Received: 21 February 2024 | Revised: 26 April 2024 | Accepted: 29 April 2024 | Published online: 6 May 2024

RESEARCH ARTICLE

Medinformatics
2024, Vol. 1(3) 112–121
DOI: 10.47852/bonviewMEDIN42022652

BON VIEW PUBLISHING

Exploring Disrupted Gene Networks in Human 22q11.2 Microdeletion

Camila Cristina de Oliveira Alves[1], Ivan Rodrigo Wolf[2], Bruno Faulin Gamba[3], Lucilene Arilho Ribeiro Bicudo[3] and Guilherme Targino Valente[1,*]

[1]*Laboratory of Applied Biotechnology, São Paulo State University, Brazil*

[2]*Department of Structural and Functional Biology, São Paulo State University, Brazil*

[3]*Biological Science Institute, Federal University of Goiás, Brazil*

Abstract: Several deletions are observed at the 22q11 locus and are responsible for 22q11.2 deletion syndrome (22q11DS), also known as DiGeorge syndrome, conotruncal anomaly face syndrome, or velocardiofacial syndrome. These microdeletions on human chromosome 22 range from 0.7 to 3 Mb. Many genes are affected by 22q11.2 deletion. However, despite the well-established clinical signs for the diagnosis of 22q11.2 deletion syndrome, the interactome background of 22q11.2 deletion syndrome is unknown. Here, we analyzed protein–protein interaction networks (PPIs) to assess the influences of 3 Mb 22q11.2 deletion on this network. We compared the general human PPI network against a network without 48 genes of the 3 Mb 22q11.2 locus in a homozygous condition: we compared topological metrics, enrichment of gene ontology terms, community assignments, and edge rewiring. The PPI networks revealed that this deletion affected the relevance of hundreds of non-deleted genes. Additionally, this 22q11.2 deletion induces intense rewiring of subnetworks, promoting an accumulation of proteins associated with DiGeorge clinical signs (CTCF, YY1, TFAP2A, PPARG, PAX6, RAX, and E2F3) in a single community (community 1). Therefore, we identified new genes that may be associated with the 22q11.2 deletion syndrome. Altogether, the systemic approaches used here yielded new insights into the 22q11.2 deletion syndrome.

Keywords: DiGeorge syndrome, SD22q11, velocardiofacial syndrome, conotruncal anomaly face syndrome, systems biology

1. Introduction

The 22q11.2 region is one of the most genetically unstable areas in the human genome [1]. 22q11.2 deletion syndrome (22q11DS), also known as DiGeorge syndrome, conotruncal anomaly face syndrome, or velocardiofacial syndrome (OMIM numbers 188400 and 192430), is a condition characterized by microdeletions on human chromosome 22 ranging from 0.7 to 3 Mb. The microdeletion 22q11.2 is the most frequently observed microdeletion in humans linked to a clinical phenotype [2, 3]. 22q11.2DS affects one in every 3,000 to 6,000 live births [2, 4–7].

The 22q11.2 region has a cluster with eight low-copy repeat sequences (LCRs). LCR22 ranges from LCR22-A to LCR22-H, and this cluster facilitates nonallelic homologous recombination leading to several types of deletion [8] with a variety of chromosomal disorders [1]. Many of these deletions comprise proximal (A–B, A–D, A–E, and A–F), central (B–D, and C–D), and distal deletions (C–E, D–E, D–F, and E–F) [4]. A 3 Mb deletion that covers LCR22A to D occurs in approximately 85% of 22q11.2 deletion cases (LCR22A-D), affecting many genes.

COMT, PRODH, GNB1L, TBX1, SEPT5/GP1BB, ZDHHC8, PI4KA, and ARVC are genes linked to the 22q11.2 deletion syndrome [9–11]. TBX1 is crucial for the development of the craniofacial region, the thymus, parathyroid, aortic arch, and the cardiac outflow tract [2, 9, 12–14]. The heterozygosity of the DGCR8 gene causes neuronal abnormalities, while CRKL haploinsufficiency develops cardiac defects. Other genes in the 22q11.2 area, such as HIRA, COMT, and PRODH, are also involved in the development of primary DiGeorge clinical signs, such as congenital malformations and cognitive and/or behavioral deficits [2, 4, 15–19].

The graph theory is used in many fields such as web, airline connections, language networks, telecommunications, social networks, and others. In molecular biology, graph theory is used to model biological networks by connecting biological entities (e.g., proteins, DNA, RNA, and metabolites). The most common types of molecular biological networks analyzed are protein-protein interaction networks (PPIs), gene regulatory networks, signal transduction networks, and metabolic networks [20, 21]. PPIs are undirect graphs where each link connects proteins. PPI networks are widely studied because they are involved in many biological processes, cell structures, and biochemical reactions [22–25].

Despite well-established clinical signs for the diagnosis of 22q11.2 deletion syndrome, the interactome background of LCR22A-D 22q11.2 deletion syndrome is unknown. Therefore, we analyzed PPI networks to test the hypothesis that deleting genes from the 3 Mb 22q11.2 locus in a homozygous condition influences molecular networks by rearranging connections between proteins encoded by the remaining genes. Analysis of a PPI network simulating a homozygous 3 Mb deletion

*Corresponding author: Guilherme Targino Valente, Laboratory of Applied Biotechnology, São Paulo State University, Brazil. Email: valentegt@gmail.com

(LCR22A-D) revealed that this deletion increased the influence of hundreds of coding genes, including those associated with 22q11.2 deletion syndrome. The deletion had a notable impact on network communities, aggregating multiple proteins with ontologies relevant to the clinical signs of the 22q11.2 deletion syndrome in a single community. Finally, network analysis allowed us to rank additional genes that could be linked to the features of the 22q11.2 deletion syndrome.

2. Research Methodology

2.1. Research design

We conducted a number of PPI network analyses to evaluate the interactome of 22q11.2 deletion. Patients with 3 Mb 22q11.2 (LCR22A-D) deletion exhibit a heterozygous condition. However, due to the limitation of network analysis using graph theory, we simulated a homozygous condition with 3 Mb 22q11.2, although this is a condition not observed in patients with this deletion: we performed this simulation in an attempt to investigate how a simulated homozygous 3 Mb 22q11.2 deletion could impact the network.

The initial network analysis determined the influence of the deletion of 3 Mb 22q11.2 on the interactome. Subsequently, we evaluated the influence of the deletion of 3 Mb 22q11.2 at the community network level: communities are subnets whose nodes (in our case, proteins) are densely connected to each other rather than to nodes outside the community [22]. We also analyzed the significance of the remaining genes after the 3 Mb 22q11.2 deletion to identify new genes likely associated with the DiGeorge phenotype by analyzing communities.

An overview of our workflow is presented in Figure 1.

2.2. Network construction and analysis of the systemic impact of the 22q11.2 deletion

Overall, we modeled a PPI network to emulate the 3 Mb 22q11.2 deletion in a homozygous condition (hereafter referred to as the homozygous network, HN) by removing these genes and connections from the global human interactome (hereafter referred to as global network (GN), which emulates an individual unaffected by deletion) (further detailed).

Human PPI networks with >3 million interactions were retrieved from the public databases HuRI [26], HumanNet (v. 1) [27], ComPPi [28], and Biogrid (v. 4.4.210) [29]: the data were obtained at June 2022. Experimentally validated interactions were selected, protein identities were standardized to UniProt ID, redundant interactions were removed, and pseudogene were not filtered out to create the GN. To simulate HN, we eliminated from GN a total of 48 nodes related to proteins encoded by genes at the −3 Mb 22q11.2 locus (chr22:18,168,234 – 21,206,711, human genome version Hg38) [30], including their connections (the GN and HN networks and the genes removed are presented at the Supplementary Table S1, retrieved from https://figshare.com/s/3ba9c70ea933f19b238c).

To determine whether the 3 Mb 22q11.2 deletion significantly altered HN, we compared the degree (number of connections of each protein in the network), betweenness, and the eigenvalue of GN and HN. Centralities of these nodes were evaluated using Igraph [31], the normal distribution of metrics was estimated by applying the Shapiro–Wilk test, and the Mann–Whitney U test was used to evaluate the significance of differences between comparisons.

2.3. Analyzing GN and HN communities

The GN and HN communities were delimited using the fastgreedy.community algorithm from Igraph [31] (the largest communities are shown in Supplementary Data 1A, retrieved from https://figshare.com/s/bd2d409096a90ba3a57c).

We examined the effect of deletion within each community by evaluating the relevance of the remaining proteins in the HN. First, we identified in GN the first proteins connected to proteins encoded by genes from the 3 Mb 22q11.2 locus (henceforth referred to as "neighbor nodes") and subsequently mapped neighbor nodes in the HN communities. Using Igraph, the degree and betweenness centralities were estimated for each neighbor node in the GN and HN communities [31], and the Log2-fold-change for each parameter was obtained when comparing GN with HN. Using WEKA's "SimpleKmeans" algorithm, each Log2-fold-change was

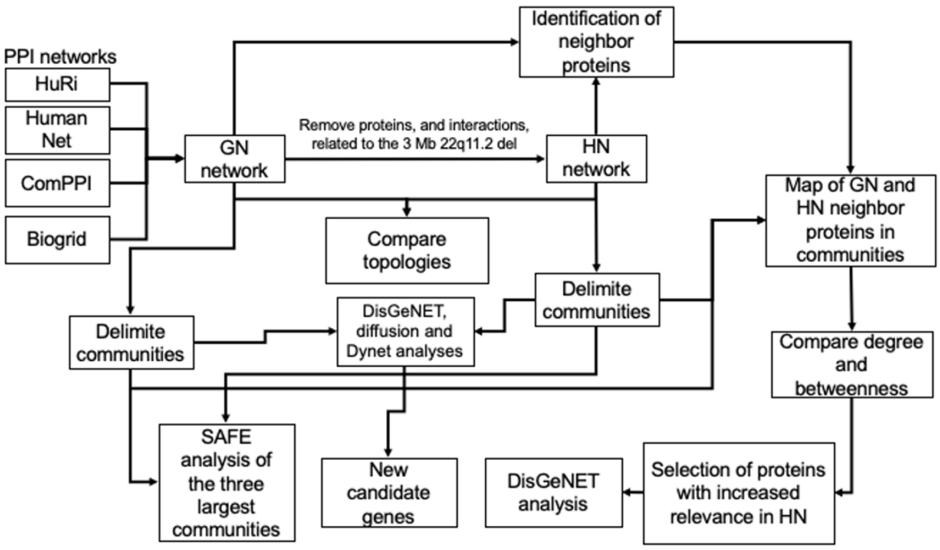

Figure 1. An overview of the analysis workflow performed here

grouped into one of three groups (clusters) signifying low, medium, or large fold-change (raising the features in the HN) [32]. As previously noted, the normal distribution of metrics and statistical tests were performed to identify the significance of the differences between the GN and HN neighbor proteins. We evaluated whether neighbor nodes related to genes with increased relevance in HN communities were associated with 22q11.2 deletion syndrome traits using the DisGeNET database in Cytoscype (gene–disease network, selected source = curated, evidence level = strong) [33]. Additionally, we determined a group of genes in HN whose relevance was reduced to 0 (betweenness = 0) and vice versa.

SAFE was used to independently assess the gene ontology (GO) enrichment of the three largest GN and HN communities [34]. We designed a matrix that includes the terms of the biological process of each protein for each community (Supplementary Data 2, retrieved from https://figshare.com/s/4d4ea6783504fde27c44). Using the mapper tool in Cytoscape [35], we mapped the proteins' UniProt IDs to HGNC IDs to facilitate protein identification in subsequent phases. The SAFE was independently executed for each community using the matrices mentioned (Supplementary Data 2, retrieved from https://figshare.com/s/4d4ea6783504fde27c44) and assuming undirected edges, distance = "Map-weighted", threshold = "0.5 percentile", background = "all nodes in the network", multi-regional landscapes = "remove", minimum landscape size = "10", landscape similarity = "Jaccard". SAFE reports GO enrichment terms as domains of each investigated network: the domain is a classification for a collection of proteins that share comparable GO terms based on their network structure [34]. REVIGO (small mode, removing obsolete GO terms, *Homo sapiens* species, and semantic similarity measured by SimRel) was used to summarize the GO enriched terms of each domain identified by SAFE for each community [36]: the main GO was regarded as the most representative one for each domain. The GO terms were accessed in June 2022.

2.4. Identified new candidate genes related to the 22q11.2 deletion-like phenotype

Initial tests revealed that the two largest communities (communities 1 and 2) had a considerably increased number of proteins in HN than in GN (further detailed). Then, we explored whether genes in communities 1 and 2 not directly related to the 3 Mb 22q11.2 deletion may play a role in the DiGeorge clinical signs.

Communities 1 and 2 of HN were categorized into two groups: 1) the "HN-shared" subnetwork, which contains proteins and interactions present in communities 1 and 2 from both HN and GN, and 2) the "HN-exclusive" subnetwork, which contains proteins and interactions present only in communities 1 and 2 of HN. Using DisGeNET implemented in Cytoscype (gene–disease association, selected source = curated, evidence level = strong) [37], we investigated which HN-shared and HN-exclusive proteins were associated with any human disorders. Using Cytoscape [38], we calculated the node's degree, betweenness, and edge's betweenness for proteins associated with human disorders from HN-shared and HN-exclusive networks; these centralities obeyed a power-law distribution. Then, we selected the most important interactions and proteins as those with the highest edge betweenness, followed by the proteins with the highest degree and betweenness (cutoffs are presented in Supplementary Table S2, retrieved from https://figshare.com/s/3ba9c70ea933f19b238c). The disorders associated with each chosen protein were then depicted within the network.

The CTCF, YY1, and TFAP2A proteins of the HN-shared community 1 are among the selected disorder-associated proteins mentioned, according to the preliminary analysis: these three proteins were associated with traits "intellectual disabilities", "congenital small ears", and "cleft upper lip" (further detailed), features of DiGeorge syndrome. Therefore, we sought additional proteins associated with disorders exhibiting DiGeorge-like traits and whether their subnetwork structure was significantly altered in HN. First, CTCF, YY1, and TFAP2A were jointly searched in community 1 of GN and HN using the diffusion algorithm implemented in Cytoscape [39]. From GN and HN, the top ~100 ranked adjacent proteins and adjacent connections were chosen (Supplementary Data 3A, retrieved from https://figshare.com/s/951db97a3f4ab4c4f977). These subnetworks were compared using Dynet (undirected networks and prefuse force-directed layout) [40], and proteins and interactions only present in HN were selected (Supplementary Data 3B-C, retrieved from https://figshare.com/s/951db97a3f4ab4c4f977). Human diseases associated with selected proteins were annotated using DisGeNET implemented in Cytoscype (gene–disease association, selected source = curated) [37]. Then, we selected proteins associated with disorders whose phenotypes are similar to those of 22q11.2 deletion syndrome (e.g., heart failure, craniofacial abnormalities, and intellectual disabilities) (Supplementary Data 3D, retrieved from https://figshare.com/s/951db97a3f4ab4c4f977). Finally, we retrieved the selected proteins from both GN and HN communities to identify to which communities they were assigned.

3. Results

The GN comprised 21,492 proteins (nodes), whereas the HN lacked 48 nodes related to proteins encoded by the 3 Mb 22q11.2 locus genes (Table 1; Supplementary Table S1, retrieved from https://figshare.com/s/3ba9c70ea933f19b238c). The networks exhibited a power-law degree distribution, and the specific topological properties of GN and HN were comparable (Table 1; Figure 2A).

Table 1. Topological properties of the global network (GN) and the homozygous network (HN)

Topological property	GN	HN
Number of interactions	682,911	679,817
Number of proteins	21,492	21,444
Avg. degree	63.5	63.4
Avg. betweenness	18,693.76	18,660.42
Avg. eigenvalue	0.031	0.031
Avg. diameter	8	8
Avg. path length	2.74	2.74

Note: Avg: average

GN and HN had 13 and 27 communities, respectively, and had 2,404 and 2,389 neighbor nodes (Table 2). The distribution of Log2-fold-changes comparing GN versus HN was skewed to the right side, indicating that the HN communities exhibited greater degree and betweenness than GN communities (Figure 2B–C). A total of 361 proteins had a significant increase in relevance in HN, of which 111 proteins were related to the 22q11.2 deletion syndrome (Supplementary Figure S1; Supplementary Table S3, retrieved from https://figshare.com/s/3ba9c70ea933f19b238c). Moreover, 112 neighbor nodes are proteins associated with human diseases (Supplementary Table S4, retrieved from https://figshare.com/s/3ba9c70ea933f19b238c). Additionally, three proteins related to neighbor nodes (nodes present in both GN and HN linked to proteins encoded by the 3 Mb 22q11.2 loci) had null betweenness in the HN (SEPTIN 10, SLC37A1 and SOWAHC): SEPTIN 10 interacts with SEPTIN 7, SLC37A1 interacts with GABPA, SP1, and SOWAHC

Table 2. Features of GN and HN communities

Community	Num. Proteins		Num. Proteins of 22q11.2 loci	Neighbor nodes	
	GN	HN	GN	GN	HN
1	7,734	10,882	25	497	675
2	4,338	5,517	6	464	667
3	4,217	3,884	7	468	1,025
4	2,819	343	7	699	18
5	1,630	741	4	274	3
6	723	19	–	2	1
7	9	10	–	–	–
8	6	8	–	–	–
9	7	9	–	–	–
10	3	3	–	–	–
11	2	6	–	–	–
12	2	7	–	–	–
13	2	3	–	–	–

interacts with PEX12. Conversely, 2 proteins had a substantial increase in importance in HN (Table 3).

GO enrichment analysis of the three largest communities of HN and GN revealed that network rewiring caused by 3 Mb 22q11.2 deletion altered the functional properties of the communities (Supplementary

Table 3. Proteins with high changes in betweenness when comparing GC vs. HC

UniProt id: Entry name	GN			HN		
	com.	D	BET	com.	D	BET
B5ME97: SEPTIN 10	3rd	2	33.6	1st	1	0
J3KNL2: SEPTIN 1	3rd	3	0	1st	4	19.5
P57057: SLC37A1	2nd	2	95.9	1st	2	0
Q53LP3: SOWAHC	1st	6	234.5	4th	1	0
Q9Y5A7: NUB1	6th	1	0	1st	11	386.5

Note: Com: the community number where a specific protein is located. D: degree centrality. BET: betweenness centrality

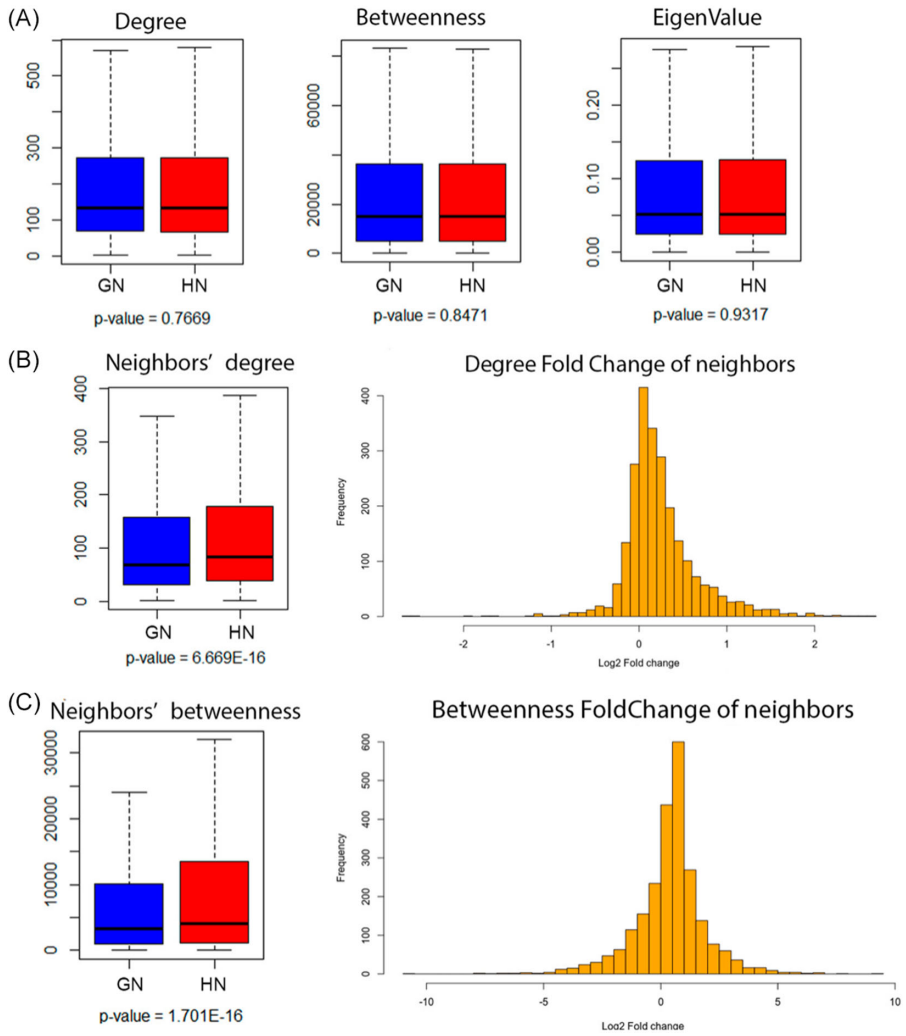

Figure 2. The main topological features of GN and HN networks. (A) Boxplot comparing the degree, betweenness, and eigenvalue of GN and HN. (B–C) Boxplots on the left report the degree and betweenness of the neighbor nodes from the GN and HN communities. The frequency distribution on the right represents the Log2-fold-change comparing the degree and betweenness of neighbor nodes from GN and HN communities.

Figure S2; Supplementary Table S5, retrieved from https://figshare.com/s/3ba9c70ea933f19b238c). Interestingly, the two largest HN communities contain more terms (absolute values) associated with the DiGeorge phenotypes than the GN communities 1–3, including terms associated with signal propagation between neurons, the heart, and skeletal development (Figure 3). We highlight the presence of TSSK2 among proteins with terms enriched in GN (establishment of cell polarity, cell division, regulation of mitotic cell cycle, cell cycle, intracellular signal transduction, protein autophosphorylation, and protein phosphorylation).

The subsequent analysis focused on the two largest communities because they had significantly more proteins in HN than in GN (Figure 4A–B; Table 2; Supplementary Data 1A, retrieved from https://figshare.com/s/bd2d409096a90ba3a57c). The HN-shared and HN-exclusive subnetworks are derived from HN communities 1 and 2, and their protein counts differ significantly (Figure 4C; Supplementary Data 1B, retrieved from https://figshare.com/s/bd2d409096a90ba3a57c). Few HN-shared and HN-exclusive proteins are associated with human disease, with CACNA1B, CTNND1, CTCF, YY1, and TFAP2A being the most relevant genes (Figure 4D–E; Supplementary Data 1C, retrieved from https://figshare.com/s/bd2d409096a90ba3a57c): these genes are involved in DiGeorge-like phenotypes. Other proteins that interact with CTCF, YY1, and TFAP2A and are associated with disorders with DiGeorge-like traits include PPARG, PAX6, RAX, and E2F3 (Supplementary Data 3D, retrieved from https://figshare.com/s/951db97a3f4ab4c4f977). Interestingly, the PPARG, PAX6, RAX, and E2F3 proteins were relocated to a single community in HN (Figure 4F).

4. Discussion

The network as a whole was not affected by 3 Mb 22q11.2 deletion, as both GN and HN displayed a power-law degree distribution; this distribution fits the Barabasí-Alberts model, reflecting on the robustness against random perturbation (in this case, random effects unlikely perturbs highly connected proteins), and indicating the reliability of our networks [41]. However, the simulated 3 Mb 22q11.2 homozygous deletion altered the significance of the remaining proteins, the number of proteins/connections within each network community, and consequently the main functions of each community. Finally, 3 Mb 22q11.2 homozygous deletion had a huge impact on community-level systems rather than on the systems as a whole.

The marked increase in systemic significance of numerous proteins demonstrates the profound network impact of the 3 Mb 22q11.2 deletion. These proteins bind to proteins encoded by 22q11.2 3 Mb genes in GN, of which ~30% are associated with 22q11.2 deletion syndrome (e.g., TBX1 and DGCR6, corroborating previous findings [13, 42–44]). However, some increased relevance proteins, such as SEPTIN 1 and NUB1, were never associated with the 22q11.2 deletion syndrome; NUB1 is associated with Huntington's disease [45], which also causes some traits observed in patients with 22q11.2 deletion. Interestingly, the simulated 3 Mb 22q11.2 homozygous deletion nullified the relevance of other genes, such as SEPTIN 10, SLC37A1, and SOWAHC. These genes were never associated with 22q11.2 deletion syndrome, although deletion of

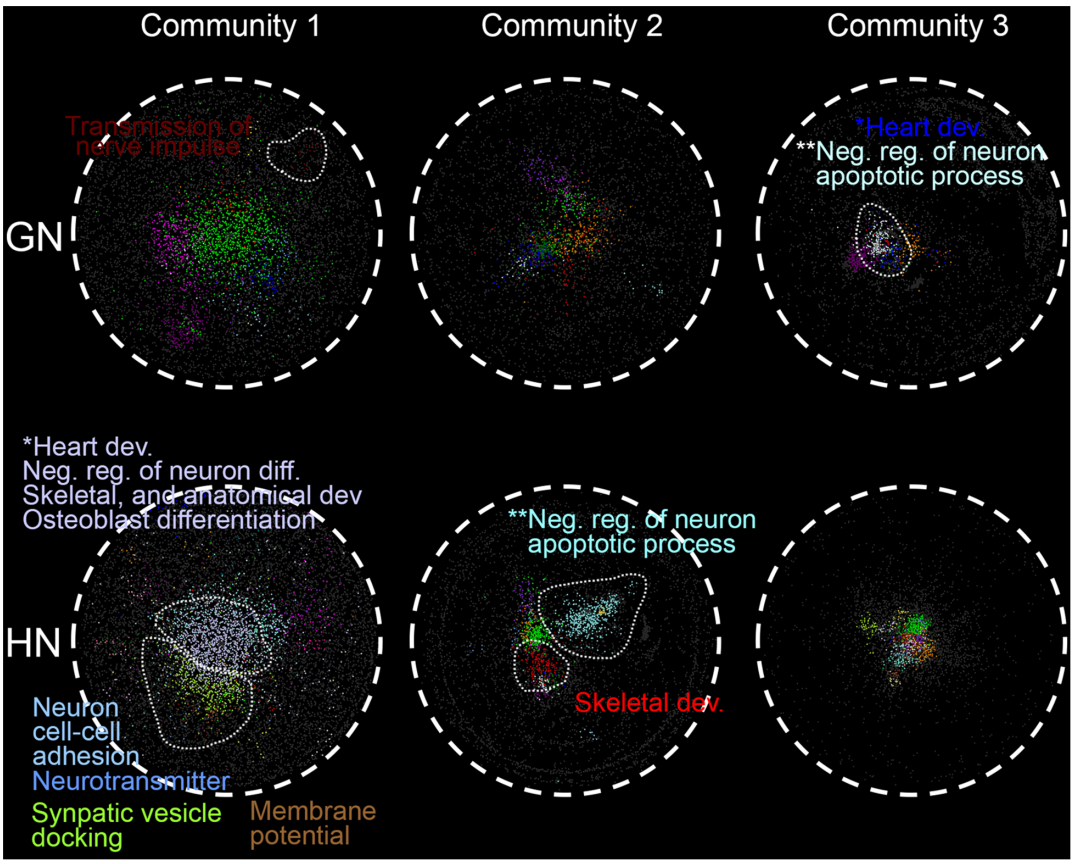

Figure 3. Annotation of GO terms associated with the DiGeorge phenotype in the three largest HN and GN communities

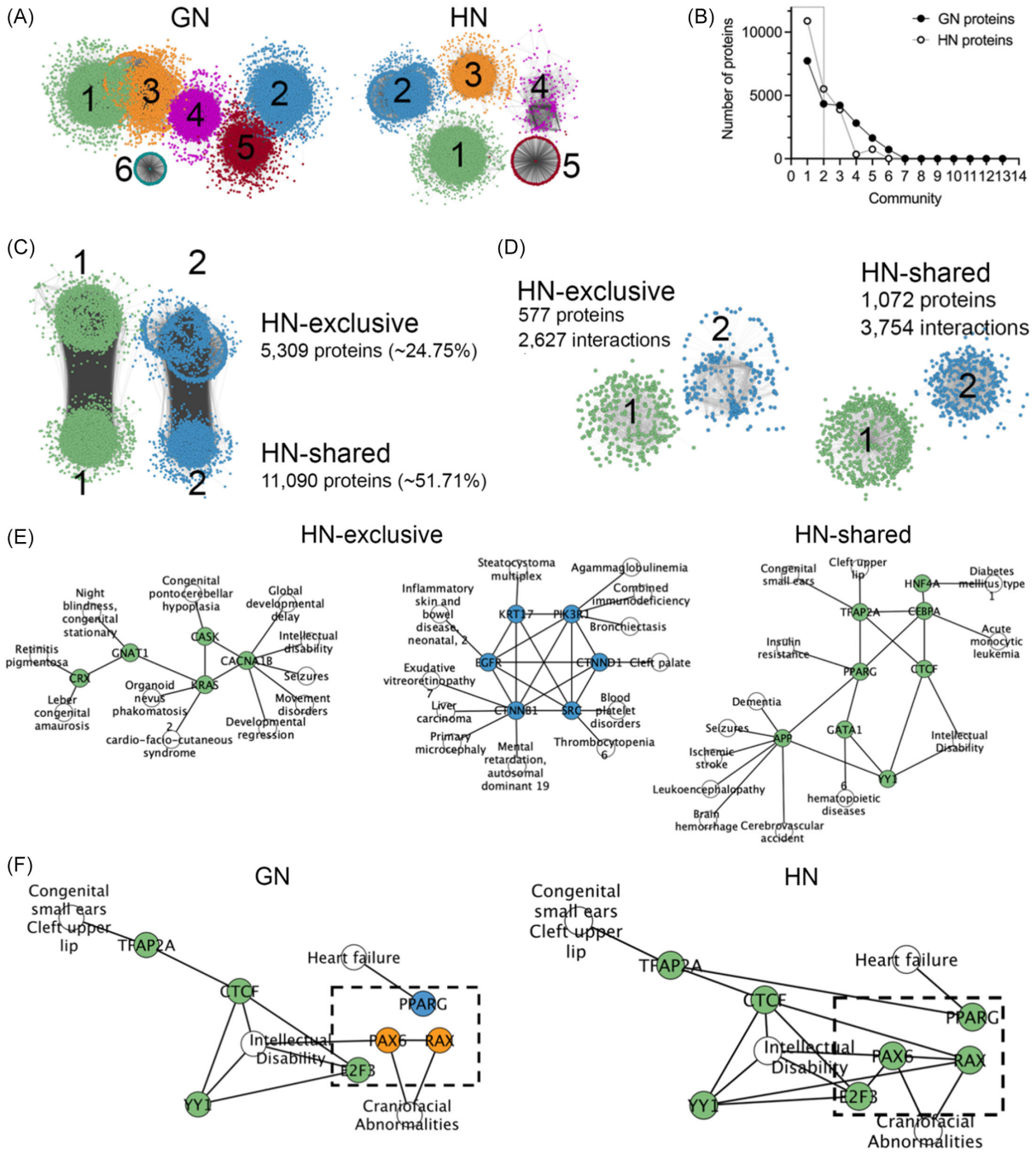

Figure 4. Analyses of the HN and GN communities. (A) The largest HN and GN communities. (B) Distribution of protein counts of the first thirteen GN and HN communities. The dashed box outlines the communities in which the number of proteins in HN is higher than that in GN. (C) Substructures of HN communities 1 and 2. The proteins of the HN-exclusive subnetwork are not present in GN communities 1 and 2, whereas the proteins of the shared subnetwork are fully present in GN communities 1 and 2. The percentages are the number of proteins in relation to the full HN. (D) The subnetworks containing proteins associated with any human disorder described in DisGeNET. (E) The main proteins (and their interactions) associated with human disorders. The node's colors correspond to the communities depicted in A. (F) Analysis CTCF, YY1, and TFAP2A interactors to identify additional proteins associated with diseases displaying DiGeorge-like traits. The node's colors correspond to the communities depicted in A.

SLC37A1 (G6PT2) causes failure to thrive, doll-shaped facial appearance, and short stature [46]. SEPTIN 10 interacts with SEPTIN 7, SLC37A1 interacts with GABPA, SP1, and SOWAHC interacts with PEX12. Interestingly, SEPTIN 7 is related to the impaired angiogenesis [47], and GABPA is related to cognitive disorders [48].

Furthermore, the impact of homozygous deletion appears to significantly alter the functionalities of communities. Indeed, we found that in the presence of homozygous deletion, terms associated with DiGeorge traits, such as skeletal, heart, and neuron development and the neuron activity terms, shifted from multiple communities to communities 1 and 2.

Altogether, to maintain relatively regular network activity in the presence of 3 Mb 22q11.2 homozygous deletion, we propose that the systems undergo intensive community rewiring that alters the relevance of hundreds of proteins. This rewiring is responsible for the repositioning between bulky proteins, creating simpler subnets, and shortening the information flow among these proteins, which might facilitate the emergence of DiGeorge clinical signs.

Deeper analysis of communities 1 and 2 of the HN revealed that the CTNND1, CTCF, YY1, TFAP2A, PPARG, PAX6, RAX, E2F3, and CACNA1B non-deleted genes might be associated with the phenotypes of 22q11.2 deletion, such as intellectual disability, craniofacial abnormalities, heart failure, congenital small ears, and upper lip cleft. Most of the proteins related to these genes were found in community 1 of HN, whereas they were dispersed in 3 communities of GN as a result of the intensive rewiring discussed. Overall, the analysis revealed that the CACNA1B, CTNND1, CTCF, YY1, TFAP2A, PPARG, PAX6, RAX, and E2F3 genes might be associated with DiGeorge traits, although experimental evidence is needed.

The expression of the CACNA1B gene is widespread throughout the central nervous system. This gene is associated with global developmental delay, intellectual disability, and movement disorders in progressive epilepsy-dyskinesia [49], and the lack of CACNA1B in mice manifests neurodevelopmental abnormalities [49, 50]. The TFAP2A mutant alleles lead to a syndrome characterized by branchial sinus defects and lip or palate cleft [51]. CTCF is associated with microcephaly, growth retardation, autistic traits, neurodevelopmental disorders, and intellectual disabilities [52–54]. Yy1 heterozygous mice exhibit growth retardation, neurulation defects, intellectual disability, and brain abnormalities [53, 55]. PPARG regulates circadian cardiovascular rhythms and appears to have a cardioprotective effect [56–58]. PAX6 regulates the development of the central nervous system, and a patient with a mental disorder had a deletion in the adjacent region of this gene [59, 60]. The lack of the RAX gene in mice resulted in the loss of the structures of the ventral forebrain and the palate. A patient with a bilateral lip and palate cleft presented a mutation in this gene [61]. A patient with mental retardation had a deletion of E2F3 [62]. Finally, CTNND1 plays an important role in tooth development in mice [63–65].

A patient with infertility presented a gonosomal mosaic of chromosome 22q11.2 deletion [66], a male with mild dysmorphic features, hypernasal voice, and mental retardation had azoospermia [67], and a patient with supernumerary inv dup (22) (q11.1) had infertility and hypogonadotropic hypogonadism [68]. TSSK2 is one of the proteins with enriched biological process terms in GN. The TSSK2 gene is located in the 22q11.2 3 Mb deleted region [69]; it is expressed in human sperm, and polymorphisms in this gene are associated with spermatogenesis impairment [70, 71]. Then, we speculate that there is an association between the 3 Mb 22q11.2 deletion syndrome, TSSK2, and male fertility, although numerous cases must be evaluated to test this hypothesis because the phenotypic spectrum of this syndrome is broad.

5. Conclusion, Limitations, and Recommendations

Overall, the 3 Mb 22q11.2 deletion had a huge impact on community-level systems rather than on the systems as a whole. In this case, intense rewiring of the network promoted the accumulation of proteins associated with DiGeorge clinical signs in a single community. The importance of several proteins was altered in the presence of 3 Mb 22q11.2 deletion. Furthermore, we found a set of new genes that might be related to the characteristics of the 22q11.2 deletion syndrome.

Concerning the limitations of our studies, because of the limitations of graph theory, we could not model the patient network under the heterozygous condition. Then, additional studies must be performed to improve graph theory to study genetic disorders in heterozygous condition. Furthermore, the new candidate genes here ranked as possibly relevant for the 22q11.2 deletion syndrome must be evaluated in model species with homologous 22q11.2 deletion, by analyzing the expression of these genes in patients affected by this deletion and by genetic population studies coupled to the expression of quantitative trait loci analysis.

Funding Support

Sao Paulo Research Foundation (2015/12093-9, 2015/19211-7, and 2017/08463-0) and the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES, Brasil, Finance Code 001).

Conflicts of Interest

The authors declare that they have no conflicts of interest to this work.

Data Availability Statement

The data that support the findings of this study are openly available in HuRI at http://www.interactome-atlas.org/, reference number [26]. The data that support the findings of this study are openly available in HumanNet at https://staging2.inetbio.org/humannetv3/, reference number [27]. The data that support the findings of this study are openly available in ComPPi at https://comppi.linkgroup.hu/, reference number [28]. The data that support the findings of this study are openly available in Biogrid at https://thebiogrid.org/, reference number [29].

Supplementary Information

The supplementary files are available at https://doi.org/10.47852/bonviewMEDIN42022652.

References

[1] Vervoort, L., & Vermeesch, J. R. (2022). The 22q11.2 low copy repeats. *Genes*, *13*(11), 2101. https://doi.org/10.3390/genes13112101

[2] McDonald-McGinn, D. M., Sullivan, K. E., Marino, B., Philip, N., Swillen, A., Vorstman, J. A. S., . . . , & Bassett, A. S. (2015). 22q11.2 deletion syndrome. *Nature Reviews Disease Primers*, *1*(1), 15071. https://doi.org/10.1038/nrdp.2015.71

[3] Szczawińska-Popłonyk, A., Schwartzmann, E., Chmara, Z., Głukowska, A., Krysa, T., Majchrzycki, M., . . . , & Babik, J. (2023). Chromosome 22q11.2 deletion syndrome: A comprehensive review of molecular genetics in the context of multidisciplinary clinical approach. *International Journal of Molecular Sciences*, *24*(9), 8317. https://doi.org/10.3390/ijms24098317

[4] Burnside, R. D. (2015). 22q11.21 deletion syndromes: A review of proximal, central, and distal deletions and their associated features. *Cytogenetic and Genome Research*, *146*(2), 89–99. https://doi.org/10.1159/000438708

[5] Dugoff, L., Mennuti, M. T., & McDonald-McGinn, D. M. (2017). The benefits and limitations of cell-free DNA screening for 22q11.2 deletion syndrome. *Prenatal Diagnosis*, *37*(1), 53–60. https://doi.org/10.1002/pd.4864

[6] Fiksinski, A. M., Schneider, M., Zinkstok, J., Baribeau, D., Chawner, S. J. R. A., & Vorstman, J. A. S. (2021). Neurodevelopmental trajectories and psychiatric morbidity: Lessons learned from the 22q11.2 deletion syndrome. *Current Psychiatry Reports*, *23*(3), 13. https://doi.org/10.1007/s11920-021-01225-z

[7] Panamonta, V., Wichajarn, K., Chaikitpinyo, A., Panamonta, M., Pradubwong, S., & Chowchuen, B. (2016). Birth prevalence of chromosome 22q11.2 deletion syndrome: A systematic review of population-based studies. *Journal of the Medical Association of Thailand*, *99*, S187–S193.

[8] Gavril, E. C., Popescu, R., Nucă, I., Ciobanu, C. G., Butnariu, L. I., Rusu, C., & Pânzaru, M. C. (2022). Different types of deletions created by low-copy repeats sequences location in 22q11.2 deletion syndrome: Genotype–phenotype correlation. *Genes*, *13*(11), 2083. https://doi.org/10.3390/genes13112083

[9] Baldini, A., Fulcoli, F. G., & Illingworth, E. (2017). Tbx1: Transcriptional and developmental functions. In M. Frasch (Ed.), *Current topics in developmental biology*, *122* (pp. 223–243). Elsevier. https://doi.org/10.1016/bs.ctdb.2016.08.002

[10] Du, Q., de la Morena, M. T., & van Oers, N. S. (2020). The genetics and epigenetics of 22q11.2 deletion syndrome. *Frontiers in Genetics*, *10*, 1365. https://doi.org/10.3389/fgene.2019.01365

[11] Motahari, Z., Moody, S. A., Maynard, T. M., & LaMantia, A. S. (2019). In the line-up: Deleted genes associated with DiGeorge/22q11.2 deletion syndrome: Are they all suspects? *Journal of Neurodevelopmental Disorders*, *11*(1), 7. https://doi.org/10.1186/s11689-019-9267-z

[12] Chen, J., Zhang, X., Li, J., Song, C., Jia, Y., & Xiong, W. (2016). Identification of a novel ENU-induced mutation in mouse Tbx1 linked to human DiGeorge syndrome. *Neural Plasticity*, *2016*(1), 5836143. https://doi.org/10.1155/2016/5836143

[13] Gao, S., Li, X., & Amendt, B. A. (2013). Understanding the role of Tbx1 as a candidate gene for 22q11.2 deletion syndrome. *Current Allergy and Asthma Reports*, *13*(6), 613–621. https://doi.org/10.1007/s11882-013-0384-6

[14] Sullivan, K. E. (2019). Chromosome 22q11.2 deletion syndrome and DiGeorge syndrome. *Immunological Reviews*, *287*(1), 186–201. https://doi.org/10.1111/imr.12701

[15] Ju, Z. R., Wang, H. J., Ma, X. J., Ma, D., & Huang, G. Y. (2016). HIRA gene is lower expressed in the myocardium of patients with tetralogy of Fallot. *Chinese Medical Journal*, *129*(20), 2403–2408. https://doi.org/10.4103/0366-6999.191745

[16] Morrow, B. E., McDonald-McGinn, D. M., Emanuel, B. S., Vermeesch, J. R., & Scambler, P. J. (2018). Molecular genetics of 22q11.2 deletion syndrome. *American Journal of Medical Genetics Part A*, *176*(10), 2070–2081. https://doi.org/10.1002/ajmg.a.40504

[17] Racedo, S. E., McDonald-McGinn, D. M., Chung, J. H., Goldmuntz, E., Zackai, E., Emanuel, B. S., . . . , & Morrow, B. E. (2015). Mouse and human CRKL is dosage sensitive for cardiac outflow tract formation. *The American Journal of Human Genetics*, *96*(2), 235–244. https://doi.org/10.1016/j.ajhg.2014.12.025

[18] Yang, J. H., Song, T. Y., Jo, C., Park, J., Lee, H. Y., Song, I., . . . , & Cho, E. J. (2016). Differential regulation of the histone chaperone HIRA during muscle cell differentiation by a phosphorylation switch. *Experimental & Molecular Medicine*, *48*(8), e252. https://doi.org/10.1038/emm.2016.68

[19] Zeitz, M. J., Lerner, P. P., Ay, F., van Nostrand, E., Heidmann, J. D., Noble, W. S., & Hoffman, A. R. (2013). Implications of COMT long-range interactions on the phenotypic variability of 22q11.2 deletion syndrome. *Nucleus*, *4*(6), 487–493. https://doi.org/10.4161/nucl.27364

[20] Koutrouli, M., Karatzas, E., Paez-Espino, D., & Pavlopoulos, G. A. (2020). A guide to conquer the biological network era using graph theory. *Frontiers in Bioengineering and Biotechnology*, *8*, 34. https://doi.org/10.3389/fbioe.2020.00034

[21] Pavlopoulos, G. A., Secrier, M., Moschopoulos, C. N., Soldatos, T. G., Kossida, S., Aerts, J., . . . , & Bagos, P. G. (2011). Using graph theory to analyze biological networks. *BioData Mining*, *4*, 10. https://doi.org/10.1186/1756-0381-4-10

[22] Junker, B. H., & Schreiber, F. (2008). *Analysis of biological networks*. USA: Wiley.

[23] Martino, E., Chiarugi, S., Margheriti, F., & Garau, G. (2021). Mapping, structure and modulation of PPI. *Frontiers in Chemistry*, *9*, 718405. https://doi.org/10.3389/fchem.2021.718405

[24] Nelson, D. L., & Cox, M. M. (2021). *Lehninger principles of biochemistry*. USA: W. H. Freeman.

[25] Pizzuti, C., & Rombo, S. E. (2014). Algorithms and tools for protein–protein interaction networks clustering, with a special focus on population-based stochastic methods. *Bioinformatics*, *30*(10), 1343–1352. https://doi.org/10.1093/bioinformatics/btu034

[26] Rolland, T., Taşan, M., Charloteaux, B., Pevzner, S. J., Zhong, Q., Sahni, N., . . . , & Vidal, M. (2014). A proteome-scale map of the human interactome network. *Cell*, *159*(5), 1212–1226. https://doi.org/10.1016/j.cell.2014.10.050

[27] Lee, I., Blom, U. M., Wang, P. I., Shim, J. E., & Marcotte, E. M. (2011). Prioritizing candidate disease genes by network-based boosting of genome-wide association data. *Genome Research*, *21*, 1109–1121. https://doi.org/10.1101/gr.118992.110

[28] Veres, D. V., Gyurkó, D. M., Thaler, B., Szalay, K. Z., Fazekas, D., Korcsmáros, T., & Csermely, P. (2015). ComPPI: A cellular compartment-specific database for protein–protein interaction network analysis. *Nucleic Acids Research*, *43*(D1), D485–D493. https://doi.org/10.1093/nar/gku1007

[29] Oughtred, R., Stark, C., Breitkreutz, B. J., Rust, J., Boucher, L., Chang, C., . . . , & Tyers, M. (2019). The BioGRID interaction database: 2019 update. *Nucleic Acids Research*, *47*(D1), D529–D541. https://doi.org/10.1093/nar/gky1079

[30] Bertini, V., Azzarà, A., Legitimo, A., Milone, R., Battini, R., Consolini, R., & Valetto, A. (2017). Deletion extents are not the cause of clinical variability in 22q11.2 deletion syndrome: Does the interaction between DGCR8 and miRNA-CNVs play a major role? *Frontiers in Genetics*, *8*, 47. https://doi.org/10.3389/fgene.2017.00047

[31] Kolaczyk, E. D., & Csárdi, G. (2014). *Statistical analysis of network data with R*. USA: Springer.

[32] Witten, I. H., Frank, E., & Hall, M. A. (2011). *Data mining: Practical machine learning tools and techniques*. Netherlands: Elsevier.

[33] Piñero, J., Ramírez-Anguita, J. M., Saüch-Pitarch, J., Ronzano, F., Centeno, E., Sanz, F., & Furlong, L. I. (2020). The DisGeNET knowledge platform for disease genomics: 2019 update. *Nucleic Acids Research*, *48*(D1), D845–D855. https://doi.org/10.1093/nar/gkz1021

[34] Baryshnikova, A. (2016). Spatial analysis of functional enrichment (SAFE) in large biological networks. *bioRxiv Preprint*. https://doi.org/10.1101/094904

[35] Shannon, P., Markiel, A., Ozier, O., Baliga, N. S., Wang, J. T., Ramage, D., . . . , & Ideker, T. (2003). Cytoscape: A software environment for integrated models of biomolecular interaction networks. *Genome Research*, *13*, 2498–2504. https://doi.org/10.1101/gr.1239303

[36] Supek, F., Bošnjak, M., Škunca, N., & Šmuc, T. (2011). REVIGO summarizes and visualizes long lists of Gene Ontology terms. *PLOS ONE*, *6*(7), e21800. https://doi.org/10.1371/journal.pone.0021800

[37] Piñero, J., Bravo, À., Queralt-Rosinach, N., Gutiérrez-Sacristán, A., Deu-Pons, J., Centeno, E., ..., & Furlong, L. I. (2017). DisGeNET: A comprehensive platform integrating information on human disease-associated genes and variants. *Nucleic Acids Research*, *45*(D1), D833–D839. https://doi.org/10.1093/nar/gkw943

[38] Doncheva, N. T., Assenov, Y., Domingues, F. S., & Albrecht, M. (2012). Topological analysis and interactive visualization of biological networks and protein structures. *Nature Protocols*, *7*(4), 670–685. https://doi.org/10.1038/nprot.2012.004

[39] Carlin, D. E., Demchak, B., Pratt, D., Sage, E., & Ideker, T. (2017). Network propagation in the cytoscape cyberinfrastructure. *PLOS Computational Biology*, *13*(10), e1005598. https://doi.org/10.1371/journal.pcbi.1005598

[40] Goenawan, I. H., Bryan, K., & Lynn, D. J. (2016). DyNet: Visualization and analysis of dynamic molecular interaction networks. *Bioinformatics*, *32*(17), 2713–2715. https://doi.org/10.1093/bioinformatics/btw187

[41] Barzel, B., Sharma, A., & Barabási, A. L. (2013). Graph theory properties of cellular networks. In A. J. M. Walhout, M. Vidal & J. Dekker (Eds.), *Handbook of systems biology* (pp. 177–193). Elsevier. https://doi.org/10.1016/B978-0-12-385944-0.00009-5

[42] Fulcoli, F. G., Franzese, M., Liu, X., Zhang, Z., Angelini, C., & Baldini, A. (2016). Rebalancing gene haploinsufficiency *in vivo* by targeting chromatin. *Nature Communications*, *7*, 11688. https://doi.org/10.1038/ncomms11688

[43] Chakraborty, D., Bernal, A. J., Schoch, K., Howard, T. D., Ip, E. H., Hooper, S. R., ..., & Shashi, V. (2012). Dysregulation of DGCR6 and DGCR6L: Psychopathological outcomes in chromosome 22q11.2 deletion syndrome. *Translational Psychiatry*, *2*(4), e105. https://doi.org/10.1038/tp.2012.31

[44] Scambler, P. J. (2000). The 22q11 deletion syndromes. *Human Molecular Genetics*, *9*(16), 2421–2426. https://doi.org/10.1093/hmg/9.16.2421

[45] Bonacci, T., Audebert, S., Camoin, L., Baudelet, E., Iovanna, J. L., & Soubeyran, P. (2017). Regulation of NUB1 activity through non-proteolytic Mdm2-mediated ubiquitination. *PLOS ONE*, *12*(1), e0169988. https://doi.org/10.1371/journal.pone.0169988

[46] Mameesh, M., Ganesh, A., Harikrishna, B., Al Zuhaibi, S., Scott, P., Al Kalbani, S., & Al Thihli, K. (2017). Co-inheritance of the membrane frizzled-related protein ocular phenotype and glycogen storage disease type Ib. *Ophthalmic Genetics*, *38*(6), 544–548. https://doi.org/10.1080/13816810.2017.1323340

[47] Liu, Z., Vong, Q. P., Liu, C., & Zheng, Y. (2014). Borg5 is required for angiogenesis by regulating persistent directional migration of the cardiac microvascular endothelial cells. *Molecular Biology of the Cell*, *25*(6), 841–851. https://doi.org/10.1091/mbc.e13-09-0543

[48] Perdomo-Sabogal, A., Nowick, K., Piccini, I., Sudbrak, R., Lehrach, H., Yaspo, M. L., ..., & Querfurth, R. (2016). Human lineage-specific transcriptional regulation through GA-binding protein transcription factor alpha (GABPa). *Molecular Biology and Evolution*, *33*(5), 1231–1244. https://doi.org/10.1093/molbev/msw007

[49] Gorman, K. M., Meyer, E., Grozeva, D., Spinelli, E., McTague, A., Sanchis-Juan, A., ..., & Kurian, M. A. (2019). Bi-allelic loss-of-function CACNA1B mutations in progressive epilepsy-dyskinesia. *The American Journal of Human Genetics*, *104*(5), 948–956. https://doi.org/10.1016/j.ajhg.2019.03.005

[50] Nakagawasai, O., Onogi, H., Mitazaki, S., Sato, A., Watanabe, K., Saito, H., ..., & Tadano, T. (2010). Behavioral and neurochemical characterization of mice deficient in the N-type Ca^{2+} channel α_{1B} subunit. *Behavioural Brain Research*, *208*(1), 224–230. https://doi.org/10.1016/j.bbr.2009.11.042

[51] Meshcheryakova, T. I., Zinchenko, R. A., Vasilyeva, T. A., Marakhonov, A. V., Zhylina, S. S., Petrova, N. V., ..., & Mutovin, G. R. (2015). A clinical and molecular analysis of branchio-oculo-facial syndrome patients in Russia revealed new mutations in TFAP2A. *Annals of Human Genetics*, *79*(2), 148–152. https://doi.org/10.1111/ahg.12098

[52] Bastaki, F., Nair, P., Mohamed, M., Malik, E. M., Helmi, M., Al-Ali, M. T., & Hamzeh, A. R. (2017). Identification of a novel CTCF mutation responsible for syndromic intellectual disability – A case report. *BMC Medical Genetics*, *18*(1), 68. https://doi.org/10.1186/s12881-017-0429-0

[53] Chen, F., Yuan, H., Wu, W., Chen, S., Yang, Q., Wang, J., ..., & Shen, Y. (2019). Three additional de novo CTCF mutations in Chinese patients help to define an emerging neurodevelopmental disorder. *American Journal of Medical Genetics Part C: Seminars in Medical Genetics*, *181*(2), 218–225. https://doi.org/10.1002/ajmg.c.31698

[54] McGill, B. E., Barve, R. A., Maloney, S. E., Strickland, A., Rensing, N., Wang, P. L., ..., & Milbrandt, J. (2018). Abnormal microglia and enhanced inflammation-related gene transcription in mice with conditional deletion of Ctcf in Camk2a-Cre-expressing neurons. *Journal of Neuroscience*, *38*(1), 200–219. https://doi.org/10.1523/JNEUROSCI.0936-17.2017

[55] Vissers, L. E. L. M., de Ligt, J., Gilissen, C., Janssen, I., Steehouwer, M., de Vries, P., ..., & Veltman, J. A. (2010). A *de novo* paradigm for mental retardation. *Nature Genetics*, *42*(12), 1109–1112. https://doi.org/10.1038/ng.712

[56] Abd Alla, J., Graemer, M., Fu, X., & Quitterer, U. (2016). Inhibition of G-protein-coupled receptor kinase 2 prevents the dysfunctional cardiac substrate metabolism in fatty acid synthase transgenic mice. *Journal of Biological Chemistry*, *291*(6), 2583–2600. https://doi.org/10.1074/jbc.M115.702688

[57] Cao, R., Dong, Y., & Kural, K. C. (2020). Integrating literature-based knowledge database and expression data to explore molecular pathways connecting PPARG and myocardial infarction. *PPAR Research*, *2020*(1), 1892375. https://doi.org/10.1155/2020/1892375

[58] Wang, N., Yang, G., Jia, Z., Zhang, H., Aoyagi, T., Soodvilai, S., ..., & Yang, T. (2008). Vascular PPARγ controls circadian variation in blood pressure and heart rate through Bmal1. *Cell Metabolism*, *8*(6), 482–491. https://doi.org/10.1016/j.cmet.2008.10.009

[59] Davis, L. K., Meyer, K. J., Rudd, D. S., Librant, A. L., Epping, E. A., Sheffield, V. C., & Wassink, T. H. (2008). Pax6 3′ deletion results in aniridia, autism and mental retardation. *Human Genetics*, *123*(4), 371–378. https://doi.org/10.1007/s00439-008-0484-x

[60] St-Onge, L., Sosa-Pineda, B., Chowdhury, K., Mansouri, A., & Gruss, P. (1997). Pax6 is required for differentiation of glucagon-producing α-cells in mouse pancreas. *Nature*, *387*(6631), 406–409. https://doi.org/10.1038/387406a0

[61] Brachet, C., Kozhemyakina, E. A., Boros, E., Heinrichs, C., Balikova, I., Soblet, J., ..., & Mathers, P. H. (2019). Truncating RAX mutations: Anophthalmia, hypopituitarism,

diabetes insipidus, and cleft palate in mice and men. *The Journal of Clinical Endocrinology & Metabolism*, *104*(7), 2925–2930. https://doi.org/10.1210/jc.2018-02316

[62] Izu, A., Yanagida, H., Sugimoto, K., Fujita, S., Sakata, N., Wada, N., . . . , & Takemura, T. (2011). Pathogenesis of focal segmental glomerular sclerosis in a girl with the partial deletion of chromosome 6p. *The Tohoku Journal of Experimental Medicine*, *223*(3), 187–192. https://doi.org/10.1620/tjem.223.187

[63] Ghoumid, J., Stichelbout, M., Jourdain, A. S., Frenois, F., Lejeune-Dumoulin, S., Alex-Cordier, M. P., . . . , & Manouvrier-Hanu, S. (2017). Blepharocheilodontic syndrome is a CDH1 pathway–related disorder due to mutations in CDH1 and CTNND1. *Genetics in Medicine*, *19*(9), 1013–1021. https://doi.org/10.1038/gim.2017.11

[64] Li, C. Y., Cha, W., Luder, H. U., Charles, R. P., McMahon, M., Mitsiadis, T. A., & Klein, O. D. (2012). E-cadherin regulates the behavior and fate of epithelial stem cells and their progeny in the mouse incisor. *Developmental Biology*, *366*(2), 357–366. https://doi.org/10.1016/j.ydbio.2012.03.012

[65] Smith, A. L., Dohn, M. R., Brown, M. V., & Reynolds, A. B. (2012). Association of Rho-associated protein kinase 1 with E-cadherin complexes is mediated by p120-catenin. *Molecular Biology of the Cell*, *23*(1), 99–110. https://doi.org/10.1091/mbc.E11-06-0497

[66] Liu, Y., Zhu, H., Zhang, X., Hu, T., Zhang, Z., Wang, J., . . . , & Sun, H. (2018). Infertility in a man with oligoasthenozoospermia associated with mosaic chromosome 22q11 deletion. *Molecular Genetics & Genomic Medicine*, *6*(6), 1249–1254. https://doi.org/10.1002/mgg3.487

[67] Özcan, A., & Şahin, Y. (2017). DiGeorge syndrome associated with azoospermia: First case in the literature. *Turkish Journal of Urology*, *43*(3), 390–392. https://doi.org/10.5152/tud.2017.08555

[68] Mikelsaar, R., Lissitsina, J., & Bartsch, O. (2011). Small supernumerary marker chromosome (sSMC) derived from chromosome 22 in an infertile man with hypogonadotropic hypogonadism. *Journal of Applied Genetics*, *52*(3), 331–334. https://doi.org/10.1007/s13353-011-0041-5

[69] Zhang, H., Su, D., Yang, Y., Zhang, W., Liu, Y., Bai, G., . . . , & Zhang, S. (2010). Some single-nucleotide polymorphisms of the TSSK2 gene may be associated with human spermatogenesis impairment. *Journal of Andrology*, *31*(4), 388–392. https://doi.org/10.2164/jandrol.109.008466

[70] Hawkinson, J. E., Sinville, R., Mudaliar, D., Shetty, J., Ward, T., Herr, J. C., & Georg, G. I. (2017). Potent pyrimidine and pyrrolopyrimidine inhibitors of testis-specific serine/threonine kinase 2 (TSSK2). *ChemMedChem*, *12*(22), 1857–1865. https://doi.org/10.1002/cmdc.201700503

[71] Shetty, J., Sinville, R., Shumilin, I. A., Minor, W., Zhang, J., Hawkinson, J. E., . . . , & Herr, J. C. (2016). Recombinant production of enzymatically active male contraceptive drug target hTSSK2-localization of the TSKS domain phosphorylated by TSSK2. *Protein Expression and Purification*, *121*, 88–96. https://doi.org/10.1016/j.pep.2016.01.009

How to Cite: de Oliveira Alves, C. C., Wolf, I. R., Gamba, B. F., Bicudo, L. A. R., & Valente, G. T. (2024). Exploring Disrupted Gene Networks in Human 22q11.2 Microdeletion. *Medinformatics*, *1*(3), 112–121. https://doi.org/10.47852/bonviewMEDIN42022652