

Brief Report

Pathway Analysis of Genome Wide Association Studies (GWAS) Data Associated with Male Infertility

Rupashree Salvi ¹, Ulka Gawde ², Susan Idicula-Thomas ² and Barnali Biswas ^{1,*}

¹ Department of Innate Immunity, ICMR-National Institute for Research in Reproductive and Child Health, Mumbai 400012, India

² Biomedical Informatics Centre, ICMR-National Institute for Research in Reproductive and Child Health, Mumbai 400012, India

* Correspondence: biswasb@nirrh.res.in or barnali.biotech@gmail.com

Abstract: Background: Infertility is a common condition affecting approximately 10–20% of the reproductive age population. Idiopathic infertility cases are thought to have a genetic basis, but the underlying causes are largely unknown. However, the genetic basis underlying male infertility in humans is only partially understood. The Purpose of the study is to understand the current state of research on the genetics of male infertility and its association with significant biological mechanisms. Results: We performed an Identify Candidate Causal SNPs and Pathway (ICSN Pathway) analysis using a genome-wide association study (GWAS) dataset, and NCBI-PubMed search which included 632 SNPs in GWAS and 451 SNPs from the PubMed server, respectively. The ICSN Pathway analysis produced three hypothetical biological mechanisms associated with male infertility: (1) rs8084 and rs7192→*HLA-DRA*→inflammatory pathways and cell adhesion; rs7550231 and rs2234167→*TNFRSF14*→TNF Receptor Superfamily Member 14→T lymphocyte proliferation and activation; rs1105879 and rs2070959→*UGT1A6*→UDP glucuronosyltransferase family 1 member A6→Metabolism of Xenobiotics, androgen, estrogen, retinol, and carbohydrates. Conclusions: We believe that our results may be helpful to study the genetic mechanisms of male infertility. Pathway-based methods have been applied to male infertility GWAS datasets to investigate the biological mechanisms and reported some novel male infertility risk pathways. This pathway analysis using GWAS dataset suggests that the biological process related to inflammation and metabolism might contribute to male infertility susceptibility. Our analysis suggests that genetic contribution to male infertility operates through multiple genes affecting common inflammatory diseases interacting in functional pathways.



Citation: Salvi, R.; Gawde, U.; Idicula-Thomas, S.; Biswas, B. Pathway Analysis of Genome Wide Association Studies (GWAS) Data Associated with Male Infertility. *Reprod. Med.* **2022**, *3*, 235–245. <https://doi.org/10.3390/reprodmed3030018>

Academic Editor: Berthold Huppertz

Received: 1 June 2022

Accepted: 14 July 2022

Published: 12 August 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

Keywords: male infertility; GWAS; SNPs; ICSN; *HLA-DRA*; *TNFRSF14*; *UGT1A6*

1. Introduction

Infertility is an emerging global problem affecting men and women equally [1]. Several causes contribute to both male and female infertility, and male infertility is responsible for 50% of the total number of cases [2,3]. The molecular genetics of male infertility are multifactorial and numerous [4–6]. Although several causes have been attributed to impaired male infertility, the majority of cases with inadequate spermatogenesis or defective sperm remain to be understood. Factors contributing are the genetics of an individual, the damaging effect of oxidative stress, sperm DNA Integrity, environmental conditions, lifestyle, patient age, and body mass index (BMI) responsible for the onset of this disorder [6–12]. Anomalies in genetic factors account for 15–30% of cases of impaired spermatogenesis and male infertility [13,14]. Oxidative stress is the major cause of DNA fragmentation in spermatozoa [7,15]. Antioxidants play a protective role, although a balance of reduction and oxidation is required for essential functions, including fertilization [15]. Studies have confirmed the assumption that polymorphisms in oxidative stress genes might increase the

risk of sperm DNA damage and unexplained male infertility [16]. During spermatogenesis, the sperm nucleus undergoes a marked rearrangement, which involves the removal of histones and their replacement by various basic nuclear proteins, including highly positively charged protamines [17,18]. It has been suggested that disturbances in the nuclear condensation of sperm might affect genes involved in the production of functional germ cells [19]. In order to investigate whether changes in the amount or structure of protamine could be the cause of human infertility, sequencing of the protamine gene was performed [20]. The authors first identified a heterozygous SNP in the protamine 1 gene in three infertile men [20]. This SNP causes an amino acid change in a highly conserved arginine cluster. Moreover, the SNP created improper phosphorylation, which could alter both DNA binding and protamine to protamine interaction in the sperm nucleus [20].

In a case control study with the Indian population, two rare novel mutations in sperm nuclear protein genes (protamine and transition nuclear protein) are associated with higher sperm DNA fragmentation [21]. In another study, the involvement of sperm nuclear basic proteins in DNA oxidative damage has been elucidated [7]. The altered ratio of protamines and histones determines an unstable binding to DNA and reduced DNA protection from external agents such as heavy metals [7]. They have shown the importance of arginine residues in the Copper ion induced DNA breakage of sperm histones inducing oxidative damage [7]. In another study, the authors reported molecular alteration of small nuclear basic proteins, DNA, and semen parameters in transgenerational inherited effects due to environmental contaminants [22]. They have found involvement in DNA oxidative damage of human small nuclear basic proteins exposed to pollutants, giving new insights into the toxicity mechanisms of some heavy metals [22].

Moreover, little is known about environmental chemical exposure in young women where kallikrein related serine peptidase is regulated [23]. Environmental factors play role in kallikrein related serine peptidase 3 (KLK3) secretion [23]. Some environmental pollutants could activate KLK3 secretion by prostatic epithelial cells and identify as a new biomarker to assess the early effect of environmental exposure on female reproductive health [23]. It is assumed that genetic makeup contributes to 10–15% of spermatogenic failure for in vitro fertilization [24].

It is helpful to understand sperm DNA integrity and unexplained infertility because DNA fragmentation is present in normozoospermic sub-fertile men [25,26]. Previous studies have shown the involvement of genetic polymorphism in specific genes associated with male fertility [27–29]. In one of the study, the authors have reported evidence of a single nucleotide polymorphism in the histidine-rich glycoprotein (HRG) named HRG C633T, which is required for male fertility [30].

Although no studies to prove the relationship between HRG C633T SNP and sperm DNA integrity was observed. It is still not clear that tests of sperm DNA integrity with a high predictive value can determine the result of *in-vivo* fertilization although the outcome of IVF and ICSI require more analysis and clarification [31]. Previous findings did not support the relationship between sperm DNA integrity and assisted reproduction technologies (ART) success [32–34]. While others reported a significant reduction in pregnancy rates following IVF or ICSI among couples with high sperm DNA fragmentation [35–38].

The current understanding of sperm DNA damage and its effects on reproductive outcomes requires further investigation. Thus, association studies from larger genome wide analysis to identify polymorphisms (SNPs) affecting spermatogenesis, which might lead to oligozoospermia or azoospermia induced male infertility [39]. Genotyping helps us in the identification of genetic factors, and the variants can be explored by single nucleotide polymorphism (SNP) array, NGS next-generation sequencing (NGS), microarrays, comparative genomic hybridization-array (array-CGH) to understand rare variants and disease etiology on a molecular level [13,40].

In this study, we used Genome Wide Associations Studies (GWAS) platform to identify potential SNPs which are related to male infertility. The study of the GWAS dataset is an emerging and powerful source to understand the variants associated with particular

diseases [41,42]. This study will help us search for hidden prognostic marker genes which are related to male infertility. There are various modifications associated with GWAS studies which include a pathway-based approach (ICSN pathway). This helps us to understand causal variants associated with complex diseases [43,44]. The combination of the GWAS dataset and ICSN pathway analysis will help us to find out candidate SNPs that target key pathways related to complex diseases. We have collected SNPs from the GWAS dataset and PubMed literature related to male infertility and performed an ICSN analysis to find possible mechanisms associated with male infertility. The consideration of these SNPs and pathways associated with these SNPs can help us screen infertile patients and improve prognosis.

HLA is a large, highly polymorphic gene set located on human chromosome 6 and is subdivided into class I and class II regions (MHCI and MHCII, respectively) [45]. HLA-DRA is mostly present on the cell surface of B-Cell dendritic cells, and macrophages. It plays important role in generating an immune response against pathogen derived peptides by presenting the antigenic peptides to T Cell [46]. HLA-DRA molecule gives the ability to individuals to respond to a variety of pathogens [47]. The polymorphism in the subunit region of the HLA-DRA molecule plays important role in generating an immune response against a variety of pathogens [48,49].

TNFRSF14 belongs to TNF (tumor necrosis factor) receptor superfamily. It plays important role in the inflammatory and inhibitory signaling network. It also plays an important role in the proliferation of T cells and the production of antimicrobial agents. In the case of adaptive immune response interaction of TNFRSF14 with T cell provide survival cell for effector T-cell [50]. UGT1A6 encodes for the gene UDP-glucuronosyltransferase 1–6 enzyme, which plays an important role in the metabolism of xenobiotics [46]. It helps in linking glucuronic acid to its substrate, due to which excretion of xenobiotics from cells becomes easier [46].

2. Methods

Candidate SNPs as input: The SNPs used in the study were extracted from the NHGRI-EBI catalog of human genome wide association studies. We selected 632 SNPs from the GWAS dataset, which are associated with male infertility. We have selected the most common SNPs having Risk Allele Frequency (RAF) > 0.02 and p -value $< 5 \times 10^{-2}$. We also selected significant SNPs from NCBI PubMed (pubmed.ncbi.nlm.nih.gov) associated with male infertility. Finally, 1083 SNPs were used for pathway analysis

Pathway Analysis associated with male infertility: Pathway Analysis was applied to understand male infertility related SNPs associated with other complex diseases. Thus, we conducted ICSN (Identify Candidate Causal SNPs and Pathway) analysis and utilized SNPs from the GWAS dataset and NCBI PubMed as input. The selection of candidate causal SNPs by LD analysis and most significant functional SNPs were annotated. The second step includes understanding the biological processes for the preselected candidate SNPs using Pathway Based Analysis (PBA). The False Discovery Rate (FDR) and p -value cut-off for PBA were set as 0.05. There are various options available we selected KEGG, GO Biological process, Biocarta and GO Molecular function to get properly defined and descriptive pathways. We selected the parameters for LD analysis CEU: European American (CEPH) as the HapMap population, LD measurement cut off of $r^2 = 0.8$, and distance search LD neighborhood = 200 kb. Functional SNPs defined can alter protein or gene expression in close association with pathway. Functional SNPs include essential splice site and intronic, non-synonymous coding, and regulatory region. Two conditions were applied for ICSN analysis. the first parameter applied was “within gene” only, and the second parameter was “gene set size minimum 1 and maximum 500”. Pathways that were enriched from the ICSN based webtool were drawn using a Bubble plot. The plot was generated using the ggplot2 package of R (version 4.0.5) [51]. Moreover, the network of genes associated with enriched pathways was created using the igraph package of R (version 4.0.5) [52].

3. Results

To perform pathway enrichment analysis, we utilized the most significant SNPs from GWAS (632) and PubMed survey of male infertility related SNPs (451) having p -value $< 5 \times 10^{-2}$ as input. We performed in-silico analysis using the ICSN pathway web server and identified six candidate SNPs and thirty-two candidate pathways. The SNPs rs8084, rs7550231, rs2234167, rs1105879, and rs2070959 are in linkage disequilibrium (LD) with rs7192, rs4486391, rs10910093, rs2070959, respectively ($r^2 = 0.80$ – 0.96), ($-\log_{10}(P) = 5.523, 1.867, 1.745$ respectively). The SNP rs7192, which was represented in the original GWAS ($-\log_{10}(P) = 5.523$), was not in LD with any SNP (Table 1). In our study, the SNPs were associated with the three most significant biological pathways which were as follows: (i) SNP rs8084 and rs7192 encoding for the gene *HLA-DRA* is a Class II MHC molecule displays antigenic peptides on the surface of APCs and is involved in several inflammatory pathways and cell adhesion. *HLA-DRA* does not have polymorphism in the peptide binding part and acts as a sole alpha chain for *DRB1*, *DRB3*, *DRB4*, and *DRB5*. SNPs rs7550231 and rs2234167 encode for the gene *TNFRSF14* (TNF Receptor Superfamily Member 14) that are expressed in T cells and activate a cascade of T cell signaling in response to ligand binding. SNPs rs1105879 and rs2070959 encoding for *UGT1A6* gene (UDP glucuronosyl-transferase family 1 member A6) involved in metabolic pathways. *UGT1A6* metabolizes Xenobiotics, androgen, estrogen, retinol, and carbohydrates. The candidate SNPs identified after analysis are thus involved in biological and cellular processes including inflammation, metabolism, and cell cycle regulation.

Table 1. Candidate SNPs associated with male fertility.

Candidate Causal SNP	Functional Class	Gene	Candidate Causal Pathway ^a	$-\log_{10}(P)$ ^b	In LD with	r^2	D'	$-\log_{10}(P)$ ^c
rs8084	Essential splice site & intronic	<i>HLA-DRA</i>	1 2 3 4 5 6 7 8 9 10 11 12 13 14 16 17 24 26 27 32	-	rs7192	0.818	1.0	5.523
rs7192	Non synonymous coding	<i>HLA-DRA</i>	1 2 3 4 5 6 7 8 9 10 11 12 13 14 16 17 24 26 27 32	5.523	rs7192	-	-	5.523
rs7550231	Regulatory region	<i>TNFRSF14</i>	15 29	-	rs4486391	0.96	1.0	1.867
rs2234167	Non synonymous coding	<i>TNFRSF14</i>	15 29	-	rs10910093	0.89	1.0	1.745
rs1105879	Non synonymous coding	<i>UGT1A6</i>	18 19 20 21 22 23 25 28 30 31	-	rs2070959	0.825	1.0	47.000
rs2070959	Non synonymous coding	<i>UGT1A6</i>	18 19 20 21 22 23 25 28 30 31	47.000	rs2070959	-	-	47.000

Abbreviations: SNP Single nucleotide polymorphism; LD Linkage disequilibrium. ^a the numbers indicate the indices of pathway which are ranked in order of statistical significance (false discovery rate). ^b $-\log_{10}(P)$ for the candidate SNPs in the original GWAS dataset; —indicate that this SNPs are not indicated in the original GWAS dataset. ^c $-\log_{10}(P)$ for the SNPs (in LD with the candidate SNPs) in the original GWAS dataset.

The six candidate causal SNPs and 32 candidate causal pathways in Table 1 are associated with several hypothetical biological processes as represented in Supplementary Materials Table S1 (a) rs8084 (essential splice site and intronic) encoding for the gene *HLA-DRA* (Major Histocompatibility Complex, Class II, DR Alpha) is influencing 20 inflammatory pathways such as MHC pathway, T Cell Receptor activation, Interleukin 5 signaling, B cell activation, Eosinophil signaling, T.Cell Activation, Inflammatory response, B-Cell

surface molecule signaling, Graft Versus Host disease signaling, Allograft rejection, Autoimmune thyroid disease, Type 1 Diabetes, Asthma. (b) rs7192 (non-synonymous coding) also encoding for the gene *HLA-DRA* (Major Histocompatibility Complex, Class II, DR Alpha) is involved in the above mentioned signaling pathways, (c) rs7550231 (regulatory region) and rs2234167 (non-synonymous coding) encoding for the gene *TNFRSF14* (TNF Receptor Superfamily Member 14) are involved in 2 inflammatory pathways such as Lymphocyte proliferation and T Cell activation signaling pathways, (d) rs1105879 (non-synonymous coding) and rs2070959 (non-synonymous coding) encoding for the gene *UGT1A6* (UDP Glucuronosyltransferase Family 1 Member A6) are involved in 10 metabolic pathways such as Drug metabolism, Xenobiotic metabolism, Androgen, and estrogen metabolism, Pentose and Glucuronate conversion, Porphyrin metabolism, Retinol metabolism, UDP-Glycosyl Transferase Activity, Transferase activity, starch, and sucrose metabolism, and response to xenobiotic stimulus signaling pathways (Supplementary Materials Table S1).

We represented causal pathways based on FDR and the number of variants associated with a pathway in a graphical representation (bubble plot). The plot is drawn with the X axis as enriched pathways and the Y axis as $-\log_{10}(p)$ value. The bubbles represent the false discovery rate (FDR) and the number of genes mapped with variants. The genes and the mapped variants in each pathway are shown in Supplementary Materials Table S1. In pathways such as UDP Glycosyl Transferase Activity and Transferase activity Transferring Hexosyl Group, the highest number of genes mapped with variants is nine and FDR 0.024. The nine variants are *UGT1A8*, *UGT1A9*, *UGT1A10*, *UGT1A6*, *UGT1A7*, *UGT1A4*, *UGT1A5*, *UGT1A3*, *CDA*, *EXTL1*, and *ATP13A*. Furthermore, the most significant FDR (0.001) having 2 variant genes (*HLA-DRB1* and *HLA-DRA*) influences inflammatory pathways such as the MHC pathway, TCRA pathway, IL5 signaling, B-cell Pathway, Eosinophils Pathway, CTLA pathway, Inflammatory pathway, BB cell pathway and B-Lymphocyte pathway (Figure 1).

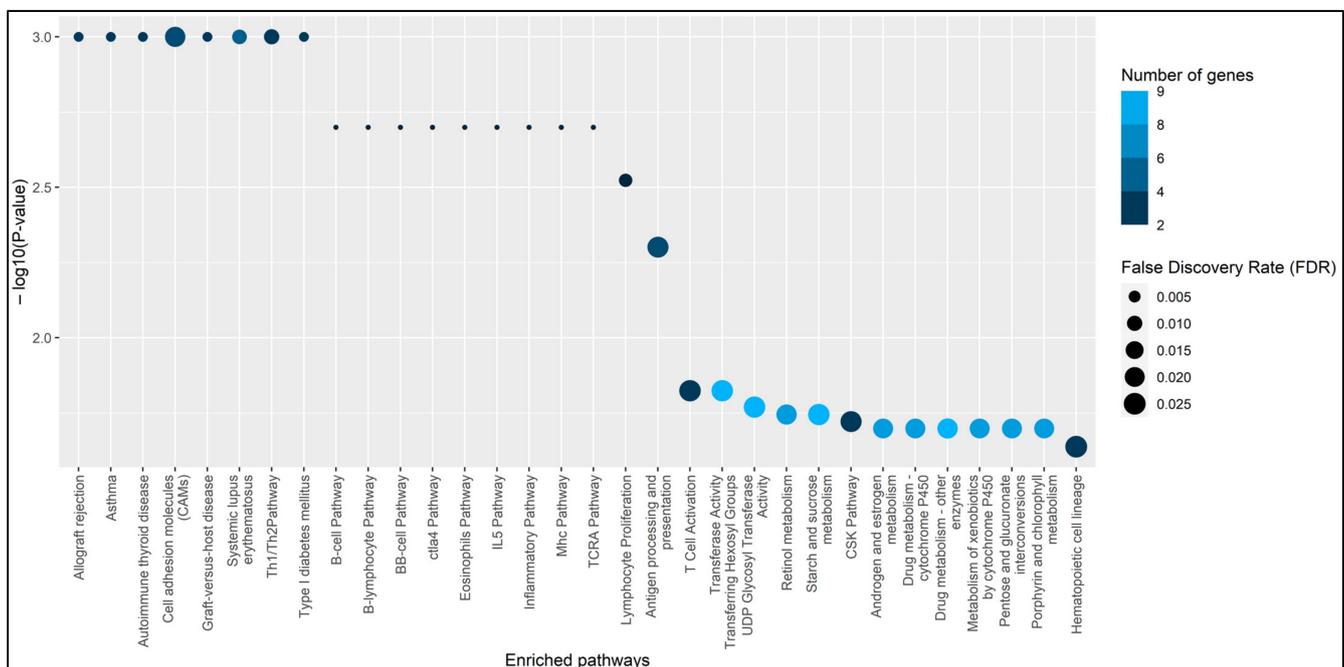


Figure 1. Bubble plot depicting enriched pathways. X-axis represents enriched pathways and Y-axis represents negative log of p -values. The bubble size is proportional to the false discovery rate (FDR) and color represents the number of genes associated with the pathways. The plot was generated using ggplot2 R package [51].

We have also represented the causal pathways in the network form. We started with the GWAS summary table for 24 genes or variants and 32 pathways downloaded from ICSN Pathway (Supplementary Materials Table S1). Figure 2 illustrates our analysis framework.

We grouped this network into 2 clusters (i) 22 immune related traits and (ii) 10 metabolic traits. The yellow and green circular nodes come under the category of immune related pathways; Blue circular nodes represent metabolic pathways. The immune related genes had more associated pathways individually and more shared pathways. We summarized the features of these datasets in Supplementary Materials Table S1.

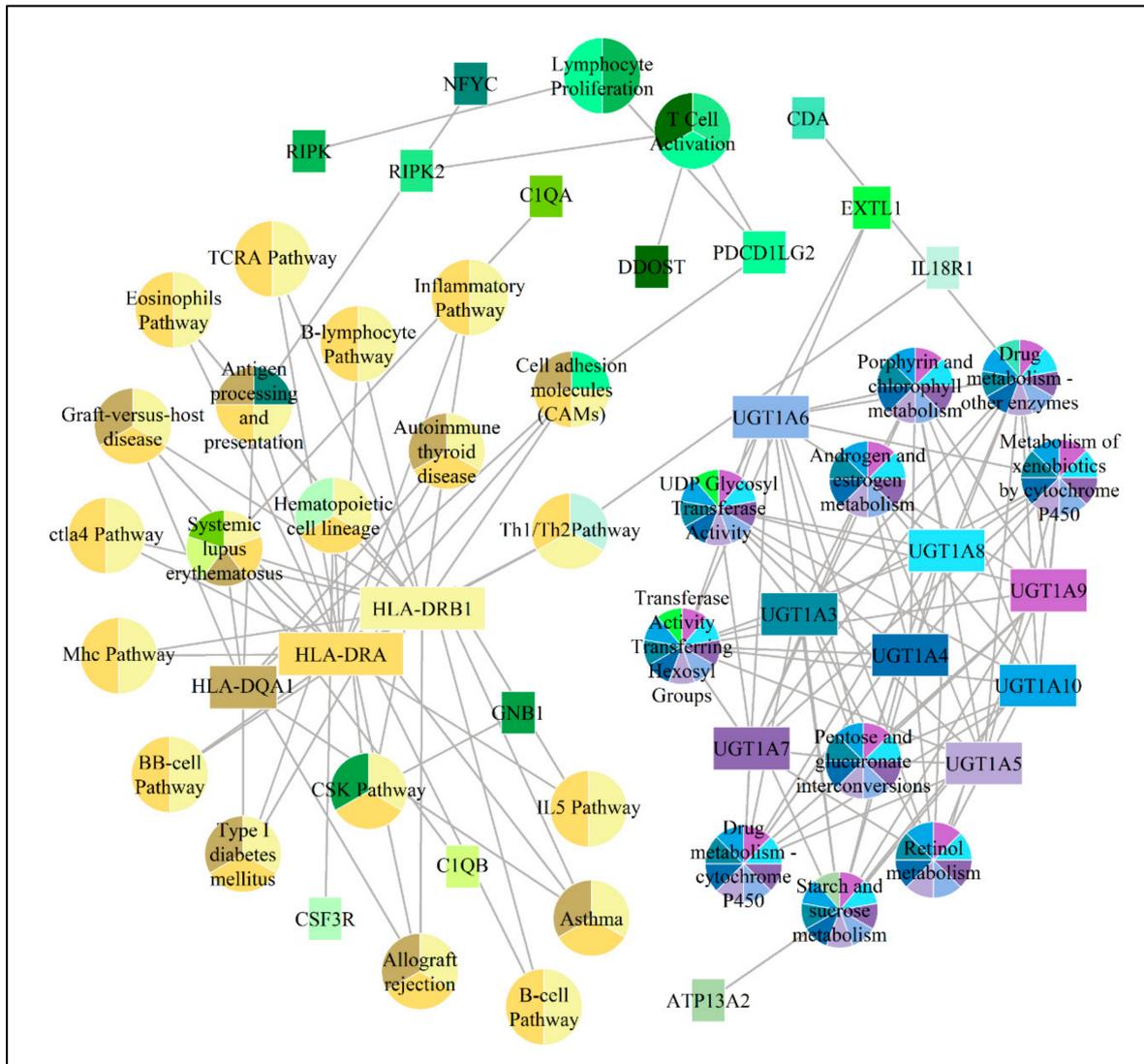


Figure 2. Network of genes associated with enriched pathways. Square and circular nodes represent the genes and enriched pathways respectively. Size of square node is proportional to the number of pathways associated with each gene. The circular node is segmented into colored pies to illustrate the associated genes with each pathway. Human leukocyte antigen (HLA) and UDP glucuronosyltransferase (UGT) genes are colored respectively with shades of yellow and blue. Remaining genes are indicated in shades of green. Edge represents a gene-pathway association. The network was created using igraph R package [52].

Our goal is to investigate infertility related genetic components which share multiple traits at the pathway level. We employed the igraph R package to evaluate and investigate pathways based on the combined effect of multiple variants and genes which may individually represent associations with the diseases or traits and pathways. Although the pathway itself did not show a clear relationship with male infertility related traits, several genes in this pathway are observed to be associated with metabolism and inflammation. For

immune related traits such as MHC pathway, CTLA4 pathway, Eosinophil pathway, TCRA pathway, Inflammatory pathway, B lymphocyte pathway, Th1/Th2 Pathway, IL5 Pathway, and CSK pathway. For metabolism such as Drug metabolism, Metabolism of xenobiotics, Androgen and estrogen metabolism, pentose and glucuronate metabolism, Porphyrin and chlorophyll metabolism, Starch and Sucrose metabolism, and retinol metabolism are associated with the variants.

By exploring the diseases associated with the immune related pathways such as Systemic Lupus Erythematosus, Autoimmune thyroid disease, Graft versus host disease, Type I diabetes mellitus, and Asthma we found that in most cases there is a combined effect of multiple genes rather than driven by a single gene (Figure 2 and Supplementary Materials Table S1).

4. Discussion

In this study, GWAS analysis has been applied with significant thresholds (5×10^{-2}) to select genetic variants for ICSN pathway-based analysis. We explored GWAS and male infertility SNP data to examine functional related genes for their combined effects in a pathway based study. Our results provide preliminary knowledge and insights into functional association related to each gene. One of the advantages of pathway-based analysis is that the combined effect of multiple variants instead of the effects of each individual gene, was examined for disease associations. We thus asked whether the same genes contributed to the disease association at the pathway level. We found individual variants are associated with multiple traits to share biological pathways related to inflammation and metabolism. We hypothesized that indirect pathways are likely due to the pathogenicity of male infertility related genetic variants. This finding was consistent with previous studies that immune-related traits (variants in the *TGFB3* gene) may be associated with male infertility [53].

Oxidative stress, sperm DNA integrity, changes in the structure of sperm nuclear proteins, and exposure to environmental pollutants are some elements that can impair fertility in men [7,8,14,17,19,22,23]. Previous studies have shown the association of oxidative stress with male infertility [54]. However, no significant associations between the seven polymorphic genes (*NQO1* rs1800566, *SOD2* rs4880, *GSTM3* rs1571858, rs3814309, rs7483, *GSTM5* rs11807, and *GSTP1* rs1695) and risk of male infertility has been noted [55]. But the authors have reported by gene-gene interactions a decreased risk of male infertility in *GSTM3* and azoospermia [55]. Genome Wide Association studies have identified the SNPs which are associated with complex genetic traits or diseases. From our current study, we have identified 6 SNPs and 32 significant causal pathways. We identified polymorphic variants in the *HLA-DRA* gene which are rs8084 and rs7192 located on chromosome 6p21.32 and associated with male infertility. rs8084 and rs7192 are found to control spermatogenesis in humans [56]. It is also expressed on spermatocytes and spermatids which helps in cell adhesion and motility [57]. rs7194 which is 834bp apart from rs7192 is found to be associated with non-obstructive azoospermia (NOA) [39]. This key gene *HLA-DRA* (MHC class II) is responsible for the selection of T cells, antigen sampling, and activation of an immune response. Previous studies have shown variations of certain alleles in the *HLA* region to be associated with the development of a wide range of diseases from rheumatoid arthritis [58] to narcolepsy [59] and Parkinson's Disease [60]. Krausz and coworkers reviewed meta-analysis data, case-control, and GWAS studies performed on more than 370,000 SNP in key genes related to cell maintenance (general cell functions), apoptotic process, DNA repair, response to reactive oxygen species, etc., in regard to expression in germ cells [13]. The study concluded that a few SNPs are correlating with infertility phenotype. The five genes that were positively correlated are *MTHFR*, *ESR1/ESR2*, *NOS3*, *SIRPA/SIRPG* and *HLA-DRA13* [40]. For *HLA-DRA* gene, which is involved in inflammation and responsive to immune regulation, shows a significant decrease in male infertility. Concretely, c.82 + 926A > C, c.*63G > A, and c.724 T > G are the susceptibility factors for autoimmune diseases [54]. They assumed that these changes in the nucleotide and variants

may have a role in inflammatory responses of patients with antecedents of urogenital inflammation⁵⁴. TNFRSF14 (TNF Receptor Superfamily Member 14) is a member of the TNF receptor superfamily also known as herpes virus entry mediator (HVEM) [61]. Mutation and Polymorphism in the genes cause multiple sclerosis, celiac diseases, and rheumatoid arthritis [62–64]. The encoded protein activates inflammatory and inhibitory T cell responses [65]. In the context of male infertility, no evidence has been found.

UGT1A6 (UDP Glucuronosyltransferase Family 1 Member A6) plays an important role in the metabolism of xenobiotics and their conversion into water soluble metabolites [66]. Genetic Polymorphism in UGT1A6 has been observed in Pediatric Epileptic Patients and people with an increased risk of cancer, especially lung and breast cancer [67–69]. In the context of male infertility, no association studies have directly been reported. However, we also cannot exclude the possibility that these SNPs may not directly cause infertility. To reach a conclusion, further studies with large cohort and molecular studies would be essential. Thus, there is an urgent requirement to understand the multiple dysregulated effects or new associations of these key genes in terms of male fertility. This study will help in understanding complex conditions and various factors that may be essential for the prevention of adverse effects of a considerable number of drugs and to predict infertility risks. Our ICSN pathway analysis may provide the bases for new hypotheses that could not explain the interaction of genes and pathways in the pathogenesis of male infertility. Futuristic studies are needed to validate the causal nature of SNPs and genes underlying the association between identified pathways and human male infertility. Functional studies will identify the potent SNPs in males having oligozoospermia, asthenozoospermia, and male infertility related disorders. Our studies have indicated a better chance to understand polymorphism and various pathways that are involved in multi-systemic diseases.

5. Conclusions

In conclusion, we analyzed SNPs related to male infertility to identify genetic association with male infertility susceptibility at both SNP and pathway levels. Additional studies are needed to explore the genetic variation of pathways underlying the pathogenesis of male infertility.

Supplementary Materials: The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/reprodmed3030018/s1>.

Author Contributions: R.S. has identified SNPs from GWAS and NCBI-PubMed, used the ICSN web tool, analyzed the data and written the manuscript. B.B. has interpreted the data and finalized the manuscript. U.G. has generated the bubble plot and pathway figure. S.I.-T. has analyzed the figures and reviewed the manuscript. All authors have read and agreed to the published version of the manuscript.

Funding: This study is supported by DST INPIRE Faculty research grant and intramural grant to the department of Innate Immunity, ICMR-NIRRH (accession number RA/1285/07-2022). B.B. is a recipient of DST INPIRE Faculty Fellowship. S.I.-T. and U.G. are grateful to receive grants from Department of Biotechnology (DBT), India [BT/PR40165/BTIS/137/12/2021].

Institutional Review Board Statement: This article does not contain any studies with human participants or animals performed by any of the authors.

Informed Consent Statement: Not applicable.

Data Availability Statement: The datasets used and analyzed in this study are included in the manuscript.

Acknowledgments: The authors gratefully acknowledge the investigators for sharing their valuable SNP related information in GWAS dataset. We express our gratitude towards Bioinformatics Centre, ICMR-NIRRH for helping us in doing the analysis.

Conflicts of Interest: The authors declare no conflict of interest.

References

- Callahan, T.; Caughey, A.B. Infertility and Assisted Reproductive Technologies. In *Blueprints Obstetrics and Gynecology*; Lippincott Williams & Wilkins: Philadelphia, PA, USA, 2008; Volume 5.
- Agarwal, A.; Mulgund, A.; Hamada, A.; Chyatte, M.R. A Unique View on Male Infertility around the Globe. *Reprod. Biol. Endocrinol.* **2015**, *13*, 37. [[CrossRef](#)] [[PubMed](#)]
- Poongothai, J.; Gopenath, T.S.; Manonayaki, S. Genetics of Human Male Infertility. *Singap. Med. J.* **2009**, *50*, 336–347.
- Singh, K.; Jaiswal, D. Human Male Infertility: A Complex Multifactorial Phenotype. *Reprod. Sci.* **2011**, *18*, 418–425. [[CrossRef](#)] [[PubMed](#)]
- Karamzade, A.; Mirzapour, H.; Kheirollahi, M. Genetics Aspects of Male Infertility. *J. Isfahan Med. Sch.* **2013**, *31*, 1149–1162.
- Choy, J.T.; Eisenberg, M.L. Male Infertility as a Window to Health. *Fertil. Steril.* **2018**, *110*, 810–814. [[CrossRef](#)]
- Lettieri, G.; D’Agostino, G.; Mele, E.; Cardito, C.; Esposito, R.; Cimmino, A.; Giarra, A.; Trifuoggi, M.; Raimondo, S.; Notari, T.; et al. Discovery of the Involvement in DNA Oxidative Damage of Human Sperm Nuclear Basic Proteins of Healthy Young Men Living in Polluted Areas. *Int. J. Mol. Sci.* **2020**, *21*, 4198. [[CrossRef](#)]
- Panner Selvam, M.K.; Sengupta, P.; Agarwal, A. Sperm DNA Fragmentation and Male Infertility. In *Genetics of Male Infertility*; Springer International Publishing: Cham, The Switzerland, 2020; pp. 155–172. [[CrossRef](#)]
- Lettieri, G.; Marra, F.; Moriello, C.; Prisco, M.; Notari, T.; Trifuoggi, M.; Giarra, A.; Bosco, L.; Montano, L.; Piscopo, M. Molecular Alterations in Spermatozoa of a Family Case Living in the Land of Fires—A First Look at Possible Transgenerational Effects of Pollutants. *Int. J. Mol. Sci.* **2020**, *21*, 6710. [[CrossRef](#)]
- Sharpe, R.M. Environmental/Lifestyle Effects on Spermatogenesis. *Philos. Trans. R. Soc. B Biol. Sci.* **2010**, *365*, 1697–1712. [[CrossRef](#)]
- Kumar, N.; Singh, A. Trends of Male Factor Infertility, an Important Cause of Infertility: A Review of Literature. *J. Hum. Reprod. Sci.* **2015**, *8*, 191–196. [[CrossRef](#)]
- Sadeghi-Bazargani, H.; Hajshafiha; Ghareaghaji; Salemi; Sadeghi-asadi. Association of Body Mass Index with Some Fertility Markers among Male Partners of Infertile Couples. *Int. J. Gen. Med.* **2013**, *6*, 447–451. [[CrossRef](#)]
- Krausz, C.; Riera-Escamilla, A. Genetics of Male Infertility. *Nat. Rev. Urol.* **2018**, *15*, 369–384. [[CrossRef](#)] [[PubMed](#)]
- Dada, R.; Shamsi, M.; Kumar, K. Genetic and Epigenetic Factors: Role in Male Infertility. *Indian J. Urol.* **2011**, *27*, 110–120. [[CrossRef](#)] [[PubMed](#)]
- Wright, C.; Milne, S.; Leeson, H. Sperm DNA Damage Caused by Oxidative Stress: Modifiable Clinical, Lifestyle and Nutritional Factors in Male Infertility. *Reprod. BioMedicine Online* **2014**, *28*, 684–703. [[CrossRef](#)] [[PubMed](#)]
- Karam, Z.M.; Baba Salari, M.; Anjom Shooa, A.; Dehghan Kouhestani, S.; Bahram Nejad, A.; Ashourzadeh, S.; Zangouyee, M.R.; Bazrafshani, M.R. Impact of Oxidative Stress SNPs on Sperm DNA Damage and Male Infertility in a South-East Iranian Population. *Reprod. Fertil. Dev.* **2022**, *34*, 633–643. [[CrossRef](#)] [[PubMed](#)]
- Wouters-Tyrou, D.; Martinage, A.; Chevaillier, P.; Sautière, P. Nuclear Basic Proteins in Spermiogenesis. *Biochimie* **1998**, *80*, 117–128. [[CrossRef](#)]
- Sassone-Corsi, P. Unique Chromatin Remodeling and Transcriptional Regulation in Spermatogenesis. *Science* **2002**, *296*, 2176–2178. [[CrossRef](#)]
- Kimmins, S.; Sassone-Corsi, P. Chromatin Remodelling and Epigenetic Features of Germ Cells. *Nature* **2005**, *434*, 583–589. [[CrossRef](#)]
- Iguchi, N.; Yang, S.; Lamb, D.J.; Hecht, N.B. An SNP in Protamine 1: A Possible Genetic Cause of Male Infertility? *J. Med. Genet.* **2005**, *43*, 382–384. [[CrossRef](#)]
- Venkatesh, S.; Kumar, R.; Deka, D.; Deecaraman, M.; Dada, R. Analysis of Sperm Nuclear Protein Gene Polymorphisms and DNA Integrity in Infertile Men. *Syst. Biol. Reprod. Med.* **2011**, *57*, 124–132. [[CrossRef](#)]
- Guerrero-Bosagna, C.; Skinner, M.K. Environmentally Induced Epigenetic Transgenerational Inheritance of Male Infertility. *Curr. Opin. Genet. Dev.* **2014**, *26*, 79–88. [[CrossRef](#)]
- Raimondo, S.; Gentile, M.; Esposito, G.; Gentile, T.; Ferrara, I.; Crescenzo, C.; Palmieri, M.; Cuomo, F.; de Filippo, S.; Lettieri, G.; et al. Could Kallikrein-Related Serine Peptidase 3 Be an Early Biomarker of Environmental Exposure in Young Women? *Int. J. Environ. Res. Public Health* **2021**, *18*, 8833. [[CrossRef](#)] [[PubMed](#)]
- Ferlin, A.; Arredi, B.; Foresta, C. Genetic Causes of Male Infertility. *Reprod. Toxicol.* **2006**, *22*, 133–141. [[CrossRef](#)] [[PubMed](#)]
- Erenpreiss, J.; Elzanaty, S.; Giwercman, A. Sperm DNA Damage in Men from Infertile Couples. *Asian J. Androl.* **2008**, *10*, 786–790. [[CrossRef](#)] [[PubMed](#)]
- Oleszczuk, K.; Giwercman, A.; Bungum, M. Intra-Individual Variation of the Sperm Chromatin Structure Assay DNA Fragmentation Index in Men from Infertile Couples. *Hum. Reprod.* **2011**, *26*, 3244–3248. [[CrossRef](#)] [[PubMed](#)]
- Ji, G.; Long, Y.; Zhou, Y.; Huang, C.; Gu, A.; Wang, X. Common Variants in Mismatch Repair Genes Associated with Increased Risk of Sperm DNA Damage and Male Infertility. *BMC Med.* **2012**, *10*, 49. [[CrossRef](#)] [[PubMed](#)]
- Miyamoto, T.; Tsujimura, A.; Miyagawa, Y.; Koh, E.; Namiki, M.; Sengoku, K. Male Infertility and Its Causes in Human. *Adv. Urol.* **2012**, *2012*, 384520. [[CrossRef](#)]
- Krausz, C.; Giachini, C. Genetic Risk Factors in Male Infertility. *Arch. Androl.* **2007**, *53*, 125–133. [[CrossRef](#)]
- Lindgren, K.E.; Nordqvist, S.; Kårehed, K.; Sundström-Poromaa, I.; Åkerud, H. The Effect of a Specific Histidine-Rich Glycoprotein Polymorphism on Male Infertility and Semen Parameters. *Reprod. BioMedicine Online* **2016**, *33*, 180–188. [[CrossRef](#)]

31. Schulte, R.T.; Ohl, D.A.; Sigman, M.; Smith, G.D. Sperm DNA Damage in Male Infertility: Etiologies, Assays, and Outcomes. *J. Assist. Reprod. Genet.* **2010**, *27*, 3–12. [[CrossRef](#)]
32. Morris, I.D. The Spectrum of DNA Damage in Human Sperm Assessed by Single Cell Gel Electrophoresis (Comet Assay) and Its Relationship to Fertilization and Embryo Development. *Hum. Reprod.* **2002**, *17*, 990–998. [[CrossRef](#)]
33. Gandini, L.; Lombardo, F.; Paoli, D.; Caruso, F.; Eleuteri, P.; Leter, G.; Ciriminna, R.; Culasso, F.; Dondero, F.; Lenzi, A.; et al. Full-Term Pregnancies Achieved with ICSI despite High Levels of Sperm Chromatin Damage. *Hum. Reprod.* **2004**, *19*, 1409–1417. [[CrossRef](#)] [[PubMed](#)]
34. Li, Z.; Wang, L.; Cai, J.; Huang, H. Correlation of Sperm DNA Damage with IVF and ICSI Outcomes: A Systematic Review and Meta-Analysis. *J. Assist. Reprod. Genet.* **2006**, *23*, 367–376. [[CrossRef](#)] [[PubMed](#)]
35. Benchaib, M. Sperm DNA Fragmentation Decreases the Pregnancy Rate in an Assisted Reproductive Technique. *Hum. Reprod.* **2003**, *18*, 1023–1028. [[CrossRef](#)]
36. Larson, K.L. Sperm Chromatin Structure Assay Parameters as Predictors of Failed Pregnancy Following Assisted Reproductive Techniques. *Hum. Reprod.* **2000**, *15*, 1717–1722. [[CrossRef](#)]
37. Virro, M.R.; Larson-Cook, K.L.; Evenson, D.P. Sperm Chromatin Structure Assay (Sca[®]) Parameters Are Related to Fertilization, Blastocyst Development, and Ongoing Pregnancy in in Vitro Fertilization and Intracytoplasmic Sperm Injection Cycles. *Fertil. Steril.* **2004**, *81*, 1289–1295. [[CrossRef](#)] [[PubMed](#)]
38. Simon, L.; Brunborg, G.; Stevenson, M.; Lutton, D.; McManus, J.; Lewis, S.E.M. Clinical Significance of Sperm DNA Damage in Assisted Reproduction Outcome. *Hum. Reprod.* **2010**, *25*, 1594–1608. [[CrossRef](#)] [[PubMed](#)]
39. Hu, Z.; Li, Z.; Yu, J.; Tong, C.; Lin, Y.; Guo, X.; Lu, F.; Dong, J.; Xia, Y.; Wen, Y.; et al. Association Analysis Identifies New Risk Loci for Non-Obstructive Azoospermia in Chinese Men. *Nat. Commun.* **2014**, *5*, 3857. [[CrossRef](#)]
40. Lee, Y.H.; Kim, J.-H.; Song, G.G. Pathway Analysis of a Genome-Wide Association Study in Schizophrenia. *Gene* **2013**, *525*, 107–115. [[CrossRef](#)]
41. Mortezaei, Z.; Tavallaee, M. Recent Innovations and In-Depth Aspects of Post-Genome Wide Association Study (Post-GWAS) to Understand the Genetic Basis of Complex Phenotypes. *Heredity* **2021**, *127*, 485–497. [[CrossRef](#)]
42. Jia, P.; Wang, L.; Meltzer, H.Y.; Zhao, Z. Pathway-Based Analysis of GWAS Datasets: Effective but Caution Required. *Int. J. Neuropsychopharmacol.* **2011**, *14*, 567–572. [[CrossRef](#)]
43. Yoon, S.; Nguyen, H.C.T.; Yoo, Y.J.; Kim, J.; Baik, B.; Kim, S.; Kim, J.; Kim, S.; Nam, D. Efficient Pathway Enrichment and Network Analysis of GWAS Summary Data Using GSA-SNP2. *Nucleic Acids Res.* **2018**, *46*, e60. [[CrossRef](#)] [[PubMed](#)]
44. Matern, B.M.; Olieslagers, T.I.; Voorter, C.E.M.; Groeneweg, M.; Tilanus, M.G.J. Insights into the Polymorphism in HLA—DRA and Its Evolutionary Relationship with HLA Haplotypes. *HLA* **2020**, *95*, 117–127. [[CrossRef](#)] [[PubMed](#)]
45. Mosaad, Y.M. Clinical Role of Human Leukocyte Antigen in Health and Disease. *Scand. J. Immunol.* **2015**, *82*, 283–306. [[CrossRef](#)] [[PubMed](#)]
46. Jones, N.R.; Sun, D.; Freeman, W.M.; Lazarus, P. Quantification of Hepatic UDP Glucuronosyltransferase 1A Splice Variant Expression and Correlation of UDP Glucuronosyltransferase 1A1 Variant Expression with Glucuronidation Activity. *J. Pharmacol. Exp. Ther.* **2012**, *342*, 720–729. [[CrossRef](#)]
47. Zhang, S.; Wu, Z.; Chang, W.; Liu, F.; Xie, J.; Yang, Y.; Qiu, H. Classification of Patients With Sepsis According to Immune Cell Characteristics: A Bioinformatic Analysis of Two Cohort Studies. *Front. Med.* **2020**, *7*, 598652. [[CrossRef](#)]
48. Shiina, T.; Hosomichi, K.; Inoko, H.; Kulski, J.K. The HLA Genomic Loci Map: Expression, Interaction, Diversity and Disease. *J. Hum. Genet.* **2009**, *54*, 15–39. [[CrossRef](#)]
49. Muñoz-Carrillo, J.L.; Castro-García, F.P.; Chávez-Rubalcaba, F.; Chávez-Rubalcaba, I.; Martínez-Rodríguez, J.L.; Hernández-Ruiz, M.E. Immune System Disorders: Hypersensitivity and Autoimmunity. In *Immunoregulatory Aspects of Immunotherapy*; InTech: Vienna, Austria, 2018. [[CrossRef](#)]
50. Ślebioda, T.J.; Kmieć, Z. Tumour Necrosis Factor Superfamily Members in the Pathogenesis of Inflammatory Bowel Disease. *Mediat. Inflamm.* **2014**, *2014*, 325129. [[CrossRef](#)]
51. Wickham, H. *Ggplot2*; Springer New York: New York, NY, USA, 2009. [[CrossRef](#)]
52. Csardi, G.; Nepusz, T. The Igraph Software Package for Complex Network Research. *Inter J. Complex Syst.* **2006**, *1695*, 1–9.
53. Drożdżik, M.; Kaczmarek, M.; Malinowski, D.; Broś, U.; Kazienko, A.; Kurzawa, R.; Kurzawski, M. TGFβ3 (TGFB3) Polymorphism Is Associated with Male Infertility. *Sci. Rep.* **2015**, *5*, 17151. [[CrossRef](#)]
54. Vendrell, X. New Genetic Point Mutations in Male Infertility. In *Reproductomics*; Elsevier: Amsterdam, The Netherlands, 2018; pp. 47–62. [[CrossRef](#)]
55. Yin, Y.; Zhu, P.; Luo, T.; Xia, X. Association of Single-nucleotide Polymorphisms in Antioxidant Genes and Their Gene-gene Interactions with Risk of Male Infertility in a Chinese Population. *Biomed. Rep.* **2020**, *13*, 49–54. [[CrossRef](#)]
56. van der Ven, K.; Fimmers, R.; Engels, G.; van der Ven, H.; Krebs, D. Evidence for Major Histocompatibility Complex-Mediated Effects on Spermatogenesis in Humans. *Hum. Reprod.* **2000**, *15*, 189–196. [[CrossRef](#)] [[PubMed](#)]
57. Zhao, H.; Xu, J.; Zhang, H.; Sun, J.; Sun, Y.; Wang, Z.; Liu, J.; Ding, Q.; Lu, S.; Shi, R.; et al. A Genome-Wide Association Study Reveals That Variants within the HLA Region Are Associated with Risk for Nonobstructive Azoospermia. *Am. J. Hum. Genet.* **2012**, *90*, 900–906. [[CrossRef](#)] [[PubMed](#)]
58. Stastny, P. HLA-D and IA Antigens in Rheumatoid Arthritis and Systemic Lupus Erythematosus. *Arthritis Rheum.* **1978**, *21*, S139–S143. [[CrossRef](#)]

59. Mignot, E.; Hayduk, R.; Black, J.; Grumet, F.C.; Guilleminault, C. HLA DQB1*0602 Is Associated with Cataplexy in 509 Narcoleptic Patients. *Sleep* **1997**, *20*, 1012–1020. [[PubMed](#)]
60. Hamza, T.H.; Zabetian, C.P.; Tenesa, A.; Laederach, A.; Montimurro, J.; Yearout, D.; Kay, D.M.; Doheny, K.F.; Paschall, J.; Pugh, E.; et al. Common Genetic Variation in the HLA Region Is Associated with Late-Onset Sporadic Parkinson's Disease. *Nat. Genet.* **2010**, *42*, 781–785. [[CrossRef](#)] [[PubMed](#)]
61. Blanco-Kelly, F.; Alvarez-Lafuente, R.; Alcina, A.; Abad-Grau, M.M.; de las Heras, V.; Lucas, M.; de la Concha, E.G.; Fernández, O.; Arroyo, R.; Matesanz, F.; et al. Members 6B and 14 of the TNF Receptor Superfamily in Multiple Sclerosis Predisposition. *Genes Immun.* **2011**, *12*, 145–148. [[CrossRef](#)]
62. Baranzini, S.E.; Khankhanian, P.; Patsopoulos, N.A.; Li, M.; Stankovich, J.; Cotsapas, C.; Søndergaard, H.B.; Ban, M.; Barizzone, N.; Bergamaschi, L.; et al. Network-Based Multiple Sclerosis Pathway Analysis with GWAS Data from 15,000 Cases and 30,000 Controls. *Am. J. Hum. Genet.* **2013**, *92*, 854–865. [[CrossRef](#)] [[PubMed](#)]
63. Dubois, P.C.A.; Trynka, G.; Franke, L.; Hunt, K.A.; Romanos, J.; Curtotti, A.; Zhernakova, A.; Heap, G.A.R.; Ádány, R.; Aromaa, A.; et al. Multiple Common Variants for Celiac Disease Influencing Immune Gene Expression. *Nat. Genet.* **2010**, *42*, 295–302. [[CrossRef](#)]
64. Ruysen-Witrand, A.; Constantin, A.; Cambon-Thomsen, A.; Thomsen, M. New Insights into the Genetics of Immune Responses in Rheumatoid Arthritis. *Tissue Antigens* **2012**, *80*, 105–118. [[CrossRef](#)]
65. Cheung, T.C.; Steinberg, M.W.; Osborne, L.M.; Macauley, M.G.; Fukuyama, S.; Sanjo, H.; D'Souza, C.; Norris, P.S.; Pfeffer, K.; Murphy, K.M.; et al. Unconventional Ligand Activation of Herpesvirus Entry Mediator Signals Cell Survival. *Proc. Natl. Acad. Sci. USA* **2009**, *106*, 6244–6249. [[CrossRef](#)]
66. Tukey, R.H.; Strassburg, C.P. Human UDP-Glucuronosyltransferases: Metabolism, Expression, and Disease. *Annu. Rev. Pharmacol. Toxicol.* **2000**, *40*, 581–616. [[CrossRef](#)] [[PubMed](#)]
67. Nandith, P.B.; Adiga, U.; Shenoy, V.; Adiga, M.N.S. UGT1A6 and UGT2B7 Gene Polymorphism and Its Effect in Pediatric Epileptic Patients on Sodium Valproate Monotherapy. *Indian J. Pediatrics* **2021**, *88*, 764–770. [[CrossRef](#)] [[PubMed](#)]
68. Kua, L.-F.; Ross, S.; Lee, S.-C.; Mimura, K.; Kono, K.; Goh, B.-C.; Yong, W.-P. UGT1A6 Polymorphisms Modulated Lung Cancer Risk in a Chinese Population. *PLoS ONE* **2012**, *7*, e42873. [[CrossRef](#)] [[PubMed](#)]
69. Justenhoven, C.; Winter, S.; Dünnebier, T.; Hamann, U.; Baisch, C.; Rabstein, S.; Spickenheuer, A.; Harth, V.; Pesch, B.; Brüning, T.; et al. Combined UGT1A1 and UGT1A6 Genotypes Together with a Stressful Life Event Increase Breast Cancer Risk. *Breast Cancer Res. Treat.* **2010**, *124*, 289–292. [[CrossRef](#)]