Mendelian randomization study of the association between telomere length and risk of cancer and non-neoplastic diseases

The Telomeres Mendelian Randomization Collaboration

Correspondence: Philip Haycock

MRC Integrative Epidemiology Unit

University of Bristol

Bristol

UK

philip.haycock@bristol.ac.uk

Tel: +44 1173 310 088

2995 words [word limit 3000]

3 figures, 2 tables, 132 references; 7 supplementary figures / 6 supplementary tables

ABSTRACT 349 WORDS

Importance The causal direction and magnitude of the association between telomere length and incidence of cancer and non-neoplastic diseases is uncertain, due to the susceptibility of observational studies to confounding and reverse causation.

Objective To conduct a Mendelian randomization study, using germline genetic variants as instrumental variables, to appraise the causal relevance of telomere length for risk of cancer and non-neoplastic diseases.

Data Sources Genome-wide association studies (GWAS) published up to January 15 2015.

Study Selection GWAS of non-communicable diseases that assayed germline genetic variation and did not select cohort or control participants on the basis of pre-existing diseases. Of 163 GWAS of non-communicable diseases identified, summary data from 103 were available.

Data Extraction Summary association statistics for single nucleotide polymorphisms (SNPs) that are strongly associated with telomere length in the general population.

Main Outcomes Odds ratios (ORs) for disease per standard deviation (SD) higher telomere length due to germline genetic variation.

Results Summary data were available for 35 cancers and 48 non-neoplastic diseases, corresponding to 420,081 cases (median 2,526 per disease) and 1,093,105 controls (median 6,789 per disease). Increased telomere length due to germline genetic variation was generally associated with increased risk for site-specific cancers. The strongest associations were observed for (ORs per 1-SD change in genetically increased telomere length): glioma 5.27 (3.15-8.81), serous low-malignant-potential ovarian cancer 4.35 (2.39-7.94), lung adenocarcinoma 3.19 (2.40-4.22), neuroblastoma 2.98 (1.92-4.62), bladder cancer 2.19 (1.32-3.66), melanoma 1.87 (1.55-2.26), testicular cancer 1.76 (1.02-3.04), kidney cancer 1.55

(1.08-2.23) and endometrial cancer 1.31 (1.07-1.61). Associations were stronger for rarer cancers and at tissue sites with lower rates of stem cell division ($P<0.05$). There was generally little evidence of association between genetically increased telomere length and risk of psychiatric, autoimmune, inflammatory, diabetic and other non-neoplastic diseases, except for coronary heart disease (0.78 [0.67-0.90]), abdominal aortic aneurysm (0.63 [0.49-0.81]), celiac disease (0.42 [0.28-0.61]) and interstitial lung disease (0.09 [0.05- 0.15]).

Conclusions It is likely that longer telomeres increase risk for several cancers but reduce risk for some non-neoplastic diseases, including cardiovascular diseases.

INTRODUCTION

At the ends of chromosomes, telomeres are DNA-protein structures that protect the genome from damage, shorten progressively over time in most somatic tissues[1] and are proposed physiological markers of ageing.[2,3] Shorter leukocyte telomeres are correlated with older age, male sex and other known risk factors for non-communicable diseases[4–6] and are generally associated with higher risk for cardiovascular diseases[7,8], type 2 diabetes[9] and non-vascular non-neoplastic causes of mortality.[8] Whether these associations are causal, however, is unknown. Telomere length has also been implicated in risk of cancer but the direction and magnitude of the association is uncertain and contradictory across observational studies.[10–14] The uncertainty reflects the considerable difficulty of designing observational studies of telomere length and cancer incidence that are robust to reverse causation, confounding and measurement error.

The aim of the present report was to conduct a Mendelian randomization study, using germline genetic variants as instrumental variables for telomere length, to help clarify the nature of the association between telomere length and risk of cancer and non-neoplastic diseases. The approach, which mimics the random allocation of individuals to the placebo and intervention arms of a randomized controlled trial, allowed us to: (1) estimate the direction and broad magnitude of the association of telomere length with risk of multiple cancer and non-neoplastic diseases; (2) appraise the evidence for causality in the estimated etiological associations; (3) investigate potential sources of heterogeneity in findings for site-specific cancers; and (4) compare genetic estimates to findings based on directly measured telomere length in prospective observational studies.

METHODS

Study design

The design of our study, illustrated in Figure S1, had three key components: 1) the identification of genetic variants to serve as instruments for telomere length; 2) the acquisition of summary data for the genetic instruments from genome wide association studies (GWASs) of diseases and risk factors for non-communicable diseases; and 3) the classification of diseases and risk factors into primary or secondary outcomes based on *a priori* statistical power. As a first step, we searched the GWAS catalog[15,16] on the 15 January 2015, to identify single nucleotide polymorphisms (SNPs) associated with telomere length. To supplement the list with additional potential instruments, we also searched the original study reports curated by the GWAS catalog (using a P-value threshold of $5x10^{-8}$).[17–25] We acquired summary data for all SNPs identified by our search from a meta-analysis of GWASs of telomere length, involving 9,190 participants of European ancestry.[18]

The second key component of our design strategy involved the acquisition of summary data, corresponding to the selected genetic instruments for telomere length, from GWASs of non-communicable diseases and risk factors (Fig. S1). As part of this step, we invited principal investigators of non-communicable disease studies curated by the GWAS catalog[15,26] to share summary data for our study (see Fig. S1 for further details). We also downloaded summary data for diseases and risk factors from publically available sources, including study-specific websites, dbGAP, ImmunoBase and the GWAS catalog (Fig. S1).

The third key component of our design strategy was the classification of diseases and risk factors into either primary or secondary outcomes, which we defined on the basis of *a priori* statistical power to detect associations with telomere length. Primary outcomes were defined

as diseases with sufficient cases and controls for >50% statistical power and secondary outcomes defined as diseases with <50% statistical power to detect odds ratios ≥2.0 per standard deviation (SD) change in genetically increased telomere length (alpha assumed to be 0.01). All risk factors were defined as secondary outcomes. Risk factors with <50% statistical power were excluded.

Further details on our design strategy can be found in the supplement.

Comparison with prospective observational studies

We searched PubMed for prospective observational studies of the association between telomere length and disease (see Tables S3 and S4 for details of the search strategy and inclusion criteria). Study-specific relative risks for disease per unit change or quantile comparison of telomere length were transformed to a SD scale using previously described methods.[27] Hazard ratios, risk ratios and odds ratios were assumed to approximate the same measure of relative risk. Where multiple independent studies of the same disease were identified, these were combined by fixed effects meta-analysis, unless there was strong evidence of between-study heterogeneity ($P_{Cochran's\ Q}<0.001$), in which case they were kept separate.

Statistical analysis

We combined summary data across SNPs into a single instrument, using maximum likelihood to estimate the slope of the relationship between β_{GD} and β_{GP} and a variance-covariance matrix to make allowance for linkage disequilibrium between SNPs,[28] where β_{GD} is the change in disease log odds or risk factor levels per copy of the effect allele and β_{GP} is the SD change in telomere length per copy of the effect allele (see supplementary methods

144 for technical details). The slope from this approach can be interpreted as the log odds ratio for
145 binary outcomes, or the unit change for continuous risk factors, per SD change in genetically
146 increased telomere length. P-values for heterogeneity amongst SNPs, in the estimated
147 associations of genetically increased telomere length with disease and risk factors, were
148 estimated by likelihood ratio tests.[28] Associations between genetically increased telomere
149 length and continuous risk factors were transformed into SD units. For five secondary disease
150 outcomes where only a single SNP was available for analysis, we estimated associations
151 using the Wald ratio: $\beta_{GD}/\beta_{GP,}$ with standard errors approximated by the delta method.[29]

152 Inference of causality in the estimated etiological associations between telomere length and
153 disease depends on satisfaction of Mendelian randomization assumptions (Fig. S7; see Table
154 S6 for a glossary of terms).[30,31] The assumptions are: 1) the selected SNPs are associated with
155 telomere length; 2) the selected SNPs are not associated with confounders; and 3) the selected
156 SNPs are associated with disease exclusively through their effect on telomere length. If these
157 assumptions are satisfied, the selected SNPs are valid instrumental variables and their
158 association with disease can be interpreted as a causal effect of telomere length. We modeled
159 the impact of violations of these assumptions through two sets of sensitivity analyses: a
160 weighted median function[32] and MR-Egger regression[30] (see supplementary methods for
161 technical details). We restricted our sensitivity analyses to diseases showing the strongest
162 evidence of association with genetically increased telomere length (defined as
163 $P_{Bonferroni} \leq 0.05$).
164
165 We used meta-regression to appraise potential sources of heterogeneity in our findings for
166 cancer. The association of genetically increased telomere length with the log odds of cancer
167 was regressed on cancer incidence, survival time and median age-at-diagnosis, downloaded
168 from the National Cancer Institute Surveillance, Epidemiology, and End Results (SEER)

Program,[33] and tissue-specific rates of stem cell division from Tomasetti and Vogelstein.[34] As the downloaded cancer characteristics from SEER correspond to the United States population, 77% of which was of white ancestry in 2015[35], the meta-regression analyses excluded genetic studies conducted in East Asian populations.

All analyses were performed in R version 3.1.2[36] and Stata release 13.1 (StataCorp, College Station, TX). P-values were two-sided and evidence of association was declared at $P<0.05$. Where indicated, Bonferroni corrections were used to make allowance for multiple testing, although this is likely to be overly conservative given the non-independence of many of the outcomes tested.

RESULTS

We selected 16 SNPs as instruments for telomere length (Fig. S1 & Table 1). The selected SNPs correspond to 10 independent genomic regions that collectively account for 2-3% of the variance in leukocyte telomere length, which is equivalent to an F statistic of ~18. This indicates that the genetic instrument, constructed from these 10 independent genomic regions, is strongly associated with telomere length (details in supplementary discussion).[37] Summary data for the genetic instruments were available for 83 non-communicable diseases, corresponding to 420,081 cases (median 2,526 per disease) and 1,093,105 controls (median 6,789 per disease), and 44 risk factors (Fig. S1, Table 2 and Table S1). The median number of SNPs available across diseases was 11 (min=1, max=13) and across risk factors was 12 (min=11, max=13). Of the 83 diseases, 56 were classified as primary outcomes and 27 as secondary outcomes (Table 2, Fig. S1 and Table S1). For 9 of the 83 non-communicable diseases, additional summary data were available from 10 independent studies for replication

194 analyses, corresponding to 40,465 cases (median 1,416 per disease) and 52,306 controls
195 (median 3,537 per disease) (Table S1).

196 The results from primary analyses of non-communicable diseases are presented in Figure 1;
197 results from secondary analyses of risk factors and diseases with low *a priori* power are
198 presented in the supplement (Fig. S2, S5 and S6). Genetically increased telomere length was
199 associated with higher odds of disease for 9 of 22 primary cancers ($P<0.05$), including (odds
200 ratio [95% confidence interval]): glioma (5.27 [3.15-8.81]), endometrial cancer (1.31 [1.07-
201 1.61]), kidney cancer (1.55 [1.08-2.23]), testicular germ cell cancer (1.76 [1.02-3.04]),
202 melanoma (1.87 [1.55-2.26]), bladder cancer (2.19 [1.32-3.66]), neuroblastoma (2.98 [1.92-
203 4.62]), lung adenocarcinoma (3.19 [2.40-4.22]) and serous low-malignancy-potential (LMP)
204 ovarian cancer (4.35 [2.39-7.94]) (Fig. 1). The associations were, however, highly variable
205 across cancer types, varying from an odds ratio of 0.86 (0.50-1.48) for head and neck cancer
206 to 5.27 (3.15-8.81) for glioma. Substantial variability was also observed within tissue sites.
207 For example, the odds ratio for lung adenocarcinoma was 3.19 (2.40-4.22) compared to 1.07
208 (0.82-1.39) for squamous cell lung cancer. For serous LMP ovarian cancer the odds ratio was
209 4.35 (2.39-7.94) compared to odds ratios of 1.21 (0.87-1.68) for endometrioid ovarian cancer,
210 1.12 (0.94-1.34) for serous invasive ovarian cancer, 1.04 (0.66-1.63) for clear cell ovarian
211 cancer and 1.04 (0.73-1.47) for mucinous ovarian cancer. The strongest evidence of
212 association was observed for glioma, lung adenocarcinoma, neuroblastoma and serous LMP
213 ovarian cancer ($P_{Bonferroni}<0.05$). Results for glioma and bladder cancer showed evidence for
214 replication in independent datasets (independent datasets were not available for other
215 cancers) (Fig. S3).

216 Genetically increased telomere length was associated with reduced odds of disease for 6 of 32
217 primary non-neoplastic diseases ($P<0.05$), including coronary heart disease (0.78 [0.67-0.9]),
218 abdominal aortic aneurysm (0.63 [0.49-0.81]), Alzheimer's disease (0.84 [0.71-0.98]), celiac

disease (0.42 [0.28-0.61]), interstitial lung disease (0.09 [0.05-0.15]) and type 1 diabetes (0.71 [0.51-0.98]) ($P<0.05$) (Figure 1). The strongest evidence of association was observed for coronary heart disease ($P_{Bonferroni}=0.05$) and abdominal aortic aneurysm, celiac disease and interstitial lung disease ($P_{Bonferroni}<0.05$). The associations with coronary heart disease and interstitial lung disease showed evidence for replication in independent datasets (Fig. S3).

Our genetic findings were generally similar in direction and magnitude to estimates based on observational prospective studies of leukocyte telomere length and disease (Figure 3). Our genetic estimates for lung adenocarcinoma, melanoma, kidney cancer and glioma, were, however, stronger in comparison to observational estimates.

In sensitivity analyses, we appraised the potential impact of confounding by pleiotropic pathways on our results. Associations estimated by the weighted median and MR-Egger were broadly similar to the main results for glioma, lung adenocarcinoma, serous LMP ovarian cancer, neuroblastoma, abdominal aortic aneurysm, coronary heart disease and interstitial lung disease (Fig. S4). In the second set of sensitivity analyses, implemented by MR-Egger regression, we found little evidence for the presence of pleiotropy ($P_{intercept}\geq0.27$) (Fig. S4). The MR-Egger analyses were, however, generally underpowered, as reflected by the wide confidence intervals in the estimated odds ratios.

In meta-regression analyses, we observed that genetically increased telomere length tended to be more strongly associated with rarer cancers ($P=0.02$) and cancers at tissue-sites with lower rates of stem cell division ($P=0.02$) (Figure 2). The associations showed little evidence of varying by percentage survival five years after diagnosis or median age-at-diagnosis ($P\geq37$).

DISCUSSION

In this report we show that genetically increased telomere length is associated with increased risk of several cancers and with reduced risk of some non-neoplastic diseases. Given the random distribution of genotypes in the general population with respect to lifestyle and other environmental factors, as well as the fixed nature of germline genotypes, these results should be less susceptible to confounding and reverse causation in comparison to observational studies. Our results are therefore compatible with causality. On the other hand, our results could reflect violations of Mendelian randomization assumptions, such as confounding by pleiotropy, population stratification or ancestry.[38] Although we cannot entirely rule out this possibility, the majority of our results persisted in sensitivity analyses that made allowance for violations of Mendelian randomization assumptions. Confounding by population stratification or ancestry is also unlikely, given the adjustments made for ancestry in the disease GWASs (see supplementary discussion).

Comparison with previous studies

Our findings for cancer are generally contradictory to those based on retrospective studies, which tend to report increased risk for cancer in individuals with shorter telomeres.[11,12,39–42] The contradictory findings may reflect reverse causation in the retrospective studies, whereby shorter telomeres arise as a result of disease, or of confounding effects, e.g. due to cases being slightly older than controls even in age-matched analyses. Our findings for cancer are generally more consistent with those based on prospective observational studies, which tend to report weak or null associations of longer leukocyte telomeres with overall and site-specific risk of cancer,[10–13,41,43–62] with some exceptions.[63] Our results are also similar to previously reported Mendelian randomization studies of telomere length and risk of

melanoma, lung cancer, chronic lymphocytic leukemia and glioma.[64–67] The shape of the association with cancer may not, however, be linear over the entire telomere length distribution. For example, individuals with dyskeratosis congenita, a disease caused by germline loss-of-function mutations in the telomerase component genes *TERC* and *TERT,* have chronically short telomeres and are at increased risk of some cancers, particularly acute myeloid leukemia and squamous cell carcinomas arising at sites of leukoplakia,[68,69] presumably due to increased susceptibility to genome instability and chromosomal end-to-end fusions.[70] Our results should therefore be interpreted as reflecting the average association at the population level and may not be generalizable to the extreme ends of the telomere length distribution.

Mechanisms of association

Our cancer findings are compatible with known biology.[70] By limiting the proliferative potential of cells, telomere shortening may serve as a tumour suppressor; and individuals with longer telomeres may be more likely to acquire somatic mutations owing to increased proliferative potential.[70] Rates of cell division are, however, highly variable amongst tissues[34] and thus the relative gain in cell proliferative potential, conferred by having longer telomeres, may also be highly variable across tissues. This could explain the ~6-fold variation in odds ratios observed across cancer types in the present study, as well as the tendency of our results to be stronger at tissue sites with lower rates of stem cell division. For example, the association was strongest for glioma (OR=5.27) and comparatively weak for colorectal cancer (OR=1.09) and the rates of stem cell division in the tissues giving rise to these cancers differ by several orders of magnitude. In neural stem cells, which give rise to gliomas, the number of divisions is ~270 million and for colorectal stem cells is ~1.2 trillion over the average lifetime of an individual.[34] The observation that genetically increased telomere

length was more strongly associated with rarer cancers potentially reflects the same mechanism, since rarer cancers also tend to show lower rates of stem cell division.[34] For example, the incidence of glioma is 0.4 and for colorectal cancer is 42.4 per 100,000 per year in the United States.[33]

The inverse associations observed for some non-neoplastic diseases may reflect the impact of telomere shortening on tissue degeneration and an evolutionary trade-off for greater resistance to cancer at the cost of greater susceptibility to degenerative diseases, particularly cardiovascular diseases.[71,72]

Study limitations

Our study is subject to some limitations, in addition to the Mendelian randomization assumptions already considered above. First, our method assumes that the magnitude of the association between SNPs and telomere length is consistent across tissues. Second, our study assumed a linear shape of association between telomere length and disease risk, whereas the shape could be "J" or "U" shaped.[44,57,68] Third, our results assume that the samples used to define the genetic instrument for telomere length[18] and the various samples used to estimate the SNP-disease associations are representative of the same general population, practically defined as being of similar ethnicity, age and sex distribution.[73] This assumption would, for example, not apply in the case of the SNP-disease associations derived from East Asian or pediatric populations. Generally speaking, violation of the aforementioned assumptions could bias the magnitude of the association between genetically increased telomere length and disease; but would be unlikely to increase the likelihood of false positives (i.e. incorrectly inferring an association when none exists).[74] Our results should therefore remain informative for the direction and broad magnitude of the average association at the population level, even

in the presence of such violations. Fourth, we cannot rule out chance in explaining some of the weaker findings. Fifth, our results may not be fully representative of non-communicable diseases (since not all studies shared data and our analyses were underpowered for the secondary disease outcomes). The diseases represented in our primary analyses probably account for >60% of all causes of death in American adults.[75]

Clinical relevance of findings

Our findings suggest that potential clinical applications of telomere length, e.g. as a tool for risk prediction or as an intervention target for disease prevention, may have to consider a trade-off in risk between cancer and non-neoplastic diseases. For example, a number of companies have been established that offer telomere length measurement services to the public (via a requesting physician), under the claim that shorter telomeres are a general indicator of poorer health status and older biological age and that such information can be used to motivate healthy lifestyle choices in individuals. However, the conflicting direction of association between telomere length and risk of cancer and non-neoplastic diseases, indicated by our findings, suggests that such services to the general public may be premature.

Conclusion

It is likely that longer telomeres increase risk for several cancers but reduce risk for some non-neoplastic diseases, including cardiovascular diseases. Further research is required to resolve whether telomere length is a useful predictor of risk that can help guide therapeutic interventions, to clarify the shape of any dose-response relationships and to characterise the nature of the association in population subgroups.

342 The Telomeres Mendelian Randomization Collaboration

343

344 Philip Haycock[1]; Stephen Burgess[2]; Aayah Nounu[1]; Jie Zheng[1]; George N Okoli[3]; Jack
345 Bowden[1,4]; Kaitlin Wade[1]; Nicholas Timpson[1]; David M. Evans[1,5]; Peter Willeit[2,6]; Abraham
346 Aviv[7]; Tom R. Gaunt[1]; Gibran Hemani[1]; Massimo Mangino[8,9]; Hayley Patricia Ellis[10];
347 Kathreena Mary Kurian[10]; Karen Pooley[11] on behalf of the BCAC and OCAC consortia;
348 Rosalind Eeles[12] on behalf of the PRACTICAL consortium; Jeffrey E Lee[13]; Shenying
349 Fang[13]; Wei Chen[13]; Matthew H Law[14], Lisa M Bowdler[15] and Mark M Iles[16] on behalf of the
350 Melanoma meta-analysis consortium; Qiong Yang[17], Bradford B. Worrall[18] and Hugh
351 Stephen Markus[19] on behalf of the METASTROKE project of the ISGC; Rayjean J. Hung[20,21]
352 and Chris I Amos[22] on behalf of the ILCCO consortium; Amanda Spurdle[23], Deborah J
353 Thompson[24] and Tracy O'Mara[23] on behalf of the ECAC consortium; Brian Wolpin[25], Laufey
354 Amundadottir[26] and Rachael Stolzenberg-Solomon[27] on behalf of the PanScan consortium;
355 Antonia Trichopoulou[29,30], Charlotte Onland-Moret[31], Eiliv Lund[32], Eric J Duell[33], Federico
356 Canzian[34], Gianluca Severi[35,36,37,38], Kim Overvad[39], Marc J Gunter[40], Rosario Tumino[41] and
357 Ulrika Svenson[42] on behalf of EPIC; Andre van Rij[43], Annette F Baas[44], Matthew J Bown[45],
358 Nilesh J Samani[45], Paul IW de Bakker[44], Femke NG van t'Hof[44], Gerard Tromp[46,47], Gregory
359 T Jones[43], Helena Kuivaniemi[46,47] and James R Elmore[48] on behalf of the Aneurysm
360 Consortium; Mattias Johansson[49]; James Mckay[49]; Ghislaine Scelo[49]; Robert Carreras-
361 Torres[49]; Valerie Gaborieau[49]; Paul Brennan[49]; Paige M. Bracci[50], Rachel E Neale[15], Sara H
362 Olson[51], Steven Gallinger[20], Donghui Li[52], Gloria M. Petersen[54], Harvey Risch[55], and Alison
363 P. Klein[56] on behalf of PanC[4]; Jiali Han[57,58]; Christian C. Abnet[59]; Neal D. Freedman[59]; Philip
364 R. Taylor[59]; John M Maris[60]; Katja K Aben[61,62]; Lambertus A Kiemeney[61]; Sita H
365 Vermeulen[61]; John K Wiencke[63,64]; Kyle M Walsh[63,64]; Margaret Wrensch[63,64]; Terri Rice[63];
366 Clare Turnbull[65]; Kevin Litchfield[66]; Lavinia Paternoster[1] and Marie Standl[67] on behalf of the
367 EAGLE consortium; Gonçalo R Abecasis[68]; John Paul SanGiovanni[69]; Lars G Fritsche[68];
368 Yong Li[70] and Vladan Mijatovic[71] on behalf of the CKDGen consortium; Yadav Sapkota[15];
369 Siew-Kee Low[72]; Krina T Zondervan[73,74]; Grant W Montgomery[15]; Dale R. Nyholt[75,15];
370 David A van Heel[76]; Karen Hunt[76]; Dan E. Arking[77], Foram N. Ashar[77] and Nona
371 Sotoodehnia[78] on behalf of the CHARGE-Sudden Cardiac Arrest Working Group; Daniel
372 Woo[79]; Jonathan Rosand[80]; Mary Comeau[81]; W. Mark Brown[82]; Edwin K. Silverman[83], John
373 E Hokanson[84] and Michael Cho[83] on behalf of COPDGene; Jennie Hui[85,86,87,88], Manuel
374 Ferreira[15] and Philip J. Thompson[89] on behalf of the AAGC consortium; Alanna C.

Morrison[90], Janine F Felix[91] and Nicholas L Smith[92] on behalf of the CHARGE-Heart Failure Working Group; Angela M Christiano[93]; Lynn Petukhova[94]; Regina C. Betz[95]; Xing Fan[96]; Xuejun Zhang[96]; Caihong Zhu[96]; Carl Langefeld[97]; Susan D. Thompson[98]; Feijie Wang[99]; Xu Lin[99,100]; David A. Schwartz[101]; Tasha Fingerlin[102]; Jerome I. Rotter[103,104], Mary Frances Cotch[105] and Richard A Jensen on behalf of the CHARGE-Eye Working Group[106,107]; Matthias Munz[108], Henrik Dommisch[108] and Arne S Schaefer[108] on behalf of the European Periodontitis Genetics Group; Fang Han[109]; Hanna M Ollila[110]; Ryan P. Hillary[110]; Omar Albagha[111]; Stuart H. Ralston[112]; Chenjie Zeng[113]; Wei Zheng[113]; Xiao-Ou Shu[113]; Andre Reis[114]; Steffen Uebe[114]; Ulrike Hüffmeier[114]; Yoshiya Kawamura[115], Takeshi Otowa[116,117] and Tsukasa Sasaki[118] on behalf of the Japanese Collaboration Team for GWAS of Panic Disorder; Martin Lloyd Hibberd[119]; Michael Levin[120]; Sonia Davila[121]; Gang Xie1[22,20]; Katherine Siminovitch[122,20]; Jin-Xin Bei[123]; Yi-Xin Zeng[123,124]; Asta Försti[125,126]; Bowang Chen[125]; Stefano Landi[127]; Andre Franke[128]; Annegret Fischer[128,129]; David Ellinghaus[128]; Carlos Flores[130,131]; Imre Noth[132]; Shwu-Fan Ma[132]; Jia Nee Foo[133]; Jianjun Liu[133]; Jong-Won Kim[134]; David G. Cox[135]; Olivier Delattre[136]; Olivier Mirabeau[136]; Christine F. Skibola[137]; Clara S. Tang[138]; Merce Garcia-Barcelo[138]; Kai-Ping Chang[139]; Wen-Hui Su[140]; Yu-Sun Chang[141]; Nicholas G Martin[15]; Scott Gordon[15]; Tracey Wade[142]; Chaeyoung Lee[143]; Michiaki Kubo[144]; Pei-Chieng Cha[145]; Yusuke Nakamura[146]; Daniel Levy[147]; Masayuki Kimura[7]; Shih-Jen Hwang[147]; Steven Hunt[148]; Tim Spector[8]; Nicole Soranzo[149]; Ani W Manichaikul[150]; R Graham Barr[151]; Bratati Kahali[152], Elizabeth Speliotes[152] and Laura M Yerges-Armstrong[153] on behalf of the GOLD Consortium; Ching-Yu Cheng[154,155,156], Jost B. Jonas[157,158] and Tien Yin Wong[154,155,156] on behalf of the SEED consortium; Isabella Fogh[159], Kuang Lin[159] and John F. Powell[159] on behalf of the SLAGEN and ALSGEN consortia; Kenneth Rice[160] on behalf of the ICBP; Caroline Relton[1]; Richard M Martin[1,3,161]; George Davey Smith[1]

[1] MRC Integrative Epidemiology Unit, University of Bristol, Bristol, UK
[2] Department of Public Health and Primary Care, University of Cambridge
[3] School of Social and Community Medicine, University of Bristol, Bristol, UK
[4] MRC Biostatistics Unit, Cambridge, UK.
[5] University of Queensland Diamantina Institute, Translational Research Institute, Brisbane, Queensland, Australia
[6] Department of Neurology, Innsbruck Medical University, Austria
[7] Center of Human Development and Aging, Department of Pediatrics, New Jersey Medical School, Rutgers, The State University of New Jersey
[8] Department of Twin Research and Genetic Epidemiology, King's College London, London UK
[9] NIHR Biomedical Research Centre at Guy's and St. Thomas' Foundation Trust, London, UK
[10] Brain Tumour Research Group, Institute of Clinical Neuroscience, Learning and Research Building, Southmead Hospital, University of Bristol
[11] Strangeways Research Laboratory, University of Cambridge, Cambridge, UK
[12] The Institute of Cancer Research, London, UK
[13] Department of Surgical Oncology, The University of Texas MD Anderson Cancer Center, Houston, TX.
[14] Statistical Genetics, QIMR Berghofer Medical Research Institute, Brisbane, Australia
[15] QIMR Berghofer Medical Research Institute, Brisbane, Australia
[16] Section of Epidemiology and Biostatistics, Leeds Institute of Cancer and Pathology, University of Leeds, Leeds, UK
[17] Department of Biostatistics, Boston University School of Public Health, Boston, Massachusetts, United States of America and the Framingham Heart Study, Framingham, Massachusetts, United States of America
[18] Departments of Neurology and Public Health Sciences University of Virginia Charlottesville, Virginia 22908
[19] Department of Clinical Neurosciences, University of Cambridge, UK
[20] Lunenfeld-Tanenbaum Research Institute of Mount Sinai Hospital, Toronto, Ontario, Canada
[21] Division of Epidemiology, Dalla Lana School of Public Health, University of Toronto, 60 Murray St. Rm L5-215, Box 18, Toronto, ON M5T 3L9, Canada
[22] Geisel School of Medicine, Dartmouth College
[23] Genetics and Computational Biology Division, QIMR Berghofer Medical Research Institute, Brisbane, QLD 4006, Australia
[24] Centre for Cancer Genetic Epidemiology, Department of Public Health and Primary Care, University of Cambridge, Cambridge, CB1 8RN, UK.
[25] Dana-Farber Cancer Institute
[26] Laboratory of Translational Genomics, Division of Cancer Epidemiology & Genetics, National Cancer Institute
[27] Division of Cancer Epidemiology and Genetics, National Cancer Institute, NIH, DHHS.

[29] Hellenic Health Foundation, Athens, Greece
[30] WHO Collaborating Center for Nutrition and Health, Unit of Nutritional Epidemiology and Nutrition in Public Health, Dept. of Hygiene, Epidemiology and Medical Statistics, University of Athens Medical School, Greece
[31] Dept of Epidemiology, Julius Center for Health Sciences and Primary Care, University Medical Center Utrecht
[32] Institute of community medicine, UiT The Artcic University of Norway
[33] Unit of Nutrition and Cancer, Cancer Epidemiology Research Program, Bellvitge Biomedical Research Institute (IDIBELL), Catalan Institute of Oncology (ICO), Avda Gran Via 199-203, 08908 L'Hospitalet de Llobregat, Barcelona, Spain
[34] Genomic Epidemiology Group, German Cancer Research Center (DKFZ), Heidelberg, Germany
[35] Université Paris-Saclay, Université Paris-Sud, UVSQ, CESP, INSERM, Villejuif, France
[36] Gustave Roussy, F-94805, Villejuif, France
[37] Human Genetics Foundation (HuGeF), Torino, Italy

[38] Cancer Council Victoria and University of Melbourne, Australia

[39] Department of Public Health, Section for Epidemiology, Aarhus University, Aarhus, Denmark

[40] School of Public Health, Imperial College London, London W2 1PG

[41] Cancer Registry, Azienda Ospedaliera "Civile M.P.Arezzo", via Dante 109, Ragusa, IT

[42] Department of Medical Biosciences, Umea University, Umea, Sweden

[43] Surgery Department, University of Otago, Dunedin, New Zealand

[44] Department of Medical Genetics, Center for Molecular Medicine and Department of Epidemiology, Julius Center for Health Sciences and Primary Care, University Medical Center Utrecht, Utrecht, The Netherlands

[45] The Department of Cardiovascular Sciences and the NIHR Leicester Cardiovascular Biomedical Research Unit, University of Leicester, Leicester, LE2 7LX, UK.

[46] Division of Molecular Biology and Human Genetics, Department of Biomedical Sciences, Faculty of Medicine and Health Sciences, Stellenbosch University, Tygerberg, 7505, South Africa

[47] The Sigfried and Janet Weis Center for Research, Geisinger Health System, Danville, PA, USA

[48] Department of Vascular and Endovascular Surgery, Geisinger Health System, Danville, PA 17822, USA

[49] Genetic Epidemiology Group, International Agency for Research on Cancer, Lyon, France

[50] Department of Epidemiology and Biostatistics, University of California San Francisco, San Francisco, California

[51] Department of Epidemiology and Biostatistics, Memorial Sloan Kettering Cancer Center, New York, New York, USA.

[52] Department of Gastrointestinal Medical Oncology, University of Texas MD Anderson Cancer Center, Houston, Texas, USA

[54] Department of Health Sciences Research, Mayo Clinic College of Medicine, Rochester, Minnesota, USA

[55] Yale School of Public Health & Yale School of Medicine & Yale Cancer Center, 60 College St., PO Box 208034, New Haven, CT 06520-8034

[56] Departments of Oncology, Pathology and Epidemiology, Johns Hopkins School of Medicine, Baltimore Maryland 21231

[57] Department of Epidemiology, Fairbanks School of Public Health, Indiana University

[58] Indiana University Melvin and Bren Simon Cancer Center, Indianapolis, IN, USA

[59] Division of Cancer Epidemiology and Genetics, National Cancer Institute, Rockville MD USA

[60] Children's Hospital of Philadelphia, Perelman School of Medicine at the University of Pennsylvania

[61] Radboud university medical center, Radboud Institute for Health Sciences, Nijmegen, The Netherlands

[62] Netherlands Comprehensive Cancer Organization, Utrecht, The Netherlands

[63] Department of Neurological Surgery, University of California, San Francisco, San Francisco, CA

[64] Institute of Human Genetics, University of California, San Francisco, San Francisco, CA

[65] William Harvey Research Institute, Queen Mary University, London, UK

[66] Division of Genetics and Epidemiology, The Institute of Cancer Research, London, UK

[67] Institute of Epidemiology I, Helmholtz Zentrum München - German Research Center for Environmental Health, Neuherberg, Germany

[68] Center for Statistical Genetics, Department of Biostatistics, University of Michigan, Ann Arbor, Michigan, USA

[69] National Eye Institute, Division of Epidemiology and Clinical Research Clinical Trials Branch National Institutes of Health, Bethesda, USA

[70] Department of Internal Medicine IV, University Hospital Freiburg

[71] Department of Life and Reproduction Sciences, University of Verona

[72] Laboratory of Statistical Analysis, Centre for Integrative Medical Sciences, The Institute of Physical and Chemical Research (RIKEN], Yokohama, Japan

[73] Genetic and Genomic Epidemiology Unit, Wellcome Trust Centre for Human Genetics, University of Oxford, Oxford, UK
[74] Nuffield Department of Obstetrics and Gynecology, University of Oxford, John Radcliffe Hospital, Oxford, UK
[75] Institute of Health and Biomedical Innovation, Queensland University of Technology, Brisbane, Australia
[76] Blizard Institute, Barts and The London School of Medicine and Dentistry, Queen Mary University of London, London E1 2AT, UK
[77] McKusick-Nathans Institute of Genetic Medicine, Johns Hopkins University School of Medicine, Baltimore, MD USA
[78] Division of Cardiology and Cardiovascular Health Research Unit, Department of Medicine, University of Washington, Seattle, Washington 98101
[79] University of Cincinnati College of Medicin, Department of Neurology, Cincinnati, OH, USA 45267
[80] Massachusetts General Hospital, Neurology, Center for Human Genetic Research, MA, USA
[81] Center for Public Health Genomics, Department of Biostatistical Sciences, Division of Public Health Sciences, Wake Forest School of Medicine, Medical Center Blvd, Winston-Salem, NC 27157
[82] Department of Biostatistical Sciences, Division of Public Health Sciences, Wake Forest School of Medicine, Medical Center Blvd, Winston-Salem, NC 27157
[83] Channing Division of Network Medicine, Brigham and Women's Hospital, Boston, MA 02115
[83] Channing Division of Network Medicine, Brigham and Women's Hospital, Boston, MA, 02115
[84] Department of Epidemiology, University of Colorado Anschutz Medical Campus, Aurora, Colorado, USA
[85] Busselton Population Medical Research Institute Inc, Sir Charles Gairdner Hospital, Perth, Australia
[86] PathWest Laboratory Medicine of Western Australia (WA), Perth, Australia
[87] School of Pathology and Laboratory Medicine, University of WA, Perth, Australia
[88] School of Population Health, University of WA, Perth, Australia
[89] The Lung Health Clinic and Institute for Respiratory Health, University of Western Australia, Perth, Australia
[90] Department of Epidemiology, Human Genetics, and Environmental Sciences, University of Texas Health Science Center at Houston, Houston, TX 77030 USA
[91] Department of Epidemiology, Erasmus MC, University Medical Center Rotterdam, Rotterdam, the Netherlands
[92] Department of Epidemiology, University of Washington, Seattle WA 98101 USA
[93] Departments of Dermatology and Genetics & Development, Columbia University, New York, NY, US
[94] Departments of Dermatology and Epidemiology, Columbia University, New York, NY, US
[95] Institute of Human Genetics, University of Bonn, Bonn D-53127, Germany
[96] Institute of Dermatology & Department of Dermatology, First Affiliated Hospital of Anhui Medical University
[97] Director, Center for Public Health Genomics, Department of Biostatistical Sciences, Division of Public Health Sciences, Wake Forest School of Medicine, Medical Center Blvd, Winston-Salem, NC 27157
[98] Center for Autoimmune Genomics and Etiology, Cincinnati Children's Hospital Medical Center, Department of Pediatrics,University of Cincinnati College of Medicine, Cincinnati, OH, USA
[99] Institute for Nutritional Sciences, SIBS, Chinese Academy of Sciences, Shanghai, 200031, PR China
[100] Key Laboratory of Nutrition and Metabolism, Chinese Academy of Sciences, Shanghai, 200031, PR China
[101] University of Colorado, 12631 East 17th Avenue, B178, Aurora, CO 80045
[102] Department of Biomedical Research, National Jewish Health Hospital
[103] Institute for Translational Genomics and Population Sciences, Los Angeles Biomedical Research Institute at Harbor-UCLA Medical Center
[104] Departments of Pediatrics and Medicine, 1124 W. Carson Street, Harbor-UCLA Medical Center, Torrance, CA 90502

[105] Epidemiology Branch, Division of Epidemiology and Clinical Applications, NIH Intramural Research Program, National Eye Institute, National Institutes of Health, Clinical Research Center 3A2521,
[106] Cardiovascular Health Research Unit, University of Washington, Seattle, Washington, USA
[107] Department of Medicine, University of Washington, Seattle, Washington, USA
[108] Charité – University Medicine Berlin, CC 03, Institute of Dental, Oral and Maxillary Medicine, Dept. of Periodontology and Synoptic Dentistry, Aßmannshauser Str. 4-6, 14197 Berlin, Germany
[109] Department of Pulmonary Medicine, Peking University People's Hospital, 100044 Beijing, China
[110] Stanford University, Center for Sleep Sciences, Palo Alto, CA, USA
[111] Centre for Genomic and Experimental Medicine, Institute of Genetics and Molecular Medicine, University of Edinburgh, Western General Hospital, Edinburgh EH4 2XU, UK
[112] Institute of Genetics and Molecular Medicine, University of Edinburgh, Edinburgh, UK
[113] Division of Epidemiology, Department of Medicine, Vanderbilt Epidemiology Center, Vanderbilt-Ingram Cancer Center, Vanderbilt University Medical Center, Nashville, Tennessee
[114] Institute of Human Genetics, Friedrich-Alexander-Universität Erlangen-Nürnberg, Erlangen, Germany
[115] Department of Psychiatry, Shonan Kamakura General Hospital, Kanagawa, Japan
[116] Department of Neuropsychiatry, Graduate School of Medicine, The University of Tokyo, Tokyo, Japan
[117] Graduate School of Clinical Psychology, Teikyo Heisei University Major of Professional Clinical Psychology, Tokyo, Japan
[118] Department of Physical and Health Education, Graduate School of Education, The University of Tokyo, Tokyo, Japan
[119] Infectious Diseases, Genome Institute of Singapore, Singapore
[120] Division of Infectious diseases, Department of medicine, Imperial College London, UK
[121] Human genetics, Genome Institute of Singapore, Singapore
[122] Departments of Medicine, Immunology, Molecular Genetics, University of Toronto
[123] Sun Yat-sen University Cancer Center, State Key Laboratory of Oncology in South China, Collaborative Innovation Center for Cancer Medicine Guangzhou 510060, P. R. China
[124] Peking Union Medical College, Beijing 100730, P.R. China
[125] Molecular Genetic Epidemiology, German Cancer Research Center (DKFZ), Heidelberg, Germany
[126] Center for Primary Health Care Research, Clinical Research Center, Lund University, Malmö, Sweden
[127] Department of Biology, University of Pisa, Pisa, Italy
[128] Institute of Clinical Molecular Biology, Christian-Albrechts-University of Kiel, Kiel, Germany.
[129] University Hospital Schleswig-Holstein, Kiel, Germany
[130] Research Unit, Hospital Universitario N.S. de Candelaria, Tenerife, Spain
[131] CIBER de Enfermedades Respiratorias, Instituto de Salud Carlos III, Madrid, Spain
[132] Section of Pulmonary and Critical Care Medicine, University of Chicago, 5841 S. Maryland Ave., Chicago IL 60637-6076
[133] Human Genetics, Genome Institute of Singapore, A*STAR, Singapore 138672, Singapore
[134] Dept. of Laboratory Medicine and Genetics, Samsung Medical Center, Sungkyunkwan, University School of Medicine, Ilwon-dong 50, Gangnam-gu, Seoul, Korea, 135-710
[135] Cancer Research Center of Lyon, INSERM U1052, Lyon, France
[136] Inserm U830, Institut Curie, PSL University, 26 rue d'Ulm, 75248 Paris Cedex 05 France, France.
[137] Department of Epidemiology, University of Alabama at Birmingham. 1665 University Boulevard, Birmingham, AL 35294-0022, USA
[138] Department of Surgery, Li Ka Shing Faculty of Medicine, The University of Hong Kong, Hong Kong, SAR China
[139] Department of Otolaryngology - Head and Neck Surgery, Chang Gung Memorial Hospital at Lin-Kou, Taoyuan, Taiwan,
[140] Department of Biomedical Sciences, Graduate Institute of Biomedical Sciences, College of Medicine, Molecular Medicine Research Center, Chang Gung University, Taoyuan, Taiwan

[141] Molecular Medicine Research Center, Chang Gung University, Taoyuan, Taiwan
[142] School of Psychology, Flinders University
[143] School of Systems Biomedical Science, Soongsil University, 369 Sangdo-ro, Dongjak-gu, Seoul 156-743, Korea
[144] RIKEN Center for Integrative Medical Science, 1-7-22, Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, JAPAN
[145] Division of Molecular Brain Science, Kobe University Graduate School of Medicine, 7-5-1 Kusunoki-chou, Chuo-ku, Kobe 650-0017, Japan
[146] Center for Personalized Therapeutics, The University of Chicago, 900E 57th Street, Chicago IL 60637 USA
[147] The NHLBI's Framingham Heart Study, Framingham, MA, Population Sciences Branch of the National Heart, Lung, and Blood Institute, Bethesda, MD.
[148] Department of Genetic Medicine, Weill Cornell Medicine in Qatar, Doha, Qatar
[149] Human Genetics, Wellcome Trust Sanger Institute, Genome Campus, Hinxton Cambridge
[150] Center for Public Health Genomics, Department of Public Health Sciences, University of Virginia, Charlottesville, VA USA
[151] Department of Medicine and Department of Epidemiology, Columbia University Medical Center, New York, NY 10032, USA
[152] Department of Internal Medicine, Division of Gastroenterology and Department of Computational Medicine and Bioinformatics, University of Michigan, Ann Arbor, MI 48109, USA
[153] Department of Medicine, University of Maryland, Baltimore, MD 21201, USA
[154] Singapore Eye Research Institute, Singapore National Eye Center, Singapore 168751, Singapore
[155] Department of Ophthalmology, National University of Singapore and National University Health System, Singapore
[156] Duke-National University of Singapore Graduate Medical School, Singapore
[157] Beijing Institute of Ophthalmology, Beijing Tongren Eye Center, Beijing Tongren Hospital, Capital Medical University, Beijing Ophthalmology and Visual Science Key Lab, Beijing, China
[158] Department of Ophthalmology, Medical Faculty Mannheim of the Ruprecht-Karls-University Heidelberg, Mannheim, Germany
[159] Department of Basic and Clinical Neuroscience, Maurice Wohl Clinical Neuroscience Institute, Institute of Psychiatry, Psychology and Neuroscience, King's College London, London, United Kingdom
[160] Dept of Biostatistics, University of Washington, Seattle, USA
[161] University of Bristol / University Hospitals Bristol NHS Foundation Trust National Institute for Health Research Bristol Nutrition Biomedical Research Unit, Bristol, UK

Affiliations of the Telomeres Mendelian Randomization Collaboration

Acknowledgements

This work was supported by CRUK grant number C18281/A19169 (the Integrative Cancer Epidemiology Programme). Dr Haycock is supported by CRUK Population Research Postdoctoral Fellowship C52724/A20138. The MRC Integrative Epidemiology Unit is supported by grants MC_UU_12013/1 and MC_UU_12013/2. Dr Martin is supported by the

National Institute for Health Research (NIHR), the Bristol Nutritional Biomedical Research Unit and the University of Bristol.

We gratefully acknowledge all the studies and databases that made GWAS summary data available (see supplementary materials for detailed acknowledgements): **AC** (the aneurysm consortium), **ALSGEN (**the International Consortium on Amyotrophic Lateral Sclerosis Genetics), **AMD Gene (**Age-related Macular Degeneration Gene Consortium), **BCAC** (Breast Cancer Association Consortium), **C4D** (Coronary Artery Disease Genetics Consortium), **CARDIoGRAM** (Coronary ARtery DIsease Genome wide Replication and Meta-analysis), **CHARGE-HF** (Cohorts for Heart and Aging Research in Genomic Epidemiology Consortium – Heart Failure Working Group), **COPDGene** (The Genetic Epidemiology of Chronic Obstructive Pulmonary Disease), **CORECT** (ColoRectal Transdisciplinary Study), **CKDGen** (Chronic Kidney Disease Genetics consortium), **dbGAP** (database of Genotypes and Phenotypes), **DIAGRAM** (DIAbetes Genetics Replication And Meta-analysis), **EAGLE** (EArly Genetics & Lifecourse Epidemiology Eczema Consortium, excluding 23andMe), **ECAC** (Endometrial Cancer Association Consortium), **EGG** (Early Growth Genetics Consortium), **EPG** (European Periodontitis Genetics Group), **GABRIEL** (A Multidisciplinary Study to Identify the Genetic and Environmental Causes of Asthma in the European Community), **GCAN** (Genetic Consortium for Anorexia Nervosa), **GECCO** (Genetics and Epidemiology of Colorectal Cancer Consortium), **GIANT** (Genetic Investigation of ANthropometric Traits), **GLGC** (Global Lipids Genetics Consortium), **GUGC** (Global Urate and Gout consortium), **ICBP** (International Consortium for Blood Pressure), **IGAP** (International Genomics of Alzheimer's Project), **HPFS** (Health Professionals Follow-Up Study), **JCTGPD** (Japanese Collaboration Team for GWAS of Panic Disorder), **ILCCO** (International Lung Cancer Consortium), **ImmunoBase** (genetic

433 database of immunologically related human diseases), **IMSGC** (International Multiple
434 Sclerosis Genetic Consortium), **IIBDGC** (International Inflammatory Bowel Disease
435 Genetics Consortium); **KIDRISK** (Kidney cancer consortium), **MAGIC** (Meta-Analyses of
436 Glucose and Insulin-related traits Consortium), **MC** (the melanoma meta-analysis
437 consortium), **MESA** (Multi-Ethnic Study of Atherosclerosis), **METASTROKE/ISGC**
438 (METASTROKE project of the International Stroke Genetics Consortium), **NBCS** (Nijmegen
439 Bladder Cancer Study), **NHGRI-EBI GWAS catalog** (National Human Genome Research
440 Institute and European Bioinformatics Institute Catalog of published genome-wide
441 association studies), **NHS** (Nurses' Health Study), **OCAC** (Ovarian Cancer Association
442 Consortium), **PanScan** (Pancreatic Cancer Cohort Consortium), **PGC** (Psychiatric Genomics
443 Consortium), **PRACTICAL** (Prostate Cancer Association Group to Investigate Cancer
444 Associated Alterations in the Genome), **SEEDS** (the Singapore Epidemiology of Eye
445 Diseases Study), **SLAGEN** (Italian Consortium for the Genetics of Ayotrophic Lateral
446 Sclerosis), **SSGAC (**Social Science Genetics Association Consortium), **TAG** (Tobacco and
447 Genetics Consortium), **T1Dbase** (type 1 diabetes database), **TICG** (Tourette International
448 Collaborative-Genetics); **TSAICG** (Tourette Syndrome Association International Consortium
449 for Genetics).

450 We gratefully acknowledge the assistance and contributions of Dr Julia Gumy, Ms Lisa
451 Wright, Dr Georg B. Ehret (ICBP), Dr Louise V. Wain (ICBP), Dr Caroline Fox (CKDGen),
452 Dr Stephan Ripke (IIBDGC), Dr Jimmy Liu (IIBDGC), Dr Carl Anderson (IIBDGC) and Dr
453 Jeremiah Scharf (TSAICG and TICG).

Table 1. Single nucleotide polymorphisms associated with telomere length

SNPs	Chr	Pos	Gene	EA	OA	EAF*	Beta*	SE*	P-value*	Phet*	No. studies*	Sample size*	Discovery p-value	% variance explained	Discovery study
rs11125529	2	54248729	*ACYP2*	A	C	0.16	0.065	0.012	0.000606	0.313	6	9177	8.00E-10	0.080	Codd[21]
rs6772228	3	58390292	*PXK*	T	A	0.87	0.041	0.014	0.049721	0.77	6	8630	3.91E-10	0.200	Pooley[17]
rs12696304	3	169763483	*TERC*	C	G	0.74	0.090	0.011	5.41E-08	0.651	6	9012	4.00E-14	0.319	Codd[22]
rs10936599	3	169774313	*TERC*	C	T	0.76	0.100	0.011	1.76E-09	0.087	6	9190	3.00E-31	0.319	Codd[21]
rs1317082	3	169779797	*TERC*	A	G	0.71	0.097	0.011	4.57E-09	0.029	6	9176	1.00E-08	0.319	Mangino[18]
rs10936601	3	169810661	*TERC*	C	T	0.74	0.087	0.011	8.64E-08	0.433	6	9150	4.00E-15	0.319	Pooley[17]
rs7675998	4	163086668	*NAF1*	G	A	0.80	0.048	0.012	0.008912	0.077	6	9161	4.35E-16	0.190	Codd[21]
rs2736100	5	1286401	*TERT*	C	A	0.52	0.085	0.013	2.14E-05	0.54	4	5756	4.38E-19	0.310	Codd[21]
rs9419958	10	103916188	*OBFC1*	T	C	0.13	0.129	0.013	5.26E-11	0.028	6	9190	9.00E-11	0.171	Mangino[18]
rs9420907	10	103916707	*OBFC1*	C	A	0.14	0.142	0.014	1.14E-11	0.181	6	9190	7.00E-11	0.171	Codd[21]
rs4387287	10	103918139	*OBFC1*	A	C	0.14	0.120	0.013	1.40E-09	0.044	6	8541	2.00E-11	0.171	Levy[25]
rs3027234	17	8232774	*CTC1*	C	T	0.83	0.103	0.012	2.75E-08	0.266	6	9108	2.00E-08	0.292	Mangino[18]
rs8105767	19	22032639	*ZNF208*	G	A	0.25	0.064	0.011	0.000169	0.412	6	9096	1.11E-09	0.090	Codd[21]
rs412658	19	22176638	*ZNF676*	T	C	0.35	0.086	0.010	1.83E-08	0.568	6	9156	1.00E-08	0.484	Mangino[18]
rs6028466	20	39500359	*DHX35*	A	G	0.17	0.058	0.013	0.003972	0.533	6	9190	2.57E-08†	0.041	Mangino[18] & Gu
rs755017	20	63790269	*ZBTB46*	G	A	0.17	0.019	0.0129	0.339611	0.757	5	8026	6.71E-09	0.090	Codd[21]

*Summary data from Mangino et al[18]; Chr, chromosome; pos, base-pair position (GRCh38.p3); EA, effect allele, OA, other allele, Beta, standard deviation change in telomere length per copy of the effect allele; SE, standard error; EAF - effect allele frequency; Phet - p value for between-study heterogeneity in association between SNP and telomere length; †from a meta-analysis of Mangino[18] and Gu[20] performed in the present study.

Table 2. Study characteristics for primary non-communicable diseases

	No. cases	No. controls	No. SNPs	Statistical power	Pop.	Study / First author
Cancer						
Bladder cancer	1601	1819	10	0.62	EUR	NBCS[76]
Breast cancer	48155	43612	13	1.00	EUR	BCAC[17,77]
Estrogen receptor –ve	7465	42175	13	1.00	EUR	BCAC[17,77]
Estrogen receptor +ve	27074	41749	13	1.00	EUR	BCAC[17,77]
Colorectal cancer	14537	16922	9	1.00	EUR	CORECT/GECCO[64,78]
Endometrial cancer	6608	37925	12	1.00	EUR	ECAC[79,80]
Esophageal SCC	1942	2111	11	0.64	EA	Abnet[81]
Glioma	1130	6300	12	0.72	EUR	Wrensch[82] & Walsh[66]
Head & neck cancer	2082	3477	12	1.00	EUR	McKay et al[83]
Kidney cancer	2461	5081	12	0.99	EUR	KIDRISK[84]
Lung cancer	11348	15861	13	1.00	EUR	ILCCO[85]
Adenocarcinoma	3442	14894	13	1.00	EUR	ILCCO[85]
Squamous cell carcinoma	3275	15038	13	1.00	EUR	ILCCO[85]
Skin cancer						
Melanoma	12814	23203	13	1.00	EUR	MC[86]
Basal cell carcinoma	3361	11518	13	1.00	EUR	NHS/HPFS[87]
Neuroblastoma	2101	4202	12	0.87	EUR	Diskin[88]
Ovarian cancer	15397	30816	13	1.00	EUR	OCAC[17,89]
Clear cell	1016	30816	13	0.76	EUR	OCAC[17,89]
Endometriod	2154	30816	13	0.98	EUR	OCAC[17,89]
Mucinous	1643	30816	13	0.94	EUR	OCAC[17,89]
Serous invasive	9608	30816	13	1.00	EUR	OCAC[17,89]
Serous LMP	972	30816	13	0.73	EUR	OCAC[17,89]
Pancreatic cancer	5105	8739	12	1.00	EUR	PanScan (incl. EPIC)[90]
Prostate cancer	22297	22323	11	1.00	EUR	PRACTICAL[91,92]
Testicular germ cell cancer	986	4946	11	0.52	EUR	Turnbull[93] & Rapley[94]
Autoimmune/inflammatory diseases						
Alopecia areata	2332	5233	7	0.60	EUR	Betz[95]
Atopic dermatitis	10788	30047	13	1.00	EUR	EAGLE[96]
Celiac disease	4533	10750	3	0.82	EUR	Dubois[97]
Inflammatory bowel disease						
Crohn's disease	5956	14927	11	1.00	EUR	IIBDGC[98]
Ulcerative colitis	6968	20464	12	1.00	EUR	IIBDGC[98]
Juvenile idiopathic arthritis	1866	14786	11	0.87	EUR	Thompson[99]†
Multiple sclerosis	14498	24091	3	1.00	EUR	IMSGC[100]
Aggressive periodontitis	888	6789	13	0.63	EUR	Schaefer[101]
Rheumatoid arthritis	5538	20163	11	1.00	EUR	Stahl[102]
Cardiovascular diseases						
Abdominal aortic aneurysm	4972	99858	13	1.00	EUR	AC[103–108]
Coronary heart disease	22233	64762	13	1.00	EUR	CARDIoGRAM[109]
Heart failure	2526	20926	13	0.99	EUR	CHARGE-HF[110]
Hemorrhagic stroke	2963	5503	12	0.96	EUR	METASTROKE/ISGC[111]
Ischemic stroke	12389	62004	13	1.00	EUR	METASTROKE/ISGC[112,113]
large vessel disease	2167	62004	13	0.99	EUR	METASTROKE/ISGC[112,113]
small vessel disease	1894	62004	13	0.97	EUR	METASTROKE/ISGC[112]
cardioembolic	2365	62004	13	0.99	EUR	METASTROKE/ISGC[112]
Sudden cardiac arrest	3954	21200	13	1.00	EUR	Unpublished
Diabetes						
Type 1 diabetes	7514	9045	6	0.95	EUR	T1DBase[114115]
Type 2 diabetes	10415	53655	11	1.00	EUR	DIAGRAM[116]
Eye disease						

AMD	7473	51177	13	1.00	EUR	AMD Gene[117]
Retinopathy	1122	18289	12	0.75	EUR	Jensen[118]
Lung diseases						
Asthma	13034	20638	4	1.00	EUR	Ferreira/GABRIEL[119,120]
COPD	2812	2534	12	0.85	EUR	COPDGene[121]
Interstitial lung disease	1616	4683	9	0.60	EUR	Fingerlin[122]
Neurological / psychiatric diseases						
ALS	6100	7125	12	1.00	EUR	SLAGEN/ALSGEN[123]
Alzheimer's disease	17008	37154	12	1.00	EUR	IGAP[124]
Anorexia nervosa	2907	14860	9	0.93	EUR	GCAN[125]
Autism	4949	5314	7	0.82	EUR	PGC[126]
Bipolar disorder	7481	9250	9	1.00	EUR	PGC[127]
Major depressive disorder	9240	9519	8	0.99	EUR	PGC[128]
Schizophrenia	35476	46839	12	1.00	EUR	PGC[129]
Tourette syndrome	1177	4955	13	0.74	EUR	TICG/TSAICG[130]
Other						
Chronic kidney disease	5807	56430	13	1.00	EUR	CKDGen[131]
Endometriosis	4604	9393	11	1.00	Mix	Nyholt[132]

Study acronyms: AC, the aneurysm consortium; **ALSGEN,** the International Consortium on Amyotrophic Lateral Sclerosis Genetics**; AMD Gene**, Age-related Macular Degeneration Gene Consortium; **BCAC**, Breast Cancer Association Consortium; **CARDIoGRAM**, Coronary ARtery DIsease Genome wide Replication and Meta-analysis; **CHARGE-HF**, Cohorts for Heart and Aging Research in Genomic Epidemiology Consortium – Heart Failure Working Group; **COPDGene**, The Genetic Epidemiology of Chronic Obstructive Pulmonary Disease; **CKDGen**, Chronic Kidney Disease Genetics consortium; **CORECT**, ColoRectal Transdisciplinary Study; **DIAGRAM**, DIAbetes Genetics Replication And Meta-analysis; **EAGLE**, EArly Genetics & Lifecourse Epidemiology Eczema Consortium (excluding 23andMe); **ECAC**, Endometrial Cancer Association Consortium; **EPIC**, European Prospective Investigation into Cancer and Nutrition study; **GABRIEL**, Multidisciplinary Study to Identify the Genetic and Environmental Causes of Asthma in the European Community; **GCAN**, Genetic Consortium for Anorexia Nervosa; **GECCO**, Genetics and Epidemiology of Colorectal Cancer Consortium; **IGAP**, International Genomics of Alzheimer's Project; **HPFS**, Health Professionals Follow-Up Study; **ILCCO**, International Lung Cancer Consortium; **IMSGC**, International Multiple Sclerosis Genetic Consortium; **IIBDGC**, International Inflammatory Bowel Disease Genetics Consortium; **KIDRISK**, Kidney cancer consortium; **MC**, the melanoma meta-analysis consortium; **METASTROKE/ISGC**, METASTROKE project of the International Stroke Genetics Consortium; **NBCS**, Nijmegen Bladder Cancer Study; **NHS**, Nurses' Health Study; **OCAC**, Ovarian Cancer Association Consortium; **PanScan**, Pancreatic Cancer Cohort Consortium; **PGC**, Psychiatric Genomics Consortium; **PRACTICAL**, Prostate Cancer Association Group to Investigate Cancer Associated Alterations in the Genome; **SLAGEN**, Italian Consortium for the Genetics of Ayotrophic Lateral Sclerosis; **T1DBase**, type 1 diabetes database; **TICG** (Tourette International Collaborative-Genetics); **TSAICG** (Tourette Syndrome Association International Consortium for Genetics);. **Abbreviations**: **ALS**, amyotrophic lateral sclerosis; **AMD**, age-related macular degeneration; **COPD**, chronic obstructive pulmonary disease; **EUR**, European; **EA**, East Asian; **LMP**, low malignant potential; **No.**, number; **Pop**., population; **SCC**, squamous cell carcinoma; **SNP**, single nucleotide polymorphism; **-ve**, negative; **+ve**, positive; †plus previously unpublished data.

Figure 1. The association between genetically increased telomere length and odds of primary non-communicable diseases

Legend to Figure 1

*P value for association between genetically increased telomere length and disease from maximum likelihood; the effect estimate for heart failure is a hazard ratio (all others are odds ratios); P_{het}, P-value for heterogeneity amongst SNPs within the instrument; COPD, chronic obstructive pulmonary disease; SNP, single nucleotide polymorphism; CI, confidence interval; LMP, low malignancy potential; ER, estrogen receptor; -VE, negative; +VE, positive.

Figure 2. The association between genetically increased telomere length and odds of cancer as a function of selected characteristics

Legend to Figure 2

The plotted data show how the strength of the relationship between genetically increased telomere length and cancer varies by the selected characteristic. The R^2 statistic indicates how much of the variation between cancers can be explained by the selected characteristic. P-values are from meta-regression models. Circle sizes are proportional to the inverse of the variance of the log odds ratio. The hashed line indicates the null of no association between telomere length and cancer (i.e. an odds ratio of 1). Data for percentage survival 5 years after diagnosis, cancer incidence and median age-at-diagnosis was downloaded from the Surveillance, Epidemiology, and End Results Program.[33] Data for average lifetime number of stem cell divisions was downloaded from Tomasetti and Vogelstein.[34] Not all cancers had information available for the selected characteristics (hence the number of cancers varies across the subplots). Information was available for 9 cancers for tissue-specific rates of stem cell division, 13 cancers for percentage surviving 5 years post-diagnosis, 17 cancers for cancer incidence and 13 cancers for median age-at-diagnosis. SD, standard deviation; OR, Odds ratio.

Figure 3. Comparison of genetic and prospective observational studies[†] of the association between telomere length and disease

Legend to Figure 3

*from fixed-effects meta-analysis of independent observational studies described in Table S3; [†]search strategy and characteristics for observational studies are described in Tables S3 and S4; ‡CCHS and CGPS; +PLCO, ATBC & SWHS (acronyms explained in Table S3); CI, confidence interval

REFERENCES

1. Blackburn EH, Epel ES, Lin J. Human telomere biology: A contributory and interactive factor in aging, disease risks, and protection. Science (80-) 2015;350(6265):1193–8.

2. López-Otín C, Blasco MA, Partridge L, Serrano M, Kroemer G. The hallmarks of aging. Cell 2013;153(6):1194–217.

3. Samani NJ, van der Harst P. Biological ageing and cardiovascular disease. Heart 2008;94(5):537–9.

4. Weischer M, Bojesen SE, Nordestgaard BG. Telomere shortening unrelated to smoking, body weight, physical activity, and alcohol intake: 4,576 general population individuals with repeat measurements 10 years apart. PLoS Genet 2014;10(3):e1004191.

5. Houben JMJ, Moonen HJJ, van Schooten FJ, Hageman GJ. Telomere length assessment: biomarker of chronic oxidative stress? Free Radic Biol Med 2008;44(3):235–46.

6. Marchesi V. Risk factors: Short telomeres: association with cancer survival and risk. Nat Rev Clin Oncol 2013;10(5):247.

7. Haycock PC, Heydon EE, Kaptoge S, Butterworth a. S, Thompson A, Willeit P. Leucocyte telomere length and risk of cardiovascular disease: systematic review and meta-analysis. BMJ 2014;349(jul08 3):g4227–g4227.

8. Rode L, Nordestgaard BG, Bojesen SE. Peripheral blood leukocyte telomere length and mortality among 64,637 individuals from the general population. J Natl Cancer Inst 2015;107(6):djv074.

9. Zhao J, Miao K, Wang H, Ding H, Wang DW. Association between telomere length and type 2 diabetes mellitus: a meta-analysis. PLoS One 2013;8(11):e79993.

10. Weischer M, Nordestgaard BG, Cawthon RM, Freiberg JJ, Tybjærg-Hansen A, Bojesen SE. Short telomere length, cancer survival, and cancer risk in 47102 individuals. J Natl Cancer Inst 2013;105(7):459–68.

11. Ma H, Zhou Z, Wei S, et al. Shortened telomere length is associated with increased risk of cancer: a meta-analysis. PLoS One 2011;6(6):e20466.

12. Wentzensen IM, Mirabello L, Pfeiffer RM, Savage SA. The association of telomere length and cancer: a meta-analysis. Cancer Epidemiol Biomarkers Prev 2011;20(6):1238–50.

13. Pooley KA, Sandhu MS, Tyrer J, et al. Telomere length in prospective and retrospective cancer case-control studies. Cancer Res 2010;70(8):3170–6.

14. Hou L, Joyce BT, Gao T, et al. Blood Telomere Length Attrition and Cancer Development in the Normative Aging Study Cohort. EBioMedicine 2015;2(6):591–6.

15. Welter D, MacArthur J, Morales J, et al. The NHGRI GWAS Catalog, a curated resource of SNP-trait associations. Nucleic Acids Res 2014;42(Database issue):D1001-6.

16. Burdett T, Hall P, Hastings E, et al. The NHGRI-EBI Catalog of published genome-wide association studies [Internet]. [cited 2015 Jan 15];Available from: www.ebi.ac.uk/gwas

17. Pooley KA, Bojesen SE, Weischer M, et al. A genome-wide association scan (GWAS) for mean telomere length within the COGS project: identified loci show little association with hormone-related cancer risk. Hum Mol Genet 2013;22(24):5056–64.

18. Mangino M, Hwang S-J, Spector TD, et al. Genome-wide meta-analysis points to CTC1 and ZNF676 as genes regulating telomere homeostasis in humans. Hum Mol Genet 2012;21(24):5385–94.

19. Prescott J, Kraft P, Chasman DI, et al. Genome-wide association study of relative telomere length. PLoS One 2011;6(5):e19635.

20. Gu J, Chen M, Shete S, et al. A genome-wide association study identifies a locus on chromosome 14q21 as a predictor of leukocyte telomere length and as a marker of susceptibility for bladder cancer. Cancer Prev Res (Phila) 2011;4(4):514–21.

21. Codd V, Nelson CP, Albrecht E, et al. Identification of seven loci affecting mean telomere length and their association with disease. Nat Genet 2013;45(4):422–7.

22. Codd V, Mangino M, van der Harst P, et al. Common variants near TERC are associated with mean telomere length. Nat Genet 2010;42(3):197–9.

23. Liu Y, Cao L, Li Z, et al. A genome-wide association study identifies a locus on TERT for mean telomere length in Han Chinese. PLoS One 2014;9(1):e85043.

24. Saxena R, Bjonnes A, Prescott J, et al. Genome-wide association study identifies variants in casein kinase II (CSNK2A2) to be associated with leukocyte telomere length in a Punjabi Sikh diabetic cohort. Circ Cardiovasc Genet 2014;7(3):287–95.

25. Levy D, Neuhausen SL, Hunt SC, et al. Genome-wide association identifies OBFC1 as a locus involved in human leukocyte telomere biology. Proc Natl Acad Sci U S A 2010;107(20):9293–8.

26. Hindorff LA LA, MacArthur J, Morales J, et al. A catalog of published genome-wide association studies [Internet]. [cited 2015 Jan 15];Available from: www.genome.gov/gwastudies

27. Chene G, Thompson SG. Methods for Summarizing the Risk Associations of Quantitative Variables in Epidemiologic Studies in a Consistent Form. Am J Epidemiol 1996;144(6):610–21.

28. Burgess S, Scott RA, Timpson NJ, Davey Smith G, Thompson SG, EPIC-InterAct Consortium. Using published data in Mendelian randomization : a blueprint for efficient identification of causal risk factors. Eur J Epidemiol 2015;30(7):543–52.

29. Thomas DC, Lawlor DA, Thompson JR. Re: Estimation of bias in nongenetic observational studies using “Mendelian triangulation” by Bautista et al. Ann Epidemiol 2007;17(7):511–3.

30. Bowden J, Davey Smith G, Burgess S. Mendelian randomization with invalid instruments: effect estimation and bias detection through Egger regression. Int J Epidemiol 2015;44(2):512–25.

31. VanderWeele TJ, Tchetgen Tchetgen EJ, Cornelis M, Kraft P. Methodological challenges in mendelian randomization. Epidemiology 2014;25(3):427–35.

32. Bowden J, Davey Smith G, Haycock PC, Burgess S. Consistent estimation in Mendelian randomization with some invalid instruments using a weighted median estimator. Genet Epidemiol

33. National Cancer Institute. Surveillance, Epidemiology, and End Results Program [Internet]. [cited 2015 Aug 1];Available from: www.seer.cancer.gov

34. Tomasetti C, Vogelstein B. Variation in cancer risk among tissues can be explained by the number of stem cell divisions. Science 2015;347(6217):78–81.

35. U.S. Census Bureau [Internet]. [cited 2016 Jul 11];Available from: U.S. Census Bureau

36. R Core Team. A language and environment for statistical computing. 2013;

37. Burgess S, Thompson SG. Bias in causal estimates from Mendelian randomization studies with weak instruments. Stat Med 2011;30(11):1312–23.

38. Davey Smith G, Ebrahim S. What can mendelian randomisation tell us about modifiable behavioural and environmental exposures? BMJ Br Med J 2005;330(7499):1076–9.

39. Anic GM, Sondak VK, Messina JL, et al. Telomere length and risk of melanoma, squamous cell carcinoma, and basal cell carcinoma. Cancer Epidemiol 2013;37(4):434–9.

40. Pellatt AJ, Wolff RK, Torres-Mejia G, et al. Telomere length, telomere-related genes, and breast cancer risk: the breast cancer health disparities study. Genes Chromosomes Cancer 2013;52(7):595–609.

41. Caini S, Raimondi S, Johansson H, et al. Telomere length and the risk of cutaneous melanoma and non-melanoma skin cancer: a review of the literature and meta-analysis. J Dermatol Sci 2015;80(3):168–74.

42. Sanchez-Espiridion B, Chen M, Chang JY, et al. Telomere length in peripheral blood leukocytes and lung cancer risk: a large case-control study in Caucasians. Cancer Res 2014;74(9):2476–86.

43. Campa D, Mergarten B, De Vivo I, et al. Leukocyte telomere length in relation to pancreatic cancer risk: a prospective study. Cancer Epidemiol Biomarkers Prev 2014;23(11):2447–54.

44. Cui Y, Cai Q, Qu S, et al. Association of leukocyte telomere length with colorectal cancer risk: nested case-control findings from the Shanghai Women's Health Study. Cancer Epidemiol Biomarkers Prev 2012;21(10):1807–13.

45. De Vivo I, Prescott J, Wong JYY, Kraft P, Hankinson SE, Hunter DJ. A prospective study of relative telomere length and postmenopausal breast cancer risk. Cancer Epidemiol Biomarkers Prev 2009;18(4):1152–6.

46. Han J, Qureshi AA, Prescott J, et al. A prospective study of telomere length and the risk of skin cancer. J Invest Dermatol 2009;129(2):415–21.

47. Hofmann JN, Lan Q, Cawthon R, et al. A prospective study of leukocyte telomere

length and risk of renal cell carcinoma. Cancer Epidemiol Biomarkers Prev 2013;22(5):997–1000.

48. Julin B, Shui I, Heaphy CM, et al. Circulating leukocyte telomere length and risk of overall and aggressive prostate cancer. Br J Cancer 2015;112(4):769–76.

49. Kim S, Sandler DP, Carswell G, et al. Telomere length in peripheral blood and breast cancer risk in a prospective case-cohort analysis: results from the Sister Study. Cancer Causes Control 2011;22(7):1061–6.

50. Lan Q, Cawthon R, Shen M, et al. A prospective study of telomere length measured by monochrome multiplex quantitative PCR and risk of non-Hodgkin lymphoma. Clin Cancer Res 2009;15(23):7429–33.

51. Lee I-M, Lin J, Castonguay AJ, Barton NS, Buring JE, Zee RYL. Mean leukocyte telomere length and risk of incident colorectal carcinoma in women: a prospective, nested case-control study. Clin Chem Lab Med 2010;48(2):259–62.

52. Liang G, Qureshi AA, Guo Q, De Vivo I, Han J. No association between telomere length in peripheral blood leukocytes and the risk of nonmelanoma skin cancer. Cancer Epidemiol Biomarkers Prev 2011;20(5):1043–5.

53. Lynch SM, Major JM, Cawthon R, et al. A prospective analysis of telomere length and pancreatic cancer in the alpha-tocopherol beta-carotene cancer (ATBC) prevention study. Int J Cancer 2013;133(11):2672–80.

54. McGrath M, Wong JYY, Michaud D, Hunter DJ, De Vivo I. Telomere length, cigarette smoking, and bladder cancer risk in men and women. Cancer Epidemiol Biomarkers Prev 2007;16(4):815–9.

55. Nan H, Du M, De Vivo I, et al. Shorter telomeres associate with a reduced risk of melanoma development. Cancer Res 2011;71(21):6758–63.

56. Prescott J, McGrath M, Lee I-M, Buring JE, De Vivo I. Telomere length and genetic analyses in population-based studies of endometrial cancer risk. Cancer 2010;116(18):4275–82.

57. Qu S, Wen W, Shu X-O, et al. Association of leukocyte telomere length with breast cancer risk: nested case-control findings from the Shanghai Women's Health Study. Am J Epidemiol 2013;177(7):617–24.

58. Risques RA, Vaughan TL, Li X, et al. Leukocyte telomere length predicts cancer risk in Barrett's esophagus. Cancer Epidemiol Biomarkers Prev 2007;16(12):2649–55.

59. Seow WJ, Cawthon RM, Purdue MP, et al. Telomere length in white blood cell DNA and lung cancer: a pooled analysis of three prospective cohorts. Cancer Res 2014;74(15):4090–8.

60. Shen M, Cawthon R, Rothman N, et al. A prospective study of telomere length measured by monochrome multiplex quantitative PCR and risk of lung cancer. Lung Cancer 2011;73(2):133–7.

61. Walcott F, Rajaraman P, Gadalla SM, et al. Telomere length and risk of glioma. Cancer Epidemiol 2013;37(6):935–8.

62. Zee RYL, Castonguay AJ, Barton NS, Buring JE. Mean telomere length and risk of

617 incident colorectal carcinoma: a prospective, nested case-control approach. Cancer
618 Epidemiol Biomarkers Prev 2009;18(8):2280–2.

619 63. Willeit P, Willeit J, Mayr A, et al. Telomere length and risk of incident cancer and
620 cancer mortality. Jama 2010;304(1538–3598 (Electronic)):69–75.

621 64. Zhang C, Doherty J a., Burgess S, et al. Genetic determinants of telomere length and
622 risk of common cancers: a Mendelian randomization study. Hum Mol Genet
623 2015;24(18):5356–66.

624 65. Iles MM, Bishop DT, Taylor JC, et al. The effect on melanoma risk of genes
625 previously associated with telomere length. J Natl Cancer Inst 2014;106(10).

626 66. Walsh KM, Codd V, Rice T, et al. Longer genotypically-estimated leukocyte telomere
627 length is associated with increased adult glioma risk. Oncotarget 2015;6(40):42468–
628 77.

629 67. Ojha J, Codd V, Nelson CP, et al. Genetic Variation Associated with Longer Telomere
630 Length Increases Risk of Chronic Lymphocytic Leukemia. Cancer Epidemiol
631 Biomarkers Prev 2016;25(7):1043–9.

632 68. Armanios M, Blackburn EH. The telomere syndromes. Nat Rev Genet
633 2012;13(10):693–704.

634 69. Armanios M. Syndromes of telomere shortening. Annu Rev Genomics Hum Genet
635 2009;10(46):45–61.

636 70. Hanahan D, Weinberg RA. Hallmarks of cancer: the next generation. Cell
637 2011;144(5):646–74.

638 71. Blasco MA. Telomere length, stem cells and aging. Nat Chem Biol 2007;3(10):640–9.

639 72. Stone RC, Horvath K, Kark JD, Susser E, Tishkoff SA, Aviv A. Telomere Length and
640 the Cancer-Atherosclerosis Trade-Off. PLoS Genet 2016;12(7):e1006144.

641 73. Pierce BL, Burgess S. Efficient design for Mendelian randomization studies:
642 subsample and 2-sample instrumental variable estimators. Am J Epidemiol
643 2013;178(7):1177–84.

644 74. Burgess S, Butterworth AS, Thompson JR. Beyond Mendelian randomization: how to
645 interpret evidence of shared genetic predictors. J Clin Epidemiol 2015;1–9.

646 75. Centers for Disease Control and Prevention [Internet]. [cited 2016 Jul 14];Available
647 from: http://www.cdc.gov/nchs/fastats/deaths.htm

648 76. Rafnar T, Sulem P, Thorleifsson G, et al. Genome-wide association study yields
649 variants at 20p12.2 that associate with urinary bladder cancer. Hum Mol Genet
650 2014;23(20):5545–57.

651 77. Michailidou K, Hall P, Gonzalez-Neira A, et al. Large-scale genotyping identifies 41
652 new loci associated with breast cancer risk. Nat Genet 2013;45(4):353–61, 361-2.

653 78. Schumacher FR, Schmit SL, Jiao S, et al. Genome-wide association study of colorectal
654 cancer identifies six new susceptibility loci. Nat Commun 2015;6:7138.

655 79. Spurdle AB, Thompson DJ, Ahmed S, et al. Genome-wide association study identifies
656 a common variant associated with risk of endometrial cancer. Nat Genet

2011;43(5):451–4.

80. Painter JN, O'Mara TA, Batra J, et al. Fine-mapping of the HNF1B multicancer locus identifies candidate variants that mediate endometrial cancer risk. Hum Mol Genet 2015;24(5):1478–92.

81. Abnet CC, Freedman ND, Hu N, et al. A shared susceptibility locus in PLCE1 at 10q23 for gastric adenocarcinoma and esophageal squamous cell carcinoma. Nat Genet 2010;42(9):764–7.

82. Wrensch M, Jenkins RB, Chang JS, et al. Variants in the CDKN2B and RTEL1 regions are associated with high-grade glioma susceptibility. Nat Genet 2009;41(8):905–8.

83. McKay JD, Truong T, Gaborieau V, et al. A genome-wide association study of upper aerodigestive tract cancers conducted within the INHANCE consortium. PLoS Genet 2011;7(3):e1001333.

84. Purdue MP, Johansson M, Zelenika D, et al. Genome-wide association study of renal cell carcinoma identifies two susceptibility loci on 2p21 and 11q13.3. Nat Genet 2010;43(1):60–5.

85. Wang Y, McKay JD, Rafnar T, et al. Rare variants of large effect in BRCA2 and CHEK2 affect risk of lung cancer. Nat Genet 2014;46(7).

86. Law MH, Bishop DT, Lee JE, et al. Genome-wide meta-analysis identifies five new susceptibility loci for cutaneous malignant melanoma. Nat Genet 2015;47(9):987–95.

87. Zhang M, Song F, Liang L, et al. Genome-wide association studies identify several new loci associated with pigmentation traits and skin cancer risk in European Americans. Hum Mol Genet 2013;22(14):2948–59.

88. Diskin SJ, Capasso M, Schnepp RW, et al. Common variation at 6q16 within HACE1 and LIN28B influences susceptibility to neuroblastoma. Nat Genet 2012;44(10):1126–30.

89. Pharoah PDP, Tsai Y-Y, Ramus SJ, et al. GWAS meta-analysis and replication identifies three new susceptibility loci for ovarian cancer. Nat Genet 2013;45(4):362–70, 370-2.

90. Wolpin BM, Rizzato C, Kraft P, et al. Genome-wide association study identifies multiple susceptibility loci for pancreatic cancer. Nat Genet 2014;46(9):994–1000.

91. Eeles RA, Olama AA Al, Benlloch S, et al. Identification of 23 new prostate cancer susceptibility loci using the iCOGS custom genotyping array. Nat Genet 2013;45(4):385–91, 391-2.

92. Al Olama AA, Kote-Jarai Z, Berndt SI, et al. A meta-analysis of 87,040 individuals identifies 23 new susceptibility loci for prostate cancer. Nat Genet 2014;46(10):1103–9.

93. Turnbull C, Rapley E a, Seal S, et al. Variants near DMRT1, TERT and ATF7IP are associated with testicular germ cell cancer. Nat Genet 2010;42(7):604–7.

94. Rapley EA, Turnbull C, Al Olama AA, et al. A genome-wide association study of testicular germ cell tumor. Nat Genet 2009;41(7):807–10.

95. Betz RC, Petukhova L, Ripke S, et al. Genome-wide meta-analysis in alopecia areata resolves HLA associations and reveals two new susceptibility loci. Nat Commun 2015;6:5966.

96. EArly Genetics and Lifecourse Epidemiology (EAGLE) Eczema Consortium, Australian Asthma Genetics Consortium(AAGC), Australian Asthma Genetics Consortium AAGC. Multi-ancestry genome-wide association study of 21,000 cases and 95,000 controls identifies new risk loci for atopic dermatitis. Nat Genet 2015;47(12):1449–56.

97. Dubois PC a, Trynka G, Franke L, et al. Multiple common variants for celiac disease influencing immune gene expression. Nat Genet 2010;42(4):295–302.

98. Liu JZ, van Sommeren S, Huang H, et al. Association analyses identify 38 susceptibility loci for inflammatory bowel disease and highlight shared genetic risk across populations. Nat Genet 2015;47(9):979–86.

99. Thompson SD, Marion MC, Sudman M, et al. Genome-wide association analysis of juvenile idiopathic arthritis identifies a new susceptibility locus at chromosomal region 3q13. Arthritis Rheum 2012;64(8):2781–91.

100. Beecham AH, Patsopoulos NA, Xifara DK, et al. Analysis of immune-related loci identifies 48 new susceptibility variants for multiple sclerosis. Nat Genet 2013;45(11):1353–60.

101. Schaefer AS, Richter GM, Nothnagel M, et al. A genome-wide association study identifies GLT6D1 as a susceptibility locus for periodontitis. Hum Mol Genet 2010;19(3):553–62.

102. Stahl EA, Raychaudhuri S, Remmers EF, et al. Genome-wide association study meta-analysis identifies seven new rheumatoid arthritis risk loci. Nat Genet 2010;42(6):508–14.

103. Bown MJ, Jones GT, Harrison SC, et al. Abdominal aortic aneurysm is associated with a variant in low-density lipoprotein receptor-related protein 1. Am J Hum Genet 2011;89(5):619–27.

104. Gretarsdottir S, Baas AF, Thorleifsson G, et al. Genome-wide association study identifies a sequence variant within the DAB2IP gene conferring susceptibility to abdominal aortic aneurysm. Nat Genet 2010;42(8):692–7.

105. Jones GT, Bown MJ, Gretarsdottir S, et al. A sequence variant associated with sortilin-1 (SORT1) on 1p13.3 is independently associated with abdominal aortic aneurysm. Hum Mol Genet 2013;22(14):2941–7.

106. Harrison SC, Smith AJP, Jones GT, et al. Interleukin-6 receptor pathways in abdominal aortic aneurysm. Eur Heart J 2013;34(48):3707–16.

107. Elmore JR, Obmann MA, Kuivaniemi H, et al. Identification of a genetic variant associated with abdominal aortic aneurysms on chromosome 3p12.3 by genome wide association. J Vasc Surg 2009;49(6):1525–31.

108. Borthwick K, Smelser D, Bock J, et al. Ephenotyping for Abdominal Aortic Aneurysm in the Electronic Medical Records and Genomics (eMERGE) Network: Algorithm Development and Konstanz Information Miner Workflow. Int J Biomed Data Min 2015;4(1).

741 109. Schunkert H, König IR, Kathiresan S, et al. Large-scale association analysis identifies
742 13 new susceptibility loci for coronary artery disease. Nat Genet 2011;43(4):333–8.

743 110. Smith NL, Felix JF, Morrison AC, et al. Association of genome-wide variation with
744 the risk of incident heart failure in adults of European and African ancestry: a
745 prospective meta-analysis from the cohorts for heart and aging research in genomic
746 epidemiology (CHARGE) consortium. Circ Cardiovasc Genet 2010;3(3):256–66.

747 111. Woo D, Falcone GJ, Devan WJ, et al. Meta-analysis of genome-wide association
748 studies identifies 1q22 as a susceptibility locus for intracerebral hemorrhage. Am J
749 Hum Genet 2014;94(4):511–21.

750 112. Malik R, Freilinger T, Winsvold BS, et al. Shared genetic basis for migraine and
751 ischemic stroke: A genome-wide analysis of common variants. Neurology
752 2015;84(21):2132–45.

753 113. Traylor M, Farrall M, Holliday EG, et al. Genetic risk factors for ischaemic stroke and
754 its subtypes (the METASTROKE collaboration): a meta-analysis of genome-wide
755 association studies. Lancet Neurol 2012;11(11):951–62.

756 114. Barrett JC, Clayton DG, Concannon P, et al. Genome-wide association study and
757 meta-analysis find that over 40 loci affect risk of type 1 diabetes. Nat Genet
758 2009;41(6):703–7.

759 115. Burren OS, Adlem EC, Achuthan P, Christensen M, Coulson RMR, Todd JA.
760 T1DBase: update 2011, organization and presentation of large-scale data sets for type
761 1 diabetes research. Nucleic Acids Res 2011;39(Database):D997–1001.

762 116. Morris ADPDPA, Voight BFB, Teslovich TMT, et al. Large-scale association analysis
763 provides insights into the genetic architecture and pathophysiology of type 2 diabetes.
764 Nat Genet 2012;44(9):981–90.

765 117. Fritsche LG, Chen W, Schu M, et al. Seven new loci associated with age-related
766 macular degeneration. Nat Genet 2013;45(4):433–9, 439-2.

767 118. Jensen RA, Sim X, Li X, et al. Genome-wide association study of retinopathy in
768 individuals without diabetes. PLoS One 2013;8(2):e54232.

769 119. Moffatt MF, Kabesch M, Liang L, et al. Genetic variants regulating ORMDL3
770 expression contribute to the risk of childhood asthma. Nature 2007;448(7152):470–3.

771 120. Ferreira MAR, Matheson MC, Duffy DL, et al. Identification of IL6R and
772 chromosome 11q13.5 as risk loci for asthma. Lancet (London, England)
773 2011;378(9795):1006–14.

774 121. Cho MH, McDonald M-LN, Zhou X, et al. Risk loci for chronic obstructive pulmonary
775 disease: a genome-wide association study and meta-analysis. Lancet Respir Med
776 2014;2(3):214–25.

777 122. Fingerlin TE, Murphy E, Zhang W, et al. Genome-wide association study identifies
778 multiple susceptibility loci for pulmonary fibrosis. Nat Genet 2013;45(6):613–20.

779 123. Fogh I, Ratti A, Gellera C, et al. A genome-wide association meta-analysis identifies a
780 novel locus at 17q11.2 associated with sporadic amyotrophic lateral sclerosis. Hum
781 Mol Genet 2014;23(8):2220–31.

124. Lambert JC, Ibrahim-Verbaas CA, Harold D, et al. Meta-analysis of 74,046 individuals identifies 11 new susceptibility loci for Alzheimer's disease. Nat Genet 2013;45(12):1452–8.

125. Boraska V, Franklin CS, Floyd JAB, et al. A genome-wide association study of anorexia nervosa. Mol Psychiatry 2014;19(10):1085–94.

126. Smoller JW, Hospital MG. Identification of risk loci with shared effects on five major psychiatric disorders: a genome-wide analysis. Lancet 2013;381(9875):1371–9.

127. Sklar P, Ripke S, Scott LJ, et al. Large-scale genome-wide association analysis of bipolar disorder identifies a new susceptibility locus near ODZ4. Nat Genet 2011;43(10):977–83.

128. Ripke S, Wray NR, Lewis CM, et al. A mega-analysis of genome-wide association studies for major depressive disorder. Mol Psychiatry 2013;18(4):497–511.

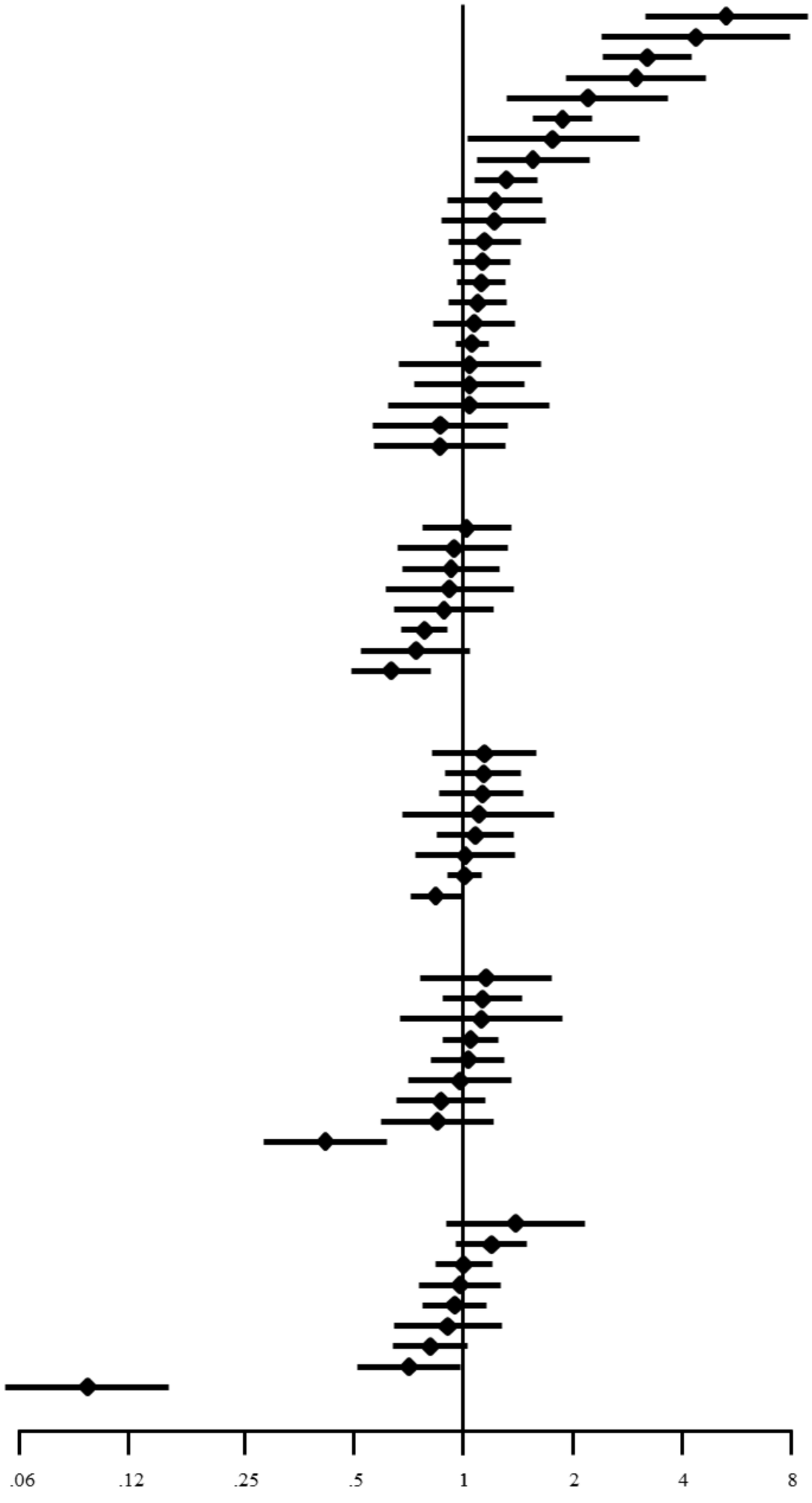
129. Ripke S, Neale BM, Corvin A, et al. Biological insights from 108 schizophrenia-associated genetic loci. Nature 2014;511(7510):421–7.

130. Scharf JM, Yu D, Mathews CA, et al. Genome-wide association study of Tourette's syndrome. Mol Psychiatry 2013;18(6):721–8.

131. Köttgen A, Pattaro C, Böger C a, et al. New loci associated with kidney function and chronic kidney disease. Nat Genet 2010;42(5):376–84.

132. Nyholt DR, Low S-K, Anderson CA, et al. Genome-wide association meta-analysis identifies new endometriosis risk loci. Nat Genet 2012;44(12):1355–9.

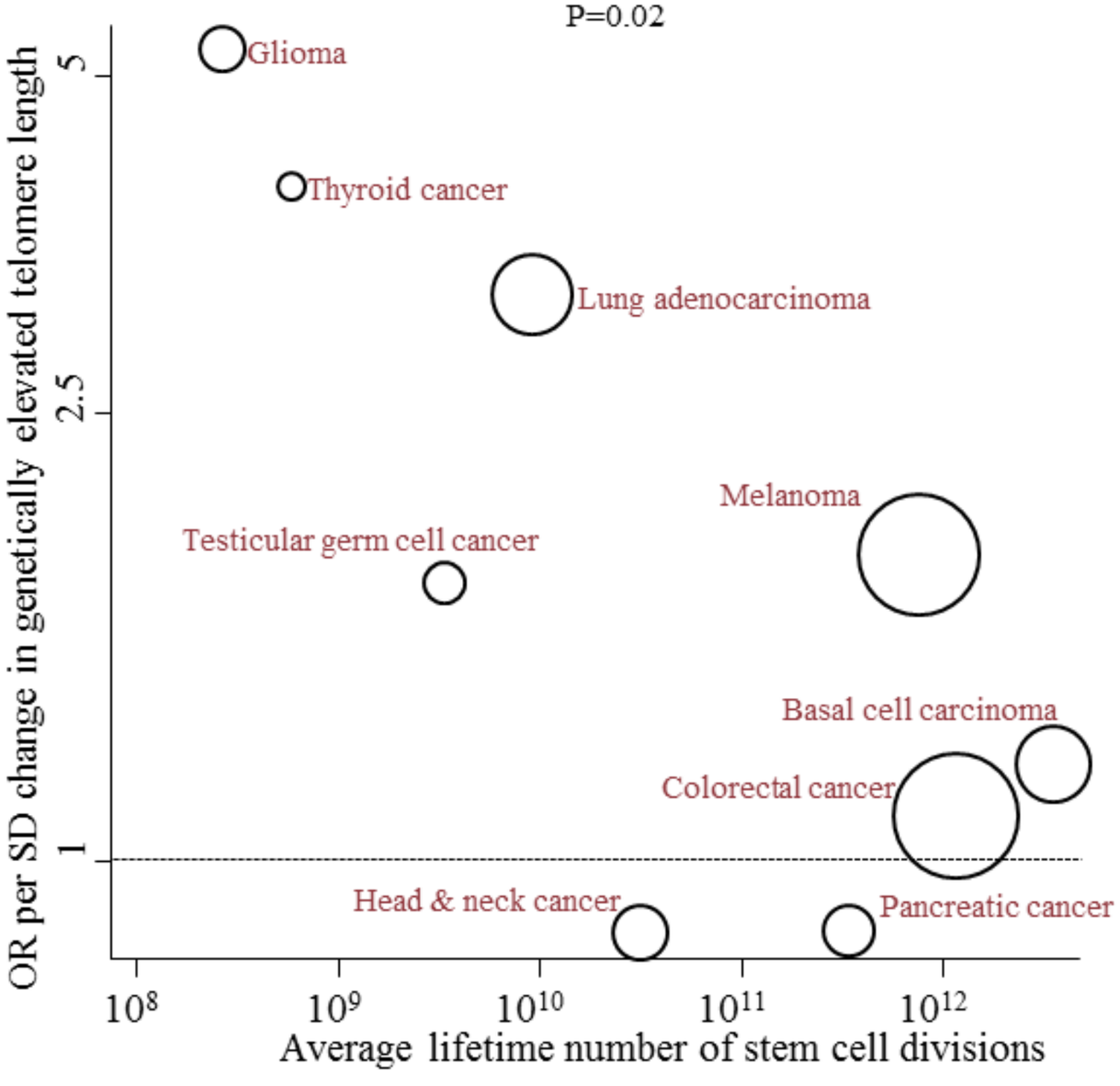
	Tumor/disease subtype	No. of cases	No. of SNPs	P*	P_{het}
Cancer					
Glioma	-	1130	12	$2.45x10^{-10}$	.0135
Ovarian cancer	Serous LMP	972	13	$1.62x10^{-6}$	.082
Lung cancer	Adenocarcinoma	3442	13	$6.66x10^{-16}$	.0001
Neuroblastoma	-	2101	12	$1.11x10^{-6}$	.3799
Bladder cancer	-	1601	10	.0026	.2541
Skin cancer	Melanoma	12814	13	$5.79x10^{-11}$	.086
Testicular germ cell cancer	-	985	11	.0423	.3002
Kidney cancer	-	2461	12	.0164	.251
Endometrial cancer	-	6607	12	.0091	.1436
Skin cancer	Basal cell carcinoma	3361	13	.203	.0398
Ovarian cancer	Endometriod	2154	13	.2499	.3319
Breast cancer	ER-ve	7465	13	.2542	<0.0001
Ovarian cancer	Serous invasive	9608	13	.2088	.0407
Prostate cancer	-	22297	11	.1533	<0.0001
Colorectal cancer	-	14537	9	.3436	.0156
Lung cancer	Squamous cell carcinoma	3275	13	.6212	.9042
Breast cancer	ER+	27074	13	.2912	.0228
Ovarian cancer	Clear cell	1016	13	.8676	.3226
Ovarian cancer	Mucinous	1643	13	.8396	.3147
Esophageal cancer	Squamous cell carcinoma	1942	8	.9055	.4513
Pancreatic cancer	Adenocarcinoma	5105	12	.5009	.0016
Head & neck cancer	-	2082	12	.4767	.0763
Cardiovascular diseases					
Heart failure	-	2526	13	.8931	.5348
Ischemic stroke	Small vessel disease	1894	13	.7141	.2358
Sudden cardiac arrest	-	3954	13	.62	.5400
Intracerebral stroke	-	5503	12	.6719	.0002
Ischemic stroke	Cardioembolic	2365	13	.4464	.0374
Coronary heart disease	-	22233	13	.0009	.2441
Ischemic stroke	Large vessel disease	2167	13	.0801	.1578
Abdominal aortic aneurysm	-	4972	13	.0003	.0617
Neurological / psychiatric diseases					
Anorexia Nervosa	-	2907	9	.4406	.1797
Bipolar disorder	-	7481	9	.315	.1427
Amyotrophic lateral sclerosis	-	6100	12	.4049	.6331
Tourette syndrome	-	1177	13	.6936	.4978
Major depressive disorder	-	9240	8	.5646	.254
Autism	-	4949	7	.9514	.8232
Schizophrenia	-	9379	12	.9245	.0053
Alzheimer's disease	-	17008	12	.0296	.1451
Autoimmune/inflammatory diseases					
Alopecia areata	-	2332	7	.4994	.9505
Inflammatory bowel disease	Crohn's disease	5956	11	.3586	.0585
Periodontitis	-	888	13	.6702	.0066
Atopic dermatitis	-	10788	13	.6094	.1194
Inflammatory bowel disease	Ulcerative colitis	6968	12	.8303	.0984
Multiple sclerosis	-	14498	3	.8885	.1455
Rheumatoid arthritis	-	5538	11	.3184	.1828
Juvenile idiopathic arthritis	-	1866	11	.3669	.922
Celiac disease	-	4533	3	$1.00x10^{-5}$	.0028
Other diseases					
Retinopathy	-	1126	12	.1453	.3831
Age-related macular degeneration		7473	13	.1206	.209
Type 2 diabetes	-	10415	11	.9837	.6811
Endometriosis	-	4604	11	.8606	.0576
Chronic kidney disease	-	5807	13	.5859	.6972
COPD	-	2812	12	.5598	.3807
Asthma	-	13034	4	.0779	.3213
Type 1 diabetes	-	7514	6	.0378	.0586
Interstitial lung disease	-	1616	9	$2.02x10^{-19}$	.0014



Average lifetime number of stem cell divisions

Adjusted R^2=52.63%

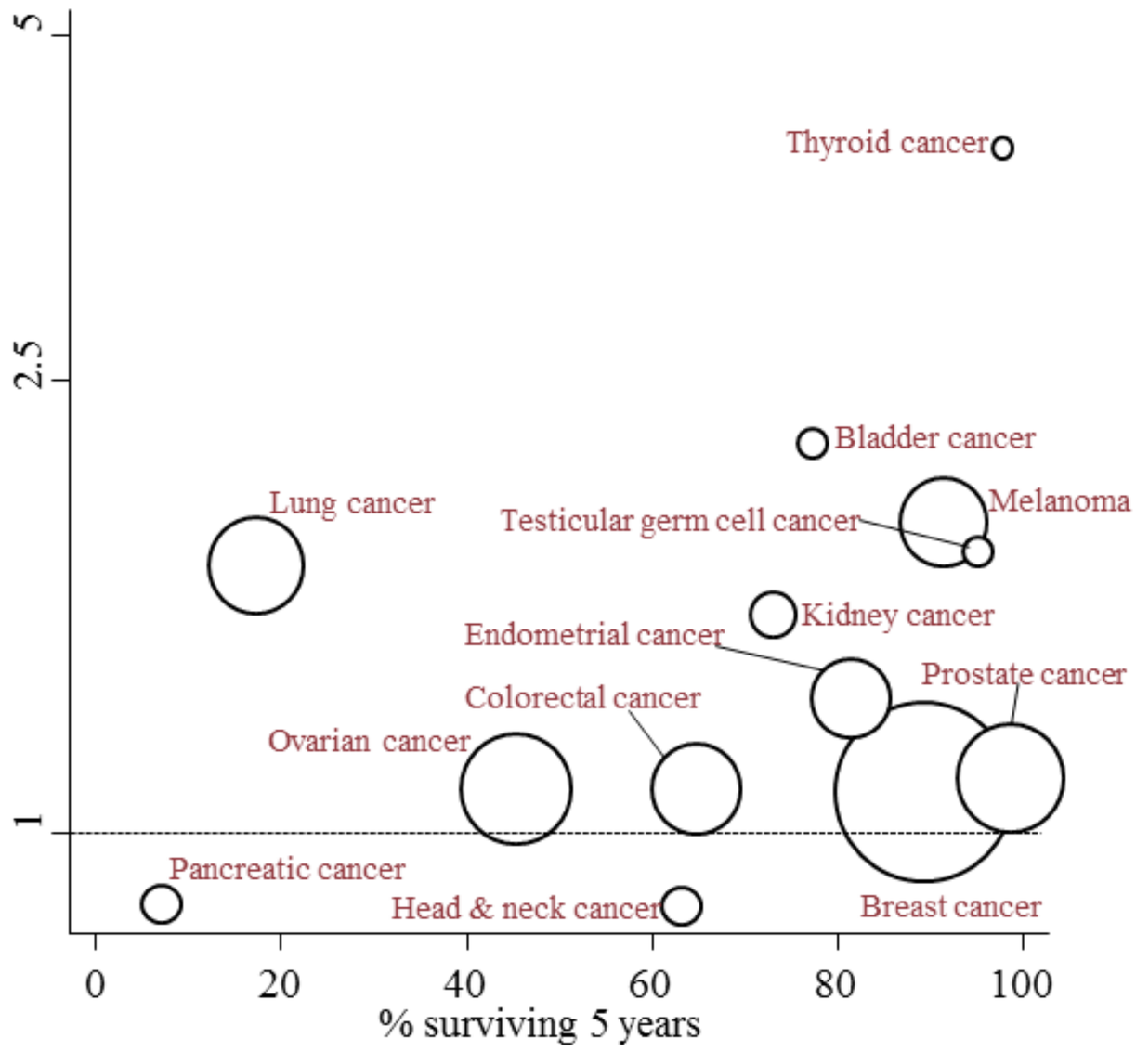
P=0.02



% surviving 5 years

Adjusted R^2=-19.49%

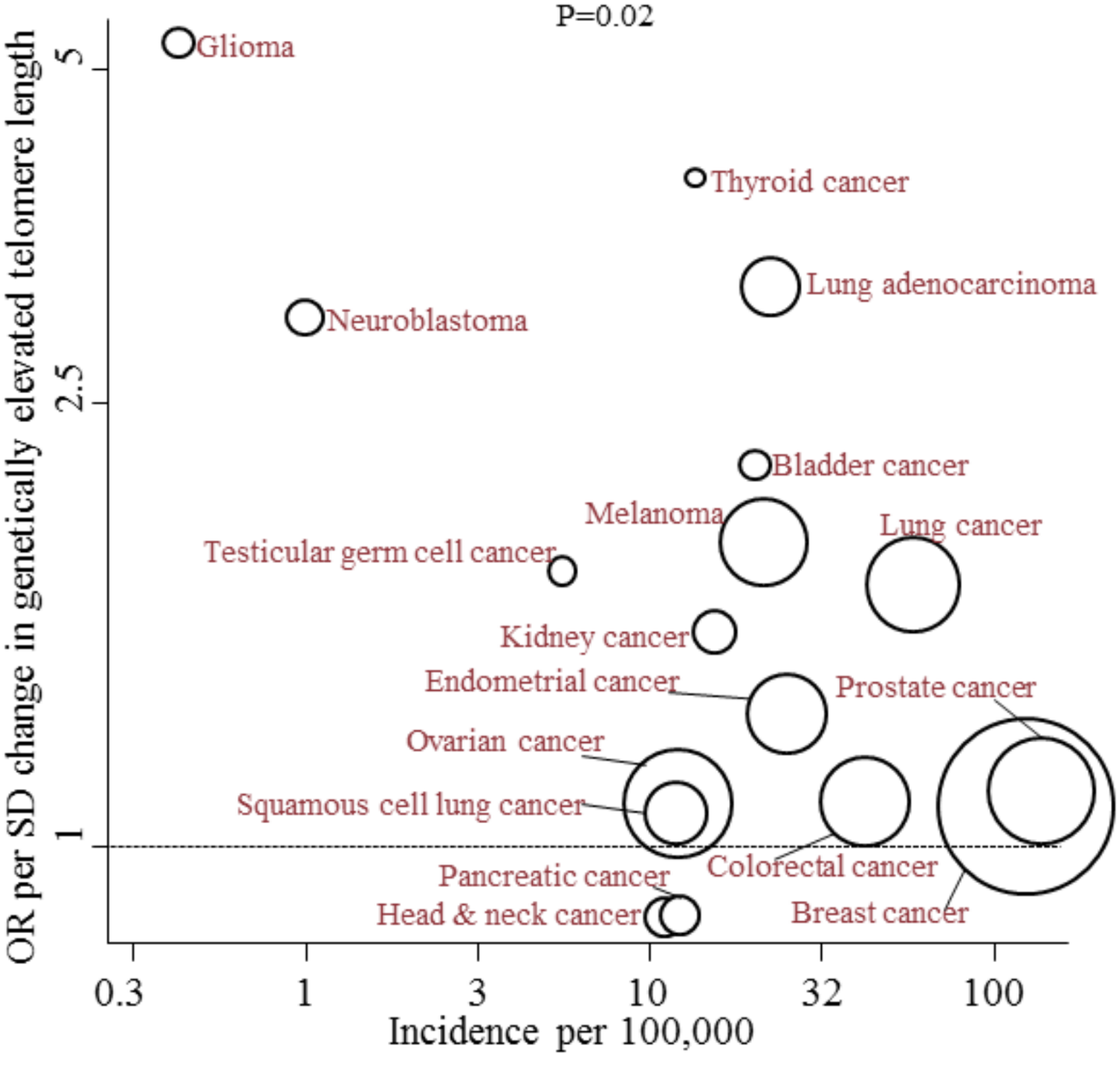
P=0.37



Cancer incidence

Adjusted R^2=32.12%

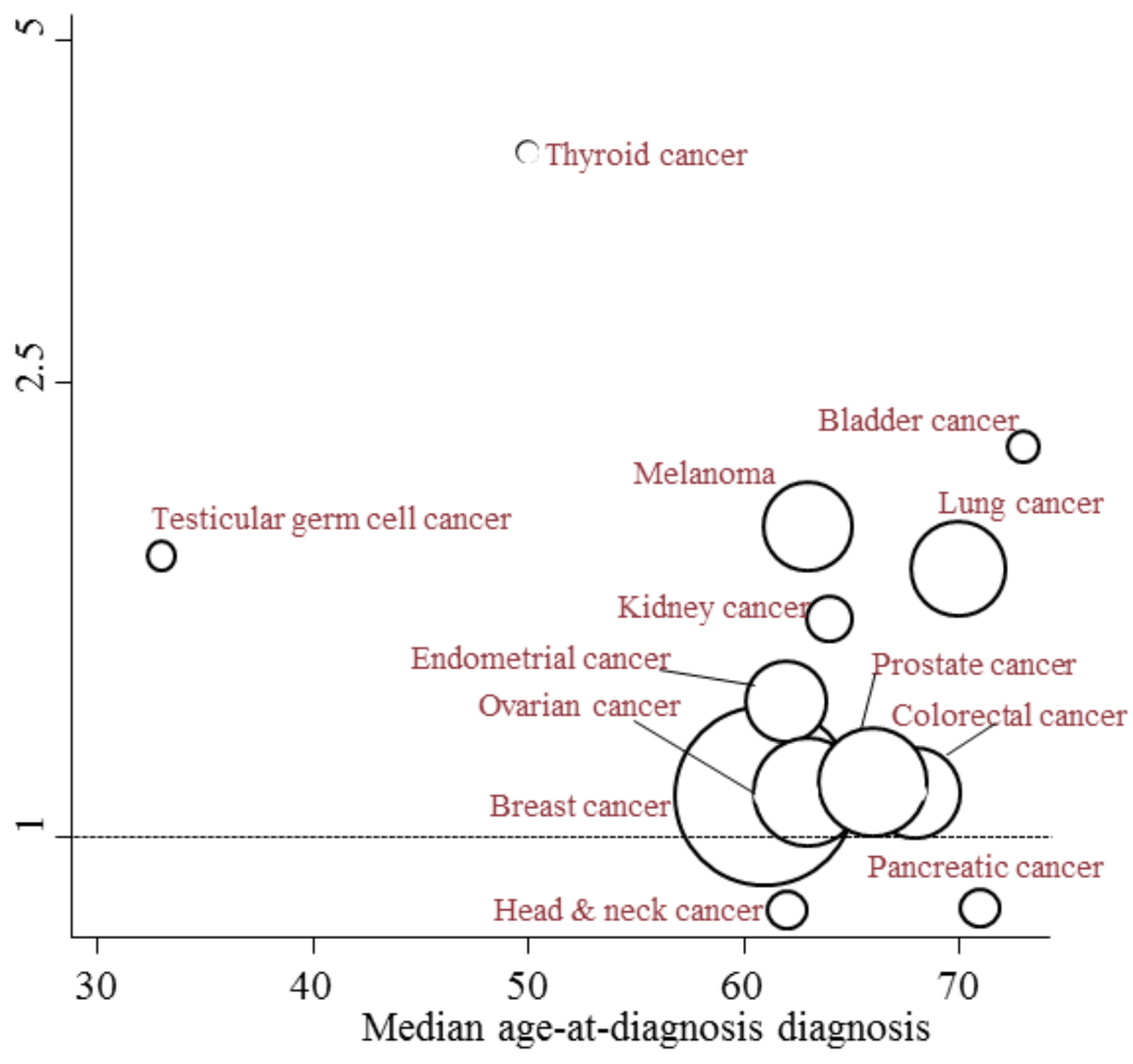
P=0.02



Median age-at-diagnosis diagnosis

Adjusted R^2=-8.86%

P=0.41



	No. of cases	Odds ratio (95% CI) per standard deviation increase in telomere length
Cancer		
Breast cancer		
Genetic study	48155	1.08 (0.99, 1.19)
Observational study*	1716	1.02 (0.99, 1.05)
Prostate cancer		
Genetic study	22297	1.12 (0.96, 1.30)
Observational study*	1340	1.07 (1.01, 1.14)
Ovarian cancer		
Genetic study	15397	1.09 (0.94, 1.27)
Observational study	96	1.13 (0.98, 1.32)
Colorectal cancer		
Genetic study	14537	1.09 (0.91, 1.31)
Observational study*	1447	1.04 (0.97, 1.11)
Lung cancer		
Genetic study	11348	1.71 (1.44, 2.04)
Observational study‡	522	0.94 (0.87, 1.02)
Observational study+	847	1.28 (1.12, 1.46)
Endometrial cancer		
Genetic study	6608	1.31 (1.07, 1.61)
Observational study*	382	1.06 (0.95, 1.19)
Pancreatic cancer		
Genetic study	5105	0.86 (0.56, 1.32)
Observational study*	648	1.05 (0.95, 1.17)
Lung adenocarcinoma		
Genetic study	3442	3.19 (2.40, 4.22)
Observational study	288	1.44 (1.14, 1.82)
Skin basal cell carcinoma		
Genetic study	3361	1.22 (0.90, 1.65)
Observational study	363	0.96 (0.85, 1.09)
Lung squamous cell carcinoma		
Genetic study	3275	1.07 (0.82, 1.39)
Observational study	163	1.05 (0.78, 1.42)
Kidney cancer		
Genetic study	2461	1.55 (1.08, 2.23)
Observational study*	268	0.94 (0.81, 1.10)
Head & neck cancer		
Genetic study	2082	0.86 (0.57, 1.30)
Observational study	76	0.89 (0.72, 1.09)
Melanoma		
Genetic study	12814	1.88 (1.54, 2.28)
Observational study*	734	1.17 (1.06, 1.29)
Bladder cancer		
Genetic study	1601	2.19 (1.32, 3.66)
Observational study	184	1.28 (1.02, 1.61)
Testicular cancer		
Genetic study	986	1.76 (1.02, 3.04)
Observational study	10	0.94 (0.56, 1.55)
Glioma		
Genetic study	1130	5.27 (3.15, 8.81)
Observational study	101	0.90 (0.68, 1.18)
Non-neoplastic diseases		
Coronary heart disease		
Genetic study	22233	0.78 (0.67, 0.90)
Observational study	2272	0.86 (0.78, 0.94)
Ischemic stroke		
Genetic study	12389	0.85 (0.73, 1.00)
Observational study	824	0.94 (0.82, 1.08)
Type 2 diabetes		
Genetic study	10415	1.00 (0.84, 1.20)
Observational study	2011	0.90 (0.83, 0.97)

Supplementary material

Mendelian randomization study of the association between telomere length and risk of cancer and non-neoplastic diseases

The Telomeres Mendelian Randomization Collaboration

Correspondence: Philip Haycock

MRC Integrative Epidemiology Unit

University of Bristol

Bristol

UK

philip.haycock@bristol.ac.uk

Tel: +44 1173 310 088

7 supplementary figures / 6 supplementary tables

Contents

SUPPLEMENTARY METHODS 4
Additional details on the design strategy 4
Identification of genetic instruments for telomere length 4
Acquisition of summary data from disease and risk factor studies 5
Power calculations 5
Estimating the association between genetically increased telomere length and outcome traits 6
Likelihood approach 7
The weighted median approach[14] 8
The MR-Egger approach 8
SUPPLEMENTARY RESULTS 9
SUPPLEMENTARY DISCUSSION 10
Mechanisms of association between SNPs and telomere length 10
Bias from sample overlap and strength of the association between SNPs and telomere length 10
Misconceptions about Mendelian randomization 10
Potential for confounding by population stratification, ancestry and age 11
Associations with non-neoplastic diseases 12
Supplementary Table S1. Study characteristics for secondary non-communicable diseases and diseases from independent studies for replication analyses 13
Supplementary Table S2. Study characteristics of 44 risk factors for non-communicable diseases 14
Supplementary Table S3. Selected prospective observational studies of the association between leukocyte telomere length and disease 16
Supplementary Table S4. PubMed search strategy for prospective observational studies of association between telomere length* and disease 19
Supplementary Table S6. Glossary of terms 20
Supplementary Figure S1. Study design 21
Supplementary Figure S3. Replication of association between genetically increased telomere length and odds of non-communicable diseases in independent datasets 23
Supplementary Figure S4. Sensitivity analyses of association between genetically increased telomere length and odds of non-communicable diseases 24
Supplementary Figure S5. Association between genetically increased telomere length and risk factors for non-communicable diseases 25
Supplementary Figure S6. Association between genetically increased telomere length and smoking 26
Supplementary Figure S7. Causal diagram illustrating the assumptions of Mendelian randomization 27
ACKNOWLEDGEMENTS OF THE CONTRIBUTING STUDIES AND CONSORTIA 28
Amyotrophic lateral sclerosis GWAS consortium 28
The Aneurysm Consortium 29
Australian Asthma Genetics Consortium 32

Coronary ARtery DIsease Genome wide Replication and Meta-analysis (CARDIoGRAM) consortium and The Coronary Artery Disease (C4D) Genetics consortium 33
The Cohorts for Heart and Aging Research in Genomic Epidemiology (CHARGE) – Heart Failure Working Group 34
CHARGE - Sudden Cardiac Arrest Working Group 34
The Genetic Epidemiology of Chronic Obstructive Pulmonary Disease (COPDGene) 35
Early Growth Genetics (EGG) Consortium 38
The EArly Genetics and Lifecourse Epidemiology (EAGLE) consortium 38
Endometrial Cancer Association Consortium (ECAC) 47
Glioma GWAS 49
Endometriosis GWA meta-analysis 50
European Periodontitis Genetics Group (EPG) 52
The International Genomics of Alzheimer's Project (IGAP) 55
The Japanese Collaboration Team for GWAS of Panic Disorder 56
Meta-Analyses of Glucose and Insulin-related traits Consortium (MAGIC) 57
Melanoma meta-analysis consortium (MC) 57
The Multi-Ethnic Study of Atherosclerosis (MESA) 69
The Nurses' Health Study (NHS) and the Health Professionals Follow-Up Study (HPFS) 70
GWAS of non-alcoholic fatty liver disease (hepatic steatosis) 70
Pancreatic cancer case-control consortium (PanC4) 70
Pancreatic Cancer Cohort Consortium (PanScan) 71
The European Prospective Investigation into Cancer and Nutrition (EPIC) study 71
The PRACTICAL Consortium 72
Sarcoidosis GWAS 75
The Singapore Epidemiology of Eye Diseases Study (SEED) 75
Acknowledgements of studies that contributed to the GWAS meta-analysis of telomere length[4] 76
The Framingham Heart Study 76
TwinsUK 76
References 77

SUPPLEMENTARY METHODS

Additional details on the design strategy

Identification of genetic instruments for telomere length

To identify genetic variants to serve as instruments for telomere length, we searched the genome-wide association study (GWAS) catalog[1,2] on the 15 January 2015, to identify reported single nucleotide polymorphisms (SNPs) associated with telomere length. To supplement the list with additional potential instruments, we also searched the original study reports curated by the GWAS catalog.[3–11] We included all 'telomere length' SNPs in the GWAS catalog as potential proxies, regardless of their reported P-value, but used a P-value threshold of $<5x10^{-8}$ (the conventional threshold for declaring association in GWAS) for SNPs identified from original study reports (if these were not already curated by the GWAS catalog). We acquired summary data for all SNPs identified by the above strategy from a meta-analysis of six GWASs of leukocyte telomere length, conducted in 9,190 participants of European ancestry.[4] Telomere length in the six studies was measured by Southern blotting. GWAS analyses in the 6 studies were adjusted for age, sex, body mass index and smoking history. The genomic control inflation factor (λ_{GC}) ranged from 0.995 to 1.076 across the six studies, indicating little evidence for confounding by population stratification.[4] The following summary data were acquired for each SNP from each of the six studies: the regression coefficient (beta) and its standard error, where the beta reflects the change in telomere length (in base pair units) per copy of the effect allele; the effect allele; the non-effect allele; and effect allele frequency. We combined the effect estimates from the six separate studies by fixed effects meta-analysis. We then excluded SNPs if they lacked strong evidence of association with telomere length. We defined strong evidence of association as a P value $<5x10^{-8}$ in: i) the discovery stage of at least one published GWAS of telomere length[3–10] or ii) a meta-analysis of summary data

from Mangino et al[4] and other GWASs of telomere length,[3,5–10] with any overlapping studies excluded from Mangino et al.[4] We also excluded SNPs with a minor allele frequency <0.05 or showing strong evidence of between-study heterogeneity in associations with telomere length (P≤0.001).

Acquisition of summary data from disease and risk factor studies

We extracted the following summary data for each genetic instrument for telomere length from GWASs of diseases and risk factors: the regression coefficient (beta) and its standard error, the effect allele, the non-effect allele and effect allele frequency. For binary traits, the beta corresponded to the log odds ratio per copy of the effect allele. For quantitative traits, the beta corresponded to the unit change in the trait per copy of the effect allele. We harmonized the summary data for diseases and risk factors so that the effect allele reflected the allele associated with longer telomeres. When SNPs were palindromic, i.e. A/T or G/C, we used information on allele frequency to resolve strand ambiguity. We also requested the following metrics of SNP genotype quality: P-values for Hardy-Weinberg equilibrium (HWE), imputation quality scores and P-values for between-study heterogeneity. We also estimated the percentage overlap in participants amongst the telomere length and disease and risk factor GWASs. When reported, statistics on between-study heterogeneity, Hardy–Weinberg equilibrium and imputation quality were used to exclude low quality SNPs from disease and risk factor studies, using the following criteria: strong evidence of between-study heterogeneity in the SNP-phenotype association (P≤0.001), Hardy–Weinberg disequilibrium (P≤0.001) or imputation quality metric (info or r^2) ≤0.90.

Power calculations

Power calculations for disease outcomes were implemented using the method described by Burgess[12] and assumed an odds ratio of ≥2.0 per standard deviation higher telomere length and an alpha of 0.01. Power calculations for risk factors for non-communicable diseases were similar,

except that a ≥0.5 standard deviation change in quantitative risk factors and an odds ratio of ≥1.5 for binary risk factors was assumed for each standard deviation change in telomere length. When more than one study was available for the same outcome trait, priority was given to the study with the higher statistical power. Power calculations took into account the variance explained in telomere length by each SNP, inferred from published reports,[3–10] and the sample size available for each outcome.

Estimating the association between genetically increased telomere length and outcome traits

We employed three general approaches for estimating the association between genetically increased telomere length and outcome traits. Our main results are based on a likelihood-approach.[13] Sensitivity analyses were based on two approaches: the weighted median[14] and MR-Egger regression.[15] The technical details of these approaches are described below.

Prior to calculating the associations of genetically increased telomere length with diseases and risk factors, we estimated the pairwise r^2 for all telomere-associated SNPs residing on the same chromosome using PLINK[16] and 1000 Genomes phase 3 data for European samples.[17] SNPs residing on separate chromosomes or separated by more than 50 megabases on the same chromosome were assumed to be in linkage equilibrium. The genetic instruments for telomere length were pruned so that no SNP pair had an $r^2>0.9$ (strong linkage disequilibrium), using the 'indep' command in PLINK.[16] The base pair position and chromosome id for each SNP, in GCRCh38 format, was extracted from Ensembl through the R biomart package.[18–20] Linkage disequilibrium between the remaining SNPs was taken into account using a variance-covariance matrix (described below). For analyses in which SNP-disease associations were derived from East Asian populations, genetic instruments were further pruned so that no SNP pair had an $r^2>0.1$ (because the variance-covariance matrix used to model the correlation between SNPs was based on a European population).

Likelihood approach

We combined summary data across SNPs into a single instrument, using maximum likelihood to estimate the slope of the relationship between β_{GD} and β_{GP} and a variance-covariance matrix to make allowance for linkage disequilibrium between SNPs, where β_{GD} is the change in the outcome trait per copy of the effect allele and β_{GP} is the standard deviation change in telomere length per copy of the effect allele.[13] The standard deviation of telomere length corresponds to approximately 650 base pairs.[4] The variance-covariance matrix was estimated using 1000 Genomes phase 3 data for Europeans.[13] The model that is fitted is:

$$\begin{pmatrix} \boldsymbol{\beta_{GP}} \\ \boldsymbol{\beta_{GD}} \end{pmatrix} \sim N_{2K}\left(\begin{pmatrix} \boldsymbol{\xi} \\ \beta_{IV}\boldsymbol{\xi} \end{pmatrix}, \begin{pmatrix} \Sigma_{PP} & \Sigma_{PD} \\ \Sigma_{DP} & \Sigma_{DD} \end{pmatrix}\right)$$

where $\boldsymbol{\beta_{GP}}$ is a vector of the SNP-telomere-length associations, $\boldsymbol{\beta_{GD}}$ is a vector of the SNP-disease associations, β_{IV} is the causal effect parameter, K is the number of SNPs, Σ_{PP} is a variance-covariance matrix with elements $(\Sigma_{PP})_{ij} = se(\beta_{GPi})se(\beta_{GPj})\rho_{ij}$ where $se(\beta_{GPi})$ is the standard error of the SNP-telomere-length association for the *i*th genetic variant, and ρ_{ij} is the correlation between the *i*th and *j*th variants due to linkage disequilibrium. Components of Σ_{DD} are similarly defined as $(\Sigma_{DD})_{ij} = se(\beta_{GDi})se(\beta_{GDj})\rho_{ij}$, and $\Sigma_{PD} = \Sigma_{DP} = 0$ due to the two-sample setting (sensitivity analyses in a previous study[13] suggested results were robust to some correlation between the gene-phenotype and gene-outcome associations that may arise due to sample overlap). The slope estimated by maximum likelihood can be interpreted as the log odds ratio for disease per standard deviation change in genetically increased telomere length. The slope can further be interpreted as the causal effect of telomere length on disease if Mendelian randomization assumptions hold. The assumptions are: the SNPs are associated with telomere length (IV1); the SNPs are independent of confounders (IV2); and the SNPs are independent of disease adjusted for telomere length and confounders (IV3). See Supplementary Figure S7 for further details of the Mendelian randomization assumptions and Supplementary Table S6 for a glossary of terms.

198

199 *The weighted median approach*[14]

200 Let $\hat{\beta}_{(1)},\ldots,\hat{\beta}_{(J)}$ represent the J causal effect estimates ordered from smallest ($\hat{\beta}_{(1)}$) to largest ($\hat{\beta}_{(J)}$).

201 Now define

202 $$w^*_{(j)} = \frac{w_j}{S_J}, \quad \text{where} \quad S_J = \sum_j w_j ,$$

203 where w_j is the inverse variance of $\hat{\beta}_{(j)}$,

204 and equate $\hat{\beta}_{(j)}$ with a quantile, $p^w_{(j)}$, defined as

205 $$p^w_{(j)} = \frac{100}{S_J}\left(S_{(j)} - \frac{w_{(j)}}{2} \right).$$

206 $p^w_{(j)}$ represents the quantile from the weighted empirical distribution function of the ordered

207 estimates $\hat{\beta}_{(1)},\ldots,\hat{\beta}_{(J)}$. The weighted median estimate, $\hat{\beta}_{WM}$ is defined as the 50th percentile of this

208 weighted distribution. Typically the 50th percentile will lie between two estimates ($\hat{\beta}_{(l)}$ and $\hat{\beta}_{(m)}$,

209 say), in which case $\hat{\beta}_{WM}$ is found by linear interpolation. $\hat{\beta}_{WM}$ is a consistent estimate for β provided

210 that at least 50% of the 'weight' making up S_J comes from genetic variants that are valid

211 instruments. In other words, the weighted median function provides a valid estimate of the causal

212 effect of telomere length on disease if at least half of the genetic information comes from valid

213 instruments (assumptions illustrated in Supplementary Figure S7; glossary of terms in

214 Supplementary Table S6).[14]

215

216 *The MR-Egger approach*

217 The MR-Egger method[15] performs a weighted linear regression of the SNP-disease coefficients on

218 the SNP-exposure coefficients (where exposure in this study is telomere length):

219 $$\frac{\hat{\Gamma}_j}{\sigma_{Yj}} = \frac{\beta_{0E}}{\sigma_{Yj}} + \beta_{1E}\frac{\hat{\gamma}_j}{\sigma_{Yj}}$$

where Γ corresponds to the SNP-disease coefficients, γ corresponds to the SNP-exposure coefficients and σ_{yj} is the standard error of $\hat{\Gamma}_j$. If all SNPs are valid instruments, then $\beta_{0E} = 0$. The value of $\hat{\beta}_{0E}$ can be interpreted as an estimate of the average pleiotropic effect across the SNPs. An intercept term that differs from zero is indicative of overall directional pleiotropy. The MR-Egger estimate for β , $\hat{\beta}_{1E}$, is consistent even if *all* SNPs are invalid, provided that

- Across all SNPs, the magnitude of the SNP-exposure associations are independent of their pleiotropic effects (also known as the InSIDE [Instrument Strength Independent of Direct Effect] assumption)
- The number of SNPs, J, grows large (i.e. tends to infinity).

See Supplementary Figure S7 for further details on the assumptions and Supplementary Table S6 for a glossary of terms.

SUPPLEMENTARY RESULTS

In analyses of secondary cancer outcomes, genetically increased telomere length was associated with thyroid cancer, chronic lymphocytic leukemia and multiple myeloma (P<0.05) (Supplementary Figure S2). In analyses of secondary non-neoplastic diseases, genetically increased telomere length was associated with reduced odds of panic disorder (P<0.05) (Supplementary Figure S2). In secondary analyses of 44 risk factors for non-communicable diseases (Supplementary Table S2), genetically increased telomere length was associated with increased pulse pressure, systolic blood pressure, diastolic blood pressure, mean arterial pressure, triglycerides, uric acid and education and with decreased HDL cholesterol, mean corpuscular haemoglobin and mean corpuscular volume (P<0.05) (Supplementary Figure S5). There was some evidence for an association between genetically increased telomere length and ever smoking status (P=0.03, Supplementary Figure S6) but this association is unlikely to be reliable given that the SNP-telomere-length associations were adjusted for smoking history; the association may therefore reflect collider bias.[21]

SUPPLEMENTARY DISCUSSION

Mechanisms of association between SNPs and telomere length

The mechanisms of the underlying associations between the selected SNPs and telomere length are generally unknown. Some of the SNPs are located in or near the *TERC* or *TERT* genes, suggesting that the mechanism could involve the telomerase enzyme, as well as the *OBFC1* and *CTC1* genes, which have known roles in regulation of telomere length biology (Table 1). OBFC1 is an enzyme involved in initiating DNA replication and is involved in the telomere-associated CST complex.[22] *CTC1* encodes a component of the CST complex, which plays a role in protecting telomeres from degradation.

Bias from sample overlap and strength of the association between SNPs and telomere length

The selected genetic instruments for telomere length correspond to 10 independent genomic loci and collectively account for 2-3% of the variance in leukocyte telomere length. The corresponding F statistic is around 18, which means that bias due to weak instruments is unlikely to be substantial even if there were considerable overlap amongst the telomere length and disease and risk factor GWASs.[23] The estimated overlap in participants amongst the telomere length and outcome GWASs was less than 11% for all diseases and risk factors, except for hepatic steatosis, for which overlap was around 51%, indicating that the vast majority of our results should be robust to weak instrument bias.

Misconceptions about Mendelian randomization

A common misconception about Mendelian randomization studies is that genetic instruments should explain a substantial proportion of the variation in target exposures (e.g. telomere length in this study) in order to provide robust inferences about exposure-disease associations. However, if the genetic instruments are valid (i.e. conform to Mendelian randomization assumptions,

Supplementary Figure S7), the variation explained by the instrument only affects statistical power and does not generally affect validity of the causal inference. In this sense, genotype assignment in a Mendelian randomization study is analogous to treatment assignment in a randomized controlled trial, e.g. of blood pressure lowering drugs.[24] Although experimental interventions to reduce blood pressure may only explain a small fraction of the total variation in blood pressure in a typical RCT, we can still make causal inferences about blood pressure as a whole (and not just the proportion of variation in blood pressure due to the experimental intervention). Moreover, the aim of Mendelian randomization studies is to make inferences at the population level and not the individual level (for which genetic proxies of substantial explanatory power would be required).[24] If Mendelian randomization assumptions were violated, however, then the limited variation explained by our genetic instruments might not behave in similar manner to other sources of variation in telomere length, which would undermine our ability to draw causal inferences. See the above section 'Estimating the association between genetically increased telomere length and outcome traits' and Supplementary Figure S7 for details on the assumptions. See Supplementary Table S6 for an explanation of Mendelian randomization terminology. See Haycock et al[25] and Davey Smith and Hemani[26] for reviews on Mendelian randomization.

Potential for confounding by population stratification, ancestry and age

It is unlikely that confounding by population stratification, ancestry or age (an important confounder of observational studies of telomere length) can account for our results. The 15 primary diseases showing some evidence of association with telomere length (defined as a P value<0.05) were 100% European, on the basis of self reported ancestry or genetic analyses (individuals showing genetic evidence of non-European ancestry were excluded).[3,27–44] In addition, these studies all made some allowance for population stratification in their analyses: 12 adjusted for principal component scores of genetic variation in their models or applied genomic control corrections to their results; and 3 concluded there was little evidence for population stratification, on the basis of

visual inspection of Quantile-Quantile plots of GWAS results (i.e. lambdas for genomic inflation were close to 1). The GWAS we used to defined genetic instruments for telomere length[4] also adjusted for principal component scores; and lambdas for genomic inflation were close to 1. Since our MR analyses will have inherited any adjustments made in the original analyses, it is therefore unlikely that confounding by ancestry or population stratification can explain our results.

Confounding by age is also unlikely, given the random distribution of genotypes in the general population with respect to lifestyle and other environmental factors, as well as the fixed nature of germline genotypes. Consistent with this expectation, we did not observe an association between subject age and their genetically predicted telomere length values in our previous studies.[44,45]

Associations with non-neoplastic diseases

The inverse associations observed for coronary heart disease, abdominal aortic aneurysm, celiac disease and interstitial lung disease are compatible with findings based on observational and Mendelian randomization studies of telomere length as well as dyskeratosis congenita (a congenital disease characterized by chronically short telomeres).[46–50]

320

321

322

323

324

325

326

327

328

Supplementary Table S1. Study characteristics for secondary non-communicable diseases and diseases from independent studies for replication analyses

	No. cases	No. controls	No. SNPs	Statistical power	Pop.	First author /database
Cancer						
Chronic lymphocytic leukemia	2883	8350	1	0.22	EUR	Speedy/GWAS cat.[51]
Chronic myeloid leukemia	201	497	8	0.07	EA	Kim[52]
Ewing's sarcoma	401	684	4	0.06	EUR	Postel-Vinay[53]
Follicular lymphoma	212	748	3	0.04	EUR	Conde[54]
Gallbladder cancer	41	866	2	0.01	EA	Cha[55]
Gastric cancer						
Cardia adenocarcinoma	1126	2111	11	0.47	EA	Abnet[56]
Noncardia adenocarcinoma	632	2111	11	0.29	EA	Abnet[56]
Multiple myeloma	4692	10990	1	0.37	EUR	Chubb/GWAS cat.[57]
Nasopharyngeal carcinoma	1583	1894	2	0.17	EA	Bei[58]
B-cell Non-Hodgkin lymphoma	253	1438	10	0.13	EA	Tan[59]
Skin squamous cell carcinoma	449	11518	13	0.34	EUR	Zhang[60]
Thyroid cancer	649	431	12	0.16	EUR	Kohler[61]
Upper gastrointestinal cancers	3523	2100	2	0.28	EA	Li/dbGAP[62]
Autoimmune/inflammatory diseases						
Inflammatory psoriatic arthritis	609	990	13	0.29	EUR	Huffmeier[63]
Kawasaki disease	405	6252	11	0.26	EUR	Khor[64]
Narcolepsy	1188	1985	9	0.46	EA	Han[65]
Psoriasis	1139	1132	9	0.34	EA	Zhang[66]
Sarcoidosis	564	1575	9	0.16	EUR	Fischer[67]
Systemic lupus erythematosus	1311	1783	4	0.20	EUR	Hom/dbGAP[68]
Vitiligo	1117	1429	2	0.12	EA	Quan[69]
Wegener's granulomatosis	459	1503	10	0.20	EUR	Xie[70]
Neurological / psychiatric diseases						
Bulimia nervosa	151	2291	8	0.07	EUR	Wade[71]
Panic disorder	718	1717	8	0.28	EA	**JCTGPD**[72]
Parkinson's disease	1713	3978	4	0.35	EUR	Simón-Sánchez/dbGAP[73]

Other						
Hirschsprung's disease	173	615	6	0.04	EA	Tang[74]
Paget's disease	741	2699	12	0.43	EUR	Albagha[75]
Vascular dementia	84	200	8	0.03	EA	Kim[76]
Independent disease studies for replication analyses						
Bladder cancer	7712	13125	1	0.56	EUR	Figueroa/GWAS cat.[77]
Colorectal cancer	728	3282	9	0.39	EA	Zhang[78]
Coronary heart disease	15399	15050	4	1.00	Mix	C4D[79]
Glioma	1854	4955	1	0.12	EUR	GliomaScan/GWAS cat.[80]
Interstitial lung disease†	542	542	11	0.15	EUR	Noth[81]
Interstitial lung disease‡	242	1469	1	0.02	EA	Mushiroda/GWAS cat.[82]
Pancreatic cancer	4164	3792	10	0.90	EUR	PanC4[83]
Multiple sclerosis	978	883	4	0.11	EUR	Baranzini/dbGAP[84]
Nasopharyngeal carcinoma	277	285	2	0.03	EA	Tse[85]
Type 2 diabetes	8569	8923	10	1.00	EA	Li[86]

†≤17% cases overlapped with cases from Fingerlin et al[31] and 77% of cases had idiopathic pulmonary fibrosis; ‡all cases had idiopathic pulmonary fibrosis.
Study/database acronyms: C4D, Coronary Artery Disease Genetics Consortium; **dbGAP**, summary data downloaded from the database of Genotypes and Phenotypes; **GWAS cat.**, data downloaded from the National Human Genome Research Institute/European Bioinformatics Institute Catalog of published genome wide association studies; **JCTGPD**, Japanese Collaboration Team for GWAS of Panic Disorder. **Abbreviations**: EUR, European; EA, East Asian; **No.**, number; **Pop.**, population; **SNP**, single nucleotide polymorphism.

Supplementary Table S2. Study characteristics of 44 risk factors for non-communicable diseases

	Sample size	SD	Units	No. of SNPs	Stat. power	Pop.	First author / study
Anthropometric							
Birth length	22557	2.0	cm	12	1.00	EUR	EGG[87]
Birth weight	26836	547.5	g	12	1.00	EUR	EGG[88]
Body mass index	241253	4.8	kg/m^2	13	1.00	EUR	GIANT[89]
Childhood obesity	13848	NA	log_e odds	12	0.78	EUR	EGG[90]
Head circumference	10705	1.5	cm	13	1.00	EUR	EGG[91]
Height	253288	0.1	m	13	1.00	EUR	GIANT[92]
Hip circumference	224459	8.5	cm	13	1.00	EUR	GIANT[93]
Waist circumference	224459	12.5	cm	13	1.00	EUR	GIANT[93]
Waist-to-hip ratio	224459	0.1	ratio	13	1.00	EUR	GIANT[93]
Smoking behaviors							
Age of smoking initiation	47961	0.3	log_e years	13	1.00	EUR	TAG[94]
Cigarettes smoked per day	68028	11.7	CPD	13	1.00	EUR	TAG[94]
Ever smoker	74035	NA	log_e odds	13	1.00	EUR	TAG[94]
Ex smoker	41969	NA	log_e odds	13	1.00	EUR	TAG[94]
Blood pressure							
Diastolic blood pressure	66466	10.7	mm Hg	12	1.00	EUR	ICBP[95]
Mean arterial pressure	27803	12.8	mm Hg	13	1.00	EUR	ICBP[96]
Pulse pressure	70903	13.5	mm Hg	13	1.00	EUR	ICBP[96]
Systolic blood pressure	66473	18.2	mm Hg	12	1.00	EUR	ICBP[95]
Education							
College completion	95427	NA	log_e odds	13	1.00	EUR	SSGAC[97]
Years of educational attainment	126559	1.2	years	13	1.00	EUR	SSGAC[97]
Glycemic							
2 hr glucose	15234	1.27	mmol/L	11	1.00	EUR	MAGIC[98]
Beta-cell function (HOMA-B)	46186	0.96	log_e HOMA	12	1.00	EUR	MAGIC[99]
Fasting glucose	46186	0.73	mmol/L	12	1.00	EUR	MAGIC[99]
Fasting insulin	38238	0.79	log_e pmol/L	12	1.00	EUR	MAGIC[99]

Fasting proinsulin	10701	0.81	$\log_e$ pmol/L	12	1.00	EUR	MAGIC[99]
Gycated hemoglobin (HbA1c)	46368	0.53	%	12	1.00	EUR	MAGIC[100]
Insulin resistance (HOMA-IR)	46186	0.67	$\log_e$ HOMA	12	1.00	EUR	MAGIC[99]
Hemotological							
Hemoglobin	54287	1.3	g/dL	12	1.00	EUR	van der Harst[101]
Mean cell hemoglobin	45969	1.99	pg	12	1.00	EUR	van der Harst[101]
Mean cell hemoglobin concentration	49632	1.01	g/dL	12	1.00	EUR	van der Harst[101]
Mean cell volume	51277	5.2	fl	12	1.00	EUR	van der Harst[101]
Packed cell volume	46848	5.9	%	12	1.00	EUR	van der Harst[101]
Red blood cell count	47873	0.5	10^{12}/L	12	1.00	EUR	van der Harst[101]
Lipids							
HDL cholesterol	103019	15.51	mg/dL	11	1.00	EUR	GLGC[102]
LDL cholesterol	97562	38.67	mg/dL	11	1.00	EUR	GLGC[102]
Total cholesterol	103266	41.75	mg/dL	11	1.00	EUR	GLGC[102]
Triglycerides	99050	90.72	mg/dL	11	1.00	EUR	GLGC[102]
Renal function							
Microalbuminuria	30482	NA	$\log_e$ odds	13	0.82	EUR	CKDGen[103]
Serum creatinine	67093	0.24	$\log_e$ ml/min/1.73m²	13	1.00	EUR	CKDGen[103]
Serum cystatin	20957	0.23	$\log_e$ ml/min/1.73m²	13	1.00	EUR	CKDGen[103]
Urinary albumin-to-creatinine ratio	31580	1.0	$\log_e$ mg/g	13	1.00	EUR	CKDGen[103]
Other							
Grade of nuclear cataract	7140	0.8	grade	11	1.00	ASN	SEEDS[104]
Hepatic steatosis	7176	5.6	Hounsfield units	12	1.00	EUR	Speliotes[105]
Percent emphysema	7914	0.71	$\log_e$ %+1	12	1.00	ME	MESA[106]
Uric acid	42742	1.3	mg/dL	12	1.00	EUR	GUGC[107]

Study acronyms: **CKDGen**, chronic kidney disease genetics consortium; **EGG,** Early Growth Genetics Consortium; **GIANT**, Genetic Investigation of ANthropometric Traits; **GUGC**, Global Urate and Gout consortium; **TAG**, Tobacco and Genetics Consortium; **ICBP**, International Consortium for Blood Pressure; **SSGAC**, Social Science Genetics Association Consortium; **MAGIC**, Meta-Analyses of Glucose and Insulin-related traits Consortium; **MESA**, Multi-Ethnic Study of Atherosclerosis; **GLGC**, Global Lipids Genetics Consortium; **SEEDS,** the Singapore Epidemiology of Eye Diseases Study.
Abbreviations: **ASN,** Asian; **Con**., concentration; **EUR,** European population; **ME,** multi-ethnic; **SD** - standard deviation; **$\log_e$**, natural log; **Stat.,** statistical

Supplementary Table S3. Selected prospective observational studies of the association between leukocyte telomere length and disease

Cohort / first author	Disease	Year	Design	No. of controls / cohort size	No. of cases	RR (95% CI) as reported by study	Scale of RR reported by study	Conversion factor[§]	RR (95% CI) per SD increase in TL	Adjusted[‡]	Pop.	P_{het}	Search strategy†
Cancer outcomes													
NHS, HPFS[108]	Bladder cancer	2007	NCC	192	184	1.88 (1.05 to 3.36)	shortest vs. longest quartile	2.54	1.28 (1.02 to 1.61)	++	EUR	NA	2
CCHS, CGPS[109]	Breast cancer	2013	PC	24588	574	0.99 (0.95 to 1.03)	per 1000 bp (1.29 SD) decrease	-1.29	1.01 (0.98 to 1.04)	+++++	EUR		1
SWHS[110]	Breast cancer	2013	NCC	695	601	1.77 (1.02 to 3.06)	shortest vs. longest quintile	2.80	1.23 (1.01 to 1.49)	++	EA	0.17	2
Sister Study[111]	Breast cancer	2011	Case-cohort	735	342	0.93 (0.64 to 1.35)	shortest vs. longest quartile	-2.54	1.03 (0.89 to 1.19)	+	EUR (92%)		1
EPIC[112]	Breast cancer	2010	NCC	420	199	1.58 (0.75 to 3.31)	shortest vs. longest quartile	2.54	1.2 (0.89 to 1.6)	+	EUR		1
WHS[113]	Colorectal cancer	2010	NCC	357	134	0.94 (0.65 to 1.38)	per unit (1.30 SD) decrease	-1.30	1.05 (0.78 to 1.4)	+++++	EUR		3
PHS[114]	Colorectal cancer	2009	NCC	306	191	0.8 (0.55 to 1.16)	per unit (1.72 SD) decrease	-1.72	1.14 (0.92 to 1.41)	++++	EUR		3
CCHS, CGPS[109]	Colorectal cancer	2013	PC	46748	496	0.97 (0.88 to 1.07)	per 1000 bp (1.29 SD) decrease	-1.29	1.02 (0.95 to 1.1)	++++	EUR	0.47	1
SWHS[115]	Colorectal cancer	2012	NCC	549	441	1.61 (0.94 to 2.75)	longest vs. 3rd shortest quintile	1.40	1.4 (0.96 to 2.06)	+	EA		1
EPIC[112]	Colorectal cancer	2010	NCC	406	185	1.13 (0.54 to 2.36)	shortest vs. longest quartile	-2.54	0.95 (0.71 to 1.27)	+	EUR		1
NHS[116]	Endometrial cancer	2010	NCC	791	279	1.2 (0.73 to 1.96)	shortest vs. longest quartile	-2.54	0.93 (0.77 to 1.13)	+++++	EUR	0.11	2
CCHS, CGPS[109]	Endometrial cancer	2013	PC	25262	103	0.85 (0.71 to 1.02)	per 1000 bp (1.29 SD)	-1.29	1.13 (0.99 to 1.31)	+++++	EUR		1

							decrease						
PLCO[117]	Glioma	2013	NCC	198	101	1.26 (0.69 to 2.29)	shortest vs. longest tertile	-2.18	0.9 (0.68 to 1.18)	++	EUR	NA	1
CCHS, CGPS[109]	Head & neck cancer	2013	PC	47036	76	1.17 (0.9 to 1.53)	per 1000 bp (1.29 SD) decrease	-1.29	0.89 (0.72 to 1.09)	++++	EUR	NA	1
CCHS, CGPS[109]	Kidney cancer	2013	PC	47063	59	1.04 (0.78 to 1.39)	per 1000 bp (1.29 SD) decrease	-1.29	0.97 (0.77 to 1.21)	++++	EUR	NA	1
PLCO[118]	Kidney cancer	2013	NCC	410	209	0.8 (0.5 to 1.5)	longest vs. shortest quartile	2.54	0.92 (0.74 to 1.14)	+++	EUR (89.5%)	NA	1
PLCO, ATBC, SWHS[119]	Lung adenocarcinoma	2014	NCC	288	288	2.52 (1.38 to 4.6)	longest vs. shortest quartile	2.54	1.44 (1.14 to 1.82)	++	EUR (75%)	NA	1
CCHS, CGPS[109]	Lung cancer	2013	PC	47035	522	1.08 (0.98 to 1.2)	per 1000 bp (1.29 SD) decrease	-1.29	0.94 (0.87 to 1.02)	++++	EUR	<0.001	1
PLCO, ATBC, SWHS[119]	Lung cancer	2014	NCC	847	847	1.86 (1.33 to 2.62)	longest vs. shortest quartile	2.54	1.28 (1.12 to 1.46)	++	EUR (75%)		1
PLCO, ATBC, SWHS[119]	Lung SCC	2014	NCC	163	163	1.14 (0.53 to 2.45)	longest vs. shortest quartile	2.54	1.05 (0.78 to 1.42)	++	EUR (75%)	NA	1
CCHS, CGPS[109]	Melanoma	2013	PC	46805	177	0.89 (0.77 to 1.03)	per 1000 bp (1.29 SD) decrease	-1.29	1.09 (0.98 to 1.23)	++++	EUR		1
WHI, HPFS, NHS[120]	Melanoma	2011	NCC	579	557	0.43 (0.27 to 0.7)	shortest vs. longest quartile	-2.54	1.39 (1.16 to 1.68)	+	EUR	0.03	2
CCHS, CGPS[109]	Ovarian cancer	2013	PC	25367	96	0.85 (0.7 to 1.03)	per 1000 bp (1.29 SD) decrease	-1.29	1.13 (0.98 to 1.32)	+++++	EUR	NA	1
CCHS, CGPS[109]	Pancreatic cancer	2013	PC	47091	124	1.14 (0.93 to 1.41)	per 1000 bp (1.29 SD) decrease	-1.29	0.9 (0.77 to 1.06)	++++	EUR		1
ATBC[121]	Pancreatic cancer	2013	NCC	660	193	1.58 (1.02 to 2.46)	longest vs. shortest quartile	2.54	1.2 (1.01 to 1.42)	++	EUR	0.05	1

EPIC[122]	Pancreatic cancer	2014	NCC	331	331	1.38 (0.8 to 2.41)	longest vs. shortest quartile	2.54	1.13 (0.91 to 1.41)	+	EUR		1
CCHS, CGPS[109]	Prostate cancer	2013	PC	21387	418	0.94 (0.85 to 1.04)	per 1000 bp (1.29 SD) decrease	-1.29	1.05 (0.97 to 1.13)	++++	EUR	0.37	1
HPFS[123]	Prostate cancer	2015	NCC	935	922	1.11 (1.01 to 1.22)	per SD increase	1.00	1.11 (1.01 to 1.22)	++++	EUR		1
NHS[124]	Skin BCC	2011	NCC	1683	363	0.91 (0.66 to 1.25)	longest vs. shortest quartile	2.54	0.96 (0.85 to 1.09)	+	EUR	NA	1
CCHS, CGPS[109]	Testicular cancer	2013	PC	21568	10	1.09 (0.57 to 2.09)	per 1000 bp (1.29 SD) decrease	-1.29	0.94 (0.56 to 1.55)	++++	EUR	NA	1
Non-neoplastic diseases													
Haycock[‖125]	Coronary heart disease	2014	MA	27352	2272	1.4 (1.15 to 1.7)	shortest vs. longest tertile	-2.18	0.86 (0.78 to 0.94)	*	EUR	NA	4
Haycock[#125]	Ischemic stroke	2014	MA	5300	824	1.14 (0.85 to 1.54)	shortest vs. longest tertile	-2.18	0.94 (0.82 to 1.08)	*	EUR	NA	4
Bruneck, SHFS, WHI[126]	Type 2 diabetes	2014	MA	6991	2011	1.31 (1.07 to 1.6)	shortest vs. longest quartile	-2.54	0.9 (0.83 to 0.97)	**	Mix	NA	4

†Search strategy used to identify the study (see Table S4 for details). [‖]Meta-analysis of 11 prospective studies; [#]Meta-analysis of 6 prospective studies (90% of cases were ischemic stroke, 10% were unclassified cerebrovascular disease); [§]To convert reported log RR to log RR per SD increase in telomere length; [‡]**Adjustment for confounders**: +adjusted for age and sex; ++plus smoking; +++plus body mass index; ++++plus alcohol and/or physical activity ; +++++plus hormone replacement therapy, menopause and/or parity; *most studies adjusted for age, sex and non-lipid vascular risk factors; **adjusted for age, sex and body mass index.
Acronyms/abbreviations: **BCC**, basal cell carcinoma**; bp**, base pairs; **CI**, confidence interval; **EA**, East Asian; **EUR**, European; **MA**, random-effects meta-analysis of prospective studies; **NCC**, nested case-control study; **PC**, prospective cohort; **Phet**, p value for heterogeneity between studies; **Pop.**, population; **RR**, relative risk; **SD**, standard deviation; **SCC**, squamous cell carcinoma; **vs.**, versus; **TL**, telomere length. **Study acronyms: ATBC,** Alpha-Tocopherol, Beta-Carotene Cancer Prevention Study**; CCHS**, Copenhagen City Heart Study; **CGPS**, Copenhagen General Population Study; **EPIC**, European Prospective Investigation into Cancer and Nutrition study; **HPFS**, Health Professionals Follow-Up Study; **NHS**, Nurses Health Study; **PHS**, Physicians' Health Study; **PLCO**, Prostate, Lung, Colorectal, and Ovarian; **SHFS**, Strong Heart Family Study; the Sister Study; **SWHS**, Shanghai Women's Health Study; **WHI**, Women's Health Initiative; **WHS**, Women's Health Study

Supplementary Table S4. PubMed search strategy for prospective observational studies of association between telomere length* and disease

Search strategy	Search terms or meta-analysis	No. of studies identified	No. meeting inclusion criteria	Reasons for further exclusions	No. of studies included
Inclusion criteria: prospective study of primary cancer outcome and telomere length†					
Strategy 1	25 February 2015: cancer[TIAB] AND telomere length[TIAB] AND (meta analysis[TIAB] OR prospective[TIAB] OR meta-analysis[TIAB])	54	11	NA	11‡
Strategy 2	25 March 2015: telomere length[Title/Abstract] AND (retrospective[Title/Abstract] OR case-control[Title/Abstract] OR case control[Title/Abstract] OR meta-analysis[Title/Abstract] OR meta analysis[Title/Abstract] OR prospective[Title/Abstract] OR cohort[Title/Abstract] OR cross-sectional[Title/Abstract] OR cross sectional[Title/Abstract]) AND (B-cell non-Hodgkin lymphoma[Title/Abstract] OR breast cancer[Title/Abstract] OR chronic myeloid leukemia[Title/Abstract] OR esophageal adenocarcinoma[Title/Abstract] OR endometrial cancer[Title/Abstract] OR esophageal cancer[Title/Abstract] OR gastric cancer[Title/Abstract] OR gallbladder cancer[Title/Abstract] OR glioma[Title/Abstract] OR head cancer[Title/Abstract] OR neck cancer[Title/Abstract] OR oesophageal adenocarcinoma[Title/Abstract] OR kidney cancer[Title/Abstract] OR melanoma[Title/Abstract] OR nasopharyngeal carcinoma[Title/Abstract] OR neuroblastoma[Title/Abstract] OR non-melanoma skin cancer[Title/Abstract] OR basal cell carcinoma[Title/Abstract] OR squamous cell carcinoma[Title/Abstract] OR ovarian cancer[Title/Abstract] OR pancreatic cancer[Title/Abstract] OR prostate cancer[Title/Abstract] OR testicular germ cell cancer[Title/Abstract] OR Wilm's tumour[Title/Abstract] OR Bladder cancer[Title/Abstract] OR Breast cancer[Title/Abstract] OR Chronic lymphocytic leukemia[Title/Abstract] OR Colorectal cancer[Title/Abstract] OR Multiple myeloma[Title/Abstract] OR Lung adenocarcinoma[Title/Abstract] OR Lung squamous cell cancer[Title/Abstract] OR cancer[Title/Abstract] OR osteosarcoma[Title/Abstract] OR leukemia[Title/Abstract] OR leukaemia[Title/Abstract] OR Ewing sarcoma[Title/Abstract]	209	17	13 duplicates	4
Strategy 3	Ma et al[127] (2011) and Wentzensen et al[128] (2011)	48	10	8 duplicates	2
Inclusion criteria: prospective study of primary disease outcome and telomere length†					
Strategy 4	8 January 2016: (meta-analysis OR "meta analysis") AND "telomere length"	42	7	2 did not report relative risks§; 3 duplicates	2‖

*all identified eligible studies were studies of leukocyte telomere length; ‡1 study reported findings for 2 primary cancer outcomes and 1 study reported findings for 11 primary cancer outcomes; ‖1 meta-analysis reported findings for 2 primary non-neoplastic diseases; †primary outcomes were diseases where a priori statistical power was >50% to detect associations with telomere length (see supplementary text for technical details); see table S1 for a list of the primary disease outcomes; §relative risks were defined as odds ratios, hazard ratios and risk ratios

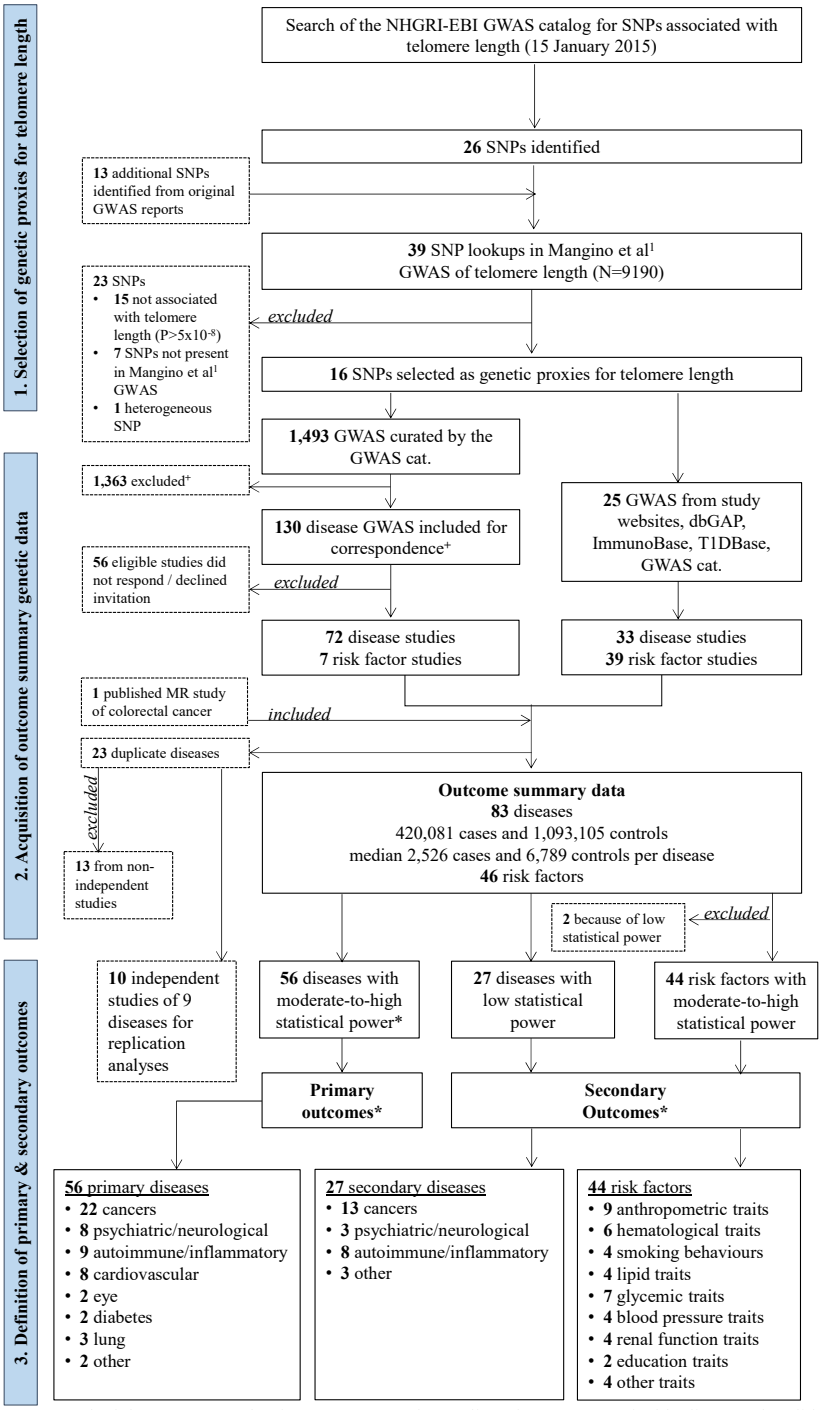
333 **Supplementary Table S6.** Glossary of terms

Formatted: Numbering: Continuous

Mendelian randomization	A technique to appraise causality in observational studies using genetic variants as 'unconfounded' instruments for risk factors or modifiable exposures of interest.
Instrumental variable	A 'proxy' variable used in place of the hypothesized risk factor or exposure in a Mendelian randomization analysis. A valid instrumental variable is associated with the exposure of interest but is not associated with confounders; and is associated with the outcome (e.g. disease) exclusively via its effect on the hypothesized exposure (see Supplementary Figure S7 for an illustration of these assumptions).
Reverse causation	When the outcome causes variation in the hypothesized exposure and not *vice versa*.
Confounding	When the association between exposure and outcome is not due to a causal relationship between the two variables but arises as a result of the separate effects of a third variable (the confounder) on the exposure and the outcome. Mendelian randomization studies are less susceptible to confounding in comparison to observational studies (but confounding by pleiotropy or population stratification is possible).
Pleiotropy	Occurs when a genetic variant is associated with multiple traits or phenotypes. Vertical pleiotropy occurs when the phenotypes are on the same causal pathway (and is less problematic for Mendelian randomization studies). Horizontal pleiotropy occurs if the phenotypes are associated with the genetic variant via separate pathways and can introduce confounding into a Mendelian randomization analysis. Sensitivity analyses, such as MR-Egger, the weighted median, scatter plots and funnel plots, can be used to test and, in some instances, adjust for pleiotropy.
Collider bias	The phenomenon by which statistical adjustment for a variable, M (known as the collider), that is a downstream consequence of both the exposure X and the outcome Y, induces an association between X and Y that was not previously present, and therefore leads to bias. In MR, if published genetic associations with the exposure and/or outcome are adjusted for a collider, this may lead to collider bias.
Weak instrument bias	Occurs when the instrument is only weakly associated with the exposure. Can introduce confounding into a Mendelian randomization analysis when the exposure and outcome data come from the same sample. When exposure and outcome data come from separate samples, as in two-sample Mendelian randomization, bias is towards the null. An F statistic > 10, for the association between the instrument and exposure, is sometimes used as a threshold for defining strong instruments, although weak instrument bias varies continuously with the strength of the F statistic.

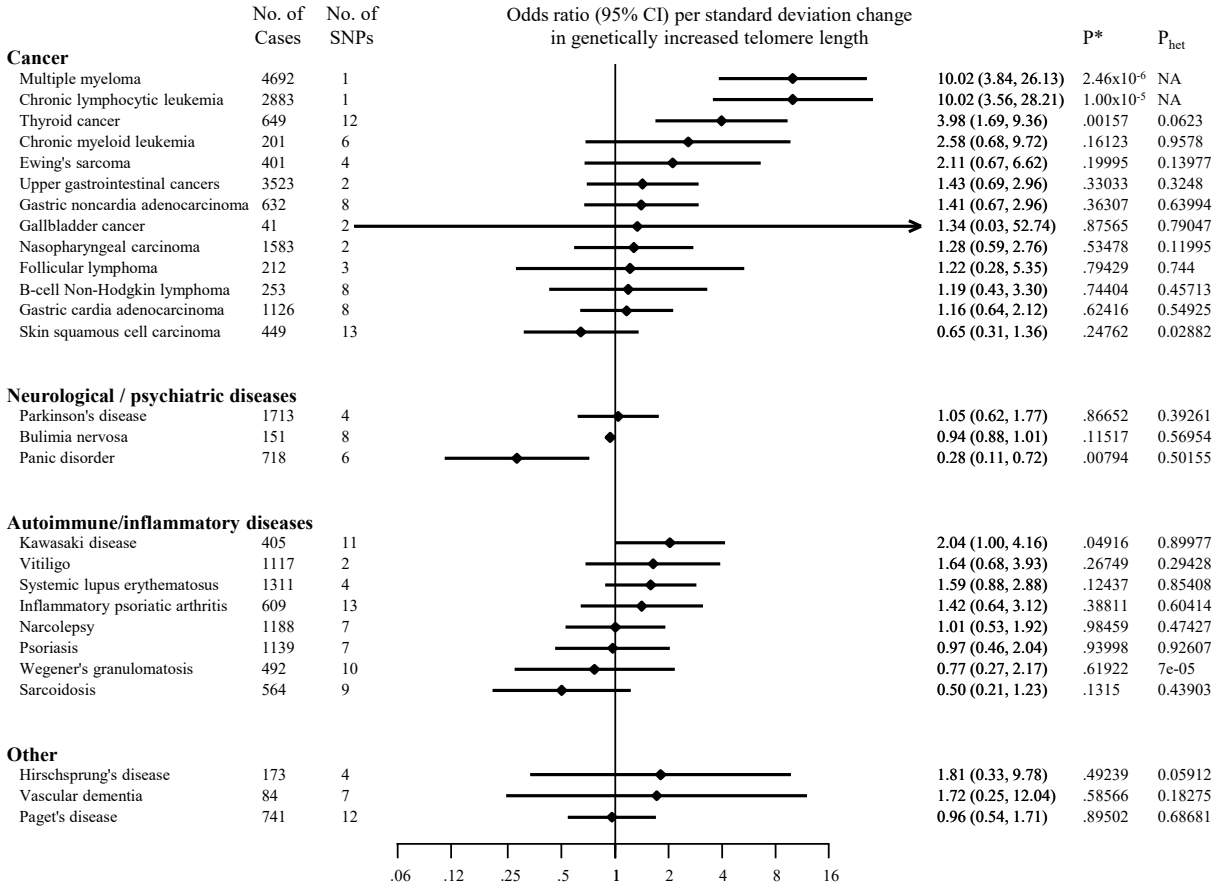
334

Supplementary Figure S1. Study design



+We searched the GWAS catalog in January 2015 for studies of non-communicable diseases that did not select controls on the basis of pre-existing conditions. Of the 1493 studies in the GWAS catalog with unique PubMed reference numbers, we classified 773 as disease studies (the excluded non-disease studies were typically studies of risk factors for disease, biomarkers or response to treatments). A further 103 studies were excluded for the following reasons: studies of infectious diseases, studies of congenital abnormalities, studies of (not-cause specific) mortality, studies nested within disease populations and studies using pooled DNA samples. Of the 670 remaining non-communicable disease studies, 130 were identified for correspondence. Our objective was to obtain the single largest available study for each non-communicable disease, so as to avoid unnecessary correspondence with duplicate studies and to avoid including studies with overlapping samples. *Primary outcomes were diseases with sufficient cases and controls for >50% power and secondary outcomes were diseases with <50% power to detect odds ratios ≥2.0 per standard deviation change in genetically increased telomere length (alpha assumed to be 0.01). All risk factors were classified as secondary outcomes. **GWAS**, genome-wide association study; **GWAS Cat.**, NHGRI-EBI GWAS catalogue; **SNP**, single nucleotide polymorphism; **NHGRI**, National Human Genome Research Institute; **EBI,** European Bioinformatics Institute

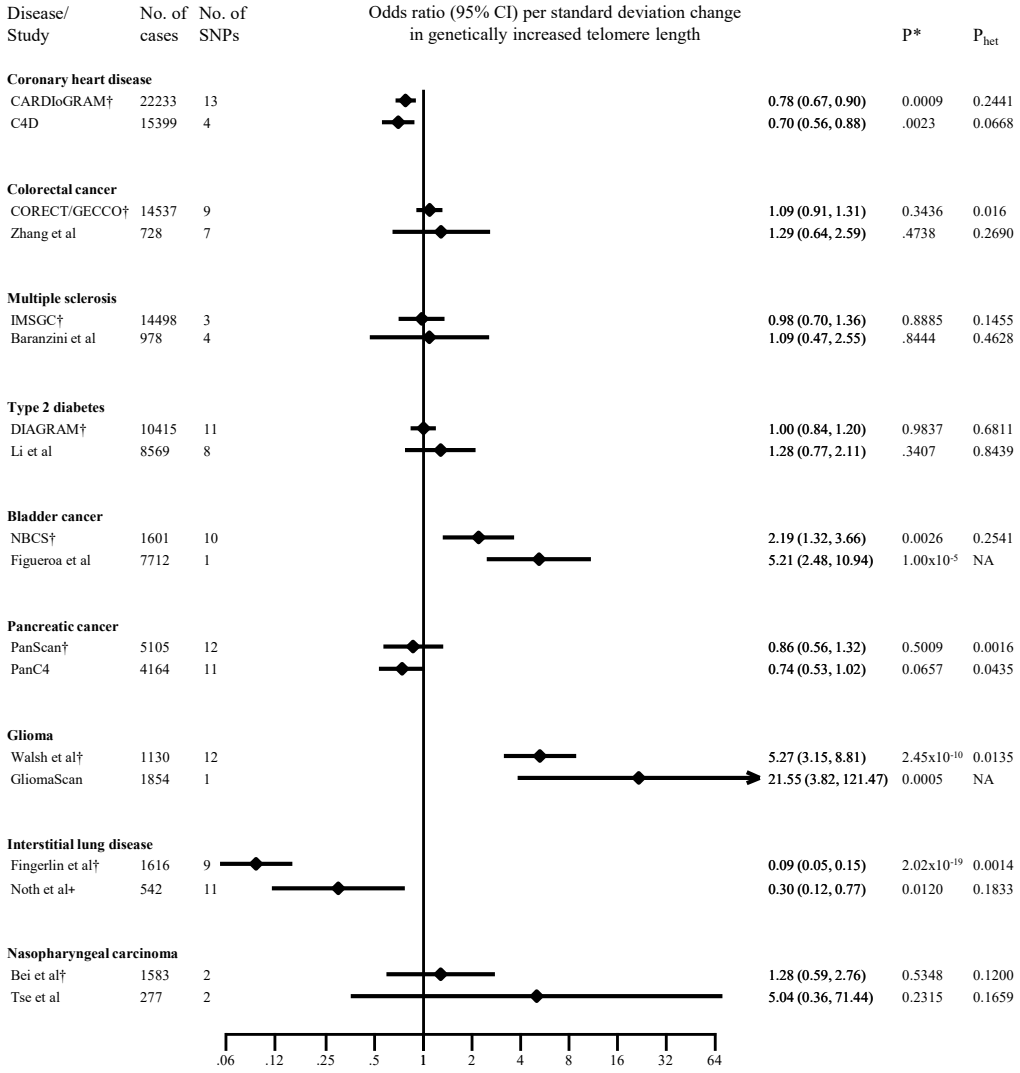
Supplementary Figure S2. Association between genetically increased telomere length and odds of secondary non-communicable diseases



*P value for association between genetically increased telomere length and disease from maximum likelihood; P_{het}, P value for heterogeneity amongst SNPs within the genetic risk score; SNP, single nucleotide polymorphism; CI, confidence interval

Formatted: Numbering: Continuous

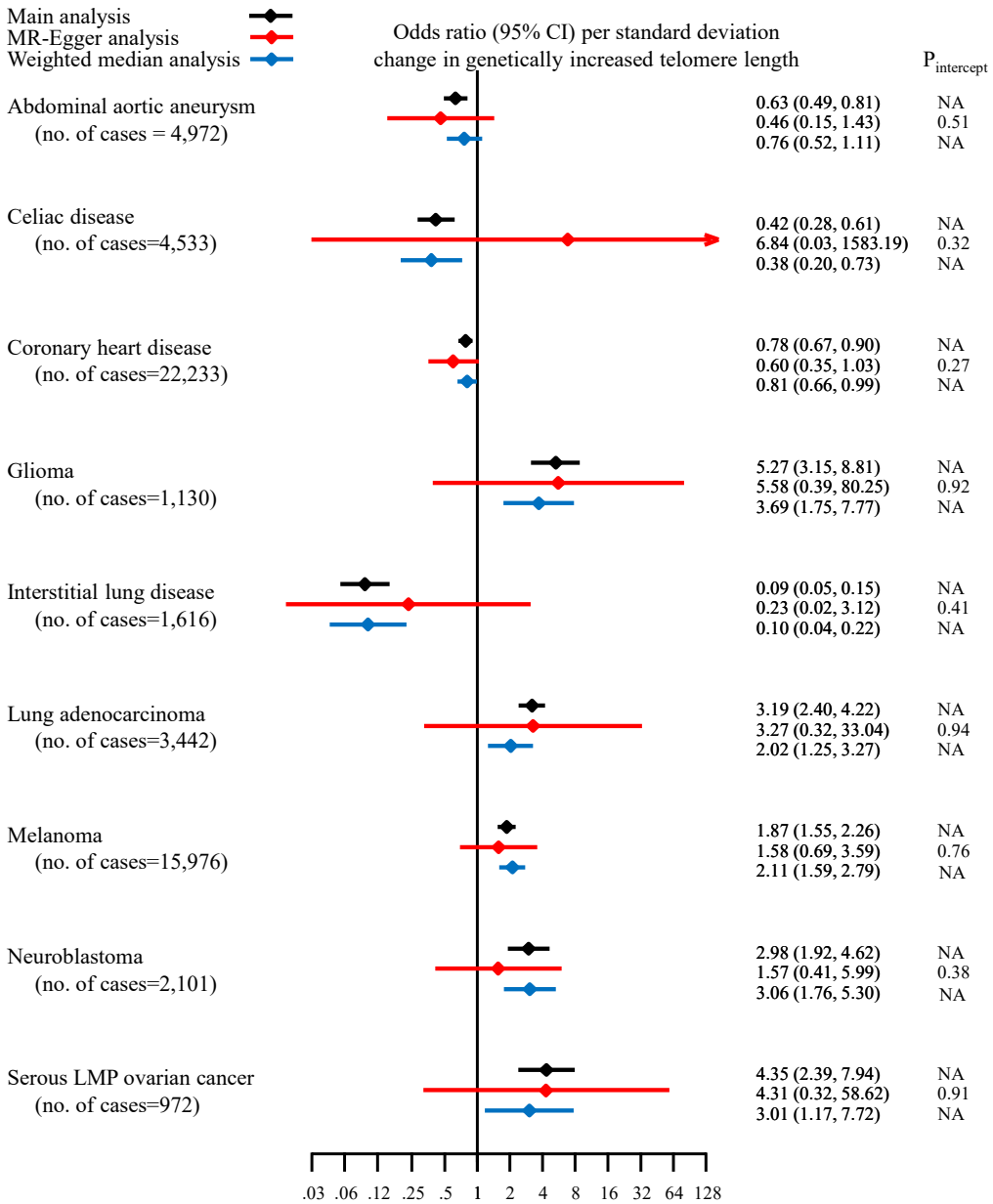
Supplementary Figure S3. Replication of association between genetically increased telomere length and odds of non-communicable diseases in independent datasets



*P value for association between genetically increased telomere length and disease from maximum likelihood. †Primary or secondary study from Fig. 1 or Fig. S2. +Noth et al[81]: ≤17% of the cases overlapped with cases from Fingerlin et al[31] and 77% of cases had idiopathic pulmonary fibrosis; ‡An inverse association was also observed in Mushiroda et al[82]. P_{het}, p value for heterogeneity amongst SNPs in the genetic risk score (NA when only a single SNP available); SNP, single nucleotide polymorphism; CI, confidence interval. **Study abbreviations**: **C4D**, Coronary Artery Disease Genetics Consortium; **CARDIoGRAM**, Coronary ARtery DIsease Genome wide Replication and Meta-analysis; **CORECT**, ColoRectal Transdisciplinary Study; **GECCO**, Genetics and Epidemiology of Colorectal Cancer Consortium; **IMSGC**, International Multiple Sclerosis Genetic Consortium; **NBCS**, Nijmegen Bladder Cancer Study; **IMSGC**, International Multiple Sclerosis Genetic Consortium.

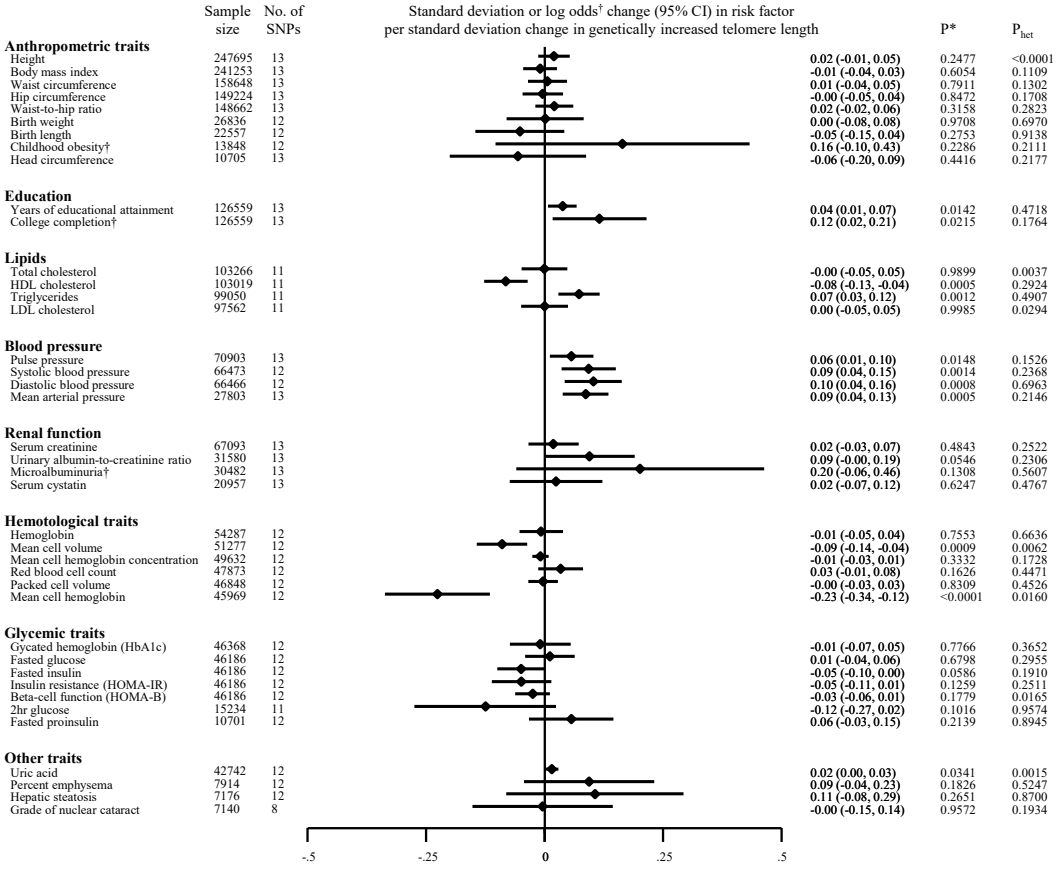
Formatted: Numbering: Continuous

Supplementary Figure S4. Sensitivity analyses of association between genetically increased telomere length and odds of non-communicable diseases



LMP, low malignancy potential; CI, confidence interval. The $P_{intercept}$ from MR-Egger regression tests the null hypothesis that the intercept is zero and can be interpreted as a statistical test for the presence of directional (bias inducing) pleiotropy; the smaller the $P_{intercept}$ value the stronger the evidence for directional pleiotropy.

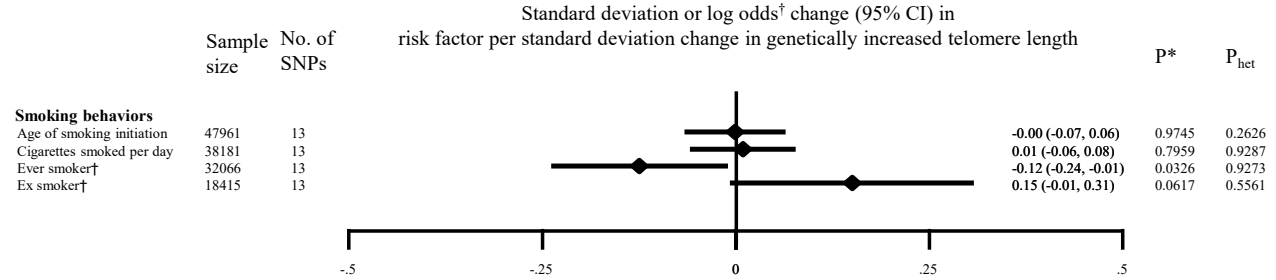
Supplementary Figure S5. Association between genetically increased telomere length and risk factors for non-communicable diseases



*P value for association between genetically increased telomere length and risk factor from maximum likelihood; P_{het}, p value for heterogeneity amongst SNPs within the genetic risk score; SNP, single nucleotide polymorphism; CI, confidence interval; HbA1c, hemoglobin A1c; HOMA-B, homeostatic model assessment β-cell function; IR, insulin resistance; †for binary risk factors results reflect the log odds ratio for the risk factor, all other results reflect the standard deviation change in the risk factor

Formatted: Numbering: Continuous

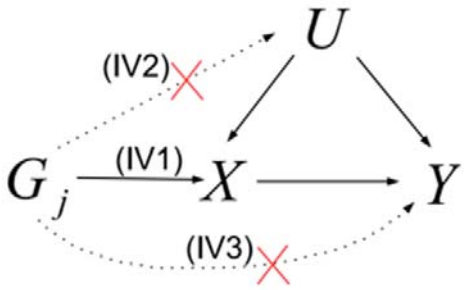
Supplementary Figure S6. Association between genetically increased telomere length and smoking



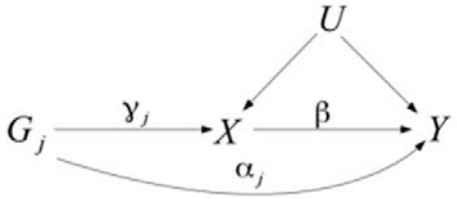
*P value for association between genetically increased telomere length and risk factor from maximum likelihood; P_{het}, P value for heterogeneity amongst SNPs within the genetic risk score; SNP, single nucleotide polymorphism; CI, confidence interval; †for binary risk factors results reflect the log odds ratio for the risk factor, all other results reflect the standard deviation change in the risk factor

Supplementary Figure S7. Causal diagram illustrating the assumptions of Mendelian randomization

a)



b)



IV, instrumental variable assumption; G_j, single nucleotide polymorphism j; X, telomere length; Y, outcome (disease or risk factor); U, confounder; α, G-Y association not mediated by telomere length (often described as a horizontal pleiotropic or direct effect); γ, SNP-telomere-length association.

a) Key assumptions of Mendelian randomization. G_j is associated with X (IV1); G_i is independent of confounders (IV2); G_i is independent of Y given X and U (IV3). The weighted median approach assumes that IV1-IV3 hold for genetic variants making up at least 50% of the weight in the analysis; MR-Egger relaxes assumption IV3 (see InSIDE assumption below).

b) Assumptions underlying the MR-Egger approach. IV3 is replaced with the InSIDE assumption (Instrument Strength Independent of Direct Effect): the strength of the pleiotropic effect (αj) does not correlate with the strength of the G-X association (γj). Under the InSIDE assumption, MR-Egger can consistently estimate the causal effect of X on Y, represented by the parameter β in (b).

Formatted: Numbering: Continuous

ACKNOWLEDGEMENTS OF THE CONTRIBUTING STUDIES AND CONSORTIA

Amyotrophic lateral sclerosis GWAS consortium

Isabella Fogh[1], Kuang Lin[1], John F. Powell[1], the SLAGEN Consortium, Vincenzo Silani[2], the ALSGEN consortium, Orla Hardiman[3], Robert H. Brown[4], Ammar Al-Chalabi[1], Jan H. Veldink[5].

1. Department of Basic and Clinical Neuroscience, Maurice Wohl Clinical Neuroscience Institute, Institute of Psychiatry, Psychology and Neuroscience, King's College London, London, United Kingdom

2. Department of Neurology and Laboratory of Neuroscience, IRCCS Istituto Auxologico Italiano, Milano, Italy

3. Population Genetics Laboratory, Smurfit Institute of Genetics, Trinity College Dublin, Dublin, Republic of Ireland

4. Department of Neurology, University of Massachusetts Medical School, Worcester, Massachusetts, United States of America

5. Department of Neurology and Neurosurgery, Brain Center Rudolf Magnus, University Medical Center Utrecht, The Netherlands

Funding/Support

I. Fogh was supported by funds from Motor Neurone Disease Association of Great Britain and Northern Ireland (grant n.905-793, 6058).

J.Powell, A.Al-Chalabi and I.Fogh received salary support from the National Institute for Health Research (NIHR) Dementia Biomedical Research Unit at South London and Maudsley NHS Foundation Trust and King's College London. The UK National DNA Bank for MND Research was funded by the Motor Neurone Disease Association (grant 3/3), the Wellcome Trust (grant

070122/A/02/Z) and the NIHR Dementias and Neurodegenerative Diseases Research Network (DeNDRoN).

V. Silani was supported by Agenzia Italiana per la Ricerca sulla SLA-AriSLA (grant NOVALS 2012 cofinanced with the contribution of 5 x 1000, Healthcare Research support of the Ministry of Health), the Italian Ministry of Health (Grant ALS-FTD, Ric. Finalizzata 2009 no.276) and Associazione Amici "Centro Dino Ferrari".

J.H. Veldink was supported by the Netherlands Organisation for Health Research and Development.

The Aneurysm Consortium

GWAS data on abdominal aortic aneurysm (AAA) studies

All known studies with AAA genome-wide genotyping were invited to join the International Aneurysm Consortium. All studies agreed to participate in the meta-GWAS, with cohort case control descriptions and inclusion/exclusion criteria having been previously reported.[28,129,130] All AAA cases shared a common definition of infra-renal aortic diameter >30 mm.

Descriptions of AAA cohorts

In the present report, the Aneurysm Consortium consists of the original Aneurysm Consortium plus the NZ AAA Genetics Study (two separate cohorts), the Geisinger Vascular Clinic AAA study, the Iceland study and the Netherlands study.

Original Aneurysm Consortium (1846 cases and 5605 controls): The original Aneurysm Consortium recruited cases of AAA from centres across the United Kingdom and Western Australia. Cases were defined as an infra-renal aortic diameter ≥ 30 mm proven on ultrasound or computerized tomography (CT) scan. Controls were taken from the WTCCC2 common control group[28,131] and were therefore unscreened for AAA.

NZ AAA Genetics Study (with two separate cohorts: set 1 with 608 cases and 612 controls; set 2 with 397 cases and 384 controls): The Vascular Research Consortium of New Zealand recruited

New Zealand men and women with a proven history of AAA (infra-renal aortic diameter ≥ 30 mm proven on ultrasound or CT scan). Approximately 80% had undergone surgical AAA repair (typically AAA's > 50-55 mm in diameter). The vast majority of cases (>97%) were of Anglo-European ancestry. The control group underwent an abdominal ultrasound scan to exclude (>25 mm) concurrent abdominal aortic aneurysm and Anglo-European ancestry was required for inclusion. Controls were also screened for peripheral artery disease (PAD; using ankle brachial index), carotid artery disease (ultrasound) and other cardiovascular risk factors.

Geisinger Vascular Clinic AAA Study, Pennsylvania, USA: AAA patients (n=724) were enrolled through the Department of Vascular Surgery at Geisinger Medical Center, Danville, PA. Details of this case-control set have been reported previously, and the samples have been used in previous association studies.[129,132] To identify cases and controls from the electronic medical records, an ePhenotyping algorithm was developed[29]. AAA cases were defined as infrarenal aortic diameter ≥ 30 mm as revealed by abdominal imaging. Approximately 20% of individuals with AAA had a family history of AAA. A control group (n=1231) was obtained through the Geisinger MyCode® Project, a cohort of Geisinger Clinic patients recruited for genomic studies. The MyCode® controls were matched for age distribution and sex to the Geisinger Vascular Clinic AAA cases. Based on electronic medical records, controls had no ICD-9 codes for AAA in their records, but they were not screened by ultrasonography for AAA. Both cases and controls from the Geisinger Clinic were of European descent. The eMERGE Network Imputed GWAS for 41 Phenotypes (the dbGaP eMERGE Phase 1 and 2 Merged data Submission) accession number is: phs000888.v1.p1 which includes the Geisinger AAA data.

Iceland, deCODE Genetics: Icelandic individuals with AAA (defined as infra-renal aortic diameter ≥ 30 mm) were recruited from a registry of individuals who were admitted at Landspitali University Hospital, in Reykjavik, Iceland, 1980 – 2006. AAA patients were either followed up or treated by

intervention for emergency repair of symptomatic or ruptured AAA or for an elective repair by surgery or endovascular intervention. In total, whole genome data from 557 subjects with AAA, enrolled as part of the CVD genetics program at deCODE, were included in the metaGWAS. The Icelandic controls used (n=89,235) were selected from among individuals who have participated in various GWA studies and who were recruited as part of genetic programs at deCODE. Individuals with known cardiovascular disease were excluded as controls[129] but controls were unscreened for AAA.

The Netherlands: The AAA sample set from Utrecht was recruited in 2007-2009 from eight centres in The Netherland[129], mainly when individuals visited their vascular surgeon in the polyclinic or, in rare cases, during hospital admission for elective or emergency AAA surgery. An AAA was defined as an infrarenal aorta ≥ 30 mm. The sample set (n=840) comprised 89.9% males, with a mean AAA diameter of 58.4 mm, 61.7% had received surgery, of which 8.1 % was after rupture. The Dutch controls (n=2791) used in the AAA GWAS were recruited as part of the Nijmegen Biomedical Study and the Nijmegen Bladder Cancer Study (see http://dceg.cancer.gov/icbc/membership.html).

Meta-analysis of AAA GWASs

Data from the six cohorts detailed above, comprising 4972 AAA cases and 99,858 controls, that were genotyped with a variety of genome-wide SNP arrays. All cohorts underwent quality control filtering using the manufacturers' array-specific guidelines but with consistently applied inclusion criteria of SNP or sample call rates >95% and Hardy-Weinberg equilibrium $P>5x10^{-5}$ in controls.[28,129,130,132] Each cohort then underwent imputation (Impute 2.2) to a shared reference panel from the 1000 Genomes project (Phase I integrated variant set release (v3), March 2012, NCBI build 37(hg19 Following imputation SNPs were quality controlled by quality score (Q>0.9) and minor allele frequency (MAF>0.05 in controls) filtering, resulting in a common set of 5331120 SNPs across all discovery phase participants.

The metaGWAS analysis was conducted using the METAL software package[133] on the BCISNPmax database platform (version 3.5, BCI Platforms, Espoo, Finland). METAL was implemented using the sample size scheme with weighting for each cohort being two times the case number. The analysis was adjusted for genomic inflation (λ) in each cohort.

Acknowledgements on AAA GWAS studies:

Data provided by the original Aneurysm Consortium was funded by the Wellcome Trust (award number 084695) and made use of data generated by the WTCCC. A full list of the investigators who contributed to the generation of the data is available from www.wtccc.org.uk. Funding for the WTCCC project was provided by the Wellcome Trust under award 076113 and 085475. Funding for the New Zealand project was provided by the Health Research Council of New Zealand (08-75, 14-155). The Geisinger sample collection was funded in part by the Pennsylvania Commonwealth Universal Research Enhancement program, the Geisinger Clinical Research Fund, the American Heart Association, and the Ben Franklin Technology Development Fund of Pennsylvania. The generation and management of GWAS genotype data for the Rotterdam Study (control samples for the Dutch GWAS) is supported by the Netherlands Organization of Scientific Research NWO Investments (nr. 175.010.2005.011, 911-03-012). This study is funded by the Research Institute for Diseases in the Elderly (014-93-015; RIDE2), the Netherlands Genomics Initiative (NGI)/NWO project nr. 050-060-810.

Australian Asthma Genetics Consortium

David L. Duffy[a] , Dale R. Nyholt[a], John Beilby[b-d], Svetlana Baltic,[e] Loren Price[e], Faang Cheah[e], Desiree Mészáros[f], Scott D. Gordon[a], Melissa C. Southey[g], Margaret J. Wright[a], James Markos[h], Li P. Chung[e], Anjali K. Henders[a], Graham Giles[i], Suzanna Temple[e], John Whitfield[a], Brad Shelton[e], Chalermchai Mitrpant[e], Minh Bui, PhD,[j] Mark Jenkins[j], Haydn Walters[f], Michael J. Abramson[k],

521 Michael Hunter[l,d], Bill Musk [l,d,m,n], Peter Le Souëf,[o] Shyamali C. Dharmage[j], Grant W.
522 Montgomery,[a] Alan James,[c,m,d], Nicholas G. Martin[a] , Melanie C. Mathesonj
523
524 [a] QIMR Berghofer Medical Research, Brisbane, Australia.
525 [b] PathWest Laboratory Medicine of Western Australia (WA), Nedlands, Australia.
526 [c] School of Pathology and Laboratory Medicine, The University of WA, Nedlands, Australia.
527 [d] Busselton Population Medical Research Foundation, Sir Charles Gairdner Hospital, Perth,
528 Australia.
529 [e] Institute of Respiratory Health, University of WA, Perth, Australia.
530 [f] Menzies Research Institute, Hobart, Australia.
531 [g] Department of Pathology, The University of Melbourne, Melbourne, Australia.
532 [h] Launceston General Hospital, Launceston, Australia.
533 [i] Cancer Epidemiology Centre, The Cancer Council Victoria, Melbourne, Australia.
534 [j] Centre for Molecular, Environmental, Genetic and Analytic Epidemiology, University of
535 Melbourne, Melbourne, Australia.
536 [k] Department of Epidemiology & Preventive Medicine, Monash University, Melbourne, Australia
537 [l] School of Population Health, The University of WA, Nedlands, Australia
538 [m] School of Medicine and Pharmacology, University of Western Australia, Nedlands, Australia
539 [n] Department of Respiratory Medicine, Sir Charles Gairdner Hospital, Perth, Australia
540 [o] School of Paediatrics and Child Health, Princess Margaret Hospital for Children, Perth, Australia
541

542 **Coronary ARtery DIsease Genome wide Replication and Meta-analysis (CARDIoGRAM)**
543 **consortium and The Coronary Artery Disease (C4D) Genetics consortium**
544 We thank the CARDIoGRAM and C4D consortia for making summary data available to the
545 research community. Data on coronary artery disease / myocardial infarction have been contributed
546 by CARDIoGRAMplusC4D investigators and have been downloaded from

www.CARDIOGRAMPLUSC4D.ORG. The investigators within CARDIoGRAM and C4D did not participate in the analysis, writing or interpretation of this report.

The Cohorts for Heart and Aging Research in Genomic Epidemiology (CHARGE) – Heart Failure Working Group

For a full list of CHARGE – Heart Failure working group members contributing to this work and for CHARGE – Heart Failure acknowledgements please see PMID 20445134.

CHARGE - Sudden Cardiac Arrest Working Group

Aravinda Chakravarti[1], Anna Moes[1], Dan E. Arking[1], Foram N. Ashar[1], Georg Ehret[1], Josef Coresh[2], Man Li[2], Ronald Prineas[3], Angel Mak[4], Pui-Yan Kwok[4], Catherine O. Johnson[5], Nona Sotoodehnia[5], David S. Siscovick[6], Oscar H. Franco[7], Thomas Lumley[8], Florence Dumaso[9], Xavier Jouven[9], Martina Muller-Nurasyid[10], Stefan Kaab[10], Barbara M. McKnight[5], Bruce M. Psaty[5], Jennifer A. Brody[5], Jerome I. Rotter[11], Ken Rice[5], Rozenn N. Lemaitre[5], Christopher J. O'Donnell[12], Christopher Newton-Cheh[13], Shih-Jen Hwang[12], Heikki Huikuri[14], Marja-Leena Kortelainen[14],M Juhani Junttila[14], Jean-Claude Tardif[15], John D. Rioux[15],Philippe Goyette[15], Christine M. Albert[16], Martin VanDenBurgh[16], Sara Pulit[17], Andre G Uitterlinden[2],Albert Hofman[2],Bruno H Stricker[2],Mark Eijgelsheim[2]

1. Institute of Genetic Medicine,Johns Hopkins,Baltimore,USA,21205
2. Department of Epidemiology,Johns Hopkins University,Baltimore,USA,21205
3. Public Health Sciences, Wake Forest University,Winston-Salem,USA,27157
4. Cardiovascular Research Institute and Institute for Human Genetics, University of California, San Francisco,San Francisco,USA,
5. Cardiovascular Health Research Unit, Department of Biostatistics,University of Washington,Seattle,USA,98101

6. New York Academy of Medicine,New York,USA,
7. Department of Epidemiology,Erasmus MC,Erasmus,The Netherlands,
8. Department of Statistics, University of Auckland,Auckland,NZ,
9. Paris sudden Death Expertise Center,University Paris Sorbonne cité,Paris,France,
10. Department of Medicine I ,Ludwig-Maximilians University,Munich,Germany,
11. Institute for Translational Genomics and Population Sciences, Los Angeles Biomedical Research Institute, Departments of Pediatrics and Medicine,Harbor-UCLA Medical Center,Los Angeles,USA
12. NHLBI Framingham Heart Study,Boston,USA,
13. Center for Human Genetic Research & Cardiovascular Research Center,Massachusetts General Hospital,Boston,USA,
14. Internal Medicine,University of Oulu,Oulu,Finland,
15. Montreal Heart Intitute,University of Montreal,Quebec,Canada,
16. Divisions of Preventive Medicine and Cardiovascular Medicine, Department of Medicine,Brigham and Women's Hospital,Boston,USA,
17. Department of Genetics,University Medical Centre Utrecht,Utrecht,The Netherlands,

The Genetic Epidemiology of Chronic Obstructive Pulmonary Disease (COPDGene)

The COPDGene project was supported by Award Number R01HL089897 and Award Number R01HL089856 from the National Heart, Lung, And Blood Institute. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Heart, Lung, And Blood Institute or the National Institutes of Health. The COPDGene project is also supported by the COPD Foundation through contributions made to an Industry Advisory Board comprised of AstraZeneca, Boehringer Ingelheim, Novartis, Pfizer, Siemens, Sunovion, and GlaxoSmithKline.

598 *COPDGene Administrative Core*: James Crapo, MD (PI), Edwin Silverman, MD, PhD (PI), Barry
599 Make, MD, Elizabeth Regan, MD, PhD
600
601 *COPDGene Genetic Analysis Core*: Terri Beaty, PhD, Nan Laird, PhD, Christoph Lange, PhD,
602 Michael Cho, MD, Stephanie Santorico, PhD, John Hokanson, MPH, PhD, Dawn DeMeo, MD,
603 MPH, Nadia Hansel, MD, MPH, Craig Hersh, MD, MPH, Peter Castaldi, MD, MSc, Merry-Lynn
604 McDonald, PhD, Emily Wan, MD, Megan Hardin, MD, Jacqueline Hetmanski, MS, Margaret
605 Parker, MS, Marilyn Foreman, MD, Brian Hobbs, MD, Robert Busch, MD, Adel El-Boueiz, MD,
606 Peter Castaldi, MD, Megan Hardin, MD, Dandi Qiao, PhD, Elizabeth Regan, MD, Eitan Halper-
607 Stromberg, Ferdouse Begum, Sungho Won, Sharon Lutz, PhD
608 *COPDGene Imaging Core*: David A Lynch, MB, Harvey O Coxson, PhD, MeiLan K Han, MD,
609 MS, MD, Eric A Hoffman, PhD, Stephen Humphries MS, Francine L Jacobson, MD, Philip F Judy,
610 PhD, Ella A Kazerooni, MD, John D Newell, Jr., MD, Elizabeth Regan, MD, James C Ross, PhD,
611 Raul San Jose Estepar, PhD, Berend C Stoel, PhD, Juerg Tschirren, PhD, Eva van Rikxoort, PhD,
612 Bram van Ginneken, PhD, George Washko, MD, Carla G Wilson, MS, Mustafa Al Qaisi, MD,
613 Teresa Gray, Alex Kluiber, Tanya Mann, Jered Sieren, Douglas Stinson, Joyce Schroeder, MD,
614 Edwin Van Beek, MD, PhD
615 *COPDGene PFT QA Core, Salt Lake City, UT*: Robert Jensen, PhD
616 *COPDGene Data Coordinating Center and Biostatistics*, *National Jewish Health, Denver, CO*:
617 Douglas Everett, PhD, Anna Faino, MS, Matt Strand, PhD, Carla Wilson, MS
618 *COPDGene Epidemiology Core*, *University of Colorado Anschutz Medical Campus, Aurora, CO*:
619 John E. Hokanson, MPH, PhD, Gregory Kinney, MPH, PhD, Sharon Lutz, PhD, Kendra Young
620 PhD, Katherine Pratte, MSPH, Lindsey Duca, MS
621
622 *COPDGene Clinical Centers*
623 *Ann Arbor VA:* Jeffrey L. Curtis, MD, Carlos H. Martinez, MD, MPH, Perry G. Pernicano, MD

Baylor College of Medicine, Houston, TX: Nicola Hanania, MD, MS, Philip Alapat, MD, Venkata Bandi, MD, Mustafa Atik, MD, Aladin Boriek, PhD, Kalpatha Guntupalli, MD, Elizabeth Guy, MD, Amit Parulekar, MD, Arun Nachiappan, MD

Brigham and Women's Hospital, Boston, MA: Dawn DeMeo, MD, MPH, Craig Hersh, MD, MPH, George Washko, MD, Francine Jacobson, MD, MPH

Columbia University, New York, NY: R. Graham Barr, MD, DrPH, Byron Thomashow, MD, John Austin, MD, Belinda D'Souza, MD, Gregory D.N. Pearson, MD, Anna Rozenshtein, MD, MPH, FACR

Duke University Medical Center, Durham, NC: Neil MacIntyre, Jr., MD, Lacey Washington, MD, H. Page McAdams, MD

Health Partners Research Foundation, Minneapolis, MN: Charlene McEvoy, MD, MPH, Joseph Tashjian, MD

Johns Hopkins University, Baltimore, MD: Robert Wise, MD, Nadia Hansel, MD, MPH, Robert Brown, MD, Karen Horton, MD, Nirupama Putcha, MD, MHS,

Los Angeles Biomedical Research Institute at Harbor UCLA Medical Center, Torrance, CA: Richard Casaburi, PhD, MD, Alessandra Adami, PhD, Janos Porszasz, MD, PhD, Hans Fischer, MD, PhD, Matthew Budoff, MD, Harry Rossiter, PhD

Michael E. DeBakey VAMC, Houston, TX: Amir Sharafkhaneh, MD, PhD, Charlie Lan, DO

Minneapolis VA: Christine Wendt, MD, Brian Bell, MD

Morehouse School of Medicine, Atlanta, GA: Marilyn Foreman, MD, MS, Gloria Westney, MD, MS, Eugene Berkowitz, MD, PhD

National Jewish Health, Denver, CO: Russell Bowler, MD, PhD, David Lynch, MD

Reliant Medical Group, Worcester, MA: Richard Rosiello, MD, David Pace, MD

Temple University, Philadelphia, PA: Gerard Criner, MD, David Ciccolella, MD, Francis Cordova, MD, Chandra Dass, MD, Gilbert D'Alonzo, DO, Parag Desai, MD, Michael Jacobs, PharmD, Steven Kelsen, MD, PhD, Victor Kim, MD, A. James Mamary, MD, Nathaniel Marchetti, DO,

Aditi Satti, MD, Kartik Shenoy, MD, Robert M. Steiner, MD, Alex Swift, MD, Irene Swift, MD, Maria Elena Vega-Sanchez, MD

University of Alabama, Birmingham, AL: Mark Dransfield, MD, William Bailey, MD, J. Michael Wells, MD, Surya Bhatt, MD, Hrudaya Nath, MD

University of California, San Diego, CA: Joe Ramsdell, MD, Paul Friedman, MD, Xavier Soler, MD, PhD, Andrew Yen, MD

University of Iowa, Iowa City, IA: Alejandro Cornellas, MD, John Newell, Jr., MD, Brad Thompson, MD

University of Michigan, Ann Arbor, MI: MeiLan Han, MD, Ella Kazerooni, MD, Carlos Martinez, MD

University of Minnesota, Minneapolis, MN: Joanne Billings, MD, Tadashi Allen, MD

University of Pittsburgh, Pittsburgh, PA: Frank Sciurba, MD, Divay Chandra, MD, MSc, Joel Weissfeld, MD, MPH, Carl Fuhrman, MD, Jessica Bon, MD

University of Texas Health Science Center at San Antonio, San Antonio, TX: Antonio Anzueto, MD, Sandra Adams, MD, Diego Maselli-Caceres, MD, Mario E. Ruiz, MD

Early Growth Genetics (EGG) Consortium

Summary data on birth anthropometrics has been contributed by the EGG Consortium and has been downloaded from www.egg-consortium.org.The investigators within the EGG did not participate in the analysis, writing or interpretation of this paper.

The EArly Genetics and Lifecourse Epidemiology (EAGLE) consortium

Lavinia Paternoster [1, 2, 112], Marie Standl [3, 112], Johannes Waage [4], Hansjörg Baurecht [5], Melanie Hotze [5], David P Strachan [6], John A Curtin [7], Klaus Bønnelykke [4], Chao Tian [8], Atsushi Takahashi [9], Jorge Esparza-Gordillo [10, 11], Alexessander Couto Alves [12], Jacob P Thyssen [13], Herman T den Dekker [14, 15, 16], Manuel A Ferreira [17], Elisabeth Altmaier [18, 19, 20], Patrick MA

Sleiman [21, 22], Feng Li Xiao [23], Juan R Gonzalez [24], Ingo Marenholz [10, 11], Birgit Kalb [10, 25], Maria Pino-Yanes [26, 27, 28], Cheng-Jian Xu [29, 30], Lisbeth Carstensen [31], Maria M Groen-Blokhuis [32], Cristina Venturini [33], Craig E Pennell [34], Sheila J Barton [35], Albert M Levin [36], Ivan Curjuric [37, 38], Mariona Bustamante [24, 39, 40, 41], Eskil Kreiner-Møller [4], Gabrielle A Lockett [42], Jonas Bacelis [43], Supinda Bunyavanich [44], Rachel A Myers [45], Anja Matanovic [10, 11], Ashish Kumar [37, 38, 46, 47], Joyce Y Tung [8], Tomomitsu Hirota [48], Michiaki Kubo [49], Wendy L McArdle [2], A J Henderson [2], John P Kemp [1, 2, 50], Jie Zheng [1, 2], George Davey Smith [1, 2], Franz Rüschendorf [10], Anja Bauerfeind [10], Min Ae Lee-Kirsch [51], Andreas Arnold [52], Georg Homuth [53], Carsten O Schmidt [54], Elisabeth Mangold [55], Sven Cichon [55, 56, 57, 58], [59], Thomas Keil [60, 61], Elke Rodríguez [5], Annette Peters [19, 62], Andre Franke [63], Wolfgang Lieb [64], Natalija Novak [65], Regina Fölster-Holst [5], Momoko Horikoshi [47], Juha Pekkanen [66, 67], Sylvain Sebert [68, 69], Lise L Husemoen [70], Niels Grarup [71], Johan C de Jongste [14], Fernando Rivadeneira [15, 16, 72], Albert Hofman [15], Vincent WV Jaddoe [14, 15, 16], Suzanne GMA Pasmans [73], Niels J Elbert [16, 73], André G Uitterlinden [15, 72], Guy B Marks [74], Philip J Thompson [75, 76], Melanie C Matheson [77], Colin F Robertson [78], Australian Asthma Genetics Consortium (AAGC) [79], Janina S Ried [20], Jin Li [21], Xian Bo Zuo [23], Xiao Dong Zheng [23], Xian Yong Yin [23], Liang Dan Sun [23], Maeve A McAleer [80, 81], Grainne M O'Regan [81], Caoimhe MR Fahy [82], Linda E Campbell [83], Milan Macek [84], Michael Kurek [85], Donglei Hu [26], Celeste Eng [26], Dirkje S Postma [29], Bjarke Feenstra [31], Frank Geller [31], Jouke Jan Hottenga [32], Christel M Middeldorp [32], Pirro Hysi [33], Veronique Bataille [33], Tim Spector [33], Carla MT Tiesler [3, 86], Elisabeth Thiering [3, 86], Badri Pahukasahasram [87], James J Yang [88], Medea Imboden [37, 38], Scott Huntsman [26], Natàlia Vilor-Tejedor [24, 40, 41], Caroline L Relton [1, 89], Ronny Myhre [90], Wenche Nystad [90], Adnan Custovic [7], Scott T Weiss [91], Deborah A Meyers [92], Cilla Söderhäll [93, 94], Erik Melén [46, 95], Carole Ober [45], Benjamin A Raby [91], Angela Simpson [7], Bo Jacobsson [43, 90], John W Holloway [42, 96], Hans Bisgaard [4], Jordi Sunyer [24, 40, 41, 97], Nicole M Probst-Hensch [37, 38], L Keoki Williams [87, 98], Keith M Godfrey [35, 99], Carol A Wang [34], Dorret I Boomsma [32, 100], Mads Melbye [31, 101, 102], Gerard H Koppelman [103], Deborah Jarvis [104, 105], WH Irwin McLean [83], Alan D Irvine [80, 81, 82], Xue Jun

Zhang [23], Hakon Hakonarson [21, 22], Christian Gieger [18, 19, 20], Esteban G Burchard [26, 106], Nicholas G Martin [17], Liesbeth Duijts [14, 15, 16,] Allan Linneberg [70, 101, 107], Marjo-Riitta Jarvelin [69, 108, 109, 110], Markus M Noethen [55, 56], Susanne Lau [25], Norbert Hübner [10], Young-Ae Lee [10, 11], Mayumi Tamari [48], David A Hinds [8], Daniel Glass [33], Sara J Brown [83, 111], Joachim Heinrich [3], David M Evans [1, 2, 50, 113], Stephan Weidinger [5, 113] for the EArly Genetics & Lifecourse Epidemiology (EAGLE) eczema consortium[114].

1 Medical Research Council (MRC) Integrative Epidemiology Unit, University of Bristol, Bristol, UK.

2 School of Social and Community Medicine, University of Bristol, Bristol, UK.

3 Institute of Epidemiology I, Helmholtz Zentrum München - German Research Center for Environmental Health, Neuherberg, Germany.

4 Copenhagen Prospective Studies on Asthma in Childhood (COPSAC), Herlev and Gentofte Hospital, University of Copenhagen, Copenhagen, Denmark.

5 Department of Dermatology, Allergology and Venereology, University Hospital Schleswig-Holstein, Campus Kiel, Kiel, Germany.

6 Population Health Research Institute, St George's, University of London, London, UK.

7 Centre for Respiratory Medicine and Allergy, Institute of Inflammation and Repair, Manchester Academic Health Science Centre, The University of Manchester and University Hospital of South Manchester National Health Service (NHS) Foundation Trust, Manchester, United Kingdom.

8 23andMe, Inc., Mountain View, CA, USA.

9 Laboratory for Statistical Analysis, Center for Integrative Medical Sciences, Institute of Physical and Chemical Research (RIKEN), Yokohama, Japan.

10 Max-Delbrück-Center (MDC) for Molecular Medicine, Berlin, Germany.

11 Clinic for Pediatric Allergy, Experimental and Clinical Research Center, Charité - Universitätsmedizin Berlin, Berlin, Germany.

12 Department of Epidemiology and Biostatistics, School of Public Health, Imperial College London, London, UK.
13 National Allergy Research Centre, Department of Dermatology and Allergology, Herlev and Gentofte Hospital, University of Copenhagen, Copenhagen, Denmark.
14 Department of Pediatrics, Erasmus MC, Rotterdam, the Netherlands.
15 Department of Epidemiology, Erasmus MC, Rotterdam, the Netherlands.
16 The Generation R Study Group, Erasmus MC, Rotterdam, the Netherlands.
17 QIMR Berghofer Medical Research Institute, Brisbane, Australia.
18 Research Unit of Molecular Epidemiology, Helmholtz Zentrum München, German Research Center for Environmental Health, Neuherberg, Germany.
19 Institute of Epidemiology II, Helmholtz Zentrum München - German Research Center for Environmental Health, Neuherberg, Germany.
20 Institute of Genetic Epidemiology, Helmholtz Zentrum München, German Research Center for Environmental Health, Neuherberg, Germany.
21 The Center for Applied Genomics, The Children's Hospital of Philadelphia, PA, USA.
22 Department of Pediatrics, The Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA, USA.
23 Institute of Dermatology, Anhui Medical University, Hefei, Anhui, China.
24 Centre for Research in Environmental Epidemiology (CREAL), Barcelona, Spain.
25 Pediatric Pneumology and Immunology, Charité - Universitätsmedizin Berlin, Berlin, Germany.
26 Department of Medicine, University of California, San Francisco, CA, USA.
27 Centro de Investigación Biomédica en Red (CIBER) de Enfermedades Respiratorias, Instituto de Salud Carlos III, Madrid, Spain.
28 Research Unit, Hospital Universitario Nuestra Señora de Candelaria, Santa Cruz de Tenerife, Spain.

753 29 University of Groningen, University Medical Center Groningen, Department of Pulmonology,
754 Groningen Research Institute for Asthma and COPD (GRIAC), Groningen, the Netherlands.
755 30 University of Groningen, University Medical Center Groningen, Department of Genetics,
756 Groningen Research Institute for Asthma and COPD (GRIAC), Groningen, the Netherlands.
757 31 Department of Epidemiology Research, Statens Serum Institut, Copenhagen, Denmark.
758 32 Dept Biological Psychology, Netherlands Twin Register, VU University, Amsterdam, the
759 Netherlands.
760 33 KCL Department of Twin Research and Genetic Epidemiology, King's College London,
761 London, UK.
762 34 School of Women's and Infants' Health, The University of Western Australia (UWA), Perth,
763 Australia.
764 35 Medical Research Council (MRC) Lifecourse Epidemiology Unit, University of Southampton,
765 Southampton, UK.
766 36 Department of Public Health Sciences, Henry Ford Health System, Detroit, MI, USA.
767 37 Department of Epidemiology and Public Health, Swiss Tropical and Public Health Institute,
768 Basel, Switzerland.
769 38 University of Basel, Basel, Switzerland.
770 39 Centre for Genomic Regulation (CRG), Barcelona, Spain.
771 40 Pompeu Fabra University (UPF), Barcelona, Spain.
772 41 Centro de Investigación Biomédica en Red de Epidemiología y Salud Pública (CIBERESP),
773 Barcelona, Spain.
774 42 Human Development and Health, Faculty of Medicine, University of Southampton,
775 Southampton, UK.
776 43 Department of Obstetrics and Gynecology, Institute of Clinical Sciences, Sahlgrenska Academy,
777 Sahlgrenska University Hosptial, Gothenburg, Sweden.

44 Department of Genetics and Genomic Sciences, Icahn School of Medicine at Mount Sinai, New York, NY, USA.
45 Department of Human Genetics, University of Chicago, Chicago, IL, USA.
46 Institute of Environmental Medicine, Karolinska Institutet, Stockholm, Sweden.
47 Wellcome Trust Centre for Human Genetics, University of Oxford, Oxford, UK.
48 Laboratory for Respiratory and Allergic Diseases, Center for Integrative Medical Sciences, Institute of Physical and Chemical Research (RIKEN), Yokohama, Japan.
49 Laboratory for Genotyping Development, Center for Integrative Medical Sciences, Institute of Physical and Chemical Research (RIKEN), Yokohama, Japan.
50 University of Queensland Diamantina Institute, Translational Research Institute, University of Queensland, Brisbane, Australia.
51 Klinik für Kinder- und Jugendmedizin, Technical University Dresden, Dresden, Germany.
52 Clinic and Polyclinic of Dermatology, University Medicine Greifswald, Greifswald, Germany.
53 Department of Functional Genomics, Interfaculty Institute for Genetics and Functional Genomics, University Medicine and Ernst-Moritz-Arndt-University Greifswald, Greifswald, Germany.
54 Institute for Community Medicine, Study of Health in Pomerania/KEF, University Medicine Greifswald, Greifswald, Germany.
55 Institute of Human Genetics, University of Bonn, Bonn, Germany.
56 Department of Genomics, Life & Brain Center, University of Bonn, Bonn, Germany.
57 Division of Medical Genetics, University Hospital Basel, Basel, Switzerland.
58 Department of Biomedicine, University of Basel, Basel, Switzerland.
59 Institute of Neuroscience and Medicine (INM-1), Structural and Functional Organisation of the Brain, Genomic Imaging, Research Centre Jülich, Jülich, Germany.
60 Institute of Social Medicine, Epidemiology and Health Economics, Charité - Universitätsmedizin Berlin, Berlin, Germany.

804 61 Institute of Clinical Epidemiology and Biometry, University of Würzburg, Würzburg, Germany.
805 62 Deutsches Forschungszentrum für Herz-Kreislauferkrankungen (DZHK) (German Research
806 Centre for Cardiovascular Research), Munich Heart Alliance, Munich, Germany.
807 63 Institute of Clinical Molecular Biology, Christian-Albrechts-University of Kiel, Kiel, Germany.
808 64 Institute of Epidemiology, Christian-Albrechts University Kiel, Kiel, Germany.
809 65 Department of Dermatology and Allergy, University of Bonn Medical Center, Bonn, Germany.
810 66 Unit of Living Environment and Health, National Institute for Health and Welfare, Kuopio,
811 Finland.
812 67 Department of Public Health, University of Helsinki, Helsinki, Finland.
813 68 Center for Life-course and Systems Epidemiology, Faculty of Medicine, University of Oulu,
814 Finland.
815 69 Biocenter Oulu, University of Oulu, Finland.
816 70 Research Centre for Prevention and Health, Capital Region of Denmark, Copenhagen, Denmark.
817 71 The Novo Nordisk Foundation Center for Basic Metabolic Research, Faculty of Health and
818 Medical Sciences, University of Copenhagen, Copenhagen, Denmark.
819 72 Department of Internal Medicine, Erasmus MC, Rotterdam, the Netherlands.
820 73 Department of Dermatology, Erasmus MC, Rotterdam, the Netherlands.
821 74 Woolcock Institute of Medical Research, University of Sydney, Sydney, Australia.
822 75 Lung Institute of Western Australia, QE II Medical Centre Nedlands , Western Australia,
823 Australia.
824 76 School of Medicine and Pharmacology, University of Western Australia, Perth, Australia.
825 77 Melbourne School of Population and Global Health, University of Melbourne, Melbourne,
826 Australia.
827 78 Murdoch Children's Research Institute, Melbourne, Australia.
828 79 A full list of consortium members is provided in Supplementary Note 1, page 4.
829 80 National Children's Research Centre, Crumlin, Dublin, Ireland.

81 Our Lady's Children's Hospital, Crumlin, Dublin, Ireland.
82 Clinical Medicine, Trinity College Dublin, Dublin, Ireland.
83 Centre for Dermatology and Genetic Medicine, University of Dundee, Dundee, UK.
84 Department of Biology and Medical Genetics, University Hospital Motol and 2nd Faculty of Medicine of Charles University, Prague, Czech Republic.
85 Department of Clinical Allergology, Pomeranian, Pomeranian Medical University, Szczecin, Poland.
86 Ludwig-Maximilians-University of Munich, Dr. von Hauner Children's Hospital, Division of Metabolic Diseases and Nutritional Medicine, Munich, Germany.
87 Center for Health Policy and Health Services Research, Henry Ford Health System, Detroit, MI, USA.
88 School of Nursing, University of Michigan, Ann Arbor, MI, USA.
89 Institute of Genetic Medicine, Newcastle University, Newcastle upon Tyne, UK.
90 Division of Epidemiology, Norwegian Institute of Public Health, Oslo, Norway.
91 Channing Division of Network Medicine, Brigham & Women's Hospital and Harvard Medical School, Boston, MA, USA.
92 Center for Genomics and Personalized Medicine Research, Wake Forest School of Medicine, Winston-Salem, NC, USA.
93 Department of Biosciences and Nutrition, Karolinska Institutet, Stockholm, Sweden.
94 Center for Innovative Medicine (CIMED), Karolinska Institutet, Stockholm, Sweden.
95 Sachs' Children's Hospital, Stockholm, Sweden.
96 Clinical and Experimental Sciences, Faculty of Medicine, University of Southampton, Southampton, UK.
97 Hospital del Mar Medical Research Institute (IMIM), Barcelona, Spain.
98 Department of Internal Medicine, Henry Ford Health System, Detroit, MI, USA.

855 99 National Institute for Health Research (NIHR) Southampton Biomedical Research Centre,
856 University of Southampton and University Hospital Southampton National Health Service (NHS)
857 Foundation Trust, Southampton, UK.
858 100 Institute for Health and Care Research (EMGO), VU University, Amsterdam, the Netherlands.
859 101 Department of Clinical Medicine, Faculty of Health and Medical Sciences, University of
860 Copenhagen, Copenhagen, Denmark.
861 102 Department of Medicine, Stanford School of Medicine, Stanford, California, USA.
862 103 University of Groningen, University Medical Center Groningen, Beatrix Children's Hospital,
863 Department of Pediatric Pulmonology and Pediatric Allergology, Groningen Research Institute for
864 Asthma and COPD (GRIAC), Groningen, the Netherlands.
865 104 Respiratory Epidemiology, Occupational Medicine and Public Health; National Heart and Lung
866 Institute; Imperial College; London, UK.
867 105 Medical Research Council-Public Health England Centre for Environment and Health, School
868 of Public Health, Imperial College London, London, UK.
869 106 Department of Bioengineering and Therapeutic Sciences, University of California, San
870 Francisco, CA, USA.
871 107 Department of Clinical Experimental Research, Rigshospitalet, Glostrup, Denmark.
872 108 Department of Epidemiology and Biostatistics, Medical Research Council (MRC) Health
873 Protection Agency (HPE) Centre for Environment and Health, School of Public Health, Imperial
874 College London, London, UK.
875 109 Center for Life Course Epidemiology, Faculty of Medicine, University of Oulu, Oulu, Finland.
876 110 Unit of Primary Care, Oulu University Hospital,
877 Oulu, Finland.
878 111 Department of Dermatology, Ninewells Hospital and Medical School, Dundee, UK.
879 112 These authors contributed equally to this work.
880 113 These authors jointly directed this work.

114 All authors.

L.Paternoster is supported by an MRC Population Health Scientist Fellowship (MR/J012165/1).

Endometrial Cancer Association Consortium (ECAC)

Amanda B Spurdle[1], Tracy A O'Mara[1], Jodie N Painter[1], The Australian National Endometrial Cancer Study Group (ANECS)[1], Mark McEvoy[2], John Attia[2, 3], Elizabeth G Holliday[2, 3], Rodney J Scott[3-6], Deborah J Thompson[7], Douglas F Easton[7, 8], Alison M Dunning[8], Paul D P Pharoah[8], Mitul Shah[8], Shahana Ahmed[8], Catherine S Healey[8], Ian Tomlinson[9], Timothy HT Cheng[9], Lynn Martin[9], Maggie Gorman[9], Shirley Hodgson[10], National Study of Endometrial Cancer Genetics Group (NSECG)[9], Peter A Fasching[11, 12], Alexander Hein[12], Matthias W Beckmann[12], Arif B Ekici[13], Matthias Rübner[12], Per Hall[14], Kamila Czene[14], Jingmei Li[14], Hatef Darabi[14], Thilo Dörk[15], Ingo Runnebaum[16], Matthias Dürst[16], Peter Hillemanns[17], Diether Lambrechts[18, 19], Frederic Amant[20], Stefanie Schrauwen[20], Jeroen Depreeuw[18-20], Ellen L Goode[21], Sean C Dowdy[22], Stacey J Winham[21], Brooke L Fridley[23], Helga B Salvesen[24, 25], Henrica MJ Werner[24, 25], Tormund S Njølstad[24, 25], Jone Trovik[24, 25], Katie Ashton[3, 5, 6], Tony Proietto[26], Geoffrey Otton[26], Emma Tham[27], Miriam Mints[28], RENDOCAS[27]

[1] Department of Genetics and Computational Biology, QIMR Berghofer Medical Research Institute, Brisbane, QLD, 4006, Australia. [2] Centre for Clinical Epidemiology and Biostatistics, School of Medicine and Public Health, University of Newcastle, NSW, 2305, Australia. [3] Hunter Medical Research Institute, John Hunter Hospital, Newcastle, NSW, 2305, Australia. [4] Hunter Area Pathology Service, John Hunter Hospital, Newcastle, NSW, 2305, Australia. [5] Centre for Information Based Medicine, University of Newcastle, NSW, 2308, Australia. [6] School of Biomedical Sciences and Pharmacy, University of Newcastle, Newcastle, NSW, 2308, Australia. [7] Centre for Cancer Genetic Epidemiology, Department of Public Health and Primary Care, University of Cambridge, Cambridge, CB1 8RN, UK. [8] Centre for Cancer Genetic Epidemiology,

907 Department of Oncology, University of Cambridge, Cambridge, CB1 8RN, UK. [9] Wellcome Trust
908 Centre for Human Genetics, University of Oxford, Oxford, OX3 7BN, UK. [10] Department of
909 Clinical Genetics, St George's, University of London, London, SW17 0RE, UK. [11] University of
910 California at Los Angeles, Department of Medicine, Division of Hematology/Oncology, David
911 Geffen School of Medicine, Los Angeles, CA, 90095, USA.
912 [12] Department of Gynecology and Obstetrics, University Hospital Erlangen, Friedrich-Alexander
913 University Erlangen-Nuremberg, Erlangen, 91054, Germany. [13] Institute of Human Genetics,
914 University Hospital Erlangen, Friedrich-Alexander-University Erlangen-Nuremberg, Erlangen,
915 91054, Germany. [14] Department of Medical Epidemiology and Biostatistics, Karolinska Institutet,
916 Stockholm, SE-171 77, Sweden. [15] Hannover Medical School, Gynaecology Research Unit,
917 Hannover, 30625, Germany. [16] Department of Gynaecology, Jena University Hospital - Friedrich
918 Schiller University, Jena, 07743, Germany. [17] Hannover Medical School, Clinics of Gynaecology
919 and Obstetrics, Hannover, 30625, Germany. [18] Vesalius Research Center, Leuven, 3000, Belgium.
920 [19] Laboratory for Translational Genetics, Department of Oncology, University Hospitals Leuven,
921 Leuven, 3000, Belgium. [20] Department of Obstetrics and Gynecology, Division of Gynecologic
922 Oncology, University Hospitals, KU Leuven - University of Leuven, 3000, Belgium. [21] Department
923 of Health Sciences Research, Mayo Clinic, Rochester, MN, 55905, USA. [22] Department of
924 Obstetrics and Gynecology, Division of Gynecologic Oncology, Mayo Clinic, Rochester, MN,
925 55905, USA. [23] Department of Biostatistics, University of Kansas Medical Center, Kansas City,
926 KS, 66160, USA. [24] Centre for Cancerbiomarkers, Department of Clinical Science, The University
927 of Bergen, 5020, Norway.
928 [25] Department of Obstetrics and Gynecology, Haukeland University Hospital, Bergen, 5021,
929 Norway.
930 [26] School of Medicine and Public Health, University of Newcastle, Newcastle, NSW, 2308,
931 Australia.

[27] Department of Molecular Medicine and Surgery, Karolinska Institutet, Stockholm, SE-171 77, Sweden. [28] Department of Women's and Children's Health, Karolinska Institutet, Karolinska University Hospital, Stockholm, SE-171 77, Sweden.

Glioma GWAS

Work at University of California, San Francisco was supported by the National Institutes of Health (grant numbers R01CA52689, P50CA097257, R01CA126831, R01CA139020 and R25CA112355), as well as the National Brain Tumor Foundation, the Stanley D. Lewis and Virginia S. Lewis Endowed Chair in Brain Tumor Research, the Robert Magnin Newman Endowed Chair in Neuro-oncology, and by donations from families and friends of John Berardi, Helen Glaser, Elvera Olsen, Raymond E. Cooper, and William Martinusen.

The collection of cancer incidence data used in this study was supported by the California Department of Public Health as part of the statewide cancer reporting program mandated by California Health and Safety Code Section 103885; the National Cancer Institute's Surveillance, Epidemiology and End Results Program under contract HHSN261201000140C awarded to the Cancer Prevention Institute of California, contract HHSN261201000035C awarded to the University of Southern California, and contract HHSN261201000034C awarded to the Public Health Institute; and the Centers for Disease Control and Prevention's National Program of Cancer Registries, under agreement # U58DP003862-01 awarded to the California Department of Public Health. The ideas and opinions expressed herein are those of the author(s) and endorsement by the State of California Department of Public Health, the National Cancer Institute, and the Centers for Disease Control and Prevention or their Contractors and Subcontractors is not intended nor should be inferred.

957

958 The results published here are in whole or part based upon data generated by The Cancer Genome
959 Atlas managed by the NCI and NHGRI. Information about TCGA can be found
960 athttp://cancergenome.nih.gov

961

962 **Endometriosis GWA meta-analysis**

963 Dale R Nyholt[1,16], Siew-Kee Low[2,16], Carl A Anderson[3], Jodie N Painter[1], Satoko Uno[2,4], Andrew P
964 Morris[5], Stuart MacGregor[1], Scott D Gordon[1], Anjali K Henders[1], Nicholas G Martin[1], John
965 Attia[6,7], Elizabeth G Holliday[6,7], Mark McEvoy[6,8,9], Rodney J Scott[7,10,11], Stephen H Kennedy[12],
966 Susan A Treloar[13], Stacey A Missmer[14], Sosuke Adachi[15], Kenichi Tanaka[15], Yusuke Nakamura[2],
967 Krina T Zondervan[5,12,17], Hitoshi Zembutsu[2,17] & Grant W Montgomery[1,17]

968

969 [1]Queensland Institute of Medical Research, Brisbane, Queensland, Australia. [2]Laboratory of
970 Molecular Medicine, Human Genome Center, Institute of Medical Science, University of Tokyo,
971 Tokyo, Japan.
972 [3]Wellcome Trust Sanger Institute, Wellcome Trust Genome Campus, Hinxton, Cambridge, UK.
973 [4]First Department of Surgery, Sapporo Medical University, School of Medicine, Sapporo, Japan.
974 [5]Genetic and Genomic Epidemiology Unit, Wellcome Trust Centre for Human Genetics, University
975 of Oxford, Oxford, UK.
976 [6]Centre for Clinical Epidemiology and Biostatistics, School of Medicine and Public Health,
977 University of Newcastle, Newcastle, New South Wales, Australia.
978 [7]Centre for Bioinformatics, Biomarker Discovery and Information-Based Medicine, Hunter Medical
979 Research Institute, Newcastle, New South Wales, Australia.
980 [8]School of Medicine and Public Health, University of Newcastle, Newcastle, New South Wales,
981 Australia.

[9]Public Health Research Program, Hunter Medical Research Institute, Newcastle, New South Wales, Australia.

[10]School of Biomedical Sciences and Pharmacy, University of Newcastle, Newcastle, New South Wales, Australia.

[11]Division of Genetics, Hunter Area Pathology Service, Newcastle, New South Wales, Australia.

[12]Nuffield Department of Obstetrics and Gynaecology, University of Oxford, John Radcliffe Hospital, Oxford, UK.

[13]Centre for Military and Veterans' Health, University of Queensland, Mayne Medical School, Brisbane, Queensland, Australia.

[14]Department of Obstetrics, Gynecology and Reproductive Biology, Brigham and Women's Hospital and Harvard Medical School, Boston, Massachusetts, USA. [15]Department of Obstetrics and Gynecology, Niigata University Graduate School of Medical and Dental Sciences, Niigata, Japan.

[16]These authors contributed equally to this work. [17]These authors jointly directed this work.

Funding

QIMR: The QIMR study was supported by grants from the National Health and Medical Research Council (NHMRC) of Australia (241944, 339462, 389927, 389875, 389891, 389892, 389938, 443036, 442915, 442981, 496610, 496739, 552485 and 552498), the Cooperative Research Centre for Discovery of Genes for Common Human Diseases (CRC), Cerylid Biosciences (Melbourne) and donations from N. Hawkins and S. Hawkins. D.R.N. was supported by the NHMRC Fellowship (339462 and 613674) and Australian Research Council (ARC) Future Fellowship (FT0991022) schemes. S.M. was supported by NHMRC Career Development Awards (496674 and 613705). E.G.H. (631096) and G.W.M. (339446 and 619667) were supported by the NHMRC Fellowship scheme. The HCS was funded by the University of Newcastle, the Gladys M Brawn Fellowship scheme and the Vincent Fairfax Family Foundation in Australia. ***OX:*** The work presented here was

supported by a grant from the Wellcome Trust (WT084766/Z/08/Z) and makes use of Wellcome Trust Case Control Consortium 2 (WTCCC2) control data generated by the WTCCC. A full list of the investigators who contributed to the generation of these data is available at the Wellcome Trust website (http://www.wtccc.org.uk/). Funding for the WTCCC project was provided by the Wellcome Trust under awards 076113 and 085475. C.A.A. was supported by a grant from the Wellcome Trust (098051). A.P.M. was supported by a Wellcome Trust Senior Research Fellowship. S.H.K. is supported by the Oxford Partnership Comprehensive Biomedical Research Centre, with funding from the Department of Health National Institute for Health Research (NIHR) Biomedical Research Centres funding scheme. K.T.Z. is supported by a Wellcome Trust Research Career Development Fellowship (WT085235/Z/08/Z). ***BBJ:*** We thank the members of the Rotary Club of Osaka-Midosuji District 2660 Rotary International in Japan for supporting our study. This work was conducted as part of the BioBank Japan Project that was supported by the Ministry of Education, Culture, Sports, Science and Technology of the Japanese government.

European Periodontitis Genetics Group (EPG)

The GWAS of aggressive periodontitis (AgP) was supported by a research grant of the Deutsche Forschungsgemeinschaft DFG (GZ: SCHA 1582/3-1). The cohort case description has been previously reported in Schaefer A.S. *et al*. Genetic evidence for PLASMINOGEN as a shared genetic risk factor of coronary artery disease and periodontitis. *Circ Cardiovasc Genet* **8**, 159-67 (2015). The investigators who contributed to the generation of this case sample are: Henrik Dommisch[1], Christian Graetz[2], Inga Harks[3], Yvonne Jockel-Schneider[4], Jörg Eberhardt[5], Joerg Meyle[6], Peter Eickholz[7], Mathias Folwaczny[8], Barbara Noack[9], Wolfgang Lieb[10], Christof Doerfer[2], Corinna Bruckmann[11], Søren Jepsen[12]

Author affiliations

[1]Charité – University Medicine Berlin, Institute of Dental, Oral and Maxillary Medicine, Department of Periodontology, Berlin, Germany

[2]Department of Operative Dentistry and Periodontology, University Medical Center Schleswig-Holstein, Kiel, Germany

[3]Center of Periodontology, Operative and Preventive Dentistry, University Medical Center Münster, Germany

[4]Department of Periodontology, Clinic of Preventive Dentistry and Periodontology, University Medical Center of the Julius-Maximilians-University, Würzburg, Germany

[5]Department of Conservative Dentistry, Periodontology and Preventive Dentistry, Hannover Medical School, Hannover, Germany

[6]Department of Periodontology, University Medical Center Giessen and Marburg, Germany

[7]Department of Periodontology, Centre for Dental, Oral, and Maxillofacial Medicine (Carolinum), Johann Wolfgang Goethe-University, Frankfurt am Main, Germany

[8]Department of Preventive Dentistry and Periodontology, University of Munic, Germany

[9]Center of Periodontology, Operative and Preventive Dentistry, Clinic of Preventive Dentistry, University Medical Center Carl Gustav Carus der Technischen Universität Dresden, Germany

[10]Biobank popgen, University Medical Center Schleswig-Holstein, Campus Kiel, Germany

[11]Department of Conservative Dentistry and Periodontology, Clinic of Dentistry, Bernhard Gottlieb University, Vienna, Austria

[12]Department of Periodontology, Operative and Preventive Dentistry, University of Bonn, Bonn, Germany

Genotyping of the AgP cases was performed on an IScan system with HumanOmni BeadChips (Illumina) at the Institute of Clinical Molecular Biology, Christian-Albrechts-University Kiel, Germany. We specially thank Andre Franke and Stefan Schreiber.

The aggressive periodontitis control sample consists of three independent studies:

1. The Heinz-Nixdorff-Recall (HNR) was described in Schmermund, A., *et al.* Assessment of clinically silent atherosclerotic disease and established and novel risk factors for predicting

myocardial infarction and cardiac death in healthy middle-aged subjects: rationale and design of the Heinz Nixdorf RECALL Study. Risk Factors, Evaluation of Coronary Calcium and Lifestyle. *Am Heart J* **144**, 212-18 (2002). The HNR study was supported by the Heinz Nixdorf Foundation (Germany). Additionally, the study was funded by the German Ministry of Education and Science and the German Research Council (DFG; Project SI 236/8-1, SI236/9-1, ER 155/6-1). The genotyping of the Illumina HumanOmni-1 Quad BeadChips of the HNR subjects was financed by the German Centre for Neurodegenerative Disorders (DZNE), Bonn. We are extremely grateful to all investigators who contributed to the generation of this dataset.

The HNR study is represented by Per Hoffmann[1,2] and Bastian Krone[3]

Authors Affiliations

[1]Institute of Human Genetics, University of Bonn, Germany

[2]Human Genomics Research Group, Department of Biomedicine, University Hospital of Basel, Switzerland

[3]Bastian Krone, Institute of Medical Informatics, Biometry and Epidemiology, University Clinic Essen, Germany.

2. The Dortmund Health Study (DOGS) is described in Berger, K. *et. al*. DHS: The Dortmund health study. *Bundesgesundheitsblatt, Gesundheitsforschung, Gesundheitsschutz* **55**, 816-21 (2012). DOGS is supported by the German Migraine & Headache Society (DMKG) and by unrestricted grants of equal share from Almirall, Astra Zeneca, Berlin Chemie, Boehringer, Boots Health Care, Glaxo-Smith-Kline, Janssen Cilag, McNeil Pharma, MSD Sharp & Dohme and Pfizer to the University of Muenster (collection of sociodemographic and clinical data). Blood collection in the Dortmund Health Study was done through funds from the Institute of Epidemiology and Social Medicine University of Muenster.

DOGS is represented by Klaus Berger[1] and Jürgen Wellmann[1]

Authors Affiliations

[1]Institute of Epidemiology and Social Medicine, University Münster, Germany.

3. The FOCUS (Food chain plus) control sample is described in Muller, N., *et al.* IL-6 blockade by monoclonal antibodies inhibits apolipoprotein (a) expression and lipoprotein (a) synthesis in humans. *J Lipid Res* **56**, 1034-42 (2015). FOCUS was supported by the Federal Ministry of Education and Research BMBF (FKZ 0315540A). FOCUS is represented by Matthias Laudes[1]

[1]Clinic of Internal Medicine I, University Medical Center Schleswig-Holstein, Kiel, Germany

The International Genomics of Alzheimer's Project (IGAP)

We thank the International Genomics of Alzheimer's Project (IGAP) for providing summary results data for these analyses. The investigators within IGAP contributed to the design and implementation of IGAP and/or provided data but did not participate in analysis or writing of this report. IGAP was made possible by the generous participation of the control subjects, the patients, and their families. The i–Select chips was funded by the French National Foundation on Alzheimer's disease and related disorders. EADI was supported by the LABEX (laboratory of excellence program investment for the future) DISTALZ grant, Inserm, Institut Pasteur de Lille, Université de Lille 2 and the Lille University Hospital. GERAD was supported by the Medical Research Council (Grant n° 503480), Alzheimer's Research UK (Grant n° 503176), the Wellcome Trust (Grant n° 082604/2/07/Z) and German Federal Ministry of Education and Research (BMBF): Competence Network Dementia (CND) grant n° 01GI0102, 01GI0711, 01GI0420. CHARGE was partly supported by the NIH/NIA grant R01 AG033193 and the NIA AG081220 and AGES contract N01–AG–12100, the NHLBI grant R01 HL105756, the Icelandic Heart Association, and the Erasmus Medical Center and Erasmus University. ADGC was supported by the NIH/NIA grants: U01 AG032984, U24 AG021886, U01 AG016976, and the Alzheimer's Association grant ADGC–10–196728.

Material and methods

International Genomics of Alzheimer's Project (IGAP) is a large two-stage study based upon genome-wide association studies (GWAS) on individuals of European ancestry. In stage 1, IGAP used genotyped and imputed data on 7,055,881 single nucleotide polymorphisms (SNPs) to meta-analyse four previously-published GWAS datasets consisting of 17,008 Alzheimer's disease cases and 37,154 controls (The European Alzheimer's disease Initiative – EADI the Alzheimer Disease Genetics Consortium – ADGC The Cohorts for Heart and Aging Research in Genomic Epidemiology consortium – CHARGE The Genetic and Environmental Risk in AD consortium – GERAD). In stage 2, 11,632 SNPs were genotyped and tested for association in an independent set of 8,572 Alzheimer's disease cases and 11,312 controls. Finally, a meta-analysis was performed combining results from stages 1 & 2.

The Japanese Collaboration Team for GWAS of Panic Disorder

Tsukasa Sasaki[1], Yoshiya Kawamura[2], Takeshi Otowa[3,4,] Mamoru Tochigi[5], Fumichika Nishimura[4], Hisashi Tanii, Katsushi Tokunaga[7], Hisanobu Kaiya[8], Yuji Okazaki[9]

1 Department of Physical and Health Education, Graduate School of Education, The University of Tokyo, Japan

2 Department of Psychiatry, Shonan Kamakura General Hospital, Kanagawa, Japan

3 Major of Professional Clinical Psychology, Graduate School of Clinical Psychology, Teikyo Heisei University, Tokyo, Japan

4 Department of Neuropsychiatry, Graduate School of Medicine, The University of Tokyo, Japan

5 Department of Psychiatry, Teikyo University School of Medicine, Tokyo, Japan

6 Department of Psychiatry, Graduate School of Medicine, Mie University,Tsu, Japan

7 Department of Human Genetics, Graduate School of Medicine, The University of Tokyo, Japan

8 Anxiety Disorder Research Center, Warakukai Medical Cooporation, Tokyo, Japan

9 Metropolitan Matsuzawa Hospital, Tokyo, Japan

Meta-Analyses of Glucose and Insulin-related traits Consortium (MAGIC)

Data on glycaemic traits have been contributed by MAGIC investigators and have been downloaded from www.magicinvestigators.org. The investigators within MAGIC did not participate in the analysis, writing or interpretation of this paper.

Melanoma meta-analysis consortium (MC)

Matthew H. Law[1], D. Timothy Bishop[2], Jeffrey E. Lee[3], Myriam Brossard[4,5,66], Nicholas G. Martin[6], Eric K. Moses[7], Fengju Song[8], Jennifer H. Barrett[2], Rajiv Kumar[9], Douglas F. Easton[10], Paul D. P. Pharoah[11], Anthony J. Swerdlow[12,13], Katerina P. Kypreou[14], , Lisa Bowdler[42], Leanne Wallace[42], Anjali Henders[42], John C. Taylor[2], Mark Harland[2], Juliette Randerson-Moor[2], Lars A. Akslen[15,16], Per A. Andresen[17], Marie-Françoise Avril[18], Esther Azizi[19,20], Giovanna Bianchi Scarrà[21,22], Kevin M. Brown[23], Tadeusz Dębniak[24], David L. Duffy[6], David E. Elder[25,] Shenying Fang[3], Eitan Friedman[20], Pilar Galan[26], Paola Ghiorzo[21,22], Elizabeth M. Gillanders[27], Alisa M. Goldstein[23], Nelleke A. Gruis[28], Johan Hansson[29], Per Helsing[30], Marko Hočevar[31], Veronica Höiom[29], Christian Ingvar[32], Peter A. Kanetsky[33], Wei V. Chen[34], GenoMEL Consortium[35], Essen-Heidelberg Investigators[35], The SDH Study Group[35], Q-MEGA and QTWIN Investigators[35], AMFS Investigators[35], ATHENS Melanoma Study Group[35],Maria Teresa Landi[23], Julie Lang[36], G. Mark Lathrop[37], Jan Lubiński[24], Rona M. Mackie[38,39], Graham J. Mann[40], Anders Molven[16,41] , Grant W. Montgomery[42], Srdjan Novaković[43], Håkan Olsson[44,45], Susana Puig[46,47] , Joan Anton Puig-Butille[46,47,], Abrar A. Qureshi[48], Graham L. Radford-Smith[49,50,51], Nienke van der Stoep[52], Remco van Doorn[28], David C. Whiteman[53], Jamie E. Craig[54], Dirk Schadendorf[55,56], Lisa A. Simms[49], Kathryn P. Burdon[57], Dale R. Nyholt[58,42], Karen A. Pooley[10], Nicholas Orr[59], Alexander J. Stratigos[14], Anne E. Cust[60], Sarah V. Ward[7], Nicholas K. Hayward[61], Jiali Han[62,63], Hans-Joachim Schulze[64], Alison M. Dunning[11], Julia A. Newton Bishop[2], Florence Demenais[,66] , Christopher I. Amos[65,66], Stuart MacGregor[1,67], Mark M. Iles[2,67]

[1] Statistical Genetics, QIMR Berghofer Medical Research Institute, Brisbane, Australia

[2] Section of Epidemiology and Biostatistics, Leeds Institute of Cancer and Pathology, University of Leeds, Leeds, UK

[3] Department of Surgical Oncology, The University of Texas MD Anderson Cancer Center, Houston, Texas, USA

[4] Institut National de la Santé et de la Recherche Médicale (INSERM), UMR-946, Genetic Variation and Human Diseases Unit, Paris, France

[5] Institut Universitaire d'Hématologie, Université Paris Diderot, Sorbonne Paris Cité, Paris, France

[6] Genetic Epidemiology, QIMR Berghofer Medical Research Institute, Brisbane, Australia

[7] Centre for Genetic Origins of Health and Disease, Faculty of Medicine, Dentistry and Health Sciences, The University of Western Australia, Western Australia, Australia

[8] Departments of Epidemiology and Biostatistics, Key Laboratory of Cancer Prevention and Therapy, Tianjin, National Clinical Research Center of Cancer, Tianjin Medical University Cancer Institute and Hospital, Tianjin, P. R. China

[9] Division of Molecular Genetic Epidemiology, German Cancer Research Center, Im Neuenheimer Feld 580, Heidelberg Germany

[10] Centre for Cancer Genetic Epidemiology, Department of Public Health and Primary Care, University of Cambridge, Cambridge, UK

[11] Centre for Cancer Genetic Epidemiology, Department of Oncology, University of Cambridge, Cambridge, UK

[12] Division of Genetics and Epidemiology, The Institute of Cancer Research, London, UK

[13] Division of Breast Cancer Research, The Institute of Cancer Research, London, UK

[14] Department of Dermatology, University of Athens School of Medicine, Andreas Sygros Hospital, Athens, Greece

[15] Centre for Cancer Biomarkers CCBIO, Department of Clinical Medicine, University of Bergen, Bergen, Norway

[16] Department of Pathology, Haukeland University Hospital, Bergen, Norway

[17] Department of Pathology, Molecular Pathology, Oslo University Hospital, Rikshospitalet, Oslo, Norway

[18] Assistance Publique–Hôpitaux de Paris, Hôpital Cochin, Service de Dermatologie, Université Paris Descartes, Paris, France

[19] Department of Dermatology, Sheba Medical Center, Tel Hashomer, Sackler Faculty of Medicine, Tel Aviv, Israel

[20] Oncogenetics Unit, Sheba Medical Center, Tel Hashomer, Sackler Faculty of Medicine, Tel Aviv University, Tel Aviv, Israel

[21] Department of Internal Medicine and Medical Specialties, University of Genoa, Genoa, Italy

[22] Laboratory of Genetics of Rare Cancers, Istituto di ricovero e cura a carattere scientifico Azienda Ospedaliera Universitaria (IRCCS AOU) San Martino-IST Istituto Nazionale per la Ricerca sul Cancro, Genoa, Italy

[23] Division of Cancer Epidemiology and Genetics, National Cancer Institute, National Institutes of Health, Bethesda, Maryland, USA

[24] International Hereditary Cancer Center, Pomeranian Medical University, Szczecin, Poland

[25] Department of Pathology and Laboratory Medicine, Perelman School of Medicine at the University of Pennsylvania, Philadelphia, Pennsylvania, USA

[26] Université Paris 13, Equipe de Recherche en Epidémiologie Nutritionnelle (EREN), Centre de Recherche en Epidémiologie et Statistiques, Institut National de la Santé et de la Recherche Médicale (INSERM U1153), Institut National de la Recherche Agronomique (INRA U1125), Conservatoire National des Arts et Métiers, Communauté d'Université Sorbonne Paris Cité, F-93017 Bobigny, France

[27] Inherited Disease Research Branch, National Human Genome Research Institute, National Institutes of Health, Baltimore, Maryland, USA

[28] Department of Dermatology, Leiden University Medical Centre, Leiden, The Netherlands

[29] Department of Oncology-Pathology, Karolinska Institutet, Karolinska University Hospital, Stockholm, Sweden

[30] Department of Dermatology, Oslo University Hospital, Rikshospitalet, Oslo, Norway

[31] Department of Surgical Oncology, Institute of Oncology Ljubljana, Ljubljana, Slovenia

[32] Department of Surgery, Clinical Sciences, Lund University, Lund, Sweden

[33] Department of Cancer Epidemiology, H. Lee Moffitt Cancer Center and Research Institute, Tampa, Florida, USA

[34] Department of Genetics, The University of Texas MD Anderson Cancer Center, Houston, Texas, USA

[35] A full list of members and affiliations appears in the Supplementary Note.

[36] Department of Medical Genetics, University of Glasgow, Glasgow, UK

[37] McGill University and Genome Quebec Innovation Centre, Montreal, Canada

[38] Department of Public Health, University of Glasgow, Glasgow UK

[39] Department of Medical Genetics, University of Glasgow, Glasgow, UK

[40] Centre for Cancer Research, University of Sydney at Westmead, Millennium Institute for Medical Research and Melanoma Institute Australia, Sydney, Australia

[41] Gade Laboratory for Pathology, Department of Clinical Medicine, University of Bergen, Bergen, Norway

[42] Molecular Epidemiology, QIMR Berghofer Medical Research Institute, Brisbane, Australia

[43] Department of Molecular Diagnostics, Institute of Oncology Ljubljana, Ljubljana, Slovenia

[44] Department of Oncology/Pathology, Clinical Sciences, Lund University, Lund; Sweden

[45] Department of Cancer Epidemiology, Clinical Sciences, Lund University, Lund, Sweden

[46] Melanoma Unit, Dermatology Department & Biochemistry and Molecular Genetics Departments, Hospital Clinic, Institut de Investigacó Biomèdica August Pi Suñe, Universitat de Barcelona, Barcelona, Spain

[47] Centro de Investigación Biomédica en Red (CIBER) de Enfermedades Raras, Instituto de Salud Carlos III, Barcelona, Spain

[48] Department of Dermatology, The Warren Alpert Medical School of Brown University, Rhode Island, USA

[49] Inflammatory Bowel Diseases, QIMR Berghofer Medical Research Institute, Brisbane, Australia

[50] Department of Gastroenterology and Hepatology, Royal Brisbane & Women's Hospital, Brisbane, Australia

[51] University of Queensland School of Medicine, Herston Campus, Brisbane, Australia

[52] Department of Clinical Genetics, Center of Human and Clinical Genetics, Leiden University Medical Center, Leiden, The Netherlands

[53] Cancer Control Group, QIMR Berghofer Medical Research Institute, Brisbane, Australia

[54] Department of Ophthalmology, Flinders University, Adelaide, Australia

[55] Department of Dermatology, University Hospital Essen, Essen, Germany

[56] German Consortium Translational Cancer Research (DKTK), Heidelberg, Germany

[57] Menzies Institute for Medical Research, University of Tasmania, Hobart, Tasmania, Australia

[58] Institute of Health and Biomedical Innovation, Queensland University of Technology, Brisbane, Queensland, Australia

[59] Breakthrough Breast Cancer Research Centre, The Institute of Cancer Research, London, UK

[60] Cancer Epidemiology and Services Research, Sydney School of Public Health, The University of Sydney, Australia

[61] Oncogenomics, QIMR Berghofer Medical Research Institute, Brisbane, Australia

[62] Department of Epidemiology, Richard M. Fairbanks School of Public Health, Indiana University, Indianapolis, Indiana, USA

[63] Melvin and Bren Simon Cancer Center, Indiana University, Indianapolis, Indiana, USA

[64] Department of Dermatology, Fachklinik Hornheide, Institute for Tumors of the Skin at the University of Münster, Münster, Germany

[65] Department of Community and Family Medicine, Geisel School of Medicine, Dartmouth College, Hanover, New Hampshire, USA

[66] These authors contributed equally to this work

[67] These authors jointly supervised this work

The melanoma meta-analysis consortium was supported by CRUK Programme grants (C588/A19167 C8197/A10123, C8197/A10865), NIH grant (R01CA083115, RO1CA001833) NIH NCI (CA88363, CA83115, CA122838, CA87969, CA055075, CA100264, CA133996 and CA49449), the NHMRC (200071, 241944, 339462, 380385, 389927, 389875, 389891, 389892, 389938, 443036, 442915, 442981, 496610, 496675, 496739, 552485, 552498, 66946, 107359, 211172, 402761, 535074, 1023911).

Groups contributing to the melanoma meta-analysis

GenoMEL Consortium

Australian Melanoma Family Study: see below.

Barcelona: Paula Aguilera, Beatriz Alejo, Celia Badenas, Abel Caño, Cristina Carrera, Francisco Cuellar, Mireia Dominguez, Daniel Gabriel, Pol Gimenez-Xavier, Pablo Iglesias, Josep Malvehy, Rosa Marti-Laborda, Montse Mila, Zighe Ogbah, Miriam Potrony, Joan-Anton Puig Butille, Susana Puig, Gemma Tell and Other members of the Melanoma Unit: Llúcia Alós, Ana Arance, Pedro Arguís, Antonio Bennassar, Oscar Chirife, Carlos Conill, Ramon Rull, Marcelo Sánchez, Sergi Vidal-Sicart, Antonio Vilalta.

Brisbane: See Q-MEGA authorship below.

Emilia-Romagna: Maria Teresa Landi, Donato Calista, Giorgio Landi, Paola Minghetti, Fabio Arcangeli, Pier Alberto Bertazzi.

Genoa: Department of Internal Medicine and Medical Specialties, Laboratory of Genetics of Rare Cancers, University of Genoa/ IRCCS AOU San Martino-IST Istituto Nazioanle epr la Ricerca sul Cancro: Giovanna Bianchi Scarrà, Paola Ghiorzo, Lorenza Pastorino, William Bruno,Virginia Andreotti, Claudia Martinuzzi, Linda Battistuzzi, Paola Origone. Medical Oncology Unit, IRCCS AOU San Martino-IST Istituto Nazionale per la Ricerca sul Cancro, Genoa :Paola Queirolo.

Glasgow: Rona Mackie, Julie Lang.

Leeds: Julia A Newton Bishop, Paul Affleck, Jennifer H Barrett, D Timothy Bishop, Jane Harrison, Mark M Iles, Juliette Randerson-Moor, Mark Harland, John C Taylor, Linda Whittaker, Kairen Kukalizch, Susan Leake, Birute Karpavicius, Sue Haynes, Tricia Mack, May Chan, Yvonne Taylor, John Davies, Paul King.

Leiden: Department of Dermatology, Leiden University Medical Centre: Nelleke A Gruis, Frans A van Nieuwpoort, Coby Out, Clasine van der Drift, Wilma Bergman, Nicole Kukutsch, Jan Nico Bouwes Bavinck. *Department of Clinical Genetics, Centre of Human and Clinical Genetics, Leiden University Medical Centre:* Bert Bakker, Nienke van der Stoep, Jeanet ter Huurne. *Department of Dermatology, HAGA Hospital, The Hague:* Han van der Rhee. *Department of Dermatology, Reinier de Graaf Groep, Delft:* Marcel Bekkenk. *Department of Dermatology, Sint Franciscus Gasthuis, Rotterdam:* Dyon Snels, Marinus van Praag. *Department of Dermatology, Ghent University Hospital, Ghent, Belgium:* Lieve Brochez and colleagues. *Department of Dermatology, St. Radboud University Medical Centre, Nijmegen:* Rianne Gerritsen and colleagues. *Department of Dermatology, Rijnland Hospital, Leiderdorp:* Marianne Crijns and colleagues. *Dutch patient organisation, Stichting Melanoom, Purmerend. The Netherlands Foundation for the detection of Hereditary Tumors, Leiden:* Hans Vasen. *ServiceXS:* Wilbert van Workum, Bart Janssen, Marjolein Janssen and Suzanne Mulder

Lund: Lund Melanoma Study Group: Håkan Olsson, Christian Ingvar, Göran Jönsson, Åke Borg, Anna Måsbäck, Lotta Lundgren, Katja Baeckenhorn, Kari Nielsen, Anita Schmidt Casslén.

Norway: Oslo University Hospital: Per Helsing, Per Arne Andresen, Helge Rootwelt. *University of Bergen:* Lars A. Akslen, Anders Molven.

Paris (MELARISK study): Florence Demenais**,** Marie-Françoise Avril, Brigitte Bressac-de Paillerets, Eve Maubec, Myriam Brossard, Amaury Vaysse, Hamida Mohamdi, Patricia Jeannin, Valérie Chaudru, Nicolas Chateigner, Eve Corda, Fabienne Lesueur, Mahaut de Lichy and the French Family Study Group including the following oncogeneticists and Dermatologists: Pascale Andry-Benzaquen, Bertrand Bachollet, Frédéric Bérard, Pascaline Berthet, Françoise Boitier, Valérie Bonadona, Jean-Louis Bonafé, Jean-Marie Bonnetblanc, Frédéric Cambazard, Olivier Caron, Frédéric Caux, Jacqueline Chevrant-Breton, Agnès Chompret (deceased), Stéphane Dalle, Liliane Demange, Olivier Dereure, Martin-Xavier Doré, Marie-Sylvie Doutre, Catherine Dugast, Laurence Faivre, Florent Grange, Philippe Humbert, Pascal Joly, Delphine Kerob, Christine Lasset, Marie Thérèse Leccia, Gilbert Lenoir, Dominique Leroux, Julien Levang, Dan Lipsker, Sandrine Mansard, Ludovic Martin, Tanguy Martin-Denavit, Christine Mateus, Jean-Loïc Michel, Patrice Morel, Laurence Olivier-Faivre, Jean-Luc Perrot, Caroline Robert, Sandra Ronger-Savle, Bruno Sassolas, Pierre Souteyrand, Dominique Stoppa-Lyonnet, Luc Thomas, Pierre Vabres, Eva Wierzbicka.

Philadelphia: David Elder, Peter Kanetsky, Jillian Knorr, Michael Ming, Nandita Mitra, Althea Ruffin, Patricia Van Belle

Poland: Tadeusz Dębniak, Jan Lubiński, Aneta Mirecka, Sławomir Ertmański. *Slovenia:* Srdjan Novakovic, Marko Hocevar, Barbara Peric, Petra Cerkovnik. *Stockholm:* Veronica Höiom, Johan Hansson. *Sydney:* Graham J. Mann, Richard F. Kefford, Helen Schmid, Elizabeth A. Holland

Tel Aviv: Esther Azizi, Gilli Galore-Haskel, Eitan Friedman, Orna Baron-Epel, Alon Scope, Felix Pavlotsky, Emanuel Yakobson, Irit Cohen-Manheim, Yael Laitman, Roni Milgrom, Iris Shimoni, Evgeniya Kozlovaa.

Australian Melanoma Family Study investigators

Anne E. Cust[1], Helen Schmid[2], Elizabeth A. Holland[2], Joanne F. Aitken[3], Bruce K. Armstrong[1], Graham G. Giles[3,4], Richard F. Kefford[2], John L. Hopper[5] Mark A. Jenkins[5], Graham J. Mann[2]

1) Cancer Epidemiology and Services Research, Sydney School of Public Health, The University of Sydney, Australia

2) Centre for Cancer Research, University of Sydney at Westmead Millennium Institute for Medical Research and Melanoma Institute Australia, Sydney, Australia

3) Viertel Centre for Research in Cancer Control, Cancer Council Queensland, Spring Hill, Brisbane, Australia

4) Cancer Epidemiology Centre, Cancer Council Victoria, Melbourne, Australia

5) Centre for Molecular, Environmental, Genetic and Analytic (MEGA) Epidemiology, Melbourne School of Population Health, University of Melbourne, Melbourne, Australia

IBD investigators

Lisa Simms[1], Grant W. Montgomery[2], Peter Visscher[3]

1) Inflammatory Bowel Diseases Laboratory, QIMR Berghofer Medical Research Institute, Brisbane, Australia

2) Molecular Epidemiology, QIMR Berghofer Medical Research Institute, Brisbane, Australia

3) The Queensland Brain Institute, The University of Queensland, QBI Building, St Lucia, Queensland 4071, Australia.

Q-MEGA and QTWIN investigators

The Queensland study of Melanoma: Environmental and Genetic Associations (Q- MEGA) Principal Investigators are: Nicholas G. Martin[1], Grant W. Montgomery[1], David Duffy[1], David C. Whiteman[1], Matthew H. Law[1], Stuart MacGregor[1], Nicholas K. Hayward[1]. The Australian Cancer

Study (ACS) Principal Investigators are: David C. Whiteman[1], Penny Webb[1], Adele Green[1], Peter Parsons[1], David Purdie[1], Nicholas K.Hayward[1].

QTWIN: Zhen Zhen Zhao[1], Joanne F Aitken[2], Anjali K.Henders[1], Mitchell Stark[1], David L. Duffy[1], Jodie N. Painter[1]

1 QIMR Berghofer, Brisbane, QLD 4029, Australia

2 Viertel Centre for Research in Cancer Control, Cancer Council Queensland, Spring Hill, Brisbane, Australia

The SDH Study Group

Study of Digestive Health (SDH) Team

Chief Investigators

David Whiteman, Adele Green, Nicholas Hayward, Peter Parsons, Sandra Pavey, David Purdie, Penny Webb (Queensland Institute of Medical Research)

David Gotley, Mark Smithers (University of Queensland / Princess Alexandra Hospital)

Paul Drew, Glyn Jamieson (University of Adelaide)

Paul Drew, David Watson (Flinders University of South Australia)

Andrew Clouston (Mayne Pathology)

Research Staff

D Nancarrow

D Hussey

E Smith

G Mayne

Project Manager S O'Brien (QIMR)

Data Manager T Sadkowsky (QIMR)

Research Nurses

QLD-

A McMurtrie, L Terry, M Connard, L Jackman, S Perry, M Davis

SA-

D Roffe, M Martin, L Smith

Clinical Collaborators

QLD-

A Clouston (Envoi Pathology)

I Brown (S&N Pathology)

N Walker (QML Pathology)

SA-

Justin Bessell (Flinders Medical Centre)

William Tam (Royal Adelaide Hospital)

Andrew Ruskowicz (Institute of Medical and Veterinary Science)

Essen-Heidelberg investigators

Rajiv Kumar (r.kumar@dkfz.de)

Division of Molecular Genetic Epidemiology, German Cancer Research Center, Im Neuenheimer Feld 580, 69120 Heidelberg Germany

Dirk Schadendorf (Dirk.Schadendorf@uk-essen.de)

Department of Dermatology, University Hospital Essen, 45122 Essen, Germany

and

German Consortium Translational Cancer Research (DKTK), 69120 Heidelberg, Germany

Hans-Joachim Schulze(Schulze@fachklinik-hornheide.de)

Department of Dermatology, Fachklinik Hornheide, Institute for Tumors of the Skin at the University of Münster, Germany

Kari Hemminki (k.hemminki@dkfz.de)

Division of Molecular Genetic Epidemiology, German Cancer Research Center, Im Neuenheimer Feld 580, 69120 Heidelberg, Germany

Antje Sucker (antje.sucker@uk-essen.de)

Department of Dermatology, University Hospital Essen, 45122 Essen, Germany, German Consortium Translational Cancer Research, 69120 Heidelberg, Germany

Thomas Vogt (thomas.vogt@uks.eu)

University Hospital Saarland, Department of Dermatology, Venerology and Allergology, Building 18, Kirrberger Straße, D - 66424 Homburg/Saar, Germany

Johan Hansson (johan.hansson@ki.se)

Department of Oncology Pathology, Karolinska Institutet, Karolinska University Hospital, Solna S 171 76, Stockholm, Sweden.

Ralf Gutzmer (Gutzmer.Ralf@mh-hannover.de)

Department of Dermatology and Allergy, Skin Cancer Center Hannover, Hannover Medical School, Carl-Neuberg-Str. 1, 30625 Hannover, Germany

Helen Gogas (hgogas@hol.gr)

1st Department of Medicine, University of Athens Medical School, Laiko Hospital,

PO 14120, 11510, Athens, Greece

Dave Hoon (hoon@jwci.org)

John Wayne Cancer Institute, 2200 Santa Monica Blvd, Santa Monica, CA 90404, USA

Eduardo Nagore (eduardo_nagore@ono.com)

Department of Dermatology,Instituto Valenciano de Oncologia, Valencia 46009, Spain

John Kirkwood (kirkwoodjm@upmc.edu)

Dermatology & Translational Science, Melanoma and Skin Cancer Program, 5117 Centre Avenue, Suite 1.32, Pittsburgh, PA 15213, USA

Benjamin Weide (benjamin.weide@med.uni-tuebingen.de)

Department of Dermatology, University Medical Center, Liebermeisterstr. 25, 72076 Tübingen, Germany

Piotr Rutkowski (rutkowskip@coi.waw.pl)

Department of Soft Tissue/Bone Sarcoma and Melanoma, Maria Sklodowska-Curie Memorial Cancer Center and Institute of Oncology, Roentgena 5, 02-781 Warsaw, Poland

Selma Ugurel (Selma.Ugurel@uk-essen.de)

Department of Dermatology, University Hospital Essen, 45122 Essen, Germany, German Consortium Translational Cancer Research, 69120 Heidelberg, Germany

ATHENS Melanoma Study Group - investigators

Katerina Kypreou, Fani Karagianni, Kyriaki Antonopoulou, Dorothea Polydorou, Vasiliki Hasapi, Michaela Plaka, Nelli Gousetti, Othon Papadopoulos, Christina Antoniou, Alexander Stratigos (Department of Dermatology, Andreas Sygros Hospital, Athens, Greece)

Helen Gogas (Department of Internal Medicine, Laikon Hospital, University of Athens, Athens, Greece)

Vangelis Evangelou (Department of Epidemiology, University of Ioannina, Ioannina, Greece)

The Multi-Ethnic Study of Atherosclerosis (MESA)

The MESA and the MESA SHARe project are conducted and supported by the National Heart, Lung, and Blood Institute (NHLBI) in collaboration with MESA investigators. Support for MESA is provided by contracts N01-HC-95159, N01-HC-95160, N01-HC-95161, N01-HC-95162, N01-HC-95163, N01-HC-95164, N01-HC-95165, N01-HC-95166, N01-HC-95167, N01-HC-95168, N01-HC-95169, UL1-TR-001079, UL1-TR-000040, and DK063491. Funding for SHARe genotyping was provided by NHLBI Contract N02-HL-64278. Genotyping was performed at

Affymetrix (Santa Clara, California, USA) and the Broad Institute of Harvard and MIT (Boston, Massachusetts, USA) using the Affymetrix Genome-Wide Human SNP Array 6.0. Funding support for the Lung CT dataset was provided by grants R01-HL077612 and RC1-HL100543.NIH Intramural award ZIAEY00403 supported the collection of eye-related data in MESA.

The Nurses' Health Study (NHS) and the Health Professionals Follow-Up Study (HPFS)

We would like to thank the participants and staff of the Nurses' Health Study, the Health Professionals Follow-Up Study for their valuable contributions as well as the following state cancer registries for their help: AL, AZ, AR, CA, CO, CT, DE, FL, GA, ID, IL, IN, IA, KY, LA, ME, MD, MA, MI, NE, NH, NJ, NY, NC, ND, OH, OK, OR, PA, RI, SC, TN, TX, VA, WA, WY. The authors assume full responsibility for analyses and interpretation of these data. This work was supported by NIH R01 CA49449, P01 CA87969, UM1 CA186107, and UM1 CA167552.

GWAS of non-alcoholic fatty liver disease (hepatic steatosis)

Bratati Kahali and Elizabeth K Speliotes were supported by the Doris Duke Medical Foundation, NIH grant R01DK106621-01, the University of Michigan Internal Medicine Department, Division of Gastroenterology, the University of Michigan Biological Sciences Scholars Program and The Central Society for Clinical Research.

Pancreatic cancer case-control consortium (PanC4)

The Mayo Clinic Molecular Epidemiology of Pancreatic Cancer study was supported by the Mayo Clinic SPORE in Pancreatic Cancer (P50CA102701). The Yale University study was supported by grant number 5R01CA098870 from the NCI. The work at Johns Hopkins University was supported by NCI Grants P50CA62924 and R01CA97075 and the Lustgarten Foundation for Pancreatic Cancer Research. The Pancreas Tumor Registry at Memorial Sloan Kettering Cancer Center was supported by NIH P30CA008748 and the Goldstein Fund for Prevention, Control and Population

Research. The work at MD Anderson was supported by NIH Grant R01CA98380. The UCSF study was supported in part by NCI Grants CA59706, CA108370, CA109767, CA89726, and CA98889 and by the Rombauer Pancreatic Cancer Research Fund. The University of Toronto study was supported by NIH Grant R01CA97075, the Lustgarten Foundation for Pancreatic Cancer Research, and the Ontario Cancer Research Network.

Pancreatic Cancer Cohort Consortium (PanScan)

PanScan is the NCI cohort consortium genome-wide association study for pancreatic cancer. This research was supported by the Intramural Research Program of the National Institutes of Health, Division of Cancer Epidemiology and Genetics, National Cancer Institute, National Institutes of Health, Department of Health and Human Services.

The European Prospective Investigation into Cancer and Nutrition (EPIC) study

The coordination of EPIC is financially supported by the European Commission (DG-SANCO) andthe International Agency for Research on Cancer. The national cohorts are supported by the Health Research Fund (FIS) of the Spanish Ministry of Health,Regional Governments of Andalucía, Asturias, Basque Country, Murcia (no.6236), Navarra and the Catalan Institute of Oncology, La Caixa (BM 06-130),Red Temática de Investigación Cooperativa en Cáncer (RD12/0036/0018; RD06/0020/0091; Spain); Danish Cancer Society (Denmark); Ligue contre le Cancer, Institut Gustave Roussy, Mutuelle Générale de l'Education Nationale,Institut National de la Santé et de la Recherche Médicale (INSERM; France); Deutsche Krebshilfe, Deutsches Krebsforschungszentrum (DKFZ) and Federal Ministry of Education and Research (Germany); the Hellenic Health Foundation (Greece); Associazione Italiana per la Ricerca sul Cancro (AIRC) and National Research Council (Italy); Dutch Ministry of Public Health, Welfare, and Sports (VWS), Netherlands Cancer Registry (NKR), LK Research Funds, Dutch Prevention Funds, Dutch ZON

(Zorg Onderzoek Nederland), World Cancer Research Fund (WCRF), and Statistics Netherlands (The Netherlands); Nordic Center of Excellence in Food, Nutrition, and Health Helga (Norway); Swedish Cancer Society, Swedish Scientific Council and Regional Government of Skane and Vasterbotten (Sweden); Cancer Research UK (C570/A16491, R.C. Travis; 14136, K.T. Khaw) and Medical Research Council (G1000143, K.T. Khaw; United Kingdom).

The PRACTICAL Consortium

(http://practical.ccge.medschl.cam.ac.uk/)

Rosalind Eeles [1, 2], Doug Easton [3], Zsofia Kote-Jarai [1], Ali Amin Al Olama [3], Sara Benlloch [3], Kenneth Muir [4], Graham G. Giles [5, 6], Fredrik Wiklund [7], Henrik Gronberg [7], Christopher A. Haiman [8], Johanna Schleutker [9, 10], Maren Weischer [11], Ruth C. Travis [12], David Neal [13], Paul Pharoah [14], Kay-Tee Khaw [15], Janet L. Stanford [16, 17], William J. Blot [18], Stephen Thibodeau [19], Christiane Maier [20, 21], Adam S. Kibel [22, 23], Cezary Cybulski [24], Lisa Cannon-Albright [25], Hermann Brenner [26,27], Jong Park [28], Radka Kaneva [29], Jyotsna Batra [30], Manuel R. Teixeira [31], Hardev Pandha[32]

[1] The Institute of Cancer Research, London, SM2 5NG, UK, [2] Royal Marsden NHS Foundation Trust, London, SW3 6JJ, UK, [3] Centre for Cancer Genetic Epidemiology, Department of Public Health and Primary Care, University of Cambridge, Strangeways Research Laboratory, Worts Causeway, Cambridge, UK, [4] University of Warwick, Coventry, UK, [5] Cancer Epidemiology Centre, Cancer Council Victoria, 615 St Kilda Road, Melbourne Victoria, Australia, [6] Centre for Epidemiology and Biostatistics, Melbourne School of Population and Global Health, The University of Melbourne, Victoria, Australia, [7] Department of Medical Epidemiology and Biostatistics, Karolinska Institute, Stockholm, Sweden, [8] Department of Preventive Medicine, Keck School of Medicine, University of Southern California/Norris Comprehensive Cancer Center, Los Angeles, California, USA, [9] Department of Medical Biochemistry and Genetics, University of

Turku, Turku, Finland, [10] Institute of Biomedical Technology/BioMediTech, University of Tampere and FimLab Laboratories, Tampere, Finland, [11] Department of Clinical Biochemistry, Herlev Hospital, Copenhagen University Hospital, Herlev Ringvej 75, DK-2730 Herlev, Denmark, [12] Cancer Epidemiology Unit, Nuffield Department of Clinical Medicine, University of Oxford, Oxford, UK, [13] Surgical Oncology (Uro-Oncology: S4), University of Cambridge, Box 279, Addenbrooke's Hospital, Hills Road, Cambridge, UK and Cancer Research UK Cambridge Research Institute, Li Ka Shing Centre, Cambridge, UK, [14] Centre for Cancer Genetic Epidemiology, Department of Oncology, University of Cambridge, Strangeways Research Laboratory, Worts Causeway, Cambridge, UK, [15] Cambridge Institute of Public Health, University of Cambridge, Forvie Site, Robinson Way, Cambridge CB2 0SR, [16] Division of Public Health Sciences, Fred Hutchinson Cancer Research Center, Seattle, Washington, USA, [17] Department of Epidemiology, School of Public Health, University of Washington, Seattle, Washington, USA, [18] International Epidemiology Institute, 1455 Research Blvd., Suite 550, Rockville, MD 20850, [19] Mayo Clinic, Rochester, Minnesota, USA, [20] Department of Urology, University Hospital Ulm, Germany, [21] Institute of Human Genetics University Hospital Ulm, Germany, [22] Brigham and Women's Hospital/Dana-Farber Cancer Institute, 45 Francis Street- ASB II-3, Boston, MA 02115, [23] Washington University, St Louis, Missouri, [24] International Hereditary Cancer Center, Department of Genetics and Pathology, Pomeranian Medical University, Szczecin, Poland, [25] Division of Genetic Epidemiology, Department of Medicine, University of Utah School of Medicine[26] Division of Clinical Epidemiology and Aging Research & Division of Preventive Oncology, German Cancer Research Center, Heidelberg Germany, [27]German Cancer Consortium (DKTK), German Cancer Research Center (DKFZ), Heidelberg Germany, [28]Division of Cancer Prevention and Control, H. Lee Moffitt Cancer Center, 12902 Magnolia Dr., Tampa, Florida, USA, [29] Molecular Medicine Center and Department of Medical Chemistry and Biochemistry, Medical University - Sofia, 2 Zdrave St, 1431, Sofia, Bulgaria, [30] Australian Prostate Cancer Research Centre-Qld, Institute of Health and Biomedical Innovation and Schools of Life Science and Public

Health, Queensland University of Technology, Brisbane, Australia, [31] Department of Genetics, Portuguese Oncology Institute, Porto, Portugal and Biomedical Sciences Institute (ICBAS), Porto University, Porto, Portugal, [32]The University of Surrey, Guildford, Surrey, GU2 7XH, UK

COGS acknowledgement and funding:

This study would not have been possible without the contributions of the following: Per Hall (COGS); Douglas F. Easton, Paul Pharoah, Kyriaki Michailidou, Manjeet K. Bolla, Qin Wang (BCAC), Andrew Berchuck (OCAC), Rosalind A. Eeles, Douglas F. Easton, Ali Amin Al Olama, Zsofia Kote-Jarai, Sara Benlloch (PRACTICAL), Georgia Chenevix-Trench, Antonis Antoniou, Lesley McGuffog, Fergus Couch and Ken Offit (CIMBA), Joe Dennis, Alison M. Dunning, Andrew Lee, and Ed Dicks, Craig Luccarini and the staff of the Centre for Genetic Epidemiology Laboratory, Javier Benitez, Anna Gonzalez-Neira and the staff of the CNIO genotyping unit, Jacques Simard and Daniel C. Tessier, Francois Bacot, Daniel Vincent, Sylvie LaBoissière and Frederic Robidoux and the staff of the McGill University and Génome Québec Innovation Centre, Stig E. Bojesen, Sune F. Nielsen, Borge G. Nordestgaard, and the staff of the Copenhagen DNA laboratory, and Julie M. Cunningham, Sharon A. Windebank, Christopher A. Hilker, Jeffrey Meyer and the staff of Mayo Clinic Genotyping Core Facility

Funding for the iCOGS infrastructure came from: the European Community's Seventh Framework Programme under grant agreement n° 223175 (HEALTH-F2-2009-223175) (COGS), Cancer Research UK (C1287/A10118, C1287/A 10710, C12292/A11174, C1281/A12014, C5047/A8384, C5047/A15007, C5047/A10692, C8197/A16565), the National Institutes of Health (CA128978) and Post-Cancer GWAS initiative (1U19 CA148537, 1U19 CA148065 and 1U19 CA148112 - the GAME-ON initiative), the Department of Defence (W81XWH-10-1-0341), the Canadian Institutes of Health Research (CIHR) for the CIHR Team in Familial Risks of Breast Cancer, Komen

Foundation for the Cure, the Breast Cancer Research Foundation, and the Ovarian Cancer Research Fund.

Sarcoidosis GWAS

This work was supported by the German Federal Ministry of Education and Research (BMBF) within the framework of the e:Med research and funding concept (SysInflame grant 01ZX1306A). This project received infrastructure support from the DFG Excellence Cluster No. 306 "Inflammation at Interfaces". Andre Franke receives an endowment professorship by the Foundation for Experimental Medicine (Zuerich, Switzerland).

The Singapore Epidemiology of Eye Diseases Study (SEED)

Jia Yu Koh,[1] Qiao Fan,[1] Wanting Zhao,[1] Blanche Lim,[1,2] Jacqueline Chua,[1,3] Paul Mitchell,[4] Jie Jin Wang,[4,5] Yik-Ying Teo,[6,7] Tien Yin Wong,[1,2,3] Ching-Yu Cheng[1,2,3]

1. Singapore Eye Research Institute, Singapore National Eye Center, Singapore
2. Department of Ophthalmology, Yong Loo Lin School of Medicine, National University of Singapore, Singapore
3. Duke-NUS Medical School, Singapore
4. Department of Ophthalmology, Centre for Vision Research, Westmead Millennium Institute, University of Sydney, Sydney, New South Wales, Australia
5. Centre for Eye Research Australia (CERA), University of Melbourne, Royal Victorian Eye and Ear Hospital, Melbourne, Victoria, Australia
6. Saw Swee Hock School of Public Health, National University Health System, National University of Singapore, Singapore
7. Department of Statistics and Applied Probability, National University of Singapore, Singapore

Acknowledgements of studies that contributed to the GWAS meta-analysis of telomere length[4]

The Framingham Heart Study

The Framingham Heart Study is funded by National Institutes of Health contract N01-HC-25195. The Framingham GWAS component of this project was funded by the Division of Intramural Research, National Heart, Lung, and Blood Institute, National Institutes of Health, Bethesda, MD.

TwinsUK

The study was funded by the Wellcome Trust; European Community's Seventh Framework Programme (FP7/2007-2013). The study also receives support from the National Institute for Health Research (NIHR) BioResource Clinical Research Facility and Biomedical Research Centre based at Guy's and St Thomas' NHS Foundation Trust and King's College London.

References

1. Welter D, MacArthur J, Morales J, et al. The NHGRI GWAS Catalog, a curated resource of SNP-trait associations. Nucleic Acids Res 2014;42(Database issue):D1001-6.

2. Hindorff LA LA, MacArthur J, Morales J, et al. A catalog of published genome-wide association studies [Internet]. [cited 2015 Jan 15];Available from: www.genome.gov/gwastudies

3. Pooley KA, Bojesen SE, Weischer M, et al. A genome-wide association scan (GWAS) for mean telomere length within the COGS project: identified loci show little association with hormone-related cancer risk. Hum Mol Genet 2013;22(24):5056–64.

4. Mangino M, Hwang S-J, Spector TD, et al. Genome-wide meta-analysis points to CTC1 and ZNF676 as genes regulating telomere homeostasis in humans. Hum Mol Genet 2012;21(24):5385–94.

5. Prescott J, Kraft P, Chasman DI, et al. Genome-wide association study of relative telomere length. PLoS One 2011;6(5):e19635.

6. Gu J, Chen M, Shete S, et al. A genome-wide association study identifies a locus on chromosome 14q21 as a predictor of leukocyte telomere length and as a marker of susceptibility for bladder cancer. Cancer Prev Res (Phila) 2011;4(4):514–21.

7. Codd V, Nelson CP, Albrecht E, et al. Identification of seven loci affecting mean telomere length and their association with disease. Nat Genet 2013;45(4):422–7.

8. Codd V, Mangino M, van der Harst P, et al. Common variants near TERC are associated with mean telomere length. Nat Genet 2010;42(3):197–9.

9. Liu Y, Cao L, Li Z, et al. A genome-wide association study identifies a locus on TERT for mean telomere length in Han Chinese. PLoS One 2014;9(1):e85043.

10. Saxena R, Bjonnes A, Prescott J, et al. Genome-wide association study identifies variants in casein kinase II (CSNK2A2) to be associated with leukocyte telomere length in a Punjabi Sikh diabetic cohort. Circ Cardiovasc Genet 2014;7(3):287–95.

11. Levy D, Neuhausen SL, Hunt SC, et al. Genome-wide association identifies OBFC1 as a locus involved in human leukocyte telomere biology. Proc Natl Acad Sci U S A 2010;107(20):9293–8.

12. Burgess S. Sample size and power calculations in Mendelian randomization with a single instrumental variable and a binary outcome. Int J Epidemiol 2014;43(3):922–9.

13. Burgess S, Scott RA, Timpson NJ, Davey Smith G, Thompson SG, EPIC-InterAct Consortium. Using published data in Mendelian randomization : a blueprint for efficient identification of causal risk factors. Eur J Epidemiol 2015;30(7):543–52.

14. Bowden J, Davey Smith G, Haycock PC, Burgess S. Consistent estimation in Mendelian randomization with some invalid instruments using a weighted median estimator. Genet Epidemiol

15. Bowden J, Davey Smith G, Burgess S. Mendelian randomization with invalid instruments: effect estimation and bias detection through Egger regression. Int J Epidemiol 2015;44(2):512–25.

16. Purcell S, Neale B, Todd-Brown K, et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. Am J Hum Genet 2007;81(3):559–75.

17. Abecasis GR, Auton A, Brooks LD, et al. An integrated map of genetic variation from 1,092 human genomes. Nature 2012;491(7422):56–65.

18. Durinck S, Moreau Y, Kasprzyk A, et al. BioMart and Bioconductor: a powerful link between biological databases and microarray data analysis. Bioinformatics 2005;21(16):3439–40.

19. Durinck S, Spellman PT, Birney E, Huber W. Mapping identifiers for the integration of genomic datasets with the R/Bioconductor package biomaRt. Nat Protoc 2009;4(8):1184–91.

20. Kinsella RJ, Kahari A, Haider S, et al. Ensembl BioMarts: a hub for data retrieval across taxonomic space. Database 2011;2011:bar030-bar030.

21. Aschard H, Vilhjálmsson BJ, Joshi AD, Price AL, Kraft P. Adjusting for heritable covariates can bias effect estimates in genome-wide association studies. Am J Hum Genet 2015;96(2):329–39.

22. Miyake Y, Nakamura M, Nabetani A, et al. RPA-like mammalian Ctc1-Stn1-Ten1 complex binds to single-stranded DNA and protects telomeres independently of the Pot1 pathway. Mol Cell 2009;36(2):193–206.

23. Burgess S, Thompson SG. Bias in causal estimates from Mendelian randomization studies with weak instruments. Stat Med 2011;30(11):1312–23.

24. Davey Smith G. Use of genetic markers and gene-diet interactions for interrogating population-level causal influences of diet on health. Genes Nutr 2011;6(1):27–43.

25. Haycock PC, Burgess S, Wade KH, Bowden J, Relton C, Davey Smith G. Best (but oft-forgotten) practices: the design, analysis, and interpretation of Mendelian randomization studies. Am J Clin Nutr 2016;103(4):965–78.

26. Davey Smith G, Hemani G. Mendelian randomization: genetic anchors for causal inference in epidemiological studies. Hum Mol Genet 2014;23(R1):R89-98.

27. Schunkert H, König IR, Kathiresan S, et al. Large-scale association analysis identifies 13 new susceptibility loci for coronary artery disease. Nat Genet 2011;43(4):333–8.

28. Bown MJ, Jones GT, Harrison SC, et al. Abdominal aortic aneurysm is associated with a variant in low-density lipoprotein receptor-related protein 1. Am J Hum Genet 2011;89(5):619–27.

29. Borthwick K, Smelser D, Bock J, et al. Ephenotyping for Abdominal Aortic Aneurysm in the Electronic Medical Records and Genomics (eMERGE) Network: Algorithm Development and Konstanz Information Miner Workflow. Int J Biomed Data Min 2015;4(1).

30. Dubois PC a, Trynka G, Franke L, et al. Multiple common variants for celiac disease influencing immune gene expression. Nat Genet 2010;42(4):295–302.

31. Fingerlin TE, Murphy E, Zhang W, et al. Genome-wide association study identifies multiple susceptibility loci for pulmonary fibrosis. Nat Genet 2013;45(6):613–20.

32. Wang Y, McKay JD, Rafnar T, et al. Rare variants of large effect in BRCA2 and CHEK2 affect risk of lung cancer. Nat Genet 2014;46(7).

33. Diskin SJ, Capasso M, Schnepp RW, et al. Common variation at 6q16 within HACE1 and LIN28B influences susceptibility to neuroblastoma. Nat Genet 2012;44(10):1126–30.

34. Pharoah PDP, Tsai Y-Y, Ramus SJ, et al. GWAS meta-analysis and replication identifies three new susceptibility loci for ovarian cancer. Nat Genet 2013;45(4):362–70, 370-2.

35. Spurdle AB, Thompson DJ, Ahmed S, et al. Genome-wide association study identifies a common variant associated with risk of endometrial cancer. Nat Genet 2011;43(5):451–4.

36. Painter JN, O'Mara TA, Batra J, et al. Fine-mapping of the HNF1B multicancer locus identifies candidate variants that mediate endometrial cancer risk. Hum Mol Genet 2015;24(5):1478–92.

37. Purdue MP, Johansson M, Zelenika D, et al. Genome-wide association study of renal cell carcinoma identifies two susceptibility loci on 2p21 and 11q13.3. Nat Genet 2010;43(1):60–5.

38. Turnbull C, Rapley E a, Seal S, et al. Variants near DMRT1, TERT and ATF7IP are associated with testicular germ cell cancer. Nat Genet 2010;42(7):604–7.

39. Rapley EA, Turnbull C, Al Olama AA, et al. A genome-wide association study of testicular germ cell tumor. Nat Genet 2009;41(7):807–10.

40. Amos CI, Wang L-E, Lee JE, et al. Genome-wide association study identifies novel loci predisposing to cutaneous melanoma. Hum Mol Genet 2011;20(24):5012–23.

41. Rafnar T, Sulem P, Thorleifsson G, et al. Genome-wide association study yields variants at 20p12.2 that associate with urinary bladder cancer. Hum Mol Genet 2014;23(20):5545–57.

42. Lambert JC, Ibrahim-Verbaas CA, Harold D, et al. Meta-analysis of 74,046 individuals identifies 11 new susceptibility loci for Alzheimer's disease. Nat Genet 2013;45(12):1452–8.

43. Barrett JC, Clayton DG, Concannon P, et al. Genome-wide association study and meta-analysis find that over 40 loci affect risk of type 1 diabetes. Nat Genet 2009;41(6):703–7.

44. Walsh KM, Codd V, Rice T, et al. Longer genotypically-estimated leukocyte telomere length is associated with increased adult glioma risk. Oncotarget 2015;6(40):42468–77.

45. Ojha J, Codd V, Nelson CP, et al. Genetic Variation Associated with Longer Telomere Length Increases Risk of Chronic Lymphocytic Leukemia. Cancer Epidemiol Biomarkers Prev 2016;25(7):1043–9.

46. Snetselaar R, van Moorsel CHM, Kazemier KM, et al. Telomere length in interstitial lung diseases. Chest 2015;148(4):1011–8.

47. Haycock PC, Heydon EE, Kaptoge S, Butterworth a. S, Thompson A, Willeit P. Leucocyte telomere length and risk of cardiovascular disease: systematic review and meta-analysis. BMJ 2014;349(jul08 3):g4227–g4227.

48. Cottliar A, Palumbo M, La Motta G, et al. Telomere length study in celiac disease. Am J Gastroenterol 2003;98(12):2727–31.

49. Atturu G, Brouilette S, Samani NJ, London NJM, Sayers RD, Bown MJ. Short leukocyte

telomere length is associated with abdominal aortic aneurysm (AAA). Eur J Vasc Endovasc Surg 2010;39(5):559–64.

50. Armanios M. Syndromes of telomere shortening. Annu Rev Genomics Hum Genet 2009;10(46):45–61.

51. Speedy HE, Di Bernardo MC, Sava GP, et al. A genome-wide association study identifies multiple susceptibility loci for chronic lymphocytic leukemia. Nat Genet 2014;46(1):56–60.

52. Kim DHD, Lee S-T, Won H-H, et al. A genome-wide association study identifies novel loci associated with susceptibility to chronic myeloid leukemia. Blood 2011;117(25):6906–11.

53. Postel-Vinay S, Véron AS, Tirode F, et al. Common variants near TARDBP and EGR2 are associated with susceptibility to Ewing sarcoma. Nat Genet 2012;44(3):323–7.

54. Conde L, Halperin E, Akers NK, et al. Genome-wide association study of follicular lymphoma identifies a risk locus at 6p21.32. Nat Genet 2010;42(8):661–4.

55. Cha P-CC, Zembutsu H, Takahashi A, Kubo M, Kamatani N, Nakamura Y. A genome-wide association study identifies SNP in DCC is associated with gallbladder cancer in the Japanese population. J Hum Genet 2012;57(4):235–7.

56. Abnet CC, Freedman ND, Hu N, et al. A shared susceptibility locus in PLCE1 at 10q23 for gastric adenocarcinoma and esophageal squamous cell carcinoma. Nat Genet 2010;42(9):764–7.

57. Chubb D, Weinhold N, Broderick P, et al. Common variation at 3q26.2, 6p21.33, 17p11.2 and 22q13.1 influences multiple myeloma risk. Nat Genet 2013;45(10):1221–5.

58. Bei J-X, Li Y, Jia W-H, et al. A genome-wide association study of nasopharyngeal carcinoma identifies three new susceptibility loci. Nat Genet 2010;42(7):599–603.

59. Tan DEK, Foo JN, Bei J-X, et al. Genome-wide association study of B cell non-Hodgkin

lymphoma identifies 3q27 as a susceptibility locus in the Chinese population. Nat Genet 2013;45(7):804–7.

60. Zhang M, Song F, Liang L, et al. Genome-wide association studies identify several new loci associated with pigmentation traits and skin cancer risk in European Americans. Hum Mol Genet 2013;22(14):2948–59.

61. Köhler A, Chen B, Gemignani F, et al. Genome-wide association study on differentiated thyroid cancer. J Clin Endocrinol Metab 2013;98(10):E1674-81.

62. Li W-Q, Hu N, Hyland PL, et al. Genetic variants in DNA repair pathway genes and risk of esophageal squamous cell carcinoma and gastric adenocarcinoma in a Chinese population. Carcinogenesis 2013;34(7):1536–42.

63. Hüffmeier U, Uebe S, Ekici AB, et al. Common variants at TRAF3IP2 are associated with susceptibility to psoriatic arthritis and psoriasis. Nat Genet 2010;42(11):996–9.

64. Khor CC, Davila S, Breunis WB, et al. Genome-wide association study identifies FCGR2A as a susceptibility locus for Kawasaki disease. Nat Genet 2011;43(12):1241–6.

65. Han F, Faraco J, Dong XS, et al. Genome wide analysis of narcolepsy in China implicates novel immune loci and reveals changes in association prior to versus after the 2009 H1N1 influenza pandemic. PLoS Genet 2013;9(10):e1003880.

66. Zhang X-J, Huang W, Yang S, et al. Psoriasis genome-wide association study identifies susceptibility variants within LCE gene cluster at 1q21. Nat Genet 2009;41(2):205–10.

67. Fischer A, Schmid B, Ellinghaus D, et al. A novel sarcoidosis risk locus for Europeans on chromosome 11q13.1. Am J Respir Crit Care Med 2012;186(9):877–85.

68. Hom G, Graham RR, Modrek B, et al. Association of Systemic Lupus Erythematosus with C8orf13–BLK and ITGAM–ITGAX. N Engl J Med 2008;358(9):900–9.

69. Quan C, Ren Y-Q, Xiang L-H, et al. Genome-wide association study for vitiligo identifies susceptibility loci at 6q27 and the MHC. Nat Genet 2010;42(7):614–8.

70. Xie G, Roshandel D, Sherva R, et al. Association of granulomatosis with polyangiitis (Wegener's) with HLA-DPB1*04 and SEMA6A gene variants: evidence from genome-wide analysis. Arthritis Rheum 2013;65(9):2457–68.

71. Wade TD, Gordon S, Medland S, et al. Genetic variants associated with disordered eating. Int J Eat Disord 2013;46(6):594–608.

72. Otowa T, Yoshida E, Sugaya N, et al. Genome-wide association study of panic disorder in the Japanese population. J Hum Genet 2009;54(2):122–6.

73. Simón-Sánchez J, Schulte C, Bras JM, et al. Genome-wide association study reveals genetic risk underlying Parkinson's disease. Nat Genet 2009;41(12):1308–12.

74. Tang CS, Sribudiani Y, Miao XP, et al. Fine mapping of the 9q31 Hirschsprung's disease locus. Hum Genet 2010;127(6):675–83.

75. Albagha OME, Wani SE, Visconti MR, et al. Genome-wide association identifies three new susceptibility loci for Paget's disease of bone. Nat Genet 2011;43(7):685–9.

76. Kim Y, Kong M, Lee C. Association of intronic sequence variant in the gene encoding spleen tyrosine kinase with susceptibility to vascular dementia. World J Biol Psychiatry 2013;14(3):220–6.

77. Figueroa JD, Ye Y, Siddiq A, et al. Genome-wide association study identifies multiple loci associated with bladder cancer risk. Hum Mol Genet 2014;23(5):1387–98.

78. Zhang B, Jia W-H, Matsuda K, et al. Large-scale genetic study in East Asians identifies six new loci associated with colorectal cancer risk. Nat Genet 2014;46(6):533–42.

79. Coronary Artery Disease (C4D) Genetics Consortium. A genome-wide association study in

Europeans and South Asians identifies five new loci for coronary artery disease. Nat Genet 2011;43(4):339–44.

80. Rajaraman P, Melin BS, Wang Z, et al. Genome-wide association study of glioma and meta-analysis. Hum Genet 2012;131(12):1877–88.

81. Noth I, Zhang Y, Ma S-F, et al. Genetic variants associated with idiopathic pulmonary fibrosis susceptibility and mortality: a genome-wide association study. Lancet Respir Med 2013;1(4):309–17.

82. Mushiroda T, Wattanapokayakit S, Takahashi a, et al. A genome-wide association study identifies an association of a common variant in TERT with susceptibility to idiopathic pulmonary fibrosis. J Med Genet 2008;45(10):654–6.

83. Childs EJ, Mocci E, Campa D, et al. Common variation at 2p13.3, 3q29, 7p13 and 17q25.1 associated with susceptibility to pancreatic cancer. Nat Genet 2015;47(8):911–6.

84. Baranzini SE, Wang J, Gibson RA, et al. Genome-wide association analysis of susceptibility and clinical phenotype in multiple sclerosis. Hum Mol Genet 2009;18(4):767–78.

85. Tse KP, Su WH, Chang KP, et al. Genome-wide Association Study Reveals Multiple Nasopharyngeal Carcinoma-Associated Loci within the HLA Region at Chromosome 6p21.3. Am J Hum Genet 2009;85(2):194–203.

86. Li H, Gan W, Lu L, et al. A genome-wide association study identifies GRK5 and RASGRP1 as type 2 diabetes loci in Chinese Hans. Diabetes 2013;62(1):291–8.

87. van der Valk RJP, Kreiner-Møller E, Kooijman MN, et al. A novel common variant in DCST2 is associated with length in early life and height in adulthood. Hum Mol Genet 2015;24(4):1155–68.

88. Horikoshi M, Yaghootkar H, Mook-Kanamori DO, et al. New loci associated with birth weight identify genetic links between intrauterine growth and adult height and metabolism.

Nat Genet 2013;45(1):76–82.

89. Locke AE, Kahali B, Berndt SI, et al. Genetic studies of body mass index yield new insights for obesity biology. Nature 2015;518(7538):197–206.

90. Bradfield JP, Taal HR, Timpson NJ, et al. A genome-wide association meta-analysis identifies new childhood obesity loci. Nat Genet 2012;44(5):526–31.

91. Taal HR, St Pourcain B, Thiering E, et al. Common variants at 12q15 and 12q24 are associated with infant head circumference. Nat Genet 2012;44(5):532–8.

92. Wood AR, Esko T, Yang J, et al. Defining the role of common variation in the genomic and biological architecture of adult human height. Nat Genet 2014;(November).

93. Shungin D, Winkler TW, Croteau-Chonka DC, et al. New genetic loci link adipose and insulin biology to body fat distribution. Nature 2015;518(7538):187–96.

94. Tobacco and Genetics Consortium. Genome-wide meta-analyses identify multiple loci associated with smoking behavior. Nat Genet 2010;42(5):441–7.

95. Ehret GB, Munroe PB, Rice KM, et al. Genetic variants in novel pathways influence blood pressure and cardiovascular disease risk. Nature 2011;478(7367):103–9.

96. Wain L V, Verwoert GC, O'Reilly PF, et al. Genome-wide association study identifies six new loci influencing pulse pressure and mean arterial pressure. Nat Genet 2011;43(10):1005–11.

97. Rietveld CA, Medland SE, Derringer J, et al. GWAS of 126,559 individuals identifies genetic variants associated with educational attainment. Science 2013;340(6139):1467–71.

98. Saxena R, Hivert M-F, Langenberg C, et al. Genetic variation in GIPR influences the glucose and insulin responses to an oral glucose challenge. Nat Genet 2010;42(2):142–8.

99. Dupuis J, Langenberg C, Prokopenko I, et al. New genetic loci implicated in fasting glucose

homeostasis and their impact on type 2 diabetes risk. Nat Genet 2010;42(2):105–16.

100. Soranzo N, Sanna S, Wheeler E, et al. Common variants at 10 genomic loci influence hemoglobin A_1(C) levels via glycemic and nonglycemic pathways. Diabetes 2010;59(12):3229–39.

101. van der Harst P, Zhang W, Mateo Leach I, et al. Seventy-five genetic loci influencing the human red blood cell. Nature 2012;492(7429):369–75.

102. Willer CJ, Schmidt EM, Sengupta S, et al. Discovery and refinement of loci associated with lipid levels. Nat Genet 2013;45(11):1274–83.

103. Köttgen A, Pattaro C, Böger C a, et al. New loci associated with kidney function and chronic kidney disease. Nat Genet 2010;42(5):376–84.

104. Liao J, Su X, Chen P, et al. Meta-analysis of genome-wide association studies in multiethnic Asians identifies two loci for age-related nuclear cataract. Hum Mol Genet 2014;23(22):6119–28.

105. Speliotes EK, Yerges-Armstrong LM, Wu J, et al. Genome-wide association analysis identifies variants associated with nonalcoholic fatty liver disease that have distinct effects on metabolic traits. PLoS Genet 2011;7(3).

106. Manichaikul A, Hoffman E a, Smolonska J, et al. Genome-wide study of percent emphysema on computed tomography in the general population. The Multi-Ethnic Study of Atherosclerosis Lung/SNP Health Association Resource Study. Am J Respir Crit Care Med 2014;189(4):408–18.

107. Huffman JE, Albrecht E, Teumer A, et al. Modulation of genetic associations with serum urate levels by body-mass-index in humans. PLoS One 2015;10(3):e0119752.

108. McGrath M, Wong JYY, Michaud D, Hunter DJ, De Vivo I. Telomere length, cigarette

smoking, and bladder cancer risk in men and women. Cancer Epidemiol Biomarkers Prev 2007;16(4):815–9.

109. Weischer M, Nordestgaard BG, Cawthon RM, Freiberg JJ, Tybjærg-Hansen A, Bojesen SE. Short telomere length, cancer survival, and cancer risk in 47102 individuals. J Natl Cancer Inst 2013;105(7):459–68.

110. Qu S, Wen W, Shu X-O, et al. Association of leukocyte telomere length with breast cancer risk: nested case-control findings from the Shanghai Women's Health Study. Am J Epidemiol 2013;177(7):617–24.

111. Kim S, Sandler DP, Carswell G, et al. Telomere length in peripheral blood and breast cancer risk in a prospective case-cohort analysis: results from the Sister Study. Cancer Causes Control 2011;22(7):1061–6.

112. Pooley KA, Sandhu MS, Tyrer J, et al. Telomere length in prospective and retrospective cancer case-control studies. Cancer Res 2010;70(8):3170–6.

113. Lee I-M, Lin J, Castonguay AJ, Barton NS, Buring JE, Zee RYL. Mean leukocyte telomere length and risk of incident colorectal carcinoma in women: a prospective, nested case-control study. Clin Chem Lab Med 2010;48(2):259–62.

114. Zee RYL, Castonguay AJ, Barton NS, Buring JE. Mean telomere length and risk of incident colorectal carcinoma: a prospective, nested case-control approach. Cancer Epidemiol Biomarkers Prev 2009;18(8):2280–2.

115. Cui Y, Cai Q, Qu S, et al. Association of leukocyte telomere length with colorectal cancer risk: nested case-control findings from the Shanghai Women's Health Study. Cancer Epidemiol Biomarkers Prev 2012;21(10):1807–13.

116. Prescott J, McGrath M, Lee I-M, Buring JE, De Vivo I. Telomere length and genetic analyses in population-based studies of endometrial cancer risk. Cancer 2010;116(18):4275–

82.

117. Walcott F, Rajaraman P, Gadalla SM, et al. Telomere length and risk of glioma. Cancer Epidemiol 2013;37(6):935–8.

118. Hofmann JN, Lan Q, Cawthon R, et al. A prospective study of leukocyte telomere length and risk of renal cell carcinoma. Cancer Epidemiol Biomarkers Prev 2013;22(5):997–1000.

119. Seow WJ, Cawthon RM, Purdue MP, et al. Telomere length in white blood cell DNA and lung cancer: a pooled analysis of three prospective cohorts. Cancer Res 2014;74(15):4090–8.

120. Nan H, Du M, De Vivo I, et al. Shorter telomeres associate with a reduced risk of melanoma development. Cancer Res 2011;71(21):6758–63.

121. Lynch SM, Major JM, Cawthon R, et al. A prospective analysis of telomere length and pancreatic cancer in the alpha-tocopherol beta-carotene cancer (ATBC) prevention study. Int J Cancer 2013;133(11):2672–80.

122. Campa D, Mergarten B, De Vivo I, et al. Leukocyte telomere length in relation to pancreatic cancer risk: a prospective study. Cancer Epidemiol Biomarkers Prev 2014;23(11):2447–54.

123. Julin B, Shui I, Heaphy CM, et al. Circulating leukocyte telomere length and risk of overall and aggressive prostate cancer. Br J Cancer 2015;112(4):769–76.

124. Liang G, Qureshi AA, Guo Q, De Vivo I, Han J. No association between telomere length in peripheral blood leukocytes and the risk of nonmelanoma skin cancer. Cancer Epidemiol Biomarkers Prev 2011;20(5):1043–5.

125. Haycock PC, Heydon EE, Kaptoge S, Butterworth AS, Thompson A, Willeit P. Leucocyte telomere length and risk of cardiovascular disease: systematic review and meta-analysis. BMJ 2014;349:g4227.

126. Willeit P, Raschenberger J, Heydon EE, et al. Leucocyte Telomere Length and Risk of Type

2 Diabetes Mellitus: New Prospective Cohort Study and Literature-Based Meta-Analysis. PLoS One 2014;9(11):e112483.

127. Ma H, Zhou Z, Wei S, et al. Shortened telomere length is associated with increased risk of cancer: a meta-analysis. PLoS One 2011;6(6):e20466.

128. Wentzensen IM, Mirabello L, Pfeiffer RM, Savage SA. The association of telomere length and cancer: a meta-analysis. Cancer Epidemiol Biomarkers Prev 2011;20(6):1238–50.

129. Gretarsdottir S, Baas AF, Thorleifsson G, et al. Genome-wide association study identifies a sequence variant within the DAB2IP gene conferring susceptibility to abdominal aortic aneurysm. Nat Genet 2010;42(8):692–7.

130. Jones GT, Bown MJ, Gretarsdottir S, et al. A sequence variant associated with sortilin-1 (SORT1) on 1p13.3 is independently associated with abdominal aortic aneurysm. Hum Mol Genet 2013;22(14):2941–7.

131. Harrison SC, Smith AJP, Jones GT, et al. Interleukin-6 receptor pathways in abdominal aortic aneurysm. Eur Heart J 2013;34(48):3707–16.

132. Elmore JR, Obmann MA, Kuivaniemi H, et al. Identification of a genetic variant associated with abdominal aortic aneurysms on chromosome 3p12.3 by genome wide association. J Vasc Surg 2009;49(6):1525–31.

133. Willer CJ, Li Y, Abecasis GR. METAL: fast and efficient meta-analysis of genomewide association scans. Bioinformatics 2010;26(17):2190–1.