Review

Phenotype-Based Genetic Association Studies (PGAS)—Towards Understanding the Contribution of Common Genetic Variants to Schizophrenia Subphenotypes

Hannelore Ehrenreich [1,2],* and Klaus-Armin Nave [1,2],*

[1]   Max Planck Institute of Experimental Medicine, Hermann-Rein-Str.3, 37075 Göttingen, Germany
[2]   DFG Center for Nanoscale Microscopy and Molecular Physiology of the Brain (CNMPB), Hermann-Rein-Str.3, 37075 Göttingen, Germany

*   Authors to whom correspondence should be addressed; E-Mails: ehrenreich@em.mpg.de (H.E.); nave@em.mpg.de (K.-A.N.); Tel.: +49-551-389-9628 (H.E.); Fax: +49-551-389-9670 (H.E.).

Abstract: Neuropsychiatric diseases ranging from schizophrenia to affective disorders and autism are heritable, highly complex and heterogeneous conditions, diagnosed purely clinically, with no supporting biomarkers or neuroimaging criteria. Relying on these "*umbrella diagnoses*", genetic analyses, including genome-wide association studies (GWAS), were undertaken but failed to provide insight into the biological basis of these disorders. "Risk genotypes" of unknown significance with low odds ratios of mostly <1.2 were extracted and confirmed by including ever increasing numbers of individuals in large multicenter efforts. Facing these results, we have to hypothesize that thousands of genetic constellations in highly variable combinations with environmental co-factors can cause the individual disorder in the sense of a final common pathway. This would explain why the prevalence of mental diseases is so high and why mutations, including copy number variations, with a higher effect size than SNPs, constitute only a small part of variance. Elucidating the contribution of normal genetic variation to (disease) phenotypes, and so re-defining disease entities, will be extremely labor-intense but crucial. We have termed this approach PGAS ("phenotype-based genetic association studies"). Ultimate goal is the definition of biological subgroups of mental diseases. For that purpose, the GRAS (Göttingen Research Association for Schizophrenia) data collection was initiated in 2005. With >3000 phenotypical data points per patient, it comprises the world-wide largest currently available schizophrenia database (N > 1200), combining genome-wide SNP coverage

and deep phenotyping under highly standardized conditions. First PGAS results on normal genetic variants, relevant for e.g., cognition or catatonia, demonstrated *proof-of-concept*. Presently, an autistic subphenotype of schizophrenia is being defined where an unfortunate accumulation of normal genotypes, so-called pro-autistic variants of synaptic genes, explains part of the phenotypical variance. Deep phenotyping and comprehensive clinical data sets, however, are expensive and it may take years before PGAS will complement conventional GWAS approaches in psychiatric genetics.

1. Schizophrenia Is a Heterogeneous Group of Diseases Diagnosed Purely Clinically

The diagnosis of schizophrenia (as of most mental diseases) is to this day an exclusively clinical one which, based on the leading classification systems, DSM and ICD, demands the simultaneous presence of a number of symptoms that are labeled "positive" or "negative". In addition, persistence of these symptoms for at least 6 months is requested. As such, schizophrenia is the unifying term of a highly complex and heterogeneous group of multigenetic disorders, with an array of environmental hazards having influenced onset and course. Common to this group of disorders are merely some gross phenotypical traits. Biomarkers in unequivocal support of the diagnosis are missing and also modern imaging technologies, despite revealing brain matter loss or various functional alterations, have not yet assisted in better understanding disease etiology or pathogenesis, or in defining objective diagnostic criteria. This conglomerate of uncertainties, however, is the groundwork on which genetic studies have been and are being built on.

2. Despite High Heritability of Schizophrenia No "Disease Genes" Have Been Uncovered

There is no doubt that schizophrenia has a major genetic root. Family studies yielded heritability estimates of up to 80% for schizophrenia [1,2]. Monozygotic twins manifest concordance rates of ~50%, thus—on top of a clear genetic origin—pointing to a considerable influence of non-genetic causes, *i.e.*, environmental and/or epigenetic factors [3–6]. Searching for the genetic part of schizophrenia etiology, a considerable number of explorations have been undertaken, ranging from segregation or linkage analyses to association studies based on candidate genes [7,8]. Most of the so identified "disease genes", e.g., *DISC1* [9], turned out not to respect disease borders. In fact, the *DISC1* translocation was found associated as much with mental health as with affective diseases, schizophrenia, autism, or personality disorders [10]. *DISC1* and others may thus at best deserve the label of a global risk gene for mental disorders.

So far, no definite "schizophrenia genes" of general significance could be identified. Recently reported genome-wide association studies (GWAS) of schizophrenia identified a number of genetic risk markers significantly associated with the disease, but unfortunately could not extract any universally convincing "disease genes". The initially limited reproducibility over studies and ethnicities has

improved with increasing numbers to ~60,000 individuals (PGC, Psychiatric Genetic Consortium) [11]. Associations, however, remain hampered by very low odds ratios (OR distinctly <1.2) [12,13]. Altogether, it appears highly unlikely that even larger GWAS, based on the umbrella term "schizophrenia", will succeed in unraveling the genetic basis of these conditions or in identifying relevant disease genes.

3. "GWAS Hits" Do Not Predict Disease Severity or Other Relevant Schizophrenia Phenotypes

Which information do we then gain from the GWAS-identified risk genotypes? In fact, even an accumulation of the GWAS-identified "top 10" risk genotypes [14] does not lead to a more severe disease phenotype. In other words: If an individual possessed all "top 10" risk markers at the same time ("accumulated risk"), his disease will not be more grave than if he carried none of the risk genotypes.

Another attempt to make use of the GWAS-derived information has been the definition of polygenic schizophrenia risk scores (PSS). Besides the genome-wide significant risk loci, a substantial proportion of schizophrenia risk has been hypothesized to lie in markers not achieving genome-wide significance. Thus, quantitative PSS were calculated based on nominal associations of each SNP from the PGC GWAS. These PSS explained up to 6% of variance regarding the diagnosis schizophrenia in an independent sample [11]. Subsequently, PSS effects on various disease-relevant phenotypes, e.g., brain matter dimensions, were explored with variable success [15,16]. Employing the GRAS data collection of >1200 well characterized schizophrenic subjects (see below), we could not uncover the slightest association of PSS with any lead features of schizophrenia. In contrast, we found in the same schizophrenic population dramatic effects of accumulated environmental risk on age at onset of the disease [17].
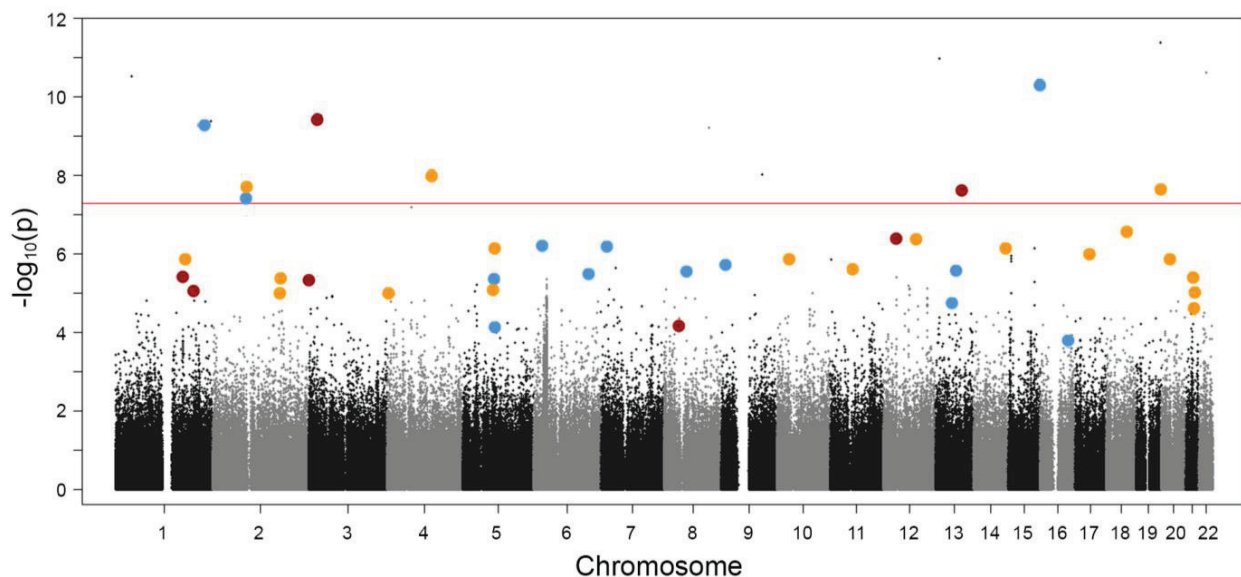
4. Definition of Biological Subgroups of Schizophrenia: The Essential Next Step

In summary, we note that all genetic "schizophrenia markers" discovered by GWAS up to date do not contribute more to the entirety of this disease than a minimally heightened probability (OR <1.2) of receiving the "umbrella diagnosis schizophrenia". This whole picture could, however, look quite different when considering biological subgroups of schizophrenia as illustrated in Figure 1. GWAS on well-defined disease-relevant subphenotypes may ultimately not only lead to the identification of entirely new risk genotypes but, importantly, also to a re-distribution of some of the identified GWAS risk markers to certain subgroups, resulting then in much higher odds ratios.

In order to define biological subgroups of schizophrenia, the GRAS (*Göttingen Research Association for Schizophrenia*) data collection was initiated in 2004 [18,19]. The lead hypothesis of GRAS states that schizophrenia is caused by thousands of possible combinations of genetic markers interacting with a large array of different environmental risk factors. There may be rare cases where a 'per se' critical genetic load exists and the disease onset is independent of additional external factors. In the vast majority of cases, however, only the interaction between certain genetic predispositions and environmental factors will lead to disease onset. The overlap of genetic risk factors between individual schizophrenic patients/families probably is fairly low. This in turn might explain why it is obviously impossible to obtain common risk genes of schizophrenia with convincing odds ratios. According to

the primary GRAS hypothesis, schizophrenia is the result of a combination of 'unfortunate' normal genetic variants exposed to unfavorable environmental influence. Apart from its scientific significance, undoubtedly, this hypothesis strongly supports any anti-stigma campaign.

Figure 1. Schematic presentation of three hypothetical subphenotypes of schizophrenia embedded in a sample Manhattan plot. The non-overlapping red, blue or yellow dots comprise genetic constellations ("assemblies") suggested to account for disease subforms with respective subphenotypes. Note that typical genome-wide association studies (GWAS) "top hits", as defined by the highest significance levels for the clinical endpoint diagnosis ("schizophrenia"), most likely fall into different assemblies. This can explain their low odd ratios in a large and diverse patient group and the apparent lack of interactions. The depicted schema is purely hypothetical and shall illustrate the difficulties to define disease genes by conventional GWAS approaches based on endpoint diagnoses.

We hypothesize that a considerable proportion of the population across cultures (likely ~50%) may harbor a principal genetic make-up for developing mental diseases whereas the remaining 50% of individuals could never acquire them due to absence of respective genetic prerequisites. Only in a small fraction of risk carriers the disease will break out (~0.5%–1% of the population across cultures) [20], co-induced by a multitude of potential external inputs, while the bulk of carriers will stay healthy and forward their predisposition to their offspring.

5. Deep Phenotyping Is a Prerequisite for PGAS: The GRAS Data Collection

If all human genetic approaches to schizophrenia as a "classic" genetic disease apparently failed, how will it be possible to learn more about the contribution of genes/genotypes to relevant disease subphenotypes? Motivated by this central question we started an alternative and at the same time complementary approach to conventional GWAS. We call this approach PGAS, which stands for *phenotype-based genetic association study*. PGAS enables us to explore the contribution of genes/genetic markers to schizophrenic subphenotypes. To start PGAS, we had to establish a

comprehensive standardized phenotypical characterization of schizophrenic patients, the above introduced GRAS data collection. Within a few years we compiled the presently world-wide largest phenotypical database of schizophrenic patients comprising presently ~1200 subjects, with slow steady-state recruitment ongoing. Although much larger data bases of schizophrenic subjects exist, unfortunately, they do not even come close to the deep phenotyping information of GRAS. A total of 23 psychiatric hospitals all over Germany were involved as hosting centers in this effort [18,19]. Inclusion criteria for the patients were confirmed diagnosis of schizophrenia or schizoaffective disorder, and the ability to cooperate at least to a minimal degree. The GRAS population is representative for those schizophrenic patients in central Europe/Germany who are in contact with the healthcare system. Unrivalled in the GRAS data collection is the fact that all cross-sectional examinations of all patients were performed by one and the same travelling team. This fact contributes in a highly significant manner to the homogeneity of the resulting database, containing at present 82% schizophrenic and 18% schizoaffective subjects [18,19].

The GRAS data collection does not only encompass sociodemographic and basic clinical parameters but also, for example, a very comprehensive cognitive test battery, an extensive investigation of neurological signs and symptoms, a thorough medication history, several tests of extrapyramidal side effects of antipsychotic medication, complete information on drug abuse/addiction or other comorbidities, to name just a few of the covered data modules. Apart from the comprehensive cross-sectional analyses, we collected nearly all psychiatric medical reports and discharge letters of all GRAS patients, allowing also for fairly solid longitudinal data analysis, e.g., information on disease course and outcome. Altogether, we have about 3000 data points per patient, which allow us to perform PGAS [18,19]. Importantly, most GRAS patients agreed to be re-contacted for follow-up studies, thereby enabling targeted analyses in the future, e.g., detailed morphometrical or functional magnetic resonance imaging studies, or inducible pluripotent stem cell approaches to study cellular consequences of complex genotypes.

6. Phenotype-Based Genetic Association Studies (PGAS): Proof-of-Principle

First *proof-of-principle* studies for the PGAS approach were performed over the last years. We could show that genetic variants within the *complexin2* gene (6 single nucleotide polymorphisms) influence the cognitive capability of schizophrenic patients [18]. This gene encodes a synaptic protein that is crucial for the regulation of neurotransmitter release. Similarly, the calcium-activated potassium channel SK3 is involved in the modulation of neuroplasticity and influences cognition of schizophrenic individuals genotype-dependently (as a function of the length of a CAG repeat in the *N*-terminal region) [21]. Moreover, a distinct combination of erythropoietin (EPO) and EPO receptor (EPOR) genotypes leads to a remarkable cognitive benefit in schizophrenic subjects as compared to all other possible EPO/EPOR genotype combinations [22]. Of note, the association of certain genetic variants with cognition in schizophrenia may likewise hold true for healthy individuals and other disease populations [22]. This clearly emphasizes the importance of investigating the genotype contribution to phenotypes in general, rather than immediately focusing on risk genes in complex diseases.

Among the discussed candidate risk genes of schizophrenia is *NRG1* [23–25]. Analyzing its disease-relevance in the GRAS population yielded for the first time an association with age at disease

onset and severity of positive symptoms [26]. A myelin-associated gene, *CNP*, turned out to co-determine the occurrence of a catatonia-depression syndrome upon aging [27]. Importantly, in most of these PGAS approaches, either replication in an independent human cohort and/or behavioral phenotypes in corresponding mouse mutants were demonstrated. In fact, the translation to mouse models, and even to cultured cells, plays an important part in the search for mechanistic comprehension of the observed phenomena.

7. PGAS Will Be Instrumental for Definition of Biological Schizophrenia Subgroups

Even though the above mentioned first *proof-of-principle* studies were important to set the stage for PGAS, it is now mandatory to work on the "bigger picture". The availability of genome-wide genetic data on the GRAS sample enables us to better understand the genetic basis of neuropsychiatric phenotypes. Using the Axiom^{TM} Caucasian European array as a backbone, we have enriched it by including additional markers of putative functional importance [28]. At present, we are in the process of defining biological subphenotypes of schizophrenia which are less complex as compared to the 'umbrella diagnosis schizophrenia' and can be described and well quantified by respective phenotype scores. Prerequisite for applying these scores to genetic studies is a high internal consistency of the selected score items, reflected by a demanded Cronbach's alpha of >80%. As a first example, we have operationalized an autistic subphenotype of schizophrenia. This subphenotype, which comprises lead features of high-functioning autism (deficits in social interaction and communication, as well as repetitive behaviors/stereotypies), shows a normal distribution in the GRAS population, thus allowing for extreme group comparisons with respect to the associated genotypical information. Our primary target genes for the autistic subphenotype are synaptic genes where we find an unfortunate accumulation of normal genotypes, that we consider as "pro-autistic" variants of these genes, to explain an appreciable part of the phenotypical variance. A sample of Asperger autists is presently recruited with the aim to replicate the genotype-phenotype associations found in the autistic subgroup of schizophrenic individuals. Similarly, a psychomotor and a schizoaffective subphenotype of schizophrenia will be phenotypically as well as—subsequently—genetically characterized, targeting myelin- and ion channel-associated genes, respectively [29].

8. Conclusions

To conclude, the lack of objective diagnoses and the non-existence of classical disease genes have forced us to re-consider the genetic approach to mental diseases. Deep clinical phenotyping as prerequisite for PGAS, combined with genome-wide SNP coverage, makes it possible to improve our understanding of the molecular-genetic architecture of schizophrenia and likely other mental diseases via systematic phenotype-based approaches. These in turn necessitate highly labor-intense groundwork since they depend on the availability of comprehensive phenotypical databases comparable to the GRAS data collection. Some efforts in a similar direction are launched in other disciplines, for instance in a first trial to integrate available electronic medical record data and GWAS information [30]. Regarding the depth of phenotyping, however, this approach certainly needs to be substantially improved. Even if "deep phenotyping" is much more tedious than collecting thousands of subjects with merely a certain diagnostic label versus healthy individuals, it will be the only way to reach the goal,

i.e., novel definitions of biologically sound disease subgroups. It can only be hoped that big consortia and sponsors will follow this rationale such that the number of deeply phenotyped individuals will grow. At the end of the road, PGAS will permit better insights into the complex genotype-phenotype interactions of schizophrenia (and other neuropsychiatric diseases) and in this way open up more targeted therapeutic strategies for biological subphenotypes of the disease.

Acknowledgments

Author Contributions

Both authors discussed the contents, designed the paper outline, and jointly wrote the manuscript.

Conflicts of Interest

The authors declare no conflict of interest.

References

1. Cardno, A.G.; Gottesman, I.I. Twin studies of schizophrenia: From bow-and-arrow concordances to star wars mx and functional genomics. *Am. J. Med. Genet.* **2000**, *97*, 12–17.
2. Lichtenstein, P.; Yip, B.H.; Bjork, C.; Pawitan, Y.; Cannon, T.D.; Sullivan, P.F.; Hultman, C.M. Common genetic determinants of schizophrenia and bipolar disorder in swedish families: A population-based study. *Lancet* **2009**, *373*, 234–239.
3. Cardno, A.G.; Marshall, E.J.; Coid, B.; Macdonald, A.M.; Ribchester, T.R.; Davies, N.J.; Venturi, P.; Jones, L.A.; Lewis, S.W.; Sham, P.C.; *et al.* Heritability estimates for psychotic disorders: The maudsley twin psychosis series. *Arch. Gen. Psychiatry* **1999**, *56*, 162–168.
4. Franzek, E.; Beckmann, H. Different genetic background of schizophrenia spectrum psychoses: A twin study. *Am. J. Psychiatry* **1998**, *155*, 76–83.
5. Svrakic, D.M.; Zorumski, C.F.; Svrakic, N.M.; Zwir, I.; Cloninger, C.R. Risk architecture of schizophrenia: The role of epigenetics. *Curr. Opin. Psychiatry* **2013**, *26*, 188–195.
6. Van Os, J.; Kenis, G.; Rutten, B.P. The environment and schizophrenia. *Nature* **2010**, *468*, 203–212.
7. Giusti-Rodriguez, P.; Sullivan, P.F. The genomics of schizophrenia: Update and implications. *J. Clin. Invest.* **2013**, *123*, 4557–4563.
8. Sullivan, P.F. The genetics of schizophrenia. *PLoS Med.* **2005**, *2*, e212.
9. St. Clair, D.; Blackwood, D.; Muir, W.; Carothers, A.; Walker, M.; Spowart, G.; Gosden, C.; Evans, H.J. Association within a family of a balanced autosomal translocation with major mental illness. *Lancet* **1990**, *336*, 13–16.
10. Brandon, N.J.; Millar, J.K.; Korth, C.; Sive, H.; Singh, K.K.; Sawa, A. Understanding the role of disc1 in psychiatric disease and during normal development. *J. Neurosci.* **2009**, *29*, 12768–12775.

11. Ripke, S.; O'Dushlaine, C.; Chambert, K.; Moran, J.L.; Kahler, A.K.; Akterin, S.; Bergen, S.E.; Collins, A.L.; Crowley, J.J.; Fromer, M.; *et al.* Genome-wide association analysis identifies 13 new risk loci for schizophrenia. *Nat. Genet.* **2013**, *45*, 1150–1159.

12. Consortium, I.S.; Purcell, S.M.; Wray, N.R.; Stone, J.L.; Visscher, P.M.; O'Donovan, M.C.; Sullivan, P.F.; Sklar, P. Common polygenic variation contributes to risk of schizophrenia and bipolar disorder. *Nature* **2009**, *460*, 748–752.

13. O'Donovan, M.C.; Craddock, N.; Norton, N.; Williams, H.; Peirce, T.; Moskvina, V.; Nikolov, I.; Hamshere, M.; Carroll, L.; Georgieva, L.; *et al.* Identification of loci associated with schizophrenia by genome-wide association and follow-up. *Nat. Genet.* **2008**, *40*, 1053–1055.

14. Papiol, S.; Malzahn, D.; Kästner, A.; Sperling, S.; Begemann, M.; Stefansson, H.; Bickeboller, H.; Nave, K.A.; Ehrenreich, H. Dissociation of accumulated genetic risk and disease severity in patients with schizophrenia. *Transl. Psychiatry* **2011**, *1*, e45.

15. Papiol, S.; Mitjans, M.; Assogna, F.; Piras, F.; Hammer, C.; Caltagirone, C.; Arias, B.; Ehrenreich, H.; Spalletta, G. Polygenic determinants of white matter volume derived from gwas lack reproducibility in a replicate sample. *Transl. Psychiatry* **2014**, doi:10.1038/tp.2013.126.

16. Terwisscha van Scheltinga, A.F.; Bakker, S.C.; van Haren, N.E.; Derks, E.M.; Buizer-Voskamp, J.E.; Boos, H.B.; Cahn, W.; Hulshoff Pol, H.E.; Ripke, S.; Ophoff, R.A.; *et al.* Genetic schizophrenia risk variants jointly modulate total brain and white matter volume. *Biol. Psychiatry* **2013**, *73*, 525–531.

17. Stepniak, B.; Papiol, S.; Hammer, C.; Ramin, A.; Everts, S.; Hennig, L.; Begemann, M.; Ehrenreich, H. Accumulated environmental risk determining the onset of schizophrenia. Submitted for publication, 2014.

18. Begemann, M.; Grube, S.; Papiol, S.; Malzahn, D.; Krampe, H.; Ribbe, K.; Friedrichs, H.; Radyushkin, K.A.; El-Kordi, A.; Benseler, F.; *et al.* Modification of cognitive performance in schizophrenia by complexin 2 gene polymorphisms. *Arch. Gen. Psychiatry* **2010**, *67*, 879–888.

19. Ribbe, K.; Friedrichs, H.; Begemann, M.; Grube, S.; Papiol, S.; Kästner, A.; Gerchen, M.F.; Ackermann, V.; Tarami, A.; Treitz, A.; *et al.* The cross-sectional gras sample: A comprehensive phenotypical data collection of schizophrenic patients. *BMC Psychiatry* **2010**, *10*, 91.

20. Van Os, J.; Kapur, S. Schizophrenia. *Lancet* **2009**, *374*, 635–645.

21. Grube, S.; Gerchen, M.F.; Adamcio, B.; Pardo, L.A.; Martin, S.; Malzahn, D.; Papiol, S.; Begemann, M.; Ribbe, K.; Friedrichs, H.; *et al.* A cag repeat polymorphism of kcnn3 predicts sk3 channel function and cognitive performance in schizophrenia. *EMBO Mol. Med.* **2011**, *3*, 309–319.

22. Kästner, A.; Grube, S.; El-Kordi, A.; Stepniak, B.; Friedrichs, H.; Sargin, D.; Schwitulla, J.; Begemann, M.; Giegling, I.; Miskowiak, K.W.; *et al.* Common variants of the genes encoding erythropoietin and its receptor modulate cognitive performance in schizophrenia. *Mol. Med.* **2012**, *18*, 1029–1040.

23. Gong, Y.G.; Wu, C.N.; Xing, Q.H.; Zhao, X.Z.; Zhu, J.; He, L. A two-method meta-analysis of neuregulin 1(nrg1) association and heterogeneity in schizophrenia. *Schizophr. Res.* **2009**, *111*, 109–114.

24. Li, D.; Collier, D.A.; He, L. Meta-analysis shows strong positive association of the neuregulin 1 (nrg1) gene with schizophrenia. *Hum. Mol. Genet.* **2006**, *15*, 1995–2002.

25. Munafo, M.R.; Thiselton, D.L.; Clark, T.G.; Flint, J. Association of the nrg1 gene and schizophrenia: A meta-analysis. *Mol. Psychiatry* **2006**, *11*, 539–546.

26. Papiol, S.; Begemann, M.; Rosenberger, A.; Friedrichs, H.; Ribbe, K.; Grube, S.; Schwab, M.H.; Jahn, H.; Gunkel, S.; Benseler, F.; *et al.* A phenotype-based genetic association study reveals the contribution of neuregulin1 gene variants to age of onset and positive symptom severity in schizophrenia. *Am. J. Med. Genet. Part B Neuropsychiatry Genet.* **2011**, *156*, 340–345.

27. Hagemeyer, N.; Goebbels, S.; Papiol, S.; Kästner, A.; Hofer, S.; Begemann, M.; Gerwig, U.C.; Boretius, S.; Wieser, G.L.; Ronnenberg, A.; *et al.* A myelin gene causative of a catatonia-depression syndrome upon aging. *EMBO Mol. Med.* **2012**, *4*, 528–539.

28. Hammer, C.; Stepniak, B.; Schneider, A.; Papiol, S.; Tantra, M.; Begemann, M.; Siren, A.L.; Pardo, L.A.; Sperling, S.; Mohd Jofrry, S.; *et al.* Neuropsychiatric disease relevance of circulating anti-nmda receptor autoantibodies depends on blood-brain barrier integrity. *Mol. Psychiatry* **2013**, doi:10.1038/mp.2013.110.

29. Ehrenreich, H.; Nave, K.-A. Max Planck Institute of Experimental Medicine, and DFG Center for Nanoscale Microscopy and Molecular Physiology of the Brain (CNMPB), Hermann-Rein-Str.3, 37075 Göttingen, Germany. Unpublished work, 2014.

30. Denny, J.C.; Bastarache, L.; Ritchie, M.D.; Carroll, R.J.; Zink, R.; Mosley, J.D.; Field, J.R.; Pulley, J.M.; Ramirez, A.H.; Bowton, E.; *et al.* Systematic comparison of phenome-wide association study of electronic medical record data and genome-wide association study data. *Nat. Biotechnol.* **2013**, *31*, 1102–1111.