Mapping Rheumatoid Arthritis Susceptibility through Integrative Bioinformatics and Genomics

by Universitas Ahmad Dahlan Yogyakarta 28

Submission date: 11-Jan-2024 11:27AM (UTC+0700)
Submission ID: 2259452164
File name: 3.Nining_15_23.pdf (1,019.96K)
Word count: 4663
Character count: 26304

Media Farmasi: Jurnal Ilmu Farmasi (Journal of Pharmaceutical Science) e-ISSN 2503-5223
Vol. 20, No. 1, March 2023, pp. 15-23
http://journal.uad.ac.id/index.php/Media-Farmasi/index

Mapping Rheumatoid Arthritis Susceptibility through Integrative Bioinformatics and Genomics

Nining Medi Sushanti[1], Wirawan Adikusuma[2,3], Arief Rahman Afief[1], Anita Silas La'ah[4], Firdayani[3], Rockie Chong[5], Zainul Amiruddin Zakaria[7], Barkah Djaka Purwanto[8,9], Rahmat Dani Satria[10,11], Riat El Khair[10,11], Abdi Wira Septama[6], Lalu Muhammad Irham[1,6]*

[1]Faculty of Pharmacy, Universitas of Ahmad Dahlan, Yogyakarta, 55164, Indonesia
[2]Departement of Pharmacy, University of Muhammadiyah Mataram, Mataram, 83127, Indonesia
[3]Research Center for Vaccine and Drugs, National Research and Innovation Agency (BRIN), South Tangerang, Indonesia.
[4]Taiwan International Graduate Program in Molecular Medicine, National Yang Ming Chiao Tung University and Academia Sinica, Taipei, Taiwan
[5]Department of Chemistry and Biochemistry, University of California, Los Angeles, USA
[6]Traditional Medicine, National Research and Innovation Agency (BRIN), South Tangerang, Indonesia
[7]Borneo Research on Algesia, Inflammation and Neurodegeneration (BRAIN) Group, Department of Biomedical Sciences, Faculty of Medicines and Health Sciences, University Malaysia Sabah, Malaysia
[8]Faculty of Medicine, University of Ahmad Dahlan, Yogyakarta, 55191, Indonesia
[9]PKU Muhammadiyah Bantul Hospital, Bantul, Yogyakarta, 55711, Indonesia
[10]Department of Clinical Pathology and Laboratory Medicine, Faculty of Medicine, Public Health and Nursing, Universitas Gadjah Mada, Yogyakarta 55281, Indonesia
[11]Clinical Laboratory Installation, Dr. Sardjito Central General Hospital, Yogyakarta 55281, Indonesia
*Correspondence author: e-mail: lalu.irham@pharm.uad.ac.id

ARTICLE INFO

Article history
Received: 25-09-2022
Revised: 19-10-2022
Accepted: 17-01-2023

Keywords

Rheumatoid arthritis
Autoimmune
Genome
Bioinformatics
Genetic variation

ABSTRACT

Rheumatoid arthritis (RA) is an autoimmune disease that influences several organs and tissues, especially the synovial joints, and is associated with multiple genetic and environmental factors. Numerous databases provide information on the relationship between a specific gene and the disease pathogenesis. However, it is important to further prioritize biological risk genes for downstream development and validation. This study aims to map RA-association genetic variation using genome-wide association study (GWAS) databases and prioritize influential genes in RA pathogenesis based on functional annotations. These functional annotations include missense/nonsense mutations, cis-expression quantitative trait locus (cis-eQTL), overlap knockout mouse phenotype (KMP), protein-protein interaction (PPI), molecular pathway analysis (MPA), and primary immunodeficiency (PID). 119 genetic variants mapped had a potential high risk for RA based on functional scoring. The top eight risk genes of RA are *TYK2* and *IFNGR2*, followed by *TNFRSF1A, IL12RB1* and *CD40, C5, NCF2*, and *IL6R*. These candidate genes are potential biomarkers for RA that can aid drug discovery and disease diagnosis.

This is an open access article under the CC–BY-SA license.

1. Introduction

Rheumatoid arthritis (RA) is the most common autoimmune disease in the world. It is an uncontrollable chronic inflammatory disease with various multi-system modalities followed by the proliferation of the synovial tissue that causes pain, joint erosion, and functional disorder (Garner et al., 2014). This synovial tissue proliferation can cause chronic impairments (Song & Lin, 2017). One sign of RA is persistent inflammation in the synovial joints. RA prevalence is within the age range of 40 – 60 which is between 0.5% and 1%, with greater prevalence in women than in men (Song & Lin, 2017). The main risk factors of RA are gender, family history, advanced age, exposure to

silicate, and smoking cigarette (Garner et al., 2014). The primary influential determinants for RA vulnerability are genetic and environmental factors (Deane et al., 2017).

In this study, Genome-Wide Association Study (GWAS) is used to identify genetic variants that are at risk for disease pathogenesis by applying functional annotation criteria. This can also be useful in medicine through the bioinformatics-based approach (Narayanan, 2016 ; Reay & Cairns, 2021). GWAS is a research approach used in finding the genetic variation associated with certain diseases, this involves the entire genome of several individuals. GWAS has been widely developed as a screening method not only in early-stage patients but also for people who are at risk based on family history. This research approach is useful for finding Single Nucleotide Polymorphisms (SNPs) that contribute to complex diseases such as cancer, infectious diseases (AIDS, leprosy, hepatitis), autoimmune diseases, neuropsychiatric, and various diseases (Fareed & Afzal, 2013).

Meta-Analysis and GWAS have identified more than 100 RA loci, and GWAS is known as a database with genetic architecture also for RA (Kwon et al., 2020). GWAS is recognized as a powerful approach to map genes responsible for various diseases (Hettiarachchi & Komar, 2022). For example, earlier studies have analyzed Chinese population SNPs that affect RA by using GWAS (Ma et al., 2013). In addition, the whole genome case-control of the linkage disequilibrium (LD) mapping was conducted by utilizing SNPs and RA-associated polymorphisms in *PADI4* and *SLC22A4/A5* cluster was identified (Too et al., 2012). These findings showed that *PADI4* is a risk gene for RA while *SLC22A4/A5* has several polymorphisms associated with many autoimmune diseases. Thus, large-scale LD mapping seems effective at identifying RA polymorphisms (Ren et al., 2014).

Information on the relationship between a given gene and the pathogenesis of a disease is largely available on the GWAS database. The priority given to the most influential genes (biological risk genes) is scarcely examined. Thus, this research was conducted based on six functional annotations resulting in the prioritization of the most influential genes in RA pathogenesis. Biomarker research arises based on the need to understand the pathogenesis that causes RA (Gavrilă et al., 2016). Hence, data from GWAS could help us to identify novel RA-associated genes as biomarkers and provide new insight into managing and treating RA. The search for new biomarkers for RA plays an important role in identifying candidate genes at risk for RA.

2. Materials and Methods

2.1. RA-associated genes

Genomic variants of RA were collected from the GWAS database, and the most influential genes were prioritized based on six functional annotations (Figure 1). The variants related to RA were expanded using the criterion $r^2 > 0.8$ of HaploReg v4.1.The SNPs associated with RA were denoted as "RA-associated genes". The genomic data were then prioritized based on six functional annotations, where genes with scores ≥ 2 were classified as "biological RA risk genes". HaploReg v4.1 was used to dictate RA-associated SNPs (Ward & Kellis, 2016). The variants were expanded on high LD ($r^2 > 0.8$) in an Asian population, based on the 1000 Genome Project. We used the LD criterion as $r^2 > 0.8$, based on Okada et al, studies (Okada et al., 2014).

2.2. Functional annotations of RA-associated genes

Candidate genes most likely to be RA targets were prioritized based on their biological function using six functional annotations. Missense or nonsense mutations are single amino acid changes that produce different protein functions and are selected as functional annotation criteria (Zhang et al., 2012). In accordance with HaploReg v4.1, the first functional annotation is missense or nonsense mutation because it contains functional annotations from the SNPs database (DB) (Leng et al., 2020). HaploReg v4.1 was used to link cis-expression quantitative trait loci (cis-eQTL) to genetic variants in blood target tissues (Ward & Kellis, 2016). We also used WebGestalt 2019 to understand the relationship between mutant genes and phenotypes (Wang et al., 2017). Phenotype information of mouse and other mammals was derived from the ontology of the Mammalian Phenotype (MP) (Smith & Eppig, 2012). To further understand the phenotypes of mouse, we used BioMart to convert geneswith human Ensembl ID into mouse Ensembl ID (Drost & Paszkowski, 2017). Genes with False Discovery Rates (FDRs) in mouse phenotypes <0.05 were considered to be significant results. RA-associated genes from biological process networks were identified using protein-protein

interactions (PPIs) (Qiu et al.,2021). In addition, we used WebGestalt 2019 to conduct reinforcement analysis and to investigate whether the genes were collected on certain functional annotations (Liao et al., 2019). Kyoto Encyclopedia of Genes and Genomes (KEGG) was used to determine what molecular pathways were gained from the list of RA-associated genes and what genes were involved. Further analysis of the biochemical pathway was carried out by WebGestalt 2019 database. The last annotation criterion was primary immunodeficiency (PID). PID refers to inborn immunity diseases and is often reported to be associated with cancer (Mortaz et al., 2016). PID attacks the immune system and causes inflammation in RA (Dimitriades & Sorensen, 2016). The hypergeometric test was employed for reinforcement analysis of these data, with the significance criterion being p-value < 0.05.

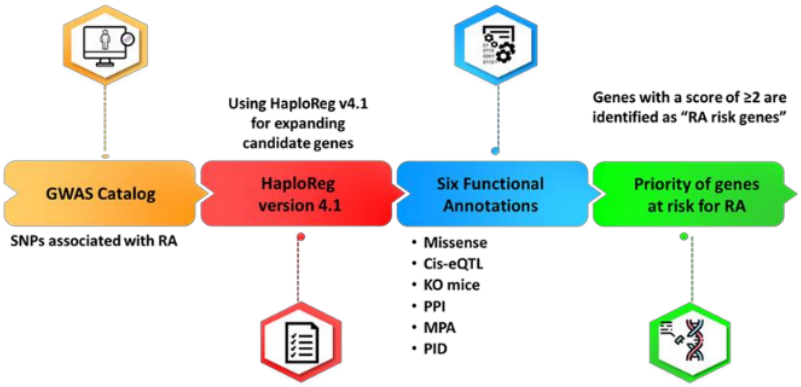

Fig 1. Schematic model showing how genome-based information can be integrated to map susceptibility genes for RA [Missense/nonsense; cis-expression quantitative trait locus (cis-eQTL); protein-protein interaction (PPI); molecular pathway analysis (MPA); and primary immunodeficiency (PID)]

3. Results and Discussion

3.1. Prioritized SNPs potentially cause RA

In this study we obtained 2,732 RA-associated SNPs which were dictated in the GWAS database (Table S1) and 2012 SNPs were obtained after filtering based on p-value < 10^{-8}. Importantly, 396 SNPs were obtained after checking for duplicates, and further filtering based on the inclusion criterion OR >2, led to a total of 360 RA-associated SNPs (Table S2). The SNPs collected from GWAS were expanded using HaploReg V4.1 while the RA- associated genes were prioritized based on six functional annotations, namely missense/nonsense mutation, cis-eQTL, knockout mouse phenotype (KMP), protein-protein interaction (PPI), molecular pathway analysis (MPA), and primary immunodeficiency (PID).

3.2. RA risk SNPs Functional Annotations

This study was designed to prioritize the best candidate genes by using six functional annotations with a scoring system. In the functional annotation process, candidate genes with higher scores represent greater biological influences on RA pathogenesis (Figure 2). Following the six annotations,we discovered 39 SNPs of missense/nonsense mutation variants from the total 360 SNPs (Figure 3).We further examined the cis-eQTL effect using HaploReg v4.1, 101 SNPs were gained from the total360 SNPs (Figure 3). Phenotype data was taken from the MP ontology which contains information about the phenotype of mice and other mammals. WebGestalt 2019 was applied to perform over- representation analysis (ORA) and we found 102 genes overlapping with RA risk genes (FDR <0.05).Gene Ontology (GO) annotations derived from WebGestalt 2019 were used to evaluate PPIs. Following the analysis on PPIs, 122 genes overlap with other RA risk genes (FDR < 0.05). KEGG was used to perform ORA on the molecular pathway. Thus, over 61 RA-

associated genes on KEGG pathways were identified. We also used the PID data from The International Union of Immunological Societies (IUIS) to analyze and confirm the overlapping genes. Sixteen overlapping genes were statistically significant ($p < 0.05$) from PID analysis.

We also scored the risk genes from 0 to 6 and obtained 156 genes with a score of 0. We noted that 85 genes havea score of 1, 45 genes have a score of 2, 40 genes have a score of 3, 26 genes have a score of 4, 6 genes have a score of 5, and 2 genes have a score of 6 (Figure 4). In total, we obtained 119 genes witha score of ≥ 2, that were categorized as 'biological RA risk genes' (Table S2). Using the aforementioned scoring criteria, the top eight RA risk genes are: *Tyrosine kinase 2* (*TYK2*), followed by *interferon-gamma receptor 2* (*IFNGR2*). Members of the *tumor necrosis factor receptor superfamily 1A* (*TNFRSF1A*), *interleukin receptor subunit 12 beta 1* (*IL12RB1*) and *cluster of differentiation 40* (*CD40*), *complement 5* (*C5*), *neutrophil cytosolic factor 2* (*NCF2*), and *interleukinreceptor 6* (*IL6R*).

GENCODE_id	GENCODE_name	missense	cis-eQTL	KO mice	PPI	KEGG	PID	Total Score
ENSG00000105397	TYK2	1	1	1	1	1	1	6
ENSG00000159128	IFNGR2	1	1	1	1	1	1	6
ENSG00000067182	TNFRSF1A	0	1	1	1	1	1	5
ENSG00000096996	IL12RB1	0	1	1	1	1	1	5
ENSG00000101017	CD40	0	1	1	1	1	1	5
ENSG00000106804	C5	0	1	1	1	1	1	5
ENSG00000116701	NCF2	1	0	1	1	1	1	5
ENSG00000160712	IL6R	1	1	1	1	1	0	5
ENSG00000003400	CASP10	0	1	0	1	1	1	4
ENSG00000020633	RUNX3	1	0	1	1	1	0	4
ENSG00000056558	TRAF1	0	1	1	1	1	0	4
ENSG00000064012	CASP8	0	0	1	1	1	1	4
ENSG00000076662	ICAM3	1	1	0	1	1	0	4
ENSG00000081237	PTPRC	0	0	1	1	1	1	4
ENSG00000081985	IL12RB2	0	1	1	1	1	0	4
ENSG00000100906	NFKBIA	0	0	1	1	1	1	4
ENSG00000103313	MEFV	1	0	1	1	0	1	4
ENSG00000103811	CTSH	1	1	0	1	1	0	4
ENSG00000110448	CD5	1	0	1	1	1	0	4
ENSG00000112486	CCR6	0	1	1	1	1	0	4
ENSG00000115267	IFIH1	1	0	1	1	1	0	4
ENSG00000115604	IL18R1	0	1	1	1	1	0	4
ENSG00000117560	FASLG	0	0	1	1	1	1	4
ENSG00000118503	TNFAIP3	1	0	1	1	1	0	4

Fig. 2. Functional annotations for rheumatoid arthritis (RA) with scores ≥ 2 based on defined scoring criteria

Tyrosine kinase 2 (*TYK2*) is a member of the *Janus kinase* (*JAK*) family of transforming growth factors and cytokines. The *TYK2* loss analysis function reveals its importance toward immunity against infections, inflammation (automatic), and immunity (automatic). Many genome-wide association study (GWAS) in humans propose relationships between *TYK2* genetic variants and many autoimmune diseases, inflammatory diseases, and tumors. Hence, *TYK2* emerges as a target of interest for therapeutic interventions (Leitner et al., 2017). The human *interferon-gamma receptor* is a multimer of *IFN-γR1* (coded by *IFNGR1*) and *IFN-γR2* (two chains). *Interferon-gamma receptor 2* (*IFN-γR2*) is a gene containing a protein and in humans, this gene is coded by the *IFNGR2* gene. The *IFNGR2* gene codes for beta chains that are not bound to the ligands of the *IFN-γR* (Kong et al., 2013).

Tumor necrosis factor receptor 1A (*TNFRSF1A*) is a *CD120a* which is a member of *the superfamily tumor necrosis factor alpha*-binding membrane receptor. The *TNFRSF1A* gene instructs the making of *tumor necrosis factor receptor 1* (*TNFR1*) protein (Chen et al., 2015). This protein is found stretching on the cell membrane, both extracellularly and intracellularly (Chen et al., 2015). Extracellularly, the *TNFR1* protein enchains *tumor necrosis factor (TNF)*. The *TNF* and the *TNFR1* protein interaction catalyzes the latter to chain to two other *TNFR1* proteins, constituting trimer, a three-protein complex (Wajant & Siegmund, 2019). This formation of the trimer is important for the *TNFR1* protein to function. The chains of the protein of *TNF* and *TNFR1* prompt the latter to deliver a signalinto the cell (Li et al., 2013). The *TNFR1* protein signal can stimulate inflammation or cell self- destruction (apoptosis). The signal into the cell starts a routeway activating the nuclear factor kappaB protein triggering inflammation and directing to the cytokine production, the immune system protein (Chen et al., 2015).

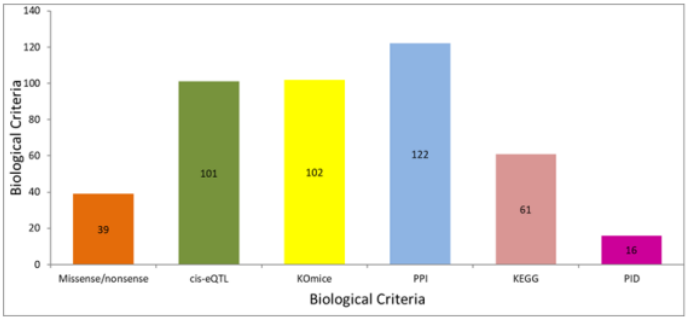

Fig. 3. Number of genes for each functional annotation that is prioritized for RA [Missense/nonsense; cis-expression quantitative trait locus (cis-eQTL); Knockout mice (KOmice); protein- protein interaction (PPI); Kyoto Encyclopedia of Genes and Genomes (KEGG); and primary immunodeficiency (PID)]

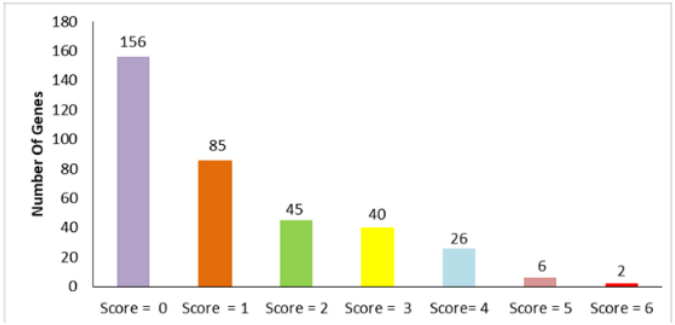

Fig. 4. Gene frequency distribution across defined functional scores for RA, showing 156 genes with a score of 0.85, genes with a score of 1, and 119 (45 + 40 + 26 + 6 + 2) genes with a total score of ≥ 2 are classified as RA risk genes

Interleukin-12 receptor beta-1 subunit (*IL12RB1*) functions as an interleukin receptor that binds to *interleukin-12* with low affinity which is involved in *IL12* transduction (Floss et al., 2016). *IL12RB2* makes a functional high-affinity receptor for *IL12* (Chen et al., 2015). *IL23R* forms the receptor of *interleukin-23* functioning in *IL23* signal transduction via *JAK/STAT* signaling cascade activation (Pastor-Fernández et al., 2020). *D40* is a member of the *tumor necrosis factor* (*TNF*) superfamily found in antigen-presenting cells, including dendritic cells (DC), B cells, monocytes, and macrophages, which play an important role in activating the immune system (Laha et al., 2021). The expression expands to various non-hematopoietic cells, including nerve cells, epithelial cells, endothelial cells, and fibroblasts (Karnell et al., 2019). Typical genetic variants at the *TRAF1-C5* locus on the chromosome correlate with an increased risk of anti-ccp-positive RA (Huang et al., 2019).

Genetic variants *rs7021206* and *rs3761847* at the *TRAF1-C5* locus are associated with RA in the Han Chinese population, indicating that *TRAF1-C5* may play an important role in the development RA thereby reducing the pathogenic role of *TRAF1-C5* in ethnic variations (van Steenbergen et al., 2015). *NCF2* is carried to the membrane of the cell to be combined with other components to build aNOX system that is active through microbial stimulation. In other studies, the polymorphism *NCF2rs10911363* was linked to SLE risk in Sweden and US populations. The variant of *NCF2 rs789181* was discovered to have a relationship with mild RA in men (T.-P. Zhang et al., 2020). *IL6* is an important cytokine in mediating inflammation and RA systemic overview, including synovitis, fatigue, anemia, anorexia, and osteoporosis (Narazaki et al., 2017).

In this study, we used the GWAS database to obtain the genomic variants for RA. The GWAS database is easily accessible and effective because it is diverse and structured. The GWAS method effectively identifies genetic loci associated with the pathogenesis of a disease, including prioritizing candidate loci, exploring disease mechanisms, predicting disease risk, identifying new drug targets, and gathering information about the desired trait or population (Caliskan et al., 2021). However, the GWAS database has drawbacks in that detailed phenotypic descriptions of the study's ethnic background are not provided by many GWAS papers, and even nowadays, there is no standard way to make an ethnicity report (Hart & Kranzler, 2015). The prospect for further research is to develop a wider use of functional annotation criteria to get more results to determine genetic variations that affect RA pathogenesis and identify drug candidates that target biological genes for RA risk. The impact of this research is to find out the genetic variation that influences the pathogenesis of RA and gain an understanding by utilizing genetic variation that can be used for drug development by using a bioinformatics approach.

4. Conclusion

Our research showed that genomic information is highly important for mapping RA genomic variants based on defined functional annotations from this study. Two candidate genes, *TYK2* and *IFNGR2*, were obtained based on defined scoring criteria, and we propose these genes as potential RA biomarkers to be prioritized for downstream clinical development and validation.

Author Contributions: Nining Medi Sushanti, Lalu Muhammad Irham, and Lalu Muhammad Irham conceived and designed the study. Nining Medi Sushanti and Lalu Muhammad Irham performed the computational analysis. Nining Medi Sushanti wrote the manuscript. Nining Medi Sushanti, Lalu Muhammad Irham., Wirawan Adikusuma, Anita Silas La'ah, Arief Rahman Afief, Zainul Amiruddin Zakaria, Firdayani, Rahmat Dani Satria, Riat El Khair, Abdi Wira Septama and Rickie Chong. revised the manuscript. Lalu Muhammad Irham supervised and coordinated this study. All authors have read and approved the manuscript and have made significant contributions to this study.

Funding

No funding was provided for this research.

Competing Interests

The authors have not declared a conflict of interest for this article.

Acknowledgment

I would like to thanks to my alma mater Universitas Ahmad Dahlan as always support my study as well as my school SMK Semesta Bumiayu for giving me opportunity to accomplish my master degree.

References

Caliskan, M., Brown, C. D., & Maranville, J. C. (2021). A catalog of GWAS fine-mapping efforts in autoimmune disease. *American Journal of Human Genetics*, *108*(4), 549–563. https://doi.org/10.1016/j.ajhg.2021.03.009.

Chen, W., Xu, H., Wang, X., Gu, J., Xiong, H., & Shi, Y. (2015). The tumor necrosis factor receptor superfamily member 1B polymorphisms predict response to anti-TNF therapy in patients with autoimmune disease: A meta-analysis. *International Immunopharmacology*, *28*(1), 146–153. https://doi.org/10.1016/j.intimp.2015.05.049.

Deane, K. D., Demoruelle, M. K., Kelmenson, L. B., Kuhn, K. A., Norris, J. M., & Holers, V. M. (2017). Genetic and environmental risk factors for rheumatoid arthritis. *Best Practice & Research. Clinical Rheumatology*, *31*(1), 3–18. https://doi.org/10.1016/j.berh.2017.08.003.

Dimitriades, V. R., & Sorensen, R. (2016). Rheumatologic manifestations of primary

immunodeficiency diseases. *Clinical Rheumatology*, *35*(4), 843–850. https://doi.org/10.1007/s10067-016-3229-6.

Drost, H. G., & Paszkowski, J. (2017). Biomartr: Genomic data retrieval with R. *Bioinformatics*, *33*(8), 1216–1217. https://doi.org/10.1093/bioinformatics/btw821.

Fareed, M., & Afzal, M. (2013). Single nucleotide polymorphism in genome-wide association of human population: A tool for broad spectrum service. *Egyptian Journal of Medical Human Genetics*, *14*(2), 123–134. https://doi.org/10.1016/j.ejmhg.2012.08.001.

Floss, D. M., Klöcker, T., Schröder, J., Lamertz, L., Mrotzek, S., Strobl, B., Hermanns, H., & Scheller, J. (2016). Defining the functional binding sites of interleukin 12 receptor β1 and interleukin 23 receptor to Janus kinases. *Molecular Biology of the Cell*, *27*(14), 2301–2316. https://doi.org/10.1091/mbc.E14-12-1645.

Garner, R., Ding, T., & Deighton, C. (2014). Management of rheumatoid arthritis. *Medicine (United Kingdom)*, *42*(5), 237–242. https://doi.org/10.1016/j.mpmed.2014.02.004.

Gavrilă, B. I., Ciofu, C., & Stoica, V. (2016). Biomarkers in Rheumatoid Arthritis, what is new? *Journal of Medicine and Life*, *9*(2), 144–148. http://www.ncbi.nlm.nih.gov/pubmed/27453744.

Hart, A. B., & Kranzler, H. R. (2015). Alcohol Dependence Genetics: Lessons Learned From Genome-Wide Association Studies (GWAS) and Post-GWAS Analyses. *Alcoholism: Clinical and Experimental Research*, *39*(8), 1312–1327. https://doi.org/10.1111/acer.12792.

Hettiarachchi, G., & Komar, A. A. (2022). GWAS to Identify SNPs Associated with Common Diseases and Individual Risk: Genome Wide Association Studies (GWAS) to Identify SNPs Associated with Common Diseases and Individual Risk. In *Single Nucleotide Polymorphisms* (pp. 51–76). Springer International Publishing. https://doi.org/10.1007/978-3-031-05616-1_4.

Huang, S.-C., Hua, D.-J., Sun, Q.-Q., Zhang, L.-N., Cen, H., & Zhou, L. (2019). Associations of TRAF1/C5 rs10818488 and rs3761847 polymorphisms with genetic susceptibility to rheumatoid arthritis: a case-control study and updated meta-analysis. *Central European Journal of Immunology*, *44*(2), 159–173. https://doi.org/10.5114/ceji.2019.87067.

Karnell, J. L., Rieder, S. A., Ettinger, R., & Kolbeck, R. (2019). Targeting the CD40-CD40L pathway in autoimmune diseases: Humoral immunity and beyond. *Advanced Drug Delivery Reviews*, *141*, 92–103. https://doi.org/10.1016/j.addr.2018.12.005.

Kong, X.-F., Vogt, G., Itan, Y., Macura-Biegun, A., Szaflarska, A., Kowalczyk, D., Chapgier, A., Abhyankar, A., Furthner, D., Djambas Khayat, C., Okada, S., Bryant, V. L., Bogunovic, D., Kreins, A., Moncada-Vélez, M., Migaud, M., Al-Ajaji, S., Al-Muhsen, S., Holland, S. M., … Boisson-Dupuis, S. (2013). Haploinsufficiency at the human IFNGR2 locus contributes to mycobacterial disease. *Human Molecular Genetics*, *22*(4), 769–781. https://doi.org/10.1093/hmg/dds484.

Kwon, Y. C., Lim, J., Bang, S. Y., Ha, E., Hwang, M. Y., Yoon, K., Choe, J. Y., Yoo, D. H., Lee, S. S., Lee, J., Chung, W. T., Kim, T. H., Sung, Y. K., Shim, S. C., Choi, C. B., Jun, J. B., Kang, Y. M., Shin, J. M., Lee, Y. K., … Bae, S. C. (2020). Genome-wide association study in a Korean population identifies six novel susceptibility loci for rheumatoid arthritis. *Annals of the Rheumatic Diseases*, *79*(11), 1438–1445. https://doi.org/10.1136/annrheumdis-2020-217663.

Leitner, N. R., Witalisz-Siepracka, A., Strobl, B., & Müller, M. (2017). Tyrosine kinase 2 – Surveillant of tumours and bona fide oncogene. *Cytokine*, *89*, 209–218. https://doi.org/10.1016/j.cyto.2015.10.015.

Leng, R. X., Di, D. S., Ni, J., Wu, X. X., Zhang, L. L., Wang, X. F., Liu, R. S., Huang, Q., Fan, Y. G., Pan, H. F., Wang, B., & Ye, D. Q. (2020). Identification of new susceptibility loci associated with rheumatoid arthritis. *Annals of the Rheumatic Diseases*, *79*(12), 1565–1571. https://doi.org/10.1136/annrheumdis-2020-217351.

Liao, Y., Wang, J., Jaehnig, E., Shi, Z., & Zhang, B. (2019). WebGestalt 2019: gene set analysis toolkit with revamped UIs and APIs. *Nucleic Acids Research*, *47*. https://doi.org/10.1093/nar/gkz401.

Ma, R. C. W., Hu, C., Tam, C. H., Zhang, R., Kwan, P., Leung, T. F., Thomas, G. N., Go, M. J., Hara, K., Sim, X., Ho, J. S. K., Wang, C., Li, H., Lu, L., Wang, Y., Li, J. W., Wang, Y., Lam, V. K. L., Wang, J., … Jia, W. (2013). Genome-wide association study in a Chinese population identifies a susceptibility locus for type 2 diabetes at 7q32 near PAX4. *Diabetologia*, *56*(6), 1291–1305. https://doi.org/10.1007/s00125-013-2874-4.

Mortaz, E., Tabarsi, P., Mansouri, D., Khosravi, A., Garssen, J., Velayati, A., & Adcock, I. M. (2016). Cancers related to immunodeficiencies: Update and perspectives. *Frontiers in Immunology*, *7*(SEP), 1–13. https://doi.org/10.3389/fimmu.2016.00365.

Narayanan, R. (2016). Repurposing Drugs in the Genomics Era: Bioinformatics Approaches. *MOJ Proteomics & Bioinformatics*, *3*(4). https://doi.org/10.15406/mojpb.2016.03.00092.

Narazaki, M., Tanaka, T., & Kishimoto, T. (2017). The role and therapeutic targeting of IL-6 in rheumatoid arthritis. *Expert Review of Clinical Immunology*, *13*(6), 535–551. https://doi.org/10.1080/1744666X.2017.1295850.

Okada, Y., Wu, D., Trynka, G., Raj, T., Terao, C., Ikari, K., Kochi, Y., Ohmura, K., Suzuki, A., Yoshida, S., Graham, R. R., Manoharan, A., Ortmann, W., Bhangale, T., Denny, J. C., Carroll, R. J., Eyler, A. E., Greenberg, J. D., Kremer, J. M., … Plenge, R. M. (2014). Genetics of rheumatoid arthritis contributes to biology and drug discovery. *Nature*, *506*(7488), 376–381. https://doi.org/10.1038/nature12873.

Qiu, S., Munir, A., Malik, S. I., Khan, S., & Hassan, A. (2021). Identification of differentially expressed genes and pathways crosstalk analysis in Rheumatoid and Osteoarthritis using next-generation sequencing and protein-protein networks. *Saudi Journal of Biological Sciences*, *28*(8), 4656–4663. https://doi.org/10.1016/j.sjbs.2021.04.076.

Reay, W. R., & Cairns, M. J. (2021). Advancing the use of genome-wide association studies for drug repurposing. *Nature Reviews Genetics*, *22*(10), 658–671. https://doi.org/10.1038/s41576-021-00387-z.

Ren, T. li, Han, Z. jun, Yang, C. jian, Hang, Y. xing, Fang, D. yu, Wang, K., Zhu, X., Ji, X. jing, & Zhou, F. fan. (2014). Association of SLC22A4 gene polymorphism with rheumatoid arthritis in the chinese population. *Journal of Biochemical and Molecular Toxicology*, *28*(5), 206–210. https://doi.org/10.1002/jbt.21554.

Smith, C. L., & Eppig, J. T. (2009). The mammalian phenotype ontology: Enabling robust annotation and comparative analysis. *Wiley Interdisciplinary Reviews: Systems Biology and Medicine*, *1*(3), 390–399. https://doi.org/10.1002/wsbm.44.

Song, X., & Lin, Q. (2017). Genomics, transcriptomics and proteomics to elucidate the pathogenesis of rheumatoid arthritis. *Rheumatology International*, *37*(8), 1257–1265. https://doi.org/10.1007/s00296-017-3732-3.

Too, C. L., Murad, S., Dhaliwal, J. S., Larsson, P., Jiang, X., Ding, B., Alfredsson, L., Klareskog, L., & Padyukov, L. (2012). Polymorphisms in peptidylarginine deiminase associate with rheumatoid arthritis in diverse Asian populations: evidence from MyEIRA study and meta-analysis. *Arthritis Research & Therapy*, *14*(6), R250. https://doi.org/10.1186/ar4093.

Van Steenbergen, H. W., Rodríguez-Rodríguez, L., Berglin, E., Zhernakova, A., Knevel, R., Ivorra-Cortés, J., Huizinga, T. W. J., Fernández-Gutiérrez, B., Gregersen, P. K., Rantapää-Dahlqvist, S., & van der Helm-van Mil, A. H. M. (2015). A genetic study on C5-TRAF1 and progression of joint damage in rheumatoid arthritis. *Arthritis Research and Therapy*, *17*(1), 1–12. https://doi.org/10.1186/s13075-014-0514-0.

Wang, J., Vasaikar, S., Shi, Z., Greer, M., & Zhang, B. (2017). WebGestalt 2017: A more comprehensive, powerful, flexible and interactive gene set enrichment analysis toolkit. *Nucleic Acids Research*, *45*(W1), W130–W137. https://doi.org/10.1093/nar/gkx356.

Ward, L. D., & Kellis, M. (2016). HaploReg v4: Systematic mining of putative causal variants, cell types, regulators and target genes for human complex traits and disease. *Nucleic Acids Research*, *44*(D1), D877–D881. https://doi.org/10.1093/nar/gkv1340.

Zhang, T.-P., Li, R., Huang, Q., Pan, H.-F., Ye, D.-Q., & Li, X.-M. (2020). Association of NCF2, NCF4, and CYBA Gene Polymorphisms with Rheumatoid Arthritis in a Chinese Population. *Journal of Immunology Research*, *2020*, 8528976. https://doi.org/10.1155/2020/8528976.

Zhang, Z., Miteva, M. A., Wang, L., & Alexov, E. (2012). Analyzing effects of naturally occurring missense mutations. *Computational and Mathematical Methods in Medicine*, *2012*. https://doi.org/10.1155/2012/805827.

Mapping Rheumatoid Arthritis Susceptibility through Integrative Bioinformatics and Genomics

ORIGINALITY REPORT

15%	10%	12%	5%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	www.hindawi.com Internet Source	1%
2	www.ncbi.nlm.nih.gov Internet Source	1%
3	"Cellular Primary Immunodeficiencies", Springer Science and Business Media LLC, 2021 Publication	1%
4	rdw.rowan.edu Internet Source	1%
5	Kuojun Zhang, Ke Ye, He Tang, Zhihao Qi, Tianyu Wang, Jie Mao, Xiangyu Zhang, Sheng Jiang. "Development and Therapeutic Implications of Tyrosine Kinase 2 Inhibitors", Journal of Medicinal Chemistry, 2023 Publication	1%
6	www.jci.org Internet Source	1%
7	pubs.rsc.org	

	Internet Source	1%
8	Lalu Muhammad Irham, Wan-Hsuan Chou, Yu-Shiuan Wang, Wirawan Adikusuma et al. "Evaluation for the Genetic Association between Store-Operated Calcium Influx Pathway (STIM1 and ORAI1) and Human Hepatocellular Carcinoma in Patients with Chronic Hepatitis B Infection", Biology, 2020 Publication	1%
9	Textbook of Personalized Medicine, 2015. Publication	1%
10	"Primary Immunodeficiency Diseases", Springer Science and Business Media LLC, 2017 Publication	1%
11	brevets-patents.ic.gc.ca Internet Source	<1%
12	ijracs.rumahpeneleh.or.id Internet Source	<1%
13	Submitted to University of Salford Student Paper	<1%
14	Xinxin Chi, Miaozhen Huang, Hailin Tu, Bin Zhang, Xin Lin, Huji Xu, Chen Dong, Xiaoyu Hu. "Innate and adaptive immune abnormalities underlying autoimmune	<1%

	diseases: the genetic connections", Science China Life Sciences, 2023 Publication	
15	Submitted to Oregon State University Student Paper	<1 %
16	Submitted to SDM Universitas Gadjah Mada Student Paper	<1 %
17	biomedpharmajournal.org Internet Source	<1 %
18	journal.uny.ac.id Internet Source	<1 %
19	www.hortidaily.com Internet Source	<1 %
20	www.sciencegate.app Internet Source	<1 %
21	www.frontiersin.org Internet Source	<1 %
22	Kaipeng Xie, Mengxi Chen, Meng Zhu, Cheng Wang et al. "A polymorphism in miR-1262 regulatory region confers the risk of lung cancer in Chinese population", International Journal of Cancer, 2017 Publication	<1 %
23	admin.lidsen.com Internet Source	<1 %

24	"Gut Microbiota in Aging and Chronic Diseases", Springer Science and Business Media LLC, 2023 Publication	<1 %
25	journals.biologists.com Internet Source	<1 %
26	Amanda Hefton, Shu Ying Liang, Kathryn Ni, Victoria Carter, Kennedy Ukadike, Christian Lood, Tomas Mustelin. "Autoantibodies against citrullinated serum albumin in patients with rheumatoid arthritis", Journal of Translational Autoimmunity, 2019 Publication	<1 %
27	Guolian Yuan, Pan Luo, Ke Xu, Wensen Jing, Feng Zhang. "A large-scale genetic correlation scan between rheumatoid arthritis and human blood metabolites", Annals of Human Genetics, 2022 Publication	<1 %
28	The Genetics of Type 2 Diabetes and Related Traits, 2016. Publication	<1 %
29	Yanan Pang, Minglu Ma, Dong Wang, Xun Li, Li Jiang. "TANK Promotes Pressure Overload Induced Cardiac Hypertrophy via Activating AKT Signaling Pathway", Frontiers in Cardiovascular Medicine, 2021 Publication	<1 %

30	bmcmedgenomics.biomedcentral.com Internet Source	<1 %
31	scholarworks.umass.edu Internet Source	<1 %
32	smb.org.mx Internet Source	<1 %
33	Imaniar Noor Faridah, Haafizah Dania, Yen-Hsu Chen, Woro Supadmi et al. "Dynamic Changes of Platelet and Factors Related Dengue Haemorrhagic Fever: A Retrospective Study in Indonesian", Diagnostics, 2022 Publication	<1 %
34	Veronica Ferrucci, Fatemeh Asadzadeh, Francesca Collina, Roberto Siciliano et al. "Prune-1 drives polarization of tumor-associated macrophages (TAMs) within the lung metastatic niche in triple-negative breast cancer", iScience, 2020 Publication	<1 %
35	Wirawan Adikusuma, Wan-Hsuan Chou, Min-Rou Lin, Jafit Ting et al. "Identification of Druggable Genes for Asthma by Integrated Genomic Network Analysis", Biomedicines, 2022 Publication	<1 %
36	doi.org Internet Source	<1 %

Exclude quotes	On	Exclude matches	Off
Exclude bibliography	On		