Indian J Med Res 145, January 2017, pp 39-50
DOI: 10.4103/ijmr.IJMR_965_14

Quick Response Code:

Identifying the risk of producing aneuploids using meiotic recombination genes as biomarkers: A copy number variation approach

Raviraj V. Suresh, Kusuma Lingaiah, Avinash M. Veerappa & Nallur B. Ramachandra

Department of Studies in Genetics & Genomics, University of Mysore, Mysuru, India

Received July 2, 2014

***Background & objectives*: Aneuploids are the most common chromosomal abnormality in liveborns and are usually the result of non-disjunction (NDJ) in meiosis. Copy number variations (CNVs) are large structural variations affecting the human genome. CNVs influence critical genes involved in causing NDJ by altering their copy number which affects the clinical outcome. In this study influence of CNVs on critical meiotic recombination was examined using new computational technologies to assess their role in causing aneuploidy.**

***Methods*: This investigation was based on the analysis of 12 random normal populations consisting of 1714 individuals for aneuploid causing genes under CNV effect. To examine the effect of CNVs on genes causing aneuploidy, meiotic recombination genes were analyzed using EnrichR, WebGestalt and Ingenuity Pathway Analysis (IPA).**

***Results*: Forty three NDJ genes were found under CNV burden; IPA (Ingenuity Pathway Analysis) and KEGG (Kyoto Encyclopedia of Genes and Genomes) pathway analysis of CNV in meiotic recombination genes revealed a significant role of breast cancer gene 1, amyloid protein precursor, mitogen-activated protein kinase and nerve growth factor as key molecular players involved in causing aneuploidy. Interaction between these genes with other CNV-overlapping genes involved in cell cycle, recombination and meiosis might lead to increased incidences of aneuploidy.**

***Interpretation & conclusions*: The findings of this study implied that the effect of CNVs on normal genome contributed in amplifying the occurrences of chromosomal aneuploidies. The normal individuals consisting of variations in the susceptible genes causing aneuploids in the population remain undetected until the disorder genes express in the succeeding generations.**

Key words Aneuploidy - copy number variations - meiotic recombination - non-disjunction

Aneuploidy is a complex chromosomal disorder which is the most frequently occurring human anomaly worldwide[1]. Genetic and epidemiological studies have revealed the prevalence of aneuploidies to 0.3 per cent[2]. Many molecules involved in the dysregulation of chromosome segregation and recombination have been identified. This anomaly is due to chromosomal non-disjunction (NDJ) which can be defined as the malsegregation of chromosomes during meiosis. Multiple risk factors are likely involved in chromosome NDJ; these act at different times in the meiotic process and can be either of genetic or environmental origin.

The two well-established risk factors for causing NDJ are maternal age and altered recombination[1,3-5].

Extensive genetic studies have identified numerous genes associated with NDJ, but these studies have focussed entirely on cytogenetic studies, single nucleotide variations and microsatellite analysis, ignoring the contribution from the other form of large variation, such as copy number variants (CNVs). In addition, large variations in regions of genome which create complex genomic architecture making it susceptible towards the generation of aneuploids, are yet to be identified. It has been implied that large structural variations across the human genome such as CNVs contribute to human diseases and population diversity[5].

CNVs are structural variations with either duplications or deletions affecting human genome by 1 kb to several Mb and present at variable copy number in comparison with a reference genome[6]. CNVs are a common phenomenon which occurs in both normal and in disease condition; however, the frequency is more in latter than the former. The pathogenicity of CNVs is correlated with their presence, location and size. Genes mapping near/to CNVs have been shown to alter the expression levels compared to those transcripts that are devoid of CNVs, and since normal individuals consisting of variations in the susceptible genes causing aneuploids in the population remain undetected until the disorder genes express in the succeeding generations, it is necessary to understand the frequency burden of CNVs in regions containing genes susceptible to aneuploids, particularly Down's syndrome (DS).

Studies have been conducted on leukaemia and lymphoma where the prevalence of such chromosomal rearrangement bearing *BCL2-IgH* genes in the healthy population from Europe and America was determined to be around 40-60 per cent[7]. It was shown that the incidence of rearrangements in Japanese population was lower (16%) when compared to that in the German population (52%)[8]. Rearrangements have also been reported amongst the healthy population in India[9]. These findings emphasize the role of such variations present in normal, healthy groups and would require contributions from other factors. Therefore, for the prediction of NDJ, it is necessary to understand the role of candidate gene in normal population. The present study was an attempt towards the identification of large variations burden on meiotic recombination genes in 12 randomly selected normal populations and to also assess whether CNVs played a role in NDJ and represented the functional DS susceptibility.

Material & Methods

Study populations: For this study, a total of 1774 individuals were selected from 12 populations across the globe, which included (*i*) 43 samples from randomly selected 12 families residing in Mysuru, Chikmagalur and Davangere across Karnataka, India (13-73 yr), (*ii*) 270 samples from HapMap covering four populations such as CEU [The Centre de'Etude du Polymorphism Humain (CEPH) collection] (18-100 yr), CHB (Han Chinese in Beijing, China), JPT (Japanese in Tokyo, Japan) and YRI (Yoruba in Ibadan, Nigeria) populations (45 yr mean age), (*iii*) 31 Tibetan samples, (*iv*) 155 Chinese samples, (*v*) 472 of Ashkenazi Jews (AJ) replicate I, (*vi*) 480 of AJ replicate II, (*vii*) 204 individuals from Taiwan, (*viii*) 55 from Australia, and (*ix*) 64 from New World population (Totonacs and Bolivians)[10-14].

The 270 HapMap individuals' sample data from the four populations were obtained from the International HapMap Consortium[15]. The samples (270) from the HapMap included 30 both-parent-and-adult-child trios from the Yoruba people in Ibadan, Nigeria, 45 unrelated Japanese individuals in Tokyo, 45 unrelated individuals Han Chinese in Beijing and the 30 both-parent-and-adult-child trios from CEPH.

Molecular studies: Five millilitre blood in ethylenediamine tetraacetic acid (EDTA) was collected from each member of the Indian study group (*i.e.* 43 members from Mysuru, Karnataka) and genomic DNA was extracted using Promega Wizard® Genomic DNA purification kit (Promega, USA). The isolated DNA was quantified by biophotometer and gel electrophoresis. Later, the DNA was sent for whole genome scan and the result obtained was processed using Affymetrix array. This study was conducted in the department of Studies in Genetics and Genomics, University of Mysore, Mysuru, India, and the study protocol was approved by the University of Mysore Institutional Human Ethics Review Committee (IHEC). Written informed consent was obtained from all sample donors. Written informed consent was obtained from parents/guardians in the cases of participants being minors. Samples were collected during 2009-2012; however, the study was carried out from 2012 to 2013.

The raw, unprocessed data from Affymetrix Genome-Wide SNP 6.0 array (Licensed version) for all the remaining populations, except India were obtained

from the ArrayExpress Archive of the European Bioinformatics Institute.

Data analysis: Data analysis was done using methods described in our earlier publication[16].

Gene ontology and pathway analysis: EnrichR is a state-of-the-art gene set enrichment analysis which includes application for gene-set libraries rank-based enriched terms and various interactive (*http://amp.pharm.mssm.edu/Enrichr/*). NDJ pathway analysis was performed using WebGestalt 'WEB-based GEne SeT AnaLysis Toolkit' a web-based enrichment analysis tool (*http://webgestalt.org/option.php*). KEGG pathway analysis was performed using WebGestalt web service (*http://bioinfo.vanderbilt.edu/webgestalt/*) which provided information about reactome and OMIM.

Ingenuity Pathway Analysis (IPA) software (Ingenuity Systems, CA, USA)[17] was used to identify the interaction between genes, protein-protein interactions, biological mechanisms, location and functions of target genes. Genes and chemical search were used to explore the information on protein families, protein signalling and metabolic pathways along with normal cellular activity of protein. Genes and their protein products were shown based on their location. An edge (line) was used to represent the relationship between two nodes. Each edge in the network was represented underpinning the information in Ingenuity Pathways Knowledge Base[17] and references from the literatures available. Ingenuity Pathways Knowledge Base was considered for the analysis. A cascade of protein-protein interaction or protein binding, molecular cleavage, activation, upregulation, downregulation and expression of mRNA by targeting mature miRNA network was analyzed. GeneMANIA tool (*http://genemania.org/*) was used to identify the interacting genes. Toenjes *et al*[18] predicted the possible associated genes in both cardiac diseases and neurodevelopmental diseases. Statistical analysis for enriched genes was done with EnrichR analysis web-based tool, and Chi-square test was performed for all CNV-NDJ genes across the populations.

Breakpoint validation: Polymerase chain reaction (PCR) amplification was performed on four recurring CNV breakpoints in 500 randomly chosen individuals from India to validate the presence of CNV. The primers used for amplifying the breakpoint regions were - AGGTCTGTTATGTGGCTGAGCCGCA on 3q29 for breakpoints 195276060 bp - 195446910 bp; ACTCTAGCCAACACATCCTCTGCGC on 15q14 for 34695310 bp - 34857998 bp; GAGTAAAGAAACAAAGGCCATCT on 21q11.2 for 14594223 bp - 15101046 bp and AGGGATCCACCCCCTGGCTGTGGA on 16p13.11 for 16377650 bp - 16635603 bp at annealing temperatures of 64, 73, 62 and 71°C, respectively (Integrated DNA Technologies, USA). Samples not containing these specific CNVs failed to amplify.

Results

Copy number variation (CNV) burden on non-disjunction (NDJ) genes: From 1714 individuals, a total of 44,109 CNVs consisting of about 126,190 genes (of which, 15,185 shared genes) were identified. The analysis was done by choosing 178 genes known to be causal or associated with NDJ. A total of 44 NDJ genes were identified which had a direct impact of CNVs in 190 individuals across 12 populations with a mean of 0.22 per cent and were found over-represented in duplication (81.08%) compared to deletion regions (19.91%). The 43 NDJ genes overlapped by CNVs identified in the study are shown in Table I with their functions and copy number state; depending on the variation (duplication or deletion), severity of the chromosomal disorders and phenotypic expressions. *SLC19A1*, *SMC1A* involved in congenital heart diseases and DNA repair were hit by CNVs more frequently across populations and affected more individuals. The CNV-NDJ genes were found distributed across populations in varying concentrations; AJI and AJII populations showed more burden of CNVs on NDJ genes. Sex bias was observed as these CNV-NDJ genes were present more in males compared to female (Table II) which implied more males acting as carriers of CNVs in NDJ genes. Chi-square test was performed for all CNV-NDJ genes across populations, and the results were found to be significant; except AJI, AJII and Australian populations (Table III).

A total of 29 shared CNV breakpoints were found across all populations which were distributed in all chromosomes except chromosome 3, 8, 13, 14, 18 and Y (Fig. 1). Amongst all the chromosomes, chromosome 21 showed the highest concentration, followed by chromosomes 16 and 4. The highest frequency was observed for the CNV breakpoint 46847759 - 46966180 bp in chromosome 21q22.3, and the global frequency was observed at 4.98 per cent for the same breakpoint. The highest frequency for this breakpoint in AJ (>33%) and the least for YRI (1.8%).

Copy number variation-non-disjunction (CNV-NDJ) gene classification based on gene ontology: Functional

Table I. Copy number variant overlapping genes involved in non-disjunction and recombination identified in normal cohorts

CNV-NDJ gene	CN state	Pathway/process
NGF	3, 1	Development and maintenance of sensory nervous systems
MAPK8/3/4/7	1	Cell proliferation, differentiation and development
NQO1	1	Quinone reductase in detoxification pathways
AUTS2	3	Expressed in brain, skeletal muscle and kidney
CNN2	3	Regulation and modulation of smooth muscle
DSCR9	3	Non-coding protein
DSCR3	3	Pathogenesis of DS
SLC19A1	3, 1	Transportation of folate
COL6A1	3, 1	Integrity of various tissues
COL6A2	3	This gene encodes collagen
CD2AP	3, 1	Dynamic actin remodelling and membrane trafficking
OLIG1/2	3	Adult neurogenesis
CBR3	4	Oxidoreductase activity
MORC3	1	Protein for nuclear matrix
CHAF1B	1	Histone octamers onto newly-replicated DNA
TTC3	3	Neuronal differentiation inhibition
BRWD1	3	Cellular processes
HMGN1	3	Altering the interaction of DNA and histone
WRB	3	Pathogenesis of DS congenital heart disease
MTRR	3	One-carbon metabolism
RFC1	1	Telomere stability
DOPEY2	4, 1	Protein traffic between late Golgi and early endosomes
RNF212	3	Meiotic recombination
SMC1A	3, 1	Functional kinetochores
SMC1B	4	Chromatid cohesion and recombination
STAG2	3, 1	Regulates the separation of sister chromatids
SYCP1	3	Transverse filaments of synaptonemal complexes
BUB1	3	Activating the spindle checkpoint
RAD51	3	Homologous recombination and repair of DNA
SMC2	3	Chromatin into condense chromosomes
NCAPG	3	Condensation and stabilization of chromosomes
NCAPD3	3	Mitotic chromosome assembly and segregation
MAD1L1	3	Cell cycle control and tumour suppression
DYNLL1	3, 1	Intracellular transport and motility
DMC1	3, 1	Homologous recombination
PMS2	1	DNA mismatch repair
RAD52	1	DNA repair
SLX1A/SLX1B	3	Regulator of genome stability

CNV, copy number variant; NDJ, non-disjunction; CN, copy number; DS, Down's syndrome; MAPK, mitogen-activated protein kinase; SMC; structural maintenance of chromosome; *BRWD1*, bromodomain and WD repeat-containing protein 1; NGF, nerve growth factor; *NQO1*, NAD (P) H quinone dehydrogenase; *AUTS2*, activator of transcription and developmental regulator; *CNN2*, calponin 2; *DSCR9*, Down syndrome critical region 9; *DSCR3*, Down syndrome critical region 3; *SLC19A1*, solute carrier family 19 member 1; *COL6A1*, collagen type VI alpha 1 chain; *COL6A2*, collagen type VI alpha 2 chain; *CD2AP*, CD2 associated protein; *OLIG1/2*, oligodendrocyte transcription factor 1/2; *CBR3*, carbonyl reductase 3; *MORC3*, MORC family CW-type zinc finger 3; *CHAF1B*, chromatin assembly factor 1 subunit B; *TTC3*, tetratricopeptide repeat domain 3; *BRWD1*, bromodomain and WD repeat domain containing 1; *HMGN1*, high mobility group nucleosome binding domain 1; *WRB*, tryptophan rich basic protein; *MTRR*, 5-methyltetrahydrofolate-homocysteine methyltransferase reductase; *RFC1*, Reduced folate carrier 1; *DOPEY2*, dopey family member 2; *RNF212*, ring finger protein 212; *STAG2*, stromal antigen 2; *SYCP1*, synaptonemal complex protein 1; *BUB1*, BUB1 mitotic checkpoint serine/threonine kinase; *RAD51*, RAD51 recombinase; *NCAPG*, non-SMC condensin I complex subunit G; *MAD1L1*, MAD1 mitotic arrest deficient like 1; *DYNLL1*, dynein light chain LC8-type 1; *DMC1*, DNA meiotic recombinase 1; *PMS2*, PMS1 homolog 2, mismatch repair system component; *RAD52*, recombinase RAD52; *SLX1A/SLX1B*, SLX1 homolog A, structure-specific endonuclease subunit

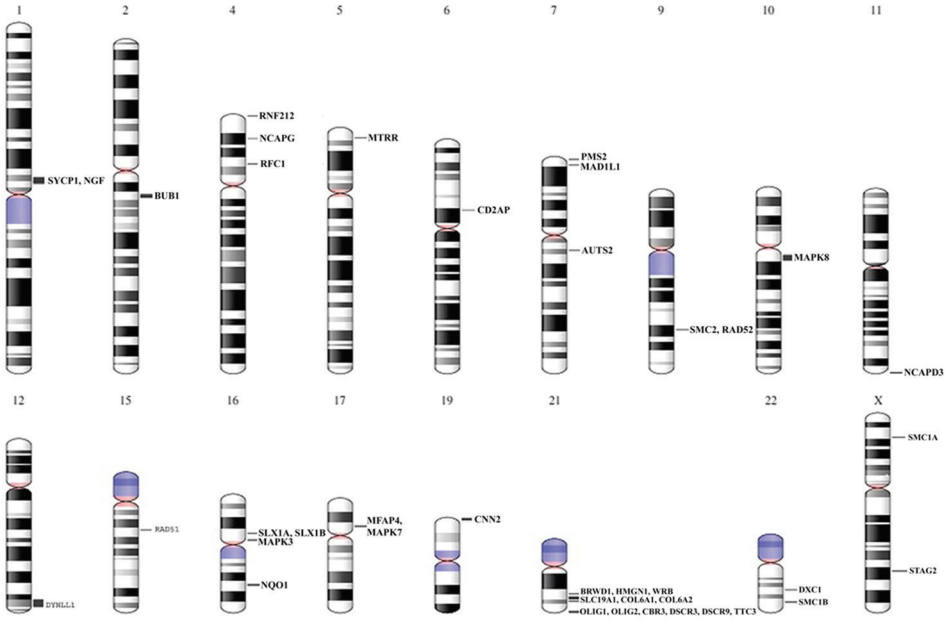

Fig. 1. Karyogram showing burden of copy number variation overlapping genes identified in the normal individuals across 12 populations which are possibly involved in causing aneuploidy. These genes were distributed across all chromosomes except 3, 8, 13, 14, 18 and Y; while copy number variation-enriched genes were found more in chromosome 21.

categories of CNV-enriched NDJ genes were identified and classified into biological process, cellular components and molecular function. In biological process, the CNV-NDJ genes were highly enriched in reciprocal meiotic recombination, while the remaining included organelle, chromosome organization, DNA metabolic process and DNA-packaging. Cellular components included CNV-NDJ genes that were part of nuclear chromosome, nucleus, nucleoplasm and extracellular matrix; while genes in molecular functions were involved in mitogen-activated protein kinase (MAPK) activity, chromatin binding, DNA binding, nucleoside binding and nucleotide binding (Table IV). CNV-NDJ genes were found enriched in signalling pathways such as MAPK signalling pathways, gonadotropin-releasing hormone (GNRH) signalling pathway, focal adhesion, pancreatic cancer (Fig. 2A) and interaction of these genes was found linked to cell cycle checkpoints, apoptosis, DNA replication and repair (Fig. 2B). These CNV-enriched genes were known to cause diseases such as myopathy, muscular dystrophy, colorectal cancer, breast cancer and anaemia (Fig. 2C).

Pathway mapping using WebGestalt: CNV-enriched NDJ genes were subjected to WebGestalt analysis resulting in several KEGG pathways; human oocyte meiosis pathway was picked as NDJ event (Fig. 3) and thought to play a central role. This pathway showed *MAPK1, MAPK3, SMC1* and Budding uninhibited by benzimidazoles 1(*BUB1*) (*R*=36.67; rawP=4.47e-06; adjP=3.72e-05; Here, *R*=ratio of enrichment, rawP=*P* value from hypergeometric test and adjP=*P* value adjusted by the multiple test adjustment) playing major role in human oocyte meiosis and these genes were hit by CNV in our study. This indicated that the CNVs on these genes would alter their expression level and dysregulate their function in normal process of oocyte meiosis depending on their CN state (Table I).

Interaction of biomolecules and their influence on aneuploidy: We examined the interaction of biomolecules and their possible role in causing NDJ using IPA. Forty four CNV-NDJ genes were subjected to IPA along with other biomolecules which were not directly related in causing NDJ. This interactive analyses led to a complex network that showed four major hubs, namely, breast cancer gene 1 (*BRCA1*), amyloid protein precursor (*APP*) complex, *MAPK* complex and Nerve Growth Factor (*NGF*). *BRCA1* was found interacting with many proteins which were involved in recombination, spindle assembly checkpoint mechanisms and microtubule formation; while *APP* was found to be interacting with *RAD*

Table II. Burden of non-disjunction genes under copy number variants in males and females of the study populations

Genes	YRI		CEU		AJI		AJII		China		Tibet		India		Australia		Taiwan		New World		JPT	
	Male	Female	Male	Female	Male	Female	Male	Female	Male	Female	Male	Female	Male	Female	Male	Female	Male	Female	Male	Female	Male	Female
NGF	-	2.7	-	2.2	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	N/A	-	-
MAPK8	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	N/A	-	-
MAPK3	-	-	-	-	-	-	-	0.41	-	-	-	-	-	-	-	-	-	-	-	N/A	-	-
MFAP4	-	-	-	-	0.42	-	-	-	-	-	-	-	-	-	-	-	-	-	-	N/A	-	-
MAPK7	-	-	-	-	0.42	-	-	-	-	-	-	-	-	-	-	-	-	-	-	N/A	-	-
NQO1	-	-	-	-	0.42	-	-	-	-	-	-	-	-	-	-	-	-	-	-	N/A	-	-
AUTS2	-	-	-	-		-	0.84	1.23	1.25	-	-	-	-	-	-	-	-	-	-	N/A	-	-
CNN2	-	-	-	-	0.42	-	-	-	2.5	-	-	-	-	-	-	-	-	-	-	N/A	-	-
DSCR9	-	-	-	-	0.42	-	-	-	-	-	-	-	-	-	-	-	-	-	-	N/A	-	-
DSCR3	-	-	-	-	0.42	-	-	-	-	-	-	-	-	-	-	-	-	-	-	N/A	-	-
SLC19A1	-	-	-	-	4.23	3.98	15.81	9.05	-		-	-	-	-	-	-	9.78	8.69	-	N/A	-	-
COL6A1	-	-	-	-	0.42	-	0.42	-	1.25	1.31	-	-	-	-	6.06	-	-	-	-	N/A	-	-
COL6A2	-	-	-	-	-	-	0.42	-	-	-	-	-	-	-	-	-	-	-	-	N/A	-	-
CD2AP	-	-	-	-	-	-	-	-	-	-	-	5	5.5	-	-	-	-	-	-	N/A	-	-
OLIG2	-	-	-	-	-	-	0.42	-	-	-	-	-	-	-	-	-	-	-	-	N/A	-	-
OLIG1	-	-	-	-	-	-	0.42	-	-	-	-	-	-	-	-	-	-	-	-	N/A	-	-
CBR3	-	-	-	-	0.84	-	-	-	-	-	-	-	-	-	-	-	-	-	-	N/A	-	-
MORC3	-	-	-	-	-	-	-	-	-	-	-	-	-	-	3.03	-	-	-	-	N/A	-	-
CHAF1B	-	-	-	-	-	-	-	-	-	-	-	-	-	-	3.03	-	-	-	-	N/A	-	-
TTC3	-	-	-	-	0.42	0.44	-	-	-	-	-	-	-	-	-	-	-	-	-	N/A	-	-
BRWD1	-	-	-	-	0.84	0.44	-	-	-	-	-	-	-	-	-	-	-	-	-	N/A	-	-
HMGN1	-	-	-	-	0.84	0.44	-	-	-	-	-	-	-	-	-	-	-	-	-	N/A	-	-
WRB	-	-	-	-	0.84	0.44	-	-	-	-	-	-	-	-	-	-	-	-	-	N/A	-	-
MTRR	-	-	-	-	0.42	0.44	-	-	-	-	-	-	-	-	-	-	-	-	-	N/A	-	-
RFC1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	N/A	-	-
DOPEY2	-	-	-	-	0.42	-	-	-	-	-	-	-	-	-	3.03	-	-	-	-	N/A	-	-
RNF212	-	-	-	-	-	-	-		-	-	-	-	-	-	-	-	-	-	2.43	N/A	-	-
SMC1A	-	-	-	-	-	4.4	0.85	1.6	-	-	-	-	-	-	3.03	26.31	-	-	-	N/A	-	4.5
SMC1B	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	N/A	-	4.5
STAG2	-	-	-	-	-	-	-	-	-	-	-	-	-	10	-	-	-	-	-	N/A	-	-
SYCP1	-	0.1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	N/A	-	-
BUB1	-	-	-	-	0.42	-	-	-	-	-	-	-	-	-	-	-	-	-	-	N/A	-	-
RAD51	-	-	-	-	-	0.44	-	-	-	-	-	-	-	-	-	-	-	-	-	N/A	-	-
SMC2	-	-	-	-	0.42	-	-	-	-	-	-	-	-	-	-	-	-	-	-	N/A	-	-
NCAPG	-	-	-	-	-	0.44	-	-	-	-	-	-	-	5	-	-	-	-	-	N/A	-	-
NCAPD3	-	-	-	-	-	-	0.42	-	-	-	-	-	-	-	-	-	-	-	-	N/A	-	-
MAD1L1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	5.2	-	-	-	N/A	-	-
DYNLL1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	3.03	-	-	-	-	N/A	-	-
DMC1	-	-	-	-	-	-	-	-	-	-	-	-	-	5	-	-	-	-	-	N/A	-	-
PMS2	-	-	-	-	-	-	-	0.4	-	-	-	-	-	-	-	-	-	-	-	N/A	-	-
SMC2	-	-	-	-	-	-	-	-	-	-	-	-	-	-	3.03	5.2	-	-	-	N/A	-	-
SLX1A	-	-	-	-	3.5	1.1	-	-	-	-	-	-	-	-	-	-	-	-	-	N/A	-	-
SLX1B	-	-	-	-	3.3	1.0	-	-	-	-	-	-	-	-	-	-	-	-	-	N/A	-	-

N/A, not available

Table III. Chi-square test analysis for the copy number variant-non-disjunction genes across the study populations

Populations	Individuals†	Number of genes	Observed	χ^2	*P*
HapMap-YRI-Africa	90	2	1.12	22.521	0.02
Ashkenazi Jews I	464	18	10.11	8.626	0.65
Ashkenazi Jews II	480	14	7.86	11.49	0.40
HapMap-CHB-China	44	1	0.56	23.603	0.01
China	155	2	1.12	22.521	0.02
Tibet	31	2	1.12	23.521	0.02
India	38	4	2.24	20.433	0.04
HapMap-JPT-Japan	45	1	0.56	23.603	0.01
Australia	53	10	5.61	14.748	1.19
New World	41	2	1.12	23.521	0.02
Taiwan	184	1	0.56	23.603	0.01

†Individuals whose data passed quality control test were only analyzed

(RAD51 recombinase), bromodomain and WD repeat-containing protein 1 (*BRWD1*) and *CHAF1B* (Chromatin assembly factor 1 subunit B), *NGF* and *MAPK* molecules, which are complexed in causing DS, also affects neuronal growth and signalling cascades. *MAPK8* was involved in activating *DYNLL1, MTHFR* and *MTRR* which are involved in folate metabolism, and also involved in causing DS, is the most frequently occurring aneuploidy. A few small hubs were also formed in nucleus between structural maintenance of chromosome (*SMC*), *STAG* and *RAD* molecules (Fig. 4).

Our analysis revealed that CNVs in these genes played a crucial role in causing NDJ as these altered expression level of these genes. Further, these CNV hit NDJ genes might interact with other genes, dysregulating their expression and causing severity in diseases or disorders. Our analysis also revealed these CNV-NDJ genes potentially dysregulated the molecular pathway in normal individuals, thus increasing the risk of giving birth to aneuploids.

Discussion

Cataloguing the nature and pattern of genome variations in the general population is fundamental in understanding human phenotypic diversity[19]. Identification of the CNVs across diverse populations helps understand distribution pattern of CNVs, their organization and evolutionary dynamics of the human genome[6, 20]. Since CNVs embedded within the regions of chromosome may create imbalance and affect the clinical outcomes by altering the local copy number of important genes or regulatory regions, this could alleviate or exacerbate certain phenotypes. Thus, CNVs contribute to the clinical variability seen in many disorders caused by chromosomal abnormalities[21]. Therefore, this study was taken up to address the issue of CNV burden on genes causing NDJ in normal population.

Selection and pooling of NDJ genes from earlier studies through literature embarked our investigation; these genes were searched and screened for CNV overlaps in them amongst the 12 study populations. Genes overlapped by CNVs were selected for EnrichR, WebGestalt analysis and for IPA. EnrichR analysis revealed CNV genes to be involved in reciprocal meiotic recombination, while the cellular components were involved in nuclear chromosome and molecular functions were involved in MAPK activity and chromatin binding. WebGestalt pathway results also indicated CNVs role in MAPK signalling pathway during oocyte meiosis. The impact of these deletion CNVs during MAPK signalling may lead to abnormal oocyte meiosis. IPA revealed *BRCA1* and *APP* to be crucial players and showed an interaction with the already identified genes *NGF* and *MAPK* in causing aneuploidy, which has been found to fall under CNV in our study. Genes overlapped by these CNVs alter the expression of the genes or nearby genes by either increasing or lowering their functions, resulting in abnormal processes.

WebGestalt pathway analysis indicated the participation of four CNV-overlapping genes *MAPK3*, *MAPK8, SMC1* and *BUB1* in human oocyte meiosis. Since NDJ may occur at any stage during meiosis, the above pathway was chosen for our study and was

Table IV. Copy number variant-non-disjunction (CNV- NDJ) genes involved in cellular, molecular and biological processes with their *P* value and *Z*-score by enrich analysis tool

Function	CNV-NDJ genes	*Z*-score	*P*
Cellular category			
Nuclear chromosome (GO: 0000228)	*SMC2, SMC1A, DMC1*	-2.3571	0.0001
Chromosome (GO: 0005694)	*SMC2, SMC1A, DMC1*	-2.2152	0.0003
Condensed chromosome (GO: 0000793)	*DMC1, SMC1A*	-1.5353	0.0023
Condensed nuclear chromosome (GO: 0000794)	*DMC1, SMC1A*	-1.2572	0.0012
Promyelocytic leukaemia body (GO: 0016605)	*MORC3, RAD51*	-1.7030	0.0061
Nuclear body (GO: 0016604)	*MORC3, RAD51*	-1.5696	0.0168
Chromatin (GO: 0000785)	*HMGN1, STAG2*	-1.3785	0.0148
Extracellular matrix part (GO: 0044420)	*MFAP4, COL6A1*	-1.5773	0.0253
Protein complex (GO: 0043234)	*NCAPD3, MAD1L1, CHAF1B, NCAPG, DYNLL1, SMC2, CD2AP, RFC1, SMC1A*	-1.6864	0.0373
Molecular category			
MAPK activity (GO: 0004707)	*MAPK8, MAPK7*	-2.2330	4.4721
Threonine kinase activity (GO: 0004702)	*MAPK8, MAPK7*	-2.4458	0.0048
Chromatin binding (GO: 0003682)	*SMC1A, CHAF1B*	-2.3165	0.0040
Single-stranded DNA binding (GO: 0003697)	*PMS2, RAD51*	-2.2365	0.0042
Double-stranded DNA binding (GO: 0003690)	*PMS2, RAD51*	-2.0979	0.0040
Nucleoside binding (GO: 0001882)	*RFC1, MTRR, DMC1*	-2.2970	0.0179
Structure-specific DNA binding (GO: 0043566)	*PMS2, RAD51*	-1.9978	0.0123
Purine nucleotide binding (GO: 0017076)	*RFC1, MTRR, DMC1*	-2.1533	0.0298
Nucleotide binding (GO: 0000166)	*RFC1, MTRR, DMC1*	-2.1033	0.0381
Biological category			
Organelle organization (GO: 0006996)	*NCAPD3, DYNLL1, SYCP1, DOPEY2, CNN2, SMC2, RFC1, CHAF1B, NCAPG, SMC1A*	-2.3230	3.0765
Chromosome organization (GO: 0051276)	*NCAPD3, RFC1, CHAF1B, NCAPG, SYCP1, SMC1A, SMC2*	-2.2879	4.5812
DNA packaging (GO: 0006323)	*SMC2, NCAPD3, NCAPG*	-2.3833	5.1512
Reciprocal meiotic recombination (GO: 0007131)	*SYCP1, RAD51, DMC1*	-2.2929	3.9311
Chromosome condensation (GO: 0030261)	*SMC2, NCAPD3, NCAPG*	-2.0875	2.4911
Mitotic chromosome condensation (GO: 0007076)	*NCAPG, SMC2, NCAPD3*	-1.8666	9.4515
DNA metabolic process (GO: 0006259)	*RFC1, RAD51, PMS2, SYCP1, SMC1A, DMC1*	-2.3138	2.4634
DNA recombination (GO: 0006310)	*RAD51, DMC1, SYCP1*	-2.0310	6.6013
Female gamete generation (GO: 0007292)	*DYNLL1, DMC1*	-1.9826	9.5212
Cell cycle checkpoint (GO: 0000075)	*MAD1L1, BUB1, SMC1A*	-1.8940	0.0011

Significance value *P*=0.05. Full forms of the genes are as given in Table I

further verified for the influence of CNVs on NDJ genes in oocyte meiosis. A complex interplay exists between MAPK and MPF in the meiosis regulation[22]. *MAPK* plays a pivotal role in the regulation of oocyte meiosis and is also an important component of microtubule organizing centre helping in functioning of microtubules. CN state of 1, deletion, for *MAPK* in the present study indicated that this lower expression level (normal as per gene atlas 48 vs. 24 in the present study) may lead to aberrant chromosome condensation and spindle organization due to a lower concentration of kinase. *BUB1* and *SMC1* are involved in chromosome segregation and act as downstream molecules in the pathway. An increased dosage of *BUB1* and *SMC1* may

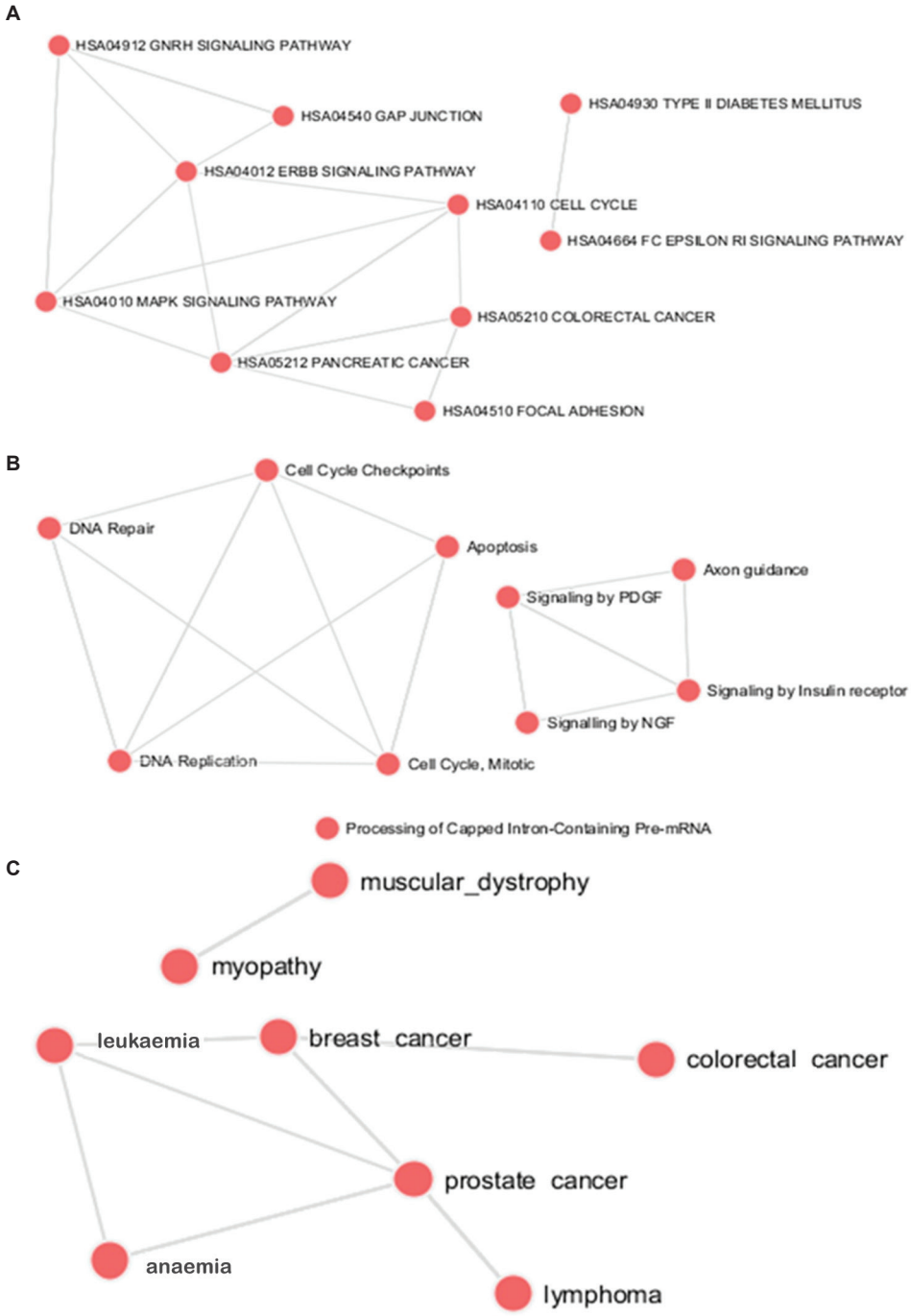

Fig. 2. Representation of disease network linked to the copy number variation-non-disjunction (CNV- NDJ) genes carried out by enrichment analysis performed using EnrichR web service. (**A**) Major Kyoto Encyclopedia of Genes and Genomes (KEGG) grid network. (**B**) Reactome grid network. (**C**) Major Online Mendelian Inheritance in Man (OMIM) disease network (*http://amp.pharm.mssm.edu/Enrichr/*).

promote the chromosome segregation, but the lack of MAPK prevents any of the downstream processes, thus leading to NDJ. Therefore, any such complex interactions created by the influence of CNVs on these NDJ genes will lead to the disruption of normal pathway.

IPA revealed several major and minor hubs linking to each other. Minor hubs contained genes involved in chromosome segregation, cohesion and SMC and a few amongst them were under CNV influence. These genes and their function would be affected by the impact of CNV leading to deformity in cell cycle machinery.

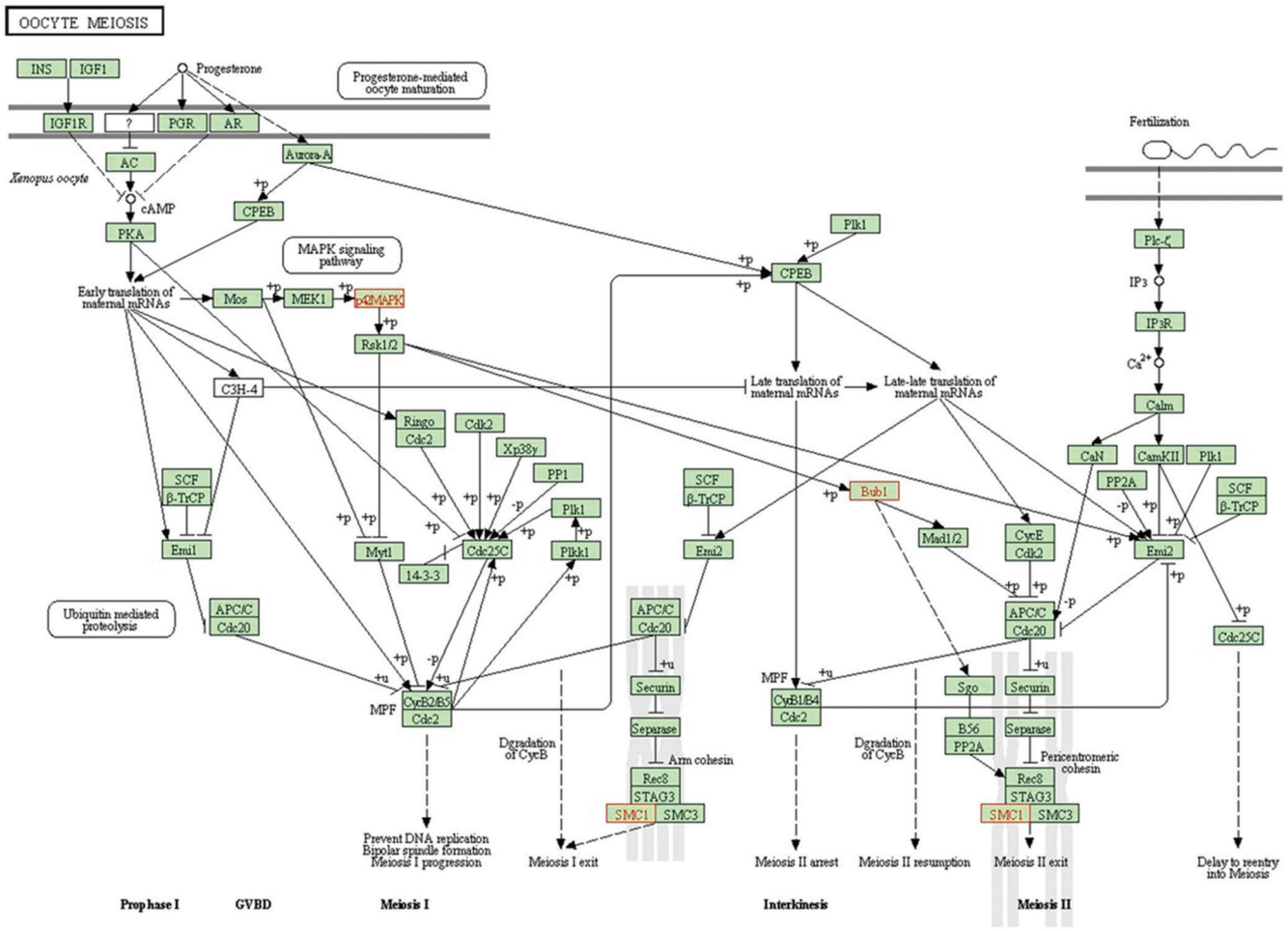

Fig. 3. The KEGG pathway of normal oocyte meiosis. The pathway highlights three copy number variation-enriched genes (mitogen-activated protein kinase, structural maintenance of chromosome 1 and Bub1) indicated in red blocks. The upregulation/downregulation of these genes will affect the normal oocyte pathway. These genes are identified in the major steps *i.e.*, meiosis I and meiosis II.

Four major hubs were formed with *BRCA1*, *APP*, *MAPK* and *NGF* as central molecules. *MAPK* and *NGF* were found overlapping with CNVs in the study and were involved in a wide variety of cellular processes; growth and maintenance, respectively. *BRCA1* showed communication with *XRCC2* (X-ray repair cross complementing 2), *PMS2*, *RAD51*, *RFC1* (reduced-folate-carrier 1), *MRE11A*, *SMC1A*, *MAPK8*, *ERG1*, *OLIG1* and *PSMG1* forming a clustered network, genes in this network were found to be involved in recombination, DNA repair, cellular processes and in causing diseases and disorders[23]. *APP* showed an interaction with *DYRK1* (Dual specificity tyrosine phosphorylation regulated kinase 1A), *MAPK* (mitogen-activated protein kinase), *NGF*, *MTHFR* (Methylene tetrahydrofolate reductase), *BRWD1*, *MTHFD1L* [Methylenetetrahydrofolate dehydrogenase (NADP$^+$ dependent) 1], *RAD51C*, *ERG1* (*ETS*-related gene), *CHAF1B* (Chromatin assembly factor 1 subunit B), *SYCP2* (Synaptonemal complex protein 2) and *RCAN1* (Regulator of calcineurin 1) leading to curiosity of its role in aneuploidy as *APP* is known to cause Alzheimer's disease which is also one of the phenotypes in DS[24].

Increase in *APP* molecules has been shown to downregulate NGF growth factor which marks the pathogenesis of aneuploids[25]. As *MAPK* plays a major role in the normal process of oocyte meiosis, any alterations in the interactions between *APP*, *NGF* and *MAPK* create dysregulation in the normal oocyte meiosis and cell cycle. Our analysis findings supported the possible role of *APP* and *NGF* as key players in causing abnormal cell cycle mechanism

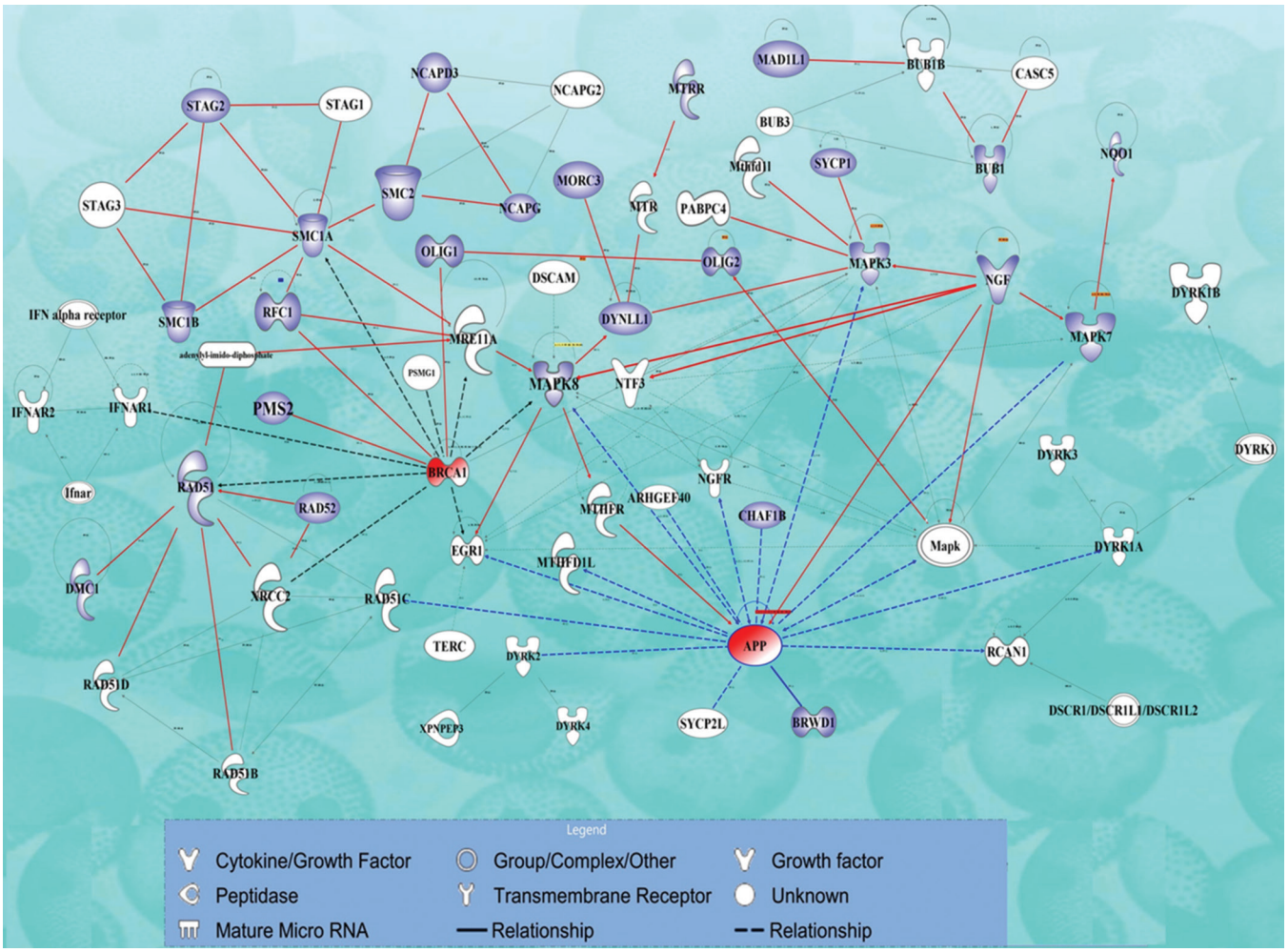

Fig. 4. Representation of the molecular interaction pathway analysis of the associated and causal genes linked to aneuploidy. Purple coloured molecules indicate the copy number variation overlapping genes identified in the study. Breast cancer gene 1 (*BRCA1*), amyloid protein precursor (*APP*) formed major hubs, interacting with other CNV- NDJ genes involved in meiotic recombination and DNA repair and hence are indicated in red colour. Paths between genes represent protein-protein interactions, regulation and phosphorylation activities. This complex network analysis was performed using Ingenuity Pathway Analysis tool (*http://www.ingenuity.com/products/ipa*).

and pathogenesis of aneuploids. It also suggests that effect of CNVs on NDJ genes and interaction between one CNV hit NDJ gene with another will influence extent of aneuploidy. Thus, we hypothesize that other secondary mutations are required to commence the events of aneuploidy in addition to CNVs on NDJ genes.

In conclusion, the effect and influence of CNVs on normal genome contributes in amplifying the occurrence of diseases and disorders such as chromosomal aneuploidies. CNV-overlapping critical genes involved in cell cycle regulation, meiosis and recombination would increase the incidences of aneuploidy in the offspring.

Acknowledgment

Authors acknowledge the Chairman, DOS in Zoology, University of Mysore, for providing facility to conduct this work; and Drs K. S. Rangappa and H. A. Ranganath for their support and encouragement. Authors thank Dr Prakash P, DOS in Psychology, University of Mysore, for diagnosing dyslexia patients, and all our laboratory colleagues, Genetics and Genomics Laboratory for their support.

***Conflicts of Interest*:** None.

References

1. Hassold T, Hunt P. To err (meiotically) is human: the genesis of human aneuploidy. *Nat Rev Genet* 2001; *18* : 69-97.
2. O'Connor C. Chromosomal abnormalities: aneuploidies. *Nat Educ* 2008; *1* : 172.

3. Malini SS, Ramachandra NB. Influence of advanced age of maternal grandmothers on Down syndrome. *BMC Med Genet* 2006; *7* : 4.
4. Malini SS, Ramachandra NB. Young mothers produce more chromosomal syndrome babies in Mysore, South India. *Int J Hum Genet* 2010; *10* : 105-12.
5. Girirajan S, Campbell CD, Eichler EE. Human copy number variation and complex genetic disease. *Annu Rev Genet* 2011; *45* : 203-26.
6. Redon R, Ishikawa S, Fitch KR, Feuk L, Perry GH, Andrews TD, *et al.* Global variation in copy number in the human genome. *Nature* 2006; *444* : 444-54.
7. Bäsecke J, Griesinger F, Trümper L, Brittinger G. Leukemia- and lymphoma-associated genetic aberrations in healthy individuals. *Ann Hematol* 2002; *81* : 64-75.
8. Yasukawa M, Bando S, Dölken G, Sada E, Yakushijin Y, Fujita S, *et al.* Low frequency of *BCL-2*/J(H) translocation in peripheral blood lymphocytes of healthy Japanese individuals. *Blood* 2001; *98* : 486-8.
9. Nambiar M, Raghavan SC. How does DNA break during chromosomal translocations? *Nucleic Acids Res* 2011; *39* : 5813-25.
10. McElroy JP, Nelson MR, Caillier SJ, Oksenberg JR. Copy number variation in African Americans. *BMC Genet* 2009; *10* : 15.
11. Chen W, Hayward C, Wright AF, Hicks AA, Vitart V, Knott S, *et al.* Copy number variation across European populations. *PLoS One* 2011; *6* : E23087.
12. Zhang YB, Li X, Zhang F, Wang DM, Yu J. A preliminary study of copy number variation in Tibetans. *PLoS One* 2012; *7* : E41768.
13. Lin CH, Li LH, Ho SF, Chuang TP, Wu JY, Chen YT, *et al.* A large-scale survey of genetic copy number variations among Han Chinese residing in Taiwan. *BMC Genet* 2008; *9* : 92.
14. Liu X, Cheng R, Ye X, Verbitsky M, Kisselev S, Mejia-Santana H, *et al.* Increased rate of sporadic and recurrent rare genic copy number variants in Parkinson's disease among Ashkenazi Jews. *Mol Genet Genomic Med* 2013; *1* : 142-54.
15. International HapMap Consortium. The international HapMap project. *Nature* 2003; *426* : 789-96.
16. Veerappa AM, Vishweswaraiah S, Lingaiah K, Murthy M, Suresh RV, Manjegowda DS, *et al*. Global spectrum of copy number variations reveals genome organizational plasticity and proposes new migration routes. *PloS One* 2015; *10* : e0121846.
17. Ingenuity pathway analysis software (IPA). Available from: *https://www.qiagenbioinformatics.com/products/ingenuity-pathway-analysis/*, accessed on April 2, 2014.
18. Toenjes M, Schueler M, Hammer S, Pape UJ, Fischer JJ, Berger F, *et al.* Prediction of cardiac transcription networks based on molecular data and complex clinical phenotypes. *Mol Biosyst* 2008; *4* : 589-98.
19. Pinto D, Marshall C, Feuk L, Scherer SW. Copy-number variation in control population cohorts. *Hum Mol Genet* 2007; *16* : R168-73.
20. Conrad DF, Pinto D, Redon R, Feuk L, Gokcumen O, Zhang Y, *et al.* Origins and functional impact of copy number variation in the human genome. *Nature* 2010; *464* : 704-12.
21. de Smith AJ, Trewick AL, Blakemore AI. 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* 2010; *4* : 1-9.
22. Fan HY, Tong C, Chen DY, Sun QY. Roles of MAP kinase signaling pathway in oocyte meiosis. *Chin Sci Bull* 2002; *47* : 1-7.
23. Titus S, Li F, Stobezki R, Akula K, Unsal E, Jeong K, *et al.* Impairment of BRCA1-related DNA double-strand break repair leads to ovarian aging in mice and humans. *Sci Transl Med* 2013; *5* : 172ra21.
24. Granic A, Padmanabhan J, Norden M, Potter H. Alzheimer A beta peptide induces chromosome mis-segregation and aneuploidy, including trisomy 21 : requirement for tau and APP. *Mol Biol Cell* 2010; *21* : 511-20.
25. Zekanowski C, Wojda U. Aneuploidy, chromosomal missegregation, and cell cycle reentry in Alzheimer's disease. *Acta Neurobiol Exp (Wars)* 2009; *69* : 232-53.

Reprint requests: Dr Nallur B. Ramachandra, Department of Studies in Zoology, Genetics & Genomics Laboratory, University of Mysore, Mysuru 570 006, Karnataka, India
e-mail: nallurbr@gmail.com