International Journal of Computing and Digital Systems
ISSN (2210-142X)
Int. J. Com. Dig. Sys.13, No.1 (May-23)

http://dx.doi.org/10.12785/ijcds/130190

Scrutinizing the Homo Sapiens Gene-Gene Interactions in Sex Chromosomes: A Network Model

Basim Mahmood[1], Marwah M. A. Dabdawb[2] and Dhafar Sami Hammadi[3]

[1]*Department of Computer Science, University of Mosul, Mosul, IRAQ/ BioComplex Laboratory, Exeter, UK*
[2]*Department of Software, University of Mosul, Mosul, IRAQ*
[3]*Department of Computer Science, University of Mosul, Mosul, IRAQ*

Received 9 Nov. 2022, Revised 26 Mar. 2023, Accepted 23 Apr. 2023, Published 30 May. 2023

Abstract: The sex chromosomes X and Y participate in determining the sex of humans. These chromosomes contain many genes that contribute to a variety of body functions. Homo sapiens (human) genes have been widely studied in the literature. Most of the studies have been conducted on the interactions among genes and their relations to diseases. Therefore, the genes of a chromosome or across chromosomes may have relations to each other. Understanding these relations is important insofar as they contribute to a wider view of diseases. This work tries to scrutinize the homo sapiens gene-gene interactions in the sex chromosomes. More precisely, we generate a network model of the interactions between the genes in the sex chromosomes and the genes of other chromosomes. The nodes of the generated network represent the genes in chromosomes X and Y. An edge between two genes is generated if there exists an interaction between both. The generated model demonstrates interesting facts about the relations among genes. Finally, it can be of interest for biologists to look at such models for a more understanding of the interactions among genes and their relations to diseases.

Keywords: Homo Sapiens Genes, Sex Chromosomes, Gene Modelling, Genes Interactions, Diseases and Genes, Complex Networks

1. INTRODUCTION

Bioinformatics is a field of study that integrates biology and computer science aiming to have a better understanding of particular phenomena [1]. This field has been significantly developed with the development of computer science field. Modelling biological issues can be performed using computerized models (e.g., statistical models, mathematical models, etc.) [1].

One of the interesting bioinformatics areas is genetics modelling. In this area of study, genetic systems are implemented as computerized models [2]. These models can be simulated and analyzed using computer systems that can provide more accurate results. Also, modeling genetic systems is not an easy task due to the complexity of such systems. To overcome this issue, it is needed to perform the modeling process using concepts inspired by computer science theories [3][4]. Furthermore, this work deals with the human genetic system, which is complicated. This system includes 23 pairs of chromosomes and each of which contains hundreds of genes that, in turn, have a lot of interactions among them [5]. Modelling this kind of system needs a lot of attention since the computerized model should simulate the same facts in the real system [6].

The traditional analysis of the genetic systems is hard to be conducted due to the huge and complex nature of the data involved [6][7]. Therefore, the main goal of modeling genetic systems is to address, comprehend, or understand particular issues in these systems [8][9][10]. One of the modern approach that can be used in modelling genetic systems is the use of Complex Networks concepts, which are based on the Graph Theory [9][10][11]. This kind of approach assumes the system as nodes and edges connecting them [12][13][14].

The literature includes a lot of studies that modelled genetic systems in the form of gene-gene interaction networks, protein-protein interaction networks, and disease networks. Lim et al. [15] studied the interactions among 54 human proteins using the graph theory concepts. In the model they suggested, each protein is represented as a node, and two nodes are connected if there is an interaction between them. In the results, the authors identified 770 novel protein-protein interactions. In the same context, Perkins et al. [16] analyzed the functional, structural, and network characteristics of transient protein interactions. They also studied the issue of detecting the different types of protein-protein interactions. Their findings showed that the transient interactions of proteins are necessary for the functional processes in biological systems. The study also mentioned that identifying transient interaction among

E-mail address: bmahmood@(uomosul.edu.iq/ biocomplexlab.org), marwa_marwan21@uomosul.edu.iq, dhafar_jm@uomosul.edu.iq

proteins needs to have a well understanding of the system's structural features. The relations among human genes were investigated in the literature aiming to predict their functions/interactions. The study of Goh et al. [17] used two network models to investigate the relations between human genes and their related diseases. They built a human disease network and disease gene network. The authors found that many disease genes are nonessential with a weak tendency to encode hub genes. In another study performed by Barabasi et al. [18], the authors studied the relations among human genes as a network model. This study showed interesting results such that "Disease genes have a high propensity to interact with each other, forming disease modules. The identification of these disease modules can help us to identify disease pathways and predict other disease genes.", the authors said. In contrast, some studies showed that genes/proteins interactions models are not enough to improve the prediction of disease genes such as the study of Hwang et al. [19]. However, studying the interactions among genes can be extended not only to include diseases but also in investigating the social behavior of humans. Quintana et al. [20] investigated the genes that affect the social behavior of humans. They deeply studied the interactions of genes and the human brain and found that specific genes' interactions may impact behavior. In the same context, Sinha et al. [21] studied the behavior-related genes regulatory networks. They also examined the relationships among these networks and provided some insight into the behavior-related genes. Lopes-Ramos et al. [22] studied the sex differences manifest in many diseases. They constructed a network of gene regularity using 29 human healthy tissues retrieved from (GTEx) project. The analysis of the network showed a sex-biased in each tissue. This result was obtained after analyzing male and female networks using a variety of measurements. Banik et al. [23] used a network analysis approach to investigate NCBI database using 988 genes of different chromosomes. The edges creation strategy was based on protein-protein interactions. The generated network was clustered and the influential genes were extracted. They found that it was efficient to use network analysis methods to understand genes and their interactions, and their relations to some diseases. The same approach was used by Mathavan et al. [24] for identifying potential genes that impact lips and cavity cancers. Their study involved 472 genes retrieved from DisGeNET database. Other recent works were also focused on the network science approach for analyzing gene networks and diseases such as [25][26][27].

According to the literature, each study considers a particular issue in human genes (e.g., a disease or a few interactions). Also, it is difficult, in this field of research, to perform a general or comprehensive study that includes all human genes. Therefore, this work focuses on one of the largest communities of genes, which is the Homo Sapiens (human) sex chromosomes including all the available and verified genes in the literature. Hence, the contribution of this work is modelling the gene-gene interactions of the human sex chromosomes. The goal of this model is to scrutinize the structural features of the suggested model and extract facts about the diseases and the relations (interactions) among the genes. Besides, the complete model will be freely available to researchers who can quickly retrieve information about gene-gene interactions. It should be mentioned that, in this work, the gene interactions dataset was not directly used from NCBI, instead, it was manually collected, double-checked with their corresponding references, and then verified as an entry in the dataset. Therefore, collecting the dataset was accurately performed for all the available gene interactions in the sex chromosome. Our model is highly reliable in terms of the accuracy of the collected data, which makes it easier and faster for researchers to investigate a particular disease, find a particular interaction, or look at the relations of a gene to other genes within the network model.

Human sex chromosomes are represented by two chromosomes X and Y. These chromosomes have a different form, size, and behavior compared to the other 22 human chromosomes [28]. Chromosome X has 800 to 900 genes, while chromosome Y has 50-60 genes [29]. However, the number of genes in each chromosome may vary from one reference to another and this is due to the approach used to predict the number of genes in each chromosome [30]. Figure 1 demonstrates the sex chromosomes and their form. Moreover, each chromosome has diseases associated with its genes [29]. Table I presents the chromosomes and their main associated diseases.

This article is organized as follows: Section 2 describes the research method including the data collection process, network generation, and the analysis approach. Section 3 presents the results obtained and the discussions of the results. Finally, the article is concluded in Section 4.

2. RESEARCH METHOD

A. *Dataset Collection and Network Creation*

The dataset was collected from the National Center for Biotechnology Information (NCBI) [29]. This source is considered one of the most accredited references in biomedical and genomic information. According to NCBI, chromosome X has approximately 900 genes. However, this work considers the genes that interacted with other genes from the same chromosome or with genes across chromosomes. Therefore, the dataset includes 732 genes in total, where 177 genes of them are from chromosome X and 1 gene from chromosome Y and the genes of both chromosomes (178) interact with 554 genes from other chromosomes. The information about the interactions among genes was collected from NCBI.

The next step is to generate the gene-gene interaction undirected network. To this end, a network model is built based on the collected data. The network is represented as a graph $G(N, E)$ that contains nodes (N) and edges (E). Each gene in the dataset is represented as a node, and two nodes are connected by an undirected edge if and only

TABLE I. Summary of the human chromosomes and their associated diseases and genes [29]

Chro.	Main Disease	Associated Genes
1	"Porphyria Cutanea Tarda" (Skin Disease)/ "Gaucher"/ "Glaucoma" (Blinding Eye)/ "Prostate Cancer"/ "Alzheimer"	UROD/ GBA/ GLC1A/ HPC1/ PS2/
2	"Essential Tremor"/ "Colon Cancer"/ "Waardenberg Syndrome"	ETM2/ MSH2/ MSH6/ PAX3
3	"von Hippel-Lindau"/ "Lung Cancer"/ "Colon Cancer"/ "Essential Tremor"/	VHL/ SCLC1/ MLH1/ ETM1
4	"Ellis-van-Creveld"/ "Huntington"/ "Achondroplasia"/ "Parkinson"/ "Narcolepsy"/ "Fibrodysplasia Ossificans Progreeiv"	EVC/ HD/ FGFR3/ SNCA/ NRCLP/ FOP
5	"Steriod 5-Alpha Reductase 1"/ "Cockayne Syndrome"/ "Spinal Muscular Atrophy"/ "Asthma"/ "Diastrophic Dysplasia"	SRD5A1/ CKN1/ SMN1/ DTD
6	"Spinocerebellar Ataxia"/ "Diabetes"/ "Hemochromatosis"/ "Congenital Adrenal Hyperplasia"/ "Epilepdy"	SCA1/ IDDM1/ HFE/ CYP21A/ EPM2A
7	"Diabetes"/ "Williams Syndrome"/ "Pendred Syndrome"/ "Cystic Fibrosis"/ "Obesity"	GCK/ ELN/ Pendrin/ CFTR/ OB
8	"Werner Syndrome"/ "Burkitt Lymphoma"	WRN/ MYC
9	"Malignant Melanoma"/ "Friedrich's Araxia"/ "Tangier"/ "Chronic Myeloid Leukemia"/ "Tuberous Sclerosis"	CDKN2/ FRDA/ ABCI/ ABL/ TSC1
10	"Refsum"/ "Gyrate Atrophy"	PAHX/ OAT
11	("Harvey Ras onocogene, Diabetes, Long QT syndrome")/ "Best Disease"/ "Multiple Endpcrine Neoplasia"/ "Ataxia Telangiectasia"	HRAS/ IDDM2/ LQT/ VMD2/ MEN1/ ATM
12	"Zellweger Syndrome"/ "Phenylketonuria"	PXR1/ PAH
13	"Breast Cancer" / "Autosomal Recessive Neurosensory Deafness"/ "Wilson"/ "Retinoblastoma"	BRCA2/ CX26/ ATP7B/ RB1
14	"Alzheimer"/ "Alpha-1-Antitrypsin Deficiency"	PS1 (AD3)/ SERPINA1
15	"Prader-Willi Syndrome"/ "Angelman Syndrome"/ "Marfan Syndrome"/ "Tay-Sachs"	SNRPN/ UBE3A/ FBN1/ HEXA
16	"Alpha Thalassemia"/ "Famili Mediterranean Fever"/ "Polycystic Kidney"/ "Crohn's Disease"	HBA1, HBA2/ FMF/ PKD1
17	"Breast Cancer"/ "Tumor Suppressor Protein"/ "Charcot-Marie-Tooth Syndrome"	BRCA1/ p53/ CMT1A
18	"Pancreatic Cancer"/ "Niemann-Pick"	DPC4/ NPCI
19	"Severe Combined Immunodeficiency"/ "Maple Syrup Urine"/ "Atherosclerosis"/ "Myotonic Dystrophy"	Jak3/ BCKDHA/ APOE/ DMPK
20	"Severe Combined Immunodeficiency"	ADA
21	"Amylotrophic Lateral Sclerosis"/ "Autoimmune Polyglandular Syndrome"	SOD1/ APS1
22	"Glucose Galactase Malabsorption"/ "Neurofibromatosis"/ "DiGeorge Syndrome"/ "Chronic Myeloid Leukemia"	SGLT1/ NF2/ DGS/ BCR
X	"Duchenne Muscular Dystrophy"/ "Paroxysmal Nocturnal Hemoglobinuria"/ "Menkes Syndrome"/ "Alport Syndrome"/ "X-linked Severe Combined Immunodeficiency" (SCID)/ "Lesch-Nyhan Syndrome"/ "Immunodeficiency with Hyper-IgM"/ "Fragile X Syndrome"/ "Adrenoleukodystrophy"/ "Rett Syndrome"/ "Hemophilia A"	DMD/ PIG-A/ ATP7A/ COL4A5/ IL2RG/ HPRT1/ TNFSF5/ FMR1/ ALD/ MECP2/ HEMA
Y	"Testes-determining Factor"	SRY

if their corresponding genes interact with each other. The dataset includes two sets; Set_1 contains the genes (nodes) and their attributes and Set_2 contains the interactions (edges) among genes. Both sets were used to generate the network model. The following example illustrates the network generation strategy:

Assume the following set of gene-gene interaction:

#1: Gene_1("Chromosome_X") **interacts with** Gene_2("Chromosome_X")
#2: Gene_1("Chromosome_X") **interacts with** Gene_3("Chromosome_X")
#3: Gene_1("Chromosome_X") **interacts with** Gene_1("Chromosome_Y")
#4: Gene_2("Chromosome_X") **interacts with** Gene_1("Chromosome_20")
#5: Gene_3("Chromosome_X") **interacts with** Gene_1("Chromosome_5")
#6: Gene_1("Chromosome_X") **interacts with** Gene_1("Chromosome_20")

Then, the gene-gene interactions network of the above set is depicted in Figure 2.

B. Analysis Approach and Tools

The analysis approach of this work is based on network visualization and network measurements. The former enables researchers to understand the general structure of the network in terms of the relations among the genes of chromosomes X and Y as well as across chromosomes. In the latter, network measurements are involved [31]. *The Shortest Path Length* $\sigma_{(i,j)}$ measures the minimum number of edges that can connect two genes (the shortest path between them). The degree centrality (D_c) measurement, for a gene, provides information about the frequency of interactions with other genes. The betweenness centrality (B_c) measurement, provides information on how well-positioned a gene is in the network. This measurement can be used in distinguishing the most influential genes in the network. A particular gene (G_i) can be calculated as follows:

$$B_c(G_i) = \sum_{i \neq j \neq k} \frac{\sigma_{i,k}(G_i)}{\sigma_{i,k}} \quad (1)$$

where $\sigma_{i,k}(G_i)$ is the number of shortest paths between gene (G_i) and (G_k) passing through the gene (G_j). The above equation is used for all the available pairs of genes in the network.

The other measurement is the closeness centrality (C_c) measurement, which reveals how close a particular gene from other network genes. For particular genes (G_j), it can be calculated as follows:

$$C_c(G_i) = \frac{N-1}{\sum_k \sigma_{i,k}} \quad (2)$$

where N is the number of genes in the network, $\sigma_{j,k}$ is the shortest path between the genes j and k. Moreover, the Diameter (δ) of a network reflects the long distance between the farthest genes in the network. The Density (S) of a network shows the ratio of the number of actual edges

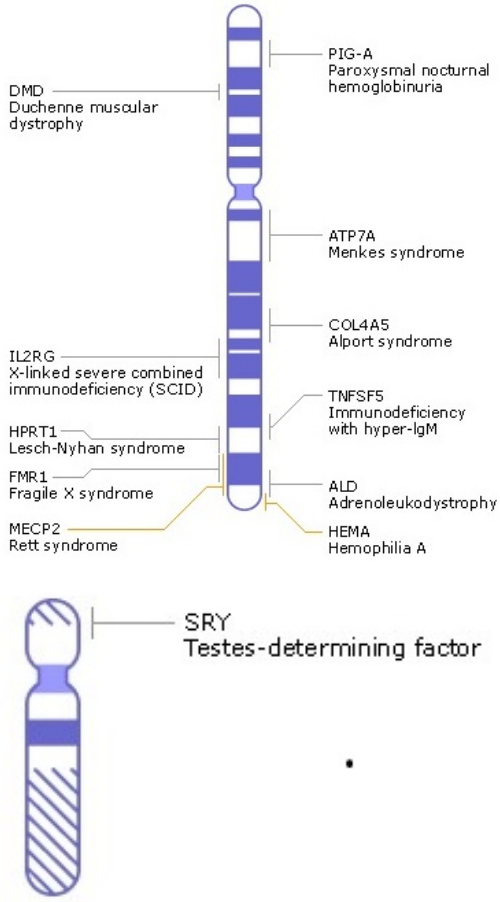

Figure 1. Human sex chromosomes (top: X & bottom: Y) [29].

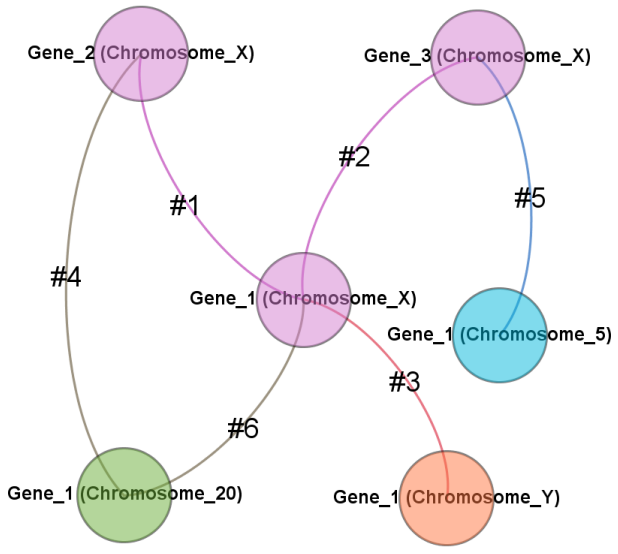

Figure 2. An example on how the gene-gene interactions network is generated Different node colors reflect different chromosomes.

among network genes to the potential (possible) edges in the network.

The generated network is termed **G**ene-g**E**ne **I**nteractions **N**etwork (**GEIN**) and will be evaluated using the aforementioned measurements as well as other analysis indicators.

Furthermore, the main tool involved in this work is Gephi, which is a network visualization and analysis open-source tool. The other visualization tool is Cytoscape, which is an open-source bioinformatics network analysis and can visualize gene interactions. The Python programming language was also used for formatting the dataset with the support of Excel workbooks. The Flourish tool was also used in generating a basic interactive visualization (*https://public.flourish.studio/visualisation/5927511/*), which makes it easier to follow by researchers.

3. RESULTS AND DISCUSSIONS

This section presents the obtained results based on the GEIN network. The strategy followed in presenting the results is to show the visualizations first, then, perform the analysis of the GEIN network. Figure 3 demonstrates the visualization of the GEIN network. It can be observed that chromosome X has relations to all other chromosomes in the GEIN network. Also, chromosome Y has one gene (SRY) in the network and only interacts with NR3C4 gene in chromosome X, which, in turn, is highly interacted with the other genes. It is also observed that the two genes IKBKG and IRAK1 in chromosome X are strongly interacted.

The general characteristics of the GEIN network are shown in Table II. According to the table, the frequency of interactions of the sex chromosomes is 572, which makes the density to be 0.002. This means the frequency of interactions among the GEIN network genes is low and follows a power-law distribution (see Figure 4). In other words, most of the interactions are performed by a few genes (e.g., NR3C4). The diameter of the GEIN network shows that there exist long distances among the genes compared to its size (732 genes). This is also confirmed when observing the shortest path length.

The strengths of the relations (levels of interactions) between the sex chromosomes and the other chromosomes are not clear in Figure 3. Therefore, another visualization is performed aiming to clarify the levels of interactions between the aforementioned chromosomes (see Figure 5). The figure shows that the genes in chromosome X are highly interacted with the genes in chromosome 17, followed by the genes of chromosome 1, and, interestingly, ending with the lowest level of interactions with the other sex chromosome (Y) includes one interaction between the NR3C4 gene in chromosome X and SRY gene in chromosome Y.

Furthermore, another visualization is performed on the GEIN network that is based on the betweenness centrality measurement. Figure 6 depicts the visualization of the

network where the high levels of betweenness centrality are represented by the bigger node sizes.

Based on Figure 6, it can be seen that the top well-positioned genes in the GEIN network are NR3C4 and WAS which belong to chromosome X. It should be mentioned that this work is focused on the sex chromosomes and the genes of other chromosomes are not of the concern. Therefore, the Src gene (chromosome 20), EGFR gene (chromosome 7), and FYN gene (chromosome 7) are not discussed even with their high levels of betweenness centrality. Figure 6 shows the distributions of the betweenness and closeness centralities of all the genes. As can be observed in Figure 7 (A), the distribution of the betweenness levels follows a power-law distribution. This means few genes are well-positioned and dominate the interactions among the genes of the same chromosome and across the chromosomes in the GEIN network. In contrast, Figure 6 (B) demonstrates that close levels of closeness centrality are shown in all the genes of the network.

In the same context, the top 10 well-positioned genes (according to their betweenness levels) in the GEIN network are presented in Table III. The reason behind considering the top 10 is that the values of betweenness significantly decreased after the top 10.

According to Table III, it is clear that NR3C4 gene is the most interacted with and the most well-positioned in the GEIN network. The NR3C4 gene interacts with genes from twenty chromosomes. This means it has relations to almost all the human chromosomes except two (chromosome 19 and chromosome 21). Biologically, the NR3C4 (also called AR) is important in male sexual activities and cell regulations [32]. The second well-positioned gene is the WAS gene which supports the immune system and cell growth [33]. According to the GEIN network, this gene interacts with genes from eleven chromosomes (chromosomes 1, 2, 3, 5, 6, 7, 15, 17, 19, 20, and 22). Based on the GEIN network, the NR3C4 and WAS genes interact with 63 and 15 genes respectively. Most of these interactions are related to controlling and regulating cell growth. The other gene SH3KBP1 is interacted with 12 genes in the GEIN network and the most frequent interactions are concerned to control cell shape and signaling cells. The FAM155B is also well-positioned in the GEIN network and most of its interactions are contributed to the immune system. The other well-positioned genes have interactions related to a variety of functions such as DNA maintenance and correction, cell cycles regulations,

The aforementioned analysis was performed based on the characteristics of the whole network. Now, we move to an in-depth analysis based on two levels; *chromosome-level* and *gene-level*.

The sex chromosomes X and Y are contributed to determining a person's sex. Biologically, females have two X chromosomes (XX), while males have one X chromosome and one Y chromosome (XY) [29]. According to this fact and the facts presented in Table I, the GEIN network is investigated in terms of the main diseases in chromosome X that are associated with the main diseases in the other chromosomes. The GEIN network shows five chromosomes with six diseases in common with chromosome X. This case is depicted in Figure 8, which is a subnetwork that is extracted from the GEIN network. More details are provided in Table IV, which presents the chromosomes that have diseases in common with chromosome X. It can be observed that even though chromosome 17 is the most frequently interacted with chromosome X, it has one main disease in common (according to the GEIN network), which is breast cancer. The authors of [37] suggested that "BRCC3 is accountable for cell radioresistance and has potential clinical relevance in breast cancer", this suggestion is approved in the GEIN network since the BRCC3 gene in chromosome X interacts with BRCC1 gene in chromosome 17, and both associated in breast cancer disease. Breast cancer is commonly found in females. The genes associated with this disease are BRCA1 in chromosome 17 and BRCC3 in chromosome X. Also, the same associated disease is found in chromosome 13 through the BRCA2 genes. Table V describes the other diseases and their corresponding genes. It is observed that the majority of the diseases in the table are more common in males than females. Also, most diseases are inherited and caused by mutations.

Although the NR3C4 gene is the most interactive in chromosome X, most of its interactions are with genes from other chromosomes. Therefore, more investigations are performed regarding the interactions across chromosomes. It is found that some pairs of genes (together) in chromosome X have participated in the interactions of the genes from other chromosomes. More precisely, there exist genes in the other chromosomes that have interactions with more than one gene in chromosome X at the same time as demonstrated in Figure 9. This can be interpreted as an indicator of the importance of particular genes in chromosome X in the interactions of the genes in the other chromosomes. The figure shows the pairs of genes in chromosome X and their interactions as pairs with individual genes in the other chromosomes. The figure also depicts two interesting genes in chromosome X (ARAF and AIFM1) that interact with three genes in chromosome 19, two genes in chromosome 20, and one gene in chromosome 2. The frequency of interactions of the two genes is six for both ARAF and AIFM1 genes. The ARAF gene's biological function is not deeply studied [43]. The other gene is AIFM1 which has a significant role in oxidative phosphorylation and redox control in healthy cells [44]. Table 5 summarizes the pairs and their contributions to other chromosomes. Finally, the presented results are of interest to the researchers in this field. Many facts about the gene interactions can be extracted using the obtained network. Also, the network can be used as a reference to quickly retrieve information about particular gene interactions.

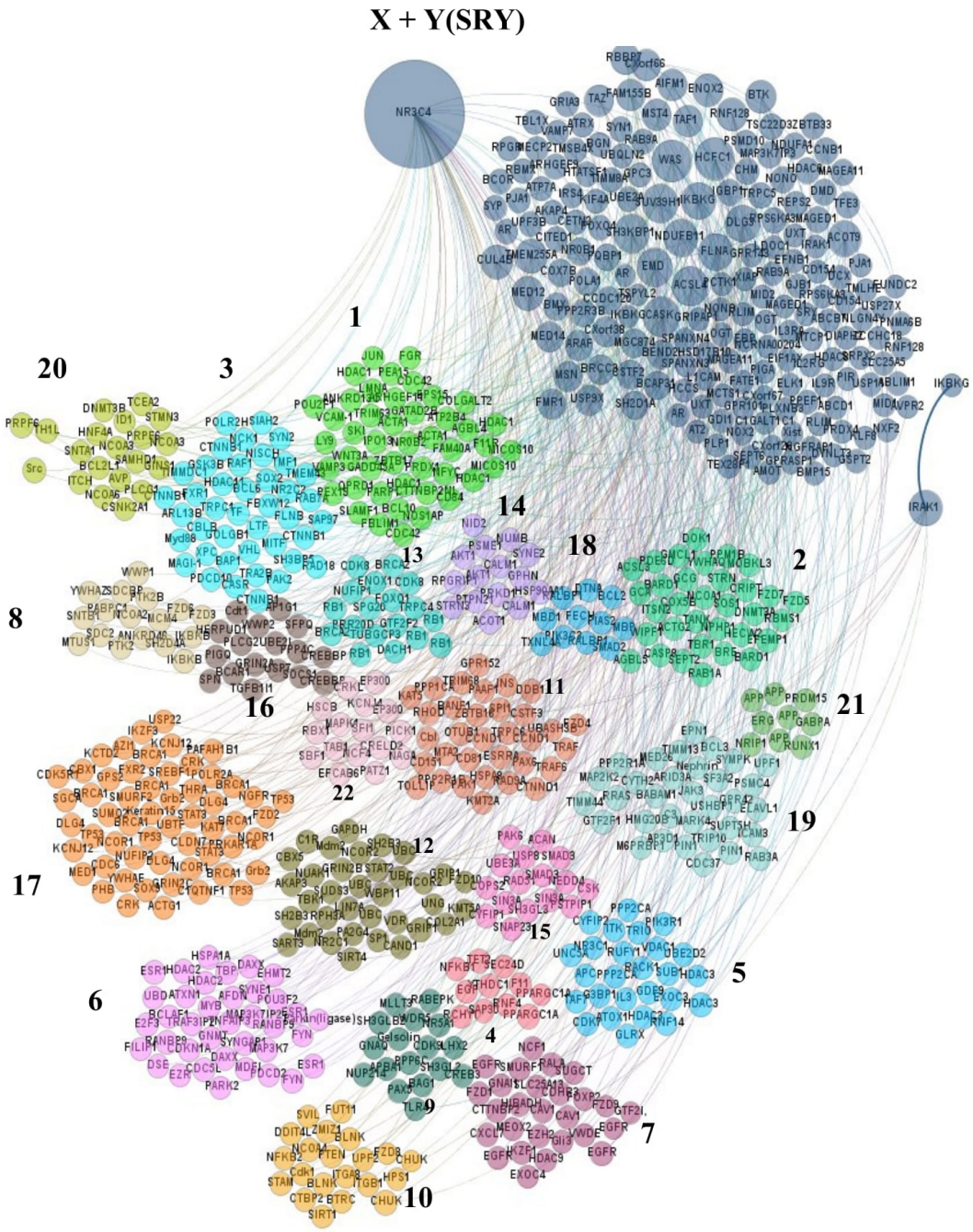

Figure 3. Visualization of the GEIN network. Different chromosomes reflected by different colors. Node size reflects the degree (frequency of interactions).

TABLE II. Characteristics of GEIN network.

Number of Nodes	**Number of Edges**	**Average Degree**	**Diameter** (δ)	**Density** (S)	**Average Shortest Path** (σ)
732	572	1.563	8	0.002	3.384

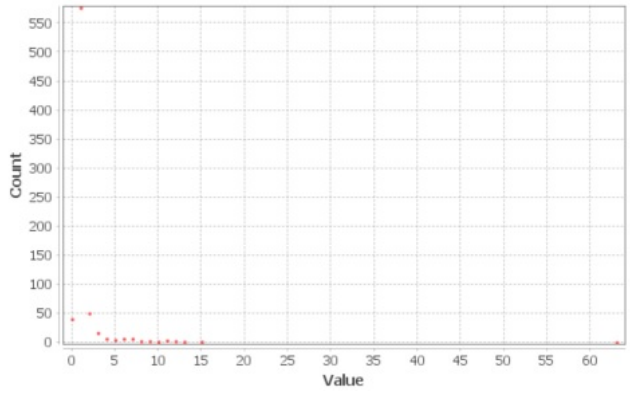

Figure 4. The distribution of the frequency of interactions among GEIN network genes.

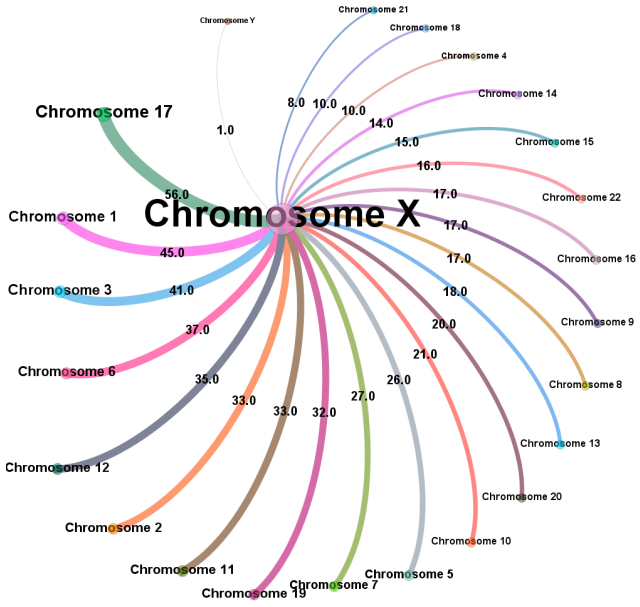

Figure 5. Frequency of interactions among the chromosomes in the GEIN network. Edges weights reflect the frequency of interactions, and the colors distinguish different chromosomes.

TABLE III. Top well-positioned genes of chromosome X along with their frequency of interactions (D_c) and closeness levels (C_c).

Rank	Gene	B_c	D_c	C_c
1^{st}	NR3C4	0.0175	63	0.425
2^{nd}	WAS	0.0117	15	0.367
3^{rd}	SH3KBP1	0.00408	12	0.234
4^{th}	FAM155B	0.00228	7	0.224
5^{th}	SUV39H1	0.00179	12	0.375
6^{th}	SH2D1A	0.00153	5	0.220
7^{th}	TAF1	0.00136	7	0.333
8^{th}	BTK	0.00084	8	0.243
9^{th}	RPS6KA3	0.000776	3	0.244
10^{th}	RBBP7	0.000731	7	0.257

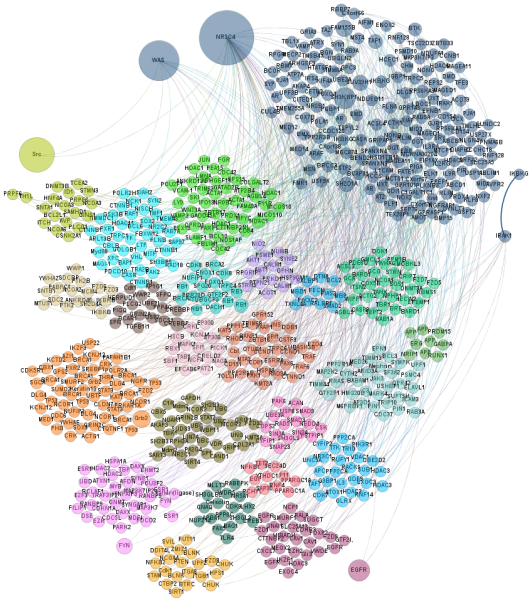

Figure 6. Visualization of the GEIN network. Node size reflects the levels of betweenness centrality.

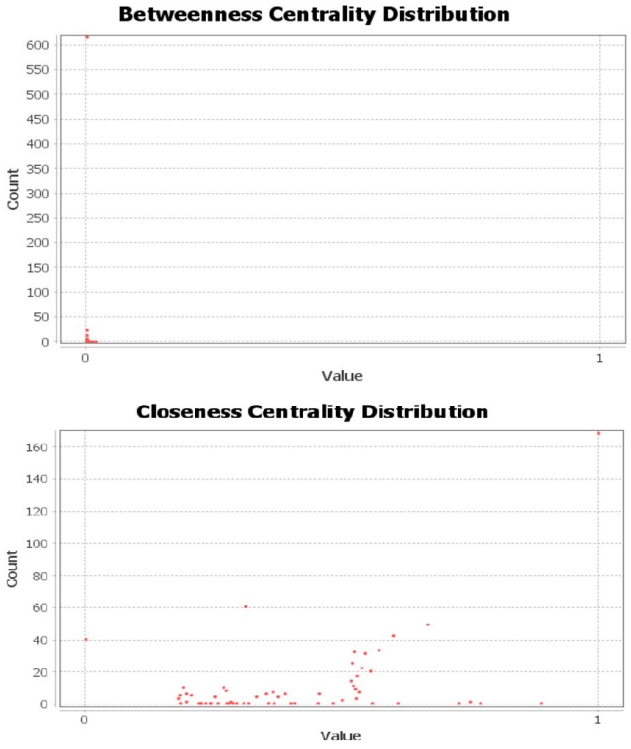

Figure 7. Distributions of betweenness and closeness centrality levels.

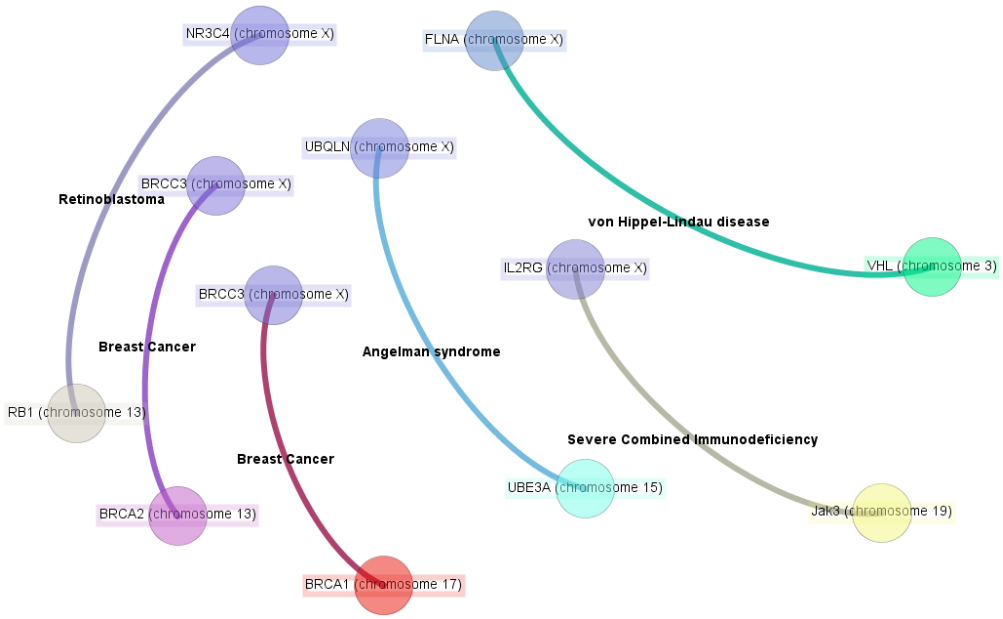

Figure 8. Visualization of associated diseases between chromosome X genes and the other chromosomes.

TABLE IV. The most common diseases between chromosome X and the other chromosomes. The table also shows the description of each associated disease along with the participated genes and the most affected sex.

Chro.	Associated Diseases with X	Interactions	Disease Description	Affected Sex
3	von Hippel-Lindau	VHL (chro. 3) **interacts with** FLNA (chro. X) [34][35]	Inherited: Blood vessel tumors of the brain, spinal cord, retina and Kidney Cancer	Males
13	Breast Cancer	BRCA2 (chro. 13) **interacts with** BRCC3 (chro. X) [36][37]	Breast Cancer	Females
13	Retinoblastoma	RB1 (chro. 13) **interacts with** NR3C4 (chro. X) [38][39]	Eye cancer that begins in retina.	Males
15	Angelman Syndrome	UBE3A (chro. 15) **interacts with** UBQLN2 (chro. X) [40]	Causes delayed development, speech and balance issues, intellectual disability, seizures.	Same
17	Breast Cancer	BRCA1 (chro. 17) **interacts with** BRCC3 (chro. X) [36][37]	Breast Cancer	Females
19	Severe Combined Immunodeficiency	Jak3 (chro. 19) **interacts with** IL2RG (chro. X) [41][42]	A group of rare disorders caused by mutations in genes involved in the development of infection-fighting immune cells.	Males

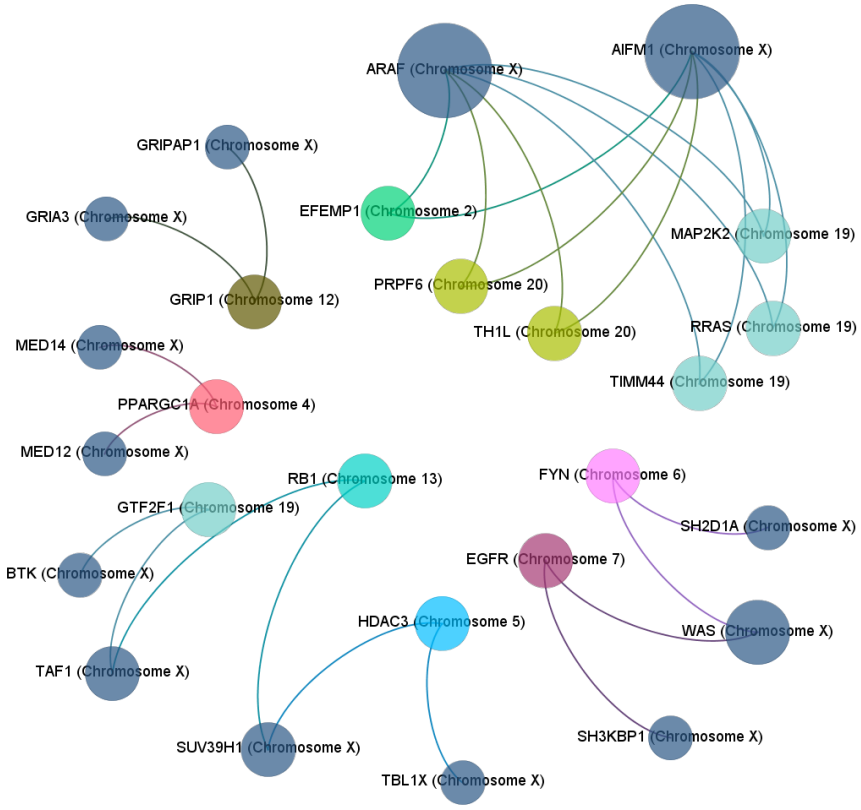

Figure 9. Visualization of chromosome X pairs of genes that interact together with genes from other chromosomes.

TABLE V. Best connected gene pairs in the GIEN network.

Pair	Interacts with	Chromosome
(ARAF-AIFM1)	MAP2K2, RRAS, and TIMM44	Chromosome 19
(ARAF-AIFM1)	PRPF6 and TH1L	Chromosome 20
(ARAF-AIFM1)	EFEMP1	Chromosome 2
(GRIPAP1-GRIA3)	GRIP1	Chromosome 12
(MED12-MED14)	PPARGC1A	Chromosome 4
(BTK-TAF1)	GTF2F1	Chromosome 19
(TAF1-SUV39H1)	RB1	Chromosome 13
(SUV39H1-TBL1X)	HDAC3	Chromosome 5
(WAS-SH3KBP1)	EGFR	Chromosome 7
(WAS-SG2D1A)	FYN	Chromosome 6

4. CONCLUSIONS

This work investigated the interactions among the genes of the human sex chromosomes. The interactions among the genes (gene-gene interactions) were modeled using concepts inspired by complex networks and the graph theory. The structural features of the whole network and its genes were measured using network science measurements. The network is also used in scrutinizing diseases and their related genes. We also measured the strength of the relations between chromosome X and the other chromosomes. The analysis showed a strong tendency for the genes of chromosome X to interact with the genes in chromosome 17. Besides, our model highlighted the best-positioned genes in chromosome X in terms of their interactions with genes of the same chromosome and across chromosomes. According to the diseases literature and the generated network model, we extracted the top interacted pairs along with their associated diseases. We strongly believe that the presented analysis can support researchers in having a wider view of the interactions among genes and their relations to diseases.

As future works, we are working on building a network model that includes the gene-gene interactions of all the human chromosomes (approximately 20,000 genes). The planned network will be useful in comprehending the genes' interactions as well as in predicting the interactions and diseases.

ACKNOWLEDGMENT

We would like to thank and appreciate all the support provided by the Departments of Computer Science and Software/University of Mosul. We also would like to thank the National Center for Biotechnology Information (NCBI) for making their data available to worldwide researchers.

REFERENCES

[1] A. D. Baxevanis, G. D. Bader, and D. S. Wishart, *Bioinformatics*. John Wiley & Sons, 2020.

[2] J. Ausborn, N. A. Shevtsova, and S. M. Danner, "Computational modeling of spinal locomotor circuitry in the age of molecular genetics," *International Journal of Molecular Sciences*, vol. 22, no. 13, p. 6835, 2021.

[3] C. J. Garroway, J. Bowman, D. Carr, and P. J. Wilson, "Applications of graph theory to landscape genetics," *Evolutionary Applications*, vol. 1, no. 4, pp. 620–630, 2008.

[4] M. Hawrylycz, J. A. Miller, V. Menon, D. Feng, T. Dolbeare, A. L. Guillozet-Bongaarts, A. G. Jegga, B. J. Aronow *et al.*, "Canonical genetic signatures of the adult human brain," *Nature neuroscience*, vol. 18, no. 12, pp. 1832–1844, 2015.

[5] E. Therman and M. Susman, *Human chromosomes: structure, behavior, and effects*. Springer Science & Business Media, 2012.

[6] B. N. Hidirov, "Mathematical and computer modelling regulatorika of hierarchical molecular-genetic systems," *Scientiae Mathematicae Japonicae*, vol. 67, no. 2, pp. 229–240, 2008.

[7] S. Petoukhov and M. He, *Symmetrical Analysis Techniques for Genetic Systems and Bioinformatics: Advanced Patterns and Applications: Advanced Patterns and Applications*. IGI Global, 2009.

[8] A. Zhang, H. Song, Y. Shen, and Y. Liu, "Construction of a gene-gene interaction network with a combined score across multiple approaches," *Genet Mol Res*, vol. 14, pp. 7018–7030, 2015.

[9] D. S. Hammadi, B. Mahmood, and M. M. Dabdawb, "Approaches on modelling genes interactions: A review," *Technium BioChemMed*, vol. 2, no. 4, pp. 38–52, 2021.

[10] M. Vidal, M. E. Cusick, and A.-L. Barabasi, "Interactome networks and human disease," *Cell*, vol. 144, no. 6, pp. 986–998, 2011.

[11] L. Salwinski and D. Eisenberg, "Computational methods of analysis of protein–protein interactions," *Current opinion in structural biology*, vol. 13, no. 3, pp. 377–382, 2003.

[12] S. Wuchty, E. Ravasz, and A.-L. Barabasi, "The architecture of biological networks," in *Complex systems science in biomedicine*. Springer, 2006, pp. 165–181.

[13] A.-L. Barabasi, Z. N. Oltvai, and S. Wuchty, "Characteristics of biological networks," in *Complex networks*. Springer, 2004, pp. 443–457.

[14] R. Albert and A.-L. Barabasi, "Statistical mechanics of complex networks," *Reviews of modern physics*, vol. 74, no. 1, p. 47, 2002.

[15] J. Lim, T. Hao, C. Shaw, A. J. Patel, G. Szabo, J.-F. Rual, C. J. Fisk, N. Li, A. Smolyar *et al.*, "A protein–protein interaction network for human inherited ataxias and disorders of purkinje cell degeneration," *Cell*, vol. 125, no. 4, pp. 801–814, 2006.

[16] J. R. Perkins, I. Diboun, B. H. Dessailly, J. G. Lees, and C. Orengo, "Transient protein-protein interactions: structural, functional, and network properties," *Structure*, vol. 18, no. 10, pp. 1233–1243, 2010.

[17] K.-I. Goh, M. E. Cusick, D. Valle, B. Childs, M. Vidal, and A.-L. Barabasi, "The human disease network," *Proceedings of the National Academy of Sciences*, vol. 104, no. 21, pp. 8685–8690, 2007.

[18] A.-L. Barabasi, N. Gulbahce, and J. Loscalzo, "Network medicine: a network-based approach to human disease," *Nature reviews genetics*, vol. 12, no. 1, pp. 56–68, 2011.

[19] S. Hwang, C. Y. Kim, S. Yang, E. Kim, T. Hart, E. M. Marcotte, and I. Lee, "Humannet v2: human gene networks for disease research," *Nucleic acids research*, vol. 47, no. D1, pp. D573–D580, 2019.

[20] D. S. Quintana, J. Rokicki, D. van der Meer, D. Alnæs, T. Kaufmann, A. Cordova-Palomera, I. Dieset, O. A. Andreassen, and L. T. Westlye, "Oxytocin pathway gene networks in the human brain," *Nature communications*, vol. 10, no. 1, pp. 1–12, 2019.

[21] S. Sinha, B. M. Jones, I. M. Traniello, S. A. Bukhari, M. S. Halfon, H. A. Hofmann, S. Huang, P. S. Katz, J. Keagy, V. J. Lynch *et al.*, "Behavior-related gene regulatory networks: A new level of organization in the brain," *Proceedings of the National Academy of Sciences*, vol. 117, no. 38, pp. 23 270–23 279, 2020.

[22] C. M. Lopes-Ramos, C.-Y. Chen, M. L. Kuijjer, J. N. Paulson, A. R. Sonawane, M. Fagny, J. Platig, K. Glass, J. Quackenbush, and D. L. DeMeo, "Sex differences in gene expression and regulatory

networks across 29 human tissues," *Cell reports*, vol. 31, no. 12, p. 107795, 2020.

[23] S. K. Banik, S. Baishya, A. Das Talukdar, and M. D. Choudhury, "Network analysis of atherosclerotic genes elucidates druggable targets," *BMC medical genomics*, vol. 15, no. 1, pp. 1–12, 2022.

[24] B. Chen, Y. Han, X. Shang, and S. Zhang, "Identifying disease related genes by network representation and convolutional neural network," *Frontiers in Cell and Developmental Biology*, vol. 9, p. 214, 2021.

[25] Y. Yu and M. Hamaneh, "Mechanism-based disease similarity: with random walks starting from a set of disease genes, traveling through the protein-protein interaction network, and then back." *Bulletin of the American Physical Society*, 2022.

[26] M. A. Scelsi, V. Napolioni, M. D. Greicius, A. Altmann, A. D. N. I. (ADNI), and the Alzheimer's Disease Sequencing Project (ADSP), "Network propagation of rare variants in alzheimer's disease reveals tissue-specific hub genes and communities," *PLoS Computational Biology*, vol. 17, no. 1, p. e1008517, 2021.

[27] W.-F. Guo, S.-W. Zhang, Y.-H. Feng, J. Liang, T. Zeng, and L. Chen, "Network controllability-based algorithm to target personalized driver genes for discovering combinatorial drugs of individual patients," *Nucleic acids research*, vol. 49, no. 7, pp. e37–e37, 2021.

[28] Q. Tan, "The sex chromosomes of the aging epigenome," *Aging (Albany NY)*, vol. 12, no. 17, p. 16667, 2020.

[29] E. W. Sayers, J. Beck, E. E. Bolton, D. Bourexis, J. R. Brister, K. Canese, D. C. Comeau, K. Funk, S. Kim, W. Klimke *et al.*, "Database resources of the national center for biotechnology information," *Nucleic acids research*, vol. 49, no. D1, p. D10, 2021. [Online]. Available: https://www.ncbi.nlm.nih.gov/gene

[30] M. Pertea and S. L. Salzberg, "Between a chicken and a grape: estimating the number of human genes," *Genome biology*, vol. 11, no. 5, pp. 1–7, 2010.

[31] B. M. Mahmood, N. A. Sultan, K. H. Thanoon, and D. S. Khadhim, "Collaboration networks: university of mosul case study," *AL-Rafidain Journal of Computer Sciences and Mathematics*, vol. 14, no. 1, pp. 117–133, 2020.

[32] S. C. Manolagas, C. A. O'brien, and M. Almeida, "The role of estrogen and androgen receptors in bone health and disease," *Nature Reviews Endocrinology*, vol. 9, no. 12, p. 699, 2013.

[33] A. J. Thrasher and S. O. Burns, "Wasp: a key immunological multitasker," *Nature Reviews Immunology*, vol. 10, no. 3, pp. 182–192, 2010.

[34] H. Tsuchiya, T. Iseda, and O. Hino, "Identification of a novel protein (vbp-1) binding to the von hippel-lindau (vhl) tumor suppressor gene product," *Cancer Research*, vol. 56, no. 13, pp. 2881–2885, 1996.

[35] M. I. Zhou, H. Wang, J. J. Ross, I. Kuzmin, C. Xu, and H. T. Cohen, "The von hippel-lindau tumor suppressor stabilizes novel plant homeodomain protein jade-1," *Journal of Biological Chemistry*, vol. 277, no. 42, pp. 39 887–39 898, 2002.

[36] Y. Dong, M.-A. Hakimi, X. Chen, E. Kumaraswamy, N. S. Cooch, A. K. Godwin, and R. Shiekhattar, "Regulation of brcc, a holoenzyme complex containing brca1 and brca2, by a signalosome-like subunit and its role in dna repair," *Molecular cell*, vol. 12, no. 5, pp. 1087–1099, 2003.

[37] Z. Tu, B. Xu, C. Qu, Y. Tao, C. Chen, W. Hua, G. Feng, H. Chang, Z. Liu, G. Li *et al.*, "Brcc3 acts as a prognostic marker in nasopharyngeal carcinoma patients treated with radiotherapy and mediates radiation resistance in vitro," *Radiation oncology*, vol. 10, no. 1, pp. 1–10, 2015.

[38] A. Sharma, W.-S. Yeow, A. Ertel, I. Coleman, N. Clegg, C. Thangavel, C. Morrissey, X. Zhang, C. E. Comstock, A. K. Witkiewicz *et al.*, "The retinoblastoma tumor suppressor controls androgen signaling and human prostate cancer progression," *The Journal of clinical investigation*, vol. 120, no. 12, pp. 4478–4492, 2010.

[39] S. Gao, Y. Gao, H. H. He, D. Han, W. Han, A. Avery, J. A. Macoska, X. Liu, S. Chen, F. Ma *et al.*, "Androgen receptor tumor suppressor function is mediated by recruitment of retinoblastoma protein," *Cell reports*, vol. 17, no. 4, pp. 966–976, 2016.

[40] M. F. Kleijnen, A. H. Shih, P. Zhou, S. Kumar, R. E. Soccio, N. L. Kedersha, G. Gill, and P. M. Howley, "The hplic proteins may provide a link between the ubiquitination machinery and the proteasome," *Molecular cell*, vol. 6, no. 2, pp. 409–419, 2000.

[41] T. Miyazaki, A. Kawahara, H. Fujii, Y. Nakagawa, Y. Minami, Z.-J. Liu, I. Oishi, O. Silvennoinen, B. A. Witthuhn, J. N. Ihle *et al.*, "Functional activation of jak1 and jak3 by selective association with il-2 receptor subunits," *Science*, vol. 266, no. 5187, pp. 1045–1047, 1994.

[42] S. M. Russell, J. A. Johnston, M. Noguchi, M. Kawamura, C. M. Bacon, M. Friedmann, M. Berg, D. W. McVicar, B. A. Witthuhn, O. Silvennoinen *et al.*, "Interaction of il-2r beta and gamma c chains with jak1 and jak3 implications for xscid and xcid," *Science*, vol. 266, no. 5187, pp. 1042–1045, 1994.

[43] E. Nekhoroshkova, S. Albert, M. Becker, and U. R. Rapp, "A-raf kinase functions in arf6 regulated endocytic membrane traffic," *PLoS One*, vol. 4, no. 2, p. e4647, 2009.

[44] P. Ferreira, R. Villanueva, M. Martinez-Julvez, B. Herguedas, C. Marcuello, P. Fernandez-Silva, L. Cabon, J. A. Hermoso, A. Lostao, S. A. Susin *et al.*, "Structural insights into the coenzyme mediated monomer–dimer transition of the pro-apoptotic apoptosis inducing factor," *Biochemistry*, vol. 53, no. 25, pp. 4204–4215, 2014.

Basim Mahmood received his M.Sc. degree in Computer Science from the University of Mosul in 2009. His Ph.D. degree was in 2015 from the College of Engineering/ Florida Institute of Technology/ USA. He is working as an associate professor at the Department of Computer Science, University of Mosul, Iraq and a member at the BioComplex Laboratory, UK. His research interests include Complex Networks and Big Data Analysis.

Marwah M. A. Dabdawb received her Bachelor degree in Software Engineering, College of Computer Science and Mathematics, University of Mosul, Mosul, Iraq in 2008. Her M.Sc. degree was from the same university in 2018. She currently works as a senior lecturer at Software Department, College of Computer Science and Mathematics, University of Mosul, Mosul Iraq. Her main research interests include Software Engineering, Artificial Intelligence, and Network Science.

Dhafar Sami Hammadi graduated from the Department of Computer Science, Collage of Computer Science and Mathematics, University of Mosul, Iraq. She worked as a programmer at the same college till 2013 when she started studying masters of science at the same college. She finished her M.Sc. degree in 2016 in the field of Artificial Intelligence. Now, she works as a senior lecturer at the College of Computer Science and Mathematics/ University of Mosul. Her research interests are Artificial Intelligence, Data Mining, and Network Science.