SUPPLEMENTARY MATERIAL

Supplementary methods 1

Genetic data in the ALSPAC cohort Genotypic information was available on 9912 of the study participants using the Illumina HumanHap550 quad genome-wide SNP genotyping platform (8). Following quality control checks on the genotypic data, participants were removed from further analysis if they had incorrect sex assignments; minimal or excessive heterozygosity; disproportionate levels of individual missingness; evidence of cryptic relatedness (i.e. kinship among participants that is not known to the investigator) or being of non-European ancestry (as detected by a multidimensional scaling analysis seeded with HapMap 2 participants) as reported previously (8, 9). SNPs that had a minor allele frequency (MAF) of <1% and call rate of <95% were removed, along with SNPs that did not pass an exact test of Hardy-Weinberg equilibrium ($P>5x10^{-7}$). The resulting dataset contained 8237 individuals.

We estimated haplotypes using ShapeIT (v2.r644), which utilizes relatedness during phasing. We obtained a phased version of the 1000 genomes reference panel (Phase 1, Version 3) from the Impute2 reference data repository (phased using ShapeIT v2.r644 (36), haplotype release date Dec 2013). Imputation of the target data was performed using Impute V2.2.2 (37, 38) against the reference panel (all polymorphic SNPs excluding singletons), using all 2186 reference haplotypes (including non-Europeans).

Supplementary methods 2: Additional sample descriptions

1000 Genomes European (EUR)

Data were generally from the 1000 Genomes project (http://www.1000genomes.org/) using the European subpopulations (Utah residents with Northern and Western European Ancestry (CEU), British in England and Scotland (GBR), Toscani in Italia (TSI) and Finnish in Finland (FIN)) with a sample size of 503 unrelated individuals (18). SNPs in the GWAS catalog were selected from this dataset (n=14,421) and then the data were pruned for LD using the –indep-pairwise function in PLINKv1.9 (14, 39) and the same parameters that were used for the ALSPAC sample (see main

manuscript). Areas of known long range LD were removed (15). This resulted in 9902 LD independent GWAS catalogue SNPs.

Wellcome Trust Case Control Consortium (WTCCC)

Data were from the WTCCC using only control individuals that were genotyped on 500K Affymetrix chip (17) (Affymetrix, Inc., Santa Clara, CA, United States) (http://www.wtccc.org.uk/). Data were combined using all SNPs in common. Data underwent routine quality control, removing SNPs genotype completeness less than 95%, minor allele frequency less than 1% and Hardy-Weinberg equilibrium p-value less than $1x10^{-6}$. Genotypes were imputed using the pre-phasing/imputation stepwise approach implemented in IMPUTE2/SHAPEIT (38, 40) and 1000Genomes (December 2013, release 1000 Genomes haplotypes Phase I integrated variant set) as the reference dataset. Data was convert to best guess genotypes using PLINKv1.9 (14) and subsequently filtered to have an imputation info score greater than 0.8 and minor allele frequency great than 1%. Only control individuals were used for the analysis. The resulting dataset contained 5,355,033 SNPs and 2938 individuals. SNPs in the GWAS catalog were selected from this dataset (n=10,134) and then the data was pruned for LD using the –indep-pairwise function in PLINKv1.9 (14) and the same parameters that were used for the ALSPAC sample (see main manuscript). Areas of known long range LD were removed (15). This resulted in 6943 LD independent GWAS catalogue SNPs.

Supplementary Table 1. Summary of included phenotypes.

Source and Additional Information	Phenotype	Variable Type	Total N	N cases (%) [a]	Mean (sd) [b]	N categories (N range) [c]
Child Age 7 Clinic	Leg length (cm)	Continuous	6,082	-	57.89 (3.42)	-
	Head circumference (cm)	Continuous	6,084	-	52.51 (1.45)	-
http://www.bristol.ac.uk/alspac/researchers/resources-available/data-details/data-tables/documents/focusclinicsessions.pdf	Arm circumference (cm)	Continuous	6,074	-	18.96 (2.11)	-
	Waist circumference (cm)	Continuous	6,079	-	56.42 (5.17)	-
	Hip circumference (cm)	Continuous	6,077	-	65.50 (5.30)	-
	Forearm (cm)	Continuous	6,059	-	18.33 (1.13)	-
	BMI	Continuous	6,027	-	16.23 (2.04)	-
	Scoliometer measure	Ordinal	5,969	-	1.54 (1.15)	-
	Any flexural dermatitis	Binary	6,065	468 (7.7)	-	-
	Allergy: Reaction to DP	Binary	5,399	667 (12.4)	-	-
	Allergy: Reaction to dog	Binary	1,895	89 (4.7)	-	-
	Air conduction right ear average 1,2,3 kHz	Continuous	5,482	-	8.18 (7.07)	-
	Air conduction left ear average 1,2,3 kHz	Continuous	5,416	-	7.90 (7.28)	-
	Bone conduction hearing threshold level (dBHL) 1kHz	Continuous	3,111	-	3.81 (5.38)	-
	Right air bone gap 1kHz	Continuous	5,393	-	11.20 (8.37)	-
	Hearing impairment	Categorical	5,696	-	-	4 (133-5,248)
	Sensorineural hearing loss	Categorical	5,407	-	-	9 (3-5,287)
	High frequency hearing loss	Categorical	5,687	-	-	4 (20-5,687)
	Bilateral tympanometry	Categorical	5,791	-	-	25 (4-4,203)
	Handedness	Binary	6,002	716 (11.9)	-	-
	Reading score	Continuous	5,986	-	28.45 (9.25)	-
	Spelling score	Continuous	5,890	-	7.88 (4.39)	-
	Phoneme task score	Continuous	6,014	-	20.22 (9.51)	-
	Mean diastolic blood pressure	Continuous	6,013	-	56.32 (6.62)	-
	Mean pulse	Continuous	6,011	-	82.99 (10.80)	-
	Hemoglobin value	Continuous	4,761	-	124.48 (7.86)	-
	Heel to Toe test	Ordinal	5,414	-	8.70 (5.71)	-
	Peg game test	Categorical	5,436	-	-	6 (18-4,955)
	String game test	Continuous	4,950	-	23.16 (7.71)	-
	Bean bag game test	Ordinal	5,415	-	5.83 (2.08)	-
	Observable vision abnormality	Binary	5,896	139 (2.4)	-	-
	Summary of eye preference	Categorical	5,772	-	-	6 (11-3,343)
	Manifest strabismus	Categorical	5,922	-	-	5 (9-5,774)
	Binocular function group	Categorical	5,931	-	-	6 (10-4,163)
	Combined pursuit eye movements	Categorical	5,922	-	-	4 (90-5,481)

	Mean convergence (vision)	Continuous	5,883	-	6.34 (1.21)	-
	Near vision (unaided)	Categorical	5,877	-	-	5 (8-5,807)
	Mean accom (unaided)	Continuous	5,849	-	8.11 (1.79)	-
	1st degree relative has a lazy eye	Binary	5,931	619 (10.4)	-	-
Mother questionnaires during pregnancy http://www.bristol.ac.uk/alspac/researchers/resources-available/data-details/questionnaires/#carer	Eat or drink health foods	Binary	5,851	485 (8.3)	-	-
	Highest education qualification	Binary	6,285	2,712 (43.2)	-	-
	Contraceptive pill use	Binary	6,280	5,955 (94.8)	-	-
	History of diabetes	Binary	6,271	63 (1.0)	-	-
	History of hypertension	Binary	6,183	884 (14.3)	-	-
	Had bulimia	Binary	6,293	146 (2.3)	-	-
	Had asthma	Binary	6,293	715 (11.4)	-	-
	Had eczema	Binary	6,293	1,464 (23.3)	-	-
	Had epilepsy	Binary	6,293	70 (1.1)	-	-
	Had kidney disease	Binary	6,293	258 (4.1)	-	-
	Had rheumatism	Binary	6,293	263 (4.2)	-	-
	Had arthritis	Binary	6,293	203 (3.2)	-	-
	Had psoriasis	Binary	6,293	211 (3.4)	-	-
	Had alcoholism	Binary	6,293	49 (0.8)	-	-
	Had anorexia nervosa	Binary	6,293	119 (1.9)	-	-
	Had severe depression	Binary	6,293	456 (7.3)	-	-
	Any allergies	Binary	6,212	2,797 (45.0)	-	-
	Height (cm)	Continuous	6,196	-	164.28 (6.67)	-
	Dwelling type	Categorical	6,308	-	-	6 (51-2,323)
	Number of medications during pregnancy	Ordinal	6,375	-	1.49 (1.44)	-
	Cups of coffee per week	Continuous	5,969	-	7.84 (11.18)	-
	Measures of alcohol per week	Continuous	5,743	-	1.62 (3.73)	-
	Partner smokes	Binary	6,243	1,951 (31.3)	-	-
	Self-induced vomiting during pregnancy	Binary	5,980	60 (1.1)	-	-
	Alcohol use in second trimester of pregnancy	Binary	6,138	1,165 (19.0)	-	-
	Mother had influenza during pregnancy	Binary	4,493	785 (17.5)	-	-
	Mother had shock or freight during pregnancy	Binary	4,499	578 (12.9)	-	-
	Mother smoked during last 2 months of pregnancy	Binary	4,759	621 (13.1)	-	-
	Mother used cannabis during pregnancy	Binary	4,594	93 (2.0)	-	-
	Maximum postnatal depression score during pregnancy (EPND)	Continuous	4,756	-	7.54 (4.62)	-
Mother questionnaires and clinics after pregnancy http://www.bristol.ac.uk/alspac/researchers/resources-	Mothers postnatal weight (lb.)	Continuous	5,173	-	143.77 (23.18)	-
	Mother had alcohol problem since child age 5	Binary	5,541	65 (1.2)	-	-
	Mother had cancer since child age 5	Binary	5,498	40 (0.7)	-	-
	Used cannabis since child age 5	Binary	5,535	238 (4.3)	-	-

available/data-details/questionnaires/#carer	Amphetamine use since child age 5	Binary	5,520	42 (0.8)	-	-
	Number of cigarettes partner smokes per day	Continuous	5,064	-	3.16 (7.21)	-
	Partners alcohol consumption	Binary	5,165	1,810 (35.0)	-	-
Child based questionnaires completed by the Mother when the child was around 7 years of age	Parenting score	Continuous	5,749	-	48.19 (7..44)	
	Child had diarrhea in past 12 months	Binary	5,931	1,971 (33.7)	-	-
	Child had blood in stools in the past 12 months	Binary	5,512	36 (0.7)	-	-
	Child had vomiting in the past 12 months	Binary	5,558	2,401 (43.2)	-	-
	Child had a cough in the past 12 months	Binary	5,590	4,239 (75.8)	-	-
http://www.bristol.ac.uk/alspac/researchers/resources-available/data-details/questionnaires/documents/questionnaire-topic-guide-dec11.pdf	Child had a high temperature in the past 12 months	Binary	5,576	2,890 (51.8)	-	-
	Child had a rash in past 12 months	Binary	5,718	991 (17.3)	-	-
	Child had wheezing in the past 12 months	Binary	5,728	605 (10.6)	-	-
	Child had breathlessness in past 12 months	Binary	5,723	323 (5.6)	-	-
	Child had urinary infection in past 12 months	Binary	5,709	166 (2.9)	-	-
	Child had headaches in past 12 months	Binary	5,723	2,673 (46.7)	-	-
	Child had constipation in past 12 months	Binary	5,720	561 (9.8)	-	-
	Child had eczema in past 12 months	Binary	5,720	941 (16.5)	-	-
	Child had hay fever in past 12 months	Binary	5,704	496 (8.7)	-	-
	Doctor has ever said child has asthma	Binary	5,677	1,132 (19.9)	-	-
	Child had an eye infection in past 12 months	Binary	5,721	183 (3.2)	-	-
	Child had an ear infection in past 12 months	Binary	5,719	792 (13.9)	-	-
	Child had a chest infection in past 12 months	Binary	5,712	339 (5.9)	-	-
	Child had tonsillitis/laryngitis in past 12 months	Binary	5,722	403 (7.0)	-	-
	Child had scarlet fever in past 12 months	Binary	5,717	21 (0.4)	-	-
	Child had influenza in past 12 months	Binary	5,696	303 (5.3)	-	-
	Any compulsions	Binary	5,652	510 (9.0)	-	-
	Any general anxieties	Binary	5,704	2,877 (50.4)	-	-
	Any attention symptoms	Binary	5,693	2,922 (51.3)	-	-
	Any awkward behaviors	Binary	5,691	1,968 (34.6)	-	-
	Any troublesome behaviors	Binary	5,714	1,856 (32.5)	-	-
	Any educational problems	Binary	5,618	741 (13.2)	-	-
	Clinical diagnosis of ADHD	Binary	5,727	114 (2.0)	-	-
	Clinical diagnosis of conduct disorder	Binary	5,727	181 (3.2)	-	-
	Clinical diagnosis of anxiety disorder	Binary	5,747	165 (2.9)	-	-
	Clinical diagnosis of depressive disorder	Binary	5,747	24 (0.4)	-	-
	Child had an accident in the past year	Binary	5,799	669 (11.5)	-	-
	Child had tonsils removed	Binary	5,799	194 (3.4)	-	-
	Child had teeth extracted	Binary	5,799	521 (9.0)	-	-
	Child snored in past year	Binary	5,454	3,424 (63.3)	-	-
	Conduct problems score	Ordinal	5,615	-	1.56 (1.42)	-
	Child moved home since 5th birthday	Binary	5,772	922 (16.0)	-	-
	Fine motor score	Continuous	5,762	-	1.98 (0.09)	-

	Cognitive score	Continuous	5,775	-	1.96 (0.11)	-
	Communication score	Continuous	5,772	-	1.96 (0.13)	-
	Child activity score	Continuous	5,799	-	17.68 (4.92)	-
	Child has tried alcohol	Binary	5,401	2,498 (46.3)	-	-
	Parity	Categorical	6,494	-	-	4 (302-2,935)
	Number of awkward behaviors	Ordinal	5,773	-	1.19 (2.16)	

[a] binary variables only
[b] continuous and ordinal variables only
[c] categorical variables, number of categories and range of individuals in each of those categories

Supplementary Table 2. Summary of top phenotype-genotype associations

SNP	Chr	Gene/closet gene	Trait[1]	Phenotype[1]	p-value	Effect size (95% CI)	MAF
Associations from pruned dataset							
rs9568856	13	*OLFM4*	Obesity	Hip circumference	3.698×10^{-8}	coef (cm) = 0.79 (0.51, 1.08)	0.129
rs4820268	22	*TMPRSS6*	Variety of hematology traits	Hemoglobin levels	3.705×10^{-8}	coef (grams per liter) = -0.89 (-1.20, -0.57)	0.482
rs9568856	13	*OLFM4*	Obesity	Arm circumference	4.662×10^{-8}	coef (cm) = 0.31 (0.20, 0.42)	0.129
Additional associations from unpruned dataset							
rs855791	22	*TMPRSS6*	Variety of hematology traits	Hemoglobin levels	8.241×10^{-12}	coef (grams per liter) = -1.10 (-1.42, -0.79)	0.458
rs9568867	13	*OLFM4*	BMI/Obesity	Arm circumference	2.191×10^{-10}	coef (cm) = 0.37 (0.26, 0.48)	0.125
rs9568867	13	*OLFM4*	BMI/Obesity	BMI	9.890×10^{-9}	coef (kg/m^2) = 0.32 (0.21, 0.43)	0.125
rs9568867	13	*OLFM4*	BMI/Obesity	Hip circumference	1.059×10^{-8}	coef (cm) = 0.84 (0.55, 1.12)	0.125
rs2413450	22	*TMPRSS6*	Variety of hematology traits	Hemoglobin levels	2.899×10^{-8}	coef (grams per liter) = -0.89 (-1.21, -0.58)	0.482
rs2290400	17	*GSDMB*	Type I diabetes	Asthma	4.511×10^{-8}	OR (of having asthma) = 1.30 (1.18, 1.42)	0.483

[1] Trait relates to the broad phenotype reported in the NHGRI-EBI GWAS catalog; phenotype relates to the phenotype examined in the study reported in this manuscript

Supplementary Table 3. Summary of phenotypic and genotypic pairwise associations observed between 121 phenotypes and genotypes (pruned set, genome-wide significant SNPs only).

Number of SNPs	Associations	Cut off *P* value	Observed N (%)	Expected N (%)	χ^2 (d.f, *P* value)	O:E ratio ((95% CI), *P* value)
2,759	333,839	0.05	16,967 (5.1)	16,692 (5.0)	4.77 (1, 0.029)	1.02 ((1.00,1.04), 0.124)
		0.01	3,661 (1.1)	3,338 (1.0)	31.57 (1, 1.92×10^{-8})	1.10 ((1.05,1.15), 0.0001)
		0.0001	66 (0.02)	33 (0.01)	33.00 (1, 9.2×10^{-9})	2.00 ((1.32,3.04), 0.0011)

Supplementary Table 4. Number of Bonferroni significant genotype-genotype correlations across the different datasets.

Dataset	Bonferroni P value	# Bonferroni significant	# in total	Percentage at Bonferroni P value (%)
ALSPAC only	$1.89x10^{-4}$	8876	46,885,086	0.0189
EUR only	$3.03x10^{-5}$	1534	50,566,596	0.0030
WTCCC only	$1.26x10^{-4}$	3042	24,099,153	0.0126
ALSPAC & EUR	$2.56x10^{-5}$	1090	42,573,378	0.0026
ALSPAC & WTCCC	$1.17x10^{-4}$	2397	20,419,245	0.0117
EUR & WTCCC	$3.09x10^{-4}$	624	20,177,128	0.0031
ALSPAC & EUR & WTCCC	$1.41x10^{-5}$	569	19,012,861	0.0030

ALSPAC is the ALSPAC sample results, EUR is the 1000 Genomes European populations and WTCCC is the Wellcome Trust Case Control Consortium.
* The percentage of overlapping genotype-genotype correlations that were Bonferroni significant in each individual dataset from the total of overlapping genotype-genotype correlations between the three datasets.

Supplementary Table 5. Summary of analysis within and between 177 phenotypes (including biomedical and serological phenotypes*) and genotypes (unpruned and pruned for LD)

Number of SNPs	Associations	Cut off *P* value	Observed N (%)	Expected N (%)	χ^2 (d.f, *P* value)	O:E ratio ((95% CI), *P* value)
WITHIN PHENOTYPES						
N/A	15,576	0.05	4,818 (30.93)	779 (5.0)	22044.11 (1, $\leq 1x10^{-50}$)	6.18 ((5.75, 6.65), $\leq 1x10^{-50}$)
		0.01	3,557 (22.77)	156 (1.0)	74869.28 (1, $\leq 1x10^{-50}$)	22.80 ((19.45, 26.72), $\leq 1x10^{-50}$)
		0.0001	2,270 (14.57)	2 (0.01)	2572242.3 (1, $\leq 1x10^{-50}$)	1135.00 ((283.74, 4540.19), $\leq 1x10^{-50}$)
BETWEEN ALL CATALOG SNPS AND PHENOTYPES						
13,720	2,428,440	0.05	131,972 (5.43)	121,442 (5.0)	961.10 (1, $\leq 1x10^{-50}$)	1.09 ((1.08, 1.09), $\leq 1x10^{-50}$)
		0.01	29,685 (1.22)	24,284 (1.0)	1213.37 (1, $\leq 1x10^{-50}$)	1.22 ((1.20, 1.24), $\leq 1x10^{-50}$)
		0.0001	1,486 (0.06)	243 (0.01)	6358.86 (1, $\leq 1x10^{-50}$)	6.12 ((5.34, 7.00), $\leq 1x10^{-50}$)
BETWEEN LD PRUNED CATALOG SNPS AND PHENOTYPES						
9,684	1,714,068	0.05	89,912 (5.25)	85,703 (5.0)	217.59 (1, $3.04x10^{-49}$)	1.05 ((1.04, 1.06), $\leq 1x10^{-50}$)
		0.01	18,941 (1.11)	17,141 (1.0)	190.93 (1, $1.99x10^{-43}$)	1.11 ((1.08, 1.13), $\leq 1x10^{-50}$)
		0.0001	373 (0.02)	171 (0.01)	238.64 (1, $\leq 1x10^{-50}$)	2.18 ((1.82, 2.61), $\leq 1x10^{-50}$)

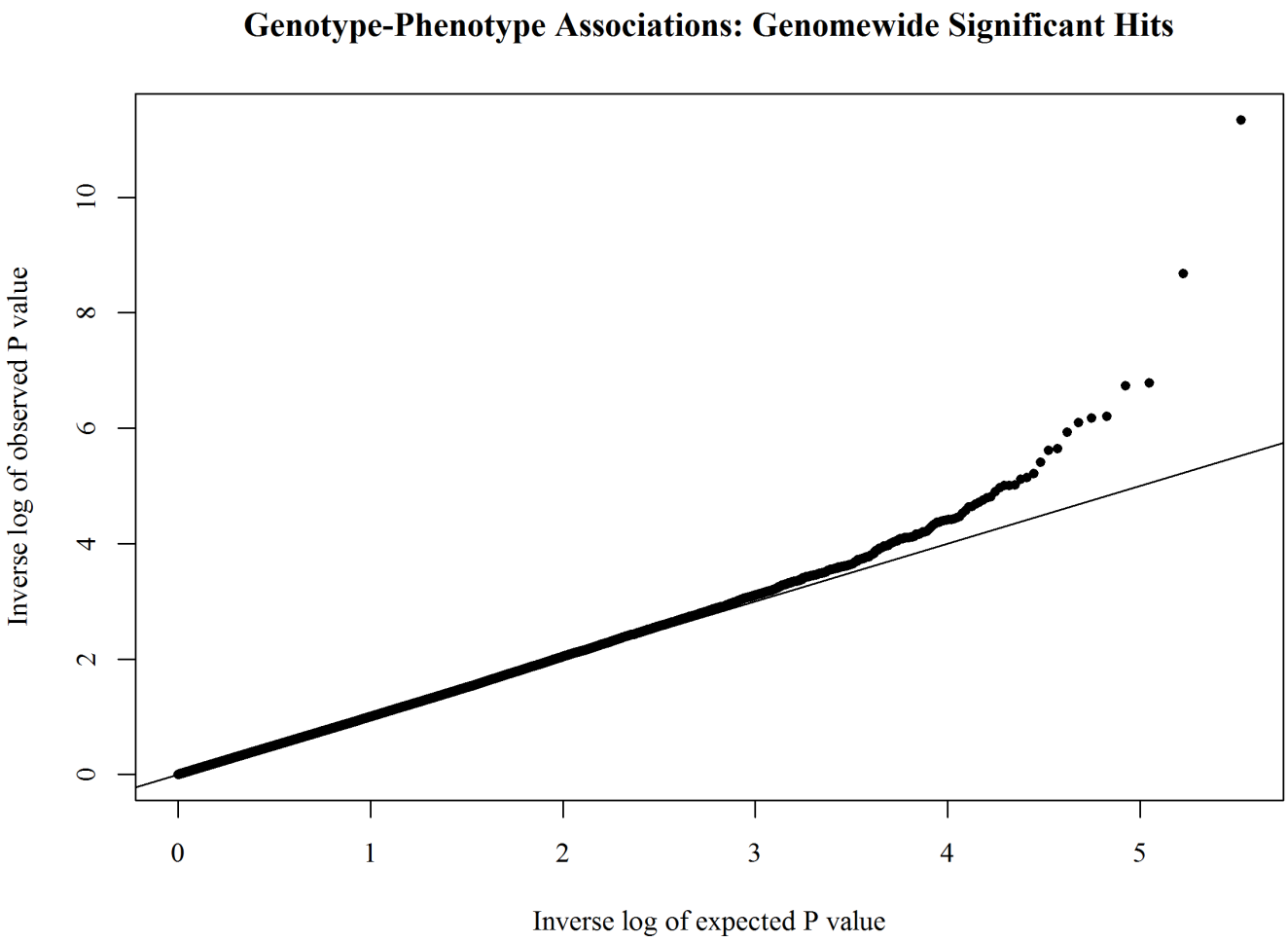
*177 phenotypes include the 121 used in primary analysis enriched for 56 biomedical and serological phenotypes used in the UK10K consortium

Supplementary Table 6. Summary of analysis within and between 56 biomedical and serological phenotypes used in the uk10k consortium and genotypes (unpruned and pruned for LD)

Number of SNPs	Associations	Cut off *P* value	Observed N (%)	Expected N (%)	χ^2 (d.f, *P* value)	O:E ratio ((95% CI), *P* value)
WITHIN PHENOTYPES						
N/A	1,540	0.05	905 (58.77)	77 (5.0)	9372.30 (1, $\leq 1x10^{-50}$)	11.75 ((9.42, 14.67), $\leq 1x10^{-50}$)
		0.01	804 (52.21)	15 (1.0)	41909.61 (1, $\leq 1x10^{-50}$)	53.60 ((32.32, 88.89), $\leq 1x10^{-50}$)
		0.0001	664 (43.12)	1 (0.0001)	439854.62 (1, $\leq 1x10^{-50}$)	664.00 ((93.51, 4714.74), $\leq 1x10^{-50}$)
BETWEEN ALL CATALOG SNPS AND PHENOTYPES						
13,720	768,320	0.05	45,224 (5.88)	38,416 (5.0)	1270.00 (1, $\leq 1x10^{-50}$)	1.17 ((1.16, 1.19), $\leq 1x10^{-50}$)
		0.01	11,388 (1.48)	7,683 (1.0)	3639.50 (1, $\leq 1x10^{-50}$)	1.48 ((1.44, 1.53), $\leq 1x10^{-50}$)
		0.0001	1,120 (0.15)	77 (0.01)	14129.33 (1, $\leq 1x10^{-50}$)	14.55 ((11.55, 18.32), $\leq 1x10^{-50}$)
BETWEEN LD PRUNED CATALOG SNPS AND PHENOTYPES						
9,684	542,304	0.05	29,776 (5.49)	27,115 (5.0)	274.89 (1, $\leq 1x10^{-50}$)	1.10 ((1.08, 1.12), $\leq 1x10^{-50}$)
		0.01	6,603 (1.22)	5,423 (1.0)	259.32 (1, $\leq 1x10^{-50}$)	1.22 ((1.17, 1.26), $\leq 1x10^{-50}$)
		0.0001	189 (0.03)	54 (0.01)	337.53 (1, $\leq 1x10^{-50}$)	3.50 ((2.59, 4.74), $4.44x10^{-16}$)

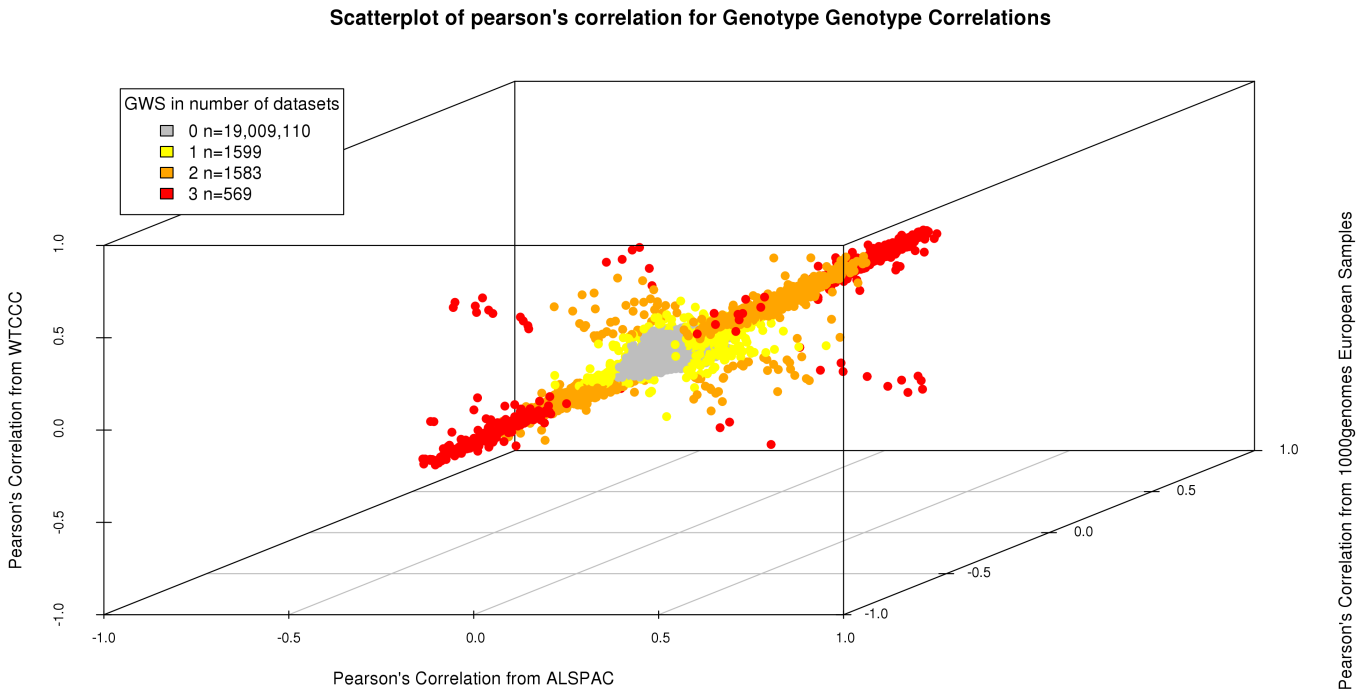
*177 phenotypes include the 121 used in primary analysis enriched for 56 biomedical and serological phenotypes used in the UK10K consortium

Supplementary Figure 1. QQ plot of genotypic and phenotypic associations (pruned dataset, genome-wide significant SNPs only).



QQ plot of observed associations between 6,175 NHGRI-EBI-GWAS catalog genomewide significant SNPs pruned for LD and 121 phenotypes included in analysis against expected P values. Inverse log of P values is the -log10 P value.

Supplementary Figure 2. Scatterplot of genotype-genotype Pearson's correlations from ALSPAC, WTCCC and 1000genomes European samples.



The colors indicate if the Genotype-Genotype correlation was Bonferroni significant in the samples, with grey being not significantly associated in any of the samples, yellow is significantly associated in one sample, orange is significantly associated in two samples and red is significantly associated in all three samples