

Supplemental materials of

Active cigarette smoking is associated with an exacerbation of genetic susceptibility to diabetes

Wan-Yu Lin ^{1,2*}, Yu-Li Liu ³, Albert C. Yang ^{4,5}, Shih-Jen Tsai ^{5,6,7}, Po-Hsiu Kuo ^{1,2*}

¹ Institute of Epidemiology and Preventive Medicine, College of Public Health, National Taiwan University, Taipei, Taiwan

² Department of Public Health, College of Public Health, National Taiwan University, Taipei, Taiwan

³ Center for Neuropsychiatric Research, National Health Research Institutes, Miaoli, Taiwan

⁴ Division of Interdisciplinary Medicine and Biotechnology, Beth Israel Deaconess Medical Center/Harvard Medical School, Boston, MA, USA

⁵ Institute of Brain Science, National Yang-Ming University, Taipei, Taiwan

⁶ Division of Psychiatry, National Yang-Ming University, Taipei, Taiwan

⁷ Department of Psychiatry, Taipei Veterans General Hospital, Taipei, Taiwan

Short title: Smoking and genetic risk of diabetes

* Corresponding authors: Po-Hsiu Kuo, Ph.D. and Wan-Yu Lin, Ph.D.

Po-Hsiu Kuo, Ph.D. (<http://orcid.org/0000-0003-0365-3587>)

Room 521, No. 17, Xu-Zhou Road, Taipei 100, Taiwan

Phone: +886-2-33668015; Fax: +886-2-23511955; E-mail: phkuo@ntu.edu.tw

Wan-Yu Lin, Ph.D. (<http://orcid.org/0000-0002-3385-4702>)

Room 501, No. 17, Xu-Zhou Road, Taipei 100, Taiwan

Phone/Fax: +886-2-33668106; E-mail: linwy@ntu.edu.tw

(Discovery cohort / Replication cohort) (D/R)	Fasting glucose (mg/dL)		HbA1c (%)		Diabetes (dichotomous trait)	
	Percent change (%)	p-value	Percent change (%)	p-value	Odds ratio	p-value
Explanatory variables in regression model (1)						
Sex (1: female vs. 0: male)	-3.71 ² (-3.33)	4.5E-77 (2.8E-124)	-0.61 (-0.81)	3.4E-5 (5.9E-15)	0.73 (0.71)	9.7E-10 (1.3E-24)
Age (in years, continuous variable)	0.34 (0.36)	4.8E-272 (0)	0.28 (0.29)	0 (0)	1.08 (1.08)	3.5E-177 (0)
Body mass index (in kg/m ² , continuous variable)	0.87 (0.85)	1.3E-239 (0)	0.70 (0.73)	3.1E-297 (0)	1.18 (1.18)	1.7E-155 (0)
The number of pack-years (mean±sd, D: 19.6±17.0; R: 19.5±17.6)	0.06 (0.05)	2.3E-7 (8.8E-9)	0.06 (0.06)	4.0E-12 (5.8E-22)	1.01 (1.01)	5.9E-9 (6.5E-12)
Drinking status (1: yes vs. 0: no)	1.18 (1.48)	1.6E-3 (1.2E-7)	-1.07 (-1.32)	7.0E-5 (5.1E-11)	0.98 (0.88)	0.83 (0.05)
Regular exercise (1: yes vs. 0: no)	-0.62 (-0.87)	1.8E-3 (2.8E-11)	-0.88 (-0.95)	1.1E-9 (1.3E-23)	0.89 (0.87)	0.02 (1.4E-5)
Educational attainment (a value ranging from 1 to 7)	-0.57 (-0.51)	6.6E-8 (7.0E-14)	-0.48 (-0.19)	3.9E-10 (1.0E-4)	0.92 (0.89)	3.2E-4 (2.7E-15)
R-square	13.8 % (13.1%)		14.2 % (13.9%)		13.8 % (13.3 %)	

Supplemental Table S1 Results of regression model (1) when the smoking measurement is “the number of pack-years” (prior to GRS analysis)

(Discovery cohort / Replication cohort) (D/R)	Fasting glucose (mg/dL)		HbA1c (%)		Diabetes (dichotomous trait)	
	Percent change (%)	<i>p</i> -value	Percent change (%)	<i>p</i> -value	Odds ratio	<i>p</i> -value
Sex (1: female vs. 0: male)	-3.76 (-3.34)	1.2E-77 (5.0E-122)	-0.65 (-0.78)	1.2E-5 (1.3E-13)	0.73 (0.71)	1.5E-9 (1.0E-22)
Age (in years, continuous variable)	0.34 (0.36)	4.6E-272 (0)	0.28 (0.29)	0 (0)	1.08 (1.08)	3.8E-177 (0)
Body mass index (in kg/m ² , continuous variable)	0.87 (0.85)	4.3E-240 (0)	0.70 (0.73)	7.2E-298 (0)	1.18 (1.18)	3.0E-156 (0)
Years as a smoker (mean±sd, D: 25.6±10.1; R: 26.3±10.8)	0.04 (0.04)	8.6E-4 (3.2E-6)	0.04 (0.05)	2.9E-6 (3.4E-20)	1.01 (1.01)	1.1E-6 (3.5E-11)
Drinking status (1: yes vs. 0: no)	1.28 (1.52)	6.8E-4 (5.5E-8)	-0.99 (-1.32)	2.5E-4 (5.5E-11)	0.99 (0.88)	0.87 (0.05)
Regular exercise (1: yes vs. 0: no)	-0.66 (-0.88)	9.8E-4 (1.2E-11)	-0.92 (-0.96)	2.5E-10 (4.0E-24)	0.88 (0.87)	0.01 (1.0E-5)
Educational attainment (a value ranging from 1 to 7)	-0.60 (-0.51)	1.5E-8 (3.5E-14)	-0.50 (-0.19)	4.7E-11 (8.8E-5)	0.91 (0.89)	1.5E-4 (1.9E-15)
R-square	13.7 % (13.1%)		14.1 % (13.9%)		13.8 % (13.3 %)	

Supplemental Table S2 Results of regression model (1) when the smoking measurement is “the number of years as a smoker” (prior to GRS analysis)

(Discovery cohort / Replication cohort) (D/R)	Fasting glucose (mg/dL)		HbA1c (%)		Diabetes (dichotomous trait)	
	Percent change (%)	<i>p</i> -value	Percent change (%)	<i>p</i> -value	Odds ratio	<i>p</i> -value
Sex (1: female vs. 0: male)	-3.70 (-3.35)	1.5E-76 (4.1E-125)	-0.61 (-0.82)	3.6E-5 (2.6E-15)	0.73 (0.71)	1.5E-9 (9.0E-25)
Age (in years, continuous variable)	0.35 (0.36)	1.7E-276 (0)	0.29 (0.29)	0 (0)	1.08 (1.08)	3.1E-181 (0)
Body mass index (in kg/m ² , continuous variable)	0.87 (0.85)	7.2E-239 (0)	0.70 (0.73)	3.5E-296 (0)	1.18 (1.18)	1.3E-154 (0)
Packs smoked per day (mean±sd, D: 0.72±0.51; R: 0.69±0.52)	1.70 (1.33)	4.9E-7 (3.6E-7)	1.63 (1.70)	3.4E-11 (7.0E-19)	1.51 (1.41)	1.1E-9 (2.3E-11)
Drinking status (1: yes vs. 0: no)	1.17 (1.49)	1.9E-3 (9.3E-8)	-1.08 (-1.31)	6.0E-5 (6.9E-11)	0.97 (0.88)	0.76 (0.05)
Regular exercise (1: yes vs. 0: no)	-0.63 (-0.88)	1.6E-3 (1.6E-11)	-0.89 (-0.96)	7.1E-10 (3.9E-24)	0.89 (0.87)	0.02 (1.2E-5)
Educational attainment (a value ranging from 1 to 7)	-0.56 (-0.50)	1.0E-7 (1.1E-13)	-0.47 (-0.19)	7.0E-10 (1.6E-4)	0.92 (0.89)	4.8E-4 (4.6E-15)
R-square	13.8 % (13.1%)		14.2 % (13.9%)		13.8 % (13.3 %)	

Supplemental Table S3 Results of regression model (1) when the smoking measurement is “packs smoked per day” (prior to GRS analysis)

(Discovery cohort / Replication cohort) (D/R)	Fasting glucose (mg/dL)		HbA1c (%)		Diabetes (dichotomous trait)	
	Percent change (%)	p-value	Percent change (%)	p-value	Odds ratio	p-value
Sex (1: female vs. 0: male)	-3.92 (-3.51)	1.2E-89 (4.6E-145)	-0.82 (-1.03)	1.1E-8 (1.4E-24)	0.68 (0.67)	9.7E-14 (1.8E-33)
Age (in years, continuous variable)	0.34 (0.36)	1.2E-272 (0)	0.28 (0.29)	0 (0)	1.08 (1.08)	1.6E-177 (0)
Body mass index (in kg/m ² , continuous variable)	0.87 (0.85)	5.3E-240 (0)	0.70 (0.73)	8.7E-298 (0)	1.18 (1.18)	4.2E-156 (0)
Hours as a passive smoker per week (mean±sd, D: 5.6±11.0; R: 5.4±10.6)	0.03 (0.04)	0.18 (0.02)	0.01 (0.03)	0.52 (0.03)	1.01 (1.01)	0.04 (0.01)
Drinking status (1: yes vs. 0: no)	1.47 (1.71)	8.0E-5 (5.0E-10)	-0.78 (-1.01)	3.6E-3 (3.1E-7)	1.04 (0.94)	0.67 (0.29)
Regular exercise (1: yes vs. 0: no)	-0.70 (-0.92)	4.0E-4 (1.2E-12)	-0.97 (-1.02)	2.3E-11 (3.9E-27)	0.87 (0.85)	0.005 (9.3E-7)
Educational attainment (a value ranging from 1 to 7)	-0.63 (-0.53)	2.5E-9 (2.8E-15)	-0.54 (-0.23)	1.7E-12 (3.6E-6)	0.91 (0.88)	3.7E-5 (8.0E-17)
R-square	13.7 % (13.1 %)		14.0 % (13.8 %)		13.6 % (13.2 %)	

Supplemental Table S4 Results of regression model (1) when the smoking measurement is “hours as a passive smoker per week” (prior to GRS analysis)

Chr.	SNP / SNV	Base pair	Effect allele	Other allele	Mapped gene	MAF	P_{INT} with “active smoking status”	P_{INT} with “the number of pack-years”	P_{INT} with “years as a smoker”	P_{INT} with “packs smoked per day”	P_{INT} with “hours as a passive smoker per week”
2	rs340515	45188370	G	T	-----	0.439523	0.421814	0.369612	0.248795	0.508034	0.237799
2	rs34177044	169754485	A	G	-----	0.403827	0.390463	0.011082	0.022235	0.117793	0.472982
2	rs2232326	169764491	C	T	<i>G6PC2</i>	0.041328	0.202882	0.203321	0.221588	0.3382	0.805891
2	rs60415045	169771124	C	A	-----	0.286759	0.18333	0.019855	0.032892	0.142232	0.241038
6	rs9356744	20685486	C	T	<i>CDKAL1</i>	0.348777	0.199681	0.20419	0.191177	0.083728	0.983701
7	rs730497*	44223721	A	G	<i>GCK</i>	0.196822	0.21977	0.225264	0.512717	0.051395	0.553832
7	rs13229610	44258964	T	G	<i>CAMK2B</i>	0.459106	0.335828	0.049154	0.029717	0.274366	0.644279
7	rs2233580**	127253550	T	C	<i>PAX4</i>	0.106884	0.02893	0.077768	0.046995	0.097878	0.783261
7	rs61342118*	127761917	A	C	Near <i>PAX4</i>	0.086016	0.026518	0.088693	0.04746	0.037248	0.549475
8	rs75199970	2546494	T	A	<i>LOC101927815</i>	0.000230	0.040606	0.03539	0.041408	0.035423	0.67093
8	rs35859536*	118191475	T	C	Near <i>SLC30A8</i>	0.462279	0.672247	0.068205	0.105456	0.318498	0.109532
9	rs10811661*	22134094	C	T	-----	0.416026	0.430898	0.124122	0.424267	0.347333	0.361638
10	rs2399794	12248800	G	A	<i>CDC123</i>	0.473194	0.028171	0.000861	0.038414	9.68E-05	0.98143
11	rs163177	2838413	C	T	<i>KCNQ1</i>	0.465963	0.366676	0.101701	0.100647	0.126057	0.316126
11	rs60808706	2857233	A	G	<i>KCNQ1</i>	0.358092	0.409672	0.131914	0.49631	0.143097	0.147599
11	rs10830963	92708710	G	C	<i>MTNR1B</i>	0.431628	0.900442	0.440297	0.274092	0.899981	0.535962

Supplemental Table S5 16 FG-associated SNPs discovered by TWB1 ($P_{SNP} < 6.25 \times 10^{-9}$ in model 2)

** rs2233580 was also HbA1c-associated ($P_{SNP} < 6.25 \times 10^{-9}$ in model 2) and diabetes-associated ($P_{SNP} < 6.25 \times 10^{-9}$ in model 2).

* rs61342118 was also diabetes-associated ($P_{SNP} < 6.25 \times 10^{-9}$ in model 2).

* rs730497, rs35859536, and rs10811661 were also HbA1c-associated ($P_{SNP} < 6.25 \times 10^{-9}$ in model 2).

P_{INT} is the p -value of testing $H_0: \beta_{INT} = 0$ vs. $H_1: \beta_{INT} \neq 0$ from the following model:

$$g[E(Y)] = \beta_0 + \beta_{SNP}SNP + \beta_{Smoking}Smoking + \beta_{INT}SNP \times Smoking + \sum_{v=1}^{16} \beta_{C_v}Covariate_v ,$$

where Y is natural log transformed fasting glucose with an identity link $g[\cdot]$, and $Smoking$ is one of the five smoking measurements.

Chr.	SNP / SNV	Base pair	Effect allele	Other allele	Mapped gene	MAF	P_{INT} with “active smoking status”	P_{INT} with “the number of pack-years”	P_{INT} with “years as a smoker”	P_{INT} with “packs smoked per day”	P_{INT} with “hours as a passive smoker per week”
2	rs78529720	169777297	G	A	<i>ABCB11</i>	0.289754	0.072579	0.024974	0.019681	0.158984	0.235235
6	rs1012635	20675295	G	A	<i>CDKAL1</i>	0.46723	0.768481	0.923266	0.788748	0.920955	0.140758
6	rs35612982	20682622	C	T	<i>CDKAL1</i>	0.346437	0.542043	0.466119	0.466379	0.310865	0.576941
7	rs730497*	44223721	A	G	<i>GCK</i>	0.196822	0.024183	0.017333	0.010552	0.008819	0.383881
7	rs12536382	127228728	C	T	<i>ARF5</i>	0.255785	0.192881	0.156799	0.151725	0.371714	0.184555
7	rs2233580**	127253550	T	C	<i>PAX4</i>	0.106884	0.033878	0.003083	0.010772	0.022765	0.255327
8	rs35859536*	118191475	T	C	Near <i>SLC30A8</i>	0.462279	0.662187	0.014467	0.036535	0.117965	0.342106
9	rs10811661*	22134094	C	T	-----	0.416026	0.289478	0.009493	0.092448	0.050075	0.263594
10	rs7896600	12255175	G	C	<i>CDC123</i>	0.442577	7.25E-07	3.94E-08	8.48E-07	2.02E-09	0.384392
10	rs1174605899	12313728	A	C	Near <i>CDC123</i>	0.489238	1.71E-06	6.37E-06	6.52E-06	9.81E-07	0.295997
11	rs11024175	2830565	T	C	<i>KCNQ1</i>	0.406234	0.04272	0.005488	0.012317	0.033745	0.008735
17	rs117948248	80759687	T	C	<i>TBCD</i>	0.1562	0.035671	0.000255	0.000855	0.009197	0.083431

Supplemental Table S6 12 HbA1c-associated SNPs discovered by TWB1 ($P_{SNP} < 6.25 \times 10^{-9}$ in model 2)

** rs2233580 was also FG-associated ($P_{SNP} < 6.25 \times 10^{-9}$ in model 2) and diabetes-associated ($P_{SNP} < 6.25 \times 10^{-9}$ in model 2).

* rs730497, rs35859536, and rs10811661 were also FG-associated ($P_{SNP} < 6.25 \times 10^{-9}$ in model 2).

P_{INT} is the p -value of testing $H_0: \beta_{INT} = 0$ vs. $H_1: \beta_{INT} \neq 0$ from the following model:

$$g[E(Y)] = \beta_0 + \beta_{SNP}SNP + \beta_{Smoking}Smoking + \beta_{INT}SNP \times Smoking + \sum_{v=1}^{16} \beta_{Cv}Covariate_v ,$$

where Y is natural log transformed HbA1c with an identity link $g[\cdot]$, and $Smoking$ is one of the five smoking measurements.

Chr.	SNP / SNV	Base pair	Effect allele	Other allele	Mapped gene	MAF	P_{INT} with “active smoking status”	P_{INT} with “the number of pack-years”	P_{INT} with “years as a smoker”	P_{INT} with “packs smoked per day”	P_{INT} with “hours as a passive smoker per week”
6	rs3749925	20712461	G	C	<i>CDKAL1</i>	0.201348	0.317065	0.317987	0.238203	0.326624	0.030372
7	rs2233580**	127253550	T	C	<i>PAX4</i>	0.106884	0.462396	0.310362	0.381483	0.502645	0.801273
7	rs61342118*	127761917	A	C	<i>Near PAX4</i>	0.086016	0.046082	0.159406	0.047092	0.223836	0.823951
8	rs13266634***	118184783	T	C	<i>SLC30A8</i>	0.461692	0.111917	0.123203	0.017727	0.298204	0.823684
10	rs11257655	12307894	C	T	<i>Near CDC123</i>	0.444304	0.020900	0.109184	0.097288	0.022917	0.042687
11	rs163184	2847069	G	T	<i>KCNQ1</i>	0.443249	0.131492	0.974283	0.063704	0.945790	0.662608

Supplemental Table S7 6 diabetes-associated SNPs discovered by TWB1 ($P_{SNP} < 6.25 \times 10^{-9}$)

** rs2233580 was also FG-associated ($P_{SNP} < 6.25 \times 10^{-9}$ in model 2) and HbA1c-associated ($P_{SNP} < 6.25 \times 10^{-9}$ in model 2).

* rs61342118 was also FG-associated ($P_{SNP} < 6.25 \times 10^{-9}$ in model 2).

*** rs13266634 is also among the 38 diabetes-associated SNPs identified from European GWASs, as listed in the following Tables S18-S20.

P_{INT} is the p -value of testing $H_0: \beta_{INT} = 0$ vs. $H_1: \beta_{INT} \neq 0$ from the following model:

$$g[E(Y)] = \beta_0 + \beta_{SNP}SNP + \beta_{Smoking}Smoking + \beta_{INT}SNP \times Smoking + \sum_{v=1}^{16} \beta_{C_v}Covariate_v ,$$

where Y is the diabetes status with a logit link $g[\cdot]$, and $Smoking$ is one of the five smoking measurements.

Explanatory variables in regression model (4)	Fasting glucose (mg/dL)		HbA1c (%)		Diabetes (dichotomous trait)	
	Percent change (%)	<i>p</i> -value	Percent change (%)	<i>p</i> -value	Odds ratio	<i>p</i> -value
Genetic risk score (GRS) (z-score standardized)	1.73	8.4E-172	1.29	7.6E-190	1.27	1.3E-35
Sex (1: female vs. 0: male)	-3.33	2.0E-120	-0.85	7.7E-16	0.71	6.3E-23
Age (in years, continuous variable)	0.36	0	0.29	0	1.08	0
Body mass index (in kg/m ² , continuous variable)	0.85	0	0.74	0	1.18	0
The number of pack-years (mean=19.5 pack-years, sd=17.6 pack-years)	0.03	0.19	0.03	0.03	1.01	0.002
GRS × The number of pack-years	0.06	1.6E-7	0.02	7.3E-4	1.002	0.15
Drinking status (1: yes vs. 0: no)	1.81	8.8E-10	-1.00	2.1E-6	0.87	0.05
Regular exercise (1: yes vs. 0: no)	-0.91	2.4E-12	-1.01	1.3E-26	0.87	9.4E-6
Educational attainment (a value ranging from 1 to 7)	-0.49	2.4E-13	-0.20	3.8E-5	0.89	2.1E-15
R-square	14.4 %		15.4 %		14.0 %	

Supplemental Table S8 Results of regression model (4) when the smoking measurement is “The number of pack-years” (including GRS and GRS-smoking interaction)

Explanatory variables in regression model (4)	Fasting glucose (mg/dL)		HbA1c (%)		Diabetes (dichotomous trait)	
	Percent change (%)	<i>p</i> -value	Percent change (%)	<i>p</i> -value	Odds ratio	<i>p</i> -value
Genetic risk score (GRS) (z-score standardized)	1.72	4.4E-170	1.30	1.4E-190	1.27	4.4E-35
Sex (1: female vs. 0: male)	-3.32	4.5E-120	-0.79	3.7E-14	0.71	1.7E-21
Age (in years, continuous variable)	0.36	0	0.29	0	1.08	0
Body mass index (in kg/m ² , continuous variable)	0.86	0	0.74	0	1.18	0
Years as a smoker (mean=26.3 years, sd=10.8 years)	0.02	0.20	0.03	0.002	1.01	0.003
GRS × Years as a smoker	0.04	8.5E-6	0.02	2.5E-3	1.004	0.0048
Drinking status (1: yes vs. 0: no)	2.03	3.6E-11	-0.93	2.3E-5	0.89	0.11
Regular exercise (1: yes vs. 0: no)	-0.92	1.3E-12	-1.01	8.6E-27	0.86	5.1E-6
Educational attainment (a value ranging from 1 to 7)	-0.51	4.0E-14	-0.21	1.9E-5	0.89	1.1E-15
R-square	14.4 %		15.4 %		14.0 %	

Supplemental Table S9 Results of regression model (4) when the smoking measurement is “the number of years as a smoker” (including GRS and GRS-smoking interaction)

Explanatory variables in regression model (4)	Fasting glucose (mg/dL)		HbA1c (%)		Diabetes (dichotomous trait)	
	Percent change (%)	<i>p</i> -value	Percent change (%)	<i>p</i> -value	Odds ratio	<i>p</i> -value
Genetic risk score (GRS) (z-score standardized)	1.74	2.5E-172	1.30	3.8E-190	1.27	1.4E-35
Sex (1: female vs. 0: male)	-3.33	2.9E-122	-0.84	6.2E-16	0.70	1.3E-23
Age (in years, continuous variable)	0.36	0	0.29	0 ²	1.08	0
Body mass index (in kg/m ² , continuous variable)	0.85	0	0.74	0 ²	1.18	0
Packs smoked per day (mean=0.69 packs, sd=0.52 packs)	0.89	0.08	1.02	0.006	1.38	0.01
GRS × Packs smoked per day	1.68	1.9E-7	0.69	4.5E-4	1.09	0.12
Drinking status (1: yes vs. 0: no)	1.77	2.4E-9	-1.05	9.0E-7	0.87	0.04
Regular exercise (1: yes vs. 0: no)	-0.91	2.1E-12	-1.01	1.2E-26	0.86	7.8E-6
Educational attainment (a value ranging from 1 to 7)	-0.49	5.4E-13	-0.20	6.0E-5	0.89	3.2E-15
R-square	14.4 %		15.4 %		14.0 %	

Supplemental Table S10 Results of regression model (4) when the smoking measurement is “Packs smoked per day” (including GRS and GRS-smoking interaction)

Explanatory variables in regression model (4)	Fasting glucose (mg/dL)		HbA1c (%)		Diabetes (dichotomous trait)	
	Percent change (%)	p-value	Percent change (%)	p-value	Odds ratio	p-value
Genetic risk score (GRS) (z-score standardized)	1.71	2.2E-167	1.29	1.6E-188	1.27	9.2E-36
Sex (1: female vs. 0: male)	-3.48	1.1E-144	-1.03	6.5E-25	0.67	1.4E-32
Age (in years, continuous variable)	0.36	0	0.29	0	1.08	0
Body mass index (in kg/m ² , continuous variable)	0.86	0	0.74	0	1.18	0
Hours as a passive smoker per week (mean=5.4 hours, sd=10.6 hours)	0.05	0.03	0.04	0.02	1.006	0.33
GRS × Hours as a passive smoker per week	0.02	0.25	-0.007	0.58	0.996	0.30
Drinking status (1: yes vs. 0: no)	1.79	8.8E-11	-0.99	5.8E-7	0.92	0.21
Regular exercise (1: yes vs. 0: no)	-0.95	2.0E-13	-1.06	1.7E-29	0.85	5.9E-7
Educational attainment (a value ranging from 1 to 7)	-0.52	1.2E-14	-0.23	1.6E-6	0.88	9.0E-17
R-square	14.3 %		15.2 %		13.9 %	

Supplemental Table S11 Results of regression model (4) when the smoking measurement is “the number of hours as a passive smoker per week” (including GRS and GRS-smoking interaction)

Smoking measurements in regression model (4)	Fasting glucose (mg/dL)		HbA1c (%)	
	$\hat{\Phi}_{INT}$	P_{INT}	$\hat{\Phi}_{INT}$	P_{INT}
GRS × Active smoking status	1.28	1.6E-4 *	0.0438	2.7E-4 *
GRS × The number of pack-years	0.07	2.3E-7 *	0.0018	7.0E-5 *
GRS × Years as a smoker	0.05	4.3E-5 *	0.0015	3.6E-4 *
GRS × Packs smoked per day	2.17	6.4E-7 *	0.0572	4.5E-5 *
GRS × Hours as a passive smoker per week	0.04	0.14	-0.00025	0.78

Supplemental Table S12 Results of regression model (4) (without natural log transformation on fasting glucose or HbA1c)

*A total of 15 GRS-smoking interactions were tested in the main manuscript (although the results of dichotomous “diabetes status” were not listed here). A GRS-smoking interaction is significant if $P_{INT} < \frac{0.05}{15} = 0.0033$.

Smoking measurements in regression model (4)	Fasting glucose (mg/dL)		HbA1c (%)		Diabetes (dichotomous trait)	
	Percent change (%)	P_{INT}	Percent change (%)	P_{INT}	Odds ratio	P_{INT}
GRS × Active smoking status	1.22	1.8E-7 *	0.51	0.001 *	1.13	0.0063
GRS × The number of pack-years	0.07	1.0E-12 *	0.03	3.5E-6 *	1.003	0.07
GRS × Years as a smoker	0.05	4.6E-9 *	0.02	1.7E-4 *	1.004	0.0032 *
GRS × Packs smoked per day	2.02	2.8E-11 *	0.79	1.8E-5 *	1.08	0.11
GRS × Hours as a passive smoker per week	0.02	0.29	-0.004	0.76	0.996	0.34

Supplemental Table S13 Results of regression model (4) (assess the GRS-smoking interaction without controlling for the other interaction terms)

*A total of 15 GRS-smoking interactions were tested here. A GRS-smoking interaction is significant if $P_{INT} < \frac{0.05}{15} = 0.0033$.

Smoking measurements in regression model (4)	Sex	Fasting glucose (mg/dL)		HbA1c (%)		Diabetes (dichotomous trait)	
		Percent change (%)	P_{INT}	Percent change (%)	P_{INT}	Odds ratio	P_{INT}
GRS × Active smoking status	Males	1.21	2.1E-4 *	0.80	4.6E-4 *	1.18	0.0018
	Females	0.32	0.52	-0.16	0.58	1.08	0.77
GRS × The number of pack-years	Males	0.06	3.2E-6 *	0.03	2.1E-4 *	1.003	0.033
	Females	-0.03	0.37	-0.02	0.23	0.99	0.25
GRS × Years as a smoker	Males	0.04	6.2E-5 *	0.03	5.1E-4 *	1.005	0.0013 *
	Females	-0.00056	0.98	-0.01	0.24	1.003	0.64
GRS × Packs smoked per day	Males	1.86	3.6E-6 *	0.95	1.2E-4 *	1.12	0.032
	Females	-0.009	0.99	-0.43	0.37	1.01	0.58
GRS × Hours as a passive smoker per week	Males	0.09	0.03	0.04	0.16	1.004	0.57
	Females	-0.002	0.92	-0.02	0.05	0.985	0.24

Supplemental Table S14 Results of GRS-smoking interactions while stratified by sex

*A total of 30 GRS-smoking interactions were tested here. A GRS-smoking interaction is significant if $P_{INT} < \frac{0.05}{30} = 0.0016$.

	Overall	Smokers	Non-smokers
Total, <i>n</i>	2,091	248	1,843
Males, <i>n</i> (%)	1,036 (49.5)	208 (83.9)	828 (44.9)
Age (years), mean (s.d.)	49.9 (11.1)	48.6 (10.5)	50.1 (11.1)
BMI (kg/m ²), mean (s.d.)	24.4 (3.7)	25.4 (3.8)	24.3 (3.6)
Drinking, <i>n</i> (%)	145 (6.9)	64 (25.8)	81 (4.4)
Regular exercise, <i>n</i> (%)	924 (44.2)	85 (34.3)	839 (45.5)
Educational attainment, mean (s.d.)	5.6 (0.9)	5.5 (0.8)	5.6 (0.9)
Fasting glucose (mg/dL), mean (s.d.)	95.9 (20.3)	101.5 (34.4)	95.1 (17.4)
HbA1c (%), mean (s.d.)	5.72 (0.74)	5.83 (1.01)	5.71 (0.69)
HbA1c (mmol/mol), mean (s.d.)	39 (8.1)	40 (11.0)	39 (7.5)
Fasting glucose > 126 mg/dL, <i>n</i> (%)	75 (3.6)	15 (6.0)	60 (3.3)
HbA1c > 6.5 % (48 mmol/mol), <i>n</i> (%)	123 (5.9)	21 (8.5)	102 (5.5)
Subjects with physician-diagnosed diabetes, <i>n</i> (%)	86 (4.1)	10 (4.0)	76 (4.1)
Subjects with diabetes, <i>n</i> (%) ¹	159 (7.6)	26 (10.5)	133 (7.2)

Supplemental Table S15 Basic characteristics of the 2,091 subjects with blood DNA methylation data

- 1 Subjects with diabetes included those with physician-diagnosed diabetes, or those having FG > 126 mg/dL or HbA1c > 6.5 % (48 mmol/mol) according to the TWB test results.

Regression coefficient (<i>p</i> -value)	cg06335123 (in <i>CDC123</i>)	cg26963277 (in <i>KCNQ1</i>)	cg01744331 (in <i>KCNQ1</i>)	cg16556677 (in <i>KCNQ1</i>)
Regressing DNA methylation level on “Packs smoked per day”, sex, age, and BMI: Negative regression coefficients => More packs smoked per day, aging, and a larger BMI are associated with decreased levels of DNA methylation at these four sites.				
Packs smoked per day (mean±sd: 0.68±0.51)	-0.01349 (0.00025)	-0.02033 (3.0E-17)	-0.02141 (4.5E-10)	-0.01328 (2.7E-5)
Sex (1: female vs. 0: male)	0.00577 (0.00645)	0.01151 (1.1E-16)	0.01790 (1.8E-19)	0.01014 (2.7E-8)
Age (in years, continuous variable)	-0.00054 (3.5E-9)	-0.00043 (2.9E-13)	-5.6E-4 (4.2E-11)	-0.00024 (1.7E-3)
Body mass index (in kg/m ² , continuous variable)	-0.00116 (4.6E-5)	-6.5E-6 (0.97)	-8.7E-5 (0.74)	-0.00021 (0.40)
(Logistic regression) Regressing diabetes status on DNA methylation level, while adjusting for sex, age, and BMI: “Odds ratio < 1.0” indicates that an increased DNA methylation level is associated with a lower risk of diabetes, and a decreased DNA methylation level is associated with a higher risk of diabetes.				
Odds ratio (<i>p</i> -value)	0.0082 (0.0413)	0.0009 (0.0045)	0.0075 (0.0118)	0.0700 (0.2083)

Supplemental Table S16 Analysis results of blood DNA methylation data

SNP rsID	Chr	Position	Effect allele	Other allele	Mapped gene	MAF	P_{INT} with “active smoking status” ¹	P_{INT} with “the number of pack-years” ¹	P_{INT} with “years as a smoker” ¹	P_{INT} with “packs smoked per day” ¹	P_{INT} with “hours as a passive smoker per week” ¹
SNP-smoking interaction analysis on natural log transformed fasting glucose:											
rs1444261	2	55354466	C	T	<i>C2orf63</i>	0.062213	0.647064	0.067076	0.340325	0.191402	0.263129
rs1801232	10	16870912	T	G	<i>CUBN</i>	0.156081	0.294669	0.759270	0.453210	0.882073	0.960374
rs4132670	10	114767771	A	G	<i>TCF7L2</i>	0.005402	0.120335	0.915558	0.287107	0.839836	0.019209
rs12243326	10	114788815	C	T	<i>TCF7L2</i>	0.005368	0.102001	0.931361	0.261229	0.850150	0.019397
rs140637	15	48766855	G	A	<i>FBN1</i>	0.000009	No results because this variant is too rare in Taiwan Biobank.				
SNP-smoking interaction analysis on natural log transformed HbA1c:											
rs1444261	2	55354466	C	T	<i>C2orf63</i>	0.062213	0.515846	0.132088	0.951349	0.380867	0.728838
rs1801232	10	16870912	T	G	<i>CUBN</i>	0.156081	0.057120	0.113832	0.064861	0.167234	0.043065
rs4132670	10	114767771	A	G	<i>TCF7L2</i>	0.005402	0.029237	0.654357	0.155104	0.513931	0.095402
rs12243326	10	114788815	C	T	<i>TCF7L2</i>	0.005368	0.025216	0.703848	0.154910	0.557736	0.097694
rs140637	15	48766855	G	A	<i>FBN1</i>	0.000009	No results because this variant is too rare in Taiwan Biobank.				
SNP-smoking interaction analysis on diabetes status:											
rs1444261	2	55354466	C	T	<i>C2orf63</i>	0.062213	0.759288	0.113310	0.515076	0.321233	0.624802
rs1801232	10	16870912	T	G	<i>CUBN</i>	0.156081	0.396722	0.713981	0.620736	0.664798	0.733262
rs4132670	10	114767771	A	G	<i>TCF7L2</i>	0.005402	0.014853	0.579142	0.079367	0.371905	0.077710
rs12243326	10	114788815	C	T	<i>TCF7L2</i>	0.005368	0.012956	0.573007	0.074901	0.368543	0.079007
rs140637	15	48766855	G	A	<i>FBN1</i>	0.000009	No results because this variant is too rare in Taiwan Biobank.				

Supplemental Table S17 The 5 SNP-smoking interactions on diabetes identified from subjects of European and African ancestry

¹ P_{INT} is the p -value of testing $H_0: \beta_{INT} = 0$ vs. $H_1: \beta_{INT} \neq 0$ from the following model:

$$g[E(Y)] = \beta_0 + \beta_{SNP}SNP + \beta_{Smoking}Smoking + \beta_{INT}SNP \times Smoking + \sum_{v=1}^{16} \beta_{C_v}Covariate_v ,$$

where Y is natural log transformed FG or HbA1c with an identity link $g[\cdot]$, or the diabetes status with a logit link $g[\cdot]$. $Smoking$ is one of the five smoking measurements. $P_{INT} < 0.05$ was highlighted in yellow.

SNP rsID	Chr	Position	Effect allele	Other allele	Mapped gene	MAF	P_{INT} with “active smoking status” ¹	P_{INT} with “the number of pack-years” ¹	P_{INT} with “years as a smoker” ¹	P_{INT} with “packs smoked per day” ¹	P_{INT} with “hours as a passive smoker per week” ¹
rs10923931	1	120517959	T	G	NOTCH2	0.0266	0.497297	0.758139	0.733951	0.657158	0.500067
rs340874	1	214159256	C	T	PROX1	0.3866	0.34053	0.047393	0.51152	0.050668	0.93393
rs780094	2	27741237	C	T	GCKR	0.4805	0.663052	0.885067	0.922033	0.630088	0.066742
rs7578597	2	43732823	T	C	THADA	0.0052	0.073969	0.215989	0.09459	0.147412	0.535595
rs243021	2	60584819	A	G	-----	0.3388	0.140963	0.195474	0.160076	0.144135	0.580446
rs13389219	2	165528876	C	T	COBLL1	0.1028	0.110077	0.110058	0.15842	0.037917	0.732636
rs7578326	2	227020653	A	G	LOC646736	0.1510	0.191067	0.483716	0.588048	0.158477	0.752026
rs17036101	3	12277845	G	A	-----	0.0339	0.81851	0.894187	0.919896	0.90745	0.003235
rs4607103	3	64711904	C	T	ADAMTS9-AS2	0.3451	0.782297	0.930636	0.977063	0.57864	0.245092
rs11708067	3	123065778	A	G	ADCY5	0.0006	0.32134	0.603048	0.565434	0.513541	0.469027
rs459193	5	55806751	G	A	C5orf67	0.4687	0.559281	0.342269	0.676071	0.138188	0.640302
rs4457053	5	76424949	G	A	ZBED3-AS1	0.0559	0.038005	0.000692	0.008547	0.002559	0.145399
rs9472138	6	43811762	T	C	LOC107986598	0.1390	0.045065	0.000116	0.003241	0.002329	0.870623
rs2191349	7	15064309	T	G	-----	0.3098	0.59694	0.221766	0.265209	0.342553	0.218164
rs864745	7	28180556	T	C	JAZF1	0.2134	0.438802	0.559807	0.724187	0.456061	0.21298
rs972283	7	130466854	G	A	LOC105375508	0.3136	0.058095	0.018156	0.073993	0.004794	0.562678
rs516946	8	41519248	C	T	ANK1	0.1267	0.346076	0.103098	0.048341	0.561153	0.215356
rs896854	8	95960511	T	C	TP53INP1	0.2746	0.858037	0.676124	0.94125	0.504252	0.852747
rs13266634***	8	118184783	C	T	SLC30A8 ²	0.4617	0.648976	0.065385	0.09736	0.312004	0.127494
rs13292136	9	81952128	C	T	-----	0.0870	0.357974	0.818348	0.917207	0.170779	0.283483
rs2796441	9	84308948	G	A	LOC101927502	0.3837	0.068336	0.11528	0.088995	0.048266	0.796523
rs12779790	10	12328010	G	A	Near CDC123²	0.1824	0.224513	0.112216	0.311603	0.049073	0.95917
rs12571751	10	80942631	A	G	ZMIZ1	0.4227	0.565342	0.792829	0.210464	0.87981	0.007788

SNP rsID	Chr	Position	Effect allele	Other allele	Mapped gene	MAF	P_{INT} with “active smoking status” ¹	P_{INT} with “the number of pack-years” ¹	P_{INT} with “years as a smoker” ¹	P_{INT} with “packs smoked per day” ¹	P_{INT} with “hours as a passive smoker per week” ¹
rs7903146	10	114758349	T	C	<i>TCF7L2</i>	0.0236	0.53338	0.031689	0.226831	0.42343	0.276035
rs231362	11	2691471	G	A	<i>KCNQ1</i> ²	0.0958	0.808066	0.602967	0.846658	0.583126	0.386619
rs1552224	11	72433098	A	C	<i>ARAP1</i>	0.0663	0.440084	0.877019	0.603174	0.952758	0.965577
rs1387153	11	92673828	T	C	-----	0.4563	0.864188	0.134995	0.22974	0.561622	0.216203
rs10842994	12	27965150	C	T	<i>LOC105369709</i>	0.2070	0.450971	0.411314	0.919561	0.39939	0.781914
rs1153188	12	55098996	A	T	-----	0.0145	0.716316	0.125885	0.732642	0.167496	0.756387
rs1531343	12	66174894	C	G	<i>RPSAP52</i>	0.1091	0.267408	0.265706	0.219831	0.350382	0.331831
rs7961581	12	71663102	C	T	-----	0.2287	0.991809	0.401535	0.701329	0.388973	0.851531
rs7957197	12	121460686	T	A	<i>OASL</i>	0.0006	0.977098	0.921786	0.985963	0.901942	0.776688
rs7177055	15	77832762	A	G	-----	0.3341	0.819691	0.999054	0.565509	0.759913	0.64185
rs11634397	15	80432222	G	A	-----	0.0836	0.225901	0.008178	0.115465	0.024652	0.594659
rs8042680	15	91521337	A	C	<i>PRC1</i>	0.0029	0.558726	0.499235	0.410369	0.438919	0.746225
rs7202877	16	75247245	T	G	-----	0.2050	0.927582	0.184688	0.84189	0.248119	0.882796
rs12970134	18	57884750	A	G	-----	0.1738	0.435106	0.12862	0.555369	0.117711	0.309806
rs10401969	19	19407718	C	T	<i>SUGP1</i>	0.0965	0.214083	0.028897	0.223583	0.022962	0.408376

Supplemental Table S18 (The 38 diabetes-associated SNPs identified from European GWASs) SNP-smoking interaction analysis on natural log transformed fasting glucose

¹ P_{INT} is the p -value of testing $H_0: \beta_{INT} = 0$ vs. $H_1: \beta_{INT} \neq 0$ from the following model:

$$g[E(Y)] = \beta_0 + \beta_{SNP}SNP + \beta_{Smoking}Smoking + \beta_{INT}SNP \times Smoking + \sum_{v=1}^{16} \beta_{C_v}Covariate_v ,$$

where Y is natural log transformed fasting glucose with an identity link $g[\cdot]$, and $Smoking$ is one of the five smoking measurements. $P_{INT} < 0.05$

was highlighted in yellow.

² *CDC123*, *KCNQ1*, and *SLC30A8* were also associated with FG/HbA1c/diabetes according to our discovery cohort (TWB1), as listed in the above Tables S5-S7.

*** SNP rs13266634 (in the *SLC30A8* gene) is also among the 6 diabetes-associated SNPs discovered by TWB1, as listed in the above Table S7.

SNP rsID	Chr	Position	Effect allele	Other allele	Mapped gene	MAF	P_{INT} with “active smoking status” ¹	P_{INT} with “the number of pack-years” ¹	P_{INT} with “years as a smoker” ¹	P_{INT} with “packs smoked per day” ¹	P_{INT} with “hours as a passive smoker per week” ¹
rs10923931	1	120517959	T	G	NOTCH2	0.0266	0.098816	0.974414	0.081733	0.727306	0.80535
rs340874	1	214159256	C	T	PROX1	0.3866	0.615122	0.745585	0.642277	0.673305	0.299826
rs780094	2	27741237	C	T	GCKR	0.4805	0.027893	0.060585	0.002821	0.162519	0.273771
rs7578597	2	43732823	T	C	THADA	0.0052	0.020788	0.014977	0.003526	0.037148	0.610847
rs243021	2	60584819	A	G	-----	0.3388	0.991047	0.796397	0.935671	0.879511	0.626304
rs13389219	2	165528876	C	T	COBLL1	0.1028	0.170474	0.64622	0.371401	0.287757	0.697176
rs7578326	2	227020653	A	G	LOC646736	0.1510	0.218471	0.324675	0.416461	0.252683	0.953437
rs17036101	3	12277845	G	A	-----	0.0339	0.606802	0.873054	0.631549	0.953848	0.016144
rs4607103	3	64711904	C	T	ADAMTS9-AS2	0.3451	0.695618	0.764223	0.598439	0.948124	0.14818
rs11708067	3	123065778	A	G	ADCY5	0.0006	0.48117	0.819303	0.724267	0.742881	0.628109
rs459193	5	55806751	G	A	C5orf67	0.4687	0.416742	0.981739	0.623708	0.547031	0.625168
rs4457053	5	76424949	G	A	ZBED3-AS1	0.0559	0.516869	0.101914	0.282282	0.133313	0.046199
rs9472138	6	43811762	T	C	LOC107986598	0.1390	0.070302	0.000355	0.002948	0.008385	0.556876
rs2191349	7	15064309	T	G	-----	0.3098	0.372583	0.494297	0.742857	0.642247	0.404233
rs864745	7	28180556	T	C	JAZF1	0.2134	0.724627	0.993206	0.483087	0.986243	0.464115
rs972283	7	130466854	G	A	LOC105375508	0.3136	0.031582	0.055947	0.044097	0.01254	0.438507
rs516946	8	41519248	C	T	ANK1	0.1267	0.424793	0.993553	0.798159	0.39221	0.093636
rs896854	8	95960511	T	C	TP53INP1	0.2746	0.95641	0.831707	0.956592	0.907038	0.517093
rs13266634***	8	118184783	C	T	SLC30A8 ²	0.4617	0.658035	0.014351	0.036456	0.117095	0.388746
rs13292136	9	81952128	C	T	-----	0.0870	0.473158	0.264458	0.149703	0.929171	0.036055
rs2796441	9	84308948	G	A	LOC101927502	0.3837	0.551078	0.908062	0.618745	0.641946	0.50278
rs12779790	10	12328010	G	A	Near CDC123²	0.1824	0.00151	0.000427	0.003391	8.91E-05	0.918916
rs12571751	10	80942631	A	G	ZMIZ1	0.4227	0.79652	0.933041	0.355454	0.858318	0.076816

SNP rsID	Chr	Position	Effect allele	Other allele	Mapped gene	MAF	P_{INT} with “active smoking status” ¹	P_{INT} with “the number of pack-years” ¹	P_{INT} with “years as a smoker” ¹	P_{INT} with “packs smoked per day” ¹	P_{INT} with “hours as a passive smoker per week” ¹
rs7903146	10	114758349	T	C	<i>TCF7L2</i>	0.0236	0.117409	0.02642	0.07501	0.267786	0.915459
rs231362	11	2691471	G	A	<i>KCNQ1</i> ²	0.0958	0.683558	0.382904	0.792207	0.505264	0.827145
rs1552224	11	72433098	A	C	<i>ARAP1</i>	0.0663	0.290744	0.498728	0.19708	0.798587	0.156948
rs1387153	11	92673828	T	C	-----	0.4563	0.648617	0.449784	0.496341	0.886364	0.149319
rs10842994	12	27965150	C	T	<i>LOC105369709</i>	0.2070	0.560089	0.777197	0.098758	0.941118	0.493545
rs1153188	12	55098996	A	T	-----	0.0145	0.195607	0.853368	0.282086	0.702558	0.634745
rs1531343	12	66174894	C	G	<i>RPSAP52</i>	0.1091	0.375011	0.090797	0.140622	0.195879	0.613751
rs7961581	12	71663102	C	T	-----	0.2287	0.713713	0.839775	0.401286	0.549239	0.160869
rs7957197	12	121460686	T	A	<i>OASL</i>	0.0006	0.915054	0.832918	0.936515	0.864064	0.625215
rs7177055	15	77832762	A	G	-----	0.3341	0.678558	0.183532	0.61204	0.163101	0.404036
rs11634397	15	80432222	G	A	-----	0.0836	0.983663	0.072363	0.703859	0.123287	0.626227
rs8042680	15	91521337	A	C	<i>PRC1</i>	0.0029	0.124564	0.356929	0.028107	0.602057	0.406163
rs7202877	16	75247245	T	G	-----	0.2050	0.953117	0.106242	0.784282	0.113571	0.39607
rs12970134	18	57884750	A	G	-----	0.1738	0.002284	0.001549	0.0032	0.001553	0.353752
rs10401969	19	19407718	C	T	<i>SUGP1</i>	0.0965	0.618241	0.366786	0.901497	0.25566	0.786735

Supplemental Table S19 (The 38 diabetes-associated SNPs identified from European GWASs) SNP-smoking interaction analysis on natural log transformed HbA1c

¹ P_{INT} is the p -value of testing $H_0: \beta_{INT} = 0$ vs. $H_1: \beta_{INT} \neq 0$ from the following model:

$$g[E(Y)] = \beta_0 + \beta_{SNP}SNP + \beta_{Smoking}Smoking + \beta_{INT}SNP \times Smoking + \sum_{v=1}^{16} \beta_{C_v}Covariate_v ,$$

where Y is natural log transformed HbA1c with an identity link $g[\cdot]$, and $Smoking$ is one of the five smoking measurements. $P_{INT} < 0.05$ was

highlighted in yellow.

² *CDC123*, *KCNQ1*, and *SLC30A8* were also associated with FG/HbA1c/diabetes according to our discovery cohort (TWB1), as listed in the above Tables S5-S7.

*** SNP rs13266634 (in the *SLC30A8* gene) is also among the 6 diabetes-associated SNPs discovered by TWB1, as listed in the above Table S7.

SNP rsID	Chr	Position	Effect allele	Other allele	Mapped gene	MAF	P_{INT} with “active smoking status” ¹	P_{INT} with “the number of pack-years” ¹	P_{INT} with “years as a smoker” ¹	P_{INT} with “packs smoked per day” ¹	P_{INT} with “hours as a passive smoker per week” ¹
rs10923931	1	120517959	T	G	NOTCH2	0.0266	0.179347	0.720227	0.427992	0.975307	0.787074
rs340874	1	214159256	C	T	PROX1	0.3866	0.80797	0.717182	0.970264	0.742546	0.880438
rs780094	2	27741237	C	T	GCKR	0.4805	0.177751	0.10087	0.052343	0.249735	0.270362
rs7578597	2	43732823	T	C	THADA	0.0052	0.045987	0.225273	0.054377	0.26392	0.548494
rs243021	2	60584819	A	G	-----	0.3388	0.961979	0.309409	0.657621	0.291143	0.837461
rs13389219	2	165528876	C	T	COBLL1	0.1028	0.896581	0.324198	0.456718	0.80187	0.622696
rs7578326	2	227020653	A	G	LOC646736	0.1510	0.401498	0.587061	0.648855	0.516609	0.425919
rs17036101	3	12277845	G	A	-----	0.0339	0.333526	0.614036	0.399858	0.56662	0.059133
rs4607103	3	64711904	C	T	ADAMTS9-AS2	0.3451	0.724647	0.745187	0.746284	0.727742	0.329993
rs11708067	3	123065778	A	G	ADCY5	0.0006	0.926012	0.920493	0.932579	0.913228	0.948399
rs459193	5	55806751	G	A	C5orf67	0.4687	0.633853	0.182762	0.48204	0.151906	0.462609
rs4457053	5	76424949	G	A	ZBED3-AS1	0.0559	0.283105	0.248024	0.128189	0.354533	0.080364
rs9472138	6	43811762	T	C	LOC107986598	0.1390	0.382594	0.592256	0.524271	0.604028	0.55964
rs2191349	7	15064309	T	G	-----	0.3098	0.89061	0.836876	0.706284	0.8688	0.182585
rs864745	7	28180556	T	C	JAZF1	0.2134	0.723531	0.41901	0.661041	0.577499	0.904833
rs972283	7	130466854	G	A	LOC105375508	0.3136	0.02176	0.061506	0.039392	0.023175	0.786707
rs516946	8	41519248	C	T	ANK1	0.1267	0.1877	0.554055	0.357718	0.348253	0.823944
rs896854	8	95960511	T	C	TP53INP1	0.2746	0.95521	0.979715	0.963165	0.857873	0.616662
rs13266634***	8	118184783	C	T	SLC30A8 ²	0.4617	0.111917	0.123203	0.017727	0.298204	0.823684
rs13292136	9	81952128	C	T	-----	0.0870	0.716811	0.682723	0.58285	0.992995	0.31485
rs2796441	9	84308948	G	A	LOC101927502	0.3837	0.705856	0.426133	0.540604	0.84844	0.668982
rs12779790	10	12328010	G	A	Near CDC123²	0.1824	0.031129	0.080318	0.134924	0.017302	0.675323
rs12571751	10	80942631	A	G	ZMIZ1	0.4227	0.298901	0.737266	0.481161	0.769763	0.401545

SNP rsID	Chr	Position	Effect allele	Other allele	Mapped gene	MAF	P_{INT} with “active smoking status” ¹	P_{INT} with “the number of pack-years” ¹	P_{INT} with “years as a smoker” ¹	P_{INT} with “packs smoked per day” ¹	P_{INT} with “hours as a passive smoker per week” ¹
rs7903146	10	114758349	T	C	<i>TCF7L2</i>	0.0236	0.337546	0.828909	0.468614	0.731794	0.430069
rs231362	11	2691471	G	A	<i>KCNQ1</i> ²	0.0958	0.425727	0.524149	0.459015	0.490526	0.839291
rs1552224	11	72433098	A	C	<i>ARAP1</i>	0.0663	0.769243	0.910252	0.66893	0.787785	0.317376
rs1387153	11	92673828	T	C	-----	0.4563	0.384234	0.306898	0.417016	0.198646	0.39214
rs10842994	12	27965150	C	T	<i>LOC105369709</i>	0.2070	0.597631	0.827573	0.502432	0.904211	0.766048
rs1153188	12	55098996	A	T	-----	0.0145	0.555762	0.943046	0.524607	0.856577	0.491495
rs1531343	12	66174894	C	G	<i>RPSAP52</i>	0.1091	0.484723	0.694244	0.524799	0.863056	0.657571
rs7961581	12	71663102	C	T	-----	0.2287	0.543227	0.640958	0.216439	0.910725	0.642499
rs7957197	12	121460686	T	A	<i>OASL</i>	0.0006	0.944019	0.942204	0.942659	0.950889	0.933128
rs7177055	15	77832762	A	G	-----	0.3341	0.458439	0.262651	0.502393	0.154147	0.315986
rs11634397	15	80432222	G	A	-----	0.0836	0.369702	0.637139	0.599593	0.613005	0.376819
rs8042680	15	91521337	A	C	<i>PRC1</i>	0.0029	0.410296	0.259713	0.229001	0.488553	0.377059
rs7202877	16	75247245	T	G	-----	0.2050	0.53142	0.17209	0.278141	0.284517	0.767809
rs12970134	18	57884750	A	G	-----	0.1738	0.274048	0.781098	0.37251	0.629756	0.444281
rs10401969	19	19407718	C	T	<i>SUGP1</i>	0.0965	0.884363	0.886034	0.825834	0.666951	0.945

Supplemental Table S20 (The 38 diabetes-associated SNPs identified from European GWASs) SNP-smoking interaction analysis on diabetes status

¹ P_{INT} is the p -value of testing $H_0: \beta_{INT} = 0$ vs. $H_1: \beta_{INT} \neq 0$ from the following model:

$$g[E(Y)] = \beta_0 + \beta_{SNP}SNP + \beta_{Smoking}Smoking + \beta_{INT}SNP \times Smoking + \sum_{v=1}^{16} \beta_{C_v}Covariate_v ,$$

where Y is the diabetes status with a logit link $g[\cdot]$, and $Smoking$ is one of the five smoking measurements. $P_{INT} < 0.05$ was highlighted in

yellow.

² *CDC123*, *KCNQ1*, and *SLC30A8* were also associated with FG/HbA1c/diabetes according to our discovery cohort (TWB1), as listed in the above Tables S5-S7.

*** SNP rs13266634 (in the *SLC30A8* gene) is also among the 6 diabetes-associated SNPs discovered by TWB1, as listed in the above Table S7.

Explanatory variables in regression model (4) ¹	Fasting glucose (mg/dL)		HbA1c (%)		Diabetes (dichotomous trait)	
	Percent change	p-value	Percent change	p-value	Odds ratio	p-value
European-derived GRS (EuGRS) (z-score standardized)	1.17	9.1E-83	0.79	1.5E-72	1.23	1.3E-30
Sex (1: female vs. 0: male)	-3.36	1.5E-123	-0.81	9.8E-15	0.70	2.7E-23
Age (in years, continuous variable)	0.36	0	0.29	0	1.08	0
Body mass index (in kg/m ² , continuous variable)	0.85	0	0.73	0	1.18	0
Active smoking status (1: yes vs. 0: no)	0.51	0.13	0.87	4.2E-4	1.21	0.06
EuGRS × active smoking status (continuous variable)	0.4836	0.0305	0.4738	0.0036	1.0782	0.1474
Drinking status (1: yes vs. 0: no)	2.08	3.3E-11	-0.91	5.9E-5	0.91	0.21
Regular exercise (1: yes vs. 0: no)	-0.90	4.7E-12	-0.98	5.7E-25	0.87	6.9E-6
Educational attainment (a value ranging from 1 to 7)	-0.52	1.5E-14	-0.20	3.9E-5	0.89	7.3E-16
R-square²	13.7 %		14.6 %		13.9 %	

Supplemental Table S21 EuGRS-smoking interaction analysis based on the TWB2 cohort

1. Natural log transformed fasting glucose (or HbA1c), or diabetes status, was regressed by model (4). European-derived GRS was calculated by $EuGRS = \sum_{j=1}^{38} w_j SNP_j$, where the weights ($w_j, j = 1, \dots, 38$) were the effect sizes reported by Said *et al.* (from eTable 5 of Said MA, Verweij N, van der Harst P. Associations of Combined Genetic and Lifestyle Risks With Incident Cardiovascular Disease and Diabetes in the UK Biobank

Study. *Jama Cardiol* 2018;3:693-702). To save space, we here omit the results of the 10 PCs, *EuGRS* × **Covariates**, and *Smoking* × **Covariates**.

- For continuous traits, R-square is the proportion of variance in natural log transformed fasting glucose (or HbA1c) that can be explained by the explanatory variables shown in model (4). For the dichotomous trait (diabetes status), we present pseudo R-square, defined as one minus the ratio of the log likelihood with intercepts only, and the log likelihood with all predictors.

Smoking measurements in regression model (4)	Fasting glucose (mg/dL)		HbA1c (%)		Diabetes (dichotomous trait)	
	Percent change (%)	P_{INT}	Percent change (%)	P_{INT}	Odds ratio	P_{INT}
EuGRS × Active smoking status	0.4836	0.0305	0.4738	0.0036	1.0782	0.1474
EuGRS × The number of pack-years	0.0188	0.0224	0.0135	0.0244	0.9986	0.3611
EuGRS × Years as a smoker	0.0182	0.0215	0.0178	0.0020	1.0010	0.5098
EuGRS × Packs smoked per day	0.4685	0.0582	0.3017	0.0939	0.9796	0.6819
EuGRS × Hours as a passive smoker per week	0.0096	0.5720	-0.0040	0.7463	0.9993	0.8559

Supplemental Table S22 EuGRS-smoking interaction analysis based on the TWB2 cohort (replacing “active smoking status” with the 4 continuous smoking measurements)

Smoking measurements in regression model (4)	Fasting glucose (mg/dL)		HbA1c (%)	
	$\hat{\Phi}_{INT}$	P_{INT}	$\hat{\Phi}_{INT}$	P_{INT}
EuGRS × Active smoking status	0.8416	0.0055	0.0351	0.0026
EuGRS × The number of pack-years	0.0351	0.0017	0.0012	0.0051
EuGRS × Years as a smoker	0.0323	0.0027	0.0014	0.0009
EuGRS × Packs smoked per day	0.8505	0.0111	0.0267	0.0383
EuGRS × Hours as a passive smoker per week	0.0189	0.4143	-1.6E-6	0.9985

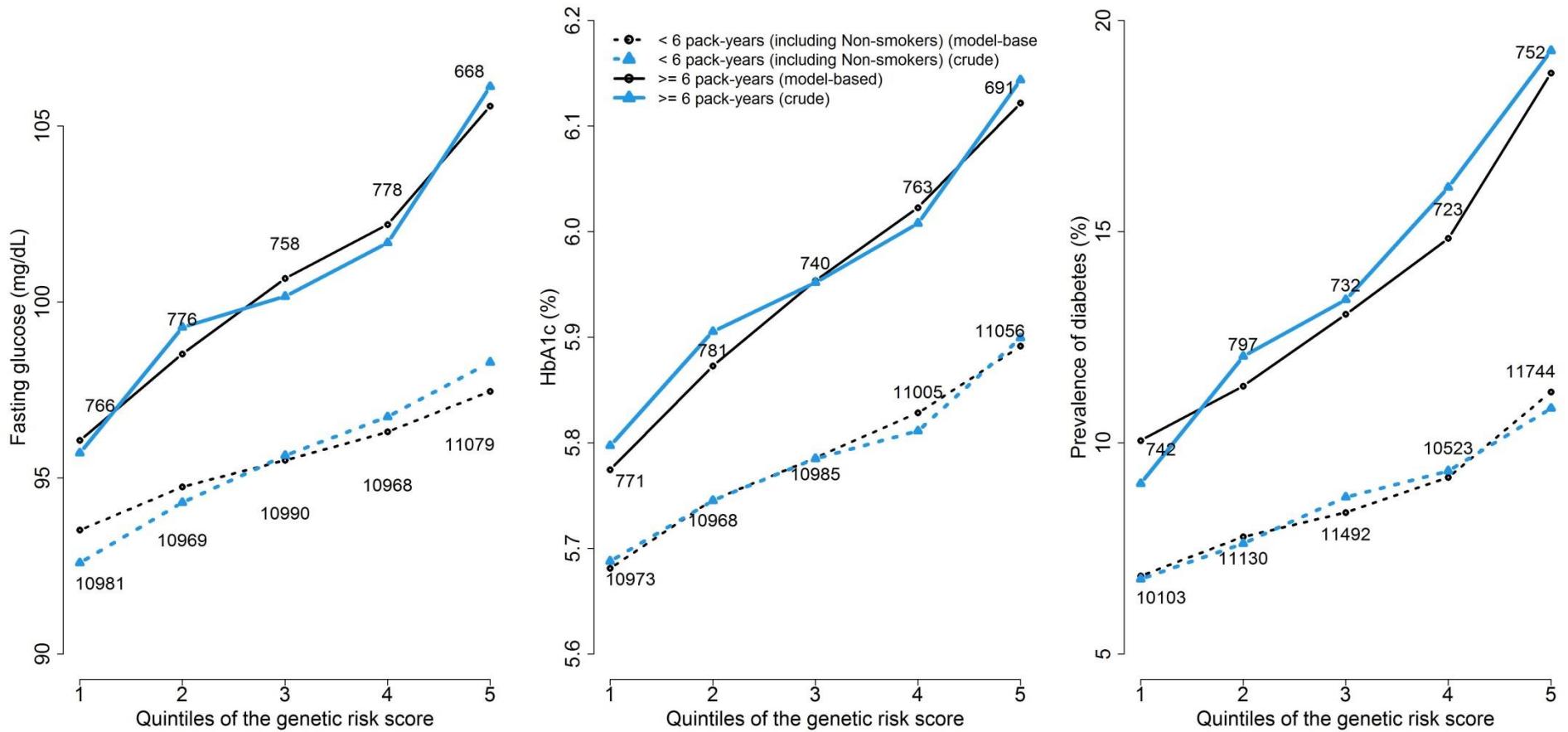
Supplemental Table S23 EuGRS-smoking interaction analysis based on the TWB2 cohort (without natural log transformation on fasting glucose or HbA1c)

Smoking measurements in regression model (4)	Fasting glucose (mg/dL)		HbA1c (%)		Diabetes (dichotomous trait)	
	Percent change (%)	P_{INT}	Percent change (%)	P_{INT}	Odds ratio	P_{INT}
EuGRS × Active smoking status	0.4780	0.0243	0.4472	0.0038	1.0637	0.1982
EuGRS × The number of pack-years	0.0264	0.0008	0.0191	0.0009	0.9991	0.5279
EuGRS × Years as a smoker	0.0227	0.0026	0.0210	0.0001	1.0012	0.4091
EuGRS × Packs smoked per day	0.6280	0.0080	0.4363	0.0114	0.9882	0.8005
EuGRS × Hours as a passive smoker per week	0.0041	0.8054	-0.0073	0.5500	0.9984	0.6701

Supplemental Table S24 EuGRS-smoking interaction analysis based on the TWB2 cohort (assess the EuGRS-smoking interaction without controlling for the other interaction terms)

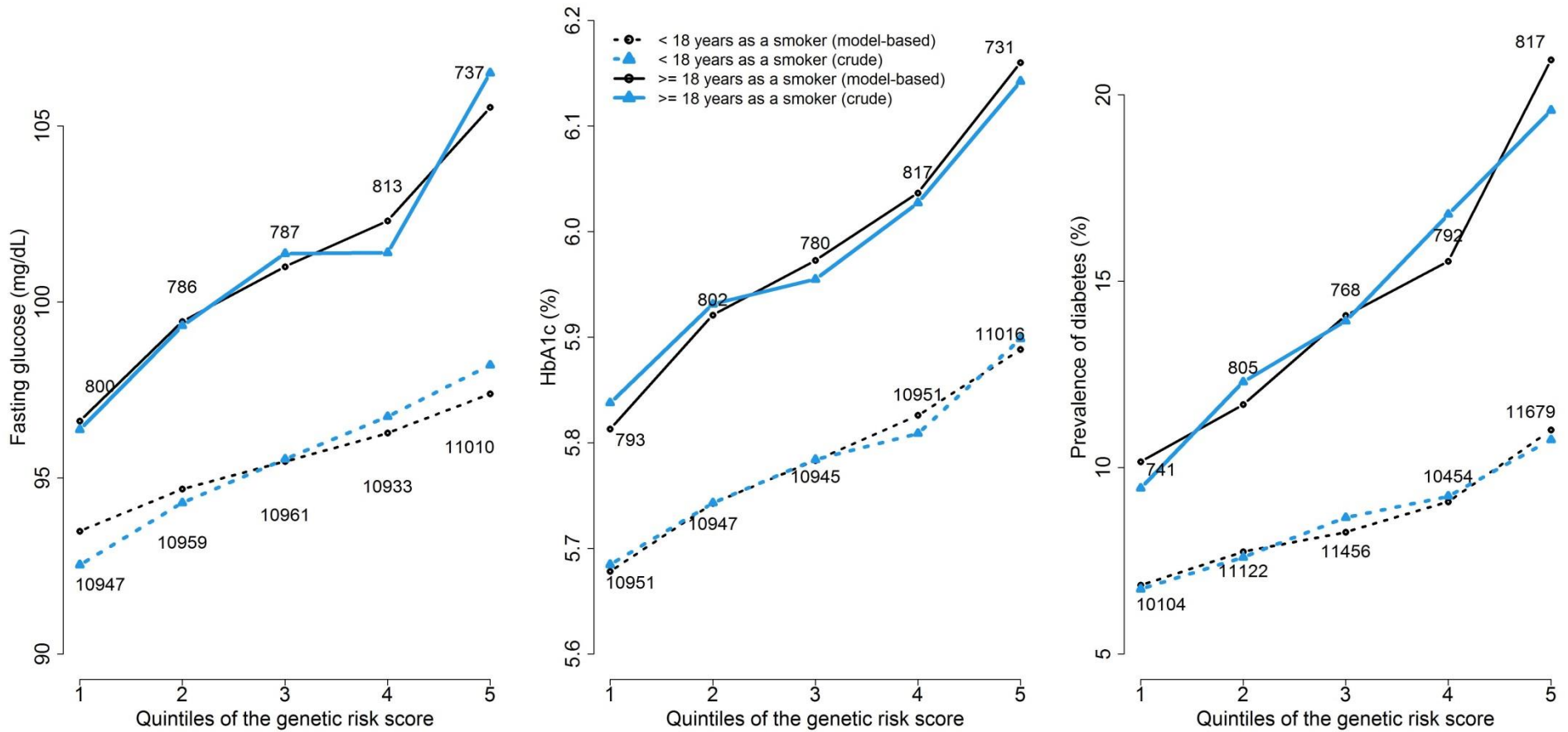
Smoking measurements in regression model (4)	Sex	Fasting glucose (mg/dL)		HbA1c (%)		Diabetes (dichotomous trait)	
		Percent change (%)	P_{INT}	Percent change (%)	P_{INT}	Odds ratio	P_{INT}
EuGRS × Active smoking status	Males	0.6478	0.0289	0.4822	0.0275	1.0625	0.2719
	Females	0.0326	0.9355	0.4313	0.1389	1.0801	0.5509
EuGRS × The number of pack-years	Males	0.0267	0.0065	0.0170	0.0185	0.9995	0.7611
	Females	-0.0449	0.0980	-0.0123	0.5314	0.9890	0.1028
EuGRS × Years as a smoker	Males	0.0224	0.0227	0.0180	0.0127	1.0013	0.4231
	Females	-0.0052	0.7794	0.0140	0.2951	0.9999	0.9782
EuGRS × Packs smoked per day	Males	0.7485	0.0148	0.4098	0.0701	0.9954	0.9286
	Females	-0.7615	0.2201	-0.1275	0.7767	0.9733	0.2943
EuGRS × Hours as a passive smoker per week	Males	0.0338	0.3495	0.0181	0.4968	1.0004	0.9446
	Females	0.0027	0.8911	-0.0129	0.3560	0.9978	0.6485

Supplemental Table S25 EuGRS-smoking interaction analysis based on the TWB2 cohort (Results of EuGRS-smoking interactions while stratified by sex)



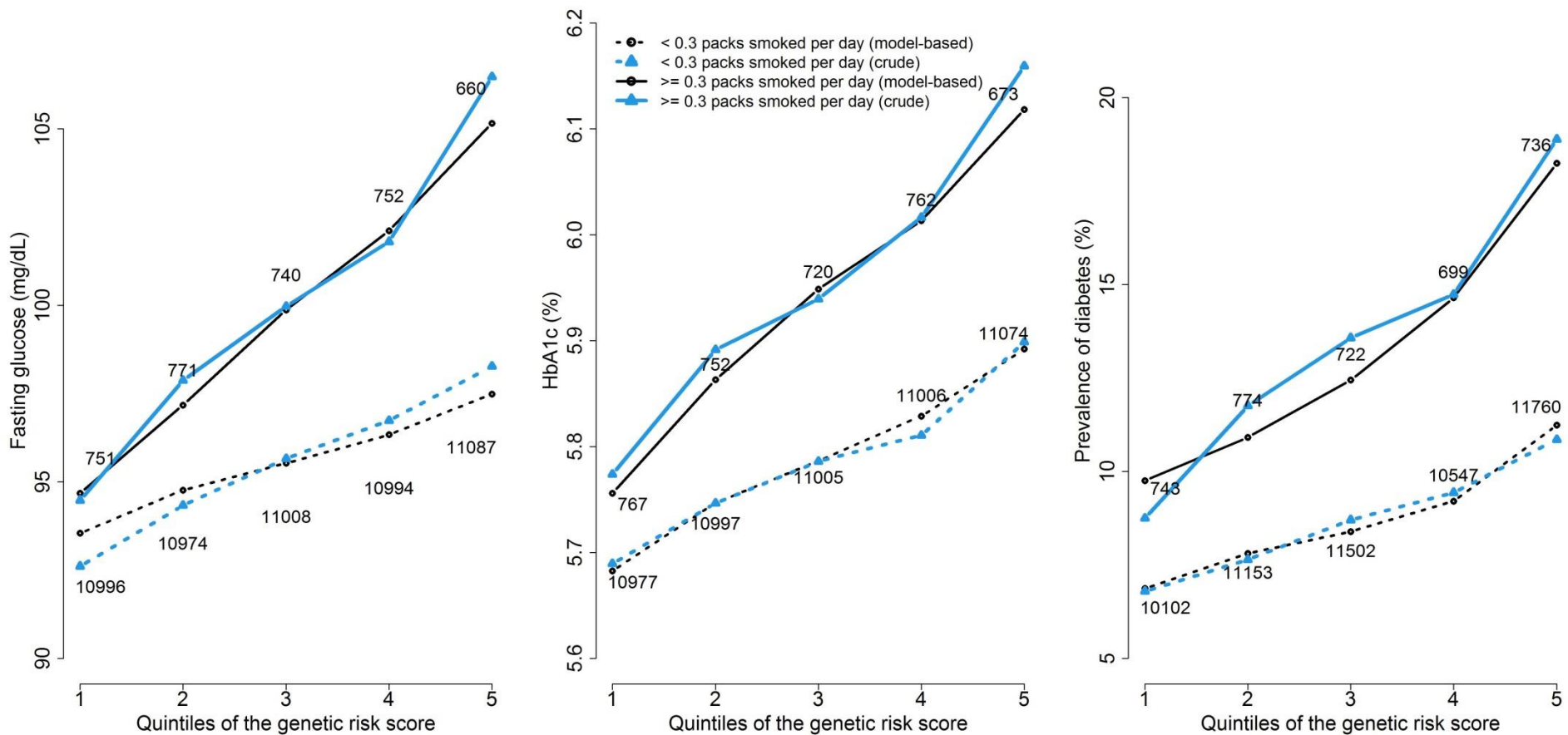
Supplemental Figure S1 Average of FG/HbA1c and the prevalence of diabetes stratified by the number of pack-years and the quintiles of the FG/HbA1c/diabetes genetic risk score

The solid lines are for smokers with ≥ 6 pack-years, where 6 is the first quartile (Q1) of the number of pack-years in TWB2 smokers (we chose Q1 as the dichotomized point, to prevent a too small sample size in the higher pack-year group). The dotted lines are for non-smokers and smokers with < 6 pack-years. The black lines depict predicted mean FG/HbA1c or predicted prevalence of diabetes based on model (4). Only subjects without any missing in covariates can be predicted. The blue lines mark crude mean FG/HbA1c or crude prevalence of diabetes, without adjusting for any covariates. The number shown around each point represents the sample size of that category.



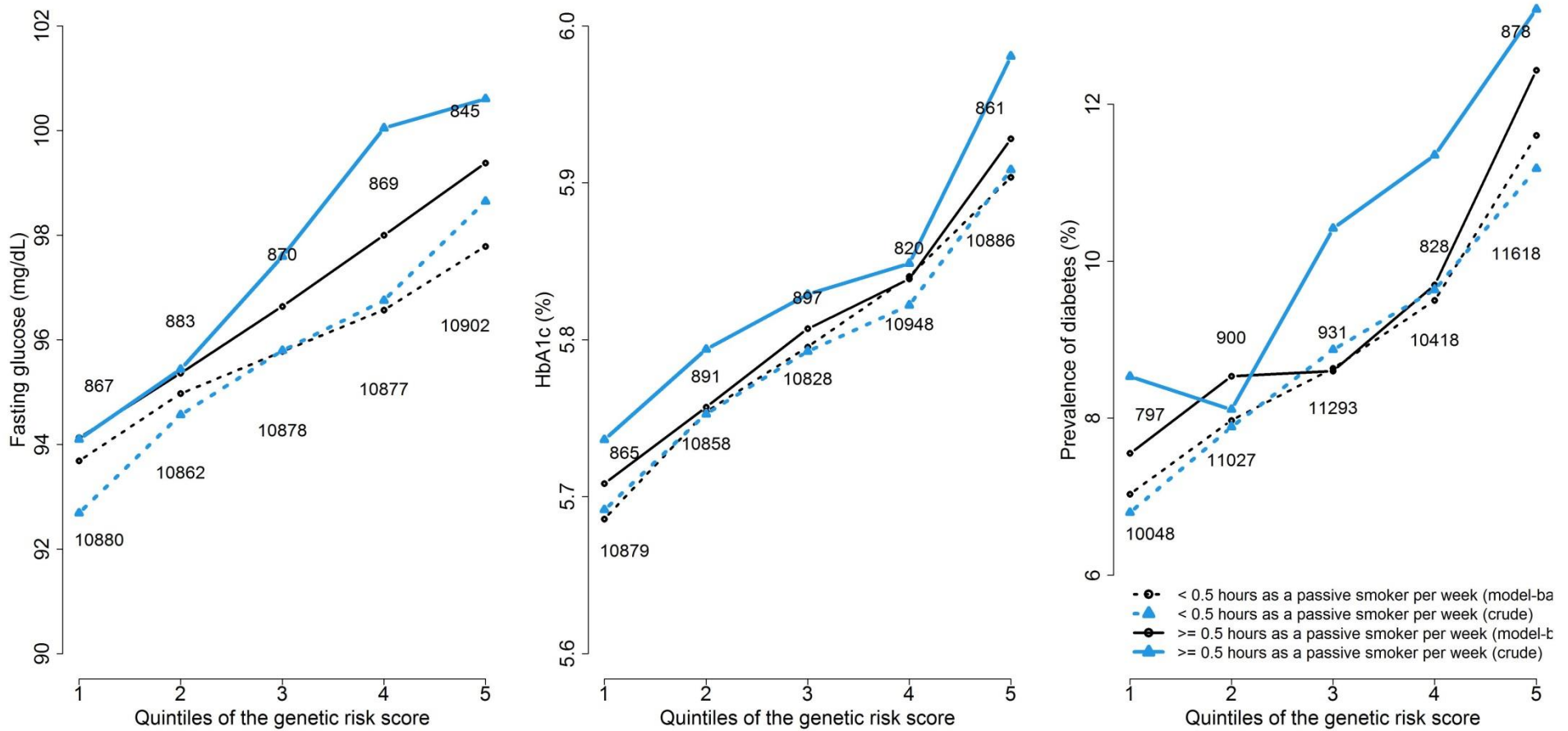
Supplemental Figure S2 Average of FG/HbA1c and the prevalence of diabetes stratified by the number of years as a smoker and the quintiles of the FG/HbA1c/diabetes genetic risk score

The solid lines are for smokers with ≥ 18 years as a smoker, where 18 is the first quartile (Q1) of the number of years as a smoker in TWB2 smokers (we chose Q1 as the dichotomized point, to prevent a too small sample size in the higher smoking-year group). The dotted lines are for non-smokers and smokers with < 18 years as a smoker. The black lines depict predicted mean FG/HbA1c or predicted prevalence of diabetes based on model (4). Only subjects without any missing in covariates can be predicted. The blue lines mark crude mean FG/HbA1c or crude prevalence of diabetes, without adjusting for any covariates. The number shown around each point represents the sample size of that category.



Supplemental Figure S3 Average of FG/HbA1c and the prevalence of diabetes stratified by the number of packs smoked per day and the quintiles of the FG/HbA1c/diabetes genetic risk score

The solid lines are for smokers with ≥ 0.3 packs smoked per day, where 0.3 is the first quartile (Q1) of the number of packs smoked per day in TWB2 smokers (we chose Q1 as the dichotomized point, to prevent a too small sample size in the higher smoking-amount group). The dotted lines are for non-smokers and smokers with < 0.3 packs smoked per day. The black lines depict predicted mean FG/HbA1c or predicted prevalence of diabetes based on model (4). Only subjects without any missing in covariates can be predicted. The blue lines mark crude mean FG/HbA1c or crude prevalence of diabetes, without adjusting for any covariates. The number shown around each point represents the sample size of that category.



Supplemental Figure S4 Average of FG/HbA1c and the prevalence of diabetes stratified by the number of hours as a passive smoker per week and the quintiles of the FG/HbA1c/diabetes genetic risk score

The solid lines are for smokers with ≥ 0.5 hours as a passive smoker per week, where 0.5 is the first quartile (Q1) of the number of hours as a passive smoker per week in TWB2 passive smokers (we chose Q1 as the dichotomized point, to prevent a too small sample size in the higher passive-smoking-amount group). The dotted lines are for non-passive-smokers and passive smokers with < 0.5 hours as a passive smoker per week. The black lines depict predicted mean FG/HbA1c or predicted prevalence of diabetes based on model (4). Only subjects without any missing in covariates can be predicted. The blue lines mark crude mean FG/HbA1c or crude prevalence of diabetes, without adjusting for any covariates. The

number shown around each point represents the sample size of that category.