

Multiple *ADH* genes are associated with upper aerodigestive cancers

Mia Hashibe¹, James D McKay¹, Maria Paula Curado^{1,2}, Jose Carlos Oliveira², Sergio Koifman³, Rosalina Koifman³, David Zaridze⁴, Oxana Shangina⁴, Victor Wunsch-Filho⁵, Jose Eluf-Neto⁵, Jose Eduardo Levi⁵, Elena Matos⁶, Pagona Lagiou⁷, Areti Lagiou⁷, Simone Benhamou⁸, Christine Bouchardy⁹, Neonilia Szeszenia-Dabrowska¹⁰, Ana Menezes¹¹, Marinel Mór Dall'Agnol¹¹, Franco Merletti¹², Lorenzo Richiardi¹², Leticia Fernandez¹³, Juan Lence¹³, Renato Talamini¹⁴, Luigi Barzan¹⁵, Dana Mates¹⁶, Ioan Nicolae Mates¹⁷, Kristina Kjaerheim¹⁸, Gary J Macfarlane¹⁹, Tatiana V Macfarlane¹⁹, Lorenzo Simonato²⁰, Cristina Canova²⁰, Ivana Holcátová²¹, Antonio Agudo²², Xavier Castellsagué²², Ray Lowry²³, Vladimir Janout²⁴, Helena Kollarova²⁴, David I Conway²⁵, Patricia A McKinney^{26,27}, Ariana Znaor²⁸, Eleonora Fabianova²⁹, Vladimir Bencko²¹, Jolanta Lissowska³⁰, Amelie Chabrier¹, Rayjean J Hung^{1,31}, Valerie Gaborieau¹, Paolo Boffetta¹ & Paul Brennan¹

Alcohol is an important risk factor for upper aerodigestive cancers and is principally metabolized by alcohol dehydrogenase (ADH) enzymes. We have investigated six *ADH* genetic variants in over 3,800 aerodigestive cancer cases and 5,200 controls from three individual studies. Gene variants rs1229984 (*ADH1B*) and rs1573496 (*ADH7*) were significantly protective against aerodigestive cancer in each individual study and overall ($P = 10^{-10}$ and 10^{-9} , respectively). These effects became more apparent with increasing alcohol consumption (P for trend = 0.0002 and 0.065, respectively). Both gene effects were independent of each other, implying that multiple *ADH* genes may be involved in upper aerodigestive cancer etiology.

The alcohol dehydrogenase (ADH) pathway includes seven distinct *ADH* genes, a key candidate gene group for aerodigestive cancers^{1–3}.

Studies of aerodigestive cancer in populations of European origin have focused on *ADH1C* with little evidence of any effect⁴. We previously reported an association for *ADH1B* R48H (rs1229984) in a central European (CE) population⁵ and now consider the effect of this and five other *ADH* variants in an expanded study comprising 809 aerodigestive cancer cases and 2,586 controls from the CE study as well as a further 3,067 aerodigestive cancer cases and 2,692 controls from two other studies in Europe (ARCAGE study) and Latin America (LA study) (total of 3,876 cases and 5,278 controls). All three studies were coordinated by the International Agency for Research on Cancer (IARC) and followed a similar protocol (**Supplementary Methods** online). Of the 3,876 cases, 1,790 were cancers of the oral cavity or pharynx, 1,659 were cancers of the hypopharynx or larynx and 427 were cancers of the esophagus (**Supplementary Table 1** online). Cases with a histology other than squamous cell were excluded.

The HapMap Consortium has genotyped 163 SNPs in the vicinity of the *ADH* gene cluster with a minor allele frequency (MAF) of 4% or more in the CEPH Utah (CEU) population⁶. Inspection of the linkage disequilibrium (LD) pattern across this region indicates that *ADH1A*, *ADH1B*, *ADH1C*, *ADH4*, *ADH5* and *ADH6* are relatively highly correlated, whereas *ADH7* showed little correlation with the remaining six (**Supplementary Fig. 1a** online). From all verified missense SNPs in the seven *ADH* genes found in both the NCBI SNP and SNP500 databases⁷, we selected eight that had a MAF > 4% in the CEU population. Three missense SNPs in *ADH4* (rs1126671, rs1126673 and rs1042364) were in strong LD, and thus were genotyped by the highly correlated tagging SNP rs1984362 ($r^2 > 0.89$). In total, we genotyped six genetic variants (five missense SNPs and one tagging SNP) in all three studies (**Table 1** and **Supplementary Table 2** online).

In the pooled analysis on all 3,876 cases and 5,278 controls, four variants reported a significant association (**Supplementary Table 3** online). The most prominent was with rs1229984 (in *ADH1B*; OR for codominant model = 0.59 (95% CI = 0.50–0.69); P under codominant model = 8×10^{-10}). This variant was significant in each of the three individual studies (CE: $P = 5 \times 10^{-5}$; ARCAGE: $P = 1 \times 10^{-4}$; LA: $P = 0.002$). A second strongly significant finding in the pooled analysis was with rs1573496 (in *ADH7*; OR = 0.69 (0.61–0.78); $P = 3 \times 10^{-9}$), a new potential susceptibility gene for this cancer. This variant was also significant in each of the three individual studies

¹International Agency for Research on Cancer, Lyon, France. ²Hospital Araujo Jorge, Goiania, and Populational Cancer Register of Goiania, Brazil. ³Escola Nacional de Saude Publica, Rio de Janeiro, Brazil. ⁴Cancer Research Centre, Moscow, Russia. ⁵Universidade de Sao Paulo, Sao Paulo, Brazil. ⁶Institute of Oncology Angel H. Roffo, University of Buenos Aires, Argentina. ⁷University of Athens School of Medicine, Athens, Greece. ⁸Institut National de la Santé et de la Recherche Médicale U794, Evry, France. ⁹Geneva Cancer Registry, Geneva, Switzerland. ¹⁰Institute of Occupational Medicine, Lodz, Poland. ¹¹Universidade Federal de Pelotas, Pelotas, Brazil. ¹²Unit of Cancer Epidemiology, CeRMS and University of Turin, Turin, Italy. ¹³Institute of Oncology and Radiobiology, Havana, Cuba. ¹⁴Aviano Cancer Centre, Aviano, Italy. ¹⁵General Hospital of Pordenone, Pordenone, Italy. ¹⁶Institute of Public Health, Bucharest, Romania. ¹⁷University of Medicine and Pharmacy "Carol Davila", Bucharest, Romania. ¹⁸Cancer Registry of Norway, Oslo, Norway. ¹⁹University of Aberdeen School of Medicine, Aberdeen, UK. ²⁰University of Padua, Padova, Italy. ²¹Institute of Hygiene and Epidemiology, Prague, Czech Republic. ²²Institut Català d'Oncologia, Barcelona, Spain. ²³University of Newcastle Dental School, Newcastle, UK. ²⁴Palacky University, Olomouc, Czech Republic. ²⁵University of Glasgow Dental School, Glasgow, Scotland. ²⁶University of Leeds Centre for Epidemiology and Biostatistics, Leeds, UK. ²⁷NHS National Services Scotland, Information Services Division, Edinburgh, Scotland. ²⁸Croatian National Cancer Registry, Zagreb, Croatia. ²⁹Specialized State Health Institute, Banská Bystrica, Slovakia. ³⁰The Maria Skłodowska-Curie Memorial Cancer Center and Institute of Oncology, Warsaw, Poland. ³¹School of Public Health, University of California at Berkeley, Berkeley, California 94720, USA. Correspondence should be addressed to P.B. (Brennan@iarc.fr).

Table 1 Minor allele frequency (%) of six *ADH* candidate variants within each recruitment center

Gene	<i>ADH1B</i> G/A		<i>ADH1B</i> A/T		<i>ADH1C</i> G/A		<i>ADH1C</i> A/G		<i>ADH4</i> C/T		<i>ADH7</i> G/C			
rs number	rs1229984		rs6413413		rs1693482		rs698		rs1984362		rs1573496			
Coding change	R48H		T60S		R272Q		I350V		Tagging		G92A			
	Number of cases and controls													
	Ca	Co	Ca	Co	Ca	Co	Ca	Co	Ca	Co	Ca	Co		
Central Europe														
Overall	809	2,586	3.44	5.94	0.56	0.36	43.69	42.16	43.30	41.12	29.72	30.16	7.67	12.77
Moscow	365	797	4.03	5.41	0.69	0.52	46.91	44.69	47.47	46.02	31.22	30.82	7.12	12.42
Lodz	204	804	3.19	5.79	0.49	0.39	44.74	41.74	45.05	42.50	31.19	30.56	8.50	13.60
Bucharest	142	178	4.35	9.66	0.71	0.29	32.09	29.36	32.37	30.18	24.64	30.70	8.03	11.11
Olomouc	58	614	0.86	5.19	0.00	0.18	41.96	40.19	42.86	41.15	23.68	29.23	6.25	12.76
Banska Bystrica	40	193	0.00	7.59	0.00	0.27	43.75	37.50	43.75	38.25	35.53	28.00	9.21	12.30
ARCAGE														
Overall	1,356	1,407	3.74	6.55	1.06	1.05	41.45	37.90	41.61	37.66	31.71	29.34	8.10	10.67
Paris	215	128	2.40	4.40	1.19	1.56	41.51	40.23	41.35	40.00	26.67	25.20	6.76	6.75
Prague	100	124	3.80	8.70	1.05	1.27	53.80	36.55	54.76	37.38	33.85	36.07	9.38	9.24
Athens	187	160	11.41	16.01	0.27	1.28	33.33	28.57	32.97	27.63	32.88	32.91	11.62	10.78
Aviano	138	140	3.26	6.72	0.75	1.80	40.46	31.02	41.54	31.85	35.71	29.50	7.14	10.22
Padova	110	118	2.29	4.70	1.89	0.43	36.79	32.91	37.62	32.89	32.11	29.74	6.25	13.56
Torino	144	178	7.09	11.78	1.41	0.28	39.36	27.62	38.32	27.51	32.75	28.74	10.71	14.00
Oslo	109	135	0.48	0.38	0.92	0.37	45.24	49.62	44.29	49.62	31.60	24.63	5.19	10.74
Edinburgh	47	51	0.00	3.00	3.19	1.00	54.35	52.94	54.26	52.94	34.04	27.00	4.35	7.84
Manchester	127	156	0.83	1.94	1.19	1.30	46.64	45.21	46.34	46.47	27.64	27.81	8.20	10.26
Newcastle	61	95	0.00	2.20	0.00	1.60	45.76	44.21	44.17	45.11	31.36	24.21	8.20	7.53
Barcelona	77	80	1.30	6.96	0.66	1.25	34.46	36.25	35.06	35.44	40.54	32.69	6.67	11.88
Zagreb	41	42	1.22	4.76	1.28	0.00	43.75	47.62	43.75	47.44	23.75	38.75	8.75	15.00
Latin America														
Overall	1,711	1,285	3.91	6.89	0.56	0.39	29.90	27.06	29.56	26.11	22.23	21.68	5.72	8.09
Cuba	150	132	7.27	7.45	0.00	0.45	29.09	23.63	29.69	26.67	23.91	23.12	8.62	7.65
Bueno Aires	292	203	4.07	6.71	0.38	0.00	35.55	30.74	37.43	31.50	24.07	25.96	5.29	8.81
Goianna	414	205	3.59	5.99	0.29	0.30	30.38	26.67	30.29	29.11	21.27	20.51	5.87	7.42
Pelotas	169	222	5.07	7.33	0.00	0.48	32.09	29.30	36.07	30.68	20.65	20.83	7.14	9.66
Rio de Janeiro	389	208	3.62	6.70	1.00	0.82	26.54	19.74	26.45	20.10	23.65	20.54	5.25	7.25
Sao Paulo	297	315	2.66	7.17	1.06	0.33	26.88	26.43	26.77	26.79	20.28	21.03	4.58	7.77

Ca, cases; Co, controls.

(CE: $P = 1 \times 10^{-7}$; ARCAGE: $P = 0.015$; LA: $P = 0.008$). Significant effects were also observed for both *ADH1C* variants: rs1693482 (OR = 1.17 (1.09–1.26); $P = 2 \times 10^{-5}$) and rs698 (OR = 1.14 (1.06–1.23); $P = 3 \times 10^{-4}$). These two variants were highly correlated in all three studies ($D' > 0.99$, $r^2 > 0.97$, **Supplementary Fig. 1b**), so we considered only rs1693482 for further analysis. A moderate increase in risk for this variant was observed in each of the three individual studies (CE: $P = 0.02$; ARCAGE: $P = 0.002$; LA: $P = 0.04$).

Both rs1229984 (*ADH1B*) and rs1693482 (*ADH1C*) were in LD, with $D' > 0.75$ in all three studies (**Supplementary Fig. 1b**). In order to determine the independence of each SNP, we repeated analysis of rs1693482 (*ADH1C*) excluding rs1229984[A] (*ADH1B*) carriers and vice versa (**Supplementary Table 4a** online). This indicated that the effect of rs1229984 (*ADH1B*) was not influenced by rs1693482 (*ADH1C*) and that the effect observed with rs1693482 (*ADH1C*) could not be explained solely by LD with rs1229984 (*ADH1B*). In contrast, there was little LD between rs1573496 (*ADH7*) and rs1229984 (*ADH1B*), with $D' < 0.21$ in all three studies (**Supplementary Fig. 1b**). To illustrate their independence, we repeated analysis of rs1229984 (*ADH1B*) excluding carriers of rs1573496[C] (*ADH7*) and vice versa. As expected considering the lack of LD between these two SNPs, the risk estimates remained very similar to those based on the complete dataset (**Supplementary Table 4b**).

Subsequently, we focused on the effects of rs1229984[A] (*ADH1B*) and rs1573496[C] (*ADH7*) after stratifying by site of cancer, age,

alcohol consumption, tobacco smoking and study (**Fig. 1**). Significant heterogeneity ($P = 0.001$) for rs1229984 (*ADH1B*) was observed by cancer site with a reduction in risk of between two- and threefold for oral and pharyngeal cancer (OR = 0.45 (0.35–0.57)) and esophageal cancer (OR = 0.34 (0.20–0.56)), as opposed to a 30% decrease in risk for larynx cancer (OR = 0.71 (0.57–0.88)). We observed significant heterogeneity by site for rs1573496 (*ADH7*) ($P = 0.023$), with the most pronounced protective effect being for esophageal cancer (OR = 0.45 (0.32–0.64)). For both gene variants, there was an increasing protective effect with increasing alcohol consumption. For rs1229984 (*ADH1B*), we did not observe any effect among never drinkers (OR = 1.02 (0.66–1.56)), whereas we observed an over twofold effect among those who drank above the median level of alcohol in each study (OR = 0.42 (0.31–0.56); P for trend = 0.0002). Similarly, for rs1573496 (*ADH7*), no effect was detected among never drinkers and a 40% decrease in risk was observed for heavy drinkers (OR = 0.61 (0.50–0.75); P for trend = 0.065). The effect of rs1229984 (*ADH1B*) was consistent across the three studies, whereas the effect of rs1573496 (*ADH7*) varied from OR = 0.54 (0.43–0.69) in the CE study to 0.79 (0.64–0.98) for the ARCAGE study (P for heterogeneity = 0.08). Any potential heterogeneity seemed to be explained by a higher proportion of esophageal cancer cases in the CE study (21%) as opposed to the LA study (9%) and the ARCAGE study (8%). When esophageal cancer cases were excluded from this analysis, no between-study heterogeneity was apparent for rs1573496 (*ADH7*) (P for heterogeneity = 0.38).

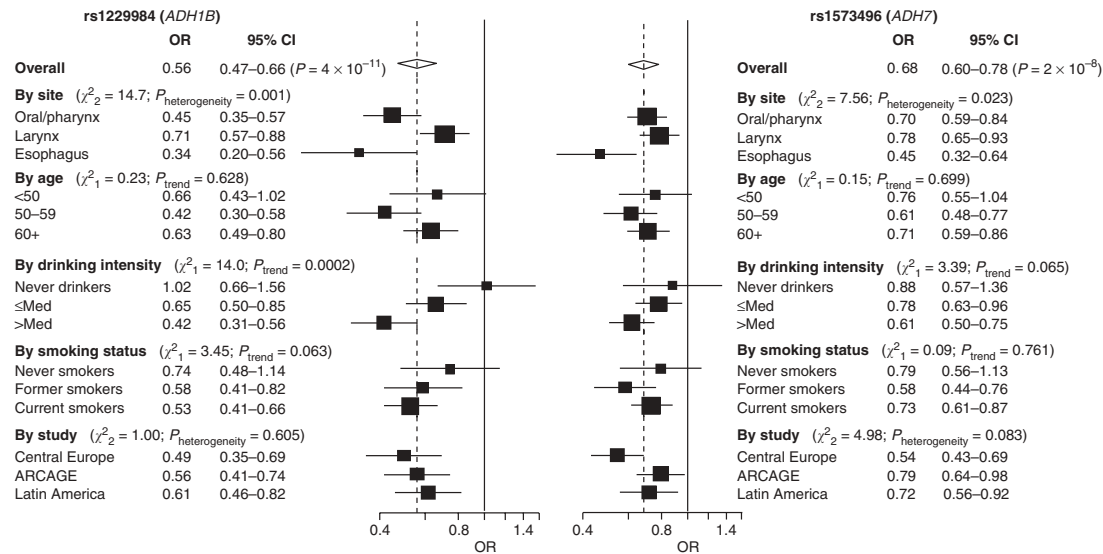


Figure 1 Odds ratio (OR) of upper aerodigestive cancer by rs1229984 (*ADH1B*) and rs1573496 (*ADH7*) genotypes. Rare allele carriers (dominant model) versus common allele homozygous genotype. ORs are standardized by age, sex, center, cumulative alcohol consumption and, when relevant, smoking. ORs and 95% CI are derived from fixed effects models.

We also analyzed the effect of carrying the rs1693482 (*ADH1C*) G/A or A/A variant after stratifying by site of cancer, age, alcohol consumption, tobacco smoking and study (**Supplementary Fig. 2** online). No notable heterogeneity by any of these factors was observed.

Finally, we assessed the combined effect of carrying either one or both of the rare gene variants from rs1229984 (*ADH1B*) and rs1573496 (*ADH7*). For those who possessed only rs1229984[A] (*ADH1B*), the OR for aerodigestive cancer was 0.56 (0.45–0.70), and for those who possessed only rs1573496[C] (*ADH7*), the OR was 0.70 (0.61–0.82), whereas for those who possessed both rare gene variants, the OR was 0.45 (0.34–0.60; P for trend in possessing zero, one or two variants = 10^{-16}).

These results provide strong evidence that both *ADH1B* and *ADH7* have an important association with susceptibility to aerodigestive cancer. The strong similarity of the results from different studies argues against population stratification or other biases. These effects seem to be relevant for all aerodigestive tract subsites, although they may be more prominent for esophageal cancer, and they both seem to be dependent on alcohol consumption—that is, among nondrinkers, the gene variants have little or no effect on disease risk, whereas among alcohol drinkers, the protective effect is more apparent at higher alcohol intake. Furthermore, neither gene variant seemed to be consistently associated with the amount of alcohol consumed in controls (data not shown), indicating that any protective effect from these gene–environment interactions is likely to be due to their role in changing the carcinogenic effect of alcohol beverages.

Whether we have studied the causal variants in these two genes or whether our associations are secondary to other causal variants is unknown. rs1229984 (*ADH1B*) G/A heterozygotes and A/A homozygotes are known to metabolize ethanol up to 100 times quicker than the common rs1229984 (*ADH1B*) G/G homozygote⁴, providing support that quick eradication of ethanol, and therefore lower local exposure, may be protective⁵. The potential role of *ADH7* is unclear, however, although variants in this gene may also influence ethanol metabolism⁸. Given the limited association between *ADH7* and other *ADH* genes (**Supplementary Fig. 1a**), it would be expected that the causal association for this variant resides in the region of *ADH7*.

In summary, our analysis of six *ADH* genetic variants in over 3,800 cases and 5,000 controls has identified two that are independently and strongly associated with aerodigestive cancers. Elucidation of these findings and further detailed characterization of this pathway in a large series of subjects seems warranted.

Note: Supplementary information is available on the Nature Genetics website.

ACKNOWLEDGMENTS

First and foremost we thank all participants of the CE, LA and ARCAGE studies and the staff in the collaborating centers. Subject recruitment was supported by four grants from the European Commission (IC15-CT96-0313, IC15-CT98-0322, IC18-CT97-0222 and QLK1-CT2001-00182) and one grant from the World Cancer Research Fund (WCRF 99A28). Genotyping and analysis was funded by a National Cancer Institute R01 grant (contract no. CA 092039-01A2). The funding agencies did not have any involvement in the study design, collection, analysis and interpretation of the data or in the writing of report and decision to submit the paper.

AUTHOR CONTRIBUTIONS

P. Boffetta, P. Brennan, M.P.C., J.C.O., S.K., R.K., V.W.-F., J.E.N., E.M., A.M., M.M.D., L.F. and J. Lence jointly designed the LA study and organized the recruitment of participants. P. Brennan, P. Boffetta, D.Z., O.S., N.S.-D., D.M., I.N.M., I.H., V.J., H.K., E.F. and J. Lissowska jointly designed the CE study and organized the recruitment of participants. P. Brennan, M.H., P.L., A.L., S.B., C.B., F.M., L.R., R.T., L.B., K.K., G.J.M., T.V.M., L.S., C.C., A.A., X.C., R.L., D.I.C., P.A.M., A.Z. and V.B. jointly designed the ARCAGE study and organized the recruitment of participants. J.D.M., M.H., A.C., R.J.H., V.G., J.E.L. and P. Brennan organized biological sample storage. J.D.M., M.H., V.G. and P. Brennan conducted SNP selection and the statistical analysis. M.H., J.D.M. and P. Brennan drafted the manuscript, and all coauthors contributed to the final draft.

Published online at <http://www.nature.com/naturegenetics/>
Reprints and permissions information is available online at <http://npg.nature.com/reprintsandpermissions/>

1. Carr, L.G. *et al.* *Alcohol. Clin. Exp. Res.* **13**, 594–596 (1989).
2. Hoog, J.O. *et al.* *Eur. J. Biochem.* **159**, 215–218 (1986).
3. Smith, M. *Adv. Hum. Genet.* **15**, 249–290 (1986).
4. Brennan, P. *et al.* *Am. J. Epidemiol.* **159**, 1–16 (2004).
5. Hashibe, M. *et al.* *Cancer Epidemiol. Biomarkers Prev.* **15**, 696–703 (2006).
6. The International HapMap Project. *Nature* **426**, 789–796 (2003).
7. Smigielski *et al.* *Nucleic Acids Res.* **28**, 352–255 (2000).
8. Birley, A.J. *et al.* *Hum. Mol. Genet.* **17**, 179–189 (2008).

Copyright of Nature Genetics is the property of Nature Publishing Group and its content may not be copied or emailed to multiple sites or posted to a listserv without the copyright holder's express written permission. However, users may print, download, or email articles for individual use.