



UNIVERSIDADE FEDERAL DE SANTA CATARINA
CENTRO DE CIÊNCIAS DA SAÚDE
PROGRAMA DE PÓS-GRADUAÇÃO EM FARMÁCIA

VINICIUS MATHEUS KNIHS

**GLP-1R (RECEPTOR DA INCRETINA *GLUCAGON-LIKE-PEPTIDE-1*)
E OS FENÓTIPOS DE SOBREPESO E OBESIDADE: UMA
ABORDAGEM *IN SILICO***

Florianópolis - SC

2023

Vinicius Matheus Knih

**GLP-1R (RECEPTOR DA INCRETINA *GLUCAGON-LIKE-PEPTIDE-1*) E OS
FENÓTIPOS DE SOBREPESO E OBESIDADE: UMA ABORDAGEM *IN SILICO***

Dissertação submetida ao Programa de Pós-Graduação em Farmácia da Universidade Federal de Santa Catarina para a obtenção do título de Mestre em Farmácia.

Orientador: Prof^ª. Dr^ª. Fabíola Branco Filippin Monteiro.

Florianópolis - SC

2023

Ficha de identificação da obra

Knihs, Vinicius Matheus

GLP-1R (RECEPTOR DA INCRETINA GLUCAGON-LIKE-PEPTIDE-1) E OS FENÓTIPOS DE SOBREPESO E OBESIDADE : UMA ABORDAGEM IN SILICO / Vinicius Matheus Knihs ; orientadora, Fabíola Branco Filippin Monteiro, 2023.

64 p.

Dissertação (mestrado) - Universidade Federal de Santa Catarina, Centro de Ciências da Saúde, Programa de Pós-Graduação em Farmácia, Florianópolis, 2023.

Inclui referências.

1. Farmácia. 2. Obesidade e diabetes. 3. GLP1R. 4. Metodologia In-Silico. 5. SNP. I. Monteiro, Fabíola Branco Filippin. II. Universidade Federal de Santa Catarina. Programa de Pós-Graduação em Farmácia. III. Título.

Vinicius Matheus Knihs

GLP-1R (Receptor da incretina *Glucagon-Like-Peptide-1*) e os fenótipos de sobrepeso e obesidade: Uma abordagem *in silico*

O presente trabalho em nível de mestrado foi avaliado e aprovado por banca examinadora composta pelos seguintes membros:

Prof^ª. Dr^ª. Tânia Beatriz Creczynski Pasa
Universidade Federal de Santa Catarina

Dr^ª. Valdelúcia Maria Alves de Souza
Universidade Federal de Santa Catarina

Certificamos que esta é a **versão original e final** do trabalho de conclusão que foi julgado adequado para obtenção do título de mestre em Farmácia.

Profa. Dra. Thaís Marques Sincero
Coordenadora do Programa

Profa. Dra. Fabíola Branco Filippin Monteiro
Orientadora

Florianópolis, 31 de maio de 2023.

Este trabalho é dedicado 100% aos meus amados pais, Gerson e Denise, que me apoiaram ao longo de todo o caminho desde o início, a toda a minha família que me acompanhou ao longo desses anos, especialmente aos meus avós Valdir e Marlene, Sílvio e Enedina, mas também às minhas tias Giana, Claudete e Rosélis, e aos meus queridos amigos Vinicius, Bruna e Tainá.

AGRADECIMENTOS

O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Código de Financiamento 001.

Gostaria de estender também meus agradecimentos a coordenadoria do Programa de Pós-Graduação em Farmácia por terem me proporcionado a oportunidade; à minha professora orientadora Fabíola Filippin, pois sem seu vasto conhecimento eu jamais teria chegado onde eu cheguei, devo muito do que o trabalho hoje é a você; agradeço também aos demais colegas do Laboratório de Imunologia do Tecido Adiposo (LAITA) por todas as contribuições ao longo da minha pesquisa; agradeço aos meus queridos amigos da vida toda, Vinicius Benatti e Bruna Leoni por terem me apoiado de maneiras que vocês nem sabem, terem me ajudado a manter um pouco da minha saúde mental e não ter pirado completamente nesses 2 anos. Apesar de toda a zoeira ainda amo vocês e não trocaria os nossos momentos por nada!!; agradeço a minha queridíssima amiga “da graduação para a vida” Tainá Lubschinski, por nunca ter desistido de mim mesmo quando eu fui chato desde os tempos de univali e também por ter me dado o toque de “tentar esse negócio de mestrado”, dois anos depois eu provavelmente devo o futuro da minha carreira à você, por isso agradeço demais, amo você!!; agradeço inusitadamente aos meus colegas jogadores de RPG, Shinkel, Paulinho, Júlio, Thalles, Juni, Erick e Jaques, sim galera, vocês me ajudaram a manter a saúde mental nesses 2 anos obrigado demais; agradeço, e agora de modo especial, a minha família, meus avós, Valdir Luís Knihs, o “Vadico” e Marlene Zucco Knihs e a minha tia Giana Carla Knihs, por toda a preocupação com o meu bem-estar, pelas caronas até Floripa quando eu precisei e por me ajudarem com tudo que eu precisei durante a minha vida toda; aos meus avós Sílvio Francisco Gonçalves, que infelizmente não me acompanhou em vida até aqui, mas tenho certeza de que continua olhando por mim onde quer que esteja, e a Maria Enedina Gonçalves, que ao longo da minha vida se fez muito presente pela minha mãe e pelas palavras de sabedoria que ela sempre teve, e às minhas tias, primas e primo, Claudete Gonçalves, Rosélis Gonçalves, Carolina Gonzaga, Jéssica Gonzaga Brogni e Lucas Francisco Nunes, vocês me apoiaram de maneiras que nem imaginam, por isso obrigado!!; agradeço aos queridíssimos amigos Rubens e Silvana por terem me cedido moradia quando eu precisei ficar um tempo em Florianópolis, sem vocês teria sido muito difícil esse período, por isso agradeço imensamente; e de modo mais do que especial, agradeço imensamente aos meus pais Gerson e Denise, se não fosse por vocês eu não só não estaria aonde eu estou, como também não seria quem eu sou, agradeço por todo o apoio ao longo de todos os meus estudos e da minha

vida, agradeço pela educação que me deram, pelas palavras de conforto e também pelos puxões de orelha quando necessário, por serem quem são, por existirem e por me apoiarem em todas as minhas decisões. Amo vocês além do que palavras podem expressar; agradeço aos meus pets queridos também, ao Max e a Brisa que já se foram, ao Tom meu gatinho companheiro e ao recém-chegado Simon, não queria no começo, mas acabei me apegando também; por fim gostaria de agradecer de todo o meu coração a Deus por ter me guiado durante toda a minha vida, por ter iluminado o meu caminho e a Nossa Senhora por ter intercedido por mim durante todo esse tempo, devo minha vida a Ti.

RESUMO

O GLP-1R é um membro proteico da família B dos receptores acoplados à proteína G, a maior proteína de ligação expressa na membrana celular e amplamente expressa nos tecidos humanos. Na estimulação do seu agonista (GLP-1), as funções do GLP-1R estão ligadas ao metabolismo da glicose, aumentando a secreção de insulina nas células beta pancreáticas e a regulação do apetite no hipotálamo e consequentemente associadas a distúrbios como a obesidade e o diabetes. Portanto, mutações nesse gene podem levar a alterações fisiológicas que podem explicar diferenças fenotípicas entre os indivíduos. Nesse contexto, este trabalho teve como objetivo avaliar sistematicamente as variantes missense do GLP-1R. Inicialmente, a mineração de dados foi realizada no dbSNP (*Single Nucleotide Polymorphism Database*) que apresentou 16.399 variantes para GLP-1R. Entre eles, 356 eram SNPs não sinônimos localizados em regiões codificantes (missense). Esses 356 nsSNPs foram analisados na interface *Web PolyPhen-2*, uma ferramenta *in silico* que prevê o impacto de uma substituição de aminoácidos na estrutura e função da proteína. As variantes também foram filtradas pela frequência alélica global (*cut-off* de 0,5) resultando em 6 potenciais variantes patogênicas: D344E, A239T, R310Q, R227H, R421P e R176G. Essas 6 variantes foram analisadas em quatro outras ferramentas de predição (*Align GVD*, *SNAP2*, *PANTHER* e *PROVEAN*), e todas confirmaram as variantes como sendo deletérias. Por fim, as variantes foram analisadas nas bases de dados *HOPE Prediction* e *SwissModdel*, plataformas que preveem a estrutura tridimensional de proteínas modificadas. Os nsSNPs D344E e A239T resultaram em resíduos maiores do que o tipo selvagem. D344E resultou em uma estrutura ligeiramente desestabilizada afetando alguns graus da funcionalidade do receptor e A239T pareceu afetar as hélices transmembrana com a membrana lipídica. No entanto, as variantes R310Q, R227H, R421P e R176G promoveram resíduos menores do que o tipo selvagem promovendo uma perda de carga positiva e um aumento da hidrofobicidade levando a distúrbios nas interações com outras hélices transmembranares do receptor. Particularmente, na variante R421P, a prolina desestabilizou a α -hélice e causou graves danos à estrutura 3D do receptor. Em conjunto, o esclarecimento do cenário mutacional das variantes do GLP-1R e o estudo do potencial dano desses nsSNPs na funcionalidade do receptor podem contribuir para a elucidação de doenças metabólicas relacionadas ao GLP-1R.

Palavras-chave: Obesidade. Diabetes. GLP1R. SNP. Genes. Receptores acoplados à proteína G.

ABSTRACT

GLP-1R is a protein member of the B family of G protein-coupled receptors, the major binding protein expressed in the cell membrane and widely diffused by human tissues. On stimulation of its agonist (GLP-1), the functions of GLP-1R are strongly linked to glucose metabolism, increasing insulin secretion in pancreatic beta cells and appetite regulation in the hypothalamus. Therefore, mutations in this gene can lead to physiological changes that can explain phenotypic differences between individuals. In this context, this work aimed to systematically evaluate the missense variants of GLP-1R. Initially, data mining was performed on dbSNP (Single Nucleotide Polymorphism Database) which retrieved 16,399 variants for GLP-1R. Among them, 356 were missense SNPs. These 356 nsSNPs were analyzed in the PolyPhen-2 web interface. Variants were also filtered by global allele frequency (0.5 cut-offs) resulting in 6 potential pathogenic variants: D344E, A239T, R310Q, R227H, R421P and R176G. They were analyzed in four other prediction tools (Align GVD, SNAP2, PANTHER and PROVEAN), and all confirmed the variants as deleterious. Finally, these 6 variants were analyzed in the HOPE Prediction and SwissModel databases, platforms that predict the three-dimensional structure of proteins. The D344E and A239T nsSNPs resulted in larger residues than wild-type ones. D344E resulted in a slightly destabilized structure affecting some degree of receptor functionality, and A239T affected the transmembrane helices with the lipid membrane. The R310Q, R227H, R421P and R176G variants resulted in smaller residues than the wild-type, promoting a loss of positive charge, and an increase in hydrophobicity leading to disturbances in interactions with other transmembrane helices of the receptor. Particularly in the R421P variant, proline destabilized the α -helix and caused severe damage to the 3D structure of the receptor. Together, clarifying the mutational scenario of GLP-1R variants and studying the potential damage of these nsSNPs on receptor functionality may contribute to the elucidation of metabolic diseases related to GLP-1R.

Keywords: Obesity. Diabetes. GLP1R. SNP. Genes. G protein-coupled receptors.

LISTA DE FIGURAS

Figura 1 – Dados da expressão do receptor da <i>glucagon-like peptide-1</i> (ENSG00000112164.5) fornecidos pela plataforma de transcriptoma GTex Portal (https://gtexportal.org/).....	21
Figura 2 - Estruturas químicas das moléculas de GLP1, Liraglutida e Semaglutida.....	22
Figura 3 - Representação esquemática das localizações das vias reguladoras do GPCR	23
Figura 4 - Fluxograma das fases da pesquisa e as bases de dados que foram utilizadas em cada uma	26
Figura 5 - Fluxograma das etapas seguidas para a obtenção das 6 variantes finais	31
Figura 6 - Seis variantes encontradas ao final do processo, destacadas no modelo de proteína do sistema SNAKE (https://gpcrdb.org/protein/glp1r_human/).....	32
Figura 7 - Visão geral da proteína na representação em fita da plataforma Swiss-Moddel® e da plataforma AlphaFold®.....	33
Figura 8 - Detalhes do resíduo do tipo selvagem em comparação com o resíduo mutante D344E.....	34
Figura 9 - Detalhes do resíduo do tipo selvagem em comparação com o resíduo mutante A239T.....	35
Figura 10 - Detalhes do resíduo do tipo selvagem em comparação com o resíduo mutante R310Q.....	36
Figura 11 - Detalhes do resíduo do tipo selvagem em comparação com o resíduo mutante R421P	36
Figura 12 - Detalhes do resíduo do tipo selvagem em comparação com o resíduo mutante R227H.....	37
Figura 13 - Detalhes do resíduo do tipo selvagem em comparação com o resíduo mutante R176G.....	38
Figura 14 - Rede de interações da proteína GLP-1R com outras proteínas ao seu redor em esquema de tipo de interação por evidência científica	39
Figura 15 - Rede de interações funcionais da proteína GLP-1R em esquema de confiabilidade	40
Figura 16 - Rede de interações físicas da proteína GLP-1R em esquema de confiabilidade	41

Figura 17 - Rede de interações físicas da proteína GLP-1R em esquema de confiabilidade com destaque para processos biológicos específicos das proteínas	42
Figura 18 - Rede de interações físicas da proteína GLP-1R em esquema de confiabilidade com destaque para função molecular das proteínas	43
Figura 19 - Rede de interações físicas da proteína GLP-1R em esquema de confiabilidade com destaque para associação entre gene das proteínas e doença	43
Figura 20 - Rede de interações físicas da proteína GLP-1R em esquema de confiabilidade com destaque para tecidos em que as proteínas são atuantes	44
Figura 21 - Rede de interações físicas da proteína GLP-1R em esquema de confiabilidade com destaque para tecidos em que as proteínas são atuantes	44

LISTA DE TABELAS

Tabela 1 – Informações básicas das mutações encontradas	32
Tabela 2 - Informações da proteína retiradas da plataforma HOPE® prediction.	38

LISTA DE ABREVIATURAS E SIGLAS

ADA – *American Diabetes Association*

BDNF - *Brain Derived Neurotrophic Factor*

dbSNP - *Single Nucleotide Polymorphism Database*

DM – *Diabetes Mellitus* GLP-1 – *Glucagon-like peptide 1*

DM1 – *Diabetes Mellitus Tipo 1* GLP-1R - *Glucagon-like peptide 1 receptor*

DM2 – *Diabetes Mellitus Tipo 2*

ERK – *Extracellular signal-regulated kinases*

FDA – *Food and Drug Administration*

FTO - *Fat mass and Obesity-associated protein* ou *Alpha-Ketoglutarate-Dependent Dioxygenase FTO*

GLP1R – *Gene codificador do Glucagon-like peptide 1 receptor*

GPCRs – *G Protein Coupled Receptors*

HbA1c – *Hemoglobina Glicada*

IMC – *Índice de Massa corpórea*

MAF – *Minor Allele Frequency*

MC4R – *Melanocortin-4 Receptor*

MYTH - *Membrane Yeast Two-Hybrid*

NCBI - *National Center for Biotechnology Information*

NLM – *National Library of Medicine*

nsSNP – *Non-synonymous Single Nucleotide Polymorphism*

OMS – *Organização Mundial da Saúde*

PPI – *(Protein-Protein Interaction)*

SNP – *Single Nucleotide Polymorphism*

UCP3 – *Uncoupling Protein 3*

SUMÁRIO

1	INTRODUÇÃO	15
1.1	Objetivos.....	25
1.1.1	Objetivo Geral	25
1.1.2	Objetivos Específicos.....	25
2	METODOLOGIA	26
2.1	Obtenção da listagem de variantes e triagem.....	26
2.2	Obtenção de dados da plataforma ClinVar	28
2.3	Algoritmos de predição para dano estrutural das variantes do tipo <i>missense</i>....	29
3	RESULTADOS	31
3.1	O gene GLP1R.....	31
3.2	Recuperação das variantes	31
3.2.1	Análise de variantes.....	34
3.2.1.1	Variante D344E.....	34
3.2.1.2	Variante A239T.....	34
3.2.1.3	Variante R310Q.....	35
3.2.1.4	Variante R421P	36
3.2.1.5	Variante R227H.....	37
3.2.1.6	Variante R176G.....	37
3.3	Plataforma STRING.....	39
3.4	Plataforma ClinVar.....	46
4	DISCUSSÃO	48
5	CONCLUSÃO	52
	REFERÊNCIAS	53
	APÊNDICE – Tabelas contendo conteúdo relacionado à obtenção das variantes patogênicas.....	64

1 INTRODUÇÃO

As doenças crônicas e a obesidade surgiram como principais preocupações de saúde no século XX, por meio de mudanças ambientais compartilhadas. Em 1900, observava-se uma maior incidência de mortes por condições agudas, como doenças infecciosas, que eram a maior causa de morte na época. Agora essas doenças são amplamente controladas e a medida que a virada do século foi se aproximando, a incidência de mortes por doenças crônicas, como obesidade, foram aumentando (MILLS, 2000).

Fatores que favorecem um balanço energético positivo e ganho de peso têm sido apontados como a causa das principais doenças metabólicas nas últimas décadas, sendo que esses fatores incluem (I) o aumento da oferta e consumo de alimentos *per capita*, principalmente de alimentos altamente calóricos e palatáveis que são frequentemente servidos em grandes porções, prática que têm aumentado conforme a passagem dos anos, levando à um concomitante aumento de peso corporal (HALL *et al*, 2009); (II) diminuição do tempo gasto em atividades físicas ocupacionais, tendência que tem sido observada a partir dos anos 60, com a mudança gradativa de ocupações que demandam atividade física moderada em favor de ocupações que favorecem comportamento sedentário; e diminuição de atividades físicas, como uso de aparelhos eletrônicos para fins de lazer, especialmente em países mais industrializados, nos quais é comum pessoas passarem metade dos seus dias sentadas, por conta dos avanços tecnológicos em ambientes domésticos e de trabalho (CHURCH *et al*, 2011; LOEFFELHOLZ *et al*, 2022); (III) uso crescente de medicamentos que têm como efeito colateral o ganho de peso, como medicamentos psicotrópicos, fármacos para reposição hormonal, contraceptivos esteroidais, anti-histamínicos e inibidores de protease e sono inadequado, por levar a, entre outras coisas, diminuição nos níveis de leptina circulante, aumento na secreção de grelina, mas também por causar cansaço, o que leva a uma diminuição da atividade corporal em geral, contribuindo para o aumento de peso (MCALLISTER *et al*, 2009). Esses e muitos outros fatores estabeleceram as bases para as epidemias conjuntas de doenças crônicas não transmissíveis como diabetes e obesidade (FIELD *et al*, 2001; HEYMSFIELD; WADDEN, 2017; BURKI, 2021).

No entanto, nem todas as pessoas expostas aos comportamentos de risco em ambientes urbanos e rurais desenvolvem obesidade ou diabetes, o que sugere a existência de mecanismos genéticos subjacentes que operam no nível individual, como por exemplo mutações no receptor da melanocortina-4, visto que a ativação desse receptor estimula a

produção de cAMP, o que leva a uma diminuição na ingestão de alimentos (BROWERS *et al*, 2021). Outros exemplos incluem o gene BDNF, que codifica o Fator Neurotrófico Derivado do Cérebro (BDNF - *Brain Derived Neurotrophic Factor*) da proteína. Mutações no gene BDNF e/ou em seu receptor estão associadas à obesidade em modelos humanos e animais. O gene GHRL, que codifica os pré-propeptídeos grelina e obestatina, e o gene UCP3, que codifica duas formas de transcrições, um mensageiro (UCP31) e uma isoforma curta (UCP3s). A proteína UCP3 (*Uncoupling Protein 3*) é um transportador transmembranar mitocondrial específico para músculos, e algumas variantes foram associadas à obesidade infantil (LOOS; YEO, 2021; BRATTI *et al*, 2023).

Onze formas monogênicas raras de obesidade são atualmente reconhecidas, incluindo uma deficiência dos receptores de leptina e melanocortina-4, que são expressos principalmente no hipotálamo e estão envolvidos em circuitos neurais que regulam a homeostase energética (KOHLSDORF, 2018). Ainda, uma estratégia amplamente utilizada para descobrir mecanismos poligênicos que conferem suscetibilidade à obesidade comum envolve a triagem de todo o genoma em altas quantidades de amostras com o objetivo de identificar polimorfismos de nucleotídeo único (SNP) associados ao índice de massa corporal (IMC) e outras características relacionadas à obesidade, como a circunferência da cintura. Estudos de associação em todo o genoma relacionados ao IMC e ao índice de circunferência da cintura, revelaram mais de 100 loci relacionados aos mesmos, mas apenas uma pequena porção de hereditariedade na obesidade foi explicada por essas variantes genéticas (WU *et al*, 2018).

Mais de 300 loci foram identificados em estudos de associação genômica ampla, embora coletivamente esses locais representem menos de 5% da variação individual no IMC e nas características de adiposidade do indivíduo (KHERA *et al*, 2019; BROUWERS *et al*, 2021). Os sinais mais proeminentes desta abordagem metodológica são observados ao analisar os achados das variantes do gene FTO, codificador da proteína associada a massa gorda e obesidade ou dioxigenase dependente de alfa-cetoglutarato (*fat mass and obesity-associated protein* ou *alpha-ketoglutarate-dependent dioxygenase FTO*), visto que o mesmo é associado ao alto IMC. Pessoas portadoras de uma ou duas cópias do alelo de risco têm um aumento de peso de 1,2 kg ou 3 kg, respectivamente, em comparação com pessoas sem cópias do alelo (MENG *et al*, 2014; DA SILVA *et al*, 2013).

A obesidade é marcada por ganho de peso não saudável, excesso de adiposidade geral e central, inflamação sistêmica de baixo grau, hipóxia tecidual e, em muitos casos, dislipidemia, podendo também ser marcada por um desbalanço energético de longo termo, entre calorias gastas e consumidas. Portanto, por definição, obesidade é uma doença inflamatória e encontra-se um desafio no seu diagnóstico, que é majoritariamente feito por

meio do IMC (BAKBAK *et al*, 2021; LIN; LI, 2021). Em adultos, a Organização Mundial de Saúde (OMS) define sobrepeso quando o indivíduo apresenta IMC entre 25,0 e 29,9 kg/m² e obesidade com IMC \geq 30,0 kg/m². A obesidade é ainda classificada em três níveis de gravidade: classe I (IMC 30,0-34,9 kg/m²), classe II (IMC 35,0-39,9 kg/m²) e classe III (IMC \geq 40,0 kg/m²). O IMC é calculado usando o peso em kg dividido pelo quadrado da altura em metros e é mais bem associado com o percentual e massa de gordura corporal na maioria dos grupos. Embora não seja a métrica mais precisa para o cálculo da proporção de gordura corporal, existem grandes diferenças individuais na porcentagem de gordura corporal para um determinado valor de IMC, dependente de sexo, etnia e idade, conforme observado no estudo de Moretto *et al* (2016), que avaliou o efeito da cor ou raça no IMC conforme o gênero, e observou uma prevalência maior em homens e mulheres de raça/cor branca, por exemplo (MORETTO *et al*, 2016). Entretanto é um parâmetro de fácil avaliação, portanto, representa o primeiro passo e o mais comumente aceito para determinar o grau de sobrepeso ou obesidade (LIN; LI, 2021; BOUTARI; MANTZOROS, 2022).

A obesidade aumenta a probabilidade de várias doenças e condições que estão ligadas ao aumento da mortalidade. Estes incluem diabetes mellitus tipo 2 (DM2), doenças cardiovasculares, síndrome metabólica, doença renal crônica, hiperlipidemia, hipertensão, doença hepática gordurosa não alcoólica, certos tipos de câncer, apneia obstrutiva do sono, osteoartrite e depressão (PANTALONE *et al*, 2017). O tratamento dessas condições pode sobrecarregar financeiramente os sistemas de saúde e, como os custos totais de saúde relacionados à doença dobram a cada década, o tratamento das consequências da obesidade representa um desafio caro aos órgãos públicos (LIN; LI, 2021).

A obesidade é um importante problema de saúde pública ao redor do globo e a dependência apenas das medidas do IMC provou ser inadequada para ajudar os médicos a avaliar e gerenciar os riscos de saúde relacionados à obesidade em seus pacientes. Isso está parcialmente relacionado ao fato de que o IMC isoladamente é um biomarcador insuficiente de adiposidade abdominal ou corporal (BUCHHOLZ; BUGARESTI, 2005). A circunferência da cintura é um método simples de avaliação da adiposidade abdominal, de fácil padronização e aplicação clínica. A circunferência da cintura está fortemente associada às causas e a mortalidade por doenças cardiovasculares, no entanto, a maior correlação da associação entre circunferência da cintura com morbidade e mortalidade só é percebida após ajuste com o IMC. Assim, a circunferência da cintura possibilita um maior refinamento do risco adverso à saúde caracterizado pelo IMC e essa medida deve ser incluída na estratificação do risco à saúde relacionado à obesidade (ROSS *et al*, 2020).

A obesidade mundial quase triplicou desde 1975 e, em 2016, mais de 1,9 bilhão de adultos, com 18 anos ou mais, estavam acima do peso. Destes, mais de 650 milhões eram obesos, o que corresponde a 39% dos adultos com 18 anos ou mais com sobrepeso e 13% com obesidade. Trinta e nove milhões de crianças menores de 5 anos estavam com sobrepeso ou obesidade em 2020. Mais de 340 milhões de crianças e adolescentes de 5 a 19 anos estavam com sobrepeso ou obesidade em 2016, segundo dados mais atualizados da OMS (WORLD HEALTH ORGANIZATION, 2016).

Uma outra doença crônica que frequentemente está associada a obesidade é o Diabetes Mellitus. O DM2 mais especificamente, é caracterizado pela diminuição da sensibilidade à insulina resultando em hiperglicemia e aumento do estresse oxidativo (RODEN; SHULMAN, 2019). O diabetes não é uma doença única, mas sim um grupo de condições amplamente categorizadas por um único critério diagnóstico: hiperglicemia. Cada vez mais, as evidências sugerem que mesmo o DM2, que representa 90-95% dos casos, é heterogêneo em termos de mecanismos de ação e relações com sinais clínicos. As abordagens usando biomarcadores clínicos ou genéticos identificaram subtipos de DM2 que são clinicamente distintos e diferencialmente associados a complicações diabéticas (COLE; FLOREZ, 2020; BAKBAK *et al*, 2021).

O diabetes pode ser classificado em 4 categorias, sendo elas diabetes tipo 1 (DM1), que ocorre geralmente por mecanismos autoimunes, o que leva a deficiência absoluta de insulina por conta da destruição das células β -pancreáticas; diabetes tipo 2 (DM2), que ocorre geralmente em casos de resistência insulínica por parte do indivíduo (por conta dos altos níveis de insulina circulante livre), mecanismo que causa uma perda progressiva da secreção adequada da mesma; diabetes gestacional, caracterizado somente quando diagnosticado no segundo ou terceiro trimestre da gravidez em pacientes que não possuíam diabetes previamente; e ainda tipos específicos de diabetes por outras causas, como síndromes de diabetes monogênicas, doenças do pâncreas exócrino que podem afetar a produção de insulina (como fibrose cística e pancreatite) e diabetes induzidas por fármacos como glicocorticoides. (AMERICAN DIABETES ASSOCIATION - ADA, 2022).

A classificação do tipo de diabetes é importante para determinar a terapia, mas alguns indivíduos não podem ser claramente classificados como tendo diabetes tipo 1 ou tipo 2 no momento do diagnóstico, devido a similaridade das condições, pois ambos os tipos apresentam ineficiência de células β -pancreáticas. Alguns casos em que ambos os tipos são reportadas ao mesmo tempo já foram documentados, e para tanto foi criada uma nova categoria, a Diabetes Dupla ou híbrida (DD) que, entre outros sintomas, inclui além de resistência a

insulina, autoanticorpos contra células β e obesidade, conforme reportado por Khawandanah (2019). A presença de obesidade nesses casos, no entanto, geralmente é prévia ao caso, pois as principais complicações da doença incluem diabetes, bem como doenças cardiovasculares e também distúrbios músculo-esqueléticos (LIN *et al*, 2021; KHAWANDANAH, 2019).

As características mais úteis na discriminação do DM1 incluem idade mais jovem no diagnóstico (<35 anos) com IMC mais baixo (<25 kg/m²), perda de peso não intencional, cetoacidose e glicose > 360 mg/dL (20 mmol/L) na apresentação (KATSAROU *et al*, 2017). A classificação do tipo de diabetes nem sempre é simples na apresentação e o diagnóstico incorreto é comum. Embora as dificuldades em distinguir o tipo de diabetes possam ocorrer em todas as faixas etárias no início, o diagnóstico torna-se mais evidente ao longo do tempo em pacientes com deficiência de células β (ILLONEN; LEMPAINEN; VEIJOLA, 2019).

Fatores genéticos e ambientais são importantes de se levar em consideração quando se trata da avaliação de casos de DM1 e DM2, pois acarretam perda progressiva de massa e/ou função das células β -pancreáticas, o que é observado clinicamente como hiperglicemia. Embora as taxas de progressão sejam diferentes de um tipo para o outro, todos os pacientes com diabetes, de ambos os tipos, correm o risco igualmente de desenvolver as mesmas complicações crônicas. O diagnóstico de diabetes é baseado principalmente na glicemia de jejum, ou no teste oral de tolerância à glicose para construção da curva glicêmica. Pode-se ainda utilizar o critério de hemoglobina glicada (HbA1C), embora seja um exame mais comumente utilizado para controle do que para diagnóstico (HARREITER; RODEN, 2019). A hiperglicemia desenvolve-se continuamente e os distúrbios da glicemia de jejum e pós-prandial têm diferentes cursos de tempo. Conforme a prevalência mundial tem aumentado, cada vez mais se faz necessário triagens direcionadas para a detecção de diabetes e pré-diabetes o quanto antes em grupos de risco, para dessa forma retardar a progressão, ou até mesmo prevenir a manifestação da doença (HARREITER; RODEN, 2019; ADA, 2022).

O número de pessoas com diabetes aumentou de 108 milhões em 1980 para 422 milhões em 2014 e um aumento mais rápido dessa prevalência tem sido observado em países de baixa e média renda, em relação a países de alta renda. A diabetes é uma das principais causas de cegueira, insuficiência renal, ataques cardíacos, acidente vascular cerebral e amputação de membros inferiores. Entre 2000 e 2016, houve um aumento de 5% na mortalidade

prematura por diabetes e, em 2019 foi a nona principal causa de morte, com uma estimativa de 1,5 milhão de mortes causadas diretamente, segundo dados mais atualizados da OMS (WORLD HEALTH ORGANIZATION, 2016).

Indivíduos que vivem com obesidade têm maior massa adiposa, aumento do débito cardíaco e pressão arterial elevada devido a alterações hemodinâmicas (KHAODHIAR; MCCOWEN; BLACKBURN, 1999). À medida que o estado de obesidade progride, com hipertrofia do tecido adiposo, a hipóxia promove a infiltração de células imunes no tecido adiposo (DE OLIVEIRA et al, 2013) desencadeando a secreção de citocinas que contribuem para a inflamação sistêmica (BRATTI *et al*, 2021; FOTI; BRUNETTI, 2017). Essas características também contribuem para o desenvolvimento de DM2 (BAKBAK *et al*, 2021).

Além dos já citados, um outro fator que ambas as doenças têm em comum é o envolvimento da proteína *glucagon-like peptide-1* (GLP-1), uma incretina multifacetada com cerca de 30 ou 31 amino-ácidos, vários sítios de ação e um potencial farmacêutico abrangente (incluindo no tratamento de DM2, como é o caso da exedina-4 e da liraglutida). O GLP-1 é principalmente secretado pelas células L enteroendócrinas no intestino distal, pelas células α no pâncreas e pelo sistema nervoso central, portanto participa, entre outras coisas, na regulação da glicose sanguínea (ZHAO *et al*, 2021). O GLP-1 atua em diversos tecidos com diversas funções por se tratar de um hormônio cujo receptor é acoplado à proteína G, a maior família de receptores do corpo humano, possuindo diversas funções fisiológicas e entre os inúmeros efeitos metabólicos da incretina GLP-1, destacam-se a estimulação da secreção de insulina dependente de glicose, diminuição do esvaziamento gástrico, inibição da ingestão de alimentos, aumento da natriurese e diurese e modulação da proliferação de células- β (MÜLLER et al, 2019).

Dados da plataforma GTEx®, uma plataforma de dados de transcriptoma (<https://gtexportal.org/home/gene/GLP1R> - Estados Unidos, acesso em 16/05/2023, ENSG00000112164.5), apontam os tecidos onde o gene GLP1R é mais expresso, em transcritos por milhão (TPM), que incluem o pâncreas (4,68); o estômago (0,41); tecido adiposo visceral (0,058); o hipotálamo no cérebro (0,39), que é responsável pela regulação do apetite, entre outras funções, juntamente com coração (1,74), testículos (0,77) e nervo tibial (0,91).

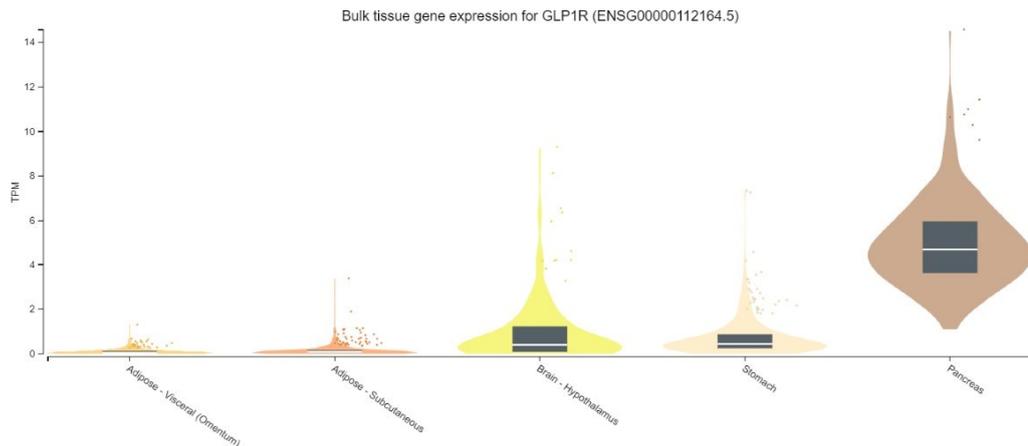


Figura 1. Dados da expressão do receptor da *glucagon-like peptide-1* (ENSG00000112164.5) fornecidos pela plataforma de transcriptoma GTex Portal (<https://gtexportal.org/>). Os tecidos de pâncreas (n= 328, TPM = 4,680), estômago (n=359, TPM= 0,4197), hipotálamo (n= 202, TPM= 0,3931) e tecidos adiposos subcutâneo (n= 663, TPM = 0,07048) e visceral (n= 541, TPM = 0,05823) foram coletados. N = número de amostras analisadas e compiladas no banco de dados. TPM= transcritos por milhão.

A incretina GLP-1 ainda possui outras funções como neuro e cardioproteção, diminuição de inflamação e apoptose, entre outras atividades reportadas pela literatura. Para os propósitos do presente trabalho, o foco será nos efeitos relacionados ao diabetes e a obesidade, pois os agonistas do receptor do GLP-1 (GLP-1R – *Glucagon-Like Peptide-1 Receptor*) são um importante alvo para tratamentos farmacoterapêuticos de obesidade e diabetes tipo 2 (MÜLLER et al, 2019).

Esses tratamentos farmacoterapêuticos para obesidade e diabetes, se dão por meio dos análogos do GLP-1 (principalmente liraglutida e semaglutida - figura 2). Os análogos do GLP-1 são amplamente utilizados atualmente para redução de peso e controle glicêmico em casos de DM2 e são os mais comumente associados a esse tipo de tratamento. A eficácia clínica de tratamentos para controle glicêmico fazendo uso de injeção subcutânea diária nas concentrações de 1,2–1,8 mg de liraglutida e tratamentos com dosagens mais altas, de 3,0 mg para tratamentos de redução de peso, já é comprovada. Casos que utilizam a dose mais alta diariamente, demonstram uma perda de peso de cerca de 6,0% entre pacientes com DM2 e obesidade (TILINCA et al, 2021; TRENSON et al, 2022).

A semaglutida, também recomendada para tratamento e manutenção de pacientes com obesidade ou com alguma comorbidade em conjunto com quadro de sobrepeso, com ou sem DM2, é ministrada em dose menor que a máxima de liraglutida, de 2,4 mg. Ambos os medicamentos são análogos do GLP-1 e possuem tempo de ação prolongado. A liraglutida tem um tempo de meia-vida que vai de 13 a 15 horas, devido a adição de uma cadeia lateral de ácido graxo C16. Já a semaglutida possui um tempo de meia-vida de cerca de 165 horas, por conta da reposição de aminoácidos e da adição de diácido graxo C18 na sua composição, conforme a figura 2 (RUBINO *et al*, 2021).

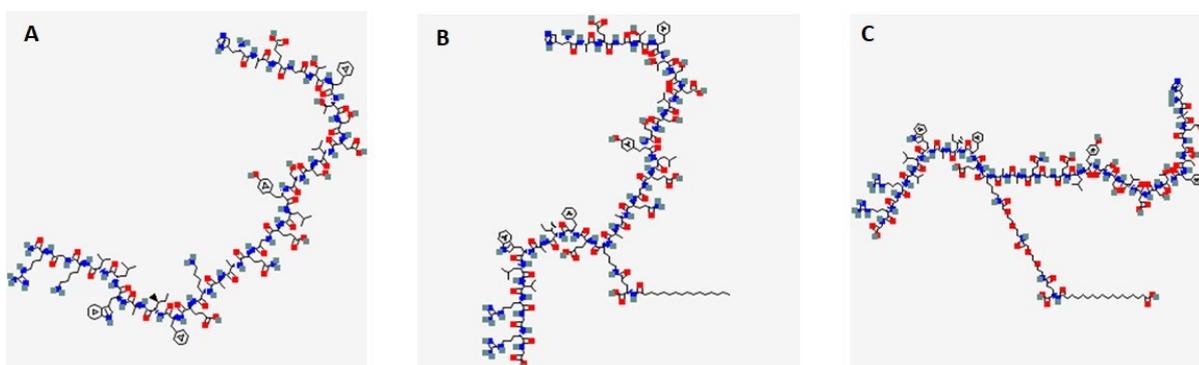


Figura 2. Estruturas químicas das moléculas de (A) GLP-1 (PubChem CID 16135499 <https://pubchem.ncbi.nlm.nih.gov/compound/16135499>), (B) Liraglutida (PubChem CID 16134956 - <https://pubchem.ncbi.nlm.nih.gov/compound/16134956>) e (C) Semaglutida (PubChem CID 56843331 - <https://pubchem.ncbi.nlm.nih.gov/compound/56843331>). <https://pubchem.ncbi.nlm.nih.gov/> Acesso em 22 de maio de 2023.

O GLP-1R faz parte da família de receptores acoplados à proteína G (*G protein-coupled receptors*, GPCRs), uma superfamília de receptores que participam ativamente de diversos processos fisiológicos e são encontrados na maioria dos tecidos do corpo, o que a caracteriza como a maior superfamília de proteínas codificadas pelo genoma humano, constituída de 826 genes. Quando ativados, os GPCRs atuam em conjunto com proteínas citosólicas, traduzindo o sinal do meio extracelular para o meio intracelular, dando assim início à cascata de sinalizações intracelulares, levando à resposta celular (HOELZ *et al*, 2013).

Estes receptores tem uma estrutura que é semelhante de forma geral, embora possuam uma diferença considerável de tamanho molecular. São metabotrópicos, com 7 domínios transmembrana, expondo a região amino-terminal (N-terminal) e três alças de interligação (EL1, EL2 e EL3) no meio extracelular, e a região carboxi-terminal (C-terminal) com outras três alças de interligação (CL1, CL2 e CL3), no meio intracelular. A grande família GPCR

constitui o alvo de fármacos mais explorados no genoma humano, justamente por ser tão abrangente. Por exemplo, os GPCRs regulam a sinalização de neurotransmissores que são alvo da ação de muitos antidepressivos, como os tratamentos com inibidores seletivos da recaptação da serotonina e inibidores da recaptação da norepinefrina, que resultam em uma estimulação indireta do GPCR da serotonina (5-HT) e da norepinefrina (WATKINS; ORLANDI, 2020; HOELZ et al, 2013).

A superfamília das GPCRs humanas é dividida em subfamílias de classes A (rodopsina), B (secretina e adesão), C (glutamato) e F (Frizzled – “crespo”) de acordo com suas seqüências de aminoácidos. Dos 826 GPCRs humanos, aproximadamente 350 membros são considerados alvos terapêuticos importantes e 165 deles são alvos de fármacos validados. Dados estatísticos mais recentes indicam que 527 medicamentos aprovados pela *Food and Drug Administration* (FDA) e cerca de 60 candidatos a medicamentos atualmente em ensaios clínicos visam os GPCRs. Para o caso do GLP-1R, principal alvo de estudo deste trabalho, ele é um membro da família B, mais especificamente da subfamília B1, de receptores de secretinas (YANG et al, 2021)(Figura 3).

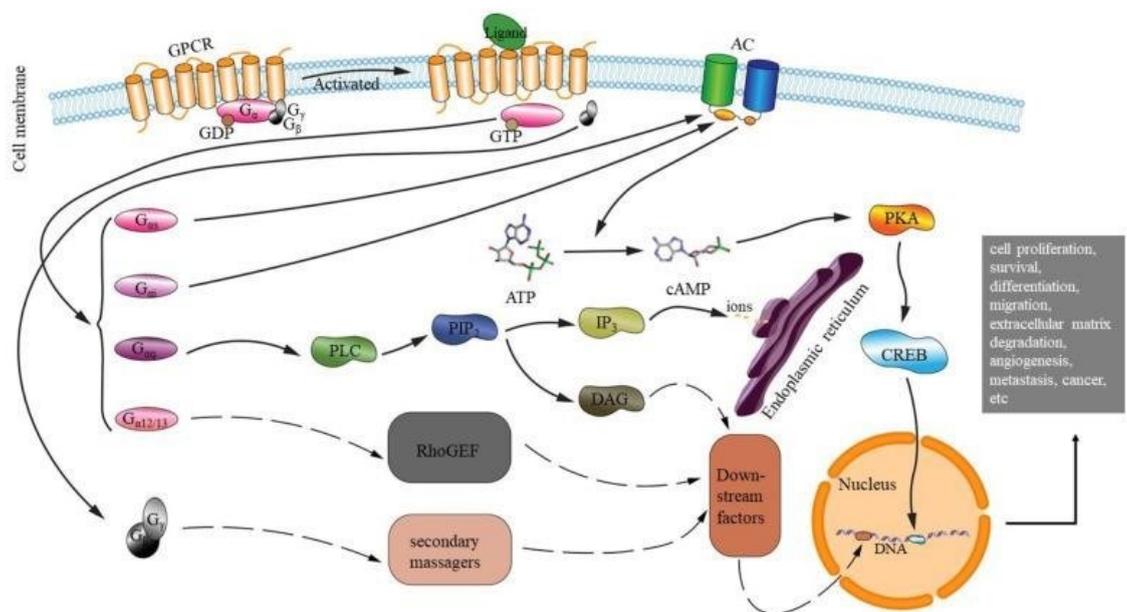


Figura 3. Representação esquemática das localizações das vias reguladoras do GPCR. Os GPCRs podem ser ativados por uma variedade de ligantes, interagindo com proteínas G heterotriméricas compostas por três subunidades ($G\alpha$, $G\beta$ e $G\gamma$) e ativando várias moléculas efetoras (LIU *et al*, 2021).

Em seu estado natural, o GLP-1R é importante no metabolismo de glicose e mecanismos de saciedade, e está diretamente associado tanto à DM2 quanto à obesidade. Levando isso em consideração, um importante foco do presente trabalho foram as mutações envolvidas no gene GLP1R. Portanto, devido à relação desse receptor e do gene que o codifica com as doenças já mencionadas, torna-se importante verificar os impactos que as mutações no mesmo podem causar.

Uma mutação pode ser definida como qualquer mudança no arranjo do DNA, afetando porções maiores de um cromossomo ou na sequência de nucleotídeos de um gene que codifica uma proteína específica, podendo ser classificadas em três grandes grupos: mutações genômicas, mutações cromossômicas e mutações gênicas (THOMPSON et al, 2002).

As mutações genômicas são aquelas que afetam o número de cromossomos totais nas células; as mutações cromossômicas são aquelas que afetam a estrutura de algum cromossomo individual e as mutações gênicas, que foram o principal alvo de estudo deste trabalho, são aquelas em que somente um gene individual é alterado, neste caso, o gene GLP1R. Entre as mutações gênicas, existem três principais tipos, chamadas mutações de ponto, são as inserções, deleções ou substituições de pares de bases. Para este trabalho, foram consideradas as mutações *missense*, que são do tipo substituição de pares de bases, na qual a troca do nucleotídeo gera uma troca de aminoácido no processo de síntese proteica, o que pode afetar a estrutura da proteína (THOMPSON et al, 2002).

Recentemente foram previstas interações polares (ligações de hidrogênio) entre o GLP-1R e os seus agonistas (ZHANG et al, 2020) e que alguns polimorfismos do gene GLP1R estavam relacionados a alterações na afinidade do agonista em diversos fenótipos, como no DM2 (BEINBORN et al, 2005; LIN et al, 2015). As sinalizações de saciedade e sensibilidade à insulina podem ser prejudicadas ou intensificadas diante de alterações funcionais no GLP-1R e que podem ser explicadas por variantes genéticas. As variantes causadoras de fenótipos patogênicos são geralmente raras, entretanto, alelos de alta frequência (> 0,5%) existem, mas seu número é bastante pequeno (ZASTROW et al, 2018). A interpretação precisa dessas variantes tem um impacto significativo no entendimento de diversas doenças e, é claro, nos fenótipos de obesidade e baixa responsividade aos tratamentos farmacológicos e nutricionais (XIANG et al, 2020).

1.1 Objetivos

1.1.1 Objetivo Geral

Compilar as variantes do gene GLP1R dentre as diversas bases de dados e classificá-las quanto ao grau de dano à proteína e que reflitam fenótipos de obesidade e diabetes.

1.1.2 Objetivos Específicos

Buscar as variantes do gene GLP1R em bases de dados genômicos e classificar como patogênicas ou benignas de acordo com bases de dados interpretativos;

Identificar as variantes *missense* que causam disfunção estrutural ou molecular na proteína a partir de algoritmos de predição;

Identificar as interações proteína-proteína do GLP-1R em bases de dados de interactoma a fim de entender como se estabelece o *network* da proteína;

Identificar variantes que proporcionam perda ou ganho de função do GLP-1R.

2 METODOLOGIA

2.1 Obtenção da listagem de variantes e triagem

A recuperação do conjunto de variantes do gene *GLP1R* foi inicialmente realizada na base de dados genômica dbSNP®, do *National Center for Biotechnology Information* (NCBI - <https://www.ncbi.nlm.nih.gov/snp/> - Estados Unidos), que contém variações de nucleotídeo único (*Single Nucleotide Polymorphisms* - SNPs) humanos, microssatélites, inserções e exclusões em pequena escala, juntamente com informações de publicação da variante específica, frequência populacional, consequência molecular provocada pela variante específica, informações de mapeamento genômico, variações comuns e mutações clínicas (HOSSAIN *et al*, 2020). Recuperou-se as informações do ID da variante, sequência FASTA (formato básico de armazenamento de sequências de nucleotídeos), localização no cromossomo e alteração de resíduos na proteína. A pesquisa foi separada em 3 fases, conforme a Figura 4.

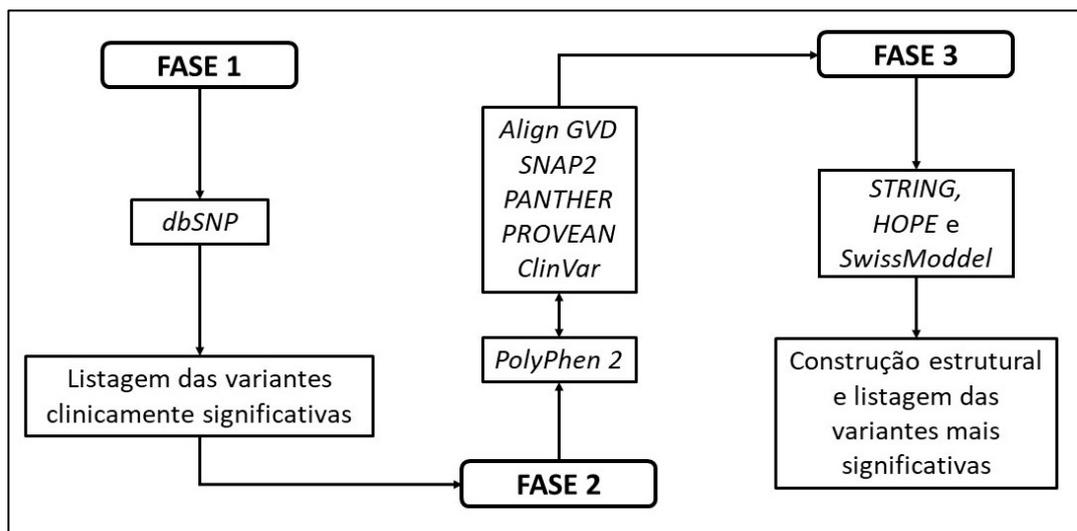


Figura 4. Fluxograma das fases da pesquisa e as bases de dados que foram utilizadas em cada uma. Fase 1, recuperação e listagem das variantes *missense*; fase 2, predição de dano estrutural; e fase 3, homologia estrutural e classificação do *network* proteico (Desenvolvida pelo autor).

O filtro para variantes *missense* foi aplicado e as variantes resultantes foram analisadas na ferramenta de predição de dano estrutural *PolyPhen-2*® (*Polymorphism Phenotyping v2*; <http://genetics.bwh.harvard.edu/pph2/> - Estados Unidos), que é uma ferramenta que prevê o possível impacto de uma substituição de aminoácidos na estrutura e função de uma proteína humana usando considerações físicas e comparativas diretas (ADZHUBEI, 2010). Descrito pela primeira vez em 2002, por Vasily Ramensky, Peer Bork, e Shamil Sunyaev, a plataforma

PolyPhen surgiu como uma forma de acelerar o estudo de um grande número de SNPs em populações em busca de variantes ao mesmo tempo. A principal função do servidor é ajudar a priorizar os SNPs de acordo com sua significância funcional, identificar a base estrutural de mutações de doenças e explicar a causa molecular das mesmas. A variante alélica causadora de uma doença pode, assim, ser identificada e associada a um locus específico (RAMENSKY *et al*, 2002). O *PolyPhen* é definido como um servidor *online* dedicado à anotação funcional automatizada de nsSNPs (*Non-Synonymous Single Nucleotide Polymorphisms*) de codificação, e essa varredura é feita através do *input*, que é a sequência de aminoácidos de uma proteína juntamente com a posição da sequência e duas variantes de aminoácidos que caracterizam o polimorfismo. A análise é baseada na caracterização da proteína, combinando informações sobre características de sequência, alinhamento múltiplo com proteínas homólogas, parâmetros estruturais e contatos para fazer uma previsão do efeito do nsSNPs na funcionalidade estrutural da proteína codificada (RAMENSKY *et al*, 2002; ADZHUBEI *et al*, 2013).

A plataforma *PolyPhen-2*®, um *upgrade* da plataforma *PolyPhen*, é uma ferramenta que possui um conjunto de recursos preditivos, *pipeline* de alinhamento e método de classificação, usando oito recursos preditivos baseados em sequência e três baseados em estrutura, selecionados por métodos de *machine-learning* (ADZHUBEI, 2010). O *PolyPhen-2*® realiza a anotação funcional de polimorfismos de nucleotídeo único (SNPs), mapeia SNPs codificantes para transcrições de genes, extrai anotações de sequências de proteínas e atributos estruturais e constrói perfis de conservação, para então predizer se a mutação *missense* tem o potencial de causar danos à proteína ou não. A previsão feita pela plataforma é baseada em uma série de características sequenciais, filogenéticas e estruturais que caracterizam a substituição e, com base nesses critérios, avalia o nível de dano na estabilidade e na função de proteínas específicas humanas. Para uma determinada substituição de aminoácidos, o *PolyPhen-2*® extrai várias características baseadas em sequência e estrutura do local de substituição e as alimenta a um classificador probabilístico que realiza essa análise através de métodos de *machine-learning* (ADZHUBEI *et al*, 2013).

A plataforma *PolyPhen-2*® possui um critério de pontuação no intervalo de 0.0 até 1.0, sendo que as mutações mais próximas de 0.0 foram consideradas benignas ou com dano insignificante e as mais próximas de 1.0 foram consideradas provavelmente danosas. Utilizou-se o critério de pontuação na faixa de 0.9 até 1.0, com o intuito de selecionar apenas as variantes

com danos potenciais. Ainda, a estas variantes, foi aplicado um último filtro na plataforma dbSNP, selecionando apenas variantes com frequência alélica inferior a 0,5%.

As frequências alélicas específicas de cada mutação foram retiradas da plataforma GnomAD® (*Genome Agregation Database*; <https://gnomad.broadinstitute.org/> - Estados Unidos) e da plataforma 1000G® (1000 *genomes project*; <https://www.internationalgenome.org/home> - Estados Unidos). Variantes causadoras de doença são raras, de uma forma geral e, embora possam ser observadas em frequências maiores, seus números são significativamente pequenos e, por esse motivo, uma frequência menor que 0,5% foi definida (XIANG *et al*, 2020).

2.2 Obtenção de dados da plataforma ClinVar

Outra plataforma que foi utilizada esta, porém, para dados complementares, foi a plataforma ClinVar (<https://www.ncbi.nlm.nih.gov/clinvar/>). O ClinVar é um arquivo público disponível gratuitamente de variantes genômicas humanas e interpretações de suas relações com doenças e outras condições. É parte do NCBI, assim como o dbSNP, dentro da *National Library of Medicine* (NLM). Interpretações de variantes no ClinVar podem ser enviadas por usuários da plataforma, o que acaba incluindo laboratórios de testes clínicos, laboratórios de pesquisa, bancos de dados específicos de locus gênicos, profissionais médicos, registros de pacientes, painéis de especialistas e outras organizações. Cada registro enviado inclui uma descrição da variante ou conjunto de variantes que foi interpretado, a condição para a qual a variante foi interpretada, a interpretação do significado clínico da variante e as evidências do remetente para essa interpretação, estruturadas como observações da variante (LANDRUM *et al*, 2020).

Uma submissão ao ClinVar é estruturada para representar um alelo ou genótipo, uma interpretação do significado clínico dessa variante em relação a uma condição e evidências clínicas de apoio. O remetente pode incluir casos em que a variante foi identificada, os fenótipos dos casos associados, bem como evidências adicionais e citações da literatura que apoiem a interpretação. O ClinVar mapeia variantes enviadas para sequências de referência apropriadas, adiciona identificadores e outros dados disponíveis de plataformas públicas, como frequências alélicas, além de agregar todos os envios de uma determinada variante. O ClinVar não avalia a precisão das interpretações enviadas, porém, atribui um *status* de revisão para envios com base no tipo de remetente e calcula a interpretação agregada com base em todas as interpretações enviadas (HARRISON *et al*, 2016).

A plataforma ClinVar apresenta os resultados da busca por um gene em forma de tabela, contendo informações básicas como a localização da variação no gene; o(s) gene(s) envolvido(s) no caso; a mudança que ocorre na proteína (troca de aminoácido); caso disponível, as condições em que as mutações foram observadas; a significância clínica (se for patogênica, benigna ou se for desconhecida) e o *status* da revisão acerca daqueles casos específicos. A grande desvantagem da plataforma, é que como ela apresenta casos clínicos, frequentemente os pacientes relatados apresentam mais de uma mutação, o que torna difícil a identificação do efeito isolado da mutação de interesse, visto que a classificação de uma condição como patogênica ou não, leva em consideração a condição clínica do paciente como um todo, e não somente cada variação genética isoladamente. Dessa forma, não há meios de saber exatamente qual mutação está causando malignidade, o caso de um paciente com mais de uma. A pesquisa na plataforma foi feita utilizando o termo “glp1r”, que resultou em uma lista de casos em que mutações nesse gene foram observadas.

2.3 Algoritmos de predição para dano estrutural das variantes do tipo *missense*

As variantes encontradas foram então analisadas em 4 ferramentas de predição: *Align GVD*® (<http://agvgd.hci.utah.edu/about.php> - *Huntsman Cancer Institute, University of Utah* Estados Unidos), *SNAP2*® (<https://roslab.org/services/snap2web/> - ROSTLAB, Alemanha), *PANTHER*® (*Protein Analysis Through Evolutionary Relationships*; <http://www.pantherdb.org/> - *Geneontology Unifying Biology*, Estados Unidos) e *PROVEAN*® (*Protein Variation Effect Analyzer*; <https://www.jcvi.org/research/provean> - *J. Craig Venter Institute*, Estados Unidos). A aplicação dos diferentes filtros de uso de diversas ferramentas de predição de dano tem por objetivo somente recuperar as variantes de alto risco, classificadas como deletérias, ou seja, que tem a possibilidade de causar dano estrutural (HOSSAIN *et al*, 2020).

2.3.1 Modelos de predição de estrutura

Para a construção dos modelos de predição de estrutura, a análise foi realizada nas plataformas *HOPE Prediction*® (<https://www3.cmbi.umcn.nl/hope/> - *Centre for Molecular and Biomolecular Informatics* CMBI; Radboud University, Holanda) e *Swiss-Model*® (<https://swissmodel.expasy.org/> - BIOZENTRUM; *University of Basel*, Suíça). A plataforma HOPE é um serviço de acesso *online* que analisa, a partir de informações estruturais de diversas

fontes, anotações de sequência da base de dados proteicos UniProt® (<https://www.uniprot.org/>) e previsão do *software* Reprof, os efeitos estruturais de uma mutação pontual em uma sequência de aminoácidos. A ferramenta combina essas informações para analisar os efeitos de uma certa mutação específica na estrutura da proteína (EMADI *et al.*, 2020; SHAH *et al.*, 2022; KHIMSURIYA; CHAUHAN, 2018). A plataforma Swiss-Model® foi utilizada para a modelagem estrutural, que é baseada em bancos de dados de sequência e estrutura proteicas. Os modelos são calculados através de modelagem por homologia de servidor *Swiss-Model* (WATERHOUSE *et al.*, 2018) que se baseia no *ProMod3* (STUDER *et al.*, 2021).

A análise de interações preditas da proteína GLP-1R com outras proteínas foi realizada a partir das interações proteína-proteína, utilizando as ferramentas disponíveis na plataforma STRING®, cruciais na avaliação de todas as interações funcionais da proteína intracelularmente. A ferramenta de busca para a obtenção dos dados de proteínas ou genes que interagem entre si, foi a plataforma de rede de integração STRING® (<https://string-db.org/>), que possibilita a interação da proteína GLP-1R humana com outras proteínas. A STRING pode prever interações proteína-proteína diretas (físicas) ou indiretas (funcionais), através de associações entre uma proteína conhecida e outras proteínas, utilizando seu banco de dados de 5.214.234 proteínas de 1.113 organismos (ROTIMI *et al.*, 2018).

3. RESULTADOS

3.1 O gene GLP1R

O gene de interesse foi o GLP1R – *glucagon like peptide 1 receptor*, *Homo sapiens* (humano) Gene ID: 2740, GRCh38.p14, que codifica o receptor do *glucagon-like peptide 1* (GLP-1R).

3.2 Recuperação das variantes

A pesquisa inicial foi realizada no dia 23 de novembro de 2021 na plataforma dbSNP que apresentou 16.399 variantes para o gene GLP1R (chave de busca: `glp1r[All Fields]`), que incluíram mutações sinonímias (222), de deleção (4), variantes de transcrição não codificante (3.820), mutações de íntron (13.837) e do tipo *missense* (356). Para o presente trabalho foram utilizadas somente variantes do tipo *missense*, cuja alteração do nucleotídeo causa alteração no aminoácido sintetizado, conseqüentemente alterando a proteína final (Material Suplementar 1). As 356 variantes do tipo *missense* foram analisadas individualmente na ferramenta de predição *PolyPhen-2*. Após a análise, 164 variantes apresentaram escores entre 0,9 e 1,0, sugerindo dano estrutural significativo a proteína (Material suplementar 2). Somente as variantes com frequência alélica acima de 0,5 foram consideradas para as análises subsequentes. Desta forma, restaram 6 variantes. A Figura 5 ilustra os critérios e a obtenção das 6 variantes finais: **D344E, A239T, R310Q, R227H, R421P e R176G**.

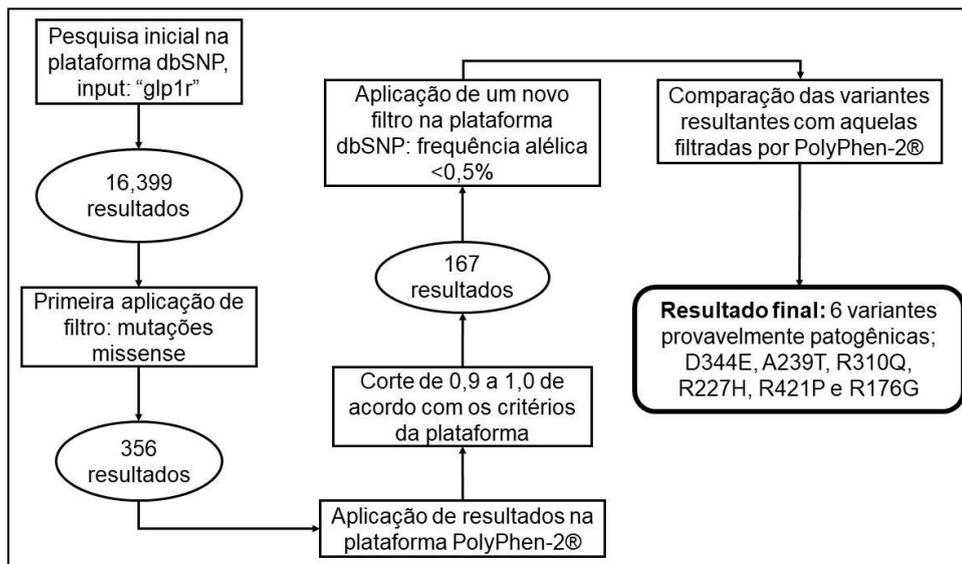


Figura 5. Fluxograma das etapas seguidas para a obtenção das 6 variantes finais.

As características das variantes estão listadas da Tabela 1 e distribuídas esquematicamente no sistema SNAKE, conforme Figura 6.

Tabela 1. Informações básicas das mutações selecionadas

ID das mutações	Troca de aminoácidos	Troca de nucleotídeos	Frequência Alélica (GnomAD)	Frequência Alélica (1000G)
rs202171972	Asp344Glu	D [GAC] > E [GAA]	A=0.000262	A=0.0006
rs140642887	Ala239Thr	A [GCC] > T [ACC]	A=0.000378	A=0.0014
rs149578908	Arg310Gln	R [CGG] > Q [CAG]	A=0.000008	A=0.0002
rs10305510	Arg421Pro	R [CGG] > P [CCG]	A=0.006088	A=0.0074
rs182447758	Arg227His	R [CGC] > H [CAC]	A=0.000016	A=0.0004
rs2235868	Arg176Gly	R [AGG] > G [GGG]	-	C=0.4453

Os dados compilados foram o número de identificação das mutações conforme o dbSNP (ID), a troca de aminoácidos, troca de nucleotídeos, frequência alélica segundo o banco GnomAD e frequênciaalélica segundo a plataforma 1000 Genomes ou 1000G.

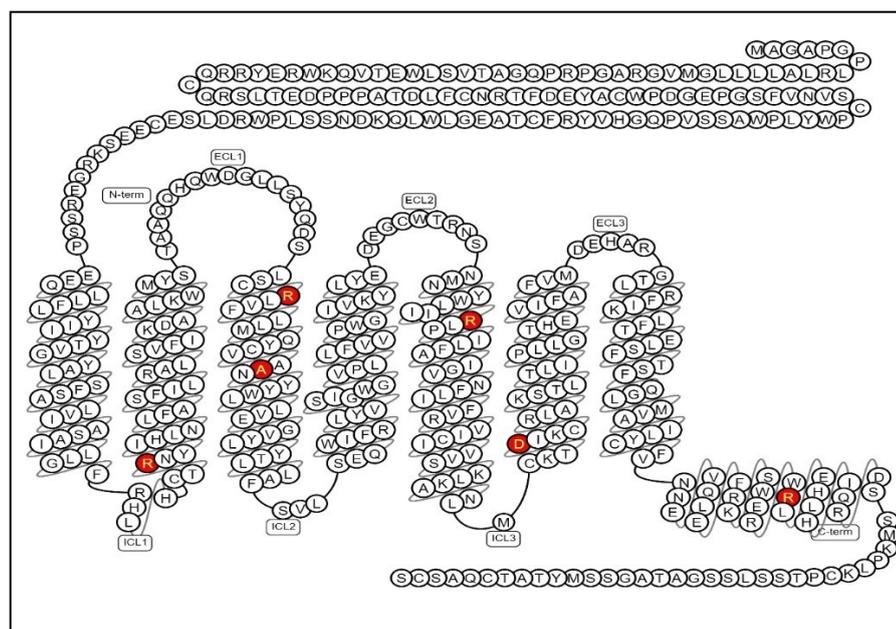


Figura 6. Seis variantes encontradas ao final do processo, destacadas no modelo de proteína do sistema SNAKE (https://gpcrdb.org/protein/glp1r_human/). As letras representam os aminoácidos da proteína GLP-1R. As alterações não foram observadas ao mesmo tempo, a figura é apenas uma representação bidimensional da localização de cada mutação.

Após a obtenção das variantes de interesse, as mesmas foram submetidas a análises na plataforma *HOPE® Prediction*. A proteína GLP-1R, ID P43220 (UniProt®), com 463 aminoácidos foi utilizada para base de análise de homologia estrutural na plataforma Hope. A estrutura tridimensional da proteína está ilustrada na Figura 7.

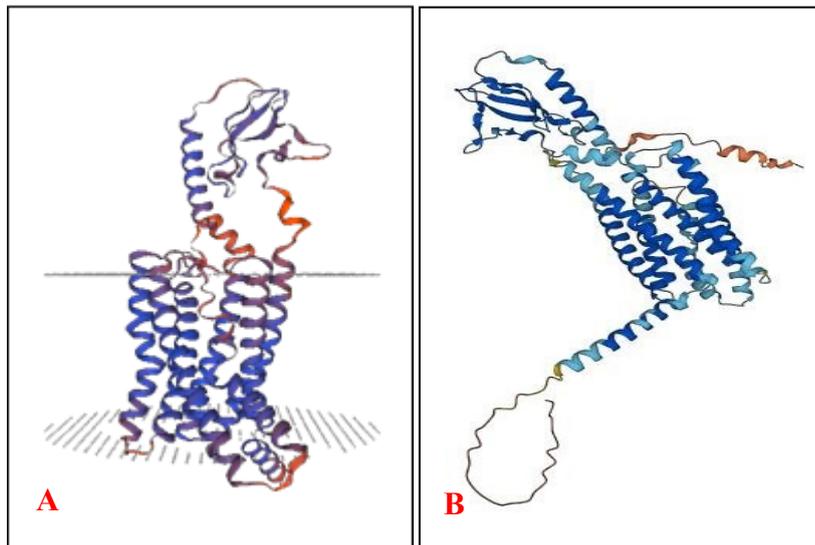


Figura 7. Visão geral da proteína na representação em fita. As imagens foram criadas, usando o ID da proteína UNIPROT® GLP-1R (463 aminoácidos)-p43220. **(A)** Imagem da plataforma Swiss-Model®, com a representação da membrana plasmática. **(B)** Imagem da plataforma AlphaFold®, na qual as cores da proteína representam diferentes níveis de confiabilidade. As porções azuis escuras são de alta fidelidade, as porções azuis claras são de fidelidade normal, as porções amarelas são de baixa fidelidade e as porções vermelhas são de fidelidade muito baixa. O AlphaFold® produz uma pontuação de confiança por resíduo (PLDDT) entre 0 e 100, e quanto maior a fidelidade, mais próximo o modelo é da realidade. Algumas regiões abaixo de 50 PLDDT podem ser destrutivas isoladamente.

3.2.1 Análise de variantes

3.2.1.1 Variante D344E

Para a variante D344E (uma troca de **ácido aspártico** para o **ácido glutâmico**), observou-se que o resíduo mutante foi maior que o resíduo do tipo selvagem (Figuras 8). Além disso, foram observadas ligações iônicas com um aminoácido de lisina na posição 346 e com uma arginina na posição 348, deslocando o resíduo mutante. O deslocamento do resíduo mutante altera a estrutura secundária da proteína, causando desestabilização na região. A análise também sugere que o resíduo do tipo selvagem está localizado em um ponto importante para o funcionamento da proteína, portanto, o resíduo mutado pode prejudicar a funcionalidade deste ponto específico. Outro ponto importante é que o resíduo do tipo selvagem foi considerado bem conservado e, embora a plataforma não tenha apresentado o resíduo mutante nessa posição em outras proteínas homólogas, foi sugerido que outros resíduos com propriedades em comum com o mutante ocorreram e isso indica que, em alguns casos raros, o resíduo mutado pode ocorrer sem prejudicar a função da proteína.

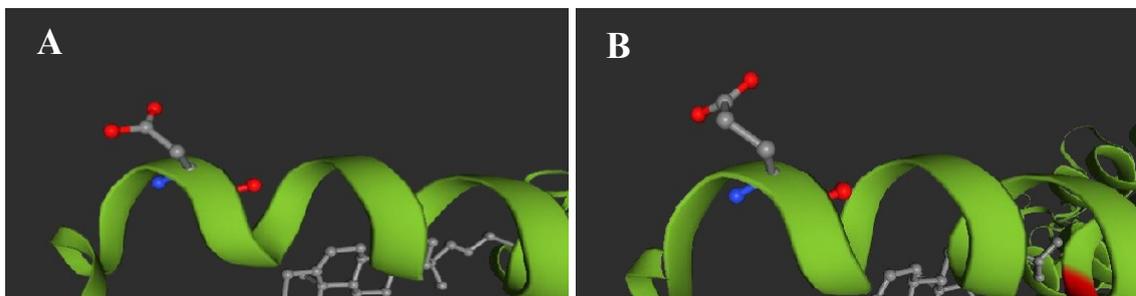


Figura 8. Detalhes do resíduo do tipo selvagem (A) em comparação com o resíduo mutante D344E (B). Em destaque a estrutura molecular química do aminoácido presente na posição de interesse, colorida em cinza (átomos de carbono), azul (átomos de nitrogênio) e vermelho (átomos de oxigênio). O restante da proteína (α -hélices) é colorido em verde.

3.2.1.2 Variante A239T

Para a variante A239T (uma troca de uma **alanina** para uma **treonina**), a ferramenta HOPE indicou que o resíduo mutante **239** foi maior e mais hidrofóbico que o resíduo do tipo selvagem. Ainda, o resíduo está localizado em um ponto importante da proteína e compromete seu correto funcionamento. A análise também revelou que o resíduo do tipo selvagem está no *core* proteico. Nesse sentido, o resíduo mutante de tamanho maior e maior hidrofobicidade pode

comprometer as interações hidrofóbicas no núcleo da proteína. Além disso, a análise revelou que o resíduo 239 está localizado no terceiro domínio transmembranar, entre as hélices extracelulares 1 e 2, sugerindo que a mutação pode comprometer a interação com outras hélices transmembranares ou mesmo com a membrana plasmática. O resíduo mutante T239 foi localizado próximo a uma região de alta conservação (Figura 9).

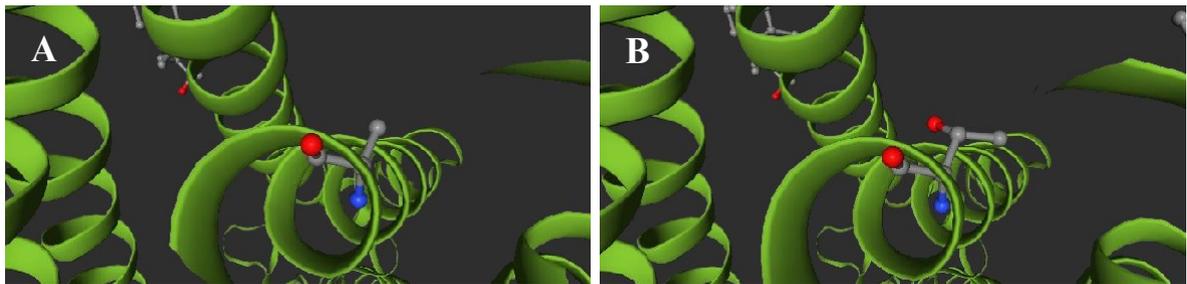


Figura 9. Detalhes do resíduo do tipo selvagem (A) em comparação com o resíduo mutante A239T (B). Em destaque a estrutura molecular química do aminoácido presente na posição de interesse, colorida em cinza (átomos de carbono), azul (átomos de nitrogênio) e vermelho (átomos de oxigênio). O restante da proteína (α -hélices) é colorido em verde.

3.2.1.3 Variante R310Q

Para a variante R310Q (uma troca de uma **arginina** para a **glutamina**), a ferramenta HOPE indicou que o resíduo mutante é menor que o resíduo do tipo selvagem e tem uma carga neutra, enquanto o resíduo do tipo selvagem tem uma carga positiva e, essa diferença pode comprometer interações iônicas deste resíduo (Figura 10). Está localizado em uma parte importante da proteína e por causar um resíduo menor, causa um espaço vazio no núcleo da proteína. O resíduo é muito bem conservado e nenhum outro resíduo semelhante foi encontrado na mesma região sendo caracterizado como uma mutação que provavelmente causa danos à proteína.

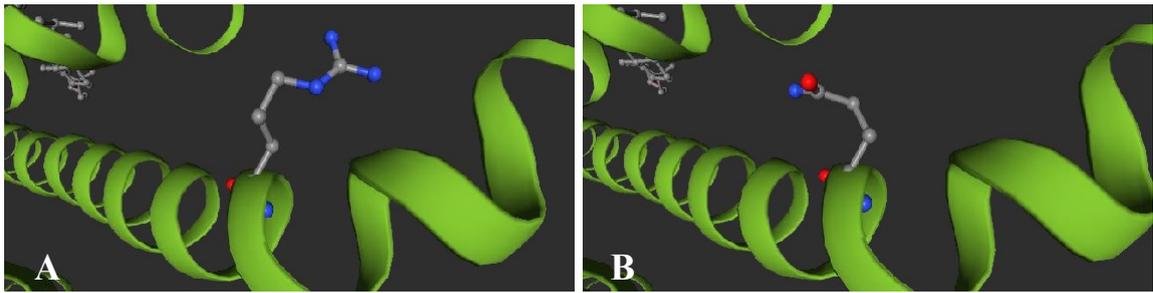


Figura 10. Detalhes do resíduo do tipo selvagem (A) em comparação com o resíduo mutante R310Q (B). Em destaque a estrutura molecular química do aminoácido presente na posição de interesse, colorida em cinza (átomos de carbono), azul (átomos de nitrogênio) e vermelho (átomos de oxigênio). O restante da proteína (α -hélices) é colorido em verde.

3.2.1.4 Variante R421P

Para a variante R421P (uma troca de uma **arginina** por **prolina**), a plataforma indicou que o resíduo mutante foi menor que o resíduo do tipo selvagem e tem uma carga neutra, enquanto o resíduo do tipo selvagem tem uma carga positiva (Figura 11). O resíduo mutante também é mais hidrofóbico. Esta mutação específica pode causar uma importante mudança estrutural, pois o resíduo do tipo selvagem é encontrado em uma α -hélice e, se a prolina não estiver localizada em posições específicas, sua conformação estrutural seria desestabilizada e, no caso dessa mutação específica, a α -hélice seria perturbada. Além disso, a diferença de tamanho pode prejudicar as interações externas da mesma. O local da mutação é bem conservado e, embora a plataforma não tenha observado este resíduo mutante especificamente nessa posição, encontrou outros similares, portanto há a possibilidade de o resíduo mutante ocorrer neste local da proteína.

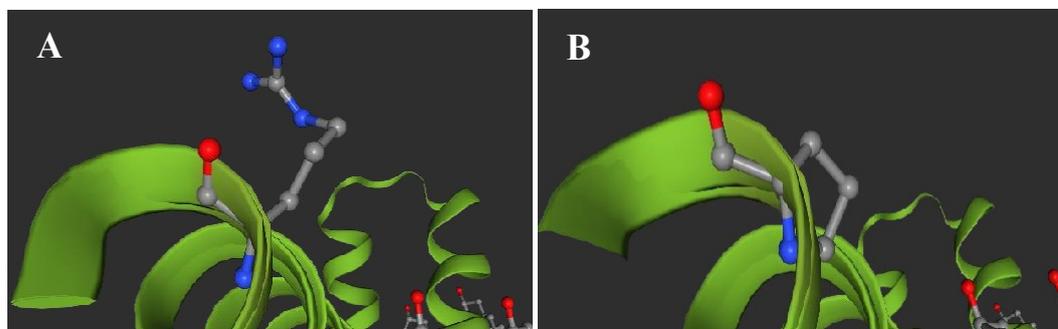


Figura 11. Detalhes do resíduo do tipo selvagem (A) em comparação com o resíduo mutante R421P (B). Em destaque a estrutura molecular química do aminoácido presente na posição de interesse, colorida em cinza (átomos de carbono), azul (átomos de nitrogênio) e vermelho (átomos de oxigênio). O restante da proteína (α -hélices) é colorido em verde.

3.2.1.5 Variante R227H

Para a variante R227H (uma troca de uma **arginina** para uma **histidina**), a plataforma indicou que o resíduo mutante é menor que o tipo selvagem e tem uma carga neutra, enquanto o tipo selvagem tem uma carga positiva, o que pode prejudicar as interações externas (Figura 12). A mutação está próxima de um local que forma uma ligação cisteína que não é diretamente afetada, mas a mutação próxima pode afetar a estabilidade dessa ligação. O resíduo é bem conservado, mas outras mutações com características semelhantes foram observadas no mesmo local, o que indica que em ocorrências raras, essa mutação pode ocorrer sem prejudicar a função da proteína.

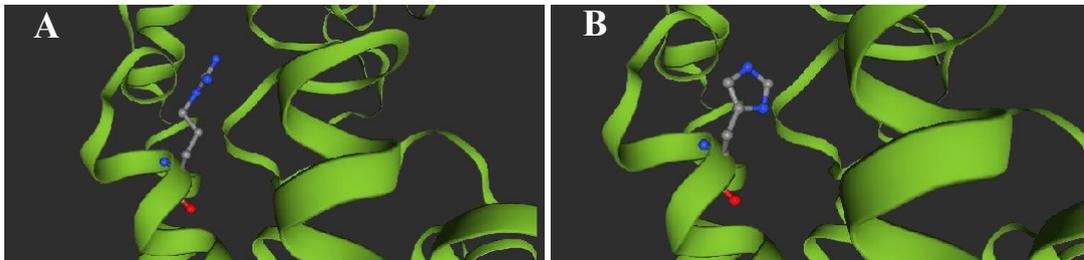


Figura 12. Detalhes do resíduo do tipo selvagem (A) em comparação com o resíduo mutante R227H (B). Em destaque a estrutura molecular química do aminoácido presente na posição de interesse, colorida em cinza (átomos de carbono), azul (átomos de nitrogênio) e vermelho (átomos de oxigênio). O restante da proteína (α -hélices) é colorido em verde.

3.2.1.6 Variante R176G

Para a variante R176G (uma troca de uma **arginina** por uma **glicina**), a plataforma indicou que o resíduo mutante foi menor e mais hidrofóbico que o tipo selvagem e tem uma carga neutra, enquanto o tipo selvagem tem uma carga positiva, que pode prejudicar interações externas (Figura 13). Esta região forma várias pontes salinas com outras regiões da proteína, que contribuem para a rigidez local e estabilidade. Esse resíduo está localizado em um domínio transmembranar e suas diferenças de carga e tamanho podem afetar as interações iônicas com outras hélices transmembranares, com outros domínios transmembranares, com a própria membrana e até perturbar as interações hidrofóbicas com o núcleo da proteína ou membrana lipídica. Para este caso especificamente, segundo a plataforma HOPE, a glicina é um aminoácido muito flexível, portanto, essa região da proteína não teria a rigidez normal, o que poderia prejudicar as interações externas. Essa mutação está em uma região de conservação muito alta, pois apenas essa variante foi identificada. Alterações nas regiões de alta conservação

tendem a ser prejudiciais à proteína.

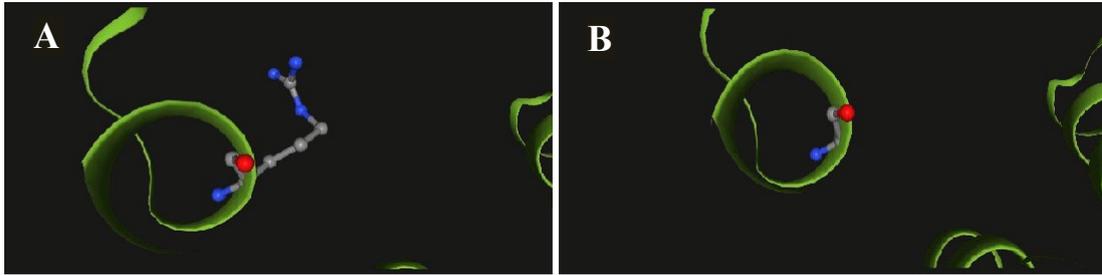


Figura 13. Detalhes do resíduo do tipo selvagem (A) em comparação com o resíduo mutante R176G (B). Em destaque a estrutura molecular química do aminoácido presente na posição de interesse, colorida em cinza (átomos de carbono), azul (átomos de nitrogênio) e vermelho (átomos de oxigênio). O restante da proteína (α -hélices) é colorido em verde.

Com base nos dados coletados, foi construída uma tabela (Tabela 2), sintetizando as informações mais relevantes sobre cada variante, para facilitar a visualização de informações específicas sobre as mesmas.

Identificação da Mutação	Tamanho do resíduo em comparação com WT	Carga	Conservação da posição
D344E	Maior	Não informado	Bem conservada
A239T	Maior	Não informado	Alta conservação
R310Q	Menor	Neutra (WT positiva)	Bem conservada
R421P	Menor	Neutra (WT positiva)	Bem conservada
R227H	Menor	Neutra (WT positiva)	Bem conservada
R176G	Menor	Neutra (WT positiva)	Alta conservação

Tabela 2. Informações da proteína retiradas da plataforma HOPE® prediction. As informações presentes na tabela são da região mutada em relação à *Wild Type*.

3.3 Plataforma STRING

Após a avaliação das variantes utilizando a plataforma HOPE, a proteína foi submetida a análise na plataforma STRING (<https://string-db.org/>), que permite avaliar as interações do modelo *Wild Type* com outras proteínas. A plataforma STRING, no entanto, não permite a avaliação de interações dos modelos com mutação, então não podemos afirmar como cada interação seria prejudicada, por cada variante, mas é possível que as interações apresentadas pela plataforma (Figura 14), sejam as mesmas que estariam prejudicadas conforme os achados na plataforma HOPE.

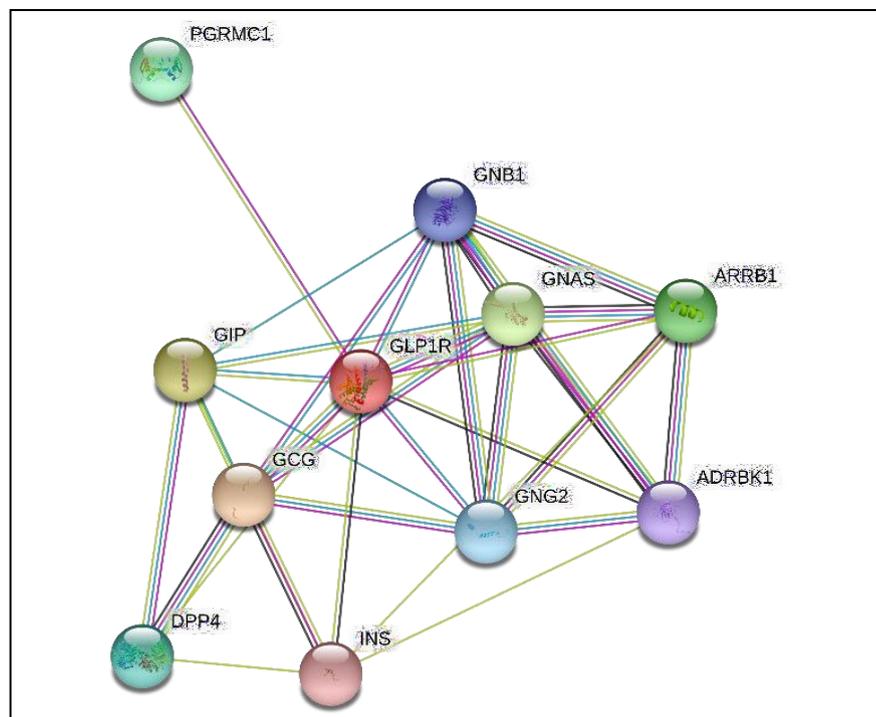


Figura 14. Rede de interações da proteína GLP-1R com outras proteínas ao seu redor em esquema de tipo de interação por evidência científica. A GLP-1R pode ser encontrada no centro, em vermelho, e cada linha leva a outra proteína com a qual a GLP-1R interage direta ou indiretamente (configuração da plataforma para não mais do que 10 interactores. Ponto de corte: 0.700).

Na plataforma STRING é possível também avaliar interações funcionais e interações físicas, que afetam a funcionalidade da proteína direta ou indiretamente. Nas figuras 15 e 16 é possível observar dois esquemas, um que mostra o interactoma funcional do GLP-1R (Figura

15) e um que mostra o interactoma físico (Figura 16). Foi definido ainda um ponto de corte de interações com nível de confiabilidade acima de 0.700, sendo que, quanto mais próximo de 1.0, mais confiável é a interação predita.

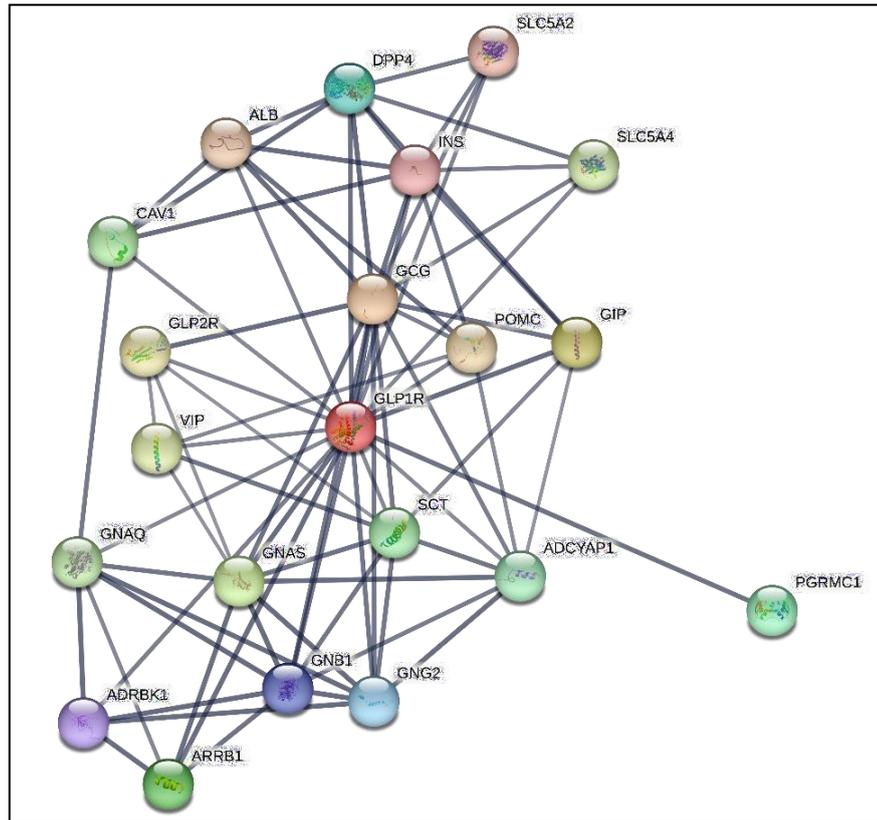


Figura 15. Rede de interações funcionais da proteína GLP-1R em esquema de confiabilidade. A proteína-alvo encontra-se colorida em vermelho, com seus parceiros funcionais ao seu redor (configuração da plataforma para não mais do que 20 interactivos. Ponto de corte: 0.700).

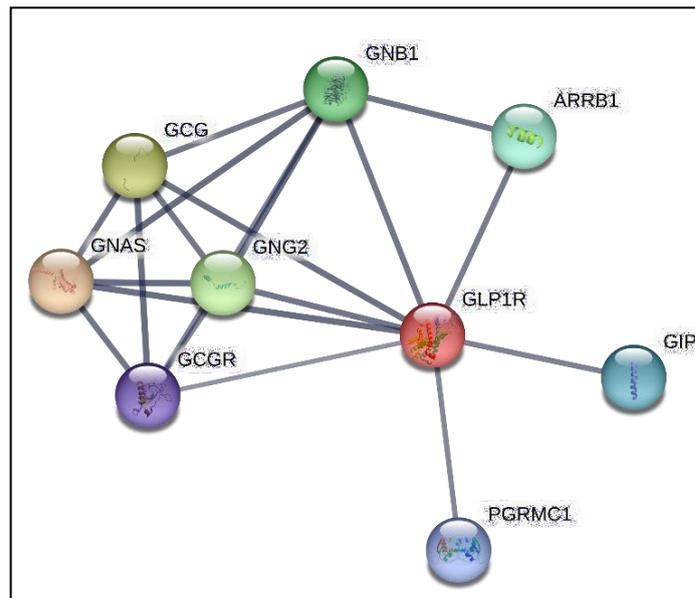


Figura 16. Rede de interações físicas da proteína GLP-1R em esquema de confiabilidade. A proteína-alvo encontra-se colorida em vermelho, com seus parceiros físicos ao seu redor (configuração da plataforma para não mais do que 10 interactores. Ponto de corte: 0.700).

A plataforma STRING® além de fornecer os dados já mencionados, permitiu analisar detalhadamente outros tipos de interação, entre as mesmas proteínas, por exemplo, interações que estão envolvidas em certos processos biológicos, conforme a figura 17. Nela é possível ver o mesmo interactoma de proteínas com interação física, porém destacando quais delas estão envolvidas em diversos processos biológicos que acontecem no corpo humano. É importante salientar que outros processos biológicos que não foram incluídos na imagem também podem ser avaliados na plataforma, porém foram desconsiderados por não serem relevantes ao tema do presente trabalho. Para a construção das figuras 17, 18, 19, 20 e 21, foram utilizados os mesmos critérios de construção das figuras 14 e 16 (não mais do que 10 interactores e ponto de corte: 0.700).

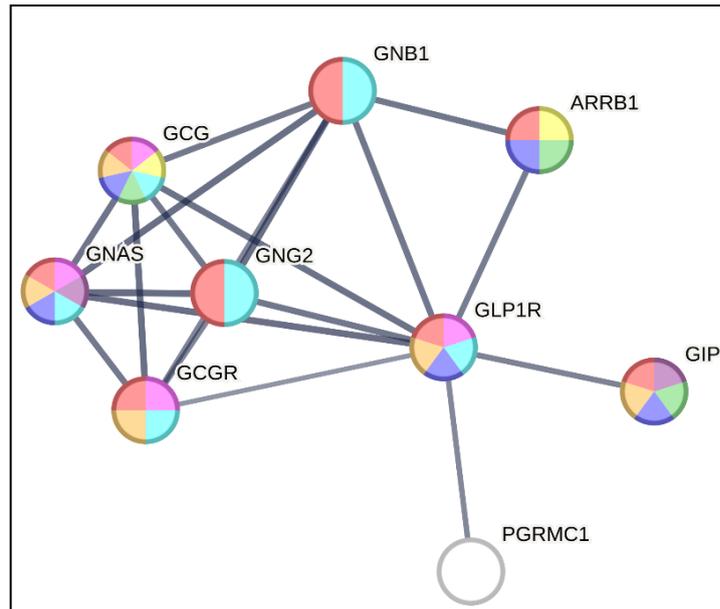


Figura 17. Rede de interações físicas da proteína GLP-1R em esquema de confiabilidade com destaque para processos biológicos específicos das proteínas (configuração da plataforma para não mais do que 10 interatores. Ponto de corte: 0.700). Em **vermelho**: Proteínas envolvidas no caminho de sinalização celular dos receptores acoplados à proteína G; em **azul escuro**: Proteínas envolvidas na regulação da secreção de insulina; em **verde**: Proteínas envolvidas na regulação positiva da secreção de insulina; em **amarelo**: Proteínas envolvidas na regulação positiva de secreção de insulina em resposta à estímulo de glicose; em **rosa**: Proteínas envolvidas na resposta celular ao estímulo por glucagon; em **azul claro**: Proteínas envolvidas no caminho de sinalização celular dos receptores acoplados à proteína G ativados por adenilato ciclase; em **laranja**: Proteínas envolvidas na resposta à hormônios peptídicos; e em **roxo**: Proteínas envolvidas na regulação positiva de sinalização celular mediada por cAMP.

Da mesma forma, outros destaques podem ser feitos no interactoma da proteína, relacionados a outros fatores, como função molecular, associação de gene com doença, tecido de expressão e até mesmo referência em publicação, a qual destaca artigos e pesquisas que reportaram a(s) interação(ões) entre as proteínas do interactoma. Com base nessas informações, foram criadas as Figuras 18, 19, 20 e 21, que são referentes aos esquemas citados acima, novamente tendo enfoque nos itens da plataforma relevantes ao tema do trabalho.

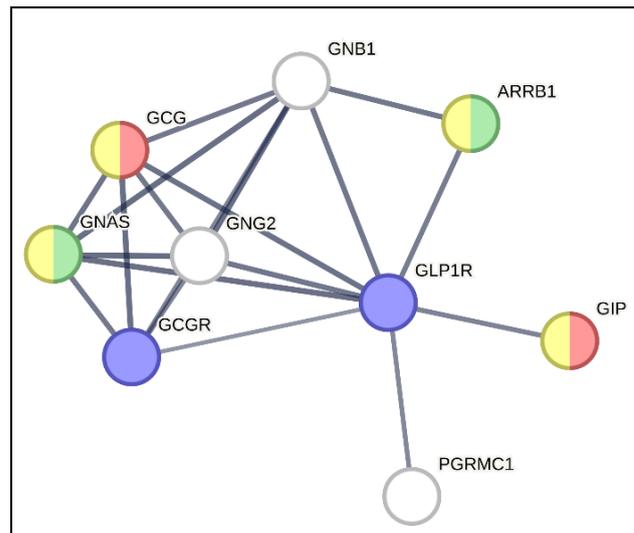


Figura 18. Rede de interações físicas da proteína GLP-1R em esquema de confiabilidade com destaque para função molecular das proteínas (configuração da plataforma para não mais do que 10 interactores. Ponto de corte: 0.700). Em **vermelho**: Ligação com o receptor de glucagon; em **azul**: Atividade do receptor de glucagon; em **verde**: Ligação com o receptor do fator de crescimento semelhante à insulina; e em **amarelo**: Ligação com o receptor acoplado à proteína G.

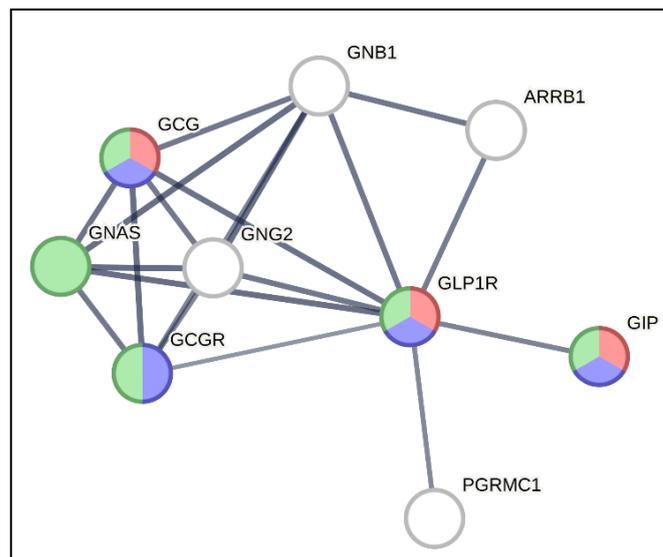


Figura 19. Rede de interações físicas da proteína GLP-1R em esquema de confiabilidade com destaque para associação entre gene das proteínas e doença (configuração da plataforma para não mais do que 10 interactores. Ponto de corte: 0.700). Em **vermelho**: Diabetes Mellitus tipo 2; em **azul**: Diabetes Mellitus; e em **verde**: Doença metabólica adquirida.

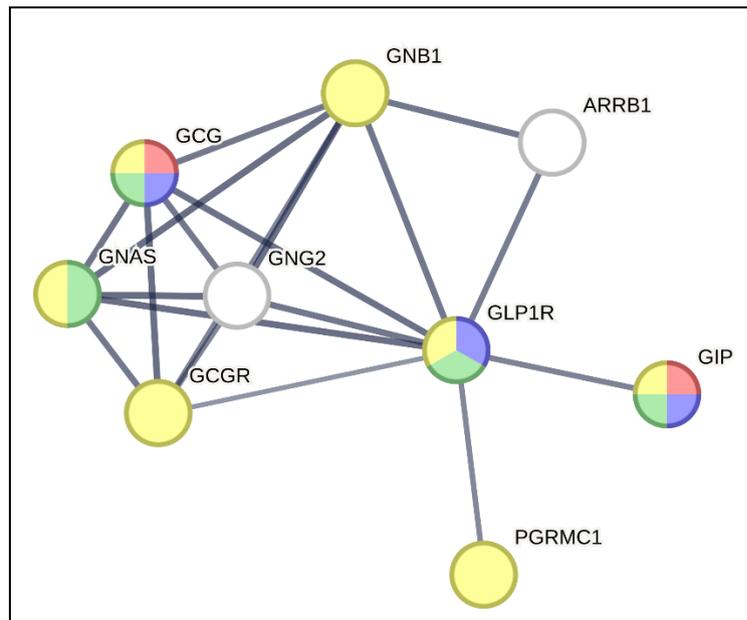


Figura 20. Rede de interações físicas da proteína GLP-1R em esquema de confiabilidade com destaque para tecidos em que as proteínas são atuantes (configuração da plataforma para não mais do que 10 interatores. Ponto de corte: 0.700). Em **vermelho**: Célula β -pancreática; em **azul**: Ilhota pancreática; em **verde**: Pâncreas; e em **amarelo**: Glândula digestiva.

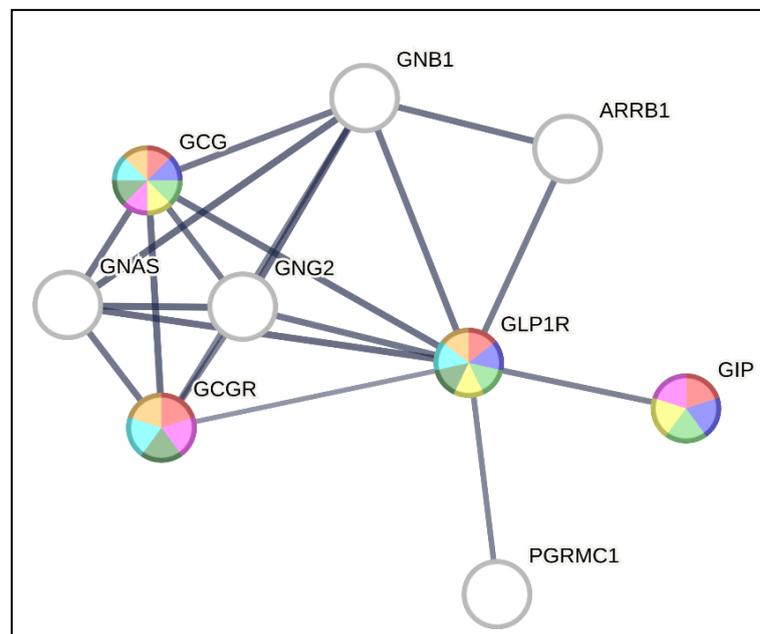


Figura 21. Rede de interações físicas da proteína GLP-1R em esquema de confiabilidade com destaque para interações reportadas em literatura (configuração da plataforma para não mais do que 10 interatores. Ponto de corte: 0.700). Em **vermelho**: (2022) *Discovery of MK-1462: GLP-1 and Glucagon Receptor Dual Agonist for the Treatment of Obesity and Diabetes*; em **azul**: (2023) *Gastrointestinal hormones: their increasing pharmacotherapeutic relevance in metabolic diseases*; em **verde**: (2022) *Glucagon-Like Peptide-1 Receptor Agonists and Dual Glucose-Dependent Insulinotropic Polypeptide Glucagon-Like Peptide-1 Receptor Agonists in the Treatment of Obesity Metabolic Syndrome, Prediabetes, Diabetes and Non-Alcoholic Fatty Liver*; em **amarelo**: (2022) *Clinical perspectives on the use of the GIP/GLP-1 receptor agonist*

tirzepatide for the treatment of type-2 diabetes and obesity; em **rosa:** (2022) *Glucagon-like peptide-1, glucose-dependent insulinotropic polypeptide, and glucagon receptor poly-agonists: a new era in obesity pharmacotherapy*; em **verde escuro:** (2022) *Efficacy and safety of high-dose glucagon-like peptide-1, glucagon-like peptide-1-glucose-dependent insulinotropic peptide, and glucagon-like peptide-1-glucagon receptor agonists in type 2 diabetes*; em **azul claro:** (2022) *Anti-obesity therapeutics with complementary dual-agonist activities at glucagon and glucagon-like peptide 1 receptors*; e em **laranja:** (2021) *Imaging of the Glucagon Receptor in Subjects with Type 2 Diabetes*.

Em meio a todas as interações proteicas destacadas no interactoma construído na plataforma STRING, destacam-se as proteínas GCG (Glucagon – escore de confiabilidade de 0.998), GIP (*Gastric Inhibitory Polypeptide* – escore de confiabilidade de 0.990) e GNAS (*Guanine Nucleotide-binding Protein Gs* – escore de confiabilidade de 0.988) no caso do interactoma de interações diretas e indiretas. O interactoma que leva em consideração apenas interações físicas diretas, no entanto, além da proteína GNAS (nesse caso com escore de 0.977) e da proteína GCG (escore de 0.959), foi observada também a proteína GNG2 (*Guanine Nucleotide-binding Protein G(I)/G(S)/G(O) subunit gamma-2* – escore de confiabilidade de 0.929). A plataforma STRING® compila informações a partir da plataforma UniProt, e traz as informações referentes a função de cada proteína que interage com o GLP-1R.

A proteína GCG, por exemplo, que possui o maior escore na plataforma, possui apenas uma evidência experimental que explora a sua interação com o GLP-1R, na plataforma BioGRID em um trabalho feito por Xiao *et al* em 2000. A plataforma em questão apontou que a interação foi observada apenas nas proteínas (GLP-1R e GCG) purificadas *in vitro*, com um *throughput* (taxa/quantidade de informações que passam por um sistema específico, nesse caso a plataforma BioGRAD) baixo, porém com um alto número de co-menções em *abstracts* da plataforma PubMed, o que acaba sendo responsável pelo seu escore alto.

A proteína GIP, possui apenas co-menção em artigos, como por exemplo Yang *et al* (2020) que apenas menciona que a proteína GIP tem um efeito similar ao GLP-1 na estimulação da secreção de insulina, o que já foi suficiente para a plataforma considerar para a construção do interactoma, e associação em bases de dados curadas (dados a partir de outras bases construtoras de interactoma, como por exemplo a base de dados Reactome® (Estados Unidos, <https://reactome.org/content/detail/R-HSA-420092>) que também apresenta alguma forma de interação das duas proteínas, sendo o suficiente para a plataforma STRING considerar na construção do interactoma), não tendo dados experimentais e/ou bioquímicos reportados na plataforma STRING. Em contrapartida, a proteína GNG2, não possui dados de co-menção em artigos, porém possui dados experimentais reportados na plataforma *Protein Data Bank* (PDB®) e também dados de associação reportados na base Reactome®.

A proteína GNAS teve um ensaio experimental reportado na plataforma BioGRID, de Dai *et al* (2015) que utilizou um sistema duplo híbrido de ubiquitina de levedura baseado em membrana (*Membrane Yeast Two-Hybrid* - MYTH), listou uma série de proteínas que interagem com GLP-1R no cérebro fetal humano que demonstraram atenuar a sua atividade, identificados a partir da biblioteca de cDNA de ilhotas humanas por MYTH. As proteínas interatoras incluíam a GNAS, que mostrou ativação da atividade da adenilato ciclase, e da GTPase quando em contato com GLP-1R (DAI *et al*, 2015). A proteína apresentou também dados de associação de interactoma da plataforma Reactome®, porém os dados de co-menção em *abstracts* do PubMed não encontraram publicações relevantes em organismos humanos, somente em outros organismos que não *Homo sapiens*.

3.4 Plataforma ClinVar

A pesquisa na plataforma ClinVar pelo termo “glp1r”, obteve 18 resultados, dos quais apenas 9 se tratavam de casos em que somente uma mutação no gene GLP1R foi observada. Em um dos casos, além de uma mutação no gene GLP1R, o paciente apresentava uma mutação no gene DNAH8 (*dynein axonemal heavy chain 8*), e mutações heterozigóticas compostas em DNAH8 podem ser responsáveis por múltiplas anormalidades morfológicas dos flagelos espermáticos (WENG *et al*, 2021), e nos 8 casos restantes, os pacientes apresentaram diversas outras alterações genéticas, em diversos outros genes, em conjunto com alguma mutação no gene GLP1R.

Todos os 9 casos com mutações observadas apenas no gene GLP1R, foram classificados como benignos, à exceção de um, que foi classificado como “provavelmente benigno”. Esta variante pode ser encontrada na plataforma dbSNP, sob o ID: rs61754624, e trata-se de uma mutação sinonímia, ou seja, a troca de nucleotídeo não altera o aminoácido encontrado na proteína, portanto não se observa malignidade. A variante foi citada em um trabalho sobre a ferramenta Sherlock, uma estrutura de classificação de variantes, que foi utilizada para refinar as diretrizes de 2015 do *American College of Medical Genetics and Genomics–Association for Molecular Pathology* (ACMG–AMP) (NYKAMP *et al*, 2017). No caso em que a mutação no gene GLP1R foi observada em conjunto com a mutação no DNAH8, a classificação foi incerta, pelo motivo de que não foram encontrados dados suficientes na plataforma para, tendo sido identificada apenas uma variante sob o ID: rs1340130, porém não ficou claro a qual proteína ela pertence e nem qual a troca de aminoácidos que ocorre.

Os 8 casos restantes, que apresentam diversas mutações em conjunto, foram classificados como patogênicos, porém poucas informações existem acerca dos mesmos. Por exemplo, dos 8, apenas 3 apresentam qual a proteína afetada pela mutação na plataforma

dbSNP, tratam-se dos IDs: rs155430, rs625706 e rs608764. São mutações em 3 proteínas diferentes, porém nenhuma delas relacionadas com o gene GLP1R. As 5 restantes não especificam qual proteína sofre a mutação na plataforma dbSNP, mas podem ser encontradas na mesma sob os IDs: rs443496, rs443497, rs608767, rs608768 e rs58427. Em se tratando das variantes benignas, podem ser encontradas na plataforma dbSNP sob os IDs: rs774310, rs780621, rs789746, rs712621, rs787828, rs779910 e rs730218, e todas são mutações sinonímias, com exceção da rs730218, que se trata da troca de uma arginina por uma histidina na posição 44, porém, ao passar por análise na plataforma PolyPhen-2®, foi classificada como benigna com escore de 0.2.

4 DISCUSSÃO

Este estudo compilou 6 variantes SNP *missense* do gene GLP1R, rs202171972 (D344E), rs140642887 (A239T), rs149578908 (R310Q), rs10305510 (R421P), rs182447758 (R227H) e rs2235868 (R176G). Alguns SNPs específicos referentes ao gene GLP1R foram investigados anteriormente, porém, os SNPs foram diferentes dos encontrados neste estudo. Michalowska e colaboradores, por exemplo, investigaram a associação de dois SNPs do gene GLP1R, rs2268641 e rs6923761 (ambos não apresentados em nossos resultados) com excesso de peso e síndrome metabólica. O estudo constatou que portadores de rs6923761 apresentavam maior risco de excesso de massa corporal, em comparação com portadores de rs2268641, que apresentam menor risco (MICHALOWSKA et al, 2022).

Outros três estudos avaliando a mutação rs6923761 (uma troca entre uma glicina e uma serina na posição 168 da proteína) chegaram a conclusões diferentes. Sathananthan e colaboradores, em estudo piloto, observaram que a variante não teve impacto significativo na concentração de glicose e na secreção de insulina após 120min. Javorski e colaboradores, seis anos depois, observaram que a variante estava significativamente associada à redução dos níveis de HbA1c. Dawed e colaboradores também observaram que a variante rs6923761 foi significativamente associada à resposta glicêmica aos agonistas de GLP-1R, mais especificamente o fato de que em meta-análises compostas por 4.571 indivíduos, cada cópia do alelo do gene correspondente à serina foi associada a uma redução menor de HbA1c. Eles também descobriram que essa associação era consistente entre os conjuntos de dados da meta-análise. Outra variante menos relevante, rs10305420, também foi estudada pelo mesmo grupo, mas eles não encontraram associação com a redução dos níveis de HbA1c na referida mutação (SATHANANTHAN et al, 2010; JAVORSKI et al, 2016; DAWED et al, 2023).

Em outro estudo foi revelado que a variante rs3765467 (R131Q) está significativamente associada ao inibidor de DPP-4. A dipeptidil peptidase-4 (DPP-4) é uma enzima peptidase que desativa o GLP-1, e o mesmo estudo apontou que a eficácia dos inibidores de DPP-4 pode ser afetada pelo GLP-1R, mais especificamente, essa alteração na estrutura ou afinidade de GLP-1R pode determinar a diferença nas respostas aos inibidores de DPP-4. Eles observaram uma redução nos níveis de HbA1c ao comparar pacientes que fizeram tratamento com inibidor de DPP-4 com pacientes que não fizeram (HAN et al, 2016).

Recentemente, outro estudo realizado em torno das variantes do gene GLP1R encontrou inúmeras mutações em relação ao receptor e como elas afetam a atividade da proteína. Eid e colaboradores fizeram uma lista com 21 variantes do gene, mas consideraram um MAF (*Minor Allele Frequency*) significativamente maior para construir sua lista, e de todos os SNPs obtidos, apenas dois são citados nesta pesquisa, rs140642887 (A239T) e

rs202171972 (D344E), ambos inicialmente encontrados por Lagou e colaboradores, em 2021. A lista também apresentava duas variantes na posição 421, sendo uma delas a troca de uma arginina por triptofano e outra de uma arginina por uma glutamina, já a variante da mesma posição citada neste trabalho (rs10305510), trata-se de uma arginina para uma prolina. Eid e colaboradores mencionam que ainda não está claro se esses SNPs (e até mesmo como) modificam o efeito da incretina GLP-1 e seus agonistas, e mais experimentos dedicados são necessários, mas é importante mencionar que uma quantidade crescente de dados na literatura documentou SNPs de ocorrência natural no locus GLP1R, que estão associados a resultados clínicos específicos, como níveis de glicose, respostas glicêmicas, perda de peso, DM2 e risco cardiovascular (EID *et al*, 2022).

Tendo em vista que o GLP-1R é um alvo estabelecido para perda de peso e hipoglicemia em doenças como obesidade e diabetes via tratamento com semaglutida e exenatida, por exemplo, outro estudo para avaliar se as variações podem ser prejudiciais ou não foi realizado. Lagou e colaboradores mediram respostas a vários agonistas endógenos e farmacológicos de GLP-1R e, concentrando-se em dois SNPs específicos, rs10305492 (A316T) e rs146868158 (R421W), observaram que a variante A316T apresentou respostas aumentadas aos referidos agonistas enquanto que a variante R421W apresentou respostas reduzidas a todos os ligantes, com exceção da exendina-4, que não teve atividade em ambas as variantes, e a semaglutida, que não teve efeito apenas para a variante A316T. A variante R421W apresentou uma leve perda de função, o que é significativo para os resultados encontrados nesta pesquisa, visto que uma das mutações listadas é a variante rs10305510, correspondente à mutação R421P (LAGOU *et al*, 2021).

Especificamente sobre as mutações filtradas pelo presente estudo, Gao *et al* (2022) em uma pesquisa que avaliou os efeitos de 60 variantes de codificação de GLP-1R na sinalização celular, separando-a em quatro vias de sinalização, acúmulo de cAMP, mobilização de Ca²⁺, ativação de ERK (*Extracellular signal-regulated kinases*) e o recrutamento de β -arrestina 2 (β -arr2), que pode servir como via secundária para a ativação do GLP-1R. Foi descoberto pelo mesmo estudo que as variantes A239T e D344E apresentavam um defeito na via de ativação de ERK, afetando o correto funcionamento do GPCR, e estas duas variantes foram consideradas prejudiciais ao funcionamento correto do GLP-1R. Outra variante, R310Q, foi incapaz de recrutar β -arr2, mas mostrou ativação significativa da via cAMP dependente de Gs, assim como outras 5 variantes não apresentadas aqui. Outros achados importantes em relação ao R310Q foram que a via de sinalização de cAMP não respondeu ao tratamento com exenatida-4 e também que foi totalmente inativa a 0,1 nM de exenatida-4, que é uma concentração que promove a secreção máxima de insulina para o gene *wild type* do GLP1R. O mesmo estudo também descobriu que duas variantes, R176W e

R310W (que apresentam substituições diferentes das filtradas por esta pesquisa, mas ocorrem nos mesmos loci, R176G e R310Q) diminuíram significativamente a produção de cAMP, classificando ambas as mutações como “expressão de superfície severamente defeituosa” (GAO *et al*, 2022).

Esses achados comprovam que o GLP-1R está ligado diretamente à obesidade e ao diabetes, daí a importância de estudá-lo. Nesta pesquisa, foi possível filtrar 6 mutações que segundo dados recolhidos nas plataformas online são susceptíveis de causar danos à estruturada proteína, e 3 delas, como mostrado anteriormente, podem ser prejudiciais para o funcionamento correto da proteína, pelo que devem ser observados com mais cuidado e atenção em possíveis estudos futuros. Considerando os dados obtidos na plataforma HOPE Prediction®, um fato interessante observado foi que, das 6 variantes filtradas, 4 apresentam trocas dos aminoácidos arginina (R) por outros aminoácidos, sendo que a mesma prevalência foi observada em outras mutações que não foram listadas. Por isso, levanta-se a possibilidade de correlação de dano aumentado no caso de mutações que apresentam troca de arginina, mas são necessários mais dados para confirmar essa hipótese.

Sobre os dados encontrados na plataforma STRING, é de extrema importância frisar que ela utiliza um algoritmo para coletar, pontuar e integrar todas as fontes publicamente disponíveis de informações de interação proteína-proteína e complementá-las com previsões computacionais e, a partir disso, formular os escores de confiabilidade das proteínas que interagem com o GLP-1R, com base em *text mining* automático da literatura científica, previsões de interação computacional de co-expressão, contexto genômico conservado, bancos de dados de experimentos de interação e complexos ou caminhos de interação proteica conhecidos de outras bases de pesquisa curadas (SZKLARCZYK *et al*, 2018; SZKLARCZYK *et al*, 2022).

Os bancos de dados de construção de interactomas proteicos são divididos em três categorias: as plataformas de interação primária, que anotam evidências de interação experimental diretamente das publicações de origem e acabam por fornecer serviços agregados altamente valiosos, como curadoria de metadados, manutenção de espaços de nomes comuns e criação de ontologias e padrões; as plataformas de previsão computacional, que são hospedadas por bancos de dados dedicados, como a PrePPI (<http://bhapp.c2b2.columbia.edu/PrePPI> - Columbia University Medical Center, Estados Unidos) que combina interações proteína-proteína (PPIs) previstas e determinadas experimentalmente e a PIPs Database (<http://www.compbio.dundee.ac.uk/www-pips> - University of Dundee, Escócia), que contém previsões de >37.000 interações de alta probabilidade que foram calculadas por meio de um método bayesiano que combina informações de expressão, ortologia, co-ocorrência de domínio, modificações pós-traducionais e localização subcelular da proteína (MCDOWAL *et al*, 2008;

ZHANG *et al*, 2012); e por fim, os bancos de dados dedicados a interações de proteínas no escopo mais amplo, que integram interações primárias e previstas, mas incluem também resultados de *text mining* (co-menção em *abstracts*), transferências entre organismos ou outras informações adicionais, categoria à qual o banco de dados STRING pertence (SZKLARCZYK *et al*, 2018; SZKLARCZYK *et al*, 2022). Por se tratar de uma plataforma de construção de interactoma abrangente, os escores não são construídos somente a partir de dados experimentais sobre as proteínas, mas também a partir de menção em artigos de bases de dados, portanto as possíveis interações apontadas pela ferramenta precisam ser observadas com muita cautela, pois existe a possibilidade de que as proteínas não interajam de fato, então, a depender do objetivo da pesquisa, os dados da plataforma precisam ser analisados minuciosamente.

5 CONCLUSÃO

Diabetes e obesidade são sérios problemas de saúde pública em todo o mundo, como mencionado anteriormente, e é por isso que mais pesquisas sobre essas doenças são importantes. Como já destacado, o fator genético é muito importante no desenvolvimento de ambas as patologias e deve ser considerado em estudos futuros sobre o assunto. Existem mais genes envolvidos com diabetes e obesidade do que GLP1R, genes que devem ser estudados para elucidar os diferentes mecanismos genéticos por trás dessas doenças. São necessárias pesquisas mais aprofundadas não apenas sobre o gene GLP1R, mas também sobre outros genes possivelmente envolvidos com diabetes e obesidade, não apenas para melhor elucidar os mecanismos por trás dessas doenças, mas também para encontrar novas e formas mais eficientes de garantir maior eficácia terapêutica. Embora as 6 variantes listadas neste trabalho não tenham relatos clínicos de sua ocorrência, fica claro o possível impacto que elas podem causar, considerando a atividade do GLP-1R em relação ao diabetes e obesidade, aumentando assim a importância do estudo desse gene para ajudar na prevenção o desenvolvimento dessas doenças em pessoas suscetíveis a elas desde a infância.

Os relatórios que pesquisaram anteriormente as variantes do GLP1R deixaram claro que há muito mais a ser estudado em relação às mutações genéticas que podem causar doenças ligadas ao referido receptor. Não apenas o diabetes e a obesidade, que são o foco principal deste estudo, mas várias outras patologias que estão ligadas a incretina GLP-1 e seus agonistas, dada a abundância deste receptor nos tecidos do corpo humano. Uma conclusão comum entre todas as pesquisas anteriores sobre o GLP-1R, é que existem poucos estudos sobre seu funcionamento exato, e mais pesquisas são necessárias para acessar exatamente como essas mutações afetam o gene e como elas podem ser tratadas. Em última análise, considerando os achados de outros estudos sobre variantes de GLP1R, e especificamente os encontrados neste estudo, o método de filtragem on-line e pesquisa de mutações prováveis nocivas aplicado por este trabalho se mostrou eficaz, considerando que a maioria das mutações listadas já foram previamente documentadas na literatura, e, mesmo para as variantes que não foram, foram relatadas outras variantes nos mesmos loci, mais uma vez comprovando a eficácia do método proposto neste estudo. Portanto, os resultados obtidos sugerem que o método pode servir com confiabilidade como uma diretriz segura para futuros estudos na área.

REFERÊNCIAS

MILLS, Dora Anne. Chronic Disease: The Epidemic of The Twentieth Century. **Maine Policy Review**, Orono, v. 9, n. 1, p. 50-65, jan. 2000. Disponível em:

<https://digitalcommons.library.umaine.edu/cgi/viewcontent.cgi?article=1287&context=mpr>.

Acesso em: 11 set. 2022.

HALL, Kevin D. *et al.* The Progressive Increase of Food Waste in America and Its Environmental Impact. **Plos One**, [S.L.], v. 4, n. 11, p. 7940-7946, 25 nov. 2009. Public Library of Science (PLoS). <http://dx.doi.org/10.1371/journal.pone.0007940>. Disponível em:

<https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0007940>. Acesso em: 18

maio 2023.

CHURCH, Timothy S. *et al.* Trends over 5 Decades in U.S. Occupation-Related Physical Activity and Their Associations with Obesity. **Plos One**, [S.L.], v. 6, n. 5, p. 19657-19664, 25 maio 2011. Public Library of Science (PLoS).

<http://dx.doi.org/10.1371/journal.pone.0019657>. Disponível em:

<https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0019657#s4>. Acesso em: 18

maio 2023.

VON LOEFFELHOLZ, Christian *et al.* Non-Exercise Activity Thermogenesis in Human Energy Homeostasis. In: FEINGOLD, Kenneth *et al.* **Endotext [Internet]**. South Dartmouth (Ma): Mdtex.Com, Inc., 2000. Disponível em: <https://pubmed.ncbi.nlm.nih.gov/25905303/>.

Acesso em: 18 maio 2023.

MCALLISTER, Emily J. *et al.* Ten Putative Contributors to the Obesity Epidemic. **Critical Reviews In Food Science And Nutrition**, [S.L.], v. 49, n. 10, p. 868-913, 2 dez. 2009.

Informa UK Limited. <http://dx.doi.org/10.1080/10408390903372599>.

HEYMSFIELD, Steven B.; WADDEN, Thomas A.. Mechanisms, Pathophysiology, and Management of Obesity. **New England Journal Of Medicine**, [S.L.], v. 376, n. 3, p. 254-266, 19 jan. 2017. Massachusetts Medical Society. <http://dx.doi.org/10.1056/nejmra1514009>.

BURKI, Talha. European Commission classifies obesity as a chronic disease. **The Lancet Diabetes & Endocrinology**, [S.L.], v. 9, n. 7, p. 418, jul. 2021. Elsevier BV. [http://dx.doi.org/10.1016/s2213-8587\(21\)00145-5](http://dx.doi.org/10.1016/s2213-8587(21)00145-5).

BROUWERS, Bas *et al.* Human MC4R variants affect endocytosis, trafficking and dimerization revealing multiple cellular mechanisms involved in weight regulation. **Cell Reports**, [S.L.], v. 34, n. 12, p. 108862-108883, mar. 2021. Elsevier BV. <http://dx.doi.org/10.1016/j.celrep.2021.108862>. Disponível em: <https://www.sciencedirect.com/science/article/pii/S2211124721001765?via%3Dihub>. Acesso em: 18 maio 2023.

FIELD, Alison E.; COAKLEY, Eugenie H.; MUST, Aviva *et al.* Impact of Overweight on the Risk of Developing Common Chronic Diseases During a 10-Year Period. **Archives Of Internal Medicine**, [S.L.], v. 161, n. 13, p. 1581-1596, 9 jul. 2001. American Medical Association (AMA). <http://dx.doi.org/10.1001/archinte.161.13.1581>.

LOOS, Ruth J. F.; YEO, Giles S. H.. The genetics of obesity: from discovery to biology. **Nature Reviews Genetics**, [S.L.], v. 23, n. 2, p. 120-133, 23 set. 2021. Springer Science and Business Media LLC. <http://dx.doi.org/10.1038/s41576-021-00414-z>.

KOHLSDORF, Katja; NUNZIATA, Adriana; FUNCKE, Jan-Bernd *et al.* Early childhood BMI trajectories in monogenic obesity due to leptin, leptin receptor, and melanocortin 4 receptor deficiency. **International Journal Of Obesity**, Berlim, v. 42, n. 1, p. 1602-1609, fev. 2018.

WU, Yili *et al.* Genetics of Obesity Traits: a bivariate genome-wide association analysis. **Frontiers In Genetics**, [S.L.], v. 9, n. [], p. 179-188, 16 maio 2018. Frontiers Media SA. <http://dx.doi.org/10.3389/fgene.2018.00179>. Disponível em:

<https://www.frontiersin.org/articles/10.3389/fgene.2018.00179/full>. Acesso em: 18 maio 2023.

KHERA, Amit V.; CHAFFIN, Mark; WADE, Kaitlin H. *et al.* Polygenic Prediction of Weight and Obesity Trajectories from Birth to Adulthood. **Cell**, [S.L.], v. 177, n. 3, p. 587-596, abr. 2019. Elsevier BV. <http://dx.doi.org/10.1016/j.cell.2019.03.028>.

BROUWERS, Bas; OLIVEIRA, Edson Mendes de; MARTI-SOLANO, Maria *et al.* Human MC4R variants affect endocytosis, trafficking and dimerization revealing multiple cellular mechanisms involved in weight regulation. **Cell Reports**, [S.L.], v. 34, n. 12, p. 1-17, mar. 2021. Elsevier BV. <http://dx.doi.org/10.1016/j.celrep.2021.108862>.

MENG, Xiang-Rui; SONG, Jie-Yun; MA, Jun *et al.* Association study of childhood obesity with eight genetic variants recently identified by genome-wide association studies. **Pediatric Research**, [S.L.], v. 76, n. 3, p. 310-315, 23 jun. 2014. Springer Science and Business Media LLC. <http://dx.doi.org/10.1038/pr.2014.88>.

SILVA, Carmela Farias da; ZANDONÁ, Marília Remuzzi; VITOLO, Márcia Regina *et al.* Association between a frequent variant of the FTO gene and anthropometric phenotypes in Brazilian children. **Bmc Medical Genetics**, [N.I.], v. 14, n. 34, p. 1-8, mar. 2013.

MÜLLER, T.D.; FINAN, B.; BLOOM, S.R. *et al.* Glucagon-like peptide 1 (GLP-1). **Molecular Metabolism**, [S.L.], v. 30, p. 72-130, dez. 2019. Elsevier BV. <http://dx.doi.org/10.1016/j.molmet.2019.09.010>.

TRENSON, Liesbet *et al.* Liraglutide for Weight Management in the Real World: significant weight loss even if the maximal daily dose is not achieved. **Obesity Facts**, [S.L.], v. 15, n. 1, p. 83-89, 22 nov. 2021. S. Karger AG. <http://dx.doi.org/10.1159/000520217>.

TILINCA, Mariana Cornelia *et al.* A 2021 Update on the Use of Liraglutide in the Modern Treatment of ‘Diabesity’: a narrative review. **Medicina**, [S.L.], v. 57, n. 7, p. 669-710, 29 jun. 2021. MDPI AG. <http://dx.doi.org/10.3390/medicina57070669>.

RUBINO, Domenica M. *et al.* Effect of Weekly Subcutaneous Semaglutide vs Daily Liraglutide on Body Weight in Adults With Overweight or Obesity Without Diabetes. **Jama**, [S.L.], v. 327, n. 2, p. 138-151, 11 jan. 2022. American Medical Association (AMA). <http://dx.doi.org/10.1001/jama.2021.23619>.

PANTALONE, Kevin M; HOBBS, Todd M; CHAGIN, Kevin M. *et al.* Prevalence and recognition of obesity and its associated comorbidities: cross-sectional analysis of electronic health record data from a large us integrated health system. **Bmj Open**, [S.L.], v. 7, n. 11, p. 1-12, nov. 2017. BMJ. <http://dx.doi.org/10.1136/bmjopen-2017-017583>.

HOELZ, L. V. B.; FREITAS, G. B. L. de; TORRES, P. H. M. *et al.* G Protein-Coupled Receptors. **Revista Virtual de Química**, [S.L.], v. 5, n. 5, p. 981-1000, set. 2013. Sociedade Brasileira de Química (SBQ). <http://dx.doi.org/10.5935/1984-6835.20130071>.

WATKINS, Lyndsay R.; ORLANDI, Cesare. Orphan G Protein Coupled Receptors in Affective Disorders. **Genes**, [S.L.], v. 11, n. 6, p. 694-724, 24 jun. 2020. MDPI AG. <http://dx.doi.org/10.3390/genes11060694>.

YANG, Dehua; ZHOU, Qingtong; LABROSKA, Viktorija. *et al.* G protein-coupled receptors: structure- and function-based drug discovery. **Signal Transduction And Targeted Therapy**, [S.L.], v. 6, n. 1, p. 1-9, 8 jan. 2021. Springer Science and Business Media LLC. <http://dx.doi.org/10.1038/s41392-020-00435-w>.

LIU, Nannan *et al.* G-Protein Coupled Receptors (GPCRs): signaling pathways, characterization, and functions in insect physiology and toxicology. **International Journal Of Molecular Sciences**, [S.L.], v. 22, n. 10, p. 5260-5279, 17 maio 2021. MDPI AG. <http://dx.doi.org/10.3390/ijms22105260>.

BAKBAK, Ehab; TERENZI, Daniella C.; TRAC, Justin Z. *et al.* Lessons from bariatric surgery: can increased glp-1 enhance vascular repair during cardiometabolic-based chronic disease?. **Reviews In Endocrine And Metabolic Disorders**, [S.L.], v. 22, n. 4, p. 1171-1188, 6 jul. 2021. Springer Science and Business Media LLC. <http://dx.doi.org/10.1007/s11154-021-09669-7>.

MORETTO, Maria Clara *et al.* Associação entre cor/raça, obesidade e diabetes em idosos da comunidade: dados do estudo fibra. **Cadernos de Saúde Pública**, [S.L.], v. 32, n. 10, out. 2016. FapUNIFESP (SciELO). <http://dx.doi.org/10.1590/0102-311x00081315>.

ZHAO, Xin *et al.* GLP-1 Receptor Agonists: beyond their pancreatic effects. **Frontiers In Endocrinology**, [S.L.], v. 12, n. [], p. 721135-721154, 23 ago. 2021. Frontiers Media SA. <http://dx.doi.org/10.3389/fendo.2021.721135>.

BOUTARI, Chrysoula; MANTZOROS, Christos S.. A 2022 update on the epidemiology of obesity and a call to action: as its twin covid-19 pandemic appears to be receding, the obesity and dysmetabolism pandemic continues to rage on. **Metabolism**, [S.L.], v. 133, n. 11, p. 126-133, ago. 2022. Elsevier BV. <http://dx.doi.org/10.1016/j.metabol.2022.155217>.

ROSS, Robert; NEELAND, Ian J.; YAMASHITA, Shizuya. *et al.* Waist circumference as a vital sign in clinical practice: a Consensus Statement from the IAS and ICCR Working Group on Visceral Obesity. **Nature Reviews Endocrinology**, [N.I.], v. 16, n. [], p. 177-189, fev. 2020.

LIN, Xihua; LI, Hong. Obesity: Epidemiology, Pathophysiology, and Therapeutics. **Frontiers In Endocrinology**, [N.I.], v. 78, n. [], p. 1-15, set. 2021.

WORLD HEALTH ORGANIZATION. **Obesity**. Disponível em: https://www.who.int/health-topics/obesity#tab=tab_1. Acesso em: 12 set. 2022.

WORLD HEALTH ORGANIZATION. **Diabetes**. Disponível em: <https://www.who.int/news-room/fact-sheets/detail/diabetes> 1. Acesso em: 12 set. 2022.

COLE, Joanne B.; FLOREZ, Jose C.. Genetics of diabetes mellitus and diabetes complications. **Nature Reviews Nephrology**, [N.I.], v. 16, n. [], p. 377-390, maio 2020.

HARREITER, Jürgen; RODEN, Michael. Diabetes mellitus – Definition, Klassifikation, Diagnose, Screening und Prävention (Update 2019). **Wiener Klinische Wochenschrift**, [S.L.], v. 131, n. 1, p. 6-15, 12 abr. 2019. Springer Science and Business Media LLC. <http://dx.doi.org/10.1007/s00508-019-1450-4>.

AMERICAN DIABETES ASSOCIATION, **Classification and Diagnosis of Diabetes: standards of medical care in diabetes**.:2022. Standards of Medical Care in Diabetes—2022. Disponível em: https://diabetesjournals.org/care/article/45/Supplement_1/S17/138925/2-Classification-and-Diagnosis-of-Diabetes. Acesso em: 13 set. 2022.

LIN, Xiling; XU, Yufeng; PAN, Xiaowen. *et al.* Global, regional, and national burdens and trend of diabetes in 195 countries and territories: an analysis from 1990 to 2025. **Nature: Scientific Reports**, [s. l], v. 10, n. 14790, p. 1-11, 2020.

KHAWANDANAH, Jomana. Double or hybrid diabetes: a systematic review on disease prevalence, characteristics and risk factors. **Nutrition & Diabetes**, [S.L.], v. 9, n. 1, p. 1-9, 4 nov. 2019. Springer Science and Business Media LLC. <http://dx.doi.org/10.1038/s41387-019-0101-1>.

ILONEN, Jorma; LEMPAINEN, Johanna; VEIJOLA, Riitta. The heterogeneous pathogenesis of type 1 diabetes mellitus. **Nature Reviews Endocrinology**, [S.L.], v. 15, n. 11, p. 635-650, 18 set. 2019. Springer Science and Business Media LLC. <http://dx.doi.org/10.1038/s41574-019-0254-y>.

KATSAROU, Anastasia; GUDBJÖRNSDOTTIR, Soffia; RAWSHANI, Araz. *et al.* Type 1 diabetes mellitus. **Nature Reviews Disease Primers**, [S.L.], v. 3, n. 1, p. 1-10, 30 mar. 2017. Springer Science and Business Media LLC. <http://dx.doi.org/10.1038/nrdp.2017.16>.

ELAGIZI, Andrew; KACHUR, Sergey; CARBONE, Salvatore. *et al.* A Review of Obesity, Physical Activity, and Cardiovascular Disease. **Current Obesity Reports**, [S.L.], v. 9, n. 4, p. 571-581, 1 set. 2020. Springer Science and Business Media LLC. <http://dx.doi.org/10.1007/s13679-020-00403-z>.

KHAODHIAR, Lalita; MCCOWEN, Karen C.; BLACKBURN, George L.. Obesity and its comorbid conditions. **Clinical Cornerstone**, [S.L.], v. 2, n. 3, p. 17-31, jan. 1999. Elsevier BV. [http://dx.doi.org/10.1016/s1098-3597\(99\)90002-9](http://dx.doi.org/10.1016/s1098-3597(99)90002-9).

OLIVEIRA, Edson Mendes de; SANDRI, Silvana; KNEBEL, Franciele Hinterholz. *et al.* Hypoxia Increases Serum Amyloid A3 (SAA3) in Differentiated 3T3-L1 Adipocytes. **Inflammation**, [S.L.], v. 36, n. 5, p. 1107-1110, 20 abr. 2013. Springer Science and Business Media LLC. <http://dx.doi.org/10.1007/s10753-013-9644-9>.

BRATTI, Letícia de Oliveira Souza; CARMO, Ícaro Andrade Rodrigues do; VILELA, Taís Ferreira; SOUZA, Liliete Canes; MORAES, Ana Carolina Rabello de; FILIPPIN-MONTEIRO, Fabíola Branco. Bariatric surgery improves clinical outcomes and adiposity biomarkers but not inflammatory cytokines SAA and MCP-1 after a six-month follow-up. **Scandinavian Journal Of Clinical And Laboratory Investigation**, [S.L.], v. 81, n. 3, p. 230-236, 7 abr. 2021. Informa UK Limited. <http://dx.doi.org/10.1080/00365513.2021.1904278>.

FOTI, Daniela P.; BRUNETTI, Antonio. Editorial: ∴linking hypoxia to obesity∴. **Frontiers In Endocrinology**, [S.L.], v. 8, p. 1-15, 22 fev. 2017. Frontiers Media SA. <http://dx.doi.org/10.3389/fendo.2017.00034>.

THOMPSON, James; THOMPSON, Margaret; NUSSBAUM, Robert; MCINNES, Roderick; WILLARD, Huntington. Variação genética em indivíduos: mutação e polimorfismo. In: THOMPSON, James; THOMPSON, Margaret; NUSSBAUM, Robert; MCINNES, Roderick; WILLARD, Huntington. **Genética Médica**. 6. ed. Campo Grande: Guanabara Koogan, 2002. Cap. 6. p. 69-71.

ADZHUBEI, Ivan; SCHMIDT, Steffen; PESHKIN, Leonid *et al.* A method and server for predicting damaging missense mutations. **Nature Methods**, [N.I.], v. 7, n. [], p. 248-259, abr. 2010.

WATERHOUSE, Andrew; BERTONI, Martino; BIENERT, Stefan; STUDER, Gabriel; TAURIELLO, Gerardo; GUMIENNY, Rafal; HEER, Florian T; DE€ BEER, Tjaart A P; REMPFER, Christine; BORDOLI, Lorenza. SWISS-MODEL: homology modelling of protein structures and complexes. **Nucleic Acids Research**, [S.L.], v. 46, n. 1, p. 296-303, 21 maio 2018. Oxford University Press (OUP). <http://dx.doi.org/10.1093/nar/gky427>.

STUDER, Gabriel; TAURIELLO, Gerardo; BIENERT, Stefan; BIASINI, Marco; JOHNER, Niklaus; SCHWEDE, Torsten. ProMod3—A versatile homology modelling toolbox. **Plos Computational Biology**, [S.L.], v. 17, n. 1, p. 5-15, 28 jan. 2021. Public Library of Science (PLoS). <http://dx.doi.org/10.1371/journal.pcbi.1008667>.

XU, J.; JIAO, F.; BERGER, B.. A tree-decomposition approach to protein structure prediction. **2005 Ieee Computational Systems Bioinformatics Conference (Csb'05)**, [S.L.], v. 5, n. 7, p. 247-256, set. 2005. IEEE. <http://dx.doi.org/10.1109/csb.2005.9>.

KRIVOV, Georgii G.; SHAPOVALOV, Maxim V.; DUNBRACK, Roland L.. Improved prediction of protein side-chain conformations with SCWRL4. **Proteins: Structure, Function, and Bioinformatics**, [S.L.], v. 77, n. 4, p. 778-795, 29 maio 2009. Wiley. <http://dx.doi.org/10.1002/prot.22488>.

ROTIMI, Solomon O.; PETER, Olamide; OGUNTADE, Oluwayomi; ROTIMI, Oluwakemi A.. In silico analysis of the functional non-synonymous single nucleotide polymorphisms in the human CYP27B1 gene. **Egyptian Journal Of Medical Human Genetics**, [S.L.], v. 19, n. 4, p. 367-378, out. 2018. Springer Science and Business Media LLC. <http://dx.doi.org/10.1016/j.ejmhg.2018.03.001>.

XIAO, Q *et al.* Characterization of glucagon-like peptide-1 receptor-binding determinants. **Journal Of Molecular Endocrinology**, [S.L.], v. 25, n. 3, p. 321-335, 1 dez. 2000. Bioscientifica. <http://dx.doi.org/10.1677/jme.0.0250321>.

HOSSAIN, Md. Shahadat; ROY, Arpita Singha; ISLAM, Md. Sajedul. In silico analysis predicting effects of deleterious SNPs of human RASSF5 gene on its structure and

functions. **Scientific Reports**, [S.L.], v. 10, n. 1, p. 1-14, 3 set. 2020. Springer Science and Business Media LLC. <http://dx.doi.org/10.1038/s41598-020-71457-1>.

XIANG, Jiale; YANG, Jiyun; CHEN, Lisha; CHEN, Qiang; YANG, Haiyan; SUN, Chengcheng; ZHOU, Qing; PENG, Zhiyu. Reinterpretation of common pathogenic variants in ClinVar revealed a high proportion of downgrades. **Scientific Reports**, [S.L.], v. 10, n. 1, p. 1-5, 15 jan. 2020. Springer Science and Business Media LLC. <http://dx.doi.org/10.1038/s41598-019-57335-5>.

EMADI, Elaheh; AKHOUNDI, Fatemeh; KALANTAR, Seyed Mehdi; EMADI-BAYGI, Modjtaba. Predicting the most deleterious missense nsSNPs of the protein isoforms of the human HLA-G gene and in silico evaluation of their structural and functional consequences. **Bmc Genetics**, [S.L.], v. 21, n. 1, p. 1-27, 31 ago. 2020. Springer Science and Business Media LLC. <http://dx.doi.org/10.1186/s12863-020-00890-y>.

SHAH, Hania; KHAN, Khushbukhat; KHAN, Naila; BADSHAH, Yasmin; ASHRAF, Naeem Mahmood; SHABBIR, Maria. Impact of deleterious missense PRKCI variants on structural and functional dynamics of protein. **Scientific Reports**, [S.L.], v. 12, n. 1. 8 mar. 2022. Springer Science and Business Media LLC. <http://dx.doi.org/10.1038/s41598-022-07526-4>.

KHIMSURIYA, Yashvant M.; CHAUHAN, Jenabhai B.. Pathogenic predictions of non-synonymous variants and their impacts: a computational assessment of arhgef6 gene. **Egyptian Journal Of Medical Human Genetics**, [S.L.], v. 19, n. 4, p. 333-344, out. 2018. Springer Science and Business Media LLC. <http://dx.doi.org/10.1016/j.ejmhg.2018.05.002>.

MICHAŁOWSKA, Joanna *et al.* Association of GLP1R variants rs2268641 and rs6923761 with obesity and other metabolic parameters in a Polish cohort. **Frontiers In Endocrinology**, [S.L.], v. 13, p. 01-10, 19 out. 2022. Frontiers Media SA. <http://dx.doi.org/10.3389/fendo.2022.1000185>.

SATHANANTHAN, Airani *et al.* Common Genetic Variation in GLP1R and Insulin Secretion in Response to Exogenous GLP-1 in Nondiabetic Subjects. **Diabetes Care**, [S.L.], v. 33, n. 9, p. 2074-2076, 1 set. 2010. American Diabetes Association. <http://dx.doi.org/10.2337/dc10-0200>.

JAVORSKÝ, M. *et al.* A missense variant in GLP1R gene is associated with the glycaemic response to treatment with gliptins. **Diabetes, Obesity And Metabolism**, [S.L.], v. 18, n. 9, p. 941-944, 7 jun. 2016. Wiley. <http://dx.doi.org/10.1111/dom.12682>.

DAWED, Adem y *et al.* Pharmacogenomics of GLP-1 receptor agonists: a genome-wide analysis of observational data and large randomised controlled trials. **The Lancet Diabetes & Endocrinology**, [S.L.], v. 11, n. 1, p. 33-41, jan. 2023. Elsevier BV. [http://dx.doi.org/10.1016/s2213-8587\(22\)00340-0](http://dx.doi.org/10.1016/s2213-8587(22)00340-0).

HAN, Eugene *et al.* A genetic variant in GLP1R is associated with response to DPP-4 inhibitors in patients with type 2 diabetes. **Medicine**, [S.L.], v. 95, n. 44, p. 01-06, nov. 2016. Ovid Technologies (Wolters Kluwer Health). <http://dx.doi.org/10.1097/md.0000000000005155>.

EID, Liliane El *et al.* Biased agonism and polymorphic variation at the GLP-1 receptor: implications for the development of personalised therapeutics. **Pharmacological Research**, [S.L.], v. 184, p. 106411, out. 2022. Elsevier BV. <http://dx.doi.org/10.1016/j.phrs.2022.106411>.

LAGOU, Vasiliki *et al.* Random glucose GWAS in 493,036 individuals provides insights into diabetes pathophysiology, complications and treatment stratification. **Medrxiv**: The preprint server for health sciences, [S.L.], 20 abr. 2021. Cold Spring Harbor Laboratory. <http://dx.doi.org/10.1101/2021.04.17.21255471>.

DAI, Feihan F. *et al.* A Novel GLP1 Receptor Interacting Protein ATP6ap2 Regulates Insulin Secretion in Pancreatic Beta Cells. **Journal Of Biological Chemistry**, [S.L.], v. 290, n. 41, p. 25045-25061, out. 2015. Elsevier BV. <http://dx.doi.org/10.1074/jbc.m115.648592>.

SZKLARCZYK, Damian *et al.* STRING v11: protein-protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. **Nucleic Acids Research**, [S.L.], v. 47, n. 1, p. 607-613, 22 nov. 2018. Oxford University Press (OUP). <http://dx.doi.org/10.1093/nar/gky1131>.

SZKLARCZYK, Damian *et al.* The STRING database in 2023: protein-protein association networks and functional enrichment analyses for any sequenced genome of interest. **Nucleic Acids Research**, [S.L.], v. 51, n. 1, p. 638-646, 12 nov. 2022. Oxford University Press (OUP). <http://dx.doi.org/10.1093/nar/gkac1000>.

ZHANG, Qiangfeng Cliff *et al.* PrePPI: a structure-informed database of protein-protein interactions. **Nucleic Acids Research**, [S.L.], v. 41, n. 1, p. 828-833, 26 nov. 2012. Oxford University Press (OUP). <http://dx.doi.org/10.1093/nar/gks1231>.

MCDOWALL, M. D. *et al.* PIPs: human protein-protein interaction prediction database. **Nucleic Acids Research**, [S.L.], v. 37, n. , p. 651-656, 1 jan. 2009. Oxford University Press (OUP). <http://dx.doi.org/10.1093/nar/gkn870>.

MICHAŁOWSKA, Joanna *et al.* Association of GLP1R variants rs2268641 and rs6923761 with obesity and other metabolic parameters in a Polish cohort. **Frontiers In Endocrinology**, [S.L.], v. 13, n. 1, p. 1-10, 19 out. 2022. Frontiers Media SA. <http://dx.doi.org/10.3389/fendo.2022.1000185>.

GAO, Wenwen *et al.* Human GLP1R variants affecting GLP1R cell surface expression contribute to impaired glucose control and increased adiposity. **Research Square**, [S.L.], 7 nov. 2022. Research Square Platform LLC. <http://dx.doi.org/10.21203/rs.3.rs-2114208/v1>.

APÊNDICE – Tabelas contendo conteúdo relacionado à obtenção das variantes patogênicas.

Tabelas formuladas ao longo do desenvolvimento do trabalho contendo informações sobre demais variantes missense encontradas na plataforma dbSNP e também com os resultados de análises em outras bases de dados com finalidade de confirmação de resultados.

Mutation ID	Amn. Exchange	Results	Total
rs2295006	Arg44His	Benign	186
rs1042044	Leu260Phe	Benign	
rs1126475	Ala316Gly	Benign	
rs3765467	Arg131Gln	Benign	
rs3765467	Arg131Pro	Benign	
rs6918287	Glu133Asp	Benign	
rs6923761	Gly168Ser	Benign	
rs6923761	Gly168Arg	Benign	
rs10305420	Pro7Leu	Benign	
rs10305492	Ala316Thr	Benign	
rs10305510	Arg421Gln	Benign	
rs138405478	Lys434Thr	Benign	
rs138965775	Ala209Thr	Benign	
rs138965775	Ala209Pro	Benign	
rs138965775	Ala209Ser	Benign	
rs144690684	Val249Met	Benign	
rs145832308	Ser432Arg	Benign	
rs145955078	Phe80Leu	Benign	
rs146340667	Val194Ile	Benign	
rs146868158	Arg421Gln	Benign	
rs147627784	Gly132Arg	Benign	
rs148862435	Gly450Ala	Benign	
rs150253529	Arg48His	Benign	
rs185053350	Ser261Pro	Benign	
rs185053350	Ser261Ala	Benign	
rs189969666	Ala449Glu	Benign	
rs189969666	Ala449Val	Benign	
rs199523424	Val270Met	Benign	
rs199526561	Tyr101His	Benign	
rs199641766	Val365Ile	Benign	
rs199790397	Ile147Met	Benign	
rs199796313	Arg414Trp	Benign	
rs199825566	Val160Ile	Benign	
rs199825566	Val160Phe	Benign	
rs199959395	Leu339Phe	Benign	
rs200053872	Leu118Val	Benign	

rs200551011	Pro137Leu	Benign
rs200691429	Arg419His	Benign
rs200721844	Pro5Leu	Benign
rs200841153	Ser441Ile	Benign
rs200843058	Ala238Ser	Benign
rs200849157	Arg40Gln	Benign
rs200890494	Cys341Ser	Benign
rs201001424	Ile427Asn	Benign
rs201223341	Thr440Ala	Benign
rs201338954	Asp53Ala	Benign
rs201338954	Asp53Gly	Benign
rs201338954	Asp53Val	Benign
rs201355669	Pro96Leu	Benign
rs201431571	Val259Ile	Benign
rs201672448	Ser445Thr	Benign
rs201675748	Arg44Ser	Benign
rs367903915	Arg43Gln	Benign
rs368940624	Arg414Gln	Benign
rs371228307	Leu141Phe	Benign
rs375821449	Trp33Arg	Benign
rs529033290	Thr298Ala	Benign
rs546151635	Met17Thr	Benign
rs554064290	Met453Thr	Benign
rs746502272	Ile161Phe	Benign
rs748354939	Gly6Val	Benign
rs748354939	Gly6Asp	Benign
rs748771635	Leu425Ser	Benign
rs748825486	Ala199Thr	Benign
rs749307293	Ser225Gly	Benign
rs749764829	Ser258Leu	Benign
rs750122674	Pro56Leu	Benign
rs754900903	Ser84Asn	Benign
rs757507531	Glu418Asp	Benign
rs758868498	Gly132Ala	Benign
rs758868498	Gly132Val	Benign
rs759060644	Ala164Val	Benign
rs760617935	Ala238Gly	Benign
rs760617935	Ala238Val	Benign
rs761302689	Arg429Lys	Benign
rs761383264	Ile286Val	Benign
rs761383378	Lys130Thr	Benign
rs763589430	Leu443Pro	Benign
rs763952154	Glu127Lys	Benign
rs768143333	Ile147Val	Benign
rs769668902	Ser124Leu	Benign

rs770806756	Trp264Arg	Benign
rs771765004	Asp430Glu	Benign
rs772199865	Leu232Ile	Benign
rs772199865	Leu232Val	Benign
rs772199865	Leu232Phe	Benign
rs773168576	Thr391Ala	Benign
rs773657617	Gln26Lys	Benign
rs773786089	Gly285Ser	Benign
rs774790092	Leu167Phe	Benign
rs776194642	Glu139Asp	Benign
rs776311029	Arg267Thr	Benign
rs776311029	Arg267Met	Benign
rs780507821	Phe257Ile	Benign
rs780507821	Phe257Val	Benign
rs903979009	Glu107Lys	Benign
rs905725891	Thr455Ala	Benign
rs907706489	Gln140His	Benign
rs936741749	Leu109Phe	Benign
rs937676680	Arg24Pro	Benign
rs961473826	Thr448Met	Benign
rs972506083	Val229Ala	Benign
rs976295207	Gly22Ser	Benign
rs985149155	Phe413Ile	Benign
rs985294169	Met397Leu	Benign
rs1042461152	Ile196Val	Benign
rs1156801898	Ala2Val	Benign
rs1158382194	His426Pro	Benign
rs1158382194	His426Arg	Benign
rs1159561886	Glu107Asp	Benign
rs1160861834	Ala153Thr	Benign
rs1183091621	Val332Ile	Benign
rs1183821389	Gly295Ser	Benign
rs1193326947	Thr207Ile	Benign
rs1196193414	Ala92Val	Benign
rs1199763316	Gly108Ala	Benign
rs1228779181	Asp222Glu	Benign
rs1230482807	Ser163Phe	Benign
rs1231292755	Ala92Thr	Benign
rs1231379442	Val405Ile	Benign
rs1231876915	Glu138Gly	Benign
rs1235863894	Ala4Asp	Benign
rs1238963152	Ser93Asp	Benign
rs1245724616	Ala28Val	Benign
rs1249883851	Ala447Gly	Benign
rs1253028244	Gly27Ala	Benign

rs1253028244	Gly27Val	Benign
rs1279957419	Leu118Gln	Benign
rs1279957419	Leu118Pro	Benign
rs1289964395	Glu387Asp	Benign
rs1290269769	Gln45Arg	Benign
rs1291437179	Arg134Thr	Benign
rs1307781425	Ile286Thr	Benign
rs1317906920	Asp53Asn	Benign
rs1317906920	Asp53Tyr	Benign
rs1335072951	Ser124Thr	Benign
rs1344505759	Asp430Asn	Benign
rs1344505759	Asp430His	Benign
rs1348279169	Pro96Ser	Benign
rs1366151858	Leu278Val	Benign
rs1372183942	Glu133Gln	Benign
rs1375037526	Ser463Gly	Benign
rs1390194943	Ser49Cys	Benign
rs1391602479	Ser431Gly	Benign
rs1394576392	Lys437Gln	Benign
rs1410795040	Val95Glu	Benign
rs1412444748	Gln211Arg	Benign
rs1415707619	Ser79Leu	Benign
rs1421643787	Glu34Lys	Benign
rs1436677729	Pro5Ala	Benign
rs1454776777	Thr39Ser	Benign
rs1455995986	Ser441Arg	Benign
rs1456463113	Thr58Ile	Benign
rs1462886581	Ile146Leu	Benign
rs1472091574	Gly6Ser	Benign
rs1473730004	Gly216Trp	Benign
rs1474396123	Glu127Asp	Benign
rs1562009540	Ala158Thr	Benign
rs1562021040	Ser452Thr	Benign
rs1583627938	Asn115Thr	Benign
rs1583628661	Glu139Lys	Benign
rs1768021501	Ala2Thr	Benign
rs1768021603	Gly3Ser	Benign
rs1768021603	Gly3Arg	Benign
rs1768022624	Ala11Glu	Benign
rs1768024088	Pro25Thr	Benign
rs1768024231	Gln26Arg	Benign
rs1768216877	Ser49Pro	Benign
rs1768218116	Thr58Pro	Benign
rs1768241801	Phe61Tyr	Benign
rs1768485089	His99Tyr	Benign

rs1768488422	Asp122Asn	Benign
rs1768489629	Ser129Pro	Benign
rs1768510179	Leu154Phe	Benign
rs1768510774	Ser157Phe	Benign
rs1768715788	Ala199Val	Benign
rs1768715958	Leu201Val	Benign
rs1768716193	Ser206Arg	Benign
rs1768733941	Tyr252His	Benign
rs1768737563	Ile272Val	Benign
rs1768889776	Val282Ile	Benign
rs1768920331	Leu322Phe	Benign
rs1768940618	Phe381Leu	Benign
rs1769132939	Glu418Lys	Benign
rs1769136580	Gly450Ser	Benign
rs1769137301	Ala460Thr	Benign
rs3765467	Arg131Leu	Poss. Damaging
rs10305421	Arg20Lys	Poss. Damaging
rs61733062	Arg376Gln	Poss. Damaging
rs61733062	Arg376Leu	Poss. Damaging
rs145619754	Leu268Phe	Poss. Damaging
rs147207725	Ile330Val	Poss. Damaging
rs148158586	Thr35Met	Poss. Damaging
rs150729240	Asp293Tyr	Poss. Damaging
rs200423130	His171Gln	Poss. Damaging
rs200437436	Phe385Leu	Poss. Damaging
rs201672448	Ser445Asn	Poss. Damaging
rs202103886	Asn338Ser	Poss. Damaging
rs202152738	Glu294Lys	Poss. Damaging
rs527991362	Ala375Thr	Poss. Damaging
rs747487240	Lys342Gln	Poss. Damaging
rs751670297	Pro54Ser	Poss. Damaging
rs753705639	Val81Leu	Poss. Damaging
rs759948924	Gln45His	Poss. Damaging
rs760402748	Tyr289Phe	Poss. Damaging
rs762655612	Gly22Asp	Poss. Damaging
rs772533032	Asn338Lys	Poss. Damaging
rs937676680	Arg24His	Poss. Damaging
rs1019126528	Val332Ala	Poss. Damaging
rs1242484640	Leu314Phe	Poss. Damaging
rs1249883851	Ala447Val	Poss. Damaging
rs1324858877	Cys347Arg	Poss. Damaging
rs1343256236	Ala399Ser	Poss. Damaging
rs1375235181	Gly248Asp	Poss. Damaging
rs1454776777	Thr391Ile	Poss. Damaging
rs1768023220	Gly19Arg	Poss. Damaging

rs1768215403	Val36Ala	Poss. Damaging	
rs1768915205	Met303Thr	Poss. Damaging	
rs1768941018	Glu387Lys	Poss. Damaging	
rs1769134274	Met433Ile	Poss. Damaging	
rs1126475	Ala316Asp	Prob. Damaging	164
rs2235868	Arg176Gly	Prob. Damaging	
rs2235868	Arg176Trp	Prob. Damaging	
rs10305493	Ser333Cys	Prob. Damaging	
rs10305510	Arg421Pro	Prob. Damaging	
rs140642887	Ala239Thr	Prob. Damaging	
rs147627784	Gly132Trp	Prob. Damaging	
rs148543734	Gly361Arg	Prob. Damaging	
rs149578908	Arg310Gln	Prob. Damaging	
rs182447758	Arg227His	Prob. Damaging	
rs190133144	His374Asn	Prob. Damaging	
rs190133144	His374Tyr	Prob. Damaging	
rs199745579	Asp198Glu	Prob. Damaging	
rs199783730	Thr353Met	Prob. Damaging	
rs199818129	Pro358Leu	Prob. Damaging	
rs200065197	Arg376Trp	Prob. Damaging	
rs200118342	Arg326Gly	Prob. Damaging	
rs200118342	Arg326Trp	Prob. Damaging	
rs200765138	Arg64Trp	Prob. Damaging	
rs200792917	Tyr291Cys	Prob. Damaging	
rs201054657	Ile357Thr	Prob. Damaging	
rs201231115	Cys46Tyr	Prob. Damaging	
rs201235247	Val331Met	Prob. Damaging	
rs201498303	Val100Met	Prob. Damaging	
rs201617183	Gly377Arg	Prob. Damaging	
rs201634613	Trp87Arg	Prob. Damaging	
rs201669667	Leu360Pro	Prob. Damaging	
rs201675748	Arg44Cys	Prob. Damaging	
rs201899163	Gln394Arg	Prob. Damaging	
rs202075228	Arg64Gln	Prob. Damaging	
rs202171972	Asp344Glu	Prob. Damaging	
rs202222263	Arg170Ser	Prob. Damaging	
rs367543060	Thr149Met	Prob. Damaging	
rs375865648	Ile272Thr	Prob. Damaging	
rs745320716	Ile272Met	Prob. Damaging	
rs745360395	His173Pro	Prob. Damaging	
rs745360395	His173Arg	Prob. Damaging	
rs746598122	Gly78Ser	Prob. Damaging	
rs747086141	Tyr101Cys	Prob. Damaging	
rs748911530	Thr175Asn	Prob. Damaging	
rs748911530	Thr175Ile	Prob. Damaging	

rs749386030	Arg380Cys	Prob. Damaging
rs750772227	Leu254Met	Prob. Damaging
rs750782093	Glu408Gly	Prob. Damaging
rs751693672	Trp417Leu	Prob. Damaging
rs752432456	Asn320Tyr	Prob. Damaging
rs753705639	Val81Met	Prob. Damaging
rs754117921	Thr355Ile	Prob. Damaging
rs754315702	Ser219Phe	Prob. Damaging
rs754650010	Phe367Leu	Prob. Damaging
rs755813574	Ser333Thr	Prob. Damaging
rs756154483	Val194Ala	Prob. Damaging
rs759380877	Trp120Arg	Prob. Damaging
rs760402748	Tyr289Cys	Prob. Damaging
rs762048338	Leu231Pro	Prob. Damaging
rs763793276	Leu401Ile	Prob. Damaging
rs764381278	Arg190Gln	Prob. Damaging
rs764752566	Leu244Phe	Prob. Damaging
rs764763043	Glu292Val	Prob. Damaging
rs765821452	Pro435Leu	Prob. Damaging
rs765868899	His363Tyr	Prob. Damaging
rs766288520	Phe393Leu	Prob. Damaging
rs767188992	Val287Asp	Prob. Damaging
rs767632966	Ala70Thr	Prob. Damaging
rs768340419	His171Tyr	Prob. Damaging
rs768822503	Arg227Cys	Prob. Damaging
rs770196509	Arg48Cys	Prob. Damaging
rs771324929	Arg176Lys	Prob. Damaging
rs772055021	Asn302Tyr	Prob. Damaging
rs774182779	Ile357Phe	Prob. Damaging
rs774357734	Asp67Asn	Prob. Damaging
rs776464173	Ile317Asn	Prob. Damaging
rs776464173	Ile317Thr	Prob. Damaging
rs777092905	Arg310Trp	Prob. Damaging
rs777241387	His180Tyr	Prob. Damaging
rs777494808	Arg9His	Prob. Damaging
rs778024660	Trp417Gly	Prob. Damaging
rs866312541	Ser31Phe	Prob. Damaging
rs868577885	Pro439Leu	Prob. Damaging
rs955304580	Leu396Pro	Prob. Damaging
rs1034202463	Ala162Thr	Prob. Damaging
rs1040512515	Gln394His	Prob. Damaging
rs1160797040	Pro277Thr	Prob. Damaging
rs1164414663	Glu373Lys	Prob. Damaging
rs1174675097	Leu172Arg	Prob. Damaging
rs1178460976	Cys403Phe	Prob. Damaging

rs1183821389	Gly295Cys	Prob. Damaging
rs1188229383	Leu60Ser	Prob. Damaging
rs1189829779	Arg419Cys	Prob. Damaging
rs1191786225	Pro23Ser	Prob. Damaging
rs1191933005	Tyr88His	Prob. Damaging
rs1191933005	Tyr88Asp	Prob. Damaging
rs1201653746	Arg102Trp	Prob. Damaging
rs1206180510	Leu181Pro	Prob. Damaging
rs1225236582	Ile330Asn	Prob. Damaging
rs1226111511	Val327Ala	Prob. Damaging
rs1227265756	Val398Met	Prob. Damaging
rs1231745755	Arg102Gln	Prob. Damaging
rs1232817319	Asp372Asn	Prob. Damaging
rs1234543862	Leu32Pro	Prob. Damaging
rs1236873018	Ser442Asn	Prob. Damaging
rs1249582122	Arg170Gly	Prob. Damaging
rs1268052644	Tyr148Asn	Prob. Damaging
rs1268052644	Tyr148His	Prob. Damaging
rs1268234905	Asp198Gly	Prob. Damaging
rs1290185930	Phe404Leu	Prob. Damaging
rs1301081484	Ser193Phe	Prob. Damaging
rs1304693608	Cys296Tyr	Prob. Damaging
rs1306846366	Leu50Pro	Prob. Damaging
rs1317393764	Pro358Ala	Prob. Damaging
rs1329654429	Val370Ala	Prob. Damaging
rs1329654429	Val370Gly	Prob. Damaging
rs1339732764	Val409Asp	Prob. Damaging
rs1339732764	Val409Gly	Prob. Damaging
rs1352145872	Arg326Gln	Prob. Damaging
rs1361251401	Ala399Gly	Prob. Damaging
rs1363680514	Leu12Gln	Prob. Damaging
rs1364791940	Ser206Cys	Prob. Damaging
rs1369172279	Asp344Tyr	Prob. Damaging
rs1388310962	Gly248Asp	Prob. Damaging
rs1392727369	Leu172Met	Prob. Damaging
rs1395888713	Leu8Met	Prob. Damaging
rs1402525983	Gly275Ser	Prob. Damaging
rs1440906113	Gln459Arg	Prob. Damaging
rs1441241286	Ala21Val	Prob. Damaging
rs1462258571	Leu349Arg	Prob. Damaging
rs1463954816	Lys415Asn	Prob. Damaging
rs1464933937	Ile165Thr	Prob. Damaging
rs1483005423	Gln234Arg	Prob. Damaging
rs1484122515	Ser186Ala	Prob. Damaging
rs1484789101	Leu159Gln	Prob. Damaging

rs1485302982	Trp87Ser	Prob. Damaging
rs1562017482	Arg348Lys	Prob. Damaging
rs1562017784	Ile366Asn	Prob. Damaging
rs1583616783	Tyr42Ser	Prob. Damaging
rs1583637698	Asp198Tyr	Prob. Damaging
rs1583648155	Ser352Pro	Prob. Damaging
rs1753959402	Asp240Ser	Prob. Damaging
rs1768022466	Arg9His	Prob. Damaging
rs1768241869	Cys62Tyr	Prob. Damaging
rs1768242404	Tyr69His	Prob. Damaging
rs1768243737	Pro86Leu	Prob. Damaging
rs1768510278	Leu154Pro	Prob. Damaging
rs1768510670	Phe156Ser	Prob. Damaging
rs1768714044	Ser186Phe	Prob. Damaging
rs1768714631	Ala193Val	Prob. Damaging
rs1768717151	Tyr220Cys	Prob. Damaging
rs1768733080	Trp243Arg	Prob. Damaging
rs1768733454	Gly248Ser	Prob. Damaging
rs1768733795	Leu251Pro	Prob. Damaging
rs1768734240	Leu255Val	Prob. Damaging
rs1768736815	Tyr269His	Prob. Damaging
rs1768736919	Tyr269Cys	Prob. Damaging
rs1768889185	Pro277His	Prob. Damaging
rs1768915135	Met303Val	Prob. Damaging
rs1768920170	Asn320Ser	Prob. Damaging
rs1768920170	Asn320Ile	Prob. Damaging
rs1768920733	Cys329Ser	Prob. Damaging
rs1768938541	Ala368Val	Prob. Damaging
rs1768982961	Gly395Arg	Prob. Damaging
rs1768984322	Cys403Gly	Prob. Damaging
rs1769136760	Ser452Cys	Prob. Damaging
rs1769137045	Ala456Asp	Prob. Damaging
rs1769137045	Ala456Val	Prob. Damaging

Material complementar 1 – Lista completa de mutações encontradas na plataforma dbSNP.

A:

rs1126475	Ala316Asp
rs2235868	Arg176Gly
rs2235868	Arg176Trp
rs10305493	Ser333Cys
rs10305510	Arg421Pro
rs140642887	Ala239Thr
rs147627784	Gly132Trp
rs148543734	Gly361Arg
rs149578908	Arg310Gln
rs182447758	Arg227His

rs190133144	His374Asn
rs190133144	His374Tyr
rs199745579	Asp198Glu
rs199783730	Thr353Met
rs199818129	Pro358Leu
rs200065197	Arg376Trp
rs200118342	Arg326Gly
rs200118342	Arg326Trp
rs200765138	Arg64Trp
rs200792917	Tyr291Cys
rs201054657	Ile357Thr
rs201231115	Cys46Tyr
rs201235247	Val331Met
rs201498303	Val100Met
rs201617183	Gly377Arg
rs201634613	Trp87Arg
rs201669667	Leu360Pro
rs201675748	Arg44Cys
rs201899163	Gln394Arg
rs202075228	Arg64Gln
rs202171972	Asp344Glu
rs202222263	Arg170Ser
rs367543060	Thr149Met
rs375865648	Ile272Thr
rs745320716	Ile272Met
rs745360395	His173Pro
rs745360395	His173Arg
rs746598122	Gly78Ser
rs747086141	Tyr101Cys
rs748911530	Thr175Asn
rs748911530	Thr175Ile
rs749386030	Arg380Cys
rs750772227	Leu254Met
rs750782093	Glu408Gly
rs751693672	Trp417Leu
rs752432456	Asn320Tyr
rs753705639	Val81Met
rs754117921	Thr355Ile
rs754315702	Ser219Phe
rs754650010	Phe367Leu
rs755813574	Ser333Thr
rs756154483	Val194Ala
rs759380877	Trp120Arg
rs760402748	Tyr289Cys
rs762048338	Leu231Pro
rs763793276	Leu401Ile

rs764381278	Arg190Gln
rs764752566	Leu244Phe
rs764763043	Glu292Val
rs765821452	Pro435Leu
rs765868899	His363Tyr
rs766288520	Phe393Leu
rs767188992	Val287Asp
rs767632966	Ala70Thr
rs768340419	His171Tyr
rs768822503	Arg227Cys
rs770196509	Arg48Cys
rs771324929	Arg176Lys
rs772055021	Asn302Tyr
rs774182779	Ile357Phe
rs774357734	Asp67Asn
rs776464173	Ile317Asn
rs776464173	Ile317Thr
rs777092905	Arg310Trp
rs777241387	His180Tyr
rs777494808	Arg9His
rs778024660	Trp417Gly
rs866312541	Ser31Phe
rs868577885	Pro439Leu
rs955304580	Leu396Pro
rs1034202463	Ala162Thr
rs1040512515	Gln394His
rs1160797040	Pro277Thr
rs1164414663	Glu373Lys
rs1174675097	Leu172Arg
rs1178460976	Cys403Phe
rs1183821389	Gly295Cys
rs1188229383	Leu60Ser
rs1189829779	Arg419Cys
rs1191786225	Pro23Ser
rs1191933005	Tyr88His
rs1191933005	Tyr88Asp
rs1201653746	Arg102Trp
rs1206180510	Leu181Pro
rs1225236582	Ile330Asn
rs1226111511	Val327Ala
rs1227265756	Val398Met
rs1231745755	Arg102Gln
rs1232817319	Asp372Asn
rs1234543862	Leu32Pro
rs1236873018	Ser442Asn

rs1249582122	Arg170Gly
rs1268052644	Tyr148Asn
rs1268052644	Tyr148His
rs1268234905	Asp198Gly
rs1290185930	Phe404Leu
rs1301081484	Ser193Phe
rs1304693608	Cys296Tyr
rs1306846366	Leu50Pro
rs1317393764	Pro358Ala
rs1329654429	Val370Ala
rs1329654429	Val370Gly
rs1339732764	Val409Asp
rs1339732764	Val409Gly
rs1352145872	Arg326Gln
rs1361251401	Ala399Gly
rs1363680514	Leu12Gln
rs1364791940	Ser206Cys
rs1369172279	Asp344Tyr
rs1388310962	Gly248Asp
rs1392727369	Leu172Met
rs1395888713	Leu8Met
rs1402525983	Gly275Ser
rs1440906113	Gln459Arg
rs1441241286	Ala21Val
rs1462258571	Leu349Arg
rs1463954816	Lys415Asn
rs1464933937	Ile165Thr
rs1483005423	Gln234Arg
rs1484122515	Ser186Ala
rs1484789101	Leu159Gln
rs1485302982	Trp87Ser
rs1562017482	Arg348Lys
rs1562017784	Ile366Asn
rs1583616783	Tyr42Ser
rs1583637698	Asp198Tyr
rs1583648155	Ser352Pro
rs1753959402	Asn240Ser
rs1768022466	Arg9His
rs1768241869	Cys62Tyr
rs1768242404	Tyr69His
rs1768243737	Pro86Leu
rs1768510278	Leu154Pro
rs1768510670	Phe156Ser
rs1768714044	Ser186Phe
rs1768714631	Ser193Pro

rs1768717151	Tyr220Cys
rs1768733080	Trp243Arg
rs1768733454	Gly248Ser
rs1768733795	Leu251Pro
rs1768734240	Leu255Val
rs1768736815	Tyr269His
rs1768736919	Tyr269Cys
rs1768889185	Pro277His
rs1768915135	Met303Val
rs1768920170	Asn320Ser
rs1768920170	Asn320Ile
rs1768920733	Cys329Ser
rs1768938541	Ala368Val
rs1768982961	Gly395Arg
rs1768984322	Cys403Gly
rs1769136760	Ser452Cys
rs1769137045	Ala456Asp
rs1769137045	Ala456Val

B:

Mutações		Align GVD	SNAP2	PANTHER	PROVEAN	
rs10305510	Arg421Pro	C65	Effect	Possibly damaging	-3.709	Deleterious
rs140642887	Ala239Thr	C35	Neutral	Probably damaging	-3.685	Deleterious
rs149578908	Arg310Gln	C35	Effect	Probably damaging	-3.859	Deleterious
rs182447758	Arg227His	C25	Effect	Probably damaging	-4.495	Deleterious
rs202171972	Asp344Glu	C35	Effect	Possibly damaging	-4.000	Deleterious
rs2235868	Arg176Gly	C65	Effect	Probably damaging	-6.685	Deleterious

Align GVD	
Class C65	Mais provável de alterar a função
Class C55	
Class C45	
Class C35	
Class C25	
Class C15	
Class C0	Menos provável de alterar a função

Material complementar 2: (A) Mutações provavelmente danosas encontradas na plataforma dbSNP com ponto de corte 1.0 aplicado na plataforma PolyPhen 2. Em vermelho: mutações abaixo de 1.0; Em verde: Mutações dentro de 1.0; Em amarelo: Mutações filtradas para análise nas demais plataformas. (B) Tabela contendo as análises feitas das seis mutações filtradas em outras plataformas de predição de dano juntamente com legenda explicativa dos resultados encontrados na plataforma Align GVD.

Mutations		Global Allele Frequency	Latino/Admixed American	African/African American	East Asian	South Asian
rs10305510	Arg421Pro	0.0002024	0.000	0.000	0.000	0.000
rs140642887	Ala239Thr	0.0001627	0.0003953	0.001123	0.000	0.000
rs149578908	Arg310Gln	0.000007954	0.000	0.000	0.000	0.00006533
rs182447758	Arg227His	0.00001596	0.00002894	0.000	0.0001087	0.00003266
rs202171972	Asp344Glu	0.0002625	0.000	0.000	0.00005437	0.002123
rs2235868	Arg176Gly	-	-	-	-	-

Mutations		European (Finnish)	European (non-Finnish)	Other	XX	XY
rs10305510	Arg421Pro	0.000	0.000	0.000	0.00004369	0.00001482
rs140642887	Ala239Thr	0.000	0.00003099	0.000	0.0001468	0.0001761
rs149578908	Arg310Gln	0.000	0.000	0.000	0.000	0.00001472
rs182447758	Arg227His	0.000	0.000	0.000	0.000008688	0.00002214
rs202171972	Asp344Glu	0.000	0.000	0.000	0.0001817	0.0003312
rs2235868	Arg176Gly	-	-	-	-	-

Material complementar 3: Tabelas formuladas com dados da plataforma GnomAD contendo as frequências alélicas global, por região/etnia e por sexo das seis mutações filtradas.