Meta-analyses identify 13 novel loci associated with age at menopause and highlights DNA repair and immune pathways

L, Stolk, JRB Perry, ... et al.

Nature Genetics: doi:10.1038/ng.1051

Supplementary Table 1. Phenotype information for the discovery and replication stage studies

a. Discovery studies

Study	N	mean age at which menopause age was collected (SD)	mean age at natural menopause (SD)	Specific menopause questions
AGES	1,315	76.34 (5.55)	48.89 (6.84)	At what age did your menstrual periods permanently stop? Do not include menstrual period bleeding resulting from using female hormone pills. If you are not sure, please make your best guess. Ever had hysterectomy?
ARIC	2,576	55.19 (5.40)	48.43 (4.03)	Have you reached menopause? Age when menopause began? Cause of menopause‘
CHS	958	72.4 (5.5)	49.3 (4.3)	How old were you at the time of your last natural period (menopause)? Have you ever had a hysterectomy, that is surgery to remove your uterus or womb? If yes, how old were you when you had this surgery? Have you ever had an ovary removed? If yes, how many ovaries were removed? At what age did you have this done? Have you ever taken Premarin for hot flashes or other symptoms of menopause? If yes, at what age did you start taking Premarin? At what age did you stop taking Premarin?
deCODE	5,857	birth year 1926.9 (8.9)	48.2(4.0)	How old were you when your menstruation ceased? Did you have your ovaries or uterus removed ? If yes, when did you have surgery? Did you use hormone replacement therapy ? If yes, When did you start and when did you stop the hormone replacement therapy?
EGCUT	279	61.1 (9.20)	49.5 (3.84)	Have your periods stopped for 1 year or more? Age periods stopped? Cause periods stopped? Have you used hormonal medicaments due to menopause? When did you start using them?
ERF	373	47.83 (14.33)	49.35 (3.872)	Q1) At what age did the menstruation stopped (and began menopause)? Q2) Why did the menstruation stopped? Q3) Have you used medication (mostly HRT) due to menopause‘
FHS	1,452	NA	49.9 (3.51)	Have your periods stopped for 1 year or more? Age periods stopped? Cause periods stopped? (natural, surgical, other) Hysterectomy (yes, no) Number of ovaries removed (0, 1, 2, unknown)
Amish	240	63.1 yrs (9.1)	49.0 yrs (3.9)	Have you reached menopause? Was your menopause natural or the result of surgery, radiation or chemotherapy? In what year or how old were you when you reached menopause?
InChianti	460	73.35 (8.65)	50.2 (4.09)	At what age did you go through the menopause. Was the menopause surgical? Have you ever used HRT? If used HRT, at what age did you start and stop?
NHS-cgems	1,344	56.78 (6.55)	50.78 (3.08)	"Have your menstrual periods ceased permanently?" If yes, "At what age did your natural periods cease?" and "For what reason did your periods cease?" Response categories were "Surgical; Radiation or Chemotherapy; Natural."
NHS-Hu	1,772	55.12 (6.71)	50.52 (3.36)	"Have your menstrual periods ceased permanently?" If yes, "At what age did your natural periods cease?" and "For what reason did your periods cease?" Response categories were "Surgical; Radiation or Chemotherapy; Natural."
NTR	331	58.24 (5.96)	49.00 (3.69)	Have you reached the menopause (no menstrual period in the last 12 months)? Was this spontaneously? At what age did the menopause start?
QIMR	430	31.3 (10.3)	48.3 (4.4)	Have your periods stopped for at least one year? If they have stopped, was this due to a) menopause b) hysterectomy c) other complications? What age were you when your periods stopped?
RSI	2,196	70.35 (9.38)	49.87 (3.89)	Did you have a menstrual period in the past 12 months? Age at last menstrual period? For what reasons did the periods stop?
RSII	665	65.65 (8.70)	50.52 (3.97)	
RSIII	597	58.44 (5.45)	50.27 (3.83)	
SardiNIA	828	61,2 (10.12)	49.8 (4.13)	How old were you at the time of your last natural period (menopause)? Cause periods stopped?
SHIP	4,310	55.0 (3.75)	49.7 (3.85)	Do you have a menstrual period? How old were you at the time of your last natural period? Have your periods stopped for natural reasons or following treatment or diseases?
TwinsUKI	605	55.6 (6.70)	48.5(3.8)	what was your age at last regular period? Have you ever had hysterectomy and/or ovary removal? Have you ever taken Hormonal replacement Therapy (HRT)? Are you currently taking HRT?
TwinsUKII	258	68.51(8.21)	47.67 (4.17)	
TwinsUKIII	743	67.51(7.87)	48.72 (4.33)	
WGHS	11379	54.68 (7.19)	50.58 (3.64)	‘‘Have your menstrual periods ceased permanently?’’ If yes, ‘‘At what age did your natural periods cease?’’ and ‘‘For what reason did your periods cease?’’ Response categories were ‘‘Surgical; Radiation or Chemotherapy; Natural.’’
Total	**38,968**			

b. Replication studies

Study	N	mean age at which menopause age was collected (SD)	mean age at natural menopause (SD)	Specific menopause questions
BWHHS	2971	69.2 (5.5)	49.4 (4.0)	Women were asked to report their age, in years, at which they had experienced their last menstrual period. In a separate question women were asked to report all operations they had ever undergone and the timing of these operations. These data were used to identify women who had had a hysterectomy or oophorectomy and the timing of these operations. Women were asked if they had ever used hormone replacement therapy and if so the name of the therapy and their age at commencing and stopping (if no longer using) the therapy. At the research nurse interview women were asked to bring all of their current medications and for women who were currently using hormone replacement their self-report of hormone replacement use was verified at the nurse interview drugs history.
COLAUS	1013	61.1 (7.2)	49.6 (4.1)	At about what age was your last menstrual period? Did you already have a hysterectomy combined with an ovariectomy? Have you ever taken hormone replacement therapy (oestrogenes)?
EGCUT	396	80.5 (9.83)	49.8 (3.84)	Have your periods stopped for 1 year or more? Age periods stopped? Cause periods stopped? Have you used hormonal medicaments due to menopause? When did you start using them?
EPOS	903	50.8 (2.41)	49.8 (3.2)	How long ago was your last menstrual period? What was the month and year of your last menstrual period? Which gynaecological surgery did you have? Did you use femoale hormones, if yes when?
GENOA	283	62.77 (8.15)	50.17 (4.23)	Have you reached menopause? Was your menopause natural, or the result of surgery, radiation, or chemotherapy? In what year or how old were you when you reached menopause?
HBCS	556	61.51 (3.09)	50.65 (3.90)	At what age did you have last menstrual period? Has your uterus been removed; and if it has; at what age? Do you use estrogen replacement therapy?
INGI - CARL	134	62.39(8.61)	49.02(4.03)	Did you have a menstrual period in the past 12 months? Age at last menstrual period? Do you take hormones? Was your menopause natural, or the result of surgery (ovariectomy, hysterectomy, annessiectomy), radiation, or chemotherapy? If yes, What age?
INGI - FVG	254	65.46(9.6)	50.65(3.70)	Did you have a menstrual period in the past 12 months? Age at last menstrual period? Do you take hormones? Was your menopause natural, or the result of surgery (ovariectomy, hysterectomy, annessiectomy), radiation, or chemotherapy? If yes, What age?
INGI - Val Borbera	476	67.22 (10.87)	50.56 (3.47)	Did you undergo menopause? If yes, at what age? Was it natural, or consequence of surgery, radiotherapy or chemotherapy, or other?
KORA F3	391	65.44 (7.33)	50.16 (4.09)	Did you have a menstrual period in the past 12 months? How old were you when you had your last menstrual period? Did you ever take hormones? How old were you then you took these hormones for the first time? How many months or years in total did you take these hormones? Do you currently take hormones? Did you have hysterectomy or ovarectomy?
KORA S4	201	60.48 (5.75)	49.19 (4.09)	Did you have a menstrual period in the past 12 months? How old were you when you had your last natural menstrual period? Did you ever take hormones? How old were you then you took these hormones for the first time? How many months or years in total did you take these hormones? Do you currently take hormones? Did you have hysterectomy or ovarectomy?
KORCULA	333	62.28 (9.07)	49.60 (4.12)	Are you having a regular menstrual cycle? Age at menopause? Have you had any type of surgery that could have induced menopause? Do you use HRT?
LBC1936	337	69.54 (0.83)	50.00 (3.99)	Can you remember how old you were when you had your last period? . Have you ever had a hysterectomy? Have you ever had an oophorectomy? Do you or have you ever used hormone replacement therapy?If yes, give details of the drug taken and duration of use
LifeLines	622	59.96 (7.64)	50.24 (3.79)	Do you still have periods (menstruation)? If you no longer have periods, how old were you when you last had a period? Has your uterus (womb) and/or ovary/ovaries been removed? In the last 5 years before you stopped menstruating, did you use hormonal contraception (also Depo Provera or IUD device)? In the last 5 years before you stopped menstruating, did you have hormonal treatment for any other reason other than contraception?
ORCADES	145	62.46(7.81)	49.50 (5.05)	Women were asked if they were still having periods. If they answered "no" they were asked their age in years when they had their last period. Women were asked if they had ever used hormone replacement therapy (HRT) and, if they had, when they started using it either in terms of age or calendar year. Women were asked if they were using HRT now and, if so what type they were using (tablets, patches, other and its name). Women no longer using HRT were asked when they stopped using HRT either in terms of age or calendar year. Women were asked whether they had had a hysterectomy and the year in which the operation was performed. They were asked if they had had both ovaries removed and in which year (the year in which the second ovary was removed if two operations were performed).
OSTEOS	498	61.97 (9.84)	48.71 (4.17)	Age at last menstrual period? Was it natural or the result of surgery or any other clinical situtation or medications? Did you receive HRT?

PROSPECT-EPIC	3424	63.03 (3.74)	50.53 (3.74)	Do you still have menstrual periods? If not, at what age did the periods stop? Did you menstruate in the previous 12 months? Did you use the Pill or other hormones for menopausal complaints? At what age did you start with these hormones? How long did you use these hormones? At what age did you stop with these hormones? Is your uterus surgically removed? If yes, at what age? Are one or both ovaries removed (can answer no, one, both or don't know)? At what age were they removed?
SASBAC	764	62.97 (6.27)	50.29 (3.36)	Do you still have menstrual periods? If not, what is the reason for this, and at what age did the periods stop? Answers could be: - It stopped by itself (naturally) - The uterus was removed - The ovaries were removed - Hormone replacement therapy was terminated - Other reason - Don't know - Both uterus and ovaries were removed
SPLIT	141	60.60(7.09)	49.99 (3.82)	Are you having a regular menstrual cycle? Age at menopause? Have you had any type of surgery that could have induced menopause? Do you use HRT?
TWINGENE	280	65.27(5.82)	50.3 (3.99)	Have you had menstrual periods during the last year? If not, why? Answers could be: Removal of uterus/ovaries Pregnancy Menopause Anorexia/eating disorder Physical training at elite level Medication Others
VIS	313	65.65(10.60)	48.78 (4.13)	How old were you when you had your last menstrual period? Are you having a regular menstrual cycle? Age at menopause? Have you had any type of surgery that could have induced menopause? Do you use HRT?
Total	**14435**			

Supplementary Table 2. Genotyping information for the discovery and replication stage studies

a. Discovery Studies

			Genotyping			Imputation		
Study	**N (samples)**	**Array**	**Callrate cut-off**	**MAF* cut-off**	**HWE cut-off**	**total N (SNPs)**	**Software**	**Analysis program**
AGES	1315	Illumina HumanHap 370K CNV	98%	0.01	1.0E-06	2,543,887	MACH	PLINK
Amish	240	Affymetrix 500K	95%	0.01	1.0E-06	2,404,474	MACH	MMAP (J. O'Connell)
ARIC	2576	Affymetrix 6.0	90%	0.01	1.0E-06	2,500,000	MACH	ProbABEL
CHS	958	Illumina HumanHap 370K CNV	97%		1.0E-05	2,543,887	BimBam	R-packages
deCODE	5857	Illumina HumanHap 300K and 370K CNV	95%	NA	1.0E-06	2,542,879	IMPUTE	Logstic regression using allele count as a covariate
ERF	373	Illumina 6K, Illumina 318K, Illumina 370K, Afyymetrix 250K	98%	0.01	1.0E-06	2543887	MACH	ProbABEL
Estonia	279	Illumina HumanHap 370K CNV	98%	0.01	1.0E-06	2551007	IMPUTE	SNPTEST
FHS	1452	Affymetrix 500K + Affymetrix 50K	97%	0.01	1.0E-06	2,539,029	MACH	R-packages
InChianti	460	Illumina HumanHap 550K	98%	0.01	1.0E-04	2,543,887	IMPUTE	SNPTEST
NHS-cgems	1344	Illumina HumanHap 550K	90%	0.01	NA	2,619,698	MACH	ProABEL
NHS-Hu	1772	Affymetrix 6.0	98%	0.02	1.0E-04	2,641,555	MACH	ProABEL
NTR	331	Affymetrix 500K Perlegen	95%	0.01	NA	2441556	IMPUTE	SNPTEST
QIMR	433	Illumina 317k + Illumina 370k + Illumina 610k	95%	0.01	1.0E-06	2,397,036	MACH	MERLIN --fastassoc
RSI / RSII / RSI	2196 / 597	Illumina HumanHap 550K	98%	0.01	1.0E-06	2,543,887	MACH	MACH2QTL
SardiNIA	828	Affymetrix 500K + Affymetrix 10K	90.00%	0.05	1.0E-06	2,252,228	MACH	MERLIN --fastassoc
SHIP	230	Affymetrix 6.0	92%	NA	NA	2,748,910	IMPUTE	SNPTEST
TwinsUKI	605	Illumina HumanHap 300K	95% (MAF 5%) / 99% (MAF 1-5%)	0.01	5.7E-05	2,544,233	IMPUTE	GenABEL
TwinsUKII	258	Illumina Hap610Quad	95% (MAF 5%) / 99% (MAF 1-5%)	0.01	5.7E-05	2,557,509	IMPUTE	GenABEL
TwinsUKIII	743	Illumina Hap610Quad	95% (MAF 5%) / 99% (MAF 1-5%)	0.01	5.7E-05	2,548,895	IMPUTE	GenABEL
WGHS	22054	Illumina HumanHap300 Duo "+"	98%	NA	1.00E-06	2,621,896	MACH	MACH2QTL

b. In silico Replication Studies

			Genotyping			Imputation		
Study	**N (samples)**	**Array**	**Callrate cut-off**	**MAF* cut-off**	**HWE cut-off**	**total N (SNPs)**	**Software**	**Analysis program**
CARL	134	Illumina 370K	97%	0.01	1.0E-06	2,401,930	MACH	R-packages
COLAUS	1013	Affymetrix 500K	70%	0.01	1.00E-07	2,557,249	IMPUTE	Matlab
EGCUT	396	Illumina HumanHap 370K	98%	0.01	1.0E-06	2,548,526	IMPUTE	SNPTEST
FVG	254	Illumina 370K	97%	0.01	1.0E-06	2,526,221	MACH	R-packages
GENOA	283	Affymetrix 6.0	95%	0.01	-	2,513,174	MACH	R-packages
HBCS	556	modified Illumina 610k	95%	0.01	1.0E-05	2,543,887	MACH	Plink and ProbABEL
INGI-Val Borbera	476	Illumina HumanHap 370K	95%	0.01	1.0E-06	2,443,906	MACH	ProbABEL
KORA F3	391	Affymetrix 500K	93%	none	none	2,557,252	MACH v1.0.9	R
KORA S4	201	Affymetrix 6.0	93%	none	none	2,543,887	MACH v1.0.15	R
KORCULA	333	Illumina Hap370CNV	98%	0.01	1.0E-06	2543888	MACH v1.0.15	GenABEL/ProbABEL
LBC1936	337	Illumina Human 610_Quadv1	95%	0.01	0.001	2,543,887	Mach v1.0.16	MACH2QTL
LifeLines	622	Illumina CytoSNP v 2.0- 300K	99%	0.01	1.0E-04	3235181	IMPUTE v0.3.2	SNPTEST v1.1.3
ORCADES	145	Illumina HumanHap300	98%	0.01	1.0E-06	2543888	MACH v1.0.15	GenABEL/ProbABEL
SASBAC	764	Illumina HumanHap550	90%	3%	1.0E-07	2,499,358	IMPUTE	SNPTEST
SPLIT	141	Illumina Hap370CNV	98%	0.01	1.0E-06	2543888	MACH v1.0.15	GenABEL/ProbABEL
TWINGENE	280	Illumina 317K	95%	5%	1.00E-07	2 552 337	IMPUTE	PLINK, R
VIS	313	Illumina HumanHap300v1	95%	0.01	1.0E-06	2543888	MACH v1.0.15	GenABEL/ProbABEL

c. De novo genotyped replication studies

Study	**N (samples)**	**Genotyping Technique**	**% samples in duplicate**	**HWE cut-off**	**Analysis program**
BWHHS	2971	KASPar	-	0.01	STATA
EPOS	903	Taqman & Sequenom iPLEX	5%	0.001	PLINK
OSTEOS	498	22 assay Sequenom iPlex	1.4%	0.01	PLINK
Prospect-Epic	3424	KASPar	0%	0.001	PLINK

*MAF: Minor allele frequency
NA: Not Applicable

Supplementary Table 3. Odds ratio estimates and 95% confidence bounds (lower,upper) for the menopause decreasing allele on dichotomized age at menopause in the WGHS study women

	Age < 45 vs >=45			Age >54 vs <=54		
SNP	**OR**	**lower**	**upper**	**OR**	**lower**	**upper**
rs4246511	**1.13**	1.00	1.29	**0.93**	0.85	1.01
rs1635501	**1.06**	0.93	1.20	**0.90**	0.82	0.99
rs2303369	**1.12**	1.01	1.24	**0.89**	0.82	0.96
rs10183486	**1.11**	1.00	1.24	**0.90**	0.83	0.97
rs4693089	**1.10**	0.99	1.23	**0.90**	0.84	0.98
rs890835	**1.11**	0.93	1.32	**0.84**	0.75	0.95
rs365132	**1.21**	1.09	1.35	**0.84**	0.78	0.90
rs2153157	**1.09**	0.98	1.21	**0.90**	0.83	0.97
rs1046089	**1.22**	1.09	1.36	**0.88**	0.81	0.95
rs2517388	**1.24**	1.05	1.47	**0.89**	0.80	0.99
rs12294104	**1.15**	1.00	1.33	**0.96**	0.87	1.05
rs2277339	**1.12**	0.95	1.32	**0.78**	0.68	0.89
rs3736830	**1.22**	1.06	1.40	**0.87**	0.78	0.96
rs4886238	**1.18**	1.05	1.32	**0.85**	0.79	0.92
rs2307449	**1.18**	1.06	1.32	**0.91**	0.84	0.98
rs10852344	**1.25**	1.12	1.40	**0.89**	0.83	0.96
rs11668344	**1.29**	1.16	1.43	**0.85**	0.78	0.92
rs12461110	**1.01**	0.91	1.13	**0.96**	0.89	1.04
rs16991615	**2.03**	1.51	2.73	**0.52**	0.46	0.60

Nature Genetics: doi:10.1038/ng.1051

Supplementary Table 4: Functional networks and their relevant functions for candidate genes associated with age of menopause

	Age of menopause associated genes in Network	Other related genes/molecules in network	Associated network functions
Network 1: p=1E-30	14 genes: APOM, BAT2, BAT4, BAT5, EXO1, GCKR, GSPT1, HELQ, MCM8, MRPS18C, NACA, RRAGC, TDRD3, TRMT6	CCDC90B, EXOSC2, FGA, HNF4A, HNRNPA1, HNRNPM, HSPA2, JUN, KHSRP, KLF6, MAPKSP1, NR5A2, PLA2G4A, PTGDS, RNF5 (includes EG:6048), ROBLD3, RRAGA, RRAGB, SAFB, SSR3, XRN1	Lipid Metabolism, Molecular Transport, Small Molecule Biochemistry
Network 2: p=1E-24	12 genes: ASH2L, C6ORF47, CHGB, CSNK2B, EIF4EBP1, FAM175A, FANCI, LSM1, LTB, MYCBP, TLK1, UIMC1,	ASGR2, BCL2, BRCC3, C16ORF53, CCNH, CHD8, CHEK1, DPY30, ESR1, FGFR3, HSP90AA1, IER3, ITPR2, MYC, NARS, RELA, SAFB2, SLPI, TLK1/2, TP53, UBA3, UXT, VRK1	Cell Cycle, Cell Death, Cancer
Network 3: p=1E-19	10 genes: BAG4, BAT3, BRSK1, GCM2, HK3, HSPBP1, NLRP4, NRBP1, TNF, TNFRSF17	CALM2, CHUK, DNAJB6, DNAJB11, EDN1, HSF1, Hsp70, HSPA2, HSPA4, HSPA1A, IKBKAP, MBP, MMP7, MMP12, NFKBIA, POU2F1, PP1-C, PTGDS, PTH, SHC1, ST13, Tnf receptor, TNFRSF13B, TNFSF13, TNFSF13B	Cell Death, Hematological System Development and Function, Cellular Development
Network 4: p=1E-12	7 genes: COX6B2, IL11, LST1, POLG, PRIM1, STAR, SUV420H2	Ap1, ATP5B, ATXN2, BCL3, C/ebp, CCNA2, CEBPA, Cytochrome c oxidase, DLG4, DNA-directed DNA polymerase, FEZ1, HTT, MCM10, NFATC3, NFATC4, NR1H3, PIK3R3, POLA1, POLA2, POLD3, POLG2, PRIM2, PRMT2, RAE1,	Infection Mechanism, DNA Replication, Recombination, and Repair, Gene Expression
		p-value	***# focus genes***
Diseases and disorders	Endocrine System Disorders	1.34E-05 - 1.24E-02	19
	Immunological Disease	1.34E-05 - 1.02E-02	20
	Metabolic Disease	1.34E-05 - 1.24E-02	19
	Infectious Disease	5.69E-05 - 1.24E-02	3
	Inflammatory Response	5.69E-05 - 1.55E-02	4
Molecular and cellular functions	Antigen Presentation	2.85E-05 - 1.55E-02	2
	Cell Death	2.85E-05 - 1.55E-02	8
	Cellular Development	2.85E-05 - 1.41E-02	8
	Cellular Function and Maintenance	2.85E-05 - 1.55E-02	3
	Cellular Movement	2.85E-05 - 1.55E-02	3
Canonical Pathways	Altered T Cell and B Cell Signaling in Rheumatoid Arthritis	2.43E-03	
	Crosstalk between Dendritic Cells and Natural Killer Cells	3.03E-03	
	NF-kB Signaling	1.54E-02	
	Airway Inflammation in Asthma	1.55E-02	
	Airway Pathology in Chronic Obstructive Pulmonary Disease	2.47E-02	

Nature Genetics: doi:10.1038/ng.1051

Supplementary Table 5. Genes in the significant pathways for MAGENTA analysis

Database	Gene Set	Gene Symbol	Entrez ID	Gene p-value	Gene Chr Num	Gene Start Pos	Gene End Pos	Gene Size kb	Num SNPs per Gene	Best SNP rs	Best SNP Chr Num	Best SNP Chr Pos	Best SNP Z	Best SNP pval
PANTHER_MOLECULAR_FUNCT	Exodeoxyribonuclease	EXO1	9156	4.64E-07	1	240078157	240119671	42	270	rs1635501	1	240107398	5.8507	4.90E-09
PANTHER_MOLECULAR_FUNCT	Exodeoxyribonuclease	POLG	5428	2.27E-06	15	87660539	87679030	18	134	rs2307449	15	87664932	5.5674	2.59E-08
PANTHER_MOLECULAR_FUNCT	Exodeoxyribonuclease	APEX1	328	2.69E-04	14	19993129	19995766	3	218	rs1760940	14	20008091	5.1431	2.70E-07
PANTHER_MOLECULAR_FUNCT	Exodeoxyribonuclease	REV3L	5980	4.18E-04	6	111726926	111911107	184	279	rs7776184	6	111916611	4.7622	1.92E-06
PANTHER_MOLECULAR_FUNCT	Exodeoxyribonuclease	WRN	7486	6.65E-02	8	31010319	31150819	140	285	rs1882928	8	31143364	3.5947	3.25E-04
PANTHER_MOLECULAR_FUNCT	Exodeoxyribonuclease	RAD50	10111	1.58E-01	5	131920528	132007494	87	199	rs1023518	5	131821671	3.1798	1.47E-03
PANTHER_MOLECULAR_FUNCT	Exodeoxyribonuclease	MRE11A	4361	1.67E-01	11	93790114	93866688	77	264	rs541472	11	93798531	3.1594	1.58E-03
PANTHER_MOLECULAR_FUNCT	Exodeoxyribonuclease	FEN1	2237	2.46E-01	11	61316725	61321286	5	92	rs102275	11	61314379	2.8776	4.01E-03
PANTHER_MOLECULAR_FUNCT	Exodeoxyribonuclease	POLE	5426	2.63E-01	12	131710420	131774018	64	154	rs10870504	12	131823410	2.8444	4.45E-03
PANTHER_MOLECULAR_FUNCT	Exodeoxyribonuclease	POLN	353497	6.18E-01	4	2043442	2200756	157	203	rs497881	4	2094612	2.336	1.95E-02
PANTHER_MOLECULAR_FUNCT	Exodeoxyribonuclease	POLD1	5424	6.64E-01	19	55579404	55613083	34	101	rs3218775	19	55610041	2.1525	3.14E-02
PANTHER_MOLECULAR_FUNCT	Exodeoxyribonuclease	RAD1	5810	8.92E-01	5	34941122	34954139	13	127	rs10521020	5	35022621	1.5552	1.20E-01
PANTHER_MOLECULAR_FUNCT	Exodeoxyribonuclease	POLA1	5422	NaN	23	24621976	24925023	303	0	NaN	NaN	NaN	NaN	NaN
PANTHER_MOLECULAR_FUNCT	Exodeoxyribonuclease	APEX2	27301	NaN	23	55043504	55050937	7	0	NaN	NaN	NaN	NaN	NaN
Ingenuity	Mitochondrial.Dysfunctio	NDUFS5	4725	7.23E-06	1	39264592	39272874	8	110	rs7537437	1	39212125	5.33	9.82E-08
Ingenuity	Mitochondrial.Dysfunction	NDUFAF1	51103	1.60E-04	15	39466842	39481934	15	49	rs11070328	15	39438861	4.748	2.05E-06
Ingenuity	Mitochondrial.Dysfunction	GPX7	2882	3.57E-03	1	52840631	52847310	7	106	rs17374258	1	52882590	4.2187	2.46E-05
Ingenuity	Mitochondrial.Dysfunction	UQCRH	7388	4.57E-03	1	46541966	46555034	13	81	rs12142240	1	46519888	4.0703	4.70E-05
Ingenuity	Mitochondrial.Dysfunction	SDHB	6390	3.08E-02	1	17217811	17253252	35	154	rs2746533	1	17266633	3.9375	8.23E-05
Ingenuity	Mitochondrial.Dysfunction	NDUFA6	4700	1.49E-02	22	40811475	40816834	5	82	rs5751229	22	40875165	3.7745	1.60E-04
Ingenuity	Mitochondrial.Dysfunction	SDHD	6392	2.03E-02	11	111462831	111471727	9	65	rs360726	11	111494802	3.7616	1.69E-04
Ingenuity	Mitochondrial.Dysfunction	NDUFA11	126328	2.86E-02	19	5845685	5855025	9	106	rs778800	19	5851247	3.7096	2.08E-04
Ingenuity	Mitochondrial.Dysfunction	PARK2	5071	4.97E-01	6	161688579	163068824	1380	2038	rs12194653	6	162512485	3.6569	2.55E-04
Ingenuity	Mitochondrial.Dysfunction	NDUFS7	374291	5.29E-02	19	1334882	1346588	12	89	rs3786982	19	1366875	3.583	3.40E-04
Ingenuity	Mitochondrial.Dysfunction	NDUFS1	4719	4.31E-02	2	206696047	206732432	36	116	rs13410751	2	206794997	3.5084	4.51E-04
Ingenuity	Mitochondrial.Dysfunction	NCSTN	23385	4.31E-02	1	158579686	158595366	16	111	rs12404657	1	158617853	3.498	4.69E-04
Ingenuity	Mitochondrial.Dysfunction	CASP3	836	6.99E-02	4	185785843	185807623	22	168	rs12108497	4	185808551	3.4962	4.72E-04
Ingenuity	Mitochondrial.Dysfunction	NDUFA4L2	56901	3.83E-02	12	55914952	55920742	6	50	rs11830804	12	56021159	3.4716	5.17E-04
Ingenuity	Mitochondrial.Dysfunction	NDUFA10	4705	1.33E-01	2	240548828	240613471	65	251	rs2099727	2	240628632	3.4428	5.76E-04
Ingenuity	Mitochondrial.Dysfunction	NDUFA3	4696	1.04E-01	19	59297971	59302080	4	100	rs10424816	19	59322020	3.4174	6.32E-04
Ingenuity	Mitochondrial.Dysfunction	PRDX5	25824	1.44E-01	11	63842144	63845859	4	58	rs915987	11	63784464	3.0009	2.69E-03
Ingenuity	Mitochondrial.Dysfunction	NDUFA7	4701	1.94E-01	19	8282233	8292280	10	71	rs2232778	19	8276120	2.9774	2.91E-03
Ingenuity	Mitochondrial.Dysfunction	UQCRFS1	7386	2.95E-01	19	34390006	34395976	6	156	rs2193516	19	34449537	2.898	3.76E-03
Ingenuity	Mitochondrial.Dysfunction	NDUFB2	4708	1.84E-01	7	140042949	140052915	10	50	rs2930315	7	140038053	2.875	4.04E-03
Ingenuity	Mitochondrial.Dysfunction	CAT	847	3.71E-01	11	34417053	34450183	33	318	rs4756138	11	34364874	2.8206	4.79E-03
Ingenuity	Mitochondrial.Dysfunction	UQCRB	7381	3.35E-01	8	97308479	97317038	9	131	rs17781326	8	97424842	2.8168	4.85E-03
Ingenuity	Mitochondrial.Dysfunction	PRDX3	10935	2.98E-01	10	120917204	120928335	11	105	rs4237510	10	120983651	2.8068	5.00E-03
Ingenuity	Mitochondrial.Dysfunction	GSR	2936	3.02E-01	8	30655976	30704985	49	112	rs9721100	8	30627769	2.7725	5.56E-03
Ingenuity	Mitochondrial.Dysfunction	SOD2	6648	3.44E-01	6	160020138	160034343	14	138	rs7754103	6	159981080	2.6806	7.35E-03
Ingenuity	Mitochondrial.Dysfunction	UQCRC2	7385	2.75E-01	16	21872109	21902169	30	17	rs11554798	16	21884263	2.6445	8.18E-03
Ingenuity	Mitochondrial.Dysfunction	NDUFB9	4715	4.20E-01	8	125620523	125631408	11	158	rs11778396	8	125556947	2.6193	8.81E-03
Ingenuity	Mitochondrial.Dysfunction	SDHA	6389	4.41E-01	5	271355	309814	38	178	rs6884461	5	206609	2.5697	1.02E-02
Ingenuity	Mitochondrial.Dysfunction	TXNRD2	10587	5.34E-01	22	18243039	18309359	66	230	rs405342	22	18417315	2.5514	1.07E-02
Ingenuity	Mitochondrial.Dysfunction	APP	351	5.96E-01	21	26174731	26465003	290	472	rs1981369	21	26171874	2.5421	1.10E-02
Ingenuity	Mitochondrial.Dysfunction	NDUFB8	4714	4.35E-01	10	102273486	102279626	6	111	rs4573623	10	102361371	2.5232	1.16E-02
Ingenuity	Mitochondrial.Dysfunction	CYB5R3	1727	5.30E-01	22	41344757	41375349	31	171	rs1009433	22	41383757	2.4854	1.29E-02
Ingenuity	Mitochondrial.Dysfunction	GPD2	2820	5.15E-01	2	157000210	157151161	151	189	rs1432573	2	157155967	2.4829	1.30E-02
Ingenuity	Mitochondrial.Dysfunction	PARK7	11315	4.69E-01	1	7944300	7967929	24	88	rs11585581	1	7888598	2.4671	1.36E-02
Ingenuity	Mitochondrial.Dysfunction	UCP2	7351	4.79E-01	11	73363363	73371537	8	114	rs2632723	11	73390777	2.4568	1.40E-02
Ingenuity	Mitochondrial.Dysfunction	PSEN1	5663	4.95E-01	14	72672931	72756862	84	158	rs7160247	14	72795226	2.4469	1.44E-02
Ingenuity	Mitochondrial.Dysfunction	OGDH	4967	5.14E-01	7	44612695	44715194	102	151	rs2331138	7	44621283	2.4439	1.45E-02
Ingenuity	Mitochondrial.Dysfunction	NDUFB10	4716	5.08E-01	16	1949517	1951977	2	146	rs338779	16	1945564	2.4179	1.56E-02
Ingenuity	Mitochondrial.Dysfunction	TXN2	25828	7.21E-01	22	35193038	35207633	15	207	rs139954	22	35179861	2.3217	2.03E-02
Ingenuity	Mitochondrial.Dysfunction	NDUFS2	4720	5.46E-01	1	159435728	159450808	15	99	rs2501870	1	159479193	2.2862	2.22E-02
Ingenuity	Mitochondrial.Dysfunction	SDHC	6391	6.37E-01	1	159550789	159601159	50	199	rs2501870	1	159479193	2.2862	2.22E-02
Ingenuity	Mitochondrial.Dysfunction	XDH	7498	6.83E-01	2	31410691	31491115	80	221	rs992137	2	31437539	2.2508	2.44E-02
Ingenuity	Mitochondrial.Dysfunction	NDUFV2	4729	6.42E-01	18	9092724	9124336	32	142	rs7240172	18	9004753	2.2298	2.58E-02
Ingenuity	Mitochondrial.Dysfunction	NDUFA5	4698	6.18E-01	7	122968318	122985194	17	119	rs11972308	7	123021321	2.2086	2.72E-02
Ingenuity	Mitochondrial.Dysfunction	NDUFA9	4704	7.48E-01	12	4628543	4666660	38	261	rs12318978	12	4584572	2.1697	3.00E-02
Ingenuity	Mitochondrial.Dysfunction	NDUFB6	4712	6.41E-01	9	32543522	32563182	20	104	rs11795343	9	32513737	2.1442	3.20E-02
Ingenuity	Mitochondrial.Dysfunction	NDUFB4	4710	6.43E-01	3	121797817	121803865	6	89	rs4286452	3	121733033	2.1027	3.55E-02
Ingenuity	Mitochondrial.Dysfunction	SNCA	6622	7.85E-01	4	90865727	90977156	111	305	rs2737035	4	90996236	2.0718	3.83E-02
Ingenuity	Mitochondrial.Dysfunction	GPX4	2879	7.96E-01	19	1054935	1057787	3	111	rs7247087	19	982212	2.0151	4.39E-02
Ingenuity	Mitochondrial.Dysfunction	NDUFB7	4713	7.61E-01	19	14537889	14543886	6	90	rs9305041	19	14497928	2.0109	4.43E-02
Ingenuity	Mitochondrial.Dysfunction	NDUFA8	4702	8.32E-01	9	123946158	123961919	16	192	rs1124729	9	123940095	1.989	4.67E-02
Ingenuity	Mitochondrial.Dysfunction	DHODH	1723	7.44E-01	16	70600143	70616456	16	141	rs7185407	16	70592458	1.965	4.94E-02
Ingenuity	Mitochondrial.Dysfunction	NDUFA4	4697	7.69E-01	7	10939339	10946338	7	148	rs2159126	7	10899911	1.9502	5.12E-02
Ingenuity	Mitochondrial.Dysfunction	NDUFS4	4724	7.78E-01	5	52892221	53014928	123	210	rs381575	5	52948648	1.9463	5.16E-02
Ingenuity	Mitochondrial.Dysfunction	CASP8	841	7.37E-01	2	201806410	201860679	54	127	rs7571586	2	201848183	1.9258	5.41E-02
Ingenuity	Mitochondrial.Dysfunction	NDUFS8	4728	7.20E-01	11	67554669	67560690	6	57	rs479763	11	67531704	1.9241	5.43E-02
Ingenuity	Mitochondrial.Dysfunction	NDUFB5	4711	7.49E-01	3	180805268	180824982	20	67	rs2339843	3	180794675	1.9172	5.52E-02
Ingenuity	Mitochondrial.Dysfunction	PSENEN	55851	7.81E-01	19	40928333	40929743	1	85	rs12459634	19	40922014	1.8185	6.90E-02
Ingenuity	Mitochondrial.Dysfunction	NDUFA12	55967	8.33E-01	12	93889240	93921642	32	202	rs11107870	12	93958583	1.8033	7.13E-02
Ingenuity	Mitochondrial.Dysfunction	NDUFB1	4707	8.89E-01	14	91652220	91657906	6	178	rs11160044	14	91652408	1.7384	8.21E-02
Ingenuity	Mitochondrial.Dysfunction	NDUFS3	4722	7.94E-01	11	47557207	47562690	5	38	rs4752783	11	47514796	1.6907	9.09E-02
Ingenuity	Mitochondrial.Dysfunction	MAP2K4	6416	8.90E-01	17	11864859	11987776	123	164	rs2079626	17	11824053	1.6262	1.04E-01
Ingenuity	Mitochondrial.Dysfunction	UQCRC1	7384	8.72E-01	3	48611435	48622102	11	56	rs7631574	3	48638050	1.5374	1.24E-01
Ingenuity	Mitochondrial.Dysfunction	CASP9	842	9.22E-01	1	15691381	15723377	32	164	rs1800615	1	15704868	1.5078	1.32E-01
Ingenuity	Mitochondrial.Dysfunction	NDUFB3	4709	8.89E-01	2	201644706	201658718	14	44	rs16842071	2	201639975	1.4654	1.43E-01
Ingenuity	Mitochondrial.Dysfunction	NDUFV1	4723	9.24E-01	11	67130898	67136588	6	91	rs2514037	11	67172766	1.4484	1.48E-01
Ingenuity	Mitochondrial.Dysfunction	NDUFA2	4695	9.34E-01	5	140005141	140007424	2	99	rs17208187	5	140003422	1.369	1.71E-01
Ingenuity	Mitochondrial.Dysfunction	NDUFA13	51079	9.19E-01	19	19487539	19500073	13	83	rs7259773	19	19472774	1.3538	1.76E-01
Ingenuity	Mitochondrial.Dysfunction	NDUFAB1	4706	9.39E-01	16	23499835	23515140	15	112	rs7500128	16	23616527	1.3316	1.83E-01
Ingenuity	Mitochondrial.Dysfunction	GLRX2	51022	9.47E-01	1	191332222	191341867	10	82	rs10801179	1	191387156	1.2486	2.12E-01
Ingenuity	Mitochondrial.Dysfunction	HTRA2	27429	9.57E-01	2	74610039	74614191	4	52	rs3806607	2	74589122	1.0964	2.73E-01
Ingenuity	Mitochondrial.Dysfunction	HSD17B10	3028	NaN	23	53474930	53478048	3	0	NaN	NaN	NaN	NaN	NaN
Ingenuity	Mitochondrial.Dysfunction	MAOA	4128	NaN	23	43400352	43491012	91	0	NaN	NaN	NaN	NaN	NaN
Ingenuity	Mitochondrial.Dysfunction	MAOB	4129	NaN	23	43510800	43626665	116	0	NaN	NaN	NaN	NaN	NaN
Ingenuity	Mitochondrial.Dysfunction	AIFM1	9131	NaN	23	129091018	129127489	36	0	NaN	NaN	NaN	NaN	NaN
Ingenuity	Mitochondrial.Dysfunction	NDUFB11	54539	NaN	23	46886562	46889381	3	0	NaN	NaN	NaN	NaN	NaN
Ingenuity	NFKB.Signaling	TNFRSF17	608	4.78E-11	16	11966464	11969426	3	109	rs10852344	16	11924420	6.7704	1.28E-11
Ingenuity	NFKB.Signaling	CARD10	29775	3.09E-02	22	36216345	36245156	29	148	rs9610813	22	36318738	3.7826	1.55E-04
Ingenuity	NFKB.Signaling	CHUK	1147	4.65E-02	10	101938113	101979334	41	158	rs2862954	10	101902054	3.6255	2.88E-04

Ingenuity	NFKB.Signaling	MAPK8	5599	4.34E-02	10	49279692	49313189	33	139	rs11598657	10	49349923	3.5875	3.34E-04
Ingenuity	NFKB.Signaling	IRAK3	11213	7.02E-02	12	64869283	64928652	59	175	rs289068	12	64948270	3.4775	5.06E-04
Ingenuity	NFKB.Signaling	BCL10	8915	7.48E-02	1	85504047	85516171	12	165	rs10489510	1	85555159	3.4131	6.42E-04
Ingenuity	NFKB.Signaling	EGF	1950	1.34E-01	4	111053488	111152870	99	279	rs971695	4	111147377	3.3123	9.25E-04
Ingenuity	NFKB.Signaling	ZAP70	7535	8.10E-02	2	97696462	97722755	26	96	rs17033906	2	97711027	3.268	1.08E-03
Ingenuity	NFKB.Signaling	IKBKB	3551	9.71E-02	8	42247985	42309122	61	118	rs3136755	8	42332944	3.2626	1.10E-03
Ingenuity	NFKB.Signaling	EIF2AK2	5610	1.53E-01	2	37187202	37229907	43	167	rs4670185	2	37209038	3.153	1.62E-03
Ingenuity	NFKB.Signaling	PLCG2	5336	2.98E-01	16	80370430	80549400	179	400	rs11644382	16	80264753	3.1202	1.81E-03
Ingenuity	NFKB.Signaling	LTBR	4055	2.48E-01	12	6363617	6370993	7	102	rs4149573	12	6319645	3.1011	1.93E-03
Ingenuity	NFKB.Signaling	TNFSF13B	10673	2.34E-01	13	107719977	107757366	37	201	rs9559300	13	107796506	3.0498	2.29E-03
Ingenuity	NFKB.Signaling	BTRC	8945	2.24E-01	10	103103814	103307060	203	239	rs11190942	10	103089716	3.0333	2.42E-03
Ingenuity	NFKB.Signaling	CD40	958	2.41E-01	20	44180312	44191791	11	174	rs6065925	20	44167703	3.0208	2.52E-03
Ingenuity	NFKB.Signaling	MAP3K8	1326	2.77E-01	10	30762871	30790767	28	200	rs8177053	10	30769877	2.9617	3.06E-03
Ingenuity	NFKB.Signaling	NFKB1	4790	3.09E-01	4	103641517	103757507	116	269	rs10489113	4	103761249	2.9147	3.56E-03
Ingenuity	NFKB.Signaling	MAP3K7	6885	3.85E-01	6	91282073	91353628	72	355	rs2616027	6	91420783	2.8444	4.45E-03
Ingenuity	NFKB.Signaling	MALT1	10892	3.22E-01	18	54489597	54568350	79	154	rs12326179	18	54434740	2.7848	5.36E-03
Ingenuity	NFKB.Signaling	PRKCZ	5590	3.22E-01	1	1971768	2106694	135	103	rs884080	1	2016609	2.7532	5.90E-03
Ingenuity	NFKB.Signaling	MAP3K7IP1	10454	3.34E-01	22	38125704	38163078	37	140	rs2014842	22	38038955	2.7276	6.38E-03
Ingenuity	NFKB.Signaling	MYD88	4615	3.52E-01	3	38155008	38159516	5	68	rs172111	3	38162998	2.6264	8.63E-03
Ingenuity	NFKB.Signaling	LCK	3932	3.47E-01	1	32489426	32524353	35	41	rs10914530	1	32382054	2.5772	9.96E-03
Ingenuity	NFKB.Signaling	RIPK1	8737	4.82E-01	6	3022056	3060420	38	162	rs4149361	6	2951984	2.5411	1.11E-02
Ingenuity	NFKB.Signaling	TNFAIP3	7128	4.82E-01	6	138230273	138246142	16	120	rs17780048	6	138220839	2.4797	1.32E-02
Ingenuity	NFKB.Signaling	TGFA	7039	6.36E-01	2	70527924	70634613	107	400	rs404420	2	70587034	2.4474	1.44E-02
Ingenuity	NFKB.Signaling	GH1	2688	5.15E-01	17	59348294	59349930	2	80	rs2286565	17	59363964	2.3708	1.78E-02
Ingenuity	NFKB.Signaling	RELA	5970	4.68E-01	11	65178392	65186951	9	70	rs559064	11	65231705	2.3455	1.90E-02
Ingenuity	NFKB.Signaling	TIRAP	114609	6.54E-01	11	125658191	125670038	12	194	rs2962114	11	125613481	2.2102	2.71E-02
Ingenuity	NFKB.Signaling	AZI2	64343	6.11E-01	3	28339089	28365579	26	78	rs9865722	3	28369956	2.1079	3.50E-02
Ingenuity	NFKB.Signaling	TTRAP	51567	7.48E-01	6	24758183	24775094	17	157	rs12181723	6	24867854	2.0075	4.47E-02
Ingenuity	NFKB.Signaling	TRAF2	7186	NaN	9	138900785	138940888	40	97	rs3812614	9	138819515	1.9901	4.66E-02
Ingenuity	NFKB.Signaling	TNFSF11	8600	7.98E-01	13	42043794	42080148	36	204	rs9594773	13	41974695	1.9375	5.27E-02
Ingenuity	NFKB.Signaling	UBE2V1	7335	8.20E-01	20	48131067	48165901	35	101	rs3746559	20	48191598	1.8303	6.72E-02
Ingenuity	NFKB.Signaling	UBE2N	7334	8.26E-01	12	92326218	92360157	34	131	rs7969431	12	92328624	1.827	6.77E-02
Ingenuity	NFKB.Signaling	NFKB2	4791	7.70E-01	10	104144218	104152271	8	53	rs2145308	10	104140989	1.7681	7.71E-02
Ingenuity	NFKB.Signaling	GSK3B	2932	8.87E-01	3	121028235	121295203	267	275	rs13064815	3	121374811	1.761	7.82E-02
Ingenuity	NFKB.Signaling	RELB	5971	8.97E-01	19	50196551	50233292	37	91	rs4803781	19	50151511	1.6009	1.09E-01
Ingenuity	NFKB.Signaling	TRAF6	7189	9.09E-01	11	36467298	36488398	21	153	rs2133165	11	36508823	1.5137	1.30E-01
Ingenuity	NFKB.Signaling	CD40LG	959	NaN	23	135558001	135570215	12	0	NaN	NaN	NaN	NaN	NaN
Ingenuity	NFKB.Signaling	IKBKG	8517	NaN	23	153423652	153446455	23	0	NaN	NaN	NaN	NaN	NaN

Supplementary Table 6. Candidate genes

Gene	Other names	Location	Start	Stop	Article citation	Category
ADAMTS19	*A DISINTEGRIN-LIKE AND METALLOPROTEINASE WITH THROMBOSPONDIN TYPE 1 MOTIF, 19*	5q31	128796103	129074376	Knauff et al 2009, Hum Mol Genet	Ovary/oocyte expressed; Candidate SNP
AIRE	*AUTOIMMUNE REGULATOR*	21q22.3	45705763	45718110	Laml et al 2002, Human Reprod Update	Syndrome
ALOX12	*ARACHIDONATE 12-OXIDOREDUCTASE*	17p13.1	6899384	6914054	Liu et al 2010, Menopause	Menopause/early menopause
AMH	*ANTI-MULLERIAN HORMONE*	19p13.3-p13.2	2249113	2252072	Durlinger et al 1999 Endocrinology	Ovary/ooctye specific expressed
AMHR2	*ANTI-MULLERIAN HORMONE TYPE II RECEPTOR*	12q13	53817641	53825312	Kevenaar et al 2007, Hum Reprod	Menopause/early menopause
AR	*Androgen Receptor*	Xq11.2-q12	66763874	66944119	Shiina et al 2006, PNAS	Animal model
BAX	*BCL2-ASSOCIATED X PROTEIN*	19q13.3-q13.4	49458117	49464519	Greenfield et al 2007, Reproduction	Animal model
BCL2L1	Bcl-X	20q11.21	29717425	29773430	Morita et al 1999, Mol Endocrinol	Ovary/ooctye specific expressed
BDNF	*BRAIN-DERIVED NEUROTROPHIC FACTOR*	11p13	27528399	27699348	Paredes et al 2004, Dev Biol	Ovary/ooctye specific expressed
BICD	*BICAUDAL D, DROSOPHILA, HOMOLOG OF, 1*	12p11.2-p11.1	32260185	32531140	Swan & Suter 1996, Development	Ovary/ooctye specific expressed
BMP15		Xp11.22	50653784	50659606	Pasquale et al 2006, J. Clin. Endocrinol. Metab.	Menopause/early menopause
BMP2	*BONE MORPHOGENETIC PROTEIN 2*	20p12.3	6696745	6708910	Ying et al 2001, Dev Biol	Ovary/ooctye specific expressed
BMP4	*BONE MORPHOGENETIC PROTEIN 4*	14q22-q23	54416457	54421270	Pierre et al 2004, J Molec Endocrin	Animal model
BMP7	*BONE MORPHOGENETIC PROTEIN 7*	20q13.1-q13.3	55743809	55841707	Lee et al 2004, Mol Reprod Dev	Ovary/ooctye specific expressed
BMP8B	*BONE MORPHOGENETIC PROTEIN 8B*	1p34.2	39996490	40027120	Ying et al 2000, Mol Endocrinol	Ovary/ooctye specific expressed
BMPR1A	*BONE MORPHOGENETIC PROTEIN RECEPTOR, TYPE IA*	10q22.3	88516396	88684944	Silva et al 2005, Mol Reprod Dev	Ovary/ooctye specific expressed
BMPR1B	*BONE MORPHOGENETIC PROTEIN RECEPTOR, TYPE IB*	4q22-q24	95679128	96079592	Silva et al 2005, Mol Reprod Dev	Ovary/ooctye specific expressed
BMPR2	*BONE MORPHOGENETIC PROTEIN RECEPTOR, TYPE II*	2q33-q34	203241050	203432473	Silva et al 2005, Mol Reprod Dev	Ovary/ooctye specific expressed
C3orf38	MGC26717(C3orf38)	3p21	88281799	88288755	Rizzolio et al 2007, Hum Genet	Menopause/early menopause
CDKN1a	*p21*	6p21.2	36646459	36655108	Jirawatnotai et al 2003, JBC	Ovary/ooctye specific expressed
CGGBP1		3p21	88101094	88199035	Rizzolio et al 2007, Hum Genet	Menopause/early menopause
CNOT6		5q35	179921417	180005405	Rizzolio et al 2007, Hum Genet	Menopause/early menopause
Connexin 37	*Gap junction protein- a4*	1p35.1	35258599	35261346	Yin et al 2009, Zygote	Ovary/ooctye specific expressed
Connexin 43	*GAP JUNCTION PROTEIN, ALPHA-1*	6q21-q23.2	121756745	121770872	Juneja et al 1999, Biol Reprod	Ovary/ooctye specific expressed
CPEB1	*Cytoplasmic polyadenylation element-binding protein 1*	15q25.2	83211952	83316728	Welk et al 2001, Gene	Ovary/ooctye specific expressed
CXCL12	*Stromal cell-derived factor 1*	10q11.1	44871366	44880542	Ara et al 2003, PNAS	Ovary/ooctye specific expressed
CXCR4	*CHEMOKINE, CXC MOTIF, RECEPTOR 4*	2q21	136871920	136875725	Molyneaux et al 2003, Development	Ovary/ooctye specific expressed
CXCR7	*CHEMOKINE, CXC MOTIF, RECEPTOR 7*	2q37	237478380	237490992	Mahabaleshwar et al 2008, Cell Adh Migr	Ovary/ooctye specific expressed
CYP1B1	*cytochrome P450, family 1, subfamily B, polypeptide 1*		38294116	38337044	Hefler et al 2005, Hum Reprod	Menopause/early menopause
DAZAP1	*DAZ-ASSOCIATED PROTEIN 1*	19p13.3	1407584	1435680	Pan et al 2005, Fert & Steril	Ovary/ooctye specific expressed
DAZL	*Deleted in azoospermia-like*	3p24	16628303	16647006	Tung et al 2006, Reprod. Biol. & Endocrin.	Menopause/early menopause
DIAPH2		Xq22	95939662	96855596	Bione et al 1998, AJHG	Menopause/early menopause
Dicer		14q31	95552565	95608085	Lei et al 2010, Mol Cell Endocrinol	Ovary/ooctye specific expressed
DMC1	*DISRUPTED MEIOTIC cDNA 1, YEAST, HOMOLOG OF*	22q13.1	38914954	38966189	Pittman et al 1998, Mol Cell	Animal model
DNAJC8		ip35.3	28527068	28559536	Rizzolio et al 2007, Hum Genet	Menopause/early menopause
DNMT1	*DNA Methyltransferase 1*	19p13.3-p13.2	10244023	10305755	Rajkovic et al 2004, Science	Animal model
EIF2B1		12q24.3	124105571	124118247		Ovary/ooctye specific expressed
EIF2B2		14q24.3	75469612	75476292	Fogli et al 2003, AJHG	Menopause/early menopause
EIF2B3		1p34.1	45316450	45452282		Ovary/ooctye specific expressed
EIF2B4		2p23.3	27587221	27592919	Fogli et al 2003, AJHG	Menopause/early menopause
EIF2B5		3q27.1	183852810	183863098	Fogli et al 2003, AJHG	Menopause/early menopause
EIF2C2	*Argonaute2*	8q24	141541265	141645645	Pepper et al 2009, PLoS ONE	Animal model
EPB41L5		2q14.2	120493131	120578289	Rizzolio et al 2007, Hum Genet	Menopause/early menopause
EPS8		ch12p12.3	15773076	15942510	D.Toniolo personal communication	Menopause/early menopause
ESR1	*ESTROGEN RECEPTOR 1*	6q25.1	152128454	152424406	Bretherick et al 2008, Fert & Steril	Menopause/early menopause
FANCA	*FANCONI ANEMIA COMPLEMENTATION GROUP A GENE*	16q24.3	89803959	89883065	Cheng et al 2000, Hum Mol Genet	Animal model
FANCC	*FANCONI ANEMIA, COMPLEMENTATION GROUP C*	9q22.3	97861338	98079991	Whitney et al 1996, Blood	Animal model
FANCG	*FANCONI ANEMIA, COMPLEMENTATION GROUP G*	9p13	35073835	35080013	Koomen et al 2002, Hum Mol Genet	Animal model
FGF8	*Fibroblast growth factor 8*	10q24	103529887	103535827	Rajkovic et al 2004, Science	Animal model
FIGLA	*Folliculogenesis specific basic helix-loop-helix*	2p13.3	71004442	71017775	Zhao et al 2008, AJHG & Pangas et al 2006, PNAS	Menopause/early menopause
FMR1	*fragile X mental retardation 1*		146993481	147032645	Mallolas 2001	Menopause/early menopause
FMR2	*FRAXE*	Xq28	147582139	148082192	Murray et al 1998, J Med Genet	Menopause/early menopause
FOXE1	*FORKHEAD BOX E1*	9q22	100615537	100618986	Watkins et al 2006, Molec Hum Rep	Menopause/early menopause
FOXL2	*FORKHEAD BOX L2*	3q22.3	138663067	138665982	Crisponi et al 2001, Nat Gen.	Syndrome

FOXO1A	*Forkhead box O1A*	13q14.1	41129803	41240734	Watkins et al 2006, Fert.Steril.	Menopause/early menopause
FOXO3A (FKHRL1)	*Forkhead Box O3A*	6q21	108926596	109047524	Castrillon et al 2003, Science	Ovary/ooctye specific expressed
FSHB	*FOLLICLE-STIMULATING HORMONE, BETA POLYPEPTIDE*	11p13	30252563	30256823	Kumar et al 1997, Nat Genet	Animal model
FSHR		2p21-p16	49189653	49381630	Ghadami et al 2010, Mol Hum Reprod	Menopause/early menopause; animal model
FST	*Follistatin*	5q11.2	52776595	52781902	Kimura et al 2010, Endocrinology	Ovary/ooctye specific expressed
GALT	*Galactose-1-phosphate uridylytransferase*	9p13	34646635	34650571	Forges et al 2006, Hum. Reprod. Update	Menopause/early menopause
GDF9	*Growth differentiation factor 9*	5q31.1	132196878	132200477	Palmer et al 2006, J. Clin. Endocrinol. Metab.	Menopause/early menopause
GDNF	*GLIAL CELL LINE-DERIVED NEUROTROPHIC FACTOR*	5p13.1-p12	37815754	37839782	Dole et al 2008, Reproduction	Ovary/ooctye specific expressed
GHR	*GROWTH HORMONE RECEPTOR*	5p13-p12	42424026	42721925	Slot et al 2006, Reproduction	Animal model
HDC	*Histidine decarboxylase gene*		50534144	50558223	Zhang et al 2006, Biochem Biophys Res Commun.	Menopause/early menopause
INHA	*INHIBIN, ALPHA*	2q35	220436954	220440427	Shelling et al 2000, Hum Reprod	Menopause/early menopause
INHBA	*INHIBIN, BETA A*	7p14.1	41728603	41742706	Shelling et al 2000, Hum Reprod	Menopause/early menopause
INHBB	*INHIBIN, BETA B*	2q14.2	121103719	121109383	Shelling et al 2000, Hum Reprod	Menopause/early menopause
KIT		4q12	55218852	55301638	Yoshida et al 1997, Dev Biol & Gougeon et al 2010, Ann Endocrinol	Ovary/ooctye specific expressed
KITLG	Kit ligand	12q21.32	87424936	87450411	Parrott and Skinner 1999, Endocrinology, Hutt et al, 2006, MolHumRepro	Ovary/ooctye specific expressed
LHR	*LUTEINIZING HORMONE/CHORIOGONADOTROPIN RECEPTOR*	2p21	48913921	48982880	Latronico et al 1996, NEJM	Menopause/early menopause
LHX8	*LIM homeobox gene 8*	1p31.1	75594119	75627216	Pangas et al 2006, PNAS	Animal model
LHX9	*LIM HOMEOBOX GENE 9*	1q31-q32	197886517	197899273	Mazaud et al 2002, Gene Express Patterns	Animal model
LSH	*HELICASE, LYMPHOID-SPECIFIC*	10q23.3-q24.2	96305574	96361855	De La Fuente et al 2006, Nat Cell Biol	Animal model
MGAT1		5q35	180217541	180242541	Rizzolio et al 2007, Hum Genet	Menopause/early menopause
MOS	*V-mos Moloney murine sarcoma viral oncogene homolog*	8q11	57025502	57026541	Rajkovic et al 2004, Science	Ovary/ooctye specific expressed
MSH4	*MutS, E. COLI, HOMOLOG OF, 4*	1p31	76262630	76378923	Kneitz et al 2000, Genes Dev	Animal model
MSH5	*MutS, E. COLI, HOMOLOG OF, 5*	chr6_mcf_hap5	3087458	3110642	de Vries et al 1999, Genes Dev	Animal model
MSY2	*Y-box binding protein 2*	17p11.2-p13.1	7191571	7197876	Yang et al 2005, PNAS	Animal model
NANOS3	*nanos homolog 3 (Drosophila)*	19p13.12	13987950	13991570	Yingying et al 2007, Fert. Steril.	Ovary/ooctye specific expressed
NGF	*NERVE GROWTH FACTOR, BETA SUBUNIT*	1p13.1	115828537	115880857	Dissen et al 2001, Endocrinolgy	Animal model
NOBOX	*NOBOX oogenesis homeobox*	7q35	144096041	144107320	Qin et al 2007, AJHG & Pangas et al 2006, PNAS	Menopause/early menopause
NOG	*Noggin*	17q23.2	54671060	54672951	Kosaki et al 2004, Fert & Steril	Menopause/early menopause
NR5A1	*NUCLEAR RECEPTOR SUBFAMILY 5, GROUP A, MEMBER 1*	9q33	127243515	127269699	Lourenço et al 2010, New Eng J Med	Menopause/early menopause
NRF2		15q26			D.Toniolo personal communication	Menopause/early menopause
NT4	*NEUROTROPHIN 4*	19q13.3	49564399	49567119	Paredes et al 2004, Dev Biol	Animal model
NTRK1	*NEUROTROPHIC TYROSINE KINASE, RECEPTOR, TYPE 1*	1q21-q22	156830671	156851642	Kerr et al 2009, Reproduction	Animal model
NTRK2	*NEUROTROPHIC TYROSINE KINASE, RECEPTOR, TYPE 2*	9q22.1	87283466	87638505	Kerr et al 2009, Reproduction	Animal model
NXF5	*NUCLEAR RNA EXPORT FACTOR 5*	Xq22.1	101087085	101112549	Bertini et al 2010, Fert & Steril	Menopause/early menopause
OCT4	*Octomer binding protein 4 or POU5F1*	6p21.31	2514038	2520393	Rajkovic et al 2004, Science	Ovary/ooctye specific expressed
p27kip1	*CDKN1B*	12p13.1-p12	12870302	12875305	Rajareddy et al 2007, Molec. Endo.	Animal model
PDPK1	*3-@PHOSPHOINOSITIDE-DEPENDENT PROTEIN KINASE 1*	16p13.3	2587970	2653188	Reddy et al 2009, Hum Mol Genet	Animal model
PGRMC1	*Progesterone Receptor Membrane Component-1*	Xq22-q24	118370211	118378428	Mansouri et al 2008, Hum Mol Gen	Menopause/early menopause
PIN1	*PEPTIDYL-PROLYL CIS/TRANS ISOMERASE, NIMA-INTERACTING, 1*	19p13	9945999	9960358	Atchison et al 2003, Development	Animal model
PMM2	*PHOSPHOMANNOMUTASE 2*	16p13.3-p13.2	8891670	8943193	Laml et al 2002, Human Reprod Update	Syndrome
POF1B		Xq21	84532395	84634748	Lacombe et al 2006	Menopause/early menopause
POG	*PHD FINGER PROTEIN 9*	2p16.1	58386380	58468515	Agoulnik et al 2002, Hum Mol Genet	Ovary/ooctye specific expressed
POLG	*POLYMERASE, DNA, GAMMA*	15q25	89859537	89878026	Luoma et al 2004, The Lancet	Syndrome
PTEN	*Phosphate and tensin homolog*	10q23.31	89623195	89728531	Reddy et al 2008, Science	Animal model
PTHB1	*PARATHYROID HORMONE-RESPONSIVE B1 GENE*	7p14	33169152	33645680	Kang et al 2008, Hum Reprod	Menopause/early menopause
PTPRO		ch12p12.3	15475451	15750335	D.Toniolo personal communication	Menopause/early menopause
RPL10	*RIBOSOMAL PROTEIN L10*	Xq28	153626571	153630680	Massad-Costa et al 2007, Maturitas	Menopause/early menopause
RAD51C		17q	56769963	56811690	Meindl et al 2010, Nat Genet	Ovary/ooctye specific expressed
RALB		2q14.2	120997640	121052284	Rizzolio et al 2007, Hum Genet	Menopause/early menopause
REC8	*REC8, S. POMBE, HOMOLOG OF*	14q11.2-q12	24641234	24649462	Xu et al 2005, Devel Cell	Animal model
RERG		ch12p12.3	15260718	15374304	D.Toniolo personal communication	Menopause/early menopause
RFPL4	*RET FINGER PROTEIN-LIKE 4A*	19q13.4	56270507	56274539	Suzumori et al 2003, PNAS	Ovary/ooctye specific expressed
RPA2		ip35.3	28218035	28241257	Rizzolio et al 2007, Hum Genet	Menopause/early menopause
RSPO1	*R-Spondin-1*	1p34.3	38076951	38100491	Chassot et al, + Tomizuka et al, 2008, HMG	Animal model
SALL4	*SAL-LIKE 4*	20q13.13-q13.2	50400585	50419048	Wang et al 2009, Mol Hum Reprod	Menopause/early menopause

SMAD1	*MOTHERS AGAINST DECAPENTAPLEGIC, DROSOPHILA, HOMOLOG OF, 1*	4q28	146402951	146480323	Tremblay et al 2001, Development	Animal model
SMAD3	*MOTHERS AGAINST DECAPENTAPLEGIC, DROSOPHILA, HOMOLOG OF, 3*	15q21-q22	67358195	67487532	Tomic et al 2002, Biol Reprod	Animal model
SMAD5	*MOTHERS AGAINST DECAPENTAPLEGIC, DROSOPHILA, HOMOLOG OF, 5*	5q31	135468536	135518422	Chang & Matzuk 2001, Mech Dev	Animal model
SMPDL3B		1p35.3	28261504	28285668	Rizzolio et al 2007, Hum Genet	Menopause/early menopause
SOHLH1	*Spermatogenesis-and oogenesis specific basic helix-loop-helix protein 1*	9q34.3	138585257	138591374	Pangas et al 2006, PNAS	Animal model
SOHLH2	*Spermatogenesis-and oogenesis specific basic helix-loop-helix protein 2*	13q13.3	36742347	36788752	Suzumori et al 2007, Curr. Med. Chem.	Ovary/ooctye specific expressed
SPO11	*SPO11, S. CEREVISIAE, HOMOLOG OF*	20q13.2-q13.3	55904831	55919048	Romanienko & Camerini-Otero 2000, Mol Cell	Animal model
STRAP		ch12p12.3	16035288	16056403	D.Toniolo personal communication	Menopause/early menopause
TAF4B	*TAF4b RNA polymerase II, TATA box binding protein- associated factor*	18q11.2	23806409	23971647	Suzumori et al 2007, Curr. Med. Chem.	Ovary/ooctye specific expressed
TIAR	*TIA1 CYTOTOXIC GRANULE-ASSOCIATED RNA-BINDING PROTEIN-LIKE 1*	10q	121332978	121356541	Beck et al 1998, PNAS	Animal model
TMEM35	*Xp18 (FLJ14084)*	Xq22.1	100333836	100351353	Cachot et al 2003, Gene	Ovary/ooctye specific expressed
TSC2	*TUBEROUS SCLEROSIS 2*	16p13	2097990	2138712	Adhikari et al 2009, Mol Hum Reprod	Animal model
Wnt4	*WINGLESS-TYPE MMTV INTEGRATION SITE FAMILY, MEMBER 4*	1p35	22443800	22469519	Jeays-Ward, 2003, Development	Ovary/ooctye specific expressed
XPNPEP2	*X-prolylaminopeptidase 2 or APP2*	Xq26.1	128872946	128903525	Prueitt et al 2002, Cytogenet Genome Res	Menopause/early menopause
ZAR1	*Zygote arrest 1*	4p11	48492309	48496420	Rajkovic et al 2004, Science	Ovary/ooctye specific expressed
ZFX	*ZINC FINGER PROTEIN, X-LINKED*	Xp22.2-p21.3	24169808	24232627	Luoh et al 1997, Development	Animal model
ZNF654		3p11	88188254	88193815	Rizzolio et al 2007, Hum Genet	Menopause/early menopause

Supplementary Table 7. Pleietropy of Menopause associations.

							Menopause Primary Analysis					Other Phenotypes				
Hit number	SNPID	Chrom	Location (bp)	Region	Minor/ Major	MAF	Effect per minor allele (years)	SE	P-value	Absolute Effect per minor allele (weeks)	Hetero-geneity P-value	Phenotype	SNP	hapmap r2	Position(b36)	Distance
3	rs2303369	2	27568920	2p23.3	t/c	0.39	-0.174	0.030	3.80E-09	9.0	0.639	eGFRcrea	rs1260326	0.48	27584444	15524
3	rs2303369	2	27568920	2p23.3	t/c	0.39	-0.174	0.030	3.80E-09	9.0	0.639	Albumin	rs1260326	0.48	27584444	15524
3	rs2303369	2	27568920	2p23.3	t/c	0.39	-0.174	0.030	3.80E-09	9.0	0.639	2Hr Glucose	rs1260326	0.48	27584444	15524
3	rs2303369	2	27568920	2p23.3	t/c	0.39	-0.174	0.030	3.80E-09	9.0	0.639	FG/FI & T2D	rs780094	0.45	27594741	25821
3	rs2303369	2	27568920	2p23.3	t/c	0.39	-0.174	0.030	3.80E-09	9.0	0.639	CRP	rs780094	0.45	27594741	25821
3	rs2303369	2	27568920	2p23.3	t/c	0.39	-0.174	0.030	3.80E-09	9.0	0.639	Serum Urate	rs780093	0.45	27596107	27187
3	rs2303369	2	27568920	2p23.3	t/c	0.39	-0.174	0.030	3.80E-09	9.0	0.639	Triglycerides	rs1260333	0.52	27602128	33208
5	rs7606918	2	172603695	2q31.1	g/a	0.16	-0.228	0.041	2.89E-08	11.8	0.374	Prostate Cancer	rs12621278	0.01	173019799	416104
5	rs7606918	2	172603695	2q31.1	g/a	0.16	-0.228	0.041	2.89E-08	11.8	0.374	Longevity	rs6433379	-	173328711	725016
7	rs890835	5	175888877	5q35.2	a/c	0.11	0.266	0.047	1.17E-08	13.8	0.003	eGFRcrea	rs6420094	-	176750242	861365
7	rs890835	5	175888877	5q35.2	a/c	0.11	0.266	0.047	1.17E-08	13.8	0.003	F12	rs2731672	-	176775080	886203
8	rs365132	5	176311180	5q35.2	t/g	0.49	0.275	0.029	1.90E-21	14.3	0.115	eGFRcrea	rs6420094	0.001	176750242	439062
8	rs365132	5	176311180	5q35.2	t/g	0.49	0.275	0.029	1.90E-21	14.3	0.115	F12	rs2731672	0.01	176775080	463900
10	rs1046089	6	31710946	6p21.33	a/g	0.35	-0.226	0.031	1.31E-13	11.8	0.426	Lung adenocarcinoma	rs3117582	0.21	31728499	17553
10	rs1046089	6	31710946	6p21.33	a/g	0.35	-0.226	0.031	1.31E-13	11.8	0.426	Weight	rs2844479	0.13	31680935	30011
10	rs1046089	6	31710946	6p21.33	a/g	0.35	-0.226	0.031	1.31E-13	11.8	0.426	Neonatal lupus	rs3099844	0.13	31556955	153991
10	rs1046089	6	31710946	6p21.33	a/g	0.35	-0.226	0.031	1.31E-13	11.8	0.426	HIV-1 control	rs2395029	0.03	31539759	171187
10	rs1046089	6	31710946	6p21.33	a/g	0.35	-0.226	0.031	1.31E-13	11.8	0.426	Drug-induced liver injury	rs2395029	0.03	31539759	171187
10	rs1046089	6	31710946	6p21.33	a/g	0.35	-0.226	0.031	1.31E-13	11.8	0.426	Height	rs13437082	0.001	31462539	248407
10	rs1046089	6	31710946	6p21.33	a/g	0.35	-0.226	0.031	1.31E-13	11.8	0.426	Ankylosing spondylitis	rs7743761	0.01	31444079	266867
10	rs1046089	6	31710946	6p21.33	a/g	0.35	-0.226	0.031	1.31E-13	11.8	0.426	AMD	rs429608	0.01	32038441	327495
10	rs1046089	6	31710946	6p21.33	a/g	0.35	-0.226	0.031	1.31E-13	11.8	0.426	Nasopharyngeal carcinoma	rs2894207	0.004	31371730	339216
10	rs1046089	6	31710946	6p21.33	a/g	0.35	-0.226	0.031	1.31E-13	11.8	0.426	Psoriasis	rs12191877	0	31360904	350042
10	rs1046089	6	31710946	6p21.33	a/g	0.35	-0.226	0.031	1.31E-13	11.8	0.426	CD4:CD8 lymphocyte ratio	rs2524054	0.01	31360375	350571
10	rs1046089	6	31710946	6p21.33	a/g	0.35	-0.226	0.031	1.31E-13	11.8	0.426	Ulcerative colitis	rs9263739	0.04	31219335	491611
10	rs1046089	6	31710946	6p21.33	a/g	0.35	-0.226	0.031	1.31E-13	11.8	0.426	White Blood Cell Count	rs3094212	-	31193749	517197
12	rs12294104	11	30339475	11p14.1	t/c	0.17	0.226	0.040	1.63E-08	11.8	0.721	serum magnesium	rs3925584	0.02	30716911	377436
13	rs2277339	12	55432336	12q13.3	g/t	0.10	-0.394	0.051	5.99E-15	20.5	0.088	Psoriasis	rs2066808	0.01	55024240	408096
13	rs2277339	12	55432336	12q13.3	g/t	0.10	-0.394	0.051	5.99E-15	20.5	0.088	urate	rs1106766	-	56095723	663387
13	rs2277339	12	55432336	12q13.3	g/t	0.10	-0.394	0.051	5.99E-15	20.5	0.088	T1D	rs2292239	-	54768447	663889
13	rs2277339	12	55432336	12q13.3	g/t	0.10	-0.394	0.051	5.99E-15	20.5	0.088	Alopecia areata	rs1701704	-	54698754	733582
14	rs3736830	13	49204222	13q14.3	g/c	0.16	-0.243	0.040	1.75E-09	12.6	0.859	Height	rs3118914	-	50014902	810680
17	rs10852344	16	11924420	16p13.13	c/t	0.42	0.198	0.029	1.28E-11	10.3	0.014	QT interval	rs8049607	0	11599254	325166
17	rs10852344	16	11924420	16p13.13	c/t	0.42	0.198	0.029	1.28E-11	10.3	0.014	Celiac disease	rs12928822	-	11311394	613026
17	rs10852344	16	11924420	16p13.13	c/t	0.42	0.198	0.029	1.28E-11	10.3	0.014	T1D	rs12708716	-	11087374	837046
18	rs11668344	19	60525476	19q13.42	g/a	0.36	-0.416	0.030	5.94E-43	21.6	0.112	Platelet aggregation	rs1671152	0	60218157	307319
19	rs12461110	19	61012475	19q13.42	a/g	0.36	-0.174	0.030	9.49E-09	9.1	0.835	Platelet aggregation	rs1671152	-	60218157	794318
20	rs16991615	20	5896227	20p12.3	a/g	0.07	0.971	0.062	1.16E-54	50.5	0.356	Colorectal cancer	rs961253	0.01	6352281	456054
20	rs16991615	20	5896227	20p12.3	a/g	0.07	0.971	0.062	1.16E-54	50.5	0.356	Height	rs967417	-	6568893	672666

Nature Genetics: doi:10.1038/ng.1051

Supplementary Note

Methods

Study subjects – Discovery stage

AGES-Reykjavik Study, the Reykjavik Study cohort originally comprised a random sample of 30,795 men and women born in 1907-1935 and living in Reykjavik in 1967. A total of 19,381 people participated in the Reykjavik Study examination, a 71% recruitment rate. The study sample was divided into six groups by birth year and birth date within month. One group was invited to participate in all subsequent examinations, while one group was designated as a control group and was not included in examinations until 1991. Other groups were invited to participate in specific examinations of the study. Between 2002 and 2006, the AGES-Reykjavik Study re-examined 5764 survivors of the original Reykjavik Study[1]. Successful genotyping was available for 1849 AGES women participants who were eligible for this study. The AGES-Reykjavik Study GWAS was approved by the National Bioethics (VSN: 00-063) and the Data Protection Authority and also was covered under the MedStar Institutional Review Board. All subjects provided written informed consent.

The **ARIC** study is a population-based, prospective cohort study of cardiovascular disease and its risk factors sponsored by National Heart, Lung and Blood Institute (NHLBI)[2]. ARIC recruited 15,792 individuals aged 45-64 years at baseline (1987-89), chosen by probability sampling from four US communities. Cohort members completed four clinic examinations, conducted three years apart between 1987 and 1998. Annual follow-up by telephone is on-going.

The **Cardiovascular Health Study** (**CHS**) is a population-based study of risk factors for cardiovascular disease in older adults, sponsored by the National Heart, Lung and Blood Institute. Men and women aged 65 and older were recruited from random samples of Health Care Financing Administration eligibility lists in four U. S. communities[3,4].

The information for **deCODE** was originating from two sources. It was collected in a nationwide cancer screening program through the Cancer Detection Clinic at the Icelandic Cancer society since 1964 to 1989 and in a study on bone mineral density led by deCODE Genetics with a recruitment period between 2000 and 2006. The question referred to age at previous birthday before the last menstruations. Of these individuals, 5,857 were genotyped on an Illumina 317K/ 370 K SNP chip in one of several genome-wide association studies recently conducted by deCODE Genetics and had a reported age at last menstruation between 36 and 57 years. All the women were born before 1948 and were not known to have a cancer of the ovary or of the uterus according to

Nature Genetics: doi:10.1038/ng.1051

the Icelandic Cancer Register (period covered: 1955-2007) All of these studies were approved by the Data Protection Commission of Iceland and the National Bioethics Committee of Iceland. Written informed consent was obtained from all participants. Personal identifiers associated with phenotypic information and blood samples were encrypted using a third-party encryption system as previously described. Only individuals with a genotype yield over 98% were included in the study.

The **EGCUT** cohort is from the population-based biobank of the **Estonian Genome Project of University of Tartu**. The whole project is conducted according to Estonian Gene Research Act and all participants have signed the broad informed consent (www.geenivaramu.ee and Metspalu, 2004[5]). The cohort size is 51,515 from 18 years of age and up which reflects closely the age distribution in the Estonian population, 33% male, 67% female, 83% Estonians, 14% Russians, 3% other. Subjects are recruited by the general practitioners (GP) and physicians in the hospitals were randomly selected from individuals visiting GP offices or hospitals. Computer Assisted Personal interview (CAPI) is filled during 1-2 hours at doctors office including personal data (place of birth, place(s) of living, nationality etc.), genealogical data (family history, four generations), educational and occupational history, lifestyle data (physical activity, dietary habits, smoking, alcohol consumption, women´s health, quality of life), also anthropometric and physiological measurents are taken. Subjects for the GWAS were selected randomly all over the country[6].

The **Erasmus Rucphen Family study** (**ERF**) is part of the Genetic Research in Isolated Population program. The study population essentially consists of one extended family of descendents from 20 related couples who lived in the isolate between 1850 and 1900 and had at least 6 children baptized in the community church. The detailed information about ERF isolate can be found elsewhere[7]. The Medical Ethical Committee of the Erasmus Medical Center, Rotterdam approved the study and informed consent was obtained from all participants.

The **Framingham Heart Study** (**FHS**) was initiated in 1948 to study determinants of cardiovascular disease and other major illnesses. The Original Cohort included 2873 women, aged 28-62 years at enrollment who have undergone routine biennial examinations[8,9]. In 1971, the offspring of the Original Cohort participants and offspring spouses, including 2641 women, aged 5 to 70 years, were enrolled into the Framingham Offspring Study. Offspring participants have been examined approximately every 4 years[10,11]. In the 1990s, DNA was obtained for genetic studies from surviving Original Cohort and Offspring participants.

The **Heredity and Phenotype Intervention (HAPI) Heart Stud**y was initiated in 2002. Participants of the HAPI Heart Study comprised adults from the Old Order Amish community of Lancaster County, PA, who were recruited over a three-year period. Study participants were included if they were aged 20 years and older and considered to be relatively healthy based on exclusion criteria of severe hypertension (blood pressure >180/105 mm Hg), malignancy, and kidney, liver or untreated thyroid disease. The study aims and recruitment

Nature Genetics: doi:10.1038/ng.1051

details, including ascertainment criteria, have been described previously[12]. Physical examinations were conducted at the Amish Research Clinic in Strasburg, PA and a reproductive health questionnaire was completed by female participants. Women presenting pregnant or within 6 months postpartum were excluded from the study.

The **InCHIANTI** study is a population-based epidemiological study aimed at evaluating factors that influence mobility in the older population living in the Chianti region of Tuscany, Italy. Details of the study have been previously reported[13]. Briefly, 1616 residents were selected from the population registry of Greve in Chianti (a rural area: 11,709 residents with 19.3% of the population greater than 65 years of age) and Bagno a Ripoli (Antella village near Florence; 4704 inhabitants, with 20.3% greater than 65 years of age). The participation rate was 90% (n= 1453) and participants ranged between 21–102 years of age. The study protocol was approved by the Italian National Institute of Research and Care of Aging Institutional Review. There were 85 parent-offspring pairs, 6 sib-pairs and 2 halfsibling pairs documented. We investigated any further familial relationships using IBD of 10,000 random SNPs using RELPAIR and uncovered 1 parent-offspring, 79 siblings and 13 half-sibling[14]. We utilized the correct family structure inferred from genetic data for all analyses.

Details of the **Nurses' Health Study** (**NHS**) cohorts have been described previously[15]. Briefly, the NHS was initiated in 1976, when 121,700 United States registered nurses between the ages of 30 and 55, residing in 11 larger U.S. states, returned an initial questionnaire reporting medical histories and baseline health-related exposures, including information related to reproductive history (age at menarche, age at first birth, parity, age at menopause etc.), and exposure to exogenous hormones (oral contraception or post-menopausal hormone replacement therapy). Biennial questionnaires with collection of exposure information on risk factors have been collected prospectively, and outcome data with follow-up of reported disease events are collected. From May 1989 through September 1990, we collected blood samples from 32,826 participants in the NHS cohort. Subsequent follow-up has been greater than 99% for this subcohort. Informed consent was obtained from all participants. The study was approved by the Institutional Review Board of the Brigham and Women's Hospital, Boston, MA, USA. **NHS breast cancer GWAS**: The NHS nested breast cancer case-control study was derived from the 32,826 women in the blood subcohort who were free of diagnosed breast cancer at blood collection and followed for incidence disease until June 1, 2004. Breast cancer follow-up in the NHS was conducted by personal mailings and searches of the National Death Index. Controls were women not diagnosed with breast cancer during follow-up, and were one-to-one matched to cases based on age at diagnosis, blood collection variables (time of day, season, and year of blood collection, as well as recent (<3 months) use of postmenopausal hormones), ethnicity (all cases and controls are self-reported Caucasians), and menopausal status (all cases were postmenopausal at diagnosis). The 2,287 NHS participants included in the present analysis were from this nested breast cancer case-control study and were self-described Caucasians with genotype data

Nature Genetics: doi:10.1038/ng.1051

available from the National Cancer Institute's Cancer Genetic Marker of Susceptibility (CGEMS) project[16]. **NHS type 2 diabetes (T2D) GWAS**: NHS participants for the current T2D GWAS were also selected among those with a blood sample using a nested case-control design[17]. Diabetes cases were defined as self-reported diabetes confirmed by a validated supplementary questionnaire. For cases before 1998, diagnosis was made using criteria consistent with those proposed by the National Diabetes Data Group (NDDG)[18]. We used the American Diabetes Association diagnostic criteria for diagnosis of diabetes cases during the 1998 and 2000 cycles[19]. 98% of self-reported cases were confirmed by medical records review in this cohort[20]. Controls were defined as those free of diabetes at the time of diagnosis of the case and remained unaffected through follow-up (2006). Although controls were originally matched per case (by gender, year of birth, month of blood collection, and fasting status), matched pairs were broken because not all subjects gave informed consent for submission of their GWAS data to dbGaP.

As part of a longitudinal survey study of health, lifestyle and personality, twins and their family members registered with the **Netherlands Twin Register** (**NTR**) are approached every 2 to 3 years[21]. As part of a case-control study for major depression disorder, genotype information was obtained in 1940 NTR participants[22]. In the surveys, women were asked whether they had reached menopause. If yes, participants were asked at what age the menopause started. Inconsistencies over time were checked. Age of menopause was available for 331 women, who reached menopause due to natural causes.

QIMR, data on age at menopause comes from a population-based cohort initially recruited from the Australian Twin Registry in 1980. At the time of recruitment, twins were aged between 17 and 88. We estimate that this represents 10-20% of living twins in Australia. Initially, female participants were mailed a health questionnaire that included questions about whether their periods had stopped for at least one year, and the reason that they had stopped. The cohort was followed up 8 years later for a study of drinking habits and participants were asked the same questions. A subsample of this cohort was followed up again in 1993 to answer a questionnaire relating to women's health. The questions related to risk factors for hysterectomy. Women who reported using hormone replacement therapy prior to menopause (n =116 cases) or had undergone a hysterectomy (n = 134) were removed from the analysis. The follow-up studies facilitated checking of inconsistencies from the original questionnaires. A total of 430 women provided both phenotypic and genotype information. The mean age at menopause in the sample was 48.3 with a standard deviation of 4.4 years. The mean age of women in the genotyped sample was 58.3 (5.7) .A detailed description of the genotyping, QC and imputation procedures are given elsewhere[23]. Detailed information on the phenotype collection is given in[24] Informed consent was obtained from all participants.

Rotterdam Study I, II, and **III** (**RSI, RSII, RSIII**) are ongoing prospective population-based cohort studies of Caucasian subjects aged 45 years and over, living in the Ommoord district of Rotterdam, the

Nature Genetics: doi:10.1038/ng.1051

Netherlands. The study was designed to investigate the incidence and determinants of chronic disabling diseases in the elderly. Rationale and design have been described previously[25-27]. For RSI, all 10,275 inhabitants aged 55 years and over were invited for baseline examination between August 1990 and June 1993, of those, 7,983 participated[27]. In 1999, 3,011 participants (out of 4,472 invitees) who had become 55 years of age or moved into the study district since the start of the study were added to the cohort (RSII). In 2006, a further extension of the cohort was initiated in which about 6,000 subjects aged 45–54 years, living in the Ommoord district, were invited, of which 3,932 participated (RSIII)[25,26]. Questionnaires including menopause related questions were filled out by study nurses during the home interview, while blood samples were taken of over 70% of the participants at the research centre. The Rotterdam Study was approved by the medical ethics committee of the Erasmus University Medical School, and written informed consent was obtained from each subject. The current study is based on 2,196 women from RSI, 665 from RSII and 597 from RSIII for whom GWAS data and age at natural menopause was available.

The **SardiNIA** genome wide association study has been described in detail previously[28,29]. Briefly, the GWA study examined a total of 4,305 related individuals participating in a longitudinal study of aging-related quantitative traits in the Ogliastra region of Sardinia, Italy. Genotyped individuals had four Sardinian grandparents and were selected for genotyping without regard to their phenotypes. Among the individuals examined, 1,412 were genotyped with the Affymetrix Mapping 500K Array Set. A total of 356,359 autosomal SNPs met the quality control criteria and were used as input for the imputation procedure using the software MACH[30,31]. The remaining 2,893 individuals were genotyped with the Affymetrix Mapping 10K Array. These individuals were mostly offspring and siblings of the 1,412 individuals that were genotyped with the Affymetrix Mapping 500K Array Set. We took advantage of the relatedness among individuals to impute missing genotypes in these additional individuals; we identified large stretches of chromosome shared within each family and probabilistically "filled-in" genotypes within each stretch whenever one or more of its carriers was genotyped with the 500K Array Set[30,32]. In order to more efficiently evaluate identity-by-descendent states at non-overlapping markers, 436 individuals out of the 1,412 were also genotyped with the 10K Array. Among the 4,305 genotyped individuals, 828 women were in menopause and their phenotype was used for analysis.

The **Study of Health in Pomerania** (**SHIP**) is a population-based study in West Pomerania, a region in Northeast Germany. Baseline examination started in October 1997 and was finished in March 2001. The net sample comprised 6267 eligible subjects, and 4310 (68.8% of eligible subjects) participated[33]. All participants gave written informed consent. The study conformed to the principles of the Declaration of Helsinki as reflected by an *a priori* approval of the Ethics Committee of the University of Greifswald. The 4105 SHIP samples were genotyped using the Affymetrix Human SNP Array 6.0. Hybridization of genomic DNA was done in accordance with the manufacturer's standard recommendations. Genotypes were determined using the

Nature Genetics: doi:10.1038/ng.1051

Birdseed2 clustering algorithm. For quality control purposes, several control samples where added. On the chip level, only subjects with a genotyping rate on QC probesets (QC call rate) of at least 86% were included. The overall genotyping efficiency of the GWA was 98.56%. The genetic data analysis workflow was created using the Software InforSense. Genetic data were stored using the database Caché (InterSystems).

The **TwinsUK** cohorts consisted of a group of twins ascertained to study the heritability and genetics of age-related diseases (www.twinsUK.ac.uk). These unselected twins were recruited from the general population through national media campaigns in the UK and shown to be comparable to age-matched population singletons in terms of disease-related and lifestyle characteristics[34,35].

The **Women's Genome Health Study** (**WGHS**) is a prospective cohort of female healthcare professionals, aged 45 or older at baseline, who provided baseline blood sample and consent for blood based analysis in the Women's Health Study (WHS), a randomized, placebo controlled trial of aspirin and vitamin E in the primary prevention of cardiovascular disease and cancer[36].

Study subjects - Replication stage

Full details of the selection of participants and measurements for **BWHHS** have been previously reported[37]. Women aged 60-79 years were randomly selected from general practitioner lists in 23 British towns. A total of 4286 women (60% of those invited) participated and baseline data (self-completed questionnaire, research nurse interview, physical examination and primary care medical record review) were collected between April 1999 and March 2001. Local ethics committee approvals were obtained. Age at menopause, use of hormone replacement therapy and history of a hysterectomy or oophorectomy were obtained from the self completed questionnaire and/or the research nurse interview. Women were asked to report their age, in years, at which they had experienced their last menstrual period. In a separate question women were asked to report all operations they had ever undergone and the timing of these operations. These data were used to identify women who had had a hysterectomy or oophorectomy and the timing of these operations. Women were asked if they had ever used hormone replacement therapy and if so the name of the therapy and their age at commencing and stopping (if no longer using) the therapy. At the research nurse interview women were asked to bring all of their current medications and for women who were currently using hormone replacement their self-report of hormone replacement use was verified at the nurse interview drugs history.

Full details of the **CoLaus** study have been published by Firmann et al.[38]. The CoLaus study is a population-based study aimed at investigating the epidemiology and genetic determinants of cardiovascular risk factors and metabolic syndrome. The cohort is a random sample of the Lausanne population aged between 35 and 75. Age at menopause was self-reported some time between 2003 and 2006. Available sample size is 1013 with mean sample age 61 and mean age at menopause 50 years.

Nature Genetics: doi:10.1038/ng.1051

The **EGCUT** cohort is from the population-based biobank of the **Estonian Genome Project of University of Tartu**. The whole project is conducted according to Estonian Gene Research Act and all participants have signed the broad informed consent (www.geenivaramu.ee and Metspalu, 2004[5]). The cohort size is 51,515 from 18 years of age and up which reflects closely the age distribution in the Estonian population, 33% male, 67% female, 83% Estonians, 14% Russians, 3% other. Subjects are recruited by the general practitioners (GP) and physicians in the hospitals were randomly selected from individuals visiting GP offices or hospitals. Computer Assisted Personal interview (CAPI) is filled during 1-2 hours at doctors office including personal data (place of birth, place(s) of living, nationality etc.), genealogical data (family history, four generations), educational and occupational history, lifestyle data (physical activity, dietary habits, smoking, alcohol consumption, women´s health, quality of life), also anthropometric and physiological measurents are taken. Subjects for the GWAS were selected randomly all over the country[6].

The **EPOS** study is a cross-sectional study and includes 5,896 pre-, peri- and post-menopausal women all living in the city of Eindhoven, the Netherlands[39]. For 1,500 women DNA was available. Women who had hysterectomy, ovariectomy or were using hormones at time of menopause were excluded for this study. After follow-up, 903 women with natural menopause between 40 and 60 were included in this study, based on self-reported age at natural menopause. The menopausal age for this study was collected both retrospectively and prospectively.

The **Genetic Epidemiology Network of Arteriopathy** (**GENOA**) study is a community-based study of hypertensive sibships that aims to identify genes influencing blood pressure[40,41]. In the initial phase of the GENOA study (9/1995 to 6/2001), sibships containing ≥ 2 individuals with essential hypertension diagnosed before age 60 years were selected for participation. At the Rochester, MN field center, 1583 non-Hispanic whites were enrolled. Participants returned in a second phase of GENOA (12/2000 to 6/2004) for physical examination and measurement of non-conventional and novel risk factors.

The **Helsinki Birth Cohort Study** (**HBCS**) is composed of 8 760 individuals born between the years 1934-44 in one of the two main maternity hospitals in Helsinki, Finland. Between 2001 and 2003, a randomly selected sample of 928 males and 1075 females participated in a clinical follow-up study with a focus on cardiovascular, metabolic and reproductive health, cognitive function and depressive symptoms. There were 908 women with valid genotype data and data on self-reported last menstrual period. All women with menopause before age 40 or after age 60 (n=5), surgical menopause before last menstrual period (n=217), hormone replacement therapy before last menstrual period (n=91), cancer before last menstrual period (n=26), and incomplete data (n=13) were removed from the analyses. After these exclusions 556 women were included in the analyses. The mean age of the participants was 61.5 years (SD=3.1) and mean age of menopause was 50.7 years (SD=3.9). Detailed information on the selection of the HBCS participants and on the study design

Nature Genetics: doi:10.1038/ng.1051

can be found elsewhere[42-44]. Research plan of of the HBCS was approved by the Institutional Review Board of the National Public Health Insitute and all participants have signed an informed consent.

INGI-Carlantino project (**INGI-CARL**), the Carlantino population, located in Southern Italy, in the extreme Northern part of Puglia, was invited to participate. We enrolled 1417 people. We genotyped all people using Illumina 370K, the data were imputated with MACH. All analysis were performed using GenABEL. A consent form either for clinical and genetic studies has been signed by each participant in the study.

INGI-Friuli Venezia Giulia project, (**INGI-FVG**) The Friuli Venezia Giulia Project was initiated in 2008. It is aimed to study six different isolated villages situated in North-Eastern Italy (FVG). The cohort included 1739 people. The partecipation was random not family-based. We genotyped all people using Illumina 370K, the data were imputated with MACH. All analysis were performed using GenABEL. Subjects gave their written informed consent for partecipating in this project.

INGI-Val Borbera project was initiated in 2005 and represents the collection of phenotypic and genotypic data from a geographically isolated population of Northern Italy living in the Val Borbera Valley in Piedmont. Inhabitants of the valley were invited to participate in the study by public advertisements through local authorities, televisions and newspapers as well as local physicians and mailings. Meetings were organized in all villages to present the project and its aims. The importance of the participation of entire families was underscored in all instances, nevertheless all people that volunteered to participate were included in the study, providing they had at least one grand parent from the valley. The study, including the overall plan and the informed consent form was reviewed and approved by the institutional review boards of San Raffaele Hospital in Milan and by the Regione Piemonte ethical committee. Information and biological samples were obtained from 1803 inhabitants between 18 and 102 years of age[45]. 1664 DNAs were genotyped with the genome-wide 370k Illumina chip. Only individuals aged 18 years or older were eligible. The cohort includes 930 women.

KORA F3 and **KORA S4.** The Cooperative Health Research in the Region of Augsburg (KORA) study is a series of independent population-based epidemiological surveys and follow-up studies of participants living in the region of Augsburg, Southern Germany45. All participants are residents of German nationality identified through the registration office and were examined in 1994/95 (KORA S3) and 1999/2001 (KORA S4). In the KORA S3 study 4,856 subjects (response 75%), and in KORA S4 in total 4,261 subjects have been examined (response 67%). 3,006 subjects participated in a 10-year follow-up examination of S3 in 2004/05 (KORA F3). For KORA F3 we selected 1,644 subjects of these participants while for KORA S4 we randomly selected 1,814 subjects for genotyping. Informed consent has been given by all participants. The study has been approved by the local ethics committee. The present analysis includes 391 postmenopausal women from KORA F3 and 201 from KORA S4 for whom age at menopause was known and whose samples passed the exclusion criteria described in the methods.

Nature Genetics: doi:10.1038/ng.1051

The **KORCULA** study, is a family-based, cross-sectional study in the isolated Croatian island of Korcula that included 909 genotyped examinees aged 18-95. The cohort includes 573 women of which 333 underwent physiological menopause between 40 and 60 years of age, with a mean age of 49.6.

The **LBC1936** consists of 1,091 relatively healthy individuals assessed on cognitive and medical traits at 70 years of age. They were born in 1936, most took part in the Scottish Mental Survey of 1947, and almost all lived independently in the Lothian region of Scotland. The sample of 548 men and 543 women had a mean age 69.6 years (SD = 0.8). A full description of participant recruitment and testing can be found elsewhere[46]. Ethics permission for the study was obtained from the Multi-Centre Research Ethics Committee for Scotland (MREC/01/0/56) and from Lothian Research Ethics Committee (LREC/2003/2/29). The research was carried out in compliance with the Helsinki Declaration. All subjects gave written, informed consent. Females were included in the study if they had genome-wide genetic data and age at natural menopause data available (n = 484). Individuals were excluded on the following; age at natural menopause < 40 years (n = 32) and > 60 years (n = 5), hysterectomy (n = 96), ovariectomy (n = 9), hormone replacement therapy use before menopause (n = 5). The final sample comprised 337 females with a mean age of natural menopause = 50 years (SD = 3.99, Maximum = 59 years, Minimum = 40 years). The age of natural menopause was validated as the participants were recruited for a second wave of the study aged ~72 years. 229 individuals were asked the same question and the age of natural menopause correlated highly ($r^2 = 0.77$, $P < 0.01$). The genome-wide association methods have previously been published[47].

The **LifeLines** Cohort Study is a multi-disciplinary prospective population-based cohort study examining in a unique three-generation design the health and health-related behaviours of 165,000 persons living in the North East region of The Netherlands[48]. It employs a broad range of investigative procedures in assessing the biomedical, socio-demographic, behavioural, physical and psychological factors which contribute to the health and disease of the general population, with a special focus on multimorbidity. In addition, the LifeLines project comprises a number of cross-sectional sub-studies, which investigate specific age-related conditions. These include investigations into metabolic and hormonal diseases, including obesity, cardiovascular and renal diseases, pulmonary diseases and allergy, cognitive function and depression, and musculoskeletal conditions. All survey participants are between 18 and 90 years old at the time of enrollment. Recruitment has been going on since the end of 2006, and in August 2010 more than 30,000 participants had been included.

The Orkney Complex Disease Study (**ORCADES**) is an ongoing family-based, cross-sectional study in the isolated Scottish archipelago of Orkney. Genetic diversity in this population is decreased compared to Mainland Scotland, consistent with the high levels of endogamy historically. Data for participants aged 18-100 years, from a subgroup of ten islands, were used for this analysis. Fasting blood samples were collected and

Nature Genetics: doi:10.1038/ng.1051

over 200 health-related phenotypes and environmental exposures were measured in each individual. All participants gave informed consent and the study was approved by Research Ethics Committees in Orkney and Aberdeen.

For the **Osteoporosis: SNPs to Environment Study (OSTEOS)** six hundred unrelated women were recruited for this cross-sectional study within December 2006 and January 2008. Participants were consecutive, unselected postmenopausal Caucasian women of Greek origin who were asked to voluntarily participate in the study through advertisement in 4 randomly selected Centers of Open Protection for the Elderly in Athens region (community centers which aimed to primary health prevention and social support of elderly persons)[49]. Natural menopause was defined as amenorrhea for at least 12 consecutive months. Age at natural menopause was self-reported[50]. Exclusion criteria included surgical or medicational induced menopause and use of HRT before menopause (self reported). After all exclusions, 102 women were excluded and the final sample consisted of 498 women. The study protocol was approved by the Bioethics Committee of Harokopio University of Athens and all subjects signed a volunteer consent form.

The **Prospect-Epic** cohort is one of the two Dutch contributions to the European Prospective Investigation into Cancer and Nutrition (EPIC). The design and rationale of this study has been described previously[51]. In brief, this cohort consists of 17,357 white women living in Utrecht and surroundings, aged 49 – 70 years, who were invited to participate in the study through the national breast cancer screening program between 1993 and 1997. All women filled out detailed questionnaires about dietary, reproductive, and medical history and underwent a physical examination at enrollment. In addition, women donated a 30-mL non-fasting blood sample which was fractioned into serum, citrated plasma, buffy coat and erythrocyte aliquots of 0.5 mL each. The samples were stored under liquid nitrogen at -196º C for future research. Natural menopause was defined according to the World Health Organization as amenorrhea for at least 12 consecutive months without other obvious reasons. A total of 3497 women were premenopausal or perimenopausal at time of enrollment and therefore excluded. All women who experienced a surgical menopause (N= 4,449), used hormones during the menopausal transition (N=2,161) or women with an unknown menopausal status or age (N=1,194) were excluded. Next, all women who were younger than 58 years at inclusion in the Prospect-Epic cohort were excluded to avoid bias due to differential inclusion of women with an early menopause N=2,248). 192 women were excluded because of missing buffy coat samples or failed DNA extraction. Finally, 62 women were excluded because they experienced menopause before 40 years or after 60 years of age, leaving a total of 3,524 women available for analysis.

SASBAC: The study base included all Swedish-born women between 50 and 74 years of age who were resident in Sweden between October 1993 and March 1995. During that period, virtually all breast cancer cases in Sweden were identified, and randomly selected controls, who matched the cases in 5-year age strata, were

Nature Genetics: doi:10.1038/ng.1051

selected from the Swedish registry of the total population. Of the eligible women, 3,345 (84%) breast cancer cases and 3,454 (82%) controls participated in this initial questionnaire-based study, providing detailed information on their use of menopausal hormone therapy, their reproductive history and other lifestyle factors. From these women, a random subsample has undergone a genome-wide association scan (803 cases, 764 controls). For the present study, only controls were considered eligible[52,53].

The **SPLIT** study, Croatia, is an ongoing population-based, cross-sectional study in the Dalmatian City of Split that included 499 genotyped examinees aged 18-95. The cohort includes 286 women of which 141 underwent physiological menopause between 40 and 60 years of age, with a mean age of 49.99.

TWINGENE, between the years 2004 and 2008 a population wide collection of blood on 12 600 twins born 1958 or earlier was undertaken in Sweden[54]. About 200 twins were contacted each month until the data collection was completed in 2008. When the signed consent forms where returned, the subjects were sent blood-sampling equipment and asked to contact a local health facility for blood sampling. The study population was recruited among twins participating in the Screening Across the Lifespan Twin Study (SALT) which was a population based telephone interview study conducted in 1998-2002. Other inclusion criteria were that both twins in the pair had to be alive and living in Sweden. Subjects were excluded from the study if they preciously declined participation in future studies or if they had been enrolled in other DNA sampling projects of the Swedish Twin Register. Menopausal information was collected from the SALT interview. Among the 302 monozygous pairs that had been genotyped with genome-wide array (Illumina 317K), menopausal information was available for at least on of the twins in the pair for 280 pairs. For pairs in which information about menopause were available for both the average within-pair value was used.

The **VIS** study is a family-based, cross-sectional study in the isolated Croatian island of Vis that included 924 genotyped examinees aged 18-93. The cohort included 536 women of which 313 underwent physiological menopause between 40 and 60 years of age, with a mean age of 48.78

Genotyping and Imputation (Discovery and in silico replication cohorts)

Eight different genotyping platform were used by the discovery and in silico replication cohort studies: Illumina Human CNV 370 (AGES, CHS, deCODE, EGCUT (discovery and replication), ERF, INGI-CARl, INGI-FVG, INGI-Val Borbera, KORCULA, QIMR, and SPLIT), the Illumina HumanHap 300K (deCODE, ERF, ORCADES, TwinsUKI, VIS, and WGHS (HumanHap 300K Duo Plus)), the Illumina Infinium II Human Hap 550 SNP chip array (InCHIANTI, NHS-cgems, RSI, RSII, RSIII, and SASBAC), the Illumina Hap610Quad (HBCS, LBC1936, and TwinsUKII and III), the Illumina 317K (TWINGENE), the Illumina CytoSNP v2 300K (LifeLines), the Affymetrix Genome-Wide Human SNP Array 6.0 (ARIC, GENOA, KORA S4, NHS-Hu, and SHIP), the Affymetrix 500K mapping array (CoLaus, HAPI, KORA F3, and NTR), the

Nature Genetics: doi:10.1038/ng.1051

Affymetrix 500K in combination with the 50K supplemental array (FHS), and the Affymetrix 500K in combination with the 10K supplemental array (SardiNIA). Each study performed genotyping quality control checks based on duplicate sample genotyping, SNP call rate, Hardy-Weinberg equilibrium, Mendelian inconsistencies, and sex mismatch, and principle components methods were used to evaluate the presence of population stratification (details provided in Table 2). Because there were only a small number overlapping SNPs from the eight genotyping platforms, each study imputed 2.5 million HapMap SNPs for each participant using currently available imputation methods. deCODE, EGCUT, InCHIANTI, NTR, SHIP and TwinsUK from the discovery stage and CoLaus, EGCUT, LifeLines, SASBAC and TWINGENE from the replication cohorts used IMPUTE (http://www.stats.ox.ac.uk/~marchini/software/gwas /impute), CHS used BimBam[55] and all other cohorts used a MACH algorithm (http://www.sph.umich.edu/csg/ abecasis/MaCH/) (Supplementary Table SM4). All studies imputed the genotype "dosage" for the expected number of minor alleles. Imputation quality was determined by either the r^2 value produced by MACH or calculated empirical variance divided by the expected variance (oevar) and for SNPTEST the 'proper info' output variable was used to determine imputation quality. SNP imputation methods and quality control procedures for each cohort are included in Table M2 for the discovery cohorts and M4 for the in silico replication cohorts.

Genotyping replication cohorts

The samples of the genotyping replication were genotyped using Sequenom iPLEX genotyping (EPOS), Taqman Allelic Discrimination (EPOS), or KaspAR (BWHHS, Prospect-Epic).

Sequenom iPLEX genotyping - Multiplex PCR assays were designed for the Sequenom iPLEX genotying using Assay Designer on the website (https://mysequenom.com/tools/genotyping-/default.aspx). For this, sequences containing the SNP site and at least 100 bp of flanking sequence on either side of the SNP were used. Briefly, 2 ng genomic DNA was amplified in a 5 ul reaction containing 1 × Taq PCR buffer (Sequenom), 2 mM $MgCl_2$, 500 uM each dNTP, 100 nM each PCR primer, 0.5 U Taq (Sequenom). The reaction was incubated at 94°C for 4 minutes followed by 45 cycles of 94°C for 20 seconds, 56°C for 30 seconds, 72°C for 1 minute, followed by 3 minutes at 72°C. Excess dNTPs were then removed from the reaction by incubation with 0.3 U shrimp alkaline phosphatase (Sequenom) at 37°C for 40 minutes followed by 5 minutes at 85°C to deactivate the enzyme. Single primer extension over the SNP was carried out in a final concentration of between 0.731 uM and 1.462 uM for each extension primer (depending on the mass of the probe), iPLEX termination mix (Sequenom), 10x iPLEX Buffer Plus and iPLEX enzyme (Sequenom) and cycled using the following program; 94°C for 30 seconds followed by 94°C for 5 seconds, 5 cycles of 52°C for 5 seconds, and 80°C for 5 seconds, the last three steps were repeated 40 times, then 72°C for 3 minutes. The reaction was then desalted by addition of 6 mg clear resin (Sequenom) followed by mixing (15 minutes) and centrifugation (5

Nature Genetics: doi:10.1038/ng.1051

min, 3,000rpm) to settle the contents of the tube. The extension product was then spotted onto a 384 well spectroCHIP using the SEQUENOM MassARRAY Nanodispenser RS1000 before analysis on the MassARRAY Compact System (Sequenom). Data collection was performed using SpectroACQUIRE 3.3.1.13 and clustering was called using TYPER Analyzer 4.0.3.18 (Sequenom). Additionally, to ensure data qualities, genotypes for each subject were also checked manually.

Taqman Allelic Discrimination - Assays for Taqman Allelic Discrimination were available at www.appliedbiosystems.com as pre-designed assays. The PCR reaction mixture included 1-2 ng of genomic DNA in a 2 ul volume and the following reagents: FAM and VIC probes (200 nM), primers (0.9 uM), 2x Taqman PCR master mix (Applied Biosystems Inc., Foster City, CA, USA). Reagents were dispensed in a 384-well plate using the Deerac Equator NS808 (Deerac Fluidics, Dublin, Ireland). PCR cycling reaction were performed in 384 wells PCR plates in an ABI 9700 PCR system (Applied Biosystems Inc., Foster City, CA, USA) and consisted of initial denaturation for 15 minutes at 95° C, and 40 cycles with denaturation of 15 seconds at 95° C and annealing and extension for 60 seconds at 60° C. Results were analysed by the ABI Taqman 7900HT using the sequence detection system 2.22 software (Applied Biosystems Inc., Foster City, CA, USA).

KASPar - Genotyping was performed by K Biosciences (Herts, U.K.), who designed and used assays based on their proprietary competitive allele-specific PCR (KASPar) method (details of which are available on their website www.kbioscience.co.uk/).

Statistical analysis

Cohort specific assocation analyses

AGES: Analysis was preformed using linear regression against the imputed genotype dosage with the ProbABEL package.

ARIC: Population stratification was estimated using principal component methods (EIGENSTRAT)[56], after removing few related individuals. We used linear regression models and assumed additive genetic effects to study the association of imputed and genotyped SNPs (dosage data) and age of menopause. The analyses were implemented in the ProbABEL package from the ABEL set of programs (http://mga.bionet.nsc.ru/yurii/ABEL/)[57].

FHS: SNP weights for 10 principal components (PCs) were inferred using a maximal set of independent individuals; the PCs for the remaining individuals were computed using the SNP weights obtained from the unrelated set of individuals. The sixth PC was significantly associated with age at natural menopause ($P<0.01$), and therefore was included as a covariate in all SNP association analyses. Linear mixed effects models were

Nature Genetics: doi:10.1038/ng.1051

used to account for familial correlations. Each SNP was tested for association with age at menopause using an additive genetic model.

GENOA: Linear mixed effects models were used to account for family structure in all SNP association tests. An additive genetic model was assumed using either directly genotyped SNPs (when available) or imputed SNP dosages

HAPI Heart Study: Analysis was performed using in house developed software (J O'Connell). In brief, we performed a measured genotype approach utilizing a t-test of the beta coefficient for the SNP variable. The polygenic component was modeled using the relationship matrix derived from the complete 14-generation pedigree structure, to properly control for the relatedness of all subjects in the study.

InCHIANTI: Analysis performed using linear regression allele dosage in SNPTEST (http://www.stats.ox.ac.uk/~marchini/software/gwas/snptest).

NHS: For both NHS breast cancer and T2D GWAS analyses, we performed linear regression to analyze the association between each of the SNPs (the imputed genotype dosage) and age at natural menopause using ProABEL software [57]. SNPs with low MAF (< 1%) in samples were excluded from analysis. To control for potential confounding by population stratification, we adjusted for the top principal components of genetic variation chosen for each study after excluding any admixed individuals clearly not of European ancestry[56,58-60]. Controlling for breast cancer or T2D case-control status in corresponding study made no material difference to the results.

Netherlands Twin Register (NTR)

Analysis was performed in unrelated Caucasian individuals using linear regression allele dosage in SNPTEST. To account for residual population stratification, the 10 principal components calculated with the EIGENSTRAT[56] software were added as covariates in the regression model. After analyses SNPs were excluded if they failed to meet one of these criteria: MAF > 0.01, Call rate > 0.95, HWE p-value > 0.00001.

QIMR: Due to the correlated nature of family data, we used the program Merlin-offline to analyse the GWAS data (http://www.sph.umich.edu/csg/abecasis/-merlin/index.html). The analysis protocol uses an allelic score-test that facilitates analysis of dosage data in families.

RSI, RSII and RSIII: Adjusted linear regression analysis was done using MACH2QTL (http://www.sph.umich.edu/csg/abecasis/MaCH/) implemented in GRIMP[61].

TwinsUK: Because of the relatedness in the TwinsUK cohort, we utilized the GenABEL software package[57] which is designed for GWAS analysis of family-based data by incorporating pair-wise kinship matrix calculated using genotyping data in the polygenic model to correct relatedness and hidden population stratification. The score test implemented in the software was used to test the association between a given SNP and the age at menopause.

Nature Genetics: doi:10.1038/ng.1051

WGHS: Association testing for reproductive aging phenotypes was performed with Mach2Qtl v. 1.0.4.

Extended Acknowledgements

The Age Gene/Environment Susceptibility Reykjavik (AGES Reykjavik) study has been funded by NIH contract N01-AG-1-2100, the NIA Intramural Research Program, Hjartavernd (the Icelandic Heart Association), and the Althingi (the Icelandic Parliament). The researchers are indebted to the participants for their willingness to participate in the study.

The **Atherosclerosis Risk in Communities Study** is carried out as a collaborative study supported by National Heart, Lung, and Blood Institute contracts N01-HC-55015, N01-HC-55016, N01-HC-55018, N01-HC-55019, N01-HC-55020, N01-HC-55021, N01-HC-55022, R01HL087641, R01HL59367 and R01HL086694; National Human Genome Research Institute contract U01HG004402; and National Institutes of Health contract HHSN268200625226C. The authors thank the staff and participants of the ARIC study for their important contributions. Infrastructure was partly supported by Grant Number UL1RR025005, a component of the National Institutes of Health and NIH Roadmap for Medical Research.

Cardiovascular Health Study: This CHS research was supported by NHLBI contracts N01-HC-85239, N01-HC-85079 through N01-HC-85086; N01-HC-35129, N01 HC-15103, N01 HC-55222, N01-HC-75150, N01-HC-45133 and NHLBI grants HL080295, HL075366, HL087652, HL105756 with additional contribution from NINDS. Additional support was provided through AG-023629, AG-15928, AG-20098, and AG-027058 from the NIA. See also http://www.chs-nhlbi.org/pi.htm. DNA handling and genotyping was supported in part by National Center for Research Resources grant UL1RR033176 to the UCLA/Cedars-Sinai/Harbor-UCLA/Drew Clinical Translation Science Institute and National Institute of Diabetes and Digestive and Kidney Diseases grant DK063491 to the Southern California Diabetes Endocrinology Research Center.

EGCUT: EGCUT authors received financing by FP7 grants (201413, 245536), also received targeted financing from Estonian Government SF0180142s08, SF0180044s09 (TL, ASa only), from University of Tartu in the frame of Center of Translational Genomics and from the European Union through the European Regional Development Fund, in the frame of Centre of Excellence in Genomics. TL and ASa were also supported by Enterprise Estonia (EU30200). Authors acknowledge EGCUT and Estonian Biocentre personnel, especially Ms. M. Hass and Mr V. Sood. EGCUT data analyzes were carried out in part in the High Performance Computing Center of University of Tartu.

Erasmus Rucphen Family (ERF): The study was supported by grants from The Netherlands Organisation for Scientific Research (NWO), Erasmus MC and the Centre for Medical Systems Biology (CMSB). We are grateful to all general practitioners for their contributions, to Petra Veraart for her help in

Nature Genetics: doi:10.1038/ng.1051

genealogy, Jeannette Vergeer for the supervision of the laboratory work and Peter Snijders for his help in data collection.

The **Framingham Heart Study** phenotype-genotype analyses and the meta-analysis for this work were supported by the National Institute of Aging (Genetics of Reproductive Life Period and Health Outcomes, R21AG032598; JMM, KLL, DK, DPK). The National Institute of Arthritis, Musculoskeletal and Skin Diseases and National Institute on Aging, R01 AR/AG 41398 provided additional support (DPK, DEK). The Framingham Heart Study of the National Heart Lung and Blood Institute of the National Institutes of Health and Boston University School of Medicine was supported by the National Heart, Lung and Blood Institute's Framingham Heart Study Contract No. N01-HC-25195 and its contract with Affymetrix, Inc for genotyping services (Contract No. N02-HL-6-4278). Analyses reflect intellectual input and resource development from the Framingham Heart Study investigators participating in the SNP Health Association Resource (SHARe) project. A portion of this research was conducted using the Linux Cluster for Genetic Analysis (LinGA-II) funded by the Robert Dawson Evans Endowment of the Department of Medicine at Boston University School of Medicine and Boston Medical Center.

The **HAPI Heart Study** was supported by NIH research grants U01 HL72515, U01 HL84756 and R01 088119, the University of Maryland General Clinical Research Center, grant M01 RR 16500, the Clinical Nutrition Research Unit of Maryland (P30 DK072488), and the Baltimore Veterans Administration Medical Center Geriatrics Research. LMY-A was supported by NIH training grant F32AR059469.

The **InCHIANTI** study baseline (1998-2000) was supported as a “targeted project” (ICS110.1/RF97.71) by the Italian Ministry of Health and in part by the U.S. National Institute on Aging (Contracts: 263 MD 9164 and 263 MD 821336); the InCHIANTI Follow-up 1 (2001-2003) was funded by the U.S. National Institute on Aging (Contracts: N.1-AG-1-1 and N.1-AG-1-2111); the InCHIANTI Follow-ups 2 and 3 studies (2004-2010) were financed by the U.S. National Institute on Aging (Contract: N01-AG-5-0002);supported in part by the Intramural research program of the National Institute on Aging, National Institutes of Health, Baltimore, Maryland.

The **NHS** breast cancer GWA scan was performed as part of the Cancer Genetic Markers of Susceptibility initiative of the NCI. We particularly acknowledge the contributions of R. Hoover, A. Hutchinson, K. Jacobs and G. Thomas. The current research is supported by CA 40356, CA 87969, and U01-CA98233 from the National Cancer Institute. The NHS/HPFS type 2 diabetes GWAS (U01HG004399) is a component of a collaborative project that includes 13 other GWAS funded as part of the Gene Environment-Association Studies (GENEVA) under the NIH Genes, Environment and Health Initiative (GEI) (U01HG004738, U01HG004422, U01HG004402, U01HG004729, U01HG004726, U01HG004735, U01HG004415, U01HG004436, U01HG004423, U01HG004728, RFAHG006033) with additional support

Nature Genetics: doi:10.1038/ng.1051

from individual NIH Institutes (NIDCR: U01DE018993, U01DE018903; NIAAA: U10AA008401, NIDA: P01CA089392, R01DA013423; NCI: CA63464, CA54281, CA136792, Z01CP010200). Assistance with phenotype harmonization and genotype cleaning, as well as with general study coordination, was provided by the GENEVA Coordinating Center (U01HG004446). Assistance with data cleaning was provided by the National Center for Biotechnology Information. Genotyping was performed at the Broad Institute of MIT and Harvard, with funding support from the NIH GEI (U01HG04424), and Johns Hopkins University Center for Inherited Disease Research, with support from the NIH GEI (U01HG004438) and the NIH contract "High throughput genotyping for studying the genetic contributions to human disease"(HHSN268200782096C). Additional funding for the current research was provided by the National Cancer Institute (NCI, P01CA087969, P01CA055075), and the National Institute of Diabetes and Digestive and Kidney Diseases (NIDDK, R01DK058845). We acknowledge the study participants in the NHS for their contribution in making this study possible.

NTR - Funding was obtained from the Netherlands Organization for Scientific Research (NWO: MagW/ZonMW): Genetic basis of anxiety and depression (904-61-090); Genetics of individual differences in smoking initiation and persistence (NWO 985-10-002); Resolving cause and effect in the association between exercise and well-being (904-61-193); Twin family database for behavior genomics studies (480-04-004); Twin research focusing on behavior (400-05-717); Genetic determinants of risk behavior in relation to alcohol use and alcohol use disorder (Addiction-31160008); Genotype/phenotype database for behavior genetic and genetic epidemiological studies (911-09-032); Spinozapremie (SPI 56-464-14192); CMSB: Center for Medical Systems Biology (NWO Genomics); NBIC/BioAssist/RK/2008.024); BBMRI –NL: Biobanking and Biomolecular Resources Research Infrastructure; the VU University: Institute for Health and Care Research (EMGO+) and Neuroscience Campus Amsterdam (NCA); the 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 (ERC) Genetics of Mental Illness (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.

QIMR: We thank the twins and their families for their participation. We also thank Dixie Statham, Ann Eldridge, Marlene Grace, Kerrie McAloney,Lisa Bowdler, Steven Crooks, Peter Visscher and Allan McRae. Funding was provided by the Australian National Health and Medical Research Council (241944, 339462, 389927, 389875, 389891, 389892, 389938, 442915, 442981, 496739, 552485, 552498), the Australian Research Council (A7960034, A79906588, A79801419, DP0770096, DP0212016, DP0343921), the FP-5 GenomEUtwin Project (QLG2-CT-2002-01254), and the U.S. National Institutes of Health (NIH grants AA07535, AA10248, AA13320, AA13321, AA13326, AA14041, MH66206). A portion of the genotyping on which this study was

based (Illumina 370K scans on 4300 individuals) was carried out at the Center for Inherited Disease Research, Baltimore (CIDR), through an access award to our late colleague Dr. Richard Todd (Psychiatry, Washington University School of Medicine, St Louis). Statistical analyses were carried out on the Genetic Cluster Computer, which is financially supported by the Netherlands Scientific Organization (NWO 480-05-003). G.W.M, D.R.N and S.E.M. are supported by the National Health and Medical Research Council (NHMRC) Fellowship Scheme.

The generation and management of GWAS genotype data for the **Rotterdam Study** are supported by the Netherlands Organisation 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)/Netherlands Organisation for Scientific Research (NWO) project nr. 050-060-810, and funding from the European Commision (HEALTH-F2-2008-201865, GEFOS; HEALTH-F2-2008-35627, TREAT-OA). The Rotterdam Study is funded by Erasmus Medical Center and Erasmus University, Rotterdam, Netherlands Organization for the Health Research and Development (ZonMw), the Research Institute for Diseases in the Elderly (RIDE), the Ministry of Education, Culture and Science, the Ministry for Health, Welfare and Sports, the European Commission (DG XII), and the Municipality of Rotterdam. We thank Pascal Arp, Mila Jhamai, Dr Michael Moorhouse, Marijn Verkerk, and Sander Bervoets for their help in creating the GWAS database. The authors are grateful to the study participants, the staff from the Rotterdam Study and the participating general practioners and pharmacists. We would like to thank Dr. Tobias A. Knoch, Karol Estrada, Luc V. de Zeeuw, Anis Abuseiris, and Rob de Graaf as well as their institutions the Erasmus Computing Grid, Rotterdam, The Netherlands, and especially the national German MediGRID and Services@MediGRID part of the German D-Grid, both funded by the German Bundesministerium fuer Forschung und Technology under grants #01 AK 803 A-H and # 01 IG 07015 G for access to their grid resources.

The **SardiNIA** study is supported by the National Institute on Aging, NIH, contract NO1-AG-1-2109 to the SardiNIA ("ProgeNIA") team. The authors are grateful to all of the volunteers who participated in this study, Monsignore Piseddu, Bishop of Ogliastra, the mayors and citizens of the Sardinian towns (Lanusei, Ilbono, Arzana, and Elini), the head of the Public Health Unit ASL4 for their volunteerism and cooperation, and team of physicians, nurses, biologists and the recruitment personnel.

SHIP is part of the Community Medicine Research net of the University of Greifswald, Germany, which is funded by the Federal Ministry of Education and Research (grants no. 01ZZ9603, 01ZZ0103, and 01ZZ0403), the Ministry of Cultural Affairs as well as the Social Ministry of the Federal State of Mecklenburg-West Pomerania. Genome-wide data have been supported by the Federal Ministry of Education and Research (grant no. 03ZIK012) and a joint grant from Siemens Healthcare, Erlangen, Germany and the Federal State of Mecklenburg- West Pomerania. The University of Greifswald is a member of the 'Center of Knowledge

Nature Genetics: doi:10.1038/ng.1051

Interchange' program of the Siemens AG. This work is also part of the research project Greifswald Approach to Individualized Medicine (GANI_MED). The GANI_MED consortium is funded by the Federal Ministry of Education and Research and the Ministry of Cultural Affairs of the Federal State of Mecklenburg – West Pomerania (03IS2061A).

TwinsUK is supported by the Wellcome Trust from the Department 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 and King's College Hospital NHS Foundation Trust. We thank the staff from the TwinsUK, the DNA Collections and Genotyping Facilities at the Wellcome Trust Sanger Institute for sample preparation; Quality Control of the Twins UK cohort for genotyping (in particular Amy Chaney, Radhi Ravindrarajah, Douglas Simpkin, Cliff Hinds, and Thomas Dibling); Paul Martin and Simon Potter of the DNA and Genotyping Informatics teams for data handling; Le Centre National de Génotypage, France, led by Mark Lathrop, for genotyping; Duke University, North Carolina, USA, led by David Goldstein, for genotyping; and the Finnish Institute of Molecular Medicine, Finnish Genome Center, University of Helsinki, led by Aarno Palotie.

The **WGHS** is funded by the Donald W. Reynolds Foundation (Las Vegas, NV), the Fondation LeDucq (Paris, France), the National Heart, Lung and Blood Institute (NHLBI; HL043851) and the National Cancer Institute (NCI; CA047988). Funding for genotyping and collaborative scientific support was provided by Amgen.

BWHHS - We thank all of the staff and participants in all of the studies. BWHHS is funded by the Department of Health (England) Policy Research Programme and the British Heart Foundation. DA Lawlor works in a Centre that receives infrastructure funding from the UK Medical Research Council and University of Bristol. The views expressed in the publication are those of the authors and not necessarily those of the funders.

CoLaus - The CoLaus study received financial contributions from GlaxoSmithKline, the Faculty of Biology and Medicine of Lausanne, and the Swiss National Science Foundation (33CSCO-122661). The authors thank Peter Vollenweider and Dawn Waterworth, Co-PIs of the CoLaus study. Special thanks to Murielle Bochud, Yolande Barreau, Mathieu Firmann, Vladimir Mayor, Anne-Lise Bastian, Binasa Ramic, Martine Moranville, Martine Baumer, Marcy Sagette, Jeanne Ecoffey and

EGCUT: EGCUT authors received financing by FP7 grants (201413, 245536), also received targeted financing from Estonian Government SF0180142s08, SF0180044s09 (TL, ASa only), from University of Tartu in the frame of Center of Translational Genomics and from the European Union through the European Regional Development Fund, in the frame of Centre of Excellence in Genomics. TL and ASa were also supported by Enterprise Estonia (EU30200). Authors acknowledge EGCUT and Estonian Biocentre personnel, especially Ms.

M. Hass and Mr V. Sood. EGCUT data analyzes were carried out in part in the High Performance Computing Center of University of Tartu.

EPOS was made possible by a grant of the Praeventiefonds (currently known as Zon-Mw): Grant no. 28-2401. The authors want to thank Luc Martens and Geraline Leusink for their participation (at that time) as a PhD student into EPOS. Finally, they want to thank all the nurses who participated in interviewing over 6000 women.

GENOA - The Genetic Epidemiology Network of Arteriopathy phenotyping and genome-wide genotyping is supported by the National Heart Lung and Blood Institute of the National Institutes of Health (HL54457, HL68737, and HL087660). We would like to thank Stephen Turner, Mariza de Andrade, and Julie Cunningham at the Mayo Clinic (Field Center and Genotyping Center), Eric Boerwinkle and Megan L. Grove-Gaona at the University of Texas Health Sciences Center (DNA Laboratory), Patricia Peyser and Lawrence Bielak at the University of Michigan (Analysis Center), and the families that participated in the GENOA study.

HBCS - We thank all study participants as well as everybody involved in the Helsinki Birth Cohort Study. Helsinki Birth Cohort Study has been supported by grants from the Academy of Finland, the Finnish Diabetes Research Society, Folkhälsan Research Foundation, Novo Nordisk Foundation, Finska Läkaresällskapet, Signe and Ane Gyllenberg Foundation, University of Helsinki, European Science Foundation (EUROSTRESS), Ministry of Education, Ahokas Foundation, Emil Aaltonen Foundation, Juho Vainio Foundation, and Wellcome Trust (grant number WT089062).

INGI-CARL/FVG - The INGI-Carlantino and INGI-FVG studies was supported by grants from Telethon, FVG region and Fondo Trieste.

INGI-Val Borbera

The Val Borbera study was supported by grants from Foundation Compagnia San Paolo Torino, Foundation Cassa di Risparmio di Alessandria and by Ministry of Health (Ricerca Corrente 07and 2008 and Ricerca Finalizzata 2007)

The **KORA** research platform was initiated and financed by the Helmholtz Center Munich, German Research Center for Environmental Health, which is funded by the German Federal Ministry of Education and Research (BMBF) and by the State of Bavaria. Part of this work was financed by the German National Genome Research Network (NGFN-2 and NGFNPlus: 01GS0823). Our research was supported within the Munich Center of Health Sciences (MC Health) as part of LMUinnovativ.

The **KORCULA** study was funded by grants from the Medical Research Council (UK), European Commission Framework 6 project EUROSPAN (Contract No. LSHG-CT-2006-018947) and Republic of Croatia Ministry of Science, Education and Sports research grants to I.R. (108-1080315-0302). We would like to acknowledge the staff of several institutions in Croatia that supported the field work, including but not

limited to The University of Split and Zagreb Medical Schools and Croatian Institute for Public Health. The SNP genotyping for the KORCULA cohort was performed in Helmholtz Zentrum München, Neuherberg, Germany

LBC1936 - We thank the LBC1936 participants. We thank Alan Gow, Janie Corley, Caroline Brett, Caroline Cameron, Michelle Taylor, and Alison Pattie for data collection and data entry. We thank the study secretary Paula Davies. We thank the nurses and staff at the Wellcome Trust Clinical Research Facility, where subjects were tested and the genotyping was performed. We thank the staff at the Lothian Health Board, and the staff at the SCRE Centre, University of Glasgow. This whole genome association study was funded by the Biotechnology and Biological Sciences Research Council (BBSRC). The LBC1936 research was supported by a programme grant from Research Into Ageing and continues with programme grants from Help the Aged/Research Into Ageing (Disconnected Mind). The work was undertaken by The University of Edinburgh Centre for Cognitive Ageing and Cognitive Epidemiology, part of the cross council Lifelong Health and Wellbeing Initiative (G0700704/84698). Funding from the Biotechnology and Biological Sciences Research Council (BBSRC), Engineering and Physical Sciences Research Council (EPSRC), Economic and Social Research Council (ESRC) and Medical Research Council (MRC) is gratefully acknowledged. Lorna M Lopez is the beneficiary of a post-doctoral grant from the AXA Research Fund.

The **LifeLines Cohort Study**, and generation and management of GWAS genotype data for the LifeLines Cohort Study is supported by the Netherlands Organization of Scientific Research NWO (grant 175.010.2007.006), the Economic Structure Enhancing Fund (FES) of the Dutch government, the Ministry of Economic Affairs, the Ministry of Education, Culture and Science, the Ministry for Health, Welfare and Sports, the Northern Netherlands Collaboration of Provinces (SNN), the Province of Groningen, University Medical Center Groningen, the University of Groningen, Dutch Kidney Foundation and Dutch Diabetes Research Foundation.

We thank Behrooz Z. Alizadeh, Annemieke Boesjes, Marcel Bruinenberg, Noortje Festen, Ilja Nolte, Lude Franke, Mitra Valimohammadi for their help in creating the GWAS database, and Rob Bieringa, Joost Keers, René Oostergo, Rosalie Visser, Judith Vonk for their work related to data-collection and validation. The authors are grateful to the study participants, the staff from the LifeLines Cohort Study and Medical Biobank Northern Netherlands, and the participating general practitioners and pharmacists. LifeLines Scientific Protocol Preparation: Rudolf de Boer, Hans Hillege, Melanie van der Klauw, Gerjan Navis, Hans Ormel, Dirkje Postma, Judith Rosmalen, Joris Slaets, Ronald Stolk, Bruce Wolffenbuttel; LifeLines GWAS Working Group: Behrooz Alizadeh, Marike Boezen, Marcel Bruinenberg, Noortje Festen, Lude Franke, Pim van der Harst, Gerjan Navis, Dirkje Postma, Harold Snieder, Cisca Wijmenga, Bruce Wolffenbuttel.

Nature Genetics: doi:10.1038/ng.1051

ORCADES was supported by the Chief Scientist Office of the Scottish Government (CZB/4/276), the Royal Society and the European Union framework program 6 EUROSPAN project (contract no. LSHG-CT-2006-018947). DNA extractions were performed at the Wellcome Trust Clinical Research Facility in Edinburgh. We would like to acknowledge the invaluable contributions of Lorraine Anderson and the research nurses in Orkney, the administrative team in Edinburgh and the people of Orkney.

OSTEOS - The study was supported by the Greek State Scholarships Foundation and the Hellenic Society for Research of Metabolic Bone Diseases. We thank all participants.

The **Prospect-EPIC** study was funded by 'Europe against Cancer' Programme of the European Commission (SANCO); Dutch Ministry of Health, Welfare and Sports (VWS); and ZONMw. We thank Statistics Netherlands (CBS) and Netherlands Cancer Registry (NKR) for data on cancer and death.

The **SASBAC** study was supported by funding from the Agency for Science, Technology and Research of Singapore (A*STAR), the United States National Institute of Health (NIH) and the Susan G. Komen Breast Cancer Foundation. E.I. was supported by grants from the Swedish Research Council, the Swedish Heart-Lung Foundation, the Swedish Society of Medicine, the Swedish Foundation for Strategic Research, and the Royal Swedish Academy of Science while working with this article.

The **SPLIT** study is funded by grants from the Medical Research Council (UK) and Republic of Croatia Ministry of Science, Education and Sports research grants to I.R. (108-1080315-0302). We would like to acknowledge the staff of several institutions in Croatia that supported the field work, including but not limited to The University of Split and Zagreb Medical Schools and Croatian Institute for Public Health. The SNP genotyping for the SPLIT cohort was performed by AROS Applied Biotechnology, Aarhus, Denmark.

The **TWINGENE** study was supported by the Swedish Research Council, GenomEUtwin, The United States National Institute of Health (NIH), Swedish Foundation for Strategic Research (SSF)

The **VIS** study was funded by grants from the Medical Research Council (UK), European Commission Framework 6 project EUROSPAN (Contract No. LSHG-CT-2006-018947) and Republic of Croatia Ministry of Science, Education and Sports research grants to I.R. (108-1080315-0302). We would like to acknowledge the staff of several institutions in Croatia that supported the field work, including but not limited to The University of Split and Zagreb Medical Schools, Institute for Anthropological Research in Zagreb and Croatian Institute for Public Health.

The SNP genotyping for the VIS cohort was performed in the core genotyping laboratory of the Wellcome Trust Clinical Research Facility at the Western General Hospital, Edinburgh, Scotland

Nature Genetics: doi:10.1038/ng.1051

References

1. Harris, T.B. et al. Age, Gene/Environment Susceptibility-Reykjavik Study: multidisciplinary applied phenomics. *Am J Epidemiol* **165**, 1076-87 (2007).
2. The Atherosclerosis Risk in Communities (ARIC) Study: design and objectives. The ARIC investigators. *Am J Epidemiol* **129**, 687-702 (1989).
3. Fried, L.P. et al. The Cardiovascular Health Study: design and rationale. *Ann Epidemiol* **1**, 263-76 (1991).
4. Tell, G.S. et al. Recruitment of adults 65 years and older as participants in the Cardiovascular Health Study. *Ann Epidemiol* **3**, 358-66 (1993).
5. Metspalu, A. The Estonian Genome Project. *Drug Development Research* **62**, 97-101 (2004).
6. Nelis, M. et al. Genetic structure of Europeans: a view from the North-East. *PLoS One* **4**, e5472 (2009).
7. Aulchenko, Y.S. et al. Linkage disequilibrium in young genetically isolated Dutch population. *Eur J Hum Genet* **12**, 527-34 (2004).
8. Dawber, T.R., Kannel, W.B. & Lyell, L.P. An approach to longitudinal studies in a community: the Framingham Study. *Ann N Y Acad Sci* **107**, 539-56 (1963).
9. Dawber, T.R., Meadors, G.F. & Moore, F.E., Jr. Epidemiological approaches to heart disease: the Framingham Study. *Am J Public Health Nations Health* **41**, 279-81 (1951).
10. Feinleib, M., Kannel, W.B., Garrison, R.J., McNamara, P.M. & Castelli, W.P. The Framingham Offspring Study. Design and preliminary data. *Prev Med* **4**, 518-25 (1975).
11. Kannel, W.B., Feinleib, M., McNamara, P.M., Garrison, R.J. & Castelli, W.P. An investigation of coronary heart disease in families. The Framingham offspring study. *Am J Epidemiol* **110**, 281-90 (1979).
12. Mitchell, B.D. et al. The genetic response to short-term interventions affecting cardiovascular function: rationale and design of the Heredity and Phenotype Intervention (HAPI) Heart Study. *Am Heart J* **155**, 823-8 (2008).
13. Ferrucci, L. et al. Subsystems contributing to the decline in ability to walk: bridging the gap between epidemiology and geriatric practice in the InCHIANTI study. *J Am Geriatr Soc* **48**, 1618-25 (2000).
14. Epstein, M.P., Duren, W.L. & Boehnke, M. Improved inference of relationship for pairs of individuals. *Am J Hum Genet* **67**, 1219-31 (2000).
15. Colditz, G.A. & Hankinson, S.E. The Nurses' Health Study: lifestyle and health among women. *Nat Rev Cancer* **5**, 388-96 (2005).
16. Hunter, D.J. et al. A genome-wide association study identifies alleles in FGFR2 associated with risk of sporadic postmenopausal breast cancer. *Nat Genet* **39**, 870-4 (2007).
17. Cornelis, M.C. et al. Joint effects of common genetic variants on the risk for type 2 diabetes in U.S. men and women of European ancestry. *Ann Intern Med* **150**, 541-50 (2009).
18. Classification and diagnosis of diabetes mellitus and other categories of glucose intolerance. National Diabetes Data Group. *Diabetes* **28**, 1039-57 (1979).
19. Report of the Expert Committee on the Diagnosis and Classification of Diabetes Mellitus. *Diabetes Care* **20**, 1183-97 (1997).
20. Hu, F.B. et al. Physical activity and television watching in relation to risk for type 2 diabetes mellitus in men. *Arch Intern Med* **161**, 1542-8 (2001).
21. Boomsma, D.I. et al. Netherlands Twin Register: from twins to twin families. *Twin Res Hum Genet* **9**, 849-57 (2006).
22. 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**, 335-42 (2008).
23. Medland, S.E. et al. Common variants in the trichohyalin gene are associated with straight hair in Europeans. *Am J Hum Genet* **85**, 750-5 (2009).

Nature Genetics: doi:10.1038/ng.1051

24. Do, K.A. et al. Predictive factors of age at menopause in a large Australian twin study. *Hum Biol* **70**, 1073-91 (1998).
25. Hofman, A. et al. The Rotterdam Study: 2010 objectives and design update. *Eur J Epidemiol* **24**, 553-72 (2009).
26. Hofman, A. et al. The Rotterdam Study: objectives and design update. *Eur J Epidemiol* **22**, 819-29 (2007).
27. Hofman, A., Grobbee, D.E., de Jong, P.T. & van den Ouweland, F.A. Determinants of disease and disability in the elderly: the Rotterdam Elderly Study. *Eur J Epidemiol* **7**, 403-22 (1991).
28. Pilia, G. et al. Heritability of cardiovascular and personality traits in 6,148 Sardinians. *PLoS Genet* **2**, e132 (2006).
29. Willer, C.J. et al. Newly identified loci that influence lipid concentrations and risk of coronary artery disease. *Nat Genet* **40**, 161-9 (2008).
30. Chen, W.M. & Abecasis, G.R. Family-based association tests for genomewide association scans. *Am J Hum Genet* **81**, 913-26 (2007).
31. Li, Y., Willer, C., Sanna, S. & Abecasis, G. Genotype imputation. *Annu Rev Genomics Hum Genet* **10**, 387-406 (2009).
32. Burdick, J.T., Chen, W.M., Abecasis, G.R. & Cheung, V.G. In silico method for inferring genotypes in pedigrees. *Nat Genet* **38**, 1002-4 (2006).
33. John, U. et al. Study of Health In Pomerania (SHIP): a health examination survey in an east German region: objectives and design. *Soz Praventivmed* **46**, 186-94 (2001).
34. Andrew, T. et al. Are twins and singletons comparable? A study of disease-related and lifestyle characteristics in adult women. *Twin Res* **4**, 464-77 (2001).
35. Spector, T.D. & Williams, F.M. The UK Adult Twin Registry (TwinsUK). *Twin Res Hum Genet* **9**, 899-906 (2006).
36. Ridker, P.M. et al. Rationale, design, and methodology of the Women's Genome Health Study: a genome-wide association study of more than 25,000 initially healthy american women. *Clin Chem* **54**, 249-55 (2008).
37. Lawlor, D.A., Bedford, C., Taylor, M. & Ebrahim, S. Geographical variation in cardiovascular disease, risk factors, and their control in older women: British Women's Heart and Health Study. *J Epidemiol Community Health* **57**, 134-40 (2003).
38. Firmann, M. et al. The CoLaus study: a population-based study to investigate the epidemiology and genetic determinants of cardiovascular risk factors and metabolic syndrome. *BMC Cardiovasc Disord* **8**, 6 (2008).
39. Smeets-Goevaers, C.G. et al. The prevalence of low bone mineral density in Dutch perimenopausal women: the Eindhoven perimenopausal osteoporosis study. *Osteoporos Int* **8**, 404-9 (1998).
40. Daniels, P.R. et al. Familial aggregation of hypertension treatment and control in the Genetic Epidemiology Network of Arteriopathy (GENOA) study. *Am J Med* **116**, 676-81 (2004).
41. The Family Blood Pressure Program Investigators. Multi-center genetic study of hypertension: The Family Blood Pressure Program (FBPP). *Hypertension* **39**, 3-9 (2002).
42. Barker, D.J., Osmond, C., Forsen, T.J., Kajantie, E. & Eriksson, J.G. Trajectories of growth among children who have coronary events as adults. *N Engl J Med* **353**, 1802-9 (2005).
43. Eriksson, J.G., Osmond, C., Kajantie, E., Forsen, T.J. & Barker, D.J. Patterns of growth among children who later develop type 2 diabetes or its risk factors. *Diabetologia* **49**, 2853-8 (2006).
44. Rikkonen, K. et al. Infant growth and hostility in adult life. *Psychosom Med* **70**, 306-13 (2008).
45. Traglia, M. et al. Heritability and demographic analyses in the large isolated population of Val Borbera suggest advantages in mapping complex traits genes. *PLoS One* **4**, e7554 (2009).
46. Deary, I.J. et al. The Lothian Birth Cohort 1936: a study to examine influences on cognitive ageing from age 11 to age 70 and beyond. *BMC Geriatr* **7**, 28 (2007).
47. Houlihan, L.M. et al. Common variants of large effect in F12, KNG1, and HRG are associated with activated partial thromboplastin time. *Am J Hum Genet* **86**, 626-31 (2010).

Nature Genetics: doi:10.1038/ng.1051

48. Stolk, R.P. et al. Universal risk factors for multifactorial diseases: LifeLines: a three-generation population-based study. *Eur J Epidemiol* **23**, 67-74 (2008).
49. Stathopoulou MG, D.G., Trovas G, Theodoraki EV, Katsalira A, Hammond N, Deloukas P, Lyritis GP. . The role of vitamin D receptor gene polymorphisms in the bone mineral density of Greek postmenopausal women with low calcium intake. *J Nutr Biochem* (in press).
50. Stathopoulou, M.G. et al. Low-density lipoprotein receptor-related protein 5 polymorphisms are associated with bone mineral density in Greek postmenopausal women: an interaction with calcium intake. *J Am Diet Assoc* **110**, 1078-83 (2010).
51. Boker, L.K. et al. Prospect-EPIC Utrecht: study design and characteristics of the cohort population. European Prospective Investigation into Cancer and Nutrition. *Eur J Epidemiol* **17**, 1047-53 (2001).
52. Einarsdottir, K. et al. Comprehensive analysis of the ATM, CHEK2 and ERBB2 genes in relation to breast tumour characteristics and survival: a population-based case-control and follow-up study. *Breast Cancer Res* **8**, R67 (2006).
53. Magnusson, C. et al. Breast-cancer risk following long-term oestrogen- and oestrogen-progestin-replacement therapy. *Int J Cancer* **81**, 339-44 (1999).
54. Rahman, I. et al. Genetic dominance influences blood biomarker levels in a sample of 12,000 Swedish elderly twins. *Twin Res Hum Genet* **12**, 286-94 (2009).
55. Servin, B. & Stephens, M. Imputation-based analysis of association studies: candidate regions and quantitative traits. *PLoS Genet* **3**, e114 (2007).
56. Price, A.L. et al. Principal components analysis corrects for stratification in genome-wide association studies. *Nat Genet* **38**, 904-9 (2006).
57. Aulchenko, Y.S., Struchalin, M.V. & van Duijn, C.M. ProbABEL package for genome-wide association analysis of imputed data. *BMC Bioinformatics* **11**, 134.
58. He, C. et al. Genome-wide association studies identify loci associated with age at menarche and age at natural menopause. *Nat Genet* (2009).
59. Patterson, N., Price, A.L. & Reich, D. Population structure and eigenanalysis. *PLoS Genet* **2**, e190 (2006).
60. Pritchard, J.K. & Rosenberg, N.A. Use of unlinked genetic markers to detect population stratification in association studies. *Am J Hum Genet* **65**, 220-8 (1999).
61. Estrada, K. et al. GRIMP: a web- and grid-based tool for high-speed analysis of large-scale genome-wide association using imputed data. *Bioinformatics* **25**, 2750-2 (2009).

Nature Genetics: doi:10.1038/ng.1051