Connecting myelin-related and synaptic dysfunction in schizophrenia with SNP-rich gene expression hubs

Hedi Hegyi

Combining genome-wide mapping of SNP-rich regions in schizophrenics and gene expression data in all brain compartments across the human life span revealed that genes with promoters most frequently mutated in schizophrenia are expression hubs interacting with far more genes than the rest of the genome. We summed up the differentially methylated “expression neighbors” of genes that fall into one of 108 distinct schizophrenia-associated loci with high number of SNPs. Surprisingly, the number of expression neighbors of the genes in these loci were 35 times higher for the positively correlating genes (32 times higher for the negatively correlating ones) than for the rest of the ~16000 genes. While the genes in the 108 loci have little known impact in schizophrenia, we identified many more known schizophrenia-related important genes with a high degree of connectedness (e.g. MOBP, SYNGR1 and DGCR6), validating our approach. Both the most connected positive and negative hubs affected synapse-related genes the most, supporting the synaptic origin of schizophrenia. At least half of the top genes in both the correlating and anti-correlating categories are cancer-related, including oncogenes (RRAS and ALDOA), providing further insight into the observed inverse relationship between the two diseases.

Gene expression correlation, protein-protein interaction and other high-throughput experiments in the post-genomic era have revealed that genes tend to form complex, scale-free networks where most genes have a few connections with others and a few have a high number of interactions, commonly referred to as “hubs”, establishing them as important central genes in these gene networks. These highly interconnected genes have become the targets of intense research expecting them to play prominent roles in genetic diseases. However, measurements found only a weak correlation between disease genes and hubs, e.g. Barabasi et al. found that disease genes have 32% more interactions with other proteins than non-disease genes, arguing that genetic mutations in topologically central, widely expressed genes are more likely to result in severe impairment of normal development, leading to lethality in utero and eventual deletion from the population.

This is apparently not the case in schizophrenia, a disease that steadily affects about one percent of the population despite the lower fecundity of the affected individuals. In most cases of schizophrenia no known mutations exist in protein-coding genes and the role of gene expression dysregulation in schizophrenia is increasingly recognized.

A recent meta-analysis, examining the mutations in the genomes of 36,000 schizophrenics, identified 108 distinct genomic regions with significantly higher mutation rates. We integrated this set with gene expression data from the Allen Brain Atlas. After determining the correlating and anti-correlating neighbors of genes in the 108 loci we found that the median number of correlating neighbors were about 35 times higher for the genes in the 108 regions than in the rest of the genome (32 times for the anti-correlating pairs).

We also integrated the data with a recent methylome study in schizophrenics. Ranking the genes for the hypermethylated probes of the positively correlating gene pairs identified the top gene as SYNGR1, a synapse-related gene whose regulatory region overlaps with one of the 108 SNP-rich loci in ref. 7. Ranking the genes for the hypermethylated probes of the negatively correlating gene pairs identified MBP and MOBP. They

CEITEC - Central European Institute of Technology, Masaryk University, 62500 Brno, Czech Republic. Correspondence and requests for materials should be addressed to H.H. (email: hegyihedi@gmail.com)
are both myelin-related, anti-correlating with a large number of synapse-related and glutamate receptor genes, offering a model that connects these frequently observed but so far disjoint pathologies in schizophrenia.

Results

**Genes in the most mutated genomic regions have the highest number of correlating partners.** In the first step we identified all human gene expression pairs with a Pearson correlation of $>0.8$ or $<-0.7$ in human brain tissues using the Allen Brain Atlas (website: brain-map.org$^8,9$) and a recent database with pre-calculated Pearson correlations for all relevant gene pairs at http://www.szdb.org/download.html#coexpression by Wu et al.$^{11}$ This resulted in 1,257,407 positively correlating and 1,108,585 negatively correlating unique gene pairs belonging to 16829 and 16761 individual genes, respectively (Supplementary Tables 1 and 2).

Subsequently we mapped all known human genes in 108 highly mutated human genomic regions identified in ref. 7 separately for genes whose promoters or only cis regulatory elements fall into these 108 loci. Filtering for genes present in the correlating or anti-correlating pairs by Wu et al.$^{11}$ resulted in 254 promoter-selected and 462 cis-selected genes (indicated by 1p/1c in Supplementary Tables 1 and 2, respectively).

Counting the positively and negatively correlating partners for all the genes except in the 108 loci in ref. 7 resulted in a median number of 71 positively and 63 negatively correlating pairs, respectively (Fig. 1). Unexpectedly, the median numbers of correlating pairs for the promoter-selected genes were 2472 and 2013, corresponding to a 35-fold and 32-fold increase for the positive and negative pairs, respectively.

The number of correlating pairs for the cis-selected genes also proved to be higher (386 and 320 pairs on average for positively and negatively correlating pairs, respectively), corresponding to an approximately 2.5-fold increase when compared to the general gene population. The median values were similar to the average gene population in this category (Fig. 1) but an unpaired t-test showed that the cis-derived genes and the general gene population (excluding the 108-loci genes) are significantly different ($p$-value $=0.012$).

We also investigated the number of differentially methylated gene pairs using a table of 56001 differentially methylated probes from a genome-wide methylome study in schizophrenics$^{10}$. We counted separately the number of hypermethylated and hypomethylated genes for the highly correlating and anti-correlating gene pairs and also the total sum of hyper- and hypomethylated probes belonging to these genes. The sums of scores for both the number of differentially methylated gene pairs and the total sum of their differentially methylated probes are shown in Supplementary Tables 1 and 2, for the positively and negatively correlating pairs, respectively.

To determine which measurement distinguishes the best between schizophrenia-related and non-specific genes we calculated the effect size for the 108-loci genes and also for genes annotated by Genecards and Malacards$^{12,13}$ as schizophrenia-related (Fig. 2), using Cliff’s delta for five different measurements (see legend for details).

Figure 2a and b shows the effect sizes for all relevant pairs of complementary gene sets (i.e. promoter-derived vs. non-promoter-derived, Genecards vs. non-Genecards genes, etc.) comparing the numbers of correlating gene partners in one set to the numbers of the gene partners in the other gene set. As expected the promoter-derived 108-loci genes had the greatest effect size, followed by Malacards and Genecards, the 108-loci cis-derived genes having the smallest effect. The results were similar for the negatively correlating pairs (Fig. 2b). Here we also filtered for gene pairs that were both differentially methylated in ref. 10, which increased the effect size by 0.039 for "pairs" in Fig. 2 for the promoter-derived positive set and by 0.061 for the same for the negative set. (Unfiltered charts shown in Supplementary Figure 1A,B).

**Specific genes with high numbers of correlating partners.** We ranked all genes according to the number of hypermethylated probes (taken from ref. 10) of the correlating gene partners. Tables 1 and 2 show the
The top 20 genes with the most hypermethylated probes, in the positively and negatively correlating partners, respectively. The top-ranking gene in Table 1 is SYNGR1, synaptogyrin, followed by GRIN1, a glutamate receptor and CHRN2, a cholinergic receptor. All three genes have synapse-related functions, SYNGR1 and GRIN1 regulate synaptic plasticity whereas CHRN2 regulates synapse assembly.

SYNGR1 is also present in the 108 most mutated loci in ref. 7 although only for its regulatory region, not for its promoter. Further functional analysis identified 5 more synapse-related genes in Table 1: NISCH regulates synaptic transmission, RIMS1 and STX1A regulate postsynaptic potential, CPLX2 regulates synaptic plasticity and GRIA1 regulates long-term synaptic depression.

Altogether 11 out of the 20 genes in Table 1 are annotated in Genecards as schizophrenia-related and 14 appear in the 108 genomic regions in ref. 7. The only gene in Table 1 that appears in neither is ARHGEF11, a glutamate transport enhancer, however it had a significantly higher expression in the thalamus of schizophrenics in ref. 15.

Table 2 lists the top 20 genes with the highest number of hypermethylated probes for the negatively correlating partners. The top-ranking gene, DBI, is a GABA receptor modulator; its role having been contemplated in schizophrenia the authors concluded that DBI might have a symptom modulatory rather than an etiological role in schizophrenia. Unexpectedly, MOBP, myelin-associated oligodendrocyte basic protein and MBP, myelin basic protein, both appear in Table 2. While myelin-related abnormalities are one of the hallmarks of schizophrenia, surprisingly, not much functional information is available about MOBP beyond its role in the formation of the myelin sheath and a knockout mouse was phenotypically indistinguishable from the wild type. Nevertheless, the authors argue that MOBP probably has a so far undiscovered function, due to the conservation of several alternatively spliced variants in rat and mouse.

Altogether, 11 of the 20 genes with the greatest numbers of hypermethylated probes for their anti-correlating gene partners (Table 2) are annotated as schizophrenia-related in Genecards and 10 appear in the 108 loci in ref. 7.
| gene | pairs hypo meth | pairs hyper meth | pairs hypo Sum | pairs hyper Sum | probes hypo | probes hyper | genes cards Score | malacards Score | prom/cis? | prot len | description |
|------|----------------|-----------------|---------------|----------------|-------------|-------------|-----------------|----------------|-----------|---------|------------|
| SYNGR1 | 4336 | 1437 | 2086 | 3553 | 8290 | 0 | 8 | 8.57 | NA | 1c | 233 | synaptogyrin 1 |
| GRIN1 | 3800 | 1261 | 1915 | 3302 | 7959 | 2 | 1 | 9.07 | 7.55 | NA | 959 | glutamate receptor, ionotropic, NMDA 1 |
| CHRNB2 | 4070 | 1391 | 1992 | 3302 | 7880 | 1 | 3 | 9.07 | 5.10 | NA | 502 | cholinergic receptor, nicotinic beta 2 |
| ARHGEF11* | 4024 | 1394 | 1951 | 3506 | 7832 | 0 | 3 | NA | NA | NA | 1562 | Rho guanine nucleotide exchange factor 11 |
| ALDOA* | 4949 | 1558 | 2208 | 3983 | 7811 | 2 | 5 | −2.01 | NA | 1p | 418 | aldolase, fructose-bisphosphate A |
| SREBF2* | 3750 | 1295 | 1810 | 3188 | 7626 | 2 | 18 | −1.79 | NA | 1c | 1141 | sterol regulatory element binding TF 2 |
| PITPNM2 | 3308 | 1202 | 1734 | 3188 | 7626 | 2 | 18 | −2.01 | NA | 1p | 1349 | phosphatidylinositol transfer protein |
| GRAMD1B | 4159 | 1428 | 1980 | 3418 | 7569 | 2 | 6 | NA | NA | NA | 745 | GRAM domain containing 1B |
| NISCH* | 4152 | 1331 | 1951 | 3506 | 7832 | 0 | 3 | −1.60 | NA | 1c | 1504 | nischarin |
| RIMS1* | 4679 | 1474 | 2067 | 3461 | 7465 | 1 | 0 | 9.07 | 3.83 | NA | 677 | chromogranin B |
| CPLX2 | 3433 | 1149 | 1677 | 3139 | 7345 | 3 | 5 | 9.07 | 2.48 | NA | 134 | complex 2 |
| CHGB* | 3465 | 1169 | 1746 | 3154 | 7316 | 1 | 0 | 9.07 | 3.83 | NA | 677 | chromogranin B |
| STX1A | 3941 | 1259 | 1880 | 3023 | 7283 | 1 | 6 | −0.60 | NA | NA | 288 | syntaxin 1A |
| SEZ6L2* | 3406 | 1169 | 1747 | 3046 | 7265 | 0 | 1 | NA | NA | NA | 923 | seizure related 6 homolog (mouse)-like 2 |
| LJMIB2L2 | 4406 | 1361 | 1972 | 3127 | 7227 | 1 | 0 | NA | NA | 1p | 705 | l(3)mbt-like 2 (Drosophila) |
| SLC45A1* | 3454 | 1160 | 1740 | 2988 | 7218 | 1 | 4 | NA | NA | NA | 782 | solute carrier family 45 member 1 |
| GRIA1* | 4337 | 1377 | 1950 | 3256 | 7141 | 0 | 3 | 9.06 | NA | 1c | 916 | glutamate receptor, ionotropic, AMPA 1 |
| RANGAP1 | 3704 | 1232 | 1772 | 3074 | 7091 | 0 | 1 | NA | NA | 1p | 587 | Ran GTPase activating protein 1 |
| TCF20* | 3497 | 1242 | 1700 | 3258 | 7044 | 0 | 1 | NA | NA | NA | 1960 | transcription factor 20 (ARI1) |
| EPC2 | 5124 | 1514 | 2084 | 3436 | 6909 | 1 | 0 | NA | NA | NA | 807 | enhancer of polycomb homolog 2 |

Table 1. The top 20 positively correlating genes, ranked according to the hypermethylated probes (listed in column 6) in the correlating gene partners. The total number of positively correlating genes (column 2), of which hypo- or hypermethylated (columns 3 and 4), are also listed. Genecards and Malacards scores are listed in columns 7 and 8. The column “prom/cis?” indicates if the promoter of the gene ("1p") or the cis regulatory region of the gene ("1c") falls into the 108 highly mutated loci in ref. 7. A “*” next to the gene identifier indicates a role in cancer.

Interestingly, several genes in both tables are cancer-related (indicated with a “*” next to the gene identifiers in Tables 1 and 2). Among the top 20 genes for hypermethylated probes of correlating partners are ARHGEF1119,20 and SREPP2, both associated with prostate and breast cancer21,22, while ALDOA functions as an oncogene in highly metastatic pancreatic cancer23. NISCH has a tumor-suppressive function in breast cancer24,25, CHGB is associated with aggressive VHL-associated pancreatic neuroendocrine tumors26. SEZ6L2 is a prognostic marker for lung cancer27, SLC45A1 is deleted in neuroblastoma28. GRIA1 might play a role in glioma proliferation29 while TCF20 is again associated with prostate30 and breast cancer31.

Among the top 20 genes for hypermethylated probes of anti-correlating partners are RAS, an oncogene, also involved in neuronal axon guidance32 and MBP, associated with oligodendrogliomas33. TMEM219 regulates apoptosis34, CDK2AP1 is a putative oral cancer suppressor35, PSMB10, a proteasome subunit, is upregulated via the NFKB1 pathway in cancer cells36. SI001A1 is also associated with several tumor types and inhibits apoptosis in ventricular cardiomyocytes37.

D GCR65 is associated with DiGeorge syndrome, a consequence of microdeletions in chromosomal region 22q11.2 and also has increased levels in metastatic mammary tumour cells38. FABP7 plays a role in neurogenesis and is a marker of glioma stem cells39. Epithelial splicing regulatory protein 2 (ESRP2) suppresses cancer cell motility39. ALDH1A1 is a marker and prognostic factor for several human cancers, bladder40, pancreatic41, lung42, colorectal and breast cancer43, among others.

Altogether, 10 genes among the top 20 for the greatest number of hypermethylated probes for the positively correlating gene partners (Table 1) and 13 of the 20 genes with the greatest numbers of hypermethylated probes for their anti-correlating gene partners (Table 2) are annotated as cancer-related in Genecards.
The connection between the positive and negative hubs. The existence of the two kinds of hub genes with high numbers of positive or negative correlating partners raises the question about their functionality and the underlying neurobiological pathways: are they related or do they form mostly separate networks? To answer this question we constructed three gene networks (Fig. 3): two representing the positive and negative correlations only for SYNGR1 and MOBP, respectively (Fig. 3a,b), and one (Fig. 3c) showing a combined network of the two. We chose SYNGR1 for the positive hub for it is the top ranking gene in Table 1 and is also located in one of the PGC regions whereas we chose MOBP because it ranks highly among the anti-correlating genes, indicating a putative regulatory function, which, however, has not been observed yet experimentally.

To reduce the size of the networks we selected only Malacards-annotated schizophrenia-related genes that are hypermethylated in ref. 10. SYNGR1 correlates positively with 49 genes (Fig. 3a); MOBP also has 49 negatively correlating partners (Fig. 3b). The combined network is shown in Fig. 3c. Strikingly, 41 genes are shared interacting partners between MOBP and SYNGR1. Each of the two hub genes interacts with only 8 genes that are not shared with the other hub. Clearly, the positive and negative correlations - and such interactions - form highly interconnected networks that provide synaptic functions, the core functionality of the human brain.

The shared genes reflect remarkably on the nature and most consistently observed features of schizophrenia: 6 shared genes are glutamate receptors (GRIA4, GRIK3, GRIK5, GRIN1, GRM1, GRM2), one GABA receptor (GABBR1), a cannabinoid receptor (CNR1), GAD2, a glutamate decarboxylase, two synapse-related (SYN2, SYT11) and several neuron-specific genes are also present among the shared genes (Fig. 3c).

We repeated the selection process replacing the Malacards-genes with the 108-loci genes. This resulted in a similar number of interacting genes that correlate positively with SYNGR1 (50 genes) and negatively with MOBP (48 genes). They share 39 common genes (Fig. 4). Biological processes derived from a Gene Ontology** analysis of the shared genes either in the Malacards set or the 108-loci set and in both cases shared between SYNGR1 and MOBP...
and significantly enriched (p-value < 0.05) for both sets are shown in Table 3. For both gene sets the top two biological processes with the highest significance are “synaptic transmission” and “modulation of synaptic transmission”.

Functional analysis of schizophrenia genes correlating with the top 20 hub genes. To get a further insight into the biological processes the top ranking genes and their network neighbors (i.e. their correlating and anti-correlating gene partners) partake in we took the top 20 genes with the most hypermethylated probes for the correlating (and anti-correlating) gene partners and selected those gene neighbors that correlated with minimum 19 of the top 20 genes. This resulted in 421 genes for the positively correlating set and 460 genes for the negatively correlating one.

For this in silico functional analysis we used an online tool provided by STRING to identify Gene Ontology terms that are over-represented in our gene set. The positive gene set was associated with 278 significantly enriched GO terms for biological processes while the negative set had 151 such GO terms. Interestingly, while the two gene sets share only 117 common genes i.e. about 1/4th of the total gene number for either set, they also share 117 common biological processes (Table 4), a much higher fraction for either set (42% of the positive set-derived biological processes and 77% of the negative set-derived processes are shared).

This result also shows that gene networks as a whole are remarkably redundant, i.e. similar functions are often carried out by several functional homologs in the brain. This network view of the genes associated with schizophrenia is also revelatory for the polygenic nature of schizophrenia: once a mutation perturbs the expression of a highly connected hub gene with many interacting partners, this in turn will lead to the perturbation of several hundred or even thousand interacting genes. The differential methylation of several thousands of genes in the prefrontal cortex of schizophrenics might be one such manifestation of the complexity of the disease.

The shared GO terms ranked by significance are listed in Table 4. The top process is “nervous system development”, followed by “synaptic transmission”, reflecting two major aspects of schizophrenia, i.e. it is considered a (neuro)developmental disease affecting mostly the synapse. On a different note, the most frequently occurring word in Table 4 is “regulation” (occurring 29 times), followed by “transport” (24x), perhaps underlining somewhat less pronounced aspects of schizophrenia.

Positive hub genes are longer, negative hub genes are shorter than average. We also calculated the protein length and protein disorder statistics for both the top 20 positive and negative hubs. Surprisingly, while the top 20 positive hub proteins were significantly longer (p-value < 10^-5) than the average human protein
Figure 4. (A) All positively correlating genes in the 108 loci. The network is again centered on SYNGR1. The red-colored edges show correlations with SYNGR1. (B) All the genes in the 108 loci that correlate negatively with MOBP (edges colored red) or one another (grey edges). (C) The combined network in (A and B). All light green-colored genes correlate positively with SYNGR1 and negatively with MOBP (edges show in red in both cases). The purple-colored and turquoise-colored genes correlate only with MOBP or SYNGR1 (of the two), respectively.

| GO_identifier | GO_Biological Process                        | shared Genes_108 loci | p_val (108 loci) | shared Genes Malacards | p_val (Malac Genes) |
|---------------|---------------------------------------------|------------------------|------------------|------------------------|---------------------|
| GO:0050804    | modulation of synaptic transmission         | 7                      | 0.00359          | 14                     | 1.23e-13            |
| GO:0007268    | synaptic transmission                        | 9                      | 0.00359          | 16                     | 8.05e-12            |
| GO:0061564    | axon development                             | 7                      | 0.0318           | 11                     | 2.14e-06            |
| GO:0031175    | neuron projection development                | 8                      | 0.0289           | 11                     | 1.08e-05            |
| GO:0048167    | regulation of synaptic plasticity           | 4                      | 0.0469           | 6                      | 4.71e-05            |
| GO:0048666    | neuron development                            | 9                      | 0.0233           | 11                     | 4.71e-05            |
| GO:0048812    | neuron projection morphogenesis              | 7                      | 0.0318           | 9                      | 0.000146            |
| GO:0007411    | axon guidance                                | 6                      | 0.0415           | 7                      | 0.00128             |
| GO:0030182    | neuron differentiation                        | 9                      | 0.0318           | 9                      | 0.00545             |
| GO:0044765    | single-organism transport                    | 15                     | 0.0318           | 15                     | 0.0062              |
| GO:0007399    | nervous system development                   | 14                     | 0.0138           | 12                     | 0.011               |
| GO:1902578    | single-organism localization                 | 16                     | 0.0289           | 15                     | 0.0111              |

Table 3. Biological processes shared between the Malacards gene set interacting with both SYNGR1 and MOBP in Fig. 3C (41 light green-colored genes) and the 108-loci genes interacting with both SYNGR1 and MOBP in Fig. 4C (39 light green-colored genes). 12 biological processes in the table are common in the two shared sets, despite the very limited number of actually shared genes (only NRGN is shared besides MOBP and SYNGR1 but these two were not included in the GO analysis).
| GO_identifier | GO_Biological Process | Pos Neighbors | p_val (Pos Neigh) | Neg Neighbors | p_val (Neg Neigh) |
|---------------|-----------------------|---------------|------------------|---------------|------------------|
| GO:0007399    | nervous system development | 88            | 5.21e-12         | 105           | 4.23e-17         |
| GO:0007268    | synaptic transmission   | 48            | 2.95e-14         | 49            | 7.45e-14         |
| GO:0048699    | generation of neurons  | 68            | 1.57e-11         | 70            | 3.36e-10         |
| GO:0048666    | neuron development     | 52            | 6.61e-12         | 52            | 3.78e-10         |
| GO:0031175    | neuron projection development | 48          | 2.83e-12         | 46            | 9.22e-10         |
| GO:0022008    | neurogenesis           | 68            | 1.62e-10         | 70            | 2.03e-09         |
| GO:0007267    | cell-cell signaling    | 57            | 2.83e-12         | 53            | 1.32e-08         |
| GO:0048731    | system development     | 99            | 0.000594         | 125           | 2.02e-08         |
| GO:0030182    | neuron differentiation | 56            | 5.4e-11          | 54            | 2.31e-08         |
| GO:0048667    | cell morphogenesis involved in neuron differentiation | 39    | 3.77e-10         | 38            | 2.98e-08         |
| GO:0048812    | neuron projection morphogenesis | 39  | 7.5e-10          | 38            | 5.45e-08         |
| GO:0061564    | axon development       | 37            | 1.74e-09         | 36            | 1.16e-07         |
| GO:0007409    | axonogenesis           | 36            | 2.41e-09         | 35            | 1.49e-07         |
| GO:0007275    | multicellular organismal development | 106 | 0.00347         | 132           | 7.65e-07         |
| GO:0030030    | cell projection organization | 51    | 2.12e-08         | 50            | 1.66e-06         |
| GO:0007610    | behavior               | 32            | 1.34e-06         | 34            | 1.8e-06          |
| GO:0048858    | cell projection morphogenesis | 42   | 1.1e-07          | 41            | 5.72e-06         |
| GO:0007156    | homophilic cell adhesion via plasma membrane adhesion molecules | 10    | 0.0277         | 17            | 7.86e-06         |
| GO:0048856    | anatomical structure development | 105 | 0.00619         | 128           | 1.01e-05         |
| GO:0050803    | regulation of synapse structure or activity | 28    | 2.83e-12         | 20            | 1.43e-05         |
| GO:0044767    | single-organism developmental process | 118 | 0.00225         | 138           | 2.61e-05         |
| GO:0098742    | cell-cell adhesion via plasma-membrane adhesion molecules | 11    | 0.0369         | 18            | 2.65e-05         |
| GO:0050804    | modulation of synaptic transmission | 27    | 2.94e-10         | 21            | 3.04e-05         |
| GO:0044708    | single-organism behavior | 26    | 1.07e-05         | 27            | 3.14e-05         |
| GO:1902580    | single-organism cellular localization | 37    | 0.000174        | 41            | 8.75e-05         |
| GO:0007626    | locomotory behavior    | 19            | 3.27e-06         | 18            | 0.000103         |
| GO:0032502    | developmental process   | 116           | 0.00628         | 136           | 0.000108         |
| GO:0030001    | metal ion transport    | 36            | 3.9e-07          | 33            | 0.00011          |
| GO:0051650    | establishment of vesicle localization | 16    | 7.43e-05         | 17            | 0.000111         |
| GO:0051640    | organelle localization | 25            | 5.23e-06         | 24            | 0.000151         |
| GO:1902578    | single-organism localization | 116 | 2.83e-12         | 98            | 0.000153         |
| GO:0096811    | ion transport          | 46            | 8.77e-05         | 49            | 0.000157         |
| GO:0007611    | learning or memory     | 16            | 0.000443         | 18            | 0.000177         |
| GO:0000902    | cell morphogenesis     | 48            | 1.22e-07         | 43            | 0.000182         |
| GO:0030534    | adult behavior         | 15            | 3.73e-05         | 15            | 0.000195         |
| GO:0044765    | single-organism transport | 110           | 7.01e-12         | 93            | 0.000195         |
| GO:0050890    | cognition              | 17            | 0.000535         | 19            | 0.000241         |
| GO:0007417    | central nervous system development | 35    | 0.0013         | 40            | 0.000252         |
| GO:0098660    | inorganic ion transmembrane transport | 34    | 1.76e-06         | 31            | 0.000405         |
| GO:0051641    | cellular localization  | 76            | 1.56e-06         | 73            | 0.000436         |
| GO:0031344    | regulation of cell projection organization | 28    | 7.75e-06         | 26            | 0.000577         |
| GO:0051656    | establishment of organelle localization | 22    | 6.8e-06         | 20            | 0.000577         |
| GO:0061024    | membrane organization  | 37            | 0.000377         | 39            | 0.000832         |
| GO:0051960    | regulation of nervous system development | 38    | 6.8e-07         | 33            | 0.000962         |
| GO:0048468    | cell development       | 65            | 5.66e-07         | 59            | 0.00104          |

Continued
| GO_identifier  | GO_Biological Process                        | Pos Neighbors | p_val (Pos Neigh) | Neg Neighbors | p_val (Neg Neigh) |
|----------------|---------------------------------------------|---------------|-------------------|---------------|-------------------|
| GO:0032989     | cellular component morphogenesis            | 49            | 5.34e-07          | 43            | 0.00118           |
| GO:0007420     | brain development                           | 28            | 0.00376           | 32            | 0.0012            |
| GO:0006810     | transport                                    | 125           | 2.83e-12          | 102           | 0.00145           |
| GO:0043087     | regulation of GTPase activity               | 29            | 6.67e-05          | 28            | 0.00148           |
| GO:0010975     | regulation of neuron projection development  | 24            | 6.8e-06           | 21            | 0.00152           |
| GO:0045664     | regulation of neuron differentiation         | 28            | 2.54e-05          | 26            | 0.00152           |
| GO:0007613     | memory                                      | 8             | 0.032            | 11            | 0.00167           |
| GO:0051234     | establishment of localization               | 126           | 6.68e-12          | 104           | 0.0017            |
| GO:0030154     | cell differentiation                         | 89            | 0.00066          | 94            | 0.00201           |
| GO:0048869     | cellular developmental process              | 91            | 0.00124          | 97            | 0.00259           |
| GO:0068586     | neurotransmitter transport                  | 11            | 0.00842          | 13            | 0.00268           |
| GO:0044700     | single organism signaling                   | 139           | 1.56e-06          | 134           | 0.00275           |
| GO:0051179     | localization                                | 146           | 2.83e-12          | 120           | 0.00276           |
| GO:0071702     | organic substance transport                 | 66            | 3.73e-05          | 64            | 0.00276           |
| GO:0060322     | head development                            | 28            | 0.00956          | 32            | 0.00331           |
| GO:0065008     | regulation of biological quality            | 94            | 1.07e-05          | 90            | 0.004             |
| GO:1902582     | single-organism intracellular transport      | 41            | 0.00307          | 44            | 0.00444           |
| GO:0043547     | positive regulation of GTPase activity      | 27            | 9.52e-05          | 25            | 0.00452           |
| GO:0007154     | cell communication                          | 143           | 6.53e-07          | 135           | 0.00462           |
| GO:0007411     | axon guidance                               | 25            | 3.26e-05          | 22            | 0.00462           |
| GO:0034220     | ion transmembrane transport                 | 41            | 3.01e-06          | 35            | 0.00483           |
| GO:0050767     | regulation of neurogenesis                  | 31            | 4.27e-05          | 28            | 0.00483           |
| GO:0007269     | neurotransmitter secretion                  | 9             | 0.0209           | 11            | 0.00491           |
| GO:0070838     | divalent metal ion transport                | 16            | 0.00343          | 17            | 0.00515           |
| GO:0023061     | signal release                              | 14            | 0.00137          | 14            | 0.00516           |
| GO:0043269     | regulation of ion transport                 | 39            | 7.33e-09          | 28            | 0.00569           |
| GO:0051649     | establishment of localization in cell       | 68            | 9.18e-07          | 59            | 0.00675           |
| GO:0068112     | cation transport                            | 35            | 0.000182         | 33            | 0.00683           |
| GO:0035637     | multicellular organismal signaling          | 10            | 0.00233          | 10            | 0.00764           |
| GO:0050806     | positive regulation of synaptic transmission| 13            | 1.78e-05          | 10            | 0.00826           |
| GO:0044093     | positive regulation of molecular function   | 55            | 0.00211          | 57            | 0.00829           |
| GO:0051336     | regulation of hydrolase activity            | 51            | 2.23e-05          | 46            | 0.00829           |
| GO:0048167     | regulation of synaptic plasticity           | 17            | 1.64e-07          | 11            | 0.00978           |
| GO:0098662     | inorganic cation transmembrane transport    | 30            | 4.56e-06          | 24            | 0.0106            |
| GO:0008088     | axon cargo transport                         | 9             | 5.65e-06          | 6             | 0.0132            |
| GO:0072384     | organelle transport along microtubule       | 7             | 0.000645         | 6             | 0.0132            |
| GO:0008104     | protein localization                         | 62            | 2.35e-05          | 56            | 0.0139            |
| GO:0045665     | negative regulation of neuron differentiation| 10            | 0.041            | 12            | 0.0139            |
| GO:0007264     | small GTPase mediated signal transduction   | 33            | 7.44e-05          | 29            | 0.0144            |
| GO:0023051     | regulation of signaling                     | 76            | 0.00514          | 80            | 0.0146            |
| GO:0031345     | negative regulation of cell projection      | 10            | 0.00473          | 10            | 0.0146            |
| GO:0001764     | neuron migration                            | 11            | 0.00124          | 10            | 0.0154            |
| GO:0015031     | protein transport                            | 51            | 1.07e-05          | 44            | 0.0154            |
| GO:0043085     | positive regulation of catalytic activity   | 49            | 0.00174          | 49            | 0.0159            |
| GO:0098655     | cation transmembrane transport              | 33            | 1.16e-05          | 27            | 0.0159            |
| GO:0010666     | regulation of cell communication            | 88            | 5.43e-05          | 83            | 0.0163            |
| GO:0010977     | negative regulation of neuron projection    | 8             | 0.0226           | 9             | 0.0163            |

Continued
(894 and 442 amino acids, on average, respectively), the top 20 negative hub proteins were significantly shorter (327 amino acids on average). We did not find their disorder to be significantly different from that of the general human protein population. Likewise, the 108 region-derived proteins among the top 20 positive and negative hubs in Tables 1 and 2, were also longer (14 proteins in Table 1a with an average length of 983 amino acids) and shorter (10 proteins in Table 1b with an average length of 261), respectively, than the average human protein (442 aa).

Discussion
It has been known for at least a decade that myelin has an inhibitory role in axonal regeneration in the CNS. Myelin is dysfunctional in schizophrenia and this dysfunctionality leads to changes in synaptic formation and function, another hallmark of schizophrenia. Several studies have identified genes whose expression is abnormal in schizophrenic brains affecting myelin-related and synapse-related biochemical pathways. However, this is the first time, to our knowledge, that a model for a complete network of gene interactions is presented that would account for both of these recurring anomalies in schizophrenia.

Our gene network has two hubs, one, SYN1, with a synaptic function, correlates positively and apparently interacts with a high number of genes that also interact with one another and at the same time interact negatively with MOBP, a myelin gene. Myelin is known to inhibit axonal sprouting, a step considered important for synaptic formation. While neither MOBP, nor MBP are known to have such inhibitory functions, another myelin-related protein, RTN4 (also called Nogo-A), ranking 88th among the negatively correlating genes (Supplementary Table 2), does have such a function, inhibiting axon growth. It is tempting to hypothesize that either MOBP or MBP also have such a - so far undiscovered - inhibitory function. MOBP is the 3rd most abundant protein in the CNS myelin and has several alternatively spliced variants. As highlighted by Montague et al., a physiological function for MOBP has not been found yet. Our model with two antagonistic hub genes would also account for the antagonistic relationship between myelin genes and synapse-related functions.

We also found a high representation of cancer-related genes in both the top correlating and anti-correlating genes (Tables 1 and 2). While only limited association has been found so far connecting schizophrenia- and cancer-related gene hubs, there are several studies pinpointing the essential role of hub proteins in cancer. Only a handful of papers have been published on gene networks in schizophrenia, e.g. a study highlighting shared biological processes shared between the positively correlating and negatively correlating gene neighbors of the top 20 genes in Table 1 and Table 2.

Table 4. Biological processes shared between the positively correlating and negatively correlating gene neighbors of the top 20 genes in Table 1 and Table 2.

| GO_identifier | GO_Biological Process                  | Pos Neighbors | p_val (Pos Neigh) | Neg Neighbors | p_val (Neg Neigh) |
|---------------|---------------------------------------|---------------|-------------------|--------------|-------------------|
| GO:0045184    | establishment of protein localization | 53            | 1.32e-05          | 46           | 0.0166            |
| GO:001500     | action potential                       | 9             | 0.00602           | 9            | 0.0172            |
| GO:0006816    | calcium ion transport                  | 14            | 0.00545           | 14           | 0.019             |
| GO:0044802    | single-organism membrane organization | 28            | 0.00842           | 29           | 0.0227            |
| GO:0033036    | macromolecule localization             | 69            | 5.82e-05          | 63           | 0.0246            |
| GO:0055085    | transmembrane transport                | 49            | 1.53e-06          | 39           | 0.0251            |
| GO:0035556    | intracellular signal transduction      | 65            | 3.26e-05          | 58           | 0.0265            |
| GO:0046907    | intracellular transport                | 47            | 0.00101           | 45           | 0.0265            |
| GO:0051899    | membrane depolarization                | 7             | 0.01             | 7            | 0.0265            |
| GO:0065009    | regulation of molecular function       | 80            | 0.00064           | 78           | 0.027             |
| GO:0007612    | learning                               | 10            | 0.001            | 10           | 0.0298            |
| GO:004849     | synaptic vesicle transport             | 10            | 0.00296           | 9            | 0.0324            |
| GO:0050773    | regulation of dendrite development     | 11            | 0.00066           | 9            | 0.0324            |
| GO:0032940    | secretion by cell                     | 22            | 0.00835           | 22           | 0.0356            |
| GO:0051345    | positive regulation of hydrolase activity | 41           | 3.94e-06          | 32           | 0.0358            |
| GO:0060284    | regulation of cell development         | 36            | 5.66e-05          | 30           | 0.0365            |
| GO:0097479    | synaptic vesicle localization          | 11            | 0.000782          | 9            | 0.0365            |
| GO:0050790    | regulation of catalytic activity       | 67            | 0.00203           | 66           | 0.0383            |
| GO:0050770    | regulation of axonogenesis             | 10            | 0.014            | 10           | 0.0405            |
| GO:0034765    | regulation of ion transmembrane transport | 31            | 1.05e-08          | 19           | 0.0418            |
| GO:0050807    | regulation of synapse organization     | 14            | 6.24e-06          | 9            | 0.0426            |
| GO:0035725    | sodium ion transmembrane transport     | 12            | 0.00121           | 10           | 0.0438            |
| GO:0016043    | cellular component organization        | 125           | 0.000226          | 121          | 0.0456            |
| GO:0047496    | vesicle transport along microtubule     | 4             | 0.0215           | 4            | 0.0457            |
| GO:0051128    | regulation of cellular component organization | 76           | 3.59e-06          | 64           | 0.0495            |
hubs and gene networks between schizophrenia and diabetes and our own work on microRNA-driven networks and hub genes in schizophrenia. Perhaps the most relevant study connecting cancer- and schizophrenia-related gene networks is by Ibanez et al. who found an inverse comorbidity between the two diseases, observing expression deregulations in opposite directions of the same genes and gene networks.

Our model is also supported by the extraordinary enrichment of correlating partners for the genes encoded in the 108 genomic regions with the highest mutation rates in the genomes of schizophrenics identified in ref. 7. As mentioned above, the median number of correlating partners for those genes whose promoters are located in the 108 loci is 35 times higher for the positively correlating genes than for the rest of the genes in this study and 32 times higher for the negatively correlating genes. The effect size for these comparisons is in the range of 0.8–0.9, making the number of interacting partners, reflecting the centrality of a gene, the single most important factor when considering the biological significance of the individual schizophrenia-related genes and their contribution to the disease. There is a significant but smaller enrichment (2.5 times on average) for those 108-loci genes whose promoters are not, only their enhancers are present in these loci, which raises the intriguing possibility that for most genes in this set the mutations fall into or close to the promoter regions, compromising their functionality as in most cases of schizophrenia the protein sequences are not corrupted by mutations.

The assumption that it is the regulatory regions, not the protein-coding regions that are affected mostly in schizophrenia is also apparent in the fact that despite the strong centrality (“hubness”) of the affected genes we do find surviving phenotypes, which are the patients, exactly. As raised in the introduction, hub genes with a mutation in the coding region would make these mutations lethal in most instances.

The robustness of our findings is also supported by the fact that when we replace the Malacards genes with the 108-loci genes in Fig. 3 (sharing only SYN, SRR and NRGN as indicated in Fig. 3) and analyze the Gene Ontology terms for the 39 genes correlated positively by SYN, SRR and negatively by MOBP, we find that the most significant biological process associated with this gene set is again “synaptic transmission”. This also shows the remarkable redundancy of the gene networks in the human brain. Altogether, beyond providing an intriguing new model for schizophrenia, with more details for the underlining gene networks than before, it is also fascinating and quite fitting that synaptic transmission, perhaps the most complex and dynamic part of the human brain also entails the most genetic complexity, i.e. the most connected gene networks with the highest number of correlating/interacting gene partners.

Methods

Determining pairwise gene expression correlations in the human brain. In the first step all pairwise Pearson correlations were determined for those genes expressed in the brain that have expression data in the Allen Brain Atlas (website: brain-map.org). Gene expression was measured for 50,000 genes in 524 different tissues taken from several compartments of the brains of several individuals spanning the human lifetime between 2 weeks of post-conception and 40 years of age. We used an in-house Perl script to calculate pairwise correlations complemented by correlation data taken from SZDB.org. We filtered the results keeping only pairwise Pearson correlation that were either minimum 0.8 or maximum −0.7. If there were several values for the same gene pair we used the most extreme ones, the highest and lowest values for the positive and negative correlations, respectively.

We mapped all known human genes in 108 highly mutated human genomic regions identified in ref. 7 separately for genes whose promoters or only cis regulatory elements fall into these 108 loci. We determined the location of the promoters and cis elements in the human genome using a study by Thurman et al.

Counting correlating gene pairs and their differentially methylated probes. For each gene whose expression correlated (r ≥ 0.8) or anti-correlated (r ≤ −0.7) with other genes in our data set (16830 and 16762 genes with correlating and anti-correlating partners, respectively) we counted the number of correlating and anti-correlating partners. Using a methylome data set in ref. 10 that recorded 56001 differentially methylated probes between two subgroups of schizophrenia patients we also counted gene partners for each gene that were hypermethylated or hypomethylated (with at least one hypermethylated or hypomethylated probe, respectively) and also the total number of hyper- or hypomethylated probes for the correlating and anti-correlating gene partners. The genes were ranked for the total sum of hypermethylated probes of the ‘neighboring’ (i.e. either correlating or anti-correlating) genes as this ranking placed SYN, a gene present with its regulatory region in the 108 SNP rich regions in schizophrenics, identified by the Psychiatric Genomics Consortium (PGC) to the top of the list (in Table 1). This ranking also produced 14 genes out of the top 20 genes in Table 1 with either its promoter or SNP rich regions in schizophrenics, identified by the Psychiatric Genomics Consortium (PGC) to the top of the

Determining the effect size to distinguish between schizophrenia-related and unrelated genes. To determine which measurement distinguishes the best between schizophrenia-related and non-specific genes we calculated the effect size for the 108-loci genes and also for genes annotated by Genecards and Malacards as schizophrenia-related (Fig. 2) regarding each specific feature mentioned in the previous paragraph (i.e. the number of correlating genes, the number of hypermethylated, hypomethylated correlating genes and the total sum of the differentially methylated probes of the correlating pairs). We repeated the calculations for the negatively correlating pairs as well. To calculate the effect size we used Cliff’s delta. Cliff’s delta is defined as

\[
d = (\#(x_i > x_j) - \#(x_i < x_j))/m \cdot n
\]
where the two distributions are of size m and n with items $x_i$ and $x_j$ (with $i$ running from 1 to m and j running from 1 to n) respectively, and # is defined as the number of times. Cliff’s delta shows that out of all possible pairwise comparisons between the numbers in set A and set B in what proportion will the numbers in set A be bigger than in set B. Cliff’s delta does not require any assumptions about distribution types.

**Functional analysis with the Gene Ontology module of STRING.** To carry out functional analysis of the top 20 genes and their network neighbors we used the Gene Ontology (GO) module of the STRING webserver. The server lists all functional categories that are significantly enriched in the provided gene set, and supplies the corresponding p-values. We recorded the biological processes for the top 20 positive and negative genes and also their network neighbors.

**Network visualization, statistical calculations.** To visualize the gene correlation networks in Figs 3 and 4 Cytoscape was used. To carry out t-test calculations and calculate the corresponding p-values we used cpan’s Statistics package, a Perl library. Whenever not mentioned explicitly, calculations and data manipulation was carried out with in-house Perl script (available on request from the author).

**References**

1. Barabasi, A. L. & Oltvai, Z. N. Network biology: understanding the cell’s functional organization. *Nat Rev Genet* **5**, 101–113, doi: 10.1038/nrg1272 (2004).
2. Golb, K. I. et al. The human disease network. *Proceedings of the National Academy of Sciences of the United States of America* **104**, 8685–8690, doi: 10.1073/pnas.0701361104 (2007).
3. Terzian, A. C., Andreoli, S. B., Razouk, D., Chaves, A. C. & Mari Jde. J. Fertility and fecundity of an outpatien sample with schizophrenia. *Rev Bras Psiquiatr* **28**, 305–307 (2006).
4. Kenny, E. M. et al. Excess of rare novel loss-of-function variants in synaptic genes in schizophrenia and autism spectrum disorders. *Mol Psychiatry* **19**, 872–879, doi: 10.1038/mp.2013.127 (2014).
5. Maynard, T. M., Sikich, L., Lieberman, J. A. & LaMantia, A. S. Neural development, cell-cell signaling, and the “two-hit” hypothesis of schizophrenia. *Schizophr Bull* **27**, 457–476 (2001).
6. Takata, A., Ionita-Laza, I., Gogos, J. A., Xu, B. & Karayiorgou, M. De Novo Synonymous Mutations in Regulatory Elements Contribute to the Genetic Etiology of Autism and Schizophrenia. *Neuron* **89**, 940–947, doi: 10.1016/j.neuron.2016.02.024 (2016).
7. Schizophrenia Working Group of the Psychiatric Genomics, C. Biological insights from 108 schizophrenia-associated genetic loci. *Nature* **511**, 421–427, doi: 10.1038/nature13595 (2014).
8. Hawrylycz, M. J. et al. An anatomically comprehensive atlas of the adult human brain transcriptome. *Nature* **489**, 391–399, doi: 10.1038/nature11405 (2012).
9. Miller, J. A. et al. Transcriptional landscape of the prenatal human brain. *Nature* **508**, 199–206, doi: 10.1038/nature13185 (2014).
10. Wockner, L. F. et al. Genome-wide DNA methylation analysis of human brain tissue from schizophrenia patients. *Transl Psychiatry* **4**, e339, doi: 10.1038/tp.2013.111 (2014).
11. Wu, Y., Yao, Y. G. & Luo, X. J. SZDB: A Database for Schizophrenia Genetic Research. *Schizophr Bull*, doi: 10.1093/schbul/sbw102 (2016).
12. Rappaport, N. et al. MacaCards: A Comprehensive Automatically-Mined Database of Human Diseases. *Curr Protoc Bioinformatics* **47**, 1 24–21–19, doi: 10.1002/9780471250953.bic12447 (2014).
13. Safran, M. et al. GeneCards Version 3—the human gene integrator. *Database: the journal of biological databases and curation* **29**, baq020, doi: 10.1093/database/baq020 (2010).
14. Binns, D. et al. QuickGO: a web-based tool for Gene Ontology searching. *Bioinformatics* **25**, 3045–3046, doi: 10.1093/bioinformatics/btp536 (2009).
15. Davidkova, G., McCullumsmith, R. E. & Meador-Woodruff, J. H. Expression of ARHGEF11 mRNA in schizphrenic thalamus. *Ann N Y Acad Sci* **1003**, 373–377 (2003).
16. van Kammen, D. P. et al. CSF diastepend binding inhibitor and schizophrenia: clinical and biochemical relationships. *Biol Psychiatry* **34**, 515–522 (1993).
17. Yamamoto, Y. et al. Cloning and expression of myelin-associated oligodendrocytic basic protein. A novel basic protein constituting the central nervous system myelin. *J Biol Chem* **269**, 31725–31730 (1994).
18. Montague, P., McCallion, A. S., Davies, R. W. & Griffiths, I. L. Myelin-associated oligodendrocytic basic protein: a family of abundant CNS myelin proteins in search of a function. *Dev Neurosci* **28**, 479–487, doi: 10.1159/000095110 (2006).
19. Zheng, R. et al. Neuropeptide-stimulated cell migration in prostate cancer cells is mediated by RhoA kinase signaling and inhibited by neutral endopeptidase. *Oncogene* **25**, 5942–5952, doi: 10.1038/sj.onc.1208586 (2006).
20. Bourguignon, L. Y., Singleton, P. A., Zhu, H. & Diedrich, F. Hyaluronan-mediated CD44 interaction with RhoGEF and Rho kinase promotes Gemb2-associated binder-1 phosphorylation and phosphatidylinositol 3-kinase signaling leading to cytokine (macrophage-colony stimulating factor) production and breast tumor progression. *J Biol Chem* **279**, 29420–29434, doi: 10.1074/jbc.M301885200 (2003).
21. Ettlinger, S. L. et al. Dysregulation of steroid response element-binding proteins and downstream effectors in prostate cancer during progression to androgen independence. *Cancer Res* **64**, 2212–2221 (2004).
22. Yang, X. et al. Regulation of fatty acid synthase expression in breast cancer by sterol regulatory element binding protein-1c. *Exp Cell Res* **282**, 132–137 (2003).
23. Ji, S. et al. ALDOA functions as an oncogene in the highly metastatic pancreatic cancer. *Cancer Lett* **374**, 127–135, doi: 10.1016/j.canlet.2016.01.054 (2016).
24. Baranwal, S. et al. Molecular characterization of the tumor-suppressive function of nischarin in breast cancer. *J Natl Cancer Inst* **103**, 1513–1528, doi: 10.1093/jnci/djt1350 (2011).
25. Weisbord, A. B. et al. Altered PTEN, ATRX, CHGA, CHGB, and TP53 expression are associated with aggressive VHL-associated pancreatic neuroendocrine tumors. *Horm Cancer* **4**, 165–175, doi: 10.1007/s12672-013-0134-1 (2013).
26. Ishikawa, N. et al. Characterization of SEZL2 cell-surface protein as a novel prognostic marker for lung cancer. *Cancer Sci* **97**, 737–745, doi: 10.1111/j.1349-7006.2006.00258.x (2006).
27. Amer, L. C. et al. Identification and characterization of novel genes located at the t(1;15)(p36.2;q24) translocation breakpoint in the neuroblastoma cell line NGP. *Genomics* **64**, 195–202, doi: 10.1006/geno.1999.6097 (2000).
28. de Groot, J. F., Piao, Y., Yu, L., Fuller, G. N. & Yung, W. K. Knockdown of GluR1 expression by RNA interference inhibits glioma proliferation. *J Neurosci Res* **88**, 121–133, doi: 10.1002/jnr.20552-2 (2008).
29. Levin, A. M. et al. Chromosome 17q12 variants contribute to risk of early-onset prostate cancer. *Cancer Res* **68**, 6492–6495, doi: 10.1158/0008-5472.CAN-08-0348 (2008).
30. Liu, Z., Ballinger, J. R., Rauth, A. M., Bendayan, R. & Wu, X. Y. Delivery of an anticancer drug and a chemosensitizer to murine breast sarcoma by intratumoral injection of sulfopropyl dextran microspheres. J Pharm Pharmacol 55, 1063–1073, doi: 10.1211/0022357021567 (2003).

31. Negishi, M., Oinuma, I. & Katoh, H. Plexins: axon guidance and signal transduction. Cell Mol Life Sci 62, 1363–1371, doi: 10.1007/s00018-005-5018-2 (2005).

32. Gottkins, J. G. et al. Expression of the genes encoding myelin basic protein and proteolipid protein in human malignant gliomas. Clin Cancer Res 3, 799–804 (1997).

33. Ingermann, A. R. et al. Identification of a novel cell death receptor mediating IGFBP-3-induced anti-tumor effects in breast and prostate cancer. J Biol Chem 285, 30233–30246, doi: 10.1074/jbc.M110.122226 (2010).

34. Daigo, Y. et al. Isolation, mapping and mutation analysis of a human cDNA homologous to the doc-1 gene of the Chinese hamster, a candidate tumor suppressor for oral cancer. Genes Chromosomes Cancer 20, 204–207 (1997).

35. Moschonas, A. et al. CD40 induces antigen transporter and immunoproteasome gene expression in carcinomas via the coordinated action of NF-kappaB and of NF-kappaB-mediated de novo synthesis of IRF-1. Mol Cell Biol 28, 6208–6222, doi: 10.1128/MCB.00611-08 (2008).

36. Most, P. et al. Extracellular S100A1 protein inhibits apoptosis in ventricular cardiomyocytes via activation of the extracellular signal-regulated protein kinase 1/2 (ERK1/2). J Biol Chem 278, 48404–48412, doi: 10.1074/jbc.M308372200 (2003).

37. Pfuhl, T. et al. Biochemical characterisation of the proteins encoded by the DiGeorge critical region 6 (DGC6R) genes. Hum Genet 117, 70–80, doi: 10.1007/s00439-005-1267-2 (2005).

38. Morihoro, Y. et al. Fatty acid binding protein 7 as a marker of glioma stem cells. Pathol Int 63, 546–553, doi: 10.1111/pin.12109 (2013).

39. Ishii, H. et al. Epithelial splicing regulatory proteins 1 (ESRP1) and 2 (ESRP2) suppress cancer cell motility via different mechanisms. J Biol Chem 289, 27386–27399, doi: 10.1074/jbc.M114.589432 (2014).

40. Su, Y. et al. Aldehyde dehydrogenase 1A1-positive cell population is enriched in tumor-initiating cells and associated with progression of bladder cancer. Cancer Epidemiol Biomarkers Prev 19, 327–337, doi: 10.1158/1055-9965.EPI-09-0865 (2010).

41. Walsh, N. et al. Aldehyde dehydrogenase 1A1 and gelosin identified as novel invasion-modulating factors in conditioned medium of pancreatic cancer cells. J Proteomics 71, 561–571, doi: 10.1016/j.jprot.2008.09.002 (2008).

42. Gao, F. et al. The role of LGR5 and ALDH1A1 in non-small cell lung cancer: Cancer progression and prognosis. Biochem Biophys Res Commun 462, 91–98, doi: 10.1016/j.bbrc.2015.04.029 (2015).

43. Zhou, Y. et al. Clinicalpathological significance of ALDH1A1 in lung, colorectal, and breast cancers: a meta-analysis. Biomark Med 9, 777–790, doi: 10.2217/BMM.15.49 (2015).

44. Gene Ontology, C. Gene Ontology Consortium: going forward. Nucleic acids research 43, D1049–1056, doi: 10.1093/nar/gku1179 (2015).

45. Szklarczyk, D. et al. STRING v10: protein–protein interaction networks, integrated over the tree of life. Nucleic acids research 43, D447–452, doi: 10.1093/nar/gku1003 (2015).

46. Volk, D. W. & Lewis, D. A. Early developmental disturbances of cortical inhibitory neurons: contribution to cognitive deficits in schizophrenia. Schizophr Bull 40, 952–957, doi: 10.1093/schbul/sbu111 (2014).

47. Mirnics, K., Middleton, F. A., Lewis, D. A. & Levitt, P. Analysis of complex brain disorders with gene expression microarrays: Implications for schizophrenia, autism and cancer. Prog Neurobiol 93, 13–24, doi: 10.1016/j.pneurobio.2010.09.004 (2011).

48. Tkachev, D. et al. Oligodendrocyte dysfunction in schizophrenia and bipolar disorder. Lancet 362, 798–805, doi: 10.1016/S0140-6736(03)14289-4 (2003).

49. Fields, R. D. White matter in learning, cognition and psychiatric disorders. Trends Neurosci 31, 361–370, doi: 10.1016/j.tins.2008.04.001 (2008).

50. Raiker, S. J. et al. Oligodendrocyte–myelin glycoprotein and Nogo negatively regulate activity-dependent synaptic plasticity. J Neurosci 30, 12432–12445, doi: 10.1523/JNEUROSCI.0895-10.2010 (2010).

51. Tabares-Seisdedos, R. & Rubenstein, J. L. Chromosome 8p as a potential hub for developmental neuropsychiatric disorders: implications for schizophrenia, autism and cancer. Mol Psychiatry 14, 563–589, doi: 10.1038/mp.2009.2.2009 (2009).

52. Kar, G., Gursoy, A. & Keskin, O. Human cancer protein–protein interaction network: a structural perspective. PLoS computational biology 5, e1000601, doi: 10.1371/journal.pcbi.1000601 (2009).

53. Liu, Y. et al. Exploring the pathogenetic association between schizophrenia and type 2 diabetes mellitus diseases based on pathway analysis. BMC Med Genomics 6 Suppl 1, S17, doi: 10.1186/1755-8794-6-S1-S17 (2013).

54. Gumerov, V. & Hegyi, H. MicroRNA-derived network analysis of differentially methylated genes in schizophrenia, implicating GABA receptor B1 [GABBR1] and protein kinase B [AKT1]. Biology direct 10, 59, doi: 10.1186/s13006-015-0089-y (2015).

55. Izan, K., Boullou, C., Tabares-Seisdedos, R., Baudot, A. & Valencia, A. Molecular evidence for the inverse comorbidity between central nervous system disorders and cancers detected by transcriptomic meta-analyses. PLoS Genet 10, e1004173, doi: 10.1371/journal.pgen.1004173 (2014).

56. Thurman, R. E. et al. The accessible chromatin landscape of the human genome. Nature 489, 75–82, doi: 10.1038/nature11232 (2012).

57. Shannon, P. et al. Cytoscape: a software environment for integrated models of biomolecular interaction networks. Genome research 13, 2498–2504, doi: 10.1101/gr.1239303 (2003).

Acknowledgements
This research was supported by the project “CEITEC - Central European Institute of Technology” (CZ.1.05/1.1.00/02.0068) from the European Regional Development Fund.

Additional Information
Supplementary information accompanies this paper at http://www.nature.com/srep
Competing Interests: The author declares no competing financial interests.

How to cite this article: Hegyi, H. Connecting myelin-related and synaptic dysfunction in schizophrenia with SNP-rich gene expression hubs. Sci. Rep. 7, 45494; doi: 10.1038/srep45494 (2017).

Publisher’s note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.
