

UNIVERSIDADE DO VALE DO PARAÍBA
INSTITUTO DE PESQUISA E DESENVOLVIMENTO
PROGRAMA DE PÓS GRADUAÇÃO EM ENGENHARIA BIOMÉDICA

ISABELLA SOUZA SILVA

**ESPECTROSCOPIA FTIR ASSOCIADA AO *MACHINE LEARNING* NA
CLASSIFICAÇÃO DOS MARCADORES DE PREDISPOSIÇÃO A DIABETES
MELLITUS TIPO 2**

**FTIR SPECTROSCOPY COMBINED WITH MACHINE LEARNING IN THE
CLASSIFICATION OF PREDISPOSITION MARKERS FOR TYPE 2 DIABETES
MELLITUS**

São José dos Campos, SP
2025

ISABELLA SOUZA SILVA

**ESPECTROSCOPIA FTIR ASSOCIADA AO *MACHINE LEARNING* NA
CLASSIFICAÇÃO DOS MARCADORES DE PREDISPOSIÇÃO A DIABETES
MELLITUS TIPO 2**

Dissertação de Mestrado apresentada ao Programa de Pós-graduação em Engenharia Biomédica, como complementação dos créditos necessários para obtenção do grau de Mestre em Engenharia Biomédica.

Orientador: Prof. Dra. Renata de Azevedo Canevari
Coorientador: Prof. Dr. Leandro José Raniero

São José dos Campos, SP
2025

TERMO DE AUTORIZAÇÃO DE DIVULGAÇÃO DA OBRA

Ficha catalográfica

Silva, Isabella Souza
Espectroscopia FTIR associada ao Machine Learning na
classificação dos marcadores de predisposição a DIABETES MELLITUS
tipo 2 / Isabella Souza Silva; orientadora, Renata de Azevedo
Canevari; co-orientadora Leandro José Raniero. - São José dos
Campos, SP, 2025.

103 p.

Dissertação (Mestrado Acadêmico) - Universidade do Vale do
Paraíba, São José dos Campos. Programa de Pós-Graduação em
Engenharia Biomédica.

Inclui referências

1. Engenharia Biomédica. 2. Diabetes tipo 2. 3.
Espectroscopia ATR-FTIR. 4. Machine Learning. 5. Polimorfismo
866G/A. I. Canevari, Renata de Azevedo, orient. II. Raniero,
Leandro José, co-orient. III. Universidade do Vale do Paraíba.
Programa de Pós-Graduação em Engenharia Biomédica. IV. Título.

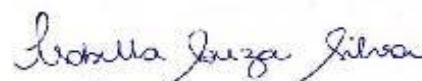
Eu, Isabella Souza Silva, autor(a) da obra acima referenciada:

Autorizo a divulgação total ou parcial da obra impressa, digital ou fixada em
outro tipo de mídia, bem como, a sua reprodução total ou parcial, devendo o
usuário da reprodução atribuir os créditos ao autor da obra, citando a fonte.

Declaro, para todos os fins e efeitos de direito, que o Trabalho foi elaborado
respeitando os princípios da moral e da ética e não violou qualquer direito de
propriedade intelectual sob pena de responder civil, criminal, ética e
profissionalmente por meus atos.

São José dos Campos, 14 de Abril de 2025.

Autor(a) da Obra


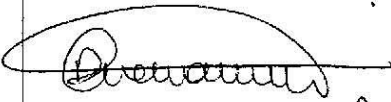
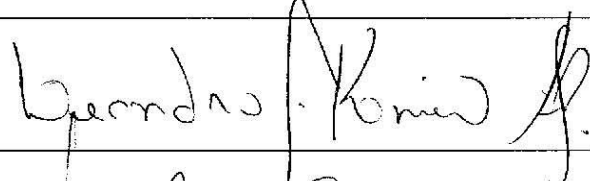
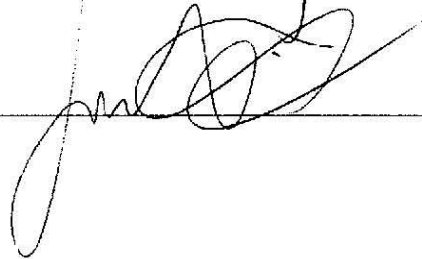


Data da defesa: 24/02/2025

ISABELLA SOUZA SILVA

**“ESPECTROSCOPIA FTIR ASSOCIADA AO MACHINE LEARNING NA CLASSIFICAÇÃO DOS
MARCADORES DE PREDISPOSIÇÃO A DIABETES MELLITUS TIPO 2.”**

Dissertação aprovada como requisito parcial à obtenção do grau de Mestre, do Programa de Pós-Graduação em Engenharia Biomédica, do Instituto de Pesquisa e Desenvolvimento da Universidade do Vale do Paraíba - Univap, pela seguinte banca examinadora:

Prof. Dr. Luis Eduardo Silva Soares	
Prof.ª Dr.ª Renata de Azevedo Canevari	
Prof. Dr. Leandro José Raniero	
Prof. Dr. Luís Felipe das Chagas e Silva Carvalho - Unitau	

Prof.ª Dr.ª Juliana Ferreira Strixino

Diretora do IP&D – Univap

São José dos Campos, 24 de fevereiro de 2025.

DECLARAÇÃO

Declaro que estou ciente de todas as informações contidas nesta dissertação de mestrado. O texto foi revisado por mim, Isabella Souza Silva e pela minha orientadora Prof^a Dr^a Renata de Azevedo Canevari e coorientador Prof. Dr. Leandro Raniero. O conteúdo está segundo as normas do Programa de Pós graduação em Engenharia Biomédica e o Português foi revisado adequadamente.

São José dos Campos, 24 de fevereiro de 2025.

Isabella Souza Silva

Profa. Dra. Renata de Azevedo Canevari

Prof. Dr. Leandro José Raniero

RESUMO

A glicose é um carboidrato essencial para o metabolismo energético celular, cuja homeostase é regulada por insulina e glucagon. Alterações nesse equilíbrio podem levar ao desenvolvimento de diabetes mellitus do tipo 2 (DM2), cuja identificação precoce é fundamental para mitigar complicações. O objetivo deste estudo é avaliar a viabilidade da técnica de espectroscopia ATR-FTIR aliada ao algoritmo *machine learning* como um método alternativo na detecção do SNP -866G/A do gene *UCP2* e na análise glicêmica de soro humano para identificação de predisposição a DM2. Foram coletados e analisados dados clínicos e bioquímicos de 59 pacientes, incluindo glicemia, hemoglobina glicada, colesterol total e triglicérides. A técnica de Espectroscopia no Infravermelho por Transformada de Fourier com Reflexão Total Atenuada (ATR-FTIR) foi utilizada para a obtenção dos espectros de 59 amostras de soro e 52 amostras de DNA amplificado pela PCR (reação em cadeia da polimerase). A genotipagem foi realizada pela técnica de PCR-RFLP (Polimorfismo de comprimento de fragmentos de restrição em 52 amostras de DNA obtido do sangue total dos pacientes). A análise de cluster hierárquico e redes neurais artificiais foram utilizadas para a separação dos grupos em relação às concentrações séricas de glicose. As técnicas de análise de componentes principais (PCA) e *machine learning* (KNN e SVM) foram utilizadas para a classificação dos grupos em relação ao tipo de SNP (AA, GA ou GG) do gene *UCP2*. A análise da genotipagem revelou uma distribuição equilibrada entre as variáveis bioquímicas e os diferentes genótipos, apresentando diferença significativa apenas nos níveis de insulina, que foram mais elevados no genótipo GA (19,75 $\mu\text{U/mL}$) em comparação aos genótipos GG (11,00 $\mu\text{U/mL}$) e AA (14,48 $\mu\text{U/mL}$). A análise de cluster hierárquico demonstrou eficácia na separação das amostras de soro em grupos com concentrações de glicose acima e abaixo de 100 mg/dL, atingindo uma acurácia de 78%. Modelo redes neurais artificiais (ANN) aplicados a essas amostras obtiveram 83% de precisão, com recall de 91% para a classe 1. A análise de *machine learning* como o KNN-Cosine, alcançou acurácia de 74% na classificação entre os genótipos AA e GG, e de 73,5% entre GA e GG, na faixa espectral de 1170-1800 cm^{-1} . A análise de ATR-FTIR combinada com o algoritmo *machine learning* foi demonstrada como uma abordagem promissora, menos invasiva e mais acessível que os métodos tradicionais. Essa técnica demonstrou capacidade de distinguir os grupos baseados nos diferentes níveis de glicose e genótipos do SNP -866G/A.

Palavras-chave: Diabetes tipo II; predição; espectroscopia ATR-FTIR; *machine learning*; gene *UCP2*, polimorfismo -866G/A.

ABSTRACT

Glucose is an essential carbohydrate for cellular energy metabolism, whose homeostasis is regulated by insulin and glucagon. Alterations in this balance can lead to the development of type 2 diabetes mellitus (T2DM), and early identification is crucial to mitigating complications. This study aims to evaluate the feasibility of the ATR-FTIR spectroscopy technique combined with a machine learning algorithm as an alternative method for detecting the -866G/A SNP of the UCP2 gene and analyzing human serum glucose levels to identify predisposition to T2DM. Clinical and biochemical data from 59 patients were collected and analyzed, including glucose, glycated hemoglobin, total cholesterol, and triglycerides. The Attenuated Total Reflectance Fourier Transform Infrared Spectroscopy (ATR-FTIR) technique was used to obtain spectra from 59 serum samples and 52 DNA samples amplified by PCR (polymerase chain reaction). Genotyping was performed using the PCR-RFLP (Restriction Fragment Length Polymorphism) technique on 52 DNA samples obtained from whole blood. Hierarchical cluster analysis and artificial neural networks were used to separate groups based on serum glucose concentrations. Principal component analysis (PCA) and machine learning techniques (KNN and SVM) were applied to classify groups according to the SNP type (AA, GA, or GG) of the UCP2 gene. Genotyping analysis revealed a balanced distribution between biochemical variables and different genotypes, with a significant difference observed only in insulin levels, which were higher in the GA genotype (19.75 $\mu\text{U/mL}$) compared to the GG (11.00 $\mu\text{U/mL}$) and AA (14.48 $\mu\text{U/mL}$) genotypes. Hierarchical cluster analysis demonstrated effectiveness in separating serum samples into groups with glucose concentrations above and below 100 mg/dL, achieving 78% accuracy. Artificial neural network (ANN) models applied to these samples achieved 83% precision, with a recall of 91% for class 1. Machine learning analysis using KNN-Cosine achieved 74% accuracy in classifying AA and GG genotypes and 73.5% accuracy for GA and GG genotypes in the spectral range of 1170-1800 cm^{-1} . The ATR-FTIR analysis combined with machine learning algorithms proved to be a promising, less invasive, and more accessible approach than traditional methods. This technique demonstrated the ability to distinguish groups based on different glucose levels and the -866G/A SNP genotypes.

Keywords: Type II Diabetes; prediction; ATR-FTIR spectroscopy; machine learning; UCP2 gene; -866G/A polymorphism.

LISTA DE ILUSTRAÇÕES

- Figura 1 - Amostras de DNA visualizadas pelas bandas de DNA genômico total em eletroforese em gel de agarose a 1% corado com brometo de etídio.....49
- Figura 2 - Eletroforese em gel de agarose (1,6%) com marcador de peso molecular 100pb (Ladder) seguido da amplificação de sequência de DNA que contém o SNP -866G/A do gene UCP2 das amostras de DNA obtida do sangue total.50
- Figura 3 - Gel de agarose com ladder de padrão de peso molecular de 100pb e três amostras digeridas de PCR-RFLP sendo a primeira GA a segunda GG e a terceira GA do polimorfismo -866G/A do gene UCP2.51
- Figura 4 - Dendrograma resultante da análise de cluster utilizando o algoritmo de Ward, aplicado aos dados espectrais do soro sanguíneo na faixa de 1139-1024 cm^{-1} , após normalização vetorial.53
- Figura 5 - Evolução da perda e da acurácia durante o treinamento e validação ao longo das épocas.....55
- Figura 6 - Espectros médios de absorvância FTIR dos genótipos AA, GA e GG.....56
- Figura 7 - Desvio médios dos espectros de absorvância FTIR dos genótipos AA, GA e GG....57
- Figura 8 - Gráficos de dispersão com combinações diferentes entre os componentes principais (PC1, PC2, PC3, PC4) sendo preto: grupo AA, azul: grupo GA e vermelho: grupo GG.....58
- Figura 9 - Gráficos de dispersão com combinações diferentes entre os componentes principais (PC1, PC2, PC3, PC4) sendo preto: grupo AA, azul: grupo GA.....59
- Figura 10 - Precisão (%) dos diferentes métodos de aprendizado de máquina aplicados nas faixas espectrais 1170-1800 cm^{-1} , 2800-3000 cm^{-1} , 400-4000 cm^{-1} e 800-1800 cm^{-1} para a classificação dos genótipos GA, AA e GG.60

LISTA DE TABELAS

Tabela 1 - Critérios laboratoriais para diagnóstico de DM.	18
Tabela 2 - Parâmetros laboratoriais utilizados para dosagem de glicose, hemoglobina glicada, insulina, colesterol total e triglicerídeos.	40
Tabela 3 - Média das concentrações bioquímicas de glicose, hemoglobina glicada, insulina, colesterol total e triglicerídeos de acordo com o genótipo do SNP -866G/A do gene UCP2.....	52
Tabela 4 - Desempenho do modelo de rede neural artificial na classificação da concentração de glicose.	54

SUMÁRIO

1	INTRODUÇÃO	12
2	OBJETIVOS	13
2.1	OBJETIVO GERAL	14
2.2	OBJETIVOS ESPECÍFICOS.....	14
3	REVISÃO DE LITERATURA	15
3.1	DIABETES	15
3.2	DIAGNÓSTICO DA DM2	17
3.3	FATORES DE RISCO.....	18
3.4	POLIMORFISMOS GENÉTICOS	22
3.5	GENE UCP2	25
3.6	ATR-FTIR APLICADA A ÁREAS MÉDICAS	29
3.7	ATR-FTIR APLICADA AO DM2	34
4	MATERIAIS E MÉTODOS	37
4.1	DELINEAMENTO DO ESTUDO E COLETA DAS AMOSTRAS.....	38
4.2	ANÁLISE BIOQUÍMICA	39
4.3	GENOTIPAGEM.....	40
4.3.1	Extração do DNA	40
4.3.2	Análise da integridade e qualidade do DNA	41
4.3.3	PCR do gene UCP2	42
4.3.4	Análise de PCR-RFLP	42
4.4	ANÁLISE DA GENOTIPAGEM VS BIOQUÍMICA	43
4.5	ANÁLISE DE ATR-FTIR	44
4.5.3	Análise de cluster hierárquico para separação dos grupos em relação aos níveis de glicose	45
4.5.4	ATR-FTIR machine learning para separação dos grupos em relação aos níveis de glicose	45
4.5.5	ATR-FTIR machine learning do polimorfismo -866G/A do gene UCP2	47
5	RESULTADOS	47
5.1	CARACTERÍSTICAS CLÍNICAS.....	48
5.4	ANÁLISE DE CLUSTER HIERÁRQUICO PARA CONCENTRAÇÃO DE GLICOSE	53
5.5	ATR-FTIR ALIADA AO MACHINE LEARNING NA CONCENTRAÇÃO DE GLICOSE.....	54
5.6	ATR-FTIR ALIADA AO MACHINE LEARNING DO POLIMORFISMO -866G/A DO GENE UCP2	55

6	DISCUSSÃO	61
7	CONCLUSÃO	70
	REFERÊNCIAS	71
	APÊNDICE A: Dados bioquímicos referente a níveis de glicose, HbA1c, insulina, colesterol total e triglicerídeos obtidos de amostras de sangue total coletados de 67 pacientes	83
	APÊNDICE B: Análises realizadas em amostras obtidas de sangue periférico de 67 participantes: análise molecular pela PCR-RFLP no DNA extraído da papa leucocitária; análises bioquímicas no plasma; análise de espectroscopia FTIR nas amostras de soro e as análises de FTIR aliada ao algoritmo <i>machine learning</i> realizada nas amostras de soro e amostras de DNA amplificado pela PCR.....	85
	APÊNDICE C: Informações pessoais e histórico familiar dos 67 participantes pertencentes ao estudo	87
	APÊNDICE D: Análise quantitativa amostras de DNA e respectivos genotipos obtidos de 52 participantes e resultado da PCR-RFLP das amostras analisadas por ATR-FTIT.....	96
	ANEXO A: Parecer do Comitê de Ética em Pesquisa	98
	ANEXO B: Termo de Consentimento Livre e Esclarecido	100
	ANEXO C: Questionário	102

1 INTRODUÇÃO

A glicose é um carboidrato essencial para o corpo, pois constitui uma das principais fontes de energia para as células, sendo obtida principalmente pela digestão de carboidratos. A manutenção dos níveis de glicose no sangue é realizada pela ação de dois hormônios, ambos produzidos pelo pâncreas: a insulina e o glucagon (Liemburg-Apers *et al.*, 2011; Kuo *et al.*, 2015). A insulina promove a absorção celular da glicose, enquanto o glucagon atua como um agente hiperglicemiante ao ocasionar a quebra do glicogênio hepático e liberação de glicose na corrente sanguínea (Brown, 2000).

Anormalidades na homeostase da glicose, especialmente níveis elevados, podem favorecer o desenvolvimento da Diabetes Mellitus (DM), especialmente o tipo 2 (DM2), resistência à insulina, pré-diabetes e síndrome metabólica. Essas doenças, quando diagnosticadas precocemente, podem ser tratadas, o que aumenta a sobrevivência do indivíduo, melhora sua qualidade de vida e promove, em alguns casos, como na condição da pré diabetes, a atenuação parcial ou completa dos sintomas, caracterizando a remissão da doença (Lee; Park; Choi, 2022).

As manifestações clínicas da DM2, de acordo com a OMS incluem sede excessiva, polaciúria e cansaço, além de sintomas sutis que dificultam o diagnóstico precoce. Conforme diretrizes estabelecidas pela Sociedade Brasileira de Diabetes (2024), o diagnóstico da doença deve ser realizado por meio de exames de rotina, como glicemia em jejum, hemoglobina glicada e teste de tolerância oral a glicose. Entretanto, esses testes possuem limitações, como desconforto gástrico, com possível sensação de desmaio, ânsia de vômito ou até mesmo vômito, e custos elevados (Sociedade Brasileira de Diabetes, 2024).

Fatores ambientais e comportamentais, tais como sedentarismo e alimentação inadequada, são relevantes para a patogênese da DM2 (Bandeira *et al.*, 2013), aumentando o risco de desenvolvimento desta doença (Shamoon, 1992). A atividade física regular melhora a sensibilidade à insulina, enquanto uma alimentação equilibrada ajuda a prevenir a doença (Bird; Hawley, 2016). No entanto, é importante ressaltar que a predisposição genética pode tornar o indivíduo ainda mais propenso ao desenvolvimento dessa condição (Morcillo *et al.*, 2008; Sanchez *et al.*, 2019).

Dentre as alterações genéticas que estão associadas ao desenvolvimento da DM2, destaca-se o polimorfismo de nucleotídeo único (SNP) -866G/A localizado no gene *UCP2*, cuja variante do alelo A está associada a níveis mais baixos de insulina e maior risco dessa doença, (Morcillo *et al.*, 2008; Sanchez *et al.*, 2019; Gomathi *et al.*, 2019; Hou *et al.*, 2020).

A análise do SNP -866G/A do *gene UCP2* pode detectar uma alteração pontual no DNA dos indivíduos e prever a propensão ao desenvolvimento da DM2 antes mesmo que alterações bioquímicas nas concentrações de glicose no sangue possam ser diagnosticadas. Dessa forma, esse SNP constitui um potencial biomarcador molecular para predisposição da DM2 (Sasahara *et al.*, 2004; Hou *et al.*, 2020; Lapik; Ranjit; Galchenko, 2021).

Recentemente, a análise de Espectroscopia no Infravermelho por Transformada de Fourier (FTIR) associada a técnicas de aprendizado de máquina (do inglês: *machine learning*) tem mostrado resultados promissores na classificação de genótipos do tipo SNPs e identificação de predisposição doenças, tais como a DM2, oferecendo uma alternativa econômica e eficiente para diagnóstico e acompanhamento de diversas condições de saúde (Gomes Rios *et al.*, 2021; Caixeta *et al.*, 2023).

FTIR é uma técnica inovadora, acessível e rápida para a caracterização bioquímica de amostras biológicas. As suas aplicações requerem pequenas quantidades de amostra, permitindo diagnósticos em tempo real e a identificação de alterações na composição química amostral antes que se tornem visíveis (Kahn *et al.*, 2009). Além disso, os métodos como a Reflectância Total Atenuada (ATR) associados ao FTIR tem conseguido facilitar a análise de amostras líquidas e sólidas (Kazarian; Chan, 2006).

Desse modo, a espectroscopia ATR-FTIR é uma excelente ferramenta que pode ser utilizada na classificação dos indivíduos de acordo com as concentrações glicêmicas e, quando associada ao algoritmo *machine learning*, na classificação genotípica dos indivíduos de acordo com o seus tipos de SNP, tais como o SNP -866G/A do *gene UCP2*. As vantagens de ser uma técnica mais acessível e rápida em comparação com a análise molecular, aliada ao seu potencial de identificar a predisposição DM2, pode incentivar mudanças no comportamento do indivíduo, no sentido de levá-lo a adquirir hábitos mais saudáveis e reduzir os riscos para desenvolvimento dessa patologia.

2 OBJETIVOS

2.1 OBJETIVO GERAL

Avaliar a viabilidade da técnica de espectroscopia ATR-FTIR aliada ao algoritmo *machine learning* como um método alternativo na detecção do SNP -866G/A do *gene UCP2* e na análise glicêmica de soro humano para identificação de predisposição a Diabetes Mellitus tipo 2.

2.2 OBJETIVOS ESPECÍFICOS

- avaliar, por meio da técnica de PCR-RFLP, o SNP rs659366 (-866G/A) localizado no *gene UCP2*;
- comparar os dados bioquímicos de insulina, glicemia, hemoglobina glicada, colesterol total e triglicerídeos dos indivíduos com os resultados da análise genotípica do SNP rs659366 (-866G/A);
- verificar a eficiência da técnica de ATR-FTIR na diferenciação de indivíduos de acordo com a concentração sérica de glicose;
- avaliar a eficiência da técnica de ATR-FTIR associada a algoritmos *machine learning* na diferenciação de indivíduos de acordo com a concentração sérica de glicose e na diferenciação de indivíduos de acordo com o tipo de SNP rs659366 (-866G/A).

3 REVISÃO DE LITERATURA

3.1 DIABETES

A glicose é uma das moléculas mais importantes para o funcionamento celular por ser o principal produto para formação de energia em forma de adenosina trifosfato (ATP) (Liemburg-Apers *et al.*, 2011).

A glicose é um monossacarídeo obtido por meio da digestão de carboidratos, compostos ricos em carbono e hidrogênio, que quando consumidos passam pela ação das enzimas amilase salivar, presente na boca, e da amilase pancreática, presente no intestino delgado; sendo absorvida no jejuno, porção intermediária do intestino delgado, onde há maior concentração de transportadores de glicose-sódio dependentes (SGLT-1) e transportadores de glicose (GLUT) (Ghezzi *et al.*, 2018) Após a absorção intestinal, a glicose passa pelo fígado, órgão responsável por armazená-la em forma de glicogênio ou distribuí-la para outros tecidos do organismo que a utilizam como fonte de energia. A homeostasia das moléculas glicêmicas é controlada por dois hormônios, insulina e glucagon, ambos secretados no pâncreas e sintetizados pelas células das ilhotas de Langerhans (Kuo *et al.*, 2015).

O aumento da glicose no sangue leva ao estímulo de secreção de insulina nas células beta das ilhotas de Langerhans no pâncreas, o que aumenta a captação desse carboidrato pela família de transportadores GLUT para o interstício das células do músculo, fígado, tecido adiposo, cérebro, e demais órgãos. O fígado, músculo e o tecido adiposo são os que mais sofrem a ação da insulina. No fígado, há estoque de glicogênio para necessidades futuras, no músculo também há estoque de glicogênio, porém sua função é garantir energia intramuscular, já o tecido adiposo recebe o excesso de glicose circulante para síntese de triacilglicerol (Petersen; Shulman, 2018).

Assim que a concentração de glicose se normaliza na corrente sanguínea, as células betas pancreáticas passam a reduzir a liberação de insulina, isso porque a concentração de glicose se encontra menor que no estágio inicial de sua obtenção por meio da digestão de carboidratos. A concentração de ATP dentro desta célula passa a ser menor, sinalizando assim o silenciamento das células beta pancreáticas (Tokarz; McDonald; Klip, 2018). Quando as concentrações glicêmicas ficam reduzidas, uma nova célula passa a produzir outro hormônio, responsável pelo controle positivo da glicose circulante, o glucagon. Esse hormônio é liberado pelas células alfa pancreáticas e sua síntese começa a partir do momento em que a concentração de ATP é reduzida, fazendo com que o glicogênio estocado passe a circular para os demais

tecidos. Além disso, outros mecanismos fisiológicos são acionados e o tecido adiposo passa a realizar quebra de gordura para formação de novas moléculas de glicose, processo esse denominado gliconeogênese (Zeigerer *et al.*, 2021).

O desequilíbrio metabólico ocorre quando o corpo não consegue regular adequadamente os níveis de glicose no sangue, que pode ser causado pela resistência a insulina, onde a capacidade das células em captar a glicose estão comprometidas em resposta ao hormônio insulina, levando à hiperglicemia e à disfunção metabólica em diversos tecidos. Essa condição complexa envolve múltiplos mecanismos patogênicos, incluindo alterações na sinalização da insulina, inflamação crônica de baixo grau e disfunção mitocondrial, e está intimamente associada ao desenvolvimento de diabetes mellitus tipo 2 (Lee; Park; Choi, 2022).

A diabetes mellitus (DM) é uma condição caracterizada pela ausência da produção de insulina de forma fisiológica pelas células beta pancreáticas ou pela incapacidade do corpo em usar a insulina de forma eficiente na captação de glicose na corrente sanguínea (American Diabetes Association, 2014). De acordo com Federação Internacional de Diabetes (IDF), a incidência mundial da doença em 2021 foi de 537 milhões de adultos diagnosticados, sendo que 3 em cada 4 indivíduos diabéticos vivem em países de baixa renda, levando a 4 milhões de mortes por ano (International Diabetes Federation, 2021).

No Brasil, aproximadamente 10,2% da população é afetada pelo diabetes mellitus, com uma maior prevalência entre as mulheres (11,1%) em relação aos homens (9,1%). Destacam-se os estados de: São Paulo e o Distrito Federal, onde 12,1% da população está diagnosticada com a doença, seguidos por Porto Alegre com 12% e Natal com 11,8% (Brasil, 2023). Os tipos mais comuns de diabetes incluem diabetes mellitus do tipo 1 (DM1), pré diabetes, diabetes gestacional e diabetes mellitus do tipo 2 (DM2) (Muzy *et al.*, 2021).

De acordo com a Sociedade Brasileira de Diabetes (SBD), a DM2 é caracterizada pela resistência à insulina e pela secreção insuficiente de insulina pelas células beta pancreáticas, além das disfunções na secreção de incretinas (Ahlqvist *et al.*, 2018). De acordo com o Ministério da Saúde, cerca de 90% das pessoas com diabetes têm o tipo 2 da doença e a manifestação clínica geralmente aparece na fase adulta (Peer; Balakrishna; Duraio, 2020).

O controle das taxas de glicose são essenciais para o metabolismo normal do organismo, onde o descontrole dos níveis glicêmicos leva a inúmeras disfunções celulares, como: doenças cardiovasculares, doenças renais, doenças oculares, neuropatia sistema imune comprometido, má circulação, e problemas de cicatrização (Sociedade Brasileira de Diabetes, 2024).

De acordo com a Organização Mundial da Saúde, os sintomas da DM2 são: sede excessiva e boca seca, micção frequente, falta de energia para atividades simples, feridas com

cicatrização lenta, infecções recorrentes na pele, visão turva e formigamento ou dormência nas mãos e nos pés. Esses sintomas podem também ser silenciosos, causando um atraso no diagnóstico da doença e até mesmo complicações que podem ser desencadeadas pela hiperglicemia frequente. As complicações causadas por uma DM2 descompensada, onde o indivíduo se mantém com as concentrações de glicose acima do esperado, consistem em perda de visão, doença renal diabética, riscos cardiovasculares, hipertensão arterial, insuficiência cardíaca, neuropatia periférica, pé diabético, úlceras, doença hepática gordurosa, hipogonadismo e doenças arteriais periféricas (Sociedade Brasileira de Diabetes, 2024).

3.2 DIAGNÓSTICO DA DM2

O diagnóstico da DM2 é concretizado pela identificação de hiperglicemia, e para isso podem ser usados glicemia plasmática em jejum (GJ), hemoglobina glicada (HbA1C) e teste de tolerância a glicose por via oral (TTGO) (Harreiter; Roden, 2023). Para a análise da glicemia plasmática em jejum (GJ) é necessário que o paciente se apresente para a coleta com pelo menos 8 hora de jejum, assim é possível analisar níveis séricos de glicose na corrente sanguínea do indivíduo (Filho et al., 2002). Para a avaliação da HbA1c, o jejum não é necessário, pois a concentração medida corresponde a uma proteína presente nos glóbulos vermelhos que interage com a glicose. Como a vida média dessas células é de aproximadamente 120 dias, o exame reflete o controle glicêmico ao longo desse período (Kahn *et al.*, 2013). Contudo, para a análise do TTGO o indivíduo passa pela coleta de sangue inicialmente em jejum, como na análise de GJ, e logo em seguida é administrado 75g de glicose via oral. Uma hora após é realizada uma nova coleta de sangue (TTGO-1) ou duas horas depois da sobrecarga glicêmica (TTGO-2) (Bergman *et al.*, 2024).

Os resultados dos exames laboratoriais são avaliados de acordo com a concentração do marcador solicitado. Na tabela 1 estão ilustrados os critérios laboratoriais para o diagnóstico da DM de acordo com a Diretriz da Sociedade Brasileira de Diabetes.

Tabela 1 - Critérios laboratoriais para diagnóstico de DM.

Critério	Normal	Pré-diabetes	DM
GJ (mg/dL)	<100	100-125	≥126
TTGO-1 (mg/dL)	<155	155-208	≥209
TTGO-2 (mg/dL)	<140	140-199	≥200
HbA1C (%)	<5,7	5,7-6,4	≥6,5

Legenda: GJ: glicose em jejum; TTGO: teste de tolerância à glicose oral; HbA1c: hemoglobina glicada.

Fonte: Sociedade Brasileira de Diabetes (2024).

O diagnóstico preciso da DM2 é obtido pelo conjunto de marcadores alterados, não bastando apenas um sendo estabelecido quando a taxa de GJ for maior que 126 mg/dL, juntamente com os resultados de HbA1c superiores a 6,5%. Caso o exame de TTGO for solicitado, valores de TTGO-1 acima de 209 mg/dL diagnosticam DM2 e são importantes também para controle de pré-diabetes. Já TTGO-2 é um teste pouco utilizado, já que o anterior demonstra boa sensibilidade (Selvin *et al.*, 2007). Apesar dos métodos descritos acima apresentarem uma boa acurácia no diagnóstico de DM2, eles possuem algumas limitações que devem ser consideradas, tais como a necessidade de jejum, custo alto e baixa precisão de alguns exames, como a HbA1C e TTGO. O jejum é um fator que gera desconforto para a maioria dos pacientes, além de poder causar hipoglicemia e desmaios quando realizado por períodos prolongados. Cabe ressaltar, que os resultados podem sofrer interferência quando o jejum não é realizado conforme solicitado, resultando em perda total da amostra e a necessidade de uma nova coleta.

Mesmo que a HbA1C seja um parâmetro que possibilita a avaliação a longo prazo das concentrações glicêmicas, a análise da hemoglobina glicada é um exame de custo mais elevado em relação aos outros exames bioquímicos, não leva em consideração a variabilidade do produto de glicação proteica, além de ser menos sensível para o diagnóstico. O TTGO é um teste mais oneroso e desconfortável pela ingestão de 75g de glicose em solução, podendo causar enjoos, vômitos e desconforto gástrico, além de tornar necessário que o paciente fique no local da coleta por um período de tempo (Sociedade Brasileira de Diabetes, 2024).

3.3 FATORES DE RISCO

Os fatores ambientais e comportamentais exercem um papel crucial no desenvolvimento do DM2. Entre esses fatores estão a alimentação rica em carboidratos, alimentos industrializados, obesidade (IMC elevado) e sedentarismo. Além disso, outros fatores de risco

também podem aumentar a chance de desenvolvimento da doença tais como, stress, sono inadequado e predisposição genética (Bandeira *et al.*, 2013), onde indivíduos que já apresentam predisposição genética podem desencadear a DM2 com mais facilidade ou até mesmo agravá-la se tiverem um estilo de vida ruim (Moore; Florez, 2008).

O consumo excessivo de carboidratos simples, como monossacarídeos e dissacarídeos que possuem baixo peso molecular, constitui risco para o desenvolvimento da DM2 em relação a indivíduos que possuem uma dieta rica em fibras (Reynolds *et al.*, 2019). De acordo com o Guia Alimentar da População Brasileira, publicado em 2014, a alimentação saudável respeita a região e os costumes de uma população, porém traz a conscientização da importância do consumo de frutas, verduras e legumes diariamente e a baixa frequência no consumo de alimentos industrializados (alimentos processados e ultraprocessados) para garantir a saúde e a longevidade.

Um estudo demonstrou que o consumo médio de calorias da população brasileira entre 2008 a 2018 era de 1754kcal, sendo 11,28% alimentos processados e 19,69% alimentos ultraprocessados, justamente a classe de alimentos que demonstra um risco maior para o desenvolvimento de DM2 (Louzada *et al.*, 2023). Durante a pandemia de Covid-19 houve aumento de consumo de alimentos ultraprocessados e bebidas açucaradas no Brasil, fatores que associados à inatividade física podem desencadear aumento no índice glicêmico e consequentemente levar a condição de pré-diabetes ou até mesmo a DM2 (Andrade *et al.*, 2023). De acordo com o Atlas do Diabetes, o número de pacientes diagnosticados com diabetes no mundo aumentou de 463 milhões em 2019 para 537 milhões em 2021, sendo o sedentarismo, má alimentação uns dos principais fatores (International Diabetes Federation, 2021)

Outro trabalho realizado com 492 pessoas pertencentes a comunidade indígena do Canadá demonstrou que os indivíduos que consumiam com alta frequência carne bovina e alimentos processados apresentavam risco aumentado do desenvolvimento do DM2 em relação aos indivíduos que possuíam uma alimentação balanceada (Reeds *et al.*, 2016). Sob a mesma perspectiva, Ley *et al.* (2016) relatou que a qualidade das gorduras e dos carboidratos consumidos são mais importantes que as quantidades destes no desenvolvimento do DM2, mostrando que manter uma alimentação focando em carboidratos complexos, fibras e gorduras insaturadas é uma forma de reduzir os riscos de desenvolver o diabetes, mas claro que as quantidades devem ser fracionadas de acordo com cada indivíduo. Dietas desequilibradas, pobres em fibras e nutrientes e com altas concentrações de ultraprocessados levam ao ganho de peso, fator que influencia na baixa sensibilidade à insulina. Este fato foi evidenciado em muitos estudos, como os de Chen *et al.*, (2016), Barengolts (2016) e Guiducci; Iervasi e Quinones-

Galvan (2014). Estes autores, ao avaliar o TTGO de indivíduos obesos ou propensos a obesidade, observaram aumento da glicemia pós prandial tardia avaliada três horas após carga glicêmica em 32% dos casos. Neste estudo verificou-se também que em 12% dos indivíduos os níveis de insulina estavam baixos, aumentando o risco do desenvolvimento da DM2.

A obesidade é um dos principais fatores que aumentam o risco do desenvolvimento da DM2. Os indivíduos considerados obesos possuem o IMC (índice de massa corporal) com valores maiores que 30 kg/m^2 (Organização Mundial da Saúde, 1997). Além do IMC, a medida da circunferência abdominal, muitas vezes é utilizada para caracterizar a obesidade, consistindo em valores maiores que 102 cm para os homens e maiores que 88 cm para as mulheres (1 Diretriz Brasileira de Diagnóstico e Tratamento da Síndrome Metabólica, 2004).

Malone e Hansen (2019) associaram a obesidade ao risco aumentado de desenvolver DM2, levando em consideração que os hábitos alimentares do obeso podem levar o indivíduo para um diagnóstico mais rápido do DM2, principalmente aqueles que apresentam predisposição genética. Contudo, esses autores observaram que a obesidade em si não foi um fator que causa alterações metabólicas quando não está associada a predisposição a DM2, mas a condição de resistência inerente a insulina pode resultar em alterações metabólicas que favorecem o armazenamento de gordura e conseqüentemente a obesidade.

O sedentarismo também tem um impacto em um maior risco de desenvolvimento da DM2, considerando que a falta de atividade física contribui para a diminuição da sensibilidade a insulina (Etchegoyen *et al.*, 2018; Arrieta-Cruz *et al.*, 2022; Pippi *et al.*, 2022). O estudo de León-Latre *et al.* (2014) mostrou que indivíduos que passam mais de 6,6 horas sentados apresentam níveis mais altos de glicose, insulina e HOMA-IR, independentemente da atividade física. Esse efeito ocorre devido à baixa ativação muscular, que reduz a captação de glicose e aumenta a resistência à insulina, além de estar associado a um estado inflamatório crônico, contribuindo para a disfunção metabólica e o risco de diabetes tipo 2.

Biddle *et al.* (2018) analisaram 435 indivíduos e acompanharam os níveis de glicose, insulina e outros marcadores bioquímicos para associar o comportamento físico dos indivíduos com os marcadores metabólicos e concluiu que quanto mais tempo os indivíduos passavam em pé e em movimento melhor eram os resultados dos marcadores bioquímicos em relação aos participantes que passavam mais tempo sentados. Além disso, sabe-se que a atividade física contribui para o ajuste das concentrações glicêmicas, uma vez que a captação de glicose pelo músculo ocorre por difusão facilitada, mediada pelos transportadores de glicose GLUT4. Em contrapartida, em repouso, a captação de glicose é predominantemente regulada pela ação da

insulina nos tecidos. Essa análise demonstra que a atividade física recorrente pode melhorar a sensibilidade a insulina (Bird; Hawley, 2016).

Outros estudos associaram o *stress* com uma maior predisposição a DM2. Em situações de altos níveis de estresse a produção descontrolada de cortisol influencia diretamente na sensibilidade a insulina. Isso acontece porque receptores de glicocorticoides são expressos em células beta pancreáticas, causando resistência à insulina quando as concentrações de cortisol se mantêm acima do normal por um longo período de tempo (Hackett; Steptoe, 2017). Um estudo que incluiu 3500 pessoas relatou que indivíduos diagnosticados com DM2 mantinham níveis normais de cortisol durante o dia enquanto durante a noite as concentrações do hormônio se elevavam em comparação aos indivíduos saudáveis (Hackett; Steptoe; Kumari, 2014). Ortiz *et al.* (2019) analisaram a relação entre o cortisol sérico matinal e o metabolismo da glicose em uma amostra de 4.206 afro-americanos. Os resultados indicaram que níveis mais altos de cortisol estavam associados a maior glicemia de jejum (FPG) e menor função das células beta em participantes sem diabetes, e a níveis mais altos de FPG e hemoglobina glicada (HbA1c) em participantes com diabetes. A análise revelou que o cortisol elevado aumentava a probabilidade de DM2, sugerindo que o cortisol matinal pode ser um marcador relevante no estudo da glicemia e do risco de diabetes.

Outro fator relevante é o sono inadequado, a privação crônica de sono e até mesmo o excesso, podem ser fatores importantes no desenvolvimento do DM2. Antza *et al.* (2021) realizaram uma meta-análise que envolveu 12 artigos de populações diferentes entre 2010 à 2018 que demonstrou que pessoas que dormem menos que 6 horas possuem duas vezes mais chances de serem diagnosticadas com DM2. Por outro lado, esse estudo demonstrou que pessoas que dormem mais que 9 horas por dia possuem 60% de chances de desenvolver DM2 em relação a pessoas que dormem uma quantidade de sete a oito horas de sono por noite.

Desse modo, fatores ambientais e o estilo de vida desempenham um papel crucial no desenvolvimento do DM2, no entanto, quando associados a fatores genéticos, a suscetibilidade a essa patologia aumenta significativamente como demonstrado nos achados de Fan *et al.* (2016) onde conduziram um estudo prospectivo com 2.293 indivíduos na Finlândia, analisando a capacidade de polimorfismos genéticos nos genes *PPARG* e *CAPN10* de prever o risco de DM2. Após até 12 anos de acompanhamento, foi observado que variantes nesses genes elevavam de forma expressiva o risco de DM2, especialmente em indivíduos com glicemia de jejum elevada e alto índice de massa corporal. Esses achados sugerem que, embora fatores ambientais e o estilo de vida sejam essenciais, polimorfismos genéticos específicos podem ser valiosos para prever o risco de DM2 em determinadas populações.

Ligthart *et al.* (2021) analisaram 15.671 indivíduos de ascendência europeia com 45 anos ou mais, participantes do ARIC (*Atherosclerosis Risk in Communities*) e do Rotterdam Study, para estimar o risco vitalício de diabetes tipo 2 com base em 403 variantes genéticas. Os participantes foram categorizados em risco genético baixo (menos de 20% de variantes genéticas), intermediário (20-80% de variantes genéticas) e alto (mais de 80% de variantes genéticas), e o impacto do IMC foi avaliado. O risco de diabetes foi maior entre aqueles com alto risco genético, chegando a 47,6% no ARIC e 35,5% no *Rotterdam Study*.

Salauddin *et al.* (2023) investigaram a associação dos polimorfismos rs12255372 (G/T) e rs7903146 (C/T) no gene *TCF7L2* com a suscetibilidade ao DM2 em uma amostra de 300 pacientes diabéticos e 234 controles saudáveis em Bangladesh. Os resultados mostraram que ambos os polimorfismos estavam significativamente associados ao DM2, com o alelo de risco "T" de rs12255372 apresentando uma associação mais forte. Foi observado que o genótipo TT de rs12255372 estava mais presente em diabéticos em comparação aos controles. Sendo assim, fatores ambientais e genéticos se encontram intimamente correlacionados com o desenvolvimento da DM2.

3.4 POLIMORFISMOS GENÉTICOS

Polimorfismos genéticos são variações na sequência de DNA que ocorrem dentro de genes, podendo estar presentes em suas regiões promotoras, codificantes (éxons), não codificantes (íntrons) ou até mesmo em regiões intergênicas. Diferente das mutações, que possuem uma frequência alélica inferior a 1%, uma variante só é classificada como polimorfismo se sua frequência alélica for superior a 1% na população mundial, independentemente do tipo de mutação que a originou, do tamanho do segmento de DNA envolvido ou se ela leva ao desenvolvimento de uma patologia ou não. Dependendo do tipo do polimorfismo, da sua localização e da influência de fatores ambientais, ele pode predispor a doenças ou alterações metabólicas, tais como a DM2; como também não levar a qualquer alteração fenotípica e surgimento de doenças (Nussbaum; McInnes; Willard, 2016).

Diversos tipos de polimorfismos são encontrados no genoma humano. Entre os menos frequentes estão os polimorfismos de comprimento de fragmento de restrição (RFLP), os polimorfismos de número variável de repetições em tandem (VNTR), também conhecidos como minissatélites, e as repetições em série curtas (STR), chamadas de microssatélites. O tipo mais comum, no entanto, é o polimorfismo de nucleotídeo único (SNP), que corresponde a uma alteração em apenas um nucleotídeo da sequência de DNA. Aproximadamente metade dessas

alterações polimórficas afeta a sequência de aminoácidos que compõem uma proteína, sendo classificadas como alterações não-sinônimas.

As variações dos SNPs podem ser silenciosas, quando a troca de nucleotídeo não modifica o aminoácido codificado; de sentido trocado (*missense*), quando o códon alterado gera um aminoácido diferente, o que pode ou não resultar em alterações funcionais significativas na proteína; ou sem sentido (*nonsense*), em que a modificação transforma o códon em um sinal de término. Além dos SNPs, existem os polimorfismos de inserção/deleção, ou *indels*, que envolvem a adição ou remoção de nucleotídeos, podendo variar de uma única base até cerca de mil nucleotídeos (Nussbaum; McInnes; Willard, 2016).

Vários estudos da literatura tem relacionado SNPs específicos com a predisposição a alterações metabólicas e distúrbios multifatoriais (Martinez-Lopes *et al.*, 2007; XIE *et al.*, 2017; Ishimaru *et al.*, 2017; Ghafar *et al.*, 2019; Uaska Sartori *et al.*, 2020). Martinez-Lopes *et al.*, (2007) estudaram a associação entre o polimorfismo A54T do gene FABP2 e os fatores de risco cardiovascular em indivíduos obesos mexicanos. Foram analisados 114 adultos obesos, sendo 47% resistentes à insulina, 49% com hipercolesterolemia e 45% com hipertrigliceridemia com a técnica molecular PCR-RFLP resultando em uma frequência do genotípica: 39% AA, 54,8% AT e 6,2% TT. O grupo AT/TT apresentou maior IMC, circunferência abdominal, níveis de triglicerídeos, colesterol total, LDL e VLDL em comparação ao grupo AA, indicando maior risco cardiovascular, o que atribuiu ao alelo T54 do gene FABP2 uma maior associação ao aumento do risco cardiovascular em obesos.

Xie *et al.* (2017) investigaram a associação entre quatro SNPs (rs10738760, rs6921438, rs6993770 e rs4416670) mapeados no gene *VEGF* (fator de crescimento endotelial vascular) com o desenvolvimento da depressão, uma patologia que afeta indivíduos globalmente. O estudo foi conduzido na população chinesa, composta por 437 pacientes diagnosticados com depressão e 477 indivíduos saudáveis que serviram como grupo controle. A técnica molecular utilizada foi PCR alelo-específica competitiva (KASP) acoplada a um sistema de genotipagem baseado em FRET e foi observado entre os SNPs analisados, que apenas o SNP rs4416670 mostrou associação significativa com o aumento do risco de desenvolver depressão nesses indivíduos. Ishimaru *et al.* (2017) avaliaram a relação do polimorfismo de deleção no gene *UGT2B17* com o prognóstico do câncer pediátrico na população japonesa pela técnica de PCR (Reação em Cadeia da Polimerase) em 145 amostras de DNA de tecidos tumorais. Os resultados deste estudo mostraram que as crianças com a ausência do polimorfismo de deleção apresentaram taxas de recidiva significativamente maiores em comparação às que possuíam o

polimorfismo, indicando que esse SNP desempenha um papel relevante no prognóstico de cânceres pediátricos.

O SNP -344C/T do gene *CYP11B2* foi relacionado com a hipertensão e hipertrofia ventricular esquerda na população egípcia em um estudo que avaliou 100 indivíduos hipertensos e 50 indivíduos saudáveis como grupo controle. As amostras de DNA extraídas do sangue foram genotipadas pela técnica de PCR-RFLP e determinou-se que a frequência alélica do alelo mutante (-344T) mostrou alta associação com o grupo de pacientes hipertensos quando comparados com o alelo selvagem (-344C) mais presentes no grupo controle, demonstrando que este SNP pode ser relacionado com a predisposição a hipertensão (Ghafar *et al.*, 2019).

Os estudos de análise de SNPs e sua associação com distúrbios multifatoriais na população brasileira estão em menor número quando comparado com as demais populações da maioria dos países (RIOS *et al.*, 2010; SARTORI *et al.*, 2020); WOBETO *et al.*, (2009) investigaram a relação entre os polimorfismos da haptoglobina (Hp) e a nefropatia diabética (ND) em 265 pacientes diabéticos brasileiros, sendo 95 com diabetes tipo 1 e 170 com tipo 2, todos com histórico da doença por pelo menos 10 e 5 anos, respectivamente. Os resultados mostraram que não houve associação significativa entre os genótipos da Hp e a presença de ND, nem correlação com a excreção de albumina urinária ou hipertensão arterial sistêmica.

Wollinger *et al.* (2015) avaliaram a associação entre os polimorfismos genéticos do gene da enzima conversora de angiotensina (*ACE*) e do gene do angiotensinogênio (*AGT*) com a suscetibilidade ao DM2 em uma população do sul do Brasil. A pesquisa, um estudo de caso-controle, incluiu 228 pacientes com DM2 e 183 indivíduos saudáveis como controles. Não foi encontrada associação significativa entre o polimorfismo InDel do gene *ACE* e a DM2, contudo, o polimorfismo rs699 do gene *AGT* mostrou uma associação significativa com a doença, com o genótipo AG sendo mais frequente nos pacientes com DM2.

Em uma amostra da população brasileira, Sartori *et al.* (2020) com o objetivo de avaliar o perfil genético dos indivíduos diagnosticados com hanseníase analisaram 19 marcadores genéticos do tipo SNPS, distribuídos entre 11 loci pré associados com a hanseníase. A amostra foi constituída por seis indivíduos brasileiros, sendo dois casos com relato de recidiva, um de reinfecção e três pacientes já infectados por hanseníase e curados. O DNA foi obtido pelo sangue periférico dos pacientes e os SNPs analisados por Real Time-PCR. Este estudo concluiu que a combinação de alelos entre os 19 diferentes genes pode resultar na predisposição aumentada para recorrência da doença.

Um outro estudo realizado com a população brasileira buscou a relação da doença multifatorial arterial coronariana (DAC), com os SNPs 511C>T e -174G>C, localizados na

região promotora dos genes IL-1B e IL-6, respectivamente, utilizando a técnica de PCR-RFLP. Os autores analisaram 667 indivíduos, sendo 253 afro-brasileiros e 414 caucasianos-brasileiros e identificaram que esses SNPs eram mais prevalentes em homens com DAC, diagnosticados com DM2 e que faziam uso de medicamentos para controle de colesterol, em comparação com os respectivos grupos de controle (Rios *et al.*, 2010).

A associação de SNPs específicos com o desenvolvimento da DM2 foi observada em muitas populações, como Lyssenko *et al.* (2005) revelaram que variantes nos genes *PPARG* e *CAPN10* em participantes finlandeses estavam associadas a um risco aumentado de DM2, especialmente em indivíduos com níveis elevados de glicose em jejum e alto índice de massa corporal (IMC). Morcillo *et al.* (2008) observaram que os SNPs do alelo -75A do gene *APOA* (-75G/A) e do alelo Arg64 do gene *ADRB3* aumenta o risco do desenvolvimento da DM2 em uma população do sul da Espanha. Dora *et al.* (2010) identificaram que a homozigose Ala92Ala do SNP Thr92Ala do gene *D2* está associada ao aumento do risco de desenvolver DM2 em participantes do sul do Brasil. Perez-Luque *et al.* (2012) detectaram que o SNP -308G/A do gene *TNF- α* é um fator de risco para o desenvolvimento da DM2 em mexicanos. Ainda na população mexicana, Sanchez *et al.* (2019) detectaram que o alelo G do SNP +45T/G do gene *ADIPOC* apresentou efeito protetor contra distúrbios metabólicos relacionados a DM2. Salauddin *et al.* (2023) desmonstraram que o polimorfismo genético rs7903146 do gene *TCF7L2* está associado ao risco de DM2 em indivíduos da população de Bangladesh.

3.5 GENE UCP2

O gene *UCP2* (*Uncoupling Protein 2*), localizado no braço longo do cromossomo 11 no genoma humano, pertence à família de genes *UCPs* e codifica uma proteína homônima que regula o transporte de íons pela membrana mitocondrial, influenciando diretamente a cadeia respiratória dessa organela e, conseqüentemente, a síntese de ATP (Hashemi *et al.*, 2014). O *UCP2* é expresso em diversos tecidos do corpo, como o tecido adiposo, músculos, cérebro, pulmões, rins e pâncreas, desempenhando um papel crucial no metabolismo de lipídios e na regulação da secreção de insulina (Liu *et al.*, 2013; Rezapour *et al.*, 2021). A atuação da proteína codificada pelo *UCP2* na membrana mitocondrial impacta a produção e liberação de adenosina trifosfato (ATP), uma vez que ela controla o transporte de ânions para fora e de prótons para dentro da mitocôndria, fatores que favorecem a cadeia respiratória e a geração de ATP como produto final. Quando sua expressão está aumentada, a proteína *UCP2* pode reduzir a síntese

de ATP, o que afeta tanto a função mitocondrial quanto as atividades celulares normais (Cadenas, 2018).

Além disso, o *gene UCP2* também regula a formação e liberação de espécies reativas de oxigênio (ROS), frequentemente subprodutos tóxicos resultantes da cadeia respiratória mitocondrial. As ROS, como superóxidos (O_2^-) e peróxidos de hidrogênio (H_2O_2), estão associadas ao estresse oxidativo e a processos inflamatórios, sendo responsáveis por danos intracelulares (Brand, 2016). Estudos de Mailloux e Harper (2011) mostraram que o aumento da concentração de ROS liberadas pela mitocôndria ativa as proteínas UCP2 e UCP3, reduzindo seus níveis. De forma semelhante, Bai *et al.* (2005) observaram que a expressão aumentada do *gene UCP2* em camundongos diminuiu a liberação de ROS, enquanto a sua redução da expressão aumentou a liberação dessas espécies reativas.

Diversos polimorfismos do tipo SNPs associados a patologias têm sido identificados no *gene UCP2*, com destaque para o -866G/A (rs659366), que tem sido associado ao desenvolvimento do diabetes tipo 2 (Shen *et al.*, 2014; Cadenas, 2018; XU *et al.*, 2021). O SNP -866G/A está localizado na região promotora do *UCP2*, onde o alelo G é considerado o tipo selvagem, enquanto o alelo A é a variante mutante. A presença do alelo A, seja em homozigose ou heterozigose tem sido associada ao aumento da expressão do *gene UCP2*, como nos achados de Boschini *et al.* (2005) que revisaram a regulação da expressão gênica das proteínas desacopladoras UCP2, com foco em como fatores como dieta, jejum e exercício físico influenciam essa expressão. O estudo inclui dados de resultados em humanos e animais, destacando que o alelo A do polimorfismo -866G/A do *gene UCP2* pode estar associado a um aumento na expressão do gene, impactando funções metabólicas e adaptativas relacionadas à homeostase celular.

A quantidade de pesquisas que buscam a relação entre o polimorfismo -866G/A do *gene UCP2* e o desenvolvimento de DM2, ou patologias relacionadas a diabetes, como retinopatia diabética e síndrome metabólica, têm aumentado consideravelmente nos últimos anos, com um maior destaque para a população asiática. Sasahara *et al.* (2004) foram os pioneiros a estudar a relação desse polimorfismo com a DM2 na população da Ásia. Este trabalho analisou na população japonesa a diferença genotípica entre 413 participantes com diagnóstico de DM2 e 172 indivíduos não-diabéticos. A presença do SNP -866G/A foi avaliada pela técnica de PCR-RFLP com a enzima de restrição *MluI* em amostras de DNA extraídas de sangue periférico. Estes autores não observaram diferenças significativas entre os genótipos mutante e selvagem entre os grupos, contudo foi possível identificar que os participantes portadores do alelo A

possuíam níveis de insulina mais baixos que os indivíduos não-diabéticos portadores do alelo G.

Wang; Ke e Lin (2012) investigaram 370 participantes chineses com 2 DM2 e 166 indivíduos saudáveis, demonstrando que o tratamento com repaglinida foi mais eficaz nos pacientes portadores do alelo selvagem G em comparação aos portadores do alelo mutante A, que apresentaram maior resistência ao medicamento. De forma complementar, Chai *et al.*, (2012). exploraram a associação desse mesmo polimorfismo com o risco de desenvolvimento de ataque isquêmico transitório (AIT), conhecido como "mini AVC". O estudo incluiu 844 pacientes chineses com DM2, divididos em dois grupos: 404 que sofreram AIT e 440 que não apresentaram histórico dessa condição. Por meio da análise de PCR quantitativo em tempo real (qPCR) com sondas TaqMan específicas, os autores observaram que o alelo mutante A estava significativamente associado a um maior risco de AIT em indivíduos com DM2. Em contraste, Qin *et al.* (2013) obtiveram resultados divergentes, não encontrando associação significativa entre o polimorfismo -866G/A e o desenvolvimento de DM2 na população chinesa. Nesse estudo, foram analisadas amostras de DNA de 715 indivíduos, extraídas de sangue periférico e submetidas à técnica de PCR-RFLP, seguida de digestão enzimática com *MluI*. Os autores concluíram que não houve evidências de relação entre o polimorfismo -866G/A e a predisposição ao DM2.

O estudo de Hou *et al.* (2020) identificaram que os indivíduos chineses portadores do alelo A (mutante) do SNP -866G/A possuíam baixos níveis de insulina e aumento do risco do desenvolvimento de pré-diabetes ou DM2. Lapik, Ranjit e Galchenko (2021), por meio da comparação entre a presença dos alelos selvagens G ou mutantes A com os exames bioquímicos dos participantes, concluíram que nas mulheres russas com diagnóstico de DM2 a presença do genótipo mutante AA estava associado a níveis glicêmicos e lipídicos elevados e o genótipo selvagem GG estava associado a níveis normais de glicemia. Os resultados de Gomathi *et al.* (2019) mostraram que na população indiana, o genótipo AA estava significativamente associado a altos níveis de hemoglobina glicada e baixos níveis de insulina no sangue quando comparado com os genótipos GA e GG, concordando com os resultados de estudos prévios sobre a relação do genótipo AA com o desenvolvimento da DM2.

Os achados de Din *et al.* (2023) demonstraram a relação entre a polimorfismo do *gene UCP2* e a predisposição a DM2 na população de Caxemira, norte da Índia, onde foram incluídos 850 participantes, sendo 425 com DM2 e 425 controles saudáveis. Os resultados mostraram que as mutações no *gene UCP2*, particularmente o genótipo AA no locus -866, estão associadas a uma redução expressiva na expressão do *UCP2*, com impacto negativo nos parâmetros

metabólicos. Pacientes com idade superior a 50 anos e $IMC \geq 25$ apresentaram uma diminuição significativa na expressão do *UCP2*, sugerindo que essas características aumentam o risco para DM2.

Por outro lado, os achados de Bulotta *et al* (2005) que analisaram 746 pacientes com DM2 e 327 controles saudáveis de origem caucasiana, todos residentes da região de Gargano, Itália mostraram que o alelo -866A estava associado a um risco reduzido de DM2, com os portadores apresentando risco progressivamente menor conforme o número de alelos A aumentava. Por outro lado, o genótipo -866G/G conferiu risco aumentado, atribuindo cerca de 12% de risco populacional ao desenvolvimento de DM2. Andersen *et al.* (2012) investigaram a relação entre o polimorfismo -866G>A do *gene UCP2* e DM2 em uma amostra de 17.636 dinamarqueses. A genotipagem foi realizada utilizando a técnica KASPar (KBioscience), com uma taxa de sucesso de 97% e sem discrepâncias em 965 amostras duplicadas aleatoriamente. Os resultados não mostraram diferenças significativas nas frequências alélicas ou na distribuição dos genótipos entre os casos de DM2 e os controles, nem em análises prospectivas de conversão para diabetes após cinco anos de acompanhamento. Conclusivamente, o SNP -866G/A não apresentou associação significativa com o risco de desenvolvimento de DM2.

Huang *et al.* (2021) avaliaram a associação dos polimorfismos UCP1-3826A/G, UCP2-866G/A e Ala55Val, e UCP3-55C/T com a suscetibilidade ao DM2. A análise incluiu 38 estudos de caso-controle envolvendo diferentes etnias e utilizou odds ratios (ORs) com intervalos de confiança de 95% para estimar a força da associação. Os resultados mostraram que o polimorfismo *UCP2* -866G/A estava associado a uma redução do risco de DM2 em asiáticos, enquanto o polimorfismo *UCP2* Ala55Val estava relacionado a um aumento do risco em populações asiáticas.

Na população brasileira, a investigação do SNP -866G/A do *gene UCP2* em relação à predisposição a DM2 até o momento somente foi realizada em apenas um estudo, o qual se limitou a amostra exclusivamente de indivíduos de descendência européia. Neste estudo, não foram identificadas diferenças significativas entre os alelos A e G do SNP -866G/A em relação à regulação da insulina, da glicose ou de outros marcadores diagnósticos para DM2 (De Souza *et al.*, 2013).

Atualmente, diversas técnicas são utilizadas para identificar SNPs, incluindo PCR-RFLP, *qPCR Genotyping*, *Tetra-primer ARMS-PCR*; *SNP Array*, entre outras. Embora essas metodologias sejam precisas, seguras, não invasivas e capazes de detectar alterações relacionadas à predisposição ao desenvolvimento de doenças antes mesmo do seu surgimento, elas apresentam limitações significativas. Entre essas limitações, destacam-se o alto custo, o

tempo prolongado para obtenção dos resultados, a complexidade da técnica que exige profissionais especializados e o elevado risco de contaminação devido aos diversos processamentos amostrais realizados durante as etapas experimentais (Medrano; De Oliveira, 2014; Matsuda, 2017; Cheung; Bi, 2018; Gomathi *et al.*, 2019). Assim, a implementação de técnicas mais rápidas e simples, como a análise de espectroscopia no Infravermelho por Transformada de Fourier (FTIR), podem oferecer resultados complementares que sejam mais precisos e robustos em relação às análises moleculares e bioquímicas.

3.6 ATR-FTIR APLICADA A ÁREAS MÉDICAS

A técnica de FTIR possui o potencial de caracterizar diferentes moléculas de diversos materiais biológicos, como os tecidos, plasma sanguíneo, ácidos nucleicos, entre outros. O diagnóstico é baseado nas variações espectrais observadas entre as diferentes amostras, nas quais os modos vibracionais das moléculas são absorvidos em comprimentos de onda específicos da radiação infravermelha. Essas diferenças espectrais permitem a identificação e caracterização das amostras analisadas. As alterações podem ser visualizadas por meio de um espectro médio da faixa 4000 a 400 cm^{-1} , que revela a identidade espectral característica de cada amostra (Lewis *et al.*, 2010). Uma das principais vantagens desta técnica é sua simplicidade de execução, requerendo apenas pequenas quantidades de amostra, além de ser de fácil manuseio, proporcionar resultados rápidos e apresentar um custo reduzido na obtenção dos resultados (Kahn *et al.*, 2009).

A Reflectância Total Atenuada (ATR) é uma das tecnologias de amostragem mais utilizadas na espectroscopia FTIR. O método ATR baseia-se na reflexão interna total, onde a onda evanescente penetra superficialmente na amostra. A profundidade de penetração do infravermelho é determinada pelo comprimento de onda da radiação e pelos índices de refração do cristal ATR e da amostra. A variação de medida entre as amostras sólidas acaba resultando em espectros com a mesma aparência. Por isso, a razão do uso da técnica ATR combinada com a espectroscopia FTIR permite a avaliação de amostras líquidas ou sólidas de uma forma mais limpa, simplificando a análise de praticamente todas as amostras (Kazarian; Chan, 2006).

O uso da espectroscopia no infravermelho possui muitas vantagens em relação a técnicas tradicionais de diagnóstico, com resultados em tempo real ou que podem ser caracterizados por meio de processamento de dados por softwares automatizados e com a possibilidade de diagnóstico não invasivos ou minimamente invasivos. Além disso, esta técnica possui especificidade na determinação das mudanças bioquímicas que acontecem no indivíduo, que

podem indicar futuras alterações morfológicas quando as mesmas ainda não se manifestaram micro e macroscopicamente (Rein *et al.*, 2020).

Vários estudos têm aplicado a técnica de ATR-FTIR na análise de tecidos e no diagnóstico de vários tipos de neoplasias humanas. Li *et al.* (2005) realizaram o diagnóstico de câncer colorretal em quatro pacientes chineses durante procedimentos cirúrgicos, utilizando um cristal ATR esterilizado previamente. O cristal foi colocado diretamente sobre a superfície do tecido irregular e analisado em tempo real por meio da técnica de FTIR ao longo da cirurgia. Além disso, o estudo incluiu a análise de um quinto paciente, também chinês, que foi avaliado antes da cirurgia por FTIR em um tecido irregular localizado a 3 cm de distância do ânus. Os resultados, indicaram que essa técnica permite a identificação de cânceres colorretais de forma significativamente mais simples e eficiente. Os achados de Dong *et al.* (2023) possibilitaram a identificação da metástase em linfonodos de câncer gástrico, com a análise de 80 amostras de linfonodos metastáticos e 80 amostras de linfonodos não metastáticos, utilizando espectros obtidos por meio da espectroscopia FTIR. Após a Análise dos Componentes Principais (PCA) e a Análise Discriminante de Fisher (FDA) foi possível diferenciar de maneira eficaz os linfonodos metastáticos dos não metastáticos.

Adicionalmente, além de sua aplicação em tecidos, a espectroscopia FTIR tem sido utilizada para fins diagnósticos em diversos materiais biológicos, incluindo saliva, soro, urina, sangue total. Um estudo recente de Kino *et al.* (2024) coletou amostras de plasma de 84 pacientes com glioma antes do início do tratamento e de 72 participantes saudáveis no Japão. Os espectros foram processados pelo método Savitzky-Golay (segunda derivada), revelando diferenças de absorção nos comprimentos de onda 1700–1500 cm^{-1} (associados aos grupos amida I e II) e 1200–840 cm^{-1} (relacionados a metabólitos de lipídios e açúcares). Essas diferenças entre os grupos em determinados comprimentos de onde foram submetidas à Análise Discriminante de Mínimos Quadrados Parciais (PLS-DA), resultando em um potencial mecanismo para o prognóstico do glioma.

Elmi *et al.* (2017) coletaram 43 amostras de soro de pacientes com câncer de mama e 43 amostras de pacientes controle. Após a coleta, as amostras de soro foram armazenadas a -80°C até a análise por espectroscopia ATR-FTIR. O processamento dos espectros foi realizado utilizando o método elástico e a normalização foi aplicada pelo método vetorial. Para a discriminação dos grupos, a Análise Discriminante Linear (LDA) foi combinada com a PCA com o objetivo de agrupar os dados de acordo com variáveis independentes. Esse processo permitiu a separação clara entre os dois grupos. O critério mais relevante para essa distinção foi a faixa de comprimento de onda de 3090 – 3700 cm^{-1} caracterizada pelo perfil proteico, que

apresentou alterações significativas nas pacientes com câncer de mama em comparação às mulheres saudáveis.

Em um estudo de triagem diagnóstica para endometriose avançada, utilizando o soro como amostra biológica, as amostras foram secas e analisadas por meio da espectroscopia ATR-FTIR. O estudo incluiu 29 pacientes previamente diagnosticadas com endometriose, 24 pacientes sem endometriose, mas com outras condições patológicas, como cistos, leiomiomas ou neoplasias, e um grupo controle composto por 18 participantes saudáveis, todas de origem polonesa. Os dados espectrais foram processados por PCA e PLS-DA. Os resultados apresentaram padrões de sensibilidade, especificidade e precisão superiores a 80%, demonstrando uma clara separação entre os grupos analisados (Kokot *et al.*, 2022).

Para o diagnóstico de paracoccidiodomicose, uma micose sistêmica granulomatosa, foram utilizadas amostras de soro de 132 pacientes infectados pelo fungo causador da doença, sendo 50 pacientes saudáveis e 42 pacientes diagnosticados com outras infecções fungicas. As amostras foram analisadas pela espectroscopia ATR-FTIR e os dados espectrais foram submetidos a análise quimiométrica, utilizando a PCA e a análise discriminante de mínimos quadrados parciais ortogonais (OPLS-DA). Essa abordagem mostrou-se significativamente mais rápida e sensível para o diagnóstico precoce da paracoccidiodomicose, alcançando 100% de especificidade na separação dos dois grupos analisados (Koehler *et al.*, 2022).

Nurdalila *et al.*, (2015) avaliaram o gene 16S de peixes para classificação taxonômica de duas localizações geográficas distintas, utilizando 125 amostras de sequências isoladas de DNA. A análise foi conduzida por espectroscopia ATR-FTIR e os dados foram processados por meio da PCA para uma primeira identificação dos grupos. Três réplicas de cada amostra foram agrupadas em um cluster discreto, demonstrando que o PCA foi capaz de diferenciar os indivíduos com base em suas amostras. No gráfico de dispersão das pontuações foi possível identificar claramente dois grupos distintos, relacionados às diferentes localizações dos peixes, evidenciando a eficácia da técnica para essa aplicação.

Além das análises convencionais por espectroscopia FTIR para discriminação dos grupos, algumas amostras requerem métricas individualizadas para otimizar a distinção entre dados complexos. Nesse contexto, mais recentemente algoritmos de *machine learning* têm sido aplicados em vários estudos da literatura para otimizar os resultados obtidos pela análise de FTIR. No estudo de Du *et al.* (2023) foram avaliadas 383 amostras de novas substâncias psicoativas de diferentes categorias, utilizando técnicas de aprendizado de máquina para detecção de alterações estruturais das substâncias. Para otimização, foram empregados diversos algoritmos, como o k-Nearest Neighbors (KNN), que prevê a classe de um pico espectral com

base na proximidade de outros e o *Support Vector Machines* (SVM), utilizado para classificar e diferenciar amostras em um plano específico. Algoritmos baseados em árvores de decisão, como Random Forest (RF) e Extra Trees (ET), também foram aplicados. No algoritmo RF, o modelo é treinado para gerar decisões a partir da agregação de múltiplos agrupamentos e, posteriormente, selecionar uma média das previsões. O algoritmo ET, por sua vez, realiza uma análise aleatória dos dados, utilizando um conceito semelhante ao do RF.

Outro estudo realizado por Du *et al.* (2023), testou individualmente e em combinação seis modelos de *machine learning* KNN, SVM, RF, ET, voting e artificial neural networks (ANNs) para classificar substâncias psicoativas, considerando sua grande diversidade estrutural. Embora alguns algoritmos não tenham conseguido separar 100% das amostras, os resultados demonstram um grande potencial para a classificação dessas substâncias, especialmente com o modelo Voting, que apresentou os melhores índices de precisão e recall. Novos estudos são necessários para aprimorar a análise, mas as abordagens de *machine learning* aplicadas revelam-se promissoras para a otimização da discriminação de compostos químicos.

Zhang *et al.* (2024) utilizaram a técnica de espectroscopia FTIR para analisar pupários vazios de cinco espécies diferentes de moscas, conseguindo distinguir com sucesso os espectros entre as espécies, confirmando a capacidade da técnica em discriminar taxonomicamente insetos com base em seus espectros. Para a classificação das espécies, foram utilizados três modelos de aprendizado de máquina: SVM, KNN e RF. O modelo SVM apresentou o melhor desempenho, atingindo uma precisão de 100% na identificação das cinco espécies analisadas, especialmente ao utilizar a região de impressão digital biológica (1800–1300 cm^{-1}), que se mostrou a mais eficiente para a identificação das espécies de pupários vazios.

Zhang *et al.* (2024) utilizaram a espectroscopia de ATR-FTIR e os algoritmos de *machine learning* SVM, RF, KNN e ANN para a otimização na identificação de casos de morte cardíaca súbita. Para isso, os autores utilizaram sangue total de 77 pacientes com o diagnóstico da causa de morte. O primeiro grupo consistiu em 34 indivíduos cuja causa principal de morte foi doença cardíaca coronária, enquanto o segundo grupo incluiu 43 participantes com laudos forenses de outras causas de morte. Dentre eles, o algoritmo SVM apresentou a maior eficiência na classificação dos grupos distintos, com 95.83% de acurácia. No entanto, os autores destacam que novos estudos com um maior número de amostras são necessários para validar Os resultados obtidos e o método utilizado.

A espectroscopia FTIR combinada com *machine learning* tem sido aplicada na identificação de genótipos, como no estudo de Hârța *et al.* (2020), que analisaram genitores de *Streptocarpus* × *hybridus* e suas progênies F1 usando marcadores SRAP e FTIR. O SRAP

(*Sequence-Related Amplified Polymorphism*) é um marcador molecular que amplifica seletivamente regiões codificantes do DNA, permitindo identificar polimorfismos por meio de perfis de bandas gerados na eletroforese. No estudo, o tipo III do SRAP foi o mais eficiente na identificação dos híbridos, pois revelou bandas específicas herdadas exclusivamente do genitor paterno. Isso permitiu confirmar a identidade dos híbridos, já que a presença dessas bandas indica a transmissão de segmentos do DNA do pai para a progênie. A análise FTIR (800–1800 cm^{-1}) detectou variações estruturais no DNA, como vibrações do esqueleto fosfato-açúcar e interações de bases nucleotídicas. A análise multivariada (PCA e HCA) diferenciou genitores e híbridos com base nas diferenças químicas, evidenciando variações genéticas úteis no melhoramento de *Streptocarpus*.

Na literatura, apenas o estudo de Gomes Rios *et al.* (2021) utilizaram a técnica ATR-FTIR como ferramenta para estratificar diferentes genótipos de SNPs. Estes autores utilizaram material genético bovino na avaliação do SNP C > T na posição 1547 do gene *bGH*. A técnica PCR-RFLP foi utilizada para a identificação dos genótipos, em um total de 60 amostras, das quais 20 eram homozigotas CC, 20 heterozigotas CT e 20 homozigotas TT. O produto da PCR foi aplicado no cristal de ATR pela técnica de sobreposição. No espectrômetro FTIR, foram realizadas 10 varreduras, e o tratamento inicial dos espectros foi realizado utilizando o algoritmo *machine learning* pelo método SVM.

Após a análise PCA e separação dos grupos, um algoritmo de *machine learning* supervisionado foi aplicado para classificar as amostras conforme seus grupos correspondentes. As análises realizadas incluíram a análise discriminante linear e quadrática, SVM e KNN. Para avaliar a precisão do modelo de classificação, foi utilizada a validação cruzada *leave-one-out* (LOO-CV). A acurácia total do modelo foi de 75% na distinção entre os três grupos genotípicos (CC, CT e TT). Os resultados mais precisos foram obtidos pelos modelos KNN e SVM, alcançando acurácia de até 95% na distinção entre os genótipos CD e DD, 90% para CC e CD, e 82,5% para CC e DD. Assim, o estudo demonstrou que a combinação de ATR-FTIR e *machine learning* pode ser uma abordagem eficaz para estratificar os diferentes genótipos bovinos.

Os resultados desses estudos demonstram que a espectroscopia FTIR é uma técnica eficaz na diferenciação bioquímica entre diferentes grupos amostrais. Quando combinada com algoritmos de *machine learning*, apresenta potencial para discriminar genótipos, permitindo o treinamento do sistema para uma classificação mais precisa e eficiente das amostras. Além disso, essa abordagem se destaca como uma alternativa mais econômica, reduzindo os custos elevados associados a métodos tradicionais de análise genética.

3.7 ATR-FTIR APLICADA AO DM2

A técnica de espectroscopia ATR-FTIR também tem sido eficaz na estratificação de indivíduos com diagnóstico de DM2 e sem DM2, utilizando saliva e sangue como amostra biológica. Caixeta *et al.* (2023) analisou por espectroscopia FTIR 68 amostras de saliva de participantes, dos quais 23 eram não diabéticos e 45 apresentavam DM2 não controlada. O teste de Kolmogorov–Smirnov foi aplicado para verificar a normalidade das variáveis e o teste *t de Student* foi utilizado para análise estatística das áreas das bandas espectrais. Os resultados indicaram um aumento significativo nas bandas em 2962 cm^{-1} (CH_3 de lipídios), 1641 cm^{-1} (amida I) e 1073 cm^{-1} (carboidratos e proteínas glicosiladas) nos indivíduos com DM2 não controlada em comparação aos não diabéticos. As análises revelaram taxas de sensibilidade de 93,3% (42/45), especificidade de 74% (17/23) e precisão de 87%, evidenciando o potencial da técnica de espectroscopia ATR-FTIR como uma alternativa promissora para a triagem e monitoramento de pacientes com DM2 (Caixeta *et al.* 2023).

Sánchez-Brito *et al.* (2021) desenvolveram uma metodologia baseada na espectroscopia ATR-FTIR em amostras de saliva para caracterizar pacientes com diagnóstico de DM2 controlada com pacientes com DM não controlada, além de estimar valores específicos de glicose. Foram analisadas 540 amostras de saliva obtidas de pacientes com DM2 entre 30 e 80 anos. Os espectros foram analisados na faixa de $4000\text{--}2000\text{ cm}^{-1}$ (região de ligação de hidrogênio) e $1800\text{--}800\text{ cm}^{-1}$ (região de impressão digital biológica) usando três modelos de aprendizado de máquina: ANN, SVM e regressão linear (LR). Os melhores resultados foram obtidos com ANN, que alcançou precisão de 100% na classificação entre os pacientes com diagnóstico de DM2 controlada e dos pacientes com DM2 não controlada e obtenção de alta correlação ($R^2 > 99,5\%$) na estimativa dos valores de glicose.

Andrade *et al.* (2023) propuseram o uso de uma técnica baseada em Evolutionary Neural Architecture Search (NAS) para otimizar arquiteturas de redes neurais convolucionais (CNN) com o objetivo de classificar espectros ATR-FTIR de saliva de pacientes com DM2 e indivíduos não diabéticos. O estudo incluiu 68 participantes (23 não diabéticos e 45 com DM2 não controlada). O modelo foi treinado com 51 espectros e validado em 17 amostras. A metodologia utilizou um Algoritmo Genético (GA) para ajustar hiperparâmetros como número de canais, número de camadas conectadas, taxa de dropout e taxa de aprendizado. A melhor configuração, composta por duas camadas convolucionais com 2 e 64 neurônios, e duas camadas totalmente conectadas com 128 e 16 neurônios, alcançou uma acurácia balanceada de 83%, macro F1-

score de 86% e área sob a curva (AUC) de 91%. A análise de SHAP indicou que as bandas vibracionais associadas às proteínas foram as mais relevantes para a classificação.

Hoşafçı *et al.* (2006) desenvolveram um método baseado em espectroscopia ATR-FTIR para análise de glicose em amostras de sangue, plasma sanguíneo e urina sem a necessidade de reagentes. Foram adquiridos espectros na faixa de 1760 a 800 cm^{-1} de 367 amostras de plasma sanguíneo, 372 amostras de sangue total e 231 amostras de urina. Para as amostras de sangue total, foi utilizado fluoreto de sódio (NaF) para inibir a glicólise e evitar a metabolização da glicose após a coleta. A calibração foi realizada por regressão por mínimos quadrados parciais (PLS), com erro quadrático médio de previsão (RMSEP) de 7,5 mg/dL para plasma, 7,6 mg/dL para sangue total com NaF e 5,2 mg/dL para urina. Os resultados para as amostras de urina indicaram um coeficiente de determinação (R^2) de 0,985 e erro de previsão de 5,2 mg/dL, demonstrando alta precisão na determinação dos níveis de glicose.

Caixeta *et al.* (2023) e Farooq e Zezell (2023) avaliaram a concentração de glicose em amostras de urina de ratos Wistar por meio de espectroscopia FTIR-ATR, demonstrando o potencial da técnica para monitoramento de DM2. Caixeta *et al.* (2023) obtiveram uma separação dos grupos com 99% de precisão, identificando o pico em 1074 cm^{-1} como o mais relevante para glicose urinária, com correlação de 78,9% com métodos enzimáticos tradicionais. Farooq e Zezell (2023) analisaram 149 amostras de urina e aplicaram análise discriminante 3D-PCA-LDA, obtendo acurácia, sensibilidade e especificidade de 100%, com picos em 1643 cm^{-1} , 1587 cm^{-1} e 1075 cm^{-1} associados a biomarcadores metabólicos. Os resultados reforçam a eficiência da espectroscopia FTIR-ATR para o diagnóstico e monitoramento de DM2.

No contexto clínico, Richardson *et al.* (2023) aplicaram a espectroscopia FTIR-ATR para detecção precoce da doença renal diabética (DKD) em pacientes humanos, evidenciando a aplicabilidade da técnica em condições patológicas associadas ao DM2. Foram analisadas 177 amostras de urina, sendo 64 de pacientes com normoalbuminúria, 61 com microalbuminúria, 30 com macroalbuminúria e 22 de indivíduos saudáveis. A análise espectral foi realizada na faixa de 4000 a 800 cm^{-1} , utilizando um cristal de diamante ATR, com 40 varreduras e resolução espectral de 4 cm^{-1} . Os espectros foram processados por PCA para identificar padrões associados à albuminúria. Para classificação, foram aplicados modelos de PLS-DA e SVM-DA. O modelo SVM-DA apresentou o melhor desempenho, alcançando sensibilidade de 97%, especificidade de 100% e área sobre a curva de 1.0 para diferenciação entre normoalbuminúria e micro/macroalbuminúria. Além disso, foi realizada análise de regressão para quantificação de

albumina, com erro médio quadrático (RMSE) de 17 mg/L e 44 mg/L para pacientes com normo e microalbuminúria, respectivamente.

Li *et al.* (2017) utilizaram o sangue total obtido de 60 pacientes chineses para determinar a concentração de glicose e colesterol usando a técnica de espectroscopia no infravermelho médio com equipamento de resolução de 6 cm^{-1} acoplada ao cristal de ATR. Na análise dos dados foi realizado o método de espectros derivativos para detectar características sutis entre as bandas, o método de fusão de Mínimos Quadrados Parciais (PLS) e algoritmo SPXY para otimizar a separação dos grupos. Os resultados apresentaram alta precisão, com coeficiente de determinação (R^2) de 0.995 para glicose e 0.947 para colesterol, além de baixos erros de previsão (RMSEP de 0.365 e 0.194, respectivamente), demonstrando-se promissores no sentido de propiciar um diagnóstico mais rápido e de baixo custo.

Huang *et al.* (2021) utilizaram a espectroscopia FTIR-ATR em amostras de sangue total para o diagnóstico de DM2, combinando a técnica com o algoritmo de extreme gradient boosting (XGBoost) para classificação dos espectros. O estudo incluiu 51 pacientes com DM2 e 55 voluntários saudáveis, totalizando 113 amostras coletadas na China. Os espectros foram

obtidos na faixa de $4500\text{ a }600\text{ cm}^{-1}$ com resolução espectral de 2 cm^{-1} , sendo submetidos ao método de suavização de Savitzky–Golay e à PCA para redução de ruído e seleção de variáveis. O modelo XGBoost foi otimizado usando um algoritmo de busca em grade, alcançando sensibilidade de 95,23%, especificidade de 96% e acurácia de 95,65% no conjunto de teste. Os picos mais relevantes para diferenciação foram identificados na região de $1000\text{ a }1500\text{ cm}^{-1}$, associados a ligações fosfodiéster e vibrações de C–O e C–N.

Pais e Araujo *et al.* (2019) desenvolveram um sistema óptico baseado na técnica de infravermelho próximo (NIR) para medição não invasiva dos níveis de glicose. A metodologia baseou-se em um sistema composto por emissores e receptores de faixas de frequência infravermelha, configurado para medir a concentração de glicose em soluções preparadas com valores de 40, 70, 120 e 200 mg/dL. A calibração foi realizada utilizando a Lei de Beer-Lambert, com uma curva de resposta em tensão e regressão linear para determinar a tendência do sistema em relação a cada valor de glicose. O dispositivo desenvolvido, baseado em um módulo Orange Pi Zero e um conversor analógico-digital MCP3208, alcançou uma variação de níveis de glicose de aproximadamente 17 mg/dL.

Os estudos apresentados demonstram que a técnica de espectroscopia ATR-FTIR tem se consolidado como uma ferramenta promissora para o diagnóstico e monitoramento de DM2. A aplicação da técnica em diferentes amostras biológicas, como saliva, sangue e urina, associada a métodos avançados de análise de dados, como aprendizado de máquina e análise

multivariada, tem proporcionado resultados de alta acurácia, sensibilidade e especificidade. A capacidade de identificar biomarcadores específicos de DM2 de forma rápida, precisa e não invasiva destaca o potencial da espectroscopia ATR-FTIR para complementar os métodos diagnósticos tradicionais, oferecendo uma alternativa viável e de baixo custo para a triagem e acompanhamento clínico de pacientes com DM2. Esses avanços reforçam a relevância da técnica na prática clínica e sugerem que futuras pesquisas podem expandir ainda mais sua aplicabilidade, melhorando a estratificação e o tratamento personalizado de pacientes com DM2.

4 MATERIAIS E MÉTODOS

Este estudo é constituído uma parte do projeto auxílio pesquisa intitulado “Polimorfismos genéticos e espectroscopia FTIR como ferramentas na identificação da predisposição a distúrbios multifatoriais”, aprovado pela agência FAPESP (processo nº 2023/16314-6) e pelo Comitê de Ética em Pesquisa em seres humanos da Universidade do Vale do Paraíba - UNIVAP (CAAE nº 62978022.0.0000.5503 , parecer nº 7.294.487) (Anexo A). A coleta das amostras e a análise bioquímica foram realizadas no Centro de Diagnóstico Laboratorial – CDLab pertencente a Faculdade de Ciências da Saúde, UNIVAP e toda a parte experimental foi desenvolvida no Laboratório de Genética Molecular – GeneLab e Central Multiusuário pertencentes ao Instituto de Pesquisa e Desenvolvimento, IP&D, UNIVAP.

4.1 DELINEAMENTO DO ESTUDO E COLETA DAS AMOSTRAS

Amostras de sangue periférico obtidas de participantes voluntários que frequentam o Centro de Práticas Supervisionadas (CPS) da Univap foram coletadas após estes autorizarem e assinarem o Termo de Consentimento Livre e Esclarecido (TCLE) de acordo com a Resolução 510/2016 (Anexo B). O acesso ao histórico pessoal e aos dados clínicos referentes ao diagnóstico de DM2 foi obtido pela análise das respostas de um questionário aplicado a todos os participantes (Anexo C) e pelos dados da análise bioquímica de exames de glicose, hemoglobina glicada, insulina, colesterol total e triglicérides (Apêndice A). Os participantes apresentavam idade entre 18 a 80 anos e dependendo do histórico pessoal e clínico de cada um deles, pertenceram a um dos grupos estabelecidos do estudo.

Neste estudo foram utilizadas amostras de sangue periférico obtidas de 67 participantes onde foram realizadas as seguintes análises: análises bioquímicas realizadas no sangue total em 67 amostras; genotipagem pela PCR-RFLP no DNA extraído da papa leucocitária obtida do sangue total em 52 amostras; espectroscopia por ATR-FTIR nas amostras de soro obtidas do sangue total dos pacientes em 59 amostras e as análises de ATR-FTIR aliada ao algoritmo *machine learning* realizada tanto nas amostras de soro obtidas do sangue total em 59 amostras quanto nas amostras de DNA amplificado pela PCR em 52 amostras (Apêndice B).

O procedimento operacional padrão para a coleta de sangue venoso foi realizado conforme as recomendações da Sociedade Brasileira de Patologia Clínica (SBPC). Foram coletados e identificados quatro tubos de sangue, cada um com volume entre 3 a 4 mL, por punção venosa. Dois tubos continham ácido etilenodiamino tetra-acético (EDTA) (tampa roxa),

sendo um destinado à extração de DNA para a realização das análises moleculares e o outro utilizado para análises bioquímicas (hemoglobina glicada). Os outros dois tubos, sem anticoagulante (tampa amarela), foram utilizados para as demais análises: um tubo foi destinado as outras análise bioquímicas (glicose sérica, insulina, colesterol total e triglicerídeos), enquanto o outro foi utilizado para a separação do soro, posteriormente analisado por espectroscopia ATR-FTIR.

Primeiramente, após a realização das análises bioquímicas, foi realizada a análise de espectroscopia ATR-FTIR nas amostras de soro obtidas de 59 participantes distribuídos em dois grupos de acordo com a concentração de glicose, o grupo A constituído por 40 indivíduos com glicemia menor a 100 mg/dL e hemoglobina glicada abaixo de 6,5%, e o o grupo B constituído por 19 indivíduos com glicemia superior a 100 mg/dL e hemoglobina glicada superior a 6,5%. Esta análise tem o objetivo de separar os indivíduos por meio da concentração glicêmica. Vinte e duas amostras não foram consideradas nesta análise pois os pacientes tomavam medicamentos para o controle da diabetes, o que pode levar a classificação errônea dos grupos que foi realizada pelo nível de glicose no sangue. Na amostra de soro de 59 indivíduos também foi realizada a análise de FTIR aliada ao *machine learning* seguindo a mesma divisão dos grupos A e B.

Na segunda etapa do estudo, os participantes foram classificados com base em seus genótipos do SNP -866G/A do *gene UCP2*, obtidos previamente pela técnica de PCR-RFLP, em 52 amostras de produto amplificado pela PCR para posterior análise de ATR-FTIR aliada ao algoritmo *machine learning*. O grupo 1 incluiu 18 indivíduos genotipados pela PCR-RFLP como AA, o grupo 2 incluiu 21 indivíduos que apresentam o genótipo G/A e o grupo 3 foi constituído por 13 indivíduos que apresentam o genótipo G/G.

4.2 ANÁLISE BIOQUÍMICA

Para a realização da análise bioquímica o sangue coletado foi centrifugado a 3.500 rpm durante 15 minutos, utilizando uma centrífuga (HITACHI – CF16RN) para posicionar o gel dentro do tubo de coleta no centro, separando assim as células sanguíneas da fração líquida (soro).

A análise bioquímica foi conduzida no equipamento LAB MAX PLENNO, da marca LABTEST, utilizando a metodologia enzimática colorimétrica. As análises incluíram glicemia (Glicose oxidase PAP Liquiform, ref. 84 - Labtest), HbA1c (Imunoensaio turbidimétrico de

inibição), insulina (Eletroquimioluminescência), colesterol total (Colesterol liquiform, ref.76 - Labtest) e triglicérides (Triglicérides liquiform, ref.87 – Labtest). Os parâmetros laboratoriais de referência seguiram os valores adotados rotineiramente pelo CPS (Centro de Práticas Supervisionadas – UNIVAP), conforme detalhado na Tabela 2.

Tabela 2 - Parâmetros laboratoriais utilizados para dosagem de glicose, hemoglobina glicada, insulina, colesterol total e triglicérides.

Exame Laboratorial	Valor de normalidade (em ambos os sexos)
Glicose	70 a 99 mg/dL
Hemoglobina Glicada	Inferior a 5,7 %
Insulina	De 2,6 a 24,9 uUI/mL
Colesterol Total	Superior a 19 anos de idade: até 190 mg/dL 2 a 19 anos de idade: até 170 mg/dL
Triglicérides	Superior a 19 anos de idade: até 175 mg/dL 2 a 19 anos de idade: até 100 mg/dL

Fonte: Autora (2024).

4.3 GENOTIPAGEM

4.3.1 Extração do DNA

A extração do DNA das amostras de sangue foi realizada pelo método fenol:clorofórmio (Kirby, 1957). Após a coleta do sangue total, as amostras de sangue foram centrifugadas e, após a separação das fases, o plasma sanguíneo e as hemácias foram descartados, enquanto a papa leucocitária foi transferida para um novo tubo. Na papa leucocitária, foi adicionada a solução de tampão Hemálise-RBC 1X para degradar as hemácias excedentes, e os tubos foram levados ao vórtex para agitação constante, sendo posteriormente incubados à temperatura ambiente, protegidos da luz. Os tubos foram centrifugados novamente, e o sobrenadante foi descartado, mantendo-se apenas o pellet. Essa etapa foi repetida tantas vezes quanto necessário até que o pellet ficasse com pouca ou nenhuma presença de hemácias.

No pellet, foram adicionadas soluções de acetato de sódio 0,2M para estabilização, SDS 10% para remover as camadas lipídicas, e proteinase K para degradação de proteínas, sendo então incubados a 37°C durante a noite. Após a digestão enzimática, o sobrenadante foi descartado, e a solução de fenol: clorofórmio: álcool isoamílico, na proporção de 25:24:1, foi adicionada para remover componentes celulares indesejáveis, eliminar reagentes do processo e

evitar a formação de espuma durante a agitação no vórtex. As amostras foram centrifugadas, e os sobrenadantes transferidos para um novo tubo. Essa etapa e outras com apenas clorofórmio foram repetidas, se necessário.

Em seguida, foi adicionado álcool etílico absoluto gelado ao sobrenadante, promovendo a precipitação do DNA, que pôde ser observado na forma de pellet no fundo do tubo Eppendorf. As amostras foram armazenadas a -20°C durante a noite e, em seguida, o álcool foi removido, e o pellet foi seco em um dessecador a vácuo. Após essa etapa, o pellet foi ressuscitado em solução tampão TE ou água livre de nucleases. Se necessário, uma nova etapa de precipitação e lavagem do DNA foi realizada

4.3.2 Análise da integridade e qualidade do DNA

A quantificação e a qualidade do DNA de cada amostra foram avaliadas por espectroscopia de absorção ultravioleta, utilizando o equipamento NanoDrop (ND-1000 Spectrophotometer v.3.0.1, Labtrade) e pela eletroforese em gel de agarose a 1%, respectivamente. As razões 260/280 e 260/230 foram analisadas para verificar contaminação por proteínas e reagentes, considerando que o DNA absorve luz a 260 nm, as proteínas a 280 nm, e os reagentes a 230 nm. As razões consideradas aceitáveis variaram entre 1,8 e 2,0 para a primeira, e entre 1,7 e 2,2 para a segunda (Apêndice D).

Para a avaliação da integridade, foi utilizado aproximadamente 500ng de DNA para a eletroforese em gel de agarose a 1,0% em Tris-Borato-EDTA 1X (TBE) corado com brometo de etídeo, reagente no qual se intercala entre as bases nitrogenadas do DNA permitindo a visualização da banda de DNA genômico total, de acordo com o marcador de peso molecular de 100pb (*Invitrogen*).

A partir dos valores obtidos, foi possível avaliar a adequação do processo de extração, garantindo a integridade e pureza das amostras de DNA. A ausência ou baixa concentração de proteínas e reagentes, bem como a não degradação do DNA, indicam que o procedimento de extração foi realizado corretamente, assegurando a qualidade das amostras para análises subsequentes.

4.3.3 PCR do gene UCP2

A PCR foi utilizada para amplificar uma região de 363 pares de bases (pb) contendo o SNP rs659366 (-866G/A), localizado na região promotora do *gene UCP2*. Os iniciadores que flanqueiam este SNP possuem as seguintes sequências: 5'-CACGCTGCTTCTGCCAGGAC-3' (*forward*) e 5'-AGGCGTCAGGAGATGGACCG-3' (*reverse*). A reação de PCR foi realizada em um termociclador (Veriti™ 96-Well Thermal Cycler) com um volume final de 25 µL, contendo 100 ng de DNA. A mistura de reação incluiu 50 mM de KCl, 10 mM de Tris-HCl (pH 8,3), 1,5 mM de MgCl₂, 200 µM de dNTP, 0,6 µM de cada primer e 2U/uL de Taq DNA Polimerase.

O protocolo de amplificação consistiu em uma desnaturação inicial de 5 minutos a 94°C, seguida por 30 ciclos de 1 minuto a 94°C, anelamento por 30 segundos a 70°C, e extensão de 1 minuto a 72°C, com uma extensão final de 6 minutos a 72°C.

Os produtos amplificados foram analisados em eletroforese em gel de agarose a 1,6% eo marcador de peso molecular utilizado foi de 100 pb. A obtenção dos resultados foi realizada por meio da visualização dos produtos de DNA amplificados em forma de banda no gel de agarose após a eletroforese. A incidência da luz ultravioleta (UV), que permite a visualização das bandas no gel, foi controlada pelo equipamento Transilluminator-D Pro MiniBIS (DNR Bio-Imaging Systems Ltd) e as imagens foram capturadas com o software GelCapture para a análise posterior.

4.3.4 Análise de PCR-RFLP

Após a amplificação da sequência de DNA contendo o SNP -866G/A, o produto amplificado foi digerido pela enzima de restrição *MluI*, que reconhece a sequência alvo e rompe as ligações fosfodiéster do DNA no sítio A↓CGCGT. A solução para digestão enzimática foi preparada utilizando 5 µL de produto de PCR, 12 µL de H₂O, 1,5 µL de tampão 10X e 1,5 µL de enzima de restrição, e a mistura foi incubada a 37°C por 1 hora.

A identificação do SNP de cada amostra foi realizada por meio da análise da atividade enzimática, com a clivagem ou não do produto amplificado pela enzima MLUI visualizada em gel de agarose a 1,6% sob luz ultravioleta, utilizando o equipamento Transilluminator-D Pro MiniBIS (DNR Bio-Imaging Systems Ltd).

A presença do sítio de restrição para a enzima *MluI* nos dois alelos, ou seja, indivíduo homocigoto para a presença do alelo A, resultará em dois fragmentos (duas bandas no gel) de 295 pb e 68 pb visualizados no gel de agarose após a PCR-RFLP e eletroforese. A ausência do sítio de restrição nos dois alelos, indivíduo homocigoto para a presença do alelo G, resultará em apenas um fragmento de 363 pb, pois não ocorrerá a digestão pela enzima. A presença do sítio de restrição em um alelo e a ausência no outro alelo resultará em três fragmentos de 295pb, 68pb e 363pb, formando três bandas no gel de agarose. Assim, o indivíduo homocigoto apresentará o genótipo AA, o indivíduo heterocigoto o genótipo GA, e o indivíduo homocigoto apresentará o genótipo GG. As bandas visualizadas no gel foram analisadas, correspondendo aos diferentes tipos de polimorfismos AA, GA ou GG.

4.4 ANÁLISE DA GENOTIPAGEM VS BIOQUÍMICA

Os genótipos GG, GA e AA do SNP -866G/A no *gene UCP2* foram determinados por meio da técnica de PCR-RFLP. Em seguida, os dados bioquímicos, incluindo glicose de jejum, hemoglobina glicada, insulina, triglicérides e colesterol total, foram comparados entre os grupos genotípicos. As médias e os desvios padrão foram calculados para cada variável, estratificados pelos genótipos, de forma a descrever as características metabólicas de cada grupo.

Para avaliar a significância estatística das diferenças entre os grupos genotípicos, foram aplicados testes estatísticos apropriados. Inicialmente, utilizou-se o teste ANOVA (Análise de Variância) para verificar a existência de diferenças nas variáveis quantitativas entre os grupos. Posteriormente, para identificar quais grupos apresentavam diferenças significativas entre si, foi aplicado o teste post hoc de Tukey. Ambos os testes consideraram um nível de significância de $p < 0,05$.

Os dados obtidos foram organizados utilizando o programa Microsoft Office Excel®. Para comparação dos resultados, foi realizado ANOVA com teste de Tukey ($p < 0,05$), realizado com o software SISVAR versão 5.6 (Ferreira, 2014; Ferreira, 2019).

4.5 ANÁLISE DE ATR-FTIR

4.5.1 Amostras

Neste estudo, como já descrito, as análises de ATR-FTIR foram realizadas em dois tipos de amostras: amostras obtidas do sangue total, onde o produto utilizado foi o soro, e amostras de DNA amplificadas previamente pela PCR (amplicon) do SNP rs659366 (-866G/A) do *gene UCP2*. No primeiro caso, foi realizada a análise de ATR-FTIR em 59 participantes, sendo 40 do grupo A com glicemia menor a 100 mg/dL e hemoglobina glicada abaixo de 6,5%, e 19 participantes no grupo B com glicemia superior a 100 mg/dL e hemoglobina glicada superior a 6,5%. Além disso, foi realizada a análise aplicando-se o algoritmo machine learning (escrever os grupos e em quantas amostras). No segundo caso, um total de 52 amostras foi utilizada para a análise de genotipagem, sendo 13 amostras homozigotas GG, 21 amostras heterozigotas GA e 18 amostras homozigotas AA.

Após a obtenção das amostras de soro e as amostras de DNA amplificado (amplicon), estas foram armazenadas em freezer a -80°C e -20°C , respectivamente, onde depois de coletadas, as amostras de soro foram separadas da parte coagulada do sangue em microtubos de 2 ml, e o amplicon em microtubos de 0,2 ml. Posteriormente, estas foram avaliadas no equipamento ATR-FTIR.

4.5.2 Obtenção dos espectros

O equipamento para obtenção dos espectros foi o espectrofotômetro FTIR (*Spotlight 400 FTIR Imaging System, Perkin Elmer*) acoplado a unidade de ATR (superfície de diamante) (*Attenuated Total Reflection*) com uma resolução de 4 cm^{-1} . As amostras do soro foram descongeladas simultaneamente em temperatura ambiente, posicionando $3\text{ }\mu\text{L}$ de cada amostra sob a superfície do diamante, e secas completamente utilizando um jato de ar ultra puro para garantir a secagem completa. Para obtenção de leituras espectrais precisas, foram realizadas 10 varreduras pelo equipamento de espectroscopia, onde a média resultava no espectro final.

Os procedimentos para a análise de FTIR com as amostras do amplicon foram praticamente os mesmos, com exceção que as mesmas foram descongeladas unidade por unidade, evitando assim a contaminação da amostra. O *backgrund* de 10 *scans* foi realizado antes de cada análise.

4.5.3 Análise de cluster hierárquico para separação dos grupos em relação aos níveis de glicose

Nos 59 espectros obtidos das amostras de soro, separadas em grupo A com 40 espectros (glicose abaixo de 100mg/dL e HbA1c abaixo de 6,5%) e grupo B com 19 espectros (glicose acima de 100mg/dL e HbA1c acima de 6,5%) foi realizado um tratamento prévio com o objetivo de garantir a qualidade e consistência dos dados espectrais.

Inicialmente, *outliers* (valores atípicos) foram identificados e removidos para minimizar interferências que poderiam prejudicar a análise. Em seguida, aplicou-se um filtro Savitzky-Golay de nove pontos para o cálculo das derivadas de primeira e segunda ordem dos espectros, o que permitiu aprimorar a resolução das bandas espectrais e corrigir deslocamentos da linha de base. Esses passos foram essenciais para garantir a integridade e comparabilidade dos dados.

Os espectros tratados foram então submetidos à análise de cluster hierárquica (HCA), utilizando o algoritmo de Ward para agrupar amostras com base na similaridade espectral. A métrica de distância foi obtida pelo método *scaling to the first range of scale* (escalonamento para o primeiro intervalo de escala), o que permitiu padronizar a escala e otimizar a discriminação entre as amostras. Todo o processo de processamento e análise dos espectros foi realizado no software OPUS 8.0 (Bruker Optik GmbH, Alemanha).

4.5.4 ATR-FTIR machine learning para separação dos grupos em relação aos níveis de glicose

Os 59 espectros, 40 do grupo A e 19 do grupo B obtidos pela análise de ATR-FTIR foram convertidos para *arrays NumPy* para melhor aplicação no algoritmos de aprendizagem. A técnica SMOTE (Técnica de Sobre amostragem Sintética de Minorias) foi aplicada para gerar amostras sintéticas da classe minoritária. Isso garantiu um balanceamento adequado das classes e evitou que o modelo apresentasse viés para a classe majoritária durante o treinamento. Em seguida, os dados foram divididos aleatoriamente em dois subconjuntos: 80% para treinamento e 20% para teste, utilizando a função *train-test-split* com uma semente aleatória (*Random-state*) para assegurar a reprodutibilidade dos resultados. Os dados foram normalizados e padronizados utilizando o *Standard Scaler*, garantindo que todas as características tivessem média 0 e desvio padrão 1, o que é essencial para a estabilidade do treinamento do modelo.

A arquitetura do modelo de Rede Neural Artificial (ANN) foi definida com várias camadas densas (fully connected), pois esse tipo de arquitetura é eficaz para capturar relações complexas em dados. A camada de entrada, com 128 neurônios, utiliza a função de ativação ReLU (Unidade Linear Retificada), que permite a aprendizagem de padrões complexos ao ativar apenas valores positivos e introduzir não-linearidade ao modelo. As camadas intermediárias, com 64 e 32 neurônios e ativação ReLU, reduzem progressivamente a dimensionalidade dos dados, o que ajuda o modelo a extrair representações relevantes de maneira eficiente. Foram incluídas camadas de Dropout, com uma taxa de 50%, que desligam aleatoriamente metade das unidades durante o treinamento, sendo essenciais para evitar o overfitting, situação em que o modelo memoriza os detalhes e ruídos dos dados de treinamento e perde a capacidade de generalizar para novos dados.

Além disso, foi aplicada regularização L2 em todas as camadas densas, pois ela penaliza pesos elevados, limitando a complexidade do modelo e reduzindo ainda mais o risco de sobreajuste.

A camada de saída é composta por um único neurônio com ativação sigmoide, ideal para problemas de classificação binária. Essa função transforma as saídas em valores entre 0 e 1, permitindo interpretar as respostas como probabilidades de pertencimento à classe positiva, o que facilita a classificação. O modelo foi compilado com o otimizador Adam, que se destaca por sua capacidade de adaptar a taxa de aprendizado durante o treinamento, acelerando a convergência. A taxa de aprendizado inicial foi configurada em 0,001 para equilibrar estabilidade e eficiência. A função de perda binary-crossentropy foi utilizada por ser apropriada para medir o erro em problemas de classificação binária, enquanto a acurácia foi definida como métrica principal para avaliar a proporção de previsões corretas durante o treinamento.

Para garantir uma avaliação robusta e evitar vieses, foi aplicada a técnica de K-Fold Cross-Validation com cinco divisões. Essa abordagem permite que o modelo seja treinado e validado em todas as amostras, utilizando partes diferentes dos dados para validação em cada iteração, o que melhora a confiabilidade das métricas de desempenho. Durante cada iteração do K-Fold, o modelo foi treinado por até 100 épocas e os históricos de perda e acurácia foram registrados, o que auxilia no monitoramento do desempenho do modelo ao longo do treinamento e permite identificar a iteração com melhor desempenho. Para unificar os resultados, os históricos foram alinhados com base no menor número de épocas entre as iterações e suas médias foram calculadas, oferecendo uma visão consolidada do desempenho geral do modelo.

Após selecionar o melhor modelo durante a validação, ele foi avaliado em um conjunto de teste contendo dados inéditos. As previsões foram classificadas com base em um limiar de 0,5, sendo que valores iguais ou maiores foram atribuídos à classe positiva (1), enquanto valores menores foram atribuídos à classe negativa (0). Esse processo foi fundamental para verificar a capacidade do modelo de generalizar para novos dados e garantir sua eficácia tanto no treinamento quanto no teste (Rufino *et al.*, 2016, Tomas *et al.*, 2022). Todo o processamento de *machine learning* foi realizado no Google Colaboratory.

4.5.5 ATR-FTIR machine learning do polimorfismo -866G/A do gene UCP2

As 52 amostras de DNA amplificado, classificadas em 13 amostras homozigotas GG, 21 amostras heterozigotas GA e 18 amostras homozigotas AA, foram analisadas utilizando técnicas de aprendizado de máquina. Inicialmente, os dados espectrais foram normalizados aplicando a técnica *Standard Normal Variate* (SNV), que elimina variações sistemáticas indesejadas e padroniza os espectros, garantindo a comparabilidade entre as amostras.

Para reduzir a dimensionalidade dos dados e destacar as variáveis espectrais mais relevantes, foi utilizada a Análise de Componentes Principais (PCA). Essa técnica identificou os componentes principais mais informativos, permitindo representar as características espectrais de forma compacta e eficiente, reduzindo o ruído e otimizando a etapa subsequente de classificação.

Na etapa de classificação, diversos algoritmos de aprendizado de máquina foram empregados para desenvolver modelos robustos, capazes de distinguir os genótipos com alta precisão. O modelo *K-Nearest Neighbors* (KNN) foi testado utilizando diferentes métricas de distância, incluindo cosine, cubic e fine, para avaliar qual abordagem proporcionava melhor desempenho. Paralelamente, o modelo Máquina de Vetor de Suporte (SVM) foi explorado com diferentes funções kernel, como linear, cúbico e quadrático, a fim de identificar a configuração mais adequada para a classificação dos genótipos.

Os modelos foram treinados com os componentes principais extraídos pela PCA e validados por meio da técnica de validação cruzada *leave-one-out* (LOO-CV). Essa abordagem consiste em iterar sobre todas as amostras, removendo uma amostra do conjunto de treinamento e utilizando-a como teste, até que todas as amostras sejam avaliadas. Esse método assegura uma avaliação rigorosa do desempenho do modelo, garantindo sua robustez e capacidade de generalização.

5 RESULTADOS

5.1 CARACTERÍSTICAS CLÍNICAS

Das 67 amostras obtidos em nosso estudo, 64,43% são do sexo feminino e 33,33% do sexo masculino, com idade média de 49,52 anos. Do total dos 67 participantes, 24,69% já haviam sido diagnosticados com DM2 antes da pesquisa, 67,90% não se declararam diabéticos ou pré-diabéticos, e 7,41% relataram ser pré-diabéticos. Em relação ao histórico familiar, 70,37% dos participantes confirmaram que pai, mãe, irmãos ou avós apresentam DM2 e 29,63% não possuem familiares próximos com o diagnóstico da doença. Dos 19 participantes que compuseram o grupo experimental, 52,63% responderam não ter diabetes tipo 2 (DM2) ou pré-diabetes no questionário, enquanto 31,58% relataram ser portadores de DM2 e 15,79% se identificaram como pré-diabéticos

Em relação ao estilo de vida foi levantado que 87,65% dos participantes não fumam, 7,41% são fumantes e dois indivíduos relataram ter fumado no passado, mas não fumam mais atualmente. Quanto ao consumo de bebidas alcoólicas, 87,65% afirmaram não consumir nenhum tipo de álcool, enquanto 10 participantes possuem o hábito de tomar bebida alcoólica.

Sobre a prática de atividade física e hábitos alimentares, 59,26% declararam realizar atividades físicas regularmente, enquanto 40,74% não praticam. Em relação à alimentação, 82,72% afirmaram seguir uma dieta saudável com inclusão de frutas, legumes e verduras, enquanto 17,28% indicaram não adotar esse padrão alimentar.

No que diz respeito aos dados antropométricos, 22,22% dos participantes apresentaram eutrofia, com IMC abaixo de 24,9 kg/m²; 28,40% estavam em sobrepeso, com IMC entre 25 e 29,9 kg/m²; e 49,38% foram classificados como obesos, com IMC acima de 30 kg/m² (Apêndice C).

Na avaliação bioquímica, 25,93% apresentaram glicose acima de 100 mg/dL, 17,28% tinham hemoglobina glicada (HbA1c) superior a 6,5%, 16,05% apresentaram colesterol total acima de 200 mg/dL e 20,99% tiveram triglicerídeos acima de 150 mg/dL, valores que ultrapassam os limites de referência estabelecidos.

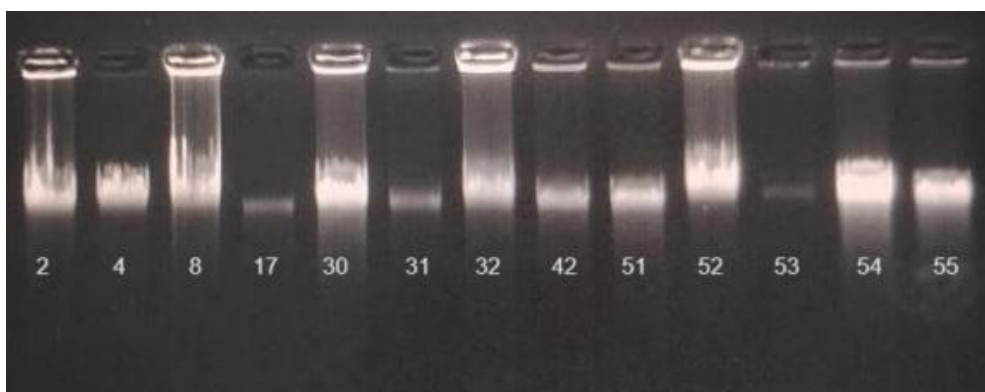
5.2 GENOTIPAGEM

No total, das 67 amostras, 67 amostras de sangue passaram pela extração do DNA, porém 52 delas foram utilizadas para a etapa de genotipagem do polimorfismo -866G/A do

gene UCP2, onde a média das concentrações de quantificação do DNA por espectroscopia de absorção do ultravioleta do grupo genotipado foi de 662,16 ng/ μ L. A média de absorbância 260/280 obtida no total de amostras de DNA analisadas, que é o parâmetro indicativo da qualidade do DNA em relação a contaminação por proteínas, foi de 1,82. Este resultado está dentro da faixa recomendável de 1,8 a 2,0, sugerindo que as amostras de DNA obtidas são de ótima qualidade, com mínima contaminação proteica. A média de absorbância 260/230, que é o parâmetro indicativo da qualidade do DNA em relação a a presença de contaminantes como carboidratos, solventes orgânicos ou reagentes utilizados na extração, foi de 1,69. Este valor está ligeiramente abaixo do intervalo ideal de 2,0 a 2,2, o que indica uma possível presença de alguns contaminantes que podem interferir na qualidade do DNA obtido. No entanto, o valor é considerado aceitável e que não interferirá na análise genotípica, considerando que a análise é pontual realizada em alguns SNPs específicos (Apêndice D).

Na figura 1 é possível observar a alta integridade dos DNAs extraídos de algumas amostras por meio da visualização das bandas de DNA genômico total em eletroforese em gel de agarose a 1% corado com brometo de etídio. As bandas resultantes mostram uma distribuição consistente de DNA, com presença de bandas nítidas e intensas, indicando que as amostras estão íntegras e sem degradação significativa (Figura 1).

Figura 1 - Amostras de DNA visualizadas pelas bandas de DNA genômico total em eletroforese em gel de agarose a 1% corado com brometo de etídio.



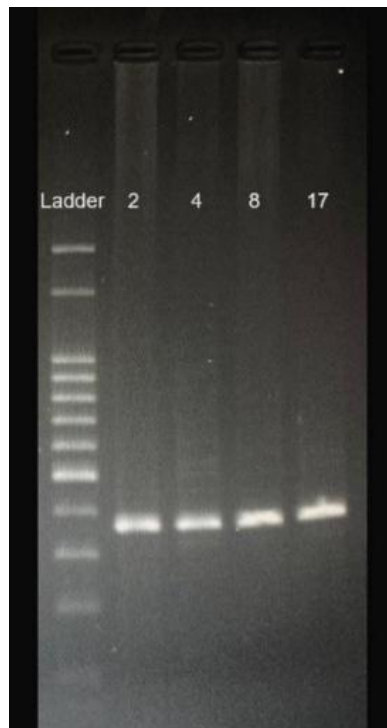
Legenda: amostra de DNA extraído da papa leucocitária do sangue total

Fonte: Autora (2024).

Na figura 2, observa-se a presença de bandas distintas que correspondem aos produtos amplificados. O gel inclui padrão de peso molecular (*Ladder* 100 pb) à esquerda, que foi utilizada como referência para confirmar o tamanho esperado dos fragmentos amplificados.

Na figura 2 está apresentada a eletroforese em gel de agarose da amplificação de sequência de DNA que contém o SNP -866G/A do *gene UCP2* de quatro amostras de DNA do sangue dos pacientes, mostrando o alto rendimento da reação de PCR. Vale ressaltar que a sensibilidade da PCR foi extremamente alta, sendo capaz de amplificar a região flanqueada pelos *primers* mesmo em amostras com concentrações mais baixas de DNA, como a amostra 10 que apresentou apenas 41,2 ng/μL de DNA. A intensidade homogênea das bandas mostra que a eficiência da amplificação foi consistente entre as diferentes amostras, indicando a alta reprodutibilidade da reação de PCR realizada nas 52 amostras do estudo.

Figura 2 - Eletroforese em gel de agarose (1,6%) com marcador de peso molecular 100pb (Ladder) seguido da amplificação de sequência de DNA que contém o SNP -866G/A do gene *UCP2* das amostras de DNA obtida do sangue total.



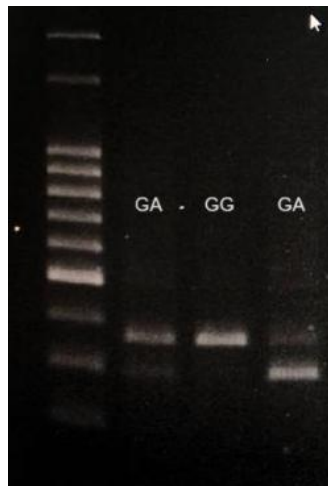
Legenda: ampliação do gene *UCP2*

Fonte: Autora (2024).

A análise do polimorfismo -866G/A do *gene UCP2* realizada por PCR-RFLP em 52 amostras identificou os três tipos de genótipos. A presença do sítio de restrição da enzima *MluI* em ambos os alelos do *gene UCP2* (genótipo homocigoto AA) foi observada em 18 amostras, detectada pela presença de dois fragmentos (295 pb e 68 pb), correspondentes aos dois alelos do gene, visualizados após eletroforese em gel de agarose.

A ausência do sítio de restrição em ambos os alelos do *gene UCP2* (genótipo homozigoto GG) foi observada em 13 amostras, identificada pela presença de um único fragmento de 363 pb em ambos os alelos. Os genótipos heterozigotos (genótipo GA), caracterizados pela presença do sítio de restrição em um alelo e sua ausência no outro, foram observados em 21 amostras. Esse genótipo foi detectado pela presença de três fragmentos: uma banda de 363 pb, relativa a um alelo, e duas bandas de 295 pb e 68 pb, referentes ao outro alelo (Apêndice D).

Figura 3 - Gel de agarose com ladder de padrão de peso molecular de 100pb e três amostras digeridas de PCR-RFLP sendo a primeira GA a segunda GG e a terceira GA do polimorfismo -866G/A do gene UCP2.



Fonte: Autora (2024).

Os resultados obtidos demonstram a eficiência da técnica de PCR-RFLP para a genotipagem do polimorfismo -866G/A no *gene UCP2*, evidenciando a capacidade de distinguir com precisão os genótipos homozigotos AA e GG, bem como os heterozigotos GA. A alta integridade e qualidade do DNA extraído, aliada à consistência dos produtos amplificados e digeridos visualizados em gel de agarose, reforça a confiabilidade do protocolo utilizado.

A distribuição equilibrada dos genótipos entre as amostras genotipadas é consistente com estudos populacionais prévios, e a identificação dos fragmentos característicos de cada genótipo confirma a robustez da técnica.

5.3 Análise estatística: bioquímica vs molecular

Na comparação entre os diferentes genótipos do SNP -866G/A (AA, GA e GG) e cada um dos índices bioquímicos avaliados (insulina, glicose, hemoglobina glicada (HbA1c), colesterol total, triglicérides). A análise estatística realizada pelo teste ANOVA indicou que há uma diferença significativa nos níveis médios de insulina entre os três grupos de genótipos ($F = 4,191$, $p = 0,0205$). A análise realizada revelou diferença estatística significativa ($p < 0,05$) onde o valor de F refletiu a razão entre a variação média das diferenças entre os grupos e a variação dentro de cada grupo. As médias de insulina foram de $11,00 \mu\text{U/mL}$ para o grupo GG, $14,48 \mu\text{U/mL}$ para o grupo AA e $19,75 \mu\text{U/mL}$ para o grupo GA, sugerindo que os genótipos influenciam os níveis de insulina. Para identificar o grupo que apresentou a diferença estatística, foi realizado o teste de Tukey, onde o grupo GA apresentou uma média de nível de insulina significativamente maior quando comparado aos grupos com os genótipos GG e AA.

Em contrapartida, na avaliação das médias obtidas para os valores da análise de glicose, hemoglobina glicada, colesterol e triglicérides não foram observadas diferenças estatisticamente significativas na comparação entre os grupos de acordo com os diferentes genótipos (AA, GA e GG) do SNP -866G/A do gene *UCP2*. As médias das concentrações das análises bioquímicas estão descritas abaixo e na tabela 3.

Tabela 3 - Média das concentrações bioquímicas de glicose, hemoglobina glicada, insulina, colesterol total e triglicérides de acordo com o genótipo do SNP -866G/A do gene *UCP2*.

Genótipo	Glicose (mg/dL)	HbA1c (%)	Insulina (uUI/mL)	Colesterol Total (mg/dL)	Triglicérides (mg/dL)
G/G	95,96	5,86	11,00	163,06	123,50
A/A	96,01	5,67	14,47	177,67	104,22
G/A	106,35	6,14	19,74	172,52	126,33

Legenda: HbA1c: hemoglobina glicada

Fonte: Autora (2024).

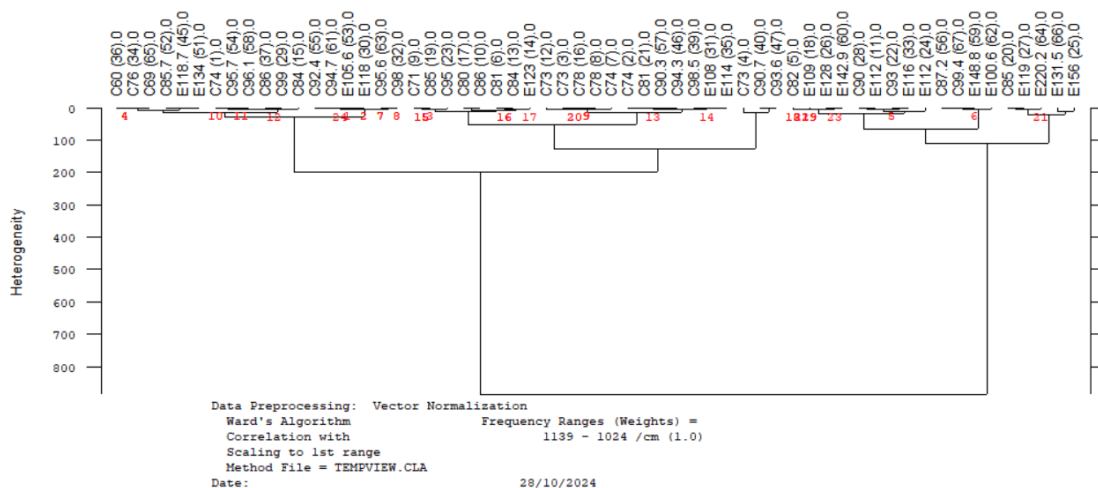
Esses resultados sugerem que, embora haja variações nas médias, não existem diferenças significativas entre esses grupos para essas variáveis quando realizado o teste ANOVA. As médias obtidas da análise de glicemia foram de $95,96 \text{ mg/dL}$ no grupo GG, $96,01 \text{ mg/dL}$ no grupo AA e $106,35 \text{ mg/dL}$ no grupo GA resultou em um $p = 0,4629$, não apresentando significância estatística, assim como nas médias de hemoglobina glicada foram $5,67\%$ para AA, $5,86\%$ para GG e $6,14\%$ para GA onde $p = 0,5029$. Com relação ao colesterol, as médias foram

163,06 mg/dL para o genótipo GG, 172,52 mg/dL para o genótipo GA e 177,67 mg/dL para o genótipo AA com resultado de $p = 0,5998$. Os valores das médias para a análise dos triglicerídeos foram 104,22 mg/dL para o genótipo AA, 123,50 mg/dL para o genótipo GG e 126,33 mg/dL para o genótipo GA e $p = 0,5541$.

5.4 Análise de cluster hierárquico para concentração de glicose

A análise de cluster foi aplicada para avaliar os níveis de glicose sérica dos 59 pacientes, sendo 40 deles com concentração de glicose abaixo de 100mg/dL e com HbA1c menor que 6,5% e 19 deles com concentração de glicose acima de 100mg/dL e HbA1c acima de 6,5%. As amostras foram agrupando com base nas similaridades dos dados espectrais obtidos a partir do soro sanguíneo. Os resultados indicaram um desempenho significativo, com uma acurácia de 78% na faixa espectral de 1139-1024 cm^{-1} , mostrando que o modelo foi eficiente em separar pacientes com diferentes concentrações de glicose (figura 4).

Figura 4 - Dendrograma resultante da análise de cluster utilizando o algoritmo de Ward, aplicado aos dados espectrais do soro sanguíneo na faixa de 1139-1024 cm^{-1} , após normalização vetorial.



Fonte: Autora (2024).

5.5 ATR-FTIR aliada ao *machine learning* na concentração de glicose

O modelo de rede neural desenvolvido apresentou uma acurácia de 83% no conjunto de teste, evidenciando sua capacidade de realizar previsões precisas na tarefa de classificação binária. O relatório de classificação revelou que a performance foi equilibrada entre as classes, com uma precisão de 83% tanto para a classe 0 quanto para a classe 1. No entanto, o comportamento em termos de *recall* variou entre as classes: a classe 0 alcançou 71% de *recall*, enquanto a classe 1 atingiu 91%, indicando que o modelo teve maior facilidade para identificar corretamente as instâncias da classe 1.

O F1-score foi de 77% para a classe 0 e 87% para a classe 1, refletindo um desempenho mais consistente para a segunda classe. O valor médio das métricas, calculado como macro average, apresentou 82% para F1-score e 81% para *recall*, sugerindo que o modelo lidou bem com a distribuição das classes. A média ponderada, que considera o número de amostras em cada classe, mostrou valores de 83% tanto para precisão quanto para F1-score e acurácia, reforçando que o desempenho foi consistente no geral.

Tabela 4 - Desempenho do modelo de rede neural artificial na classificação da concentração de glicose.

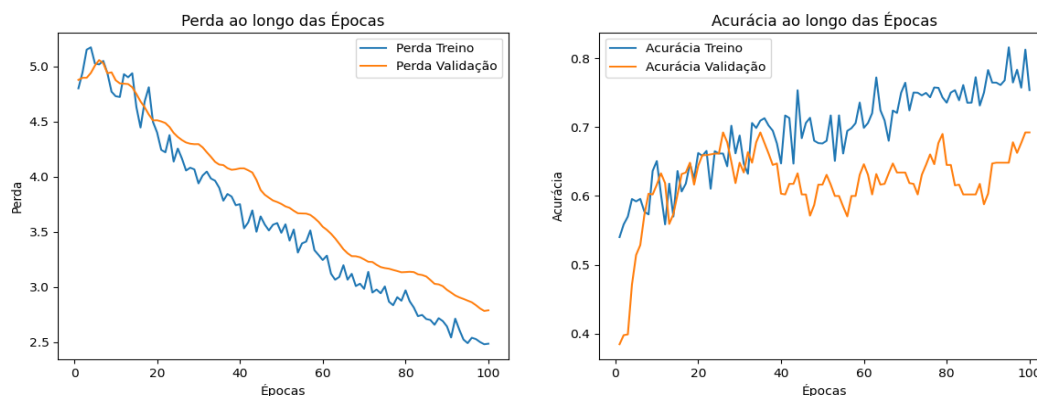
	<i>Precisão</i>	<i>Recall</i>	<i>F1-score</i>	<i>Suporte</i>
Grupo B	0.75	0.86	0.80	7
Grupo A	0.90	0.82	0.86	11
Acurácia			0.83	18
Macro média	0.82	0.84	0.83	18
Média ponderadaa	0.84	0.83	0.83	18

Legenda: Classe 0 referente ao grupo experimental, portanto o grupo B e a classe 1 representada pelo grupo controle, sendo assim o grupo A.

Fonte: Autora (2024).

Os gráficos gerados durante o treinamento e validação indicaram que as curvas de perda e acurácia se comportaram de maneira próxima, com uma diminuição progressiva na perda e aumento na acurácia ao longo das épocas. A convergência dos valores entre treino e validação sugere que o modelo foi capaz de generalizar bem, minimizando o risco de *overfitting*. A aplicação de técnicas como *Dropout* e regularização L2 foi essencial para evitar sobre ajuste, enquanto o uso do SMOTE garantiu que ambas as classes fossem representadas de forma justa durante o treinamento.

Figura 5 - Evolução da perda e da acurácia durante o treinamento e validação ao longo das épocas.



Fonte: Autora (2024).

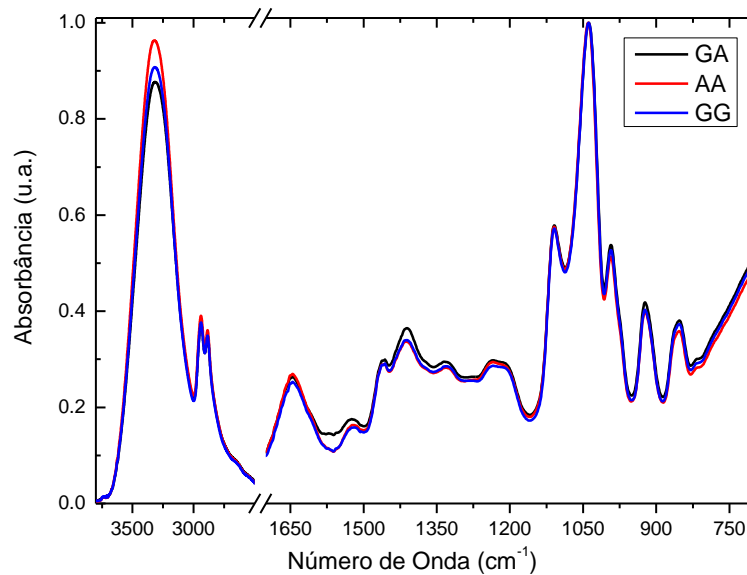
O modelo se mostrou eficiente em capturar padrões relevantes nos dados espectrais e utilizar esse conhecimento para classificar corretamente novas amostras. Durante o treinamento, o modelo passou por múltiplas épocas, que são ciclos completos em que o algoritmo processa todos os dados de treinamento e ajusta seus parâmetros, refinando gradualmente sua capacidade de identificar padrões complexos. A combinação das abordagens aplicadas permitiu que o modelo equilibrasse as métricas entre precisão, *recall* e F1-score, resultando em um desempenho robusto e confiável, com destaque para sua capacidade de identificar corretamente a classe 1.

5.6 ATR-FTIR aliada ao *machine learning* do polimorfismo -866G/A do gene *UCP2*

Uma avaliação prévia da média dos espectros de cada grupo genotípico (Figura 6) apresenta os espectros médios de absorção no infravermelho para os grupos de pacientes com genótipo GG, GA e AA do polimorfismo -866G/A do gene *UCP2*. Os espectros mostram padrões semelhantes entre os grupos, com algumas variações sutis em regiões específicas. Na faixa entre 3500–3000 cm^{-1} , o grupo AA apresentou maior intensidade de absorção, o que pode indicar uma maior presença de ligações de hidrogênio, como em grupos hidroxila ou amina. Na região de 1650–1200 cm^{-1} , o grupo GA demonstrou um leve aumento na absorção em torno de 1350 cm^{-1} , sugerindo diferenças estruturais em proteínas ou compostos nitrogenados. Já na

região de 1050–750 cm^{-1} , o grupo GG apresentou uma absorção ligeiramente maior em torno de 950 cm^{-1} , o que pode estar relacionado a diferenças em estruturas de carboidratos ou compostos semelhantes.

Figura 6 - Espectros médios de absorvância FTIR dos genótipos AA, GA e GG.

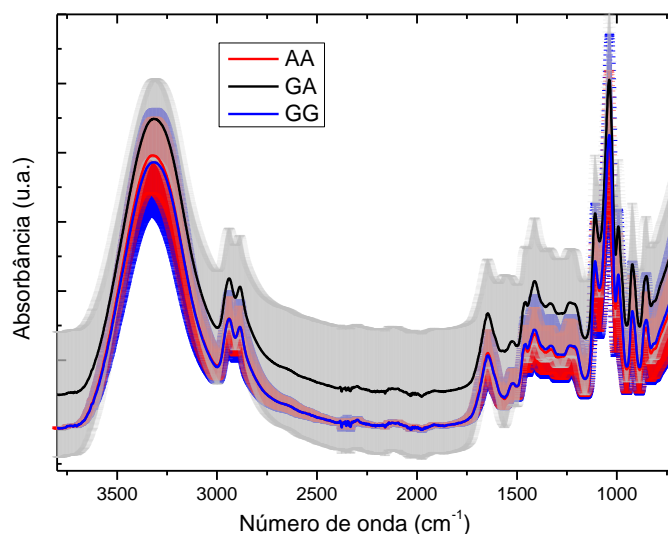


Legenda: Espectros médios de absorvância FTIR dos genótipos AA (vermelho), GA (preto) e GG (azul)

Fonte: Autora (2025).

A análise dos desvios médios sugere que, apesar das variações observadas, há sobreposição dos intervalos de desvio padrão entre os grupos (Figura 7). Essa sobreposição indica que as diferenças observadas podem não ser estatisticamente significativas, sugerindo que as variações detectadas nos espectros podem resultar de flutuações naturais ou variação biológica entre os indivíduos.

Figura 7 - Desvio médios espectros de absorvância FTIR dos genótipos AA, GA e GG.

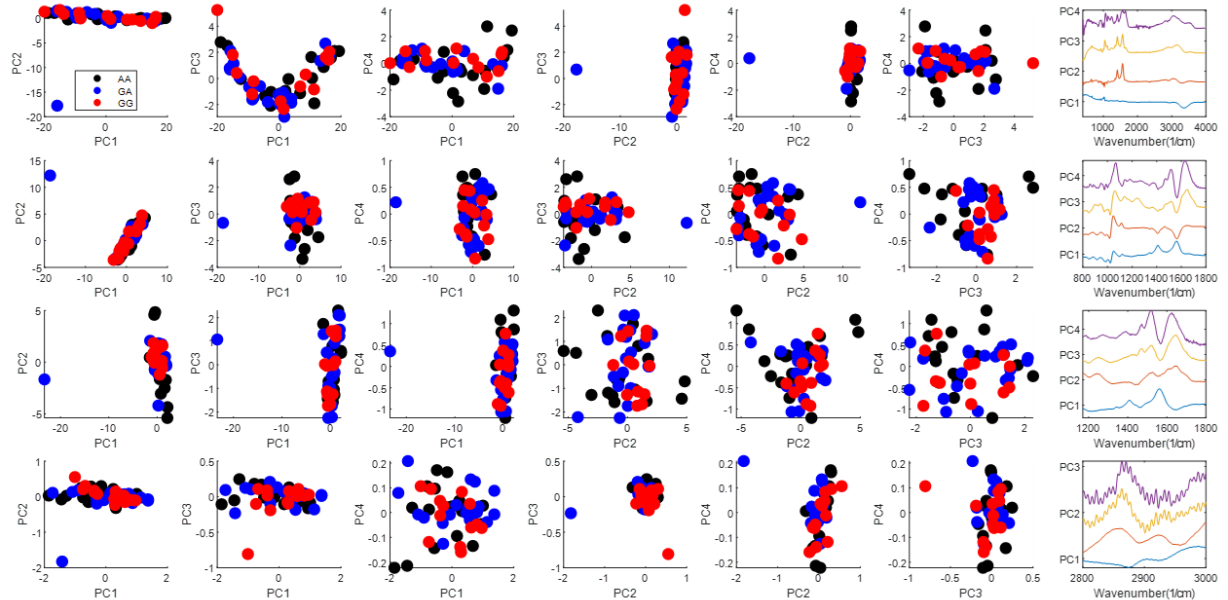


Legenda: Espectros médios de absorvância FTIR dos genótipos AA (vermelho), GA (preto) e GG (azul), com desvio padrão indicado pelas áreas sombreadas.

Fonte: Autora (2025).

O pré-processamento dos espectros foi realizado utilizando a técnica SNV (Variação Normal Padrão), com o objetivo de eliminar variações indesejadas e normalizar os dados, garantindo maior consistência entre as amostras analisadas. Em seguida, foi aplicada a Análise de Componentes Principais (PCA) para identificar variações espectrais que pudessem separar os grupos genéticos AA, GA e GG. Para isso foram avaliadas quatro faixas espectrais: 400-4000 cm⁻¹ espectro completo para uma análise ampla (PC1); 800-1800 cm⁻¹ faixa associada a proteínas e ácidos nucleicos (PC2); 1170-1800 cm⁻¹ região crítica para biomarcadores proteicos (PC3) e 2800-3000 cm⁻¹ região associada a lipídios e compostos orgânicos (PC4). A PCA permite reduzir a dimensionalidade dos dados espectrais, destacando os componentes principais responsáveis pelas maiores variações entre as amostras como demonstra a figura 8.

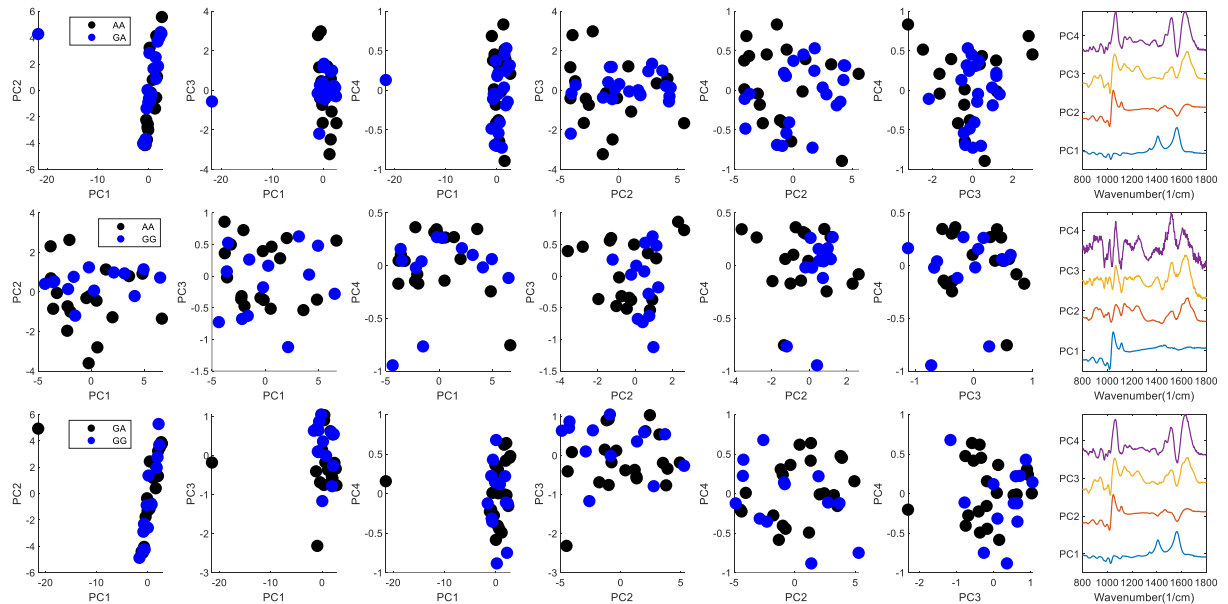
Figura 8 - Gráficos de dispersão com combinações diferentes entre os componentes principais (PC1, PC2, PC3, PC4) sendo preto: grupo AA, azul: grupo GA e vermelho: grupo GG.



Fonte: Autora (2024).

A análise de PC1 vs PC2 revelou que não houve uma separação clara entre os grupos AA, GA e GG. Embora alguns gráficos tenham mostrado a separação entre alguns pontos, a maioria das amostras apresentou sobreposição significativa, indicando que os grupos permanecem misturados nos componentes principais analisados. Esses resultados sugerem que os componentes principais 1 e 2 não são capazes de identificar padrões espectrais suficientes para distinguir de forma consistente os três grupos.

Figura 9 - Gráficos de dispersão com combinações diferentes entre os componentes principais (PC1, PC2, PC3, PC4) sendo preto: grupo AA, azul: grupo GA.



Fonte: Autora (2024).

Na figura 9, outras correlações foram realizadas agora comparando dois genótipos por vez e foi possível identificar que PC2 e PC4 (2ª linha, 5ª coluna), apresentou uma leve separação entre os pontos AA e GG, sugerindo que esses dois componentes são capazes de distinguir parcialmente esses grupos. Além disso, o gráfico espectral (2ª linha, 7ª coluna), onde PC3 é destacado em amarelo, revelou três picos importantes: um em torno de 1000 cm^{-1} e dois entre $1400\text{--}1800\text{ cm}^{-1}$, faixas associadas a proteínas.

De forma semelhante, o gráfico PC2 vs PC3 (3ª linha, 4ª coluna) mostrou uma separação parcial entre os grupos GA e GG. Esses resultados sugerem que PC3, em conjunto com PC2, possui potencial para distinguir os grupos GA e GG. Novamente, no gráfico de PC3 (3ª linha, 7ª coluna), as assinaturas associadas a proteínas ficaram evidentes, confirmando que essas assinaturas bioquímicas são as principais responsáveis pela separação observada entre os grupos.

Para complementar a PCA, foram aplicados algoritmos de aprendizado de máquina, como KNN e SVM. Essas técnicas foram avaliadas em diversas faixas espectrais para verificar sua capacidade de classificação, sendo que o KNN-Cosine apresentou a maior acurácia global de 54% ao ser aplicado na faixa mais ampla de $400\text{--}4000\text{ cm}^{-1}$. Em termos de desempenho específico, o modelo SVM Linear demonstrou uma boa performance na faixa $2800\text{--}3000\text{ cm}^{-1}$,

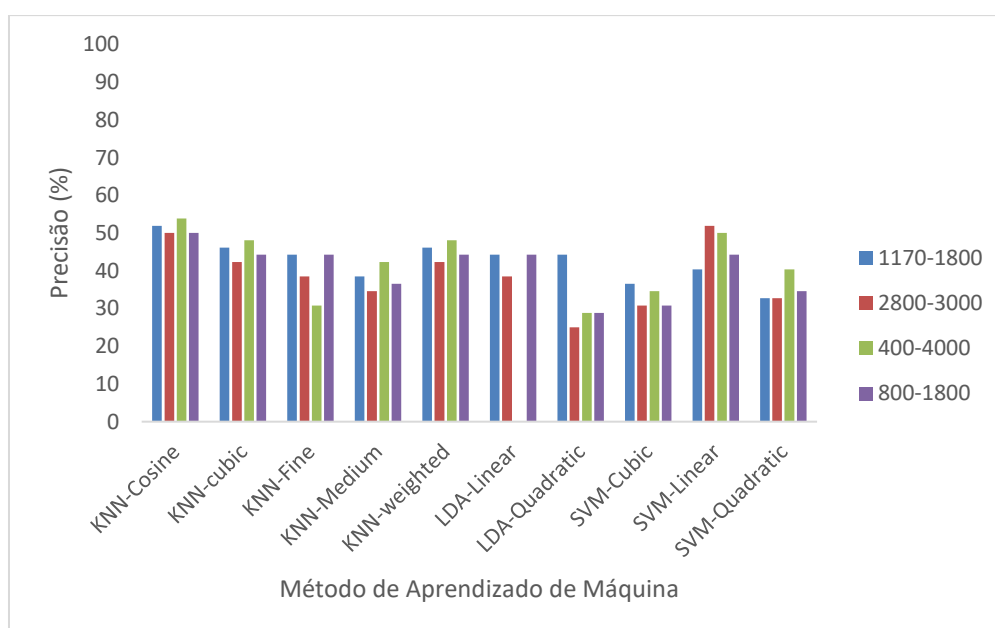
indicando que essa região do espectro pode conter informações importantes para a diferenciação dos genótipos, possivelmente relacionadas a padrões específicos de grupos funcionais.

Entretanto, os dados não apresentaram a melhor performance na faixa 800-1800 cm^{-1} para a classificação global. Para garantir uma análise abrangente, aplicamos todas as técnicas de ML em todas as faixas espectrais, e alguns resultados específicos se destacaram. Na faixa de 1170-1800 cm^{-1} , o modelo KNN-Cosine atingiu a maior acurácia para os pares AA vs GG (74%) e GA vs GG (73,5%). Esses resultados sugerem que essa faixa espectral é particularmente eficaz para a discriminação dos genótipos AA e GG, assim como GA e GG, indicando que pode haver padrões estruturais específicos capturados nessa região.

Já na faixa espectral entre 800-1800 cm^{-1} embora não tenha apresentado a melhor performance geral, ela mostrou ser útil na discriminação entre AA vs GA, com uma acurácia de 72% utilizando KNN-Cubic. Esse desempenho indica que, para esse par específico de genótipos, a aplicação do modelo com função de kernel cúbica conseguiu capturar as variações relevantes nos dados espectrais.

Para representar o grau de precisar das técnicas empregadas foi realizado um gráfico que ilustra o desempenho dos modelos em separar os genótipos GA, GG e AA (figura 10).

Figura 10 - Precisão (%) dos diferentes métodos de aprendizado de máquina aplicados nas faixas espectrais 1170-1800 cm^{-1} , 2800-3000 cm^{-1} , 400-4000 cm^{-1} e 800-1800 cm^{-1} para a classificação dos genótipos GA, AA e GG.



Fonte: Autora (2024).

Esses resultados revelam que a escolha da técnica de *machine learning* e da faixa espectral analisada tem impacto direto na precisão da classificação. A técnica KNN-Cosine se destacou como a mais eficiente, especialmente nas comparações entre AA vs GG e GA vs GG. A utilização da métrica de cosseno para a medida de similaridade parece ser eficaz na detecção de padrões relevantes nas faixas espectrais analisadas. A SVM Linear, por outro lado, apresentou um bom desempenho em faixas espectrais mais restritas, sugerindo que essa técnica é adequada para padrões específicos e concentrados. Já o KNN-Cubico se mostrou útil na faixa de 800-1800 cm^{-1} , onde outras técnicas não alcançaram o mesmo nível de acurácia.

6 DISCUSSÃO

O presente estudo investigou o potencial da combinação ATR-FTIR com algoritmos de aprendizado de máquina como uma alternativa rápida e acessível aos métodos moleculares e bioquímicos estabelecidos para a identificação de biomarcadores de risco para o desenvolvimento da DM2. Na análise do perfil bioquímico dos participantes, a observação de valores médio de glicose (130,97 mg/dL) e HbA1c (7,08%) no grupo experimental, constituído por indivíduos diabéticos, confirmam a presença da desregulação glicêmica, considerando que os valores de referência para o diagnóstico de DM2 estabelecido pela diretriz da Sociedade Brasileira de Diabetes é de glicose 120mg/dL e 5,7% de HbA1c. Neste grupo amostral de indivíduos com DM2 fica claro a necessidade de intervenções terapêuticas e tratamento clínico adequado para mitigar os riscos associados à hiperglicemia crônica como preconizado pela Sociedade Brasileira de Diabetes (2024).

Entre os participantes do nosso estudo que declararam ter DM2, quatro deles utilizam medicação de rotina, especificamente o medicamento Glifage 500 mg, para controle. Contudo, apesar do tratamento, as suas concentrações séricas de glicose permaneceram acima dos valores recomendados. A média da glicose no grupo foi de 126,77 mg/dL, e a hemoglobina glicada apresentou um valor médio de 7,1%, ambos acima dos limites de normalidade. Todos os quatro participantes medicados utilizavam. Os demais participantes com diagnóstico de DM2 que estavam sob tratamento medicamentoso e apresentaram concentrações de glicose dentro da normalidade não entraram na amostragem utilizada para a análise de FTIR, pois esta análise estratificou os grupos de acordo com as concentrações de glicemia.

A presença dos três perfis genotípicos do SNP -866G/A do *gene UCP2* identificados em nosso estudo, sendo AA observado em 34,61% das amostras (fragmentos de 295pb e 68pb), GA (fragmentos de 363pb, 295 pb e 68pb) detectado em 40,38% das amostras e GG (fragmento de 363pb) do SNP -866G/A do *gene UCP2* em 25% das amostras, confirmou a heterogeneidade desse polimorfismo na nossa população. Na comparação com estudos publicados em outras populações de vários países (Wang; Ke; Lin, 2012; de Souza *et al.*, 2013; Shen *et al.*, 2014; Qin *et al.*, 2013; Gomathi *et al.*, 2019; Bulotta *et al.*, 2005; Andersen *et al.*, 2013), observa-se que a frequência do genótipo AA é maior em nossa população.

Andersen *et al.* (2013), pela avaliação de uma amostra de 17.636 indivíduos dinamarqueses, observaram a frequência genotípica de AA: 17% para o genótipo AA, 48% para o genótipo GA e 35% para o genótipo GG, com menor prevalência do genótipo AA quando comparado com nosso estudo que foi de 34,61%. Bulotta *et al.* (2005), em um estudo com 1.073

indivíduos da população italiana, observaram a frequência de 25% para o genótipo AA 50% para o genótipo GA e 25% para o genótipo GG, valores que também diferem da nossa amostra, especialmente na menor frequência do genótipo AA e maior frequência do genótipo GA.

Em populações asiáticas, variações significativas também foram observadas. Gomathi *et al.* (2019) analisaram 630 indivíduos do sul da Índia e encontraram AA: 14%, GA: 46% e GG: 40%, revelando a menor prevalência do genótipo AA dentre todos os estudos revisados. Já Hou *et al.* (2020), estudando 1.476 chineses do norte, reportaram AA: 15%, GA: 48% e GG: 37%, com uma distribuição semelhante a de Qin *et al.* (2013), que avaliou 715 chineses de Hubei e encontrou AA: 21%, GA: 50% e GG: 29%.

Em Taiwan, Wang; Ke e Lin (2012) analisaram 604 indivíduos, relatando AA: 23%, GA: 47% e GG: 30%, enquanto Sasahara *et al.* (2004), em um estudo com 585 japoneses, identificaram uma maior proporção do genótipo AA (29%), seguida de GA: 51% e GG: 20%, aproximando-se do nosso estudo no que diz respeito à frequência do genótipo AA. Xu; Chen; e Zhan (2021), em uma meta-análise de diversas populações asiáticas, reportou AA: 22%, GA: 49% e GG: 29%, enquanto Shen *et al.* (2014), analisando 958 chineses, encontrou AA: 20%, GA: 50% e GG: 30%.

No Brasil, até o momento, apenas um estudo de Souza *et al.* (2013) investigou o SNP - 866G/A do *gene UCP2*. Estes autores analisaram indivíduos brasileiros e encontraram uma distribuição de 17,2% para o genótipo de AA, 47,8% para o genótipo GA e 35% para o genótipo GG: Esses valores diferem significativamente dos observados em nosso estudo, onde a frequência do genótipo AA observada foi quase o dobro da relatada por de Souza *et al.*, (2013), enquanto a frequência dos genótipos GA e o GG apresentaram distribuições mais próximas.

Em comparação, a maior frequência do genótipo AA observada em nosso estudo (34,61%) se aproxima mais dos valores encontrados na população japonesa (29%) (Sasahara *et al.*, 2004), sendo superior às proporções relatadas em todas as demais populações analisadas (Wang; Ke e Lin, 2012; de Souza *et al.*, 2013 ; Shen *et al.*, 2014; Qin *et al.*, 2013; Gomathi *et al.*, 2019). A frequência de 40,38% para o genótipo GA observada em nossa amostra foi inferior à observada em estudos de outras populações, com a frequência desse genótipo variando entre 46% e 51%. Para o genótipo GG (25%) a frequência foi similar ao encontrado na população italiana (Bulotta *et al.* 2005) e chinesa (Qin *et al.* 2013), mas inferior ao reportado na população do sul da Índia (40%) (Gomathi *et al.* 2019).

As variações observadas entre os diferentes estudos podem estar associadas a diferenças étnicas e geográficas, além de fatores ambientais e amostrais que influenciam a distribuição dos

genótipos do *gene UCP2* em diferentes populações. O estudo de de Souza *et al.* (2013) representa, até o momento, a única investigação que busca relacionar esse SNP com a predisposição a DM2 realizado no Brasil, evidenciando a necessidade da realização de pesquisas adicionais para esclarecer a sua distribuição e impacto na população brasileira.

A maior concentração de insulina observada no grupo com o genótipo GA do SNP -866G/A do *gene UCP2* quando comparada com os grupos com os genótipos GG e AA sugere que indivíduos que possuem o genótipo GA apresentem maiores níveis de insulina. Alguns estudos (Krempler *et al.*, 2002; Sesti *et al.*, 2003; Gomathi *et al.*, 2019; Hou *et al.*, 2020; Andersen *et al.*, 2013) também pesquisaram essa relação, contudo os resultados foram diferentes quando comparados com o presente estudo. Resultados semelhantes aos resultados encontrados neste estudo foram os de Gomathi *et al.* (2019), que observaram concentrações mais elevadas de insulina em indivíduos com os genótipos GA e GG do SNP -866G/A em comparação com aqueles portadores do genótipo AA, neste estudo, realizado com 318 pacientes com diagnóstico de DM2 e 312 indivíduos controle na população do sul da Índia, os níveis médios de insulina em jejum foram $17,6 \pm 1,9 \mu\text{U/mL}$ para o genótipo GG, $15,8 \pm 2,2 \mu\text{U/mL}$ para GA e $13,4 \pm 1,7 \mu\text{U/mL}$ para AA ($P < 0,005$), evidenciando uma redução significativa nos níveis de insulina nos indivíduos portadores do genótipo AA. Os resultados de Hou *et al.* (2020) corroboram essa tendência, demonstrando que indivíduos com o genótipo AA apresentaram menores concentrações de insulina em jejum e pós-prandial em comparação com os portadores dos genótipos GG e GA. No grupo com pré-diabetes, os níveis médios de insulina em jejum foram de $9,22 \pm 3,18 \text{ mU/mL}$ para GG, $8,56 \pm 2,72 \text{ mU/mL}$ para GA e $7,65 \pm 3,83 \text{ mU/mL}$ para AA ($P < 0,001$). No grupo com diagnóstico de DM2, os valores foram de $8,13 \pm 2,74 \text{ mU/mL}$ para GG, $7,92 \pm 2,73 \text{ mU/mL}$ para GA e $6,38 \pm 2,50 \text{ mU/mL}$ para AA ($P < 0,001$). A redução da insulina nos portadores do alelo A sugere um comprometimento na secreção pancreática, possivelmente associado a uma maior expressão do *gene UCP2*.

Krempler *et al.* (2002) ao avaliarem uma população branca europeia, demonstraram que o genótipo GG foi associado a uma maior resposta insulínica aguda ($314,1 \pm 145,2 \text{ mU/L}$) em comparação com GA ($264,3 \pm 187,9 \text{ mU/L}$) e AA ($233,4 \pm 160,2 \text{ mU/L}$) ($P = 0,015$), além de um maior índice de disposição, indicando uma melhor função das células β pancreáticas. O estudo de Sesti *et al.* (2003) também reforça essa relação, evidenciando que o polimorfismo *UCP2* -866G/A está diretamente ligado às concentrações de insulina, sendo que os portadores do genótipo AA apresentam uma secreção reduzida em comparação aos indivíduos GG e GA. Durante o teste oral de tolerância à glicose, os indivíduos AA apresentaram menores índices de secreção na primeira ($P = 0,03$) e na segunda fase ($P = 0,04$), sugerindo uma redução na

capacidade das células β pancreáticas de responder à glicose. Essa redução foi confirmada na análise de ilhotas pancreáticas isoladas, onde as células de indivíduos AA liberaram menos insulina ($39 \mu\text{U/mL}$) em resposta à glicose ($16,7 \text{ mmol/L}$) do que as de indivíduos GG ($89 \pm 19 \mu\text{U/mL}$) e GA ($50 \pm 10 \mu\text{U/mL}$). Já os estudo de Andersen *et al.* (2013), ao analisarem o polimorfismo *UCP2* -866G>A em 17.636 indivíduos da população da Dinamarca, constataram que o genótipo GG foi associado a níveis mais elevados de insulina em jejum ($P = 0,002$) e maior resistência à insulina ($P = 0,0007$). Esses achados sugerem que a menor expressão do gene *UCP2* pode afetar a função mitocondrial e aumentar o estresse oxidativo, contribuindo para a resistência insulínica.

A análise espectral da concentração de glicose nos indivíduos levou em consideração a região espectral de $1300\text{--}900 \text{ cm}^{-1}$, que contém o pico de absorção característico da glicose, conforme relatado por diversos estudos (Li *et al.* 2005; Li *et al.* 2017; Caixeta *et al.* 2023). Essa região se destaca por sua relevância na investigação das diferenças na composição bioquímica entre os dois grupos de pacientes analisados. Por esse motivo, foi selecionada para a análise de cluster, permitindo a classificação dos grupos de acordo com os níveis de glicose.

Li *et al.* (2005) destacou a região em torno de 1035 cm^{-1} como crucial para a absorção de glicose no infravermelho. Além disso, os achados de Guang *et al.* (2020) sugeriram que a faixa espectral de $1000\text{--}1500 \text{ cm}^{-1}$ é essencial para diferenciar grupos de pacientes com DM2 e indivíduos saudáveis. Neste estudo, foram analisadas amostras de sangue total de 51 pacientes diagnosticados com DM2 e 55 voluntários saudáveis por meio de espectroscopia FTIR-ATR, combinada com o algoritmo XGBoost para classificação. A análise espectral revelou que os indivíduos com DM2 apresentaram intensidade de absorção reduzida na região de $1000\text{--}1500 \text{ cm}^{-1}$, indicando alterações metabólicas significativas. Mateus Pereira de Souza *et al.* (2023), ao avaliarem 74 amostras de plasma sanguíneo (14 com síndrome metabólica e 60 amostras do grupo controle) para a diferenciação de pacientes com síndrome metabólica na população brasileira, identificaram vibrações específicas entre $1020\text{--}1050 \text{ cm}^{-1}$, associadas à concentração de glicose. O estudo utilizou espectroscopia ATR-FTIR acoplada à modelagem quimiométrica supervisionada (OPLS-DA), atingindo 100% de sensibilidade e especificidade na diferenciação entre síndrome metabólica e indivíduos controle. As principais diferenças espectrais ocorreram nas regiões de $1717\text{--}1703 \text{ cm}^{-1}$ (C=O e N-H), $1166\text{--}1137 \text{ cm}^{-1}$ (C-OH e C-O), $1113\text{--}1040 \text{ cm}^{-1}$ (PO_2^- e C-OH) e $1027\text{--}1008 \text{ cm}^{-1}$ (C-O e CH_2OH), refletindo alterações bioquímicas associadas à síndrome metabólica

A região espectral determinada em nosso estudo é similar as regiões escolhidas de absorção de glicose dos trabalhos prévios (Li *et al.*, 2005; Li *et al.*, 2017; Caixeta *et al.*, 2023)

fortalecendo assim a validade de nossa abordagem e reforçando a importância da análise da região espectral de 1139-1024 cm^{-1} na diferenciação entre os grupos de participantes que apresentam alto e baixo nível de glicose.

Com o objetivo de aumentar a acurácia de 78% da análise de clusterização observada em nosso estudo na estratificação do grupos dos pacientes com diferentes concentrações de glicose na faixa espectral da glicose (glicose acima e abaixo de 100mg/dL) foi aplicado o modelo de treinamento ANN, que aumentou a acurácia dessa classificação para 81% com excelente equilíbrio entre as classes, conforme demonstrado pelo F1-score (77% para classe 0 e 87% para classe 1). Alguns modelos de treinamento foram empregados em estudos publicados recentemente para a classificação dos indivíduos baseado nos níveis glicose e para a classificação de patologias como o câncer (Guevara *et al.*, 2018; NÃO CONSTA; Sánchez-Brito *et al.*, 2021; Tomas *et al.*, 2022; Caixeta *et al.*, 2023; Song *et al.*, 2023; NÃO CONSTA; MITURA *et al.*, 2024).

Guevara *et al.* (2018) utilizaram espectroscopia Raman combinada a modelos de aprendizado supervisionado, como ANN e SVM, para discriminar pacientes com DM2 de indivíduos saudáveis. Os autores relataram acurácia entre 88,9% e 90,9% para ANN e entre 76,0% e 82,5% para SVM, destacando o melhor desempenho da ANN, resultado que está alinhado com a acurácia de 83% obtida pelo modelo de rede neural desenvolvido neste estudo. Assim como nos resultados apresentados por Guevara *et al.* (2018), o modelo desenvolvido neste trabalho também apresentou um desempenho superior na identificação da classe associada a DM2 (classe 1), com recall de 91%, evidenciando que a ANN foi mais eficiente na detecção de casos positivos.

Sánchez-Brito *et al.* (2021) também utilizaram espectroscopia FTIR-ATR em amostras de saliva associada a ANN para classificar pacientes com DM2 controlada e não controlada, atingindo acurácia de 100%. Embora o modelo desenvolvido neste estudo tenha apresentado acurácia inferior (83%), o desempenho em termos de precisão e F1-score de 83% reforça a capacidade de discriminação equilibrada entre as classes. A diferença na acurácia pode ser explicada pela natureza das amostras e pelas variações nos padrões espectrais associados a diferentes condições metabólicas. No entanto, o valor médio do F1-score (82%) e do recall (81%) sugere que o modelo lidou bem com a distribuição das classes, o que reforça sua robustez mesmo em um cenário mais complexo.

Caixeta *et al.* (2023) obtiveram uma acurácia de 87% na discriminação entre indivíduos com DM2 e não diabéticos por meio de espectroscopia FTIR em amostras de saliva, com sensibilidade de 93,3% e especificidade de 74%. O modelo desenvolvido neste estudo

apresentou um comportamento semelhante, com recall de 91% para a classe positiva e 71% para a classe negativa, indicando uma tendência a uma maior capacidade de identificação de casos positivos. Essa diferença na identificação entre as classes reflete o desafio inerente à classificação de condições metabólicas, sendo parcialmente compensada pelo equilíbrio geral observado nos valores de F1-score e precisão (83%).

Song *et al.* (2023) relataram resultados mais elevados ($R^2 = 0,9928$) na previsão de glicose em amostras de sangue por espectroscopia Raman combinada com aprendizado de máquina, indicando uma correlação quase perfeita entre os valores previstos e os valores reais. Embora o modelo desenvolvido neste trabalho tenha apresentado valores inferiores em termos de correlação direta, a capacidade de classificação binária com valores consistentes de precisão, recall e F1-score reforça a aplicabilidade da técnica para análise metabólica.

Caixeta *et al.* (2023) também utilizaram espectroscopia FTIR-ATR em amostras de urina de ratos Wistar para análise de DM2, com acurácia, sensibilidade e especificidade de 100%, superando o desempenho do modelo desenvolvido neste estudo. No entanto, o equilíbrio entre as métricas obtidas neste trabalho (F1-score de 77% para a classe 0 e 87% para a classe 1) sugere que o modelo foi eficiente na discriminação entre os grupos, mesmo em um contexto de menor controle sobre a variabilidade das amostras.

Hoşafçı *et al.* (2006) relataram coeficiente de determinação (R^2) de 0,985 e erro de previsão de 5,2 mg/dL na análise de glicose em urina, utilizando espectroscopia FTIR associada a modelos quimiométricos. Embora os métodos utilizados sejam diferentes, a capacidade do modelo desenvolvido neste estudo de alcançar uma precisão consistente de 83% sugere que o uso de espectroscopia FTIR combinada com aprendizado de máquina é promissor para monitoramento e diagnóstico metabólico.

Bhandare *et al.* (1993) relataram que a combinação de PLS com ANN melhorou a precisão na determinação de glicose em sangue total, com erro padrão de predição de 15,6 mg/dL, superando os modelos PLS (21,5 mg/dL) e PCR (24,0 mg/dL). A alta precisão e o equilíbrio obtidos pelo modelo desenvolvido neste estudo reforçam o papel das redes neurais na modelagem de dados espectroscópicos complexos, permitindo uma análise robusta e confiável dos níveis de glicose.

Em síntese, o modelo de rede neural desenvolvido neste estudo apresentou desempenho consistente e equilibrado entre as classes, com acurácia e precisão de 83% e F1-score médio de 82%. Esses resultados estão alinhados com os estudos prévios, evidenciando que a combinação de espectroscopia FTIR-ATR com aprendizado de máquina é uma ferramenta promissora para o diagnóstico e monitoramento de DM2 em diferentes matrizes biológicas.

Os resultados obtidos em nosso estudo também mostraram a eficácia da combinação da espectroscopia ATR-FTIR com técnicas de aprendizado de máquina, como PCA, KNN e SVM em amostras de amplicons da sequência de DNA que contem o SNP -866G/A do *gene UCP2*, com o objetivo de discriminar os diferentes genótipos presentes na população brasileira referente a este SNP: GG, GA e AA. Entre os algoritmos testados, o modelo KNN-cosine obteve o melhor desempenho, com uma acurácia média de 74% na faixa espectral de 1170–1800 cm^{-1} na diferenciação entre AA e GG, e 73,5% na separação entre GA e GG.

Em nosso estudo, o SVM também apresentou bom desempenho, capturando padrões espectrais relacionados a variações no DNA na faixa de 2800-3000 cm^{-1} . A separação entre os genótipos AA e GA atingiu 69,2% de acurácia, enquanto a discriminação entre AA e GG foi de 58,1%. Além disso, a análise de PCA permitiu a redução dimensional dos espectros, destacando regiões espectrais críticas associadas a alterações estruturais no DNA.

Os achados do nosso estudo estão em concordância com os resultados do trabalho de Taillandier e Liquier (1992), que aplicaram a espectroscopia FTIR na caracterização das estruturas secundárias do DNA em diferentes estados físicos, como soluções, filmes hidratados e cristais. Os autores identificaram as conformações A, B e Z do DNA, associadas a bandas espectrais entre 1800 e 700 cm^{-1} , abrangendo bases nitrogenadas, açúcares e grupos fosfato. Já em nosso estudo, regiões espectrais específicas foram relacionadas à estrutura do DNA, como 1220–1260 cm^{-1} e 1070–1090 cm^{-1} , associadas às vibrações dos grupos fosfato (PO_4^{3-}), permitindo a distinção entre as diferentes conformações estruturais do DNA. A região entre 1710–1660 cm^{-1} foi atribuída às ligações C=O das bases nitrogenadas, enquanto a faixa de 1490–1450 cm^{-1} indicou variações nas interações açúcar-base. A análise espectral revelou que a transição da conformação B para A ocorre em baixa umidade, enquanto a indução da forma Z é favorecida pela presença de íons metálicos (Taillandier e Liquier, 1992). Esses achados reforçam a aplicabilidade da espectroscopia FTIR na avaliação de modificações estruturais do DNA, incluindo interações com proteínas, fármacos e fatores ambientais.

Resultados consistentes também foram observados no estudo de Brasileira *et al.* (2003), que utilizou modelos supervisionados, como SVM e KNN, para discriminar genótipos de arroz geneticamente modificados. O SVM apresentou acurácia de 90,5%, seguido pela análise discriminante linear (87,2%) e KNN (84,6%) na diferenciação entre plantas resistentes e sensíveis. Embora o material genético seja distinto, o uso das mesmas técnicas de aprendizado de máquina reforça a aplicabilidade desses algoritmos na discriminação de genótipos com base em espectros FTIR, o que é corroborado pelo desempenho do KNN e do SVM em nosso estudo.

Outro paralelo pode ser traçado com o trabalho de Leskovar *et al.* (2020), que investigou a preservação de DNA em restos ósseos humanos utilizando espectroscopia ATR-FTIR combinada com algoritmos de aprendizado de máquina, como RF e ANN. Os autores identificaram regiões espectrais críticas, incluindo 1035 cm^{-1} (fosfatos), 1240 cm^{-1} (fosfodiéster), 1400 cm^{-1} (carbonatos), 1545 cm^{-1} (amida II) e 1650 cm^{-1} (amida I), que foram fundamentais para discriminar amostras com DNA bem preservado daquelas degradadas, atingindo acurácia de até 82%. Em nosso estudo, a região entre $1170\text{--}1800\text{ cm}^{-1}$ também se destacou na diferenciação genotípica, indicando que essas bandas espectrais podem refletir a integridade estrutural do DNA, similar ao que foi observado por Leskovar *et al.* (2020).

Em relação à aplicação de aprendizado de máquina em espectros FTIR de DNA amplificado por PCR, o número de estudos publicados é limitado. Gomes Rios *et al.* (2021) analisaram polimorfismos do gene do hormônio de crescimento bovino (bGH) utilizando PCA, SVM e KNN, alcançando acurácia de 90% na separação entre os genótipos CC e TT, 76% entre CC e CT e 64% entre CT e TT. Apesar de os algoritmos e análises serem semelhantes, a acurácia obtida em nosso estudo foi inferior, o que pode ser explicado pela diferença nas sequências genômicas, no número de amostras e na composição dos grupos genotípicos (13 GG, 21 GA e 18 AA). A menor quantidade de amostras em alguns grupos pode ter reduzido a capacidade dos modelos de capturar padrões espectrais consistentes, impactando diretamente na acurácia final.

Adicionalmente, o estudo de Kahn *et al.* (2009) sugere que a composição das regiões flanqueadoras pode influenciar a resposta espectral do DNA. Os autores demonstraram que sequências ricas em citosina e guanina conferem maior rigidez estrutural ao DNA, enquanto sequências ricas em adenina e timina aumentam a flexibilidade e distorções conformacionais, o que impacta diretamente nas assinaturas espectrais capturadas pelos modelos de aprendizado de máquina. O modelo SVM empregado por Kahn *et al.* (2009) apresentou uma acurácia de 91,2%, reforçando a influência das sequências flanqueadoras na estabilidade estrutural e nos padrões espectrais. Assim, é plausível que variações na composição das regiões flanqueadoras do SNP -866G/A do gene *UCP2* também tenham influenciado a resposta espectral em nosso estudo, contribuindo para a diferença na acurácia entre os modelos.

Portanto, os resultados obtidos em nosso estudo não apenas confirmam a aplicabilidade da espectroscopia ATR-FTIR na discriminação de genótipos com base em assinaturas espectrais, como também demonstram que fatores genômicos, estruturais e metodológicos podem impactar o desempenho dos modelos de aprendizado de máquina. A combinação de espectroscopia FTIR e aprendizado de máquina revela-se, assim, uma abordagem promissora

para a análise genotípica, com potencial para aplicações em estudos de variações genéticas e suas implicações biológicas.

7 CONCLUSÃO

O presente estudo conseguiu genotipar o polimorfismo -866G/A do gene *UCP2* por meio da técnica de PCR-RFLP, além de demonstrar que indivíduos com o genótipo GA apresentam concentrações de insulina estatisticamente maiores em comparação aos demais genótipos (GG e AA). Além disso, a técnica ATR-FTIR mostrou-se eficiente na diferenciação de indivíduos com diferentes concentrações séricas de glicose por meio da análise de cluster hierárquico, com acurácia de 78%.

A análise espectral de DNA, associada a algoritmos de aprendizado de máquina, demonstrou o potencial da técnica ATR-FTIR para análise do polimorfismo -866G/A do gene *UCP2* e para a predição de concentrações glicêmicas, evidenciando a aplicabilidade dessa abordagem como ferramenta preditiva e diagnóstica para diabetes tipo 2 (DM2). Entre os modelos testados para as amostras de soro, o algoritmo ANN destacou-se ao alcançar uma precisão de 83%, com *recall* de 91% na separação entre grupos com glicose elevada e grupos com níveis normais de glicose. No caso da classificação dos genótipos, o KNN-Cosine obteve até 74% de acurácia na distinção entre os genótipos AA e GG, e 73,5% para GA versus GG na faixa espectral de 1170–1800 cm^{-1} , associada a proteínas.

Esses resultados reforçam a capacidade da integração entre espectroscopia ATR-FTIR e algoritmos de aprendizado de máquina para análises rápidas e de menor custo, tanto para a estratificação de pacientes com base nos níveis de glicose em amostras de soro obtidas de sangue humano quanto para a distinção de polimorfismos genéticos em amostras de DNA amplificado por PCR.

Apesar de limitações, como o tamanho reduzido da amostra e a sobreposição espectral em alguns genótipos, a eficiência da técnica ATR-FTIR associada a algoritmos de aprendizado de máquina na discriminação de polimorfismos genéticos específicos e na estratificação das amostras com base em concentrações glicêmicas indica um caminho promissor para aplicações clínicas. A simplicidade e o custo reduzido dessa metodologia destacam sua vantagem em relação aos métodos diagnósticos tradicionais de análise bioquímica e molecular.

Este trabalho contribui para o campo da engenharia biomédica ao propor uma solução tecnológica inovadora e acessível, ressaltando a importância da interseção entre biologia molecular, análise espectral e inteligência artificial.

REFERÊNCIAS

AHLQVIST, E.; STORM, P.; KÄRÄJÄMÄKI, A.; MARTINELL, M.; DORKHAN, M.; CARLSSON, A.; VIKMAN, P.; PRASAD, R. B.; ALY, D. M.; ALMGREN, P.; WESSMAN, Y.; SHAAT, N.; SPÉGEL, P.; MULDER, H.; LINDHOLM, E.; MELANDER, O.; HANSSON, O.; MALMQVIST, U.; LERNMARK, Å.; LAHTI, K.; FORSÉN, T.; TUOMI, T.; ROSENGREN, A. H.; GROOP, L. Novel subgroups of adult-onset diabetes and their association with outcomes: a data-driven cluster analysis of six variables. **The Lancet Diabetes & Endocrinology**, v. 6, n. 5, p. 361-369, 2018.

AI, Y.; GU, B.; QIU, J. R.; YI, H. G.; ZHU, Q.; ZHANG, L.; HU, G. The uncoupling protein 2 -866G > A polymorphism is associated with the risk of ischemic stroke in Chinese type 2 diabetic patients. **CNS Neuroscience & Therapeutics**, v. 18, n. 8, p. 636-640, 2012.

AMERICAN DIABETES ASSOCIATION. Standards of medical care in diabetes—2014. **Diabetes Care**, v. 37, suppl. 1, p. S14-S80, 2014. Disponível em: <https://doi.org/10.2337/dc14-S014>. Acesso em: 2 dez. 2024.

ANDERSEN, G.; DALGAARD, L. T.; JUSTESEN, J. M.; ANTHONSON, S.; NIELSEN, T.; THØRNER, L. W.; WITTE, D.; JØRGENSEN, T.; CLAUSEN, J. O.; LAURITZEN, T.; HOLMKVIST, J.; HANSEN, T.; PEDERSEN, O. The frequent UCP2-866G>A polymorphism protects against insulin resistance and is associated with obesity: a study of obesity and related metabolic traits among 17,636 Danes. **International Journal of Obesity**, v. 37, n. 2, p. 175–181, 2013.

ANDRADE, G. C.; LEVY, R. B.; LEITE, M. A.; RAUBER, F.; CLARO, R. M.; COUTINHO, J. G.; MAIS, L. A. Mudanças nos marcadores da alimentação durante a pandemia de covid-19 no Brasil. **Revista de Saúde Pública**, v. 57, n. 1, 2023.

ANTZA, C.; KOSTOPOULOS, G.; MOSTAFA, S.; NIRANTHARAKUMAR, K.; TAHRANI, A. The links between sleep duration, obesity and type 2 diabetes mellitus. **Journal of Endocrinology**, v. 252, n. 2, p. 125–141, 2022.

ARRIETA-CRUZ, I.; TORRES-ÁVILA, B. S.; MARTÍNEZ-CORIA, H.; LÓPEZ-VALDÉS, H. E.; GUTIÉRREZ-JUÁREZ, R. Diet-induced metabolic dysfunction of hypothalamic nutrient sensing in rodents. **International Journal of Molecular Sciences**, v. 23, n. 7, 2022.

BAI, Y. *et al.* Persistent Nuclear Factor- κ B Activation in Ucp2^{-/-} Mice Leads to Enhanced Nitric Oxide and Inflammatory Cytokine Production. **The Journal of Biological Chemistry**, v. 290, n. 9, p. 19062-19069, mai. 2005.

BANDEIRA, N.; MARINHO, P.; ALVES DE VASCONCELOS, H. C.; PARENTE, A. M.; ALENCAR, G.; CÉSAR DE ALMEIDA, P.; COELHO DAMASCENO, M. M. Risk for type 2 diabetes mellitus and associated factors. **Acta Paul Enferm.**, v. 26, n. 6, 2013.

BARENGOLTS, E. Gut microbiota, prebiotics, probiotics, and synbiotics in management of obesity and prediabetes: review of randomized controlled trials. **Endocrine Practice**, v. 22, n. 10, p. 1224–1234, 2016.

BERGMANN, A.; LI, J.; REIMANN, M.; HENTRICH, T.; HANEFELD, M.; BORNSTEIN, S. R.; SCHWARZ, P. E. H. Polymorphisms in perilipin gene (PLIN) are not associated with obesity and weight variation in people with high risk of type 2 diabetes. **Experimental and Clinical Endocrinology and Diabetes**, v. 116, supl. 1, p. 56–59, 2008.

BHANDARE, P.; MENDELSON, Y.; PEURA, R. A.; JANATSCH, G.; KRUSE-JARRES, J. D.; MARBACH, R.; HEISE, H. M. Multivariate determination of glucose in whole blood using partial least-squares and artificial neural networks based on mid-infrared spectroscopy. **Vibrational Spectroscopy**, v. 4, n. 2, p. 239–250, 1993.

BIDDLE, G. J. H.; EDWARDSON, C. L.; HENSON, J.; DAVIES, M. J.; KHUNTI, K.; ROWLANDS, A. V.; YATES, T. Associations of physical behaviours and behavioural reallocations with markers of metabolic health: a compositional data analysis. **International Journal of Environmental Research and Public Health**, v. 15, n. 10, 2018.

BIRD, S. R.; HAWLEY, J. A. Update on the effects of physical activity on insulin sensitivity in humans. **BMJ Seminars**, v. 2, p. 143, 2017.

BOSCHINI, R. P.; RODRIGUES, J.; JÚNIOR, G. Regulação da expressão gênica das UCP2 e UCP3 pela restrição energética, jejum e exercício físico. **Revista de Nutrição**, Campinas, v. 18, n. 6, 2005.

BRAND, M. D. Mitochondrial generation of superoxide and hydrogen peroxide as the source of mitochondrial redox signaling. **Free Radical Biology and Medicine**, v. 100, p. 14-31, 2016.

BRASIL. Ministério da Saúde. **Vigitel Brasil 2023**: vigilância de fatores de risco e proteção para doenças crônicas por inquérito telefônico. Brasília, DF: Ministério da Saúde, 2023.

BRASILEIRA, R.; LILGE, C. G.; MACHADO, Â. M.; TILLMANN, A.; VILLELA, F. A.; DODE, L. B. Identificação de genótipos de arroz geneticamente modificados: identificação de sementes de arroz transformado geneticamente resistente ao herbicida glufosinato de amônio. **Revista Brasileira de Sementes**, v. 25, 2003.

BROWN, G. K. Glucose transporters: structure, function and consequences of deficiency. **J. Inherit. Metab. Dis.**, v. 23, 2000

BULOTTA, A.; LUDOVICO, O.; COCO, A.; DI PAOLA, R.; QUATTRONE, A.; CARELLA, M.; PELLEGRINI, F.; PRUDENTE, S.; TRISCHITTA, V. The common -866G/A polymorphism in the promoter region of the UCP-2 gene is associated with reduced risk of type 2 diabetes in Caucasians from Italy. **Journal of Clinical Endocrinology and Metabolism**, v. 90, n. 2, p. 1176–1180, 2005.

CADENAS, S. Mitochondrial uncoupling, ROS generation and cardioprotection. **Biochimica et Biophysica Acta (BBA)-Bioenergetics**, v. 1859, n. 9, p. 940-950, 2018..

CAIXETA, D. C.; CARNEIRO, M. G.; RODRIGUES, R.; ALVES, D. C. T.; GOULART, L. R.; CUNHA, T. M.; ESPINDOLA, F. S.; VITORINO, R.; SABINO-SILVA, R. Salivary ATR-FTIR spectroscopy coupled with support vector machine classification for screening of type 2 diabetes mellitus. **Diagnostics**, v. 13, n. 8, 2023.

CAIXETA, D. C.; LIMA, C.; XU, Y.; GUEVARA-VEGA, M.; ESPINDOLA, F. S.; GOODACRE, R.; ZECELL, D. M.; SABINO-SILVA, R. Monitoring glucose levels in urine using FTIR spectroscopy combined with univariate and multivariate statistical methods. **Spectrochimica Acta - Part A: Molecular and Biomolecular Spectroscopy**, v. 290, p. 122259, 2023.

CHAI, Y. *et al.* The Uncoupling Protein 2 -866G > A Polymorphism is Associated with the Risk of Ischemic Stroke in Chinese Type 2 Diabetic Patients. **CNS Neuroscience & Therapeutics**, v. 18, n. 8, p. 636-640, mai. 2012.

CHEN, C.; ZENG, Y.; XU, J.; ZHENG, H.; LIU, J.; FAN, R.; ZHU, W.; YUAN, L.; QIN, Y.; CHEN, S.; ZHOU, Y.; WU, Y.; WAN, J.; MI, M.; WANG, J. Therapeutic effects of soluble dietary fiber consumption on type 2 diabetes mellitus. **Experimental and Therapeutic Medicine**, v. 12, n. 2, p. 1232–1242, 2016

CHEUNG, S. W.; BI, W. Novel applications of array comparative genomic hybridization in molecular diagnostics. **Expert Review of Molecular Diagnostics**, v. 18, n. 6, p. 531–542, 2018.

DA SILVA, R. M.; PUPIN, B.; BHATTACHARJEE, T. T.; VAMONDES KULCSAR, M. A.; UNO, M.; CHAMMAS, R.; DE AZEVEDO CANEVARI, R. ATR-FTIR spectroscopy and CDKN1C gene expression in the prediction of lymph nodes metastases in papillary thyroid carcinoma. **Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy**, v. 228, p. 117693, 2020.

DE SOUZA, B. M.; BRONDANI, L. A.; BOUÇAS, A. P.; SORTICA, D. A.; KRAMER, C. K.; CANANI, L. H.; LEITÃO, C. B.; CRISPIM, D. Associations between UCP1 -3826A/G, UCP2 -866G/A, Ala55Val and Ins/Del, and UCP3 -55C/T Polymorphisms and Susceptibility to Type 2 Diabetes mellitus: case-control study and meta-analysis. **PLoS ONE**, v. 8, n. 1, 2013.

DIN, I.; MAJID, S.; RASHID, F. *et al.* Mitochondrial uncoupling protein 2 (UCP2) gene polymorphism -866 G/A in the promoter region is associated with type 2 diabetes mellitus among Kashmiri population of Northern India. **Molecular Biology Reports**, v. 50, p. 475–483, 2023

DONG, L.; DUAN, X.; BIN, L.; WANG, J.; GAO, Q.; SUN, X.; XU, Y. Evaluation of Fourier transform infrared (FTIR) spectroscopy with multivariate analysis as a novel diagnostic tool for lymph node metastasis in gastric cancer. **Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy**, v. 289, p. 122209, 2023.

DORA, J. M.; MACHADO, W. E.; RHEINHEIMER, J.; CRISPIM, D.; MAIA, A. L. Association of the type 2 deiodinase Thr92Ala polymorphism with type 2 diabetes: Case-control study and meta-analysis. **European Journal of Endocrinology**, v. 163, n. 3, p. 427–434, 2010.

DU, Y.; HUA, Z.; LIU, C.; LV, R.; JIA, W.; SU, M. ATR-FTIR combined with machine learning for the fast non-targeted screening of new psychoactive substances. **Forensic Science International**, v. 349, p. 111761, 2023.

ELMI, F.; MOVAGHAR, A. F.; ELMI, M. M.; ALINEZHAD, H.; NIKBAKHS, N. Application of FTIR spectroscopy on breast cancer serum analysis. **Spectrochimica Acta - Part A: Molecular and Biomolecular Spectroscopy**, v. 187, p. 87–91, 2017..

ETCHEGOYEN, M.; NOBILE, M. H.; BAEZ, F.; POSESORSKI, B.; GONZÁLEZ, J.; LAGO, N.; MILEI, J.; OTERO-LOSADA, M. Metabolic syndrome and neuroprotection. **Frontiers in Neuroscience**, v. 12, p. 196, 2018.

FAN, Y.; LI, X.; ZHANG, Y.; FAN, X.; ZHANG, N.; ZHENG, H.; SONG, Y.; SHEN, C.; SHEN, J.; REN, F.; YANG, J. Genetic variants of TPCN2 associated with type 2 diabetes risk in the Chinese population. **PLoS ONE**, v. 11, n. 2, 2016.

FAROOQ, S.; ZECELL, D. Diabetes monitoring through urine analysis using ATR-FTIR spectroscopy and machine learning. **Chemosensors**, v. 11, p. 565, 2023.

FERREIRA, D. F. SISVAR. A guide for its bootstrap procedures in multiple comparisons. **Ciênc. agrotec.**, v.38, n. 2, 2014.

FERREIRA, D. F. SISVAR: A computer analysis system to fixed effects split plot type designs. **Revista Brasileira de Biometria**, v. 37, n. 4, p. 529-535, 2019.

GHAFFAR, M. T. A. Association of aldosterone synthase CYP11B2 (-344C/T) gene polymorphism with essential hypertension and left ventricular hypertrophy in the Egyptian population. **Clinical and Experimental Hypertension**, v. 41, n. 8, p. 779-786, dez. 2018.

GHEZZI, C.; LOO, D. D. F.; WRIGHT, E. M. Physiology of renal glucose handling via SGLT1, SGLT2 and GLUT2. **Diabetologia**, v. 61, n. 10, p. 2087–2097, out. 2018.

GOMATHI, P.; SAMARTH, A. P.; RAJ, N. B. A. J.; SASIKUMAR, S.; MURUGAN, P. S.; NALLAPERUMAL, S.; SELVAM, G. S. The -866G/A polymorphism in the promoter of the UCP2 gene is associated with risk for type 2 diabetes and with decreased insulin levels. **Gene**, v. 701, p. 125–130, 2019.

GOMES RIOS, T.; LARIOS, G.; MARANGONI, B.; OLIVEIRA, S. L.; CENA, C.; ALBERTO DO NASCIMENTO RAMOS, C. FTIR spectroscopy with machine learning: A new approach to animal DNA polymorphism screening. **Spectrochimica Acta - Part A: Molecular and Biomolecular Spectroscopy**, v. 261, p. 120036, 2021.

GUANG, P.; HUANG, W.; GUO, L.; YANG, X.; HUANG, F.; YANG, M.; WEN, W.; LI, L.; YAN, L. J. Blood-based FTIR-ATR spectroscopy coupled with extreme gradient boosting for the diagnosis of type 2 diabetes: a STARD compliant diagnosis research. **Medicine (United States)**, v. 99, n. 15, e19657, 2020.

GUEVARA, E.; TORRES-GALVÁN, J. C.; RAMÍREZ-ELÍAS, M. G.; LUEVANO-CONTRERAS, C.; GONZÁLEZ, F. J. Use of Raman spectroscopy to screen diabetes mellitus with machine learning tools. **Biomedical Optics Express**, v. 9, n. 10, p. 4998, 2018.

GUIDUCCI, L.; IERVASI, G.; QUINONES-GALVAN, A. On the paradox insulin resistance/insulin hypersensitivity and obesity: Two tales of the same history. **Expert Review of Cardiovascular Therapy**, v. 12, n. 6, p. 637–642, 2014.

HACKETT, R. A.; STEPTOE, A. Type 2 diabetes mellitus and psychological stress: a modifiable risk factor. **Nature Reviews Endocrinology**, v. 13, n. 9, p. 547–560, 2017.

HACKETT, R. A.; STEPTOE, A.; KUMARI, M. Association of diurnal patterns in salivary cortisol with type 2 diabetes in the Whitehall II study. **Journal of Clinical Endocrinology and Metabolism**, v. 99, n. 12, p. 4625–4631, 2014.

HARREITER, J.; RODEN, M. Diabetes mellitus: definition, classification, diagnosis, screening and prevention (Update 2023). **Wiener Klinische Wochenschrift**, v. 135, p. 7–17, 2023.

HÂRȚA, M.; BORSAL, O.; MUNTEAN, C. M.; DINA, N. E.; FĂLĂMAȘ, A.; OLAR, L. E.; SZABO, K.; PAMFIL, D.; ȘTEFAN, R. Assessment of genetic relationships between *Streptocarpus x hybridus* v. parents and F1 progenies using SRAP markers and FTIR spectroscopy. **Plants**, v. 9, n. 2, 2020.

HASHEMI, M.; REZAEI, H.; KAYKHAEI, M. A.; TAHERI, M. A 45-bp insertion/deletion polymorphism of UCP2 gene is associated with metabolic syndrome. **Journal of Diabetes & Metabolic Disorders**, v. 13, n. 1, p. 12, 2014.

HOŞAFÇI, G.; KLEIN, O.; OREMEK, G.; MÄNTELE, W. Clinical chemistry without reagents? An infrared spectroscopic technique for determination of clinically relevant constituents of body fluids. **Analytical Bioanalytical Chemistry**, v. 387, n. 5, p. 1815–1822, 2007.

HOU, G.; JIN, Y.; LIU, M.; WANG, C.; SONG, G. UCP2–866G/A Polymorphism is associated with prediabetes and type 2 diabetes. **Archives of Medical Research**, v. 51, n. 6, p. 556–563, 2020.

HUANG, R.; CAI, T.; ZHOU, Y.; WANG, Y.; WANG, H.; SHEN, Z.; XIA, W.; LIU, X.; DING, B.; LUO, Y.; YAN, R.; LI, H.; WU, J.; MA, J. Ethnicity differences in the association of UCP1-3826A/G, UCP2-866G/A and Ala55Val, and UCP3-55C/T polymorphisms with type 2 diabetes mellitus susceptibility: An updated meta-analysis. **BioMed Research International**, v. 2021, 2021.

INTERNATIONAL DIABETES FEDERATION. **Diabetes Atlas 2021**. Disponível em: <https://diabetesatlas.org/>. Acesso em: 20 dez. 2024.

ISHIMARU, S. *et al.* Effect of UGT2B17 deletion polymorphism on prognosis in pediatric cancer. **Pediatrics International**, v. 59, n. 4, p. 427-431.

KAHN, T. R.; FONG, K. K.; JORDAN, B.; LEK, J. C.; LEVITAN, R.; MITCHELL, P. S.; WOOD, C.; HATCHER, M. E. An FTIR investigation of flanking sequence effects on the structure and flexibility of DNA binding sites. **Biochemistry**, v. 48, n. 6, p. 1315–1321, 2009.

KAZARIAN, S. G.; CHAN, K. L. A. Applications of ATR-FTIR spectroscopic imaging to biomedical samples. **Biochimica et Biophysica Acta - Biomembranes**, v. 1758, n. 7, p. 858–867, 2006.

KINO, S.; KANAMORI, M.; SHIMODA, Y.; NIIZUMA, K.; ENDO, H.; MATSUURA, Y. Distinguishing IDH mutation status in gliomas using FTIR-ATR spectra of peripheral blood plasma indicating clear traces of protein amyloid aggregation. **BMC Cancer**, v. 24, n. 1, 2024.

KIRBY, K. S. A new method for the isolation of deoxyribonucleic acids; evidence on the nature of bonds between deoxyribonucleic acid and protein. **Biochemical Journal**, v. 66, n. 3, p. 495-504, 1957.

KOEHLER, A.; SCROFERNEKER, M. L.; PEREIRA, B. A. S.; PEREIRA DE SOUZA, N. M.; DE SOUZA CAVALCANTE, R.; MENDES, R. P.; CORBELLINI, V. A. Using infrared spectroscopy of serum and chemometrics for diagnosis of paracoccidioidomycosis. **Journal of Pharmaceutical and Biomedical Analysis**, v. 221, p. 115021, 2022.

KOKOT, I.; MAZUREK, S.; PIWOWAR, A.; SZOSTAK, R.; JĘDRYKA, M.; KRATZ, E. M. ATR-IR Spectroscopy application to diagnostic screening of advanced endometriosis. **Oxidative Medicine and Cellular Longevity**, v. 2022, p. 4777434, 2022.

KREMPLER, F.; ESTERBAUER, H.; WEITGASSER, R.; EBENBICHLER, C.; PATSCH, J. R.; MILLER, K.; XIE, M.; LINNEMAYR, V.; OBERKOFER, H.; PATSCH, W. A functional polymorphism in the promoter of UCP2 enhances obesity risk but reduces type 2 diabetes risk in obese middle-aged humans. **Diabetes**, v. 51, 2002

KUO, T.; MCQUEEN, A.; CHEN, T. C.; WANG, J. C. Regulation of glucose homeostasis by glucocorticoids. **Advances in Experimental Medicine and Biology**, v. 872, p. 99–126, 2015.

LAPIK, I. A.; RANJIT, R.; GALCHENKO, A. V. Impact of KCNJ11 rs5219, UCP2 rs659366, and MTHFR rs1801133 Polymorphisms on Type 2 Diabetes: A Cross-Sectional Study. **The Review of Diabetics Studies**, v. 17, n. 1, p. 21-29, 2021.

LEE, S. H.; PARK, S. Y.; CHOI, C. S. Insulin resistance: from mechanisms to therapeutic strategies. **Diabetes and Metabolism Journal**, v. 46, n. 1, p. 15–37, 2022.

LESKOVAR, T.; ZUPANIČ PAJNIČ, I.; GERŠAK, Ž. M.; JERMAN, I.; ČREŠNAR, M. ATR-FTIR spectroscopy combined with data manipulation as a pre-screening method to assess DNA preservation in skeletal remains. **Forensic Science International: Genetics**, v. 44, 2020.

LEÓN-LATRE, M.; MORENO-FRANCO, B.; ANDRÉS-ESTEBAN, E. M.; LEDESMA, M.; LACLAUSTRA, M.; ALCALDE, V.; PEÑALVO, J. L.; ORDOVÁS, J. M.; CASASNOVAS, J. A. Sedentary lifestyle and its relation to cardiovascular risk factors, insulin resistance and inflammatory profile. **Revista Española de Cardiología (English Edition)**, v. 67, n. 6, p. 449–455, 2014.

LEWIS, P. D.; LEWIS, K. E.; GHOSAL, R.; BAYLISS, S.; LLOYD, A. J.; WILLS, J.; GODFREY, R.; KLOER, P.; MUR, L. A. J. Evaluation of FTIR Spectroscopy as a diagnostic tool for lung cancer using sputum. **BMC Cancer**, v. 10, p. 640, 2010.

LEY, S. H.; HAMDY, O.; MOHAN, V.; HU, F. B. Prevention and management of type 2 diabetes: dietary components and nutritional strategies. **The Lancet**, v. 383, n. 9933, p. 1999–2007, 2014.

LI, Q.-B.; XU, Z.; ZHANG, N.-W.; ZHANG, L.; WANG, F.; YANG, L.-M.; WANG, J.-S.; ZHOU, S.; ZHANG, Y.-F.; ZHOU, X.-S.; SHI, J.-S.; WU, J.-G.; LI, B. In vivo and in situ detection of colorectal cancer using fourier transform infrared spectroscopy. **World Journal of Gastroenterology**, v. 11, 2005.

LI, Z.; LV, H.; LI, T.; SI, G.; WANG, Q.; LV, J.; HU, X. Reagent-free simultaneous determination of glucose and cholesterol in whole blood by FTIR-ATR. **Spectrochimica Acta - Part A: Molecular and Biomolecular Spectroscopy**, v. 178, p. 192–197, 2017.

LIEMBURG-APERS, D. C.; IMAMURA, H.; FORKINK, M.; NOOTEBOOM, M.; SWARTS, H. G.; BROCK, R.; SMEITINK, J. A. M.; WILLEMS, P. H. G. M.; KOOPMAN, W. J. H. Quantitative glucose and ATP sensing in mammalian cells. **Pharmaceutical Research**, v. 28, n. 11, p. 2745–2757, 2011

LIGTHART, S.; HASBANI, N. R.; AHMADIZAR, F.; VAN HERPT, T. T. W.; LEENING, M. J. G.; UITTERLINDEN, A. G.; SIJBRANDS, E. J. G.; MORRISON, A. C.; BOERWINKLE, E.; PANKOW, J. S.; SELVIN, E.; ARFAN IKRAM, M.; KAVOUSI, M.; DE VRIES, P. S.; DEGHAN, A. Genetic susceptibility, obesity and lifetime risk of type 2 diabetes: The ARIC study and Rotterdam Study. **Diabetic Medicine**, v. 38, n. 10, 2021.

LIU, J.; LI, J.; LI, W. J.; WANG, C. M. The role of uncoupling proteins in diabetes mellitus. **Journal of Diabetes Research**, v. 2013, p. 585897, 2013.

LOUZADA, M. L. C. ; CRUZ, G. L.; SILVA, K. A. A. N.; GRASSI, A. G. F.; ANDRADE, G. C.; RAUBER, F.; LEVY, R. B.; MONTEIRO, C. A. Consumption of ultra-processed foods in Brazil: distribution and temporal evolution 2008–2018. **Revista de Saúde Pública**, v. 57, p. 12, 2023.

LYSSENKO, V.; ALMGREN, P.; ANEVSKI, D.; ORHO-MELANDER, M.; SJÖGREN, M.; SALORANTA, C.; TUOMI, T.; GROOP, L. Genetic prediction of future type 2 diabetes. **PLoS Medicine**, v. 2, n. 12, p. 1299–1308, 2005.

MAILLOUX, R. J., HARPER, M. -E. Uncoupling proteins and the control of mitochondrial reactive oxygen species production. **Free Radical Biology and Medicine**, v. 51, n. 6, p. 1106–1115, set. 2011.

MALONE, J. I.; HANSEN, B. C. Does obesity cause type 2 diabetes mellitus (T2DM)? or is it the opposite? **Pediatric Diabetes**, v. 20, n. 1, p. 5–9, 2019.

MARTÍNEZ-LÓPEZ, E.; RUÍZ-MADRIGAL, B.; HERNÁNDEZ-CAÑAVERAL, I.; PANDURO, A. Association of the T54 allele of the FABP2 gene with cardiovascular risk factors in obese Mexican subjects. **Diabetes and Vascular Disease Research**, v. 4, n. 3, p. 235–236, 2007.

MATEUS PEREIRA DE SOUZA, N.; HUNTER MACHADO, B.; KOCHE, A.; BEATRIZ FERNANDES DA SILVA FURTADO, L.; BECKER, D.; ANTONIO CORBELLINI, V.; RIEGER, A. Detection of metabolic syndrome with ATR-FTIR spectroscopy and chemometrics in blood plasma. **Spectrochimica Acta - Part A: Molecular and Biomolecular Spectroscopy**, v. 288, 2023.

MATSUDA, K. PCR-based detection methods for single-nucleotide polymorphism or mutation: real-time PCR and its substantial contribution toward technological refinement. **Advances in Clinical Chemistry**, v. 80, p. 45-72, 2017.

MEDRANO, R. F. V.; DE OLIVEIRA, C. A. Guidelines for the tetra-primer ARMS-PCR technique development. **Molecular Biotechnology**, v. 56, n. 7, p. 599–608, 2014.

MITURA, P.; PAJA, W.; KLEBOWSKI, B.; PŁAZA, P.; BAR, K.; MŁYNARCZYK, G.; DEPCIUCH, J. Urine analysed by FTIR, chemometrics and machine learning methods in determination spectroscopy marker of prostate cancer in urine. **Journal of Biophotonics**.2024

MOORE, A. F.; FLOREZ, J. C. Genetic susceptibility to type 2 diabetes and implications for antidiabetic therapy. **Annual Review of Medicine**, v. 59, p. 95–111, 2008.

MUZY, J.; CAMPOS, M. R.; EMMERICK, I.; DA SILVA, R. S.; DE ANDRADE SCHRAMM, J. M. Prevalence of diabetes mellitus and its complications and characterization of healthcare gaps based on triangulation of studies. **Cadernos de Saúde Pública**, v. 37, n. 5, 2021.

MORCILLO, S.; CARDONA, F.; ROJO-MARTÍNEZ, G.; ALMARAZ, M. C.; ESTEVA, I.; RUIZ-DE-ADANA, M. S.; OLVEIRA, G.; GARCÍA-FUENTES, E.; GÓMEZ-ZUMAQUERO, J. M.; SORIGUER, F. Effect of the combination of the variants -75G/A APOA1 and Trp64Arg ADRB3 on the risk of type 2 diabetes (DM2). **Clinical Endocrinology**, v. 68, n. 1, p. 102–107, 2008.

NURDALILA, A. A.; BUNAWAN, H.; KUMAR, S. V.; RODRIGUES, K. F.; BAHARUM, S. N. Homogeneous nature of Malaysian Marine fish *Epinephelus fuscoguttatus* (Perciformes; Serranidae): Evidence based on molecular markers, morphology and Fourier transform infrared analysis. **International Journal of Molecular Sciences**, v. 16, n. 7, p. 14884–14900, 2015.

NUSSBAUM, R. L.; MCINNES, R. R.; WILLARD, H. F. **Thompson & Thompson genética médica**. 8. ed. Rio de Janeiro: Elsevier, 2016.

ORGANIZAÇÃO MUNDIAL DA SAÚDE. **Obesity**: preventing and managing the global epidemic. Report of a WHO consultation on obesity. Geneva: World Health Organization, 1997.

ORTIZ, R.; KLUWE, B.; ODEI, J. B.; ECHOUFFO TCHEUGUI, J. B.; SIMS, M.; KALYANI, R. R.; BERTONI, A. G.; GOLDEN, S. H.; JOSEPH, J. J. The association of morning serum cortisol with glucose metabolism and diabetes: The Jackson Heart Study. **Psychoneuroendocrinology**, v. 103, p. 25–32, 2019.

PAIS, F. C.; ARAUJO, G. Aplicações ópticas: infravermelho próximo para medição de glicose / optical applications: near infrared for measuring glucose.

PEER, N.; BALAKRISHNA, Y.; DURAO, S. Screening for type 2 diabetes mellitus. **Cochrane Database of Systematic Reviews**, v. 5, n. 5, p. CD005266, 2020.

PEREZ-LUQUE, E.; MALACARA, J. M.; GARAY-SEVILLA, M. E.; FAJARDO, M. E. Association of the TNF- α -308G/A polymorphism with family history of type 2 diabetes mellitus in a Mexican population. **Clinical Biochemistry**, v. 45, n. 1–2, p. 12–15, 2012.

PETERSEN, M. C.; SHULMAN, G. I. Mechanisms of insulin action and insulin resistance. **Physiological Reviews**, v. 98, p. 2133–2223, 2018.

PIPPI, R.; CUGUSI, L.; BERGAMIN, M.; BINI, V.; FANELLI, C. G.; BULLO, V.; GOBBO, S.; DI BLASIO, A. Impact of BMI, physical activity, and sitting time levels on health-related outcomes in a group of overweight and obese adults with and without Type 2 Diabetes. **Journal of Functional Morphology and Kinesiology**, v. 7, n. 1, 2022.

QIN, L. J.; WEN, J.; QU, Y. L.; HUANG, Q. Y. Lack of association of functional UCP2 - 866G/A and Ala55Val polymorphisms and type 2 diabetes in the Chinese population based on a case-control study and a meta-analysis. **Genetics and Molecular Research**, v. 12, n. 3, p. 3324–3334, 2013.

REEDS, J.; MANSURI, S.; MAMAKEESICK, M.; HARRIS, S. B.; ZINMAN, B.; GITTELSON, J.; WOLEVER, T. M. S.; CONNELLY, P. W.; HANLEY, A. Dietary patterns and type 2 diabetes mellitus in a first nations community. **Canadian Journal of Diabetes**, v. 40, n. 4, p. 304–310, 2016.

REIN, L. C. D. S.; SIQUEIRA, D. E. D.; GUILLAUMON, A. T.; AVELAR, W. M.; CENDES, F.; MESQUITA, R. C. Evaluation of the brain hemodynamic response by means of near-infrared spectroscopy (NIRS) monitoring in patients with atherosclerotic carotid disease undergoing endarterectomy. **Jornal Vascular Brasileiro**, v. 19, 2020.

REYNOLDS, A.; MANN, J.; CUMMINGS, J.; WINTER, N.; METE, E.; TE MORENGA, L. Carbohydrate quality and human health: a series of systematic reviews and meta-analyses. **The Lancet**, v. 393, n. 10170, p. 434–445, 2019.

REZAPOUR, S. *et al.* Association of 45-bp ins/del polymorphism of uncoupling protein 2 (UCP2) and susceptibility to nonalcoholic fatty liver and type 2 diabetes mellitus in North-West of Iran. **BMC Research Notes**, dez. 2021.

RICHARDSON, Z.; KINCSES, A.; EKINCI, E.; PEREZ-GUAITA, D.; JANDELEIT-DAHM, K.; WOOD, B. R. ATR-FTIR spectroscopy for early detection of diabetic kidney disease. **Analysis & Sensing**, v. 3, n. 4, e202200094, 27 dez. 2022.

RIOS, D. L. S. *et al.* Interleukin-1 beta and interleukin-6 gene polymorphism associations with angiographically assessed coronary artery disease in Brazilians. **Cytokine**, v. 50, n. 3, p. 292–296, jun. 2010.

RUFINO, H. L. P.; VEIGA, A. C. P.; NAKAMOTO, P. T. SMOTE_EASY: um algoritmo para tratar o problema de classificação em bases de dados reais. **JISTEM - Journal of Information Systems and Technology Management**, v. 13, p. 61–80, 2016.

SALAUDDIN, A.; CHAKMA, K.; HASAN, M. M.; AKTER, F.; CHOWDHURY, N. A.; CHOWDHURY, S. R.; MANNAN, A. Association between TCF7L2 polymorphism and type 2 diabetes mellitus susceptibility: a case–control study among the Bangladeshi population. **Molecular Biology Reports**, v. 50, p. 609–619, 2023.

SANCHEZ, M. P.; PRIETO, C.; MUJICA, E.; VERGARA, K.; VALENCIA, E.; VILLALOBOS, E.; MEDINA, M.; PARRA, M.; D'ADDOSIO, R.; HOEDEBECKE, K.; RODRÍGUEZ, J. E.; BERMUDEZ, V. Association between +45t>g adiponectin polymorphism gene and type 2 diabetes mellitus and metabolic syndrome in a Venezuelan population. **F1000Research**, v. 8, p. 292, 2019.

SÁNCHEZ-BRITO, M.; LUNA-ROSAS, F. J.; MENDOZA-GONZÁLEZ, R.; MATA-MIRANDA, M. M.; MARTÍNEZ-ROMO, J. C.; VÁZQUEZ-ZAPIÉN, G. J. A machine-learning strategy to evaluate the use of FTIR spectra of saliva for a good control of type 2 diabetes. **Talanta**, v. 221, 2021.

SASAHARA, M.; NISHI, M.; KAWASHIMA, H.; UEDA, K.; SAKAGASHIRA, S.; FURUTA, H.; MATSUMOTO, E.; HANABUSA, T.; SASAKI, H.; NANJO, K. Uncoupling protein 2 promoter polymorphism 866G/A affects its expression in cells and modulates clinical profiles of Japanese type 2 diabetic patients. **Diabetes**, v. 53, p. 482, 2004.

SELVIN, E.; CRAINICEANU, C. M.; BRANCATI, F. L.; CORESH, J. Short-term variability in measures of glycemia and implications for the classification of diabetes. **Arch. Intern. Med.** v. 167, n. 14, p. 1545-1551, 2007.

SESTI, G.; CARDELLINI, M.; MARINI, M. A.; FRONTONI, S.; ADAMO, M. D.; GUERRA, S. DEL; LAURO, D.; DE NICOLAIS, P.; SBRACCIA, P.; PRATO, S. DEL; GAMBARDELLA, S.; FEDERICI, M.; MARCHETTI, P.; LAURO, R. A common polymorphism in the promoter of UCP2 contributes to the variation in insulin secretion in glucose-tolerant subjects. **Diabetes**, v. 52, 2003.

SHAMOON, H. Pathophysiology of diabetes: a review of selected recent developments and their impact on treatment. **Drugs**, v. 44, 1992.

SHEN, Y.; WEN, Z.; WANG, N.; ZHENG, Z.; LIU, K.; XIA, X.; GU, Q.; SHI, Y.; XU, X. Investigation of variants in UCP2 in Chinese type 2 diabetes and diabetic retinopathy. **PLoS ONE**, v. 9, n. 11, 2014.

SOCIEDADE BRASILEIRA DE CARDIOLOGIA. Diretriz Brasileira de Diagnóstico e Tratamento da Síndrome Metabólica, 1. **Arquivos Brasileiros de Cardiologia**, v. 84, supl. 1, p. 1-28, 2005.

SOCIEDADE BRASILEIRA DE DIABETES. **Diretrizes da Sociedade Brasileira de Diabetes: 2024-2025**. 2024. Disponível em: <https://www.diabetes.org.br/>. Acesso em: 10 nov. 2024.

SONG, S.; WANG, Q.; ZOU, X.; LI, Z.; MA, Z.; JIANG, D.; FU, Y. Q.; LIU, Q. High-precision prediction of blood glucose concentration utilizing Fourier transform Raman spectroscopy and an ensemble machine learning algorithm. **Spectrochimica Acta - Part A: Molecular and Biomolecular Spectroscopy**, v. 303, 2023.

TAILLANDIER, E.; LIQUIER, J. Infrared spectroscopy of DNA. **Methods in Enzymology**, v. 211, p. 307-335, 1992.

TOKARZ, V. L.; MACDONALD, P. E.; KLIP, A. The cell biology of systemic insulin function. **Journal of Cell Biology**, v. 217, n. 7, p. 2273–2289, 2018.

TOMAS, R. C.; SAYAT, A. J.; ATIENZA, A. N.; DANGANAN, J. L.; RAMOS, M. R.; FELLIZAR, A.; NOTARTE, K. I.; ANGELES, L. M.; BANGAOIL, R.; SANTILLAN, A.; ALBANO, P. M. Detection of breast cancer by ATR-FTIR spectroscopy using artificial neural networks. **PLoS ONE**, v. 17, n. 1, 2022.

UASKA SARTORI, P. V.; PENNA, G. O.; BÜHRER-SÉKULA, S.; PONTES, M. A. A.; GONÇALVES, H. S.; CRUZ, R.; VIRMOND, M. C. L.; DIAS-BAPTISTA, I. M. F.; ROSA, P. S.; PENNA, M. L. F.; MEDEIROS FAVA, V.; STEFANI, M. M. A.; TÁVORA MIRA, M. Human genetic susceptibility of leprosy recurrence. **Scientific Reports**, v. 10, n. 1, 2020.

WANG, C.-H.; KE, W.-S.; LIN, E. Evaluation of the ENPP1 and PLIN single nucleotide polymorphisms with type 2 diabetes in a Taiwanese population. **Journal of Investigative Medicine**, v. 60, n. 8, p. 1169–1173, 2012.

WOBETO, V. P. A.; GARCIA, P. M. D.; ZACCARIOTTO, T. R.; SONATI, M. D. F. Haptoglobin polymorphism and diabetic nephropathy in Brazilian diabetic patients. **Annals of Human Biology**, v. 36, n. 4, p. 437–441, 2009.

WOLLINGER, L. M.; DAL BOSCO, S. M.; REMPEL, C.; ALMEIDA, S. E. M.; BERLESE, D. B.; CASTOLDI, R. P.; ARNDT, M. E.; CONTINI, V.; GENRO, J. P. Role of ACE and AGT gene polymorphisms in genetic susceptibility to diabetes mellitus type 2 in a Brazilian sample. **Genetics and Molecular Research**, v. 14, n. 4, p. 19110–19116, 2015..

XIE, T.; STATHOPOULOU, M. G.; DE ANDRÉS, F.; SIEST, G.; MURRAY, H.; MARTIN, M.; COBALEDA, J.; DELGADO, A.; LAMONT, J.; PENÑAS-LIEDÓ, E.; LLERENA, A.; VISVIKIS-SIEST, S. VEGF-related polymorphisms identified by GWAS and risk for major depression. **Translational Psychiatry**, v. 7, n. 3, 2017.

XU, L.; CHEN, S.; ZHAN, L. Association of uncoupling protein-2 -866G/A and Ala55Val polymorphisms with susceptibility to type 2 diabetes mellitus: A meta-analysis of case-control studies. **Medicine (Baltimore)**, v. 100, n. 6, p. e24464, 2021.

ZEIGERER, A.; SEKAR, R.; KLEINERT, M.; NASON, S.; HABEGGER, K. M.; MÜLLER, T. D. Glucagon's metabolic action in health and disease. **Comprehensive Physiology**, v. 11, n. 2, p. 1759–1783, 2021.

ZHANG, X.; XIAO, J.; YANG, F.; QU, H.; YE, C.; CHEN, S.; GUO, Y. Identification of sudden cardiac death from human blood using ATR-FTIR spectroscopy and machine learning. **International Journal of Legal Medicine**, v. 138, n. 3, p. 1139–1148, 2024.

ZHANG, X.; YANG, F.; XIAO, J.; QU, H.; JOCELIN, N. F.; REN, L.; GUO, Y. Analysis and comparison of machine learning methods for species identification utilizing ATR-FTIR spectroscopy. **Spectrochimica Acta - Part A: Molecular and Biomolecular Spectroscopy**, v. 308, p. 123713, 2024.

APÊNDICE A: Dados bioquímicos referente a níveis de glicose, HbA1c, insulina, colesterol total e triglicérides obtidos de amostras de sangue total coletados de 67 pacientes

Amostras	Glicemia (mg/dL)	HbA1c (%)	Insulina (uUI/mL)	Colesterol Total (mg/dL)	Triglicérides (mg/dL)
1	74	4,9	6	152	87
2	74	4,7	8	150	92
3	73	5,1	14,6	106	26
4	73	4,9	7,2	158	101
5	82	5,6	13,6	155	74
6	81	5,6	10,1	126	71
7	74	4,8	20,4	230	138
8	78	4,6	16,3	159	117
9	71	5	9,9	177	70
10	86	5,1	9,4	133	32
11	112	6	20,7	148	114
12	73	5,2	21	155	237
13	84	5,8	16,5	113	56
14	123	7,8	7,5	179	92
15	84	5,2	12,8	116	63
16	78	5	9	104	40
17	80	5,2	21,9	143	60
18	109	6,5	17,8	264	85
19	85	4,8	11,2	189	140
20	85	5,2	20,5	150	150
21	81	4,7	1,7	183	93
22	93	5,5	12	159	52
23	95	4,9	6,7	242	182
24	112	7,3	10,1	102	71
25	156	7,6	9	241	67
26	128	8,5	5,1	170	133
27	119	7,3	26,4	120	159
28	90	5,9	8,1	220	236
29	99	5,8	12,3		57
30	118	6,5	8,5	182	103
31	108	6,3	20,2	154	179
32	98	5,9		263	235
33	116	6,2	23,8	145	43
34	76	5,3	4,5	188	75
35	114	6,9	36,6	142	112
36	60	5,8	10,5	137	96
37	86	6,1	25,4	182	86
38	192,5	8,3	19,6	110	166

39	98,5	5,5	5,6	161	113
40	90,7	5,1	9,6	151	91
41	98	5,6	14,1	199	118
42	95,7	6	7	132	103
43	94,3	5	6,4	127	98
44	93,6	5,2	15	190	179
45	118,7	6,6	8	150	50
46	94,3	6	12,9		
47	93,6	5,1	18,7	194	79
48	97,5	5,3	32,2	166	90
49	140,6	8,9	30,9	245	265
50	88	4,9	7,9	175	115
51	134	7,7	12,8	208	277
52	85,7	5,4	17,1	245	117
53	105,6	5,1	9,7	167	75
54	95,7	4,8	22,9	161	65
55	92,4	5,4	17	162	185
56	87,2	4,5	17,3	153	282
57	90,3	4,6	15,1	167	81
58	96,1	5,3	19,6	167	137
59	148,8	9	14,9	197	207
60	142,9	7,3	25,8	217	164
61	94,7	6,1	28,5		
62	100,6	5,7	16,9	284	168
63	95,6	5,8	18,3	205	90
64	220,2	9,7	61,2	166	270
65	69	5,3	6,4	164	87
66	131,5	8,3	35,1	186	97
67	99,4	5,1	11,9	228	239

APÊNDICE B: Análises realizadas em amostras obtidas de sangue periférico de 67 participantes: análise molecular pela PCR-RFLP no DNA extraído da papa leucocitária; análises bioquímicas no plasma; análise de espectroscopia FTIR nas amostras de soro e as análises de FTIR aliada ao algoritmo *machine learning* realizada nas amostras de soro e amostras de DNA amplificado pela PCR.

Amostras	Análise molecular	Análise bioquímica	FTIR soro	FTIR <i>machine learning</i> soro	FTIR <i>machine learning</i> DNA
1	x	x	x	x	x
2	x	x	x	x	x
3	x	x	x	x	x
4	x	x	x	x	x
5	x	x			x
6	x	x	x	x	x
7	x	x			x
8	x	x	x	x	x
9	x	x	x	x	x
10	x	x	x	x	x
11	x	x	x	x	x
12	x	x	x	x	x
13	x	x	x	x	x
14	x	x			x
15	x	x	x	x	x
16	x	x			x
17	x	x	x	x	x
18	x	x			x
19	x	x			x
20	x	x			x
21	x	x	x	x	x
22	x	x	x	x	x
23	x	x	x	x	x
24	x	x	x	x	x
25	x	x	x	x	x
26	x	x	x	x	x
27	x	x	x	x	x
28	x	x	x	x	x
29	x	x	x	x	x
30	x	x	x	x	x
31	x	x	x	x	x
32	x	x	x	x	x
33	x	x			x
34	x	x			x
35	x	x			x
36	x	x			x
37	x	x			x
38		x			

39	x	x			x
40	x	x			x
41		x			
42		x	x	x	
43		x	x	x	
44		x	x	x	
45	x	x	x	x	x
46	x	x	x	x	x
47	x	x	x	x	x
48		x	x	x	
49		x	x	x	
50		x	x	x	
51	x	x	x	x	x
52	x	x	x	x	x
53	x	x	x	x	x
54	x	x	x	x	x
55	x	x	x	x	x
56	x	x	x	x	x
57	x	x	x	x	x
58	x	x	x	x	x
59	x	x	x	x	x
60	x	x	x	x	x
61	x	x	x	x	x
62	x	x	x	x	x
63	x	x	x	x	x
64	x	x	x	x	x
65	x	x	x	x	x
66	x	x	x	x	x
67	x	x	x	x	x

APÊNDICE C: Informações pessoais e histórico familiar dos 67 participantes pertencentes ao estudo

Amostra	Idade (anos)	Gênero	Raça/Cor	Descendência
1	18	Feminino	Branco	Brasileira
2	19	Feminino	Branco	Brasileira
3	18	Feminino	Parda	Brasileira
4	20	Feminino	Branco	Brasileira
5	18	Feminino	Branco	Brasileira
6	19	Masculino	Branco	Italiana
7	20	Feminino	Amarelo	Japonesa e brasileira
8	19	Feminino	Branco	Brasileira
9	18	Feminino	Branco	Brasileira
10	19	Feminino	Branco	Brasileira
11	67	Feminino	Branco	Brasileira
12	70	Feminino	Branco	Brasileira
13	48	Feminino	Branco	Brasileira
14	63	Feminino	Branco	Brasileira
15	21	Masculino	Branco	Brasileira
16	23	Feminino	Branco	Brasileira
17	23	Feminino	Branco	Brasileira
18	37	Feminino	Parda	Brasileira
19	23	Feminino	Branca	Brasileira
20	31	Feminino	Parda	Brasileira
21	28	Feminino	Branca	
22	47	Masculino	Branca	Italiana
23	46	Masculino	Branca	Italiana
24	70	Feminino	Branca	Português
25	75	Feminino	Branca	Português
26	76	Feminino	Branca	Italiana
27	74	Masculino	Branca	Não sabe informar
28	77	Masculino	Branca	Brasileira
29	72	Feminino	Branca	Italiana (mãe) e portuguesa (pai)
30	54	Feminino	Branca	Não sabe informar

31	66	Feminino	Branca	Italiana (avós paternos) e brasileira (Avós maternos)
32	54	feminino	Branca	Libanes
33	65	Masculino	Branca	Indiana
34	50	Feminino	Branca	Brasileira
35	67	Masculino	Branca	Brasileira
36	51	Feminino	Branca	Brasileira
37	50	Feminino	Pardo	Não sabe informar
38	34	Feminino	Pardo	Italiana
39	61	Feminino	Branca	Espanhol, Brasileira, Francês
40	65	Feminino	Branca	Portuguêsa
41	54	Feminino	Branca	Indígena, Portuguesa, Italiana
42	72	Masculino	Preto	Brasileiro
43	47	Feminino	Branco	Arabe
44	54	Feminino	Branco	Portugues e Frances
45	73	Feminino	Branco	Avó materna indígena, avô materno brasileiro
46	77	Feminino	X	Russo, espanhol, portuguesa, italiana
47	36	Feminino	Branco	Pai Europeu
48	19	Feminino	Branco	Brasileiro
49	50	Masculino	Branco	Brasileiro
50	20	Feminino	Branco	Alemão

51	70	Feminino	Branco	Materna(Polonês e Português) Paterna (Português e índio)
52	46	Feminino	Pardo	Paterna (Indígena) e Materna (Português)
53	21	Masculino	Branco	Não
54	18	Feminino	Branco	Polonesa e Espanhola
55	57	Masculino	Pardo	Indígena, Espanhola, Portuguesa
56	20	Feminino	Branco	Indígena
57	20	Feminino	Branco	Portugues e espanhol
58	62	Masculino	Branco	Libanesa (materna) e portuguesa (paterno)
59	52	Masculino	Branco	Italiano
60	60	Feminino	Branco	Brasileiro
61	41	Feminino	Branco	Indiígena (pai) e portuguesa (mãe)
62	37	Feminino	Branco	Não
63	40	Masculino	Pardo	Brasileiros
64	47	Masculino	Branco	Italiano (paterno)
65	32	Masculino	Branco	Italiano
66	34	Masculino	Branco	Italiano
67	65	Feminino	Branco	Italiana

Amostra	Exercício Físico	Alimentação Saudável	Fumante	Etilista
1	Sim	Sim	Não	Não
2	Sim	Sim	Não	Não
3	Sim	Sim	Não	Não
4	Sim	Sim	Não	Sim
5	Sim	Não	Não	Não
6	Sim	Não	Não	Não
7	Não	Sim	Não	Não
8	Não	Não	Não	Não
9	Sim	Sim	Não	Não
10	Sim	Não	Não	Não
11	Sim	Não	Sim	Não
12	Sim	Não	Não	Não
13	Sim	Sim	Não	Não
14	Não	Sim	Não	Não
15	Não	Não	Não	Não
16	Sim	Sim	Não	Não
17	Sim	Sim	Não	Não
18	Não	Sim	Não	Não
19	Não	Sim	Não	Não
20	Não	Sim	Não	Não
21	Sim	Sim	Não	Sim
22	Não	Sim	Não	Não
23	Sim	Sim	Não	Não
24	Sim	Sim	Não	Não
25	Não	Sim	Não	Não
26	Sim	Sim	Não	Não
27	Não	Sim	Não	Não
28	Não	Não	Não	Não
29	Não	Sim	Não	Não
30	Não	Sim	Não	Não
31	Sim	Sim	Não	Não
32	não	sim	não	não
33	sim	sim	não	não
34	Sim	Sim	Não	Não
35	Não	Sim	Não	Não
36	Sim	Sim	Não	Não
37	Não	Sim	Não	Não
38	Sim	Não	Não	Não
39	Sim	Sim	Sim	Sim
40	sim	Sim	Não	Não
41	Sim	Sim	Não	Sim
42	Sim	Sim	Não	Não
43	Sim	Sim	Não	Não

44	Não	Sim	Sim	Não
45	Sim	Sim	Não	Não
46	Sim	Sim	Não	Não
47	Não	Sim	Não	Não
48	Sim	Sim	Não	Não
49	Não	Não	Não	Não
50	Não	Não	Não	Não
51	Não	Sim	Não	Não
52	Não	Sim	Não	Não
53	Sim	Sim	Não	Não
54	Não	Sim	Não	Não
55	Não	Sim	Não	Não
56	Não	Sim	Não	Não
57	Sim	Sim	Sim	Não
58	Não	Sim	Não	Não
59	Sim	Sim	Não	Não
60	Não	Não	Não	Não
61	Não	Não	Não	Não
62	Sim	Sim	Sim	Não
63	Não	Sim	Não	Sim
64	Não	Sim	Não	Não
65	Sim	Sim	Não	Não
66	Sim	Não	Não	Não
67	Não	Sim	Não	Não

Amostra	Diabetes Individual	Diabetes na Família	Histórico de Glicose Alta	Medicamento para Pressão Alta	Obesidade na Família	Câncer Individual	Câncer na Família	Outras Doenças	Medicamento ou Suplementação Regular
1	Não	Sim (Pai DM2)	Não	Não	Sim	Não	Sim (Pai intestino)	Não	Não
2	Não	Sim (Avô materno DM2)	Não	Não	Sim (Avó materna)	Não	Não	Não	Não
3	Não	Não	Não	Não	Não	Não	Não	Não	Não
4	Não	Sim (Pai DM2)	Não	Não	Sim (Pai)	Não	Não	Não	Creatina, anticoncepcional
5	Não	Não	Não	Não	Sim	Não	Não	Não	Não
6	Não	Não	Sim	Não	Não	Não	Sim (Avô e avó paternos)	Não	Creatina, Whey
7	Não	Sim (Avó materna DM2)	Não	Não	Sim	Não	Não	Não	Não
8	Não	Sim (Avó materna DM2)	Não	Não	Não	Não	Sim (Avó materna estômago)	Tremor essencial	Não
9	Não	Não	Não	Não	Não	Não	Não	Não	Anticoncepcional
10	Não	Sim (Avó paterna DM2)	Não	Não	Não	Não	Não	Não	Não
11	Pré-diabética	Sim (Pai DM2)	Sim	Sim	Não	Não	Não	Hipertensão, dislipidemia, osteoporose	Rosuvastatina, bizoprolol, holmes
12	Não	Não	Não	Sim	Não	Não	Sim (utero e estômago)	Não	Não

13	Não	Não	Não	Não	Sim	Não	Não	Não	Não
14	Tipo 2 (Glifage 500mg, Jardiance 25 (1/dia))	Não	Sim	Sim	Não	Sim (Mama em 2007 e em 2013 - 2 tipos diferentes)	Não	Não há resposta	Não há resposta
15	Não	Sim (avó materna DM2)	Não	Não	Não	Não	Sim (avó materno de prostata)	Rinite	Não
16	Não	Sim (Avós materno e paterno DM2)	Não	Não	Não	Não	Não	Hipotiroidismo, Colesterol Alto (Genético?)	Puran T4 50mg, Desve, Plenance EZE
17	Não	Sim (Avó materno DM2 e pai pré-diabético)	Não	Não	Sim (Avó materno)	Não	Sim (Avó paterno garganta)	Não	Ômega 3, Rhodiola, L-Teanina, Coenzima Q10
18	Não	Sim (mãe e avó DM2)	Sim	Sim	Não	Não	Avó paterna (mama)	Rinite	-
19	Não	Sim (pai DM2 insulino dependente)	Não	Não	Não	Não	Não	Não (mãe pressão alta e pai glaucoma)	Fluoxetina 60mg e Vit. D
20	Não	Não	Não	Não	Sim	Não	Sim	Não	-
21	Tipo 2 (Metformina)	Sim	Sim	Não	Sim	Não	Sim	-	-
22	Não	Sim (Pai DM2)	Não	Sim	Sim (Pai, avó paterna)	Não	Sim (avó materno e 3 tios maternos próstata; Avó materno boca;intestino avó paterno)	Ansiedade, depressão, pressão alta, distúrbios do sono.	Alprazolam; Sertralina; Atenolol; Melatonina; Ômega 3
23	Não	Sim (mãe DM2)	Não	Não	Não	Não	Não	-	-
24	Tipo 2 (Glifage 500mg, glibenclamida, insulina)	Não sabe informar	Sim	Sim	Sim (pai)	Não	Sim (avó intestino, mama e pele mãe)	-	-
25	Tipo 2 (Glifage 500mg)	Não	Sim	Sim	Não	Não	Não	Não há resposta	Artrodar, losartana, hidroclorotiazida, levodopa, vitamina D)
26	Tipo 2 (Suganon 5mg, Glifage 500mg, insulina 12U)	Sim (Mãe e avó DM2)	Não	Sim	Não	Não	Não	-	janumet, selozok, glifage, insulina 10u, sinvastatina (dia s/n), pregabalina, mirtazapina, vit D 1x na semana
27	Tipo 2 (Glifage, glicazida, dapaglifozina, insulina NPH) - aos 70 anos	Sim (pai)	Sim	Sim	Sim	Sim	Não (mas irmãos sim)	Hipertensão, crise de gota, disfunção tireoide, gastrite/úlceras	Vitamina D, Losartana, Alopunirol, Levodopa, Omeoprazol

28	Tipo 2 (Glifage, Nesina) - aos 45 anos	Não sabe informar	Sim	Sim	Não	Não	Não	Guillain-Barré	Glucerna, Impere
29	Tipo 2 (Glicosida) - aos 45 anos	Sim (pai)	Sim	Não	Não	Sim (mama em 2019)	Não (mas irmã sim)	Não	Anastrozol, Glucerna
30	Tipo 2 - aos 46 anos	Sim (pai)	Sim	Não	Não	Não	Não	-	-
31	Tipo 2 (Trayenta e glifage)	Sim (mãe)	Sim	Sim (Losartana)	Não	Sim (mama em 2012)	Sim (avô paterno prostata)	Hipotiroidismo, hipercolesterolemia	Euthyrox 75; Pitavastatina
32	Tipo 2	não	sim	não	sim, mãe	não	não	não	met-morfina, clifage, insulina
33	Não, pré-diabético	sim, mãe	sim	não	sim	sim (Cérebro)	sim	não	vit D, sinvastatina
34	Não	Sim (mão e avó paterna)	Não	Não	Não	Sim (Tireoide)	Sim (mãe - pulmão, vô - intestino, pai e avô - próstata)	Não	Vit D, ferro, whey, synthroide
35	Tipo 2 - aos 55 anos	Sim (mãe)	Sim	Sim	Sim (mãe)	Sim (Tireoide)	Sim (pai tireoide, irmão tireoide, irmã intestino)	Não	Não
36	Não	Sim (Avô materno e mãe)	Sim	Sim (Corus 50mg; Novanlo 2,5mg)	Não sabe informar	Não	Sim (Pai prostata e Tia materna mama)	Hipertensão	Trezor 10mg; Glifage 500mg; Estradial, Estrial, Progesterona, Vagifen, DHEA; Ômega 3
37	Pré-diabétes	Não	Sim	Sim (Metiformina, alopimol, losartana)	Sim (mãe)	Não	Sim (Mãe útero)	Hipertensão, ácido urico elevado	Ômega 3, Vitamina D, Magnésio, curcuma
38	sim	sim (avô paterno)	Sim	Sim	Não	não	Não		
39	Não	Não sabe informar	Sim	Não	Não	Não	Sim (mãe mama, pai estômago)	Não	Óleo de borragem, cloridrato de vortioxetina
40	Não	Não	Não	Sim	Sim	Não	Não	Hepatite C, Cirurgia bariátrica, Retirada de visícula, Cirurgia da tireóide, Histerectomia	-
41	Pré-diabétes	Sim (pai, mãe e irmão)	Não	Sim	Sim	Não	Não sabe informar	Não	-
42	Não	Não	Não	Sim	Não	Não	Não	Não	-
43	Não	Sim (mãe e avós)	Não	Não	Sim	Não	Sim (avós)	Inflamações nos tendões	Creatina, Omega, vit. C, vit E, colageno tipo 2, complexo B, magnésio, anticoncepcional

44	Não	Sim (pai)	Não	Sim (losartana)	Sim (pai com sobrepeso)	Não	Sim (avó com cancer de pele)	Não	Antidepressivo (sertralina, omega 3)
45	Não	Não	Sim	Sim (losartana)	Não	Não	Não	Não	Ômega 3, vit E e suplementação para ossos
46	Não	Não sabe informar	Não	Não	Não	Não	Não	Não	Medicamentos naturais
47	Não	Sim (Avó materno e mãe)	Sim (durante gestação)	Não	Sim (Mãe)	Não	Avó materna (mama)	Não	Ômega 3
48	Não	Sim, (Avô paterno)	Não	Não	Não	Não	Não	Não	Não
49	Sim	Sim (Mãe)	Sim	Não	Sim	Não	Não	Ansiedade	Ansiolítico (Zoloft)
50	Não	Sim (Avô paterno)	Sim	Não	Não	Não	Avô - pulmão	Não	Não
51	Sim	Sim (Pai)	Sim	Sim (Losartana)	Não	Não	Mãe (câncer de mama e de pele) e Avó materna (Aparelho digestivo)	Osteoporose	Pregabalina, Losartana, Atenalol, Anlodipino, AAS.
52	Não	Não	Não	Não	Sim, avó paterna	Não	Mãe (câncer de esôfago), pai (câncer de próstata)	Não	Sertralina
53	Não	Sim (Pai)	Não	Não	Não	Não	Não	Não	Não
54	Não	Não	Sim	Sim	Sim	Não	Sim (Avó materna - fígado)	Não	Não
55	Não	Não	Sim	Sim	Não	Sim (Próstata)	Pai (Próstata e Pulmão)	Não	Losartana, Cetralina, Indapamina
56	Não	Não	Não	Não	Sim	Não	Sim (avó materna - mama)	Não	Não
57	Não	Sim (pai)	Sim	Não	Não	Não	Não	TDHA	Denvans
58	Sim	Não	Sim	Sim (Benicar)	Não	Sim (intestino)	Sim (mãe)	Não	Não
59	Sim (tipo 2)	Sim (mãe)	Sim	Não	Sim	Não	Não	Não	Não
60	Sim (tipo 2)	Sim (pai)	Sim	Sim (losartana)	Sim (mãe)	Sim (mama)	Sim (mãe)	Pressão alta, distúbio da tireoide	Levoid, glotalgona, fenofibrada e losartana
61	Não	Sim (mãe)	Não	Não	Sim	Não	Sim (mãe - mama, avó - ovário)	Colesterol alto e endometriose	Sinvastatina 20mg (colesterol) e anticoncepcional (endometriose)
62	Não	Não	Sim (durante gestação)	Não	Não	Não	Não	Ansiedade e Crise de pânico	Certalina e diazepam
63	Não	Sim (tia e avó - tipo 1)	Não	Não	Não	Não	Não	Não	Não
64	Sim (tipo 2)	Sim (mãe)	Sim	Sim (Lozartana)	Não	Não	Sim (irmã - intestino)	Cálculos renais, pressão alta e esteatose hepática	Lozartana, metiformina e insulina
65	Pré - diabético	Sim (Mãe - tipo 2)	Sim	Não	Não	Não	Sim (mãe - tumor gastrointestinal)	Esteatose hepática	

66	Não	Sim (Mãe - tipo 2)	Não	Não	Não	Não	Sim (mãe - tumor gastrointestinal)	Não	
67	Sim (tipo 2)	Sim (pai e irmão - tipo 2)	Sim	Não	Não	Sim (Gist no estômago)	Não	Anemia, artrite e artrose, hipotireoidismo	Ferro intravenoso, levotiroxina, metiformina, omeprazol e vitamina D3

APÊNDICE D: Análise quantitativa amostras de DNA e respectivos genótipos obtidos de 52 participantes e resultado da PCR-RFLP das amostras analisadas por ATR-FTIT

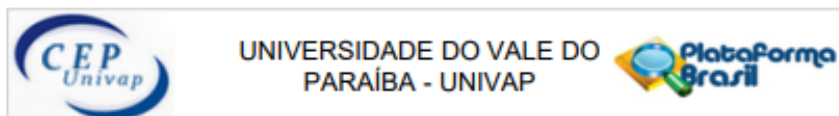
Amostra	Concentração do DNA (ng/μL)	Razão 260/280	Razão 260/230	Genótipo
1	202,98	1,62	1,42	G/G
2	145,27	1,42	1,14	G/A
3	162,66	1,83	1,82	G/G
4	210,24	1,51	1,27	G/A
6	57,5	1,41	0,88	G/G
8	216,48	1,71	1,62	G/G
9	128,72	1,47	0,94	A/A
10	41,2	1,69	1,46	A/A
11	754,5	1,54	1,66	G/A
12	1101,7	1,73	1,95	G/A
13	1422,7	1,72	1,93	A/A
15	411,7	1,81	1,74	G/A
17	1113,5	1,83	1,84	G/A
21	1098,6	1,92	2,32	A/A
22	139	1,91	1,88	A/A
23	178,4	1,85	2	G/G
24	401,1	1,88	1,43	G/G
25	535,8	1,74	1,7	G/G
26	465,6	1,77	1,53	G/G
27	170,8	1,79	1,38	G/G
28	236,8	1,77	1,72	G/G
29	291,4	1,73	1,74	G/A
30	237,5	1,82	1,49	A/A
31	483,7	1,83	1,17	G/G
32	514,5	1,72	0,91	A/A
42	670,24	1,92	2,21	A/A
43	118,99	1,99	1,87	A/A
44	114,37	1,85	1,77	G/A
45	470,86	1,89	2,18	A/A
46	1171,63	1,96	2,16	G/A
47	945,18	1,77	2,25	G/A
48	1575,39	1,9	2,27	A/A
49	984,01	1,99	1,89	A/A
50	1.479,04	1,96	2,11	A/A
51	177,5	1,83	1,31	G/A
52	2550,4	1,88	2,14	A/A
53	244,35	1,8	1,42	A/A
54	168,48	1,86	1,25	A/A
55	926,3	1,98	2,07	G/A
56	142,46	1,85	1,21	G/G
57	531,04	1,92	1,9	G/A

58	1671,26	2	2,12	G/A
59	1884,09	2,02	2,12	A/A
60	706,86	1,92	1,84	G/A
61	294,16	1,85	1,58	G/A
62	311,39	1,84	1,34	G/A
63	309,851	1,78	1,25	G/A
64	46,592	1,98	0,58	G/A
65	1313,897	2,03	2,01	G/G
66	310,372	1,94	1,57	G/A
67	2260,577	2,04	2,27	A/A

Legenda: Razões 260/280 e 260/230 na extração de DNA, utilizadas para avaliar a pureza da amostra.

A razão 260/280 indica contaminação por proteínas (ideal: 1,8–2,0), enquanto a 260/230 avalia contaminantes orgânicos (ideal: 2,0–2,2).

ANEXO A: Parecer do comitê de ética em Pesquisa



PARECER CONSUBSTANCIADO DO CEP

DADOS DA EMENDA

Título da Pesquisa: Relação do polimorfismo -866G/A do gene UCP2 com o desenvolvimento de distúrbios multifatoriais na população brasileira

Pesquisador: Renata de Azevedo Canevari

Área Temática: Genética Humana:

(Trata-se de pesquisa envolvendo Genética Humana que não necessita de análise ética por parte da CONEP.);

Versão: 4

CAAE: 62978022.0.0000.5503

Instituição Proponente: Universidade do Vale do Paraíba - UNIVAP

Patrocinador Principal: FUNDACAO DE AMPARO A PESQUISA DO ESTADO DE SAO PAULO

DADOS DO PARECER

Número do Parecer: 7.294.487

Apresentação do Projeto:

Trata-se de emenda a projeto aprovado por este CEP em 13/12/2022, parecer n.: 5.810.129. A emenda visa acrescentar no projeto seis membros, excluir dois membros e acrescentar as técnicas qPCR SNP genotyping e FTIR. Informa que não haverá nova coleta de dados e de amostras biológicas, visto que serão utilizadas as mesmas amostras de sangue. O tema do projeto, as formas de reduzir os riscos e os critérios de inclusão e exclusão descritos no projeto permanecerão os mesmos.

Objetivo da Pesquisa:

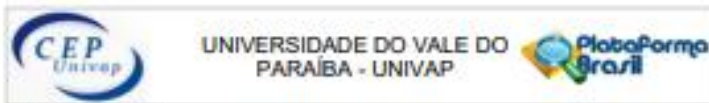
Objetivo Primário:

Analisar o polimorfismo -866G/A (rs659366) da região promotora do gene UCP2 pelas técnicas de PCR-RFLP, qPCR SNP genotyping e espectroscopia ATR-FTIR para investigar por meio de análises bioquímicas e clínicas, se o mesmo está relacionado com a predisposição ao desenvolvimento da DM2 e obesidade e a suscetibilidade ao câncer na população brasileira.

Objetivo Secundário:

- Analisar o polimorfismo -866G/A nos participantes do sexo feminino entre os grupos de baixo risco (grupo DA), risco moderado (grupo DB) e alto risco (grupo DC) para a DM2, segundo o

Endereço: Av. Shishima Hifumi, 2911 - Bloco 11 (IP&D), Sala 13
Bairro: Urbanova **CEP:** 12.244-000
UF: SP **Município:** SAO JOSE DOS CAMPOS
Telefone: (12)3947-1111 **E-mail:** cep@univap.br



Confirmação de Parecer: 7.290.837

Este parecer foi elaborado baseado nos documentos abaixo relacionados:

Tipo Documento	Arquivo	Postagem	Autor	Situação
Informações Básicas do Projeto	PB_INFORMAÇÕES_BÁSICAS_2413413_51.pdf	28/11/2024 18:09:33		Aceito
Outros	CartaRespostaCEP_28_11_2024.pdf	28/11/2024 18:57:34	Renata de Azevedo Canevari	Aceito
Outros	RelatórioParcial_28_11_2024.pdf	28/11/2024 18:56:53	Renata de Azevedo Canevari	Aceito
Outros	ConvitePesquisa28_11_2024.pdf	28/11/2024 18:54:45	Renata de Azevedo Canevari	Aceito
Outros	Termodeconfidencialidade_28_11_2024.pdf	28/11/2024 18:53:16	Renata de Azevedo Canevari	Aceito
Projeto Detalhado / Brochura Investigador	ProjetoCEP_28_11_2024.pdf	28/11/2024 18:51:36	Renata de Azevedo Canevari	Aceito
Cronograma	Cronogramadeexecução_28_11_2024.pdf	28/11/2024 18:51:34	Renata de Azevedo Canevari	Aceito
Orçamento	Orçamentox_28_11_2024.pdf	28/11/2024 18:50:55	Renata de Azevedo Canevari	Aceito
TCLE / Termos de Assentamento / Justificativa de Assinância	TCLCECP_28_11_24.pdf	28/11/2024 18:50:22	Renata de Azevedo Canevari	Aceito
Outros	Questionario_CEP_28_11.docx	25/11/2022 10:57:39	Renata de Azevedo Canevari	Aceito
Declaração de Instituição e Infraestrutura	Cartadeanuência_FC5.pdf	01/11/2022 17:24:39	Renata de Azevedo Canevari	Aceito
Outros	Convite_CEP_01_11.docx	01/11/2022 17:21:19	Renata de Azevedo Canevari	Aceito
Folha de Rosto	Folhaderocto.pdf	02/09/2022 12:01:34	Renata de Azevedo Canevari	Aceito

Situação do Parecer:

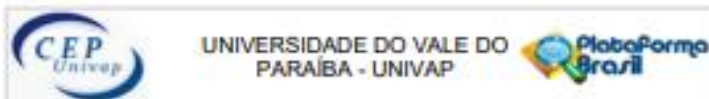
Aprovado

Necessita Apreciação da CONEP:

Não

Endereço: Av. São João Hilari, 2911 - Bloco 11 (PRAD), Sala 12
 Bairro: Ulsalóvia CEP: 12.240-000
 UF: SP Município: SÃO JOSÉ DOS CAMPOS
 Telefone: (12)2847-1111 E-mail: univap@univap.br

Página 02



Confirmação de Parecer: 7.290.837

SÃO JOSÉ DOS CAMPOS, 13 de Dezembro de 2024

Assinado por:
 Maurício Martins Alves
 (Coordenador(a))

ANEXO B: Termo de Consentimento Livre e Esclarecido

Dados de Identificação da Pesquisa

Título do projeto: Relação do polimorfismo -866G/A do *gene UCP2* com o desenvolvimento de distúrbios multifatoriais na população brasileira.

Pesquisador Responsável: Prof.^a Dra. Renata de Azevedo Canevari.

Equipe: Prof. MsC. Leoberto Lima e Prof. MsC. Matheus de Oliveira

Graduandos: Bianca Espindola Martins, Letícia Rodrigues Sobrinho, Clara Ribeiro de Oliveira, Aleph Vaz Arruda, João Victor Castro de Souza e Rayssa Gomes dos Santos.

Mestrandos: Igor Martins Alves Melo, Ramon Varella Diniz e Isabella Souza Silva.

Instituições: UNIVERSIDADE DO VALE DO PARAÍBA – Univap e Instituto de Pesquisa e Desenvolvimento – IP&D. Avenida Shishima Hifumi, 2911. CEP: 12244-000. São José dos Campos - SP Telefones para contato: 12-39471165 e 12-981591607

Nome do participante: _____

Data de nascimento: _____ CPF nº _____ Idade: _____ Gênero: _____

Telefone/celular: _____

E-mail (para receber os resultados dos exames e análises): _____

O (A) Sr.(a) está sendo convidado(a) a participar do projeto de pesquisa que tem como tema/título “Relação do polimorfismo -866g/a do *gene UCP2* com o desenvolvimento de distúrbios multifatoriais na população brasileira em desenvolvimento” de responsabilidade da pesquisadora Prof.^a Dra. Renata de Azevedo Canevari e financiamento da Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP Processo 2023/16314-6). O objetivo do estudo é coletar uma amostra do seu sangue (3 a 4 ml), identificá-la por código, armazená-la em geladeira ou freezer e neste material analisar o DNA como também realizar exames de sangue. No DNA será pesquisado a presença de uma alteração específica que pode ter relação com uma maior chance de se desenvolver a Diabetes Mellitus tipo 2, ou ter facilidade para ganho de peso ou talvez uma maior possibilidade de desenvolver futuramente um câncer em relação a população geral. Os exames de sangue realizados serão para avaliar a glicemia, insulina, hemoglobina glicada, colesterol e triglicérides. Como auxílio à esta pesquisa, com a sua permissão e após a assinatura deste termo, será aplicado à você um questionário para obtermos suas informações de histórico pessoal e familiar em relação as doenças que serão avaliadas neste estudo. Para a coleta de sangue, que será realizada por um biomédico no laboratório pertencente a universidade, será necessário que você esteja em jejum de 8 horas. Durante a sua participação, podem ocorrer hematoma, inflamação e alguma dor no local do braço em que foi coletado o sangue. Para minimizar isto, medidas de segurança serão tomadas, como a coleta de sangue em ambiente adequado e no local correto do corpo, higienização e descontaminação correta, ângulo e lado adequado da agulha corretos, uso de materiais estéreis e descartáveis, além do uso correto e completo dos equipamentos de proteção individual (EPI) pelo profissional que faz parte da Equipe de Pesquisa. Esperamos, que esta pesquisa possa auxiliar a adoção de hábitos de vida mais saudáveis, no sentido de prevenir o surgimento de algumas dessas doenças. Esclarecemos que os usos das informações por você oferecidas estão submetidas às normas éticas destinadas à pesquisa envolvendo seres humanos, da Comissão Nacional de Ética em Pesquisa (CONEP) do Conselho Nacional de Saúde, do Ministério da Saúde. A sua participação é de livre escolha e de forma anônima, em que permitirá a coleta de amostras do seu sangue para que sejam realizadas as análises descritas a partir da assinatura desta autorização. Informamos que há risco relativo a quebra de sigilo e confidencialidade na manipulação e publicação das informações obtidas. Contudo, para minimizar este risco iremos utilizar um código numérico correspondente ao seu nome e que será dado a você no momento da coleta de sangue no box do laboratório. Este código estará disponível em uma tabela que será consultada Página 2 de 2 apenas pelos pesquisadores do projeto para garantirmos a confidencialidade das informações e a sua privacidade. Contudo, vale ressaltar que durante o atendimento na recepção da CPS da FCS todos as informações serão coletadas como qualquer outro cliente

seguindo as normativas da Vigilância Sanitária citadas na RDC no 302/2005 que são: nome completo, endereço, sexo, idade, telefone para contato e e-mail para envio do laudo, que consta os resultados laboratoriais, se assim você desejar. As informações e materiais obtidos nesta pesquisa não poderão ser utilizados para outras finalidades que não sejam a desta pesquisa científica. Garantimos a você de se retirar desta pesquisa a qualquer momento, sem prejuízos ou sofrer quaisquer sanções ou constrangimentos. No caso de gastos decorrentes da participação nesta pesquisa como, por exemplo, transporte e alimentação, você e o seu acompanhante serão imediatamente e integralmente ressarcidos de todos os gastos. No caso de algum dano, imediato ou tardio, decorrente desta pesquisa, você também tem direito de ser indenizado pelo pesquisador desta pesquisa, bem como a ter assistência gratuita, integral e imediata. Sempre que desejar, você poderá entrar em contato para obter informações sobre este projeto de pesquisa, sobre sua participação ou resultados da pesquisa, com a pesquisadora responsável pelo telefone 12-39471165 ou email rcanevari@univap.br. Você também pode entrar em contato com o CEP – Comitê de Ética em Pesquisa da Universidade do Vale do Paraíba (UNIVAP), corresponsável por garantir e zelar pelos direitos do participante da pesquisa, pelo telefone (12) 3947-1111, pelo e-mail cep@univap.br ou pessoalmente na Av. Shishima Hifumi, 2911, Urbanova – Bloco 11 – Instituto de Pesquisa e Desenvolvimento II, sala 13, de segunda a sexta-feira, das 08:00h às 12:00h. Este termo está elaborado em duas vias, rubricadas em todas as suas páginas e assinadas, ao seu término, pelo participante da pesquisa e pelo pesquisador, sendo uma das vias entregue ao participante.

Eu, _____, fui informado e concordo em participar do projeto de pesquisa acima descrito.

São José dos Campos – SP, _____ de _____ de _____

Participante

Pesquisador

Prof.^a Dra. Renata de Azevedo Canevari

Testemunha nº 1
Prof. MsC. Leoberto de Lima

Testemunha nº 2
Prof.^a Dra. Sônia Khouri Sibe

ANEXO C: Questionário

PROJETO: RELAÇÃO DO POLIMORFISMO -866G/A DO *GENE UCP2* COM O DESENVOLVIMENTO DE DISTÚRBIOS MULTIFATORIAIS NA POPULAÇÃO BRASILEIRA

Data da coleta e preenchimento do questionário: ___/___/___

DADOS PESSOAIS

Nome: _____

Data de Nascimento: ___/___/___

Idade: _____

Gênero:

Masculino

Feminino

Raça/Cor:

Branco

Pardo

Preto

Indígena

Amarelo

Descendência(s) paterna e materna (italiana, indígena, africana e outras)

 Não sabe informar

DADOS FÍSICOS (preenchimento pelos pesquisadores)

Índice de Massa Corporal (IMC)? _____

Abaixo do peso (abaixo de 18,5Kg/m²)

Peso normal (18,5 - 24,9Kg/m²)

Sobrepeso (25 - 30Kg/m²)

Obesidade (acima de 30Kg/m²)

Peso: _____

Altura: _____

ESTILO DE VIDA

Você pratica pelo menos 30 minutos de atividade física diária no trabalho e/ou durante o horário de lazer?

Sim

Não

Com que frequência você come legumes, verduras, frutas ou grãos?

- Todos os dias
- Pelo menos 4 a 5 vezes por semana
- Pelo menos 2 a 3 vezes por semana
- Não consumo

Você é fumante (fuma ao menos um cigarro por dia todos os dias)?

- Sim
- Não

Você é etilista?

- Sim, bebo com frequência
- Bebo ocasionalmente
- Não consumo bebidas alcoólicas

C. INFORMAÇÕES CLÍNICAS

Você já foi diagnosticado com diabetes?

- Sim, Tipo 1
- Sim, Tipo 2
- Não, mas sou pré-diabético
- Não sou diabético
- Não sabe informar

Se sim, com qual idade foi diagnosticado? _____

Se sim, realiza qual tratamento? _____

Algum membro de sua família (irmãos, pais e avós) ou parente próximo já foi diagnosticado com diabetes (tipo 1 ou tipo 2)?

- Sim, há casos na família
- Não há casos na família
- Não sabe informar

Se sim, quem? (informar partes paterna e materna) _____

Você já apresentou glicose alta em algum exame médico?

- Sim
- Não

Prefere não informar

Se sim, qual (is) os exames? _____

Você já tomou regularmente algum medicamento para pressão alta?

Sim

Não

Você possui casos de pessoas com obesidade na família (pais e avós)?

Obesidade: pessoas com IMC acima de 30Kg/m², segundo a OMS

Sim, há casos na família

Não há casos na família

Não sabe informar

Prefere não informar

Você já teve câncer durante a sua vida?

Nunca

Sim

Se sim, qual(is) o(s) tipo(s)? _____

Você tem casos de câncer na família (pais e avós somente)?

Sim, há casos na família

Não há casos na família

Não sabe informar

Se sim, informar qual(is) o(s) tipo(s) e o membro familiar: _____

Espaço para descrever outras doenças e medicamentos e/ou suplementos em uso regular:
