Supporting Information

Direct sequencing of *ADH1B* and *ALDH2* genes

ADH1B rs1229984 and *ALDH2 rs671* polymorphisms were confirmed by direct sequencing. The primers were listed in **Supplemental Table 3**. The PCR was carried out with 50 ng/μL genomic DNA. The PCR condition was performed by initial denaturation at 95℃ for 10min, followed by 35 cycles for at 95℃ for 30 second, 60℃ for 30 second, 72℃ for 30 second and a final extension at 72℃ for 10 min. Sequencing reactions were custom performed by Genomics Biotechnology (Taiwan).

Supplemental Table 1 TaqMan probes used of *ALDH2* and *ADH1B* genes

Genes	**SNP**	**Assay ID**	**Polymorphism**	**Context Sequence [VIC/FAM]**
ALDH2	*rs671*	C_11703892_10	A/G	CGAGTACGGGCTGCAGGCATACACT **[A/G]** AAGTGAAAACTGTGAGTGTGGGACC
ADH1B	*rs1229984*	C_2688467_20	C/T, Transition Substitution	GCCACTAACCACGTGGTCATCTGTG **[C/T]** GACAGATTCCTACAGCCACCATCTA

Supplemental Table 2 Sequences of PCR primers

Genes	Primer name	Sequence (5'→ 3')	Product length (bp)
ALDH2	*ALDH2*-Forward	5'-TCCTGGGAGTGTAACCCATA-3'F	234
	ALDH2-Reverse	5'-CACCAGCAGACCCTCAAG-3'R	
ADH1B	*ADH1B*-Forward	5'-TCTGAATCTGAACAGCTTCTCTT-3'F	194
	ADH1B-Reverse	5'-TCCCTGAGTGTGAATCCTGT-3'R	

Supplemental Table 3 Allele frequencies of TWB control group consisted of 1:4 frequency-matching (TWB-1, N =600) and the residual TWB subjects (TWB-2, N=29,683)

Gene (SNP)	Genotypes	TWB-1 (N = 600)		TWB-2 (N = 29,683)	
		N (%)	HWE[1] *p-value*	N (%)	HWE[1] *p-value*
ADH1B rs1229984	***2/*2**(AA)	311(51.8)	0.1567	16,248(54.7)	0.2837
	***1/*2**(GA)	232(38.7)		11,378(38.3)	
	***1/*1**(GG)	57(9.5)		2,057(6.9)	
ALDH2 rs671	***1/*1**(GG)	341(56.8)	0.9139	15,213(51.3)	0.6796
	***1/*2**(GA)	222(37.0)		12,054(40.6)	
	***2/*2**(AA)	37(6.2)		2,416(8.1)	

[1]HWE, Hardy-Weinberg equilibrium**.**

Supplemental Table 4 Sensitivity analysis of 106 LAD patients and 424 controls from TWB1 with individual-matching under the restriction of age between 30 and 70 years old.

N (%)	LAD (N=106)	TWB-1 (N=424)	*p*-value[1]
Gender			*1.000*
Male	35 (33.0)	140 (33.0)	
Female	71 (67.0)	284 (67.0)	
Age in years			*0.912*
Mean ± SD	58.21±7.26	58.29±7.29	
Median	58.0	58.0	
IQR	54.0-65.0	54.0-65.0	
Education			*1.000*
≤ Elementary school	17 (16.0)	68 (16.0)	
Junior high school	11 (10.4)	44 (10.4)	
Senior high school	41 (38.7)	164 (38.7)	
≥ University	37 (34.9)	148 (34.9)	
Cigarette smoking			*0.819*
Non-smoker	87 (82.1)	358 (84.4)	
Former smoker	8 (7.5)	26 (6.1)	
Current smoker	11 (10.4)	40 (9.4)	
Alcohol drinking			*0.481*
Non-drinker	97 (91.5)	396 (93.4)	
Former drinker	2 (1.9)	11 (2.6)	
Current drinker	7 (6.6)	17 (4.0)	

Abbreviations: N, numbers; LAD, lung adenocarcinoma; TWB, Taiwan Biobank; IQR, interquartile range; SD, standard deviation.
[1]*p-values* were obtained using the Chi-square test or t-test

Supplemental Table 5 Sensitivity analysis of the effect of *ADH1B* and *ALDH2* genotypes of 106 lung adenocarcinoma patients and 424 controls in the conditional logistic regression models.

SNPs, N (%)	LAD (N=106)	TWB-1 (N=424)	Crude OR	95% CI		AOR[1]	95% CI		*P-value*[2]
				LL	UL		LL	UL	
ADH1B -rs1229984									
****2 (A)***	146 (68.9)	613 (72.3)	1			1			
****1 (G)***	66 (31.1)	235 (27.7)	1.17	0.84	1.64	1.18	0.84	1.65	0.344
ADH1B -rs1229984									
****2/*2(AA)***	48 (45.3)	225 (53.1)	1			1			
****1/*2(GA)***	50 (47.2)	163 (38.4)	1.37	0.87	2.17	1.34	0.85	2.12	0.212
****1/*1(GG)***	8 (7.5)	36 (8.5)	1.09	0.47	2.54	1.16	0.49	2.72	0.736
ADH1B -rs1229984									
****2/*2(AA)***	48 (45.3)	225 (53.1)	1			1			
****1/*2+*1/*1(GA +GG)***	58 (54.7)	199 (46.9)	1.33	0.85	2.06	1.31	0.84	2.04	0.230
ALDH2-rs671									
****1 (G)***	138 (65.1)	634 (74.8)	1			1			
****2 (A)***	74 (34.9)	214 (25.2)	**1.57**	**1.13**	**2.19**	**1.58**	**1.13**	**2.21**	**0.007****
ALDH2-rs671									
****1/*1(GG)***	47 (44.3)	238 (56.1)	1			1			
****1/*2(GA)***	44 (41.5)	158 (37.3)	1.40	0.87	2.25	1.33	0.82	2.16	0.242
****2/*2(AA)***	15 (14.2)	28 (6.6)	**2.90**	**1.39**	**6.07**	**3.19**	**1.51**	**6.75**	**0.002****
ALDH2-rs671									
****1/*1+*1/*2(GG+GA)***	91 (85.8)	396 (93.4)	1			1			
****2/*2 (AA)***	15 (14.2)	28 (6.6)	**2.42**	**1.22**	**4.81**	**2.75**	**1.36**	**5.54**	**0.005****

Abbreviations: *ADH1B, alcohol dehydrogenase 1B*; *ALDH2, aldehyde dehydrogenase 2 family member gene*; OR, odd ratio; AOR, adjusted OR; LL, lower limit of 95% CI; UL, upper limit of 95% CI.

[1]Adjusting for smoking status (former smoker *vs*. non-smoker; current smoker *vs*. non-smoker) and alcohol drinking (former drinker *vs*. non-drinker; current drinker *vs*. non-drinker).

[2]*p-values* were bolded under the 0.025 at the significant levels with Bonferroni adjustment

Supplemental Table 6 Characteristics of studies related with ***ALDH2*** **allele frequency** and lung cancer and lung adenocarcinoma

Author/Country/Year	Study design	Cancer type	Allele frequency (N, %)						Crude OR (95% CI)
			Cases			Controls			
			Total	*1	*2	Total	*1	*2	
Minegishi, Japan, 2007	HB Ca/Co	LC	1010	806 (79.80)	204 (25.31)	512	376 (73.44)	136 (36.17)	0.70 (0.55-0.90)
		LAD	660	526 (79.70)	134 (25.48)	512			0.70 (0.54-0.93)
Park, Japan, 2010	HB Ca/Co	LC	1,436	970 (67.55)	466 (48.04)	2,832	1,981 (69.95)	851 (42.96)	1.12 (0.98-1.28)
Shimizu, Japan, 2013	HB Ca/Co	LC	1,144	796 (69.58)	348 (43.72)	930	720 (77.42)	210 (29.17)	1.50 (1.23-1.83)
Our study	TWB-1[1] Ca/Co	LAD	300	198 (66.0)	102 (34.0)	1,200	904 (75.3)	296 (24.7)	**1.57 (1.20-2.07)**
	TWB-2[2] Ca/Co					59,366	42,480 (71.6)	16,886 (28.4)	**1.30 (1.02-1.65)**

Abbreviations: HB, hospital-based; Ca, case; Co, control; LC, lung cancer; LAD, lung adenocarcinoma.

[1]TWB-1 using 1:4 frequency-matching of age group (30-60; >60 years), sex, and education level (four levels: ≤ elementary school, junior high school, senior high school, and ≥ university) from the original 30,283 subjects of TWB dataset.

[2] TWB-2 consists of the remaining subjects from the original 30,283 participants in the TWB dataset, after the exclusion of the 600 subjects of TWB-1.

Supplemental Table 7 Allele frequencies of *ADH1B rs1229984* and *ALDH2 rs671* genotypes among various ethnic groups.

Regions	Populations	*ADH1B, rs1229984*			*ALDH2, rs671*			References
		N	*1 (%)	*2 (%)	N	*1 (%)	*2 (%)	
America	Mexican	97	100.0	0.0	101	99.5	0.5	[1]
South America	Colombian	148	91.55	8.45	148	100.0	0.0	[2]
Europe	Danish	1,188	99.2	0.80	1,188	99.96	0.04	[3]
South Asia	Indian	108	99.54	0.46	108	98.15	1.85	[4]
Southwest Asia	Turkish	81	99.38	0.62	81	100.0	0.0	[5]
Southeast Asia	Vietnamese	235	31.49	68.51	235	79.57	20.43	[6]
East Asia	Chinese	9,339	31.27	68.73	9,339	78.74	21.26	[7]
	Han Chinese in Beijing	206	29.13	70.87	206	83.98	16.02	[8,9]
	Southern Han Chinese	**210**	**24.29**	**75.71**	**210**	**72.86**	**27.14**	[8,9]
	Korean	796	24.81	75.19	796	83.79	16.21	[10]
	Japanese	914	22.05	77.95	1,935	76.54	23.46	[11]
	Taiwanese	**30,283**	**25.58**	**72.44**	**30,283**	**70.14**	**27.88**	**Our study**

Supplemental Table 8 Summarized the frequency of *ADH1B* and *ALDH2* genes in Taiwan population from literature review (focus on case control study, only shown the control groups)

References/Author/Year	Database sources	Total numbers (N)	Age (years ± SD)	Female (%)	*ADH1B, rs1229984* N (%)			HWE *P* value	*ALDH2, rs671* N (%)			HWE *P* value
					***2/*2**	***1/*2**	***1/*1**		***1/*1**	***1/*2**	***2/*2**	
[12] Chung, 2014	NTUH [3]	276	57.6±10.4	5.8	125 (47.9)	111 (42.5)	25 (9.6)	0.960	128 (49.0)	113 (43.3)	20 (7.7)	0.469
[13] Tsai, 2014	NCKUH [4]	514	53.7±0.4 (SE)	4.3	268 (52.1)	221 (43.0)	25 (4.9)	**0.015**	267 (51.9)	203 (39.5)	44 (8.6)	0.541
[14] Liu, 2016	REVEAL-HBV cohort [5]	3,213	30-65	41.2	1,748 (54.4)	1,229 (38.3)	236 (7.3)	0.325	1,617 (50.2)	1,354 (42.0)	250 (7.8)	0.149
[15] Wu, 2016	Population-based	267	45.2±12.0	NA	152 (56.9)	99 (37.1)	16 (6.0)	0.982	130 (47.8)	116 (42.6)	26 (9.6)	0.986
[16] Huang, 2017	NCKUH [4]	940	54.6±0.3 (SE)	5.3	484 (52.9)	381 (41.6)	50 (5.5)	**0.024**	447 (48.8)	386 (42.2)	82 (9.0)	0.918
[17] Chien, 2019	TGP [6]	427	40.0±17.7	44.3	235 (55)	166 (38.9)	26 (6.1)	0.642	215 (50.4)	172 (40.3)	40 (9.4)	0.512
[18] Huang, 2020	CGMH [7]	244	60±10.5	47.5	129 (52.9)	105 (43)	10 (4.1)	**0.044**	118 (48.4)	105 (43)	21 (8.6)	0.729
[19] Wu, 2021	TWB [8]	2,505	67.1±6.3	49.2	1,363 (54.4)	984 (39.3)	158 (6.3)	0.268	1,267 (50.6)	1,022 (40.8)	214 (8.5)	0.698
[20] Yang, 2021	TWB [8]	1,512	NA	NA	789 (52.2)	611 (40.4)	112 (7.4)	0.673	785 (52.1)	603 (40)	119 (7.9)	0.831
[21] Lin, 2021	TWB [8]	16,837	48.3±10.9	51.8	9,159 (54.4)	6,568 (39)	1,110 (6.6)	0.142	8,554 (50.8)	6,832 (40.6)	1,451 (8.6)	0.098
[22] Shan, 2021	NCKUH [4]	940	58.3±0.4 (SE)	59.0	502 (54.6)	334 (36.3)	83 (9)	**0.013**	450 (47.9)	407 (43.3)	83 (8.8)	0.506
Our study	TWB-1[1]	600	61.4±7.8	67.0	311 (51.8)	232 (38.7)	57 (9.5)	0.157	341 (56.8)	222 (37.0)	37 (6.2)	0.914
	TWB-2[2]	29,683	49.3±10.5	67.0	16,248 (54.7)	11,378 (38.8)	2,057 (6.9)	0.284	15,213 (51.3)	12054 (40.6)	2,416 (8.1)	0.680

Abbreviations: HWE, Hardy-Weinberg equilibrium; NA, not available.

[1]TWB-1 using 1:4 frequency-matching of age group (30-60; >60 years), sex, and education level (four levels: ≤ elementary school, junior high school, senior high school, and ≥ university) from the original 30,283 subjects of TWB dataset.

[2] TWB-2 consists of the remaining subjects from the original 30,283 participants in the TWB dataset, after the exclusion of the 600 subjects of TWB-1.

[3] NTUH=National Taiwan University Hospital.

[4] NCKUH=National Cheng Kung University Hospital.

[5] REVEAL-HBV=Associated Liver Disease/Cancer in HBV in Taiwan.

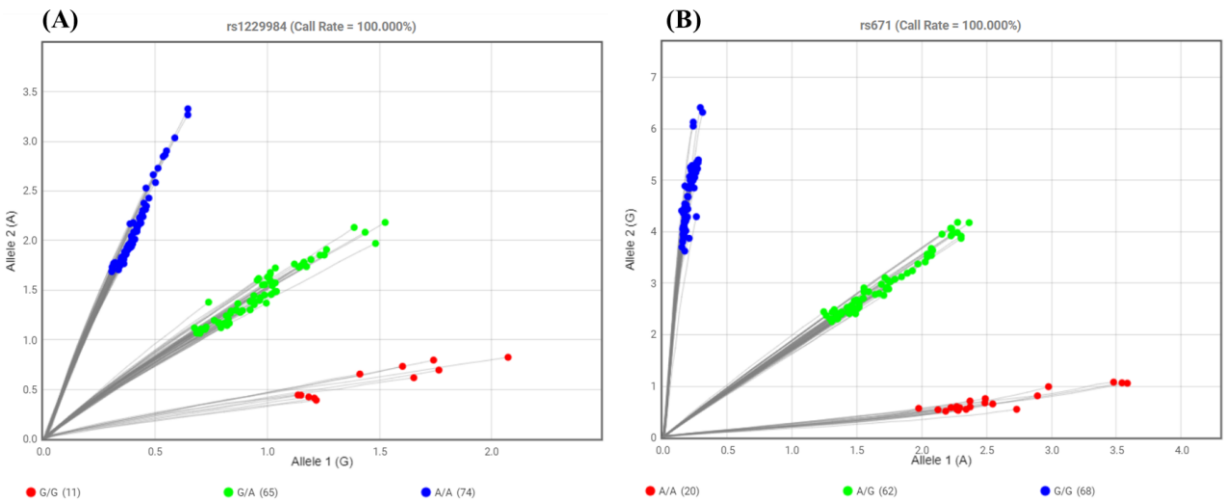
[6] TGP=Taiwanese General Population.

[7] CGMH=Chang Gung Memory Hospital.

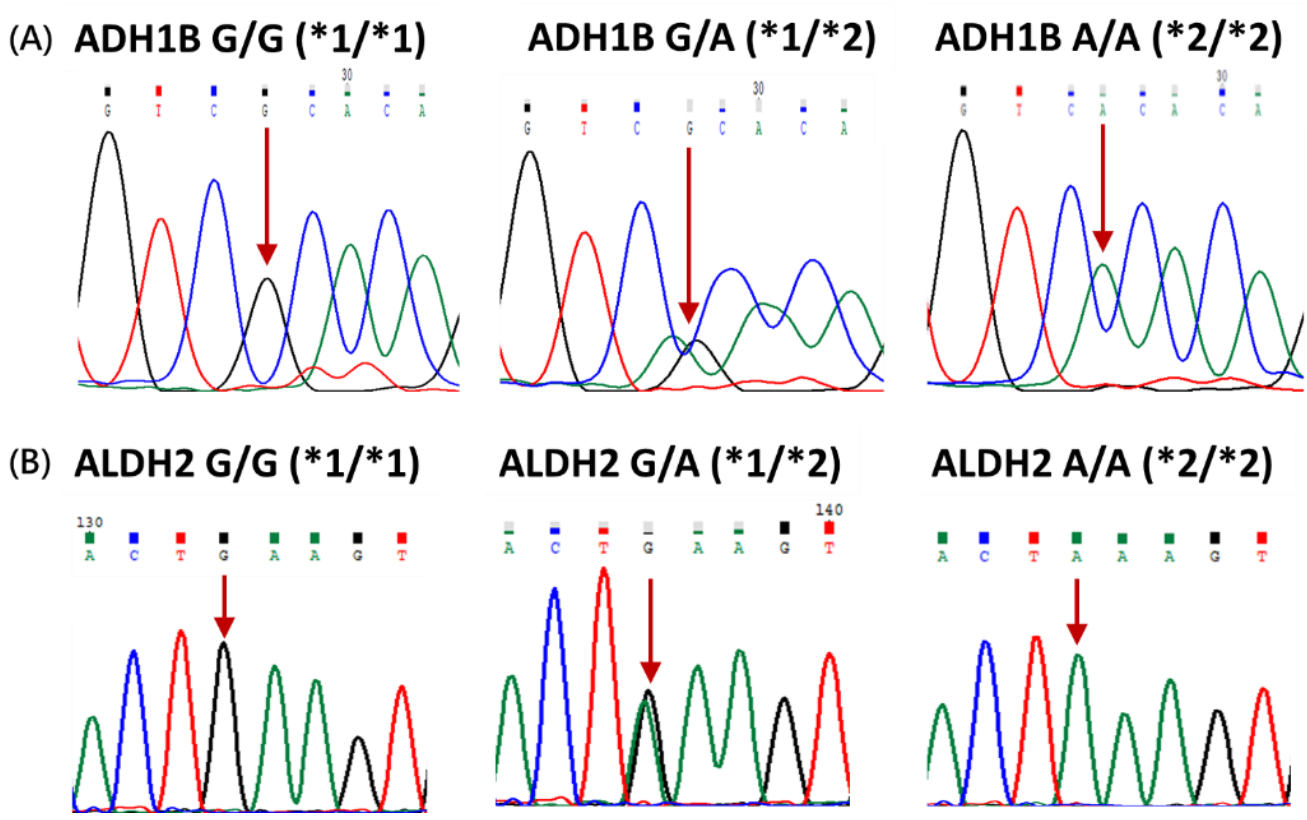
[8] TWB=Taiwan biobank.

Supplemental Table 9 Allele frequencies of *ADH1B* and *ALDH2* genotypes by age in 10-year range of the TWB cohort

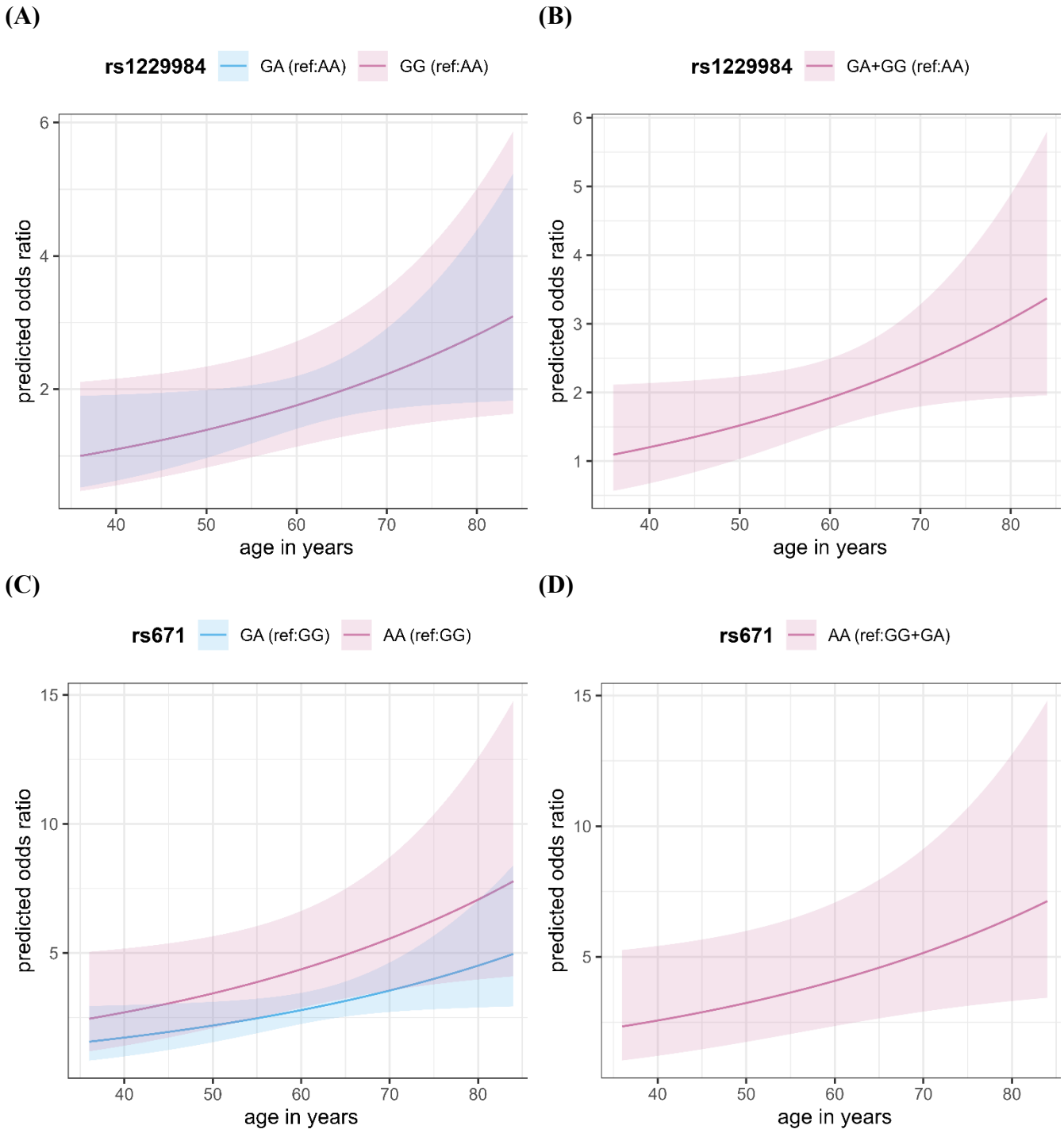
Age range (years)	Total numbers (N=86,322)	*ADH1B, rs1229984* N (%)			*ALDH2, rs671* N (%)		
		***2/*2**	***1/*2**	***1/*1**	***1/*1**	***1/*2**	***2/*2**
30-39	18,502	9,968 (53.88)	7,219 (39.02)	1,315 (7.11)	9,584 (51.80)	7,430 (40.16)	1,488 (8.04)
40-49	20,248	10,977 (54.21)	7,831 (38.68)	1,440 (7.11)	10,442 (51.57)	8,152 (40.26)	1,654 (8.17)
50-59	27,095	14,657 (54.09)	10,505 (38.77)	1,933 (7.13)	14,074 (51.94)	10,896 (40.21)	2,125 (7.84)
60-69	20,142	10,876 (54.00)	7,832 (38.88)	1,434 (7.12)	10,381 (51.54)	8,193 (40.68)	1,568 (7.78)
70-79	335	167 (49.85)	138 (41.19)	30 (8.96)	172 (51.34)	141 (42.09)	22 (6.57)



Supplemental Fig. 1. The allelic discrimination plots of TaqMan genotyping assays included 150 case patients. (A) *ADH1B rs1229984* (n = 150); (B) *ALDH2 rs671* (n = 150). Red and blue dots represent the homozygous genotypes and green circles represent the heterozygous genotypes. (*rs1229984**1/*1: homozygous for G/G; *rs1229984**1/*2: heterozygous for G/A; *rs1229984**2/*2: homozygous for A/A); *rs671* *1/*1: homozygous for G/G; *rs671**1/*2: heterozygous for A/G; *rs671**2/*2: homozygous for A/A).



Supplemental Fig. 2. The representative sequencing data of *ADH1B rs1229984* and *ALDH2 rs671*. (A) *ADH1B rs1229984* (*ADH1B*1/*1*: homozygous for G/G; *ADH1B*1/*2*: heterozygous for G/A; *ADH1B*2/*2*: homozygous for A/A); (B) *ALDH2 rs671* (*ALDH2*1/*1*: homozygous for G/G; *ALDH2*1/*2*: heterozygous for A/G; *ALDH2*2/*2*: homozygous for A/A).



Supplemental Fig. 3. Predicted odds ratios for *ADH1B rs1229984* and *ALDH2 rs671* across age in the 150 lung adenocarcinoma patients and 600 controls from TWB-1. (A-B) *ADH1B rs1229984*; (C-D) *ALDH2 rs671*.

References

1. Gordillo- Bastidas E, Panduro A, Gordillo- Bastidas D, Zepeda- Carrillo EA, García- Bañuelos JJ, Muñoz-Valle JF, *et al*. Polymorphisms of alcohol metabolizing enzymes in indigenous Mexican population: unusual high frequency of CYP2E1* c2 allele. Alcoholism: Clinical and Experimental Research. 2010; 34 (1):142-149.
2. Méndez C, Rey M. Characterization of polymorphisms of genes ADH2, ADH3, ALDH2 and CYP2E1 and relationship to the alcoholism in a Colombian population. Colombia Médica. 2015; 46 (4):176-182.
3. Husemoen LLN, Fenger M, Friedrich N, Tolstrup JS, Beenfeldt Fredriksen S, Linneberg A. The association of ADH and ALDH gene variants with alcohol drinking habits and cardiovascular disease risk factors. Alcoholism: Clinical and Experimental Research. 2008; 32 (11):1984-1991.
4. Ghosh S, Bankura B, Ghosh S, Saha ML, Pattanayak AK, Ghatak S, *et al*. Polymorphisms in ADH1B and ALDH2 genes associated with the increased risk of gastric cancer in West Bengal, India. BMC cancer. 2017; 17:1-11.
5. Vatansever S, Tekin F, Salman E, Altintoprak E, Coskunol H, Akarca US. Genetic polymorphisms of ADH1B, ADH1C and ALDH2 in Turkish alcoholics: lack of association with alcoholism and alcoholic cirrhosis. Bosnian Journal of Basic Medical Sciences. 2015; 15 (2):37.
6. Hoang YTT. Single Nucleotide Polymorphisms of ADH1B, ADH1C and ALDH2 Genes in 235 People Living in Thai Nguyen Province of Vietnam. Asian Pacific Journal of Cancer Prevention. 2018; 23 (12):4243-4251.
7. Im PK, Yang L, Kartsonaki C, Chen Y, Guo Y, Du H, *et al*. Alcohol metabolism genes and risks of site-specific cancers in Chinese adults: An 11- year prospective study. International journal of cancer. 2022; 150 (10):1627-1639.
8. Harrison PW, Amode MR, Austine-Orimoloye O, Azov AG, Barba M, Barnes I, *et al*. Ensembl 2024. Nucleic acids research. 2024; 52 (D1):D891-D899.
9. Ensembl release 112. Available at: https://asia.ensembl.org/ [Accessed October 11, 2024 2024].
10. Park SK, Park C-S, Lee H-S, Park KS, Park BL, Cheong HS, *et al*. Functional polymorphism in aldehyde dehydrogenase-2 gene associated with risk of tuberculosis. BMC Medical Genetics. 2014; 15 (1):1-5.
11. Nakano Y, Ochi H, Onohara Y, Sairaku A, Tokuyama T, Matsumura H, *et al*. Genetic variations of aldehyde dehydrogenase 2 and alcohol dehydrogenase 1B are associated with the etiology of atrial fibrillation in Japanese. Journal of Biomedical Science. 2016; 23 (1):1-9.
12. Chung C-S, Lee Y-C, Liou J-M, Wang C-P, Ko J-Y, Lee J-M, *et al*. Tag single nucleotide polymorphisms of alcohol-metabolizing enzymes modify the risk of upper aerodigestive tract cancers: HapMap database analysis. Diseases of the Esophagus. 2014; 27 (5):493-503.
13. Tsai ST, Wong TY, Ou CY, Fang SY, Chen KC, Hsiao JR, *et al*. The interplay between alcohol consumption, oral hygiene, ALDH2 and ADH1B in the risk of head and neck cancer. International journal of cancer. 2014; 135 (10):2424-2436.
14. Liu J, Yang H-I, Lee M-H, Jen C-L, Hu H-H, Lu S-N, *et al*. Alcohol Drinking Mediates the Association between Polymorphisms ofADH1BandALDH2and Hepatitis B–Related Hepatocellular Carcinoma.
15. Wu LSH, Lee CS, Weng TY, Wang KHT, Cheng ATA. Association study of gene polymorphisms in GABA, serotonin, dopamine, and alcohol metabolism pathways with alcohol dependence in Taiwanese han men. Alcoholism: clinical and experimental research. 2016; 40 (2):284-290.
16. Huang C-C, Hsiao J-R, Lee W-T, Lee Y-C, Ou C-Y, Chang C-C, *et al*. Investigating the association between alcohol and risk of head and neck cancer in Taiwan. Scientific reports. 2017; 7 (1):1-13.
17. Chien H-T, Young C-K, Chen T-P, Liao C-T, Wang H-M, Cheng S-D, *et al*. Alcohol-metabolizing Enzymes' Gene Polymorphisms and Susceptibility to Multiple Head and Neck CancersALDH2, ADH1B, and ADH1C SNPs in Multiple Primaries. Cancer Prevention Research. 2019; 12 (4):247-254.
18. Huang Y-H, Chang K-H, Lee Y-S, Chen C-M, Chen Y-C. Association of alcohol dehydrogenase and aldehyde dehydrogenase polymorphism with spontaneous deep intracerebral haemorrhage in the Taiwan population. Scientific Reports. 2020; 10 (1):3641.
19. Wu Y-Y, Lee Y-S, Liu Y-L, Hsu W-C, Ho W-M, Huang Y-H, *et al*. Association study of alcohol dehydrogenase and aldehyde dehydrogenase polymorphism with Alzheimer disease in the Taiwanese population. Frontiers in Neuroscience. 2021; 15:625885.
20. Yang P-W, Lin M-C, Huang P-M, Wang C-P, Chen T-C, Chen C-N, *et al*. Risk factors and genetic biomarkers of multiple primary cancers in esophageal cancer patients. Frontiers in Oncology. 2021; 10:585621.

21. Lin C-H, Nfor ON, Ho C-C, Hsu S-Y, Tantoh DM, Liaw Y-C, *et al*. Association of ADH1B polymorphism and alcohol consumption with increased risk of intracerebral hemorrhagic stroke. Journal of Translational Medicine. 2021; 19 (1):1-8.
22. Shan Y-S, Chen L-T, Wu C-H, Chang Y-F, Lee C-T, Chiang N-J, *et al*. No association between alcohol consumption and pancreatic cancer even among individuals genetically susceptible to the carcinogenicity of alcohol. Scientific reports. 2021; 11 (1):14567.