Title Page

Generative network models identify biological mechanisms of altered structural brain connectivity in schizophrenia

Xiaolong Zhang [a] MSc[*], Urs Braun [a] MD[*], Anais Harneit [a] MSc, Zhenxiang Zang [a] MSc, Lena S. Geiger [a] MSc, Richard F. Betzel [b] PhD, Junfang Chen [a] MSc, Janina Schweiger [a] MD, Kristina Schwarz [a] MSc, Jonathan Rochus Reinwald [a] MD, Stefan Fritze [a] MD, Stephanie Witt [c] PhD, Marcella Rietschel [c] MD, Markus M. Nöthen [d] MD, Franziska Degenhardt [d] MD, Emanuel Schwarz [a] PhD, Dusan Hirjak [a] MD, Andreas Meyer-Lindenberg [a] MD, Danielle S. Bassett [e] PhD[*], Heike Tost [a] MD[*]

[a] Department of Psychiatry and Psychotherapy, Central Institute of Mental Health, Medical Faculty Mannheim, University of Heidelberg, Germany

[b] Department of Psychological and Brain Sciences, Indiana University, Bloomington, IN, USA

[c] Department of Genetic Epidemiology in Psychiatry, Central Institute of Mental Health, Medical Faculty Mannheim, University of Heidelberg, Mannheim, Germany

[d] Institute of Human Genetics, University of Bonn, School of Medicine & University Hospital Bonn, Bonn, Germany

[e] Department of Bioengineering, Department of Psychiatry, Department of Electrical & Systems Engineering, Department of Neurology, and Department of Physics & Astronomy, University of Pennsylvania, Philadelphia, PA, USA

[*] These authors contributed equally to this work.

Corresponding Author: Urs Braun M.D., Department of Psychiatry and Psychotherapy, Central Institute of Mental Health, Medical Faculty Mannheim, University of Heidelberg, J5, 68159 Mannheim, Germany; Tel.: +49 621 1703 6519; e-mail: urs.braun@zi-mannheim.de

Key Words: generative network model, structural connectivity, genetic risk, cognition, schizophrenia

Short title: Generative models of brain networks in schizophrenia

Word Count (Abstract): 220/250

Word Count (Article Body): 3999/4000

Number of Figures: 3

Number of Tables: 1

Number of References: 62

This paper contains Supplementary Materials.

Abstract

Background: Alterations in the structural connectome of schizophrenia patients have been widely characterized, but the mechanisms leading to those alterations remain largely unknown. Generative network models have recently been introduced as a tool to test the biological underpinnings of the formation of altered structural brain networks.

Methods: We evaluated different generative network models to investigate the formation of structural brain networks in healthy controls (n=152), schizophrenia patients (n=66) and their unaffected first-degree relatives (n=32), and we identified spatial and topological factors contributing to network formation. We further investigated the association of these factors to cognition and to polygenic risk for schizophrenia.

Results: Structural brain networks can be best accounted for by a two-factor model combining spatial constraints and topological neighborhood structure. The same wiring model explained brain network formation for all groups analyzed. However, relatives and schizophrenia patients exhibited significantly lower spatial constraints and lower topological facilitation compared to healthy controls. The model parameter for spatial constraint was correlated with the polygenic risk for schizophrenia and predicted reduced cognitive performance.

Conclusions: Our results identify spatial constraints and local topological structure as two interrelated mechanisms contributing to normal brain development as well as altered connectomes in schizophrenia. Spatial constraints were linked to the genetic risk for schizophrenia and general cognitive functioning, thereby providing insights into their biological basis and behavioral relevance.

INTRODUCTION

Schizophrenia is a highly heritable neurodevelopmental disorder (1-3) characterized by abnormalities in perception, cognition, affect, behavior and social functioning (4). Converging evidence supports the notion that wiring disruptions of brain networks may partially underlie these abnormalities (5-7). Indeed, previous studies have found marked differences in the brain network architecture in schizophrenia (8, 9) and delineated alterations in their structural development (10). More generally, human brain networks show a complex architecture favoring topologically advantageous properties while still being sparsely connected, in line with the proposal that the developmental architecture of the human brain connectome results from an economic trade-off between minimizing wiring costs and allowing adaptively valuable topological features (11). Indeed, there is evidence for alterations in both connection distance (12, 13) and network topology including reduced local clustering and modularity (14, 15) in schizophrenia, consistent with a biased trade-off between wiring cost and topology (11). However, the potential processes by which brain networks develop and their disturbances in schizophrenia are poorly understood.

Current network neuroscience predominantly approaches focus on *descriptive* individual or population-level differences, offering little insight into the mechanisms that give rise to network alterations in brain disorders (16-18). A recent adaption of these methods from the branch of generative network models (GNMs) may help address some of these limitations. GNM-based approaches model the stepwise formation of networks based on wiring rules reflecting potential neurodevelopmental constraints. One can then compare synthetic networks generated by the model to empirical brain networks reconstructed from neuroimaging data, thereby explicitly testing different mechanistic explanations that might govern their (disordered) structural formation (16, 19). Two recent examples have tested different wiring rules in the formation of healthy brain networks, finding converging evidence for a two factor model where one factor accounts for the spatial embedding of brain networks by penalizing spatially distant connections while the other factor enhances a complex local topological organization (20, 21).

4

The model parameters have a high degree of biological plausibility as they account for the metabolic cost of wiring and the strength of a „Hebbian-like" wiring rule, in part buttressed by the fact that they undergo progressive changes over the lifespan and show alterations in disease states (20, 21).

In addition to their biological plausibility and developmental sensitivity, an appropriate model of brain network architecture might help to illuminate genetic aspects underlying network alterations and formation in mental disorders. An important strategy is the examination of unaffected first-degree relatives of patients, who have an increased familial risk for developing the disorder (22-24). This allows for the identification of intermediate brain phenotypes linked to psychiatric risk independent of potential disease-related confounders (25). In addition, the genetic contributions to these phenotypes can be studied with modern genetic approaches utilizing the potential of cumulative genetic risk scores.

Here, we combined GNMs with imaging genetics analyses to identify potential developmental mechanisms promoting the altered formation of structural brain networks in schizophrenia (Figure 1). Building on a family of previously described and validated generative models (20), we first replicated their best-fitting model in a healthy sample, and we subsequently applied it to a group of unaffected first-degree relatives and schizophrenia patients. Following the hypothesis that there exists an aberrant balance between wiring cost and topological properties during network formation in schizophrenia, we tested whether these model parameters show the quality of an intermediate phenotype, relate to schizophrenia polygenic risk and are relevant for cognitive function.

METHODS AND MATERIALS

Participants

We studied 152 healthy controls (HC) without a first-degree relative with mental illness (mean [SD] age, 30.32 [10.28] years; 94 women), 32 unaffected first-degree relatives (REL) of

5

schizophrenia patients (33.25 [11.50] years, 19 women) and 66 unrelated patients satisfying DSM-IV-TR criteria (see Supplement) for schizophrenia (SZ, 32.77 [9.26] years; 20 women) in Mannheim, Germany. All participants provided written informed consent for the protocols approved by the local Ethics Committee of the University of Heidelberg.

Neuroimaging data acquisition and processing

Diffusion Tensor Imaging (DTI) data were acquired with a 3-T Siemens Trio scanner using two echo planar imaging (EPI) sequences with different parameters (see Supplement for details). DTI data were preprocessed with standard routines implemented in the software package FSL (https://fsl.fmrib.ox.ac.uk/fsl/). The full pipeline included the following steps: correction for head motion and eddy currents by affine registration to $b0$ image, extraction of non-brain tissues (26), and linear diffusion tensor fitting. After estimating the diffusion tensor, we performed deterministic whole-brain fiber tracking using a modified FACT algorithm (27). Further methods details are provided in the Supplement.

Construction of generative network models

We constructed synthetic networks using generative models. After defining a seed network consisting of those edges common across all subjects, edges were added one at a time until the number of edges in the synthetic network conformed to that of the observed network. The relative probability of edge formation was evaluated at each step according to the equation:

$$(1) \qquad P(u, v) = E(u, v)^{\eta} * K(u, v)^{\gamma}.$$

Here $E(u, v)$ denotes the fiber distance between brain areas u and v, and η controls the edge length. When η is negative, short-distance edges are favored, while a positive η favors long-distance edges. The term $K(u, v)$ represents the topological relationship between brain areas u and v, and γ represents its relative importance. Notably, $K(u, v)$ can be varied to realize different wiring rules.

In this study, we limited our analysis to four generative models, each representing one of four previously-studied classes: the geometric model, the degree-product model, the clustering-product model and the matching index (MI) model (20). In the geometric model, the probability of connectivity $P(u, v)$ between regions u and v is a function of the distance $E(u, v)$ between them. In the degree-product and the clustering-product model, the connection probability function includes an additional topological term $K(u, v)$, which is the product of degrees (number of connections of a brain region) or clustering coefficients (fraction of connection triangles around a brain region) between nodes u and v respectively. In the matching index model, $K(u, v)$ denotes the normalized number of nearest neighbors in common between two nodes (homophily). The topological parameters were computed using the Brain Connectivity Toolbox (https://sites.google.com/site/bctnet/Home) as implemented in MATLAB.

To evaluate the fitness of synthetic networks and to optimize models, we define an energy function that measures how dissimilar a synthetic network is to the observed network:

$$(2) \qquad E = \max \left(KS_k,\ KS_c,\ KS_b,\ KS_e \right),$$

Here, each term is a Kolmogorov-Smirnov statistic that compares degree (k), clustering coefficient (c), betweenness centrality (b) and edge length (e) distributions of synthetic and observed networks. Since we defined energy as the maximum of the four statistics, smaller energy indicated greater fitness.

We used classical Monte Carlo methods to find the parameters (η, γ) that generated networks with minimal energy. The procedure starts from randomly sampling 2000 points from the defined parameter space. Then, by computing the energy at each point and dividing the whole parameter space into Voronoi cells, we sample points preferentially within Voronoi cells with low energy. We repeated this procedure five times until it converged to a (locally) optimal solution (20).

Cognitive assessment and factor construction

In a subset of 120 individuals from all three groups, we assessed a range of cognitive subdomains frequently impaired in schizophrenia (including attention and psychomotor speed, executive function, memory, impulsivity and social emotional cognition) using the Cambridge Neuropsychological Test Automated Battery (CANTAB) (28, 29). Considering that the CANTAB measures were correlated with each other, we performed a principal component analysis (PCA) to reduce the redundancies and minimize potential for Type I error (30). The first component accounted for 27.1% of the variance and factor loadings were consistently negatively correlated with correct response rates and positively correlated with latency (or reaction times) across the seven test domains, suggesting that *lower* factor values indicate better individual cognitive performance. The detailed description of methods and a full list of included outcome measures across tasks and the resulting cognitive factors are provided in the Supplementary Information.

Polygenic risk score

We used standard methods to extract genomic DNA from EDTA blood to perform genome-wide SNP (single nucleotide polymorphism) genotyping of all individuals using the Infinium PsychArray (Illumina Inc). Quality control and imputation was performed with Gimpute (31) (see Supplement for details). To control for population stratification, we performed a PCA on the linkage-disequilibrium pruned set of autosomal SNPs using GCTA (32). Then we excluded outliers and used the first five principal components as covariates in the following association analyses of model parameters. We computed the polygenic risk score with PRSice v-2, while the expected value of the missing genotypes were imputed based on the sample allele frequency (33). In this study, genome-wide association (34) nominal $P < 0.05$, was used to achieve a balance between the number of false-positive and true-positive risk alleles (35, 36). The association analyses were repeated for thresholds of nominal $P < 0.01$ and of nominal $P < 0.1$.

Statistical analysis

8

For each participant, we tuned the parameters (η, γ) to the space where the generative model always produced synthetic networks with near-lowest energy. Within this space, we analyzed the top 1% minimal-energy synthetic networks. We compared the individual averaged energy of 1% lowest-energy synthetic networks between different types of generative models and different groups of participants using a one-way analysis of variance (ANOVA) with the Statistical Package for the Social Sciences (SPSS) 24. Then, we compared the parameters of the best-fitting model between groups using a general linear model. To investigate the genetic association of the model factors and to circumvent the potential effect of confounding factors not related to the genetic risk for the disorder in patient populations (37), we assessed the correlation between the parameters of the best-fitting model and polygenic risk scores in healthy controls only while controlling for age, sex, DTI protocol, temporal signal to noise ratio (tSNR) (38) and the first five principal components of population structure. To investigate the behavioral relevance of the identified model parameters, we assessed the correlation between the parameters of the best-fitting model and individual cognitive factor loadings while controlling for age, sex and tSNR in the three groups separately. To investigate whether our best-fitting model could capture other structural network abnormalities in schizophrenia, we also computed the global efficiency, modularity and hub degree (the degree of the top 10% highest-degree nodes) using the Brain Connectivity Toolbox. We also investigated the effect of antipsychotics on our model parameters (see Supplement).

RESULTS

Sample characterization

The groups were matched for age, education, tSNR and head motion, but not for sex and acquisition protocol (see detailed demographic and clinical characteristics as well as image quality control parameters in Table 1). To account for the group differences in these latter variables, we included sex and DTI protocol as covariates of no-interest in all analyses that included multiple groups.

9

Generative Network models

Comparison between the four network models revealed significant differences in mean energy (repeated measure ANOVA: $F(3,453) = 2964.277$, $p < 0.001$) in HC, with the matching index (MI) model showing the lowest energy level (see Figure 2A). Importantly, when including all diagnostic groups in the analysis, the group by model type interaction was not significant (repeated measure ANOVA with group as between-subjects factor, sex and DTI protocol as covariates of no-interest: $F(6,735) = 0.870$, $p = 0.516$), arguing for the same pattern across all groups. Hence, in our following investigation we focused on the analysis of the MI model, as it provided the best fit to the individual, experimentally-derived structural networks across diagnostic groups.

In Figure 3B, we illustrate the synthetic network structure of a single subject at different parameter values, thereby offering an intuition regarding the roles of each parameter in the MI model. As expected, the two parameters η and γ are strongly anti-correlated (Pearson correlation: $r = -0.613$, $p < 0.001$) in HC, a relation that was conserved across diagnostic groups (relatives: $r = -0.535$, $p = 0.002$, patients: $r = -0.492$, $p < 0.001$). Investigating the between-group difference of the two parameters, we found a significant between-group effect on the distance parameter η (ANOVA, sex and DTI protocol as covariates of no-interest; $F(2,245) = 4.777$, $p = 0.009$) and also on the topological parameter γ (ANOVA, same covariates of no-interest; $F(2,245) = 3.054$, $p = 0.049$, see Figure 2B). Post-hoc analysis confirmed significant differences between HC and SZ (η: $F(1,214) = 3.956$, $p = 0.048$; γ: $F(1,214) = 4.707$, $p = 0.031$) as well as between HC and REL in η ($F(1,180) = 8.970$, $p = 0.003$), but not in γ ($F(1,180) = 2.609$, $p = 0.108$). We found no significant correlation between individual olanzapine equivalents and model parameters (η: $r = -0.121$, $p = 0.393$; γ: $r = 0.092$, $p = 0.517$).

When comparing global efficiency, modularity and hub degree between HC and SZ, we detected significant between-group differences in global efficiency (ANOVA, same covariates of no-interest; $F(1,214) = 11.143$, $p = 0.001$) and in hub degree (ANOVA; $F(1,214) = 8.602$, $p = 0.004$) in the observed networks. In the synthetic networks, we also detected a trend-wise

10

significant between-group difference in hub degree (ANOVA; $F(1,214) = 3.218$, $p = 0.074$), and a significant difference in global efficiency (ANOVA; $F(1,214) = 8.082$, $p = 0.005$). No group differences were found in modularity for neither the observed data (ANOVA; $F(1,214) = 0.100$, $p = 0.752$) nor the synthetic network (ANOVA; $F(1,214) = 0.251$, $p = 0.617$, see Figure 3A).

Polygenic risk score

To characterize the influence of genetic risk for schizophrenia on both model parameters, we correlated the individual participants' risk scores with η and γ. We found a significant positive association for the distance parameter η ($r_{par} = 0.173$, $p = 0.045$) and a weaker, trend-wise negative association for the topological parameter γ ($r_{par} = -0.154$, $p = 0.073$, Figure 2C) in HC for all genetic variants with a nominal genome-wide significant association to schizophrenia ($P < 0.05$, uncorrected). Supplemental analyses confirmed the robustness of this finding to the choice of significance threshold used for polygenic risk score computation: nominal $P < 0.01$: η ($r_{par} = 0.154$, $p = 0.074$), γ ($r_{par} = -0.156$, $p = 0.071$); $P < 0.1$: η ($r_{par} = 0.196$, $p = 0.022$), γ ($r_{par} = -0.174$, $p = 0.043$).

CANTAB

Exploring the correlation between the individual participants' scores of the first component and the distance parameter η (or the topological parameter γ), we found a significantly negative association for η ($r = -0.261$, $p = 0.029$) and no association for γ ($r = 0.108$, $p = 0.374$) in HC. No association was found in relatives (η: $r = -0.137$, $p = 0.589$; γ: $r = 0.035$, $p = 0.890$) and in patients (η: $r = 0.261$, $p = 0.240$; γ: $r = -0.133$, $p = 0.556$, Figure 2D). As expected, healthy controls showed lower factor loadings compared to relatives and patients ($F(2, 115) = 7.680$, $p = 0.001$). We did not detect any correlation between the other four cognitive components and the network model parameters.

DISCUSSION

Converging evidence points towards alterations in the structural connectome of schizophrenia patients, but the mechanistic disturbances in the process of network formation leading to those remain poorly understood. Here we replicate prior work on the performance of generative network models (20) and thereby demonstrate that this approach yields synthetic networks that simulate many properties of structural brain networks in health and disease. We further identify significant differences between healthy controls, first-degree relatives and schizophrenia patients for the two best-fitting model parameters in a pattern that aligns with the increasing risk for schizophrenia. Specifically, these differences imply lesser geometrical constraints and lesser complexity of topological facilitation of network formation in schizophrenia. Moreover, by demonstrating differential associations between polygenetic risk score and the network formation parameters, we provide a potential mechanistic explanation of how genetic risk contributes to the malformation of brain networks. These parameters also demonstrated behavioral relevance by predicting latent features of cognitive functioning.

Firstly, our data replicate previous accounts of a superior performance of a model – parameterized by spatial distance and a matching index – in capturing several topological features of structural brain networks in comparison to other models (20). This observation aligns well with the current theory that brain networks are shaped by a mixture of geometrical constraints and topological complexity. The human brain is physically expensive and the metabolic cost of building and maintaining the axonal connections increases with connection length (39). Therefore, cost minimization seems to be an important evolutionary rule for many aspects of brain anatomy (40-42). However, cost minimization alone only poorly accounts for some of the adaptive properties of the human connectome, such as the capacity for information processing (43), which should therefore result from a trade-off between wiring cost and formation of topological features (11). In our best-fitting model, this trade-off is represented by the two generative factors: the distance parameter η penalizes the formation of long connections by imposing a spatial constraint, while the topological parameter γ favors connections between regions sharing similar neighbors, meaning that cortical areas that have similar inputs and outputs tend to be connected, compatible with Hebb's law. Interestingly and

12

in contrast to a previous study (20), we observe a negative correlation between our model parameters η and γ, suggesting that connectomes that are less constrained by spatial distance also exhibit a lower topological neighborhood structure. Together, our findings thus rather argue against a simple trade-off between both parameters. Instead our data may indicate that connectomes may be better modeled with a single, to-be-discovered latent factor that could account for both parameters simultaneously.

Secondly, when extending this framework to model mechanisms of connectome formation in schizophrenia patients and first-degree relatives, we detected no significant between-group differences in the fit between the synthetic networks and the observed networks across models. Importantly, these results suggest that the same wiring rules can equally well describe both normal and abnormal brain network formation. Previous studies have found increased connection distance of brain networks (12, 13) and changed network topology including reduced clustering and modularity (14, 15, 44) in schizophrenia (see Figure 3A). In line with this, theoretical accounts suggest that the abnormal organization of brain networks in schizophrenia may result from a biased trade-off between generative factors of homophilic attraction and distance penalization in the process of brain network formation (11). To probe further, we tested whether individual model parameters contributed differently to network formation in all three study groups. We identify smaller values of the distance parameter η in HC than in relatives and patients, while values of the topological parameter γ were higher in HC than in relatives and patients. This said, larger η values in relatives and patients indicate a lower distance penalization, thus increasing the edge length distribution, which is consistent with the increased connection distance of brain networks in schizophrenia (12, 13). Since the topological parameter mainly influences the degree distribution (21) (see Figure 3B), smaller γ values in patients and relatives indicate the presence of fewer and smaller hubs. This corroborates previous findings that suggested brain networks in schizophrenia are less clustered and have fewer hubs (9, 13). Notably, the presence of decreased spatial constraints and homophilic association in a sample of unaffected first-degree relatives suggests that these network mechanisms may resemble intermediate phenotypes, i.e., brain phenotypes that

13

relate to the increased familial risk for schizophrenia rather than being epiphenomena of potential confounds such antipsychotic medication.

Thirdly, we examined the associations between model parameters and schizophrenia polygenic risk in healthy individuals. We identified a positive correlation between polygenic risk and the distance parameter η. In addition, we showed a trend-wise negative association between polygenic risk and the topological parameter γ, suggesting that increasing genetic risk load for schizophrenia leads to a diminished distance penalization and local homophily of the structural brain connectome. In general, two interconnected factors are thought to contribute to the formation of long distance connections in the brain: a) neurons connect to each other at an early stage of neurodevelopment when the neural system is small in scale, and b) at later stages, axons follow developmental pathways already established by earlier pioneer neurons (fasciculation) (45). Studies in *C. elegans* show that 70% of long distance connections are already present at the time of hatching (46). This observation suggests that most neurons already form early connections in a spatially localized system where guidance through differential expression of guidance molecules, such as *netrin* (47) and *slit* (48), is feasible. In contrast, these molecules will not be the preferred driver to guide axons at later stages of development because source and projection neurons are further away (45). Importantly, the genes coding for such guidance molecules have been repeatedly implicated in the pathophysiology of schizophrenia (49-51). It is interesting to speculate that altered spatial expression of guidance molecules imposes less spatial constraints on brain network formation, and these fine-scale mechanisms are reflected in our large-scale model parameters.

Moreover, we found a negative correlation between the distance parameter η and individual scores of the first principal cognitive factor, predominantly capturing converging aspects of executive function and memory, in HC. In particular, we observed that larger values of η (reflecting less geometric constraints, thus a higher probability of long distance connections) were associated with better cognitive performance. Healthy human brain networks usually contain only a small fraction of long-distance shortcuts preferentially linking hub regions (12,

52). Although these long-distance connections are expensive in terms of material and metabolic cost, they greatly reduce the path-length of information transfer between spatially remote regions, thus increasing the potential for efficient information processing (11, 53), but not in weighted networks with weights based on streamline density (54). A number of previous studies have shown that more topologically efficient structural and functional networks are associated with increased IQ (55, 56) and cognitive performance (57, 58). Therefore, while long-range connections are beneficial for brain function (59), they may also be subject to increased vulnerability to disease (60). Indeed, previous studies have shown an increased proportion of long-range connections in schizophrenia resulting in brain networks shifted towards random networks (61). The association of the distance parameter η and cognition was not detectable in schizophrenia patients and first-degree relatives, suggesting an optimum in the number of long-range connections potentially exceeded in those populations.

There are a number of methodological considerations that deserve discussion. Firstly, since our goal was to investigate the mechanisms underlying the abnormal brain network formation in schizophrenia with GNM, we restricted our analysis to an already validated model framework (20). In principle, other wiring rules may be used to model connectome formation and evaluate the fitness of the resulting models. However, such an exploration is beyond the scope of this paper and would require extensive prior validation in healthy control datasets. Secondly, even complete correspondence of two networks does not necessarily imply that both models have been shaped by the same biological mechanism(s). While we have attempted to limit interpretational restraints by external validation with other well-established network features, it is important to note that generative models can be used to offer candidate mechanisms for an observed topology, but cannot conclusively prove that a given candidate mechanism actually occurred in the developing organism (16). Thirdly, while GNMs can provide insights into the formation of structural brain networks, they do not explicitly model neurodevelopmental processes. Such investigations are warranted, and will require longitudinal datasets as well as the use of an advanced GNM framework explicitly modeling variant developmental processes within subjects.

15

In conclusion, we show that the distinct wiring rules can simulate normal and abnormal network formation in humans, identify intermediate connectomic phenotypes for schizophrenia familial risk manifesting as altered spatial and topological characteristics of brain connectome formation in schizophrenia and first-degree relatives. We further corroborate the link of these network mechanisms to schizophrenia polygenic risk and demonstrate their relevance for individual cognitive function in a domain frequently disturbed in schizophrenia. Together, these data suggest that brain network formation is under strong genetic control, is optimized to support cognitive functioning and disturbed in heritable developmental disorders such as schizophrenia. While these results provide important insight into the wiring mechanisms in health and schizophrenia, longitudinal studies in developmental cohorts are needed to further elucidate successful and aberrant brain connectome formation.

16

FINANCIAL DISCLOSURES

A.M.-L. has received consultant fees from Blueprint Partnership, Boehringer Ingelheim, Daimler und Benz Stiftung, Elsevier, F. Hoffmann-La Roche, ICARE Schizophrenia, K. G. Jebsen Foundation, L.E.K Consulting, Lundbeck International Foundation (LINF), R. Adamczak, Roche Pharma, Science Foundation, Synapsis Foundation – Alzheimer Research Switzerland, System Analytics, and has received lectures including travel fees from Boehringer Ingelheim, Fama Public Relations, Institut d'investigacions Biomèdiques August Pi i Sunyer (IDIBAPS), Janssen-Cilag, Klinikum Christophsbad, Göppingen, Lilly Deutschland, Luzerner Psychiatrie, LVR Klinikum Düsseldorf, LWL PsychiatrieVerbund Westfalen-Lippe, Otsuka Pharmaceuticals, Reunions i Ciencia S. L., Spanish Society of Psychiatry, Südwestrundfunk

Fernsehen, Stern TV, and Vitos Klinikum Kurhessen. MMN is a shareholder of the Life & Brain GmbH and receives a salary from Life & Brain GmbH. MMN has received support from Shire for attending conferences. MMN has received financial remuneration from the Lundbeck Foundation, the Robert Bosch Foundation and the Deutsches Ärzteblatt for participation in scientific advisory boards. The remaining authors reported no biomedical financial interests or potential conflicts of interest.

REFERENCES

1. Sullivan PF, Daly MJ, O'Donovan M (2012): Genetic architectures of psychiatric disorders: the emerging picture and its implications. *Nature reviews Genetics*. 13:537-551.
2. Lee SH, DeCandia TR, Ripke S, Yang J, Sullivan PF, Goddard ME, et al. (2012): Estimating the proportion of variation in susceptibility to schizophrenia captured by common SNPs. *Nature genetics*. 44:247-250.
3. Sullivan PF, Kendler KS, Neale MC (2003): Schizophrenia as a complex trait: evidence from a meta-analysis of twin studies. *Archives of general psychiatry*. 60:1187-1192.
4. van Os J, Kapur S (2009): Schizophrenia. *Lancet (London, England)*. 374:635-645.
5. Fornito A, Zalesky A, Pantelis C, Bullmore ET (2012): Schizophrenia, neuroimaging and connectomics. *Neuroimage*. 62:2296-2314.
6. Fitzsimmons J, Kubicki M, Shenton ME (2013): Review of functional and anatomical brain connectivity findings in schizophrenia. *Current opinion in psychiatry*. 26:172-187.
7. van den Heuvel MP, Fornito A (2014): Brain networks in schizophrenia. *Neuropsychology review*. 24:32-48.
8. van den Heuvel MP, Sporns O, Collin G, Scheewe T, Mandl RC, Cahn W, et al. (2013): Abnormal rich club organization and functional brain dynamics in schizophrenia. *JAMA psychiatry*. 70:783-792.
9. van den Heuvel MP, Mandl RC, Stam CJ, Kahn RS, Hulshoff Pol HE (2010): Aberrant frontal and temporal complex network structure in schizophrenia: a graph theoretical analysis. *J Neurosci*. 30:15915-15926.
10. Zalesky A, Pantelis C, Cropley V, Fornito A, Cocchi L, McAdams H, et al. (2015): Delayed Development of Brain Connectivity in Adolescents With Schizophrenia and Their Unaffected Siblings. *JAMA psychiatry*. 72:900-908.
11. Bullmore E, Sporns O (2012): The economy of brain network organization. *Nat Rev Neurosci*. 13:336-349.
12. Alexander-Bloch AF, Vertes PE, Stidd R, Lalonde F, Clasen L, Rapoport J, et al. (2013): The anatomical distance of functional connections predicts brain network topology in health and schizophrenia. *Cereb Cortex*. 23:127-138.
13. Bassett DS, Bullmore E, Verchinski BA, Mattay VS, Weinberger DR, Meyer-Lindenberg A (2008): Hierarchical organization of human cortical networks in health and schizophrenia. *J Neurosci*. 28:9239-9248.
14. Liu Y, Liang M, Zhou Y, He Y, Hao Y, Song M, et al. (2008): Disrupted small-world networks in schizophrenia. *Brain : a journal of neurology*. 131:945-961.
15. Alexander-Bloch AF, Gogtay N, Meunier D, Birn R, Clasen L, Lalonde F, et al. (2010): Disrupted modularity and local connectivity of brain functional networks in childhood-onset schizophrenia. *Front Syst Neurosci*. 4:147.

16.   Betzel RF, Bassett DS (2017): Generative models for network neuroscience: prospects and promise. *J R Soc Interface*. 14.

17.   Vertes PE, Bullmore ET (2015): Annual research review: Growth connectomics--the organization and reorganization of brain networks during normal and abnormal development. *Journal of child psychology and psychiatry, and allied disciplines*. 56:299-320.

18.   Braun U, Schaefer A, Betzel RF, Tost H, Meyer-Lindenberg A, Bassett DS (2018): From Maps to Multi-dimensional Network Mechanisms of Mental Disorders. *Neuron*. 97:14-31.

19.   Bassett DS, Sporns O (2017): Network neuroscience. *Nat Neurosci*. 20:353-364.

20.   Betzel RF, Avena-Koenigsberger A, Goni J, He Y, de Reus MA, Griffa A, et al. (2016): Generative models of the human connectome. *Neuroimage*. 124:1054-1064.

21.   Vertes PE, Alexander-Bloch AF, Gogtay N, Giedd JN, Rapoport JL, Bullmore ET (2012): Simple models of human brain functional networks. *Proc Natl Acad Sci U S A*. 109:5868-5873.

22.   Erk S, Meyer-Lindenberg A, Schmierer P, Mohnke S, Grimm O, Garbusow M, et al. (2014): Hippocampal and frontolimbic function as intermediate phenotype for psychosis: evidence from healthy relatives and a common risk variant in CACNA1C. *Biol Psychiatry*. 76:466-475.

23.   Rasetti R, Weinberger DR (2011): Intermediate phenotypes in psychiatric disorders. *Current opinion in genetics & development*. 21:340-348.

24.   Collin G, Scholtens LH, Kahn RS, Hillegers MHJ, van den Heuvel MP (2017): Affected Anatomical Rich Club and Structural-Functional Coupling in Young Offspring of Schizophrenia and Bipolar Disorder Patients. *Biol Psychiatry*. 82:746-755.

25.   Rasetti R, Sambataro F, Chen Q, Callicott JH, Mattay VS, Weinberger DR (2011): Altered cortical network dynamics: a potential intermediate phenotype for schizophrenia and association with ZNF804A. *Arch Gen Psychiatry*. 68:1207-1217.

26.   Smith SM (2002): Fast robust automated brain extraction. *Human brain mapping*. 17:143-155.

27.   Yeh FC, Verstynen TD, Wang Y, Fernandez-Miranda JC, Tseng WY (2013): Deterministic diffusion fiber tracking improved by quantitative anisotropy. *PloS one*. 8:e80713.

28.   Mesholam-Gately RI, Giuliano AJ, Goff KP, Faraone SV, Seidman LJ (2009): Neurocognition in first-episode schizophrenia: a meta-analytic review. *Neuropsychology*. 23:315-336.

29.   Barch DM, Ceaser A (2012): Cognition in schizophrenia: core psychological and neural mechanisms. *Trends Cogn Sci*. 16:27-34.

30.   Levin HS, Li X, McCauley SR, Hanten G, Wilde EA, Swank P (2013): Neuropsychological outcome of mTBI: a principal component analysis approach. *Journal of neurotrauma*. 30:625-632.

31.   Chen J, Lippold D, Frank J, Rayner W, Meyer-Lindenberg A, Schwarz E (2018): Gimpute: An efficient genetic data imputation pipeline. *Bioinformatics (Oxford, England)*.

32.   Yang J, Lee SH, Goddard ME, Visscher PM (2011): GCTA: a tool for genome-wide complex trait analysis. *American journal of human genetics*. 88:76-82.

33.   Euesden J, Lewis CM, O'Reilly PF (2015): PRSice: Polygenic Risk Score software. *Bioinformatics (Oxford, England)*. 31:1466-1468.

34.   Schizophrenia Working Group of the Psychiatric Genomics C (2014): Biological insights from 108 schizophrenia-associated genetic loci. *Nature*. 511:421-427.

35.   Wray NR, Lee SH, Mehta D, Vinkhuyzen AA, Dudbridge F, Middeldorp CM (2014): Research review: Polygenic methods and their application to psychiatric traits. *Journal of child psychology and psychiatry, and allied disciplines*. 55:1068-1087.

36.   Agerbo E, Sullivan PF, Vilhjalmsson BJ, Pedersen CB, Mors O, Borglum AD, et al. (2015): Polygenic Risk Score, Parental Socioeconomic Status, Family History of Psychiatric Disorders, and the Risk for Schizophrenia: A Danish Population-Based Study and Meta-analysis. *JAMA psychiatry*. 72:635-641.

37.   Chen Q, Ursini G, Romer AL, Knodt AR, Mezeivtch K, Xiao E, et al. (2018): Schizophrenia polygenic risk score predicts mnemonic hippocampal activity. *Brain : a journal of neurology*. 141:1218-1228.

38.   Roalf DR, Quarmley M, Elliott MA, Satterthwaite TD, Vandekar SN, Ruparel K, et al. (2016): The impact of quality assurance assessment on diffusion tensor imaging outcomes in a large-scale population-based cohort. *Neuroimage*. 125:903-919.

39.   Chklovskii DB (2004): Exact solution for the optimal neuronal layout problem. *Neural computation*. 16:2067-2078.

40.   Chklovskii DB, Schikorski T, Stevens CF (2002): Wiring optimization in cortical circuits. *Neuron*. 34:341-347.

41.   Kaiser M, Hilgetag CC (2006): Nonoptimal component placement, but short processing paths, due to long-distance projections in neural systems. *PLoS Comput Biol*. 2:e95.

42.   Chklovskii DB, Koulakov AA (2004): Maps in the brain: what can we learn from them? *Annual review of neuroscience*. 27:369-392.

43.   Costa Lda F, Kaiser M, Hilgetag CC (2007): Predicting the connectivity of primate cortical networks from topological and spatial node properties. *BMC systems biology*. 1:16.

44.   Lynall ME, Bassett DS, Kerwin R, McKenna PJ, Kitzbichler M, Muller U, et al. (2010): Functional connectivity and brain networks in schizophrenia. *J Neurosci*. 30:9477-9487.

45.   Kaiser M (2017): Mechanisms of Connectome Development. *Trends Cogn Sci*. 21:703-717.

46.   Varier S, Kaiser M (2011): Neural development features: spatio-temporal development of the Caenorhabditis elegans neuronal network. *PLoS computational biology*. 7:e1001044.

47.   Serafini T, Kennedy TE, Galko MJ, Mirzayan C, Jessell TM, Tessier-Lavigne M (1994): The netrins define a family of axon outgrowth-promoting proteins homologous to C. elegans UNC-6. *Cell*. 78:409-424.

48.   Farmer WT, Altick AL, Nural HF, Dugan JP, Kidd T, Charron F, et al. (2008): Pioneer longitudinal axons navigate using floor plate and Slit/Robo signals. *Development (Cambridge, England)*. 135:3643-3653.

49.   Aoki-Suzuki M, Yamada K, Meerabux J, Iwayama-Shigeno Y, Ohba H, Iwamoto K, et al. (2005): A family-based association study and gene expression analyses of netrin-G1 and -G2 genes in schizophrenia. *Biol Psychiatry*. 57:382-393.

50.   Eastwood SL, Harrison PJ (2008): Decreased mRNA expression of netrin-G1 and netrin-G2 in the temporal lobe in schizophrenia and bipolar disorder. *Neuropsychopharmacology : official publication of the American College of Neuropsychopharmacology*. 33:933-945.

51.   Shi Y, Zhao X, Yu L, Tao R, Tang J, La Y, et al. (2004): Genetic structure adds power to detect schizophrenia susceptibility at SLIT3 in the Chinese Han population. *Genome research*. 14:1345-1349.

52.   Bullmore ET, Bassett DS (2011): Brain graphs: graphical models of the human brain connectome. *Annual review of clinical psychology*. 7:113-140.

53.   Buzsaki G, Geisler C, Henze DA, Wang XJ (2004): Interneuron Diversity series: Circuit complexity and axon wiring economy of cortical interneurons. *Trends in neurosciences*. 27:186-193.

54.   Betzel RF, Bassett DS (2018): Specificity and robustness of long-distance connections in weighted, interareal connectomes. *Proceedings of the National Academy of Sciences*. 115:E4880-E4889.

55.   Langer N, Pedroni A, Gianotti LR, Hanggi J, Knoch D, Jancke L (2012): Functional brain network efficiency predicts intelligence. *Hum Brain Mapp*. 33:1393-1406.

56.   van den Heuvel MP, Stam CJ, Kahn RS, Hulshoff Pol HE (2009): Efficiency of functional brain networks and intellectual performance. *J Neurosci*. 29:7619-7624.

57.   Giessing C, Thiel CM, Alexander-Bloch AF, Patel AX, Bullmore ET (2013): Human brain functional network changes associated with enhanced and impaired attentional task performance. *J Neurosci*. 33:5903-5914.

58.   Breckel TP, Thiel CM, Bullmore ET, Zalesky A, Patel AX, Giessing C (2013): Long-term effects of attentional performance on functional brain network topology. *PLoS One*. 8:e74125.

59.   Hofman MA (2014): Evolution of the human brain: when bigger is better. *Frontiers in neuroanatomy*. 8:15.

60.   Crow TJ (1997): Is schizophrenia the price that Homo sapiens pays for language? *Schizophrenia research*. 28:127-141.

61.   Lo CY, Su TW, Huang CC, Hung CC, Chen WL, Lan TH, et al. (2015): Randomization and resilience of brain functional networks as systems-level endophenotypes of schizophrenia. *Proc Natl Acad Sci U S A*. 112:9123-9128.
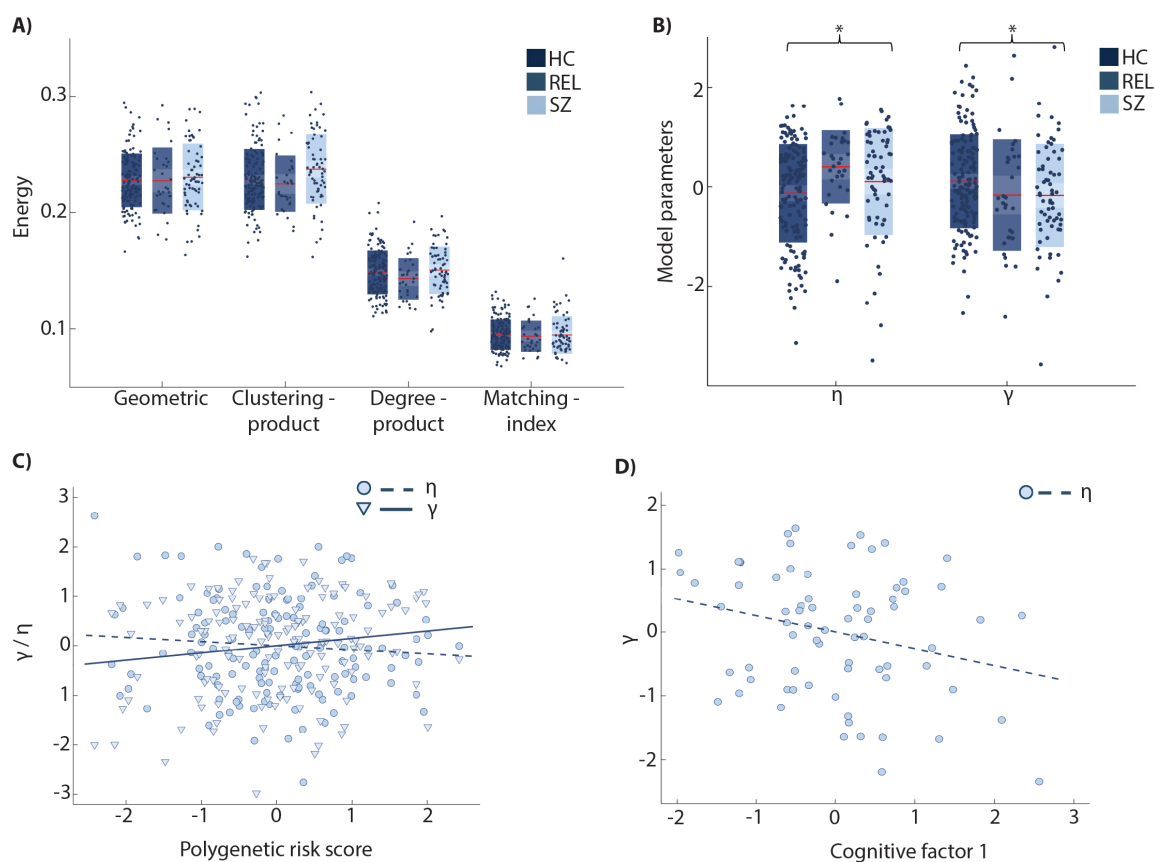
FIGURES:
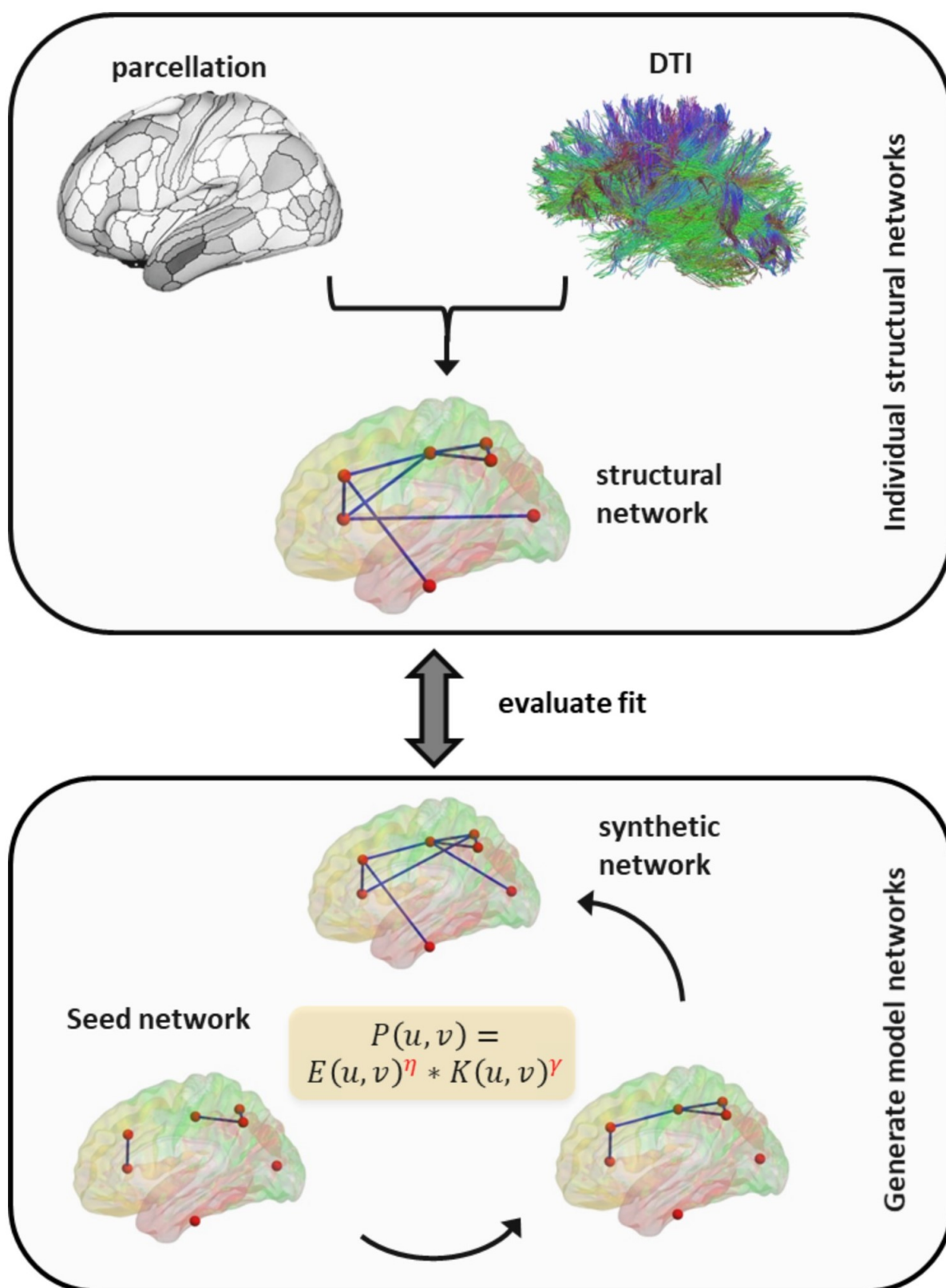
Figure 1. Overview of generative network models. Deterministic whole-brain fiber tracking was

performed to reconstruct white mater pathways, from which we constructed structural networks

linking 180 regions of interest. By retaining edges present in all subjects, a seed network was created and then edges were added stepwise with the probability of edge formation $P(u, v)$ until the number of connections in the synthetic network was the same as that in the observed structural network. The fitness of the synthetic network was evaluated by comparing the degree, clustering coefficient, betweenness centrality and edge length distributions between the synthetic network and the observed structural network.

Figure 2. Generative model networks in health and disease. (A) Individual energy was significantly different among the four types of models (repeated-measure ANOVA: $F(3,453)$ = 2964.277, $p < 0.001$) without a significant model by group interaction ($F(6,735)$ = 0.870, p = 0.516). The matching index model showed the lowest energy. (B) In the matching index model, there was a significant between-group effect on the distance parameter η (ANOVA, $F(2,245)$ = 4.777, p = 0.009) and on the topological parameter γ (ANOVA, $F(2,245)$ = 3.054, p = 0.049) correcting for sex and DTI protocol. Red lines indicate mean values and boxes indicate one standard deviation of the mean. Asterisks denote significant difference between all diagnostic groups. (C) Individual polygenic risk score for schizophrenia was significantly positively associated with the distance parameter η (r_{par} = 0.173, p = 0.045) and trend-wise negatively associated with the topological parameter γ (r_{par} = -0.154, p = 0.073). (D) After performing principal component analysis on 13 main outcome measures of the neuropsychological test battery, we obtained five components whose eigenvalues were larger than 1, and we found a

significant negative correlation between the first component and η (r = -0.261, p = 0.029) in

healthy controls.
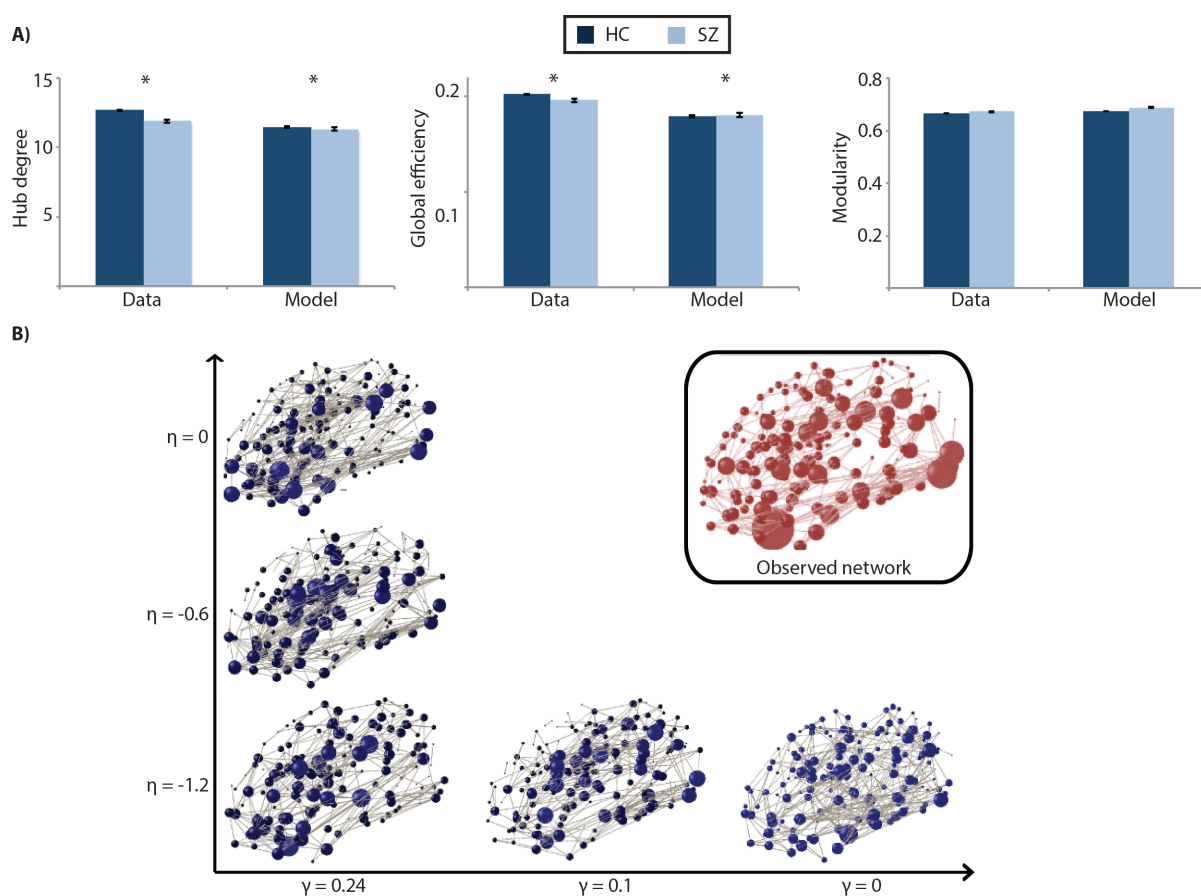
Figure 3. Topological characteristics in health and disease. (A) Comparison of several network characteristics in data and model between healthy controls (HC) and schizophrenia patients (SZ). The matching index model captured the abnormal hubness and global efficiency in schizophrenia patients well, while also adequately modeling no modularity difference between groups. The deep blue bars represent the clustering coefficient, modularity and hub degree of the synthetic network (model) and observed network (data) in healthy controls, respectively, while the light blue bars represent schizophrenia patients. Bars indicate mean values. Error bars indicate standard errors. Asterisks denote significant difference between diagnostic groups. Note that raw data is displayed. (B) Visualization of parameter effects on network structure for a single subject. The topological parameter γ mainly influences the degree distribution with larger γ corresponding to the occurrence of more and larger hubs, while the

distance parameter η appears to tune the spatial position of hubs in the network. Here, η = -1.2 and γ = 0.24 correspond to the best-fitting model.

TABLES

Table 1: Demographic, clinical and neuroimaging characteristics

	Healthy controls (n = 152)	First-degree relatives (n = 32)	Schizophrenia patients (n = 66)	F or χ^2 value	*P* value
Demographic characteristics					
Age (years)	30.32 ± 10.28	33.25 ± 11.50	32.77 ± 9.26	1.977	0.141
Sex (male / female)	58/94	13/19	46/20	18.95	< 0.001
Acquisition protocol (32/12 channel coil)	76/76	21/11	66/0	50.71	< 0.001
Education (years)	15.40 ± 1.60	15.19 ± 2.34	14.86 ± 2.17	1.881	0.155
Clinical characteristics					
PANSS positive	n.a	n.a.	14.61 ± 7.54	-	-
PANSS negative	n.a	n.a	14.51 ± 8.27	-	-
PANSS general	n.a	n.a	31.70 ± 11.42	-	-
PANSS total	n.a	n.a	60.82 ± 23.51	-	-
Duration of illness (years)	n.a.	n.a.	10.44 ± 8.34	-	-
Olanzapine equivalents (n= 52)	n.a.	n.a.	15.04 ± 8.91		
QC parameters					
DTI: mean relative root-mean-square displacement (mm)	0.31 ± 0.10	0.33 ± 0.12	0.35 ± 0.18	1.844	0.160
DTI: tSNR	5.85 ± 0.29	5.76 ± 0.25	5.85 ± 0.26	1.491	0.227

PANSS = Positive and Negative Syndrome Scale (62). tSNR = temporal signal to noise ratio, QC = quality control