




## IDENTIFICAÇÃO GENÔMICA DE VARIANTES NÃO SINÔNIMAS DO GENE IL1A ASSOCIADAS A DOENÇAS HUMANAS POR ANÁLISE DE BIOINFORMÁTICA

 <https://doi.org/10.56238/levv15n41-105>

Data de submissão: 29/09/2024

Data de publicação: 29/10/2024

**Arthur Felipe Ferreira de Freitas**

Graduando em Ciências Biológicas  
Universidade Federal Rural de Pernambuco  
E-mail: arthur.ffreitas@ufrpe.br

**Nara Suzy Aguiar de Freitas**

Doutora em Genética pela Universidade Federal de Pernambuco  
Universidade Federal Rural de Pernambuco  
E-mail: nara.safreitas@ufrpe.br

**Maria Helena Queiroz de Araújo Mariano**

Doutora em Ciências Biológicas pela Universidade Federal de Pernambuco  
Universidade de Pernambuco  
E-mail: irpe.diretoria@gmail.com

**Eliézer Rushansky**

Pós-graduação em Clínica Médica pela Universidade de Pernambuco  
Universidade de Pernambuco  
E-mail: eliezer.rushansky@upe.pe.gov.br

**Maria de Mascena Diniz Maia**

Doutora em Ciências Biológicas pela Universidade Federal de Pernambuco  
Universidade Federal Rural de Pernambuco  
E-mail: maria.dmaia@ufrpe.br

### RESUMO

A interleucina-1 alfa é uma citocina que se destaca por seu papel pró-inflamatório essencial que atua ativando células-chave para combate às infecções. Algumas mutações genéticas podem promover a substituição de aminoácidos nesta proteína, levando a consequências que podem estar associadas a doenças. Assim, esse trabalho buscou identificar Polimorfismos de Nucleotídeo Único (SNPs) de alto risco à saúde humana e prever seus efeitos morfofuncionais na proteína. As variantes foram localizadas no transcrito IL1A-201, disponível no banco de dados genômico ENSEMBL. Após a seleção das variantes missense, utilizamos os programas SIFT, PolyPhen-2, MetaLR e Mutation Assessor para prever os impactos das substituições dos aminoácidos na função e estrutura da proteína. Nossa análise revelou a presença da variante rs1190431689, que se trata da mudança de uma Alanina por uma Valina, a rs997926639, que corresponde a mudança de uma Prolina por uma Serina, e a variante rs1681201328, que representa a substituição de um Ácido Glutâmico por uma Glicina. Todos estes polimorfismos apresentaram resultado Deletério (SIFT), Provavelmente Prejudicial (PolyPhen-2) e Prejudicial (MetaLR) nos softwares de predição, o que indica que estas são variantes



que podem ser de alto risco para a saúde humana e podem estar associadas ao desenvolvimento de diversas patologias.

**Palavras-chave:** Citocina. Polimorfismos Genéticos. Predição in Sílico.

## 1 INTRODUÇÃO

O sistema imunológico humano é composto por uma rede complexa de células e moléculas que atuam sinergicamente para proteger o organismo contra invasores exógenos. Dentro desse contexto, o sistema imunológico inato, representado por componentes como os neutrófilos e células epiteliais de barreira, desempenha um papel fundamental como a primeira linha de defesa contra patógenos (Heeb *et al.*, 2020). No entanto, a resposta imunológica não se limita apenas à imunidade inata, mas é complementada pelo sistema imunológico adaptativo, que garante uma resposta mais específica e de memória contra ameaças recorrentes (Heeb *et al.*, 2020).

Nesta perspectiva, as citocinas, que são moléculas sinalizadoras cruciais no sistema imunológico, atuam como mediadores intercelulares que regulam a inflamação, a resposta imune e diversos processos fisiológicos e patológicos (Varella, 2001). Essas proteínas, produzidas por uma variedade de células, como macrófagos, células T, e células epiteliais, desempenham um papel central na comunicação entre as células do sistema imune e na coordenação de respostas imunes tanto inatas quanto adaptativas (Freitas *et al.*, 2022).

Dentro da classe das citocinas, a interleucina-1 alfa (IL-1 $\alpha$ ) se destaca por seu papel como uma citocina pró-inflamatória essencial. A IL-1 $\alpha$ , juntamente com a IL-1 beta (IL-1 $\beta$ ), forma parte da família IL-1, que inclui também a proteína antagonista IL-1ra (Moreira *et al.*, 2006). Essas citocinas são críticas na mediação de respostas inflamatórias, sendo produzidas principalmente por macrófagos, células epiteliais e outros tipos celulares em resposta a sinais de estresse ou infecção (Graves *et al.*, 1990; Moreira *et al.*, 2006). A IL-1 $\alpha$ , especificamente, é uma proteína que age tanto de forma autócrina, ou seja, agindo sobre as próprias células que a produzem, quanto parácrina, agindo sobre as células vizinhas às que a produzem, facilitando a sinalização celular em processos inflamatórios e participando na ativação de linfócitos e neutrófilos, células-chave no combate às infecções (Becerra *et al.*, 2023).

Do ponto de vista molecular, a IL-1 $\alpha$  é codificada pelo gene IL1A, localizado no cromossomo 2 humano, e sintetiza uma proteína constituída por uma cadeia de 271 aminoácidos, apresentando uma estrutura altamente conservada evolutivamente (Moreira *et al.*, 2006). Estudos estruturais, como os realizados por Graves *et al.* (1990), revelam que a IL-1 $\alpha$  é formada por um núcleo de folhas beta antiparalelas organizadas em um barril beta, uma característica estrutural essencial para sua estabilidade e interação com receptores celulares. Essa organização molecular complexa permite que a IL-1 $\alpha$  se ligue a receptores específicos na superfície das células-alvo, desencadeando uma cascata de sinalização que amplifica a resposta inflamatória.

A IL-1 $\alpha$  não só desempenha um papel vital na ativação e regulação de respostas imunológicas, como também está associada a diversas condições patológicas, como em doenças autoimunes e inflamações crônicas (Kubaski *et al.*, 2013). Por isso, mutações em genes que codificam proteínas

imunológicas, como IL1A, é um campo de investigação genômica que tem revelado a correlação positiva entre variações genéticas e o desenvolvimento de doenças (Kubaski *et al.*, 2013).

A análise de polimorfismos de nucleotídeo único (*SNPs*) é fundamental para a compreensão das consequências moleculares das substituições de aminoácidos e suas implicações clínicas. Neste contexto, o presente trabalho visou realizar uma análise abrangente das variantes genéticas do gene IL1A disponíveis em bancos de dados biológicos, com o objetivo de identificar e prever seus efeitos morfofuncionais na proteína.

## 2 MATERIAL E MÉTODOS

### 2.1 CONJUNTO DE DADOS

Os dados genéticos foram coletados a partir da plataforma *Ensembl*, um repositório de acesso público que integra dados genômicos de referência (Harrison *et al.*, 2024). Foi selecionado o transcrito IL1A-201 (ENST00000263339.4) como base para a análise, o que permitiu obter informações sobre a sequência de nucleotídeos, a estrutura proteica e os domínios da interleucina-1 alfa (IL-1 $\alpha$ ). Além disso, utilizamos o banco de dados para examinar os Polimorfismos de Nucleotídeo Único (*SNPs*) associados ao gene *IL1A*.

### 2.2 PREDIÇÃO DO IMPACTO DAS VARIANTES

O impacto funcional das variantes não sinônimas identificadas foi previsto utilizando um conjunto robusto de ferramentas de bioinformática, que avaliam as substituições de aminoácidos com base em critérios estruturais, funcionais e evolutivos:

- *SIFT (Sorting Intolerant From Tolerant)*: Avalia a intolerância de aminoácidos substituídos, baseando-se na homologia das sequências e nas características físico-químicas. Classifica as variantes como "Deletério" (score 0) ou "Tolerado" (score > 0,05).
- *PolyPhen-2 (Polymorphism Phenotyping v2)*: Estima a probabilidade de impacto funcional das variantes, categorizando-as em "Provavelmente prejudicial" (score 0,85–1,0), "Possivelmente prejudicial" (score 0,15–0,85) ou "Benigno" (score < 0,15).
- *MetaLR*: Prediz o grau de patogenicidade das variantes com base em modelos de aprendizado de máquina, classificando-as em "Prejudicial" (score próximo de 1) ou "Benigno" (score próximo de 0).
- *Mutation Assessor*: Avalia o impacto das variantes levando em consideração a conservação evolutiva das posições afetadas em homólogos da proteína. Classifica as variantes como "Neutro", "Baixo", "Médio" ou "Alto", com base no *score* de impacto funcional.

### 3 RESULTADOS

O transcrito IL1A-203 disponível no *ENSEMBL* apresenta um total de 250 variantes não sinônimas, das quais 5 estão na região de *splicing* do gene. Além disso, foram identificadas 8 variantes que introduzem um códon de parada (*Stop gained*), 7 variantes com mudança de fase de leitura (*frameshift variant*), 2 variantes que perdem um códon de parada (*Stop lost*), 2 variantes que perdem um códon de iniciação (*Start lost*), 1 inserção que mantém a fase de leitura (*inframe insertion*), 3 deleções que mantêm a fase de leitura (*inframe deletion*), 14 variantes na região de *splicing*, 1 variante que mantém o códon de parada (*Stop retained*), 91 variantes sinônimas e 110 variantes na sequência codificadora.

A classificação das variantes *missenses* é mostrada na Tabela 1 de acordo com os preditores de informática usados. A análise das variantes no SIFT resultou em 113 variantes *missenses* classificadas como deletérias, no entanto, duas dessas predições apresentam baixa confiabilidade. Além disso, 137 variantes foram consideradas de efeitos tolerados. No *software* PolyPhen-2, 42 variantes foram classificadas como provavelmente prejudiciais, 47 como possivelmente prejudiciais e 36 como benignas usando o MetaLR, 3 variantes foram consideradas prejudiciais, enquanto 247 foram classificadas como benignas. No *Mutation Assessor*, 89 variantes apresentaram médio impacto funcional na proteína, 123 apresentaram baixo impacto e 38 foram classificadas como neutras.

Na Tabela 2, são mostradas as variantes não sinônimas com maior impacto potencial na proteína, filtradas com base nos resultados dos programas de predição, resultando na identificação de três SNPs. A variante rs1190431689 corresponde à substituição de Alanina (Ala) por Valina (Val) na posição 2 da cadeia polipeptídica. Esta variante foi classificada como Deletéria (score: 0) no SIFT, Provavelmente Prejudicial (score: 0,921) no PolyPhen-2, Prejudicial (score: 0,547) no MetaLR e de Médio impacto (score: 0,553) no Mutation Assessor. A variante rs997926639 refere-se à substituição de Prolina (Pro) por Serina (Ser) na posição 5 da sequência de aminoácidos, com impactos previstos como Deletério (score: 0) no SIFT, Provavelmente Prejudicial (score: 0,961) no PolyPhen-2, Prejudicial (score: 0,683) no MetaLR e de Médio impacto (score: 0,553) no Mutation Assessor. A variante rs1681201328, localizada na região de *splicing*, envolve a substituição de Ácido Glutâmico (Glu) por Glicina (Gly) na posição 106 da sequência proteica, com impactos previstos como Deletério (score: 0) no SIFT, Provavelmente Prejudicial (score: 0,917) no PolyPhen-2, Prejudicial (score: 0,503) no MetaLR e de Médio Impacto (score: 0,553) no Mutation Assessor (Tabelas 2 e 3).

A Tabela 3 apresenta a análise das variantes *missenses* inseridas na região de *splicing* do gene *IL1A*. A variante rs200686721, que envolve a substituição de Serina (Ser) por Asparagina (Asn) na posição 16, foi classificada como "Tolerada" pelo SIFT, com um score de 0,05. No PolyPhen-2, esta variante foi avaliada como "Possivelmente Prejudicial" com um score de 0,794. O MetaLR também a classificou como "Tolerada", com um score de 0,406. Já o Mutation Assessor atribuiu um impacto

"Médio", com um score de 0,553. A variante rs1350937353, onde a Glutamina (Gln) é substituída por Histidina (His) na posição 32, apresentou classificações divergentes entre as ferramentas. No SIFT, recebeu a classificação "Deletério" com um score de 0, indicando um impacto negativo significativo na função da proteína. O PolyPhen-2 a classificou como "Provavelmente prejudicial", com um score de 0,936, apontando para um efeito adverso substancial. Em contraste, o MetaLR a considerou "Tolerada" com um score de 0,306, e o Mutation Assessor a avaliou com um impacto "Médio" de 0,553. A variante rs1681201211, na qual o Ácido Glutâmico (Glu) é substituído por Lisina (Lys) na posição 107, também recebeu classificações variadas. No SIFT, foi considerada "Deletério" com um score de 0,03, indicando um impacto negativo significativo. O PolyPhen-2 a classificou como "Possivelmente prejudicial" com um score de 0,879. Em contraste, o MetaLR a avaliou como "Tolerada" com um score de 0,374, e o Mutation Assessor indicou um impacto "Baixo" com um score de 0,553. Já a variante rs777688692, que envolve a substituição de Lisina (K) por Arginina (R) na posição 205, foi classificada como "Deletério" pelo SIFT, com um score de 0,03, mas recebeu uma classificação "Tolerada" do MetaLR com um score de 0,154 e "Baixo" impacto pelo Mutation Assessor. Não foi fornecida uma classificação pelo PolyPhen-2 para esta variante.

Tabela 1. Classificação das variantes *missenses* do gene *IL1A* humano e sua predição de impacto funcional utilizando diferentes ferramentas de bioinformática.

SIFT	Qntd	PolyPhen-2	Qntd	MetaLR	Qntd	Mutation Assessor	Qntd
Deletério	111	Provavelmente prejudicial	42	Prejudicial	3	Médio	89
Deletério (Baixa confiabilidade)	2	Possivelmente prejudicial	47	Benigno	247	Baixo	123
Tolerado	137	Benigno	36	-	-	Neutro	38
Total	250		250		250		250

Tabela 2. Variantes *missenses* do gene *IL1A* preditos com elevado risco para o desenvolvimento de doenças humanas.

Variante Missense	Resíduo	Predição			
		SIFT	PolyPhen-2	MetaLR	Mutation Assessor
rs1190431689	2 Ala/Val	Deletério 0	Provavelmente prejudicial 0,921	Prejudicial 0,547	Médio 0,553

rs997926639	5 Pro/ Ser	Deletério 0	Provavelmente prejudicial 0,961	Prejudicial 0,683	Médio 0,553
rs1681201328*	106 Glu/Gly	Deletério 0	Provavelmente prejudicial 0,917	Prejudicial 0,503	Médio 0,553

\*Variante localizada na região de *splicing* do gene.

Tabela 3. Predição das variantes *missenses* na região de *splicing*.

Variante	Resíduo	Predição			
		SIFT	PolyPhen-2	MetaLR	Mutation Assessor
rs200686721	16 S(Ser)/ N (Asn)	Tolerado 0,05	Possivelmente prejudicial 0,794	Tolerado 0,406	Médio
rs1350937353	32 Q(Gln)/ H(His)	Deletério 0	Provavelmente prejudicial 0,936	Tolerado 0,306	Médio
rs1681201328	106 E(Glu) / G(Gly)	Deletério 0	Provavelmente prejudicial 0,917	Prejudicial 0,683	Médio
rs1681201211	107 E(Glu)/ K(Lys)	Deletério 0,03	Possivelmente prejudicial 0,879	Tolerado 0,374	Baixo
rs777688692	205 K (Lys)/R(Arg)	Deletério 0,03	-	Tolerado 0,154	Baixo

#### 4 DISCUSSÃO

A simetria tridimensional interna de três dobras do núcleo estrutural é notavelmente precisa e com alta fidelidade (Graves *et al.*, 1990). Sua estabilidade estrutural é uma grande preocupação porque somente proteínas estáveis podem desempenhar sua função de forma eficiente (Kulshreshtha *et al.*, 2016). Portanto, mudanças de aminoácidos com conformidades físico-químicas muito diferentes, podem resultar na desestabilização da molécula e afetar a atuação e o desempenho correto da *IL-1α* no organismo.

Ao examinarmos os aminoácidos substituídos em cada variante disponível, podemos inferir que as predições da rs1190431689 é altamente prejudicial, visto que a substituição do aminoácido Alanina (Ala) por uma Valina (Val) pode apresentar impactos severos na estrutura da proteína (Guzzi, 2018). Embora ambas as moléculas sejam hidrofóbicas e possuam estruturas químicas semelhantes, o que pode não afetar significativamente a estrutura tridimensional da proteína ou sua função biológica, a mudança de um aminoácido pode alterar as propriedades físicas da proteína e afetar sua atividade biológica (Redler *et al.*, 2016.). A valina, por exemplo, é uma cadeia lateral mais longa e volumosa

que a alanina, o que pode afetar a conformação da proteína, a interação com outras proteínas ou receptores e sua capacidade de ser processada e secretada pelas células produtoras.

A variante rs997926639 é resultado da substituição de um aminoácido Prolina (Pro) por Serina (Ser) e pode apresentar diferentes impactos, dependendo do contexto em que ocorre. A prolina é um aminoácido não polar e sua substituição por uma serina, que é um aminoácido polar, pode alterar a estrutura tridimensional da proteína e afetar sua função biológica. A presença da prolina na estrutura da proteína ajuda a estabilizar a hélice alfa, que é uma estrutura secundária importante na proteína. A substituição por serina pode interromper essa estabilização e afetar a estrutura tridimensional da proteína. Além disso, a serina possui uma cadeia lateral hidroxila que pode alterar a polaridade local da proteína e afetar sua interação com outras moléculas. Em relação à função biológica, estudos mostraram que a substituição de prolina por serina na *IL-1α* pode afetar sua capacidade de se ligar a receptores específicos e modular a resposta inflamatória. <https://www.studocu.com/pt-br/document/instituto-federal-de-educacao-ciencia-e-tecnologia-do-rio-grande-do-norte/biologia/tipos-de-mutacao/17397090>. A substituição de ácido glutâmico (Glu), um aminoácido ácido com uma cadeia lateral longa, por glicina (Gly), um aminoácido neutro e com uma cadeia lateral muito curta, pode ter um impacto significativo. O ácido glutâmico tem um grupo carboxila na sua cadeia lateral, que contribui para a carga negativa da proteína e pode interagir com outras moléculas de forma específica. A glicina, sendo muito pequena e neutra, pode alterar a carga e a capacidade de interação da proteína. A mudança pode afetar a estabilidade da estrutura tridimensional da proteína, bem como suas propriedades bioquímicas e funcionais.

Com estes resultados, fica evidente que os *softwares* de predição podem ser extremamente úteis para a compreensão de diversas patologias, pois permitem identificar variantes associadas às doenças e como estas podem interferir na estabilidade molecular (Freitas *et al.*, 2022). Além disso, recomendamos que estudos futuros sejam direcionados a compreender a história evolutiva do gene *IL1A* para determinar o grau de conservação do gene e a sua importância evolutiva para a saúde humana, visto que, conhecer a evolução das interleucinas permite o entendimento da origem e mecanismos regulatórios de diversas patologias associadas às citocinas em humanos (Freitas *et al.*, 2024).

## 5 CONCLUSÃO

Este trabalho buscou estudar e identificar as mutações *missenses* associadas ao gene humano *IL1A*, destacando variantes com alto risco para o desenvolvimento de doenças humanas e compreender seus possíveis impactos físico-químicos na molécula e seus efeitos na atuação da citocina no organismo. Assim, servindo como base para auxiliar novos estudos na identificação de genes de interesse associados ao desenvolvimento de doenças humanas.





## REFERÊNCIAS

- BECERRA, Carlos Roberto et al. A phase I study of isunakinra, an IL-1 alfa/beta inhibitor, in combination with nivolumab for patients with solid tumors refractory to standard therapies. 2023.
- FREITAS, A.F.F. et al. In silico analysis of the impact of non-synonymous Single Nucleotide Polymorphisms (nsSNPs) in the human Il-6 gene related to autoimmune diseases. DOI:10.29327/229003.3.1-1. 2022.
- FREITAS, A. F. F.; DE FREITAS, N. S. A.; MARIANO, M. H. Q. de A.; RUSHANSK, E.; MAIA, M. de M. D. Padrões evolutivos da interleucina-6 (il-6) e seu impacto para a saúde humana. *Brazilian Journal of Health Review*, [S. l.], v. 7, n. 1, p. 4624–4637, 2024. DOI: 10.34119/bjhrv7n1-374. Disponível em: <https://ojs.brazilianjournals.com.br/ojs/index.php/BJHR/article/view/66998>. Acesso em: 14 oct. 2024.
- GRAVES, Bradford J. et al. Structure of interleukin 1. alpha. at 2.7-Å resolution. *Biochemistry*, v. 29, n. 11, p. 2679-2684, 1990.
- GUZZI, ANETE FERAZ. Caracterização in silico do polimorfismo A122V na proteína receptora de quimiocina do tipo 1 (CXCR1) associado com suscetibilidade à mastite em bovinos. Thesis em Pt | VETTESES | ID: vtt-217109. 2018. Biblioteca responsável: [BR68.1](#)
- HEEB, Lukas EM; EGHOLM, Cecilie; BOYMAN, Onur. Evolution and function of interleukin-4 receptor signaling in adaptive immunity and neutrophils. *Genes & Immunity*, v. 21, n. 3, p. 143-149, 2020.
- KUBASKI, Francyne et al. Análises in silico para predição do fenótipo na substituição de aminoácidos no gene da GALNS. *Revista HCPA*. Porto Alegre, 2013.
- KULSHRESHTHA, Shweta et al. Computational approaches for predicting mutant protein stability. *Journal of computer-aided molecular design*, v. 30, p. 401-412, 2016.
- MOREIRA, P. R. et al. The IL1A (– 889) gene polymorphism is associated with chronic periodontal disease in a sample of Brazilian individuals. *Journal of periodontal research*, v. 42, n. 1, p. 23-30, 2006.
- REDLER, Rachel L. et al. Protein destabilization as a common factor in diverse inherited disorders. *Journal of molecular evolution*, v. 82, p. 11-16, 2016.
- VARELLA, Pedro PV; FORTE, Wilma C. Neves. Citocinas: revisão. *Rev. bras. alergias imunopatol*, p. 146-154, 2001.  
<https://www.studocu.com/pt-br/document/instituto-federal-de-educacao-ciencia-e-tecnologia-do-rio-grande-do-norte/biologia/tipos-de-mutacao/17397090>.  
Visitado em 20 de outubro de 2024.