



Análise dos polimorfismos do gene CCR5 em populações globais: um caminho para compreender a imunogenética da infecção pelo HIV e outras doenças

Isabella Martins Thomaz¹, Flávia Peixoto da Silva Guimarães², Millena Cardoso Sales Santos²,
Isabella Isaac Lopes², Iane de Oliveira Pires Porto³

¹ Graduanda do curso de Medicina, participante do Programa de Iniciação Científica – PIVIC/UnirV, Universidade de Rio Verde.

² Graduanda do curso de Medicina, Universidade de Rio Verde.

³ Orientadora, Prof. Dr. da Faculdade de Medicina, Universidade de Rio Verde.

E-mail: isabellamthomaz@academico.unirv.edu.br

Reitor:

Prof. Me. Alberto Barella Netto

Pró-Reitor de Pesquisa e Inovação:

Prof. Dr. Carlos César E. de Menezes

Editor Geral:

Prof. Dr. Fábio Henrique Baia

Editor de Seção:

Profa. Dra. Andrea Sayuri
Silveira Dias Terada
Prof. Dr. Hidélberto Matos Silva

Correspondência:

Isabella Martins Thomaz

Fomento:

Programa PIBIC/PIVIC UnirV/
CNPq 2021-2022

Resumo: O receptor de quimiocinas CC 5 (CCR5) é responsável por regular funções de efeito funcional intracelular e tráfego dos linfócitos T, macrófagos e células dendríticas imaturas, mas se destacou por sua grande importância na regulação da fisiopatologia da infecção pelo vírus da imunodeficiência humana (HIV). Esse trabalho tem por objetivo analisar a variabilidade do gene CCR5, utilizando dados de amostras genéticas internacionais de 2504 indivíduos, abrangendo 26 populações mundiais, e 1.171 amostras brasileiras. Verificou-se as mutações de maior frequência global e nas respectivas populações estudadas. Quanto ao caráter das mutações, 64% geram mutação não sinônima, 25% são sinônimas, 7% resultam em um códon de parada e cerca de 5% mudam o quadro de leitura. Os principais polimorfismos com significância publicada são rs1799987, rs41469351, rs1800560 e rs333, dentre outros. A partir deste estudo, foi possível elucidar os polimorfismos do gene CCR5 e as consequências a proteína final, patogenicidade e suas correlações com inúmeras doenças, em especial, o HIV.

Palavras-chave: Genética humana. Receptores CCR5. Polimorfismo genético.

Analysis of CCR5 gene polymorphisms in global populations: a path to understanding the immunogenetics of HIV infection and other diseases

Abstract: The CC 5 (CCR5) chemokine receptor is responsible for regulating functions of intracellular functional effect and T lymphocytes, macrophages and immature dendritic cells traffic, but it stands out for its great importance in regulating the pathophysiology of human immunodeficiency virus infection. (HIV). This paper aims to analyze the variability of the CCR5 gene, using data from international genetic samples from 2504 individuals, covering 26 populations worldwide, and 1171 Brazilian samples. The most frequent mutations were found globally and in the respective populations analyzed. As for the characterization of the mutations, 64% result in a non-synonymous mutation, 25% are synonymous, 7% result in a stop codon and about 5% change the reading frame. The main polymorphisms with published significance are rs1799987, rs41469351, rs1800560 and rs333, among others. From this paper, it was possible to elucidate the main

polymorphisms of the CCR5 gene and its consequences in the final protein, pathogenicity and its correlations with numerous diseases, especially HIV.

Key words: Human genetics. CCR5 receptors. Genetic polymorphism.

Introdução

Na década de 90, ao analisar a genética de populações de infectados pelo vírus da imunodeficiência humana (HIV), vários estudos trouxeram o Receptor 5 de quimiocinas CC (CCR5) como um importante regulador da fisiopatologia da infecção pelo HIV. Estudos posteriores elucidam que isso se deve ao CCR5 ser um cofator de fusão para a entrada cepas de HIV-1 com tropismo para macrófagos (LIN, 2021; MCNICHOLL, 1997).

Os receptores de quimiocina pertencem a superfamília de receptores acoplados à proteína ligadora de trifosfato de guanossina, proteína G, (GPCR) de sete alças transmembrana (7TM), sendo responsáveis por regular funções de efeito funcional intracelular e tráfego dos linfócitos T, macrófagos e células dendríticas imaturas (MUMMIDI, 1997; OPPERMAN, 2004). Muitos polimorfismos encontrados no gene CCR5 parecem estar diretamente relacionados ao curso da infecção por HIV e pela imunodeficiência adquirida. Os portadores de genótipo homozigoto para a deleção de 32pb (CCR5 Δ 32/CCR5 Δ 32), no éxon 3, são relativamente imunes a infecção ao HIV, mesmo com comportamento de risco (LIU, 1996; MICHAEL, 1997). Por outro lado, indivíduos heterozigotos de genótipo CCR5/CCR5 Δ 32 podem desenvolver a infecção por HIV-1, porém com progressão mais branda da doença (MUMMIDI, 1997). No entanto, mesmo nos indivíduos sem a mutação, a expressão desse receptor na superfície celular é altamente variável (WU, 1997). Outras mutações também são notadas por sua influência na progressão do HIV e associados, também, com a resistência/susceptibilidade a outras doenças como Hepatite C, e Vírus do Nilo Ocidental, Diabetes Mellitus, dentre outras (WANG, 2014; COENEN, 2010; GLASS, 2006).

A terapia genética se tornou uma grande esperança para uma possível cura definitiva do HIV, devido aos promissores polimorfismos encontrados. Até o momento, a literatura carece de estudos publicados atualizados, que utilizem técnicas modernas de sequenciamento, sobre o gene CCR5 e sua variabilidade estrutural. Portanto, esse trabalho tem por objetivo analisar a variabilidade e consequências dos principais polimorfismos das diferentes regiões

do gene CCR5, através de sequências genômicas de bancos de dados internacionais e brasileiros.

Material e Métodos

As sequências do gene CCR5 foram obtidas na base de dados de amostras genéticas internacionais, 1000 Genomes (<https://www.ncbi.nlm.nih.gov/variation/tools/1000genomes/>), em formato “.vcf”. O arquivo foi convertido para os respectivos arquivos de entrada “.arp” pelo software PGD Spider. O projeto 1000 genomes conta com o sequenciamento genético de 2504 indivíduos, de 213 etnias, agrupadas em 26 populações mundiais distribuídas por 5 continentes, com diferentes ancestralidades (africana, asiática, europeia e ameríndia). Para avaliar a frequência dessas mutações na população brasileira foi utilizado a base de dados ABraOM, a qual é composta de 1.171 indivíduos saudáveis, de região urbana do estado de São Paulo. (NASLAVSKY, 2022).

Através dos arquivos obtidos, foi possível utilizar o software Arlequin (EXCOFFIER; LINSCHER, 2010) e VCFtools (DENECEK, 2011) para computar dados das sequências e apuração de informações significativas para compreender os polimorfismos, além de frequência genotípica e frequência de haplótipos. As sequências foram analisadas como um todo e, posteriormente, em suas respectivas populações.

As consequências dos polimorfismos foram imputadas pelo software Sift (sift.bii.a-star.edu.sg) e PolyPhen (genetics.bwh.harvard.edu/pph2) e correlacionados com o encontrado na ferramenta VEP (Preditor de Efeito de Variante) do banco de dados Ensembl (ensembl.org/info/docs/tools/vep/). Dessa forma, será possível quantificar os diferentes tipos de mutações encontradas com base em seu efeito estrutural e extrapolar suas consequências funcionais na proteína com a sua patogenicidade. Por fim, observamos os achados com os dados disponíveis em literatura a respeito dos polimorfismos do gene CCR5 e sua importância clínica, principalmente na infecção pelo vírus HIV-1. Para este fim, também foram utilizadas bases de dados com o OMIN (Online Mendelian Inheritance in Men) e ClinVar (ncbi.nlm.nih.gov/clinvar/), com o objetivo de retificar as importâncias clínicas das variantes encontradas.

Resultados e Discussão

O gene CCR5 ocupa uma região de 5,353 bp do cromossomo 3 (3p21.31), localizado na região genômica do grupo de receptores de quimiocinas, possui 3 éxons e 2 íntrons. Diferente da maioria

dos GPCRs, nas regiões 5' não traduzidas (5'UTR, do inglês, 5'Untranslated Region) do CCR5 residem em múltiplos sítios de splicing alternativos que podem gerar diferentes isoformas de RNAm. Das 2504 amostras analisadas foram encontrados 240 haplótipos, sendo o mais predominante, das

populações, estudadas o de frequência 0,25463. Quanto a análise primária do desequilíbrio de ligação, foi possível identificar pares de alelos em desequilíbrio (dados não representados). A tabela 1 representa os principais polimorfismos (frequência alélica maior que 1%) encontrados globalmente.

Tabela 1 - polimorfismos do gene CCR5 e suas respectivas frequências alélicas.

SNPid	AR	Freq. AR	AA	Freq. AA	SNPid	AR	Freq. AR	AA	Freq. AA
rs41475349	G	0.9248	A	0.0751	rs71619644	CACA	0.6568	C	0.3376
rs2227010	G	0.2781	A	0.7219	rs41515644	A	0.6688	G	0.3311
rs2856758	A	0.9455	G	0.0544	rs3176763	G	0.9345	T	0.0654
rs2734648	G	0.5870	T	0.4130	rs333	Δ32	0.9708	-	0.029
rs1799987	A	0.4507	G	0.5493	rs1800452	G	0.9879	A	0.0120
rs1799988	C	0.4507	T	0.5493	rs1800944	C	0.9871	T	0.0128
rs41469351	C	0.9290	T	0.0710	rs41495153	G	0.9351	A	0.0648
rs1800023	A	0.6621	G	0.3379	rs17765882	C	0.9282	T	0.0717
rs1800024	C	0.8321	T	0.1678	rs41418945	G	0.9842	A	0.0157
rs9282632	A	0.9353	C	0.0646	rs41418945	G	0.9842	A	0.0157
rs3181037	A	0.9353	G	0.0646	rs1800874	G	0.6629	T	0.3370
rs3216198	CAG	0.6623	C	0.3376	rs41535253	T	0.9865	C	0.0134
rs2856762	C	0.9280	T	0.0719	rs3188094	A	0.9861	C	0.0138
rs2254089	C	0.6619	T	0.3380	rs41442546	C	0.9857	A	0.0142
rs41395249	G	0.9895	A	0.0104	rs746492	G	0.4445	T	0.5554
rs2856764	C	0.6621	T	0.3378	rs41403149	C	0.9871	T	0.0128
rs2856765	G	0.6623	A	0.3376	rs746492	G	0.4445	T	0.5554

SNPid: Identificador do polimorfismo. AR: Alelo de Referência. AA: Alelo Alternativo. Freq: Frequência alélica. Δ32 corresponde a sequência de alelos: A(CAGT)²ATCAATTCTGGAAGAATTTCCAG. Fonte: Autoria própria.

Quanto à análise populacional das amostras globais - referente às 2504 sequências do 1000Genomes - e brasileiras - referentes às 1.171 sequências do ABra-OM, foram selecionados os polimorfismos de frequência >2% (por questões gráficas), apresentados na Tabela 2 e 3. É possível verificar alguns polimorfismos, como o rs41475349, rs41469351, rs9282632,

rs3181037, rs3176763, rs333, rs41495153, que embora tenham frequência global considerável, são prevalentes em populações específicas. Na população brasileira estudada, os polimorfismos mais frequentes são o rs746492 (0,475), rs1799987 (0,470), rs2734648 (0,38), rs201063877 (0,350), rs2254089 (0,347), dentre outros.

Tabela 2 – polimorfismos do gene e suas respectivas frequências globais e por população (parte 1).

	rs41475349	rs2227010	rs2856758	rs2734648	rs1799987	rs1799988	rs41469351	rs1800023	rs1800024	rs9282632	rs3181037	rs3216198
global	0,075	0,721	0,054	0,412	0,548	0,548	0,071	0,337	0,168	0,071	0,071	0,336
ACB	0,197	0,802	0,031	0,333	0,541	0,541	0,192	0,135	0,239	0,197	0,197	0,135
ASW	0,155	0,844	0,065	0,360	0,573	0,573	0,147	0,204	0,245	0,196	0,196	0,204

BEB	-	0,779	0,005	0,436	0,616	0,616	-	0,436	0,162	-	-	0,436
CDX	-	0,833	-	0,774	0,629	0,629	-	0,612	0,204	-	-	0,774
CEU	-	0,555	0,121	0,363	0,469	0,469	-	0,363	0,085	-	-	0,363
CHB	-	0,839	-	0,538	0,592	0,592	-	0,538	0,247	-	-	0,538
CHS	-	0,776	-	0,566	0,590	0,590	-	0,566	0,185	-	-	0,566
CLM	0,021	0,696	0,079	0,414	0,494	0,494	0,021	0,393	0,202	0,021	0,021	0,393
ESN	0,252	0,803	0,040	0,338	0,570	0,570	0,242	0,085	0,222	0,232	0,232	0,085
FIN	-	0,545	0,156	0,297	0,434	0,434	-	0,292	0,111	-	-	0,292
GBR	-	0,500	0,208	0,318	0,428	0,428	-	0,318	0,071	-	-	0,318
GIH	-	0,669	0,019	0,359	0,587	0,587	-	0,359	0,082	-	-	0,359
GWD	0,380	0,823	0,053	0,415	0,659	0,659	0,362	0,035	0,141	0,243	0,243	0,035
IBS	-	0,509	0,112	0,369	0,448	0,448	-	0,369	0,060	0,009	0,009	0,369
ITU	-	0,656	0,024	0,406	0,593	0,593	-	0,406	0,063	-	-	0,406
JPT	-	0,769	-	0,451	0,475	0,475	-	0,451	0,293	-	-	0,451
KHV	-	0,828	-	0,631	0,651	0,651	-	0,631	0,176	-	-	0,631
LWK	0,166	0,818	0,030	0,257	0,510	0,510	0,151	0,095	0,308	0,252	0,252	0,090
MSL	0,329	0,894	0,017	0,394	0,676	0,676	0,3	0,064	0,194	0,282	0,282	0,064
MXL	0,007	0,609	0,125	0,406	0,453	0,453	0,007	0,398	0,156	0,015	0,015	0,398
PEL	0,011	0,664	0,147	0,282	0,335	0,335	0,011	0,270	0,329	0,011	0,011	0,270
PJL	-	0,718	0,026	0,395	0,619	0,619	-	0,395	0,088	-	-	0,395
PUR	0,028	0,615	0,062	0,370	0,447	0,447	0,028	0,341	0,168	0,028	0,028	0,341
STU	-	0,725	-	0,411	0,612	0,612	-	0,411	0,112	-	-	0,411
TSI	-	0,630	0,093	0,481	0,528	0,528	-	0,481	0,102	0,004	0,004	0,481
YRI	0,376	0,870	0,041	0,449	0,671	0,671	0,356	0,069	0,180	0,222	0,222	0,069
BRA ^a	-	-	0,090	0,381	0,471	0,471	0,03	0,347	0,141	0,045	0,045	-

^a Frequência obtida através da base de dados brasileira (ABraOM), não incluída na tabela de frequência global; Finlandeses na Finlândia (FIN), chineses Han do Sul (CHS), Kinh na cidade de Ho Chi Minh, Vietnã (KHV), bengali em Bangladesh (BEB), porto-riquenhos em Porto Rico (PUR), afro-caribenhos em Barbados (ACB), residentes de Utah com ascendência do norte e da Europa Ocidental (CEU), povos Esan na Nigéria (ESN), colombianos em Medellín, Colômbia (CLM), peruanos em Lima, Peru (PEL), povos Ibéricos na Espanha (IBS), pessoas de ascendência mexicana em Los Angeles, Califórnia (MXL), índios Gujarati em Houston, Texas (GIH), telugos indianos no Reino Unido (ITU), britânicos na Inglaterra e Escócia (GBR), povos de ascendência africana no sudoeste dos EUA (ASW), lorubás em Ibadan, Nigéria (YRI), Gambianos na Divisão Ocidental, Gâmbia (GWD), japoneses em Tóquio, Japão (JPT), chineses Han em Pequim, China (CHB), Mandês em Serra Leoa (MSL), chineses Dai em Xishuangbanna, China (CDX), Punjabis em Lahore, Paquistão (PJL), toscanos na Itália (TSI), Luha em Webuye, Quênia (LWK), tâmeis do Sri Lanka no Reino Unido (STU). Fonte: Autoria própria.

Quando avaliamos as consequências dos polimorfismos através do Preditor de Efeito de Variante (VEP) encontramos o caráter das mutações, sendo possível verificar que 64% geram mutação não sinônima, 25% são sinônimas, 7% resultam em um

códon de parada e cerca de 5% mudam o quadro de leitura. Todas as variantes que causam efeitos deletérios – alteram a estrutura da proteína final – segundo a análise pelo PolyPhen e Sift, possuem baixa frequência global (<0,01)

Tabela 3 – polimorfismos do gene e suas respectivas frequências globais e por população (Parte 2).

	rs2856762	rs2254089	rs2856764	rs2856765	rs201063877	rs41515644	rs3176763	rs333	rs41495153	rs17765882	rs1800874	rs746492
Global	0,071	0,337	0,337	0,337	0,336	0,330	0,065	0,029	0,065	0,071	0,336	0,554
ACB	0,010	0,135	0,135	0,135	0,135	0,135	0,203	0,005	0,197	0,010	0,135	0,552
ASW	0,016	0,204	0,204	0,204	0,204	0,204	0,196	0,024	0,196	0,016	0,204	0,590

BEB	0,180	0,436	0,436	0,436	0,436	0,424	-	-	-	0,180	0,436	0,616
CDX	0,021	0,612	0,612	0,612	0,612	0,586	-	-	-	0,016	0,774	0,629
CEU	0,106	0,368	0,363	0,363	0,363	0,358	-	0,106	-	0,106	0,363	0,474
CHB	0,053	0,538	0,538	0,538	0,538	0,533	-	-	-	0,053	0,538	0,592
CHS	0,023	0,566	0,566	0,566	0,566	0,552	-	-	-	0,023	0,566	0,590
CLM	0,058	0,393	0,393	0,393	0,393	0,388	0,021	0,042	0,021	0,058	0,393	0,494
ESN	-	0,085	0,085	0,085	0,085	0,085	0,232	-	0,232	-	0,085	0,601
FIN	0,141	0,292	0,292	0,292	0,292	0,292	-	0,156	-	0,141	0,292	0,434
GBR	0,109	0,318	0,318	0,318	0,318	0,313	-	0,142	-	0,109	0,318	0,428
GIH	0,228	0,359	0,364	0,359	0,359	0,359	-	0,014	-	0,228	0,359	0,597
GWD	-	0,035	0,035	0,035	0,035	0,035	0,247	-	0,243	-	0,035	0,668
IBS	0,070	0,369	0,369	0,369	0,369	0,359	0,009	0,088	0,009	0,070	0,369	0,453
ITU	0,406	0,406	0,406	0,406	0,406	0,406	-	0,004	-	0,186	0,406	0,598
JPT	0,024	0,451	0,451	0,451	0,451	0,442	-	-	-	0,024	0,447	0,475
KHV	0,202	0,631	0,631	0,631	0,631	0,601	-	-	-	0,020	0,631	0,651
LWK	-	0,090	0,090	0,090	0,090	0,090	0,252	-	0,257	-	0,085	0,525
MSL	-	0,064	0,064	0,064	0,064	0,064	0,294	-	0,282	-	0,064	0,694
MXL	0,031	0,398	0,398	0,398	0,398	0,390	0,015	0,039	0,015	0,031	0,398	0,453
PEL	0,041	0,276	0,270	0,270	0,270	0,270	0,011	0,017	0,011	0,041	0,270	0,347
PJL	0,223	0,395	0,395	0,395	0,395	0,390	-	0,026	-	0,223	0,395	0,619
PUR	0,048	0,341	0,341	0,341	0,341	0,336	0,028	0,028	0,028	0,048	0,341	0,451
STU	0,200	0,411	0,411	0,411	0,411	0,387	-	-	-	0,200	0,411	0,617
TSI	0,042	0,481	0,481	0,481	0,481	0,476	0,004	0,065	0,004	0,042	0,481	0,528
YRI	-	0,069	0,069	0,069	0,069	0,069	0,222	-	0,222	-	0,069	0,694
BRA ^a	0,043	0,347	0,347	0,347	0,350	0,346	0,045	0,004	0,046	0,043	0,346	0,475

^a Frequência obtida através da base de dados brasileira (ABraOM), não incluída na tabela de frequência global. Fonte: Autoria própria.

Acerca da significância clínica das variantes estudadas, encontramos que o rs1799987 tem 69 relatos em literatura por sua capacidade de proteção a infecção por HIV e aumento da sobrevivência de pacientes com linfoma não-Hodgkin, bem como efeito patogênico e fator de risco na nefropatia na diabetes mellitus tipo 1 e tromboembolismo venoso, dentre outras correlações publicadas (JOUBERT, 2010; ASCHEBROOK-KILFOY, 2012; TRÉGOUET, 2008). Já o polimorfismo rs41469351 foi relatado como patogênico, uma vez que está relacionado a susceptibilidade a transmissão perinatal de HIV-1 (KOSTRIKIS, 1999), além disso, conforme tabela 2, é notória sua alta prevalência em populações de ancestralidade africana (ACB: 0,192; ASW: 0,147; MSL:0,300), Gâmbia (GWD: 0,362) e Nigéria (YRI:0,356), enquanto é ausente ou pouco prevalente em outras etnias. A variante rs1800560 resulta em um códon de parada e tem valor protetivo contra a infecção do HIV-1 (QUIL-LENT, 1998), porém tem uma frequência global de 0,0004.

Os estudos de variabilidade do gene CCR5 tendem a evidenciar apenas seus éxons e região promotora (QUIL-LENT, 1998; SOLLOCK, 2017; BALWI, 2017), com foco especial a uma inserção/deleção de 32 pares de bases (32pb) devido à sua influência na invasão celular pelo vírus HIV. A homozigose da deleção de 32-bp na região codificadora do gene CCR5 (CCR5 Δ 32 – rs333) foi reportada como um importante protetor contra a infecção do HIV. Essa mutação causa uma mudança no quadro de leitura do RNA mensageiro (frame-shift) que resulta em um códon de parada prematura e gera uma forma truncada do CCR5, com 215 aminoácidos, ao invés de 352. Dessa forma, a proteína truncada não é expressa na superfície celular, impedindo a ligação do vírus HIV (RONSARD, 2019). Considerando o rs333, 115 estudos demonstram seu valor protetivo para a infecção por HIV-1, hepatite C, flavivírus, varíola e facilita a recuperação após acidente vascular cerebral, estudos recentes com outras populações observaram que, quando em homozigose (Δ 32/ Δ 32), essa mutação pode ter

efeito deletério e resultam em menor expectativa de vida, maior mortalidade por Influenza – que é uma doença mais prevalente nas comunidades globais, comparado ao HIV - e quatro vezes mais susceptibilidade a certas doenças infecciosas (WEI, 2019). No ano de 2020, houve o nascimento de gêmeas chinesas que tiveram a edição gênica, pelo método CRISPR, para obter a CCR5 Δ 32, com o objetivo de ser resistente ao HIV (XU, 2020). Embora seja uma forma promissora de proteção, ainda não é possível prever todas as consequências desta mutação no indivíduo, sendo necessário mais estudos funcionais para compreender o papel do CCR5 nas patologias humanas.

Conclusão

A partir deste estudo, foi possível elucidar os principais polimorfismos do gene CCR5 com parâmetros globais e populacionais. As consequências das mutações e sua análise foi essencial para compreender os resultados na proteína final, patogenicidade e suas correlações com inúmeras doenças. No contexto ético, podemos verificar que há uma grande tentativa pela comunidade científica de avançar no contexto da terapia de edição gênica pelas novas tecnologias de engenharia genética e um dos principais polimorfismos visados é a deleção rs333. No entanto, uma deleção em homozigose em um gene de função tão importante como o CCR5 pode trazer malefícios tão significativos quanto os benefícios. Dessa forma, são necessários mais estudos ponderando as consequências destes polimorfismos.

Agradecimentos

À Faculdade de Medicina da Universidade de Rio Verde, ao Programa Institucional de Bolsas de Iniciação Científica e à minha orientadora Prof. Dr. Iane Porto pelo apoio, treinamento e bolsa.

Referências Bibliográficas

MCNICHOLL, J. et al. Host Genes and HIV: The Role of the Chemokine Receptor Gene CCR5 and Its Allele (Δ 32 CCR5). **Emerg. Infect. Dis.**, 3, n.3, 1997.

LIN, D. et al. Increased Efficiency for Biallelic Mutations of the CCR5 Gene by CRISPR-Cas9 Using Multiple Guide RNAs As a Novel Therapeutic Option for Human Immunodeficiency Virus. **The CRISPR Journal**, v. 4, n. 1, 2021.

MUMMIDI, S.; AHUJA, S. The Human CC Chemokine Receptor 5 (CCR5) Gene. **The Journal of Biological Chemistry**, 272, n.49, 1997.

LIU, R.; PAXTON, W. Homozygous defect in HIV-1 co-receptor accounts for resistance of some multiply-exposed individuals to HIV-1 infection. **Cell**, 36, p.367-377, 1996.

MICHAEL, N. et al. The role of viral phenotype and CCR-5 gene defects in HIV-1 transmission and disease progression. **Nat. Med.**, 3, p.338-340, 1997.

WU, L. et al. CCR5 levels and expression pattern correlate with infectability by macrophage-tropic HIV-1, in vitro. **J. Exp. Med.**, v. 185, n. 9, p. 1681-1691, 1997.

WANG, W.; YE, C. Contrasting Genetic Influence of CCR2 and CCR5 Variants on HIV-1 Infection and Disease Progression. **PLoS ONE**, 9, n.12, 2014.

COENEM, M. NATTERMANN, J. The role of CCR5 in HCV infection. **Eur. J. Med. Res.**, v. 15, n. 3, 2010.

GLASS, W. et al. CCR5 deficiency increases risk of symptomatic West Nile virus infection. **J. Exp. Med.**, v. 203, n. 1, 2006.

NASLAVSKY, M. et al. Whole-genome sequencing of 1,171 elderly admixed individuals from Brazil. **Nature**, v. 1005, 2022.

DENECK, P. et al. The variant call format and VCF-tools. **Bioinformatics**, v. 27, n. 15, 2011.

EXCOFFIER L. LINSCHER, H. Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. **Mol. Ecol. Resour.**, v. 10, n. 3, 2010.

AMBERGER, J. et al. OMIM.org: leveraging knowledge across phenotype-gene relationships. **Nucleic. Acids. Res.**, v. 47, 2019.

JOUBERT, B. et al. A whole genome association study of mother-to-child transmission of HIV in Malawi. **Genome Medicine**, v. 2, n. 3, 2010.

ASCHEBROOK-KILFOY, B. et al. Polymorphisms in immune function genes and non-Hodgkin lymphoma survival. **Journal of Cancer Survivorship**, v. 6, n. 1, 2012.

TRÉGOUET, D. et al. G/T substitution in intron 1 of the UNC13B gene is associated with increased risk of nephropathy in patients with type 1 diabetes. **Diabetes**, v. 57, n. 10, 2008.

ZEE, R. et al. Genetic risk factors in recurrent venous thromboembolism: A multilocus, population-based, prospective approach. **International Journal of Clinical Chemistry**, v. 402, n. 1, 2009.

GOULART, A. REXRODE, K. CHENG. Association of genetic variants with the metabolic syndrome in 20,806 white women: The Women's Health Genome Study. **American Heart Journal**, v. 158, n. 2, 2009.

KOSTRIKIS, L. A polymorphism in the regulatory region of the CC-chemokine receptor 5 gene influences perinatal transmission of human immunodeficiency virus type 1 to African-American infants. **J. Virol.**, v. 73, n. 12, 1999.

BALWI, M. Analysis of CCR5 gene polymorphisms in 321 healthy Saudis using Next Generation Sequencing. **Hum. Immunol.**, 2017.

QUILLET, C. *et al.* HIV-1-resistance phenotype conferred by combination of two separate inherited mutations of CCR5 gene. **The Lancet**, 351, 1998.

SOLLOCH, U.; LANG, K.; LANGE, V. *et al.* Frequencies of gene variant CCR5-Δ32 in 87 countries based on next-generation sequencing of 1.3 million individuals sampled from 3 national DKMS donor centers. **Human Immunology**, 2017.

WU, M. CCR5-Δ32 biology, gene editing, and warnings for the future of CRISPR-Cas9 as a human and humane gene editing tool. **Cell & Biology**, v. 48, 2020.

WEI, X. RASMUS, N. CCR5-Δ32 is deleterious in the homozygous state in humans. **Nature Medicine**, v. 25, 2019.