Sequencing* 392 pacientes com câncer de colo de útero, negativas para HPV por PCR pacientes, encontrou DNA viral de 30 tipos de HPV em 169 casos (Arroyo Mühr et al 2020).

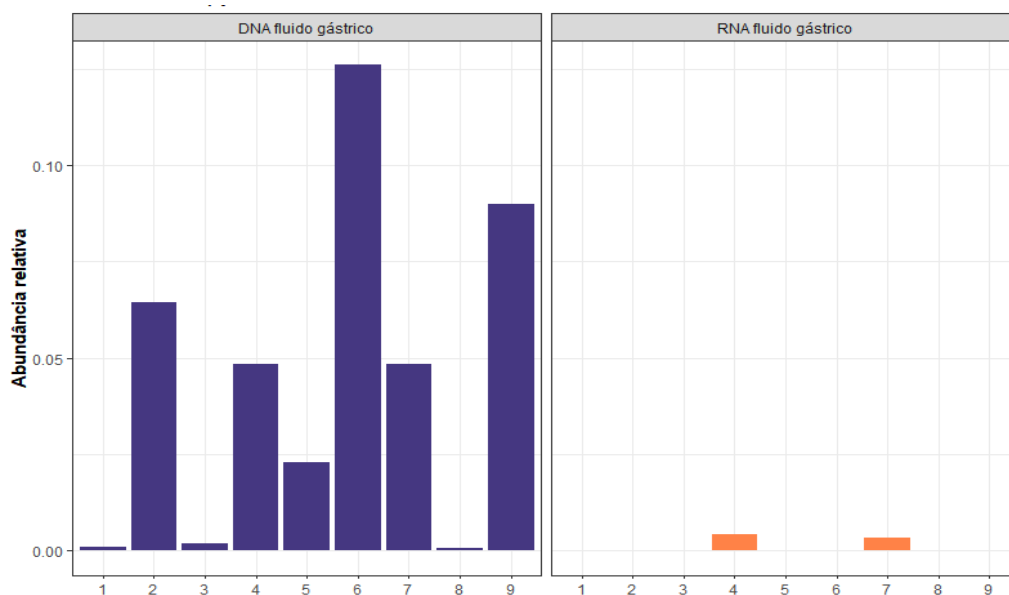


Figura 14 - Abundância relativa de *Helicobacter pylori* na microbiota bacteriana de fluidos gástricos (DNA x RNA).

De modo geral, as espécies encontradas na microbiota ao nível de DNA parecem ser diferentes das encontradas na microbiota ativa transcricionalmente, visto que encontramos espécies diferencialmente abundantes, tais como *B. cereus*, *E. anophelis*, *L. pneumophila* e *V. senegalensis* mais presentes em *reads* de RNA, versus *Solobacterium moorei* e *Streptococcus pseudopneumoniae* mais abundantes a nível de DNA. As amostras no geral tiveram uma baixa correlação entre as variáveis, e através dos índices de diversidade elas foram consideradas populações distintas. Com exceção de *S. pneumoniae*, já foi relatada como aumentada em pacientes com CG em um estudo usando a abordagem de *shotgun* conforme já citado (Hu et al. 2018), as bactérias identificadas no nosso estudo como diferencialmente abundantes na microbiota transcricionalmente ativa, não foram até o momento corroboradas outros trabalhos, sendo que o presente estudo é o primeiro no qual temos conhecimento a avaliar a microbiota de fluidos gástricos pela abordagem de RNAseq.

4.6 CULTIVO MICROBIANO E ESPÉCIES IDENTIFICADAS NAS AMOSTRAS DE FLUIDO GÁSTRICO

Utilizamos para experimento de cultivo microbiológico, amostras de fluido gástrico coletadas de três pacientes com AdG. Para um destes pacientes (#8) não foi observado crescimento microbiano em nenhum meio sob nenhuma das condições de cultivo utilizadas. Do fluido gástrico cultivado do paciente #7 somente nos meios em condições com oxigênio foi observado crescimento, enquanto que para o fluido gástrico proveniente do paciente #9 houve crescimento em todos os meios e condições utilizadas. A **Tabela 13** exemplifica os meios e condições utilizados e em quais houve crescimento bacteriano.

Tabela 13 - Meios de cultura, condições de oxigênio e número de colônias observadas através do fluido gástrico de três pacientes do estudo.

		paciente 7		paciente 8		paciente 9	
Meio	Oxigênio	crescimento	colônias (n)	crescimento	colônias	Crescimento	colônias
BHI	O ₂ +	✓	1	X	0	✓	1
	O ₂ -	X	0	X	0	✓	1
Chocolate	O ₂ +	✓	2	X	0	✓	1
	O ₂ -	X	0	X	0	✓	3
Sangue	O ₂ +	✓	1	X	0	✓	2
	O ₂ -	X	0	X	0	✓	1

Dentre as informações clínicas coletadas, podemos destacar que os pacientes #7 e #9 possuíam lesão localizada no corpo gástrico. Ambos foram considerados respondedores ao tratamento neoadjuvante, tinham pH estomacal ácido (pH 3 e 4 respectivamente). Já para o paciente #8, a lesão foi na transição esofagogástrica, sem resposta à QTneo, com pH estomacal neutro no momento da coleta (pH=7). Ainda, os pacientes #8 e #9 relataram estar fazendo uso de inibidores de bomba de próton.

Foram selecionadas todas as quatro colônias obtidas do paciente #7 enquanto que do paciente #9, visto a semelhança entre algumas colônias entre os diferentes meios, foram

selecionadas as duas colônias obtidas através do meio de cultura sangue na presença de oxigênio, enquanto que das colônias que cresceram em anaerobiose, selecionamos duas colônias que cresceram no meio chocolate e a colônia obtida através do meio BHI. A **Figura 15** mostra o crescimento microbiano obtido nas amostras após 48h.

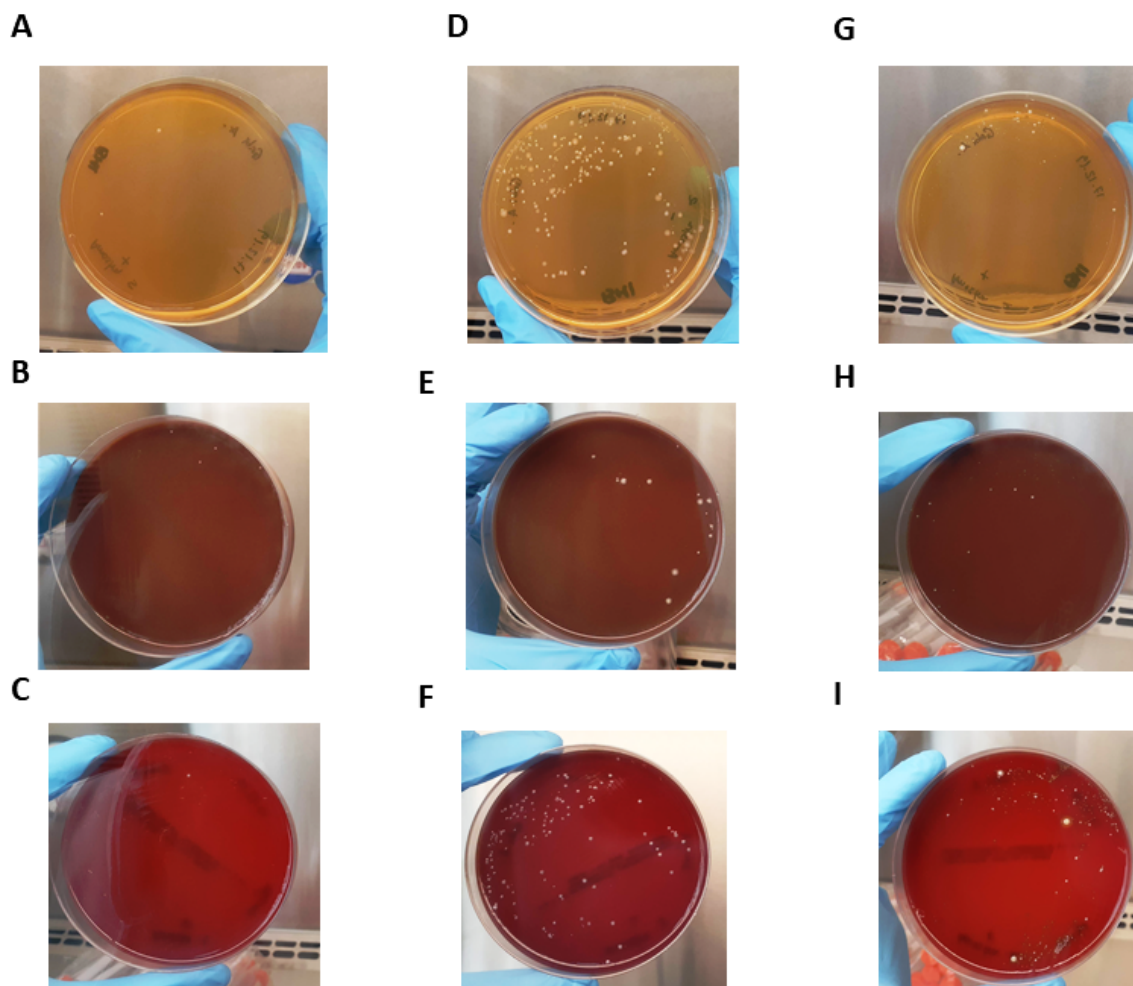


Figura 15 - Crescimento microbiano de fluidos gástricos avaliado após 48h. Na primeira coluna temos as colônias provenientes do paciente #7 nos meios (A) BHI/O₂ +, (B) chocolate/O₂+ e (C) sangue/O₂+. Nas colunas dois e três encontram-se as colônias obtidas do paciente #9 nos meios (D) BHI/O₂-, (E) chocolate/O₂-, (F) sangue/O₂-, (G) BHI/O₂+, (H) chocolate/O₂+ e (I) sangue/O₂+. A indicação O₂- ou O₂+ equivalem respectivamente à ausência ou presença de oxigênio no cultivo.

As colônias isoladas foram transferidas para meio líquido, onde todas cresceram de modo adequado, seguido de extração de DNA total e posterior sequenciamento, a **Tabela 14** resume as bactérias identificadas a partir das colônias isoladas sequenciadas. Quando observamos os microrganismos que cresceram a partir do fluido gástrico, no paciente #7

observamos apenas espécies de fungo. Os dados de sequenciamento das duas colônias distintas obtidas a partir do meio chocolate na presença de O₂ indicaram se tratar das espécies *Candida parapsilosis* e *Trichosporon asahii*, sendo que para a *C. parapsilosis* 95% das *reads* foram anotadas para este fungo. Em contraste, apenas 10% das *reads* foram associadas à espécie *T. asahii*, mas considerando que apenas 14% das *reads* foram anotadas no domínio fúngico, o resultado obtido parece viável. As colônias obtidas através dos meios sangue e BHI, ambos também na presença de O₂, acabaram indicando se tratar da mesma espécie que cresceu em uma das colônias do meio chocolate, a *C. parapsilosis*. Cabe, no entanto, notar que ambos fungos *C. parapsilosis* e *T. asahii* são conhecidos por habitarem a pele humana (Jautová et al. 2001; Rafat et al. 2017), fato que não nos faz descartar a possibilidade de que sejam contaminantes do meio de cultura, apesar de não termos tido problema de contaminação ou crescimento de colônias nos controles negativos utilizados.

Tabela 14 - Identificação dos microrganismos sequenciados, *reads* totais geradas e *reads* associadas ao microrganismo potencialmente identificado.

Colônia	Meio	Oxigênio	Microrganismo identificado	<i>reads</i> totais	<i>reads</i> do microrganismo	(%)
paciente 7						
1	chocolate	+	<i>Candida parapsilosis</i>	515.015	491.602	95,45
2	chocolate	+	<i>Trichosporon asahii</i>	301.227	30.311	10,06
3	sangue	+	<i>Candida parapsilosis</i>	133.775	125.719	93,98
4	BHI	+	<i>Candida parapsilosis</i>	191.583	180.415	94,17
paciente 9						
1	sangue	+	<i>Streptococcus parasanguinis</i>	694.839	234.285	33,72
			<i>Rothia dentocariosa</i>		227.820	32,79
2	sangue	+	<i>Staphylococcus aureus</i>	616.145	548.102	88,96
3	chocolate	-	<i>Bacillus subtilis</i>	951.525	223.037	23,44
			<i>Streptococcus salivarius</i>		206.645	21,72
4	chocolate	-	<i>Streptococcus</i>	557.137	490.719	88,08
			<i>Streptococcus vestibularis</i>		115.200	20,68
			<i>Streptococcus gordonii</i>		96.318	17,29
5	BHI	-	<i>Streptococcus</i>	598.662	536.452	89,61
			<i>Streptococcus sp, LPB0220</i>		210.177	35,11
			<i>Streptococcus parasanguinis</i>		192.110	32,09

Com relação aos microrganismos identificados no paciente #9, todos são bactérias, sendo que a colônia 1 obtida no meio sangue O₂+ na verdade eram duas colônias distintas, sendo classificadas como as espécies *S. parasanguinis* e *R. dentocariosa*. Já a segunda colônia obtida deste meio foi da espécie *S. aureus*. A colônia três, que cresceu em meio chocolate sem a presença de O₂, também resultou em duas espécies distintas, a *B. subtilis* e a *S. salivarius*, mostrando que mais uma vez foram selecionadas duas colônias em conjunto. A colônia 4, por sua vez, obtida do mesmo meio da anterior, identificou duas espécies do gênero *Streptococcus*, *S. vestibularis* e *S. gordonii*. Por fim, na colônia cinco, extraída do meio BHI O₂,

também foram identificadas duas espécies do gênero *Streptococcus*, porém distintas das obtidas do meio chocolate, foram estas a *S. sp. LPB0220* e *S. parasanguinis*.

Em geral, evidências na literatura corroboram as bactérias encontradas vivas no nosso trabalho. Em um estudo no qual foram cultivadas em meio ágar sangue biópsias e fluido gástrico de indivíduos saudáveis e com gastrite crônica, foram identificadas colônias das espécies bacterianas *S. aureus*, *S. salivarius* e *S. parasanguinis*, similares às encontradas no fluido gástrico do paciente #9, além de outras espécies bacterianas do gênero *Streptococcus* e *Rothia*, e uma espécie de fungo (*Candida albicans*), mesmo gênero encontrado no nosso estudo (Liu et al. 2018). Em outro estudo no qual foram cultivadas biópsias de indivíduos com e sem infecção por *H. pylori*, a espécie *Rothia dentocariosa* foi uma das espécies mais frequentes, que interessantemente só foi detectada na ausência de *H. pylori*, além disto, diversas espécies do gênero *Streptococcus* também foram identificadas neste estudo (Troncoso et al. 2020). Curiosamente, em ambos os estudos citados acima também foi relatado um grande número de espécies do gênero *Neisseria* que cresceram a partir do material estomacal, no nosso estudo, porém, não foi detectada esta bactéria.

Já com relação a microbiota fúngica, apesar de ser menos estudada na literatura e de não ter sido o foco deste projeto, sabe-se que esta contribui de forma relevante para a microbiota humana, com estudos avaliando a presença de fungos em diferentes sítios anatômicos, tais como cavidade oral, nasal, trato gastrointestinal (principalmente intestino), pele, pulmão, etc., além de relacionando-os a diversas doenças, como alergias, doenças inflamatórias intestinais, dermatite, hepatite B, fibrose cística, entre outros (Ott et al. 2008; Nelson et al. 2010; Chen et al. 2011; Cui et al. 2013). Mais recentemente, um estudo identificou que a microbiota fúngica está relacionada ao surgimento do câncer de pâncreas, mostrando a importância de mais estudos relacionados a estes organismos (Aykut et al. 2019).

Revisando os fungos encontrados no intestino saudável, Hallen-Adams e Suhr (2017) observaram que *Candida* foi um dos gêneros mais detectados, estando presente em 86% dos estudos avaliados, enquanto que, em contraste, o gênero *Trichosporon* foi o menos citado dentre os fungos que se encontram nesta região, presente apenas em 17% dos estudos avaliados por eles. Dentre as espécies do gênero *Candida*, *C. albicans* foi a mais encontrada, seguida de *C. tropicalis* e *C. parapsilosis*. Na microbiota fecal de indivíduos saudáveis também já foi relatada uma associação positiva entre o gênero *Candida* e a presença de bactérias do gênero *Prevotella*, ambos associados a uma dieta rica em carboidratos (Hoffmann et al. 2013).

Em um estudo mais antigo, mas que abordou a região estomacal, foram identificadas diversas espécies do gênero *Candida*, inclusive a *C. parapsilosis*, tanto em biópsias como em fluidos gástricos (Zwolińska-Wcisło et al. 1998). Além disto, já foi observado que na candidemia, infecção da corrente sanguínea causada por fungos do gênero *Candida*, a *C. parapsilosis* é o principal fungo, fora a *C. albicans*, associado a esta infecção em pacientes com diversos tipos tumorais, sendo que o recebimento prévio de anti-fúngico foi um fator associado à presença desta espécie (Sun et al. 2019). A espécie *Trichosporon asahii*, por sua vez, apesar de já ter sido encontrada tanto fezes como na região perineal de brasileiros saudáveis (Silvestre et al. 2010; Gouba et al. 2014), é um patógeno oportunista, e que pode infectar severamente pacientes imunossuprimidos, incluindo pacientes oncológicos (Colombo et al. 2011; Magee et al. 2015).

Ao compararmos as espécies que cresceram em meio de cultura com a abundância relativa destas nas amostras de fluido gástrico sequenciadas, vemos que estas espécies foram identificadas tanto na microbiota avaliada por DNA como também nos dados de RNA, porém, ambas em frequências muito baixas (**Tabela 15**). Dentre os pacientes deste estudo, o paciente #9 foi o que apresentou menor porcentagem de bactérias do gênero *Streptococcus*, e a espécie mais frequente deste gênero foi a *S. pneumoniae* ao nível de RNA (1%) e de DNA (14%).

Tabela 15 - Bactérias cultivadas e suas frequências relativas na microbiota sequenciada de fluidos gástricos (DNA x RNA).

paciente 9	RNAseq (%)	DNA total (%)
<i>Streptococcus parasanguinis</i>	0,6	0,8
<i>Rothia dentocariosa</i>	0,4	0,1
<i>Staphylococcus aureus</i>	0,07	11,0
<i>Bacillus subtilis</i>	X	X
<i>Streptococcus salivarius</i>	0,1	0,1
<i>Streptococcus vestibularis</i>	X	0,006
<i>Streptococcus gordonii</i>	0,03	0,05
<i>Streptococcus sp LPB0220</i>	X	X
<i>Streptococcus parasanguinis</i>	0,9	0,8

Os dados de cultura, quando contrastados com os de sequenciamento nos mostram alguns fatos interessantes: i) as espécies mais abundantes nos dados de sequenciamento não vão necessariamente crescer em determinado meio de cultura; ii) espécies abundantes em termos de crescimento em cultura, não são necessariamente abundantes nas amostras originais, como indicado por dados de sequenciamento e iii) as espécies com baixa frequência também devem ser consideradas na análise de dados de sequenciamento, visto que nos fluidos gástricos, espécies com frequência de até 0,006% no sequenciamento cresceram em meio de cultura, o que já se destaca como um viés do nosso trabalho, visto que no geral focamos em bactérias com frequência maior que 1%, pensando que estas poderiam ter maior relevância na cavidade gástrica.

A abordagem de cultura neste estudo foi utilizada com meios enriquecidos mais gerais, sem termos procurado selecionar condições mais similares ao encontrado na região estomacal, talvez por isso certas bactérias mais abundantes mensuradas nas análises de sequenciamento não cresceram, visto que elas devem requerer condições ótimas de crescimento que não foram providas no nosso experimento. Além disto, é válido ressaltar que dos três pacientes que tiveram o fluido gástrico cultivado, obtivemos crescimento bacteriano apenas de um deles, e consequentemente só foi possível fazer comparações com as análises de *shotgun* e RNAseq para este, sendo necessário um maior número amostral para inferir achados mais relevantes.

De modo geral, a parte de cultura de bactérias neste estudo foi um *screening* inicial, visando trabalhos futuros do nosso grupo utilizando esta abordagem, com o intuito de confirmar possíveis achados via sequenciamento de microrganismos ativos transcricionalmente com potencial relevante em diversas neoplasias

4.6 RESULTADOS DO SEQUENCIAMENTO: rRNA 16S

4.6.1 Processamento das *reads*

O processamento dos dados de sequenciamento de *amplicons* (16S) teve início com a contagem do número total de sequências obtidas para cada amostra. Como é possível observar na **Tabela 16**, atingimos um número superior a 50 mil *reads* brutos para todas as amostras de fluido gástrico e de saliva, dentro do esperado nesta etapa inicial para atingir a saturação, conforme experiência prévia de nosso grupo. Um dos primeiros passos neste

processamento é a retirada de *reads* contaminantes, não-bacterianos, derivados do genoma humano. Para as amostras de fluidos gástricos, houve grande variação na porcentagem de *reads* mapeadas no genoma humano, indo de apenas cerca 8% nos pacientes #6, #7 e #9 até 60% no paciente #3. Já para as amostras de saliva, os pacientes #6 e #7 tiveram menos de 1% de *reads* humanas, e a porcentagem máxima deste contaminante foi cerca de 17% no paciente #9. É interessante ressaltar que a maioria das *reads* foi eliminada após filtro de qualidade com o *deblur*, sendo que só duas amostras de fluido gástrico tiveram menos de 50% de *reads* removidas por este filtro (pacientes #3 e #4) e uma de saliva (#9). O *deblur* é específico para erros de sequenciamento gerados pelo Illumina, sendo importante para avaliar a qualidade dos *reads* finais. Após todo o processamento, a porcentagem de *reads* finais variou entre 5 a 26% em relação ao número inicial sequenciado. Mesmo assim, todas as amostras atingiram mais de 3 mil *reads* com uma mediana maior que 11,5 mil sequências, o que permitiu uma boa cobertura da diversidade bacteriana, conforme as curvas de saturação das amostras (**Figura 16**).

Tabela 16 - Processamento das *reads*: amostras sequenciadas (16S) a partir de fluido gástrico e saliva.

Paciente	<i>reads</i> totais	<i>reads</i> com adaptador (%)	Removidas (humano) (%)	<i>reads</i> pareadas (%)	Removidas (deblur) (%)	<i>reads</i> finais	(%)
fluido gástrico							
1	149.143	97,97	11,04	85,55	59,54	38.792	26,01
2	96.441	97,46	14,81	73,49	54,53	18.287	18,96
3	55.666	97,64	60,81	24,03	18,50	3.076	5,53
4	71.234	92,16	30,21	61,56	38,06	16.736	23,49
5	103.428	97,10	10,12	84,28	71,41	13.313	12,87
6	78.342	93,60	8,69	68,37	53,63	11.547	14,74
7	75.927	94,27	8,94	67,65	54,41	10.054	13,24
8	81.025	94,62	11,45	60,87	50,88	8.091	9,99
9	80.608	94,12	8,71	67,97	54,63	10.754	13,34
Média	87.979.33					14.516.67	
Mediana	80.608.00					11.547.00	
Saliva							
6	77.109	94,29	0,02	78,73	61,68	13.148	17,05
7	99.037	94,21	0,02	77,63	60,17	17.297	17,47
8	93.600	94,48	6,88	72,13	57,53	13.663	14,60
9	103.565	94,10	16,96	56,90	44,85	12.481	12,05
Média	93.327.75					14.147.25	
Mediana	96.318.5					13.405.5	

Ainda de acordo com a **Figura 16**, é possível observar que em geral as amostras sequenciadas permitiram identificar de 50 a 100 OTUs, e que as biópsias cujo cDNA foi sequenciado no geral saturaram antes, ou seja, menos bactérias foram identificadas. Além disto, curiosamente o paciente #3 teve um número bem maior de OTUs na biópsia (DNA) do que nos fluidos (saliva e fluido gástrico).

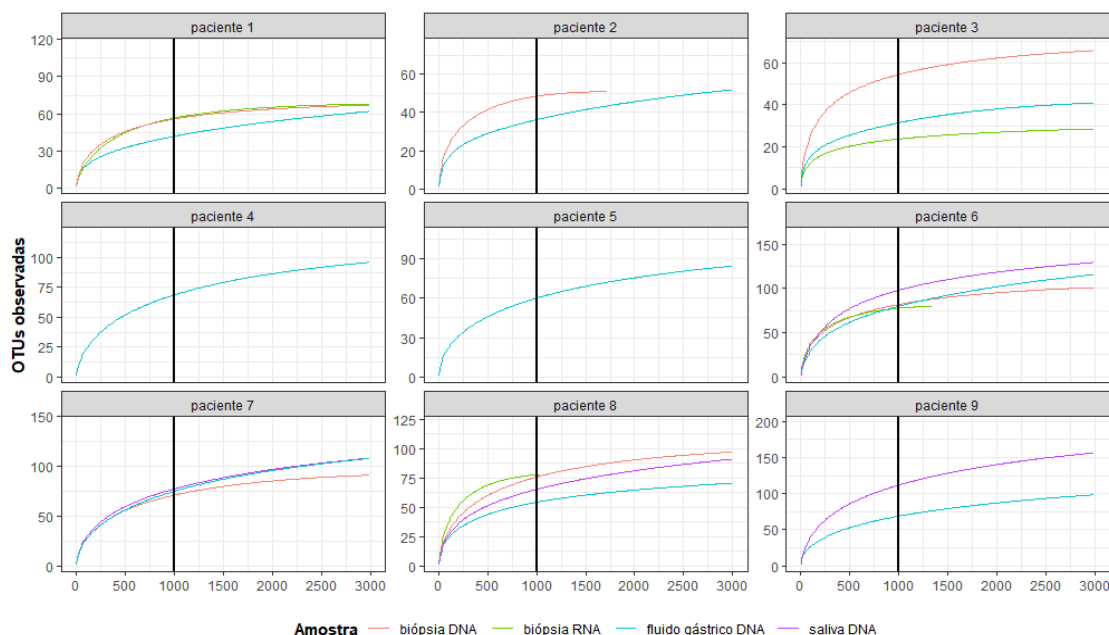


Figura 16 - Curvas de rarefação das amostras sequenciadas por 16S. Para esta abordagem de sequenciamento foram avaliados fluidos gástricos (n=9), salivas (n=4) e biópsias (DNA, n=6 e RNA, n=4). O eixo y do gráfico indica o número de OTUs (*Operational Taxonomic Units*) identificadas por número de *read* sequenciada.

Em concordância com os resultados obtidos previamente pela quantificação do DNA bacteriano por qPCR, quando avaliamos o sequenciamento 16S das biópsias, observamos que o número de *reads* após processamento foi menor do que o obtido para os fluidos gástricos, Na **Tabela 17**, vemos que para a maioria das amostras mais de 50% das *reads* foram mapeadas no genoma humano, atingindo 84% na biópsia-DNA do paciente #9. Somente as biópsias do paciente #3 ficaram abaixo de 50%, com 47% de *reads* humanos ao avaliarmos a partir de DNA e 23% a partir do RNA/cDNA, o que vai de encontro ao resultado da qPCR, visto que este paciente era o que possuía a maior razão entre genomas bacterianos/humano dentre as biópsias (1,17). Quando observamos o número de *reads* finais, com relação às biópsias (DNA), três destas ficaram com menos de um mil *reads*, não atingindo a saturação mínima estabelecida para identificação de todas as bactérias presentes na amostra. Para as leituras de 16S a partir de RNA, este número foi ainda pior, sendo que para 5/9 amostras não foi possível obtermos *reads* em quantidade suficiente para seguirmos com as análises (amostras em destaque na **Tabela 17**). Diante disso, seguimos com as análises com seis biópsias sequenciadas a nível de DNA, e com apenas quatro ao nível de RNA.

Tabela 17 - Processamento das *reads*: amostras sequenciadas (16S) obtidas a partir de biópsias (DNA e RNA).

Paciente	<i>reads</i> totais	<i>reads</i> com adaptador (%)	Removidas (humano) (%)	<i>reads</i> pareadas (%)	Removidas (deblur) (%)	<i>reads</i> finais	(%)
Biópsia DNA							
1	107.485	97,88	69,65	27,05	23,07	4.276	3,98
2	97.766	97,22	73,02	16,61	14,85	1.719	1,76
3	78.549	97,68	47,21	47,14	41,40	4.504	5,73
4	120.145	97,43	78,46	4,96	4,63	395	0,33
5	121.399	97,56	78,2	6,45	5,83	754	0,62
6	93.410	97,22	52,18	42,36	37,82	4.243	4,54
7	85.972	97,90	61,35	34,28	30,52	3.232	3,76
8	85.353	97,89	51,15	44,13	39,01	4.371	5,12
9	85.474	97,63	84,25	6,69	6,00	591	0,69
média	97.283.67					2.676.11	
mediana	93.410.00					3.232.00	
Biópsia RNA							
1	70.954	97,22	51,29	36,32	31,99	3.070	4,33
2	82.481	97,68	76,8	5,23	5,06	136	0,16
3	76.597	97,31	23,75	68,98	60,66	6.377	8,33
4	94.478	97,13	81,61	3,03	3,00	30	0,03
5	81.057	97,51	78,95	5,52	5,45	57	0,07
6	72.810	97,41	72,47	18,91	17,06	1.342	1,84
7	60.102	97,35	77,36	11,46	10,93	319	0,53
8	84.128	97,45	74,76	14,50	13,27	1.040	1,24
9	87.928	97,47	77,42	3,94	3,79	133	0,15
média	78.948.33					1.389.33	
mediana	81.057.00					319.00	

Um estudo anterior que avaliou a região V3-V4 do gene 16S em biópsias de pacientes com CG, porém em uma casuística de 80 pacientes, também estabeleceu um mínimo de 1.000 *reads* após filtros de qualidade para seguirem nas análises. Apesar de ter foco principal nesta abordagem, 26 pacientes foram descartados (32,5% das amostras), por não atingirem este valor mínimo de cobertura (Yu et al. 2017). Tais achados demonstram que mesmo sequenciando *amplicons* específicos, ainda existe uma perda importante de amostras, seja pela qualidade das *reads* sequenciadas ou por contaminantes (principalmente DNA humano). Cabe notar que no caso de biópsias, por mais que tenhamos especificidade dos iniciadores usados, a presença de DNA do hospedeiro em grande quantidade - em paralelo com quantidade muito baixa do DNA alvo - vai comumente gerar *amplicons* inespecíficos.

4.6.2 Composição da microbiota bacteriana das biópsias avaliada por tipo de molécula sequenciada (DNA x RNA)

No total identificamos 50 gêneros nas amostras de biópsias, sendo que destes, 52% (26 gêneros) estavam presentes tanto no DNA quanto no RNA, enquanto 32% foram encontrados apenas na microbiota ao nível de DNA e 16% apenas na microbiota de RNA. Quando observamos a abundância relativa dos gêneros exclusivos de cada material, porém, vemos que eles não chegam a 2% da abundância total, sugerindo que a falha em encontrar certos microrganismos nas *reads* de DNA e de RNA, seja um problema de amostragem. No geral, 98% da composição bacteriana é similar entre as biópsias independente da molécula sequenciada, demonstrando uma concordância entre as abordagens utilizadas, visto que a maioria das bactérias é encontrada tanto ao nível de DNA como de RNA. Tal achado sugere que as bactérias encontradas nas biópsias são residentes e estejam vivas transcricionalmente neste tecido.

Dentre as bactérias exclusivas de cada material, os gêneros *Abiotrophia*, *Chryseobacterium*, *Dialister*, *Kingella*, *Massilia* e *Megasphaera*, presentes exclusivamente no DNA, já foram relatados como contaminantes de sequenciamento - fato que não descarta que sejam de fato relevantes em nosso contexto, mas levanta um ponto de atenção - além dos gêneros *Corynebacterium* e *Methylobacterium*, exclusivos no RNA (Salter et al. 2014; Eisenhofer et al. 2019). Assim, é possível que tais bactérias também possam ser fatores de confusão e potenciais contaminantes em nossas análises. Tais contaminantes são sempre detectadas mais facilmente em amostras com menor abundância de bactérias (Glassing et al.

2016) e podem ser decorrentes do processo de extração de DNA/RNA ou ainda dos reagentes de amplificação e sequenciamento. Apesar disto, gêneros em alta frequência em ambos os materiais, tais como *Veillonella*, *Prevotella*, *Haemophilus* e até mesmo o gênero *Streptococcus*, também já foram relatados como contaminantes, mostrando que mais do que a presença de determinado organismo na amostra, a sua frequência e o contexto biológico no qual ela se encontra também deve ser levada em consideração para discriminarmos uma possível contaminação, ou até mesmo algum erro de anotação do sequenciamento. Diante disto, considerando ser menos provável que bactérias em frequências tão baixas sejam relevantes no contexto biológico destas amostras, focamos as nossas análises em bactérias que apresentaram frequências mais elevadas (>1%).

Ao avaliarmos a composição bacteriana das amostras (DNA x RNA) com abundância maior que 1% em pelo menos um material, vemos que de modo geral os gêneros que mais se destacam são similares (**Figura 17**). Na microbiota ao nível de DNA, *Prevotella* (20%) é o gênero mais abundante, seguido de *Streptococcus* (17%), *Actinobacillus* (14%) e *Haemophilus* (10%). Na microbiota transcricionalmente ativa, porém, o gênero *Treponema* representou 25% da amostra, seguido de *Streptococcus* (21%), *Alloprevotella* (13%) e *Haemophilus* (8%). O gênero *Prevotella*, abundante na microbiota total, representou apenas 5% ao nível de RNA, enquanto o gênero *Alloprevotella*, representando cerca de 13% da composição bacteriana ativa, constitui apenas 4% da microbiota ao nível de DNA.

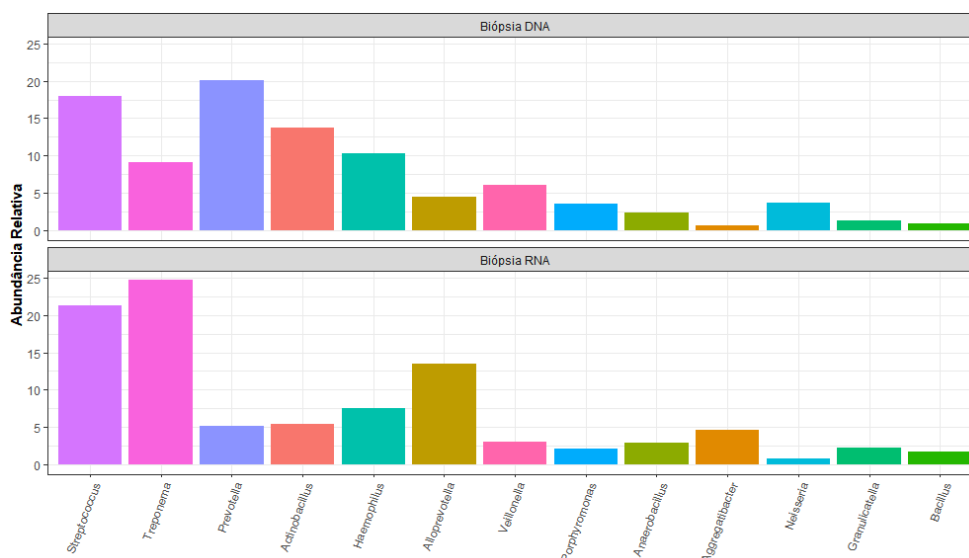


Figura 17 - Abundância relativa dos gêneros bacterianos com frequência superior a 1% nas amostras de biópsia, estratificado por material sequenciado (DNA x RNA).

De modo geral, os gêneros mais abundantes observados neste estudo vão de acordo com o já observado na literatura em biópsias gástricas de indivíduos sem infecção por *H. pylori*, com destaque para o gênero *Streptococcus* e *Prevotella*, altamente frequentes em indivíduos com CG (Dicksved et al. 2009; Ferreira et al. 2018) e sem câncer, tanto no seu DNA (Li et al. 2009; Wurm et al. 2018) quanto no RNA (Schmieder et al. 2011). Já a predominância do gênero *Treponema* a nível de RNA, deve-se ao fato de que no paciente #3, 99% da sua microbiota parece representar quase unicamente este gênero, causando um viés ao avaliar todos os pacientes de forma conjunta, visto que nosso número amostral é de apenas de quatro pacientes a nível de RNA para as biópsias.

O gênero *Treponema*, predominante no paciente #3 é conhecido por estar associado a doenças periodontais (You et al. 2013; Chen et al. 2018). Já quando investigada a relação entre o gênero *Treponema* e a biópsia estomacal, alguns trabalhos associam a presença da espécie *Treponema pallidum* em biópsias de estômago como causadora da sífilis gástrica, forma rara desta doença sexualmente transmissível, representando apenas 1% dos casos de sífilis (Guimarães et al. 2016; Okamoto et al. 2018). Entretanto, como a abordagem por 16S geralmente não permite chegar ao nível de espécie, não podemos inferir que o paciente tenha de fato esta bactéria específica ou alguma outra espécie deste mesmo gênero. Também vale ressaltar que este gênero só foi detectado neste único paciente e não encontramos no prontuário clínico deste caso qualquer evidência de sífilis gástrica neste indivíduo.

Ao avaliarmos diferenças significativas entre a abundância relativa dos gêneros com frequência maior que 1% nos diferentes tipos de moléculas, foi possível observar mesmo com o pequeno número amostral do estudo que o gênero *Prevotella* está mais abundante na microbiota ao nível de DNA quando comparado à microbiota transcricionalmente ativa ($p=0,04$) (**Figura 18**). No demais, não foram encontrados outros gêneros diferencialmente abundantes.

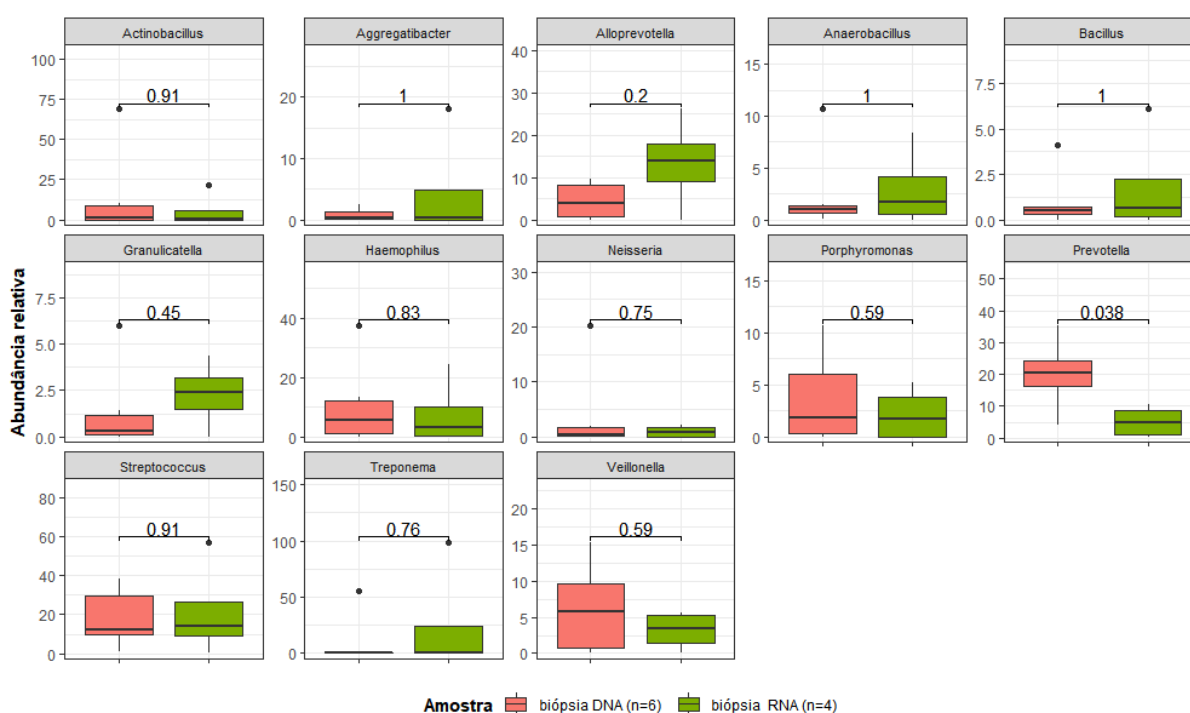


Figura 18 - Análise comparativa da abundância dos gêneros bacterianos com frequência superior a 1% nas amostras de biópsia (DNA x RNA). Foram selecionados os gêneros que possuíam abundância relativa maior que 1% em pelo menos um material sequenciado de biópsia (DNA ou RNA). Para análise estatística foi utilizado o teste não paramétrico *U de Mann-Whitney*, para amostras não pareadas. Os gêneros foram considerados como diferencialmente abundantes quando $p \leq 0,05$.

Recentemente, Wurm et al. (2018) também compararam a microbiota da biópsia estomacal através da abordagem RNA e DNA e identificaram o filo Firmicutes mais abundante na microbiota ativa transcricionalmente, do qual o gênero *Streptococcus* faz parte, além do filo *Bacteroidetes* como mais abundante na microbiota ao nível de DNA, no qual o gênero *Prevotella* se encontra, e que foi identificado como mais abundante na microbiota de DNA no nosso estudo. O gênero *Prevotella* já foi extensamente relatado como presente tanto na

microbiota oral, associado a doenças periodontais (Tanaka et al. 2008; Ibrahim et al. 2017; Zhang et al. 2017) e até mesmo relacionada com câncer oral (Gholizadeh et al. 2016), e visto a diferença significativa deste gênero entre as moléculas analisadas (DNA *versus* RNA) este pode ser um bom candidato a ser encontrado na saliva, e que por isso foi detectado em maior frequência ao nível de DNA. Além disso, devido a poucos relatos na literatura abordando a microbiota comparativa entre DNA e RNA, de modo ainda mais específico no estômago, não foi possível encontrar mais suporte para nossos achados na literatura.

Quando avaliamos a existência de uma correlação entre a abundância relativa das bactérias ao nível de DNA versus RNA, por sua vez, observamos que existe uma associação positiva entre os dois tipos de amostra ($r\hat{o}=0,82$ e $p<0,05$), ou seja, a abundância de determinado gênero é similar em ambos, e quanto maior a abundância deste no DNA, maior é a tendência de também ser observada maior abundância ao nível de RNA (**Figura 19**), o que sugere que a maioria dos gêneros encontrados a partir do DNA está de fato transcricionalmente ativa no tecido estomacal.

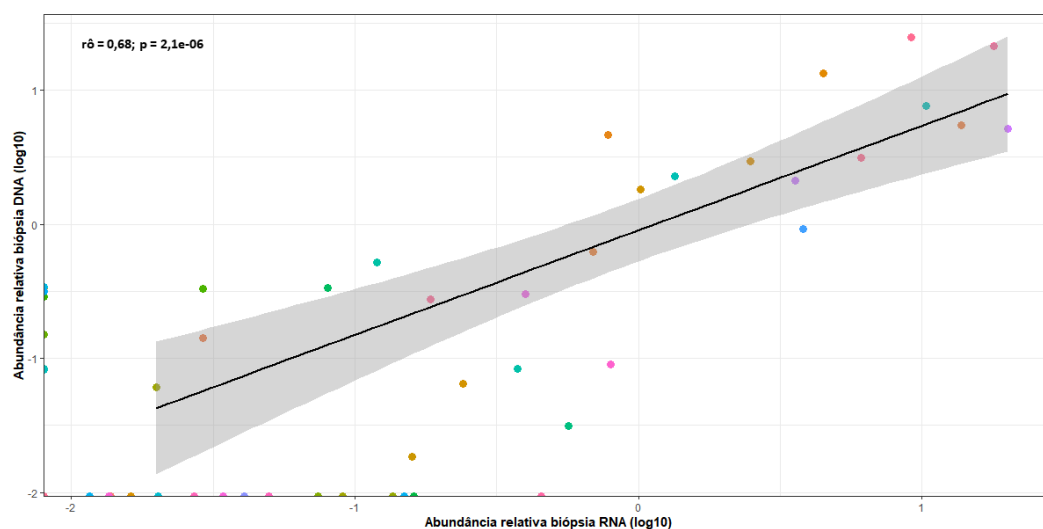


Figura 19 - Correlação da abundância relativa (log10) dos gêneros bacterianos presentes nas amostras de biópsias (DNA x RNA). Cada círculo representa um gênero bacteriano. A linha diagonal indica abundâncias similares nos conjuntos de dados de DNA e RNA, sendo que os gêneros mais próximos da linha possuem maior correlação entre os materiais. Para análise estatística foi utilizada a correlação de *Spearman*, sendo que quanto $r\hat{o}$ mais perto de 1, maior é a correlação entre as variáveis. A área sombreada representa o intervalo de confiança. A correlação foi considerada como estatisticamente significativa quando $p\leq 0,05$.

Para comparar a diversidade bacteriana das amostras, utilizamos mais uma vez as métricas de diversidade alfa e beta. Como mostra a **Figura 20**, não houve diferença significativa entre a riqueza de bactérias ao nível de DNA *versus* RNA nas biópsias dos pacientes (gráfico *Observed*), nem quando levamos em conta suas abundâncias relativas (índices de *Shannon* e *Simpson*). Com relação a diversidade beta, também não observamos diferenças significativas entre as amostras, o que apesar de poder estar relacionado ao nosso limitado número amostral, também pode nos indicar que as amostras tendem a apresentar diversidades similares.

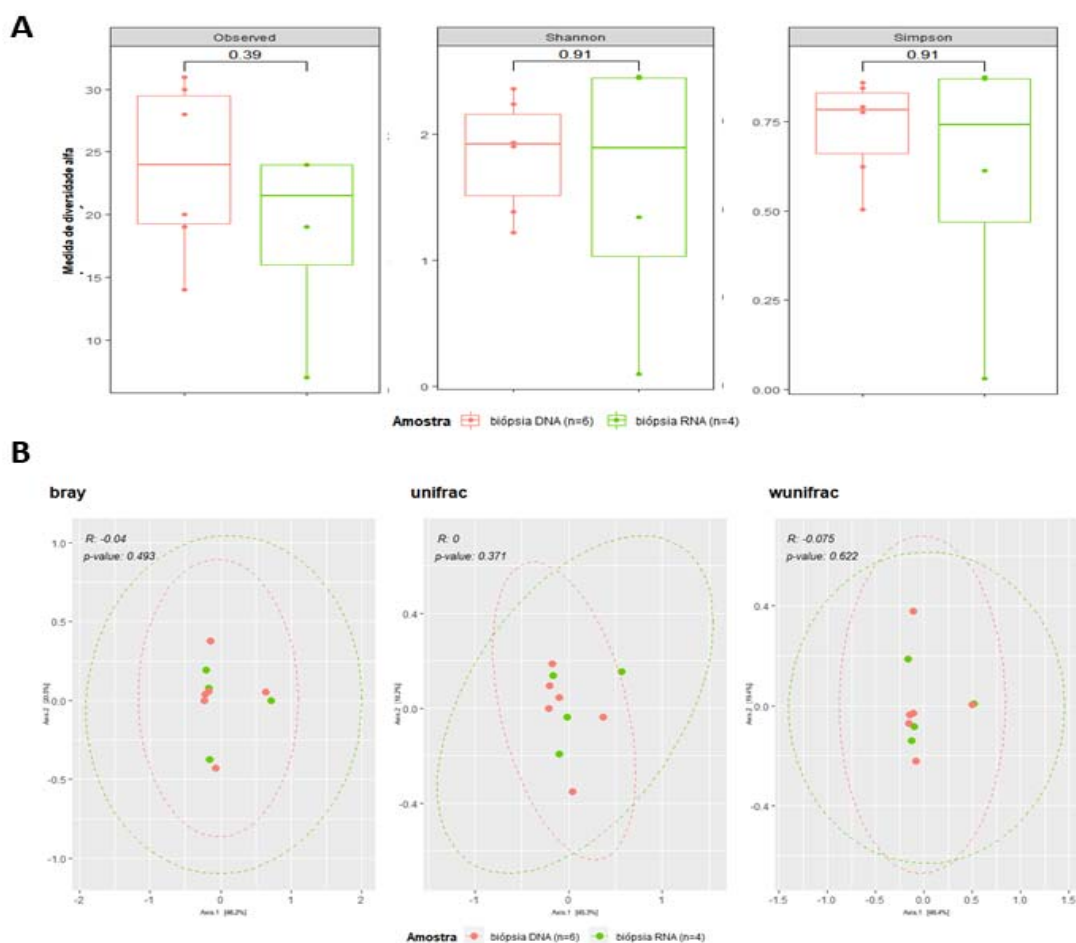


Figura 20 - Medidas de diversidade nas amostras de biópsia (DNA x RNA). (A) A diversidade alfa foi avaliada pelos índices *Observed*, *Shannon* e *Simpson* respectivamente, sendo que o teste não paramétrico *U de Mann-Whitney* foi utilizado para verificar a existência de diferença significativa entre as amostras. (B) A diversidade beta foi avaliada pelos índices de *Bray-Curtis*, e *Unifrac* (*unweighted* e *weighted*) respectivamente, e o teste não paramétrico *ANOSIM* foi utilizado para avaliar a similaridade entre as amostras, sendo o *R* seu coeficiente de correlação. A diferença foi considerada como estatisticamente significativa quando $p \leq 0,05$.

Observamos em linhas gerais que as análises da microbiota das biópsias a partir de DNA ou de RNA apresentam resultados similares em termos de composição, visto tanto por sua alta correlação, como observado pela similaridade de suas diversidades. Dentre as bactérias mais frequentes, apenas o gênero *Prevotella* se destacou como diferencialmente abundante. Em outros tipos de amostra, como mucosa intestinal, já foram observadas diferenças entre a microbiota ao nível de DNA e RNA (Heinsen et al. 2015; Moen et al. 2018), porém visto que a biópsia estomacal já é um tecido com baixo número de bactérias, as que estão geralmente presentes parecem estar, de fato, ativas e colonizando o tecido gástrico.

4.6.3 Composição da microbiota bacteriana: região estomacal x cavidade oral

Com o intuito de avaliar a similaridade entre a microbiota bacteriana ao nível de DNA da cavidade estomacal versus oral, foram selecionados os sequenciamentos apenas provenientes da abordagem de *amplicons* (16S), já que a abordagem de sequenciamento por DNA total poderia gerar algum viés nas análises, visto que não só o método de sequenciamento foi diferente, mas também as etapas de processamento das *reads* e o banco usado para classificação taxonômica.

No total, detectamos 62 gêneros em fluido gástrico, seguido de 58 na saliva e 42 na biópsia. Ao compararmos os gêneros exclusivos dos diferentes materiais, observamos que fluido gástrico e saliva tiveram mais gêneros exclusivamente encontrados quando comparados com a biópsia, além de terem um número maior de gêneros em comum (**Figura 21**). Quando observamos a frequência dos gêneros exclusivamente encontrados em cada material, em geral observamos mais uma vez baixa frequência destas bactérias exclusivas, sendo que comparando fluido gástrico e saliva os gêneros exclusivos representam menos de 1% da composição total de cada material, já quando comparamos biópsia e fluido gástrico os gêneros exclusivos compõem menos de 4% de cada amostra, enquanto que ao avaliar biópsia e saliva os gêneros exclusivos da biópsia representam 13% de sua abundância (sendo que na saliva estes representam apenas 4%). Ou seja, o fluido gástrico parece ser mais similar em composição com a saliva, e a biópsia por sua vez parece ser mais similar ao fluido gástrico, o que faz sentido visto que a saliva deglutida fica em constante transição pelo estômago, ficando em contato com os fluidos da região.

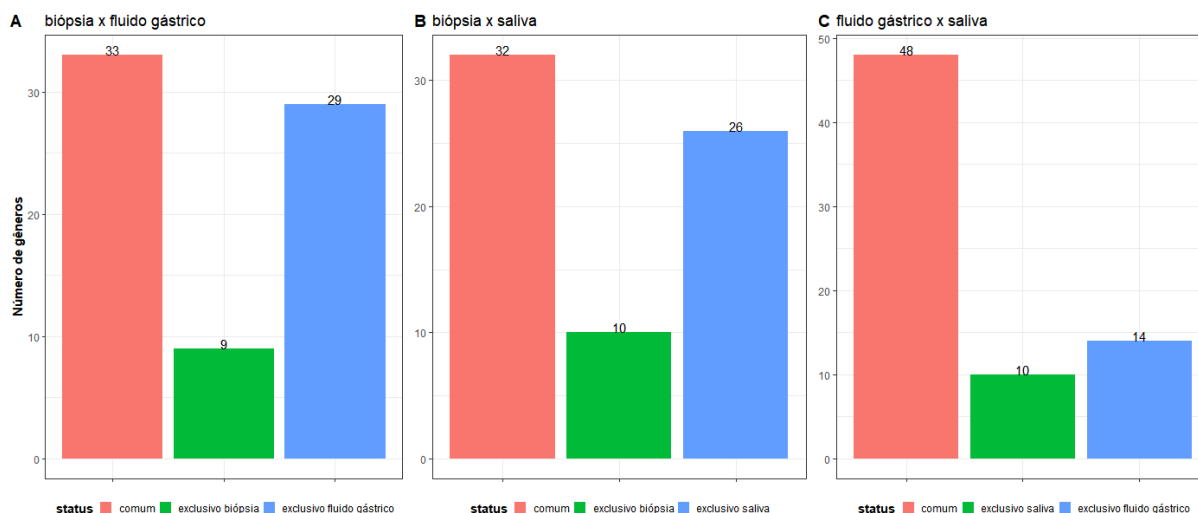


Figura 21 - Comparação do número de gêneros encontrados (A) na biópsia x fluido gástrico, (B) biópsia x saliva e (C) fluido gástrico x saliva.

Quando observamos os gêneros mais frequentes nas amostras (>1%), podemos ver que no fluido gástrico predomina o gênero *Streptococcus*, compondo mais de 50% da amostra, já na biópsia este gênero se encontra em torno de 18% e na saliva 12%. Curiosamente, o gênero *Prevotella*, o mais presente na saliva, com 24%, foi mais presente na biópsia (20%) do que no fluido gástrico (apenas 8%). Este dado parece corroborar o que foi observado quando as biópsias foram comparadas (DNA x RNA) pois o gênero *Prevotella* se destacou ao nível de DNA, e parece realmente vir da cavidade oral. Já os gêneros *Veillonella* e *Neisseria*, também se encontraram com maior frequência na saliva quando comparado com as amostras estomacais (**Figura 22**).

Mesmo com nosso número amostral limitante, os gêneros mais abundantes encontrados na saliva, como *Streptococcus*, *Veillonella*, *Prevotella*, *Alloprevotella*, *Neisseria*, *Haemophilus*, *Porphyromonas* e *Fusobacterium*, vão de acordo com estudos que definem tanto o core da microbiota oral saudável (Zaura et al. 2009), como também já foram detectados em pacientes com doenças na cavidade oral, como doenças periodontais e câncer (Chen et al. 2015; Zhao et al. 2017).

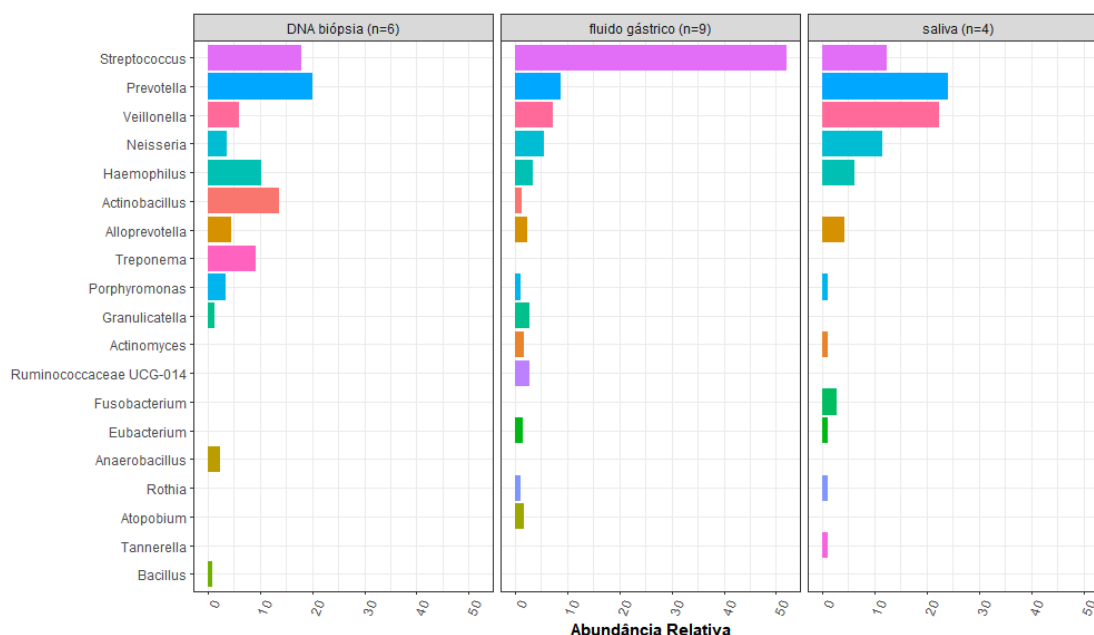


Figura 22 - Abundância relativa dos gêneros bacterianos com frequência superior a 1% nas amostras de biópsia, fluido gástrico e saliva.

Dentre os gêneros mais abundantes, os únicos que apresentaram frequência significativamente diferente entre as amostras foram *Anaerobacillus*, *Bacillus*, *Prevotella* e *Streptococcus* (**Figura 23**). Os gêneros *Anaerobacillus* e *Bacillus* só foram detectados nas amostras de biópsia, com frequências medianas de 1% e 0,5% respectivamente. O gênero *Prevotella*, por sua vez, está menos abundante nas amostras de fluido gástrico com relação às biópsias (mediana de 4% *versus* 20%), enquanto que o gênero *Streptococcus* está mais abundante nas amostras de fluido gástrico, com mediana de 51% nestas amostras, *versus* apenas 12% nas amostras de biópsia e saliva. Curiosamente nenhum gênero foi detectado como mais abundante na saliva, isso pode se dever tanto ao fato do nosso número amostral limitado, porém também pode indicar que praticamente todas as bactérias presentes na saliva também são encontradas no estômago, levando a esta alta similaridade entre ambos.

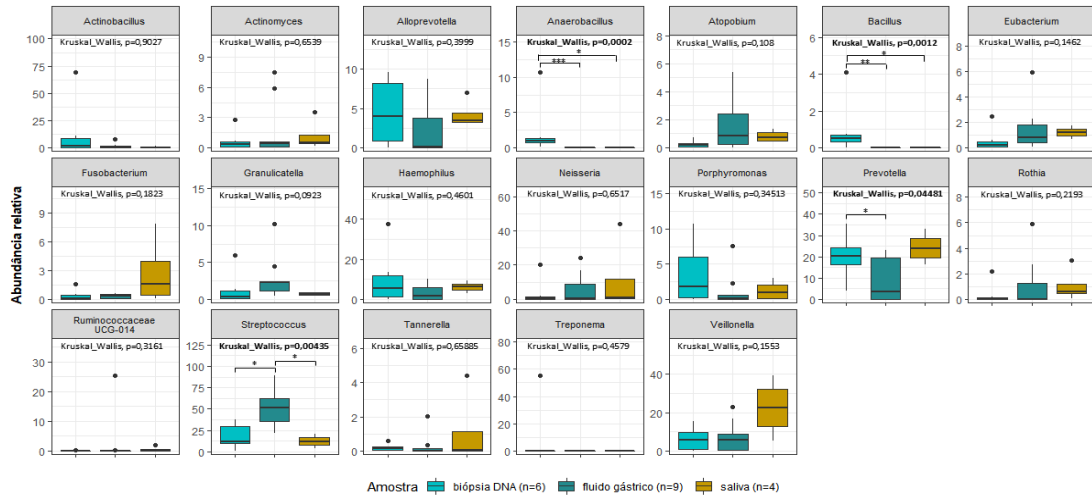


Figura 23 - Análise comparativa da abundância dos gêneros bacterianos com frequência superior a 1% nas amostras de biópsia, fluido gástrico e saliva. Foram selecionados os gêneros que possuíam abundância relativa maior que 1% em pelo menos um material sequenciado (biópsia, fluido gástrico ou saliva). Para análise estatística foi utilizado o teste não-paramétrico de comparação múltipla *Kruskal-Wallis* para amostras não-pareadas, seguido do teste *U de Mann-Whitney* quando necessário. Os gêneros foram considerados como diferencialmente abundantes quando $p \leq 0,05$

Ao avaliar se existe uma associação entre as bactérias encontradas na cavidade oral x estomacal, constatamos que existe uma correlação mais positiva entre as amostras de saliva x fluido gástrico ($r = 0,8$ e $p < 0,05$), do que quando comparamos biópsia x saliva, e até mesmo biópsia x fluido gástrico, indicando que a microbiota dos fluidos é mais similar entre si (saliva x fluido gástrico) do que a microbiota estomacal (biópsia x fluido gástrico) (**Figura 24**).

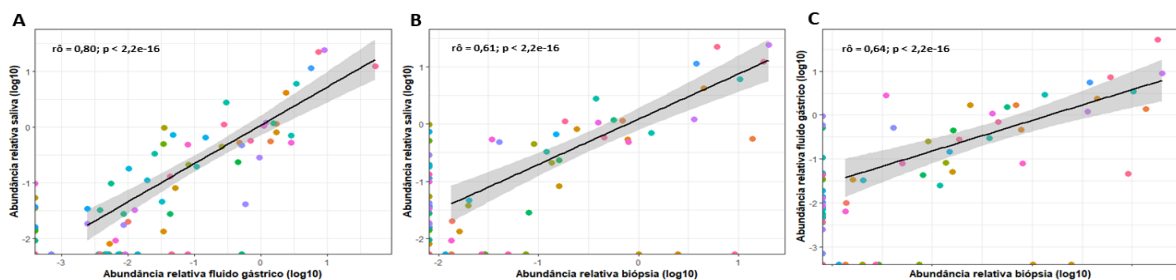


Figura 24 - Correlação da abundância relativa (log10) dos gêneros bacterianos presentes nas amostras de (A) saliva x fluido gástrico, (B) saliva x biópsia e (C) fluido gástrico x biópsia. A linha diagonal indica abundâncias similares nos conjuntos de dados das variáveis analisadas, sendo que os gêneros mais próximos da linha possuem maior correlação entre as amostras. Para análise estatística foi utilizada a correlação de Spearman, sendo que quanto r mais perto de 1, maior é a correlação entre as variáveis. A área sombreada representa o intervalo de confiança. A correlação foi considerada como estatisticamente significativa quando $p \leq 0,05$.

Avaliando a diversidade das amostras, destaca-se maior riqueza de gêneros na saliva, seguido do fluido gástrico e biópsia (diversidade alfa, índice *Observed* – **Figura 25A**), porém quando a abundância relativa dos gêneros é levada em consideração, não houve diferença estatisticamente significativa entre as amostras. Já quando avaliamos a diversidade beta, foi observado que a composição bacteriana entre os materiais é diferente, sendo que quando comparamos biópsia versus fluido gástrico isto foi verdade para todos os índices analisados (**Figura 25B**). Já saliva e biópsia foram consideradas como diferenciadas quando levado em conta apenas a distância filogenética entre as bactérias (índice *Unifrac unweighted*), enquanto que saliva e fluido gástrico se agrupam de modo distinto nas métricas no qual levamos em consideração a abundância relativa das bactérias (*Bray-curtis* e *Unifrac weighted*), o que nos leva a inferir que embora a saliva tenha gêneros em comum com as amostras provenientes do estômago, os materiais não são tão similares entre si. É importante ressaltar, mais uma vez, que o número amostral tanto da saliva como da biópsia é pequeno (n=4 e n=6 respectivamente), o que pode causar algum tipo de viés nas nossas análises, visto que a correlação entre as amostras de saliva e fluido gástrico foi maior, porém com relação a diversidade beta a saliva foi mais similar à biópsia.

Estudos já mostraram que o muco parece permitir o crescimento e colonização de certas bactérias, visto que seu pH não é tão ácido como o do estômago, podendo explicar em partes a semelhança entre estes materiais (Yang et al. 2013; Leung et al. 2019; Herath et al. 2020). Tal fato pode sugerir um mecanismo onde certas bactérias são deglutidas e atravessam a camada de muco, indo colonizar o tecido gástrico onde passam a habitar.

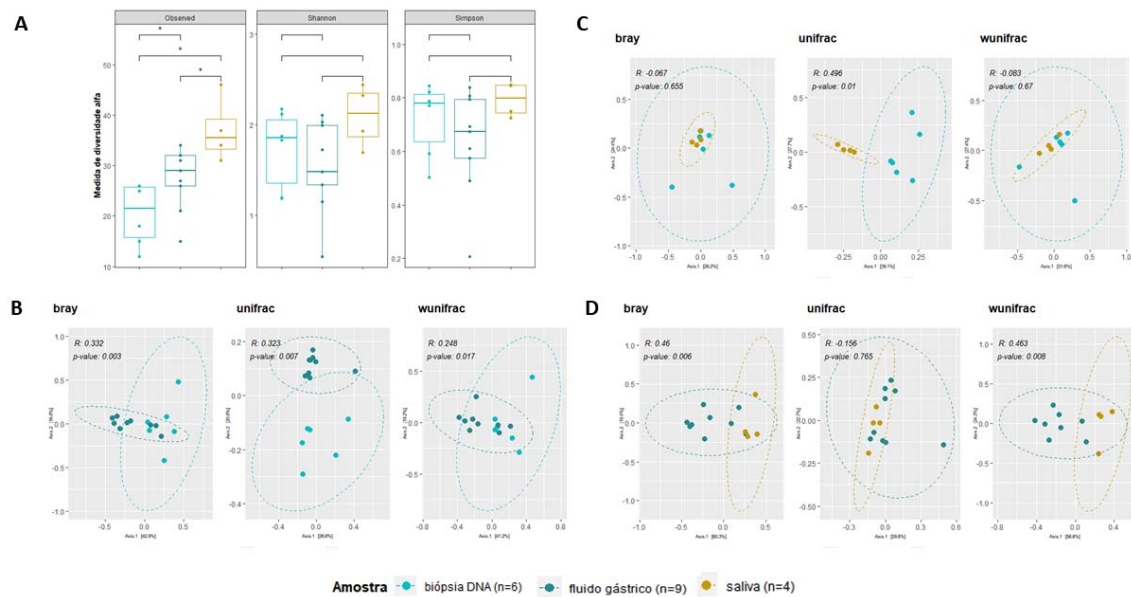


Figura 25 - Medidas de diversidade nas amostras de biópsia, fluido gástrico e saliva. (A) A diversidade alfa foi avaliada pelos índices *Observed*, *Shannon* e *Simpson* respectivamente, sendo que o teste não paramétrico *U de Mann-Whitney* foi utilizado para verificar a existência de diferença significativa entre as amostras. A diversidade beta comparando (B) biópsia x fluido gástrico, (C) biópsia x saliva e (D) fluido gástrico x saliva foi avaliada pelos índices de *Bray-Curtis*, e *Unifrac* (*unweighted* e *weighted*) respectivamente, e o teste não paramétrico ANOSIM foi utilizado para avaliar a similaridade entre as amostras, sendo o R seu coeficiente de correlação. A diferença foi considerada como estatisticamente significativa quando $p \leq 0,05$.

No geral, as amostras de saliva e fluido gástrico possuem maior número de gêneros em comum, além de maior correlação entre a proporção dos gêneros encontrados ($r = 0,8$; $p < 0,05$). Dentre os gêneros com frequência $> 1\%$, apenas *Streptococcus* está mais abundante nos fluidos gástricos ($p < 0,05$) enquanto o gênero *Prevotella* é mais abundante nas amostras de saliva e biópsia, indicando que esta bactéria migra para o estômago e possivelmente se associa ao tecido ao ser deglutida. Além disso, ao compararmos biópsia x saliva também observamos uma associação positiva mediana significativa ($r = 0,61$; $p < 0,05$). É provável que alguns gêneros bacterianos estão em proporções mais similares entre saliva e fluido gástrico, enquanto outros se assemelham entre saliva e biópsia, indicando que algumas bactérias da cavidade oral podem migrar para a região do tecido, enquanto outras permanecem no lúmen estomacal. Tal associação diferencial deve estar relacionada às características fisiológicas de cada bactéria e sua capacidade de sobreviver perante diferentes nutrientes e condições ambientais.

5 CONCLUSÕES

Análise quantitativa do genoma humano e bacteriano por qPCR

- Observamos para as amostras de fluido gástrico mediana de 12 mil cópias de genomas bacterianos *versus* apenas 12 cópias de genoma humano por ng de amostra, indicando o alto número de bactérias neste material e confirmando a viabilidade dos estudos de metagenômica por DNA total;
- As amostras de biópsias mostraram uma mediana de 10 cópias de genoma bacteriano *versus* 31 cópias de genoma humano por ng de amostra, indicando baixos números de bactéria neste material, e que a melhor abordagem de metagenômica para este caso é o sequenciamento de *amplicons*.

Microbiota bacteriana de fluidos gástricos

- A bactéria *Streptococcus pneumoniae* se destacou como a mais abundante tanto a nível de DNA (11%) como a nível de RNA (28%), indicando sua alta atividade transcricional na região estomacal e seu papel como provável bactéria residente;
- As bactérias *Bacillus cereus*, *Elizabethkingia anophelis*, *Legionella pneumophila* e *Virgibacillus senegalensis* estão mais abundantes ao nível de RNA, indicando que apesar de pouco frequentes ao nível de DNA, elas devem estar muito ativas transcionalmente, sendo possivelmente bactérias residentes;
- As bactérias *Solobacterium moorei* e *Streptococcus pseudopneumoniae* foram mais abundantes no DNA, indicando que estas possam estar apenas passando pela região (bactérias transientes);
- O sequenciamento de DNA e RNA apresentou populações distintas, indicando que as duas abordagens são necessárias e complementares na avaliação da microbiota estomacal;
- Observamos crescimento de espécies bacterianas com baixa frequência no sequenciamento, mostrando que as bactérias pouco frequentes também são relevantes e estão vivas na região estomacal.

Microbiota bacteriana de biópsias

- Os gêneros mais abundantes tanto a nível de DNA como de RNA foram *Streptococcus*, *Treponema*, *Actinobacillus*, *Haemophilus* e *Alloprevotella* (bactérias residentes);
- O gênero *Prevotella* foi mais abundante nas amostras de DNA e nas de saliva, indicando que este gênero deve estar apenas em trânsito pela região (bactéria transiente) sem atividade transcricional significativa no microambiente estomacal;
- Para a biópsia estomacal, que já é um tecido com menor número de bactérias, nossos dados sugerem que as bactérias encontradas a nível de DNA são de fato ativas e residentes, colonizando este tecido.

Microbiota da cavidade estomacal x oral

- Nenhum dos gêneros bacterianos avaliados apresentou maior abundância na saliva quando comparado com biópsia ou fluido gástrico, indicando similaridade na composição bacteriana entre as regiões;
- As amostras da cavidade estomacal se parecem mais com as da saliva do que com o outro material da mesma região (biópsia), indicando que algumas bactérias provenientes da região estomacal se deslocam para a região do tecido, enquanto outras permanecem no lúmen estomacal.

6 REFERÊNCIAS

Abbott A. Scientists bust myth that our bodies have more bacteria than human cells. *Nature*. 2016.

Achilli P, De Martini P, Ceresoli M, Mari GM, Costanzi A, Maggioni D, et al. Tumor response evaluation after neoadjuvant chemotherapy in locally advanced gastric adenocarcinoma: a prospective, multi-center cohort study. *J Gastrointest Oncol*. 2017 Dec;8(6):1018-25.

Acton DS, Plat-Sinnige MJ, van Wamel W, de Groot N, van Belkum A. Intestinal carriage of *Staphylococcus aureus*: how does its frequency compare with that of nasal carriage and what is its clinical impact? *Eur J Clin Microbiol Infect Dis*. 2009;28(2):115-27.

Aguiar-Pulido V, Huang W, Suarez-Ulloa V, Cickovski T, Mathee K, Narasimhan G. Metagenomics, Metatranscriptomics, and Metabolomics Approaches for Microbiome Analysis. *Evol Bioinform Online*. 2016 May 12;12(Suppl 1):5-16.

Ajani JÁ, Lee J, Sano T, Janjigian YY, Fan D, Song S. Gastric adenocarcinoma. *Nat Rev Dis Primers*. 2017;3:17036.

Al-Batran SE, Homann N, Pauligk C, Goetze TO, Meiler J, Kasper S, et al. Perioperative chemotherapy with fluorouracil plus leucovorin, oxaliplatin, and docetaxel versus fluorouracil or capecitabine plus cisplatin and epirubicin for locally advanced, resectable gastric or gastro-oesophageal junction adenocarcinoma (FLOT4): a randomised, phase 2/3 trial. *Lancet*. 2019 May 11;393(10184):1948-57.

Allentoft ME, Collins M, Harker D, Haile J, Oskam CL, Hale ML, et al. The half-life of DNA in bone: measuring decay kinetics in 158 dated fossils. *Proc Biol Sci*. 2012 Dec 7;279(1748):4724-33.

Alsina M, Moehler M, Hierro C, Guardado R, Tabernero J. Immunotherapy for Gastric Cancer: A Focus on Immune Checkpoints. *Target Oncol*. 2016;11(4):469-77.

Amann RI, Ludwig W, Schleifer KH. Phylogenetic identification and in situ detection of individual microbial cells without cultivation. *Microbiol Rev*. 1995;59:143-69.

Amin MB. American Joint Committee on Cancer AJCC Cancer Staging Manual. 8th ed. New York: Springer; 2017. Stomch cancer; p.117-21.

Amir A, McDonald D, Navas-Molina JA, Kopylova E, Morton JT, Zech Xu Z, et al. Deblur Rapidly Resolves Single-Nucleotide Community Sequence Patterns. *mSystems*. 2017 Mar 7;2(2):e00191-16.

Ansari S, Gantuya B, Tuan VP, Yamaoka Y. Diffuse Gastric Cancer: A Summary of Analogous Contributing Factors for Its Molecular Pathogenicity. *Int J Mol Sci*. 2018 Aug 16;19(8):2424.

Arbique JC, Poyart C, Trieu-Cuot P, Quesne G, Carvalho Mda G, Steigerwalt AG, et al. Accuracy of phenotypic and genotypic testing for identification of *Streptococcus pneumoniae* and description of *Streptococcus pseudopneumoniae* sp. nov. *J Clin Microbiol*. 2004 Oct;42(10):4686-96.

Arnesen LPS, Fagerlund A, Granum PE. From soil to gut: *Bacillus cereus* and its food poisoning toxins. *FEMS Microbiol Rev*. 2008 Jul;32(4):579-606.

Arroyo Mühr LS, Lagheden C, Lei J, Eklund C, Nordqvist Kleppe S, Sparén P, et al. Deep sequencing detects human papillomavirus (HPV) in cervical cancers negative for HPV by PCR. *Br J Cancer*. 2020 Dec;123(12):1790-5.

Assumpção PP, Barra WF, Ishak G, Coelho LGV, Coimbra FJF, Freitas HC, et al. The diffuse-type gastric cancer epidemiology enigma. *BMC Gastroenterol*. 2020 Jul 13;20(1):223.

Aykut B, Pushalkar S, Chen R, Li Q, Abengozar R, Kim JI, et al. The fungal mycobiome promotes pancreatic oncogenesis via activation of MBL. *Nature*. 2019 Oct;574(7777):264-7.

Bai C, Li D, Zhang Q, Zheng S, Zheng L, Wang H, et al. Prognostic analysis of cancer patients with staphylococcus aureus infection: five-year experience at a comprehensive cancer center. *Int J Clin Exp Med* 2018;11(8):8640-5.

Balasubramanian D, Harper L, Shopsin B, Torres VJ. *Staphylococcus aureus* pathogenesis in diverse host environments. *Pathog Dis*. 2017;75(1):ftx005.

Bang YJ, Van Cutsem E, Feyereislova A, Chung HC, Shen L, Sawaki A, et al. Trastuzumab in combination with chemotherapy versus chemotherapy alone for treatment of HER2-positive advanced gastric or gastro-oesophageal junction cancer (ToGA): a phase 3, open-label, randomised controlled trial. *Lancet*. 2010 Aug 28;376(9742):687-97. Erratum in: *Lancet*. 2010 Oct 16;376(9749):1302.

Barko PC, McMichael MA, Swanson KS, Williams DA. The Gastrointestinal Microbiome: A Review. *J Vet Intern Med*. 2018;32(1):9-25.

Barrak I, Stájer A, Gajdács M, Urbán E. Small. but smelly: the importance of *Solobacterium moorei* in halitosis and other human infections. *Heliyon*. 2020;6(10):e05371.

Bashiardes S, Zilberman-Schapira G, Elinav E. Use of Metatranscriptomics in Microbiome Research. *Bioinform Biol Insights*. 2016;10:19-25.

Bassis CM, Erb-Downward JR, Dickson RP, Freeman CM, Schmidt TM, Young VB, et al. Analysis of the upper respiratory tract microbiotas as the source of the lung and gastric microbiotas in healthy individuals. *mBio*. 2015 Mar 3;6(2):e00037.

Becker K, Mueller JD, Schulmacher C, Ott K, Fink U, Busch R, et al. Histomorphology and grading of regression in gastric carcinoma treated with neoadjuvant chemotherapy. *Cancer*. 2003 Oct 1;98(7):1521-30.

Bik EM, Eckburg PB, Gill SR, Nelson KE, Purdom EA, Francois F, et al. Molecular analysis of the bacterial microbiota in the human stomach. *Proc Natl Acad Sci U S A*. 2006 Jan 17;103(3):732-7.

Bikel S, Valdez-Lara A, Cornejo-Granados F, Rico K, Canizales-Quinteros S, Soberón X, et al. Combining metagenomics, metatranscriptomics and viromics to explore novel microbial interactions: towards a systems-level understanding of human microbiome. *Comput Struct Biotechnol J*. 2015 Jun 9;13:390-401.

Brar G, Shah MA. The role of pembrolizumab in the treatment of PD-L1 expressing gastric and gastroesophageal junction adenocarcinoma. *Therap Adv Gastroenterol*. 2019;12:1756284819869767.

Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries [published correction appears in CA Cancer J Clin. 2020 Jul;70(4):313]. CA Cancer J Clin. 2018;68(6):394-424.

Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics. 2014;30(15):2114-20.

Bolyen E, Rideout JR, Dillon MR, Bokulich NA, Abnet CC, Al-Ghalith GA, et al. Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. Nat Biotechnol. 2019;37:852-7.

Bonequi P, Meneses-González F, Correa P, Rabkin CS, Camargo MC. Risk factors for gastric cancer in Latin America: a meta-analysis. Cancer Causes Control. 2013;24(2):217-31.

Brito BB, da Silva FAF, Soares AS, Pereira VA, Santos MLC, Sampaio MM, et al. Pathogenesis and clinical management of *Helicobacter pylori* gastric infection. World J Gastroenterol. 2019 Oct 7;25(37):5578-89.

Buchfink B, Xie C, Huson DH. Fast and sensitive protein alignment using DIAMOND. Nat Methods. 2015 Jan;12(1):59-60.

Bullman S, Pedamallu CS, Sicinska E, Clancy TE, Zhang X, Cai D, et al. Analysis of *Fusobacterium* persistence and antibiotic response in colorectal cancer. Science. 2017 Dec 15;358(6369):1443-8.

Burkitt MD, Duckworth CA, Williams JM, Pritchard DM. *Helicobacter pylori*-induced gastric pathology: insights from in vivo and ex vivo models. Dis Model Mech. 2017;10(2):89-104.

Campylobacter pylori becomes *Helicobacter pylori*. Lancet. 1989 Oct 28;2(8670):1019-20.

Cancer Genome Atlas Research Network. Comprehensive molecular characterization of gastric adenocarcinoma. Nature. 2014;513(7517):202-209.

Carcas LP. Gastric cancer review. J Carcinog. 2014;13:14.

Castaño-Rodríguez N, Goh KL, Fock KM, Mitchell HM, Kaakoush NO. Dysbiosis of the microbiome in gastric carcinogenesis. Sci Rep. 2017;7(1):15957.

Cervantes A, Roda D, Tarazona N, Roselló S, Pérez-Fidalgo JA. Current questions for the treatment of advanced gastric cancer. *Cancer Treat Rev*. 2013;39(1):60-7.

Chen C, Hemme C, Beleno J, Shi ZJ, Ning D, Qin Y, et al. Oral microbiota of periodontal health and disease and their changes after nonsurgical periodontal therapy. *ISME J*. 2018 May;12(5):1210-24.

Chen X, Winckler B, Lu M, Cheng H, Yuan Z, Yang Y, et al. Oral Microbiota and Risk for Esophageal Squamous Cell Carcinoma in a High-Risk Area of China. *PLoS One*. 2015 Dec 7;10(12):e0143603.

Chen Y, Chen Z, Guo R, Chen N, Lu H, Huang S, et al. Correlation between gastrointestinal fungi and varying degrees of chronic hepatitis B virus infection. *Diagn Microbiol Infect Dis*. 2011 Aug;70(4):492-8.

Chou SH, Tung YC, Wu LS, Chang CJ, Kung S, Chu PH. Severity of chronic periodontitis and risk of gastrointestinal cancers: A population-based follow-up study from Taiwan. *Medicine (Baltimore)*. 2018;97(27):e11386.

Cisło M, Filip AA, Arnold Offerhaus GJ, Ciseł B, Rawicz-Pruszyński K, Skierucha M, Polkowski WP. Distinct molecular subtypes of gastric cancer: from Laurén to molecular pathology. *Oncotarget*. 2018 Apr 10;9(27):19427-42.

Coker OO, Dai Z, Nie Y, Zhao G, Cao L, Nakatsu G, et al. Mucosal microbiome dysbiosis in gastric carcinogenesis. *Gut*. 2018 Jun;67(6):1024-32.

Cole ST, Eiglmeier K, Parkhill J, James KD, Thomson NR, Wheeler PR, et al. Massive gene decay in the leprosy bacillus. *Nature*. 2001 Feb 22;409(6823):1007-11.

Colombo AL, Padovan AC, Chaves GM. Current knowledge of *Trichosporon* spp. and Trichosporonosis. *Clin Microbiol Rev*. 2011;24(4):682-700.

Cong J, Zhang X. How human microbiome talks to health and disease. *Eur J Clin Microbiol Infect Dis*. 2018;37(9):1595-601.

Correa P, Haenszel W, Cuello C, Tannenbaum S, Archer M. A model for gastric cancer epidemiology. *Lancet*. 1975;2(7924):58-60. 5

Cristescu R, Lee J, Nebozhyn M, Kim KM, Ting JC, Wong SS, et al. Molecular analysis of gastric cancer identifies subtypes associated with distinct clinical outcomes. *Nat Med*. 2015 May;21(5):449-56.

Christoff AP, Sereia AFR, Boberg DR, Moraes RLV, Oliveira LFV. Bacterial identification through accurate library preparation and highthroughput sequencing. *Neoprospecta Microbiome Technol*. 2017:1–5.

Crowe SE. *Helicobacter pylori* Infection. *N Engl J Med*. 2019;380(12):1158-65.

Cui L, Morris A, Ghedin E. The human mycobiome in health and disease. *Genome Med*. 2013;5(7):63.

Culligan EP, Sleator RD, Marchesi JR, Hill C. Metagenomics and novel gene discovery: promise and potential for novel therapeutics. *Virulence*. 2014;5(3):399-412.

Cunningham D, Allum WH, Stenning SP, Thompson JN, Van de Velde CJ, Nicolson M, et al. Perioperative chemotherapy versus surgery alone for resectable gastroesophageal cancer. *N Engl J Med*. 2006 Jul 6;355(1):11-20.

DeGruttola AK, Low D, Mizoguchi A, Mizoguchi E. Current Understanding of Dysbiosis in Disease in Human and Animal Models. *Inflamm Bowel Dis*. 2016;22(5):1137-50.

Delgado S, Cabrera-Rubio R, Mira A, Suárez A, Mayo B. Microbiological survey of the human gastric ecosystem using culturing and pyrosequencing methods. *Microb Ecol*. 2013;65(3):763-72.

Derks S, Liao X, Chiaravalli AM, Xu X, Camargo MC, Solcia E, et al. Abundant PD-L1 expression in Epstein-Barr Virus-infected gastric cancers. *Oncotarget*. 2016 May 31;7(22):32925-32.

Diakite A, Dubourg G, Dione N, Afouda P, Bellali S, Ngom II, et al. Optimization and standardization of the culturomics technique for human microbiome exploration. *Sci Rep*. 2020 Jun 15;10(1):9674.

Dicken BJ, Bigam DL, Cass C, Mackey JR, Joy AA, Hamilton SM. Gastric adenocarcinoma: review and considerations for future directions. *Ann Surg*. 2005;241(1):27-39.

Dicksved J, Lindberg M, Rosenquist M, Enroth H, Jansson JK, Engstrand L. Molecular characterization of the stomach microbiota in patients with gastric cancer and in controls. *J Med Microbiol*. 2009;58(Pt 4):509-16.

Dolcetti R, De Re V, Canzonieri V. Immunotherapy for Gastric Cancer: Time for a Personalized Approach?. *Int J Mol Sci*. 2018;19(6):1602.

Downes J, Wade WG. *Peptostreptococcus stomatis* sp. nov.. isolated from the human oral cavity. *Int J Syst Evol Microbiol*. 2006;56(Pt 4):751-4.

Duport C, Jobin M, Schmitt P. Adaptation in *Bacillus cereus*: From Stress to Disease. *Front Microbiol*. 2016;7:1550.

Dzutsev A, Goldszmid RS, Viaud S, Zitvogel L, Trinchieri G. The role of the microbiota in inflammation, carcinogenesis, and cancer therapy. *Eur J Immunol*. 2015;45(1):17-31.

Emerson JB, Adams RI, Román CMB, Brooks B, Coil DA, Dahlhausen K, et al. Schrödinger's microbes: Tools for distinguishing the living from the dead in microbial ecosystems. *Microbiome*. 2017 Aug 16;5(1):86.

Eisenhofer R, Minich JJ, Marotz C, Cooper A, Knight R, Weyrich LS. Contamination in Low Microbial Biomass Microbiome Studies: Issues and Recommendations. *Trends Microbiol*. 2019;27(2):105-17.

Fang X, Wei J, He X, An P, Wang H, Jiang L, et al. Landscape of dietary factors associated with risk of gastric cancer: A systematic review and dose-response meta-analysis of prospective cohort studies. *Eur J Cancer*. 2015 Dec;51(18):2820-32.

Ferreira RM, Pereira-Marques J, Pinto-Ribeiro I, Costa JL, Carneiro F, Machado JC, Figueiredo C. Gastric microbial community profiling reveals a dysbiotic cancer-associated microbiota. *Gut*. 2018 Feb;67(2):226-36.

Franzosa EA, Morgan XC, Segata N, Waldron L, Reyes J, Earl AM, et al. Relating the metatranscriptome and metagenome of the human gut. *Proc Natl Acad Sci U S A*. 2014 Jun 3;111(22):E2329-38.

Gao Z, Guo B, Gao R, Zhu Q, Qin H. Microbiota disbiosis is associated with colorectal cancer. *Front Microbiol*. 2015;6:20.

Freitas HC. Avaliação do perfil de mutações somáticas em uma coorte de pacientes com adenocarcinoma gástrico. [Doutorado] São Paulo; Fundação Antônio Prudente; 2019.

Fuchs CS, Doi T, Jang RW, Muro K, Satoh T, Machado M, et al. Safety and Efficacy of Pembrolizumab Monotherapy in Patients With Previously Treated Advanced Gastric and Gastroesophageal Junction Cancer: Phase 2 Clinical KEYNOTE-059 Trial. *JAMA Oncol.* 2018 May 10;4(5):e180013. Erratum in: *JAMA Oncol.* 2019 Apr 1;5(4):579.

Geller LT, Barzily-Rokni M, Danino T, Jonas OH, Shental N, Nejman D, et al. Potential role of intratumor bacteria in mediating tumor resistance to the chemotherapeutic drug gemcitabine. *Science.* 2017 Sep 15;357(6356):1156-60.

Gholizadeh P, Eslami H, Yousefi M, Asgharzadeh M, Aghazadeh M, Kafil HS. Role of oral microbiome on oral cancers. a review. *Biomed Pharmacother.* 2016;84:552-8.

Glassing A, Dowd SE, Galandiuk S, Davis B, Chiodini RJ. Inherent bacterial DNA contamination of extraction and sequencing reagents may affect interpretation of microbiota in low bacterial biomass samples. *Gut Pathog.* 2016;8:24.

González CA, Pera G, Agudo A, Palli D, Krogh V, Vineis P, et al. Smoking and the risk of gastric cancer in the European Prospective Investigation Into Cancer and Nutrition (EPIC). *Int J Cancer.* 2003 Nov 20;107(4):629-34.

Goodman B, Gardner H. The microbiome and cancer. *J Pathol.* 2018;244(5):667-76.

Gouba N, Raoult D, Drancourt M. Eukaryote culturomics of the gut reveals new species. *PLoS One.* 2014;9(9):e106994.

Guimarães TF, Novis CF, Bottino CB, D'Acri AM, Lima RB, Martins CJ. Gastric syphilis - Case report. *An Bras Dermatol.* 2016;91(5):670-2.

Guo F, Ju F, Cai L, Zhang T. Taxonomic precision of different hypervariable regions of 16S rRNA gene and annotation methods for functional bacterial groups in biological wastewater treatment. *PLoS One.* 2013;8(10):e76185.

Guo Y, Zhang Y, Gerhard M, Gao JJ, Mejias-Luque R, Zhang L, et al. Effect of *Helicobacter pylori* on gastrointestinal microbiota: a population-based study in Linqu, a high-risk area of gastric cancer. *Gut.* 2020 Sep;69(9):1598-607.

Hallen-Adams HE, Suhr MJ. Fungi in the healthy human gastrointestinal tract. *Virulence*. 2017;8(3):352-8.

Haraszthy VI, Zambon JJ, Sreenivasan PK, Zambon MM, Gerber D, Rego R, Parker C. Identification of oral bacterial species associated with halitosis. *J Am Dent Assoc*. 2007 Aug;138(8):1113-20.

Hatakeyama M. *Helicobacter pylori* and gastric carcinogenesis. *J Gastroenterol*. 2009;44(4):239-48.

Helmink BA, Khan MAW, Hermann A, Gopalakrishnan V, Wargo JA. The microbiome, cancer, and cancer therapy. *Nat Med*. 2019;25(3):377-88.

Heinsen FA, Knecht H, Neulinger SC, Schmitz RA, Knecht C, Kühbacher T, et al. Dynamic changes of the luminal and mucosa-associated gut microbiota during and after antibiotic therapy with paromomycin. *Gut Microbes*. 2015 Jul 4;6(4):243-54.

Henson DE, Dittus C, Younes M, Nguyen H, Albores-Saavedra J. Differential trends in the intestinal and diffuse types of gastric carcinoma in the United States, 1973-2000: increase in the signet ring cell type. *Arch Pathol Lab Med*. 2004 Jul;128(7):765-70.

Herath M, Hosie S, Bornstein JC, Franks AE, Hill-Yardin EL. The Role of the Gastrointestinal Mucus System in Intestinal Homeostasis: Implications for Neurological Disorders. *Front Cell Infect Microbiol*. 2020;10:248.

Hoffmann C, Dollive S, Grunberg S, Chen J, Li H, Wu GD, Lewis JD, Bushman FD. Archaea and fungi of the human gut microbiome: correlations with diet and bacterial residents. *PLoS One*. 2013 Jun 17;8(6):e66019.

Hsiao EY, McBride SW, Hsien S, Sharon G, Hyde ER, McCue T, et al. Microbiota modulate behavioral and physiological abnormalities associated with neurodevelopmental disorders. *Cell*. 2013 Dec 19;155(7):1451-63.

Hu B, El Hajj N, Sittler S, Lammert N, Barnes R, Meloni-Ehrig A. Gastric cancer: Classification, histology and application of molecular pathology. *J Gastrointest Oncol*. 2012;3(3):251-61.

Hu S, Jiang T, Zhang X, Zhou Y, Yi Z, Wang Y, et al. Elizabethkingia anophelis Isolated from Patients with Multiple Organ Dysfunction Syndrome and Lower Respiratory Tract Infection: Report of Two Cases and Literature Review. *Front Microbiol.* 2017 Mar 8;8:382.

Hu YL, Pang W, Huang Y, Zhang Y, Zhang CJ. The Gastric Microbiome Is Perturbed in Advanced Gastric Adenocarcinoma Identified Through Shotgun Metagenomics. *Front Cell Infect Microbiol.* 2018;8:433.

Human Microbiome Project Consortium. Structure, function and diversity of the healthy human microbiome. *Nature.* 2012;486(7402):207-14.

Ibarbalz FM, Pérez MV, Figuerola EL, Erijman L. The bias associated with amplicon sequencing does not affect the quantitative assessment of bacterial community dynamics. *PLoS One.* 2014;9(6):e99722.

Iida N, Dzutsev A, Stewart CA, Smith L, Bouladoux N, Weingarten RA, et al. Commensal bacteria control cancer response to therapy by modulating the tumor microenvironment. *Science.* 2013 Nov 22;342(6161):967-70.

Ibrahim M, Subramanian A, Anishetty S. Comparative pan genome analysis of oral Prevotella species implicated in periodontitis. *Funct Integr Genomics.* 2017;17(5):513-36.

In H, Ravetch E, Langdon-Embry M, Palis B, Ajani JA, Hofstetter WL, Kelsen DP, Sano T. The newly proposed clinical and post-neoadjuvant treatment staging classifications for gastric adenocarcinoma for the American Joint Committee on Cancer (AJCC) staging. *Gastric Cancer.* 2018 Jan;21(1):1-9.

Janda JM, Abbott SL. 16s rRNA gene sequencing for bacterial identification in the diagnostic laboratory: pluses, perils and pitfalls. *J Clin Microbiol.* 2007;45:2761-4.

Janda JM, Lopez DL. Mini review: New pathogen profiles: Elizabethkingia anophelis. *Diagn Microbiol Infect Dis.* 2017;88(2):201-5.

Jautová J, Virágová S, Ondrasovic M, Holoda E. Incidence of Candida species isolated from human skin and nails: a survey. *Folia Microbiol (Praha).* 2001;46(4):333-7.

Job CK, Jayakumar J, Kearney M, Gillis TP. Transmission of leprosy: a study of skin and nasal secretions of household contacts of leprosy patients using PCR. *Am J Trop Med Hyg.* 2008;78(3):518-21.

Johnston FM, Beckman M. Updates on Management of Gastric Cancer. *Curr Oncol Rep*. 2019 Jun 24;21(8):67.

Kadioglu A, Weiser JN, Paton JC, Andrew PW. The role of *Streptococcus pneumoniae* virulence factors in host respiratory colonization and disease. *Nat Rev Microbiol*. 2008;6(4):288-301.

Kämpfer P, Matthews H, Glaeser SP, Martin K, Lodders N, Faye I. *Elizabethkingia anophelis* sp. nov.. isolated from the midgut of the mosquito *Anopheles gambiae* [published correction appears in *Int J Syst Evol Microbiol*. 2012 Apr;62(Pt 4):1016]. *Int J Syst Evol Microbiol*. 2011;61(Pt 11):2670-5.

Karimi P, Islami F, Anandasabapathy S, Freedman ND, Kamangar F. Gastric cancer: descriptive epidemiology. risk factors. screening. and prevention. *Cancer Epidemiol Biomarkers Prev*. 2014;23(5):700-13.

Kolonel LN, Hirohata T. Leprosy and cancer: a retrospective cohort study in Hawaii. *J Natl Cancer Inst*. 1977 Jun;58(6):1577-81.

Kopylova E, Noé L, Touzet H. SortMeRNA: fast and accurate filtering of ribosomal RNAs in metatranscriptomic data. *Bioinformatics*. 2012;28(24):3211-7.

Kostic AD, Chun E, Robertson L, Glickman JN, Gallini CA, Michaud M, et al. *Fusobacterium nucleatum* potentiates intestinal tumorigenesis and modulates the tumor-immune microenvironment. *Cell Host Microbe*. 2013 Aug 14;14(2):207-15.

Lagier JC, Dubourg G, Million M, Cadoret F, Bilen M, Fenollar F, et al. *Nat Rev Microbiol*. 2018 May 1;16:540-50.

Lande R. Statistics and Partitioning of Species Diversity, and Similarity among Multiple Communities. *Oikos*. 1996;76(1):5.

Langmead. B, Salzberg. S. Fast gapped-read alignment with Bowtie 2. *Nat Methods*. 2012;9:357-9.

Larkin J, Hodi FS, Wolchok JD. Combined Nivolumab and Ipilimumab or Monotherapy in Untreated Melanoma. *N Engl J Med*. 2015 Sep 24;373(13):1270-1. Erratum in: *N Engl J Med*. 2018 Nov 29;379(22):2185.

Lau SK, Chow WN, Foo CH, Curreem SO, Lo GC, Teng JL, et al. *Elizabethkingia anophelis* bacteremia is associated with clinically significant infections and high mortality. *Sci Rep*. 2016 May 17;6:26045.

Lauren P. The two histological main types of gastric carcinoma: Diffuse and so-called intestinal-type carcinoma. An attempt at a histo-clinical classification. *Acta Pathol Microbiol Scand*. 1965;64(1):31-49. 1965.

Lauwers GY, Carneiro F, Graham DY. Gastric carcinoma. In: Bowman FT, Carneiro F, Hruban RH. eds. *Classification of Tumours of the Digestive System*. Lyon: IARC; 2010

Leschner S, Weiss S. Salmonella-allies in the fight against cancer. *J Mol Med (Berl)*. 2010;88(8):763-73.

Lehouritis P, Cummins J, Stanton M, Murphy CT, McCarthy FO, Reid G, et al. Local bacteria affect the efficacy of chemotherapeutic drugs. *Sci Rep*. 2015 Sep 29;5:14554.

Leung PHM, Subramanya R, Mou Q, Lee KT, Islam F, Gopalan V, Lu CT, Lam AK. Characterization of Mucosa-Associated Microbiota in Matched Cancer and Non-neoplastic Mucosa From Patients With Colorectal Cancer. *Front Microbiol*. 2019 Jun 12;10:1317.

Li XX, Wong GL, To KF, Wong VW, Lai LH, Chow DK, et al. Bacterial microbiota profiling in gastritis without *Helicobacter pylori* infection or non-steroidal anti-inflammatory drug use. *PLoS One*. 2009 Nov 24;4(11):e7985.

Liao Y, Wu F, Dai F, Huang Q, Feng Y, Ling Y, Lu H. *Mycobacterium tuberculosis* and *Pseudoramibacter alactolyticus* coinfection in brain after dental extraction: A case report. *Medicine (Baltimore)*. 2019 Dec;98(50):e18289.

Liu J, Xue Y, Zhou L. Detection of gastritis-associated pathogens by culturing of gastric juice and mucosa. *Int J Clin Exp Pathol*. 2018;11(4):2214-20.

Liu Y, Zhang Y, Dong P, An R, Xue C, Ge Y, Wei L, Liang X. Digestion of Nucleic Acids Starts in the Stomach. *Sci Rep*. 2015 Jul 14;5:11936.

Lo CH, Kwon S, Wang L, Polychronidis G, Knudsen MD, Zhong R, et al. Periodontal disease, tooth loss, and risk of oesophageal and gastric adenocarcinoma: a prospective study. *Gut*. 2020 Jul 20;gutjnl-2020-321949.

Lorenzen S, Thuss-Patience P, Al-Batran SE, Lordick F, Haller B, Schuster T, et al. Impact of pathologic complete response on disease-free survival in patients with esophagogastric adenocarcinoma receiving preoperative docetaxel-based chemotherapy. *Ann Oncol*. 2013 Aug;24(8):2068-73.

Lozupone CA, Hamady M, Kelley ST, Knight R. Quantitative and qualitative β diversity measures lead to different insights into factors that structure microbial communities. *Appl Environ Microbiol*. 2007;73(5):1576-85.

Magee K, Hata DJ, Meza D. Trichosporon asahii Infection in a Patient with Metastatic Prostate Cancer as an Example of an Emerging Fungal Pathogen. *Lab Med*. 2015;46(3):e74-e8.

Ma K, Baloch Z, He TT, Xia X. Alcohol Consumption and Gastric Cancer Risk: A Meta-Analysis. *Med Sci Monit*. 2017;23:238-46.

Mansour JC, Tang L, Shah M, Bentrem D, Klimstra DS, Gonen M, et al. Does graded histologic response after neoadjuvant chemotherapy predict survival for completely resected gastric cancer? *Ann Surg Oncol*. 2007 Dec;14(12):3412-8.

Marchesi JR, Ravel J. The vocabulary of microbiome research: a proposal. *Microbiome*. 2015;3:31.

Marshall BJ, Warren JR. Unidentified curved bacilli in the stomach of patients with gastritis and peptic ulceration. *Lancet*. 1984 Jun 16;1(8390):1311-5.

Martel C, Ferlay J, Franceschi S, Vignat J, Bray F, Forman D, Plummer M. Global burden of cancers attributable to infections in 2008: a review and synthetic analysis. *Lancet Oncol*. 2012 Jun;13(6):607-15.

Martiny AC. High proportions of bacteria are culturable across major biomes. *ISME J*. 2019;13(8):2125-8.

McMurdie PJ, Holmes S. phyloseq: an R package for reproducible interactive analysis and graphics of microbiome census data. *PLoS One*. 2013;8(4):e61217.

Miao R, Wan C, Wang Z. The relationship of gastric microbiota and *Helicobacter pylori* infection in pediatrics population. *Helicobacter*. 2020;25(1):e12676.

Ministério da Saúde. Instituto Nacional de Câncer Jose Alencar Gomes da Silva. Estimativa/2020. Incidência de câncer no Brasil. Rio de Janeiro; 2020. Disponível em: <https://bit.ly/38hQrXu>. [2020 ago 19].

Moeller AH, Caro-Quintero A, Mjungu D, Georgiev AV, Lonsdorf EV, Muller MN, et al. Cospeciation of gut microbiota with hominids. *Science*. 2016 Jul 22;353(6297):380-2.

Moen AEF, Lindstrøm JC, Tannæs TM, Vatn S, Ricanek P, Vatn MH, Jahnsen J; IBD-Character Consortium. The prevalence and transcriptional activity of the mucosal microbiota of ulcerative colitis patients. *Sci Rep*. 2018 Nov 22;8(1):17278.

Morihiro T, Kuroda S, Kanaya N, Kakiuchi Y, Kubota T, Aoyama K, et al. PD-L1 expression combined with microsatellite instability/CD8+ tumor infiltrating lymphocytes as a useful prognostic biomarker in gastric cancer. *Sci Rep*. 2019 Mar 15;9(1):4633.

Motzer RJ, Escudier B, McDermott DF, George S, Hammers HJ, Srinivas S, et al. Nivolumab versus Everolimus in Advanced Renal-Cell Carcinoma. *N Engl J Med*. 2015 Nov 5;373(19):1803-13.

Mueller O., Lightfoot S, Schroeder A. RNA Integrity Number (RIN)—Standardization of RNA Quality Control. 2016. Agilent Technologies. Technical Report 5989-1165EN.

Mulcahy-O'Grady H, Workentine ML. The Challenge and Potential of Metagenomics in the Clinic. *Front Immunol*. 2016;7:29.

Nagini S. Carcinoma of the stomach: A review of epidemiology. pathogenesis. molecular genetics and chemoprevention. *World J Gastrointest Oncol*. 2012;4(7):156-69.

Nardone G, Compare D. The human gastric microbiota: Is it time to rethink the pathogenesis of stomach diseases? *United European Gastroenterol J*. 2015;3(3):255-60.

National Cancer Institute. The Surveillance, Epidemiology, and End Results-SEER. Cancer Statistics: SEER Cancer Stat Fact Sheets: Stomach Cancer. 2010-2016. Disponible em: <https://bit.ly/3nI0YFM>. [2020 ago 19].

National Center for Biotechnology-NCBI. Genome Browse by Organism. Disponible em: <https://www.ncbi.nlm.nih.gov/assembly/organism>. [2020 ago 19].

Nelson A, De Soyza A, Bourke SJ, Perry JD, Cummings SP. Assessment of sample handling practices on microbial activity in sputum samples from patients with cystic fibrosis. *Lett Appl Microbiol*. 2010;51(3):272-7.

Nishikawa J, Iizasa H, Yoshiyama H, Shimokuri K, Kobayashi Y, Sasaki S, et al. Clinical Importance of Epstein-Barr Virus-Associated Gastric Cancer. *Cancers (Basel)*. 2018 May 29;10(6):167.

Okamoto K, Hatakeyama S, Umezawa M, Hayashi S. Gastric syphilis: The great imitator in the stomach. *IDCases*. 2018;12:97-8.

Olsen I, Yamazaki K. Can oral bacteria affect the microbiome of the gut?. *J Oral Microbiol*. 2019;11(1):1586422.

Ott SJ, Kühbacher T, Musfeldt M, Rosenstiel P, Hellmig S, Rehman A, et al. Fungi and inflammatory bowel diseases: Alterations of composition and diversity. *Scand J Gastroenterol*. 2008;43(7):831-41.

Overmann J, Abt B, Sikorski J. Present and Future of Culturing Bacteria. *Annu Rev Microbiol*. 2017;71:711-30.

Park DJ, Thomas NJ, Yoon C, Yoon SS. Vascular endothelial growth factor a inhibition in gastric cancer. *Gastric Cancer*. 2015;18(1):33-42.

Peek RM Jr, Blaser MJ. *Helicobacter pylori* and gastrointestinal tract adenocarcinomas. *Nat Rev Cancer*. 2002;2(1):28-37.

Peng Z, Cheng S, Kou Y, Wang Z, Jin R, Hu H, et al. The Gut Microbiome Is Associated with Clinical Response to Anti-PD-1/PD-L1 Immunotherapy in Gastrointestinal Cancer. *Cancer Immunol Res*. 2020 Oct;8(10):1251-61.

Petrovchich I, Ford JM. Genetic predisposition to gastric cancer. *Semin Oncol*. 2016;43(5):554-9.

Pinto MP, Córdova-Delgado M, Retamal IN, Muñoz-Medel M, Bravo ML, Durán D, et al. A Molecular Stratification of Chilean Gastric Cancer Patients with Potential Clinical Applicability. *Cancers (Basel)*. 2020 Jul 10;12(7):1863.

Prichard W, Fick L. When Diarrhea Can Become Deadly: Legionnaires' Disease Complicated by Bowel Obstruction. *Case Rep Gastroenterol*. 2016;10(3):781-6.

Purtilo DT, Pangi C. Incidence of cancer in patients with leprosy. *Cancer*. 1975;35(4):1259-61.

Quast C, Pruesse E, Yilmaz P, Gerken J, Schweer T, Yarza P, Peplies J, Glöckner FO. The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic Acids Res*. 2013 Jan;41(Database issue):D590-6.

Rafat Z, Hashemi SJ, Ahamdikia K, Daie Ghazvini R, Bazvandi F. Study of skin and nail *Candida* species as a normal flora based on age groups in healthy persons in Tehran-Iran. *J Mycol Med*. 2017;27(4):501-5.

Rajilic-Stojanovic M, Figueiredo C, Smet A, Hansen R, Kupcinskas J, Rokkas T, et al. Systematic review: gastric microbiota in health and disease. *Aliment Pharmacol Ther*. 2020 Mar;51(6):582-602.

Reck M, Rodríguez-Abreu D, Robinson AG, Hui R, Csósz T, Fülöp A, et al. Pembrolizumab versus Chemotherapy for PD-L1-Positive Non-Small-Cell Lung Cancer. *N Engl J Med*. 2016 Nov 10;375(19):1823-33.

Ribeiro. HSC. Expressão proteica de PD-L1 em pacientes com adenocarcinoma gástrico antes e após tratamento neoadjuvante e correlação com desfechos clínicos. Tese (Doutorado em Oncologia) – Fundação Antonio Prudente. São Paulo. 2020.

Richman DM, Tirumani SH, Hornick JL, Fuchs CS, Howard S, Krajewski K, Ramaiya N, Rosenthal M. Beyond gastric adenocarcinoma: Multimodality assessment of common and uncommon gastric neoplasms. *Abdom Radiol (NY)*. 2017 Jan;42(1):124-40.

Ronnachit A, Ellenberger KA, Gray TJ, Liu EY, Gilbey TL, Cheong EY, Gottlieb T, McKew GL. Streptococcus Pneumoniae Causing Intra-abdominal and Pelvic Infection: A Case Series. Cureus. 2017 Dec 19;9(12):e1967.

Rosner J. Ten Times More Microbial Cells than Body Cells in Humans?. Microbe Magazine. 2014;9:47-7.

Roy S, Trinchieri G. Microbiota: a key orchestrator of cancer therapy. Nat Rev Cancer. 2017;17(5):271-85.

Salama NR, Hartung ML, Müller A. Life in the human stomach: persistence strategies of the bacterial pathogen Helicobacter pylori. Nat Rev Microbiol. 2013;11(6):385-99.

Salter SJ, Cox MJ, Turek EM, Calus ST, Cookson WO, Moffatt MF, et al. Reagent and laboratory contamination can critically impact sequence-based microbiome analyses. BMC Biol. 2014 Nov 12;12:87.

Sampson TR, Debelius JW, Thron T, Janssen S, Shastri GG, Ilhan ZE, et al. Gut Microbiota Regulate Motor Deficits and Neuroinflammation in a Model of Parkinson's Disease. Cell. 2016 Dec 1;167(6):1469-80.e12.

Savage DC. Microbial ecology of the gastrointestinal tract. Annu Rev Microbiol. 1977;31:107-33.

Schmieder R, Edwards R. Fast identification and removal of sequence contamination from genomic and metagenomic datasets. PLoS One. 2011;6(3):e17288.

Schulz C, Schütte K, Koch N, Vilchez-Vargas R, Wos-Oxley ML, Oxley APA, et al. The active bacterial assemblages of the upper GI tract in individuals with and without Helicobacter infection. Gut. 2018 Feb;67(2):216-25.

Scollard DM, Adams LB, Gillis TP, Krahenbuhl JL, Truman RW, Williams DL. The continuing challenges of leprosy. Clin Microbiol Rev. 2006;19(2):338-81.

Seck E, Rathored J, Khelaifia S, Croce O, Robert C, Couderc C, et al. *Virgibacillus senegalensis* sp. nov., a new moderately halophilic bacterium isolated from human gut. *New Microbes New Infect.* 2015 Oct 3;8:116-26.

Sender R, Fuchs S, Milo R. Revised Estimates for the Number of Human and Bacteria Cells in the Body. *PLoS Biol.* 2016;14(8):e1002533.

Sharpton TJ. An introduction to the analysis of shotgun metagenomic data. *Front Plant Sci.* 2014;5:209.

Shang FM, Liu HL. *Fusobacterium nucleatum* and colorectal cancer: A review. *World J Gastrointest Oncol.* 2018;10(3):71-81.

Shreiner AB, Kao JY, Young VB. The gut microbiome in health and in disease. *Curr Opin Gastroenterol.* 2015;31(1):69-75.

Silvestre AJM, Miranda MABR, Camargo ZP. *Trichosporon* species isolated from the perigenital region. urine and catheters of a brazilian population. *Braz J Microbiol.* 2010;41(3):628-34.

Siqueira JF Jr, Rôças IN. *Pseudoramibacter alactolyticus* in primary endodontic infections. *J Endod.* 2003;29(11):735-8.

Siqueira JF Jr, Rôças IN. Distinctive features of the microbiota associated with different forms of apical periodontitis. *J Oral Microbiol.* 2009;1:10.3402/jom.v1i0.2009.

Sivan A, Corrales L, Hubert N, Williams JB, Aquino-Michaels K, Earley ZM, et al. Commensal *Bifidobacterium* promotes antitumor immunity and facilitates anti-PD-L1 efficacy. *Science.* 2015 Nov 27;350(6264):1084-9.

Smyth EC, Verheij M, Allum W, Cunningham D, Cervantes A, Arnold D; ESMO Guidelines Committee. Gastric cancer: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up. *Ann Oncol.* 2016 Sep;27(suppl 5):v38-v49.

Sohn SH, Kim N, Jo HJ, Kim J, Park JH, Nam RH, et al. Analysis of Gastric Body Microbiota by Pyrosequencing: Possible Role of Bacteria Other Than *Helicobacter pylori* in the Gastric Carcinogenesis. *J Cancer Prev.* 2017 Jun;22(2):115-25. Epub 2017 Jun 30. Erratum in: *J Cancer Prev.* 2017 Dec;22(4):267.

Song S, Vuai MS, Zhong M. The role of bacteria in cancer therapy - enemies in the past. but allies at present. *Infect Agent Cancer*. 2018;13:9.

Staley JT, Konopka A. Measurement of in situ activities of non-photosynthetic microorganisms in aquatic and terrestrial habitats. *Annu Rev Microbiol*. 1985;39:321-46.

Stewart EJ. Growing unculturable bacteria. *J Bacteriol*. 2012;194(16):4151-60.

Suau A, Bonnet R, Sutren M, Godon JJ, Gibson GR, Collins MD, Doré J. Direct analysis of genes encoding 16S rRNA from complex communities reveals many novel molecular species within the human gut. *Appl Environ Microbiol*. 1999 Nov;65(11):4799-807.

Sun CH, Li BB, Wang B, Zhao J, Zhang XY, Li TT, et al. The role of *Fusobacterium nucleatum* in colorectal cancer: from carcinogenesis to clinical management. *Chronic Dis Transl Med*. 2019 Oct 1;5(3):178-87.

Sun JH, Li XL, Yin J, Li YH, Hou BX, Zhang Z. A screening method for gastric cancer by oral microbiome detection. *Oncol Rep*. 2018 May;39(5):2217-24.

Sun M, Chen C, Xiao W, Chang Y, Liu C, Xu Q. Increase in *Candida Parapsilosis* Candidemia in Cancer Patients. *Mediterr J Hematol Infect Dis*. 2019;11(1):e2019012.

Takahashi S, Tomita J, Nishioka K, Hisada T, Nishijima M. Development of a prokaryotic universal primer for simultaneous analysis of Bacteria and Archaea using next-generation sequencing. *PLoS One*. 2014;9(8):e105592.

Takeuchi K, Tsuzuki Y, Ando T, Sekihara M, Hara T, Yoshikawa M, Kudo M, Kuwano H. Clinical studies of enteritis caused by methicillin-resistant *Staphylococcus aureus*. *Eur J Surg*. 2001 Apr;167(4):293-6.

Tanaka S, Yoshida M, Murakami Y, Ogiwara T, Shoji M, Kobayashi S, et al. The relationship of *Prevotella intermedia*, *Prevotella nigrescens* and *Prevotella melaninogenica* in the supragingival plaque of children, caries and oral malodor. *J Clin Pediatr Dent*. 2008 Spring;32(3):195-200.

Tani H, Akimitsu N. Genome-wide technology for determining RNA stability in mammalian cells: historical perspective and recent advantages based on modified nucleotide labeling. *RNA Biol*. 2012;9(10):1233-8.

Tatusova T, Ciufo S, Fedorov B, O'Neill K, Tolstoy I. RefSeq microbial genomes database: new representation and annotation strategy. *Nucleic Acids Res.* 2014;42(Database issue):D553–9.

Thomas AM, Jesus EC, Lopes A, Aguiar S Jr, Begnami MD, Rocha RM, et al. Tissue-Associated Bacterial Alterations in Rectal Carcinoma Patients Revealed by 16S rRNA Community Profiling. *Front Cell Infect Microbiol.* 2016 Dec 9;6:179.

Thomas AM, Manghi P, Asnicar F, Pasolli E, Armanini F, Zolfo M, et al. Metagenomic analysis of colorectal cancer datasets identifies cross-cohort microbial diagnostic signatures and a link with choline degradation. *Nat Med.* 2019 Apr;25(4):667-678. Epub 2019 Apr 1. Erratum in: *Nat Med.* 2019 Dec;25(12):1948.

Thorell K, Bengtsson-Palme J, Liu OH, Palacios Gonzales RV, Nookaew I, Rabeneck L, et al. In Vivo Analysis of the Viable Microbiota and *Helicobacter pylori* Transcriptome in Gastric Infection and Early Stages of Carcinogenesis. *Infect Immun.* 2017 Sep 20;85(10):e00031-17.

Tong SY, Davis JS, Eichenberger E, Holland TL, Fowler VG Jr. *Staphylococcus aureus* infections: epidemiology. pathophysiology. clinical manifestations. and management. *Clin Microbiol Rev.* 2015;28(3):603-61.

Torsvik V, Øvreås L. Microbial diversity and function in soil:from genes to ecosystems. *Curr Opin Microbiol.* 2002;5:240-5.

Treiber ML, Taft DH, Korf I, Mills DA, Lemay DG. Pre- and post-sequencing recommendations for functional annotation of human fecal metagenomes. *BMC Bioinformatics.* 2020;21(1):74.

Troncoso C, Pavez M, Cerda A, Oporto M, Villarroel D, Hofmann E, et al. MALDI-TOF MS and 16S RNA Identification of Culturable Gastric Microbiota: Variability Associated with the Presence of *Helicobacter pylori*. *Microorganisms.* 2020 Nov 10;8(11):1763.

Tsuda A, Suda W, Morita H, Takanashi K, Takagi A, Koga Y, Hattori M. Influence of Proton-Pump Inhibitors on the Luminal Microbiota in the Gastrointestinal Tract. *Clin Transl Gastroenterol.* 2015 Jun 11;6(6):e89.

Valles-Colomer M, Falony G, Darzi Y, Tigchelaar EF, Wang J, Tito RY, et al. The neuroactive potential of the human gut microbiota in quality of life and depression. *Nat Microbiol*. 2019 Apr;4(4):623-32.

Valles-Colomer M, Falony G, Darzi Y, Tigchelaar EF, Wang J, Tito RY, et al. The neuroactive potential of the human gut microbiota in quality of life and depression. *Nat Microbiol*. 2019 Apr;4(4):623-32.

Orditura M, Galizia G, Sforza V, Gambardella V, Fabozzi A, Laterza MM, et al. Treatment of gastric cancer. *World J Gastroenterol*. 2014 Feb 21;20(7):1635-49.

Vartoukian SR, Palmer RM, Wade WG. Strategies for culture of 'unculturable' bacteria. *FEMS Microbiol Lett*. 2010;309(1):1-7.

Vétizou M, Pitt JM, Daillère R, Lepage P, Waldschmitt N, Flament C, et al. Anticancer immunotherapy by CTLA-4 blockade relies on the gut microbiota. *Science*. 2015 Nov 27;350(6264):1079-84.

Von Rosenvinge EC, Song Y, White JR, Maddox C, Blanchard T, Fricke WF. Immune status. antibiotic medication and pH are associated with changes in the stomach fluid microbiota. *ISME J*. 2013;7(7):1354-66.

Waldum HL, Fossmark R. Types of Gastric Carcinomas. *Int J Mol Sci*. 2018;19(12):4109.

Wang L, Zhou J, Xin Y, Geng C, Tian Z, Yu X, Dong Q. Bacterial overgrowth and diversification of microbiota in gastric cancer. *Eur J Gastroenterol Hepatol*. 2016 Mar;28(3):261-6.

Wang Q, Liu G, Hu C. Molecular Classification of Gastric Adenocarcinoma. *Gastroenterology Res*. 2019;12(6):275-282.

Wang Z, Liu L, Ji J, Zhang J, Yan M, Zhang J, et al. ABO blood group system and gastric cancer: a case-control study and meta-analysis. *Int J Mol Sci*. 2012 Oct 17;13(10):13308-21.

Wang ZK, Yang YS. Upper gastrointestinal microbiota and digestive diseases. *World J Gastroenterol*. 2013;19(10):1541-1550.

Weiser JN, Ferreira DM, Paton JC. *Streptococcus pneumoniae*: transmission. colonization and invasion. *Nat Rev Microbiol*. 2018;16(6):355-67.

Westreich ST, Treiber ML, Mills DA, Korf I, Lemay DG. SAMSA2: a standalone metatranscriptome analysis pipeline. *BMC Bioinformatics*. 2018;19(1):175.

Wei MQ, Ellem KA, Dunn P, West MJ, Bai CX, Vogelstein B. Facultative or obligate anaerobic bacteria have the potential for multimodality therapy of solid tumours. *Eur J Cancer*. 2007;43(3):490-6.

Wilke H, Muro K, Van Cutsem E, Oh SC, Bodoky G, Shimada Y, et al. Ramucirumab plus paclitaxel versus placebo plus paclitaxel in patients with previously treated advanced gastric or gastro-oesophageal junction adenocarcinoma (RAINBOW): a double-blind, randomised phase 3 trial. *Lancet Oncol*. 2014 Oct;15(11):1224-35.

White C, Franco-Paredes C. Leprosy in the 21st century. *Clin Microbiol Rev*. 2015;28(1):80-94.

Wood DE, Lu J, Langmead B. Improved metagenomic analysis with Kraken 2. *Genome Biol*. 2019;20(1):257.

Wong SH, Zhao L, Zhang X, Nakatsu G, Han J, Xu W, et al. Gavage of Fecal Samples From Patients With Colorectal Cancer Promotes Intestinal Carcinogenesis in Germ-Free and Conventional Mice. *Gastroenterology*. 2017 Dec;153(6):1621-33.e6.

World Health Organization-WHO. Fact sheets. 2020. Disponível em: <https://bit.ly/3pXQE8y>. [2020 ago 19].

World Health Organization. 2018. Global leprosy update. 2018: moving towards a leprosy-free world. *Wkly Epidemiol Rec* 94:389–412.

Wroblewski LE, Peek RM Jr, Wilson KT. *Helicobacter pylori* and gastric cancer: factors that modulate disease risk. *Clin Microbiol Rev*. 2010;23(4):713-739.

Wurm P, Dörner E, Kremer C, Spranger J, Maddox C, Halwachs B, et al. Qualitative and Quantitative DNA-and RNA-Based Analysis of the Bacterial Stomach Microbiota in Humans, Mice, and Gerbils. *mSystems*. 2018 Nov 20;3(6):e00262-18.

Yang I, Nell S, Suerbaum S. Survival in hostile territory: the microbiota of the stomach. *FEMS Microbiol Rev*. 2013;37(5):736-61.

Yang I, Woltemate S, Piazuolo MB, Bravo LE, Yepez MC, Romero-Gallo J, et al. Different gastric microbiota compositions in two human populations with high and low gastric cancer risk in Colombia. *Sci Rep*. 2016 Jan 5;6:18594.

Yu G, Torres J, Hu N, Medrano-Guzman R, Herrera-Goepfert R, Humphrys MS, et al. Molecular Characterization of the Human Stomach Microbiota in Gastric Cancer Patients. *Front Cell Infect Microbiol*. 2017 Jul 6;7:302.

Yu J, Feng Q, Wong SH, Zhang D, Liang QY, Qin Y, et al. Metagenomic analysis of faecal microbiome as a tool towards targeted non-invasive biomarkers for colorectal cancer. *Gut*. 2017 Jan;66(1):70-8.

Yusefi AR, Lankarani KB, Bastani P, Radinmanesh M, Kavosi Z. Risk Factors for Gastric Cancer: A Systematic Review. *Asian Pac J Cancer Prev*. 2018 Mar 27;19(3):591-603.

You M, Mo S, Leung WK, Watt RM. Comparative analysis of oral treponemes associated with periodontal health and disease. *BMC Infect Dis*. 2013;13:174.

Zaura E, Keijser BJ, Huse SM, Crielaard W. Defining the healthy "core microbiome" of oral microbial communities. *BMC Microbiol*. 2009;9:259.

Zeller G, Tap J, Voigt AY, Sunagawa S, Kultima JR, Costea PI, et al. Potential of fecal microbiota for early-stage detection of colorectal cancer. *Mol Syst Biol*. 2014 Nov 28;10(11):766.

Zeng ZM, Luo FF, Zou LX, He RQ, Pan DH, Chen X, et al. Human papillomavirus as a potential risk factor for gastric cancer: a meta-analysis of 1,917 cases. *Onco Targets Ther*. 2016 Nov 17;9:7105-14.

Zhang C, Cleveland K, Schnoll-Sussman F, McClure B, Bigg M, Thakkar P, et al. Identification of low abundance microbiome in clinical samples using whole genome sequencing. *Genome Biol*. 2015 Nov 27;16:265.

Zhang Y, Zhen M, Zhan Y, Song Y, Zhang Q, Wang J. Population-Genomic Insights into Variation in *Prevotella intermedia* and *Prevotella nigrescens* Isolates and Its Association with Periodontal Disease. *Front Cell Infect Microbiol*. 2017;7:409.

Zhao H, Chu M, Huang Z, Yang X, Ran S, Hu B, Zhang C, Liang J. Variations in oral microbiota associated with oral cancer. *Sci Rep*. 2017 Sep 18;7(1):11773.

Zhu K, Hölzel CS, Cui Y, Mayer R, Wang Y, Dietrich R, et al. Probiotic *Bacillus cereus* Strains, a Potential Risk for Public Health in China. *Front Microbiol*. 2016 May 23;7:718.

Zhao Q, Cao L, Guan L, Bie L, Wang S, Xie B, et al. Immunotherapy for gastric cancer: dilemmas and prospect. *Brief Funct Genomics*. 2019 Mar 22;18(2):107-12.

Zilberstein B, Quintanilha AG, Santos MA, Pajecki D, Moura EG, Alves PR, et al. Digestive tract microbiota in healthy volunteers. *Clinics (Sao Paulo)*. 2007 Feb;62(1):47-54.

Zwolińska-Wcisło M, Budak A, Trojanowska D, Bogdał J, Stachura J. Fungal colonization of the stomach and its clinical relevance. *Mycoses*. 1998;41(7-8):327-34.

Anexo 1 - Carta de aprovação do Comitê de Ética em Pesquisa do projeto temático gástrico



FUNDAÇÃO ANTÔNIO
PRUDENTE-HOSPITAL DO
CÂNCER A C CAMARGO



PARECER CONSUBSTANCIADO DO CEP

DADOS DO PROJETO DE PESQUISA

Título da Pesquisa: Genômica de Adenocarcinomas Gástricos

Pesquisador: Emmanuel Dias-Neto

Área Temática:

Versão: 2

CAAE: 50523015.8.0000.5432

Instituição Proponente: Fundação Antônio Prudente-Hospital do Câncer-A C Camargo

Patrocinador Principal: FUNDAÇÃO DE AMPARO A PESQUISA DO ESTADO DE SÃO PAULO

DADOS DO PARECER

Número do Parecer: 1.373.212

Apresentação do Projeto:

Conforme parecer anterior

Objetivo da Pesquisa:

Conforme parecer anterior

Avaliação dos Riscos e Benefícios:

Conforme parecer anterior

Comentários e Considerações sobre a Pesquisa:

Conforme parecer anterior

Considerações sobre os Termos de apresentação obrigatória:

As considerações feitas no parecer anterior foram acatadas.

Conclusões ou Pendências e Lista de Inadequações:

Foi acrescentado no Termo de Consentimento Livre e Esclarecido a perda de confidencialidade.

Considerações Finais a critério do CEP:

Nota: Informações a respeito do andamento do referido projeto deverão ser encaminhadas ao CEP dentro de 08 meses a partir desta data em relatório (modelo CEP).

Endereço: Rua Professor Antônio Prudente, 211

Bairro: Liberdade

CEP: 01.509-900

UF: SP

Município: SÃO PAULO

Telefone: (11)2189-5020

Fax: (11)2189-5020

E-mail: cep_hcancer@accamargo.org.br

Continuação do Parecer: 1.373.212

Este parecer foi elaborado baseado nos documentos abaixo relacionados:

Tipo Documento	Arquivo	Postagem	Autor	Situação
Informações Básicas do Projeto	PB_INFORMAÇÕES_BÁSICAS_DO_PROJETO_616050.pdf	03/12/2015 16:38:44		Aceito
Outros	reply_CEP_gastric_031215.pdf	03/12/2015 16:38:06	Emmanuel Dias-Neto	Aceito
TCLE / Termos de Assentimento / Justificativa de Ausência	TCLE_tematico_estomago_031215.doc	03/12/2015 16:35:17	Emmanuel Dias-Neto	Aceito
Declaração de Pesquisadores	onco_clinica_ciencia.PDF	28/10/2015 14:11:52	Emmanuel Dias-Neto	Aceito
Declaração de Pesquisadores	declaracao_infraestrutura_onco_clinica.PDF	28/10/2015 14:10:10	Emmanuel Dias-Neto	Aceito
Folha de Rosto	folha_rosto_dirigente.pdf	27/10/2015 15:46:59	Emmanuel Dias-Neto	Aceito
Declaração de Pesquisadores	declaracao_cirurgia_abdominal.pdf	27/10/2015 15:39:18	Emmanuel Dias-Neto	Aceito
Orçamento	orcamento_CEP_151027.pdf	27/10/2015 13:57:54	Emmanuel Dias-Neto	Aceito
Declaração de Pesquisadores	declaracao_comprometimento_CIPE.pdf	27/10/2015 10:55:14	Emmanuel Dias-Neto	Aceito
Outros	formulario_submissao_CEP_151026.pdf	26/10/2015 19:13:50	Emmanuel Dias-Neto	Aceito
Cronograma	cronograma_CEP_151026.pdf	26/10/2015 19:02:43	Emmanuel Dias-Neto	Aceito
TCLE / Termos de Assentimento / Justificativa de Ausência	TCLE_tematico_estomago_221015_VF.doc	26/10/2015 18:45:04	Emmanuel Dias-Neto	Aceito
Declaração de Manuseio Material Biológico / Biorepositório / Biobanco	Declaracao_sobre_uso_e_destino_do_material_biologico.PDF	26/10/2015 16:32:33	Emmanuel Dias-Neto	Aceito
Outros	Declaracao_sobre_dados_coletados.PDF	26/10/2015 16:31:57	Emmanuel Dias-Neto	Aceito
Outros	Declaracao_sobre_a_publicacao_dos_dados_de_pesquisa.PDF	26/10/2015 16:30:19	Emmanuel Dias-Neto	Aceito
Outros	Identificacao_dos_Curriculos_na_Plataforma_Lattes.PDF	26/10/2015 16:28:27	Emmanuel Dias-Neto	Aceito
Outros	Declaracao_sobre_plano_de_recrutamento_dos_participantes_de_pesquisas.PDF	26/10/2015 16:26:31	Emmanuel Dias-Neto	Aceito
Outros	Termo_de_compromisso_do_pesquisador_com_resolucoes_do_Conselho	26/10/2015 16:24:48	Emmanuel Dias-Neto	Aceito

Endereço: Rua Professor Antônio Prudente, 211

Bairro: Liberdade

CEP: 01.509-900

UF: SP

Município: SÃO PAULO

Telefone: (11)2189-5020

Fax: (11)2189-5020

E-mail: cep_hcancer@accamargo.org.br



FUNDAÇÃO ANTÔNIO
PRUDENTE-HOSPITAL DO
CÂNCER A C CAMARGO



Continuação do Parecer: 1.373.212

Outros	Nacional_da_Saude.PDF	26/10/2015 16:24:48	Emmanuel Dias-Neto	Aceito
Declaração de Pesquisadores	Declaracao_de_Ciencia_e_Comprometimento_Oncogenetica.PDF	26/10/2015 16:19:32	Emmanuel Dias-Neto	Aceito
Declaração de Pesquisadores	Declaracao_de_Ciencia_e_Comprometimento_Endoscopia.PDF	26/10/2015 16:18:46	Emmanuel Dias-Neto	Aceito
Declaração de Pesquisadores	Declaracao_de_Ciencia_e_Comprometimento_Cirurgia_Abdominal.PDF	26/10/2015 16:18:15	Emmanuel Dias-Neto	Aceito
Declaração de Pesquisadores	Declaracao_de_Ciencia_e_Comprometimento_Biobanco.PDF	26/10/2015 16:17:49	Emmanuel Dias-Neto	Aceito
Declaração de Instituição e Infraestrutura	Declaracao_de_Infraestrutura_Anatomia_Patologica.pdf	26/10/2015 16:14:20	Emmanuel Dias-Neto	Aceito
Declaração de Instituição e Infraestrutura	Declaracao_de_Infraestrutura_CIPE.pdf	26/10/2015 16:11:22	Emmanuel Dias-Neto	Aceito
Projeto Detalhado / Brochura Investigador	Tematico_para_CEP_221015.docx	26/10/2015 16:06:06	Emmanuel Dias-Neto	Aceito

Situação do Parecer:

Aprovado

Necessita Apreciação da CONEP:

Não

SAO PAULO, 17 de Dezembro de 2015

Assinado por:
Jefferson Luiz Gross
(Coordenador)

Endereço: Rua Professor Antônio Prudente, 211

Bairro: Liberdade

CEP: 01.509-900

UF: SP

Município: SAO PAULO

Telefone: (11)2189-5020

Fax: (11)2189-5020

E-mail: cep_hcancer@accamargo.org.br

Anexo 2 - Carta de aprovação do Comitê de Ética em Pesquisa-CEP



A.C. Camargo Cancer Center

Centro Integrado de Diagnóstico, Tratamento, Ensino e Pesquisa

**COMITÊ DE ÉTICA
EM PESQUISA - CEP**

São Paulo, 25 de julho de 2018.

Ao

Dr. Emmanuel Dias-Neto

Aluna: Gabriela Estrela de Albuquerque

Ref.: Projeto de Pesquisa nº. 2134/15E

"Caracterização da microbiota gástrica residente e transiente em pacientes portadores de Adenocarcinoma Gástrico."

Os membros do Comitê de Ética em Pesquisa em Seres Humanos da Fundação Antonio Prudente – A.C. Camargo Cancer Center, em sua última reunião de **17/07/2018** **tomaram conhecimento e aprovaram** o seguinte documento:

- Solicitação de dispensa da submissão da documentação obrigatória e análise ética do projeto acima mencionado por se tratar de um projeto afiliado ao temático intitulado "Genômica de Adenocarcinomas Gástricos", registrado neste CEP sob nº 2134/15. O projeto afiliado em referência será Mestrado da aluna Gabriela Estrela de Albuquerque, sob orientação do Dr. Emmanuel Dias-Neto.

Atenciosamente,

Dra. Sandra Caires Serrano

2ª. Vice-Coordenadora do Comitê de Ética em Pesquisa

Apêndice 1 - Termo de Consentimento Livre e Esclarecido-TCLE



Projeto: Aspectos Genômicos de Adenocarcinomas Gástricos

Termo de Consentimento Livre e Esclarecido

Número do Protocolo:
Título do Estudo: Aspectos Genômicos de Adenocarcinomas Gástricos
Nome completo do Pesquisador: Dr. Emmanuel Dias-Neto
Nome da Instituição: A.C. Camargo Cancer Center
Telefone do pesquisador: (11) 2189-5023

Nome do Participante: _____

RGH: _____ N° de Identificação no estudo: _____

DADOS SOBRE O ESTUDO

Este é um projeto de pesquisa do Centro Internacional de Pesquisas e do Núcleo de Cirurgia Abdominal do AC Camargo Cancer Center, está sob responsabilidade do Dr. Emmanuel Dias-Neto.

Todos os estudos de pesquisa devem ser aprovados por um comitê de ética em pesquisa antes que qualquer sujeito possa participar. O comitê de ética ajuda a proteger os interesses dos sujeitos em estudos de pesquisa. Este estudo foi aprovado pelo Comitê de Ética em Pesquisa do Centro de Tratamento e Pesquisa A. C. Camargo Cancer Center.

EXPLICAÇÕES DO PESQUISADOR AO PARTICIPANTE OU SEU REPRESENTANTE LEGAL

Convidamos você a participar deste estudo, que visa compreender as bases do desenvolvimento do câncer de estômago. Apesar de vários estudos já terem sido feitos, ainda faltam análises específicas que poderão ajudar a compreender por que o câncer se desenvolve, quão agressivo ele será e quais os pacientes que devem se beneficiar do tratamento quimioterápico antes da cirurgia. Ainda não sabemos também sobre a importância ou influência da mistura de raças no comportamento dos tumores, mas hoje dispomos de tecnologias modernas que permitem avaliar estas questões e contribuir para o melhor conhecimento desta doença.

Objetivos da pesquisa

Este é um estudo sobre a importância de diversos fatores que podem contribuir para o desenvolvimento do câncer gástrico. Neste estudo, usando amostras de pessoas saudáveis e de pessoas com câncer, avaliaremos as *Aspectos Genômicos de Adenocarcinomas Gástricos*

TCLE versão 3.0 Abril/19

bactérias presentes no estômago, as alterações em substâncias importantes (como proteínas, e o material genético presente dentro das células, DNA e RNA) e aspectos relacionados à etnia (“raça” - européia, africana, asiática ou outras, que podem ter contribuído para a sua formação).

Em indivíduos saudáveis e também naqueles com de câncer de estômago, iremos avaliar estas alterações e como elas variam entre indivíduos dos grupos com e sem doença. Iremos comparar ainda as alterações características de indivíduos que tenham uma ocorrência freqüente de câncer na família, indivíduos com melhor ou pior resposta ao tratamento por quimioterapia e indivíduos com tumores de comportamento mais agressivo versus menos agressivo. Análises destes dados podem ajudar na compreensão de um possível papel destes fatores no desenvolvimento do câncer.

Procedimentos que serão utilizados e propósitos

Caso aceite, você poderá participar de todos os estudos, ou de apenas algum deles, dependendo do momento em que você tiver sido contatado e conforme você se enquadre nos grupos a serem avaliados. Para o projeto que estuda as bactérias presentes no estômago é necessário que você tenha um pedido de um exame de endoscopia (um exame onde uma sonda é passada por sua boca, até atingir o estômago, e permite a sua visualização), que deve ter sido solicitado pelo seu médico pessoal. Esclarecemos que o exame não é solicitado por um médico deste estudo e os seus custos não serão cobertos por este estudo. Neste exame o interior do seu estômago será analisado para ver se existem alterações. É feito um procedimento de lavagem do interior do estômago. Após a lavagem, o líquido é normalmente aspirado e jogado em lixo adequado. Se obtivermos a sua autorização, ao invés de descartarmos este líquido no lixo, ele será usado para os estudos. Para os demais estudos você pode participar na condição de um controle (sem diagnóstico de câncer) ou ainda caso o câncer de estômago tenha sido diagnosticado em você ou em membros da sua família. Neste caso, de acordo com o grupo do estudo em que você se enquadre, solicitamos sua autorização para a coleta de uma pequena amostra de saliva, de sangue (usando procedimentos de rotina) ou, quando for o caso, de pequenos fragmentos do tumor obtidos na rotina do seu tratamento ou de seus exames. Em nenhum caso a coleta destas amostras irá prejudicar o seu diagnóstico nem o seu tratamento.

Caso aceite participar do estudo você deverá:

- Ter mais que 18 anos de idade;

Quando você receber este formulário, explicaremos a você quais as razões da pesquisa e seu objetivo. Possivelmente outros pesquisadores do AC Camargo devem te contatar para obter outras informações tais como alguns hábitos comuns tais como o que você geralmente come, bebe, se faz uso de cigarro ou bebidas com álcool, medicamentos e outros.

Caso concorde em participar, por favor, assine este Termo de Consentimento em 2 vias. Uma ficará em seu poder e a outra será arquivada pelo pesquisador.

Esclarecemos ainda que você tem toda a liberdade de participar ou não do estudo, e sua decisão, seja ela qual for, não afetará de nenhum modo o seu tratamento. Caso você concorde, esta será uma colaboração voluntária, não havendo nenhuma forma de remuneração por isto. Do mesmo modo, esta decisão não afeta de nenhum modo o que será cobrado de você ou de sua seguradora de saúde durante o tratamento. A qualquer momento, mesmo que você já tenha concordado em participar e fornecido a sua assinatura, você tem toda a liberdade de

desistir e solicitar o descarte da amostra coletada, sem que isso ocasione qualquer prejuízo a você ou a seu tratamento.

Estudos a serem realizados e material necessário (marque nos quadros se você aceita participar, se não deseja participar, ou se o estudo não se aplica – caso o procedimento necessário não se aplique ao seu caso, conforme orientação que você terá recebido no momento de preenchimento deste documento)

1) Avaliação das bactérias presentes no estômago e sua relação com o desenvolvimento do câncer de estômago. Amostra necessária: suco gástrico normalmente coletado durante a endoscopia, e que seria jogado fora. O que faremos: iremos determinar quais as bactérias podem ser encontradas em seu estômago e avaliaremos se elas podem ser associadas ao surgimento do câncer. ☐ aceito; ☐ não aceito; ☐ não se aplica.

2) Avaliação de alternativas de diagnóstico usando compostos voláteis (elementos que evaporam) derivados da respiração ou que podem exalar a partir do suco gástrico. Amostra necessária: 1) suco gástrico normalmente coletado durante a endoscopia (como descrito acima), e 2) um pouco do ar que você exala quando respira. O que faremos: 1) serão avaliados os compostos que evaporam do suco gástrico, usando um aparelho específico; 2) pediremos para você soprar através de um canudo estéril e novo, para coletarmos e analisarmos compostos presentes na respiração. Os resultados talvez ajudem no desenvolvimento de testes simples e baratos para diagnosticar pacientes no futuro. ☐ aceito; ☐ não aceito; ☐ não se aplica.

3) Avaliação de alterações de DNA relacionadas ao desenvolvimento do câncer (incluindo câncer familiar) ou aspectos de etnia (raça). Amostra necessária: uma pequena quantidade de saliva (cerca de 2mL) ou 15 mL de sangue (equivalente a uma colher de sopa) de sangue coletado de uma veia do braço. O que faremos: no DNA obtido buscaremos alterações relacionadas com o desenvolvimento do câncer em função da raça (etnia) determinada por análise de DNA. ☐ aceito; ☐ não aceito; ☐ não se aplica.

4) Avaliação de marcadores de resposta à quimioterapia. Amostra necessária: pequenos pedaços do tumor (coletados antes ou depois da quimioterapia) após diagnóstico; e 15 mL (equivalente a uma colher de sopa) de sangue coletado de uma veia do braço. As coletas de sangue poderão ser feitas em diferentes momentos do tratamento/acompanhamento médico. O que faremos: buscaremos alterações de substâncias (DNA, RNA ou proteínas) no material coletado, e no sangue buscaremos células tumorais. ☐ aceito; ☐ não aceito; ☐ não se aplica.

5) Determinação de agressividade tumoral. Amostra necessária: pequenos pedaços do tumor, coletados normalmente no exame de endoscopia antes da quimioterapia, 15 mL (equivalente a uma colher de sopa) de sangue coletado de uma veia do braço. O que faremos: buscaremos alterações de substâncias (DNA, RNA ou proteínas) no material coletado. ☐ aceito; ☐ não aceito; ☐ não se aplica.

Observação: caso você aceite participar de um ou mais estudos acima, que envolvam a coleta de sangue (estudos 3, 4 e 5), esclarecemos que uma única coleta inicial já atende a todos estes estudos. No entanto, coletas em outros momentos que você vier ao A.C. Camargo podem ser necessárias para os estudos 4 e 5.

Complicações e riscos esperados

Caso você aceite participar deste estudo, esclarecemos que não haverá nenhum custo além daqueles necessários para cobrir os exames solicitados por seu médico para a avaliação do seu caso. Caso você participe doando uma pequena amostra de sangue, as raras complicações associadas a esta coleta compreendem: dor pela picada da agulha, pequeno sangramento ou hematoma no local onde foi usada a agulha para coletar o sangue, mas este não deve causar nenhum mal. Todas as perguntas que serão feitas a você, e também o método de coleta de células e amostras que serão utilizados neste estudo foram previamente empregados em outros estudos.

Confidencialidade das informações

Todos os dados levantados serão usados apenas para pesquisa científica, visando a melhor compreensão do câncer de estômago. Todos os esforços serão feitos no sentido de manter a privacidade das informações, que ficarão restritas aos médicos, pesquisadores e enfermeiros diretamente envolvidos neste projeto. Neste sentido, todo e qualquer dado que possa levar à sua identificação será codificado, fazendo com que o risco de perda de confidencialidade seja mínimo.

Benefícios que poderão ser obtidos

Os participantes não se beneficiarão diretamente com a participação nesse estudo. Contudo, o estudo contribuirá para um melhor entendimento dos fatores que podem causar o surgimento ou a evolução do câncer de estômago, desta forma ajudando no desenvolvimento de métodos de prevenção.

CONSENTIMENTO PÓS-INFORMADO

Declaro que, após ter sido adequadamente esclarecido dos riscos e benefícios deste estudo, conforme definidos acima, eu concordo em participar do projeto de pesquisa intitulado: “Aspectos Genômicos de Adenocarcinomas Gástricos”.

São Paulo, de de 20_____

Assinatura do participante ou responsável legal

Assinatura do pesquisador ou responsável pela
aplicação do termo

Em caso de dúvida, entre em contato com o Dr. Emmanuel Dias-Neto, no Laboratório de Genômica Médica do A.C. Camargo Cancer Center, pelo telefone (11) 2189-5023.

Caso o pesquisador responsável não forneça informações/esclarecimentos suficientes, entre em contato com o Comitê de Ética em Pesquisa da Fundação Antônio Prudente/A.C. Camargo Cancer Center pelo telefone (11) 2189-5020. Este Comitê de Ética se localiza à Rua Prof. Antônio Prudente, 211, Liberdade; horário de atendimento: segunda-feira à sexta-feira das 8 às 17 horas. O Comitê de Ética em Pesquisa é um grupo formado por cientistas e não-cientistas que realizam a revisão ética inicial e contínua do estudo clínico para manter sua segurança e proteger seus direitos.

A quantitative view of bacteria and fungi in human samples.

Albuquerque, GE¹, et al.,

1 - Laboratory of Medical Genomics, A.C.Camargo Cancer Center, Rua Taguá 440, 01508-010, São Paulo, SP, Brazil

1. Introduction

The human microbiota is a community of billions of microorganisms, such as viruses, archaea, bacteria, fungi, protists and other eukaryotes which seem to be important for health maintenance and tissue homeostasis. The disbalance of these communities, known as dysbiosis, may lead to diverse pathological states. Research on this field has recently flourished after the reduction in sequencing costs and the development of new high throughput bioinformatics pipelines. These advances allowed the generation and analysis of millions of sequencing reads that go beyond the human-associated microorganisms and have permitted a deeper view of the microbiota associated with plants (Hardoim et al., 2008; Khan et al., 2009), soils (Baldrian et al., 2012; Auffret et al., 2016) urban-environments (MetaSUB International Consortium., 2016), hospitals (Lax et al., 2017; Chng et al., 2020), the international space-station (Chęcinska et al., 2019) and many other sources. Remarkably, advances in human-associated microbiota allowed the investigation of health and disease aspects, going from the impact of c-section versus vaginal delivery (Biasucci et al., 2010; Mueller et al., 2015) to studies of the impact of carcinogens (Thomas et al., 2019), neurodegenerative diseases (Poole et al., 2013; Scheperjans et al., 2015), and how the microbiota may impact cancer treatment (Lehouritis et al., 2015).

In most cases, the desired and more informative microbiota-investigating protocols would include the straightforward extraction of the total DNA from a given sample, followed by library construction and sequencing, known as shotgun metagenomics. The shotgun approach, depending on the efficiency of the DNA/RNA extraction methods, would allow the capture of all nucleic acids available, allowing a less skewed view of the microbiota (Quince et al., 2017). However, as the human genome is much larger (10-10,000X) as compared to viruses, bacteria, fungi, or parasites, the total DNA pool can easily become dominated by host DNA. This will not only affect proper inferences of abundance but may even impede the effective microbiota analysis when the non-human:human ratios

are less favorable. In this case, shotgun sequencing would be severely impacted, leading to a dramatic reduction in the proportion of microbiota-derived reads, because most reads (>99.9%) would come from the host (Marotz et al., 2018).

For this reason, studies of the tissue-associated microbiota remain limited to cultivation protocols or the study of targeted regions of the genome, such as 16s rRNA or ITS (Gophna et al., 2006; Esposito et al., 2017; Aykut et al., 2019). While the first is impacted by the lack of media that would allow the unbiased growth of all microorganisms as a means to reflect their abundance in the sample of interest, the later imposes limitations such as representation bias and lower phylogenetic classification (Donachie et al., 2007; Campanaro et al., 2018). These limitation biases have certainly affected the selection of sample sources and experimental approaches for the investigation of the microbiota and, as a consequence, most publications have focused on the study of the gut/fecal microbiota, where the bacteria:human ratios are very favorable. In particular, the gut microbiota has been shown to impact local immunomodulation or chemotherapy response in oncology (revised by Anfossi & Calin 2020), and these studies would benefit from more comprehensive shotgun metagenomics approaches. Notwithstanding, other body tissues contain important and unique microbiota, whose function and interest have been recently revealed (Chng et al., 2016; Poore et al., 2020). In this sense, if the non-human:human DNA ratios are known, this knowledge may indicate situations where shotgun metagenomic studies would be feasible, allowing a more comprehensive and less biased microbiota evaluation. The quantitative analysis of the microbiota would reveal the relative proportions of bacteria: or fungi:human allowing the amounts of these organisms to be considered during microbiome analysis, and also giving support to select the most appropriate methods for assessing the population of microorganisms in specific tissue types.

Here we give an approach for a quantitative view of bacteria: and fungi:human DNA ratios for different sample types. For some samples the results were compared to shotgun sequencing to further confirm the proportions of each DNA source.

2. Materials and methods

2.1 Sample types and DNA extraction methods

A total of 87 samples were included, representing nine different tissue types: oral mucosa swabs (n=10), saliva (n=10), gastric fluids collected during endoscopy (n=16), stool (n=5), tissue biopsies of salivary glands (n=6), stomach (n=10), rectum (n=10), penis (n=10), and skin (n=10). DNA from oral swabs and skin samples were extracted with DNeasy PowerSoil Kit (Qiagen); saliva and gastric fluids with Phenol:Chloroform:Isoamyl Alcohol 25:24:1 (Merck); salivary gland, rectum, penis and stool-DNA was obtained with E.Z.N.A.[®] Bacterial DNA Kit (Omega), and stomach DNA was extracted

with AllPrep DNA/RNA Mini Kit (Qiagen). DNA samples were quantified with Qubit™ dsDNA HS Assay Kit (ThermoFisher) and qPCR amplified. The collection and study of these samples was previously approved by the A.C.Camargo Cancer Center ethics committee.

2.2 Primer pairs and qPCR DNA amplification

Bacteria- and fungi-derived DNA, were respectively evaluated by DNA amplification of a short fragment of the V1 region of the 16S rRNA (V1_F: 5'-AGAGTTTGATCMTGGCTCAG-3' and V1_R: 5'-TTACTCACCCGTICGCCRCT-3') or a portion of the 5.8S rRNA region (5.8S_F: 5'-CARCAAYGGATCTCTTGG-3' and 5.8S_R: 5'-TGTGCGTTCAAAGATTCGAT-3'). For evaluating the amounts of human DNA, a normalizer for tissue derived microbiota, we used a fragment of similar size of the human *ACTB* gene (ACTB_F: 5'-CCATCTACGAGGGGTATGC-3' and ACTB_R: 5'-GGTGAGGATCTTCATGAGGTA-3'). The qPCR mix for *ACTB* and 16S-V1 genes consisted of 0.2μM of each primer, 5μl of Fast SYBR Green 2x (ThermoFisher) and 3μl of the eluted DNA in various concentrations, in a final volume of 10μl. Cycling conditions consisted of an initial denaturation (95°C for 20 secs) followed by 40 cycles at 95°C for 3 secs and 60°C for 30 secs. For 5.8S gene fragment, the qPCR mix included 0.2μM of each primer, 5μl of the GoTaq Master Mix 2x (Promega) and 4.2μl of eluted DNA in a final volume of 10μl. Cycling conditions for qPCR were: initial denaturation at 95°C for 3mins, followed by 40 cycles at 95°C for 15 secs and 58°C for 30 secs. All experiments were conducted in duplicates and qPCR was performed in a 7500 Real-Time PCR System (ThermoFisher).

2.3 Standard curves

For absolute quantitative analysis by qPCR, we utilized a commercially-available human genomic DNA (Promega - G1471) to generate a standard curve of amplicons from the human β-actin gene. The ZymoBIOMICS™ Microbial Community DNA Standard (Zymo Research - D6305/D6306), composed of five bacteria and two fungi species was used to prepare standard curves for bacteria and fungi, starting with a DNA mass of 5ng. The mean CTs corresponding to each dilution point were used for plotting the standard curves.

The number of genomes per ng of DNA was calculated for bacteria, fungi and human DNA (**Fig. S1**). Considering genome sizes of 6.4Gb (human diploid genome), 19.78Mb (fungi) and 3.95Mb (bacteria). The average genome sizes used for bacteria and fungi were determined according to complete genomes sizes of 16,031 bacteria and 64 fungi available at the National Center for Biotechnology (NCBI) "Genome Browse by Organism" (<https://www.ncbi.nlm.nih.gov/assembly/organism/>). Standard curves correlating the mean CTs with the total genome copies/ng of DNA were constructed. Each standard curve is given as: i) expected CTs - considered according to the lowest CT-value, under the assumption that a difference of one CT would

correspond to a 2X difference in the mass of template DNA and ii) observed CTs - values observed after qPCR amplification.

$\text{number of copies} = \frac{\text{mass (ng)} * 6.022 \times 10^{23}}{\text{length} * 1 \times 10^9 * 650}$ Were: 6.022x10²³ number of molecules/mol= Avogadro's number 650 = average weight of a base pair (bp) 1x10⁹ = to convert g to ng length = genome length

Figure S1. Calculations used to determine the number of genomes per 1ng of genomic DNA (<http://cels.uri.edu/gsc/cndna.html>).

2.4 Shotgun sequencing

As another metric to evaluate bacteria:human and fungi:human ratios, we compared qPCR ratios to shotgun metagenomic sequencing of 10 samples of gastric fluids and two stool samples. For that, we used 100 ng of DNA-template for library preparation for each sample (Nextera DNA Flex kit, Illumina), followed by sequencing in the NextSeq 500 platform (Illumina). The results were analyzed by Kraken2 (Wood et al 2019), and the number of reads for bacteria, fungi and human was used to calculate the relative frequency of each group, and to compare with qPCR data.

3. Results

3.1 Standard curves correlating genome copies and CTs

Standard curves were plotted, allowing to correlate the number of genome copies for bacteria and fungi according to the respective CTs achieved in qPCR amplification (**Table S1**). As shown in **Fig. 1**, all standard curves showed very high correlations (R^2 equal to 0.99 or 1), despite some delay for observed CTs for fungi, especially in the more diluted points of the curve.

Table S1. Calculated correlations between DNA mass, mean CTs and genome copies.

Human (ACTB)					
Mass qPCR (ng)	Mass ACTB (ng)	Mean CT	Number of genome copies in 1ng	Number of copies in Xng	Log10
5	5	25.39	144.77	723.85	2.86
0.8333	0.8333	27.82	144.77	120.64	2.081
0.1389	0.1389	31.08	144.77	20.11	1.303
0.0231	0.0231	33.82	144.77	3.35	0.525
0.0039	0.0039	36.9	144.77	0.56	-0.253
0.0006	0.0006	Undetermined	144.77	0.09	-1.031
Mock - Bacteria (16S)					
Mass qPCR (ng)	Mass bacteria (ng)	Mean CT	Number of genome copies in 1ng	Number of copies in Xng	Log10
5	4.8	12.06	234275.4	1124522	6.051
0.45045	0.432432	15.78	234275.4	101308.3	5.006
0.040581	0.038958	19.81	234275.4	9126.87	3.96
0.003656	0.00351	23.84	234275.4	822.24	2.915
0.000329	0.000316	27.14	234275.4	74.08	1.87
0.00003	0.000028	30.01	234275.4	6.67	0.824
0.000003	0.000003	32.44	234275.4	0.6	-0.221
Mock - Fungi (5.8S)					
Mass qPCR (ng)	Mass fungus (ng)	Mean CT	Number of genome copies in 1ng	Number of copies in Xng	Log10
5	0.2	17.09	46844.62	9368.92	3.972
0.5	0.02	21.06	46844.62	936.89	2.972
0.05	0.002	25.33	46844.62	93.69	1.972
0.005	0.0002	29.22	46844.62	9.37	0.972
0.0005	0.00002	32.56	46844.62	0.94	-0.028
0.00005	0.000002	36.43	46844.62	0.09	-1.028

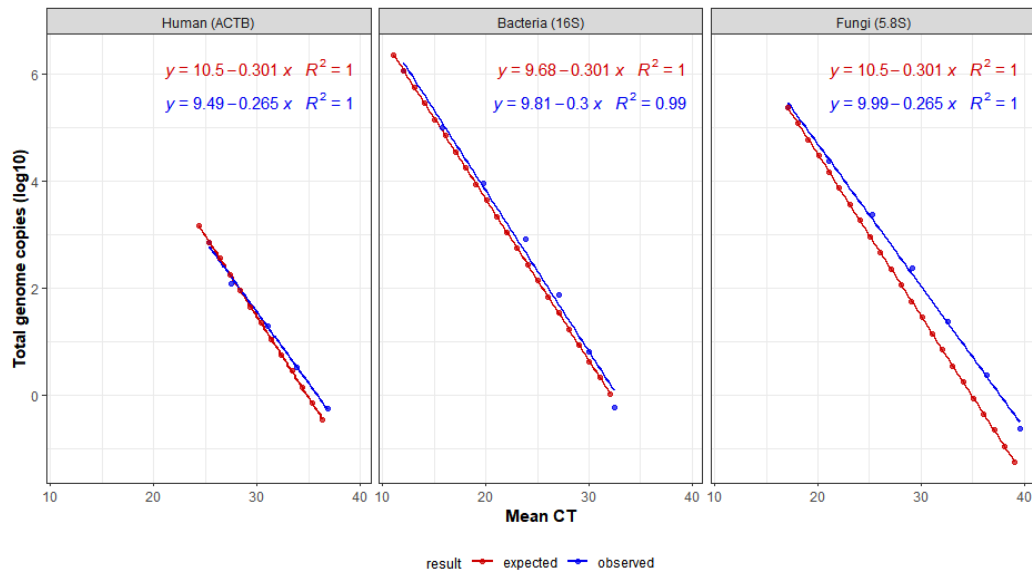


Figure 1. Standard DNA dilution curves for Human (ACTB), bacteria (V1-16S) and fungi (5.8S) qPCR, showing expected (red) and observed (blue) CT-values.

3.2 Impacts of varying bacteria:human and fungi:human DNA-ratios over amplification/quantification capabilities

In order to simulate real-life analysis of sample types where significant variations in the ratios of target microorganisms and host-derived DNA are common, we evaluated the CT fluctuations in face of different DNA ratios and their possible impact over bacteria/fungi quantification. When using a fixed amount of human DNA (1ng) and up to 500X variation (5ng to 0.01ng) of the mock DNA, the CT for the human ACTB gene remained within the expected range of variation. The very same finding was observed when the amount of mock DNA was fixed at 1ng, and the human DNA varied 500X (0.01 ng to 5 ng), with no significant impacts seen for the bacteria 16S- or the fungi 5.8S-detection (**Figure 2**).

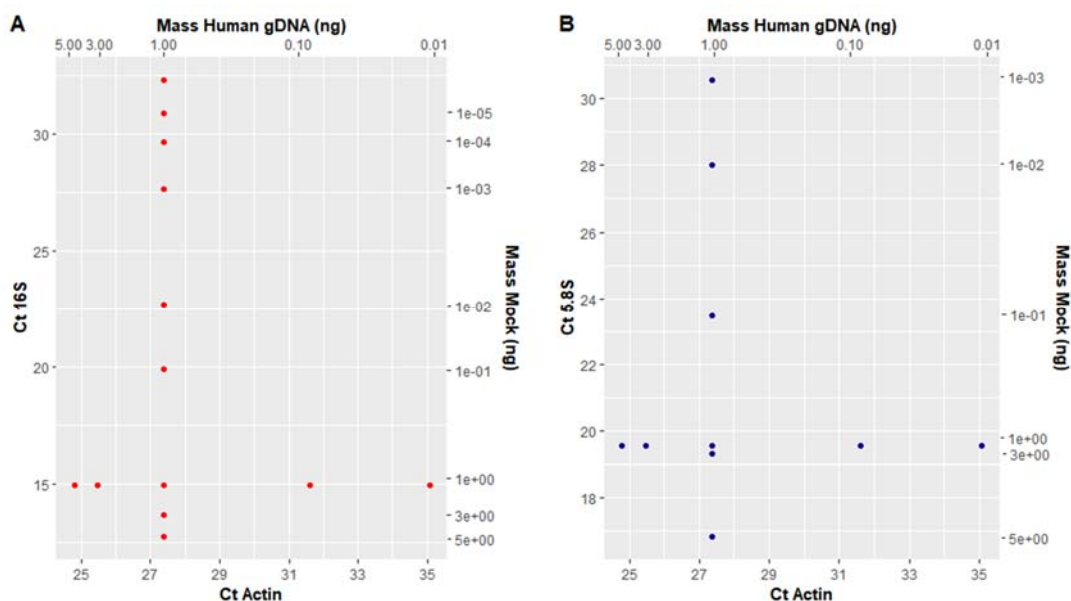


Figure 2. Effects of varying DNA mass over amplification capabilities of bacteria (A) and fungi (B), as compared to human *ACTB*.

3.3 Genome copies in different sample-types

After running the standard curves for different amounts of DNA, and showing no significant variations of background human DNA to affect the expected CTs, we were able to determine the approximate number of bacteria or fungi genome copies for nine different types of samples. **Figure 3A** shows, as expected, that stool samples have the highest bacteria:human ratios - making this a very good source for metagenome shotgun sequencing, as published by many different groups. Our data also shows that skin, oral swabs, saliva and gastric fluids also contain more bacteria than human DNA and, therefore, shall provide informative microbiota data using this same sequencing approach, even considering the much larger length of the human genome. On the other hand, human-derived DNA prevails in tissue-derived samples, still with high amounts of bacteria DNA being observed in rectal and penile biopsies, suggesting that metagenomics shotgun could still produce informative data from these samples. As for the mycobiome analysis, we see high amounts of fungi-derived DNA in stool and skin samples, but also in gastric fluids of some samples (**Figure 3B**).

After comparing the approximate number of human and bacteria genomes in all samples, we observe that only salivary gland and stomach biopsies have less bacteria than human DNA (**Fig. 3C**). Rectum and penile biopsies appear to have similar amounts of human and bacterial genome copies, whereas the other samples types have significantly more bacteria, being stool the sample with most bacterial DNA. From our limited set of samples, only stool and skin samples seem to have higher fungi:human genome-ratios. Oral swabs, saliva and gastric fluids apparently have similar proportions of genomes from each organism, which makes sense as the gastric microbiota is heavily influenced by the swallowing of large amounts of microorganisms present in the saliva (Bassis et al., 2015; Schulz et al., 2018).

The quantitative analysis of fungi by qPCR showed stool and skin to be the samples with more fungi genome copies, with a reduction in saliva, oral swabs, rectal biopsies, all with relatively similar amounts of fungi genomes. Interestingly, very small amounts of human genomes were found in the skin and stool, as compared with the other DNA sources evaluated here (**Figure 3B, C**).

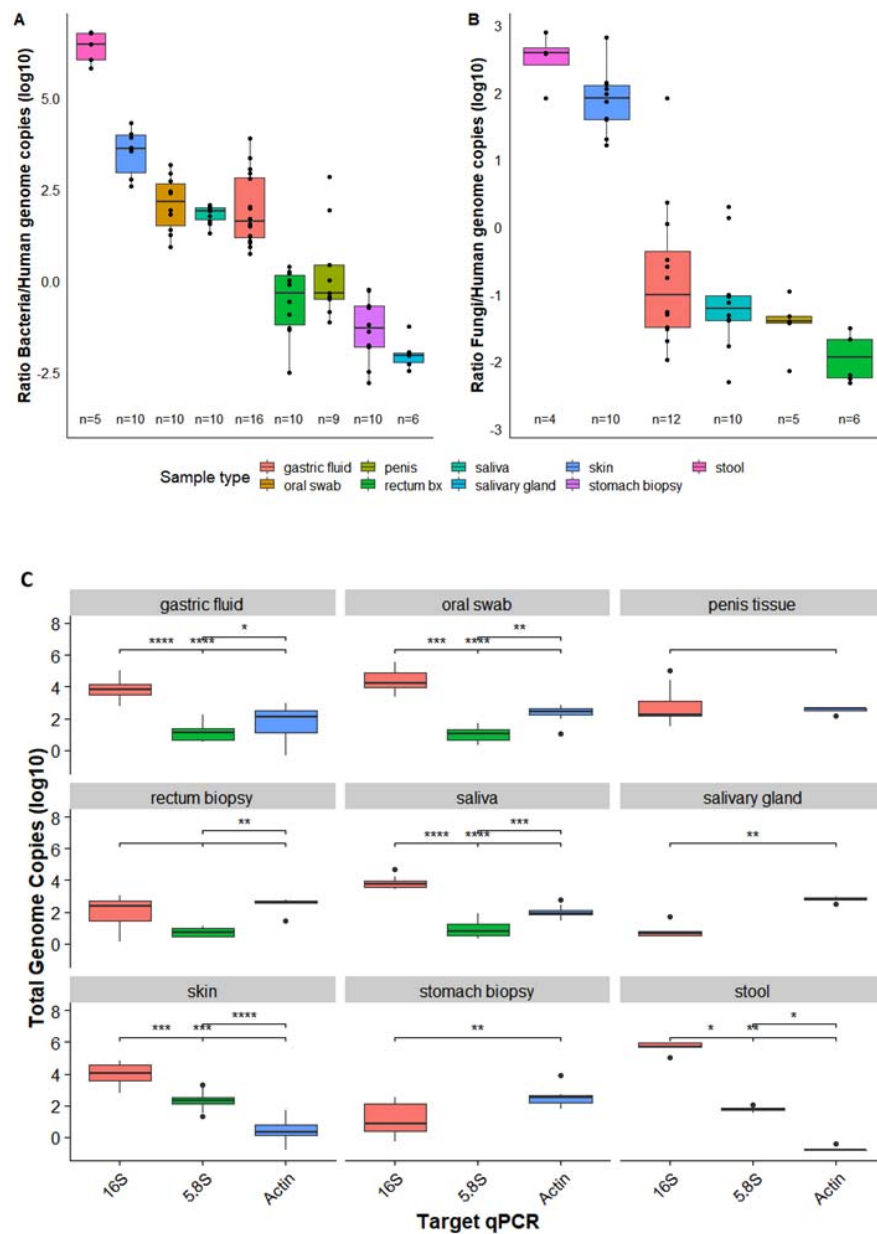


Figure 3. Genome-ratios for bacteria:human (A) and fungi:human (B) and relative amounts of both (C) in different types of human-derived samples.

3.4 Shotgun analysis

As shotgun sequencing is a less biased approach, we have also used this method to compare the data with our qPCR results, performing an analysis of relative abundance between bacteria, fungi and human-derived DNA from 18 samples, including 16 samples of gastric fluids and two stool samples. This analysis showed a percentage of bacteria-derived reads varying from 2 to 93% in gastric fluids, and was 85% and 87% in stool samples. As expected, we observed for both sample types, a smaller contribution of fungi-DNA, as compared to bacteria. The presence of fungi was confirmed in all samples

but with low frequency, varying from 0.04 to 0.86% in gastric fluids and of 0.5% in the two samples of stool.

Our shotgun sequencing results were also compared in terms of bacteria:human and fungi:human genome ratios as calculated by our qPCR analysis from the same samples. As can be seen in **Figure 4**, very good correlations were reached between qPCR data and the ratio of bacteria:human reads (**Fig 4A**) and fungi:human reads (**Figure 4B**), obtained by shotgun. This validation reinforces the validity of the qPCR approach designed here as well as the supporting calculations that appear to be reliable to determine the relative amounts of these organisms in stool and in gastric samples, and possibly in other samples as well.

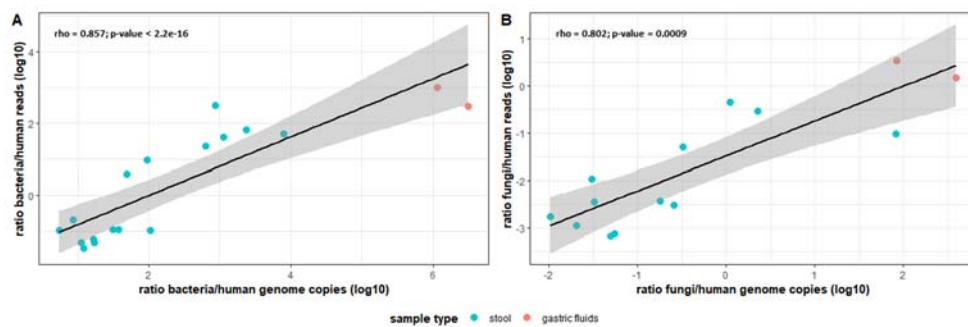


Figure 4. Correlation between (A) ratio bacteria:human reads versus ratio bacteria:human genome copies and (B) ratio fungi:human reads versus ratio fungi:human genome copies. Each circle represents a sample. For statistical analysis, Spearman's correlation was used, and the closer it is to 1, the greater the correlation between the variables. The shaded area represents the confidence interval. The correlation was considered to be statistically significant when $p \leq 0.05$

3.5 Analysis of bacteria:human ratios over time

As an exercise to evaluate the applicability of this approach, we have followed a small cohort of five patients diagnosed with squamous cell carcinoma in the oral cavity. These patients have been treated with radiotherapy and swabs of the neoplastic lesion were collected at diagnosis, before the start of radiotherapy and after the end of this treatment. In **Figure 5** we observe that individuals had varied responses in terms of bacterial ratios along time and at the end of this treatment and this tool was efficient to capture this variation along time.

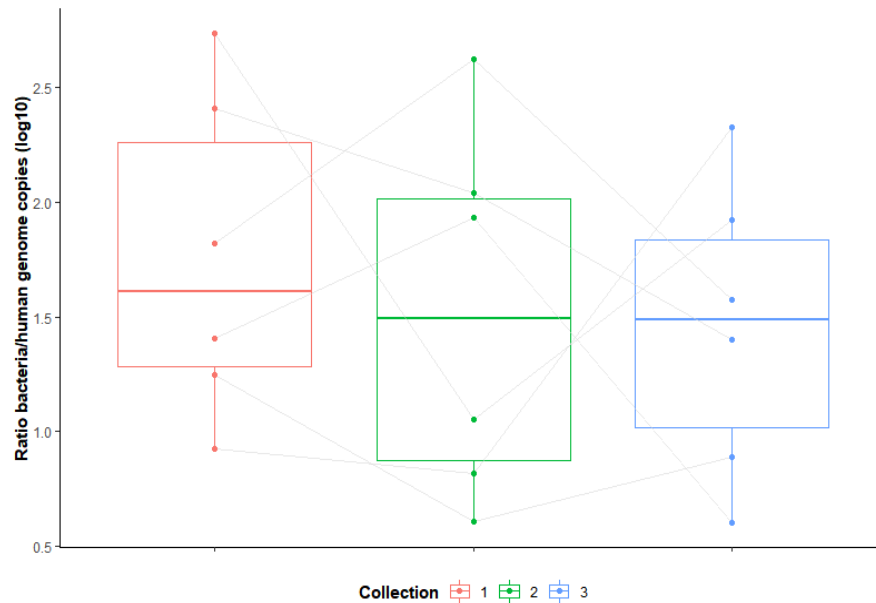


Figure 5. Ratio bacteria:human genome copies of swab samples in the same patient in different times. The first collection point was at the first contact with the patient, the second point was before the radiotherapy and the third point was after the radiotherapy.

4. Discussion

The ratio of bacteria to human cells in a sample has not been considered for most microbiota analysis of human diseases, as the main focus of the majority of the manuscripts has been on the composition and diversity of the microbiota and amount of bacteria in the sample is usually not considered. Additionally, very few authors have focused on the mycobiome composition, not to mention the abundance of fungi in sample types and diseases, and possible impacts of this biomass.

However, it seems clear that in some physiologic situations and for different tissues, important variations can be seen in the number of microorganisms capable of colonizing a certain tissue. It is also possible that variations in the microbiota composition, as well as the ratio of microorganisms:human cells in a certain tissue or sample could be a proxy of the immune surveillance status.

Other applications are more technical and include the use of this tool to effectively indicate the best approach for determining the microbiota composition, indicating with good precision the number of expected reads for bacteria and human-derived DNA in shotgun metagenomics, or the inadequacy of this approach due to unachievable coverage when very little amounts of bacteria is present. A good demonstration of this is the evaluation of the microbiota of gastric fluids by shotgun metagenomics, as predicted by our quantitative analysis. This application can also be visualized for a number of medical/biological situations of interest, as demonstrated here when we investigated bacteria:human ratios for head and neck cancer patients during their treatment radiotherapy. In this case, some authors have suggested the use of specific microorganisms in the protection of the

intestinal epithelium against radiation-induced injuries (He et al., 2018; Rhiel et al., 2019). Our approach could be applied to investigate a series of treatment-related scenarios including the impact of bacteria amounts in pre-radiotherapy x post-radiotherapy stage and outcomes such as the effectiveness of this treatment and the later development of mucositis in different grades. Whereas radiotherapy and chemotherapy efficacy, and the role of the microbiota, including biofilms, as mucosal protective agents or drug metabolizers that could possibly impact cancer therapy and treatment side effects goes beyond the scope of this work, it is possible to envisage the use of this tool to monitor bacterial levels in a fast and inexpensive fashion in diverse scenarios of interest.

References

- Anfossi S, Calin GA. Gut microbiota: a new player in regulating immune- and chemo-therapy efficacy. *Cancer Drug Resist* 2020;3:[Online First].
- Aykut, B., Pushalkar, S., Chen, R. *et al.* The fungal mycobiome promotes pancreatic oncogenesis via activation of MBL. *Nature* 574, 264–267 (2019).
- Auffret MD, Karhu K, Khachane A, Dungait JAJ, Fraser F, Hopkins DW, et al. (2016) The Role of Microbial Community Composition in Controlling Soil Respiration Responses to Temperature. *PLoS ONE* 11(10): e0165448.
- Baldrian, P., Kolařík, M., Štursová, M. *et al.* Active and total microbial communities in forest soil are largely different and highly stratified during decomposition. *ISME J* 6, 248–258 (2012).
- Bassis CM, Erb-Downward JR, Dickson RP, et al. Analysis of the upper respiratory tract microbiotas as the source of the lung and gastric microbiotas in healthy individuals. *mBio*. 2015;6(2):e00037. Published 2015 Mar 3. doi:10.1128/mBio.00037-15
- Biasucci G, Rubini M, Riboni S, Morelli L, Bessi E, Retetangos C. Mode of delivery affects the bacterial community in the newborn gut. *Early Hum Dev*. 2010;86 Suppl 1:13-15.
- Campanaro, S., Treu, L., Kougias, P.G. *et al.* Taxonomy of anaerobic digestion microbiome reveals biases associated with the applied high throughput sequencing strategies. *Sci Rep* 8, 1926 (2018).
- Chng KR, Chan SH, Ng AHQ, et al. Tissue Microbiome Profiling Identifies an Enrichment of Specific Enteric Bacteria in *Opisthorchis viverrini* Associated Cholangiocarcinoma. *EBioMedicine*. 2016;8:195-202.
- Chng KR, Li C, Bertrand D, et al. Cartography of opportunistic pathogens and antibiotic resistance genes in a tertiary hospital environment. *Nat Med*. 2020;26(6):941-951.
- Checinska Sielaff, A., Urbaniak, C., Mohan, G.B.M. *et al.* Characterization of the total and viable bacterial and fungal communities associated with the International Space Station surfaces. *Microbiome* 7, 50 (2019).
- Donachie SP, Foster JS, Brown MV. Culture clash: challenging the dogma of microbial diversity. *ISME J*. 2007;1(2):97-99.
- Esposito S, De Simone G, Gioia R, et al. Deep tissue biopsy vs. superficial swab culture, including microbial loading determination, in the microbiological assessment of Skin and Soft Tissue Infections (SSTIs). *J Chemother*. 2017;29(3):154-158.

Genome Browse by Organism. NCBI. <<https://www.ncbi.nlm.nih.gov/genome/>>

Gophna, Uri et al. "Differences between tissue-associated intestinal microfloras of patients with Crohn's disease and ulcerative colitis." *Journal of clinical microbiology* vol. 44,11 (2006): 4136-41.

Hardoim PR, van Overbeek LS, Elsas JD. Properties of bacterial endophytes and their proposed role in plant growth. *Trends Microbiol.* 2008;16(10):463-471.

He JY, Wang WZ, Qi HZ, Ma Y, He SY. Use of recombinant *Lactobacillus sakei* for the prevention and treatment of radiation-induced enteritis. *Med Hypotheses*. 2018 Oct;119:37-40.

Khan, Z., Doty, S.L. Characterization of bacterial endophytes of sweet potato plants. *Plant Soil* **322**, 197–207 (2009).

Lax S, Sangwan N, Smith D, et al. Bacterial colonization and succession in a newly opened hospital. *Sci Transl Med.* 2017;9(391):eaah6500.

Lehouritis, P., Cummins, J., Stanton, M. et al. Local bacteria affect the efficacy of chemotherapeutic drugs. *Sci Rep* **5**, 14554 (2015).

Marotz CA, Sanders JG, Zuniga C, Zaramela LS, Knight R, Zengler K. Improving saliva shotgun metagenomics by chemical host DNA depletion. *Microbiome*. 2018;6(1):42.

MetaSUB International Consortium. "The Metagenomics and Metadesign of the Subways and Urban Biomes (MetaSUB) International Consortium inaugural meeting report." *Microbiome* vol. 4,1 24. 3 Jun. 2016,

Mueller NT, Whyatt R, Hoepner L, et al. Prenatal exposure to antibiotics, cesarean section and risk of childhood obesity. *Int J Obes (Lond)*. 2015;39(4):665-670.

Oberauner L, Zachow C, Lackner S, Högenauer C, Smolle KH, Berg G. The ignored diversity: complex bacterial communities in intensive care units revealed by 16S pyrosequencing. *Sci Rep*. 2013;3:1413.

Poole S, Singhrao SK, Kesavalu L, Curtis MA, Crean S. Determining the presence of periodontopathic virulence factors in short-term postmortem Alzheimer's disease brain tissue. *J Alzheimers Dis*. 2013;36(4):665-677.

Poore, G.D., Kopylova, E., Zhu, Q. et al. Microbiome analyses of blood and tissues suggest cancer diagnostic approach. *Nature* **579**, 567–574 (2020).

Quince C, Walker AW, Simpson JT, Loman NJ, Segata N. Shotgun metagenomics, from sampling to analysis [published correction appears in *Nat Biotechnol*. 2017 35:1211]. *Nat Biotechnol*. 2017;35(9):833-844.

Riehl TE, Alvarado D, Ee X, Zuckerman A, Foster L, Kapoor V, Thotala D, Ciorba MA, Stenson WF. *Lactobacillus rhamnosus* GG protects the intestinal epithelium from radiation injury through release of lipoteichoic acid, macrophage activation and the migration of mesenchymal stem cells. *Gut*. 2019 Jun;68(6):1003-1013.

Scheperjans F, Aho V, Pereira PA, et al. Gut microbiota are related to Parkinson's disease and clinical phenotype. *Mov Disord*. 2015;30(3):350-358.

Schulz C, Schütte K, Koch N, et al. The active bacterial assemblages of the upper GI tract in individuals with and without *Helicobacter* infection. *Gut*. 2018;67(2):216-225.

Thomas, AM, Manghi, P, Asnicar, F. et al. Metagenomic analysis of colorectal cancer datasets identifies cross-cohort microbial diagnostic signatures and a link with choline degradation. *Nat Med* **25**, 667–678 (2019).

Wood DE, Lu J, Langmead B. Improved metagenomic analysis with Kraken 2. *Genome Biol*. 2019;20(1):257.

Note

Other authors (order to be defined) include: Serpa, MS¹., Moda, BS^{1,2}., Takenaka I¹., Miola E¹., Defelicibus, A²., Amorim MG¹, Branco, G¹., Nunes, DN¹., Bartelli TF¹, Dias-Neto, E¹.

Acknowledgements

ED-N is a researcher from Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq - Brasil). This project was funded in part by Fundação de Amparo À Pesquisa do Estado de São Paulo (FAPESP) and Programa Nacional de Atenção ao Paciente Oncológico (PRONON), departamento de Ciência e Tecnologia, Ministério da Saúde (DECIT, MS), Governo Federal.

1 - Laboratory of Medical Genomics, A.C.Camargo Cancer Center, Rua Taguá 440, 01508-010, São Paulo, SP, Brazil

2 - Laboratory of Computational Biology and Bioinformatics, A.C.Camargo Cancer Center, Rua Taguá 440, 01508-010, São Paulo, SP, Brazil



A.C. Camargo Cancer Center

Centro Integrado de Diagnóstico, Tratamento, Ensino e Pesquisa

São Paulo, 27 de julho de 2021

À FAPESP - Fundação de Amparo à Pesquisa do Estado de São Paulo

Ref. Errata Dissertação

Informamos que recebemos a errata da Dissertação da aluna **Gabriela Estrela de Albuquerque**, orientanda do Dr. Emmanuel Dias Neto, e consideramos a versão corrigida em anexo.

Atenciosamente,

Rubens Chojniak

Coordenador do Programa de Pós-Graduação

233/2021 PPG

Página de assinaturas



Rubens Chojniak
A.C.Camargo Cancer Center
Signatário

HISTÓRICO

- | | | |
|-------------------------|---|--|
| 27 jul 2021
14:33:21 |  | Suely Francisco criou este documento. (E-mail: suely.francisco@accamargo.org.br , CPF: 737.484.398-49) |
| 27 jul 2021
15:10:28 |  | Rubens Chojniak (E-mail: chojniak@accamargo.org.br , CPF: 151.387.888-33) visualizou este documento por meio do IP 189.102.192.116 localizado em São Paulo - Sao Paulo - Brazil. |
| 27 jul 2021
15:10:36 |  | Rubens Chojniak (E-mail: chojniak@accamargo.org.br , CPF: 151.387.888-33) assinou este documento por meio do IP 189.102.192.116 localizado em São Paulo - Sao Paulo - Brazil. |

