ESCOLA DE DOUTORAMENTO
INTERNACIONAL DA USC

Laila
Al-Soufi Novo

Tese de doutoramento

Unha aproximación xenómica á susceptibilidade do tabaquismo e outros trazos de consumo de substancias en pacientes con esquizofrenia e poboación xeral

Santiago de Compostela, 2024

Programa de Doutoramento en Medicina Molecular

ESCOLA DE DOUTORAMENTO
INTERNACIONAL DA USC

TESE DE DOUTORAMENTO

UNHA APROXIMACIÓN XENÓMICA Á SUSCEPTIBILIDADE DO TABAQUISMO E OUTROS TRAZOS DE CONSUMO DE SUBSTANCIAS EN PACIENTES CON ESQUIZOFRENIA E POBOACIÓN XERAL

Autora

Laila Al-Soufi Novo

Directores:

Javier Costas Costas

Laura Sánchez Piñón

Titora: Laura Sánchez Piñón

PROGRAMA DE DOUTORAMENTO EN MEDICINA MOLECULAR

SANTIAGO DE COMPOSTELA, 2024

DECLARACIÓN DE CONFLITOS DE INTERESE

Eu, Laila Al-Soufi Novo, con DNI 34280596Q, presento a miña Tese de Doutoramento, titulada "Unha aproximación xenómica á susceptibilidade do tabaquismo e outros trazos de consumo de substancias en pacientes con esquizofrenia e poboación xeral", seguindo o procedemento adecuado ao Regulamento, e declaro non ter ningún conflito de interese en relación coa tese de doutoramento. Todas as figuras incluídas na tese son de elaboración propia.

En Santiago de Compostela, a 16 de xuño de 2024.

Fdo. Laila Al-Soufi Novo

USC
UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

AGRADECEMENTOS

Ao meu director de tese, Xabi Costas, pola túa inmensa dedicación a esta tese. Grazas por terme ensinado tanto ao longo destes xa seis anos. E, sobre todo, grazas por transmitirme, co teu exemplo, o importante que é a ética e o rigor na investigación.

Á miña directora de tese, Laura Sánchez, polo teu enorme apoio en cada paso desta tese. Grazas por orientarme buscando sempre o mellor para min. Grazas por acollerme tan ben no Departamento de Xenética de Lugo, onde sempre me sentín como na casa.

Aos meus compañeiros do grupo de Xenética Psiquiátrica, por ser tan bos compañeiros. Grazas pola vosa xenerosidade e por buscar sempre un ambiente de apoio e axuda mutua. A Pablo, por facer das horas no traballo un desfrute e por alegrarte dos meus logros como se fosen teus. Estou segura de que as nosas conversas infinitas seguirán por moitos anos máis. A Fer, por terme dado a oportunidade de coñecer en persoa o que só lera en artigos. Grazas pola túa orientación cando as cousas non eran fáciles. A Fer, por ter sempre unha historia que contar e por facer a fase de escritura da tese algo máis levadía ao ser en compañía. Ás persoas que xa non están no grupo mais que seguen estando sempre que o preciso. Só teño boas palabras do meu paso polo grupo de Xenética Psiquiátrica.

Ás nosas compañeiras do Lab 15, Laura, Eva, Miguel, Raquel e Diego. Grazas por facer do laboratorio un sitio onde me sinto moi querida. Ás compañeiras do laboratorio 20, especialmente a Laura, Soraya, Nerea, Jen e María, por acollerme tan ben dende o principio, aínda que fose a única que non sabía empregar unha pipeta. Grazas a todas as persoas que coñecín nestes anos e que me fixeron pasar tan bos momentos. As amizades que fixen durante a tese é do mellor que me levo desta etapa.

Ás pacientes, que tan xenerosamente empregaron parte do seu valioso tempo en colaborar neste proxecto.

A cada unha das compañeiras da Asemblea de Investigadoras de Compostela, por ser sempre un apoio e por facer da USC un lugar mellor. Da AIC lévome unha lección de por vida: a importancia e a ledicia de traballar xuntas.

USC
UNIVERSIDADE
DE SANTIAGO
DE COMPOSTELA

A todas as miñas amigas, grazas por acompañarme con tanto cariño en cada paso que dou. Grazas ás marabillosas amigas que fixen en Bioloxía e que seguen comigo anos despois. A Carme, por comezar buscando piso xuntas e acabar sendo indispensable nesta etapa. Grazas a todas as amigas de Compos, por facer a miña vida moito máis feliz. Non me podo imaxinar estes anos sen a vosa compañía. Aprendo moito de cada unha de vós cada día.

Á miña Cuadrilla lucense: Anita, Pauli, Celti, Gómez, Basan, Samu, Soto, Cori, Aldao, Marta, Mariita, Samu, Mimi, Duarte e Laura, e todas aquelas persoas que forman parte de nós. Síntome moi afortunada de tervos ao meu lado e de vervos crecer. É incrible como unhas nenas e adolescentes que se fixeron amigas se sigan entendendo e querendo tan ben anos despois. Facedes de Lugo o sitio ao que sempre quero volver e de Compos a miña casa.

Á miña familia. Grazas aos meus pais, polos seus valores e o seu cariño incondicional. Grazas ás miñas irmás, a nosa relación é do que máis orgullosa estou na vida, grazas por sosterme sempre tan ben. Grazas a Dani, por ser o meu maior apoio e a calma que preciso. Grazas por ser o meu sitio seguro.

USC
UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

ÍNDICE

ABSTRACT **7**

RESUMO **11**

PUBLICACIÓNS INCLUÍDAS NA TESE **16**

ABREVIATURAS **17**

1. INTRODUCIÓN **19**

1.1. Trazos de consumo de substancias: comportamentos de consumo de substancias e trastornos polo uso de substancias 19

1.1.1. Epidemioloxía dos trazos de consumo de substancias *19*

1.1.2. Definición e diagnóstico dos trazos de consumo de substancias *21*

1.1.3. O ciclo da adicción *24*

1.1.4. Base xenética dos trazos de consumo de substancias *24*

1.1.5. A vía externalizante e a vía internalizante cara o consumo de substancias *26*

1.1.5.1. Evidencia epidemiolóxica da vía externalizante e a vía internalizante 26

1.1.5.2. Diferenzas no efecto da vía externalizante e a vía internalizante sobre o consumo de substancias 26

1.2. Asociación entre trazos de consumo de substancias e trastornos psiquiátricos 27

1.2.1. Asociación entre trastornos polo uso de substancias e outros trastornos psiquiátricos *27*

1.2.2. Asociación entre consumo de cannabis e psicose *28*

1.2.3. Asociación entre tabaquismo e esquizofrenia *28*

1.2.3.1. Evidencia epidemiolóxica da asociación entre tabaquismo e esquizofrenia 28

1.2.3.2. Consecuencias do tabaquismo en pacientes con esquizofrenia 30

1.2.3.3. Posibles explicacións á asociación entre tabaquismo e esquizofrenia 30

1.2.3.3.1. Hipótese da automedicación 31

1.2.3.3.2. Hipótese do tabaco como factor de risco a esquizofrenia 32

1.2.3.3.3. Combinación das hipóteses para explicar a asociación entre tabaquismo e esquizofrenia 34

1.3. Principais métodos xenómicos para o estudo de trazos complexos 35

1.3.1. Estudos de asociación do xenoma completo *35*

1.3.2. Estimas baseadas en LD score regression *38*

1.3.3. Puntuacións de risco polixénico *39*

1.3.4. Aleatorización mendeliana *41*

1.3.5. Modelos de ecuacións estruturais xenómicas *43*

1.3.6. Identificación de mecanismos biolóxicos subxacentes aos trazos complexos *44*

1.3.7. Da xenómica aos estudos funcionais: o peixe cebra como modelo animal *46*

1.4. Aplicación da metodoloxía xenómica ao estudo de trazos de consumo de substancias 47

1.4.1. Estudos de asociación do xenoma completo para trazos de consumo de substancias *47*

1.4.2. Susceptibilidade xenética compartida entre as distintas fases de consumo da mesma substancia *48*

1.4.3. Susceptibilidade xenética compartida entre trazos de consumo de distintas substancias *51*

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

1.4.4. Evidencia xenómica da vía externalizante e a vía internalizante cara o consumo de substancias ..52
1.5. Aplicación da metodoloxía xenómica ao estudo da asociación entre trazos de consumo de substancias e trastornos psiquiátricos ..52
1.5.1. Susceptibilidade xenética compartida entre trazos de consumo de substancias e trastornos psiquiátricos ..52
1.5.2. Abordaxe xenómica da asociación entre tabaquismo e esquizofrenia..55
1.5.2.1. Susceptibilidade xenética compartida entre tabaquismo e esquizofrenia..55
1.5.2.2. Relacións de causalidade entre tabaquismo e esquizofrenia ..58
2. HIPÓTESES E OBXECTIVOS..60
3. METODOLOXÍA..64
3.1. Estudos de asociación do xenoma completo empregados ..64
3.2. Metodoloxías xenómicas empregadas ..70
3.2.1. Obtención de módulos de xenes co-expresados ..71
3.2.2. Identificación de loci significativos e mapeado xénico..71
3.2.3. Análise de enriquecemento en conxuntos de xenes ..72
3.2.4. Estimación de puntuacións de risco polixénico ..72
3.2.4.1. Mostras diana..73
3.2.4.2. Mostras de descubrimento..74
3.2.4.3. Métodos de estimación das puntuacións de risco polixénico ..74
3.2.4.4. Análises estatísticas das puntuacións de risco polixénico ..75
3.2.5. Análise do patrón de solapamento xenético independente da correlación xenética ..76
3.2.6. Análise de colocalización..76
3.2.7. Análise de aleatorización mendeliana ..77
3.2.8. Análise condicional de estudos de asociación do xenoma completo ..77
3.2.9. Aplicacións de LD score regression..78
4. RESULTADOS..79
4.1. Estudo 1: Identification of gene co-expression modules from zebrafish brain data: applications in psychiatry illustrated through alcohol-related traits ..79
4.2. Estudo 2: Polygenic overlap of substance use behaviors and disorders with externalizing and internalizing problems independent of genetic correlations ..100
4.3. Estudo 3: Colocalization of association signals at nicotinic acetylcholine receptor genes between schizophrenia and smoking traits..115
4.4. Estudo 4: A polygenic approach to the association between smoking and schizophrenia..125
4.5. Estudo 5: Genetic susceptibility for schizophrenia after adjustment by genetic susceptibility for smoking: implications in identification of risk genes and genetic correlation with related traits ..142
5. DISCUSIÓN XERAL..161
5.1. Susceptibilidade aos trazos de consumo de substancias..164
5.1.1. Modelización da susceptibilidade xenética aos trazos de consumo de substancias ..164
5.1.2. Efecto da vía externalizante e a vía internalizante nos comportamentos de consumo de substancias e nos trastornos polo uso de substancias ..167

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

5.2. Análise do tabaquismo en pacientes con esquizofrenia 168
5.2.1. Susceptibilidade xenética a tabaquismo nas persoas con esquizofrenia 169
5.2.2. Susceptibilidade xenética compartida entre tabaquismo e esquizofrenia 170
5.2.3. Asociación dos xenes de subunidades de nAchR con tabaquismo e esquizofrenia 177
5.2.4. Hipótese proposta para explicar a asociación entre tabaquismo e esquizofrenia 179
5.2.5. Diferenzas na asociación entre tabaquismo e esquizofrenia segundo o trazo de consumo de tabaco 180
5.3. Breves consideracións clínicas 181
5.4. Limitacións xerais dos estudos 182

6. CONCLUSIONS 185

BIBLIOGRAFÍA 187

ANEXOS 247

Anexo I: Informes favorables dos comités de ética 247
Anexo II: Artigos publicados incluídos na tese de doutoramento 248
Anexo III: Permisos de utilización dos artigos publicados 252

USC
UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

ABSTRACT

Substance use represents a major global health problem. Within substance use traits, a distinction can be made between substance use behaviors that are not necessarily problematic or indicative of addiction, and problematic use that can lead to a substance use disorder (SUD). Substance use susceptibility is mainly explained by two pathways: an externalizing pathway, characterized by high levels of impulsivity, aggression, risk-taking, and sensation-seeking; and an internalizing pathway, which has been less studied, characterized by using substances to cope with negative affect such as stress, depression, or anxiety. Genetic susceptibility to these traits, which are highly polygenic, is primarily studied through genome-wide association studies (GWAS). These studies have indicated differences in genetic susceptibility among substance use traits, with a moderate genetic correlation between substance use behaviors and SUDs. GWAS also found a relatively high genetic overlap of substance use traits, both substance use behaviors and SUDs, with psychiatric disorders. Shared genetic susceptibility between psychiatric disorders and SUDs has been explained by a general genetic factor of psychopathology that increases the risk to any psychiatric disorder, as well as through correlated latent genetic factors that increase the risk to a subset of traits.

Among the different psychiatric disorders, there is a particularly strong association between schizophrenia and tobacco use. This higher prevalence of smoking in patients with schizophrenia significantly contributes to the higher mortality of this population group. GWAS have shown that the high comorbidity between smoking and schizophrenia is reflected in a significant genetic correlation between the two traits. Furthermore, GWAS have identified, in addition to the significant associations of nearly all genes coding for nicotinic acetylcholine receptor (nAchR) subunits with smoking, a significant association of some of these genes with schizophrenia. This association between smoking and schizophrenia was mainly explained by three non-exclusive hypotheses: 1) the self-medication hypothesis, which states that patients with schizophrenia may smoke to reduce symptoms of the disorder or side effects of antipsychotic medication; 2) the hypothesis of tobacco as a risk factor for developing schizophrenia; and 3) the hypothesis that the same genetic and non-genetic factors could independently increase the risk of both schizophrenia and smoking. Previous Mendelian randomization studies have provided evidence of a causal effect of smoking on schizophrenia.

USC UNIVERSITAT DE CANTABRIA

However, it is necessary to approach the causal relationship between smoking and schizophrenia from different perspectives.

The main objective of this thesis was to study the genetic factors influencing susceptibility to substance use in general population and, more specifically, to smoking in patients with schizophrenia. To this end, a battery of genomic tools based on GWAS results was used in the five studies presented in this thesis.

In Study 1, 34 gene modules with a correlated expression in the brain were identified from publicly available zebrafish gene expression data, using Weighted Gene Co-expression Network Analysis (WGCNA). The data for this analysis included gene expression in the adult zebrafish brain under different experimental conditions, including substance withdrawal. Characterization of these gene modules through enrichment analyses in gene sets of interest suggested that these co-expression modules may represent gene networks in the human brain, with some potentially involved in psychiatric traits. Particularly, certain gene modules with an expression pattern specifically associated with substance withdrawal in the zebrafish brain were identified. Moreover, a polygenic risk score (PRS) based on genetic variants of genes from these co-expression modules was significantly more associated with alcohol dependence in humans than a PRS based on genetic variants of other genes expressed in the brain. Therefore, specific gene sets that may be involved in alcohol dependence in humans were identified from zebrafish expression data, demonstrating the potential of zebrafish as an animal model in psychiatry.

In Study 2, bivariate causal mixture model (MiXeR) was used to characterize the pattern of genetic overlap of substance use traits with externalizing and internalizing traits. Results were consistent with an effect of both the externalizing and internalizing pathways on substance use traits, but with differences between substance use behaviors and SUDs. Substance use behaviors showed a high genetic overlap with internalizing traits, but with a particularly low genetic correlation of effect sizes of shared variants, indicating a divergent effect distribution of shared genetic susceptibility for both traits. In contrast, these traits showed a higher genetic correlation of shared variants with externalizing traits, also in the case of considerable genetic overlap, suggesting a stronger influence of the externalizing pathway on substance use behaviors than the internalizing pathway. On the other hand, SUDs, even in the case of high genetic overlap, showed a high genetic correlation of shared variants with both externalizing and internalizing traits, although more pronounced with the former. This result is consistent with an effect of both pathways on the development of SUDs.

In Study 3, different methods were used to assess whether the GWAS association signals for smoking and schizophrenia at those nAchR genes for which there was evidence of association with both traits are due to the same causal variant, i.e., if they colocalize. For the *CHRNA2* gene, there was evidence of colocalization between smoking and schizophrenia, as well as with cannabis use disorder (CUD). This result is consistent with a vertical pleiotropy where the association of this gene with schizophrenia is mediated by smoking acting as a risk factor. However, the other two analyzed nAchR loci, specifically, the *CHRNA4* gene and the *CHRNA5-CHRNA3-CHRNB4* gene cluster, seem to be associated with smoking and schizophrenia through independent signals.

In Study 4, a PRS for smoking initiation estimated in general population was found to explain a similar proportion of the variance of smoking status in patients with schizophrenia and in controls. Therefore, genetic factors that increase the risk of smoking in patients with schizophrenia seem to be the same as in the general population. To confirm this result, PRSs for traits genetically correlated with smoking initiation with enough power in this analysis were also examined. For some of these correlated traits, specifically, educational attainment and body mass index, PRSs were also associated with smoking status in both patients with schizophrenia and controls. However, PRSs for other traits genetically correlated with smoking initiation, specifically, attention-deficit/hyperactivity disorder (ADHD) and age when having the first child, both related to externalizing behavior, were only associated with smoking in patients with schizophrenia after correcting for multiple tests. These results indicate that genetic susceptibility to externalizing problems seems to explain smoking susceptibility in patients with schizophrenia to a greater extent than in the general population. Moreover, a PRS combining all the analyzed traits explained better the smoking status of both patients with schizophrenia and controls than the individual PRS for smoking initiation, indicating that genetic susceptibility to these other traits has an additional effect on smoking risk. In this same study, the possible explanation for the shared genetic susceptibility between smoking and schizophrenia was also analyzed. As expected according to a causal effect of tobacco on schizophrenia, smoking initiation PRS was no longer significantly associated with schizophrenia after adjustment by smoking status. The same pattern was found for the most significant genetic variant for schizophrenia at the nAchR gene cluster *CHRNA5-CHRNA3-CHRNB4*.

USC UNIVERSIDADE DE SANTIAGO

In Study 5, the largest schizophrenia GWAS was conditioned on the causal effect of smoking initiation using multi-trait-based conditional and joint analysis (mtCOJO). After discarding the genetic susceptibility to schizophrenia mediated by the causal pathway of

smoking, 19 new risk loci for schizophrenia were identified whose association might have been masked by their effect on smoking initiation in the opposite direction, and 42 loci were no longer significantly associated with schizophrenia, as their association might be partially mediated by the causal pathway of smoking. Additionally, this study assessed the impact of conditioning schizophrenia GWAS by smoking on the genetic correlation between schizophrenia and a broad range of traits of interest, including substance use traits, psychiatric disorders, behavioral traits, and neuropsychological traits. Among all analyzed traits, genetic correlation of schizophrenia significantly decreased with substance use behaviors, SUDs, ADHD, and behavioral traits related to risk-taking. Therefore, these results indicate that genetic correlation of schizophrenia with externalizing behavior seems to be mediated, at least partially, by smoking.

Results of the last three studies provided further evidence regarding the cause of the association between smoking and schizophrenia. On one hand, genetic factors that explain smoking in patients with schizophrenia seem to be the same as in the general population, contrary to the self-medication hypothesis. On the other hand, there is consistent evidence indicating that part of the high prevalence of smoking in patients with schizophrenia is due to a causal effect of smoking initiation on schizophrenia, although this effect could be indirect through cannabis use. In relation to this, and in line with the externalizing pathway effect on smoking initiation, the externalizing genetic factor may be associated with schizophrenia partially through the smoking causal pathway. Additionally, at nAchR genes, which are involved in nicotine pharmacogenetics, there could be different association signals with varying patterns of association with smoking and schizophrenia.

Therefore, the results of the five studies in this thesis provided more information on genetic susceptibility to substance use, especially to tobacco use in patients with schizophrenia. This information could have important clinical implications, allowing the development of more effective strategies to predict and treat substance use.

USC
UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

RESUMO

O consumo de substancias representa un grave problema de saúde global. Dentro dos trazos de consumo de substancias pódese distinguir entre comportamentos de consumo de substancias que non son necesariamente problemáticos nin indicativos de adicción, e un consumo problemático que pode derivar nun trastorno polo uso de substancias (SUD). A susceptibilidade ao consumo de substancias explícase principalmente por dúas vías: unha vía externalizante, caracterizada por altos niveis de impulsividade, agresividade, asunción de riscos e procura de novas sensacións; e unha vía internalizante, a cal foi menos estudada, caracterizada polo consumo de substancias para afrontar un afecto negativo como estrés, depresión ou ansiedade. A susceptibilidade xenética a estes trazos, altamente polixénicos, é principalmente estudada mediante estudos de asociación do xenoma completo (GWAS). Estes estudos indicaron diferenzas na susceptibilidade xenética aos distintos trazos de consumo de substancias, sendo a correlación xenética entre comportamentos de consumo de substancias e SUD moderada. Ademais, atoparon un solapamento xenético relativamente alto dos trazos de consumo de substancias, tanto dos comportamentos de consumo de substancias como dos SUD, con trastornos psiquiátricos. A susceptibilidade xenética compartida entre trastornos psiquiátricos e SUD foi explicada mediante un factor xenético de psicopatoloxía xeral que predispón a calquera trastorno psiquiátrico, así como a través de factores xenéticos latentes correlacionados que aumentan o risco a un subconxunto de trazos.

Dentro dos distintos trastornos psiquiátricos, a esquizofrenia presenta unha asociación particularmente forte co consumo de tabaco. Esta maior prevalencia do consumo de tabaco en persoas con esquizofrenia contribúe notablemente á maior mortalidade deste grupo poboacional. Os GWAS mostraron que a alta comorbididade entre tabaquismo e esquizofrenia reflíctese nunha correlación xenética significativa entre ambos trazos. Ademais, os GWAS identificaron, ademais da agardada asociación dos xenes codificantes das subunidades de receptores colinérxicos nicotínicos (nAchR) co consumo de tabaco, unha asociación significativa dalgúns destes xenes con esquizofrenia. Existen principalmente tres hipóteses non excluíntes que tratan de explicar esta asociación entre tabaquismo e esquizofrenia: 1) a hipótese da automedicación, segundo a cal, os/as pacientes con esquizofrenia poderían fumar para contrarrestar os seus síntomas ou os efectos secundarios da medicación antipsicótica; 2) a hipótese do tabaco como un factor de risco para desenvolver esquizofrenia; e 3) a hipótese

USC UNIVERSIDADE DE SANTIAGO

segundo a cal os mesmos factores xenéticos e non xenéticos poderían incrementar de forma independente o risco a padecer esquizofrenia e tabaquismo. Estudos previos de aleatorización mendeliana aportaron evidencia dun efecto causal do tabaco sobre a esquizofrenia. Porén, é necesario abordar a relación causal entre tabaquismo e esquizofrenia dende diferentes perspectivas.

O obxectivo principal desta tese foi estudar os factores xenéticos que inflúen na susceptibilidade aos trazos de consumo de substancias en poboación xeral e, máis especificamente, no consumo de tabaco en pacientes con esquizofrenia. Con este fin, nos cinco estudos presentados nesta tese, empregouse unha batería de ferramentas xenómicas que se basean nos resultados de GWAS.

No Estudo 1, identificáronse 34 módulos de xenes cunha expresión correlacionada no cerebro a partir de datos de expresión xénica do peixe cebra dispoñibles publicamente, empregando *Weighted Gene Co-expression Network Analysis* (WGCNA). Estes datos nos que se baseou a análise recollían a expresión xénica no cerebro do peixe cebra adulto baixo diferentes condicións experimentais, entre elas, a abstinencia a substancias. A caracterización destes módulos mediante análises de enriquecemento en conxuntos de xenes de interese suxeriu que estes módulos de co-expresión poderían representar redes xénicas do cerebro humano, podendo estar, algúns deles, involucrados en trazos relacionados coa psiquiatría. En concreto, identificáronse módulos de xenes co-expresados cun patrón de expresión especificamente asociado á abstinencia a substancias no cerebro de peixe cebra. A continuación, comprobouse que unha puntuación de risco polixénico (PRS) baseada en variantes xenéticas destes xenes explica significativamente máis a dependencia a alcol en humanos que unha PRS baseada en variantes xenéticas doutros xenes expresados en cerebro. Polo tanto, a partir de datos de expresión de peixe cebra, puidéronse identificar uns conxuntos de xenes que parecen estar involucrados na dependencia a alcol en humanos, demostrando a utilidade do peixe cebra como modelo animal en psiquiatría.

No Estudo 2, empregouse a ferramenta *bivariate causal mixture model* (MiXeR) para caracterizar o patrón de solapamento xenético dos trazos de consumo de substancias con trazos externalizantes e internalizantes. Os resultados foron consistentes cun efecto da vía externalizante e a vía internalizante sobre os trazos de consumo de substancias, mais con diferenzas entre os comportamentos de consumo de substancias e os SUD. Os comportamentos de consumo de substancias presentaron un alto solapamento xenético con trazos internalizantes, mais cunha correlación xenética do efecto das variantes compartidas particularmente baixa,

indicando unha distribución diverxente do efecto da susceptibilidade xenética compartida para ambos trazos. Pola contra, estes trazos presentaron unha maior correlación xenética das variantes compartidas con trazos externalizantes, tamén no caso dun solapamento xenético considerable, suxerindo unha maior influencia da vía externalizante sobre os comportamentos de consumo de substancias que a vía internalizante. Por outro lado, os SUD presentaron, mesmo no caso dun solapamento xenético alto, unha correlación xenética das variantes compartidas elevada tanto con trazos externalizantes como con trazos internalizantes, aínda que máis acentuada cos primeiros. Este resultado é consistente cun efecto de ambas vías no desenvolvemento de SUD.

No Estudo 3, empregáronse diferentes métodos para avaliar se os sinais de asociación de GWAS para tabaquismo e esquizofrenia naqueles xenes de nAchR para os cales existe evidencia de asociación con ambos trazos son debidos á mesma variante causal, é dicir, se colocalizan. No xene *CHRNA2*, obtívose evidencia de colocalización entre consumo de tabaco e esquizofrenia, así como co trastorno polo uso de cannabis (CUD). Este resultado é consistente cunha pleiotropía vertical onde a asociación deste xene con esquizofrenia está mediada polo tabaquismo actuando como un factor de risco. Porén, os outros dous loci de nAchR analizados, en concreto, o xene *CHRNA4* e a agrupación de xenes *CHRNA5-CHRNA3-CHRNB4*, parecen estar asociados a tabaquismo e esquizofrenia por vías independentes.

No Estudo 4, comprobouse que a PRS estimada en poboación xeral para o inicio do consumo de tabaco explicou unha proporción da varianza do hábito tabáquico similar en persoas con esquizofrenia e en controis. Polo tanto, os factores xenéticos que aumentan o risco a tabaquismo en persoas con esquizofrenia parecen ser os mesmos que os da poboación xeral. Para confirmar este resultado, analizáronse tamén PRS para aqueles trazos xeneticamente correlacionados co inicio do consumo de tabaco con suficiente potencia nesta análise. Para algúns destes trazos correlacionados, en concreto, para o nivel de formación académica e o índice de masa corporal, as PRS estiveron tamén asociadas ao hábito tabáquico tanto de pacientes con esquizofrenia como de controis. Porén, PRS para outros trazos xeneticamente correlacionados co inicio do consumo de tabaco, especificamente, para o trastorno por déficit de atención e hiperactividade (ADHD) e para a idade á que se tivo o/a primeiro/a fillo/a, dous trazos relacionados cunha conduta externalizante, estiveron só asociadas a tabaquismo, tras corrixir por múltiples tests, en persoas con esquizofrenia. Estes resultados indican que a susceptibilidade xenética a problemas externalizantes parece explicar en maior medida a susceptibilidade a tabaquismo en pacientes con esquizofrenia que en poboación xeral. Ademais, unha PRS combinando todos os trazos analizados explicou mellor o hábito tabáquico de

USC UNIVERSIDADE DE SANTIAGO

pacientes con esquizofrenia e controis que a PRS individual para o inicio do consumo de tabaco, indicando que a susceptibilidade xenética a estes outros trazos ten un efecto adicional no risco a tabaquismo. Neste mesmo estudo, tamén se analizou a posible explicación para a susceptibilidade xenética compartida entre tabaquismo e esquizofrenia. Tal e como se agarda segundo un efecto causal do tabaco sobre a esquizofrenia, atopouse que a PRS para o inicio do consumo de tabaco deixou de estar significativamente asociada con esquizofrenia tras axustar a mostra por hábito tabáquico. O mesmo patrón foi atopado para a variante xenética máis significativa para esquizofrenia da agrupación de xenes de nAchR *CHRNA5-CHRNA3-CHRNB4*.

No Estudo 5, condicionouse o GWAS máis amplo para esquizofrenia polo efecto causal do inicio do consumo de tabaco empregando *multi-trait-based conditional and joint analysis* (mtCOJO). Ao descartar a susceptibilidade xenética a esquizofrenia mediada pola vía causal do tabaquismo, identificáronse 19 novos loci de risco para esquizofrenia cuxa asociación podería ter quedado oculta polo seu efecto sobre o inicio do consumo de tabaco en dirección contraria e 42 loci que deixaron de estar significativamente asociados a esquizofrenia, xa que a súa asociación no GWAS orixinal podería estar parcialmente mediada pola vía causal do tabaquismo. Ademais, neste estudo avaliouse o impacto de condicionar o GWAS de esquizofrenia por tabaquismo na correlación xenética da esquizofrenia cunha ampla variedade de trazos de interese, incluíndo trazos de consumo de substancias, trastornos psiquiátricos, trazos condutuais e trazos neuropsicolóxicos. Entre todos os trazos avaliados, a correlación xenética da esquizofrenia diminuíu significativamente cos comportamentos de consumo de substancias, SUD, ADHD e trazos condutuais relacionados coa asunción de riscos. Estes resultados indican, polo tanto, que a correlación xenética da esquizofrenia con conduta externalizante parece estar mediada, polo menos parcialmente, polo tabaquismo.

Os resultados destes tres últimos estudos permitiron afondar máis na causa da asociación entre tabaquismo e esquizofrenia. Por un lado, suxeriron que os factores xenéticos que explican o consumo de tabaco de persoas con esquizofrenia son os mesmos que os da poboación xeral, en contra da hipótese da automedicación. Por outro lado, proporcionaron probas consistentes de que parte da elevada prevalencia do tabaquismo nas persoas con esquizofrenia se debe a un efecto causal do inicio do consumo de tabaco sobre a esquizofrenia, aínda que este efecto podería ser indirecto a través do consumo de cannabis. En relación a isto, e en consonancia co efecto da vía externalizante sobre o inicio do consumo de tabaco, apoiaron que o factor xenético externalizante está asociado con esquizofrenia parcialmente a través da vía causal do tabaquismo. Ademais, nos xenes de nAchR, os cales están implicados na farmacoxenética da

USC UNIVERSIDADE DE SANTIAGO

nicotina, podería haber diferentes sinais de asociación cun patrón de asociación con tabaquismo e esquizofrenia variable.

Polo tanto, os resultados dos cinco estudos desta tese permitiron obter máis información sobre a susceptibilidade xenética ao consumo de substancias, especialmente ao tabaquismo en pacientes con esquizofrenia. Esta información podería ter importantes implicacións a nivel clínico, permitindo desenvolver estratexias máis eficaces para prever e tratar o consumo de substancias.

USC
UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

PUBLICACIÓNS INCLUÍDAS NA TESE

1. **Estudo 1**: Al-Soufi, L., Arana, A. J., Facal, F., Flórez, G., Vázquez, F. L., Arrojo, M., the GenPol Study Group, Sánchez, L. e Costas, J. (2024). Identification of gene co-expression modules from zebrafish brain data: applications in psychiatry illustrated through alcohol-related traits [Traballo non publicado].
2. **Estudo 2:** Al-Soufi, L., Hindley, G., Rødevand, L., Shadrin, A. A., Jaholkowski, P., Fominykh, V., Icick, R., Tesfaye, M., Costas, J. e Andreassen, O. A. (2024). Polygenic overlap of substance use behaviors and disorders with externalizing and internalizing problems independent of genetic correlations [Traballo non publicado].
3. **Estudo 3:** Al-Soufi, L. e 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.
4. **Estudo 4:** Al-Soufi, L., Martorell, L., Moltó, M. D., González-Peñas, J., García-Portilla, M. P., Arrojo, M., Rivero, O., Gutiérrez-Zotes, A., Nácher, J., Muntané, G., Paz, E., Páramo, M., Bobes, J., Arango, C., Sanjuan, J., Vilella, E. e Costas, J. (2021). A polygenic approach to the association between smoking and schizophrenia. *Addiction biology, 27*(1), e13104. https://doi.org/10.1111/adb.13104.
5. **Estudo 5:** Al-Soufi, L. e Costas, J. (2023). Genetic susceptibility for schizophrenia after adjustment by genetic susceptibility for smoking: implications in identification of risk genes and genetic correlation with related traits. *Psychological Medicine, 53*(14), 6806–6816. https://doi.org/10.1017/S0033291723000326.

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

ABREVIATURAS

ADHD:	trastorno por déficit de atención e hiperactividade
AUD:	trastorno polo uso de alcol
CEIC:	Comité Ético de Investigación Clínica
CUD:	trastorno polo uso de cannabis
DSM-5:	Manual Diagnóstico e Estatístico dos Trastornos Mentais 5
eQTL:	loci de trazos cuantitativos de expresión
GCTA:	*Genome-wide Complex Trait Analysis*
Genomic SEM:	modelos de ecuacións estruturais xenómicas
GPC:	*Genetics of Personality Consortium*
GSCAN:	*GWAS & Sequencing Consortium of Alcohol and Nicotine use*
GSMR:	*generalized summary data-based Mendelian randomization*
GTEx:	*Genotype-Tissue Expression*
GWAS:	estudos de asociación do xenoma completo
LD:	desequilibrio de ligamento
LDSC:	*LD score regression*
MAGMA:	*Multi-marker Analysis of GenoMic Annotation*
MiXeR:	*bivariate causal mixture model*
MR:	aleatorización mendeliana
MTAG:	*multi-trait analysis of GWAS*
mtCOJO:	*multi-trait-based conditional and joint analysis*
nAchR:	receptores colinérxicos nicotínicos
OUD:	trastorno polo uso de opioides

USC UNIVERSIDADE DE SANTIAGO DE

PGC:	*Psychiatric Genomics Consortium*
PRS:	puntuacións de risco polixénico
SNP:	polimorfismos dun único nucleótido
SSGAC:	*Social Science Genetic Association Consortium*
SUD:	trastorno polo uso de substancias
TUD:	trastorno polo uso de tabaco
WGCNA:	*Weighted Gene Co-expression Network Analysis*

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

1. INTRODUCIÓN

1.1. TRAZOS DE CONSUMO DE SUBSTANCIAS: COMPORTAMENTOS DE CONSUMO DE SUBSTANCIAS E TRASTORNOS POLO USO DE SUBSTANCIAS

1.1.1. Epidemioloxía dos trazos de consumo de substancias

O consumo de substancias, tanto legais como o tabaco ou o alcol, como ilegais como o cannabis, a cocaína, os opioides ou as anfetaminas, representa un grave problema de saúde a nivel global. Os datos epidemiolóxicos destacan a alta prevalencia do consumo de substancias en todo o mundo. Unha meta-análise de estudos epidemiolóxicos realizados en 195 países mostrou que, en 2016, 100 millóns de persoas foron diagnosticadas con dependencia a alcol (Degenhardt, Charlson, et al., 2018). A dependencia a opioides e a cannabis situouse a continuación, con ao redor de 26 e 22 millóns de casos, respectivamente. En Australia, un estudo de 2007 mostrou que o 94,1% da poboación consumira alcol algunha vez e o 64,5% era consumidora regular (Degenhardt, Glantz, et al., 2018). No caso do cannabis, o 19,8% consumírao algunha vez e o 4,4% de forma regular. Nos Estados Unidos, aproximadamente o 30,3% da poboación experimenta algún trastorno polo uso de alcol (AUD, do inglés *alcohol use disorder*) ao longo da súa vida, o 17,7% desenvolve dependencia a nicotina e o 10,3% presenta algún trastorno polo uso de substancias ilegais (D. S. Hasin & Grant, 2015).

Esta problemática tamén está presente a nivel máis local. En Galicia, o Observatorio Español das Drogas e as Adiccións, informou que, en 2022, o 94,7% da poboación de 15 a 64 anos consumira bebidas alcohólicas algunha vez na vida (Observatorio Español de las Drogas y las Adicciones, 2023). O tabaco foi a segunda substancia máis consumida en Galicia en 2022, co 70,6% da poboación galega declarando ter fumado algunha vez. O cannabis foi a substancia ilegal máis estendida en Galicia, consumindo cannabis algunha vez o 41,5% da poboación. A nivel estatal, o tabaco foi a substancia máis consumida diariamente en 2022, co 33,1% da poboación española sendo fumadora diaria; seguido do alcol, co 9% da poboación sendo consumidora de alcol diaria, aínda que esta porcentaxe ascende ao 64,5% se consideramos o consumo nos últimos 30 días. Finalmente, o 2,8% e o 1,9% da poboación española foi consumidora diaria de cannabis e opioides, respectivamente. Ademais, este informe proporcionou datos sobre a prevalencia dun consumo problemático de substancias en España. En 2021, rexistráronse en España 69.487 admisións a tratamento por abuso ou dependencia a

substancias, excluíndo o tabaco, sendo o alcol a substancia que deu lugar a un maior número de admisións, o 36,2% delas.

Este consumo de substancias ten graves consecuencias para a saúde, contribuíndo notablemente á carga mundial de enfermidades (Degenhardt, Charlson, et al., 2018; Schulte & Hser, 2013; Whiteford et al., 2015). Só o alcol foi responsable de preto de 3 millóns de mortes en 2016 a nivel global (Degenhardt, Charlson, et al., 2018). Aínda que a maior parte da carga mundial de enfermidades atribúeselle ao consumo de alcol en comparación coas substancias ilegais, ambas representan importantes factores de risco a enfermidades. O tabaquismo é tamén unha das principais causas de enfermidades evitables e mortalidade prematura (Mokdad, 2004; G. M. J. Taylor & Munafò, 2018). Nos Estados Unidos, mentres que o consumo de alcol foi responsable do 3,5% das mortes en 2000, o tabaco foi a principal causa modificable de morte, explicando o 18,1% das mortes (Mokdad, 2004). En poboación española, arredor do 13% das mortes producidas en 2016 e 2017 foron debidas ao tabaco, sendo o cancro de pulmón a principal causa de morte dentro daquelas atribuíbles a esta substancia (Rey-Brandariz et al., 2021). Confirmando estas observacións a nivel epidemiolóxico, unha variante xenética asociada cunha maior intensidade do consumo de tabaco incrementou o risco de mortalidade nun 9% dentro das persoas fumadoras (Vie et al., 2019). O consumo regular de substancias tamén pode levar ao desenvolvemento dun trastorno polo uso de substancias (SUD, do inglés *substance use disorder*), provocando un elevado sufrimento nas persoas afectadas. Datos poboacionais de Australia demostraron que o 22,2% e o 19,9% dos individuos que consumen cannabis e alcol, respectivamente, desenvolven un abuso de substancias (Degenhardt, Glantz, et al., 2018). Unha ampla meta-análise corroborou esta alta prevalencia dun consumo problemático de cannabis unha vez que se consume de forma regular (Leung et al., 2020).

Os estudos epidemiolóxicos tamén revelaron unha alta comorbididade entre o consumo de diferentes substancias. Moitos individuos consumen dúas ou máis substancias nun mesmo período de tempo, o que se coñece como policonsumo (Connor et al., 2014). Nunha mostra de 472.624 veteranos con, polo menos, un SUD, o 24% tiña dous ou tres SUD e o 2,7% tiña máis de tres SUD (Bhalla et al., 2017). En España, en 2022, o 33,7% da poboación consumiu dúas ou máis substancias diferentes no último mes, ascendendo esta porcentaxe ao 40,9% se se teñen en conta os últimos 12 meses (Observatorio Español de las Drogas y las Adicciones, 2023). Este policonsumo está asociado a unha maior gravidade do mesmo (Bhalla et al., 2017).

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

En resumo, o consumo de substancias é un problema de saúde pública global cun impacto considerable na saúde e benestar das persoas e na sociedade en xeral. Para abordar esta problemática, son necesarias estratexias eficaces de prevención e tratamento.

1.1.2. Definición e diagnóstico dos trazos de consumo de substancias

O proceso de consumo de substancias comeza cun inicio do consumo, progresando cara un consumo regular e crónico que pode chegar a ser problemático e evolucionar cara unha adicción (Agrawal et al., 2012). Así, dentro dos trazos de consumo de substancias podemos distinguir entre comportamentos relacionados cun consumo de substancias que non son necesariamente problemáticos nin indicativos dunha adicción (denominados neste traballo como comportamentos de consumo de substancias) e comportamentos de consumo problemáticos que poden derivar nun SUD (englobados neste traballo dentro do termo SUD). Os principais comportamentos de consumo de substancias que se analizan neste traballo, así como algúns exemplos á hora de definilos na bibliografía, aparecen recollidos na Táboa 1.

Táboa 1. Principais comportamentos de consumo de substancias e algunhas formas de definir estes comportamentos de consumo de substancias para tabaco, alcol ou cannabis na bibliografía a modo de exemplo.

Comportamentos de consumo de substancias	Formas de definir os comportamentos de consumo de substancias
Consumo ocasional da substancia	Consumir bebidas alcólicas menos dunha vez á semana (Marees et al., 2019)
	Ter fumado máis dunha vez, mais non no último mes ou semana (Mayhew et al., 2000)
	Ter fumado máis dun cigarro pero menos de 100 cigarros na vida (Mayhew et al., 2000)
Consumo regular da substancia	Ter fumado máis de 100 cigarros ao longo da vida (Hartz et al., 2014; M. Liu et al., 2019)
	Fumar todos os días durante, polo menos, un mes (Hartz et al., 2014; M. Liu et al., 2019)
	Consumir bebidas alcólicas, polo menos, unha ou dúas veces á semana (Marees et al., 2019)
Consumo grave da substancia	Fumar 30 ou máis cigarros ou outros produtos de tabaco ao día (Koyanagi et al., 2016)
	Consumir máis de 4 bebidas alcólicas nun mesmo día (Hartz et al., 2014)
	Segundo o test de identificación de trastornos polo uso de alcol (AUDIT), consumir 6 ou máis bebidas alcólicas nunha mesma ocasión (J. B. Saunders et al., 1993)
	Consumir cannabis máis de 21 veces nun mesmo ano (Hartz et al., 2014)
Inicio do consumo da substancia	Ter sido algunha vez na vida fumador/a regular fronte a nunca ter sido fumador/a regular (M. Liu et al., 2019)
	Ter consumido algunha vez na vida cannabis fronte a nunca ter consumido cannabis (Pasman et al., 2018)
Idade de inicio do consumo da substancia	Idade á que se iniciou o consumo de tabaco de forma regular (M. Liu et al., 2019)

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

Intensidade do consumo da substancia	Número medio de cigarros fumados por día, tanto sendo fumador/a actualmente como ex-fumador/a (M. Liu et al., 2019)
	Número medio de bebidas alcólicas de todo tipo consumidas por semana (M. Liu et al., 2019)
	Maior número de bebidas alcólicas consumidas nun día nun mes típico (Deak et al., 2022; Kendler et al., 2010)
	Segundo o test de identificación de trastornos polo uso de alcol (AUDIT), a combinación da frecuencia coa que se consume unha bebida alcólica, do número de bebidas alcólicas consumidas nun día típico cando bebes e da frecuencia coa que se consumen seis ou máis bebidas nunha ocasión (J. B. Saunders et al., 1993)
	Frecuencia coa que se consumen bebidas alcólicas (Marees et al., 2019)
	Número medio de vasos de bebidas alcólicas consumidos por semana ponderado por un factor dependendo do tipo de alcol (Marees et al., 2019)
	Número medio de unidades de bebidas alcólicas de todo tipo consumidas por semana en consumidores actuais (Clarke et al., 2017)
	Número de días nos que se consumen bebidas alcólicas nun mes típico do ano pasado (Kendler et al., 2010)
	Frecuencia no ano pasado coa que se consumiron suficientes bebidas alcólicas como para sentirse ebrio (Kendler et al., 2010)
	Frecuencia coa que se consume alcol, dende 1 (nunca) a 6 (diariamente ou case diariamente) (Polimanti et al., 2019)
	Naqueles que consumen alcol polo menos unha ou dúas veces por semana, número de porcións de bebidas alcólicas consumidas de media por semana (Polimanti et al., 2019)
Gravidade do consumo da substancia	Medida que recolle o inicio, a duración e a intensidade do consumo de tabaco, así como o tempo transcorrido dende que se deixou de consumir tabaco (Wootton et al., 2020)
Deixar de consumir a substancia	Ser fumador/a actualmente fronte a ser ex-fumador/a (M. Liu et al., 2019)

O consumo regular dunha substancia pode converterse en problemático cando se manifestan síntomas como unha busca compulsiva da substancia, unha falta de control no consumo, un deterioro das relacións sociais, unha asunción de riscos para consumir a substancia e síntomas de abstinencia (Prom-Wormley et al., 2017). A experimentación destes problemas asociados ao consumo de substancias recibe o diagnóstico de SUD, o cal é considerado un trastorno psiquiátrico e, como tal, aparece definido no Manual Diagnóstico e Estatístico dos Trastornos Mentais 5 (DSM-5). A adicción é a fase máis grave e crónica dun SUD, caracterizada pola perda do autocontrol e un comportamento compulsivo de busca da substancia a pesar dos esforzos por interromper o consumo.

Segundo o DSM-5, o trastorno polo uso de tabaco (TUD, do inglés *tobacco use disorder*), por exemplo, aínda que o diagnóstico doutro SUD sería similar, é definido como un patrón problemático de consumo de tabaco que provoca un deterioro ou un malestar clinicamente significativo e no cal se manifestan polo menos dous dos seguintes feitos nun prazo de 12 meses (American Psychiatric Association, 2013):

USC UNIVERSIDADE DE SANTIAGO

1) Consúmese tabaco con frecuencia en cantidades superiores ou durante un tempo máis prolongado do previsto.
2) Existe un desexo persistente ou esforzos fracasados por deixar ou controlar o consumo de tabaco.
3) Invístese moito tempo nas actividades necesarias para a obtención ou consumo de tabaco.
4) Existen ansias ou un poderoso desexo ou necesidade de consumir tabaco.
5) Existe un consumo recorrente de tabaco que leva ao incumprimento das obrigas importantes no traballo, escola ou fogar.
6) Existe un consumo continuado de tabaco a pesar de sufrir problemas sociais ou interpersonais persistentes ou recorrentes, provocados ou exarcebados polos efectos do tabaco. Por exemplo, discusións con outras persoas sobre o consumo de tabaco.
7) O consumo de tabaco provoca o abandono ou a redución de importantes actividades sociais, profesionais ou de ocio.
8) Existe un consumo recorrente de tabaco en situacións nas que provoca un risco físico como, por exemplo, na cama.
9) Continúase co consumo de tabaco a pesar de saber que se sofre un problema físico ou psicolóxico persistente ou recorrente, probablemente causado ou exarcebado polo tabaco.
10) Existe tolerancia ao tabaco, definida por algún dos seguintes feitos:
    a. Unha necesidade de consumir cantidades cada vez maiores de tabaco para conseguir o efecto desexado.
    b. Un efecto notablemente reducido tras o consumo continuado da mesma cantidade de tabaco.
11) Existe abstinencia, manifestada nalgún dos seguintes feitos:
    a. Presenza do síndrome de abstinencia.
    b. Consúmese tabaco para aliviar ou evitar os síntomas da abstinencia.

O DSM-5 clasifica os SUD segundo a súa gravidade en leve (presenza de 2 ou 3 síntomas), moderado (presenza de 4 ou 5 síntomas) ou grave (presenza de 6 ou máis síntomas), en función dos criterios cumpridos. Anteriormente, o DSM-IV clasificaba aos individuos cun consumo problemático de substancias en dúas categorías diferentes: abuso de substancias e dependencia a substancias (American Psychiatric Association, 1994). Mentres que o abuso de substancias implicaba o consumo disfuncional de substancias cun deterioro das interaccións sociais e un incumprimento das responsabilidades laborais, escolares ou sociais; a dependencia a

USC DE COMPOSTELA

substancias era unha forma máis grave con síntomas relacionados coa tolerancia e a abstinencia. Actualmente, o termo SUD engloba o abuso de substancias, a dependencia a substancias e o consumo problemático de substancias. Ademais, equivale aos trastornos mentais e do comportamento debidos ao uso de substancias definidos na Clasificación Internacional de Enfermidades (https://icd.who.int/en).

1.1.3. O ciclo da adicción

Un ciclo de tres fases foi proposto para explicar o proceso de desenvolvemento dun SUD (Goldstein & Volkow, 2002). Este modelo está formado por unha fase de atacadura/intoxicación, unha fase de abstinencia/afecto negativo e outra fase de preocupación/anticipación (Koob & Volkow, 2010). Na primeira destas fases, o consumo de substancias produce un reforzo positivo inicial debido á estimulación do circuíto neuronal de recompensa, o cal dá lugar a un comportamento impulsivo de busca da substancia. A medida que o consumo de substancias continúa e se avanza cara un SUD, pásase dun reforzo positivo a un reforzo negativo. O consumo da substancia pasa a ser compulsivo para aliviar o afecto negativo e o síndrome de abstinencia. Este comportamento leva a unha preocupación cognitiva polo consumo da substancia e un deterioro da función executiva. A función executiva comprende os diferentes procesos cognitivos que son necesarios para o control cognitivo do comportamento. Polo tanto, o seu deterioro dificulta o control dos impulsos e a toma de decisións. Durante as distintas fases do modelo, como consecuencia da exposición crónica á substancia, prodúcese unha neuroadaptación que provoca déficits no sistema de recompensa, contribuíndo ao consumo compulsivo; déficits no sistema de estrés durante a fase de abstinencia/afecto negativo; e cambios no sistema de función executiva (Uhl et al., 2019).

1.1.4. Base xenética dos trazos de consumo de substancias

Os trazos de consumo de substancias presentan unha etioloxía complexa, caracterizada por unha arquitectura xenética altamente polixénica e unha forte influencia de factores ambientais. Nestes trazos polixénicos, numerosas variantes xenéticas comúns, é dicir, variantes relativamente frecuentes na poboación, contribúen ao risco ao trazo con tamaños de efecto pequenos (Sullivan et al., 2018). Estas variantes xenéticas comúns son principalmente polimorfismos dun único nucleótido (SNP, do inglés *single nucleotide polymorphism*), variantes que afectan a un único nucleótido do ADN e que están presentes, polo menos, no 1% da poboación (Andreassen et al., 2023).

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

Os estudos de xemelgos mostraron que os factores xenéticos teñen unha influencia considerable nos distintos aspectos do consumo de substancias, dende a probabilidade do

consumo de substancias ata a intensidade do consumo de substancias e o desenvolvemento dun SUD (Dick, 2016; Goldman et al., 2005; Prom-Wormley et al., 2017). Segundo estes estudos, a variación xenética explica sobre o 33-71% da variabilidade da dependencia a nicotina, o 48-66% da dependencia a alcol e o 51-59% da dependencia a cannabis (Agrawal et al., 2012). A herdabilidade dos comportamentos de consumo de substancias, aínda que tamén é significativa, é menor (Goldman et al., 2005; Vink et al., 2005).

A menor herdabilidade dos comportamentos de consumo de substancias en comparación cos SUD mostra que, ao longo do curso do consumo de substancias, a influencia dos factores xenéticos varía (Agrawal et al., 2012; Kendler et al., 2003; Pagan et al., 2006). Mentres que as primeiras etapas do consumo de substancias parecen estar máis influenciadas por factores ambientais, os factores xenéticos parecen cobrar maior importancia nas fases posteriores de consumo problemático e dependencia. Por exemplo, un estudo realizado a partir dos datos de 1572 xemelgos holandeses evidenciou que, mentres que os factores de risco xenéticos explicaron só o 44% das súas diferencias no inicio do consumo de tabaco, o 75% das súas diferencias na dependencia a nicotina foron explicadas por factores xenéticos (Vink et al., 2005). Este estudo tamén descartou que os mesmos factores xenéticos e ambientais aumenten o risco ao inicio de consumo de tabaco e a dependencia a nicotina. En lugar diso, parece que existe tanto unha compoñente compartida como unha compoñente específica a cada trazo, as cales interactúan para determinar se un individuo é fumador/a.

A influencia dos factores xenéticos sobre os trazos de consumo de substancias tamén parece variar considerablemente dependendo do momento do desenvolvemento no que se estime, sendo menor na adolescencia e aumentando durante a idade adulta (Kendler et al., 2008; Prom-Wormley et al., 2017). Durante a adolescencia, factores ambientais como un consumo de substancias do grupo de iguais, unha escasa calidade dos vínculos familiares, unha baixa cohesión social ou unha alta dispoñibilidade de drogas explican máis da variabilidade do consumo de substancias que factores xenéticos (Kendler et al., 2008).

Os datos de xemelgos tamén indicaron que a maior parte dos factores de risco xenéticos e ambientais para os comportamentos de consumo de substancias e SUD non son específicos do tipo de substancia, o cal apoia a existencia dunha susceptibilidade xenética e ambiental común que aumenta o risco ao consumo de diferentes substancias (Kendler et al., 2003; Ystrom et al., 2014). Porén, tamén se identificaron factores xenéticos e ambientais específicos para cada substancia que confiren predisposición ao consumo desa substancia en particular (Haberstick et al., 2011; Palmer et al., 2012; Ystrom et al., 2014).

1.1.5. A vía externalizante e a vía internalizante cara o consumo de substancias

Para mellorar a prevención e o tratamento do consumo de substancias, é fundamental comprender a súa orixe. Propuxéronse principalmente dúas vías principais polas que os individuos comezan a consumir substancias, mantéñense neste consumo de substancias e progresan ata un SUD: unha vía externalizante, caracterizada por altos niveis de impulsividade, agresividade, asunción de riscos e procura de novas sensacións; e unha vía internalizante, caracterizada polo consumo de substancias para afrontar un afecto negativo, depresión e ansiedade (Hussong et al., 2011; Krueger et al., 2002).

1.1.5.1. Evidencia epidemiolóxica da vía externalizante e a vía internalizante

Os estudos lonxitudinais apoian a existencia destas dúas vías cara o consumo de substancias, atopando unha asociación dos problemas externalizantes e internalizantes cun posterior consumo de substancias. Un estudo lonxitudinal con 10.533 participantes atopou que tanto problemas externalizantes severos como internalizantes predín independentemente o inicio do tabaquismo (Green et al., 2018). Outra meta-análise de 12 estudos lonxitudinais confirmou que os síntomas externalizantes e internalizantes durante a infancia e a adolescencia aumentan nun 21% e nun 62%, respectivamente, o risco de desenvolver AUD na adultez temperá (Meque et al., 2019).

A existencia destas dúas vías cara o consumo de substancias tamén se ve apoiada pola evidencia de estudos lonxitudinais dun efecto causal dos trazos externalizantes e internalizantes sobre os trazos de consumo de substancias. Por un lado, existe forte evidencia de que o diagnóstico na infancia do trastorno por déficit de atención e hiperactividade (ADHD, do inglés *attention-deficit hyperactivity disorder*), un trastorno psiquiátrico externalizante, predí futuros comportamentos de consumo de substancias e SUD (Charach et al., 2011; S. S. Lee et al., 2011). Por outro lado, trastornos psiquiátricos internalizantes, como a depresión ou a ansiedade, están asociados con posteriores trazos de consumo de substancias (M. Fluharty et al., 2017; Swendsen et al., 2010).

1.1.5.2. Diferenzas no efecto da vía externalizante e a vía internalizante sobre o consumo de substancias

A evidencia dunha vía internalizante é menos robusta que para unha vía externalizante (Edwards et al., 2016; S. M. King et al., 2004; Ning et al., 2020). Mentres que os estudos lonxitudinais atoparon unha clara asociación entre problemas externalizantes e o consumo de substancias, os problemas internalizantes mostraron asociacións menos consistentes (Edwards et al., 2016; Pedersen et al., 2018). Unha revisión sistemática de estudos lonxitudinais nos cales

USC UNIVERSIDADE DE SANTIAGO

se analizou o efecto de problemas externalizantes e internalizantes antes dos 18 anos na presenza posterior de distintos niveis de consumo de alcol atopou posibles explicacións a esta menor evidencia dunha vía internalizante (Ning et al., 2020). Os problemas internalizantes estiveron positivamente asociados a un consumo excesivo e problemático de alcol e a un AUD en individuos adolescentes ou na transición á idade adulta, especialmente tras axustar por problemas externalizantes, mais estiveron negativamente asociados a un consumo leve de alcol. Polo tanto, a vía internalizante parece ter un maior efecto nos niveis máis severos de consumo de substancias. Ademais, en liña co resultado de Ning et al. (2020), a sintomatoloxía internalizante parece non dar lugar a un maior consumo de substancias cando convive con síntomas externalizantes elevados, mesmo tendo un efecto protector neste caso (Colder et al., 2018). Estas dúas vías tamén parecen ter un efecto diferente segundo o xénero, sendo a vía externalizante a principal en homes e a internalizante en mulleres (Dir & Hulvershorn, 2019; D. S. Hasin & Grant, 2015; Heitzeg et al., 2018). Polo tanto, a evidencia mixta en canto ao impacto da vía internalizante no consumo de substancias pode deberse a variacións segundo o momento do desenvolvemento, o xénero, a comorbididade con problemas externalizantes ou o tipo de consumo de substancias analizado.

1.2. ASOCIACIÓN ENTRE TRAZOS DE CONSUMO DE SUBSTANCIAS E TRASTORNOS PSIQUIÁTRICOS

1.2.1. Asociación entre trastornos polo uso de substancias e outros trastornos psiquiátricos

Múltiples estudos epidemiolóxicos apoian que unha parte substancial das persoas que padecen un trastorno psiquiátrico presentan tamén, nalgún momento das súas vidas, un SUD (Buckley, 2006; D. Hasin & Kilcoyne, 2012; Kessler, 2004). Esta comorbididade é tan alta que se creou un termo específico: trastorno dual, que fai referencia a padecer, ao mesmo tempo, un SUD e un trastorno psiquiátrico (Buckley, 2006; Szerman et al., 2022). Un estudo realizado con datos recollidos nas unidades de asistencia a trastornos afectivos de Galicia atopou que o 56,3% dos/as pacientes cun diagnóstico de dependencia ou abuso de substancias presentan algún outro trastorno psiquiátrico (Pereiro et al., 2013). Outro estudo indicou que a probabilidade de transición dun consumo de substancias normalizado a un SUD é maior para individuos diagnosticados cun trastorno psiquiátrico en comparación con individuos sen un diagnóstico, sendo particularmente forte esta asociación para nicotina (Lev-Ran et al., 2013). Dentro dunha mostra de 190 adolescentes estadounidenses de orixe hispánico cun diagnóstico de abuso de substancias, só o 16,8% non presentaron síntomas consistentes con algunha outra condición psiquiátrica (Gattamorta et al., 2017). Ademais, aqueles/as adolescentes cun trastorno dual comezaron a consumir cannabis e alcol de forma máis temperá que os/as adolescentes que só

USC DE SANTIAGO DE COMPOSTELA

estaban diagnosticados con SUD. Nunha cohorte de 2.494 individuos mozos/as suecos/as cunha hospitalización por un primeiro episodio psicótico, o 39% foi tamén hospitalizado, polo menos unha vez na vida, por SUD (Strålin & Hetta, 2020).

Esta alta comorbididade entre trastornos psiquiátricos e SUD ten graves consecuencias para os/as pacientes (Buckley, 2006). O consumo de substancias agrava os síntomas do trastorno psiquiátrico xa que desestabiliza aos/ás pacientes e provoca unha baixa adherencia á medicación. Esta asociación tamén aumenta os comportamentos de risco nos/as pacientes, así como a súa agresividade e o seu risco de suicidio. Ademais, ao igual que na poboación xeral, o consumo de substancias en pacientes con trastornos psiquiátricos asóciase a unha peor saúde.

1.2.2. Asociación entre consumo de cannabis e psicose

Dentro dos trazos de consumo de substancias, os cales presentan unha alta prevalencia en pacientes con trastornos psiquiátricos, o consumo de cannabis presenta unha clara asociación con psicose (Fergusson et al., 2015). Por exemplo, un estudo realizado nunha mostra de 109.308 participantes sen un diagnóstico de trastorno psicótico mostrou unha asociación significativa entre a frecuencia de consumo de cannabis e experiencias psicóticas (Wainberg et al., 2021).

A evidencia apunta claramente a un efecto causal do consumo de cannabis no risco de desenvolver psicose. Os estudos lonxitudinais demostran que o consumo de cannabis aumenta o risco posterior de psicose (Moore et al., 2007; Vaucher et al., 2018). Un estudo lonxitudinal realizado en 5.300 participantes da mostra *Avon Longitudinal Study of Parents and Children* atopou evidencia sólida de que o consumo de cannabis, tanto cun comezo temperá como tardío, incrementa o risco de experiencias psicóticas futuras, cun efecto mínimo de posibles factores de confusión (Jones et al., 2018). Ademais, observouse unha relación dependente de dose, cunha maior asociación naqueles individuos cun consumo de cannabis máis grave e unha menor asociación naqueles que consumen cannabis ocasionalmente (Moore et al., 2007). En concreto, o consumo diario de cannabis de alta potencia aumenta 3,2 veces o risco de psicose (Di Forti et al., 2019). As probas son suficientemente sólidas para que, a nivel de saúde pública, se asuma que o consumo de cannabis pode aumentar o risco de trastornos psicóticos (Gage, Hickman, et al., 2016; Ortiz-Medina et al., 2018).

1.2.3. Asociación entre tabaquismo e esquizofrenia

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

1.2.3.1. Evidencia epidemiolóxica da asociación entre tabaquismo e esquizofrenia

O tabaquismo é común entre as persoas con trastornos mentais (Centers for Disease Control and Prevention, 2013; Lawrence et al., 2009; Prochaska et al., 2017). Dentro dos

distintos trastornos psiquiátricos, a esquizofrenia e outros trastornos psicóticos presentan unha asociación particularmente forte co consumo de tabaco (de Leon & Diaz, 2005; Hughes et al., 1986; McClave et al., 2010). A prevalencia do tabaquismo é alta en pacientes con esquizofrenia e non se está vendo reducida nos últimos anos, a pesar da tendencia xeral dunha redución do tabaquismo na poboación (Dickerson, Schroeder, et al., 2018). Na mostra española CLAMORS, con 1.704 pacientes diagnosticados con esquizofrenia, a prevalencia do tabaquismo foi do 54,54%, considerablemente maior que na poboación xeral, cunha prevalencia do 31,51% (Bobes et al., 2010). Noutra mostra de 588 pacientes hospitalizados, tamén se observou unha diferenza notable na prevalencia do tabaquismo en pacientes con esquizofrenia e controis, sendo do 75% e o 55%, respectivamente (de Leon et al., 2002).

Unha meta-análise de estudos realizados en diversos contextos socioculturais e países confirma unha asociación sólida entre tabaquismo e esquizofrenia, que persiste ao empregar pacientes con trastornos mentais severos como controis (de Leon & Diaz, 2005). Segundo este estudo, en comparación coa poboación xeral, os/as pacientes con esquizofrenia presentan un maior risco a comezar a fumar de forma diaria, un tabaquismo máis severo, unha maior dependencia a nicotina e unha menor taxa de cesamento do consumo de tabaco. Outro estudo posterior tamén demostrou que os/as pacientes con esquizofrenia teñen unha menor probabilidade de deixar de fumar en comparación coa poboación xeral e con pacientes con outros trastornos psiquiátricos (Zeng et al., 2020). Unha meta-análise baseada nun total de 14.555 persoas fumadoras e 273.162 persoas non fumadoras confirmou que o consumo diario de tabaco se asocia cun maior risco a padecer un trastorno psicótico e a desenvolvelo a unha idade máis temperá (Gurillo et al., 2015). Outro estudo tamén evidenciou esta asociación, atopando que os individuos con trastornos psicóticos presentan un maior risco a fumar de forma regular, mais tamén a un consumo grave de alcol e cannabis e a un consumo de drogas recreativas (Hartz et al., 2014). Porén, a evidencia suxire que a asociación entre tabaquismo e esquizofrenia non pode explicarse totalmente pola dependencia a alcol ou a outras substancias actuando como un factor de confusión (de Leon et al., 2002).

O tabaquismo tamén está asociado con experiencias psicóticas subclínicas na poboación xeral, incluíndo alucinacións, delirios e interferencias do pensamento (Barkhuizen et al., 2019; Bhavsar et al., 2018; Gage et al., 2014; Koyanagi et al., 2016; Saha et al., 2011; Van Gastel et al., 2013; Wolfe et al., 2017). Nun destes estudos, realizado a partir dos datos recollidos na enquisa de saúde internacional da Organización Mundial da Saúde, comprobouse que ter experimentado polo menos un síntoma psicótico está asociado a ser fumador/a actualmente e a un tabaquismo grave entre os/as fumadores/as diarios (Koyanagi et al., 2016). Noutros estudos,

USC UNIVERSIDADE DE SANTIAGO

atopouse que esta asociación é dependente de dose, incrementando o número de experiencias psicóticas coa intensidade do consumo de tabaco (Bhavsar et al., 2018; Saha et al., 2011). Non obstante, esta relación dependente de dose non foi corroborada por outro estudo (Wolfe et al., 2017). Varios destes estudos comprobaron que esta asociación entre fumar e ter experimentado experiencias psicóticas permanece despois de axustar polo consumo doutras drogas, entre elas, o cannabis (Bhavsar et al., 2018; Van Gastel et al., 2013; Wolfe et al., 2017).

1.2.3.2. Consecuencias do tabaquismo en pacientes con esquizofrenia

A maior prevalencia do tabaquismo en persoas con esquizofrenia ten graves consecuencias para elas, contribuíndo notablemente á maior mortalidade deste grupo poboacional, cunha esperanza de vida case 15 anos menor que a da poboación xeral (Dickerson, Origoni, et al., 2018; Hjorthøj et al., 2017). Nun estudo realizado en 1.213 persoas cun trastorno psicótico relacionado coa esquizofrenia, comprobouse que a mortalidade é significativamente maior entre as persoas fumadoras en comparación coas non fumadoras, principalmente debido a unha maior mortalidade por causas cardíacas (Kelly et al., 2011). Do mesmo xeito, noutro estudo estimouse que os/as pacientes con esquizofrenia teñen un risco do 30% de mortalidade nun período de 20 anos (Stolz et al., 2019). As enfermidades cardiovasculares foron a principal causa de mortalidade en persoas con esquizofrenia, explicando o tabaquismo o 86% dos casos de mortalidade por enfermidade cardiovascular. Na mostra española CLAMORS, estimouse que, se os/as pacientes con esquizofrenia fumadores/as deixasen de fumar, reduciríase o seu risco dun evento cardiovascular en 10 anos en case o 90% (Bobes et al., 2010). Todos estes estudos confirman o impacto do tabaquismo na maior mortalidade dos/as pacientes con esquizofrenia. Polo tanto, comprender as bases desta asociación entre tabaquismo e esquizofrenia é crucial para mellorar a saúde das persoas con esquizofrenia.

1.2.3.3. Posibles explicacións á asociación entre tabaquismo e esquizofrenia

Aínda que a asociación entre tabaquismo e esquizofrenia é clara, a súa causa segue sendo obxecto de debate na comunidade científica (Firth et al., 2023; Sagud et al., 2019). Principalmente, existen tres hipóteses non excluíntes que tratan de explicar esta asociación (Sagud et al., 2019): 1) a hipótese da automedicación, segundo a cal, os/as pacientes con esquizofrenia poderían fumar para contrarrestar os seus síntomas ou os efectos secundarios da medicación antipsicótica; 2) a hipótese do tabaco como un factor de risco a desenvolver esquizofrenia, segundo a cal, consumir tabaco podería aumentar o risco a desenvolver o trastorno; e 3) a hipótese de factores de risco comúns, segundo a cal, os mesmos factores de

USC DE COMPOSTELA

risco xenéticos e non xenéticos poderían incrementar de forma independente o risco a padecer tanto esquizofrenia como tabaquismo.

1.2.3.3.1. Hipótese da automedicación

Tradicionalmente, dentro das distintas hipóteses propostas, a hipótese máis xeneralizada para explicar a relación observada entre o tabaquismo e a esquizofrenia foi a da automedicación (Hughes et al., 1986; Khantzian, 1997). A defensa desta hipótese baséase en estudos que apoian que a nicotina reduce os síntomas da esquizofrenia e os efectos adversos da medicación antipsicótica (Kumari & Postma, 2005). Por un lado, observouse que fumar cigarros de xeito agudo podería reducir de forma transitoria os síntomas negativos da esquizofrenia, aínda que non está claro se este efecto é debido directamente á nicotina (R. C. Smith et al., 2002). Ademais, varios estudos suxeriron que a nicotina mellora, en pacientes con esquizofrenia, déficits cognitivos (D'Souza & Markou, 2012; George et al., 2002; Jacobsen et al., 2004; Sacco et al., 2005) e síntomas neurolóxicos menores como aqueles asociados ao filtrado sensorial auditivo e aos movementos oculares (Adler et al., 1993; Dépatie et al., 2002). Por outro lado, a nicotina parece ter un efecto na taxa de metabolización dos fármacos antipsicóticos, diminuíndo a súa concentración en plasma (Desai et al., 2001). Isto podería reducir os efectos adversos destes medicamentos, como as deficiencias cognitivas (Levin et al., 1996) ou o parkinsonismo (Decina et al., 1990; Goff et al., 1992).

En contraposición á hipótese da automedicación, numerosos estudos evidencian que fumar non só non mellora o déficit cognitivo propio da esquizofrenia e do seu tratamento, senón que o pode agravar nunha relación dependente de dose (Coustals et al., 2020; A. C. Reed et al., 2016; Ringin et al., 2022; Sánchez-Gutiérrez et al., 2018; Vermeulen et al., 2018). Un estudo realizado nunha mostra española demostrou que os/as pacientes cun primeiro episodio psicótico presentan unha peor función cognitiva que os controis independentemente do consumo de tabaco, tendo tanto os/as pacientes fumadores/as como non fumadores/as unha función cognitiva similar (Sánchez-Gutiérrez et al., 2018). O estudo de Vermeulen et al. (2018) mostrou que o tabaquismo está incluso asociado a un peor rendemento cognitivo tanto nos/as pacientes con psicose non afectiva como nos individuos sans e que este deterioro é máis pronunciado a medida que aumenta o consumo de tabaco. Ademais, deixar de fumar parece ter un impacto positivo no rendemento cognitivo dos/as pacientes.

A ampla asimilación da hipótese da automedicación debeuse en parte á intromisión da industria tabaqueira. Un estudo revelador de 2008 que revisou documentos da industria tabaqueira previamente secretos mostrou como esta industria supervisou e financiou

investigacións que apoiasen a idea de que as persoas con esquizofrenia son menos susceptibles aos danos do tabaco e que realmente se benefician do seu consumo para contrarrestar os síntomas do trastorno (Prochaska et al., 2008).

Esta influencia da industria tabaqueira levou a unha cultura de normalización do tabaquismo nos contextos psiquiátricos (Hall & Prochaska, 2009; Prochaska et al., 2017). Esta normalización do tabaquismo tradúcese en varios obstáculos que impiden que as persoas con esquizofrenia deixen de fumar (Lum et al., 2018). Por un lado, os/as pacientes con esquizofrenia parecen ser menos conscientes dos riscos para a saúde asociados ao tabaquismo, estando menos motivados/as a deixar de fumar (Kelly et al., 2012). As súas familias tamén adoitan ser tolerantes ao tabaquismo, o que supón unha barreira para que deixen de fumar (Aschbrenner et al., 2017; Missen et al., 2013). Por outro lado, os/as profesionais da saúde mental parecen ter a percepción de que o tabaquismo é beneficioso para os/as pacientes e que eles/as non están interesados ou son incapaces de deixar de fumar (Sheals et al., 2016). Ademais, adoitan ser reticentes á hora de prescribir tratamentos específicos para deixar de fumar en pacientes con trastornos mentais, a pesar de que se demostrou a súa eficacia nestes/as pacientes (G. M. J. Taylor et al., 2020).

Estas barreiras coas que se atopan os/as pacientes con trastornos mentais para deixar de fumar chocan coa evidencia clara de que as intervencións para deixar de fumar en persoas con esquizofrenia son efectivas e non teñen efectos prexudiciais na saúde mental dos/as pacientes con esquizofrenia (Stubbs et al., 2015; Tsoi et al., 2010). Unha meta-análise de estudos lonxitudinais mesmo mostrou que aquelas persoas que deixan de fumar, tanto persoas con trastornos psiquiátricos como poboación xeral, teñen menores niveis de ansiedade, depresión e estrés en comparación con aquelas que continúan fumando (G. Taylor et al., 2014). Outro estudo tamén confirmou que reducir o tabaquismo non está asociado a cambios na saúde mental en comparación con continuar fumando (G. Taylor et al., 2015). Estes datos subliñan a importancia de superar as barreiras existentes para que as persoas con esquizofrenia poidan acceder a tratamentos efectivos para deixar de fumar, o cal podería mellorar a súa calidade de vida.

1.2.3.3.2. Hipótese do tabaco como factor de risco a esquizofrenia

A evidencia acumulada en diversos estudos epidemiolóxicos apoia a idea de que o tabaco actúa como un factor de risco para o desenvolvemento de esquizofrenia (Alderson & Lawrie, 2015; Firth et al., 2020, 2023; Gage & Munafò, 2015a, 2015b; Laursen & McGrath, 2016; MacCabe, 2018; G. M. J. Taylor & Munafò, 2019). Por un lado, estes estudos indicaron

USC UNIVERSIDADE DE SANTIAGO

que iniciar o consumo de tabaco a idades temperás aumenta o risco de desenvolver psicose, apoiando, deste xeito, un efecto causal do tabaco sobre a esquizofrenia (Koyanagi et al., 2016; McGrath et al., 2016; Mustonen et al., 2018; Weiser et al., 2004). As persoas que comezan a fumar diariamente antes dos 15 anos mostran un maior risco de síntomas psicóticos (Koyanagi et al., 2016). Esta asociación entre a idade de inicio do consumo de tabaco e o risco a psicose persiste despois de excluír todas aquelas persoas consumidoras actuais de cannabis ou cunha historia de trastorno polo uso de cannabis (CUD, do inglés *cannabis use disorder*), descartando que esta asociación estea mediada polo consumo de cannabis (McGrath et al., 2016).

Estudos prospectivos tamén contribuíron a esta cuestión, atopando que o consumo de tabaco é anterior ao desenvolvemento da esquizofrenia. Estímase que, de media, o inicio do consumo de tabaco precede ao inicio da psicose en 5,3 anos (Myles et al., 2012). Varios estudos mostraron que un tabaquismo diario aos 15-16 anos de idade está asociado cunha posterior psicose (Gage et al., 2014; Mustonen et al., 2018). Unha meta-análise baseada nos cinco estudos prospectivos publicados ata a data evidenciou que as persoas fumadoras teñen case o dobre de risco posterior a desenvolver esquizofrenia en comparación coas non fumadoras (Hunter et al., 2020). Ademais, na maior parte dos estudos analizados nesta meta-análise, atopouse unha relación dependente de dose, cun maior risco posterior a esquizofrenia a medida que a cantidade de tabaco consumido aumenta. Un destes estudos prospectivos realizado por Kendler et al. (2015), baseado en máis de 1.600.000 persoas, descartou que factores como o nivel socioeconómico ou o consumo doutras substancias sexan responsables deste maior risco posterior a esquizofrenia das persoas fumadoras. Tamén se descartou que o consumo de cannabis estea explicando o efecto causal do tabaco sobre a psicose, atopando que a asociación entre o consumo de tabaco e a posterior psicose permanece significativa despois de axustar polo consumo de cannabis (Gage et al., 2014; M. King et al., 2021; Mustonen et al., 2018). Porén, factores xenéticos comúns si que parecen explicar parte da asociación entre o tabaquismo e a posterior esquizofrenia (Kendler et al., 2015).

Polo tanto, os estudos epidemiolóxicos ofrecen forte evidencia de que o tabaquismo en idades temperás aumenta o risco posterior de esquizofrenia. Este resultado non é consistente coa hipótese da automedicación, segundo a cal, agardariamos unha asociación entre psicose e tabaquismo principalmente naquelas persoas que comezan a fumar máis tarde, posteriormente á idade habitual de inicio dos síntomas da esquizofrenia (Alderson & Lawrie, 2015). Porén, este resultado dos estudos epidemiolóxicos podería explicarse se aspectos preclínicos da esquizofrenia desen lugar a un maior risco a tabaquismo. Esta posibilidade non é apoiada polos estudos epidemiolóxicos, segundo os cales, o tabaquismo iníciase, principalmente, antes da fase

USC UNIVERSIDADE DE SANTIAGO

prodrómica da esquizofrenia ou, polo menos, durante este período (Gogos et al., 2019; Kendler et al., 2015). Ademais, a asociación existente entre un tabaquismo grave na adolescencia e un risco posterior a psicose non se explica por experiencias psicóticas previas ao diagnóstico de psicose (Mustonen et al., 2018).

Pola contra, estes resultados dos estudos epidemiolóxicos son consistentes cun efecto causal do tabaco sobre a esquizofrenia. Non obstante, a identificación desta asociación en estudos epidemiolóxicos non proba causalidade, xa que podería ser debida a factores de confusión non considerados. Un estudo lonxitudinal evidenciou o posible efecto de factores de confusión, mostrando que o axuste por unha serie de potenciais factores de confusión reduciu a asociación entre o consumo de tabaco a unha idade temperá, sen consumo de cannabis, e experiencias psicóticas posteriores en aproximadamente un 60% (Jones et al., 2018). Os potenciais factores de confusión que se examinaron foron o sexo, unha historia familiar de esquizofrenia, unha historia familiar de consumo de substancias, un tabaquismo parental durante o embarazo, a educación da nai, a clase social parental, o cociente intelectual, un trauma durante a infancia, o acoso escolar, problemas emocionais ou condutuais e o consumo de alcol. Ademais, tal e como MacCabe (2018) xustifica, esta asociación entre un consumo de tabaco a idades temperás e un maior risco posterior a desenvolver psicose non só é consistente cun efecto causal do tabaco senón que tamén sería compatible cunha traxectoria de desenvolvemento atípica debida a factores de risco do neurodesenvolvemento, xenéticos ou psicosociais que dá lugar tanto a un consumo de tabaco na adolescencia como, posteriormente, a un maior risco a psicose na idade adulta. Polo tanto, aínda que os estudos epidemiolóxicos proporcionan forte evidencia de que a asociación entre tabaquismo e esquizofrenia sexa debida, polo menos parcialmente, a un efecto causal do tabaquismo sobre a esquizofrenia, debemos ter en conta estas limitacións.

1.2.3.3.3. Combinación das hipóteses para explicar a asociación entre tabaquismo e esquizofrenia

Ademais das dúas hipóteses explicadas previamente, a hipótese da automedicación e a hipótese do tabaco como un factor de risco para a esquizofrenia, existiría tamén a terceira hipótese duns factores de risco comúns que aumentan de forma independente o risco a esquizofrenia e tabaquismo. Estes factores de risco comúns poden ser tanto xenéticos, como se explicará no apartado 1.5.2.1, como non xenéticos. De feito, identificáronse factores ambientais de risco comúns a tabaquismo e a esquizofrenia, como un ambiente familiar desestruturado, unha clase social baixa ou experiencias traumáticas na infancia (MacCabe, 2018).

USC UNIVERSIDADE DE SANTIAGO

As tres hipóteses propostas para explicar a asociación entre tabaquismo e esquizofrenia non son excluíntes. Polo tanto, a relación entre tabaquismo e esquizofrenia podería ser debida a unha combinación das tres hipóteses propostas, con variacións segundo o paciente (Sagud et al., 2019). Esta hipótese concorda cun estudo realizado a partir da enquisa de saúde internacional da Organización Mundial da Saúde, o cal atopou unha asociación bidireccional entre psicose e tabaquismo, estando asociado o consumo de tabaco cun maior risco posterior de experiencias psicóticas e viceversa (Degenhardt, Saha, et al., 2018).

Neste senso, propúxose unha hipótese que unifica as tres hipóteses propostas, na cal unha base xenética común entre a esquizofrenia e o consumo de substancias convive cun efecto causal do consumo de substancias sobre o risco de esquizofrenia e viceversa (Khokhar et al., 2018). Segundo esta hipótese, factores xenéticos que dan lugar a un sistema de recompensa anómalo contribúen tanto ao risco a psicose como ao consumo de substancias, dando lugar a unha maior vulnerabilidade ao consumo de substancias dentro dos/as pacientes con esquizofrenia. Como esta vulnerabilidade ao consumo de substancias adoita aparecer antes da aparición de síntomas psicóticos, o propio consumo de substancias durante a adolescencia podería ser un factor de risco adicional á aparición de psicose. Ademais, a propia neuropatoloxía da esquizofrenia tamén afectaría ao sistema de recompensa neuronal, dando lugar a un maior risco a comezar a usar substancias e a manter este consumo, a pesar das súas consecuencias negativas (Chambers et al., 2001).

1.3. PRINCIPAIS MÉTODOS XENÓMICOS PARA O ESTUDO DE TRAZOS COMPLEXOS

Nos últimos anos, houbo importantes avances no estudo dos trazos de consumo de substancias e da súa asociación cos trastornos psiquiátricos grazas á información acadada con métodos xenómicos. A continuación, descríbense as principais ferramentas xenómicas empregadas para o estudo destes e outros trazos complexos.

1.3.1. Estudos de asociación do xenoma completo

As variantes xenéticas comúns que confiren un maior risco aos trazos complexos identifícanse, principalmente, mediante estudos de asociación do xenoma completo (GWAS, do inglés *genome-wide association study*). Os GWAS analizan simultaneamente o efecto, nun determinado trazo, de millóns de SNP ao longo de todo o xenoma (Andreassen et al., 2023; Sullivan & Geschwind, 2019). Estes estudos examinan se existe unha asociación entre cada variante xenética e o trazo. Debido ao gran número de SNP analizados nos GWAS, é necesario realizar unha corrección por múltiples comparacións para diminuír a probabilidade de falsos

positivos. En concreto, estableceuse arbitrariamente que se o nivel de significación é inferior a $5x10^{-8}$ (o correspondente a corrixir por 1 millón de comparacións) rexéitase a hipótese nula de non asociación e acéptase que existe evidencia de asociación.

Nos GWAS, para evitar xenotipar todos os SNP do xenoma, escóllense uns SNP representativos que capturan gran parte da variación xenética común nunha poboación (Sullivan & Geschwind, 2019). A proporción de variación xenética común detectada depende do *array* empregado para xenotipar os SNP. A selección de SNP xenotipados baséase no desequilibrio de ligamento (LD, do inglés *linkage disequilibrium*) (Dehghan, 2018). O LD é a asociación non aleatoria entre dúas variantes xenéticas próximas no xenoma nunha poboación determinada. Este LD entre dúas variantes describe o grao en que estas dúas variantes xenéticas se herdan de forma conxunta na poboación e dependerá do nivel de recombinación entre elas durante a meiose. Deste xeito, o xenoma organízase en bloques haplotípicos, rexións do xenoma en LD que se herdan en bloque, os cales son diferentes segundo a ascendencia do individuo. Os SNP xenotipados nos GWAS funcionan como marcadores de cada un destes bloques e proporcionan información sobre o xenotipo dos SNP incluídos nesa rexión. Desta forma, os GWAS informan principalmente sobre loci xenómicos asociados a un fenotipo, isto é, rexións con múltiples variantes xenéticas en LD asociadas co trazo.

A capacidade dun GWAS para identificar SNP asociados a un trazo depende, en parte, da arquitectura xenética do trazo a nivel de variación común (Andreassen et al., 2023). En concreto, depende da súa polixenicidade, o número de variantes xenéticas comúns que inflúen no fenotipo; da súa herdabilidade baseada en SNP, que se refire á proporción de varianza fenotípica explicada por variantes xenéticas comúns; e da distribución do tamaño de efecto das variantes xenéticas. Ademais, está condicionada polo tamaño da mostra empregada no GWAS, sendo necesarias mostras de gran tamaño para alcanzar a potencia estatística necesaria para poder detectar variantes xenéticas asociadas ao trazo cun efecto mínimo. As colaboracións internacionais en grandes consorcios foron esenciais para acadar estes tamaños de mostra. O *Psychiatric Genomics Consortium* (PGC) e o *GWAS & Sequencing Consortium of Alcohol and Nicotine use* (GSCAN) destacan por proporcionar mostras amplas para trastornos psiquiátricos (Sullivan et al., 2018) e para trazos de consumo de tabaco e alcol (G. R. B. Saunders et al., 2022), respectivamente. Datos de cohortes nacionais tamén contribuíron a este reto de obter tamaños de mostra elevados. Dentro das distintas cohortes empregadas nos GWAS, destaca o *UK Biobank* por ser das primeiras cun elevado tamaño e alcance (Bycroft et al., 2018). Esta cohorte prospectiva está formada por ao redor de 500.000 individuos do Reino Unido cunha idade entre 40 e 69 anos. Para estes participantes, recolléronse datos

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

fenotípicos relacionados coa saúde, incluíndo medidas biolóxicas, indicadores do estilo de vida, biomarcadores no sangue e na urina e datos de imaxe, e datos xenotípicos de todo o xenoma, os cales se puxeron a disposición da comunidade científica para realizar as súas investigacións. Actualmente, existen tamén datos para outras cohortes máis amplas como *23andMe* ou *Million Veteran Program* (Abbasi, 2017; Gaziano et al., 2016).

O aumento do tamaño das mostras analizadas nos GWAS permitiu identificar un número cada vez maior de variantes xenéticas comúns asociadas cos trastornos psiquiátricos e trazos relacionados como os trazos de consumo de substancias ou os trazos condutuais. Porén, aínda non se identificaron a maior parte das variantes xenéticas de risco (Andreassen et al., 2023). Por exemplo, o GWAS para esquizofrenia máis potente ata a data, o cal analizou o xenotipo de 76.755 casos e 243.649 controis, identificou 287 loci significativos a nivel xenómico, os cales representan menos do 10% da herdabilidade a nivel de SNP da esquizofrenia (Schizophrenia Working Group of the Psychiatric Genomics Consortium et al., 2022).

A diferenza existente entre a herdabilidade explicada polos loci significativos detectados en GWAS e a herdabilidade estimada en estudos de xemelgos recibe o nome de "herdabilidade perdida" (Wray et al., 2014; Young, 2019). Por un lado, esta diferenza parece ser debida a que gran parte da variación común que afecta aos trazos complexos non foi aínda identificada nos GWAS. Isto, á súa vez, pode ser explicado por problemas de potencia dos GWAS que levan a non detectar variantes comúns cun efecto moi pequeno e pola existencia de variación común que non é identificada polos *arrays*, tanto SNP que non son incluídos no *array* nin están en LD cun SNP do *array* como outros tipos de variantes comúns que non son incluídas. Por outro lado, a "herdabilidade perdida" tamén se debe ao efecto de variantes raras que non son detectadas nos GWAS. Finalmente, tamén se considera que está contribuíndo a esta diferenza unha sobreestimación da herdabilidade en xemelgos debido a un ambiente compartido, interaccións xene-ambiente ou interaccións xénicas.

Os resultados xenómicos obtidos nos GWAS recóllense, principalmente, en forma de estatísticas resumo. Estas estatísticas resumo inclúen información sobre os SNP analizados, sobre o seu efecto no trazo e sobre o nivel de significación da súa asociación co trazo. A distribución ao público das estatísticas resumo dos GWAS é esencial xa que a maior parte das ferramentas xenómicas empregadas actualmente baséanse na súa análise. Consorcios como o PGC ou o GSCAN poñen a disposición pública as estatísticas resumo dos seus GWAS. Ademais, a iniciativa *NHGRI-EBI GWAS Catalog* (Sollis et al., 2023) permite obter as estatísticas resumo de máis de 65.000 GWAS publicados (datos consultados o 14/06/2024).

1.3.2. Estimas baseadas en *LD score regression*

O método *LD score regression* (LDSC) é amplamente empregado nos estudos xenómicos debido á súa versatilidade, permitindo estimar, a partir das estatísticas resumo de GWAS, a herdabilidade baseada en SNP dun trazo, a correlación xenética entre trazos, a cal representa a proporción da varianza compartida entre dous trazos que é debida a factores xenéticos, e o grao de solapamento das mostras de distintos GWAS (B. Bulik-Sullivan et al., 2015). Este método baséase na estimación da puntuación de LD dos SNP, a cal reflicte o LD de cada SNP con outros SNP da súa rexión xenómica (B. K. Bulik-Sullivan & Neale, 2015). A idea central do método é que, dado que o tamaño de efecto dun SNP estimado no GWAS incorpora os efectos dos SNP en LD con este e dado que un SNP cunha maior puntuación de LD terá unha maior probabilidade de estar en LD cunha variante causal asociada co trazo, o nivel de asociación dos SNP no GWAS será proporcional á puntuación de LD dos mesmos. A partir da pendente da regresión entre estas dúas magnitudes poderase estimar a herdabilidade do trazo. Ademais, desviacións no intercepto desta regresión serán principalmente debidas a un solapamento entre as mostras de ambos GWAS analizados, permitindo estimar o grao de solapamento entre eles.

Este método tamén permite calcular a correlación xenética do tamaño de efecto dos SNP para dous trazos, sen un nesgo no caso de solapamento de mostras (B. Bulik-Sullivan et al., 2015). Neste caso, a variable dependente da regresión será o produto do tamaño de efecto dos SNP sobre os dous trazos estimado nos GWAS. De forma similar, esta variable será proporcional á puntuación de LD dos SNP e a pendente desta regresión representará a covarianza xenética entre ambos trazos. A partir desta covarianza xenética, poderase estimar a correlación xenética entre os trazos.

Aínda que a correlación xenética é unha medida informativa sobre a susceptibilidade xenética compartida entre dous trazos, non captura todas as dimensións do solapamento xenético. A existencia dunha mestura de variantes xenéticas compartidas entre dous trazos cunha dirección de efecto concordante, é dicir, cun efecto na mesma dirección para ambos trazos, e cunha dirección de efecto discordante, é dicir, cun efecto en direccións opostas para cada trazo, enmascara parcialmente o solapamento xenético real entre os trazos (Hindley et al., 2022). Deste xeito, un valor similar de correlación xenética a nivel de todo o xenoma pode deberse tanto a un solapamento xenético elevado cunha mestura de SNP concordantes e discordantes como a un solapamento xenético menor cunha correlación xenética elevada dos SNP compartidos. Para poder caracterizar o patrón de solapamento xenético entre dous trazos de forma máis detallada, incluíndo o número de variantes causais compartidas e a correlación

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

xenética destas variantes xenéticas compartidas, desenvolveuse a ferramenta *bivariate causal mixture model* (MiXeR) (Frei et al., 2019).

1.3.3. Puntuacións de risco polixénico

Unha metodoloxía que permite capturar o efecto conxunto de SNP distribuídos por todo o xenoma a partir das estimas de GWAS son as puntuacións de risco polixénico (PRS, do inglés *polygenic risk score*, tamén coñecido como *polygenic score*), as cales representan a susceptibilidade xenética dun individuo a un trazo (Lewis & Vassos, 2020; Martin et al., 2019; Wray et al., 2014). Esta metodoloxía permite avaliar o efecto conxunto de SNP, tanto daqueles SNP significativos no GWAS orixinal como daqueles SNP que non alcanzasen a significación por problemas de potencia. Para estimar a PRS, en primeiro lugar, selecciónanse SNP en base ao seu LD e á súa asociación co trazo de interese no GWAS orixinal, o cal se denomina mostra de descubrimento. Estes SNP son posteriormente empregados para construír a PRS nunha mostra independente, denominada mostra diana, para a cal coñecemos os datos xenotípicos a nivel individual. Así, estímase a PRS de cada un dos individuos desta mostra diana como a suma do número de alelos de risco que porta ese individuo para cada SNP ponderado polo seu efecto no trazo estimado no GWAS. A continuación, pódese analizar a asociación desa PRS e un trazo de interese para o cal existan datos na mostra diana. Este trazo de interese pode ser o mesmo que o analizado na mostra de descubrimento ou outro. Un resumo do procedemento empregado nos estudos de PRS recóllese na Figura 1.

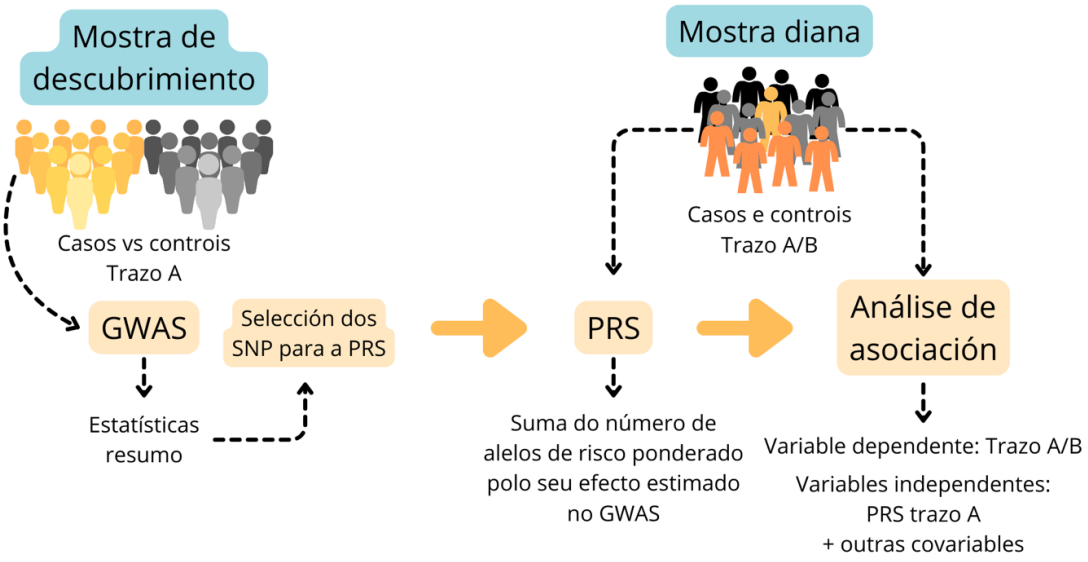


Figura 1. Representación gráfica do procedemento empregado nos estudos de PRS. A PRS calcúlase a partir dun GWAS, denominado mostra de descubrimento, que analiza a asociación dos SNP cun trazo A. A partir das estatísticas resumo dese GWAS selecciónanse os SNP que se van incorporar no cálculo da PRS. Se se emprega o método tradicional de *clumping* e *thresholding* selecciónanse SNP independentes a distintos limiares de significación na mostra de descubrimento. A PRS calcúlase para unha mostra diana para a cal

coñecemos os datos xenotípicos a nivel individual e a cal foi fenotipada para un trazo de interese, que pode ser o mesmo que o da mostra de descubrimento (trazo A) ou outro (trazo B). A PRS calcúlase como a suma do número de alelos de risco dos SNP seleccionados ponderados polo seu efecto no trazo A estimado na mostra de descubrimento. Finalmente, pódese analizar a asociación da PRS para o trazo A co trazo A ou B na mostra diana.

Existen múltiples métodos para estimar PRS, os cales difiren no criterio para seleccionar os SNP finalmente empregados na PRS e para ponderar cada un deses SNP na PRS final (Ni et al., 2021). O método tradicional de cálculo de PRS inclúe os pasos de *clumping* e *thresholding*. No *clumping*, selecciónanse SNP independentes segundo o seu LD e a súa asociación co trazo na mostra de descubrimento. O *thresholding* fai referencia a que a PRS se calcula a partir dun subconxunto dos SNP seleccionados cunha significación na mostra de descubrimento menor a un ou máis limiares (Purcell et al., 2009). Nos últimos anos, desenvolvéronse metodoloxías alternativas que permiten reter todas as variantes xenéticas na PRS, co fin de maximizar a información capturada no modelo (Ni et al., 2021; Pain et al., 2021). A pesar de empregar todas as variantes xenéticas, estes métodos teñen tamén en conta o LD entre as variantes, para evitar unha sobreestimación debida ao efecto non independente de variantes en LD, e os diferentes niveis de polixenicidade dos trazos. Con este fin, estes métodos modelan o LD entre as variantes xenéticas analizadas e, xeralmente, aplican un parámetro de contracción ao tamaño de efecto dos SNP para reducir o sobreaxuste e permitir diferentes niveis de polixenicidade. Dentro destes métodos atópanse os métodos bayesianos, os cales infiren a distribución do efecto dos SNP na mostra de descubrimento e, a continuación, contraen o tamaño de efecto dos SNP para que se axuste a esa distribución. Algúns destes métodos bayesianos son, por exemplo, *EB-PRS* (S. Song et al., 2020), *LDpred2* (Privé et al., 2020), *SBayesR* (Lloyd-Jones et al., 2019) ou *PRS-CS* (Ge et al., 2019). Demostrouse que cada un destes métodos novos incrementa a predición das PRS en comparación co método tradicional de *clumping* e *thresholding* (Ni et al., 2021; Pain et al., 2021). Ademais, recentemente, desenvolvéronse métodos que, aproveitando a alta pleiotropía entre trazos, empregan a información contida en GWAS de trazos correlacionados para mellorar a fiabilidade preditiva das PRS. Por exemplo, *SMTPred* permite combinar diferentes PRS correlacionadas nunha mesma PRS (Maier et al., 2018). Outras metodoloxías como *multi-trait analysis of GWAS* (MTAG) permite obter novas estimas do efecto dos SNP para un trazo de interese tendo en conta o seu efecto noutros trazos relacionados, aumentando, deste xeito, a potencia estatística para detectar asociacións xenéticas co trazo (Turley et al., 2018). Este novo GWAS obtido con MTAG pode ser logo empregado para estimar PRS.

USC DE SANTIAGO DE COMPOSTELA

As utilidades dos estudos de PRS son múltiples. Por un lado, permiten obter evidencia dun efecto polixénico nun trazo cando non se detectan SNP significativos no GWAS, acadando unha asociación significativa entre a PRS e o trazo de interese. A asociación entre a PRS para un trazo e un trazo diferente tamén permite demostrar a existencia dun solapamento xenético entre ambos trazos. Ademais, as PRS constitúen un primeiro paso para desenvolver modelos de predición de risco, aínda que as PRS actuais en psiquiatría están lonxe de ter unha fiabilidade en clínica a nivel de predición (Murray et al., 2021). Porén, son prometedoras á hora de estratificar a poboación en grupos de risco, o cal permitiría prever e tratar de forma precoz o trastorno naquelas persoas cun maior risco xenético, axudaría a establecer un diagnóstico naquelas persoas con síntomas do trastorno e contribuiría a unha elección personalizada do tratamento (Lewis & Vassos, 2020).

1.3.4. Aleatorización mendeliana

Tanto a correlación xenética como as PRS proporcionan información sobre o nivel de solapamento xenético entre dous trazos, mais isto non pode ser tomado como evidencia dunha relación de causalidade entre eles. Para abordar esta cuestión, xurdiron os métodos de aleatorización mendeliana (MR, do inglés *Mendelian randomization*), os cales empregan a variación xenética para avaliar a relación causal entre un trazo de exposición e un trazo resultante (Davies et al., 2018; Katikireddi et al., 2018; Treur et al., 2021). Estes estudos empregan SNP asociados ao trazo de exposición como variables instrumentais para investigar o posible efecto causal do trazo de exposición sobre o trazo resultante. En concreto, a asociación entre o trazo resultante e a variante xenética que é indicadora de exposición ao posible factor de risco permítenos estimar a magnitude da relación causal entre os dous trazos.

A MR foi desenvolvida para superar algunhas das principais limitacións dos estudos observacionais tradicionais á hora de establecer relacións causais (Davies et al., 2018; Katikireddi et al., 2018). Os estudos observacionais lonxitudinais poden atopar certa evidencia de causalidade se un trazo aumenta o risco posterior dun segundo trazo. Porén, factores de confusión que incrementan o risco a ambos trazos poderían dar lugar a unha evidencia nesgada. Teoricamente, se todos os factores de confusión fosen medidos de forma precisa, os estudos observacionais poderían establecer o efecto dun factor de risco de forma fiable. Mais na práctica, é imposible identificar todos os posibles factores de confusión. Ademais, é difícil que os estudos observacionais poidan distinguir a causalidade dunha causalidade inversa, onde o trazo resultante ou aspectos preclínicos do mesmo teñen un efecto no trazo de exposición. Finalmente, os estudos observacionais poden estar afectados polo nesgo de colisión (en inglés *collider bias*), segundo o cal pode aparecer unha correlación espuria entre dúas variables que

USC DE SANTIAGO DE COMPOSTELA

son independentes ao estratificar a mostra por unha terceira variable asociada a ambas. Por exemplo, unha diferente participación no estudo segundo se o individuo presenta os dous trazos analizados pode introducir un nesgo na asociación entre ambos trazos.

A MR emprega variantes xenéticas como indicadores de exposición ao posible factor de risco (Richmond & Smith, 2022). Estas variantes xenéticas, ao herdarse de forma aleatoria e, en principio, independentemente de factores de confusión, e ao ser fixas no individuo dende a fecundación, van ser máis robustas aos problemas descritos anteriormente. Os SNP empregados nos estudos de MR deben cumprir tres condicións principais para ser unhas variables instrumentais válidas: 1) estar asociados ao trazo de exposición de forma robusta; 2) ser independentes de factores de confusión que afecten á relación entre ambos trazos; e 3) non afectar directamente ao trazo resultante máis que a través dun efecto causal mediado polo trazo de exposición (Figura 2) (Davies et al., 2018; Katikireddi et al., 2018; Treur et al., 2021).

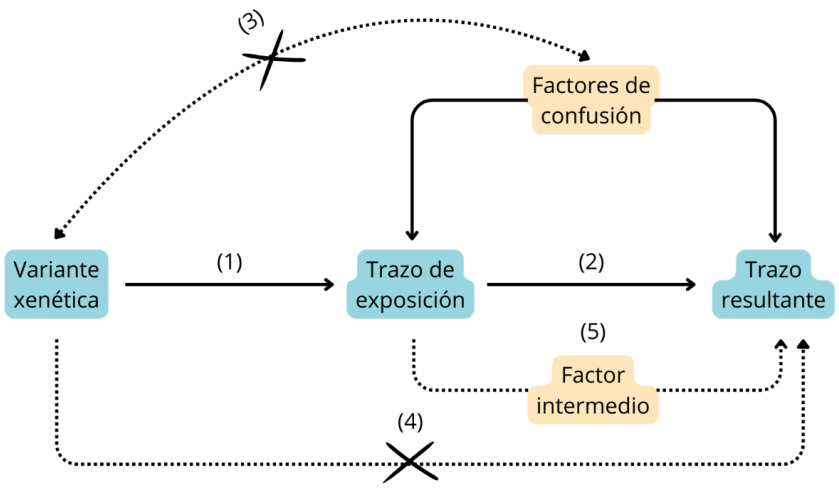


Figura 2. Representación gráfica das condicións necesarias nos estudos de MR. (1): A variable instrumental debe estar asociada de forma robusta ao trazo de exposición. (2): Se existe un efecto causal do trazo de exposición sobre o trazo resultante, a variante xenética estará tamén asociada ao trazo resultante (pleiotropía vertical). (3): A variable instrumental debe ser independente de factores de confusión. (4): A variable instrumental só pode estar asociada ao trazo resultante a través dunha vía que involucre ao trazo de exposición, non de forma independente (pleiotropía horizontal). (5): A variable instrumental pode estar asociada co trazo resultante a través doutros factores intermedios na mesma vía causal que o trazo de exposición.

Os GWAS, ao proporcionar información da asociación de millóns de SNP cun trazo de interese, facilitaron os estudos de MR. As metodoloxías de MR empregadas actualmente estiman o efecto causal do trazo de exposición sobre o trazo resultante tendo en conta a asociación de múltiples variantes xenéticas recollida nas estatísticas resumo de GWAS (Boehm & Zhou, 2022). Deste xeito, incrementouse a potencia estatística das análises de MR. Estes

USC DE COMPOSTELA

métodos de MR baseados nas estatísticas resumo de GWAS que empregan múltiples variables instrumentais diferéncianse no criterio á hora de elixir o subconxunto de SNP que se emprega na análise, na forma de modelar e controlar a pleiotropía horizontal e no procedemento á hora de estimar o efecto causal.

A pesar de representar un avance na avaliación de relacións de causalidade, como non é posible verificar totalmente que non se estea producindo ningunha violación das condicións dos estudos de MR expostas anteriormente, as conclusións ás que se poden chegar con este método son limitadas (Katikireddi et al., 2018). Por un lado, non se pode descartar a influencia de factores de confusión, aínda que se demostrase unha independencia estatística das variables instrumentais empregadas e uns determinados factores de confusión (Hartwig et al., 2018). Ademais, pode producirse un nesgo de colisión ou un emparellamento non aleatorio en base ao fenotipo, denominado apareamento orientado, que tamén producirían un nesgo na estimación da MR. Por outro lado, as variantes xenéticas elixidas pode que non afecten ao trazo resultante a través de vías nas cales participa o trazo de exposición, o cal é coñecido como pleiotropía vertical (tamén denominada pleiotropía mediada), senón que poderían afectar ao trazo resultante de forma independente ao trazo de exposición, o cal é coñecido como pleitropía horizontal (tamén denominada pleiotropía verdadeira ou pleiotropía biolóxica) (Hemani et al., 2018). Esta situación é altamente probable tendo en conta a alta pleiotropía existente entre trastornos psiquiátricos e trazos relacionados, afectando moitas variantes xenéticas a varios destes trazos (Polushina et al., 2021).

Recentemente, desenvolvéronse métodos de MR que permiten avaliar se os resultados son robustos mesmo no caso dunha posible violación dalgunha das condicións da MR como, por exemplo, asumindo a existencia de pleiotropía horizontal (Davies et al., 2018). Tamén se desenvolveron métodos que permiten detectar e eliminar da análise variantes cun posible efecto pleiotrópico e métodos que permiten modelar e corrixir por pleiotropía (Richmond & Smith, 2022). A combinación de diferentes métodos de MR, cada un coas súas propias características, permite explorar ata que punto a evidencia dun efecto causal é robusta ás asuncións da MR.

1.3.5. Modelos de ecuacións estruturais xenómicas

O alto grao de pleiotropía observado entre trastornos psiquiátricos e trazos relacionados suxire que están parcialmente afectados polas mesmas fontes de susceptibilidade xenética. Polo tanto, é esencial dispor de ferramentas que poidan modelar a complexa relación xenética entre estes trazos. Neste contexto, nace o método de modelos de ecuacións estruturais xenómicas (genomic SEM, do inglés *genomic structural equation modelling*) (Grotzinger et al., 2019).

USC UNIVERSIDADE DE SANTIAGO

Esta metodoloxía introduce a información acadada nos GWAS aos modelos de ecuacións estruturais. Os modelos de ecuacións estruturais permiten avaliar a validez de hipóteses previas sobre a relación entre diferentes variables empregando datos experimentais (Beran & Violato, 2010). Por un lado, permiten analizar o efecto dun ou varios trazos sobre outros, avaliando as hipotéticas relacións causais entre eles. Por outro lado, permiten avaliar a existencia de factores latentes, factores non medidos directamente que representan o risco común a diferentes trazos.

No caso de genomic SEM, modélase a estrutura de covarianza xenética entre diferentes trazos complexos empregando as estatísticas resumo de GWAS (Grotzinger et al., 2019). Esta metodoloxía permítenos avaliar a validez dun modelo para explicar a relación xenética entre os trazos analizados construído a partir de hipóteses previas. Neste modelo determínase o número de factores xenéticos latentes e a súa relación cos distintos trazos analizados. A continuación, avalíase se o modelo proposto se axusta ao patrón de covarianza xenética observado entre os distintos trazos. Ademais, pódense comparar diferentes estruturas de factores para elixir o modelo máis apropiado para explicar os datos xenéticos. Unha vez identificado un modelo, é posible incorporar SNP para identificar variantes xenéticas asociadas aos factores xenéticos latentes e, polo tanto, cun efecto nesas dimensións xerais de risco xenético a diferentes trazos. Polo tanto, genomic SEM é unha valiosa ferramenta para obter información sobre a arquitectura xenética compartida entre distintos fenotipos.

1.3.6. Identificación de mecanismos biolóxicos subxacentes aos trazos complexos

A identificación dos mecanismos biolóxicos subxacentes aos trazos complexos a partir dos descubrimentos a nivel xenómico é un proceso complexo. Os GWAS identifican variantes xenéticas asociadas aos trazos, mais é preciso pasar destas asociacións xenéticas a variantes xenéticas funcionais involucradas neses trazos (Schaid et al., 2018). É pouco probable que o SNP máis significativamente asociado a un trazo nun GWAS sexa a variante causal responsable do risco a ese trazo, senón que seguramente estará en LD coa variante causal real. Como os patróns de LD entre SNP poden ser complexos, é difícil determinar cal das múltiples variantes xenéticas en LD dentro dun locus xenómico asociado significativamente a un trazo é a verdadeira variante causal. Para abordar esta cuestión, propuxéronse unha serie de procedementos estatísticos de mapeado fino que axudan a priorizar aquelas variantes xenéticas candidatas a ser causais e responsables da asociación xenética observada co trazo.

Unha forma de priorizar aquelas variantes xenéticas con máis probabilidade de ser causais é identificar aquelas que poderían influír na función xénica, obtendo tamén, deste xeito, aqueles xenes posiblemente involucrados no trazo (Andreassen et al., 2023). Para as variantes xenéticas

que están localizadas nunha rexión codificante do xenoma, é relativamente sinxelo prever se teñen un impacto na estrutura e función da proteína codificada por ese locus (Wong et al., 2021). Porén, a maior parte das asociacións de GWAS detéctanse en rexións non codificantes (Jaffe et al., 2018; Maurano et al., 2012). Polo tanto, estas variantes poderían exercer o seu efecto a nivel fenotípico regulando a expresión dun ou máis xenes próximos ou localizados a unha gran distancia xenómica. Nos últimos tempos, obtivéronse datos de expresión xénica a nivel xenómico que poden ser empregados para anotar funcionalmente as variantes causais reguladoras da expresión e identificar os xenes sobre os cales exercen un efecto nun determinado tecido (Uffelmann & Posthuma, 2021). Integrando datos de expresión xénica con datos de xenotipado das mesmas mostras pódense identificar variantes xenéticas que inflúen na expresión xénica, coñecidas como loci de trazos cuantitativos de expresión (eQTL, do inglés *expression quantitative trait loci*). Con este fin naceu o Proxecto *Genotype-Tissue Expression* (GTEx), o cal emprega datos de secuenciación de ARN para 54 tecidos, incluíndo tecidos do cerebro, para identificar eQTL dos xenes codificantes de proteínas e dos xenes de ARN non codificante en diferentes tecidos (GTEx Consortium, 2020). Empregando estes datos, pódese avaliar se as variantes xenéticas identificadas en GWAS funcionan como eQTL nun determinado tecido. Para chegar a unha evidencia máis robusta de que un locus significativo nun GWAS exerce o seu efecto regulando a expresión xénica, pódese realizar unha análise de colocalización, a cal permite comparar a distribución de dous sinais de asociación, neste caso, o sinal de asociación do GWAS e do eQTL para identificar se son debidas á mesma variante causal tendo en conta o LD entre as variantes (Giambartolomei et al., 2014; B. Liu et al., 2019). A mellor opción para priorizar aquelas posibles variantes causais é non só ter en conta o seu posible efecto funcional sobre os xenes próximos, senón combinar esta información con datos de eQTL e outros datos xenómicos funcionais como a conformación e interacción da cromatina, tal e como propoñen recursos como *Open Targets Genetics* (Ghoussaini et al., 2021). Esta ferramenta permite identificar para cada variante xenética o xene sobre o que máis probablemente teña un efecto, proporcionando unha única puntuación da robustez desa asociación tendo en conta unha gran variedade de datos funcionais.

Para chegar a unha comprensión máis profunda dos mecanismos moleculares subxacentes aos trazos complexos, é interesante analizar funcionalmente os xenes identificados nun GWAS de forma integrada. Isto pode incluír examinar se están enriquecidos en conxuntos de xenes con determinadas funcións biolóxicas, en xenes diferencialmente expresados en determinados tipos celulares e tecidos ou en grupos de xenes cunha expresión correlacionada e, probablemente, cunha relación a nivel funcional (Andreassen et al., 2023; Sullivan & Geschwind, 2019). Desta

USC DE COMPOSTELA

forma, pódese obter información sobre os mecanismos subxacentes ás asociacións xenéticas observadas e os procesos moleculares que poderían estar alterados nos trazos.

1.3.7. Da xenómica aos estudos funcionais: o peixe cebra como modelo animal

Unha ferramenta valiosa para investigar o efecto funcional de xenes identificados mediante métodos xenómicos é o uso de modelos animais, os cales permiten obter datos experimentais en condicións específicas e controladas. O peixe cebra (*Danio rerio*) destaca como un modelo especialmente relevante, presentando unha serie de vantaxes a nivel experimental en comparación cos modelos mamíferos (McCammon & Sive, 2015). Dentro das súas vantaxes inclúese o seu pequeno tamaño, a súa fertilización externa, a súa facilidade para xerar un gran número de individuos, o seu rápido desenvolvemento embrionario, o seu xenoma totalmente secuenciado, as súas facilidades para a edición xénica e o seu potencial para a visualización óptica de todo o corpo e o cribado de alto rendemento (Kalueff et al., 2014; Stewart et al., 2014). Ademais, o peixe cebra presenta unha gran homoloxía a nivel fisiolóxico e xenético cos humanos, cun 70% dos xenes humanos tendo un ortólogo en peixe cebra, proporción que incrementa ao 82% no caso de xenes relacionados con enfermidades humanas (Howe et al., 2013). O peixe cebra, coa axuda da metodoloxía CRISPR-Cas (do inglés *clustered regularly interspaced short palindromic repeat sequences/CRISPR-associated protein*), tamén proporciona unha ferramenta eficaz para validar variantes xenéticas atopadas en GWAS analizando a perda de función ou mutación de xenes diana (L. M. Saunders et al., 2023; Wang et al., 2023). Esta metodoloxía foi empregada en peixe cebra, por exemplo, para confirmar o papel do xene *CHMP7* en ADHD (Dark et al., 2020).

De interese para esta tese, o peixe cebra tamén destaca como un modelo prometedor para o estudo de trastornos psiquiátricos e trazos relacionados como os trazos de consumo de substancias (Cabana-Domínguez et al., 2023; Fontana et al., 2018; Mathur & Guo, 2010; Stewart et al., 2015). O sistema nervioso humano está altamente conservado no peixe cebra, cunha arquitectura a nivel cerebral semellante e unha conservación de todos os tipos neuronais (McCammon & Sive, 2015; Wang et al., 2023). Procesos relacionados coa adicción a substancias, como a activación da vía da recompensa polo consumo de substancias, a busca compulsiva de substancias ou as neuroadaptacións provocadas pola exposición crónica a substancias, están conservados en peixe cebra (Müller et al., 2020; Stewart et al., 2015). Ademais, o alto grao de conservación dos sistemas endocannabinoide e opioide entre o peixe cebra e os humanos permite estudar mecanismos relacionados coa ansiedade e a recompensa (Demin et al., 2018). Nos últimos anos, o desenvolvemento dunha gran variedade de probas condutuais no peixe cebra permitiu analizar dende funcións motoras e sensoriais básicas a

USC DE SANTIAGO DE COMPOSTELA

comportamentos máis complexos asociados á cognición, ás emocións e á adicción a substancias (Cabana-Domínguez et al., 2023; Kalueff et al., 2014; Müller et al., 2020). Por último, cabe destacar que os xenes asociados con trastornos psiquiátricos superan a conservación xeral de xenes en peixe cebra (Franklin & Dwyer, 2021; Kasap et al., 2018). Deste xeito, o peixe cebra ofrece unha plataforma robusta para estudar os xenes relacionados con trastornos psiquiátricos e trazos de consumo de substancias.

1.4. APLICACIÓN DA METODOLOXÍA XENÓMICA AO ESTUDO DE TRAZOS DE CONSUMO DE SUBSTANCIAS

1.4.1. Estudos de asociación do xenoma completo para trazos de consumo de substancias

Os GWAS para trazos de consumo de substancias proporcionaron máis información sobre a arquitectura xenética destes trazos. Co aumento de tamaño das mostras analizadas, os GWAS detectaron un número cada vez maior de loci significativamente asociados aos trazos de consumo de substancias (Figura 3) (Gerring et al., 2024). Porén, este aumento no número de loci de risco identificados non foi uniforme para todos os trazos, xa que se analizaron mostras máis amplas para substancias legais como o tabaco e o alcol en comparación con substancias ilegais como o cannabis ou os opioides (Figura 3).

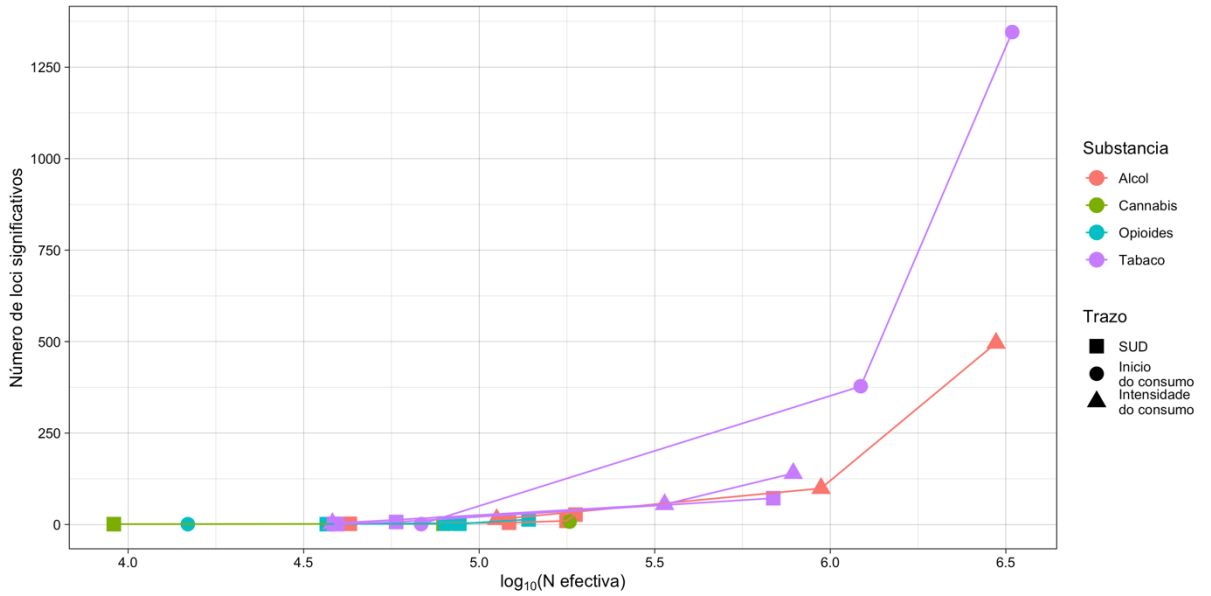


Figura 3. Número de loci significativamente asociados a trazos de consumo de alcol, cannabis, opioides e tabaco en función do tamaño de mostra efectivo do GWAS. Considéranse tanto comportamentos de consumo de substancias, en concreto, o inicio e a intensidade do consumo, como SUD. Os datos para SUD foron obtidos de Gerring & Thorp (2024). O tamaño de mostra efectivo para trazos dicotómicos foi calculado como 4/((1/N casos)+(1/N controis)).

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

O GWAS máis recente para comportamentos de consumo de tabaco e alcol atopou que a herdabilidade baseada en SNP para estes trazos variou dende o 4,0% para o número medio de bebidas alcólicas consumidas á semana ata o 8,0% para o inicio do consumo regular de tabaco e o número medio de cigarros fumados ao día (G. R. B. Saunders et al., 2022). A herdabilidade baseada en SNP para AUD estimouse nun 8,5% en europeos (Zhou et al., 2023). Unha herdabilidade baseada en SNP similar (8,6%) foi atopada para a dependencia a nicotina en europeos (Quach et al., 2020).

Os GWAS tamén permitiron identificar determinados xenes que teñen un elevado impacto na susceptibilidade xenética aos trazos de consumo de determinadas substancias (Prom-Wormley et al., 2017). Por un lado, xenes que codifican para enzimas implicados no metabolismo e procesado do alcol, como a agrupación de xenes do cromosoma 4 que codifica para as alcol deshidroxenases e o xene *ALDH2* que codifica para a aldehido deshidroxenase, están fortemente asociados ao consumo de alcol (Tawa et al., 2016). Estes enzimas están implicados na principal vía do metabolismo do alcol, na cal, nun primeiro paso, as alcol deshidroxenases catalizan a conversión do alcol en aceltaldehido e, nun segundo paso, a aldehido deshidroxenase oxida o acetaldehido a acetato (Edenberg, 2007). Por outro lado, os xenes que codifican para as subunidades dos receptores colinérxicos nicotínicos (nAchR) están fortemente asociados ao consumo de tabaco. Por exemplo, a agrupación de xenes do cromosoma 15 *CHRNA5-CHRNA3-CHRNB4*, a cal codifica para as subunidades de nAchR α5, α3 e β4, está fortemente asociada á intensidade do consumo de tabaco (Ware et al., 2011). A nicotina exerce o seu efecto no sistema nervioso central interactuando con estes nAchR presinápticos e influíndo, deste xeito, nunha variedade de funcións neuronais incluíndo a liberación de neurotransmisores (Marchi & Grilli, 2010). Os xenes de subunidades de nAchR codifican para 9 subunidades α (α2-α10) e 3 subunidades β (β2-β4) (Gotti et al., 2009). Cinco destas subunidades dispóñense en diferentes combinacións para orixinar diferentes subtipos de nAchR funcionais.

1.4.2. Susceptibilidade xenética compartida entre as distintas fases de consumo da mesma substancia

Os estudos xenómicos proporcionaron información sobre a susceptibilidade xenética compartida entre as distintas fases do consumo dunha mesma substancia, indicando que existen diferenzas na susceptibilidade xenética aos diferentes estadíos. A correlación xenética entre comportamentos de consumo de substancias e SUD para unha mesma substancia é moderada (Táboa 2). En consonancia coa súa correlación xenética, a intensidade do consumo de alcol e o consumo problemático de alcol presentaron diferenzas na súa asociación con outros trazos,

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

estando o primeiro asociado con condicións metabólicas e o segundo con trastornos mentais (Sanchez-Roige et al., 2019, 2021). De forma similar, atopouse un patrón de correlación xenética con outros trazos diferente para o inicio do consumo de cannabis e CUD (E. C. Johnson et al., 2020). Polo tanto, as distintas fases do consumo dunha mesma substancia, dende os comportamentos de consumo normalizados ata o diagnóstico dun SUD, parecen presentar unha base xenética parcialmente diferente.

Táboa 2. Exemplos na bibliografía de correlacións xenéticas significativas entre trazos de interese.

	Trazo 1	Trazo 2	r_g (SE)[a]	Fonte
Entre trazos de consumo da mesma substancia	Inicio do consumo de tabaco	Dependencia a nicotina	0,40 (0,05)	Quach et al., 2020
	Intensidade do consumo de tabaco	Dependencia a nicotina	0,95 (0,05)	Quach et al., 2020
	Intensidade do consumo de alcol	Dependencia a alcol	0,37 (0,09)	Walters et al., 2018
	Inicio do consumo de cannabis	CUD	0,50 (0,05)	E. C. Johnson et al., 2020
Entre trazos de consumo de distintas substancias	Inicio do consumo de tabaco	Intensidade do consumo de alcol	0,27 (0,02)	G. R. B. Saunders et al., 2022
	Inicio do consumo de tabaco	Consumo problemático de alcol	0,53 (0,02)	Zhou et al., 2023
	Inicio do consumo de cannabis	Inicio do consumo de tabaco	0,60 (0,05)	M. Liu et al., 2019
	Inicio do consumo de cannabis	Intensidade do consumo de alcol	0,36 (0,06)	M. Liu et al., 2019
	Dependencia a nicotina	Dependencia a alcol	0,57 (0,13)	Quach et al., 2020
	Consumo problemático de alcol	CUD	0,61 (0,04)	Zhou et al., 2023
	Consumo problemático de alcol	Trastorno polo uso de opioides	0,77 (0,03)	Zhou et al., 2023
Entre trazos de consumo de substancias e trazos externalizantes	Inicio do consumo de tabaco	ADHD	0,41 (0,10)	M. Liu et al., 2019
	Inicio do consumo de tabaco	Tolerancia xeral ao risco	0,28 (0,03)	M. Liu et al., 2019
	Inicio do consumo de tabaco	Idade á que se tivo o/a primeiro/a fillo/a	-0,48 (0,02)	M. Liu et al., 2019
	Intensidade do consumo de tabaco	ADHD	0,20 (0,09)	M. Liu et al., 2019
	Intensidade do consumo de tabaco	Tolerancia xeral ao risco	0,15 (0,04)	M. Liu et al., 2019
	Intensidade do consumo de tabaco	Idade á que se tivo o/a primeiro/a fillo/a	-0,40 (0,03)	M. Liu et al., 2019
	Intensidade do consumo de alcol	Tolerancia xeral ao risco	0,20 (0,04)	M. Liu et al., 2019
	Inicio do consumo de cannabis	ADHD	0,29 (0,07)	Soler Artigas et al., 2020
	Dependencia a nicotina	ADHD	0,54 (0,07)	Abdellaoui et al., 2021
	Dependencia a alcol	ADHD	0,46 (0,09)	Abdellaoui et al., 2021
	CUD	ADHD	0,36 (0,11)	Abdellaoui et al., 2021

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

Entre trazos de consumo de substancias e trazos internalizantes	Inicio do consumo de tabaco	Depresión maior	0,19 (0,04)	M. Liu et al., 2019
	Inicio do consumo de tabaco	Neuroticismo	0,20 (0,03)	M. Liu et al., 2019
	Intensidade do consumo de tabaco	Depresión maior	0,12 (0,06)	M. Liu et al., 2019
	Intensidade do consumo de tabaco	Neuroticismo	0,13 (0,04)	M. Liu et al., 2019
	Dependencia a nicotina	Depresión maior	0,45 (0,07)	Abdellaoui et al., 2021
	Dependencia a alcol	Depresión maior	0,52 (0,08)	Abdellaoui et al., 2021
	CUD	Depresión maior	0,14 (0,11)	Abdellaoui et al., 2021
Entre trazos de consumo de substancias e trastornos psiquiátricos[b]	Inicio do consumo de tabaco	Esquizofrenia	0,14 (0,02)	M. Liu et al., 2019
	Inicio do consumo de tabaco	Trastorno bipolar	0,05 (0,03)	M. Liu et al., 2019
	Intensidade do consumo de tabaco	Esquizofrenia	0,10 (0,03)	M. Liu et al., 2019
	Inicio do consumo de cannabis	Esquizofrenia	0,26 (0,03)	Jang et al., 2020
	Inicio do consumo de cannabis	Trastorno bipolar	0,27 (0,04)	Jang et al., 2020
	Dependencia a nicotina	Esquizofrenia	0,18 (0,05)	Abdellaoui et al., 2021
	Dependencia a nicotina	Trastorno bipolar	0,18 (0,06)	Abdellaoui et al., 2021
	Dependencia a nicotina	Trastorno de estrés postraumático	0,72 (0,15)	Quach et al., 2020
	Dependencia a alcol	Esquizofrenia	0,35 (0,05)	Abdellaoui et al., 2021
	Dependencia a alcol	Trastorno bipolar	0,24 (0,07)	Abdellaoui et al., 2021
	Consumo problemático de alcol	Esquizofrenia	0,31 (0,02)	Zhou et al., 2023
	Consumo problemático de alcol	Trastorno bipolar	0,32 (0,02)	Zhou et al., 2023

[a]: Correlación xenética entre ambos trazos e, entre parénteses, a súa desviación estándar.

[b]: Como xa aparecen indicadas nos apartados anteriores, neste apartado non se mostran as correlacións xenéticas existentes dos trazos de consumo de substancias con ADHD e con depresión maior.

Esta diferente susceptibilidade xenética, xunto coa maior herdabilidade dos SUD en comparación cos comportamentos de consumo de substancias atopada en estudos de xemelgos, suxire que, aínda que o consumo de substancias é necesario para desenvolver dependencia, existen factores xenéticos propios dos SUD. Un estudo de PRS indicou que, ademais dunha susceptibilidade xenética ao consumo de alcol, unha susceptibilidade xenética a psicopatoloxía xeral contribúe á dependencia a alcol (Facal et al., 2021). Ademais, un factor xenético de risco xeral a SUD construído mediante genomic SEM presentou unha varianza única significativa, mesmo despois de ter en conta a susceptibilidade xenética ao consumo de substancias en poboación xeral e a outros trastornos psiquiátricos (Hatoum et al., 2022). Este resultado suxire que a susceptibilidade a SUD non se explica só pola combinación da susceptibilidade xenética ao consumo de substancias e a psicopatoloxía.

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

A diferenza doutros comportamentos de consumo de tabaco, a intensidade do consumo de tabaco definida como o número medio de cigarros fumados ao día é case idéntica xeneticamente á dependencia a nicotina (Táboa 2) (Quach et al., 2020). Debido a esta alta correlación xenética e ao seu patrón similar de asociación con outros trazos, a intensidade do consumo de tabaco foi proposta como un indicador da susceptibilidade xenética para a dependencia a nicotina (Sanchez-Roige et al., 2021). Pola contra, a intensidade do consumo de alcol non se considera un bo indicador xenético dun consumo problemático de alcol, debido á súa diverxencia na asociación con outros trazos e a súa correlación xenética moderada (Táboa 2). Estes resultados parecen indicar que, dependendo da substancia en cuestión, un fenotipo que mida a intensidade do consumo da substancia e o SUD poden ser máis ou menos similares a nivel xenético.

1.4.3. Susceptibilidade xenética compartida entre trazos de consumo de distintas substancias

Ao igual que os estudos de xemelgos, os estudos xenómicos evidenciaron a existencia dunha susceptibilidade xenética parcialmente compartida entre o consumo e dependencia a distintas substancias, a cal podería explicar a elevada frecuencia do policonsumo. Por un lado, a correlación xenética entre trazos de consumo de diferentes substancias é relativamente alta (Táboa 2). Evidencia desta susceptibilidade xenética compartida tamén provén de estudos de PRS, estando PRS para comportamentos de consumo de tabaco significativamente asociadas ao consumo de alcol (L.-H. Chang et al., 2019).

Esta susceptibilidade xenética compartida podería reflectir a existencia dun factor de risco subxacente común. Tal e como se mencionou previamente, aplicando genomic SEM a GWAS para consumo problemático de alcol, consumo problemático de tabaco, CUD e trastorno polo uso de opioides (OUD, do inglés *opioid use disorder*), conseguiuse identificar un factor xenético que aumenta o risco a múltiples SUD (Hatoum et al., 2023). Este factor común caracterizouse por unha maior polixenicidade que a susceptibilidade xenética a cada substancia específica. Ademais, mentres que este factor xenético común estivo asociado con psicopatoloxía, enfermidades somáticas e comportamentos externalizantes; os loci asociados especificamente a cada substancia foron xenes metabólicos e codificantes de receptores.

Este solapamento xenético tamén podería ser debido á coñecida como "hipótese de entrada", a cal suxire que o consumo de substancias legais como o tabaco e o alcol pode levar ao consumo de substancias ilegais como o cannabis, a cocaína ou os opioides (Kleinig, 2015). Certa evidencia desta hipótese foi atopada mediante MR, a cal demostrou que o inicio do consumo de tabaco ten un efecto causal sobre o posterior consumo de cannabis, aínda que sen

USC DE COMPOSTELA

poder descartar a existencia dun efecto causal en sentido inverso nin dun factor de risco común a ambos trazos (Z. E. Reed et al., 2022).

1.4.4. Evidencia xenómica da vía externalizante e a vía internalizante cara o consumo de substancias

Os datos de estudos de correlacións xenéticas apoiaron a existencia dunha susceptibilidade xenética compartida entre os trazos de consumo de substancias e trazos externalizantes e internalizantes (Táboa 2). Por exemplo, o estudo de Abdellaoui et al. (2021) suxeriu que, dentro dos distintos trastornos psiquiátricos avaliados, ADHD e depresión maior presentan a maior correlación xenética con SUD. Outros estudos confirmaron esta correlación xenética moderada destes dous trazos con SUD e con comportamentos de consumo de substancias. Do mesmo xeito, comportamentos de toma de riscos considerados como externalizantes e o neuroticismo, unha característica común á ansiedade e á depresión altamente correlacionada cun factor internalizante (Griffith et al., 2010), estiveron correlacionados xeneticamente con comportamentos de consumo de substancias.

A existencia da vía externalizante e a vía internalizante cara o consumo de substancias tamén se ve apoiada por estudos de MR que evidencian un efecto causal de trazos externalizantes e internalizantes sobre trazos de consumo de substancias. Por un lado, existe forte evidencia dun efecto causal de ADHD sobre comportamentos de consumo de tabaco e cannabis (Treur et al., 2019; Vilar-Ribó et al., 2020). Por outro lado, os estudos de MR tamén indican que a depresión é un factor de risco para comportamentos de consumo de tabaco e dependencia a alcol (Polimanti et al., 2019; Wootton et al., 2020; Yao et al., 2021). Ademais, mostran un efecto causal do neuroticismo sobre a intensidade do consumo de tabaco dentro dos/as fumadores/as actuais (Sallis et al., 2019).

1.5. APLICACIÓN DA METODOLOXÍA XENÓMICA AO ESTUDO DA ASOCIACIÓN ENTRE TRAZOS DE CONSUMO DE SUBSTANCIAS E TRASTORNOS PSIQUIÁTRICOS

1.5.1. Susceptibilidade xenética compartida entre trazos de consumo de substancias e trastornos psiquiátricos

A alta comorbididade entre o consumo de substancias e os trastornos psiquiátricos reflíctese nun notable solapamento xenético entre ambos trazos. Estudos de correlacións xenéticas mostraron unha correlación xenética significativa dos SUD con trastornos psiquiátricos como esquizofrenia, depresión maior, trastorno bipolar ou ADHD (Táboa 2). O factor xenético que recolle o risco común a SUD tamén mostrou unha forte correlación xenética con trastornos psiquiátricos como ADHD, depresión maior ou ansiedade (Hatoum et al., 2023).

USC UNIVERSIDADE DE SANTIAGO

Esta correlación xenética con trastornos psiquiátricos tamén se estende aos comportamentos de consumo de substancias (Táboa 2). Jang et al. (2020) detectou, ademais dunha forte correlación xenética dos comportamentos de consumo de tabaco, alcol e cannabis cos trastornos psiquiátricos, un alto número de loci pleiotrópicos para ambos trazos.

Os estudos de PRS confirmaron o solapamento xenético entre trastornos psiquiátricos e trazos de consumo de substancias. Unha PRS a psicopatoloxía, baseada nunha meta-análise de GWAS para diferentes trastornos psiquiátricos realizada polo PGC, mostrou unha asociación significativa cun maior consumo de substancias como alcol, cannabis, cocaína, nicotina e opioides, explicando, aproximadamente, o 1% da varianza no consumo xeral de substancias (Carey et al., 2016). Do mesmo xeito, unha PRS baseada nesta mesma meta-análise e outra PRS específica para esquizofrenia estiveron significativamente asociadas a padecer un SUD (Gurriarán et al., 2019).

Os estudos máis recentes sobre a estrutura da psicopatoloxía permiten contextualizar mellor esta susceptibilidade xenética compartida entre trastornos psiquiátricos e trazos de consumo de substancias. A alta comorbididade entre SUD e outros trastornos psiquiátricos podería explicarse mediante un factor de psicopatoloxía xeral que predispón a calquera trastorno psiquiátrico e que capta o risco compartido entre todos eles (Caspi et al., 2014). Orixinalmente, este factor de psicopatoloxía xeral foi identificado aplicando modelos de ecuacións estruturais a datos de diagnóstico de trastornos e síntomas psiquiátricos (Lahey et al., 2012). Posteriormente, o factor de psicopatoloxía xeral foi tamén identificado a nivel xenético (Allegrini et al., 2020; Grotzinger et al., 2019; Neumann et al., 2016; Pettersson et al., 2016; Selzam et al., 2018). Por exemplo, identificouse un factor xenómico latente de risco a psicopatoloxía aplicando genomic SEM aos GWAS de cinco trastornos psiquiátricos (Grotzinger et al., 2019). Un estudo baseado nos datos de irmáns/ás de rexistros nacionais de Suecia mostrou que o factor xenético de psicopatoloxía xeral contribúe a varios trastornos psiquiátricos, incluíndo os SUD (Pettersson et al., 2016). En concreto, sobre o 35% e o 20% da varianza a nivel fenotípico do abuso de substancias ilegais e do AUD, respectivamente, pode ser explicado por este factor xenético compartido entre trastornos psiquiátricos. Do mesmo xeito, aplicando unha análise de compoñentes principais á correlación xenética entre trastornos psiquiátricos, estimada por diferentes métodos, conseguiuse identificar un factor de psicopatoloxía xeral polixénico que contribuíu a todos os trastornos psiquiátricos analizados, entre eles, o abuso de alcol (Selzam et al., 2018). Ademais, unha PRS para este factor de psicopatoloxía xeral identificada mediante análise de compoñentes principais mostrou unha contribución significativa á dependencia a alcol (Facal et al., 2021). Polo tanto, este factor de

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

psicopatoloxía xeral parece ter un impacto substancial tanto na susceptibilidade xenética a SUD como a outros trastornos psiquiátricos.

Porén, tal e como demostran varios estudos epidemiolóxicos e xenómicos, a psicopatoloxía observada parece explicarse mellor empregando varios factores latentes específicos que un único factor de psicopatoloxía xeral (Figura 4A) (Caspi et al., 2014; Gard et al., 2021; Krueger et al., 2001). Á hora de engadir estes factores latentes específicos, existen principalmente dúas propostas. Por un lado, un modelo que inclúe só estes factores latentes específicos cunha alta correlación entre eles (Figura 4B). Por outro lado, os modelos bifactorial ou factorial xerárquico que inclúen, ademais destes factores máis específicos, o factor de psicopatoloxía xeral (Figura 4C e Figura 4D). Dentro destas dúas propostas para explicar a estrutura de psicopatoloxía, algúns estudos apoiaron os modelos bifactorial e factorial xerárquico (Caspi et al., 2014; Castellanos-Ryan et al., 2016; Laceulle et al., 2015) e outros estudos apoiaron o modelo de factores correlacionados (Waldman et al., 2020). Ademais, outros estudos atopan evidencia de ambos modelos (Grotzinger et al., 2022; Rosenström et al., 2019). Por exemplo, Grotzinger et al. (2022) atoparon que tanto un modelo con catro factores latentes correlacionados como un modelo factorial xerárquico, onde, ademais destes catro factores latentes, está incluído o factor de psicopatoloxía xeral, explican ben a correlación xenética entre 11 trastornos psiquiátricos. Porén, outro estudo de genomic SEM que analizou GWAS para trastornos psiquiátricos e trazos de consumo de substancias apoiou que un modelo de catro factores latentes correlacionados é capaz de explicar mellor os datos xenómicos que un modelo bifactorial (Waldman et al., 2020). De todos os xeitos, incluíndo ou non o factor de psicopatoloxía xeral, os trazos de consumo de substancias aparecen xeralmente incluídos nun factor específico externalizante, de acordo coa existencia dunha vía externalizante cara o consumo de substancias (Caspi et al., 2014; Gard et al., 2021; Jang et al., 2022; Rosenström et al., 2019; Waldman et al., 2020). Non obstante, a diferencia deses outros estudos, o estudo de Abdellaoui et al. (2021) atopou que os SUD non forman parte dun factor latente externalizante xunto con ADHD nin dun factor latente único separado dos outros trastornos psiquiátricos.

USC
UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

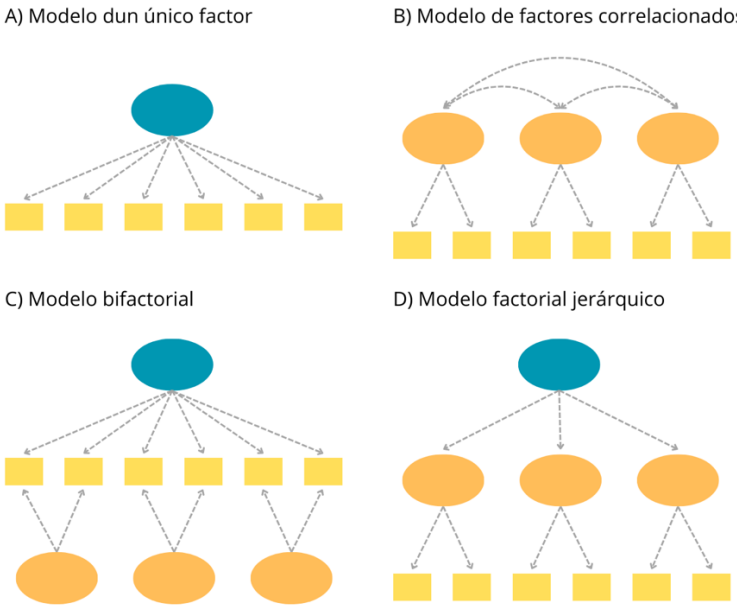


Figura 4. Posibles modelos para explicar a estrutura de psicopatoloxía (Caspi & Moffitt, 2018). A) Modelo cun único factor latente, o factor de psicopatoloxía xeral, que explica a varianza compartida entre os trastornos psiquiátricos. B) Modelo con varios factores latentes correlacionados entre si que inflúen, cada un deles, nun subconxunto de trastornos psiquiátricos. C) Modelo bifactorial co factor de psicopatoloxía xeral que explica a varianza compartida entre os trastornos psiquiátricos e outros factores latentes independentes que inflúen nun subconxunto máis pequeno de trastornos psiquiátricos. D) Modelo factorial xerárquico cun factor xeral de segundo orde, o factor de psicopatoloxía xeral, que explica a varianza compartida entre os factores latentes de primeiro orde.

1.5.2. Abordaxe xenómica da asociación entre tabaquismo e esquizofrenia

1.5.2.1. Susceptibilidade xenética compartida entre tabaquismo e esquizofrenia

GWAS recentes, baseados en amplas mostras para tabaquismo e esquizofrenia, mostraron a existencia dunha susceptibilidade xenética compartida entre ambos trazos. Por un lado, confirmaron a agardada asociación dos xenes que codifican para subunidades dos nAchR con comportamentos de consumo de tabaco e dependencia a nicotina (J. Chen et al., 2020; M. Liu et al., 2019; Quach et al., 2020; G. R. B. Saunders et al., 2022). Mais tamén identificaron unha asociación significativa da agrupación de xenes de nAchR *CHRNA5-CHRNA3-CHRNB4* e do xene de nAchR *CHRNA2* con esquizofrenia (Schizophrenia Working Group of the Psychiatric Genomics Consortium et al., 2014, 2022). Polo tanto, os xenes de nAchR parecen estar asociados a ambos trazos. A nivel funcional, as variantes xenéticas de *CHRNA5-CHRNA3-CHRNB4* asociadas tanto á cantidade de consumo de tabaco como a esquizofrenia parecen estar regulando a expresión destes xenes en rexións do cerebro (Ohi et al., 2019). Ademais dos xenes de nAchR, identificáronse outros loci pleiotrópicos asociados significativamente a través da mesma variante tanto a esquizofrenia como ao inicio de consumo de tabaco e ao número medio de cigarros fumados ao día (Jang et al., 2022).

USC DE SANTIAGO DE COMPOSTELA

A nivel xenómico, tamén se atopou evidencia dunha susceptibilidade xenética compartida entre tabaquismo e esquizofrenia. Por un lado, numerosos estudos identifican unha correlación xenética significativa entre trazos de consumo de tabaco e esquizofrenia (Abdellaoui et al., 2021; Barkhuizen et al., 2021; Hartz et al., 2018; Jang et al., 2022). Por exemplo, a intensidade do consumo de tabaco, o inicio do consumo regular de tabaco e a dependencia a nicotina presentan unha correlación xenética significativa con esquizofrenia (Táboa 2) (M. Liu et al., 2019; Quach et al., 2020).

Os estudos de PRS tamén reforzan esta idea. Os resultados destes estudos mostran unha asociación recíproca entre esquizofrenia e tabaquismo. Nun destes estudos, tanto a PRS para esquizofrenia estivo asociada cunha maior dependencia a nicotina e cunha maior intensidade do consumo de tabaco como a PRS para os niveis séricos de cotinina, un metabolito da nicotina que se emprega como indicador do consumo recente de nicotina, estivo asociada con esquizofrenia (J. Chen et al., 2016). A PRS a esquizofrenia está asociada con todos os niveis de consumo de tabaco, dende un consumo non regular ata unha dependencia a nicotina severa (Carey et al., 2016; Miller et al., 2019; Reginsson et al., 2018). Do mesmo xeito, a PRS para tabaquismo, en concreto, para o inicio do consumo de tabaco e para a intensidade do consumo de tabaco, son significativamente maiores en pacientes con esquizofrenia que en controis (Peterson et al., 2021). Porén, no caso da PRS para a intensidade do consumo de tabaco, esta asociación estivo parcialmente mediada pola agrupación de nAchR do cromosoma 15.

É importante sinalar que os estudos de PRS mostran tamén unha susceptibilidade xenética compartida entre esquizofrenia e outros SUD. Por exemplo, un estudo, aínda que atopou unha asociación significativa entre a PRS a esquizofrenia e TUD, mostrou unha relación aínda máis consistente entre a PRS a esquizofrenia e ter calquera tipo de SUD: TUD, AUD, CUD ou trastorno polo uso de estimulantes (Hartz et al., 2017). Empregando genomic SEM identificouse un factor xenético común a esquizofrenia, ao inicio do consumo de tabaco e ao inicio do consumo de cannabis que explicou o 8,6% da variación de esquizofrenia e o 37,2% da variación do inicio do consumo de tabaco (W. Song et al., 2022). Estes resultados apoian que o risco xenético compartido entre a esquizofrenia e o consumo de substancias é, polo menos parcialmente, común a todas as substancias, en lugar de ser totalmente específico para cada substancia individual. Polo tanto, a comorbididade entre tabaquismo e SUD pode estar contribuíndo á asociación xenética entre esquizofrenia e tabaquismo.

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

Este solapamento xenético entre esquizofrenia e tabaquismo non pode ser explicado pola hipótese da automedicación xa que, segundo esta hipótese, as variantes xenéticas que confiren

risco a tabaquismo nas persoas con esquizofrenia serían diferentes ás da poboación xeral. Esta susceptibilidade xenética compartida entre esquizofrenia e tabaquismo si que podería deberse a unha pleiotropía horizontal, é dicir, os mesmos factores xenéticos poderían facer aos individuos máis vulnerables a ambos trazos de forma independente (Chambers et al., 2001). Un estudo atopou evidencia desta hipótese analizando unha cohorte de 1.805 participantes, incluíndo pacientes cun diagnóstico de psicose non afectiva, os seus irmáns non afectados e controis non relacionados (Lin et al., 2023). Este traballo revelou que os/as irmáns/ás de pacientes con psicose presentan un comportamento de consumo de tabaco intermedio entre os/as pacientes e os controis. Ademais, as PRS para comportamentos de consumo de tabaco estiveron asociadas a síntomas psicóticos subclínicos nos/as irmáns/ás non afectados e en controis. Polo tanto, estes resultados suxiren unha pleiotropía horizontal entre tabaquismo e psicose.

A susceptibilidade xenética compartida entre esquizofrenia e tabaquismo tamén podería ser explicada por unha pleiotropía vertical onde o tabaco actúa como factor de risco para a esquizofrenia (Solovieff et al., 2013). Neste escenario, como os GWAS identifican tamén aquelas variantes xenéticas cun efecto no trazo de interese a través dos seus factores de risco, atoparíase un solapamento xenético entre o trazo de interese e os seus factores de risco (Gage, Davey Smith, et al., 2016; McGrath, 2015). Neste caso, se o tabaco é un factor de risco para a esquizofrenia, habería unha maior prevalencia de fumadores/as entre os/as pacientes con esquizofrenia que nos controis, o cal faría que as variantes xenéticas asociadas a tabaquismo estivesen tamén asociadas a esquizofrenia no correspondente GWAS de esquizofrenia (Gage, Davey Smith, et al., 2016; Gage & Munafò, 2015a).

Para identificar cal destas dúas hipóteses explica o solapamento xenético entre esquizofrenia e tabaquismo, é necesario estratificar ás pacientes con esquizofrenia polo seu hábito tabáquico (Alderson & Lawrie, 2015; Gage, Davey Smith, et al., 2016; Gage & Munafò, 2015a). Desta forma, atopariamos evidencia dunha pleiotropía horizontal se as variantes xenéticas asociadas a tabaquismo estivesen asociadas a esquizofrenia tanto en fumadores/as como en non fumadores/as, xa que estarían asociadas a esquizofrenia independentemente do tabaquismo. Pola contra, se o tabaquismo está na vía causal cara a esquizofrenia, as variantes xenéticas asociadas á intensidade do consumo de tabaco estarían asociadas a esquizofrenia só nos/as pacientes que fuman e non nos/as pacientes non fumadores/as, xa que aumentarían o risco a esquizofrenia a través do tabaquismo. Aínda que esta estratificación segundo o hábito tabáquico pode introducir outros nesgos, existiría evidencia convincente dunha vía causal cara a esquizofrenia na que está involucrado o tabaquismo se o nivel de asociación é moi diferente

USC UNIVERSIDADE DE SANTIAGO

entre os/as fumadores/as e os/as non fumadores/as (Davies et al., 2018; Gage, Davey Smith, et al., 2016). Wium-Andersen et al. (2015) aplicaron este enfoque á hora de estudar 63.000 persoas danesas, atopando unha asociación entre unha variante do xene *CHRNA3*, asociada coa intensidade do consumo de tabaco, e a prescrición de tratamento antipsicótico só en persoas fumadoras, sendo esta asociación inexistente en persoas non fumadoras. Polo tanto, os seus resultados apoian a existencia dunha pleiotropía vertical entre tabaquismo e esquizofrenia. Porén, este estudo non controlou polo consumo de cannabis, o cal podería estar mediando esta asociación.

1.5.2.2. Relacións de causalidade entre tabaquismo e esquizofrenia

Os estudos de MR contribuíron ao debate da causa da asociación entre tabaquismo e esquizofrenia, investigando se existe unha relación causal entre ambos trazos. Aqueles estudos de MR baseados en GWAS cunha potencia limitada non atoparon evidencia de causalidade entre esquizofrenia e tabaquismo, suxerindo que a comorbididade atopada entre estes trazos xurde dunha pleiotropía horizontal (J. Chen et al., 2021; Gage et al., 2017). Porén, estudos de MR máis recentes, baseados en GWAS cunha maior potencia e en diversos métodos estatísticos robustos ás asuncións da MR, atoparon evidencia dun efecto causal do inicio do consumo de tabaco no risco a desenvolver esquizofrenia, aínda que non todas as análises de sensibilidade confirmaron estes resultados, suxerindo que poderían estar nesgados pola presenza de pleiotropía horizontal (Barkhuizen et al., 2021; Wootton et al., 2020; Yuan et al., 2020; M. Zhang et al., 2024). Algúns destes estudos de MR tamén apoiaron un efecto causal bidireccional, presentando forte evidencia dun efecto causal dos comportamentos de consumo de tabaco sobre a esquizofrenia pero tamén dun efecto causal pequeno pero consistente da esquizofrenia sobre comportamentos de consumo de tabaco (Barkhuizen et al., 2021; Wootton et al., 2020). O último estudo de MR que empregou os últimos GWAS para comportamentos de consumo de tabaco do GSCAN atopou non só un efecto causal do inicio do consumo de tabaco sobre esquizofrenia, senón que tamén da intensidade do consumo de tabaco e da idade de inicio do consumo de tabaco (M. Zhang et al., 2024). Ademais, estes resultados foron confirmados empregando múltiples métodos de MR e eliminando variables instrumentais asociadas con posibles factores de confusión. Este estudo, ademais, descartou unha causalidade inversa da esquizofrenia sobre calquera dos comportamentos de consumo de tabaco. En contraposición a todos estes estudos de MR, un estudo que empregou outras metodoloxías robustas á existencia de pleiotropía horizontal para avaliar relacións de causalidade, diferentes dos métodos de MR convencionais, en concreto, a metodoloxía de MR bidireccional desenvolvida por Pickrell et al. (2016) e o *Latent Causal Variable model*, non atopou unha

relación causal entre distintos comportamentos de consumo de tabaco e esquizofrenia (Jang et al., 2022). Este estudo suxeriu, polo tanto, que a asociación entre estes trazos reflicte unha pleiotropía horizontal, onde os mesmos factores xenéticos aumentan o risco a ambos trazos de forma independente.

Estas contradicións atopadas nos estudos de MR subliñan a necesidade de abordar a relación causal entre esquizofrenia e tabaquismo dende diferentes perspectivas (Lawlor et al., 2016; Ohlsson & Kendler, 2020). Se diferentes métodos que avalían causalidade, cada un coas súas propias fontes de nesgo potenciais, obteñen resultados consistentes, aumentaría a confianza nesa conclusión.

USC
UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

2. HIPÓTESES E OBXECTIVOS

Esta tese aborda tres temáticas centrais. Para cada un destes eixos, propóñense unha serie de hipóteses específicas que tratan de explicar a susceptibilidade dos trazos de consumo de substancias, tanto na poboación xeral como no contexto dos/as pacientes con esquizofrenia.

Hipóteses do Estudo 1, centradas no estudo da susceptibilidade xenética aos trazos de consumo de substancias:

H1.1. As variantes xenéticas de risco para os trazos de consumo de substancias exercen o seu efecto, principalmente, regulando a expresión de xenes cuxa expresión alterada confire unha maior susceptibilidade a estes trazos. Polo tanto, as variantes xenéticas próximas a xenes de módulos de co-expresión cunha expresión alterada no cerebro baixo o consumo de substancias contribuirán máis á susceptibilidade xenética a trazos de consumo de substancias que variantes xenéticas próximas a outros xenes expresados no cerebro.

H1.2. Os datos de expresión xénica do cerebro de peixe cebra permiten modelar as redes xénicas do cerebro humano cunha expresión alterada nos trazos de consumo de substancias e noutros trazos relacionados coa psiquiatría. Polo tanto, os conxuntos de xenes cunha expresión correlacionada no cerebro de peixe cebra estarán enriquecidos en conxuntos de xenes relevantes en psiquiatría e, en concreto, no consumo de substancias.

Hipóteses do Estudo 2, relacionadas co estudo da vía externalizante e a vía internalizante cara o consumo de substancias:

H2.1. Os problemas externalizantes e internalizantes inflúen no desenvolvemento de trazos de consumo de substancias. Polo tanto, os trazos externalizantes e internalizantes presentarán un solapamento xenético con trazos de consumo de substancias.

H2.2. A vía externalizante e a vía internalizante teñen un efecto diferente nos comportamentos de consumo de substancias e nos SUD. Polo tanto, os comportamentos de consumo de substancias e os SUD mostrarán diferenzas no patrón de solapamento xenético (número de variantes causais compartidas e correlación xenética das variantes xenéticas causais compartidas) cos trazos externalizantes e internalizantes.

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

Hipóteses dos Estudos 3, 4 e 5, enfocadas no estudo das causas da asociación entre tabaquismo e esquizofrenia:

H3.1. A maior prevalencia do tabaquismo en pacientes con esquizofrenia que na poboación xeral non se debe á hipótese da automedicación. Polo tanto, os/as pacientes con esquizofrenia son propensos a fumar polas mesmas razóns que a poboación en xeral, é dicir, por unha maior conduta externalizante e/ou internalizante. Se isto é certo:

- **H3.1.1.** A PRS para tabaquismo, estimada en poboación xeral, explicará aproximadamente igual o tabaquismo en pacientes con esquizofrenia que en poboación xeral.

- **H3.1.2.** PRS para trazos xeneticamente correlacionados con tabaquismo, estimadas en poboación xeral, explicarán aproximadamente igual o tabaquismo en pacientes con esquizofrenia que en poboación xeral.

H3.2. Unha parte da maior prevalencia do tabaquismo en pacientes con esquizofrenia que na poboación xeral débese a que o tabaco é un factor de risco para desenvolver esquizofrenia. As variantes xenéticas asociadas a tabaquismo están tamén asociadas a esquizofrenia, en parte, debido á maior proporción de fumadores/as dentro das persoas con esquizofrenia que dentro dos controis. Se isto é certo:

- **H3.2.1.** Os sinais de asociación con esquizofrenia e tabaquismo dos xenes de subunidades de nAchR colocalizarán, é dicir, serán sinais de asociación comúns a ambos trazos, debidas á mesma variante causal.

- **H3.2.2.** A asociación da PRS para tabaquismo con esquizofrenia, así como a asociación de variantes xenéticas nos xenes de subunidades de nAchR con esquizofrenia, diminuirá ao axustar ou estratificar a mostra por consumo de tabaco.

- **H3.2.3.** O axuste do GWAS de esquizofrenia polo efecto causal do tabaquismo sobre esquizofrenia terá un impacto na susceptibilidade xenética a esquizofrenia, permitindo identificar novos loci asociados significativamente con esquizofrenia e loci que deixan de estar asociados significativamente con esquizofrenia.

- **H3.2.4.** O axuste do GWAS de esquizofrenia polo efecto causal do tabaquismo sobre esquizofrenia permitirá identificar factores xenéticos específicos que incrementan o risco a esquizofrenia a través do tabaquismo.

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

H3.3. Outra parte da maior prevalencia do tabaquismo en pacientes con esquizofrenia que na poboación xeral débese a que os mesmos factores xenéticos aumentan a susceptibilidade xenética a tabaquismo e a esquizofrenia de forma independente. Polo tanto, a asociación da PRS para tabaquismo con esquizofrenia, así como a asociación de variantes xenéticas nos xenes de nAchR con esquizofrenia, permanecerá ao axustar ou estratificar a mostra por consumo de tabaco.

O obxectivo principal da tese é afondar nos factores xenéticos que inflúen na susceptibilidade aos trazos de consumo de substancias, tanto aos comportamentos de consumo de substancias como aos SUD, estudando, con máis detalle, as causas da elevada prevalencia do tabaquismo en pacientes con esquizofrenia. Para abordar cada unha das hipóteses descritas anteriormente, establécense os seguintes obxectivos específicos:

Obxectivos do Estudo 1:

O1.1. Identificar módulos de xenes co-expresados potencialmente involucrados en trazos relacionados coa psiquiatría, entre eles, o consumo e a dependencia a substancias, a partir de datos de expresión xénica do cerebro de peixe cebra.

O1.2. Avaliar se a dependencia a alcol dunha mostra caso-control de orixe galega se explica mellor por unha PRS baseada en variantes xenéticas próximas aos xenes de módulos de co-expresión cunha expresión alterada baixo a abstinencia a substancias en peixe cebra que por unha PRS baseada en variantes xenéticas escollidas aleatoriamente dentro daquelas próximas a xenes expresados en cerebro.

Obxectivo do Estudo 2:

O2.1. Avaliar se o patrón de solapamento xenético de diferentes trazos externalizantes e internalizantes con distintos comportamentos de consumo de substancias e SUD, incluíndo o número de variantes xenéticas causais compartidas entre cada par de trazos e a correlación xenética do efecto das variantes xenéticas compartidas, é consistente cunha relación causal.

Obxectivo do Estudo 3:

O3.1. Realizar unha análise de colocalización para determinar se os sinais de asociación a tabaquismo e esquizofrenia de xenes de subunidades de nAchR son comúns a ambos trazos, tendo en conta tamén a súa colocalización co consumo de cannabis.

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

Obxectivos do Estudo 4:

O4.1. Analizar a contribución de diferentes PRS para tabaquismo e trazos correlacionados, estimadas na poboación xeral, no consumo de tabaco dos/as pacientes con esquizofrenia e dos controis dunha mostra para a que se determinou o hábito tabáquico.

O4.2. Analizar, na mostra anterior, a asociación da PRS para tabaquismo con esquizofrenia, tras axustar e estratificar a mostra por consumo de tabaco. Así mesmo, analizar, nesta mesma mostra, a asociación de variantes xenéticas situadas en xenes de nAchR con esquizofrenia axustando e estratificando a mostra por hábito tabáquico.

Obxectivos do Estudo 5:

O5.1. Condicionar o GWAS de esquizofrenia polo efecto causal do consumo de tabaco sobre esquizofrenia para descartar, dentro da susceptibilidade xenética a esquizofrenia, aquela mediada pola vía causal do tabaquismo.

O5.2. Comparar os loci significativos no GWAS de esquizofrenia antes e despois de axustar por tabaquismo, identificando novos xenes asociados significativamente a esquizofrenia e xenes que estaban asociados significativamente a esquizofrenia a través do tabaquismo.

O5.3. Caracterizar a susceptibilidade xenética a esquizofrenia mediada pola vía causal do tabaquismo, identificando aqueles trazos de consumo de substancias, trastornos psiquiátricos, trazos condutuais e trazos neuropsicolóxicos que presentan unha correlación xenética con esquizofrenia diferente tras condicionar por tabaquismo.

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

3. METODOLOXÍA

3.1. ESTUDOS DE ASOCIACIÓN DO XENOMA COMPLETO EMPREGADOS

Gran parte das ferramentas xenómicas empregadas nesta tese baseáronse na análise das estatísticas resumo dunha gran variedade de GWAS. En concreto, incluíronse GWAS para trastornos psiquiátricos, trazos de consumo de substancias, trazos condutuais, trazos da personalidade, trazos cognitivos, trazos relacionados co sono, trazos sociais e trazos fisiolóxicos. Para cada trazo, elixiuse aquel GWAS co maior tamaño de mostra dispoñible publicamente no momento da análise. Os GWAS empregados aparecen recollidos na Táboa 3. Parte dos GWAS seleccionados proveñen de consorcios a nivel internacional como o PGC, o GSCAN, o *Social Science Genetic Association Consortium* (SSGAC) ou o *Genetics of Personality Consortium* (GPC). Ademais, empregáronse GWAS de iniciativas como o *Neale lab* (http://www.nealelab.is/uk-biobank/) ou o *GeneATLAS* (http://geneatlas.roslin.ed.ac.uk) (Canela-Xandri et al., 2018), que analizaron centos de trazos na mostra de *UK Biobank*.

Para a maior parte das análises, empregáronse GWAS restrinxidos a mostras de ascendencia europea. Isto foi debido a varias razóns. Por un lado, a maior parte das ferramentas xenómicas empregadas nesta tese, debido ás diferenzas existentes na estrutura de LD e na frecuencia alélica entre poboacións, só permiten analizar GWAS da mesma ascendencia. Dado que actualmente a maior parte da mostra analizada nos GWAS é de ascendencia europea, os GWAS foron restrinxidos a esta ascendencia para maximizar a potencia das análises. Ademais, en moitos destes métodos xenómicos, como, por exemplo, LDSC, é necesario un panel de referencia para modelar o LD, o cal debe ser da mesma ascendencia que os GWAS analizados. Os paneis de referencia dispoñibles publicamente están principalmente compostos por individuos de ascendencia europea, dificultando, deste xeito, a análise de GWAS doutras poboacións (O'Connell et al., 2021). Do mesmo xeito, para aumentar a capacidade preditiva das PRS, elixíronse mostras de descubrimento da mesma ascendencia que as mostras dianas, é dicir, europeas (Peterson et al., 2019).

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

Táboa 3. Descrición dos GWAS empregados para cada trazo analizado nos estudos desta tese.

Abreviatura empregada[a]	Descrición do trazo	Fonte	Tamaño de mostra[b,c]
Trazos de consumo de substancias[d]			
SmkInit/EverSmk	Inicio do consumo de tabaco: ter sido algunha vez fumador/a regular fronte a nunca ter sido fumador/a regular	G. R. B. Saunders et al., 2022	805.431 (393.707)
		M. Liu et al., 2019	632.802 (311.655)
Cigsperday/CigDay	Número medio de cigarros consumidos por día	G. R. B. Saunders et al., 2022	326.497
		M. Liu et al., 2019	263.954
AgeSmk	Idade á que se iniciou o consumo de tabaco regularmente	G. R. B. Saunders et al., 2022	323.386
		M. Liu et al., 2019	262.990
SmkCes	Ser fumador/a regular actualmente fronte a ser ex-fumador/a	G. R. B. Saunders et al., 2022	388.313 (121.754)
		M. Liu et al., 2019	312.821 (92.564)
LifetimeSmk	Índice do tabaquismo ao longo da vida que recolle o inicio, a duración, a intensidade e o tempo transcorrido dende que se deixou de fumar	Wootton et al., 2020	462.690
FTND	Test de Fagerström para a dependencia a nicotina	Hancock et al., 2015	17.074
Frequency	Frecuencia actual de consumo de tabaco: nunca, ocasionalmente ou case todos os días	Canela-Xandri et al., 2018	501.649
Attempts	Número de intentos fallados para deixar de fumar	Neale lab[e]	81.026
TUD	Diagnóstico de trastorno polo uso de tabaco	GWAS propio empregando a mostra de *UK Biobank*[f]	409.535 (19.085)
		Canela-Xandri et al., 2018	452.264 (12.310)
DrnkWk	Número medio de bebidas alcólicas bebidas cada semana	G. R. B. Saunders et al., 2022	666.978
		M. Liu et al., 2019	537.349
AUD	Diagnóstico de trastorno polo uso de alcol	Icick et al., 2023	981.384 (75.635)
AD	Diagnóstico de dependencia a alcol	Walters et al., 2018	46.568 (11.569)

AUDIT-T	Puntuación total do test de identificación de trastornos polo uso de alcol (AUDIT)	Sanchez-Roige et al., 2019	121.604
AUDIT-C	Puntuación nos items relacionados co consumo de alcol do test de identificación de trastornos polo uso de alcol (AUDIT)	Sanchez-Roige et al., 2019	121.604
AUDIT-P	Puntuación nos items relacionados coas consecuencias problemáticas do consumo de alcol do test de identificación de trastornos polo uso de alcol (AUDIT)	Sanchez-Roige et al., 2019	121.604
CANN/CAN	Inicio do consumo de cannabis: ter consumido algunha vez cannabis fronte a nunca ter consumido cannabis	Pasman et al., 2018	162.082 (43.380)
CUD	Diagnóstico de trastorno polo uso de cannabis	Johnson et al., 2020	357.806 (14.080)
		Demontis, Rajagopal, et al., 2019	51.372 (2.387)
Opioid use	Consumo de opioides, polo menos unha vez na vida, en individuos sen un diagnóstico de dependencia a opioides	Polimanti et al., 2020	28.313 (2.876)
Opioid dep	Diagnóstico de dependencia a opioides	Polimanti et al., 2020	28.709 (3.272)
Trastornos psiquiátricos[d]			
ADHD	Diagnóstico de trastorno por déficit de atención e hiperactividade	Demontis et al., 2023	225.534 (38.691)
		Demontis, Walters, et al., 2019	53.293 (19.099)
SCZ	Diagnóstico de esquizofrenia ou un trastorno esquizoafectivo	Schizophrenia Working Group of the Psychiatric Genomics Consortium et al., 2022	130.644 (53.381)
		Schizophrenia Working Group of the Psychiatric Genomics Consortium et al., 2014	77.096 (33.637)
DEP	Diagnóstico de depresión maior ou depresión definida sen criterios diagnósticos reportada polo propio individuo	Howard et al., 2019	500.199 (170.756)
MDD	Diagnóstico de depresión maior	Wray et al., 2018	173.005 (59.851)
BIP	Diagnóstico de trastorno bipolar	Mullins et al., 2021	413.466 (41.917)

ANX	Ter un diagnóstico de trastorno de ansiedade fronte a ter poucos ou ningún síntoma clínico de ansiedade	Otowa et al., 2016	17.310[g]
	Diagnóstico por ansiedade, nervios ou trastorno de ansiedade xeneralizada[h]	Neale lab[e]	117.751 (16.730)
GAD	Resposta polo propio individuo á escala do trastorno de ansiedade xeneralizada de 2 items	Levey et al., 2020	175.163
ASD	Diagnóstico dun trastorno do espectro autista	Grove et al., 2019	46.351 (18.382)
AN	Diagnóstico de anorexia nerviosa ou anorexia nerviosa reportada polo propio individuo	Watson et al., 2019	72.517 (16.992)
PTSD	Diagnóstico de trastorno de estrés postraumático	Nievergelt et al., 2019	34.529 (9.354)
OCD	Diagnóstico de trastorno obsesivo-compulsivo	Arnold et al., 2018	9.725 (2.688)
TS	Diagnóstico do síndrome de Tourette	Yu et al., 2019	14.307 (4.819)
Cross-disorder	Meta-análise de GWAS para oito trastornos psiquiátricos: anorexia nerviosa, trastorno do espectro autista, ADHD, trastorno bipolar, depresión maior, trastorno obsesivo-compulsivo, síndrome de Tourette e esquizofrenia	Cross-Disorder Group of the Psychiatric Genomics Consortium et al., 2019	438.997 (162.165)
Trazos condutuais[d]			
Risk PC1	Primeira compoñente principal de catro comportamentos de toma de risco do *UK Biobank*: propensión a un exceso de velocidade conducindo, número medio de bedidas alcólicas consumidas á semana, inicio do consumo de tabaco e número de parellas sexuais ao longo da vida	Karlsson Linnér et al., 2019	315.894
RISK/Risk tolerance	Tolerancia xeral ao risco[h]	Karlsson Linnér et al., 2019	466.571
NSEX/Sexual partners	Número de parellas sexuais ao longo da vida[h]	Karlsson Linnér et al., 2019	370.711
Automobile speeding	Propensión a un exceso de velocidade conducindo[h]	Karlsson Linnér et al., 2019	404.291
AFB	Idade á que se tivo o primeiro fillo/a vivo/a[h]	Mills et al., 2021	542.901
		Neale lab[e]	131.987
AFS	Idade na primeira relación sexual[h]	Mills et al., 2021	397.338
AGG	Conduta agresiva infantil avaliada de forma continua	Ip et al., 2021	151.741

Trazos da personalidade[d]			
NEURO/Neuroticism	Puntuación de neuroticismo baseada en diferentes escalas	Nagel, Jansen, et al., 2018	390.278
		de Moor et al., 2015	63.661
WORRY-NEURO/Worry	Subconxunto de preocupación: puntuación nos items relacionados coa preocupación dentro da puntuación de neuroticismo	Nagel, Jansen, et al., 2018	348.219
DEP-NEURO/Depressed	Subconxunto de afecto depresivo: puntuación nos items relacionados co afecto depresivo dentro da puntuación de neuroticismo	Nagel, Jansen, et al., 2018	357.957
MOOD	Experimentar cambios de humor frecuentemente[h]	Nagel, Watanabe, Stringer, Posthuma, & Van Der Sluis, 2018	373.733 (168.180)
Extraversion	Puntuación de extraversión baseada na resposta a diferentes items	van den Berg et al., 2016	63.030
SWB	Benestar subxectivo: combinación de afecto positivo e satisfacción coa vida en xeral	Kim et al., 2022	311.126
		Okbay et al., 2016	204.966
LONE/Loneliness	Sentirse só/a con frecuencia[h]	Nagel, Watanabe, et al., 2018	376.352 (67.743)
		Neale lab[e]	355.583 (63.508)
CONF/Confide	Frecuencia coa que se pode confiar en alguén próximo[h]	Neale lab[e]	350.618
Highly irritable	Sentirse algunha vez irritable durante dous días[h]	Neale lab[e]	117.359 (20.930)
Irritability	Ser unha persoa irritable[h]	Neale lab[e]	345.231 (96.862)
Extreme irritability	Ter tido algunha vez un período de tempo sentindo unha irritabilidade extrema[h]	Neale lab[e]	114.422 (29.747)
Trazos cognitivos[d]			
Intelligence	Intelixencia avaliada mediante distintas probas neurocognitivas	Savage et al., 2018	269.867
Cognitive	Rendemento cognitivo avaliado mediante distintas probas neurocognitivas	J. J. Lee et al., 2018	257.828
Trazos relacionados co sono[d]			
Insomnia	Insomnio: frecuencia coa que se teñen problemas para quedarse durmido pola noite ou coa que se esperta en medio da noite[h]	Jansen et al., 2019	386.533 (109.402)

Chronotype	Cronotipo: ser unha persoa máis activa polas mañás ou polas noites[h]	Neale lab[e]	322.488
Sleep	Horas de sono cada 24 horas[h]	Neale lab[e]	359.020
Narcolepsy	Narcolepsia: frecuencia coa que se adormece ou queda durmido durante o día sen intención[h]	Neale lab[e]	359.752
Trazos sociais[d]			
EA	Anos de formación académica	J. J. Lee et al., 2018	766.345
Social deprivation	Nivel socioeconómico da área na que o individuo vive medido a través do índice de privación social de Townsend	Hill et al., 2016	112.151
LONE-MTAG/Loneliness and isolation	GWAS combinado dos GWAS de soidade e de illamento social empregando MTAG	Day, Ong, & Perry, 2018	445.024
Trazos fisiolóxicos para os cales o tabaco é un factor de risco[d]			
BMI	Índice de masa corporal	Pulit et al., 2019	806.834
LungCancer	Cancro de pulmón	McKay et al., 2017	85.716 (29.266)
FVC	Función pulmonar medida como a capacidade vital forzada	Shrine et al., 2019	321.047
-	Obesidade de tipo 1 definida segundo o índice de masa corporal	Berndt et al., 2013	98.697 (32.856)
-	Enfermidade das arterias coronarias	Nikpay et al., 2015	184.305 (60.802)[i]
-	Niveis de colesterol de lipoproteínas de alta densidade (HDL)	Hoffmann et al., 2018	76.627

[a]: No caso de habelas, abreviatura/s empregada/s nos estudos desta tese para referirse a ese GWAS.

[b]: Para os GWAS de trazos dicotómicos, móstrase o número de casos entre parénteses.

[c]: Tamaño de mostra do GWAS restrinxido a poboación europea.

[d]: Clasificación en grupos para facilitar a busca mais algúns trazos poderían incluírse en varios grupos.

[e]: http://www.nealelab.is/uk-biobank/; resultados de GWAS da ronda 2, publicados o 1 de Agosto de 2018.

[f]: Pódese atopar unha descrición detallada deste GWAS no Estudo 2 da tese.

[g]: O número de casos e de controis non se indican no artigo orixinal.

[h]: Trazos determinados a partir da resposta do propio individuo a unha ou unhas poucas preguntas de UK Biobank.

[i]: Neste caso, non existe GWAS restrinxido a ascendencia europea, aínda que o 77% da mostra do GWAS é europea.

3.2. METODOLOXÍAS XENÓMICAS EMPREGADAS

Nesta tese, empregáronse multitude de ferramentas xenómicas para caracterizar a susceptibilidade xenética a un único trazo e analizar a susceptibilidade xenética compartida entre diferentes trazos. A continuación, descríbese de maneira resumida a metodoloxía seguida para acadar cada un dos obxectivos da tese. Ademais, na Táboa 4 recóllense as principais aplicacións informáticas empregadas. Unha explicación máis técnica pode atoparse en cada un dos estudos específicos.

Táboa 4. Principais aplicacións informáticas empregadas nesta tese.

Software	Referencia[a]	Ligazón[b]	Aplicación[c]	Estudo[d]
R	-	https://www.r-project.org	Manexo de datos, creación de figuras e análises estatísticas	Estudos 1, 2, 3, 4 e 5
WGCNA	Langfelder & Horvath, 2008	-	Obtención de módulos de xenes co-expresados	Estudo 1
Biomart de Ensembl	Smedley et al., 2009	http://www.ensembl.org/biomart/martview	Conversión de sondas a xenes e cambios na nomenclatura de xenes	Estudo 1
SNP2GENE de FUMA	Watanabe et al., 2017	https://fuma.ctglab.nl	Identificación de loci significativos e mapeado xénico	Estudos 1 e 5
WebGestaltR	Liao et al., 2019	https://cran.r-project.org/web/packages/WebGestaltR/index.html	Análise de enriquecemento en conxuntos de xenes de *Gene Ontology*	Estudo 1
PRSice-2	Choi & O'Reilly, 2019	https://choishingwan.github.io/PRSice/	Estimación de PRS co método de *clumping* e *thresholding* e análise de asociación das PRS co trazo de interese	Estudo 1
PRSet de PRSice-2	Choi et al., 2023	https://choishingwan.github.io/PRSice/	Estimación de PRS baseadas en conxuntos de xenes e obtención do nivel de enriquecemento das PRS na asociación co trazo diana	Estudo 1
MiXeR	Frei et al., 2019	https://github.com/precimed/mixer	Análise do patrón de solapamento xenético entre dous trazos	Estudo 2
LDSC	B. Bulik-Sullivan et al., 2015	https://github.com/bulik/ldsc	Estimación de herdabilidades baseadas en SNP, correlación xenética entre trazos e grao de solapamento das mostras de GWAS	Estudos 3, 4 e 5
coloc	Giambartolomei et al., 2014	https://chr1swallace.github.io/coloc/articles/a01_intro.html	Análise de colocalización asumindo a existencia dunha única variante causal no locus e GWAS sen solapamento de mostras	Estudo 3
	Wallace, 2020		Análise de colocalización para varios sinais de asociación nun locus	
LocusCompare	B. Liu et al., 2019	http://locuscompare.com	Visualización do sinal de asociación de dous GWAS tendo en conta o LD	Estudo 3
fgwas	Pickrell et al., 2016	https://github.com/joepickrell/fgwas	Análise de colocalización para GWAS con solapamento de mostras	Estudos 3 e 5
gwas-pw		https://github.com/joepickrell/gwas-pw		
AVENGEME	Palla & Dudbridge, 2015	https://github.com/DudbridgeLab/avengeme	Estimación da potencia estatística das PRS	Estudo 4

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

EB-PRS	S. Song et al., 2020	https://github.com/shuangsong0110/EBPRS	Estimación das ponderacións dos SNP para PRS baseándose en inferencia bayesiana	Estudo 4
PLINK	C. C. Chang et al., 2015	https://www.cog-genomics.org/plink/	Análises xenómicos, entre eles, *clumping* e estimación de PRS	Estudo 4
SMTPred	Maier et al., 2018	https://github.com/uqrmaie1/smtpred	Combinación de distintas PRS correlacionadas nunha única PRS	Estudo 4
LDlink	Machiela & Chanock, 2015	https://ldlink.nih.gov/?tab=home	Consulta do LD entre SNP	Estudo 4
GSMR implementado en GCTA	Zhu et al., 2018	https://yanglab.westlake.edu.cn/software/gsmr/	Estimación do efecto causal dun trazo sobre outro mediante MR	Estudo 5
mtCOJO implementado en GCTA	Byrne et al., 2021; Zhu et al., 2018	https://yanglab.westlake.edu.cn/software/gcta/#mtCOJO	Análise condicional dun GWAS para un trazo polo efecto causal doutro trazo	Estudo 5
MAGMA implementado en FUMA	de Leeuw et al., 2015	https://fuma.ctglab.nl	Análise de enriquecemento en conxuntos de xenes	Estudo 5

[a]: Estudo no que se describe a aplicación informática e a metodoloxía asociada.

[b]: Ligazón para acceder ou descargar a aplicación informática.

[c]: Análises para as que se empregou a aplicación informática.

[d]: Estudos da tese nos que se empregou a aplicación informática.

3.2.1. Obtención de módulos de xenes co-expresados

No Estudo 1, empregouse a ferramenta *Weighted Gene Co-expression Network Analysis* (WGCNA) para identificar grupos de xenes cunha expresión altamente correlacionada en diferentes mostras (B. Zhang & Horvath, 2005). Esta análise foi levada a cabo co paquete de R desenvolvido para este fin (Langfelder & Horvath, 2008). WGCNA constrúe módulos de xenes co-expresados a partir de datos de expresión ao longo de todo o xenoma, como aqueles obtidos mediante secuenciación de ARN ou *microarray*, identificando grupos de xenes cun patrón de expresión similar entre eles e diferente dos outros módulos. Neste estudo, aplicouse WGCNA a datos de expresión de *microarray* de peixe cebra baixo diferentes condicións. A partir do resultado de WGCNA, identificáronse módulos de xenes co-expresados cuxo patrón de expresión estaba asociado a determinadas condicións nas mostras analizadas.

3.2.2. Identificación de loci significativos e mapeado xénico

Nos estudos 1 e 5, empregouse a ferramenta *SNP2GENE* de *FUMA* para identificar loci independentes, cun LD menor a un limiar predefinido, e significativos nun determinado GWAS (Watanabe et al., 2017). No Estudo 5 tamén se empregou *FUMA* para obter os xenes que mapean con cada un dos loci significativos do GWAS. Os xenes foron seleccionados baseándose en dous criterios: por posición e por expresión xénica. En concreto, identificáronse

USC DE COMPOSTELA

tanto os xenes próximos aos SNP dos loci de interese, como aqueles xenes para os cales os SNP deses loci actúan como eQTL en tecidos cerebrais.

3.2.3. Análise de enriquecemento en conxuntos de xenes

Co fin de caracterizar a nivel funcional as asociacións identificadas nun GWAS ou uns determinados xenes, nos Estudos 1 e 5 da tese avaliouse o seu nivel de enriquecemento en conxuntos de xenes de interese. Os principais conxuntos de xenes analizados foron, entre outros, xenes diferencialmente expresados en cada un dos tecidos do cerebro humano, identificados por GTEx (GTEx Consortium, 2020); xenes diferencialmente expresados en cada un dos estadios do desenvolvemento do cerebro humano, identificados por *BrainSpan* (http://www.brainspan.org); e xenes implicados nun mesmo proceso biolóxico, compoñente celular ou función molecular segundo os termos *Gene Ontology* (Aleksander et al., 2023).

No Estudo 1, examinouse se os xenes dos distintos módulos de co-expresión identificados por WGCNA están máis representados en cada un dos conxuntos de xenes de interese en comparación con outros xenes expresados no cerebro realizando unha proba hiperxeométrica. Ademais, neste estudo, a análise de enriquecemento en conxuntos de xenes de *Gene Ontology* realizouse co paquete de R *WebGestaltR*, avaliando tanto conxuntos de xenes do peixe cebra como do humano (Liao et al., 2019).

No Estudo 5, analizouse se o nivel de asociación dos xenes nun GWAS de interese está significativamente correlacionado coa pertenza a cada un dos conxuntos de xenes. Con este fin, neste estudo, empregouse a metodoloxía *Multi-marker Analysis of GenoMic Annotation* (MAGMA) (de Leeuw et al., 2015), implementada na ferramenta *SNP2GENE* de *FUMA* (Watanabe et al., 2017), a cal asigna cada SNP do GWAS a un xene e, a partir desa información, realiza a análise de enriquecemento tendo en conta o nivel de asociación do SNP no GWAS.

3.2.4. Estimación de puntuacións de risco polixénico

Nesta tese, estimáronse PRS nos Estudos 1 e 4, empregando, en cada un deles, diferentes mostras de descubrimento, mostras diana e métodos para a súa estimación. No Estudo 1, analizouse a asociación dunha PRS baseada en xenes de módulos de co-expresión cunha expresión alterada baixo a abstinencia a substancias coa dependencia a alcol dunha mostra diana. No Estudo 4, examinouse a asociación de diferentes PRS con tabaquismo ou con esquizofrenia dividindo a mostra diana en distintos subgrupos.

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

3.2.4.1. Mostras diana

A mostra de dependencia a alcol empregada no Estudo 1 estivo formada por 1.253 individuos de Galicia e Asturias, dentro dos cales, 524 foron diagnosticados por médicos especialistas en psiquiatría con dependencia a alcol seguindo os criterios diagnósticos do DSM-IV, e 729 foron controis. No Estudo 4, analizouse unha mostra diana formada por 2.692 individuos recollida entre 1999 e 2017 en diferentes lugares de España: Santiago de Compostela, Oviedo, Reus, Valencia e Madrid. Desta mostra, 1.108 individuos foron diagnosticados con esquizofrenia, segundo os criterios do DSM-IV, e 1.584 foron controis, dos cales o 75% eran doadores de sangue non cribados e o resto individuos cribados para descartar trastornos psiquiátricos. Ademais, para esta mostra, determinouse o hábito tabáquico actual dos individuos, é dicir, se son fumadores/as actualmente ou non, mediante entrevista clínica ou, no caso de non dispor desta información, mediante a cuantificación da metilación do xene *AHRR*, a cal está fortemente asociada co hábito tabáquico. Así, para esta mostra coñecemos tanto o diagnóstico de esquizofrenia como o hábito tabáquico, algo esencial para este estudo.

Todos os participantes destes dous estudos firmaron un consentimento informado. Os estudos foron aprobados polo Comité Ético de Investigación Clínica (CEIC) de Galicia e realizáronse conforme ao marco ético da última versión da Declaración de Helsinki. As mostras de ADN de dependencia a alcol e de esquizofrenia pertencen ás coleccións de Javier Costas, director desta tese. A colección de dependencia a alcol está rexistrada co número 0003122 no rexistro nacional de biobancos. A súa creación foi aprobada polo CEIC de Galicia (código de rexistro: 2012/368) para o seu uso en investigación sobre as bases xenéticas das drogodependencias. A colección de esquizofrenia está rexistrada co número C.0000877 no rexistro nacional de biobancos, dende o 23 de novembro de 2012, para o estudo das bases xenéticas da esquizofrenia. As mostras recolléronse no marco dos seguintes proxectos que recibiron ditame favorable do CEIC (Anexo I):

Título: Secuenciación exónica de pacientes esquizofrénicos con microdelecións e microduplicacións que confiren risco elevado a padecer esquizofrenia e outros trastornos do neurodesenvolvemento. Entidade financiadora: ISCIII. Código: CP11/00163. Duración: 2012-2014. Código de Rexistro do CEIC de Galicia: 2011/276. Data de aprobación do CEIC de Galicia: 26/07/2011.

Título: Caracterización do papel de microdelecións/microduplicacións do xenoma como determinantes biolóxicos do policonsumo de drogas. Sen financiación. Código de rexistro do CEIC de Galicia: 2012/368. Data de aprobación do CEIC de Galicia: 27/11/2012.

Título: Risco polixénico a desenvolver esquizofrenia a partir de módulos de co-expresión espazo-temporal en cerebro: implicacións etiopatoxénicas e clínicas. Entidade financiadora: ISCIII. Código: PI14/01020. Duración: 2015-2017. Código de Rexistro do CEIC de Galicia: 2014/369. Data de aprobación do CEIC de Galicia: 21/07/2014.

Título: Contribución da variación polixénica común á predición do risco a dependencia de alcol. Entidade financiadora: Delegación do Goberno para o Plan Nacional Sobre Drogas. Código: 2014|075. Duración: 01/02/2015-28/02/2018. Código de Rexistro do CEIC de Galicia: 2014/428. Data de aprobación do CEIC de Galicia: 20/10/2014.

Título: Unha aproximación xenómica e epixenómica á asociación entre tabaco e esquizofrenia. Entidade financiadora: ISCIII. Código: PI17/01424. Duración: 2018-2020. Código de Rexistro do CEIC de Santiago-Lugo: 2018/095. Data de aprobación do CEIC de Santiago-Lugo: 22/02/2018.

3.2.4.2. Mostras de descubrimento

Para ambos estudos, empregáronse GWAS daqueles recollidos na Táboa 3 como mostras de descubrimento. No Estudo 1, analizáronse tres PRS baseadas en GWAS para trazos relacionados co consumo de alcol. En concreto, avaliouse a PRS para dependencia a alcol, para a puntuación total do test de identificación de trastornos polo uso de alcol e para a intensidade do consumo de alcol. Entre estes GWAS, elixiuse aquel cuxa PRS estivese máis asociada a dependencia a alcol na mostra diana como mostra de descubrimento para a PRS baseada nos xenes de módulos de co-expresión específicos. No Estudo 4, empregouse un GWAS para o inicio do consumo de tabaco como principal mostra de descubrimento para tabaquismo, debido á súa similitude co fenotipo de hábito tabáquico recollido na mostra diana. Ademais, empregáronse GWAS para trazos xeneticamente correlacionados co inicio do consumo de tabaco cuxa PRS tivese unha potencia estatística suficientemente elevada na análise, o cal foi avaliado empregando o paquete de R *AVENGEME* (Palla & Dudbridge, 2015).

3.2.4.3. Métodos de estimación das puntuacións de risco polixénico

No Estudo 1, a PRS da mostra diana foi estimada empregando o software *PRSice-2* (Choi & O'Reilly, 2019). Este software estima PRS aplicando o método tradicional de *clumping,* para identificar aqueles SNP independentes cunha asociación no GWAS máis significativa, e *thresholding,* para testar a asociación de múltiples PRS para diferentes limiares de significación na mostra de descubrimento co fenotipo de interese, neste caso, a dependencia a alcol, na mostra diana. Neste mesmo estudo, empregouse unha nova ferramenta de *PRSice-2*, denominada

PRSet, que permite estimar PRS baseadas en conxuntos de xenes, é dicir, baseadas nos SNP dunhas rexións xénicas concretas (Choi et al., 2023). PRSet realiza un proceso de *clumping* característico que permite ter en conta o LD maximizando o sinal dentro de cada conxunto de xenes. Ademais, por defecto, non establece limiares de significación para o cálculo de PRS, senón que considera todos os SNP independentes dentro das rexións xénicas especificadas para non empregar un número demasiado baixo de SNP na análise. Neste caso, esta metodoloxía permitiu estimar PRS baseadas en SNP próximos ou dentro de xenes de módulos de co-expresión específicos.

No Estudo 4, as PRS foron estimadas empregando *EB-PRS*, un método de estimación de PRS bayesiano (S. Song et al., 2020). Este método infire a distribución xeral do tamaño de efecto dos SNP na mostra de descubrimento e logo emprega esta distribución para estimar o tamaño de efecto dos SNP maximizando a precisión da predición. Ao contrario doutros métodos que permiten mellorar a predición das PRS, *EB-PRS* non precisa información externa nin a optimización de parámetros. Para estimar as PRS, combinouse este método con *PLINK*, unha ferramenta bioinformática para a análise de GWAS (C. C. Chang et al., 2015). *PLINK* foi empregada para realizar o procedemento de *clumping* e, desta forma, elixir os SNP que se incluirían na PRS, así como para calcular a PRS da mostra diana tendo en conta o tamaño de efecto dos SNP estimado por *EB-PRS*. No Estudo 4, tamén se empregou *SMTPred* para combinar, nunha mesma PRS, a PRS para o inicio do consumo de tabaco e PRS para trazos xeneticamente correlacionados (Maier et al., 2018).

3.2.4.4. Análises estatísticas das puntuacións de risco polixénico

Tanto no Estudo 1 como no Estudo 4, analizouse a asociación das PRS estimadas co fenotipo de interese na mostra diana mediante regresión loxística, na cal se incluíron outras covariables como o sexo, a idade ou compoñentes principais de ascendencia. Ademais, estimouse a proporción de varianza do fenotipo de interese explicada pola PRS como a diferenza na R^2 de Nagelkerke entre o modelo completo e o modelo sen a PRS. Esta R^2 de Nagelkerke converteuse á R^2 proposta por S. H. Lee, Goddard, Wray, & Visscher (2012) para estimar a proporción de varianza explicada pola PRS na escala de risco corrixida polo nesgo derivado da sobremostraxe de casos na mostra diana en comparación coa prevalencia do trazo na poboación. Deste xeito, esta estima permite comparar a porcentaxe de varianza explicada por diferentes PRS. Para estimar esta R^2 na escala de risco, precisouse a proporción de casos na mostra diana e a prevalencia do trazo de interese na poboación, a cal foi obtida da bibliografía. Todas estas análises foron levadas a cabo en *R* (https://www.r-project.org).

USC DE SANTIAGO DE COMPOSTELA

No Estudo 1, *PRSet* permitiu comparar a asociación da PRS de interese cunha PRS de fondo que considera o mesmo número de SNP independentes, realizando permutacións para obter o nivel de enriquecemento da PRS de interese na asociación co trazo diana en comparación coa PRS de fondo (Choi et al., 2023). No noso caso, comparamos a PRS baseada en SNP próximos ou dentro dos xenes de determinados módulos de co-expresión cunha PRS baseada en SNP escollidos ao azar dentro daqueles próximos ou dentro de xenes expresados no cerebro.

3.2.5. Análise do patrón de solapamento xenético independente da correlación xenética

No Estudo 2, empregouse a ferramenta MiXeR para caracterizar o patrón de solapamento xenético entre dous trazos a partir das estatísticas resumo de GWAS (Frei et al., 2019). Primeiro, MiXeR estima o número de variantes causais involucradas en cada trazo analizado. A continuación, estima o número de variantes causais compartidas entre os dous trazos e a correlación do tamaño de efecto das variantes compartidas. Desta forma, a diferenza dos estudos de correlación xenética, MiXeR permite identificar a existencia de solapamento xenético entre dous trazos independentemente da dirección de efecto das variantes xenéticas para ambos trazos.

3.2.6. Análise de colocalización

Nos Estudos 3 e 5, empregáronse distintos métodos para avaliar se existe colocalización nun locus para dous trazos, é dicir, se o sinal de asociación de GWAS para os dous trazos nese locus é debido á mesma variante causal. Estes métodos permiten diferenciar a colocalización da existencia de sinais de asociación independentes asociadas a cada un dos trazos no locus. Ademais, todos eles poden analizar directamente as estatísticas resumo dos GWAS. Estes métodos de colocalización computan a probabilidade posterior de cinco hipóteses alternativas: H_0, o locus non está asociado con ningún dos dous trazos; H_1, existe un sinal de asociación no locus co trazo 1 pero non co trazo 2; H_2, existe un sinal de asociación no locus co trazo 2 pero non co trazo 1; H_3, existen dous sinais de asociación independentes no locus, unha para cada trazo; e H_4, existe un sinal de asociación común aos dous trazos no locus.

No Estudo 3, empregouse como metodoloxía principal o método bayesiano *coloc* (Giambartolomei et al., 2014). Porén, esta metodoloxía presenta algunhas limitacións. *Coloc* non permite avaliar a existencia de colocalización para variantes xenéticas secundarias en LD coa variante xenética máis significativa do locus. Para superar esta limitación, no Estudo 3, empregáronse metodoloxías adicionais. Por un lado, *LocusCompare* foi empregado para

visualizar o sinal de asociación dos dous GWAS no locus tendo en conta, ao mesmo tempo, o LD (B. Liu et al., 2019). Esta representación permitiu diferenciar a existencia de varios sinais de asociación nun mesmo locus. Por outro lado, empregouse o método *masking* dentro da versión de *coloc* 4.0-2 que permite avaliar a existencia de colocalización para diferentes SNP en LD (Wallace, 2020). Ademais, *coloc* só permite comparar GWAS da mesma ascendencia e sen solapamento das mostras. Esta segunda restrición limita moito as análises xa que moitos dos GWAS empregados baséanse, polo menos parcialmente, en mostras comúns como *UK Biobank*. Polo tanto, tanto no Estudo 3 como unha análise de sensibilidade, como no Estudo 5 como principal metodoloxía de colocalización, fíxose uso de dúas ferramentas: *fgwas* e *gwas-pw* que, conxuntamente, nos permiten avaliar a existencia de colocalización do sinal de asociación entre GWAS con solapamento de mostras (Pickrell et al., 2016).

3.2.7. Análise de aleatorización mendeliana

No Estudo 5, empregouse o método de MR *generalized summary data-based Mendelian randomization* (GSMR) para identificar, dentro de diferentes comportamentos de consumo de tabaco, aquel co maior efecto causal sobre a esquizofrenia (Zhu et al., 2018). GSMR, implementado en *Genome-wide Complex Trait Analysis* (GCTA) (J. Yang et al., 2011), emprega varios SNP independentes asociados ao trazo de exposición como variables instrumentais. A principal vantaxe desta metodoloxía é que descarta aqueles SNP cun efecto pleiotrópico sobre ambos trazos co test *HEIDI-outlier*. Este test baséase na idea de que, en ausencia de pleiotropía horizontal, o efecto causal estimado empregando calquera dos SNP debe ser idéntico. Polo tanto, descarta como variables instrumentais aqueles SNP que estiman un efecto causal desviado do proporcionado polo resto de SNP. Para poder comparar o efecto causal de diferentes comportamentos de consumo de tabaco sobre a esquizofrenia, o efecto causal transformouse á escala de risco (Byrne et al., 2021).

3.2.8. Análise condicional de estudos de asociación do xenoma completo

No Estudo 5, empregouse como metodoloxía principal *multi-trait-based conditional and joint analysis* (mtCOJO) (Byrne et al., 2021; Zhu et al., 2018), implementada en GCTA (J. Yang et al., 2011). O obxectivo deste estudo foi condicionar o GWAS orixinal de esquizofrenia polo comportamento de consumo de tabaco elixido na análise de MR previa. En concreto, mtCOJO proporciona novas estimas dos efectos dos SNP do GWAS condicionando o seu efecto sobre a esquizofrenia polo efecto causal do consumo de tabaco sobre esta. Polo tanto, estas novas estimas de efecto dos SNP modificaranse nun factor proporcional ao efecto causal do tabaquismo sobre esquizofrenia. Desta forma, descartamos, dentro da susceptibilidade xenética

USC DE COMPOSTELA

a esquizofrenia, aquela susceptibilidade xenética asociada a esquizofrenia a través da vía causal do tabaquismo.

3.2.9. Aplicacións de LD score regression

O método LDSC empregouse ao longo de toda a tese para estimar a herdabilidade baseada en SNP dos trazos, a correlación xenética entre trazos e o grao de solapamento das mostras de distintos GWAS (B. Bulik-Sullivan et al., 2015). Esta metodoloxía foi utilizada de forma máis extensa no Estudo 5 da tese para avaliar o cambio na correlación xenética da esquizofrenia con diferentes trazos de interese tras condicionar por tabaquismo. En todas as análises de LDSC, empregáronse as puntuacións de LD calculadas na mostra europea da fase 3 de *The 1000 Genomes*. Ademais, a herdabilidade baseada en SNP transformouse á escala de risco para os trazos dicotómicos, empregando a proporción de casos no GWAS e a prevalencia na poboación.

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

4. RESULTADOS

4.1. ESTUDO 1: IDENTIFICATION OF GENE CO-EXPRESSION MODULES FROM ZEBRAFISH BRAIN DATA: APPLICATIONS IN PSYCHIATRY ILLUSTRATED THROUGH ALCOHOL-RELATED TRAITS

Referencia:

Al-Soufi, L., Arana, A. J., Facal, F., Flórez, G., Vázquez, F. L., Arrojo, M., the GenPol Study Group, Sánchez, L. e Costas, J. (2024). Identification of gene co-expression modules from zebrafish brain data: applications in psychiatry illustrated through alcohol-related traits [Traballo non publicado].

Información dos/as autores/as:

Laila Al-Soufi[1,2], Álvaro J. Arana[2], Fernando Facal[1,3], Gerardo Flórez[4,5], Fernando L. Vázquez[6], Manuel Arrojo[3], the GenPol Study Group, Laura Sánchez[2] e Javier Costas[1,7]

[1]Grupo de Xenética Psiquiátrica, Instituto de Investigación Sanitaria de Santiago de Compostela (IDIS), Santiago de Compostela, Galicia, España. Rede de Investigación en Atención Primaria de Adiccións (RIAPAd).

[2]Departamento de Zooloxía, Xenética e Antropoloxía Física, Facultade de Veterinaria, Universidade de Santiago de Compostela, Lugo, España.

[3]Servizo de Psiquiatría, Complexo Hospitalario Universitario de Santiago de Compostela, Servizo Galego de Saúde (SERGAS), Santiago de Compostela, Galicia, España.

[4]Unidade para o Tratamento de Condutas Adictivas, Hospital Universitario de Ourense, Ourense, Galicia, España.

[5]Centro de Investigación Biomédica en Rede de Saúde Mental (CIBERSAM), Oviedo, España.

[6]Departamento de Psicoloxía Clínica e Psicobioloxía. Universidade de Santiago de Compostela, Santiago de Compostela, España.

[7]Complexo Hospitalario Universitario de Santiago de Compostela (CHUS), Servizo Galego de Saúde (SERGAS), Santiago de Compostela, Galicia, España.

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

Contribución da doutoranda na publicación:

Realización de todas as análises salvo o tratamento dos datos de xenotipado da mostra de dependencia a alcol e a revisión de modelos de peixe cebra. Redacción do manuscrito inicial e revisión da versión final.

USC
UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

ABSTRACT

Cumulative evidence suggests that zebrafish is a useful model in psychiatric research. Weighted Gene Co-expression Network Analysis (WGCNA) enables the reduction of genome-wide expression data to modules of highly co-expressed genes, which are hypothesized to interact within molecular networks. In this study, we first applied WGCNA to zebrafish brain expression data across different experimental conditions. Then, we characterized the different co-expression modules by gene-set enrichment analysis and hub gene-phenotype association. Finally, we analyzed association of polygenic risk scores (PRSs) based on genes of some interesting co-expression modules with alcohol dependence in 524 patients and 729 controls from Galicia, using competitive tests. Our approach revealed 34 co-expression modules in the zebrafish brain, with some showing enrichment in human synaptic genes, brain tissues, or brain developmental stages. Moreover, certain co-expression modules were enriched in psychiatry-related GWAS and comprised hub genes associated with psychiatry-related traits in both human GWAS and zebrafish models. Expression patterns of some co-expression modules were associated with the tested experimental conditions, mainly with substance withdrawal and cold stress. Notably, a PRS based on genes from co-expression modules exclusively associated with substance withdrawal in zebrafish showed a stronger association with human alcohol dependence than PRSs based on randomly selected brain-expressed genes. In conclusion, our analysis led to the identification of co-expressed gene modules that may model human brain gene networks involved in psychiatry-related traits. Specifically, we detected a cluster of co-expressed genes whose expression was exclusively associated with substance withdrawal in zebrafish, which significantly contributed to alcohol dependence susceptibility in humans.

USC
UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

1. INTRODUCTION

Genetic susceptibility to psychiatric disorders exhibit a complex etiology characterized by a highly polygenic architecture, i.e., numerous common genetic variants contribute to the disorder risk with a small effect size (Sullivan et al., 2018). Over the past decade, the study of genetic risk factors underlying psychiatric disorders has largely relied on genome-wide association studies (GWAS). These studies analyze the effect of common genetic variation across the genome on a given trait (Andreassen et al., 2023). With the increasing sample sizes of GWAS, a growing number of common genetic variants associated with psychiatric disorders have been found. The information gathered in GWAS can be used to assess an individual's genetic susceptibility to a psychiatric disorder through polygenic risk scores (PRS), i.e., the cumulative sum of an individual's genetic risk variants, each weighted by its estimated effect in a GWAS on an independent discovery sample. However, understanding the molecular pathways linking GWAS variants to psychiatric disorders remains challenging. This is primarily because a considerable proportion of GWAS associations are detected in non-coding elements that regulate gene expression, rather than in protein coding sequences (Jaffe et al., 2018; Maurano et al., 2012).

One potential strategy to link the effect of GWAS regulatory variants with the underlying molecular mechanism causing the psychiatric disorder involves the identification of those genes with altered expression in that psychiatric disorder (Y. Chen et al., 2008). To unravel the intricate interactions among the multitude of genes expressed in the brain and identify group of genes involved in the same molecular pathways, gene networks acting in the brain can be constructed. Weighted Gene Co-expression Network Analysis (WGCNA) is a widely used method for analyzing co-expression modules, allowing the reduction of genome-wide expression data to clusters of highly co-expressed genes (B. Zhang & Horvath, 2005). These co-expression modules, characterized by similar expression patterns across samples, are hypothesized to be influenced by the same regulatory system, comprising genes that interact in a molecular network and are functionally related (H. K. Lee et al., 2004). The validity of this approach has been confirmed by previous studies that constructed gene networks with functional interpretations in the human brain (Hawrylycz et al., 2012; Kang et al., 2011; Oldham et al., 2008). WGCNA has also been successfully used to identify brain co-expression modules with altered expression in schizophrenia (Fromer et al., 2016), gene modules displaying dynamic spatiotemporal expression in the human brain enriched in genes involved in neuropsychiatry traits (Li et al., 2018), and co-expressed genes associated with autism (Voineagu et al., 2011). WGCNA was also used to investigate potential new drug targets for

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

schizophrenia by identifying highly interconnected genes within co-expression modules, i.e., hub genes, associated with this trait, focusing on those with expression during adolescence and adulthood (Rodriguez-López et al., 2020). However, this approach is constrained by the challenge to obtain transcriptomic data from the human brain under diverse conditions.

Zebrafish stands out as a promising model for studying mental disorders, including psychiatric disorders, substance use and dependence, and behavioral traits, thanks to its numerous advantages (Fontana et al., 2018; Mathur & Guo, 2010; Stewart et al., 2015). With the aid of Clustered Regularly Interspaced Short Palindromic Repeat sequences/CRISPR-associated protein (CRISPR-Cas), zebrafish offers an effective tool for validating GWAS genetic variants through loss of function or mutation of target genes (L. M. Saunders et al., 2023; Wang et al., 2023). Human nervous system is highly conserved at zebrafish, with a remarkably similar brain architecture between both species and a preservation of all neuron subtypes (McCammon & Sive, 2015; Wang et al., 2023). Processes related to substance addiction are also conserved in zebrafish, including the activation of the reward pathway by drugs of abuse, a drug-seeking behavior, and neuroadaptations following chronic drug exposure (Müller et al., 2020; Stewart et al., 2015). Furthermore, a variety of behavioral tests have been developed in zebrafish, covering complex behaviors associated with cognition, drug addiction and emotion (Kalueff et al., 2014; Müller et al., 2020). Notably, the conservation of genes associated with psychiatric disorders in zebrafish surpasses the general rate of gene conservation (Franklin & Dwyer, 2021; Kasap et al., 2018). Thus, transcriptomic data from the zebrafish brain may be useful to model human brain gene networks underling psychiatry-related traits.

The main objective of the present study was to identify sets of co-expressed genes potentially implicated in psychiatry-related traits, including psychiatric disorders, substance use and dependence traits and personality traits, through the analysis of zebrafish brain expression data. Specifically, our approach involved: 1) the identification of brain co-expression modules in zebrafish using WGCNA; 2) the characterization of these co-expression modules by gene-set enrichment analysis and hub gene - phenotype association; 3) the identification of co-expression modules whose expression pattern was associated with specific experimental conditions; and 4) as a proof of concept, given that substance withdrawal was one of the experimental conditions in our data, the investigation of whether co-expression modules associated with substance withdrawal significantly contribute to the genetic susceptibility to alcohol dependence (AD) in humans.

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

2. METHODS

2.1. Gene expression data

Bibliography was reviewed to identify studies with publicly available gene expression data of the adult zebrafish brain. These studies were required to assess gene expression in at least four samples, including both samples under some experimental condition and control samples. We have included brain expression data from various studies assessing gene expression under diverse conditions to increase the variability of our expression data and the statistical power of our co-expression analysis (H. K. Lee et al., 2004). To facilitate data combination, studies using microarray technology to assess gene expression were limited to those using the Affymetrix GeneChip Zebrafish Genome Array, given its widespread use.

2.2. Data preparation

For WGCNA analysis, it is highly recommended to filter probes based on expression levels to avoid correlations attributed to the background signal of non-expressed genes and ensure sufficient variation across samples (Eising et al., 2019; Rodriguez-López et al., 2020). In our case, variation in the expression data could not be accurately estimated due to the low number of samples in each study. To exclude low-expressed genes that may introduce noise in the analysis, we identified probes with expression levels surpassing the median expression level of the sample in at least half of the samples for any study and given condition. The resulting set of probes expressed in any study was then used for WGCNA analysis.

To integrate expression data from different studies, a normalization process was implemented. First, expression data were transformed to logarithmic scale. Following this, each sample was standardized based on its mean and standard deviation. To address systematic biases introduced by combining data from distinct studies ("batch effects"), we applied the Combat function from the sva R package (W. E. Johnson et al., 2007). ComBat eliminates batch effects by adjusting for variations between conditions to preserve differences across conditions. Finally, for each condition and study independently, quantile normalization was performed using the preprocessCore R package (Bolstad et al., 2003).

2.3. Weighted gene co-expression network analysis (WGCNA)

WGCNA was conducted using the WGCNA R package (Langfelder & Horvath, 2008). The analysis focused on the expression levels of all the filtered probes. Following WGCNA guidelines, we first chose a soft-thresholding power of 7 based on the criterion of approximate scale-free topology (scale-free topology fit $R^2 = 0.9373$) (Supplementary Figure 1) (B. Zhang & Horvath, 2005). We then performed network construction and module detection. As

USC UNIVERSIDADE DE SANTIAGO

recommended, we considered a signed network type, a minimal module size of 30, a dendrogram cut height for module merging of 0.25 and bidweight midcorrelation with a maximum proportion of outliers of 0.05. Expression level of all the probes was assessed simultaneously.

WGCNA identified co-expression modules, i.e., groups of probes exhibiting highly positive correlation of expression profiles across samples and sharing a similar co-expression pattern with other probes (B. Zhang & Horvath, 2005). Each probe is assigned to a co-expression module or to a group of probes (M0) without enough correlation with any module. Module membership (MM) for each probe in each co-expression module is calculated as the correlation between the expression level of the probe across the samples and the module eigengene. The module eigengene corresponds to the first principal component of the expression levels of all probes within the module, which is representative of the expression profiles in the co-expression module. MM measures the connectivity of the probe within a module, indicating the similarity between its expression level and the characteristic pattern of expression of the co-expression module. Probes with an absolute MM > 0.85 were considered hub probes of each co-expression module.

2.4. Probe to zebrafish gene and human gene conversion

Following the WGCNA analysis, probes were converted to zebrafish and human gene symbols for subsequent analyses. This conversion was carried out using the Biomart tool from Ensembl (Smedley et al., 2009). For zebrafish gene symbols, the probes were directly mapped to ZFIN gene symbol (Bradford et al., 2023). For human gene symbols, the probes were initially converted to zebrafish gene stable IDs from Ensembl. Following this, zebrafish gene IDs were transformed into human gene symbols.

2.5. Co-expression modules enrichment analysis

We assessed the enrichment of the co-expression modules in non-redundant Gene Ontology (GO) terms, including biological processes, cellular components, and molecular functions, using the WebGestaltR R package (Liao et al., 2019). As zebrafish GO terms are less extensively annotated, we analyzed module enrichment in both zebrafish and human GO terms. We also examined enrichment in 1605 synaptic genes annotated by synGO (Koopmans et al., 2019); differentially expressed genes in human brain tissues using GTEx v8 data available in FUMA (Watanabe et al., 2017); differentially expressed genes in human brain developmental stages using BrainSpan data available in FUMA (Watanabe et al., 2017); and genes mapping with GWAS associations for traits relevant to psychiatry. Specifically, enrichment in psychiatric disorders GWAS from the Psychiatric Genomics Consortium (PGC) and public

USC DE SANTIAGO DE COMPOSTELA

GWAS for substance use, substance dependence and personality traits was assessed (Supplementary Table 1). Significant independent SNPs at $r^2 < 0.1$ and $P < 1x10^{-5}$ were obtained for these GWAS using FUMA (Watanabe et al., 2017). The nearest gene to each of these SNPs within a distance less than 250 kb was determined using the vautils R package.

Co-expression module enrichment analysis was conducted by calculating the enrichment of the genes in the co-expression modules in the gene-sets of interest using a one-sided Fisher exact test. False discovery rate (FDR) correction for multiple-testing was applied for each module and category, i.e., biological processes GO terms, cellular components GO terms, molecular functions GO terms, brain tissues, brain developmental stages and GWAS, separately. In all enrichment analyses, the background genes consisted of those corresponding to the probes used for WGCNA analysis, excluding those belonging to the tested module. Gene-sets with fewer than 10 annotated genes were excluded from analysis. In the case of GO terms, gene-sets with fewer than 500 annotated genes were selected to avoid considering overly broad gene-sets.

2.6. Co-expression modules association with experimental conditions

Association between the expression pattern of each co-expression module, measured as the module eigengene, and each experimental condition was assessed by a logistic regression model. The proportion of variance explained by the module eigengene was estimated as the Nagelkerke's R^2 and FDR correction for multiple-testing of the P-value of the logistic regression model was applied for each module for the number of conditions tested.

2.7. Co-expression modules based polygenic risk score analysis

We evaluated the variance explained of AD by PRSs based on genes from specific co-expression modules in comparison with a background PRS of brain-expressed genes. PRS association with AD was analyzed in a sample of 1253 individuals of Northwest Iberian origin, comprising 524 cases diagnosed with AD according to DSM-IV criteria and 729 controls. Our target sample showed a relatively high comorbidity with other substance dependence and psychiatric disorders. Genotyping was performed using the Infinium PsychArray Bead-Chip (Illumina, San Diego, California), and genotype imputation was carried out at the Michigan Imputation Server (Das et al., 2016). Detailed information of the target sample, as well as genotyping, imputation, and quality control procedures, can be found in Facal et al. (2021). All participants signed an informed consent for this study. The study was performed in accordance with the latest version of the Helsinki Declaration of 1975, as revised in 2008, and has been approved by the Galician Ethical Committee for Clinical Research.

USC DE SANTIAGO DE COMPOSTELA

In order to select which GWAS to use for the estimation of effect sizes in PRS, three distinct whole-genome PRSs were analyzed based on different alcohol-related phenotypes, including AD (N cases = 11,569, N controls = 34,999) (Walters et al., 2018), the total score of the Alcohol Use Disorder Identification Test (AUDIT-T) (N = 121,604) (Sanchez-Roige et al., 2019) and the average number of drinks drunk each week (DrnkWk) (N = 666,978) (G. R. B. Saunders et al., 2022). Among these GWAS, the GWAS with the most significant association with AD in our target sample was selected for estimating PRS based on co-expression modules.

PRS of the target sample was estimated using PRSice-2 with default parameters (Choi & O'Reilly, 2019). PRS was calculated as the sum of effect alleles carried by each individual weighted by their effect sizes estimated in the GWAS. PRSice-2 implemented clumping to identify independent SNPs with a more significant association in the GWAS, using a linkage disequilibrium threshold of $r^2 = 0.1$ and a window size of 250 kb. PRSice-2 calculated PRS considering those SNPs exceeding a P-value threshold in the GWAS. Multiple P-value thresholds were assessed, and the most significant P-value threshold was identified. The association of PRS with AD was assessed using a logistic regression model including sex, age, study phase, and the first 10 principal components of ancestry as in Facal et al. (2021). The percentage of variance explained by each PRS was estimated as the increase in Nagelkerke's R^2 between a full model and a model without the PRS. To adjust the R^2 for ascertainment bias, a prevalence of AD of 13.8% was assumed (D. S. Hasin et al., 2007).

PRSet, a novel feature of PRSice-2 for gene-set based PRS analysis, was used to estimate PRS based on the genes comprising specific co-expression modules using the previously selected GWAS (Choi et al., 2023). PRSet performed set-based clumping to account for linkage disequilibrium and maximize signal within each gene-set. PRS estimation was restricted to SNPs located in the gene regions (±10 kb) at a P-value threshold of 1, as recommended (Choi et al., 2023). The association of the gene-set PRS with the target trait was calculated. To assess enrichment, association of our PRS was compared with a background PRS considering the same number of independent SNPs. PRSet conducts 10,000 permutations to obtain a competitive P-value, indicating the level of enrichment of the gene-set PRS in association with the target trait compared to the background PRS.

2.8. Hub genes revision

Hub genes demonstrating a significant association with psychiatry-related traits at $P < 5x10^{-8}$ were identified using data from the GWAS Catalog. An extensive literature research was carried out using resources such as ZFIN (https://zfin.org), ALLIANCE (https://www.alliancegenome.org), and keyword searches in NCBI to identify relevant

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

publications describing genetic depletion zebrafish models for our hub genes associated with neuropsychiatry phenotypes. We included studies using knockout models by CRISPR/Cas9 or ENU mutagenesis, as well as knockdown models by Morpholinos, to investigate these associations.

3. RESULTS

3.1. Source of brain expression data

Four microarray studies met our requirements (Arslan-Ergul & Adams, 2014; Drew et al., 2008; Kily et al., 2008; R. Yang et al., 2011) (Table 1). These studies examined the effect on gene expression of ethanol and nicotine withdrawal (Study 1 (S1) (Kily et al., 2008)), starvation (Study 2 (S2) (Drew et al., 2008)), cold stress (Study 3 (S3) (R. Yang et al., 2011)) and aging in both sexes (Study 4 (S4) (Arslan-Ergul & Adams, 2014)) (Table 1). Gene expression datasets corresponding to these studies are available in the ArrayExpress repository (S1: E-MEXP-1301) or the GEO database (S2: GSE11107, S3: GSE32092, and S4: GSE53430). In these microarray data, expression levels of 15,617 probes in a total of 31 samples were analyzed (Supplementary Table 2). Given that sex was not specified for most samples, it was not considered as a distinct condition in S4 (Table 1). Two additional studies using RNAseq also met our criteria (Banu et al., 2020; Mao et al., 2022), but they were discarded as it was not technically possible to combine them with the microarray studies, and they analyzed a lower number of samples (23 samples) compared to microarray data.

Table 1. Microarray studies used in WGCNA analysis. Experimental conditions and samples included in each study are provided.

Study	Source	Experimental conditions	Number of conditions	Samples	Number of samples	Sex
Study 1	E-MEXP-1301 (Kily et al., 2008)	Conditioned to ethanol or nicotine for 20 min each day over a 4-week followed by 3 weeks of withdrawal	3	3 control samples, 2 ethanol withdrawal samples, 2 nicotine withdrawal samples	7	Not provided
Study 2	GSE11107 (Drew et al., 2008)	Fed three times daily or received no food for 21 days	2	4 control samples and 4 starved samples	8	Female
Study 3	GSE32092 (R. Yang et al., 2011)	Raised under the normal laboratory breeding temperature (28C) or subjected to cold stress (10C) for 10 days	2	2 samples under normal conditions and 2 samples under cold stress	4	Not provided
Study 4	GSE53430 (Arslan-Ergul & Adams, 2014)	Adult (7.5-8.5 months old) male and female and old adult (31-36 months old) male and female zebrafish	4	3 control male samples, 3 control female samples, 3 old adult male samples and 3 old adult female samples	12	6 Male and 6 female

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

3.2. Weighted gene co-expression network construction

After filtering based on expression levels, a total of 10,629 probes out of the initial 15,617 were retained for the subsequent WGCNA analysis. Data normalization effectively harmonized the expression data from the different studies, eliminating study effects while preserving condition-specific differences, as demonstrated by hierarchical clustering of the samples (Figure 1). The 31 samples included in WGCNA analysis clustered by study prior to data normalization (Figure 1A). After normalization, those samples subjected to ethanol and nicotine withdrawal exhibited the most distinct expression profiles from control samples, while samples from starvation and aging showed similar expression patterns to controls (Figure 1B). The two samples subjected to cold stress clustered together, close to control samples.

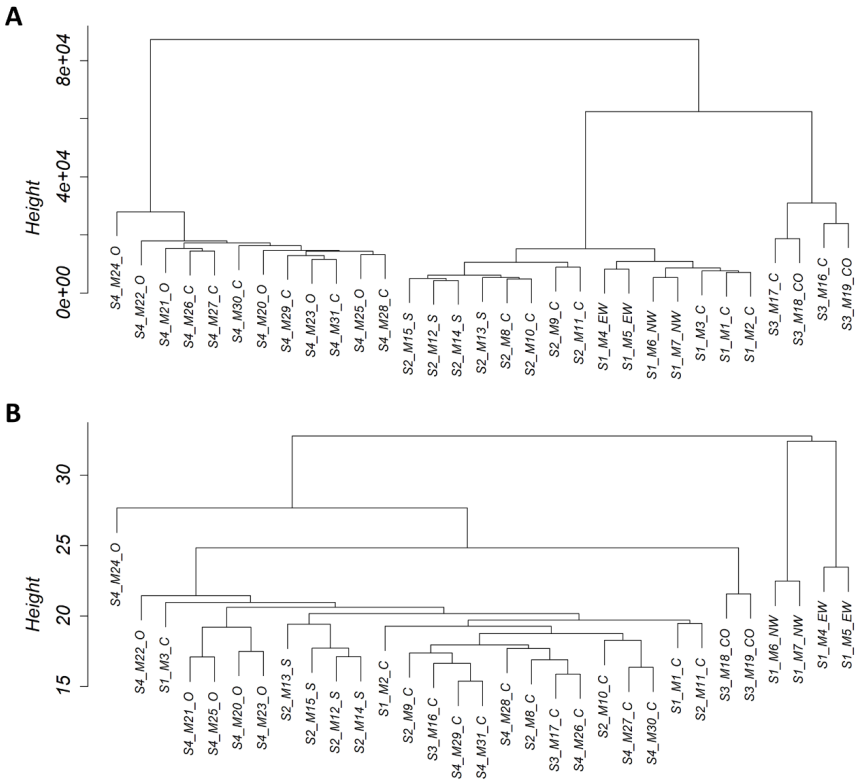


Figure 1. Sample hierarchical clustering using A) raw microarray data and B) normalized microarray data. Samples are described in Supplementary Table 2.

The application of WGCNA to the normalized expression levels resulted in the identification of 34 co-expression modules (M1-M34), ranging in size from 1,425 (M1) to 36 (M34) probes (Table 2, Supplementary Figure 2, Supplementary Table 3). Notably, 2,673 of the 10,629 probes were not assigned to any module (M0). 33 of the 34 co-expression modules had at least one hub probe. Only M28 did not have any hub probe.

Following the detection of co-expression modules, probes were converted to their corresponding zebrafish and human gene using Biomart. This conversion process identified

8,561 probes, with 8,368 and 7,221 probes assigned to zebrafish and human genes, respectively. Some probes mapped to the same gene, resulting in a total of 7,234 zebrafish genes and 6,195 human genes (Table 2, Supplementary Table 3).

Table 2. Co-expression modules detected by WGCNA analysis.

Module	N probes[a]	N human genes[b]	N zebrafish genes[c]	N hub probes[d]	N human hub genes[e]	N zebrafish hub genes[f]	Zebrafish hub genes
M1	1,425	1,061	1,089	63	54	58	*inka1a, pgm1, ythdf1, bcl2l1, osbpl2b, papola, ncam2, slc49a4, mef2d, ehf, faf2, ube2r2, ptn, hnf1ba, si:ch1073-157b13.1, dctpp1, si:ch211-114l13.1, dedd, tshz2, rnf146, dgcr6, ncor2, gspt1, iqsec3a, wdr44, isoc1, gria2a, ccnt1, tmem67, nlgn1, ptpn13, ppp3r1b, ftr90, ftr72, ftr73, ftr76, ammecr1, gna11b, pcdh10a, scd, psmc3ip, psmf1, lxn, si:dkey-21c19.3, txnrd3, ppp1r7, sympk, dcps, mxi1, hoxb5b, agpat2, poll, pafah1b1b, pias4a, fbxl3b, zgc:113208, ptpro, cyp4t8*
M2	1,272	860	960	105	38	57	*oprm1, itm2ba, slc9a6a, akt2l, taf1, wdr34, mcm5, si:dkeyp-98a7.5, si:dkeyp-98a7.7, si:dkeyp-98a7.4, si:dkeyp-98a7.8, si:dkeyp-98a7.9, si:dkeyp-98a7.3, smarcc1a, gngt1, sh3bgr, pcmtd1, eif4g3a, tdp2b, gkap1, paip1, dlgap5, mt-co1, mt-co2, tbx16l, alpl, chd1l, ctslb, si:dkey-239j18.3, zgc:174154, si:dkey-239j18.2, si:dkey-269i1.4, si:dkey-26g8.5, zgc:174153, zgc:174855, stim1a, vav2, srcap, fundc2, sh3gl1a, cpa5, tjp2b, epb41l4a, zgc:163022, wu:fj05g07, ubn2b, zdhhc24, zp3.2, zgc:173556, zp3, crestin, cnot8, birc5b, dync1i2b, plcg1, grnas, cyp1b1*
M3	600	402	409	26	2	4	*txndc5, si:dkey-260c8.6, si:dkeyp-4f2.1, plod3*
M4	462	371	374	6	8	5	*psma4, sptssa, atp2b2, rab40c, slc1a2b*
M5	412	307	319	4	6	4	*higd1a, vsnl1a, si:ch211-134m17.9, fzd8a*
M6	358	288	306	3	3	3	*lgals2a, rnf8, tfap2a*
M7	325	252	253	5	5	5	*b3galt2, nnt, rpl36a, lrrn1, cntn1a*
M8	287	205	219	6	5	6	*psmb4, lsm8, camk2b1, mrpl47, sin3ab, psmb2*
M9	271	204	222	4	2	2	*eif4ebp1, sqstm1*
M10	268	228	211	18	12	14	*cry2, hsf2, ciarta, cry-dash, arg2, si:dkey-177p2.6, bhlhe41, ddb2, lonrf1l, cfap298, per2, cry1b, arntl1a, hsd17b12a*
M11	191	135	167	10	3	9	*krt91, krt5, krt96, krt97, si:dkey-193p11.2, si:dkey-7c18.24, epcam, krt4, cldnb*
M12	170	121	149	20	11	20	*actc1b, acta1b, tnnt3b, col1a1a, tpma, mylz3, nme2b.2, mylpfa, pvalb2, myhc4, myhz2, myhz1.2, myhz1.3, myhz1.1, myha, col1a2, col10a1b, col1a1b, mybphb, tnnc2*
M13	157	127	133	4	3	3	*aldh7a1, ppt2, rab5aa*
M14	147	94	111	10	3	5	*hif1al, foxg1b, fam107b, fkbp5, nr1d2a*
M15	146	111	111	1	1	1	*spata5l1*
M16	124	105	108	6	7	6	*rps12, rps26l, rps17, rpl23, lgals9l1, rpl39*
M17	114	96	105	3	3	3	*mrpl36, ndufb5, mrps36*

USC UNIVERSIDADE DE COMPOSTELA

M18	110	78	90	4	4	2	*si:dkey-102m7.3, asah1b*
M19	108	95	92	5	5	6	*hibadhb, crip2l, ndufa9b, ndufa9a, dlat, atp6v1ab*
M20	102	76	81	15	10	12	*rbp4l, rcvrn3, arr3a, rbp3, exorh, pde6gb, gngt1, pde6c, pde6a, gngt2a, gnat1, unc119.2*
M21	89	60	72	1	1	1	*dusp23b*
M22	88	71	72	1	1	1	*sypa*
M23	83	56	69	4	4	4	*dlgap4a, maco1a, tnrc6a, mdc1*
M24	78	53	65	13	0	0	-
M25	78	67	54	2	2	2	*c1qc, flna*
M26	66	61	55	4	3	3	*tbl2, ddx47, dlg5a*
M27	63	44	50	7	6	7	*tbr1b, eomesa, foxg1a, bhlhe22, c1ql3b, cpne2, ets2*
M28	62	48	51	0	0	0	-
M29	56	46	49	2	2	2	*cib1, mrps31*
M30	55	36	43	2	2	2	*sept5a, fbxo34*
M31	54	31	36	9	3	3	*samd1a, frmd7, plxnb2a*
M32	53	44	43	2	1	1	*cdh2*
M33	46	37	37	6	4	5	*ier2a, fosab, npas4a, ier2b, junbb*
M34	36	24	33	3	1	2	*dnajb6a, stard4*

[a]: Number of probes assigned to each module.

[b]: Number of human genes that map with the probes of each module.

[c]: Number of zebrafish genes that map with the probes of each module.

[c]: Number of hub probes (probes with an absolute module membership (MM) > 0.85) of each module.

[d]: Number of human genes that map with the hub probes of each module.

[e]: Number of zebrafish genes that map with the hub probes of each module.

3.3. Characterization of co-expression modules by Gene Ontology enrichment analysis

Enrichment analysis revealed that 11 and 19 co-expression modules were significantly enriched at FDR < 0.05 in GO terms for different biological processes, molecular functions, or cellular components in zebrafish and human, respectively (Supplementary Figures 3-4 and Supplementary Tables 4-5). Enrichment in GO terms was dependent on the module size, with smaller modules exhibiting lower enrichment in GO terms. As expected, a higher level of enrichment was observed in human gene-sets compared to zebrafish gene-sets. Notably, co-expression modules M12 and M17 were significantly enriched in a higher number of gene-sets in both zebrafish and human. Some co-expression modules were enriched in zebrafish GO terms relevant to the study, such as M6 in central nervous system and sensory organ development and M17 in drug metabolic processes (Supplementary Figure 3 and Supplementary Table 4). In human, modules M9, M20 and M27 were enriched in myelin sheath, sensory system and

USC DE COMPOSTELA

stimulus detection, and forebrain development, respectively (Supplementary Figure 4 and Supplementary Table 5).

3.4. Enrichment analysis in gene-sets of relevance for psychiatry

Among the 34 co-expression modules, six were significantly enriched in synaptic genes from SynGO: M4, M7, M8, M16, M19 and M28 (Supplementary Table 6). Additionally, modules M14, M22 and M28 exhibited significant enrichment at FDR < 0.05 in different brain tissues (Figure 2). Specifically, M14 was significantly enriched in the cerebellum and hypothalamus; M22 in the amygdala, caudate basal ganglia, and substantia nigra; and M28 in the frontal cortex (BA9), hypothalamus, and substantia nigra. In terms of brain developmental stages, while M27 and M33 were significantly enriched at FDR < 0.05 in different prenatal stages; M6 was significantly enriched in early childhood, M7 in young adulthood, and M8 in adolescence (Figure 2). None of the modules demonstrated enrichment at FDR < 0.05 with genes associated with the GWAS of interest, but 11 modules exhibited enrichment at P < 0.05 (Figure 2). For instance, M1 was enriched in smoking initiation (SmkInit) and cigarettes smoked per day (Cigsperday); M3 in AD; and M32 in age at smoking initiation (AgeSmk) and DrnkWk.

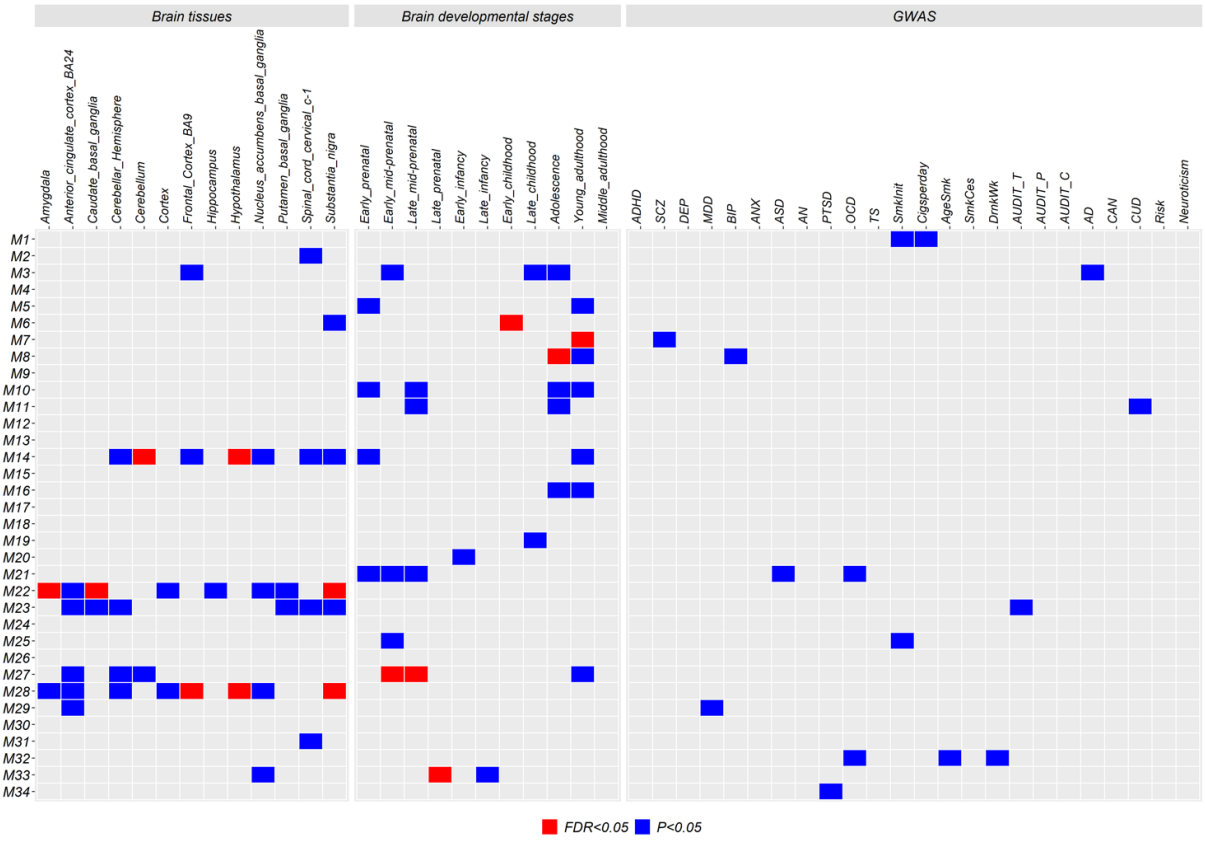


Figure 2. Enrichment analysis of each co-expression module in differentially expressed genes in brain tissues; in differentially expressed genes in brain developmental stages; and in genes associated with psychiatric traits, substance use and dependence traits and personality traits at GWAS. Significant enrichment at FDR < 0.05 is colored in red and significant enrichment at P < 0.05 is colored in blue. FDR

correction for multiple-testing was applied for each module for the number of gene-sets in each category separately. GWAS are described in Supplementary Table 1.

3.5. Hub genes association with phenotypes of relevance in psychiatry

Hub genes of 25 of our co-expression modules were significantly associated at $P < 5x10^{-8}$ with a relevant phenotype for psychiatry in GWAS Catalog (Supplementary Table 7). Module M1 demonstrated the highest association with psychiatry-related traits, specifically, 67 different traits. Additionally, modules M2, M8, M10, and M23 were related with a high number of relevant traits, each associated with more than 30 traits. Among the 171 identified traits, educational attainment was the trait associated with the highest number of co-expression modules, with hub genes from 15 modules associated with this trait. SmkInit was also associated with a high number of modules, with hub genes from 12 modules associated with this trait. Other traits such as insomnia, highest math taken, general risk tolerance, intelligence, brain morphology, Cigsperday, cognitive performance, and cortical surface area were also associated with hub genes from more than 5 of our modules.

Zebrafish models linked to neuropsychiatric traits were identified for 12 specific hub genes corresponding to 8 of our co-expression modules (Supplementary Table 8). These experimental models were associated with a spectrum of neuropsychiatric traits, including alterations in circadian rhythms, locomotor behaviors, daytime hyperactivity, anorexia, as well as brain and neurodevelopmental defects, and epilepsy. Various experimental tests were employed to assess those behavioral phenotypes, including dark-light tests, tank position tests, mating tests, feeding tests, and swimming tests. Modules M1 and M2 had the most hub genes studied using genetic models, with 4 and 3 genes, respectively. While traits in M1 were more related to behavior and circadian rhythms, those in M2 were more related to neurodevelopmental defects. Additionally, the most studied genes were *slc1a2b* (M4) and *aldh7a1* (M13), both related to locomotor behavior anomalies associated with epilepsy.

3.6. Co-expression modules association with the experimental conditions

Association of each co-expression module with experimental condition status (ethanol withdrawal, nicotine withdrawal, starved, cold stress, and old adult) was assessed. To consider the variation of a higher number of samples and analyze the genetic susceptibility to drug dependence across substances, association of our co-expression modules with some type of substance withdrawal was also assessed. Among the co-expression modules, 32 exhibited significant associations at FDR < 0.05 with at least one of the analyzed experimental conditions (Figure 3 and Supplementary Table 9). A higher number of the co-expression modules were associated with ethanol withdrawal (20 significant modules), nicotine withdrawal (14

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

significant modules), substance withdrawal (17 significant modules), and cold stress (16 significant modules) than with starvation (3 significant modules) and aging (6 significant modules) (Figure 3). In general, expression of the module explained a moderate proportion of the experimental condition variance (mean Nagelkerke's R^2=0.32). However, in several instances, module expression explained all the variance associated with the experimental condition (Nagelkerke's R^2 = 1.00). Modules M2, M15, M26, M28, and M31 were significantly associated with both substance withdrawal and ethanol and nicotine withdrawal, but not with the rest of the experimental conditions (Figure 3).

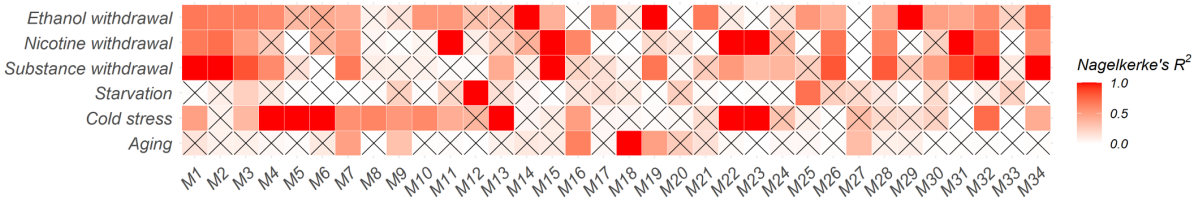


Figure 3. Association of each co-expression module with each experimental condition included in our data. Proportion of variance explained (Nagelkerke's R2) by each module eigengene is shown. FDR correction of the P-value of the logistic regression model was applied for each module for the number of conditions tested. Non-significant associations at FDR > 0.05 are indicated with a cross.

3.7. Contribution of co-expression modules polygenic risk score to alcohol dependence

PRS from the three tested GWAS explained a similar percentage of AD variance in our target sample at the most significant P-value threshold (Supplementary Figure 5). However, for AD and AUDIT-T GWAS, that threshold ($P = 5x10^{-8}$) only involved 1 SNP, located at the alcohol dehydrogenase cluster, and 15 SNPs, probably related to alcohol metabolism, respectively. The other P-value thresholds for AD and AUDIT-T were much less associated with our target trait (Supplementary Figure 5). By contrast, the most significant threshold for DrnkWk (P = 0.0574) involved 21,193 SNPs; and the percentage of variance explained by the PRS does not show a drastic reduction at other P-value thresholds, consistent with a polygenic effect (Supplementary Figure 5). Therefore, given our focus on assessing common polygenic susceptibility to AD, we selected DrnkWk GWAS for PRSet analysis.

We assessed the association between AD and a PRS based on all the genes belonging to the five co-expression modules exclusively associated with substance withdrawal, ethanol withdrawal and nicotine withdrawal. We compared the level of association of our PRS with a background PRS based on those genes corresponding to the probes used for WGCNA analysis, excluding those belonging to co-expression modules significantly associated with ethanol withdrawal, nicotine withdrawal or substance withdrawal. This PRS based on M2, M15, M26, M28, and M31 modules was significantly associated with AD (increase in Nagelkerke's R^2 =

USC DE COMPOSTELA

0.89%, P = 0.0061) and significantly enriched in association with this trait compared to the background PRS of randomly selected brain-expressed genes (P = 0.0060) (Table 3). We then analyzed the contribution of each module PRS individually. Among the five tested module PRSs, M2 and M31 PRSs were significantly associated with our target trait (Table 3). Moreover, the PRS based on genes from module M2 was significantly enriched in association with AD (Table 3).

To test whether we are finding co-expression modules specifically involved in each experimental condition instead of co-expression modules involved in general response to external stimulus, we assessed the association with AD of a PRS based on genes belonging to co-expression modules exclusively associated with cold stress. This PRS based on M5, M6, and M8 modules was significantly associated with AD (increase in Nagelkerke's R^2 = 0.58%, P = 0.0260) and significantly enriched in association with AD (P = 0.0357), but it was less enriched than the PRS based on co-expression modules exclusively associated with substance withdrawal (Table 3).

Table 3. Results for a substance withdrawal PRS based on all the genes assigned to the co-expression modules exclusively associated with substance withdrawal, for the PRSs of each of these co-expression modules, and for a cold stress PRS based on all the genes assigned to the co-expression modules exclusively associated with cold stress. Results of association analysis between each PRS and AD in our target sample are shown. Enrichment of each PRS in association with AD is also provided.

PRS	Independent SNPs[a]	Increase in R^2 (%)[b]	Coefficient (SE)	Association P-value[c]	Enrichment P-value[d]
Substance withdrawal PRS[e]	4,170	0.8870	9,538.83 (3,475.83)	0.0061	0.0060
M2	3,340	0.6240	7,026.04 (3,048.29)	0.0212	0.0350
M15	366	0.0020	-144.17 (1,109.48)	0.8966	0.9040
M26	195	0.2095	930.47 (695.15)	0.1807	0.2093
M28	230	0.0099	233.30 (801.65)	0.7710	0.7873
M31	145	0.4895	1,365.54 (668.59)	0.0411	0.0522
Cold stress PRS[f]	2,857	0.5816	6,163.11 (2,768.43)	0.0260	0.0357

[a]: Number of independent SNPs after clumping.

[b]: Increase in Nagelkerke's R^2 adjusted for ascertainment bias by a prevalence of AD of 13.8%.

[c]: P-value indicating the level of association of each PRS with AD in our target sample. Adjusted by sex, age, study phase, and the first 10 principal components of ancestry as covariates.

[d]: Competitive P-value comparing the level of association of each PRS with a background PRS.

[e]: PRS based on all the genes assigned to modules M2, M15, M26, M28, and M31.

[f]: PRS based on all the genes assigned to modules M5, M6, and M8.

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

4. DISCUSSION

Our study revealed 34 co-expression gene modules in the zebrafish brain based on gene expression data. Notably, certain co-expression modules showed enrichment in human synaptic genes, specific brain tissues, or distinct brain developmental stages, supporting their potential to model huan brain gene networks. Moreover, our analysis identified co-expression modules that may play a role in psychiatry-related traits, as evidenced by both GWAS enrichment analysis and hub genes revision. We also identified specific co-expression modules whose expression patterns were associated with experimental conditions. Finally, we confirmed that a PRS based on these co-expression modules exclusively associated with substance withdrawal significantly contributes to the genetic susceptibility of AD.

In this study, we used brain expression data collected under diverse conditions to enhance variability of our expression data. Notably, hierarchical clustering of the samples revealed that under two of these conditions, i.e., starvation and aging, gene expression patterns were similar to control samples, although they tended to cluster separately. In contrast, substance withdrawal and cold stress produced distinct expression patterns, emerging as the primary sources of variation in our expression data. Consequently, most co-expression modules were found to be associated with substance withdrawal and/or cold stress.

Notably, several of our co-expression modules were associated with both substance withdrawal and cold stress, indicating enrichment in molecular pathways involved in both conditions rather than being specific to either (F. He & Maslov, 2016). Previous studies suggest that gene expression changes induced by both substance withdrawal and cold stress may be related to stress response. Elevated cortisol levels and anxiety-like behaviors, characteristic of this stress response, may arise in zebrafish as a consequence of substance withdrawal (Cachat et al., 2010; López-Patiño et al., 2008; Mathur & Guo, 2011; Ponzoni et al., 2021) and cold stress (Nonnis et al., 2021; Reid et al., 2022). This response may be partially induced by oxidative stress, which can occur in zebrafish due to substance withdrawal (Agostini et al., 2020; Mocelin et al., 2019) or cold stress (Reid et al., 2022; Tseng et al., 2011). Moreover, both substance withdrawal and cold stress lead to changes in brain gene expression that likely contribute to the increase of these anxiety-like behaviors (Nonnis et al., 2021; Ponzoni et al., 2021). Consistently, our analysis identified two co-expression modules, namely M4 and M7, enriched in synapse-related genes from synGO and associated with both cold stress and substance withdrawal.

USC UNIVERSIDADE DE SANTIAGO

Substance withdrawal has been proposed as one of the three stages constituting the addiction cycle, namely binge/intoxication, withdrawal/negative affect, and preoccupation/anticipation

(Koob & Volkow, 2010, 2016). While reward mechanisms play a crucial role in the initial phases of substance dependence; in the later stages, as a consequence to chronic substance exposure, there is a neuroadaptation of the reward, stress, and executive function systems, which is crucial for sustaining substance addiction. Consequently, gene expression changes in zebrafish during substance withdrawal after chronic substance exposure may be indicative of this neuroadaptation during substance addiction. The original study used in this work, which assessed gene expression under substance withdrawal, found that chronic exposure to substances results in behavior changes in zebrafish, including increased time spent in the treatment side despite punishment that lasted up to 3 weeks after the last exposure to the substance (Kily et al., 2008). This drug seeking behavior despite adverse consequences confirmed its suitability to model substance addiction and dependence. Therefore, our co-expression modules associated with substance withdrawal in zebrafish may be used to model molecular pathways implicated in human substance dependence.

We used PRSs to verify whether the genes of those co-expression modules whose expression is exclusively associated with substance withdrawal in zebrafish are involved in substance use disorders in humans. Specifically, we assessed the contribution of a PRS based on genetic variants around these genes to AD risk. Our PRS was constructed based on a DrnkWk GWAS, rather than a substance withdrawal GWAS or an AD GWAS, which would be more suited to assess the effect of these genes associated with substance withdrawal on AD. Unfortunately, available GWAS for substance withdrawal have limited sample sizes (Hällfors et al., 2019; A. H. Smith et al., 2018), and our exploratory analysis discarded the AD GWAS. However, there is a considerable genetic correlation between alcohol use disorder and DrnkWk (r_g=0.67, SE=0.02) (Zhou et al., 2020). Moreover, a previous study using the same target sample observed independent contributions of both AD and DrnkWk PRS to AD risk (Facal et al., 2021). This evidence supports the suitability of DrnkWk GWAS to predict AD. Our PRS, based on the subset of genes at co-expression modules exclusively associated with substance withdrawal, was more strongly associated with AD than PRSs of the same size based on randomly selected brain-expressed genes included in our WGCNA analysis. Moreover, among all co-expression modules exclusively associated with substance withdrawal, we identified one that was individually enriched in AD. Six hub genes of this module have already been identified as associated with different substance use behaviors in GWAS Catalog, aligning with the common genetic susceptibility across substances (Hatoum et al., 2023). Furthermore, one of the hub genes of this module is *OPRM1*, which encodes the μ-opioid receptor involved in reward-related behaviors (Thorsell, 2013). Naltrexone, an antagonist for this receptor, is used in the pharmacological treatment of alcohol use disorder and opioid use disorder, underscoring

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

the potential of our approach in identifying novel pharmacological targets (Bell & Strang, 2020; Thorsell, 2013). Additionally, a PRS based on genes exclusively associated with cold stress also showed significant enrichment in association with AD, but to a lesser extent than the substance withdrawal PRS. This may be probably related to common response to stress, involving anxiety-like behaviors (see above).

Our approach facilitates the integration of zebrafish expression data under specific conditions with human GWAS data. A similar strategy was previously applied in a study where a PRS based on genes associated with nicotine exposure in zebrafish and other animal models significantly predicted human tobacco consumption in an independent sample (Palmer et al., 2021). Other studies have followed a comparable approach to assess the contribution of gene-sets derived from differentially expressed genes identified in murine models to the genetic susceptibility to psychiatry-related traits, not always finding evidence for enrichment (Brasher et al., 2023; Huggett et al., 2021; Mize et al., 2023). It is noteworthy that all these previous studies analyzed sets of genes exhibiting differential expression under specific conditions. In our case, we take a step further by assessing the contribution of sets of functionally related genes that are highly co-expressed under a certain condition and whose common expression pattern is associated with that condition. Our approach may also aid in the validation of GWAS associations using a zebrafish model, which has already been successfully utilized for this purpose in validating GWAS associations with psychiatric disorders (Antón-Galindo et al., 2024; Dark et al., 2020). However, modeling psychiatric disorders becomes intricate due to their polygenic nature, with multiple genes interacting to determine the genetic risk for the trait. Future studies could benefit from our approach, potentially enabling a reduction of the high number of genes associated with a psychiatric disorder to a subset of genes representing the molecular networks involved in that specific psychiatric disorder, i.e., the hub genes of the co-expression modules associated with that disorder.

The present study has certain limitations that should be considered in the interpretation of the results. Firstly, the sample size in our WGCNA analysis was relatively low, limited by publicly available expression data of the zebrafish brain and the difficulty in combining data from various sources. Secondly, the conversion of all probes used in the WGCNA analysis to the corresponding zebrafish and human genes was not feasible, reducing the power of our subsequent analyses. Thirdly, hub genes characterization was constrained by the limited number of experimental studies investigating the effects of gene depletion models on psychiatry-related traits in zebrafish, despite the existence of numerous behavioral tests for zebrafish (Kalueff et al., 2014). Finally, as previously mentioned, PRS estimation relied on a

USC DE COMPOSTELA

DrnkWk GWAS, rather than on a substance withdrawal or substance dependence GWAS, which might be more appropriate.

5. CONCLUSION

In summary, our analysis of zebrafish brain expression data led to the identification of sets of co-expressed genes potentially implicated in psychiatry-related traits. Specifically, we detected a group of co-expressed genes whose expression was exclusively associated with substance withdrawal in zebrafish and significantly contributed to AD susceptibility in humans. This novel approach detects relevant hub genes for psychiatry disorders, facilitating further investigation of underlying molecular mechanisms and the identification of potential pharmacological targets.

USC
UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

4.2. ESTUDO 2: POLYGENIC OVERLAP OF SUBSTANCE USE BEHAVIORS AND DISORDERS WITH EXTERNALIZING AND INTERNALIZING PROBLEMS INDEPENDENT OF GENETIC CORRELATIONS

Referencia:

Al-Soufi, L., Hindley, G., Rødevand, L., Shadrin, A. A., Jaholkowski, P., Fominykh, V., Icick, R., Tesfaye, M., Costas, J. e Andreassen, O. A. (2024). Polygenic overlap of substance use behaviors and disorders with externalizing and internalizing problems independent of genetic correlations [Traballo non publicado].

Información dos/as autores/as:

Laila Al-Soufi[1,2], Guy Hindley[3], Linn Rødevand[3], Alexey A. Shadrin[3], Piotr Jaholkowski[3], Vera Fominykh[3], Romain Icick[3,4], Markos Tesfaye[3,5], Javier Costas[1,6,7] e Ole A. Andreassen[3,7]

[1]Grupo de Xenética Psiquiátrica, Instituto de Investigación Sanitaria de Santiago de Compostela (IDIS), Santiago de Compostela, Galicia, España. Rede de Investigación en Atención Primaria de Adiccións (RIAPAd).

[2]Departamento de Zooloxía, Xenética e Antropoloxía Física, Universidade de Santiago de Compostela, Galicia, España.

[3]Centre for Precision Psychiatry, Division of Mental Health and Addiction, Institute of Clinical Medicine, University of Oslo, Oslo, Noruega.

[4]INSERM UMR-S1144, Université de Paris Cité, Paris F-75006

[5]Department of Clinical Science, University of Bergen, Bergen, Noruega.

[6]Complexo Hospitalario Universitario de Santiago de Compostela (CHUS), Servizo Galego de Saúde (SERGAS), Santiago de Compostela, Galicia, España.

[7]Estes autores contribuíron por igual: Javier Costas, Ole A. Andreassen

Contribución da doutoranda na publicación:

Realización de todas as análises salvo o GWAS de TUD e a análise de MiXeR para o GWAS de ansiedade de Levey et al. (2020). Redacción do manuscrito inicial e revisión da versión final.

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

ABSTRACT

Externalizing and internalizing pathways may lead to the development of substance use behaviors (SUBs) and substance use disorders (SUDs), which are all heritable phenotypes. Genetic correlation studies have indicated differences in the genetic susceptibility between SUBs and SUDs. We investigated whether these substance use phenotypes are differently related to externalizing and internalizing problems at a genetic level. We analyzed data from genome-wide association studies (GWAS) of four SUBs and SUDs, five externalizing traits, and five internalizing traits using the bivariate causal mixture model (MiXeR) to estimate genetic overlap beyond genetic correlation. Two distinct patterns were found. SUBs demonstrated high genetic overlap but low genetic correlation of shared variants with internalizing traits, suggesting a pattern of mixed effect directions of shared genetic variants. Conversely, SUDs and externalizing traits exhibited considerable genetic overlap with moderate to high positive genetic correlation of shared variants, suggesting concordant effect direction of shared risk variants. These results highlight the importance of the externalizing pathway in SUDs as well as the limited role of the internalizing pathway in SUBs. Further studies are needed to reveal the underlying neuronal mechanisms of these traits.

USC
UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

1. INTRODUCTION

Substance use stands as a major risk factor contributing significantly to the global burden of disease (Degenhardt, Charlson, et al., 2018). To improve prevention and treatment, it is important to increase our understanding of the pathogenesis of substance use behaviors (SUBs) (e.g., initiation, lifetime use and severity of consumption) and substance use disorders (SUDs). This includes identifying factors influencing different aspects of substance use, from severity of substance consumption to pathological substance use.

Evidence suggests two main pathways for substance use initiation and escalation towards SUDs: an *externalizing* pathway characterized by high levels of impulsivity, aggression, and sensation-seeking, and an *internalizing* pathway characterized by using substances to cope with negative affect, depression, and anxiety (Hussong et al., 2011; Krueger et al., 2002). This is supported by prospective evidence of distinct externalizing and internalizing pathways leading to future SUBs and SUDs (Meque et al., 2019; Ning et al., 2020). However, evidence is weaker for the internalizing pathway. This may be due to comorbidity between internalizing and externalizing problems and differences regarding gender, developmental timing or type of substance use (Heitzeg et al., 2018; Ning et al., 2020). For instance, internalizing traits seem to be related to reduced alcohol consumption in middle adulthood subjects, but to increased consumption in more severe outcomes (i.e., heavy drinking and alcohol use disorder (AUD)) during transition to adulthood (Ning et al., 2020).

Recent studies have examined the extent of shared genetic susceptibility between different SUBs and SUDs by estimating genetic correlations (r_g) and cross-trait polygenic risk scores. These have shown that smoking severity, defined as the number of cigarettes smoked per day (Cigsperday), and nicotine dependence are almost genetically identical (r_g=0.95) (Quach et al., 2020). In fact, Cigsperday was proposed as a "genetic proxy" for nicotine dependence, considering their strong genetic correlation and their similar pattern of associations with multiple diseases (Sanchez-Roige et al., 2021). By contrast, the genetic correlation between alcohol consumption, measured as the number of drinks per week (DrnkWk), and alcohol dependence was weaker (r_g=0.37) (Walters et al., 2018). Similarly, there is a low to moderate genetic correlation between smoking initiation and nicotine dependence (r_g=0.40) (Quach et al., 2020). Further, in contrast to normal alcohol use, problematic alcohol use (e.g., AUD, alcohol dependence and problematic consequences of drinking identified using the Alcohol Use Disorders Identification Test (AUDIT-P)) seems to be associated with increased genetic predisposition for general psychopathology (Facal et al., 2021; Sanchez-Roige et al., 2019; Walters et al., 2018; Zhou et al., 2023). These studies indicated variation in the genetic

susceptibility between SUBs and SUDs. Thus, the involvement of externalizing and internalizing pathways may differ between SUBs and SUDs, which represent heterogeneous substance-related phenotypes at the biological and clinical levels.

Significant genetic correlations have been identified between both externalizing and internalizing traits and both SUBs and SUDs across different substances. For example, attention-deficit/hyperactivity disorder (ADHD) was genetically correlated with SUDs such as nicotine dependence, alcohol dependence and cannabis use disorder (CUD) (Abdellaoui et al., 2021; Vink et al., 2020; Walters et al., 2018); as well as with SUBs such as tobacco and cannabis use (M. Liu et al., 2019; Pasman et al., 2018; Soler Artigas et al., 2020). Other risky-behavior traits considered as externalizing were associated with SUBs. For example, age at first sexual intercourse (AFS) and age at first birth (AFB) showed a significant genetic correlation with age of smoking initiation and cannabis use (Mills et al., 2021). Further, other studies reported a significant genetic correlation of depression with both SUBs and SUDs, including smoking initiation, smoking cessation, nicotine dependence and alcohol dependence (Abdellaoui et al., 2021; M. Liu et al., 2019; Vink et al., 2020; Walters et al., 2018).

Support for a causal effect of both externalizing and internalizing problems on SUBs and SUDs comes from Mendelian randomization studies. There is evidence for a causal effect of ADHD on tobacco and cannabis use behaviors, but also for reverse causation (Treur et al., 2019; Vilar-Ribó et al., 2020). In addition, Mendelian randomization studies indicated that depression and neuroticism are risk factors for tobacco use and alcohol dependence and discarded reverse causation (Polimanti et al., 2019; Sallis et al., 2019; Yao et al., 2021). However, there is also some evidence for a bidirectional causal relationship between depression and smoking behavior (Wootton et al., 2020).

Studies using the bivariate causal mixture model (MiXeR) have provided further evidence for the genetic association between mental traits (Hindley et al., 2022). MiXeR characterizes the pattern of genetic overlap between two traits beyond genome-wide genetic correlation (Frei et al., 2019). First, MiXeR estimates the number of causal variants involved in each trait. Next, MiXeR estimates the number of shared variants between the two traits and the genetic correlation of shared variants. In contrast to genetic correlation studies, MiXeR enables the discovery of genetic overlap even if the shared variants possess a mixture of concordant and discordant effect directions.

USC UNIVERSIDADE DE SANTIAGO

In the current study, we aimed to characterize the genetic relationship of SUBs and SUDs with externalizing and internalizing traits using MiXeR. Specifically, we investigated if: i) SUBs and SUDs differ in their pattern of polygenic overlap with externalizing and internalizing traits,

including the proportion of genetic overlap and genetic correlation of shared variants, and ii) these differences are present regardless of the particular substance involved. This approach can improve our understanding of the contribution of the externalizing and internalizing pathways to SUBs and SUDs.

2. METHODS

2.1. Genome-wide association studies (GWAS)

We included summary statistics for SUBs and SUDs and externalizing and internalizing traits from large GWAS (Table 1 and Supplementary Table 1). These GWAS were selected according to their large sample size and public availability. All GWAS summary statistics were based on samples of European ancestry in order to share similar linkage disequilibrium pattern, a requirement for MiXeR analysis.

For the three most studied drugs: tobacco, alcohol, and cannabis, we selected at least one SUB and one SUD or SUD-proxy. As tobacco use behavior, we chose ever been a regular smoker (EverSmk) (G. R. B. Saunders et al., 2022), which reflects smoking initiation. To assess tobacco use disorder (TUD), we generated our own GWAS of TUD using UK Biobank data (further information in Supplementary Methods). UK Biobank data was obtained under accession number 27412. We also used Cigsperday (G. R. B. Saunders et al., 2022), which is a good "genetic proxy" of nicotine dependence (Sanchez-Roige et al., 2021). As alcohol use behaviors, we selected DrnkWk (G. R. B. Saunders et al., 2022), a score comprising the items related to alcohol consumption of the Alcohol Use Disorders Identification Test (AUDIT-C) (Sanchez-Roige et al., 2019) and AUDIT-P (Sanchez-Roige et al., 2019). AUD was assessed by an AUD GWAS (Icick et al., 2023). For cannabis, we selected lifetime cannabis use (CANN) (Pasman et al., 2018) and CUD (E. C. Johnson et al., 2020).

As externalizing traits, we chose those four traits that were previously used to construct an externalizing genetic factor and do not involve substance use (Karlsson Linnér et al., 2021): ADHD (Demontis et al., 2023), AFS (Mills et al., 2021), number of sexual partners (NSEX) (Karlsson Linnér et al., 2019) and general risk tolerance (RISK) (Karlsson Linnér et al., 2019). In addition, AFB (Mills et al., 2021) and childhood aggressive behavior (AGG) (Ip et al., 2021), related to externalizing behavior (Baselmans et al., 2022; Mills et al., 2021), were also included.

For internalizing traits, different phenotypes of depression and anxiety were selected: a broader diagnosis of depression (DEP) (Howard et al., 2019), major depressive disorder (MDD) (Wray et al., 2018), a continuous trait for anxiety based on the Generalized Anxiety Disorder 2-item scale (GAD) (Levey et al., 2020) and self-reported anxiety (ANX) (http://www.nealelab.is/uk-

biobank/). Moreover, neuroticism, a core feature of internalizing behavior (Griffith et al., 2010), measured by a neuroticism score based on 12 items (NEURO) was also included. Specific subclusters and items of NEURO were also analyzed: depressed affect subcluster of NEURO (DEP-NEURO) (Nagel, Jansen, et al., 2018), worry subcluster of NEURO (WORRY-NEURO) (Nagel, Jansen, et al., 2018), mood instability (MOOD) (Nagel, Watanabe, et al., 2018) and often feeling lonely (LONE) (Nagel, Watanabe, et al., 2018). Subjective well-being (SWB) (Kim et al., 2022) was also selected as it is highly correlated with anxious/depressive behavior (Bartels et al., 2013). Other aspects of loneliness were also included: frequency of being able to confide in someone close (CONF) (http://www.nealelab.is/uk-biobank/) and a combined multi-trait GWAS of loneliness and social isolation (LONE-MTAG) (Day et al., 2018), which is strongly genetically correlated with neuroticism and depressive symptoms.

Table 1. GWAS summary statistics used in MiXeR analysis.

Trait	Description	Source	Sample size (N cases)
SUBs and SUDs or SUD-proxies			
EverSmk	Ever been a regular smoker, i.e., ever versus never regular smokers	Saunders et al., 2022	805,431 (393,707)
Cigsperday	Average number of cigarettes smoked per day, either as a current or as a former smoker	Saunders et al., 2022	326,497
DrnkWk	Average number of drinks drunk each week	Saunders et al., 2022	666,978
AUD	Alcohol use disorder	Icick et al., 2023	981,384 (75,635)
Externalizing traits			
ADHD	Attention deficit hyperactivity disorder	Demontis et al., 2023	225,534 (38,691)
RISK	General risk tolerance	Karlsson Linnér et al., 2019	466,571
NSEX	Lifetime number of sexual partners	Karlsson Linnér et al., 2019	370,711
AFB	Age at first birth	Mills et al., 2021	542,901
AFS	Age at first sexual intercourse	Mills et al., 2021	397,338
Internalizing traits			
DEP	Lifetime diagnosis of major depressive disorder or broad self-reported depression	Howard et al., 2019	500,199 (170,756)
NEURO	Weighted neuroticism score based on the response to 12 items	Nagel et al., 2018	390,278
WORRY-NEURO	Worry subcluster of NEURO	Nagel et al., 2018	348,219
DEP-NEURO	Depressed affect subcluster of NEURO	Nagel et al., 2018	357,957
MOOD	Experiencing mood swings	Nagel et al., 2018	373,733 (168,180)

2.2. Data analysis

We applied MiXeR, version 1.2.0, to characterize the pattern of genetic overlap between SUBs and SUDs and externalizing and internalizing traits using GWAS summary statistics (Frei et al., 2019). First, univariate MiXeR was applied to each trait to estimate their polygenicity, i.e., the fraction of causal variants in the genome, and discoverability, i.e., the causal effect size

variance, maximizing likelihood estimation (Holland et al., 2020). Polygenicity is presented as the number of causal variants required to explain 90% of h^2_{SNP}. This threshold of 90% is applied to exclude those variants with infinitesimally small effects. Next, bivariate MiXeR was applied to assess the pattern of genetic overlap between each pair of traits. Using the parameters estimated in univariate MiXeR for each trait, bivariate MiXeR estimates the number of shared causal variants, the number of unique causal variants specific to each trait and the correlation of effect sizes of shared variants (r_{gs}). To enhance result interpretation, we assessed the degree of genetic overlap between each pair of traits by quantifying the percentage of shared variants with respect to the SUB/SUD. To assess model fit, the difference between the Akaike information criterion (AIC) for the MiXeR-fitted model and a reference model was calculated. Positive AIC differences indicate that the MiXeR-fitted estimates are more informative than the reference model. Further information about MiXeR is provided in Supplementary Methods and the original publication (Frei et al., 2019).

To limit bivariate MiXeR analyses to the most reliable results, we selected GWAS according to their MiXeR model fit. Specifically, we selected those traits with the most reliable univariate MiXeR model fit, using an arbitrary threshold of univariate AIC higher than 50. For Cigsperday GWAS, we excluded the cluster of neuronal nicotinic acetylcholine receptor (nAchR) genes at chromosome 15 because its strong effect on this phenotype violates MiXeR's assumption that the effect sizes of causal variants are normally distributed. Further details about the exclusion of the cluster of nAchR genes at chromosome 15 for Cigsperday are shown in Supplementary Methods.

3. RESULTS

After excluding GWAS with poor MiXeR model fit, we included the following tobacco and alcohol use behaviors, respectively: EverSmk and DrnkWk; and a SUD or SUD-proxy for each of these substances: Cigsperday and AUD (Figure 1). We included five externalizing traits: AFS, NSEX, AFB, RISK and ADHD; and five internalizing traits: NEURO, DEP-NEURO, WORRY-NEURO, DEP and MOOD.

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

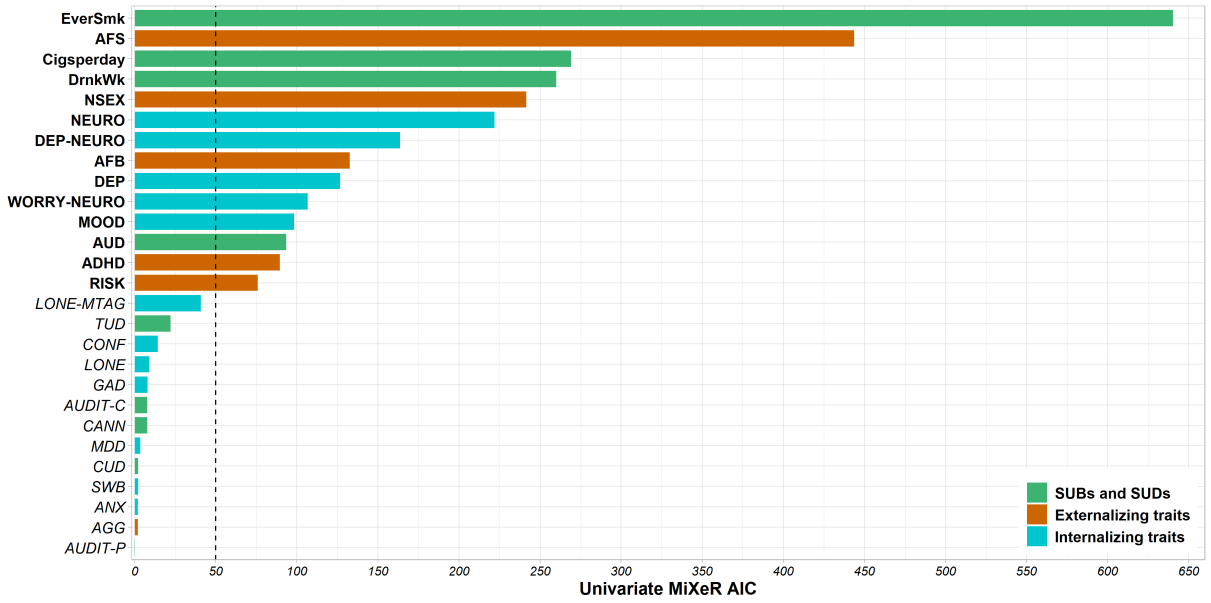


Figure 1. GWAS filtering according to their univariate MiXeR AIC. GWAS selected for MiXeR analyses, i.e., those with a univariate AIC higher than 50, are highlighted in bold. Dashed line represents univariate MiXeR AIC=50. Traits are described in Table 1 and Supplementary Table 1.

3.1. Univariate MiXeR results

Univariate MiXeR estimated the polygenicity and discoverability of each of the selected traits (Supplementary Table 2). Among SUBs and SUDs, EverSmk was the most polygenic trait (N=11.6K causal variants explaining 90% of the h^2_{SNP}, SE=0.7K). DrnkWk and AUD had similar polygenicity (DrnkWk: N=7.7K, SE=0.6K; AUD: N=7.7K, SE=0.3K). Cigsperday was the least polygenic trait with 4.3K (SE=0.6K) variants involved. All selected externalizing and internalizing traits were highly polygenic with 9.7K-12.2K variants explaining 90% of the h^2_{SNP}, except for ADHD with 7.7K (SE=0.4K) causal variants.

3.2. Bivariate MiXeR results

Bivariate MiXeR was used to estimate the pattern of genetic overlap between each of the four selected SUBs and SUDs and the 10 selected externalizing and internalizing traits. Specifically, the percentage of shared variants with respect to the SUB/SUD and the genetic correlation of shared variants were estimated for each pair of traits. Full results can be found in Supplementary Table 3.

3.2.1.Bivariate MiXeR model fit

For almost all bivariate analyses, both AIC compared to minimum and maximum overlap were positive, indicating a good model fit (Supplementary Table 3). Minimum AIC was negative for AUD with NSEX and DEP, indicating that the overlap modeled by MiXeR was not distinguishable from the minimum overlap taking into account genetic correlation between both traits. Maximum AIC was negative for some analyses for which there was a high overlap

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

between both traits, indicating that the overlap estimated by MiXeR was not distinguishable from complete genetic overlap.

3.2.2.Pattern of genetic overlap between externalizing/internalizing traits and SUBs

The percentage of SUB causal variants shared with externalizing traits ranged from 52.0% to 95.9% (Figure 2 and Supplementary Figure 1). The genetic correlation of shared variants was moderate-high with almost all externalizing traits (absolute values of r_{gs}=0.28-0.87), except between DrnkWk and AFB (absolute value of r_{gs}=0.05) (Figure 3).

In the case of internalizing traits, genetic overlap with SUBs was high (Figure 2 and Supplementary Figure 1). EverSmk and DrnkWk shared more than 82% and 93% of their causal variants with internalizing traits, respectively (Figure 2). The genetic correlation of shared variants with internalizing traits was clearly lower than with externalizing traits (absolute values of r_{gs}=0.02-0.35) (Figure 3).

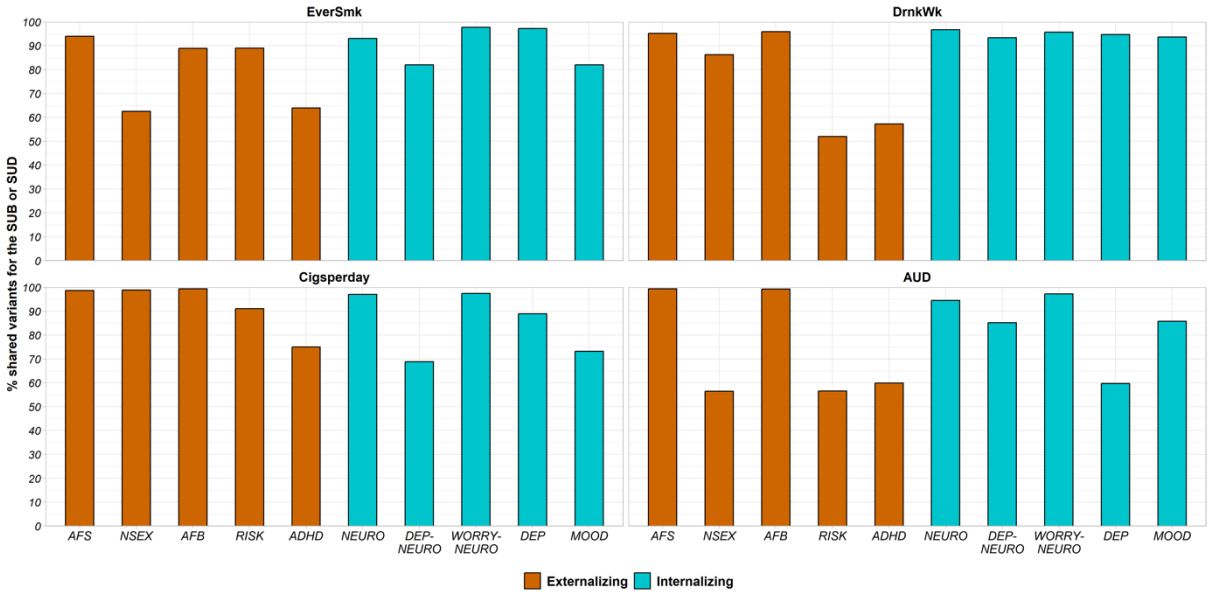


Figure 2. Percentage of variants of each SUB (EverSmk and DrnkWk) or SUD/SUD-proxy (Cigsperday and AUD) shared with each externalizing (colored in orange) or internalizing (colored in blue) trait. Traits are described in Table 1.

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

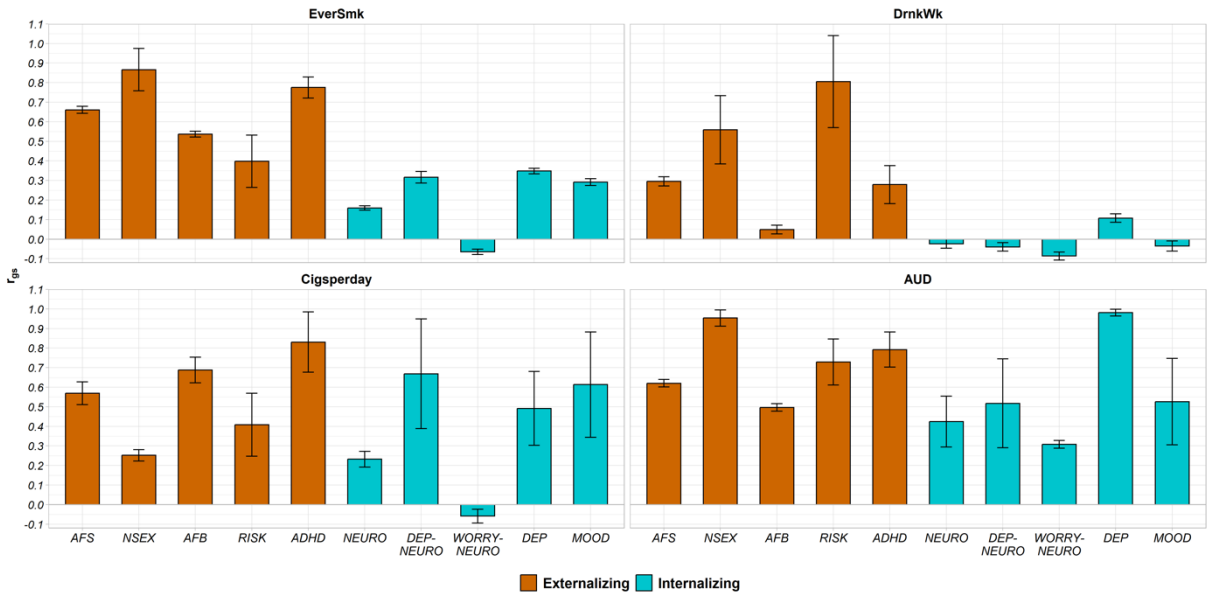


Figure 3. Genetic correlation of the shared variants (r_{gs}) between each SUB (EverSmk and DrnkWk) or SUD/SUD-proxy (Cigsperday and AUD) and each externalizing (colored in orange) or internalizing (colored in blue) trait. Traits are described in Table 1. Error bars represent ±1 SD. Direction was reversed for AFB and AFS to indicate genetic correlation with a higher externalizing behavior.

Despite following the same general pattern of genetic overlap, some differences emerged between the tobacco use behavior (EverSmk) and alcohol use behavior (DrnkWk). EverSmk exhibited a slightly higher genetic correlation of shared variants (absolute values of r_{gs}=0.40-0.87) with externalizing traits compared to DrnkWk (absolute values of r_{gs}=0.05-0.81) (Figure 3). Similarly, internalizing traits exhibited a higher genetic correlation of shared variants with EverSmk (absolute values of r_{gs}=0.06-0.35) compared to DrnkWk (absolute values of r_{gs}=0.02-0.11). Notably, DrnkWk showed a negative genetic correlation of shared variants, although close to 0, with all internalizing traits except for DEP, while EverSmk only displayed a negative genetic correlation of shared variants with WORRY-NEURO among internalizing traits.

3.2.3. Pattern of genetic overlap between externalizing/internalizing traits and SUDs/SUD-proxies

The percentage of causal variants for SUDs and SUD-proxies shared with externalizing traits ranged from 56.5% to 99.4% (Figure 2 and Supplementary Figure 1). The genetic correlation of shared variants for SUDs and SUD-proxies with externalizing traits was relatively high (absolute values of r_{gs}=0.25-0.95) (Figure 3). In all pairs, a higher value for externalizing behavior was positively correlated with both SUBs and SUDs among the shared variants (Figure 3).

Genetic overlap with internalizing traits for SUDs and SUD-proxies (59.7%-97.4%) was more variable than for SUBs (82.0%-97.8%) (Figure 2 and Supplementary Figure 1). Cigsperday and AUD demonstrated a higher genetic correlation of shared variants with internalizing traits,

except WORRY-NEURO, (absolute values of r_{gs}=0.23-0.98) than SUBs (absolute values of r_{gs}=0.02-0.35) (Figure 3). In contrast to the remaining internalizing traits, WORRY-NEURO showed a similar pattern of genetic overlap with both SUBs and SUDs. WORRY-NEURO shared a large number of variants (> 95%) with all SUBs and SUDs, despite possessing low genetic correlation of shared variants (Figure 2 and Figure 3). The direction of the genetic correlation of shared variants was positive for pairs involving AUD and internalizing traits, contrary to DrnkWk (Figure 3). By contrast, the direction of the genetic correlation of shared variants with internalizing traits was positive for both tobacco traits.

3.2.4.Relationship between the genetic correlation of shared variants and the number of shared causal variants

An inverse relationship was observed between the genetic correlation of shared variants and the number of shared causal variants for each pair of traits, with a significant negative correlation between these variables (r=-0.66, P=4.32x10^{-6}) (Figure 4). Those pairs of traits exhibiting a lower genetic correlation of shared variants than expected given the number of shared variants comprised predominantly SUBs and internalizing traits (Figure 4). In contrast, pairs showing a genetic correlation of shared variants higher than expected based on the number of shared variants involved SUDs or SUD-proxies with externalizing traits.

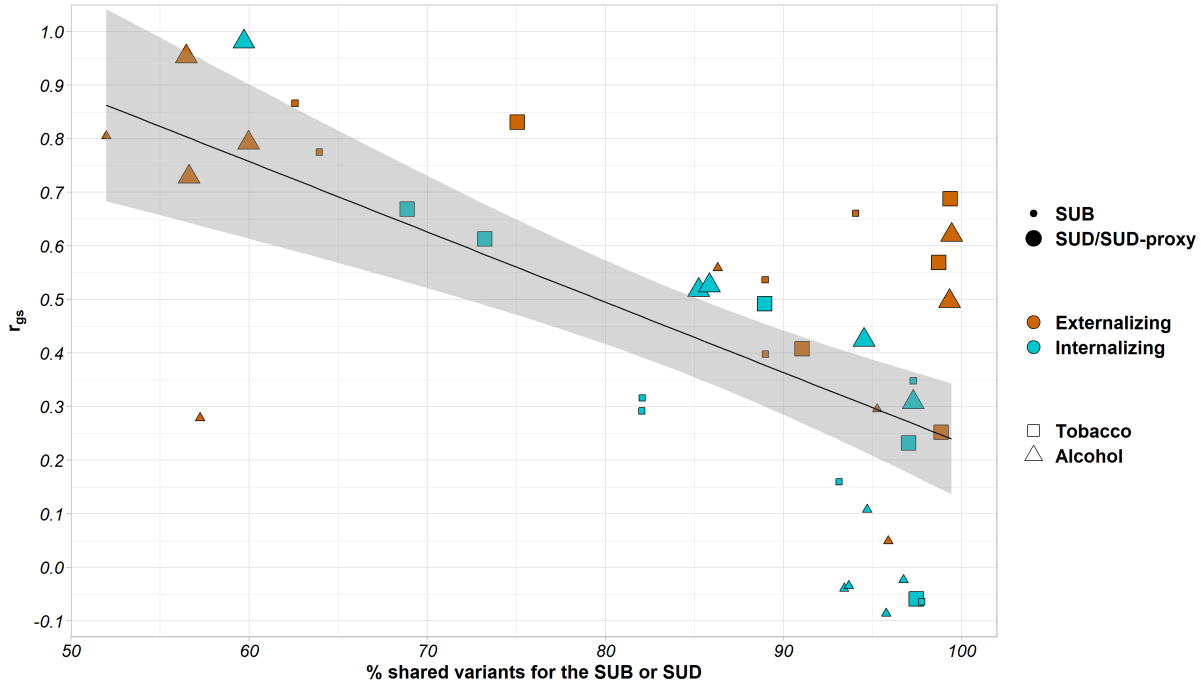


Figure 4. Scatter plot illustrating the negative relationship between the genetic correlation of shared variants (r_{gs}) and the percentage of shared variants across analyzed trait pairs. The solid line represents the linear regression, while the shaded region depicts the 95% confidence interval. On the x-axis, the percentage of shared variants between each SUB or SUD and each externalizing or internalizing trait with respect to the SUB/SUD is displayed. The y-axis shows the genetic correlation of shared variants between each SUB or SUD and each externalizing or internalizing trait. For pairs involving AFB and AFS, direction of

the genetic correlation of shared variants was reversed to indicate genetic correlation with a higher externalizing behavior. Pairs involving SUBs (EverSmk and DrnkWk) are denoted by smaller dots, while pairs involving SUDs or SUD-proxies (Cigsperday and AUD) are represented by larger dots. Pairs involving externalizing or internalizing traits are colored in orange or blue, respectively. Pairs involving tobacco use behaviors or disorders (EverSmk and Cigsperday) are depicted by squares and pairs involving alcohol use behaviors or disorders (DrnkWk and AUD) are represented by triangles.

Among the analyzed trait pairs, a subgroup exhibited a comparable degree of genetic overlap (> 90% of shared causal variants) (Figure 4). Some pairs involving SUDs and SUD-proxies with externalizing traits demonstrated a relatively high genetic correlation of shared variants (absolute values of r_{gs}=0.50-0.69), despite their high degree of genetic overlap. Conversely, pairs with a similar degree of genetic overlap comprising SUDs or SUD-proxies with internalizing traits demonstrated a lower genetic correlation of shared variants (absolute values of r_{gs}=0.06-0.42). However, it was much lower for SUBs, notably DrnkWk, and internalizing traits (absolute values of r_{gs}=0.02-0.16, except for EverSmk and DEP (absolute value of r_{gs}=0.35)).

4. DISCUSSION

Using MiXeR, we detected a distinct pattern of genetic overlap between SUBs and SUDs with externalizing and internalizing traits. SUBs (i.e., EverSmk and DrnkWk) and internalizing traits demonstrated a notably high genetic overlap but a particularly low genetic correlation of shared variants. This finding may be explained by a divergence in the effect distribution of the numerous shared risk variants between SUBs and internalizing traits, suggesting a limited role of the genetics of internalizing problems at SUBs. By contrast, SUBs showed a higher genetic correlation of shared variants with externalizing traits, suggesting that externalizing problems may have a stronger genetic influence on SUBs, albeit with variations depending on the substance. With regards to SUDs or SUD-proxies (i.e., Cigsperday and AUD), we observed notably high genetic correlation of shared variants with externalizing traits, also in cases of high genetic overlap. This observation suggests that most risk variants for SUDs and SUD-proxies are shared with externalizing problems, and these variants show a similar effect distribution on both traits. A similar pattern, although not as marked, was found for pairs involving SUDs and SUD-proxies with internalizing problems. These findings suggest that while both externalizing and internalizing pathways may be involved in SUDs, externalizing pathway seems to play a more consistent role.

Bivariate MiXeR revealed a substantial genetic overlap between SUBs and SUDs and both externalizing and internalizing traits. A similar pattern was already identified in a previous study from our group between mental disorders and other traits such as intelligence, educational

USC DE COMPOSTELA

attainment, neuroticism, and SWB (Hindley et al., 2022). This study concluded that a large number of pleiotropic variants are involved in all these traits with differences in the effect size distribution of these pleiotropic variants playing a more prominent role in the genetic predisposition for a given trait than trait-specific genetic variants. Considering our results, this conclusion can also be extended to the genetic overlap of externalizing and internalizing traits with SUBs and SUDs. Our study identified two distinct patterns within this general framework. SUBs and internalizing traits showed a high genetic overlap but a genetic correlation of shared variants close to zero, suggesting a distinctly divergent effect distribution within both traits. Conversely, certain pairs involving SUDs or SUD-proxies with externalizing traits displayed a similar distribution of effects as well as a high genetic overlap, suggesting a shared genetic basis among these traits.

Our results are consistent with previous studies suggesting that, while the internalizing pathway plays a limited role at the initial stages of substance consumption, it has a stronger impact at more severe levels of use (Hussong et al., 2011; Ning et al., 2020). A different genetic relationship of internalizing traits with SUBs and SUDs has already been suggested using MiXeR (Icick et al., 2022). In this previous study, mood instability showed a different pattern of genetic overlap with alcohol consumption, defined as AUDIT-C, and AUD. Here, we have expanded on these findings by using several internalizing traits in addition to mood instability and a different alcohol consumption trait (DrnkWk). Our findings are also consistent with previous studies suggesting a role of the externalizing pathway at all substance use stages. An effect of the externalizing pathway on substance use initiation was suggested in a previous study which found that the shared genetic susceptibility between schizophrenia and smoking initiation is related with externalizing behavior (Al-Soufi & Costas, 2023). Importantly, our results emphasize the substantial role of the externalizing pathway in SUDs, consistent with previous studies supporting a clearer association of externalizing problems with severe substance use compared to substance consumption and to internalizing problems (Edwards et al., 2016; Ning et al., 2020). Previous studies have implicated the genetic susceptibility to both externalizing and internalizing problems in SUDs. A recent study found that an addiction risk factor based on problematic alcohol use, problematic tobacco use, CUD and opioid use disorder GWAS correlated genetically with externalizing behavior (absolute r_g=0.35-0.52), as well as with substance use to cope with anxiety (r_g=0.63) and depression (r_g=0.60), suggesting a strong association of externalizing and internalizing behavior with SUDs (Hatoum et al., 2023). Another study based on a similar addiction risk factor found that the association between this addiction risk factor and risk-taking significantly decreased after taking substance use into account (Hatoum et al., 2022). In contrast, genetic susceptibility to substance use had a small

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

effect on the association between this addiction risk factor and neuroticism. Our findings are compatible with these results suggesting that, in contrast to the internalizing pathway, the effect of the externalizing pathway on SUDs may be partially mediated by substance use.

For SUDs and SUD-proxies, both tobacco and alcohol showed comparable patterns of genetic overlap with externalizing and internalizing traits. However, differences emerged between the tobacco (EverSmk) and alcohol (DrnkWk) use behaviors. EverSmk displayed a higher genetic correlation of shared variants with externalizing and internalizing traits compared to DrnkWk. Moreover, the genetic correlation of shared variants of internalizing traits with DrnkWk was close to 0, suggesting that the internalizing pathway may have a minimal effect on alcohol use. In contrast to tobacco use, DrnkWk is a highly heterogenous trait, combining frequency of use and quantity per drinking occasion (Mallard et al., 2022; Sanchez-Roige et al., 2021). These dimensions show opposite genetic correlations with traits related to socioeconomic status, psychiatric disorders, and psychological traits (Marees et al., 2019). As a consequence of this heterogeneity, DrnkWk is expected to be less influenced by the internalizing and externalizing pathways than use of tobacco or other substances.

The present study has some limitations. First, we cannot conclude that there is a causal effect of externalizing and internalizing problems on SUBs and SUDs on the basis of MiXeR results alone. We can only assess whether there is a shared genetic susceptibility between both traits and its effect distribution on both traits. Second, some of the MiXeR models were not distinguishable from a complete genetic overlap or the minimum genetic overlap allowed for a given genetic correlation. Therefore, larger GWAS based on a more precise phenotype definition are required to achieve more accurate estimates of the number of shared and unique genetic variants. Third, a lack of well-powered GWAS for other substances only allowed the analysis of substance use for two legal substances: alcohol and tobacco. Thus, our conclusions may not generalize to other substances. Fourth, a continuous trait for pathological alcohol use would assess more accurately the level of substance dependence of an individual. Fifth, there is considerable comorbidity between externalizing and internalizing behavior, which makes it difficult to distinguish the separate effect of each analyzed pathway. Moreover, many GWAS rely on samples like UK Biobank, which may not accurately represent the general population in terms of gender, socio-economic status or ancestry, potentially biasing our results (Fry et al., 2017). In addition, since MiXeR assumes that the effect size of causal variants is normally distributed, the cluster of nAchR genes at chromosome 15 was excluded from our Cigsperday GWAS. Furthermore, we did not provide statistical significance for our results, as formal significance tests based on MiXeR results are not feasible. Finally, although we used several

USC DE COMPOSTELA

sexual behaviors as representative traits of genetic susceptibility for externalizing behavior, these complex phenotypes are affected by gene-environment correlations attributable to social, economic and political processes (Abdellaoui et al., 2022). Therefore, a cautionary remainder is needed to avoid misinterpretation of genomic findings that may lead to stigmatization.

In summary, we found a different genetic relationship between SUBs and SUDs with externalizing and internalizing traits, suggesting a larger involvement of the externalizing and internalizing pathways in severe substance use. In addition, the externalizing pathway seems to play a more prominent role throughout the progression of substance use than the internalizing pathway. These findings improve our understanding of the origins of substance use and SUDs, which may inform attempts to advance prevention and treatment strategies.

USC
UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

4.3. ESTUDO 3: COLOCALIZATION OF ASSOCIATION SIGNALS AT NICOTINIC ACETYLCHOLINE RECEPTOR GENES BETWEEN SCHIZOPHRENIA AND SMOKING TRAITS

Referencia:

Al-Soufi, L. e 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

Información dos/as autores/as:

Laila Al-Soufi[1,2] e Javier Costas[1,3]

[1]Grupo de Xenética Psiquiátrica, Instituto de Investigación Sanitaria de Santiago de Compostela (IDIS), Galicia, España.

[2]Universidade de Santiago de Compostela (USC), Galicia, España.

[3]Complexo Hospitalario Universitario de Santiago de Compostela (CHUS), Servizo Galego de Saúde (SERGAS), Santiago de Compostela, Galicia, España.

Información do artigo:

DOI: 10.1016/j.drugalcdep.2021.108517

PMID: 33454625

Ligazón: https://doi.org/10.1016/j.drugalcdep.2021.108517

Versión orixinal publicada no Anexo II.

Información da revista:

Drug and alcohol dependence

ISSN: 1879-0046

Índices de calidade segundo o JCR en 2021:

- Factor de impacto: 4,852
- Factor de impacto nos últimos cinco anos: 5,166
- Psychiatry: Q2 (61/155)
- Substance abuse: Q2 (6/21), mais estivo situada no Q1 de 2013 a 2020.
- Psychiatry (SSCI): Q2 (44/143)
- Substance abuse (SSCI): Q1 (8/37)

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

Contribución da doutoranda na publicación:

Realización de todas as análises, redacción de parte do manuscrito inicial e revisión da versión final.

Autorización da revista:

A revista *Drug and alcohol dependence*, pertencente á editorial Elsevier, permite que a autora reutilice o artigo nesta tese (Anexo III).

Material suplementario do artigo:

O material suplementario deste artigo pode atoparse na seguinte ligazón: https://ars.els-cdn.com/content/image/1-s2.0-S0376871621000120-mmc1.docx

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

ABSTRACT

Background: Epidemiological data suggest that smoking may be a risk factor for schizophrenia (SCZ), but more evidence is needed. Two regions coding nicotinic acetylcholine receptor (nAchR) subunits, at *CHRNA2* and the *CHRNA5/A3/B4* cluster, were associated with SCZ in genome-wide association studies (GWAS). Additionally, a signal at *CHRNA4* is near significance. *CHRNA2* was also associated with cannabis use disorder (CUD). These regions were also associated with smoking behaviors. If tobacco is a risk factor, the GWAS signals at smoking behaviors and SCZ have to be due to the same causal variants, i.e. they have to colocalize, although colocalization does not necessarily imply causality. Here, we present colocalization analysis at these loci between SCZ and smoking behaviors.

Methods: The Bayesian approach implemented in coloc was used to check for posterior probability of colocalization versus independent signals at the three loci with some evidence of association with SCZ and smoking behaviors, using GWAS summary statistics. Colocalization was also assessed for positive control traits and CUD. Three different sensibility analyses were used to confirm the results. A visualization tool, LocusCompare, was used to facilitate interpretation of the coloc results.

Results: Evidence for colocalization of GWAS signals between SCZ and smoking behaviors was found for *CHRNA2*. Evidence for independent causal variants was found for the other two loci. CUD GWAS signal at *CHRNA2* colocalizes with SCZ and smoking behaviors.

Conclusions: Overall, the results indicate that the association between some nAchR subunit genes and SCZ cannot be solely explained by their effect on smoking behaviors.

USC
UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

1. INTRODUCTION

Epidemiological studies have found that smoking increases the posterior risk of developing schizophrenia (SCZ) and is associated with an earlier onset of psychosis in a dose-dependent manner, suggesting a role of smoking on risk for psychotic disorders (Gurillo et al., 2015; Kendler et al., 2015; Koyanagi et al., 2016). Genome-wide association studies (GWAS) have also contributed to this topic. GWAS of smoking behaviors identified genome-wide significant associations at almost all genes coding for neuronal nicotinic acetylcholine receptor (nAchR) subunits (M. Liu et al., 2019). A SCZ GWAS also identified a genome-wide significant association at the cluster of nAchR genes at chromosome 15, herein referred to as *CHRNA5/A3/B4* (Schizophrenia Working Group of the Psychiatric Genomics Consortium et al., 2014). This suggests that the association of *CHRNA5/A3/B4* with SCZ may be mediated by smoking, acting as a risk factor (Alderson & Lawrie, 2015; Gage & Munafò, 2015a; Laursen & McGrath, 2016; McGrath, 2015). Remarkably, taking into account that cannabis acts as a risk factor for SCZ (Belbasis et al., 2018; Di Forti et al., 2019), a GWAS on cannabis use disorder (CUD) identified just one GWAS significant hit, at *CHRNA2* (Demontis, Rajagopal, et al., 2019), a locus also associated with SCZ (Schizophrenia Working Group of the Psychiatric Genomics Consortium, 2014). We must bear in mind that while GWAS identify associated regions, the existence of high linkage disequilibrium (LD) precludes the identification of the specific causal SNP (or SNPs) responsible for the association. However, one SNP associated with several traits, rs16969968, deserves attention. It is a missense SNP at *CHRNA5*, changing the evolutionary conserved amino acid aspartate at position 398 of the encoded protein by asparagine. This change has functional consequences, both in vitro and in vivo (Maskos, 2020).

However, identification of an association did not prove causality, due to problems such as confounding or reverse causality. Further evidence is provided by the Mendelian randomization (MR) approach. Recently, a large MR study found a causal effect of smoking on SCZ (Wootton et al., 2020). However, although MR is a step forward towards causality, it presents several limits that preclude definitive conclusions (Davies et al., 2018; Hartwig et al., 2018; Hemani et al., 2018; Katikireddi et al., 2018). Different approaches are needed to triangulate the evidence pointing to a causal role of smoking on SCZ (Lawlor et al., 2016; Ohlsson & Kendler, 2020). To this aim, we applied for the first time colocalization analysis to this problem. Colocalization analysis is generally used to assess colocalization of GWAS signals between diseases (or related traits) and expression quantitative trait loci (eQTLs) (Giambartolomei et al., 2014; X. He et al., 2013; Hormozdiari et al., 2016), but may be used to compare the association signals at any two traits (Giambartolomei et al., 2014). If smoking is a risk factor for SCZ, i.e.,

USC DE COMPOSTELA

mediated (or vertical) pleiotropy (Hemani et al., 2018), variants at genes associated with both smoking and SCZ have to colocalize, although colocalization does not necessarily imply mediated pleiotropy.

2. METHODS

2.1. Datasets

Several publicly available GWAS summary statistics of different traits were analyzed (Table 1 and Supplementary Material). Colocalization between tobacco use and SCZ was assessed using GWAS data for seven smoking traits. GWAS on lung cancer and lung function (forced vital capacity) were used as positive controls, and a GWAS of CUD was used to assess a possible confounding effect.

Table 1. Datasets used in the present study.

Trait name[a]	Description	Source[b]	Sample size[c]	p *CHRNA5/A3/B4*[d]	p *CHRNA2*[d]	p *CHRNA4*[d]
CigDay	Average number of cigarettes smoked per day	GSCAN	337 334	4.8×10^{-211}	2.44×10^{-10}	2.47×10^{-18}
SmkCes	Smoking cessation, current vs former smokers	GSCAN	139 453 / 407 766	7.57×10^{-11}	$p > 1 \times 10^{-5}$	2.04×10^{-16}
AgeSmk	Age started smoking cigarettes regularly	GSCAN	341 427	$p > 1 \times 10^{-5}$	7.00×10^{-11}	$p > 1 \times 10^{-5}$
SmkInit	Ever vs never regular smokers	GSCAN	557 337 / 674 754	$p > 1 \times 10^{-5}$	3.42×10^{-17}	$p > 1 \times 10^{-5}$
FTND	Fagerström Test for Nicotine Dependence	Hancock et al., 2015	17 074	3.50×10^{-17}	$p > 1 \times 10^{-5}$	3.80×10^{-8}
Frequency	Current smoking frequency	GeneAtlas	501 649	8.35×10^{-7}	5.82×10^{-13}	1.52×10^{-13}
Attempts	Number of unsuccessful stop smoking attempts	Neale's lab[e]	81 026	6.45×10^{-10}	$p > 1 \times 10^{-5}$	$p > 1 \times 10^{-5}$
SCZ	Schizophrenia	PGC	33 640 / 43 456	2.26×10^{-13}	8.76×10^{-9}	9.45×10^{-8}
CUD	Cannabis use disorder	iPSYCH	2387 / 48 985	$p > 1 \times 10^{-5}$	9.31×10^{-12}	$p > 1 \times 10^{-5}$
LungCancer	Overall lung cancer	TRICL-ILCCO and LC3	44 069 / 68 712	3.08×10^{-103}	1.69×10^{-8}	$p > 1 \times 10^{-5}$
FVC	Lung function measured as the forced vital capacity	UKBB	321 047	6.95×10^{-10}	$p > 1 \times 10^{-5}$	3.04×10^{-6}

[a]Name used along the text.

[b]GSCAN, GWAS and Sequencing Consortium of Alcohol and Nicotine use (M. Liu et al., 2019); GeneAtlas (Canela-Xandri et al., 2018); PGC, Psychiatric Genomics Consortium (Schizophrenia Working Group of the Psychiatric Genomics Consortium, 2014); iPSYCH, Lundbeck Foundation Initiative for Integrative Psychiatric Research (Demontis, Rajagopal, et al., 2019); TRICL-ILCCO and LC3, Transdisciplinary Research of Cancer in Lung of the International Lung Cancer Consortium and Lung Cancer Cohort Consortium (McKay et al., 2017); UKBB, UK BioBank (Shrine et al., 2019).

[c]For dichotomous traits, case / control sample sizes are shown.

[d]p value of the most significant SNP at each locus.

[e]http://www.nealelab.is/uk-biobank/; GWAS round 2 results, released 1st August 2018.

2.2. Colocalization analysis

Coloc is a Bayesian statistical method to test for colocalization that works with the summary statistics of two association studies (Giambartolomei et al., 2014). This method enables the differentiation between colocalization, i.e., one common association signal for two traits, and a situation where independent signals are associated with each trait. It computes the posterior probability for five alternative hypotheses: H_0, no association with either trait; H_1, association with trait 1 but no with trait 2; H_2, association with trait 2 but no with trait 1; H_3, two independent association signals, one for each trait; or H_4, a common association signal for both traits. Assuming that the analyzed loci are associated with both tested traits, the power of the colocalization analysis to detect an association with both traits is indicated by the sum of the posterior probability for H_3 (PP_3) and the posterior probability for H_4 (PP_4). A threshold of $PP_3 + PP_4 > 0.8$ was set to filter comparisons with a high support for association with both traits. Among these comparisons, support for hypothesis H_3 and H_4 was assessed as the ratio of posterior probabilities, considering evidence for support if the posterior probability of one hypothesis was more than three times the posterior probability for the other hypothesis, as previously done (Bhalala et al., 2018). The prior probabilities for SNP association with either of the two traits or with both traits were set to 1×10^{-4} and 1×10^{-5}, the default values, respectively.

Comparisons were restricted to traits from the same ethnic background without overt overlapping samples, owing to the assumptions of the method (Giambartolomei et al., 2014), and to loci with any SNP in the ±200 kb region around the transcribed region of nAchR subunits genes with $p < 1x10^{-5}$ in their respective GWAS (Table 1).

Colocalization events were visualized with LocusCompare which generates a useful plot to differentiate between two distinct signals and a common signal for both traits (B. Liu et al., 2019). LD score regression (B. Bulik-Sullivan et al., 2015) was used to check the existence of cryptic samples overlap (Supplementary Material).

2.3. Sensitivity analyses

Three different sensitivity analyses were performed considering different values for prior probabilities, correcting for putative samples overlap, and considering different causal variants (Supplementary Methods).

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

3. RESULTS

3.1. Association of genes coding nAchR subunits with traits under study

Searching for significant SNPs at ± 200kb of the transcribed region of nAchR subunits genes in the GWAS summary statistics from the Psychiatric Genomics Consortium identified three loci associated with SCZ at $p < 1x10^{-5}$, involving *CHRNA2*, *CHRNA4*, and *CHRNA5/A3/B4* (Table 1). These three GWAS association signals were the focus of the colocalization analysis. Table 1 also shows the GWAS results at these loci for smoking behaviors, control traits and CUD.

3.2. Colocalization between control traits and smoking traits

Table 2. Results of colocalization analysis for comparisons showing evidence for colocalization (H_4) or two independent signals (H_3) at the main analysis.

Locus	Trait 1	Trait 2	Coloc main test[a] H (PP_3+PP_4, PP_4/PP_3)	Sensitivity I[b] H (PP_3+PP_4, PP_4/PP_3)	Sensitivity II[c] H (PP_3+PP_4, PP_4/PP_3)	Sensitivity III[d] H (PP_3+PP_4, PP_4/PP_3)
CHRNA5/A3/B4	LungCancer	SmkCes	H_4 (>0.99, 32.50)	H_4 (>0.99, 3.25)	-	H_3 (0.88, 3.3e-5)
CHRNA5/A3/B4	LungCancer	Attempts	H_4 (>0.99, 25.34)	-	-	H_3 (>0.99, 1.0e-5)
CHRNA5/A3/B4	LungCancer	FTND	H_4 (>0.99, 7.32)	-	-	-
CHRNA5/A3/B4	LungCancer	Frequency	H_3 (0.89, 2.6e-3)	H_3 (0.89, 2.6e-4)	-	H_3 (0.89, 3.5e-3)
CHRNA5/A3/B4	LungCancer	CigDay	H_3 (>0.99, 5.2e-3)	H_3 (>0.99, 5.2e-4)	-	H_3 (0.88, 1.66e-202)
CHRNA2	LungCancer	AgeSmk	H_4 (>0.99, 249.12)	H_4 (>0.99, 24.91)	-	H_3 (0.93, 3.5e-7)
CHRNA2	LungCancer	SmkInit	H_4 (>0.99, 10.82)	-	-	H_3 (0.94, 1.9e-9)
CHRNA2	LungCancer	Frequency	H_4 (>0.99, 5.56)	-	-	H_3 (0.94, 1.7e-8)
CHRNA2	LungCancer	CigDay	H_4 (>0.99, 4.02)	-	-	H_3 (0.94, 3.5e-4)
CHRNA5/A3/B4	FVC	FTND	H_4 (>0.99, 8.76)	-	-	-
CHRNA5/A3/B4	SCZ	Frequency	H_4 (0.99, 7.84)	-	H_4 (0.94, 141.90)	none
CHRNA5/A3/B4	SCZ	CigDay	H_3 (>0.99, 3.4e-5)	H_3 (>0.99, 3.4e-6)	H_3 (>0.99, 2.0e-3)	H_3 (>0.99, 1.2e-8)
CHRNA5/A3/B4	SCZ	FTND	H_3 (>0.99, 1.2e-4)	H_3 (>0.99, 1.1e-5)	-	-
CHRNA5/A3/B4	SCZ	SmkCes	H_3 (>0.99, 2.8e-1)	H_3 (>0.99, 2.8e-2)	-	none
CHRNA2	SCZ	CigDay	H_4 (>0.99, 376.05)	H_4 (>0.99, 37.61)	H_4 (>0.99, 1596.63)	none
CHRNA2	SCZ	SmkInit	H_4 (>0.99, 33.42)	H_4 (>0.99, 3.34)	H_4 (>0.99, 246.22)	none
CHRNA2	SCZ	Frequency	*H_4 (>0.99, 15.24)*	-	H_4 (>0.99, 272.74)	none
CHRNA2	SCZ	AgeSmk	H_3 (>0.99, 2.1e-1)	H_3 (>0.99, 2.1e-2)	-	none
CHRNA4	SCZ	SmkCes	H_3 (0.95, 4.5e-4)	H_3 (0.95, 4.5e-4)	H_3 (>0.99, 1.2e-3)	H_3 (0.93, 9.4e-4)
CHRNA4	SCZ	CigDay	H_3 (0.95, 1.2e-3)	H_3 (0.95, 1.2e-4)	H_3 (>0.99, 5.7e-3)	none
CHRNA4	SCZ	Frequency	H_3 (0.95, 6.1e-3)	H_3 (0.95, 6.1e-4)	H_3 (>0.99, 4.6e-2)	none
CHRNA2	CUD	SCZ	H_4 (>0.99, 81.44)	H_4 (>0.99, 8.14)	H_4 (>0.99, 2044.49)	none

CHRNA2	CUD	SmkInit	H_4 (>0.99, 55.29)	H_4 (>0.99, 5.53)	H_4 (>0.99, 9047.79)	none
CHRNA2	CUD	Frequency	H_4 (>0.99, 11.88)	-	H_4 (>0.99, 884.99)	none
CHRNA2	CUD	CigDay	H_4 (>0.99, 5.47)	-	H_4 (>0.99, 24.86)	none
CHRNA2	CUD	AgeSmk	H_3 (>0.99, 2.4e-1)	H_3 (>0.99, 2.4e-2)	-	none

[a]Those comparisons with LD score regression intercepts higher than one standard error away from zero, suggesting cryptic samples overlap, are in italics.

[b]Results for sensitivity analysis using a different prior probability for colocalization.

[c]Results for sensitivity analysis using gwas-pw to correct for samples overlap. This analysis was not applied to comparisons involving LungCancer or FTND because it requires full GWAS summary statistics. The different hypotheses were named following coloc nomenclature.

[d]Results for sensitivity analysis using masking to consider the existence of more than one causal variant. Results are shown for additional pair of variants supporting any hypothesis, after masking of signals of the main analysis. "none" means that there were no additional signals at these comparisons. In case of more than one additional pair of signals, the one with the strongest support is shown. Full results are shown in Supplementary Table 1. The analysis was not applied to comparisons involving FTND because the required beta values and standard errors were not supplied in the FTND GWAS summary statistics.

Coloc identified colocalization between LungCancer and three smoking traits at *CHRNA5/A3/B4* (Table 2, Figure S1). For one additional trait, CigDay, while coloc supported independent signals, visual inspection of the comparison of these two GWAS results using LocusCompare suggested that there was colocalization, but the strong LD among many SNPs led coloc to discard colocalization (Table 2, Figure S1). The shared colocalization signal was related to rs16969968. There was evidence of two independent signals in the comparison with Frequency, a trait that did not show association with rs16969968 (Table 2, Figure S1). Regarding FVC, comparison with the only smoking trait without overlapping samples, FTND, revealed evidence of colocalization (Table 2, Figure S2).

For *CHRNA2*, there was evidence of colocalization between LungCancer and all the smoking behaviors tested (Table 2, Figure S1).

Regarding *CHRNA4*, the only comparison with non-overlapped samples was FVC versus FTND. Analysis of this comparison suggested colocalization ($PP_4/PP_3 = 35.9$), but with limited power ($PP_4 + PP_3 = 0.73$) (Figure S2).

3.3. Colocalization between SCZ and smoking traits

In the case of *CHRNA5/A3/B4*, there were four comparisons between SCZ and smoking traits with enough power. Three of them supported independent signals, while that involving Frequency supported colocalization (Table 2). LocusCompare plots revealed that independent signals were due to the strong association of smoking traits with rs16969968, lacked in the case of SCZ (Figure S3).

For *CHRNA2*, three of the four tested comparisons supported colocalization (Table 2, Figure S3). The remaining analysis, involving AgeSmk, supported the existence of two different signals (Table 2, Figure S3). Finally, there was evidence for independent signals at *CHRNA4*, supported by three different smoking traits (Table 2, Figure S3).

3.4. Colocalization analysis with CUD

Taking into account that tobacco and cannabis use are correlated, we performed colocalization analysis between CUD, SCZ and smoking traits at *CHRNA2*. The signal at this locus colocalized with SCZ (Table 2, Figure S4). In addition, there was evidence for colocalization between CUD and the same smoking traits that colocalized with SCZ (Table 2, Figure S4). The comparison between CUD and AgeSmk revealed independent signals, as was the case for SCZ (Table 2, Figure S4).

3.5. Sensitivity analyses

Sensitivity analyses confirmed the main results for comparisons involving SCZ, although additional causal variants with different evidence of colocalization appeared for LungCancer (Table 2, Table S1).

4. DISCUSSION

In this work, we analyzed colocalization at nAchR subunit genes to provide new data on the debate about the relationship between smoking and SCZ based on a different approach. Our results at *CHRNA5/A3/B4* are against the causal relationship, as we found strong evidence supporting the existence of independent association signals at SCZ and smoking traits. In the case of several conditional independent associations at one locus, as in *CHRNA5/A3/B4* for CigDay (M. Liu et al., 2019), coloc considers only the strongest signal (Giambartolomei et al., 2014). Therefore, we can conclude that the strongest signal at this locus responsible for association with smoking traits such as CigDay, SmkCes, or FTND, is not the strongest signal associated with SCZ. Moreover, additional signals at this locus in smoking traits also did not colocalize with SCZ. By contrast, we have found strong evidence of colocalization between the same smoking traits and LungCancer, used as positive control. As clearly revealed by LocusCompare, the signal shared between smoking traits and LungCancer is related to the missense SNP rs16969968 at *CHRNA5*. Therefore, contrary to smoking traits, the strongest association signal for SCZ at *CHRNA5/A3/B4* is not lead by rs16969968.

Regarding *CHRNA4*, coloc supports the existence of two independent signals, one associated with SCZ and another with smoking traits. In fact, this locus was not associated at genome-

wide significance with LungCancer, suggesting that the effect of this locus mediating the role of smoking on other traits may be very low.

Finally, colocalization of GWAS signals between SCZ, LungCancer, CUD and smoking traits is the most probable hypothesis at *CHRNA2*. Therefore, we cannot discard a role of cannabis use as a confounding factor on the association between SCZ and smoking (Gage et al., 2014; Jones et al., 2018).

Our approach presents several strengths and limitations. Among the strengths, the approach is orthogonal to other approaches, providing independent results to address causality. Additionally, we have used several datasets based on meta-analysis of many different cohorts with different ascertainment schemes, bias and demographic characteristics, and covering a large range of smoking behaviors to add confidence to our findings. Finally, the use of coloc has the advantage of easy interpretation of results, comparing evidence among different hypotheses, while LocusCompare helps to visually inspect congruence of coloc results.

Regarding limitations, coloc assumes that there is just one causal variant per locus. However, coloc is robust to the existence of several causal variants if they did not explain similar proportion of the variance of the trait (Giambartolomei et al., 2014). Moreover, coloc only compares GWAS from independent samples, precluding the use of additional positive controls. There is also a problem in calculation of posterior probabilities if there are many SNPs at strong LD with the causal variants, as exemplified in the case of LungCancer and CigDay at *CHRNA5/A3/B4*. Other limitation is that, as a Bayesian approach, prior probabilities have to be defined. We have used positive control traits and several sensitivity analyses to deal with these limitations, suggesting that the results are robust. Finally, colocalization does not imply mediated pleiotropy. Therefore, horizontal pleiotropy and reverse causality cannot be discarded.

In summary, the shared association signal at *CHRNA2* is compatible with putative mediated pleiotropy, i.e., its effect on SCZ risk may be due to tobacco and/or cannabis use. The lack of colocalization at the *CHRNA5/A3/B4* cluster and *CHRNA4* suggests additional involvement of nAchRs in SCZ risk, irrespective of smoking. These findings may have implications on the potential of nAchRs as novel drug targets for treatment of SCZ (Terry et al., 2015; Wilkerson et al., 2020).

USC
UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

4.4. ESTUDO 4: A POLYGENIC APPROACH TO THE ASSOCIATION BETWEEN SMOKING AND SCHIZOPHRENIA

Referencia:

Al-Soufi, L., Martorell, L., Moltó, M. D., González-Peñas, J., García-Portilla, M. P., Arrojo, M., Rivero, O., Gutiérrez-Zotes, A., Nácher, J., Muntané, G., Paz, E., Páramo, M., Bobes, J., Arango, C., Sanjuan, J., Vilella, E. e Costas, J. (2021). A polygenic approach to the association between smoking and schizophrenia. *Addiction biology, 27*(1), e13104. https://doi.org/10.1111/adb.13104

Información dos/as autores/as:

Laila Al-Soufi[1,2], Lourdes Martorell[3,4], M. Dolores Moltó[4,5,6], Javier González-Peñas[4,7], Mª Paz García-Portilla[4,8], Manuel Arrojo[1,9], Olga Rivero[4,5,6], Alfonso Gutiérrez-Zotes[3,4], Juan Nácher[4,5,10], Gerard Muntané[3,4], Eduardo Paz[1,9], Mario Páramo[1,9], Julio Bobes[4,8], Celso Arango[4,7], Julio Sanjuan[4,5,11], Elisabet Vilella[3,4] e Javier Costas[1,12]

[1]Grupo de Xenética Psiquiátrica, Instituto de Investigación Sanitaria de Santiago de Compostela (IDIS), Santiago de Compostela, Galicia, España.

[2]Departamento de Zooloxía, Xenética e Antropoloxía Física, Universidade de Santiago de Compostela (USC), Santiago de Compostela, Galicia, España.

[3]Hospital Universitari Institut Pere Mata (HUIPM); Institut d'Investigació Sanitària Pere Virgili (IISPV); Universitat Rovira i Virgili (URV), Reus, España.

[4]Centro de Investigación Biomédica en Rede de Saúde Mental (CIBERSAM), Madrid, España.

[5]INCLIVA Instituto de Investigación Sanitaria, Fundación Investigación Hospital Clínico de Valencia, Valencia, España.

[6]Departament de Genètica, Universitat de València, Valencia, España.

[7]Departamento de Psiquiatría del Niño y del Adolescente, Instituto de Psiquiatría y Salud Mental, Hospital General Universitario Gregorio Marañón, Facultad de Medicina, Universidad Complutense de Madrid, Instituto de Investigación Sanitaria Gregorio Marañón (IiSGM), Madrid, España.

[8]Departamento de Psiquiatría, Universidad de Oviedo, Instituto de Investigación Sanitaria del Principado de Asturias (ISPA), Instituto Universitario de Neurociencias del Principado de Asturias (INEUROPA), Servicio de Salud del Principado de Asturias (SESPA) Oviedo, España.

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

[9]Servizo de Psiquiatría, Complexo Hospitalario Universitario de Santiago de Compostela, Servizo Galego de Saúde (SERGAS), Santiago de Compostela, Galicia, España.

[10]Departament de Biologia Celular, Institut Universitari de Biotecnologia i Biomedicina (BIOTECMED), Universitat de València, Valencia, España.

[11]Departamento de Psiquiatria, Facultat de Medicina, Universitat de València, Valencia, España.

[12]Complexo Hospitalario Universitario de Santiago de Compostela (CHUS), Servizo Galego de Saúde (SERGAS), Santiago de Compostela, Galicia, España.

Información do artigo:

DOI: 10.1111/adb.13104

PMID: 34779080

Ligazón: https://doi.org/10.1111/adb.13104

Versión orixinal publicada no Anexo II.

Información da revista:

Addiction biology

ISSN: 1369-1600

Índices de calidade segundo o JCR en 2021:

- Factor de impacto: 4,093
- Factor de impacto nos últimos cinco anos: 4,323
- Biochemistry and molecular biology: Q3 (156/297)
- Substance abuse: Q3 (11/21), mais estivo situada en Q1 dende 2011 a 2020.

Contribución da doutoranda na publicación:

Realización de todas as análises, salvo o tratamento dos datos de xenotipado. Redacción de gran parte do manuscrito inicial e revisión da versión final.

Autorización da revista:

A revista *Addiction biology*, pertencente á editorial *John Wiley and Sons*, deu licenza á doutoranda a reutilizar este artigo na tese (Anexo III).

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

Material suplementario do artigo:

O material suplementario deste artigo pode atoparse na seguinte ligazón: https://onlinelibrary.wiley.com/action/downloadSupplement?doi=10.1111%2Fadb.13104&file=adb13104-sup-0001-SupportingInformation_Resub.docx

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

ABSTRACT

Smoking prevalence in schizophrenia is considerably larger than in general population, playing an important role in early mortality. We compared the polygenic contribution to smoking in schizophrenic patients and controls to assess if genetic factors may explain the different prevalence. Polygenic risk scores (PRSs) for smoking initiation and four genetically correlated traits were calculated in 1108 schizophrenic patients (64.4% smokers) and 1584 controls (31.1% smokers). PRSs for smoking initiation, educational attainment, body mass index and age at first birth were associated with smoking in patients and controls, explaining a similar percentage of variance in both groups. Attention-deficit hyperactivity disorder (ADHD) PRS was associated with smoking only in schizophrenia. This association remained significant after adjustment by psychiatric cross-disorder PRS. A PRS combining all the traits was more explanative than smoking initiation PRS alone, indicating that genetic susceptibility to the other traits plays an additional role in smoking behavior. Smoking initiation PRS was also associated with schizophrenia in the whole sample, but the significance was lost after adjustment for smoking status. This same pattern was observed in the analysis of specific SNPs at the *CHRNA5-CHRNA3-CHRNB4* cluster associated with both traits. Overall, the results indicate that the same genetic factors are involved in smoking and in general population, and are compatible with smoking acting, directly or indirectly, as a risk factor for schizophrenia that contributes to the high prevalence of smoking in these patients. The contrasting results for ADHD PRS may be related to higher ADHD symptomatology in schizophrenic patients.

USC
UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

1. INTRODUCTION

The life expectancy in schizophrenia is reduced compared with the general population by almost 15 years (Hjorthøj et al., 2017). Smoking is an important contributor to this mortality gap, due to its higher prevalence in schizophrenic patients (SzP) than in general population (Bobes et al., 2010). This association has been usually explained by the self-medication hypothesis, which states that SzP may smoke to reduce symptoms of the disorder or side effects of antipsychotic medication (Goff et al., 1992; Kumari & Postma, 2005). However, prospective studies have recently suggested a risk effect of smoking on the development of schizophrenia, finding an association between tobacco use and subsequent psychotic experiences (Gurillo et al., 2015; Koyanagi et al., 2016) and schizophrenia (Kendler et al., 2015; Weiser et al., 2004).

Interestingly, genome-wide association studies (GWAS) revealed that the *CHRNA5-CHRNA3-CHRNB4* cluster on chromosome 15, coding for neuronal nicotinic acetylcholine receptor subunits, is associated with schizophrenia and different smoking traits (J. Chen et al., 2020; Hancock et al., 2015; M. Liu et al., 2019; Ohi et al., 2019; Schizophrenia Working Group of the Psychiatric Genomics Consortium et al., 2014). Consequently, the reason for the high prevalence of smoking in SzP is currently under debate and the consideration of tobacco as a risk factor for schizophrenia is becoming strongly supported (Gage & Munafò, 2015a; MacCabe, 2018; G. M. J. Taylor & Munafò, 2019). Mendelian randomization (MR) studies, which use genetic variants associated with the risk factor of interest as instrumental variables to assess causality, are an alternative to observational studies. The largest MR study on this topic supported both nonexclusive hypotheses, presenting strong evidence for a casual effect of smoking behavior on schizophrenia as well as some evidence for a causal effect of schizophrenia on lifetime smoking (Wootton et al., 2020).

Smoking behavior, assessed by different smoking phenotypes, is genetically correlated with schizophrenia (Hartz et al., 2018). Polygenic risk scores (PRSs), as a measure of individual genetic susceptibility to a trait, were previously used to demonstrate a shared genetic susceptibility, i.e. pleiotropy, between schizophrenia and tobacco use based on the association of the PRS for schizophrenia or smoking with the other trait in a target sample (J. Chen et al., 2016; Ohi et al., 2020; Reginsson et al., 2018). However, these studies are not able to determine if the shared genetic variants independently contribute to each trait, i.e. horizontal pleiotropy, or if both traits are in the same causal path, i.e. vertical pleiotropy. Information on smoking behavior in both SzP and controls is required to differentiate between both types of pleiotropy (Gage, Davey Smith, et al., 2016; Gage & Munafò, 2015a). Specifically, the association of

smoking behavior PRS with schizophrenia will persist after adjustment for smoking status only in case of horizontal pleiotropy.

The aim of the present study was to assess if genetic susceptibility to smoke differs between SzP and general population. Specifically, we compared the percentage of variance explained in smoking status by smoking PRS from GWAS of general population in SzP versus controls. Additionally, we compared the percentage of variance explained in smoking status by PRSs for traits genetically correlated with smoking. If SzP smoke mainly due to different reasons than general population, association of PRSs for these traits with smoking status will be weaker in SzP than in controls. We also assessed the association of smoking PRS with schizophrenia after adjustment by smoking status, as well as the specific effect of the cluster of neuronal nicotinic acetylcholine receptor subunits at chromosome 15, in order to add new data on the debate about the cause behind shared genetic susceptibility.

2. MATERIALS AND METHODS

2.1. Target samples

The initial target sample comprised 1239 patients with a diagnosis of schizophrenia according to DSM-IV and 1803 controls from five different recruitment sites of Spain at Santiago de Compostela, Oviedo, Reus, Valencia and Madrid. All subjects were from Spain and of European origin. Around 75% of controls from Santiago de Compostela were unscreened blood donors. The remaining controls were screened subjects to discard main psychiatric disorders. Adolescence/young adulthood is the relevant period for initiation of smoking and transition to regular smoking. Therefore, subjects under 20 years of age were excluded from further analysis. Recruitment of subjects was carried out at different periods of time from 1999 to 2017, of relevance taking into account that the first smoking-free law in Spain was implemented at the beginning of 2006. Subjects were recruited after informed consent; research was approved by the appropriate ethics committees and followed the principles of the Declaration of Helsinki.

Smoking status of the subjects, considered as a binary trait at the time of assessment, was determined by clinical interview or, if this information is missed, by a biological measure of smoking, methylation analysis of the aryl hydrocarbon receptor repressor *(AHRR)* gene. Quantitative methylation analysis was performed using the Agena (formerly Sequenom) MassARRAY technology and the assay described in Tian et al. (2017) (see Supplementary Text, Figure S1, Figure S2 and Table S1 for more information). The Youden's index (specificity + sensitivity - 1) was used to determine the optimal cut-off point for classification based on methylation levels in a training sample formed by 514 individuals for which smoking status by clinical interview was also known. A gray zone with low prediction ability was defined as the

USC DE SANTIAGO DE COMPOSTELA

95% CI of the optimal cut-off point, estimated by 5000 bootstrap resampling replicates, using pROC v.1.16.2 (Robin et al., 2011). Those samples with a methylation level within the gray zone (N = 219) were removed from the analysis. By this procedure, the diagnostic accuracy of the test, measured by the area under the receiver operating characteristic curve (AUC ROC) was 0.9214.

Genotyping was performed using the Illumina Infinium PsychArray at the Broad Institute or the Spanish National Genotyping Center. General quality control filters were applied using PLINK v1.9 including the percentage of failed genotypes by sample (2.5%) and by SNP (1%), departure for Hardy-Weinberg equilibrium ($P < 1x10^{-5}$), sex incongruencies, excess of heterozygosity (> 3 sd), cryptic relatedness (proportion of genome identical by descent > 0.2), or a difference in missing genotype rates between cases and controls ($P < 0.001$) (C. C. Chang et al., 2015). Autosomal SNPs at minor allele frequency (MAF) higher than 1% were retained for imputation, using the Michigan Imputation Server v1.0.4 (Das et al., 2016). Prior to imputation, data were prepared and checked following the McCarthy's lab instructions (https://www.well.ox.ac.uk/~wrayner/tools/). Human Reference Consortium dataset v.1.1 was used as haplotype references (McCarthy, 2016). After imputation, SNPs with imputation quality Rsq > 0.8 and MAF > 0.01 were retained for analysis.

PLINK v1.9 was used for principal component analysis to obtain the first 10 principal components to exclude outliers and control for population structure. Regions of long-range linkage disequilibrium that determined the existence of specific principal components in a preliminary analysis, such as the major histocompatibility complex at chromosome 6, the inversion polymorphism at 8p23, or complex rearrangement at the odorant receptors cluster at 11q11, were removed prior to the final principal components analysis.

The final sample comprised 2692 subjects. Of them, 1205 individuals were classified as smokers, 713 SzP and 492 controls, and 1487 individuals were classified as non-smokers, 395 SzP and 1092 controls. The main characteristics of the samples are shown in Table 1.

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

Table 1. Characteristics of the samples included in the main analysis.

Recruitment site	Group[a]	Sample size	Females (%)	Age (median, IQR)[b]	Smokers (%)	Methylation (%)[c]	Recruitment after 2006 (%)
Santiago	Controls	676	44.4	44.0, 35.0 - 51.0	20.6	68.5	27.2
	SzP	453	31.3	46.0, 37.0 - 59.0	60.5	76.4	0
Reus	Controls	538	53.2	41.0, 31.0 - 53.0	36.8	5.8	0
	SzP	322	35.1	45.5, 33.0 - 63.0	65.8	4.7	6.5
Valencia	Controls	252	32.5	37.0, 29.0 - 44.0	46.8	100	72.2
	SzP	272	31.6	40.5, 33.0 - 48.0	72.1	28.3	38.2
Madrid	Controls	118	55.9	38.0, 29.0 - 47.0	31.4	55.1	100
	SzP	6	16.7	27.5, 24.0 - 34.8	50.0	33.3	100
Oviedo	SzP	55	32.7	40.0, 33.0 - 49.5	50.9	100	0
All sites	Controls	1584	46.4	41.0, 32.0 - 51.0	31.1	51.2	30.6
	SzP	1108	32.5	44.0, 34.0 - 56.2	64.4	44.7	11.8

[a]SzP: schizophrenic patients.

[b]Age at recruitment.

[c]Percentage of subjects whose smoking status was determined by methylation as information of smoking status by clinical interview was missing.

2.2. Discovery samples

Several discovery samples were used for PRS calculation (Table S2). Smoking status risk estimation was based on smoking initiation (SmkInit) GWAS data from the GWAS and Sequencing Consortium of Alcohol and Nico tine use (GSCAN) (M. Liu et al., 2019). SmkInit was defined as a binary phenotype that compared ever versus never regular smokers (M. Liu et al., 2019). The choice of this discovery sample was based on its large sample size and its similarity with the smoking phenotype collected in our target sample, taking into account the low specificity of PRSs based on different smoking behaviors (Deutsch & Selya, 2020; Sanchez-Roige et al., 2021).

GWAS on additional traits were also used as discovery sample in PRS analysis. Specifically, all traits with a significant genetic correlation at the Bonferroni-correction level with SmkInit according to M. Liu et al. (2019) and with publicly available GWAS summary statistics data were selected for power calculation, estimated using the R package AVENGEME (Palla & Dudbridge, 2015). Those traits with power > 80% in SzP and/or controls were used in analyses, specifically age at first birth (AFB), i.e., age when having the first live child, attention-deficit hyperactivity disorder (ADHD), educational attainment (EA) and body mass index (BMI) (Table S2).

The last cross-disorder GWAS meta-analysis from the Psychiatric Genomics Consortium (PGC), not tested by M. Liu et al. (2019), was also used as a discovery sample to adjust by general psychopathology risk (P. H. Lee et al., 2019). In this cross-disorder meta-analysis,

GWAS data for eight psychiatric disorders (anorexia nervosa, autism spectrum disorder, ADHD, bipolar disorder, major depression, obsessive-compulsive disorder, Tourette syndrome and schizophrenia) were analyzed.

2.3. PRS calculation

EB-PRS, a novel PRS estimation method based on Empirical Bayes theory, was used for PRS calculation (S. Song et al., 2020). This approach estimates the effect sizes from the SNPs inferring the prior overall distribution of effect sizes of all SNPs in the discovery sample and leveraging this distribution to minimize prediction error. External information or parameter selection are not needed. For quantitative traits, the required case and control sample sizes were set as half of the sample size. In order to deal with linkage disequilibrium (LD), PRS models were built with the clumping algorithm of PLINK v.1.9, considering an $r^2 = 0.1$ and a window size of 500 kb to select those independent SNPs with a more significant association in the discovery sample. Prior to clumping, uncommon SNPs (MAF < 1%) and ambiguous SNPs (A/T or C/G alleles) were removed. Finally, PLINK v.1.9 was used to calculate the PRS for each individual in target sample as the sum of the number of effect alleles for all SNPs in the PRS model carried by that individual weighted by its effect size estimated with EB-PRS.

Due to the novelty of the EB-PRS method, EB-PRS was compared with the widely used method of thresholding, considering 31 significance thresholds ranging from 5×10^{-8} to 1. The weight of the SNPs was defined as the logarithm of the odds ratio. Correlation between the PRSs estimated at each threshold and with EB-PRS is shown in Figure S3.

2.4. Multitrait PRS calculation

SMTpred was used to combine SmkInit PRS and PRSs for those genetically correlated traits with enough power (AFB, ADHD, EA and BMI) in a multitrait PRS to improve prediction (Maier et al., 2018). SMTpred combines the different PRSs using weights based on sample size and SNP heritability of the GWAS discovery sample as well as genetic correlation between the different individual PRSs. SNP heritability and genetic correlation from GWAS summary statistics were calculated using LD Score regression (B. Bulik-Sullivan et al., 2015). LD Scores based on 1000 Genomes European data were used and GWAS summary statistics were filtered to HapMap3 SNPs. In the case of SNP heritability, proportion of cases and prevalence in discovery sample were used for the conversion to liability scale (Table S2).

2.5. Statistical analyses

All statistical analyses were conducted after the standardization of the PRSs to a mean of 0 and a standard deviation of 1 for easy interpretation. The main analysis tested the PRS association

with smoking status (smoker/non-smoker) of the target sample stratified by schizophrenic status (SzP/controls). PRS association with smoking status was assessed by a logistic regression model including sex, age at recruitment, recruitment site, recruitment year (coded as a binary variable considering recruitment before or after the first Spanish smoking-free law in January 2006), the method through which smoking status was determined and 10 principal components of ancestry as covariates. The percentage of variance explained by the PRS was estimated as the increase in R^2 between the full model and the model without the PRS. Age was missing for 3.0% of samples. In these subjects, the mean value of subjects at the same recruitment size, sex, smoking status and schizophrenic status was used as age. Recruitment year was missing for 3.2% of subjects. The mode for the same recruitment site and schizophrenic status was used in those instances.

The R^2 proposed by S. H. Lee et al. (2012) was used to estimate the proportion of variance explained by the PRS on the liability scale corrected for ascertainment bias due to oversampling of cases relative to the population prevalence. The main advantage of this estimate is its validity to compare the goodness-of-fit of different models. Because of that, it is the main statistic presented across the text. A prevalence of schizophrenia in general population of 1% was assumed for the calculation of the R^2 on the liability scale. The prevalences of smoking in SzP (54.54%) and in controls (31.51%) were taken from the CLAMORS schizophrenia cohort, a multicenter cross-sectional study based on Spanish population. As a sensitivity analysis, we estimated the R^2 on the liability scale using different prevalence values. CI for the increase in R^2 was estimated through bootstrap resampling (N = 5000), applying the Bias-Corrected and Accelerated Interval (BCa) method.

Significance of the interaction between schizophrenic status and each PRS on smoking status was assessed by inclusion of the interaction term in the logistic regression model. The model also included interaction terms between all covariates and PRS and between all covariates and schizophrenic status to control for the effects that these covariates might have on the interaction term of interest (Keller, 2014). Post-hoc power analysis for the test of interaction was assessed by simulation in R, as indicated (see Figure S4 for details) (Kleinman & Horton, 2014).

Association of SNPs at *CHRNA5-CHRNA3-CHRNB4* cluster was analyzed by logistic regression, using an additive genetic model. Taking into account that the strongest signals in the GWAS of schizophrenia and smoking traits did not colocalized (Al-Soufi & Costas, 2021), two SNPs were considered: the functional SNP D398N at *CHRNA5* (dbSNP ID: rs16969968) (Maskos, 2020), associated with several smoking traits (J. Chen et al., 2020; M. Liu et al., 2019), and the most significantly associated SNP with schizophrenia at this cluster, rs8042374

(Schizophrenia Working Group of the Psychiatric Genomics Consortium et al., 2014). Both SNPs are in partial LD ($r^2 = 0.18$ in European samples of the 1000 Genomes Phase 3 according to LDlink) (Machiela & Chanock, 2015).

3. RESULTS

3.1. Contribution of SmkInit PRS to smoking status in SzP and controls

Our study was well powered to analyze the effect of SmkInit PRS on smoking status in our samples (Table S2). SmkInit PRS was positively associated with smoking status in both SzP and controls (Table 2). The percentage of variance explained, corrected by prevalence and ascertainment bias, was slightly higher in controls (increase in R^2 (95% CI) = 2.05 (0.76 – 4.47)) than in SzP (increase in R^2 (95% CI) = 1.87 (0.54 – 4.52)) (Figure 1 and Table S3). Interaction term between SmkInit PRS and schizophrenic status had no significant effect on smoking status (P = 0.15). Sensitivity analysis using different prevalence values confirmed this finding (Table S4).

Table 2. Results of association analysis between polygenic risk scores (PRSs) and smoking status in schizophrenic patients (SzP) and controls.

PRS[a]	Target sample	Increase in Nagelkerke's R^2	Odds ratio (95% CI)	P value	Adjusted by SmkInit PRS	
					Odds ratio (95% CI)	P value
SmkInit	SzP	1.54	1.30 (1.13 - 1.49)	1.65×10^{-4}[b]	-	-
	Controls	1.68	1.30 (1.16 - 1.45)	6.97×10^{-6}[b]	-	-
AFB	SzP	0.89	0.82 (0.71 - 0.94)	4.07×10^{-3}[b]	0.85 (0.74 - 0.99)	0.0313
	Controls	0.47	0.87 (0.78 - 0.98)	0.0169	0.92 (0.82 - 1.03)	0.1560
ADHD	SzP	1.07	1.24 (1.08 - 1.42)	1.65×10^{-3}[b]	1.19 (1.04 - 1.36)	0.0129
	Controls	0.22	1.10 (0.98 - 1.23)	0.0991	1.06 (0.94 - 1.19)	0.3230
EA	SzP	1.10	0.80 (0.70 - 0.92)	1.45×10^{-3}[b]	0.85 (0.74 - 0.98)	0.0206
	Controls	0.88	0.82 (0.73 - 0.93)	1.13×10^{-3}[b]	0.87 (0.77 - 0.98)	0.0238
BMI	SzP	0.51	1.16 (1.02 - 1.33)	0.0291	1.11 (0.97 - 1.27)	0.1355
	Controls	0.42	1.14 (1.02 - 1.27)	0.0244	1.10 (0.98 - 1.23)	0.1007
Multitrait	SzP	2.43	1.39 (1.21 - 1.59)	2.41×10^{-6}[b]	-	-
	Controls	2.20	1.35 (1.20 - 1.52)	2.76×10^{-7}[b]	-	-

[a]SmkInit: smoking initiation, AFB: age at first birth, ADHD: attention-deficit hyperactivity disorder, EA: educational attainment and BMI: body mass index.

[b]Significant association at the Bonferroni's correction level ($P < 4.17 \times 10^{-3}$, corrected for 12 tests).

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

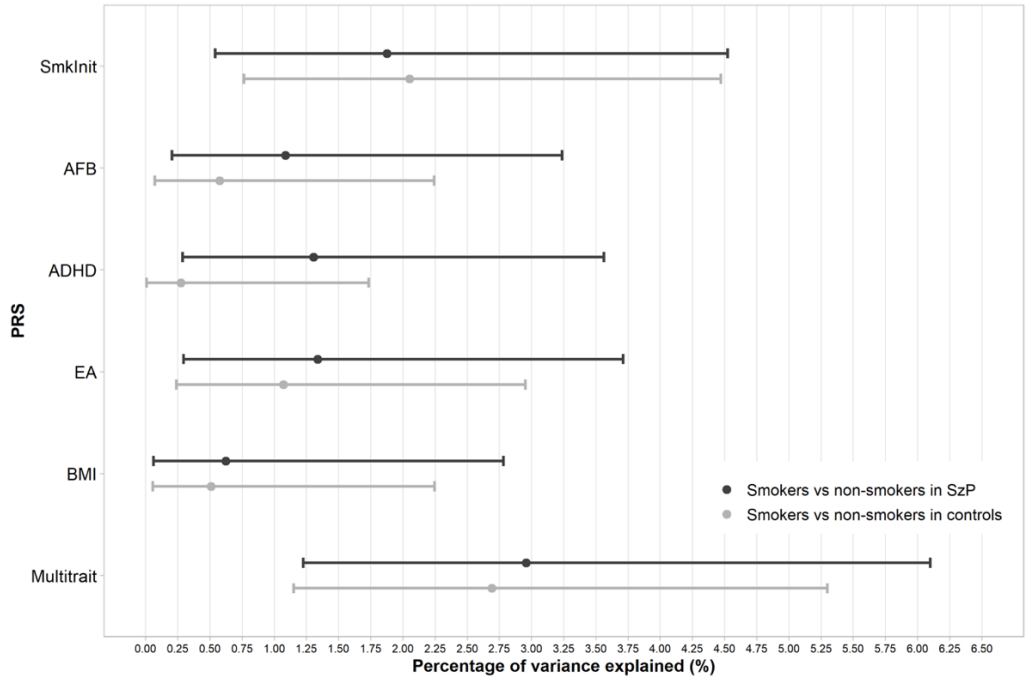


Figure 1. Percentage of variance explained of smoking status (and 95% CI) in schizophrenic patients (SzP) and controls by the different polygenic risk scores (PRSs). SmkInit: smoking initiation, AFB: age at first birth, ADHD: attention-deficit hyperactivity disorder, EA: educational attainment and BMI: body mass index.

3.2. Contribution of PRSs for SmkInit genetically correlated traits to smoking status in SzP and controls

Among those traits genetically correlated with SmkInit, statistical power to analyze the contribution of the PRS to smoking status was higher than 80% in case of AFB and ADHD in both SzP and controls and EA and BMI in controls. Statistical power in SzP was 74% and 69% for EA PRS and BMI PRS, respectively (Table S2). All the PRSs were significant in SzP. AFB PRS, EA PRS and BMI PRS were also significant in controls (Table 2). The different PRSs explained a higher percentage of variance in SzP than in controls (Figure 1 and Table S3). Interaction terms between each PRS and schizophrenic status had no significant effect on smoking status (P = 0.17, P = 0.40, P = 0.41, and P = 0.36, for AFB PRS, ADHD PRS, EA PRS and BMI PRS, respectively), although the tests were substantially underpowered (Figure S4). As expected, they explained less variance in smoking status than SmkInit PRS. All associations were in the expected direction, that is, PRS was negatively associated with smoking in case of AFB and EA, and positively associated in case of the remaining traits (Table 2).

To help interpretation of ADHD PRS results, PRS for a psychiatric cross-disorder meta-analysis was analyzed. Correlation between ADHD PRS and cross-disorder PRS was moderate (r_g = 0.2257). Statistical power to analyze the effect of cross-disorder PRS on smoking status was higher than 80% in both SzP and controls (Table S2). Cross-disorder PRS was not significantly associated with smoking status in SzP or in controls (Table S5). ADHD PRS remained

USC DE COMPOSTELA

significant in SzP and not significant in controls after including cross-disorder PRS in regression model (SzP: OR (95% CI) = 1.23 (1.07 – 1.41), P = 4.13 x 10^{-3}; controls: OR (95% CI) = 1.12 (0.99 – 1.25), P = 0.0633).

3.3. Joint contribution of the different PRSs to smoking status in SzP and controls

The analyzed PRSs were moderately correlated among them, and the pattern of correlation was similar in SzP and controls (Figure 2). Inclusion of SmkInit PRS in the regression models attenuated the effect of the PRSs for SmkInit genetically correlated traits in all cases, although EA PRS was still nominally significant in both SzP and controls, and AFB and ADHD PRSs remained nominally significant in SzP (Table 2). Multitrait prediction, combining the five PRSs using weights estimated by SMTpred, substantially improved the prediction based on SmkInit PRS alone (Table 2 and Table S6). The percentage of variance explained in smoking status by the multitrait PRS was higher in SzP than in controls (Figure 1 and Table S3), although the interaction term between multitrait PRS and schizophrenic status did not reach significance (P = 0.06).

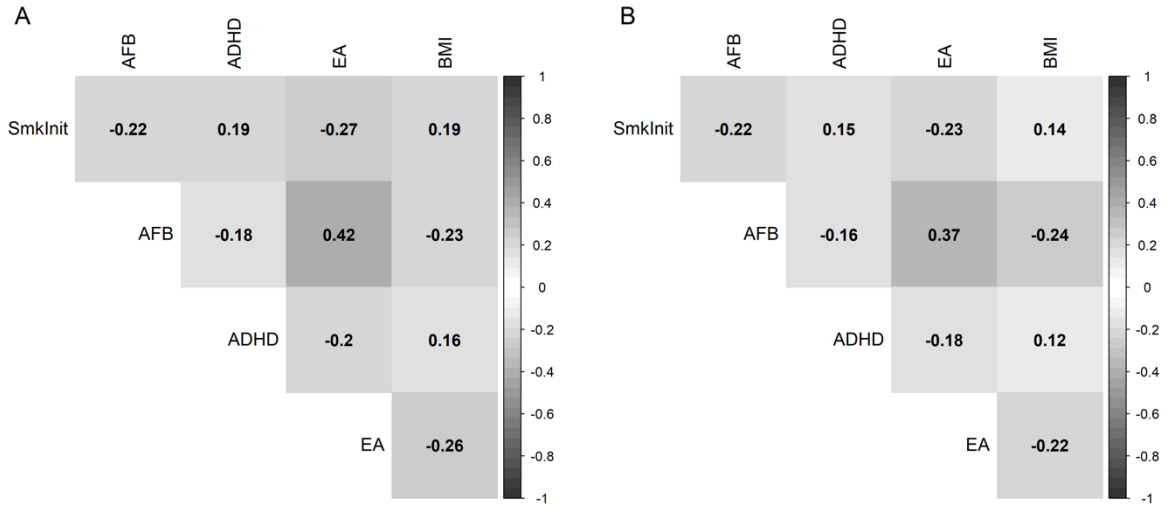


Figure 2. Pearson's correlation between the polygenic risk scores (PRSs) analyzed in this study in schizophrenic patients (SzP) (A) and controls (B). P < 5 x 10^{-3} in all comparisons. SmkInit: smoking initiation, AFB: age at first birth, ADHD: attention-deficit hyperactivity disorder, EA: educational attainment and BMI: body mass index.

3.4. Contribution of SmkInit PRS to schizophrenic status

Statistical power to analyze the contribution of SmkInit PRS to schizophrenic status in the whole sample (SzP and controls) was close to 80%. SmkInit PRS had a nominally significant effect in schizophrenic status in the whole sample (P = 0.0439) but significance was lost after adjustment by smoking status (P = 0.8599). Statistical power to analyze the association between SmkInit PRS and schizophrenic status in the sample stratified by smoking status was low.

USC
DE SANTIAGO
DE COMPOSTELA

SmkInit PRS was not associated with schizophrenic status neither in smokers nor in non-smokers and the percentage of variance explained was similar in both groups (Table S7).

3.5. Contribution of the CHRNA5-CHRNA3-CHRNB4 cluster to schizophrenic status and smoking status

rs16969968 was significantly associated with smoking status, independently of adjustment by schizophrenic status or sample stratification by schizophrenic status. This SNP was not significantly associated with schizophrenic status. rs8042374 was associated with schizophrenic status in the whole sample but significance was lost after sample stratification by smoking status or adjustment by smoking status. This SNP was associated with smoking status in the whole sample, also after adjustment by schizophrenic status, and in SzP (Table 3).

Table 3. Association analysis of SNPs at the *CHRNA5-CHRNA3-CHRNB4* cluster with smoking status and schizophrenic status.

SNP	Trait	Sample[a]	Neff[b]	Odds ratio (95% IC)	P value	Adjusted[c]	
						Odds ratio (95% IC)	P value
rs16969968	Smoking status	Total	2662	1.23 (1.10 - 1.39)	4.85×10^{-4}	1.23 (1.09 - 1.39)	1.13×10^{-3}
		SzP	1017	1.22 (1.00 - 1.48)	0.0462	-	-
		Controls	1357	1.23 (1.04 - 1.44)	0.0152	-	-
	Schizophrenic status	Total	2608	1.10 (0.97 - 1.25)	0.1516	1.03 (0.90 - 1.18)	0.6460
		Smokers	1164	1.00 (0.83 - 1.22)	0.9722	-	-
		Non-smokers	1160	1.06 (0.88 - 1.29)	0.5314	-	-
rs8042374	Smoking status	Total	2662	1.25 (1.09 - 1.44)	1.55×10^{-3}	1.22 (1.05 - 1.41)	8.71×10^{-3}
		SzP	1017	1.28 (1.01 - 1.61)	0.0387	-	-
		Controls	1357	1.17 (0.96 - 1.42)	0.1135	-	-
	Schizophrenic status	Total	2608	1.19 (1.02 - 1.38)	0.0283	1.10 (0.94 - 1.30)	0.2294
		Smokers	1164	1.15 (0.90 - 1.46)	0.2609	-	-
		Non-smokers	1160	1.05 (0.84 - 1.32)	0.6474	-	-

[a]SzP: schizophrenic patients.

[b]Neff: effective sample size, calculated as twice the harmonic mean.

[c]Association analysis for smoking status adjusted by schizophrenic status and association analysis for schizophrenic status adjusted by smoking status.

4. DISCUSSION

Our analysis revealed that SmkInit PRS explained a similar amount of variance in susceptibility to regular smoking in SzP and in controls. Additional PRSs based on traits genetically correlated with SmkInit also contribute to smoking susceptibility, as reflected by the larger percentage of variance explained by a multitrait PRS. As expected, SmkInit PRS was the PRS

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

that explained the highest variability in controls. This PRS also explained the highest variability in SzP. The same PRSs involved in susceptibility to regular smoking in controls, such as EA PRS, AFB PRS, and BMI PRS, affect susceptibility to regular smoking in SzP. The only exception was ADHD PRS which is associated with smoking status only in SzP. Among the PRSs for traits genetically correlated with SmkInit, EA PRS explained the highest percentage of variance, both in SzP and controls. These findings point to the existence of common genetic susceptibility to smoking in SzP and in general population.

There was just one PRS for a psychiatric trait with enough power to be tested, that for ADHD. The association of this PRS with smoking in SzP but not in controls was the main difference between both groups, although the difference in effect was not statistically significant. Both ADHD as well as ADHD symptoms in non-clinical samples are associated with smoking (Charach et al., 2011; Kollins et al., 2005). MR studies suggested a casual effect of ADHD on smoking behavior (Treur et al., 2019; Vilar-Ribó et al., 2020), but not on nicotine dependence, probably due to lack of power (Vink et al., 2020). ADHD PRS is associated with current smoking in the UK Biobank sample, explaining 0.33% (Nagelkerke's R^2) of the variance (Du Rietz et al., 2018). This value is similar to that detected in our controls, 0.22%, and considerably lower than that detected in our schizophrenia sample, 1.07%. The higher variance explained in SzP, pending on confirmation in additional samples, might be related to the considerably higher prevalence of ADHD symptoms in SzP than in general population, reaching about a quarter of patients (Arican et al., 2019). Alternatively, it may be due to pleiotropic genetic effects responsible for the high genetic correlation among psychiatric disorders (P. H. Lee et al., 2019), including substance use disorders (Abdellaoui et al., 2021; Vink et al., 2020). Thus, nicotine dependence was strongly correlated at the genetic level with ADHD (r_g = 0.53), but also with major depressive disorder (r_g = 0.42) and to a lesser extent, with schizophrenia (r_g = 0.18) and bipolar disorder (r_g = 0.13) (Vink et al., 2020). This shared genetic susceptibility may reflect the existence of a general genetic factor of psychopathology (Neumann et al., 2016; Pettersson et al., 2016). To analyze if the association of ADHD PRS with smoking in SzP was due to this general genetic psychopathology factor, we tested the association of cross-disorder PRS with smoking, as well as the effect of adjustment of ADHD PRS by cross-disorder PRS. Cross-disorder PRS was not associated and ADHD PRS remained significant in SzP, suggesting that the association is due to genetic factors specific to ADHD symptomatology. Therefore, future studies should be performed evaluating heterogeneity of ADHD symptoms in SzP, and its involvement in tobacco and other toxic substances use and dependence.

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

A related question is why SmkInit is genetically correlated with schizophrenia. One explanation may be horizontal pleiotropy, considering substance use disorders as different manifestations of the general genetic factor of psychopathology, as suggested previously (Facal et al., 2021; Vink et al., 2020). Alternatively, taking into account that the same genetic factors are involved in smoking behavior in SzP and general population, as revealed by our work, this genetic correlation may be consequence of the higher prevalence of smoking in SzP versus controls in the schizophrenia GWAS. The higher prevalence is compatible with smoking acting as risk factor for schizophrenia, i.e. vertical pleiotropy. As expected under this scenario, we detected an association of the SmkInit PRS with schizophrenia in our sample that was lost when the model was adjusted by smoking status. Similarly, there was no association between SmkInit PRS and schizophrenia when the sample was split in smokers or non-smokers. However, we cannot distinguish between a direct effect of smoking on schizophrenia risk and an indirect effect mediated, for instance, by increasing use of additional substances.

Additionally, we assessed this question focusing on the cluster of neuronal nicotinic cholinergic receptor subunits at chromosome 15. Two SNPs were tested, the missense SNP rs16969968 at *CHRNA5*, strongly associated with several smoking traits (J. Chen et al., 2020; M. Liu et al., 2019), and rs8042374, the top association signal at this locus in the schizophrenia GWAS (Schizophrenia Working Group of the Psychiatric Genomics Consortium et al., 2014). On the one hand, both SNPs were associated with smoking in the full sample. When stratified by schizophrenic status, the effect of rs8042374 was restricted to SzP, perhaps due to a type II error because of low power. On the other hand, association with schizophrenia was present only for rs8042374, and this association was lost after adjustment for smoking status. This is similar to the SmkInit PRS results, strengthen the conclusion that the association between this locus and schizophrenia is due to the higher prevalence of smoking in SzP, probably as a consequence of its role as risk factor.

Recently, the PGC performed a similar analysis on the subset of SzP with smoking data (~5000 patients, no controls) (Peterson et al., 2021). They also concluded that smoking in SzP may be explained, at least partially, by the same genetic factors than in general population, although they did not compare the effect of SmkInit PRS in SzP versus controls. The proportion of variance in SmkInit explained in SzP by the SmkInit PRS was considerably larger in our case (Nagelkerke's R^2 = 1.54% vs 0.58%), probably as a consequence of the larger sample size of the GWAS summary statistics used in our study, GSCAN, in comparison with the Tobacco and Genetics (TAG) Consortium data used by the PGC. In addition to SmkInit, their conclusion was also based on cigarettes-per-day (CPD). The CPD PRS from TAG Consortium predicted

USC DE COMPOSTELA

CPD in SzP. Finally, genetic correlation between SmkInit in SzP from PGC and data from the TAG Consortium did not differ significantly from 1. Our analysis of PRSs based on traits genetically correlated with SmkInit represents independent evidence strengthening the general conclusion that genetic susceptibility to smoking is shared between SzP and general population.

Our study presents several limitations. First of all, characterization of smoking behavior is incomplete. As a cross-sectional study, we did not take into account temporal changes in smoking status. Furthermore, nicotine dependence or intensity of smoking were not considered. In addition, determination of smoking status was not homogeneous across subjects. More controls than cases were classified as smokers/non-smokers using methylation in our study. In order to reduce this bias, we defined a gray zone with ambiguous methylation levels and discarded the samples at this zone and we used the method for smoking status determination as a covariate in our regression models. Furthermore, taking into account that case-control samples did not match by age, we cannot exclude the possibility that a fraction of subjects, different in cases versus controls, did not smoke at recruitment because they were too young for initiation of smoking. According to the Spanish survey on alcohol and other drugs, the average age of initiation of smoking among 15-64 years old people remained stable at ~16.5 from 1995 to 2013. Regular smoking begins two years later (Observatorio Español de las Drogas y las Adicciones, 2015). Therefore, we removed subjects under 20 years of age to reduce this putative bias. Another limitation was the low power to test for interaction between PRS and disease status and to test PRSs for other traits genetically correlated with SmkInit, including that for schizophrenia, which might inform about severity of symptoms. Finally, we did not get access to additional samples for replication.

In summary, genetic susceptibility to smoking in SzP seems to be due to the same factors than in general population, although we have detected a larger effect of ADHD PRS in SzP, perhaps related to higher ADHD symptomatology in SzP. Therefore, the higher prevalence of smoking in SzP than in general population is compatible with smoking, directly or indirectly, acting as a risk factor for schizophrenia. Hence, smoking may be considered an important modifiable factor in selective and indicated prevention targeting vulnerable individuals.

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

4.5. ESTUDO 5: GENETIC SUSCEPTIBILITY FOR SCHIZOPHRENIA AFTER ADJUSTMENT BY GENETIC SUSCEPTIBILITY FOR SMOKING: IMPLICATIONS IN IDENTIFICATION OF RISK GENES AND GENETIC CORRELATION WITH RELATED TRAITS

Referencia:

Al-Soufi, L. e Costas, J. (2023). Genetic susceptibility for schizophrenia after adjustment by genetic susceptibility for smoking: implications in identification of risk genes and genetic correlation with related traits. *Psychological Medicine, 53*(14), 6806–6816. https://doi.org/10.1017/S0033291723000326

Información dos/as autores/as:

Laila Al-Soufi[1,2] e Javier Costas[1,3]

[1]Grupo de Xenética Psiquiátrica, Instituto de Investigación Sanitaria de Santiago de Compostela (IDIS), Santiago de Compostela, Galicia, España. Rede de Investigación en Atención Primaria de Adiccións (RIAPAd).

[2]Departamento de Zooloxía, Xenética e Antropoloxía Física, Universidade de Santiago de Compostela (USC), Galicia, España.

[3]Complexo Hospitalario Universitario de Santiago de Compostela (CHUS), Servizo Galego de Saúde (SERGAS), Santiago de Compostela, Galicia, España.

Información do artigo:

DOI: 10.1017/S0033291723000326

PMID: 36876478

Ligazón: https://doi.org/10.1017/s0033291723000326

Versión orixinal publicada no Anexo II.

Información da revista:

Psychological Medicine

ISSN: 1469-8978

Índices de calidade segundo o JCR en 2022:

- Factor de impacto: 6,9
- Factor de impacto nos últimos cinco anos: 7,6
- Psychology: Q1 (9/81)
- Psychiatry: Q1 (27/155)

- Psychology, clinical (SSCI): D1 (10/131)
- Psychiatry (SSCI): Q1 (21/144)

Contribución da doutoranda na publicación:

Realización de todas as análises. Redacción de gran parte do manuscrito inicial e revisión da versión final.

Autorización da revista:

A revista *Psychological Medicine*, pertencente á editorial *Cambridge University Press*, permite reutilizar o artigo nesta tese mais non a súa distribución ao público xeral (Anexo III).

Material suplementario do artigo:

Os materiais suplementarios deste artigo poden atoparse nas seguintes ligazóns:

https://static.cambridge.org/content/id/urn:cambridge.org:id:article:S0033291723000326/resource/name/S0033291723000326sup001.xlsx

https://static.cambridge.org/content/id/urn:cambridge.org:id:article:S0033291723000326/resource/name/S0033291723000326sup002.docx

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

5. DISCUSIÓN XERAL

Nesta tese, a través de cinco estudos nos que se empregaron distintas metodoloxías xenómicas, estudáronse os factores xenéticos que inflúen nos trazos de consumo de substancias en poboación xeral e, máis especificamente, no consumo de tabaco en pacientes con esquizofrenia. Por unha banda, estudouse a susceptibilidade xenética aos trazos de consumo de substancias e outros trazos relacionados, combinando PRS con datos de expresión xénica. Ademais, examinouse o efecto dos problemas externalizantes e internalizantes no consumo de substancias, analizando o solapamento xenético de trazos externalizantes e internalizantes con diferentes trazos de consumo de substancias. Por outra banda, caracterizouse a susceptibilidade xenética compartida entre tabaquismo e esquizofrenia, o cal permitiu afondar máis nas relacións de causalidade entre ambos trazos.

No Estudo 1, aplicouse a metodoloxía WGCNA a datos de expresión xénica do cerebro do peixe cebra adulto baixo diferentes condicións experimentais, obtendo 34 módulos de xenes co-expresados no cerebro do peixe cebra. Algúns destes módulos de xenes estiveron enriquecidos en xenes implicados en sinapse e xenes diferencialmente expresados en tecidos do cerebro e estadios do desenvolvemento do cerebro, suxerindo que estes módulos de co-expresión identificados en peixe cebra poderían modelar as redes xénicas do cerebro humano. Ademais, identificáronse módulos de xenes co-expresados que poderían estar involucrados en trazos relacionados coa psiquiatría, analizando o seu enriquecemento en GWAS e a asociación dos seus xenes centrais. Finalmente, avaliouse a utilidade destes módulos de co-expresión para modelar a susceptibilidade xenética a trazos de consumo de substancias, en particular, a dependencia a substancias. Dentro dos distintos módulos de co-expresión, identificáronse cinco cuxo patrón de expresión xénica estivo asociado á abstinencia a substancias, e non a outras condicións experimentais, no peixe cebra. Comprobouse que unha PRS baseada en variantes xenéticas dentro ou próximas aos xenes destes módulos de co-expresión especificamente asociados á abstinencia a substancias explica significativamente máis a dependencia a alcol en humanos que unha PRS baseada en variantes xenéticas dentro ou próximas a outros xenes expresados en cerebro.

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

No Estudo 2, analizouse a evidencia dun efecto causal dos problemas externalizantes e internalizantes sobre os trazos de consumo de substancias, o cal se reflectiría nun solapamento xenético entre ambos trazos e unha distribución similar do efecto das variantes xenéticas causais

compartidas para ambos trazos. Ademais, examinouse se esta relación xenética é diferente para comportamentos de consumo de substancias e SUD. Con este fin, empregouse MiXeR para caracterizar o patrón de solapamento xenético de catro trazos de consumo de substancias con cinco trazos externalizantes e cinco trazos internalizantes. Os comportamentos de consumo de substancias presentaron un alto solapamento xenético con trazos internalizantes, mais cunha correlación xenética do efecto das variantes compartidas particularmente baixa, indicando unha distribución diverxente do efecto da susceptibilidade xenética compartida para ambos trazos. Pola contra, estes trazos presentaron unha maior correlación xenética das variantes compartidas con trazos externalizantes, tamén no caso dun solapamento xenético considerable. Estes resultados indican que a vía externalizante parece ter unha maior influencia en comportamentos de consumo de substancias que a vía internalizante. Por outro lado, os SUD presentaron, mesmo no caso dun solapamento xenético alto, unha correlación xenética das variantes compartidas elevada tanto con trazos externalizantes como con trazos internalizantes, aínda que máis acentuada cos primeiros. Esta observación suxire que ambas vías inflúen no desenvolvemento de SUD.

Nos seguintes estudos investigouse, en concreto, a asociación existente entre tabaquismo e esquizofrenia, analizando e caracterizando a susceptibilidade xenética compartida entre ambos trazos. No Estudo 3 da tese, analizouse esta susceptibilidade xenética compartida entre tabaquismo e esquizofrenia a nivel dos xenes de subunidades de nAchR para atopar evidencias dun efecto causal do tabaquismo sobre a esquizofrenia. Se esta hipótese é certa, os sinais de asociación nestes xenes de GWAS para trazos de consumo de tabaco e para esquizofrenia deberían ser debidos ás mesmas variantes causais, é dicir, deberían colocalizar. Empregando a metodoloxía *coloc*, obtívose a probabilidade de colocalización entre esquizofrenia e unha serie de trazos de consumo de tabaco en tres loci de nAchR para os que existe evidencia de asociación con ambos trazos. Para o xene *CHRNA2*, obtívose forte evidencia de colocalización entre esquizofrenia e o consumo de tabaco, así como con CUD. Porén, para os outros dous loci, o xene *CHRNA4* e a agrupación de xenes *CHRNA5-CHRNA3-CHRNB4*, *coloc* apoiou a existencia de sinais de asociación independentes.

No Estudo 4, analizouse a asociación entre tabaquismo e esquizofrenia a través de PRS. Por un lado, comprobouse se os/as pacientes con esquizofrenia comezan a fumar polas mesmas razóns que a poboación xeral. PRS estimadas en poboación xeral para o inicio do consumo de tabaco, así como para algúns trazos xeneticamente correlacionados, explicaron unha proporción da varianza do consumo de tabaco similar en persoas con esquizofrenia e en controis. Porén, PRS para outros trazos correlacionados co inicio do consumo de tabaco, en concreto, para

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

ADHD e para a idade á que se tivo o/a primeiro/a fillo/a, soamente estiveron significativamente asociadas a tabaquismo, tras corrixir por múltiples tests, en persoas con esquizofrenia. Esta asociación permaneceu significativa despois de incluír a PRS para o inicio do consumo de tabaco. Ademais, a asociación da PRS para ADHD non cambiou ao introducir a PRS para unha meta-análise de trastornos psiquiátricos. Por outro lado, neste mesmo estudo, analizouse a posible explicación á susceptibilidade xenética compartida entre tabaquismo e esquizofrenia. Tal e como se espera por un efecto causal do tabaco sobre o desenvolvemento da esquizofrenia, atopouse que a asociación da PRS para o inicio do consumo de tabaco con esquizofrenia deixa de ser significativa tras axustar ou estratificar polo hábito tabáquico da mostra. Do mesmo xeito, a asociación da variante xenética máis significativa para esquizofrenia da agrupación de xenes de nAchR *CHRNA5-CHRNA3-CHRNB4* perdeuse ao axustar ou estratificar polo hábito tabáquico da mostra.

No Estudo 5 da tese, condicionouse o último GWAS de esquizofrenia do PGC polo efecto causal do inicio do consumo de tabaco sobre esquizofrenia, eliminando, deste xeito, dentro da susceptibilidade xenética a esquizofrenia, aquela mediada pola vía causal do tabaquismo. Deste xeito, identificáronse 19 novos loci de risco para esquizofrenia cuxa asociación podería ter quedado oculta polo seu efecto sobre o inicio do consumo de tabaco na outra dirección e 42 loci que deixaron de estar significativamente asociados a esquizofrenia xa que a súa asociación con esquizofrenia no GWAS orixinal podería estar parcialmente mediada pola vía causal do tabaquismo. A colocalización do sinal de asociación nalgúns destes loci para esquizofrenia e o inicio do tabaquismo apoiou que fosen loci parcialmente asociados a esquizofrenia a través da mesma vía que o tabaquismo. Ademais, tras axustar o GWAS de esquizofrenia polo efecto causal do inicio do consumo de tabaco, a súa correlación xenética con fenotipos externalizantes, como comportamentos de consumo de substancias, SUD, ADHD e trazos condutuais relacionados coa asunción de riscos, cambiou de forma significativa. En concreto, a correlación xenética da esquizofrenia cunha conduta externalizante diminuíu tras eliminar aquela susceptibilidade xenética mediada pola vía causal do tabaquismo. Finalmente, dentro dos loci cuxa asociación con esquizofrenia deixou de ser significativa tras condicionar por tabaquismo, a análise de colocalización identificou 14 xenes cun sinal de asociación común a esquizofrenia e a trazos externalizantes. Ademais, tal e como se agardaba, estes loci, a excepción do xene *CHRNA2*, presentaron un efecto na mesma dirección para esquizofrenia e conduta externalizante.

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

5.1. SUSCEPTIBILIDADE AOS TRAZOS DE CONSUMO DE SUBSTANCIAS

5.1.1. Modelización da susceptibilidade xenética aos trazos de consumo de substancias

A susceptibilidade xenética identificada en GWAS para trazos complexos, como os trazos de consumo de substancias, é xeralmente recollida en forma de PRS. Porén, as PRS normalmente non se basean nunha hipótese biolóxica á hora de modelar o risco xenético a un trazo complexo, senón que representan de forma xeral o número de alelos de risco herdados por un individuo (Restrepo-Lozano et al., 2023). Para obter máis información sobre os procesos e funcións biolóxicas que están involucrados no trazo avaliado e aumentar o potencial das PRS como biomarcadores da susceptibilidade a un trazo no futuro, é conveniente que as PRS se constrúan en base a información funcional obtida previamente.

Os GWAS mostraron que a maior parte do risco aos trazos polixénicos é explicado por variantes que regulan a expresión xénica (Jaffe et al., 2018). Polo tanto, a expresión xénica compórtase como un fenotipo molecular intermedio que une a susceptibilidade xenética co risco á enfermidade (Restrepo-Lozano et al., 2023). Empregando datos de expresión xénica, pódense identificar aqueles xenes cunha expresión alterada nun determinado trazo que poderían estar implicados na susceptibilidade ao mesmo (Gerring et al., 2022). Ademais, pódense diferenciar cales variantes xenéticas identificadas en GWAS están involucradas na regulación da expresión deses xenes (Restrepo-Lozano et al., 2023). Con esta información, pódense estimar PRS máis específicas que recollan o risco a unha expresión alterada duns determinados xenes posiblemente implicados no trazo.

Ademais, é sabido que os xenes non funcionan de forma illada, senón que forman parte de redes xénicas complexas cunha gran especificidade espazo-temporal (Li et al., 2018). Os xenes cunha expresión correlacionada en distintas condicións posiblemente estean regulados polo mesmo sistema de regulación da expresión e, polo tanto, seguramente estean relacionados funcionalmente (H. K. Lee et al., 2004). Isto foi atopado no Estudo 1, onde o enriquecemento dos módulos de co-expresión en diferentes procesos biolóxicos, funcións moleculares e compoñentes celulares confirmou que recollen xenes relacionados a nivel funcional. Por conseguinte, os módulos de co-expresión cun patrón de expresión alterado nun determinado trazo poden estar reflectindo aqueles procesos moleculares implicados no trazo (Gaiteri et al., 2014). Deste xeito, a información de redes xénicas de co-expresión pode ser aplicada a PRS que recollan procesos moleculares específicos desregulados nun trazo (Restrepo-Lozano et al., 2023). Este enfoque foi previamente empregado para estimar unha PRS asociada á rede de co-expresión do xene *DRD2*, o cal codifica para o receptor de dopamina D2 e está fortemente

asociado a fenotipos relacionados coa esquizofrenia (Pergola et al., 2017). Neste estudo, confirmouse que esta rede de co-expresión do xene *DRD2* estaba enriquecida en xenes de risco para esquizofrenia. Ademais, unha PRS formada polos SNP que, en forma de eQTL, afectan á expresión dos xenes dentro desta rede estivo asociada a un procesamento ineficiente da memoria de traballo, un trazo relacionado coa esquizofrenia. Do mesmo xeito, outro estudo mostrou que unha PRS baseada nos eQTL dos xenes da rede de co-expresión do xene *DCC* nas rexións corticolímbicas estaba asociada a impulsividade durante a infancia, explicando significativamente máis que unha PRS do mesmo tamaño derivada do GWAS de ADHD (Restrepo-Lozano et al., 2022).

No Estudo 1, seguiuse este enfoque co fin de construír unha PRS que recollese o efecto de variantes xenéticas cun efecto a nivel funcional naqueles conxuntos de xenes funcionalmente relacionados e especificamente alterados na dependencia a substancias. En concreto, estimouse esta PRS a partir das variantes xenéticas localizadas dentro ou próximas aos xenes de módulos de co-expresión cun patrón de expresión alterado no cerebro de peixe cebra baixo a abstinencia a substancias. Comprobouse que esta PRS contribúe significativamente máis á susceptibilidade a dependencia a alcol en humanos que unha PRS construída a partir do mesmo número de SNP, escollidos aleatoriamente dentro daqueles próximos a outros xenes expresados no cerebro. Deste xeito, puidéronse atopar módulos de co-expresión que parecen estar implicados no risco a dependencia a alcol, ademais de identificar unha PRS que parece recoller o risco a unha alteración funcional deses xenes.

Nesta tese, a PRS foi estimada a partir daquelas variantes xenéticas localizadas dentro dos xenes dos módulos de co-expresión de interese ou próximas a eles, xa que é probable que estas variantes teñan un efecto a nivel funcional neses xenes. Alternativamente, para estimar PRS que midan a susceptibilidade xenética a ter unha expresión alterada duns determinados xenes, poderíanse empregar aqueles SNP para os que existise evidencia de actuar como un *cis*-eQTL para os xenes de interese no tecido avaliado, é dicir, que afecten á expresión deses xenes próximos. Esta información está dispoñible en recursos como GTEx, o cal nos permite identificar os principais eQTL dos xenes en distintos tecidos (GTEx Consortium, 2020). Non obstante, existe evidencia de que unha importante proporción da varianza dos trazos complexos é explicada polo efecto de *trans*-eQTL, é dicir, polo efecto acumulativo de eQTL sobre xenes afastados no xenoma, a pesar de que o seu efecto individual é moito menor que o dos *cis*-eQTL (X. Liu et al., 2019). Ademais, estes *trans*-eQTL parecen ser os principais encargados de regular conxuntamente a expresión dunha rede xénica. Polo tanto, sería importante empregar ferramentas que permitan identificar *trans*-eQTL dos módulos de co-expresión relevantes para

USC UNIVERSIDADE DE SANTIAGO

o trazo (Wang et al., 2024). Outra alternativa á metodoloxía empregada nesta tese sería o uso de ferramentas como PrediXcan ou FUSION (Gamazon et al., 2015; Gusev et al., 2016). Empregando como referencia bases de datos como GTEx onde se mediu tanto a expresión xénica como a variación xenética a nivel de SNP, estas metodoloxías permiten predicir a expresión xénica debida ao efecto agregado de *cis*-eQTL nunha mostra nun tecido determinado a partir dos datos xenotípicos dos individuos da mostra. A continuación, esta información pode ser empregada para avaliar se existe asociación entre os niveis de expresión xénica e un fenotipo de interese nesa mesma mostra. Neste caso, empregando esta metodoloxía, poderíase testar se a expresión xénica predita dos individuos da mostra diana nos xenes de interese está asociada coa dependencia a alcol. Porén, a diferenza do método empregado nesta tese, estas estratexias alternativas dependen da existencia dos datos de expresión xénica necesarios para o tipo celular e a etapa do desenvolvemento relevante no fenotipo estudado.

Os modelos animais facilitan a obtención de datos de expresión xénica en tipos celulares e etapas do desenvolvemento específicos. No Estudo 1, comprobamos que os módulos de co-expresión identificados a partir de datos de expresión do cerebro do peixe cebra adulto teñen potencial para modelar as redes xénicas do cerebro humano. Ademais, atopamos certa evidencia de que algúns destes módulos poderían estar involucrados tanto no consumo de substancias como noutros trazos relacionados coa psiquiatría. Esta alta pleiotropía a nivel de expresión xénica entre trastornos psiquiátricos e trazos de consumo de substancias foi xa demostrada anteriormente (Gerring et al., 2022).

Debido á súa natureza polixénica con múltiples xenes interactuando para determinar o risco ao trazo, resulta complicado xerar modelos animais para trazos de consumo de substancias e outros trazos complexos. Polo tanto, é esencial identificar xenes representativos das redes xénicas implicadas nestes trazos. Nun estudo previo atopouse que, ao reducir a expresión do xene *CHD8*, fortemente asociado ao trastorno do espectro autista, se produce un cambio na expresión de múltiples xenes relacionados con este trastorno, indicando que este xene participa nunha rede xénica na que converxen moitos xenes asociados co trastorno do espectro autista (Sugathan et al., 2014). Unha vez constatado isto, viuse que a supresión deste xene en peixe cebra daba lugar a unha macrocefalia comparable á atopada nas persoas con trastorno do espectro autista que portan unha mutación inactivadora dese xene. O enfoque seguido nesta tese vai un paso máis alá, conseguindo reducir o elevado número de variantes xenéticas asociadas a un trazo en GWAS a unha serie de xenes representativos dos módulos de co-expresión posiblemente involucrados nese trazo. Deste xeito, este enfoque podería facilitar a xeración de modelos animais para trazos complexos en futuros traballos.

USC UNIVERSIDADE DE SANTIAGO

5.1.2. Efecto da vía externalizante e a vía internalizante nos comportamentos de consumo de substancias e nos trastornos polo uso de substancias

O estudo da susceptibilidade aos trazos de consumo de substancias baséase principalmente na evidencia de que problemas externalizantes e internalizantes aumentan a susceptibilidade a un posterior consumo de substancias (Green et al., 2018; Meque et al., 2019). Porén, dado que os comportamentos de consumo de substancias e os SUD presentan diferenzas na súa susceptibilidade xenética, sería posible que estas dúas vías tivesen un efecto diferente en cada un destes trazos.

Os resultados do Estudo 2 apoian unha diferente implicación da vía externalizante e a vía internalizante ao longo do proceso de consumo de substancias. Por un lado, a vía internalizante parece ter principalmente un efecto nos niveis máis severos do consumo de substancias. Esta conclusión concorda co ciclo da adicción, segundo o cal, un consumo compulsivo de substancias para evitar o afecto negativo propio do síndrome de abstinencia é clave na dependencia a substancias (Koob et al., 2014). Ademais, é consistente cos resultados de Ning et al. (2020), segundo os cales, os problemas internalizantes predín un maior nivel de consumo severo de alcol, mais un menor consumo leve de alcol. Polo tanto, no consumo de substancias non patolóxico, parece ter un efecto principalmente a vía externalizante. Isto concorda con estudos de MR que apoian que ADHD actúa como un factor de risco sobre comportamentos de consumo de substancias como o inicio do consumo de tabaco, a idade de inicio do consumo de tabaco, a intensidade do consumo de tabaco, o cesamento do tabaquismo ou o consumo de cannabis (M. E. Fluharty et al., 2018; Jang et al., 2022; Soler Artigas et al., 2020; Treur et al., 2019; Vilar-Ribó et al., 2020).

Por outro lado, segundo os resultados deste estudo, tanto a vía internalizante como a vía externalizante parecen presentar un efecto nos SUD, cunha implicación especialmente acentuada da vía externalizante. Este resultado, de novo, foi o atopado na revisión sistemática de estudos lonxitudinais de Ning et al. (2020). Neste estudo, atopouse maior evidencia dunha asociación positiva do consumo de alcol máis severo con problemas externalizantes que con problemas internalizantes. Do mesmo xeito, un estudo lonxitudinal mostrou que unha vía externalizante caracterizada por comportamentos externalizantes na adolescencia predí un posterior consumo problemático de alcol na idade adulta temperá (Edwards et al., 2016). Porén, dentro dun amplo rango de síntomas internalizantes a diferentes idades, só os síntomas depresivos aos 16 anos predixeron un consumo problemático de alcol aos 20 anos.

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

Tendo en conta esta hipótese, os problemas externalizantes, a diferenza dos problemas internalizantes, poderían estar parcialmente asociados a un consumo problemático de substancias a través do seu efecto nas primeiras fases do consumo de substancias. Neste sentido, atopouse que a asociación do factor xenético de risco a adiccións construído con genomic SEM cun GWAS de toma de riscos diminuíu considerablemente tras ter en conta, como covariable, a susceptibilidade xenética a un consumo de substancias en poboación xeral (Hatoum et al., 2022). Porén, o consumo de substancias non tivo gran efecto na asociación entre este factor de risco a adiccións e o neuroticismo.

A nosa conclusión non é consistente con propostas previas, nas cales se presentaba a vía internalizante como a principal vía de risco aos SUD. Por exemplo, unha revisión suxeriu que, mentres que a impulsividade característica da vía externalizante tiña un claro efecto no inicio do consumo de tabaco, o mantemento do tabaquismo e a recaída ao mesmo debíase principalmente a unha impulsividade por fumar cando se experimenta un afecto negativo (Bloom et al., 2014). Unha meta-análise baseada en datos epidemiolóxicos de 79.757 individuos foi consistente con esta hipótese, atopando que maiores niveis de neuroticismo predín un menor cesamento do tabaquismo e menores niveis de conciencia, un trazo da personalidade caracterizado por un maior autocontrol e, polo tanto, inversamente asociado coa impulsividade, predín un maior inicio do consumo de tabaco (Hakulinen et al., 2015). Ademais, estudos de MR previos non detectaron claramente un efecto causal de ADHD sobre a dependencia a substancias como alcol, nicotina ou cocaína (Treur et al., 2019; Vilar-Ribó et al., 2020; Vink et al., 2020). Un recente estudo de MR bidireccional atopou principalmente evidencia dun efecto causal do uso problemático de alcol e tabaco sobre ADHD, con só un pequeno efecto causal no sentido inverso (Koller et al., 2024). Porén, si que atopou un efecto bidireccional entre ADHD e CUD, cunha magnitude similar en ambas direccións. Estas discrepancias entre os distintos estudos en canto ao efecto destas dúas vías de risco poderían deberse a diferenzas segundo o momento no que se avalíen os síntomas externalizantes e internalizantes e o consumo de substancias (Edwards et al., 2016; Ning et al., 2020).

5.2. ANÁLISE DO TABAQUISMO EN PACIENTES CON ESQUIZOFRENIA

Tras unha análise da susceptibilidade ao consumo de substancias na poboación xeral, avaliouse a causa da elevada prevalencia do consumo de tabaco nas persoas con esquizofrenia. Empregando unha batería de metodoloxías xenómicas, examinouse a evidencia para cada unha das hipóteses propostas para explicar esta asociación. Por un lado, analizouse se os factores xenéticos de risco a tabaquismo nas persoas con esquizofrenia son equivalentes aos da

USC DE SANTIAGO DE COMPOSTELA

poboación xeral. Por outro lado, examinouse a causa da susceptibilidade xenética compartida entre tabaquismo e esquizofrenia.

5.2.1. Susceptibilidade xenética a tabaquismo nas persoas con esquizofrenia

Segundo a hipótese da automedicación, a hipótese tradicionalmente proposta para explicar a asociación entre o tabaquismo e a esquizofrenia, os factores xenéticos que levan ás persoas con esquizofrenia a consumir tabaco son diferentes dos da poboación xeral, xa que son particularidades do trastorno as que aumentan o seu risco a tabaquismo (Kumari & Postma, 2005). Os resultados obtidos no Estudo 4 da tese non apoian esta hipótese, suxerindo que os factores xenéticos que fan que as persoas con esquizofrenia fumen son os mesmos que os da poboación xeral. Estes resultados son consistentes cun estudo previo, o cal tamén atopou que PRS para tabaquismo estimadas na poboación xeral están significativamente asociadas a tabaquismo nos máis de 5.000 casos de esquizofrenia do PGC con datos de hábito tabáquico (Peterson et al., 2021). O Estudo 4 desta tese dá un paso máis ao atopar unha asociación da PRS para o inicio do consumo de tabaco co hábito tabáquico tanto en pacientes con esquizofrenia como en controis, sendo a varianza explicada pola PRS similar en ambos grupos.

Se as persoas con esquizofrenia fuman polas mesmas razóns que a poboación xeral, é de agardar que a vía externalizante e/ou a vía internalizante teñan un efecto no seu consumo de tabaco. No Estudo 5 da tese, ao eliminar do GWAS de esquizofrenia a susceptibilidade xenética mediada pola vía causal do inicio do consumo de tabaco, a correlación xenética da esquizofrenia con trazos externalizantes diminuíu. En contraposición, a súa correlación xenética con trazos internalizantes mantívose. Este resultado suxire que os factores xenéticos que incrementan o risco ao inicio do consumo de tabaco nas persoas con esquizofrenia están capturando susceptibilidade xenética a conduta externalizante. Isto concorda co Estudo 2 da tese, segundo o cal, a vía externalizante presenta un maior efecto no inicio do consumo de substancias que a vía internalizante. Polo tanto, ao igual que na poboación xeral, a conduta externalizante parece aumentar o risco dos/as pacientes con esquizofrenia a comezar a consumir substancias, como xa mostraron estudos epidemiolóxicos anteriores (Dervaux et al., 2001; Gut-Fayand et al., 2001).

Porén, o Estudo 4 da tese, ao examinar a asociación de PRS para trazos correlacionados co inicio do consumo de tabaco co hábito tabáquico en persoas con esquizofrenia, atopou algunhas diferencias na susceptibilidade xenética a tabaquismo nas persoas con esquizofrenia en comparación coa poboación xeral. As PRS para ADHD, un trastorno psiquiátrico externalizante, e para a idade á que se tivo o/a primeiro/a fillo/a, un trazo condutual asociado a

USC UNIVERSIDADE DE SANTIAGO

unha maior asunción de riscos (Therriault et al., 2024), explicaron o hábito tabáquico das persoas con esquizofrenia pero non o dos controis. Polo tanto, estas PRS para trazos externalizantes parecen estar recollendo unha susceptibilidade xenética, non totalmente capturada pola propia PRS para o inicio do consumo de tabaco nin pola susceptibilidade xenética a psicopatoloxía xeral, medida a través dunha PRS para unha meta-análise de trastornos psiquiátricos, que explica mellor o consumo de tabaco nas persoas con esquizofrenia que na poboación xeral.

Estes resultados indican que a susceptibilidade xenética a problemas externalizantes parece explicar en maior medida a susceptibilidade a tabaquismo en pacientes con esquizofrenia que en poboación xeral. Unha posible explicación a isto poderían ser os maiores niveis de impulsividade e de sintomatoloxía relacionada con ADHD en pacientes con esquizofrenia en comparación con controis (Arican et al., 2019; Heerey et al., 2007; Nolan et al., 2011; Ouzir, 2013). Non obstante, outros estudos suxiren que esta asociación entre impulsividade e esquizofrenia depende da dimensión da impulsividade analizada (Fischer et al., 2015; Reddy et al., 2014). Nestes estudos, os/as pacientes con esquizofrenia, en comparación cos controis, parecen presentar unha maior impulsividade cognitiva e unha menor percepción de dominio da súa vida. Porén, non mostran diferenzas noutros aspectos como o nesgo optimista á hora de avaliar os riscos ou a busca de novas sensacións, e mesmo presentan unha menor conduta de asunción de riscos que os controis.

Se os/as pacientes con esquizofrenia efectivamente presentan unha maior conduta externalizante que a poboación xeral, poderían ter un maior risco a consumir tabaco por esta vía de risco. Estes maiores niveis de conduta externalizante existentes en persoas con esquizofrenia que en controis poderían ser un síntoma da propia esquizofrenia. Porén, nun estudo previo, a PRS para esquizofrenia estivo asociada á toma de riscos en poboación xeral, descartando que esta conduta externalizante en persoas con esquizofrenia sexa unha consecuencia do trastorno (Socrates et al., 2021). Esta conduta externalizante en persoas con esquizofrenia tamén podería explicarse polo efecto dun factor xenético externalizante na etioloxía a esquizofrenia aumentando, tal e como discuten Socrates et al. (2021), o risco a migrar ou a vivir nun ambiente urbano, os cales son factores de risco a psicose amplamente recoñecidos (Stilo & Murray, 2019).

5.2.2. Susceptibilidade xenética compartida entre tabaquismo e esquizofrenia

Unha vez confirmado que, ao igual que a poboación xeral, as persoas con esquizofrenia comezan a fumar principalmente a través da vía externalizante, púidose desbotar a hipótese de

que a elevada comorbididade existente entre tabaquismo e esquizofrenia se deba a unha automedicación con tabaco. Tal e como se comentou previamente, esta alta comorbididade reflíctese nunha susceptibilidade xenética compartida entre ambos trazos (W. Song et al., 2022). Este solapamento xenético podería explicarse tanto por unha pleiotropía vertical, onde o tabaco provoca alteracións neuropatolóxicas que incrementan o risco a esquizofrenia, como por unha pleiotropía horizontal, onde os mesmos factores xenéticos causan alteracións que incrementan o risco a ambos trazos de xeito independente (Solovieff et al., 2013). Nesta tese, analizouse a orixe desta susceptibilidade xenética compartida dende diferentes perspectivas co fin de atopar evidencias para cada unha destas dúas hipóteses.

O Estudo 3 da tese mostrou que o mesmo sinal de asociación do xene *CHRNA2* é responsable da súa asociación con tabaquismo e esquizofrenia, desbotando a existencia de pleiotropía espuria neste xene, é dicir, unha asociación deste xene con tabaquismo e esquizofrenia a través de variantes causais diferentes. Este resultado é consistente cunha pleiotropía vertical, actuando o tabaco como un factor de risco para desenvolver esquizofrenia (Figura 5). Non obstante, como este sinal de asociación do xene tamén colocaliza con CUD, podería ser o cannabis o que teña un efecto causal sobre a esquizofrenia. Este resultado tamén sería compatible cunha pleiotropía vertical inversa, coa esquizofrenia promovendo o consumo de tabaco. Porén, tendo en conta que a prevalencia da esquizofrenia na poboación é moito menor que a do tabaquismo, este locus non podería estar asociado a tabaquismo a través da esquizofrenia.

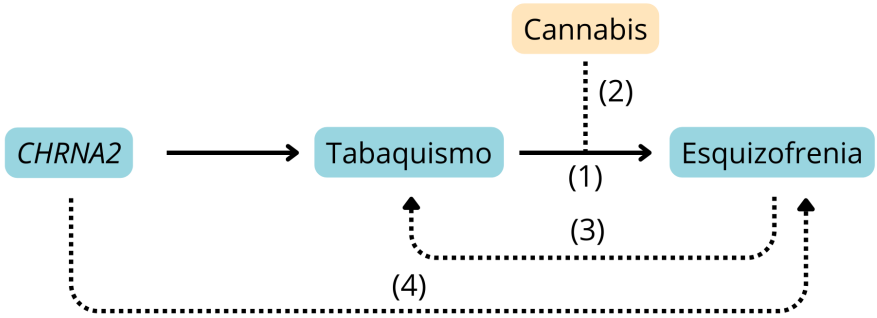


Figura 5. Posible explicación á asociación do xene *CHRNA2* con tabaquismo e esquizofrenia segundo a análise de colocalización. (1): Como este xene está asociado a tabaquismo e esquizofrenia a través do mesmo sinal de asociación, a súa asociación con esquizofrenia podería estar mediada por tabaquismo actuando como un factor de risco a esquizofrenia. (2): Como o mesmo sinal de asociación tamén colocaliza con CUD, podería ser que o tabaquismo teña un efecto causal indirecto sobre a esquizofrenia a través do consumo de cannabis. (3): A presenza de colocalización tamén é compatible coa esquizofrenia promovendo o consumo de tabaco. (4): A presenza de colocalización non asegura unha relación causal entre ambos trazos, a mesma variante causal podería estar asociada a tabaquismo e a esquizofrenia por vías independentes.

Porén, que a mesma variante causal explique a asociación do xene *CHRNA2* con tabaquismo e esquizofrenia non implica necesariamente que estea asociada a ambos trazos pola mesma vía, senón que esta variante tamén podería estar asociada a ambos trazos de xeito independente (Figura 5). Non obstante, o Estudo 4 da tese obtivo evidencias de que a susceptibilidade xenética asociada a tabaquismo está tamén asociada a esquizofrenia debido a un efecto causal do tabaco sobre a esquizofrenia e non debido a unha asociación independente con ambos trazos. Esta diferenciación entre ambas hipóteses foi posible grazas á estratificación da mostra de pacientes con esquizofrenia segundo o seu hábito tabáquico, o cal xa fora proposto anteriormente (Alderson & Lawrie, 2015; Gage, Davey Smith, et al., 2016; Gage & Munafò, 2015a). Esta conclusión baseouse en que a asociación con esquizofrenia da PRS para o inicio do consumo de tabaco e da principal variante xenética para esquizofrenia da agrupación de xenes de nAchR *CHRNA5-CHRNA3-CHRNB4* deixou de ser significativa tras axustar ou estratificar polo hábito tabáquico da mostra. Porén, non podemos descartar que estes resultados sexan debidos a unha falta de potencia ao estratificar a mostra diana.

Unha proba máis da hipótese do tabaco como factor de risco da esquizofrenia é o efecto causal significativo atopado na análise de MR do Estudo 5 do inicio do consumo de tabaco, e outros comportamentos de consumo de tabaco, sobre esquizofrenia. Este efecto causal foi tamén identificado por outros estudos de MR (Barkhuizen et al., 2021; Wootton et al., 2020; Yuan et al., 2020; M. Zhang et al., 2024). Nesta tese, empregouse a metodoloxía GSMR, a cal presenta a vantaxe de detectar e descartar aqueles SNP pleiotrópicos como variables instrumentais (Zhu et al., 2018). Ademais, este estudo baseouse na análise dun total de 99 variables instrumentais. Polo tanto, a evidencia dun efecto causal do inicio do consumo de tabaco sobre esquizofrenia parece bastante robusta.

O Estudo 5 tamén permitiu analizar o impacto na suceptibilidade xenética a esquizofrenia deste efecto causal do inicio do consumo de tabaco sobre esquizofrenia. Neste estudo, examinouse a consecuencia de eliminar da susceptibilidade xenética a esquizofrenia aquela susceptibilidade xenética tamén asociada ao inicio do consumo de tabaco, ponderada polo efecto causal deste trazo de tabaquismo sobre esquizofrenia calculado con MR. Deste xeito, descartouse da susceptibilidade xenética a esquizofrenia aquela derivada de factores xenéticos que incrementan o risco a esquizofrenia a través da vía causal do tabaquismo. Tras eliminar da susceptibilidade xenética a esquizofrenia aquela mediada pola vía causal do tabaquismo, a súa correlación xenética co inicio do consumo de tabaco pasou a ser case inexistente, suxerindo que a suceptibilidade xenética compartida entre ambos trazos é debida, principalmente, a un efecto causal do tabaquismo sobre a esquizofrenia. Polo tanto, en base a

USC UNIVERSIDADE DE SANTIAGO

este resultado, unha pleiotropía horizontal, cuns factores xenéticos aumentando o risco tanto a tabaquismo como a esquizofrenia de forma independente, non parece explicar a susceptibilidade xenética compartida entre ambos trazos.

A principal consecuencia de condicionar o GWAS de esquizofrenia polo inicio do consumo de tabaco foi unha redución da súa correlación xenética con conduta externalizante. Polo tanto, esta susceptibilidade xenética asociada a esquizofrenia a través da vía causal do tabaquismo parece estar capturando, entre outra susceptibilidade xenética asociada a tabaquismo, o efecto dun factor xenético externalizante. Tal como se expuxo anteriormente, isto concorda co Estudo 2 da tese, segundo o cal a vía externalizante ten un maior efecto no inicio do consumo de tabaco que a vía internalizante. Polo tanto, a conduta externalizante parece estar aumentando o risco a esquizofrenia, polo menos parcialmente, a través da vía causal do tabaquismo. Un esquema desta hipótese atópase na Figura 6.

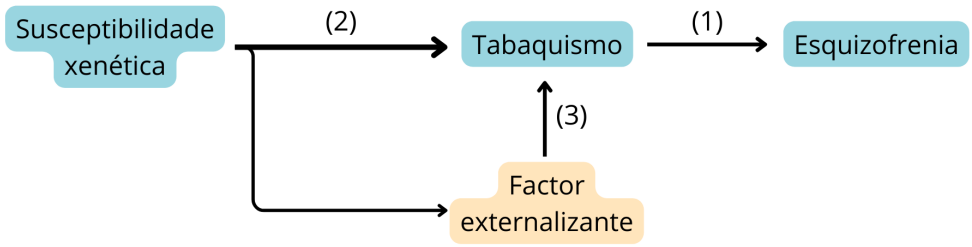


Figura 6. Posible explicación á redución da correlación xenética do GWAS de esquizofrenia con conduta externalizante tras condicionar por tabaquismo. (1): O tabaquismo presenta un efecto causal sobre a esquizofrenia. (2): A susceptibilidade xenética asociada a tabaquismo estará tamén asociada a esquizofrenia a través desta vía causal. (3): Como a etioloxía do tabaquismo é parcialmente explicada por un factor xenético externalizante, a susceptibilidade xenética a conduta externalizante estará asociada a esquizofrenia de forma indirecta a través do tabaquismo.

Esta hipótese de que a base xenética compartida entre tabaquismo e esquizofrenia está relacionada, dentro dos distintos factores xenéticos específicos, co factor xenético externalizante é consistente con estudos previos de genomic SEM. Un destes estudos mostrou que a correlación xenética existente entre o inicio do consumo de tabaco e esquizofrenia deixa de ser significativa ao ter en conta covariables como a asunción de riscos, o consumo de cannabis e alcol ou o insomnio (Barkhuizen et al., 2021). Outro estudo atopou que o factor xenético común a esquizofrenia, inicio do consumo de tabaco e inicio do consumo de cannabis está fortemente correlacionado con trazos externalizantes, mais tamén con trastornos psiquiátricos (W. Song et al., 2022). Porén, en liña coa nosa hipótese, este factor común estivo máis fortemente asociado a ADHD que á propia esquizofrenia.

USC DE COMPOSTELA

Este factor externalizante é un dos factores latentes habitualmente empregados para explicar a estrutura de psicopatoloxía. Tal e como se comentou anteriormente, existe forte evidencia de que este factor externalizante teña un efecto sobre os trazos de consumo de substancias (Caspi et al., 2014; Gard et al., 2021; Jang et al., 2022; Rosenström et al., 2019; Waldman et al., 2020). Porén, nestas análises da estrutura de psicopatoloxía, a esquizofrenia ou outros trazos relacionados coa psicose non aparecen englobados dentro deste factor externalizante. Este factor externalizante si que podería ter un efecto indirecto na esquizofrenia, por exemplo, a través da súa correlación con outros factores latentes. De tódolos xeitos, existe evidencia dunha asociación entre esquizofrenia e conduta externalizante. Por un lado, tal e como se apuntou anteriormente, as persoas con esquizofrenia presentan maiores niveis de conduta externalizante que controis (Arican et al., 2019; Heerey et al., 2007; Nolan et al., 2011; Ouzir, 2013). A nivel xenético, existe unha correlación xenética significativa da esquizofrenia coa tolerancia a condutas de risco (r_g=0,17; SE=0,02) e a toma de riscos (r_g=0,27; SE=0,04) (Karlsson Linnér et al., 2019; Strawbridge et al., 2018). Do mesmo xeito, empregando a ferramenta MiXeR, atopouse un alto solapamento xenético entre a esquizofrenia e comportamentos de toma de riscos (Hindley et al., 2021). Ademais, a esquizofrenia e estes trazos externalizantes presentaron unha correlación xenética das variantes causais compartidas moderada e positiva, variando entre 0,23 e 0,27. Outro estudo tamén mostrou que unha maior PRS para esquizofrenia estivo asociada a unha maior asunción de riscos en pacientes con esquizofrenia, principalmente naqueles/as non medicados/as (Socrates et al., 2021). Polo tanto, existe forte evidencia dunha base xenética compartida entre esquizofrenia e conduta externalizante, a cal, segundo os resultados desta tese, parece estar mediada, polo menos parcialmente, polo consumo de tabaco.

O Estudo 5 da tese tamén identificou loci específicos cuxa asociación con esquizofrenia cambiou tras eliminar aquela susceptibilidade xenética mediada pola vía causal do inicio do consumo de tabaco. Por un lado, atopáronse loci que deixaron de estar significativamente asociados a esquizofrenia, o cal é indicativo de que estaban parcialmente asociados a esquizofrenia a través da vía causal do inicio do consumo de tabaco. Ademais, nalgúns destes "loci perdidos" atopouse que o mesmo sinal estivo asociado con esquizofrenia e conduta externalizante na mesma dirección, o cal suxire que son loci asociados a esquizofrenia a través do factor externalizante que aumenta o risco a tabaquismo e este, á súa vez, a esquizofrenia. Por outro lado, identificáronse loci que pasaron a estar asociados significativamente a esquizofrenia, o que denota que a súa asociación con esquizofrenia no GWAS orixinal podería ter quedado oculta por estar tamén asociados a esquizofrenia en dirección oposta a través da vía

USC DE COMPOSTELA

causal do inicio do consumo de tabaco. Porén, os valores de significación destes loci non cambiaron dramaticamente, estando os "loci perdidos" e os "loci novos" preto da significación no GWAS de esquizofrenia despois de condicionar por tabaquismo e no GWAS de esquizofrenia orixinal, respectivamente. Isto é indicativo de que, ademais de estar asociados a esquizofrenia a través da vía causal do tabaquismo, tamén teñen un efecto directo sobre esquizofrenia por vías independentes ao tabaquismo, na mesma dirección ou en dirección oposta. Esta hipótese concorda coa alta pleiotropía atopada entre os trastornos psiquiátricos e trazos relacionados, habendo unha mestura de variantes xenéticas asociadas na mesma dirección e en direccións opostas a cada un dos trazos dentro das variantes pleiotrópicas (Hindley et al., 2022). Na Figura 7 esquematízase esta hipótese.

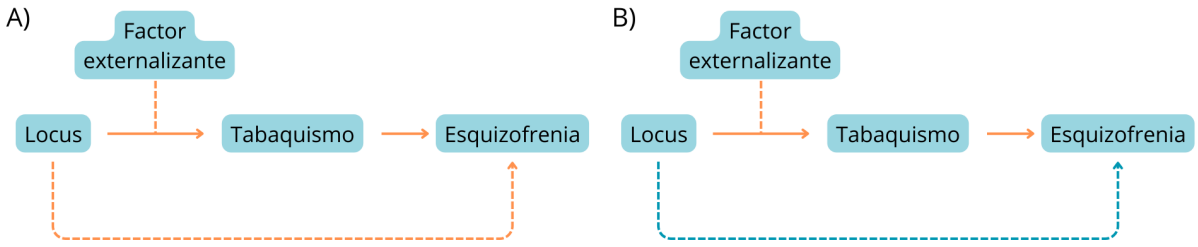


Figura 7. Representación esquemática dos loci cuxa asociación con esquizofrenia cambiou tras condicionar por tabaquismo. Unha mesma cor das frechas representa unha mesma dirección de efecto, cores diferentes representan un efecto en direccións opostas. A) "Loci perdidos" cuxa asociación con esquizofrenia deixa de ser significativa tras condicionar por tabaquismo: loci parcialmente asociados a esquizofrenia a través da vía causal do tabaquismo, na cal pode estar actuando o factor externalizante, e, ademais, de forma independente desta vía causal na mesma dirección. B) "Loci novos" cuxa asociación con esquizofrenia pasa a ser significativa tras condicionar por tabaquismo: loci parcialmente asociados a esquizofrenia a través da vía causal do tabaquismo, na cal pode estar actuando o factor externalizante, e, ademais, de forma independente desta vía causal en dirección oposta.

Polo tanto, nesta tese, empregando diferentes metodoloxías para abordar a relación causal entre tabaquismo e esquizofrenia, púidose atopar evidencia consistente dun efecto causal do inicio do consumo de tabaco sobre a esquizofrenia. Esta triangulación das cuestións científicas, empregando múltiples ferramentas, cada unha coas súas propias asuncións, fortalezas e limitacións, é esencial para chegar a conclusións robustas (Munafò & Davey Smith, 2018). Os resultados desta tese, xunto coa evidencia acumulada nos últimos anos, aumentan a confianza de que o tabaco é un factor de risco para desenvolver trastornos psiquiátricos (Firth et al., 2023).

Non obstante, o efecto causal do inicio do consumo de tabaco sobre esquizofrenia podería estar mediado, polo menos parcialmente, polo consumo de cannabis, o cal se recoñece como un importante factor de risco a esquizofrenia (Fergusson et al., 2015). Existen evidencias de que comezar a consumir tabaco incrementa o risco a consumir cannabis, o cal, á súa vez, podería

aumentar o risco a esquizofrenia (Figura 8) (Z. E. Reed et al., 2022). Esta situación non violaría as condicións da MR xa que o consumo de tabaco e de cannabis estarían na mesma vía causal cara a esquizofrenia. Tamén podería darse un caso de pleiotropía horizontal se as variables instrumentais empregadas na análise de MR estivesen asociadas independentemente a tabaco e a cannabis (Figura 8). Neste caso, o consumo de cannabis actuaría como un factor de confusión na asociación entre tabaquismo e esquizofrenia. Porén, para que esta confusión do cannabis fose a responsable do efecto causal atopado, a maior parte das variables instrumentais empregadas na análise de MR terían que estar asociadas ao consumo de tabaco e de cannabis de forma independente o cal, tendo en conta que a análise de MR se sustentou en 99 variables instrumentais, é pouco probable. Do contrario, a metodoloxía de MR empregada, tal e como se comentou previamente, desbotaría como variables instrumentais aqueles SNP cun efecto pleiotrópico sobre o consumo de cannabis.

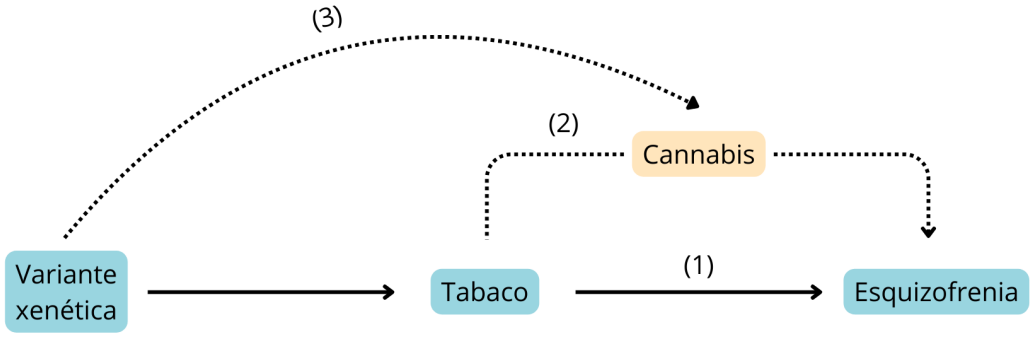


Figura 8. O consumo de cannabis podería estar mediando o efecto causal do tabaco sobre a esquizofrenia atopado en MR. (1): A MR atopa un evidencia dun efecto causal do tabaco sobre a esquizofrenia. (2): O consumo de tabaco podería estar aumentando o risco a consumir cannabis o cal, á súa vez, ten un efecto causal sobre a esquizofrenia. Como ambos trazos están na mesma vía causal cara a esquizofrenia, esta situación non suporía unha violación das condicións da MR. (3): As variantes xenéticas empregadas para avaliar o efecto causal do tabaquismo sobre esquizofrenia poderían estar asociadas de forma independente co consumo de cannabis o cal, á súa vez, ten un efecto causal sobre a esquizofrenia. Este caso de pleiotropía horizontal suporía unha violación das condicións da MR.

De acordo coa hipótese do consumo de cannabis mediando a relación causal entre tabaquismo e esquizofrenia, estudos anteriores mostraron que a base xenética común entre tabaquismo e esquizofrenia relacionada con conduta externalizante podería estar tamén asociada ao consumo de cannabis. Por un lado, un estudo recente revelou que, mentres que aquelas variantes xenéticas de risco asociadas na mesma dirección a CUD, ao inicio do consumo de tabaco e a esquizofrenia estiveron asociadas na mesma dirección a un comportamento de toma de riscos, aquelas concordantes para esquizofrenia e o inicio do consumo de tabaco e discordantes para CUD estiveron negativamente correlacionadas coa

USC UNIVERSIDADE DE SANTIAGO

asunción de riscos (E. C. Johnson et al., 2024). Este resultado concorda co Estudo 2 da tese, no cal se atopou un efecto especialmente importante da vía externalizante sobre os SUD. Se CUD se comporta igual que o resto de SUD analizados neste estudo, podería ser que estas variantes xenéticas cun efecto concordante para o inicio do consumo de tabaco e CUD estean recollendo a susceptibilidade xenética a esta vía externalizante. Pola contra, aquelas discordantes para CUD poderían estar recollendo aquela susceptibilidade xenética asociada a tabaquismo que non depende da vía externalizante. Noutro estudo, a asociación entre o inicio do consumo de tabaco e a esquizofrenia pasou a ser negativa tras ter en conta os efectos do consumo de cannabis, CUD e a dependencia a nicotina (E. C. Johnson et al., 2021). Este resultado concorda co valor obtido no Estudo 5, no cal, tras condicionar o GWAS de esquizofrenia pola vía causal do inicio do consumo de tabaco, a correlación co inicio do consumo de tabaco pasou a ser negativa e, aínda que pequena, significativa. Polo tanto, parece que ao eliminar da susceptibilidade xenética a esquizofrenia aquela compartida co consumo de cannabis e o inicio do consumo de tabaco, segue habendo un pequeno solapamento xenético entre o tabaquismo e a esquizofrenia mais en direccións opostas.

5.2.3. Asociación dos xenes de subunidades de nAchR con tabaquismo e esquizofrenia

Os xenes codificantes para subunidades de nAchR poden ofrecer importante información sobre a orixe da asociación entre tabaquismo e esquizofrenia, xa que, ademais de estar fortemente implicados na resposta á nicotina no cerebro, algúns deles tamén están asociados a esquizofrenia. Polo tanto, se o tabaco presenta un efecto causal sobre a esquizofrenia, tal e como apoian os estudos desta tese, esperaríase que, ao exercer o seu efecto a través destes receptores, estes xenes estivesen asociados a esquizofrenia a través do consumo de tabaco.

Por un lado, no xene *CHRNA2*, a análise de colocalización do Estudo 3 atopou que o mesmo sinal é responsable da súa asociación con diferentes trazos de consumo de tabaco, como a intensidade, o inicio, a frecuencia ou a idade de inicio, e con esquizofrenia, mais tamén con CUD. Polo tanto, en liña cos resultados previamente expostos, a asociación deste xene con esquizofrenia podería estar mediada polo consumo de tabaco, directa ou indirectamente a través do consumo de cannabis. Apoiando esta hipótese, o Estudo 5 da tese mostrou que a asociación deste locus con esquizofrenia se perde tras axustar polo efecto causal do inicio do consumo de tabaco. Ademais, o sinal de asociación deste xene para tabaquismo e esquizofrenia tamén colocaliza cun trazo externalizante, o número de parellas sexuais, mais está asociado con este trazo en dirección oposta. Polo tanto, parece que este locus aumenta o risco a esquizofrenia a través da vía causal do inicio do consumo de tabaco, mais non a través do factor xenético externalizante. Tería sido interesante confirmar esta hipótese no Estudo 4 analizando, ao igual

USC DE SANTIAGO DE COMPOSTELA

que para a agrupación de xenes *CHRNA5-CHRNA3-CHRNB4,* se a asociación deste xene con esquizofrenia se perde ao estratificar a mostra por hábito tabáquico.

Por outro lado, a asociación da agrupación de xenes de nAchR *CHRNA5-CHRNA3-CHRNB4* con tabaquismo e esquizofrenia, aínda que leva tempo sendo estudada, parece ser máis complexa (Hong et al., 2011). A análise de colocalización do Estudo 3 mostrou que a asociación deste locus de nAchR con esquizofrenia non pode ser explicada por unha pleiotropía vertical a través do consumo de tabaco, xa que a asociación deste locus con ambos trazos é debida a sinais de asociación diferentes, o cal se coñece como pleiotropía espuria. Isto non concorda cos resultados presentados previamente que suxiren que a susceptiblidade xenética compartida entre tabaquismo e esquizofrenia é principalmente debida a un efecto causal do inicio do consumo de tabaco sobre esquizofrenia. Unha posible explicación a esta incongruencia podería ser que os xenes de nAchR implicados na farmacoxenética da nicotina cun elevado efecto no tabaquismo se comportasen de xeito diferente do resto da susceptibilidade xenética ao trazo, altamente polixénica (Gelernter & Polimanti, 2021). É dicir, o efecto causal do inicio do consumo de tabaco sobre esquizofrenia atopado a nivel de todo o xenoma podería non ser responsable da asociación deste locus con tabaquismo e esquizofrenia. Outra posible explicación a este resultado podería provir da propia limitación de *coloc* á hora de avaliar a existencia de colocalización para variantes xenéticas secundarias en LD coa variante xenética máis significativa do locus (Giambartolomei et al., 2014). *Coloc* non apoiou a existencia de colocalización para o sinal de asociación máis forte da agrupación de xenes *CHRNA5-CHRNA3-CHRNB4*. Este sinal está relacionada co SNP rs16969968, o cal dá lugar a un cambio aminoacídico na subunidade de nAchR α5 con consecuencias a nivel funcional (Maskos, 2020). Porén, neste locus poderían existir outros sinais de asociación secundarios que fosen comúns para tabaquismo e esquizofrenia. No Estudo 4 volveuse avaliar se a asociación deste locus con esquizofrenia está mediada por tabaquismo. Non obstante, neste estudo, analizouse tamén a asociación doutro SNP diferente ao rs16969968, o rs8042374, que é o SNP deste locus máis asociado a esquizofrenia. En contraposición á análise de colocalización, a asociación deste SNP con esquizofrenia dependeu do hábito tabáquico da mostra, apoiando que está asociado a esquizofrenia a través do tabaquismo. Polo tanto, nesta agrupación de xenes de nAchR parecen existir dous sinais de asociación diferentes: un sinal relacionado coa amplamente estudada variante rs16969968 que está fortemente asociado a trazos de consumo de tabaco, como a intensidade, a frecuencia, o cesamento ou a dependencia e que non é común a esquizofrenia, e un sinal no que está involucrada a variante rs8042374 que seguramente estea asociado a tabaquismo e a esquizofrenia a través da mesma vía causal. Outra posible

USC DE COMPOSTELA

explicación á incongruencia atopada neste locus é que, para esta agrupación de nAchR, se analizou a existencia de colocalización para trazos de consumo de tabaco diferentes ao inicio do consumo de tabaco, xa que non había evidencia de asociación deste locus con este trazo no GWAS empregado nese estudo. Porén, no último GWAS do GSCAN si que se atopou unha asociación desta agrupación de xenes *CHRNA5-CHRNA3-CHRNB4* co inicio do consumo de tabaco, a través de dúas variantes xenéticas en LD moderado coa variante rs16969968 (G. R. B. Saunders et al., 2022). Sería interesante analizar se este sinal de asociación colocaliza para o inicio do tabaquismo e a esquizofrenia, o cal apoiaría a hipótese dun efecto causal do inicio do consumo de tabaco sobre a esquizofrenia.

5.2.4. Hipótese proposta para explicar a asociación entre tabaquismo e esquizofrenia

Reunindo os resultados dos distintos estudos da tese, púidose formular a seguinte hipótese para explicar a asociación entre o inicio do consumo de tabaco e a esquizofrenia (Figura 9). Por un lado, existe evidencia de que a elevada prevalencia de fumadores/as entre as persoas con esquizofrenia se debe a un efecto causal do inicio do consumo de tabaco sobre a esquizofrenia, aínda que podería ser indirecto a través do cannabis. Este efecto causal do tabaquismo sobre esquizofrenia provocaría que a susceptibilidade xenética a tabaquismo estea tamén asociada a esquizofrenia. Dado que o risco ao inicio do consumo de tabaco é en parte explicado por unha maior conduta externalizante, o factor xenético externalizante tamén estaría asociado indirectamente a esquizofrenia a través do inicio do consumo de tabaco. Por outro lado, os factores xenéticos que explican o consumo de tabaco en persoas con esquizofrenia parecen ser os mesmos que os da poboación xeral, en contra da hipótese da automedicación. Porén, as persoas con esquizofrenia parecen comezar a fumar máis por un conduta externalizante que a poboación xeral. Isto podería deberse aos seus maiores niveis de conduta externalizante, os cales poderían ser consecuencia dunha vía externalizante que ten un efecto na etioloxía da esquizofrenia ou poderían ser un síntoma da propia esquizofrenia, o cal sería un caso de causalidade inversa.

Polo tanto, en base a esta hipótese, a asociación entre tabaquismo e esquizofrenia parece ser explicada principalmente por un efecto causal do inicio do tabaquismo sobre a esquizofrenia. Isto contrasta cun estudio anterior onde se atopou evidencia de que a asociación existente entre a esquizofrenia, o inicio do consumo de tabaco e o inicio do consumo de cannabis se explica tanto por unha relación causal como por uns factores xenéticos que aumentan o risco aos trazos de xeito independente (W. Song et al., 2022). Outro estudo chegou a esta mesma conclusión, atopando evidencia de causalidade bidireccional entre esquizofrenia,

USC DE SANTIAGO DE COMPOSTELA

CUD e o inicio do consumo de tabaco mais tamén de pleiotropía horizontal entre estes tres trazos (E. C. Johnson et al., 2024).

A importancia destes distintos factores xenéticos que explican o tabaquismo nas persoas con esquizofrenia podería depender do momento do desenvolvemento. Por exemplo, o tabaco podería actuar como un factor de risco a desenvolver esquizofrenia principalmente cando o seu consumo se produce a idades temperás, alterando procesos do neurodesenvolvemento durante a adolescencia (MacCabe, 2018).

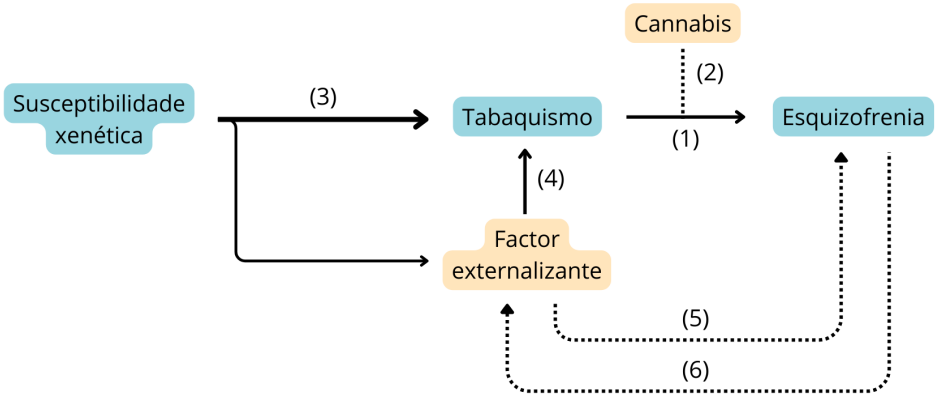


Figura 9. Hipótese para explicar a asociación entre o tabaquismo e a esquizofrenia que unifica os resultados da tese. (1): O consumo de tabaco presenta un efecto causal sobre a esquizofrenia, aumentando a prevalencia de tabaquismo en pacientes con esquizofrenia. (2): O consumo de tabaco podería incrementar o risco a esquizofrenia indirectamente a través do cannabis. (3): A susceptibilidade xenética asociada a tabaquismo estará tamén asociada a esquizofrenia a través desta vía causal. (4): Dado que o risco ao inicio do consumo de tabaco é en parte explicado por unha maior conduta externalizante, o factor xenético externalizante tamén estaría asociado indirectamente a esquizofrenia a través do tabaquismo. (5): Os factores xenéticos que aumentan o risco a conduta externalizante poderían tamén ter un efecto na etioloxía da esquizofrenia de forma independente ao tabaquismo, o cal aumentaría a prevalencia de tabaquismo en pacientes con esquizofrenia. (6): A propia esquizofrenia podería dar lugar a maiores niveis de conduta externalizante nas persoas que a sofren o cal, á súa vez, incrementaría o seu risco a comezar a fumar.

5.2.5. Diferenzas na asociación entre tabaquismo e esquizofrenia segundo o trazo de consumo de tabaco

Nos estudos 4 e 5 da tese, os cales proporcionan as principais evidencias sobre a asociación entre tabaquismo e esquizofrenia, o trazo de consumo de tabaco avaliado foi o inicio do consumo de tabaco. Tal e como mostrou o Estudo 2 da tese, os distintos trazos de consumo de tabaco poden presentar unha etioloxía diferente, con distintos factores xenéticos participando na súa susceptibilidade. En base a isto, a asociación doutros trazos de consumo de tabaco con esquizofrenia podería ter outra explicación. Por exemplo, a susceptibilidade xenética compartida entre esquizofrenia e outros trazos de consumo de tabaco podería recoller outros factores, como o factor de psicopatoloxía xeral, e non só o factor externalizante. Apoiando esta

hipótese, un dos estudos comentados anteriormente mostrou que, mentres que o solapamento xenético entre o inicio do consumo de tabaco e a esquizofrenia podería ser explicado por outras covariables como un comportamento de toma de riscos, o consumo de cannabis e alcol, ou o insomnio, a intensidade do consumo de tabaco seguiu estando correlacionada xeneticamente con esquizofrenia tras axustar por estas covariables (Barkhuizen et al., 2021). Este resultado concorda co Estudo 2 da tese, no cal se mostrou que a vía internalizante, ademais da vía externalizante, presenta un efecto na intensidade do consumo de tabaco, comportándose este como un indicador xenético de TUD.

Do mesmo xeito, no Estudo 4 da tese constatouse que os factores xenéticos que incrementan o risco ao inicio do consumo de tabaco na poboación xeral tamén aumentan o risco a ser fumador/a en pacientes con esquizofrenia. Porén, podería ser que outros trazos de consumo de tabaco en pacientes con esquizofrenia non se expliquen polos mesmos factores xenéticos que os da poboación xeral. Por exemplo, un estudo anterior mostrou que os factores xenéticos asociados á intensidade do consumo de tabaco na poboación xeral e en pacientes con esquizofrenia non parecen estar significativamente correlacionados (Peterson et al., 2021). Non obstante, este resultado podería deberse a unha falta de potencia. De feito, neste mesmo estudo, unha PRS para a intensidade do consumo de tabaco estimada en poboación xeral explicou este fenotipo en pacientes con esquizofrenia. Esta variabilidade, polo tanto, suxire que a relación entre tabaquismo e esquizofrenia é complexa, implicando diferentes vías e factores xenéticos dependendo do trazo de consumo de tabaco considerado.

5.3. BREVES CONSIDERACIÓNS CLÍNICAS

Os estudos desta tese permitiron obter máis evidencia sobre aqueles factores xenéticos que incrementan o risco ao consumo de substancias, especialmente ao tabaquismo en pacientes con esquizofrenia. Aínda que é necesaria a replicación destes resultados noutros estudos, esta información podería ter certas implicacións a nivel clínico, permitindo desenvolver estratexias máis eficaces para prever e tratar o consumo de substancias.

Nesta tese confirmouse a existencia de dúas vías de risco que actúan ao longo do curso do consumo de substancias. Por un lado, podería ser importante a detección de problemas externalizantes, xa durante a infancia, para prever un posterior comezo de consumo de substancias. Esta conclusión tamén se estende ás persoas con esquizofrenia. Por outro lado, unha vez iniciado o consumo da substancia, o tratamento de problemas externalizantes e internalizantes podería evitar un consumo problemático da mesma.

USC UNIVERSIDADE DE SANTIAGO

O tabaquismo contribúe notablemente á mortalidade das persoas con esquizofrenia polo que comprender as bases da asociación entre tabaquismo e esquizofrenia podería ser importante para mellorar a saúde das persoas con esquizofrenia (Bobes et al., 2010). A evidencia suxire que o tabaco actúa como un factor de risco para o desenvolvemento da esquizofrenia, polo que, tanto prever o comezo do consumo de tabaco como promover o cesamento do seu consumo, terá efectos beneficiosos para a saúde mental das persoas (G. M. J. Taylor & Munafò, 2019). Tal e como se comentou previamente, estratexias de prevención do consumo de tabaco naquelas persoas con problemas externalizantes poderían ser tamén eficaces para diminuír o risco a tabaquismo nas persoas con esquizofrenia. En todo caso, esta tese non atopa evidencias da hipótese da automedicación, a cal actúa como unha barreira á hora de implementar medidas para reducir o consumo de tabaco neste grupo poboacional (Lum et al., 2018). Polo tanto, é importante que as persoas con esquizofrenia poidan acceder aos mesmos tratamentos que a poboación xeral para deixar de fumar, o cal melloraría substancialmente a súa calidade de vida.

5.4. LIMITACIÓNS XERAIS DOS ESTUDOS

Os estudos da tese presentan unha serie de limitacións comúns que deben ser tidas en conta á hora de interpretar os resultados. Principalmente, estas limitacións están relacionadas cos GWAS empregados ao longo da tese. Por un lado, só puideron utilizarse aqueles GWAS dispoñibles publicamente, os cales, para algúns trazos, non presentan o maior tamaño de mostra. Por exemplo, en varios casos, o GWAS máis amplo incluía a mostra de *23andMe*, mais para esta mostra non é posible distribuír publicamente os resultados de GWAS. Isto limitou a potencia estatística dalgunhas análises. Ademais, moitos dos GWAS empregados baséanse en mostras, como o *UK Biobank*, que poderían non ser representativas da poboación xeral en termos de xénero, nivel socioeconómico ou ascendencia, o cal introduciría un nesgo nos nosos resultados (Fry et al., 2017; MacDermod et al., 2022). Por outro lado, moitos dos GWAS analizados, principalmente aqueles para comportamentos de consumo de substancias e outros trazos condutuais e da personalidade, están baseados no que se coñece como "fenotipado mínimo" (do inglés, *minimal phenotyping*), é dicir, no uso dun pequeno número de preguntas respondidas polo propio individuo para identificar o trazo. Existe evidencia de que este enfoque, aínda que permite aumentar o tamaño de mostra dos GWAS, nesga os resultados dos mesmos, reducindo a súa especificidade (Cai et al., 2020). Ademais, os GWAS non só capturan os efectos directos do xenotipo sobre o trazo analizado, senón que tamén efectos indirectos como aqueles debidos á correlación xene-ambiente, á estratificación da poboación ou ao apareamento orientado segundo o trazo analizado ou trazos correlacionados, o cal pode afectar á interpretación dos resultados (Abdellaoui et al., 2022; Plomin & von Stumm, 2022). Esta

influencia de efectos indirectos é particularmente importante para trazos como o éxito educativo, onde a asociación da PRS co fenotipo se reduce á metade cando se controla polas PRS dos/as pais/nais e para o cal parece producirse un apareamento orientado non só segundo o éxito académico senón que tamén por outros factores correlacionados (Okbay et al., 2022). Polo tanto, é preciso ser cautos á hora de transmitir os resultados derivados de GWAS para trazos complexos como o éxito educativo ou trazos relacionados con comportamentos sexuais (Abdellaoui et al., 2022). Ao estar afectados por correlacións xene-ambiente debidas a procesos sociais, económicos ou políticos, poden realizarse interpretacións erróneas que dean lugar a unha estigmatización de determinados grupos poboacionais. Finalmente, tal e como se especificou anteriormente, para a maior parte das análises, empregáronse GWAS restrinxidos a mostras de ascendencia europea. Ata agora, a maior parte da mostra dos GWAS, estímase que máis do 78%, é de ascendencia europea, estando outras poboacións claramente infrarrepresentadas neles (Sirugo et al., 2019). Isto supón unha limitación á hora de xeneralizar os resultados dos estudos a outras ascendencias (Martin et al., 2017; Peterson et al., 2019).

Por outra banda, os resultados obtidos con ferramentas como MiXeR e *coloc* non implican necesariamente a existencia dunha relación causal entre os trazos analizados. Estas metodoloxías só nos permiten avaliar se certas características da susceptibilidade xenética compartida entre os trazos son compatibles cunha relación causal entre eles. Por un lado, MiXeR permitiu analizar se había un solapamento xenético dos trazos externalizantes e internalizantes con trazos de consumo de substancias, así como unha distribución do efecto das variantes xenéticas compartidas similar para ambos trazos, o cal é compatible cun efecto causal dos trazos externalizantes ou internalizantes sobre os trazos de consumo de substancias. Por outro lado, *coloc* permitiu obter evidencia de que o sinal de asociación no xene *CHRNA2* é común a tabaquismo e a esquizofrenia, condición necesaria para afirmar que este locus estea asociado a esquizofrenia a través do tabaquismo e, polo tanto, que o tabaquismo teña un efecto causal sobre a esquizofrenia. Porén, outras situacións como unha pleiotropía horizontal entre os trazos ou unha causalidade no sentido inverso tamén poderían explicar estes resultados.

Ademais, aínda que existe forte evidencia de que o cannabis é un factor de risco a desenvolver esquizofrenia, non se tivo en conta o seu efecto na asociación entre o tabaquismo e a esquizofrenia. O tabaquismo podería ter un efecto indirecto sobre a esquizofrenia promovendo o consumo de cannabis. Polo tanto, sería interesante analizar se as mesmas vías de risco cara o consumo de tabaco actúan no consumo de cannabis. Porén, a análise das vías de risco cara o consumo de substancias focalizouse só no consumo de dúas substancias legais: o tabaco e o alcol, debido á potencia limitada dos GWAS para cannabis ou outras substancias.

USC UNIVERSIDADE DE SANTIAGO

Poderían existir tamén outros factores de confusión na asociación entre tabaquismo e esquizofrenia. Por exemplo, existe unha clara asociación entre o tabaquismo materno durante o embarazo e o desenvolvemento de esquizofrenia, o cal podería estar dando lugar ao efecto causal atopado entre tabaquismo e esquizofrenia (Hunter et al., 2020; Niemelä et al., 2016). Con todo, un estudo de datos poboacionais de Suecia evidenciou que esta asociación entre o tabaquismo materno e a esquizofrenia, máis que reflectir unha relación causal, parece ser debida a factores compartidos cos/as irmáns/ás (Quinn et al., 2017). Polo tanto, este trazo non parece ter un efecto importante na asociación entre tabaquismo e esquizofrenia.

Tal e como se mencionou anteriormente, o trazo de tabaquismo principalmente analizado foi o inicio do consumo de tabaco, polo que as conclusións se deben estender a outros trazos de consumo de tabaco con cautela. A razón pola que se analizou principalmente este trazo foi que é o comportamento de consumo de tabaco para o cal o GWAS do GSCAN presenta un maior tamaño de mostra (G. R. B. Saunders et al., 2022). Ademais, na mostra diana de pacientes con esquizofrenia e controis empregada para as análises de PRS, só se coñecía se os individuos eran fumadores/as ou non nese momento, sen considerar a intensidade do seu tabaquismo ou se presentaban dependencia a nicotina. Polo tanto, a PRS para o inicio do consumo de tabaco era a que mellor se adaptaba a este fenotipo.

Outra importante limitación das análises foi non ter en conta o efecto do sexo e o xénero máis que como covariable nalgunhas análises. Porén, factores relacionados co sexo e o xénero teñen unha importante influencia na variabilidade de trazos condutuais e trastornos psiquiátricos, seguramente existindo unha interacción entre o xenotipo e o xénero (Bernabeu et al., 2021; Bölte et al., 2023). Por exemplo, é coñecido que as vías de risco cara o consumo de substancias externalizante e internalizante presentan un diferente efecto segundo o xénero (Dir & Hulvershorn, 2019; Heitzeg et al., 2018). Polo tanto, tería sido importante avaliar o impacto desta variable nos resultados, aínda que seguramente non se contaría coa suficiente potencia estatística. Finalmente, outra limitación desta tese foi que non se levou a cabo unha replicación dos resultados noutras mostras. Porén, a maior parte da evidencia presentada susténtase en diferentes análises empregando distintas metodoloxías e mostras.

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

6. CONCLUSIONS

C1. Gene modules with correlated expression in the zebrafish brain are enriched in genes involved in synapses, genes differentially expressed in human brain tissues and developmental stages, as well as genes associated with psychiatry-related traits in GWAS. Therefore, zebrafish brain co-expression modules can model human brain gene networks.

C2. A PRS based on genes from co-expression modules associated with substance withdrawal in zebrafish brain better explains alcohol dependence in humans compared to a PRS based on randomly selected brain-expressed genes. Therefore, specific gene sets with putative altered expression in alcohol dependence were identified from zebrafish expression data, supporting the utility of zebrafish as an animal model in psychiatry.

C3. The pattern of genetic overlap between externalizing traits and substance use traits is consistent with a causal effect of externalizing problems on substance use behaviors and SUD, with this causal effect being particularly strong for SUD. The pattern of genetic overlap between internalizing traits and substance use traits is consistent with a causal effect of internalizing problems on SUD, but not on substance use behaviors. Therefore, the externalizing and internalizing pathways to substance use seem to have variable effects depending on the substance use phase, acting differently on substance use behaviors and SUD.

C4. Genetic factors that increase smoking risk in the general population also explain smoking status in patients with schizophrenia. Therefore, the self-medication hypothesis does not seem to explain the higher prevalence of tobacco use in patients with schizophrenia compared to controls.

C5. PRSs for externalizing traits explain smoking status in patients with schizophrenia better than in controls. Therefore, patients with schizophrenia seem to have a higher risk of smoking through the externalizing pathway compared to the general population.

C6. The association signal at the *CHRNA2* gene with smoking and schizophrenia is due to the same causal variant. Moreover, smoking initiation PRS and a specific genetic variant at the nAchR gene cluster *CHRNA5-CHRNA3-CHRNB4* are no longer significantly associated with schizophrenia after adjustment by smoking status. Similarly, MR analysis indicates a causal

USC UNIVERSIDADE DE SANTIAGO

effect of smoking on schizophrenia. Therefore, part of the high comorbidity between smoking and schizophrenia seems to be due to tobacco being a risk factor for developing schizophrenia.

C7. Main association signals for smoking and schizophrenia at some nAchR loci, specifically, the *CHRNA4* gene and the *CHRNA5-CHRNA3-CHRNB4* gene cluster, do not colocalize. Therefore, these loci also seem to be associated with both traits through independent pathways.

C8. Conditioning the schizophrenia GWAS by the causal effect of smoking initiation reveals new loci significantly associated with schizophrenia and identifies loci that are no longer significantly associated with schizophrenia. Additionally, genetic correlation of schizophrenia with externalizing behavior decreases after conditioning genetic susceptibility to schizophrenia on smoking initiation. Therefore, the causal effect of smoking on schizophrenia has an impact on genetic susceptibility to schizophrenia. An externalizing genetic factor seems to be associated with schizophrenia partially through the causal pathway of smoking.

USC
UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

BIBLIOGRAFÍA

Abbasi, J. (2017). *23andMe, Big Data, and the Genetics of Depression*. *317*(1), 14-16.

Abdellaoui, A., Dolan, C. V., Verweij, K. J. H., & Nivard, M. G. (2022). Gene–environment correlations across geographic regions affect genome-wide association studies. *Nature Genetics*, *54*(9), 1345-1354. https://doi.org/10.1038/s41588-022-01158-0

Abdellaoui, A., Smit, D. J. A., van den Brink, W., Denys, D., & Verweij, K. J. H. (2021). Genomic relationships across psychiatric disorders including substance use disorders. *Drug and Alcohol Dependence*, *220*, 108535. https://doi.org/10.1016/j.drugalcdep.2021.108535

Adler, E., Hoffer, L. D., & Wiser, A. (1993). Normalization of Auditory Physiology by Cigarette Smoking in Schizophrenic Patients. *Am J Psychiatry*, *150*(12), 1856-1861.

Agostini, J. F., Costa, N. L. F., Bernardo, H. T., Baldin, S. L., Mendes, N. V., de Pieri Pickler, K., Manenti, M. C., & Rico, E. P. (2020). Ceftriaxone Attenuated Anxiety-Like Behavior and Enhanced Brain Glutamate Transport in Zebrafish Subjected to Alcohol Withdrawal. *Neurochemical Research*, *45*(7), 1526-1535. https://doi.org/10.1007/s11064-020-03008-z

Agrawal, A., Verweij, K. J. H., Gillespie, N. A., Heath, A. C., Lessov-Schlaggar, C. N., Martin, N. G., Nelson, E. C., Slutske, W. S., Whitfield, J. B., & Lynskey, M. T. (2012). The genetics of addiction—a translational perspective. *Translational Psychiatry*, *2*(7), e140-e140. https://doi.org/10.1038/tp.2012.54

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

Al-Soufi, L., & Costas, J. (2023). Genetic susceptibility for schizophrenia after adjustment by genetic susceptibility for smoking: implications in identification of risk genes and genetic correlation with related traits. *Psychological Medicine*, *53*(14), 6806-6816. https://doi.org/10.1017/S0033291723000326

Al-Soufi, L., Martorell, L., Moltó, M. D., González-Peñas, J., García-Portilla, M. P., Arrojo,

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

M., Rivero, O., Gutiérrez-Zotes, A., Nácher, J., Muntané, G., Paz, E., Páramo, M., Bobes, J., Arango, C., Sanjuan, J., Vilella, E., & Costas, J. (2022). A polygenic approach to the association between smoking and schizophrenia. *Addiction Biology*, *27*(1), 1-10. https://doi.org/10.1111/adb.13104

Alderson, H. L., & Lawrie, S. M. (2015). Does cigarette smoking cause psychosis? *The Lancet Psychiatry*, *2*(8), 672-673. https://doi.org/10.1016/S2215-0366(15)00239-4

Aleksander, S. A., Balhoff, J., Carbon, S., Cherry, J. M., Drabkin, H. J., Ebert, D., Feuermann, M., Gaudet, P., Harris, N. L., Hill, D. P., Lee, R., Mi, H., Moxon, S., Mungall, C. J., Muruganugan, A., Mushayahama, T., Sternberg, P. W., Thomas, P. D., Van Auken, K., … Westerfield, M. (2023). The Gene Ontology knowledgebase in 2023. *Genetics*, *224*(1), 1-14. https://doi.org/10.1093/genetics/iyad031

Allegrini, A. G., Cheesman, R., Rimfeld, K., Selzam, S., Pingault, J., Eley, T. C., & Plomin, R. (2020). The p factor: genetic analyses support a general dimension of psychopathology in childhood and adolescence. *Journal of Child Psychology and Psychiatry*, *61*(1), 30-39. https://doi.org/10.1111/jcpp.13113

Alnæs, D., Kaufmann, T., Van Der Meer, D., Córdova-Palomera, A., Rokicki, J., Moberget, T., Bettella, F., Agartz, I., Barch, D. M., Bertolino, A., Brandt, C. L., Cervenka, S., Djurovic, S., Doan, N. T., Eisenacher, S., Fatouros-Bergman, H., Flyckt, L., Di Giorgio, A., Haatveit, B., … Westlye, L. T. (2019). Brain Heterogeneity in Schizophrenia and Its Association with Polygenic Risk. *JAMA Psychiatry*, *76*(7), 739-748. https://doi.org/10.1001/jamapsychiatry.2019.0257

American Psychiatric Association. (1994). *Diagnostic and Statistical Manual of Mental Disorders (4th ed)*. American Psychiatric Association. https://doi.org/10.1176/appi.books.9780890420249.dsm-iv-tr

American Psychiatric Association. (2013). *Diagnostic and Statistical Manual of Mental Disorders (5th ed)*. American Psychiatric Association. https://doi.org/10.1176/appi.books.9780890425596

Andreassen, O. A., Hindley, G. F. L., Frei, O., & Smeland, O. B. (2023). New insights from the last decade of research in psychiatric genetics: discoveries, challenges and clinical implications. *World Psychiatry*, *22*(1), 4-24. https://doi.org/10.1002/wps.21034

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

Antón-Galindo, E., Adel, M. R., García-González, J., Leggieri, A., López-Blanch, L., Irimia, M., Norton, W. H. J., Brennan, C. H., Fernàndez-Castillo, N., & Cormand, B. (2024). Pleiotropic contribution of rbfox1 to psychiatric and neurodevelopmental phenotypes in two zebrafish models. *Translational Psychiatry*, *14*(1), 99. https://doi.org/10.1038/s41398-024-02801-6

Arican, I., Bass, N., Neelam, K., Wolfe, K., McQuillin, A., & Giaroli, G. (2019). Prevalence of attention deficit hyperactivity disorder symptoms in patients with schizophrenia. *Acta Psychiatrica Scandinavica*, *139*(1), 89-96. https://doi.org/10.1111/acps.12948

Arnold, P. D., Askland, K. D., Barlassina, C., Bellodi, L., Bienvenu, O. J., Black, D., Bloch, M., Brentani, H., Burton, C. L., Camarena, B., Cappi, C., Cath, D., Cavallini, M., Conti, D., Cook, E., Coric, V., Cullen, B. A., Cusi, D., Davis, L. K., … 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

Arslan-Ergul, A., & Adams, M. M. (2014). Gene expression changes in aging Zebrafish (Danio rerio) brains are sexually dimorphic. *BMC Neuroscience*, *15*(1), 29. https://doi.org/10.1186/1471-2202-15-29

Aschbrenner, K. A., Dixon, L. B., Naslund, J. A., Bienvenida, J. C. M., McManus, K. L., Bartels, S. J., & Brunette, M. F. (2017). An Online Survey of Family Members' Beliefs and Attitudes About Smoking and Mental Illness. *Journal of Dual Diagnosis*, *13*(3), 179-183. https://doi.org/10.1080/15504263.2017.1326651

Banu, S., Srivastava, S., Mohammed, A., Kushawah, G., Sowpati, D. T., & Mishra, R. K. (2020). Transcriptomic dataset of zebrafish tissues following chronic alcohol exposure and withdrawal. *Data in Brief*, *33*, 106442. https://doi.org/10.1016/j.dib.2020.106442

Barkhuizen, W., Dudbridge, F., & Ronald, A. (2021). Genetic overlap and causal associations between smoking behaviours and mental health. *Scientific Reports*, *11*(1), 1-13. https://doi.org/10.1038/s41598-021-93962-7

Barkhuizen, W., Taylor, M. J., Freeman, D., & Ronald, A. (2019). A Twin Study on the Association Between Psychotic Experiences and Tobacco Use During Adolescence. *Journal of the American Academy of Child and Adolescent Psychiatry*, *58*(2), 267-276.e8. https://doi.org/10.1016/j.jaac.2018.06.037

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

Bartels, M., Cacioppo, J. T., Van Beijsterveldt, T. C. E. M., & Boomsma, D. I. (2013). Exploring the association between well-being and psychopathology in adolescents. *Behavior Genetics*, *43*(3), 177-190. https://doi.org/10.1007/s10519-013-9589-7

Baselmans, B., Hammerschlag, A. R., Noordijk, S., Ip, H., van der Zee, M., de Geus, E., Abdellaoui, A., Treur, J. L., & van 't Ent, D. (2022). The Genetic and Neural Substrates of Externalizing Behavior. *Biological Psychiatry Global Open Science*, *2*(4), 389-399. https://doi.org/10.1016/j.bpsgos.2021.09.007

Belbasis, L., Köhler, C. A., Stefanis, N., Stubbs, B., van Os, J., Vieta, E., Seeman, M. V., Arango, C., Carvalho, A. F., & Evangelou, E. (2018). Risk factors and peripheral biomarkers for schizophrenia spectrum disorders: an umbrella review of meta-analyses. *Acta Psychiatrica Scandinavica*, *137*(2), 88-97. https://doi.org/10.1111/acps.12847

Bell, J., & Strang, J. (2020). Medication Treatment of Opioid Use Disorder. *Biological Psychiatry*, *87*(1), 82-88. https://doi.org/10.1016/j.biopsych.2019.06.020

Beran, T. N., & Violato, C. (2010). Structural equation modeling in medical research: a primer. *BMC Research Notes*, *3*(1), 267. https://doi.org/10.1186/1756-0500-3-267

Bernabeu, E., Canela-Xandri, O., Rawlik, K., Talenti, A., Prendergast, J., & Tenesa, A. (2021). Sex differences in genetic architecture in the UK Biobank. *Nature Genetics*, *53*(9), 1283-1289. https://doi.org/10.1038/s41588-021-00912-0

Berndt, S. I., Gustafsson, S., Mägi, R., Ganna, A., Wheeler, E., Feitosa, M. F., Justice, A. E., Monda, K. L., Croteau-Chonka, D. C., Day, F. R., Esko, T., Fall, T., Ferreira, T., Gentilini, D., Jackson, A. U., Luan, J., Randall, J. C., Vedantam, S., Willer, C. J., … Ingelsson, E. (2013). Genome-wide meta-analysis identifies 11 new loci for anthropometric traits and provides insights into genetic architecture. *Nature Genetics*, *45*(5), 501-512. https://doi.org/10.1038/ng.2606

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

Bhalla, I. P., Stefanovics, E. A., & Rosenheck, R. A. (2017). Clinical Epidemiology of Single Versus Multiple Substance Use Disorders. *Medical Care*, *55*(9), S24-S32.

USC DE COMPOSTELA

https://doi.org/10.1097/MLR.0000000000000731

Bhavsar, V., Jauhar, S., Murray, R. M., Hotopf, M., Hatch, S. L., McNeill, A., Boydell, J., & MacCabe, J. H. (2018). Tobacco smoking is associated with psychotic experiences in the general population of South London. *Psychological Medicine*, *48*(1), 123-131. https://doi.org/10.1017/S0033291717001556

Bloom, E. L., Matsko, S. V., & Cimino, C. R. (2014). The relationship between cigarette smoking and impulsivity: A review of personality, behavioral, and neurobiological assessment. *Addiction Research and Theory*, *22*(5), 386-397. https://doi.org/10.3109/16066359.2013.867432

Bobes, J., Arango, C., Garcia-Garcia, M., & Rejas, J. (2010). Healthy lifestyle habits and 10-year cardiovascular risk in schizophrenia spectrum disorders: An analysis of the impact of smoking tobacco in the CLAMORS schizophrenia cohort. *Schizophrenia Research*, *119*(1-3), 101-109. https://doi.org/10.1016/j.schres.2010.02.1030

Boehm, F. J., & Zhou, X. (2022). Statistical methods for Mendelian randomization in genome-wide association studies: A review. *Computational and Structural Biotechnology Journal*, *20*, 2338-2351. https://doi.org/10.1016/j.csbj.2022.05.015

Bolstad, B. M., Irizarry, R. A., Åstrand, M., & Speed, T. P. (2003). A comparison of normalization methods for high density oligonucleotide array data based on variance and bias. *Bioinformatics*, *19*(2), 185-193. https://doi.org/10.1093/bioinformatics/19.2.185

Bölte, S., Neufeld, J., Marschik, P. B., Williams, Z. J., Gallagher, L., & Lai, M.-C. (2023). Sex and gender in neurodevelopmental conditions. *Nature Reviews Neurology*, *19*(3), 136-159. https://doi.org/10.1038/s41582-023-00774-6

Bradford, Y. M., Van Slyke, C. E., Howe, D. G., Fashena, D., Frazer, K., Martin, R., Paddock, H., Pich, C., Ramachandran, S., Ruzicka, L., Singer, A., Taylor, R., Tseng, W. C., & Westerfield, M. (2023). From multiallele fish to nonstandard environments, how ZFIN assigns phenotypes, human disease models, and gene expression annotations to genes. *Genetics*, *224*(1), 1-10. https://doi.org/10.1093/genetics/iyad032

Brasher, M. S., Mize, T. J., Thomas, A. L., Hoeffer, C. A., Ehringer, M. A., & Evans, L. M. (2023). Testing associations between human anxiety and genes previously implicated by mouse anxiety models. *Genes, Brain and Behavior*, *22*(6), 1-12.

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

https://doi.org/10.1111/gbb.12851

Buckley, P. F. (2006). Prevalence and consequences of the dual diagnosis of substance abuse and severe mental illness. *Journal of Clinical Psychiatry*, *67*(Suppl 7), 5-9.

Bulik-Sullivan, B., Finucane, H. K., Anttila, V., Gusev, A., Day, F. R., Loh, P.-R., Duncan, L., Perry, J. R. B., Patterson, N., Robinson, E. B., Daly, M. J., Price, A. L., & Neale, B. M. (2015). An atlas of genetic correlations across human diseases and traits. *Nature Genetics*, *47*(11), 1236-1241. https://doi.org/10.1038/ng.3406

Bulik-Sullivan, B. K., & Neale, B. M. (2015). LD Score Regression Distinguishes Confounding from Polygenicity in GWAS. *Nature Genetics*, *47*(3), 291-295. https://doi.org/10.1038/ng.3211.LD

Bycroft, C., Freeman, C., Petkova, D., Band, G., Elliott, L. T., Sharp, K., Motyer, A., Vukcevic, D., Delaneau, O., O'Connell, J., Cortes, A., Welsh, S., Young, A., Effingham, M., McVean, G., Leslie, S., Allen, N., Donnelly, P., & Marchini, J. (2018). The UK Biobank resource with deep phenotyping and genomic data. *Nature*, *562*(7726), 203-209. https://doi.org/10.1038/s41586-018-0579-z

Byrne, E. M., Ferreira, M. A. R., Xue, A., Lindström, S., Jiang, X., Yang, J., Easton, D. F., Wray, N. R., & Chenevix-Trench, G. (2019). Is Schizophrenia a Risk Factor for Breast Cancer?-Evidence from Genetic Data. *Schizophrenia Bulletin*, *45*(6), 1251-1256. https://doi.org/10.1093/schbul/sby162

Byrne, E. M., Zhu, Z., Qi, T., Skene, N. G., Bryois, J., Pardinas, A. F., Stahl, E., Smoller, J. W., Rietschel, M., Owen, M. J., Walters, J. T. R., O'Donovan, M. C., McGrath, J. G., Hjerling-Leffler, J., Sullivan, P. F., Goddard, M. E., Visscher, P. M., Yang, J., & 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

Cabana-Domínguez, J., Antón-Galindo, E., Fernàndez-Castillo, N., Singgih, E. L., O'Leary, A., Norton, W. H., Strekalova, T., Schenck, A., Reif, A., Lesch, K.-P., Slattery, D., & Cormand, B. (2023). The translational genetics of ADHD and related phenotypes in model organisms. *Neuroscience & Biobehavioral Reviews*, *144*, 104949. https://doi.org/10.1016/j.neubiorev.2022.104949

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

Cachat, J., Canavello, P., Elegante, M., Bartels, B., Hart, P., Bergner, C., Egan, R., Duncan, A., Tien, D., Chung, A., Wong, K., Goodspeed, J., Tan, J., Grimes, C., Elkhayat, S., Suciu, C., Rosenberg, M., Chung, K. M., Kadri, F., … Kalueff, A. V. (2010). Modeling withdrawal syndrome in zebrafish. *Behavioural Brain Research*, *208*(2), 371-376. https://doi.org/10.1016/j.bbr.2009.12.004

Cai, N., Revez, J. A., Adams, M. J., Andlauer, T. F. M., Breen, G., Byrne, E. M., Clarke, T.-K., Forstner, A. J., Grabe, H. J., Hamilton, S. P., Levinson, D. F., Lewis, C. M., Lewis, G., Martin, N. G., Milaneschi, Y., Mors, O., Müller-Myhsok, B., Penninx, B. W. J. H., Perlis, R. H., … Flint, J. (2020). Minimal phenotyping yields genome-wide association signals of low specificity for major depression. *Nature Genetics*, *52*(4), 437-447. https://doi.org/10.1038/s41588-020-0594-5

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

Carey, C. E., Agrawal, A., Bucholz, K. K., Hartz, S. M., Lynskey, M. T., Nelson, E. C., Bierut, L. J., & Bogdan, R. (2016). Associations between Polygenic Risk for Psychiatric Disorders and Substance Involvement. *Frontiers in Genetics*, *7*, 149. https://doi.org/10.3389/fgene.2016.00149

Caspi, A., Houts, R. M., Belsky, D. W., Goldman-Mellor, S. J., Harrington, H., Israel, S., Meier, M. H., Ramrakha, S., Shalev, I., Poulton, R., & Moffitt, T. E. (2014). The p factor: One general psychopathology factor in the structure of psychiatric disorders? *Clinical Psychological Science*, *2*(2), 119-137. https://doi.org/10.1177/2167702613497473

Caspi, A., & Moffitt, T. E. (2018). All for one and one for all: Mental disorders in one dimension. *American Journal of Psychiatry*, *175*(9), 831-844. https://doi.org/10.1176/appi.ajp.2018.17121383

Castellanos-Ryan, N., Briere, F. N., O'Leary-Barrett, M., Banaschewski, T., Bokde, A., Bromberg, U., Biichel, C., Flor, H., Frouin, V., Gallinat, J., Garavan, H., Martinot, J. L., Nees, F., Paus, T., Pausova, Z., Rietschel, M., Smolka, M. N., Robbins, T. W., Whelan, R., … The IMAGEN Consortium. (2016). The structure of psychopathology in adolescence and its common personality and cognitive correlates. *Journal of Abnormal Psychology*, *125*(8), 1039-1052. https://doi.org/10.1037/abn0000193

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

Catania, E. H., Pimenta, A., & Levitt, P. (2008). Genetic deletion of Lsamp causes exaggerated behavioral activation in novel environments. *Behavioural Brain Research*, *188*(2), 380-390. https://doi.org/10.1016/j.bbr.2007.11.022

Centers for Disease Control and Prevention. (2013). Vital signs: Current cigarette smoking among adults aged ≥18 years with mental illness - United States, 2009-2011. *Morbidity and Mortality Weekly Report*, *62*(5), 81-87.

Chambers, R. A., Krystal, J. H., & Self, D. W. (2001). A neurobiological basis for substance abuse comorbidity in schizophrenia. *Biological Psychiatry*, *50*(2), 71-83. https://doi.org/10.1016/S0006-3223(01)01134-9

Chang, C. C., Chow, C. C., Tellier, L. C. A. M., Vattikuti, S., Purcell, S. M., & Lee, J. J. (2015). Second-generation PLINK: rising to the challenge of larger and richer datasets. *GigaScience*, *4*, 7. https://doi.org/10.1186/s13742-015-0047-8

Chang, L.-H., Whitfield, J. B., Liu, M., Medland, S. E., Hickie, I. B., Martin, N. G., Verhulst, B., Heath, A. C., Madden, P. A., Statham, D. J., & Gillespie, N. A. (2019). Associations between polygenic risk for tobacco and alcohol use and liability to tobacco and alcohol use, and psychiatric disorders in an independent sample of 13,999 Australian adults. *Drug and Alcohol Dependence*, *205*, 107704. https://doi.org/10.1016/j.drugalcdep.2019.107704

Charach, A., Yeung, E., Climans, T., & Lillie, E. (2011). Childhood Attention-Deficit/Hyperactivity Disorder and Future Substance Use Disorders: Comparative Meta-Analyses. *Journal of the American Academy of Child & Adolescent Psychiatry*, *50*(1), 9-21. https://doi.org/10.1016/j.jaac.2010.09.019

Chen, J., Bacanu, S.-A., Yu, H., Zhao, Z., Jia, P., Kendler, K. S., Kranzler, H. R., Gelernter, J., Farrer, L., Minica, C., Pool, R., Milaneschi, Y., Boomsma, D. I., Penninx, B. W. J. H., Tyndale, R. F., Ware, J. J., Vink, J. M., Kaprio, J., Munafò, M., & Chen, X. (2016). Genetic Relationship between Schizophrenia and Nicotine Dependence. *Scientific Reports*, *6*, 25671. https://doi.org/10.1038/srep25671

Chen, J., Chen, R., Xiang, S., Li, N., Gao, C., Wu, C., Zhang, Q., Zhao, Y., Liao, Y., Stewart, R., Xu, Y., Shi, Y., & Li, Z. (2021). Cigarette smoking and schizophrenia: Mendelian randomisation study. *The British Journal of Psychiatry*, *218*(2), 98-103. https://doi.org/10.1192/bjp.2020.116

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

Chen, J., Loukola, A., Gillespie, N. A., Peterson, R., Jia, P., Riley, B., Maes, H., Dick, D. M., Kendler, K. S., Damaj, M. I., Miles, M. F., Zhao, Z., Li, M. D., Vink, J. M., Minica, C. C., Willemsen, G., Boomsma, D. I., Qaiser, B., Madden, P. A. F., … Chen, X. (2020). Genome-Wide Meta-Analyses of FTND and TTFC Phenotypes. *Nicotine & Tobacco Research*, *22*(6), 900-909. https://doi.org/10.1093/ntr/ntz099

Chen, Y., Zhu, J., Lum, P. Y., Yang, X., Pinto, S., MacNeil, D. J., Zhang, C., Lamb, J., Edwards, S., Sieberts, S. K., Leonardson, A., Castellini, L. W., Wang, S., Champy, M. F., Zhang, B., Emilsson, V., Doss, S., Ghazalpour, A., Horvath, S., … Schadt, E. E. (2008). Variations in DNA elucidate molecular networks that cause disease. *Nature*, *452*(7186), 429-435. https://doi.org/10.1038/nature06757

Choi, S. W., García-González, J., Ruan, Y., Wu, H. M., Porras, C., Johnson, J., Hoggart, C. J., & OReilly, P. F. (2023). PRSet: Pathway-based polygenic risk score analyses and software. *PLoS Genetics*, *19*(2), 1-26. https://doi.org/10.1371/journal.pgen.1010624

Choi, S. W., & O'Reilly, P. F. (2019). PRSice-2: Polygenic Risk Score software for biobank-scale data. *GigaScience*, *8*(7), 1-6. https://doi.org/10.1093/gigascience/giz082

Clarke, T. K., Adams, M. J., Davies, G., Howard, D. M., Hall, L. S., Padmanabhan, S., Murray, A. D., Smith, B. H., Campbell, A., Hayward, C., Porteous, D. J., Deary, I. J., & McIntosh, A. M. (2017). Genome-wide association study of alcohol consumption and genetic overlap with other health-related traits in UK biobank (N=112117). *Molecular Psychiatry*, *22*(10), 1376-1384. https://doi.org/10.1038/mp.2017.153

Colder, C. R., Frndak, S., Lengua, L. J., Read, J. P., Hawk, L. W., & Wieczorek, W. F. (2018). Internalizing and Externalizing Problem Behavior: a Test of a Latent Variable Interaction Predicting a Two-Part Growth Model of Adolescent Substance Use. *Journal of Abnormal Child Psychology*, *46*(2), 319-330. https://doi.org/10.1007/s10802-017-0277-6

Connor, J. P., Gullo, M. J., White, A., & Kelly, A. B. (2014). Polysubstance use: Diagnostic challenges, patterns of use and health. *Current Opinion in Psychiatry*, *27*(4), 269-275. https://doi.org/10.1097/YCO.0000000000000069

Coustals, N., Martelli, C., Brunet-Lecomte, M., Petillion, A., Romeo, B., & Benyamina, A. (2020). Chronic smoking and cognition in patients with schizophrenia: A meta-analysis. *Schizophrenia Research*, *222*, 113-121. https://doi.org/10.1016/j.schres.2020.03.071

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

Creamer, M. R., Wang, T. W., Babb, S., Cullen, K. A., Day, H., Willis, G., Jamal, A., & Neff, L. (2019). Tobacco Product Use and Cessation Indicators Among Adults — United States, 2018. *MMWR. Morbidity and Mortality Weekly Report*, *68*(45), 1013-1019. https://doi.org/10.15585/mmwr.mm6845a2

D'Souza, M. S., & Markou, A. (2012). Schizophrenia and tobacco smoking comorbidity: NAChR agonists in the treatment of schizophrenia-associated cognitive deficits. *Neuropharmacology*, *62*(3), 1564-1573. https://doi.org/10.1016/j.neuropharm.2011.01.044

Dark, C., Williams, C., Bellgrove, M. A., Hawi, Z., & Bryson-Richardson, R. J. (2020). Functional validation of CHMP7 as an ADHD risk gene. *Translational Psychiatry*, *10*(1), 385. https://doi.org/10.1038/s41398-020-01077-w

Das, S., Forer, L., Schönherr, S., Sidore, C., Locke, A. E., Kwong, A., Vrieze, S. I., Chew, E. Y., Levy, S., McGue, M., Schlessinger, D., Stambolian, D., Loh, P.-R., Iacono, W. G., Swaroop, A., Scott, L. J., Cucca, F., Kronenberg, F., Boehnke, M., … Fuchsberger, C. (2016). Next-generation genotype imputation service and methods. *Nature Genetics*, *48*(10), 1284-1287. https://doi.org/10.1038/ng.3656

Davies, N. M., Holmes, M. V., & Davey Smith, G. (2018). Reading Mendelian randomisation studies: a guide, glossary, and checklist for clinicians. *BMJ*, *362*, k601. https://doi.org/10.1136/bmj.k601

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

de Leeuw, C. A., Mooij, J. M., Heskes, T., & Posthuma, D. (2015). MAGMA: Generalized Gene-Set Analysis of GWAS Data. *PLoS Computational Biology*, *11*(4), 1-19. https://doi.org/10.1371/journal.pcbi.1004219

de Leon, J., & Diaz, F. J. (2005). A meta-analysis of worldwide studies demonstrates an association between schizophrenia and tobacco smoking behaviors. *Schizophrenia Research*, *76*(2-3), 135-157. https://doi.org/10.1016/j.schres.2005.02.010

USC DE COMPOSTELA

de Leon, J., Tracy, J., McCann, E., McGrory, A., & Diaz, F. J. (2002). Schizophrenia and tobacco smoking: A replication study in another US psychiatric hospital. *Schizophrenia*

Research, *56*(1-2), 55-65. https://doi.org/10.1016/S0920-9964(01)00192-X

de Moor, M. H. M., van den Berg, S. M., Verweij, K. J. H., Krueger, R. F., Luciano, M., Arias Vasquez, A., Matteson, L. K., Derringer, J., Esko, T., Amin, N., Gordon, S. D., Hansell, N. K., Hart, A. B., Seppälä, I., Huffman, J. E., Konte, B., Lahti, J., Lee, M., Miller, M., … Boomsma, D. I. (2015). Meta-analysis of Genome-wide Association Studies for Neuroticism, and the Polygenic Association With Major Depressive Disorder. *JAMA Psychiatry*, *72*(7), 642. https://doi.org/10.1001/jamapsychiatry.2015.0554

Deak, J. D., Levey, D. F., Wendt, F. R., Zhou, H., Galimberti, M., Kranzler, H. R., Gaziano, J. M., Stein, M. B., Polimanti, R., Gelernter, J., Muralidhar, S., Moser, J., Deen, J. E., Gaziano, J. M., Beckham, J., Chang, K.-M., Tsao, P. S., Luoh, S.-W., Casas, J. P., … Ramoni, R. B. (2022). Genome-Wide Investigation of Maximum Habitual Alcohol Intake in US Veterans in Relation to Alcohol Consumption Traits and Alcohol Use Disorder. *JAMA Network Open*, *5*(10), e2238880. https://doi.org/10.1001/jamanetworkopen.2022.38880

Decina, P., Caracci, G., Sandik, R., Berman, W., Mukherjee, S., & Scapicchio, P. (1990). Cigarette Smoking and Neuroleptic-Induced Parkinsonism. *Biological Psychiatry*, *28*, 502-508.

Degenhardt, L., Charlson, F., Ferrari, A., Santomauro, D., Erskine, H., Mantilla-Herrara, A., Whiteford, H., Leung, J., Naghavi, M., Griswold, M., Rehm, J., Hall, W., Sartorius, B., Scott, J., Vollset, S. E., Knudsen, A. K., Haro, J. M., Patton, G., Kopec, J., … Vos, T. (2018). The global burden of disease attributable to alcohol and drug use in 195 countries and territories, 1990–2016: a systematic analysis for the Global Burden of Disease Study 2016. *The Lancet Psychiatry*, *5*(12), 987-1012. https://doi.org/10.1016/S2215-0366(18)30337-7

Degenhardt, L., Glantz, M., Bharat, C., Peacock, A., Lago, L., Sampson, N., & Kessler, R. C. (2018). The impact of cohort substance use upon likelihood of transitioning through stages of alcohol and cannabis use and use disorder: Findings from the Australian National Survey on Mental Health and Wellbeing. *Drug and Alcohol Review*, *37*(4), 546-556. https://doi.org/10.1111/dar.12679

Degenhardt, L., Saha, S., Lim, C. C. W., Aguilar-Gaxiola, S., Al-Hamzawi, A., Alonso, J., Andrade, L. H., Bromet, E. J., Bruffaerts, R., Caldas-de-Almeida, J. M., de Girolamo, G.,

Florescu, S., Gureje, O., Haro, J. M., Karam, E. G., Karam, G., Kovess-Masfety, V., Lee, S., Lepine, J. P., … McGrath, J. J. (2018). The associations between psychotic experiences and substance use and substance use disorders: findings from the World Health Organization World Mental Health surveys. *Addiction*, *113*(5), 924-934. https://doi.org/10.1111/add.14145

Dehghan, A. (2018). Genome-Wide Association Studies. *Methods in Molecular Biology*, *1793*, 37-49. https://doi.org/10.1007/978-1-4939-7868-7_4

Demin, K. A., Meshalkina, D. A., Kysil, E. V., Antonova, K. A., Volgin, A. D., Yakovlev, O. A., Alekseeva, P. A., Firuleva, M. M., Lakstygal, A. M., de Abreu, M. S., Barcellos, L. J. G., Bao, W., Friend, A. J., Amstislavskaya, T. G., Rosemberg, D. B., Musienko, P. E., Song, C., & Kalueff, A. V. (2018). Zebrafish models relevant to studying central opioid and endocannabinoid systems. *Progress in Neuro-Psychopharmacology and Biological Psychiatry*, *86*, 301-312. https://doi.org/10.1016/j.pnpbp.2018.03.024

Demontis, D., Rajagopal, V. M., Thorgeirsson, T. E., Als, T. D., Grove, J., Leppälä, K., Gudbjartsson, D. F., Pallesen, J., Hjorthøj, C., Reginsson, G. W., Tyrfingsson, T., Runarsdottir, V., Qvist, P., Christensen, J. H., Bybjerg-Grauholm, J., Bækvad-Hansen, M., Huckins, L. M., Stahl, E. A., Timmermann, A., … Børglum, A. D. (2019). Genome-wide association study implicates CHRNA2 in cannabis use disorder. *Nature Neuroscience*, *22*(7), 1066-1074. https://doi.org/10.1038/s41593-019-0416-1

Demontis, D., Walters, G. B., Athanasiadis, G., Walters, R., Therrien, K., Nielsen, T. T., Farajzadeh, L., Voloudakis, G., Bendl, J., Zeng, B., Zhang, W., Grove, J., Als, T. D., Duan, J., Satterstrom, F. K., Bybjerg-Grauholm, J., Bækved-Hansen, M., Gudmundsson, O. O., Magnusson, S. H., … Børglum, A. D. (2023). Genome-wide analyses of ADHD identify 27 risk loci, refine the genetic architecture and implicate several cognitive domains. *Nature Genetics*, *55*(2), 198-208. https://doi.org/10.1038/s41588-022-01285-8

Demontis, D., Walters, R. K., Martin, J., Mattheisen, M., Als, T. D., Agerbo, E., Baldursson, G., Belliveau, R., Bybjerg-Grauholm, J., Bækvad-Hansen, M., Cerrato, F., Chambert, K., Churchhouse, C., Dumont, A., Eriksson, N., Gandal, M., Goldstein, J. I., Grasby, K. L., Grove, J., … 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

Dépatie, L., O'Driscoll, G. A., Holahan, A. L. V., Atkinson, V., Thavundayil, J. X., Kin, N. N.

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

Y., & Lal, S. (2002). Nicotine and behavioral markers of risk for schizophrenia: A double-blind, placebo-controlled, cross-over study. *Neuropsychopharmacology*, *27*(6), 1056-1070. https://doi.org/10.1016/S0893-133X(02)00372-X

Dervaux, A., Baylé, F. J., Laqueille, X., Bourdel, M., Le Borgne, M.-H., Olié, J., & Krebs, M. (2001). Is Substance Abuse in Schizophrenia Related to Impulsivity, Sensation Seeking, or Anhedonia? *American Journal of Psychiatry*, *158*(3), 492-494. https://doi.org/10.1176/appi.ajp.158.3.492

Desai, H. D., Seabolt, J., & Jann, M. W. (2001). Smoking in Patients Receiving Psychotropic Medications. *CNS Drugs*, *15*(6), 469-494. https://doi.org/10.2165/00023210-200115060-00005

Deutsch, A. R., & Selya, A. S. (2020). Stability in effects of different smoking-related polygenic risk scores over age and smoking phenotypes. *Drug and Alcohol Dependence*, *214*, 108154. https://doi.org/10.1016/j.drugalcdep.2020.108154

Di Forti, M., Quattrone, D., Freeman, T. P., Tripoli, G., Gayer-Anderson, C., Quigley, H., Rodriguez, V., Jongsma, H. E., Ferraro, L., La Cascia, C., La Barbera, D., Tarricone, I., Berardi, D., Szöke, A., Arango, C., Tortelli, A., Velthorst, E., Bernardo, M., Del-Ben, C. M., … van der Ven, E. (2019). The contribution of cannabis use to variation in the incidence of psychotic disorder across Europe (EU-GEI): a multicentre case-control study. *The Lancet Psychiatry*, *6*(5), 427-436. https://doi.org/10.1016/S2215-0366(19)30048-3

Dick, D. M. (2016). The Genetics of Addiction: Where Do We Go From Here? *Journal of Studies on Alcohol and Drugs*, *77*(5), 673-675. https://doi.org/10.15288/jsad.2016.77.673

Dickerson, F., Origoni, A., Schroeder, J., Adamos, M., Katsafanas, E., Khushalani, S., Savage, C. L. G., Schweinfurth, L. A. B., Stallings, C., Sweeney, K., & Yolken, R. (2018). Natural cause mortality in persons with serious mental illness. *Acta Psychiatrica Scandinavica*, *137*(5), 371-379. https://doi.org/10.1111/acps.12880

Dickerson, F., Schroeder, J., Katsafanas, E., Khushalani, S., Origoni, A. E., Savage, C., Schweinfurth, L., Stallings, C. R., Sweeney, K., & Yolken, R. H. (2018). Cigarette smoking by patients with serious mental illness, 1999-2016: An increasing disparity. *Psychiatric Services*, *69*(2), 147-153. https://doi.org/10.1176/appi.ps.201700118

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

Dir, A. L., & Hulvershorn, L. A. (2019). Sex Differences in Adolescent Neurobiological Risk

for Substance Use and Substance Use Disorders. *Current Addiction Reports*, *6*(4), 514-521. https://doi.org/10.1007/s40429-019-00276-w

Drew, R. E., Rodnick, K. J., Settles, M., Wacyk, J., Churchill, E., Powell, M. S., Hardy, R. W., Murdoch, G. K., Hill, R. A., & Robison, B. D. (2008). Effect of starvation on transcriptomes of brain and liver in adult female zebrafish (Danio rerio). *Physiological Genomics*, *35*(3), 283-295. https://doi.org/10.1152/physiolgenomics.90213.2008

Du Rietz, E., Coleman, J., Glanville, K., Choi, S. W., O'Reilly, P. F., & Kuntsi, J. (2018). Association of Polygenic Risk for Attention-Deficit/Hyperactivity Disorder With Co-occurring Traits and Disorders. *Biological Psychiatry: Cognitive Neuroscience and Neuroimaging*, *3*(7), 635-643. https://doi.org/10.1016/j.bpsc.2017.11.013

Edenberg, H. J. (2007). The genetics of alcohol metabolism: Role of alcohol dehydrogenase and aldehyde dehydrogenase variants. *Alcohol Research and Health*, *30*(1), 5-13.

Edwards, A. C., Gardner, C. O., Hickman, M., & Kendler, K. S. (2016). A prospective longitudinal model predicting early adult alcohol problems: evidence for a robust externalizing pathway. *Psychological Medicine*, *46*(5), 957-968. https://doi.org/10.1017/S0033291715002457

Eising, E., Carrion-Castillo, A., Vino, A., Strand, E. A., Jakielski, K. J., Scerri, T. S., Hildebrand, M. S., Webster, R., Ma, A., Mazoyer, B., Francks, C., Bahlo, M., Scheffer, I. E., Morgan, A. T., Shriberg, L. D., & Fisher, S. E. (2019). A set of regulatory genes co-expressed in embryonic human brain is implicated in disrupted speech development. *Molecular Psychiatry*, *24*(7), 1065-1078. https://doi.org/10.1038/s41380-018-0020-x

Facal, F., Flórez, G., Blanco, V., Rodríguez, J., Pereiro, C., Fernández, J. M., Fariñas, E., Estévez, V., Gómez-Trigo, J., Gurriarán, X., Sáiz, P., Vázquez, F. L., Arrojo, M., Costas, J., & the GenPol Study Group. (2021). Genetic predisposition to alcohol dependence: The combined role of polygenic risk to general psychopathology and to high alcohol consumption. *Drug and Alcohol Dependence*, *221*, 108556. https://doi.org/10.1016/j.drugalcdep.2021.108556

Fergusson, D. M., Hall, W., Boden, J. M., & Horwood, L. J. (2015). Rethinking cigarette smoking, cannabis use, and psychosis. *The Lancet Psychiatry*, *2*(7), 581-582. https://doi.org/10.1016/S2215-0366(15)00208-4

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

Firth, J., Solmi, M., Wootton, R. E., Vancampfort, D., Schuch, F. B., Hoare, E., Gilbody, S., Torous, J., Teasdale, S. B., Jackson, S. E., Smith, L., Eaton, M., Jacka, F. N., Veronese, N., Marx, W., Ashdown-Franks, G., Siskind, D., Sarris, J., Rosenbaum, S., … Stubbs, B. (2020). A meta-review of "lifestyle psychiatry": the role of exercise, smoking, diet and sleep in the prevention and treatment of mental disorders. *World Psychiatry*, *19*(3), 360-380. https://doi.org/10.1002/wps.20773

Firth, J., Wootton, R. E., Sawyer, C., & Taylor, G. M. (2023). Clearing the air: clarifying the causal role of smoking in mental illness. *World Psychiatry*, *22*(1), 151-152. https://doi.org/10.1002/wps.21023

Fischer, B. A., McMahon, R. P., Kelly, D. L., Wehring, H. J., Meyer, W. A., Feldman, S., Carpenter, W. T., & Gorelick, D. A. (2015). Risk-taking in schizophrenia and controls with and without cannabis dependence. *Schizophrenia Research*, *161*(2-3), 471-477. https://doi.org/10.1016/j.schres.2014.11.009

Fluharty, M. E., Sallis, H., & Munafò, M. R. (2018). Investigating possible causal effects of externalizing behaviors on tobacco initiation: A Mendelian randomization analysis. *Drug and Alcohol Dependence*, *191*, 338-342. https://doi.org/10.1016/j.drugalcdep.2018.07.015

Fluharty, M., Taylor, A. E., Grabski, M., & Munafò, M. R. (2017). The association of cigarette smoking with depression and anxiety: A systematic review. *Nicotine and Tobacco Research*, *19*(1), 3-13. https://doi.org/10.1093/ntr/ntw140

Fontana, B. D., Mezzomo, N. J., Kalueff, A. V., & Rosemberg, D. B. (2018). The developing utility of zebrafish models of neurological and neuropsychiatric disorders: A critical review. *Experimental neurology*, *299*(Pt A), 157-171. https://doi.org/10.1016/j.expneurol.2017.10.004

Franklin, C., & Dwyer, D. S. (2021). Candidate risk genes for bipolar disorder are highly conserved during evolution and highly interconnected. *Bipolar Disorders*, *23*(4), 400-408. https://doi.org/10.1111/bdi.12996

Frei, O., Holland, D., Smeland, O. B., Shadrin, A. A., Fan, C. C., Maeland, S., O'Connell, K. S., Wang, Y., Djurovic, S., Thompson, W. K., Andreassen, O. A., & Dale, A. M. (2019). Bivariate causal mixture model quantifies polygenic overlap between complex traits beyond genetic correlation. *Nature Communications*, *10*(1), 1-11. https://doi.org/10.1038/s41467-019-10310-0

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

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., Klei, L. L., Kramer, R., Pinto, D., Gümüş, Z. H., Cicek, A. E., Dang, K. K., Browne, A., Lu, C., Xie, L., … Sklar, P. (2016). Gene expression elucidates functional impact of polygenic risk for schizophrenia. *Nature Neuroscience*, *19*(11), 1442-1453. https://doi.org/10.1038/nn.4399

Fry, A., Littlejohns, T. J., Sudlow, C., Doherty, N., Adamska, L., Sprosen, T., Collins, R., & Allen, N. E. (2017). Comparison of Sociodemographic and Health-Related Characteristics of UK Biobank Participants with Those of the General Population. *American Journal of Epidemiology*, *186*(9), 1026-1034. https://doi.org/10.1093/aje/kwx246

Gage, S. H., Davey Smith, G., Ware, J. J., Flint, J., & Munafò, M. R. (2016). G = E: What GWAS Can Tell Us about the Environment. *PLOS Genetics*, *12*(2), e1005765. https://doi.org/10.1371/journal.pgen.1005765

Gage, S. H., Hickman, M., Heron, J., Munafò, M. R., Lewis, G., Macleod, J., & Zammit, S. (2014). Associations of cannabis and cigarette use with psychotic experiences at age 18: Findings from the Avon Longitudinal Study of Parents and Children. *Psychological Medicine*, *44*(16), 3435-3444. https://doi.org/10.1017/S0033291714000531

Gage, S. H., Hickman, M., & Zammit, S. (2016). Association between cannabis and psychosis: Epidemiologic evidence. *Biological Psychiatry*, *79*(7), 549-556. https://doi.org/10.1016/j.biopsych.2015.08.001

Gage, S. H., Jones, H. J., Taylor, A. E., Burgess, S., Zammit, S., & Munafò, M. R. (2017). Investigating causality in associations between smoking initiation and schizophrenia using Mendelian randomization. *Scientific reports*, *7*, 40653. https://doi.org/10.1038/srep40653

Gage, S. H., & Munafò, M. R. (2015a). Rethinking the association between smoking and schizophrenia. *The Lancet Psychiatry*, *2*(2), 118-119. https://doi.org/10.1016/S2215-0366(14)00057-1

Gage, S. H., & Munafò, M. R. (2015b). Smoking as a causal risk factor for schizophrenia. *The lancet. Psychiatry*, *2*(9), 778-779. https://doi.org/10.1016/S2215-0366(15)00333-8

Gaiteri, C., Ding, Y., French, B., Tseng, G. C., & Sibille, E. (2014). Beyond modules and hubs: The potential of gene coexpression networks for investigating molecular mechanisms of complex brain disorders. *Genes, Brain and Behavior*, *13*(1), 13-24.

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

https://doi.org/10.1111/gbb.12106

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. (2015). A gene-based association method for mapping traits using reference transcriptome data. *Nature Genetics*, *47*(9), 1091-1098. https://doi.org/10.1038/ng.3367

Gard, A. M., Ware, E. B., Hyde, L. W., Schmitz, L. L., Faul, J., & Mitchell, C. (2021). Phenotypic and genetic markers of psychopathology in a population-based sample of older adults. *Translational psychiatry*, *11*(1), 239. https://doi.org/10.1038/s41398-021-01354-2

Gattamorta, K. A., Mena, M. P., Ainsley, J. B., & Santisteban, D. A. (2017). The Comorbidity of Psychiatric and Substance Use Disorders Among Hispanic Adolescents. *Journal of Dual Diagnosis*, *13*(4), 254-263. https://doi.org/10.1080/15504263.2017.1343965

Gaziano, J. M., Concato, J., Brophy, M., Fiore, L., Pyarajan, S., Breeling, J., Whitbourne, S., Deen, J., Shannon, C., Humphries, D., Guarino, P., Aslan, M., Anderson, D., Lafleur, R., Hammond, T., Schaa, K., Moser, J., Huang, G., Muralidhar, S., … O'Leary, T. J. (2016). Million Veteran Program: A mega-biobank to study genetic influences on health and disease. *Journal of Clinical Epidemiology*, *70*, 214-223. https://doi.org/10.1016/j.jclinepi.2015.09.016

Ge, T., Chen, C.-Y., Ni, Y., Feng, Y.-C. A., & Smoller, J. W. (2019). Polygenic prediction via Bayesian regression and continuous shrinkage priors. *Nature Communications*, *10*, 1776. https://doi.org/10.1038/s41467-019-09718-5

Gelernter, J., & Polimanti, R. (2021). Genetics of substance use disorders in the era of big data. *Nature Reviews Genetics*, *22*(11), 712-729. https://doi.org/10.1038/s41576-021-00377-1

George, T. P., Vessicchio, J. C., Termine, A., Sahady, D. M., Head, C. A., Pepper, W. T., Kosten, T. R., & Wexler, B. E. (2002). Effects of smoking abstinence on visuospatial working memory function in schizophrenia. *Neuropsychopharmacology*, *26*(1), 75-85. https://doi.org/10.1016/S0893-133X(01)00296-2

Gerring, Z. F., Thorp, J. G., Gamazon, E. R., & Derks, E. M. (2022). An analysis of genetically regulated gene expression and the role of co-expression networks across 16 psychiatric and substance use phenotypes. *European Journal of Human Genetics*, *30*(5), 560-566. https://doi.org/10.1038/s41431-022-01037-6

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

Gerring, Z. F., Thorp, J. G., Treur, J. L., Verweij, K. J. H., & Derks, E. M. (2024). The genetic landscape of substance use disorders. *Molecular psychiatry*, *March*. https://doi.org/10.1038/s41380-024-02547-z

Ghoussaini, M., Mountjoy, E., Carmona, M., Peat, G., Schmidt, E. M., Hercules, A., Fumis, L., Miranda, A., Carvalho-Silva, D., Buniello, A., Burdett, T., Hayhurst, J., Baker, J., Ferrer, J., Gonzalez-Uriarte, A., Jupp, S., Karim, M. A., Koscielny, G., MacHlitt-Northen, S., … Dunham, I. (2021). Open Targets Genetics: Systematic identification of trait-associated genes using large-scale genetics and functional genomics. *Nucleic Acids Research*, *49*(D1), D1311-D1320. https://doi.org/10.1093/nar/gkaa840

Giambartolomei, C., Vukcevic, D., Schadt, E. E., Franke, L., Hingorani, A. D., Wallace, C., & Plagnol, V. (2014). Bayesian Test for Colocalisation between Pairs of Genetic Association Studies Using Summary Statistics. *PLoS Genetics*, *10*(5), e1004383. https://doi.org/10.1371/journal.pgen.1004383

Goff, D. C., Henderson, D. C., & Amico, E. (1992). Cigarette smoking in schizophrenia: relationship to psychopathology and medication side effects. *American Journal of Psychiatry*, *149*(9), 1189-1194. https://doi.org/10.1176/ajp.149.9.1189

Gogos, A., Skokou, M., Ferentinou, E., & Gourzis, P. (2019). Nicotine consumption during the prodromal phase of schizophrenia – a review of the literature. *Neuropsychiatric Disease and Treatment*, *15*, 2943-2958. https://doi.org/10.2147/NDT.S210199

Goldman, D., Oroszi, G., & Ducci, F. (2005). The genetics of addictions: Uncovering the genes. *Nature Reviews Genetics*, *6*(7), 521-532. https://doi.org/10.1038/nrg1635

Goldstein, R. Z., & Volkow, N. D. (2002). Drug addiction and its underlying neurobiological basis: Neuroimaging evidence for the involvement of the frontal cortex. *American Journal of Psychiatry*, *159*(10), 1642-1652. https://doi.org/10.1176/appi.ajp.159.10.1642

Gotti, C., Clementi, F., Fornari, A., Gaimarri, A., Guiducci, S., Manfredi, I., Moretti, M., Pedrazzi, P., Pucci, L., & Zoli, M. (2009). Structural and functional diversity of native brain neuronal nicotinic receptors. *Biochemical Pharmacology*, *78*(7), 703-711. https://doi.org/10.1016/j.bcp.2009.05.024

Green, V. R., Conway, K. P., Silveira, M. L., Kasza, K. A., Cohn, A., Cummings, K. M., Stanton, C. A., Callahan-Lyon, P., Slavit, W., Sargent, J. D., Hilmi, N., Niaura, R. S.,

USC DE COMPOSTELA

Reissig, C. J., Lambert, E., Zandberg, I., Brunette, M. F., Tanski, S. E., Borek, N., Hyland, A. J., & Compton, W. M. (2018). Mental Health Problems and Onset of Tobacco Use Among 12- to 24-Year-Olds in the PATH Study. *Journal of the American Academy of Child & Adolescent Psychiatry*, *57*(12), 944-954. https://doi.org/10.1016/j.jaac.2018.06.029

Griffith, J. W., Zinbarg, R. E., Craske, M. G., Mineka, S., Rose, R. D., Waters, A. M., & Sutton, J. M. (2010). Neuroticism as a common dimension in the internalizing disorders. *Psychological Medicine*, *40*(7), 1125-1136. https://doi.org/10.1017/S0033291709991449

Grotzinger, A. D., Mallard, T. T., Akingbuwa, W. A., Ip, H. F., Adams, M. J., Lewis, C. M., McIntosh, A. M., Grove, J., Dalsgaard, S., Lesch, K. P., Strom, N., Meier, S. M., Mattheisen, M., Børglum, A. D., Mors, O., Breen, G., Mors, O., Lee, P. H., Kendler, K. S., … Nivard, M. G. (2022). Genetic architecture of 11 major psychiatric disorders at biobehavioral, functional genomic and molecular genetic levels of analysis. *Nature Genetics*, *54*(5), 548-559. https://doi.org/10.1038/s41588-022-01057-4

Grotzinger, A. D., Rhemtulla, M., de Vlaming, R., Ritchie, S. J., Mallard, T. T., Hill, W. D., Ip, H. F., Marioni, R. E., McIntosh, A. M., Deary, I. J., Koellinger, P. D., Harden, K. P., Nivard, M. G., & Tucker-Drob, E. M. (2019). Genomic structural equation modelling provides insights into the multivariate genetic architecture of complex traits. *Nature Human Behaviour*, *3*(5), 513-525. https://doi.org/10.1038/s41562-019-0566-x

Grove, J., Ripke, S., Als, T. D., Mattheisen, M., Walters, R. K., Won, H., Pallesen, J., Agerbo, E., Andreassen, O. A., Anney, R., Awashti, S., Belliveau, R., Bettella, F., Buxbaum, J. D., Bybjerg-Grauholm, J., Bækvad-Hansen, M., Cerrato, F., Chambert, K., Christensen, J. 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

GTEx Consortium. (2020). The GTEx Consortium atlas of genetic regulatory effects across human tissues. *Science*, *369*(6509), 1318-1330. https://doi.org/10.1126/science.aaz1776

Gurillo, P., Jauhar, S., Murray, R. M., & MacCabe, J. H. (2015). Does tobacco use cause psychosis? Systematic review and meta-analysis. *The Lancet Psychiatry*, *2*(8), 718-725. https://doi.org/10.1016/S2215-0366(15)00152-2

Gurriarán, X., Rodríguez-López, J., Flórez, G., Pereiro, C., Fernández, J. M., Fariñas, E.,

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

Estévez, V., Arrojo, M., & Costas, J. (2019). Relationships between substance abuse/dependence and psychiatric disorders based on polygenic scores. *Genes, Brain and Behavior*, *18*(3), e12504. https://doi.org/10.1111/gbb.12504

Gusev, A., Ko, A., Shi, H., Bhatia, G., Chung, W., Penninx, B. W. J. H., Jansen, R., De Geus, E. J. C., Boomsma, D. I., Wright, F. A., Sullivan, P. F., Nikkola, E., Alvarez, M., Civelek, M., Lusis, A. J., Lehtimäki, T., Raitoharju, E., Kähönen, M., Seppälä, I., … Pasaniuc, B. (2016). Integrative approaches for large-scale transcriptome-wide association studies. *Nature Genetics*, *48*(3), 245-252. https://doi.org/10.1038/ng.3506

Gut-Fayand, A., Dervaux, A., Olié, J. P., Lôo, H., Poirier, M. F., & Krebs, M. O. (2001). Substance abuse and suicidality in schizophrenia: A common risk factor linked to impulsivity. *Psychiatry Research*, *102*(1), 65-72. https://doi.org/10.1016/S0165-1781(01)00250-5

Haberstick, B. C., Zeiger, J. S., Corley, R. P., Hopfer, C. J., Stallings, M. C., Rhee, S. H., & Hewitt, J. K. (2011). Common and drug-specific genetic influences on subjective effects to alcohol, tobacco and marijuana use. *Addiction*, *106*(1), 215-224. https://doi.org/10.1111/j.1360-0443.2010.03129.x

Hakulinen, C., Hintsanen, M., Munafò, M. R., Virtanen, M., Kivimäki, M., Batty, G. D., & Jokela, M. (2015). Personality and smoking: Individual-participant meta-analysis of nine cohort studies. *Addiction*, *110*(11), 1844-1852. https://doi.org/10.1111/add.13079

Hall, S. M., & Prochaska, J. J. (2009). Treatment of Smokers with Co-Occurring Disorders: Emphasis on Integration in Mental Health and Addiction Treatment Settings. *Annual Review of Clinical Psychology*, *5*(1), 409-431. https://doi.org/10.1146/annurev.clinpsy.032408.153614

Hällfors, J., Palviainen, T., Surakka, I., Gupta, R., Buchwald, J., Raevuori, A., Ripatti, S., Korhonen, T., Jousilahti, P., Madden, P. A. F., Kaprio, J., & Loukola, A. (2019). Genome-wide association study in Finnish twins highlights the connection between nicotine addiction and neurotrophin signaling pathway. *Addiction Biology*, *24*(3), 549-561. https://doi.org/10.1111/adb.12618

Hancock, D. B., Reginsson, G. W., Gaddis, N. C., Chen, X., Saccone, N. L., Lutz, S. M., Qaiser, B., Sherva, R., Steinberg, S., Zink, F., Stacey, S. N., Glasheen, C., Chen, J., Gu, F., Frederiksen, B. N., Loukola, A., Gudbjartsson, D. F., Brüske, I., Landi, M. T., …

Stefansson, K. (2015). Genome-wide meta-analysis reveals common splice site acceptor variant in CHRNA4 associated with nicotine dependence. *Translational Psychiatry*, *5*(10), e651. https://doi.org/10.1038/tp.2015.149

Hartwig, F. P., Davies, N. M., & Davey Smith, G. (2018). Bias in Mendelian randomization due to assortative mating. *Genetic Epidemiology*, *42*(7), 608-620. https://doi.org/10.1002/gepi.22138

Hartz, S. M., Horton, A. C., Hancock, D. B., Baker, T. B., Caporaso, N. E., Chen, L.-S., Hokanson, J. E., Lutz, S. M., Marazita, M. L., McNeil, D. W., Pato, C. N., Pato, M. T., Johnson, E. O., & Bierut, L. J. (2018). Genetic correlation between smoking behaviors and schizophrenia. *Schizophrenia Research*, *194*, 86-90. https://doi.org/10.1016/j.schres.2017.02.022

Hartz, S. M., Horton, A. C., Oehlert, M., Carey, C. E., Agrawal, A., Bogdan, R., Chen, L. S., Hancock, D. B., Johnson, E. O., Pato, C. N., Pato, M. T., Rice, J. P., & Bierut, L. J. (2017). Association Between Substance Use Disorder and Polygenic Liability to Schizophrenia. *Biological Psychiatry*, *82*(10), 709-715. https://doi.org/10.1016/j.biopsych.2017.04.020

Hartz, S. M., Pato, C. N., Medeiros, H., Cavazos-Rehg, P., Sobell, J. L., Knowles, J. A., Bierut, L. J., & Pato, M. T. (2014). Comorbidity of severe psychotic disorders with measures of substance use. *JAMA Psychiatry*, *71*(3), 248-254. https://doi.org/10.1001/jamapsychiatry.2013.3726

Hasin, D., & Kilcoyne, B. (2012). Comorbidity of psychiatric and substance use disorders in the United States: Current issues and findings from the NESARC. *Current Opinion in Psychiatry*, *25*(3), 165-171. https://doi.org/10.1097/YCO.0b013e3283523dcc

Hasin, D. S., & Grant, B. F. (2015). The National Epidemiologic Survey on Alcohol and Related Conditions (NESARC) Waves 1 and 2: review and summary of findings. *Social Psychiatry and Psychiatric Epidemiology*, *50*(11), 1609-1640. https://doi.org/10.1007/s00127-015-1088-0

Hasin, D. S., Stinson, F. S., Ogburn, E., & Grant, B. F. (2007). Prevalence, Correlates, Disability, and Comorbidity of DSM-IV Alcohol Abuse and Dependence in the United States: Results From the National Epidemiologic Survey on Alcohol and Related Conditions. *Archives of General Psychiatry*, *64*(7), 830-842. https://doi.org/10.1001/archpsyc.64.7.830

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

Hatoum, A. S., Colbert, S. M. C., Johnson, E. C., Huggett, S. B., Deak, J. D., Pathak, G. A., Jennings, M. V., Paul, S. E., Karcher, N. R., Hansen, I., Baranger, D. A. A., Edwards, A., Grotzinger, A. D., Adkins, D. E., Adkins, A. E., Alanne-Kinnunen, M., Alexander, J. C., Aliev, F., Bacanu, S.-A., … Agrawal, A. (2023). Multivariate genome-wide association meta-analysis of over 1 million subjects identifies loci underlying multiple substance use disorders. *Nature Mental Health*, *1*(3), 210-223. https://doi.org/10.1038/s44220-023-00034-y

Hatoum, A. S., Johnson, E. C., Colbert, S. M. C., Polimanti, R., Zhou, H., Walters, R. K., Gelernter, J., Edenberg, H. J., Bogdan, R., & Agrawal, A. (2022). The addiction risk factor: A unitary genetic vulnerability characterizes substance use disorders and their associations with common correlates. *Neuropsychopharmacology*, *47*(10), 1739-1745. https://doi.org/10.1038/s41386-021-01209-w

Hawrylycz, M. J., Lein, E. S., Guillozet-Bongaarts, A. L., Shen, E. H., Ng, L., Miller, J. A., Van De Lagemaat, L. N., Smith, K. A., Ebbert, A., Riley, Z. L., Abajian, C., Beckmann, C. F., Bernard, A., Bertagnolli, D., Boe, A. F., Cartagena, P. M., Mallar Chakravarty, M., Chapin, M., Chong, J., … Jones, A. R. (2012). An anatomically comprehensive atlas of the adult human brain transcriptome. *Nature*, *489*(7416), 391-399. https://doi.org/10.1038/nature11405

He, F., & Maslov, S. (2016). Pan- and core- network analysis of co-expression genes in a model plant. *Scientific Reports*, *6*, 38956. https://doi.org/10.1038/srep38956

He, X., Fuller, C. K., Song, Y., Meng, Q., Zhang, B., Yang, X., & Li, H. (2013). Sherlock: Detecting Gene-Disease Associations by Matching Patterns of Expression QTL and GWAS. *American Journal of Human Genetics*, *92*(5), 667-680. https://doi.org/10.1016/j.ajhg.2013.03.022

Heerey, E. A., Robinson, B. M., McMahon, R. P., & Gold, J. M. (2007). Delay discounting in schizophrenia. *Cognitive Neuropsychiatry*, *12*(3), 213-221. https://doi.org/10.1080/13546800601005900

Heitzeg, M. M., Hardee, J. E., & Beltz, A. M. (2018). Sex differences in the developmental neuroscience of adolescent substance use risk. *Current Opinion in Behavioral Sciences*, *23*, 21-26. https://doi.org/10.1016/j.cobeha.2018.01.020

Hemani, G., Bowden, J., & Davey Smith, G. (2018). Evaluating the potential role of pleiotropy

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

in Mendelian randomization studies. *Human molecular genetics*, *27*(R2), R195-R208. https://doi.org/10.1093/hmg/ddy163

Hill, W. D., Davies, N. M., Ritchie, S. J., Skene, N. G., Bryois, J., Bell, S., Di Angelantonio, E., Roberts, D. J., Xueyi, S., Davies, G., Liewald, D. C. M., Porteous, D. J., Hayward, C., Butterworth, A. S., McIntosh, A. M., Gale, C. R., & Deary, I. J. (2019). Genome-wide analysis identifies molecular systems and 149 genetic loci associated with income. *Nature Communications*, *10*, 5741. https://doi.org/10.1038/s41467-019-13585-5

Hill, W. D., Hagenaars, S. P., Marioni, R. E., Harris, S. E., Liewald, D. C. M., Davies, G., Okbay, A., McIntosh, A. M., Gale, C. R., & 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

Hindley, G., Bahrami, S., Steen, N. E., O'Connell, K. S., Frei, O., Shadrin, A., Bettella, F., Rødevand, L., Fan, C. C., Dale, A. M., Djurovic, S., Smeland, O. B., & Andreassen, O. A. (2021). Characterising the shared genetic determinants of bipolar disorder, schizophrenia and risk-taking. *Translational Psychiatry*, *11*, 466. https://doi.org/10.1038/s41398-021-01576-4

Hindley, G., Frei, O., Shadrin, A. A., Cheng, W., O'Connell, K. S., Icick, R., Parker, N., Bahrami, S., Karadag, N., Roelfs, D., Holen, B., Lin, A., Fan, C. C., Djurovic, S., Dale, A. M., Smeland, O. B., & Andreassen, O. A. (2022). Charting the Landscape of Genetic Overlap Between Mental Disorders and Related Traits Beyond Genetic Correlation. *American Journal of Psychiatry*, *179*(11), 833-843. https://doi.org/10.1176/appi.ajp.21101051

Hjorthøj, C., Stürup, A. E., McGrath, J. J., & Nordentoft, M. (2017). Years of potential life lost and life expectancy in schizophrenia: a systematic review and meta-analysis. *The Lancet Psychiatry*, *4*(4), 295-301. https://doi.org/10.1016/S2215-0366(17)30078-0

Hoffmann, T. J., Theusch, E., Haldar, T., Ranatunga, D. K., Jorgenson, E., Medina, M. W., Kvale, M. N., Kwok, P., Schaefer, C., Krauss, R. M., Iribarren, C., & Risch, N. (2018). A large electronic-health-record-based genome-wide study of serum lipids. *Nature Genetics*, *50*(3), 401-413. https://doi.org/10.1038/s41588-018-0064-5

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

Holland, D., Frei, O., Desikan, R., Fan, C. C., Shadrin, A. A., Smeland, O. B., Sundar, V. S., Thompson, P., Andreassen, O. A., & Dale, A. M. (2020). Beyond SNP heritability:

Polygenicity and discoverability of phenotypes estimated with a univariate Gaussian mixture model. *PLoS Genetics*, *16*(5), 1-30. https://doi.org/10.1371/journal.pgen.1008612

Hong, L. E., Yang, X., Wonodi, I., Hodgkinson, C. A., Goldman, D., Stine, O. C., Stein, E. S., & Thaker, G. K. (2011). A CHRNA5 allele related to nicotine addiction and schizophrenia. *Genes, Brain and Behavior*, *10*(5), 530-535. https://doi.org/10.1111/j.1601-183X.2011.00689.x

Hormozdiari, F., van de Bunt, M., Segrè, A. V., Li, X., Joo, J. W. J., Bilow, M., Sul, J. H., Sankararaman, S., Pasaniuc, B., & Eskin, E. (2016). Colocalization of GWAS and eQTL Signals Detects Target Genes. *American Journal of Human Genetics*, *99*(6), 1245-1260. https://doi.org/10.1016/j.ajhg.2016.10.003

Howard, D. M., Adams, M. J., Clarke, T.-K., Hafferty, J. D., Gibson, J., Shirali, M., Coleman, J. R. I., Hagenaars, S. P., Ward, J., Wigmore, E. M., Alloza, C., Shen, X., Barbu, M. C., Xu, E. Y., Whalley, H. C., Marioni, R. E., Porteous, D. J., Davies, G., Deary, I. J., … 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

Howe, K., Clark, M. D., Torroja, C. F., Torrance, J., Berthelot, C., Muffato, M., Collins, J. E., Humphray, S., McLaren, K., Matthews, L., McLaren, S., Sealy, I., Caccamo, M., Churcher, C., Scott, C., Barrett, J. C., Koch, R., Rauch, G. J., White, S., … Stemple, D. L. (2013). The zebrafish reference genome sequence and its relationship to the human genome. *Nature*, *496*(7446), 498-503. https://doi.org/10.1038/nature12111

Huggett, S. B., Johnson, E. C., Hatoum, A. S., Lai, D., Srijeyanthan, J., Bubier, J. A., Chesler, E. J., Agrawal, A., Palmer, A. A., Edenberg, H. J., & Palmer, R. H. C. (2021). Genes identified in rodent studies of alcohol intake are enriched for heritability of human substance use. *Alcoholism: Clinical and Experimental Research*, *45*(12), 2485-2494. https://doi.org/10.1111/acer.14738

Hughes, J. R., Hatsukami, K., Mitchell, J. E., & Dahlgren, L. A. (1986). Prevalence of smoking among psychiatric outpatients. *American Journal of Psychiatry*, *143*(8), 993-997. https://doi.org/10.1176/ajp.143.8.993

Hunter, A., Murray, R., Asher, L., & Leonardi-Bee, J. (2020). The Effects of Tobacco Smoking, and Prenatal Tobacco Smoke Exposure, on Risk of Schizophrenia: A Systematic Review

USC UNIVERSIDADE DE SANTIAGO

and Meta-Analysis. *Nicotine & Tobacco Research*, *22*(1), 3-10. https://doi.org/10.1093/ntr/nty160

Hussong, A. M., Jones, D. J., Stein, G. L., Baucom, D. H., & Boeding, S. (2011). An internalizing pathway to alcohol use and disorder. *Psychology of Addictive Behaviors*, *25*(3), 390-404. https://doi.org/10.1037/a0024519

Icick, R., Shadrin, A., Holen, B., Karadag, N., Lin, A., Hindley, G., O'Connell, K., Frei, O., Bahrami, S., Høegh, M. C., Cheng, W., Fan, C. C., Djurovic, S., Dale, A. M., Lagerberg, T. V., Smeland, O. B., & Andreassen, O. A. (2022). Genetic overlap between mood instability and alcohol-related phenotypes suggests shared biological underpinnings. *Neuropsychopharmacology*, *47*(11), 1883-1891. https://doi.org/10.1038/s41386-022-01401-6

Icick, R., Shadrin, A., Holen, B., Karadag, N., Parker, N., O'Connell, K. S., Frei, O., Bahrami, S., Høegh, M. C., L, T. V., Cheng, W., Seibert, T. M., Djurovic, S., Dale, A. M., Zhou, H., Edenberg, H. J., Gelernter, J., Smeland, O. B., Hindley, G., & Andreassen, O. A. (2023). *Identification of Novel Loci and Cross-Disorder Pleiotropy Through Multi- Ancestry Genome-Wide Analysis of Alcohol Use Disorder in Over One Million Individuals. PREPRINT (Version 1) available at Research Square.* https://doi.org/10.21203/rs.3.rs-3755915/v1

Ip, H. F., van der Laan, C. M., Krapohl, E. M. L., Brikell, I., Sánchez-Mora, C., Nolte, I. M., St Pourcain, B., Bolhuis, K., Palviainen, T., Zafarmand, H., Colodro-Conde, L., Gordon, S., Zayats, T., Aliev, F., Jiang, C., Wang, C. A., Saunders, G., Karhunen, V., Hammerschlag, A. R., … Boomsma, D. I. (2021). Genetic association study of childhood aggression across raters, instruments, and age. *Translational Psychiatry*, *11*, 413. https://doi.org/10.1038/s41398-021-01480-x

Jablensky, A. (2006). Subtyping schizophrenia: Implications for genetic research. *Molecular Psychiatry*, *11*(9), 815-836. https://doi.org/10.1038/sj.mp.4001857

Jacobsen, L. K., D'Souza, D. C., Mencl, W. E., Pugh, K. R., Skudlarski, P., & Krystal, J. H. (2004). Nicotine effects on brain function and functional connectivity in schizophrenia. *Biological Psychiatry*, *55*(8), 850-858. https://doi.org/10.1016/j.biopsych.2003.12.023

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

Jaffe, A. E., Straub, R. E., Shin, J. H., Tao, R., Gao, Y., Collado-Torres, L., Kam-Thong, T., Xi, H. S., Quan, J., Chen, Q., Colantuoni, C., Ulrich, W. S., Maher, B. J., Deep-Soboslay,

A., Cross, A. J., Brandon, N. J., Leek, J. T., Hyde, T. M., Kleinman, J. E., & Weinberger, D. R. (2018). Developmental and genetic regulation of the human cortex transcriptome illuminate schizophrenia pathogenesis. *Nature Neuroscience*, *21*(8), 1117-1125. https://doi.org/10.1038/s41593-018-0197-y

Jang, S.-K., Saunders, G., Liu, M., Jiang, Y., Liu, D. J., & Vrieze, S. (2022). Genetic correlation, pleiotropy, and causal associations between substance use and psychiatric disorder. *Psychological Medicine*, *52*(5), 968-978. https://doi.org/10.1017/S003329172000272X

Jansen, P. R., Watanabe, K., Stringer, S., Skene, N., Bryois, J., Hammerschlag, A. R., de Leeuw, C. A., Benjamins, J. S., Muñoz-Manchado, A. B., Nagel, M., Savage, J. E., Tiemeier, H., White, T., Agee, M., Alipanahi, B., Auton, A., Bell, R. K., Bryc, K., Elson, S. L., … 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., Austin-Zimmerman, I., Thorpe, H. H., Levey, D. F., Baranger, D. A., Colbert, S. M., Demontis, D., Khokhar, J. Y., Davis, L. K., Edenberg, H. J., Forti, M. Di, Sanchez-Roige, S., Gelernter, J., & Agrawal, A. (2024). *Cross-ancestry genetic investigation of schizophrenia, cannabis use disorder, and tobacco smoking. PREPRINT (Version 1) available at medRxiv*. https://doi.org/10.1101/2024.01.17.24301430

Johnson, E. C., Demontis, D., Thorgeirsson, T. E., Walters, R. K., Polimanti, R., Hatoum, A. S., Sanchez-Roige, S., Paul, S. E., Wendt, F. R., Clarke, T. K., Lai, D., Reginsson, G. W., Zhou, H., He, J., Baranger, D. A. A., Gudbjartsson, D. F., Wedow, R., Adkins, D. E., Adkins, A. E., … 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

Johnson, E. C., Hatoum, A. S., Deak, J. D., Polimanti, R., Murray, R. M., Edenberg, H. J., Gelernter, J., Di Forti, M., & Agrawal, A. (2021). The relationship between cannabis and schizophrenia: a genetically informed perspective. *Addiction*, *116*(11), 3227-3234. https://doi.org/10.1111/add.15534

Johnson, W. E., Li, C., & Rabinovic, A. (2007). Adjusting batch effects in microarray expression data using empirical Bayes methods. *Biostatistics*, *8*(1), 118-127. https://doi.org/10.1093/biostatistics/kxj037

USC DE SANTIAGO DE COMPOSTELA

Jones, H. J., Gage, S. H., Heron, J., Hickman, M., Lewis, G., Munafò, M. R., & Zammit, S. (2018). Association of combined patterns of tobacco and cannabis use in adolescence with psychotic experiences. *JAMA Psychiatry*, *75*(3), 240-246. https://doi.org/10.1001/jamapsychiatry.2017.4271

Kahn, R. S., Sommer, I. E., Murray, R. M., Meyer-Lindenberg, A., Weinberger, D. R., Cannon, T. D., O'Donovan, M., Correll, C. U., Kane, J. M., van Os, J., & Insel, T. R. (2015). Schizophrenia. *Nature Reviews Disease Primers*, *1*(1), 15067. https://doi.org/10.1038/nrdp.2015.67

Kalueff, A. V., Stewart, A. M., & Gerlai, R. (2014). Zebrafish as an emerging model for studying complex brain disorders. *Trends in Pharmacological Sciences*, *35*(2), 63-75. https://doi.org/10.1016/j.tips.2013.12.002

Kang, H. J., Kawasawa, Y. I., Cheng, F., Zhu, Y., Xu, X., Li, M., Sousa, A. M. M., Pletikos, M., Meyer, K. A., Sedmak, G., Guennel, T., Shin, Y., Johnson, M. B., Krsnik, Ž., Mayer, S., Fertuzinhos, S., Umlauf, S., Lisgo, S. N., Vortmeyer, A., … Šestan, N. (2011). Spatio-temporal transcriptome of the human brain. *Nature*, *478*(7370), 483-489. https://doi.org/10.1038/nature10523

Karlsson Linnér, R., Biroli, P., Kong, E., Meddens, S. F. W., Wedow, R., Fontana, M. A., Lebreton, M., Tino, S. P., Abdellaoui, A., Hammerschlag, A. R., Nivard, M. G., Okbay, A., Rietveld, C. A., Timshel, P. N., Trzaskowski, M., Vlaming, R. de, Zünd, C. L., Bao, Y., Buzdugan, L., … 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

Karlsson Linnér, R., Mallard, T. T., Barr, P. B., Sanchez-Roige, S., Madole, J. W., Driver, M. N., Poore, H. E., de Vlaming, R., Grotzinger, A. D., Tielbeek, J. J., Johnson, E. C., Liu, M., Rosenthal, S. B., Ideker, T., Zhou, H., Kember, R. L., Pasman, J. A., Verweij, K. J. H., Liu, D. J., … Dick, D. M. (2021). Multivariate analysis of 1.5 million people identifies genetic associations with traits related to self-regulation and addiction. *Nature Neuroscience*, *24*(10), 1367-1376. https://doi.org/10.1038/s41593-021-00908-3

Kasap, M., Rajani, V., Rajani, J., & Dwyer, D. S. (2018). Surprising conservation of schizophrenia risk genes in lower organisms reflects their essential function and the

evolution of genetic liability. *Schizophrenia Research*, *202*, 120-128. https://doi.org/10.1016/j.schres.2018.07.017

Katikireddi, S. V., Green, M. J., Taylor, A. E., Davey Smith, G., & Munafò, M. R. (2018). Assessing causal relationships using genetic proxies for exposures: an introduction to Mendelian randomization. *Addiction*, *113*(4), 764-774. https://doi.org/10.1111/add.14038

Keller, M. C. (2014). Gene × environment interaction studies have not properly controlled for potential confounders: the problem and the (simple) solution. *Biological Psychiatry*, *75*(1), 18-24. https://doi.org/10.1016/j.biopsych.2013.09.006

Kelly, D. L., McMahon, R. P., Wehring, H. J., Liu, F., MacKowick, K. M., Boggs, D. L., Warren, K. R., Feldman, S., Shim, J. C., Love, R. C., & Dixon, L. (2011). Cigarette smoking and mortality risk in people with schizophrenia. *Schizophrenia Bulletin*, *37*(4), 832-838. https://doi.org/10.1093/schbul/sbp152

Kelly, D. L., Raley, H. G., Lo, S., Wright, K., Liu, F., McMahon, R. P., Moolchan, E. T., Feldman, S., Richardson, C. M., Wehring, H. J., & Heishman, S. J. (2012). Perception of smoking risks and motivation to quit among nontreatment-seeking smokers with and without schizophrenia. *Schizophrenia Bulletin*, *38*(3), 543-551. https://doi.org/10.1093/schbul/sbq124

Kendler, K. S., Jacobson, K. C., Prescott, C. A., & Neale, M. C. (2003). Specificity of genetic and environmental risk factors for use and abuse/dependence of cannabis, cocaine, hallucinogens, sedatives, stimulants, and opiates in male twins. *American Journal of Psychiatry*, *160*(4), 687-695. https://doi.org/10.1176/appi.ajp.160.4.687

Kendler, K. S., Lönn, S. L., Sundquist, J., & Sundquist, K. (2015). Smoking and Schizophrenia in Population Cohorts of Swedish Women and Men: A Prospective Co-Relative Control Study. *American Journal of Psychiatry*, *172*(11), 1092-1100. https://doi.org/10.1176/appi.ajp.2015.15010126

Kendler, K. S., Myers, J., Dick, D., & Prescott, C. A. (2010). The relationship between genetic influences on alcohol dependence and on patterns of alcohol consumption. *Alcoholism: Clinical and Experimental Research*, *34*(6), 1058-1065. https://doi.org/10.1111/j.1530-0277.2010.01181.x

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

Kendler, K. S., Schmitt, E., Aggen, S. H., & Prescott, C. A. (2008). Genetic and Environmental

Influences on Alcohol, Caffeine, Cannabis, and Nicotine Use From Early Adolescence to Middle Adulthood. *Archives of General Psychiatry*, *65*(6), 674. https://doi.org/10.1001/archpsyc.65.6.674

Kessler, R. C. (2004). The epidemiology of dual diagnosis. *Biological Psychiatry*, *56*(10), 730-737. https://doi.org/10.1016/j.biopsych.2004.06.034

Khantzian, E. J. (1997). The Self-Medication Hypothesis of Substance Use Disorders: A Reconsideration and Recent Applications. *Harvard Review of Psychiatry*, *4*(5), 231-244. http://www.tandfonline.com/doi/abs/10.3109/10673229709030550

Khokhar, J. Y., Dwiel, L. L., Henricks, A. M., Doucette, W. T., & Green, A. I. (2018). The link between schizophrenia and substance use disorder: A unifying hypothesis. *Schizophrenia Research*, *194*(1), 78-85. https://doi.org/10.1016/j.schres.2017.04.016

Kily, L. J. M., Cowe, Y. C. M., Hussain, O., Patel, S., McElwaine, S., Cotter, F. E., & Brennan, C. H. (2008). Gene expression changes in a zebrafish model of drug dependency suggest conservation of neuro-adaptation pathways. *Journal of Experimental Biology*, *211*(10), 1623-1634. https://doi.org/10.1242/jeb.014399

Kim, S., Kim, K., Hwang, M. Y., Ko, H., Jung, S. H., Shim, I., Cha, S., Lee, H., Kim, B., Yoon, J., Ha, T. H., Kim, D. K., Kim, J., Park, W. Y., Okbay, A., Kim, B. J., Kim, Y. J., Myung, W., & Won, H. H. (2022). Shared genetic architectures of subjective well-being in East Asian and European ancestry populations. *Nature Human Behaviour*, *6*(7), 1014-1026. https://doi.org/10.1038/s41562-022-01343-5

King, M., Jones, R., Petersen, I., Hamilton, F., & Nazareth, I. (2021). Cigarette smoking as a risk factor for schizophrenia or all non-Affective psychoses. *Psychological Medicine*, *51*(8), 1373-1381. https://doi.org/10.1017/S0033291720000136

King, S. M., Iacono, W. G., & McGue, M. (2004). Childhood externalizing and internalizing psychopathology in the prediction of early substance use. *Addiction*, *99*(12), 1548-1559. https://doi.org/10.1111/j.1360-0443.2004.00893.x

Kleinig, J. (2015). Ready for Retirement: The Gateway Drug Hypothesis. *Substance Use and Misuse*, *50*(8-9), 971-975. https://doi.org/10.3109/10826084.2015.1007679

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

Kleinman, K., & Horton, N. J. (2014). *SAS and R. Data Management, Statistical Analysis and Graphics* (2nd ed.). Chapman & Hall.

Koller, D., Mitjans, M., Kouakou, M., Friligkou, E., Cabrera-Mendoza, B., Deak, J. D., Llonga, N., Pathak, G. A., Stiltner, B., Løkhammer, S., Levey, D. F., Zhou, H., Hatoum, A. S., Kember, R. L., Kranzler, H. R., Stein, M. B., Corominas, R., Demontis, D., Artigas, M. S., … Polimanti, R. (2024). Genetic contribution to the comorbidity between attention-deficit/hyperactivity disorder and substance use disorders. *Psychiatry Research*, *333*(September 2023). https://doi.org/10.1016/j.psychres.2024.115758

Kollins, S. H., McClernon, F. J., & Fuemmeler, B. F. (2005). Association Between Smoking and Attention-Deficit/Hyperactivity Disorder Symptoms in a Population-Based Sample of Young Adults. *Archives of General Psychiatry*, *62*(10), 1142-1147. https://doi.org/10.1001/archpsyc.62.10.1142

Koob, G. F., Buck, C. L., Cohen, A., Edwards, S., Park, P. E., Schlosburg, J. E., Schmeichel, B., Vendruscolo, L. F., Wade, C. L., Whitfield, T. W., & George, O. (2014). Addiction as a stress surfeit disorder. *Neuropharmacology*, *76*(PART B), 370-382. https://doi.org/10.1016/j.neuropharm.2013.05.024

Koob, G. F., & Volkow, N. D. (2010). Neurocircuitry of addiction. *Neuropsychopharmacology*, *35*(1), 217-238. https://doi.org/10.1038/npp.2009.110

Koob, G. F., & Volkow, N. D. (2016). Neurobiology of addiction: a neurocircuitry analysis. *The Lancet Psychiatry*, *3*(8), 760-773. https://doi.org/10.1016/S2215-0366(16)00104-8

Koopmans, F., van Nierop, P., Andres-Alonso, M., Byrnes, A., Cijsouw, T., Coba, M. P., Cornelisse, L. N., Farrell, R. J., Goldschmidt, H. L., Howrigan, D. P., Hussain, N. K., Imig, C., de Jong, A. P. H., Jung, H., Kohansalnodehi, M., Kramarz, B., Lipstein, N., Lovering, R. C., MacGillavry, H., … Verhage, M. (2019). SynGO: An Evidence-Based, Expert-Curated Knowledge Base for the Synapse. *Neuron*, *103*(2), 217-234. https://doi.org/10.1016/j.neuron.2019.05.002

Koyanagi, A., Stickley, A., & Haro, J. M. (2016). Psychotic symptoms and smoking in 44 countries. *Acta Psychiatrica Scandinavica*, *133*(6), 497-505. https://doi.org/10.1111/acps.12566

Krueger, R. F., Hicks, B. M., Patrick, C. J., Carlson, S. R., Iacono, W. G., & McGue, M. (2002). Etiologic connections among substance dependence, antisocial behavior, and personality: Modeling the externalizing spectrum. *Journal of Abnormal Psychology*, *111*(3), 411-424. https://doi.org/10.1037/0021-843X.111.3.411

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

Krueger, R. F., McGue, M., & Iacono, W. G. (2001). The higher-order structure of common DSM mental disorders: Internalization, externalization, and their connections to personality. *Personality and Individual Differences*, *30*(7), 1245-1259. https://doi.org/10.1016/S0191-8869(00)00106-9

Kumari, V., & Postma, P. (2005). Nicotine use in schizophrenia: The self medication hypotheses. *Neuroscience & Biobehavioral Reviews*, *29*(6), 1021-1034. https://doi.org/10.1016/j.neubiorev.2005.02.006

Laceulle, O. M., Vollebergh, W. A. M., & Ormel, J. (2015). The structure of psychopathology in adolescence: replication of a general psychopathology factor in the TRAILS study. *Clinical Psychological Science*, *3*(6), 850-860. https://doi.org/10.1177/2167702614560750

Lahey, B. B., Applegate, B., Hakes, J. K., Zald, D. H., Hariri, A. R., & Rathouz, P. J. (2012). Is There a general factor of prevalent psychopathology during adulthood? *Journal of Abnormal Psychology*, *121*(4), 971-977. https://doi.org/10.1037/a0028355

Langfelder, P., & Horvath, S. (2008). WGCNA: an R package for weighted correlation network analysis. *BMC Bioinformatics*, *9*(1), 559. https://doi.org/10.1186/1471-2105-9-559

Laursen, T. M., & McGrath, J. J. (2016). The Strange Case of Smoking and Schizophrenia - The Epidemiology Detectives Are on the Trail. En *American Journal of Psychiatry* (Vol. 173, Número 8, pp. 757-758). https://doi.org/10.1176/appi.ajp.2016.16050506

Lawlor, D. A., Tilling, K., & Smith, G. D. (2016). Triangulation in aetiological epidemiology. *International Journal of Epidemiology*, *45*(6), 1866-1886. https://doi.org/10.1093/ije/dyw314

Lawrence, D., Mitrou, F., & Zubrick, S. R. (2009). Smoking and mental illness: Results from population surveys in Australia and the United States. *BMC Public Health*, *9*, 1-14. https://doi.org/10.1186/1471-2458-9-285

Lee, H. K., Hsu, A. K., Sajdak, J., Qin, J., & Pavlidis, P. (2004). Coexpresion analysis of human genes across many microarray data sets. *Genome Research*, *14*(6), 1085-1094. https://doi.org/10.1101/gr.1910904

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

Lee, J. J., Wedow, R., Okbay, A., Kong, E., Maghzian, O., Zacher, M., Nguyen-Viet, T. A., Bowers, P., Sidorenko, J., Karlsson Linnér, R., Fontana, M. A., Kundu, T., Lee, C., Li, H.,

Li, R., Royer, R., Timshel, P. N., Walters, R. K., Willoughby, E. A., … 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

Lee, P. H., Anttila, V., Won, H., Feng, Y.-C. A., Rosenthal, J., Zhu, Z., Tucker-Drob, E. M., Nivard, M. G., Grotzinger, A. D., Posthuma, D., Wang, M. M.-J., Yu, D., Stahl, E. A., Walters, R. K., Anney, R. J. L., Duncan, L. E., Ge, T., Adolfsson, R., Banaschewski, T., … Smoller, J. W. (2019). Genomic Relationships, Novel Loci, and Pleiotropic Mechanisms across Eight Psychiatric Disorders. *Cell*, *179*(7), 1469-1482. https://doi.org/10.1016/j.cell.2019.11.020

Lee, S. H., Goddard, M. E., Wray, N. R., & Visscher, P. M. (2012). A Better Coefficient of Determination for Genetic Profile Analysis. *Genetic Epidemiology*, *36*(3), 214-224. https://doi.org/10.1002/gepi.21614

Lee, S. S., Humphreys, K. L., Flory, K., Liu, R., & Glass, K. (2011). Prospective association of childhood attention-deficit/hyperactivity disorder (ADHD) and substance use and abuse/dependence: A meta-analytic review. *Clinical Psychology Review*, *31*(3), 328-341. https://doi.org/10.1016/j.cpr.2011.01.006

Leung, J., Chan, G. C. K., Hides, L., & Hall, W. D. (2020). What is the prevalence and risk of cannabis use disorders among people who use cannabis? a systematic review and meta-analysis. *Addictive Behaviors*, *109*(April), 106479. https://doi.org/10.1016/j.addbeh.2020.106479

Lev-Ran, S., Imtiaz, S., Rehm, J., & Le Foll, B. (2013). Exploring the association between lifetime prevalence of mental illness and transition from substance use to substance use disorders: Results from the National Epidemiologic Survey of Alcohol and Related Conditions (NESARC). *American Journal on Addictions*, *22*(2), 93-98. https://doi.org/10.1111/j.1521-0391.2013.00304.x

Levey, D. F., Gelernter, J., Polimanti, R., Zhou, H., Cheng, Z., Aslan, M., Quaden, R., Concato, J., Radhakrishnan, K., Bryois, J., Sullivan, P. F., & Stein, M. B. (2020). Reproducible Genetic Risk Loci for Anxiety: Results From ~200,000 Participants in the Million Veteran Program. *American Journal of Psychiatry*, *177*(3), 223-232. https://doi.org/10.1176/appi.ajp.2019.19030256

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

Levin, E. D., Wilson, W., Rose, J. E., & McEvoy, J. (1996). Nicotine-haloperidol interactions and cognitive performance in schizophrenics. *Neuropsychopharmacology*, *15*(5), 429-436. https://doi.org/10.1016/S0893-133X(96)00018-8

Lewis, C. M., & Vassos, E. (2020). Polygenic risk scores: from research tools to clinical instruments. *Genome Medicine*, *12*(1), 44. https://doi.org/10.1186/s13073-020-00742-5

Li, M., Santpere, G., Imamura Kawasawa, Y., Evgrafov, O. V., Gulden, F. O., Pochareddy, S., Sunkin, S. M., Li, Z., Shin, Y., Zhu, Y., Sousa, A. M. M., Werling, D. M., Kitchen, R. R., Kang, H. J., Pletikos, M., Choi, J., Muchnik, S., Xu, X., Wang, D., … Li, Z. (2018). Integrative functional genomic analysis of human brain development and neuropsychiatric risks. *Science*, *362*(6420), eaat7615. https://doi.org/10.1126/science.aat7615

Liao, Y., Wang, J., Jaehnig, E. J., Shi, Z., & Zhang, B. (2019). WebGestalt 2019: gene set analysis toolkit with revamped UIs and APIs. *Nucleic Acids Research*, *47*(W1), W199-W205. https://doi.org/10.1093/nar/gkz401

Lin, B. D., Vermeulen, J. M., Bolhuis, K., Chang, X., Schirmbeck, F., van Eijk, K. R., Alizadeh, B. Z., Bartels-Velthuis, A. A., van Amelsvoort, T., Bruggeman, R., Cahn, W., de Haan, L., Rutten, B. P. F., Guloksuz, S., Luykx, J. J., van Os, J., Simons, C. J. P., van Winkel, R., Blankers, M., & van den Brink, W. (2023). Associations between genetic liabilities to smoking behavior and schizophrenia symptoms in patients with a psychotic disorder, their siblings and healthy controls. *Psychiatry Research*, *323*(December 2022). https://doi.org/10.1016/j.psychres.2023.115143

Liu, B., Gloudemans, M. J., Rao, A. S., Ingelsson, E., & Montgomery, S. B. (2019). Abundant associations with gene expression complicate GWAS follow-up. *Nature Genetics*, *51*(5), 768-769. https://doi.org/10.1038/s41588-019-0404-0

Liu, M., Jiang, Y., Wedow, R., Li, Y., Brazel, D. M., Chen, F., Datta, G., Davila-Velderrain, J., McGuire, D., Tian, C., Zhan, X., Choquet, H., Docherty, A. R., Faul, J. D., Foerster, J. R., Fritsche, L. G., Gabrielsen, M. E., Gordon, S. D., Haessler, J., … 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

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

Liu, X., Li, Y. I., & Pritchard, J. K. (2019). Trans Effects on Gene Expression Can Drive Omnigenic Inheritance. *Cell*, *177*(4), 1022-1034.e6.

https://doi.org/10.1016/j.cell.2019.04.014

Lloyd-Jones, L. R., Zeng, J., Sidorenko, J., Yengo, L., Moser, G., Kemper, K. E., Wang, H., Zheng, Z., Magi, R., Esko, T., Metspalu, A., Wray, N. R., Goddard, M. E., Yang, J., & Visscher, P. M. (2019). Improved polygenic prediction by Bayesian multiple regression on summary statistics. *Nature Communications*, *10*(1). https://doi.org/10.1038/s41467-019-12653-0

López-Patiño, M. A., Yu, L., Cabral, H., & Zhdanova, I. V. (2008). Anxiogenic effects of cocaine withdrawal in zebrafish. *Physiology and Behavior*, *93*(1-2), 160-171. https://doi.org/10.1016/j.physbeh.2007.08.013

Lum, A., Skelton, E., Wynne, O., & Bonevski, B. (2018). A systematic review of psychosocial barriers and facilitators to smoking cessation in people living with schizophrenia. *Frontiers in Psychiatry*, *9*(NOV). https://doi.org/10.3389/fpsyt.2018.00565

MacCabe, J. H. (2018). It is time to start taking tobacco seriously as a risk factor for psychosis: self-medication cannot explain the association. *Acta Psychiatrica Scandinavica*, *138*(1), 3-4. https://doi.org/10.1111/acps.12923

MacDermod, C., Pettie, M. A., Carrino, E. A., Garcia, S. C., Padalecki, S., Finch, J. E., Sanzari, C., Kennedy, H. L., Pawar, P. S., Mcgough, M. M., Iwashita, A., Takgbajouah, M., Coan, D., Szakasits, L., Goode, R. W., Wu, Y. K., Reyes-Rodríguez, M. L., Vacuán, E. M. T. C., Kennedy, M. A., … Bulik, C. M. (2022). Recommendations to encourage participation of individuals from diverse backgrounds in psychiatric genetic studies. *American Journal of Medical Genetics, Part B: Neuropsychiatric Genetics*, *189*(5), 163-173. https://doi.org/10.1002/ajmg.b.32906

Machiela, M. J., & Chanock, S. J. (2015). LDlink: a web-based application for exploring population-specific haplotype structure and linking correlated alleles of possible functional variants. *Bioinformatics*, *31*(21), 3555-3557. https://doi.org/10.1093/bioinformatics/btv402

Maier, R. M., Zhu, Z., Lee, S. H., Trzaskowski, M., Ruderfer, D. M., Stahl, E. A., Ripke, S., Wray, N. R., Yang, J., Visscher, P. M., & Robinson, M. R. (2018). Improving genetic prediction by leveraging genetic correlations among human diseases and traits. *Nature Communications*, *9*(1), 989. https://doi.org/10.1038/s41467-017-02769-6

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

Mallard, T. T., Savage, J. E., Johnson, E. C., Huang, Y., Edwards, A. C., Hottenga, J. J., Grotzinger, A. D., Gustavson, D. E., Jennings, M. V., Anokhin, A., Dick, D. M., Edenberg, H. J., Kramer, J. R., Lai, D., Meyers, J. L., Pandey, A. K., Harden, K. P., Nivard, M. G., de Geus, E. J. C., … Sanchez-Roige, S. (2022). Item-Level Genome-Wide Association Study of the Alcohol Use Disorders Identification Test in Three Population-Based Cohorts. *The American journal of psychiatry*, *179*(1), 58-70. https://doi.org/10.1176/appi.ajp.2020.20091390

Mao, K., Li, X., Chen, Z., Dong, X., Zhangsun, D., Zhu, X., & Luo, S. (2022). α-Conotoxin TxIB Improved Behavioral Abnormality and Changed Gene Expression in Zebrafish (Danio rerio) Induced by Alcohol Withdrawal. *Frontiers in Pharmacology*, *13*(February), 1-15. https://doi.org/10.3389/fphar.2022.802917

Marchi, M., & Grilli, M. (2010). Presynaptic nicotinic receptors modulating neurotransmitter release in the Central Nervous System: Functional interactions with other coexisting receptors. *Progress in Neurobiology*, *92*(2), 105-111. https://doi.org/10.1016/j.pneurobio.2010.06.004

Marees, A. T., Smit, D. J. A., Ong, J. S., Macgregor, S., An, J., Denys, D., Vorspan, F., Van Den Brink, W., & Derks, E. M. (2019). Potential influence of socioeconomic status on genetic correlations between alcohol consumption measures and mental health. *Psychological Medicine*. https://doi.org/10.1017/S0033291719000357

Martin, A. R., Daly, M. J., Robinson, E. B., Hyman, S. E., & Neale, B. M. (2019). Predicting Polygenic Risk of Psychiatric Disorders. En *Biological Psychiatry* (Vol. 86, Número 2, pp. 97-109). Elsevier USA. https://doi.org/10.1016/j.biopsych.2018.12.015

Martin, A. R., Gignoux, C. R., Walters, R. K., Wojcik, G. L., Neale, B. M., Gravel, S., Daly, M. J., Bustamante, C. D., & Kenny, E. E. (2017). Human Demographic History Impacts Genetic Risk Prediction across Diverse Populations. *American Journal of Human Genetics*, *100*(4), 635-649. https://doi.org/10.1016/j.ajhg.2017.03.004

Maskos, U. (2020). The nicotinic receptor alpha5 coding polymorphism rs16969968 as a major target in disease: Functional dissection and remaining challenges. *Journal of Neurochemistry*, *154*(3), 241-250. https://doi.org/10.1111/jnc.14989

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

Mathur, P., & Guo, S. (2010). Use of zebrafish as a model to understand mechanisms of addiction and complex neurobehavioral phenotypes. *Neurobiology of Disease*, *40*(1), 66-

72. https://doi.org/10.1016/j.nbd.2010.05.016

Mathur, P., & Guo, S. (2011). Differences of acute versus chronic ethanol exposure on anxiety-like behavioral responses in zebrafish. *Behavioural Brain Research*, *219*(2), 234-239. https://doi.org/10.1016/j.bbr.2011.01.019

Maurano, M. T., Humbert, R., Rynes, E., Thurman, R. E., Haugen, E., Wang, H., Reynolds, A. P., Sandstrom, R., Qu, H., Brody, J., Shafer, A., Neri, F., Lee, K., Kutyavin, T., Stehling-Sun, S., Johnson, A. K., Canfield, T. K., Giste, E., Diegel, M., … Stamatoyannopoulos, J. A. (2012). Systematic localization of common disease-associated variation in regulatory DNA. *Science*, *337*(6099), 1190-1195. https://doi.org/10.1126/science.1222794

Mayhew, K. P., Flay, B. R., & Mott, J. A. (2000). Stages in the development of adolescent smoking. *Drug and Alcohol Dependence*, *59*(SUPPL. 1), 61-81. https://doi.org/10.1016/S0376-8716(99)00165-9

McCammon, J. M., & Sive, H. (2015). Addressing the Genetics of Human Mental Health Disorders in Model Organisms. *Annual Review of Genomics and Human Genetics*, *16*, 173-197. https://doi.org/10.1146/annurev-genom-090314-050048

McCarthy, S. (2016). A reference panel of 64,976 haplotypes for genotype imputation. *Nature Genetics*, *48*(10), 1279-1283. https://doi.org/10.1038/ng.3643

McClave, A. K., McKnight-Eily, L. R., Davis, S. P., & Dube, S. R. (2010). Smoking characteristics of adults with selected lifetime mental illnesses: Results from the 2007 national health interview survey. *American Journal of Public Health*, *100*(12), 2464-2472. https://doi.org/10.2105/AJPH.2009.188136

McGrath, J. J. (2015). A Rosetta stone for epidemiology: Genomic risk profile scores contain clues related to modifiable risk factors. *Epidemiology and Psychiatric Sciences*, *24*(1), 1-5. https://doi.org/10.1017/S2045796014000651

McGrath, J. J., Alati, R., Clavarino, A., Williams, G. M., Bor, W., Najman, J. M., Connell, M., & Scott, J. G. (2016). Age at first tobacco use and risk of subsequent psychosis-related outcomes: A birth cohort study. *Australian & New Zealand Journal of Psychiatry*, *50*(6), 577-583. https://doi.org/10.1177/0004867415587341

McKay, J. D., Hung, R. J., Han, Y., Zong, X., Carreras-Torres, R., Christiani, D. C., Caporaso, N. E., Johansson, M., Xiao, X., Li, Y., Byun, J., Dunning, A., Pooley, K. A., Qian, D. C.,

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

Ji, X., Liu, G., Timofeeva, M. N., Bojesen, S. E., Wu, X., … Amos, C. I. (2017). Large-scale association analysis identifies new lung cancer susceptibility loci and heterogeneity in genetic susceptibility across histological subtypes. *Nature Genetics*, *49*(7), 1126-1132. https://doi.org/10.1038/ng.3892

Meque, I., Dachew, B. A., Maravilla, J. C., Salom, C., & Alati, R. (2019). Externalizing and internalizing symptoms in childhood and adolescence and the risk of alcohol use disorders in young adulthood: A meta-analysis of longitudinal studies. *Australian and New Zealand Journal of Psychiatry*, *53*(10), 965-975. https://doi.org/10.1177/0004867419844308

Miller, A. P., Gizer, I. R., Fleming, W. A., Otto, J. M., Deak, J. D., Martins, J. S., & Bartholow, B. D. (2019). Polygenic liability for schizophrenia predicts shifting-specific executive function deficits and tobacco use in a moderate drinking community sample. *Psychiatry Research*, *279*, 47-54. https://doi.org/10.1016/j.psychres.2019.06.025

Mills, M. C., Tropf, F. C., Brazel, D. M., van Zuydam, N., Vaez, A., Agbessi, M., Ahsan, H., Alves, I., Andiappan, A. K., Arindrarto, W., Awadalla, P., Battle, A., Beutner, F., Jan Bonder, M., Boomsma, D. I., Christiansen, M. W., Claringbould, A., Deelen, P., Esko, T., … 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

Missen, R. L., Brannelly, T., & Newton-Howes, G. (2013). Qualitative exploration of family perspectives of smoke-free mental health and addiction services. *International Journal of Mental Health Nursing*, *22*(4), 294-303. https://doi.org/10.1111/j.1447-0349.2012.00882.x

Mize, T. J., Funkhouser, S. A., Buck, J. M., Stitzel, J. A., Ehringer, M. A., & Evans, L. M. (2023). Testing Association of Previously Implicated Gene Sets and Gene-Networks in Nicotine Exposed Mouse Models with Human Smoking Phenotypes. *Nicotine and Tobacco Research*, *25*(5), 1030-1038. https://doi.org/10.1093/ntr/ntac269

Mocelin, R., Marcon, M., da Rosa Araujo, A. S., Herrmann, A. P., & Piato, A. (2019). Withdrawal effects following repeated ethanol exposure are prevented by N-acetylcysteine in zebrafish. *Progress in Neuro-Psychopharmacology and Biological Psychiatry*, *93*(January), 161-170. https://doi.org/10.1016/j.pnpbp.2019.03.014

Mokdad, A. H. (2004). Actual Causes of Death in the United States, 2000. *JAMA*, *291*(10),

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

1238. https://doi.org/10.1001/jama.291.10.1238

Moore, T. H., Zammit, S., Lingford-Hughes, A., Barnes, T. R., Jones, P. B., Burke, M., & Lewis, G. (2007). Cannabis use and risk of psychotic or affective mental health outcomes: a systematic review. *Lancet*, *370*(9584), 319-328. https://doi.org/10.1016/S0140-6736(07)61162-3

Müller, T. E., Fontana, B. D., Bertoncello, K. T., Franscescon, F., Mezzomo, N. J., Canzian, J., Stefanello, F. V., Parker, M. O., Gerlai, R., & Rosemberg, D. B. (2020). Understanding the neurobiological effects of drug abuse: Lessons from zebrafish models. *Progress in Neuro-Psychopharmacology and Biological Psychiatry*, *100*(January), 109873. https://doi.org/10.1016/j.pnpbp.2020.109873

Mullins, N., Forstner, A. J., O'Connell, K. S., Coombes, B., Coleman, J. R. I., Qiao, Z., Als, T. D., Bigdeli, T. B., Børte, S., Bryois, J., Charney, A. W., Drange, O. K., Gandal, M. J., Hagenaars, S. P., Ikeda, M., Kamitaki, N., Kim, M., Krebs, K., Panagiotaropoulou, G., … 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

Mullins, N., Kang, J. E., Campos, A. I., Coleman, J. R. I., Edwards, A. C., Galfalvy, H., Levey, D. F., Lori, A., Shabalin, A., Starnawska, A., Su, M. H., Watson, H. J., Adams, M., Awasthi, S., Gandal, M., Hafferty, J. D., Hishimoto, A., Kim, M., Okazaki, S., … Willour, V. (2022). Dissecting the Shared Genetic Architecture of Suicide Attempt, Psychiatric Disorders, and Known Risk Factors. *Biological Psychiatry*, *91*(3), 313-327. https://doi.org/10.1016/j.biopsych.2021.05.029

Munafò, M. R., & Davey Smith, G. (2018). Robust research needs many lines of evidence. *Nature*, *553*(7689), 399-401. https://doi.org/10.1038/d41586-018-01023-3

Murray, G. K., Lin, T., Austin, J., McGrath, J. J., Hickie, I. B., & Wray, N. R. (2021). Could Polygenic Risk Scores Be Useful in Psychiatry?: A Review. *JAMA Psychiatry*, *78*(2), 210-219. https://doi.org/10.1001/jamapsychiatry.2020.3042

Mustonen, A., Ahokas, T., Nordström, T., Murray, G. K., Mäki, P., Jääskeläinen, E., Heiskala, A., Mcgrath, J. J., Scott, J. G., Miettunen, J., & Niemelä, S. (2018). Smokin' hot: adolescent smoking and the risk of psychosis. *Acta Psychiatrica Scandinavica*, *138*(1), 5-14. https://doi.org/10.1111/acps.12863

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

Myles, N., Newall, H. D., Curtis, J., Nielssen, O., Shiers, D., & Large, M. (2012). Tobacco use before, at, and after first-episode psychosis: A systematic meta-analysis. En *Journal of Clinical Psychiatry* (Vol. 73, Número 4, pp. 468-475). Physicians Postgraduate Press Inc. https://doi.org/10.4088/JCP.11r07222

Nagel, M., Jansen, P. R., Stringer, S., Watanabe, K., De Leeuw, C. A., Bryois, J., Savage, J. E., Hammerschlag, A. R., Skene, N. G., Muñoz-Manchado, A. B., Agee, M., Alipanahi, B., Auton, A., Bell, R. K., Bryc, K., Elson, S. L., Fontanillas, P., Furlotte, N. A., Hinds, D. A., … 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

Nagel, M., Watanabe, K., Stringer, S., Posthuma, D., & Van Der Sluis, S. (2018). Item-level analyses reveal genetic heterogeneity in neuroticism. *Nature Communications*, *9*(1). https://doi.org/10.1038/s41467-018-03242-8

Neumann, A., Pappa, I., Lahey, B. B., Verhulst, F. C., Medina-Gomez, C., Jaddoe, V. W., Bakermans-Kranenburg, M. J., Moffitt, T. E., van IJzendoorn, M. H., & Tiemeier, H. (2016). Single Nucleotide Polymorphism Heritability of a General Psychopathology Factor in Children. *Journal of the American Academy of Child & Adolescent Psychiatry*, *55*(12), 1038-1045. https://doi.org/10.1016/j.jaac.2016.09.498

Ni, G., Zeng, J., Revez, J. A., Wang, Y., Zheng, Z., Ge, T., Restuadi, R., Kiewa, J., Nyholt, D. R., Coleman, J. R. I., Smoller, J. W., Ripke, S., Neale, B. M., Corvin, A., Walters, J. T. R., Farh, K. H., Holmans, P. A., Lee, P., Bulik-Sullivan, B., … Pedersen, N. L. (2021). A Comparison of Ten Polygenic Score Methods for Psychiatric Disorders Applied Across Multiple Cohorts. *Biological Psychiatry*, *90*(9), 611-620. https://doi.org/10.1016/j.biopsych.2021.04.018

Niemelä, S., Sourander, A., Surcel, H.-M., Hinkka-Yli-Salomäki, S., McKeague, I. W., Cheslack-Postava, K., & Brown, A. S. (2016). Prenatal Nicotine Exposure and Risk of Schizophrenia Among Offspring in a National Birth Cohort. *American Journal of Psychiatry*, *173*(8), 799-806. https://doi.org/10.1176/appi.ajp.2016.15060800

Nievergelt, C. M., Maihofer, A. X., Klengel, T., Atkinson, E. G., Chen, C. Y., Choi, K. W., Coleman, J. R. I., Dalvie, S., Duncan, L. E., Gelernter, J., Levey, D. F., Logue, M. W., Polimanti, R., Provost, A. C., Ratanatharathorn, A., Stein, M. B., Torres, K., Aiello, A. E.,

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

Almli, L. M., … Koenen, K. C. (2019). International meta-analysis of PTSD genome-wide association studies identifies sex- and ancestry-specific genetic risk loci. *Nature Communications*, *10*(1), 1-16. https://doi.org/10.1038/s41467-019-12576-w

Nikpay, M., Goel, A., Won, H.-H., Hall, L. M., Willenborg, C., Kanoni, S., Saleheen, D., Kyriakou, T., Nelson, C. P., Hopewell, J. C., Webb, T. R., Zeng, L., Dehghan, A., Alver, M., Armasu, S. M., Auro, K., Bjonnes, A., Chasman, D. I., Chen, S., … Farrall, M. (2015). A comprehensive 1000 Genomes–based genome-wide association meta-analysis of coronary artery disease. *Nature Genetics*, *47*(10), 1121-1130. https://doi.org/10.1038/ng.3396

Ning, K., Gondek, D., Patalay, P., & Ploubidis, G. B. (2020). The association between early life mental health and alcohol use behaviours in adulthood: A systematic review. *PLoS ONE*, *15*(2), 1-28. https://doi.org/10.1371/journal.pone.0228667

Nolan, K. A., D'Angelo, D., & Hoptman, M. J. (2011). Self-report and laboratory measures of impulsivity in patients with schizophrenia or schizoaffective disorder and healthy controls. *Psychiatry Research*, *187*(1-2), 301-303. https://doi.org/10.1016/j.psychres.2010.10.032

Nonnis, S., Angiulli, E., Maffioli, E., Frabetti, F., Negri, A., Cioni, C., Alleva, E., Romeo, V., Tedeschi, G., & Toni, M. (2021). Acute environmental temperature variation affects brain protein expression, anxiety and explorative behaviour in adult zebrafish. *Scientific Reports*, *11*(1), 1-21. https://doi.org/10.1038/s41598-021-81804-5

O'Connell, J., Yun, T., Moreno, M., Li, H., Litterman, N., Kolesnikov, A., Noblin, E., Chang, P. C., Shastri, A., Dorfman, E. H., Shringarpure, S., Aslibekyan, S., Babalola, E., Bell, R. K., Bielenberg, J., Bryc, K., Bullis, E., Coker, D., Partida, G. C., … McLean, C. Y. (2021). A population-specific reference panel for improved genotype imputation in African Americans. *Communications Biology*, *4*(1), 1-9. https://doi.org/10.1038/s42003-021-02777-9

O'dushlaine, C., Rossin, L., Lee, P. H., Duncan, L., Parikshak, N. N., Newhouse, S., Ripke, S., Neale, B. M., Purcell, S. M., Posthuma, D., Nurnberger, J. I., Lee, S. H., Faraone, S. V., Perlis, R. H., Mowry, B. J., Thapar, A., Goddard, M. E., Witte, J. S., Absher, D., … Breen, G. (2015). Psychiatric genome-wide association study analyses implicate neuronal, immune and histone pathways. *Nature Neuroscience*, *18*(2), 199-209. https://doi.org/10.1038/nn.3922

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

Observatorio Español de las Drogas y las Adicciones. (2015). *Informe 2015: Alcohol, tabaco y drogas ilegales en España*. http://crm.org.mx/PDF/INFORMES/INFORME2014.pdf

Observatorio Español de las Drogas y las Adicciones. (2023). Informe 2023. Alcohol, tabaco y drogas ilegales en España. En *Ministerio de Sanidad. Delegación del Gobierno para el Plan Nacional sobre Drogas*.

Ohi, K., Kuwata, A., Shimada, T., Kataoka, Y., Yasuyama, T., Uehara, T., & Kawasaki, Y. (2019). Genome-Wide Variants Shared Between Smoking Quantity and Schizophrenia on 15q25 Are Associated With CHRNA5 Expression in the Brain. *Schizophrenia Bulletin*, *45*(4), 813-823. https://doi.org/10.1093/schbul/sby093

Ohi, K., Nishizawa, D., Muto, Y., Sugiyama, S., Hasegawa, J., Soda, M., Kitaichi, K., Hashimoto, R., Shioiri, T., & Ikeda, K. (2020). Polygenic risk scores for late smoking initiation associated with the risk of schizophrenia. *npj Schizophrenia*, *6*, 36. https://doi.org/10.1038/s41537-020-00126-z

Ohlsson, H., & Kendler, K. S. (2020). Applying Causal Inference Methods in Psychiatric Epidemiology. *JAMA Psychiatry*, *77*(6), 637. https://doi.org/10.1001/jamapsychiatry.2019.3758

Okbay, A., Baselmans, B. M. L., De Neve, J. E., Turley, P., Nivard, M. G., Fontana, M. A., Meddens, S. F. W., Linnér, R. K., Rietveld, C. A., Derringer, J., Gratten, J., Lee, J. J., Liu, J. Z., De Vlaming, R., SAhluwalia, T., Buchwald, J., Cavadino, A., Frazier-Wood, A. C., Furlotte, N. 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

Okbay, A., Wu, Y., Wang, N., Jayashankar, H., Bennett, M., Nehzati, S. M., Sidorenko, J., Kweon, H., Goldman, G., Gjorgjieva, T., Jiang, Y., Hicks, B., Tian, C., Hinds, D. A., Ahlskog, R., Magnusson, P. K. E., Oskarsson, S., Hayward, C., Campbell, A., … Young, A. I. (2022). Polygenic prediction of educational attainment within and between families from genome-wide association analyses in 3 million individuals. *Nature Genetics*, *54*(4), 437-449. https://doi.org/10.1038/s41588-022-01016-z

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

Oldham, M. C., Konopka, G., Iwamoto, K., Langfelder, P., Kato, T., Horvath, S., & Geschwind, D. H. (2008). Functional organization of the transcriptome in human brain. *Nature Neuroscience*, *11*(11), 1271-1282. https://doi.org/10.1038/nn.2207

Ortiz-Medina, M. B., Perea, M., Torales, J., Ventriglio, A., Vitrani, G., Aguilar, L., & Roncero, C. (2018). Cannabis consumption and psychosis or schizophrenia development. *International Journal of Social Psychiatry*, *64*(7), 690-704. https://doi.org/10.1177/0020764018801690

Otowa, T., Hek, K., Lee, M., Byrne, E. M., Mirza, S. S., Nivard, M. G., Bigdeli, T., Aggen, S. H., Adkins, D., Wolen, A., Fanous, A., Keller, M. C., Castelao, E., Kutalik, Z., Der Auwera, S. V., Homuth, G., Nauck, M., Teumer, A., Milaneschi, Y., … 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

Ouzir, M. (2013). Impulsivity in schizophrenia: A comprehensive update. *Aggression and Violent Behavior*, *18*(2), 247-254. https://doi.org/10.1016/j.avb.2012.11.014

Pagan, J. L., Rose, R. J., Viken, R. J., Pulkkinen, L., Kaprio, J., & Dick, D. M. (2006). Genetic and Environmental Influences on Stages of Alcohol Use Across Adolescence and into Young Adulthood. *Behavior Genetics*, *36*(4), 483-497. https://doi.org/10.1007/s10519-006-9062-y

Pain, O., Glanville, K. P., Hagenaars, S. P., Selzam, S., Furtjes, A. E., Gaspar, H. A., Coleman, J. R. I., Rimfeld, K., Breen, G., Plomin, R., Folkersen, L., & Lewis, C. M. (2021). Evaluation of polygenic prediction methodology within a reference-standardized framework. *PLoS Genetics*, *17*(5 May), 1-22. https://doi.org/10.1371/journal.pgen.1009021

Palla, L., & Dudbridge, F. (2015). A Fast Method that Uses Polygenic Scores to Estimate the Variance Explained by Genome-wide Marker Panels and the Proportion of Variants Affecting a Trait. *The American Journal of Human Genetics*, *97*(2), 250-259. https://doi.org/10.1016/j.ajhg.2015.06.005

Palmer, R. H. C., Benca-Bachman, C. E., Huggett, S. B., Bubier, J. A., McGeary, J. E., Ramgiri, N., Srijeyanthan, J., Yang, J., Visscher, P. M., Yang, J., Knopik, V. S., & Chesler, E. J. (2021). Multi-omic and multi-species meta-analyses of nicotine consumption. *Translational Psychiatry*, *11*(1), 98. https://doi.org/10.1038/s41398-021-01231-y

Palmer, R. H. C., Button, T. M., Rhee, S. H., Corley, R. P., Young, S. E., Stallings, M. C., Hopfer, C. J., & Hewitt, J. K. (2012). Genetic etiology of the common liability to drug dependence: Evidence of common and specific mechanisms for DSM-IV dependence

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

symptoms. *Drug and Alcohol Dependence*, *123*(1), S24-S32. https://doi.org/10.1016/j.drugalcdep.2011.12.015

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., Han, J., Hubbard, L., Lynham, A., Mantripragada, K., Rees, E., MacCabe, J. H., McCarroll, S. A., Baune, B. T., Breen, G., … Walters, J. T. R. (2018). Common schizophrenia alleles are enriched in mutation-intolerant genes and in regions under strong background selection. *Nature Genetics*, *50*(3), 381-389. https://doi.org/10.1038/s41588-018-0059-2

Pasman, J. A., Verweij, K. J. H., Gerring, Z., Stringer, S., Sanchez-Roige, S., Treur, J. L., Abdellaoui, A., Nivard, M. G., Baselmans, B. M. L., Ong, J.-S., Ip, H. F., van der Zee, M. D., Bartels, M., Day, F. R., Fontanillas, P., Elson, S. L., de Wit, H., Davis, L. K., MacKillop, J., … 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

Pedersen, M. U., Thomsen, K. R., Heradstveit, O., Skogen, J. C., Hesse, M., & Jones, S. (2018). Externalizing behavior problems are related to substance use in adolescents across six samples from Nordic countries. *European Child and Adolescent Psychiatry*, *27*(12), 1551-1561. https://doi.org/10.1007/s00787-018-1148-6

Pereiro, C., Pino, C., Flórez, G., Arrojo, M., & Becoña, E. (2013). Psychiatric Comorbidity in Patients from the Addictive Disorders Assistance Units of Galicia: The COPSIAD Study. *PLoS ONE*, *8*(6). https://doi.org/10.1371/journal.pone.0066451

Pergola, G., Di Carlo, P., D'Ambrosio, E., Gelao, B., Fazio, L., Papalino, M., Monda, A., Scozia, G., Pietrangelo, B., Attrotto, M., Apud, J. A., Chen, Q., Mattay, V. S., Rampino, A., Caforio, G., Weinberger, D. R., Blasi, G., & Bertolino, A. (2017). DRD2 co-expression network and a related polygenic index predict imaging, behavioral and clinical phenotypes linked to schizophrenia. *Translational Psychiatry*, *7*(1). https://doi.org/10.1038/tp.2016.253

Peterson, R. E., Bigdeli, T. B., Ripke, S., Bacanu, S., Gejman, P. V, Levinson, D. F., Li, Q. S., Rujescu, D., Rietschel, M., Weinberger, D. R., Straub, R. E., Walters, J. T. R., Owen, M. J., O'Donovan, M. C., Mowry, B. J., Ophoff, R. A., Andreassen, O. A., Esko, T., Petryshen, T. L., … Schizophrenia Working Group of the Psychiatric Genomics

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

Consortium. (2021). Genome-wide analyses of smoking behaviors in schizophrenia: Findings from the Psychiatric Genomics Consortium. *Journal of Psychiatric Research*, *137*, 215-224. https://doi.org/10.1016/j.jpsychires.2021.02.027

Peterson, R. E., Kuchenbaecker, K., Walters, R. K., Chen, C. Y., Popejoy, A. B., Periyasamy, S., Lam, M., Iyegbe, C., Strawbridge, R. J., Brick, L., Carey, C. E., Martin, A. R., Meyers, J. L., Su, J., Chen, J., Edwards, A. C., Kalungi, A., Koen, N., Majara, L., … Duncan, L. E. (2019). Genome-wide Association Studies in Ancestrally Diverse Populations: Opportunities, Methods, Pitfalls, and Recommendations. *Cell*, *179*(3), 589-603. https://doi.org/10.1016/j.cell.2019.08.051

Pettersson, E., Larsson, H., & Lichtenstein, P. (2016). Common psychiatric disorders share the same genetic origin: a multivariate sibling study of the Swedish population. *Molecular Psychiatry*, *21*(5), 717-721. https://doi.org/10.1038/mp.2015.116

Philips, M. A., Lilleväli, K., Heinla, I., Luuk, H., Hundahl, C. A., Kongi, K., Vanaveski, T., Tekko, T., Innos, J., & Vasar, E. (2015). Lsamp is implicated in the regulation of emotional and social behavior by use of alternative promoters in the brain. *Brain Structure and Function*, *220*(3), 1381-1393. https://doi.org/10.1007/s00429-014-0732-x

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

Plomin, R., & von Stumm, S. (2022). Polygenic scores: prediction versus explanation. *Molecular Psychiatry*, *27*(1), 49-52. https://doi.org/10.1038/s41380-021-01348-y

Polimanti, R., Peterson, R. E., Ong, J.-S., MacGregor, S., Edwards, A. C., Clarke, T.-K., Frank, J., Gerring, Z., Gillespie, N. A., Lind, P. A., Maes, H. H., Martin, N. G., Mbarek, H., Medland, S. E., Streit, F., Agrawal, A., Edenberg, H. J., Kendler, K. S., Lewis, C. M., … Derks, E. M. (2019). Evidence of causal effect of major depression on alcohol dependence: findings from the psychiatric genomics consortium. *Psychological Medicine*, *49*(07), 1218-1226. https://doi.org/10.1017/S0033291719000667

Polimanti, R., Walters, R. K., Johnson, E. C., McClintick, J. N., Adkins, A. E., Adkins, D. E., Bacanu, S.-A., Bierut, L. J., Bigdeli, T. B., Brown, S., Bucholz, K. K., Copeland, W. E., Costello, E. J., Degenhardt, L., Farrer, L. A., Foroud, T. M., Fox, L., Goate, A. M., Grucza, R., … Gelernter, J. (2020). Leveraging genome-wide data to investigate differences

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

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

Polushina, T., Banerjee, N., Giddaluru, S., Bettella, F., Espeseth, T., Lundervold, A. J., Djurovic, S., Cichon, S., Hoffmann, P., Nöthen, M. M., Steen, V. M., Andreassen, O. A., & Le Hellard, S. (2021). Identification of pleiotropy at the gene level between psychiatric disorders and related traits. *Translational Psychiatry*, *11*(1), 1-9. https://doi.org/10.1038/s41398-021-01530-4

Ponzoni, L., Melzi, G., Marabini, L., Martini, A., Petrillo, G., Teh, M. T., Torres-Perez, J. V., Morara, S., Gotti, C., Braida, D., Brennan, C. H., & Sala, M. (2021). Conservation of mechanisms regulating emotional-like responses on spontaneous nicotine withdrawal in zebrafish and mammals. *Progress in Neuro-Psychopharmacology and Biological Psychiatry*, *111*(April), 110334. https://doi.org/10.1016/j.pnpbp.2021.110334

Privé, F., Arbel, J., & Vilhjálmsson, B. J. (2020). LDpred2: Better, faster, stronger. *Bioinformatics*, *36*(22-23), 5424-5431. https://doi.org/10.1093/bioinformatics/btaa1029

Prochaska, J. J., Das, S., & Young-Wolff, K. C. (2017). Smoking, Mental Illness, and Public Health. *Annual Review of Public Health*, *38*, 165-185. https://doi.org/10.1146/annurev-publhealth-031816-044618

Prochaska, J. J., Hall, S. M., & Bero, L. A. (2008). Tobacco use among individuals with schizophrenia: What role has the tobacco industry played? *Schizophrenia Bulletin*, *34*(3), 555-567. https://doi.org/10.1093/schbul/sbm117

Prom-Wormley, E. C., Ebejer, J., Dick, D. M., & Bowers, M. S. (2017). The genetic epidemiology of substance use disorder: A review. *Drug and Alcohol Dependence*, *180*(June), 241-259. https://doi.org/10.1016/j.drugalcdep.2017.06.040

Pulit, S. L., Stoneman, C., Morris, A. P., Wood, A. R., Glastonbury, C. A., Tyrrell, J., Yengo, L., Ferreira, T., Marouli, E., Ji, Y., Yang, J., Jones, S., Beaumont, R., Croteau-Chonka, D. C., Winkler, T. W., Hattersley, A. T., Loos, R. J. F., Hirschhorn, J. N., Visscher, P. M., … Lindgren, C. M. (2019). Meta-analysis of genome-wide association studies for body fat distribution in 694 649 individuals of European ancestry. *Human Molecular Genetics*, *28*(1), 166-174. https://doi.org/10.1093/hmg/ddy327

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

Purcell, S. M., Wray, N. R., Stone, J. L., Visscher, P. M., O'Donovan, M. C., Sullivan, P. F., Ruderfer, D. M., McQuillin, A., Morris, D. W., Oĝdushlaine, C. T., Corvin, A., Holmans, P. A., Oĝdonovan, M. C., MacGregor, S., Gurling, H., Blackwood, D. H. R., Craddock, N. J., Gill, M., Hultman, C. M., … Sklar, P. (2009). Common polygenic variation contributes to risk of schizophrenia and bipolar disorder. *Nature*, *460*(7256), 748-752. https://doi.org/10.1038/nature08185

Quach, B. C., Bray, M. J., Gaddis, N. C., Liu, M., Palviainen, T., Minica, C. C., Zellers, S., Sherva, R., Aliev, F., Nothnagel, M., Young, K. A., Marks, J. A., Young, H., Carnes, M. U., Guo, Y., Waldrop, A., Sey, N. Y. A., Landi, M. T., McNeil, D. W., … Hancock, D. B. (2020). Expanding the genetic architecture of nicotine dependence and its shared genetics with multiple traits. *Nature Communications*, *11*(1). https://doi.org/10.1038/s41467-020-19265-z

Quinn, P. D., Rickert, M. E., Weibull, C. E., Johansson, A. L. V., Lichtenstein, P., Almqvist, C., Larsson, H., Iliadou, A. N., & D'Onofrio, B. M. (2017). Association between maternal smoking during pregnancy and severe mental illness in offspring. *JAMA Psychiatry*, *74*(6), 589-596. https://doi.org/10.1001/jamapsychiatry.2017.0456

Reddy, L. F., Lee, J., Davis, M. C., Altshuler, L., Glahn, D. C., Miklowitz, D. J., & Green, M. F. (2014). Impulsivity and risk taking in bipolar disorder and schizophrenia. *Neuropsychopharmacology*, *39*(2), 456-463. https://doi.org/10.1038/npp.2013.218

Reed, A. C., Harris, J. G., & Olincy, A. (2016). Schizophrenia, smoking status, and performance on the matrics Cognitive Consensus Battery. *Psychiatry Research*, *246*, 1-8. https://doi.org/10.1016/j.psychres.2016.08.062

Reed, Z. E., Wootton, R. E., & Munafò, M. R. (2022). Using Mendelian randomization to explore the gateway hypothesis: possible causal effects of smoking initiation and alcohol consumption on substance use outcomes. *Addiction*, *117*(3), 741-750. https://doi.org/10.1111/add.15673

Reginsson, G. W., Ingason, A., Euesden, J., Bjornsdottir, G., Olafsson, S., Sigurdsson, E., Oskarsson, H., Tyrfingsson, T., Runarsdottir, V., Hansdottir, I., Steinberg, S., Stefansson, H., Gudbjartsson, D. F., Thorgeirsson, T. E., & Stefansson, K. (2018). Polygenic risk scores for schizophrenia and bipolar disorder associate with addiction. *Addiction Biology*, *23*(1), 485-492. https://doi.org/10.1111/adb.12496

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

Reid, C. H., Patrick, P. H., Rytwinski, T., Taylor, J. J., Willmore, W. G., Reesor, B., & Cooke, S. J. (2022). An updated review of cold shock and cold stress in fish. *Journal of Fish Biology*, *100*(5), 1102-1137. https://doi.org/10.1111/jfb.15037

Restrepo-Lozano, J. M., Flores, C., & Silveira, P. P. (2023). Novel Functional Genomics Approaches Bridging Neuroscience and Psychiatry. *Biological Psychiatry Global Open Science*, *3*(3), 351-361. https://doi.org/10.1016/j.bpsgos.2022.07.005

Restrepo-Lozano, J. M., Pokhvisneva, I., Wang, Z., Patel, S., Meaney, M. J., Silveira, P. P., & Flores, C. (2022). Corticolimbic DCC gene co-expression networks as predictors of impulsivity in children. *Molecular Psychiatry*, *27*(6), 2742-2750. https://doi.org/10.1038/s41380-022-01533-7

Rey-Brandariz, J., Pérez-Ríos, M., Santiago-Pérez, M. I., Varela-Lema, L., Giraldo-Osorio, A., Mourino, N., & Ruano-Ravina, A. (2021). Mortalidad atribuida al consumo de tabaco en España: Revisión sistemática. *Adicciones*, *35*(2), 185. https://doi.org/10.20882/adicciones.1619

Richmond, R. C., & Smith, G. D. (2022). Mendelian Randomization: Concepts and Scope. *Cold Spring Harbor Perspectives in Medicine*, *12*(1). https://doi.org/10.1101/cshperspect.a040501

Ringin, E., Cropley, V., Zalesky, A., Bruggemann, J., Sundram, S., Weickert, C. S., Weickert, T. W., Bousman, C. A., Pantelis, C., & Van Rheenen, T. E. (2022). The impact of smoking status on cognition and brain morphology in schizophrenia spectrum disorders. *Psychological Medicine*, *52*(14), 3097-3115. https://doi.org/10.1017/S0033291720005152

Robin, X., Turck, N., Hainard, A., Tiberti, N., Lisacek, F., Sanchez, J.-C., & Müller, M. (2011). pROC: an open-source package for R and S+ to analyze and compare ROC curves. *BMC Bioinformatics*, *12*, 77. https://doi.org/10.1186/1471-2105-12-77

Rodriguez-López, J., Arrojo, M., Paz, E., Páramo, M., & Costas, J. (2020). Identification of relevant hub genes for early intervention at gene coexpression modules with altered predicted expression in schizophrenia. *Progress in Neuro-Psychopharmacology and Biological Psychiatry*, *98*(November 2019), 109815. https://doi.org/10.1016/j.pnpbp.2019.109815

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

Rosenström, T., Gjerde, L. C., Krueger, R. F., Aggen, S. H., Czajkowski, N. O., Gillespie, N. A., Kendler, K. S., Reichborn-Kjennerud, T., Torvik, F. A., & Ystrom, E. (2019). Joint factorial structure of psychopathology and personality. *Psychological Medicine*, *49*(13), 2158-2167. https://doi.org/10.1017/S0033291718002982

Sacco, K. A., Termine, A., Seyal, A., Dudas, M. M., Vessicchio, J. C., Krishnan-Sarin, S., Jatlow, P. I., Wexler, B. E., & George, T. P. (2005). Effects of Cigarette Smoking on Spatial Working Memory and Attentional Deficits in Schizophrenia. *Archives of General Psychiatry*, *62*(6), 649-659. https://doi.org/10.1001/archpsyc.62.6.649

Sagud, M., Mihaljevic Peles, A., & Pivac, N. (2019). Smoking in schizophrenia: Recent findings about an old problem. *Current Opinion in Psychiatry*, *32*(5), 402-408. https://doi.org/10.1097/YCO.0000000000000529

Saha, S., Scott, J. G., Varghese, D., Degenhardt, L., Slade, T., & McGrath, J. J. (2011). The association between delusional-like experiences, and tobacco, alcohol or cannabis use: A nationwide population-based survey. *BMC Psychiatry*, *11*. https://doi.org/10.1186/1471-244X-11-202

Sallis, H. M., Davey Smith, G., & Munafò, M. R. (2019). Cigarette smoking and personality: interrogating causality using Mendelian randomisation. *Psychological Medicine*, *49*(13), 2197-2205. https://doi.org/10.1017/S0033291718003069

Sánchez-Gutiérrez, T., García-Portilla, M. P., Parellada, M., Bobes, J., Calvo, A., Moreno-Izco, L., González-Pinto, A., Lobo, A., de la Serna, E., Cabrera, B., Torrent, C., Roldán, L., Sanjuan, J., Ibáñez, Á., Sánchez-Torres, A. M., Corripio, I., Bernardo, M., Cuesta, M. J., Vieta, E., … Balanzá, V. (2018). Smoking does not impact social and non-social cognition in patients with first episode psychosis. *Schizophrenia Research*, *199*, 64-74. https://doi.org/10.1016/j.schres.2018.03.025

Sanchez-Roige, S., Cox, N. J., Johnson, E. O., Hancock, D. B., & Davis, L. K. (2021). Alcohol and cigarette smoking consumption as genetic proxies for alcohol misuse and nicotine dependence. *Drug and Alcohol Dependence*, *221*, 108612. https://doi.org/10.1016/j.drugalcdep.2021.108612

Sanchez-Roige, S., Palmer, A. A., Fontanillas, P., Elson, S. L., Adams, M. J., Howard, D. M., Edenberg, H. J., Davies, G., Crist, R. C., Deary, I. J., McIntosh, A. M., Clarke, T. K., Elson, L., Fontanillas, P., Furlotte, N. A., Hinds, D. A., Huber, K. E., Kleinman, A.,

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

Litterman, N. K., … 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

Saunders, G. R. B., Wang, X., Chen, F., Jang, S.-K., Liu, M., Wang, C., Gao, S., Jiang, Y., Khunsriraksakul, C., Otto, J. M., Addison, C., Akiyama, M., Albert, C. M., Aliev, F., Alonso, A., Arnett, D. K., Ashley-Koch, A. E., Ashrani, A. A., Barnes, K. C., … Vrieze, S. (2022). Genetic diversity fuels gene discovery for tobacco and alcohol use. *Nature*, *612*(7941), 720-724. https://doi.org/10.1038/s41586-022-05477-4

Saunders, J. B., Aasland, O. G., Babor, T. F., De la Fuente, J. R., & Grant, M. (1993). Development of the Alcohol Use Disorders Identification Test (AUDIT): WHO Collaborative Project on Early Detection of Persons with Harmful Alcohol Consumption-II. *Addiction*, *88*(6), 791-804. https://doi.org/10.1111/j.1360-0443.1993.tb02093.x

Saunders, L. M., Srivatsan, S. R., Duran, M., Dorrity, M. W., Ewing, B., Linbo, T. H., Shendure, J., Raible, D. W., Moens, C. B., Kimelman, D., & Trapnell, C. (2023). Embryo-scale reverse genetics at single-cell resolution. *Nature*, *623*(7988), 782-791. https://doi.org/10.1038/s41586-023-06720-2

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., Grasby, K. L., Hammerschlag, A. R., Kaminski, J. A., Karlsson, R., Krapohl, E., Lam, M., Nygaard, M., Reynolds, C. A., Trampush, J. W., … 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

Schaid, D. J., Chen, W., & Larson, N. B. (2018). From genome-wide associations to candidate causal variants by statistical fine-mapping. *Nature Reviews Genetics*, *19*(8), 491-504. https://doi.org/10.1038/s41576-018-0016-z

Schizophrenia Working Group of the Psychiatric Genomics Consortium, Ripke, S., Neale, B. M., Corvin, A., Walters, J. T. R., Farh, K. H., Holmans, P. A., Lee, P., Bulik-Sullivan, B., Collier, D. A., Huang, H., Pers, T. H., Agartz, I., Agerbo, E., Albus, M., Alexander, M., Amin, F., Bacanu, S. A., Begemann, M., … O'Donovan, M. C. (2014). Biological insights from 108 schizophrenia-associated genetic loci. *Nature*, *511*(7510), 421-427.

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

https://doi.org/10.1038/nature13595

Schizophrenia Working Group of the Psychiatric Genomics Consortium, Trubetskoy, V., Pardiñas, A. F., Qi, T., Panagiotaropoulou, G., Awasthi, S., Bigdeli, T. B., Bryois, J., Chen, C.-Y., Dennison, C. A., Hall, L. S., Lam, M., Watanabe, K., Frei, O., Ge, T., Harwood, J. C., Koopmans, F., Magnusson, S., Richards, A. L., … van Os, J. (2022). Mapping genomic loci implicates genes and synaptic biology in schizophrenia. *Nature*, *604*(7906), 502-508. https://doi.org/10.1038/s41586-022-04434-5

Schulte, M. T., & Hser, Y.-I. (2013). Substance Use and Associated Health Conditions throughout the Lifespan. *Public Health Reviews*, *35*(2), 3. https://doi.org/10.1007/BF03391702

Selzam, S., Coleman, J. R. I., Caspi, A., Moffitt, T. E., & Plomin, R. (2018). A polygenic p factor for major psychiatric disorders. *Translational Psychiatry*, *8*(1), 205. https://doi.org/10.1038/s41398-018-0217-4

Sheals, K., Tombor, I., McNeill, A., & Shahab, L. (2016). A mixed-method systematic review and meta-analysis of mental health professionals' attitudes toward smoking and smoking cessation among people with mental illnesses. *Addiction*, *111*(9), 1536-1553. https://doi.org/10.1111/add.13387

Shrine, N., Guyatt, A. L., Erzurumluoglu, A. M., Jackson, V. E., Hobbs, B. D., Melbourne, C. A., Batini, C., Fawcett, K. A., Song, K., Sakornsakolpat, P., Li, X., Boxall, R., Reeve, N. F., Obeidat, M., Zhao, J. H., Wielscher, M., Weiss, S., Kentistou, K. A., Cook, J. P., … Wain, L. V. (2019). New genetic signals for lung function highlight pathways and chronic obstructive pulmonary disease associations across multiple ancestries. *Nature Genetics*, *51*(3), 481-493. https://doi.org/10.1038/s41588-018-0321-7

Sirugo, G., Williams, S. M., & Tishkoff, S. A. (2019). The Missing Diversity in Human Genetic Studies. *Cell*, *177*(1), 26-31. https://doi.org/10.1016/j.cell.2019.02.048

Smedley, D., Haider, S., Ballester, B., Holland, R., London, D., Thorisson, G., & Kasprzyk, A. (2009). BioMart - Biological queries made easy. *BMC Genomics*, *10*, 1-12. https://doi.org/10.1186/1471-2164-10-22

USC DE COMPOSTELA

Smith, A. H., Ovesen, P. L., Skeldal, S., Yeo, S., Jensen, K. P., Olsen, D., Diazgranados, N., Zhao, H., Farrer, L. A., Goldman, D., Glerup, S., Kranzler, H. R., Nykjær, A., & Gelernter,

J. (2018). Risk Locus Identification Ties Alcohol Withdrawal Symptoms to SORCS 2. *Alcoholism: Clinical and Experimental Research*, *42*(12), 2337-2348. https://doi.org/10.1111/acer.13890

Smith, R. C., Singh, A., Infante, M., Khandat, A., & Kloos, A. (2002). Effects of cigarette smoking and nicotine nasal spray on psychiatric symptoms and cognition in schizophrenia. *Neuropsychopharmacology*, *27*(3), 479-497. https://doi.org/10.1016/S0893-133X(02)00324-X

Socrates, A., Maxwell, J., Glanville, K. P., Di Forti, M., Murray, R. M., Vassos, E., & O'Reilly, P. F. (2021). Investigating the effects of genetic risk of schizophrenia on behavioural traits. *npj Schizophrenia*, *7*(1). https://doi.org/10.1038/s41537-020-00131-2

Soler Artigas, M., Sánchez-Mora, C., Rovira, P., Richarte, V., Garcia-Martínez, I., Pagerols, M., Demontis, D., Stringer, S., Vink, J. M., Børglum, A. D., Neale, B. M., Franke, B., Faraone, S. V., Casas, M., Ramos-Quiroga, J. A., & Ribasés, M. (2020). Attention-deficit/hyperactivity disorder and lifetime cannabis use: genetic overlap and causality. *Molecular Psychiatry*, *25*(10), 2493-2503. https://doi.org/10.1038/s41380-018-0339-3

Sollis, E., Mosaku, A., Abid, A., Buniello, A., Cerezo, M., Gil, L., Groza, T., Güneş, O., Hall, P., Hayhurst, J., Ibrahim, A., Ji, Y., John, S., Lewis, E., Macarthur, J. A. L., Mcmahon, A., Osumi-Sutherland, D., Panoutsopoulou, K., Pendlington, Z., … Harris, L. W. (2023). The NHGRI-EBI GWAS Catalog: knowledgebase and deposition resource. *Nucleic Acids Research*, *51*(1 D), D977-D985. https://doi.org/10.1093/nar/gkac1010

Solovieff, N., Cotsapas, C., Lee, P. H., Purcell, S. M., & Smoller, J. W. (2013). Pleiotropy in complex traits: challenges and strategies. *Nature Reviews Genetics*, *14*(7), 483-495. https://doi.org/10.1038/nrg3461

Song, S., Jiang, W., Hou, L., & Zhao, H. (2020). Leveraging effect size distributions to improve polygenic risk scores derived from summary statistics of genome-wide association studies. *PLOS Computational Biology*, *16*(2), e1007565. https://doi.org/10.1371/journal.pcbi.1007565

Song, W., Lin, G. N., Yu, S., & Zhao, M. (2022). Genome-wide identification of the shared genetic basis of cannabis and cigarette smoking and schizophrenia implicates NCAM1 and neuronal abnormality. *Psychiatry Research*, *310*. https://doi.org/10.1016/j.psychres.2022.114453

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

Stewart, A. M., Braubach, O., Spitsbergen, J., Gerlai, R., & Kalueff, A. V. (2014). Zebrafish models for translational neuroscience research: From tank to bedside. *Trends in Neurosciences*, *37*(5), 264-278. https://doi.org/10.1016/j.tins.2014.02.011

Stewart, A. M., Ullmann, J. F. P., Norton, W. H. J., Parker, M. O., Brennan, C. H., Gerlai, R., & Kalueff, A. V. (2015). Molecular psychiatry of zebrafish. *Molecular Psychiatry*, *20*(1), 2-17. https://doi.org/10.1038/mp.2014.128

Stilo, S. A., & Murray, R. M. (2019). Non-Genetic Factors in Schizophrenia. *Current Psychiatry Reports*, *21*(10), 100. https://doi.org/10.1007/s11920-019-1091-3

Stolz, P. A., Wehring, H. J., Liu, F., Love, R. C., Ellis, M., DiPaula, B. A., & Kelly, D. L. (2019). Effects of Cigarette Smoking and Clozapine Treatment on 20-Year All-Cause & Cardiovascular Mortality in Schizophrenia. *Psychiatric Quarterly*, *90*(2), 351-359. https://doi.org/10.1007/s11126-018-9621-4

Strålin, P., & Hetta, J. (2020). Substance use disorders before, at and after first episode psychosis hospitalizations in a young national Swedish cohort. *Drug and Alcohol Dependence*, *209*(October 2019), 107919. https://doi.org/10.1016/j.drugalcdep.2020.107919

Strawbridge, R. J., Ward, J., Cullen, B., Tunbridge, E. M., Hartz, S., Bierut, L., Horton, A., Bailey, M. E. S., Graham, N., Ferguson, A., Lyall, D. M., Mackay, D., Pidgeon, L. M., Cavanagh, J., Pell, J. P., O'Donovan, M., Escott-Price, V., Harrison, P. J., & Smith, D. J. (2018). Genome-wide analysis of self-reported risk-taking behaviour and cross-disorder genetic correlations in the UK Biobank cohort. *Translational Psychiatry*, *8*(1), 1-11. https://doi.org/10.1038/s41398-017-0079-1

Stubbs, B., Vancampfort, D., Bobes, J., De Hert, M., & Mitchell, A. J. (2015). How can we promote smoking cessation in people with schizophrenia in practice? A clinical overview. *Acta Psychiatrica Scandinavica*, *132*(2), 122-130. https://doi.org/10.1111/acps.12412

Sugathan, A., Biagioli, M., Golzio, C., Erdin, S., Blumenthal, I., Manavalan, P., Ragavendran, A., Brand, H., Lucente, D., Miles, J., Sheridan, S. D., Stortchevoi, A., Kellis, M., Haggarty, S. J., Katsanis, N., Gusella, J. F., & Talkowski, M. E. (2014). CHD8 regulates neurodevelopmental pathways associated with autism spectrum disorder in neural progenitors. *Proceedings of the National Academy of Sciences of the United States of America*, *111*(42), E4468-E4477. https://doi.org/10.1073/pnas.1405266111

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

Sullivan, P. F., Agrawal, A., Bulik, C. M., Andreassen, O. A., Børglum, A. D., Breen, G., Cichon, S., Edenberg, H. J., Faraone, S. V., Gelernter, J., Mathews, C. A., Nievergelt, C. M., Smoller, J. W., O'Donovan, M. C., Daly, M., Gill, M., Kelsoe, J., Koenen, K., Levinson, D., … Sklar, P. (2018). Psychiatric genomics: An update and an Agenda. *American Journal of Psychiatry*, *175*(1), 15-27. https://doi.org/10.1176/appi.ajp.2017.17030283

Sullivan, P. F., & Geschwind, D. H. (2019). Defining the Genetic, Genomic, Cellular, and Diagnostic Architectures of Psychiatric Disorders. En *Cell* (Vol. 177, Número 1, pp. 162-183). Cell Press. https://doi.org/10.1016/j.cell.2019.01.015

Swendsen, J., Conway, K. P., Degenhardt, L., Glantz, M., Jin, R., Merikangas, K. R., Sampson, N., & Kessler, R. C. (2010). Mental disorders as risk factors for substance use, abuse and dependence: Results from the 10-year follow-up of the National Comorbidity Survey. *Addiction*, *105*(6), 1117-1128. https://doi.org/10.1111/j.1360-0443.2010.02902.x

Szerman, N., Torrens, M., Maldonado, R., Balhara, Y. P. S., Salom, C., Maremmani, I., Sher, L., Didia-Attas, J., Chen, J., & Baler, R. (2022). Addictive and other mental disorders: a call for a standardized definition of dual disorders. *Translational Psychiatry*, *12*(1). https://doi.org/10.1038/s41398-022-02212-5

Tawa, E. A., Hall, S. D., & Lohoff, F. W. (2016). Overview of the genetics of alcohol use disorder. *Alcohol and Alcoholism*, *51*(5), 507-514. https://doi.org/10.1093/alcalc/agw046

Taylor, G. M. J., Itani, T., Thomas, K. H., Rai, D., Jones, T., Windmeijer, F., Martin, R. M., Munafò, M. R., Davies, N. M., & Taylor, A. E. (2020). Prescribing Prevalence, Effectiveness, and Mental Health Safety of Smoking Cessation Medicines in Patients With Mental Disorders. *Nicotine & tobacco research : official journal of the Society for Research on Nicotine and Tobacco*, *22*(1), 48-57. https://doi.org/10.1093/ntr/ntz072

Taylor, G. M. J., & Munafò, M. R. (2018). What about treatment of smoking to improve survival and reduce depression? *The Lancet Psychiatry*, *5*(6), 464. https://doi.org/10.1016/S2215-0366(18)30132-9

Taylor, G. M. J., & Munafò, M. R. (2019). Does smoking cause poor mental health? *The Lancet Psychiatry*, *6*(1), 2-3. https://doi.org/10.1016/S2215-0366(18)30459-0

Taylor, G., McNeill, A., Girling, A., Farley, A., Lindson-Hawley, N., & Aveyard, P. (2014).

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

Change in mental health after smoking cessation: Systematic review and meta-analysis. *BMJ (Online)*, *348*(February), 1-22. https://doi.org/10.1136/bmj.g1151

Taylor, G., Taylor, A., Munafò, M. R., McNeill, A., & Aveyard, P. (2015). Does smoking reduction worsen mental health? A comparison of two observational approaches. *BMJ Open*, *5*(5), 1-6. https://doi.org/10.1136/bmjopen-2015-007812

Terry, A. V., Callahan, P. M., & Hernandez, C. M. (2015). Nicotinic ligands as multifunctional agents for the treatment of neuropsychiatric disorders. *Biochemical Pharmacology*, *97*(4), 388-398. https://doi.org/10.1016/j.bcp.2015.07.027

Therriault, D., Lemelin, J. P., Toupin, J., Martin-Storey, A., & Déry, M. (2024). Associations between externalizing behavior problems and risky sexual behaviors in adolescence: Attachment as a mediator. *Journal of Adolescence*, *96*(2), 394-410. https://doi.org/10.1002/jad.12285

Thorsell, A. (2013). The μ-Opioid receptor and treatment response to naltrexone. *Alcohol and Alcoholism*, *48*(4), 402-408. https://doi.org/10.1093/alcalc/agt030

Tian, F.-Y., Hivert, M.-F., Wen, X., Xie, C., Niu, Z., Fan, L., Gillman, M. W., & Chen, W.-Q. (2017). Tissue differences in DNA methylation changes at AHRR in full term low birth weight in maternal blood, placenta and cord blood in Chinese. *Placenta*, *52*, 49-57. https://doi.org/10.1016/j.placenta.2017.02.017

Treur, J. L., Demontis, D., Smith, G. D., Sallis, H., Richardson, T. G., Wiers, R. W., Børglum, A. D., Verweij, K. J. H., & Munafò, M. R. (2019). Investigating causality between liability to ADHD and substance use, and liability to substance use and ADHD risk, using Mendelian randomization. *Addiction Biology*, *26*(1), e12849. https://doi.org/10.1111/adb.12849

Treur, J. L., Munafò, M. R., Logtenberg, E., Wiers, R. W., & Verweij, K. J. H. (2021). Using Mendelian randomization analysis to better understand the relationship between mental health and substance use: a systematic review. *Psychological Medicine*, *51*(10), 1593-1624. https://doi.org/10.1017/S003329172100180X

Tseng, Y.-C., Chen, R.-D., Lucassen, M., Schmidt, M. M., Dringen, R., Abele, D., & Hwang, P.-P. (2011). Exploring Uncoupling Proteins and Antioxidant Mechanisms under Acute Cold Exposure in Brains of Fish. *PLoS ONE*, *6*(3), e18180.

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

https://doi.org/10.1371/journal.pone.0018180

Tsoi, D. T., Porwal, M., & Webster, A. C. (2010). Interventions for smoking cessation and reduction in individuals with schizophrenia. En D. T. Tsoi (Ed.), *Cochrane Database of Systematic Reviews*. John Wiley & Sons, Ltd. https://doi.org/10.1002/14651858.CD007253.pub2

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., Magnusson, P., Oskarsson, S., Johannesson, M., Visscher, P. M., Laibson, D., Cesarini, D., Neale, B. M., Benjamin, D. J., Agee, M., … Pitts, S. J. (2018). Multi-trait analysis of genome-wide association summary statistics using MTAG. *Nature Genetics*, *50*(2), 229-237. https://doi.org/10.1038/s41588-017-0009-4

Uffelmann, E., & Posthuma, D. (2021). Emerging Methods and Resources for Biological Interrogation of Neuropsychiatric Polygenic Signal. *Biological Psychiatry*, *89*(1), 41-53. https://doi.org/10.1016/j.biopsych.2020.05.022

Uhl, G. R., Koob, G. F., & Cable, J. (2019). The neurobiology of addiction. *Annals of the New York Academy of Sciences*, *1451*(1), 5-28. https://doi.org/10.1111/nyas.13989

van den Berg, S. M., de Moor, M. H. M., Verweij, K. J. H., Krueger, R. F., Luciano, M., Arias Vasquez, A., Matteson, L. K., Derringer, J., Esko, T., Amin, N., Gordon, S. D., Hansell, N. K., Hart, A. B., Seppälä, I., Huffman, J. E., Konte, B., Lahti, J., Lee, M., Miller, M., … 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

Van Gastel, W. A., MacCabe, J. H., Schubart, C. D., Vreeker, A., Tempelaar, W., Kahn, R. S., & Boks, M. P. M. (2013). Cigarette smoking and cannabis use are equally strongly associated with psychotic-like experiences:A cross-sectional study in 1929 young adults. *Psychological Medicine*, *43*(11), 2393-2401. https://doi.org/10.1017/S0033291713000202

Vaucher, J., Keating, B. J., Lasserre, A. M., Gan, W., Lyall, D. M., Ward, J., Smith, D. J., Pell, J. P., Sattar, N., Paré, G., & Holmes, M. V. (2018). Cannabis use and risk of schizophrenia: A Mendelian randomization study. *Molecular Psychiatry*, *23*(5), 1287-1292. https://doi.org/10.1038/mp.2016.252

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

Vermeulen, J. M., Schirmbeck, F., Blankers, M., Van Tricht, M., Bruggeman, R., Van Den Brink, W., De Haan, L., Van Amelsvoort, T., Alizadeh, B. Z., Bartels-Velthuis, A. A., Van Beveren, N. J., Cahn, W., Delespaul, P., Meijer, C. J., Myin-Germeys, I., Kahn, R. S., Simons, C. J. P., Van Haren, N. E., Van Os, J., & Van Winkel, R. (2018). Association between smoking behavior and cognitive functioning in patients with psychosis, siblings, and healthy control subjects: Results from a prospective 6-year follow-up study. *American Journal of Psychiatry*, *175*(11), 1121-1128. https://doi.org/10.1176/appi.ajp.2018.18010069

Vie, G. Å., Wootton, R. E., Bjørngaard, J. H., Åsvold, B. O., Taylor, A. E., Gabrielsen, M. E., Davey Smith, G., Romundstad, P. R., & Munafò, M. R. (2019). The effect of smoking intensity on all-cause and cause-specific mortality - A Mendelian randomization analysis. *International Journal of Epidemiology*, *48*(5), 1438-1446. https://doi.org/10.1093/ije/dyz081

Vilar-Ribó, L., Sánchez-Mora, C., Rovira, P., Richarte, V., Corrales, M., Fadeuilhe, C., Arribas, L., Casas, M., Ramos-Quiroga, J. A., Ribasés, M., & Soler Artigas, M. (2020). Genetic overlap and causality between substance use disorder and attention-deficit and hyperactivity disorder. *American Journal of Medical Genetics Part B: Neuropsychiatric Genetics*. https://doi.org/10.1002/ajmg.b.32827

Vink, J. M., Treur, J. L., Pasman, J. A., & Schellekens, A. (2020). Investigating genetic correlation and causality between nicotine dependence and ADHD in a broader psychiatric context. *American Journal of Medical Genetics Part B: Neuropsychiatric Genetics*. https://doi.org/10.1002/ajmg.b.32822

Vink, J. M., Willemsen, G., & Boomsma, D. I. (2005). Heritability of smoking initiation and nicotine dependence. *Behavior Genetics*, *35*(4), 397-406. https://doi.org/10.1007/s10519-004-1327-8

Voineagu, I., Wang, X., Johnston, P., Lowe, J. K., Tian, Y., Horvath, S., Mill, J., Cantor, R. M., Blencowe, B. J., & Geschwind, D. H. (2011). Transcriptomic analysis of autistic brain reveals convergent molecular pathology. *Nature*, *474*(7351), 380-386. https://doi.org/10.1038/nature10110

Wainberg, M., Jacobs, G. R., di Forti, M., & Tripathy, S. J. (2021). Cannabis, schizophrenia genetic risk, and psychotic experiences: a cross-sectional study of 109,308 participants from the UK Biobank. *Translational Psychiatry*, *11*(1). https://doi.org/10.1038/s41398-

USC DE SANTIAGO DE COMPOSTELA

021-01330-w

Waldman, I. D., Poore, H. E., Luningham, J. M., & Yang, J. (2020). Testing structural models of psychopathology at the genomic level. *World Psychiatry*, *19*(3), 350-359. https://doi.org/10.1002/wps.20772

Wallace, C. (2020). Eliciting priors and relaxing the single causal variant assumption in colocalisation analyses. *PLoS Genetics*, *16*(4), e1008720. https://doi.org/10.1371/journal.pgen.1008720

Walters, R. K., Polimanti, R., Johnson, E. C., McClintick, J. N., Adams, M. J., Adkins, A. E., Aliev, F., Bacanu, S. A., Batzler, A., Bertelsen, S., Biernacka, J. M., Bigdeli, T. B., Chen, L. S., Clarke, T. K., Chou, Y. L., Degenhardt, F., Docherty, A. R., Edwards, A. C., Fontanillas, P., … 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

Wang, L., Babushkin, N., Liu, Z., & Liu, X. (2024). Trans-eQTL mapping in gene sets identifies network effects of genetic variants. *Cell Genomics*, *4*(4), 100538. https://doi.org/10.1016/j.xgen.2024.100538

Wang, L., Liu, F., Fang, Y., Ma, J., Wang, J., Qu, L., Yang, Q., Wu, W., Jin, L., & Sun, D. (2023). Advances in Zebrafish as a Comprehensive Model of Mental Disorders. *Depression and Anxiety*, *2023*, 1-48. https://doi.org/10.1155/2023/6663141

Ware, J. J., Van den bree, M. B. M., & Munafò, M. R. (2011). Association of the CHRNA5-A3-B4 gene cluster with heaviness of smoking: A meta-analysis. *Nicotine and Tobacco Research*, *13*(12), 1167-1175. https://doi.org/10.1093/ntr/ntr118

Watanabe, K., Taskesen, E., Van Bochoven, A., & Posthuma, D. (2017). Functional mapping and annotation of genetic associations with FUMA. *Nature Communications*, *8*(1), 1-10. https://doi.org/10.1038/s41467-017-01261-5

Watson, H. J., Yilmaz, Z., Thornton, L. M., Hübel, C., Coleman, J. R. I., Gaspar, H. A., Bryois, J., Hinney, A., Leppä, V. M., Mattheisen, M., Medland, S. E., Ripke, S., Yao, S., Giusti-Rodríguez, P., Hanscombe, K. B., Purves, K. L., Adan, R. A. H., Alfredsson, L., Ando, T., … 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-

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

1214. https://doi.org/10.1038/s41588-019-0439-2

Weiser, M., Reichenberg, A., Grotto, I., Yasvitzky, R., Rabinowitz, J., Lubin, G., Nahon, D., Knobler, H. Y., & Davidson, M. (2004). Higher Rates of Cigarette Smoking in Male Adolescents Before the Onset of Schizophrenia: A Historical-Prospective Cohort Study. *American Journal of Psychiatry*, *161*(7), 1219-1223. https://doi.org/10.1176/appi.ajp.161.7.1219

Whiteford, H. A., Ferrari, A. J., Degenhardt, L., Feigin, V., & Vos, T. (2015). The global burden of mental, neurological and substance use disorders: An analysis from the global burden of disease study 2010. *PLoS ONE*, *10*(2), 1-14. https://doi.org/10.1371/journal.pone.0116820

Wilkerson, J. L., Deba, F., Crowley, M. L., Hamouda, A. K., & McMahon, L. R. (2020). Advances in the In vitro and In vivo Pharmacology of Alpha4beta2 Nicotinic Receptor Positive Allosteric Modulators. *Neuropharmacology*, *168*, 108008. https://doi.org/10.1016/j.neuropharm.2020.108008

Wium-Andersen, M. K., Ørsted, D. D., & Nordestgaard, B. G. (2015). Tobacco smoking is causally associated with antipsychotic medication use and schizophrenia, but not with antidepressant medication use or depression. *International Journal of Epidemiology*, *44*(2), 566-577. https://doi.org/10.1093/ije/dyv090

Wolfe, R. M., Reeves, L. E., Gibson, L. E., Cooper, S., & Ellman, L. M. (2017). Attenuated positive psychotic symptoms in relation to cigarette smoking in a nonclinical population. *Nicotine and Tobacco Research*, *19*(1), 124-128. https://doi.org/10.1093/ntr/ntw240

Wong, A. K., Sealfon, R. S. G., Theesfeld, C. L., & Troyanskaya, O. G. (2021). Decoding disease: from genomes to networks to phenotypes. *Nature Reviews Genetics*, *22*(12), 774-790. https://doi.org/10.1038/s41576-021-00389-x

Wootton, R. E., Richmond, R. C., Stuijfzand, B. G., Lawn, R. B., Sallis, H. M., Taylor, G. M. J., Hemani, G., Jones, H. J., Zammit, S., Davey Smith, G., & 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

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

Wray, N. R., Lee, S. H., Mehta, D., Vinkhuyzen, A. A. E., Dudbridge, F., & Middeldorp, C. M.

(2014). Research Review: Polygenic methods and their application to psychiatric traits. *Journal of Child Psychology and Psychiatry*, *55*(10), 1068-1087. https://doi.org/10.1111/jcpp.12295

Wray, N. R., Ripke, S., Mattheisen, M., Trzaskowski, M., Byrne, E. M., Abdellaoui, A., Adams, M. J., Agerbo, E., Air, T. M., Andlauer, T. M. F., Bacanu, S. A., Bækvad-Hansen, M., Beekman, A. F. T., Bigdeli, T. B., Binder, E. B., Blackwood, D. R. H., Bryois, J., Buttenschøn, H. N., Bybjerg-Grauholm, J., … Sullivan, P. F. (2018). Genome-wide association analyses identify 44 risk variants and refine the genetic architecture of major depression. *Nature Genetics*, *50*(5), 668-681. https://doi.org/10.1038/s41588-018-0090-3

Yang, J., Lee, S. H., Goddard, M. E., & Visscher, P. M. (2011). GCTA: A tool for genome-wide complex trait analysis. *American Journal of Human Genetics*, *88*(1), 76-82. https://doi.org/10.1016/j.ajhg.2010.11.011

Yang, R., Dai, Z., Chen, S., & Chen, L. (2011). MicroRNA-mediated gene regulation plays a minor role in the transcriptomic plasticity of cold-acclimated Zebrafish brain tissue. *BMC Genomics*, *12*(1), 605. https://doi.org/10.1186/1471-2164-12-605

Yao, Y., Xu, Y., Cai, Z., Liu, Q., Ma, Y., Li, A. N., Payne, T. J., & Li, M. D. (2021). Determination of shared genetic etiology and possible causal relations between tobacco smoking and depression. *Psychological Medicine*, *51*(11), 1870-1879. https://doi.org/10.1017/S003329172000063X

Young, A. I. (2019). Solving the missing heritability problem. *PLoS Genetics*, *15*(6), 1-7. https://doi.org/10.1371/journal.pgen.1008222

Ystrom, E., Reichborn-Kjennerud, T., Neale, M. C., & Kendler, K. S. (2014). Genetic and environmental risk factors for illicit substance use and use disorders: Joint analysis of self and co-twin ratings. *Behavior Genetics*, *44*(1), 1-13. https://doi.org/10.1007/s10519-013-9626-6

Yu, D., Sul, J. H., Tsetsos, F., Nawaz, M. S., Huang, A. Y., Zelaya, I., Illmann, C., Osiecki, L., Darrow, S. M., Hirschtritt, M. E., Greenberg, E., Muller-Vahl, K. R., Stuhrmann, M., Dion, Y., Rouleau, G. A., Aschauer, H., Stamenkovic, M., Schlögelhofer, M., Sandor, P., … Worbe, Y. (2019). Interrogating the genetic determinants of Tourette's syndrome and other tiC disorders through genome-wide association studies. *American Journal of Psychiatry*, *176*(3), 217-227. https://doi.org/10.1176/appi.ajp.2018.18070857

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

Yuan, S., Yao, H., & Larsson, S. C. (2020). Associations of cigarette smoking with psychiatric disorders: evidence from a two-sample Mendelian randomization study. *Scientific Reports*, *10*(1), 1-9. https://doi.org/10.1038/s41598-020-70458-4

Zeng, L. N., Zong, Q. Q., Zhang, L., Feng, Y., Ng, C. H., Ungvari, G. S., Chen, L. G., & Xiang, Y. T. (2020). Worldwide prevalence of smoking cessation in schizophrenia patients: A meta-analysis of comparative and observational studies. *Asian Journal of Psychiatry*, *54*(May), 102190. https://doi.org/10.1016/j.ajp.2020.102190

Zhang, B., & Horvath, S. (2005). A General Framework for Weighted Gene Co-Expression Network Analysis. *Statistical Applications in Genetics and Molecular Biology*, *4*(1), 17. https://doi.org/10.2202/1544-6115.1128

Zhang, M., Tang, J., Li, W., Xue, K., Wang, Z., Chen, Y., Xu, Q., Zhu, D., Cai, M., Ma, J., Yao, J., Zhang, Y., Wang, H., Liu, F., & Guo, L. (2024). Schizophrenia mediating the effect of smoking phenotypes on antisocial behavior: A Mendelian randomization analysis. *CNS Neuroscience and Therapeutics*, *30*(3), 1-9. https://doi.org/10.1111/cns.14430

Zhou, H., Kember, R. L., Deak, J. D., Xu, H., Toikumo, S., Yuan, K., Lind, P. A., Farajzadeh, L., Wang, L., Hatoum, A. S., Johnson, J., Lee, H., Mallard, T. T., Xu, J., Johnston, K. J. A., Johnson, E. C., Nielsen, T. T., Galimberti, M., Dao, C., … Gelernter, J. (2023). Multi-ancestry study of the genetics of problematic alcohol use in over 1 million individuals. *Nature Medicine*, *29*(12), 3184-3192. https://doi.org/10.1038/s41591-023-02653-5

Zhou, H., Sealock, J. M., Sanchez-Roige, S., Clarke, T., Levey, D. F., Cheng, Z., Li, B., Polimanti, R., Kember, R. L., Smith, R. V., Thygesen, J. H., Morgan, M. Y., Atkinson, S. R., Thursz, M. R., Nyegaard, M., Mattheisen, M., Børglum, A. D., Johnson, E. C., Justice, A. C., … Gelernter, J. (2020). Genome-wide meta-analysis of problematic alcohol use in 435,563 individuals yields insights into biology and relationships with other traits. *Nature Neuroscience*, *23*(7), 809-818. https://doi.org/10.1038/s41593-020-0643-5

Zhu, Z., Zheng, Z., Zhang, F., Wu, Y., Trzaskowski, M., Maier, R., Robinson, M. R., McGrath, J. J., Visscher, P. M., Wray, N. R., & Yang, J. (2018). Causal associations between risk factors and common diseases inferred from GWAS summary data. *Nature Communications*, *9*(1), 224. https://doi.org/10.1038/s41467-017-02317-2

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

ANEXOS

ANEXO I: INFORMES FAVORABLES DOS COMITÉS DE ÉTICA

USC
UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

XUNTA DE GALICIA
CONSELLERÍA DE SANIDADE
Secretaría Xeral

Comité Ético de Investigación Clínica de Galicia
Edificio Administrativo de San Lázaro
15781 SANTIAGO DE COMPOSTELA
Tlf: 881 546425 Fax: 881 541804
ceic@sergas.es

DICTAMEN DEL COMITÉ ÉTICO DE INVESTIGACIÓN CLÍNICA DE GALICIA

Dña. Paula M. López Vázquez, Secretaria del Comité Ético de Investigación Clínica de Galicia

CERTIFICA:

Que este Comité evaluó en su reunión del día 26/07/2011 el estudio:

Título: Secuenciación exómica de pacientes esquizofrénicos con microdeleciones y microduplicaciones que confieren riesgo elevado a paceder esquizofrenia y otros trastornos del neurodesarrollo
Promotor: Javier Costas Costas
Version:
Código do Promotor:
Código de Registro CEIC de Galicia: 2011/276

Y que este Comité de conformidad con sus Procedimientos Normalizados de Trabajo y tomando en cuenta los requisitos éticos, metodológicos y legales exigibles a los estudios de investigación con seres humanos, sus muestras o registros, emite un **DICTAMEN FAVORABLE** al estudio propuesto y que se llevará a cabo en:

Centros	Investigadores principales
Fundación Pública Galega de Medicina Xenómica	Javier Costas Costas

En Santiago de Compostela a 29 de julio de 2011
La Secretaria

XUNTA DE GALICIA · SECRETARÍA XERAL · COMITÉ ÉTICO DE INVESTIGACIÓN CLÍNICA DE GALICIA · CONSELLERÍA DE SANIDADE

Paula M. López Vázquez

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

XUNTA DE GALICIA
CONSELLERÍA DE SANIDADE
Secretaría Xeral

Comité Ético de Investigación Clínica de Galicia
Edificio Administrativo de San Lázaro
15781 SANTIAGO DE COMPOSTELA
Tlf: 881 546425 Fax: 881 541804
ceic@sergas.es

galicia

DICTAMEN DEL COMITÉ ÉTICO DE INVESTIGACIÓN CLÍNICA DE GALICIA

Juan M. Vázquez Lago, Secretario Suplente del Comité Ético de Investigación Clínica de Galicia

CERTIFICA:

Que este Comité evaluó en su reunión del día 27/11/2012 el estudio:

Título: Caracterización del papel de microdeleciones/microduplicaciones del genoma como determinantes biológicos del policonsumo de drogas
Promotor: Javier Costas Costas
Tipo de estúdio: No-EPA
Código de Registro CEIC de Galicia: 2012/368

Y que este Comité de conformidad con sus Procedimientos Normalizados de Trabajo y tomando en cuenta los requisitos éticos, metodológicos y legales exigibles a los estudios de investigación con seres humanos, sus muestras o registros, emite un **DICTAMEN FAVORABLE** al estudio propuesto y que se llevará a cabo en:

Centros	Investigadores principales
Fundación Pública Galega de Medicina Xenómica	Javier Costas Costas
C.H. Universitario de Santiago	María Páramo
ACLAD A Coruña	César Pereiro
UMAD Santiago	Emilio Fariñas
UAD Noia	Joaquín Palomares
UCA Ourense	Gerardo Flórez
UAD Pontevedra	Carlos Pino
C.H. Universitario de A Coruña	Manuel Serrano

En Santiago de Compostela a 03 de diciembre de 2012
El Secretario Suplente

Juan M. Vázquez Lago

USC
UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

XUNTA DE GALICIA
CONSELLERÍA DE SANIDADE
Secretaría Xeral Técnica

Secretaria Técnica
Comité Autonómico de Ética da Investigación de Galicia
Secretaria Xeral. Consellería de Sanidade
Edificio Administrativo San Lázaro
15703 SANTIAGO DE COMPOSTELA
Tel: 881 546425; ceic@sergas.es

DICTAME DO COMITÉ DE ÉTICA DA INVESTIGACIÓN DE SANTIAGO-LUGO

Juan Manuel Vázquez Lago, Secretario do Comité de Ética da Investigación de Santiago-Lugo

CERTIFICA:

Que este Comité avaliou na súa reunión do día 21/07/2014 o estudo:

Título: Riesgo poligénico a desarrollar esquizofrenia a partir de módulos de coexpresión espaciotemporal en cerebro: implicaciones etiopatogénicas y clínicas
Promotor: Javier Costas Costas
Tipo de estudo: Outros
Versión:
Código do Promotor:
Código de Rexistro: 2014/369

E, tomando en consideración as seguintes cuestións:

- A pertinencia do estudo, tendo en conta o coñecemento dispoñible, así coma os requisitos legais aplicables, e en particular a Lei 14/2007, de investigación biomédica, o Real Decreto 1716/2011, de 18 de novembro, polo que se establecen os requisitos básicos de autorización e funcionamento dos biobancos con fins de investigación biomédica e do tratamento das mostras biolóxicas de orixe humana, e se regula o funcionamento e organización do Rexistro Nacional de Biobancos para investigación biomédica, a ORDE SAS/3470/2009, de 16 de decembro, pola que se publican as Directrices sobre estudos Posautorización de Tipo Observacional para medicamentos de uso humano, e a Circular nº 07/2004, investigacións clínicas con produtos sanitarios.
- A idoneidade do protocolo en relación cos obxectivos do estudo, xustificación dos riscos e molestias previsibles para o suxeito, así coma os beneficios esperados.
- Os principios éticos da Declaración de Helsinki vixente.
- Os Procedementos Normalizados de Traballo do Comité.

Emite un **INFORME FAVORABLE** para a realización do estudo **polo/a investigador/a do centro**:

Centros	Investigadores Principais
C.H. Universitario de Santiago; IDIS	Javier Costas Costas

En Santiago de Compostela, a 23 de xullo de 2014
O secretario

juan.manuel.vazquez.lago@sergas.es

Firmado digitalmente por juan.manuel.vazquez.lago@sergas.es
Nombre de reconocimiento (DN): cn=juan.manuel.vazquez.lago@sergas.es
Fecha: 2014.07.24 12:22:18 +02'00'

Juan M. Vázquez Lago

XUNTA DE GALICIA
CONSELLERÍA DE SANIDADE
Secretaría Xeral Técnica

Secretaria Técnica
Comité Autonómico de Ética da Investigación de Galicia
Secretaria Xeral. Consellería de Sanidade
Edificio Administrativo San Lázaro
15703 SANTIAGO DE COMPOSTELA
Tel: 881 546425; ceic@sergas.es

DICTAME DO COMITÉ DE ÉTICA DA INVESTIGACIÓN DE SANTIAGO-LUGO

Juan Manuel Vázquez Lago, Secretario do Comité de Ética da Investigación de Santiago-Lugo

CERTIFICA:

Que este Comité avaliou na súa reunión do día 20/10/2014 o estudo:

Título: Contribución de la variación poligénica común a la predicción del riesgo a dependencia de alcohol
Promotor: Javier Costas Costas
Tipo de estudo: Outros
Versión:
Código do Promotor:
Código de Rexistro: 2014/428

E, tomando en consideración as seguintes cuestións:

- A pertinencia do estudo, tendo en conta o coñecemento dispoñible, así coma os requisitos legais aplicables, e en particular a Lei 14/2007, de investigación biomédica, o Real Decreto 1716/2011, de 18 de novembro, polo que se establecen os requisitos básicos de autorización e funcionamento dos biobancos con fins de investigación biomédica e do tratamento das mostras biolóxicas de orixe humana, e se regula o funcionamento e organización do Rexistro Nacional de Biobancos para investigación biomédica, a ORDE SAS/3470/2009, de 16 de decembro, pola que se publican as Directrices sobre estudos Posautorización de Tipo Observacional para medicamentos de uso humano, e a Circular nº 07/2004, investigacións clínicas con produtos sanitarios.
- A idoneidade do protocolo en relación cos obxectivos do estudo, xustificación dos riscos e molestias previsibles para o suxeito, así coma os beneficios esperados.
- Os principios éticos da Declaración de Helsinki vixente.
- Os Procedementos Normalizados de Traballo do Comité.

Emite un **INFORME FAVORABLE** para a realización do estudo **polo/a investigador/a do centro**:

Centros	Investigadores Principais
C.H. Universitario de Santiago	Javier Costas Costas

En Santiago de Compostela, a 27 de outubro de 2014

O secretario

juan.manuel.vazquez.lago@sergas.es

Firmado digitalmente por juan.manuel.vazquez.lago@sergas.es
Nombre de reconocimiento (DN): cn=juan.manuel.vazquez.lago@sergas.es
Fecha: 2014.10.27 10:14:32 +01'00'

ANEXO II: ARTIGOS PUBLICADOS INCLUÍDOS NA TESE DE DOUTORAMENTO

A continuación achégase, para aqueles artigos incluídos nesta tese de doutoramento que están publicados, a versión orixinal publicada na revista.

USC
UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

Estudo 3: Al-Soufi, L. e 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.

USC
UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

Drug and Alcohol Dependence 220 (2021) 108517

Contents lists available at ScienceDirect

Drug and Alcohol Dependence

journal homepage: www.elsevier.com/locate/drugalcdep

Short communication

Colocalization of association signals at nicotinic acetylcholine receptor genes between schizophrenia and smoking traits

Laila Al-Soufi [a,b], Javier Costas [a,c,*]

[a] Psychiatric Genetics Group, Instituto de Investigación Sanitaria de Santiago de Compostela (IDIS), Galicia, Spain
[b] Universidade de Santiago de Compostela (USC), Galicia, Spain
[c] Complexo Hospitalario Universitario de Santiago de Compostela (CHUS), Servizo Galego de Saúde (SERGAS), Santiago de Compostela, Galicia, Spain

ARTICLE INFO

Keywords:
Genetic epidemiology
Nicotinic acetylcholine receptors
Smoking risk
Psychosis
Mental health
Tobacco

ABSTRACT

Background: Epidemiological data suggest that smoking may be a risk factor for schizophrenia (SCZ), but more evidence is needed. Two regions coding nicotinic acetylcholine receptor (nAchR) subunits, at*CHRNA2* and the *CHRNA5/A3/B4* cluster, were associated with SCZ in genome-wide association studies (GWAS). Additionally, a signal at *CHRNA4* is near significance. *CHRNA2* was also associated with cannabis use disorder (CUD). These regions were also associated with smoking behaviors. If tobacco is a risk factor, the GWAS signals at smoking behaviors and SCZ have to be due to the same causal variants, i.e. they have to colocalize, although colocalization does not necessarily imply causality. Here, we present colocalization analysis at these loci between SCZ and smoking behaviors.

Methods: The Bayesian approach implemented in coloc was used to check for posterior probability of colocalization versus independent signals at the three loci with some evidence of association with SCZ and smoking behaviors, using GWAS summary statistics. Colocalization was also assessed for positive control traits and CUD. Three different sensibility analyses were used to confirm the results. A visualization tool, LocusCompare, was used to facilitate interpretation of the coloc results.

Results: Evidence for colocalization of GWAS signals between SCZ and smoking behaviors was found for *CHRNA2*. Evidence for independent causal variants was found for the other two loci. CUD GWAS signal at *CHRNA2* colocalizes with SCZ and smoking behaviors.

Conclusions: Overall, the results indicate that the association between some nAchR subunit genes and SCZ cannot be solely explained by their effect on smoking behaviors.

1. Introduction

Epidemiological studies have found that smoking increases the posterior risk of developing schizophrenia (SCZ) and is associated with an earlier onset of psychosis in a dose-dependent manner, suggesting a role of smoking on risk for psychotic disorders (Gurillo et al., 2015; Kendler et al., 2015; Koyanagi et al., 2016). Genome-wide association studies (GWAS) have also contributed to this topic. GWAS of smoking behaviors identified genome-wide significant associations at almost all genes coding for neuronal nicotinic acetylcholine receptor (nAchR) subunits (Liu et al., 2019b). A SCZ GWAS also identified a genome-wide significant association at the cluster of nAchR genes at chromosome 15, herein referred to as *CHRNA5/A3/B4* (Schizophrenia Working Group of the Psychiatric Genomics Consortium et al., 2014). This suggests that the association of *CHRNA5/A3/B4* with SCZ may be mediated by smoking, acting as a risk factor (Alderson and Lawrie, 2015; Gage and Munafò, 2015; Laursen and McGrath, 2016; McGrath, 2015). Remarkably, taking into account that cannabis acts as a risk factor for SCZ (Belbasis et al., 2018; Di Forti et al., 2019), a GWAS on cannabis use disorder (CUD) identified just one GWAS significant hit, at *CHRNA2* (Demontis et al., 2019), a locus also associated with SCZ (Schizophrenia Working Group of the Psychiatric Genomics Consortium, 2014). We must bear in mind that while GWAS identify associated regions, the existence of high linkage disequilibrium (LD) precludes the identification of the specific causal SNP (or SNPs) responsible for the association. However, one SNP associated with several traits, rs16969968, deserves attention. It is a

* Corresponding author at: Instituto de Investigación Sanitaria (IDIS) de Santiago de Compostela, Hospital Clínico Universitario, edificio Consultas, andar -2, Grupo de xenética psiquiátrica, despacho 15, E-15706, Santiago de Compostela, Galicia, Spain.
E-mail addresses: laila.alsoufi.novo@usc.es (L. Al-Soufi), javier.costas.costas@sergas.es (J. Costas).

https://doi.org/10.1016/j.drugalcdep.2021.108517
Received 10 September 2020; Received in revised form 16 December 2020; Accepted 17 December 2020
Available online 10 January 2021
0376-8716/© 2021 Elsevier B.V. All rights reserved.

Table 1
Datasets used in the present study.

Trait name[a]	Description	Source[b]	Sample size[c]	p *CHRNA5/A3/B4*[d]	p *CHRNA2*[d]	p *CHRNA4*[d]
CigDay	Average number of cigarettes smoked per day	GSCAN	337 334	4.8×10^{-211}	2.44×10^{-10}	2.47×10^{-18}
SmkCes	Smoking cessation, current vs former smokers	GSCAN	139 453 / 407 766	7.57×10^{-11}	$p > 1 \times 10^{-5}$	2.04×10^{-16}
AgeSmk	Age started smoking cigarettes regularly	GSCAN	341 427	$p > 1 \times 10^{-5}$	7.00×10^{-11}	$p > 1 \times 10^{-5}$
SmkInit	Ever vs never regular smokers	GSCAN	557 337 / 674 754	$p > 1 \times 10^{-5}$	3.42×10^{-17}	$p > 1 \times 10^{-5}$
FTND	Fagerström Test for Nicotine Dependence	Hancock et al. (2015)	17 074	3.50×10^{-17}	$p > 1 \times 10^{-5}$	3.80×10^{-8}
Frequency	Current smoking frequency	GeneAtlas	501 649	8.35×10^{-7}	5.82×10^{-13}	1.52×10^{-13}
Attempts	Number of unsuccessful stop smoking attempts	Neale's lab[e]	81 026	6.45×10^{-10}	$p > 1 \times 10^{-5}$	$p > 1 \times 10^{-5}$
SCZ	Schizophrenia	PGC	33 640 / 43 456	2.26×10^{-13}	8.76×10^{-9}	9.45×10^{-8}
CUD	Cannabis use disorder	iPSYCH	2387 / 48 985	$p > 1 \times 10^{-5}$	9.31×10^{-12}	$p > 1 \times 10^{-5}$
LungCancer	Overall lung cancer	TRICL-ILCCO and LC3	44 069 / 68 712	3.08×10^{-103}	1.69×10^{-8}	$p > 1 \times 10^{-5}$
FVC	Lung function measured as the forced vital capacity	UKBB	321 047	6.95×10^{-10}	$p > 1 \times 10^{-5}$	3.04×10^{-6}

[a] Name used along the text.
[b] GSCAN, GWAS and Sequencing Consortium of Alcohol and Nicotine use (Liu et al., 2019b); GeneAtlas (Canela-Xandri et al., 2018); PGC, Psychiatric Genomics Consortium (Schizophrenia Working Group of the Psychiatric Genomics Consortium, 2014); iPSYCH, Lundbeck Foundation Initiative for Integrative Psychiatric Research (Demontis et al., 2019); TRICL-ILCCO and LC3, Transdisciplinary Research of Cancer in Lung of the International Lung Cancer Consortium and Lung Cancer Cohort Consortium (McKay et al., 2017); UKBB, UK BioBank (Shrine et al., 2019).
[c] For dichotomous traits, case / control sample sizes are shown.
[d] p value of the most significant SNP at each locus.
[e] http://www.nealelab.is/uk-biobank/; GWAS round 2 results, released 1st August 2018.

missense SNP at *CHRNA5*, changing the evolutionary conserved amino acid aspartate at position 398 of the encoded protein by asparagine. This change has functional consequences, both in vitro and in vivo (Maskos, 2020).

However, identification of an association did not prove causality, due to problems such as confounding or reverse causality. Further evidence is provided by the Mendelian randomization (MR) approach. Recently, a large MR study found a causal effect of smoking on SCZ (Wootton et al., 2019). However, although MR is a step forward towards causality, it presents several limits that preclude definitive conclusions (Davies et al., 2018; Hartwig et al., 2018; Hemani et al., 2018; Katikireddi et al., 2018). Different approaches are needed to triangulate the evidence pointing to a causal role of smoking on SCZ (Lawlor et al., 2016; Ohlsson and Kendler, 2020). To this aim, we applied for the first time colocalization analysis to this problem. Colocalization analysis is generally used to assess colocalization of GWAS signals between diseases (or related traits) and expression quantitative trait loci (eQTLs) (Giambartolomei et al., 2014; He et al., 2013; Hormozdiari et al., 2016), but may be used to compare the association signals at any two traits (Giambartolomei et al., 2014). If smoking is a risk factor for SCZ, i.e., mediated (or vertical) pleiotropy (Hemani et al., 2018), variants at genes associated with both smoking and SCZ have to colocalize, although colocalization does not necessarily imply mediated pleiotropy.

2. Methods

2.1. Datasets

Several publicly available GWAS summary statistics of different traits were analyzed (Table 1 and Supplementary Material). Colocalization between tobacco use and SCZ was assessed using GWAS data for seven smoking traits. GWAS on lung cancer and lung function (forced vital capacity) were used as positive controls, and a GWAS of CUD was used to assess a possible confounding effect.

2.2. Colocalization analysis

Coloc is a Bayesian statistical method to test for colocalization that works with the summary statistics of two association studies (Giambartolomei et al., 2014). This method enables the differentiation between colocalization, i.e., one common association signal for two traits, and a situation where independent signals are associated with each trait. It computes the posterior probability for five alternative hypotheses: H_0, no association with either trait; H_1, association with trait 1 but no with trait 2; H_2, association with trait 2 but no with trait 1; H_3, two independent association signals, one for each trait; or H_4, a common association signal for both traits. Assuming that the analyzed loci are associated with both tested traits, the power of the colocalization analysis to detect an association with both traits is indicated by the sum of the posterior probability for H_3 (PP_3) and the posterior probability for H_4 (PP_4). A threshold of $PP_3 + PP_4 > 0.8$ was set to filter comparisons with a high support for association with both traits. Among these comparisons, support for hypothesis H_3 and H_4 was assessed as the ratio of posterior probabilities, considering evidence for support if the posterior probability of one hypothesis was more than three times the posterior probability for the other hypothesis, as previously done (Bhalala et al., 2018). The prior probabilities for SNP association with either of the two traits or with both traits were set to 1×10^{-4} and 1×10^{-5}, the default values, respectively.

Comparisons were restricted to traits from the same ethnic background without overt overlapping samples, owing to the assumptions of the method (Giambartolomei et al., 2014), and to loci with any SNP in the ±200 kb region around the transcribed region of nAchR subunits genes with $p < 1 \times 10^{-5}$ in their respective GWAS (Table 1).

Colocalization events were visualized with LocusCompare which generates a useful plot to differentiate between two distinct signals and a common signal for both traits (Liu et al., 2019a). LD score regression (Bulik-Sullivan et al., 2015) was used to check the existence of cryptic samples overlap (Supplementary Material).

2.3. Sensitivity analyses

Three different sensitivity analyses were performed considering different values for prior probabilities, correcting for putative samples overlap, and considering different causal variants (Supplementary Methods).

3. Results

3.1. Association of genes coding nAchR subunits with traits under study

Searching for significant SNPs at ±200 kb of the transcribed region of nAchR subunits genes in the GWAS summary statistics from the Psychiatric Genomics Consortium identified three loci associated with SCZ at $p < 1 \times 10^{-5}$, involving *CHRNA2*, *CHRNA4*, and *CHRNA5/A3/B4* (Table 1). These three GWAS association signals were the focus of the colocalization analysis. Table 1 also shows the GWAS results at these loci for smoking behaviors, control traits and CUD.

Table 2
Results of colocalization analysis for comparisons showing evidence for colocalization (H_4) or two independent signals (H_3) at the main analysis.

Locus	Trait 1	Trait 2	Coloc main test[a] H (PP_3+PP_4, PP_4/PP_3)	Sensitivity I[b] H (PP_3+PP_4, PP_4/PP_3)	Sensitivity II[c] H (PP_3+PP_4, PP_4/PP_3)	Sensitivity III[d] H (PP_3+PP_4, PP_4/PP_3)
CHRNA5/A3/B4	LungCancer	SmkCes	H_4 (>0.99, 32.50)	H_4 (>0.99, 3.25)	–	H_3 (0.88, 3.3e-5)
CHRNA5/A3/B4	LungCancer	Attempts	H_4 (>0.99, 25.34)	–	–	H_3 (>0.99, 1.0e-5)
CHRNA5/A3/B4	LungCancer	FTND	H_4 (>0.99, 7.32)	–	–	–
CHRNA5/A3/B4	LungCancer	Frequency	H_3 (0.89, 2.6e-3)	H_3 (0.89, 2.6e-4)	–	H_3 (0.89, 3.5e-3)
CHRNA5/A3/B4	LungCancer	CigDay	H_3 (>0.99, 5.2e-3)	H_3 (>0.99, 5.2e-4)	–	H_3 (0.88, 1.66e-202)
CHRNA2	LungCancer	AgeSmk	H_4 (>0.99, 249.12)	H_4 (>0.99, 24.91)	–	H_3 (0.93, 3.5e-7)
CHRNA2	LungCancer	SmkInit	H_4 (>0.99, 10.82)	–	–	H_3 (0.94, 1.9e-9)
CHRNA2	LungCancer	Frequency	H_4 (>0.99, 5.56)	–	–	H_3 (0.94, 1.7e-8)
CHRNA2	LungCancer	CigDay	H_4 (>0.99, 4.02)	–	–	H_3 (0.94, 3.5e-4)
CHRNA5/A3/B4	FVC	FTND	H_4 (>0.99, 8.76)	–	–	–
CHRNA5/A3/B4	SCZ	Frequency	*H_4 (0.99, 7.84)*	–	H_4 (0.94, 141.90)	none
CHRNA5/A3/B4	SCZ	CigDay	H_3 (>0.99, 3.4e-5)	H_3 (>0.99, 3.4e-6)	H_3 (>0.99, 2.0e-3)	H_3 (>0.99, 1.2e-8)
CHRNA5/A3/B4	SCZ	FTND	H_3 (>0.99, 1.2e-4)	H_3 (>0.99, 1.1e-5)	–	–
CHRNA5/A3/B4	SCZ	SmkCes	*H_3 (>0.99, 2.8e-1)*	H_3 (>0.99, 2.8e-2)	–	none
CHRNA2	SCZ	CigDay	H_4 (>0.99, 376.05)	H_4 (>0.99, 37.61)	H_4 (>0.99, 1596.63)	none
CHRNA2	SCZ	SmkInit	H_4 (>0.99, 33.42)	H_4 (>0.99, 3.34)	H_4 (>0.99, 246.22)	none
CHRNA2	SCZ	Frequency	*H_4 (>0.99, 15.24)*	–	H_4 (>0.99, 272.74)	none
CHRNA2	SCZ	AgeSmk	*H_3 (>0.99, 2.1e-1)*	H_3 (>0.99, 2.1e-2)	–	none
CHRNA4	SCZ	SmkCes	*H_3 (0.95, 4.5e-4)*	H_3 (0.95, 4.5e-4)	H_3 (>0.99, 1.2e-3)	H_3 (0.93, 9.4e-4)
CHRNA4	SCZ	CigDay	H_3 (0.95, 1.2e-3)	H_3 (0.95, 1.2e-4)	H_3 (>0.99, 5.7e-3)	none
CHRNA4	SCZ	Frequency	*H_3 (0.95, 6.1e-3)*	H_3 (0.95, 6.1e-4)	H_3 (>0.99, 4.6e-2)	none
CHRNA2	CUD	SCZ	H_4 (>0.99, 81.44)	H_4 (>0.99, 8.14)	H_4 (>0.99, 2044.49)	none
CHRNA2	CUD	SmkInit	H_4 (>0.99, 55.29)	H_4 (>0.99, 5.53)	H_4 (>0.99, 9047.79)	none
CHRNA2	CUD	Frequency	H_4 (>0.99, 11.88)	–	H_4 (>0.99, 884.99)	none
CHRNA2	CUD	CigDay	H_4 (>0.99, 5.47)	–	H_4 (>0.99, 24.86)	none
CHRNA2	CUD	AgeSmk	H_3 (>0.99, 2.4e-1)	H_3 (>0.99, 2.4e-2)	–	none

[a] Those comparisons with LD score regression intercepts higher than one standard error away from zero, suggesting cryptic samples overlap, are in italics.
[b] Results for sensitivity analysis using a different prior probability for colocalization.
[c] Results for sensitivity analysis using gwas-pw to correct for samples overlap. This analysis was not applied to comparisons involving LungCancer or FTND because it requires full GWAS summary statistics. The different hypotheses were named following coloc nomenclature.
[d] Results for sensitivity analysis using masking to consider the existence of more than one causal variant. Results are shown for additional pair of variants supporting any hypothesis, after masking of signals of the main analysis. "none" means that there were no additional signals at these comparisons. In case of more than one additional pair of signals, the one with the strongest support is shown. Full results are shown in Supplementary Table 1. The analysis was not applied to comparisons involving FTND because the required beta values and standard errors were not supplied in the FTND GWAS summary statistics.

3.2. Colocalization between control traits and smoking traits

Coloc identified colocalization between LungCancer and three smoking traits at *CHRNA5/A3/B4* (Table 2, Fig. S1). For one additional trait, CigDay, while coloc supported independent signals, visual inspection of the comparison of these two GWAS results using LocusCompare suggested that there was colocalization, but the strong LD among many SNPs led coloc to discard colocalization (Table 2, Fig. S1). The shared colocalization signal was related to rs16969968. There was evidence of two independent signals in the comparison with Frequency, a trait that did not show association with rs16969968 (Table 2, Fig. S1). Regarding FVC, comparison with the only smoking trait without overlapping samples, FTND, revealed evidence of colocalization (Table 2, Fig. S2).

For *CHRNA2*, there was evidence of colocalization between LungCancer and all the smoking behaviors tested (Table 2, Fig. S1).

Regarding *CHRNA4*, the only comparison with non-overlapped samples was FVC versus FTND. Analysis of this comparison suggested colocalization ($PP_4/PP_3 = 35.9$), but with limited power ($PP_4 + PP_3 = 0.73$) (Fig. S2).

3.3. Colocalization between SCZ and smoking traits

In the case of *CHRNA5/A3/B4*, there were four comparisons between SCZ and smoking traits with enough power. Three of them supported independent signals, while that involving Frequency supported colocalization (Table 2). LocusCompare plots revealed that independent signals were due to the strong association of smoking traits with rs16969968, lacked in the case of SCZ (Fig. S3).

For *CHRNA2*, three of the four tested comparisons supported colocalization (Table 2, Fig. S3). The remaining analysis, involving AgeSmk, supported the existence of two different signals (Table 2, Fig. S3). Finally, there was evidence for independent signals at *CHRNA4*, supported by three different smoking traits (Table 2, Fig. S3).

3.4. Colocalization analysis with CUD

Taking into account that tobacco and cannabis use are correlated, we performed colocalization analysis between CUD, SCZ and smoking traits at *CHRNA2*. The signal at this locus colocalized with SCZ (Table 2, Fig. S4). In addition, there was evidence for colocalization between CUD and the same smoking traits that colocalized with SCZ (Table 2, Fig. S4). The comparison between CUD and AgeSmk revealed independent signals, as was the case for SCZ (Table 2, Fig. S4).

3.5. Sensitivity analyses

Sensitivity analyses confirmed the main results for comparisons involving SCZ, although additional causal variants with different evidence of colocalization appeared for LungCancer (Tables 2 and S1).

4. Discussion

In this work, we analyzed colocalization at nAchR subunit genes to provide new data on the debate about the relationship between smoking and SCZ based on a different approach. Our results at *CHRNA5/A3/B4* are against the causal relationship, as we found strong evidence supporting the existence of independent association signals at SCZ and smoking traits. In the case of several conditional independent associations at one locus, as in *CHRNA5/A3/B4* for CigDay (Liu et al., 2019b), coloc considers only the strongest signal (Giambartolomei et al., 2014). Therefore, we can conclude that the strongest signal at this locus

responsible for association with smoking traits such as CigDay, SmkCes, or FTND, is not the strongest signal associated with SCZ. Moreover, additional signals at this locus in smoking traits also did not colocalize with SCZ. By contrast, we have found strong evidence of colocalization between the same smoking traits and LungCancer, used as positive control. As clearly revealed by LocusCompare, the signal shared between smoking traits and LungCancer is related to the missense SNP rs16969968 at *CHRNA5*. Therefore, contrary to smoking traits, the strongest association signal for SCZ at *CHRNA5/A3/B4* is not lead by rs16969968.

Regarding *CHRNA4*, coloc supports the existence of two independent signals, one associated with SCZ and another with smoking traits. In fact, this locus was not associated at genome-wide significance with LungCancer, suggesting that the effect of this locus mediating the role of smoking on other traits may be very low.

Finally, colocalization of GWAS signals between SCZ, LungCancer, CUD and smoking traits is the most probable hypothesis at *CHRNA2*. Therefore, we cannot discard a role of cannabis use as a confounding factor on the association between SCZ and smoking (Gage et al., 2014; Jones et al., 2018).

Our approach presents several strengths and limitations. Among the strengths, the approach is orthogonal to other approaches, providing independent results to address causality. Additionally, we have used several datasets based on meta-analysis of many different cohorts with different ascertainment schemes, bias and demographic characteristics, and covering a large range of smoking behaviors to add confidence to our findings. Finally, the use of coloc has the advantage of easy interpretation of results, comparing evidence among different hypotheses, while LocusCompare helps to visually inspect congruence of coloc results.

Regarding limitations, coloc assumes that there is just one causal variant per locus. However, coloc is robust to the existence of several causal variants if they did not explain similar proportion of the variance of the trait (Giambartolomei et al., 2014). Moreover, coloc only compares GWAS from independent samples, precluding the use of additional positive controls. There is also a problem in calculation of posterior probabilities if there are many SNPs at strong LD with the causal variants, as exemplified in the case of LungCancer and CigDay at *CHRNA5/A3/B4*. Other limitation is that, as a Bayesian approach, prior probabilities have to be defined. We have used positive control traits and several sensitivity analyses to deal with these limitations, suggesting that the results are robust. Finally, colocalization does not imply mediated pleiotropy. Therefore, horizontal pleiotropy and reverse causality cannot be discarded.

In summary, the shared association signal at *CHRNA2* is compatible with putative mediated pleiotropy, i.e., its effect on SCZ risk may be due to tobacco and/or cannabis use. The lack of colocalization at the *CHRNA5/A3/B4* cluster and *CHRNA4* suggests additional involvement of nAchRs in SCZ risk, irrespective of smoking. These findings may have implications on the potential of nAchRs as novel drug targets for treatment of SCZ (Terry et al., 2015; Wilkerson et al., 2020).

Contributors

JC conceived and designed the study. LA-S performed the analyses under the supervision of JC. Both authors have written the manuscript. All authors have read and approved the final manuscript.

Role of the funding source

Nothing declared.

Declaration of Competing Interest

No conflict declared.

Acknowledgements

This work was supported by the Instituto de Salud Carlos III [grant number ISCIII/PI17/01424] to JC, cofounded by FEDER. LA-S was the recipient of a research fellowship by the Ministerio de Ciencia, Innovación y Universidades [grant number FPU18/03243]. The authors wish to thank all the Consortia and research groups providing the summary statistics datasets used in this work.

Appendix A. Supplementary data

Supplementary material related to this article can be found, in the online version, at doi:https://doi.org/10.1016/j.drugalcdep.2021.108517.

References

Alderson, H.L., Lawrie, S.M., 2015. Does cigarette smoking cause psychosis? Lancet Psychiatry 2, 672–673. https://doi.org/10.1016/S2215-0366(15)00239-4.

Belbasis, L., Köhler, C.A., Stefanis, N., Stubbs, B., van Os, J., Vieta, E., Seeman, M.V., Arango, C., Carvalho, A.F., Evangelou, E., 2018. Risk factors and peripheral biomarkers for schizophrenia spectrum disorders: an umbrella review of meta-analyses. Acta Psychiatr. Scand. 137, 88–97. https://doi.org/10.1111/acps.12847.

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 Genet. 14, e1007607 https://doi.org/10.1371/journal.pgen.1007607.

Bulik-Sullivan, B., Finucane, H.K., Anttila, V., Gusev, A., Day, F.R., Loh, P.R., Duncan, L., Perry, J.R.B., Patterson, N., Robinson, E.B., Daly, M.J., Price, A.L., Neale, B.M., 2015. An atlas of genetic correlations across human diseases and traits. Nat. Genet. 47, 1236–1241. https://doi.org/10.1038/ng.3406.

Canela-Xandri, O., Rawlik, K., Tenesa, A., 2018. An atlas of genetic associations in UK Biobank. Nat. Genet. 50, 1593–1599. https://doi.org/10.1038/s41588-018-0248-z.

Davies, N.M., Holmes, M.V., Davey Smith, G., 2018. Reading Mendelian randomisation studies: a guide, glossary, and checklist for clinicians. BMJ 362, k601. https://doi.org/10.1136/bmj.k601.

Demontis, D., Rajagopal, V.M., Thorgeirsson, T.E., Als, T.D., Grove, J., Leppälä, K., Gudbjartsson, D.F., Pallesen, J., Hjorthøj, C., Reginsson, G.W., Tyrfingsson, T., Runarsdottir, V., Qvist, P., Christensen, J.H., Bybjerg-Grauholm, J., Bækvad-Hansen, M., Huckins, L.M., Stahl, E.A., Timmermann, A., Agerbo, E., Hougaard, D. M., Werge, T., Mors, O., Mortensen, P.B., Nordentoft, M., Daly, M.J., Stefansson, H., Stefansson, K., Nyegaard, M., Børglum, A.D., 2019. Genome-wide association study implicates CHRNA2 in cannabis use disorder. Nat. Neurosci. 22, 1066–1074. https://doi.org/10.1038/s41593-019-0416-1.

Di Forti, M., Quattrone, D., Freeman, T.P., Tripoli, G., Gayer-Anderson, C., Quigley, H., Rodriguez, V., Jongsma, H.E., Ferraro, L., La Cascia, C., La Barbera, D., Tarricone, I., Berardi, D., Szöke, A., Arango, C., Tortelli, A., Velthorst, E., Bernardo, Miguel, Del-Ben, C.M., Menezes, P.R., Selten, J.P., Jones, P.B., Kirkbride, J.B., Rutten, B.P., de Haan, L., Sham, P.C., et al., 2019. The contribution of cannabis use to variation in the incidence of psychotic disorder across Europe (EU-GEI): a multicentre case-control study. Lancet Psychiatry 6, 427–436. https://doi.org/10.1016/S2215-0366(19)30048-3.

Gage, S.H., Munafò, M.R., 2015. Rethinking the association between smoking and schizophrenia. Lancet Psychiatry 2, 118–119. https://doi.org/10.1016/S2215-0366(14)00057-1.

Gage, S.H., Hickman, M., Heron, J., Munafò, M.R., Lewis, G., Macleod, J., Zammit, S., 2014. Associations of cannabis and cigarette use with psychotic experiences at age 18: findings from the Avon Longitudinal Study of Parents and Children. Psychol. Med. 44, 3435–3444. https://doi.org/10.1017/S0033291714000531.

Giambartolomei, C., Vukcevic, D., Schadt, E.E., Franke, L., Hingorani, A.D., Wallace, C., Plagnol, V., 2014. Bayesian test for colocalisation between pairs of genetic association studies using summary statistics. PLoS Genet. 10, e1004383 https://doi.org/10.1371/journal.pgen.1004383.

Gurillo, P., Jauhar, S., Murray, R.M., MacCabe, J.H., 2015. Does tobacco use cause psychosis? Systematic review and meta-analysis. Lancet Psychiatry 2, 718–725. https://doi.org/10.1016/S2215-0366(15)00152-2.

Hancock, D.B., Reginsson, G.W., Gaddis, N.C., Chen, X., Saccone, N.L., Lutz, S.M., Qaiser, B., Sherva, R., Steinberg, S., Zink, F., Stacey, S.N., Glasheen, C., Chen, J., Gu, F., Frederiksen, B.N., Loukola, A., Gudbjartsson, D.F., Brüske, I., Landi, M.T., Bickeböller, H., Madden, P., Farrer, L., Kaprio, J., Kranzler, H.R., Gelernter, J., Baker, T.B., Kraft, P., Amos, C.I., Caporaso, N.E., Hokanson, J.E., Bierut, L.J., Thorgeirsson, T.E., Johnson, E.O., Stefansson, K., 2015. Genome-wide meta-analysis reveals common splice site acceptor variant in CHRNA4 associated with nicotine dependence. Transl. Psychiatry 5, e651. https://doi.org/10.1038/tp.2015.149.

Hartwig, F.P., Davies, N.M., Davey Smith, G., 2018. Bias in Mendelian randomization due to assortative mating. Genet. Epidemiol. 42, 608–620. https://doi.org/10.1002/gepi.22138.

He, X., Fuller, C.K., Song, Y., Meng, Q., Zhang, B., Yang, X., Li, H., 2013. Sherlock: detecting gene-disease associations by matching patterns of expression QTL and GWAS. Am. J. Hum. Genet. 92, 667–680. https://doi.org/10.1016/j.ajhg.2013.03.022.

Hemani, G., Bowden, J., Davey Smith, G., 2018. Evaluating the potential role of pleiotropy in Mendelian randomization studies. Hum. Mol. Genet. 27, R195–R208. https://doi.org/10.1093/hmg/ddy163.

Hormozdiari, F., van de Bunt, M., Segrè, A.V., Li, X., Joo, J.W.J., Bilow, M., Sul, J.H., Sankararaman, S., Pasaniuc, B., Eskin, E., 2016. Colocalization of GWAS and eQTL signals detects target genes. Am. J. Hum. Genet. 99, 1245–1260. https://doi.org/10.1016/j.ajhg.2016.10.003.

Jones, H.J., Gage, S.H., Heron, J., Hickman, M., Lewis, G., Munafò, M.R., Zammit, S., 2018. Association of combined patterns of tobacco and cannabis use in adolescence with psychotic experiences. JAMA Psychiatry 75, 240–246. https://doi.org/10.1001/jamapsychiatry.2017.4271.

Katikireddi, S.V., Green, M.J., Taylor, A.E., Davey Smith, G., Munafò, M.R., 2018. Assessing causal relationships using genetic proxies for exposures: an introduction to Mendelian randomization. Addiction 113, 764–774. https://doi.org/10.1111/add.14038.

Kendler, K.S., Lönn, S.L., Sundquist, J., Sundquist, K., 2015. Smoking and schizophrenia in population cohorts of Swedish women and men: a prospective co-relative control study. Am. J. Psychiatry 172, 1092–1100. https://doi.org/10.1176/appi.ajp.2015.15010126.

Koyanagi, A., Stickley, A., Haro, J.M., 2016. Psychotic symptoms and smoking in 44 countries. Acta Psychiatr. Scand. 133, 497–505. https://doi.org/10.1111/acps.12566.

Laursen, T.M., McGrath, J.J., 2016. The strange case of smoking and schizophrenia - the epidemiology detectives are on the trail. Am. J. Psychiatry 173, 757–758. https://doi.org/10.1176/appi.ajp.2016.16050506.

Lawlor, D.A., Tilling, K., Smith, G.D., 2016. Triangulation in aetiological epidemiology. Int. J. Epidemiol. 45, 1866–1886. https://doi.org/10.1093/ije/dyw314.

Liu, B., Gloudemans, M.J., Rao, A.S., Ingelsson, E., Montgomery, S.B., 2019a. Abundant associations with gene expression complicate GWAS follow-up. Nat. Genet. 51, 768–769. https://doi.org/10.1038/s41588-019-0404-0.

Liu, M., Jiang, Y., Wedow, R., Li, Y., Brazel, D.M., Chen, F., Datta, G., Davila-Velderrain, J., McGuire, D., Tian, C., Zhan, X., Agee, M., Alipanahi, B., Auton, A., Bell, R.K., Bryc, K., Elson, S.L., Fontanillas, P., Furlotte, N.A., Hinds, D.A., Hromatka, B.S., Huber, K.E., Kleinman, A., Litterman, N.K., McIntyre, M.H., et al., 2019b. Association studies of up to 1.2 million individuals yield new insights into the genetic etiology of tobacco and alcohol use. Nat. Genet. 51, 237–244. https://doi.org/10.1038/s41588-018-0307-5.

Maskos, U., 2020. The nicotinic receptor alpha5 coding polymorphism rs16969968 as a major target in disease: functional dissection and remaining challenges. J. Neurochem. 154, 241–250. https://doi.org/10.1111/jnc.14989.

McGrath, J.J., 2015. A Rosetta stone for epidemiology: genomic risk profile scores contain clues related to modifiable risk factors. Epidemiol. Psychiatr. Sci. 24, 1–5. https://doi.org/10.1017/S2045796014000651.

McKay, J.D., Hung, R.J., Han, Y., Zong, X., Carreras-Torres, R., Christiani, D.C., Caporaso, N.E., Johansson, Mattias, Xiao, X., Li, Y., Byun, J., Dunning, A., Pooley, K. A., Qian, D.C., Ji, X., Liu, G., Timofeeva, M.N., Bojesen, S.E., Wu, X., Le Marchand, L., Albanes, D., Bickeböller, H., Aldrich, M.C., Bush, W.S., Tardon, A., Rennert, G., et al., 2017. Large-scale association analysis identifies new lung cancer susceptibility loci and heterogeneity in genetic susceptibility across histological subtypes. Nat. Genet. 49, 1126–1132. https://doi.org/10.1038/ng.3892.

Ohlsson, H., Kendler, K.S., 2020. Applying causal inference methods in psychiatric epidemiology: a review. JAMA Psychiatry 77, 637–644. https://doi.org/10.1001/jamapsychiatry.2019.3758.

Schizophrenia Working Group of the Psychiatric Genomics Consortium, Ripke, S., Neale, B.M., Corvin, A., Walters, J.T.R., Farh, K.H., Holmans, P.A., Lee, P., Bulik-Sullivan, B., Collier, D.A., Huang, H., Pers, T.H., Agartz, I., Agerbo, E., Albus, M., Alexander, M., Amin, F., Bacanu, S.A., Begemann, M., Belliveau, R.A., Bene, J., Bergen, S.E., Bevilacqua, E., Bigdeli, T.B., Black, D.W., et al., 2014. Biological insights from 108 schizophrenia-associated genetic loci. Nature 511, 421–427. https://doi.org/10.1038/nature13595.

Shrine, N., Guyatt, A.L., Erzurumluoglu, A.M., Jackson, V.E., Hobbs, B.D., Melbourne, C. A., Batini, C., Fawcett, K.A., Song, K., Sakornsakolpat, P., Li, Xingnan, Boxall, R., Reeve, N.F., Obeidat, M., Zhao, J.H., Wielscher, M., Weiss, S., Kentistou, K.A., Cook, J.P., Sun, B.B., Zhou, J., Hui, J., Karrasch, S., Imboden, M., Harris, S.E., Marten, J., et al., 2019. New genetic signals for lung function highlight pathways and chronic obstructive pulmonary disease associations across multiple ancestries. Nat. Genet. 51, 481–493. https://doi.org/10.1038/s41588-018-0321-7.

Terry, A.V., Callahan, P.M., Hernandez, C.M., 2015. Nicotinic ligands as multifunctional agents for the treatment of neuropsychiatric disorders. Biochem. Pharmacol. 97, 388–398. https://doi.org/10.1016/j.bcp.2015.07.027.

Wilkerson, J.L., Deba, F., Crowley, M.L., Hamouda, A.K., McMahon, L.R., 2020. Advances in the in vitro and in vivo pharmacology of Alpha4beta2 nicotinic receptor positive allosteric modulators. Neuropharmacology 168, 108008. https://doi.org/10.1016/j.neuropharm.2020.108008.

Wootton, R.E., Richmond, R.C., Stuijfzand, B.G., Lawn, R.B., Sallis, H.M., Taylor, G.M.J., Hemani, G., Jones, H.J., Zammit, S., Davey Smith, G., Munafò, M.R., 2019. Evidence for causal effects of lifetime smoking on risk for depression and schizophrenia: a Mendelian randomisation study. Psychol. Med. 6, 1–9. https://doi.org/10.1017/S0033291719002678.

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

Estudo 4: Al-Soufi, L., Martorell, L., Moltó, M. D., González-Peñas, J., García-Portilla, M. P., Arrojo, M., Rivero, O., Gutiérrez-Zotes, A., Nácher, J., Muntané, G., Paz, E., Páramo, M., Bobes, J., Arango, C., Sanjuan, J., Vilella, E. e Costas, J. (2021). A polygenic approach to the association between smoking and schizophrenia. *Addiction biology, 27*(1), e13104. https://doi.org/10.1111/adb.13104.

USC
UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

Received: 23 April 2021 | Revised: 18 August 2021 | Accepted: 20 September 2021

DOI: 10.1111/adb.13104

ORIGINAL ARTICLE

Addiction Biology SSA WILEY

A polygenic approach to the association between smoking and schizophrenia

Laila Al-Soufi[1,2] | Lourdes Martorell[3,4] | M.Dolores Moltó[4,5,6] | Javier González-Peñas[4,7] | Ma Paz García-Portilla[4,8] | Manuel Arrojo[1,9] | Olga Rivero[4,5,6] | Alfonso Gutiérrez-Zotes[3,4] | Juan Nácher[4,5,10] | Gerard Muntané[3,4] | Eduardo Paz[1,9] | Mario Páramo[1,9] | Julio Bobes[4,8] | Celso Arango[4,7] | Julio Sanjuan[4,5,11] | Elisabet Vilella[3,4] | Javier Costas[1,12]

[1]Psychiatric Genetics Group, Instituto de Investigación Sanitaria de Santiago de Compostela (IDIS), Santiago de Compostela, Spain

[2]Department of Zoology, Genetics and Physical Anthropology, Universidade de Santiago de Compostela (USC), Santiago de Compostela, Spain

[3]Hospital Universitari Institut Pere Mata (HUIPM); Institut d'Investigació Sanitària Pere Virgili (IISPV); Universitat Rovira i Virgili (URV), Reus, Spain

[4]Spanish Mental Health Research Network (CIBERSAM), Madrid, Spain

[5]INCLIVA Biomedical Research Institute, Fundación Investigación Hospital Clínico de Valencia, Valencia, Spain

[6]Department of Genetics, Universitat de València, Valencia, Spain

[7]Department of Child and Adolescent Psychiatry, Institute of Psychiatry and Mental Health, Hospital General Universitario Gregorio Marañón, School of Medicine, Universidad Complutense de Madrid, Instituto de Investigación Sanitaria Gregorio Marañón (IiSGM), Madrid, Spain

[8]Department of Psychiatry, Universidad de Oviedo; Instituto de Investigación Sanitaria del Principado de Asturias (ISPA); Instituto Universitario de Neurociencias del Principado de Asturias (INEUROPA); Servicio de Salud del Principado de Asturias (SESPA), Oviedo, Spain

[9]Servizo de Psiquiatría, Complexo Hospitalario Universitario de Santiago de Compostela, Servizo Galego de Saúde (SERGAS), Santiago de Compostela, Spain

[10]Department of Cell Biology, Interdisciplinary Research Structure for Biotechnology and Biomedicine (BIOTECMED), Universitat de València, Valencia, Spain

[11]Department of Psychiatric, School of Medicine, Universitat de València, Valencia, Spain

[12]Servizo Galego de Saúde (SERGAS), Complexo Hospitalario Universitario de Santiago de Compostela (CHUS), Santiago de Compostela, Spain

Correspondence
Javier Costas, Psychiatric Genetics Group, Instituto de Investigación Sanitaria de Santiago de Compostela (IDIS), Santiago de Compostela, Spain.
Email: javier.costas.costas@sergas.es

Funding information
Government of the Principality of Asturias, Grant/Award Number: PCTI-2018-2022 IDI/2018/235; Cerca Programme (Generalitat de Catalunya); Broad Institute; ERDF; Ministerio de Ciencia, Innovación y Universidades, Grant/Award Number: FPU18/03243; Instituto de Salud Carlos III, Grant/Award Numbers: ISCIII/PI17/01424, PT17/0019

Abstract

Smoking prevalence in schizophrenia is considerably larger than in general population, playing an important role in early mortality. We compared the polygenic contribution to smoking in schizophrenic patients and controls to assess if genetic factors may explain the different prevalence. Polygenic risk scores (PRSs) for smoking initiation and four genetically correlated traits were calculated in 1108 schizophrenic patients (64.4% smokers) and 1584 controls (31.1% smokers). PRSs for smoking initiation, educational attainment, body mass index and age at first birth were associated with smoking in patients and controls, explaining a similar percentage of variance in both groups. Attention-deficit hyperactivity disorder (ADHD) PRS was associated with smoking only in schizophrenia. This association remained significant after adjustment by psychiatric cross-disorder PRS. A PRS combining all the traits was more explanative than smoking initiation PRS alone, indicating that genetic susceptibility to the other traits plays an additional role in smoking behaviour. Smoking initiation PRS was also associated with schizophrenia in the whole sample, but the significance was lost after adjustment for smoking status. This same pattern was observed in the analysis of specific SNPs at the *CHRNA5–CHRNA3–CHRNB4* cluster associated with both

 © 2021 Society for the Study of Addiction

traits. Overall, the results indicate that the same genetic factors are involved in smoking susceptibility in schizophrenia and in general population and are compatible with smoking acting, directly or indirectly, as a risk factor for schizophrenia that contributes to the high prevalence of smoking in these patients. The contrasting results for ADHD PRS may be related to higher ADHD symptomatology in schizophrenic patients.

KEYWORDS
genetic correlation, genetic susceptibility, nicotine, pleiotropy, psychosis, tobacco

1 | INTRODUCTION

The life expectancy in schizophrenia is reduced compared with the general population by almost 15 years.[1] Smoking is an important contributor to this mortality gap, due to its higher prevalence in schizophrenic patients (SzP) than in general population.[2] This association has been usually explained by the self-medication hypothesis, which states that SzP may smoke to reduce symptoms of the disorder or side effects of antipsychotic medication.[3,4] However, prospective studies have recently suggested a risk effect of smoking on the development of schizophrenia, finding an association between tobacco use and subsequent psychotic experiences[5,6] and schizophrenia.[7,8]

Interestingly, genome-wide association studies (GWAS) revealed that the *CHRNA5–CHRNA3–CHRNB4* cluster on chromosome 15, coding for neuronal nicotinic acetylcholine receptor subunits, is associated with schizophrenia and different smoking traits.[9–13] Consequently, the reason for the high prevalence of smoking in SzP is currently under debate, and the consideration of tobacco as a risk factor for schizophrenia is becoming strongly supported.[14–16] Mendelian randomization (MR) studies, which use genetic variants associated with the risk factor of interest as instrumental variables to assess causality, are an alternative to observational studies. The largest MR study on this topic supported both non-exclusive hypotheses, presenting strong evidence for a casual effect of smoking behaviour on schizophrenia as well as some evidence for a causal effect of schizophrenia on lifetime smoking.[17]

Smoking behaviour, assessed by different smoking phenotypes, is genetically correlated with schizophrenia.[18] PRSs, as a measure of individual genetic susceptibility to a trait, were previously used to demonstrate a shared genetic susceptibility, that is, pleiotropy, between schizophrenia and tobacco use based on the association of the PRS for schizophrenia or smoking with the other trait in a target sample.[19–21] However, these studies are not able to determine if the shared genetic variants independently contribute to each trait, that is, horizontal pleiotropy, or if both traits are in the same causal path, that is, vertical pleiotropy. Information on smoking behavior in both SzP and controls is required to differentiate between both types of pleiotropy.[16,22] Specifically, the association of smoking behaviour PRS with schizophrenia will persist after adjustment for smoking status only in case of horizontal pleiotropy.

The aim of the present study was to assess if genetic susceptibility to smoke differs between SzP and general population. Specifically, we compared the percentage of variance explained in smoking status by smoking PRS from GWAS of general population in SzP versus controls. Additionally, we compared the percentage of variance explained in smoking status by PRSs for traits genetically correlated with smoking. If SzP smoke mainly due to different reasons than general population, association of PRSs for these traits with smoking status will be weaker in SzP than in controls. We also assessed the association of smoking PRS with schizophrenia after adjustment by smoking status, as well as the specific effect of the cluster of neuronal nicotinic acetylcholine receptor subunits at chromosome 15, in order to add new data on the debate about the cause behind shared genetic susceptibility.

2 | MATERIALS AND METHODS

2.1 | Target samples

The initial target sample comprised 1239 patients with a diagnosis of schizophrenia according to DSM-IV and 1803 controls from five different recruitment sites of Spain at Santiago de Compostela, Oviedo, Reus, Valencia and Madrid. All subjects were from Spain and of European origin. Around 75% of controls from Santiago de Compostela were unscreened blood donors. The remaining controls were screened subjects to discard main psychiatric disorders. Adolescence/young adulthood is the relevant period for initiation of smoking and transition to regular smoking. Therefore, subjects under 20 years of age were excluded from further analysis. Recruitment of subjects was carried out at different periods of time from 1999 to 2017, of relevance taking into account that the first smoking-free law in Spain was implemented at the beginning of 2006. Subjects were recruited after informed consent; research was approved by the appropriate ethics committees and followed the principles of the Declaration of Helsinki.

Smoking status of the subjects, considered as a binary trait at the time of assessment, was determined by clinical interview or, if this information is missed, by a biological measure of smoking, methylation analysis of the aryl hydrocarbon receptor repressor *(AHRR)* gene. Quantitative methylation analysis was performed using the Agena

(formerly Sequenom) MassARRAY technology and the assay described in Tian et al.[23] (see Figures S1 and S2 and Table S1 for more information). Youden's index (specificity + sensitivity − 1) was used to determine the optimal cut-off point for classification based on methylation levels in a training sample formed by 514 individuals for which smoking status by clinical interview was also known. A grey zone with low prediction ability was defined as the 95% CI of the optimal cut-off point, estimated by 5000 bootstrap resampling replicates, using pROC v.1.16.2.[24] Those samples with a methylation level within the grey zone ($N = 219$) were removed from the analysis. By this procedure, the diagnostic accuracy of the test, measured by the area under the receiver operating characteristic curve (AUC ROC), was 0.9214.

Genotyping was performed using the Illumina Infinium PsychArray at the Broad Institute or the Spanish National Genotyping Center. General quality control filters were applied using PLINK v1.9[25] including the percentage of failed genotypes by sample (2.5%) and by single nucleotide polymorphism (SNP) (1%), departure for Hardy–Weinberg equilibrium ($p < 1 \times 10^{-5}$), sex incongruencies, excess of heterozygosity (>3 sd), cryptic relatedness (proportion of genome identical by descent >0.2) or a difference in missing genotype rates between cases and controls ($p < 0.001$). Autosomal SNPs at minor allele frequency (MAF) higher than 1% were retained for imputation using the Michigan Imputation Server v1.0.4.[26] Prior to imputation, data were prepared and checked following McCarthy's lab instructions (https://www.well.ox.ac.uk/~wrayner/tools/). Human Reference Consortium dataset v.1.1 was used as haplotype references.[27] After imputation, SNPs with imputation quality Rsq > 0.8 and MAF > 0.01 were retained for analysis.

PLINK v1.9 was used for principal component analysis to obtain the first 10 principal components to exclude outliers and control for population structure. Regions of long-range linkage disequilibrium (LD) that determined the existence of specific principal components in a preliminary analysis, such as the major histocompatibility complex at chromosome 6, the inversion polymorphism at 8p23 or complex rearrangement at the odorant receptors cluster at 11q11, were removed prior to the final principal component analysis.

The final sample comprised 2692 subjects. Of them, 1205 individuals were classified as smokers, 713 SzP and 492 controls, and 1487 individuals were classified as non-smokers, 395 SzP and 1092 controls. The main characteristics of the samples are shown in Table 1.

2.2 | Discovery samples

Several discovery samples were used for PRS calculation (Table S2). Smoking status risk estimation was based on smoking initiation (SmkInit) GWAS data from the GWAS and Sequencing Consortium of Alcohol and Nicotine use (GSCAN).[11] SmkInit was defined as a binary phenotype that compared ever versus never regular smokers.[11] The choice of this discovery sample was based on its large sample size and its similarity with the smoking phenotype collected in our target sample, taking into account the low specificity of PRSs based on different smoking behaviours.[28,29]

GWAS on additional traits were also used as discovery sample in PRS analysis. Specifically, all traits with a significant genetic correlation at the Bonferroni-correction level with SmkInit according to Liu et al.[11] and with publicly available GWAS summary statistics data were selected for power calculation, estimated using the R package AVENGEME.[30] Those traits with power >80% in SzP and/or controls were used in analyses, specifically age at first birth (AFB), that is, age when having the first live child, attention-deficit hyperactivity disorder (ADHD), educational attainment (EA) and body mass index (BMI) (Table S2).

The last cross-disorder GWAS meta-analysis from the Psychiatric Genomics Consortium (PGC), not tested by Liu et al.,[11] was also used as a discovery sample to adjust by general psychopathology risk.[31] In this cross-disorder meta-analysis, GWAS data for eight psychiatric disorders (anorexia nervosa, autism spectrum disorder, ADHD, bipolar

TABLE 1 Characteristics of the samples included in the main analysis

Recruitment site	Group[a]	Sample size	Females (%)	Age (median, IQR)[b]	Smokers (%)	Methylation (%)[c]	Recruitment after 2006 (%)
Santiago	Controls	676	44.4	44.0, 35.0–51.0	20.6	68.5	27.2
	SzP	453	31.3	46.0, 37.0–59.0	60.5	76.4	0
Reus	Controls	538	53.2	41.0, 31.0–53.0	36.8	5.8	0
	SzP	322	35.1	45.5, 33.0–63.0	65.8	4.7	6.5
Valencia	Controls	252	32.5	37.0, 29.0–44.0	46.8	100	72.2
	SzP	272	31.6	40.5, 33.0–48.0	72.1	28.3	38.2
Madrid	Controls	118	55.9	38.0, 29.0–47.0	31.4	55.1	100
	SzP	6	16.7	27.5, 24.0–34.8	50.0	33.3	100
Oviedo	SzP	55	32.7	40.0, 33.0–49.5	50.9	100	0
All sites	Controls	1584	46.4	41.0, 32.0–51.0	31.1	51.2	30.6
	SzP	1108	32.5	44.0, 34.0–56.2	64.4	44.7	11.8

[a]SzP: schizophrenic patients.
[b]Age at recruitment.
[c]Percentage of subjects whose smoking status was determined by methylation as information of smoking status by clinical interview was missing.

disorder, major depression, obsessive–compulsive disorder, Tourette syndrome and schizophrenia) were analysed.

2.3 | PRS calculation

EB-PRS, a novel PRS estimation method based on empirical Bayes theory, was used for PRS calculation.[32] This approach estimates the effect sizes from the SNPs inferring the prior overall distribution of effect sizes of all SNPs in the discovery sample and leveraging this distribution to minimize prediction error. External information or parameter selection are not needed. For quantitative traits, the required case and control sample sizes were set as half of the sample size. In order to deal with LD, PRS models were built with the clumping algorithm of PLINK v.1.9, considering an $r^2 = 0.1$ and a window size of 500 kb to select those independent SNPs with a more significant association in the discovery sample. Prior to clumping, uncommon SNPs (MAF < 1%) and ambiguous SNPs (A/T or C/G alleles) were removed. Finally, PLINK v.1.9 was used to calculate the PRS for each individual in target sample as the sum of the number of effect alleles for all SNPs in the PRS model carried by that individual weighted by its effect size estimated with EB-PRS.

Due to the novelty of the EB-PRS method, EB-PRS was compared with the widely used method of thresholding, considering 31 significance thresholds ranging from 5×10^{-8} to 1. The weight of the SNPs was defined as the logarithm of the odds ratio. Correlation between the PRSs estimated at each threshold and with EB-PRS is shown in Figure S3.

2.4 | Multitrait PRS calculation

SMTpred was used to combine SmkInit PRS and PRSs for those genetically correlated traits with enough power (AFB, ADHD, EA and BMI) in a multitrait PRS to improve prediction.[33] SMTpred combines the different PRSs using weights based on sample size and SNP heritability of the GWAS discovery sample as well as genetic correlation between the different individual PRSs. SNP heritability and genetic correlation from GWAS summary statistics were calculated using LD score regression.[34] LD scores based on 1000 Genomes European data were used, and GWAS summary statistics were filtered to HapMap3 SNPs. In the case of SNP heritability, proportion of cases and prevalence in discovery sample were used for the conversion to liability scale (Table S2).

2.5 | Statistical analyses

All statistical analyses were conducted after the standardization of the PRSs to a mean of 0 and a standard deviation of 1 for easy interpretation. The main analysis tested the PRS association with smoking status (smoker/non-smoker) of the target sample stratified by schizophrenic status (SzP/controls). PRS association with smoking status was assessed by a logistic regression model including sex, age at recruitment, recruitment site, recruitment year (coded as a binary variable considering recruitment before or after the first Spanish smoking-free law in January 2006), the method through which smoking status was determined and 10 principal components of ancestry as covariates. The percentage of variance explained by the PRS was estimated as the increase in R^2 between the full model and the model without the PRS. Age was missing for 3.0% of samples. In these subjects, the mean value of subjects at the same recruitment size, sex, smoking status and schizophrenic status was used as age. Recruitment year was missing for 3.2% of subjects. The mode for the same recruitment site and schizophrenic status was used in those instances.

The R^2 proposed by Lee et al.[35] was used to estimate the proportion of variance explained by the PRS on the liability scale corrected for ascertainment bias due to oversampling of cases relative to the population prevalence. The main advantage of this estimate is its validity to compare the goodness of fit of different models. Because of that, it is the main statistic presented across the text. A prevalence of schizophrenia in general population of 1% was assumed for the calculation of the R^2 on the liability scale. The prevalences of smoking in SzP (54.54%) and in controls (31.51%) were taken from the CLAMORS schizophrenia cohort, a multicentre cross-sectional study based on Spanish population.[2] As a sensitivity analysis, we estimated the R^2 on the liability scale using different prevalence values. CI for the increase in R^2 was estimated through bootstrap resampling ($N = 5000$), applying the bias-corrected and accelerated interval (BCa) method.

Significance of the interaction between schizophrenic status and each PRS on smoking status was assessed by inclusion of the interaction term in the logistic regression model. The model also included interaction terms between all covariates and PRS and between all covariates and schizophrenic status to control for the effects that these covariates might have on the interaction term of interest.[36] Post hoc power analysis for the test of interaction was assessed by simulation in R, as indicated (see Figure S4 for details).[37]

Association of SNPs at *CHRNA5–CHRNA3–CHRNB4* cluster was analysed by logistic regression using an additive genetic model. Taking into account that the strongest signals in the GWAS of schizophrenia and smoking traits did not co-localized,[38] two SNPs were considered: the functional SNP D398N at *CHRNA5* (dbSNP ID: rs16969968),[39] associated with several smoking traits[11,12] and the most significantly associated SNP with schizophrenia at this cluster, rs8042374.[9] Both SNPs are in partial LD ($r^2 = 0.18$ in European samples of the 1000 Genomes Phase 3 according to LDlink[40]).

3 | RESULTS

3.1 | Contribution of SmkInit PRS to smoking status in SzP and controls

Our study was well powered to analyse the effect of SmkInit PRS on smoking status in our samples (Table S2). SmkInit PRS was positively associated with smoking status in both SzP and controls (Table 2). The

TABLE 2 Results of association analysis between polygenic risk scores (PRSs) and smoking status in schizophrenic patients (SzP) and controls

					Adjusted by SmkInit PRS	
PRS	**Target sample**	**Increase in Nagelkerke's R^2**	**Odds ratio (95% CI)**	***P*-value**	**Odds ratio (95% CI)**	***P*-value**
SmkInit	SzP	1.54	1.30 (1.13–1.49)	1.65×10^{-4a}	-	-
	Controls	1.68	1.30 (1.16–1.45)	6.97×10^{-6a}	-	-
AFB	SzP	0.89	0.82 (0.71–0.94)	4.07×10^{-3a}	0.85 (0.74–0.99)	0.0313
	Controls	0.47	0.87 (0.78–0.98)	0.0169	0.92 (0.82–1.03)	0.1560
ADHD	SzP	1.07	1.24 (1.08–1.42)	1.65×10^{-3a}	1.19 (1.04–1.36)	0.0129
	Controls	0.22	1.10 (0.98–1.23)	0.0991	1.06 (0.94–1.19)	0.3230
EA	SzP	1.10	0.80 (0.70–0.92)	1.45×10^{-3a}	0.85 (0.74–0.98)	0.0206
	Controls	0.88	0.82 (0.73–0.93)	1.13×10^{-3a}	0.87 (0.77–0.98)	0.0238
BMI	SzP	0.51	1.16 (1.02–1.33)	0.0291	1.11 (0.97–1.27)	0.1355
	Controls	0.42	1.14 (1.02–1.27)	0.0244	1.10 (0.98–1.23)	0.1007
Multitrait	SzP	2.43	1.39 (1.21–1.59)	2.41×10^{-6a}	-	-
	Controls	2.20	1.35 (1.20–1.52)	2.76×10^{-7a}	-	-

Abbreviations: ADHD, attention-deficit hyperactivity disorder; AFB, age at first birth; BMI, body mass index; EA, educational attainment; SmkInit, smoking initiation.

[a]Significant association at Bonferroni's correction level ($P < 4.17 \times 10^{-3}$, corrected for 12 tests).

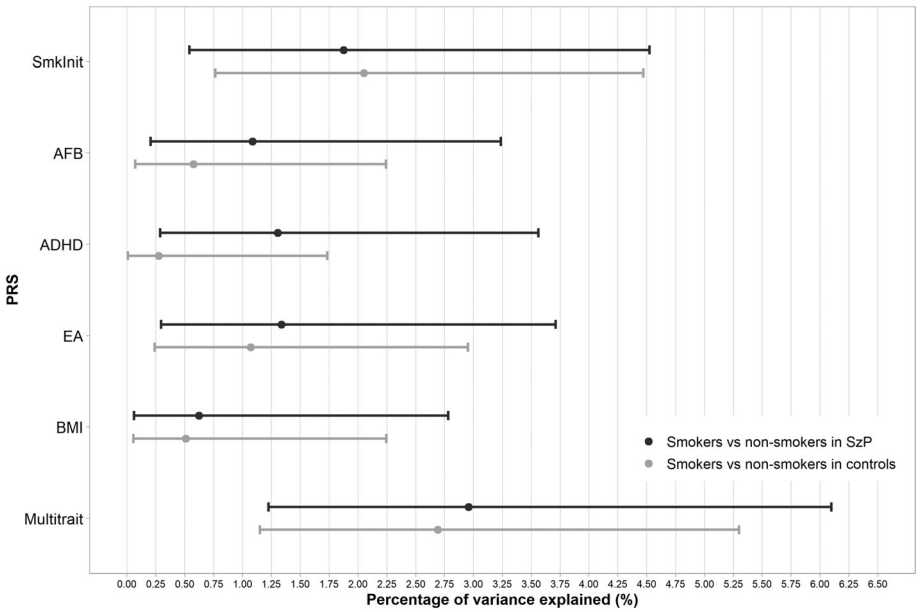


FIGURE 1 Percentage of variance explained of smoking status (and 95% CI) in schizophrenic patients (SzP) and controls by the different polygenic risk scores (PRSs). ADHD, attention-deficit hyperactivity disorder; AFB, age at first birth; BMI, body mass index; EA, educational attainment; SmkInit, smoking initiation

percentage of variance explained, corrected by prevalence and ascertainment bias, was slightly higher in controls (increase in R^2 (95% CI) $= 2.05$ $(0.76 - 4.47)$) than in SzP (increase in R^2 (95% CI) $= 1.87$ $(0.54 - 4.52)$) (Figure 1 and Table S3). Interaction term between SmkInit PRS and schizophrenic status had no significant effect on smoking status ($p = 0.15$). Sensitivity analysis using different prevalence values confirmed this finding (Table S4).

3.2 | Contribution of PRSs for SmkInit genetically correlated traits to smoking status in SzP and controls

Among those traits genetically correlated with SmkInit, statistical power to analyse the contribution of the PRS to smoking status was higher than 80% in case of AFB and ADHD in both SzP and controls and EA and BMI in controls. Statistical power in SzP was 74% and 69% for EA PRS and BMI PRS, respectively (Table S2). All the PRSs were significant in SzP. AFB PRS, EA PRS and BMI PRS were also significant in controls (Table 2). The different PRSs explained a higher percentage of variance in SzP than in controls (Figure 1 and Table S3). Interaction terms between each PRS and schizophrenic status had no significant effect on smoking status ($p = 0.17$, 0.40, 0.41 and 0.36 for AFB PRS, ADHD PRS, EA PRS and BMI PRS, respectively), although the tests were substantially underpowered (Figure S4). As expected, they explained less variance in smoking status than SmkInit PRS. All associations were in the expected direction, that is, PRS was negatively associated with smoking in case of AFB and EA, and positively associated in case of the remaining traits (Table 2).

To help interpretation of ADHD PRS results, PRS for a psychiatric cross-disorder meta-analysis was analysed. Correlation between ADHD PRS and cross-disorder PRS was moderate ($r_g = 0.2257$). Statistical power to analyse the effect of cross-disorder PRS on smoking status was higher than 80% in both SzP and controls (Table S2). Cross-disorder PRS was not significantly associated with smoking status in SzP or in controls (Table S5). ADHD PRS remained significant in SzP and not significant in controls after including cross-disorder PRS in regression model (SzP: OR (95% CI) = 1.23 (1.07 − 1.41), $p = 4.13 \times 10^{-3}$; controls: OR (95% CI) = 1.12 (0.99 − 1.25), $p = 0.0633$).

3.3 | Joint contribution of the different PRSs to smoking status in SzP and controls

The analysed PRSs were moderately correlated among them, and the pattern of correlation was similar in SzP and controls (Figure 2). Inclusion of SmkInit PRS in the regression models attenuated the effect of the PRSs for SmkInit genetically correlated traits in all cases, although EA PRS was still nominally significant in both SzP and controls, and AFB and ADHD PRSs remained nominally significant in SzP (Table 2). Multitrait prediction, combining the five PRSs using weights estimated by SMTpred, substantially improved the prediction based on SmkInit PRS alone (Tables 2 and S6). The percentage of variance explained in smoking status by the multitrait PRS was higher in SzP than in controls (Figure 1 and Table S3), although the interaction term between multitrait PRS and schizophrenic status did not reach significance ($p = 0.06$).

3.4 | Contribution of SmkInit PRS to schizophrenic status

Statistical power to analyse the contribution of SmkInit PRS to schizophrenic status in the whole sample (SzP and controls) was close to 80%. SmkInit PRS had a nominally significant effect in schizophrenic status in the whole sample ($p = 0.0439$), but significance was lost after adjustment by smoking status ($p = 0.8599$). Statistical power to analyse the association between SmkInit PRS and schizophrenic status in the sample stratified by smoking status was low. SmkInit PRS was not associated with schizophrenic status neither in smokers nor in non-smokers, and the percentage of variance explained was similar in both groups (Table S7).

3.5 | Contribution of the CHRNA5–CHRNA3–CHRNB4 cluster to schizophrenic status and smoking status

rs16969968 was significantly associated with smoking status, independently of adjustment by schizophrenic status or sample stratification by schizophrenic status. This SNP was not significantly associated with schizophrenic status. rs8042374 was associated with schizophrenic status in the whole sample, but significance was lost after sample stratification by smoking status or adjustment by smoking status. This SNP was associated with smoking status in the whole sample, also after adjustment by schizophrenic status, and in SzP (Table 3).

4 | DISCUSSION

Our analysis revealed that SmkInit PRS explained a similar amount of variance in susceptibility to regular smoking in SzP and in controls. Additional PRSs based on traits genetically correlated with SmkInit also contribute to smoking susceptibility, as reflected by the larger percentage of variance explained by a multitrait PRS. As expected, SmkInit PRS was the PRS that explained the highest variability in controls. This PRS also explained the highest variability in SzP. The same PRSs involved in susceptibility to regular smoking in controls, such as

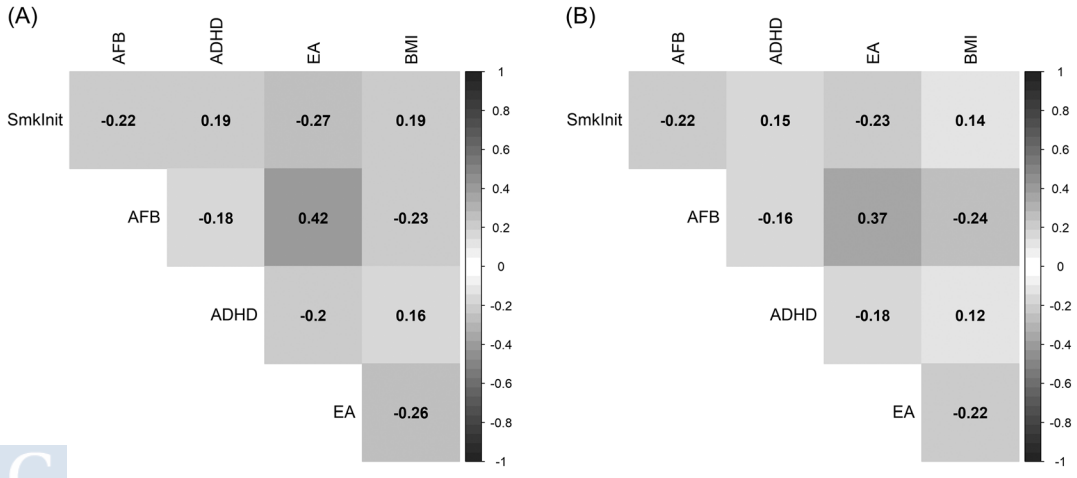


FIGURE 2 Pearson's correlation between the polygenic risk scores (PRSs) analysed in this study in schizophrenic patients (SzP) (A) and controls (B). $P < 5 \times 10^{-3}$ in all comparisons. ADHD, attention-deficit hyperactivity disorder; AFB, age at first birth; BMI, body mass index; EA, educational attainment; SmkInit, smoking initiation

TABLE 3 Association analysis of SNPs at the *CHRNA5–CHRNA3–CHRNB4* cluster with smoking status and schizophrenic status

						Adjusted[b]	
SNP	**Trait**	**Sample**	**Neff[a]**	**Odds ratio (95% IC)**	***P*-value**	**Odds ratio (95% IC)**	***P*-value**
rs16969968	Smoking status	Total	2662	1.23 (1.10–1.39)	4.85×10^{-4}	1.23 (1.09–1.39)	1.13×10^{-3}
		SzP	1017	1.22 (1.00–1.48)	0.0462	-	-
		Controls	1357	1.23 (1.04–1.44)	0.0152	-	-
	Schizophrenic status	Total	2608	1.10 (0.97–1.25)	0.1516	1.03 (0.90–1.18)	0.6460
		Smokers	1164	1.00 (0.83–1.22)	0.9722	-	-
		Non-smokers	1160	1.06 (0.88–1.29)	0.5314	-	-
rs8042374	Smoking status	Total	2662	1.25 (1.09–1.44)	1.55×10^{-3}	1.22 (1.05–1.41)	8.71×10^{-3}
		SzP	1017	1.28 (1.01–1.61)	0.0387	-	-
		Controls	1357	1.17 (0.96–1.42)	0.1135	-	-
	Schizophrenic status	Total	2608	1.19 (1.02–1.38)	0.0283	1.10 (0.94–1.30)	0.2294
		Smokers	1164	1.15 (0.90–1.46)	0.2609	-	-
		Non-smokers	1160	1.05 (0.84–1.32)	0.6474	-	-

Abbreviation: SzP, schizophrenic patients.
[a]Neff: effective sample size, calculated as twice the harmonic mean.
[b]Association analysis for smoking status adjusted by schizophrenic status and association analysis for schizophrenic status adjusted by smoking status.

EA PRS, AFB PRS and BMI PRS, affect susceptibility to regular smoking in SzP. The only exception was ADHD PRS, which is associated with smoking status only in SzP. Among the PRSs for traits genetically correlated with SmkInit, EA PRS explained the highest percentage of variance, both in SzP and controls. These findings point to the existence of common genetic susceptibility to smoking in SzP and in general population.

There was just one PRS for a psychiatric trait with enough power to be tested, that for ADHD. The association of this PRS with smoking in SzP but not in controls was the main difference between both groups, although the difference in effect was not statistically significant. Both ADHD and ADHD symptoms in non-clinical samples are associated with smoking.[41,42] MR studies suggested a casual effect of ADHD on smoking behaviour,[43,44] but not on nicotine dependence, probably due to lack of power.[45] ADHD PRS is associated with current smoking in the UK Biobank sample, explaining 0.33% (Nagelkerke's R^2) of the variance.[46] This value is similar to that detected in our controls, 0.22%, and considerably lower than that detected in our schizophrenia sample, 1.07%. The higher variance explained in SzP, pending on confirmation in additional samples, might be related to the considerably higher prevalence of ADHD symptoms in SzP than in general population, reaching about a quarter of patients.[47] Alternatively, it may be due to pleiotropic genetic effects responsible for the high genetic correlation among psychiatric disorders,[31] including substance use disorders.[45,48] Thus, nicotine dependence was strongly correlated at the genetic level with ADHD ($r_g = 0.53$), but also with major depressive disorder ($r_g = 0.42$) and to a lesser extent, with schizophrenia ($r_g = 0.18$) and bipolar disorder ($r_g = 0.13$).[45] This shared genetic susceptibility may reflect the existence of a general genetic factor of psychopathology.[49,50] To analyse if the association of ADHD PRS with smoking in SzP was due to this general genetic psychopathology factor, we tested the association of cross-disorder PRS with smoking, as well as the effect of adjustment of ADHD PRS by cross-disorder PRS. Cross-disorder PRS was not associated and ADHD PRS remained significant in SzP, suggesting that the association is due to genetic factors specific to ADHD symptomatology. Therefore, future studies should be performed evaluating heterogeneity of ADHD symptoms in SzP and its involvement in tobacco and other toxic substances use and dependence.

A related question is why SmkInit is genetically correlated with schizophrenia. One explanation may be horizontal pleiotropy, considering substance use disorders as different manifestations of the general genetic factor of psychopathology, as suggested previously.[45,51] Alternatively, taking into account that the same genetic factors are involved in smoking behaviour in SzP and general population, as revealed by our work, this genetic correlation may be consequence of the higher prevalence of smoking in SzP versus controls in the schizophrenia GWAS. The higher prevalence is compatible with smoking acting as risk factor for schizophrenia, that is, vertical pleiotropy. As expected under this scenario, we detected an association of the SmkInit PRS with schizophrenia in our sample that was lost when the model was adjusted by smoking status. Similarly, there was no association between SmkInit PRS and schizophrenia when the sample was split in smokers or non-smokers. However, we cannot distinguish between a direct effect of smoking on schizophrenia risk and an indirect effect mediated, for instance, by increasing use of additional substances.

Additionally, we assessed this question focusing on the cluster of neuronal nicotinic cholinergic receptor subunits at chromosome 15. Two SNPs were tested: the missense SNP rs16969968 at *CHRNA5*, strongly associated with several smoking traits,[11,12] and rs8042374, the top association signal at this locus in the schizophrenia GWAS.[9] On the one hand, both SNPs were associated with smoking in the full sample. When stratified by schizophrenic status, the effect of

rs8042374 was restricted to SzP, perhaps due to a Type II error because of low power. On the other hand, association with schizophrenia was present only for rs8042374, and this association was lost after adjustment for smoking status. This is similar to the SmkInit PRS results, strengthening the conclusion that the association between this locus and schizophrenia is due to the higher prevalence of smoking in SzP, probably as a consequence of its role as risk factor.

Recently, the PGC performed a similar analysis on the subset of SzP with smoking data (~5000 patients, no controls).[52] They also concluded that smoking in SzP may be explained, at least partially, by the same genetic factors than in general population, although they did not compare the effect of SmkInit PRS in SzP versus controls. The proportion of variance in SmkInit explained in SzP by the SmkInit PRS was considerably larger in our case (Nagelkerke's $R^2 = 1.54\%$ vs. 0.58%), probably as a consequence of the larger sample size of the GWAS summary statistics used in our study, GSCAN, in comparison with the Tobacco and Genetics (TAG) Consortium data used by the PGC. In addition to SmkInit, their conclusion was also based on cigarettes per day (CPD). The CPD PRS from TAG Consortium predicted CPD in SzP. Finally, genetic correlation between SmkInit in SzP from PGC and data from the TAG Consortium did not differ significantly from 1. Our analysis of PRSs based on traits genetically correlated with SmkInit represents independent evidence strengthening the general conclusion that genetic susceptibility to smoking is shared between SzP and general population.

Our study presents several limitations. First of all, characterization of smoking behavior is incomplete. As a cross-sectional study, we did not take into account temporal changes in smoking status. Furthermore, nicotine dependence or intensity of smoking was not considered. In addition, determination of smoking status was not homogeneous across subjects. More controls than cases were classified as smokers/non-smokers using methylation in our study. In order to reduce this bias, we defined a grey zone with ambiguous methylation levels and discarded the samples at this zone, and we used the method for smoking status determination as a covariate in our regression models. Furthermore, taking into account that case–control samples did not matched by age, we cannot exclude the possibility that a fraction of subjects, different in cases versus controls, did not smoke at recruitment because they were too young for initiation of smoking. According to the Spanish survey on alcohol and other drugs, the average age of initiation of smoking among 15–64 years old people remained stable at ~16.5 from 1995 to 2013. Regular smoking begins 2 years later.[53] Therefore, we removed subjects under 20 years of age to reduce this putative bias. Another limitation was the low power to test for interaction between PRS and disease status and to test PRSs for other traits genetically correlated with SmkInit, including that for schizophrenia, which might inform about severity of symptoms. Finally, we did not get access to additional samples for replication.

In summary, genetic susceptibility to smoking in SzP seems to be due to the same factors than in general population, although we have detected a larger effect of ADHD PRS in SzP, perhaps related to higher ADHD symptomatology in SzP. Therefore, the higher prevalence of smoking in SzP than in general population is compatible with smoking, directly or indirectly, acting as a risk factor for schizophrenia. Hence, smoking may be considered an important modifiable factor in selective and indicated prevention targeting vulnerable individuals.

ACKNOWLEDGEMENTS

This work was supported by the Instituto de Salud Carlos III (ISCIII) (grant number ISCIII/PI17/01424) to JC, co-founded by FEDER. LA-S was the recipient of a research fellowship by the Ministerio de Ciencia, Innovación y Universidades (grant number FPU18/03243). The authors wish to thank all the Consortia and research groups providing the summary statistics datasets used in this work. Genotyping was performed at CEGEN-PRB3-ISCIII, supported by grant PT17/0019 of the PE I + D + i 2013-2016, funded by ISCIII and ERDF and by Broad Institute. Methylation assays were performed at CEGEN-PRB3-ISCIII. Dr. Vilella's group was supported by the Cerca Programme (Generalitat de Catalunya). Dr. Bobes and García-Portilla are also grateful for the support of the Government of the Principality of Asturias PCTI-2018-2022 IDI/2018/235.

CONFLICT OF INTEREST

Julio Bobes has received research grants and served as consultant, advisor or speaker within de last 5 years for AB-Biotics, Acadia Pharmaceuticals, Alkermes, Allergan, Ambrosseti-Angelini, Biogen, Casen Recordati, D&A Pharma, Exeltis, Gilead, Indivior, GW Pharmaceuticals, Janssen-Cilag, Jazz Pharmaceuticals, Lundbeck, Mundipharma, Newron, Otsuka, Pfizer, Roche, Sage Therapeutics, Servier, Schwabe Farma Ibérica, Shire, Takeda, research funding from the Spanish Ministry of Economy and Competiveness–Centro de Investigación Biomedica en Red area de Salud Mental (CIBERSAM) and Instituto de Salud Carlos III, Spanish Ministry of Health.

Maria Paz Garcia-Portilla has been a consultant to and/or has received honoraria/grants from Angelini, Alianza Otsuka-Lundbeck, Instituto de Salud Carlos III, Janssen-Cilag, Lundbeck, Otsuka, Pfizer and SAGE Therapeutics.

AUTHOR CONTRIBUTIONS

JC was responsible for the study concept and design. LA-S performed main data analysis under supervision of JC. LA-S and JC drafted the manuscript. JG-P contributed genotyping data curation. Other authors contributed to recruitment and characterization of study participants, acquisition of data and data curation. All authors critically reviewed content and approved the final version for publication.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available from the corresponding author upon reasonable request. The data are not publicly available due to privacy or ethical restrictions.

ORCID

Laila Al-Soufi https://orcid.org/0000-0003-2975-1636
Javier Costas https://orcid.org/0000-0003-0306-3990

REFERENCES

1. Hjorthøj C, Stürup AE, McGrath JJ, Nordentoft M. Years of potential life lost and life expectancy in schizophrenia: a systematic review and meta-analysis. *Lancet Psychiatry*. 2017;4(4):295-301. doi:10.1016/S2215-0366(17)30078-0
2. Bobes J, Arango C, Garcia-Garcia M, Rejas J. Healthy lifestyle habits and 10-year cardiovascular risk in schizophrenia spectrum disorders: An analysis of the impact of smoking tobacco in the CLAMORS schizophrenia cohort. *Schizophr Res*. 2010;119(1-3):101-109. doi:10.1016/j.schres.2010.02.1030
3. Kumari V, Postma P. Nicotine use in schizophrenia: The self medication hypotheses. *Neurosci Biobehav Rev*. 2005;29(6):1021-1034. doi:10.1016/j.neubiorev.2005.02.006
4. Goff DC, Henderson DC, Amico E. Cigarette smoking in schizophrenia: relationship to psychopathology and medication side effects. *Am J Psychiatry*. 1992;149(9):1189-1194. doi:10.1176/ajp.149.9.1189
5. Gurillo P, Jauhar S, Murray RM, MacCabe JH. Does tobacco use cause psychosis? Systematic review and meta-analysis. *Lancet Psychiatry*. 2015;2(8):718-725. doi:10.1016/S2215-0366(15)00152-2
6. Koyanagi A, Stickley A, Haro JM. Psychotic symptoms and smoking in 44 countries. *Acta Psychiatr Scand*. 2016;133(6):497-505. doi:10.1111/acps.12566
7. Kendler KS, Lönn SL, Sundquist J, Sundquist K. Smoking and schizophrenia in population cohorts of Swedish Women and Men: A prospective co-relative control study. *Am J Psychiatry*. 2015;172(11):1092-1100. doi:10.1176/appi.ajp.2015.15010126
8. Weiser M, Reichenberg A, Grotto I, Yasvitzky R, Rabinowitz J, Lubin G, et al. Higher rates of cigarette smoking in Male adolescents before the onset of schizophrenia: A historical-prospective cohort study. *Am J Psychiatry*. 2004;161(7):1219-1223. doi:10.1176/appi.ajp.161.7.1219
9. Schizophrenia Working Group of the Psychiatric Genomics Consortium RipkeS, Neale BM, et al. Biological insights from 108 schizophrenia-associated genetic loci. *Nature*. 2014;511(7510):421-427. doi:10.1038/nature13595
10. Hancock DB, Reginsson GW, Gaddis NC, Chen X, Saccone NL, Lutz SM, et al. Genome-wide meta-analysis reveals common splice site acceptor variant in CHRNA4 associated with nicotine dependence. *Transl Psychiatry*. 2015;5(10):e651. doi:10.1038/tp.2015.149
11. Liu M, Jiang Y, Wedow R, et al. Association studies of up to 1.2 million individuals yield new insights into the genetic etiology of tobacco and alcohol use. *Nat Genet*. 2019;51(2):237-244. doi:10.1038/s41588-018-0307-5
12. Chen J, Loukola A, Gillespie NA, Peterson R, Jia P, Riley B, et al. Genome-wide meta-analyses of FTND and TTFC phenotypes. *Nicotine Tob Res*. 2020;22(6):900-909. doi:10.1093/ntr/ntz099
13. Ohi K, Kuwata A, Shimada T, Kataoka Y, Yasuyama T, Uehara T, et al. Genome-wide variants shared between smoking quantity and schizophrenia on 15q25 are associated with CHRNA5 expression in the brain. *Schizophr Bull*. 2019;45(4):813-823. doi:10.1093/schbul/sby093
14. Taylor GMJ, Munafò MR. Does smoking cause poor mental health? *Lancet Psychiatry*. 2019;6(1):2-3. doi:10.1016/S2215-0366(18)30459-0
15. MacCabe JH. It is time to start taking tobacco seriously as a risk factor for psychosis: self-medication cannot explain the association. *Acta Psychiatr Scand*. 2018;138(1):3-4. doi:10.1111/acps.12923
16. Gage SH, Munafò MR. Rethinking the association between smoking and schizophrenia. *Lancet Psychiatry*. 2015;2(2):118-119. doi:10.1016/S2215-0366(14)00057-1
17. Wootton RE, Richmond RC, Stuijfzand BG, Lawn RB, Sallis HM, Taylor GMJ, et al. Evidence for causal effects of lifetime smoking on risk for depression and schizophrenia: a Mendelian randomisation study. *Psychol Med*. 2020;50(14):2435-2443. doi:10.1017/S0033291719002678
18. Hartz SM, Horton AC, Hancock DB, Baker TB, Caporaso NE, Chen LS, et al. Genetic correlation between smoking behaviors and schizophrenia. *Schizophr Res*. 2018;194:86-90. doi:10.1016/j.schres.2017.02.022
19. Chen J, Bacanu S-A, Yu H, et al. Genetic Relationship between Schizophrenia and Nicotine Dependence. *Sci Rep*. 2016;6:25671. doi:10.1038/srep25671
20. Reginsson GW, Ingason A, Euesden J, Bjornsdottir G, Olafsson S, Sigurdsson E, et al. Polygenic risk scores for schizophrenia and bipolar disorder associate with addiction. *Addict Biol*. 2018;23(1):485-492. doi:10.1111/adb.12496
21. Ohi K, Nishizawa D, Muto Y, et al. Polygenic risk scores for late smoking initiation associated with the risk of schizophrenia. *NPJ Schizophr*. 2020;6:36. doi:10.1038/s41537-020-00126-z
22. Gage SH, Davey Smith G, Ware JJ, Flint J, Munafò MR. G = E: What GWAS Can Tell Us about the Environment. *PLoS Genet*. 2016;12(2):e1005765. doi:10.1371/journal.pgen.1005765
23. Tian F-Y, Hivert M-F, Wen X, Xie C, Niu Z, Fan L, et al. Tissue differences in DNA methylation changes at AHRR in full term low birth weight in maternal blood, placenta and cord blood in Chinese. *Placenta*. 2017;52:49-57. doi:10.1016/j.placenta.2017.02.017
24. Robin X, Turck N, Hainard A, Tiberti N, Lisacek F, Sanchez JC, et al. pROC: an open-source package for R and S to analyze and compare ROC curves. *BMC Bioinformatics*. 2011;12:77. doi:10.1186/1471-2105-12-77, 1
25. Chang CC, Chow CC, Tellier LCAM, Vattikuti S, Purcell SM, Lee JJ. Second-generation PLINK: rising to the challenge of larger and richer datasets. *Gigascience*. 2015;4:7. doi:10.1186/s13742-015-0047-8, 1
26. Das S, Forer L, Schönherr S, Sidore C, Locke AE, Kwong A, et al. Next-generation genotype imputation service and methods. *Nat Genet*. 2016;48(10):1284-1287. doi:10.1038/ng.3656
27. McCarthy S.A reference panel of 64,976 haplotypes for genotype imputation. *Nat Genet*. 2016;48(10):1279-1283. doi:10.1038/ng.3643
28. Deutsch AR, Selya AS. Stability in effects of different smoking-related polygenic risk scores over age and smoking phenotypes. *Drug Alcohol Depend*. 2020;214:108154. doi:10.1016/j.drugalcdep.2020.108154
29. Sanchez-Roige S, Cox NJ, Johnson EO, Hancock DB, Davis LK. Alcohol and cigarette smoking consumption as genetic proxies for alcohol misuse and nicotine dependence. *Drug Alcohol Depend*. 2021;221:108612. doi:10.1016/j.drugalcdep.2021.108612
30. Palla L, Dudbridge F. A Fast Method that Uses Polygenic Scores to Estimate the Variance Explained by Genome-wide Marker Panels and the Proportion of Variants Affecting a Trait. *Am J Hum Genet*. 2015;97(2):250-259. doi:10.1016/j.ajhg.2015.06.005
31. Lee PH, Anttila V, Won H, Feng YCA, Rosenthal J, Zhu Z, et al. Genomic relationships, novel loci, and pleiotropic mechanisms across eight psychiatric disorders. *Cell*. 2019;179(7):1469-1482.e11. doi:10.1016/j.cell.2019.11.020
32. Song S, Jiang W, Hou L, Zhao H. Leveraging effect size distributions to improve polygenic risk scores derived from summary statistics of genome-wide association studies. *PLoS Comput Biol*. 2020;16(2):e1007565. doi:10.1371/journal.pcbi.1007565
33. Maier RM, Zhu Z, Lee SH, Trzaskowski M, Ruderfer DM, Stahl EA, et al. Improving genetic prediction by leveraging genetic correlations among human diseases and traits. *Nat Commun*. 2018;9(1):989. doi:10.1038/s41467-017-02769-6
34. Bulik-Sullivan B, Finucane HK, Anttila V, et al. An atlas of genetic correlations across human diseases and traits. *Nat Genet*. 2015;47(11):1236-1241. doi:10.1038/ng.3406
35. Lee SH, Goddard ME, Wray NR, Visscher PM. A better coefficient of determination for genetic profile analysis. *Genet Epidemiol*. 2012;36(3):214-224. doi:10.1002/gepi.21614

36. Keller MC. Gene × environment interaction studies have not properly controlled for potential confounders: the problem and the (simple) solution. *Biol Psychiatry*. 2014;75(1):18-24. doi:10.1016/j.biopsych.2013.09.006
37. Kleinman K, Horton NJ. *SAS and R: Data management, statistical analysis and graphics*. 2nd ed. Boca Raton: Chapman & Hall; 2014:425.
38. Al-Soufi L, Costas J. Colocalization of association signals at nicotinic acetylcholine receptor genes between schizophrenia and smoking traits. *Drug Alcohol Depend*. 2021;220:108517. doi:10.1016/j.drugalcdep.2021.108517
39. Maskos U. The nicotinic receptor alpha5 coding polymorphism rs16969968 as a major target in disease: Functional dissection and remaining challenges. *J Neurochem*. 2020;154(3):241-250. doi:10.1111/jnc.14989
40. Machiela MJ, Chanock SJ. LDlink: a web-based application for exploring population-specific haplotype structure and linking correlated alleles of possible functional variants. *Bioinformatics*. 2015;31(21):3555-3557. doi:10.1093/bioinformatics/btv402
41. Kollins SH, McClernon FJ, Fuemmeler BF. Association between smoking and attention-deficit/hyperactivity disorder symptoms in a population-based sample of young adults. *Arch Gen Psychiatry*. 2005;62(10):1142-1147. doi:10.1001/archpsyc.62.10.1142
42. Charach A, Yeung E, Climans T, Lillie E. Childhood attention-deficit/hyperactivity disorder and future substance use disorders: Comparative meta-analyses. *J Am Acad Child Adolesc Psychiatry*. 2011;50(1):9-21. doi:10.1016/j.jaac.2010.09.019
43. Treur JL, Demontis D, Smith GD, et al. Investigating causality between liability to ADHD and substance use, and liability to substance use and ADHD risk, using Mendelian randomization. *Addict Biol*. 2019;26(1):e12849. doi:10.1111/adb.12849
44. Vilar-Ribó L, Sánchez-Mora C, Rovira P, Richarte V, Corrales M, Fadeuilhe C, et al. Genetic overlap and causality between substance use disorder and attention-deficit and hyperactivity disorder. *Am J Med Genet Part B Neuropsychiatr Genet*. 2020. doi:10.1002/ajmg.b.32827, 186, 3, 140, 150
45. Vink JM, Treur JL, Pasman JA, Schellekens A. Investigating genetic correlation and causality between nicotine dependence and ADHD in a broader psychiatric context. *Am J Med Genet Part B Neuropsychiatr Genet*. 2020. doi:10.1002/ajmg.b.32822
46. Du Rietz E, Coleman J, Glanville K, Choi SW, O'Reilly PF, Kuntsi J. Association of polygenic risk for attention-deficit/hyperactivity disorder with co-occurring traits and disorders. *Biol Psychiatry Cogn Neurosci Neuroimaging*. 2018;3(7):635-643. doi:10.1016/j.bpsc.2017.11.013
47. Arican I, Bass N, Neelam K, Wolfe K, McQuillin A, Giaroli G. Prevalence of attention deficit hyperactivity disorder symptoms in patients with schizophrenia. *Acta Psychiatr Scand*. 2019;139(1):89-96. doi:10.1111/acps.12948
48. Abdellaoui A, Smit DJA, van den Brink W, Denys D, Verweij KJH. Genomic relationships across psychiatric disorders including substance use disorders. *Drug Alcohol Depend*. 2021;220:108535. doi:10.1016/j.drugalcdep.2021.108535
49. Pettersson E, Larsson H, Lichtenstein P. Common psychiatric disorders share the same genetic origin: a multivariate sibling study of the Swedish population. *Mol Psychiatry*. 2016;21(5):717-721. doi:10.1038/mp.2015.116
50. Neumann A, Pappa I, Lahey BB, Verhulst FC, Medina-Gomez C, Jaddoe VW, et al. Single nucleotide polymorphism heritability of a general psychopathology factor in children. *J Am Acad Child Adolesc Psychiatry*. 2016;55(12):1038-1045. doi:10.1016/j.jaac.2016.09.498
51. Facal F, Flórez G, Blanco V, Rodríguez J, Pereiro C, Fernández JM, et al. Genetic predisposition to alcohol dependence: The combined role of polygenic risk to general psychopathology and to high alcohol consumption. *Drug Alcohol Depend*. 2021;221:108556. doi:10.1016/j.drugalcdep.2021.108556
52. Peterson RE, Bigdeli TB, Ripke S, Bacanu SA, Gejman PV, Levinson DF, et al. Genome-wide analyses of smoking behaviors in schizophrenia: Findings from the Psychiatric Genomics Consortium. *J Psychiatr Res*. 2021;137:215-224. doi:10.1016/j.jpsychires.2021.02.027
53. Observatorio Español de la Droga y las Toxicomanías. Informe 2015: Alcohol, Tabaco y Drogas Ilegales En España. 2015. http://crm.org.mx/PDF/INFORMES/INFORME2014.pdf

SUPPORTING INFORMATION

Additional supporting information may be found in the online version of the article at the publisher's website.

How to cite this article: Al-Soufi L, Martorell L, Moltó M, et al. A polygenic approach to the association between smoking and schizophrenia. *Addiction Biology*. 2021;e13104. doi:10.1111/adb.13104

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

ANEXO III: PERMISOS DE UTILIZACIÓN DOS ARTIGOS PUBLICADOS

Estudo 3:



CCC RightsLink

Sign in/Register

Colocalization of association signals at nicotinic acetylcholine receptor genes between schizophrenia and smoking traits

Author: Laila Al-Soufi,Javier Costas
Publication: Drug and Alcohol Dependence
Publisher: Elsevier
Date: 1 March 2021

© 2021 Elsevier B.V. All rights reserved.

Journal Author Rights

Please note that, as the author of this Elsevier article, you retain the right to include it in a thesis or dissertation, provided it is not published commercially. Permission is not required, but please ensure that you reference the journal as the original source. For more information on this and on your other retained rights, please visit: https://www.elsevier.com/about/our-business/policies/copyright#Author-rights

BACK CLOSE WINDOW

© 2024 Copyright - All Rights Reserved | Copyright Clearance Center, Inc. | Privacy statement | Data Security and Privacy | For California Residents | Terms and Conditions
Comments? We would like to hear from you. E-mail us at customercare@copyright.com

Estudo 4:



License Number	5797190050440
License date	May 27, 2024

Printable Details

Licensed Content

Licensed Content Publisher	John Wiley and Sons
Licensed Content Publication	Addiction Biology
Licensed Content Title	A polygenic approach to the association between smoking and schizophrenia
Licensed Content Author	Lourdes Martorell, M.Dolores Moltó, Javier González-Peñas, et al
Licensed Content Date	Nov 14, 2021
Licensed Content Volume	27
Licensed Content Issue	1
Licensed Content Pages	10

Order Details

Type of use	Dissertation/Thesis
Requestor type	Author of this Wiley article
Format	Print and electronic
Portion	Full article
Will you be translating?	No

About Your Work

Title of new work	Unha aproximación xenómica á susceptibilidade do tabaquismo e outros trazos de consumo de substancias en pacientes con esquizofrenia e poboación xeral
Institution name	Universidade de Santiago de Compostela
Expected presentation date	Jul 2024

Additional Data

Order reference number	1

Requestor Location

Mrs. Laila Al-Soufi Novo

Tax Details

Publisher Tax ID	EU826007151

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

Dear Mrs. Laila Al-Soufi Novo,

Thank you for placing your order through Copyright Clearance Center's RightsLink® service.

Order Summary

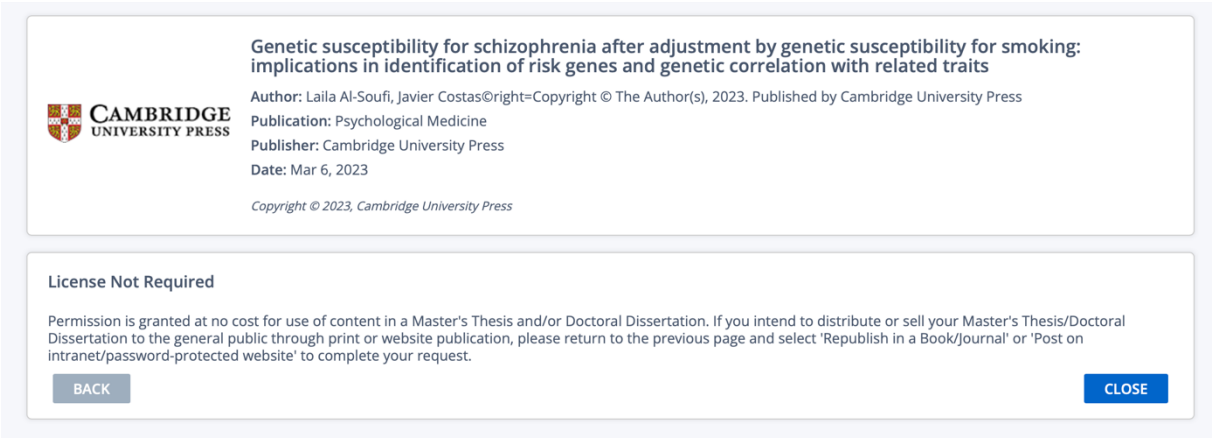
Licensee:	Mrs. Laila Al-Soufi Novo
Order Date:	May 27, 2024
Order Number:	5797190050440
Publication:	Addiction Biology
Title:	A polygenic approach to the association between smoking and schizophrenia
Type of Use:	Dissertation/Thesis
Order Ref:	1
Order Total:	0.00 USD

View or print complete details of your order and the publisher's terms and conditions.

Sincerely,

Copyright Clearance Center

Estudo 5:



USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

O consumo de substancias representa un grave problema de saúde global. A susceptibilidade xenética aos trazos de consumo de substancias, altamente polixénicos, é estudada principalmente mediante estudos de asociación do xenoma completo (GWAS). Nos cinco estudos desta tese, emprégase unha batería de ferramentas xenómicas baseadas en GWAS para investigar os factores xenéticos implicados na susceptibilidade ao consumo de substancias. Ademais, analízase a posible causa da alta prevalencia do tabaquismo en persoas con esquizofrenia, atopando evidencia dun efecto causal do consumo de tabaco sobre a esquizofrenia. Esta información podería ter implicacións a nivel clínico, destacando a necesidade de desenvolver estratexias eficaces para prever e tratar o consumo de substancias, especialmente nas persoas con esquizofrenia.