

Table S1. Summary statistics of study participants (N=772)

	Female	Male	All subjects
EWAS sample set	N=298	N=277	N=575
Age Mean (SD)	46.4 (13.7)	43.3 (15.1)	44.9 (14.5)
Smoking status			
Never Smoker	226 (75.8%)	178 (64.3%)	404 (70.3%)
Age Mean (SD)	46.7 (13.9)	45.8 (16.0)	46.3 (14.8)
Current Smoker	72 (24.2%)	99 (35.7%)	171 (29.7%)
Age Mean (SD)	45.5 (13.2)	38.7 (12.3)	41.6 (13.1)
Smoker by frequency			
Occasional	31 (43.1%)	45 (45.5%)	76 (44.4%)
Regular small amount	18 (25.0%)	17 (17.2%)	35 (20.5%)
Regular medium amount	15 (20.8%)	27 (27.3%)	42 (24.6%)
Regular large amount	8 (11.1%)	10 (10.1%)	18 (10.5%)
Modeling sample sets*	N=144	N=194	N=338
Training set			
N	100	138	238
Non Smoker	50 (34.7%)	69 (35.6%)	119 (35.2%)
Age Mean (SD)	45.5 (12.6)	38.7 (12.4)	41.6 (12.9)
Current Smoker	50 (34.7%)	69 (35.6%)	119 (35.2%)
Age Mean (SD)	45.4 (13.4)	38.8 (12.4)	41.6 (13.2)
Test set			
N	44	56	100
Non Smoker	22 (15.3%)	28 (14.4%)	50 (14.8%)
Age Mean (SD)	45.9 (13.1)	37.6 (10.9)	41.2 (12.5)
Current Smoker	22 (15.3%)	28 (14.4%)	50 (14.8%)
Age Mean (SD)	45.9 (13.1)	39.4 (12.1)	42.3 (12.8)
Validation sample sets	N=97	N=100	N=197
Former Smokers			
N	82	80	162
Age Mean (SD)	51.7 (13.4)	52.0 (16.1)	51.8 (14.7)
Time since quit			
Mean (SD)	14.1 (9.9)	15.0 (11.1)	14.6 (10.5)
Missing data	27 (32.9%)	18 (22.5%)	45 (27.8)
Children			
N	15	20	35
Age Mean (SD)	5.7 (3.2)	7.1 (3.4)	6.5 (3.3)

*The sample set used for modeling purposes is part of the sample set used for the EWAS study, but was selected to maintain appropriate sex and age matching.

Table S2. Summary statistics of lifestyle habits and sociodemographic data in the studied population (N=737)

	Female (N=380)	Male (N=357)	All subjects (N=737)
Age			
Mean (SD)	47.6 (13.8)	45.3 (15.8)	46.5 (14.8)
Socioeconomic status (SES)			
Low	21 (5.5%)	8 (2.2%)	29 (3.9%)
Medium	286 (75.3%)	251 (70.3%)	537 (72.9%)
High	73 (19.2%)	98 (27.5%)	171 (23.2%)
Education level			
No college degree	126 (33.2%)	128 (35.9%)	254 (34.5%)
University degree	254 (66.8%)	229 (64.1%)	483 (65.5%)
Area of residence			
Village	62 (16.3%)	60 (16.8%)	122 (16.6%)
City of < 100,000 residents	64 (16.8%)	58 (16.2%)	122 (16.6%)
City of > 100,000 residents	254 (66.8%)	238 (66.8%)	492 (66.8%)
Missing	0 (0%)	1 (0.3%)	1 (0.1%)
Type of work			
Manual/physical (M)	37 (9.7%)	92 (25.8%)	129 (17.5%)
Mental work with partial sedentary (SM)	117 (30.8%)	88 (24.6%)	205 (27.8%)
Mental work with long sedentary (LSM)	157 (41.3%)	119 (33.3%)	276 (37.4%)
Not applicable (unemployed, other)	69 (18.2%)	58 (16.2%)	127 (17.2%)
Harmful factors in the workplace			
High temperature (HT)	13 (3.4%)	75 (21.0%)	88 (11.9%)
Low temperature (LT)	7 (1.8%)	48 (13.4%)	55 (7.5%)
Pesticides/chemicals	37 (9.7%)	47 (13.2%)	84 (11.4%)
Environmental pollutants (EP)	20 (5.3%)	64 (17.9%)	84 (11.4%)
Ionizing Radiation (IR)	13 (3.4%)	26 (7.3%)	39 (5.3%)
Sun exposure (SE)	5 (1.3%)	46 (12.9%)	51 (6.9%)
Severe stress (SS) exposure	151 (39.7%)	168 (47.1%)	319 (43.3%)
Other	85 (22.4%)	48 (13.4%)	133 (18.0%)
Smoking status			
Non-smoker	226 (59.5%)	178 (49.9%)	404 (54.8%)
Former smoker	82 (21.6%)	80 (22.4%)	162 (22.0%)
Current smoker	72 (18.9%)	99 (27.7%)	171 (23.2%)
Frequency of drinking alcohol			
Non-drinker	53 (13.9%)	35 (9.8%)	88 (11.9%)
Occasionally	296 (77.9%)	268 (75.1%)	564 (76.5%)
Frequently (at least 3-4 time x week)	30 (7.9%)	52 (14.6%)	82 (11.1%)
Missing	1 (0.3%)	2 (0.6%)	3 (0.4%)

	Female (N=380)	Male (N=357)	All subjects (N=737)
BMI			
Mean (SD)	25.5 (4.8)	27.5 (4.1)	26.5 (5.1)
Missing	3 (0.8%)	2 (0.6%)	5 (0.7+++%)
Frequency of physical activity			
None	68 (17.9%)	41 (11.5%)	109 (14.8%)
Once a month	28 (7.4%)	22 (6.2%)	50 (6.8%)
Once a week	88 (23.2%)	55 (15.4%)	143 (19.4%)
2-3 times a week	133 (35.0%)	152 (42.6%)	285 (38.7%)
Everyday	63 (16.6%)	87 (24.4%)	150 (20.4%)
Physical activity (PA)			
Active	284 (74.7%)	294 (82.4%)	578 (78.4%)
Inactive	96 (25.3%)	63 (17.6%)	159 (21.6%)
Type of regularly practised sport			
Aerobic	151 (39.7%)	170 (47.6%)	321 (43.6%)
Strengthening	65 (17.1%)	82 (23.0%)	147 (19.9%)
Yoga and Balance	55 (14.5%)	13 (3.6%)	68 (9.2%)
Not applicable	109 (28.7%)	92 (25.8%)	201 (27.3%)
Sleep hours			
Mean (SD)	6.7 (1.2)	6.7 (1.1)	6.7 (1.1)
≤ 7 hrs	279 (73.4%)	266 (74.5%)	545 (73.9%)
> 7 hrs	101 (26.6%)	91 (25.5%)	192 (26.1%)
≤ 8 hrs	365 (96.1%)	345 (91.0%)	710 (96.3%)
> 8 hrs	15 (3.9%)	12 (3.2%)	27 (3.7%)
Diet			
Varied	340 (89.5%)	326 (91.3%)	666 (90.4%)
Vegetarian	17 (4.5%)	6 (1.7%)	23 (3.1%)
Protein	4 (1.1%)	16 (4.5%)	20 (2.7%)
Other	18 (4.7%)	8 (2.2%)	26 (3.5%)
Missing	1 (0.3%)	1 (0.3%)	2 (0.3%)
Meals per day			
2 to 3	82 (21.6%)	101 (28.3%)	183 (24.8%)
3 to 4	205 (53.9%)	177 (49.6%)	382 (51.8%)
>4	92 (24.2%)	379 (22.1%)	171 (23.2%)
Missing	1 (0.3%)	0 (0%)	1 (0.1%)
Portions of fruits per day			
0	22 (5.8%)	44 (12.3%)	66 (9.0%)
1 to 2	290 (76.3%)	268 (75.1%)	558 (75.7%)
3 to 4	60 (15.8%)	37 (10.4%)	97 (13.2%)
≥5	8 (2.1%)	4 (1.1%)	12 (1.6%)
Missing	0 (0%)	4 (1.1%)	4 (0.5%)

	Female (N=380)	Male (N=357)	All subjects (N=737)
Portions of vegetables per day			
0	11 (2.9%)	17 (4.8%)	28 (3.8%)
1 to 2	247 (65.0%)	266 (74.5%)	513 (69.6%)
3 to 4	112 (29.5%)	66 (18.5%)	178 (24.2%)
>=5	9 (2.4%)	6 (1.7%)	15 (2.0%)
Missing	1 (0.3%)	2 (0.6%)	3 (0.4%)
Cups of coffee per day			
0	60 (15.8%)	81 (22.7%)	141 (19.1%)
1 to 2	247 (65.0%)	166 (46.5%)	413 (56.0%)
3 to 4	61 (16.1%)	94 (26.3%)	155 (21.0%)
>=5	11 (2.9%)	15 (4.2%)	26 (3.5%)
Missing	1 (0.3%)	1 (0.3%)	2 (0.3%)
Frequency of meat consumption			
Never	19 (5.0%)	8 (2.2%)	27 (3.6%)
1-2 times a week	83 (21.8%)	27 (7.6%)	110 (14.9%)
3-4 times a week	90 (23.7%)	53 (14.8%)	143 (19.4%)
Frequently, but not every day	96 (25.3%)	105 (29.4%)	201 (27.3%)
Everyday	92 (24.2%)	164 (45.9%)	256 (34.7%)
Frequency of fish consumption			
Never	7 (1.8%)	6 (1.7%)	13 (1.8%)
Rarely	133 (35.0%)	117 (32.8%)	250 (33.9%)
Sometimes	167 (43.9%)	173 (48.5%)	340 (46.1%)
Frequently	73 (19.2%)	61 (17.1%)	134 (18.2%)

Table S8. Increase in the value of the AUC parameter with the inclusion of subsequent markers in the smoking status prediction model.

CpG introduced into the model	Training AUC	Test AUC
1 - cg05575921	0.75	0.76
2 - cg09594361	0.77	0.77
3 - cg21322436	0.78	0.79
4 - cg09842685	0.80	0.80

Table S10. Analysis of the association between lifestyle behaviors and DNA methylation at *AHRR* in a group of N=737 individuals. Results were adjusted for age, sex, blood cell count, smoking status and DNAm pack-years.

Lifestyle/demographic factor	Test for an association between lifestyle and DNAm at the <i>AHRR</i>			
	B	Std. Error	Std. B	p
Age	-0.001	0.0002	-0.155	1.89E-05
Sex Male	-0.031	0.006	-0.174	1.67E-06
Smoking				
Current smoker vs. never smoker	-0.104	0.007	-0.516	4.71E-41
Current smoker vs. former smoker	-0.092	0.011	-0.424	6.24E-15
Former smoker vs. never smoker	-0.028	0.005	-0.221	6.65E-08
Education				
University/No	0.002	0.003	0.010	0.595
Socioeconomic status (SES)				
High vs. low	-0.010	0.010	-0.044	0.325
High vs. medium	-0.004	0.004	-0.020	0.278
Residence-village				
City >100K vs. village	-0.006	0.004	-0.025	0.169
Type of work				
Mental vs. manual/physical work	-0.006	0.005	-0.032	0.189
Harmful factors in the workplace				
High temperature (HT)	-0.004	0.005	-0.014	0.455
Low temperature (LT)	-0.003	0.006	-0.010	0.591
Pesticides/chemicals	0.006	0.005	0.021	0.237
Environmental pollutants (EP)	0.001	0.005	0.003	0.879
Ionizing Radiation (IR)	0.007	0.007	0.017	0.340
Sun exposure (SE)	-0.001	0.006	-0.002	0.908
Severe stress (SS) exposure	0.004	0.003	0.023	0.201
Lifestyle Risk Factors				
Sleep hours $\geq$ 8 hrs	0.021	0.010	0.045	0.027
Physical Activity (PA)	0.009	0.004	0.040	0.028
PA freq. (None)				
Once a month	-0.001	0.008	-0.003	0.930
Once a week	0.010	0.005	0.051	0.047
Every day	0.012	0.005	0.058	0.028
Yoga	-0.010	0.006	-0.034	0.066
Aerobic exercises	0.002	0.004	0.013	0.530
Strengthening	0.005	0.005	0.028	0.323
BMI	0.000	0.000	0.018	0.343
Alcohol-Non				
Occasionally	-0.011	0.005	-0.042	0.027
Frequently	-0.015	0.007	-0.087	0.046
Diet-Variied				
Vegetarian	-0.006	0.010	-0.011	0.573
Protein	-0.006	0.010	-0.012	0.552
Meat (Never)				

1-2 times a week	-0.019	0.010	-0.097	0.051
3-4 times a week	-0.012	0.009	-0.054	0.184
Frequently	-0.014	0.008	-0.037	0.090
Fish				
Frequently vs. rarely	-0.001	0.003	-0.005	0.764
Portion of Veg. (0)				
1 to 2	-0.007	0.008	-0.018	0.398
>=3	0.00	0.009	-0.001	0.979
Portion of Fruit (0)				
1 to 2	0.002	0.006	0.006	0.766
>=3	0.002	0.008	0.010	0.817
Cups of coffee (0)				
1 to 2	-0.001	0.004	-0.007	0.765
3 to 4	-0.007	0.005	-0.040	0.157
>=5	-0.009	0.009	-0.039	0.352