ARTICLE

Pleiotropic Meta-Analysis of Cognition, Education, and Schizophrenia Differentiates Roles of Early Neurodevelopmental and Adult Synaptic Pathways

Max Lam,[1,2,3] W. David Hill,[4,5] Joey W. Trampush,[6] Jin Yu,[2] Emma Knowles,[7] Gail Davies,[4,5] Eli Stahl,[8,9,10] Laura Huckins,[8,9,10] David C. Liewald,[5] Srdjan Djurovic,[11,12] Ingrid Melle,[12,13] Kjetil Sundet,[13,14] Andrea Christoforou,[15] Ivar Reinvang,[14] Pamela DeRosse,[2] Astri J. Lundervold,[16] Vidar M. Steen,[12,15] Thomas Espeseth,[13,14] Katri Räikkönen,[17] Elisabeth Widen,[18] Aarno Palotie,[18,19,20] Johan G. Eriksson,[21,22,23] Ina Giegling,[24] Bettina Konte,[24] Annette M. Hartmann,[24] Panos Roussos,[8,9,10,25] Stella Giakoumaki,[26] Katherine E. Burdick,[8,25,27] Antony Payton,[28] William Ollier,[29,30] Ornit Chiba-Falek,[31,32] Deborah K. Attix,[31,32,33,34] Anna C. Need,[35]

(Author list continued on next page)

Susceptibility to schizophrenia is inversely correlated with general cognitive ability at both the phenotypic and the genetic level. Paradoxically, a modest but consistent positive genetic correlation has been reported between schizophrenia and educational attainment, despite the strong positive genetic correlation between cognitive ability and educational attainment. Here we leverage published genome-wide association studies (GWASs) in cognitive ability, education, and schizophrenia to parse biological mechanisms underlying these results. Association analysis based on subsets (ASSET), a pleiotropic meta-analytic technique, allowed jointly associated loci to be identified and characterized. Specifically, we identified subsets of variants associated in the expected ("concordant") direction across all three phenotypes (i.e., greater risk for schizophrenia, lower cognitive ability, and lower educational attainment); these were contrasted with variants that demonstrated the counterintuitive ("discordant") relationship between education and schizophrenia (i.e., greater risk for schizophrenia and higher educational attainment). ASSET analysis revealed 235 independent loci associated with cognitive ability, education, and/or schizophrenia at $p < 5 \times 10^{-8}$. Pleiotropic analysis successfully identified more than 100 loci that were not significant in the input GWASs. Many of these have been validated by larger, more recent single-phenotype GWASs. Leveraging the joint genetic correlations of cognitive ability, education, and schizophrenia, we were able to dissociate two distinct biological mechanisms—early neurodevelopmental pathways that characterize concordant allelic variation and adulthood synaptic pruning pathways—that were linked to the paradoxical positive genetic association between education and schizophrenia. Furthermore, genetic correlation analyses revealed that these mechanisms contribute not only to the etiopathogenesis of schizophrenia but also to the broader biological dimensions implicated in both general health outcomes and psychiatric illness.

Introduction

It has long been observed that impaired cognitive ability is a significant aspect of the illness in schizophrenia (MIM: 181500).[1–5] Cognitive deficits have been shown to be largely independent of clinical state and treatment status in patients with schizophrenia[1,4,6–9] and are observed (in more subtle forms) in their first-degree relatives.[10,11]

[1]Institute of Mental Health, Singapore, 539747, Singapore; [2]Division of Psychiatry Research, The Zucker Hillside Hospital, Glen Oaks, NY 11004, USA; [3]Stanley Center for Psychiatric Research, Broad Institute of Harvard and MIT, Cambridge, MA 02142, USA; [4]Centre for Cognitive Ageing and Cognitive Epidemiology, University of Edinburgh, Edinburgh, Scotland, EH8 9JZ, United Kingdom; [5]Department of Psychology, University of Edinburgh, Edinburgh, Scotland, EH8 9JZ, United Kingdom; [6]Department of Psychiatry and the Behavioral Sciences, Keck School of Medicine, University of Southern California, Los Angeles, CA 90033, USA; [7]Department of Psychiatry, Yale University School of Medicine, New Haven, CT 06511, USA; [8]Department of Psychiatry, Icahn School of Medicine at Mount Sinai, New York, NY 10029, USA; [9]Department of Genetics and Genomic Science, Icahn School of Medicine at Mount Sinai, New York, NY 10029, USA; [10]Institute for Multiscale Biology, Icahn School of Medicine at Mount Sinai, New York, NY 10029, USA; [11]Department of Medical Genetics, Oslo University Hospital, University of Bergen, Bergen 4956, Nydalen 0424, Norway; [12]Norsk Senter for Forskning på Mentale Lidelser, K.G. Jebsen Centre for Psychosis Research, University of Bergen, Bergen 4956, Nydalen 0424, Norway; [13]Division of Mental Health and Addiction, Oslo University Hospital, Oslo 1039, Blindern 0315, Norway; [14]Department of Psychology, University of Oslo, Oslo 1094, Blindern 0317, Norway; [15]Dr. Einar Martens Research Group for Biological Psychiatry, Center for Medical Genetics and Molecular Medicine, Haukeland University Hospital, Bergen 7804, N-5020 Bergen, Norway; [16]Department of Biological and Medical Psychology, University of Bergen, 7807, N-5020, Norway; [17]Institute of Behavioural Sciences, University of Helsinki, Helsinki, 00014, Finland; [18]Institute for Molecular Medicine Finland (FIMM), University of Helsinki, 00014, Finland; [19]Wellcome Trust Sanger Institute, Wellcome Trust Genome Campus, Cambridge CB10 1SA, United Kingdom; [20]Department of Medical Genetics, University of Helsinki and University Central Hospital, Helsinki, 00014, Finland; [21]Department of General Practice, University of Helsinki and Helsinki University Hospital, Helsinki, 00014, Finland; [22]National Institute for Health and Welfare, Helsinki FI-00271, Finland; [23]Folkhälsan Research Center, Helsinki 00290, Finland; [24]Department of Psychiatry, Martin Luther University of Halle-Wittenberg, Halle 06108, Germany; [25]Mental Illness Research, Education, and Clinical Center (VISN 2), James J. Peters VA Medical Center, Bronx, NY 10468, USA; [26]Department of Psychology, University of Crete, Crete 74100, Greece; [27]Department of Psychiatry, Brigham and Women's Hospital, Harvard Medical School, Boston, MA 02115; [28]Division of Informatics, Imaging, and Data Sciences, School of Health Sciences, University of Manchester, Manchester M139NT, United Kingdom; [29]Centre for Epidemiology, Division of Population Health, Health Services Research and Primary Care, University of Manchester, Manchester M139PL, United Kingdom; [30]School of Healthcare Sciences,

(Affiliations continued on next page)

Elizabeth T. Cirulli,[36] Aristotle N. Voineskos,[37] Nikos C. Stefanis,[38,39,40] Dimitrios Avramopoulos,[41,42] Alex Hatzimanolis,[37,38,39] Dan E. Arking,[41] Nikolaos Smyrnis,[37,38] Robert M. Bilder,[42] Nelson A. Freimer,[42] Tyrone D. Cannon,[44] Edythe London,[43] Russell A. Poldrack,[45] Fred W. Sabb,[46] Eliza Congdon,[43] Emily Drabant Conley,[47] Matthew A. Scult,[48] Dwight Dickinson,[49] Richard E. Straub,[50] Gary Donohoe,[52] Derek Morris,[52] Aiden Corvin,[53,54] Michael Gill,[53,54] Ahmad R. Hariri,[48] Daniel R. Weinberger,[50] Neil Pendleton,[51] Panos Bitsios,[55] Dan Rujescu,[24] Jari Lahti,[17,56] Stephanie Le Hellard,[12,15] Matthew C. Keller,[57] Ole A. Andreassen,[12,13,58] Ian J. Deary,[4,5] David C. Glahn,[7] Anil K. Malhotra,[2,59,60] and Todd Lencz[2,59,60,*]

Moreover, cognitive deficits precede illness onset by many years; they begin in early childhood[5,12–14] and thus result in reduced educational attainment.[15,16]

Unlike phenotypic correlation, which measures covariances within the distribution of the phenotypes, genetic correlation (r_g) indexes the covariance between SNP ≈ phenotype genome-wide association study (GWAS) effect sizes across a pair of phenotypes measured in separate GWAS studies. Recent advances in psychiatric and cognitive genomics have reliably demonstrated that the inverse relationship between cognitive ability and risk for schizophrenia is also observed at the molecular genetic level ($r_g \approx -.20$).[17–23] Paradoxically, genetic correlation studies have indicated a *positive* relationship between educational attainment and risk for schizophrenia ($r_g \approx .10$),[20,23–27] despite the fact that educational attainment and cognitive ability exhibit a very strong polygenic overlap ($r_g \approx .70$).[18,23,27] Educational attainment is often considered to be a proxy for cognitive ability; however, the lack of perfect genetic overlap between the two, combined with the paradoxical genetic correlation between educational attainment and schizophrenia, suggests an opportunity to decompose distinct genetic mechanisms accounting for this pattern of results.

Whereas genetic-correlation analysis has recently become widespread because of the availability of techniques such as linkage disequilibrium (LD) score regression (LDSC),[25,28] these approaches generally result in a single, genome-wide estimate of polygenic overlap. Moreover, novel meta-analytic approaches (e.g., multi-trait analysis of GWAS [MTAG][29]) for merging seemingly heterogeneous GWAS datasets tend to exploit commonalities across phenotypes rather than differences; for example, two recent studies have employed MTAG across the highly correlated cognitive and educational GWASs in order to accelerate the process of gene discovery.[18,23] In contrast, few studies have attempted to examine the counter-intuitive correlation between schizophrenia and educational attainment or to parse subsets of SNPs that might drive cross-phenotype correlations. An initial effort has successfully identified a few individual loci that act in paradoxical fashion, increasing educational attainment while simultaneously increasing risk for schizophrenia;[30] two other studies have identified loci that demonstrate other pleiotropic effects.[19,31]

To date, however, no studies have utilized pleiotropic meta-analytic techniques to comprehensively parse variance from cognitive, educational, and schizophrenia-focused GWASs that might pinpoint differential biological mechanisms. In order for the paradoxical pattern of genome-wide correlations to exist, there must be identifiable subsets of SNPs that are differentially involved in driving these genetic relationships. Therefore, we sought to identify differentially associated variants, which could yield crucial insights into the fine-grained genetic architecture of schizophrenia and in turn give us insights into the

Manchester Metropolitan University, Manchester M15 6BH, United Kingdom; [31]Department of Neurology, Bryan Alzheimer Disease Research Center, Duke University Medical Center, Durham, NC 27705, USA; [32]Center for Genomic and Computational Biology, Duke University Medical Center, Durham, NC 27705, USA; [33]Psychiatry and Behavioral Sciences, Division of Medical Psychology, Duke University Medical Center, Durham, NC 27708, USA; [34]Department of Neurology, Duke University Medical Center, Durham, NC 27708, USA; [35]Division of Brain Sciences, Department of Medicine, Imperial College, London W12 0NN, UK; [36]Helix Inc, San Diego, CA 92121, USA; [37]Campbell Family Mental Health Institute, Centre for Addiction and Mental Health, University of Toronto, Toronto M6J 1H4, Canada; [38]Department of Psychiatry, National and Kapodistrian University of Athens Medical School, Eginition Hospital, Athens, Greece; [39]University Mental Health Research Institute, Athens 115 27, Greece; [40]Neurobiology Research Institute, Theodor-Theohari Cozzika Foundation, Athens, Greece; [41]Department of Psychiatry, Johns Hopkins University School of Medicine, Baltimore, MD 21287, USA; [42]McKusick-Nathans Institute of Genetic Medicine, Johns Hopkins University School of Medicine, Baltimore, MD 21205, USA; [43]UCLA Semel Institute for Neuroscience and Human Behavior, Los Angeles, CA 90024, USA; [44]Department of Psychology, Yale University, New Haven, CT 06511, USA; [45]Department of Psychology, Stanford University, Palo Alto, CA 94305, USA; [46]Robert and Beverly Lewis Center for Neuroimaging, University of Oregon, Eugene, OR, 97401, USA; [47]23andMe, Inc., Mountain View, CA, 94041, USA; [48]Laboratory of NeuroGenetics, Department of Psychology and Neuroscience, Duke University, Durham, NC 27708, USA; [49]Clinical and Translational Neuroscience Branch, Intramural Research Program, National Institute of Mental Health, National Institute of Health, Bethesda, MD 20814, USA; [50]Lieber Institute for Brain Development, Johns Hopkins University Medical Campus, Baltimore, MD 21205, USA; [51]Division of Neuroscience and Experimental Psychology, School of Biological Sciences, University of Manchester, Manchester Academic Health Science Centre, Salford Royal NHS Foundation Trust, Manchester M13 9PL, United Kingdom; [52]Neuroimaging, Cognition, and Genomics Centre, School of Psychology and Discipline of Biochemistry, National University of Ireland, Galway, Ireland; [53]Neuropsychiatric Genetics Research Group, Department of Psychiatry, Trinity College Dublin, Dublin, Ireland; [54]Trinity College Institute of Neuroscience, Trinity College Dublin, Dublin, Ireland; [55]Department of Psychiatry and Behavioral Sciences, Faculty of Medicine, University of Crete, Heraklion, Crete GR-71003, Greece; [56]Helsinki Collegium for Advanced Studies, University of Helsinki, Helsinki 00014, Finland; [57]Institute for Behavioral Genetics, University of Colorado, Boulder, CO 80303, USA; [58]Institute of Clinical Medicine, University of Oslo, Oslo 0318, Norway; [59]Department of Psychiatry, Zucker School of Medicine at Hofstra/Northwell, Hempstead, NY 11549, USA; [60]Center for Psychiatric Neuroscience, Feinstein Institute for Medical Research, Manhasset, NY 11030, USA
*Correspondence: tlencz@northwell.edu
https://doi.org/10.1016/j.ajhg.2019.06.012.

etiopathogenic mechanisms underlying the illness—mechanisms that standard GWAS cannot detect.

In the study reported here, we first utilized a simple subsetting approach to identify SNPs that are significantly associated *either* with cognitive ability or with educational attainment, but not with both (Figure S1A). We hypothesized that these SNP subsets would demonstrate stronger genetic correlations with schizophrenia than what is observed with a simple genome-wide approach. We then employed a pleiotropic meta-analytic approach, association analysis based on subsets (ASSET),[32] which permits the characterization of each SNP with respect to its pattern of effects on multiple phenotypes (Figure S1B). For example, ASSET has previously been used to demonstrate that the minor allele of rs2736100 (at the *TERT* [MIM: 187270] locus) is positively associated with risk for pancreatic cancer (MIM: 260350), negatively associated with risk for kidney (MIM: 144700) and lung cancers (MIM: 211980), and not significantly associated with risk for cancers of the breast (MIM: 114480), bladder (MIM: 109800), or prostate (MIM: 176207); other cancer loci were demonstrated to have various other patterns of effects.[32] We utilized ASSET to identify two types of loci: (1) those SNPs that are consistently associated with all three phenotypes in the expected direction (i.e., the same allele is associated with higher cognitive ability, higher educational attainment, and lower risk for schizophrenia), which we label "concordant," and (2) SNPs that demonstrate the paradoxical association between education and schizophrenia (i.e., the same allele associated with higher educational attainment and higher risk for schizophrenia), which we labeled "discordant." Next, we compared the statistically significant ASSET results to the output of single-trait GWASs of cognitive ability, educational attainment, or schizophrenia[20,33,34] in order to identify novel loci suggested by ASSET. Subsequently, we conducted a series of pathway and transcriptome-wide analyses to biologically characterize differential mechanisms underlying concordant versus discordant loci. Finally, we performed a series of genetic correlation analyses to compare the overlap of concordant and discordant SNP subsets with other relevant traits. Further analytic details are covered in the Material and Methods section; the full analysis workflow is also represented in Figures S1 and S2.

Material and Methods

Stage 1: Simple Subsetting Approach Based on p Values in Cognition and Education GWAS

Note that for most purposes in this manuscript, we are using the largest GWAS for schizophrenia,[35] cognitive ability,[18] and educational attainment[36] published prior to 2018. For each of these phenotypes, larger GWASs have been published in 2018; these were used for validation and extension as described in subsequent sections. Before we performed wany subsetting analyses, we used genome-wide genetic correlations to confirm the earlier observed genetic correlations between schizophrenia and both cognitive ability and education. In stage 1, preliminary SNP subsets were formed simply based on p value thresholds of cognitive ability and educational attainment GWAS: (1) SNPs nominally associated with cognition ($p < 0.05$) and not associated with education ($p > 0.05$) were selected, resulting in 74,470 SNPs; (2) SNPs nominally associated with cognition ($p < 0.05$) and not associated with education using a stricter threshold ($p > 0.5$) were selected, resulting in 66,657 SNPs; (3) similar procedures were carried out for SNPs nominally associated with education ($p < 0.05$) but not with cognition ($p > 0.05$), were selected, resulting in 104,807 SNPs; and (4) SNPs nominally associated with education ($p < 0.05$) and not cognition using a stricter threshold ($p > 0.5$) were selected, resulting in 44,803 SNPs.

Next, we performed a heterogeneity test between results of cognitive and educational GWAS by using METAL[37] and generated sets of SNPs showing opposite effects between the two (i.e., the same allele predicts better cognitive performance but less educational attainment, and vice versa). We identified sets of SNPs of varying sizes on the basis of varying p value thresholds for the heterogeneity tests ($p < 0.5$; $p < 0.25$; $p < 0.1$; $p < 0.05$; $p < 0.01$; and $p < 0.001$).

To evaluate the degree of genetic correlation of these preliminary subsets of SNPs with respect to schizophrenia, we utilized GNOVA,[38] a recently published method similar to LD score regression. GNOVA is specifically designed to examine genetic correlations using SNP subsets (rather than global genome-wide summary statistics), whereas such applications have not been explicitly tested in LD score regression and may not be robust.

Stage 2: ASSET Meta-Analysis and ASSET-Generated SNP Subsets

Schizophrenia GWAS summary statistics based on a European ancestry GWAS of schizophrenia (n = 77,096, cases = 33,640, controls = 43,456; GWAS mean $\chi^2 = 1.677$) were obtained from the Psychiatric Genomics Consortium.[35] To make them compatible with effect sizes (beta weights) derived from the linear-regression-based cognition and education GWASs, we converted odds ratios from the case-control schizophrenia GWAS to beta by taking the natural logarithm of the odds ratios. Effect direction per SNP was also reversed for schizophrenia so thatit would be consistent with the interpretation of cognition and education (i.e., concordant alleles are those where the direction of effect is the same for cognitive ability, educational attainment, and schizophrenia). Summary statistics for the education GWAS were obtained from the Social Science Genomics Association Consortium (SSGAC)[36] (n = 328,917, GWAS mean $\chi^2 = 1.638$). GWAS summary statistics for cognition are available via earlier inverse-variance meta-analysis of samples[18] from the COGENT[27] consortium (n = 107,207, GWAS mean $\chi^2 = 1.245$). We applied general quality-control parameters to the schizophrenia and cognitive GWAS summary statistics but we excluded SNPs with INFO scores < 0.6 and minor-allele frequency < 0.01; multiple quality-control parameters' thresholds were previously reported for the education GWAS,[36] and summary statistics were provided by the SSGAC. Detailed quality-control and meta-analytic procedures were reported earlier.[18] Only SNPs that were present for all three phenotypes were retained as inputs to the ASSET meta-analysis, resulting in 7,306,098 SNPs for subsequent analysis.

Pooling GWAS summary statistics via conventional inverse-variance meta-analysis increases power but also poses methodological

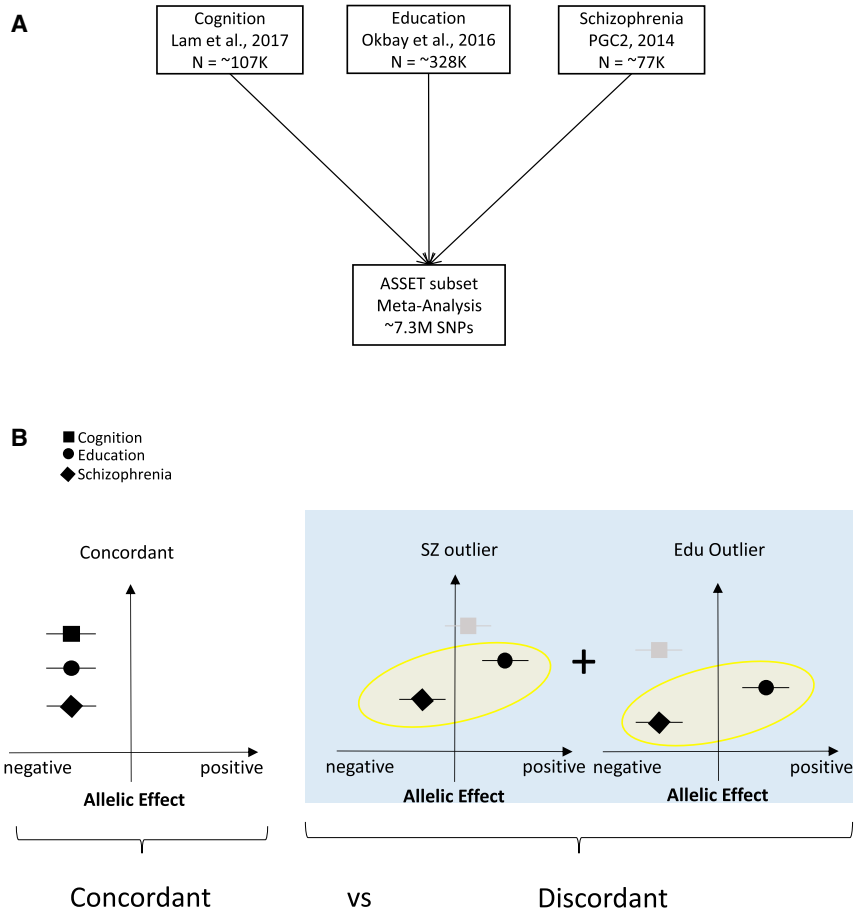

Figure 1. Design of the Present Study
(A) Input GWAS studies used for ASSET analysis.
(B) Definition of concordant and discordant SNP subsets. Concordant SNPs have alleles that demonstrate negative effects on cognitive ability, educational attainment, and schizophrenia risk (i.e., increased schizophrenia risk, reverse-coded for consistency). Discordant SNPs have alleles that demonstrate paradoxical effects on educational attainment and schizophrenia (i.e., higher educational attainment and increased schizophrenia risk, reverse-coded). The Y axes on the forest plots represent different input summary statistics for cognition, education, and schizophrenia.

challenges when different studies are capturing heterogeneous and/or pleiotropic phenotypes. In the case of pleiotropy, individual variants are likely to be associated with only a subset of the traits analyzed, or they might even demonstrate effects in different directions for the different phenotypes under analysis. ASSET meta-analysis[32] is an agnostic approach that generalizes standard fixed-effects meta-analysis by allowing a subset of the input GWASs to have no effect on a given SNP, and it exhaustively searches across all possible subsets of "non-null" GWAS inputs within a fixed-effect framework to identify the strongest association signal in both positive and negative directions. ASSET then evaluates the significance of these positive and negative associations while accounting for multiple testing. This methodology allows for a powerful pooled two-tailed Z-score test statistic that effectively combines p values for variants with strong effects in opposite directions across input GWASs. ASSET also permits the addition of a covariance term for the adjustment of overlapping samples. The genetic correlation matrix between the three input GWASs had been added to all reported ASSET analyses so that this adjustment could be performed. Recently, comparisons between cross-phenotype meta-analysis methodologies demonstrated that ASSET performed best as the number of meta-analyzed traits with null effects increased, and ASSET also did well in terms of specificity and sensitivity of the results. In addition, the ASSET approach best controlled for potential Type 1 inflation due to sample overlap and for non-uniform distribution of effect sizes.[39] As such, for the purpose of the current report, we selected ASSET for its conservative effect estimates and minimal inflation.

We combined GWAS summary statistics from schizophrenia, cognition, and education by using ASSET two-tailed meta-analysis (version 1.9.1) to obtain single cross-phenotype pleiotropic GWAS results. Default parameters were applied with the "h.traits" function. Inter-study correlations of the phenotype were first ascertained via LDSC,[25,28] which accounts for the genome-wide genetic correlation of the phenotypes and also for sample overlap. For each given SNP, ASSET generates Z scores of effect size and p values based on the strongest association from the input studies in positive and negative directions, respectively; then these p values are pooled into a single two-tailed p value for pleiotropy.[32,39] SNPs with similar relationships across the input traits (regardless of statistical significance) are then grouped into subsets identified by ASSET (see Figure 1B and Figure S1, bottom).

Again, as noted above, it is important to emphasize that the per-SNP direction of effect was reversed for schizophrenia so that it was consistent with the interpretation of cognition and education (i.e., higher scores are better, such that higher scores for schizophrenia are now coded as *decreased* risk for the disorder). Thus, in the notations to follow, ∩ represents variant subsets with the same effect directions, following the reversal of the direction of effect for the schizophrenia dataset, and | represents traits whose effect sizes are in the opposite direction of those for the other two traits (again following the reversal of the direction of effect for the schizophrenia dataset). ASSET subsets included: (1) scz ∩ edu ∩ cog (concordant, variants with an allele associated with an increase in cognitive ability and educational attainment, but a decrease in schizophrenia risk); (2) edu ∩ cog | scz (schizophrenia outliers, variants associated with an increase in cognitive ability and educational attainment but also with an increase in schizophrenia risk); (3) scz ∩ cog | edu (education outliers, variants associated with an increase in schizophrenia risk and reduced cognitive ability, but an increase in educational attainment); and (4) scz ∩ edu | cog (cognition outliers, variants associated with an increase in schizophrenia risk and reduced educational attainment, but an increase in cognitive ability) subsets. ASSET also identified SNPs where only a single trait (scz or edu or cog) was significant; these were included in a category called "single phenotype."

Finally, to generate an appropriate contrast for the "concordant" subset, we included a combined single "discordant" subset, representing the counter-intuitive genetic correlation between education and schizophrenia, where discordant = (edu ∩ cog | scz) + (scz ∩ cog | edu) [subsets 2 and 3 above], regardless of the effect for cognition. These contrasts are also represented visually in Figure 1 and Figure S1.

Consolidation of Independent Loci

Independent genome-wide-significant loci for each ASSET meta-analysis subset were identified via SNP clumping procedures that are a part of the functional mapping and annotation (FUMA) pipeline.[40] For the LD-rich MHC region, a single top SNP was retained. For all other loci, clumping procedures were carried out on the basis of the European 1000 Genomes Project phase 3 reference panel. First, "independent significant SNPs" were defined as those SNPs with a p value $< 5 \times 10^{-8}$ and with an LD of $r^2 < 0.6$ and other, more significant SNPs at the same locus. Second, candidate SNPs were then identified for subsequent annotations and were defined as all SNPs that had an MAF of 0.01 and a maximum p value of 0.05 and that were in LD, with $r^2 \geq 0.6$, with at least one of the independent significant SNPs. To ensure that biological annotation of these loci would not be hampered by poor coverage at any locus, we included SNPs that came from the 1000 Genomes reference panel but that might not have been included in the ASSET data. Third, "lead SNPs" were defined as the independent significant SNPs that had the strongest signal at a given locus and were strictly independent from each other ($r^2 < 0.1$). Finally, risk loci that were 250 kb or closer were merged into a single locus. The FUMA procedure was iterated across all ASSET SNP subsets, which were comprised (by definition) of non-overlapping SNPs. Additional variant annotations were conducted with ANNOVAR,[41] and lookups with published GWASs were conducted with the GWAS catalog. Additional SNP lookups were performed with the input summary statistics (cognition, education, and schizophrenia GWASs[18,35,36]), recent MTAG analyses of intelligence,[23] recent cognition or intelligence[20,21] GWASs, and pleiotropic analyses of cognition and schizophrenia[19] as well as education and schizophrenia.[31] RAggr[43] was utilized for extracting SNPs within 250 kb and $r^2 > 0.6$ from published reports to allow merging of loci generated from ASSET subsets.

MAGMA Gene-Based Analysis: Tissue Expression and Competitive Pathway Analysis

SNPs were mapped to 19,436 protein-coding genes via MAGMA as implemented in the FUMA[40] pipeline. MAGMA[44] gene analysis was performed with a default SNP-wide mean model that used the 1000 Genomes phase 3 reference panel and default gene annotations that are part of the FUMA pipeline. Genome-wide SNP p values and SNP-level sample sizes were included in the input files. MAGMA gene-tissue expression analysis was carried out with the Genotype-Tissue Expression (GTEx; version 7[45–47]) resource to examine the relationship between gene expression in a specific tissue type and ASSET results. The gene-property test was performed for average expression (log_2-transformed RPKM with pseudocount 1 after winsorization at 50) of 53 specific tissue types conditioning on average expression across all tissue types.

In order to identify specific biological processes linked to our subphenotypes of interest, we also used the results that emerged from the ASSET analysis to conduct MAGMA competitive pathway analysis. Gene sets that were tested included drug-related pathways,[48,49] as well as custom-curated neurodevelopmental and other brain-related gene sets that had gone through stringent quality control in a study originally designed to interrogate rare variants in schizophrenia.[50] In the latter, pathways with more than 100 genes from Gene Ontology (release 146; June 22, 2015 release), KEGG (July 1, 2011 release), PANTHER (May 18, 2015 release), REACTOME (March 23, 2015 release), DECIPHER Developmental Disorder Genotype-Phenotype (DDG2P) database (April 13, 2015 release), and the Molecular Signatures Database (MSigDB) hallmark processes (version 4, March 26, 2015 release) were initially included. Brain-level tissue expression gene sets included the Brainspan RNA-seq dataset[51] and the GTEx v7 dataset.[45] MAGMA gene-based and gene-set analysis were conducted with MAGMA v1.6.[44] Additional gene sets were selected on the basis of risk for schizophrenia and neurodevelopmental disorders, including those reported for schizophrenia rare variants[52] (translational targets of *FMRP* [MIM: 309550]; components of the post-synaptic density; ion channel proteins; components of the ARC, mGluR5, and NMDAR complexes; proteins at cortical inhibitory synapses; targets of mir-137; and genes near schizophrenia common risk loci) and autism (MIM: 613436) risk (targets of *CHD8* [MIM: 610528], splice targets of *RBFOX* [MIM: 605104], hippocampal gene expression networks, and neuronal gene lists from the Genes2Cognition database, as well as loss-of-function [LoF]-intolerant genes [pLI > 0.9 from the ExAC v0.3.1 pLI metric], ASD risk genes for FDR < 10% and 30%, and ASD or developmental disorder *de novo* genes hit by an LoF and/or missense *de novo* variant). Further details of curation of these gene sets was previously reported.[50]

S-PrediXcan: Brain-Tissue Expression Profiles and Hypergeometric Gene-Set Enrichment Analysis

Genetically regulated gene expression was imputed for the ASSET summary statistics with tissue models from GTEx v7 and the CommonMind Consortium via S-PrediXcan (formerly MetaXcan).[53–55] GTEx v7 tissue included amygdala (n = 88), anterior cingulate cortex (n = 109), basal ganglia (n = 144), cerebellar hemisphere (n = 125), cerebellum, cortex (n = 136), frontal cortex (n = 118), hippocampus (n = 111), hypothalamus (n = 108), nucleus accumbens (n = 130), putamen (n = 111), spinal cervical-1 (n = 83), and substantia nigra (n = 80). The CommonMind consortium data consist of tissue expression data derived from the dorsolateral prefrontal cortex (DLPFC, n = 279).[56] GTEX v7 Tissue expression models were trained using elastic net models that were made publicly available (see Web Resources). Elastic net models for DLPFC were contributed by collaborators from the CommonMind Consortium.[56–58] Bonferroni correction was first conducted for each ASSET subset of genes. Genes that survived multiple testing corrections were entered into GENE2FUNC, which is part of the FUMA[40] pipeline, to be tested for over-representation. This analysis differs from the MAGMA gene-set analysis in that the MAGMA gene-set analysis is used to examine whether gene sets, united by a known biological theme, are enriched for the phenotype under investigation. In a test of over-representation, as conducted with GENE2FUNC, the shared function of the genes of interest is unknown, and its elucidation is the goal of a test of over-representation. The over-representation test conducted with GENE2FUNC queries gene sets from (1) the Molecular Signature Database (MsigDB v 5.2), (2) WikiPathways (curated version 20161010), and (3) GWAS catalog (reported genes ver e91 20180206); to avoid spurious results, we required a minimum of three genes per pathway. For each gene set, hypergeometric tests were conducted

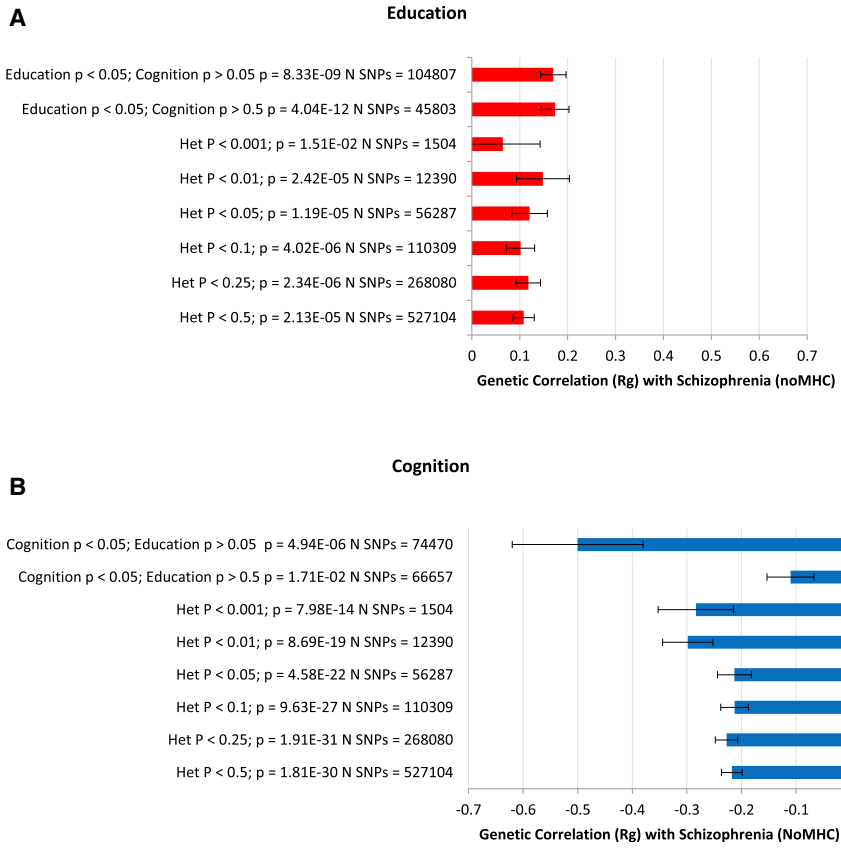

Figure 2. Genetic Correlations with Schizophrenia for SNPs Demonstrating Heterogeneity of Effects between (A) Cognitive Ability and (B) Educational Attainment

(A) Cognitive ability.

(B) Educational attainment.

Genetic correlation was carried out with GNOVA. Error bars represent standard errors. Education = Okbay PMID 27225129; cognition = Lam PMID 29186694; schizophrenia (Ripke) = PGC Schizophrenia Working Group PMID 25056061.

so that the list of genes significant in each ASSET subset could be examined for overlap with gene sets within the databases stipulated above; Bonferroni correction for multiple testing was applied. To reduce the likelihood that hypergeometric pathway analysis would be influenced by the dense number of genes within the MHC region, we removed genes within the coordinates of 28,000,000–35,000,000[59] on chromosome 6.

Genetic Correlations

To examine how our ASSET concordant and discordant SNP subsets relate to other phenotypes with available GWAS data, we conducted genetic correlation tests by using GNOVA,[38] an approach similar to LD score regression but capable of working with SNP subsets. Notably, GNOVA provides a corrected r_g that has been demonstrated to be robust to sample overlaps, and GNOVA is able to account for LD across SNPs.[38] We selected a series of neuropsychiatric, inflammatory, brain, metabolic, and cardiovascular phenotypes that have been previously demonstrated to have genetic correlations with cognitive measures and used them to interrogate the genetic overlaps of our ASSET subsets. Interpretation of GNOVA for the concordant subset was straightforward because the three input GWAS weights all follow the same direction (following the reverse coding of schizophrenia, as noted previously). In contrast, discordant SNPs have two separate potential weights (allelic β for schizophrenia versus allelic β for education); as shown in Figure 1B, a given SNP might have somewhat different effect sizes (distances from the center line) for education as compared to schizophrenia. Therefore, we weighted each SNP by the stronger value of β: variants for which the schizophrenia β was stronger than the education β were referred to as "schizophrenia type" and variants with the opposite pattern were referred to as "education type." Nevertheless, it is important to emphasize that the discordant SNPs represent a single dimension of biology, and the net effects of all "schizophrenia type" variants were equivalent to those of the "education type" SNPs, albeit with opposite signs.

Results

Stage 1: Preliminary Evaluation of Genetic Correlations

We used GWAS summary statistics for cognition (n = 107,207)[18] and education (n = 328,917)[36] to evaluate preliminary genetic correlations with schizophrenia (n = 77, 096).[35] Consistent with previous results, the inverse genetic correlation that GNOVA revealed between cognition and schizophrenia was significant ($r_g = -.21$, se = 0.03, $p = 1.12 \times 10^{-12}$), as was the counter-intuitive positive correlation between education and schizophrenia ($r_g = 0.08$, se = 0.02, $p = 2.05 \times 10^{-5}$). Note that these analyses were conducted before we reversed the direction of effect for schizophrenia.

Prior to the main ASSET analysis, we used two simple approaches to examine subsets of SNPs and their association with schizophrenia (Figure S1). First, we selected SNPs that were nominally associated with education ($p < 0.05$) and generally not associated with cognition ($p > 0.05$); GNOVA revealed a slightly stronger positive correlation, r_g of 0.17, between this subset of educational attainment SNPs and schizophrenia than did genome-wide summary statistics (Figure 2A). With a stricter threshold for SNPs not associated with cognition ($p > 0.50$), these "non-cognitive" educational attainment SNPs also attained an r_g of 0.17 with schizophrenia. GNOVA analyses were repeated for SNPs nominally associated with cognition ($p < 0.05$), but generally not associated with education ($p > 0.05$), and the analyses were repeated again with the stricter threshold for education ($p > 0.50$). Values for r_g of −.50 and −.11 were obtained between schizophrenia and these cognition subsets (Figure 2A).

The second approach involved calculating the heterogeneity p values for cognition and education and identifying SNPs that have discrepant direction of effects between

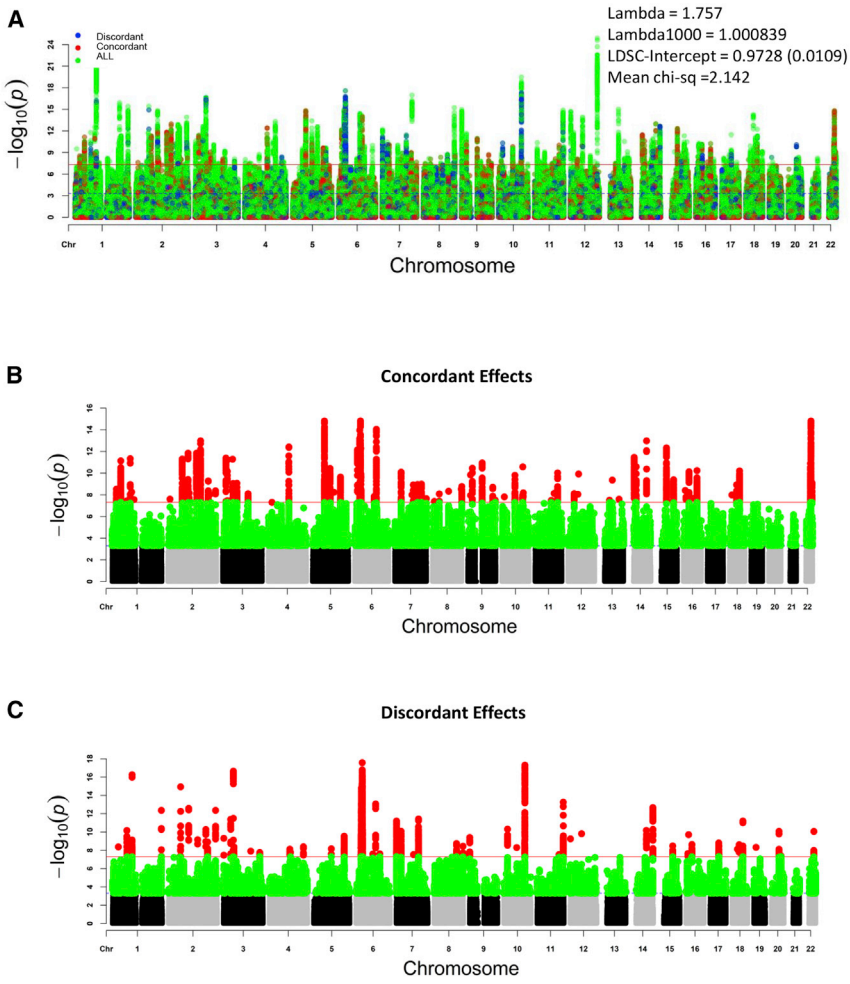

Figure 3. Manhattan Plots for ASSET Results
ASSET meta-analysis outputs:
(A) All subsets.
(B) Concordant subset.
(C) Discordant subset.

cognition and education. These SNPs were then binned, ranging from low probability ($p < 0.5$) to high probability ($p < 0.001$) for heterogeneous effect sizes between cognition and education (Figure 2B). GNOVA indicated that the greater the discrepancy in effect direction between SNP effects for cognition and education, the stronger the association between cognition and schizophrenia. However, this pattern was not observed for education and schizophrenia.

Stage 2: ASSET Meta-Analysis and SNP Subsets

Genome-wide cross-phenotype ASSET meta-analysis across 7,306,098 SNPs revealed 300 lead SNPs (across 236 independent loci) that met the genome-wide significance threshold of $p < 5 \times 10^{-8}$ for the ASSET two-tailed test (see Figure 3A and Tables S1 and S2). There were 1,381,020 SNPs that demonstrated consistent direction of effects between cognition, education, and schizophrenia (i.e., lower cognitive ability, lower educational attainment, and increased risk for schizophrenia); these were assigned to the "concordant" subset, which contained 89 genome-wide-significant loci harboring 103 independent significant SNPs. By contrast, the "discordant" subset, which consisted of SNPs with counter-intuitive allelic effects for schizophrenia vis-à-vis education, encompassed 1,891,743 SNPs, with 65 genome-wide-significant loci comprising 77 independent significant SNPs (Figures 3B and 3C and Table S1). Significant loci for other ASSET subsets are also detailed in Table S2. Other supplemental tables show FUMA-derived annotations for potential functional consequences, including CADD scores (Table S3), eQTL lookups (Table S4), and prior GWAS lookups (Table S5).

Consolidation of Independent Loci

Next, we wanted to identify which loci from our ASSET results were not previously identified with respect to the three-input GWAS. Using RAggr,[43] we extracted SNPs with $r^2 > 0.6$ within a window of 250 kb of lead SNPs in reported GWASs, i.e., we extracted 101 loci from the European-ancestry cohorts of the Psychiatric Genomics Consortium GWAS of schizophrenia,[35] 74 loci from the SSGAC educational attainment GWAS,[36] and 40 loci from the COGENT GWAS of cognitive ability.[18] These were merged with the 236 loci from ASSET. As earlier described, independent loci within 250 kb were merged, resulting in 280 independent loci being identified across ASSET and the input GWAS. As shown in the resulting Venn diagram (Figure 4), 110 loci not reported earlier were identified by the ASSET meta-analysis. In contrast, 126 loci overlapped with either education or schizophrenia, whereas 44 loci were only significant in the input GWAS and not in ASSET.

Very recently, new GWASs have been published for schizophrenia, cognitive ability, and educational attainment, and these studies are larger than the input GWAS used for our ASSET analysis.[20,33,34] This permitted us to perform a lookup of our 110 "novel" ASSET SNPs, thus providing an opportunity to validate ASSET as a tool for locus discovery (Table S6). We also performed lookup in a paper that utilized MTAG to examine intelligence[20] and several recent papers that applied pleiotropic approaches to these phenotypes.[19,23,31] We found that 75% of the loci were in fact reported as significant in the later GWASs with larger sample sizes, and 28 of the 110 loci were independent from other single-phenotype GWAS reports.

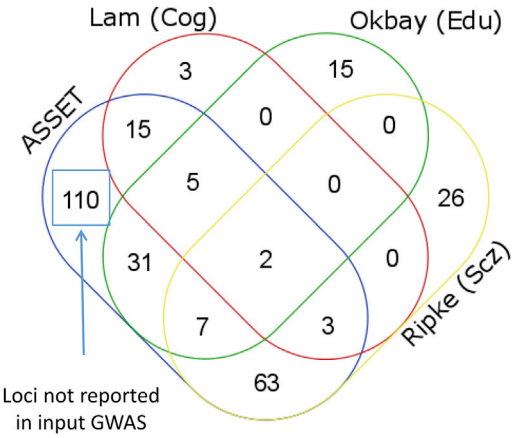

Figure 4. Venn Diagram Comparing Significant ASSET Loci to Significant Loci from Input GWASs
Education = Okbay PMID 27225129; Cognition = Lam PMID 29186694; Schizophrenia (Ripke) = PGC Schizophrenia Working Group PMID 25056061; ASSET = results from ASSET meta-analysis.

These loci are reported in Table 1. Notably, three of these loci (indexed by rs207338, rs708212, and rs11617058) were identified in secondary analyses (using MTAG) performed in the study of educational attainment,[34] and one locus (rs67652508) has been reported as being genome-wide significant for association to putamen volumes as measured on magnetic resonance imaging (MRI) scans of the brain.[60] Further ANNOVAR[41] annotations are available for novel loci (Table S7).

MAGMA Gene-Based Analysis: Tissue-Expression and Competitive-Pathway Analysis

MAGMA gene-based analysis was conducted on all ASSET subsets. 772 genes survived Bonferroni correction in the overall ASSET analysis, with 306 genes in the concordant subset and 304 genes within the discordant subset (Table S8). MAGMA gene property analysis revealed significant association ($p < 0.000926$, Bonferroni-corrected) of gene expression of ASSET SNP subsets across GTExv7 brain tissues (Figure S3 and Table S9). There were no significant differences between concordant and discordant result subsets; both subsets were significantly enriched (positive beta weights) across all brain compartments.

Because of the significant enrichment in brain tissues, we next performed MAGMA competitive-pathway analyses by using neurodevelopmental and other brain-related gene sets as curated in a recent publication;[50] full results are reported in Table S10. Although there was considerable overlap of pathway enrichment across ASSET categories, several gene sets were uniquely associated with either the concordant or the discordant result subsets (Table 2). Specifically, the *CHD8* pathway, reflecting genes involved in early neurodevelopment, was uniquely associated with the concordant subset ($p = 7.11 \times 10^{-6}$). In contrast, a number of synaptic pathways (e.g., ion-channel and synaptic-density pathways) and constrained gene sets appeared to be uniquely associated with the discordant subset. Only two gene sets were enriched in both the concordant and the discordant results, and these were rather generic: brain-enriched genes and schizophrenia GWAS results (Table S10). It is notable that when the MHC region was removed from the pathway analysis, the overall pattern of results remained (see Table S10).

To see whether the distinction between the concordant and discordant subsets harbors potential implications for drug targeting (for schizophrenia and/or cognitive enhancement), we performed drug-based and drug family competitive gene set analyses on our MAGMA results. These analyses revealed that the class of drugs associated with voltage-gated calcium channel genes was over-represented among in the results from the discordant subset (Bonferroni-corrected $p = 0.02$), as was Abacavir (nucleoside reverse transcriptase inhibitor; Bonferroni-corrected $p = 0.00018$). Although both of these sets showed similar direction of effects with respect to the concordant subset, no drug-related gene sets attained Bonferroni-corrected significance in the results from the concordant subset (Table S11).

S-Predixcan: Brain Tissue Expression Profiles and Gene-Set Enrichment Analysis

Transcriptome-wide association analysis (TWAS) was carried out via S-Predixcan to identify top expressed genes within GTEXv7[46] and CommonMind Consortium[54,56–58] brain tissue models (Figure 5 and Table S12). The top brain-expressed genes unique to the discordant subsets were *CYP21A1P* (MIM: 613815), *CFB* (MIM: 138470), and *C4A* (MIM: 120810), along with 177 additional genes that were significantly expressed in the discordant, but not the concordant, subsets. On the other hand, *ELOVL7* (MIM: 614451), *NAGA* (MIM: 609241), and 201 other genes were uniquely associated with the concordant subset. Significant genes identified by S-Predixcan were subjected to GENE2FUNC hypergeometric gene-set analysis (excluding MHC genes, which were over-represented due to significant LD; see Material and Methods for more details). The goal of this analysis was to examine whether the genes identified in the TWAS overlapped with those found in known biological systems. As shown in Table 3, the results of the TWAS consistently identified genes found in cell adhesion and membrane protein gene sets for the concordant subset. In contrast, synaptic (specifically dendritic) pathways, as well as chromosomal repair pathways, were consistently identified by the TWAS during examination of the discordant subset.

Genetic Correlations

A series of psychiatric, personality, structural-brain-imaging, metabolic, cardiovascular, and anthropometric traits were selected for GNOVA modeling with the ASSET subset results (see Figure 6 and Table S13); multiple testing was adjusted on the basis of the false discovery rate (FDR).The concordant subset demonstrated significant (FDR < .05)

Table 1. Loci Identified by ASSET

Lead SNPs	Chr	Base Position Start (bp)	Base Position End (bp)	GWAS *p* Value	Beta (β)	SE	Nearest Gene	SNP Function
rs13010104	2	208331258	208369715	9.30×10^{-9}	−0.0148	0.0026	ENSG00000223725	ncRNA_intronic
rs207338	4	19053350	19070123	4.99×10^{-8}	−0.0108	0.0020	ENSG00000248238	intergenic
rs6844280	4	31371970	31387144	3.53×10^{-8}	0.0113	0.0021	ENSG00000251434	intergenic
rs71615297	4	179013479	179210420	2.17×10^{-8}	−0.0115	0.0021	RNU1-45P	intergenic
rs73260443	5	113404038	113464825	6.30×10^{-9}	−0.0147	0.0025	ENSG00000251628	intergenic
rs9372208	6	109507533	109648130	2.96×10^{-9}	−0.0120	0.0020	ENSG00000233908	intergenic
rs9371912	6	155907483	156009420	1.06×10^{-10}	−0.0168	0.0026	RNU7-152P	intergenic
rs1870571	8	80129136	80296429	4.69×10^{-9}	0.0124	0.0021	ENSG00000253659	intergenic
rs11786685	8	81220185	81316928	2.79×10^{-8}	−0.0116	0.0021	ENSG00000253237	intergenic
rs3890699	8	110164392	110344596	7.31×10^{-9}	0.0122	0.0021	*NUDCD1*	intergenic
rs11166628	8	137022220	137091947	3.27×10^{-8}	0.0113	0.0020	ENSG00000253248	ncRNA_intronic
rs10993909	9	136924744	136942560	3.13×10^{-8}	0.0116	0.0021	*BRD3*	intronic
rs10994707	10	62919886	63096664	1.14×10^{-9}	0.0202	0.0033	*TMEM26*	intergenic
rs72945305	11	81164016	81210528	3.34×10^{-8}	−0.0128	0.0023	ENSG00000254747	intergenic
rs556587	11	92317365	92554716	1.57×10^{-8}	−0.0163	0.0029	*FAT3*	intronic
rs75261250	11	124276497	124303201	2.16×10^{-8}	−0.0188	0.0034	*OR8X1P*	intergenic
rs708212	12	31517960	31769501	7.79×10^{-9}	0.0133	0.0023	*DENND5B*	intronic
rs3741434	12	53605344	53605344	1.20×10^{-10}	0.0191	0.0030	*RARG*	UTR3
rs7321596	13	44390857	44514022	3.16×10^{-8}	−0.0109	0.0020	*LACC1*	intergenic
rs11617058	13	85176690	85305386	2.95×10^{-11}	0.0183	0.0028	*LINC00333*	intergenic
rs67652508	14	55487496	55567991	2.88×10^{-8}	−0.0140	0.0025	*MAPK1IP1L*	intergenic
rs11130	16	15687755	15837246	4.24×10^{-8}	0.0110	0.0020	*NDE1*	UTR3
rs7214058	17	9968014	9995284	6.32×10^{-9}	0.0118	0.0020	*GAS7*	intronic
rs28584904	17	68984046	69007006	2.36×10^{-8}	0.0144	0.0026	ENSG00000271101	intergenic
rs56791590	18	26259012	26496051	7.59×10^{-9}	0.0117	0.0020	ENSG00000265994	intergenic
rs12462428	19	16665215	16738369	2.20×10^{-8}	0.0143	0.0026	ENSG00000141979: *MED26*:CTC-429P9.4	intronic:intronic:intronic
rs5767976	22	48133458	48183889	8.14×10^{-10}	−0.0125	0.0020	RP11-191L9.4	ncRNA_intronic
rs68178377	22	50742346	50771464	4.28×10^{-8}	0.0122	0.0022	*DENND6B*	exonic

Abbreviations are as follows: Chr = chromosome; SE = standard error

genetic correlations, in the expected direction, with many forms of psychopathology in addition to schizophrenia (such forms included ADHD [MIM: 143465], bipolar disorder [MIM: 125480], and major depressive disorder [MDD, MIM: 608516], as well as neuroticism and smoking). This subset also demonstrated a significant (FDR < .05) positive genetic correlation (i.e., better cognition/higher education/lower risk for schizophrenia) with larger volumes of several brain regions (including total intracranial volume) as measured by structural MRI, as well as several measures of infant size and adult height. Significant positive associations were also seen with the personality dimensions of openness and conscientiousness, and (surprisingly) with self-reported cancer; significant negative associations were seen for total cholesterol and triglycerides, as well as the presence of ulcerative colitis and inflammatory bowel disease. Additionally, a negative genetic correlation was observed for the concordant subset with BMI and measures of cardiovascular disease (i.e., lower cognition/lower education/greater risk for schizophrenia associated with greater BMI and risk for cardiovascular disease).

The discordant subset was strongly associated with schizophrenia and education, by definition, in a manner demonstrating the paradoxical relationship (higher education with greater risk for schizophrenia, Figure 6). (It is important to note that the light blue bars and dark blue bars in Figure 6 are essentially mirror images of each other and are therefore providing somewhat redundant

Table 2. MAGMA Significant Gene Sets for Concordant and Discordant SNP Subsets

MAGMA Gene Sets	NGENES	p	Pbon	BETA	BETA_STD	SE
Concordant						
Cotney2015:hNSC_Chd8_prom*	8,186	4.05×10^{-9}	7.11×10^{-6}	0.115	0.0571	0.0199
Cotney2015:hNSC+human+mouse_Chd8_prom	1,902	2.81×10^{-5}	0.049458	0.123	0.0378	0.0304
Sugathan2014:Chd8_binding	5,314	1.69×10^{-5}	0.029744	0.0874	0.04	0.0211
Discordant						
Trapnell2012: constrained_genes_0_10*	1,705	3.41×10^{-7}	0.0006	0.131	0.0383	0.0263
Lek2016: constrained_genes_pLI_90*	2,972	3.33×10^{-7}	0.000585	0.104	0.0387	0.021
Darnell2011:fmrp_targets*	765	1.41×10^{-5}	0.024693	0.162	0.0327	0.0387
ddg2p:dominant_mis_all_brain	137	4.42×10^{-6}	0.007768	0.411	0.0357	0.0925
g2cdb:bayes_collins_mouse_psd_consensus*	918	8.89×10^{-6}	0.015615	0.152	0.0333	0.0353
g2cdb:bayes_collins_mouse_psd_full*	1,442	1.26×10^{-6}	0.002214	0.134	0.0363	0.0284
g2cdb:human_psd*	1,001	1.40×10^{-6}	0.002461	0.159	0.0363	0.0339
g2cdb:human_psp*	1,039	2.04×10^{-6}	0.003579	0.154	0.0358	0.0334
GOBP:membrane_depolarization	105	2.17×10^{-5}	0.038129	0.439	0.0334	0.107
GOBP:synaptic_transmission	549	6.58×10^{-6}	0.011554	0.205	0.0353	0.0471
GOMF:voltage-gated_cation_channel_activity*	144	6.43×10^{-7}	0.00113	0.45	0.0401	0.093
Sanders2015:asd_fdr10*	64	4.67×10^{-6}	0.008207	0.622	0.037	0.14
Sanders2015:asd_lof_genes*	559	3.52×10^{-6}	0.006178	0.203	0.0352	0.0451

Abbreviations are as follows: NGENES = number of genes in pathway; Pbon = Bonferroni-corrected p value; BETA_STD = standardized beta; GOBP = gene ontology biological process; GOMF = gene ontology biological molecular function; Cotney = PMID 25752243; hNSC = human neural stem cells; prom = promotor; Sugathan = PMID 25294932; Trapnell = PMID 22383036; Lek = PMID 27535533; Darnell = PMID 21784246; ddg2p = DECIPHER developmental disorder genotype-phenotype; dominant_mis_all_brain = dominant mode of effect/loss-of-function missense/brain affected; g2cdb = Genes2cognition database; psd = post-synaptic density; psp = post-synaptic proteomes; Sanders = 26402605; asd_fdr_10 = autism risk genes, FDR < 0.10; and asd_lof_genes = ASD loss-of-function genes.
*Results remaining significant after removal of MHC region variants.

information; both sets of bars are included to indicate the both sides of this dimension). Interestingly, a similar pattern was observed for bipolar disorder (higher education/greater risk for schizophrenia—greater risk for bipolar disorder). Similar relationships were also observed, at a nominally significant level, for autism spectrum disorder and eating disorders (MIM: 606788), which were not associated with the concordant subset, as well as for MDD. The reverse relationship, however, was observed with ADHD (i.e., higher education/greater risk for schizophrenia—lower risk for ADHD). This pattern was also observed for the smoking, BMI, and cardiovascular disease phenotypes. A counter-intuitive pattern was observed for the relationship between the discordant subset and neuroticism, which was the opposite of that observed for MDD (despite the fact that MDD and neuroticism are themselves highly correlated).

Discussion

A consistent finding in recent schizophrenia, cognition, and education GWASs has been the involvement of both neurodevelopmental pathways and synaptic processes;[19,20,39,61,62,92] the present study aimed to at least partially disentangle these mechanisms. In this study, we leveraged the genetic pleiotropy underlying three partially overlapping, complex phenotypes in order to identify homogeneous subsets of SNPs with distinct characteristics. Specifically, we were able to parse a subset of SNPs with alleles that were associated in the expected fashion across our three phenotypes of interest: lower cognitive ability, lower educational attainment, and greater risk for schizophrenia. These "concordant" SNPs were characterized by their association with genes and pathways relevant to early neurodevelopmental processes. By contrast, SNPs that demonstrated a counterintuitive, discordant pattern of association (higher educational attainment yet greater risk for schizophrenia) were primarily associated with genes and pathways involved in synaptic function of mature neurons.

This distinction was robustly observed across several methods of functional annotation. First, MAGMA competitive gene-set analysis revealed a significant enrichment of *CHD8*-related genes in the concordant subset (Table 2). *CHD8,* encoding a chromatin remodeling protein, is a gene that has been robustly associated with autism[62–65] but that to date has only limited or anecdotal evidence

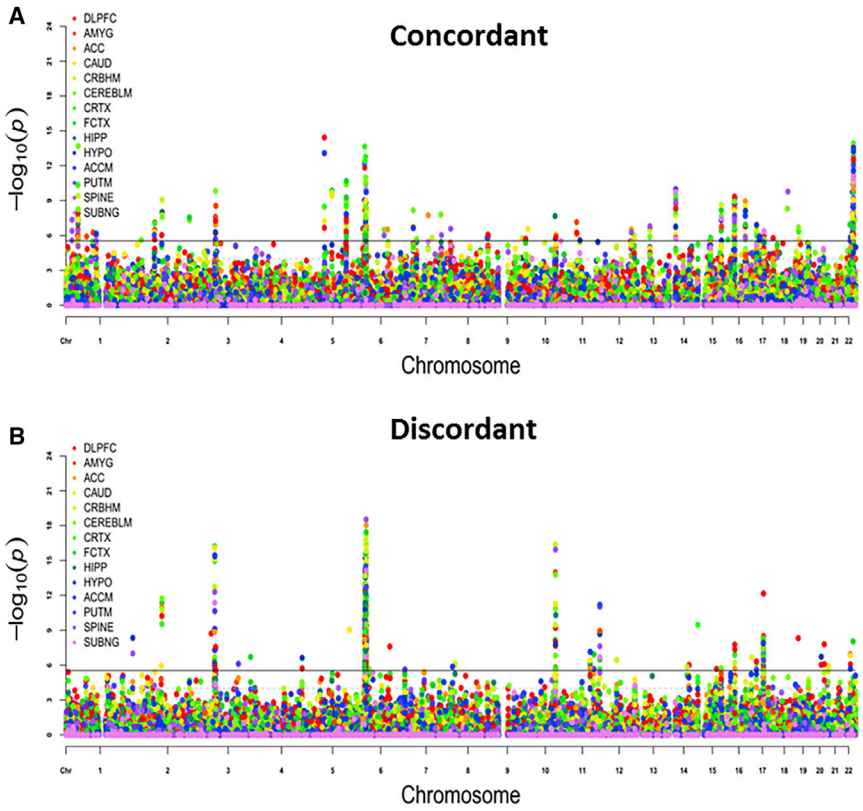

Figure 5. S-PrediXcan Transcriptome-Wide Association Results
(A) Concordant and (B) discordant subsets.

for association to schizophrenia.[66,67] Disruption of the homologous gene (*Chd8*) in animal models has demonstrated that the resulting protein plays a key role in very early neurodevelopmental processes, including neuronal proliferation and differentiation[68,69] as well as cell adhesion and axon guidance.[70] On the other hand, MAGMA competitive gene-set analysis revealed a significant enrichment of discordant genes for functions including synaptic transmission and postsynaptic density, as well as membrane depolarization and voltage-gated cation channel activity. Although these processes have been commonly associated with both schizophrenia[33,35] and cognitive phenotypes,[20,21,24,71–74] our study is the first to demonstrate that these synaptic mechanisms operate in a surprising manner: the same synaptic functions that increase risk for schizophrenia also serve to enhance educational attainment.

The linkage of early neurodevelopmental processes to SNPs associated with impaired cognition and increased risk for schizophrenia is consistent with a large body of literature demonstrating that cognitive deficits are often observed early on in the lifespans of these individuals.[5,13,14,75] At the same time, the discordant variant subset implicates more mature neuronal regulation, and synaptic pruning mechanisms that are most prominent later in childhood, adolescence, and into adulthood, ostensibly as part of a neuroplasticity mechanism for making more "efficient" connections within the brain.[77] However, the dysregulation of such mechanisms has been shown to be intricately linked to schizophrenia psychopathology.[78] It is important to note that these results are obtained from separate GWASs of two different phenotypes and do not represent a subset of highly educated individuals with schizophrenia. Rather, it is plausible to posit an inverted U relationship such that efficient synaptic pruning processes are essential mechanisms underlying academic performance but might be carried too far in disorders such as schizophrenia.

Additionally, transcriptome-wide analysis using S-Predixcan pointed toward the same distinction between concordant and discordant genes and pathways. Two of the strongest genes with differential expression in the concordant subset were *NAGA* (an enzyme cleaving specific moieties from glycoconjugates) and *NDUFAF2* (part of the mitochondrial complex [MIM: 609653]); rare mutations in each of these genes are associated with early and severe neurodevelopmental disorders.[79,80] TWAS of the discordant subset revealed synaptic genes including *C4A*, which plays a key role in synaptic pruning,[78] as well as other genes essential to synapse structure and function; such genes include *ARL3* (MIM: 604695), *FXR1* (MIM: 600819), and *CNNM2* (MIM: 607803). Moreover, pathway analysis of S-Predixcan results (Table 3) demonstrated that the strongest gene set associated with the concordant subset was cell-to-cell adhesion via plasma-membrane adhesion molecules (GO: 0098742); this gene set encompasses processes such as those necessary for neural tube closure, cerebral cortex migration, and neuronal-glial interactions. In contrast, the discordant subset transcriptome was significantly enriched for genes located at dendrites, as well as for genes associated with DNA repair. Recently, the role of DNA repair in modulating neuronal activity-induced gene expression has been shown to be crucial for synaptic plasticity and related processes of learning and memory;[81] impairments in DNA repair have been linked to neurodegeneration[82,83]

ASSET also permitted the identification of SNPs for cognition-related phenotypes. Lookups of the full ASSET results revealed that ~75% of the additional 110 loci, which were not identified in the input GWAS studies,[18,35,36] were in fact replicated in an MTAG study examining intelligence[20] and in more recent follow-up GWASs[20,33,34] that used larger samples and were better powered for variant discovery. This result strongly supports the validity of the ASSET methodology and demonstrates

Table 3. GENE2FUNC Pathway Analysis of GO Genesets with MHC Filtered

Category	GeneSet	N_genes	N_overlap	p	Pbon
Concordant					
GO_bp	GO_CELL_CELL_ADHESION_VIA_PLASMA_MEMBRANE_ADHESION_MOLECULES	202	10	6.89×10^{-9}	3.1×10^{-5}
GO_bp	GO_RESPONSE_TO_XENOBIOTIC_STIMULUS	105	7	6.22×10^{-8}	2.8×10^{-4}
GO_bp	GO_HOMOPHILIC_CELL_ADHESION_VIA_PLASMA_MEMBRANE_ADHESION_MOLECULES	151	8	7.74×10^{-8}	3.4×10^{-4}
GO_bp	GO_CELL_CELL_ADHESION	604	14	4.46×10^{-7}	2.0×10^{-3}
GO_mf	GO_OXIDOREDUCTASE_ACTIVITY_ACTING_ON_NAD_P_H_QUINONE_OR_SIMILAR_COMPOUND_AS_ACCEPTOR	52	4	6.39×10^{-6}	5.8×10^{-3}
GO_cc	GO_MEMBRANE_MICRODOMAIN	286	8	1.53×10^{-5}	8.9×10^{-3}
GO_cc	GO_MEMBRANE_PROTEIN_COMPLEX	1018	16	1.58×10^{-5}	9.2×10^{-3}
GO_bp	GO_PURINE_RIBONUCLEOSIDE_BISPHOSPHATE_METABOLIC_PROCESS	20	3	2.77×10^{-6}	1.2×10^{-2}
GO_cc	GO_MITOCHONDRION	1633	21	2.47×10^{-5}	1.4×10^{-2}
GO_bp	GO_BIOLOGICAL_ADHESION	1027	17	4.56×10^{-6}	2.0×10^{-2}
GO_bp	GO_MACROMOLECULAR_COMPLEX_ASSEMBLY	1388	20	6.87×10^{-6}	3.0×10^{-2}
GO_bp	GO_MITOCHONDRIAL_RESPIRATORY_CHAIN_COMPLEX_I_BIOGENESIS	56	4	9.25×10^{-6}	4.1×10^{-2}
Discordant					
GO_cc	GO_DENDRITIC_SHAFT	37	3	1.28×10^{-5}	7.4×10^{-3}
GO_bp	GO_DOUBLE_STRAND_BREAK_REPAIR	165	6	3.89×10^{-6}	1.7×10^{-2}
GO_cc	GO_NUCLEAR_CHROMOSOME	522	9	3.84×10^{-5}	2.2×10^{-2}
GO_mf	GO_PROTEIN_DOMAIN_SPECIFIC_BINDING	620	10	3.14×10^{-5}	2.8×10^{-2}
GO_bp	GO_NON_RECOMBINATIONAL_REPAIR	70	4	7.97×10^{-6}	3.5×10^{-2}
GO_cc	GO_DENDRITE	451	8	6.94×10^{-5}	4.0×10^{-2}

Note: N_gene = Genes identified in pathway; N_overlap = input genes overlapping with pathway; Pbon = Bonferoni-adjusted p value

that the approach indeed improves power for cross-phenotype discovery of new loci, as previously discussed by the developers of the method.[32] Notably, several of our loci were associated with eQTLs, suggesting new potential biological mechanisms for individual variation in cognitive and psychiatric phenotypes. For example, one of the loci strongly implicates variation in *PLXNB2* (MIM: 604293), a gene associated with GABA and glutamate synapses in the hippocampus.[84] Another locus shows strong eQTL signal with *NDE1* (MIM: 609449), a neurodevelopmental gene at the 16p13.11 locus, where copy-number variants have been associated with neurodevelopmental disorders.[85] Our work supports and extends a recent study by Bansal and colleagues,[30] whose paper is the only published report (to our knowledge) that has deeply examined the paradoxical relationship between educational attainment and schizophrenia. Using a proxy-phenotype approach, these investigators identified two loci, implicating the *FOXO6* (MIM: 611457) and *SLITRK1* (MIM: 609678) genes, with pleiotropic (i.e., "discordant") effects across the two phenotypes. Using ASSET, we also uncovered those genes among our 110 loci (one of which was also not identified in any of the updated single-phenotype GWAS, see Table 1). Several other studies[18,19,23,31] have employed other statistical approaches to identify pleiotropy and/or overlap across cognitive/educational and schizophrenia GWAS and have uncovered a subset of the loci identified by ASSET. By utilizing ASSET, we were able to systematically and powerfully identify concordant and discordant pleiotropic loci across the genome and to then characterize underlying biological mechanisms. Future studies could also apply ASSET and related techniques to further understand other reported polygenic overlaps such as that between schizophrenia and creativity[86] or that between cognitive ability and smoking status.[26]

In addition to functional characterization via pathway analyses, we were able to characterize the concordant and discordant SNP sets with respect to genetic overlap with other relevant phenotypes. To our knowledge, this is the first study to examine genetic correlations with dimensional sub-sets rather than global correlations with full GWASs. Although the concordant subset followed the expected patterns of genetic correlation with several forms of psychopathology, as well as brain and head size, results for the discordant subset were somewhat surprising. For example, we had anticipated that the discordant subset

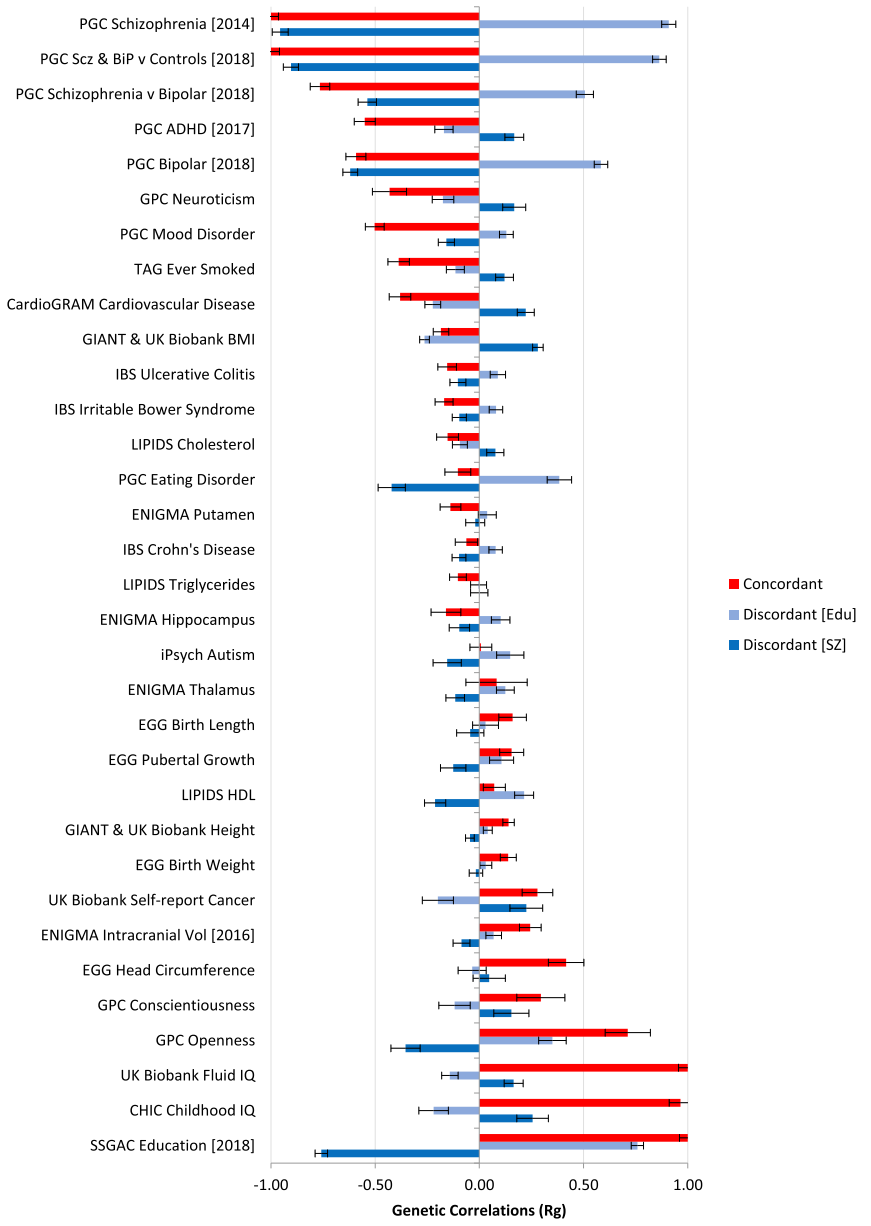

Figure 6. Genetic Correlations for Concordant and Discordant Subsets with Other Relevant Phenotypes

Genetic correlation analysis was carried out with GNOVA. Error bars represent standard errors. Summary statistics of selected phenotypes were downloaded from the LD Hub and Psychiatric Genomics Consortium websites. See Web Resources for further detail.

might be significantly related to personality as a non-cognitive trait that could promote greater educational attainment. However, correlations with conscientiousness, openness, and neuroticism were stronger for the concordant than for the discordant subset.

On the other hand, significant correlations for the discordant subset were observed with risk for autism, which has previously been shown to demonstrate a counter-intuitive positive genetic correlation with cognition.[87] Given that variants within the discordant subset tend to index regulation of synaptic function and pruning processes, our results suggest that these mechanisms should be investigated with respect to their impact on autism, eating disorders, and bipolar disorder. Moreover, it is noteworthy that autism, despite being a neurodevelopmental disorder, did not demonstrate a significant genetic correlation with the concordant subset, indicating that it does not share the specific neurodevelopmental pathways implicated in the common variant genetic overlap between schizophrenia risk and impaired cognition. It is also intriguing that bipolar disorder demonstrated a very similar pattern of GNOVA results to schizophrenia, despite prior reports that bipolar disorder is not significantly correlated at the genetic level with general cognitive ability.[87,88] Thus, our approach was able to refine components of neurodevelopment and synaptic function that are shared across cognitive phenotypes, schizophrenia, and bipolar disorder. Further research is needed to identify components of cognition that differentiate schizophrenia and bipolar disorder.

One limitation of this study is that only common SNPs (MAF $>$ 0.01) were examined. The genetic architecture of cognitive ability and educational attainment is composed of causal variants in LD with common SNPs (cognitive ability h^2 = 22.7%, education h^2 = 15.6%) as well as with causal variants in LD with rare and less-common SNPs (cognitive ability h^2 = 31.3%, education h^2 = 28.1%); rarer variants make greater contributions to cognitive differences than more common variants do.[89] Rare variants are also known to explain some of the differences in schizophrenia prevalence.[50] However, GNOVA, used in the identification of genetic correlations across data-independent datasets using summary GWAS data, can only capture the contributions made by common genetic effects. Future work aiming to investigate the concordant and discordant effect of rare variants across cognitive ability, schizophrenia, and education is needed.[90] Additionally, the input GWASs for ASSET were of somewhat different sample sizes and power, and the cognitive GWAS demonstrated smaller mean effect sizes than those for schizophrenia and educational attainment; the effects

of such differences on ASSET results are not fully understood, although ASSET has been benchmarked as the best available approach to handling non-uniform distribution of effect sizes.[39]

Now that the utility and validity of the ASSET approach has been demonstrated, future studies are planned that can further exploit this method using larger, and more varied, input GWASs. Recent studies have demonstrated that genetic correlations exist across seemingly disparate brain-related phenotypes.[91] However, such genetic correlations only describe the average genetic effect between pairs of traits. As such, they are not informative as to which variants are associated across traits, nor if a minority of these variants have effects across traits that are the opposite of what would be expected on the basis of the direction of the genetic correlation. The application of the ASSET approach to these datasets would help researchers to move beyond the analysis of shared genetic variance and begin to identify shared genetic variants that, as shown in the current study, might be composed of variants with different combinations of protective and deleterious effects. Future studies, employing additional statistical techniques and incorporating rare variants and novel annotation resources, are needed to further decompose the early neurodevelopmental and adult synaptic pathways highlighted in the present report.

Supplemental Data

Supplemental Data can be found online at https://doi.org/10.1016/j.ajhg.2019.06.012.

Declaration of Interests

The authors declare no competing interests.

Received: January 23, 2019
Accepted: June 12, 2019
Published: August 1, 2019

Web Resources

ASSET, https://dceg.cancer.gov/tools/analysis/asset
FUMA, https://fuma.ctglab.nl/
Genes2Cognition database, http://www.genes2cognition.org
LD Hub, http://ldsc.broadinstitute.org/ldhub/
LDSC, https://github.com/bulik/ldsc
MAGMA, https://ctg.cncr.nl/software/magma
NHGRI-EBI catalog of published genome-wide association studies, https://www.genome.gov/gwastudies/
OMIM, https://www.omim.org/
PredictDB, http://predictdb.org/
Psychiatric Genomics Consortium, https://www.med.unc.edu/pgc/results-and-downloads/
S-Predixcan, https://github.com/hakyimlab/MetaXcan
Social Science Genetic Association Consortium, https://www.thessgac.org/data
VEP, http://grch37.ensembl.org/Homo_sapiens/Tools/VEP

References

1. Keefe, R.S.E. (2008). Should cognitive impairment be included in the diagnostic criteria for schizophrenia? World Psychiatry *7*, 22–28.
2. Mesholam-Gately, R.I., Giuliano, A.J., Goff, K.P., Faraone, S.V., and Seidman, L.J. (2009). Neurocognition in first-episode schizophrenia: a meta-analytic review. Neuropsychology *23*, 315–336.
3. Heinrichs, R.W., and Zakzanis, K.K. (1998). Neurocognitive deficit in schizophrenia: a quantitative review of the evidence. Neuropsychology *12*, 426–445.
4. Bilder, R.M., Goldman, R.S., Robinson, D., Reiter, G., Bell, L., Bates, J.A., Pappadopulos, E., Willson, D.F., Alvir, J.M.J., Woerner, M.G., et al. (2000). Neuropsychology of first-episode schizophrenia: initial characterization and clinical correlates. Am. J. Psychiatry *157*, 549–559.
5. Reichenberg, A., Caspi, A., Harrington, H., Houts, R., Keefe, R.S.E., Murray, R.M., Poulton, R., and Moffitt, T.E. (2010). Static and dynamic cognitive deficits in childhood preceding adult schizophrenia: a 30-year study. Am. J. Psychiatry *167*, 160–169.
6. Gottesman, I.I., and Gould, T.D. (2003). The endophenotype concept in psychiatry: etymology and strategic intentions. Am. J. Psychiatry *160*, 636–645.
7. Glahn, D.C., Almasy, L., Blangero, J., Burk, G.M., Estrada, J., Peralta, J.M., Meyenberg, N., Castro, M.P., Barrett, J., Nicolini, H., et al. (2007). Adjudicating neurocognitive endophenotypes for schizophrenia. Am. J. Med. Genet. B. Neuropsychiatr. Genet. *144B*, 242–249.
8. Lencz, T., Smith, C.W., McLaughlin, D., Auther, A., Nakayama, E., Hovey, L., and Cornblatt, B.A. (2006). Generalized and specific neurocognitive deficits in prodromal schizophrenia. Biol. Psychiatry *59*, 863–871.
9. Goldberg, T.E., Burdick, K.E., McCormack, J., Napolitano, B., Patel, R.C., Sevy, S.M., Goldman, R., Lencz, T., Malhotra, A.K., Kane, J.M., and Robinson, D.G. (2009). Lack of an inverse relationship between duration of untreated psychosis and cognitive function in first episode schizophrenia. Schizophr. Res. *107*, 262–266.
10. Cornblatt, B., Obuchowski, M., Roberts, S., Pollack, S., and Erlenmeyer-Kimling, L. (1999). Cognitive and behavioral precursors of schizophrenia. Dev. Psychopathol. *11*, 487–508.
11. Snitz, B.E., Macdonald, A.W., 3rd, and Carter, C.S. (2006). Cognitive deficits in unaffected first-degree relatives of schizophrenia patients: a meta-analytic review of putative endophenotypes. Schizophr. Bull. *32*, 179–194.
12. Reichenberg, A., Weiser, M., Rabinowitz, J., Caspi, A., Schmeidler, J., Mark, M., Kaplan, Z., and Davidson, M. (2002). A population-based cohort study of premorbid intellectual, language, and behavioral functioning in patients with schizophrenia, schizoaffective disorder, and nonpsychotic bipolar disorder. Am. J. Psychiatry *159*, 2027–2035.
13. Bilder, R.M., Reiter, G., Bates, J., Lencz, T., Szeszko, P., Goldman, R.S., Robinson, D., Lieberman, J.A., and Kane, J.M. (2006). Cognitive development in schizophrenia: follow-back from the first episode. J. Clin. Exp. Neuropsychol. *28*, 270–282.
14. Lam, M., Lee, J., Rapisarda, A., See, Y.M., Yang, Z., Lee, S.-A., Abdul-Rashid, N.A., Kraus, M., Subramaniam, M., Chong, S.-A., and Keefe, R.S.E. (2018). longitudinal cognitive changes in young individuals at ultrahigh risk for psychosis. JAMA Psychiatry *75*, 929–939.

15. MacCabe, J.H., Lambe, M.P., Cnattingius, S., Torrång, A., Björk, C., Sham, P.C., David, A.S., Murray, R.M., and Hultman, C.M. (2008). Scholastic achievement at age 16 and risk of schizophrenia and other psychoses: a national cohort study. Psychol. Med. *38*, 1133–1140.
16. Chong, S.A., Subramaniam, M., Lee, I.-M., Pek, E., Cheok, C., Verma, S., and Wong, J. (2009). Academic attainment: a predictor of psychiatric disorders? Soc. Psychiatry Psychiatr. Epidemiol. *44*, 999–1004.
17. Lencz, T., Knowles, E., Davies, G., Guha, S., Liewald, D.C., Starr, J.M., Djurovic, S., Melle, I., Sundet, K., Christoforou, A., et al. (2014). Molecular genetic evidence for overlap between general cognitive ability and risk for schizophrenia: a report from the Cognitive Genomics Consortium (COGENT). Mol. Psychiatry *19*, 168–174.
18. Lam, M., Trampush, J.W., Yu, J., Knowles, E., Davies, G., Liewald, D.C., Starr, J.M., Djurovic, S., Melle, I., Sundet, K., et al. (2017). Large-scale cognitive GWAS meta-analysis reveals tissue-specific neural expression and potential nootropic drug targets. Cell Rep. *21*, 2597–2613.
19. Smeland, O.B., Frei, O., Kauppi, K., Hill, W.D., Li, W., Wang, Y., Krull, F., Bettella, F., Eriksen, J.A., Witoelar, A., et al.; NeuroCHARGE (Cohorts for Heart and Aging Research in Genomic Epidemiology) Cognitive Working Group (2017). Identification of genetic loci jointly influencing schizophrenia risk and the cognitive traits of verbal-numerical reasoning, reaction time, and general cognitive function. JAMA Psychiatry *74*, 1065–1075.
20. Savage, J.E., Jansen, P.R., Stringer, S., Watanabe, K., Bryois, J., de Leeuw, C.A., Nagel, M., Awasthi, S., Barr, P.B., Coleman, J.R.I., et al. (2018). Genome-wide association meta-analysis in 269,867 individuals identifies new genetic and functional links to intelligence. Nat. Genet. *50*, 912–919.
21. Davies, G., Lam, M., Harris, S.E., Trampush, J.W., Luciano, M., Hill, W.D., Hagenaars, S.P., Ritchie, S.J., Marioni, R.E., Fawns-Ritchie, C., et al. (2018). Study of 300,486 individuals identifies 148 independent genetic loci influencing general cognitive function. Nat. Commun. *9*, 2098.
22. Hill, W.D., Davies, G., Liewald, D.C., McIntosh, A.M., Deary, I.J.; and CHARGE Cognitive Working Group (2016). Age-dependent pleiotropy between general cognitive function and major psychiatric disorders. Biol. Psychiatry *80*, 266–273.
23. Hill, W.D., Marioni, R.E., Maghzian, O., Ritchie, S.J., Hagenaars, S.P., McIntosh, A.M., Gale, C.R., Davies, G., and Deary, I.J. (2019). A combined analysis of genetically correlated traits identifies 187 loci and a role for neurogenesis and myelination in intelligence. Mol. Psychiatry *24*, 169–181.
24. Davies, G., Marioni, R.E., Liewald, D.C., Hill, W.D., Hagenaars, S.P., Harris, S.E., Ritchie, S.J., Luciano, M., Fawns-Ritchie, C., Lyall, D., et al. (2016). Genome-wide association study of cognitive functions and educational attainment in UK Biobank (N=112 151). Mol. Psychiatry *21*, 758–767.
25. Bulik-Sullivan, B.K., Loh, P.-R., Finucane, H.K., Ripke, S., Yang, J., Patterson, N., Daly, M.J., Price, A.L., Neale, B.M.; and Schizophrenia Working Group of the Psychiatric Genomics Consortium (2015). LD Score regression distinguishes confounding from polygenicity in genome-wide association studies. Nat. Genet. *47*, 291–295.
26. Hagenaars, S.P., Harris, S.E., Davies, G., Hill, W.D., Liewald, D.C.M., Ritchie, S.J., Marioni, R.E., Fawns-Ritchie, C., Cullen, B., Malik, R., et al.; METASTROKE Consortium, International Consortium for Blood Pressure GWAS; SpiroMeta Consortium; and CHARGE Consortium Pulmonary Group, CHARGE Consortium Aging and Longevity Group (2016). Shared genetic aetiology between cognitive functions and physical and mental health in UK Biobank (N=112 151) and 24 GWAS consortia. Mol. Psychiatry *21*, 1624–1632.
27. Trampush, J.W., Yang, M.L.Z., Yu, J., Knowles, E., Davies, G., Liewald, D.C., Starr, J.M., Djurovic, S., Melle, I., Sundet, K., et al. (2017). GWAS meta-analysis reveals novel loci and genetic correlates for general cognitive function: a report from the COGENT consortium. Mol. Psychiatry *22*, 336–345.
28. Bulik-Sullivan, B., Finucane, H.K., Anttila, V., Gusev, A., Day, F.R., Loh, P.R., Duncan, L., Perry, J.R., Patterson, N., Robinson, E.B., et al.; ReproGen Consortium; Psychiatric Genomics Consortium; and Genetic Consortium for Anorexia Nervosa of the Wellcome Trust Case Control Consortium 3 (2015). An atlas of genetic correlations across human diseases and traits. Nat. Genet. *47*, 1236–1241, Bulik-Sullivan, B.
29. Turley, P., Walters, R.K., Maghzian, O., Okbay, A., Lee, J.J., Fontana, M.A., Nguyen-Viet, T.A., Wedow, R., Zacher, M., Furlotte, N.A., et al.; 23andMe Research Team; and Social Science Genetic Association Consortium (2018). Multi-trait analysis of genome-wide association summary statistics using MTAG. Nat. Genet. *50*, 229–237.
30. Bansal, V., Mitjans, M., Burik, C.A.P., Linnér, R.K., Okbay, A., Rietveld, C.A., Begemann, M., Bonn, S., Ripke, S., de Vlaming, R., et al. (2018). Genome-wide association study results for educational attainment aid in identifying genetic heterogeneity of schizophrenia. Nat. Commun. *9*, 3078.
31. Le Hellard, S., Wang, Y., Witoelar, A., Zuber, V., Bettella, F., Hugdahl, K., Espeseth, T., Steen, V.M., Melle, I., Desikan, R., et al.; Schizophrenia Working Group of the Psychiatric Genomics Consortium (2017). Identification of gene loci that overlap between schizophrenia and educational attainment. Schizophr. Bull. *43*, 654–664.
32. Bhattacharjee, S., Rajaraman, P., Jacobs, K.B., Wheeler, W.A., Melin, B.S., Hartge, P., Yeager, M., Chung, C.C., Chanock, S.J., Chatterjee, N.; and GliomaScan Consortium (2012). A subset-based approach improves power and interpretation for the combined analysis of genetic association studies of heterogeneous traits. Am. J. Hum. Genet. *90*, 821–835.
33. Pardiñas, A.F., Holmans, P., Pocklington, A.J., Escott-Price, V., Ripke, S., Carrera, N., Legge, S.E., Bishop, S., Cameron, D., Hamshere, M.L., et al.; GERAD1 Consortium; and CRESTAR Consortium (2018). Common schizophrenia alleles are enriched in mutation-intolerant genes and in regions under strong background selection. Nat. Genet. *50*, 381–389.
34. Lee, J.J., Wedow, R., Okbay, A., Kong, E., Maghzian, O., Zacher, M., Nguyen-Viet, T.A., Bowers, P., Sidorenko, J., Karlsson Linnér, R., et al.; 23andMe Research Team; COGENT (Cognitive Genomics Consortium); and Social Science Genetic Association Consortium (2018). Gene discovery and polygenic prediction from a genome-wide association study of educational attainment in 1.1 million individuals. Nat. Genet. *50*, 1112–1121.
35. Schizophrenia Working Group of the Psychiatric Genomics Consortium (2014). Biological insights from 108 schizophrenia-associated genetic loci. Nature *511*, 421–427.
36. Okbay, A., Beauchamp, J.P., Fontana, M.A., Lee, J.J., Pers, T.H., Rietveld, C.A., Turley, P., Chen, G.-B., Emilsson, V., Meddens, S.F.W., et al.; LifeLines Cohort Study (2016). Genome-wide association study identifies 74 loci associated with educational attainment. Nature *533*, 539–542.

37. Willer, C.J., Li, Y., and Abecasis, G.R. (2010). METAL: fast and efficient meta-analysis of genomewide association scans. Bioinformatics *26*, 2190–2191.
38. Lu, Q., Li, B., Ou, D., Erlendsdottir, M., Powles, R.L., Jiang, T., Hu, Y., Chang, D., Jin, C., Dai, W., et al. (2017). A powerful approach to estimating annotation-stratified genetic covariance via GWAS summary statistics. Am. J. Hum. Genet. *101*, 939–964.
39. Zhu, Z., Anttila, V., Smoller, J.W., and Lee, P.H. (2018). Statistical power and utility of meta-analysis methods for cross-phenotype genome-wide association studies. PLoS ONE *13*, e0193256.
40. Watanabe, K., Taskesen, E., van Bochoven, A., and Posthuma, D. (2017). Functional mapping and annotation of genetic associations with FUMA. Nat. Commun. *8*, 1826.
41. Wang, K., Li, M., and Hakonarson, H. (2010). ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. Nucleic Acids Res. *38*, e164.
43. Barrett, J.C., Fry, B., Maller, J., and Daly, M.J. (2005). Haploview: analysis and visualization of LD and haplotype maps. Bioinformatics *21*, 263–265.
44. de Leeuw, C.A., Mooij, J.M., Heskes, T., and Posthuma, D. (2015). MAGMA: generalized gene-set analysis of GWAS data. PLoS Comput. Biol. *11*, e1004219.
45. GTEx Consortium (2015). Human genomics. The Genotype-Tissue Expression (GTEx) pilot analysis: multitissue gene regulation in humans. Science *348*, 648–660.
46. Lonsdale, J., Thomas, J., Salvatore, M., Phillips, R., Lo, E., Shad, S., Hasz, R., Walters, G., Garcia, F., Young, N., et al.; GTEx Consortium (2013). The Genotype-Tissue Expression (GTEx) project. Nat. Genet. *45*, 580–585.
47. Gamazon, E.R., Segrè, A.V., van de Bunt, M., Wen, X., Xi, H.S., Hormozdiari, F., Ongen, H., Konkashbaev, A., Derks, E.M., Aguet, F., et al.; GTEx Consortium (2018). Using an atlas of gene regulation across 44 human tissues to inform complex disease- and trait-associated variation. Nat. Genet. *50*, 956–967.
48. Gaspar, H.A., and Breen, G. (2016). Pathways analyses of schizophrenia GWAS focusing on known and novel drug targets. bioRxiv.
49. Gaspar, H.A., and Breen, G. (2017). Drug enrichment and discovery from schizophrenia genome-wide association results: an analysis and visualisation approach. Sci. Rep. *7*, 12460.
50. Singh, T., Walters, J.T.R., Johnstone, M., Curtis, D., Suvisaari, J., Torniainen, M., Rees, E., Iyegbe, C., Blackwood, D., McIntosh, A.M., et al.; INTERVAL Study; and UK10K Consortium (2017). The contribution of rare variants to risk of schizophrenia in individuals with and without intellectual disability. Nat. Genet. *49*, 1167–1173.
51. Kang, H.J., Kawasawa, Y.I., Cheng, F., Zhu, Y., Xu, X., Li, M., Sousa, A.M., Pletikos, M., Meyer, K.A., Sedmak, G., et al. (2011). Spatio-temporal transcriptome of the human brain. Nature *478*, 483–489.
52. Purcell, S.M., Moran, J.L., Fromer, M., Ruderfer, D., Solovieff, N., Roussos, P., O'Dushlaine, C., Chambert, K., Bergen, S.E., Kähler, A., et al. (2014). A polygenic burden of rare disruptive mutations in schizophrenia. Nature *506*, 185–190.
53. Barbeira, A.N., Dickinson, S.P., Bonazzola, R., Zheng, J., Wheeler, H.E., Torres, J.M., Torstenson, E.S., Shah, K.P., Garcia, T., Edwards, T.L., et al.; GTEx Consortium (2018). Exploring the phenotypic consequences of tissue specific gene expression variation inferred from GWAS summary statistics. Nat. Commun. *9*, 1825.
54. Dobbyn, A., Huckins, L.M., Boocock, J., Sloofman, L.G., Glicksberg, B.S., Giambartolomei, C., Hoffman, G.E., Perumal, T.M., Girdhar, K., Jiang, Y., et al.; CommonMind Consortium (2018). Landscape of conditional eQTL in dorsolateral prefrontal cortex and co-localization with schizophrenia GWAS. Am. J. Hum. Genet. *102*, 1169–1184.
55. Gamazon, E.R., Wheeler, H.E., Shah, K.P., Mozaffari, S.V., Aquino-Michaels, K., Carroll, R.J., Eyler, A.E., Denny, J.C., Nicolae, D.L., Cox, N.J., Im, H.K.; and GTEx Consortium (2015). A gene-based association method for mapping traits using reference transcriptome data. Nat. Genet. *47*, 1091–1098.
56. Fromer, M., Roussos, P., Sieberts, S.K., Johnson, J.S., Kavanagh, D.H., Perumal, T.M., Ruderfer, D.M., Oh, E.C., Topol, A., Shah, H.R., et al. (2016). Gene expression elucidates functional impact of polygenic risk for schizophrenia. Nat. Neurosci. *19*, 1442–1453.
57. Senthil, G., Dutka, T., Bingaman, L., and Lehner, T. (2017). Genomic resources for the study of neuropsychiatric disorders. Mol. Psychiatry *22*, 1659–1663.
58. Hauberg, M.E., Zhang, W., Giambartolomei, C., Franzén, O., Morris, D.L., Vyse, T.J., Ruusalepp, A., Sklar, P., Schadt, E.E., Björkegren, J.L.M., Roussos, P.; and CommonMind Consortium (2017). Large-scale identification of common trait and disease variants affecting gene expression. Am. J. Hum. Genet. *100*, 885–894.
59. Trowsdale, J., and Knight, J.C. (2013). Major histocompatibility complex genomics and human disease. Annu. Rev. Genomics Hum. Genet. *14*, 301–323.
60. Chen, C.-H., Wang, Y., Lo, M.-T., Schork, A., Fan, C.-C., Holland, D., Kauppi, K., Smeland, O.B., Djurovic, S., Sanyal, N., et al. (2017). Leveraging genome characteristics to improve gene discovery for putamen subcortical brain structure. Sci. Rep. *7*, 15736.
61. O'Dushlaine, C., Kenny, E., Heron, E., Donohoe, G., Gill, M., Morris, D., Corvin, A.; and International Schizophrenia Consortium (2011). Molecular pathways involved in neuronal cell adhesion and membrane scaffolding contribute to schizophrenia and bipolar disorder susceptibility. Mol. Psychiatry *16*, 286–292.
62. Neale, B.M., Kou, Y., Liu, L., Ma'ayan, A., Samocha, K.E., Sabo, A., Lin, C.-F., Stevens, C., Wang, L.-S., Makarov, V., et al. (2012). Patterns and rates of exonic de novo mutations in autism spectrum disorders. Nature *485*, 242–245.
63. Bernier, R., Golzio, C., Xiong, B., Stessman, H.A., Coe, B.P., Penn, O., Witherspoon, K., Gerdts, J., Baker, C., Vulto-van Silfhout, A.T., et al. (2014). Disruptive CHD8 mutations define a subtype of autism early in development. Cell *158*, 263–276.
64. Krumm, N., Turner, T.N., Baker, C., Vives, L., Mohajeri, K., Witherspoon, K., Raja, A., Coe, B.P., Stessman, H.A., He, Z.-X., et al. (2015). Excess of rare, inherited truncating mutations in autism. Nat. Genet. *47*, 582–588.
65. Wang, T., Guo, H., Xiong, B., Stessman, H.A.F., Wu, H., Coe, B.P., Turner, T.N., Liu, Y., Zhao, W., Hoekzema, K., et al. (2016). De novo genic mutations among a Chinese autism spectrum disorder cohort. Nat. Commun. *7*, 13316.
66. McCarthy, S.E., Gillis, J., Kramer, M., Lihm, J., Yoon, S., Berstein, Y., Mistry, M., Pavlidis, P., Solomon, R., Ghiban, E., et al. (2014). De novo mutations in schizophrenia implicate chromatin remodeling and support a genetic overlap with

autism and intellectual disability. Mol. Psychiatry *19*, 652–658.
67. Kimura, H., Wang, C., Ishizuka, K., Xing, J., Takasaki, Y., Kushima, I., Aleksic, B., Uno, Y., Okada, T., Ikeda, M., et al. (2016). Identification of a rare variant in CHD8 that contributes to schizophrenia and autism spectrum disorder susceptibility. Schizophr. Res. *178*, 104–106.
68. Durak, O., Gao, F., Kaeser-Woo, Y.J., Rueda, R., Martorell, A.J., Nott, A., Liu, C.Y., Watson, L.A., and Tsai, L.-H. (2016). Chd8 mediates cortical neurogenesis via transcriptional regulation of cell cycle and Wnt signaling. Nat. Neurosci. *19*, 1477–1488.
69. Gompers, A.L., Su-Feher, L., Ellegood, J., Copping, N.A., Riyadh, M.A., Stradleigh, T.W., Pride, M.C., Schaffler, M.D., Wade, A.A., Catta-Preta, R., et al. (2017). Germline Chd8 haploinsufficiency alters brain development in mouse. Nat. Neurosci. *20*, 1062–1073.
70. Suetterlin, P., Hurley, S., Mohan, C., Riegman, K.L.H., Pagani, M., Caruso, A., Ellegood, J., Galbusera, A., Crespo-Enriquez, I., Michetti, C., et al. (2018). Altered neocortical gene expression, brain overgrowth and functional over-connectivity in chd8 haploinsufficient mice. Cereb. Cortex *28*, 2192–2206.
71. Sniekers, S., Stringer, S., Watanabe, K., Jansen, P.R., Coleman, J.R.I., Krapohl, E., Taskesen, E., Hammerschlag, A.R., Okbay, A., Zabaneh, D., et al. (2017). Genome-wide association meta-analysis of 78,308 individuals identifies new loci and genes influencing human intelligence. Nat. Genet. *49*, 1107–1112.
72. Davies, G., Armstrong, N., Bis, J.C., Bressler, J., Chouraki, V., Giddaluru, S., Hofer, E., Ibrahim-Verbaas, C.A., Kirin, M., Lahti, J., et al.; Generation Scotland (2015). Genetic contributions to variation in general cognitive function: a meta-analysis of genome-wide association studies in the CHARGE consortium (N=53949). Mol. Psychiatry *20*, 183–192.
73. Hill, W.D., Davies, G., van de Lagemaat, L.N., Christoforou, A., Marioni, R.E., Fernandes, C.P.D., Liewald, D.C., Croning, M.D.R., Payton, A., Craig, L.C.A., et al. (2014). Human cognitive ability is influenced by genetic variation in components of postsynaptic signalling complexes assembled by NMDA receptors and MAGUK proteins. Transl. Psychiatry *4*, e341.
74. Fernández, E., Collins, M.O., Frank, R.A.W., Zhu, F., Kopanitsa, M.V., Nithianantharajah, J., Lemprière, S.A., Fricker, D., Elsegood, K.A., McLaughlin, C.L., et al. (2017). Arc requires PSD95 for assembly into postsynaptic complexes involved with neural dysfunction and intelligence. Cell Rep. *21*, 679–691.
75. Reichenberg, A., Weiser, M., Rapp, M.A., Rabinowitz, J., Caspi, A., Schmeidler, J., Knobler, H.Y., Lubin, G., Nahon, D., Harvey, P.D., and Davidson, M. (2005). Elaboration on premorbid intellectual performance in schizophrenia: premorbid intellectual decline and risk for schizophrenia. Arch. Gen. Psychiatry *62*, 1297–1304.
77. Denève, S., Alemi, A., and Bourdoukan, R. (2017). The Brain as an efficient and robust adaptive learner. Neuron *94*, 969–977.
78. Sekar, A., Bialas, A.R., de Rivera, H., Davis, A., Hammond, T.R., Kamitaki, N., Tooley, K., Presumey, J., Baum, M., Van Doren, V., et al.; Schizophrenia Working Group of the Psychiatric Genomics Consortium (2016). Schizophrenia risk from complex variation of complement component 4. Nature *530*, 177–183.
79. Keulemans, J.L., Reuser, A.J., Kroos, M.A., Willemsen, R., Hermans, M.M., van den Ouweland, A.M., de Jong, J.G., Wevers, R.A., Renier, W.O., Schindler, D., et al. (1996). Human alpha-N-acetylgalactosaminidase (alpha-NAGA) deficiency: new mutations and the paradox between genotype and phenotype. J. Med. Genet. *33*, 458–464.
80. Ghaloul-Gonzalez, L., Goldstein, A., Walsh Vockley, C., Dobrowolski, S.F., Biery, A., Irani, A., Ibarra, J., Morton, D.H., Mohsen, A.-W., and Vockley, J. (2016). Mitochondrial respiratory chain disorders in the Old Order Amish population. Mol. Genet. Metab. *118*, 296–303.
81. Shiwaku, H., and Okazawa, H. (2015). Impaired DNA damage repair as a common feature of neurodegenerative diseases and psychiatric disorders. Curr. Mol. Med. *15*, 119–128.
82. Madabhushi, R., Pan, L., and Tsai, L.-H. (2014). DNA damage and its links to neurodegeneration. Neuron *83*, 266–282.
83. Su, Y., Ming, G.L., and Song, H. (2015). DNA damage and repair regulate neuronal gene expression. Cell Res. *25*, 993–994.
84. McDermott, J.E., Goldblatt, D., and Paradis, S. (2018). Class 4 Semaphorins and Plexin-B receptors regulate GABAergic and glutamatergic synapse development in the mammalian hippocampus. Mol. Cell. Neurosci. *92*, 50–66.
85. Grayton, H.M., Fernandes, C., Rujescu, D., and Collier, D.A. (2012). Copy number variations in neurodevelopmental disorders. Prog. Neurobiol. *99*, 81–91.
86. Power, R.A., Steinberg, S., Bjornsdottir, G., Rietveld, C.A., Abdellaoui, A., Nivard, M.M., Johannesson, M., Galesloot, T.E., Hottenga, J.J., Willemsen, G., et al. (2015). Polygenic risk scores for schizophrenia and bipolar disorder predict creativity. Nat. Neurosci. *18*, 953–955.
87. Hill, W.D., Harris, S.E., and Deary, I.J. (2018). What genome-wide association studies reveal about the association between intelligence and mental health. Curr. Opin. Psychol. *27*, 25–30.
88. Gale, C.R., Batty, G.D., McIntosh, A.M., Porteous, D.J., Deary, I.J., and Rasmussen, F. (2013). Is bipolar disorder more common in highly intelligent people? A cohort study of a million men. Mol. Psychiatry *18*, 190–194.
89. Hill, W.D., Arslan, R.C., Xia, C., Luciano, M., Amador, C., Navarro, P., Hayward, C., Nagy, R., Porteous, D.J., McIntosh, A.M., et al. (2018). Genomic analysis of family data reveals additional genetic effects on intelligence and personality. Mol. Psychiatry *23*, 2347–2362.
90. Evans, L.M., Tahmasbi, R., Vrieze, S.I., Abecasis, G.R., Das, S., Gazal, S., Bjelland, D.W., de Candia, T.R., Goddard, M.E., Neale, B.M., et al.; Haplotype Reference Consortium (2018). Comparison of methods that use whole genome data to estimate the heritability and genetic architecture of complex traits. Nat. Genet. *50*, 737–745.
91. Brainstorm Consortium, Anttila, V., Bulik-Sullivan, B., Finucane, H.K., Walters, R.K., Bras, J., Duncan, L., Escott-Price, V., Falcone, G.J., Gormley, P., et al. (2018). Analysis of shared heritability in common disorders of the brain. Science *360*, eaap8757.
92. Network and Pathway Analysis Subgroup of Psychiatric Genomics Consortium (2015). Psychiatric genome-wide association study analyses implicate neuronal, immune and histone pathways. Nat. Neurosci. *18*, 199–209.

The American Journal of Human Genetics, Volume *105*

Supplemental Data

Pleiotropic Meta-Analysis of Cognition, Education, and Schizophrenia Differentiates Roles of Early Neurodevelopmental and Adult Synaptic Pathways

Max Lam, W. David Hill, Joey W. Trampush, Jin Yu, Emma Knowles, Gail Davies, Eli Stahl, Laura Huckins, David C. Liewald, Srdjan Djurovic, Ingrid Melle, Kjetil Sundet, Andrea Christoforou, Ivar Reinvang, Pamela DeRosse, Astri J. Lundervold, Vidar M. Steen, Thomas Espeseth, Katri Räikkönen, Elisabeth Widen, Aarno Palotie, Johan G. Eriksson, Ina Giegling, Bettina Konte, Annette M. Hartmann, Panos Roussos, Stella Giakoumaki, Katherine E. Burdick, Antony Payton, William Ollier, Ornit Chiba-Falek, Deborah K. Attix, Anna C. Need, Elizabeth T. Cirulli, Aristotle N. Voineskos, Nikos C. Stefanis, Dimitrios Avramopoulos, Alex Hatzimanolis, Dan E. Arking, Nikolaos Smyrnis, Robert M. Bilder, Nelson A. Freimer, Tyrone D. Cannon, Edythe London, Russell A. Poldrack, Fred W. Sabb, Eliza Congdon, Emily Drabant Conley, Matthew A. Scult, Dwight Dickinson, Richard E. Straub, Gary Donohoe, Derek Morris, Aiden Corvin, Michael Gill, Ahmad R. Hariri, Daniel R. Weinberger, Neil Pendleton, Panos Bitsios, Dan Rujescu, Jari Lahti, Stephanie Le Hellard, Matthew C. Keller, Ole A. Andreassen, Ian J. Deary, David C. Glahn, Anil K. Malhotra, and Todd Lencz

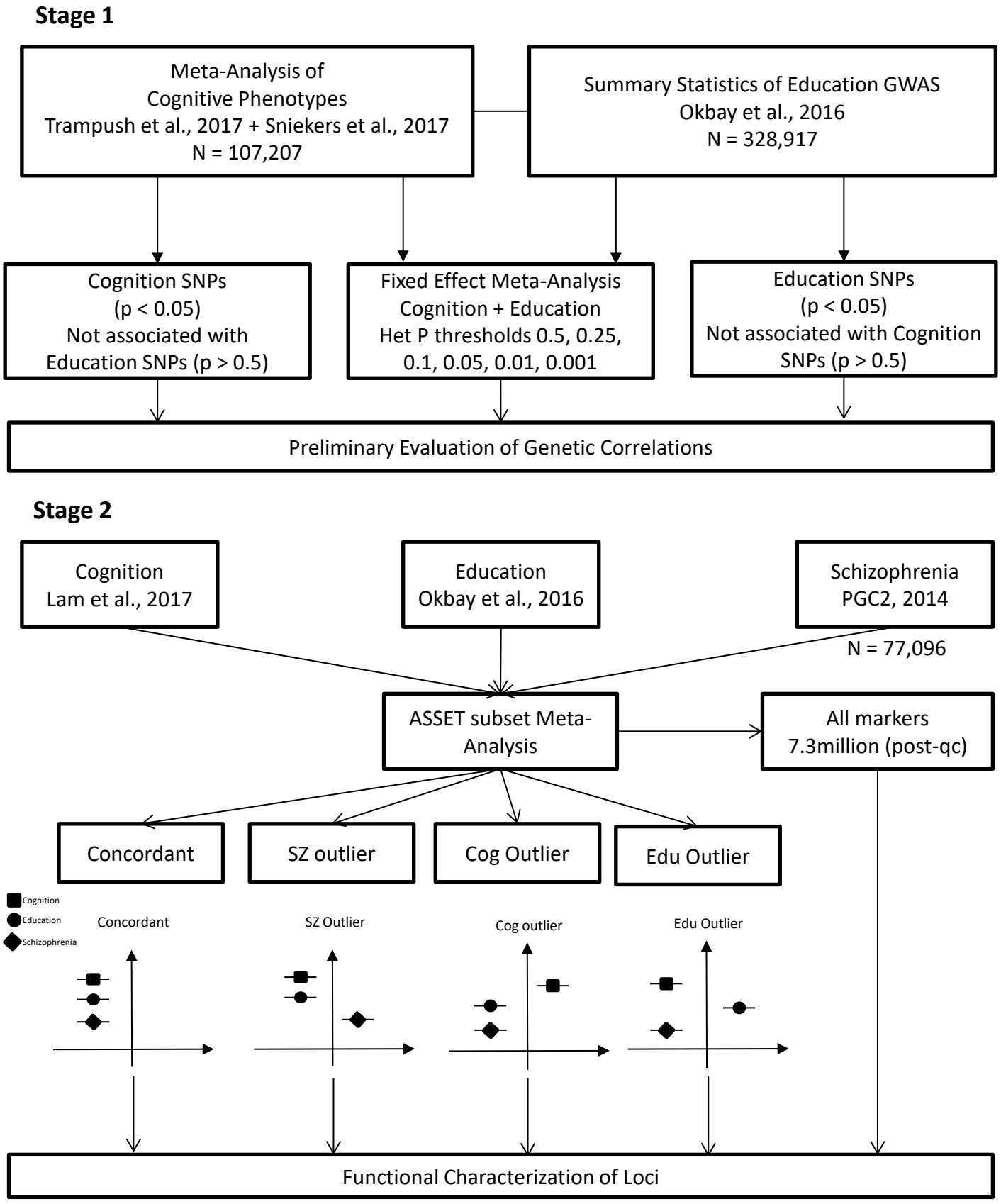

Figure S1: Preliminary Analysis and ASSET Meta Analysis Workflow

Note: Summary statistics data was first consolidated in Stage 1 of the analysis, followed by performing preliminary genetic correlations to evaluate global genetic correlations across the traits of interests, i.e. cognition, education and schizophrenia. ASSET meta-analysis was carried during Stage 2 of the analysis followed by functional characterization of loci, genes and pathways obtained from the ASSET results.

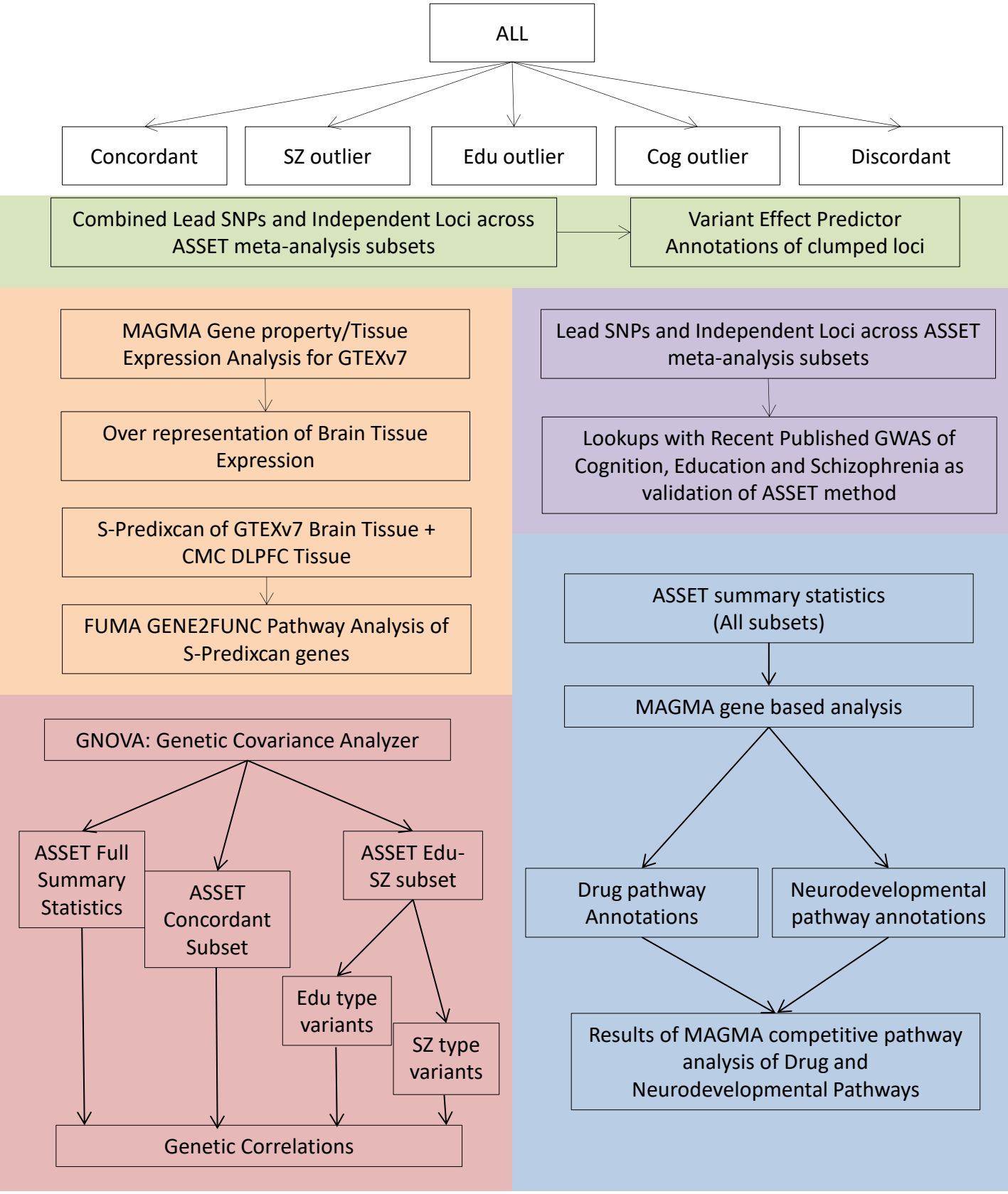

Figure S2: GWAS Loci Functional Characterization Workflow

Note: Functional characterization involved using tools such as MAGMA for gene-based and competitive pathway analysis, S-Predixcan was carried out to identify eQTL profiles in the brain, driven by ASSET meta-analysis, and GNOVA was utilized to examine localized genetic correlations of subsets identified by ASSET. A myriad of annotation databases were used including, GTEX7, CommonMind Consortium, DLPFC expression, and Drug pathway annotations. The FUMA annotation pipeline was also used to identify significant independent genomic loci from the ASSET meta-analysis, which also enabled loci lookups with more recent GWASs.

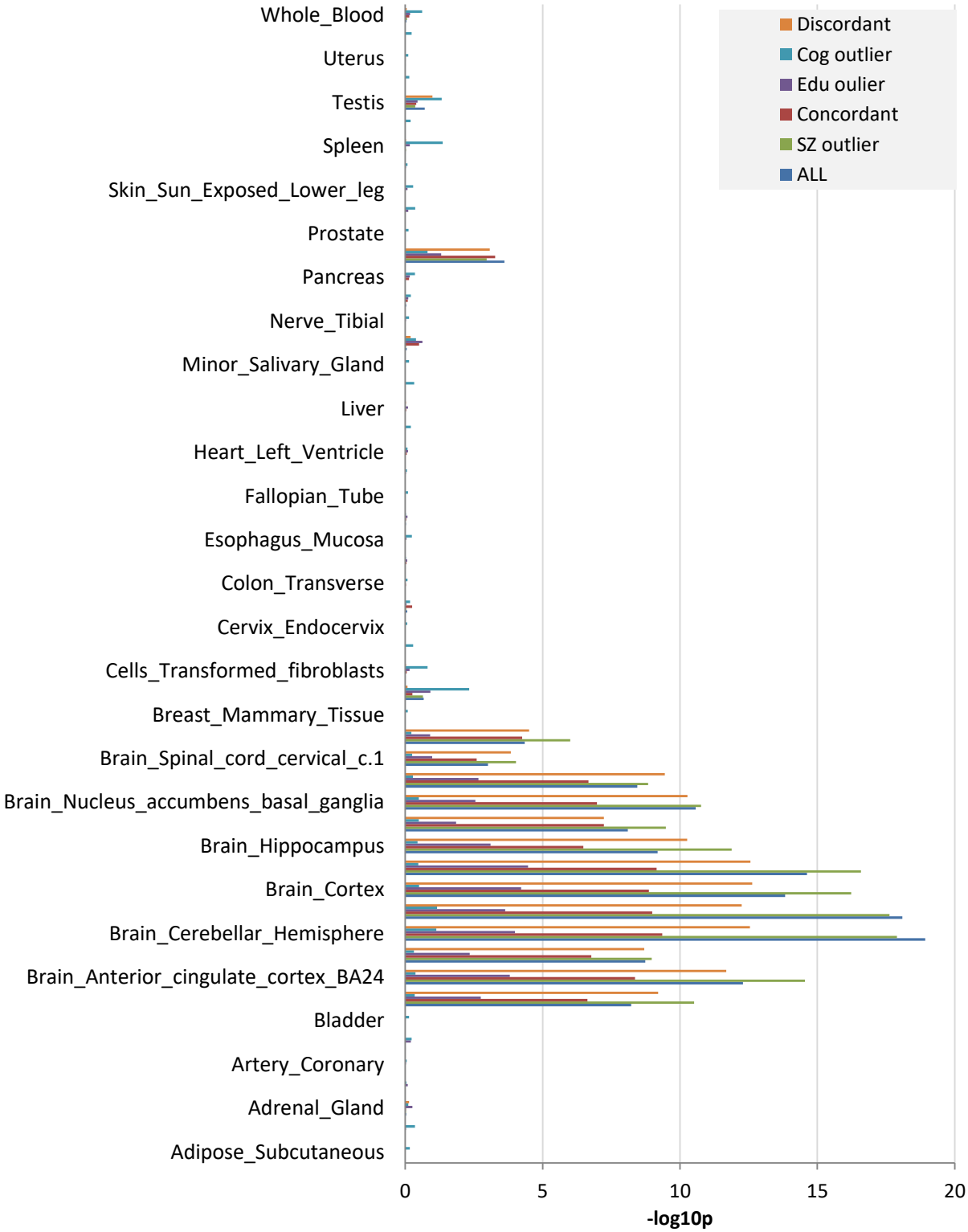

Figure S3: MAGMA Gene property tissue expression results for ASSET Subsets
Note: Colored bars indicate –log10P values for MAGMA gene property analysis. Both concordant and discordant ASSET subsets include genes highly expressed across brain tissue.

Supplemental Acknowledgments

This work has been supported by grants from the National Institutes of Health (R01 MH117646 to TL; R01 MH079800 and P50 MH080173 to AKM; R01 MH080912 to DCG; K23 MH077807 to KEB; K01 MH085812 to MCK). Data collection for the TOP cohort was supported by the Research Council of Norway, South-East Norway Health Authority, and KG Jebsen Foundation. The NCNG study was supported by Research Council of Norway Grants 154313/V50 and 177458/V50. The NCNG GWAS was financed by grants from the Bergen Research Foundation, the University of Bergen, the Research Council of Norway (FUGE, Psykisk Helse), Helse Vest RHF and Dr Einar Martens Fund. The Helsinki Birth Cohort Study has been supported by grants from the Academy of Finland, the Finnish Diabetes Research Society, Folkhälsan Research Foundation, Novo Nordisk Foundation, Finska Läkaresällskapet, Signe and Ane Gyllenberg Foundation, University of Helsinki, Ministry of Education, Ahokas Foundation, Emil Aaltonen Foundation. For the LBC1936 cohort, phenotype collection was supported by The Disconnected Mind project. Genotyping was funded by the UK Biotechnology and Biological Sciences Research Council (BBSRC grant No. BB/F019394/1). The work was undertaken by The University of Edinburgh Centre for Cognitive Ageing and Cognitive Epidemiology, part of the cross council Lifelong Health and Wellbeing Initiative, which is funded by the Medical Research Council and the Biotechnology and Biological Sciences Research Council (MR/K026992/1). WDH is supported by a grant from Age UK (Disconnected Mind Project). The CAMH work was supported by the CAMH Foundation and the Canadian Institutes of Health Research. The Duke Cognition Cohort (DCC) acknowledges K. Linney, J.M. McEvoy, P. Hunt, V. Dixon, T. Pennuto, K. Cornett, D. Swilling, L. Phillips, M. Silver, J. Covington, N. Walley, J. Dawson, H. Onabanjo, P. Nicoletti, A. Wagoner, J. Elmore, L. Bevan, J. Hunkin and R. Wilson for recruitment and testing of subjects. DCC also acknowledges the Ellison Medical Foundation New Scholar award AG-NS-0441-08 for partial funding of this study as well as the National Institute of Mental Health of the National Institutes of Health under award number K01MH098126. The UCLA Consortium for Neuropsychiatric Phenomics (CNP) study acknowledges the following sources of funding from the NIH: Grants UL1DE019580 and PL1MH083271 (RMB), RL1MH083269 (TDC), RL1DA024853 (EL) and PL1NS062410. The ASPIS study was supported by National Institute of Mental Health research grants R01MH085018 and R01MH092515 to Dr. Dimitrios Avramopoulos. Support for the Duke Neurogenetics Study was provided the National Institutes of Health (R01 DA033369 and R01 AG049789 to ARH) and by a National Science Foundation Graduate Research Fellowship to MAS. Recruitment, genotyping and analysis of the TCD healthy control samples were supported by Science Foundation Ireland (grants 12/IP/1670, 12/IP/1359 and 08/IN.1/B1916).

Data access for several cohorts used in this study was provided by the National Center for Biotechnology Information (NCBI) database of Genotypes and Phenotypes (dbGaP). dbGaP accession numbers for these cohorts were:

Cardiovascular Health Study (CHS): phs000287.v4.p1, phs000377.v5.p1, and phs000226.v3.p1

Framingham Heart Study (FHS): phs000007.v23.p8 and phs000342.v11.p8

Multi-Site Collaborative Study for Genotype-Phenotype Associations in Alzheimer's Disease (GENADA): phs000219.v1.p1

Long Life Family Study (LLFS): phs000397.v1.p1

Genetics of Late Onset Alzheimer's Disease Study (LOAD): phs000168.v1.p1

Minnesota Center for Twin and Family Research (MCTFR): phs000620.v1.p1

Philadelphia Neurodevelopmental Cohort (PNC): phs000607.v1.p1

The acknowledgment statements for these cohorts are found below:

Framingham Heart Study: The Framingham Heart Study is conducted and supported by the National Heart, Lung, and Blood Institute (NHLBI) in collaboration with Boston University (Contract No. N01-HC-25195 and HHSN2682015000001I). This manuscript was not prepared in collaboration with investigators of the Framingham Heart Study and does not necessarily reflect the opinions or views of the Framingham Heart Study, Boston University, or NHLBI. Funding for SHARe Affymetrix genotyping was provided by NHLBI Contract N02-HL-64278. SHARe Illumina genotyping was provided under an agreement between Illumina and Boston University.

Cardiovascular Health Study: This research was supported by contracts HHSN268201200036C, HHSN268200800007C, N01-HC-85079, N01-HC-85080, N01-HC-85081, N01-HC-85082, N01-HC-85083, N01-HC-85084, N01-HC-85085, N01-HC-85086, N01-HC-35129, N01 HC-15103, N01 HC-55222, N01-HC-75150, N01-HC-45133, and N01-HC-85239; grant numbers U01 HL080295 and U01 HL130014 from the National Heart, Lung, and Blood Institute, and R01 AG-023629 from the National Institute on Aging, with additional contribution from the National Institute of Neurological Disorders and Stroke. A full list of principal CHS investigators and institutions can be found at https://chs-nhlbi.org/pi. This manuscript was not prepared in collaboration with CHS investigators and does not necessarily reflect the opinions or views of CHS, or the NHLBI. Support for the genotyping through the CARe Study was provided by NHLBI Contract N01-HC-65226. Support for the Cardiovascular Health Study Whole Genome Study was provided by NHLBI grant HL087652. Additional support for infrastructure was provided by HL105756 and additional genotyping among the African-American cohort was supported in part by HL085251, DNA handling and genotyping at Cedars-Sinai Medical Center was supported in part by National Center for Research Resources grant UL1RR033176, now at the National Center for Advancing Translational Technologies CTSI grant UL1TR000124; in addition to the National Institute of Diabetes and Digestive and Kidney Diseases grant DK063491 to the Southern California Diabetes Endocrinology Research Center.

Multi-Site Collaborative Study for Genotype-Phenotype Associations in Alzheimer's Disease: The genotypic and associated phenotypic data used in the study were provided by the GlaxoSmithKline, R&D Limited. Details on data acquisition have been published previously in: Li H, Wetten S, Li L, St Jean PL, Upmanyu R, Surh L, Hosford D, Barnes MR, Briley JD, Borrie M, Coletta N, Delisle R, Dhalla D, Ehm MG, Feldman HH, Fornazzari L, Gauthier S, Goodgame N, Guzman D, Hammond S, Hollingworth P, Hsiung GY, Johnson J, Kelly DD, Keren R, Kertesz A, King KS, Lovestone S, Loy-English I, Matthews PM, Owen MJ, Plumpton M, Pryse-Phillips W, Prinjha RK, Richardson JC, Saunders A, Slater AJ, St George-Hyslop PH, Stinnett SW, Swartz JE, Taylor RL, Wherrett J, Williams J, Yarnall DP, Gibson RA, Irizarry MC, Middleton LT, Roses AD. Candidate single-nucleotide polymorphisms from a genomewide association study of Alzheimer disease. Arch Neurol., Jan;65(1):45-53, 2008 (PMID: 17998437). Filippini N, Rao A, Wetten S, Gibson RA, Borrie M, Guzman D, Kertesz A, Loy-English I, Williams J, Nichols T, Whitcher B, Matthews PM. Anatomically-distinct genetic associations of APOE epsilon4 allele load with regional cortical atrophy in Alzheimer's disease. Neuroimage, Feb 1;44(3):724-8, 2009. (PMID: 19013250).

Genetics of Late Onset Alzheimer's Disease Study: Funding support for the "Genetic Consortium for Late Onset Alzheimer's Disease" was provided through the Division of Neuroscience, NIA. The Genetic Consortium for Late Onset Alzheimer's Disease includes a genome-wide association study

funded as part of the Division of Neuroscience, NIA. Assistance with phenotype harmonization and genotype cleaning, as well as with general study coordination, was provided by Genetic Consortium for Late Onset Alzheimer's Disease. A list of contributing investigators is available at https://www.ncbi.nlm.nih.gov/projects/gap/cgi-bin/study.cgi?study_id=phs000168.v1.p1

Long Life Family Study: Funding support for the Long Life Family Study was provided by the Division of Geriatrics and Clinical Gerontology, National Institute on Aging. The Long Life Family Study includes GWAS analyses for factors that contribute to long and healthy life. Assistance with phenotype harmonization and genotype cleaning as well as with general study coordination, was provided by the Division of Geriatrics and Clinical Gerontology, National Institute on Aging. Support for the collection of datasets and samples were provided by Multicenter Cooperative Agreement support by the Division of Geriatrics and Clinical Gerontology, National Institute on Aging (UO1AG023746; UO1023755; UO1023749; UO1023744; UO1023712). Funding support for the genotyping which was performed at the Johns Hopkins University Center for Inherited Disease Research was provided by the National Institute on Aging, National Institutes of Health.

Minnesota Center for Twin and Family Research: This project was led by William G. Iacono, PhD. And Matthew K. McGue, PhD (Co-Principal Investigators) at the University of Minnesota, Minneapolis, MN, USA. Co-investigators from the same institution included: Irene J. Elkins, Margaret A. Keyes, Lisa N. Legrand, Stephen M. Malone, William S. Oetting, Michael B. Miller, and Saonli Basu. Funding support for this project was provided through NIDA (U01 DA 024417). Other support for sample ascertainment and data collection came from several grants: R37 DA 05147, R01 AA 09367, R01 AA 11886, R01 DA 13240, R01 MH 66140.

Philadelphia Neurodevelopmental Cohort: Support for the collection of the data sets was provided by grant RC2MH089983 awarded to Raquel Gur, MD, and RC2MH089924 awarded to Hakon Hakonarson, MD, PhD. All subjects were recruited through the Center for Applied Genomics at The Children's Hospital in Philadelphia.