Research Article

Integrated analysis of the critical region 5p15.3–p15.2 associated with cri-du-chat syndrome

Thiago Corrêa[1], Bruno César Feltes[2] and Mariluce Riegel[1,3*] iD

[1]*Post-Graduate Program in Genetics and Molecular Biology, Universidade Federal do Rio Grande do Sul, Porto Alegre, RS, Brazil.*
[2]*Institute of Informatics, Universidade Federal do Rio Grande do Sul, Porto Alegre, RS, Brazil.*
[3]*Medical Genetics Service, Hospital de Clínicas de Porto Alegre, Porto Alegre, RS, Brazil.*

Abstract

Cri-du-chat syndrome (CdCs) is one of the most common contiguous gene syndromes, with an incidence of 1:15,000 to 1:50,000 live births. To better understand the etiology of CdCs at the molecular level, we investigated theprotein–protein interaction (PPI) network within the critical chromosomal region 5p15.3–p15.2 associated with CdCs using systemsbiology. Data were extracted from cytogenomic findings from patients with CdCs. Based on clinical findings, molecular characterization of chromosomal rearrangements, and systems biology data, we explored possible genotype–phenotype correlations involving biological processes connected with CdCs candidate genes. We identified biological processes involving genes previously found to be associated with CdCs, such as *TERT*, *SLC6A3,* and *CTDNND2*, as well as novel candidate proteins with potential contributions to CdCs phenotypes, including CCT5, TPPP, MED10, ADCY2, MTRR, CEP72, NDUFS6, and MRPL36. Although further functional analyses of these proteins are required, we identified candidate proteins for the development of new multi-target genetic editing tools to study CdCs. Further research may confirm those that are directly involved in the development of CdCs phenotypes and improve our understanding of CdCs-associated molecular mechanisms.

Keywords: Cri-du-Chat Syndrome, 5p– cytogenomics, integrative Analysis, PPI, systems biology.

Received: June 13, 2018; Accepted: July 29, 2018.

Introduction

Cri-du-chat syndrome (CdCs, OMIM 123450) is one of the most common contiguous gene syndromes, with an incidence of 1:15,000 to 1:50,000 live births (Niebuhr, 1978; Duarte *et al.*, 2004). Although 5p deletion is clinically and genetically well described, the phenotypic variability observed among patients with the deletion suggests that additional modifying factors, including genetic and environmental factors, may impact patients' clinical manifestations (Nguyen *et al.*, 2015). The classic phenotype of CdCs encompasses a cat-like cry, facial dysmorphism, microcephaly, psychomotor delays, and intellectual disability (Overhauser *et al.*, 1994). However, the clinical spectrum and severity of the disease depend of the size of the deleted chromosomal region (Smith *et al.*, 2010). Around 80% of individuals with CdCs exhibit *de novo* terminal deletions, and 5% exhibit interstitial deletions, where the deletion is most commonly inherited (Mainardi, 2006). In this sense, approximately 10–15% of the deletions result from an un-

balanced parental translocation (Mainardi, 2006), whereas complex genomic rearrangements, such as mosaicism, *de novo* translocation, or ring chromosomes, comprise fewer than 10% of cases (Perfumo *et al.*, 2000).

Previous studies looking for phenotype–genotype correlations through determination of deleted regions on 5p have described critical regions related to increased susceptibility for cat-like cry, speech delay, facial dimorphism, and intellectual disability (Overhauser *et al.*, 1994; Church *et al.*, 1997; Marinescu *et al.*, 1999; Mainardi *et al.*, 2001; Zhang *et al.*, 2005; Elmakky *et al.*, 2014). Although studies differ in the actual contribution of these critical regions to a particular phenotype, they allow that refinement of genes under hemizygous conditions may contribute to the pathogenesis of CdCs (Mainardi, 2006; Damasceno *et al.*, 2016). Candidate genes, such as *TERT*, *MARCH6*, *CTNND2*, and *SLC6A3,* are considered dose-sensitive or conditionally haploinsufficient (i.e., a single copy of these genes is insufficient to ensure normal functioning in individuals with CdCs) (Nguyen *et al.*, 2015). Haploinsufficiency of the genes mentioned above has been implicated in telomere maintenance dysfunction, cat-like cry, intellectual disability, and attention-deficit/hyperactivity disorder, respective-

Send correspondence to Mariluce Riegel. Serviço de Genética Médica, Hospital de Clínicas de Porto Alegre, Rua Ramiro Barcelos 2350, 90035-903 Porto Alegre, RS, Brazil. E-mail: mriegel@hcpa.edu.br.

ly (Wu *et al.*, 2005; Du *et al.*, 2007; Hofmeister *et al.*, 2015; Tong *et al.*, 2015).

Even with the increasing resolution of cytogenetic techniques and the large amount of information available in databases, the investigation of contiguous gene syndromes remains a challenge. Studies have attempted to characterize genomic rearrangements and establish genotype–phenotype correlations through the identification of critical regions of susceptibility to CdCs, candidate genes, and haploinsufficiency-related altered mechanisms implicated in CdCs phenotypes (Lupski and Stankiewicz, 2005; Nguyen *et al.*, 2015). Therefore, in this study, to better understand the etiology of CdCs at the molecular level, we applied an integrative approach that combines conventional cytogenetic techniques, chromosomal microarray analysis (CMA), and systems biology tools to elucidate the probable molecular mechanisms underlying the clinical conditions present in CdCs.

Subjects and Methods

Study design and sample selection

This is a retrospective cytogenomic integrative analysis involving results of a series of cases. Clinical and cytogenomic data were extracted from six patients with CdCs enrolled in the Brazilian Network of Reference and Information in Microdeletion Syndromes (RedeBRIM) project (Riegel *et al.*, 2014, 2017; De Souza *et al.*, 2015; Dorfman *et al.*, 2015). The patients were regularly reevaluated over several years. Psychomotor development assessments were based on personal observations, school performance, and parent information. Daily abilities and skills, such as language, social interactions, concentration/attention, impulsiveness, motor control, perception, and learning and memory were recorded and published by our group elsewhere (Damasceno *et al.*, 2016). The five most frequent groups of clinical findings were selected and registered in the present study. This study has been approved by the Ethics Research Committee of Hospital de Clínicas de Porto Alegre (HCPA), followed the Declaration of Helsinki, and the standards established by the author's Institutional Review Board.

Cytogenomic Small Region of Overlap (SRO)

The deletions were mapped by whole genome array-CGH using a 60-mer oligonucleotide-based microarray with a theoretical resolution of 40 kb (8 60K, Agilent Technologies Inc., Santa Clara, CA). Labeling and hybridization were performed following the protocols provided by Agilent 2011. The arrays were analyzed using a microarray scanner (G2600D) and Feature Extraction software (version 9.5.1) (both from Agilent Technologies). Image analyses were performed using Agilent GenomicWorkbench Lite Edition 6.5.0.18 with the statistical algorithm ADM-2 at a sensitivity threshold of 6.0. The detailed cytogenomic

profiles of the patients analyzed in this study were presented by our group elsewhere (Damasceno *et al.*, 2016). Based on it, the chromosomal SRO was determined.

Network design

The protein–protein interaction (PPI) metasearch engine STRING 10.0 (http://string-db.org/) was used to create PPI networks based on genes located in the SRO. The list of genes was obtained from the human assembly of February 2009 (GRCh37/hg19) (Kent *et al.*, 1976; von Mering *et al.*, 2005). The parameters used in STRING were: (i) degree of confidence, 0.400, with 1.0 being the highest level of confidence; (ii) 500 proteins in the 1st and 2nd shell; and (iii) all prediction methods enabled, except for text mining and gene fusion. The final PPI network obtained through STRING was analyzed using Cytoscape 3.5 (Shannon *et al.*, 2003). Non-connected nodes from the networks were not included.

Clustering and GO analysis

The MCODE tool was used to identify densely connected regions in the final Cytoscape network. The analysis was based on vertex weighting by the local neighborhood density and outward traversal from a locally dense seed protein to isolate the highly clustered regions (Bader and Hogue 2003). The PPI modules generated by MCODE were further studied by focusing on major biology-associated processes using the Biological Network Gene Ontology (BiNGO) 3.0.3 Cytoscape plugin (Maere *et al.*, 2005). The degree of functional enrichment for a given cluster and category was quantitatively assessed *(p*-value) using a hypergeometric distribution. Multiple test correction was also implemented by applying the false discovery rate (FDR) algorithm (Benjamini and Hochberg 1995) at a significance level of $p < 0.05$.

Centralities

Two major parameters of network centralities (node degree and betweenness) were used to identify H-B nodes from the PPI network using the Cytoscape plugin CentiScaPe 3.2.1 (Scardoni *et al.*, 2009). The node degree centrality indicates the total number of adjacent nodes that are connected to a unique node. Nodes with a high node degree are called hubs and have central functions in a biological network (Scardoni *et al.*, 2009). Furthermore, we also analyzed the betweenness score, which corresponds to the number of shortest paths between two nodes that pass through a node of interest. Thus, nodes with high betweenness scores, compared to the average betweenness score of the network, are responsible for controlling the flow of information through the network topology (Newman, 2006; Scardoni *et al.*, 2009). These nodes are called bottlenecks and are normally related to the control of information between groups of proteins (Scardoni *et al.*, 2009).

Availability of data and materials

All data generated or analyzed during this study are included in this published article and its supplementary information files (Tables S1 - S17).

Results

The main clinical findings of six patients with CdCs selected to this study are presented in Figure 1. Intellectual disability (6/6 patients), learning difficulties (6/6 patients), multiple congenital abnormalities (6/6 patients), hyperactivity/impulsiveness (5/6 patients), and heart defects (4/6 patients) were the most frequent findings (Figure 1). Among the samples, three were from males, with ages ranging from 6 to 38 years, and three were from females, with ages ranging from 7 to 20 years.

Cytogenomic data analysis MR

Six *de novo* terminal deletions that ranged in size from approximately 11.2 Mb to 28.6 Mb, with breakpoints from 5p15.2 to 5p13 were mapped. The analysis of CMA profile data revealed a small region of overlap (SRO) of 10.8 Mb encompassing the bands 5p15.33–p15.2. The approximate genomic position of the SRO is chr5:527552–11411700, comprising 44 genes according to the UCSC genome browser human assembly of February 2009 (GRCh38/hg19) (Figure 2).

Networks and topological analysis

Overall, the scale-free network was composed of 2284 nodes (proteins) and 83340 edges (interactions) (Figure 3). Centrality analyses were carried out to identify hub-bottlenecks (H-B), the most topologically relevant no-
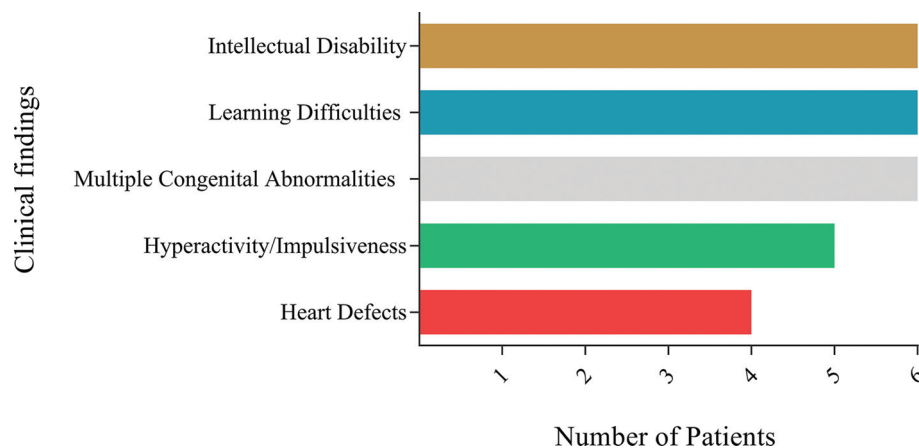

Figure 1 - Summary of clinical findings of the six individuals in the study according to Damasceno *et al.* (2016).
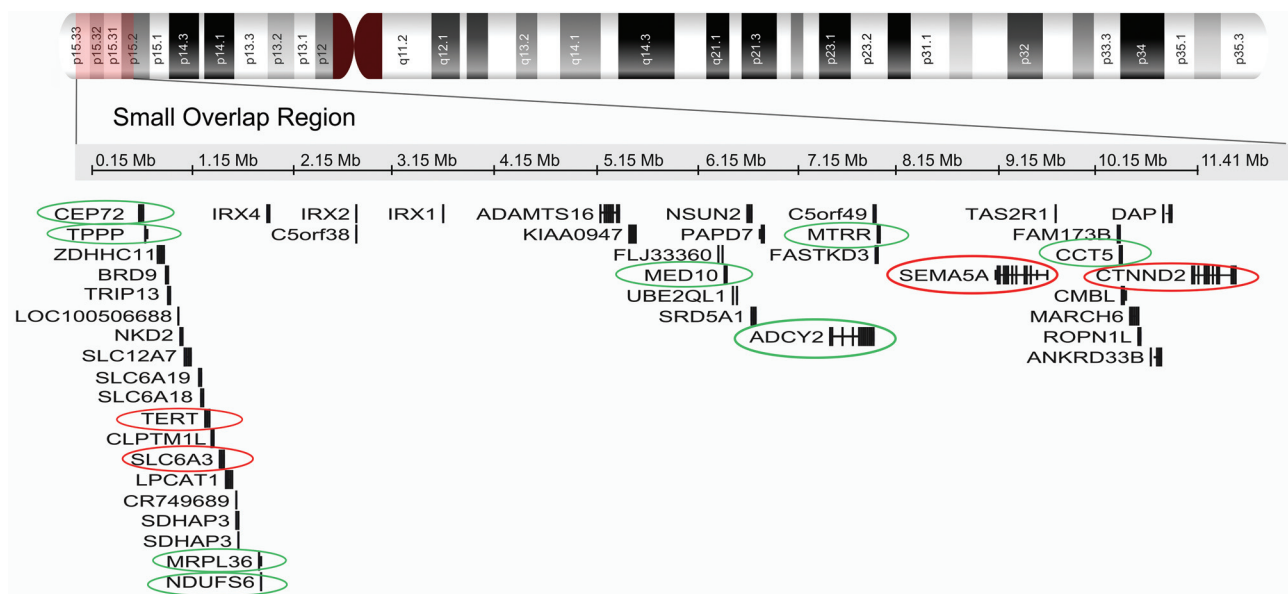

Figure 2 - Cytogenomic profile of chromosome 5. Chromosomal critical region of 5p15.33–p15.2. Genes localized to the critical region were obtained from the human assembly of February 2009 (GRCh37/hg19). Red circles show genes already associated with CdCs. Green circles show candidate genes from this study for contributing to the phenotype in CdCs.
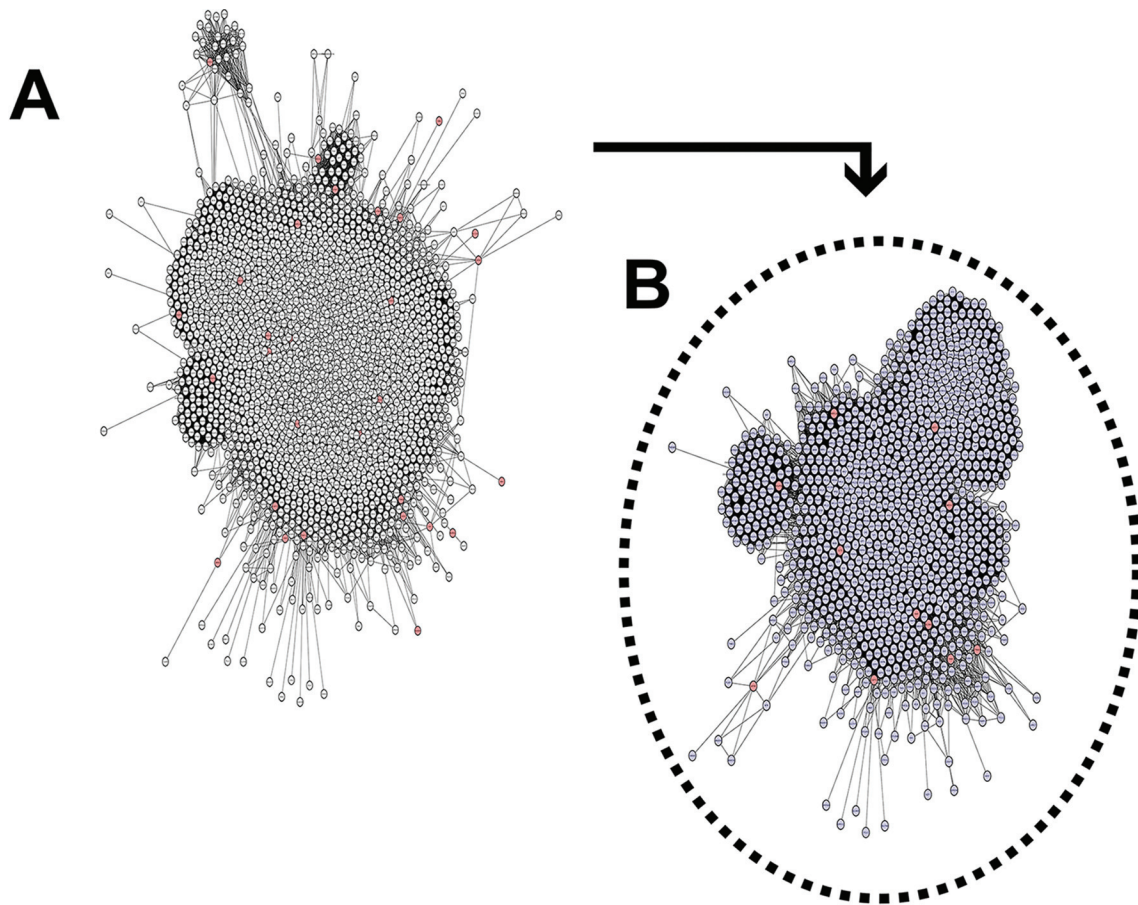
Figure 3 - The PPI network. The list of 44 genes was obtained from the human assembly of February 2009 (GRCh37/hg19). Interaction data from STRING were used to construct networks using Cytoscape software. (A) The primary network is composed of 2284 nodes and 83,340 edges. Red nodes are target proteins (SLC6A3, SRD5A1, CCT5, ADCY2, TAS2R1, MED10, MTRR, SLC12A7, CEP72, NDUFS6, MARCH6, LPCAT1, NKD2, CTNND2, TERT, CLPTM1L, MRPL18, MRPL36, UBE2QL1, PAPD7, and TPPP). (B) Secondary network composed of 1062 nodes and 41,309 edges. Red nodes are candidate proteins (CCT5, TPPP, MED10, ADCY2, MTRR, CEP72, NDUFS6, MRPL36, CTNND2, TERT, and SLC6A3) and immediate neighbors from SRO.

des. The network hubs (nodes with an above average number of connections) and betweenness (total number of non-redundant shortest paths going through a node or edge) indicate the most critical points in a biological network (Yu *et al.*, 2007). In our analysis, we observed 273 H-B nodes in the SRO network. Furthermore, we performed a cluster analysis that identified 16 major cluster regions above our cutoff score, and gene ontology (GO) analyses were performed in the identified modules.

Clusters taken into consideration for further analysis were those containing major proteins related to CdCs and deleted in all patients according to Espitiro Santo (2016), namely those containing combinations of SLC6A3, SRD5A1, CCT5, ADCY2, TAS2R1, MED10, MTRR, SLC12A7, CEP72, NDUFS6, MARCH6, LPCAT1, NKD2, CTNND2, TERT, CLPTM1L, MRPL18, MRPL36, UBE2QL1, PAPD7, and TPPP (Figure 4). In addition, the TERT protein was a commonly clusterized protein, and all clusters containing TERT were selected. Clusters that did not contain multiple combinations of the CdCs protein tar-

gets mentioned above, TERT, were excluded from the final analysis.

The most relevant GO terms are listed in Table S1. The main observed terms were: (i) nervous system-associated processes, such as development, synapsis, and learning; (ii) aging; (iii) double-strand break repair; (iv) regulation of apoptosis/cell death; (v) telomere maintenance; (vi) senescence; (vii) response to cytokine stimulus; (viii) regulation of interleukin (IL)-1; (ix) hormone biosynthetic processes, especially androgen biosynthesis; and (x) regulation of the NF-κB/IκB pathway. The relative number of GO terms associated with each cluster can be found in Figure 5. Our analysis excluded GO terms that were not associated with significant biological processes related to the disease, or that were too general (e.g., regulation of biological process, signaling process, or response to endogenous stimulus).
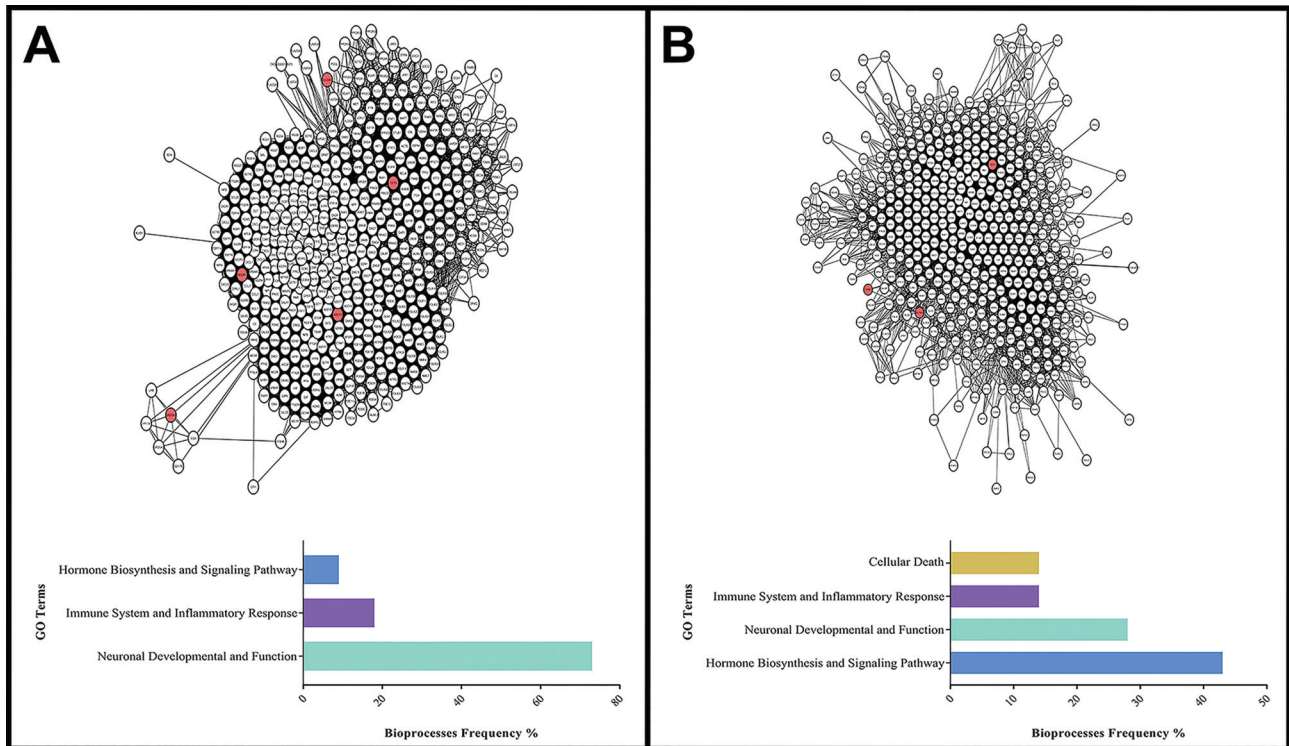
Figure 4 - Subnetworks derived from clustering analysis. Red nodes are target proteins. (A) Cluster 1, with Ci = 94,369, composed of 509 nodes and 24,064 edges. Target proteins: SLC6A3, SRD5A1, CCT5, ADCY2, and TAS2R1. Below, summary of the bioprocess frequency identified in the PPI network (B) Cluster 8, Ci = 23,208, contains 471 nodes and 5477 edges. Target proteins: SLC6A3, TERT, and TPPP. Below, summary of the Bioprocess frequency identified in the PPI network.
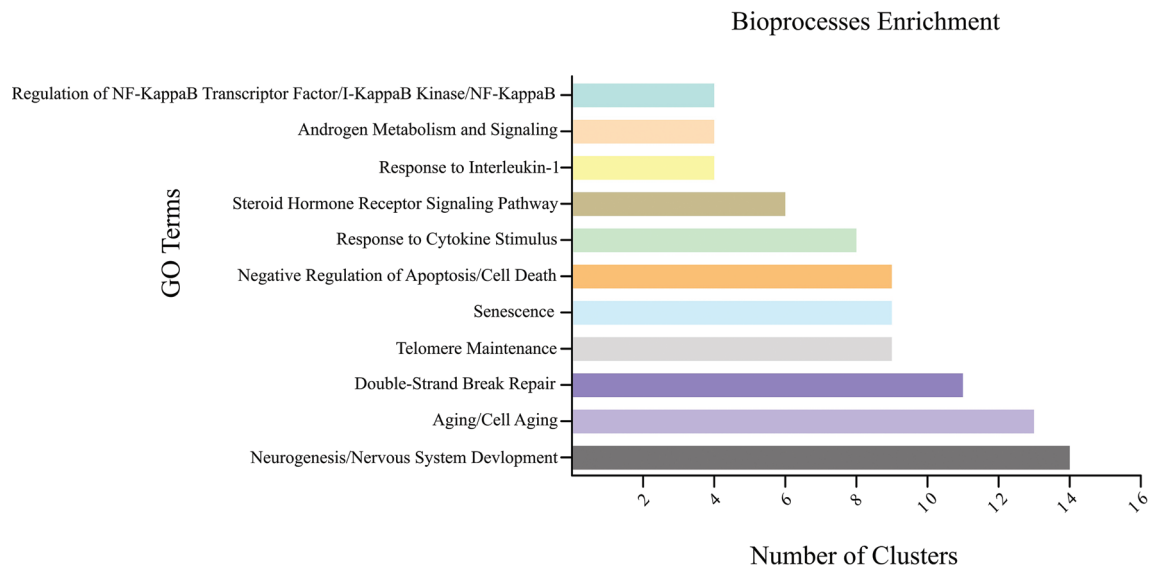

Figure 5 - Summary of the bioprocess enrichment identified in the PPI network. The colored horizontal bars show GO terms frequently present in the analyzed clusters.

Discussion

CdCs patients are traditionally diagnosed based on a detailed clinical evaluation and cytogenetic investigations. Furthermore, some studies have shown the importance of characterizing the genomic position of the critical chromo-somal region associated with CdCs for a better understanding of genotype–phenotype correlations (Wu *et al.*, 2005; Zhang *et al.*, 2005; Damasceno *et al.*, 2016). Network-based approaches may contribute to the identification of specific genes distributions in a given disease and reveal common molecular mechanisms among genes affected by the condition. Furthermore, genes associated with the same

illness have been observed to interact with each other more frequently than expected by chance (Barabási *et al.*, 2011).

Interaction between SLC6A3 TPPP and CCT5 and Processes related to neuronal development and function in CdCs

In this study, the constructed networks and topological analysis, such as those in clusters 1 and 8 (Figures 2 and 4), showed interactions between SLC6A3, TPPP, and CCT5, genes which are located in the SRO, and interactions between processes related to neuronal development and function in CdCs. The GO analysis of clusters 1 and 8 indicated the presence of proteins deleted in hemizygous individuals in our study that are related to the regulation of glutamatergic and dopaminergic synaptic transmission, catecholamine uptake involved in synaptic transmission, and norepinephrine secretion and neurogenesis. Changes in patterns of neuronal activity modulated by dopamine and noradrenaline in the cortico-striatal region of the brain are able to influence the emergence of disturbances, such as attention deficit hyperactivity disorder (ADHD) (Del Campo *et al.*, 2011; Cummins *et al.*, 2012). Interestingly, ADHD is present in about 70% of children with CdCs (Nguyen *et al.*, 2015), and, in our study, hyperactivity was present in five out of the six subjects (Figure 1). SLC6A3, a dopamine transporter, regulates extracellular dopamine, is responsible for the reuptake of dopamine, and functions to balance levels of neuronal dopamine (Gizer *et al.*, 2009). Deficiency of this protein can lead to the accumulation of dopamine in the cytosol, with deleterious effects (Sotnikova *et al.*, 2005). These effects may be associated with hyperlocomotion, stereotyped behaviors, and hyperactivity, as in *Slc6a3* KO mice (Giros *et al.*, 1996; Pogorelov *et al.*, 2005; Lohr *et al.*, 2017), or decreased immobility, as in *Slc6a3*$^{+/-}$ mice (Perona *et al.*, 2008). Therefore, SLC6A3 can be proposed as a good target on subsequent functional analyses that could increase the mechanistic knowledge related to those CdCs phenotypes. Interestingly, we observed that TPPP is among the direct neighbors of SLC6A3 in cluster 8 (Figure 4). TPPP functions in tubulin polymerization and microtubule stabilization (Vincze *et al.*, 2006). TPPP plays an important role in pathological conditions through the co-enrichment and co-localization of TPPP and α-synuclein in human brain inclusions, such as in Parkinson's disease (Oláh and Ovádi, 2014). Through the polymerization of the tubulin polymer, TPPP contributes to the extension of peripheral axons in sensory neurons (Aoki *et al.*, 2014). Changes in the expression of TPPP are associated with the phenotypes of depression and anxiety following early life stress in humans (Montalvo-Ortiz *et al.*, 2016). Therefore, these results identified by network analysis suggest an important perturbation between the proteins SLC6A3 and TPPP generating neural changes in CdCs individuals. SLC6A3 also interacts with the H-B CCT5 in cluster 1, in which processes related to cognition, memory, and learning can be found (Figure 4, Table S2). The protein CCT5 is involved in cilia morphogenesis and neurodegenerative processes, and its deficiency may cause neurodegenerative diseases, such asspastic paraplegia (Bouhouche *et al.*, 2006; Posokhova *et al.*, 2011), supporting the GO results. Individuals with spastic paraplegia may present with atrophy of the spinal cord and defects in the upper limbs. These results indicate that SLC6A3, CCT5 and TPPP show important connection. Thus, we could consider that disruption of these interactions may change the processes related to neuronal development and function underlying in some patients with CdCs.

Interplay between of genes in the SRO and behavioral and cognitive impairment

The proteins encoded by *CTNND2*, *TERT*, and *MED10*, which are located in the SRO determined in this study (Figure 2), are commonly deleted in CdCs and interact in several modules associated with neuronal development/function and cellular death, specifically clusters 3, 5, 6, 8, 10, and 11 (Tables S4, S5, S6, S7, S9, S11 and S12). This suggests an interplay between genes in the SRO and behavioral and cognitive impairment. These genes are expressed during important periods of embryonic and neuronal development (Yui *et al.*, 1998; Kwon *et al.*, 1999; Ho *et al.*, 2000). *CTNND2*, considered a bottleneck in our analysis, encodes δ-catenin, a component of adherens junction complexes (Kosik *et al.*, 2005) that regulates spine morphogenesis and synapse function in hippocampal neural cells during development (Arikkath *et al.*, 2009). δ-Catenin is stabilized by N-cadherin, which binds to PDZ domain proteins in the post-synaptic compartment at synapse junctions and regulates spine architecture during hippocampal development and the differentiation of neurons via downstream effectors that bind to actin in the cytoskeleton (Kosik *et al.*, 2005; Yuan *et al.*, 2015). Among the bioprocesses investigated in the protein interaction network, we identified the negative regulation of the Wnt receptor signaling pathway. Through Wnt signaling, δ-catenin prevents Rho GTPase signaling, modulating the Ras superfamily in cytoskeletal reorganization (Lu *et al.*, 2016). Perturbations in this pathway, observed after depletion of δ-catenin, may contribute to functional neurological alterations (Arikkath *et al.*, 2009). In this sense, the loss of a copy of *CTNND2* in CdCs may be associated with intellectual disability, reading problems (Medina *et al.*, 2000; Belcaro *et al.*, 2015; Hofmeister *et al.*, 2015), learning difficulties, and autism spectrum disorder (ASD) (Asadollahi *et al.*, 2014) (Figure 1). The interplay of δ-catenin with cadherin suggests its influence on Wnt/β-catenin signaling (Lu *et al.*, 2016), which increases keloid cell proliferation and inhibits apoptosis through its interaction with telomerase (Yu *et al.*, 2016). This mechanism perhaps explains the enrichment of the negative regulation of apoptosis process in the GO analysis

(Figure 5). In addition, reduction in MED10 levels enhances Wnt signaling and is required for the expression of developmentally regulated genes (Kwon *et al.*, 1999; Lin *et al.*, 2007). The H-B MED10 is crucial for DNA-binding factors that activate transcription via RNA polymerase II (Sato *et al.*, 2003). Lastly, the telomerase reverse transcriptase, encoded by *TERT*, which behaved as an H-B, was the most clusterized protein (Tables S15 and S17). The hemizygosity of *TERT* has been associated with shorter telomeres in lymphocytes from CdCs patients and contributes to the phenotypic changes seen in the syndrome (Zhang *et al.*, 2003). However, another study with 52 individuals affected by CdCs showed that the telomere length in CdCs patients was within the normal range, though the average was shorter than that in normal controls (Du *et al.*, 2007). These data suggest that the contribution of *TERT* to CdCs may involve alterations in other biological processes or pathways. For instance, *TERT* can exert protective effects. Under dietary restriction conditions, TERT accumulates in the mouse brain, leading to reductions in free radicals in the mitochondria, DNA damage, and apoptosis through the inhibition of the mTOR cascade (Miwa *et al.*, 2016). These processes were present in all clusters except 1 and 13 (Tables S2 and S14).

Therefore, analyses of centrality suggest that the deficiency in *CTNND2*, *TERT*, and *MED10* genes expression during important stages of development may affect processes related to neurogenesis and the regulation of apoptosis and DNA repair, being inherent in the cognitive and behavioral impairments seen in CdC patients (Figure 1).

Control of NF-kB transcription factor/interleukin 1 and inflammatory response

In several clusters, GO analysis identified processes related to the immune system and inflammatory response. Considering this, we explored the control of the NF-κB transcription factor/IL-1 and the inflammatory response. The appearance of respiratory and intestinal infections during the first years of life is common in patients with CdCs, though it has been rarely discussed (Mainardi, 2006). Processes related to immune response-activating signal transduction, response to IL-1, leukocyte activation, and regulation of the IκB kinase/NF-κB cascade, which has an important role in inflammation (Deacon and Knox 2018), were observed in our study, especially in clusters 1, 2, 3, 5, 6, 7, 8, 9, 10, 11, and 12 (Figure 5, Tables S2-S4 and S6-S13). With the use of telomerase inhibitors and telomerase-targeting small interfering RNAs, it has been found that H-B TERT reduces TNF-α-induced chemokine expression in airway smooth muscle cells (SMCs) (Deacon and Knox, 2018). Another protein involved in the immune response is adenylyl cyclase (ADCY2), which is also an H-B according to the centrality analysis. This protein catalyzes the formation of cyclic adenosine monophosphate (cAMP) from adenosine triphosphate (ATP), involving va-

rious signal transduction pathways. ADCY2 regulates the production of IL-6 in inflammatory processes and enhances its expression in SMCs (Bogard *et al.*, 2014; Jajodia *et al.*, 2016). In addition, single-nucleotide polymorphisms in ADCY2 have been associated with severe chronic obstructive pulmonary disease (Hardin *et al.*, 2012).

These data suggest that the presence of specific pathways related to the immune response can be affected by genes commonly deleted in CdCs (Figure 5). These results bring new insights into the pathogenesis of the syndrome, in an attempt to explain the emergence of recurrent respiratory and intestinal infections during the first years of life in individuals with CdCs (Mainardi, 2006).

Association between genes in SRO and congenital malformations.

Regarding the association between genes in the SRO and the multiple congenital malformations observed in CdCs, the network analysis demonstrated interactions between MTRR, CEP72, NDUFS6, MRPL36, and MED10 in clusters 2 and 4, in which the GO analysis identified processes related to DNA repair, cell cycle control, cellular death, and mitochondrial ATP synthesis, and electron transport (Figure 5). *MTRR* encodes a methionine synthase reductase that is fundamental for the remethylation of homocysteine, which regenerates functional methionine synthase via reductive methylation. Individuals with neural tube defects (NTDs) exhibit elevated homocysteine concentrations (Steegers-Theunissen *et al.*, 1993; Zhu *et al.*, 2003; Cheng *et al.*, 2015). The protein MTRR emerged as a bottleneck in our protein interaction network. Heterozygous mutations that lead to MTRR deficiency have been implicated in homocysteine accumulation, resulting in adverse reproductive outcomes and congenital heart defects in mice (Zhu *et al.*, 2003; Li *et al.*, 2005). Therefore, defects in the activity of MTRR could be associated with frequent clinical manifestations of CdCs, such as cardiac abnormalities. Furthermore, neurodevelopmental disorders such as primary microcephaly are associated with mutations in proteins that interact with the centrosomes, such as the CEP72 (Kodani *et al.*, 2015), which was considered an H-B in our analysis. CEP72 regulates the localization of centrosomal proteins and bipolar spindle formation (Oshimori *et al.*, 2009). Therefore, CEP72 is involved in centriole duplication and biological processes such as control of the cell cycle, and deficiency of this protein may contribute to dysmorphic phenotypes in CdCs (Figure 1).

Another protein in cluster 2 was the H-B NDUFS6, an accessory subunit of the mitochondrial chain NADH dehydrogenase (Murray *et al.*, 2003). Deletion of NDUFS6 or mutation of its Zn-binding residues blocks a late step in complex I assembly (Kmita *et al.*, 2015). Mutations in this protein may also cause lethal neonatal mitochondrial complex I deficiency (Kirby *et al.*, 2004) and fatal neonatal lactic acidemia (Spiegel *et al.*, 2009). Besides these proteins,

MRPL36, a component of the ribosomal subunit (Williams *et al.*, 2004), emerged as a hub in our network of protein interactions. Decreases in MRPL36 prevent the correct folding and assembly of translation products, leading to rapid degradation of these molecules and defects in the biogenesis of respiratory chain complexes in the mitochondria (Prestele *et al.*, 2009). Therefore, the hub MRPL36 may contribute to oxidative stress-related processes found in cluster 2 (Table S3) and may be associated with excess apoptosis and NTDs (Yang *et al.*, 2008).

Excessive apoptosis in fetal central nervous tissues can cause NTDs by decreasing the number of cells in the neural folds or by physical disruption of the dorsal midline, consequently resulting in embryonic dysmorphogenesis (Chen *et al.*, 2017; Lin *et al.*, 2018). Furthermore, the H-B MED10, located in clusters 2 (Figure 4), 3, and 4, regulates heart valve formation in zebrafish (Just *et al.*, 2016). In addition, network analysis demonstrated an interaction between MED10 and the protein encoded by *MED24/TRAP100*, located on chromosome 17. MED24 is necessary for enteric nervous system development in zebrafish (Pietsch, 2006). Together, these findings contribute to our understanding of the emergence of congenital heart defects, microcephaly, and occasional abnormalities such as agenesis of the corpus callosum, cerebral atrophy, and cerebellar hypoplasia, which may be present in CdCs.

Conclusion

The possibility of using microarrays to characterize chromosomal rearrangements has led to several studies aimed at establishing genotype-phenotype correlations in several contiguous gene deletion syndromes, and some of them have proposed the regions of susceptibility to each specific condition. However, no consensus has been reached on the exact identity of the genes and cell signaling pathways involved in promoting these symptoms, as e.g. in the CdCs. This is the first study to explore the interaction network of the proteins encoded in the critical region associated with CdCs by combining cytogenomic data and systems biology tools. This study identified and demonstrated the biological processes involving genes previously found to be associated with CdCs, such as *TERT*, *SLC6A3,* and *CTDNND2*. Furthermore, through analysis of the protein interaction network, we identified other possible candidate proteins, including CCT5, TPPP, MED10, ADCY2, MTRR, CEP72, NDUFS6, and MRPL36, with potential contributions to the phenotypes observed in CdCs. Further functional analysis of these proteins is required to fully understand their involvement and interplay in CdCs. Additional research in this direction may confirm those that are directly involved in the development of the CdCs phenotype and improve genotype–phenotype correlations.

Conflict of interests

There are no conflicts of interest to declare.

Author contributions

All authors contributed to the analysis and interpretation of data; all authors participated in the writing of the manuscript and approved the version submitted for publication.

References

Aoki M, Segawa H, Naito M and Okamoto H (2014) Identification of possible downstream genes required for the extension of peripheral axons in primary sensory neurons. Biochem Biophys Res Commun 445:357–362.

Arikkath J, Peng IF, Gie Ng Y, Israely I, Liu X, Ullian EM and Reichardt LF (2009) Delta-catenin regulates spine and synapse morphogenesis and function in hippocampal neurons during development. J Neurosci 29:5435–5442.

Asadollahi R, Oneda B, Joset P, Azzarello-Burri S, Bartholdi D, Steindl K, Vincent M, Cobilanschi J, Sticht H, Baldinger R *et al.* (2014) The clinical significance of small copy number variants in neurodevelopmental disorders. J Med Genet 51:677–688.

Bader GD and Hogue CWV (2003) An automated method for finding molecular complexes in large protein interaction networks. BMC Bioinformatics 4:2.

Barabási AL, Gulbahce N and Loscalzo J (2011) Network medicine: A network-based approach to human disease. Nat Rev Genet 12:56–68.

Belcaro C, Dipresa S, Morini G, Pecile V, Skabar A and Fabretto A (2015) CTNND2 deletion and intellectual disability. Gene 565:146–149.

Benjamini Y and Hochberg Y (1995) Controlling the False Discovery Rate: A practical and powerful approach to multiple testing. J R Stat Soc B 57:289–300.

Bogard AS, Birg AV and Ostrom RS (2014) Non-raft adenylyl cyclase 2 defines a cAMP signaling compartment that selectively regulates IL-6 expression in airway smooth muscle cells: Differential regulation of gene expression by AC isoforms. Naunyn Schmiedebergs Arch Pharmacol 387:329–339.

Bouhouche A, Benomar A, Bouslam N, Chkili T and Yahyaoui M (2006) Mutation in the epsilon subunit of the cytosolic chaperonin-containing t-complex peptide-1 (Cct5) gene causes autosomal recessive mutilating sensory neuropathy with spastic paraplegia. J Med Genet 43:441–443.

Chen Z, Wang J, Bai Y, Wang S, Yin X, Xiang J, Li X, He M, Zhang X, Wu T *et al.* (2017) The associations of TERT-CLPTM1L variants and TERT mRNA expression with the prognosis of early stage non-small cell lung cancer. Cancer Gene Ther 24:20–27.

Cheng H, Li H, Bu Z, Zhang Q, Bai B, Zhao H, Li RK, Zhang T and Xie J (2015) Functional variant in methionine synthase

reductase intron-1 is associated with pleiotropic congenital malformations. Mol Cell Biochem 407:51–56.

Church DM, Yang J, Bocian M, Shiang R and Wasmuth JJ (1997) A high-resolution physical and transcript map of the Cri du Chat region of human chromosome 5p. Genome Res 7:787–801.

Cummins TDR, Hawi Z, Hocking J, Strudwick M, Hester R, Garavan H, Wagner J, Chambers CD and Bellgrove MA (2012) Dopamine transporter genotype predicts behavioural and neural measures of response inhibition. Mol Psychiatry 17:1086–1092.

De Souza KR, Mergener R, Huber J, Campos Pellanda L and Riegel M (2015) Cytogenomic evaluation of subjects with syndromic and nonsyndromic conotruncal heart defects. Biomed Res Int 2015:401941.

Deacon K and Knox AJ (2018) PINX1 and TERT are required for TNF-α–induced airway smooth muscle chemokine gene expression. J Immunol 200:1283-1294.

Del Campo N, Chamberlain SR, Sahakian BJ and Robbins TW (2011) The roles of dopamine and noradrenaline in the pathophysiology and treatment of attention-deficit/hyperactivity disorder. Biol Psychiatry 69:e145–e157.

Dorfman LE, Leite JCL, Giugliani R and Riegel M (2015) Microarray-based comparative genomic hybridization analysis in neonates with congenital anomalies: Detection of chromosomal imbalances. J Pediatr (Rio J) 91:59–67.

Du HY, Idol R, Robledo S, Ivanovich J, An P, Londono-Vallejo A, Wilson DB, Mason PJ and Bessler M (2007) Telomerase reverse transcriptase haploinsufficiency and telomere length in individuals with 5p- syndrome. Aging Cell 6:689–697.

Duarte AC, Cunha E, Roth JM, Ferreira FLS, Garcias GL and Martino-Roth MG (2004) Cytogenetics of genetic counseling patients in Pelotas, Rio Grande do Sul, Brazil. Genet Mol Res 3:303–308.

Elmakky A, Carli D, Lugli L, Torelli P, Guidi B, Falcinelli C, Fini S, Ferrari F and Percesepe A (2014) A three-generation family with terminal microdeletion involving 5p15.33-32 due to a whole-arm 5;15 chromosomal translocation with a steady phenotype of atypical cri du chat syndrome. Eur J Med Genet 57:145–150.

Espirito Santo LD, Maria L, Moreira A and Riegel M (2016) Cri-Du-Chat Syndrome: Clinical profile and chromosomal microarray analysis in six patients. BioMed Res Int 2016:5467083.

Giros B, Jaber M, Jones SR, Wightman RM and Caron MG (1996) Hyperlocomotion and indifference to cocaine and amphetamine in mice lacking the dopamine transporter. Nature 379:606–612.

Gizer IR, Ficks C and Waldman ID (2009) Candidate gene studies of ADHD: A meta-analytic review. Hum Genet 126:51–90.

Hardin M, Zielinski J, Wan ES, Hersh CP, Castaldi PJ, Schwinder E, Hawrylkiewicz I, Sliwinski P, Cho MH and Silverman EK (2012) CHRNA3/5, IREB2, and ADCY2 are associated with severe chronic obstructive pulmonary disease in Poland. Am J Respir Cell Mol Biol 47:203–208.

Ho C, Zhou J, Medina M, Goto T, Jacobson M, Bhide PG and Kosik KS (2000) δ-Catenin is a nervous system-specific adherens junction protein which undergoes dynamic relocalization during development. J Comp Neurol 420:261–276.

Hofmeister W, Nilsson D, Topa A, Anderlid BM, Darki F, Matsson H, Tapia Páez I, Klingberg T, Samuelsson L, Wirta V *et al.* (2015) *CTNND2* —a candidate gene for reading problems and mild intellectual disability. J Med Genet 52:111–122.

Jajodia A, Kaur H, Kumari K, Kanojia N, Gupta M, Baghel R, Sood M, Jain S, Chadda RK and Kukreti R (2016) Evaluation of genetic association of neurodevelopment and neuro-immunological genes with antipsychotic treatment response in schizophrenia in Indian populations. Mol Genet Genomic Med 4:18–27.

Just S, Hirth S, Berger IM, Fishman MC and Rottbauer W (2016) The mediator complex subunit Med10 regulates heart valve formation in zebrafish by controlling Tbx2b-mediated Has2 expression and cardiac jelly formation. Biochem Biophys Res Commun 477:581–588.

Kent WJ, Sugnet CW, Furey TS and Roskin KM (1976) The Human Genome Browser at UCSC W. J Med Chem 19:1228–31.

Kirby DM, Salemi R, Sugiana C, Ohtake A, Parry L, Bell KM, Kirk EP, Boneh A, Taylor RW, Dahl HHM *et al.* (2004) NDUFS6 mutations are a novel cause of lethal neonatal mitochondrial complex I deficiency. J Clin Invest 114:837–845.

Kmita K, Wirth C, Warnau J, Guerrero-Castillo S, Hunte C, Hummer G, Kaila VRI, Zwicker K, Brandt U and Zickermann V (2015) Accessory NUMM (NDUFS6) subunit harbors a Zn-binding site and is essential for biogenesis of mitochondrial complex I. Proc Natl Acad Sci USA 112:5685–5690.

Kodani A, Yu TW, Johnson JR, Jayaraman D, Johnson TL, Al-Gazali L, Sztriha L, Partlow JN, Kim H, Krup AL *et al.* (2015) Centriolar satellites assemble centrosomal microcephaly proteins to recruit CDK2 and promote centriole duplication. ELife 4:e07519.

Kosik KS, Donahue CP, Israely I, Liu X and Ochiishi T (2005) δ-Catenin at the synaptic-adherens junction. Trends Cell Biol 15:172–178.

Kwon JY, Park JM, Gim BS, Han SJ, Lee J and Kim YJ (1999) *Caenorhabditis elegans* mediator complexes are required for developmental-specific transcriptional activation. Proc Natl Acad Sci USA 96:14990–14995.

Li D, Pickell L, Liu Y, Wu Q, Cohn JS and Rozen R (2005) Maternal methylenetetrahydrofolate reductase deficiency and low dietary folate lead to adverse reproductive outcomes and congenital heart defects in mice. Am Soc Clin Nutr 82:188–195.

Lin S, Ren A, Wang L, Huang Y, Wang Y, Wang C and Greene N (2018) Oxidative stress and apoptosis in benzo[a]pyrene-Induced neural tube defects. Free Radic Biol Med 116:149–158.

Lin X, Rinaldo L, Fazly AF and Xu X (2007) Depletion of Med10 enhances Wnt and suppresses Nodal signaling during zebrafish embryogenesis. Dev Biol 303:536–548.

Lohr KM, Masoud ST, Salahpour A and Miller GW (2017) Membrane transporters as mediators of synaptic dopamine dynamics: implications for disease. Eur J Neurosci 45:20–33.

Lu Q, Aguilar BJ, Li M, Jiang Y and Chen YH (2016) Genetic alterations of δ-catenin/NPRAP/Neurojungin (CTNND2): Functional implications in complex human diseases. Hum Genet 135:1107–1116.

Lupski JR and Stankiewicz P (2005) Genomic disorders: Molecular mechanisms for rearrangements and conveyed phenotypes. PLoS Genet 1:627–633.

Maere S, Heymans K and Kuiper M (2005) BiNGO: A Cytoscape plugin to assess overrepresentation of Gene Ontology categories in Biological Networks. Bioinformatics 21:3448–3449.

Mainardi PC (2006) Cri du Chat syndrome. Orphanet J Rare Dis 1:33

Mainardi PC, Perfumo C, Calì A, Coucourde G, Pastore G, Cavani S, Zara F, Overhauser J, Pierluigi M and Bricarelli FD (2001) Clinical and molecular characterisation of 80 patients with 5p deletion: genotype-phenotype correlation. J Med Genet 38:151–158.

Marinescu RC, Johnson EI, Dykens EM, Hodapp RM and Overhauser J (1999) No relationship between the size of the deletion and the level of developmental delay in cri-du-chat syndrome. Am J Med Genet 86:66–70.

Medina M, Marinescu RC, Overhauser J and Kosik KS (2000) Hemizygosity of δ-catenin (CTNND2) is associated with severe mental retardation in cri-du-chat syndrome. Genomics 63:157–164.

Miwa S, Czapiewski R, Wan T, Bell A, Hill KN, Von Zglinicki T and Saretzki G (2016) Decreased mTOR signalling reduces mitochondrial ROS in brain via accumulation of the telomerase protein TERT within mitochondria. Aging 8:2551–2567.

Montalvo-Ortiz JL, Bordner KA, Carlyle BC, Gelernter J, Simen AA and Kaufman J (2016) The role of genes involved in stress, neural plasticity, and brain circuitry in depressive phenotypes: Convergent findings in a mouse model of neglect. Behav Brain Res 315:71–74.

Murray J, Zhang B, Taylor SW, Oglesbee D, Fahy E, Marusich MF, Ghosh SS and Capaldi RA (2003) The subunit composition of the human NADH dehydrogenase obtained by rapid one-step immunopurification. J Biol Chem 278:13619–13622.

Newman MEJ (2006) Modularity and community structure in networks. Proc Natl Acad Sci USA 103:8577–8582.

Nguyen JM, Qualmann KJ, Okashah R, Reilly A, Alexeyev MF and Campbell DJ (2015) 5p deletions: Current knowledge and future directions. Am J Med Genet Part C Semin Med Genet 169:224–238.

Niebuhr E (1978) The cri du chat syndrome - epidemiology, cytogenetics, and clinical features. Hum Genet 44:227–275.

Oláh J and Ovádi J (2014) Dual life of TPPP/p25 evolved in physiological and pathological conditions. Biochem Soc Trans 42:1762–1767.

Oshimori N, Li X, Ohsugi M and Yamamoto T (2009) Cep72 regulates the localization of key centrosomal proteins and proper bipolar spindle formation. EMBO J 28:2066–2076.

Overhauser J, Huang X, Gersh M, Wilson W, Mcmahon J, Bengtsson U, Rojas K, Meyer M and Wasmuth JJ (1994) Molecular and phenotypic mapping of the short arm of chromosome 5: Sublocalization of the critical region for the cri-du-chat syndrome. Hum Mol Genet 3:247–252.

Perfumo C, Mainardi P, Cali A, Coucourde G, Zara F, Cavani S, Overhauser J, Bricarelli F and Pierluigi M (2000) The first three mosaic cri du chat syndrome patients with two rearranged cell lines. J Med Genet 37:967–972.

Perona MTG, Waters S, Hall FS, Sora I, Lesch KP, Murphy DL, Caron M and Uhl GR (2008) Animal models of depression in dopamine, serotonin, and norepinephrine transporter knockout mice: Prominent effects of dopamine transporter deletions. Behav Pharmacol 19:566–574.

Pietsch J (2006) Lessen encodes a Zebrafish Trap100 required for enteric nervous system development. Development 133:395–406.

Pogorelov VM, Rodriguiz RM, Insco ML, Caron MG and Wetsel WC (2005) Novelty seeking and stereotypic activation of behavior in mice with disruption of the Dat1 gene. Neuropsychopharmacology 30:1818–1831.

Posokhova E, Song H, Belcastro M, Higgins L, Bigley LR, Michaud NA, Martemyanov KA and Sokolov M (2011) Disruption of the chaperonin containing TCP-1 function affects protein networks essential for rod outer segment morphogenesis and survival. Mol Cell Proteomics 10:M110.000570.

Prestele M, Vogel F, Reichert AS, Herrmann JM and Ott M (2009) Mrpl36 is important for generation of assembly competent proteins during mitochondrial translation. Mol Biol Cell 20:2615–2625.

Riegel M, Barcellos N, Mergener R, Souza RS De, César J, Leite L, Gus R, Maria L, Moreira DA and Giugliani R (2014) Molecular cytogenetic evaluation of chromosomal microdeletions: The experience of a public hospital in southern Brazil. Clin Biomed Res 34:357–365.

Riegel M, Mergener R, Rosa RF and Zen P (2017) Chromosomal structural rearrangements: Characterizing interstitial deletions and duplications in the clinical practice. Arch Pediatr J 118:0–4.

Sato S, Tomomori-Sato C, Banks CAS, Sorokina I, Parmely TJ, Kong SE, Jin J, Cai Y, Lane WS, Brower CS *et al.* (2003) Identification of mammalian mediator subunits with similarities to yeast mediator subunits Srb5, Srb6, Med11, and Rox3. J Biol Chem 278:15123–15127.

Scardoni G, Petterlini M and Laudanna C (2009) Analyzing biological network parameters with CentiScaPe. Bioinformatics 25:2857–2859.

Shannon P, Markiel A, Owen Ozier 2, Baliga NS, Wang JT, Ramage D, Amin N, Schwikowski B and Ideker T (2003) Cytoscape: A software environment for integrated models of biomolecular interaction networks. Genome Res 2498–2504.

Smith AJ, Trewick AL and Blakemore AIF (2010) Implications of copy number variation in people with chromosomal abnormalities: Potential for greater variation in copy number state may contribute to variability of phenotype. Hugo J 4:1–9.

Sotnikova TD, Beaulieu JM, Barak LS, Wetsel WC, Caron MG and Gainetdinov RR (2005) Dopamine-independent locomotor actions of amphetamines in a novel acute mouse model of parkinson disease. PLoS Biol 3:e271.

Spiegel R, Shaag A, Mandel H, Reich D, Penyakov M, Hujeirat Y, Saada A, Elpeleg O and Shalev SA (2009) Mutated NDUFS6 is the cause of fatal neonatal lactic acidemia in Caucasus Jews. Eur J Hum Genet 17:1200–1203.

Steegers-Theunissen RPM, Boers GHJ, Trijbels FJM, Finkelstein JD, Blom HJ, Thomas CMG, Born GF, Wouters MGAJ and Eskes TKAB (1993) Maternal hyperhomocysteinemia: A risk factor for NTDs? Metabolism 43:1475–1480.

Tong JHS, Cummins TDR, Johnson BP, Mckinley LA, Pickering HE, Fanning P, Stefanac NR, Newman DP, Hawi Z and Bellgrove MA (2015) An association between a dopamine transporter gene (SLC6A3) haplotype and ADHD symptom measures in nonclinical adults. Am J Med Genet Part B Neuropsychiatr Genet 168:89–96.

Vincze O, Tökési N, Oláh J, Hlavanda E, Zotter Á, Horváth I, Lehotzky A, Tirián L, Medzihradszky KF, Kovács J *et al.* (2006) Tubulin polymerization promoting proteins (TPPPs): Members of a new family with distinct structures and functions. Biochemistry 45:13818–13826.

von Mering C, Jensen LJ, Snel B, Hooper SD, Krupp M, Foglierini M, Jouffre N, Huynen MA and Bork P (2005) STRING: Known and predicted protein-protein associations, integrated and transferred across organisms. Nucleic Acids Res 33:433–437.

Williams EH, Perez-Martinez X and Fox TD (2004) MrpL36p, a highly diverged L31 ribosomal protein homolog with additional functional domains in *Saccharomyces cerevisiae* mitochondria. Genetics 167:65–75.

Wu Q, Niebuhr E, Yang H and Hansen L (2005) Determination of the "critical region" for cat-like cry of Cri-du-chat syndrome and analysis of candidate genes by quantitative PCR. Eur J Hum Genet 13:475–485.

Yang P, Zhao Z and Reece EA (2008) Activation of oxidative stress signaling that is implicated in apoptosis with a mouse model of diabetic embryopathy. Am J Obstet Gynecol 198:130.e1-7.

Yu D, Shang Y, Yuan J, Ding S, Luo S and Hao L (2016) Wnt/β-catenin signaling exacerbates keloid cell proliferation by regulating telomerase. Cell Physiol Biochem 39:2001–2013.

Yu H, Kim PM, Sprecher E, Trifonov V and Gerstein M (2007) The importance of bottlenecks in protein networks: Correlation with gene essentiality and expression dynamics. PLoS Comput Biol 3:713–720.

Yuan L, Seong E, Beuscher JL and Arikkath J (2015) Delta-catenin regulates spine architecture via cadherin and PDZ-dependent interactions. J Biol Chem 290:10947–10957.

Yui J, Chiu CP and Lansdorp P (1998) Telomerase activity in candidate stem cells from fetal liver and adult bone marrow. Blood 91:3255–3262.

Zhang A, Zheng C, Hou M, Lindvall C, Li KJ, Erlandsson F, Björkholm M, Gruber A, Blennow E and Xu D (2003) Deletion of the telomerase reverse transcriptase gene and haploinsufficiency of telomere maintenance in Cri du Chat syndrome. Am J Hum Genet 72:940–948.

Zhang X, Snijders A, Segraves R, Zhang X, Niebuhr A, Albertson D, Yang H, Gray J, Niebuhr E, Bolund L *et al.* (2005) High-resolution mapping of genotype-phenotype relationships in Cri du Chat syndrome using array comparative genomic hybridization. Am J Hum Genet 76:312–326.

Zhu H, Wicker NJ, Shaw GM, Lammer EJ, Hendricks K, Suarez L, Canfield M and Finnell RH (2003) Homocysteine remethylation enzyme polymorphisms and increased risks for neural tube defects. Mol Genet Metab 78:216–221.

Supplementary material

The following online material is available for this article:

Table S1 - List of genes located in the smallest overlap region.

Table S2 - List of Go terms identified by BiNGO in the Cluster 1.

Table S3 - List of Go terms identified by BiNGO in the Cluster 2.

Table S4 - List of Go terms identified by BiNGO in the Cluster 3.

Table S5 - List of Go terms identified by BiNGO in the Cluster 4.

Table S6 - List of Go terms identified by BiNGO in the Cluster 5.

Table S7 - List of Go terms identified by BiNGO in the Cluster 6.

Table S8 - List of Go terms identified by BiNGO in the Cluster 7.

Table S9 - List of Go terms identified by BiNGO in the Cluster 8.

Table S10 - List of Go terms identified by BiNGO in the Cluster 9.

Table S11 - List of Go terms identified by BiNGO in the Cluster 10.

Table S12 - List of Go terms identified by BiNGO in the Cluster 11.

Table S13 - List of Go terms identified by BiNGO in the Cluster 12.

Table S14 - List of Go terms identified by BiNGO in the Cluster 13.

Table S15 - List of Go terms identified by BiNGO in the Cluster 14.

Table S16 - List of Go terms identified by BiNGO in the Cluster 15.

Table S17 - List of Go terms identified by BiNGO in the Cluster 16.