

DANIELLE CRISTINA FONSECA CANDIAN

**Avaliação do potencial preditivo da microbiota intestinal
associada a características fenotípicas para identificação de
doenças multifatoriais**

Tese apresentada à Faculdade de
Medicina da Universidade de São Paulo
para obtenção do título de Doutora em
Ciências

Programa - Ciências em
Gastroenterologia

Orientador: Prof. Dr. Dan Linetzky
Waitzberg

Coorientador: Prof. Dr. Gabriel da Rocha
Fernandes

São Paulo

2023

DANIELLE CRISTINA FONSECA CANDIAN

**Avaliação do potencial preditivo da microbiota intestinal
associada a características fenotípicas para identificação de
doenças multifatoriais**

Tese apresentada à Faculdade de
Medicina da Universidade de São Paulo
para obtenção do título de Doutora em
Ciências

Programa - Ciências em
Gastroenterologia

Orientador: Prof. Dr. Dan Linetzky
Waitzberg

Coorientador: Prof. Dr. Gabriel da Rocha
Fernandes

São Paulo

2023

Dados Internacionais de Catalogação na Publicação (CIP)

Preparada pela Biblioteca da
Faculdade de Medicina da Universidade de São Paulo

©reprodução autorizada pelo autor

Candian, Danielle Cristina Fonseca
Avaliação do potencial preditivo da microbiota
intestinal associada a características fenotípicas
para identificação de doenças multifatoriais /
Danielle Cristina Fonseca Candian. -- São Paulo,
2023.
Tese (doutorado)--Faculdade de Medicina da
Universidade de São Paulo.
Programa de Ciências em Gastroenterologia.
Orientador: Linetzky Waitzberg.
Coorientador: Gabriel da Rocha Fernandes.

Descritores: 1.Microbiota intestinal 2.16S rRNA,
modelos de predição 3.Árvores de decisões
4.Enfermidades 5.Reprodutibilidade dos resultados

USP/FM/DBD-112/23

Responsável: Erinalva da Conceição Batista, CRB-8 6755

Dedico essa tese de doutorado aos meus pais, Jairo e Rosana, pelo amor, suporte, incentivo, e por serem exemplo de que, com persistência, trabalho duro e honestidade é possível conquistar nossos sonhos.

Aos meus avós, que sempre foram uma fonte inesgotável de carinho.

A minha irmã, Débora, por ser minha conselheira e melhor amiga em todos os momentos.

Ao meu marido por todo amor, apoio, parceria e compreensão. Por tolerar minha bagunça e o mau humor, principalmente nos dias em que eu “só existi” para a pesquisa. Sem seu incentivo constante eu não teria conseguido.

Aos meus amigos, que resistiram ao passar do tempo e ao meu gênio forte! Obrigada por nunca me abandonarem nessa louca jornada da vida.

Agradecimentos

Agradeço ao meu orientador **Dan L. Waitzberg**, pelo suporte e ensinamentos ao longo da minha jornada na pesquisa. Obrigada por me acolher em seu laboratório quando eu mal sabia ler um artigo científico, e por continuar acreditando em mim ao longo desses quase 8 anos. Sei que cada puxão de orelha, seguido de palavras de incentivo foi importante para que eu pudesse me tornar a profissional que sou hoje. Palavras não seriam o suficiente para agradecer toda a ajuda que recebi do senhor nesses últimos anos. É uma honra enorme aprender com um dos maiores nomes da nutrição e nutrologia em nosso país.

Ao meu coorientador **Gabriel Fernandes**, agradeço o carinho, a amizade e a paciência para me ensinar tanto sobre ciência, estatística, a importância de questionar o p-valor, curiosidades aleatórias e sobre acolhimento. Obrigada por todas as vezes que você me disse “você é boa, o mundo precisa saber disso! Eu acredito em você”. Agradeço sua sensibilidade em me dar bronca quando precisava, me acalmar nos momentos de aflição, e comemorar pelas pequenas conquistas. Você foi essencial para eu conseguisse chegar até aqui. Por fim, mas não menos importante, agradeço o tempo que você investiu em nossas inúmeras reuniões, aulas e troca de ideias sobre meu trabalho e sobre carreira. Espero ter aprendido como ser menos jacu nesses anos de convivência virtual!

À **Raquel Torrinas**, pelos ensinamentos, discussões e por toda a ajuda que me deu ao longo dos meus anos em pesquisa. Tenho muita sorte por poder aprender com uma pessoa que transborda conhecimento e didática. Obrigada por compartilhar comigo sua sabedoria.

Aos **colegas e amigos do METANUTRI**, que ao longo de quase 8 anos, torceram por mim, me aconselharam, me acolheram e vibraram com minhas conquistas. Esse trabalho tem a participação de todos vocês, afinal, ninguém faz pesquisa sozinho. Levarei um pouquinho de cada um de vocês comigo!

Às minhas parceiras de **Projeto Validys**, que compartilharam comigo as dores e delícias de viver esse projeto. Nós conseguimos!

Aos professores-doutores **Claudia Pinto de Oliveira, Ulysses Ribeiro Junior, José Jukemura** e aos demais membros da comissão da Pós-Graduação em Ciências em Gastroenterologia, pelo maravilhoso trabalho à frente de um dos maiores centros de produção científica do país.

Aos funcionários da Disciplina de Cirurgia do Aparelho Digestivo do HC-FMUSP, que me acolheram no programa de pós-graduação. Em especial à **Vilma de Jesus Libério**, pelo auxílio em todas as dúvidas que surgiram pelo caminho, e por aguentar minha “falação” por quase 8 anos! Obrigada pelo carinho e pela gentileza.

A todas as equipes da Unidade de Diabetes do Serviço de Endocrinologia e Metabologia, Unidade de Doenças Reumáticas, Unidade de Psoríase do Serviço de Dermatologia, Unidade de Cirurgia Bariátrica e Metabólica, Unidade de Doenças Inflamatórias Intestinais da Divisão de Gastroenterologia e Hepatologia e da Divisão de Coloproctologia do Hospital das Clínicas da Faculdade de Medicina da Universidade de São Paulo (HC-FMUSP), agradeço o auxílio, oportunidade e ensinamentos.

Ao **serviço de Biblioteca** da Faculdade de Medicina da Universidade de São Paulo pelo auxílio.

À **Bioma4me e FQM Farma** pela gentil disponibilização dos dados de sequenciamento de microbiota intestinal para os subprojetos do trabalho VALIDYS, do grupo METANUTRI.

À Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (**CAPES**), pela bolsa de doutorado que permitiu a minha dedicação a esse trabalho.

Aos **pacientes e voluntários que participaram dessa pesquisa** por doarem seu tempo em prol desse trabalho, sem eles nada disso seria possível.

Aos **amigos e família** que gentilmente entenderam meus sumiços em virtude dos estudos, escrita da tese e alguns mini surtos. Obrigada por

segurarem a barra das minhas reclamações e por serem colo nos momentos mais desafiadores.

Por fim, agradeço a todos aqueles que fazem ou fizeram parte da minha vida e que, direta ou indiretamente, contribuíram para minha formação pessoal e profissional no caminho tortuoso de fazer pesquisa em nosso país.

Pensar é voar sobre o que não se sabe. Não existe nada mais fatal para o pensamento que o ensino das respostas certas. Para isso existem as escolas: não para ensinar as respostas, mas para ensinar as perguntas. As respostas nos permitem andar sobre a terra firme. Mas somente as perguntas nos permitem entrar pelo mar desconhecido.

Rubem Alves

Normalização Adotada

Esta tese está de acordo com as seguintes normas, em vigor no momento desta publicação:

Referências: adaptado de *International Committee of Medical Journals Editors* (Vancouver).

Universidade de São Paulo. Faculdade de Medicina. Divisão de Biblioteca e Documentação. *Guia de apresentação de dissertações, teses e monografias*. Elaborado por Anneliese Carneiro da Cunha, Maria Julia de A. L. Freddi, Maria F. Crestana, Marinalva de Souza Aragão, Suely Campos Cardoso, Valéria Vilhena. 3a ed. São Paulo: Divisão de Biblioteca e Documentação; 2011.

Abreviaturas dos títulos dos periódicos de acordo com *List of Journals Indexed in Index Medicus*.

SUMÁRIO

| | |
|---|----|
| LISTA DE ABREVIATURAS E SIMBOLOS | |
| LISTA DE FIGURAS | |
| LISTA DE GRÁFICOS | |
| LISTA DE QUADROS | |
| LISTA DE TABELAS | |
| RESUMO | |
| ABSTRACT | |
| 1 INTRODUÇÃO | 24 |
| 2 REVISÃO DA LITERATURA | 28 |
| 2.1 Características fenotípicas, microbiota intestinal e sua conversa bidirecional nos estados de saúde e doença | 29 |
| 2.1.1 Dieta..... | 30 |
| 2.1.2 Atividade física | 32 |
| 2.1.3 Tabagismo | 33 |
| 2.2 Composição da microbiota intestinal nas doenças | 34 |
| 2.2.1 Microbiota intestinal nas Doenças Inflamatórias Intestinais (DII) | 36 |
| 2.2.2 Microbiota intestinal na Psoríase em Placas..... | 38 |
| 2.2.3 Microbiota intestinal na Artrite Reumatoide..... | 39 |
| 2.2.4 Microbiota intestinal no Lupus Eritematoso Sistêmico | 41 |
| 2.2.5 Microbiota intestinal no Diabetes tipo 1 (DM 1)..... | 42 |
| 2.2.6 Microbiota intestinal no Diabetes tipo 2 (DM2)..... | 44 |
| 2.3 Microbiota intestinal como potencial biomarcador de doenças..... | 46 |
| 3 JUSTIFICATIVA | 49 |
| 4 OBJETIVOS | 51 |
| 4.1 Objetivo principal | 52 |
| 4.2 Objetivos específicos..... | 52 |
| 5 CASUÍSTICA E MÉTODOS | 53 |
| 5.1 Local de execução do estudo | 54 |

| | |
|--|----|
| 5.2 Aspectos éticos..... | 54 |
| 5.3 Seleção de participantes e pacientes | 54 |
| 5.4 Cálculo amostral | 56 |
| 5.5 Anamnese e história clínica..... | 57 |
| 5.6 Avaliação do Consumo Alimentar..... | 57 |
| 5.7 Avaliação de composição corpórea | 59 |
| 5.7.1 Peso, estatura e cálculo de IMC..... | 59 |
| 5.7.2 Composição corporal | 59 |
| 5.8 Análise fecal - composição da microbiota intestinal..... | 60 |
| 5.8.1 Coleta de amostras | 60 |
| 5.8.2 Extração de DNA..... | 61 |
| 5.8.3 Preparo de biblioteca | 62 |
| 5.8.4 Sequenciamento 16SrRNA | 62 |
| 5.9 Bioinformática e estatística | 63 |
| 5.9.1 Processamento das sequências e classificação taxonômica da microbiota intestinal | 63 |
| 5.9.2 Caracterização de α e β diversidade da microbiota intestinal | 63 |
| 5.9.3 Caracterização da microbiota intestinal..... | 63 |
| 5.9.4 Análise das informações fenotípicas..... | 64 |
| 5.9.5 Análise da ingestão habitual de nutrientes..... | 65 |
| 5.9.6 Seleção de variáveis para integrar os modelos preditivos | 65 |
| 5.9.7 Criação de modelos preditivos | 66 |
| 6 RESULTADOS | 70 |
| 6.1 Características descritivas da amostra | 71 |
| 6.2 Aspectos descritivos da Microbiota intestinal..... | 73 |
| 6.2.1 Ecossistemas microbianos (α diversidade) e macroestrutura (β diversidade) em indivíduos saudáveis e enfermos..... | 73 |

| | | |
|-----------|---|-----|
| 6.2.2 | Caracterização de filos da microbiota intestinal | 75 |
| 6.2.3 | Descrição táxons bacterianos em indivíduos saudáveis e enfermos | 77 |
| 6.3 | Características fenotípicas de indivíduos sadios e enfermos. | 78 |
| 6.4 | Modelos preditivos para identificação dos estados de saúde e doença . | 85 |
| 6.4.1 | Seleção de variáveis candidatas | 85 |
| 6.5 | Análise de performance preditiva dos modelos | 90 |
| 7 | DISCUSSÃO | 105 |
| 8 | CONCLUSÕES | 115 |
| | REFERÊNCIAS | 117 |
| APÊNDICES | | |

Apêndice 1. Número de leituras totais e filtradas de cada amostra

Apêndice 2. Parâmetros ecológicos de diversidade α – Índices Chao1, Shannon e Simpson.

Apêndice 3. Comparação de dados fenotípicos contendo: sociodemográficos, medicamentos de uso contínuo, estilo de vida e consumo dietético de nutrientes para pré-seleção de variáveis

Apêndice 4. Associação entre variáveis de fatores ambientais e estilo de vida e macroestrutura da microbiota intestinal

Apêndice 5. Boxplots dos táxons microbianos pré-selecionadas para compor os modelos preditivos

Apêndice 6. Bactérias pré-selecionadas para modelos preditivos

Apêndice 7. Curvas ROC e Índice Gini - capacidade preditiva das variáveis fenotípicas qualitativas

Apêndice 8. Curvas ROC e Índice Gini - capacidade preditiva das variáveis fenotípicas quantitativas

Apêndice 9. Curvas ROC - capacidade de modelos integrados com as variáveis fenotípicas e de microbiota

Apêndice 10. Análise paralela de associação entre variáveis fenotípicas e estrutura da microbiota intestinal

Apêndice 11. Características descritas sobre o microbioma intestinal associadas ao DM2

ANEXOS

Anexo A. Aprovação projeto principal

Anexo B. Desmembramento

Anexo C. TCLE

LISTA DE ABREVIATURAS E SIMBOLOS

| | |
|-------------------------|---|
| % | percentual |
| AGCC | Ácidos graxos de cadeia curta |
| AR | Grupo artrite reumatoide |
| AUC | Área sob a curva [<i>Area Under the Curve</i>] |
| CAPPesq | Comissão de Ética para Análise de Projetos de Pesquisa |
| cm | Centímetros |
| CT | Grupo controle |
| DM1 | Diabetes mellitus tipo 1 |
| DM2 | Diabetes mellitus tipo 2 |
| DII | Doenças inflamatórias intestinais |
| DC | Doença de Crohn |
| DP | DP Desvio padrão |
| HC-FMUSP | Hospital das Clínicas da Faculdade de Medicina da Universidade de São Paulo |
| IBP | Inibidores da bomba de prótons |
| IMC | Índice de Massa Corpórea |
| Kg | Quilogramas |
| LES | Lúpus eritematoso sistêmico |
| LIM | Laboratório de investigação médica |
| MI | Microbiota intestinal |
| mL | Mililitros |
| MSM | <i>Multiple source method</i> |
| NGS | Sequenciamento de Próxima Geração |
| DM2 | Diabetes tipo 2 |
| PCoA | Análise de coordenadas principais |
| PCR | Reação em cadeia polimerase |
| kg/m² | Quilo dividido pela altura em metros ao quadrado |
| PP | Grupo Psoríase em placas |
| R24 | Recordatório alimentar de 24 horas |
| RCU | Retocolite ulcerativa |
| Rho | Coefficiente de correlação de postos de Spearman |
| ROC | Característica de Operação do Receptor |
| TCLE | Termo de Consentimento Livre e Esclarecido |
| TGI | Trato gastrointestinal |
| ASV | Variante de sequência de amplicon |
| α | Alfa |
| β | Beta |

LISTA DE FIGURAS

| | |
|--|----|
| Figura 1. Principais funções da microbiota intestinal em homeostase | 27 |
| Figura 2. Fatores considerados como parte do expossoma. | 34 |
| Figura 3. Microbiota intestinal e doenças locais e sistêmicas..... | 35 |
| Figura 4 . Táxons bacterianos enriquecidos ou diminuídos no diabetes mellitus tipo 2 (DM2) em diversas populações mundiais..... | 45 |
| Figura 5. Diagrama consort fluxo dos participantes do estudo..... | 56 |
| Figura 6. Distribuição do índice global de Jaccard. | 64 |
| Figura 7. Fluxo da seleção de variáveis candidatas e aprendizado de máquina para criação de modelos preditivos..... | 68 |
| Figura 8. Macroestrutura da microbiota intestinal..... | 74 |
| Figura 9 Heatmap de seleção de variáveis fenotípicas – qualitativas | 87 |
| Figura 10 Heatmap de seleção de variáveis fenotípicas – quantitativas... | 88 |
| Figura 11 Heatmap de seleção de variáveis da microbiota intestinal..... | 89 |
| Figura 12. Curva ROC, considerando a microbiota intestinal isoladamente como preditor da doença de Crohn. | 91 |
| Figura 13. Valores índice Gini de diferentes táxons microbianos e seu impacto na previsão do status da doença de Crohn..... | 92 |
| Figura 14. Curva ROC, considerando a microbiota intestinal isoladamente como preditor da Retocolite Ulcerativa (RCU). | 93 |
| Figura 15. Valores índice Gini de diferentes táxons microbianos e seu impacto na previsão do status da Retocolite Ulcerativa..... | 93 |
| Figura 16. Curva ROC, considerando a microbiota intestinal isoladamente como preditor de Psoríase em Placas. | 94 |

| | |
|---|----|
| Figura 17. Valores índice Gini de diferentes táxons microbianos e seu impacto na previsão do status da Psoríase em Placas..... | 95 |
| Figura 18. Curva ROC, considerando a microbiota intestinal isoladamente como preditor de Artrite Reumatoide. | 95 |
| Figura 19. Valores índice Gini de diferentes táxons microbianos e seu impacto na previsão do status da Artrite Reumatoide..... | 96 |
| Figura 20. Curva ROC, considerando a microbiota intestinal isoladamente como preditor de Lupus Eritematoso Sistêmico. | 96 |
| Figura 21. Valores índice Gini de diferentes táxons microbianos e seu impacto na previsão do status da Lupus Eritematoso Sistêmico. | 97 |
| Figura 22. Curva ROC, considerando a microbiota intestinal isoladamente como preditor de Diabetes Tipo 1. | 97 |
| Figura 23. Valores índice Gini de diferentes táxons microbianos e seu impacto na previsão do status da Diabetes Tipo 1. | 98 |
| Figura 24. Curva ROC, considerando a microbiota intestinal isoladamente como preditor de Diabetes Tipo 2. | 98 |
| Figura 25. Valores índice Gini de diferentes táxons microbianos e seu impacto na previsão do status da Diabetes Tipo 2. | 99 |

LISTA DE GRÁFICOS

- Gráfico 1.** Indicador de diversidade alfa, índices Shannon, Simpson e Chao173
- Gráfico 2.** Abundância relativa dos 5 principais filos bacterianos..... 75
- Gráfico 3.** Táxons com abundância relativa significativamente diferente entre controles e doentes por análise par a par 77

LISTA DE QUADROS

| | |
|---|----|
| Quadro 1. Critérios de inclusão e exclusão do estudo VALIDYS. | 55 |
| Quadro 2. Classificação da capacidade preditiva das AUC | 69 |

LISTA DE TABELAS

| | |
|--|-----|
| Tabela 1. Alterações observadas no microbioma intestinal de indivíduos adultos com diabetes tipo 1 | 43 |
| Tabela 2. Características fenotípicas da população do estudo..... | 72 |
| Tabela 3. Análise de β diversidade entre grupos..... | 75 |
| Tabela 4. Características fenotípicas (fatores ambientais e variáveis de estilo de vida) de indivíduos sadios e enfermos..... | 80 |
| Tabela 5. Hábitos alimentares relacionados ao consumo de alimentos mais saudáveis, menos saudáveis e álcool..... | 82 |
| Tabela 6. Ingestão habitual de calorias, macro e micronutrientes..... | 83 |
| Tabela 7. Táxons que apresentaram abundância diferencial na comparação entre todos os grupos doentes e grupo controle..... | 85 |
| Tabela 8. Comparação do desempenho dos modelos preditivos de doença com microbiota isolada e integrada..... | 101 |
| Tabela 9. Compilado de pesquisas com aprendizado de máquina em estudos de microbioma intestinal na identificação de diversas doenças..... | 112 |

RESUMO

Candian DCF. Avaliação do potencial preditivo da microbiota intestinal associada a características fenotípicas para identificação de doenças multifatoriais. [Tese] São Paulo: Faculdade de Medicina, Universidade de São Paulo 2023.

Introdução: A microbiota intestinal (MI) é um complexo conjunto de microrganismos, tão singular que tem sido comparado com a “impressão digital” do indivíduo. Em consideração com as variadas funções fisiológicas e patofisiológicas da MI, pesquisadores têm estudado, nas últimas duas décadas, a sua participação em diversos estados de doença, tendo em vista oportunidades de previsão e tratamento das distintas enfermidades. Entretanto, a baixa concordância entre estes estudos faz com que a real participação da MI na predição de doenças multifatoriais permaneça pouco esclarecida. **Objetivo:** Estabelecer modelos preditivos que integrem variáveis fenotípicas com táxons da microbiota intestinal para distinguir indivíduos saudáveis e enfermos. **Metodologia:** Indivíduos saudáveis e enfermos (202), provenientes da casuística de um estudo maior (VALIDYS), foram avaliados quanto a variáveis fenotípicas e composição da MI por meio de sequenciamento do gene 16S rRNA. As informações sobre características fenotípicas foram coletadas por meio de história clínica, consumo alimentar (três registros alimentares de 24h) e composição corpórea (bioimpedância elétrica). Os resultados da MI foram expressos em variantes de sequência de amplicon (ASVs). Diferenças na abundância de táxons bacterianos entre os grupos foram avaliadas por DESeq2 para a pré-seleção de táxons para compor modelos preditivos. A MI também foi testada por comparação geral entre todos os grupos, para obter o maior número de táxons relevantes que pudessem compor os modelos. A criação de modelos preditivos deu-se, *a priori*, para cada conjunto de dados – quantitativos, qualitativos, e abundância de gêneros da microbiota. Toda a etapa de modelagem foi implementada em ambiente R com o algoritmo Random Forest. A performance de cada modelo foi avaliada pela sensibilidade e especificidade usando o conjunto de dados de teste. **Resultados:** 50 indivíduos saudáveis e 152 enfermos (portadores de doença de Crohn, retocolite ulcerativa, psoríase em placas, artrite reumatoide, lúpus eritematoso sistêmico, diabetes tipo 1 e tipo 2) compuseram a amostra. Diferenças das características fenotípicas foram identificadas entre os grupos e aquelas com relevância estatística foram pré-selecionadas para compor o modelo integrativo final. Diferenças ecológicas gerais da MI foram observadas apenas para o subgrupo de doença inflamatória intestinal. Foram selecionados 52 táxons bacterianos e todas as variáveis pré-selecionadas foram filtradas por Random Forest, aplicando 50 modelos, para identificar as 10 variáveis mais importantes em pelo menos 50% dos modelos. A avaliação da performance preditiva dos modelos deu-se pelas curvas ROC, criadas para mostrar a capacidade global dos modelos em preverem uma determinada doença. A análise do desempenho, entre os modelos para todas as doenças, apontou as melhores performances quando os dados microbianos e fenotípicos foram integrados. **Conclusão:** Nas condições da presente pesquisa conclui-se que existem marcadores taxonômicos pontuais que apresentaram abundância relativa diferente entre os grupos avaliados, porém isso não os torna, isoladamente, marcadores com potencial preditivo. O uso de reamostragens

(bootstrap) deve ser uma ferramenta utilizada para minimizar o efeito da grande variabilidade observada entre estudos de MI e favorecer a reprodutibilidade dos resultados obtidos. Coletivamente, nossos dados sugerem que a utilização de marcadores microbianos isoladamente é pouco capaz de prever um desfecho de saúde. A adição de características fenotípicas com os resultados da composição da MI em um modelo preditivo integrado aumenta a capacidade de previsão de desfechos de saúde.

Palavras - chave: Microbiota intestinal. 16S rRNA, modelos de predição. Árvores de Decisões . Enfermidades. Reprodutibilidade dos Resultados

ABSTRACT

Candian DCF. Evaluation of the predictive potential of the gut microbiota associated with phenotypic characteristics for the identification of multifactorial diseases. [Thesis]. São Paulo: “Faculdade de Medicina, Universidade de São Paulo” 2023.

Introduction: The gut microbiota (GM) is a complex set of microorganisms so unique that it has been compared to the “fingerprint” of the individual. In consideration of GM importance in physiological functions, researchers have studied its participation in several disease states in the last two decades with a view to opportunities for forecasting and treating different diseases. However, the low concordance between these studies means that the real participation of GM in the prediction of multifactorial diseases remains unclear. **Aim:** To establish predictive models that integrate phenotypic variables with taxa of the intestinal microbiota to distinguish healthy and sick individuals. **Methodology:** Healthy and sick individuals (202) from the sample of a larger study (VALIDYS) were evaluated for phenotypic variables and GM composition through sequencing of the 16S rRNA gene. Information on phenotypic characteristics was collected through clinical history, dietary intake (three 24-hour dietary records), and body composition (bioelectrical impedance). GM results were expressed as amplicon sequence variants (ASVs). Differences in the abundance of bacterial taxa between groups were evaluated by DESeq2 for the pre-selection of taxa to compose predictive models. GM was also tested by general comparison between all groups, to obtain the largest number of relevant taxa that could compose the models. The creation of predictive models took place, a priori, for each set of data – quantitative, qualitative, and abundance of genera of the microbiota. The entire modeling step was implemented in the R environment with the Random Forest algorithm. The performance of each model was evaluated for sensitivity and specificity using the test dataset. **Results:** 50 healthy and 152 unhealthy individuals (with Crohn's disease, ulcerative colitis, plaque psoriasis, rheumatoid arthritis, systemic lupus erythematosus, type 1 and type 2 diabetes) comprised the sample. Differences in phenotypic characteristics were identified between groups, and those with statistical significance were pre-selected to compose the final integrative model. Overall ecological differences in GM were observed only for the inflammatory bowel disease subgroup. Fifty-two bacterial taxa were selected, and all pre-selected variables were filtered by Random Forest, applying 50 models, in order to identify variables that were among the 10 most important in at least 50% of the models. The performance of predictive models was evaluated using ROC curves, created to show the global ability of the models to predict a given disease using a set of data. The performance analysis, among models for all diseases, showed the best performances when microbial and phenotypic data were integrated. **Conclusion:** Under the conditions of this research, it is

concluded that there are punctual taxonomic markers that showed different relative abundance between the evaluated groups, but this does not make them, in isolation, markers with predictive potential. The use of resampling (bootstrap) should be a tool used to minimize the effect of the great variability observed between GM studies and favor the reproducibility of the results obtained. Collectively, our data suggest that the use of microbial markers alone is poorly able to predict a health outcome. Adding phenotypic traits with GM composition results in an integrated predictive model that increases the predictability of health outcomes.

Keywords: Gut microbiota. 16S rRNA. prediction models. Decision Trees. Diseases. Reproducibility of Results.

1 INTRODUÇÃO

1 INTRODUÇÃO

O microbioma humano pode ser descrito como um enorme e complexo conjunto de microrganismos que habitam o corpo humano. Este grupo de microrganismos inclui arqueias, bactérias, fungos e vírus que residem em diferentes locais do organismo, incluindo pele, cavidade oral, trato respiratório, urinário e reprodutivo e o mais profundamente estudado, o trato gastrointestinal.¹ O microbioma refere-se à população residente do habitat estudado, mas também à coleção de elementos genômicos e seus subprodutos.^{1,2}

Embora pesquisas sobre microrganismos remontem do século XVII, foi no início dos anos 2000, após a finalização do Projeto Genoma Humano e com o aprimoramento de tecnologias de sequenciamento genético, que pesquisadores focaram sua atenção em procurar saber como as comunidades microbianas interagem com o organismo humano.² Assim, tiveram início os projetos Microbioma Humano (HMP) e Metagenômica do Trato Intestinal Humano (MetaHIT), que tiveram como objetivo pesquisar sobre genes e genomas microbianos denominados coletivamente como “microbioma”.^{3,4}

Diferente da composição genética dos seres humanos, que é extremamente semelhante, o microbioma humano e seu metagenoma (DNA dos microrganismos) é bastante variável e singular, e conta com apenas um terço de seus genes constituintes encontrados na maioria dos indivíduos saudáveis.⁴ O microbioma humano é complexo, abriga trilhões de microrganismos e coletivamente contém 100 vezes mais genes do que o genoma humano.^{1,4-8}

Cada indivíduo abriga em seu microbioma diferentes comunidades microbianas em cada local do corpo, que variam muito em termos de estrutura, composição e funções. Cada comunidade de um determinado local anatômico do organismo, por exemplo, pele ou trato gastrointestinal, é conhecido como “microbiota”.^{1,7,10} A diversidade composicional e funcional da microbiota de cada local do corpo varia dependendo de sua complexidade e as funções a serem exercidas naquele ambiente. Essa comunidade pode se adaptar a partir da

exposição a fatores externos, como fumo, estresse, exposição a pesticidas; alimentos consumidos e/ou fatores internos de um indivíduo, como deficiências nutricionais e alterações orgânicas (estado de doença).^{11,12}

A microbiota intestinal (MI) abriga aproximadamente 90% de todo o microbioma humano, sendo dominada principalmente pela população bacteriana que pode representar até 1000 diferentes espécies em um único indivíduo¹⁰. Fica assim evidente a complexidade desse ecossistema, que repercute na dificuldade em encontrar padrões específicos e homogêneos de microbiota intestinal.¹³⁻¹⁶

Nos últimos anos, análises de composição e função da microbiota intestinal têm sido realizadas em diferentes tipos de amostras, como biópsia de mucosa, fluido intestinal e amostras fecais, por ser uma forma não invasiva para avaliar a composição da microbiota do cólon distal.¹⁷

O desenvolvimento de várias novas disciplinas de conhecimento e tecnologias aplicadas, como meta-transcriptômica, metagenômica e ferramentas de bioinformática têm auxiliado na compreensão da contribuição dos diferentes microbiomas e microbiotas do ser humano e sua possível participação na saúde e doenças humanas.¹³

A microbiota intestinal tem sido referida como uma possível “segunda impressão digital”, por seu caráter único em cada indivíduo. Assim, identificar e compreender a composição e função de uma “microbiota intestinal saudável” tem sido grande desafio nas pesquisas com microbioma humano.¹⁴

Embora seja um campo desafiador, o interesse em estudos com microbiota intestinal se explica pelo fato da MI ter composição abrangente e variável, e exercer importante papel no processo de digestão e absorção, comunicação interórgãos (ex. eixo intestino-cérebro¹⁸, eixo intestino-músculo¹⁹), participação no amadurecimento e sinalização para o sistema imunológico, metabolismo de medicamentos entre outras importantes funções (**Figura 1**).^{20,21}

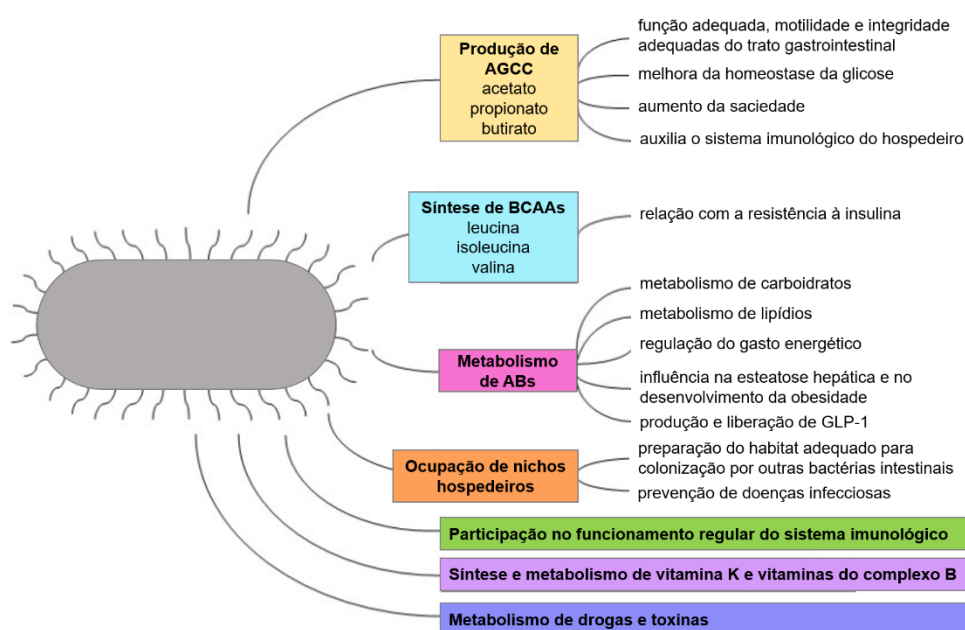


Figura 1. Principais funções da microbiota intestinal em homeostase

Adaptado e traduzido de: Laukens, D et al. e Bielka, W et al.^{20,21}

Considerando as inúmeras funções fisiológicas da MI, estudos sobre sua possível participação em estados de doença, frequentemente pautados em trabalhos experimentais e observacionais, vêm sendo realizados nas últimas duas décadas para melhor compreensão sobre a participação da MI em diversas condições clínicas. Entretanto, a real compreensão da participação da microbiota intestinal na predição de doenças multifatoriais ainda não foi completamente esclarecida.

O presente trabalho, a fim de auxiliar na construção do conhecimento, aborda a seguir, na sessão “revisão de literatura”, informações mais profundas sobre a comunicação bidirecional entre a microbiota intestinal, sua composição, e características fenotípicas em doenças multifatoriais de interesse. Nesta sessão também será abordado o atual conhecimento sobre a utilização da microbiota intestinal como biomarcador e seu valor como preditor em algumas enfermidades.

2 REVISÃO DA LITERATURA

2 REVISÃO DA LITERATURA

2.1 CARACTERÍSTICAS FENOTÍPICAS, MICROBIOTA INTESTINAL E SUA CONVERSA BIDIRECIONAL NOS ESTADOS DE SAÚDE E DOENÇA

As interações entre o microbioma intestinal e o metabolismo humano são complexas e vêm se consolidando nos últimos anos. São vários os fatores que podem influenciar favoravelmente o microbioma intestinal ou, ao mesmo tempo, serem fatores de risco e/ou consequências de enfermidades. Estes incluem genética²², idade²³, consumo dietético²⁴, uso de medicamentos²⁵, geografia²⁶, consumo de probióticos²⁷, exercícios²⁸, composição corpórea²⁹ e história clínica^{30,31}.

A microbiota intestinal é importante agente contribuinte para a variabilidade interindividual fisiológica e bioquímica nos seres humanos. A adaptabilidade da MI é um fator que, provavelmente, ajuda o organismo a manter o estado homeostático que difere de pessoa para pessoa e funciona como um sistema dinâmico em resposta a mudanças fisiológicas. A alta heterogeneidade dos perfis da microbiota intestinal entre os indivíduos impede até o momento uma definição prática de “microbiota intestinal saudável”; neste sentido existem poucas evidências para estabelecer um perfil filogenético central que diferencie populações distintas.^{23, 32, 33}

De modo geral, a MI adulta saudável participa de funções necessárias à homeostase metabólica e imunológica, mantém a estabilidade a longo prazo e resiste à colonização de patógenos. Perturbações no equilíbrio do ecossistema do microbioma intestinal que podem resultar na interrupção de tais funções tendem a ser definidos genericamente como disbiose.^{32,33}

No presente trabalho, consideramos que o uso do termo disbiose não é apropriado quando tratamos da determinação taxonômica da microbiota intestinal, visto que essa determinação não permite avaliar a função dos microrganismos intestinais, o que impede saber se alguma das funções essenciais previamente descritas foi interrompida^{34,35,36}. Ademais, em

decorrência da ausência de padrão de referência para uma “microbiota ideal”, bem como das múltiplas possibilidades de tratamento de dados, fica difícil caracterizar o termo “disbiose” de maneira única, tornando quase obrigatório a seleção de grupo controle para comparação com a doença ou condição avaliada em cada novo estudo executado. ³⁷⁻⁴⁰

Na era do microbioma, muitas pesquisas buscam utilizar as informações sobre os microrganismos que habitam o corpo humano como possíveis marcadores para estados de saúde e doença. No entanto, existe baixa concordância e reprodutibilidade entre estudos que examinam o papel da microbiota em doenças humanas, o que desafia a capacidade de identificar relações causais entre microrganismos do hospedeiro e enfermidades. ^{12, 32, 41}

O risco de obter falsos positivos ou encontrar correlações espúrias entre dados de microbiota isolados e algum estado específico de saúde é exacerbado pela heterogeneidade interindividual da composição microbiana, e deve-se frequentemente a limitações metodológicas e caracterização de “assinaturas de microbiota” baseadas exclusivamente em diferenças em abundância relativa de táxons microbianos, o que mantém ainda desconhecida a capacidade da microbiota intestinal humana em prever condições clínicas. ^{23, 37, 42, 43}

Assim, torna-se importante integrar outros aspectos fenotípicos do hospedeiro, como consumo dietético, prática de atividade física, tabagismo, características corpóreas que conversam simultaneamente com a microbiota, para compreender o papel de cada uma destas na saúde e na doença.

2.1.1 Dieta

Ao longo da vida, o tipo de dieta e a carga de nutrientes obtidos através da alimentação são frequentemente citados como fatores de grande impacto na composição da microbiota intestinal. Isto porque macro e micronutrientes, bem como grupos alimentares considerados mais ou menos saudáveis para a alimentação humana (ex. ultraprocessados ou vegetais), podem ser associados

a modificações da composição da microbiota intestinal, uma vez que bactérias da MI utilizam alguns nutrientes como substrato e/ou também metabolizam nutrientes, tornando-os mais biodisponíveis para o hospedeiro.^{13,44,45}

Pesquisas apontam que, embora dietas temporárias possam modificar a MI e sua composição, estas variações pontuais tendem a ser superadas também em curto prazo. São os hábitos alimentares de longo prazo que parecem responsáveis pelo perfil microbiano mais estável ao longo da vida.⁴⁴

O padrão da dieta ocidental tem composição variada de proteínas, carboidratos, gorduras e fibras, fatores que podem parecer influenciar a composição da microbiota intestinal. Dietas ricas em proteínas e gorduras, principalmente de origem animal, aparentam ser relacionadas com redução da diversidade e riqueza bacteriana, além da abundância maior de *Bacteroides* e *Bilophila*. Já dietas ricas em gorduras mono e poli-insaturadas têm maior associação com diversidade microbiana elevada e aumento nos membros da família *Lachnospiraceae*.⁴⁶

A dieta ocidental geralmente contém menor quantidade de fibras. As fibras, em particular as solúveis e prebióticas, são importante fonte de energia para microrganismos da microbiota intestinal, incluindo Bifidobactérias, consideradas, há décadas, indicadores de boa saúde intestinal. Fibras dietéticas são importantes substratos para reações de fermentação realizadas por espécies bacterianas específicas que possuem o repertório enzimático necessário para degradar esses tipos de carboidratos complexos.^{47,48}

O consumo de dietas pobres em fibras é geralmente associado com redução na diversidade microbiana, redução de espécies fermentadoras de fibras e, como consequência menor produção de ácidos graxos de cadeia curta (AGCC), um substrato energético importante para as células do cólon. Acetato, propionato e butirato são os principais AGCCs, com atuação sobre regulação gênica, proliferação de colonócitos e inflamação. Por outro lado, com ingestão de dietas ricas em fibras, observa-se adequada riqueza e diversidade microbiana com maior abundância de gêneros, como *Prevotella*.^{47,48}

No entanto, a dieta do ser humano não apenas impacta e é impactada pela microbiota intestinal, mas também pode ser considerada como fator de risco direto para condições clínicas do hospedeiro. Por exemplo, uma dieta muito rica em sódio é reconhecida como fator de risco para hipertensão arterial sistêmica. Já a dieta com grande concentração de gorduras saturadas e carboidratos simples, como acontece no padrão ocidental, é associada com doenças metabólicas, como diabetes tipo 2 e resistência à insulina. Neste sentido, estudos que avaliem condições de saúde e doença devem considerar a avaliação do consumo alimentar e dietético como um fator associado ao perfil microbiano intestinal.⁴³

2.1.2 Atividade física

A prática de exercício, ou sua ausência, é outra característica fenotípica que pode influenciar a estrutura da MI e distintas condições de doença. Dentre as interações entre MI e atividade física, por exemplo, a prática regular de exercícios é geralmente associada a redução da inflamação, que por sua vez se correlaciona com mudanças no perfil de citocinas.⁴⁹

O efeito do exercício sobre a MI tem sido foco de vários estudos transversais, intervencionistas e de revisões recentes. Os efeitos positivos se associam principalmente ao aumento da diversidade da microbiota e ao equilíbrio entre as comunidades bacterianas benéficas e enteropatogênicas. No entanto, diferentes resultados foram observados dependendo do nível de atividade física praticada no momento e o estado de atividade anterior do indivíduo avaliado (sedentário ou ativo). Fatores importantes, como peso e/ou IMC e dieta específica realizada pela população alvo também deve ser considerados.⁵⁰

Revisões sistemáticas recentes identificaram que o exercício físico pode modificar a abundância relativa da MI.⁵¹⁻⁵⁴ Alguns aspectos do exercício devem ser considerados, principalmente a intensidade (praticantes de atividades

extenuantes ou atletas profissionais), que diferencia o exercício físico “saudável” de um fator de estresse físico.

Assim como o consumo dietético, a atividade física contribui para prevenir e tratar várias doenças crônicas, que incluem, entre outras, doenças cardíacas, diabetes tipo 2, pressão alta, alguns tipos de câncer, depressão e demência.⁵⁵

2.1.3 Tabagismo

Outro importante fator de impacto para a saúde é o tabagismo. A ligação direta entre tabagismo e aumento de risco para alguns tipos de câncer está bem estabelecida, mas novos estudos têm explorado a influência do tabagismo sobre o microbioma oral, esofágico e gastrintestinal.⁵⁶⁻⁵⁸

A fumaça do cigarro inalada é uma mistura que inclui nicotina, aldeídos, metais pesados e outros compostos, mas parte pode ser engolida e direcionada para o trato gastrintestinal (TGI). Esses compostos tóxicos podem diminuir os antioxidantes endógenos do organismo, aumentar o estresse oxidativo e as concentrações de compostos pró-inflamatórios do hospedeiro. No TGI, estes compostos parecem levar a modificações da composição da MI por meio de diferentes mecanismos, como a alteração da atividade antimicrobiana e regulação do microambiente intestinal.⁵⁹⁻⁶¹

Estudos em animais e humanos indicam que fumar parece afetar a composição microbiana e a inflamação intestinal. Em animais, a exposição à fumaça do tabaco pode aumentar a abundância de *Clostridium clostridiforme* e diminuir *Lactococcus*, *Ruminococcus albus* e a família de *Enterobacteriaceae*. Essas mudanças foram associadas a mudanças nas proteínas de junção apertada (*tight junctions*) e sinalização inflamatória intestinal. Em humanos, o tabagismo se associa à diminuição de diversidade microbiana, redução de Firmicutes e Actinobacteria, além de aumento de *Bacteroidetes* e *Proteobacteria* no intestino.⁵⁷⁻⁶⁰

O mecanismo exato de como o fumo afeta a MI, ainda em pesquisa, pode estar ligado à alterações das junções intestinais, sinalização imunológica e / ou produção de mucina.⁶⁰ Considerando os efeitos deletérios do uso de tabaco sobre a saúde, o tabagismo deve ser considerado também um fator de controle em estudos de MI.

A análise baseada em coortes humanas sugere que a natureza dinâmica do microambiente intestinal reflete a complexa interação do hospedeiro com suas características fenotípicas, como estilo de vida, dieta, fatores ecológicos e outros, chamado por vezes de expossoma (**Figura 2**).⁶²

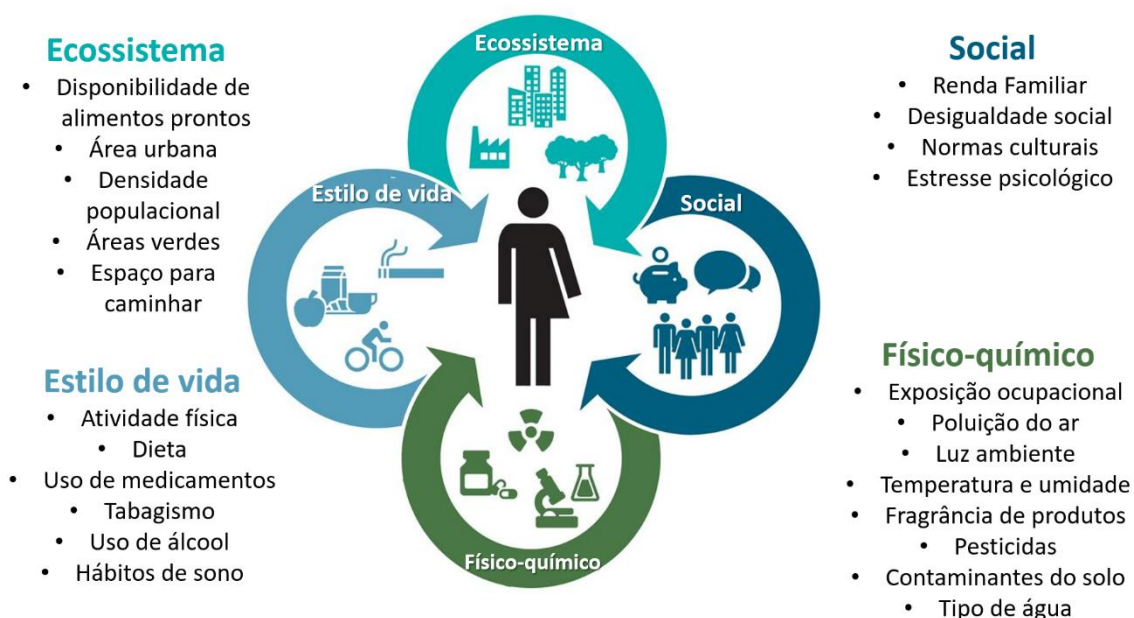


Figura 2. Fatores considerados como parte do expossoma. (adaptado de Vermeulen et al. 2020)⁶².

2.2 COMPOSIÇÃO DA MICROBIOTA INTESTINAL NAS DOENÇAS

Nas últimas duas décadas, a microbiota intestinal e sua composição, têm sido frequentemente associadas a doenças intestinais e sistêmicas (**Figura 3**).^{39,61}

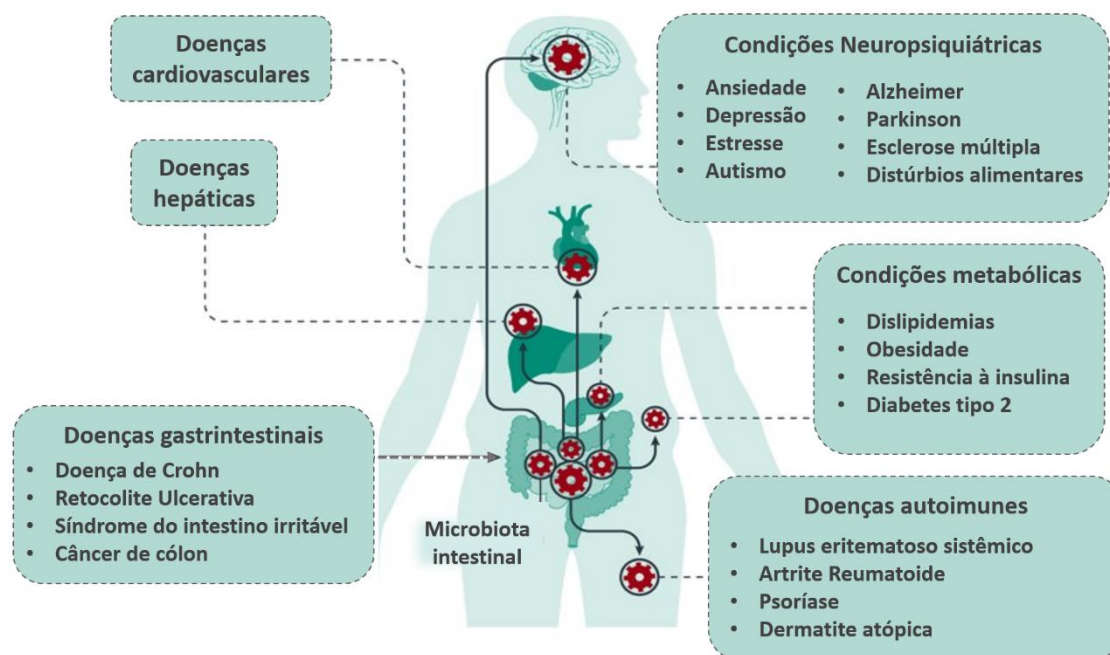


Figura 3. Microbiota intestinal e doenças locais e sistêmicas. Adaptado de Cani P, et al.³⁹ e Gui X, et al.⁶¹

Um grande desafio para a identificação de perfis de microbioma intestinal em condições específicas vem da ampla diversidade interpessoal mesmo entre indivíduos saudáveis, que dificulta o entendimento dos constituintes microbianos no estado de doença.⁶³

Com o avanço das tecnologias de sequenciamento de nova geração, a análise do microbioma intestinal foi aplicada a inúmeras coortes humanas. Descobriu-se que poucas bactérias intestinais são compartilhadas pela maioria dos indivíduos, e que espécies diferentes podem apresentar funções redundantes.³⁶ Por exemplo, em 4.000 amostras fecais de indivíduos europeus, apenas 17 gêneros bacterianos foram identificados como parte do microbioma central, presente em mais de 95% de todas as amostras, indicando que a maioria das bactérias na MI são raras, ou seja, pouco frequentes na maioria dos indivíduos. Por este motivo, os dados sobre assinaturas de microbiota associados a doenças, por vezes, apresentam dados conflitantes e merecem atenção para seleção correta de potenciais marcadores específicos e reprodutíveis.^{63,64}

Nos tópicos abaixo, detalhes sobre cada doença pertinente ao presente trabalho são descritos.

2.2.1 Microbiota intestinal nas Doenças Inflamatórias Intestinais (DII)

A doença de Crohn (DC) e a retocolite ulcerativa (RCU) são subtipos de doenças inflamatórias intestinais (DII), definidos como distúrbios crônicos de característica recidivante, relacionados ao sistema imunológico do trato gastrointestinal que, até o momento, não possuem etiologia totalmente esclarecida. Além de fatores imunológicos do TGI, outros fatores intrínsecos como genética, e extrínsecos, como fatores ambientais, podem contribuir para a fisiopatologia⁶⁵. Entre os fatores intrínsecos do organismo humano, a microbiota intestinal tem sido considerada possível participante da patogênese, remissão e recidiva das DIIs, e assim, têm sido alvo comum de estudos da composição da microbiota intestinal⁶⁶.

Ao pesquisar sobre microbiota intestinal em DIIs, é possível encontrar referências da área que abordam as enfermidades como um conjunto, identificando sua composição e alterações gerais ou em DC e RCU separadamente⁶⁷⁻⁷¹, além de momentos de atividade ou remissão clínica da doença.

As alterações gerais da microbiota intestinal mais comumente relatadas em DII, tem sido em relação a medidas ecológicas, como menor riqueza e alfa diversidade.^{66,72} Em relação aos níveis taxonômicos, como filos, gêneros e espécies, é possível observar uma ampla lista de modificações, como, por exemplo, alteração na proporção dos filos *Firmicutes* e *Bacteroidetes*, em consequência ao aumento do filo *Proteobacteria*. Aumento de *Gammaproteobacterias*, presença de *Fusobacteria*, redução de *Clostridia*, *Bacteroides*, *Suterella*, *Ruminococcaceae*, *Bifidobacterium*, *Lactobacillus* e *Eubacterium*. Observa-se também menor abundância de *bactérias produtoras de butirato*, como *Fecalibacterium prausnitzii*, e maior percentual de espécies

como *Escherichia coli*, *Ruminococcus gnavus* e *Eggerthella lenta* quando se compara com a MI de indivíduos considerados saudáveis.^{66,73}

A doença de Crohn, em particular, está associada a uma composição microbiana intestinal mais alterada e instável do que a RCU. Em DC, diferentes estudos apontam alguns possíveis perfis microbianos, Gevers et al. verificou maior abundância de Pasturellaceae, Veillonellaceae, Neisseriaceae, espécies de *Escherichia coli*; e diminuição na abundância de *Bacteroides*, Clostridiales, espécies de *Faecalibacterium*, *Roseburia*, *Blautia*, *Ruminococcus* e Lachnospiraceae em pacientes com DC recém-diagnosticados e virgens de tratamento.⁷⁴

Em pacientes brasileiros com DC em remissão clínica, Magro et al. sugeriu que, mesmo em remissão clínica, a microbiota intestinal quando comparados a controles saudáveis, ainda exibam menor α diversidade microbiana e maior abundância do filo Proteobacteria. Houve também redução nos gêneros Akkermansia e Oscillospira.⁷⁵ Por sua vez, Pascal et al., verificou perda de organismos produtores de butirato, sendo eles espécies de *Faecalibacterium*, *Peptostreptococcaceae*, *Anaerostipes*, *Methanobrevibacter*, Christensenellaceae e *Collinsella*. Também se identificou maior abundância de espécies de *Fusobacterium* e *Escherichia*. Testes de abundância diferencial entre táxons bacterianos de pacientes com DC ou RCU também foram observados no presente estudo.^{76,77}

Em RCU, Nemoto et al, verificou menor proporção de *Bacteroides* e *Clostridium XIV* em relação a indivíduos saudáveis.⁷⁸ Em outro estudo, pacientes com RCU tiveram menos bactérias pertencentes ao conjunto *Clostridium XIVa* e um número significativamente aumentado de Bacteroidetes, dado não observado comumente em DC.⁷⁹ Ademais, assim como em outras áreas de estudo da microbiota, estudos com RCU mostram alguns achados conflitantes, como o aumento⁸⁰ ou diminuição⁸¹ de Bifidobactérias, aumento⁸² ou diminuição⁸³ de *Bacteroides*. Verificou-se também que a alfa diversidade, um dos indicadores mais consistentes da relação entre MI e DIIs, pode ou não estar reduzida, o que

ressalta a importância de testar o potencial das bactérias da MI como possíveis marcadores para as enfermidades avaliadas.

Dentre as enfermidades atualmente estudadas em relação a MI, as DII são consideradas como uma das mais bem caracterizadas e com informações mais frequentemente convergentes.

2.2.2 Microbiota intestinal na Psoríase em Placas

A psoríase é caracterizada como doença inflamatória crônica da pele. Sua patogênese é complexa e ainda não foi totalmente estabelecida. No entanto, sabe-se que envolve uma grande suscetibilidade hereditária e resposta autoimune exacerbada ^{84 ref.} Grande parte dos casos de psoríase inclui a psoríase em placa (PP) ou *psoríase vulgar*. Os sintomas típicos dessa forma da doença se manifestam em placas vermelhas na pele, bem definidas e pruriginosas, decorrentes do crescimento descontrolado e diferenciação defeituosa de queratinócitos. ^{84,85}

O estudo da microbiota intestinal tem ganhado interesse científico em doenças extra intestinais devido a sua íntima comunicação com múltiplos órgãos. A literatura recente vem investigando sua participação em doenças crônicas, principalmente aquelas que envolvam o sistema imunológico, como a psoríase. ⁸⁶.

Os estudos de MI e psoríase, ao identificar assinaturas bacterianas intestinais associadas a doença e buscam, se possível, modular a sua composição como abordagem terapêutica. ⁸⁷⁻⁹⁰ Alterações da composição da MI em indivíduos com psoríase comparados a sujeitos saudáveis apontam grande heterogeneidade entre os achados. Huang L. e colaboradores encontraram maior abundância ⁹¹ do filo Bacteroidetes na microbiota de pacientes portadores de PP, enquanto outros estudos encontraram redução ⁹² de sua proporção. Outros estudos reportaram alterações na abundância de espécies específicas, como *Akkermansia muciniphila* ⁹³ e *Faecalibacterium prausnitzii* ⁹⁴, descritas

como reduzidas em pacientes com PP e associadas frequentemente a quadros de inflamação crônica. De forma controversa, Codoñer e colaboradores encontraram como característica de sua coorte, aumento do gênero *Akkermansia*⁸⁹. O mesmo cenário foi observado em relação aos níveis de *Prevotella* spp., que foram superiores⁸⁸ ou inferiores⁹² entre indivíduos com psoríase, em comparação com controles saudáveis. Em estudo caso-controle na população brasileira, observou-se aumento no gênero *Dialister* e na espécie *Prevotella copri* em pacientes com psoríase. Nestes pacientes houve também redução nos gêneros *Ruminococcus*, *Lachnospira* e *Blautia*, e na espécie *Akkermansia muciniphila*, quando comparados ao grupo controle.⁹⁵

Diante das observações discordantes entre estudos, ainda é difícil saber qual tipo bacteriano pode ser considerado o real perfil de microbiota em PP e se os táxons acima citados são capazes de auxiliar no diagnóstico dessa doença.

2.2.3 Microbiota intestinal na Artrite Reumatoide

Assim como a psoríase, a Artrite Reumatoide (AR) é uma condição autoimune crônica de característica inflamatória. AR é altamente debilitante e afeta entre 0,5% e 1% da população mundial entre jovens e adultos,⁹⁶ além de impactar de forma negativa a qualidade de vida dos enfermos.⁹⁷ Artrite reumatoide tem componente genético substancial⁹⁸, mas fatores de risco ambientais e de estilo de vida como: dieta, uso de medicamentos e exposição a infecções e tabagismo parecem ter relação com o início da doença em indivíduos geneticamente suscetíveis⁹⁹. No entanto, ainda não há consenso sobre os exatos mecanismos da doença, e nos últimos anos, considera-se a MI fator de risco adicional para ocorrência da AR^{98,99}.

Hipóteses consideram a participação dos microrganismos residentes da microbiota intestinal, que habitam a mucosa do trato gastrintestinal (TGI), como agentes contribuintes para a modulação da imunidade do hospedeiro com AR.¹⁰⁰ A eficácia clínica de medicamentos antibacterianos, como minociclina ou

salazosulfapiridina, em pacientes com AR favorece a teoria de que as populações microbianas do intestino podem estar associadas à AR ¹⁰¹

No mesmo sentido, experimentalmente, a abundância da espécie bacteriana *Prevotella copri* está aumentada em casos de AR precoce ¹⁰² e pode contribuir para o desenvolvimento de AR dependente de células Th17. Outra espécie de *Prevotella*, a *Prevotella histicola*, parece ser capaz de suprimir o desenvolvimento de artrite através da produção aumentada de butirato ¹⁰³. Outro achado comum na AR é o aumento de *Bacillus* e *Lactobacillus* ¹⁰⁴. Zhang e colaboradores detectaram aumento de *Lactobacillus salivarius* em amostras fecais de pacientes com AR e sua abundância relacionada à gravidade da doença.¹⁰⁵

Curiosamente, *L. salivarius*, uma bactéria probiótica, administrada oralmente em animais com AR levou a diminuição das erosões ósseas e infiltração sinovial, além de aumentar os níveis circulantes de IL-10 anti-inflamatórias ¹⁰⁶. Assim, devemos interpretar o papel da MI na AR com cautela. Contudo, mesmo sendo heterogêneos os achados disponíveis, torna-se de interesse investigar o papel de bactérias intestinais como potencial biomarcador diagnóstico, de gravidade ou de resposta terapêutica.¹⁰⁷

2.2.4 Microbiota intestinal no Lupus Eritematoso Sistêmico

Descrito como doença autoimune inflamatória complexa, mais prevalente em mulheres, o LES se apresenta com anormalidades das células B, autoanticorpos circulantes para antígenos nucleares e formação de imunocomplexos.¹⁰⁸ A heterogeneidade da manifestação do LES ocorre devido à variabilidade da atividade da doença que pode ser representada pela remissão (estabilidade da doença e ausência de suas manifestações) até as exacerbações, progressão e acometimento de múltiplos órgãos. Essa grande variedade de fenótipos no LES contribui para os desafios clínicos para o diagnóstico e tratamento eficaz.¹⁰⁹

De maneira similar a doenças autoimunes, como psoríase e artrite reumatoide, estudos sobre lúpus eritematoso sistêmico (LES) têm considerado a possível participação da MI na sua patogênese ou capacidade como biomarcador para a enfermidade ¹¹⁰.

Em indivíduos com LES ativo sem medicação, diferenças em nível de filo e gênero foram observadas. Em comparação com controles saudáveis, os pacientes com LES tiveram maiores abundâncias relativas do filo Bacteroidetes, Actinobacteria e Proteobacteria. ¹¹⁰ Em nível de gênero *Rhodococcus*, *Eggerthella*, *Klebsiella*, *Prevotella*, *Eubacterium*, *Flavonifractor* e *Incertae sedis* foram enriquecidos, enquanto os gêneros *Dialister* e *Pseudobutyrvibrio* foram depletados em pacientes com LES.¹¹⁰

Por outro lado, pacientes LES com alta e baixa atividade de doença, com e sem acometimento renal apresentaram alterações diferentes. Por exemplo, existe superabundância média de 5 vezes mais *Ruminococcus gnavus* em pacientes LES com nefrite lúpica em comparação com os controles. Os pacientes com nefrite, quando comparados àqueles sem nefrite, também apresentaram maiores níveis do gênero *Veillonella*.¹¹¹ O perfil microbiano, ou a identificação de marcadores microbianos exclusivamente encontrados na fase inativa da doença, comparados a indivíduos saudáveis são escassos, e poderiam ser importantes para

a compreensão se a manutenção do paciente sem atividade de doença normalizasse a composição aberrante da microbiota intestinal.

Recentemente, Toumi e colaboradores¹¹² avaliaram táxons abundantes entre pacientes com LES, ativo e inativo entre si, e combinados em comparação a indivíduos saudáveis usando DESeq2¹¹³ para identificar biomarcadores candidatos. Ao nível dos filos, verificou-se diminuição de *Tenericutes* em pacientes com LES. Os gêneros *Alistipes*, *Flintibacter* e *Parabacteroides* foram significativamente mais abundantes em pacientes com LES. Mais uma vez, é possível notar dois aspectos interessantes, o primeiro é que bactérias “comuns” ou comensais podem estar aumentadas em múltiplas enfermidades de uma só vez, o que pode ser um fator limitante para sua utilização como marcador simples, pensando no sentido de aumento ou redução. O segundo, é observar a grande variância de resultados em relação a microbiota no LES torna de grande interesse estudos que favoreçam a compreensão de sua utilidade como marcador da doença.^{107,112}

2.2.5 Microbiota intestinal no Diabetes tipo 1 (DM 1)

Diabetes tipo 1 (DM1), doença autoimune, é caracterizado pela deficiência de insulina decorrente da autodestruição das células β pancreáticas e consequente hiperglicemia. As taxas de incidência de DM1 vêm aumentando em todo o mundo nas últimas décadas.^{114,115} A patogênese de DM1 envolve principalmente a predisposição genética¹¹⁶, mas fatores ambientais¹¹⁷ são apontados entre umas das principais causas do aumento da prevalência da enfermidade.

Nas últimas décadas pesquisadores têm sugerido possível papel da microbiota intestinal como gatilho da autoimunidade do DM1¹¹⁸ Grande parcela da literatura de pacientes com DM1 concentra-se em avaliar a microbiota intestinal na infância. No entanto, pouco se sabe sobre a composição e função do microbioma intestinal em adultos enfermos.

Os dados mais frequentes sobre a MI de adultos relatam redução na diversidade da MI em pacientes com DM1. Algumas características mais específicas variam de acordo com a população estudada, sendo encontradas mudanças heterogêneas como diminuição da abundância de Clostridium e Prevotella ¹¹⁹, aumento de diversas espécies de Bacteroides e Prevotella copri, redução de Bifidobacterium, Faecalibacterium¹²⁰ e Roseburia e, em alguns casos, nenhuma alteração na composição da MI (**Tabela 1**).

Tabela 1. Alterações observadas no microbioma intestinal de indivíduos adultos com diabetes tipo 1

| AUTOR - ANO PAÍS | N= CASOS / CONTROLES | PRINCIPAIS ACHADOS |
|--|-------------------------|--|
| SHILO 2022 ¹²¹ ISRAEL | 74 / 296 | ↑ Prevotella copri; ↑ Eubacterium siraenum; ↑ Alistipes inops ↓ Firmicutes, ↓ Alistipes putredinis, ↓ Faecalibacterium prausnitzii ↓ Ruminococcus gnavus |
| HIGUCHI - 2018 ¹²² BRASIL | 20 / 28 | ↑ Bacteroides vulgatus; ↑ Bacteroides rodentium ↑ Bacteroides xylanisolvens ↑ Prevotella copri; |
| HUANG - 2018 ¹²³ CHINA | 12 / 10 | ↓ Bifidobacterium; ↓ Roseburia ↑ Razão Bacteroidetes/ Firmicutes ↓ Faecalibacterium |
| DE GROOT - 2017 ¹²⁴ HOLANDA | 53 / 50 | ↑ Bacteroidales; ↑ Christensenella; ↑ Bifidobacterium; ↑ Ruminococcus; ↑ Christensenellaceae; ↑ Chloroplast ↓ Subdoligranulum; ↓ Streptococcus; ↓ Barnesiella; ↓ Haemophilus; ↓ Roseburia; ↓ Rhodospirillales |
| STEWART- 2017 ¹²⁵ REINO UNIDO | 10 / 10 | ↔ |

Legenda: ↔ Nenhuma diferença na composição microbiana; ↑ aumento de um táxon microbiano ↓ redução de um táxon microbiano

Conhecer a composição da microbiota e sua capacidade de prever a presença do estado de doença em DM1 permanece como um campo a ser explorado.

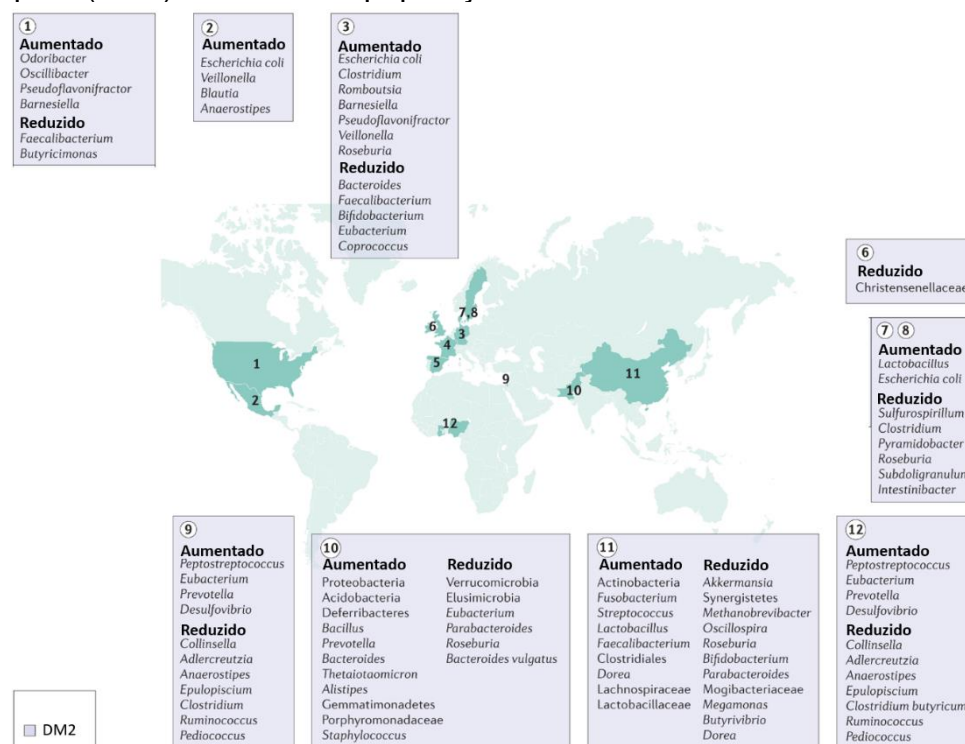
2.2.6 Microbiota intestinal no Diabetes tipo 2 (DM2)

Além do envolvimento da MI com o sistema imunológico, os microrganismos que habitam o intestino humano desempenham importantes funções metabólicas no organismo humano ¹⁹.

De fato, os estudos de microbiota baseados em sequenciamento ganharam ampla visibilidade no campo dos distúrbios metabólicos, como diabetes mellitus tipo 2 (DM2) e obesidade em diversos aspectos.^{25, 126-129} Mas, apesar dos extensos esforços para definir o papel da MI nessa área, a reprodutibilidade tem sido bastante limitada e a especificidade dos táxons bacterianos associados ao DM2 e obesidade tem sobreposição àqueles destacados para outras doenças como *Christensenellaceae*, *Collinsella* e *Escherichia coli*, que também são marcadores candidatos na doença de Crohn.
77,130

A heterogeneidade dos resultados em DM2 (**apêndice 11**), é tão grande ou maior quanto a encontrada nas enfermidades previamente citadas. O tema foi alvo de múltiplas pesquisas, incluindo revisões¹²⁶, como a recentemente publicada na revista Nature Reviews Gastroenterology and Hepatology⁷², que teve como objetivo avaliar o microbioma intestinal como biomarcador para doenças inflamatórias e metabólicas. Nesse estudo, as bactérias postuladas como candidatas a biomarcadores foram diferentes em várias localidades do mundo (**Figura 4**).

Figura 4 . Táxons bacterianos enriquecidos ou diminuídos no diabetes mellitus tipo 2 (DM2) em diversas populações mundiais



Adaptado de: Metwaly, A. e colaboradores (2022) ⁷²

O DM2, caracterizado por resistência insulínica, aumento da glicemia plasmática e inflamação crônica de baixo grau, pode impactar múltiplos órgãos, incluindo fígado, tecido adiposo, músculos e intestino. Sua origem é multifatorial e sua etiopatogenia envolve interação complexa entre suscetibilidade genética e fatores ambientais relacionados ao estilo de vida.^{131, 132}

Neste contexto, os pacientes portadores de DM2 podem apresentar múltiplas características fenotípicas diferentes, o que pode justificar as variações dos marcadores microbianos encontrados nos estudos em diferentes populações.^{72,126}

Todo o exposto até aqui, sugere que, para além da composição da microbiota intestinal, as características fenotípicas individuais podem ser importantes integrantes para melhorar a acurácia de modelos preditivos de doenças. Uma questão-chave para o aprimoramento de modelos preditivos é

identificar quais propriedades do microbioma e das características fenotípicas são candidatos mais valiosos para integrar modelos de estudos da área.

2.3 MICROBIOTA INTESTINAL COMO POTENCIAL BIOMARCADOR DE DOENÇAS

O conhecimento sobre o campo da microbiota intestinal avança rapidamente e suas importantes funções fisiológicas vêm sendo melhor compreendidas nas últimas décadas.¹³³

À medida que os pesquisadores descobrem mais detalhes sobre a composição das bactérias que habitam o intestino humano, mais questões sobre como esses microrganismos podem desempenhar um papel na saúde ou na doença do hospedeiro são suscitadas. Centenas de estudos clínicos demonstraram associações entre o microbioma humano e doenças, mas muitas questões sobre como podemos generalizar esse conhecimento permanecem sem respostas.¹³⁴

Ao mesmo tempo, tem-se observado uma busca pela aplicação translacional desse conhecimento para a prática clínica e/ou desenvolvimento de testes que forneçam a indicação da saúde de uma pessoa com base em análises específicas da microbiota intestinal.¹³⁵⁻¹³⁸ Em texto público, o pesquisador Francisco Guarner descreve o uso de testes de microbioma na prática clínica da seguinte maneira: “os testes de microbioma existentes hoje usam técnicas variadas e focam na taxonomia. A maioria deles fornece pouca informação para fins de diagnóstico ou prognóstico”. Nesse sentido, conhecer o real potencial da microbiota intestinal como biomarcador diagnóstico ou prognóstico de uma determinada condição de saúde parece importante para o momento de pesquisa atual.¹³⁹

Metwaly e seus colaboradores, em revisão recente, discutem o papel clínico dos biomarcadores do microbioma.⁷² Os autores ressaltam que um

biomarcador valioso deve contribuir com um poder de classificação adicional ao derivado da informação clínica, como é o caso da calprotectina fecal para detecção de inflamação intestinal. Neste caso, trabalhos diversos puderam confirmar com alta sensibilidade (entre 70 e 100%) e especificidade de moderada a alta (entre 44 e 100%), a capacidade da calprotectina fecal em detectar inflamação endoscópica.^{140,141}

Para o campo da microbiota intestinal, nos últimos cinco anos, aumentou o interesse por conhecer não apenas a composição da MI, em uma determinada enfermidade, mas também sua reprodutibilidade e capacidade em prever um determinado desfecho.^{142,143} O uso de algoritmos de aprendizado de máquina (AM) melhorou acentuadamente a precisão de modelos de previsão. Um exemplo foi o estudo realizado por Vila et al, que avaliou fezes de 1.792 indivíduos para distinguir pacientes com DII e SII (síndrome do intestino irritável), baseado no perfil da microbiota intestinal (sozinho ou associado a calprotectina).¹⁴² Com o uso de AM, os modelos de previsão da doença utilizando marcadores da MI (com 20 táxons principais) apresentaram área sob a curva (AUC) de 0,91, em comparação com os níveis fecais de calprotectina sozinhos AUC = 0,80. Esses resultados sugerem que a integração de outras características do hospedeiro, associadas com as características microbianas melhoram a precisão diagnóstica.

O uso da MI como biomarcador em doenças intestinais vem se estabelecendo, como nas DIIs e no câncer de cólon^{144,145}, mas não se limita a condições que acometem intestino. Estudos robustos em condições metabólicas também testaram combinações de biomarcadores tradicionais e microbianos. Salientam-se como promissores os trabalhos que visaram identificar a presença de cirrose na doença hepática gordurosa não alcoólica (DHGNA)¹⁴⁶, e que buscaram compreender o papel preditivo de marcadores microbianos na progressão de DM2¹⁴⁷ e obtiveram resultados promissores.

Para compreender esse papel preditivo, utilizou-se modelo de aprendizado de máquina baseado em microbioma para distinguir indivíduos com menor ou maior grau de resistência à insulina. Obteve-se AUC de 0,78, o que

sugere que a MI pode ser importante identificador da progressão do DM2.¹⁴⁷ Já a análise realizada por Sze e Schloss demonstrou limitada a capacidade de classificar com segurança indivíduos como obesos apenas com base na composição de seu microbioma.¹⁴⁸

Vale lembrar que um biomarcador traduz uma observação biológica que substitui e/ou prediz um ponto final clinicamente relevante, e que bons biomarcadores devem ser mensuráveis com pouca ou nenhuma variabilidade, e devem mudar de forma confiável em resposta a alterações na condição avaliada.¹⁴⁹ De tal modo, a identificação isolada de táxons microbianos com diferenças em sua abundância relativa não caracteriza um biomarcador.⁴² Esses táxons podem ser considerados candidatos, que devem ser testados quanto a sua utilidade através de modelos preditivos robustos que contemplem níveis de especificidade (valor que indica a proporção de indivíduos que genuinamente não têm a doença avaliada e recebem resultados de testes negativos) e sensibilidade (proporção de indivíduos que efetivamente têm a doença e recebem resultados de testes positivos).⁷²

Há 5 anos, o pesquisador Patrice D. Cani, comentou que até o momento não temos conhecimento completo sobre a composição da MI, e essa afirmação permanece verdadeira. Novas espécies e cepas microbianas podem ser isoladas do intestino humano, algumas com papel central no core microbiano, outras simples componentes do conjunto e identificadas recentemente.³⁹

Utilizar biomarcadores candidatos da microbiota para desenvolver diagnósticos aparenta ser hoje uma tarefa difícil, porém chegamos ao momento em que se faz necessário mostrar que as diferenças na microbiota podem, ou não, ser usadas para prever ou melhorar doenças, e não apenas mostrar que alguns táxons específicos são diferentes entre dois grupos de indivíduos, e como Olesen e Alm sugerem em seu artigo: “disbiose não é resposta”!¹⁵⁰

3 JUSTIFICATIVA

3 JUSTIFICATIVA

O perfil composicional qualitativo e quantitativo da microbiota intestinal tem sido frequentemente sugerido como potencial indicador dos estados de saúde ou doença.⁷² No entanto, dada a grande diversidade intra e interindividual existente na microbiota intestinal humana, a utilização isolada de bactérias intestinais como marcadores para predição de estados de saúde e doença permanece incerta.²¹

Diversos trabalhos têm denominado bactérias que apresentam abundância relativa diferentes entre grupo controle e grupo enfermo, como “potenciais marcadores” para os estados doença. No entanto, poucos desses estudos corroboram entre si, o que dificultaria o emprego dessas bactérias como marcadores.^{39, 61,69, 71,72, 77,98,99, 110,121,130,151}

Trabalhos mais recentes envolvendo análises robustas com aprendizado de máquina sugerem que, isoladamente, a microbiota intestinal apresenta baixa capacidade de diferenciar sadios de enfermos¹⁴⁸. Neste sentido, o presente estudo justifica-se pelo interesse em criar novos modelos preditivos que integrem características microbianas e características fenotípicas de indivíduos sadios e enfermos a fim de auxiliar na compreensão da real participação da microbiota intestinal na predição de estados de saúde e doenças multifatoriais.

4 OBJETIVOS

4 OBJETIVOS

4.1 Objetivo principal

- Estabelecer modelos preditivos que integrem variáveis fenotípicas com táxons da microbiota intestinal para distinguir indivíduos saudáveis e enfermos.

4.2 Objetivos específicos

- Descrever a composição geral da microbiota intestinal em uma população de indivíduos saudáveis e enfermos
- Verificar a existência de padrões de composição de microbiota intestinal relacionados aos estados de saúde e doenças
- Identificar as variáveis fenotípicas qualitativas e quantitativas (de composição corporal, dados antropométricos, atividade física, tipo de parto, amamentação e consumo alimentar) pertinentes para compor os modelos preditivos
- Identificar as variáveis microbianas pertinentes para compor os modelos preditivos
- Avaliar a performance dos modelos preditivos que integrem variáveis fenotípicas com a microbiota intestinal para identificação de estados de saúde e doenças.

5 CASUÍSTICA E MÉTODOS

5 CASUÍSTICA E MÉTODOS

5.1 LOCAL DE EXECUÇÃO DO ESTUDO

O presente estudo foi conduzido no Laboratório de Nutrição e Cirurgia Metabólica do Aparelho Digestivo (LIM-35) do Hospital das Clínicas da Faculdade de Medicina da Universidade de São Paulo (HC-FMUSP), em parceria com a Unidade de Diabetes do Serviço de Endocrinologia e Metabologia do HC-FMUSP (Diabetes tipo 1 e tipo 2), Unidade de Doenças Reumáticas (Artrite Reumatoide e Lupus Eritematoso Sistêmico), Unidade de Psoríase do Serviço de Dermatologia do HC-FMUSP (Psoríase em Placas) e com as Unidades de Cirurgia Bariátrica e Metabólica, Doenças Inflamatórias Intestinais da Divisão de Gastroenterologia e Hepatologia e da Divisão de Coloproctologia do Hospital das Clínicas da Faculdade de Medicina da Universidade de São Paulo (HC-FMUSP).

5.2 ASPECTOS ÉTICOS

A presente pesquisa constituiu subprojeto de estudo maior intitulado “Estudo VALIDYS: Validação de uma ferramenta para rastreamento do risco de disbiose e caracterização da microbiota intestinal sadia e enferma” aprovado pelo comitê de ética sob o nº 3.008.966 (**Anexo A**). Seu protocolo específico (como desmembramento do referido estudo) obteve aprovação da Comissão de Ética para Análise de Projetos de Pesquisa (CAPPesq) no ano de 2020, sob o nº 4.019.807 (**Anexo B**). Todos os participantes incluídos neste estudo manifestaram interesse, condições e disponibilidade de participação dos procedimentos realizados e assinaram o Termo de Consentimento Livre e Esclarecido (TCLE) (**Anexo C**).

5.3 SELEÇÃO DE PARTICIPANTES E PACIENTES

No período de janeiro de 2019 a fevereiro de 2020, duzentos e vinte participantes (n= 53 controles e n= 167 enfermos) foram recrutados para participar do projeto maior denominado “VALIDYS: Validação de uma ferramenta

para rastreamento do risco de disbiose e caracterização da microbiota intestinal sadia e enferma”, de acordo com critérios de inclusão e exclusão específicos de diferentes ambulatórios do Hospital das Clínicas da FMUSP. Após assinatura do TCLE, foram incluídos participantes, de ambos os sexos, com idade maior que 18 anos e inferior a 85 anos, e sem restrição de etnia. Constituem critérios de exclusão a existência isolada ou concomitante de enfermidades infecciosas (virais, bacterianas e parasitárias), neoplasia maligna, assim como pós-operatório imediato de intervenções cirúrgicas de médio ou grande porte (3 meses), indivíduos hospitalizados ou em cuidados domiciliares (**Quadro 2**).

Quadro 1. Critérios de inclusão e exclusão do estudo VALIDYS.

| Critérios de inclusão | |
|---|---|
| Enfermos | Sadios (controles) |
| <ul style="list-style-type: none"> - Idade \geq 18 anos e \leq 85 anos - Ter diagnóstico confirmado de alguma das doenças abaixo e realizar acompanhamento nos ambulatórios do HCFMUSP: <ul style="list-style-type: none"> • Diabetes tipo 1 (DM 1) • Diabetes tipo 2 (DM 2) • Psoríase em placas (PP) • Lupus eritematoso sistêmico (LES) em remissão clínica • Artrite reumatoide (AR) • Doenças inflamatórias intestinais (DIIs) <ul style="list-style-type: none"> - Doença de Crohn e Retocolite Ulcerativa em remissão clínica - Estar capacitado para responder, de forma consciente, perguntas realizadas na anamnese, recordatório alimentar de 24 horas (R24h) e realizar auto coleta de material fecal. - Não ter utilizado antibiótico nos 30 dias antecedentes a coleta de material fecal | <ul style="list-style-type: none"> - Idade \geq 18 anos e \leq 85 anos - Não ter nenhum diagnóstico para enfermidades - Não realizar tratamento medicamentoso para enfermidades - Não fazer uso de nenhuma medicação contínua (exceto anticoncepcional) - Estar capacitado para responder, de forma consciente, perguntas realizadas na anamnese, recordatório alimentar de 24 horas (R24h) e realizar auto coleta de material fecal. - Não ter utilizado antibiótico ou corticoides nos 30 dias antecedentes a coleta de material fecal - IMC \leq 29,99kg/m² |
| Critérios de exclusão | |
| <ul style="list-style-type: none"> - Ter consumido antibióticos nos últimos 3 meses antes da participação do estudo | |

- Presença de doenças crônicas virais, bacterianas ou parasitárias
- Presença de câncer
- Pós-operatório imediato ou de curto prazo (3 meses) de cirurgias de médio ou grande porte
- Ter realizado cirurgia bariátrica a qualquer momento antes da participação do estudo
- Indivíduos hospitalizados ou em cuidados domiciliares

5.4 CÁLCULO AMOSTRAL

O projeto maior, “VALIDYS”, apresentou amostra total de 220 indivíduos (53 sujeitos no grupo sadio (controle) e 167 sujeitos para os grupos de doença). A fim de responder o objetivo da presente pesquisa, uma estratificação deste grupo foi criada para que nenhum grupo enfermo apresentasse número amostral inferior a 20 indivíduos (número escolhido embasado em estudos anteriores da literatura comparando indivíduos saudáveis e enfermos^{88,152-154}) e tornando a amostra homogênea (**Figura 5**).

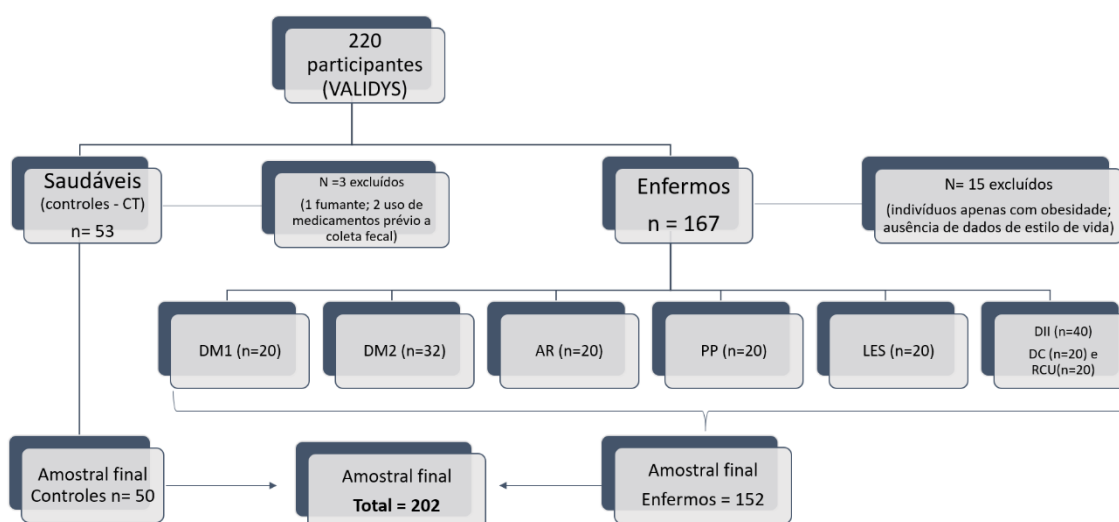


Figura 5. Diagrama consorte fluxo dos participantes do estudo

Legenda: DM1, Diabetes tipo 1; DM 2, Diabetes tipo 2; AR, artrite reumatoide; PP, Psoríase em placas; LES, Lupus eritematoso sistêmico; DII, Doença inflamatória intestinal, DC, Doença de Crohn; RCU, Retocolite Ulcerativa.

O presente estudo utilizou os dados clínicos, sociodemográficos, perfil de composição corporal e dados do consumo alimentar de 202 participantes obtidos a partir do projeto principal VALIDYS. Todos estes dados foram expressos em frequência média ou mediana, com seus respectivos desvios padrão e intervalo interquartilico.

5.5 ANAMNESE E HISTÓRIA CLÍNICA

Os participantes responderam a anamnese realizada pelo pesquisador, em que foram coletadas informações que incluíram nome, data de nascimento, etnia, sexo, estado civil, escolaridade, profissão/ocupação, histórico de doenças atuais e prévias, tabagismo, etilismo, alergia ou intolerância alimentar, prática de atividade física, consumo alimentar habitual (frutas, legumes e verduras, alimentos processados e consumo de açúcares/adoçantes), hábito intestinal (frequência de evacuações e escala de Bristol), uso de suplementos nutricionais, ingestão hídrica e medicamentos em uso (tipo, frequência e dose).

5.6 AVALIAÇÃO DO CONSUMO ALIMENTAR

O consumo alimentar foi avaliado por meio de recordatório alimentar de 24h em três momentos distintos. O primeiro no mesmo momento da aplicação da anamnese clínica. Em um segundo momento, na entrega do kit de coleta fecal para o paciente, foi realizada a segunda entrevista nutricional por meio do R24.¹⁵⁵ Na data de devolução do kit contendo as fezes coletadas, foi realizado o terceiro R24.

A aplicação do R24 se deu por meio de entrevista estruturada e semidirigida com questionamentos sobre o alimento consumido, preparações e horários das refeições, conforme preconizado em literatura.¹⁵⁶ Na aplicação dos registros de 24 horas, de modo a evitar erros na etapa de cálculo de nutrientes, foi questionado aos participantes os seguintes tópicos:

- Medidas caseiras utilizadas;

- Quando em alimentos preparados (como bolos, tortas, vitaminas, sopas, etc.) o registro detalhado dos ingredientes e forma de preparo;
- Para alimentos industrializados a marca do alimento;
- Detalhes como adição de sal, açúcar, óleos e molhos;
- Questionamento de uso de alimentos diets e lights;
- Em caso de sucos, se industrializados ou naturais; leites e derivados, detalhar se era integral, semidesnatado ou desnatado;
- Pães, arroz, macarrão se integral ou simples, além de especificar os tipos das carnes.
- Em todos os alimentos, indicar o modo de preparo.
- Uso de suplementos alimentares e multivitamínicos.

O valor energético da dieta e conteúdo de nutrientes foi calculado utilizando-se o programa Easy Diet[®], escolhido pela necessidade de utilizar informações de alimentos brasileiros, por ser sensível em identificar o déficit de ingestão de nutrientes e pela facilidade de inserir alimentos e preparações, assim como suas informações nutricionais. Posteriormente, os nutrientes calculados foram exportados para o programa Excel[®]. As tabelas de composição utilizada pelo software são: Tabela de Composição de Alimentos (TACO,¹⁵⁷ Philippi,¹⁵⁸ IBGE¹⁵⁹ e USDA¹⁶⁰).

A inserção das informações foi padronizada e, no caso das preparações não contidas na base de dados do programa, estas foram incluídos no item “receita”. A definição dos ingredientes, quantidades e rendimento das receitas, foi estabelecida consultando as publicações nacionais, Pinheiro et al. (2001)¹⁶¹ e IBGE (2011)¹⁵⁹. Alimentos industrializados tiveram seus valores nutricionais consultados no site da própria indústria produtora.

Houve padronização dos dados de consumo alimentar, no qual as medidas caseiras foram convertidas em unidades de peso (gramas) ou em unidades de volume (mililitros). Foi realizado também o controle de qualidade dos registros, onde identificou-se e corrigiu-se eventuais erros na obtenção dos dados dietéticos, tais como a descrição incompleta dos alimentos, do modo de

preparo e/ou das medidas caseiras. Utilizou-se o “Manual de críticas de inquéritos alimentares” como apoio, que contempla as principais informações necessárias para a crítica de inquéritos alimentares para a padronização das medidas caseiras.¹⁶²⁵ Além deste manual, foi necessário a escolha do próprio pesquisador quanto a padronização de alguns tipos de alimentos, utilizada quando o indivíduo pesquisado não especificou o alimento detalhadamente, por exemplo, quando foi registrado apenas “banana”, foi optado pelo registro de “banana nanica”.

5.7 AVALIAÇÃO DE COMPOSIÇÃO CORPÓREA

5.7.1 Peso, estatura e cálculo de IMC

A avaliação antropométrica incluiu medidas de peso, altura, aplicação de bioimpedância elétrica e cálculo de IMC. O peso atual (kg) do paciente foi medido em balança de plataforma eletrônica (Life Measurement Instruments, Concord, CA), calibrada conforme orientação do fabricante. Durante a avaliação, o paciente permaneceu em pé, no centro da base da balança, descalço e usando roupas leves.

A mensuração da altura foi realizada por meio do estadiômetro, modelo Sanny. Durante o procedimento, o paciente manteve-se de costas para o seu marcador, com os pés unidos, em posição ereta, olhando para frente e braços estendidos ao lado do corpo. A leitura foi feita no centímetro mais próximo, quando a haste horizontal da barra vertical da escala de estatura encostou à cabeça do indivíduo.

O IMC foi determinado pelo peso do indivíduo dividido por sua altura ao quadrado e expresso em kg/m². Estes dados foram anotados em planilha apropriada no software de captura de dados eletrônicos RedCap® para posterior criação de banco de dados e análises estatísticas.

5.7.2 Composição corporal

Foi realizada estimativa da composição corporal, utilizando aparelho de bioimpedância elétrica portátil de quatro frequências (5, 50, 100, 200 kHz) modelo Quadiscam 4000, Bodystat Ltda, Douglas, UK. Nas 24 horas

anteriores ao exame, os voluntários foram orientados a evitar consumo de cafeína e a prática de atividade física intensa. No dia do exame, antes do início do procedimento de avaliação da composição corporal, os voluntários foram orientados a esvaziar a bexiga urinária, retirar todos os objetos metálicos do corpo e permanecer 15 minutos de repouso na posição supina, com braços e pernas em abdução de 30°.

A execução do exame contou com uso de eletrodos específicos para o equipamento Quadiscam 4000, Bodystat®, que foram colocados na mão e no pé do voluntário, especificamente no calcanhar e punho do hemicorpo direito, conforme orientado no manual do fabricante para mensuração da composição corporal.

Dados referentes ao sexo, idade, peso, altura e nível de atividade física foram inseridos no aparelho. A mensuração da composição corporal foi feita por meio da passagem de correntes elétricas imperceptíveis de 5, 50, 100 e 200 kHz, via eletrodos distais, posicionados na mão e pé do paciente. Foram obtidos resultados da impedância das quatro frequências e consequente estimativa da composição corporal.

Os dados referentes a gordura corporal (% e kg) e massa magra (% e kg) obtidos por equações fornecidas pelo próprio aparelho foram anotados em planilha apropriada no software de captura de dados eletrônicos RedCap® para posterior criação de banco de dados e análises estatísticas.

5.8 ANÁLISE FECAL - COMPOSIÇÃO DA MICROBIOTA INTESTINAL

5.8.1 Coleta de amostras

Os participantes receberam um kit padronizado de coleta de amostras fecais do laboratório Bioma4me. Este kit permitiu que o paciente realizasse sua auto coleta fecal em domicílio e excluiu a necessidade de congelamento imediato. Os participantes foram instruídos a utilizar um *swab* estéril para transferir uma pequena quantidade de amostra fecal (na quantidade menor que

um grão de feijão) para o interior de microtubos plásticos de 2 ml contendo solução tampão de preservação de DNA.

Foram fornecidos 3 microtubos para a coleta de amostras, sendo dois microtubos destinados para realização da caracterização da composição da microbiota intestinal pela técnica de sequenciamento do gene 16SrRNA. O terceiro tubo foi armazenado para a eventual necessidade de repetição do sequenciamento da microbiota intestinal e / ou outras análises oportunas.

Todos os participantes foram orientados a realizar sua coleta no máximo 24 horas antes da entrega do material fecal para os pesquisadores e manter sua amostra em temperatura controlada (geladeira), até o momento da entrega do material.

5.8.2 Extração de DNA

O DNA das amostras fecais foi extraído com o kit QIAamp® PowerFecal® DNA Kit, da QUIAGEN® de acordo com protocolo padronizado.

Resumidamente, considerando peso médio de 0,25 gramas de amostra, as amostras foram homogeneizadas em um tubo de 2 ml com microesferas metálicas. Esse processo promove a lise celular das células hospedeiras, assim como das células microbianas, que juntamente com a adição de reagentes químicos apropriados, garante a extração eficiente do DNA bacteriano.

Este kit de extração de DNA bacteriano usa tecnologia de remoção de inibidores (TRI) para fezes. TRI é muito eficaz na remoção de substâncias inibitórias comumente encontradas em material fecal, como polissacarídeos, heme e sais biliares que interferem na reação em cadeia polimerase (PCR).

Após obtenção do DNA genômico total, capturado em uma membrana de sílica em um formato de coluna, foram realizados processos de lavagem e eluição, dando origem ao DNA isolado com alta pureza, para realização de sequenciamento genético da microbiota.

5.8.3 Preparo de biblioteca

O preparo de biblioteca consistiu em adicionar identificadores as amostras. Nesta etapa, cada amostra de DNA recebeu etiquetas específicas que permitem o processamento simultâneo de todas as amostras.

Neste momento foram adicionadas as sequências específicas de *primers*, seguindo o protocolo Illumina[®] 162, contendo regiões variáveis V3 e V4 descritos abaixo:

16S Amplicon PCR Forward Primer = 5 '

**TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGCCTACGGGNGGCWGC
AG**

Primer reverso de PCR de 16S Amplicon = 5 '

**GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAGGACTACHVGGGTATC
TAATCC**

O protocolo também inclui sequências de adaptadores *overhang*, que devem ser anexadas às sequências do par de primers para compatibilidade com o índice Illumina e adaptadores de sequenciamento.

Os produtos foram agrupados e as bibliotecas foram quantificadas por PCR usando os kits Quant-iT and Qubit DNA quantitation assays (PICOGREEN[®], Quant-iT[™]) e Kapa Hifi Hotstart (roche[®]), antes de introduzir as amostras no sequenciador.

5.8.4 Sequenciamento 16SrRNA

A avaliação da composição da microbiota intestinal (MI) foi realizada por técnica de sequenciamento do gene 16S, que é um método de sequenciamento de nova geração (SNG) de amplificação usado para identificar e avaliar abundância de bactérias presentes em uma amostra. O sequenciamento do gene que codifica a região 16s do rRNA é um método bem estabelecido para estudar filogenia e taxonomia de amostras de microbiomas complexos.

Após preparo das amostras e da biblioteca, as amostras de DNA bacteriano foram inseridas no sequenciador Illumina[®] MiSeq que pode fornecer leituras de 2 x 300 pb e até 70 milhões de leituras pareadas, gerando até 15 GB de dados. Para o presente estudo, foi selecionado como padrão o mínimo de 68

mil leituras por amostra para promover melhor captação dos microrganismos avaliados.

5.9 BIOINFORMÁTICA E ESTATÍSTICA

5.9.1 Processamento das sequências e classificação taxonômica da microbiota intestinal

As leituras de sequência bruta do gene 16S rRNA foram processadas, cortadas e agrupadas em variantes de sequência de amplicon (ASVs), usando DADA2.¹⁶⁴ Os iniciadores usados na amplificação foram removidos, e sequências com mais do que 2 erros esperados foram descartadas. As sequências remanescentes foram utilizadas para treinar um modelo de identificação e correção dos erros. As leituras diretas e reversas, já corrigidas, foram concatenadas para formação da ASVs, para a remoção de sequências quiméricas e quantificação das ASVs.

Cada ASV teve sua classificação taxonômica designada pelo pacote TAG.ME¹⁶⁵, usando como referência o banco de dados SILVA¹⁶⁶ utilizando o modelo específico para o amplicon, que corresponde à região 341F-805R.

5.9.2 Caracterização de α e β diversidade da microbiota intestinal

Foram realizados os cálculos de índices de Simpson, Shannon e Chao1 para alfa diversidade utilizando o pacote R vegan (versão 2.5-7). A beta diversidade foi calculada pela divergência de Jehnsen-Shannon.

5.9.3 Caracterização da microbiota intestinal

A tabela de quantificação de ASVs gerada na etapa 5.9.1 foi filtrada pelo pacote HTSfilter no ambiente R (versão 1.30.1), que eliminou os táxons com representatividade não relevante para identificação de padrões. Este pacote utiliza o índice de dissimilaridade de Jaccard para identificar um valor de corte para o número de sequências a serem filtradas para as análises seguintes **(Figura 6)**.

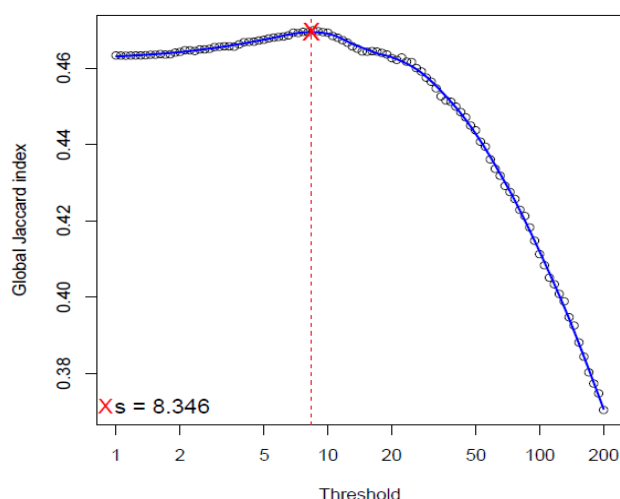


Figura 6. Distribuição do índice global de Jaccard.

O índice global de Jaccard aponta o valor de corte de sequências com menos de 8,34 leituras. Finalmente, após a filtragem de amostras inferiores a 8,34 leituras, utilizou-se o teste Wald do pacote DESeq2 (Versão 1.30.1)¹¹³ para identificar os táxons bacterianos diferencialmente abundantes entre cada grupo clínico (cada grupo enfermo *versus* grupo sadio). Da presente análise foram selecionados "potenciais marcadores microbianos" com valor p ajustado para testes múltiplos usando o método Benjamini-Hochberg (p ajustado $p < 0,05$).

5.9.4 Análise das informações fenotípicas

No presente estudo, as análises descritivas tiveram as variáveis contínuas expressas em termos de medidas de resumo (média, mediana, desvio padrão e quartis), enquanto as variáveis categóricas foram expressas em termos de frequências e porcentagens. A normalidade das amostras e definição de qual teste estatístico era adequado para o conjunto de dados foi realizada com a utilização do teste de Shapiro-Wilk.

A análise de dados (não relacionados com a metagenômica) e comparação de dois grupos nas variáveis contínuas deu-se pela utilização de teste não paramétrico U de Mann-Whitney, porque as variáveis não seguiram distribuição normal. Análises com 3 ou mais grupos que não seguiam distribuição

normal foram realizadas por meio do teste de Kruskal-Wallis. Neste caso, para qualquer significância no teste Kruskal-Wallis ($p < 0,05$) foi aplicado teste post-hoc Dunnett, comparando o grupo da enfermidade específica com o grupo sadio. Variáveis categóricas foram avaliadas utilizando teste qui-quadrado.

O nível de significância adotado nos testes foi de 0,05. Foram consideradas hipóteses bicaudais. O software JASP Team (2020). JASP (Version 0.14.1) [Computer software] e R versão 4.0.5 foram utilizados para a realização das análises estatísticas.

5.9.5 Análise da ingestão habitual de nutrientes

A ingestão habitual dietética foi calculada pelo método MSM, sigla do inglês “*Multiple Source Method*”, na plataforma desenvolvida pelo Departamento de Epidemiologia do Instituto Alemão de Nutrição Humana Potsdam-Rehbrücke (DIfE)© 2008-2021 Versão 1.0.1 online.^{167,168} Este método é caracterizado por uma técnica que estima a probabilidade de ingestão individual de cada nutriente, seguido do cálculo da ingestão habitual individual de cada um desses nutrientes nos dias de consumo e, por fim, multiplicado os valores obtidos nos passos anteriores para obter a ingestão diária habitual, reduzindo as variações inter e intraindividuais.

Após os ajustes oriundos do MSM, realizou-se comparação de dois grupos (doentes *versus* um grupo enfermo) nas variáveis, utilizando o teste não paramétrico U de Mann-Whitney.

5.9.6 Seleção de variáveis para integrar os modelos preditivos

A seleção de variáveis é muito importante para criação de modelos mais eficientes, especialmente para conjuntos de dados de alta dimensão, como dados de microbiota intestinal. No que diz respeito às variáveis fenotípicas, foram consideradas pré-selecionadas para integrar os modelos preditivos aquelas que foram significativamente diferentes na comparação entre os grupos. No que diz respeito à microbiota intestinal, tanto as bactérias que apresentaram diferença

na abundância diferencial em uma análise par a par (diferença específica entre sadios e um único grupo enfermo), quanto aquelas que foram diferentes em uma comparação entre todos os grupos, foram consideradas como pré-selecionadas (**figura 7 a**).

No momento da seleção final de variáveis, aplicou-se 50 vezes o algoritmo Random Forest (ou floresta randômica - FR) e optou-se por aquelas que, dos 50 testes realizados, figurassem em pelo menos 50% das vezes entre as 10 mais importantes. Esse processo foi feito separadamente para cada tipo de variável, a saber, categóricas, quantitativas e da microbiota intestinal (**figura 7c**).

5.9.7 Criação de modelos preditivos

5.9.7.1 Criação de modelos preditivos para discriminação de indivíduos sadios e enfermos utilizando características fenotípicas e da microbiota intestinal isoladamente

Os modelos preditivos foram construídos, a priori, para cada conjunto de dados – quantitativos, qualitativos, e abundância de gêneros da microbiota. Toda a etapa de modelagem foi implementada em ambiente R utilizando o pacote “caret”¹⁶⁹ com o algoritmo Random Forest (ou floresta randômica - FR) implementado em 100 árvores e validação cruzada de 20 vezes. O conjunto de dados foi dividido ao separarmos aleatoriamente metade dos pacientes saudáveis e doentes em um conjunto para a construção do modelo. Os demais pacientes foram utilizados para testar o modelo e avaliar sua performance (**Figura 7b**).

A informação final de uma FR é obtida usando uma votação majoritária das árvores de decisão individuais. Por serem métodos bem estudados, são usados como base para comparação em muitos estudos para estratificação de

pacientes, descoberta e confirmação de biomarcadores, e busca por “assinaturas de microbiota” para determinadas condições.^{70,151, 170-174}

A performance foi avaliada ao medir a capacidade preditiva – Sensibilidade e Especificidade – usando o conjunto de dados de teste. A importância de cada variável foi medida utilizando o índice “Mean Decrease Gini”¹⁷⁶, que representa a porcentagem de predições que o modelo deixa de acertar ao retirar aquela variável (**Figura 7d**).

O procedimento acima descrito foi repetido 50 vezes, com seleção aleatória de 50 conjuntos de dados diferentes para os dados de treino e de teste. A cada construção, os dados das importâncias das variáveis eram armazenados, e as 10 variáveis mais importantes eram identificadas. A performance de cada um dos modelos foi utilizada para a construção da curva ROC (**Figura 7e**).

5.9.7.2 Criação do modelo preditivo integrativo para discriminação de indivíduos sadios e enfermos utilizando características fenotípicas e da microbiota intestinal isoladamente

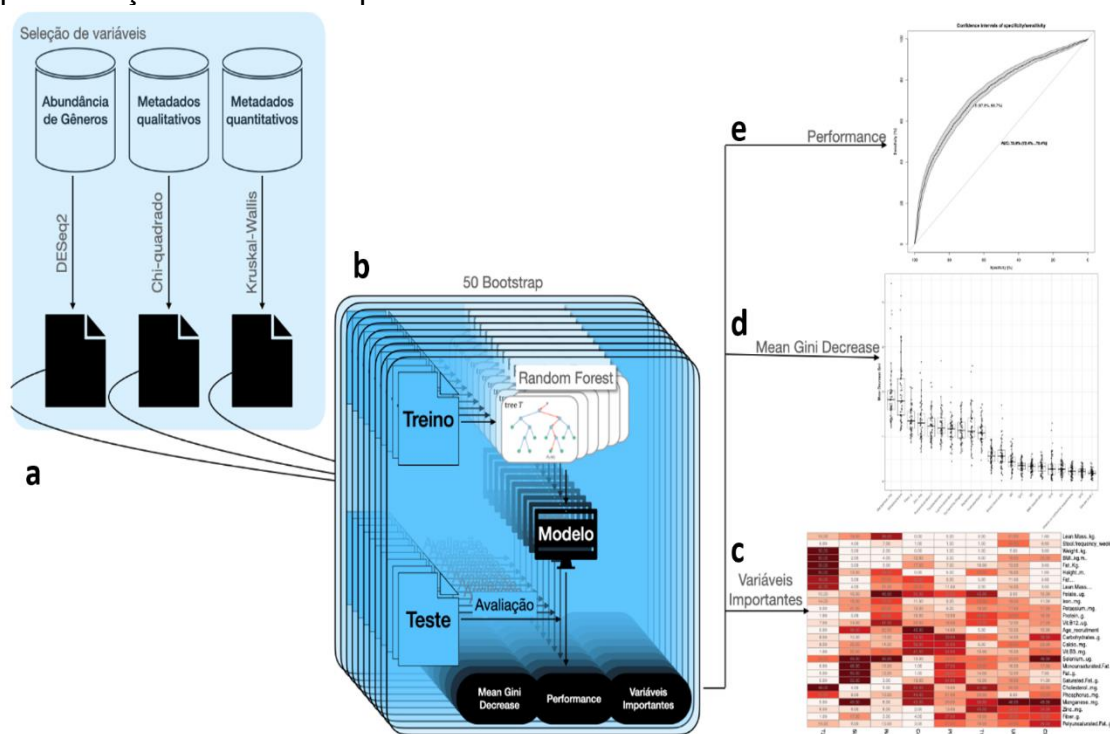
O modelo integrativo foi criado utilizando as variáveis que apareceram entre as 10 mais importantes em, pelo menos, metade (25) dos modelos de, pelo menos, uma das doenças e em cada conjunto de dados – quantitativo, qualitativo e microbiota. O modelo final foi gerado a partir do consenso de 50 reamostragens de bootstrap (amostra *bootstrap* é um conjunto de valores extraídos ao acaso, com reposição da amostra original). A importância de cada variável foi armazenada, e as performances foram utilizadas para a construção da curva ROC (**Figura 7**).

ROC (curva característica de operação do receptor) é uma curva de probabilidade e a área sob a curva (AUC) representa sua medida de separabilidade. Essas métricas dizem o quanto o modelo é capaz de distinguir entre as classes. No presente estudo, quanto maior a AUC, melhor o modelo é em distinguir entre pacientes com a doença e controles sadios.¹⁷⁵

A modelagem com FR fornece a AUC-ROC com valores de discriminação que variam de discriminação aleatória (AUC=0,5 ou 50%) a discriminação perfeita (AUC=1,0 ou 100%).¹⁷⁶

A **figura 7** resume os passos utilizados para a seleção de variáveis candidatas e criação dos modelos preditivos.

Figura 7. Fluxo da seleção de variáveis candidatas e aprendizado de máquina para criação de modelos preditivos



Legenda: O fluxo de modelagem consistiu em cinco etapas.

a) A pré-seleção das variáveis microbianas e variáveis fenotípicas qualitativas e quantitativas. **b)** A segunda etapa consistiu na criação dos modelos. O primeiro resultado da modelagem elegeu as variáveis importantes **(c)**. Com as variáveis importantes **(c)** selecionadas, criou-se um modelo para cada doença através de processos de treinamento e avaliação. Os modelos desenvolvidos tiveram seu desempenho avaliado através do apontamento de **(d)** importância da variável (mean decrease gini) e **(e)** avaliação de capacidade preditiva (performance).

Para os presentes modelos, seguiremos a classificação de Hosmer e Lemeshow (**quadro 2**) sobre a capacidade preditiva dos modelos que considera:¹⁷⁷

Quadro 2. Classificação da capacidade preditiva das AUC

- $AUC = 50\%$: discriminação aleatória,
- $50\% < AUC \leq 60\%$: discriminação fraca,
- $60\% < AUC \leq 70\%$: discriminação regular,
- $70\% < AUC \leq 80\%$: discriminação boa,
- $80\% < AUC \leq 90\%$: discriminação excelente,

Fonte: Hosmer e Lemeshow, 2000.¹⁷⁷

6 RESULTADOS

6 RESULTADOS

6.1 CARACTERÍSTICAS DESCRITIVAS DA AMOSTRA

O presente estudo incluiu 50 indivíduos sadios e 152 enfermos (amostra total de 202 indivíduos). Cada enfermidade contou com aproximadamente 15% da amostra total de enfermos, com exceção ao grupo de doenças inflamatórias intestinais (DII) que contemplou 2 subtipos de doença, agregando, assim, 30% da amostra enferma.

A mediana de idade no grupo controle (CT) foi de 32 anos, que diferiu significativamente dos grupos Psoríase em placas (PP), Artrite reumatoide (AR) e diabetes tipo 2 (DM2), que foram 55, 59,5 e 52,5 anos, respectivamente. Com relação aos gêneros, o sexo feminino foi mais predominante em todas as categorias de doenças avaliadas. Já o grupo controle apresentou distribuição similar de indivíduos de ambos os sexos. As características fenotípicas da população avaliada estão apresentadas na **Tabela 2**.

Com relação as características de composição corpórea e IMC, no grupo controle, a amostra foi representada apenas por pacientes com eutrofia ou sobrepeso, representando 56% e 42%, respectivamente. Nos grupos enfermos, o IMC e a composição corpórea diferiram significativamente nos pacientes dos grupos DII, PP, AR e DM2, comparado ao grupo controle. As características de composição corpórea e IMC da população avaliada estão apresentadas na **Tabela 2**.

Tabela 2. Características fenotípicas da população do estudo.

| Variáveis | CT (n=50) | DII (n=40) a | PP (n=20) b | AR (n=20) c | LES (n=20) d | DM1 (n=20) e | DM2 (n=32) f | P-valor |
|-----------------------------------|---------------------|----------------------|---------------------|---------------------|---------------------|---------------------|---------------------|--------------------|
| Idade (min-max) | 32,0 (19-83) | 40,0(20-72) | 55,0 (34-71) | 59,5(32-69) | 35,5 (30-45) | 35 (19-56) | 52,5 (33-73) | < ,001 * b,c,f |
| Etnia (n)(%) | | | | | | | | |
| Preto | 2,0 (4,0) | 3,0 (7,5) | 5,0 (25,0) | 3,0 (15,0) | 2 (10,0) | 1 (5,0) | 3 (9,4) | < ,001 # |
| Pardo | 4,0 (8,0) | 14,0 (35) | 6,0 (30,0) | 4,0 (20,0) | 7 (35,0) | 8 (40,0) | 19 (59,4) | |
| Branco | 39,0 (78,0) | 23,0 (57,5) | 9,0 (45,0) | 13,0 (65,0) | 11 (55,0) | 11 (55,0) | 10 (31,2) | |
| Amarelo | 5,0 (10,0) | 0,0 (0,0) | 0,0 (0,0) | 0,0 (0,0) | 0 (0,0) | 0 (0,0) | 0 (0,0) | |
| Gênero (%) | | | | | | | | |
| Feminino | 23,0 (46,0) | 24 (60,0) | 12,0 (60,0) | 15 (75,0) | 20 (100) | 11 (55,0) | 17 (53,0) | 0,003 # |
| Masculino | 27,0 (54,0) | 16 (40,0) | 8 (40,0) | 5 (25,0) | 0 (0,0) | 9 (45,0) | 15 (47,0) | |
| IMC (kg/m²) (%) | | | | | | | | |
| Desnutrição | 0,0 (0,0) | 1 (2,5) | 0 (0,0) | 0 (0,0) | 0 (0,0) | 1 (5,0) | 0 (0,0) | < ,001# a,b,c,f |
| Eutrofia | 28,0 (56,0) | 18 (45,0) | 7 (35,0) | 9 (45,0) | 9 (45,0) | 10 (50,0) | 7 (21,8) | |
| Sobrepeso | 21,0 (42,0) | 14 (35,0) | 7 (35,0) | 4 (20,0) | 10 (50,0) | 7 (35,0) | 3 (9,4) | |
| Obesidade | 0,0 (0,0) | 7 (17,5) | 6 (30,0) | 6 (30,0) | 1 (5,0) | 2 (10,0) | 22 (68,8) | |
| Composição corporal | | | | | | | | |
| Peso (kg) | 68,0 (41,5-100) | 70,0 (43-108) | 72,0 (49,1-106) | 68,0 (47,5-92) | 65,5 (51,5-82) | 71 (40-97) | 97,0 (58-201) | <,001*f |
| IMC (kg/m²) | 24,3 (18,1-29,8) | 25,4 (17,8-35,9) | 27,0 (20,2-37,5) | 25,6 (20,8-36,6) | 25,4 (19,4-31,6) | 24,9 (16,7-31,3) | 37,5 (21,6-60,6) | <,001*f |
| Massa gorda (%) | 26,8 (8,7-41,0) | 33,0 (17,0-48,6) | 39,2 (20,3-55,4) | 44,2 (31,7-61,7) | 33,2 (19,6-46,8) | 29,1 (15,6-42,9) | 47,5 (19,7-68,1) | <,001* a,b,c,f |
| Massa magra (%) | 73,2 (59,0-91,3) | 66,95 (51,4-83,0) | 60,8 (44,6-79,7) | 55,8 (38,3-68,3) | 66,9 (50,7-80,4) | 70,9 (57,1-84,4) | 52,5 (31,9-80,3) | <,001* a,b,c,f |

Legenda: Valores expressos em mediana (mínimo – máximo) ou valor absoluto (%); CT, controles; DII, doença inflamatória intestinal (a); PP, psoríase em placas (b); AR, artrite reumatoide(c); LES, Lupus eritematoso sistêmico (d); DM1, diabetes tipo 1 (e); DM2, diabetes tipo 2 (f), Teste Qui-quadrado # ; Teste Kruskal-Wallis*; Nível de significância adotado p<0,05

6.2 ASPECTOS DESCRITIVOS DA MICROBIOTA INTESTINAL

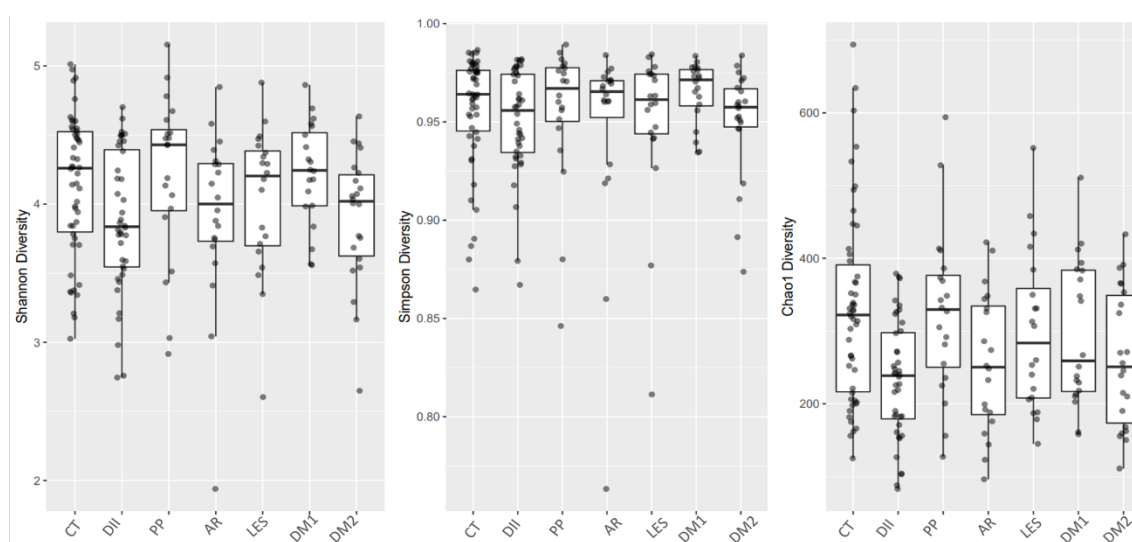
Após o pré-processamento dos dados do sequenciamento de microbiota intestinal de 202 indivíduos saudáveis e enfermos, obteve-se o total de 10.808 ASVs, das quais 95,31% (desvio padrão - DP 5,42%) foram classificadas em nível taxonômico até gênero e 54,62% (DP 15,26%) até espécie.

As amostras apresentaram em média 113 mil leituras após limpeza das sequências, e somente aquelas com mais de 68 mil foram consideradas para os próximos passos (**Apêndice 1**).

6.2.1 Ecossistemas microbianos (α diversidade) e macroestrutura (β diversidade) em indivíduos saudáveis e enfermos.

Para a comparação dos parâmetros ecológicos de diversidade α , foram realizadas análises dos índices de diversidade de Shannon e Simpson, bem como índice Chao1 para verificação da riqueza microbiana de cada um dos grupos (**Gráfico 1**).

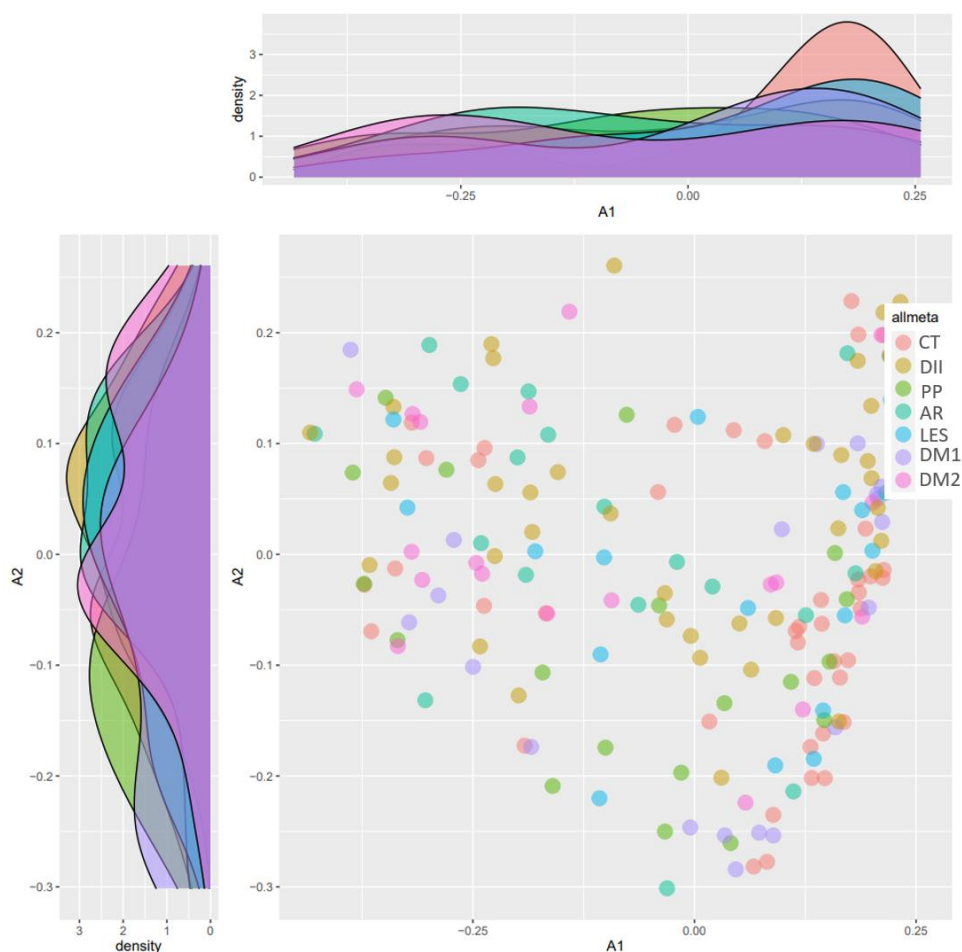
Gráfico 1. Indicador de diversidade alfa, índices Shannon, Simpson e Chao1



Foi possível observar maior riqueza microbiana (Chao1) em indivíduos controle somente quando comparados aos indivíduos do grupo DII ($p=0,007$). Os índices de Shannon e Simpson não diferiram entre nenhum dos grupos; a tabela completa com os valores de significância para cada comparação encontra-se no **Apêndice 2**.

A análise de macroestrutura da microbiota intestinal entre os diferentes grupos deu-se a partir de comparação de métricas de diversidade β (**Figura 8**). Por meio da análise de coordenadas principais (PcoA), calculadas com base na distância de Jensen-Shannon, é possível observar que a macroestrutura é bastante sobreposta entre todos os grupos, impedindo a clara separação por grupo de doença ou estado de saúde.

Figura 8. Macroestrutura da microbiota intestinal.



Legenda: PcoA da macroestrutura da microbiota intestinal de controles versus DII, PP, AR, LES, DM1 e DM2 calculadas pelo índice Jensen-Shannon.

Ao realizar análise PERMANOVA com correção por Benjamini-Hochberg para identificação de diferenças entre grupos, novamente, apenas o grupo DII apresentou uma pequena distinção comparada ao grupo controle, sendo que a doença explica uma variação de 1,65% na diversidade β entre os indivíduos (Tabela 3).

Tabela 3. Análise de β diversidade entre grupos.

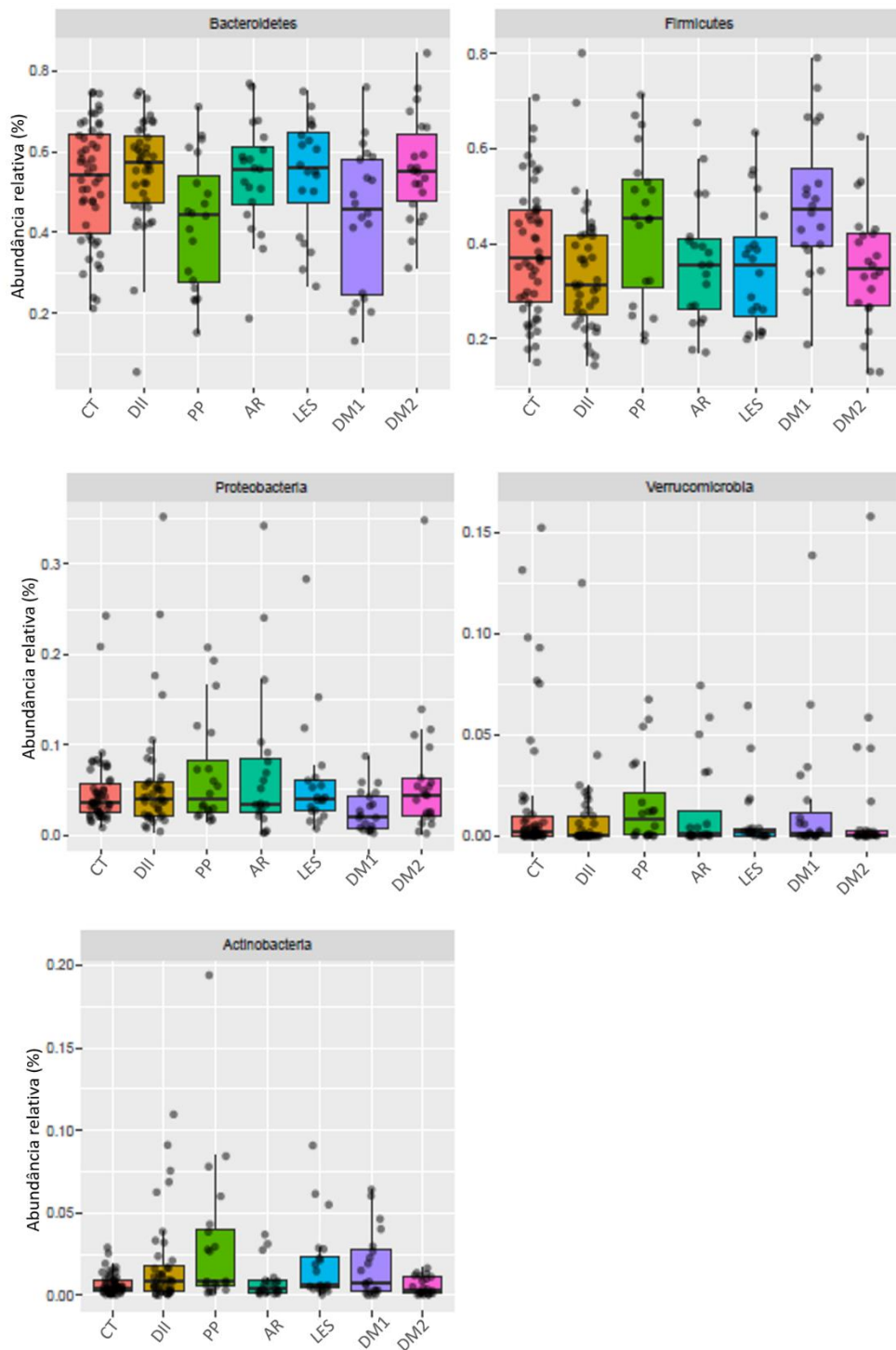
| Grupos | SumsOfSqs | F.Model | R ² | P-valor | P ajustado |
|------------------|--------------|--------------|----------------|--------------|--------------|
| CT vs PP | 0,368 | 1,436 | 0,020 | 0,007 | 0,147 |
| CT vs DM1 | 0,296 | 1,147 | 0,016 | 0,103 | 1,00 |
| CT vs DM2 | 0,356 | 1,370 | 0,019 | 0,016 | 0,336 |
| CT vs AR | 0,361 | 1,378 | 0,019 | 0,010 | 0,210 |
| CT vs DII | 0,386 | 1,476 | 0,016 | 0,001 | 0,021 |
| CT vs LES | 0,226 | 0,883 | 0,012 | 0,872 | 1,00 |

Legenda: CT, controles; DII, doença inflamatória intestinal (a); PP, psoríase em placas (b); AR, artrite reumatoide(c); LES, Lupus eritematoso sistêmico (d); DM1, diabetes tipo 1 (e); DM2, diabetes tipo 2 (f). Tamanho do efeito (R²) calculado por testes de análise de variância multivariada permutacional (PERMANOVA) e correspondente p-valores ajustados por Benjamini-Hochberg (BH); Nível de significância adotado $p < 0,05$.

6.2.2 Caracterização de filós da microbiota intestinal

Os cinco filós mais abundantes na microbiota dos indivíduos avaliados foram *Bacteroidetes*, *Firmicutes*, *Proteobacteria*, *Verrucomicrobia* e *Actinobacteria* (Gráfico 2). Não foi identificada diferença significativa entre os grupos enfermos e grupo CT.

Gráfico 2. Abundância relativa dos 5 principais filós bacterianos.

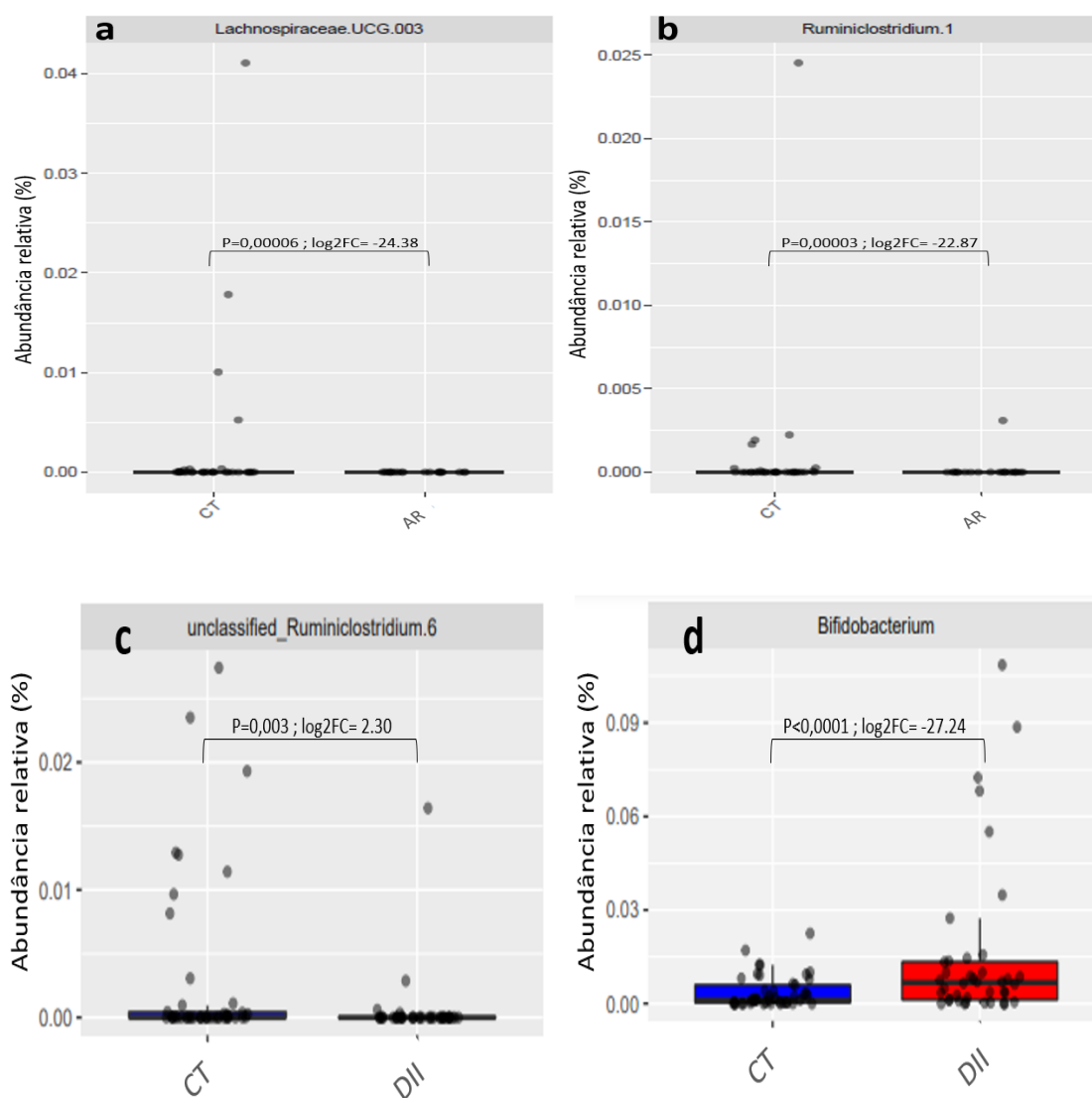


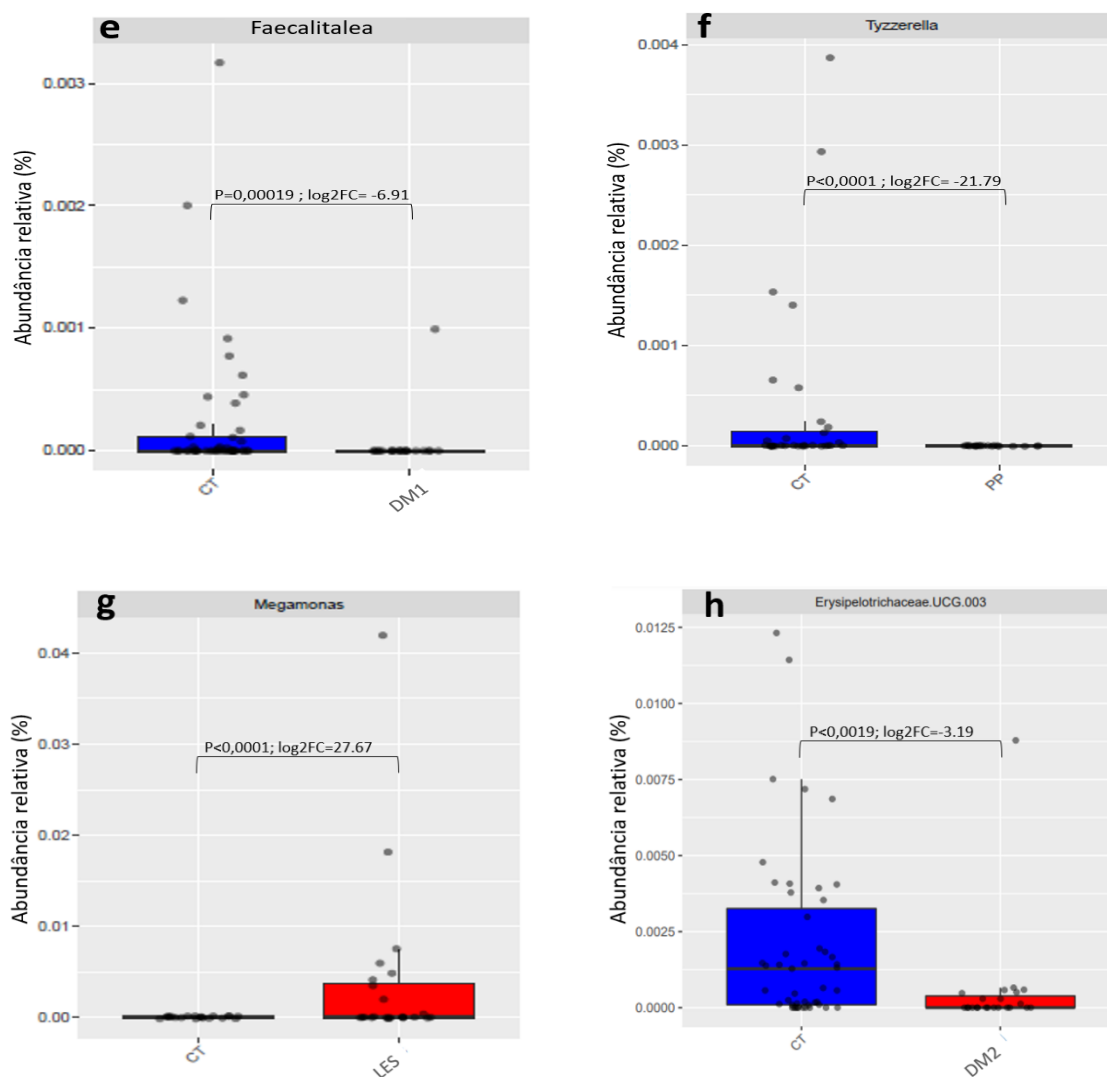
Legenda: CT, controles; DII, doença inflamatória intestinal; PP, psoríase em placas; AR, artrite reumatoide; LES, Lupus eritematoso sistêmico; DM1, diabetes tipo 1; DM2, diabetes tipo 2.

6.2.3 Descrição táxons bacterianos em indivíduos saudáveis e enfermos

Após o teste de Wald, para comparação par a par dos táxons microbianos, foram selecionados 8 táxons bacterianos diferentes entre grupo CT e alguma das enfermidades (**Gráfico 3**).

Gráfico 3. Táxons com abundância relativa significativamente diferente entre controles e doentes por análise par a par





Legenda: (a) e (b), referem-se às análises do grupo Controle versus AR, (c) e (d) grupo Controle *versus* DII, (e) Controle vs DM1, (f) Controle vs PP, (g) Controle vs LES e (h) Controle vs DM2. P-valor ajustado por Benjamini-Hochberg (BH). Log2FC, log2 Fold Change; CT, controles; DII, doença inflamatória intestinal; PP, psoríase em placas; AR, artrite reumatoide; LES, Lupus eritematoso sistêmico; DM1, diabetes tipo 1; DM2, diabetes tipo 2. Nível de significância adotado $p<0,01$.

6.3 CARACTERÍSTICAS FENOTÍPICAS DE INDIVÍDUOS SADIOS E ENFERMOS.

As perguntas da anamnese clínica, referentes a fatores ambientais ao estilo de vida, foram tabuladas e estão descritas na **Tabela 4**. Na categoria de fatores ambientais (tipo de parto e amamentação), é possível verificar maior frequência de parto normal em indivíduos enfermos (DII, PP, AR e DM2) quando

comparado ao grupo controle ($p < 0,001$). Ao que se refere a amamentação, a frequência de ter sido ou não amamentado ao seio não diferiu entre indivíduos CT ($p = 0,518$) e os indivíduos de nenhuma enfermidade avaliada.

As variáveis de estilo de vida referem-se à questões sobre o hábito intestinal (frequência e consistência caracterizada pela escala de Bristol), prática de atividade física (tempo praticado por semana e ser considerado fisicamente ativo de acordo com a OMS), tabagismo e estresse percebido.

Não houve diferença entre os grupos, referente à quantidade de minutos de exercício físico realizado por semana ($p = 0,715$). No entanto, maior número de indivíduos do grupo CT foram fisicamente ativos em comparação aos dos grupos DII, PP e AR ($p = 0,012$). Alterações na frequência e consistência das fezes foram percebidas entre CT e DII ($p = 0,002$), sendo DII caracterizado por maior frequência evacuatória e fezes de características amolecidas, a saber, Bristol 5, 6 e 7.

Nenhum indivíduo do CT foi tabagista ($p = 0,015$). Desta forma, todos os grupos enfermos apresentaram diferença significativa nesta variável, quando comparados a indivíduos saudáveis.

Em relação ao estresse percebido, apenas o grupo DM2 diferiu significativamente do grupo controle, apresentando 36% dos indivíduos com estresse alto ou muito alto e 32% com estresse baixo, contra 22% e no grupo controle nas categorias descritas.

Tabela 4. Características fenotípicas (fatores ambientais e variáveis de estilo de vida) de indivíduos sadios e enfermos.

| Variáveis | CT (n=50) | DII (n=40) a | PP (n=20) b | AR (n=20) c | LES (n=20) d | DM1 (n=20) e | DM2 (n=32) f | P-valor |
|---------------------------------------|--------------|--------------|-------------|-------------|--------------|--------------|--------------|-----------------|
| Fatores ambientais | | | | | | | | |
| Tipo de parto (%) | | | | | | | | |
| Vaginal | 23 (46%) | 33 (82,5%) | 18 (90%) | 19 (95%) | 12 (60%) | 12 (60%) | 24 (75%) | < ,001 #a,b,c,f |
| Cesariana | 24 (48%) | 7 (17,5%) | 2 (10%) | 1 (5%) | 7 (35%) | 7 (35%) | 6 (19%) | |
| Não sabe | 3 (6%) | 0 (0%) | 0 (0%) | 0 (0%) | 1 (5%) | 1 (5%) | 2 (6%) | |
| Amamentação (%) | | | | | | | | |
| Não | 3 (6%) | 2 (5%) | 1 (5%) | 3 (15%) | 1 (5%) | 0 (0%) | 2 (6%) | 0,518 # |
| Sim | 44 (88%) | 35 (87,5%) | 15 (75%) | 14 (70%) | 17 (85%) | 19 (95%) | 29 (90%) | |
| Não sabe | 3 (6%) | 3 (7,5%) | 4 (20%) | 3 (15%) | 2 (10%) | 1 (5%) | 1 (4%) | |
| Hábitos e estilo de vida | | | | | | | | |
| Escala de Bristol (%) | | | | | | | | |
| 1 | 0 (0%) | 0 (0%) | 3 (15%) | 0 (0%) | 0 (0%) | 1 (5%) | 1 (3%) | 0,002 # a |
| 2 | 4 (8%) | 2 (5%) | 2(10%) | 4 (20%) | 5 (25%) | 5 (25%) | 5 (16%) | |
| 3 | 26 (52%) | 11 (27,5%) | 10 (50%) | 7 (35%) | 8 (40%) | 7 (35%) | 13 (41%) | |
| 4 | 16 (32%) | 14 (35%) | 2 (10%) | 6 (30%) | 4 (20%) | 4 (20%) | 6 (19%) | |
| 5 | 3 (6%) | 7 (17,5%) | 0 (0%) | 0 (0%) | 1 (5%) | 2 (10%) | 2 (6%) | |
| 6 | 1 (2%) | 6 (15%) | 3 (15%) | 2 (10%) | 2 (10%) | 1 (5%) | 3 (9%) | |
| 7 | 0 (0%) | 0 (0%) | 0 (0%) | 0 (0%) | 0 (0%) | 0 (0%) | 2 (6%) | |
| Frequência de evacuação | | | | | | | | |
| Evacuações / semana mediana (min-max) | 7 (2-21) | 7 (2-35) | 7 (1-21) | 7 (2-70) | 7 (3-14) | 7 (2-14) | 14 (1-28) | 0,036*a |
| Tabagismo (%) | | | | | | | | |
| Não | 50 (100%) | 37 (92,5%) | 15 (75%) | 17 (85%) | 18 (90%) | 20 (100%) | 29 (91%) | 0,015 |
| Sim | 0 (0%) | 3 (7,5%) | 5 (25%) | 3 (15%) | 2 (10%) | 0 | 3 (9%) | #a,b,c,d,e,f |
| Atividade física | | | | | | | | |
| Minutos por semana | 180 (40-840) | 165 (50- | 90 (40- | 180 (60- | 150 (45- | 180 (60- | 180 (30- | 0,715* |

| mediana (mín-máx) | | 270) | 240) | 840) | 360) | 840) | 480) | |
|---|----------|------------|----------|----------|----------|----------|----------|--------------|
| Atividade física >150 min/sem | | | | | | | | |
| Não | 29 (58%) | 34 (85%) | 19 (95%) | 17 (85%) | 14 (70%) | 14 (70%) | 24 (75%) | 0,012 #a,b,c |
| Sim | 21 (42%) | 6 (15%) | 1 (5%) | 3 (15%) | 6 (30%) | 6 (30%) | 8 (25%) | |
| Nível de estresse (n-%) | | | | | | | | |
| Baixo | 11 (22%) | 11 (27,5%) | 5 (25%) | 8 (40%) | 5 (25%) | 3 (15%) | 8 (25%) | 0,042 #f |
| Médio | 28 (56%) | 16 (40%) | 8 (40%) | 10 (50%) | 8 (40%) | 11 (55%) | 9 (28%) | |
| Alto | 10 (20%) | 7 (17,5%) | 5 (25%) | 2 (10%) | 5 (25%) | 4 (20%) | 8 (25%) | |
| Muito alto | 1 (2%) | 6 (15%) | 2 (10%) | 0 (0%) | 2 (10%) | 2 (10%) | 7 (22%) | |

Legenda: CT, controles; DII, doença inflamatória intestinal (a); PP, psoríase em placas (b); AR, artrite reumatoide(c); LES, Lupus eritematoso sistêmico (d); DM1, diabetes tipo 1 (e); DM2, diabetes tipo 2 (f); min, minutos; sem, semana; Teste Qui-quadrado #; Teste Kruskal-Wallis *; Nível de significância adotado $p < 0,05$.

Os hábitos alimentares estão descritos na **Tabela 5** e foram divididos nas categorias: consumo de alimentos saudáveis, consumo de alimentos menos saudáveis e consumo de álcool. Neste sentido, é possível observar que houve maior consumo de alimentos saudáveis no grupo CT apenas quando comparado ao grupo DII ($p=0,019$).

O grupo CT também diferiu significativamente dos grupos enfermos em relação ao consumo de açúcar ou adoçantes artificiais, que foram mais frequentes nos grupos DII, PP e LES. Já o consumo de ultra processados e industrializados foi diferente apenas na comparação de grupo CT vs. Grupo DM1 ($P=0,033$), que consumiu menos frequentemente este tipo de alimento.

Bebidas alcoólicas foram consumidas com maior frequência por indivíduos CT, que diferiram significativamente dos grupos DII, PP, AR e DM2 ($P=0,005$).

Tabela 5. Hábitos alimentares relacionados ao consumo de alimentos mais saudáveis, menos saudáveis e álcool.

| Variáveis | CT (n=50) | DII (n=40) a | PP (n=20) b | AR (n=20) c | LES (n=20) d | DM1 (n=20) e | DM2 (n=32) f | P-valor |
|--|--------------|-----------------|----------------|----------------|-----------------|-----------------|-----------------|------------------|
| Consumo de alimentos “saudáveis” | | | | | | | | |
| Consumo de Frutas, Legumes, Verduras e Grãos integrais / dia (%) | | | | | | | | |
| Menos de 1x | 1 (2%) | 7 (17.5%) | 4 (20%) | 3 (15%) | 1 (5%) | 2 (10%) | 5 (16%) | 0.019# a,f |
| 1 a 2 x | 22 (44%) | 21 (52.5%) | 7 (35%) | 12 (60%) | 10 (50%) | 11 (55%) | 18 (56%) | |
| 3 a 4 x | 19 (38%) | 10 (25%) | 6 (30%) | 3 (15%) | 7 (35%) | 7 (35%) | 7 (22%) | |
| Mais de 5 x | 8 (16%) | 2 (5%) | 3 (15%) | 2 (10%) | 2 (10%) | 0 (0%) | 2 (6%) | |
| Consumo de alimentos “menos saudáveis” | | | | | | | | |
| Preparações caseiras adicionadas de açúcar ou adoçante artificial / semana (%) | | | | | | | | |
| Nunca | 5 (10%) | 2 (5%) | 0 (0%) | 1 (5%) | 0 (0%) | 0 (0%) | 1 (5%) | 0.005 # a,b,d |
| Menos de 1 x | 7 (14%) | 2 (5%) | 0 (0%) | 1 (5%) | 1 (5%) | 1 (5%) | 2 (9%) | |
| 1 x | 1 (2%) | 0 (0%) | 0 (0%) | 1 (5%) | 3 (15%) | 0 (0%) | 0 (0%) | |
| 2 a 3 x | 3 (6%) | 5 (12.5%) | 1 (5%) | 1 (5%) | 3 (15%) | 5 (25%) | 3 (14%) | |
| 4 a 5 x | 10 (20%) | 0 (0%) | 1 (5%) | 1 (5%) | 0 (0%) | 3 (15%) | 1 (5%) | |
| Mais de 5 x | 24 (48%) | 31 (77.5%) | 18 (90%) | 15 (75%) | 13 (65%) | 11 (55%) | 15 (68%) | |
| Consumo de ultra processados e industrializados / semana (%) | | | | | | | | |
| Nunca | 1 (2%) | 1 (2.5%) | 1 (5%) | 0 (0%) | 0 (0%) | 3 (15%) | 0 (0%) | 0.033e, |
| Menos de 1 x | 9 (18%) | 8 (20%) | 3 (15%) | 3 (15%) | 5 (25%) | 4 (20%) | 6 (27.3%) | |
| 1 x | 0 (0%) | 3 (7.5%) | 0 (0%) | 1 (5%) | 0 (0%) | 2 (10%) | 0 (0%) | |
| 2 a 3 x | 22 (44%) | 11 (27.5%) | 4 (20%) | 6 (30%) | 6 (30%) | 4 (20%) | 6 (27.3%) | |
| 4 a 5 x | 7 (14%) | 4 (10%) | 5 (25%) | 2 (10%) | 2 (10%) | 4 (20%) | 0 (0%) | |
| Mais de 5 x | 11 (22%) | 13 (32.5%) | 7 (35%) | 8 (40%) | 7 (35%) | 3 (15%) | 10 (45.4%) | |
| Consumo de álcool (%) | | | | | | | | |
| Não | 10 (20%) | 26 (65%) | 13 (65%) | 15 (75%) | 9 (45%) | 9 (45%) | 22 (60%) | < .001 |
| Sim | 40 (80%) | 14 (35%) | 7 (35%) | 5 (25%) | 11 (55%) | 11 (55%) | 10 (31%) | #a,b,c,f |

Legenda: CT, controles; DII, doença inflamatória intestinal (a); PP, psoríase em placas (b); AR, artrite reumatoide(c); LES, Lupus eritematoso sistêmico (d); DM1, diabetes tipo 1 (e); DM2, diabetes tipo 2 (f), Teste Qui-quadrado # ; Teste Kruskal-Wallis *; Nível de significância adotado p<0,05

Baseado no resultado após aplicação do método de resíduos MSM, foram obtidos os valores de ingestão habitual de nutrientes baseados na aplicação de 3 registros alimentares de 24h para cada indivíduo, de cada grupo avaliado. As médias de cada grupo encontram-se descritas da **Tabela 6**.

Os nutrientes carboidratos, gordura poli-insaturada, fibras, sódio, zinco, vitamina A e vitamina D não apresentaram diferença significativa quando comparado ao grupo controle. Ainda que em alguns casos o p-valor pelo teste Teste Kruskal-Wallis tenha sido significativo, a ausência da inicial de identificação do grupo (por exemplo 0.046 **c**) indica que a diferença foi verificada entre os grupos doentes entre si e não *versus* controles. Com relação aos demais nutrientes, verificou-se diferença significativa entre controles e algum dos grupos enfermos avaliados.

Tabela 6. Ingestão habitual de calorias, macro e micronutrientes.

| Variáveis | CT (n=50) | DII (n=40) a | PP (n=20) b | AR (n=20) c | LES (n=20) d | DM1 (20) e | DM2 (n=32) f | p-valor |
|-------------------------|--------------|-----------------|----------------|----------------|-----------------|---------------|-----------------|------------------|
| Calorias | 1773.93 | 1718.67 | 1551.06 | 1341.00 | 1698.37 | 1609.13 | 1503.83 | 0.019 c,f |
| Proteínas | 84.26 | 66.86 | 63.565 | 63.25 | 74.97 | 73.24 | 70.31 | < .001 a,b,c,f |
| Carboidratos | 182.52 | 227.14 | 193.815 | 162.45 | 212.98 | 179.60 | 185.50 | 0.076 |
| Gordura total | 69.23 | 58.48 | 53.445 | 48.03 | 65.32 | 67.92 | 53.0 | 0.046 c,f |
| Gordura saturada | 23.78 | 20.97 | 17.040 | 17.47 | 20.46 | 22.55 | 18.28 | < .001 b,c,f |
| Gordura monoinsaturada | 21.09 | 17.80 | 15.220 | 16.46 | 18.28 | 22.86 | 17.26 | < .001 a,b,c,f |
| Gordura poli-insaturada | 15.40 | 13.26 | 15.625 | 11.56 | 14.23 | 15.89 | 13.75 | < .001f |
| Colesterol | 288.77 | 237.14 | 210.94 | 190.44 | 233.60 | 290.56 | 176.26 | < .001 a,b,c,d,f |

| | | | | | | | | |
|----------|---------|---------|---------|---------|---------|---------|---------|--------------------|
| Fibra | 18.47 | 19.24 | 18.350 | 16.745 | 18.07 | 21.46 | 18.15 | 0.439 |
| Sódio | 3162.52 | 2771.97 | 3214.55 | 2862.66 | 3413.25 | 3842.85 | 3231.1 | 0.034 |
| Calcio | 682.61 | 546.97 | 471.70 | 433.32 | 600.51 | 517.66 | 424.21 | < .001 a,b,c,f |
| Ferro | 8.17 | 7.77 | 7.30 | 5.83 | 8.16 | 9.44 | 6.63 | < .001 b,c,f |
| Zinco | 10.19 | 7.92 | 7.815 | 7.94 | 9.50 | 10.74 | 10.50 | < .001 |
| Vit A | 291.26 | 299.83 | 285.91 | 166.49 | 206.23 | 283.53 | 202.40 | < .001f |
| Vit B1 | 1.00 | 0.77 | 0.75 | 1.02 | 0.97 | 1.01 | 1.10 | < .001b |
| Vit C | 137.4 | 99.87 | 76.21 | 107.78 | 75.40 | 83.63 | 122.9 | < .001 b,d,f |
| Vit E | 1.88 | 1.78 | 0.910 | 0.83 | 1.86 | 1.88 | 1.17 | < .001 c |
| Vit B6 | 0.56 | 0.73 | 0.490 | 0.44 | 0.60 | 0.82 | 0.63 | < .001 e |
| Vit D | 1.76 | 1.85 | 0.98 | 1.31 | 1.510 | 2.17 | 2.13 | 0.002 |
| Potássio | 2279.88 | 1998.63 | 1805.52 | 1762.83 | 2109.49 | 1935.77 | 1996.55 | 0.003 a,b,c,f |
| Magnésio | 249.60 | 218.11 | 243.11 | 197.47 | 228.06 | 276.73 | 193.85 | 0.001 c,f |
| Selênio | 23.46 | 24.50 | 14.96 | 11.28 | 19.94 | 26.53 | 17.10 | < .001 b,c,f |
| Folato | 98.88 | 82.14 | 59.85 | 38.58 | 82.41 | 84.32 | 55.01 | < .001 b,c,d,f |
| Fosforo | 516.31 | 456.64 | 301.67 | 260.97 | 435.99 | 315.86 | 284.74 | < .001 b,c,f |
| Cobre | 2.47 | 0.55 | 0.46 | 0.68 | 0.43 | 0.21 | 0.23 | < .001 a,b,c,d,e,f |
| Vit B3 | 5.38 | 4.73 | 2.89 | 3.14 | 4.57 | 4.41 | 3.20 | < .001 b,c,f |
| Vit B12 | 1.80 | 1.69 | 1.15 | 0.72 | 1.72 | 1.22 | 1.10 | < .001 c,e,f |
| Manganês | 0.73 | 0.80 | 0.46 | 0.41 | 0.36 | 0.32 | 0.43 | < .001 d,f |

Após teste Kruskal-Wallis ($p < 0,05$) foi aplicado teste post-hoc Dunnett para identificação do grupo específico com diferença estatística. Legenda: CT, controles; DII, doença inflamatória intestinal (a); PP, psoríase em placas (b); AR, artrite reumatoide(c); LES, Lupus eritematoso sistêmico (d); DM1, diabetes tipo 1 (e); DM2, diabetes tipo 2 (f); Vit., Vitamina; Teste Kruskal-Wallis; Nível de significância adotado $p < 0,05$

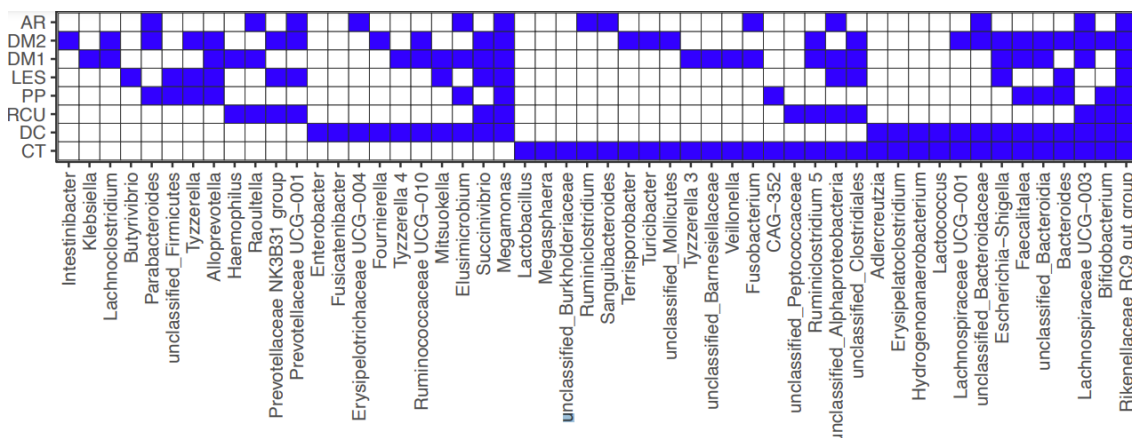
6.4 MODELOS PREDITIVOS PARA IDENTIFICAÇÃO DOS ESTADOS DE SAÚDE E DOENÇA

6.4.1 Seleção de variáveis candidatas

A identificação de marcadores microbianos que fariam parte dos modelos preditivos das doenças foi realizada utilizando-se dois critérios. O primeiro foi a pré-seleção das variáveis que apresentaram diferença significativa nos testes comparativos dos itens 6.2.2 e 6.2.3. O segundo critério foi aplicado para a seleção de marcadores mais amplos da microbiota intestinal, comparando todos os grupos, e não mais de forma par a par, com intuito de obter o maior número de táxons que pudesse compor os modelos preditivos.

Cinquenta e dois táxons bacterianos foram selecionados a partir desta comparação mais ampla. Abaixo (**tabela 8**), apresentamos de forma sumarizada qual táxon marcador pôde entrar no modelo preditivo para cada enfermidade.

Tabela 7. Táxons que apresentaram abundância diferencial na comparação entre todos os grupos doentes e grupo controle.

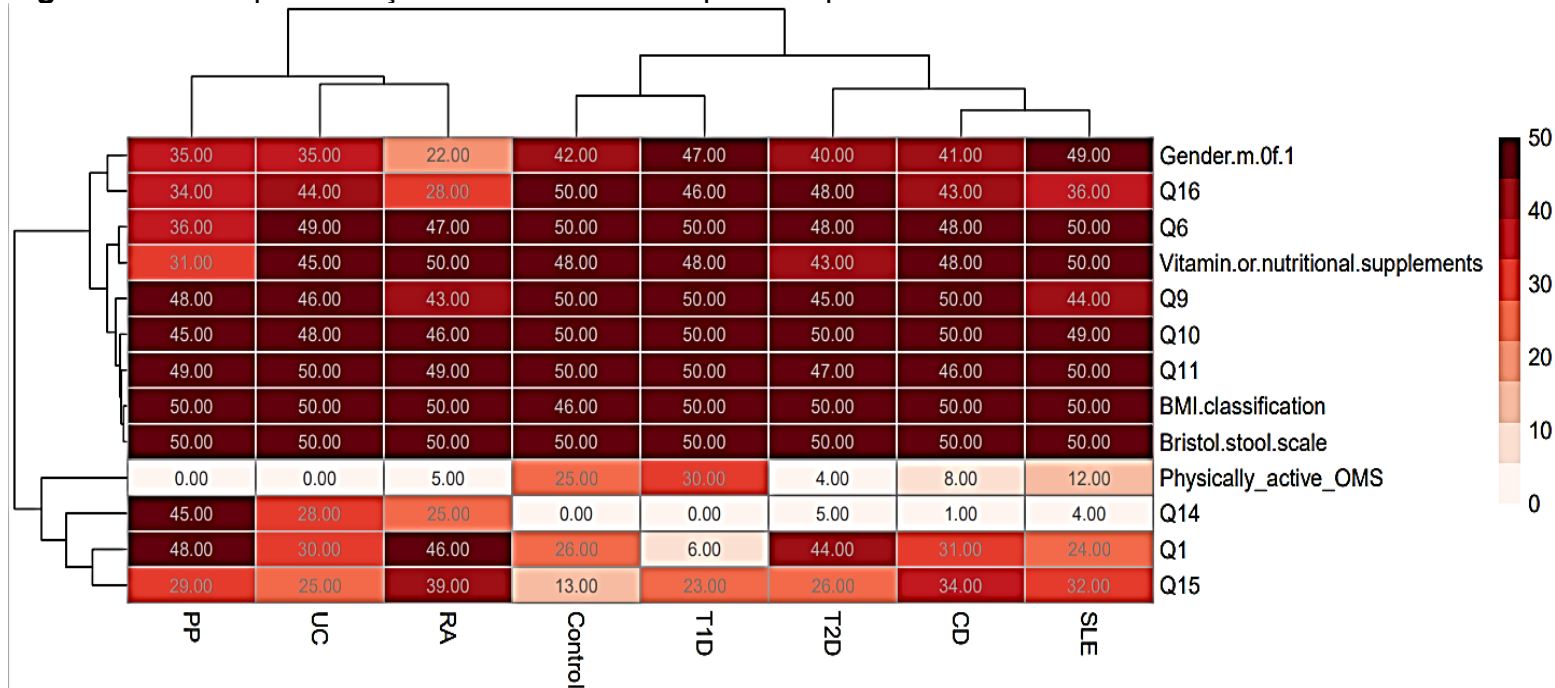


Legenda: CT, controles; DII, doença inflamatória intestinal; PP, psoríase em placas; AR, artrite reumatoide; LES, Lupus eritematoso sistêmico; DM1, diabetes tipo 1; DM2, diabetes tipo 2

Todas as variáveis pré-selecionadas nos testes par a par, e também as variáveis obtidas no teste geral da microbiota intestinal, foram então filtradas pela primeira análise de Floresta randômica. Essa etapa deu-se pela aplicação de 50 modelos, usando conjuntos de dados selecionados aleatoriamente para treinamento, a fim de identificar variáveis que figuravam entre as 10 mais importantes em pelo menos 50% dos modelos. As variáveis qualitativas, quantitativas (tabela 2 e apêndice 3) e da microbiota intestinal submetidas ao processo de seleção final encontram-se nas figuras 9, 10 e 11.

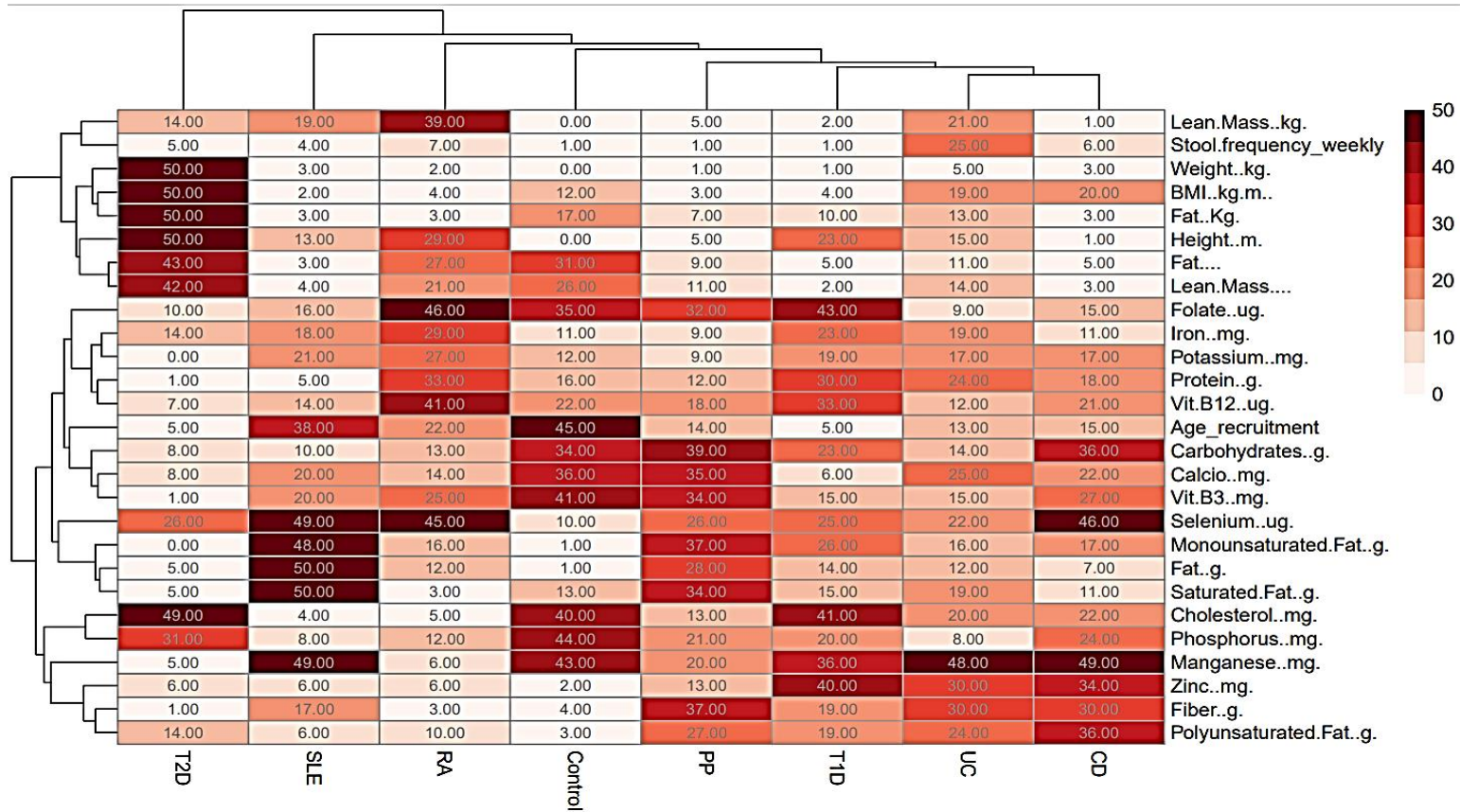
O número dentro de cada quadrado no mapa de calor indica a quantidade de vezes em que cada variável apareceu entre as mais importantes dentro das 50 rodadas de Floresta randômica. Variáveis que apareceram menos de 25 vezes não foram consideradas para o modelo preditivo final.

Figura 9 Heatmap de seleção de variáveis fenotípicas – qualitativas



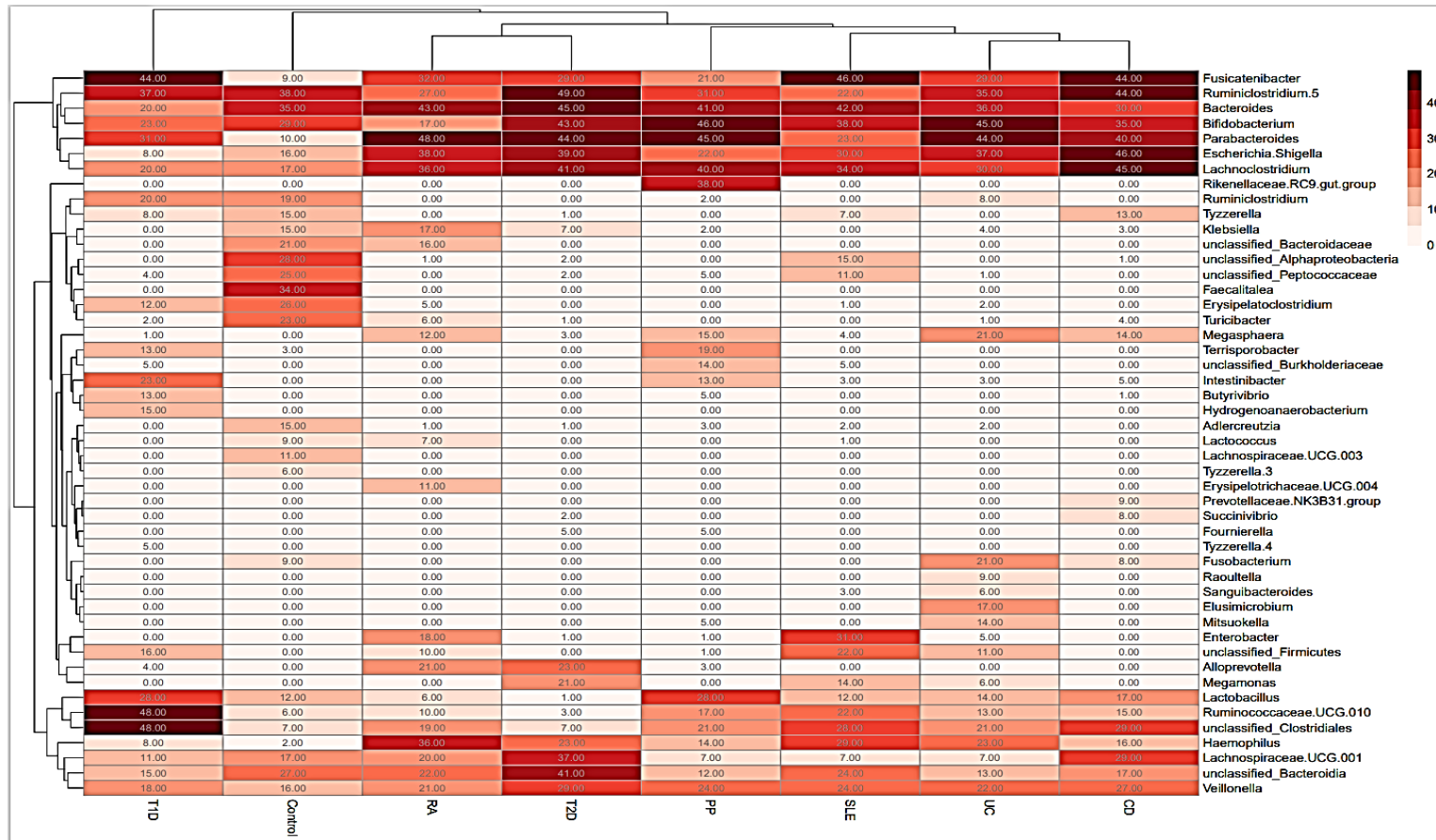
Legenda: CT, controles; CD: Doença de Crohn; UC: Colite ulcerativa; PP, psoríase em placas; AR, artrite reumatoide; LES, Lupus eritematoso sistêmico; DM1, diabetes tipo 1; DM2, diabetes tipo 2.

Figura 10 Heatmap de seleção de variáveis fenotípicas – quantitativas



Legenda: CT, controles CD: Doença de Crohn; UC: Colite ulcerativa; PP, psoríase em placas; AR, artrite reumatoide; LES, Lupus eritematoso sistêmico; DM1, diabetes tipo 1; DM2, diabetes tipo 2.

Figura 11 Heatmap de seleção de variáveis da microbiota intestinal



Legenda: CT, controles; CD: Doença de Crohn; UC: Colite ulcerativa; PP, psoríase em placas; RA, artrite reumatoide; LES, Lupus eritematoso sistêmico; DM1, diabetes tipo 1; DM2, diabetes tipo 2.

6.5 ANÁLISE DE PERFORMANCE PREDITIVA DOS MODELOS

A avaliação da performance preditiva dos modelos para cada doença foi testada em quatro etapas. **Etapa 1:** Baseada exclusivamente nos táxons bacterianos para verificar a capacidade isolada da MI em detectar alguma das doenças avaliadas. **Etapa 2 e 3:** Baseadas na capacidade das características fenotípicas qualitativas e quantitativas, respectivamente, para a detecção de alguma das doenças avaliadas. **Etapa 4:** Baseadas nas integrações das etapas 1, 2 e 3 para a detecção de uma das enfermidades avaliadas neste trabalho.

Para visualizar o potencial de táxons microbianos significativamente alterados como biomarcadores de enfermidades, modelos criados pelos algoritmos de Floresta Randômica e suas respectivas curvas de características operacionais do receptor (ROC) foram construídos. As curvas ROC criadas mostram a capacidade global de prever uma determinada doença usando um conjunto de dados. Cada curva apresenta a sensibilidade e especificidade para distinguir indivíduos com uma doença ou indivíduos do grupo controle.

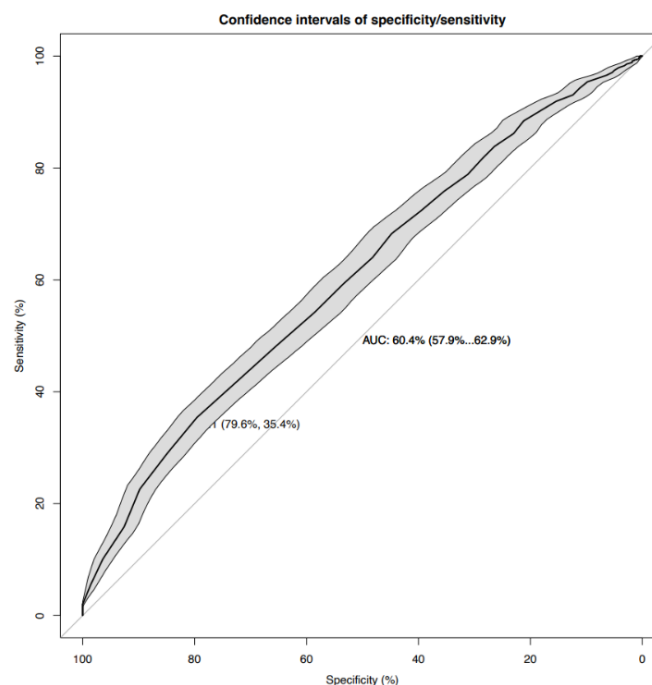
Já o índice Gini indica a importância de cada uma das variáveis no modelo global. Ou seja, cada variável apresenta seu valor específico (Mean Decrease Gini), ou valor de importância da variável no modelo, que indica quanto o modelo preditivo geral passa a errar a predição da enfermidade quando essa variável é retirada do modelo.

*6.4.4.1 ETAPA 1 - Capacidade de distinção da **microbiota intestinal isoladamente** para diferenciação de estados de doença e grupo controle.*

Para ilustrar o valor preditivo geral do microbioma associado às doenças, primeiro construímos modelos aleatórios de Floresta Randômica com base nos táxons intestinais pré-selecionados nos itens 6.2.2 e 6.2.3 e tabela 11. Cada modelo foi repetido cem vezes em cada subgrupo avaliado (ex. doença de Crohn versus controles sadios).

O modelo final levou a uma área sob a curva (AUC) das características operacionais do receptor (ROC), específica para cada enfermidade. Abaixo estão representadas as curvas ROC, utilizando dados exclusivos da microbiota intestinal para diferenciar indivíduos do grupo controle e seus respectivos Índices Gini: Doença de Crohn (**AUC = 0,60**; Sens. = 0,58; Espec.= 0,63) – figuras 12 e 13, Retocolite Ulcerativa (**AUC = 0,53**; Sens. = 0,50; Espec.= 0,56) – figuras 14 e 15, Psoríase em Placas (**AUC = 0,65**; Sens. = 0,626; Espec.= 0,68) – figuras 16 e 17, Artrite Reumatoide (**AUC = 0,54**; Sens. = 0,51; Espec.= 0,57) – figuras 18 e 19, Lupus Eritematoso sistêmico (**AUC = 0,54**; Sens. = 0,52; Espec.= 0,57) – figuras 20 e 21, Diabetes Tipo 1 (**AUC = 0,70**; Sens. = 0,67; Espec.= 0,72) - figuras 22 e 23; Diabetes Tipo 2 (**AUC = 0,672**; Sens. = 0,65; Espec.= 0,69) – figuras 24 e 25.

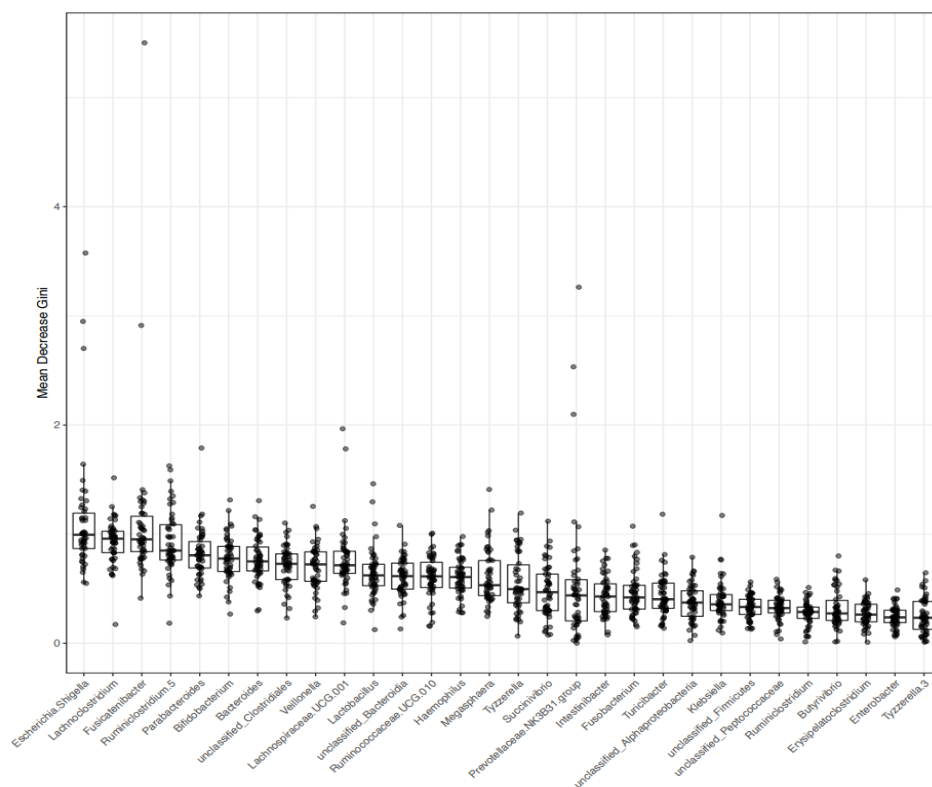
Figura 12. Curva ROC, considerando a **microbiota intestinal isoladamente** como preditor da doença de Crohn.



Legenda: AUC, área sob a curva. Os intervalos de confiança de 95% (ICs) são mostrados como áreas sombreadas. A linha diagonal no gráfico marca uma AUC de 0,5 (ou seja, classificação aleatória).

É possível observar que os 5 táxons bacterianos mais importantes no modelo preditivo de **Doença de Crohn (figura 13)** foram *Escherichia-Shiguella*; *Lachnoclostridium*; *Fusicatenibacter*; *Ruminiclostridium 5* e *Parabacteroides*. cada um contribuindo com uma importância de 1% ou menos, na capacidade do modelo.

Figura 13. Valores índice Gini de diferentes táxons microbianos e seu impacto na previsão do status da doença de Crohn.

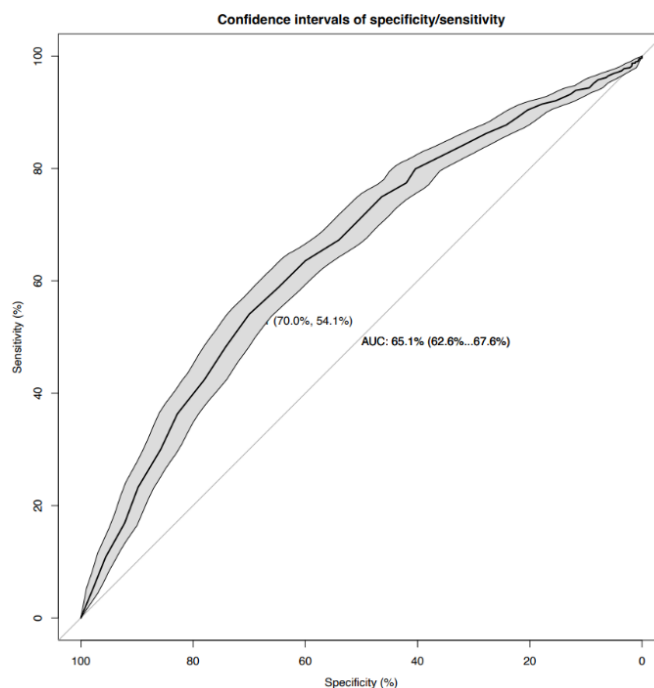


Na **figura 15** é possível observar que os 5 táxons bacterianos mais importantes no modelo preditivo de **Retocolite Ulcerativa** foram Bifidobacterium; Parabacteroides; Escherichia-Shiguella; Bacteroides e Ruminiclostridium 5, cada um contribuindo com uma importância de no máximo 1% para a capacidade do modelo.

Já para o modelo preditivo de **Psoríase em placas (figura 17)**, Bifidobacterium; Parabacteroides; Richinellaceae RC9 gut group; Lachnoclostridium e Bacteroides foram mais impactantes, sendo o primeiro responsável pela importância de 1% na capacidade do modelo e os demais contribuindo com percentuais menores que 1%.

Para os casos de **Artrite Reumatoide**, Parabacteroides, Bacteroides, Haemophilus, Lachnoclostridium, Escherichia-shiguella foram os táxons bacterianos mais relevantes (**figura 19**).

Figura 16. Curva ROC, considerando a microbiota intestinal isoladamente como preditor de Psoríase em Placas.



Legenda: AUC, área sob a curva. Os intervalos de confiança de 95% (ICs) são mostrados como áreas sombreadas. A linha diagonal no gráfico marca uma AUC de 0,5 (ou seja, classificação aleatória).

Figura 17. Valores índice Gini de diferentes táxons microbianos e seu impacto na previsão do status da Psoríase em Placas.

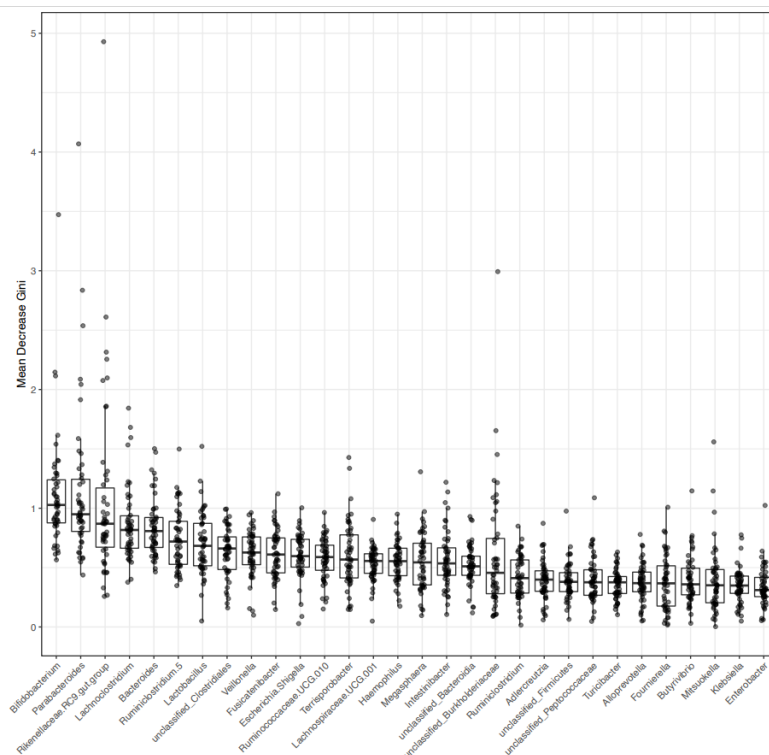
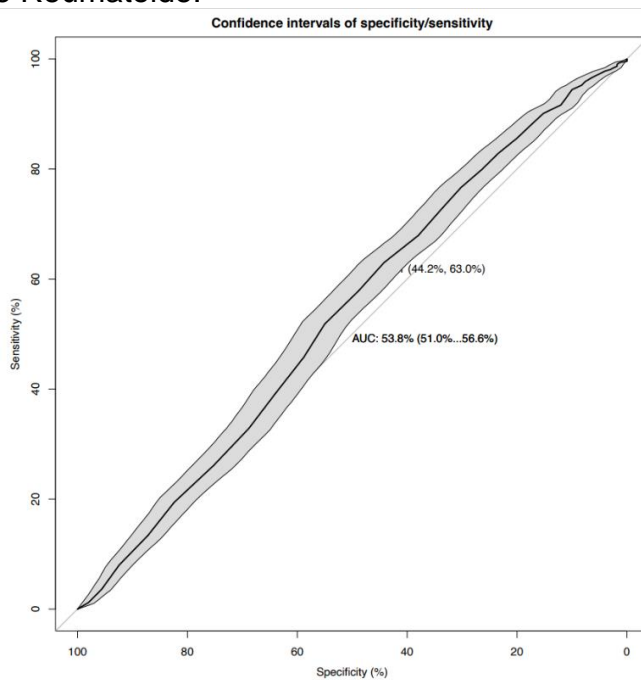


Figura 18. Curva ROC, considerando a microbiota intestinal isoladamente como preditor de Artrite Reumatoide.



Legenda: AUC, área sob a curva. Os intervalos de confiança de 95% (ICs) são mostrados como áreas sombreadas. A linha diagonal no gráfico marca uma AUC de 0,5 (ou seja, classificação aleatória).

Figura 19. Valores índice Gini de diferentes táxons microbianos e seu impacto na previsão do status da Artrite Reumatoide

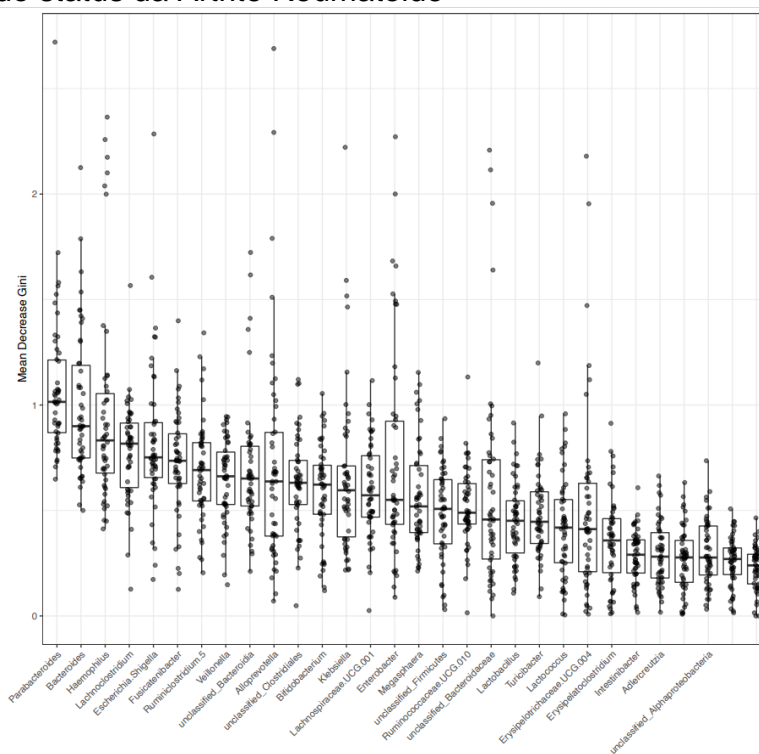
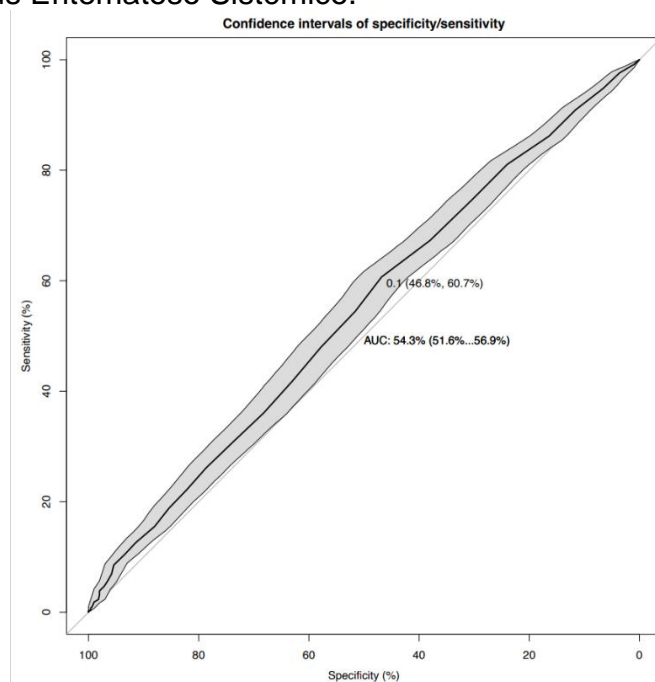


Figura 20. Curva ROC, considerando a microbiota intestinal isoladamente como preditor de Lupus Eritematoso Sistêmico.



Legenda: AUC, área sob a curva. Os intervalos de confiança de 95% (ICs) são mostrados como áreas sombreadas. A linha diagonal no gráfico marca uma AUC de 0,5 (ou seja, classificação aleatória).

Figura 21. Valores índice Gini de diferentes táxons microbianos e seu impacto na previsão do status da Lupus Eritematoso Sistêmico.

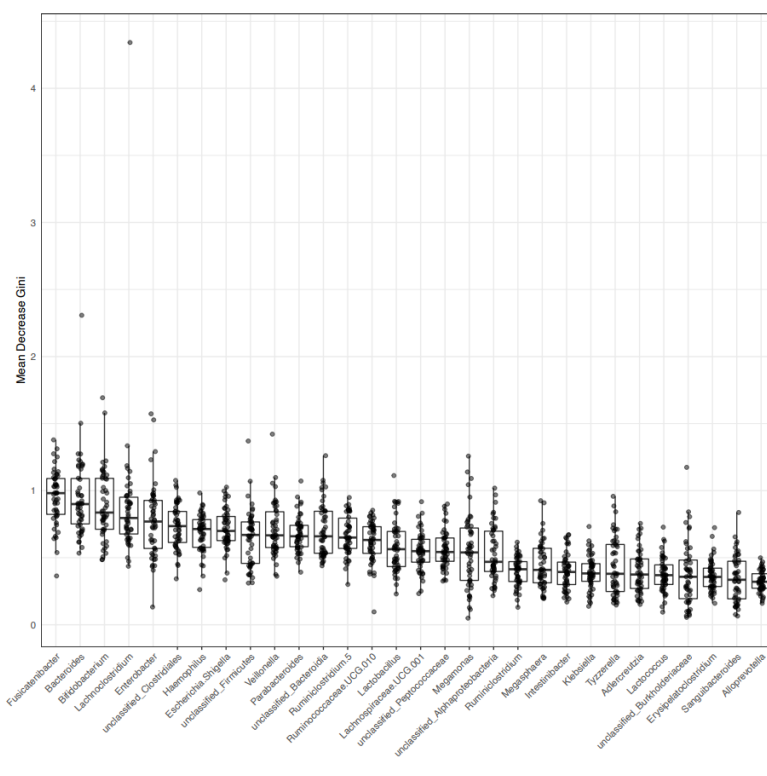
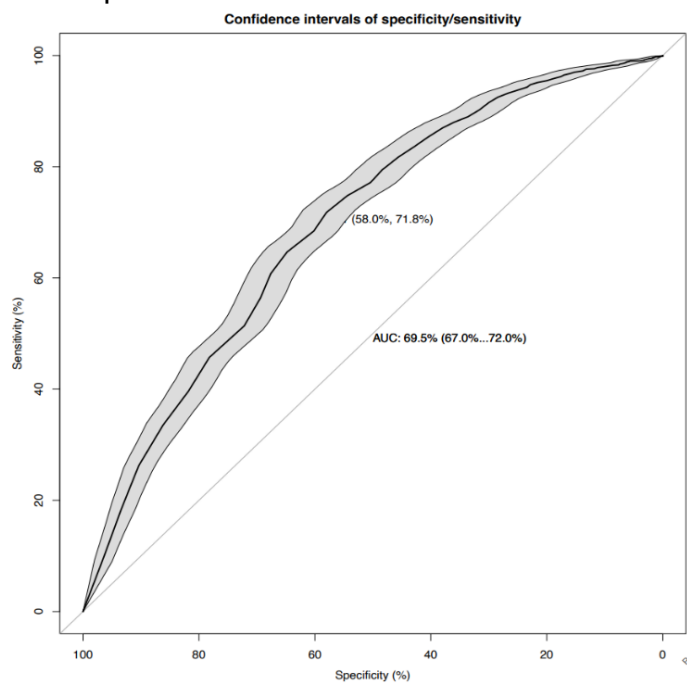


Figura 22. Curva ROC, considerando a microbiota intestinal isoladamente como preditor de Diabetes Tipo 1.



Legenda: AUC, área sob a curva. Os intervalos de confiança de 95% (ICs) são mostrados como áreas sombreadas. A linha diagonal no gráfico marca uma AUC de 0,5 (ou seja, classificação aleatória).

Figura 23. Valores índice Gini de diferentes táxons microbianos e seu impacto na previsão do status da Diabetes Tipo 1.

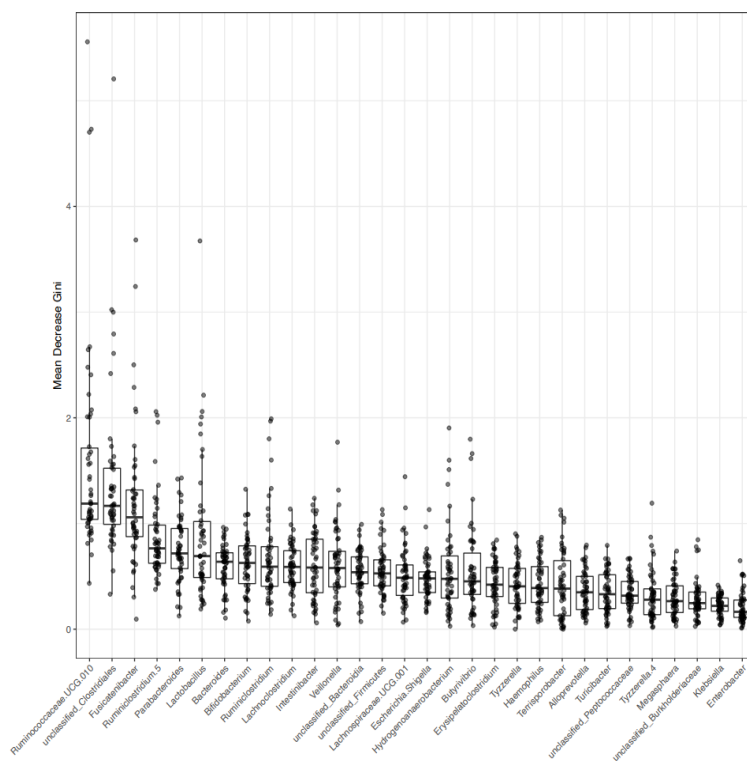
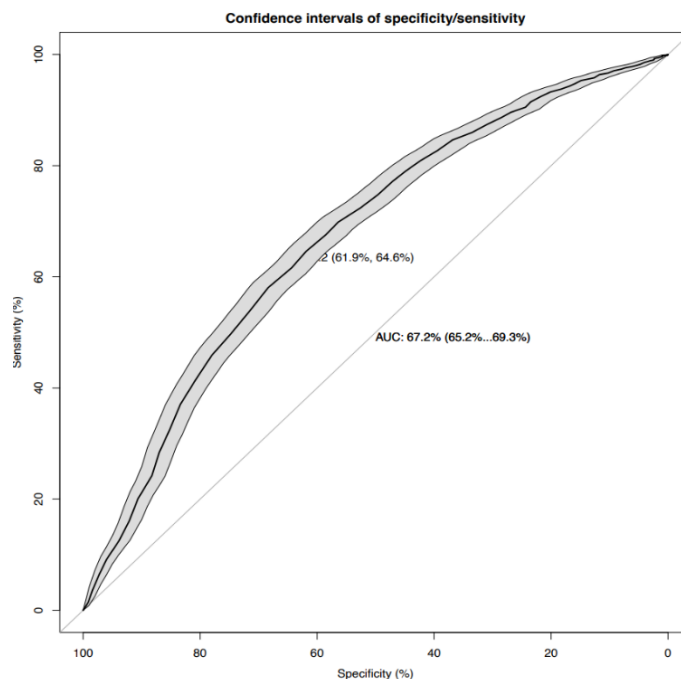
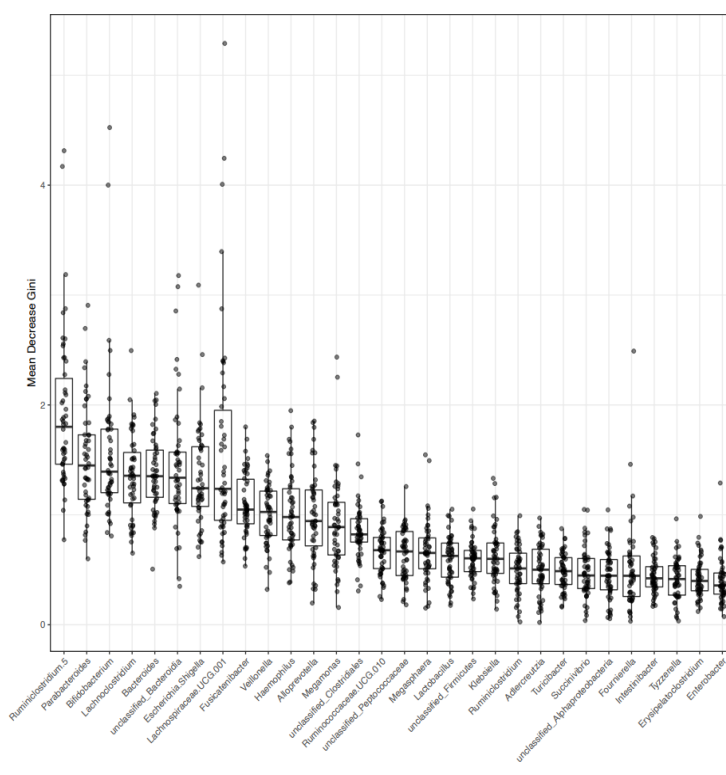


Figura 24. Curva ROC, considerando a microbiota intestinal isoladamente como preditor de Diabetes Tipo 2.



Legenda: AUC, área sob a curva. Os intervalos de confiança de 95% (ICs) são mostrados como áreas sombreadas. A linha diagonal no gráfico marca uma AUC de 0,5 (ou seja, classificação aleatória).

Figura 25. Valores índice Gini de diferentes táxons microbianos e seu impacto na previsão do status da Diabetes Tipo 2.



Nas **figuras 21, 23 e 25**, identificamos os táxons mais importantes no modelo preditivo de **LES, DM1 e DM2, respectivamente**. Para LES, Fusicatenibacter, Bacteroides, Bifidobacterium, Lachnospirillum e Enterobacter foram os cinco mais importantes. No entanto, nenhuma bactéria identificada atingiu o nível de pelo menos 1%.

No modelo preditivo de **DM1**, modelo que apresentou melhor desempenho entre todas as enfermidades quando consideramos a microbiota como fator isolado, teve como táxons mais impactantes: Ruminococcaceae UCG 010, Clostridiales, Fusicatenibacter, Ruminococcidium 5 e Parabacteroides. Os três primeiros táxons participaram com percentual acima de 1% de importância.

Por fim, no modelo de **DM2**, modelo com segundo melhor desempenho, contou com Ruminococcidium 5, Parabacteroides, Bifidobacterium, Lachnospirillum e Bacteroides. Nesse modelo, Ruminococcidium 5 contribuiu com aproximadamente 1,5% de impacto e os demais táxons contribuíram com aproximadamente 1% cada um.

De todos os modelos preditivos baseados exclusivamente em táxons bacterianos, os que obtiveram melhor performance foram Diabetes Tipo 1 (**AUC = 0,695**) seguido de Diabetes Tipo 2 (**AUC = 0,672**) e Psoríase em Placas (**AUC = 0,651**), todos com AUC com discriminação regular.

6.4.4.2 ETAPAS 2 e 3 - Capacidade de distinção das variáveis fenotípicas (qualitativas e quantitativas) isoladamente para diferenciação de estados de doença e grupo controle.

Embora não seja objetivo específico do presente trabalho conhecer a capacidade das variáveis fenotípicas em predizer os estados de doença, foi realizada análise de performance desse aspecto como um passo necessário para a integração dessas variáveis fenotípicas com a microbiota em conjunto.

Para ambas as variáveis, qualitativas e quantitativas, a enfermidade que obteve o modelo com melhor desempenho para LES (AUC = 0,87; AUC= 0,98), seguido por Diabetes tipo 2 (AUC = 0,78; AUC = 0,97), apresentando discriminação dos grupos entre boa e excelente. Os dados completos referente a curvas ROC e seus índices Gini estão graficamente representados nos apêndices 7 e 8.

*6.4.4.3 ETAPA 4 - Capacidade de distinção de todas as **variáveis fenotípicas e de microbiota intestinal integradas** para diferenciação de estados de doença e grupo controle.*

A etapa 4 é o ponto chave para responder o objetivo principal do presente estudo: “Estabelecer modelos preditivos que integrem variáveis fenotípicas com a microbiota intestinal para distinguir indivíduos saudáveis e enfermos”. Nessa etapa, foram agrupadas tanto as variáveis microbianas quanto as variáveis fenotípicas qualitativas e quantitativas.

A integração de dados microbianos e fenotípicos favoreceu o melhor desempenho de todos os modelos (**Tabela 8**), quando comparamos o desempenho versus a microbiota isolada. As variáveis que compuseram cada modelo e seus respectivos índice Gini estão representados de maneira completa no **apêndice 9**.

Tabela 8. Comparação do desempenho dos modelos preditivos de doença com microbiota isolada e integrada.

| Doença | AUC isolada (nível de discriminação) | AUC integrada (nível de discriminação) | Ordem de importância de biomarcadores |
|--------|---|---|--|
| DC | 0,604 (regular) | 0,781 (boa) | 1 Manganês (cn) 2 Selênio (cn) 3 <i>Escherichia-Shiguela</i> (tb) 4 Fusicatenibacter (tb) 5 Lachnoclostridium (tb) 6 Carboidratos (cn). 7 Gordura poliinsat. (cn) 8 Vitamina B3 (cn) 9 Lachnospiraceae.UGC.001 (tb) 10. Zinco |
| RCU | 0,527 (fraca) | 0,579 (fraca) | 1 Manganês (cn) 2 Bifidobacterium (tb) 3 Fibras (cn) 4 Zinco (cn) 5 Ruminiclostridium5 (tb) 6 Parabacteroides (tb) 7 Lachnoclostridium (tb) 8 <i>Escherichia.Shigella</i> (tb) 9 Bacteroides (tb) 10 Fusicatenibacter (tb) |
| PP | 0,651 (regular) | 0,763 (boa) | 1 Bifidobacterium (tb) 2 Parabacteroides (tb) 3 Carboidratos (cn) 4 Gordura monoinsat. (cn) 5 Rikenellaceae.RC9 (tb) 6 Gordura saturada (cn) 7 Fibra (cn) 8 Selênio (cn) 9 Vitamina B3 (cn) 10 Calcio (cn) |
| AR | 0,543 (fraca) | 0,981 (excelente) | 1 Vitamina B12 (cn) 2 Folato (cn) 3 Selênio (cn) 4 Massa Magra (cc) 5 Proteína (cn) 6 Massa gorda (cc) 7 Ferro (cn) |

| | | | |
|----------------|-----------------|-------------------|-----------------------------------|
| LES | 0,538(fraca) | 0,881 (excelente) | 8 Altura (cc) |
| | | | 9 Potássio (cn) |
| | | | 10 Bacteroides (tb) |
| | | | 1 Gordura total. (cn) |
| | | | 2 Selênio (cn) |
| | | | 3 Gordura saturada (cn) |
| | | | 4 Gordura monoinsat. (cn) |
| | | | 5 Manganês (cn) |
| | | | 6 Suplemento de vitaminas (cn) |
| | | | 7 Idade (cc) |
| DM1 | 0,695 (regular) | 0,856 (excelente) | 8 Bacteroides (tb) |
| | | | 9 Fusicatenibacter (tb) |
| | | | 10 Bifidobacterium (tb) |
| | | | 1 Folato (cn) |
| | | | 2 Colesterol (cn) |
| | | | 3 Selênio (cn) |
| | | | 4 unclassified_Clostridiales (tb) |
| | | | 5 Zinco (cn) |
| | | | 6 Manganês (cn) |
| | | | 7 Proteína (cn) |
| DM2 | 0,672 (regular) | 0,978 (excelente) | 8 Gordura monoinsat. (cn) |
| | | | 9 Fusicatenibacter (tb) |
| | | | 10 Vitamina B12 (cn) |
| | | | 1 IMC (cc) |
| | | | 2 Peso (cc) |
| | | | 3 Massa gorda (cc) |
| | | | 4 Colesterol (cn) |
| | | | 5 Massa Magra (cc) |
| | | | 6 Massa gorda (cc) |
| | | | 7 Fósforo (cn) |
| 8 Selênio (cn) | | | |
| DM2 | 0,672 (regular) | 0,978 (excelente) | 9 Parabacteroides (tb) |
| | | | 10 unclassified_Bacteroidia (tb) |

Legenda: nível de discriminação das AUC de acordo com Hosmer e Lemeshow, 2000. AUC: área sob a curva; DC: Doença de Crohn; RCU: Retocolite ulcerativa; PP, psoríase em placas; AR, artrite reumatoide; LES, Lupus eritematoso sistêmico; DM1, diabetes tipo 1; DM2, diabetes tipo 2; cn: consumo de nutriente; tb.: táxon bacteriano; cc: composição corporal ou clínica. Na tabela, são descritas as 10 variáveis de maior importância para os modelos preditivos.

É digno de nota que, independentemente do desempenho do modelo, para cada enfermidade houve uma característica diferente de táxons bacterianos e sua importância (índice Gini).

Na doença de Crohn, dez táxons compuseram o modelo integrado (Escherichia-Shiguela, Fusicatenibacter, Lachnoclostridium e Lachnospiraceae. UCG.001; Ruminiclostridium.5, Bifidobacterium, Parabacteroides, unclassified_Clostridiales, Bacteroides e Veillonella), sendo

que a variável de maior importância foi o manganês, característica do consumo de nutrientes. Em comparação ao modelo isolado da microbiota intestinal, quando integrado às outras variáveis fenotípicas, obteve-se melhor desempenho, de regular para bom.

Na RCU, seis táxons compuseram o modelo final (*Ruminiclostridium 5*, *Parabacteroides*, *Lachnoclostridium*, *Escherichia.Shigella*, *Bacteroides* e *Fusicatenibacter*). Para esse modelo novamente a variável de maior importância foi o manganês e, neste caso, o desempenho do modelo se manteve fraco.

Comparado ao modelo isolado, o modelo integrado para PP também demonstrou melhoria de desempenho de regular para bom. Os táxons bacterianos que compuseram o modelo final foram *Bifidobacterium*, *Parabacteroides* e *Rikenellaceae.RC9*, *Lactobacillus*, *Lachnoclostridium*, *Bacteroides* e *Ruminiclostridium 5*, sendo o gênero bacteriano *Bifidobacterium* a variável de maior importância para o modelo.

Para o modelo de Artrite Reumatoide, sete táxons compuseram o modelo final, entre eles: *Bacteroides*, *Parabacteroides*, *Lachnoclostridium*, *Haemophilus*, *Escherichia.Shigella*, *Fusicatenibacter* e *Ruminiclostridium.5*. Entretanto, apenas um deles figurou entre os dez itens mais importantes do modelo. O modelo integrado de AR obteve desempenho excelente, tendo como fator mais importante o consumo de vitamina B12.

O modelo preditivo de LES contou com os táxons bacterianos (*Bacteroides*, *Fusicatenibacter*, *Bifidobacterium*, *Lachnoclostridium*, *unclassified_Clostridiales*, *Enterobacter*, *Escherichia.Shigella* e *Haemophilus*). A bactéria de maior importância foi *Bacteroides* e foi a oitava variável em ordem de importância na modelagem.

Na análise do desempenho dos modelos de DM1 e DM2, ambos obtiveram melhora de performance quando integrados (o desempenho foi de regular para excelente). Em DM1, 6 táxons (*unclassified_Clostridiales*, *Ruminococcaceae.UCG.010*, *Fusicatenibacter*, *Parabacteroides*, *Ruminiclostridium.5* e *Lactobacillus*) compuseram o modelo integrado final.

Unclassified_Clostridiales alcançou importância de aproximadamente 1,5% no modelo, ocupando a terceira posição no ranking do índice Gini. Em DM2, Parabacteroides ocupou a décima colocação do índice Gini. Ao todo 10 bactérias compuseram o modelo integrado, mas as variáveis mais importantes para o modelo foram as características de composição corpórea como IMC, peso, massa gorda e massa magra e variáveis do consumo alimentar.

Em conjunto, nossos resultados apontam a utilidade de modelar de maneira integrada as informações fenotípicas do indivíduo e características microbianas como forma de melhorar a predição de condições de saúde.

7 DISCUSSÃO

7 DISCUSSÃO

No presente estudo, nosso principal interesse foi compreender a capacidade da microbiota intestinal, isolada e associada a informações fenotípicas, em identificar pacientes com diferentes condições clínicas por meio de aprendizado de máquina. Anteriormente, alguns estudos já apontaram os potenciais benefícios da utilização dessa forma de avaliação da microbiota intestinal, em vez da tradicional avaliação do estado genérico de “disbiose” onde o foco é apenas as bactérias que apresentam abundância relativa diferente na comparação do grupo controle *versus* uma determinada condição.¹⁷⁸⁻¹⁸¹

O presente trabalho incluiu uma amostra total de 202 indivíduos. Para cada enfermidade incluímos cerca de 20 participantes em cada subgrupo. O tamanho amostral em estudo com microbiota intestinal é muito variável. Quando desenhamos esse trabalho, a literatura disponível contou com essa média de pacientes em suas coortes.^{88, 152-154}

Um aspecto inusitado em nosso estudo é a variabilidade do perfil fenotípico de pacientes em nossa coorte. Habitualmente, os trabalhos optam por filtrar uma amostra pareada por idade, gênero, etnia, IMC e hábitos alimentares, por serem considerados fatores potencialmente impactantes para a microbiota intestinal.⁴³ Consideramos que a variabilidade amostral de nossa população é uma vantagem do presente trabalho, uma vez que as diferenças observadas nessas características é que permitiram a seleção de variáveis fenotípicas que compuseram nossos modelos preditivos. Essas diferenças permitiram, dentro de nosso conjunto de dados, identificar o que foi mais relevante para prever um desfecho de saúde, característica considerada importante em uma revisão sobre “considerações experimentais e computacionais para apoiar à reprodutibilidade de estudos com microbiomas”.¹⁸² Ademais, em análise paralela em nossa amostra, verificamos que o único fator com associação significativa com a macroestrutura da MI foi o estado de saúde ou doença (**apêndice 10**)

Em nosso trabalho, mostramos que em termos de estrutura ecológica (alfa diversidade), a microbiota intestinal não apresentou diferenças robustas. Nessa

análise foi possível observar diferença no índice Chao1 somente ao comparar indivíduos do grupo controle (GC) e indivíduos do grupo DII (**gráfico 1**). Isso também foi observado em relação a macroestrutura (beta diversidade) da MI, que indicou ampla sobreposição entre todos os grupos de diferentes tipos de doença (**figura 8**). Ainda que sobreposta, a macroestrutura (beta diversidade) do grupo DII apresentou diferença quando comparada ao GC. Esses achados corroboram com o estudo de Clooney e colaboradores¹⁸³, que observaram que independentemente da atividade da doença, pacientes com DII apresentam menores índices de diversidade alfa e beta.

No que se refere a outras enfermidades, a literatura é mais diversa. Por exemplo, em LES, a diversidade beta também não mostrou um padrão de agrupamento claro e distinto entre os grupos LES e GC no estudo de Toumi et al.¹¹². No entanto, a alfa diversidade apontou diferença entre os grupos, enquanto nosso trabalho não identificou esse cenário para nenhuma das medidas ecológicas aplicadas.

Avanços recentes no sequenciamento do genoma e na análise bioinformática permitiram maior compreensão das complexas associações hospedeiro-microbioma. No entanto, desde a primeira publicação sobre as diferenças do microbioma intestinal, para a maioria das condições clínicas,^{82,83,110,112,184} tem havido grande heterogeneidade de informações sobre a composição bacteriana nessas enfermidades, e a ausência de um perfil saudável da MI não facilita o cenário. O conhecimento sobre o perfil microbiano, per se, traz pouca colaboração para o cenário complexo das doenças autoimunes e metabólicas que, em sua maioria, representam doenças multifatoriais. Nesse sentido, conhecer as características fenotípicas, principalmente aquelas relacionadas à dieta, estilo de vida e características da composição corpórea se tornam importante para se relacionarem aos dados microbianos.

Em revisão recente, Wilkinson e colaboradores¹⁸⁵ destacaram o grande potencial do estudo da microbiota humana em escala populacional. Os autores destacam em seu artigo que grande parte do potencial do microbioma humano

para saúde pública está na possibilidade da descoberta de novos biomarcadores, terapias ou mecanismos moleculares que podem ser identificados em estudo observacionais com seres humanos e depois caracterizados com mais detalhes em condições experimentais.

O aprendizado de máquina, ou *Machine Learning* (ML), tem sido considerado fundamental para a análise de dados em microbioma que auxiliam a melhor reprodutibilidade dos trabalhos, criação de modelos preditivos para possivelmente futuro uso diagnóstico, bem como servirem como embasamento para intervenções clínicas baseadas em microrganismos, ou seja, direcionadas para a modulação destes.¹⁸⁶⁻¹⁹¹

Modelos preditivos em estudos com MI têm sido empregados para verificação da acurácia de sua utilização para fins diagnósticos (**tabela 9**) e predição de resposta a tratamentos. Algoritmos de árvore de decisão são abordagem comumente usada para essa finalidade, uma vez que árvores de decisão são capazes de inferir a variável de desfecho a partir de informações oferecidas para treinar o modelo.¹⁹⁰ A árvore de decisão constrói uma base para o emprego de outros métodos de aprendizado de conjunto, como a floresta randômica (ou *random forest*).

Floresta randômica (FR) é um tipo de algoritmo de agregação bootstrap, que já tem sido aplicado em estudos de microbiota intestinal^{121,192,193}. Algoritmos de FR têm sido utilizados para classificação de pacientes pediátricos com doença de Crohn (DC), de acordo com o estado da doença e a resposta ao tratamento, usando a diversidade alfa das amostras e o escore de risco genético (GRS) de cada paciente. Eles encontraram maior precisão de classificação com conjuntos de dados 16S rRNA do que a metagenômica shotgun.¹⁹⁴

Em FR, vários modelos de árvores de decisão são treinados ao mesmo tempo, com subconjunto dos dados (como uma coorte de treinamento). Em seguida, essas informações são inseridas no conjunto de dados maior para testar a sua capacidade final de predição.¹⁹⁵ Em nosso trabalho, utilizamos

algoritmos de FR para obter os modelos preditivos isolados e integrados para as enfermidades avaliadas em nossa coorte.

O trabalho de Ananthakrishnan e colaboradores,¹⁹⁶ que utilizou características do microbioma intestinal para prever resposta à terapia biológica com vedolizumabe em pacientes com doença de Crohn e retocolite ulcerativa, utilizou metodologia semelhante ao nosso trabalho. Primeiramente, Ananthakrishnan et al desenvolveu um modelo utilizando isoladamente os dados clínicos dos participantes. Esse modelo obteve desempenho apenas regular para prever os desfechos esperados (Ananthakrishnan teve AUC= 0,62 regular). Em contraste, o uso de táxons microbianos obteve AUC= 0,72 (capacidade boa), e quando o modelo foi integrado e associou dados clínicos e táxons da microbiota intestinal, obteve-se o melhor desempenho (AUC= 0,78 boa).

Embora o estudo anteriormente citado¹⁹⁶ tenha como objetivo outro desfecho, semelhanças com nosso estudo podem ser discutidas, pois em ambos os estudos a associação de dados clínicos e informações relativas a MI obtiveram melhor desempenho que o uso isolado dessas informações. Em nossa coorte, as informações microbianas isoladamente demonstram desempenho regular (AUC = 0,60). enquanto o modelo que integrou os dados fenotípicos e da microbiota obteve desempenho de AUC = 0,78, o mesmo obtido pelo outro autor.¹⁹⁶ Já os modelos preditivos para RCU em nossa coorte obtiveram desempenho próximo ao aleatório, com desempenho fraco tanto isoladamente quanto de forma integrada (AUC= 0,53 e AUC=0,58, respectivamente), sugerindo que, para o perfil de indivíduos de nossa coorte, a microbiota intestinal pode não ser o melhor marcador para diferenciar indivíduos sadios de indivíduos com todas as enfermidades.

No trabalho de Wu e colaboradores¹⁹⁷, os pesquisadores avaliaram o papel preditivo da microbiota intestinal para diferenciação entre quatro condições diferentes: artrite reumatoide, diabetes tipo 2, cirrose hepática e ausência de doença (controles saudáveis) em indivíduos chineses utilizando diversos algoritmos diferentes. A área sob a curva ROC (AUC) obtida por eles, considerando a análise metagenômica da microbiota, foi de AUC=0,81 para

DM2, AUC=0,94 para artrite reumatoide e AUC= 0,83 em casos de cirrose hepática.

Os táxons microbianos mais relevantes no modelo de AR contaram com *Veillonella*, *Haemophilus*, *Clostridium*, *Faecalibacterium*, *Blautia*, *Coprococcus*, *Enterococcus*, *Eubacterium* e *Roseburia*. Em nosso trabalho, também observamos a presença de *Veillonella* e *Haemophilus* na composição do modelo preditivo. No entanto, as demais bactérias do estudo de Wu não foram identificadas como importantes em nosso modelo. Ainda na identificação de AR¹⁹⁷, os autores ressaltam que neste trabalho nenhum biomarcador foi específico para essa enfermidade, ou seja, os mesmos táxons integraram modelos capazes de identificar diferentes doenças. o cenário ocorreu em nosso estudo. Por fim, os autores relataram que o adequado desempenho de seus modelos preditivos possivelmente deve-se ao maior emprego de recursos quando comparado a estudos anteriores, bem como pela utilização de informações clínicas integradas, o que apoia a ideia desenvolvida na presente tese.

Em Lupus, dois trabalhos ^{198,199} recentes debruçaram-se sobre o tema para investigar o papel preditivo da MI para detecção da condição clínica. Wei et al.¹⁹⁸ avaliaram a capacidade preditiva dos gêneros *Streptococcus*, *Prevotella_9*, *Roseburia*, *Ruminococcaceae_UCG-003*, *Ruminococcaceae_NK4A214_grupo*, *Paraprevotella*, *Ruminococcaceae_UCG-013*, *Ezakiella* e *Porphyromonas* como potenciais biomarcadores em LES. Esses nove gêneros compuseram um modelo preditivo que teve boa capacidade (AUC = 0,74) para diferenciar pacientes com LES de controles saudáveis. Em desacordo com o trabalho de Wei et al, nosso trabalho, ao considerar a MI como variável isolada, obteve desempenho próximo ao aleatório (AUC= 0,54). As bactérias que compuseram o nosso modelo, a saber *Fusicatenibacter*, *Bacteroides*, *Bifidobacterium*, *Lachnoclostridium*, *Enterobacter*, *unclassified_Clostridiales*, *Haemophilus*, *Escherichia.Shigella*, *unclassified_Firmicutes* e *Veillonella*, também diferem completamente das identificadas no estudo de Wei¹⁹⁸.

No trabalho de Li et al.¹⁹⁹, os autores buscaram diferenciar os pacientes com LES de controles saudáveis e de pacientes com AR baseando-se no gênero e nas espécies bacterianas. O modelo por eles obtido alcançou (AUC=0,79), performance bastante similar ao observado por Wei¹⁹⁸. O modelo de Li, considerou *Lactobacillus mucosae*, *Megasphaera* e *Streptococcus* enriquecido em LES enquanto *Faecalibacterium* estava reduzido em controles sadios. É possível que as diferenças observadas em nosso grupo sejam mais modestas comparadas a Li e Wei, devido à característica da população avaliada. Nossa população é composta por pacientes em remissão clínica da doença, com SLEDAI = 0 e composta cem por cento por mulheres que habitualmente apresentam gravidade de doença menor que homens.^{198,199}

A literatura que relaciona MI em DM1 em adultos é bastante limitada (5 artigos), até onde vai nosso conhecimento.¹²¹⁻¹²⁵ Apenas o estudo mais recente de Shilo et al.¹²¹, avaliou o papel preditivo da MI para DM1. Nesse estudo, adultos com DM1 apresentaram bactérias que geraram modelo preditivo de capacidade elevada (AUC = 0,89). Encontraram-se elevadas *Prevotella copri*, *Eubacterium siraeum*, e reduzida a bactéria *Faecalibacterium prausnitzii*. Em nosso trabalho, o modelo preditivo para DM1 obteve performance regular (AUC= 0,69), quando avaliada isoladamente. Elevado desempenho (AUC=0,85) foi obtido por nosso grupo quando se considerou os potenciais marcadores microbianos em conjunto com o consumo de nutrientes (folato, colesterol, zinco, manganês e proteína). Sabiamente, nutrientes impactam o controle glicêmico e, por vezes, o perfil inflamatório sistêmico do organismo. Ambos os fatores podem ter impacto com o manejo do DM1 e, conseqüentemente, justificam tamanha melhoria de desempenho no modelo preditivo ao inserirmos essas informações no conjunto de dados.

No campo de estudos da microbiota intestinal, uma doença de amplo interesse é o diabetes tipo 2. Inúmeros foram os trabalhos que tentaram identificar um perfil microbiano em portadores da enfermidade, sem sucesso em encontrar um único perfil convergente (**apêndice 11**). No presente trabalho, o modelo preditivo para identificação de diabetes tipo 2 apresentou o segundo

melhor desempenho ao utilizar a microbiota intestinal de forma isolada (AUC = 0,67 – regular). Por sua vez, o modelo que integrou características fenotípicas com os marcadores da microbiota teve desempenho excelente (AUC=0,98).

No interessante estudo de Gou et al (2021)²⁰⁰, as características do microbioma selecionado, a saber: as classes *Alphaproteobacteria* e *Deltaproteobacteria*, famílias *Lactobacillaceae*; *Mogibacteriaceae*; *Comamonadaceae*; *Mogibacteriaceae*; ordem *Lactobacillales*; gêneros *Clostridiaceae spp*, *Butyrivibrio*, *Roseburia*, *Megamonas*, *Dorea* e a espécie *Veillonella díspar* mostraram precisão superior na previsão do DM2 em comparação com as informações genéticas do hospedeiro e outras características fenotípicas (escore de risco cardiovascular+ estilo de vida + dieta). Gou e colaboradores²⁰⁰ identificaram que o modelo, somente com características fenotípicas, tinha capacidade preditiva de AUC=0,63. Ao somar dados de microbiota ao modelo, AUC subiu para 0,73, corroborando nosso achado de que unir informações do microbioma as características do estilo de vida é uma forma útil de potencializar a capacidade preditiva de modelos de ML.

Tabela 9. Compilado de pesquisas com aprendizado de máquina em estudos de microbioma intestinal na identificação de diversas doenças.

| Desfecho | Algoritmo | Desempenho | N= | Tipo de dados | Referência |
|--------------|----------------------------------|-------------|------|-------------------|------------|
| Risco de DM2 | SVM;mRMR | AUC = 0,81 | 345 | MG | 201 |
| | RF | AUC = 0,83 | 96 | MG | 192 |
| | Light GBM | AUC = 0,73 | 1832 | 16SrRNA | 200 |
| DII | RF | AUC > 0,86 | 155 | MG e metabolômica | 202 |
| DM1 | XGBoost | AUC = 0,89 | 74 | MG | 121 |
| LES | RF | AUC = 0,79 | 82 | 16S rRNA | 199 |
| | Regressão logística multivariada | AUC = 0.74. | 30 | 16S rRNA | 198 |
| AR | mRMR | AUC =0,94 | 130 | MG | 197 |

Adaptada de: Li, P. et al. (2022).¹⁹⁹

Legenda: DM2: diabetes tipo 2; DII: doenças inflamatórias intestinais; DM1: diabetes tipo 1; LES: lúpus eritematoso sistêmico; AR: artrite reumatoide; SVM: algoritmo de vetor de suporte; RF: Random forest (ou floresta randômica); Light GBM e XGBoost: tipos de Algoritmo de Reforço de Gradiente; MVIB: Gargalo de Informação Variacional Multimodal; mRMR: Relevância máxima e redundância mínima; metagenômica-shotgun, *MG*; *AUC*: área sob a curva.

Com base no exposto até o momento, é possível notar que a descoberta de potenciais biomarcadores microbianos para a predição de doenças e condições específicas é de grande interesse da comunidade científica. Para alguns casos, como o enriquecimento da abundância de *Fusobacterium nucleatum* em pacientes com câncer colorretal (CRC)^{66,191,203}, a literatura é convergente independentemente da região geográfica, hábitos de vida, idade e outros fatores variáveis. Para outras condições de saúde, embora a combinação de métodos baseados em ômicas e métodos computacionais venham melhorando o desempenho preditivo baseado em microbiota, a baixa reprodutibilidade dos potenciais biomarcadores microbianos pode levar à baixa confiabilidade.

As evidências crescentes que indicam que o microbioma desempenha um papel importante para o organismo humano são válidas. No entanto, o anseio por rápidas informações no campo da MI, sem a adequada padronização de coleta, manipulação de amostras, extração, sequenciamento, metodologia analítica e utilização de dados clínicos complementares dificultam conclusões científicas confiáveis e favorecem a veiculação de informações pouco reprodutíveis.^{182,185}

Deve-se refletir que existe uma grande diferença entre um achado “diferencialmente abundante” e a “identificação de um biomarcador” capaz de prever uma determinada condição, já que a identificação de táxons diferencialmente abundantes acaba por ser influenciada pelos inúmeros fatores previamente citados nesta tese.

O presente trabalho corrobora com publicações anteriores^{204,205}, que ressaltam a importância de não basear todas as conclusões de uma pesquisa exclusivamente em um valor estatístico. Em nosso caso, especialmente não

considerar uma bactéria como “marcador” de uma condição clínica somente porque esta apresentou abundância relativa diferente ($p < 0,05$) entre dois grupos.

Como sugestão para minimizar o viés em futuros estudos da área, nosso estudo destaca a importância da filtragem de variáveis, sejam elas fenotípicas ou microbianas, por reamostragem (bootstrap), a fim de minimizar o efeito da grande variabilidade observada entre estudos sobre MI e favorecer a reprodutibilidade dos resultados obtidos.^{182,197} Em concomitância, devem-se estabelecer diretrizes e procedimentos padronizados (preferencialmente em todo o mundo) para facilitar a comparabilidade entre os estudos com MI.¹⁸²

A pesquisa atual do microbioma está começando a avançar para além da mera descrição de estruturas da microbiota, reduzindo a importância que se dá a testes de correlação e p-valores, possivelmente espúrios, e progredindo na direção de compreender se esses achados da microbiota intestinal podem ser utilizados de forma reprodutível na prática.

Neste sentido, para além da padronização de metodologias, é essencial garantir a adequada interpretação dos dados gerados em estudos de microbiota e microbioma, e isso requer a participação de equipes multidisciplinares capacitadas no desenvolvimento de estudos com MI, onde a interpretação dos dados e sua associação às informações biológicas ultrapassem os “achados estatísticos”.

Táxons microbianos podem e devem ser considerados candidatos para predição de algumas condições, e para tanto devem ser testados quanto a sua reprodutibilidade e capacidade discriminatória por meio de modelos preditivos, como os expostos na presente tese.

8 CONCLUSÕES

8 CONCLUSÕES

Nas condições da presente pesquisa com adultos em estados de saúde e doença, em que se estudou a composição da microbiota intestinal e variáveis fenotípicas foi possível concluir que:

- Existem marcadores taxonômicos (gênero e espécies) pontuais que apresentam abundância relativa diferente entre os grupos avaliados, porém isso não os torna, isoladamente, marcadores com potencial preditivo.
- A avaliação da característica do ecossistema microbiano, - beta-diversidade, não permite identificar uma separação clara (padrão) na macroestrutura da microbiota capaz de discriminar entre estado de saúde e estados de doença, exceto para as doenças inflamatórias intestinais.
- Variáveis fenotípicas, como consumo de alguns nutrientes, características da composição corporal e índice de massa corpórea diferem entre indivíduos sadios e enfermos e devem integrar modelos preditivos para melhora da performance de identificação de alguns estados de doença, principalmente condições metabólicas como o DM2.
- Os modelos preditivos estabelecidos no presente trabalho, que integraram variáveis fenotípicas com táxons da microbiota intestinal para distinguir indivíduos sadios e enfermos, tiveram diferentes níveis de desempenho a depender da doença avaliada.
- Coletivamente, nossos dados sugerem que a utilização de marcadores microbianos isoladamente é pouco capaz de prever um desfecho de saúde. A adição de características fenotípicas com os resultados da composição da MI em um modelo preditivo integrado aumenta a capacidade de previsão de desfechos de saúde.

REFERÊNCIAS

REFERÊNCIAS

- 1 Dekaboruah E, Suryavanshi MV, Chettri D, Verma AK. Human microbiome: an academic update on human body site specific surveillance and its possible role. *Arch Microbiol.* 2020 Oct;202(8):2147-2167. doi: 10.1007/s00203-020-01931-x.
- 2 Marchesi JR, Ravel J. The vocabulary of microbiome research: a proposal. *Microbiome.* 2015 Jul 30;3:31. doi: 10.1186/s40168-015-0094-5.
- 3 Human Microbiome Project Consortium. Structure, function and diversity of the healthy human microbiome. *Nature.* 2012 Jun 13;486(7402):207-14. doi: 10.1038/nature11234.
- 4 Integrative HMP (iHMP) Research Network Consortium. The Integrative Human Microbiome Project. *Nature.* 2019 May;569(7758):641-648. doi: 10.1038/s41586-019-1238-8.
- 5 Almeida A, Nayfach S, Boland M, Strozzi F, Beracochea M, Shi ZJ, et al. A unified catalog of 204,938 reference genomes from the human gut microbiome. *Nat Biotechnol.* 2021;39:105–14. doi: 10.1038/s41587-020-0603-3.
- 6 Lloyd-Price J, Abu-Ali G, Huttenhower C. The healthy human microbiome. *Genome Med.* 2016 Apr 27;8(1):51. doi: 10.1186/s13073-016-0307-y.
- 7 Sender R, Fuchs S, Milo R. Revised Estimates for the Number of Human and Bacteria Cells in the Body. *PLoS Biol.* 2016 Aug 19;14(8):e1002533. doi: 10.1371/journal.pbio.1002533
- 8 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. doi: 10.4137/EBO.S36436.
- 9 Rowan-Nash AD, Korry BJ, Mylonakis E, Belenky P. Cross-Domain and Viral Interactions in the Microbiome. *Microbiol Mol Biol Rev.* 2019 Jan 9;83(1):e00044-18. doi: 10.1128/MMBR.00044-18.

10 Almeida A, Mitchell AL, Boland M, Forster SC, Gloor GB, Tarkowska A, et al. A new genomic blueprint of the human gut microbiota. *Nature*. 2019 Apr;568(7753):499-504. doi: 10.1038/s41586-019-0965-1.

11 Redinbo MR. The microbiota, chemical symbiosis, and human disease. *J Mol Biol*. 2014 Nov 25;426(23):3877-91. doi: 10.1016/j.jmb.2014.09.011.

12 Grace-Farfaglia P, Frazier H, Iversen MD. Essential Factors for a Healthy Microbiome: A Scoping Review. *Int J Environ Res Public Health*. 2022 Jul 8;19(14):8361. doi: 10.3390/ijerph19148361.

13 Gilbert JA, Blaser MJ, Caporaso JG, Jansson JK, Lynch SV, Knight R. Current understanding of the human microbiome. *Nat Med*. 2018 Apr 10;24(4):392-400. doi: 10.1038/nm.4517.

14 Davenport ER, Sanders JG, Song SJ, Amato KR, Clark AG, Knight R. The human microbiome in evolution. *BMC Biol*. 2017 Dec 27;15(1):127. doi: 10.1186/s12915-017-0454-7.

15 Heintz-Buschart A, Wilmes P. Human Gut Microbiome: Function Matters. *Trends Microbiol*. 2018 Jul;26(7):563-574. doi: 10.1016/j.tim.2017.11.002.

16 Thursby E, Juge N. Introduction to the human gut microbiota. *Biochem J*. 2017 May 16;474(11):1823-1836. doi: 10.1042/BCJ20160510.

17 Tang Q, Jin G, Wang G, Liu T, Liu X, Wang B, Cao H. Current Sampling Methods for Gut Microbiota: A Call for More Precise Devices. *Front Cell Infect Microbiol*. 2020 Apr 9;10:151. doi: 10.3389/fcimb.2020.00151

18 Margolis KG, Cryan JF, Mayer EA. The Microbiota-Gut-Brain Axis: From Motility to Mood. *Gastroenterology*. 2021 Apr;160(5):1486-1501. doi: 10.1053/j.gastro.2020.10.066.

19 de Sire R, Rizzatti G, Ingravalle F, Pizzoferrato M, Petito V, Lopetuso L, et al. Skeletal muscle-gut axis: emerging mechanisms of sarcopenia for intestinal and extra intestinal diseases. *Minerva Gastroenterol Dietol*. 2018 Dec;64(4):351-362. doi: 10.23736/S1121-421X.18.02511-4.

20 Laukens D, Brinkman BM, Raes J, De Vos M, Vandenabeele P. Heterogeneity of the gut microbiome in mice: guidelines for optimizing experimental design. *FEMS Microbiol Rev.* 2016 Jan;40(1):117-32. doi: 10.1093/femsre/fuv036.

21 Bielka W, Przekaz A, Pawlik A. The Role of the Gut Microbiota in the Pathogenesis of Diabetes. *Int J Mol Sci.* 2022 Jan 1;23(1):480. doi: 10.3390/ijms23010480.

22 Rothschild D, Weissbrod O, Barkan E, Kurilshikov A, Korem T, Zeevi D, et al. Environment dominates over host genetics in shaping human gut microbiota. *Nature.* 2018;555:210–5.

23 Wilmanski T, Diener C, Rappaport N, Patwardhan S, Wiedrick J, Lapidus J, et al. Gut microbiome pattern reflects healthy ageing and predicts survival in humans. *Nat Metab.* 2021;3:274–86.

24 Zmora N, Suez J, Elinav E. You are what you eat: diet, health and the gut microbiota. *Nat Rev Gastroenterol Hepatol.* 2019;16:35–56.

25 Wu H, Esteve E, Tremaroli V, Khan MT, Caesar R, Manneras-Holm L, et al. Metformin alters the gut microbiome of individuals with treatment-naive type 2 diabetes, contributing to the therapeutic effects of the drug. *Nat Med.* 2017;23:850–8.

26 Deschasaux M, Bouter KE, Prodan A, Levin E, Groen AK, Herrema H, et al. Depicting the composition of gut microbiota in a population with varied ethnic origins but shared geography. *Nat Med.* 2018;24:1526–31.

27 Sergeev IN, Aljutaily T, Walton G, Huarte E. Effects of Synbiotic Supplement on Human Gut Microbiota, Body Composition and Weight Loss in Obesity. *Nutrients.* 2020 Jan 15;12(1):222. doi: 10.3390/nu12010222.

28 Quiroga R, Nistal E, Estebanez B, Porrás D, Juárez-Fernández M, Martínez-Florez S, et al. Exercise training modulates the gut microbiota profile and impairs inflammatory signaling pathways in obese children. *Exp Mol Med.* 2020;52:1048–61.

- 29 Stanislowski MA, Dabelea D, Lange LA, Wagner BD, Lozupone CA. Gut microbiota phenotypes of obesity. *NPJ Biofilms Microbiomes*. 2019 Jul 1;5(1):18. doi: 10.1038/s41522-019-0091-8.
- 30 Asnicar F, Berry SE, Valdes AM, Nguyen LH, Piccinno G, Drew DA, et al. Microbiome connections with host metabolism and habitual diet from 1,098 deeply phenotyped individuals. *Nat Med*. 2021;27:321–32.
- 31 He Y, Wu W, Zheng HM, Li P, McDonald D, Sheng HF, et al. Regional variation limits applications of healthy gut microbiome reference ranges and disease models. *Nat Med*. 2018;24:1532–5.
- 32 Zhou W, Sailani MR, Contrepois K, et al. Longitudinal multi-omics of host-microbe dynamics in prediabetes. *Nature* 2019; 569(7758): 663–71.
- 33 Shin N-R, Whon TW, Bae J-W. Proteobacteria: microbial signature of dysbiosis in gut microbiota. *Trends Biotechnol* 2015; 33(9): 496–503.
- 34 Tian L, Wang XW, Wu AK, et al. Deciphering functional redundancy in the human microbiome. *Nat Commun* 2020; 11(1): 6217.
- 35 Visconti A, Le Roy CI, Rosa F, et al. Interplay between the human gut microbiome and host metabolism. *Nat Commun* 2019; 10(1): 4505.
- 36 Ching Jian. Gut microbiota in relation to diet-induced metabolic worsening and improvements in overweight and obese adults: toward prediction of personalized dietary response. [Dissertação] Helsinque, Universidade de Helsinque, 2022.
- 37 Nearing, J.T., Douglas, G.M., Hayes, M.G. et al. Microbiome differential abundance methods produce different results across 38 datasets. *Nat Commun* 13, 342 (2022). <https://doi.org/10.1038/s41467-022-28034-z>
- 38 Weersma RK, Zhernakova A, Fu J. Interaction between drugs and the gut microbiome. *Gut*. 2020 Aug;69(8):1510-1519. doi: 10.1136/gutjnl-2019-320204.
- 39 Cani PD. Gut microbiota - at the intersection of everything? *Nat Rev Gastroenterol Hepatol*. 2017 Jun;14(6):321-322. doi: 10.1038/nrgastro.2017.54.

40 Hooks KB, O'Malley MA. Dysbiosis and Its Discontents. *mBio*. 2017 Oct 10;8(5):e01492-17. doi: 10.1128/mBio.01492-17.

41 Cho I, Blaser MJ. The human microbiome: at the interface of health and disease. *Nat Rev Genet*. 2012;13(4):260–70.

42 Yang, L., Chen, J. A comprehensive evaluation of microbial differential abundance analysis methods: status and potential solutions. *Microbiome* 10, 130 (2022). <https://doi.org/10.1186/s40168-022-01320-0>

43 Vujkovic-Cvijin I, Sklar J, Jiang L, Natarajan L, Knight R, Belkaid Y. Host variables confound gut microbiota studies of human disease. *Nature*. 2020 Nov;587(7834):448-454. doi: 10.1038/s41586-020-2881-9.

44 Albenberg LG, Wu GD. Diet and the intestinal microbiome: associations, functions, and implications for health and disease. *Gastroenterology*. 2014 May;146(6):1564-72. doi: 10.1053/j.gastro.2014.01.058.

45 Zinöcker MK, Lindseth IA. The Western Diet-Microbiome-Host Interaction and Its Role in Metabolic Disease. *Nutrients*. 2018 Mar 17;10(3):365. doi: 10.3390/nu10030365.

46 Cai J, Chen Z, Wu W, Lin Q, Liang Y. High animal protein diet and gut microbiota in human health. *Crit Rev Food Sci Nutr*. 2021 Mar 16:1-13. doi: 10.1080/10408398.2021.1898336.

47 Precup G, Vodnar DC. Gut Prevotella as a possible biomarker of diet and its eubiotic versus dysbiotic roles: a comprehensive literature review. *Br J Nutr*. 2019 Jul 28;122(2):131-140. doi: 10.1017/S0007114519000680.

48 Cronin P, Joyce SA, O'Toole PW, O'Connor EM. Dietary Fibre Modulates the Gut Microbiota. *Nutrients*. 2021 May 13;13(5):1655. doi: 10.3390/nu13051655.

49 Cook MD, Allen JM, Pence BD, Wallig MA, Gaskins HR, White BA, et al. Exercise and gut immune function: evidence of alterations in colon immune cell homeostasis and microbiome characteristics with exercise training. *Immunol Cell Biol*. 2016 Feb;94(2):158-63. doi: 10.1038/icb.2015.108.

50 Gubert C, Kong G, Renoir T, Hannan AJ. Exercise, diet and stress as modulators of gut microbiota: Implications for neurodegenerative diseases. *Neurobiol Dis.* 2020 Feb;134:104621. doi: 10.1016/j.nbd.2019.104621.

51 Allen JM, Mailing LJ, Niemi GM, Moore R, Cook MD, White BA, et al. Exercise Alters Gut Microbiota Composition and Function in Lean and Obese Humans. *Med Sci Sports Exerc.* 2018 Apr;50(4):747-757. doi: 10.1249/MSS.0000000000001495.

52 Munukka E, Ahtiainen JP, Puigbó P, Jalkanen S, Pahkala K, Keskitalo A, et al. Six-Week Endurance Exercise Alters Gut Metagenome That Is not Reflected in Systemic Metabolism in Over-weight Women. *Front Microbiol.* 2018 Oct 3;9:2323. doi: 10.3389/fmicb.2018.02323.

53 Dalton A, Mermier C, Zuhl M. Exercise influence on the microbiome-gut-brain axis. *Gut Microbes.* 2019;10(5):555-568. doi: 10.1080/19490976.2018.1562268.

54 Mitchell CM, Davy BM, Hulver MW, Neilson AP, Bennett BJ, Davy KP. Does Exercise Alter Gut Microbial Composition? A Systematic Review. *Med Sci Sports Exerc.* 2019 Jan;51(1):160-167. doi: 10.1249/MSS.0000000000001760.

55 Warburton, D.E.R.; Bredin, S.S.D. Health benefits of physical activity: A systematic review of current systematic reviews. *Curr. Opin. Cardiol.* 2017, 32, 541–556.

56 Vogtmann E, Flores R, Yu G, Freedman ND, Shi J, Gail MH, et al. Association between tobacco use and the upper gastrointestinal microbiome among Chinese men. *Cancer Causes Control.* 2015 Apr;26(4):581-8. doi: 10.1007/s10552-015-0535-2.

57. Wang H, Zhao JX, Hu N, Ren J, Du M, Zhu MJ. Side-stream smoking reduces intestinal inflammation and increases expression of tight junction proteins. *World J Gastroenterol.* 2012 May 14;18(18):2180-7. doi: 10.3748/wjg.v18.i18.2180.

58. Biedermann L, Zeitz J, Mwinyi J, Sutter-Minder E, Rehman A, Ott SJ, et al. Smoking cessation induces profound changes in the composition of the intestinal

microbiota in humans. *PLoS One*. 2013;8(3):e59260. doi: 10.1371/journal.pone.0059260.

59. Biedermann L, Brülisauer K, Zeitz J, Frei P, Scharl M, Vavricka SR, et al. Smoking cessation alters intestinal microbiota: insights from quantitative investigations on human fecal samples using FISH. *Inflamm Bowel Dis*. 2014 Sep;20(9):1496-501. doi: 10.1097/MIB.000000000000129.

60. Allais L, Kerckhof FM, Verschuere S, Bracke KR, De Smet R, Laukens D, et al. Chronic cigarette smoke exposure induces microbial and inflammatory shifts and mucin changes in the murine gut. *Environ Microbiol*. 2016 May;18(5):1352-63. doi: 10.1111/1462-2920.12934.

61 Gui X, Yang Z, Li MD. Effect of Cigarette Smoke on Gut Microbiota: State of Knowledge. *Front Physiol*. 2021 Jun 17;12:673341. doi: 10.3389/fphys.2021.673341.

62 Vermeulen R, Schymanski EL, Barabási AL, Miller GW. The exposome and health: Where chemistry meets biology. *Science*. 2020 Jan 24;367(6476):392-396. doi: 10.1126/science.aay3164.

63 Falony G, Joossens M, Vieira-Silva S, Wang J, Darzi Y, Faust K, et al. Population-level analysis of gut microbiome variation. *Science*. 2016 Apr 29;352(6285):560-4. doi: 10.1126/science.aad3503.

64 Zhernakova A, Kurilshikov A, Bonder MJ, Tigchelaar EF, Schirmer M, Vatanen T, et al. Population-based metagenomics analysis reveals markers for gut microbiome composition and diversity. *Science*. 2016 Apr 29;352(6285):565-9. doi: 10.1126/science.aad3369.

65 Ananthakrishnan, A.N.; Bernstein, C.N.; Iliopoulos, D.; Macpherson, A.; Neurath, M.F.; Ali, R.A.R.; Vavricka, S.R.; Fiocchi, C. Environmental triggers in IBD: A review of progress and evidence. *Nat. Rev. Gastroenterol. Hepatol*. 2018, 15, 39–49.

66 Kostic AD, Xavier RJ, Gevers D. The microbiome in inflammatory bowel disease: current status and the future ahead. *Gastroenterology*. 2014 May;146(6):1489-99. doi: 10.1053/j.gastro.2014.02.009.

67 Glassner KL, Abraham BP, Quigley EMM. The microbiome and inflammatory bowel disease. *J Allergy Clin Immunol*. 2020 Jan;145(1):16-27. doi: 10.1016/j.jaci.2019.11.003.

68 Zhu S, Han M, Liu S, Fan L, Shi H, Li P. Composition and diverse differences of intestinal microbiota in ulcerative colitis patients. *Front Cell Infect Microbiol*. 2022 Aug 30;12:953962. doi: 10.3389/fcimb.2022.953962.

69 Sankarasubramanian J, Ahmad R, Avuthu N, Singh AB, Guda C. Gut Microbiota and Metabolic Specificity in Ulcerative Colitis and Crohn's Disease. *Front Med (Lausanne)*. 2020 Nov 27;7:606298. doi: 10.3389/fmed.2020.606298.

70 Braun T, Di Segni A, BenShoshan M, Neuman S, Levhar N, Bubis M, et al. Individualized Dynamics in the Gut Microbiota Precede Crohn's Disease Flares. *Am J Gastroenterol*. 2019 Jul;114(7):1142-1151. doi: 10.14309/ajg.0000000000000136.

71 Pittayanon R, Lau JT, Leontiadis GI, Tse F, Yuan Y, Surette M, Moayyedi P. Differences in Gut Microbiota in Patients With vs Without Inflammatory Bowel Diseases: A Systematic Review. *Gastroenterology*. 2020 Mar;158(4):930-946.e1. doi: 10.1053/j.gastro.2019.11.294.

72 Metwaly, A., Reitmeier, S. & Haller, D. Microbiome risk profiles as biomarkers for inflammatory and metabolic disorders. *Nat Rev Gastroenterol Hepatol* 19, 383–397 (2022). <https://doi.org/10.1038/s41575-022-00581-2>

73 Matsuoka K, Kanai T. The gut microbiota and inflammatory bowel disease. *Semin Immunopathol*. 2015 Jan;37(1):47-55. doi: 10.1007/s00281-014-0454-4.

74 Gevers D, Kugathasan S, Denson LA, Vázquez-Baeza Y, Van Treuren W, Ren B, et al. The treatment-naive microbiome in new-onset Crohn's disease. *Cell Host Microbe*. 2014 Mar 12;15(3):382-392. doi: 10.1016/j.chom.2014.02.005.

75 Magro DO, Santos A, Guadagnini D, de Godoy FM, Silva SHM, Lemos WJF, et al. Remission in Crohn's disease is accompanied by alterations in the gut microbiota and mucins production. *Sci Rep*. 2019 Sep 13;9(1):13263. doi: 10.1038/s41598-019-49893-5.

76 Glassner KL, Abraham BP, Quigley EMM. The microbiome and inflammatory bowel disease. *J Allergy Clin Immunol*. 2020 Jan;145(1):16-27. doi: 10.1016/j.jaci.2019.11.003.

77 Pascal V, Pozuelo M, Borruel N, Casellas F, Campos D, Santiago A, et al. A microbial signature for Crohn's disease. *Gut*. 2017 May;66(5):813-822. doi: 10.1136/gutjnl-2016-313235.

78 Nemoto H, Kataoka K, Ishikawa H, Ikata K, Arimochi H, Iwasaki T, et al. Reduced diversity and imbalance of fecal microbiota in patients with ulcerative colitis. *Dig Dis Sci*. 2012 Nov;57(11):2955-64. doi: 10.1007/s10620-012-2236-y.

79 Fuentes S, Rossen NG, van der Spek MJ, et al. Microbial shifts and signatures of long-term remission in ulcerative colitis after faecal microbiota transplantation. *ISME J*. 2017 Aug;11(8):1877-1889. doi: 10.1038/ismej.2017.44.

80 Machiels K, Joossens M, Sabino J, et al. A decrease of the butyrate-producing species *Roseburia hominis* and *Faecalibacterium prausnitzii* defines dysbiosis in patients with ulcerative colitis. *Gut*. 2014 Aug;63(8):1275-83. doi: 10.1136/gutjnl-2013-304833.

81 Sha S, Xu B, Wang X, et al. The biodiversity and composition of the dominant fecal microbiota in patients with inflammatory bowel disease. *Diagn Microbiol Infect Dis*. 2013 Mar;75(3):245-51. doi: 10.1016/j.diagmicrobio.2012.11.022.

82 Frank DN, St Amand AL, Feldman RA, Boedeker EC, Harpaz N, Pace NR. Molecular-phylogenetic characterization of microbial community imbalances in human inflammatory bowel diseases. *Proc Natl Acad Sci U S A*. 2007 Aug 21;104(34):13780-5. doi: 10.1073/pnas.0706625104.

83 Ott SJ, Plamondon S, Hart A, Begun A, Rehman A, Kamm MA, Schreiber S. Dynamics of the mucosa-associated flora in ulcerative colitis patients during

remission and clinical relapse. *J Clin Microbiol.* 2008 Oct;46(10):3510-3. doi: 10.1128/JCM.01512-08.

84 Rendon, A.; Schäkel, K. Psoriasis Pathogenesis and Treatment. *Int. J. Mol. Sci.* 2019, 20, 1475.

85 Hawkes, J.E.; Chan, T.C.; Krueger, J.G. Psoriasis Pathogenesis and the Development of Novel Targeted Immune Therapies. *J. Allergy Clin. Immunol.* 2017, 140, 645–653.

86 Kierasińska, M.; Donskow-Łysoniewska, K. Both the Microbiome and the Macrobiome Can Influence Immune Responsiveness in Psoriasis. *Central Eur. J. Immunol.* 2021, 46, 502–508.

87 Buhaş MC, Gavrilaş LI, Candrea R, Căţinean A, Mocan A, Miere D, Tătaru A. Gut Microbiota in Psoriasis. *Nutrients.* 2022; 14(14):2970. <https://doi.org/10.3390/nu14142970>

88 Hidalgo-Cantabrana C, Gómez J, Delgado S, Requena-López S, Queiro-Silva R, Margolles A, et al. Gut microbiota dysbiosis in a cohort of patients with psoriasis. *Br J Dermatol.* 2019 Dec;181(6):1287-1295. doi: 10.1111/bjd.17931.

89 Codoñer FM, Ramírez-Bosca A, Climent E, Carrión-Gutierrez M, Guerrero M, Pérez-Orquín JM, et al.. Gut microbial composition in patients with psoriasis. *Sci Rep.* 2018 Feb 28;8(1):3812. doi: 10.1038/s41598-018-22125-y

90 Widhiati S, Purnomosari D, Wibawa T, Soebono H. The role of gut microbiome in inflammatory skin disorders: A systematic review. *Dermatol Reports.* 2021 Dec 28;14(1):9188. doi: 10.4081/dr.2022.9188.

91 Huang L, Gao R, Yu N, Zhu Y, Ding Y, Qin H. Dysbiosis of gut microbiota was closely associated with psoriasis. *Sci China Life Sci.* 2019 Jun;62(6):807-815. doi: 10.1007/s11427-018-9376-6.

92 Shapiro J, Cohen NA, Shalev V, Uzan A, Koren O, Maharshak N. Psoriatic patients have a distinct structural and functional fecal microbiota compared with controls. *J Dermatol.* 2019 Jul;46(7):595-603. doi: 10.1111/1346-8138.14933.

93 Tan L, Zhao S, Zhu W, Wu L, Li J, Shen M, et al. The *Akkermansia muciniphila* is a gut microbiota signature in psoriasis. *Exp Dermatol*. 2018 Feb;27(2):144-149. doi: 10.1111/exd.13463.

94 Eppinga H, Sperna Weiland CJ, Thio HB, van der Woude CJ, Nijsten TE, et al. Similar Depletion of Protective *Faecalibacterium prausnitzii* in Psoriasis and Inflammatory Bowel Disease, but not in Hidradenitis Suppurativa. *J Crohns Colitis*. 2016 Sep;10(9):1067-75. doi: 10.1093/ecco-jcc/jjw070

95 Schade L, Mesa D, Faria AR, Santamaria JR, Xavier CA, Ribeiro D, et al. The gut microbiota profile in psoriasis: a Brazilian case-control study. *Lett Appl Microbiol*. 2022 Apr;74(4):498-504. doi: 10.1111/lam.13630. Epub 2021 Dec 20. PMID: 34897759.

96 Almutairi K, Nossent J, Preen D, Keen H, Inderjeeth C. The global prevalence of rheumatoid arthritis: a meta-analysis based on a systematic review. *Rheumatol Int*. 2021 May;41(5):863-877. doi: 10.1007/s00296-020-04731-0.

97 Gibofsky A. Epidemiology, pathophysiology, and diagnosis of rheumatoid arthritis: A Synopsis. *Am J Manag Care*. 2014 May;20(7 Suppl):S128-35. PMID: 25180621.

98 MacGregor AJ, Snieder H, Rigby AS, Koskenvuo M, Kaprio J, Aho K, Silman AJ. Characterizing the quantitative genetic contribution to rheumatoid arthritis using data from twins. *Arthritis Rheum*. 2000 Jan;43(1):30-7. doi: 10.1002/1529-0131(200001)43:1

99 Belkaid Y, Hand TW. Role of the microbiota in immunity and inflammation. *Cell*. 2014 Mar 27;157(1):121-41. doi: 10.1016/j.cell.2014.03.011.

100 Round JL, O'Connell RM, Mazmanian SK. Coordination of tolerogenic immune responses by the commensal microbiota. *J Autoimmun*. 2010 May;34(3):J220-5. doi: 10.1016/j.jaut.2009.11.007

101 O'Dell JR, Blakely KW, Mallek JA, Eckhoff PJ, Leff RD, Wees SJ, et al. Treatment of early seropositive rheumatoid arthritis: a two-year, double-blind

comparison of minocycline and hydroxychloroquine. *Arthritis Rheum.* 2001 Oct;44(10):2235-41. doi: 10.1002/1529-0131(200110)44:10

102 Pianta A, Arvikar S, Strle K, Drouin EE, Wang Q, Costello CE, Steere AC. Evidence of the Immune Relevance of *Prevotella copri*, a Gut Microbe, in Patients With Rheumatoid Arthritis. *Arthritis Rheumatol.* 2017 May;69(5):964-975. doi: 10.1002/art.40003.

103 Balakrishnan B, Luckey D, Bodhke R, Chen J, Marietta E, Jeraldo P, et al. *Prevotella histicola* Protects From Arthritis by Expansion of Allobaculum and Augmenting Butyrate Production in Humanized Mice. *Front Immunol.* 2021 May 4;12:609644. doi: 10.3389/fimmu.2021.609644.

104 Picchianti-Diamanti A, Panebianco C, Salemi S, Sorgi ML, Di Rosa R, Tropea A, Sgrulletti M, et al. Analysis of Gut Microbiota in Rheumatoid Arthritis Patients: Disease-Related Dysbiosis and Modifications Induced by Etanercept. *Int J Mol Sci.* 2018 Sep 27;19(10):2938. doi: 10.3390/ijms19102938.

105 Zhang X, Zhang D, Jia H, Feng Q, Wang D, Liang D, et al. The oral and gut microbiomes are perturbed in rheumatoid arthritis and partly normalized after treatment. *Nat Med.* 2015 Aug;21(8):895-905. doi: 10.1038/nm.3914.

106 Liu X, Zhang J, Zou Q, Zhong B, Wang H, Mou F, et al. *Lactobacillus salivarius* Isolated from Patients with Rheumatoid Arthritis Suppresses Collagen-Induced Arthritis and Increases Treg Frequency in Mice. *J Interferon Cytokine Res.* 2016 Dec;36(12):706-712. doi: 10.1089/jir.2016.0057.

107 Reyes-Castillo Z, Valdés-Miramontes E, Llamas-Covarrubias M, Muñoz-Valle JF. Troublesome friends within us: the role of gut microbiota on rheumatoid arthritis etiopathogenesis and its clinical and therapeutic relevance. *Clin Exp Med.* 2021 Feb;21(1):1-13. doi: 10.1007/s10238-020-00647-y.

108 Tsokos GC. Systemic lupus erythematosus. *N Engl J Med.* 2011 Dec 1;365(22):2110-21. doi: 10.1056/NEJMra1100359.

109 Ugarte-Gil MF, Mendoza-Pinto C, Reátegui-Sokolova C, Pons-Estel GJ, van Vollenhoven RF, Bertsias G, et al. Achieving remission or low disease activity is

associated with better outcomes in patients with systemic lupus erythematosus: a systematic literature review. *Lupus Sci Med.* 2021 Sep;8(1):e000542. doi: 10.1136/lupus-2021-000542.

110 He Z, Shao T, Li H, Xie Z, Wen C. Alterations of the gut microbiome in Chinese patients with systemic lupus erythematosus. *Gut Pathog.* 2016 Dec 8;8:64. doi: 10.1186/s13099-016-0146-9.

111 Azzouz D, Omarbekova A, Heguy A, Schwudke D, Gisch N, Rovin BH, et al. Lupus nephritis is linked to disease-activity associated expansions and immunity to a gut commensal. *Ann Rheum Dis.* 2019 Jul;78(7):947-956. doi: 10.1136/annrheumdis-2018-214856.

112 Toumi E, Goutorbe B, Plauzolles A, Bonnet M, Mezouar S, Militello M, et al. Gut microbiota in systemic lupus erythematosus patients and lupus mouse model: a cross species comparative analysis for biomarker discovery. *Front Immunol.* 2022 Aug 2;13:943241. doi: 10.3389/fimmu.2022.943241.

113 Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* 2014;15(12):550. doi: 10.1186/s13059-014-0550-8.

114 Marre ML, James EA, Piganelli JD. beta cell ER stress and the implications for immunogenicity in type 1 diabetes. *Front Cell Dev Biol.* (2015) 3:67. doi: 10.3389/fcell.2015.00067

115 Atkinson M, Gale EA. Infant diets and type 1 diabetes: too early, too late, or just too complicated? *JAMA.* (2003) 290:1771–2. doi: 10.1001/jama.290.13.1771

116 Ceccarelli F, Agmon-Levin N, Perricone C. Genetic factors of autoimmune diseases 2017. *J Immunol Res.* (2017) 2017:2789242. doi: 10.1155/2017/2789242

117 Vojdani A, Pollard KM, Campbell AW. Environmental triggers and autoimmunity. *Autoimmune Dis.* (2014) 2014:798029. doi: 10.1155/2014/798029

118 Dedrick S, Sundaresh B, Huang Q, Brady C, Yoo T, Cronin C, et al. The Role of Gut Microbiota and Environmental Factors in Type 1 Diabetes Pathogenesis. *Front Endocrinol (Lausanne)*. 2020 Feb 26;11:78. doi: 10.3389/fendo.2020.00078.

119 Zhou H, Zhao X, Sun L, Liu Y, Lv Y, Gang X, et al. Gut Microbiota Profile in Patients with Type 1 Diabetes Based on 16S rRNA Gene Sequencing: A Systematic Review. *Dis Markers*. 2020 Aug 27;2020:3936247. doi: 10.1155/2020/3936247.

120 Jamshidi P, Hasanzadeh S, Tahvildari A, Farsi Y, Arbabi M, Mota JF, Sechi LA, Nasiri MJ. Is there any association between gut microbiota and type 1 diabetes? A systematic review. *Gut Pathog*. 2019 Oct 14;11:49. doi: 10.1186/s13099-019-0332-7.

121 Shilo S, Godneva A, Rachmiel M, Korem T, Bussi Y, Kolobkov D, et al. The Gut Microbiome of Adults with Type 1 Diabetes and Its Association with the Host Glycemic Control. *Diabetes Care*. 2022 Mar 1;45(3):555-563. doi: 10.2337/dc21-1656. PMID: 35045174.

122 Higuchi BS, Rodrigues N, Gonzaga MI, Paiolo JCC, Stefanutto N, Omori WP, et al. Intestinal Dysbiosis in Autoimmune Diabetes Is Correlated With Poor Glycemic Control and Increased Interleukin-6: A Pilot Study. *Front Immunol*. 2018 Jul 25;9:1689. doi: 10.3389/fimmu.2018.01689.

123 Huang Y, Li SC, Hu J, Ruan HB, Guo HM, Zhang HH, Wang X, Pei YF, Pan Y, Fang C. Gut microbiota profiling in Han Chinese with type 1 diabetes. *Diabetes Res Clin Pract*. 2018 Jul;141:256-263. doi: 10.1016/j.diabres.2018.04.032.

124 de Groot PF, Belzer C, Aydin Ö, Levin E, Levels JH, Aalvink S, et al. Distinct fecal and oral microbiota composition in human type 1 diabetes, an observational study. *PLoS One*. 2017 Dec 6;12(12):e0188475. doi: 10.1371/journal.pone.0188475.

125 Stewart CJ, Nelson A, Campbell MD, Walker M, Stevenson EJ, Shaw JA, Cummings SP, West DJ. Gut microbiota of Type 1 diabetes patients with good

glycaemic control and high physical fitness is similar to people without diabetes: an observational study. *Diabet Med.* 2017 Jan;34(1):127-134. doi: 10.1111/dme.13140.

126 Gurung M, Li Z, You H, Rodrigues R, Jump DB, Morgun A, Shulzhenko N. Role of gut microbiota in type 2 diabetes pathophysiology. *EBioMedicine.* 2020 Jan;51:102590. doi: 10.1016/j.ebiom.2019.11.051.

127 Thingholm LB, Rühlemann MC, Koch M, Fuqua B, Laucke G, Boehm R, et al. Obese Individuals with and without Type 2 Diabetes Show Different Gut Microbial Functional Capacity and Composition. *Cell Host Microbe.* 2019 Aug 14;26(2):252-264.e10. doi: 10.1016/j.chom.2019.07.004.

128 Yamaguchi Y, Adachi K, Sugiyama T, Shimosato A, Ebi M, Ogasawara N, Funaki Y, Goto C, Sasaki M, Kasugai K. Association of Intestinal Microbiota with Metabolic Markers and Dietary Habits in Patients with Type 2 Diabetes. *Digestion.* 2016;94(2):66-72. doi: 10.1159/000447690.

129 Zhang Q, Hu N. Effects of Metformin on the Gut Microbiota in Obesity and Type 2 Diabetes Mellitus. *Diabetes Metab Syndr Obes.* 2020 Dec 16;13:5003-5014. doi: 10.2147/DMSO.S286430.

130 Reitmeier S, Kiessling S, Clavel T, List M, Almeida EL, Ghosh TS, et al. Arrhythmic Gut Microbiome Signatures Predict Risk of Type 2 Diabetes. *Cell Host Microbe.* 2020 Aug 12;28(2):258-272.e6. doi: 10.1016/j.chom.2020.06.004. Epub 2020 Jul 2. PMID: 32619440.

131 Tinajero MG, Malik VS. An Update on the Epidemiology of Type 2 Diabetes: A Global Perspective. *Endocrinol Metab Clin North Am.* 2021 Sep;50(3):337-355. doi: 10.1016/j.ecl.2021.05.013. PMID: 34399949.

132 International Diabetes Federation. *IDF Diabetes Atlas, 10th edn.* Brussels, Belgium: International Diabetes Federation, 2021.

133 Duvallet C, Gibbons SM, Gurry T, Irizarry RA, Alm EJ. Meta-analysis of gut microbiome studies identifies disease-specific and shared responses. *Nat Commun.* 2017 Dec 5;8(1):1784. doi: 10.1038/s41467-017-01973-8.

134 Gupta VK, Kim M, Bakshi U, Cunningham KY, Davis JM 3rd, Lazaridis KN, Nelson H, Chia N, Sung J. A predictive index for health status using species-level gut microbiome profiling. *Nat Commun.* 2020 Sep 15;11(1):4635. doi: 10.1038/s41467-020-18476-8

135 Allaband, C. et al. Microbiome 101: studying, analyzing, and interpreting gut microbiome data for clinicians. *Clin. Gastroenterol. Hepatol.* 17, 218 (2019).

136 Staley, C., Kaiser, T. & Khoruts, A. Clinician guide to microbiome testing. *Dig. Dis. Sci.* 63, 3167–3177 (2018).

137 McBurney, M. I. et al. Establishing what constitutes a healthy human gut microbiome: state of the science, regulatory considerations, and future directions. *J. Nutr.* 149, 1882–1895 (2019).

138 Hagerty, S. L., Hutchison, K. E., Lowry, C. A. & Bryan, A. D. An empirically derived method for measuring human gut microbiome alpha diversity: demonstrated utility in predicting health-related outcomes among a human clinical sample. *PLoS ONE* 15, e0229204 (2020).

139 <https://www.gutmicrobiotaforhealth.com/can-mapping-and-manipulating-the-gut-microbiome-based-on-a-test-be-relevant-for-clinical-practice/?search=test>

140 Lewis, J. D. The utility of biomarkers in the diagnosis and therapy of inflammatory bowel disease. *Gastroenterology* 140, 1817–1826.e2 (2011).

141 Sipponen, T. & Kolho, K.-L. Fecal calprotectin in diagnosis and clinical assessment of inflammatory bowel disease. *Scand. J. Gastroenterol.* 50, 74–80 (2015).

142 Vich Vila, A. et al. Gut microbiota composition and functional changes in inflammatory bowel disease and irritable bowel syndrome. *Sci. Transl Med.* 10, eaap8914 (2018).

143 Tian, H., Chen, Q., Yang, B. et al. Analysis of Gut Microbiome and Metabolite Characteristics in Patients with Slow Transit Constipation. *Dig Dis Sci* 66, 3026–3035 (2021). <https://doi.org/10.1007/s10620-020-06500-2>

144 Yu J, Feng Q, Wong SH, Zhang D, Liang QY, et al. Metagenomic analysis of faecal microbiome as a tool towards targeted non-invasive biomarkers for colorectal cancer. *Gut*. 2017 Jan;66(1):70-78. doi: 10.1136/gutjnl-2015-309800.

145 Zackular JP, Rogers MA, Ruffin MT 4th, Schloss PD. The human gut microbiome as a screening tool for colorectal cancer. *Cancer Prev Res (Phila)*. 2014 Nov;7(11):1112-21. doi: 10.1158/1940-6207.CAPR-14-0129.

146 Caussy, C., Tripathi, A., Humphrey, G. et al. A gut microbiome signature for cirrhosis due to nonalcoholic fatty liver disease. *Nat Commun* 10, 1406 (2019). <https://doi.org/10.1038/s41467-019-09455-9>

147 Wu H, Tremaroli V, Schmidt C, Lundqvist A, Olsson LM, Krämer M, Gummesson A, Perkins R, Bergström G, Bäckhed F. The Gut Microbiota in Prediabetes and Diabetes: A Population-Based Cross-Sectional Study. *Cell Metab*. 2020 Sep 1;32(3):379-390.e3. doi: 10.1016/j.cmet.2020.06.011.

148 Sze MA, Schloss PD. Looking for a Signal in the Noise: Revisiting Obesity and the Microbiome. *mBio*. 2016 Aug 23;7(4):e01018-16. doi: 10.1128/mBio.01018-16. Erratum in: *MBio*. 2017 Dec 5;8(6):

149 Aronson JK, Ferner RE. Biomarkers-A General Review. *Curr Protoc Pharmacol*. 2017 Mar 17;76:9.23.1-9.23.17. doi: 10.1002/cpph.19. PMID: 28306150.

150 Olesen SW, Alm EJ. Dysbiosis is not an answer. *Nat Microbiol*. 2016 Nov 25;1:16228. doi: 10.1038/nmicrobiol.2016.228. PMID: 27886190.

151 Koohi-Moghadam M, Borad MJ, Tran NL, Swanson KR, Boardman LA, Sun H, Wang J. MetaMarker: a pipeline for de novo discovery of novel metagenomic biomarkers. *Bioinformatics*. 2019 Oct 1;35(19):3812-3814. doi: 10.1093/bioinformatics/btz123.

152 Rodrigues GSP, Cayres LCF, Gonçalves FP, Takaoka NNC, Lengert AH, Tansini A, et al. Detection of Increased Relative Expression Units of Bacteroides and Prevotella, and Decreased Clostridium leptum in Stool Samples from Brazilian Rheumatoid Arthritis Patients: A Pilot Study. *Microorganisms*. 2019; 7(10):413. <https://doi.org/10.3390/microorganisms7100413>

153 Hevia A, Milani C, López P, Cuervo A, Arboleya S, Duranti S, et al. Intestinal dysbiosis associated with systemic lupus erythematosus. *mBio*. 2014 Sep 30;5(5):e01548-14. doi: 10.1128/mBio.01548-14.

154 Stewart CJ, Nelson A, Campbell MD, Walker M, Stevenson EJ, Shaw JA, et al. Gut microbiota of Type 1 diabetes patients with good glycaemic control and high physical fitness is similar to people without diabetes: an observational study. *Diabet Med*. 2017 Jan;34(1):127-134. doi: 10.1111/dme.13140.

155 Karvetti RL, Knuts LR. Validity of the 24-hour dietary recall. *J Am Diet Assoc*. 1985 Nov;85(11):1437-42. PMID: 4056262.

156 Fisberg RM, Marchioni DM, Colucci AC. Avaliação do consumo alimentar e da ingestão de nutrientes na prática clínica [Assessment of food consumption and nutrient intake in clinical practice]. *Arq Bras Endocrinol Metabol*. 2009 Jul;53(5):617-24. Portuguese. doi: 10.1590/s0004-27302009000500014.

157 UNIVERSIDADE ESTADUAL DE CAMPINAS - Tabela Brasileira de Composição de Alimentos (TACO). NEPA – UNICAMP, Campinas, ed. 4, rev. e ampl. - Campinas: NEPAUNICAMP, 2011. 161p

158 Philippi, S.T. Tabela de Composição de Alimentos: Suporte para decisão nutricional. Manole, Barueri, SP, ed, 5, 2016.

159 INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA – IBGE. Pesquisa de Orçamento Familiares. IBGE, Rio de Janeiro, ed. 1, 2011.

160 USDA. U. S. Department of Agriculture. USDA Table of Nutrient Retention Factors, Release 6. Natl Acad Press 2007:18.

161 Pinheiro, A. B. V. et al. Tabela para avaliação de consumo alimentar em medidas caseiras. 5. ed. São Paulo: Atheneu, 2004. 131p. Paulo: Atheneu, 2004. 131 p.

162 Castro, M. A. De; Marchioni, D. M. L.; Fisberg, L. M. Manual de Críticas de Inquéritos alimentares. Edição atualizada por Lopes, L. do V. C.; Fontanelli, M de M. Viçosa, MG, 2013.

163 Illumina. (2013) 16S metagenomic sequencing library preparation. https://support.illumina.com/documents/documentation/chemistry_documentation/16s/16s-metagenomic-library-prep-guide-15044223-b.pdf

164 Callahan BJ, McMurdie PJ, Rosen MJ, Han AW, Johnson AJ, Holmes SP. DADA2: High-resolution sample inference from Illumina amplicon data. *Nat Methods*. 2016 Jul;13(7):581-3. doi: 10.1038/nmeth.3869.

165 Pires DEV, Oliveira FS, Correa FB, Morais DK, Fernandes GR. TAG.ME: Taxonomic Assignment of Genetic Markers for Ecology. *bioRxiv* 263293; doi: <https://doi.org/10.1101/263293>

166 Quast C, Pruesse E, Yilmaz P, Gerken J, Schweer T, Yarza P, et al. The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic Acids Res*. 2013 Jan;41(Database issue):D590-6. doi: 10.1093/nar/gks1219. Disponível em: <https://www.arb-silva.de/>

167 Harttig U, Haubrock J, Knüppel S, Boeing H. The MSM program: web-based statistics package for estimating usual dietary intake using the Multiple Source Method. *Eur J Clin Nutr*. julho de 2011;65 Suppl 1:S87–91.

168 Multiple Source Method (MSM) está implementado na web no site: <https://msm.dife.de/> acesso em: junho de 2021.

169 Kuhn, M. Building Predictive Models in R Using the caret Package. *Journal of Statistical Software*, 2008; 28(5), 1–26. <https://doi.org/10.18637/jss.v028.i05>

170 Flemer B, Warren RD, Barrett MP, Cisek K, Das A, Jeffery IB, Hurley E, O'Riordain M, Shanahan F, O'Toole PW. The oral microbiota in colorectal cancer

is distinctive and predictive. *Gut*. 2018 Aug;67(8):1454-1463. doi: 10.1136/gutjnl-2017-314814.

171 Yachida S, Mizutani S, Shiroma H, Shiba S, Nakajima T, Sakamoto T, et al. Metagenomic and metabolomic analyses reveal distinct stage-specific phenotypes of the gut microbiota in colorectal cancer. *Nat Med*. 2019 Jun;25(6):968-976. doi: 10.1038/s41591-019-0458-7.

172 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. doi: 10.1038/s41591-019-0405-7. Epub 2019 Apr 1. Erratum in: *Nat Med*. 2019 Dec;25(12):1948.

173 Wirbel, J., Pyl, P. T., Kartal, E., Zych, K., Kashani, A., Milanese, A., et al. (2019). Meta-analysis of fecal metagenomes reveals global microbial signatures that are specific for colorectal cancer. *Nat. Med.* 25, 679–689. doi: 10.1038/s41591-019-0406-6

174 Hong Han, Xiaoling Guo and Hua Yu, "Variable selection using Mean Decrease Accuracy and Mean Decrease Gini based on Random Forest," 2016 7th IEEE International Conference on Software Engineering and Service Science (ICSESS), 2016, pp. 219-224, doi: 10.1109/ICSESS.2016.7883053.

175 Pepe MS. An interpretation for the ROC curve and inference using GLM procedures. *Biometrics*. 2000 Jun;56(2):352-9. doi: 10.1111/j.0006-341x.2000.00352.x. PMID: 10877289.

176 Aryal S, Alimadadi A, Manandhar I, Joe B, Cheng X. Machine Learning Strategy for Gut Microbiome-Based Diagnostic Screening of Cardiovascular Disease. *Hypertension*. 2020 Nov;76(5):1555-1562. doi: 10.1161/HYPERTENSIONAHA.120.15885.

177 Hosmer DW, Lemeshow S. *Applied logistic regression*. 2000. 2ª edição. Ed Wiley, pp 156–164.

178 Buffet-Bataillon S, Bouguen G, Fleury F, Cattoir V, Le Cunff Y. Gut microbiota analysis for prediction of clinical relapse in Crohn's disease. *Sci Rep.* 2022 Nov 19;12(1):19929. doi: 10.1038/s41598-022-23757-x. PMID: 36402792; PMCID: PMC9675750.

179 Mossotto E, Ashton JJ, Coelho T, Beattie RM, MacArthur BD, Ennis S. Classification of Paediatric Inflammatory Bowel Disease using Machine Learning. *Sci Rep.* 2017 May 25;7(1):2427. doi: 10.1038/s41598-017-02606-2. PMID: 28546534; PMCID: PMC5445076.

180 Ghannam RB, Techtman SM. Machine learning applications in microbial ecology, human microbiome studies, and environmental monitoring. *Comput Struct Biotechnol J.* 2021 Jan 27;19:1092-1107. doi: 10.1016/j.csbj.2021.01.028. PMID: 33680353

181 Kohli A, Holzwanger EA, Levy AN. Emerging use of artificial intelligence in inflammatory bowel disease. *World J Gastroenterol.* 2020 Nov 28;26(44):6923-6928. doi: 10.3748/wjg.v26.i44.6923.

182 Poussin C, Sierro N, Boué S, Battey J, Scotti E, Belcastro V, et al. Interrogating the microbiome: experimental and computational considerations in support of study reproducibility. *Drug Discov Today.* 2018 Sep;23(9):1644-1657. doi: 10.1016/j.drudis.2018.06.005.

183 Clooney AG, Eckenberger J, Laserna-Mendieta E, Sexton KA, Bernstein MT, Vagianos K, et al. Ranking microbiome variance in inflammatory bowel disease: a large longitudinal intercontinental study. *Gut.* 2021 Mar;70(3):499-510. doi: 10.1136/gutjnl-2020-321106.

184 Scher, J.U.; Ubeda, C.; Artacho, A.; Attur, M.; Isaac, S.; Reddy, S.M.; et al. Decreased bacterial diversity characterizes the altered gut microbiota in patients with psoriatic arthritis, resembling dysbiosis in inflammatory bowel disease. *Arthritis Rheumatol.* 2015, 67, 128–13

185 Wilkinson JE, Franzosa EA, Everett C, Li C; HCMPH researchers and trainees; HCMPH investigators; Hu FB, Wirth DF, Song M, Chan AT, Rimm E,

Garrett WS, Huttenhower C. A framework for microbiome science in public health. *Nat Med*. 2021 May;27(5):766-774. doi: 10.1038/s41591-021-01258-0.

186 Sanna S, van Zuydam NR, Mahajan A, Kurilshikov A, Vich Vila A, Vösa U, et al. Causal relationships among the gut microbiome, short-chain fatty acids and metabolic diseases. *Nat Genet*. 2019 Apr;51(4):600-605. doi: 10.1038/s41588-019-0350-x.

187 Schloss PD. Identifying and Overcoming Threats to Reproducibility, Replicability, Robustness, and Generalizability in Microbiome Research. *mBio*. 2018 Jun 5;9(3):e00525-18. doi: 10.1128/mBio.00525-18.

188 Kurilshikov A, Medina-Gomez C, Bacigalupe R, Radjabzadeh D, Wang J, Demirkan A, et al. Large-scale association analyses identify host factors influencing human gut microbiome composition. *Nat Genet*. 2021 Feb;53(2):156-165. doi: 10.1038/s41588-020-00763-1.

189 Goodswen SJ, Barratt JLN, Kennedy PJ, Kaufer A, Calarco L, Ellis JT. Machine learning and applications in microbiology. *FEMS Microbiol Rev*. 2021 Sep 8;45(5):fuab015. doi: 10.1093/femsre/fuab015.

190 Li P, Luo H, Ji B, Nielsen J. Machine learning for data integration in human gut microbiome. *Microb Cell Fact*. 2022 Nov 23;21(1):241. doi: 10.1186/s12934-022-01973-4.

191 Salim F, Mizutani S, Zolfo M, Yamada T. Recent advances of machine learning applications in human gut microbiota study: from observational analysis toward causal inference and clinical intervention. *Curr Opin Biotechnol*. 2023 Feb;79:102884. doi: 10.1016/j.copbio.2022.102884.

192 Karlsson FH, Tremaroli V, Nookaew I, Bergstrom G, Behre CJ, Fagerberg B, et al. Gut metagenome in european women with normal, impaired and diabetic glucose control. *Nature*. 2013 Jun 6;498(7452):99-103. doi: 10.1038/nature12198. Epub 2013 May 29. PMID: 23719380.

193 Zhou Y, Xu ZZ, He Y, Yang Y, Liu L, Lin Q, et al. Gut microbiota offers universal biomarkers across ethnicity in inflammatory bowel disease diagnosis

and infliximab response prediction. *mSystems*. 2018 Jan 30;3(1):e00188-17. doi: 10.1128/mSystems.00188-17.

194 Douglas GM, Hansen R, Jones CMA, Dunn KA, Comeau AM, Bielawski JP, et al. Multi-omics differentially classify disease state and treatment outcome in pediatric Crohn's disease. *Microbiome*. 2018 Jan 15;6(1):13. doi: 10.1186/s40168-018-0398-3.

195 Marcos-Zambrano LJ, Karaduzovic-Hadziabdic K, Loncar Turukalo T, Przymus P, Trajkovic V, Aasmets O, et al. Applications of Machine Learning in Human Microbiome Studies: A Review on Feature Selection, Biomarker Identification, Disease Prediction and Treatment. *Front Microbiol*. 2021 Feb 19;12:634511. doi: 10.3389/fmicb.2021.634511. PMID: 33737920; PMCID: PMC7962872.

196 Ananthakrishnan AN, Luo C, Yajnik V, Khalili H, Garber JJ, Stevens BW, et al. Gut Microbiome Function Predicts Response to Anti-integrin Biologic Therapy in Inflammatory Bowel Diseases. *Cell Host Microbe*. 2017 May 10;21(5):603-610.e3. doi: 10.1016/j.chom.2017.04.010.

197 Wu H, Cai L, Li D, Wang X, Zhao S, Zou F, Zhou K. Metagenomics Biomarkers Selected for Prediction of Three Different Diseases in Chinese Population. *Biomed Res Int*. 2018 Jan 11;2018:2936257. doi: 10.1155/2018/2936257.

198 Wei F, Xu H, Yan C, Rong C, Liu B, Zhou H. Changes of intestinal flora in patients with systemic lupus erythematosus in northeast China. *PLoS One*. 2019 Mar 14;14(3):e0213063. doi: 10.1371/journal.pone.0213063.

199 Li Y, Wang HF, Li X, Li HX, Zhang Q, Zhou HW, He Y, Li P, Fu C, Zhang XH, Qiu YR, Li JL. Disordered intestinal microbes are associated with the activity of Systemic Lupus Erythematosus. *Clin Sci (Lond)*. 2019 Apr 15;133(7):821-838. doi: 10.1042/CS20180841.

200 Gou W, Ling CW, He Y, Jiang Z, Fu Y, Xu F, Miao Z, Sun TY, Lin JS, Zhu HL, Zhou H, Chen YM, Zheng JS. Interpretable Machine Learning Framework

Reveals Robust Gut Microbiome Features Associated With Type 2 Diabetes. *Diabetes Care*. 2021 Feb;44(2):358-366. doi: 10.2337/dc20-1536.

201 Qin J, Li Y, Cai Z, Li S, Zhu J, Zhang F, et al. A metagenome-wide association study of gut microbiota in type 2 diabetes. *Nature*. 2012 Oct 4;490(7418):55-60. doi: 10.1038/nature11450.

202 Franzosa EA, Sirota-Madi A, Avila-Pacheco J, Fornelos N, Haiser HJ, Reinker S, Vatanen T, et al. Gut microbiome structure and metabolic activity in inflammatory bowel disease. *Nat Microbiol*. 2019 Feb;4(2):293-305. doi: 10.1038/s41564-018-0306-4. Epub 2018 Dec 10. Erratum in: *Nat Microbiol*. 2019 May;4(5):898.

203 Castellarin M, Warren RL, Freeman JD, Dreolini L, Krzywinski M, et al. *Fusobacterium nucleatum* infection is prevalent in human colorectal carcinoma. *Genome Res*. 2012 Feb;22(2):299-306. doi: 10.1101/gr.126516.111.

204 Ronald L. Wasserstein, Allen L. Schirm & Nicole A. Lazar (2019) Moving to a World Beyond “ $p < 0.05$ ”, *The American Statistician*, 73:sup1, 1-19, DOI: 10.1080/00031305.2019.1583913

205 Amrhein V, Greenland S, McShane B. Scientists rise up against statistical significance. *Nature*. 2019 Mar;567(7748):305-307. doi: 10.1038/d41586-019-00857-9.

206 Chávez-Carbajal A, Pizano-Zárate ML, Hernández-Quiroz F, Ortiz-Luna GF, Morales-Hernández RM, De Sales-Millán A, et al. Characterization of the Gut Microbiota of Individuals at Different T2D Stages Reveals a Complex Relationship with the Host. *Microorganisms*. 2020 Jan 10;8(1):94. doi: 10.3390/microorganisms8010094.

207 Zeevi D, Korem T, Zmora N, Israeli D, Rothschild D, Weinberger A, et al. Personalized Nutrition by Prediction of Glycemic Responses. *Cell*. 2015 Nov 19;163(5):1079-1094. doi: 10.1016/j.cell.2015.11.001.

208 Ahmad A, Yang W, Chen G, Shafiq M, Javed S, Ali Zaidi SS, Shahid R, Liu C, Bokhari H. Analysis of gut microbiota of obese individuals with type 2 diabetes and healthy individuals. *PLoS One*. 2019 Dec 31;14(12):e0226372. doi: 10.1371/journal.pone.0226372.

209 He Y, Wu W, Wu S, Zheng HM, Li P, Sheng HF, et al. Linking gut microbiota, metabolic syndrome and economic status based on a population-level analysis. *Microbiome*. 2018 Sep 24;6(1):172. doi: 10.1186/s40168-018-0557-6.

210 Li Q, Chang Y, Zhang K, Chen H, Tao S, Zhang Z. Implication of the gut microbiome composition of type 2 diabetic patients from northern China. *Sci Rep*. 2020 Mar 25;10(1):5450. doi: 10.1038/s41598-020-62224-3.

211 Doumatey AP, Adeyemo A, Zhou J, Lei L, Adebamowo SN, Adebamowo C, Rotimi CN. Gut Microbiome Profiles Are Associated With Type 2 Diabetes in Urban Africans. *Front Cell Infect Microbiol*. 2020 Feb 25;10:63. doi: 10.3389/fcimb.2020.00063.

212 Peters BA, Shapiro JA, Church TR, Miller G, Trinh-Shevrin C, Yuen E, et al. A taxonomic signature of obesity in a large study of American adults. *Sci Rep*. 2018 Jun 27;8(1):9749. doi: 10.1038/s41598-018-28126-1.

APÊNDICES

APÊNDICE 1. NÚMERO DE LEITURAS TOTAIS E FILTRADAS DE CADA AMOSTRA

| ID | Leituras totais | Leituras filtradas |
|-----------|------------------------|---------------------------|
| AMB_1 | 111.203 | 99.732 |
| AMB_2 | 103.816 | 94.075 |
| AMB_3 | 119.211 | 112.314 |
| AMB_4 | 121.671 | 118.385 |
| AMB_5 | 109.454 | 98.630 |
| AMB_6 | 100.036 | 95.226 |
| AMB_7 | 96.816 | 92.998 |
| AMB_8 | 89.823 | 86.293 |
| AMB_9 | 94.928 | 89.934 |
| AMB_10 | 119.952 | 113.942 |
| AMB_11 | 113.139 | 101.710 |
| AMB_12 | 129.833 | 124.982 |
| AMB_14 | 109.104 | 103.198 |
| AMB_17 | 122.300 | 117.615 |
| AMB_19 | 106.383 | 102.663 |
| AMB_20 | 189.422 | 178.120 |
| AMB_23 | 123.020 | 119.781 |
| AMB_24 | 91.292 | 83.955 |
| AMB_25 | 132.568 | 123.971 |
| AMB_26 | 93.526 | 86.726 |
| AMB_27 | 117.189 | 111.122 |
| AMB_28 | 107.732 | 104.268 |
| AMB_29 | 98.099 | 95.565 |
| AMB_30 | 113.263 | 105.115 |
| AMB_31 | 94.347 | 88.434 |
| AMB_32 | 115.549 | 109.574 |
| AMB_33 | 100.969 | 93.323 |
| AMB_35 | 102.132 | 94.000 |
| AMB_37 | 103.555 | 82.498 |
| AMB_38 | 210.100 | 194.380 |
| AMB_39 | 196.623 | 186.755 |
| AMB_40 | 86.280 | 81.544 |
| AMB_42 | 106.925 | 103.322 |
| AMB_43 | 117.272 | 108.577 |
| AMB_45 | 100.564 | 97.715 |
| AMB_47 | 102.441 | 91.258 |
| AMB_48 | 114.898 | 103.182 |
| AMB_49 | 127.234 | 118.067 |
| AMB_50 | 89.739 | 86.168 |
| AMB_52 | 115.236 | 112.181 |
| AMB_55 | 90.100 | 79.658 |
| AMB_56 | 116.581 | 113.890 |
| AMB_57 | 106.247 | 101.446 |
| AMB_58 | 90.085 | 85.513 |

| | | |
|---------|---------|---------|
| AMB_59 | 83.854 | 81.505 |
| AMB_60 | 95.114 | 90.621 |
| AMB_61 | 105.392 | 103.775 |
| AMB_62 | 106.060 | 96.933 |
| AMB_63 | 206.805 | 190.658 |
| AMB_64 | 90.232 | 88.069 |
| AMB_65 | 112.555 | 99.998 |
| AMB_67 | 108.767 | 103.782 |
| AMB_68 | 110.930 | 107.451 |
| AMB_70 | 139.117 | 131.816 |
| AMB_72 | 119.829 | 106.142 |
| AMB_73 | 103.088 | 95.039 |
| AMB_74 | 94.860 | 84.977 |
| AMB_75 | 117.187 | 97.832 |
| AMB_76 | 211.528 | 187.690 |
| AMB_77 | 119.068 | 116.713 |
| AMB_78 | 116.150 | 112.522 |
| AMB_80 | 123.329 | 113.972 |
| AMB_81 | 83.329 | 82.096 |
| AMB_83 | 92.759 | 88.716 |
| AMB_86 | 192.925 | 182.797 |
| AMB_87 | 107.259 | 95.345 |
| AMB_88 | 104.662 | 93.095 |
| AMB_89 | 108.630 | 107.543 |
| AMB_90 | 92.301 | 77.679 |
| AMB_91 | 124.166 | 120.709 |
| AMB_92 | 112.542 | 97.774 |
| AMB_93 | 98.151 | 77.576 |
| AMB_94 | 200.205 | 193.395 |
| AMB_95 | 111.869 | 105.892 |
| AMB_96 | 86.081 | 81.548 |
| AMB_97 | 93.518 | 87.961 |
| AMB_98 | 92.585 | 86.532 |
| AMB_99 | 100.808 | 96.275 |
| AMB_100 | 99.715 | 92.089 |
| AMB_102 | 110.779 | 106.919 |
| AMB_103 | 109.384 | 104.097 |
| AMB_104 | 112.566 | 108.979 |
| AMB_105 | 125.330 | 119.165 |
| AMB_106 | 115.908 | 112.305 |
| AMB_107 | 113.968 | 102.393 |
| AMB_108 | 127.297 | 110.045 |
| AMB_109 | 91.640 | 89.039 |
| AMB_110 | 90.319 | 84.908 |
| AMB_111 | 113.902 | 112.140 |
| AMB_112 | 121.226 | 112.723 |
| AMB_113 | 100.531 | 98.549 |
| AMB_114 | 116.227 | 106.760 |

| | | |
|---------|---------|---------|
| AMB_115 | 191.259 | 183.746 |
| AMB_116 | 106.756 | 101.688 |
| AMB_117 | 83.855 | 72.579 |
| AMB_118 | 92.649 | 90.017 |
| AMB_119 | 114.101 | 106.728 |
| AMB_120 | 114.443 | 111.873 |
| AMB_121 | 90.504 | 82.073 |
| AMB_122 | 113.818 | 107.628 |
| AMB_123 | 87.389 | 84.769 |
| AMB_125 | 105.528 | 96.392 |
| AMB_126 | 81.768 | 80.966 |
| AMB_127 | 82.362 | 79.819 |
| AMB_130 | 104.308 | 96.915 |
| AMB_131 | 106.230 | 100.148 |
| AMB_133 | 84.346 | 80.584 |
| AMB_136 | 99.889 | 95.224 |
| AMB_137 | 126.795 | 124.712 |
| AMB_138 | 79.103 | 76.651 |
| AMB_139 | 153.623 | 148.455 |
| AMB_140 | 156.340 | 147.694 |
| AMB_141 | 102.281 | 96.574 |
| AMB_142 | 119.797 | 116.618 |
| AMB_143 | 95.501 | 86.840 |
| AMB_144 | 98.810 | 90.315 |
| AMB_145 | 91.310 | 90.347 |
| AMB_146 | 109.887 | 105.686 |
| AMB_147 | 110.019 | 107.849 |
| AMB_148 | 116.770 | 108.991 |
| AMB_149 | 103.003 | 97.370 |
| AMB_150 | 98.588 | 94.222 |
| AMB_151 | 102.326 | 98.920 |
| AMB_152 | 97.648 | 88.739 |
| AMB_154 | 104.336 | 100.346 |
| AMB_155 | 126.669 | 122.757 |
| AMB_156 | 92.147 | 84.442 |
| AMB_157 | 88.578 | 84.218 |
| AMB_158 | 94.029 | 87.790 |
| AMB_159 | 100.139 | 95.157 |
| AMB_160 | 100.178 | 98.130 |
| AMB_161 | 98.790 | 91.302 |
| AMB_162 | 100.443 | 95.434 |
| AMB_163 | 108.176 | 103.808 |
| AMB_164 | 117.638 | 110.739 |
| AMB_165 | 113.401 | 101.432 |
| AMB_167 | 94.597 | 93.199 |
| AMB_168 | 121.668 | 110.601 |
| AMB_169 | 156.552 | 154.174 |
| AMB_170 | 135.981 | 123.096 |

| | | |
|---------|---------|---------|
| AMB_171 | 87.372 | 83.044 |
| AMB_172 | 198.586 | 191.550 |
| AMB_176 | 124.360 | 113.262 |
| AMB_177 | 135.773 | 125.039 |
| AMB_178 | 101.825 | 98.103 |
| AMB_179 | 103.021 | 102.008 |
| AMB_180 | 206.019 | 198.089 |
| AMB_181 | 160.649 | 156.777 |
| AMB_182 | 97.645 | 84.702 |
| AMB_183 | 179.110 | 166.130 |
| AMB_184 | 165.683 | 155.203 |
| AMB_185 | 173.327 | 156.101 |
| CT_1 | 110.121 | 93.960 |
| CT_2 | 134.033 | 124.265 |
| CT_3 | 110.352 | 102.586 |
| CT_4 | 203.841 | 193.446 |
| CT_5 | 143.748 | 130.566 |
| CT_6 | 160.765 | 148.768 |
| CT_7 | 201.043 | 186.287 |
| CT_10 | 204.821 | 194.612 |
| CT_11 | 96.200 | 87.591 |
| CT_12 | 209.303 | 198.000 |
| CT_13 | 200.813 | 186.576 |
| CT_15 | 195.972 | 190.884 |
| CT_18 | 192.038 | 181.251 |
| CT_19 | 188.936 | 176.824 |
| CT_20 | 78.977 | 71.055 |
| CT_21 | 108.303 | 100.439 |
| CT_22 | 220.063 | 208.111 |
| CT_23 | 105.415 | 93.111 |
| CT_28 | 194.213 | 176.225 |
| CT_29 | 106.087 | 101.815 |
| CT_31 | 93.195 | 90.144 |
| CT_32 | 76.980 | 69.977 |
| CT_33 | 71.560 | 69.664 |
| CT_34 | 177.282 | 160.982 |
| CT_35 | 86.418 | 85.351 |
| CT_36 | 197.368 | 176.841 |
| CT_37 | 89.466 | 83.014 |
| CT_38 | 82.780 | 79.266 |
| CT_40 | 115.244 | 107.686 |
| CT_45 | 100.660 | 95.019 |
| CT_46 | 120.722 | 110.705 |
| CT_47 | 108.692 | 107.661 |
| CT_48 | 109.967 | 107.939 |
| CT_49 | 101.053 | 100.221 |
| CT_50 | 99.755 | 92.418 |
| CT_51 | 178.011 | 164.181 |

| | | |
|---------------------------|-------------------|-------------------|
| CT_52 | 83.507 | 80.064 |
| CT_53 | 201.114 | 190.334 |
| CT_54 | 175.671 | 170.991 |
| CT_55 | 104.176 | 101.852 |
| CT_56 | 89.191 | 87.289 |
| CT_57 | 218.465 | 194.782 |
| CT_58 | 142.220 | 136.104 |
| CT_59 | 127.679 | 122.902 |
| CT_60 | 105.473 | 102.336 |
| CT_61 | 120.610 | 117.962 |
| CT_62 | 114.818 | 111.760 |
| CT_63 | 138.318 | 136.635 |
| CT_64 | 112.156 | 101.598 |
| CT_66 | 148.723 | 145.143 |
| Média de leituras | 120.193,66 | 112.976,74 |
| Mínimo de leituras | 71.560,00 | 69.664,00 |
| Máximo de leituras | 220.063,00 | 208.111,00 |

APÊNDICE 2. PARÂMETROS ECOLÓGICOS DE DIVERSIDADE A – ÍNDICES CHAO1, SHANNON E SIMPSON.

| Índices | Grupo | Média | DP | P-valor |
|----------------|-------|--------|--------|--------------|
| Chao1 | CT | 329,31 | 131,20 | - |
| | DII | 232,69 | 81,16 | 0,007 |
| | PP | 325,16 | 114,05 | 0,904 |
| | AR | 256,08 | 96,30 | 0,179 |
| | LES | 298,21 | 109,14 | 0,583 |
| | DM1 | 297,33 | 100,39 | 0,73 |
| | DM2 | 261,84 | 94,53 | 0,179 |
| Shannon | CT | 4,14 | 0,51 | - |
| | DII | 3,88 | 0,51 | 0,615 |
| | PP | 4,21 | 0,60 | 0,951 |
| | AR | 3,92 | 0,63 | 0,672 |
| | LES | 4,03 | 0,53 | 0,672 |
| | DM1 | 4,22 | 0,37 | 0,672 |
| | DM2 | 3,90 | 0,48 | 0,615 |
| Simpson | CT | 0,96 | 0,03 | - |
| | DII | 0,95 | 0,03 | 0,159 |
| | PP | 0,96 | 0,04 | 0,728 |
| | AR | 0,95 | 0,05 | 0,397 |
| | LES | 0,95 | 0,04 | 0,625 |
| | DM1 | 0,97 | 0,02 | 0,773 |
| | DM2 | 0,95 | 0,03 | 0,201 |

P-valor da comparação entre grupo CT e cada grupo enfermo. Legenda: CT, controles; DII, doença inflamatória intestinal (a); PP, psoríase em placas (b); AR, artrite reumatoide(c); LES, Lupus eritematoso sistêmico (d); DM1, diabetes tipo 1 (e); DM2, diabetes tipo 2 (f).

APÊNDICE 3. COMPARAÇÃO DE DADOS FENOTÍPICOS CONTENDO: SOCIODEMOGRÁFICOS, MEDICAMENTOS DE USO CONTÍNUO, ESTILO DE VIDA E CONSUMO DIETÉTICO DE NUTRIENTES PARA PRÉ-SELEÇÃO DE VARIÁVEIS

| Variáveis | DII | PP | AR | LES | DM1 | DM2 |
|---|-----|----|----|-----|-----|-----|
| Sociodemográfico | | | | | | |
| Idade | | | | | | |
| Gênero | | | | | | |
| Etnia | | | | | | |
| Medicamentos uso contínuos | | | | | | |
| Vitaminas ou suplementos | | | | | | |
| Antidepressivos | | | | | | |
| Antieméticos | | | | | | |
| Anti-hipertensivo | | | | | | |
| Anticonvulsivante | | | | | | |
| Analgésico | | | | | | |
| Antipsicótico | | | | | | |
| Metotrexato | | | | | | |
| Antimalárico | | | | | | |
| Antirreumático | | | | | | |
| Corticosteroides | | | | | | |
| Reposição de enzimas | | | | | | |
| Estatinas | | | | | | |
| Imunossupressor | | | | | | |
| Insulina | | | | | | |
| Miorrelaxante | | | | | | |
| Anticorpo monoclonal | | | | | | |
| AINES | | | | | | |
| Estilo de vida | | | | | | |
| Q1. Tem acima de 60 anos? | | | | | | |
| Q2. Tipo de parto? | | | | | | |
| Q3. Foi amamentado com leite materno? | | | | | | |
| Q5. Consumo de Frutas, Verduras, Legumes e Cereais integrais por dia | | | | | | |
| Q6. Consumo de preparações caseiras com adição de açúcar refinado ou adoçantes artificiais por semana | | | | | | |
| Q7. Açúcar refinado ou adoçante? | | | | | | |
| Q8. Consumo de alimentos ultraprocessados por semana? | | | | | | |
| Q9. Consome bebida alcoólica? Quantas doses por semana? | | | | | | |
| Q10. Prática de Atividade Física dias por semana | | | | | | |
| Q11. Prática de Atividade Física minutos por semana | | | | | | |
| Q12. Nível de estresse percebido | | | | | | |
| Q13. É tabagista? | | | | | | |

| | | | | | |
|--|--|--|--|--|--|
| Q14. Utilizou antibiótico nos últimos 3 meses | | | | | |
| Q15. Fez uso agudo de IBP, Laxantes, Anti-inflamatórios (AINES), Corticoides, Metformina (nos últimos 60 dias) | | | | | |
| Q16. É fisicamente ativo? (> 150 min/semana) | | | | | |
| Escala de Bristol | | | | | |
| Composição corpórea e IMC | | | | | |
| Peso (kg) | | | | | |
| IMC (kg/m ²) | | | | | |
| IMC classificado | | | | | |
| Gordura (%) | | | | | |
| Massa magra (%) | | | | | |
| Consumo dietético | | | | | |
| Calorias | | | | | |
| Proteínas | | | | | |
| Carboidratos | | | | | |
| Gordura total | | | | | |
| Gordura saturada | | | | | |
| Gordura monoinsaturada | | | | | |
| Gordura poli-insaturada | | | | | |
| Colesterol | | | | | |
| Calcio | | | | | |
| Ferro | | | | | |
| Zinco | | | | | |
| Vit B1 | | | | | |
| Vit C | | | | | |
| Vit E | | | | | |
| Vit B6 | | | | | |
| Potássio | | | | | |
| Magnésio | | | | | |
| Selênio | | | | | |
| Folato | | | | | |
| Fosforo | | | | | |
| Cobre | | | | | |
| Vit B3 | | | | | |
| Vit B12 | | | | | |
| Manganês | | | | | |

Legenda: as células pintadas em laranja indicam diferença significativa entre controles e o grupo avaliado. Vit, vitamina; CT, controles; DII, doença inflamatória intestinal (a); PP, psoríase em placas (b); AR, artrite reumatoide(c); LES, Lupus eritematoso sistêmico (d); DM1, diabetes tipo 1 (e); DM2, diabetes tipo 2 (f)

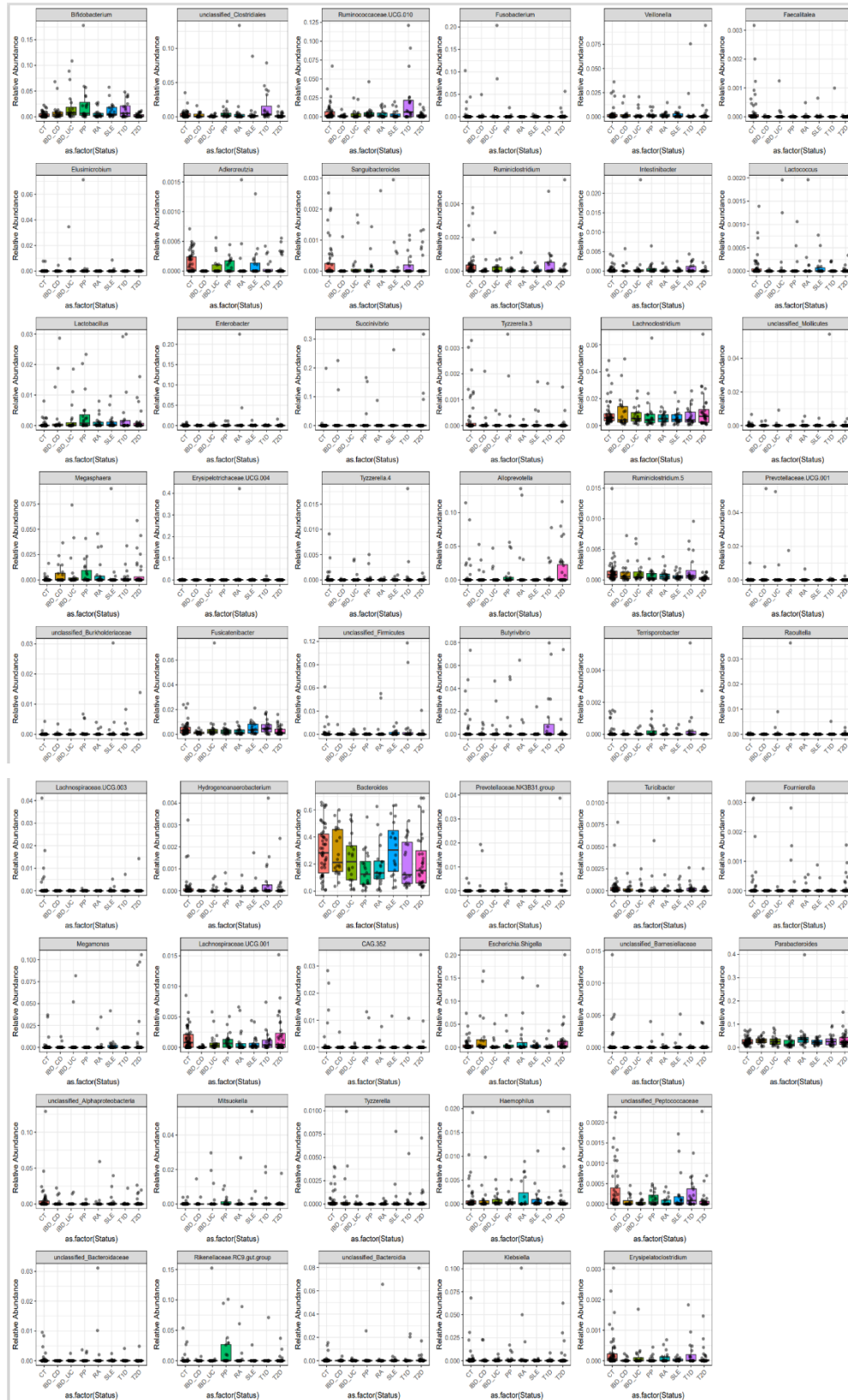
APÊNDICE 4. ASSOCIAÇÃO ENTRE VARIÁVEIS DE FATORES AMBIENTAIS E ESTILO DE VIDA E MACROESTRUTURA DA MICROBIOTA INTESTINAL

| Variáveis | SumOfSqs | R2 | P ajustado |
|---|--------------|--------------|--------------|
| Enfermidade | 1,854 | 0,044 | 0,046 |
| Metformina | 0,352 | 0,008 | 0,207 |
| Suplementos vitamínicos | 0,347 | 0,008 | 0,207 |
| Corticoides (uso agudo) | 0,349 | 0,008 | 0,207 |
| Antipsicóticos | 0,337 | 0,008 | 0,272 |
| Q12. Nível de estresse percebido | 0,897 | 0,021 | 0,272 |
| Consumo de bebida alcoólica | 0,897 | 0,021 | 0,272 |
| Antidepressivos | 0,321 | 0,008 | 0,272 |
| Q13. É tabagista? | 0,317 | 0,008 | 0,272 |
| Antirreumático | 0,313 | 0,007 | 0,272 |
| Corticosteroides (uso contínuo) | 0,311 | 0,007 | 0,272 |
| Contraceptivo oral | 0,3 | 0,007 | 0,387 |
| Q5. Consumo de Frutas, Verduras, Legumes e Cereais integrais por dia | 0,861 | 0,02 | 0,389 |
| Laxantes (uso agudo) | 0,297 | 0,007 | 0,404 |
| Sulfassalazina ou Mesalazina | 0,293 | 0,007 | 0,491 |
| Q3. Foi amamentado com leite materno? | 0,289 | 0,007 | 0,508 |
| Anti-hipertensivo | 0,286 | 0,007 | 0,508 |
| Metotrexato | 0,281 | 0,007 | 0,508 |
| Retinoide | 0,286 | 0,007 | 0,508 |
| Q1. Tem acima de 60 anos? | 0,277 | 0,007 | 0,585 |
| Antieméticos | 0,277 | 0,007 | 0,585 |
| Q16 | 0,275 | 0,007 | 0,585 |
| Insulina | 0,275 | 0,007 | 0,585 |
| Q10. Prática de Atividade Física (dias por semana) | 1,879 | 0,045 | 0,585 |
| Inibidor da bomba de prótons (uso contínuo) | 0,272 | 0,006 | 0,585 |
| Analgésico | 0,268 | 0,006 | 0,651 |
| Antidiabético Oral | 0,266 | 0,006 | 0,673 |
| Q15. Fez uso agudo de IBP, Laxantes, (AINES), Corticoides (últimos 60 dias) | 0,265 | 0,006 | 0,69 |
| Escala de Bristol | 1,578 | 0,037 | 0,694 |
| Hormônio tireoidiano | 0,258 | 0,006 | 0,694 |
| Atividade física (OMS) | 0,257 | 0,006 | 0,694 |
| Anticorpo monoclonal | 0,259 | 0,006 | 0,694 |
| Q9. Consome bebida alcoólica? Quantas doses por semana? | 0,516 | 0,012 | 0,694 |
| Relaxante muscular | 0,251 | 0,006 | 0,777 |
| Antimaláricos | 0,25 | 0,006 | 0,777 |
| Inibidor da bomba de prótons (uso agudo) | 0,245 | 0,006 | 0,838 |
| Imunossupressores | 0,239 | 0,006 | 0,865 |
| Q6. Consumo de preparações caseiras com adição de açúcar refinado ou adoçantes artificiais por semana | 1,002 | 0,024 | 0,865 |
| Q2. Tipo de parto | 0,239 | 0,006 | 0,865 |
| Consumo de alimentos Pre, Pro ou | 0,238 | 0,006 | 0,865 |

| Simbióticos | | | |
|---|-------|-------|-------|
| Hipolipemiante | 0,237 | 0,006 | 0,865 |
| Antiagregante plaquetário | 0,236 | 0,006 | 0,865 |
| Anti-inflamatório não esteroideal (AINEs) (uso contínuo) | 0,234 | 0,006 | 0,865 |
| Q8. Consumo de alimentos ultraprocessados por semana? | 1,238 | 0,029 | 0,898 |
| Anti-inflamatório não esteroideal (AINEs) (uso agudo) | 0,228 | 0,005 | 0,898 |
| Q7. Açúcar refinado ou adoçante? | 0,212 | 0,005 | 0,984 |

Tamanhos de efeito (R²) calculados por testes de análise multivariada de variância multivariada permutacional (PERMANOVA) e correspondentes p ajustados-valores por Benjamini-Hochberg (BH).

APÊNDICE 5. BOXPLOTS DOS TÁXONS MICROBIANOS PRÉ-SELECIONADAS PARA COMPOR OS MODELOS PREDITIVOS



APÊNDICE 6. BACTÉRIAS PRÉ-SELECIONADAS PARA MODELOS PREDITIVOS

| Táxon microbiano | Média dos valores | log2 Fold-change | P ajustado | Grupo teste | Grupo referência |
|----------------------------------|-------------------|------------------|------------|-------------|------------------|
| Bifidobacterium | 1.021.223 | -6.564 | 0.013 | Controle | Todos |
| Elusimicrobium | 129.212 | -1,18E-05 | 0.032 | Controle | Todos |
| Lactobacillus | 209.817 | -6,91E-03 | 0.011 | Controle | Todos |
| Megasphaera | 750.958 | -1,80E-03 | 0.004 | Controle | Todos |
| unclassified_Burkholderiaceae | 21.398 | -7,81E-04 | 0.005 | Controle | Todos |
| Lachnospiraceae UCG-003 | 9.992 | -5,09E-04 | 9,10E-04 | RCU | Todos |
| Megamonas | 139.405 | 2,65E-04 | 8,94E-17 | RCU | Todos |
| unclassified_Alphaproteobacteria | 250.016 | -3,69E-04 | 9,10E-04 | RCU | Todos |
| unclassified_Bacteroidaceae | 18.160 | 3,26E-05 | 0.039 | RCU | Todos |
| unclassified_Clostridiales | 401.122 | -4,95E-03 | 0.014 | RCU | Todos |
| Adlercreutzia | 9.091 | -1,35E-03 | 5,75E-02 | DC | Todos |
| Elusimicrobium | 18.322 | -1,17E-04 | 4,84E-07 | DC | Todos |
| Enterobacter | 61.863 | -8,03E-04 | 0.023 | DC | Todos |
| Erysipelotrichaceae UCG-004 | 14.208 | 8,89E-05 | 4,71E-13 | DC | Todos |
| Fusicatenibacter | 397.250 | -5,70E-03 | 0.034 | DC | Todos |
| Hydrogenoanaerobacterium | 9.568 | -3,69E-03 | 0.006 | DC | Todos |
| Lachnospiraceae UCG-001 | 124.135 | 2,11E-04 | 9,21E-03 | DC | Todos |
| Lachnospiraceae UCG-003 | 10.183 | -9,26E-04 | 1,47E-05 | DC | Todos |
| Megamonas | 190.022 | -1,87E-04 | 0.011 | DC | Todos |
| Mitsuokella | 41.825 | -2,02E-04 | 1,02E-14 | DC | Todos |
| Rikenellaceae RC9 gut group | 172.683 | -2,05E-04 | 4,26E-18 | DC | Todos |
| Ruminococcaceae UCG-010 | 544.855 | -2,10E-03 | 0.019 | DC | Todos |
| Sanguibacteroides | 16.885 | -3,52E-04 | 3,30E-43 | DC | Todos |
| Succinivibrio | 226.934 | 8,11E-05 | 0.032 | DC | Todos |
| Tyzzereella 4 | 9.941 | -3,00E-04 | 3,66E-17 | DC | Todos |
| unclassified_Bacteroidaceae | 19.197 | -5,28E-04 | 1,53E-32 | DC | Todos |
| unclassified_Firmicutes | 95.393 | -1,13E-03 | 0.002 | DC | Todos |
| Bacteroides | 30.193.115 | -9,10E-03 | 0.003 | PP | Todos |
| CAG-352 | 160.162 | 6,99E-05 | 8,80E-12 | PP | Todos |
| Megamonas | 201.115 | -2,65E-04 | 1,39E-18 | PP | Todos |
| Tyzzereella | 21.572 | -2,30E-04 | 3,78E-56 | PP | Todos |
| unclassified_Alphaproteobacteria | 260.303 | -5,33E-04 | 0.030 | PP | Todos |
| unclassified_Bacteroidia | 36.348 | -6,58E-04 | 0.005 | PP | Todos |
| Elusimicrobium | 19.722 | -1,37E-04 | 9,60E-07 | AR | Todos |
| Fusobacterium | 92.456 | -5,49E-04 | 0.011 | AR | Todos |
| Lachnospiraceae UCG-003 | 10.183 | 2,87E-04 | 2,55E-05 | AR | Todos |

| | | | | | |
|-----------------------------------|-----------|-----------|----------|----------|-------|
| Ruminiclostridium | 16.833 | -1,31E-03 | 0.0004 | AR | Todos |
| Sanguibacteroides | 16.572 | -4,35E-04 | 2,31E-43 | AR | Todos |
| Tyzzarella 3 | 12.117 | -1,53E-03 | 3,85E-24 | AR | Todos |
| Alloprevotella | 1.163.400 | -9,95E-04 | 0.0002 | LES | Todos |
| Butyrivibrio | 302.045 | -3,36E-04 | 0.004 | LES | Todos |
| Prevotellaceae NK3B31 group | 9.602 | -6,76E-05 | 3,33E-03 | LES | Todos |
| Rikenellaceae RC9 gut group | 146.987 | -1,79E-04 | 6,31E-18 | LES | Todos |
| Elusimicrobium | 19.717 | -1,38E-04 | 5,54E-07 | DM1 | Todos |
| Escherichia-Shigella | 1.021.279 | -3,23E-03 | 0.002 | DM1 | Todos |
| Fusobacterium | 92.220 | -5,99E-04 | 0.007 | DM1 | Todos |
| Haemophilus | 135.859 | -1,17E-03 | 0.010 | DM1 | Todos |
| Klebsiella | 145.637 | -9,12E-04 | 0.028 | DM1 | Todos |
| Megamonas | 200.969 | -2,34E-04 | 7,16E-16 | DM1 | Todos |
| Ruminococcaceae UCG-010 | 544.855 | 2,23E-03 | 0.010 | DM1 | Todos |
| unclassified_Alphaproteo bacteria | 248.915 | -3,23E-04 | 0.012 | DM1 | Todos |
| unclassified_Bacteroidia | 363.488 | 6,20E-04 | 0.004 | DM1 | Todos |
| unclassified_Clostridiales | 401.122 | 4,77E-03 | 0.007 | DM1 | Todos |
| Veillonella | 247.386 | -3,04E-03 | 0.004 | DM1 | Todos |
| Bifidobacterium | 1.021.223 | 5,85E-03 | 0.026 | DM2 | Todos |
| Intestinibacter | 46.439 | -1,42E-03 | 0.048 | DM2 | Todos |
| Lachnoclostridium | 1.069.603 | 0.869 | 0.015 | DM2 | Todos |
| Ruminiclostridium 5 | 87.751 | -8,70E-03 | 0.005 | DM2 | Todos |
| Terrisporobacter | 8.614 | -2.041 | 0.019 | DM2 | Todos |
| Turicibacter | 28.195 | -1,51E-03 | 0.048 | DM2 | Todos |
| unclassified_Bacteroidaceae | 16.467 | -7,10E-04 | 0.015 | DM2 | Todos |
| unclassified_Firmicutes | 98.502 | 5,12E-03 | 0.026 | DM2 | Todos |
| Bifidobacterium | 908.355 | -6,14E-03 | 0.033 | Controle | RCU |
| Lachnospiraceae UCG-003 | 36.765 | 3,71E-04 | 5,62E-12 | Controle | RCU |
| Tyzzarella | 37.516 | 9,65E-04 | 0.008 | Controle | RCU |
| unclassified_Alphaproteo bacteria | 227.233 | 7,81E-04 | 0.0001 | Controle | RCU |
| unclassified_Barnesiellaceae | 10.941 | 2,18E-04 | 7,34E-07 | Controle | RCU |
| unclassified_Peptococcaceae | 25.758 | 2.818 | 0.008 | Controle | RCU |
| Adlercreutzia | 10.461 | 9.268 | 2,11E-02 | Controle | DC |
| Bifidobacterium | 885.503 | -1.528 | 0.041 | Controle | DC |
| Erysipelatoclostridium | 20.349 | 2.050 | 0.021 | Controle | DC |
| Escherichia-Shigella | 1.056.279 | -2.520 | 0.028 | Controle | DC |
| Faecalitalea | 129.239 | 9,01E-04 | 0.017 | Controle | DC |
| Hydrogenoanaerobacterium | 5.654 | 1,18E-03 | 0.025 | Controle | DC |
| Lachnospiraceae UCG-001 | 115.641 | 3.600 | 4,17E-02 | Controle | DC |
| Lachnospiraceae UCG-003 | 36.839 | -1,06E-03 | 5,65E-12 | Controle | DC |
| Lactococcus | 6.441 | 1,16E-03 | 0.0182 | Controle | DC |
| Rikenellaceae RC9 gut group | 24.870 | 1,50E-04 | 3,74E-08 | Controle | DC |
| Sanguibacteroides | 18.982 | 6,85E-04 | 8,77E-04 | Controle | DC |

| | | | | | |
|----------------------------------|------------|-----------|----------|----------|-----|
| Terrisporobacter | 9.376 | 3,61E-04 | 1,10E-14 | Controle | DC |
| unclassified_Bacteroidaceae | 9.580 | 7,06E-04 | 0.022 | Controle | DC |
| unclassified_Barnesiellaceae | 15.680 | 2,21E-04 | 1,09E-06 | Controle | DC |
| unclassified_Firmicutes | 60.348 | -7,01E-04 | 0.018 | Controle | DC |
| Bacteroides | 33.625.397 | 1.261 | 0.002 | Controle | PP |
| Bifidobacterium | 998.592 | -5,46E-03 | 0.038 | Controle | PP |
| CAG-352 | 9.247 | 7,42E-05 | 6,59E-06 | Controle | PP |
| Faecalitalea | 13.741 | 9,39E-04 | 0.005 | Controle | PP |
| Tyzzereella | 37.348 | 11.779 | 9,91E-64 | Controle | PP |
| unclassified_Alphaproteobacteria | 262.640 | 9,37E-04 | 0.004 | Controle | PP |
| Fusobacterium | 137.902 | 6.162 | 0.002 | Controle | AR |
| Lachnospiraceae UCG-003 | 37.633 | -1,68E-04 | 6,68E-12 | Controle | AR |
| Ruminiclostridium | 199.439 | 1,64E-03 | 0.0009 | Controle | AR |
| Sanguibacteroides | 19.135 | 5,22E-04 | 2,70E-02 | Controle | AR |
| Rikenellaceae RC9 gut group | 29.275 | 3,13E-03 | 7,39E-08 | Controle | LES |
| Faecalitalea | 13.604 | -5,21E-04 | 0.001 | Controle | DM1 |
| Fusobacterium | 142.077 | 5.464 | 0.001 | Controle | DM1 |
| Lachnospiraceae UCG-003 | 41.589 | 2,06E-04 | 2,60E-12 | Controle | DM1 |
| Tyzzereella 3 | 29.661 | 4,69E-04 | 0.001 | Controle | DM1 |
| unclassified_Alphaproteobacteria | 256.409 | 7,44E-04 | 0.008 | Controle | DM1 |
| unclassified_Barnesiellaceae | 12.172 | 2,06E-04 | 3,90E-07 | Controle | DM1 |
| unclassified_Clostridiales | 757.546 | -1.989 | 0.014 | Controle | DM1 |
| Veillonella | 265.256 | -3,64E-04 | 0.031 | Controle | DM1 |
| Faecalitalea | 10.601 | 4.574 | 0.004 | Controle | DM2 |
| Lachnospiraceae UCG-003 | 30.324 | 2,41E-04 | 2,49E-17 | Controle | DM2 |
| Ruminiclostridium 5 | 88.265 | 1.330 | 0.0008 | Controle | DM2 |
| Terrisporobacter | 7.619 | 3,24E-04 | 2,49E-17 | Controle | DM2 |
| Turicibacter | 33.255 | -1,39E-04 | 0.010 | Controle | DM2 |
| unclassified_Bacteroidaceae | 8.553 | 6,58E-04 | 0.035 | Controle | DM2 |
| unclassified_Bacteroidia | 73.719 | -5,48E-03 | 0.016 | Controle | DM2 |
| unclassified_Mollicutes | 74.284 | 6,64E-04 | 0.036 | Controle | DM2 |
| Mitsuokella | 22.898 | 2,19E-04 | 2,90E-11 | RCU | DC |
| Prevotellaceae UCG-001 | 18.846 | -8,99E-04 | 2,61E-10 | RCU | DC |
| Raoultella | 5.595 | 4,39E-04 | 3,22E-09 | RCU | DC |
| Rikenellaceae RC9 gut group | 5.010 | 1,99E-04 | 3,86E-09 | RCU | DC |
| unclassified_Alphaproteobacteria | 27.823 | -7,48E-05 | 1,52E-12 | RCU | DC |
| Prevotellaceae UCG-001 | 29.673 | 2,46E-04 | 1,12E-11 | RCU | PP |
| Raoultella | 7.287 | -7,29E-04 | 1,25E-09 | RCU | PP |
| Megamonas | 4.673 | 2,65E-04 | 8,31E-08 | RCU | AR |
| Prevotellaceae UCG-001 | 19.553 | 2,85E-04 | 5,13E-11 | RCU | AR |
| Raoultella | 6.084 | -3,42E-05 | 2,65E-09 | RCU | AR |
| unclassified_Alphaproteobacteria | 22.200 | -3,55E-03 | 5,13E-11 | RCU | AR |
| Megamonas | 69.119 | 4,96E-04 | 0.0001 | RCU | LES |

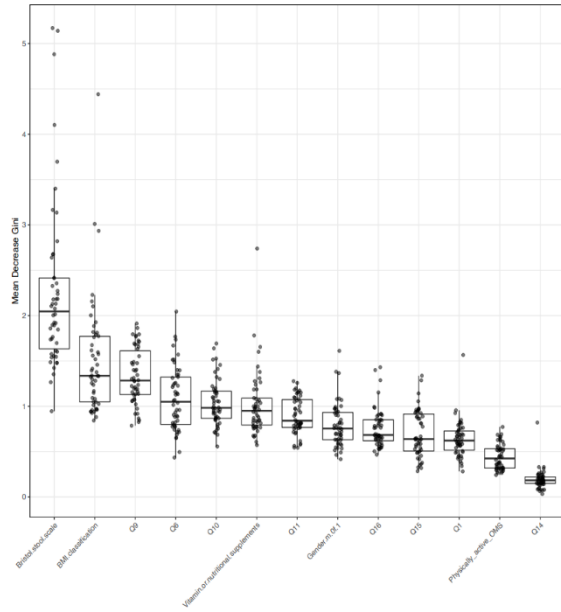
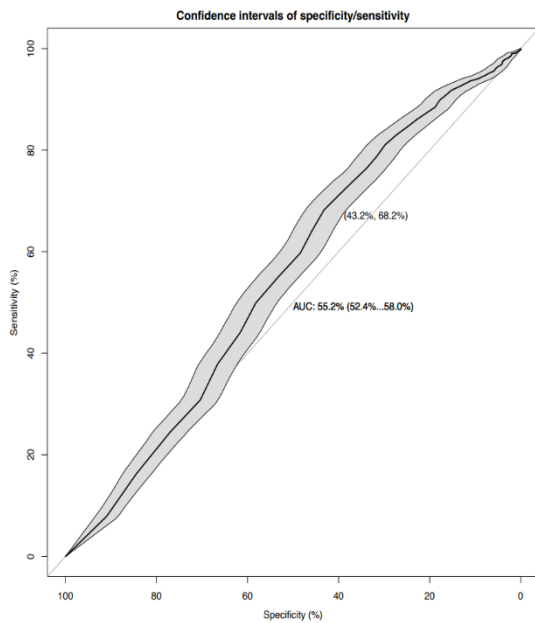
| | | | | | |
|-----------------------------------|------------|-----------|----------|-----|-----|
| Prevotellaceae UCG-001 | 21.574 | 3,80E-04 | 5,91E-10 | RCU | LES |
| Rikenellaceae RC9 gut group | 5.425 | 2,62E-04 | 3,59E-09 | RCU | LES |
| unclassified_Alphaproteo bacteria | 33.722 | -1,03E-03 | 3,20E-14 | RCU | LES |
| Haemophilus | 85.354 | 9,11E-04 | 0.031 | RCU | DM1 |
| Raoultella | 7.560 | 9,80E-05 | 2,34E-09 | RCU | DM1 |
| unclassified_Clostridiales | 732.219 | -3,26E-03 | 0.0003 | RCU | DM1 |
| Bifidobacterium | 822.804 | 1.917 | 0.035 | RCU | DM2 |
| Megamonas | 661.403 | -1,19E-04 | 3,94E-24 | RCU | DM2 |
| Prevotellaceae NK3B31 group | 11.011 | -6,36E-04 | 4,38E-10 | RCU | DM2 |
| Prevotellaceae UCG-001 | 15.549 | 4,53E-04 | 3,26E-11 | RCU | DM2 |
| Ruminiclostridium 5 | 58.889 | 1.282 | 0.0351 | RCU | DM2 |
| Succinivibrio | 427.417 | 5,01E-04 | 0.0351 | RCU | DM2 |
| Bacteroides | 22.972.209 | 1.100 | 0.0433 | DC | PP |
| Elusimicrobium | 6.343 | -3,62E-04 | 7,93E-09 | DC | PP |
| Mitsuokella | 14.189 | -2,59E-03 | 7,66E-10 | DC | PP |
| Erysipelotrichaceae UCG-004 | 7.670 | -3,60E-04 | 1,41E-09 | DC | AR |
| Rikenellaceae RC9 gut group | 35.499 | -2,86E-04 | 1,01E-11 | DC | AR |
| unclassified_Bacteroidaceae | 13.865 | -4,13E-04 | 2,18E-11 | DC | AR |
| Megamonas | 67.137 | -2,56E-04 | 0.016 | DC | LES |
| Mitsuokella | 44.375 | -2,68E-04 | 3,06E-11 | DC | LES |
| Escherichia-Shigella | 1.173.709 | 3.487 | 0.002 | DC | DM1 |
| Mitsuokella | 6.670 | -2,99E-05 | 5,77E-10 | DC | DM1 |
| Ruminococcaceae UCG-010 | 1.112.483 | -2,29E-03 | 0.002 | DC | DM1 |
| Tyzzrella 4 | 7.233 | -3,47E-04 | 5,77E-10 | DC | DM1 |
| Fournierella | 10.456 | -2,29E-04 | 6,85E-17 | DC | DM2 |
| Lachnospiraceae UCG-001 | 123.913 | -2,02E-03 | 0.002 | DC | DM2 |
| Rikenellaceae RC9 gut group | 16.475 | -1,88E-04 | 1,50E-10 | DC | DM2 |
| Succinivibrio | 399.382 | -2,36E-03 | 0.008 | DC | DM2 |
| unclassified_Bacteroidia | 70.132 | -2,88E-03 | 0.001 | DC | DM2 |
| Elusimicrobium | 6.537 | 3,91E-04 | 6,23E-09 | PP | AR |
| Megamonas | 6.157 | -3,61E-04 | 4,39E-10 | PP | AR |
| Parabacteroides | 2.475.078 | -2.503 | 0.034 | PP | AR |
| Alloprevotella | 135.542 | 6.733 | 0.012 | PP | LES |
| Bacteroides | 27.969.673 | -8,78E-03 | 0.005 | PP | LES |
| Megamonas | 85.375 | -1,82E-04 | 9,93E-19 | PP | LES |
| Rikenellaceae RC9 gut group | 763.209 | 6,58E-04 | 0.001 | PP | LES |
| Tyzzrella | 9.239 | -3,88E-04 | 3,91E-10 | PP | LES |
| unclassified_Firmicutes | 91.926 | -1,14E-03 | 0.037 | PP | LES |
| Elusimicrobium | 7.695 | 3,37E-04 | 3,68E-09 | PP | DM1 |
| Rikenellaceae RC9 gut group | 880.785 | 6,94E-04 | 0.0002 | PP | DM1 |
| Bacteroides | 24.025.339 | -0.0004 | 0.0429 | PP | DM2 |
| Megamonas | 762.835 | -3,70E-04 | 1,54E-23 | PP | DM2 |
| Parabacteroides | 3.501.956 | -7,28E-03 | 0.049 | PP | DM2 |
| Tyzzrella | 8.960 | -6,00E-04 | 1,26E-22 | PP | DM2 |
| unclassified_Bacteroidia | 40.472 | -3,60E-03 | 0.001 | PP | DM2 |

| | | | | | |
|-----------------------------|-----------|-----------|----------|-----|-----|
| Rikenellaceae RC9 gut group | 43.368 | 2,53E-04 | 1,46E-10 | AR | LES |
| Megamonas | 5.590 | 3,70E-04 | 7,50E-09 | AR | DM1 |
| Alloprevotella | 157.883 | -1,07E-03 | 0.014 | LES | DM1 |
| Megamonas | 80.668 | 6,00E-04 | 4,95E-18 | LES | DM1 |
| unclassified_Clostridiales | 885.206 | -1,57E-03 | 0.024 | LES | DM1 |
| Alloprevotella | 1.780.036 | -1,06E-03 | 2,75E-04 | LES | DM2 |
| Escherichia-Shigella | 993.099 | -1,60E-03 | 0.026 | LES | DM2 |
| Prevotellaceae NK3B31 group | 12.098 | -3,14E-04 | 3,29E-10 | LES | DM2 |
| Rikenellaceae RC9 gut group | 17.443 | -5,47E-04 | 1,71E-10 | LES | DM2 |
| Succinivibrio | 434.472 | -1,33E-04 | 0.027 | LES | DM2 |
| Escherichia-Shigella | 1.032.233 | -3,40E-03 | 0.032 | DM1 | DM2 |
| Lachnoclostridium | 1.329.394 | -5,74E-03 | 0.039 | DM1 | DM2 |
| Megamonas | 839.768 | -3,73E-04 | 9,82E-25 | DM1 | DM2 |
| Ruminiclostridium 5 | 67.212 | 1.314 | 0.032 | DM1 | DM2 |
| Ruminococcaceae UCG-010 | 1.005.870 | 2.487 | 0.032 | DM1 | DM2 |
| Succinivibrio | 463.131 | -6,78E-04 | 0.025 | DM1 | DM2 |
| unclassified_Bacteroidia | 82.930 | -6,47E-06 | 0.0001 | DM1 | DM2 |
| unclassified_Clostridiales | 664.973 | 2.576 | 0.025 | DM1 | DM2 |

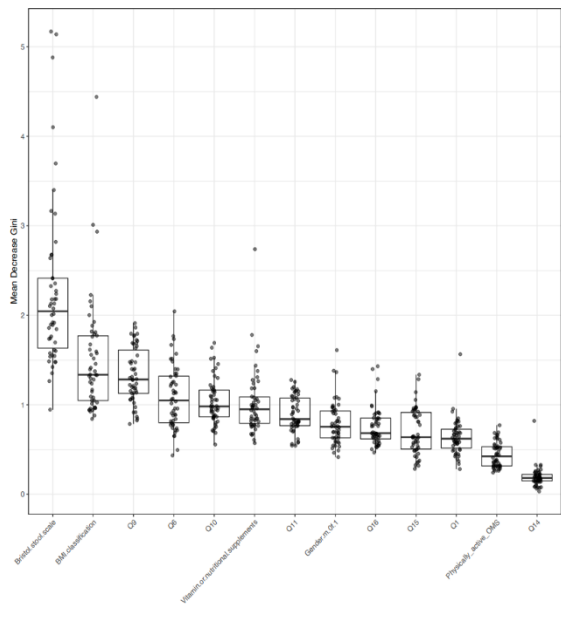
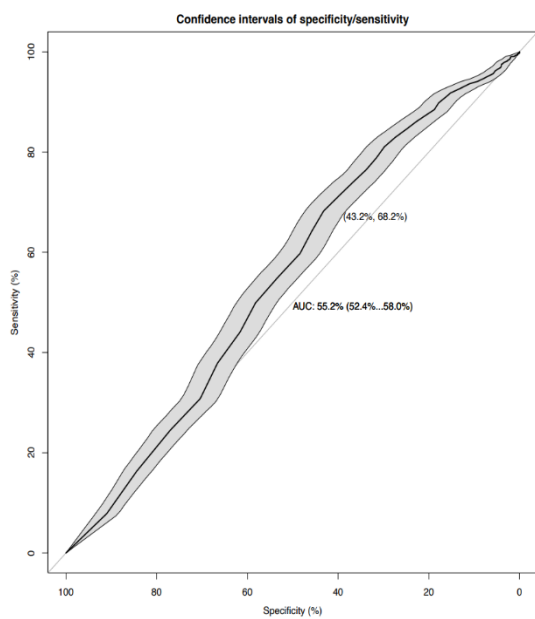
Legenda: CT, controles; DC, doença de Chron; RCU, retocolite ulcerativa PP, psoríase em placas; AR, artrite reumatoide; LES, Lupus eritematoso sistêmico; DM1, diabetes tipo 1; DM2, diabetes tipo 2

APÊNDICE 7. CURVAS ROC E ÍNDICE GINI - CAPACIDADE PREDITIVA DAS VARIÁVEIS FENOTÍPICAS QUALITATIVAS

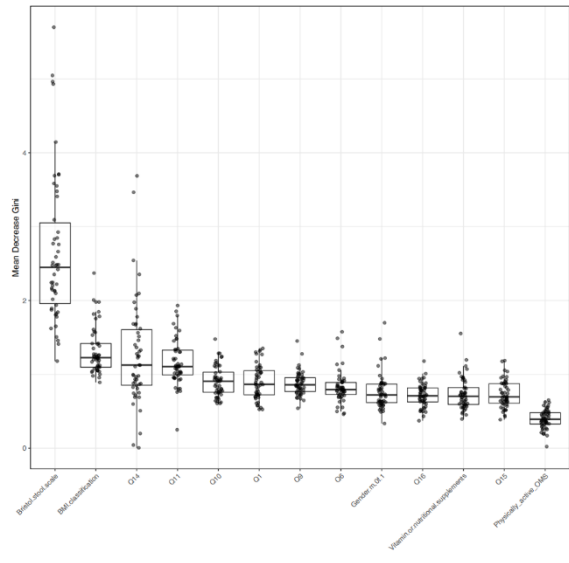
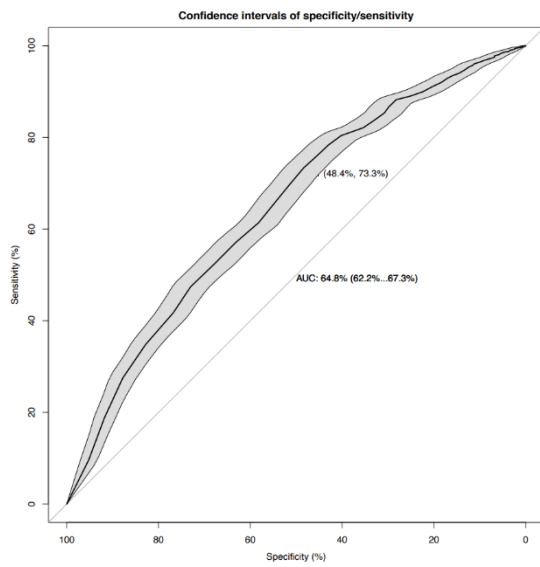
a. Doença de Crohn



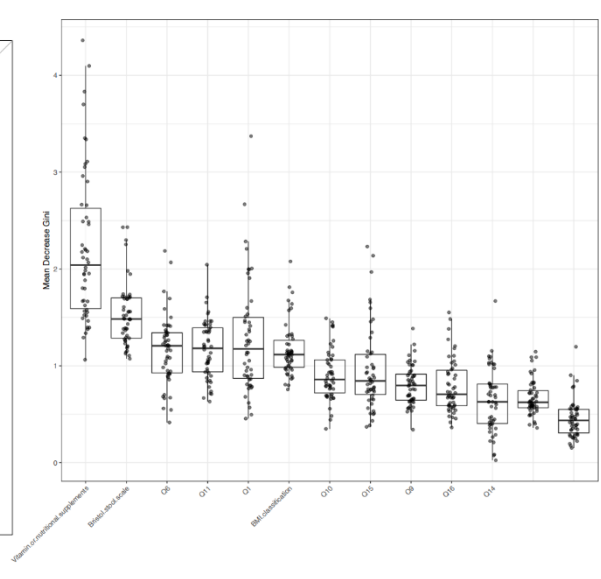
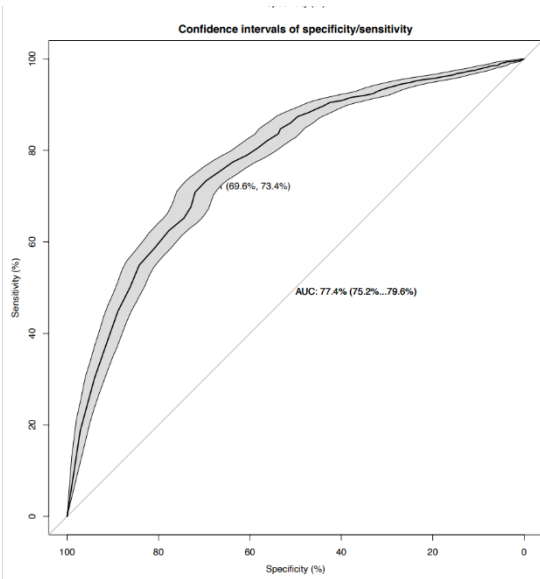
b. RCU



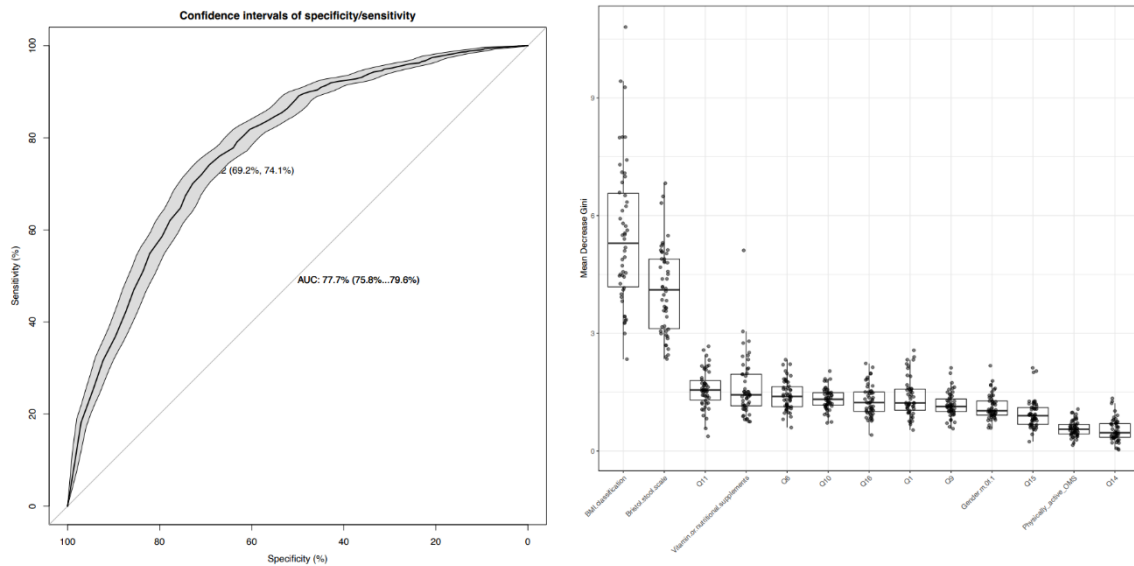
c. Psoríase em placas



d. Artrite reumatoide



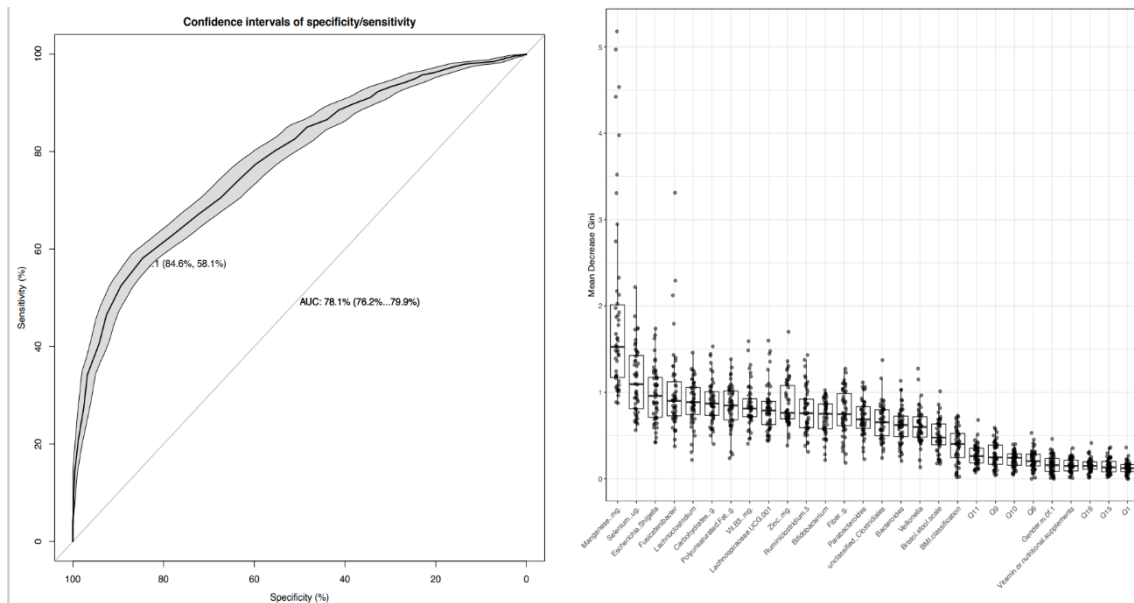
g. Diabetes tipo 2



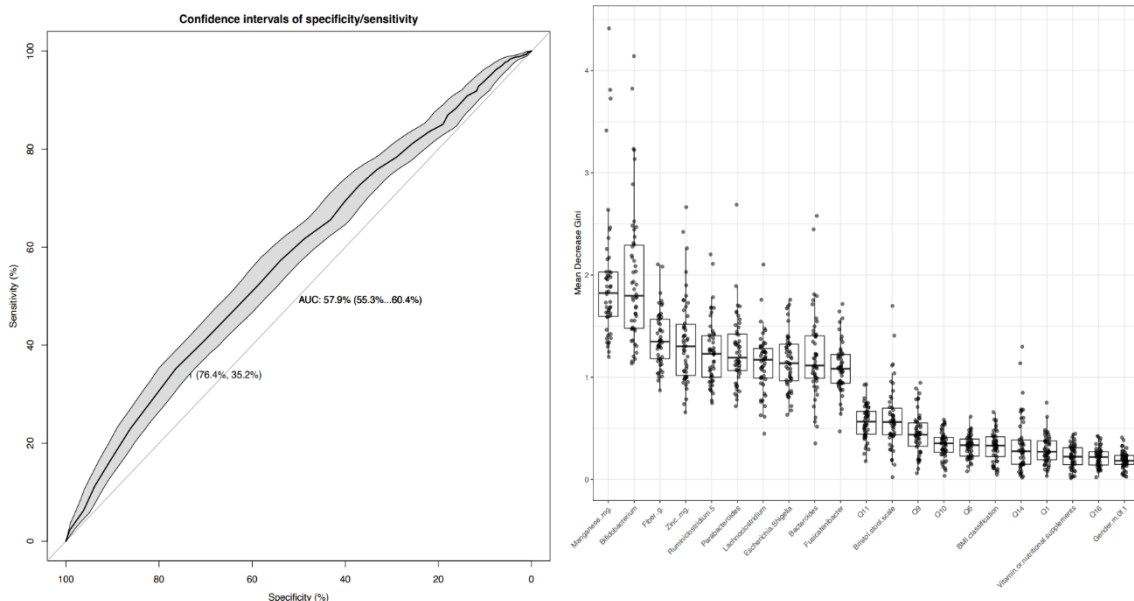
Legenda: AUC, área sob a curva. Os intervalos de confiança de 95% (ICs) são mostrados como áreas sombreadas. A linha diagonal no gráfico marca uma AUC de 0,5 (ou seja, classificação aleatória). A) doença de Crohn, B) retocolite ulcerativa, C) psoríase em placas, D) artrite reumatoride, E) Lupus eritematoso sistêmico, F) diabetes tipo 1, G) diabetes tipo 2

APÊNDICE 9. CURVAS ROC - CAPACIDADE DE MODELOS INTEGRADOS COM AS VARIÁVEIS FENOTÍPICAS E DE MICROBIOTA

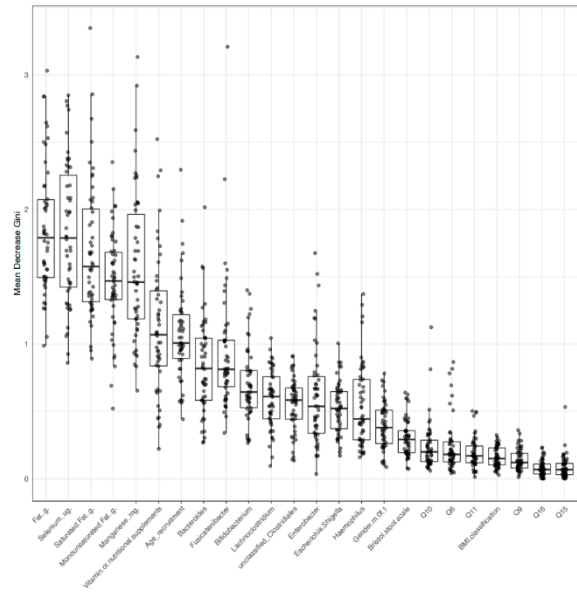
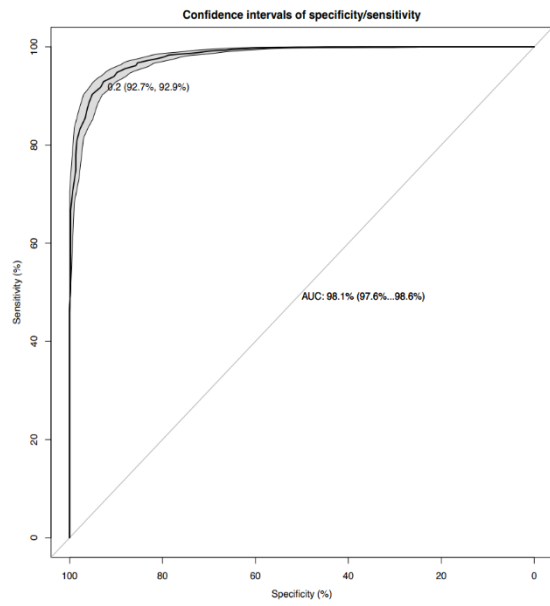
a. Doença de Crohn



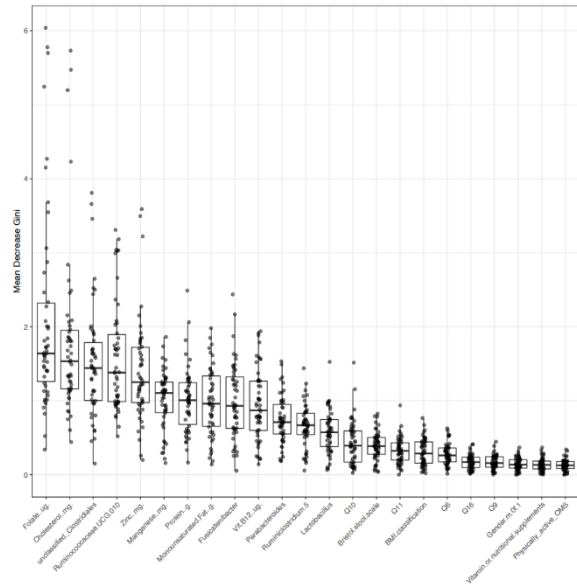
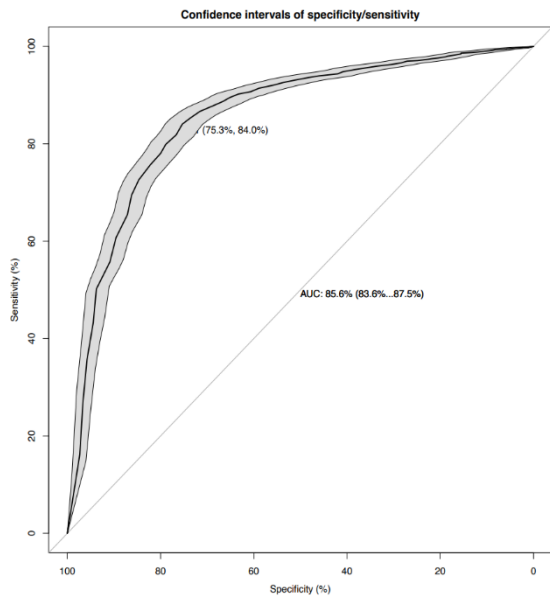
b. RCU



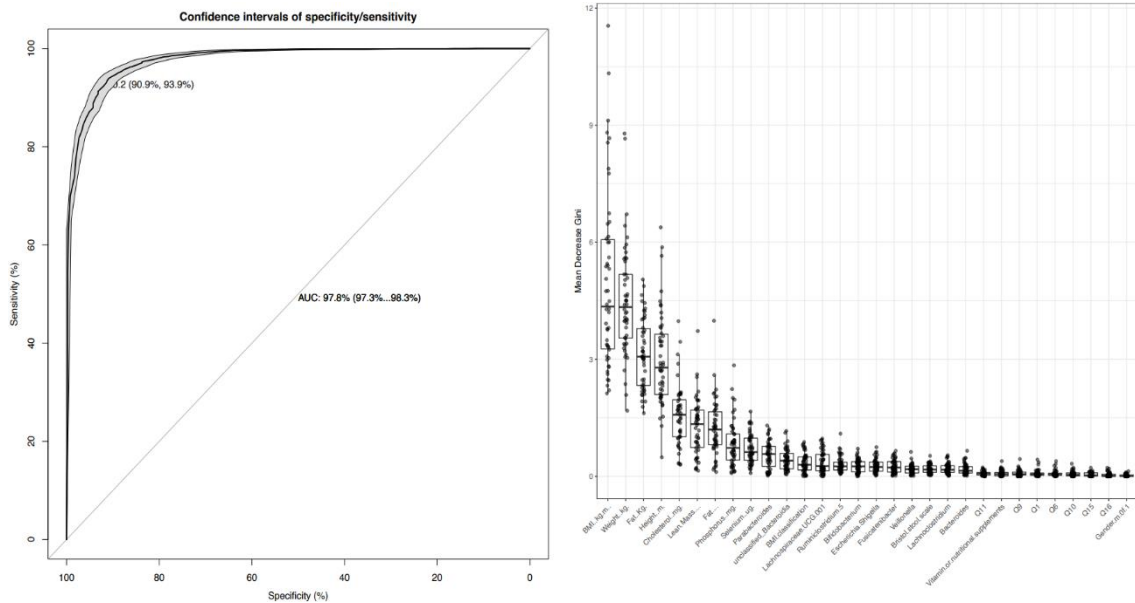
e. Lupus eritematoso sistêmico



f. Diabetes tipo 1



g. Diabetes tipo 2



Legenda: AUC, área sob a curva. Os intervalos de confiança de 95% (ICs) são mostrados como áreas sombreadas. A linha diagonal no gráfico marca uma AUC de 0,5 (ou seja, classificação aleatória). A) doença de crohn, B) retocolite ulcerativa, C) psoríase em placas, D) artrite reumatoride, E) Lupus eritematoso sistêmico, F) diabetes tipo 1, G) diabetes tipo 2

APÊNDICE 10. ANÁLISE PARALELA DE ASSOCIAÇÃO ENTRE VARIÁVEIS FENOTÍPICAS E ESTRUTURA DA MICROBIOTA INTESTINAL

| Variável | R2 | F | P | P-valor ajustado |
|-----------------------------|--------------|--------|-------|------------------|
| Grupo (CT ou enfermo) | 0,044 | 1,1811 | 0,001 | 0,046 |
| Metformina | 0,008 | 1,3458 | 0,013 | 0,207 |
| Suplemento vitamínico | 0,008 | 1,3259 | 0,017 | 0,207 |
| Corticoides (uso agudo) | 0,008 | 1,3358 | 0,018 | 0,207 |
| Antipsicóticos | 0,008 | 1,2896 | 0,034 | 0,272 |
| Nível de estresse | 0,021 | 1,1433 | 0,039 | 0,272 |
| Consumo de bebida alcoólica | 0,021 | 1,1431 | 0,042 | 0,272 |

Tamanhos de efeito (R2) calculados por testes de análise multivariada de variância multivariada permutacional (PERMANOVA) e correspondentes p ajustados-valores por Benjamini-Hochberg (BH).

APÊNDICE 11. CARACTERÍSTICAS DESCRITAS SOBRE O MICROBIOMA INTESTINAL ASSOCIADAS AO DM2

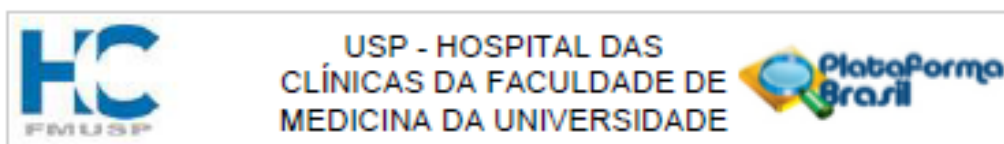
| População | Metodologia | Assinatura de risco | | Ref. |
|------------------|-----------------------|--|---|------|
| | | Aumentou | diminuiu | |
| 427 indivíduos | 16S rRNA região V4 | <i>Escherichia coli</i> , <i>Veillonella</i> , <i>Blautia</i> , <i>Anaerostipes</i> | NR | 206 |
| 800 indivíduos | 16S rRNA região V3–V4 | <i>Bacteroides thetaiotaomicron</i> , <i>Alistipes putredinis</i> | <i>Eubacterium rectale</i> , <i>Parabacteroides distasonis</i> , <i>Roseburia inulinivorans</i> , <i>Eubacterium eligens</i> , <i>Bacteroides vulgatus</i> | 207 |
| 60 indivíduos | 16S rRNA região V3–V4 | Bacteroidetes, Verrucomicrobia, Proteobacteria, Elusimicrobia, Acidobacteria, Deferribacteres, Gemmatimonadetes, Porphyromonadacea e, <i>Alistipes marseilloanorexica</i> , <i>Bacillus sporothermodurans</i> , <i>Staphylococcus</i> , <i>Prevotella</i> | Verrucomicrobia, Elusimicrobia, Arqueia metanogênica | 208 |
| 6.896 indivíduos | 16S rRNA região V4 | Actinobacteria, <i>Fusobacterium</i> , <i>Streptococcus</i> , <i>Lactobacillus</i> | <i>Akkermansia</i> , Synergistes, <i>Methanobrevibacter</i> , <i>Oscillospira</i> , <i>Roseburia</i> , <i>Bifidobacterium</i> | 209 |
| 60 indivíduos | 16S rRNA região V3–V4 | <i>Faecalibacterium</i> , Clostridiales, <i>Dorea</i> , Clostridiaceae, Lachnospiraceae | <i>Bifidobacterium</i> , <i>Parabacteroides</i> , <i>Oscillospira</i> , <i>Bacteroides</i> | 210 |
| 291 indivíduos | 16S rRNA região V4 | <i>Peptostreptococcus</i> , <i>Eubacterium</i> , <i>Prevotella</i> , <i>Desulfovibrio</i> | <i>Collinsella</i> , <i>Adlercreutzia</i> , <i>Anaerostipes</i> , <i>Epulopiscium</i> , <i>Clostridium butyricum</i> , <i>Ruminococcus</i> , <i>Pediococcus</i> | 211 |

| | | | | |
|---|--------------------------|---|---|-----|
| 1.976 pacientes na coorte KORA | 16S rRNA região V3–V4 | <i>E. coli</i> | <i>Faecalibacterium prausnitzii,</i> <i>Bifidobacterium longum,</i> <i>Intestinales bartettii,</i> <i>Coprococcus,</i> <i>Eubacterium rectale</i> | 130 |
| 436 pacientes da coorte Popgen e 844 pacientes da coorte FoCUS | 16S rRNA região V1–V2 | <i>B. thetaiotamicron</i> | <i>Clostridium sensu stricto,</i> <i>E. coli,</i> <i>Romboutsia,</i> <i>Barnesiella,</i> <i>Pseudoflavonifractor</i> <i>Veillonella,</i> <i>Roseburia</i> | 127 |
| 690 pacientes da coorte do CDC e do estudo da NYU, ^b | 16S rRNA região V4 | Gemellaceae, <i>Streptococcus,</i> <i>Blautia</i> | Parabacteroides, Clostridiaceae, Lachnospiraceae, Ruminococcaceae, Clostridiales, <i>Oscillospira</i> | 212 |

Legenda: ADA, Associação Americana de Diabetes; CDC, Centros de Controle e Prevenção de Doenças; NR, não informado; NYU, Universidade de Nova York; DM2, diabetes mellitus tipo 2;

ANEXOS

ANEXO A. APROVAÇÃO PROJETO PRINCIPAL



PARECER CONSUBSTANCIADO DO CEP

DADOS DO PROJETO DE PESQUISA

Título da Pesquisa: Estudo VALIDYS: Validação de uma ferramenta para rastreamento do risco de disbiose e caracterização da microbiota intestinal sadia e enferma.

Pesquisador: Dan Linetzky Waltzberg

Área Temática:

Versão: 1

CAAE: 01713018.0.0000.0068

Instituição Proponente: Hospital das Clínicas da Faculdade de Medicina da USP

Patrocinador Principal: Financiamento Próprio

DADOS DO PARECER

Número do Parecer: 3.008.966

Apresentação do Projeto:

Este é um projeto de grande porte e duração, voltado para o microbioma intestinal.

Objetivo da Pesquisa:

Visa analisar critérios de disbiose, com base em parâmetros clínicos e microbiológicos (a nível de filo e gênero)

Avaliação dos Riscos e Benefícios:

Os riscos não fogem do usual, uma vez que os questionários e a coleta de material fecal para sequenciamento genético já fazem parte da rotina. Poderá haver benefícios, padronizando e simplificando o diagnóstico e a conduta clínica em portadores de disbiose.

Comentários e Considerações sobre a Pesquisa:

O projeto está corretamente estruturado e dimensionado (1100 participantes). Deverá envolver custos substanciais, pelo que verbas serão solicitadas de agências de fomento.

Considerações sobre os Termos de apresentação obrigatória:

O TCLE é adequado

Recomendações:

Recomenda-se anexar a Anuência dos diversos serviços envolvidos (diabetes, enfermidades digestivas, alergia etc)

Endereço: Rua Ovídio Pires de Campos, 225 5ª andar
Bairro: Cerqueira César **CEP:** 05.403-010
UF: SP **Município:** SAO PAULO
Telefone: (11)2981-7585 **Fax:** (11)2981-7585 **E-mail:** cappelq.edm@hc.fm.usp.br



Continuação do Parecer: 3.006.906

Conclusões ou Pendências e Lista de Inadequações:

Não há pendências

Considerações Finais a critério do CEP:

Em conformidade com a Resolução CNS nº 466/12 – cabe ao pesquisador: a) desenvolver o projeto conforme delineado; b) elaborar e apresentar relatórios parciais e final; c) apresentar dados solicitados pelo CEP, a qualquer momento; d) manter em arquivo sob sua guarda, por 5 anos da pesquisa, contendo fichas individuais e todos os demais documentos recomendados pelo CEP; e) encaminhar os resultados para publicação, com os devidos créditos aos pesquisadores associados e ao pessoal técnico participante do projeto; f) justificar perante ao CEP interrupção do projeto ou a não publicação dos resultados.

Este parecer foi elaborado baseado nos documentos abaixo relacionados:

| Tipo Documento | Arquivo | Postagem | Autor | Situação |
|---|---|------------------------|---------------------------|----------|
| Informações Básicas do Projeto | PB INFORMACOES_BASICAS_DO_PROJETO_1239517.pdf | 23/10/2018 16:45:41 | | Aceito |
| Outros | avaliacao_antecipada.pdf | 23/10/2018 16:36:27 | Dan Linetzky Waltzberg | Aceito |
| Cronograma | Cronograma_valDYS_21102018.docx | 23/10/2018 16:13:19 | Dan Linetzky Waltzberg | Aceito |
| Projeto Detalhado / Brochura Investigador | Projeto_detalhado.pdf | 23/10/2018 16:12:44 | Dan Linetzky Waltzberg | Aceito |
| TCLE / Termos de Assentimento / Justificativa de Ausência | TCLE_21102018.pdf | 23/10/2018 15:58:27 | Dan Linetzky Waltzberg | Aceito |
| Folha de Rosto | folha_de_rosto_final231018.pdf | 23/10/2018 14:43:07 | Dan Linetzky Waltzberg | Aceito |

Situação do Parecer:

Aprovado

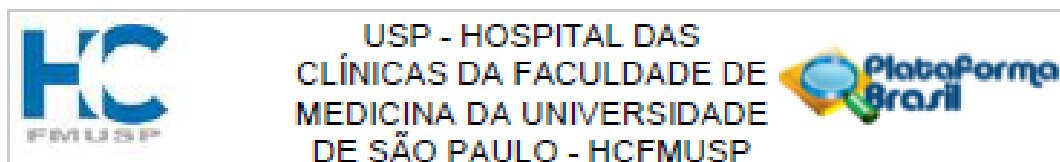
Necessita Apreciação da CONEP:

Não

Endereço: Rua Ovídio Pires de Campos, 225 5ª andar
Bairro: Cerqueira Cesar CEP: 05.403-010
UF: SP Município: SAO PAULO
Telefone: (11)2861-7585 Fax: (11)2861-7585 E-mail: cappelq.edm@hc.fm.usp.br

ANEXO B. DESMEMBRAMENTO

Liberação específica para o presente projeto



PARECER CONSUBSTANCIADO DO CEP

DADOS DA EMENDA

Título da Pesquisa: Estudo VALIDYS: Validação de uma ferramenta para rastreamento do risco de diabetes e caracterização da microbiota intestinal sadia e enferma.

Pesquisador: Dan Linetzky Waltzberg

Área Temática:

Versão: 7

CAAE: 01713018.0.0000.0068

Instituição Proponente: Hospital das Clínicas da Faculdade de Medicina da USP

Patrocinador Principal: Financiamento Próprio

DADOS DO PARECER

Número do Parecer: 4.019.807

Apresentação do Projeto:

Nova documentação foi enviada

Objetivo da Pesquisa:

Trata-se de Emenda 5, voltada para um Subprojeto de Doutorado "Identificação do Impacto de fatores ambientais ao perfil da microbiota intestinal em estados de saúde e de doença" a cargo de aluna Danielle Fonseca

Avaliação dos Riscos e Benefícios:

sem alterações

Comentários e Considerações sobre a Pesquisa:

O subprojeto encontra-se corretamente elaborado

Considerações sobre os Termos de apresentação obrigatória:

O TCLE foi atualizado, todavia falta autorização para Exame de Bioimpedância

Recomendações:

Vide item do TCLE

Conclusões ou Pendências e Lista de Inadequações:

Não há pendências, só uma recomendação

Endereço: Rua Ovídio Pires de Campos, 225 5º andar
Bairro: Cerqueira César CEP: 05403-010
UF: SP Município: SÃO PAULO
Telefone: (11)2661-7585 Fax: (11)2661-7585 E-mail: cappelq.adm@hc.fm.usp.br



USP - HOSPITAL DAS
CLÍNICAS DA FACULDADE DE
MEDICINA DA UNIVERSIDADE
DE SÃO PAULO - HCFMUSP



Continuação do Parecer: 4.019.007

Considerações Finais a critério do CEP:

Este parecer foi elaborado baseado nos documentos abaixo relacionados:

| Tipo Documento | Arquivo | Postagem | Autor | Situação |
|---|--|------------------------|---------------------------|----------|
| Informações Básicas do Projeto | PB_INFORMAÇÕES_BÁSICAS_1545189_E5.pdf | 26/04/2020 20:44:02 | | Aceito |
| Outros | FORMULARIO_PARA_SUBMISSAO_D E_EMENDAS_DanielleFonseca.pdf | 26/04/2020 20:42:11 | Dan Linetzky Waltzberg | Aceito |
| Cronograma | Cronograma_Doutorado_Danielle_Fonseca_sadlos_enfermos_240420.pdf | 26/04/2020 20:41:37 | Dan Linetzky Waltzberg | Aceito |
| Projeto Detalhado / Brochura Investigador | Projeto_Doutorado_Danielle_Fonseca_sadlos_enfermos_240420.pdf | 26/04/2020 20:41:28 | Dan Linetzky Waltzberg | Aceito |
| Outros | pacias.pdf | 26/02/2020 16:48:01 | Dan Linetzky Waltzberg | Aceito |
| Outros | Relatorio_parcial_estudovalidys.pdf | 12/12/2019 17:49:25 | Dan Linetzky Waltzberg | Aceito |
| Cronograma | Cronograma_validYS_21102018.docx | 23/10/2018 16:13:19 | Dan Linetzky Waltzberg | Aceito |
| Projeto Detalhado / Brochura Investigador | Projeto_detalhado.pdf | 23/10/2018 16:12:44 | Dan Linetzky Waltzberg | Aceito |
| TCLE / Termos de Assentimento / Justificativa de Ausência | TCLE_21102018.pdf | 23/10/2018 15:58:27 | Dan Linetzky Waltzberg | Aceito |
| Folha de Rosto | folha_de_rosto_final231018.pdf | 23/10/2018 14:43:07 | Dan Linetzky Waltzberg | Aceito |

Situação do Parecer:

Aprovado

Necessita Apreciação da CONEP:

Não

SÃO PAULO, 11 de Maio de 2020

Assinado por:
ALFREDO JOSÉ MANSUR
(Coordenador(a))

Endereço: Rua Ovídio Pires de Campos, 225 5ª andar
Bairro: Cerqueira César CEP: 05.403-010
UF: SP Município: SÃO PAULO
Telefone: (11)2661-7585 Fax: (11)2661-7585 E-mail: cappelq.adm@hc.fm.usp.br

ANEXO C. TCLE

Estudo VALIDY'S: Validação de uma ferramenta para rastreamento do risco de diabetes e caracterização da microbiota intestinal sadia e enferma (TCLE versão 1/ 15.10.2018)

TERMO DE CONSENTIMENTO LIVRE E ESCLARECIDO (TCLE)

IDENTIFICAÇÃO DO PARTICIPANTE DA PESQUISA OU RESPONSÁVEL LEGAL

1. NOME:.....
DOCUMENTO DE IDENTIDADE Nº : SEXO BIOLÓGICO: M = F =
DATA NASCIMENTO:/...../.....
ENDEREÇO Nº APTO:
BAIRRO: CIDADE
CEP:..... TELEFONE: DDD (.....)
E-MAIL:.....

2. RESPONSÁVEL LEGAL
NATUREZA (grau de parentesco, tutor, curador etc.)
DOCUMENTO DE IDENTIDADE : SEXO BIOLÓGICO: M = F =
DATA NASCIMENTO.:/...../..... IDADE:
ENDEREÇO: Nº APTO:
BAIRRO: CEP:
TELEFONE: DDD (.....).....
E-MAIL:.....

DADOS SOBRE A PESQUISA

1. TÍTULO DO PROTOCOLO DE PESQUISA: Estudo VALIDY'S: Validação de uma ferramenta para rastreamento do risco de diabetes e caracterização da microbiota intestinal sadia e enferma.

PESQUISADOR RESPONSÁVEL / PRINCIPAL: Dan Linebky Waltzberg

CARGO/FUNÇÃO: médico/ professor - CRM Nº: 22052

UNIDADE DO HCFMUSP: "Faculdade de Medicina da Universidade de São Paulo"

PESQUISADOR EXECUTANTE: Danielle C. Fonseca Candian – Nutricionista CRN3-37872

2. AVALIAÇÃO DO RISCO DA PESQUISA:

RISCO MÍNIMO (X)

RISCO MÉDIO =

RISCO BAIXO =

RISCO MAIOR =

3. DURAÇÃO DA PESQUISA : 4 anos

Prezado (a) Senhor(a):

Você está sendo convidado(a) para participar voluntariamente de uma pesquisa que prevê responder perguntas de questionários e, eventualmente, a coleta de amostra fecal e análise de composição corporal (análise de gordura e músculos corporais). Pesquisa é um conjunto de procedimentos que procura criar ou aumentar o conhecimento sobre um assunto. Embora frequentemente não tragam benefícios diretos ao participante, as descobertas de uma pesquisa podem ser úteis para muitas pessoas no futuro. Para decidir se aceita ou não participar desta pesquisa, o(a) senhor(a) precisa entender o suficiente sobre todos os seus aspectos. Este documento, chamado Termo de Consentimento Livre e Esclarecido (TCLE), descreve as razões da pesquisa, os procedimentos previstos, bem como os riscos e benefícios nela envolvidos. Pedimos que leia e, caso sinta necessidade, discuta com familiares e/ou outras pessoas de sua confiança todo o seu conteúdo. Se alguma dúvida persistir, você pode esclarecê-la diretamente com o pesquisador. Havendo interesse em participar, será solicitada a sua rubrica em todas as páginas do TCLE e sua assinatura na última página. Uma via assinada deste termo deverá ficar em posse do(a) senhor(a) ou por seu representante legal e uma cópia será arquivada pelo pesquisador responsável.

Rubrica do participante da pesquisa

Rubrica do pesquisador

Objetivo do estudo: Aplicar um questionário (DYS) sobre risco de disbiose intestinal (alterações no padrão de bactérias que vivem em nosso intestino), para verificar sua sensibilidade em detectar alterações da microbiota intestinal (bactérias que vivem em nosso intestino) em casos de saúde e doença (população geral, funcionários e pacientes do Hospital das Clínicas de São Paulo – HC-FMUSP).

Procedimentos: Sua participação poderá ser apenas com o preenchimento do questionário DYS e questionários de consumo alimentar, ou também contar com a coleta adicional de quatro pequenas amostras de fezes, avaliação da composição corporal e novamente um questionário de consumo alimentar. O referido questionário DYS, contém perguntas que podem trazer algum constrangimento, como idade; tipo de parto em que você nasceu; hábitos referentes à alimentação, prática de atividade física, consumo de álcool e/ou fumo; uso de medicação, presença de enfermidades e funcionamento intestinal. Caso o(a) senhor(a) preencha alguns critérios estabelecidos pelo estudo, os profissionais responsáveis pela pesquisa poderão entrar em contato com o(a) senhor(a) por e-mail, carta ou telefone, para o convidar a participar da etapa que necessita de amostra de fezes. As referidas amostras serão utilizadas para a realização de dois exames chamados "sequenciamento de microbiota das fezes", que avalia quais bactérias vivem em seu intestino através da análise de fezes e "metabolômica das fezes", que avalia o que essas bactérias podem produzir nas fezes. Caso seja solicitado, esse exame não terá nenhum custo para o(a) senhor(a), não trará nenhum risco à sua saúde, nem implicará em qualquer alteração do seu tratamento. Trata-se apenas da coleta de amostra de fezes, feita em sua casa.

Benefícios. Esta pesquisa trará maior conhecimento sobre o tema abordado, mas não há previsão de benefício direto para o(a) senhor(a). Estudos científicos baseiam-se na análise de grande número de amostras, sendo impossível estimar o benefício individual de cada paciente incluído no estudo.

Riscos. O preenchimento destes questionários, e/ou coleta de fezes e avaliação de composição corporal não representará qualquer risco de ordem física ou psicológica para você. Ainda assim, caso quera desistir de sua participação em qualquer momento, você não sofrerá nenhum prejuízo.

Forma de acompanhamento e assistência. A sua participação neste estudo não modifica em nada os tratamentos que o(a) senhor(a) realiza nesta instituição (aplica-se exclusivamente a pacientes do Hospital das Clínicas de São Paulo – HC-FMUSP). Em qualquer etapa do estudo, o(a) senhor(a) terá acesso aos profissionais responsáveis pela pesquisa. O investigador principal é o Dr. Dan Linetzky Waltzberg, que pode ser encontrado pelo telefone (11) 3061-7468, horário de atendimento 08-18 horas. O investigador executante é Danielle Cristina Fonseca Candian que pode ser encontrado na Faculdade de Medicina da USP, Avenida Doutor Arnaldo nº466, sala 2206/2208 – prédio principal, Telefone (11) 88286-8068, e-mail: daniellecf_25@hotmail.com.

Se você tiver alguma consideração ou dúvida sobre a ética da pesquisa, entre em contato com o Comitê de Ética em Pesquisa (CEP) – Rua Ovidio Pires de Campos, 225 – 5º andar – tel: (11) 2661-7585, (11) 2661-1548, (11) 2661-1549; e-mail: cep@cepa.admi@hc.fm.usp.br

Liberdade de recusar-se e retirar-se do estudo. A escolha de entrar ou não nesse estudo é inteiramente sua. Caso o(a) senhor(a) se recuse a participar deste estudo, o(a) senhor(a) receberá o tratamento habitual (caso seja paciente do Hospital das Clínicas de São Paulo – HC-FMUSP), sem qualquer tipo de prejuízo ou represália. O(A) senhor(a) também tem o direito de retirar-se deste estudo a qualquer momento e, se isso acontecer, seu médico continuará a tratá-lo(a) sem qualquer prejuízo ao tratamento ou represália (aplica-se a pacientes do Hospital das Clínicas de São Paulo – HC-FMUSP).

Manutenção do sigilo e privacidade. Os seus dados serão analisados em conjunto com outros pacientes, sendo assim, as informações fornecidas por você terão sua privacidade garantida pelos pesquisadores responsáveis, não sendo divulgada a identificação de nenhum paciente sob qualquer circunstância. Solicitamos sua autorização para que os dados obtidos nesta pesquisa sejam utilizados em publicações científicas, meio como os resultados de uma pesquisa são divulgados e compartilhados com a comunidade científica.

Garantia de Ressarcimento. O(A) senhor(a) não terá qualquer custo, pois o custo desta pesquisa será de responsabilidade do orçamento da pesquisa. O(A) senhor(a) tem direito a ressarcimento em caso de despesas decorrentes da sua participação na pesquisa.

Garantia de Indenização. Esta pesquisa não é invasiva e não traz riscos à saúde. O(A) senhor(a) tem direito à indenização diante de eventuais danos comprovadamente decorrentes dela.

Acredito ter sido suficientemente informado a respeito das informações que li ou que foram lidas para mim, descrevendo o estudo "Estudo VALIDYS: Validação de uma ferramenta para rastreamento do risco de disbiose e caracterização da microbiota intestinal sadia e enferma". Eu discuti com o Pesquisador Responsável Dr Dan Linetzky Waltzberg, ou pessoa(s) por ele delegada(s) _____ sobre a minha decisão em participar nesse estudo.

Ficaram claros para mim quais são os propósitos do estudo, os procedimentos a serem realizados, seus desconfortos e riscos, as garantias de confidencialidade e de esclarecimentos permanentes. Ficou claro também que minha participação é voluntária, isenta de despesas e participar deste estudo não modifica em nada meu acesso a tratamento hospitalar (aplica-se a pacientes do Hospital das Clínicas de São Paulo – HC-FMUSP), quando necessário.

Concordo voluntariamente em participar deste estudo e poderei retirar o meu consentimento a qualquer momento, antes ou durante o mesmo, sem penalidades, prejuízo ou perda de qualquer benefício que eu possa ter adquirido.

Assinatura do participante/representante legal Data ____/____/____

Declaro que obtive de forma apropriada e voluntária o Consentimento Livre e Esclarecido deste participante ou representante legal para a participação neste estudo.



Assinatura do responsável pelo estudo Data ____/____/____