**SUPPLEMENTARY MATERIAL**

**Genetic susceptibility for schizophrenia after adjustment by genetic susceptibility for smoking: implications in identification of risk genes and genetic correlation with related traits**

Laila Al-Soufi, Javier Costas

***SUPPLEMENTARY METHODS***

**Transformation of GSMR estimates to liability scale**

In order to get interpretable estimates of for binary traits, GSMR estimates were transformed to the liability scale, as described in Byrne et al. (2020). Three estimates are needed for these transformations, population prevalence of exposure trait *x* (*Kx*), population prevalence of outcome trait *y* (*Ky*) and the height of the normal distribution at the truncation point on the liability scale corresponding to risk *K* (). Used population prevalences for each binary trait are found in Table 1. If exposure trait is quantitative and outcome is binary, as in the case of CigDay and AgeSmk on SCZ, GSMR estimates and on the liability scale is calculated as . If both traits are binary, as in the case of SmkInit and SmkCes on SCZ, GSMR estimates and on the liability scale is calculated as . So, these estimates can be interpretated as the increase in phenotypic liability to trait *y* per 1 SD increase in trait *x* (quantitative) or per 1 SD increase in phenotypic liability to trait *x* (binary). If only the exposure binary trait *x* is transformed to the liability scale, is estimated. This can be interpreted as the log of the odds ratio of trait *y* per 1 SD increase in trait *x* and can be directly used to get the odds ratio.

**Colocalization analysis**

Colocalization analysis was performed following Pickrell et al. (2016). This colocalization methodology is robust to sample overlap between GWAS. The posterior probability (PP) for five alternative hypotheses was computed: H0, no association with either trait; H1, association with trait 1 but no with trait 2; H2, association with trait 2 but no with trait 1; H3, two independent association signals, one for each trait; and H4, a common association signal for both traits. Power of colocalization analysis was assessed by PP3+PP4. Those analyses with PP3+PP4>0.8 and with a posterior probability for H3 or H4 more than three times the posterior probability for the other hypothesis were considered as highly supportive for that hypothesis, as previously done (Al-Soufi & Costas, 2021; Bhalala, Nath, Inouye, UK Brain Expression Consortium, & Sibley, 2018). First, fgwas v.0.3.6 was run using default values to estimate the posterior probability of association with the traits at each approximately independent genomic block of European populations. Then, all SNPs at those blocks with a posterior probability of association lower than 0.2 were selected for each trait. Next, Pearson’s correlation between the Z-scores of these SNPs for the two traits under study was estimated. Finally, gwas-pw v.0.21 was used to estimate the posterior probability for each of the hypotheses tested, using this correlation value for correction, for those genomic blocks including the loci under study.

***REFERENCES***

Al-Soufi, L., & Costas, J. (2021). Colocalization of association signals at nicotinic acetylcholine receptor genes between schizophrenia and smoking traits. *Drug and Alcohol Dependence*, *220*, 108517. https://doi.org/10.1016/j.drugalcdep.2021.108517

Arnold, P. D., Askland, K. D., Barlassina, C., Bellodi, L., Bienvenu, O. J., Black, D., … Zai, G. (2018). Revealing the complex genetic architecture of obsessive-compulsive disorder using meta-analysis. *Molecular Psychiatry*, *23*(5), 1181–1188. https://doi.org/10.1038/mp.2017.154

Bhalala, O. G., Nath, A. P., Inouye, M., UK Brain Expression Consortium, & Sibley, C. R. (2018). Identification of expression quantitative trait loci associated with schizophrenia and affective disorders in normal brain tissue. *PLoS Genetics*, *14*(8), e1007607. https://doi.org/10.1371/journal.pgen.1007607

Byrne, E. M., Zhu, Z., Qi, T., Skene, N. G., Bryois, J., Pardinas, A. F., … Wray, N. R. (2021). Conditional GWAS analysis to identify disorder-specific SNPs for psychiatric disorders. *Molecular Psychiatry*, *26*(6), 2070–2081. https://doi.org/10.1038/s41380-020-0705-9

Canela-Xandri, O., Rawlik, K., & Tenesa, A. (2018). An atlas of genetic associations in UK Biobank. *Nature Genetics*, *50*(11), 1593–1599. https://doi.org/10.1038/s41588-018-0248-z

Day, F. R., Ong, K. K., & Perry, J. R. B. (2018). Elucidating the genetic basis of social interaction and isolation. *Nature Communications*, *9*(1), 1–6. https://doi.org/10.1038/s41467-018-04930-1

Demontis, D., Walters, R. K., Martin, J., Mattheisen, M., Als, T. D., Agerbo, E., … Neale, B. M. (2019). Discovery of the first genome-wide significant risk loci for attention deficit/hyperactivity disorder. *Nature Genetics*, *51*(1), 63–75. https://doi.org/10.1038/s41588-018-0269-7

Grove, J., Ripke, S., Als, T. D., Mattheisen, M., Walters, R. K., Won, H., … Børglum, A. D. (2019). Identification of common genetic risk variants for autism spectrum disorder. *Nature Genetics*, *51*(3), 431–444. https://doi.org/10.1038/s41588-019-0344-8

Hill, W. D., Hagenaars, S. P., Marioni, R. E., Harris, S. E., Liewald, D. C. M., Davies, G., … Deary, I. J. (2016). Molecular Genetic Contributions to Social Deprivation and Household Income in UK Biobank. *Current Biology*, *26*(22), 3083–3089. https://doi.org/10.1016/j.cub.2016.09.035

Howard, D. M., Adams, M. J., Clarke, T.-K., Hafferty, J. D., Gibson, J., Shirali, M., … McIntosh, A. M. (2019). Genome-wide meta-analysis of depression identifies 102 independent variants and highlights the importance of the prefrontal brain regions. *Nature Neuroscience*, *22*(3), 343–352. https://doi.org/10.1038/s41593-018-0326-7

Jansen, P. R., Watanabe, K., Stringer, S., Skene, N., Bryois, J., Hammerschlag, A. R., … Posthuma, D. (2019). Genome-wide analysis of insomnia in 1,331,010 individuals identifies new risk loci and functional pathways. *Nature Genetics*, *51*(3), 394–403. https://doi.org/10.1038/s41588-018-0333-3

Johnson, E. C., Demontis, D., Thorgeirsson, T. E., Walters, R. K., Polimanti, R., Hatoum, A. S., … Børglum, A. D. (2020). A large-scale genome-wide association study meta-analysis of cannabis use disorder. *The Lancet Psychiatry*, *7*(12), 1032–1045. https://doi.org/10.1016/S2215-0366(20)30339-4

Karlsson Linnér, R., Biroli, P., Kong, E., Meddens, S. F. W., Wedow, R., Fontana, M. A., … Wagner, G. G. (2019). Genome-wide association analyses of risk tolerance and risky behaviors in over 1 million individuals identify hundreds of loci and shared genetic influences. *Nature Genetics*, *51*(2), 245–257. https://doi.org/10.1038/s41588-018-0309-3

Lee, J. J., Wedow, R., Okbay, A., Kong, E., Maghzian, O., Zacher, M., … Cesarini, D. (2018). Gene discovery and polygenic prediction from a genome-wide association study of educational attainment in 1.1 million individuals. *Nature Genetics*, *50*(8), 1112–1121. https://doi.org/10.1038/s41588-018-0147-3

Liu, M., Jiang, Y., Wedow, R., Li, Y., Brazel, D. M., Chen, F., … Vrieze, S. (2019). Association studies of up to 1.2 million individuals yield new insights into the genetic etiology of tobacco and alcohol use. *Nature Genetics*, *51*(2), 237–244. https://doi.org/10.1038/s41588-018-0307-5

Mills, M. C., Tropf, F. C., Brazel, D. M., van Zuydam, N., Vaez, A., Agbessi, M., … Day, F. R. (2021). Identification of 371 genetic variants for age at first sex and birth linked to externalising behaviour. *Nature Human Behaviour*, *5*(12), 1717–1730. https://doi.org/10.1038/s41562-021-01135-3

Mullins, N., Forstner, A. J., O’Connell, K. S., Coombes, B., Coleman, J. R. I., Qiao, Z., … Andreassen, O. A. (2021). Genome-wide association study of more than 40,000 bipolar disorder cases provides new insights into the underlying biology. *Nature Genetics*, *53*(6), 817–829. https://doi.org/10.1038/s41588-021-00857-4

Nagel, M., Jansen, P. R., Stringer, S., Watanabe, K., De Leeuw, C. A., Bryois, J., … Posthuma, D. (2018). Meta-analysis of genome-wide association studies for neuroticism in 449,484 individuals identifies novel genetic loci and pathways. *Nature Genetics*, *50*(7), 920–927. https://doi.org/10.1038/s41588-018-0151-7

Okbay, A., Baselmans, B. M. L., De Neve, J. E., Turley, P., Nivard, M. G., Fontana, M. A., … Cesarini, D. (2016). Genetic variants associated with subjective well-being, depressive symptoms, and neuroticism identified through genome-wide analyses. *Nature Genetics*, *48*(6), 624–633. https://doi.org/10.1038/ng.3552

Otowa, T., Hek, K., Lee, M., Byrne, E. M., Mirza, S. S., Nivard, M. G., … Hettema, J. M. (2016). Meta-analysis of genome-wide association studies of anxiety disorders. *Molecular Psychiatry*, *21*(10), 1391–1399. https://doi.org/10.1038/mp.2015.197

Pasman, J. A., Verweij, K. J. H., Gerring, Z., Stringer, S., Sanchez-Roige, S., Treur, J. L., … Vink, J. M. (2018). GWAS of lifetime cannabis use reveals new risk loci, genetic overlap with psychiatric traits, and a causal effect of schizophrenia liability. *Nature Neuroscience*, *21*(9), 1161–1170. https://doi.org/10.1038/s41593-018-0206-1

Pickrell, J. K., Berisa, T., Liu, J. Z., Ségurel, L., Tung, J. Y., & Hinds, D. A. (2016). Detection and interpretation of shared genetic influences on 42 human traits. *Nature Genetics*, *48*(7), 709–717. https://doi.org/10.1038/ng.3570

Polimanti, R., Walters, R. K., Johnson, E. C., McClintick, J. N., Adkins, A. E., Adkins, D. E., … Gelernter, J. (2020). Leveraging genome-wide data to investigate differences between opioid use vs. opioid dependence in 41,176 individuals from the Psychiatric Genomics Consortium. *Molecular Psychiatry*, *25*(8), 1673–1687. https://doi.org/10.1038/s41380-020-0677-9

Sanchez-Roige, S., Palmer, A. A., Fontanillas, P., Elson, S. L., Adams, M. J., Howard, D. M., … Wilson, C. H. (2019). Genome-wide association study meta-analysis of the alcohol use disorders identification test (AUDIT) in two population-based cohorts. *American Journal of Psychiatry*, *176*(2), 107–118. https://doi.org/10.1176/appi.ajp.2018.18040369

Savage, J. E., Jansen, P. R., Stringer, S., Watanabe, K., Bryois, J., De Leeuw, C. A., … Posthuma, D. (2018). Genome-wide association meta-analysis in 269,867 individuals identifies new genetic and functional links to intelligence. *Nature Genetics*, *50*(7), 912–919. https://doi.org/10.1038/s41588-018-0152-6

van den Berg, S. M., de Moor, M. H. M., Verweij, K. J. H., Krueger, R. F., Luciano, M., Arias Vasquez, A., … Boomsma, D. I. (2016). Meta-analysis of Genome-Wide Association Studies for Extraversion: Findings from the Genetics of Personality Consortium. *Behavior Genetics*, *46*(2), 170–182. https://doi.org/10.1007/s10519-015-9735-5

Walters, R. K., Polimanti, R., Johnson, E. C., McClintick, J. N., Adams, M. J., Adkins, A. E., … Agrawal, A. (2018). Transancestral GWAS of alcohol dependence reveals common genetic underpinnings with psychiatric disorders. *Nature Neuroscience*, *21*(12), 1656–1669. https://doi.org/10.1038/s41593-018-0275-1

Watson, H. J., Yilmaz, Z., Thornton, L. M., Hübel, C., Coleman, J. R. I., Gaspar, H. A., … Bulik, C. M. (2019). Genome-wide association study identifies eight risk loci and implicates metabo-psychiatric origins for anorexia nervosa. *Nature Genetics*, *51*(8), 1207–1214. https://doi.org/10.1038/s41588-019-0439-2

Wootton, R. E., Richmond, R. C., Stuijfzand, B. G., Lawn, R. B., Sallis, H. M., Taylor, G. M. J., … Munafò, M. R. (2020). Evidence for causal effects of lifetime smoking on risk for depression and schizophrenia: a Mendelian randomisation study. *Psychological Medicine*, *50*(14), 2435–2443. https://doi.org/10.1017/S0033291719002678

**Supplementary Table 1. GWAS summary statistics used in genetic correlation analysis.**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Trait** | **Description** | **Sourcea** | **Cohortsa,b** | **Sample size (N cases + N controls)** |
| **LifetimeSmk** | Lifetime smoking index which captures smoking initiation, duration, heaviness and time since cessation | University of Bristol (Wootton et al., 2020) | UKBB | 462,690 |
| **Cannabis use** | Lifetime cannabis use | ICC (Pasman et al., 2018) | ICC & UKBB | 162,082 (43,380 + 118,702) |
| **DrnkWk** | Average number of drinks per week | GSCAN (Liu et al., 2019) | GSCAN (including UKBB) | 537,349 |
| **AUDIT** | Alcohol Use Disorder Identification Test Total Score | PGC (Sanchez-Roige et al., 2019) | UKBB | 121,604 |
| **Opioid use** | Opioid use | PGC (Polimanti et al., 2020) | PGC | 28,313 (2,876 + 25,437) |
| **TUD** | Mental and behavioral disorders due to use of tobacco (coded as in ICD10 F17) | GeneATLAS (Canela-Xandri, Rawlik, & Tenesa, 2018) | UKBB | 452,264 (12,310 + 439,954) |
| **CUD** | Cannabis use disorder | PGC (Johnson et al., 2020) | PGC, iPSYCH & deCODE | 357,806 (14,080 + 343,726) |
| **Alcohol dep** | Alcohol dependence | PGC (Walters et al., 2018) | PGC | 46,568 (11,569 + 34,999) |
| **Opioid dep** | Opioid dependence | PGC (Polimanti et al., 2020) | PGC | 28,709 (3,272 + 25,437) |
| **ADHD** | Attention deficit-hyperactivity disorder | PGC (Demontis et al., 2019) | PGC & iPSYCH | 53,293 (19,099 + 34,194) |
| **MDD** | Major depressive disorder | PGC (Howard et al., 2019) | PGC & UKBB | 500,199 (170,756 + 329,443) |
| **BIP** | Bipolar disorder | PGC (Mullins et al., 2021) | PGC, iPSYCH, deCODE, Estonian Biobank, HUNT and UKBB | 413,466 (41,917 + 371,549) |
| **ASD** | Autism spectrum disorder | PGC (Grove et al., 2019) | PGC & iPSYCH | 46,351 (18,382 + 27,969) |
| **AN** | Anorexia nervosa | PGC (Watson et al., 2019) | PGC, ANGI & UKBB | 72,517 (16,992 + 55,525) |
| **OCD** | Obsessive-compulsive disorder | PGC (Arnold et al., 2018) | IOCDF-GC & OCGAS | 9,725 (2,688 + 7,037) |
| **AD** | Anxiety disorders | PGC (Otowa et al., 2016) | ANGST | 17,310c |
| **Risk PC1** | First PC of the four risky behaviors in the UKBB: automobile speeding propensity, drinks per week, ever smoker and number of sexual partners | SSGAC (Karlsson Linnér et al., 2019) | UKBB | 315,894 |
| **Risk tolerance** | General risk tolerance | SSGAC (Karlsson Linnér et al., 2019) | UKBB & 10 other cohorts | 466,571 |
| **Sexual partners** | Number of sexual partners | SSGAC (Karlsson Linnér et al., 2019) | UKBB | 370,711 |
| **Automobile speeding** | Automobile speeding propensity | SSGAC (Karlsson Linnér et al., 2019) | UKBB | 404,291 |
| **AFB** | Age when having the first child | GWAS Catalog (Mills et al., 2021) | UKBB & 35 other cohorts | 542,901 |
| **AFS** | Age at first sexual intercourse | GWAS Catalog (Mills et al., 2021) | UKBB | 397,338 |
| **Extraversion** | Response to different extraversion items | GPC (van den Berg et al., 2016) | GPC | 63,030 |
| **SWB** | Subjective well being | SSGAC (Okbay et al., 2016) | UKBB & 57 other cohorts | 204,966 |
| **Neuroticism** | Response to different neuroticism items | CTG (Nagel et al., 2018) | UKBB & GPC | 390,278 |
| **Depressed** | Neuroticism subcluster: response to depressed affect items | CTG (Nagel et al., 2018) | UKBB | 357,957 |
| **Worry** | Neuroticism subcluster: response to worry items | CTG (Nagel et al., 2018) | UKBB | 348,219 |
| **Intelligence** | Intelligence assessed by various neurocognitive tests | CTG (Savage et al., 2018) | UKBB & 13 other cohorts | 269,867 |
| **Cognitive** | Cognitive performance assessed by neuropsychological tests | SSGAC (Lee et al., 2018) | COGENT & UKBB | 257,828 |
| **Insomnia** | Insomnia disorder | CTG (Jansen et al., 2019) | UKBB | 386,533 (109,402 + 386,533) |
| **Chronotype** | Morning/evening person | Neale Labd | UKBB | 322,488 |
| **Sleep** | Hours of sleep in every 24 hours | Neale Labd | UKBB | 359,020 |
| **Narcolepsy** | Dozing or falling asleep during the daytime | Neale Labd | UKBB | 359,752 |
| **Highly irritable** | Ever feeling highly irritable/argumentative for 2 days | Neale Labd | UKBB | 117,359 (20,930 + 96,429) |
| **Irritability** | Being an irritable person | Neale Labd | UKBB | 345,231 (96,862 + 248,369) |
| **Extreme irritability** | Ever had a period of time feeling extreme irritability | Neale Labd | UKBB | 114,422 (29,747 + 84,675) |
| **EA** | Years of educational attainment | SSGAC (Lee et al., 2018) | UKBB & 69 other cohorts | 766,345 |
| **Loneliness** | Often feeling lonely | Neale Labd | UKBB | 355,583 (63,508 + 292075) |
| **Confide** | Frequency of being able to confide in someone close | Neale Labd | UKBB | 350,618 |
| **Social deprivation** | Townsend Social Deprivation Index | GWAS Catalog (Hill et al., 2016) | UKBB | 112,151 |
| **Loneliness and isolation** | Combined multi-trait GWAS of loneliness and social isolation | University of Cambridge (Day, Ong, & Perry, 2018) | UKBB | 487,647 |

a: Abbreviations are as follows: UKBB: UK Biobank; ICC: International Cannabis Consortium, GSCAN: GWAS and Sequencing Consortium of Alcohol and Nicotine use; PGC: Psychiatric Genomics Consortium; iPSYCH: Lundbeck Foundation Initiative for Integrative Psychiatric Research; HUNT: Trøndelag Health Study; ANGI: Anorexia Nervosa Genetics Initiative; IOCDF-GC: International Obsessive Compulsive Disorder Foundation Genetics Collaborative, OCGAS: OCD Collaborative Genetics Association Studies, ANGST: Anxiety NeuroGenetics Study, SSGAC: Social Science Genetic Association Consortium; GPC: Genetics of Personality Consortium; CTG: Complex Traits Genetics Lab; COGENT: Cognitive Genomics Consortium.

b: Some of the original GWAS included 23andme samples, not used in the present analysis.

c: N of cases and N of controls are not indicated in the original article.

d: http://www.nealelab.is/uk-biobank/

**Supplementary Tables 2 - 7 supplied as Excel file.**

**Supplementary Figure 1. Overview of the study design.**



**Supplementary Figure 2. MAGMA enrichment analysis with original SCZ GWAS and SCZcond. The solid vertical line represents significance level at Bonferroni's correction for the number of gene expression profiles tested. A: tissues enrichment, B: brain ages enrichment.**



**Supplementary Figure 3. Significantly enriched GO terms referring to A: biological processes, B: cellular components and C: molecular functions for SCZ GWAS and/or SCZcond. The solid vertical line represents significance level at Bonferroni's correction for the number of gene-sets tested.**

