



Functional variants in *ADH1B* and *ALDH2* are non-additively associated with all-cause mortality in Japanese population

Saori Sakaue^{1,2,3} · Masato Akiyama^{1,4} · Makoto Hirata⁵ ⁵ · Koichi Matsuda⁶ · Yoshinori Murakami⁷ · Michiaki Kubo⁸ · Yoichiro Kamatani^{1,9} ^{1,9} · Yukinori Okada^{2,10} ^{2,10}

Received: 2 February 2019 / Revised: 1 August 2019 / Accepted: 10 September 2019 / Published online: 26 September 2019
© The Author(s) 2019. This article is published with open access

Abstract

The functional variants involved in alcohol metabolism, the A allele of rs1229984:A > G in *ADH1B* and the A allele of rs671:G > A in *ALDH2*, are specifically prevalent among East Asian population. They are shown to be under recent positive selection, but the reasons for the selection are unknown. To test whether these positively selected variants have beneficial effects on survival in modern population, we performed the survival analyses using the large-scale Japanese cohort ($n = 135,974$) with genotype and follow-up survival data. The rs671-A allele was significantly associated with the better survival in the additive model (HR for mortality = 0.960, $P = 1.7 \times 10^{-5}$), and the rs1229984-A had both additive and non-additive effects (HR = 0.962, $P = 0.0016$ and HR = 0.958, $P = 0.0066$, respectively), which was consistent with the positive selection. The favorable effects of these alleles on survival were independent of the habit of alcohol consumption itself. The heterogenous combinatory effect between rs1229984 and rs671 genotype was also observed (HRs for AA genotype at rs671 were 1.03, 0.80, and 0.90 for GG, GA, and AA genotype at rs1229984, respectively), supposedly reflecting the synergistic effects on survival.

Introduction

Functional variants in *ADH1B* and *ALDH2*, which are specifically prevalent among East Asian population,

substantially alter enzymatic activity involved in alcohol metabolism and make the population less tolerant to alcohol consumption [1]. The A allele of rs1229984:A > G (hg19 chr4:g.100239319 A > G; NM_000668.5:c.143 A > G [NP_000659.2:p.(His48Arg)]) causes the rapid oxidation of ethanol to acetaldehyde by *ADH1B*, which increases an aversive reaction to alcohol, while the A allele of rs671:G > A (hg19 chr12:g.112241766 G > A; NM_000690.3:

Supplementary information The online version of this article (<https://doi.org/10.1038/s41431-019-0518-y>) contains supplementary material, which is available to authorized users.

✉ Yoichiro Kamatani
yoichiro.kamatani@riken.jp

✉ Yukinori Okada
yokada@sg.med.osaka-u.ac.jp

¹ Laboratory for Statistical Analysis, RIKEN Center for Integrative Medical Sciences, Yokohama 230-0045, Japan

² Department of Statistical Genetics, Osaka University Graduate School of Medicine, Suita 565-0871, Japan

³ Department of Allergy and Rheumatology, Graduate School of Medicine, the University of Tokyo, Tokyo 113-8655, Japan

⁴ Department of Ophthalmology, Graduate School of Medical Sciences, Kyushu University, Fukuoka 812-8582, Japan

⁵ Laboratory of Genome Technology, Institute of Medical Science, The University of Tokyo, Tokyo 108-8639, Japan

⁶ Department of Computational Biology and Medical Sciences, Graduate school of Frontier Sciences, The University of Tokyo, Tokyo 108-8639, Japan

⁷ Division of Molecular Pathology, the Institute of Medical Sciences, the University of Tokyo, Tokyo 108-8639, Japan

⁸ RIKEN Center for Integrative Medical Sciences, Yokohama 230-0045, Japan

⁹ Laboratory of Complex Trait Genomics, Department of Computational Biology and Medical Sciences, Graduate School of Frontier Sciences, the University of Tokyo, Tokyo 108-8639, Japan

¹⁰ Laboratory of Statistical Immunology, Immunology Frontier Research Center (WPI-IFReC), Osaka University, Suita 565-0871, Japan

c.1510 G > A [NP_000681.2:p.(Glu504Lys)]) causes the functional deficiency of ALDH2, which slows the metabolism of acetaldehyde [2]. These alleles are common among East Asians (frequency = 0.738 and 0.255), but are rare or at low frequency in other populations (frequency = 0.047 and 0.0003, respectively [3]). These variants are also known to be highly pleiotropic, and associated with many complex human traits. Rs1229984 is associated with body mass index [4] and pulse pressure [5], while rs671 is associated with diseases such as coronary artery disease [6] and intracranial aneurysm [7], as well as affecting quantitative traits such as body mass index [8], uric acid [9], and triglycerides [10]. Intriguingly, although they are the risk alleles for some of the modern diseases, we have shown that they are under strong recent positive selection among Japanese by analyzing the whole-genome sequencing data [11]. The reasons for their positive selection in East Asians are unknown.

In order to test whether these positively selected variants have beneficial effects on survival, we performed the survival analyses using the large-scale Japanese cohort ($n = 135,974$) with genotype and follow-up survival data. We then performed further statistical analyses to estimate the non-additive and combinatory effects of these variants on survival.

Materials and methods

Clinical information, genotype, and follow-up survival data were obtained from BioBank Japan [12, 13], which collected DNA and serum samples from ~200,000 participants. We obtained informed consent from all the participants, following the protocols approved by ethics committees of RIKEN Center for Integrative Medical Sciences and the Institute of Medical Sciences, the University of Tokyo. The detailed information of participants is summarized in Supplementary Table 1. Genotyping and Quality control of participants are described elsewhere [14]. We analyzed the concordance of genotyping between the SNP array and the whole-genome sequencing (WGS) data ($n = 1638$). The genotype data in this study is deposited on the Japanese Genotype-phenotype Archive affiliated to the DDBJ (DNA Data Bank of Japan), via National Bioscience Database Center (NBDC), Japan. The data is accessible with the accession IDs hum0014 and JGAS00000000114 at <https://ddbj.nig.ac.jp/jga/viewer/view/study/JGAS00000000114>.

The survival analyses of rs1229984 and rs671 for all-cause mortality were performed by Cox proportional-hazard models, adjusted for age, sex, the disease status, the habit of cigarette smoking and alcohol consumption, and 10 principal components. Primary analysis was performed by assuming an additive model. Next, to estimate the non-

additive allelic effect, we additionally included a dominance term which is one if the genotype is heterozygous and zero otherwise [15]. Finally, to investigate the combinatory effect between rs1229984 and rs671 on survival, we obtained hazard ratios (HR) for each of the combination of genotypes at rs1229984 and rs671. We compared them to the expected HRs by assuming that they should be the products of the two HRs of each allele under the null hypothesis where there are no gene \times gene combinatory effects. The deviation in the fold change was obtained by dividing the observed HR by the expected. All the survival analyses were performed using R software, version 3.3.0.

Results

We first genotyped rs1229984 and rs671, two of the variants under positive selection [11]. As both rs1229984 and rs671 were significantly deviated from the QC threshold of Hardy–Weinberg equilibrium ($P_{\text{HWE}} < 1.0 \times 10^{-6}$), we analyzed the concordance of genotyping between the SNP array and the whole-genome sequencing (WGS) data ($n = 1638$). We confirmed the high concordance between the array-based genotype and sequenced genotype (97.6 and 100%, respectively), indicating that the observed deviation from HWE was not caused by genotyping error but by heterogeneity in allele frequency spectra among the regions of Japan.

We then associated the genotype of these variants with the all-cause mortality ($n = 135,974$) [16]. The median follow-up period was 8.08 years, and the number of deaths during the follow-up was 31,403. We observed that in both of the variants, the alleles which make their carriers less tolerant to alcohol showed significantly favorable effects on survival. When we assume an additive effect, the rs671-A allele was strongly associated with the all-cause mortality (HR = 0.960 [95% Confidence Interval: 0.942–0.978] and $P = 1.7 \times 10^{-5}$), while the effect of the rs1229984-A was modest (HR = 0.983 [0.965–1.001] and $P = 0.067$, Supplementary Table 2). We note that the effects of these alleles on survival were independent of the habit of alcohol consumption itself, as we regressed out the status of alcohol consumption in constructing the Cox proportional-hazard models.

Intriguingly, we found that the association of the heterozygous genotype GA at rs1229984 was comparable to that of the homozygous genotype AA (HR = 0.929 and 0.922) (Fig. 1a and Table 1). The GA and the AA genotype at rs1229984 were both reported to produce 40-fold faster ethanol oxidation by ADH1B than the GG genotype [17]. Thus, in order to clarify the suggested non-linear effect of rs1229984 on survival outcome, we incorporated both the additive and non-additive term into the model [15]. As

Fig. 1 The survival analysis of rs1229984 and rs671 in Japanese population. **(a)** The standardized and adjusted survival curves for all-cause mortality according to the genotype of rs1229984 and rs671. The blue, orange, and red line indicates the survival curve of participants with the GG, GA and AA genotype, respectively. The GG genotypes were adopted as baselines. **(b)** The observed (left) and expected (middle) hazard ratios for each of the combination of genotypes at rs1229984 and rs671. The expected hazard ratios are calculated by assuming that they should be the products of the two hazard ratios of each allele under the null hypothesis where there are no gene \times gene combinatory effects. The deviations of the observed hazard ratio from the expected are shown in fold change (right)

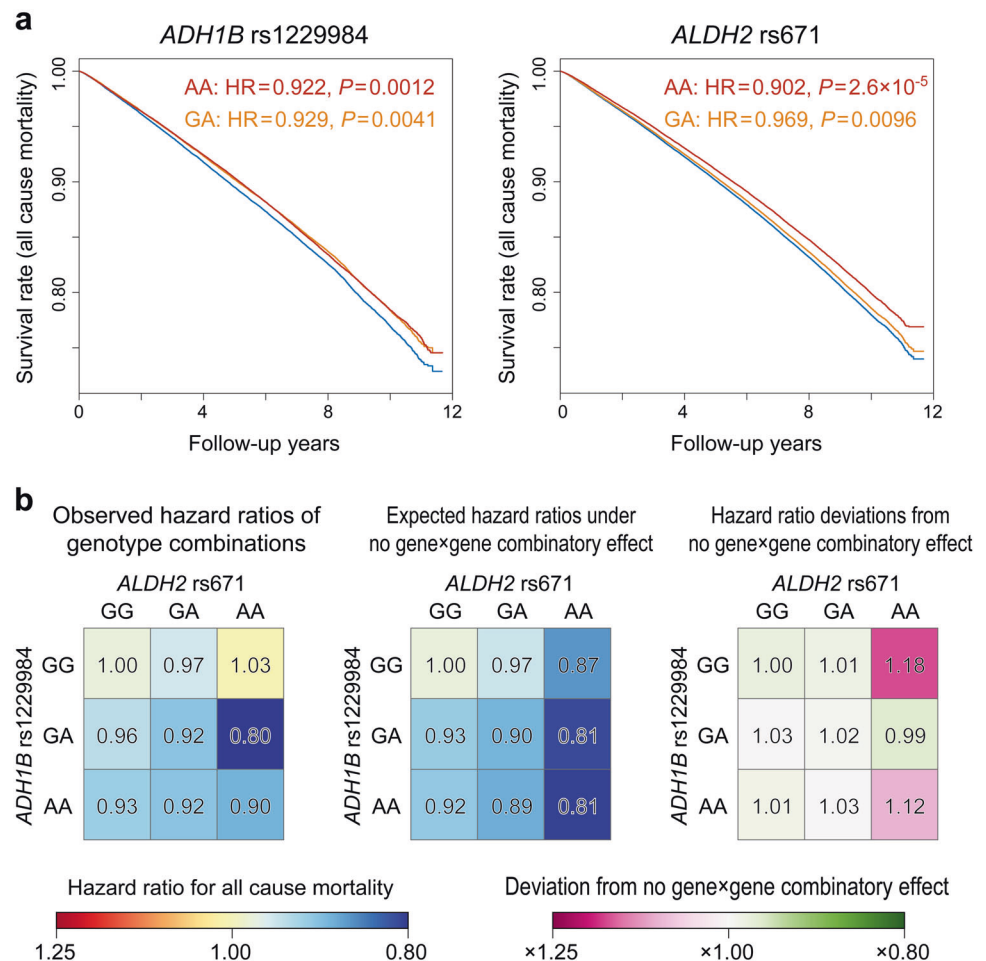


Table 1 The association of the functional variants within *ADH1B* and *ALDH2* with the all-cause mortality

SNP	Chr	Position (hg19)	Gene	Genotype	Freq.	Hazard ratio (95% CI)	<i>P</i> -value
rs1229984	4	100,239,319	<i>ADH1B</i>	GG	0.057	Reference	–
				GA	0.354	0.929 (0.884–0.977)	0.0041
				AA	0.589	0.922 (0.879–0.968)	0.0012
rs671	12	112,241,766	<i>ALDH2</i>	GG	0.568	Reference	–
				GA	0.367	0.969 (0.946–0.992)	0.0099
				AA	0.065	0.902 (0.860–0.947)	2.6×10^{-5}

Freq. frequency of the genotype

suggested in the genotype-level association test, we observed a significant additive and non-additive allelic effects of the rs1229984-A on survival ($P_{\text{Additive}} = 0.0016$ and $P_{\text{NonAdditive}} = 0.0066$), while the rs671-A only had an additive allelic effect ($P_{\text{Additive}} = 3.4 \times 10^{-5}$ and $P_{\text{NonAdditive}} = 0.20$, Table 2). Likelihood ratio tests confirmed that the incorporation of non-additive term significantly improved the model fit in rs1229984 ($P_{\text{ANOVA}} = 0.0068$), but not in rs671 ($P_{\text{ANOVA}} = 0.20$).

We have previously shown that rs671 and rs1229984 are in trans-chromosomal linkage disequilibrium, reflecting the synergistic natural selection pressure [18]. Thus, we finally

Table 2 The additive and non-additive allelic effects of rs1229984 and rs671 on the all-cause mortality

SNP allele	Mode	Hazard ratio (95% CI)	<i>P</i> -value
rs1229984 (A)	Additive	0.962 (0.939–0.985)	0.0016
rs1229984 (A)	Non-additive	0.958 (0.929–0.988)	0.0066
rs671 (A)	Additive	0.951 (0.929–0.974)	3.4×10^{-5}
rs671 (A)	Non-additive	1.020 (0.990–1.051)	0.20

tested if there existed gene \times gene combinatory effects on survival between rs671 and rs1229984. The combination of the AA genotype at rs671 and the GA genotype at

rs1229984 showed the most favorable effect (HR = 0.801 [0.719–0.892] and $P = 5.5 \times 10^{-5}$, Fig. 1b and Supplementary Table 3). Interestingly, we could observe an upward deviation of the observed HR in the combination of the AA genotype at rs671 and the GG or AA genotype at rs1229984 if we assume that there are no gene \times gene combinatory effects (fold change = 1.18 and 1.12, respectively). On the other hand, the incorporation of interactive terms between rs1229984 and rs671 failed to capture the significant effects, probably due to the co-linearity between the terms or lack of statistical power. Taken together, people with the AA genotype at rs671 had the favorable survival outcome than those with the GG or GA genotype as discussed above, and this favorable effect was mostly driven by the subset of people with the GA genotype at rs1229984. By careful examination of the effects of each genotype on survival outcome, we showed a rare example of trans-chromosomal combinatory effects, which we consider to be also essential in analyzing other pleiotropic associations of these loci.

Discussion

The common functional variants at *ADH1B* and *ALDH2* are specifically prevalent in East Asians, and thus their phenotypic landscape has been understudied. Here we described the association of rs1229984 and rs671 with the all-cause mortality by leveraging the large-scale biobank in Japanese. In both of the variants, the alleles which make their carriers less tolerant to alcohol were shown to have beneficial effects on survival. These observations are consistent with the fact that they are under recent positive selection and in different frequency spectrum among East Asians. They showed beneficial effects on survival even though they are also associated with causing various modern diseases [6, 7]. To have further insights into the model of their effects on survival, we have shown that the rs1229984-A had both the additive and non-additive allelic effect, and that there existed a combinatory effect of rs1229984 and rs671.

One of the alcohol-metabolism related loci, *BRAP-ALDH2*, was also reported to be significantly associated with the length of lifespan in European population in the large-scale study on UK Biobank [19]. The reasons for the positive selection of alcohol-related genes are indecisive, and one of the speculations is that a higher concentration of acetaldehyde was advantageous for parasitic infections endemic in East Asia, past or present [20]. Future functional studies will be awaited to reveal the reasons why these loci have been positively selected in modern Asian populations and hold the survival benefit, and how it is connected with the increasing prevalence of late-onset complex diseases in human populations. Both of the positively selected genes,

ADH1B and *ALDH2*, are not only substantially explaining the heritability of the alcohol consumption, but also associated with many complex human traits (i.e. pleiotropy) [4–10]. Our results suggested that the favorable effects of the rs1229984-A and rs671-A on survival were independent of alcohol consumption itself, because the associations were conditioned on the habit of alcohol consumption. Their biological mechanism on survival remains elusive, which warrants further studies incorporating the comprehensive phenotypic associations.

In conclusion, the survival analysis on *ADH1B* and *ALDH2* revealed that the functional variants within these loci had favorable effects on survival in Japanese population, and that their non-additive and combinatory allelic effects should be taken into consideration.

Acknowledgements This research was supported by the Tailor-Made Medical Treatment program (the BioBank Japan Project) of the Ministry of Education, Culture, Sports, Science, and Technology (MEXT), the Japan Agency for Medical Research and Development (AMED), Bioinformatics Initiative of Osaka University Graduate School of Medicine, and Integrated Frontier Research for Medical Science Division, Institute for Open and Transdisciplinary Research Initiatives, Osaka University. Y.O. was supported by the Japan Society for the Promotion of Science (JSPS) KAKENHI (15H05911, 19H01021), AMED (19gm6010001h0004, 19ek0410041h0003, 19km0405211h0001, and 19ek0109413h0001).

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

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/>.

References

1. Takeshita T, Mao XQ, Morimoto K. The contribution of polymorphism in the alcohol dehydrogenase beta subunit to alcohol sensitivity in a Japanese population. *Hum Genet.* 1996;97:409–13.
2. Edenberg HJ. The genetics of alcohol metabolism: role of alcohol dehydrogenase and aldehyde dehydrogenase variants. *Alcohol Res Health.* 2007;30:5–13.
3. Lek M, Karczewski KJ, Minikel EV, Samocha KE, Banks E, Fennell T et al. Analysis of protein-coding genetic variation in

- 60,706 humans. *Nature* 2016;536. <https://doi.org/10.1038/nature19057>.
4. Akiyama M, Okada Y, Kanai M, Takahashi A, Momozawa Y, Ikeda M, et al. Genome-wide association study identifies 112 new loci for body mass index in the Japanese population. *Nat Genet*. 2017;49:1458–67.
 5. Warren HR, Evangelou E, Cabrera CP, Gao H, Ren M, Mifsud B, et al. Genome-wide association analysis identifies novel blood pressure loci and offers biological insights into cardiovascular risk. *Nat Genet*. 2017;49:403–15.
 6. Takeuchi F, Yokota M, Yamamoto K, Nakashima E, Katsuya T, Asano H, et al. Genome-wide association study of coronary artery disease in the Japanese. *Eur J Hum Genet*. 2012;20:333–40.
 7. Low S-K, Takahashi A, Cha P-C, Zembutsu H, Kamatani N, Kubo M, et al. Genome-wide association study for intracranial aneurysm in the Japanese population identifies three candidate susceptible loci and a functional genetic variant at EDNRA. *Hum Mol Genet*. 2012;21:2102–10.
 8. Wen W, Zheng W, Okada Y, Takeuchi F, Tabara Y, Hwang JY, et al. Meta-analysis of genome-wide association studies in East Asian-ancestry populations identifies four new loci for body mass index. *Hum Mol Genet*. 2014;23:5492–504.
 9. Zhang D, Yang M, Zhou D, Li Z, Cai L, Bao Y, et al. The polymorphism rs671 at ALDH2 associated with serum uric acid levels in Chinese Han males: a genome-wide association study. *Gene*. 2018;651:62–69.
 10. Tan A, Sun J, Xia N, Qin X, Hu Y, Zhang S, et al. A genome-wide association and gene-environment interaction study for serum triglycerides levels in a healthy Chinese male population. *Hum Mol Genet*. 2012;21:1658–64.
 11. Okada Y, Momozawa Y, Sakaue S, Kanai M, Ishigaki K, Akiyama M, et al. Deep whole-genome sequencing reveals recent selection signatures linked to evolution and disease risk of Japanese. *Nat Commun*. 2018;9:1631.
 12. Hirata M, Kamatani Y, Nagai A, Kiyohara Y, Ninomiya T, Tamakoshi A, et al. Cross-sectional analysis of BioBank Japan clinical data: a large cohort of 200,000 patients with 47 common diseases. *J Epidemiol*. 2017;27:S9–S21.
 13. Nagai A, Hirata M, Kamatani Y, Muto K, Matsuda K, Kiyohara Y, et al. Overview of the BioBank Japan Project: study design and profile. *J Epidemiol*. 2017;27:S2–S8.
 14. Kanai M, Akiyama M, Takahashi A, Matoba N, Momozawa Y, Ikeda M, et al. Genetic analysis of quantitative traits in the Japanese population links cell types to complex human diseases. *Nat Genet*. 2018;50:390–400.
 15. Lenz TL, Deutsch AJ, Han B, Hu X, Okada Y, Eyre S, et al. Widespread non-additive and interaction effects within HLA loci modulate the risk of autoimmune diseases. *Nat Genet*. 2015;47:1085–90.
 16. Hirata M, Nagai A, Kamatani Y, Ninomiya T, Tamakoshi A, Yamagata Z, et al. Overview of BioBank Japan follow-up data in 32 diseases. *J Epidemiol*. 2017;27:S22–S28.
 17. Bosron WF, Crabb DW, Ting-Kai Li. Relationship between kinetics of liver alcohol dehydrogenase and alcohol metabolism. *Pharm Biochem Behav*. 1983;18:223–7.
 18. Okada Y. eLD: entropy-based linkage disequilibrium index between multiallelic sites. *Hum Genome Var*. 2018;5:29.
 19. Timmers PR, Mounier N, Lall K, Fischer K, Ning Z, Feng X et al. Genomics of 1 million parent lifespans implicates novel pathways and common diseases and distinguishes survival chances. *Elife* 2019;8. <https://doi.org/10.7554/eLife.39856>.
 20. Oota H, Pakstis AJ, Bonne-Tamir B, Goldman D, Grigorenko E, Kajuna SL, et al. The evolution and population genetics of the ALDH2 locus: random genetic drift, selection, and low levels of recombination. *Ann Hum Genet*. 2004;68:93–109.