

Article

# Identification and validation of a novel gene *ARVCF* associated with alcohol dependence among Chinese population

Xiaoqiang Shi, Yan Wang, Zhongli Yang, Wenji Yuan, Ming D. Li

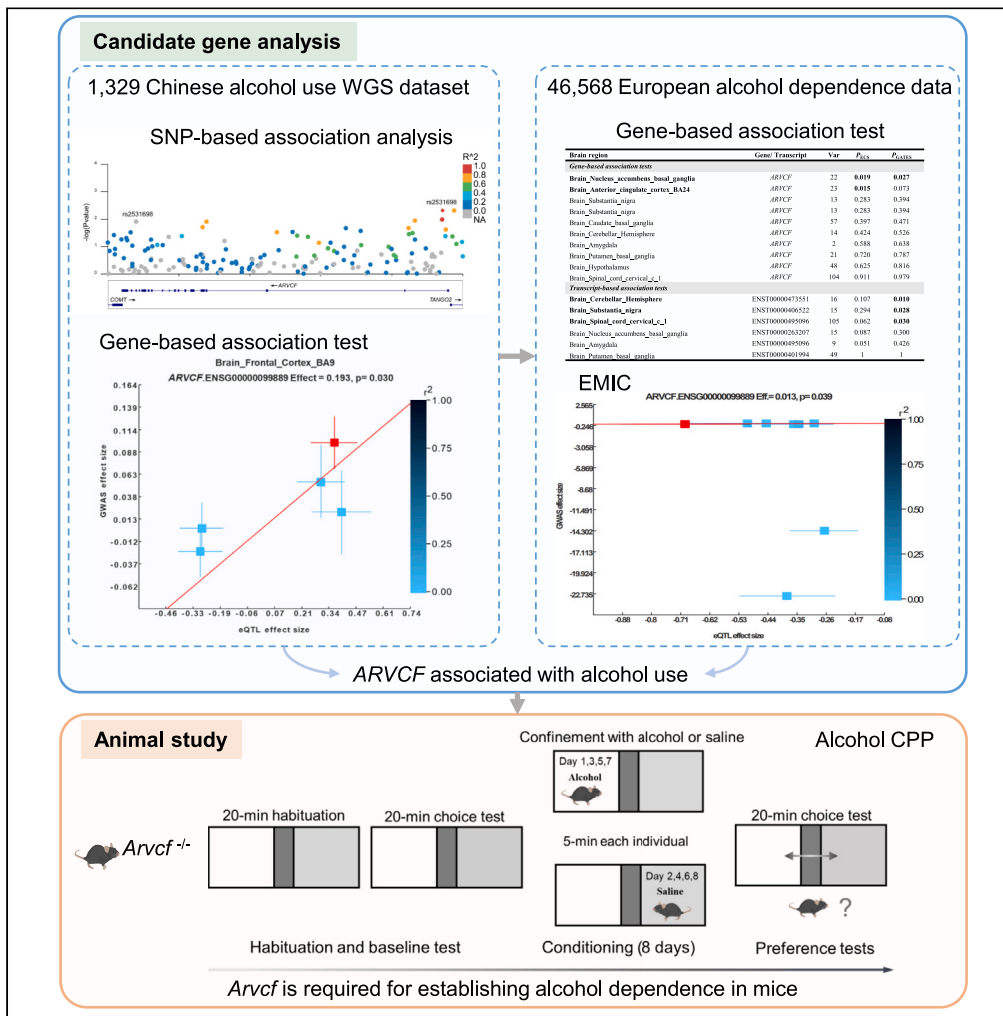
m12km@zju.edu.cn

Highlights

*ARVCF* was identified as a susceptibility gene for alcohol dependence

*ARVCF* expression in the cortex and NAc causally affects alcohol dependence

*Arvcf* is required for establishing alcohol dependence in mice



## Article

Identification and validation of a novel gene *ARVCF* associated with alcohol dependence among Chinese populationXiaoqiang Shi,<sup>1</sup> Yan Wang,<sup>1</sup> Zhongli Yang,<sup>1</sup> Wenji Yuan,<sup>1</sup> and Ming D. Li<sup>1,2,3,\*</sup>

## SUMMARY

Alcohol dependence is a heritable disorder, yet its genetic basis and underlying mechanisms remain poorly understood, especially in Chinese population. In this study, we conducted gene-based and transcript-based association tests and found a significant association between *ARVCF* expression in the cortex and hippocampus of the brain and alcohol use in a cohort of 1,329 individuals with Chinese ancestry. Further analysis using the effective-median-based Mendelian randomization framework for inferring the causal genes (EMIC) revealed a causal relationship between *ARVCF* expression in the frontal cortex and alcohol use. Moreover, leveraging extensive European alcohol dependence data, our gene association tests and EMIC analysis showed that *ARVCF* expression in the nucleus accumbens was significantly associated with alcohol dependence. Finally, animal studies indicated that *Arvcf* knockout mice lacked conditioned place preference for alcohol. Together, our combined human genetic and animal studies indicate that *ARVCF* plays a crucial role in alcohol dependence.

## INTRODUCTION

Excessive alcohol consumption is a leading risk factor for the global burden of disease, as it contributes significantly to morbidity and mortality. Moreover, alcohol use disorders (AUDs) are the most prevalent of all substance use disorders.<sup>1</sup> It was estimated that alcohol use contributed to about 3 million deaths (5.3% of all deaths) and 132.6 million (5.1%) disability-adjusted life years worldwide in 2016.<sup>2</sup> Alcohol dependence is responsible for continuing alcohol abuse-related behaviors, with approximately 12.5% of alcohol drinkers in the United States meeting the Diagnostic and Statistical Manual (DSM)-IV criteria for alcohol dependence in their lifetime.<sup>3</sup> Meta-analysis of twin and adoption studies has revealed the role of genetic and environmental factors in alcohol dependence, with an estimated heritability of 50%.<sup>4</sup> Despite the significant public health consequences, AUDs remain among the most undertreated mental disorders, partially due to the lack of effective treatment targets.<sup>5</sup>

Although a great number of genes and genetic variants are reported to be associated with alcohol use in European populations, relevant information in the Chinese population is still limited. Furthermore, only limited number of those identified susceptibility genes or variants were validated. Even alcohol is being consumed globally, but the majority studies on alcohol dependence and related clinical phenotypes such as alcohol consumption, AUD, or problematic alcohol use (PAU) are primarily focused on individuals of European ancestry and limited studies have been conducted on Asian subjects.<sup>6</sup> For example, a recent meta-analysis analyzed a PAU phenotype by combining the phenotypes of AUD and alcohol use-related problems,<sup>7</sup> which included 1,079,947 individuals, with nearly 1 million of European descent and only 13,551 of East Asian descent, indicating a substantial discrepancy in sample sizes. A similar pattern is observed for alcohol consumption phenotype. Presently, the largest genome-wide association study (GWAS) study on drinks per week included 3.4 million samples, with approximately 2.67 million of European descent and only 296,438 samples from East Asians.<sup>8</sup> Studies focused on Chinese population are even scarcer,<sup>9,10</sup> with only 533 male alcoholics in a recent alcohol dependence GWAS.<sup>10</sup> In China, alcohol consumption is recently increasing faster than in other parts of the world,<sup>11,12</sup> and this trend is projected to continue rising in the future.<sup>13</sup> Recently, a larger-scale dataset for alcohol consumption from China Kadoorie Biobank became available, encompassing 168,050 individuals, but only two East Asian genetic variants (*ALDH2*-rs671 and *ADH1B*-rs1229984) were genotyped.<sup>14</sup> Genetic architecture often varies between ancestries. According to a previous multi-ancestry study, polygenic risk scores developed in one ancestry, such as Europeans, performed poorly in other ancestral groups.<sup>8</sup> Therefore, identifying novel genes that underlie alcohol dependence in non-European ancestries is important for gaining more insights and discovering new therapeutic targets.

<sup>1</sup>State Key Laboratory for Diagnosis and Treatment of Infectious Diseases, National Clinical Research Center for Infectious Diseases, National Medical Center for Infectious Diseases, Collaborative Innovation Center for Diagnosis and Treatment of Infectious Diseases, The First Affiliated Hospital, Zhejiang University School of Medicine, Hangzhou 310009, China

<sup>2</sup>Research Center for Air Pollution and Health, Zhejiang University, Hangzhou 310058, China

<sup>3</sup>Lead contact

\*Correspondence: ml2km@zju.edu.cn

<https://doi.org/10.1016/j.isci.2024.110976>



Our previous research identified a novel functional variant, rs148582811, located in Armadillo Repeat gene deleted in Velo-Cardio-Facial syndrome (ARVCF), which regulates ARVCF expression and plays an important role in nicotine dependence.<sup>15</sup> ARVCF is among the genes implicated in 22q11.2 syndrome, a neurogenetic disorder also known as velo-cardio-facial or DiGeorge syndrome. Individuals with chromosome 22q11.2 deletions are at an elevated risk of developing psychiatric conditions.<sup>16</sup> The ARVCF protein belongs to the p120 catenin (p120ctn) subfamily, which includes four proteins encoded by four independent genes (p120ctn, encoded by *CTNND1*;  $\delta$ -catenin, encoded by *CTNND2*; p0071, encoded by *PKP4*; and ARVCF, encoded by *ARVCF*). Previous studies have reported that p120ctn,  $\delta$ -catenin, and p0071 play functional roles in dendritic and synaptic development.<sup>17,18</sup> However, relatively few studies have focused on ARVCF. A recent study utilizing a primary rat hippocampal neuronal culture model indicated that increasing ARVCF-catenin levels resulted in increased dendritic length and branching,<sup>18</sup> which was associated with alcohol addiction.<sup>19</sup> To our knowledge, the potential role of the ARVCF gene in alcohol dependence has not been reported.

Here, we first conducted a gene-based association analysis and effective-median-based Mendelian randomization framework for inferring the causal genes (EMIC) analysis to determine if ARVCF is associated with alcohol use in a Chinese genome-wide sequencing dataset, and we then replicated the findings with an independent European alcohol dependence dataset. To further validate the role of the ARVCF gene in alcohol dependence, we then conducted an alcohol-induced conditioned place preference (CPP) test in CRISPR-Cas9-generated *Arvcf* knockout (KO) mice. Taken together, we first discovered that ARVCF is a susceptible gene for alcohol use, and then we confirmed that ARVCF plays an important role in alcohol dependence in mice.

## RESULTS

### ARVCF variants associated with alcohol consumption in our Chinese cohort

Chinese participants included 834 individuals with alcohol drinking (mean age  $40 \pm 8$  years) and 470 control individuals (mean age  $42 \pm 8$  years) (Table S1). A total of 168 variants located in ARVCF (including  $\pm 2$  kb flanking regions) were analyzed. After clumping the ARVCF locus for linkage disequilibrium ( $r^2 < 0.01$ , within a 250 kb genomic window<sup>20</sup>), five single-nucleotide polymorphisms (SNPs) remained (Table S2). Our analysis indicated that the variant rs2531698 in ARVCF was significantly associated with drinks per week ( $\beta = 0.107$ , SE = 0.038,  $p = 4.79 \times 10^{-3}$ , Figure S1). Besides, another variant, rs116570619 in ARVCF, also showed an association with drinks per week ( $\beta = -0.409$ , SE = 0.163,  $p = 0.012$ , Figure S1). Notably, neither of these two variants has been reported in previous studies.

### Gene association analysis links ARVCF with alcohol use

We performed gene-based association analyses using the SNP-based association analysis results located in ARVCF from our Chinese samples and combined the Genotype-Tissue Expression (GTEx) brain expression quantitative trait loci (eQTL) data. The gene-based analysis revealed that ARVCF expression in the cervical spinal cord ( $p_{\text{GATES}} = 2.93 \times 10^{-4}$ ,  $p_{\text{ECS}} = 0.013$ ), frontal cortex ( $p_{\text{GATES}} = 0.004$ ,  $p_{\text{ECS}} = 0.012$ ), cortex ( $p_{\text{GATES}} = 0.007$ ), and cerebellum ( $p_{\text{GATES}} = 0.029$ ) was significantly associated with alcohol use (Tables 1 and S3). Additionally, transcript-based association analysis further implicated ARVCF expression in alcohol use, particularly in the hippocampus ( $p_{\text{GATES}} = 4.07 \times 10^{-4}$ ,  $p_{\text{ECS}} = 0.022$ ), cervical spinal cord ( $p_{\text{GATES}} = 6.40 \times 10^{-4}$ ,  $p_{\text{ECS}} = 0.002$ ), cortex ( $p_{\text{GATES}} = 0.022$ ), and frontal cortex ( $p_{\text{GATES}} = 0.044$ ) (Tables 1 and S4).

### ARVCF causally associated with alcohol use

Integrative analysis using EMIC, combining GWAS and brain eQTL data from GTEx, suggested a causal relationship between ARVCF expression and alcohol use. Our gene-level EMIC analysis revealed a significant causal association of ARVCF expression in the frontal cortex with alcohol use ( $\beta = 0.193$ , SE = 0.089,  $p = 0.030$ ) (Figure 1A and Table S5). The transcript-level EMIC analysis also indicated potential causal effects of ARVCF transcript expression in the cerebellum (Figure 1B), cervical spinal cord (Figure 1C), and amygdala (Figure 1D) on alcohol use ( $p < 0.05$ ) (Table S6).

### ARVCF associated with alcohol dependence in European populations

Replication analysis using data on alcohol dependence from individuals of European ancestry ( $N = 46,568$ ) further supported our findings. Our gene-based association tests showed significant correlations of ARVCF expression in the nucleus accumbens (NAc) basal ganglia ( $p_{\text{ECS}} = 0.019$ ,  $p_{\text{GATES}} = 0.027$ ) and anterior cingulate cortex ( $p_{\text{ECS}} = 0.015$ ) with alcohol dependence (Tables 2 and S7). Besides, transcript-based analysis further associated ARVCF transcript expression in the cerebellar hemisphere ( $p_{\text{GATES}} = 0.010$ ), substantia nigra ( $p_{\text{GATES}} = 0.028$ ), and cervical spinal cord ( $p_{\text{GATES}} = 0.030$ ) with alcohol dependence (Tables 2 and S8). We also performed EMIC analysis to infer whether ARVCF expression in the brain causally associated with alcohol dependence. Our gene-level EMIC analysis indicated a causal relationship between ARVCF expression in the NAc and alcohol dependence ( $\beta = 0.013$ , SE = 0.006,  $p = 0.039$ ) (Table 3). Transcript-level EMIC analysis showed that ARVCF transcript expression in the amygdala ( $p < 0.05$ ) may also be implicated in alcohol dependence (Table 3). Because most SNPs identified in Chinese population were absent in the European alcohol dependence dataset (Table S9), we primarily focused on gene-level association analysis in the European data. Additionally, we examined the association of ARVCF with other related phenotypes using phenome-wide association studies,<sup>21,22</sup> which indicated that ARVCF is associated with many psychiatric conditions, including smoking and alcohol use (Table S10).

**Table 1. Results of gene-based and transcript-based ARVCF association tests in Chinese samples**

Phenotype	Brain region	Gene/transcript	Var	$p_{ECS}$	$p_{GATES}$
<b>Gene-based association tests</b>					
Drinks per week	Brain_Spinal_cord_cervical_c_1	ARVCF	45	0.013	2.93E-04
Alcohol drink	Brain_Frontal_Cortex_BA9	ARVCF	5	0.012	0.004
Alcohol drink	Brain_Cortex	ARVCF	31	0.082	0.007
Alcohol drink	Brain_Cerebellum	ARVCF	40	0.105	0.029
<b>Transcript-based association tests</b>					
Drinks per week	Brain_Hippocampus	ENST00000406522	13	0.022	4.07E-04
Drinks per week	Brain_Spinal_cord_cervical_c_1	ENST00000487793	25	0.002	6.40E-04
Alcohol drink	Brain_Cortex	ENST00000473551	8	0.055	0.022
Alcohol drink	Brain_Frontal_Cortex_BA9	ENST00000495096	41	0.054	0.044

Var, number of variants within the gene;  $p_{ECS}$ ,  $p$  value of ECS (effective chi-squared statistics);  $p_{GATES}$ ,  $p$  value of GATES (gene-based association test using extended Simes).

### Arvcf is required for alcohol-induced CPP in mice

A major driver of alcohol dependence is drug-seeing behavior. To further examine the potential behavioral effect of *Arvcf* in alcohol dependence, we performed an ethanol-induced CPP test. In the CPP paradigm, mice were intraperitoneally injected with either saline or ethanol and placed in distinct spatial chambers distinguished by environmental cues. After conditioning, mice were given unrestricted access to the drug-paired and saline-paired chambers, and CPP was measured by calculating the difference in time spent between the conditioned chamber and the unconditioned chamber (Figure 2A). We also compared the time that mice spent in the unconditioned (exposure to saline) and conditioned (exposure to alcohol) chambers to assess the conditioned reinforcing effect of alcohol.

After conditioning, wild-type (WT) mice showed a significant preference for the chamber that had been paired with alcohol ( $p < 0.001$ ) (Figure 2B). In contrast, *Arvcf*-KO mice displayed no such preference, spending similar amounts of time in the alcohol-paired chamber and the unconditioned side (Figure 2B). Moreover, WT mice spent significantly more time in the conditioned chamber compared to *Arvcf*-KO mice ( $p < 0.01$ , Figure 2B), with no significant difference in the unconditioned chamber. In addition, a significant CPP score was observed in both male ( $p < 0.01$ ) and female ( $p < 0.001$ ) WT mice, as well as in combined sexes ( $p < 0.0001$ , Figure 2C), whereas *Arvcf*-KO mice did not exhibit a significant CPP response (Figure 2D). Altogether, these results suggest that *Arvcf* is required for the establishment of alcohol dependence in mice.

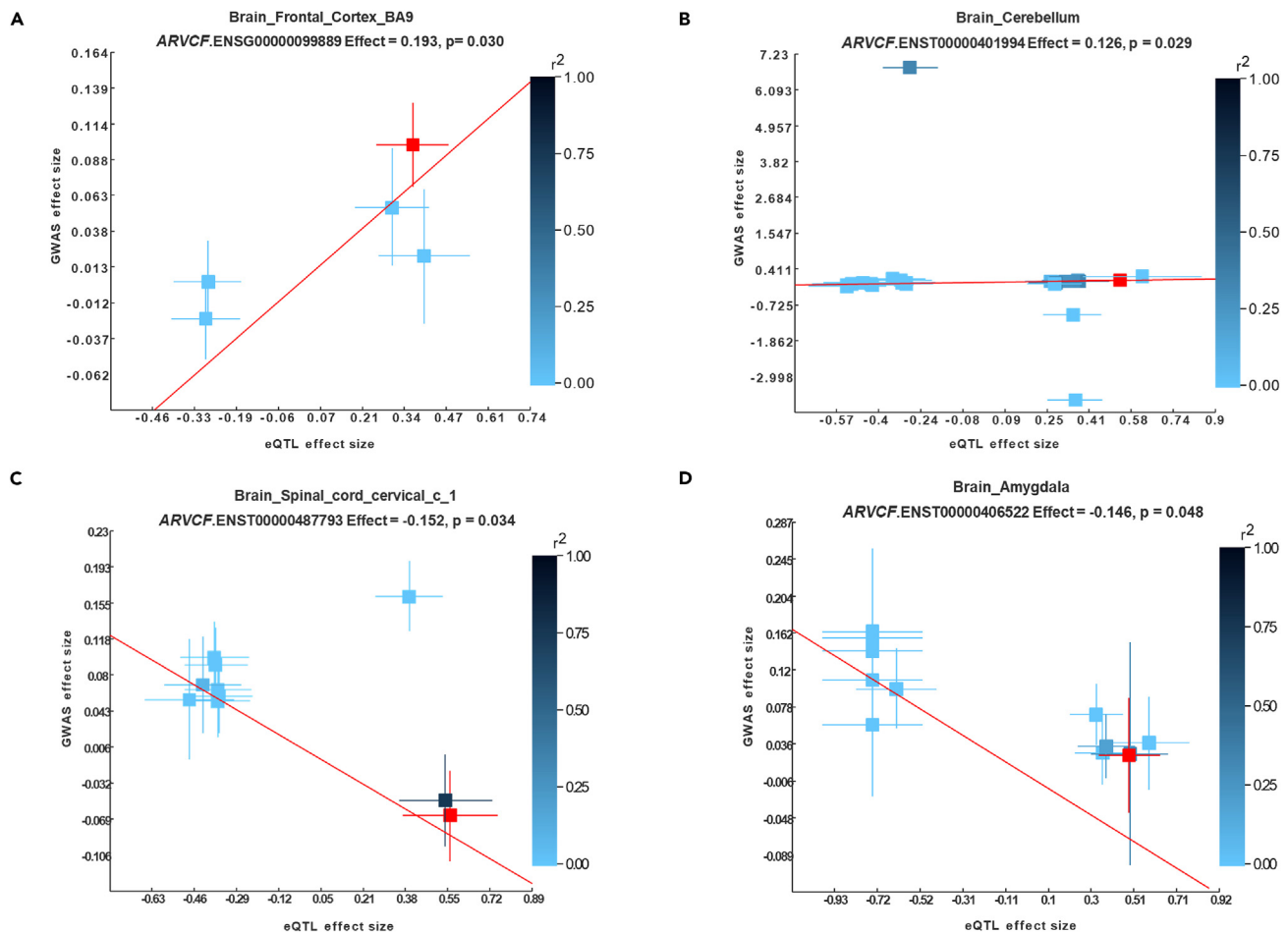
## DISCUSSION

In this study, we first revealed that the gene *ARVCF* confers susceptibility to alcohol use in Chinese samples. We further validated the causal relationship between *ARVCF* and alcohol dependence in a large European dataset. More importantly, through the utilization of alcohol-induced CPP tests with CRISPR-Cas9-generated *Arvcf*-KO mice, we demonstrated that *Arvcf* is required for the establishment of alcohol dependence in mice. The central conclusion of this study underscores the essential role of *ARVCF*, a member of the p120ctn protein family, in regulating behaviors related to alcohol dependence.

The p120ctn family of proteins, to which the *ARVCF* belongs, is implicated in the development of dendrites and synapses, both of which play crucial roles in learning, memory, addiction, and neurogenetic disorders.<sup>17,18</sup> Building on our prior discovery of a novel nicotine dependence-associated variant in *ARVCF*,<sup>15</sup> we explored whether *ARVCF* could potentially influence alcohol-related behaviors. In the Chinese population of whole-genome sequencing samples, SNP-based association analysis showed that two variants located in *ARVCF* were significantly associated with drinks per week. One of the variants, rs2531698, is an intronic variant within *ARVCF*, located in 2 kb upstream of the *TANGO2* gene, which is associated with neurodegeneration and intellectual disability.<sup>23</sup> The other variant, rs116570619, is a synonymous variant. These two variants have not been previously reported in the literature and may serve as potential targets for investigating alcohol use.

Through candidate gene association analysis utilizing Chinese samples, we discovered that *ARVCF* expression in the brain regions known to be affected by alcohol, such as the cortex and hippocampus,<sup>24</sup> was significantly associated with alcohol use. To further investigate the causal relationship between *ARVCF* expression in the brain and alcohol use, we employed the EMIC method, which is based on the Mendelian randomization framework. Notably, our EMIC analysis consistently reinforced the findings of the association test, indicating a causal relationship between *ARVCF* expression in the frontal cortex and alcohol use.

Our replication analysis, utilizing the dataset on alcohol dependence from individuals of European ancestry, further validated the causal association between *ARVCF* and alcohol use. *ARVCF* was significantly associated with alcohol dependence in both gene-based association test and EMIC analysis. Our findings highlight the significant role of *ARVCF* in alcohol dependence, particularly in its expression within the NAc, a region central to dopamine release and reward-seeking behavior. Alcohol activates the mesocorticolimbic brain reward circuit, which includes dopaminergic projections from the ventral tegmental area in the midbrain to several forebrain structures including the NAc and cortex.<sup>24,25</sup> These circuits are involved in the rewarding effects of alcohol. Both our discovery analysis using Chinese samples and our replication



**Figure 1. ARVCF expression is causally associated with alcohol use in Chinese samples**

(A) EMIC (effective-median-based Mendelian randomization framework for inferring the causal genes of complex phenotypes) infers the causal effect of ARVCF expression in the frontal cortex on alcohol use.

(B–D) EMIC analysis infers the causal effect of ARVCF transcript expression in cerebellum (B), cervical spinal cord (C), and amygdala (D) on alcohol use. Each rectangle in the plots represents a single-nucleotide polymorphism (SNP). The red rectangle denotes the most significant GWAS variant in the gene. The slope of the red line represents the estimated causal effects determined by EMIC. The varying shades of blue indicate the degree of linkage disequilibrium between the SNP and the most significant variant in association analysis. The error bars depict the standard errors of the coefficient estimates.

analysis using a European dataset consistently indicated that ARVCF affects alcohol use through brain regions within these circuits. These genetic data-based analyses collectively establish that ARVCF plays an important role in alcohol dependence.

Our animal experiments provide compelling evidence for the important role of Arvcf in alcohol dependence, further supporting the findings from human genetic studies. In the CPP test, mice lacking Arvcf displayed an absence of alcohol-evoked place preference, and the reinforcing effect of alcohol in Arvcf-KO mice was depleted, further confirming the critical role of ARVCF in alcohol dependence. These genetic and animal studies together suggested that the effect of ARVCF on alcohol dependence is likely attributable to its interaction with the brain's reward system, particularly the mesocorticolimbic brain reward circuit, which underlies the rewarding effects of alcohol.<sup>24</sup> Future studies that focus on unraveling the underlying mechanisms of how ARVCF controls alcohol dependence could be valuable.

In conclusion, our study identifies ARVCF as a susceptibility gene for alcohol dependence. ARVCF and its polymorphic sites, such as rs2531698 and rs116570619 identified in this study, may play an important role in alcohol-related disorders and could be valuable for diagnostic and preventative measures of related disorders. Future studies investigating the role of ARVCF in the mesocorticolimbic brain reward circuit and its involvement in alcohol dependence may yield further valuable insights. ARVCF emerges as a promising drug target for the treatment of alcohol and drug addiction.

### Limitations of the study

This study has several limitations. Firstly, while our dataset is substantial within the context of Chinese alcohol use studies, the sample size is still relatively small compared to European datasets. Future studies with larger Chinese populations and deep sequencing approaches may

**Table 2. Results of gene-based and transcript-based ARVCF association tests for alcohol dependence in European samples**

Brain region	Gene/transcript	Var	$p_{ECS}$	$p_{GATES}$
<b>Gene-based association tests</b>				
Brain_Nucleus_accumbens_basal_ganglia	ARVCF	22	0.019	0.027
Brain_Anterior_cingulate_cortex_BA24	ARVCF	23	0.015	0.073
Brain_Substantia_nigra	ARVCF	13	0.283	0.394
Brain_Substantia_nigra	ARVCF	13	0.283	0.394
Brain_Caudate_basal_ganglia	ARVCF	57	0.397	0.471
Brain_Cerebellar_Hemisphere	ARVCF	14	0.424	0.526
Brain_Amygdala	ARVCF	2	0.588	0.638
Brain_Putamen_basal_ganglia	ARVCF	21	0.720	0.787
Brain_Hypothalamus	ARVCF	48	0.625	0.816
Brain_Spinal_cord_cervical_c_1	ARVCF	104	0.911	0.979
<b>Transcript-based association tests</b>				
Brain_Cerebellar_Hemisphere	ENST00000473551	16	0.107	0.010
Brain_Substantia_nigra	ENST00000406522	15	0.294	0.028
Brain_Spinal_cord_cervical_c_1	ENST00000495096	105	0.062	0.030
Brain_Nucleus_accumbens_basal_ganglia	ENST00000263207	15	0.087	0.300
Brain_Amygdala	ENST00000495096	9	0.051	0.426
Brain_Putamen_basal_ganglia	ENST00000401994	49	1	1

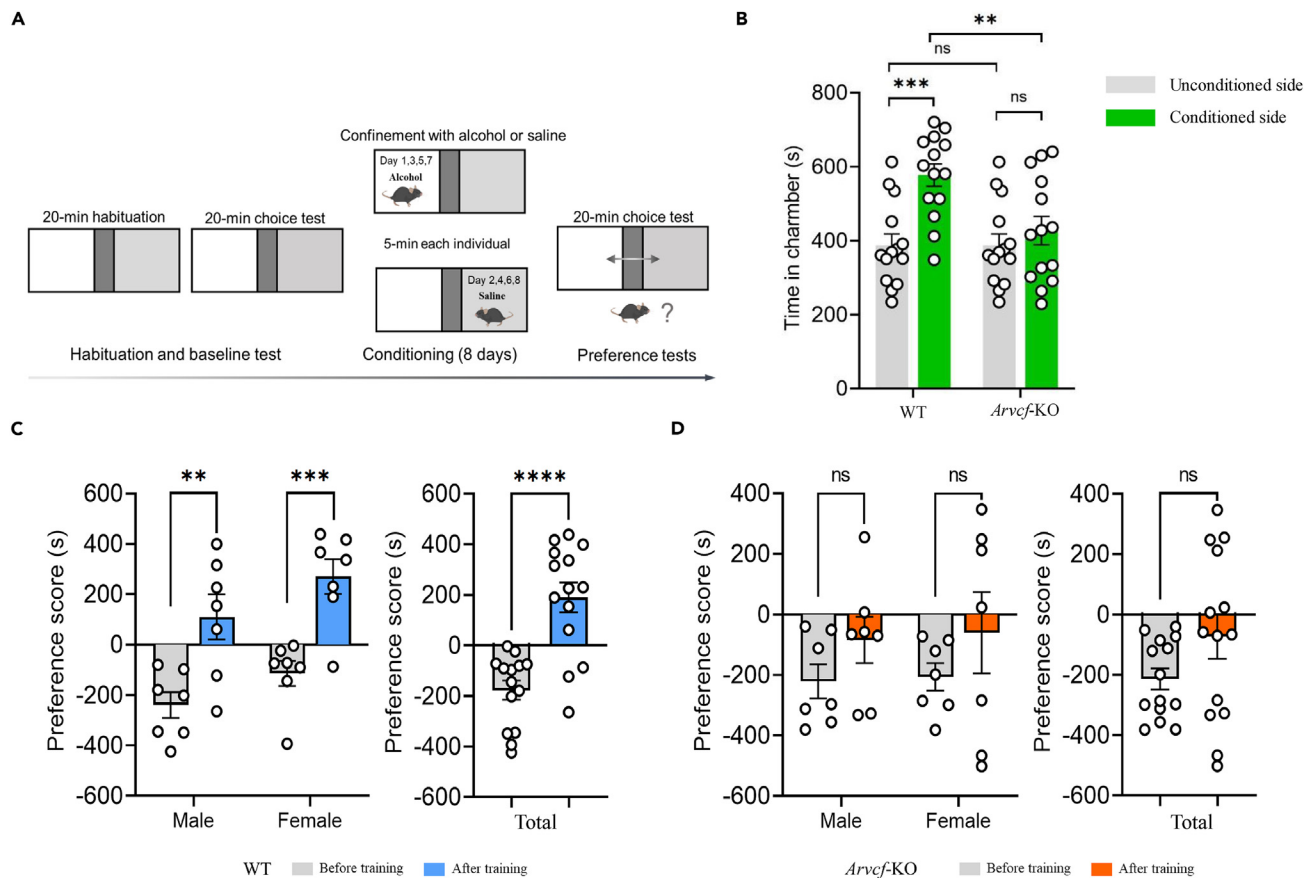
Var, number of variants within the gene;  $p_{ECS}$ ,  $p$  value of ECS (effective chi-squared statistics);  $p_{GATES}$ ,  $p$  value of GATES (gene-based association test using extended Simes).

provide more comprehensive insights. Secondly, the validation analysis in the European population showed less significant results compared to our Chinese cohort. This discrepancy might be attributed to differences in ethnic groups, as allele frequencies and the presence of certain alleles vary greatly between populations. This has also been reflected in our SNP-based analysis, where many SNPs identified in the Chinese population were absent in the European population. Lastly, although we have demonstrated that the *Arvcf* gene plays a crucial role in alcohol

**Table 3. ARVCF gene-level and transcript-level EMIC results for alcohol dependence in European samples**

Brain region	Gene/transcript	Var	Effect	SE	$p_{EMIC}$
<b>Gene-level EMIC tests</b>					
Nucleus_accumbens_basal_ganglia	ARVCF	22	0.013	0.006	0.039
Brain_Frontal_Cortex_BA9	ARVCF	16	-0.004	0.004	0.293
<b>Transcript-level EMIC tests</b>					
Brain_Amygdala	ENST00000406522	9	-0.007	0.004	0.049
Brain_Cortex	ENST00000263207	83	-0.007	0.004	0.072
Brain_Cerebellar_Hemisphere	ENST00000263207	84	-0.014	0.010	0.150
Brain_Anterior_cingulate_cortex_BA24	ENST00000473551	116	-0.004	0.004	0.322
Brain_Hippocampus	ENST00000263207	25	-4.015	4.259	0.346
Brain_Caudate_basal_ganglia	ENST00000406522	167	-0.003	0.004	0.494
Brain_Spinal_cord_cervical_c_1	ENST00000487793	105	0.001	0.004	0.838
Brain_Frontal_Cortex_BA9	ENST00000473551	58	-0.002	0.013	0.852
Brain_Cerebellum	ENST00000487793	80	0.000	0.006	0.967
Brain_Nucleus_accumbens_basal_ganglia	ENST00000473551	27	0.001	0.047	0.987
Brain_Putamen_basal_ganglia	ENST00000406522	26	-2.945	436.142	0.995

Var, number of variants within the gene; SE, standard error;  $p_{EMIC}$ ,  $p$  value of EMIC (effective-median-based Mendelian randomization framework for inferring the causal genes of complex phenotypes).



**Figure 2. *Arvcf* is required for alcohol-induced conditioned place preference in mice**

(A) Schematic of ethanol-induced conditioned place preference (CPP).

(B) A comparison of time spent in ethanol-conditioned and unconditioned chambers after ethanol-induced CPP training between wild-type (WT,  $n = 14$ , 7 males and 7 females) and *Arvcf*-KO mice ( $n = 14$ , 7 males and 7 females).

(C and D) Preference scores for the ethanol-paired chamber in wild-type mice (C) ( $n = 14$ , 7 males and 7 females; left: male and female, respectively; right: male and female pooled) and mice with knockout of *Arvcf* (D) ( $n = 14$ , 7 males and 7 females; left: male and female, respectively; right: male and female pooled). Data are presented as mean  $\pm$  SEM;  $p$  values were determined using the two-tailed unpaired  $t$  test. ns, not significant; \*\* $p < 0.01$ , \*\*\* $p < 0.001$ , \*\*\*\* $p < 0.0001$ .

dependence in mice, the molecular mechanisms underlying this association were not explored in this study. Future research should aim to elucidate the mechanisms of the *ARVCF* gene in alcohol dependence and other psychiatric conditions.

## RESOURCE AVAILABILITY

### Lead contact

Further information and request for resources should be addressed to Prof. Ming D. Li at [ml2km@zju.edu.cn](mailto:ml2km@zju.edu.cn).

### Materials availability

The study did not generate any new materials.

### Data and code availability

- Datasets used in this study are listed in the [STAR methods](#) section and supplementary files.
- This paper does not report original code.
- Any additional information required to reanalyze the data reported in this paper or reproduce the results is available from the [lead contact](#) upon request.

## ACKNOWLEDGMENTS

This study was supported in part by the National Natural Science Foundation of China (82271560), China Precision Medicine Initiative (2016YFC0906300), Research Center for Air Pollution and Health of Zhejiang University, and the State Key Laboratory for Diagnosis and Treatment of Infectious Diseases of the First Affiliated Hospital of Zhejiang University.

## AUTHOR CONTRIBUTIONS

X.S. and Y.W. conducted the experiments; X.S. and Z.Y. participated in data analysis; Z.Y., W.Y., and M.D.L. provided the resources and managed the project; X.S. and M.D.L. participated in article writing and editing; M.D.L. conceived the study and was involved in every step of the study. All authors approved the article as submitted.

## DECLARATION OF INTERESTS

The authors declare no competing interests.

## STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS
  - Subjects
  - Animals
- METHOD DETAILS
  - Alcohol phenotypes
  - Genetic data and variants association analysis
  - Gene-based association test
  - Effective-median-based mendelian randomization analysis
  - Mouse experiments
- QUANTIFICATION AND STATISTICAL ANALYSIS

## SUPPLEMENTAL INFORMATION

Supplemental information can be found online at <https://doi.org/10.1016/j.isci.2024.110976>.

Received: April 27, 2024

Revised: August 24, 2024

Accepted: September 13, 2024

Published: September 16, 2024

## REFERENCES

1. Degenhardt, L., Charlson, F., Ferrari, A., Santomauro, D., Erskine, H., Mantilla-Herrera, A., Whiteford, H., Leung, J., Naghavi, M., Griswold, M., et al. (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. *Lancet Psychiatr.* 5, 987–1012. [https://doi.org/10.1016/s2215-0366\(18\)30337-7](https://doi.org/10.1016/s2215-0366(18)30337-7).
2. Organization, W.H. (2019). *Global Status Report on Alcohol and Health 2018* (World Health Organization).
3. Hasin, D.S., Stinson, F.S., Ogburn, E., and 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. *Arch. Gen. Psychiatr.* 64, 830–842. <https://doi.org/10.1001/archpsyc.64.7.830>.
4. Verhulst, B., Neale, M.C., and Kendler, K.S. (2015). The heritability of alcohol use disorders: a meta-analysis of twin and adoption studies. *Psychol. Med.* 45, 1061–1072. <https://doi.org/10.1017/S0033291714002165>.
5. Carvalho, A.F., Heilig, M., Perez, A., Probst, C., and Rehm, J. (2019). Alcohol use disorders. *Lancet* 394, 781–792. [https://doi.org/10.1016/s0140-6736\(19\)31775-1](https://doi.org/10.1016/s0140-6736(19)31775-1).
6. Zhou, H., Kalayasiri, R., Sun, Y., Nuñez, Y.Z., Deng, H.W., Chen, X.D., Justice, A.C., Kranzler, H.R., Chang, S., Lu, L., et al. (2022). Genome-wide meta-analysis of alcohol use disorder in East Asians. *Neuropsychopharmacology* 47, 1791–1797. <https://doi.org/10.1038/s41386-022-01265-w>.
7. Zhou, H., Kember, R.L., Deak, J.D., Xu, H., Toikumoto, S., Yuan, K., Lind, P.A., Farajzadeh, L., Wang, L., Hatoum, A.S., et al. (2023). Multi-ancestry study of the genetics of problematic alcohol use in over 1 million individuals. *Nat. Med.* 29, 3184–3192. <https://doi.org/10.1038/s41591-023-02653-5>.
8. Saunders, G.R.B., Wang, X., Chen, F., Jang, S.K., Liu, M., Wang, C., Gao, S., Jiang, Y., Khunsriraksakul, C., Otto, J.M., et al. (2022). Genetic diversity fuels gene discovery for tobacco and alcohol use. *Nature* 612, 720–724. <https://doi.org/10.1038/s41586-022-05477-4>.
9. Quillen, E.E., Chen, X.D., Almasy, L., Yang, F., He, H., Li, X., Wang, X.Y., Liu, T.Q., Hao, W., Deng, H.W., et al. (2014). ALDH2 is associated to alcohol dependence and is the major genetic determinant of "daily maximum drinks" in a GWAS study of an isolated rural Chinese sample. *Am. J. Med. Genet. B Neuropsychiatr. Genet.* 165B, 103–110. <https://doi.org/10.1002/ajmg.b.32213>.
10. Sun, Y., Chang, S., Wang, F., Sun, H., Ni, Z., Yue, W., Zhou, H., Gelemtner, J., Malison, R.T., Kalayasiri, R., et al. (2019). Genome-wide association study of alcohol dependence in male Han Chinese and cross-ethnic polygenic risk score comparison. *Transl. Psychiatry* 9, 249. <https://doi.org/10.1038/s41398-019-0586-3>.
11. Tang, Y.L., Xiang, X.J., Wang, X.Y., Cubells, J.F., Babor, T.F., and Hao, W. (2013). Alcohol and alcohol-related harm in China: policy changes needed. *Bull. World Health Organ.* 91, 270–276. <https://doi.org/10.2471/BLT.12.107318>.
12. Jiang, H., Room, R., and Hao, W. (2015). Alcohol and related health issues in China: action needed. *Lancet Global Health* 3, e190–e191. [https://doi.org/10.1016/S2214-109X\(15\)70017-3](https://doi.org/10.1016/S2214-109X(15)70017-3).
13. Manthey, J., Shield, K.D., Rylett, M., Hasan, O.S.M., Probst, C., and Rehm, J. (2019). Global alcohol exposure between 1990 and 2017 and forecasts until 2030: a modelling study. *Lancet* 393, 2493–2502. [https://doi.org/10.1016/S0140-6736\(18\)32744-2](https://doi.org/10.1016/S0140-6736(18)32744-2).
14. Im, P.K., Wright, N., Yang, L., Chan, K.H., Chen, Y., Guo, Y., Du, H., Yang, X., Avery, D., Wang, S., et al. (2023). Alcohol consumption and risks of more than 200 diseases in Chinese men. *Nat. Med.* 29, 1476–1486. <https://doi.org/10.1038/s41591-023-02383-8>.
15. Yang, Z., Chen, J., Han, H., Wang, Y., Shi, X., Zhang, B., Mao, Y., Li, A.N., Yuan, W., Yao, J., and Li, M.D. (2023). Single nucleotide polymorphisms rs148582811 regulates its host gene ARVCF expression to affect nicotine-associated hippocampus-dependent memory. *iScience* 26, 108335. <https://doi.org/10.1016/j.isci.2023.108335>.
16. Sumitomo, A., Horike, K., Hirai, K., Butcher, N., Boot, E., Sakurai, T., Nucifora, F.C., Jr., Bassett, A.S., Sawa, A., and Tomoda, T. (2018). A mouse model of 22q11.2 deletions: Molecular and behavioral signatures of Parkinson's disease and schizophrenia. *Sci. Adv.* 4, eaar6637. <https://doi.org/10.1126/sciadv.aar6637>.
17. Yuan, L., and Arikkath, J. (2017). Functional roles of p120ctn family of proteins in central



- neurons. *Semin. Cell Dev. Biol.* 69, 70–82. <https://doi.org/10.1016/j.semcdb.2017.05.027>.
18. Donta, M.S., Srivastava, Y., Di Mauro, C.M., Paulucci-Holthausen, A., Waxham, M.N., and McCrea, P.D. (2023). p120-catenin subfamily members have distinct as well as shared effects on dendrite morphology during neuron development in vitro. *Front. Cell. Neurosci.* 17, 1151249. <https://doi.org/10.3389/fncel.2023.1151249>.
  19. Peterson, V.L., McCool, B.A., and Hamilton, D.A. (2015). Effects of ethanol exposure and withdrawal on dendritic morphology and spine density in the nucleus accumbens core and shell. *Brain Res.* 1594, 125–135. <https://doi.org/10.1016/j.brainres.2014.10.036>.
  20. Giannakopoulou, O., Lin, K., Meng, X., Su, M.-H., Kuo, P.-H., Peterson, R.E., Awasthi, S., Moscati, A., Coleman, J.R.I., Bass, N., et al. (2021). The Genetic Architecture of Depression in Individuals of East Asian Ancestry. *JAMA Psychiatr.* 78, 1258–1269. <https://doi.org/10.1001/jamapsychiatry.2021.2099>.
  21. Denny, J.C., Ritchie, M.D., Basford, M.A., Pulley, J.M., Bastarache, L., Brown-Gentry, K., Wang, D., Masys, D.R., Roden, D.M., and Crawford, D.C. (2010). PheWAS: demonstrating the feasibility of a phenome-wide scan to discover gene-disease associations. *Bioinformatics* 26, 1205–1210. <https://doi.org/10.1093/bioinformatics/btq126>.
  22. Denny, J.C., Bastarache, L., Ritchie, M.D., Carroll, R.J., Zink, R., Mosley, J.D., Field, J.R., Pulley, J.M., Ramirez, A.H., Bowton, E., et al. (2013). Systematic comparison of phenome-wide association study of electronic medical record data and genome-wide association study data. *Nat. Biotechnol.* 31, 1102–1110. <https://doi.org/10.1038/nbt.2749>.
  23. Dines, J.N., Golden-Grant, K., LaCroix, A., Muir, A.M., Cintrón, D.L., McWalter, K., Cho, M.T., Sun, A., Merritt, J.L., Thies, J., et al. (2019). TANGO2: expanding the clinical phenotype and spectrum of pathogenic variants. *Genet. Med.* 21, 601–607. <https://doi.org/10.1038/s41436-018-0137-y>.
  24. Egervari, G., Siciliano, C.A., Whiteley, E.L., and Ron, D. (2021). Alcohol and the brain: from genes to circuits. *Trends Neurosci.* 44, 1004–1015. <https://doi.org/10.1016/j.tins.2021.09.006>.
  25. Ben Hamida, S., Mendonça-Netto, S., Arefin, T.M., Nasseef, M.T., Boulos, L.J., McNicholas, M., Ehrlich, A.T., Clarke, E., Moquin, L., Gratton, A., et al. (2018). Increased Alcohol Seeking in Mice Lacking Gpr88 Involves Dysfunctional Mesocorticolimbic Networks. *Biol. Psychiatr.* 84, 202–212. <https://doi.org/10.1016/j.biopsych.2018.01.026>.
  26. Chang, C.C., Chow, C.C., Tellier, L.C., Vattikuti, S., Purcell, S.M., and 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>.
  27. Li, M., Jiang, L., Mak, T.S.H., Kwan, J.S.H., Xue, C., Chen, P., Leung, H.C.-M., Cui, L., Li, T., Sham, P.C., and Valencia, A. (2019). A powerful conditional gene-based association approach implicated functionally important genes for schizophrenia. *Bioinformatics* 35, 628–635. <https://doi.org/10.1093/bioinformatics/bty682>.
  28. Li, M.-X., Gui, H.-S., Kwan, J.S.H., and Sham, P.C. (2011). GATES: A Rapid and Powerful Gene-Based Association Test Using Extended Simes Procedure. *Am. J. Hum. Genet.* 88, 283–293. <https://doi.org/10.1016/j.ajhg.2011.01.019>.
  29. Jiang, L., Miao, L., Yi, G., Li, X., Xue, C., Li, M.J., Huang, H., and Li, M. (2022). Powerful and robust inference of complex phenotypes' causal genes with dependent expression quantitative loci by a median-based Mendelian randomization. *Am. J. Hum. Genet.* 109, 838–856. <https://doi.org/10.1016/j.ajhg.2022.04.004>.
  30. Dong, S.S., He, W.M., Ji, J.J., Zhang, C., Guo, Y., and Yang, T.L. (2021). LDBlockShow: a fast and convenient tool for visualizing linkage disequilibrium and haplotype blocks based on variant call format files. *Briefings Bioinf.* 22, bbaa227. <https://doi.org/10.1093/bib/bbaa227>.
  31. Ma, Y., Wen, L., Cui, W., Yuan, W., Yang, Z., Jiang, K., Jiang, X., Huo, M., Sun, Z., Han, H., et al. (2017). Prevalence of Cigarette Smoking and Nicotine Dependence in Men and Women Residing in Two Provinces in China. *Front. Psychiatr.* 8, 254. <https://doi.org/10.3389/fpsy.2017.00254>.
  32. Li, Y., Jiang, Y., Zhang, M., Yin, P., Wu, F., and Zhao, W. (2011). Drinking behaviour among men and women in China: the 2007 China Chronic Disease and Risk Factor Surveillance. *Addiction* 106, 1946–1956. <https://doi.org/10.1111/j.1360-0443.2011.03514.x>.
  33. Liu, M., Jiang, Y., Wedow, R., Li, Y., Brazel, D.M., Chen, F., Datta, G., Davila-Velderrain, J., McGuire, D., Tian, C., et al. (2019). 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>.
  34. Li, M.D., Liu, Q., Shi, X., Wang, Y., Zhu, Z., Guan, Y., He, J., Han, H., Mao, Y., Ma, Y., et al. (2024). Integrative analysis of genetics, epigenetics and RNA expression data reveal three susceptibility loci for smoking behavior in Chinese Han population. *Mol. Psychiatr.* <https://doi.org/10.1038/s41380-024-02599-1>.
  35. Xu, K., Kranzler, H.R., Sherva, R., Sartor, C.E., Almasy, L., Koesterer, R., Zhao, H., Farrer, L.A., and Gelernter, J. (2015). Genomewide Association Study for Maximum Number of Alcoholic Drinks in European Americans and African Americans. *Alcohol Clin. Exp. Res.* 39, 1137–1147. <https://doi.org/10.1111/acer.12751>.
  36. 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., et al. (2018). Transancestral GWAS of alcohol dependence reveals common genetic underpinnings with psychiatric disorders. *Nat. Neurosci.* 21, 1656–1669. <https://doi.org/10.1038/s41593-018-0275-1>.
  37. 1000 Genomes Project Consortium, Auton, A., Brooks, L.D., Durbin, R.M., Garrison, E.P., Kang, H.M., Korbel, J.O., Marchini, J.L., McCarthy, S., McVean, G.A., and Abecasis, G.R. (2015). A global reference for human genetic variation. *Nature* 526, 68–74. <https://doi.org/10.1038/nature15393>.
  38. Mews, P., Egervari, G., Nativo, R., Sidoli, S., Donahue, G., Lombroso, S.I., Alexander, D.C., Riesche, S.L., Heller, E.A., Nestler, E.J., et al. (2019). Alcohol metabolism contributes to brain histone acetylation. *Nature* 574, 717–721. <https://doi.org/10.1038/s41586-019-1700-7>.

## STAR★METHODS

## KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Deposited data		
GTEEx v8 eQTL data	KGGSEE website	<a href="https://pmglab.top/kggsee/#/download">https://pmglab.top/kggsee/#/download</a>
European alcohol dependence GWAS summary statistics	Psychiatric Genomics Consortium	<a href="https://pgc.unc.edu/for-researchers/download-results/">https://pgc.unc.edu/for-researchers/download-results/</a>
Experimental models: Organisms/strains		
Mouse: C57BL/6	GemPharmatech	N/A
Mouse: <i>Arvcf</i> knockout	GemPharmatech	N/A
Software and algorithms		
PLINK v1.9	Chang et al. <sup>26</sup>	<a href="https://www.cog-genomics.org/plink/">https://www.cog-genomics.org/plink/</a>
ECS on KGGSEE v1.0	Li et al. <sup>27</sup>	<a href="https://pmglab.top/kggsee/#/">https://pmglab.top/kggsee/#/</a>
GATES on KGGSEE v1.0	Li et al. <sup>28</sup>	<a href="https://pmglab.top/kggsee/#/">https://pmglab.top/kggsee/#/</a>
EMIC on KGGSEE v1.0	Lin et al. <sup>29</sup>	<a href="https://pmglab.top/kggsee/#/">https://pmglab.top/kggsee/#/</a>
R	N/A	<a href="https://www.r-project.org">https://www.r-project.org</a>
LDBlockShow	Dong et al. <sup>30</sup>	<a href="https://github.com/BGI-shenzhen/LDBlockShow">https://github.com/BGI-shenzhen/LDBlockShow</a>
GraphPad Prism 9.0.0	GraphPad Software Inc	<a href="https://www.graphpad.com/">https://www.graphpad.com/</a>
ANY-maze 6.0	Stoelting	<a href="https://stoeltingco.com/Neuroscience/ANY-maze">https://stoeltingco.com/Neuroscience/ANY-maze</a>

## EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS

## Subjects

DNA samples and phenotypic data were collected from Chinese participants. All participants were recruited from Shanxi Province between 2012 and 2014, as detailed in our previous study.<sup>15,31</sup> The alcohol-related questionnaires, including drinking status and drinks per week, were administered, and demographic characteristics such as age and gender were obtained from each individual. Participants diagnosed with psychiatric disorders as defined by the DSM-IV, such as schizophrenia and major depressive disorders, were excluded from the study. Only males were included in the study due to few females drinking either in our study or in China.<sup>32</sup> In total, 1,329 biologically unrelated subjects were included for whole genome-sequencing (WGS).

## Animals

The Wild-type (WT, C57BL/6) mice were purchased from GemPharmatech (Nanjing, China). The *Arvcf*-KO mice were generated using the CRISPR/Cas9 system and provided by GemPharmatech. All mice were socially housed (3–5 animals per cage) and maintained on a 12-h light-dark cycle (light on from 7:00 to 19:00) with food and water provided *ad libitum*. The project was approved by the Animal Care and Use Committee of the First Affiliated Hospital of Zhejiang University (Approval #: 2023-596).

## METHOD DETAILS

## Alcohol phenotypes

Currently consuming any alcoholic beverages was defined as being a drinker, and drinks per week was defined as the average number of drinks a participant reported drinking each week, including all types of alcohol. While a range of response for drinks per week was recorded, the midpoint was used (e.g., 2 drinks per week on average instead of a range of 1–3 drinks per week). This phenotype was left-anchored at 1 and log-transformed before analysis, as previously described.<sup>33</sup>

## Genetic data and variants association analysis

Whole genome sequencing and quality control were performed as described in our recent report.<sup>34</sup> Briefly, genomic DNA was extracted from blood and sequenced on the Illumina HiSeqX platform. After quality control, all sequencing reads were aligned to the NCBI build 37 of the human reference genome using BWA, and variants were determined using GATK, as detailed in our recent report.<sup>34</sup> The SNP-based

association analysis was performed using PLINK v1.9<sup>26</sup> under an additive genetic model. SNPs with a minor allele frequency (MAF) < 0.01, SNPs and samples with a call rate < 0.99, and P-value for Hardy-Weinberg equilibrium (HWE) <  $5 \times 10^{-6}$  were excluded from the analysis.<sup>15,35</sup> After implementing quality control procedures, we analyzed data from 470 nondrinkers and 834 alcohol drinkers, along with their corresponding drinks per week, while accounting for age and the first three multidimensional scaling (MDS) components as covariates.<sup>15</sup> We extracted all variants located in *ARVCF* gene region (including  $\pm 2$  kb flanking regions) from the whole genome-sequencing analysis data for further analyses. Regional association plot was generated using LDBlockShow.<sup>30</sup>

### Gene-based association test

For our Chinese dataset, the SNP-based association analysis results in the *ARVCF* gene region were applied to gene-based and transcript-based association testing using KGGSEE software, which employs the approaches of a rapid and powerful Gene-based Association Test using Extended Simes procedure (GATES)<sup>28</sup> and Effective Chi-square Statistics (ESC).<sup>27</sup> The gene- and transcript-based eQTL summary statistics from GTEx v8 brain tissues were utilized. SNPs with MAF  $\leq 0.01$  and eQTL summary statistics with P-value > 0.05 were excluded from the analysis. The dataset used for validation on alcohol dependence was the largest so far publicly available genome-wide association study on DSM-IV-diagnosed alcohol dependence, including 14,904 individuals with alcohol dependence and 37,944 controls from 28 case-control and family-based studies.<sup>36</sup> We applied the same procedures as those used for the Chinese data to generate gene-based association estimates for *ARVCF* in relation to alcohol dependence.

### Effective-median-based mendelian randomization analysis

We used the EMIC (Effective-median-based Mendelian randomization framework for Inferring the Causal genes of complex phenotypes) to infer the causal effect of *ARVCF* gene expression on alcohol use.<sup>29</sup> To achieve this, the pre-calculated *cis*-eQTLs ( $p \leq 0.05$ ) in 13 brain tissues with gene-level and transcript-level expression from GTEx (V8) and SNPs (MAF > 0.01) within the *ARVCF* gene were utilized for the EMIC analysis. In the EMIC analysis, variants affecting *ARVCF* expression, as well as their effect sizes and standard errors for alcohol-related phenotypes, were used as input. This allowed EMIC to estimate the causal effect of gene expression or transcript levels on alcohol use. Since allele frequencies for alcohol dependence were not reported, we utilized the allele frequencies from the European 1000 Genomes dataset.<sup>37</sup> We applied both the Chinese data and dataset of alcohol dependence from European ancestry<sup>36</sup> to EMIC analysis.

### Mouse experiments

Behavioral assays were conducted when the animals were 8 weeks old. The mice were placed in the testing room for habituation at least 30 min before the test and were handled once per day for at least 4 days before the behavioral test. The behavioral tasks were performed during the light phase. In the ethanol-induced CPP test, both adult *Arvcf*-KO and WT male and female mice were used. Ethanol was administered via intraperitoneal injection at a dose of 2 g/kg (20% solution in saline).<sup>38</sup> We used a three-chamber CPP apparatus consisting of two large boxes that act as conditioning chambers separated by a small middle chamber serving as a transit route between boxes. The two side chambers (450 × 450 × 450 mm) had different wall colors (white and black) and different floor textures, with one having a grid metal floor and the other having a floor with horizontal metal strips. The paradigm consisted of four phases: habituation, pre-conditioning, conditioning, and post-conditioning. During the phases of habituation and pre-conditioning (day 1–2), guillotine doors were removed, and the animal was placed in the middle chamber to allow it to freely explore the entire apparatus for 20 min. Animal behavior was recorded by a video-tracking system (ANY-maze 6.0), and the time spent in each chamber was automatically calculated on day 2 as the baseline for the CPP test. Mice were paired with ethanol on the least preferred side, while the other side was the saline-paired chamber. During the conditioning phase (day 3–10), mice received alternating intraperitoneal injections of saline or ethanol and were immediately confined to the saline- or drug-paired chamber for 5 min. In the post-conditioning phase (day 11), the procedure was the same as in the pre-conditioning phase, i.e., mice were allowed access to both chamber for 20 min. Preference scores were calculated by subtracting the time spent in the saline-paired side from the ethanol-paired side.<sup>38</sup>

### QUANTIFICATION AND STATISTICAL ANALYSIS

PLINK v1.9 was used for SNP-based association analysis.<sup>26</sup> The KGGSEE software was employed for gene-based and transcript-based association tests,<sup>27,28</sup> as well as for EMIC analyses.<sup>29</sup> In the animal studies, data are presented as Mean  $\pm$  Standard Error of the Mean (SEM), as indicated in the figure legends. Two-tailed unpaired t-test were used for two-group comparisons.