RESEARCH

Open Access

Risk homozygous haplotype regions for autism identifies population-specific ten genes for numerous pathways

Swati Agarwala and Nallur B. Ramachandra*

Abstract

Background: Recessive homozygous haplotype (rHH) mapping is a reliable tool for identifying recessive genes by detecting homozygous segments of identical haplotype structures. These are shared at a higher frequency amongst probands compared to parental controls. Finding out such rHH blocks in autism subjects can help in deciphering the disorder etiology.

Objectives: The study aims to detect rHH segments of identical haplotype structure shared at a higher frequency in autism subjects than controls to identify recessive genes responsible for autism manifestation.

Methods: In the present study, 426 unrelated autism genotyped probands with 232 parents (116 trios) were obtained from Gene Expression Omnibus (GEO) Database. Homozygosity mapping analyses have been performed on the samples using standardized algorithms using the Affymetrix GeneChip® 500K SNP Nsp and Sty mapping arrays datasets.

Results: A total of 38 homozygous haplotype blocks were revealed across sample datasets. Upon downstream analysis, 10 autism genes were identified based on selected autism candidate genes criteria. Further, expressive Quantitative Trait Loci (QTL) analysis of SNPs revealed various binding sites for regulatory proteins *BX3*, *FOS*, *BACH1*, *MYC*, *JUND*, *MAFK*, *POU2F2*, *RBBP5*, *RUNX3*, and *SMARCA4 impairing essential autism genes CEP290*, *KITLG*, *CHD8*, *and INS2*. Pathways and processes such as adherens junction, dipeptidase activity, and platelet-derived growth factor—vital to autism manifestation were identified with varied protein-protein clustered interactions.

Conclusion: These findings bring various population clusters with significant rHH genes. It is suggestive of the existence of common but population-specific risk alleles in related autism subjects.

Keywords: Homozygosity, Recessive genes, Autism, Mutations, Haplotype blocks

Background

Autism is a heritable, neurodevelopmental condition affecting information processing in the brain, heterogeneous with < 15% known genetic causes. It has a worldwide prevalence rate of 1 in 59 children being affected [1]. It alters connections and organization of nerve cells and their synapses, impairing the overall cognition, emotional, social, and physical health of the affected individuals [2].

The study of recessive risk gene loci is performed through extended runs of homozygosity (ROH) as a genomic feature, useful to map recessive disease genes in outbred populations [3, 4].

The authors expect to find an unusually higher number of affected individuals in complex disorders to have the identical haplotype in the region surrounding a disease [5, 6]. Therefore, a rare pathogenic variant and surrounding haplotype are often enriched in frequency in a group of affected individuals compared to a cohort of unaffected controls [7]. These variants in the haplotypes

* Correspondence: nbruom@gmail.com
Department of Studies in Genetics and Genomics, University of Mysore, Mysuru, Karnataka 06, India

indicate that shared homozygous haplotypes (HH) in multiple affected individuals result in identifying recessive genes in diseases such as autism. The recessive risk gene loci approach has been proven to help understand autism genetics and behavioral severity.

Gene mapping of rare recessive conditions in a large outbred population has been a herculean task due to the lack of multiple affected individuals in families. Homozygosity mapping is an efficient gene mapping method applicable to rare recessive disorders since small chromosomal regions tend to be transmitted whole. Affected individuals will also have identical by descent alleles at markers located near to disease locus and thus will be homozygous at these markers [8, 9]. The basic idea is thus to look for regions of homozygosity shared amongst different affected individuals. This study would identify disease-causing mutations by pursuing the hypothesis-free genome-wide search for homozygosity blocks through an efficient homozygosity scan, using single nucleotide polymorphisms (SNPs) chip-based genotyping platforms: the Affymetrix GeneChip® 500K SNP Nsp and Sty mapping arrays. Given this, the present study aims to detect homozygous segments of identical haplotype structure shared at a higher frequency in autism subjects to distill out the recessive genes essential to autism manifestation.

Methods

The study included 426 unrelated probands analyzed along with 232 parents (116 trios) obtained from an international public repository Gene Expression Omnibus (GEO) Database (accession GSE9222). The unrelated autism samples have been collected from the Hospital for Sick Children, Toronto, by the primary investigators. The datasets have been identified with chromosomal abnormalities; however, their whole etiology was unknown. These have been genotyped for autism index cases using blood-derived DNA performed on the Affymetrix GeneChip® 500K single nucleotide polymorphism (SNP) mapping array (Nsp/Sty arrays) [10]. Out of 426 probands and 232 parents, 334 probands and 122 parents (61 trios) were scrutinized based on the type of microarrays used to perform genome variation profiling and SNP genotyping. This selection for the set of 334 probands and 122 parents was based on initial screening using Autism Diagnostic Observation Schedule [11] and Autism Diagnostic Interview-Revised [12] criteria on a clinical best estimate. The detailed case history of each case was used to select cases and controls based on classical autism without any other comorbidities such as intellectual disability, epilepsy, and mild autism to avoid noise in the datasets. Information for all the comorbidities has been considered, as mentioned in the primary investigation. The controls were the parents, wherever available.

If parental samples were not available, the Yoruba population was used as control. All samples were screened for Fragile X syndrome using karyotyping techniques, and if positive, the samples were excluded. Based on selection criteria, the samples genotyped using Nsp and Sty arrays were selected and analyzed for homozygosity mapping analysis (Table 1). The study aims to identify risk homozygous regions and the encompassed recessive genes for autism manifestation.

In view of this, all the subject samples have been analyzed against the controls (parental samples) in autism trios; otherwise, the HapMap Yoruba dataset has been used as the controls for autism proband cases and divided into categories for ease of understanding. Category A comprises trios with matched parental samples with subcategory A1: 61 trios with 122 controls on Affymetrix Mapping 250K Nsp SNP Array and A2: 61 trios 122 controls using Affymetrix Mapping 250K Sty SNP Array. Category B contains individuals autism subjects with subcategory B1: 273 autism subjects with HapMap Yoruba controls using Affymetrix Mapping 250K Nsp SNP Array and B2:273 autism cases with HapMap Yoruba controls using Affymetrix Mapping 250K Sty SNP Array. The subcategories were divided based on genotyping arrays used for analysis.

Whole-genome genotyping was performed on the autism subjects and controls using Genotyping Console to generate output CEL Files using select probe set summary. The genotyping console helped integrate SNP genotyping, generate genotyping calls, loss of heterozygosity (LOH) data, and quality control metrics for sample data.

Output CEL file datasets generated by the Genotyping Console were used to analyze the homozygous regions through Homozygosity Mapper (www.homozygositymapper.org) [13, 14]. Homozygosity Mapper is a web-based approach to homozygosity mapping to analyze and detect homozygous stretches. It provides an intuitive graphical interface to visualize the results. The homozygous regions in affected subjects were visualized as peaks chromosome-wise with underlying genotypes. Homozygosity scores were plotted against the physical position with a threshold of ≥ 4000. The length of the homozygous block (in SNPs) at each marker for each sample was calculated.

Genes involved in the homozygous regions were identified, filtered, sorted, and prioritized according to criteria derived from the background knowledge and preconception of autism using GeneDistiller (http://www.genedistiller.org/) [15] and Genome Map Viewer [16] (https://www.ncbi.nlm.nih.gov/genome/gdv/).

Golden Helix GenomeBrowse® visualization tool (Version 2.x) was used to visualize and browse the entire genome with annotated data, including gene prediction and structure,

Table 1 Comparative risk homozygous haplotype regions among overall landscape and specific to autism

Sample type	Category	Risk haplotype homozygous regions overall landscape				Risk haplotype homozygous regions-autism specific			
		Chromo somal location	Max score	rHH regions count	Genes identified number	Chromosomal loci	Max score*	rHH regions count	Associated genes identified
Autism Trios using Affymetrix Mapping 250k Nsp SNP Array	A1	2,4,7,8, 11,12	4500-5700	12	29	8: 50280717-50921942	5714	6	*SNTG1*-**RELN**
						11:49038697-49228613	4645		**FOLH1**
						11:48086151-48110043	4603		*PTPRJ*-**JAK2**
						12:88808147-89252667	5154		**KITLG**
						12:88459309-88681867	4592		**CEP290, KITLG**
Autism Trios using Affymetrix Mapping 250k Sty SNP Array	A2	2,8,9,10	4000-5000	11	102	16:31716601-46710869	4827	6	**SLC6A10P, TP53TG3, TP53TG3B**
Autism cases Yoruba HapMap genotype as Controls using Affymetrix Mapping 250k Nsp SNP Array	B1	4,8,11	21000-24000	9	28	8:50280717-50808334	24275	3	*SNTG1*-**RELN**
						11:49038697-49295867	21801		**FOLH1**
						11:48168664-48293931	21778		*PTPRJ*-**JAK2**
Autism cases Yoruba HapMap genotype as Controls using Affymetrix Mapping 250k Sty SNP Array	B2	3,11,16	18268	6	145	3:51880614-52047421	18268	3	**ACY1, ABHD14A**

*Threshold level of significance ($P = 0.05$) is scored at 4000 and above; max block length = 1000

**Genes in bold are previously known autism genes

protein product, and gene variation for SNP visualization [17]. In concurrence, expression Quantitative Trait Loci (eQTL) analysis was performed. eQTl loci are genomic regions with DNA sequence variants that influence the expression level of one or more genes. It was performed on the rules of the homozygous recessive genes with readseq identifiers (rsIDs) using RegulomeDB (https://regulomedb.org/regulome-search/). The chromosomal regions, bound proteins, affected genes, and enhancers were identified. rsIDs that fall under score 1 and score 2 were selected [18]. RegulomeDb was conditioned to identify known associated functional variants to depict its functionality. It interpreted genomic regulatory variants using computational predictions and manual annotations.

Genes identified from the homozygous blocks of the cases were analyzed, which meet the following criteria for selection set for autism gene. It should be novel autism candidate gene expressed in brain; participate in neuronal development; interact with known autism genes; non-homozygous in controls; de novo in origin; overlap in two or more unrelated samples; recurrent in two or more unrelated samples; involve in the expression of brain development and participate in neuronal migration, axon growth, neuritis outgrowth, synaptic plasticity, and cell adhesion.

Physical interactors of identified affected genes were predicted through a web-based application GeneMania (https://genemania.org/) [19]. Further, pathways were constructed with the enriched candidate genes and associated genes related to autism using Ingenuity® Systems, IPA software (www. ingenuity.com) [20]. Genes and the chemical search were used to explore the information on protein families, protein signaling, and metabolic pathways, and the regular cellular activity of the protein. Genes and their protein products are shown based on their location. This analysis was performed using methods adapted as in Veerappa and colleagues, 2014 [21].

Results

The study group comprises both males and females of average intellectual level, with no significant emotional and behavioral problems. The confidence interval of category A data is considerably higher than those of category B datasets. For category B datasets, 47.36% of the homozygous regions were exclusively significant in autism subjects than in controls enriched with autism-specific genes.

An integrated approach of whole-genome genotyping and homozygosity mapping revealed 38 homozygous regions highlighted, bearing 308 genes. Genome-wide
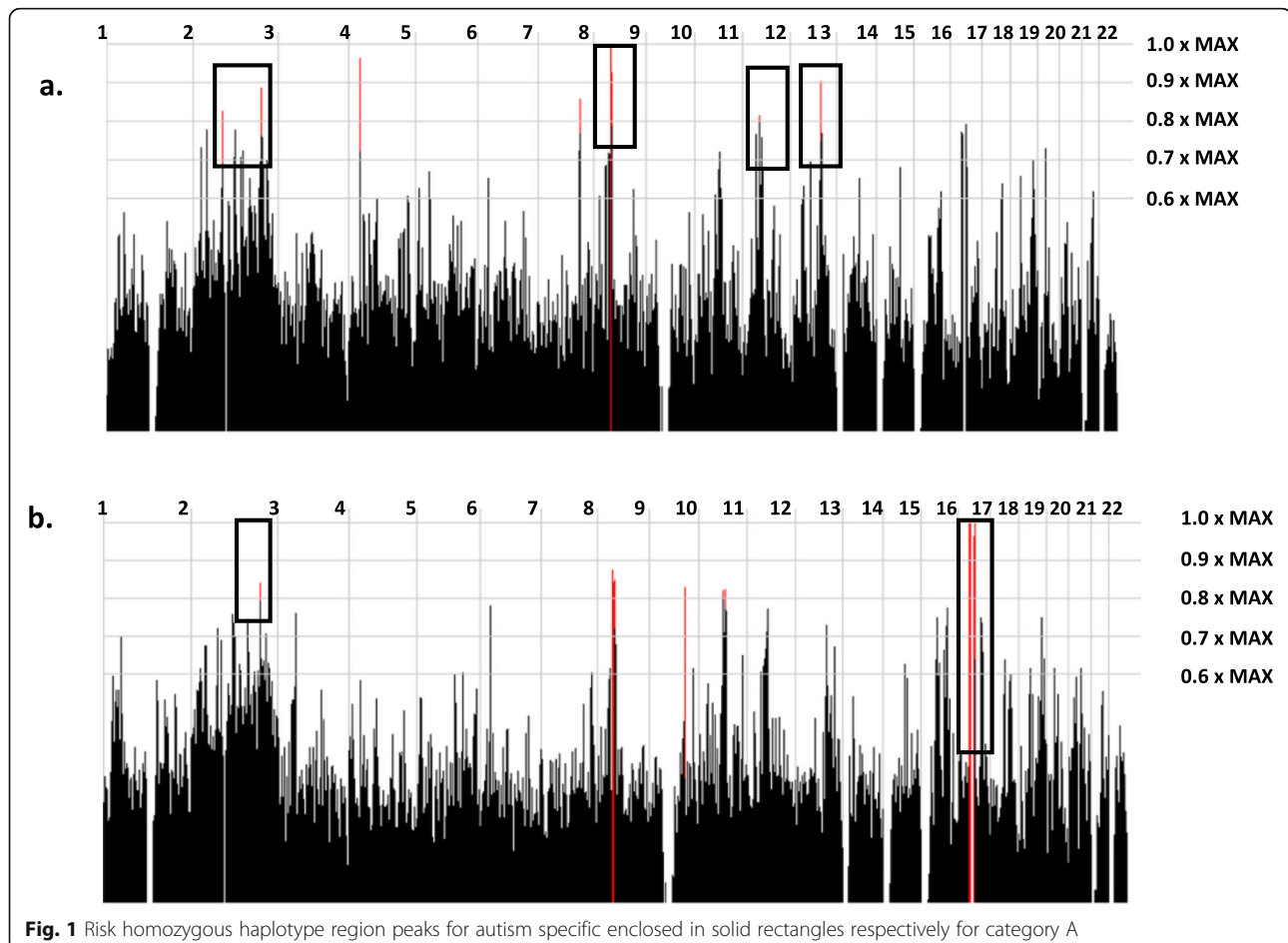
homozygosity data displayed scored blocks of haplotype compared to controls highlighted in red bars and autism-specific blocks enclosed in rectangle blocks. It depicted genomic haplotype regions for the sample cohort, passed on across generations (Figs. 1 and 2). Chromosomes 1, 13, 14, 15, 17-22 and sex chromosomes showed no homozygous regions identified.

Downstream analysis and autism gene selection criteria being put forth, the homozygosity blocks analysis showed 16 candidate autism recessive genes in cases partaking in neuronal development. Stringent filtration of risk homozygous haplotypes (rHH) based on homozygosity mapping-based pipeline revealed 12, 11, 9, and 6 homozygous regions respectively for category A1, A2, B1, and B2 datasets (max. score range 4500–25000; max. block length = 1000) using Homozygosity Mapper and applying the criteria for candidate gene selection revealed 6, 6, 3, and 3 homozygous regions with 5, 3, 3, and 2 autism genes respectively for subcategories A1, A2, B1, and B2 on varied chromosomes (Table 1).

eQTL analysis of the homozygous blocks in subcategory A1 revealed 4, 3, 2, and 3 polymorphisms

deregulating *CEP290, KITLG, and PTPRJ* genes (Table 2). Several polymorphisms were identified in subcategories A2: rs7078127, rs11866251, and rs16945839, impacting *RERE, CHD8,* and *EMSY* genes (Table 2). eQTL analysis in the homozygous blocks of subcategory B1 identified two polymorphisms: rs1681625 and rs12292520, deregulating significant autism candidate genes *PTPRJ-JAK2* and *INS2*-involved in brain development. Subcategory B2 revealed polymorphisms including rs17051043 and rs2675835 impairing *EP300* and *MAPK3* genes expressed in the brain from the developmental stage (Table 2). These contained valid RegulomeDB scores and rank affecting the chromatin state impacting motif at varied levels.

Pathway network analysis revealed a significant overrepresentation of these genes in physiological pathways: arginine and proline metabolism, protein farnesylation, protein geranylation, protein prenylation, aminoacylase activity, hydrolase activity, urea cycle, and metabolism of amino groups, GTPase binding- relevant to autism. Further, pathway analysis for significant genes found in homozygous regions with
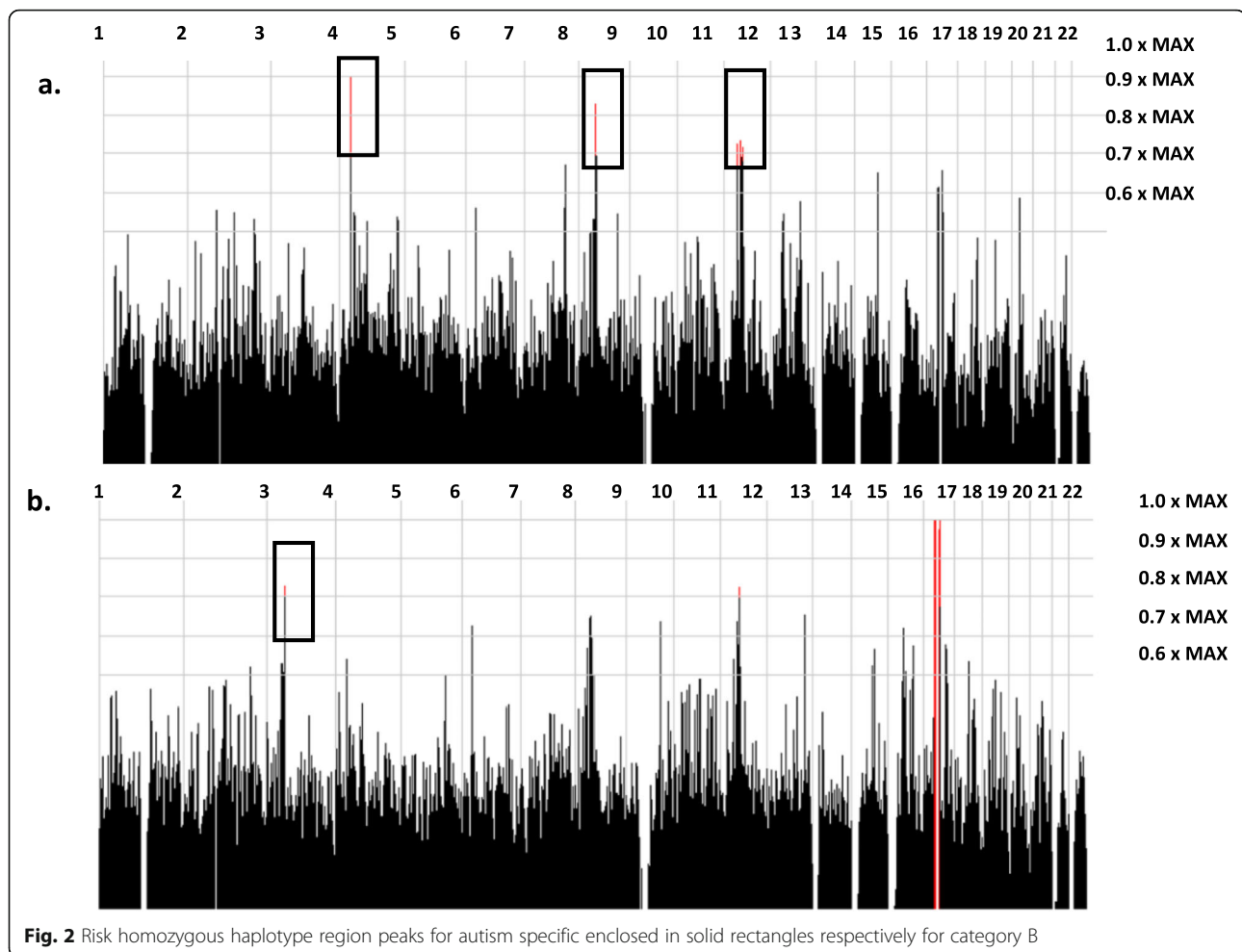

Fig. 1 Risk homozygous haplotype region peaks for autism specific enclosed in solid rectangles respectively for category A

Fig. 2 Risk homozygous haplotype region peaks for autism specific enclosed in solid rectangles respectively for category B

eQTL analysis using Ingenuity Pathway Analysis pipeline identified notable pathways and processes viz., migration of Purkinje cells, morphology, formation, and hypoplasia of brain, development of the head and central nervous system relevant to autism. Several gene hits such as *AUTS2, ADGRA2, CELSR1, FZD9, MBD1, RERE, and OPN1* were localized in various levels of these pathways (Fig. 3).

Discussion

Susceptible genetic loci identification is crucial for better understanding the underlying mechanisms of autism, thus aiding the development of their treatment and management [22]. Due to the high heritability of autism, various common genetic risk variants have surfaced, yet a long way to find rare variations showing the significance of heritability in autism [23]. The homozygous haplotype mapping approach is complimentary to genome wide association studies and Next Generation Sequencing in studying the complexity and heterogeneity of autism [6]. Homozygosity mapping analysis adds the missing pieces of the puzzle for complex disorders in

terms of heritability and recessive gene burden [6]. The present study reports the presence of rHH mapping study to identify candidate autism gene variants, particularly recessive gene loci involved in autism manifestation. The idea was to apply the concept of homozygosity mapping to the trios sample cohort and understand the role of haplotype blocks in unrelated subjects.

Studies have been performed to identify a genome-wide survey of runs of homozygosity for bone mineral density in Caucasian and Chinese populations [5]. Similarly, an integrated approach of whole-genome genotyping and homozygosity mapping revealed the presence of 38 rHH regions with a much higher degree of haplotype-sharing. Oversharing of haplotype indicates a disease locus, an observation that forms the basis of the current study. The authors observed that these cases shared 18 homozygous blocks bearing 24 recessive genes related to autism. Previous studies have reported recessive autism genes such as *ABHD14A, CADM2, CHRF AM7A, EPHA3, FGF10, GRIK2, GRM3, KCND2 PDZK1, present in the haplotype regions* [6]. *ABHD14A genes*

Table 2 eQTL analysis for the homozygous regions identified specific to autism across samples

Sample type	Category	rsID	SNP position	RDB rank*	RDB Score**	Ref/Alt	Bound proteins	Affected autism genes (predicted)
Autism Trios using Affymetrix Mapping 250kNSp SNP Array	A1	rs11104868	chr12: 88828441	1d	0.43	G/A	CBX3,FOS,BACH1, MYC,JUND, MAFK,POU2F2, RBBP5 ,RUNX3, SMARCA4	CEP290
		rs11104939	chr12: 88932363	3a	0.65	T/C	CEBPB	KITLG
		rs12292520	chr11: 48086150	2b	0.51	G/T	POLR2A	PTPRJ-JAK2
		rs17789481	chr11: 48110042	2b	0.59	A/G	EP300,FOXA1, GATA3,AR	
Autism Trios using Affymetrix Mapping 250k Sty SNP Array	A2	rs7078127	chr10: 74847633	1f	0.55	A/C	-	RERE
		rs11866251	chr16: 48227861	3a	1.00	G/A	EMSY	EMSY
		rs16945839	chr16: 48160409	3a	0.78	C/T	CTCF	CHD8
Autism cases with Yoruba population genotype as Controls using Affymetrix Mapping 250kNSp SNP Array	B1	rs1681625	chr11: 48020542	2b	0.40	A/G	HNF4A,ELF1, HNF4G,MYBL2, NFIC,POLR2A,SP1, SREBF1,CDX2, FOXA1,JUND,USF1, POLR2A,TBP	PTPRJ-JAK2, INS2
		rs12292520	chr11: 48086150	2b	0.51	G/T	POLR2A	
Autism cases with Yoruba population genotype as Controls using Affymetrix Mapping 250k Sty SNP Array	B2	rs17051043	chr3: 50693998	1d	0.94	C/T	GATA2	EP300
		rs2675835	chr3: 50862740	1f	0.55	G/A	-	MAPK3

*Threshold significant rank: 1a-3b;
**Threshold score: 0.40 and above; score 1: most likely to be a regulatory variant
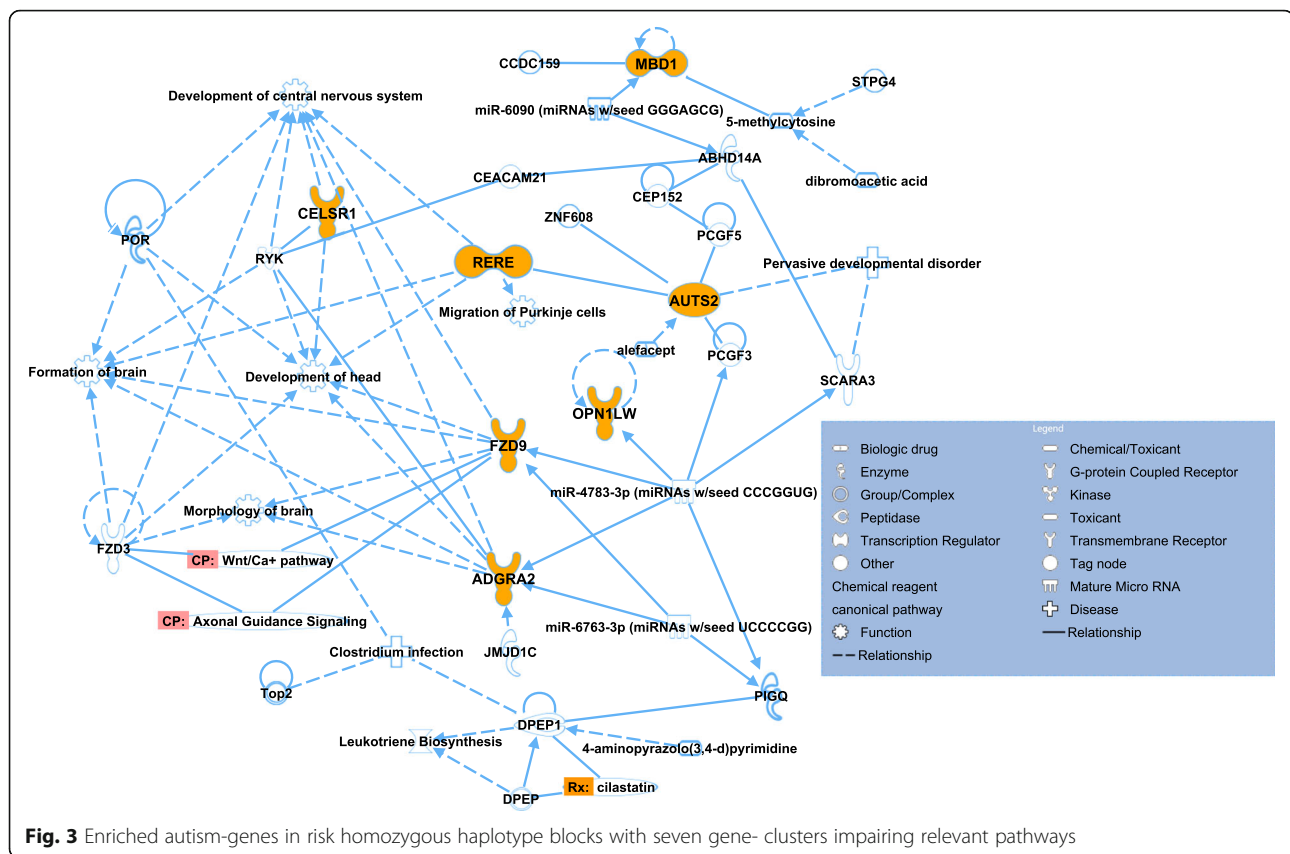*rsID* readseq identifier, *SNP* single nucleotide polymorphism, *RDB* Regulomedb, *Ref/Alt* reference/alternate

have been identified in the present study as well, impairing autism pathophysiology.

eQTL analysis of polymorphisms in the homozygous blocks affected many downstream genes such as *CEP290, KITLG, PTPRJ-JAK2, RERE, EMSY, CHD8, INS2, EP300,* and *MAPK3,* identified in subcategories A1, A2, B1, and B2, known for autism manifestation. These nine genes have been exclusively identified in the current study. However, *CNTN4, CADPS2, SUMF1, SLC9A9,* and *NTRK3* genes have been identified and implicated in autism elsewhere [24]. These genes show affected binding of regulatory proteins, which deregulates autism genes and impairs the downstream gene functionality and processes. Among the regulatory proteins, CBX3 [25] and BACH1 [26] are functioning as repressors, while FOS, BACH1, MYC, JUND, MAFK, POU2F2, RBBP5, RUNX3, and SMARCA4 are functioning as activators. These are involved in the activation of autism genes through various cellular functions. These cellular functions include chromatin organization, G protein-coupled receptor (GPCR) signaling, cell cycle regulation, homeostasis, signaling pathways, and cellular stress response, which play a vital role in

the severity of autism symptoms. Ubiquitination of calcium signaling pathway in endoplasmic reticulum across all body cells results in deficits observed in fibroblasts in autism subjects through neuronal functional impairment [27]. Intracellular GPCRs, linked to synaptogenesis, memory and learning, behavior, and cognition for pathophysiological roles in autism [28]. Cellular endoplasmic reticulum stress can lead to apoptosis, resulting in autism [29].

The establishment of the autism gene risk loci enrichment pathway led to identifying notable genes with a promising role in its pathogenesis. These include *AUTS2, JMJD1C, PCGF5, PCGF3, CEP152, ABHD14A, CEACAM21, A4GALT, OPN1MW, CELSR1, RYK, ADGR A2, FZD3, FZD9, NF608,* and *RERE*. These formed clustering at varied levels affecting downstream processes. *ABHD14A* has been implicated as a novel gene in a previous relevant autism haplotype study [6]. In yet another study, a family was reported to carry two changes in *ABHD14A,* a gene involved in cerebellar development and diagnosed with intellectual disability and a male obligate carrier [30]. Further studies need to be merited on

Fig. 3 Enriched autism-genes in risk homozygous haplotype blocks with seven gene- clusters impairing relevant pathways

each of these novel gene variants to decipher the complete cascade in haplotype analysis.

The pathway identified could be sliced down into seven gene pathway clusters: *AUTS2, ADGRA2, CELS R1, FZD9, MBD1, RERE, and OPN1.* Each of these genes physically interacts with autism-specific genes and bears receptors and ligands relevant to autism and affects the downstream cascade of processes like formation, morphology, hypoplasia, brain, and head development, development of a central nervous system, migration of Purkinje cells, and pervasive development disorder. ADGRA2 and CELSR1 share similar ectodomain structures, implicate in neural tube formation, as evident in network analysis in the current study as well [31]. Extensive studies have established their role in these autism genes directly or indirectly as follows: (1) regulation of actin cytoskeleton is performed by cytoplasmic AUTS2 to control neuronal migration and neurite extension, vital to autism [32]. (2). ADGRA2 plays a vital role during brain angiogenesis and has functionality as a WNT7A/7B-specific co-activator for beta-catenin signaling in the brain endothelium [33]. (3). CELSR1 protein is involved in processes of the neural progenitor cells in the basal compartment to decide the fate for the development of the cerebral cortex [34]. (4). Homozygous deletion in *FZD9* resulted in an acute deficiency in learning and memory, leading to apoptosis in the dentate gyrus and lowered seizure cut-off [35]. (5). Abnormal serotonin systems resulting from a deletion in MBD1 have been linked to autism due to its higher levels of receptor called Htr2c being synthesized. This system acts as an important link between MBD1 and autism-like behavior [36]. (6). Based on a discovery resource of rare copy number variations in autism individuals. De novo mutations have been identified to be associated with proximal 1p36 deletions and various processes of transcriptional, synaptic, and chromatin gene disruptions in autism [37]. (7). The OPN1 is a rho-linked protein for mental retardation, which controls synaptic vesicle endocytosis via endophilin A1, vital to autism [38].. These seven gene pathways bring out multiple population clusters with significant rHH genes suggesting the existence of common population-specific risk alleles.

Conclusion

Homozygous haplotype analysis would be an essential tool in uncovering the missing pieces of disease heritability. Replicative homozygous haplotypes in autism subjects showed enrichment for previously identified autism candidate genes, validating our approach. The

reported regions provide promising genomic regions containing highly plausible candidate genes. Further studies should be warranted in a larger cohort to validate the findings.

Abbreviations
rHH: Recessive Homozygous Haplotype; GEO: Gene Expression Omnibus; eQTL: Expression Quantitative Trait Loci; ROH: Runs of Homozygosity; HH: Homozygous Haplotypes; SNP: Single nucleotide polymorphisms; LOH: Loss of heterozygosity; *CEP290*: Centrosomal Protein 290kDa; *PTPRJ*: Protein Tyrosine Phosphatase, Receptor Type, J; *P4HA1*: Prolyl 4-Hydroxylase, Alpha Polypeptide I; *IRF7*: Interferon regulatory factor 7; *HEMK1*: HemKMethyltransferase Family Member 1; rsID: Readseq identifier; GPCR: G protein-coupled receptor

Authors' contributions
SA conceived the idea, designed the experiment, downloaded sample datasets, performed practical works including validation, analyzed data, and drafted the writing. NBR conceived the idea, helped in the design of the experiment and to finalize the manuscript. All authors have read and approved the manuscript.

Availability of data and materials
The data for the present study has been obtained from an international public repository Gene Expression Omnibus (GEO) Database (accession GSE9222). It included 426 unrelated probands analyzed along with 232 parents (116 trios).
Link for the dataset: https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc= GSE9222

Declarations

Ethics approval and consent to participate
The datasets used have been obtained from a public repository and all necessary consents and ethics have been obtained. Ethical clearances were obtained from the (1) Institutional Human Ethics Committee, University of Mysore, Mysuru, India—dated 12-05-2016 with reference number IHEC-UOM No. 128 Ph.D/2016-17 and (2) All India Institute of Speech and Hearing Ethics Committee for all research involving human subjects dated 09-02-2016 with reference no. MGG/258/2015-16. Informed written consent has been taken from all participants by the data provider and the authors have collected consent about the same from the data providers for the said publication.

Consent for publication
Not applicable

Competing interests
The authors declare that they have no competing interests.

References
1. Edition F. Diagnostic and statistical manual of mental disorders. *Am Psychiatric Assoc*. 2013;21.
2. Gilbert J, Man HY. Fundamental elements in autism: from neurogenesis and neurite growth to synaptic plasticity. *Front Cell Neurosci*. 2017;20:359.
3. Hildebrandt F, Heeringa SF, Rüschendorf F, Attanasio M, Nürnberg G, Becker C, et al. A systematic approach to mapping recessive disease genes in individuals from outbred populations. *PLos Genet*. 2009;5(1):e1000353. https://doi.org/10.1371/journal.pgen.1000353.
4. Melhem NM, Lu C, Dresbold C, Middleton FA, Klei L, Wood S, et al. Characterizing runs of homozygosity and their impact on risk for psychosis in a population isolate. *Am J Med Genet B Neuropsychiatr Genet*. 2014;165(6): 521–30. https://doi.org/10.1002/ajmg.b.32255.
5. Yang TL, Guo Y, Zhang JG, Xu C, Tian Q, Deng HW. Genome-wide survey of runs of homozygosity identifies recessive loci for bone mineral density in Caucasian and Chinese populations. *J Bone Miner Res*. 2015;30(11):2119–26. https://doi.org/10.1002/jbmr.2558.
6. Casey JP, Magalhaes T, Conroy JM, Regan R, Shah N, Anney R, et al. A novel approach of homozygous haplotype sharing identifies candidate genes in autism spectrum disorder. *Hum Genet*. 2012;131(4):565–79. https://doi.org/1 0.1007/s00439-011-1094-6.
7. Saitou M, Satta Y, Gokcumen O. Complex haplotypes of GSTM1 gene deletions harbor signatures of a selective sweep in East Asian populations. *G3*. 2018;8:2953–66.
8. Vahidnezhad H, Youssefian L, Jazayeri A, Uitto J. Research techniques made simple: genome-wide homozygosity/autozygosity mapping is a powerful tool for identifying candidate genes in autosomal recessive genetic diseases. *J Invest Dermatol*. 2018;138(9):1893–900. https://doi.org/10.1016/j. jid.2018.06.170.
9. Gholipoorfeshkecheh R, Agarwala S, Kavya G, Krishnappa S, Savitha MR, Ramachandra NB. Whole-exome sequencing and homozygosity mapping identify variants in NCOR1 and MAP2K3 associated with non-syndromic congenital heart defects. Egypt J Med Hum Genet. 2020;21:1–10.
10. Marshall CR, Scherer SW. Detection and characterization of copy number variation in autism spectrum disorder. *Methods Mol Biol*. 2012:115–35. https://doi.org/10.1007/978-1-61779-507-7_5.
11. Gotham K, Risi S, Pickles A, Lord C. The autism diagnostic observation schedule: Revised algorithms for improved diagnostic validity. *J Autism Dev Disord*. 2007;37(4):613–27. https://doi.org/10.1007/s10803-006-0280-1.
12. Rutter M, Le Couteur A, Lord C. Autism diagnostic interview-revised. Los Angeles, CA: Western. *Psychol Serv*. 2003;29:30.
13. Seelow D, Schuelke M. HomozygosityMapper2012-bridging the gap between homozygosity mapping and deep sequencing. *Nucleic Acids Res*. 2012;40(W1):W516–20. https://doi.org/10.1093/nar/gks487.
14. Seelow D, Schuelke M, Hildebrandt F, Nürnberg P. HomozygosityMapper - An interactive approach to homozygosity mapping. *Nucleic Acids Res*. 2009; 37(Web Server):W593–9. https://doi.org/10.1093/nar/gkp369.
15. Seelow D, Schwarz JM, Schuelke M. Genedistiller - Distilling candidate genes from linkage intervals. *PLoS One*. 2008;3(12):e3874. https://doi.org/10.1371/ journal.pone.0003874.
16. Wolfsberg TG. Using the NCBI map viewer to browse genomic sequence data. *Curr Protoc Bioinformatics*. 2010;29:1–5.
17. Golden HI. VarSeq. Bozeman, MT: Golden Helix, Inc; 2018.
18. Boyle AP, Hong EL, Hariharan M, Cheng Y, Schaub MA, Kasowski M, et al. Annotation of functional variation in personal genomes using RegulomeDB. *Genome Res*. 2012;22(9):1790–7. https://doi.org/10.1101/gr.137323.112.
19. Warde-Farley D, Donaldson SL, Comes O, Zuberi K, Badrawi R, Chao P, et al. The GeneMANIA prediction server: Biological network integration for gene prioritization and predicting gene function. *Nucleic Acids Res*. 2010;38(suppl_ 2):W214–20. https://doi.org/10.1093/nar/gkq537.
20. Krämer A, Green J, Pollard J, Tugendreich S. Causal analysis approaches in ingenuity pathway analysis. *Bioinformatics*. 2014;30(4):523–30. https://doi. org/10.1093/bioinformatics/btt703.
21. Veerappa AM, Murthy M, Vishweswaraiah S, Lingaiah K, Suresh RV, Nachappa SA, et al. Copy number variations burden on miRNA genes reveals layers of complexities involved in the regulation of pathways and phenotypic expression. *PLoS One*. 2014;9(2):e90391. https://doi.org/10.1371/ journal.pone.0090391.
22. Hyman SL, Levy SE, Myers SM. Identification, Evaluation, and Management of Children With Autism Spectrum Disorder. *Pediatrics*. 2020;145(1): e20193447. https://doi.org/10.1542/peds.2019-3447.
23. Grove J, Ripke S, Als TD, Mattheisen M, Walters RK, Won H, et al. Identification of common genetic risk variants for autism spectrum disorder. *Nat Genet*. 2019;51(3):431–44. https://doi.org/10.1038/s41588-019-0344-8.
24. Vardarajan BN, Eran A, Jung JY, Kunkel LM, Wall DP. Haplotype structure enables prioritization of common markers and candidate genes in autism spectrum disorder. *Transl Psychiatry*. 2013;3(5):e262. https://doi.org/10.1038/ tp.2013.38.

25. Ruan J, Ouyang H, Amaya MF, Ravichandran M, Loppnau P, Min J, et al. Structural basis of the chromodomain of Cbx3 bound to methylated peptides from histone H1 and G9a. *PLoS One*. 2012;7(4):e35376. https://doi.org/10.1371/journal.pone.0035376.
26. Reichard JF, Sartor MA, Puga A. BACH1 is a specific repressor of HMOX1 that is inactivated by arsenite. *J Biolumin Chemilumin*. 2008;283(33):22363–70. https://doi.org/10.1074/jbc.M801784200.
27. Nguyen RL, Medvedeva YV, Ayyagari TE, Schmunk G, Gargus JJ. Intracellular calcium dysregulation in autism spectrum disorder: an analysis of converging organelle signaling pathways. *Biochim Biophys Acta*. 1865;2018:1718–32.
28. Jong YJI, Harmon SK, O'Malley KL. Intracellular GPCRs play key roles in synaptic plasticity. *ACS Chem Neurosci*. 2018;9(9):2162–72. https://doi.org/10.1021/acschemneuro.7b00516.
29. Di Nardo A, Kramvis I, Cho N, Sadowski A, Meikle L, Kwiatkowski DJ, et al. Tuberous sclerosis complex activity is required to control neuronal stress responses in an mTOR-dependent manner. *J Neurosci*. 2009;29(18):5926–37. https://doi.org/10.1523/JNEUROSCI.0778-09.2009.
30. Cukier HN, Dueker ND, Slifer SH, Lee JM, Whitehead PL, Lalanne E, et al. Exome sequencing of extended families with autism reveals genes shared across neurodevelopmental and neuropsychiatric disorders. *Mol Autism*. 2014;5(1):1–10. https://doi.org/10.1186/2040-2392-5-1.
31. Langenhan T. Adhesion G protein–coupled receptors—Candidate metabotropic mechanosensors and novel drug targets. *Basic Clin Pharmacol Toxicol*. 2020;126(S6):5–16. https://doi.org/10.1111/bcpt.13223.
32. Hori K, Hoshino M. Neuronal migration and AUTS2 syndrome. *Brain Sci*. 2017;7(12):54. https://doi.org/10.3390/brainsci7050054.
33. Wang Y, Cho C, Williams J, Smallwood PM, Zhang C, Junge HJ, et al. Interplay of the Norrin and Wnt7a/Wnt7b signaling systems in blood–brain barrier and blood–retina barrier development and maintenance. *Proc Natl Acad Sci U S A*. 2018;115(50):E11827–36. https://doi.org/10.1073/pnas.1813217115.
34. Boucherie C, Boutin C, Jossin Y, Schakman O, Goffinet AM, Ris L, et al. Neural progenitor fate decision defects, cortical hypoplasia and behavioral impairment in Celsr1-deficient mice. *Mol Psychiatry*. 2018;23(3):723–34. https://doi.org/10.1038/mp.2017.236.
35. Zhao C, Avilés C, Abel RA, Almli CR, McQuillen P, Pleasure SJ. Hippocampal and visuospatial learning defects in mice with a deletion of frizzled 9, a gene in the Williams syndrome deletion interval. *Development*. 2005;132(12):2917–27. https://doi.org/10.1242/dev.01871.
36. Allan AM, Liang X, Luo Y, Pak C, Li X, Szulwach KE, et al. The loss of methyl-CpG binding protein 1 leads to autism-like behavioral deficits. *Hum Mol Genet*. 2008;17(13):2047–57. https://doi.org/10.1093/hmg/ddn102.
37. Prasad A, Merico D, Thiruvahindrapuram B, Wei J, Lionel AC, Sato D, et al. A Discovery resource of rare copy number variations in individuals with autism spectrum disorder. *G3*. 2012;2:1665–85.
38. Nakano-Kobayashi A, Kasri NN, Newey SE, van Aelst L. The Rho-Linked Mental Retardation Protein OPHN1 Controls Synaptic Vesicle Endocytosis via Endophilin A1. *Curr Biol*. 2009;19(13):1133–9. https://doi.org/10.1016/j.cub.2009.05.022.

Publisher's Note