The American Journal of Human Genetics, Volume *94*

Supplemental Data

A Common 16p11.2 Inversion Underlies the Joint Susceptibility to Asthma and Obesity

Juan R. González, Alejandro Cáceres, Tonu Esko, Ivon Cuscó, Marta Puig, Mikel Esnaola, Judith Reina, Valerie Siroux, Emmanuelle Bouzigon, Rachel Nadif, Eva Reinmaa, Lili Milani, Mariona Bustamante, Deborah Jarvis, Josep M. Antó, Jordi Sunyer, Florence Demenais, Manolis Kogevinas, Andres Metspalu, Mario Cáceres, and Luis A. Pérez-Jurado

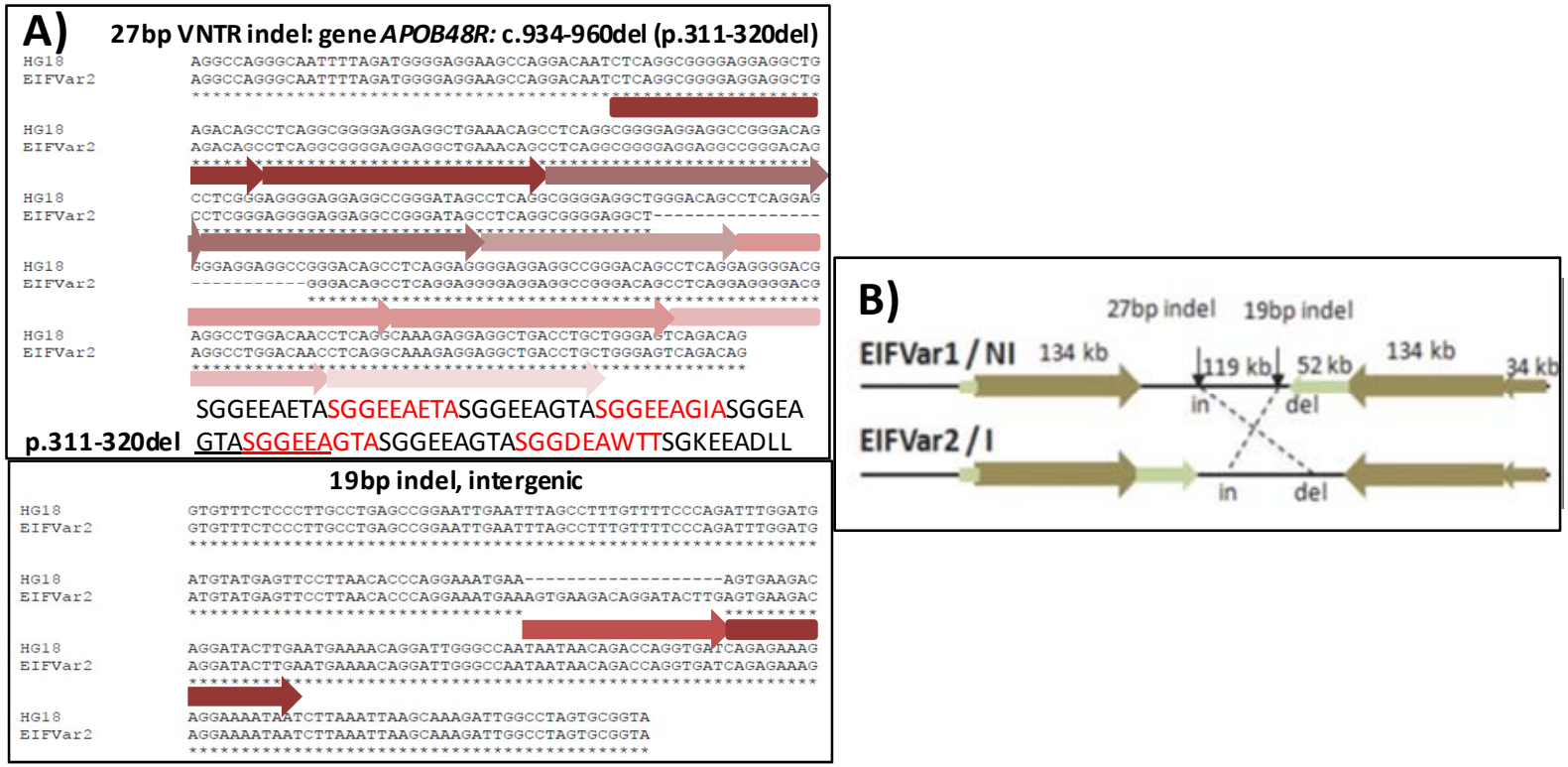

Figure S1. Identification of additional haplotype markers for the NI and I alleles. A) Sequence alignment of the reference haplotype (EIFVar1) and EIFVar2 at the two insertion-deletion polymorphisms. The arrows under the alignments show the 27bp or 19bp repeats in each sequence. The 27bp deletion in the *APOB48R* gene present in the EIFVar2 allele causes the loss of 9 amino acids of this internal repeat in the protein (underlined). **B)** Schematic representation of the EIFVar1/NI and EIFVar2/I alleles at 16p11.2 with the flanking SDs (inverted dark green arrows) and the single copy interval in between with the size of each region in Kb. Other smaller blocks of SDs are shown in light green. The location of the two insertion-deletion polymorphisms and the specific alleles per haplotype are indicated.

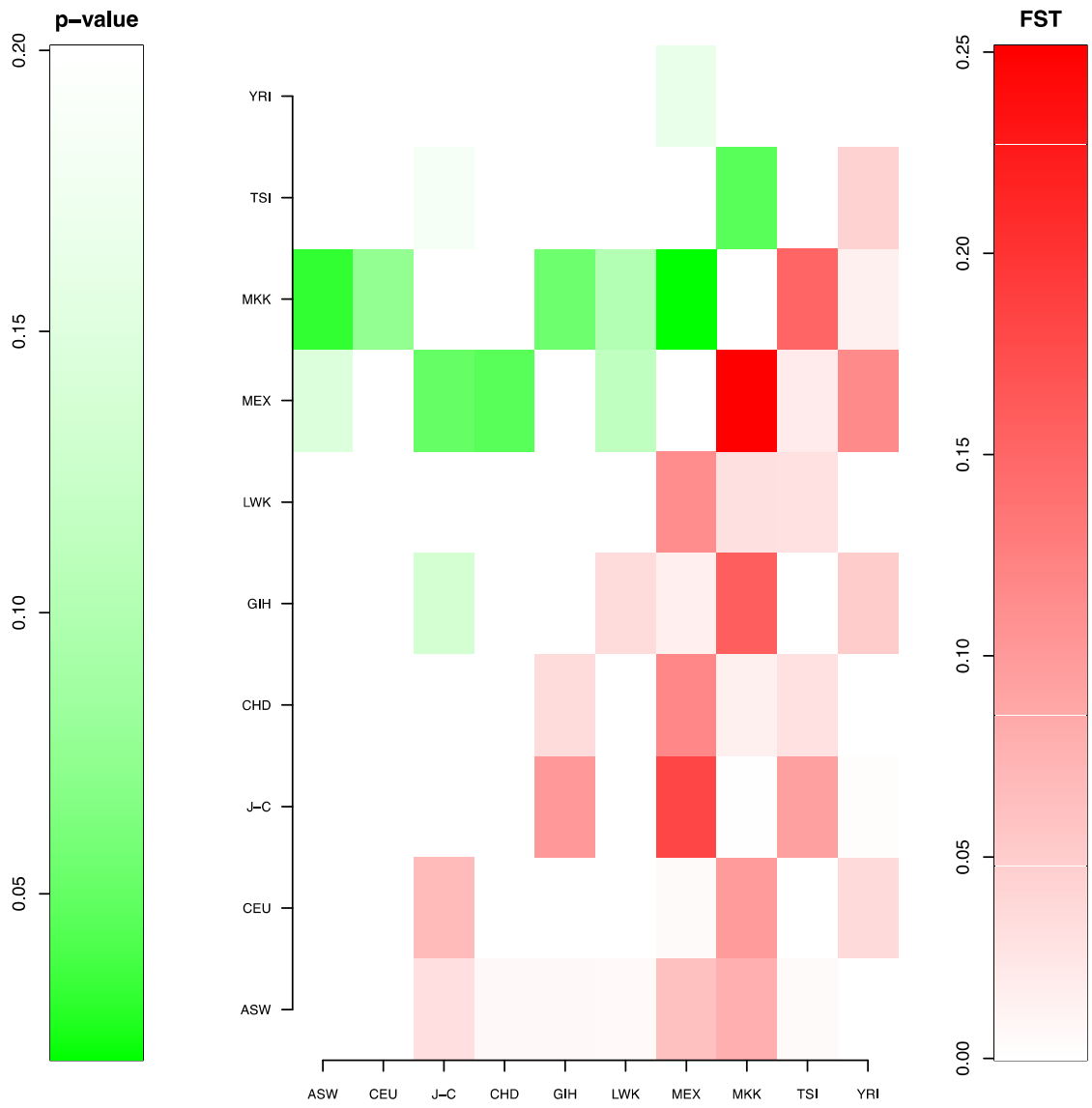

Figure S2. Population stratification of the 16p11.2 inversion. Pairwise F_{ST} between HapMap3 populations based on inversion genotypes and the corresponding significance relative to an empirical distribution of neutral SNPs. Each shaded box represents a pairwise population comparison with higher F_{ST} in red and the corresponding p-values in green, as indicated by the bars on each side. The following populations are included: ASW: African ancestry in Southwest USA; CEU: Utah residents with Northern and Western European ancestry; CHB: Han Chinese in Beijing, China; CHD: Chinese in Metropolitan Denver, Colorado; GIH: Gujarati Indians in Houston, Texas; JPT: Japanese in Tokyo, Japan; LWK: Luhya in Webuye, Kenya; MEX: Mexican ancestry in Los Angeles, California; MKK: Maasai in Kinyawa, Kenya; TSI: Toscans in Italy (n=88), YRI: Yoruba in Ibadan, Nigeria.

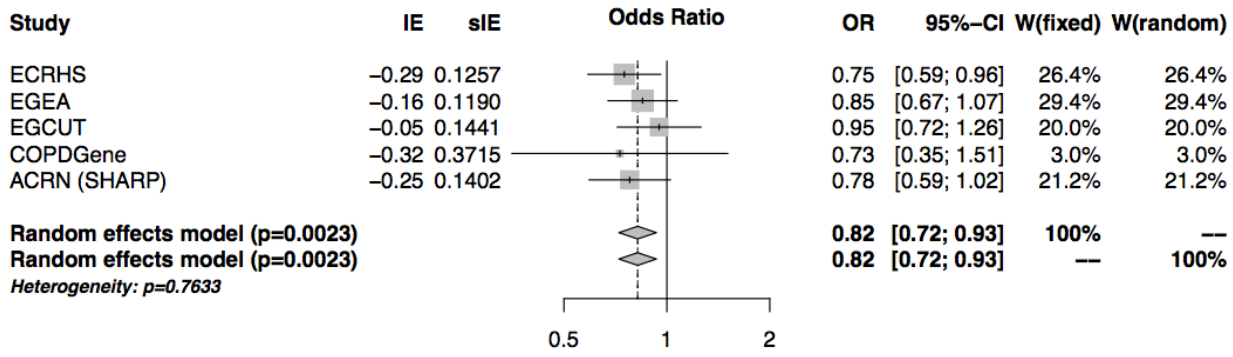

Figure S3. Meta-analysis forest plot for the association studies of the inversion with asthma alone. The inversion-allele effect (IE), the standard error of the effect (sIE), the odds ratio (OR) with confidence intervals (95%-CI) and the relative weight (W) corresponding to the inversion allele effect under a dominant model on asthma are shown for each study. Overall, there is a protective effect of the inversion allele for asthma with an OR of 0.82 (p value of 0.0026).

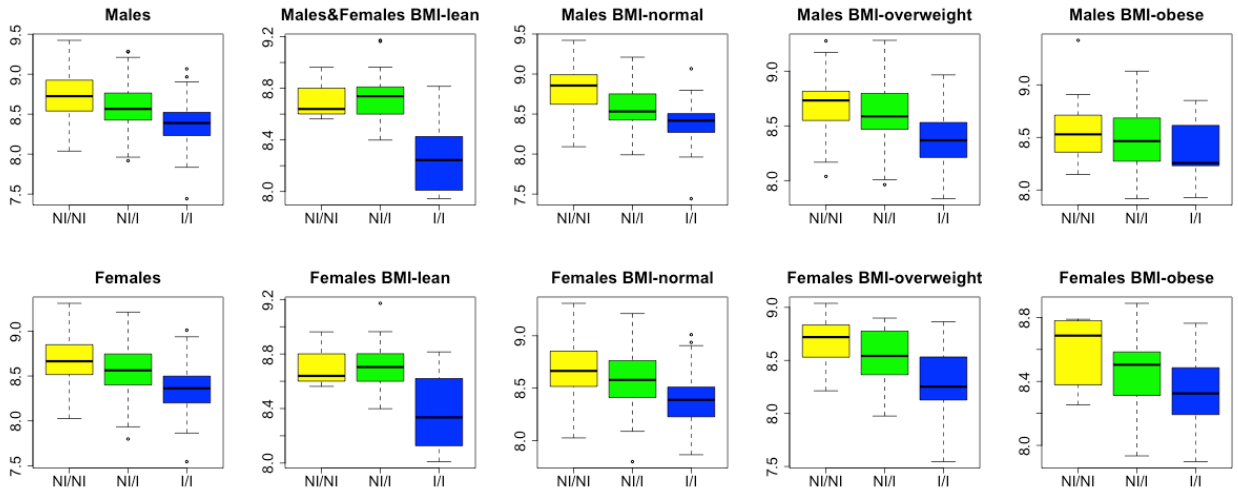

Figure S4. Expression of the *TUFM* gene in the EGCUT gene expression cohort. Boxplots of gene expression for each inversion genotype stratified by gender and BMI (lean: <18; normal: [18 − 25); overweight: [25 − 30); obese: ≥ 30.

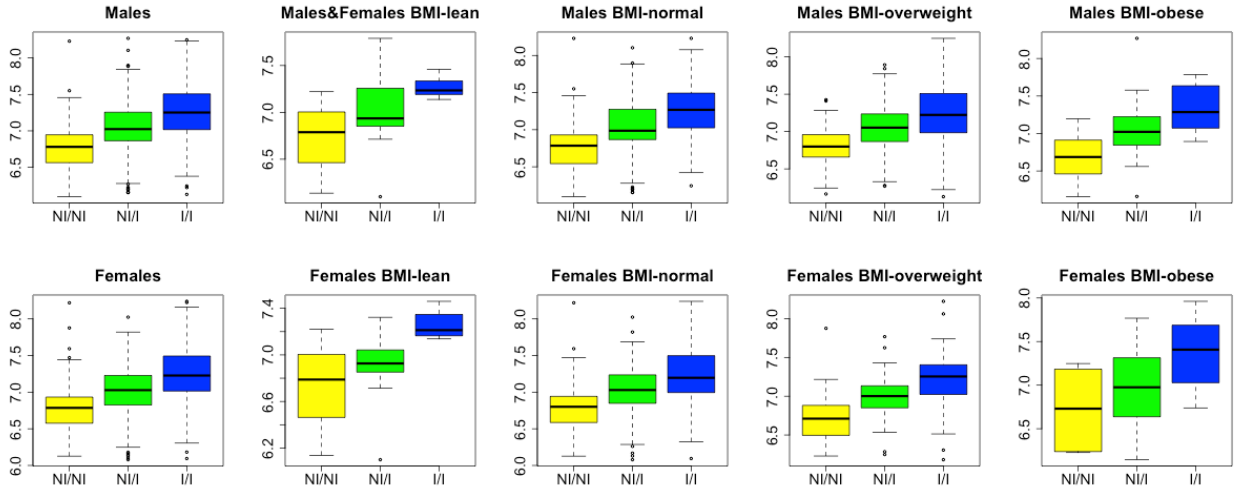

Figure S5. Expression of the *SULT1A4* gene in the EGCUT gene expression cohort. Boxplots of gene expression for each inversion genotype stratified by gender and BMI (lean: <18; normal: [18 − 25); overweight: [25 − 30); obese: ≥ 30.

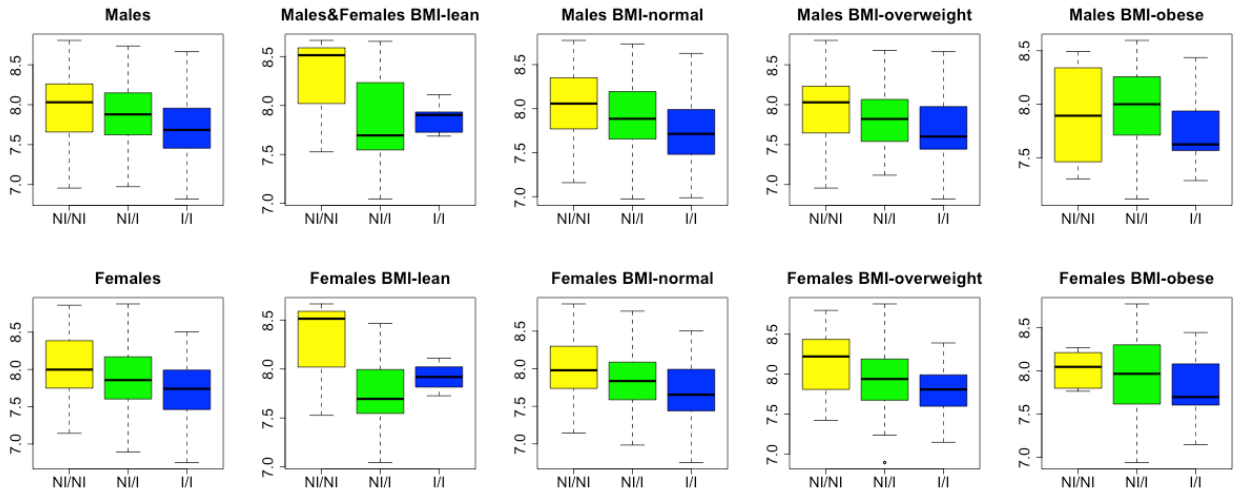

Figure S6. Expression of the *SNPS1* gene in the EGCUT gene expression cohort. Boxplots of gene expression for each inversion genotype stratified by gender and BMI (lean: <18; normal: [18 − 25); overweight: [25 − 30); obese: ≥ 30.

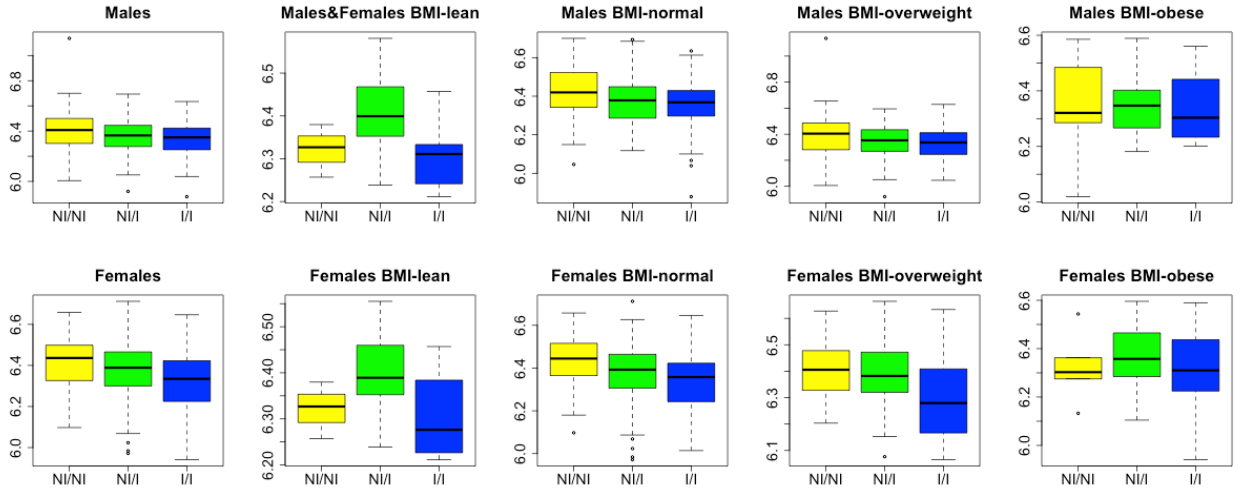

Figure S7. Expression of the *SULT1A1* gene in the EGCUT gene expression cohort. Boxplots of gene expression for each inversion genotype stratified by gender and BMI (lean: <18; normal: [18 − 25); overweight: [25 − 30); obese: ≥ 30.

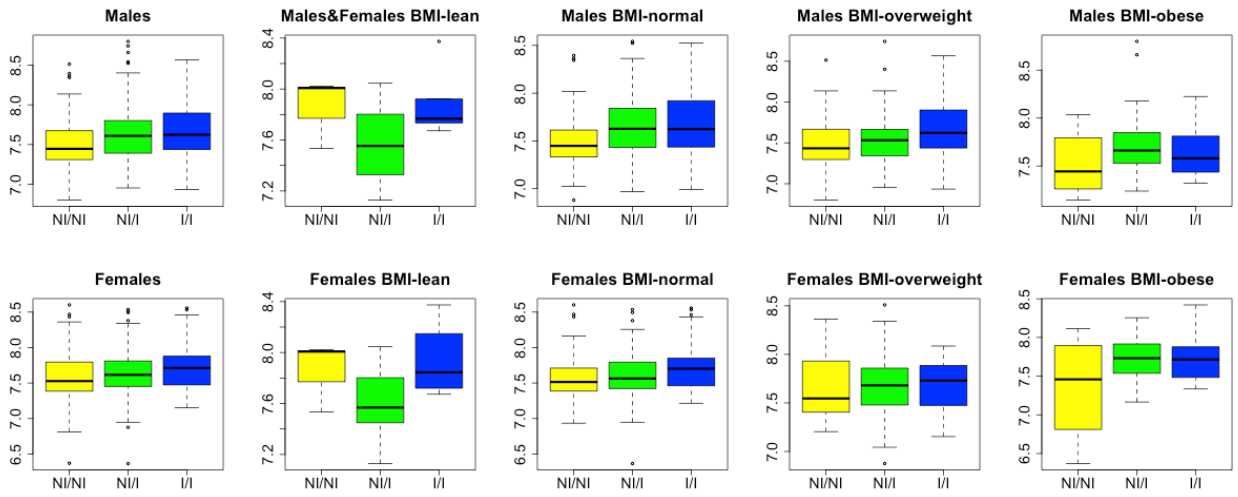

Figure S8. Expression of the *CCDC101* gene in the EGCUT gene expression cohort. Boxplots of gene expression for each inversion genotype stratified by gender and BMI (lean: <18; normal: [18 − 25); overweight: [25 − 30); obese: ≥ 30.

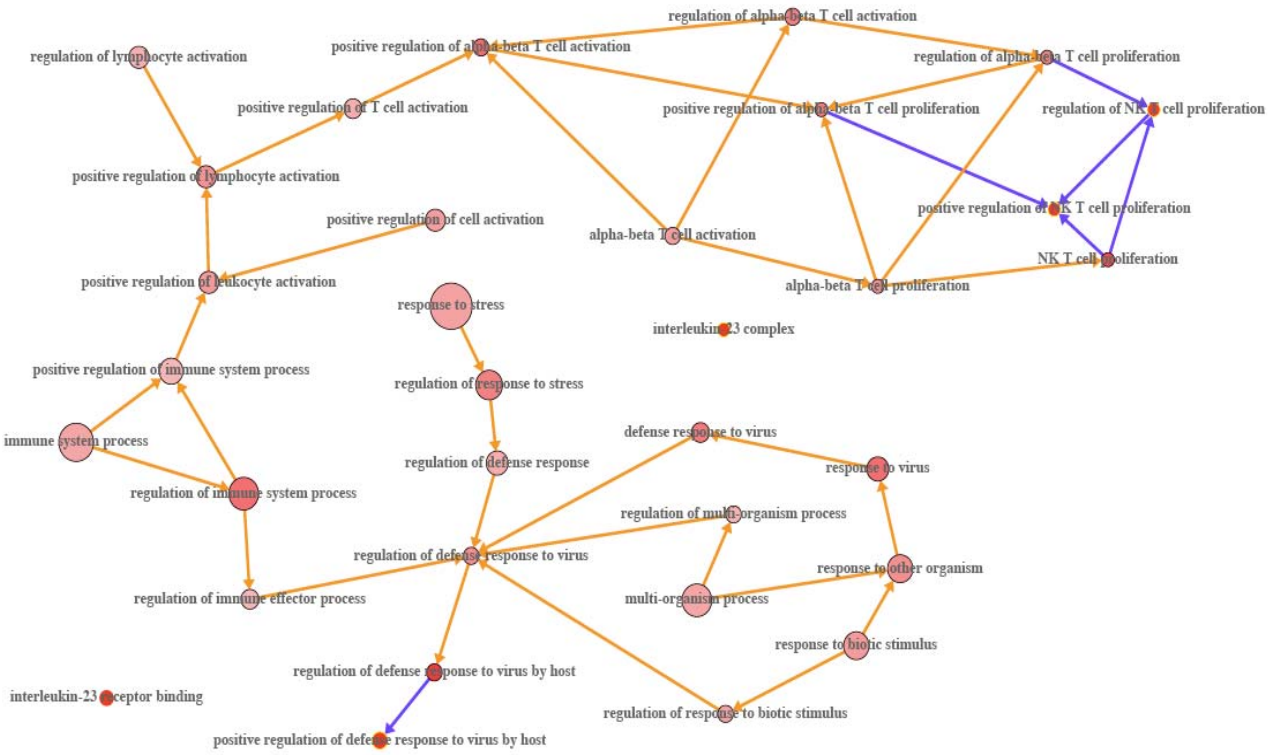

Figure S9. Concept-network resulting from the GO enrichment analysis of transcriptome data from lymphoblastoid cell lines. Gene Ontology (GO) enrichment analysis was performed including those genes (from the whole genome) that were differentially expressed with respect to the inversion at a false discovery rate < 10^{-3}. A concept-network using GeneAnswers bioconductor package was then estimated. Nodes represent the different categories to which the most expressed genes belong, while edges encode relationships between categories. The size of a node is proportional to the number of genes in its GO category. Likewise, the color stands for how relative the given genes are to the GO categories (the redder, the more relative). Finally, dark purple edges indicate the most representative connections between the GO categories.

Study	Frequency	Kappa	% of variance
EGCUT	0.49	0.96	90.00
HapMap 3 CEU	0.36	0.97	95.30
ECRHS-I	0.40	0.95	92.30
EGEA	0.34	0.97	92.90
COPDgene-SHARP	0.36	0.96	94.40

Table S1: Frequency of the 16p11.2 inversion polymorphism in the different cohorts included in the analyses. The first column provides the frequency of the inversion, the second shows the concordance of calls between inveRsion and PFIDO algorithms, and the last column indicates the percentage (%) of variance captured by the inversion.

Population	DNA	PFIDO/inveRsion genotype	19-bp indel	27-bp indel
CEU	NA07029	NI/NI	-	+
CEU	NA10856	NI/NI	-	+
CEU	NA12707	NI/NI	-	+
CEU	NA12236	NI/NI	-	+
CEU	NA12801	NI/NI	-	+
CEU	NA12875	NI/NI	-	+
CHB	NA18542	NI/NI	-	+
CHB	NA18561	NI/NI	-	+
CHB	NA18570	NI/NI	-	+
CHB	NA18573	NI/NI	-	+
CHB	NA18593	NI/NI	-	+
CHB	NA18594	NI/NI	-	+
CHB	NA18620	NI/NI	-	+
CHB	NA18632	NI/NI	-	+
CHB	NA18633	NI/NI	-	+
CHB	NA18636	NI/NI	-	+
JPT	NA18943	NI/NI	-	+
JPT	NA18951	NI/NI	-	+
JPT	NA18973	NI/NI	-	+/-
JPT	NA18995	NI/NI	-	+/-
JPT	NA18999	NI/NI	-	+
YRI	NA18507	NI/NI	-	+
YRI	NA18508	NI/NI	-	+/-
YRI	NA18517	NI/NI	-	+/-
YRI	NA18855	NI/NI	-	+
YRI	NA18861	NI/NI	-	-
YRI	NA18871	NI/NI	-	+
YRI	NA18912	NI/NI	-	+
YRI	NA19103	NI/NI	-	+/-
YRI	NA19127	NI/NI	-	+
YRI	NA19130	NI/NI	-	+
YRI	NA19142	NI/NI	-	+
YRI	NA19144	NI/NI	-	+/-
YRI	NA19173	NI/NI	-	+/-
YRI	NA19203	NI/NI	-	-
CEU	NA06994	NI/I	+/-	+/-
CEU	NA07034	NI/I	+/-	+/-
CEU	NA10846	NI/I	+/-	-
CEU	NA10855	NI/I	+/-	+/-

Population	DNA	PFIDO/inveRsion genotype	19-bp indel	27-bp indel
CEU	NA11839	NI/I	+/-	+/-
CEU	NA11882	NI/I	+/-	+/-
CEU	NA11992	NI/I	+/-	+
CEU	NA12057	NI/I	+/-	+/-
CEU	NA12249	NI/I	+/-	+/-
CEU	NA12264	NI/I	+/-	+/-
CEU	NA12716	NI/I	+/-	+/-
CEU	NA12751	NI/I	+/-	+/-
CHB	NA18564	NI/I	+/-	+/-
CHB	NA18576	NI/I	+/-	+/-
CHB	NA18579	NI/I	+/-	+/-
CHB	NA18622	NI/I	+/-	+/-
JPT	NA18959	NI/I	+/-	+/-
JPT	NA18970	NI/I	+/-	+/-
JPT	NA18987	NI/I	+/-	+/-
YRI	NA18501	NI/I	+/-	+/-
YRI	NA18503	NI/I	+/-	+/-
YRI	NA18852	NI/I	+/-	+*
YRI	NA18862	NI/I	+/-	+*
YRI	NA19119	NI/I	+/-	+*
YRI	NA19143	NI/I	+/-	-
YRI	NA19208	NI/I	+/-	+/-
YRI	NA19209	NI/I	+/-	+/-
YRI	NA19200	NI/I	+/-	+/-
YRI	NA19221	NI/I	+/-	+/-
CEU	NA10854	I/I	+	-
CEU	NA12006	I/I	+	-
CEU	NA12056	I/I	+	-
CEU	NA12145	I/I	+	-
CEU	NA12864	I/I	+/-	-
CHB	NA18608	I/I	+	-
CHB	NA18612	I/I	+	-

Concordant genotypes, both chromosomes
Both discordant chromosomes, 27bp indel
One discordant chromosome, 27bp indel
One discordant chromosome, 19bp indel
*Other genotype, probably corresponding to two copies of the 27bp

Table S2: Concordance between the predicted non-inverted (NI) or inverted (I) genotypes. The table shows the concordance between the bioinformatic methods and the experimental genotypes (+ for insertion, - for deletion) obtained at the two *indel* polymorphisms within the region in 71 Hapmap individuals (142 chromosomes). Full concordance was observed in 141/142 chromosomes at the 19bp *indel* locus and 122/136 at the 27bp *indel* locus (green). A single non-concordant chromosome at the 19bp *indel* locus (yellow) and a total of 14/136 discrepant chromosomes at the 27bp *indel* -10 cases with one discrepant chromosome (light orange) and 2 with both discrepant chromosomes (dark orange)-, were found. Three YRI individuals carried an allele different from the two standard ones with an extra copy of the 27bp repeat.

SNP ID	Coordinate	LD	EIFVar1/NI	EIFVar2/I	Chimp	Orang	Macaq	Denisova	Neand	San1029	Yor927	Han778	Pap542	French521	Gene / location
rs149299	28.392.642	0.804	T	C	T	T	T	T	T	T/T	T/T	T/C	T/T	T/T	*CLN3* intron
rs180743	28.415.145	0.841	C	G	G	G	G	G	G	C/C	G/G	C/G	C/G	C/C	*APOB48R* missense
rs26528	28.425.210	0.834	T	C	C	C	T	C		T/C	T/T	T/C	C/C	T/T	*IL27* intron
rs28698667	28.440.572	0.975	C	T	C	C	C	T	T	C/T	C/C	C/T	T/T	C/C	intergenic
rs4788084	28.447.349	0.982	C	T	C	C		T		C/T	C/C	C/T	T/T	C/C	intergenic
rs12446550	28.450.882	0.980	G	A	A	A		A	A	G/A	G/G	G/A	A/A	G/A	intergenic
rs4788073	28.502.050	0.832	A	G	A	A	A	G		A/G	A/A	A/G	A/G	A/A	*CCDC101* intron
rs2008514	28.733.106	0.962	C	T	T	T	T	T	T	T/T	C/C	C/C	T/T	C/C	intergenic
rs8049439	28.745.016	0.945	T	C	C	T	A	C	C	T/C	T/C	T/T	C/C	T/T	*ATXNL2* intron
rs12928404	28.754.747	0.903	T	C	C	C	T	C		T/C	C/C	T/T	C/C	T/T	*ATXNL2* intron
rs12325113	28.756.169	0.987	T	C	T	T	T	T	T	T/T	T/T	T/T	T/T	T/T	intergenic
rs4788099	28.763.228	0.989	A	G	G	G	G	G	G	A/G	A/G	A/G	G/G	A/A	*TUFM* intron
rs7187776	28.765.146	0.942	A	G	G		G	G		A/A	A/G	A/A	G/G	A/A	*TUFM* 3'utr
rs9972768	28.769.235	0.979	A	C	A	A	A		C	A/C	A/C	A/A	C/C	A/A	intergenic
rs4788101	28.775.305	0.991	C	T	T	T	T	T		C/T	C/C	C/C	T/T	C/C	intergenic
rs11861132	28.779.361	0.987	A	G	G		G	G	G	A/G	A/A	A/A	C/G	A/A	intergenic
rs11861174	28.779.490	0.918	A	G	G	G		G		A/G	A/A	A/G	G/G	A/A	intergenic
rs4788102	28.780.899	0.991	G	A	G	G		A		G/A	G/G	G/G	A/A	A/A	intergenic
rs7193733	28.782.983	0.976	A	G	G		G	G	G	A/A	A/A	A/A	G/G	A/G	*SH2B1* 5'UTR
rs7198606	28.782.623	0.991	T	G	G	G	G	G	G	G	T/T	T/T	G/G	G/T	*SH2B1* 5'UTR
rs7498665	28.790.742	0.939	A	G	G	G	G	G	G	A/G	A/A	A/A	A/A	A/A	*SH2B1* missense
rs11864107	28.793.432	0.970	T	C	C	T	C		C	T/C	T/T	T/T	C/C	T/T	intergenic
rs3888190	28.796.987	0.987	G	T	T	T	T		T	G/T	G/G	G/G	T/T	G/G	intergenic
rs8061590	28.802.631	0.991	A	G	G	G	G	G		A/G	A/A	A/A	G/G	A/A	*ATP2A1* intron

Table S3: Genotypes at 24 representative SNPs on LD>0.8 with inversion haplotypes in the NI and I alleles. The table shows data of Neanderthal, Denisova, non-human primates and representative individuals from modern human populations (San, Yoruba, Han, Papuan & French). The genotypes at the inversion allele coincidental with chimpanzee and Neanderthal or Denisova genotypes are labeled in yellow.

Population	Frequency (%)	95% CI	HWE p-value
ASW	18.1	(12.7-25.0)	1.0000
CEU	35.8	(30.6-41.2)	0.8668
CHB	14.9	(10.0-21.4)	0.5386
CHD	11.8	(7.5-17.8)	1.0000
GIH	28.4	(22.0-35.8)	0.0168
JPT	12.8	(8.4-18.9)	0.5386
LWK	13.3	(8.9-19.4)	0.6936
MXL	39.6	(31.9-47.8)	0.3429
MKK	10.2	(7.3-14.1)	1.0000
TSI	27.3	(34.6-21.0)	0.2804
YRI	18.6	(23.2-14.6)	0.2065

Table S4: Frequency of inversion polymorphism in 16p11.2 across HapMap 3 populations. First column provides an estimation of inversion (I) allele frequencies per population, while second column shows their 95% confidence intervals. The following populations are included in the analysis: ASW: African ancestry in Southwest USA (n=83), CEU: Utah residents with Northern and Western European ancestry from the CEPH collection (n=165), CHB: Han Chinese in Beijing, China (n=84), CHD: Chinese in Metropolitan Denver, Colorado (n=85), GIH: Gujarati Indians in Houston, Texas (n=88), JPT: Japanese in Tokyo, Japan (n=86), LWK: Luhya in Webuye, Kenya (n=90), MXL: Mexican ancestry in Los Angeles, California (n=77), MKK: Maasai in Kinyawa, Kenya (n=171), TSI: Toscans in Italy (n=88), YRI: Yoruba in Ibadan, Nigeria (n=167).

	Controls		Cases				
	n	%	n	%	OR	95%-CI	P-value
ECRHS							
Asthma/Obese							
NI/NI	27	31,8	17	51,5	1		
NI/I + I/I	58	68,2	16	48,5	0,35	(0.14, 0.84)	0,0178
Asthma							
NI/NI	531	34,8	134	40,2	1		
NI/I + I/I	996	65,2	199	59,8	0,75	(0.59, 0.96)	0,0231
Obese/non-asthma							
NI/NI	287	35	27	31,8	1		
NI/I + I/I	533	65	58	68,2	1,18	(0.73, 1.91)	0,49034
EGEA							
Asthma/Obese							
NI/NI	20	30,3	26	53,1	1		
NI/I + I/I	46	69,7	23	46,9	0,33	(0.14, 0.77)	0,0088
Asthma							
NI/NI	334	42	211	46,4	1		
NI/I + I/I	462	58	244	53,6	0,85	(0.67, 1.07)	0,1691
Obese/non-asthma							
NI/NI	115	44,6	20	30,3	1		
NI/I + I/I	143	55,4	46	69,7	1,8	(1.00, 3.24)	0,0447
EGCUT							
Asthma/Obese							
NI/NI	23	20,9	27	32,5	1		
NI/I + I/I	87	79,1	56	67,5	0,45	(0.22, 0.90)	0,0227
Asthma							
NI/NI	213	25,7	105	27,1	1		
NI/I + I/I	616	74,3	283	72,9	0,95	(0.72, 1.26)	0,7175
Obese/non-asthma							
NI/NI	97	24,9	23	20,9	1		
NI/I + I/I	292	75,1	87	79,1	1,09	(0.63, 1.90)	0,7625
COPDGene							
Asthma/Obese							
NI/NI	51	36,2	9	56,2	1		
NI/I + I/I	90	63,8	7	43,8	0,44	(0.15, 1.32)	0,1403
Asthma							
NI/NI	155	36,8	15	45,5	1		
NI/I + I/I	266	63,2	18	54,5	0,73	(0.35, 1.51)	0,3962
Obese/non-asthma							
NI/NI	104	37,1	51	36,2	1		
NI/I + I/I	176	62,9	90	63,8	1,05	(0.69, 1.61)	0,8036
ACRN (SHRAP)							
Asthma/Obese							
NI/NI	51	36,2	63	46,3	1		
NI/I + I/I	90	63,8	73	53,7	0,62	(0.38, 1.01)	0,0542
Asthma							
NI/NI	155	36,8	187	42,5	1		
NI/I + I/I	266	63,2	253	57,5	0,78	(0.59, 1.02)	0,074
Obese/non-asthma							
NI/NI	104	37,1	51	36,2	1		
NI/I + I/I	176	62,9	90	63,8	1,05	(0,69, 1.60)	0,8119

Table S5: Association analyses between the 16p11.2 inversion with asthma, obesity and the joint occurrence of both genotypes in five independent studies. The number and percentage of cases and controls with the studied phenotypes per inversion genotypes and per study are shown. The relative odds ratio (OR) under a dominant model with the confidence intervals (95%-CI) and the p-value for each association are also indicated. The threshold for significance after correcting for the analysis of different genetic models was set at 0.0227 (0.05/2.2) (the number of effective tests was 2.2, as described in the methods section).

Gene	Chr16 location	NI/NI	NI/I	I/I	p-value
SPNS1	28893596-28893596	-	ref	+	1.04e-04
EIF3CL	28298400-28298400	-	ref	+	1.09e-04
IL32	3055313-3055313	-	ref	-	3.44e-03
QPRT	29597941-29597941	-	ref	-	3.80e-03
ZNF768	30442822-30442822	-	ref	-	4.59e-03
NOL3	65761905-65761905	-	ref	-	7.02e-03
MT2A	55199978-55199978	-	ref	-	7.70e-03
UBE2I	1299154-1299154	-	ref	+	8.06e-03
IRX6	53915971-53915971	-	ref	-	9.27e-03
MT1A	55230078-55230078	-	ref	-	1.07e-02
TUFM	28761232-28761232	-	ref	+	1.11e-02
ANKS3	4686511-4686511	-	ref	-	1.17e-02
COG7	23307314-23307314	-	ref	-	1.35e-02
HERPUD1	55523248-55523248	-	ref	-	1.36e-02
FLYWCH1	2901980-2901980	-	ref	-	1.38e-02
IL27	28418183-28418183	-	ref	-	1.44e-02
ZC3H7A	11751942-11751942	-	ref	-	1.66e-02
MAPK8IP3	1696221-1696221	-	ref	-	2.04e-02
FBRS	30583278-30583278	-	ref	-	2.11e-02
SRRM2	2742330-2742330	-	ref	-	2.28e-02
PRR14	30569741-30569741	-	ref	-	2.41e-02
TAOK2	29892688-29892688	-	ref	-	2.44e-02
SLC6A2	54247042-54247042	+	ref	+	2.54e-02
XYLT1	17103681-17103681	-	ref	-	2.63e-02
FAM65A	66120217-66120217	-	ref	-	2.73e-02
SPN	29581771-29581771	-	ref	-	2.84e-02
RSL1D1	11835555-11835555	-	ref	-	2.91e-02
MT1B	55243311-55243311	-	ref	-	2.95e-02
SYNGR3	1979946-1979946	-	ref	-	3.11e-02
PMM2	8799170-8799170	-	ref	-	3.12e-02
GDPD3	30023631-30023631	-	ref	-	3.45e-02
KREMEN2	2954217-2954217	-	ref	-	3.54e-02
YPEL3	30011135-30011135	-	ref	-	3.56e-02
HEATR3	48657381-48657381	+	ref	+	3.66e-02
ZNF646	30993243-30993243	-	ref	-	3.72e-02
SPN	29581771-29581771	-	ref	-	3.89e-02
ATP6V0C	2503727-2503727	-	ref	-	3.92e-02
NDRG4	57055049-57055049	-	ref	-	3.94e-02
SETD1A	30876115-30876115	-	ref	-	3.97e-02
PRSS36	31057747-31057747	-	ref	-	3.98e-02
ZNF263	3273487-3273487	-	ref	-	4.03e-02
MMP25	3036682-3036682	-	ref	-	4.05e-02
OTOA	21597335-21597335	-	ref	-	4.18e-02
ASPHD1	29819647-29819647	+	ref	+	4.18e-02
AGRP	66073974-66073974	+	ref	-	4.40e-02
NOD2	49288550-49288550	-	ref	-	4.43e-02
ZKSCAN2	25154822-25154822	-	ref	-	4.53e-02
TBL3	1962064-1962064	-	ref	+	4.59e-02
RBL2	52025851-52025851	-	ref	-	4.64e-02
DCUN1D3	20776896-20776896	-	ref	-	4.87e-02

Table S6: Expression analysis of chromosome 16 genes of RNA from lymphoblastoid cell lines of 105 CEU samples of HapMap 3, in relation to the inversion genotypes. Only the top-deregulated genes are shown. Columns NI/NI, NI/I, and I/I show increases/decreases (+/-) in gene expression of the homozygous genotypes with respect to the heterozygous inversion (N/I) used as reference. The column 'p-value' gives the p-value of an ANOVA test comparing mean gene levels across inversion genotypes. In bold, the genes located within or immediately flanking the inversion region.

Gene	Chr16 location	NI/NI	NI/NI	I/I	p-value
TUFM	28761848-28761897	-	ref	+	2.96e-40
SULT1A4	29382990-29383039	+	ref	-	7.85e-31
SPNS1	28903301-28903350	-	ref	+	4.20e-13
SULT1A1	28538927-28538955:28541953-28541973	-	ref	+	1.87e-07
CCDC101	28510515-28510528: 28510529-28510564	+	ref	-	2.21e-07
CCDC101	28509634-28509683	+	ref	-	2.30e-05
FLJ12331	67627108-67627157	+	ref	-	4.61e-04
HS.564975	3309854-3309903	-	ref	-	5.32e-04
APOB48R	28417416-28417465	-	ref	+	6.90e-04
EIF3C	28654505-28654552	-	ref	+	1.43e-03
COQ7	18998559-18998608	-	ref	-	1.95e-03
ABCC11	46758582-46758631	+	ref	+	2.32e-03
ZNF469	87034341-87034390	-	ref	+	3.44e-03
HS.116279	79902090-79902139	+	ref	-	3.49e-03
FTSJD1	69874424-69874473	-	ref	-	3.68e-03
HS.128031	24154944-24154957: 24166654-24166689	+	ref	+	3.76e-03
LOC283867	63876501-63876550	-	ref	-	5.78e-03
OSGIN1	82557082-82557113: 82557114-82557131	-	ref	-	6.15e-03
HS.143909	64192713-64192762	-	ref	-	7.04e-03
LOC342346	4546492-4546541	+	ref	+	7.05e-03
HS.578925	9470399-9470448	+	ref	+	7.16e-03
HS.564922		+	ref	+	7.60e-03

Table S7: Expression analysis of chromosome 16 genes of RNA from peripheral blood cells of 882 Estonian samples of the EGCUT, in relation to the inversion genotypes. Only the top-deregulated genes are shown. Columns NI/NI, NI/I, and I/I show increases/decreases (+/-) in gene expression of the homozygous genotypes with respect to the heterozygous inversion (N/I) used as reference. The column 'p-value' gives the p-value of an ANOVA test comparing mean gene levels across inversion genotypes. Gene names with the 'HS.' prefix are Illumina annotated genes, not included in the refSeq database. In bold, the genes located within or immediately flanking the inversion region.

Table S8 (TableS8.xls): Gene Expression analysis stratified by gender and BMI status. Each column provides the −log p-value of association between gene expression and inversion in the EGCUT sample. P-values of association are also given stratified by gender (males and females) and BMI (lean: <18; normal: [18 − 25); overweight: [25 − 30); obese: ≥ 30).

Gene	Gender+BMI+Inversion	Gender	BMI
TUFM	18.2	0.1	1.5
SULT1A4	14.8	0.01	0.05
SPNS1	5.7	0.01	0.01
SULT1A1	3.7	0.01	0.01

Table S9: Top-5 differentially expressed genes against polymorphic inversion 16p11.2, BMI, and gender. The table lists the percentage of observed variability explained by the inversion, BMI and gender, for the top five genes expressed differentially with the inversion genotypes. Results correspond to the R2 measure of a linear regression model.

Acknowledgements by study

Estonian Genome Center, University of Tartu (EGCUT): EGCUT received financing from FP7 programs (ENGAGE, OPENGENE), targeted financing from Estonian Government SF0180142s08, Estonian Research Roadmap through Estonian Ministry of Education and Research, Center of Excellence in Genomics (EXCEGEN) and University of Tartu (SP1GVARENG). We acknowledge EGCUT technical personnel, especially Mr V. Soo and S. Smit. Data analyzes were carried out in part in the High Performance Computing Center of University of Tartu.

European Community Respiratory Health Survey I, ECRHS-I: The coordination of ECRHS II was supported by the European Commission, as part of their Quality of Life programme. The following bodies funded the local studies in ECRHS II: Albacete: Fondo de Investigaciones Santarias (FIS) (grants 97/0035-01, 99/0034-01, and 99/0034-02), Hospital Universitario de Albacete, Consejería de Sanidad; Barcelona: SEPAR, Public Health Service (grant R01 HL62633-01), Fondo de Investigaciones Santarias (FIS) (grants 97/0035-01, 99/0034-01, and 99/0034-02) CIRIT (grant 1999SGR 00241), Red Respira ISCIII; CIBER Epidemiología y Salud Pública (CIBERESP), Spain; Basel: Swiss National Science Foundation, Swiss Federal Office for Education & Science, Swiss National Accident Insurance Fund (SUVA), USC NIEHS Center grant 5P30 ES07048; Bergen: Norwegian Research Council, Norwegian Asthma & Allergy Association (NAAF), Glaxo Wellcome AS, Norway Research Fund; Erfurt: GSF-National Research Centre for Environment & Health, Deutsche Forschungsgemeinschaft (DFG) (grant code FR 1526/1-1); Galdakao: Basque Health Dept; Grenoble: Programme Hospitalier de Recherche Clinique-DRC de Grenoble 2000 no. 2610, Ministry of Health, Direction de la Recherche Clinique, CHU de Grenoble, Ministere de l'Emploi et de la Solidarite, Direction Generale de la Sante, Comite des Maladies Respiratoires de lIsere; Hamburg: GSF-National Reasearch Centre for Environment & Health, Deutsche Forschungsgemeinschaft (DFG) (grant code MA 711/4-1); Ipswich and Norwich: Asthma UK (formerly known as National Asthma Campaign); Huelva: Fondo de Investigaciones Santarias (FIS) (grant code: 97/0035-01, 99/0034-01 and 99/0034-02); Oviedo: Fondo de Investigaciones Santarias (FIS) (grant code: 97/0035-01, 99/0034-01 and 99/0034-02); Paris: Ministere de l'Emploi et de la Solidarite, Direction Generale de la Sante, UCB-Pharma (France), Aventis (France), Glaxo France, Programme Hospitalier de Recherche Clinique-DRC de Grenoble 2000 no. 2610, Ministry of Health, Direction de la Recherche Clinique, CHU de Grenoble; Tartu: Estonian Science Foundation; Umea: Swedish Heart Lung Foundation, Swedish Foundation for Health Care Sciences & Allergy Research, Swedish Asthma & Allergy Foundation, Swedish Cancer & Allergy Foundation; Uppsala: Swedish Heart Lung Foundation, Swedish Foundation for Health Care Sciences & Allergy Research, Swedish Asthma & Allergy Foundation, Swedish Cancer & Allergy Foundation; Financial support for ECRHS I for centers in ECRHS II was provided by: Ministere de la Sante, Glaxo France, Institut Pneumologique dAquitaine, Contrat de Plan Etat-Region Languedoc-Rousillon, CNMATS, CNMRT (90MR/10, 91AF/6), Ministre delegue de la sante, RNSP, France; GSF, and the Bundesminister fur Forschung und Technologie, Bonn, Germany; Norwegian Research Council project no. 101422/310; Ministero de Sanidad y Consumo FIS (grants number 91/0016060/00E-05E and 93/0393), and grants from Hospital General de Albacete, Hospital General Juan Ramón Jiménez, Consejería de Sanidad Principado de Asturias, Spain; The Swedish

Medical Research Council, the Swedish Heart Lung Foundation, the Swedish Association against Asthma and Allergy; Swiss National Science Foundation grant 4026-28099; National Asthma Campaign, British Lung Foundation, Department of Health, South Thames Regional Health Authority, UK. A.R. was supported by the Department of Health, UK and the European Commission as part of GABRIEL contract number 018996 under the Integrated Program LSH-2004-1.2.5-1.

Epidemiological study on Genetics and Environment of Asthma, EGEA The EGEA study was supported, in part, by grants from: Merck Sharp & Dohme (MSD), the Hospital Program of Clinical Research (PHRC-Paris, PHRC-Grenoble), National Research Agency, Health environment, health-work program (ANR 05-SEST-020-02/05-9-97), Biological collections for Health Program (ANR-06-CEBS), and ANR Public Health Research Program, French Agency of health safety, environment and work (AFSSET, EST-09-15), the national scientific committee of the medico-technology support at home (AGIR a‘ dom) and the Isere committee against 16 respiratory diseases (COMARES), the GA2LEN project, Global Allergy and Asthma European Network and GABRIEL, a multidisciplinary study to identify the genetic and environmental causes of asthma in the European Community (contract 018996 from the European Commission).

The NCBI database of Genotypes and Phenotypes (dbGaP) (www.ncbi.nlm.nih.gov/gap). Controlled-access genotype and phenotype data was obtained following authorization by the appropriate Data Access Committee to the following studies: the COPDGene cohort (dbGaP Study Accession number phs000179.v3.p2), and the National Heart, Lung, and Blood Institute SNP Health Association Asthma Resource Project – SHARP (dbGaP Study Accession number phs000166.v2.p1), that includes the Asthma Clinical Research Network (ACRN).