Indonesian Journal of Global Health Research
Volume 8 Number 2, April 2026
e-ISSN 2715-1972; p-ISSN 2714-9749
http://jurnal.globalhealthsciencegroup.com/index.php/IJGHR

IDENTIFICATION OF GENOMIC DATABASES AND BIOINFORMATICS ANALYSIS OF GENE VARIATIONS IN BEHÇET SYNDROME ACROSS CONTINENTAL POPULATIONS

Dwiki Fitri[1]*, Husnul Khuluq[1], Diah Kurniasih[1], Muhammad Ma'ruf[2]
[1]Pharmacy Study Program, Faculty of Health Sciences, Universitas Muhamamdiyah Gombong, Jl. Yos Sudarso 461 Gombong, Kebumen Jawa Tengah 54412, Indonesia
[2]Diploma 3 Pharmacy Study Program, Sekolah Tinggi Ilmu Kesehatan ISFI Banjarmasin, Jl. Flamboyan III No.7C, Sungai Miai, Banjarmasin Utara, Banjarmasin, Kalimantan Selatan 70123, Indonesia
*dwikifitri@unimugo.ac.id

ABSTRACT
Behçet's Syndrome is a rare multisystem inflammatory disorder with an etiology that is not yet fully understood. The disease is characterized by various clinical manifestations, including recurrent oral and genital ulcers, ocular lesions, skin abnormalities, arthritis, central nervous system involvement, and vascular complications. Understanding the genetic factors associated with Behçet's Syndrome across different populations may provide insights into its pathogenesis and potential therapeutic targets. This study aims to identify genetic variations associated with Behçet's Syndrome across populations from different continents using a bioinformatics approach. Data were analyzed from the GWAS Catalog, HaploReg v4.2, GTEx Portal, and Ensembl Genome. Single nucleotide polymorphisms (SNPs) associated with Behçet's Syndrome were identified, and their tissue-specific gene expression and allele frequency across populations were investigated. Two main SNPs, rs17482078 and rs2617170, were identified as missense mutations in the ERAP1 and KLRC4 genes, respectively, both of which have been linked to various autoimmune diseases. GTEx Portal analysis revealed that ERAP1 is expressed in 30 tissues, including skeletal muscle, blood, brain regions (putamen, caudate, hypothalamus), skin, pancreas, esophagus, and other organs. KLRC4 expression was observed in 18 tissues, including the cerebral cortex, lungs, heart, small intestine, thyroid, prostate, and additional tissues. Ensembl Genome data indicated population-specific allele frequency variation: rs17482078 in ERAP1 was highest in Africa (94.6%), followed by East Asia (94.2%), South Asia (93.5%), the Americas (88%), and Europe (78%). SNP rs2617170 showed the highest frequency in Africa (57%), followed by South Asia (46%), East Asia (45%), Europe (34%), and the Americas (33%). The findings highlight ERAP1 and KLRC4 as key genetic factors associated with Behçet's Syndrome and demonstrate population-specific allele frequency differences. These results provide valuable insights into the genetic architecture of the disease and may guide future studies on its pathogenesis and potential therapeutic strategies.

Keywords: autoimmunity; behçet syndrome; bioinformatics

How to Cite (in APA Style)
Fitri, D., Khuluq, H., Kurniasih, D., & Ma'ruf, M. (2026). Identification of Genomic Databases and Bioinformatics Analysis of Gene Variations in Behçet Syndrome Across Continental Populations. *Indonesian Journal of Global Health Research*, *8*(2), 859–866. https://doi.org/10.37287/ijghr.v8i2.751.

INTRODUCTION

Behçet disease (BD) or Behçet syndrome is classified as a systemic vasculitis due to the involvement of arterial and venous vessels in multiple organs (Sari & Setyawati, 2008). Behçet syndrome is a multisystem inflammatory disorder with an unknown etiology, characterized by recurrent episodes of oral aphthous ulcers, genital ulcers, other lesions, and ocular involvement. The disease affects nearly every system and organ, including the ocular, cardiovascular, gastrointestinal, renal, pulmonary, urological, central nervous system, and joints (Kokturk, 2012). The syndrome was first recognized in 1908 by Bluthe, who described the triad of iritis, mucocutaneous ulcers, and genital ulcers. In 1930, a Greek ophthalmologist, Benediktos Adamantiades, first reported patients with arthritis, oral and genital ulcers, phlebitis, and relapsing iritis with hypopyon. In 1937, Hulusi Behçet, a Turkish dermatologist, suspected a viral etiology and reported three patients with oral ulcers, genital ulcers, and hypopyon uveitis (Amarawardena et al., 2014). Behçet syndrome is also known as Adamantiades-Behçet's disease; however, the term Behçet syndrome was chosen by the

International Associations and Societies of Behçet (Tan et al., 2016).

Behçet disease is considered extremely rare worldwide but is endemic along the Silk Road from Asia to the Mediterranean (Sari & Setyawati, 2008). The syndrome occurs in countries bordering the Silk Road in East Asia, such as Japan, Korea, China, Iraq, and Turkey. The highest prevalence is observed in Middle Eastern countries, with Turkey reaching 370 per 100,000 population and Iran 80 per 100,000. Behçet syndrome typically begins between 30 and 40 years of age, with a nearly equal male-to-female ratio. BD is relatively rare globally but shows regional and local variation in high-incidence areas and immigrant communities, particularly affecting individuals of Southern European, North African, Middle Eastern, and East Asian descent. The prevalence is best documented in Turkey, reaching up to 400 per 100,000 population, approximately comparable to the estimated prevalence of systemic lupus erythematosus (SLE) and rheumatoid arthritis (RA) in the United States. The prevalence of BD is strongly correlated with the minor allele frequency of HLA-B*51 in the general population, which is highest among populations in Southern Europe, the Middle East, the Maghreb region of North Africa, and the Far East (Al-obeidi & Nowatzky, 2024).

Although studies collectively report a roughly equal male-to-female ratio, most patients with severe BD are male. The disease typically manifests in the third decade of life and generally subsides with age. BD can occur in children but usually not before the age of five and rarely before the age of ten. Early onset and male sex are strong predictors of severe disease (Senusi et al., 2015).

In the absence of specific laboratory diagnostic tests or histopathological findings, diagnosis relies on clinical criteria and often takes several years. Prognosis depends on clinical involvement, and the disease can lead to significant morbidity and mortality. Loss of visual acuity and neurological complications are the main causes of morbidity and disability (Kokturk, 2012). Research using Genome-Wide Association Studies (GWAS) specifically focusing on Behçet syndrome is limited. This study, by identifying gene variants through a bioinformatics approach, aims to determine the variants and expression patterns of genes across multiple tissues involved in the development of Behçet syndrome and to examine allele distribution across continental populations. The results are expected to provide valuable information for future research related to clinical treatment of Behçet syndrome from a genetic perspective.

METHOD

This study employed a bioinformatics-based approach to identify genomic variations that may contribute to the development of Behçet's Syndrome, following the methods described by (Ma'ruf et al., 2023) and (Puspitaningrum et al., 2022) The initial step involved searching the GWAS Catalog database (https://www.ebi.ac.uk/gwas/, accessed on May 15, 2025) to obtain Single Nucleotide Polymorphisms (SNPs) previously reported to be associated with Behçet's Disease (BD). The identified SNPs were then filtered based on statistical significance criteria, specifically a p-value $< 1 \times 10^{-8}$, ensuring that only strongly associated variants were included for further analysis. These selected variants were subsequently examined using HaploReg v4.2 (https://pubs.broadinstitute.org/mammals/haploreg/haploreg.php) to map the SNPs to their corresponding genes and to assess their potential impact on genomic regulatory elements. Gene expression analysis was then conducted through Expression Quantitative Trait Locus (eQTL) data using the GTEx Portal (https://www.gtexportal.org/home/, accessed on May 15, 2025) to identify expression changes across various human tissues that may be influenced by these SNPs. To understand the distribution of the identified variants across global populations, allele frequencies for European, African, American, East Asian, and South Asian populations were obtained from the Ensembl Genome database (https://www.ensembl.org/Homo_sapiens/, accessed on May 15, 2025). Together, these steps provide a comprehensive overview of genomic variations associated with BD, from the initial identification of risk SNPs to the characterization of their functional roles and their distribution across different populations worldwide.

RESULT

Based on the search results of SNP data using the GWAS Catalog for Behçet syndrome, a total of 91 SNPs were identified. Among these, 36 SNPs met the significance criterion of p-value < 1×10^{-8} (Table 1). These SNPs were subsequently analyzed using the HaploReg v4.2 database to identify validated missense variants. Two missense SNPs were identified: rs17482078 and rs2617170, corresponding to the ERAP1 and KLRC4 genes, respectively (Table 2). Gene expression profiles of ERAP1 and KLRC4 across various human tissues were examined using the GTEx Portal, which provides expression levels across multiple tissues (Figures 1 and 2). Furthermore, allele frequencies across global populations were determined using the Ensembl Genome database (Table 3; Figure 3).

Table 1.
Search results from the GWAS Catalog identified 91 SNPs, of which only 36 SNPs showed statistical significance (p-value < 1×10^{-8}).

No	SNP	p-value
1	rs9266490	2×10^{-44}
2	HLA-B*51	7×10^{-32}
3	rs4959053	2×10^{-20}
4	rs1518111	4×10^{-18}
5	rs1800871	1×10^{-14}
6	rs7616215	4×10^{-13}
7	rs76830965	3×10^{-12}
8	rs4947296	1×10^{-11}
9	rs1495965	2×10^{-11}
10	rs17482078	5×10^{-11}
11	rs17810546	1×10^{-10}
12	rs6660226	1×10^{-10}
13	rs9266406	2×10^{-10}
14	rs2848479	3×10^{-10}
15	rs3024490	3×10^{-10}
16	rs9357105	6×10^{-10}
17	rs2087726	9×10^{-10}
18	rs7574070	1×10^{-09}
19	rs2617170	1×10^{-09}
20	rs4896243	2×10^{-09}
21	rs17006292	5×10^{-09}
22	rs681343	5×10^{-09}
23	rs142459005	5×10^{-09}
24	rs897200	6×10^{-09}
25	rs4143322	6×10^{-09}
26	rs924080	7×10^{-09}
27	rs2121034	9×10^{-09}
28	rs74566205	1×10^{-08}
29	rs1874886	2×10^{-08}
30	rs1660760	3×10^{-08}
31	rs12220700	3×10^{-08}
32	rs7572482; rs7574070; rs897200	4×10^{-08}
33	rs17482078	4×10^{-08}
34	rs41268928	5×10^{-08}
35	rs2181036	8×10^{-08}
36	rs12656327	9×10^{-08}

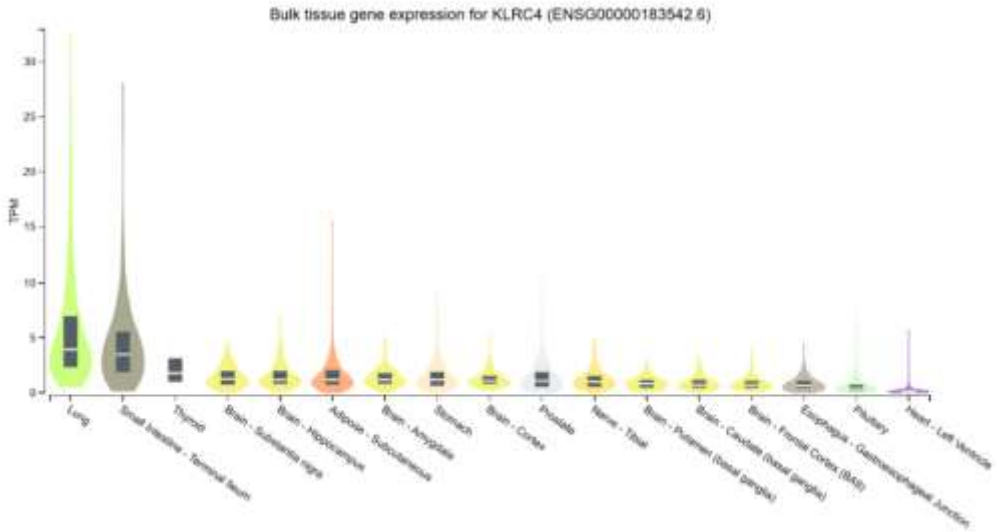

Figure 1. *Gene expression of ERAP1 associated with Behçet syndrome susceptibility across various human tissues based on the GTEx Portal database.*

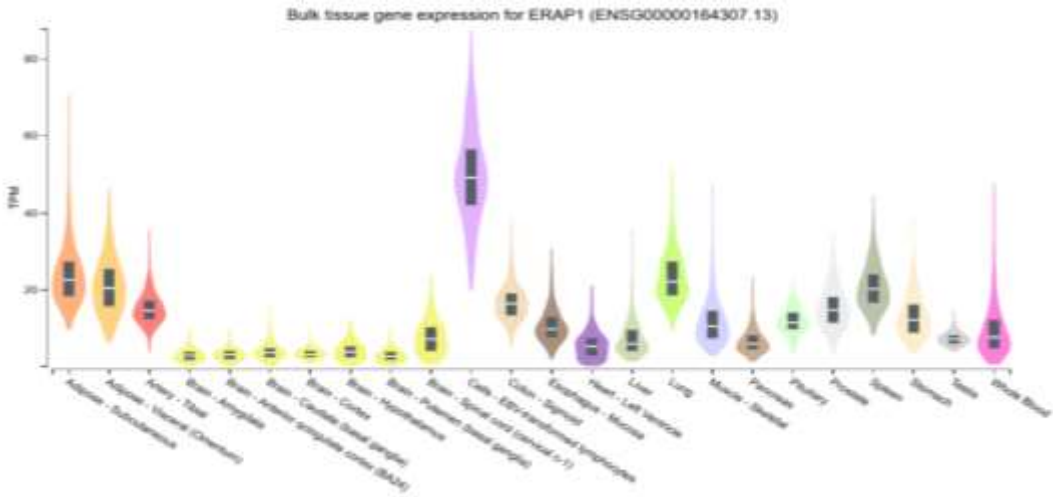

Figure 2. *Gene expression of KLRC4 associated with Behçet syndrome susceptibility across various human tissues based on the GTEx Portal database.*

Table 2.
Allele Frequency Across Continental Populations

SNP	Post (hg38)	Gen	p-value	Alel		Frekuensi Alel (N)				
				Ref	Alt	AFR	AMR	EAS	EUR	SAS
rs17482078	96783162	ERAP1	5×10^{-11}	C	T	0.946 (1251)	0.876 (608)	0.942 (950)	0.776 (781)	0.935 (914)
rs2617170	10408358	KLRC4	1×10^{-9}	T	C	0.567 (750)	0.333 (231)	0.450 (454)	0.337 (339)	0.456 (446)

DISCUSSION

Bioinformatics analysis of the genomic data used in this study revealed several genetic variations associated with Behçet syndrome, showing distinct distribution patterns among different populations. This study utilized a bioinformatics-based approach by extracting data from several genomic databases. The GWAS Catalog was used with a significance threshold of p-value $< 1 \times 10^{-8}$, which serves as a standard criterion to ensure high confidence in the results and to confirm that the identified SNPs are truly and significantly correlated with Behçet syndrome. This threshold also minimizes misinterpretation and reduces false-positive associations between SNPs and their phenotypes (Kim et al., 2021). Single Nucleotide Polymorphisms (SNPs) are important genetic markers used to identify disease-related abnormalities (Study et al., 2023)

In this study, two missense SNPs—rs17482078 and rs2617170—were identified from the GWAS Catalog and validated through HaploReg v4.2, corresponding to the ERAP1 and KLRC4 genes, respectively. Missense SNPs are known to potentially alter protein function and contribute to disease mechanisms. Although some may have mild effects, they can still influence complex disease phenotypes (Astuti et al., 2023). These findings are consistent with previous research identifying that genetic variations in ERAP1 and KLRC4 are associated with an increased risk of Behçet disease. BD is considered an autoimmune disorder belonging to the group of seronegative vasculitides, involving blood vessels of varying calibres (Pérez García, 2021).

The ERAP1 (Endoplasmic Reticulum Aminopeptidase 1) gene, located on chromosome 5q15, encodes an aminopeptidase protein with ubiquitous tissue distribution. This enzyme is responsible for trimming N-terminal peptides in the endoplasmic reticulum, an essential step in optimizing peptide length for MHC-I binding (Saric et al., 2002). Additionally, ERAP1 is involved in the cleavage of several proinflammatory cytokine receptors from the cell membrane, including the tumor necrosis factor receptor (TNFR1), IL-1RII, and IL-6α receptors. ERAP1 is therefore a strong biological candidate for BD association and has also been linked to other HLA class I–mediated immune diseases, such as ankylosing spondylitis (AS) and psoriasis (Burton et al., 2007). The KLRC4 (Killer Cell Lectin-Like Receptor C4) gene contributes to Behçet syndrome pathogenesis through the regulation of innate immune responses, particularly natural killer (NK) cell activity. Specific genetic variants in KLRC4 may cause excessive inflammatory activation, leading to increased risk of vasculitis and clinical manifestations of BD.

The Genotype-Tissue Expression (GTEx) database was used to identify genotype and gene expression profiles across human and mammalian tissues. GTEx provides tools for querying genes, variants, and tissues through comprehensive genomic and molecular phenotype datasets (Stanfill & Cao, 2021). eQTL analysis revealed functional consequences of genetic variations, demonstrating that ERAP1 and KLRC4 are expressed in multiple human tissues. Based on the bulk tissue gene expression data from the GTEx Portal, ERAP1 showed the highest relative expression in subcutaneous adipose tissue, tibial artery, EBV-transformed lymphoblastoid cells, and the colon. High expression in subcutaneous adipose tissue and arterial tissue aligns with the role of ERAP1 in vascular inflammation and connective tissue—primary targets in Behçet's systemic vasculitis. Meanwhile, expression in EBV-transformed lymphoid cells (LCLs) reflects ERAP1's activity in immune cells, supporting its role in antigen presentation with HLA-B*51, a major genetic factor in this disease (Hammam et al., 2021), Significant expression in the colon also indicates its relevance to gastrointestinal mucosal involvement, a clinical manifestation occasionally observed in Behçet's disease.

Similarly, KLRC4 expression analysis from bulk RNA GTEx data showed a distinct expression pattern in tissues relevant to mucosal immunity and inflammatory regulation. The highest expression levels were observed in the lung, small intestine, thyroid, and brain, all of which have significant local immunological activity. High expression in the lung and small intestine supports KLRC4's role in modulating mucosal immune responses to environmental antigens, while expression in the thyroid and brain may relate to NK cell–mediated immune homeostasis within these organs. This distribution pattern supports the concept that Behçet syndrome pathogenesis involves innate immune dysregulation, particularly through NK cell activity regulated by the killer cell lectin-like receptor (KLR) gene cluster, including KLRC4.

According to Table 2, allele frequencies associated with Behçet syndrome varied among populations from Africa, America, Europe, East Asia, and South Asia, extracted from the 1000 Genomes Project using the Ensembl Genome database. The purpose of this mapping was to visualize the distribution of genetic susceptibility across global populations. The allele frequency of rs17482078 (ERAP1) was highest in Africa (94.6%), followed by East Asia (94.2%), South Asia (93.5%), America (88%), and Europe (78%). Meanwhile, rs2617170 showed the highest frequency in Africa (57%), followed by South Asia (46%), East Asia (45%), Europe (34%), and America (33%). These variations likely reflect the influence of environmental, genetic, and immunological factors across populations (Shallis et al., 2018).

CONCLUSION

This study employed a bioinformatics-based approach to investigate Behçet's syndrome using several genomic databases, including the GWAS Catalog, HaploReg V4.2, GTEx Portal, and Ensembl Genome. Based on the analysis, two missense-validated variants were identified: ERAP1 and KLRC4, corresponding to SNPs rs17482078 and rs2617170, respectively. Gene expression analysis from the GTEx Portal revealed that ERAP1 and KLRC4 were expressed in 30 and 18 human tissues, including skeletal muscle, blood, brain, oesophagus, skin, stomach, pancreas, colon, subcutaneous adipose tissue, lung, lymph node, heart, liver, testis, artery, prostate, and tibial nerve. The allele frequency of rs17482078 was found to be highest in the African population (94.6%), followed by East Asia (94.2%), South Asia (93.5%), America (88%), and Europe (78%). Meanwhile, rs2617170 showed the highest distribution in Africa (57%), followed by South Asia (46%), East Asia (45%), Europe (34%), and America (33%). A higher allele frequency in a given region may indicate an increased susceptibility to Behçet's syndrome within that population. For future research, it is recommended to utilise additional bioinformatics resources such as the Weighted Gene Co-expression Network Analysis (WGCNA) and the Gene Expression Omnibus (GEO) database to further identify genes associated with Behçet's syndrome.

REFERENCES

Al-obeidi, A. F., & Nowatzky, J. (2024). IMMUNOPATHOGENESIS OF BEHÇET'S DISEASE. 1–19. https://doi.org/10.1016/j.clim.2023.109661.IMMUNOPATHOGENESIS

Amarawardena, W. K. M. G., Wijesundere, A., Muhandiram, W. M. T., & Appuhamy, H. S. D. (2014). An unusual presentation of Behçet's disease. The Ceylon Medical Journal, 59(4), 144–145. https://doi.org/10.4038/cmj.v59i4.7870

Astuti, F., Capritasari, R., Iskardyani, D., Alifiar, I., & Irham, L. M. (2023). Identifikasi Variasi Dan Ekspresi Gen Pada Stroke Iskemik Dengan Pendekatan Bioinformatika. Jurnal Kesehatan Bakti Tunas Husada: Jurnal Ilmu-Ilmu Keperawatan, Analis Kesehatan Dan Farmasi, 23(1), 9–19. https://doi.org/10.36465/jkbth.v23i1.1302

Burton, P. R., Clayton, D. G., Cardon, L. R., Craddock, N., Deloukas, P., Duncanson, A., Kwiatkowski, D. P., McCarthy, M. I., Ouwehand, W. H., Samani, N. J., Todd, J. A., Donnelly, P., Barrett, J. C., Davison, D., Easton, D., Evans, D. M., Leung, H. T., Marchini, J. L., Morris, A. P., … Weisman, M. H. (2007). Association scan of 14,500 nonsynonymous SNPs in four diseases identifies autoimmunity variants. Nature Genetics, 39(11), 1329–1337. https://doi.org/10.1038/ng.2007.17

Hammam, N., Li, J., Evans, M., Kay, J. L., Izadi, Z., Anastasiou, C., Gianfrancesco, M. A., Yazdany, J., & Schmajuk, G. (2021). Epidemiology and treatment of Behçet's disease in the USA: insights from the Rheumatology Informatics System for Effectiveness (RISE) Registry with a comparison with other published cohorts from endemic regions. Arthritis Research and Therapy, 23(1), 1–9. https://doi.org/10.1186/s13075-021-02615-7

Kokturk, A. (2012). Clinical and pathological manifestations with differential diagnosis in Behçet's disease. Pathology Research International, 2012. https://doi.org/10.1155/2012/690390

Ma'ruf, M., Irham, L. M., Adikusuma, W., Sarasmita, M. A., Khairi, S., Purwanto, B. D., Chong, R., Mazaya, M., & Siswanto, L. M. H. (2023). A genomic and bioinformatic-based approach to identify genetic variants for liver cancer across multiple continents. Genomics and Informatics, 21(4), 1–8. https://doi.org/10.5808/gi.23067

Pérez García, R. A. (2021). "Oral Manifestations of Behcet's Syndrome: Case Report". Biomedical Journal of Scientific & Technical Research, 39(1), 31079–31084. https://doi.org/10.26717/bjstr.2021.39.006259

Puspitaningrum, A. N., Perwitasari, D. A., Adikusuma, W., Djalilah, G. N., Dania, H., Maliza, R., Faridah, I. N., Sarasmita, M. A., Rezadhini, M., Cheung, R., & Irham, L. M. (2022). Integration of genomic databases and bioinformatic approach to identify genomic variants for sjogren's syndrome on multiple continents. Media Farmasi: Jurnal Ilmu Farmasi, 19(2), 71. https://doi.org/10.12928/mf.v19i2.23706

Sari, L. M., & Setyawati, T. (2008). Manifestasi Behçet Disease yang Parah dan Komplikasi Perawatannya dalam Rongga Mulut. Journal of Dentistry Indonesia, 15(2), 111–120. https://doi.org/10.14693/jdi.v15i2.67

Saric, T., Chang, S. C., Hattori, A., York, I. A., Markant, S., Rock, K. L., Tsujimoto, M., & Goldberg, A. L. (2002). An IFN-γ-induced aminopeptidase in the ER, ERAP I, trims precursors to MHC class I-presented peptides. Nature Immunology, 3(12), 1169–1176. https://doi.org/10.1038/ni859

Senusi, A., Seoudi, N., Bergmeier, L. A., & Fortune, F. (2015). Genital ulcer severity score and genital health quality of life in Behçet's disease Rare systemic diseases. Orphanet Journal of Rare Diseases, 10(1), 1–11. https://doi.org/10.1186/s13023-015-0341-7

Shallis, R. M., Ahmad, R., & Zeidan, A. M. (2018). Aplastic anemia: Etiology, molecular pathogenesis, and emerging concepts. European Journal of Haematology, 101(6), 711–720. https://doi.org/10.1111/ejh.13153

Stanfill, A. G., & Cao, X. (2021). Enhancing Research Through the Use of the Genotype-Tissue Expression (GTEx) Database. Biological Research for Nursing, 23(3), 533–540. https://doi.org/10.1177/1099800421994186

Study, O., Krishnamurthy, H. K., Balaguru, U. M., & Rajasekaran, J. J. (2023). Influence of genetic polymorphisms on serum biomarkers of cardiac health. 23(January).

Tan, S., Gunawan, L., & Reginata, G. (2016). Manifestasi klinis sindrom behcet. Cermin Dunia Kedokteran, 43(10), 756–759.