

## ARTICLE OPEN



# Polygenic risk scores for late smoking initiation associated with the risk of schizophrenia

Kazutaka Ohi<sup>1,2</sup>, Daisuke Nishizawa<sup>3</sup>, Yukimasa Muto<sup>1</sup>, Shunsuke Sugiyama<sup>1</sup>, Junko Hasegawa<sup>3</sup>, Midori Soda<sup>4</sup>, Kiyoyuki Kitaichi<sup>4</sup>, Ryota Hashimoto<sup>5,6</sup>, Toshiki Shioiri<sup>1</sup> and Kazutaka Ikeda<sup>3</sup>

Patients with schizophrenia display characteristic smoking-related behaviors and genetic correlations between smoking behaviors and schizophrenia have been identified in European individuals. However, the genetic etiology of the association remains to be clarified. The present study investigated transethnic genetic overlaps between European-based smoking behaviors and the risk of Japanese schizophrenia by conducting polygenic risk score (PRS) analyses. Large-scale European genome-wide association study (GWAS) datasets ( $n = 24,114$ – $74,035$ ) related to four smoking-related intermediate phenotypes [(i) smoking initiation, (ii) age at smoking initiation, (iii) smoking quantity, and (iv) smoking cessation] were utilized as discovery samples. PRSs derived from these discovery GWASs were calculated for 332 Japanese subjects [schizophrenia patients, their unaffected first-degree relatives (FRs), and healthy controls (HCs)] as a target sample. Based on GWASs of European smoking phenotypes, we investigated the effects of PRSs on smoking phenotypes and the risk of schizophrenia in the Japanese population. Of the four smoking-related behaviors, the PRSs for age at smoking initiation in Europeans significantly predicted the age at smoking initiation ( $R^2 = 0.049$ ,  $p = 0.026$ ) and the PRSs for smoking cessation significantly predicted the smoking cessation ( $R^2 = 0.092$ ,  $p = 0.027$ ) in Japanese ever-smokers. Furthermore, the PRSs related to age at smoking initiation in Europeans were higher in Japanese schizophrenia patients than in the HCs and those of the FRs were intermediate between those of patients with schizophrenia and those of the HCs ( $R^2 = 0.015$ ,  $p = 0.015$ ). In our target subjects, patients with schizophrenia had a higher mean age at smoking initiation ( $p = 0.018$ ) and rate of daily smoking initiation after age 20 years ( $p = 0.023$ ) compared with the HCs. A total of 60.6% of the patients started to smoke before the onset of schizophrenia. These findings suggest that genetic factors affecting late smoking initiation are associated with the risk of schizophrenia.

npj Schizophrenia (2020)6:36; <https://doi.org/10.1038/s41537-020-00126-z>

## INTRODUCTION

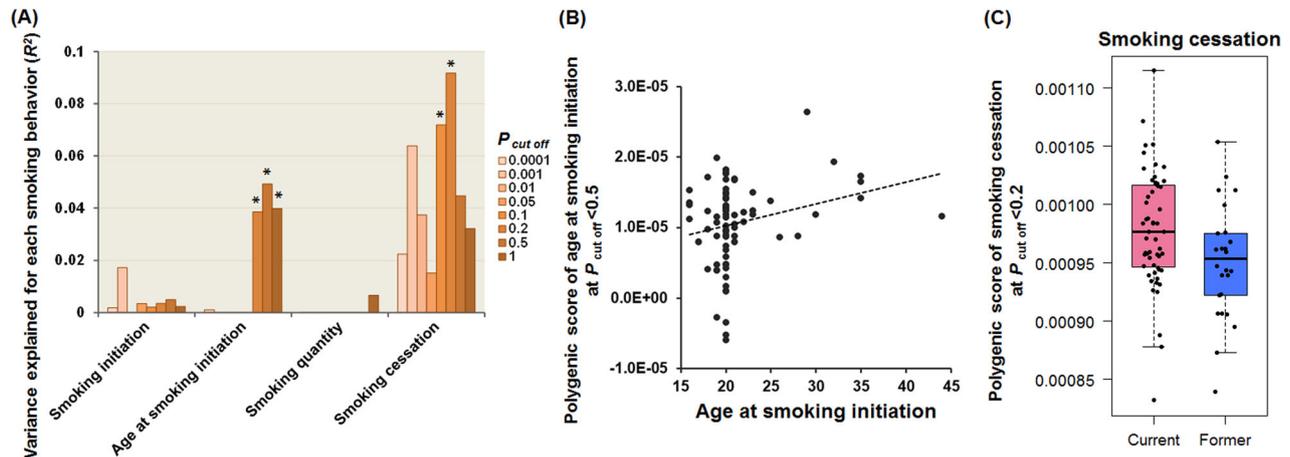
The lifespan of patients with schizophrenia is 10–20 years shorter than that of the general population<sup>1–3</sup>. Patients with schizophrenia have increased mortality and morbidity, mainly due to metabolic syndrome and premature cardiovascular disease<sup>1–3</sup>. A high rate of cigarette smoking contributes to metabolic syndrome and premature cardiovascular disease<sup>3,4</sup>. The prevalence of cigarette smoking in patients with schizophrenia is estimated to be two- to fivefold higher than in the general population worldwide<sup>5–7</sup>. Smokers with schizophrenia are generally “heavy smokers” who smoke more cigarettes and consume relatively larger total cigarette volumes<sup>8</sup>; these individuals generally experience nicotine dependence<sup>9</sup>.

The self-medication hypothesis argues that patients with schizophrenia smoke to alleviate their psychiatric symptoms, cognitive impairments, and/or antipsychotic-induced side effects<sup>10,11</sup>. In contrast, smoking itself may be a causal factor for the onset of schizophrenia<sup>12</sup>. Patients with schizophrenia display a higher risk of starting daily smoking at least 5 years before the onset of schizophrenia and starting psychiatric medication<sup>13–15</sup>, suggesting that the association between smoking and schizophrenia cannot be explained by the onset of illness or the effects of medication. Furthermore, the age at initiation of daily smoking is higher after age 20 years in

individuals who later develop schizophrenia<sup>13–15</sup>. A drastic greater exposure for a brief period despite later age at smoking initiation may be at a higher risk of schizophrenia. Therefore, patients with schizophrenia display characteristic smoking-related behaviors: a high smoking initiation rate, later age at smoking initiation, high smoking quantity, and low smoking cessation rate.

Schizophrenia and smoking-related behaviors are highly heritable, with an estimated heritability of ~50–80%<sup>16–18</sup>. The risks of the onset of schizophrenia and smoking-related behaviors are increased in unaffected first-degree relatives (FRs) of patients with schizophrenia compared with healthy controls (HCs)<sup>19–21</sup>, suggesting that the relationships between schizophrenia and smoking-related behaviors are largely mediated by shared genetic factors. The Schizophrenia Working Group of the Psychiatric Genomics Consortium (PGC) and the Tobacco and Genetics (TAG) consortium have performed large-scale genome-wide association studies (GWASs) in samples of mainly European ancestry, to find risk genes for schizophrenia and smoking-related behaviors<sup>16,22</sup>. The TAG consortium has examined four elements of smoking behavior: smoking initiation, age at smoking initiation, smoking quantity, and smoking cessation. The GWAS from the PGC has identified 108 distinct genomic loci related to the risk of schizophrenia<sup>22</sup>. In contrast, among the GWASs of smoking-related phenotypes

<sup>1</sup>Department of Psychiatry and Psychotherapy, Gifu University Graduate School of Medicine, Gifu, Japan. <sup>2</sup>Department of General Internal Medicine, Kanazawa Medical University, Ishikawa, Japan. <sup>3</sup>Addictive Substance Project, Tokyo Metropolitan Institute of Medical Science, Tokyo, Japan. <sup>4</sup>Department of Biomedical Pharmaceutics, Gifu Pharmaceutical University, Gifu, Japan. <sup>5</sup>Department of Pathology of Mental Diseases, National Institute of Mental Health, National Center of Neurology and Psychiatry, Kodaira, Tokyo, Japan. <sup>6</sup>Molecular Research Center for Children's Mental Development, United Graduate School of Child Development, Osaka University, Suita, Osaka, Japan. ✉email: k\_ohi@gifu-u.ac.jp



**Fig. 1 Effects of polygenic risk scores (PRSs) for European-based smoking-related intermediate phenotypes on smoking behaviors in a Japanese population.** Effects of PRSs for four European-based smoking-related intermediate phenotypes based on each threshold ( $P_{T \text{ cutoff}}$ ) on each smoking-related intermediate phenotype in Japanese target subjects (a). The y-axis shows the adjusted  $R^2$  or Nagelkerke's pseudo- $R^2$ , indicating the explanatory power of the model.  $*p < 0.05$ . Positive correlation between PRSs related to European-based age at smoking initiation and age at smoking initiation in the target ever-smokers (b). Differences in PRSs based on European-based smoking cessation between Japanese current and former smokers (c). Box plots are used to visually show the distribution of PRSs. The median is shown by the center line that divides the box into two parts. Lower and upper bounds of box represent first and third quartiles, respectively. The lower and upper whiskers represent scores outside the middle 50%.

from the TAG, only an association between a genomic locus on 15q25 and smoking quantity has been revealed in smoking GWAS<sup>16</sup>. We found that genome-wide significant single-nucleotide polymorphisms (SNPs) on 15q25 shared between schizophrenia and smoking quantity contributed to a common pathophysiology underlying these phenotypes via altered *CHRNA5* expression in the brain<sup>23</sup>.

On the other hand, the risks of schizophrenia and smoking-related behaviors are mediated by not only genome-wide significant SNPs but also a polygenic component comprising the additive effects of a large number of common SNPs with weak effects<sup>22,24–30</sup>. Consistent with the comorbidity of schizophrenia and cigarette smoking behaviors, genetic correlations between the risk of schizophrenia, and higher smoking initiation, later age at smoking initiation and greater smoking quantity have been indicated by linkage disequilibrium score regression (LDSC) and polygenic risk score (PRS) analyses using European-based GWASs from the PGC and TAG<sup>27,28,31</sup>. However, these genetic correlations are restricted to findings derived from the same European-based GWAS datasets. To the best of our knowledge, it is unknown whether PRSs for European-based smoking behaviors are associated with the risk of schizophrenia in independent individuals of non-European ancestry.

As cigarette smoking causes severe health impairments in patients with schizophrenia, understanding the genetic basis underlying the comorbidity is clinically important. We hypothesized that genetic variants related to smoking behaviors in Europeans would transethnically overlap with genetic risk variants in Japanese schizophrenia patients and unaffected FRs. PRS analyses can examine whether the PRSs related to discovery GWASs can predict the risk of the phenotype in an independent target GWAS sample. The present study investigated the effects of PRSs based on GWASs of four European-based smoking-related intermediate phenotypes [(i) smoking initiation, (ii) age at smoking initiation, (iii) smoking quantity, and (iv) smoking cessation] on smoking behaviors and the risk of schizophrenia in the Japanese population by PRS analyses as well.

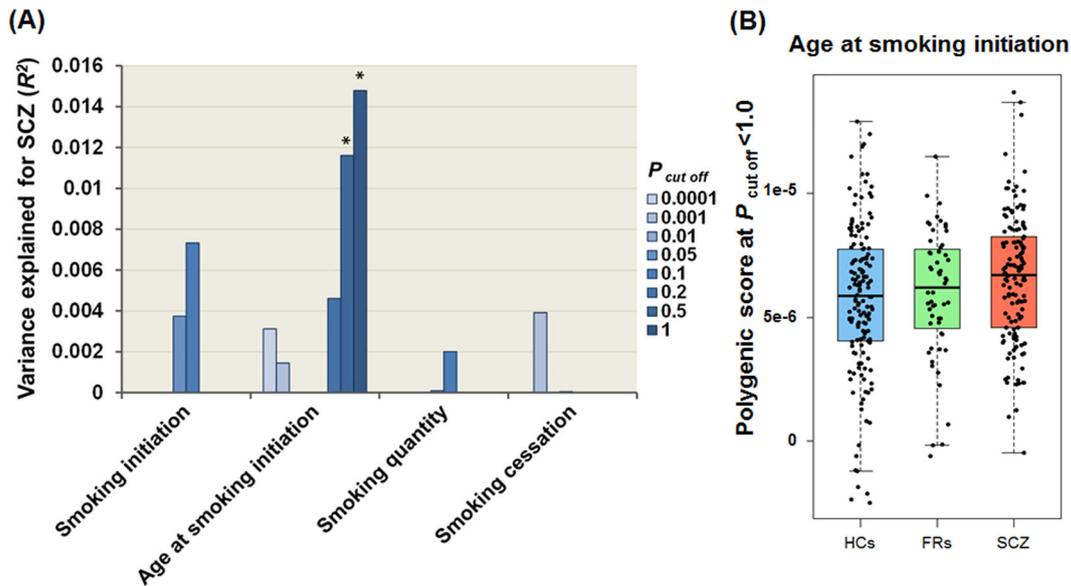
## RESULTS

Effects of PRSs for European-based smoking-related intermediate phenotypes on smoking behaviors in a Japanese population

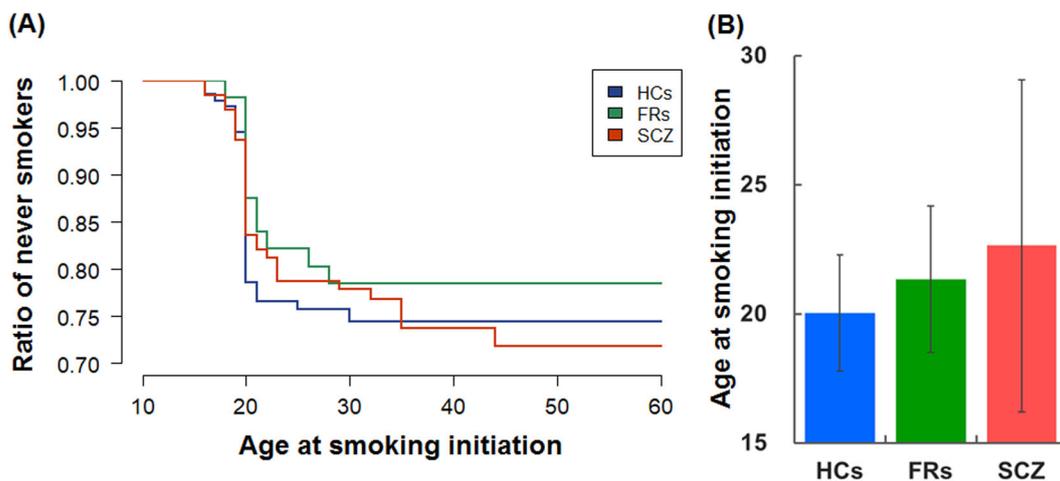
We first investigated the transethnic effect of PRSs for European-based smoking-related intermediate phenotypes (smoking initiation, age at smoking initiation, smoking quantity, and smoking cessation) on each smoking-related intermediate phenotype in Japanese target subjects at different  $P_T$  levels, respectively (Fig. 1a). Of four smoking-related intermediate phenotypes, the PRSs for European-based age at smoking initiation significantly predicted age at smoking initiation in Japanese ever-smokers ( $n = 81$ ) (Fig. 1b, a maximum at  $P_T \leq 0.5$ : adjusted  $R^2 = 0.049$ ,  $p = 0.026$ ). In contrast, the PRSs for European-based smoking cessation were significantly and unexpectedly higher in Japanese current smokers ( $n = 55$ ) than in former smokers ( $n = 26$ ) (Fig. 1c, a maximum at  $P_T \leq 0.2$ : Nagelkerke's  $R^2 = 0.092$ ,  $p = 0.027$ ). Even after including diagnostic status, age, and sex as covariates, these associations were still significant (age at smoking initiation,  $p = 0.043$ ; smoking cessation,  $p = 0.013$ ). There were no significant associations between the PRSs for European-based smoking initiation or smoking quantity and these smoking behaviors in a Japanese population, respectively (Fig. 1a,  $p > 0.05$ ).

Effects of PRSs for European-based smoking-related intermediate phenotypes on the risk of schizophrenia in a Japanese population

To reveal genetic correlations between European-based smoking phenotypes and the risk of schizophrenia in a Japanese population, we investigated the effects of PRSs based on European-based smoking GWASs on risk levels in Japanese schizophrenia patients (schizophrenia patients vs. FRs vs. HCs) at different  $P_T \text{ cutoff}$  levels (Fig. 2). The PRSs obtained from European-based GWAS for the age at smoking initiation were significantly different among Japanese HCs, FRs, and patients with schizophrenia (Fig. 2a, a maximum at  $P_T \leq 1.0$ : adjusted  $R^2 = 0.015$ ,  $p = 0.015$ ). The PRSs related to age at smoking initiation in Europeans were higher in Japanese patients with schizophrenia than in HCs and those in FRs were intermediate between those in



**Fig. 2** Effects of PRSs for European-based smoking-related intermediate phenotypes on the risk of SCZ in a Japanese population. Effects of PRSs for four European-based smoking-related intermediate phenotypes on the risk of SCZ (HCs, FRs, and patients with SCZ) (a). \* $p < 0.05$ . Differences in PRSs based on European-based age at smoking initiation among HCs, FRs, and patients with SCZ (b). HC healthy control; FR first-degree relative; SCZ schizophrenia.



**Fig. 3** Differences in age at smoking initiation among Japanese HCs, FRs, and patients with SCZ. A survival analysis of onset age for daily smoking among Japanese HCs, FRs, and patients with SCZ (a). Differences in the mean ages at smoking initiation among the diagnostic groups in ever-smokers (b).

patients with schizophrenia and those in HCs (Fig. 2b). In contrast, there were no significant differences in the PRSs obtained from European-based GWAS for the other three smoking-related intermediate phenotypes (smoking initiation, smoking quantity, or smoking cessation) among the diagnostic groups (Fig. 2a,  $p > 0.05$ ). Even after excluding FRs, PRSs related to the European age at smoking initiation were significantly higher in Japanese patients with schizophrenia than in HCs (Supplementary Fig. 1, a maximum at  $P_T \leq 1.0$ : Nagelkerke's  $R^2 = 0.028$ ,  $p = 0.017$ ). The directions of these associations between patients with schizophrenia vs. FRs vs. HCs and patients with schizophrenia vs. HCs were identical. There was no significant interaction between the PRSs for age at smoking initiation and ever/never smoking status with regard to the risk of schizophrenia in a Japanese population (Supplementary Fig. 2,  $p = 0.88$ ).

Differences in age at smoking initiation among Japanese HCs, FRs, and patients with schizophrenia

We performed a survival analysis of the age at onset for daily smoking (conversion into ever-smoker from never-smoker) among Japanese HCs, FRs, and patients with schizophrenia (Fig. 3a), but the survival curves of the age at onset for daily smoking were not significantly different among the diagnostic groups ( $p = 0.70$ ). In contrast, the mean ages at smoking initiation were significantly different among the diagnostic groups in ever-smokers (Fig. 3b,  $\beta = 0.26$ ,  $p = 0.018$ ). Even after including age and sex as covariates, differences in the mean ages at smoking initiation were still significant ( $\beta = 0.29$ ,  $p = 0.013$ ). Patients with schizophrenia showed a higher age at smoking initiation than the other groups. As schizophrenia is associated with a higher rate of smoking initiation after age 20

years (>20) in multiple ethnic populations<sup>13–15</sup>, we investigated the rate of smoking initiation after age 20 years among HCs, FRs, and patients with schizophrenia. The rates of smoking initiation after age 20 years, with age and sex as covariates, were significantly different among the diagnostic groups ( $p = 0.023$ ). The rates of smoking initiation after age 20 years were 13.9% (5/36) in HCs, 41.7% (5/12) in FRs, and 36.4% (12/33) in patients with schizophrenia. Many smokers who were HCs started to smoke at age 20 years, whereas smokers who were FRs and patients with schizophrenia gradually started to smoke after age 20 years. Furthermore, after age 20 years, smoking initiation rates have been higher in overall schizophrenia patients and in schizophrenia patients who started daily smoking at least 5 years before the onset of the disorder compared with HCs or general population<sup>13–15</sup>. Histograms of the age at smoking initiation, age at onset of schizophrenia, and the age at smoking initiation minus the age at the onset of schizophrenia in our target patients are shown in Supplementary Fig. 3. Of 33 schizophrenia patients who had ever smoked, 60.6% started to smoke before the onset of the disorder, 3.0% started to smoke at the onset of the disorder, and 36.4% started to smoke after the onset of the illness.

## DISCUSSION

The present study examined for the first time whether European-based PRSs for smoking-related intermediate phenotypes (smoking initiation, age at smoking initiation, smoking quantity, and smoking cessation) based on large-scale GWASs from the TAG consortium transethnically affect smoking behaviors and the risk of schizophrenia in a Japanese population by PRS analyses. Of the four smoking-related behaviors, the PRSs for European-based age at smoking initiation and smoking cessation significantly predicted age at smoking initiation and smoking cessation, respectively, in Japanese ever-smokers. Furthermore, the PRSs related to European-based age at smoking initiation were linearly associated with the risk of schizophrenia in the Japanese population. Among Japanese ever-smokers, patients with schizophrenia had a delayed mean age at smoking initiation, especially after age 20 years, but most patients started to smoke before the onset of the disorder.

Consistent with the positive genetic correlation between the risk of schizophrenia and age at smoking initiation in individuals of European ancestry according to LDSC analysis<sup>28</sup>, the PRSs related to age at smoking initiation in Europeans predicted the risk of schizophrenia in non-European subjects regardless of the differences in sample ethnicity and analytical approach (LDSC and PRS analyses). Considering that earlier age at smoking initiation contributes to a higher risk of nicotine dependence in the general population<sup>32</sup>, it is expected that genetic factors for age at smoking initiation would be earlier (lower) in patients with schizophrenia than in HCs. However, age at smoking initiation in patients with schizophrenia was genetically and clinically later compared with that in HCs. The PRSs related to age at smoking initiation explained 1.5% of the variance in schizophrenia status. Furthermore, even after including confounding factors, such as age, sex, and education years, as covariates, the PRSs related to age at smoking initiation were strongly associated with the diagnostic status ( $p = 5.98 \times 10^{-3}$ ). We suggest that smoking behaviors, i.e., later age at smoking initiation but initiation before age at onset of the disorder, genetically contribute to the risk of developing schizophrenia.

As unaffected FRs share approximately half of the genetic risk of schizophrenia, PRSs for schizophrenia in unaffected FRs are intermediate between those of schizophrenia patients and HCs<sup>29,33,34</sup>. In addition to this evidence, we found that the PRSs for age at smoking initiation were higher in schizophrenia patients than in HCs, and those in FRs were intermediate between those in patients with schizophrenia and those in HCs. These findings

support the concept that there are shared genetic components between the risk of schizophrenia and age at smoking initiation. In contrast, it may also be a causal effect in either direction. To explore the causal effect, further study using Mendelian randomization is warranted.

Despite positive genetic correlations between the risk of schizophrenia and smoking initiation and smoking quantity in individuals of European ancestry according to LDSC analysis<sup>27,28,31</sup>, we could not find genetic overlaps in the Japanese population. This might be the result of socioenvironmental changes surrounding smoking in Japan that occurred prior to ~2010. Governmental health insurance has approved smoking cessation treatment in Japan since 2006 and the price per pack of cigarettes increased >30% in 2010. Furthermore, cigarette smoking was gradually restricted in public places and in psychiatric hospitals in Japan. These events may have affected the decline in the rate of ever-smokers and smoking quantity since 2010.

The PRSs for European-based smoking initiation and heaviness did not significantly predict smoking initiation or heaviness in Japanese population. The PRSs for European-based smoking cessation were unexpectedly higher in Japanese current smokers than in former smokers. The reasons why we could not detect significant transethnic effects of PRSs for European-based smoking-related intermediate phenotypes on Japanese smoking behaviors might be a power issue of our samples. In contrast, the initiation PRSs were much more associated with risk taking/impulsivity and make more or less risky initiation, and might be social differences in Japan. Therefore, we might not have predicted the association by the European score. Given that these European-based PRSs did not predict Japanese smoking behaviors, it was not surprising that they did not predict risk of schizophrenia.

Cigarettes are very refined tools to deliver nicotine to the brain. Cigarette smoking modulates dopaminergic activity in the brain through inducing the release and inhibiting the degradation<sup>35</sup>. Furthermore, cigarette smoking can reduce impairments associated with dopamine hypofunction in the prefrontal cortex<sup>35</sup>. Therefore, patients with schizophrenia may practice self-medication behavior to alleviate their psychiatric symptoms, cognitive impairments, and/or antipsychotic-induced side effects<sup>10,11</sup>. Despite our evidence that patients with schizophrenia initiate smoking at a later age, patients may develop nicotine dependence through self-medicating behavior and as a result of genetic factors affecting their susceptibility to nicotine dependence.

There are some limitations of the interpretations of our findings. Recently, the GWAS and Sequencing Consortium of Alcohol and Nicotine use has performed larger-scale GWAS for smoking-related behaviors in up to 1.2 million individuals<sup>36</sup>. A critical factor in determining if the polygenic components can explain a target trait in independent participants is the sample size of the discovery GWAS<sup>37,38</sup>. Therefore, further study using the latest and largest-scale GWAS is required to confirm our findings. Our target sample size might have been insufficient to identify possible associations between the risk of schizophrenia and PRSs related to the other three smoking-related behaviors. Furthermore, the sample size of Japanese ever-smokers in our target cohort was relatively small. Therefore, negative findings should be interpreted with caution, as our sample size was small for PRS analysis. Further investigation using a larger sample is warranted. Compared with the mean ages of the FRs and patients with schizophrenia, the mean age of the HCs was younger ( $37.2 \pm 14.1$  years), potentially including future ever-smokers as never smokers. However, the possibility that HCs have a risk of smoking initiation may be low, because the mean age at smoking initiation in HCs was  $20.0 \pm 2.2$  years.

In conclusion, the common polygenic factors for European-based age at smoking initiation, which is associated with age at smoking initiation in Japanese ever-smokers, could transethnically explain susceptibility to schizophrenia in a Japanese population.

**Table 1.** Sample sizes for European-based discovery GWASs of four smoking-related intermediate phenotypes.

Smoking phenotype	Description	Authors (year)	GWS locus	Sample sizes ( <i>n</i> )
Smoking initiation	Ever vs. never smokers	Furberg et al. <sup>16</sup>	0	74,035
Age at smoking initiation	Age at started smoking cigarettes	Furberg et al. <sup>16</sup>	0	24,114
Smoking quantity	Number of cigarettes smoked per day	Furberg et al. <sup>16</sup>	1	38,181
Smoking cessation	Former vs. current smokers	Furberg et al. <sup>16</sup>	0	41,278

GWS genome-wide significant.

These findings suggest that there are common transethnic genetic factors for the risk of schizophrenia and delayed age at smoking initiation between individuals of European and non-European ancestry. As smoking behavior is a protectable and treatable cause of morbidity and mortality in schizophrenia, as well as the onset of other illnesses, understanding the genetic etiology underlying smoking behavior is important.

## METHODS

### Discovery European samples

Several publicly available smoking-related GWAS datasets [(i) smoking initiation, (ii) age at smoking initiation, (iii) smoking quantity, and (iv) smoking cessation (<https://www.med.unc.edu/pgc/results-and-downloads>)] from individuals of European ancestry from the TAG Consortium<sup>16</sup> were used as discovery samples to identify variants related to each intermediate smoking phenotype and their *p*-values and effect sizes ( $\beta$ ).

The four smoking-related intermediate phenotypes were defined as follows<sup>16</sup>: (i) smoking initiation was assessed as ever vs. never regular smokers. Regular smokers and never regular smokers were defined as those who reported having smoked  $\geq 100$  cigarettes and those who reported having smoked between 0 and 99 cigarettes during their lifetime, respectively. (ii) Age at smoking initiation in ever-smokers was assessed as the age at which the individual first tried smoking cigarettes or began smoking cigarettes regularly. The age at smoking initiation was transformed using the natural logarithm. (iii) Smoking quantity in ever-smokers was assessed as the average or maximum number of cigarettes smoked per day (CPD). (iv) Smoking cessation in ever-smokers was assessed as former vs. current smokers. Current smokers reported that they smoked at the time of the interview, whereas former smokers had quit smoking at least 1 year before the interview. Smokers who had quit smoking for <1 year at interview were excluded, because relapse is the most common in the first year of quitting smoking.

A detailed description of the sample collection, sample information, genotyping, quality control (QC), and imputation procedures applied in the discovery samples has been described previously<sup>16</sup>. Briefly, the 16 TAG studies involved genotyping on six different GWAS platforms and performed their own QC and imputation. Only subjects of European ancestry were included. Genotype imputation was performed using the HapMap-2 Utah residents with Northern and Western European ancestry (CEU) samples as a reference panel. Genotype imputation resulted in a common set of ~2.5 million genetic variants. Imputed allele dosages for each genetic variant were tested for associations with each intermediate smoking phenotype under additive logistic [(i) smoking initiation and (iv) smoking cessation] or linear [(ii) age at smoking initiation and (iii) smoking quantity] regression models. Case-control status was adjusted if the primary study was case-control in design. These analyses were performed separately for males and females. A fixed-effect meta-analysis was performed for each intermediate smoking phenotype. The sample sizes for each GWAS (Table 1) and the number of SNPs used for the PRS analyses at each liberal significance threshold ( $P_{T \text{ cutoff}}$ ) (Supplementary Table 1) are also provided.

### Target Japanese sample

The target sample ( $n = 332$ ) comprised 130 patients with schizophrenia (50 males/80 females, mean age  $\pm$  SD: 42.9  $\pm$  13.1 years), 56 of their unaffected FRs (40 parents/12 siblings/4 offspring, 18 males/38 females, 56.8  $\pm$  15.6 years), and 146 HCs (97 males/49 females, 37.2  $\pm$  14.1 years). All of these subjects participated in our previous study<sup>29</sup>. The study

sample was recruited from the Schizophrenia Non-Affected Relative Project<sup>29,39</sup>. All subjects were of Japanese descent and had no biological first- or second-degree relatives within their own diagnostic groups. A detailed description of sample recruitment and diagnosis has been provided previously<sup>7,29,39-43</sup>. Briefly, each patient was diagnosed on the basis of unstructured clinical interviews, medical records, and clinical conferences according to the criteria in the fifth edition of the Diagnostic and Statistical Manual of Mental Disorders (DSM-5). FRs and HCs were evaluated using the Structured Clinical Interview for DSM-IV-Non-Patient version. Written informed consent was obtained from all participants after the procedures had been thoroughly explained. This study was performed in accordance with the World Medical Association's Declaration of Helsinki and was approved by the Research Ethics Committees of Gifu University and Kanazawa Medical University.

The definitions of the four smoking-related intermediate phenotypes in the target sample were as follows: (i) smoking initiation was assessed as ever vs. never regular smokers. (ii) Age at smoking initiation in ever-smokers was assessed as the age at which the individual first tried smoking cigarettes. (iii) Smoking quantity in ever-smokers was assessed as the average number of CPDs. (iv) Smoking cessation in ever-smokers was assessed as former vs. current smokers. Current smokers reported that they smoked at the time of the interview, whereas former smokers had quit smoking before the interview. These definitions in the current study were not strict compared with the original GWAS<sup>16</sup>, but these definitions were similar between studies.

A detailed description of the genotyping, QC and imputation procedures applied in the target sample has been provided previously<sup>29</sup>. Briefly, venous blood was collected from the target subjects and genomic DNA was extracted from the whole blood samples. Genotyping was performed using the Infinium OmniExpressExome-8 v1.4 BeadChip (Illumina, San Diego, CA, USA). After applying QC, including the genotype call rate, genotype call frequency, sex chromosome anomalies, and sample relatedness in each diagnostic group, 332 individuals were included in the current study. The demographic information of the diagnostic groups is summarized in Table 2. Furthermore, after the exclusion of SNPs that were duplicated or ambiguous, were localized on the Y chromosome or mitochondria, deviated from Hardy-Weinberg equilibrium, or had a low minor allele frequency, 627,789 SNPs were retained. Genotype imputation was performed using the 1000 Genomes Project Phase 3 dataset [[https://mathgen.stats.ox.ac.uk/impute/1000GP\\_Phase3.html](https://mathgen.stats.ox.ac.uk/impute/1000GP_Phase3.html)]<sup>44</sup> as a reference panel. For the PRS analysis, SNPs with high imputation quality (>0.9) were retained. Finally, 8,339,066 variants were considered in the PRS analysis.

### Statistical analyses

The SNPs in the target sample were pruned based on a pairwise  $r^2$  threshold of 0.25 and a window size of 200 SNPs to remove SNPs that were in linkage disequilibrium<sup>29</sup>. After pruning, 1,354,311 independent SNPs remained. PRSs constructed from SNPs showing a nominal association with each smoking-related phenotype in the discovery GWASs according to the following eight  $P_{T \text{ cutoff}}$  values were then calculated:  $P_T \leq 0.0001$ ,  $P_T \leq 0.001$ ,  $P_T \leq 0.01$ ,  $P_T \leq 0.05$ ,  $P_T \leq 0.1$ ,  $P_T \leq 0.2$ ,  $P_T \leq 0.5$ , and  $P_T \leq 1$ . For each individual included in the target sample, a PRS was calculated by weighting the scores for the "risk SNPs" by the logarithm of the odds ratio (OR) converted from the  $\beta$  (logOR) observed in each smoking-related discovery dataset. The score, consisting of the number of risk alleles (0, 1, or 2) multiplied by the logOR, was summed over all of the SNPs in the  $P_T$ -SNP sets for each individual in the target sample. To examine the effects of the European PRSs based on each  $P_T \text{ cutoff}$  on smoking behaviors and the risk of schizophrenia in the Japanese sample population, we used R 3.6.1 (<http://www.r-project.org/>) to perform linear regression or logistic regression with diagnostic status (schizophrenia patients, FRs, and HCs) or each smoking-

**Table 2.** Demographic information for the Japanese target GWAS sample with SCZ.

	HC	FR	SCZ	<i>p</i> -values ( <i>F</i> or $\chi^2$ )
Variables	( <i>n</i> = 146)	( <i>n</i> = 56)	( <i>n</i> = 130)	
Age (years)	37.2 ± 14.1	56.8 ± 15.6	42.9 ± 13.1	2.49 × 10 <sup>-21</sup> (55.0)
Sex (male/female)	97/49	18/38	50/80	3.22 × 10 <sup>-7</sup> (30.0) <sup>a</sup>
Education (years)	16.1 ± 2.4	12.8 ± 2.1	12.6 ± 2.2	2.05 × 10 <sup>-32</sup> (91.8)
Estimated premorbid IQ	108.5 ± 7.7	99.7 ± 9.1	98.7 ± 10.4	5.38 × 10 <sup>-18</sup> (45.0)
Ever/never smoked	36/110	12/44	33/97	0.83 (0.3)
Age at smoking initiation <sup>b</sup>	20.0 ± 2.2	21.3 ± 2.8	22.6 ± 6.4	0.062 (2.9)
Cigarettes per day <sup>b</sup>	16.4 ± 10.3	19.0 ± 14.4	21.3 ± 14.3	0.29 (1.3)
Current/former smoker <sup>b</sup>	24/12	2/10	29/4	3.50 × 10 <sup>-5</sup> (20.5) <sup>a</sup>
CPZ-eq. (mg/day)	0	0	509.6 ± 512.7	–
Age at onset (years)	–	–	26.9 ± 10.6	–
DOI (years)	–	–	15.8 ± 11.3	–
PANSS-positive symptoms	–	–	16.0 ± 6.2	–
PANSS-negative symptoms	–	–	17.8 ± 6.8	–

CPZ-eq. chlorpromazine equivalent of total antipsychotics, DOI duration of illness, FR first-degree relative, HC healthy control, PANSS Positive and Negative Syndrome Scale, SCZ schizophrenia. The mean ± SD and *p*-values are shown. The significant *p*-values (*p* < 0.05) are shown in boldface and underlined. Complete demographic information was not obtained for all subjects (estimated premorbid IQ in HCs, *n* = 145). Continuous variables and categorical variables were analyzed using parametric analysis of variance (ANOVA) and Pearson's  $\chi^2$ -test, respectively.

<sup>a</sup> $\chi^2$ -test.

<sup>b</sup>Calculated among ever regular smokers (HC, *n* = 36; FR, *n* = 12; SCZ, *n* = 33).

related phenotype as a dependent variable and each PRS based on each European smoking GWAS as the independent variable. We compared the differences in the adjusted  $R^2$  for linear regression or Nagelkerke's pseudo- $R^2$  for logistic regression, which are a measure of the variance explained by the model. A survival analysis of age at onset for daily smoking among HCs, FRs and patients with schizophrenia was performed using R. The PRSs at each  $P_{T \text{ cutoff}}$  were highly correlated with each other and were not independent. Furthermore, the smoking-related intermediate phenotypes were also marginally correlated with each other and were not independent. Therefore, the *p*-values based on different  $P_{T \text{ cutoff}}$  values were not corrected and the significance level for the current study was set at *p* < 0.05.

### Reporting summary

Further information on research design is available in the Nature Research Reporting Summary linked to this article.

### DATA AVAILABILITY

The discovery data are publicly available (<https://www.med.unc.edu/pgc/results-and-downloads>). Our target data are not publicly available due to them containing information that could compromise research participant privacy/consent.

Received: 23 June 2020; Accepted: 6 October 2020;

Published online: 23 November 2020

### REFERENCES

- Capasso, R. M., Lineberry, T. W., Bostwick, J. M., Decker, P. A. & St Sauver, J. Mortality in schizophrenia and schizoaffective disorder: An Olmsted County, Minnesota cohort: 1950–2005. *Schizophr. Res.* **98**, 287–294 (2008).
- Walker, E. R., McGee, R. E. & Druss, B. G. Mortality in mental disorders and global disease burden implications: a systematic review and meta-analysis. *JAMA Psychiatry* **72**, 334–341 (2015).
- Mitchell, A. J. et al. Prevalence of metabolic syndrome and metabolic abnormalities in schizophrenia and related disorders—a systematic review and meta-analysis. *Schizophr. Bull.* **39**, 306–318 (2013).
- Vancampfort, D. et al. Relationships between physical fitness, physical activity, smoking and metabolic and mental health parameters in people with schizophrenia. *Psychiatry Res.* **207**, 25–32 (2013).
- de Leon, J. & Diaz, F. J. A meta-analysis of worldwide studies demonstrates an association between schizophrenia and tobacco smoking behaviors. *Schizophr. Res.* **76**, 135–157 (2005).
- Hartz, S. M. et al. Comorbidity of severe psychotic disorders with measures of substance use. *JAMA Psychiatry* **71**, 248–254 (2014).
- Ohi, K. et al. Smoking rates and number of cigarettes smoked per day in schizophrenia: a large cohort meta-analysis in a Japanese population. *Int. J. Neuropsychopharmacol.* **22**, 19–27 (2019).
- Tidey, J. W., Rohsenow, D. J., Kaplan, G. B. & Swift, R. M. Cigarette smoking topography in smokers with schizophrenia and matched non-psychiatric controls. *Drug Alcohol Depend.* **80**, 259–265 (2005).
- Ziedonis, D. et al. Tobacco use and cessation in psychiatric disorders: National Institute of Mental Health report. *Nicotine Tob. Res.* **10**, 1691–1715 (2008).
- Leonard, S., Mexas, S. & Freedman, R. Genetics of smoking and schizophrenia. *J. Dual Diagn.* **3**, 43–59 (2007).
- Kumari, V. & Postma, P. Nicotine use in schizophrenia: the self medication hypotheses. *Neurosci. Biobehav. Rev.* **29**, 1021–1034 (2005).
- Kendler, K. S., Lonn, S. L., Sundquist, J. & Sundquist, K. Smoking and schizophrenia in population cohorts of Swedish women and men: a prospective co-relative control study. *Am. J. Psychiatry* **172**, 1092–1100 (2015).
- Zhang, X. Y. et al. Smoking initiation and schizophrenia: a replication study in a Chinese Han population. *Schizophr. Res.* **119**, 110–114 (2010).
- Gurpegui, M. et al. Smoking initiation and schizophrenia: a replication study in a Spanish sample. *Schizophr. Res.* **76**, 113–118 (2005).
- Diaz, F. J., Velasquez, D. M., Susce, M. T. & de Leon, J. The association between schizophrenia and smoking: Unexplained by either the illness or the prodromal period. *Schizophr. Res.* **104**, 214–219 (2008).
- Furberg, H. et al. Genome-wide meta-analyses identify multiple loci associated with smoking behavior. *Nat. Genet.* **42**, 441–447 (2010).
- Sullivan, P. F. & Kendler, K. S. The genetic epidemiology of smoking. *Nicotine Tob. Res.* **1 Suppl 2**, S51–S57; discussion S69–S70 (1999).
- Sullivan, P. F., Kendler, K. S. & Neale, M. C. Schizophrenia as a complex trait: evidence from a meta-analysis of twin studies. *Arch. Gen. Psychiatry* **60**, 1187–1192 (2003).
- Tsuang, M. Schizophrenia: genes and environment. *Biol. Psychiatry* **47**, 210–220 (2000).

20. Cardno, A. G. & Gottesman, I. I. Twin studies of schizophrenia: from bow-and-arrow concordances to Star Wars Mx and functional genomics. *Am. J. Med. Genet.* **97**, 12–7 (2000).
21. Ferchiou, A. et al. Exploring the relationships between tobacco smoking and schizophrenia in first-degree relatives. *Psychiatry Res.* **200**, 674–678 (2012).
22. Ripke, S. et al. Biological insights from 108 schizophrenia-associated genetic loci. *Nature* **511**, 421–427 (2014).
23. Ohi, K. et al. Genome-wide variants shared between smoking quantity and schizophrenia on 15q25 are associated with CHRNA5 expression in the brain. *Schizophr. Bull.* **45**, 813–823 (2019).
24. Pardini, A. F. et al. Common schizophrenia alleles are enriched in mutation-intolerant genes and in regions under strong background selection. *Nat. Genet.* **50**, 381–389 (2018).
25. Purcell, S. M. et al. Common polygenic variation contributes to risk of schizophrenia and bipolar disorder. *Nature* **460**, 748–752 (2009).
26. Taniguchi, S. et al. Polygenic risk scores in schizophrenia with clinically significant copy number variants. *Psychiatry Clin. Neurosci.* **74**, 35–39 (2020).
27. Hartz, S. M. et al. Genetic correlation between smoking behaviors and schizophrenia. *Schizophr. Res.* **194**, 86–90 (2018).
28. Bulik-Sullivan, B. et al. An atlas of genetic correlations across human diseases and traits. *Nat. Genet.* **47**, 1236–1241 (2015).
29. Ohi, K. et al. Polygenic risk scores for major psychiatric disorders among schizophrenia patients, their first-degree relatives, and healthy participants. *Int. J. Neuropsychopharmacol.* **23**, 157–164 (2020).
30. Nestic, M. J., Stojkovic, B. & Maric, N. P. On the origin of schizophrenia: Testing evolutionary theories in the post-genomic era. *Psychiatry Clin. Neurosci.* **73**, 723–730 (2019).
31. Chen, J. et al. Genetic relationship between schizophrenia and nicotine dependence. *Sci. Rep.* **6**, 25671 (2016).
32. Oncel, S. Y., Dick, D. M., Maes, H. H. & Aliev, F. Risk factors influencing smoking behavior: a Turkish twin study. *Twin Res. Hum. Genet.* **17**, 563–573 (2014).
33. Ranlund, S. et al. A polygenic risk score analysis of psychosis endophenotypes across brain functional, structural, and cognitive domains. *Am. J. Med. Genet. B Neuropsychiatr. Genet.* **177**, 21–34 (2018).
34. Touloupoulou, T. et al. Polygenic risk score increases schizophrenia liability through cognition-relevant pathways. *Brain* **142**, 471–485 (2019).
35. Sagud, M. et al. Smoking and schizophrenia. *Psychiatr. Danub.* **21**, 371–375 (2009).
36. Liu, M. 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.* **51**, 237–244 (2019).
37. Dudbridge, F. Power and predictive accuracy of polygenic risk scores. *PLoS Genet.* **9**, e1003348 (2013).
38. Ohi, K. et al. Polygenic components for schizophrenia, bipolar disorder and rheumatoid arthritis predict risk of schizophrenia. *Schizophr. Res.* **175**, 226–229 (2016).
39. Ohi, K. Intelligence decline between present and premorbid IQ in schizophrenia: Schizophrenia Non-Affected Relative Project (SNARP). *Eur. Neuropsychopharmacol.* **29**, 653–661 (2019).
40. Ohi, K. et al. Meta-analysis of physical activity and effects of social function and quality of life on the physical activity in patients with schizophrenia. *Eur. Arch. Psychiatry Clin. Neurosci.* **269**, 517–527 (2019).
41. Okada, N. et al. Abnormal asymmetries in subcortical brain volume in schizophrenia. *Mol. Psychiatry* **21**, 1460–6 (2016).
42. Ohi, K. et al. Impact of familial loading on prefrontal activation in major psychiatric disorders: a near-infrared spectroscopy (NIRS) study. *Sci. Rep.* **7**, 44268 (2017).
43. Ohi, K. et al. Cognitive clustering in schizophrenia patients, their first-degree relatives and healthy subjects is associated with anterior cingulate cortex volume. *Neuroimage Clin.* **16**, 248–256 (2017).
44. Auton, A. et al. A global reference for human genetic variation. *Nature* **526**, 68–74 (2015).

## ACKNOWLEDGEMENTS

This work was supported by Grants-in-Aid for Scientific Research (C) (19K08081), Young Scientists (B) (16K19784), and KAKENHI Advanced Animal Model Support (AdAMS) (16H06276) from the Japan Society for the Promotion of Science (JSPS); a grant from the Uehara Memorial Foundation; a grant from the Takeda Science Foundation; a grant from the YOKOYAMA Foundation for Clinical Pharmacology (YRY-1807); a grant from the Smoking Research Foundation; and a Grant for Assist KAKEN (K2017-8, K2018-16 and K2018-17) and Promoted Research (S2017-3 and S2018-5) from Kanazawa Medical University. We thank all individuals who participated in this study.

## AUTHOR CONTRIBUTIONS

K.O. supervised the entire project, collected the data, wrote the manuscript, and was critically involved in the design, analysis, and interpretation of the data. K.O., Y.M., S.S., and T.S. were responsible for performing the literature review. D.N., J.H., M.S., K.K., R.H., and K.I. were involved in the genotyping. Y.M., S.S., M.S., K.K., and T.S. were heavily involved in the collection of the majority of the data and intellectually contributed to data interpretation. All authors contributed to and approved the final manuscript.

## COMPETING INTERESTS

The authors declare no competing interests.

## ADDITIONAL INFORMATION

**Supplementary information** is available for this paper at <https://doi.org/10.1038/s41537-020-00126-z>.

**Correspondence** and requests for materials should be addressed to K.O.

**Reprints and permission information** is available at <http://www.nature.com/reprints>

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



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2020