



Article

Network Analysis of Biomarkers Associated with Occupational Exposure to Benzene and Malathion

Marcus Vinicius C. Santos ^{1,*} , Arthur S. Feltrin ², Isabele C. Costa-Amaral ¹, Liliane R. Teixeira ¹,
Jamila A. Perini ³, David C. Martins, Jr. ² and Ariane L. Larentis ^{1,*}

¹ Studies Center of Worker's Health and Human Ecology (CESTEH), Sergio Arouca National School of Public Health (ENSP), Oswaldo Cruz Foundation (FIOCRUZ), Rio de Janeiro 21041-210, RJ, Brazil

² Center for Mathematics, Computation and Cognition, Federal University of ABC, Santo André 09210-580, SP, Brazil

³ Research Laboratory of Pharmaceutical Sciences (LAPESF), State University of Rio de Janeiro (West Zone-UERJ-ZO), Rio de Janeiro 23070-200, RJ, Brazil

* Correspondence: marcus.santos@ensp.fiocruz.br (M.V.C.S.); ariane.larentis@fiocruz.br (A.L.L.)

Abstract: Complex diseases are associated with the effects of multiple genes, proteins, and biological pathways. In this context, the tools of Network Medicine are compatible as a platform to systematically explore not only the molecular complexity of a specific disease but may also lead to the identification of disease modules and pathways. Such an approach enables us to gain a better understanding of how environmental chemical exposures affect the function of human cells, providing better perceptions about the mechanisms involved and helping to monitor/prevent exposure and disease to chemicals such as benzene and malathion. We selected differentially expressed genes for exposure to benzene and malathion. The construction of interaction networks was carried out using GeneMANIA and STRING. Topological properties were calculated using MCODE, BiNGO, and CentiScaPe, and a Benzene network composed of 114 genes and 2415 interactions was obtained. After topological analysis, five networks were identified. In these subnets, the most interconnected nodes were identified as: IL-8, KLF6, KLF4, JUN, SERTAD1, and MT1H. In the Malathion network, composed of 67 proteins and 134 interactions, HRAS and STAT3 were the most interconnected nodes. Path analysis, combined with various types of high-throughput data, reflects biological processes more clearly and comprehensively than analyses involving the evaluation of individual genes. We emphasize the central roles played by several important hub genes obtained by exposure to benzene and malathion.

Keywords: network medicine; occupational health; malathion; benzene



Citation: Santos, M.V.C.; Feltrin, A.S.; Costa-Amaral, I.C.; Teixeira, L.R.; Perini, J.A.; Martins, D.C., Jr.; Larentis, A.L. Network Analysis of Biomarkers Associated with Occupational Exposure to Benzene and Malathion. *Int. J. Mol. Sci.* **2023**, *24*, 9415. <https://doi.org/10.3390/ijms24119415>

Academic Editors: Youping Deng and Kurt K. Zhang

Received: 1 March 2023

Revised: 21 April 2023

Accepted: 3 May 2023

Published: 28 May 2023



Copyright: © 2023 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/>).

1. Introduction

Complex diseases are associated with the effects of multiple genes, proteins, and biological pathways [1]. Therefore, in the case of a complex disease, one should not expect that a single genetic mutation can be identified as the cause. In fact, complex diseases or disorders (e.g., cancer, AIDS, and obesity) stem from dysfunctions of different biomolecular networks and not only their isolated components (e.g., genes, proteins, and metabolites) [2].

Biological networks are powerful resources for the discovery of genes and genetic modules that drive disease. Biomolecular networks include gene and transcription regulatory networks, protein-protein interaction networks, metabolic, signaling, and hybrid networks. With advances in high-throughput measurement techniques such as microarray, RNA-seq, CHIP-on-chip, and mass spectrometry, large-scale biological datasets have been continuously produced. Such data contain detailed information to understand the mechanism of molecular biological systems and have proven to be useful in the diagnosis, treatment, and design of drugs for complex diseases or disorders [3].

In this context, the Network Medicine [4] aims to use complex network analysis to find clusters (or modules) in a biological network that could be related to a phenotype of interest

based on a few hypotheses, such as the hypothesis that a protein-protein interaction (PPI) network follows a scale-free (power law) degree distribution and that genes associated with the same (or similar) pathways are clustered closely in the PPI networks; therefore, they have a high probability of interacting with each other. This is a fundamental concept that can be used to combine and amplify signals from individual genes [5], genes with similar expression patterns [6], synthetic lethality [7], or chemical sensitivity [8], which often present similar functions. In addition, genes whose products interact physically [4,9] are part of the same complex [10], display similar three-dimensional structures [11], similar phylogenetic profiles [12], or have common protein domains [13]. Therefore, a biological pathway plays an important role in understanding the mechanisms of complex diseases, improving clinical treatment, and revealing drug targets and biomarkers [14]. Such an approach opens the possibility of a better understanding of how environmental chemical exposures affect the function of human cells, providing better insights into the mechanisms involved and assisting in the monitoring/prevention of exposure [15]. For instance, benzene is classified by the International Agency for Research on Cancer (IARC) as a carcinogen belonging to group I, i.e., carcinogenic to humans [16]. Exposure to benzene is associated with the occurrence of hematotoxicity, acute myeloid leukemia, and myelodysplastic syndromes [17–21].

Hematotoxicity effects are observed even at relatively low concentrations [22–25], although the hematopoietic toxicity mechanisms of action in benzene exposure remain unclear and are still under study, mainly by applying toxicogenomics techniques [26–28].

As is well known, benzene metabolism creates many reactive elements [16,29], and exposure to benzene and its metabolites can generate DNA mutations, chromosome insertions and/or deletions, DNA double-strand breaks, apoptosis, oxidative stress, and altered gene expression [28].

Another relevant environmental chemical exposure is related to malathion, which is a likely carcinogen due to the increased risk of cancer associated with its exposure [30,31]. Farm workers and their children, particularly, face an increased risk of developing leukemia and non-Hodgkin's lymphoma due to their exposure to malathion [31–33]. In fact, studies have shown that malathion induces chromosomal and DNA damage in humans [34,35]. Thus, the IARC—a specialized cancer agency of the World Health Organization (WHO)—also classified malathion as 'probably carcinogenic to humans' (Group 2A). However, molecular changes caused by exposure to malathion have not been extensively explored to date, although neurological malignancies are prevalent in humans.

The biological processes and molecular functions underlying such exposures constitute complex systems [36], which cannot always be designated by a simplistic view such as assigning functions to individual genes, proteins, and other cell macromolecules [37]. In this context, several studies have presented methods to analyze and identify essential genes and proteins in a biological interaction network. Luo et al. [38] proposed a method to predict essential proteins in PPI networks based on local interaction density and protein complexes. Wang et al. [39] proposed a method to identify essential proteins by combining information about protein complexes and topological features of the PPI network. Hu et al. [40] address a method to be applied to weighted networks by considering the total strength of the interaction, the number of edges of the interaction, and the distribution of the total strength of the connection at the edge of the connection in the local domain. Cinaglia and Cannataro [41] and Dai et al. [42] address the use of static network alignment methods adapted to the dynamic context by performing network discrimination and providing other additional information. By reporting how various problems can be transferred from static networks to dynamic networks, taking into account temporal information. In addition, many of the key factors for measuring nodes in complex networks are based on graph theory to quantify the topological structure and attributes of each node, and comparisons of the centrality of each node are made by using different centrality calculation methods, such as degree center, median center, proximity center, and edge clustering coefficient center. Quantitative methods are also used to find the essential nodes in networks [43–46].

Considering these issues, this study adopted a systems biology approach to reconstruct networks of molecular interactions based on differentially expressed genes (DEG) related to benzene and malathion exposure, retrieved from literature text mining. Using some of the Network Medicine hypotheses, such as the disease module and the local hypotheses, our goal is to use network analysis to evaluate possible biological pathways associated with the response to benzene and malathion. Gene interactions were observed using the GeneMANIA software (version 3.5.2), while interactions between proteins were detected using the Search Tool for Retrieval of Interacting Genes (STRING) database (version 11.0). The Molecular Complex Detection (MCODE) software (version 2.0.2) was used for the characterization of clusters of different biological processes, and the Biological Networks Gene Ontology (BiNGO) software (version 3.0.5) was used for ontological gene characterization. The Cytoscape software (version 3.6.0) was used to calculate different network centrality metrics from these genes.

2. Results

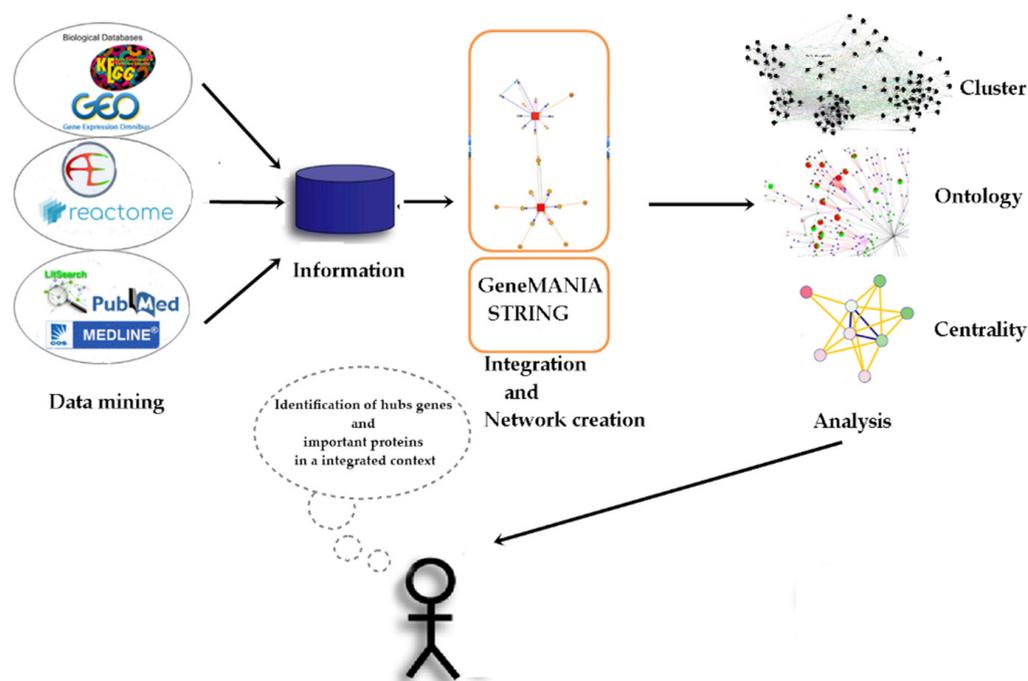
To create, analyze, and select the hub genes related to exposure to benzene and malathion, a total of 96 human DEGs were selected through a bibliographic search for articles published between 2010 and 2014, dealing with quantitative microarray data for environmental and occupational exposure to benzene (Table 1). Additionally, 57 human DEGs were selected through a bibliographic search for articles published with quantitative microarray data for malathion exposure (Table 2). Figure 1 shows the entire experimental design, and a list of all genes selected for benzene and malathion bibliographic searches, in addition to the list of genes added outside the list in the network creation stage by STRING and GeneMANIA, is present in Supplementary Table S1.

Table 1. Benzene list of the most important studies obtained herein and their selected genes.

Reference	Method	Exposure	Controls	DEGs
McHale et al. [47]	Microarray	83 cases of benzene exposure ranging from <<1 to >10 ppm.	42	16 genes with high expression (<i>SERPINB2</i> , <i>TNFAIP6</i> , <i>IL1A</i> , <i>KCNJ2</i> , <i>PTX3</i> , <i>F3</i> , <i>CD44</i> , <i>CCL20</i> , <i>ACSL1</i> , <i>PTGS2</i> , <i>CLEC5A</i> , <i>IL1RN</i> , <i>PRG2</i> , <i>SLC2A6</i> , <i>GPR132</i> , and <i>PLAUR</i>).
Bi et al. [48]	cDNA microarray	7 women were diagnosed with benzene poisoning.	7	Top 40 genes with altered expression (<i>PTGS2</i> , <i>BAI3</i> , <i>GCL</i> , <i>CYP4F3</i> , <i>MYO47</i> , <i>TRA@</i> , <i>AD022</i> , <i>PRKCH</i> , <i>RASGRP1</i> , <i>FPR1</i> , <i>TGFBR3</i> , <i>GRO1</i> , <i>SEL1L</i> , <i>CSF2RB</i> , <i>IFITM1</i> , <i>STAT4</i> , <i>IFITM2</i> , <i>ABLIM</i> , <i>KIAA1382</i> , <i>SPTBN1</i> , <i>HBB</i> , <i>PRKDC</i> , <i>S100A10</i> , <i>ITGB2</i> , <i>TKT</i> , <i>VAMP8</i> , <i>FOSB</i> , <i>ASAH1</i> , <i>CDC37</i> , <i>SLC25A6</i> , <i>CLN2</i> , <i>ACTA2</i> , <i>CST3</i> , <i>HLA-DMB</i> , <i>ALDH2</i> , <i>LGALS2</i> , <i>LGALS1</i> , <i>ARHB</i> , <i>KLF4</i> , and <i>ATF3</i>).
Xing et al. [49]	Microarray (RT-PCR)	11	People in the same sector with no symptoms of benzene poisoning.	Decrease in the expression of p15 (<i>CDKN2B</i>) and p16 (<i>CDKN2A</i>).
Sarma et al. [50]	Microarray	Culture of HL-60 cells treated with IC50 concentrations of benzene, hydroquinone, and benzoquinone.	Culture of HL-60 cells treated with dimethyl sulfoxide.	Alteration in expression of 27 genes (<i>CCL2</i> , <i>EGR1</i> , <i>GCLM</i> , <i>PMAIP1</i> , <i>SESN2</i> , <i>CD69</i> , <i>HERPUD1</i> , <i>HSPA8</i> , <i>RIT1</i> , <i>SERTAD1</i> , <i>SLC38A2</i> , <i>SLC7A11</i> , <i>DNAJB4</i> , <i>ANKFY1</i> , <i>ANLN</i> , <i>AR</i> , <i>ARHGAP19</i> , <i>CDCA2</i> , <i>DEPDC1</i> , <i>ELK1</i> , <i>FBXW9</i> , <i>HERC2</i> , <i>HTR5A</i> , <i>KIF20A</i> , <i>MKI67</i> , <i>MT1G</i> , and <i>MT1X</i>).
Gao et al. [51]	cDNA microarray	4 people were diagnosed with benzene poisoning, and 3 people from the same factory were exposed but had no symptoms.	3	Top 14 significant genes with altered expression (<i>PIK3R1</i> , <i>PIK3CG</i> , <i>PIK3R2</i> , <i>GNAI3</i> , <i>SYK</i> , <i>PTPN6</i> , <i>KRAS</i> , <i>NRAS</i> , <i>PLCG2</i> , <i>NFKB1</i> , <i>LYN</i> , <i>SOCS4</i> , <i>HLA-DMA</i> , and <i>HLA-DMB</i>).

Table 2. Malathion list of the most important studies obtained herein and their selected genes.

Reference	Method	Exposure	Controls	DEGs
Anjitha et al. [52]	Microarray	Culture of human lymphocytes treated with three concentrations of malathion (50, 100, and 150 µg/mL).	Culture of human lymphocytes treated with DMSO (1%).	57 DEGs (<i>B4GALT1</i> , <i>BMI1</i> , <i>BTG1</i> , <i>C1QBP</i> , <i>CASP4</i> , <i>CBFB</i> , <i>CD14</i> , <i>CD5</i> , <i>CD36</i> , <i>CDK2</i> , <i>CEBPB</i> , <i>COL1A2</i> , <i>DDX11</i> , <i>DUSP1</i> , <i>EPHA4</i> , <i>EPS8L1</i> , <i>FGF6</i> , <i>FGFR1</i> , <i>FOXO4</i> , <i>FSHR</i> , <i>GNAI2</i> , <i>GPS2</i> , <i>GRAMD4</i> , <i>GSTP1</i> , <i>HLA-A</i> , <i>HLA-E</i> , <i>HTATIP2</i> , <i>IL13RA2</i> , <i>IRAK1</i> , <i>LGALS1</i> , <i>LPCAT4</i> , <i>LZTS2</i> , <i>MAP2K3</i> , <i>MICA</i> , <i>NINJ1</i> , <i>NME2</i> , <i>NTRK2</i> , <i>PAX1</i> , <i>PEBP1</i> , <i>PFN1</i> , <i>PHB</i> , <i>PLAU</i> , <i>PML</i> , <i>PRKD1</i> , <i>RAB11FIP3</i> , <i>SHC1</i> , <i>SOCS3</i> , <i>TEK</i> , <i>TFRC</i> , <i>TLR8</i> , <i>TPR</i> , <i>TRIM35</i> , <i>TUSC2</i> , <i>TWIST1</i> , <i>VHL</i> , <i>VPREB3</i> , and <i>WT1</i>)

**Figure 1.** Main stages of the experimental project.

2.1. Interaction Network

The analysis of biological interactions by GeneMANIA revealed a network comprising 114 genes and 2415 interactions (Figure 2A) and a PPI network comprising 67 nodes and 134 edges by STRING.

The gene-gene network presented a predominance of co-expression interactions (Figure 2B), followed by physical interactions (Figure 2C), molecular pathways (Figure 2D), predicted (Figure 2E), co-localization (Figure 2F), genetic interactions (Figure 2G), and shared protein domains (Figure 2H), respectively.

2.2. Cluster Analysis

The MCODE software (version 2.0.2) classified and separated molecular aggregates into clusters. The cluster analysis identified five sub-networks (Figure 3), ranked based on node connectivity level (score).

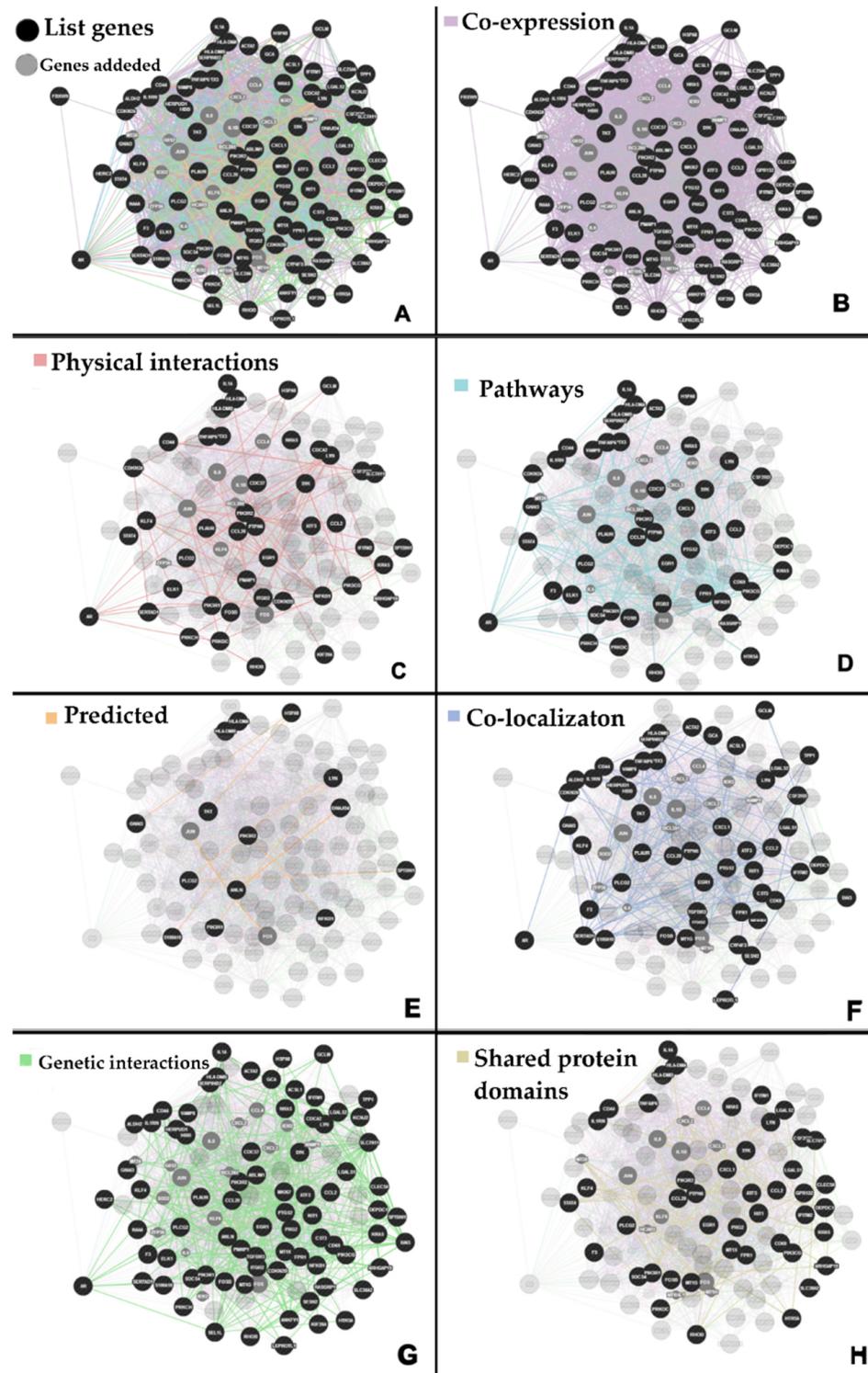


Figure 2. Benzene biological interaction network. (A) Interaction network generated using GeneMania with 114 genes comprising 2415 interactions; (B) co-expression interactions; (C) physical interactions; (D) pathways; (E) predicted; (F) co-localization; (G) genetic interactions; and (H) shared protein domains.

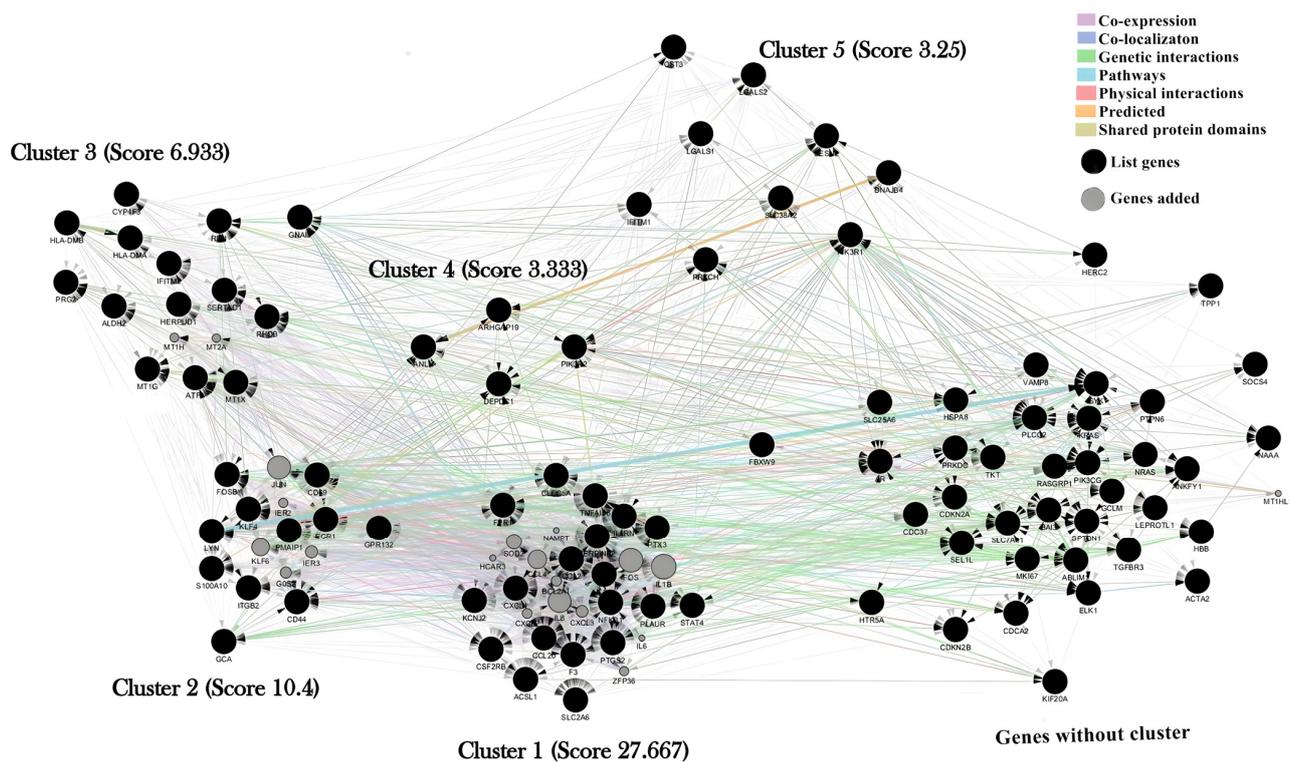


Figure 3. Overview of the five clusters obtained from the benzene gene-gene network using the MCODE software. Interaction network generated using GeneMANIA with 114 nodes (8 nodes expanded from a 96-gene list) and 2415 edges. All nodes are interconnected in a unique connected component, but 38 nodes do not belong to any cluster: Cluster 1: score = 27.667; Cluster 2: score = 10.4; Cluster 3: score = 6.933; Cluster 4: score = 3.333; and Cluster 5: score = 3.25. Nodes may represent biological elements, while the edges describe the nature of their relationships (co-expression; physical; pathways; predicted; co-localization; Genetic interactions; and shared protein domains).

2.3. Centiscape Analysis

In order to search for network biomarkers, the structure and manner of information flows (connectivity) along the networks were evaluated (Figures 4 and 5 and Table 3).

Table 3. Most relevant nodes for Benzene and Malathion networks.

GeneMania Benzene Network		
Symbol	Name	HGNC-ID
KLF4	KLF transcription factor 4	6348
SERTAD1	SERTA domain containing 1	17932
IL8	Interleukin 8	6025
JUN	Jun proto-oncogene, AP-1 transcription factor subunit	6204
KLF6	KLF transcription factor 6	2235
MT1H	Metallothionein 1H	7400
String Malathion		
Symbol	Name	HGNC-ID
HRAS	Hras proto-oncogene, GTPase	5173
STAT3	Signal transducer and activator of transcription 3	11364

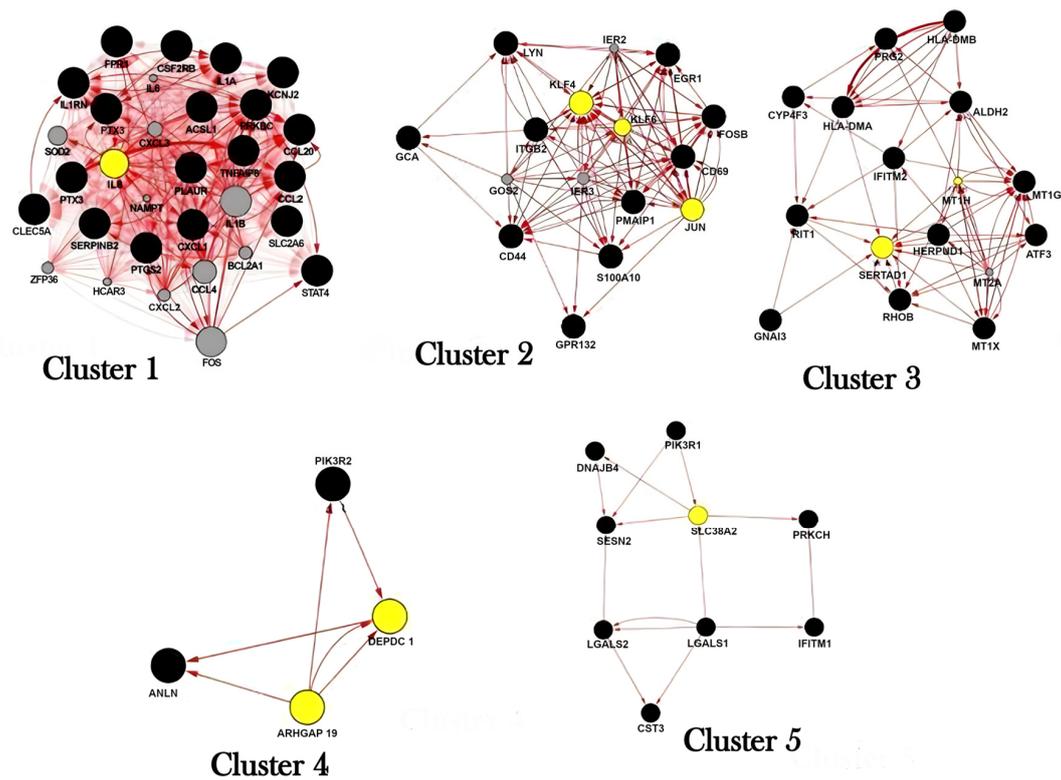


Figure 4. Most relevant nodes (in yellow) for the subnetworks predicted by the MCODE analysis in the benzene gene-gene network. Centrality is calculated by node degree, betweenness, and eigenvector. Cluster 1, IL8 was identified as the bottle-neck node; Cluster 2, KLF4, KLF6, and JUN were identified as the bottleneck nodes; Cluster 3, SERTAD1 and MT1H were identified as bottleneck nodes and Clusters 5 DEPDC1 and ARHGAP19 were identified as bottleneck nodes.

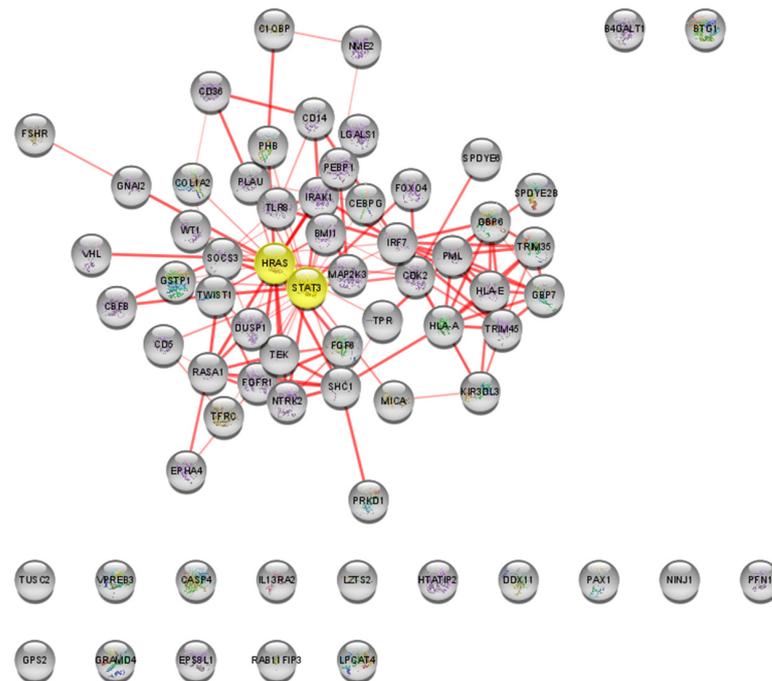


Figure 5. Most relevant nodes (in yellow) for Malathion PPI network analysis in STRING. Centrality is calculated by node degree, betweenness, and eigenvector. Interaction network generated using STRING with 67 nodes (10 nodes expanded from a 57-gene list) and 134 edges. Fifty nodes are interconnected in the main connected component, while 17 nodes are isolated.

For Benzene network exposure, in the 1st cluster, *IL8* was identified as the bottleneck node with degree = 34.12903, betweenness = 3.03225, and eigenvector = 0.18904. In the 2nd cluster, *KLF4*, *KLF6*, and *JUN* were identified as the bottleneck nodes with degree = 12.625, betweenness = 5.25, and eigenvector = 0.23.

In the 3rd cluster, *SERTAD1* and *MT1H* were identified as bottleneck nodes, with degree = 8.25, betweenness = 9.125, and eigenvector = 0.23. The 4th and 5th clusters are poorly interconnected, with central nodes in *DEPDC1* and *ARHGAP19* (degree = 3; betweenness = 0.5, and eigenvector = 0.49) and *SLC38A2* (degree = 3; betweenness = 6.6, and eigenvector = 0.5), respectively (Figure 4).

In the PPI network, *HRAS* and *STAT3* were identified as the bottleneck nodes with degree = 5.5, betweenness = 52.95, and eigenvector = 0.28 (Figure 5).

2.4. GO Overrepresentation Analysis (BiNGO)

Of total 390 ontologies (Table S2) regulated by DEGs in the Benzene network, the GO (Gene Ontology) enrichment analysis identified up-frequency of Overrepresentation (adjusted p value < 0.05) in Biological Processes for “biological regulation”, “regulation of the biological process”, and “regulation of the cellular process”, while down-frequency of Overrepresentation in Biological Processes related to the “positive regulation of nitric oxide biosynthetic process”, “regulation of vascular endothelial growth factor production”, “cytokine-mediated signaling pathway”, “leukocyte migration”, “leukocyte chemotaxis”, “cell chemotaxis”, and the most significant terms were “response to chemical stimulus”, indicating that the regulation of factors related to cellular locomotion, mainly leukocytes, plays a central role on the relationship between benzene and the development of leukemia (Figure 6).

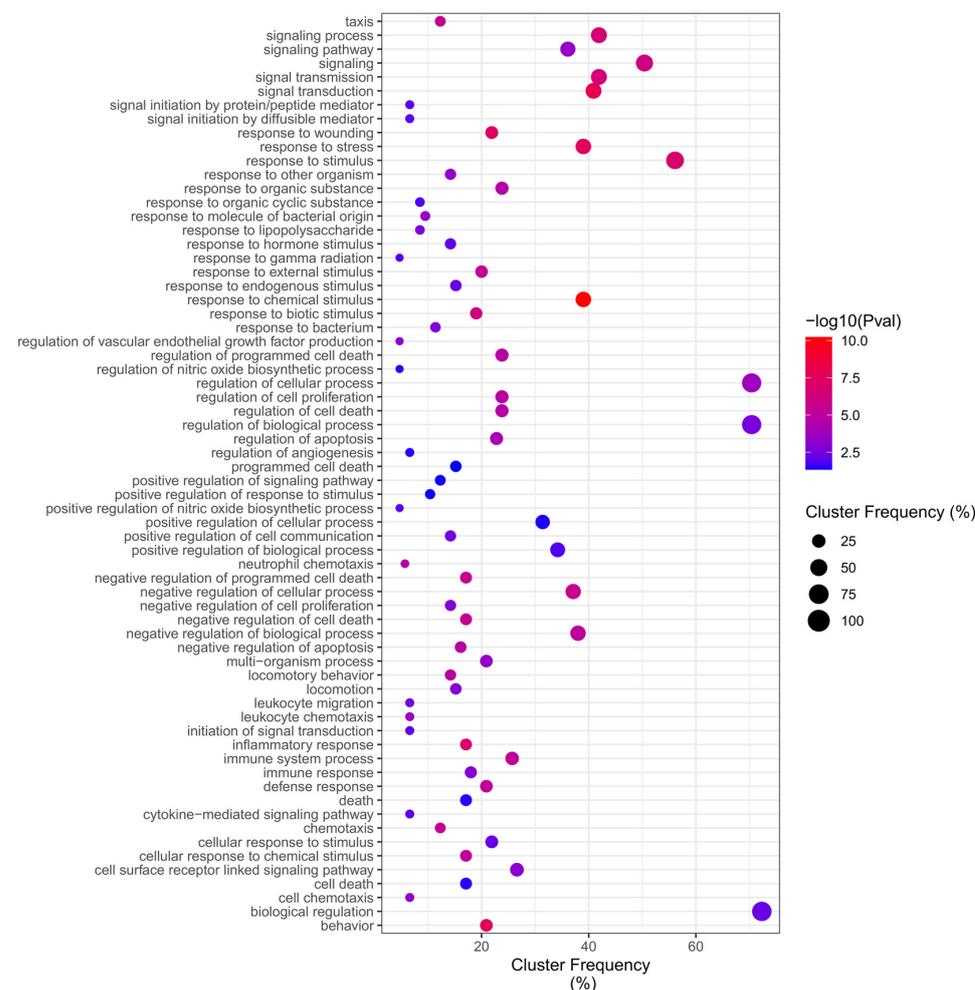


Figure 6. Benzene Biological Process Overrepresentation Analysis (BiNGO).

Of total 314 ontologies (Table S3) regulated by DEGs in the Malathion network, the GO enrichment analysis identified up-frequency of Overrepresentation (adjusted p value < 0.05) in Biological Processes for “regulation of cellular process”, “regulation of biological process”, and “biological regulation”, while down-frequency of Overrepresentation in Biological Processes related to “negative regulation of apoptosis”, “negative regulation of cell death”, “negative regulation of programmed cell death”, “transmembrane receptor protein tyrosine kinase signaling pathway”, and the most significant terms were “positive regulation of biological process”, indicating more general biological functions and cell death regulation (Figure 7).

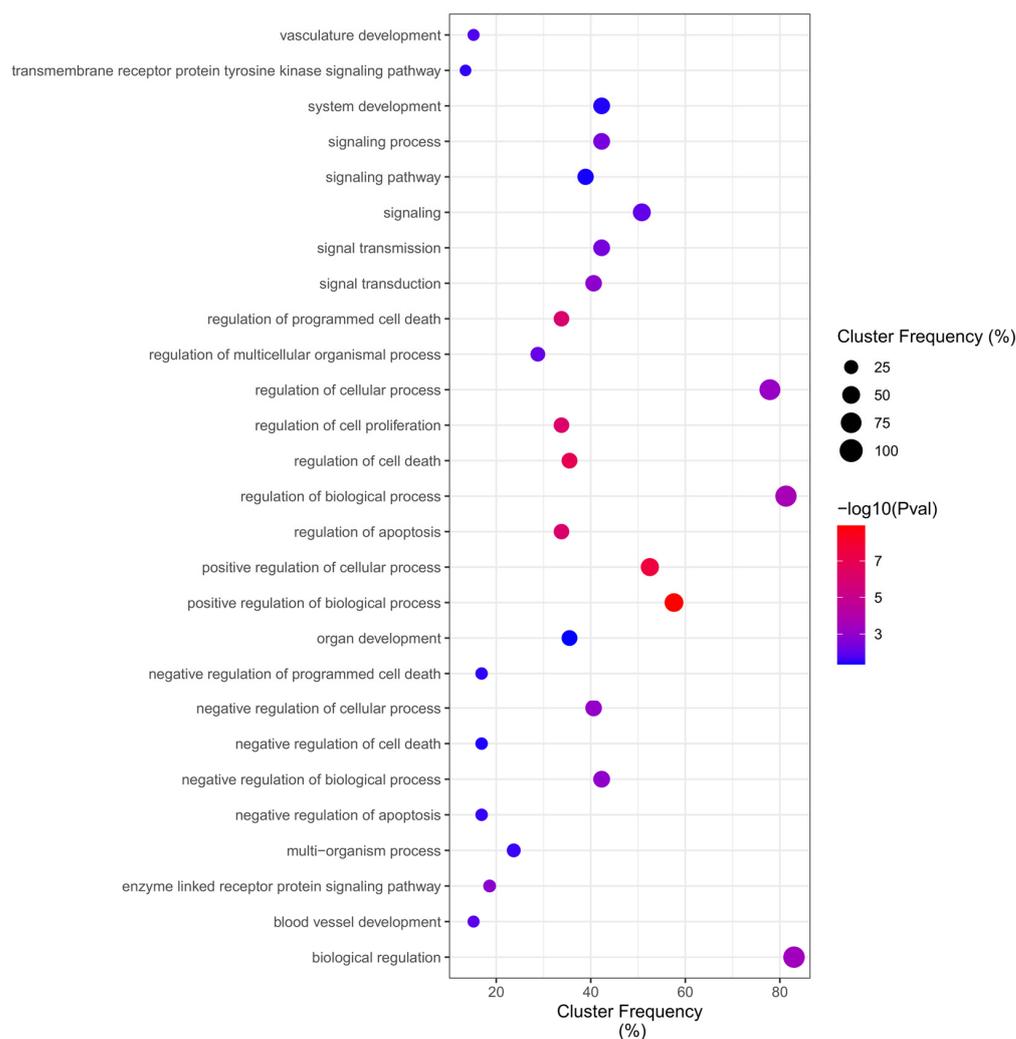


Figure 7. Malathion Biological Process Overrepresentation Analysis (BiNGO).

3. Discussion

STRING and GeneMANIA are two well-known and simple-to-use web tools, with no need for the user to have advanced knowledge in programming to use them. When compared to other tools such as the Alignment-Based Network Construction Algorithm (ANCA) [53], their simple interface, web access, and the possibility that they can produce a list of associated genes from a query based on several biological associations make these tools more attractive for users.

These two tools provide the same service for four biological associations (physical interaction, genetic interaction, co-expression, and co-citation), but other associations are unique to a particular tool. STRING, for example, has three unique biological associations (gene fusion evidence, co-occurrence, and pathway evidence), as well as GeneMANIA (co-inheritance, colocalization, and shared protein domains).

STRING provides a limited variety of association network data to be used for a given query. In contrast, GeneMANIA generates customized results for the consulted genes and the data sources selected by the user. However, GeneMANIA needs extensive literature to filter the results (e.g., the co-expression data, which is a database of cancer genomes), while STRING may be more suitable for cases where not so much data is available. This is precisely why STRING was adopted for malathion DEGs, since not many data exist for humans (according to IARC classification [8], malathion is considered 2A—“probably carcinogenic to humans”).

Among the advantages observed when using GeneMANIA is the potential to expand the search for related genes, which in this study represented 96 genes obtained from the literature, for a final network comprising 114 genes (Table S1). Some of the added genes also show altered expression upon benzene exposure (such as *JUN*, *KLF6*, *KFP36*, and *BCL2A1*) [54–56]. However, one of the limitations of using this software (version 3.5.2) is its inability to identify gene synonyms, as GeneMANIA did not identify *TRA@*, *AD022*, and its synonyms as valid names, as well as p15, although the official name *CDKN2B* was identified. In addition to its ability to detect and compensate for data redundancy, the GeneMANIA Prediction Server also displays an advantage due to the predictive accuracy and propagation of its algorithm [57,58].

With both approaches demonstrating potential in their use to identify hub genes with biological plausibility in identifying alteration/characterization of exposure and possible use as a target for cancer diagnosis, prognosis, and treatment. For example, the hub genes identified in this study are:

Benzene exposure:

IL-8 is a chemokine related to the promotion of chemotaxis and neutrophil degranulation. This chemokine activates multiple intracellular signaling pathways and is a significant regulatory factor within tumor microenvironments [54]. Increased *IL-8* expression is present in several cell types in the presence of benzene metabolites [55,59,60]; Gillis et al. [58] have demonstrated that human peripheral blood mononuclear cells (PBMCs) induce the production of cytokines in the presence of benzene metabolites, such as *IL-8*, at levels from 10 to several thousand-fold, resulting in increased cytokine levels in the medium, while the effects of benzene metabolites on the secretion of soluble cytokines are varied. For example, reductions in *IL-8* production dependent on hydroquinone and catechol concentrations but not on benzenetriol and benzoquinone have been observed. This demonstrates the relevance of network visualization as a whole, not just the possibility of direct *IL-8* changes, where the most specific Biological Process of the Benzene network are related to cytokine-mediated signaling pathways.

Krüppel-like factors (KLFs) are highly conserved zinc-finger proteins that regulate cellular transcription machinery and regulate a wide range of cellular functions, including cell proliferation, apoptosis, differentiation, and neoplastic transformation, by binding to GC-rich promoter regions [51,56]. *KLF6* and the proto-oncogene *JUN* are significant in differentiation and cell death, hematopoiesis, and cell survival. Regarding functions, all categories linked to DNA structure and transcription were present (in the case of benzene exposure and the identification of possible biomarkers, this becomes important) [51]. Unlike the *JUN* gene, which is known to regulate myeloid differentiation [61], the *KLF6* gene is a known tumor suppressor for prostate [62], colorectal [63], lung [64], ovary [65], gliomas [66], head and neck [67], and hepatocellular cancer [68]. However, the *KLF6* gene inhibits *JUN*-dependent transcription, which leads to an antagonistic effect on cell proliferation induced by the *JUN* gene [69]. This demonstrates both the non-specificity of benzene-inducing changes and the importance of a biological interaction approach.

SERTAD1 antagonizes the function of the inhibitor of apoptosis-stimulating protein p53 (iASPP), preventing its entry into the nucleus to interact with p53 in leukemic cells when iASPP is in its overproduction stage [70]. Evidence associating increased *SERTAD1* expression with the presence of benzene and its metabolite benzoquinone has been reported in the literature [51].

The *HRAS* protein is a small GTPase that cycles between inactive and active conformations. In the active state, *HRAS* binds to guanosine triphosphate (GTP) and possesses an intrinsic enzymatic activity that cleaves the terminal phosphate of this nucleotide, converting it to GDP. Upon conversion of GTP to guanosine diphosphate (GDP), *HRAS* is made inactive [70]. *STAT3* is a member of a family of cytoplasmic proteins that participate in cellular responses to cytokines and growth factors as transcription factors. Signal transducers and activators of transcription (STATs) are transcription factors that transmit signals from the extracellular surface of cells to the nucleus. *STAT3* is phosphorylated and activated in response to interleukin-6, contributing to an increased expression of genes activated in the liver during the acute phase response to inflammation [71].

Malathion network context:

Dysregulation of *HRAS* and *STAT3* pathways is frequently observed in several cancers [72–74]. Then altered *HRAS* protein is permanently activated within the cell. This overactive protein directs the cell to grow and divide in the absence of outside signals, leading to uncontrolled cell division and the formation of a tumor [2]. *STAT3* regulates basic biological processes essential to tumorigenesis, including cell-cycle progression, apoptosis, tumor angiogenesis, invasion and metastasis, and tumor-cell evasion of the immune system [75].

GO analysis and pathway enrichment provided further insights about Biological Processes related to Benzene and Malathion networks. The interrelationship between pathways and genes can be observed as the most frequent pathways represent more general processes and the least frequent pathways represent more specific gene processes. Such Benzene network functions are associated with leukemia mechanisms, and several of these functions and processes are mediated by *IL1A* and *PTGS2*, which play a central role in the characterization of gene expression associated with benzene exposure [47]. *IL1A* exhibits a single nucleotide polymorphism (SNP), which increases the expression of its mRNA and is inversely associated with granulocyte counts in benzene-exposed individuals [56], while *PTGS2* overexpression frequently occurs in pre-malignant and malignant neoplasms, including hematological malignancies [76]. The Malathion network is more associated with cellular death regulation mechanisms.

Therefore, to approach a complex disease study, a useful clue is provided by the fact that genes, gene products, and small molecules interact with each other to form a complex interaction network. Thus, an alteration in one gene might propagate through interactions, possibly affecting other genes in the network. The fact that one can observe similar disease phenotypes despite different genetic causes suggests that these different causes are not unrelated but rather jointly contribute to dysregulating the same component of the cellular system [1]. However, it is worth noting that exposure to low concentrations of environmental and occupational carcinogens does not eliminate the risk, considering that there is no safe exposure limit to carcinogenic substances such as benzene and malathion [77–79].

Moreover, one of the limitations of this study is that the representation of these complex networks of undirected biological interactions represents only a large number of possible pathways and interactions without taking into account the direction and dynamics of these processes. Although this approach shows potential to identify hub genes with biological plausibility in identifying alteration/characterization of exposure and possible use as a target for cancer diagnosis, prognosis, and treatment. In the limiting case, undirected network analysis assumes that all biological interactions and pathways represent processes occurring at the same rate, and then the observed clustering may lead to unrealistic conclusions. For example, interaction dynamics (chemical reaction kinetics) may promote some pathways while inhibiting others, just as the expression of some genes inhibits that of others [70,80]. For this reason, additional studies are recommended, mainly those involving directed networks. Once the directions in which information flows in the network are known, it can lead to a better identification of potential markers for diagnosis, treatment, and prevention.

4. Materials and Methods

4.1. Data Collection

Text mining was carried out using the bibliographic searches of NCBI (PubMed), Agilent Literature Search, Gene Expression Omnibus (GEO), Reactome, ArrayExpress, and Medline, involving two queries: one with the keywords “benzene gene expression”, “benzene microarray”, “benzene expression”, and “benzene poisoning;”, and another with the keywords “malathion gene expression”, “malathion microarray”, “malathion expression”, and “malathion poisoning”.

4.2. Construction of the Protein-Protein Interaction (PPI) Network

The Search Tool for Retrieval of Interacting Genes (STRING 11.0; <http://string-db.org/> accessed on 22 February 2022) was used to evaluate the PPI of DEGs related to malathion exposure. The PPI network was derived from proven experimental statistics such as automated text mining of scientific literature, experimental data, available signaling pathways, the PPI database, systematic coexpression, phylogenetic co-occurrence, observation of neighboring genomes, and genetic fusion events. We only extracted the edges with a minimum confidence score of 0.7 (high confidence) and the maximum number of interactors at the first shell with no more than five interactors.

The visualization of network and module analysis was carried out with the software Cytoscape (version 3.6.0; <http://www.cytoscape.org/> accessed on 15 February 2018).

4.3. Extended Interaction Network Analysis

The obtained gene list corresponding to benzene exposure was used for a functional and gene ontology (GO) analysis using GeneMANIA (version 3.5.2; <http://genemania.org/> accessed on 17 February 2018), a tool used for predicting gene function that can be implemented as a plug-in for the Cytoscape networks visualization software (version 3.6.0; <http://www.cytoscape.org/> accessed on 15 February 2018) [81].

The biological data sets searched by GeneMANIA, the characteristics considered for interaction formation, and the data source for the network creation are shown in Table 4.

Table 4. List of interaction network categories demonstrating the information for the creation of the biological network interactions and the data source researched by GeneMANIA.

Network Categories	Datasets	Network Information	Data Source
Co-expression:	Gene expression	Two genes are linked if their expression levels are similar across conditions in a gene expression study.	Most of these data are collected from the Gene Expression Omnibus (GEO).
Physical Interaction:	Protein-protein interaction	Two gene products are linked if they were found to interact in a protein-protein interaction study.	These data are collected from primary studies found in protein interaction databases, including BioGRID and PathwayCommons.
Genetic interaction:	Genetic interaction	Two genes are functionally associated if the effects of perturbing one gene are found to be modified by perturbations to a second gene.	These data are collected from primary studies and BioGRID.
Shared protein domains:	Protein domain	Two gene products are linked if they have the same protein domain.	These data are collected from domain databases such as InterPro, SMART, and Pfam.
Co-localization:	Genes expressed in the same tissue or proteins found in the same location.	Two genes are linked if they are both expressed in the same tissue or if their gene products are both identified in the same cellular location.	

Table 4. Cont.

Network Categories	Datasets	Network Information	Data Source
Pathway:	Pathway	Two gene products are linked if they participate in the same reaction within a pathway.	These data are collected from various sources, such as Reactome and BioCyc, via PathwayCommons.
Predicted:	Predicted functional relationships between genes, often protein interactions.	For instance, two proteins are predicted to interact if their orthologs are known to interact in another organism. In these cases, network names describe the original data source of experimentally measured interactions and the organism from which the interactions were mapped from (e.g. a mouse network predicted from a human network).	A major source of predicted data is mapping known functional relationships to another organism via orthology. These data are collected from various sources, such as BioGRID and I2D orthology.

Source: GeneMANIA network categories (<http://pages.genemania.org/help/> accessed on 28 February 2023).

For the analysis, we applied the default setting of 20 genes, which present the highest number of interactions. The software (version 3.5.2) standard was maintained regarding the advanced settings for physical, genetic, and interrelated paths, only modifying the co-expression data for articles on genomic leukemia and other hematological diseases.

4.4. Identification of Molecular Complexes

Highly interconnected, or dense, regions may represent molecular complexes. The Molecular Complex Detection (MCODE) algorithm, an automated method for finding clusters (highly interconnected regions), was used to identify sets of molecules that strongly interact with each other in the Benzene network. In addition, MCODE standard parameters were maintained (degree cutoff = 2, node score cutoff = 0.2, K-core = 2, and maximum depth = 100).

4.5. GO Category Representation

The biological network gene ontology (BiNGO) tool was used to perform a GO functional enrichment analysis. The hypergeometric distribution was used for the biological processes and molecular functions of functional overrepresentation categories. Only the over-represented GO categories were considered significant (adjusted p value < 0.05).

4.6. Centrality Analysis

Centralities are parameters that identify nodes with relevant positions in the global network architecture. The CentiScaPe 2.2 plug-in was used for clusters displaying an overrepresentation of GO categories of interest.

CentiScaPe calculates specific centrality parameters, describes the network topology, and assists in identifying the most important nodes in a complex network [82]. According to the connectivity of each node degree, betweenness, and eigenvector, the arithmetic mean of each centrality parameter was defined to obtain the most connected nodes.

Degree centrality indicates the number of adjacent nodes that are connected to a unique node. If the degree of centrality of a given node is much larger than the average degree of centrality of the network, such a node is classified as a hub. Hubs in interactive networks tend to be essential since their exclusion reduces the connectivity of the global network; consequently, they also represent greater biological relevance [4,83]. Betweenness, in turn, is defined by the number of shortest paths between all pairs of nodes that pass through a node of interest [82,84]. The nodes with the highest betweenness score (the 'bottleneck') control most of the information flow in the network, as they present the largest number of shortest paths (between other nodes in the network) passing through. As such,

bottleneck genes (nodes in the interactome ranked by betweenness centrality) are related to regulatory functions, representing critical points in the PPI network [85]. The eigenvector defines the node “prestige” of a network, i.e., the eigenvector centrality of a node in a network is large if this node is connected to many central and highly connected nodes [86].

5. Conclusions

Complex diseases, especially cancer, are extremely harmful to human health. Therefore, the identification of biomarkers is key to dealing with complex disease studies. Pathway analysis, combined with multiple types of high-throughput data, reflects biological processes more clearly and comprehensively than analyses involving the assessment of individual genes. For this reason, pathway-based complex disease analysis approaches have become a hot research topic.

In this study, we began with the systematic analysis of associated genes using text mining, followed by the identification of essential genes and pathways by functional annotation. We emphasize the central roles performed by several important hub genes, such as proteins from the *IL8*, *KLF*, *JUN*, and *SERTAD1* families obtained by exposure to benzene, while *HRAS* and *STAT3* are hubs for exposure to malathion.

Although the network results corroborate the literature about hub genes plausibility as potential biomarkers, additional validation studies are required to confirm this hypothesis. Given that systems biology approaches for predicting and characterizing hub genes are recent [82–85], the way in which data collection and network creation are conducted can vary according to the purpose of the study.

Supplementary Materials: The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/ijms24119415/s1>.

Author Contributions: Conceptualization, M.V.C.S. and A.L.L.; methodology, M.V.C.S.; writing—original draft preparation, M.V.C.S.; supervision—review and editing of benzene research, A.L.L. and J.A.P.; supervision—review and editing of malathion research, A.L.L., L.R.T. and I.C.C.-A.; review and editing of Network Medicine analysis, A.S.F. and D.C.M.J.; funding acquisition, A.L.L., J.A.P. and L.R.T. All authors have read and agreed to the published version of the manuscript.

Funding: This study was partially supported by the Carlos Chagas Filho Foundation for Research Support of São Paulo State —FAPESP (procs. 2018/18560-6 and 2018/21934-5), Carlos Chagas Filho Foundation for Research Support of Rio de Janeiro —FAPERJ Nota 10 fellowships (Faperj N° E_01/2017—Scholarship Note 10—Masters—Process E-26/201.144/2017), —FAPERJ N° E_04/2017 Scholarship Note 10—Pos-doctoral—Process E-26/202.437/2017. and Foundation Oswaldo Cruz (Fiocruz)—Finance Code 001.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Data are contained within the article or Supplementary Materials. The data that support the findings of our study are publicly available datasets from the GEO (<https://www.ncbi.nlm.nih.gov/geo/>), GeneMANIA (<https://genemania.org/>), and STRING (<https://string-db.org/>) databases, accessed on 22 February 2022.

Acknowledgments: This work was supported by the Brazilian agencies Carlos Chagas Filho Foundation for Research Support of Rio de Janeiro State (FAPERJ) grant number E-26/211.209/2021, National Council for Scientific and Technological Development (CNPq) grant number 309065/2021-6, Post-graduate Program in Saúde Pública e Meio Ambiente (SPMA) of the Escola Nacional de Saúde Pública/Fiocruz (ENSP/Fiocruz); and Vice-Presidência de Ambiente, Atenção e Promoção da Saúde/Fiocruz (VPAAPS/Fiocruz). The authors would like to express their gratitude to Rita de Cássia Oliveira da Costa Mattos for her contribution in benzene network and Hermano Albuquerque de Castro for his supporting.

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

References

1. Zhang, Q.; Li, J.; Xue, H.; Kong, L.; Wang, Y. Network-based methods for identifying critical pathways of complex diseases: A survey. *Mol. Biosyst.* **2016**, *12*, 1082–1089. [[CrossRef](#)] [[PubMed](#)]
2. Wu, X.Y.; Liu, W.T.; Wu, Z.F.; Chen, C.; Liu, J.Y.; Wu, G.N.; Yao, X.Q.; Liu, F.K.; Li, G. Identification of HRAS as cancer-promoting gene in gastric carcinoma cell aggressiveness. *Am. J. Cancer Res.* **2016**, *6*, 1935–1948. [[PubMed](#)]
3. Chen, L.; Wang, R.-S.; Zhang, X. *Biomolecular Networks: Methods and Applications in Systems Biology*, 1st ed.; Wiley: Hoboken, NJ, USA, 2009; p. 387.
4. Barabási, A.L.; Gulbahce, N.; Loscalzo, J. Network medicine: A network-based approach to human disease. *Nat. Rev. Genet.* **2011**, *12*, 56–68. [[CrossRef](#)] [[PubMed](#)]
5. Cowen, L.; Ideker, T.; Raphael, B.J.; Sharan, R. Network propagation: A universal amplifier of genetic associations. *Nat. Rev. Genet.* **2017**, *18*, 551–562. [[CrossRef](#)]
6. Stuart, J.M. A Gene-Coexpression Network for Global Discovery of Conserved Genetic Modules. *Science* **2003**, *302*, 249–255. [[CrossRef](#)]
7. Zhang, L.V.; King, O.D.; Wong, S.L.; Goldberg, D.S.; Tong, A.H.Y.; Lesage, G.; Andrews, B.; Bussey, H.; Boone, C.; Roth, F.P. Motifs, themes and thematic maps of an integrated *Saccharomyces cerevisiae* interaction network. *J. Biol.* **2005**, *4*, 6. [[CrossRef](#)]
8. Giaever, G.; Shoemaker, D.D.; Jones, T.W.; Liang, H.; Winzeler, E.A.; Astromoff, A.; Davis, R.W. Genomic profiling of drug sensitivities via induced haploinsufficiency. *Nat. Genet.* **1999**, *1*, 278–283. [[CrossRef](#)]
9. Uetz, P.; Giot, L.; Cagney, G.; Mansfield, T.A.; Judson, R.S.; Knight, J.R.; Lockshon, D.; Narayan, V.; Srinivasan, M.; Pochart, P.; et al. A comprehensive analysis of protein-protein interactions in *Saccharomyces cerevisiae*. *Nature* **2000**, *403*, 623–627. [[CrossRef](#)]
10. Mering, C.V.; Krause, R.; Snel, B.; Cornell, M.; Oliver, S.G.; Fields, S.; Bork, P. Comparative assessment of large-scale data sets of protein-protein interactions. *Nature* **2002**, *417*, 399–403. [[CrossRef](#)]
11. Polacco, B.J.; Babbitt, P.C. Automated discovery of 3D motifs for protein function annotation. *Bioinformatics* **2006**, *22*, 723–730. [[CrossRef](#)]
12. Pellegrini, M.; Marcotte, E.M.; Thompson, M.J.; Eisenberg, D.; Yeates, T.O. Assigning protein functions by comparative genome analysis: Protein phylogenetic profiles. *Proc. Natl. Acad. Sci. USA* **1999**, *96*, 4285–4288. [[CrossRef](#)]
13. Hegyi, H.; Gerstein, M. The relationship between protein structure and function: A comprehensive survey with application to the yeast genome. *J. Mol. Biol.* **1999**, *288*, 147–164. [[CrossRef](#)]
14. Akavia, U.D.; Litvin, O.; Kim, J.; Sanchez-Garcia, F.; Kotliar, D.; Causton, H.C.; Pochanard, P.; Mozes, E.; Garraway, L.A.; Pe'er, D. An Integrated Approach to Uncover Drivers of Cancer. *Cell* **2010**, *143*, 1005–1017. [[CrossRef](#)]
15. Zhang, L.; McHale, C.M.; Rothman, N.; Li, G.; Ji, Z.; Vermeulen, R.; Hubbard, A.E.; Ren, X.; Shen, M.; Rappaport, S.M.; et al. Systems biology of human benzene exposure. *Chem. Biol. Interact.* **2010**, *184*, 86–93. [[CrossRef](#)]
16. International Agency for Research on Cancer (IARC). *Chemical Agents and Related Occupations*; IARC: Lyon, France, 2012; Volume 100F.
17. Irons, R.D.; Chen, Y.; Wang, X.; Ryder, J.; Kerzic, P.J. Acute myeloid leukemia following exposure to benzene more closely resembles de novo than therapy related-disease. *Genes Chromosomes Cancer* **2013**, *52*, 887–894. [[CrossRef](#)]
18. Lagorio, S.; Ferrante, D.; Ranucci, A.; Negri, S.; Sacco, P.; Rondelli, R.; Cannizzaro, S.; Torregrossa, M.V.; Cocco, P.; Forastiere, F.; et al. Exposure to benzene and childhood leukaemia: A pilot case-control study. *BMJ Open* **2013**, *3*, e002275. [[CrossRef](#)]
19. Li, G.; Yin, S. Progress of epidemiological and molecular epidemiological studies on benzene in China. *Ann. N. Y. Acad. Sci.* **2006**, *1076*, 800–809. [[CrossRef](#)]
20. Snyder, R. Leukemia and Benzene. *Int. J. Environ. Res. Public Health* **2012**, *9*, 2875–2893. [[CrossRef](#)]
21. Fonseca, A.S.A.; Costa, D.F.; Dapper, V.; Machado, J.M.H.; Valente, D.; Carvalho, L.V.B.; Costa-Amaral, I.S.; Alves, S.R.; Sarcinelli, P.N.; Menezes, M.A.C.; et al. Classificação clínico-laboratorial para manejo clínico de trabalhadores expostos ao benzeno em postos de venda de combustíveis. *Rev. Bras. Saúde Ocup.* **2017**, *42*, e5s. [[CrossRef](#)]
22. Glass, D.C.; Gray, C.N.; Jolley, D.J.; Gibbons, C.; Sim, M.R.; Fritschi, L.; Adams, G.G.; Bisby, J.A.; Manuell, R. Leukemia risk associated with low-level benzene exposure. *Epidemiol. Camb. Mass.* **2003**, *14*, 569–577. [[CrossRef](#)]
23. Lan, Q.; Zhang, L.; Li, G.; Vermeulen, R.; Weinberg, R.S.; Dosemeci, M.; Rappaport, S.M.; Shen, M.; Alter, B.P.; Wu, Y.; et al. Hematotoxicity in workers exposed to low levels of benzene. *Science* **2004**, *306*, 1774–1776. [[CrossRef](#)] [[PubMed](#)]
24. Pesatori, A.C.; Garte, S.; Popov, T.; Georgieva, T.; Panev, T.; Bonzini, M.; Consonni, D.; Carugno, M.; Goldstein, B.D.; Taioli, E.; et al. Early effects of low benzene exposure on blood cell counts in Bulgarian petrochemical workers. *Med. Lav.* **2009**, *100*, 83–90. [[PubMed](#)]
25. Swaen, G.M.H.; Amelsvoort, L.V.; Twisk, J.J.; Verstraeten, E.; Slootweg, R.; Collins, J.J.; Burns, C.J. Low level occupational benzene exposure and hematological parameters. *Chem. Biol. Interact.* **2010**, *184*, 94–100. [[CrossRef](#)] [[PubMed](#)]
26. Smith, M.T. Advances in understanding benzene health effects and susceptibility. *Annu. Rev. Public Health* **2010**, *31*, 133–148. [[CrossRef](#)]
27. Costa-Amaral, I.C.; Carvalho, L.V.B.; Pimentel, J.N.S.; Pereira, A.C.; Vieira, J.A.; Castro, V.S.; Borges, R.M.; Alvez, S.R.; Nogueira, S.M.; Tabalipa, M.d.M.; et al. Avaliação ambiental de BTEX (benzeno, tolueno, etilbenzeno, xilenos) e biomarcadores de genotoxicidade em trabalhadores de postos de combustíveis. *Rev. Bras. Saúde Ocup.* **2017**, *42*, e8s. [[CrossRef](#)]

28. Valente, D.; Costa-Amaral, I.C.; Carvalho, L.V.B.; Santos, M.V.C.; Castro, V.S.; Rodrigues, D.R.F.; Falco, A.; Silva, C.B.; Nogueira, S.M.; Golçoloves, E.S.; et al. Utilização de biomarcadores de genotoxicidade e expressão gênica na avaliação de trabalhadores de postos de combustíveis expostos a vapores de gasolina. *Rev. Bras. Saúde Ocup.* **2017**, *42*, e2s. [[CrossRef](#)]
29. McHale, C.M.; Zhang, L.; Smith, M.T. Current understanding of the mechanism of benzene-induced leukemia in humans: Implications for risk assessment. *Carcinogenesis* **2012**, *33*, 240–252. [[CrossRef](#)]
30. Moore, P.D.; Yedjou, C.G.; Tchounwou, P.B. Malathion-Induced Oxidative Stress, Cytotoxicity, and Genotoxicity in Human Liver Carcinoma (HepG2) Cells. *Environ. Toxicol.* **2010**, *25*, 221–226. [[CrossRef](#)]
31. Navarrete-Meneses, M.P.; Pedraza-Meléndez, A.I.; Salas-Labadía, C.; Moreno-Lorenzana, D.; Pérez-Vera, P. Low concentrations of permethrin and malathion induce numerical and structural abnormalities in KMT2A and IGH genes in vitro: Permethrin and malathion induce abnormalities in KMT2A and IGH genes. *J. Appl. Toxicol.* **2018**, *38*, 1262–1270. [[CrossRef](#)]
32. Cabello, G.; Valenzuela, M.; Vilaxa, A.; Durán, V.; Rudolph, I.; Hrepic, N.; Calaf, G. A Rat Mammary Tumor Model Induced by the Organophosphorous Pesticides Parathion and Malathion, Possibly through Acetylcholinesterase Inhibition. *Environ. Health Perspect.* **2001**, *109*, 471–479. [[CrossRef](#)]
33. Bonner, M.R.; Coble, J.; Blair, A.; Freeman, L.E.B.; Hoppin, J.A.; Sandler, D.P.; Alavanja, M.C.R. Malathion exposure and the incidence of cancer in the Agricultural Health Study. *Am. J. Epidemiol.* **2007**, *166*, 1023–1103. [[CrossRef](#)]
34. Lerro, C.C.; Koutros, S.; Andreotti, G.; Friesen, M.C.; Alavanja, M.C.; Blair, A.; Hoppin, J.A.; Sandler, D.P.; Lubin, J.H.; Ma, X.; et al. Organophosphate insecticide use and cancer incidence among spouses of pesticide applicators in the Agricultural Health Study. *Occup. Environ. Med.* **2015**, *72*, 736–744. [[CrossRef](#)]
35. Reuber, M.D. Carcinogenicity and toxicity of malathion and malaoxon. *Environ. Res.* **1985**, *37*, 119–153. [[CrossRef](#)]
36. Smith, M.T.; Zhang, L.; McHale, C.M.; Skibola, C.F.; Rappaport, S.M. Benzene, the exposome and future investigations of leukemia etiology. *Chem. Biol. Interact.* **2011**, *192*, 155–159. [[CrossRef](#)]
37. Roy, S.; Bhattacharyya, D.K.; Kalita, J.K. Reconstruction of gene co-expression network from microarray data using local expression patterns. *BMC Bioinform.* **2014**, *15*, S10. [[CrossRef](#)]
38. Luo, J.; Qi, Y. Identification of essential proteins based on a new combination of local interaction density and protein complexes. *PLoS ONE* **2015**, *10*, e0131418. [[CrossRef](#)]
39. Li, M.; Lu, Y.; Niu, Z.; Wu, F.X.; Pan, Y. Identification of essential proteins by using complexes and interaction network. In *Bioinformatics Research and Applications*; Basu, M., Pan, Y., Wang, J., Eds.; Springer International Publishing: Cham, Switzerland, 2014; pp. 255–265.
40. Hu, P.; Mei, T. Ranking influential nodes in complex networks with structural holes. *Phys. A Stat. Mech. Appl.* **2018**, *490*, 624–631. [[CrossRef](#)]
41. Cinaglia, P.; Cannataro, M. Network alignment and motif discovery in dynamic networks. *Netw. Model. Anal. Health Inform. Bioinform.* **2022**, *11*, 38. [[CrossRef](#)]
42. Dai, C.; He, J.; Hu, K.; Ding, Y. Identifying essential proteins in dynamic protein networks based on an improved h-index algorithm. *BMC Med. Inform. Decis. Mak.* **2020**, *20*, 110. [[CrossRef](#)]
43. Zhang, W.; Xu, J.; Li, Y.; Zou, X. Detecting Essential Proteins Based on Network Topology, Gene Expression Data, and Gene Ontology Information. *IEEE/ACM Trans. Comput. Biol. Bioinform.* **2018**, *15*, 109–116. [[CrossRef](#)]
44. Lei, X.; Wang, S.; Wu, F.X. Identification of essential proteins based on improved HITS algorithm. *IEEE/ACM Trans. Comput. Biol. Bioinform. Genes* **2019**, *10*, 177. [[CrossRef](#)] [[PubMed](#)]
45. Mistry, D.; Wise, R.; Dickerson, J. DiffSLC: A graph centrality method to detect essential proteins of a protein-protein interaction network. *PLoS ONE* **2017**, *12*, e0187091. [[CrossRef](#)] [[PubMed](#)]
46. Zaki, N.; Berenguères, J.; Efimov, D. Detection of protein complexes using a protein ranking algorithm. *Proteins Struct. Funct. Bioinform.* **2012**, *80*, 2459–2468. [[CrossRef](#)] [[PubMed](#)]
47. McHale, C.M.; Zhang, L.; Lan, Q.; Vermeulen, R.; Li, G.; Hubbard, A.E.; Porter, K.E.; Thomas, R.; Portier, C.J.; Shenet, M.; et al. Global Gene Expression Profiling of a Population Exposed to a Range of Benzene Levels. *Environ. Health Perspect.* **2010**, *119*, 628–634. [[CrossRef](#)] [[PubMed](#)]
48. Bi, Y.; Li, Y.; Kong, M.; Xiao, X.; Zhao, Z.; He, X.; Ma, Q. Gene expression in benzene-exposed workers by microarray analysis of peripheral mononuclear blood cells: Induction and silencing of CYP4F3A and regulation of DNA-dependent protein kinase catalytic subunit in DNA double strand break repair. *Chem. Biol. Interact.* **2010**, *184*, 207–211. [[CrossRef](#)]
49. Xing, C.; Wang, Q.; Li, B.; Tian, H.; Ni, Y.; Yin, S.; Li, G. Methylation and expression analysis of tumor suppressor genes p15 and p16 in benzene poisoning. *Chem. Biol. Interact.* **2010**, *184*, 306–309. [[CrossRef](#)]
50. Sarma, S.N.; Kim, Y.-J.; Ryu, J.-C. Differential gene expression profiles of human leukemia cell lines exposed to benzene and its metabolites. *Environ. Toxicol. Pharmacol.* **2011**, *32*, 285–295. [[CrossRef](#)]
51. Gao, A.; Yang, J.; Yang, G.; Niu, P.; Tian, L. Differential gene expression profiling analysis in workers occupationally exposed to benzene. *Sci. Total Environ.* **2014**, *472*, 872–879. [[CrossRef](#)]
52. Anjitha, R.; Antony, A.; Shilpa, O.; Anupama, K.P.; Mallikarjunaiah, S.; Gurushankara, H.P. Malathion induced cancer-linked gene expression in human lymphocytes. *Environ. Res.* **2020**, *182*, 109–131. [[CrossRef](#)]
53. Chow, K.; Sarkar, A.; Elhesha, R.; Cinaglia, P.; Ay, A.; Kahveci, T. ANCA: Alignment-Based Network Construction Algorithm. *IEEE/ACM Trans. Comput. Biol. Bioinf.* **2021**, *18*, 512–524. [[CrossRef](#)]
54. Waugh, D.J.J.; Wilson, C. The Interleukin-8 Pathway in Cancer. *Clin. Cancer Res.* **2008**, *14*, 6735–6741. [[CrossRef](#)]

55. Moro, A.M.; Charão, M.F.; Brucker, N.; Durgante, J.; Baierle, M.; Bubols, G.; Goethel, G.; Fracasso, R.; Nascimento, S.; Bulcão, R.; et al. Genotoxicity and oxidative stress in gasoline station attendants. *Mutat. Res. Toxicol. Environ. Mutagen.* **2013**, *754*, 63–70. [[CrossRef](#)]
56. Lan, Q. Polymorphisms in Cytokine and Cellular Adhesion Molecule Genes and Susceptibility to Hematotoxicity among Workers Exposed to Benzene. *Cancer Res.* **2005**, *65*, 9574–9581. [[CrossRef](#)]
57. Bader, G.D.; Hogue, C.W.V. An automated method for finding molecular complexes in large protein interaction networks. *BMC Bioinform.* **2003**, *4*, 2. [[CrossRef](#)]
58. Mostafavi, S.; Ray, D.; Warde-Farley, D.; Grouios, C.; Morris, Q. GeneMANIA: A real-time multiple association network integration algorithm for predicting gene function. *Genome Biol.* **2008**, *9*, S4. [[CrossRef](#)]
59. Bironaite, D.; Siegel, D.; Moran, J.L.; Weksler, B.B.; Ross, D. Stimulation of endothelial IL-8 (eIL-8) production and apoptosis by phenolic metabolites of benzene in HL-60 cells and human bone marrow endothelial cells. *Chem. Biol. Interact.* **2004**, *149*, 177–188. [[CrossRef](#)]
60. Gillis, B.; Gavin, I.M.; Arbieva, Z.; King, S.T.; Jayaraman, S.; Prabhakar, B.S. Identification of human cell responses to benzene and benzene metabolites. *Genomics* **2007**, *90*, 324–333. [[CrossRef](#)]
61. Lord, K.A.; Abdollahi, A.; Hoffman-Liebermann, B.; Liebermann, D.A. Proto-oncogenes of the fos/jun family of transcription factors are positive regulators of myeloid differentiation. *Mol. Cell. Biol.* **1993**, *13*, 841–851.
62. Narla, G. KLF6, a Candidate Tumor Suppressor Gene Mutated in Prostate Cancer. *Science* **2001**, *294*, 2563–2566. [[CrossRef](#)]
63. Mukai, S. Involvement of Krüppel-like factor 6 (KLF6) mutation in the development of nonpolypoid colorectal carcinoma. *World J. Gastroenterol.* **2007**, *13*, 3932. [[CrossRef](#)]
64. Ito, G.; Uchiyama, M.; Kondo, M.; Mori, S.; Usami, N.; Maeda, O.; Kawabe, T.; Hasegawa, Y.; Shimokata, K.; Sekido, Y. Kruppel-Like Factor 6 Is Frequently Down-Regulated and Induces Apoptosis in Non-Small Cell Lung Cancer Cells. *Cancer Res.* **2004**, *64*, 3838–3843. [[CrossRef](#)] [[PubMed](#)]
65. DiFeo, A. Roles of KLF6 and KLF6-SV1 in Ovarian Cancer Progression and Intraperitoneal Dissemination. *Clin. Cancer Res.* **2006**, *12*, 3730–3739. [[CrossRef](#)] [[PubMed](#)]
66. Camacho-Vanegas, O.; Narla, G.; Teixeira, M.S.; DiFeo, A.; Misra, A.; Singh, G.; Chan, A.M.; Friedman, S.L.; Feuerstein, B.G.; Martignetti, J.A. Functional inactivation of the KLF6 tumor suppressor gene by loss of heterozygosity and increased alternative splicing in glioblastoma. *Int. J. Cancer* **2007**, *121*, 1390–1395. [[CrossRef](#)]
67. Teixeira, M.S.; Camacho-Vanegas, O.; Fernandez, Y.; Narla, G.; DiFeo, A.; Lee, B.; Kalir, T.; Friedman, S.L.; Schlecht, N.F.; Genden, E.M.; et al. KLF6 allelic loss is associated with tumor recurrence and markedly decreased survival in head and neck squamous cell carcinoma. *Int. J. Cancer* **2007**, *121*, 1976–1983. [[CrossRef](#)] [[PubMed](#)]
68. Song, J.; Kim, C.J.; Cho, Y.G.; Kim, S.Y.; Nam, S.W.; Lee, S.H.; Yoo, N.J.; Lee, J.Y.; Park, W.S. Genetic and epigenetic alterations of the KLF6 gene in hepatocellular carcinoma. *J. Gastroenterol. Hepatol.* **2006**, *21*, 1286–1289. [[CrossRef](#)]
69. Slavin, D.A.; Koritschoner, N.P.; Prieto, C.C.; López-Díaz, F.J.; Chatton, B.; Bocco, J.L. A new role for the Krüppel-like transcription factor KLF6 as an inhibitor of c-Jun proto-oncoprotein function. *Oncogene* **2004**, *23*, 8196–8205. [[CrossRef](#)]
70. Qiu, S.; Liu, S.; Yu, T.; Yu, J.; Wang, M.; Rao, Q.; Xing, H.; Tang, K.; Mi, Y.; Wang, J. Sertad1 antagonizes iASPP function by hindering its entrance into nuclei to interact with P53 in leukemic cells. *BMC Cancer* **2017**, *17*, 795. [[CrossRef](#)]
71. Kompier, L.C.; Lurkin, I.; van-der-Aa, M.N.M.; van-Rhijn, B.W.G.; van-der-Kwast, T.H.; Zwarthoff, E.C. FGFR3, HRAS, KRAS, NRAS and PIK3CA mutations in bladder cancer and their potential as biomarkers for surveillance and therapy. *PLoS ONE* **2010**, *5*, e13821. [[CrossRef](#)]
72. Raz, R.; Durbin, J.E.; Levy, D.E. Acute phase response factor and additional members of the interferon-stimulated gene factor 3 family integrate diverse signals from cytokines, interferons, and growth factors. *J. Biol. Chem.* **1994**, *269*, 24391–24395. [[CrossRef](#)]
73. van-Engen-van Grunsven, A.C.H.; van-Dijk, M.C.R.F.; Ruiters, D.J.; Klaasen, A.; Mooi, W.J.; Blokx, W.A.M. HRAS-mutated Spitz Tumors: A Subtype of Spitz Tumors With Distinct Features. *Am. J. Surg. Pathol.* **2010**, *34*, 1436–1441. [[CrossRef](#)]
74. Calaf, G.; Roy, D. Cancer genes induced by malathion and parathion in the presence of estrogen in breast cells. *Int. J. Mol. Med.* **2008**, *2*, 261–268. [[CrossRef](#)]
75. Fernandez-Medarde, A.; Santos, E. Ras in Cancer and Developmental Diseases. *Genes Cancer* **2011**, *2*, 344–358. [[CrossRef](#)]
76. Bernard, M.; Bancos, S.; Sime, P.; Phipps, R. Targeting Cyclooxygenase-2 in Hematological Malignancies: Rationale and Promise. *Curr. Pharm. Des.* **2008**, *14*, 2051–2060. [[CrossRef](#)]
77. Santos, M.V.C.; Figueiredo, V.O.; Arcuri, A.S.A.; Costa-Amaral, I.C.; Gonçalves, E.S.; Larentis, A.L. Aspectos toxicológicos do benzeno, biomarcadores de exposição e conflitos de interesses. *Rev. Bras. Saúde Ocup.* **2017**, *42*, e13s. [[CrossRef](#)]
78. Friedrich, K.; Gurgel, A.D.M.; Sarpa, M.; Bedor, C.N.G.; Siqueira, M.T.D.; Gurgel, I.G.D.; Augusto, L. Toxicologia crítica aplicada aos agrotóxicos—perspectivas em defesa da vida. *Saúde Debate* **2022**, *46*, 293–315. [[CrossRef](#)]
79. Mendes, M.P.R.; Paiva, M.J.N.; Costa-Amaral, I.C.; Carvalho, L.V.B.; Figueiredo, V.O.; Gonçalves, E.S.; Larentis, A.L.; André, L.C. Metabolomic Study of Urine from Workers Exposed to Low Concentrations of Benzene by UHPLC-ESI-QToF-MS Reveals Potential Biomarkers Associated with Oxidative Stress and Genotoxicity. *Metabolites* **2022**, *12*, 978. [[CrossRef](#)]
80. Stobiecka, M.; Chalupa, A. DNA Strand Replacement Mechanism in Molecular Beacons Encoded for the Detection of Cancer Biomarkers. *J. Phys. Chem. B* **2016**, *120*, 4782–4790. [[CrossRef](#)]
81. Montojo, J.; Zuberi, K.; Rodriguez, H.; Kazi, F.; Wright, G.; Donaldson, S.L.; Morris, Q.; Bader, G.D. GeneMANIA Cytoscape plugin: Fast gene function predictions on the desktop. *Bioinformatics* **2010**, *26*, 2927–2928. [[CrossRef](#)]

82. Scardoni, G.; Petterlini, M.; Laudanna, C. Analyzing biological network parameters with CentiScaPe. *Bioinformatics* **2009**, *25*, 2857–2859. [[CrossRef](#)]
83. Azevedo, H.; Bando, S.Y.; Bertonha, F.B.; Moreira-Filho, C.A. Redes de interação gênica e controle epigenético na transição saúde-doença. *Rev. Med.* **2015**, *94*, 223. [[CrossRef](#)]
84. Newman, M.E.J. A measure of betweenness centrality based on random walks. *Soc. Netw.* **2005**, *27*, 39–54. [[CrossRef](#)]
85. Yu, H.; Kim, P.M.; Sprecher, E.; Trifonov, V.; Gerstein, M. The Importance of Bottlenecks in Protein Networks: Correlation with Gene Essentiality and Expression Dynamics. *PLoS Comput. Biol.* **2007**, *3*, e59. [[CrossRef](#)] [[PubMed](#)]
86. Vilela, M.; Chou, I.-C.; Vinga, S.; Vasconcelos, A.; Voit, E.O.; Almeida, J.S. Parameter optimization in S-system models. *BMC Syst. Biol.* **2008**, *2*, 35. [[CrossRef](#)] [[PubMed](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.