

Title: Association and epistatic analysis of white matter related genes across the continuum Schizophrenia and Autism spectrum disorders: the joint effect of NRG1-ErbB genes

Authors: C Prats^{1,2,3*}, M Fatjó-Vilas^{1,2,4,*}, MJ Penzol^{2,5}, O Kebir^{7,15}, L Pina-Camacho^{2,5}, D Demontis⁸, B Crespo-Facorro^{2,9}, V Peralta^{10,11}, A González-Pinto^{2,12}, E Pomarol-Clotet^{2,4}, Sergi Papiol^{2,13,14}, M Parellada^{2,5}, MO Krebs^{6,7}, L Fañanás^{1,2,a}

¹ Departament de Biologia Evolutiva, Ecologia i Ciències Ambientals, Facultat de Biologia, Universitat de Barcelona, Spain. Institut de Biomedicina de la Universitat de Barcelona (IBUB), Barcelona, Spain.

² Instituto de Salud Carlos III, Centro de Investigación Biomédica en Red de Salud Mental (CIBERSAM), Madrid, Spain.

³ Institut d'Investigació Biomèdica de Bellvitge, Hospital Duran i Reynals, L'Hospitalet de Llobregat, Barcelona, Spain

⁴ FIDMAG Germanes Hospitalàries Research Foundation, Barcelona, Spain.

⁵ Department of Child and Adolescent Psychiatry, Institute of Psychiatry and Mental Health, Hospital General Universitario Gregorio Marañón, School of Medicine, Universidad Complutense, IISGM, Madrid, Spain.

⁶ University Paris Descartes, Sorbonne Paris Cité, Faculté de Médecine Paris Descartes, Service Hospitalo-Universitaire, Centre Hospitalier Sainte-Anne, Paris, France.

⁷ INSERM, U1266, Laboratory "Pathophysiology of psychiatric disorders", Institute of psychiatry and neurosciences of Paris, Paris, France.

⁸ Department of Biomedicine, Aarhus University, Aarhus, Denmark. The Lundbeck Foundation Initiative for Integrative Psychiatric Research iPSYCH, Aarhus, Denmark.

⁹ University Hospital Virgen del Rocío, IbiS Department of Psychiatry, School of Medicine, University of Sevilla, Santander.

¹⁰ Gerencia de Salud Mental. Servicio Navarro de Salud-Osasunbidea. Pamplona. Navarra

¹¹ Instituto de Investigación Sanitaria de Navarra (IdiSNa). Pamplona. Navarra.

¹² Psychiatry Service, University Hospital of Alava-Santiago. EMBREC. EHU/UPV University of the Basque Country. Kronikgune. Vitoria, Spain.

¹³ Institute of Psychiatric Phenomics and Genomics (IPPG), University Hospital, LMU Munich, Munich, Germany.

¹⁴ Department of Psychiatry, University Hospital, Ludwig Maximilian University, Munich, Germany.

¹⁵ GHU Psychiatrie et Neurosciences de Paris, Paris, France.

*Joint first authorship.

^a Corresponding author at: Departamento de Biología Evolutiva, Ecología y Ciencias Ambientales, Facultad de Biología, Universitat de Barcelona, Av. Diagonal, 643, 08028 Barcelona, Spain. E-mail address: lfananas@ub.edu (L. Fañanás)

Abstract word count: 207

Text word count: 3705

Number of Tables: 3

Number of Figures: 1

Supplementary materials: yes

ABSTRACT

Background: Schizophrenia-spectrum disorders (SSD) and Autism spectrum disorders (ASD) are neurodevelopment disorders that share clinical, cognitive and genetic characteristics, as well as particular white matter (WM) abnormalities. In this study, we aimed to investigate the role of a set of oligodendrocyte/myelin-related (OMR) genes and their epistatic effect on the risk for SSD and ASD.

Methods: We examined 108 SNPs in a set of 22 OMR genes in 1749 subjects divided into three independent samples (187 SSD trios, 915 SSD cases/control and 91 ASD trios). Genetic association and gene-gene interaction analyses were conducted with PLINK and MB-MDR, and permutation procedures were implemented in both.

Results: Some OMR genes showed an association trend with SSD, while after correction, the ones that remained significantly associated were *MBP*, *ERBB3* and *AKT1*. Significant gene-gene interactions were found between: i) *NRG1*MBP* (*perm p-value*=0.002) in the SSD trios sample, ii) *ERBB3*AKT1* (*perm p-value*=0.001) in the SSD case-control sample, and iii) *ERBB3*QKI* (*perm p-value*=0.0006) in the ASD trios sample

Discussion: Our results suggest the implication of OMR genes in the risk for both SSD and ASD and highlight the role of *NRG1* and *ERBB* genes. These are in line with the previous evidence and may suggest pathophysiological mechanisms related to *NRG1/ERBBs* signaling in these disorders.

Keywords: white matter, oligodendrocyte, myelin, schizophrenia, autism spectrum disorders, gene-gene interaction

1. INTRODUCTION

Neurodevelopmental processes are essential for the acquisition and maintenance of the human brain configuration and function efficiency. The interplay of inherent genetic programs with a wide range of environmental exposures determines the brain's diverse configurations and functional attributes. This is achieved by orchestrating different processes involved in, for example, synaptogenesis, synaptic plasticity, myelination or connectivity (Tau and Peterson 2010). These processes, which take place along the developmental timeline and generate the intrinsic human variability, are ultimately related to the ability of the brain to perceive and interpret the world and make adaptive changes (Markham and Greenough, 2005). Deviations in such processes are involved in the etiology of many psychiatric disorders (neurodevelopmental disorders, NDDs) (Paus et al. 2008); however, the specific mechanisms leading to these disorders remain unknown.

Current evidence suggests that a phenotypic continuum links neurodevelopmental disorders such as schizophrenia-spectrum disorders (SSD) and autism spectrum disorders (ASD). Although they are considered separate disease entities, emerging data has increasing recognition of these conditions overlap. Epidemiological studies have revealed the co-occurrence of these disorders in the same subject and different family members (Stahlberg et al. 2004; Daniels et al. 2008; Sullivan et al. 2012, 2013). Concerning clinical characteristics, both SSD and ASD present deficits in social interaction and communication and show impairments in similar cognitive domains (i.e. attention, memory, executive function and social cognition) (Kerns et al. 2008; Sasson et al. 2011; Martinez et al. 2019). At a neurobiological level, subjects with SSD and ASD show microscopic and macroscopic evidence of brain disruption in overlapping areas (de Lacy and King 2013). For example, recent studies focusing on induced pluripotent stem cells-derived neural cells from SSD and ASD patients support common alterations in glutamatergic synapse formation and function in both disorders (Habela et al. 2016). Also, from neuroimaging approaches, different studies have reported WM alterations in SSD and ASD (Dennis and Thompson 2013; Wheeler and Voineskos 2014). The observed differences in both disorders as compared to healthy subjects include structural connectivity deficits regarding long-distance and interhemispheric bundles in the corpus callosum and the superior longitudinal, the inferior fronto-occipital and the inferior longitudinal fasciculi (Ford *et al*, 2002; Mueller et al., 2012; Karlsgodt 2020; Katz et al. 2016).

The importance of intact WM for brain functionality and the implication of WM alterations in numerous psychiatric disorders has been extensively described (e.g., Catani and ffytche 2005; Konrad and

Winterer 2007). At the histological level, WM is composed of bundles of myelinated axons connecting different parts of grey matter regions and contains the connections between specialized processing regions. Then, the oligodendrocytes, the brain cell type responsible for the synthesis of the myelin sheath that supports and isolate neuronal axons, appear to have a critical role in the white-grey matter cooperation essential for the correct functioning of the brain. Interestingly, neuropathological studies suggest that oligodendrocytes may be altered in SZ (Uranova et al. 2001) and ASD (Morgan et al. 2010; Sheikh et al. 2012).

Regarding the genetic background of SZ and ASD, several studies have confirmed cross-diagnosis genetic correlations (Lee et al. 2013; Hammerschlag et al. 2019). Interestingly, recent molecular studies have reported evidence for overlap in SZ GWAS regions and those with de novo non-synonymous mutations in ASD (Ripke et al. 2014). Also, transcriptome data has allowed identifying gene-sets with shared dysregulated expression in cortical brain regions of individuals affected with SZ or ASD (Guan et al. 2019; Hammerschlag et al. 2019).

The candidate gene pathway and epistatic approaches have gained attention as a strategy to deep into the overlapping mechanisms in these disorders. Multiple lines of genetic evidence derived from candidate gene association studies or genome-wide approaches support the association of oligodendrocyte/myelination related (OMR) genes with SZ (Karoutzou et al. 2008; Roussos and Haroutunian 2014). OMR genes are mainly expressed in oligodendrocytes and are involved in myelination (Davis et al. 2003), trophic support to the axon (Segal et al. 2007) and axon-glia interactions (Pernet et al. 2008). For example, the *NRG1* gene, which has a central role in cortico-cortical myelination during neurodevelopment (Chen et al. 2006), has been repeatedly associated with SZ (Agim et al. 2013; Mostaid et al. 2017).

Additionally, the SZ GWAS published by Ripke et al. (2014) reported that although no OMR gene reached genome-wide significance, a set of them showed suggestive association (*ERBB4*, *NRG1*, *ANK3*). Moreover, WM volume was highlighted as a candidate endophenotype linked to SZ-associated SNPs (Terwisscha Van Scheltinga et al. 2013). Remarkably, these authors reported an effect of GWAS-identified SZ risk variants, when compiled as polygenic risk scores, on WM volumes in SZ. However, these results were not replicated in an independent study, which suggests the need for further analyses of WM genes' role in the polygenic architecture of mental disorders (Papiol et al. 2014).

Also, astrocyte and oligodendrocyte gene-sets have been associated with an increased risk for SZ (Duncan et al. 2014; Goudriaan et al. 2014), suggesting that genetic alterations underlie specific glial

cell type functions which increase susceptibility to SZ. However, other studies did not replicate such effects (O'dushlaine et al. 2015).

About ASD, gene network analyses have also shed light on the putative role of these gene-sets. In detail, Li et al. (2014) identified a functional module linked to autism, revealing a subset of genes expressed in oligodendrocytes at the corpus callosum, a structure associated with ASD (Wolff et al., 2015). Accordingly, new genetic studies involving WM related genes and their genetic interactions could provide new insights into the molecular and cellular mechanisms underlying each disorder and the shared etiological pathways between both disorders.

Based on the above mentioned, we hypothesized that WM-related genes and their epistatic effects would contribute to both SSD and ASD. To test this hypothesis, we first explored the relationship between genetic variability among a selected set of 22 WM related genes and the risk for developing SSD or ASD. Secondly, we analyzed second-order gene-gene interactions.

2. METHODS

2.1 Sample description

Patients were included in this study based on the presence of a diagnosis of SSD or ASD. When possible, healthy patients' parents were also recruited in order to have a trio-based sample. A sample of healthy subjects was also included.

The sample was drawn and coordinated from three centers: i) University of Barcelona (Coordination Center of a Spanish multi-center enrollment including samples from Complejo Hospitalario de Navarra, Hospital Santiago Apostol Vitoria, Hospital de Valdecilla Santander and Hospital Benito Menni Sant Boi Barcelona), ii) Hospital Universitario Gregorio Marañón (Madrid, Spain), iii) Centre Hospitalier Sainte Anne (Paris, France).

The global sample (1749 subjects) comprised:

- i) Sample A - SSD case-parents' trios sample -- 561 subjects in 187 nuclear families (187 cases and 374 parents).
- ii) Sample B - SSD case-control sample: 915 subjects (451 cases and 464 controls).
- iii) Sample C - ASD case-parents' trios sample -- 273 subjects in 91 nuclear families (91 cases and 182 parents).

Probands and parents were evaluated by trained psychiatrists using: i) the Comprehensive Assessment of Symptoms and History (CASH), Structured Clinical Interview for DSM Disorders and/or the Diagnostic Interview for Genetic Studies (DIGS) (Sample A and B), ii) the Autism Diagnostic Interview (Sample C).

The DSM-IV-TR diagnoses distribution for the SSD family-based sample (Sample A) were as follows: schizophrenia (54%), schizophreniform disorder (9.9%), psychotic disorder not otherwise specified (6.8%), schizoaffective disorder (6.8%), schizotypal personality disorder (0.5%), other psychotic disorders/NOS (0.6%) and bipolar disorder with psychosis features (20.9%). For the SSD case-control sample (Sample B): schizophrenia (68.7%), schizophreniform disorder (11.5%), psychotic disorder not otherwise specified (10.4%), schizoaffective disorder (5.7%), schizotypal personality disorder (0.4%), other psychotic disorders/NOS (10.4%) and bipolar disorder with psychosis features (3.3%).

Controls were recruited from university students and staff and their acquaintances, plus independent sources in the community. They were interviewed and excluded if they reported a history of mental illness and/or treatment with psychotropic medication.

The exclusion criteria included: major medical illnesses that could affect brain function, substance-induced psychotic disorder, neurological conditions, history of head trauma with loss of consciousness, and moderate or severe mental retardation.

All participants were of Caucasian origin, thereby reducing the possibility of confounding genetic differences by population stratification.

All participants provided written consent after being informed of the study procedures and implications. The study was performed following the institutions' guidelines and approved by the local ethics committee of each participating center.

2.2 Gene selection

The selection of the 22 white matter related genes was based on: i) their direct biological implication with WM structure/function, ii) their relationship with WM related biological pathways and also, iii) their previously reported association with both the risk for SSD/ASD and WM alterations.

According to its function, the selected genes are related to i) myelin structure (*MAG*, *MBP*, *PLP1*, *MOG*, *CNP*, *PTEN*, *AKT1* and *FYN*), ii) oligodendrocyte development (*QKI*), iii) synaptic plasticity and axonal regeneration (*OMG*, *CDH10*, *MAG*), iv) transcription and signaling factors (*OLIG2*, *NRG1*, *ErbB2*, *ErbB3*, *ErbB4*), v) cell adhesion molecules and receptors (*NRXN1*, *CNTNAP2* and *SPON1*), vi) calcium channels (*CACNA1C*, *CACNB2*) and coding for zinc finger binding protein (*ZNF804A*).

Gene interaction graph network from STRING v.9 was implemented (Franceschini et al. 2013). STRING integrates protein-protein interactions from literature curation, computationally predicted interactions and interactions transferred from model organisms based on orthology. As visible in Figure 1, most of the proteins included in our study are linked and form an interaction network, supporting the gene-gene interaction approach.

2.3 Genotyping

Genomic DNA was extracted from peripheral blood cells or buccal mucosa using standard methods. The genotyping was conducted at the Genomic Service of the Spanish National Cancer Research Center (CEGEN-CNIO) by using the *Open Array® Genotyping System of Applied Biosystems*.

The molecular genetics data for the sample include 108 Single Nucleotide Polymorphisms (SNPs) among the 22 selected genes (Table 1). The optimal set of SNPs that contained maximum information about surrounding variants was selected using SYSNPs (<http://www.sysnps.org/>) with a minor allele frequency (MAF) > 5%, using the pairwise option tagger (threshold of $r^2=0.8$). These SNPs were also selected due to either previous associated findings or functional implications. FASTSNP (function analysis and selection tool for single nucleotide polymorphisms) was used to identify and prioritize high-risk SNPs according to their phenotypic risks and putative functional effects (Yuan et al. 2006: <http://fastsnp.ibms.sinica.edu.tw>).

After genotyping, standard data cleaning and quality control were performed (see Supplementary Material).

2.4 Statistical methods

2.4.1 Generating case/pseudo-controls data

In the family-based samples (Sample A & C), pseudo-controls were generated (with PLINK --tucc command). This approach consists of generating pseudo-controls using the parent's untransmitted alleles, thus creating a matched case-control design where the observed case is compared to all possible genotypic combinations that could have arisen from the parental mating type (Cordell et al. 2004). For any single variant, there are three alternative genotypes for pseudo-controls that could have been transmitted to the case (case:pseudo-controls ratio is 1:3). Case/pseudocontrols files were used to conduct both association and epistasis analyses.

2.4.2 Association analyses

Association analyses were conducted using PLINK 1.07 to explore the relationship between our 22 WM related genes and the risk for developing SSD/ASD.

In those analyses with individuals from different sampling sites (Spain and France), Cochran-Mantel-Haenszel (CMH) and Breslow-Day (BD) tests were used for association analyses and heterogeneity testing, respectively. The Cochran-Mantel-Haenszel (CMH) association test controls for the potential confounding of population enrollment. This test allows the comparison of SNP allelic frequencies between groups while controlling for the country of collection. In addition, the asymptotic P-value of the Breslow-Day test was used to analyze the heterogeneity of the odds ratios (ORs). After Bonferroni correction, none of the analysis reached statistical significance, meaning that any SNP was excluded due to population differences.

CMH and BD were applied to conduct: i) Association analysis in NDDs sample (Sample A + C), ii) Association analysis in SSD (Sample A and B, separately). Since sample C was recruited from one centre, allelic association analysis was conducted directly. All the analyses were adjusted by sex. Holm-Bonferroni correction was applied to control for multiple testing in all association analyses.

2.4.3 Gene-gene interaction analysis

To capture second-order SNP-SNP interactions, we used the model-based multifactor dimensionality reduction (MB-MDR) method as implemented in mbmdr R package (Calle et al. 2010). It is an extension of the multifactor dimensionality reduction (MDR) method in which risk categories are defined using a regression model that also allows adjustments for main effects and covariates. By this approach, first, logistic regressions analyses are performed to determine the nine possible genotypic combinations as high (*H*), low (*L*), or no risk (*0*). Then, genotypes of the same risk category are merged, and two Wald statistics (*WH* and *WL*, one for each risk) are calculated. Finally, the significance of a specified model is assessed through a permutation test on the maximum Wald statistic, implemented in the function mbmdr.PermTest.

Gene-gene interactions were tested in the whole group of patients affected by NDDs (Sample A and C) and also in each sample separately. The permutation procedure (10,000 permutations) was applied to the interaction models with the threshold set at $p < 0.05$. In all interaction models (Table 3), sex was added as a covariate. In addition, when necessary (Sample A+C and Sample B), the analyses were repeated including the center as a covariate.

3. RESULTS

3.1 Sample Description

Table 2 shows the characteristics of the three independent samples of our study. As sex distribution between groups (patients vs relatives and patients vs controls) showed significant differences, sex was added as a covariate in all the analyses.

INSERT TABLE 2 HERE

3.2 Association analyses

First, when sample A and C were analyzed together (considering all patients to belong to the Neurodevelopmental Disorders group), CMH test showed a non-significant association for all SNPs (data not shown).

Second, to explore diagnosis-specific effects, we conducted family-based association analyses separately in Sample A and C. In SSD families (Sample A), the associated SNPs implicated six genes: *ERBB4*, *ERBB2*, *MOG*, *NRG1*, *MBP* and *MAG* (uncorrected p-values $p < 0.04$; Table S1). After B-H correction *MBP* gene remained significant. Analyses in ASD families (Sample C) showed a trend association for SNPs implicating four genes: *MOG*, *QKI*, *MAG* and *MBP* (uncorrected p-value $p < 0.05$; Table S2). However, after B-H correction, none of these gene associations remained significant.

Third, in the SSD case-control sample (Sample B), the association analyses detected four genes: *ERBB3*, *AKT1*, *CNP* and *PLP1* (uncorrected p-values $p < 0.04$; Table S3). After B-H correction *ERBB3* and *AKT1* genes remained significant.

3.3 Epistatic analysis effect of WM related genes

When the analysis included the whole group of NDDs patients (Sample A + C), the following epistasis was found: *NRG1* (rs6989777)**ERBB4* (rs707284), *Perm p-value*=0.0003. [This interaction reflects that some allelic combinations of these two SNPs are differentially distributed in patients compared to healthy subjects.](#)

Afterwards, we conducted the epistatic analyses in each sample separately (Table 3). In SSD (Sample A), a significant interaction was detected involving *MBP* and *NRG1* genes (*MBP* (rs9966986)**NRG1* (rs3802160), *Perm p-value*=0.002); while in ASD (Sample C) the detected interaction involved *ERBB3* and *QKI* genes (*ERBB3*(rs2271194)**QKI*(rs7772756), *Perm p-value*=0.0006). The analyses in SSD case-

control (Sample B) showed a significant joint effect of *AKT1* and *ERBB3* genes (*AKT1(rs2494746)*ERBB3(rs7971751)*, Perm *p-value*=0.001).

INSERT TABLE 3 HERE

4. DISCUSSION

The present study has focused on exploring the role of a set of white matter related genes, individually or in interaction, in the risk of developing two psychiatric disorders with a neurodevelopmental component, schizophrenia and autism spectrum disorders (SSD and ASD). The inclusion of the two conditions and the cross-disorders analyses combined with the diagnostic-specific ones represent a novel approach that has allowed studying a targeted gene-set across the continuum of these neurodevelopmental disorders.

We first developed single-gene based analyses in three samples, including individuals diagnosed with SSD or ASD and healthy subjects (relatives or non-related). On the one hand, our results suggest that some of the analyzed WM genetic risk variants seem to be shared across SSD-ASD while others seem to be diagnosis-specific. For example, *MBP* or *MAG* genes appear to be associated with either SSD or ASD, while *AKT1* and *NRG1* (and its receptors, *ERBB2*, *ERBB3* and *ERBB4*) emerge only in SSD based samples. These results should be interpreted cautiously. First, because some do not remain significant after multiple testing correction; second, because some are not replicated in the two SSD samples; and third, because there is no ASD replication sample. However, our results gain interest when considering the gene-gene interactions analyzed afterwards.

In this regard, epistasis analyses identified significant effects that include the genes previously highlighted by the single-gene approach. While the specific genes involved in the second-order interactions differ across samples, it is remarkable that in the three samples as well as in the analyses combining both disorders, the identified epistatic effects include the interaction of *NRG1* or ErbB receptors genes with OMR genes. These differences across samples could be attributed to diagnostic specificities and/or to design effects (case-control vs family-based); however, our convergent results suggest the main role of Nrg-ErbB signaling across the continuum of these neurodevelopment disorders.

On the one hand, the *ERBB3* gene appears to interact with *AKT1* (in SSD case-control sample) and with *QKI* (in ASD trio sample). Although these interactions have not been previously described, it is

interesting to note that *QKI* and *AKT1* genes are related to myelination pathways (Åberg et al. 2006; Flores et al. 2008; Lee 2009). In this line, disruption of *QKI* gene expression has been related to downstream consequences on oligodendrocyte development and myelination, eventually increasing the susceptibility for suffering psychiatric disorders (Lee et al, 2009). Of note, *QKI* expression changes have been associated with the treatment with typical neuroleptics in SZ patients (Åberg et al. 2006). Also, Akt signaling has emerged as a fundamental player in both peripheral and central nervous systems myelination (Figlia et al. 2017) due to its implication in regulating several processes during the development of myelinating Schwann cells and oligodendrocytes (Norrmen and Suter 2013).

On the other hand, a combined effect between *NRG1* and *ERBB4* has been observed in the whole group of Neurodevelopment disorders. Genetic association studies have identified *NRG1* and *ERBB4* as SZ risk genes, and altered *NRG-ERBB4* signaling has been associated with positive, negative, and cognitive symptoms (Stefansson et al. 2002; Li et al. 2006; Nicodemus et al. 2006; Harrison and Law 2006; Yin et al. 2018). In addition, *NRG1xERBB4* interaction has already been described in SZ (Norton et al. 2006), but, to our knowledge, no study has previously studied it across neurodevelopment disorders, further suggesting that.

In favor of the role of Nrg-ErbB signaling across the continuum of neurodevelopment disorders, recent studies have revealed complex Nrg/ErbB signaling networks that regulate the assembly of neural circuitry, myelination, neurotransmission, and synaptic plasticity (Mei and Nave 2014). Moreover, evidence indicates that deviances in Nrg/ErbB signaling impair brain functions critically involved in neural development, including circuitry generation, peripheral myelination, neurotransmission and homeostasis of CNS synaptic functions (Mei and Nave 2014). In this sense, OMR and Nrg/ErbB signaling pathways are highlighted as therapeutic targets for neuropsychiatric diseases (Wehr et al. 2017).

In this support, the disruption of the *NRG1-ErbB4* pathway in oligodendrocytes of animal models has been observed to lead to an alteration of the myelin sheath of the white matter tracts, reduced conduction velocity and cognitive changes (Roy et al. 2007). Moreover, postmortem human-based studies have also reported that *NRG1* expression is deregulated and *ERBB4* hyperphosphorylated in the dorsolateral prefrontal cortex and hippocampus in SZ (Hashimoto et al. 2004; Law et al. 2006; Hahn et al. 2006; Weickert et al. 2012), which can also be related to a hypothetical unbalanced mechanism of expression underlying *NRG1-ErbB4* interaction. Then, given the role of *NRG1* and ErbB receptors in oligodendrocytes development, alterations in the corresponding genes could impact the oligodendrocytes functionality and thus, affect the white-grey matter relationship.

Even though the two SNPs detected in the second-order interaction are intronic, it is of note that the vast majority of the genome has gene regulatory properties (Bernstein et al. 2012), in which intronic and intergenic variants are involved. The impact of the non-coding variants captured by this interaction *NRG1xERBB4* (rs6989777 x rs707284) can be evaluated using HaploReg (Ward and Kellis 2012). This is a tool that uses LD information from the 1000 Genomes Project to provide data on the predicted chromatin state of the queried SNPs, their sequence conservation mammals, and their effect on regulatory motifs. In this case, both SNPs are associated with putative changes in regulatory motifs (multi-tissue eQTL data). For example, *NRG1*-rs6989777, which is also associated with SSD risk (single-gene approach), is in an intronic region, and it is predicted to alter several motifs that overlap the recognition sequences of transcription factors such as AP-1, CTCF and Maf. In this sense, AP-1 and CTCF have been proposed as key players in maintaining a chromatin conformation of gene regulatory elements (Park et al. 2012). Interestingly, there is also evidence that this SNP is included in one of the top identified genes through polygenic scoring and pathway analyses in SZ (Ayalew et al. 2012). Besides the putative regulatory properties of the identified SNPs, it has also to be considered that the detected effects could reflect the involvement of other SNPs in linkage disequilibrium.

Some limitations of this study must be acknowledged. The polygenic nature of mental disorders, the relatively small sample size and the minor effect of the common genetic variants limit the power of our research. Also, the clinical heterogeneity derived from the inclusion based on diagnosis instead of more specific clinical traits restricts our analyses and could be potentially associated with error I type errors. Accordingly, as described in detail in the methodology, multiple testing corrections have been applied in each analysis. Nevertheless, although such procedures have been used, if multiple testing is addressed taking together all the analyses, probably not all the findings would remain significant. Then, further studies are required to validate our findings and determine the biological mechanisms underlying the detected gene interactions among the WM related genes.

Overall, our results contribute from a biological approach to identifying putative genetic mechanisms involved in SSD and ASD and suggest the potential role for the genes involved in the relationship between OMR and *Nrg-ErbB* signaling. These findings help in the detection and characterization of the biological pathways that underpin these disorders and add interest in investigating the interactions of genes to explain a substantial shared component of the risk.

Conflicts of interest: All authors report no conflict of interest.

Acknowledgements:

We are deeply grateful to all the participants whose generosity made this work possible. We also sincerely acknowledge the psychiatrists, psychologists, and mental health staff from all clinical and research centers who collaborated in this study. We also thank Anna Valldeperas for her assistance with the molecular laboratory tasks.

This study was supported by: i) ERA-NET-NEURON-PIM2010ERN, ANR-2010-002-01; ii) the Comissionat per a Universitats I Recerca del DIUE (2017SGR1577); iii) travel grant from Aarhus University graduate school; iii) an Ajuts de Personal Investigador Predoctoral en Formació (APIF-Universitat de Barcelona) to C Prats, and iv) a Miguel Servet contract (Instituto de Salud Carlos III) to M Fatjó-Vilas (CP20/00072).

References

- Åberg K, Saetre P, Lindholm E, et al (2006) Human QKI, a new candidate gene for schizophrenia involved in myelination. *Am J Med Genet - Neuropsychiatr Genet* 141 B:84–90. <https://doi.org/10.1002/ajmg.b.30243>
- Agim ZS, Esendal M, Briollais L, et al (2013) Discovery, Validation and Characterization of Erbb4 and Nrg1 Haplotypes Using Data from Three Genome-Wide Association Studies of Schizophrenia. *PLoS One* 8:. <https://doi.org/10.1371/journal.pone.0053042>
- Ayalew M, Le-Niculescu H, Levey DF, et al (2012) Convergent functional genomics of schizophrenia: From comprehensive understanding to genetic risk prediction. *Mol Psychiatry* 17:887–905. <https://doi.org/10.1038/mp.2012.37>
- Bernstein BE, Birney E, Dunham I, et al (2012) An integrated encyclopedia of DNA elements in the human genome. *Nature* 489:57–74. <https://doi.org/10.1038/nature11247>
- Calle ML, Urrea V, Malats N, van Steen K (2010) Mbmdr: An R package for exploring gene-gene interactions associated with binary or quantitative traits. *Bioinformatics* 26:2198–2199. <https://doi.org/10.1093/bioinformatics/btq352>
- Catani M, ffytche DH (2005) The rises and falls of disconnection syndromes. *Brain* 128:2224–2239. <https://doi.org/10.1093/brain/awh622>
- Chen S, Velardez MO, Warot X, et al (2006) Neuregulin 1-erbB Signaling Is Necessary for Normal Myelination and Sensory Function. *J Neurosci* 26:3079–3086. <https://doi.org/10.1523/JNEUROSCI.3785-05.2006>
- Cordell HJ, Barratt BJ, Clayton DG (2004) Case/pseudocontrol analysis in genetic association studies: A unified framework for detection of genotype and haplotype associations, gene-gene and gene-environment interactions, and parent-of-origin effects. *Genet Epidemiol* 26:167–185. <https://doi.org/10.1002/gepi.10307>
- Daniels JL, Forssen U, Hultman CM, et al (2008) Parental psychiatric disorders associated with autism spectrum disorders in the offspring. *Pediatrics* 121:e1357-62. <https://doi.org/10.1542/peds.2007-2296>
- Davis KL, Stewart DG, Friedman JI, et al (2003) White matter changes in schizophrenia: evidence for myelin-related dysfunction. *Arch Gen Psychiatry* 60:443–56. <https://doi.org/10.1001/archpsyc.60.5.443>
- de Lacy N, King BH (2013) Revisiting the relationship between autism and schizophrenia: toward an integrated neurobiology. *Annu Rev Clin Psychol* 9:555–87. <https://doi.org/10.1146/annurev-clinpsy-050212-185627>
- Dennis EL, Thompson PM (2013) Typical and atypical brain development: A review of neuroimaging studies. *Dialogues Clin Neurosci* 15:359–384

- Duncan LE, Holmans PA, Lee PH, et al (2014) Pathway analyses implicate glial cells in schizophrenia. *PLoS One*. <https://doi.org/10.1371/journal.pone.0089441>
- Figlia G, Norrmén C, Pereira JA, et al (2017) Dual function of the PI3K-Akt-mTORC1 axis in myelination of the peripheral nervous system. *Elife*. <https://doi.org/10.7554/elife.29241>
- Flores AI, Narayanan SP, Morse EN, et al (2008) Constitutively Active Akt Induces Enhanced Myelination in the CNS. *J Neurosci* 28:7174–7183. <https://doi.org/10.1523/JNEUROSCI.0150-08.2008>
- Ford JM, Mathalon DH, Whitfield S, et al (2002) Reduced communication between frontal and temporal lobes during talking in schizophrenia. *Biol Psychiatry* 51:485–492. [https://doi.org/10.1016/S0006-3223\(01\)01335-X](https://doi.org/10.1016/S0006-3223(01)01335-X)
- Franceschini A, Szklarczyk D, Frankild S, et al (2013) STRING v9.1: protein-protein interaction networks, with increased coverage and integration. *Nucleic Acids Res* 41:D808-15. <https://doi.org/10.1093/nar/gks1094>
- Friston KJ, Frith CD (1995) Schizophrenia: a disconnection syndrome? *Clin. Neurosci.* 3:89–97
- Goudriaan A, De Leeuw C, Ripke S, et al (2014) Specific glial functions contribute to Schizophrenia susceptibility. *Schizophr Bull* 40:925–935. <https://doi.org/10.1093/schbul/sbt109>
- Guan J, Cai JJ, Ji G, Sham PC (2019) Commonality in dysregulated expression of gene sets in cortical brains of individuals with autism, schizophrenia, and bipolar disorder. *Transl Psychiatry*. <https://doi.org/10.1038/s41398-019-0488-4>
- Habela CW, Song H, Ming G li (2016) Modeling synaptogenesis in schizophrenia and autism using human iPSC derived neurons. *Mol. Cell. Neurosci.*
- Hahn C-G, Wang H-Y, Cho D-S, et al (2006) Altered neuregulin 1-erbB4 signaling contributes to NMDA receptor hypofunction in schizophrenia. *Nat Med* 12:824–828. <https://doi.org/10.1038/nm1418>
- Hammerschlag AR, de Leeuw CA, Middeldorp CM, Polderman TJC (2019) Synaptic and brain-expressed gene sets relate to the shared genetic risk across five psychiatric disorders. *Psychol Med*. <https://doi.org/10.1017/s0033291719001776>
- Harrison PJ, Law AJ (2006) Neuregulin 1 and Schizophrenia: Genetics, Gene Expression, and Neurobiology. *Biol. Psychiatry* 60:132–140
- Hashimoto R, Straub RE, Weickert CS, et al (2004) Expression analysis of neuregulin-1 in the dorsolateral prefrontal cortex in schizophrenia. *Mol Psychiatry* 9:299–307. <https://doi.org/10.1038/sj.mp.4001434>
- Karoutzou G, Emrich HM, Dietrich DE (2008) The myelin-pathogenesis puzzle in schizophrenia: a literature review. *Mol Psychiatry* 13:245–260. <https://doi.org/10.1038/sj.mp.4002096>
- Katz J, d'Albis MA, Boisgontier J, et al (2016) Similar white matter but opposite grey matter changes in schizophrenia and high-functioning autism. *Acta Psychiatr Scand* 134:31–39. <https://doi.org/10.1111/acps.12579>
- Kerns JG, Nuechterlein KH, Braver TS, Barch DM (2008) Executive Functioning Component Mechanisms and Schizophrenia. *Biol. Psychiatry* 64:26–33
- Konrad A, Winterer G (2007) Disturbed Structural Connectivity in Schizophrenia Primary Factor in Pathology or Epiphenomenon? *Schizophr Bull* 34:72–92. <https://doi.org/10.1093/schbul/sbm034>
- Law AJ, Lipska BK, Weickert CS, et al (2006) Neuregulin 1 transcripts are differentially expressed in schizophrenia and regulated by 5' SNPs associated with the disease. *Proc Natl Acad Sci U S A* 103:6747–52. <https://doi.org/10.1073/pnas.0602002103>
- Lee PR (2009) Regulation of myelin genes implicated in psychiatric disorders by functional activity in axons. *Front Neuroanat*. <https://doi.org/10.3389/neuro.05.004.2009>
- Lee SH, Ripke S, Neale BM, et al (2013) Genetic relationship between five psychiatric disorders estimated from genome-wide SNPs. *Nat Genet* 45:984–94. <https://doi.org/10.1038/ng.2711>
- Li D, Collier DA, He L (2006) Meta-analysis shows strong positive association of the neuregulin 1 (NRG1) gene with schizophrenia. *Hum Mol Genet*. <https://doi.org/10.1093/hmg/ddl122>
- Li J, Shi M, Ma Z, et al (2014) Integrated systems analysis reveals a molecular network underlying autism spectrum disorders. *Mol Syst Biol* 10:774. <https://doi.org/10.15252/msb.20145487>

- Marín O (2016) Developmental timing and critical windows for the treatment of psychiatric disorders. *Nat Med* 22:1229–1238. <https://doi.org/10.1038/nm.4225>
- Markham JA, Greenough WT (2005) Experience-driven brain plasticity: beyond the synapse. *Neuron Glia Biol* 1:351. <https://doi.org/10.1017/S1740925X05000219>
- Martinez G, Mosconi E, Daban-Huard C, et al (2019) “A circle and a triangle dancing together”: Alteration of social cognition in schizophrenia compared to autism spectrum disorders. *Schizophr Res*. <https://doi.org/10.1016/j.schres.2019.05.043>
- Mei L, Nave KA (2014) Neuregulin-ERBB signaling in the nervous system and neuropsychiatric diseases. *Neuron* 83:27–49
- Morgan JT, Chana G, Pardo CA, et al (2010) Microglial activation and increased microglial density observed in the dorsolateral prefrontal cortex in autism. *Biol Psychiatry* 68:368–376. <https://doi.org/10.1016/j.biopsych.2010.05.024>
- Mostaid MS, Mancuso SG, Liu C, et al (2017) Meta-analysis reveals associations between genetic variation in the 5' and 3' regions of Neuregulin-1 and schizophrenia. *Transl Psychiatry* 7:. <https://doi.org/10.1038/tp.2016.279>
- Nicodemus KK, Luna A, Vakkalanka R, et al (2006) Further evidence for association between ErbB4 and schizophrenia and influence on cognitive intermediate phenotypes in healthy controls [1]. *Mol. Psychiatry*
- Norrmén C, Suter U (2013) Akt/mTOR signalling in myelination. *Biochem Soc Trans* 41:944–950. <https://doi.org/10.1042/BST20130046>
- Norton N, Moskvina V, Morris DW, et al (2006) Evidence that interaction between neuregulin 1 and its receptor erbB4 increases susceptibility to schizophrenia. *Am J Med Genet - Neuropsychiatr Genet* 141 B:96–101. <https://doi.org/10.1002/ajmg.b.30236>
- O'dushlaine C, Rossin L, Lee PH, et al (2015) Psychiatric genome-wide association study analyses implicate neuronal, immune and histone pathways. *Nat Neurosci*. <https://doi.org/10.1038/nn.3922>
- Papiol S, Mitjans M, Assogna F, et al (2014) Polygenic determinants of white matter volume derived from GWAS lack reproducibility in a replicate sample. *Transl Psychiatry* 4:e362. <https://doi.org/10.1038/tp.2013.126>
- Park J, Sarode VR, Euhus D, et al (2012) Neuregulin 1-HER axis as a key mediator of hyperglycemic memory effects in breast cancer. *Proc Natl Acad Sci U S A* 109:21058–63. <https://doi.org/10.1073/pnas.1214400109/-DCSupplemental.www.pnas.org/cgi/doi/10.1073/pnas.1214400109>
- Paus T, Keshavan M, Giedd JN (2008) Why do many psychiatric disorders emerge during adolescence? *Nat Rev Neurosci* 9:947–957. <https://doi.org/10.1038/nrn2513>
- Pernet V, Joly S, Christ F, et al (2008) Nogo-A and Myelin-Associated Glycoprotein Differently Regulate Oligodendrocyte Maturation and Myelin Formation. *J Neurosci* 28:7435–7444. <https://doi.org/10.1523/JNEUROSCI.0727-08.2008>
- Ripke S, Neale BM, Corvin A, et al (2014) Biological insights from 108 schizophrenia-associated genetic loci. *Nature* 511:421–427. <https://doi.org/10.1038/nature13595>
- Roussos P, Haroutunian V (2014) Schizophrenia: susceptibility genes and oligodendroglial and myelin related abnormalities. *Front Cell Neurosci* 8:. <https://doi.org/10.3389/fncel.2014.00005>
- Roy K, Murtie JC, El-Khodori BF, et al (2007) Loss of erbB signaling in oligodendrocytes alters myelin and dopaminergic function, a potential mechanism for neuropsychiatric disorders. *Proc Natl Acad Sci* 104:8131–8136. <https://doi.org/10.1073/pnas.0702157104>
- Sasson NJ, Pinkham AE, Carpenter KLH, Belger A (2011) The benefit of directly comparing autism and schizophrenia for revealing mechanisms of social cognitive impairment. *J Neurodev Disord* 3:87–100. <https://doi.org/10.1007/s11689-010-9068-x>
- Segal D, Koschnick JR, Slegers LHA, Hof PR (2007) Oligodendrocyte pathophysiology: a new view of schizophrenia. *Int J Neuropsychopharmacol* 10:503. <https://doi.org/10.1017/S146114570600722X>
- Sheikh A, Wen G, Cao F, et al (2012) Alteration of astrocytes and Wnt/beta-catenin signaling in the frontal cortex of autistic subjects. *J Neuroinflammation* 9:223. <https://doi.org/10.1186/1742-2094-9-223>
- Stahlberg O, Soderstrom H, Rastam M, Gillberg C (2004) Bipolar disorder, schizophrenia, and other psychotic disorders in adults with childhood onset AD/HD and/or autism spectrum disorders. *J Neural Transm* 111:891–902.

<https://doi.org/10.1007/s00702-004-0115-1>

- Stefansson H, Petursson H, Sigurdsson E, et al (2002) Neuregulin 1 and Susceptibility to Schizophrenia. *Am J Hum Genet* 71:877–892. <https://doi.org/10.1086/342734>
- Sullivan PF, Magnusson C, Reichenberg A, et al (2012) Family History of Schizophrenia and Bipolar Disorder as Risk Factors for Autism. *Arch Gen Psychiatry* 69:. <https://doi.org/10.1001/archgenpsychiatry.2012.730>
- Sullivan S, Rai D, Golding J, et al (2013) The association between autism spectrum disorder and psychotic experiences in the Avon Longitudinal Study of Parents and Children (ALSPAC) birth cohort. *J Am Acad Child Adolesc Psychiatry* 52:. <https://doi.org/10.1016/j.jaac.2013.05.010>
- Tau GZ, Peterson BS (2010) Normal Development of Brain Circuits. *Neuropsychopharmacol Rev* 35:115:147–168. <https://doi.org/10.1038/npp.2009.115>
- Terwisscha Van Scheltinga AF, Bakker SC, Van Haren NEM, et al (2013) Genetic schizophrenia risk variants jointly modulate total brain and white matter volume. *Biol Psychiatry* 73:525–531. <https://doi.org/10.1016/j.biopsych.2012.08.017>
- Uranova N, Orlovskaya D, Vikhрева O, et al (2001) Electron microscopy of oligodendroglia in severe mental illness. *Brain Res Bull* 55:597–610
- Wehr MC, Hinrichs W, Brzózka MM, et al (2017) Spironolactone is an antagonist of NRG 1- ERBB 4 signaling and schizophrenia-relevant endophenotypes in mice . *EMBO Mol Med*. <https://doi.org/10.15252/emmm.201707691>
- Weickert CS, Tiwari Y, Schofield PR, et al (2012) Schizophrenia-associated HapICE haplotype is associated with increased NRG1 type III expression and high nucleotide diversity. *Transl Psychiatry*. <https://doi.org/10.1038/tp.2012.25>
- Wheeler AL, Voineskos AN (2014) A review of structural neuroimaging in schizophrenia: from connectivity to connectomics. *Front Hum Neurosci* 8:653. <https://doi.org/10.3389/fnhum.2014.00653>
- Yin L, Cheung EFC, Chen RYL, et al (2018) Leveraging genome-wide association and clinical data in revealing schizophrenia subgroups. *J Psychiatr Res* 106:106–117. <https://doi.org/10.1016/j.jpsychires.2018.09.010>
- Yuan HY, Chiou JJ, Tseng WH, et al (2006) FASTSNP: An always up-to-date and extendable service for SNP function analysis and prioritization. *Nucleic Acids Res* 34:. <https://doi.org/10.1093/nar/gkl236>

TABLES AND FIGURES

Table 1. List of the genotyped 108 SNPs among the 22 white matter related genes. For each gene, the position and the name of the gene are reported.

Gene symbol	Selected SNPs	Gene Description	Gene location
MAG	rs6510476, rs2301600, rs3746248, rs11669734, rs11670792, rs720309, rs720308, rs1034597, rs756796	Myelin associated glycoprotein	19q13.12
MBP	rs1049004, rs470279, rs9675994, rs523243, rs12967023, rs12959006, rs11150994, rs470473, rs871673, rs1026520, rs1629089, rs9966986	Myelin basic protein	18q23
PLP1	rs521895, rs2294152	Proteolipid protein 1	Xq22.2
MOG	rs9468571, rs3130250, rs16895223, rs2535260, rs2857766, rs3130253, rs2071653, rs2535246, rs9257936	Myelin oligodendrocyte glycoprotein	6p22.1
CNP	rs8078650, rs4258677, rs12602950, rs8077391, rs2070106, rs11079028, rs11296	2',3'-cyclic nucleotide 3' phosphodiesterase	17q21.2
PTEN	rs3781195, rs1234220, rs11202596, rs1234219, rs10490920, rs2248293, rs17562384, rs2736627, rs11202607	phosphatase and tensin homolog	10q23.31
AKT1	rs2494732, rs1130233, rs2494734, rs2494739, rs2494743, rs11847866, rs2494746, rs1130214	AKT serine/threonine kinase 1	14q32.33
FYN	rs12191154, rs2344706, rs2237257, rs4947144, rs2301465, rs6901958, rs706915	FYN proto-oncogene, Src family tyrosine kinase	6q21
QKI	rs2784867, rs7772756, rs803612, rs1744926, rs6931903, rs6931903, rs9364692, rs9458853, rs11964059, rs9456869	QKI, KH domain containing RNA binding	6q26
OMG	rs11080149, rs11655238	Oligodendrocyte myelin glycoprotein	17q11.2
CDH10	rs4307059	cadherin 10	5p14.2-p14.1
OLIG2	rs1005573, rs1059004, rs6517137	Oligodendrocyte lineage transcription factor 2	21q22.11
NRG1	rs73235619, rs62510682, rs6994992, rs3802160, rs10503929, rs6989777	neuregulin 1	8p12
ErbB2	rs4252596, rs1565923, rs2952155, rs4252612	erb-b2 receptor tyrosine kinase 2	17q12
ErbB3	rs2271194, rs7971751, rs877636, rs705708, rs10783779, rs773123	erb-b2 receptor tyrosine kinase 3	12q13.2
ErbB4	rs4673628, rs7598440, rs839541, rs839523, rs707284, rs1026882	erb-b2 receptor tyrosine kinase 4	2q34
NRXN1	rs858932, rs1045881	neurexin 1	2p16.3
CNTNAP2	rs2710102	Contactin associated protein-like 2	7q35-q36.1
SPON1	rs2618516	spondin 1	11p15.2
CACNA1C	rs1024582	calcium voltage-gated channel subunit alpha1 C	12p13.33
CACNB2	rs2799573	calcium voltage-gated channel auxiliary subunit beta 2	10p12.33-p12.31
ZNF804A	rs1344706	zinc finger protein 804A	2q32.1

Table 2. Descriptive data of the three samples included in this study.

	Family-based sample		Case /control sample		Family-based sample	
	Sample A: SSD case-parents' trios		Sample B: SSD case-control		Sample C: ASD case-parents' trios	
	SSD Patients (n=451)	Controls (n=464)	SSD Patients (n=187)	1 st degree healthy relatives (n= 374)	ASD Patients (n= 91)	1 st degree healthy relatives (n= 182)
Age at interview	31.11 (11.39)	33.72 (11.76)	23.17 (7.81)	42.50 (1.05)	14.13 (7.4)	47.64 (7.7)
Sex (% males)	71.5 %	46.9 %	59.9 %	41.78 %	84.6 %	38.5 %
Age at onset	24.09 (7.93)		18.75 (5.43)	-		-
<1 year					26.6%	
1-3 years					65.5%	
>3years					7.9%	

Proportion (%) or mean scores (standard deviation) are given.

Sample A: 67.6% Spain, 32.4% Paris

Sample B: 73.7% Spain, 26.3% Paris

Sample C: 100% Spain

Table 3. Gene-gene interaction analyses on the risk for neurodevelopment disorders. Significant gene-gene interactions using Model-based multifactor dimensionality reduction method (MB-MDR) for each sample are shown. The Risk category indicates that for a given gene-gene combination, there are allelic combinations that occur more frequently in patients than in healthy individuals (High-risk) or increased in healthy subjects compared to patients (Low-risk).

Phenotype	Best multigene interaction model	Risk category	Beta	Wald	Perm P-value
Neurodevelopmental disorders (Sample A+C)	<i>NRG1</i> _rs6989777 x <i>ERBB4</i> _rs707284	High	0.86	18.26	3e-4 ⁺
		Low	-0.71	6.054	
SSD Risk (Sample A)	<i>NRG1</i> _rs3802160 * <i>MBP</i> _rs9966986	High	0.91	12.25	0.002 ⁺
		Low	-1.15	17.46	
SSD Risk (Sample B)	<i>ERBB3</i> _rs7971751 * <i>AKT1</i> _rs2494746	High	0.50	11.7	0.001
		Low	-0.67	2.91	
ASD Risk (Sample C)	<i>ERBB3</i> _rs2271194 * <i>QKI</i> _rs7772756	High	1.44	14.63	6e-4
		Low	-0.67	2.91	

Perm P-value: refers to the P-value for the interaction model after applying 10000 permutations.

All the reported analyses were obtained including sex as a covariate.

The analyses indicated (+) were repeated adjusting by center. Results were highly similar to those covaried by sex and remained significant (data not shown).

Sample A refers to the SSD family-based sample (parents and an offspring with a diagnosis of schizophrenia-spectrum disorder (SSD))

Sample B refers to the SSD case-control sample (individuals with a diagnosis of SSD and healthy unrelated subjects).

Sample C refers to the ASD family-based sample (parents and an offspring with a diagnosis of autism-spectrum disorder (ASD))

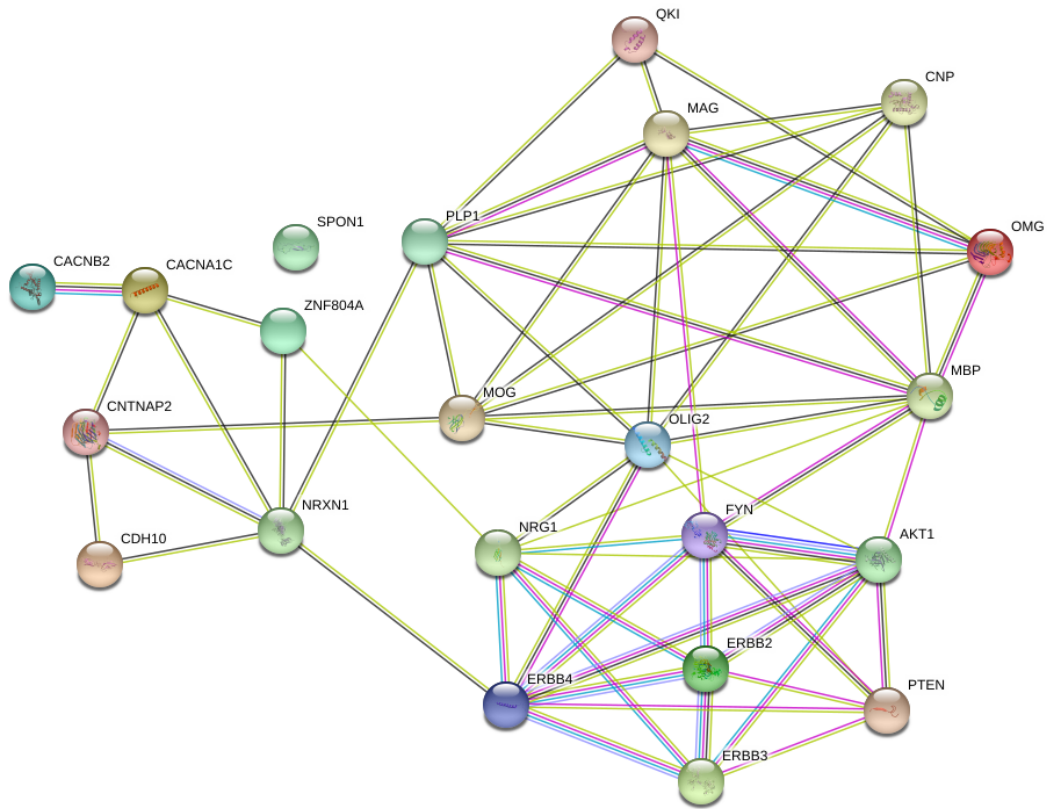


Figure 1. Protein-protein interaction network analyzed by STRING software in a set of 22 White matter genes included in our study. All 22 genes were used as input for STRING analysis and a network was built based on high confidence (0.8) evidence from experimental protein-protein interaction and curated databases. Evidence for interaction between these proteins is provided. The nodes of the network (the marbles in the figure) represent the proteins, while the edges of the network (the lines between the marbles) represent the predicted functional associations between the proteins. The color of each of the edges represents the type of evidence that exists for that interaction: a red line indicates the presence of fusion evidence, a green line indicates neighborhood evidence, a blue line indicates co-occurrence evidence, a magenta/purple line indicates experimental evidence, a yellow line indicates text-mining evidence, a light blue line indicates database evidence, and a black line indicates co-expression evidence.

SUPPLEMENTARY MATERIAL

A. METHODOLOGY

A.1. Genotyping quality control (QC)

Standard data cleaning and quality control (QC) was performed before analysis by using SPSS 21.0 software (SPSS IBM, New York, USA) and PLINK v1.07 (Purcell et al., 2007), respectively. QC of our data consisted of four steps: 1) identification of SNPs with an excessive missing genotype (--geno 0.1, which includes only SNPs with a 90% genotyping rate), 2) identification of SNPs demonstrating a significant deviation from Hardy-Weinberg equilibrium (HWE) (--hwe 0.000001), 3) identification of missing rate per person (--mind 0.2; excluding those individuals with >20% missing genotypes), and 4) the identification of makers with a low minor allele frequency (--maf 0.01).

In addition, after these steps, we performed a SNP LD pruning analysis using PLINK (which allows detecting those SNPs that are highly correlated ($r^2 > 0.8$)).

After quality control and SNP pruning, 91 SNPs survived for the ASD sample and 90 SNPs for SSD family-based and SSD case-control sample. The genotyping rates were >94%.

B. RESULTS

Table S1. Cochran-Mantel-Haenszel (CMH) association analysis in Caucasian Schizophrenia families (Sample A) for the White Matter related Genes. Note that only significant raw p-values SNPs are included.

Chr	Gene	SNP ID	Position (BP)	CMH test				
				A1 (minor allele)	X ²	P-value	OR	95% CI
2	ERBB4	rs707284	213387389	C	4.278	0.038	0.69	(0.48-0.98)
2	ERBB2	rs4252612	37863708	C	4.696	0.030	0.43	(0.19-0.93)
6	MOG	rs2535246	29636409	G	5.174	0.022	1.55	(1.06-2.25)
8	NRG1	rs73235619	31473740	A	5.335	0.020	1.75	(1.07-2.84)
8	NRG1	rs6989777	32415682	C	5.299	0.021	0.62	(0.41-0.930)
18	MBP	rs470279	74693395	T	5.828	0.015	0.56	(0.35-0.90)
18	MBP	rs523243	74714585	A	7.845	0.005*	0.52	(0.33-0.82)
19	MAG	rs6510476	35784166	G	6.625	0.010	0.57	(0.37-0.87)

*After H-B correction this SNP remained significant (corrected pvalue=0.04).

Table S2. Allelic association analysis in Caucasian Autism families (Sample C) of a set of White Matter related genes. Note that only SNPs that had a trend association are shown.

Chr	Gene	SNP ID	Position (BP)	Allelic association result				
				A1 (minor allele)	X ²	P-value	OR	95% CI
6	MOG	rs9257936	29639776	A	3.98	0.045	1.97	(1.04-3.88)
6	QKI	rs6931903	163910325	T	4.91	0.026	2.56	(1.10-6.05)
19	MAG	rs2301600	35786868	T	5.56	0.018	0.51	(0.29-0.89)
18	MBP	rs12959006	74727631	T	5.65	0.017	0.50	(0.28-0.89)

After H-B correction, no SNP remained significant.

Table S3. Cochran-Mantel-Haenszel (CMH) association analysis in SSD case-control sample (Sample B) for the White Matter related Gene-SNPs. Note that only significant raw p-values SNPs are included.

Chr	Gene	SNP ID	Position (BP)	CMH test				
				A1 (minor allele)	χ^2	P-value	OR	95% CI
12	ERBB3	rs2271194	56477694	A	9.22	0.002*	0.75	(0.19-0.93)
12	ERBB3	rs7971751	56478658	T	8.65	0.003*	0.75	(1.06-2.25)
12	ERBB3	rs877636	56480583	G	5.64	0.017	0.79	(1.07-2.84)
12	ERBB3	rs773123	56494998	T	4.75	0.029	0.72	(0.41-0.93)
14	AKT1	rs2494743	105251720	C	6.28	0.012	1.46	(0.35-0.90)
14	AKT1	rs2494746	105257719	C	7.3	0.007*	1.52	(0.33-0.82)
17	CNP	rs11296	40127060	C	4.42	0.035	0.69	(0.34-0.86)
23	PLP1	rs521895	103036412	A	5.41	0.020	0.75	(0.37-0.87)

**After H-B correction these SNPs remained significant (corrected p-values=0.018; 0.024; 0.049, respectively).*