Supplementary Information

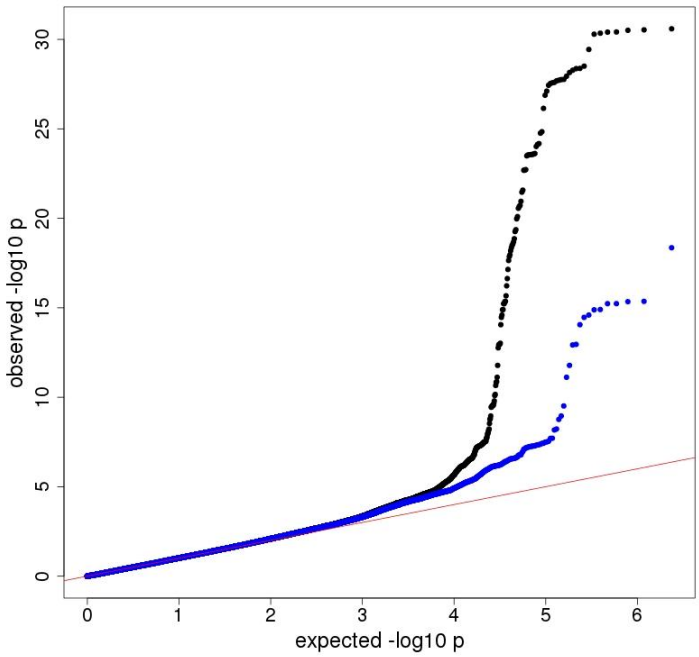
Identification of seven loci affecting mean telomere length and their association with disease

Veryan Codd, Christopher P. Nelson, Eva Albrecht, Massimo Mangino, Joris Deelen, Jessica L. Buxton, Jouke Jan Hottenga, Krista Fischer, Tõnu Esko, Ida Surakka, Linda Broer, Dale R. Nyholt, Irene Mateo Leach, Perttu Salo, Sara Hägg, Mary K. Matthews, Jutta Palmen, Giuseppe D. Norata, Paul F. O'Reilly, Danish Saleheen, Najaf Amin, Anthony J. Balmforth,Marian Beekman, Rudolf A. de Boer,Stefan Böhringer, Peter S. Braund, Paul R. Burton,Anton J. M. de Craen, Matthew Denniff, Yanbin Dong, Konstantinos Douroudis, Elena Dubinina, Johan G. Eriksson, Katia Garlaschelli, Dehuang Guo, Anna-Liisa Hartikainen, Anjali K Henders, Jeanine J. Houwing-Duistermaat, Laura Kananen, Lennart C. Karssen, Johannes Kettunen, Norman Klopp, Vasiliki Lagou, Elisabeth M. van Leeuwen, Pamela Madden, Reedik Mägi, Patrik K.E. Magnusson, Satu Männistö, Mark I. McCarthy, Sarah E. Medland, Evelin Mihailov, Grant W. Montgomery, Ben A. Oostra, Aarno Palotie, Annette Peters, Helen Pollard, Anneli Pouta, Inga Prokopenko, Samuli Ripatti, Veikko Salomaa, H. Eka D. Suchiman[5], Ana M.Valdes, Niek Verweij, Ana Vinuela, Xiaoling Wang, H.-Erich Wichmann, Elisabeth Widen, Gonneke Willemsen, Margie J. Wright, Kai Xia, Xiangjun Xiao,Dirk J. van Veldhuisen, Alberico L. Catapano, Martin D. Tobin, Alistair S. Hall, Alexandra I.F. Blakemore, Wiek H. van Gilst, Haidong Zhu, CARDIoGRAM consortium, Jeanette Erdmann, Muredach P. Reilly, Sekar Kathiresan, Heribert Schunkert, Philippa J. Talmud, Nancy L. Pedersen, Markus Perola, Willem Ouwehand, Jaakko Kaprio, Nicholas G. Martin, Cornelia M. van Duijn, Iiris Hovatta, Christian Gieger, Andres Metspalu, Dorret I. Boomsma, Marjo-Riitta Jarvelin, P. Eline Slagboom, John R. Thompson, Tim D. Spector, Pim van der Harst, Nilesh J. Samani.

Nature Genetics: doi:10.1038/ng.2528

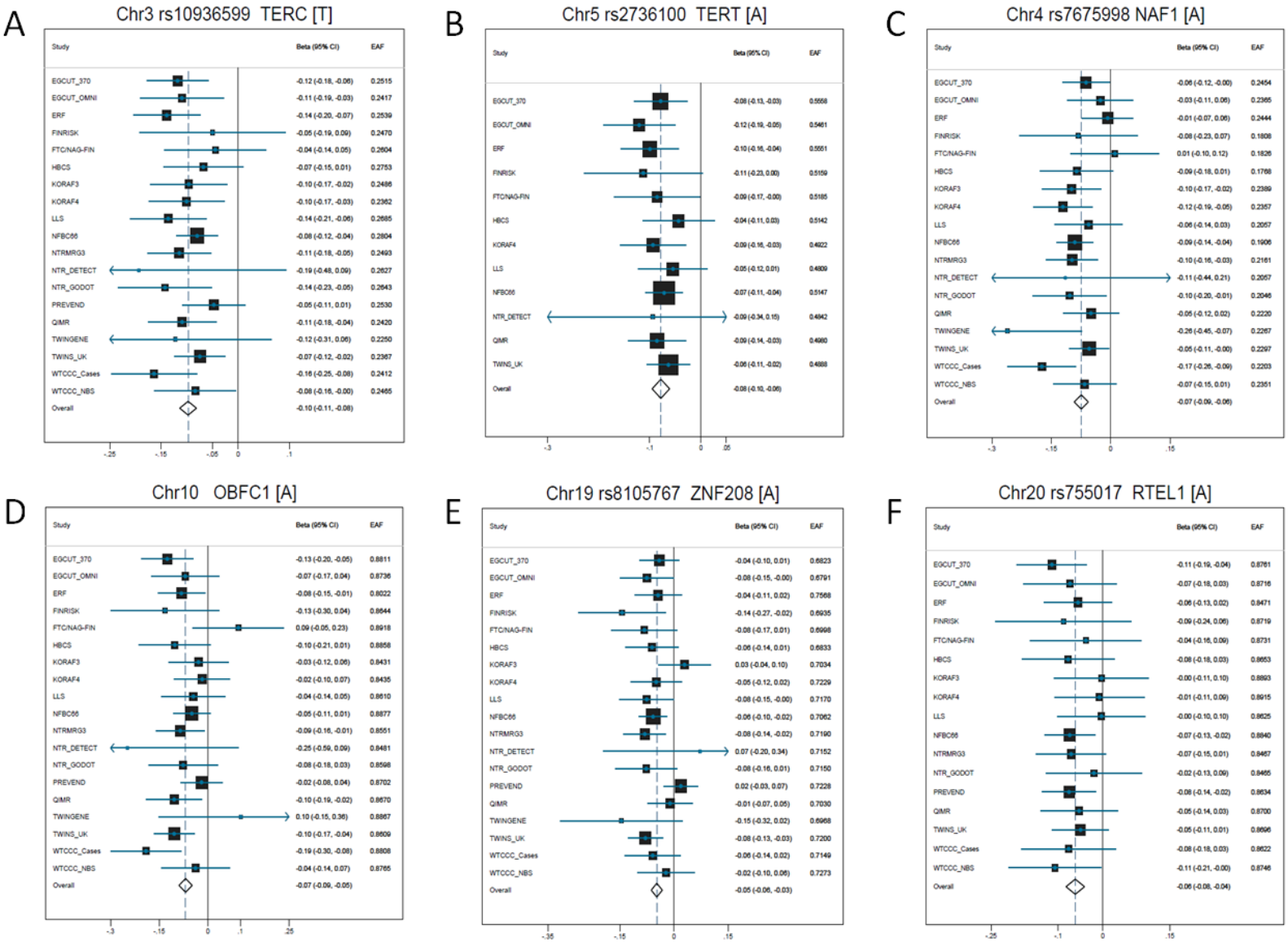
Content

Nature Genetics: doi:10.1038/ng.2528

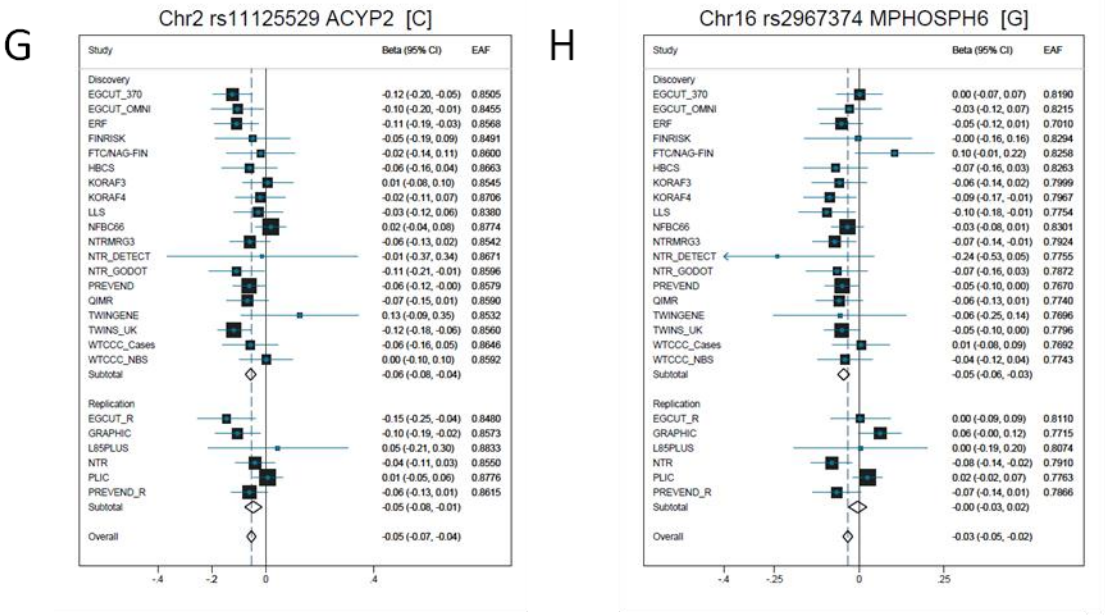


Supplementary Figure 1: Quantile-quantile plot of observed versus expected results (on the basis of a null hypothesis) of the whole association meta-analysis (black dots) and after removal of the previously described chromosome 3q26 *TERC* and 10q24.33 *OBFC1* loci (blue dots)

Nature Genetics: doi:10.1038/ng.2528

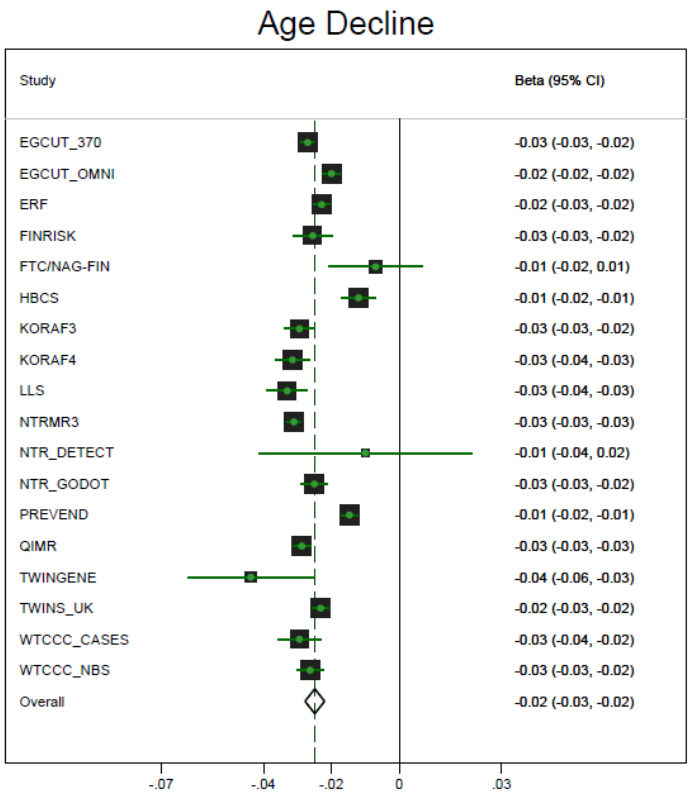


Nature Genetics: doi:10.1038/ng.2528



Supplementary Figure 2. Forest plots of the identified loci showing the association results for individual cohorts. The plots for the two SNPs taken forward for replication (G, H) also include the results from the replication cohorts. All effect sizes are plotted with 95% CI intervals for the allele associating with shorter LTL. The effect allele is specified in the title for each plot. Box sizes represent the weight attributed to each study. The frequencies for the effect allele (EAF) are given on the right hand side of each plot and good agreement can be seen across studies.

Nature Genetics: doi:10.1038/ng.2528



Supplementary Figure 3. Estimates of cross-sectional age-related decline in telomere length based on z-scored LTL. Effect sizes are plotted with 95% CI intervals for each study and box sizes represent the weight attributed to each study. The overall estimate of per-allele decline in LTL is from a random-effects meta-analysis across all studies.

Nature Genetics: doi:10.1038/ng.2528

Supplementary Table 1: Cohort demographics for studies included in the GWA meta-analysis (1A) and replication phase (1B).

Supplementary Table 1A

Cohort	Nationality	Cohort Type	N	Age distribution Mean +/- SD (Range)	Sex distribution % Male	T/S distribution Mean +/- SD (Range)	T/S change per year	Sex effect	LTL laboratory	LTL CV (%)
BHF-FHS	UK	CAD	1487	60.8±7.9 (36-82)	80.1	1.35±0.22 (0.69 - 2.13)	-0.006**	0.033*	1	3.5
EGCUT_370	Estonia	Population	2309	40.1±16.2 (17-91)	48.2	1.85±0.33 (0.94- 3.52)	-0.009**	0.085**	1	3.7
EGCUT_OMNI	Estonia	Population	1251	58.9±22.2 (18 - 103)	38.7	1.69±0.30 (0.93 - 2.84)	-0.006**	0.057*	1	3.7
ERF	Netherlands	Population Family based	2581	49.76±14.87 (17-89)	44.5	1.78 ±0.36 (0.77-3.17)	-0.008**	0.068**	1	3.5
FINRISK	Finland	Population	520	52.2±13.8 (25-74)	47.4	0.150± 0.186 (-0.37 - 0.77)	-0.0047**	0.018	2	7.7
FTC/NAG-FIN	Finland	Smokers, Twin	1054	54.9±4.5 (41- 76)	61.9	0.91±0.16 (0.52 - 1.42)	-0.001	0.056**	2	8.2
HBCS	Finland	Population	1582	61.5±2.9 (56 – 69)	43.4	1.39 - 0.29 (0.61 - 2.32)	-0.012**	0.043*	2	24.8
KORA F3	Germany	Population	1636	62.2±10.1 (34-79)	49.6	1.70±0.28 (0.92 - 2.62)	-0.008**	0.061**	1	3.6
KORA F4	Germany	Population	1801	60.9±8.9 (31-81)	48.6	1.80±0.31 (1.02 - 3.01)	-0.010**	0.112**	1	3.1
LLS	Netherlands	Population	2266	59.24± 6.8 (30 - 80)	45.7	1.46±0.27, (0.74 – 2.43)	-0.009**	0.045**	1	2.7
NFBC1966	Finland	Population Birth cohort	5146	31±0	48.2	1.22±0.48 (0.28-4.88)	N/A	0.059**	3	6.2
NTRMRG3	Netherlands	Population, Twin	2532	43.6±14.7 (17 – 82)	33.3	2.67±0.49 (1.01 – 4.47)	-0.015**	0.177**	1	3.7

Nature Genetics: doi:10.1038/ng.2528

NTR_DETECT	Netherlands	Population, Twin	158	18.7±3.9 (17 - 57)	52.5	2.99±0.38 (1.52 – 4.21)	0.000	-0.003†	1	3.7
NTR_GODOT	Netherlands	Population, Twin	1435	35.7±10.6 (17 - 76)	37.8	2.83±0.47 (0.81 – 4.61)	-0.012**	0.063*	1	3.7
PREVEND	Netherlands	Population	2926	48.0±11.1 (28-69)	51.0	0.004±0.28 (-0.68 - 1.015)	-0.004**	0.033**	4	3.9
QIMR	Australia	Population, Twin	2371	24.0±14.9 (7 -72)	49.1	3.49±0.61 (1.47 – 5.72)	-0.018**	0.093**	1	3.9
TWINGENE	Sweden	Population, Twin	300	71.7±5.9 (55 - 91)	0	1.43±0.25 (0.96 -2.26)	-0.011**	N/A	1	2.9
TWINSUK	UK	Population, Twin	4899	51.0±13.4 (16 - 99)	9.0	3.71±0.68 (0.68 – 11.40)	-0.016**	-0.008†	1	3.3
UKBS	UK	Population	1430	43.4±12.4 (17-69)	48.4	1.80±0.50 (0.80 - 3.01)	-0.009**	0.035*	1	3.5

Nature Genetics: doi:10.1038/ng.2528

Supplementary Table 1B

Cohort	Nationality	Cohort Type	N	Age distribution Mean +/- SD (Range)	Sex distribution % Male	T/S distribution Mean +/- SD (Range)	T/S change per year	Sex effect	LTL laboratory	LTL CV (%)
EGCUT_R	Estonia	Population	1087	47.3±17.5 (18-93)	57.0	1.77±0.32 (0.88-3.24)	-0.006**	0.045	1	3.8
GRAPHIC	UK	Population, Family	2020	39.3±14.5 (18-60)	50.5	1.60±0.26 (0.50 - 3.01)	-0.007**	0.051**	1	3.6
L85PLUS	Netherlands	Population (Elderly)	298	85±0	28.9	2.98±0.74 (1.66-6.52)	NA	0.061	1	3.4
PLIC	Italy	Population	1871	54.6±11.4 (18 - 82)	42.1	1.11±0.36 (0.01- 2.54)	-0.005**	0.014	5	<5
PREVEND	Netherlands	Population	2878	47.5±11.5 (28-69)	47.5	1.00±0.25 (0.47-1.65)	-0.004**	0.022*	4	3.9
NTR	Netherlands	Population, Twin	2585	48.28±15.66 (12-90)	41.5	2.53±0.46 (0.96-4.6)	-0.013*	0.016*	1	3.9

T/S distributions are given from the primary data for each study prior to z-transformation for analysis. Level of statistical significance is denoted by *P<0.01, **P<0.0001. Sex effect is shown as the effect of being female on T/S. As expected, LTL was found to be longer in females in the majority of studies although this effect was not statistically significant for all. †Two exceptions to this were the NTR_DETECT and TWINSUK cohorts, most likely due to the small sample size and relative lack of male subjects in these studies respectively. All cohorts showed the expected age-associated decline in TL, again with the exception of the very small NTR_DETECT cohort. Where non-normal distribution of LTL measurements was observed logT/S is given. For the measurement laboratory: 1, Leicester; 2, Helsinki; 3, Imperial College London; 4; Groningen; 5,UCL London. The mean inter-run coefficient of variation for each study (calculated for the T/S ratio) is also stated.

Nature Genetics: doi:10.1038/ng.2528

Supplementary Table 2: Details of genotyping platforms, imputation algorithm and analysis methods used by each study.

Study	Genotyping Platform	Genotype calling algorithm	Genotyped SNPs	Imputation algorithm	Total SNPs (after QC)	Analysis program	Study-specific covariates	Genomic inflation control
BHF-FHS	Affymetrix 500K	CHIAMO	470,454	IMPUTE	2,250,328	SNPTest	-	1.014
EGCUT_370	Illumina HumanCNV370 HumanOmniExpress	GenomeStudio	321,407	IMPUTE	2,337,450	SNPTest	First three principle components	1.031
EGCUT_OMNI	Illumina HumanCNV370 HumanOmniExpress	GenomeStudio	630,155	IMPUTE	2,434,144	SNPTest	First three principle components	1.005
ERF	Illumina6K, Illumina 318K, Illumina370K, Affymetrix 250K	Beadstudio	650,197	MACH	2,357,460	ProbABEL	Family structure	1.070
FINRISK	Illumina 610 Quad	Illuminus	554,988	MACH	2,394,977	ProbABEL	-	1.025
FTC/NAG-FIN	Illumina HumanHap670K	Illuminus	549,060	IMPUTE	2,505,753	SNPTest	Family structure	1.007
HBCS	Illumina HumanHap670K	Illuminus	546,814	MACH	2,398,230	ProbABEL	-	1.003
KORA F3	Affymetrix 500K	BRLMM	379,392	IMPUTE	2,326,768	SNPTest	-	1.006
KORA F4	Affymetrix 6.0	Birdseed2	909,622	IMPUTE	2,406,045	SNPTest	-	1.020
LLS	Illumina 660w-quad / IlluminaOmniExpress	GenomeStudio	298,538	IMPUTE	2,391,111	QT-assoc	Family Structure	1.062
L85PLUS	IlluminaOmniExpress	GenomeStudio	603,314	IMPUTE	2,470,677	QT-assoc		1.010
NFBC1966	Illumina	Beadstudio	339,629	IMPUTE	2,365,573	SNPTest	First three	1.028

Nature Genetics: doi:10.1038/ng.2528

	HumanCNV370DUO						principle components	
NTRMRG3	Illumina 370 Affy-PerlegenIllumina 660Affymetrix 6.0 Illumina Omni Express	Beadstudio Afymterix_ perlegen proprietary, Birdseed 2	1,257,594	IMPUTE	2,300,386	EMMAX	Family structure	1.028
NTR_DETECT	Affymetrix 6.0	Affy Birdseed 2	729,472	Beagle/Minimach	2,047,577	PLINK	Family structure	1.112
NTR_GODOT	Affymetrix 6.0	Affy Birdseed 2	666,284	Minimach	2,296,548	PLINK	Family structure	1.018
PREVEND	Illumina CytoSNP12 v2	GenomeStudio	244,868	Beagle v3.1.0	1,734,983	PLINK	-	1.002
QIMR	Illumina HumanHap610K	Beadstudio	529,721	MACH	2,358,027	Merlin-offline	Family structure	0.993
TWINGENE	Illumina 317K platform	Beadstudio	317,000	IMPUTE	2,324,772	PLINK	Family structure	1.012
TWINSUK	Illumina HumanHap300	Illuminus	303,940	IMPUTE	1,839,131	GenABEL (v1.6-7)	Family structure	1.012
	Illumina HumanHap610Q		553,487					
	Illumina 1M-Duo		874,733					
UKBS	Affymetrix 500K	CHIAMO	470,398	IMPUTE	2,252,636	SNPTest	-	0.998

All studies were adjusted for age and sex (unless single age or single gender). Additional study specific covariates were included where families were present or where evidence of population structure was observed.

Nature Genetics: doi:10.1038/ng.2528

Supplementary Table 3: **Sex specific results for all loci associated with LTL**

					Females				Males				Sex difference
SNP	**Chr**	**Position**	**Effect Allele**	**Other Allele**	**N**	**Beta**	**SE**	***P*-value**	**N**	**Beta**	**SE**	***P*-value**	***P*-value**
rs11125529	2	54,329,370	C	A	21679	-0.068	0.014	6.39E-07	15984	-0.040	0.015	7.53E-03	0.163
rs10936599	3	170,974,795	T	C	21693	-0.098	0.011	4.67E-18	15986	-0.101	0.012	1.42E-16	0.829
rs7675998	4	164,227,270	A	G	20222	-0.061	0.012	5.03E-07	14482	-0.093	0.014	1.18E-11	0.079
rs2736100	5	1,339,516	A	C	15478	-0.077	0.011	9.04E-12	10374	-0.082	0.014	1.99E-09	0.782
rs9420907	10	105,666,455	A	C	21683	-0.069	0.014	8.41E-07	15980	-0.070	0.016	6.89E-06	0.968
rs8105767	19	22,007,281	A	G	21564	-0.058	0.011	5.23E-08	15967	-0.032	0.011	5.92E-03	0.092
rs755017	20	61,892,066	A	G	21208	-0.050	0.014	4.50E-04	15864	-0.073	0.016	2.83E-06	0.288

Per allele effect sizes, standard errors and significance levels on a fixed-effects model are reported separately for each sex and analysed for a potential difference between them. All effect sizes are stated for the allele associating with shorter LTL.

Nature Genetics: doi:10.1038/ng.2528

Supplementary Table 4: Conditional analysis to test three loci showing potential independent signals.

		Single SNP model				Multiple SNP model				
		n	**beta**	**se**	**p**	**n**	**Beta**	**se**	**p**	**ΔBeta**
chr5	rs2736100	21698	-0.079	0.009	6.51E-17	21312	-0.059	0.010	3.56E-09	24.95%
	rs1801075	21322	-0.069	0.012	3.01E-08	21312	-0.032	0.013	1.42E-02	52.87%
	rs2853676	21709	-0.084	0.011	9.42E-14	21312	-0.029	0.013	2.44E-02	64.87%
chr10	rs9420907	35087	-0.071	0.011	4.39E-11	34767	-0.046	0.008	5.91E-09	34.47%
	rs11191849	34767	-0.041	0.007	1.56E-08	34767	-0.036	0.008	6.57E-06	13.09%
chr20	rs755017	29989	-0.066	0.011	8.64E-09	29986	-0.036	0.012	2.27E-03	45.71%
	rs6011040	29994	-0.046	0.008	5.35E-08	29986	-0.020	0.009	2.87E-02	55.89%

Effect estimates of potential independent SNPs were compared between a joint multiple SNP model and their single SNP model results within available studies.R^2 and *D'* between two SNPs is given from HapMap II release 22. Single and multiple SNP model analyses were performed in the same sub-set of studies. The sample sizes are similar but not identical due to some missing data points in one or other analyses. Studies that were not included in these analyses were: Chr5 (BHF-FHS, KORA F3, KORA F4, LLS, NBS, NTRMRG3, NTR_DETECT, NTR_GODOT, PREVEND, TWINGENE), Chr10 (LLS, TWINGENE), Chr20 (LLS, NTR_DETECT, TWINGENE, TWINSUK).

Nature Genetics: doi:10.1038/ng.2528

Supplementary Table 5: Details of key genes within each LTL-associated locus.

Chr	Lead SNP	Gene name	Distance from lead SNP (KB)	Gene function
2	rs11125529	*Acylphosphatase 2, muscle type (ACYP2)*	Within gene	This gene has roles in muscle differentiation and stress induced apoptosis in muscle[1].
		Testis-specific Y-encoded-like protein 6 (TSPYL6)	4.5	*TSPYL6* shows homology to nucleosome assembly proteins.
		Proteasome activator subunit 4 (PSME4,also known as PA200)	282	*PSME4* is a proteasome activator involved in DNA repair. It is required for normal spermatogenesis in mice[2].
3	rs10936599	*Myoneurin (MYNN)*	Within gene	BTB/POZ and zinc finger domain-containing transcription factor.
		Actin related protein M1 (ACTRT3, also known as ARPM1)	4.4	Testis-specific protein thought to play a role in nuclear organisation during spermiogenesis[3]
		Telomerase RNA component (TERC)	9.3	The RNA encoded by this gene provides the template of the telomere repeat sequence within the telomerase enzyme complex. Other components of telomerase are TERT, Dyskerin, NOP10 and NHP2[4]. It is a member of the H/ACA snoRNA family[4]. Mutations within TERC cause the autosomal dominant form of Dyskeratosis congenita[5], a disease linked with telomere biology and premature aging.
		Leucine rich repeat containing 34 (LRRC34)	19.1	The three members of the leucine rich repeat containing family (*LRRC34*, *LRRIQ4* and *LRRC31*) are of unknown function. Other members of this family function in DNA repair, cell cycle regulation, apoptosis and chromosomal stability.
		Leucine-rich repeats and IQ motif containing 4 (LRRIQ4)	47.6	
		Leucine rich repeat containing 31 (LRRC31)	64.9	

Nature Genetics: doi:10.1038/ng.2528

		SEC62 homolog (SEC62)	192	Forms part of the protein translocation apparatus of the ER. Increase levels of SEC62 have been associated with both lung and prostate cancer.
		Polyhomeotic homolog 3 (*PHC3)*	313	Part of the human polcomb complex thought to act as a tumour suppressor.
		Protein Kinase C, iota (PRKCI)	448	Shown to protect leukaemia cells against drug induced apoptosis
		MDS1 and EVI1 complex locus (MECOM)	691	Transcriptional factor and oncoprotein that may have roles in hematopoiesis, apoptosis, cell differentiation, proliferation and development
4	rs7675998	***Nuclear assembly factor 1 homolog (NAF1)***	40.0	NAF1is required for H/ACA box snoRNA assembly. It is involved in the formation of the telomerase enzyme and is replaced by GAR1after initial RNP assembly to form the mature complex[4].
		Neuropeptide Y receptor Y1 (NPY1R) *Neuropeptide Y receptor Y5 (NPY5R)*	237 257	Both NPY1R and NPY5R are receptors for neuropeptide Y, one on the most abundant neutorpeptides in the mammalian nervous system. Roles of neuropeptide Y include regulation of energy homeostasis and vasoactive effects on the cardiovascular system.
5	rs2736100	**Telomerase reverse transcriptase (TERT)**	Within gene	*TERT* encodes the reverse transcriptase subunit of the telomerase enzyme. Both deregulation of TERT expression and mutations within this gene are linked to several forms of cancer. Mutations within TERT also cause dysteratosis congenita[5].
		Cleft Lip and Palate Transmembrane Protein 1-like (CLPTM1L)	31.4	This protein is of unknown function but mutations within this gene cause cleft lip and palate.
		Naked cuticle homolog 2 (NKD2)	248	Negative regulator of WNT signalling
10	rs9420907	***Oligonucleotide/oligosaccharide-binding fold containing 1 (OBFC1)***	Within gene	Component of the telomere associated complex CST, along with CTC1 and TEN1 which bind and protect telomeres via association

Nature Genetics: doi:10.1038/ng.2528

				with the shelterin complex[6,7]. OBFC1 is also is a subunit of alpha accessory factor (AAF) that is involved in the initiation of DNA replication[8].
		STE20-like kinase (SLK)	51.0	SLK is involved in the regulation of cancer cell motility
		SH3 and PX domains 2A (SH3PXD2A)	61.3	This protein is a podosome/invadopodia scaffold protein involved in tumour growth and invasion.
		SW15-dependent recombination repair 1 (SFR1, c10orf78)	205	SFR1 is involved in homologous DNA recombination and repair
		WD repeat domain containing 96 (WDR96, c10orf79)	213	WDR96 is a neuralised homolog. It acts to increase apoptosis and is a consequently a candidate tumour suppressor.
19	rs8105767	*Zinc finger protein 208 (ZNF208)* *Zinc finger protein 43 (ZNF43)*	21.7 181	This region contains a cluster of zinc finger proteins though to be transcriptional regulators. It is a beta-satellite repeat region that has arisen later in evolution and is found only in primates. Of these ZNF208 is considered to be a novel member of the family although nothing is known of its function[9]. ZNF43 is predicted a member of the C2H2-type zinc finger proteins, a family involved in gene regulation and development.
20	rs755017	*Zinc finger and BTB domain containing 46 (ZBTB46)*	15.2	Recent studies show ZBTB46 to be a novel transcription factor that is expressed specifically in classical dentritic cells amongst the mature hematopoietic cells. It is also expressed in erythroid progenitors and endothelium[10].
		Lck interacting transmembrane adaptor 1 (LIME1)	51.2	This gene is expressed mainly in T- and B-cells and is responsible for their activation[11,12]. Activation of these cells results in their proliferation.

Nature Genetics: doi:10.1038/ng.2528

	Regulator of telomere elongation helicase 1 (RTEL1)	94.0	RTEL1 is an essential helicase that is involved in setting telomere length and functions telomere maintenance and DNA repair in mice[13,14].
	tumor necrosis factor receptor superfamily, member 6b, decoy (TNFRSF6B)	91.6	This protein is a member of the tumour necrosis factor receptor family and acts to regulate cell death. A read though transcript (RTEL1- TNFRSF6B) from the neighbouring RTEL1 gene produces a non-coding RNA.
	tumor protein D52-like 2 (*TPD52L2)*	75.0	Member of the tumour protein D52-like family.
	SRY(sex determination region Y)-box 18 (SOX18)	257	SOX18 is a transcriptional regulator involved in development and determination of cell fate.
	Regulator of G-protein signalling 19 (RGS19)	283	Over-expression of this gene results in cell proliferation via deregulation of cell cycle control factors[15].

Genes were selected on proximity to the lead SNP and potential function with relation to LTL. Those with a known function in telomere biology are given in bold type.

Nature Genetics: doi:10.1038/ng.2528

Supplementary Table 6: Association of telomere-length SNPs with telomerase activity.

SNP	Chr	Effect Allele	EAF	Raw telomerase activity			Log transformed telomerase activity			Beta	P-value
				0	1	2	0	1	2		
rs11125529	2	C	0.797	0.872±0.528	1.187±0.660	1.164±0.524	0.456±0.278	0.608±0.338	0.597±0.277	0.085	0.196
rs10936599	3	T	0.131	1.143±0.581	1.221±0.558	1.066±0.551	0.584±0.302	0.634±0.291	0.549±0.284	0.091	0.240
rs7675998	4	G	0.800	1.089±0.457	1.092±0.632	1.199±0.570	0.560±0.236	0.560±0.323	0.615±0.300	0.107	0.100
rs2736100	5	T	0.547	1.157±0.565	1.159±0.626	1.155±0.495	0.595±0.290	0.596±0.322	0.589±0.269	-0.015	0.783
rs9420907	10	C	0.398	1.068±0.525	1.225±0.534	1.209±0.715	0.542±0.276	0.635±0.277	0.620±0.368	0.058	0.312
rs8105767	19	G	0.370	1.194±0.624	1.128±0.552	1.203±0.549	0.607±0.329	0.582±0.285	0.618±0.281	0.000	0.996
rs755017	20	G	0.216	1.128±0.537	1.201±0.651	1.086±0.541	0.581±0.277	0.612±0.346	0.554±0.268	-0.008	0.895

All SNPs were tested for association with leukocyte telomerase activity in 208 subjects. Telomerase activity (raw and log-transformed) values are shown by genotype group: 0: wild type, 1: heterozygotes, 2: effect allele homozygote .The effect allele tested, the effect allele frequency (EAF) and the effect size (beta) based on log transformed telomerase activity are given for each SNP. The value of positive control HK293 (1000cells) is 1.900±0.063 and ranges from 1.810-1.991. Three thousand PBMC cells per sample were used for the TRAP reaction. Telomerase activity was expressed as the ratio of telomerase activity divided by HK293 telomerase activity (positive control).

Nature Genetics: doi:10.1038/ng.2528

Supplementary Table 7: Identification of non-synonymous SNPs within telomere-length associated loci.

Lead SNP	Chr	Non-synonymous SNPs	r^2 with lead SNP	Protein	AA substitution	PolyPhen functional prediction	SIFT prediction
rs10936599	3	rs10936600	1.0	LRRC34	L241I	Probably damaging	Tolerated
		rs6793295	0.92	LRRC34	S249G	Benign	Tolerated
rs2736100	5	None					
rs7675998	4	rs4691895	0.91	NAF1	L368V	Benign	Tolerated
		rs4691896	0.93	NAF1	I162V	Benign	Tolerated
rs9420907	10	rs2487999	0.72	OBFC1	T151A	Benign	Tolerated
	10	rs10786775	0.70	OBFC1	S248C	Benign	Tolerated
rs8105767	19	None					
rs755017	20	rs2281929	1.0	ZBTB46	T11P	Benign	Tolerated
rs11125529	2	None					

Coding variants within each loci were identified where r^2>0.7 to lead SNP. Functional predictions were carried out using PolyPhen 2[16]and SIFT[17].

Nature Genetics: doi:10.1038/ng.2528

Supplementary Table 8: Association of telomere length variants with diseases/traits.

Lead SNP	Chromosome and position	Relative SNP position	Genes in region with known function in telomere biology	Other genes in region	Association of lead SNP or proxy (r^2>0.7) with disease or associated phenotype (P<1x10^{-5})	Disease allele associated with longer or shorter telomere length
rs10936599	3: 170,974,795	Synonymous change in *MYNN*	*TERC*	*MYNN, ARPM1, LRRC34, LRRIQ4, LRRC31*	Colorectal cancer[18] Multiple sclerosis[19] Celiac disease[20]	Longer Longer Shorter
rs2736100	5: 1,339,516	Within intron of *TERT*	*TERT*	*SLC6A18, SLC6A19, CLPTM1L*	Lung cancer / adenocarcinoma[21-24] Glioma[25-27] Testicular germ cell cancer[28] Idiopathic pulmonary fibrosis[29] Higher red blood cell count[30]	Longer Shorter Shorter Shorter Longer
rs7675998	4: 164,227,270	Downstream of *NAF1*	*NAF1*	*NPY1R, NPY5R*		
rs9420907	10: 105,666,455	Within intron of *OBFC1*	*OBFC1*	*SLK, COL17A1, SH3PDX2A*		
rs8105767	19:22,007,281	Upstream of both *ZNF257* and *ZNF208*		*ZNF257 ,ZNF208 ZNF676, ZNF43 ZNF98*		
rs755017	20:61,892,066	Synonymous change in *ZBTB46*	*RTEL1*	*ZBTB46, LIME1, ZGPAT, TNFRSF6B, ARFRP1, STMN3, C20orf135 TPD52L2, DNAJC5*		
rs11125529	2: 54,329,370	Within intron of *ACYP2*		*TSPYL6, C2ORF73, PSME4, SPTBN1RPL23AP32*		

Details of the position of each lead SNP, RefSeq genes within each region and the presence of a gene with known function in telomere biology are stated. Searches of the NHGRI (http://www.genome.gov/gwastudies/) and GWAS Central (https://www.gwascentral.org/) catalogues (containing only the results of GWAS studies) were performed for lead SNPs and proxies (r^2>0.7) alongside literature searches to identify disease/phenotype associations with each locus.

Nature Genetics: doi:10.1038/ng.2528

Supplementary Note

Study Cohorts

The demographic characteristics of all study cohorts, for both discovery and replication phases are shown in **Supplementary Table 1**. All individuals included in the analysis are of European descent.

British Heart Foundation Family Heart Study (BHF-FHS) The BHF-FHS subjects are comprised of unrelated individuals who had a validated personal history of premature coronary artery disease (CAD) before the age of 66 years and a family history of CAD in a first degree relative. Recruitment took place in the UK between 1998 and 2003. Further details of this cohort are provided elsewhere[31,32].

Estonian Genome Center, University of Tartu (EGCUT) EGCUT is a population-based biobank of the Estonian Genome Project of University of Tartu (www.biobank.ee)[33]. The current cohort size is over 51,515, from 18 years of age and up, which reflects closely the age distribution in the adult Estonian population. The samples included in this study form a random subset of the cohort, with the exception of 500 female individuals aged 83+ which were specifically selected according to age and sex. Subjects are recruited by the general practitioners (GP) and physicians in the hospitals. Each participant filled out a Computer Assisted Personal interview, including personal data (place of birth, place(s) of living, nationality etc.), genealogical data (family history, three generations), educational and occupational history and lifestyle data (physical activity, dietary habits, smoking, alcohol consumption, women's health, quality of life). Anthropometric and physiological measurements were also taken.

Erasmus Rucphen Family (ERF) The ERF study is a cross-sectional cohort including 3,000 living descendants of 22 couples who had at least 6 children baptized in the community church around 1850-1900. The participants are not selected on any disease or other outcome. Details about the genealogy of the population have been described elsewhere[34,35].

FINRISK The National FINRISK Study 2007 Survey was carried out in five areas in Finland and 2000 inhabitants aged 25-74 years were invited to participate in each area[36,37]. The sample was a random sample from the Finnish Population Information System, stratified according to sex, 10-year age groups, and the six geographical areas. Only participants of the DILGOM sub-study from Helsinki were included in the analysis.

Nicotine Addiction Genetics – Finland study (FTC/NAG-FIN) The FTC/NAG-FIN sample was ascertained from the Finnish Twin Cohort study consisting of adult twins born between 1938 and 1957 (www.twinstudy.helsinki.fi). Based on earlier health questionnaires, the twin pairs concordant for ever smoking were identified and recruited along with their family members (mainly siblings) for the Nicotine Addiction Genetics Finland study (N = 2,265), as part of the consortium including Finland, Australia, and United States[38,39]. Data collection took place between 2001 and 2005.

GRAPHIC The GRAPHIC study comprises individuals from 520 white nuclear families of European descent recruited from the general population in Leicestershire UK, for the purpose of investigating the genetic determinants of blood pressure and related cardiovascular traits. Families were included if both parents aged 40-60 years and two offspring ≥18 years wished to participate. Families were

Nature Genetics: doi:10.1038/ng.2528

recruited through participating family practitioners in Leicestershire, UK, between 2003 and 2005. Further details are provided elsewhere[40].

Helsinki Birth Cohort Study (HBCS) HBCS is a birth cohort based study cohort and includes 8,760 subjects born in Helsinki between 1934 and 1944[41,42]. Between 2000 and 2003, a representative subset of 928 males and 1,075 females participated in a clinical study focusing upon cardiovascular and metabolic outcomes and cognitive function.

Cooperative Health Research in the Region of Augsburg (KORA) The KORA study is a series of independent population-based epidemiological surveys and follow-up studies of participants living in the region of Augsburg, Southern Germany[43]. All survey participants are of German nationality, identified through the registration office. Informed consent has been given by all participants. The present study includes data of the KORA F3 (2004/2005) survey which is a follow-up study of the KORA S3 survey (1994/1995), as well as data of the KORA F4 (2006-2008) study which is a follow-up study of the KORA S4 survey (1999-2001).

Leiden Longevity Study (LLS) LLS (http://www.molepi.nl) consists of offspring of nonagenarian sibling pairs of which the members are aged over 90 years. The partners of these offspring were recruited as population controls. The study has been designed to investigate biomarkers of healthy ageing and longevity[44].

Leiden 85-plus study (L85PLUS) L85P is a prospective study consisting of inhabitants of Leiden enrolled between 1997 and 1999in the month of their 85th birthday. Details about the study have been described elsewhere[45,46].

Northern Finland Birth Cohort 1966 (NFBC1966) NFBC1966 (http://kelo.oulu.fi/NFBC/) is a prospective follow-up population study of children from the two northernmost provinces of Finland[47]. Women with expected delivery dates in 1966 were recruited through maternity health centres[48]. Cohort members living in northern Finland or in the capital area were invited to a clinical examination at age 31 years.

Netherlands Twin Register (NTR) NTR (http://www.tweelingenregister.org/) recruits twins and their family members to study the causes of individual differences in health, behavior and lifestyle. Participants are followed longitudinally. Further details about the cohort can be found elsewhere[49-51]. Three subgroups from this population (NTRMRG3, NTR_DETECT and NTR_GODOT) were included in the discovery analysis and a further 2585 individuals from these subgroups had telomere length measurements performed for the replication phase.

PLIC The PLIC study is a prospective population based study designed to investigate the presence and progression of atherosclerotic lesions and intima media thickness in the common carotid artery in a local cohort of European descent origin (2141 subjects both men and women)[52-54]. The recruitment resulted from the collaboration with general practitioners, who enrolled the subjects referring to their ambulatory. 1871 subjects with available telomere length measurements were genotyped for the replication study.

Prevention of REnal and Vascular ENdstage Disease (PREVEND) Study PREVEND (www.prevend.org), is an ongoing prospective study investigating the natural course of increased levels of urinary albumin excretion and its relation to renal and cardiovascular disease. Inhabitants 28 to 75 years of

Nature Genetics: doi:10.1038/ng.2528

age in the city of Groningen, Netherlands were asked to complete a short questionnaire and individuals were then selected with a urinary albumin concentration of at least 10 mg/L and a randomly selected control group with a urinary albumin concentration less than 10 mg/L. Further details of the recruitment protocol and subjects are given elsewhere[55,56]. 2926 subjects from the PREVEND study were included in the discovery stage of this analysis, having both telomere length measurements and genotype data available. A further 2878 subjects with available telomere length measurements were genotyped for the replication study.

Queensland Institute of Medical Research (QIMR) Brisbane Adolescent Twin Study Subjects were recruited from the general population, in the context of ongoing studies of melanoma risk factors and studies of cognition[57].Twins and their singleton siblings were enlisted by contacting the principals of primary schools in the greater Brisbane area, by media appeals, and by word of mouth. It is estimated that approximately 50% of the eligible birth cohort were recruited into the study, which began in 1992[58].Most (98% by self-report) are of mixed European ancestry, mainly from British Isles. The participants are not selected on any disease or other outcome. Twins and siblings are evaluated for melanoma risk factors at ages twelve and fourteen, and for cognitive variables at age sixteen. Blood samples were collected at the end of testing sessions from participant twins and siblings and if possible from their parents. Pedigree relationships and zygosity were confirmed by genotype data.

TwinGene The TwinGene project is part of the Swedish Twin Registry (STR). Twins born before 1958 were contacted to participate in a simple health check-up, with measurement of height, weight, waist and hip circumference and blood pressure. Health and medication data were collected from self-reported questionnaires, and blood sampling materials were mailed to the subjects who then went to a local health care center for blood sampling for subsequent DNA extraction, serum collection and clinical chemistry tests. For the purpose of this study a subset of 300 female MZ twin pairs (600 individuals) were used for telomere length assessments. One member of each pair had been genotyped using the Illumina 317K SNP platform.

TwinsUK The TwinsUK cohort (www.twinsuk.ac.uk) is an adult twin British registry shown to be representative of singleton populations and the United Kingdom population[59]. A total of 4,899 subjects with telomere length measurement were included in the analysis. The age range of the cohort was 16-99. The design and methodology of the GWA study for TwinsUK is described in detail elsewhere[60].

United Kingdom Blood Service (UKBS) Subjects were recruited from healthy blood donors of European descent as part of the Wellcome Trust Case Control Consortium (WTCCC) study between 2005 and 2006[61]. The UKBS controls had a wide age range (between 17-69 years) with the majority of subjects between 40-59 years. Apart from sex and age, other phenotypic information was not available on the UKBS controls.

Telomerase activity cohort Two hundred and eight subjects (39.4% of European descent, 44.2% males, aged 19-57 years) were selected from two on-going studies[62,63]. Ethnicity was determined by self-reports of each subject. Subjects were overtly healthy, free of any acute or chronic illness on the basis of parental reports and were taking no medication that could influence the results.

Nature Genetics: doi:10.1038/ng.2528

Ethical approval for all studies was obtained from local ethics committees and all participants provided informed consent.

Additional details on methods for telomere length measurements

Leukocyte telomere length (LTL) measurements were made in five laboratories (Leicester, Helsinki, Imperial College London, Groningen and UCL London) each using a quantitative PCR assay comparing a TL PCR product (T) against a PCR product of a reference (S) gene to produce a T/S ratio, but with some modifications in relation to the reference gene and/or the calibrator samples or method used to enable inter-plate comparisons. Laboratory-specific details are given below and the laboratory is listed for each cohort in **Supplementary Table 1.** 32,446 samples (67% of total) were measured in Leicester, 3156 (6%) in Helsinki, 5146 (11%) in Imperial College, 5804 (12%) in Groningen and 1871 (4%) in UCL, London.

1. Leicester: LTL measurements were performed using a protocol described in detail elsewhere[64,65]. In brief, DNA samples were run in duplicate in 25µL reactions using a CAS-1200 liquid handling system (Qiagen, UK) and run on a Rotorgene-Q Real Time Thermal Cycler (Qiagen, UK). The single copy gene used was *36B4*. Alongside the samples, each run also contained a Calibrator sample (DNA from the K562 cell line) in duplicate and a no template control. Analysis of the PCR output was performed using Comparative Quantification (Qiagen Rotorgene analysis software, Qiagen, UK) and quantification is relative to the calibrator DNA. Samples were checked for concordance between duplicate measurements and to ensure that they ran within the established linear range of the assay. In addition, to ensure reproducibility of the assay, samples were re-run at random on different days. Inter-run coefficients of variation were between 2.7% and 3.9% for the cohorts measured using this method.

2.Helsinki: LTL measurements were performed using a protocol described in detail elsewhere[64,66]. Human beta-globin (Hgb) was used as the single copy reference gene. Samples were run in triplicate and quantification of both telomere and single copy gene amounts was made by absolute quantification against a 7-point standard curve included on each plate. In the case of Finrisk and FTC/NAG-FIN cohorts, genomic DNA was used for the standard curve while a synthetic oligomer (Sigma) dilution series[67] was used in HBCS. Samples and standard dilutions were transferred into the plates using a DNA Hydra 96 robot. Reactions were performed with CFX384 Real-Time PCR Detection System (Bio-Rad). Quality control was performed using the Bio-Rad CFX Manager software and samples with standard deviation of >0.5 between triplicates were omitted from the analysis. Plate effect was taken into account by analyzing four (HBCS) or five (Finrisk, FTC/NAG-FIN) genomic DNA control samples on every plate. The telomere and HgB signal values were normalized separately to the mean of these control samples before taking the T/S ratio. The control samples were used for calculating the inter-run coefficient of variation (CV) values and they were 24.8%, 7.7%, and 8.2% in the HBCS (n=1,582), Finrisk (n=520), and FTC/NAG-FIN (n=1,054) samples, respectively.

3. Imperial College London: LTL measurements were performed using a multiplex quantitative real-time PCR method[68], with minor modifications as described previously[69]. Human beta-globin (Hgb)was used as the single copy reference gene.Five serial dilutions of a single common reference sample (leukocyte DNA from a 42 year-old female) spanning 5-50ng were run in triplicate on each

Nature Genetics: doi:10.1038/ng.2528

plate. Any samples found to have an input DNA amount outside of this range were diluted and run again. The overall mean coefficient of variation for duplicate test samples on the same plate was 5%, and the mean inter-run CV for selected samples was 6.2% for the 5,146 samples measured.

4. Groningen:LTL was measured using a multiplex quantitative real-time PCR method described in detail elsewhere[65,68,70]. All experimental DNA samples were assayed in triplicate on different plates but in the same well position. Alongside the samples, all assay plates contained a no template control (NTC) and two standard DNA samples. The standard DNA samples consisted of a human control sample and genomic DNA of a human leukemia cell line (1301) with extreme long telomeres (kindly provided by Dr. Cesaro, IST, Genova). Albumin was used as the single-copy gene for these assays. Samples were run in 10ul reactions using a Bio-Rad CFX384 real-time system on a C1000 thermal cycler. For quantification, each plate also contained a standard curve, consisting of seven serial dilutions of a reference DNA sample (standard) spanning a ~12-fold range (5.2 to 60 ng). Analysis was performed using the Bio-Rad CFX manager software.For the reference DNA sample, each DNA concentration the Ct for albumin occurred ~ 7.2 cycles later in cycling than the Ct for the telomere. For quality control all samples were checked for concordance between triplicate values. Samples with a coefficient of variation (CV) of ≥ 10% within the triplicate were re-run. If the CV remained ≥ 10% the sample was omitted from the statistical analyses. Samples were run in triplicate and the intra-assay coefficient of variation was 2.0% (T), 1.85% (S) and 4.5% (T/S ratio). Reproducibility data was obtained for 216 subjects from PREVEND and good agreement between T/S ratios, measured on different days, was observed (r^2=0.99, P<0.0001, inter-run CV 3.9%). A total of 5,804 samples were measured in Groningen.

5.UCL London: LTL measurements of PLIC samples were performed using a multiplex quantitative real-time PCR method[68]. LTL measurements were run on a Rotor-Gene 6000(Corbett Research Ltd, Cambridge, UK) with two common reference sample on each run. All samples, including the reference sample were run in triplicate. Concordance between triplicate measures was assessed. Replication was also carried out on subsequent days of random samples to ensure reproducibility. Inter-run coefficients of variation were <5.0% for the 1,871 samples measured.

Locus by locus results of bioinformatics analyses

In order to gain better functional insight into the associated loci we undertook bioinformatics analyses to search for evidence that an associated SNP could alter protein function or gene expression. For all analyses we tested the lead SNP at a locus and all SNPs with an r^2>0.7 to the lead SNP identified through the 1000 Genomes study (http://www.1000genomes.org/).

Functional predictions of any identified coding variants were carried out using PolyPhen2[16] and SIFT[17] (**Supplementary Table 7**). In order to assess whether any variants influenced gene expression we searched two available genome-wide gene expression databases, the monocyte genome-wide gene expression data from the Gutenburg Heart Study[71] and the Genotype-Tissue Expression Project (GTEx) data base (http://www.genome.gov/gtex/), which includes liver, brain and lymphoblastoid cell types. Potential regulatory variants were identified by searching ENCODE data in the UCSC Genome Browser database (http://genome.ucsc.edu/)[72] to examine SNP locations in relation to

Nature Genetics: doi:10.1038/ng.2528

promoter, enhancer or insulator regions (Chromatin State Segmentation), methylation sites (predicted CpG islands, Methyl 450K Bead array data and Bisufite sequencing), conserved elements, conserved transcription factor binding sites and regions of known transcription factor binding (transcription factor ChIP-seq). The results seen at each locus are summarised below. Details of the key genes in each locus are given in more detail in **Supplementary Table 5**.

Chr 2p16.2 (*ACYP2*)

The lead SNP, rs11125529, along with all identified high LD SNPs, is located within the large intron 3 of *ACYP2*. A single exon gene, *TSPYL6,* is also located within intron 3 and the variants with most evidence of potential function are located *TSPYL6*. The most notable of these are rs6740641 and rs10165485 which are located within a region annotated as an active promoter in K562 cells. rs6740641 (r^2=1.0 to the lead SNP) results in a synonymous change in *TYPYL6*. It is also located within a CpG island, falling between two methylated CpG sites. rs10165485 (r^2=1.0) is located within the 3' UTR and also within an insulator region in two further cell lines. CTCF has been found to bind this insulator (transcription factor ChIP-seq). As insulators and enhancers work to regulate correct expression of neighbouring genes, it is possible that this region is important in controlling transcription of *TSPYL6/ACYP2*.

Chr 3q26 (*TERC*)

The lead SNP on 3q26, rs10936599, results in a synonymous change in MYNN (*Myoneurin*) a transcription factor which is a member of both the BTB/POZ and zinc finger domain-containing family (also referred to as *ZBTB31* and *ZNF902*).Chromatin state in this region shows this SNP to lie within an enhancer/promoter and it is also within a vertebrate conserved element. rs3821383 (r^2=0.92 to the lead SNP) is located within the promoter of MYNN with many transcription factors binding across the site.

Another SNP in the region with perfect LD to the lead SNP (rs2293607, r^2=1.0) is positioned 63bp 3' to the *TERC* RNA template sequence and lies within an active promoter, with evidence of many factors binding across the site. Although not located within a CpG island this SNP lies only a few base pairs outside of one that shows differential methylation across cell types. A previous study also reported this SNP to have a potential role in *TERC* RNA stability[73].

Two further SNPs in this locus, rs10936600 (r^2=1.0 to the lead SNP) and rs6793295 (r^2=0.92) lead to L241I and S249G changes within LRRC34 respectively. These were investigated using functional prediction in both PolyPhen2 and SIFT. Although PolyPhen2 suggests a damaging effect of L241I, SIFT suggests that this mutation would be tolerated (**Supplementary Table 7**). Both functional prediction tools suggest S249G to be of no consequence. In addition to rs10936600 being a non-synonymous SNP, it also lies within a conserved transcription factor binding site for HLF and CEBP. However, chromatin state within six cell lines does not suggest either promoter or enhancer activity and there is no evidence of transcription factor binding activity. Another SNP (rs67795055, r^2=0.79) within intron 1 of LRRC34 is located in a region of weak promoter activity and within a CpG island. Little is known about the function of LRRC34, although expression of this gene is down-regulated in mouse cells carrying a knockout of the tumour suppressor E-cadherin[74].

Nature Genetics: doi:10.1038/ng.2528

Three SNPs in high LD with the lead SNP lie within the 5' UTR and active promoter of *ACTRT3 (ARPM1* rs12637184, r^2=1.0; rs9811216, r^2=0.92 and rs9866776, r^2=0.92) with several transcription factors binding across all of these SNP sites. All three SNPs are within an annotated CpG island and lie within (rs9811216) or close to (rs12637184 and rs9866776) differentially methylated CpG sites.

Chr 4q32.2 (*NAF1*)

The lead SNP within this locus, rs7675998, is located 40.3 kb upstream of *NAF1* (*Nuclear assembly factor 1*). There is little evidence to suggest a putative function for this SNP. Two SNPs within high LD of rs7675998 (r^2=0.91) and in perfect LD to each other (rs4691895 and rs4691896, r^2=1.0) cause non-synonymous changes within NAF1 (**Supplementary Table 7**). Although neither, individually, are predicted to have a functional effect it is not known what the effect of these changes would be in combination, although it should be noted that rs4691895 is only present within the protein produced from transcript variant 2.

Other SNPs in potential regulatory regions include rs2320333 (r^2=0.77) which falls within a predicted insulator and rs936562 (r^2=0.82) which is located within the 5' UTR and promoter of *NAF1*, within a CpG island and in a region binding many transcription factors. The CpG island is unmethylated in the cell types studied, consistent with all cell types showing this region to be an active promoter.

Chr 5p15.33 (*TERT*)

rs2736100 is located within intron 2 of the telomerase reverse transcriptase gene, *TERT*. There are no high LD SNPs (r^2>0.7) to rs2736100. Two further SNPs on low LD with the lead SNP also associate with telomere length and were investigated in the conditional analyses (**Supplementary Table 4**). rs2853676 (r^2=0.17 to rs2736100 and association with LTL P=1.11x10^{-13}) is also located within intron 2 of *TERT*. This intron contains multiple CpG sites and both rs2736100 and rs2853676 lie in close proximity to such sites (<100bp) but not within them. The third SNP, rs1801075 (r^2=0.00 to rs2736100 and association to LTL P=3.70x10^{-8}) falls within the 5' UTR of the adjacent gene, *CLPTM1L*.

Chr 10q24.33 (*OBFC1*)

The lead SNP (rs9420907) is located within intron 1 of *OBFC1* and within a promoter/enhancer region. ChIP-seq data shows many factors binding across the region but not across the SNP location. Two further SNPs (rs1265164, r^2=0.90 and rs9419958, r^2=0.75 to rs9420907) are located within predicted enhancers in this region, but again there is no evidence of protein binding across the SNP sites. Another SNP in LD the lead SNP (rs4387287, r^2=0.79) falls within the 5' UTR, a CpG island and the promoter region of *OBFC1*. This region is shown to bind multiple transcription factors.

Chr 19p12 (*ZNF208*)

The lead SNP in this locus (rs8105767) lies intergenic between two genes, *ZNF208* and *ZNF257*, which are transcribed in opposite directions. It is not located within the promoter region of either. All associated SNPs at this locus fall within regions annotated as heterochomatic in the cell lines studied in ENCODE and there is little evidence for functional activity. One cluster of SNPs with

Nature Genetics: doi:10.1038/ng.2528

r^2=0.71 to the lead SNP and in perfect LD with each other fall within vertebrate conserved elements, but show no other evidence of function.

Chr 20q13.3 (*RTEL1*)

Both the lead SNP (rs755017) and a SNP in perfect LD (rs2881929, r^2=1.0) are located within exon 1 of *ZBTB46*. Whereas rs755017 results in a synonymous change, rs2881929 causes a T11P substitution. T11P is not predicted to have a damaging effect on ZBTB46 (**Supplementary Table 7**). Both SNPs also lie in a CpG island and within vertebrate conserved elements. Although chromatin state in this area does not predict promoter or enhancer function, there is some evidence of DNAseI hypersensitivity. Relaxation of stringency within the conserved transcription factor binding element search suggests both STAT4 and OCT_C binding sites. However, due to the relaxed stringency this could be considered speculative and would require experimental validation.

A group of five further SNPs in high LD to the lead SNP (r^2=0.89) and in high LD with each other (r^2=1.0) in this region all fall in potential regulatory regions. Three (rs67416152, rs6011138 and rs6011139) lie within an insulator region whilst two others (rs6011173 and rs4809367) lie within a poised promoter. rs4809367 is also within a CpG island that shows differential methylation patterns across the cell lines studied.

Conditional analysis was performed within this region as several SNPs in low LD to the lead SNP also showed association with telomere length (**Online Methods**, **Supplementary Table 4**). The SNP used for the conditional analysis rs6011040 (r^2=0.19 to the lead SNP, association with LTL P=1.58x10^{-7}) showed an eQTL with *LIME1* expression in monocytes. This SNP is located within intron 3 of *ARFRP1* (29.6KB from *LIME1*) in an enhancer region. *LIME1* is a transmembrane adaptor protein involved in T- and B-cell activation that lies within the locus[75,76]. As activation of T- and B-cells results in proliferation, it is possible that altered *LIME1* expression could affect cellular turnover, which could in turn be reflected by LTL.

Nature Genetics: doi:10.1038/ng.2528

The CARDIoGRAM Consortium – Full list of members and affiliations

Executive Committee: Sekar Kathiresan[1,2,3], Muredach P. Reilly[4], Nilesh J. Samani[5,6], Heribert Schunkert[7], Jeanette Erdmann[7]

Steering Committee: Themistocles L. Assimes[8], Eric Boerwinkle[9], Jeanette Erdmann[7], Alistair Hall[10], Christian Hengstenberg[11], Sekar Kathiresan[1,2,3], Inke R. König[12], Reijo Laaksonen[13], Ruth McPherson[14], Muredach P. Reilly[4], Nilesh J. Samani[5,6], Heribert Schunkert[7], John R. Thompson[15], Unnur Thorsteinsdottir[16,17], Andreas Ziegler[12]

Statisticians: Inke R. König[12] (chair), John R. Thompson[15] (chair), Devin Absher[18], Li Chen[19], L. Adrienne Cupples[20,21], Eran Halperin[22], Mingyao Li[23], Kiran Musunuru[1,2,3], Michael Preuss[12,7], Arne Schillert[12], Gudmar Thorleifsson[16], Benjamin F. Voight[2,3,24], George A. Wells[25]

Writing group: Themistocles L. Assimes[8], Panos Deloukas[26], Jeanette Erdmann[7], Hilma Holm[16], Sekar Kathiresan[1,2,3], Inke R. König[12], Ruth McPherson[14], Muredach P. Reilly[4], Robert Roberts[14], Nilesh J. Samani[5,6], Heribert Schunkert[7], Alexandre F. R. Stewart[14]

ADVANCE: Devin Absher[18], Themistocles L. Assimes[8], Stephen Fortmann[8], Alan Go[27], Mark Hlatky[8], Carlos Iribarren[27], Joshua Knowles[8], Richard Myers[18], Thomas Quertermous[8], Steven Sidney[27], Neil Risch[28], Hua Tang[29]

CADomics: Stefan Blankenberg[30], Tanja Zeller[30], Arne Schillert[12], Philipp Wild[30], Andreas Ziegler[12], Renate Schnabel[30], Christoph Sinning[30], Karl Lackner[31], Laurence Tiret[32], Viviane Nicaud[32], Francois Cambien[32], Christoph Bickel[30], Hans J. Rupprecht[30], Claire Perret[32], Carole Proust[32], Thomas Münzel[30]

CHARGE: Maja Barbalic[33], Joshua Bis[34], Eric Boerwinkle[9], Ida Yii-Der Chen[35], L. Adrienne Cupples[20,21], Abbas Dehghan[36], Serkalem Demissie-Banjaw[37,21], Aaron Folsom[38], Nicole Glazer[39], Vilmundur Gudnason[40,41], Tamara Harris[42], Susan Heckbert[43], Daniel Levy[21], Thomas Lumley[44], Kristin Marciante[45], Alanna Morrison[46], Christopher J. O´Donnell[47], Bruce M. Psaty[48], Kenneth Rice[49], Jerome I. Rotter[35], David S. Siscovick[50], Nicholas Smith[43], Albert Smith[40,41], Kent D. Taylor[35], Cornelia van Duijn[36], Kelly Volcik[46], Jaqueline Whitteman[36], Vasan Ramachandran[51], Albert Hofman[36], Andre Uitterlinden[52,36]

deCODE: Solveig Gretarsdottir[16], Jeffrey R. Gulcher[16], Hilma Holm[16], Augustine Kong[16], Kari Stefansson[16,17], Gudmundur Thorgeirsson[53,17], Karl Andersen[53,17], Gudmar Thorleifsson[16], Unnur Thorsteinsdottir[16,17]

GERMIFS I and II: Jeanette Erdmann[7], Marcus Fischer[11], Anika Grosshennig[12,7], Christian Hengstenberg[11], Inke R. König[12], Wolfgang Lieb[54], Patrick Linsel-Nitschke[7], Michael Preuss[12,7], Klaus Stark[11], Stefan Schreiber[55], H.-Erich Wichmann[56,58,59], Andreas Ziegler[12], Heribert Schunkert[7]

GERMIFS III (KORA): Zouhair Aherrahrou[7], Petra Bruse[7], Angela Doering[56], Jeanette Erdmann[7], Christian Hengstenberg[11], Thomas Illig[56], Norman Klopp[56], Inke R. König[12], Patrick Linsel-Nitschke[7], Christina Loley[12,7], Anja Medack[7], Christina Meisinger[56], Thomas Meitinger[57,60], Janja Nahrstedt[12,7], Annette Peters[56], Michael Preuss[12,7], Klaus Stark[11], Arnika K. Wagner[7], H.-Erich Wichmann[56,58,59], Christina Willenborg[12,7], Andreas Ziegler[12], Heribert Schunkert[7]

Nature Genetics: doi:10.1038/ng.2528

LURIC/AtheroRemo: Bernhard O. Böhm[61], Harald Dobnig[62], Tanja B. Grammer[63], Eran Halperin[22], Michael M. Hoffmann[64], Marcus Kleber[65], Reijo Laaksonen[13], Winfried März[63,66,67], Andreas Meinitzer[66], Bernhard R. Winkelmann[68], Stefan Pilz[62], Wilfried Renner[66], Hubert Scharnagl[66], Tatjana Stojakovic[66], Andreas Tomaschitz[62], Karl Winkler[64]

MIGen: Benjamin F. Voight[2,3,24], Kiran Musunuru[1,2,3], Candace Guiducci[3], Noel Burtt[3], Stacey B. Gabriel[3], David S. Siscovick[50], Christopher J. O'Donnell[47], Roberto Elosua[69], Leena Peltonen[49], Veikko Salomaa[70], Stephen M. Schwartz[50], Olle Melander[26], David Altshuler[71,3], Sekar Kathiresan[1,2,3]

OHGS: Alexandre F. R. Stewart[14], Li Chen[19], Sonny Dandona[14], George A. Wells[25], Olga Jarinova[14], Ruth McPherson[14], Robert Roberts[14]

PennCATH/MedStar: Muredach P. Reilly[4], Mingyao Li[23], Liming Qu[23], Robert Wilensky[4], William Matthai[4], Hakon H. Hakonarson[72], Joe Devaney[73], Mary Susan Burnett[73], Augusto D. Pichard[73], Kenneth M. Kent[73], Lowell Satler[73], Joseph M. Lindsay[73], Ron Waksman[73], Christopher W. Knouff[74], Dawn M. Waterworth[74], Max C. Walker[74], Vincent Mooser[74], Stephen E. Epstein[73], Daniel J. Rader[75,4]

WTCCC: Nilesh J. Samani[5,6], John R. Thompson[15], Peter S. Braund[5], Christopher P. Nelson[5], Benjamin J. Wright[76], Anthony J. Balmforth[77], Stephen G. Ball[78], Alistair S. Hall[10], Wellcome Trust Case Control Consortium

1 Cardiovascular Research Center and Cardiology Division, Massachusetts General Hospital, Boston, MA, USA; 2 Center for Human Genetic Research, Massachusetts General Hospital, Boston, MA, USA; 3 Program in Medical and Population Genetics, Broad Institute of Harvard and Massachusetts Institute of Technology (MIT), Cambridge, MA, USA; 4 The Cardiovascular Institute, University of Pennsylvania, Philadelphia, PA, USA; 5 Department of Cardiovascular Sciences, University of Leicester, Glenfield Hospital, Leicester, UK; 6 Leicester National Institute for Health Research Biomedical Research Unit in Cardiovascular Disease, Glenfield Hospital, Leicester, LE3 9QP, UK; 7 Medizinische Klinik II, Universität zu Lübeck, Lübeck, Germany; 8 Department of Medicine, Stanford University School of Medicine, Stanford, CA, USA; 9 University of Texas Health Science Center, Human Genetics Center and Institute of Molecular Medicine, Houston, TX, USA; 10 Division of Cardiovascular and Neuronal Remodelling, Multidisciplinary Cardiovascular Research Centre, Leeds Institute of Genetics, Health and Therapeutics, University of Leeds, UK; 11 Klinik und Poliklinik für Innere Medizin II, Universität Regensburg, Regensburg, Germany; 12 Institut für Medizinische Biometrie und Statistik, Universität zu Lübeck, Lübeck, Germany; 13 Science Center, Tampere University Hospital, Tampere, Finland; 14 The John & Jennifer Ruddy Canadian Cardiovascular Genetics Centre, University of Ottawa Heart Institute, Ottawa, Canada; 15 Department of Health Sciences, University of Leicester, Leicester, UK; 16 deCODE Genetics, 101 Reykjavik, Iceland; 17 University of Iceland, Faculty of Medicine, 101 Reykjavik, Iceland; 18 Hudson Alpha Institute, Huntsville, Alabama, USA; 19 Cardiovascular Research Methods Centre, University of Ottawa Heart Institute, 40 Ruskin Street, Ottawa, Ontario, Canada, K1Y 4W7; 20 Department of Biostatistics, Boston University School of Public Health, Boston, MA USA; 21 National Heart, Lung and Blood Institute's Framingham Heart Study, Framingham, MA, USA; 22 The Blavatnik School of Computer Science and the Department of Molecular Microbiology and Biotechnology, Tel-Aviv University, Tel-Aviv, Israel, and the International Computer Science Institute, Berkeley, CA, USA; 23 Biostatistics and Epidemiology, University of Pennsylvania, Philadelphia, PA, USA; 24 Department of Medicine, Harvard Medical School, Boston, MA, USA; 25 Research Methods, Univ Ottawa Heart Inst; 26

Nature Genetics: doi:10.1038/ng.2528

Department of Clinical Sciences, Hypertension and Cardiovascular Diseases, Scania University Hospital, Lund University, Malmö, Sweden; 27 Division of Research, Kaiser Permanente, Oakland, CA, USA; 28 Institute for Human Genetics, University of California, San Francisco, San Francisco, CA, USA; 29 Dept Cardiovascular Medicine, Cleveland Clinic; 30 Medizinische Klinik und Poliklinik, Johannes-Gutenberg Universität Mainz, Universitätsmedizin, Mainz, Germany; 31 Institut für Klinische Chemie und Laboratoriumsmediizin, Johannes-Gutenberg Universität Mainz, Universitätsmedizin, Mainz, Germany; 32 INSERM UMRS 937, Pierre and Marie Curie University (UPMC, Paris 6) and Medical School, Paris, France; 33 University of Texas Health Science Center, Human Genetics Center, Houston, TX, USA; 34 Cardiovascular Health Resarch Unit and Department of Medicine, University of Washington, Seattle, WA USA; 35 Cedars-Sinai Medical Center, Medical Genetics Institute, Los Angeles, CA, USA; 36 Erasmus Medical Center, Department of Epidemiology, Rotterdam, The Netherlands; 37 Boston University, School of Public Health, Boston, MA, USA; 38 University of Minnesota School of Public Health, Division of Epidemiology and Community Health, School of Public Health (A.R.F.), Minneapolis, MN, USA; 39 University of Washington, Cardiovascular Health Research Unit and Department of Medicine, Seattle, WA, USA; 40 Icelandic Heart Association, Kopavogur Iceland; 41 University of Iceland, Reykjavik, Iceland; 42 Laboratory of Epidemiology, Demography, and Biometry, Intramural Research Program, National Institute on Aging, National Institutes of Health, Bethesda MD, USA; 43 University of Washington, Department of Epidemiology, Seattle, WA, USA; 44 University of Washington, Department of Biostatistics, Seattle, WA, USA; 45 University of Washington, Department of Internal Medicine, Seattle, WA, USA; 46 University of Texas, School of Public Health, Houston, TX, USA; 47 National Heart, Lung and Blood Institute, Framingham Heart Study, Framingham, MA and Cardiology Division, Massachusetts General Hospital, Boston, MA, USA; 48 Center for Health Studies, Group Health, Departments of Medicine, Epidemiology, and Health Services, Seattle, WA, USA; 49 The Wellcome Trust Sanger Institute, The Wellcome Trust Genome Campus, Hinxton, Cambridge, UK; 50 Cardiovascular Health Research Unit, Departments of Medicine and Epidemiology, University of Washington, Seattle; 51 Boston University Medical Center, Boston, MA, USA; 52 Department of Internal Medicine, Erasmus Medical Center, Rotterdam, The Netherlands; 53 Department of Medicine, Landspitali University Hospital, 101 Reykjavik, Iceland; 54 Boston University School of Medicine, Framingham Heart Study, Framingham, MA, USA; 55 Institut für Klinische Molekularbiologie, Christian-Albrechts Universität, Kiel, Germany; 56 Institute of Epidemiology, Helmholtz Zentrum München – German Research Center for Environmental Health, Neuherberg, Germany; 57 Institut für Humangenetik, Helmholtz Zentrum München, Deutsches Forschungszentrum für Umwelt und Gesundheit, Neuherberg, Germany; 58 Institute of Medical Information Science, Biometry and Epidemiology, Ludwig-Maximilians-Universität München, Germany; 59 Klinikum Grosshadern, Munich, Germany; 60 Institut für Humangenetik, Technische Universität München, Germany; 61 Division of Endocrinology and Diabetes, Graduate School of Molecular Endocrinology and Diabetes, University of Ulm, Ulm, Germany; 62 Division of Endocrinology, Department of Medicine, Medical University of Graz, Austria; 63 Synlab Center of Laboratory Diagnostics Heidelberg, Heidelberg, Germany; 64 Division of Clinical Chemistry, Department of Medicine, Albert Ludwigs University, Freiburg, Germany; 65 LURIC non profit LLC, Freiburg, Germany; 66 Clinical Institute of Medical and Chemical Laboratory Diagnostics, Medical University Graz, Austria; 67 Institute of Public Health, Social and Preventive Medicine, Medical Faculty Manneim, University of Heidelberg, Germany; 68 Cardiology Group Frankfurt-Sachsenhausen, Frankfurt, Germany; 69 Cardiovascular Epidemiology and Genetics Group, Institut Municipal d'Investigació Mèdica, Barcelona; Ciber Epidemiología y Salud Pública (CIBERSP), Spain; 70

Nature Genetics: doi:10.1038/ng.2528

Chronic Disease Epidemiology and Prevention Unit, Department of Chronic Disease Prevention, National Institute for Health and Welfare, Helsinki, Finland; 71 Department of Molecular Biology and Center for Human Genetic Research, Massachusetts General Hospital, Harvard Medical School, Boston, USA; 72 The Center for Applied Genomics, Children's Hospital of Philadelphia, Philadelphia, Pennsylvania, USA; 73 Cardiovascular Research Institute, Medstar Health Research Institute, Washington Hospital Center, Washington, DC 20010, USA; 74 Genetics Division and Drug Discovery, GlaxoSmithKline, King of Prussia, Pennsylvania 19406, USA; 75 The Institute for Translational Medicine and Therapeutics, School of Medicine, University of Pennsylvania, Philadelphia, PA, USA; 76 Department of Cardiovascular Surgery, University of Leicester, Leicester, UK; 77 Division of Cardiovascular and Diabetes Research, Multidisciplinary Cardiovascular Research Centre, Leeds Institute of Genetics, Health and Therapeutics, University of Leeds, Leeds, LS2 9JT, UK; 78 LIGHT Research Institute, Faculty of Medicine and Health, University of Leeds, Leeds, UK

Nature Genetics: doi:10.1038/ng.2528

Acknowledgments

BHF-FHS. The British Heart Foundation Family Heart Study (BHF-FHS) was funded by the British Heart Foundation (BHF). Genotyping of the BHF-FHS study was undertaken as part of the WTCCC and funded by the Wellcome Trust. N.J.S holds a chair funded by the BHF and V.C is funded by the BHF. V.C., C.P.N. and N.J.S. are supported by the NIHR Leicester Cardiovascular Biomedical Research Unit and N.J.S. holds an NIHR Senior Investigator award.

EGCUT. EGCUT received financing from FP7 programs (ENGAGE, OPENGENE), targeted financing from Estonian Government SF0180142s08, Estonian Research Roadmap through Estonian Ministry of Education and Research, Center of Excellence in Genomics (EXCEGEN) and University of Tartu (SP1GVARENG). The work of Krista Fischer and Reedik Mägi has also been supported by Estonian Science Foundation grant EstSF ETF9353. We acknowledge EGCUT technical personnel, especially Mr V. Soo and S. Smit. Data analyses were carried out in part in the High Performance Computing Center of University of Tartu.

ERF. The ERF study was supported by grants from The Netherlands Organisation for Scientific Research (NWO), Erasmus MC, the Centre for Medical Systems Biology (CMSB) and the European Community's Seventh Framework Programme (FP7/2007-2013), ENGAGE Consortium, grant agreement HEALTH-F4-2007- 201413. We are grateful to all general practitioners for their contributions, to Petra Veraart for her help in genealogy, Jeannette Vergeer for the supervision of the laboratory work and Peter Snijders for his help in data collection.

Finnish Twin Cohort/Nicotine Addiction Genetics-Finland study. This work was supported for data collection by Academy of Finland grants (J Kaprio) and a NIH grant DA12854 (PAFM). Genome-wide genotyping in the Finnish sample was funded by Global Research Award for Nicotine Dependence / Pfizer Inc. (JK), and Wellcome Trust Sanger Institute, UK. This work was further supported by the Academy of Finland Center of Excellence in Complex Disease Genetics (grant numbers: 213506, 129680, J Kaprio) and the European Community's Seventh Framework Programme ENGAGE Consortium, (grant agreement HEALTH-F4-2007- 201413). We thank Sirli Raud and Anne Vikman for help in TL measurement.

Finrisk. The FINRISK 2007/ DILGOM-study was supported by the Academy of Finland, grant # 118065. S.M. was supported by grants #136895 and #141005 and V.S. by grants #139635 and 129494 from the Academy of Finland. We thank Jenni Lahtinen for help in TL measurement.

GRAPHIC. The GRAPHIC study was funded by the BHF. MDT is supported by an MRC Senior Clinical Fellowship (G0902313).

HBCS. The study was supported by grants to J.G.E. from the Academy of Finland (JGE: 135072, 129255, 126775), FinskaLäkaresällskapet, SamfundetFolkhälsan, JuhoVainio Foundation, Päivikki and SakariSohlberg Foundation, Signe and AneGyllenberg Foundation, and YrjöJahnsson Foundation.I.H. is supported by the Academy of Finland (140747 and 218017), Sigrid Jusélius Foundation, YrjöJahnsson Foundation, Signe and Ane Gyllenberg Foundation, and the Finnish Foundation for Alcohol Studies. S.R was supported by the Academy of Finland Center of Excellence in Complex Disease Genetics (213506 and 129680), Academy of Finland (251217), the Finnish foundation for Cardiovascular Research and the Sigrid Juselius Foundation

Nature Genetics: doi:10.1038/ng.2528

KORA. The KORA Augsburg studies were financed by the Helmholtz Zentrum München, German Research Center for Environmental Health, Neuherberg, Germany and supported by grants from the German Federal Ministry of Education and Research (BMBF). Part of this work was financed by the German National Genome Research Network (NGFN). Our research was supported within the Munich Center of Health Sciences (MC Health) as part of LMUinnovativ.

LLS The Leiden Longevity Study has received funding from the European Union's Seventh Framework Programme (FP7/2007-2011) under grant agreement n° 259679. This study was supported by a grant from the Innovation-Oriented Research Program on Genomics (Senter Novem IGE05007), the Centre for Medical Systems Biology, and the Netherlands Consortium for Healthy Ageing (grant 050-060-810), all in the framework of the Netherlands Genomics Initiative, Netherlands Organization for Scientific Research (NWO).

NFBC Northern Finland Birth Cohort 1966 (NFBC1966) NFBC1966 received financial support from the Academy of Finland (project grants 104781, 120315, 129269, 1114194, 139900/24300796, Center of Excellence in Complex Disease Genetics and SALVE), University Hospital Oulu, Biocenter, University of Oulu, Finland (75617), the European Commission (EURO-BLCS, Framework 5 award QLG1-CT-2000-01643), NHLBI grant 5R01HL087679-02 through the STAMPEED program (1RL1MH083268-01), NIH/NIMH (5R01MH63706:02), ENGAGE project and grant agreement HEALTH-F4-2007-201413, the Medical Research Council, UK (G0500539, G0600705, G0600331, PrevMetSyn/SALVE, PS0476) and the Wellcome Trust (project grant GR069224, WT089549), UK.Replication genotyping was supported in part by MRC grant G0601261, Wellcome Trust grants 085301, 090532 and 083270, and Diabetes UK grants RD08/0003704 and BDA 08/0003775.The DNA extractions, sample quality controls, biobank up-keeping and aliquotting was performed in the National Public Health Institute, Biomedicum Helsinki, Finland and supported financially by the Academy of Finland and Biocentrum Helsinki. We thank Professor (emerita) Paula Rantakallio (launch of NFBC1966 and 1986), and Ms Outi Tornwall and Ms Minttu Jussila (DNA biobanking). The authors would like to acknowledge the contribution of the late Academian of Science Leena Peltonen. JLB is supported by a Wellcome Trust fellowship grant (WT088431MA).

NTR. Funding was obtained from the Netherlands Organization for Scientific Research (NWO: MagW/ZonMW): Genotype/phenotype database for behaviour genetic and genetic epidemiological studies (NWO 911-09-032); Spinozapremie (SPI 56-464-14192); NBIC/BioAssist/RK/2008.024); BBMRI –NL: Biobanking and Biomolecular Resources Research Infrastructure (184.021.007); the Institute for Health and Care Research (EMGO+); Neuroscience Campus Amsterdam (NCA); European Science Foundation (ESF): Genomewide analyses of European twin and population cohorts (EU/QLRT-2001-01254); European Community's Seventh Framework Program (FP7/2007-2013): ENGAGE (HEALTH-F4-2007-201413); the European Science Council : Genetics of Mental Illness (ERC-230374); Rutgers University Cell and DNA Repository cooperative agreement (NIMH U24 MH068457-06); Collaborative study of the genetics of DZ twinning (NIH R01D0042157-01A); the Genetic Association Information Network, a public–private partnership between the NIH and Pfizer Inc., Affymetrix Inc. and Abbott Laboratories; and the National Institutes of Mental Health (NIMH): GODOT: Integration of Genomics & Transcriptomics in Normal Twins & Major Depression (1RC2 MH089951-01) and DETECT: Developmental TrajECtoriesin Twins (1RC2MH089995-01).

Nature Genetics: doi:10.1038/ng.2528

PLIC. The PLIC study is supported by the Italian Society for the Study of Atherosclerosis, Lombardia Chapter. ALC is supported by grants from Regione Lombardia and the Cariplo Foundation. J.P. and P.J.T. are supported by a grant from the British Heart Foundation (RG/08/014).

PREVEND. PREVEND genetics is supported by the Dutch Kidney Foundation (Grant E033), the National Institutes of Health (grant LM010098), The Netherlands organisation for health research and development (NWO VENI grant 916.761.70 & ZonMw grant 90.700.441), and the Dutch Inter University Cardiology Institute Netherlands (ICIN).

QIMR. We thank Marlene Grace, Ann Eldridge, and Kerrie McAloney for sample collection; Megan Campbell, Lisa Bowdler, Sara Smith, Steven Crooks, and staff of the Molecular Epidemiology Laboratory for sample processing and preparation; Harry Beeby, David Smyth, and Daniel Park for IT/database support; Scott Gordon for his substantial efforts involving the QC and preparation of the GWA data; and the twins and their families for their participation. We acknowledge support from the Australian Research Council (A7960034, A79906588, A79801419, DP0212016, and DP0343921). Telomere length assessment was co-funded by the European Community's Seventh Framework Programme (FP7/2007-2013), ENGAGE project, grant agreement HEALTH-F4-2007-201413 and National Health and Medical Research Council (NHMRC)-European Union Collaborative Research Grant 496739. GWM was supported by an NHMRC Fellowship (619667), DRN (FT0991022) and SEM (FT110100548) were supported by an ARC Future Fellowship. Genotyping was funded by the NHMRC (Medical Bioinformatics Genomics Proteomics Program, 389891). Genotype imputation was carried out on the Genetic Cluster Computer (http://www.geneticcluster.org), which is financially supported by the Netherlands Scientific Organization (NWO 480-05-003).

TWINGENE. The Ministry for Higher Education, the Swedish Research Council (M-2005-1112), GenomEUtwin (EU/QLRT-2001-01254; QLG2-CT-2002-01254), NIH DK U01-066134, The Swedish Foundation for Strategic Research (SSF) Heart and Lung foundation no. 20070481

TwinsUK. The study was funded by the Wellcome Trust; European Community's Seventh Framework Programme (FP7/2007-2013), ENGAGE project grant agreement (HEALTH-F4-2007-201413). The study also receives support from the Dept of Health via the National Institute for Health Research (NIHR) comprehensive Biomedical Research Centre award to Guy's & St Thomas' NHS Foundation Trust in partnership with King's College London. TDS is an NIHR senior Investigator and is holder of an ERC Advanced Principal Investigator award. Genotyping was performed by The Wellcome Trust Sanger Institute, support of the National Eye Institute via an NIH/CIDR genotyping project.

UKBS. The UK Blood Services collection of Common Controls (UKBS-CC collection) is supported by the Wellcome Trust grant 076113/C/04/Z and the collection was established as part of the Wellcome Trust Case Control Consortium. WHO is supported by a NIHR programme grant to NHSBT (RP-PG-0310-1002) and is a NIHR Senior Investigator.

Leiden 85-plus. The Leiden 85-plus Study was in part supported by an unrestricted grant from the Dutch Ministry of Health, Welfare and Sports. This current study was funded by the Netherlands Genomics Initiative/Netherlands Organization for scientific research (NGI/NWO; 05040202 and 050-060-810. NCHA) and the European Union's Seventh Framework Programme (FP7/2007-2011) under grant agreement n° 259679.

Nature Genetics: doi:10.1038/ng.2528

Telomerase activity cohort. This study was funded by the National Heart, Lung and Blood Institute HL69999 and Georgia Health Sciences University Diabetes and Obesity Discovery Institute.

Nature Genetics: doi:10.1038/ng.2528

Supplementary References

1. Kim, J.W., Kwon, O.Y., & Kim M.H. Differentially expressed genes and morphological changes during lengthened immobilization in rat soleus muscle. *Differentiation* **75**, 147-157 (2007)
2. Khor, B. et al., Proteasome activator PA200 is required for normal spermatogenesis. Mol. Cell Biol. 26, 2999-3007 (2006)
3. Hara, Y., Yamagata, K., Oguchi, K. and Baba, T. Nuclear localization of profilin III-ArpM1 complex in mouse spermiogenesis. *FEBS Lett.* **582**, 2998-3004 (2008).
4. Egan, E.D. & Collins, K. An enhanced H/ACA RNP assembly mechanism for human telomerase RNA. *Mol. Cell. Biol.32,* 2428-2439 (2012).
5. Mason, P.J. and Bessler, M. The genetics of dyskeratosis congenital. *Cancer Genet*. **204**, 635-645 (2011).
6. Miyake, Y. *et al*., RPA-like mammalian Ctc1-Stn1-Ten1 complex binds to single-stranded DNA and protects telomeres independently of the Pot1 pathway. *Mol. Cell* **36**, 193-206 (2009).
7. Wan, M., Qin, J., Songyang, Z., & Liu D. OB fold containing protein (OBFC1), a human homolog of yeast Stn1, associates with TPP1 and is implicated in telomere length regulation. *J. Biol. Chem.* **284**, 26725-26731 (2009).
8. Casteel, D.E., *et al*., A DNA polymerase-α primase cofactor with homology to replication protein A-32 regulates DNA replication in mammalian cells. *J. Biol. Chem.* **284**, 5807-5818 (2009).
9. Eichler, E.E., *et al*., Comples beta-satellite repeat structures and the expansion of the zinc finger gene cluster in 19p12. *Genome* Res. **8**, 791-818 (1998).
10. Reizis, B. Classical dendritic cells as a unique immune cell lineage. *J. Exp. Med.* **209**, 1053-1056 (2012).
11. Hur, E.M. *et al.* LIME, a novel transmembrane adaptor protein, associates with p56Ick and mediates T-cell activation. *J. Exp. Med*. **198**, 1463-1473 (2003).
12. Ahn, E., Lee, H., & Yun, Y. LIME acts as a transmembrane adaptor mediating BCR-dependent B-cell activation. *Blood* **107**, 1521-1527 (2006).
13. Ding, H. *et al.* Regulation of murine telomere length by Rtel1: an essential gene encoding a helicase-like protein. *Cell* **117**, 873-886 (2004).
14. Barber, L.J. *et al.* RTEL1 maintains genomic stability by suppressing homologous recombination. *Cell* **135**, 261-271 (2008).
15. Tso, P.H., Yung, L.Y., Wang, Y. and Wong, Y.H. RGS19 stimulates cell proliferation by deregulating cell cycle control and enhancing Akt signalling. *Cancer Lett*. **309**, 199-208 (2011).
16. Adzhubei, I.A. *et al.* A method and server for predicting damaging missense mutations. *Nat Methods* **7**, 248-249 (2010).
17. . Kumar, P., Henikoff, S. and Ng P.C. Predicting the effects of coding non-synonymous variants on protein function using the SIFT algorithm. *Nat Protoc.* **4**, 1073-81 (2009).
18. Houlston, R.S. *et al.* Meta-analysis of three genome-wide association studies identifies susceptibility loci for colorectal cancer at 1q41, 3q26.2, 12q13.13 and 20q13.33. *Nat. Genet*. **42**, 973-977 (2010)
19. Sawcer, S. *et al.* Genetic risk and a primary role for cell-mediated immune mechanisms in multiple sclerosis. *Nature* **476**, 214-219 (2011)
20. Dubois, P.C. *et al.* Multiple common variants for celiac disease influencing immune gene expression. *Nat. Genet.***42**, 295-302 (2010)
21. Landi, M.T. *et al.* A genome-wide association study of lung cancer identifies a region of chromosome 5p15 associated with risk of adenocarcinoma. *Am. J. Hum. Gene.***85**, 679-691 (2009)
22. Hsiung, C.A. *et al.*The 5p15.33 locus is associated with risk of lung adenocarcinoma in never-smoking females in Asia. *PLoS Genet.***6**, e1001051
23. Miki, D. *et al.* Variation in TP63 is associated with lung adenocarcinoma susceptibility in Japanese and Korean populations. *Nat. Genet.***42**, 893-896 (2010)

Nature Genetics: doi:10.1038/ng.2528

24. Hu, Z. *et al.* A genome-wide association study identifies two new lung cancer susceptibility loci at 13q12.12 and 22q12.2 in Han Chinese. *Nat. Genet.***43**, 792-796 (2011)
25. Shete, S. *et al.* Genome-wide association study identifies five susceptibility loci for glioma. *Nat. Genet.***41**, 899-904 (2009)
26. Sanson, M. et al. Chromosome 7p11.2 (EGFR) variation influences glioma risk. *Hum. Mol. Genet.***20**, 2897-904 (2011)
27. Yang, T.H., Kon, M., Hung, J.H., & Delisi, C. Combinations of newly formed Glioma-associated loci link regions on chromosomes 1 and 9 to increased disease risk. *BMC Med. Genomics* **4**, 63 (2011)
28. Turnbull, C. *et al.* Variants near DMRT1, TERT and ATF7IP are associated with testicular germ cell cancer. *Nat. Genet.* **42**, 604-607 (2010)
29. Mushiroda, T. *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.***45**, 654-656 (2008)
30. Kamatani, Y. *et al.* Genome-wide association study of haematological and biochemical traits in a Japanese population. *Nat. Genet.* **42**, 210-215 (2010)
31. BHF Family Heart Study Research Group. A Genomewide Linkage Study of 1,933 Families Affected by Premature Coronary Artery Disease: The British Heart Foundation (BHF) Family Heart Study. *Am J Hum Genet* **77**, 1011-1020 (2005)
32. Samani, N.J., et al. Genomewide Association Analysis of Coronary Artery Disease. *N Engl J Med* **357**, 443-453 (2007)
33. Nelis, M. et al. Genetic structure of Europeans: a view from the North-East. *PLoS One* **4**, e5472 (2009)
34. Aulchenko, Y.S. et al. Linkage disequilibrium in young genetically isolated Dutch population. *Eur J Hum Genet* **12**, 527-534 (2004)
35. Pardo, L.M. et al. The effect of genetic drift in a young genetically isolated population. *Ann Hum Genet* **69**, 288-295 (2005)
36. Peltonen, M. et al. Kansallinen FINRISKI 2007 -terveystutkimus, Helsinki 2008
37. Vrtiainen, E. *et al*. Thirty-five-year trends in cardiovascular risk factors in Finland. *Int. J. Epidemiol.***39**, 504-518 (2010)
38. Loukola, A. et al. Linkage of nicotine dependence and smoking behaviour on 10q, 7q and 11p in twins with homogeneous genetic background. *Pharmacogenomics J.***8**, 209-19 (2008)
39. Broms, U. et al. Analysis of Detailed Phenotype Profiles Reveals CHRNA5-CHRNA3-CHRNB4 Gene Cluster Association With Several Nicotine Dependence Traits. *Nicotine Tob Res.*14, 720-733 (2012)
40. Tobin, M.D., et al. Common variants in genes underlying monogenic hypertension and hypotension and blood pressure in the general population. *Hypertension.***51**, 1658-1664 (2008)
41. Yliharsila, H. et al. Body mass index during childhood and adult body composition in men and women aged 56-70 y. *Am J Clin Nutr* **87**, 1769-75 (2008)
42. Eriksson, J.G et al. Early growth and coronary artery disease in later life. *BMJ* **322**, 949-53 (2001)
43. Wichmann, H.E. KORA-gen--resource for population genetics, controls and a broad spectrum of disease phenotypes. *Gesundheitswesen* 67, S26-30 (2005)
44. Westendorp, R.G. et al. Nonagenarian siblings and their offspring display lower risk of mortality and morbidity than sporadic nonagenarians: The Leiden Longevity Study. *J Am Geriatr Soc* **57**, 1634-7 (2009)
45. Weverling-Rijnsburger, A.W.E. *et al*. Total cholesterol and mortality in the oldest old. *Lancet* **350**, 1119-1123 (1997)
46. Bootsma- van der Wiel, A. *et al*. A high response is not essential to prevent selection bias: Results from the Leiden 85-plus study. *Journal of Clinical Epidemiology* **55**, 1119-1925 (2002).
47.Rantakallio P. The longitudinal study of the northern Finland birth cohort of 1966. *Paediatr Perinat Epidemiol*. **2**, 59–88 (1988)
48. Sovio U. *et al*. Genetic determinants of height growth assessed longitudinally from infancy to adulthood in the northern Finland birth cohort 1966. *PLoS Genet*. **5**, e1000409 (2009).

Nature Genetics: doi:10.1038/ng.2528

49. Boomsma, D.I. et al. Netherlands Twin Register: from twins to twin families. *Twin Res Hum Genet***9**, 849-857 (2006)
50. Boomsma,D.I. *et al*. Genome-wide association of Major Depression: Description of samples for the GAIN Major Depressive Disorder study: NTR and NESDA Biobank Projects. *Eur J Hum Genet***16**, 335-342 (2008)
51. Willemsen, G*. et al.*The Netherlands Twin Register Biobank; a resource for genetic epidemiological studies. *Twin Res Hum Genet.***13**, 231-245 (2010)
52. Norata, G.D. *et al*. Effects of the TLR4 variants on intima media thickness and monocyte-derived macrophage response to LPS. *Journal of Internal Medicine* **257**, 1-7 (2005)
53. Norata, G.D. *et al*. Effects of fractalkine receptor variants on common carotid artery intima-media thickness. *Stroke* **37**, 1558-1561 (2006)
54. Norata, G.D. *et al*. Leptin: adiponectin ratio is an independent predictor of intima-media thickness of the common carotid artery. *Stroke* **38**, 2844-2846 (2007)
55. Hillege, H.L. *et al*. Urinary albumin excretion predicts cardiovascular and non cardiovascular mortality in general population. *Circulation* **106**, 1777-82 (2002)
56. Verhave, J.C. et al. The reliability of different formulae to predict creatinine clearance. *J Intern Med*. **253**, 563-73 (2003)
57. Medlans, S.E. *et al*. Common variants in the trichohyalin gene are associated with straight hair in Europeans. *Am. J. Hum. Genet.* **85**, 750-755 (2009)
58. Wright, M.G. and Martin, N.G. Brisbane adolescent twin study: outline of study methods and research projects. *Aust J Psychol* **56**, 65-78 (2004)
59. Moayyeri, A. *et al*. Cohort Profile: TwinsUK and Healthy Ageing Twin Study. *Int J Epidemiol.* Epub (2012)
60. Richards, J.B. et al. Bone mineral density, osteoporosis, and osteoporotic fractures: a genome-wide association study. *Lancet*. **371**, 1505-12 (2008)
61.The Wellcome Trust Case Control Consortium. Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. *Nature* **447**, 661-678 (2007)
62. Wang X, Poole JC, Treiber FA, Harshfield GA, Hanevold CD, Snieder H. Ethnic and sex differences in ambulatory blood pressure trajectories: results from a 15-year longitudinal study in youth and young adults. *Circulation*. **114**, 2780-7 (2006)
63. Zhu H, Guo D, Li K, Pedersen-White J, Stallmann-Jorgensen IS, Huang Y, et al. Increased telomerase activity and vitamin D supplementation in overweight African Americans. *Int J Obes (Lond*) (2011)
64. Cawthon, R.M. Telomere measurement by quantitative PCR. Nucleic Acids Res 30, e47 (2002)
65. Codd, V. *et al.* Common variants near TERC are associated with mean telomere length. *Nat Genet* **42**, 197-9 (2010)
66. Eerola, J. *et al.* No evidence for shorter leukocyte telomere length in Parkinson's disease patients. *J Gerontol A BiolSci Med Sci* **65**, 1181-1184 (2010)
67. O'Callaghan, N. A quantitative real-time PCR method for absolute telomere length. *Biotechniques* **44**,807-809 (2008)
68. Cawthon RM. Telomere length measurement by a novel monochrome multiplex quantitative PCR method. *Nucleic Acids Res* **37**, e21 (2009)
69. Buxton, J.L. *et al.* Childhood Obesity Is Associated with Shorter Leukocyte Telomere Length. *Journal of Clinical Endocrinology & Metabolism* **96**, 1500-1505 (2011).
70. Huzen, J.*et al.* Circulating leukocyte and carotid atherosclerotic plaque telomere length: interrelation, association with plaque characteristics, and restenosis after endarterectomy. *Arterioscler Thromb Vasc Biol* **3**,1219-25 (2011)
71. Zeller, T. *et al.* Genetics and beyond – The transcriptome of human monocytes and disease susceptibility. *PLoS One* **5** e10693 (2010)

Nature Genetics: doi:10.1038/ng.2528

72. Rosenbloom, K.R., *et al*., ENCODE whole-genome data in the UCSC Genome Browser: update 2012. *Nucleic Acids Res*. **40**, d912-d917 (2012)
73. Jones, A.M. *et al.* Terc polymorphisms are associated both with susceptibility to colorectal cancer and with longer telomeres. *Gut***61**, 248-254 (2012)
74. Soncin, F. *et al.* E-cadherin acts as a regulator of transcripts associated with a wide range of cellular processes in mouse embryonic stem cells. *PLoS One* **6**, e21463 (2011).
75. Hur, E.M. *et al.* LIME, a novel transmembrane adaptor protein, associates with p56Ick and mediates T-cell activation. *J. Exp. Med*. **198**, 1463-1473 (2003).
76. Ahn, E., Lee, H., & Yun, Y. LIME acts as a transmembrane adaptor mediating BCR-dependent B-cell activation. *Blood* **107**, 1521-1527 (2006).

Nature Genetics: doi:10.1038/ng.2528