Article

Evidence for Recent Polygenic Selection on Educational Attainment Inferred from GWAS Hits

Davide Piffer

Department of Psychology, Ben Gurion University of the Negev, Beer Sheva, Israel; Email: piffer@post.bgu.ac.il

Abstract: The genetic variants identified by three large genome-wide association studies (GWAS) of educational attainment were used to test a polygenic selection model. Average frequencies of alleles with positive effect (polygenic scores or PS) were compared across populations (N=26) using data from 1000 Genomes. The PS of 161 GWAS significant SNPs in a recent meta-analysis was highly correlated to population IQ (r=0.863) and to the polygenic score of four alleles independently associated with general cognitive ability. High correlations with PISA scores for a subsample were observed. SNP p value predicted correlation to population IQ and factors from the two previous GWAS (r= -.25). Factor analysis produced similar estimates of selection pressure for educational attainment across the three datasets. Polygenic and factor scores computed using the top 20 significant SNPs showed very high correlation to population IQ (r=0.88; 0.9). Similar findings were obtained using 52 populations from another database (ALFRED). The results together constitute a replication of preliminary findings and provide strong evidence for recent diversifying polygenic selection on educational attainment and underlying cognitive ability.

Keywords: educational attainment; polygenes; polygenic selection; IQ; GWAS

1.   Introduction

 Over the last decade, population geneticists have recognized that most traits are highly polygenic, and hence have moved away from the study of genetic evolution using the single-gene, Mendelian approach, towards models that examine many genes together (i.e. polygenic models). The origin of this trend can be traced back to Fisher (1918). Novel methods for identifying polygenic evolution were recently proposed by Piffer (2013), Berg and Coop (2014), Field et al. (2016), and Leinonen et al. (2013).

The aim of this study is to replicate findings by Piffer (2013, 2015) suggesting the action of polygenic selection on single-nucleotide polymorphisms (SNPs) that had been associated with educational attainment and cognition in genome-wide association studies (GWAS).

To date, GWA studies have identified several alleles with replicated association with intelligence and proxy phenotypes, such as educational attainment. However, each allele typically accounts for a tiny portion of the variance in IQ scores (about 0.5 IQ points per allele). Thus, more recent studies rely on larger and larger samples to identify the greatest possible number of significant associations. Piffer (2013) employed the genetic variants associated with educational attainment and general cognitive ability, showing population differences in allele frequencies that closely matched IQ and educational outcomes differences between ethnic groups and nations. Since then, two big GWAS carried out on independent samples have been published. Thus, a goal of this study is to test the results obtained by Piffer (2013, 2015) against the genetic variants found by the latest GWAS of educational attainment. The important question is whether the new GWAS findings support the results obtained by Piffer (2015) using the first GWAS of this kind (Rietveld et al., 2013) and subsequent studies on general cognitive ability (Davies et al., 2015; Ibrahim-Verbaas et al.,, 2015). In

other words, the question is: are the population frequencies of the latest GWAS hits correlated to the frequencies of the original GWAS hits employed by Piffer (2013, 2015)? Can the claim (Piffer, 2013) that they predict population IQ be substantiated?

To this end, the hits from the two most recent and largest GWAS of educational attainment, measured as completed years of education (Davies et al., 2016; Okbay et al., 2016) will be used in the analysis. The first GWAS was carried out using the UK Biobank sample (N>100,000). Over a thousand SNPs reached genome-wide significance (p< 5 x 10⁻⁸), but after controlling for linkage disequilibrium. Genotypes were pruned for linkage disequilibrium (LD) using clumping to obtain SNPs in linkage disequilibrium with an r^2<0.25 within a 200 kbp window. A few independent signals were identified (Davies et al., 2016). For the sake of simplicity, the three hits found by Rietveld et al. (2013) were lumped together with this polygenic score.

The latest GWAS was carried out on a sample of >293,000 individuals (Okbay et al., 2016) and produced 74 independent ("LD-free") hits. A replication of this study found that the polygenic score predicted 9% of the individual variance in educational achievement at age 16 (Selzam et al., 2016).

Population differences in allele frequencies are normally correlated to phenotypic differences. If there is a correlation between the population frequencies of unlinked alleles favouring a phenotype, this suggests diversifying selection on the phenotype. Recently, Robinson et al. (2015) used a similar approach and showed a correlation between average height of populations and the frequencies of height increasing alleles.

IQ will be used as the phenotype of interest and main dependent variable in the analyses. This choice can be justified by its privileged status in psychometric research and its robust genetic correlation (r= around 0.7) with educational performance (Krapohl et al., 2014) and attainment (Bulik-Sullivan et al., 2015). Moreover, the GWAS hits identified by the three educational GWAS also predict general cognitive ability in their samples (Rietveld et al., 2013; Okbay et al., 2016). A re-analysis of the Okbay et al. dataset revealed that the polygenic score also predicts general intelligence (3.6%) compared to 2% for the 2013 polygenic score (Selzam et al., 2016). Population-level polygenic scores (the average population frequency of alleles with a positive GWAS Beta) will be calculated to test the prediction that they explain variance in population IQ.

Factor analysis will be used to extract a factor accounting for cross-population variation in allele frequency, hence representing a signal of polygenic selection (Piffer, 2015). Factor loadings will be examined to ascertain the reliability of the factor (i.e. do most alleles with positive GWAS effect load positively on the factor?).

Predictive validity will be measured by computing the correlation between factor scores and population IQ. If alleles with positive GWAS beta (i.e. alleles with a positive effect on a trait within population/between individuals) load positively on a factor that is positively correlated to population IQ, this is interpreted as evidence of directional selection on the phenotype (educational attainment or related cognitive abilities).

The robustness of the findings will be tested against a null model using simulations with random SNPs. Moreover, the issue of phylogenetic autocorrelation will be addressed: genome-wide genetic distances (Fst index) representing a neutral evolutionary model will be tested against a polygenic selection model. Autocorrelation is considered problematic because it violates the assumption of independent observations, hence leads to false positive (Type I) errors. Demonstrating that the alleles predict country-level differences in cognitive ability above and beyond that predicted on the basis of migration, drift etc, can be taken as evidence for the hypothesis that these differences have been caused by diversifying polygenic selection operating on these alleles.

If scores computed from GWAS hits predict cross-population phenotypic differences independently of genome-wide genetic distances, then the result is more likely to be indicative of polygenic selection.

2.   Methods

1000 Genomes

Frequencies were calculated from VCF files belonging to the phase 3 data: ftp://ftp-trace.ncbi.nih.gov/1000genomes/ftp/release/20130502/
Rietveld et al. (2013) produced 3 SNPs reaching GWAS significance for educational attainment.
Davies et al. (2016) reported 1115 SNPs reaching GWAS significance, of which 15 were independent signals for educational attainment. 942 SNPs were found on 1000 Genomes. Among the 15 independent signals, one (2:48696432_G_A) was missing.
Okbay et al. (2016) reported 74 SNPs associated with years of education. 70 were found in 1000 Genomes (the other 4 variants were flagged because they had more than 3 different alleles).
Population IQs for the 1000 Genomes populations were obtained from Piffer (2015). Country-level average IQs were obtained from Lynn and Vanhanen (2012). The Finnish and Vietnamese IQs were both adjusted upwards to 101 (from 97 in Lynn & Vanhanen, 2012), to account for recent, more accurate estimates of the Finish population (Armstrong, Woodley, & Lynn, 2014) and a recent reanalysis of the Vietnamese data (Rindermann, Hoang, & Baumeister, 2013). IQ for Tuscany was calculatedas the average between the IQs estimated from PISA Creative
Problem Solving (Piffer & Lynn, 2014) and from PISA Math, Science, Reading. There were three populations (Chinese Dai, Gujarati Indian and Indian Telegu) for which no estimates of average IQ were available.These populations were therefore excluded from the analyses.
Polygenic score refers to the average frequency of alleles with positive effect at the individual level (i.e. GWAS beta).
PISA 2012 results for countries matching 14 of the 26 populations of 1000 Genomes were obtained (OECD, 2012; OECD, 2014). The average score of performance in Mathematics, Reading and Science was used. For the US, scores were broken down by race (Hispanics, African-American and White) and data for Puerto Rico were obtained from a separate report (OECD, 2014).

Statistical analyses were carried out using R (v. 3.2.3). Factor analysis used Ordinary Least Squares (OLS) to find the minimum residual solution (fa function in the R "psych" package) and factor scores were computed using the regression method.
Fst distances (Weir & Cockerham, 1984) for Chromosome 21 and 1 (largest and smallest) will be utilized. These were calculated using VCFtools on the 1000 Genomes, phase 3 files. VCFtools is a program for working with Variant Call Format (VCF) files, like those generated by the 1000 Genomes Project (Danecek et al., 2011).

3.1 Results (1000 Genomes).

3.1.1 Polygenic scores

Rietveld et al., 2013

A polygenic score was created using the top 3 SNPs in Rietveld et al. (2013).

Davies et al., 2016

Davies et al. (2016) reported 1115 SNPs reaching GWAS significance, of which 15 were independent signals for educational attainment. 942 SNPs were found on 1000 Genomes. Among the 15 independent signals, one (2:48696432_G_A) was missing. Thus, a polygenic score (I.S. PS) was calculated using 14 SNPs.

Okbay et al., 2016

Okbay et al. (2016) reported 74 loci independently associated with educational attainment (years of education).

Polygenic scores and population IQs are reported in table 1.

Table 1. Polygenic scores and population IQ.

Population	Rietveld et al. 2013	PS_Ed.Att. Davies	PS Ed. Att. Okbay et al. 2016	IQ SNPs	IQ	PISA 2012
Afr.Car.Barbados	0.106	0.419	0.508	0.4248	83	
US Blacks	0.129	0.447	0.517	0.4624	85	434
Bengali Bangladesh	0.227	0.516	0.507	0.6198	81	
Chinese Dai	0.418	0.610	0.547	0.7162		
Utah Whites	0.374	0.493	0.506	0.6222	99	518
Chinese, Bejing	0.434	0.671	0.563	0.772	105	613
Chinese, South	0.414	0.648	0.555	0.7456	105	561
Colombian	0.252	0.500	0.509	0.6416	83.5	376
Esan, Nigeria	0.096	0.416	0.507	0.4048	71	
Finland	0.387	0.560	0.523	0.6526	101	519
British, GB	0.397	0.526	0.512	0.6296	100	494
Gujarati Indian, Tx	0.311	0.498	0.508	0.6302		
Gambian	0.085	0.438	0.507	0.3868	62	
Iberian, Spain	0.366	0.512	0.519	0.6066	97	484
Indian Telegu, UK	0.229	0.510	0.502	0.6048		
Japan	0.417	0.652	0.556	0.6982	105	536
Vietnam	0.461	0.618	0.552	0.7314	99.4	511
Luhya, Kenya	0.079	0.425	0.502	0.369	74	
Mende, Sierra Leone	0.127	0.416	0.509	0.401	64	
Mexican in L.A.	0.237	0.499	0.505	0.6392	88	413
Peruvian, Lima	0.196	0.477	0.488	0.68	85	368
Punjabi, Pakistan	0.257	0.511	0.513	0.5926	84	
Puerto Rican	0.279	0.489	0.503	0.604	83.5	382
Sri Lankan, UK	0.222	0.506	0.501	0.5754	79	

Toscani, Italy	0.354	0.501	0.518	0.642	99	485
Yoruba, Nigeria	0.097	0.421	0.512	0.3824	71	

The polygenic scores have strong intercorrelations (r= 0.68-0.88) and are also strongly correlated to population IQ (r= 0.65-0.92) and PISA 2012 scores (r= 0.80-0.83). The correlation between PISA 2012 and population IQ is r= 0.949.

3.1.2. Factor analysis

The 17 hits from the two earlier GWAS (Rietveld et al., 2013 and Davies et al., 2016) were lumped together and one hit from Davies et al., 2016 (rs1906252) was removed because it was in LD with rs9320913 from Rietveld et al., 2013. This yielded a set of 16 LD-free SNPs.
A factor analysis (function "fa", package "psych") was carried out using Ordinary Least Squares to find the minimum residual solution. The proportion of variance explained was 0.54. This factor was correlated to population IQ (r= 0.89). 14/16 alleles loaded positively and the average loading was 0.494.

A factor analysis was carried out for 7 sets of 10 SNPs belonging to the 74 Okbay et al. (2016) independent hits (4 were missing). The number 10 was chosen for two reasons: 1) To follow the recommendation that the subject to item ratio be >2:1; 2) Because 70 (the total number of SNPs) is a multiple of 10.
These were sorted by p value, with the first group having the lowest p value (i.e. highest GWAS significance).

3.1.3 Spatial Autocorrelation (SAC)

Spatial (phylogenetic) correlation was calculated using the procedure illustrated in a previous paper (Piffer, 2015), which was based (then unknown to the author) on the Mantel test (Mantel, 1967). Regression analysis applied to the Mantel test enables estimation of polygenic selection pressures (Piffer, 2015).
A total of 325 pairwise comparisons were obtained for the 26 populations from the 1000 Genomes database utilized in the present study. Distances were calculated as the absolute value of the difference between population pairs on the selected variables (country IQ, GWAS hit metagene and polygenic score) (Piffer, 2015).Two matrices with N unique pairwise comparisons are generated, where N = n*(n − 1)/2. One matrix represents genetic distance, and the other matrix represents average population IQ. The country IQ variable had missing values, so a total of 253 distances was calculated. Fst distances were used as a measure of population differentiation due to genetic structure, which is based on the partitioning of genetic diversity within and between populations. The vast majority of random SNPs all over the genome are believed not to be associated with specific phenotypes and therefore not to have been subject to selection. They rather represent population structure resulting from random genetic drift, migration, admixture, and similar processes. Fst is calculated as the ratio of genetic variance in allele frequencies between populations to the sum of variance between gametes within individuals and within populations (Holsinger & Weir, 2009; Weir & Cockerham, 1984).
Regression of phenotypic distances on factor distances coupled with genome-wide Fst distances is carried out. If factor distances have an independent positive effect on the dependent variable (phenotypic distances), then the result is more likely to be indicative of polygenic selection.
The SAC-free effect size (standardized beta) of polygenic and factor scores on population IQ are reported in tables 2 and 3.

Table 2 . SAC control for polygenic scores: Betas.

Source	Fst	PS
P.S. Davies et al. 2016. B=	0.385	0.294
P.S. Davies et al. 2016 + Rietveld et al. 2013. B=	0.329	0.361
P.S. Okbay et al. 2016	0.540	0.154

 Table 3. SAC control for factor scores: Betas. Factor scores extracted from Okbay et al. 2016 GWAS. 7 sets of 10 SNPs sorted by p value and factor score extracted from Rietveld et al. (2013) and Davies et al. (2016).

Source	Fst	Factor
Fac_Rietveld_Davies. B=	-0.162	0.861
Fac_1. B=	0.516	0.122
Fac_2. B=	0.650	-0.076
Fac_3. B=	0.598	-0.011
Fac_4. B=	0.622	-0.090
Fac_5. B=	0.699	-0.138
Fac_6. B=	0.557	0.095
Fac_7. B=	0.428	0.204

3.1.4 MCV

The method of correlated vectors (MCV) originated in psychometric research to assess the validity of items in psychometric batteries (Jensen & Weng, 1994). It consists of a meta-correlation, that is computing the correlation between two sets of correlation coefficients. Applied to GWAS results, the correlation between allele frequencies and the average phenotypes of different populations is computed; then, the resulting correlation coefficients are correlated with the corresponding alleles' genome-wide significance (Piffer, 2016).
MCV was applied to the 70 SNPs from Okbay et al. (2016): the vector of the correlation of each SNP's GWAS p value was correlated to the vector of the correlation between each SNP's frequency and population IQ (r x IQ) and the vector of the correlation with the factor extracted from the two previous GWAS. Negative correlations of the p value were found with r x IQ and with r x Rietveld_Davies factor (r= -0.26; -0.25). Smaller, non-significant correlations (although in the right direction) were found for effect size (0.079 and 0.087 for population IQ and R_D factor).

Four indicators of factor reliability were devised: 1) Average factor loading (mean loading of the 10 SNPs on the factor calculated from all 70 SNPs); 2) correlation to the factor scores obtained from Rietveld et al. (2013) and Davies et al. (2016); 3) Correlation with population IQ; 4) SAC-free Beta
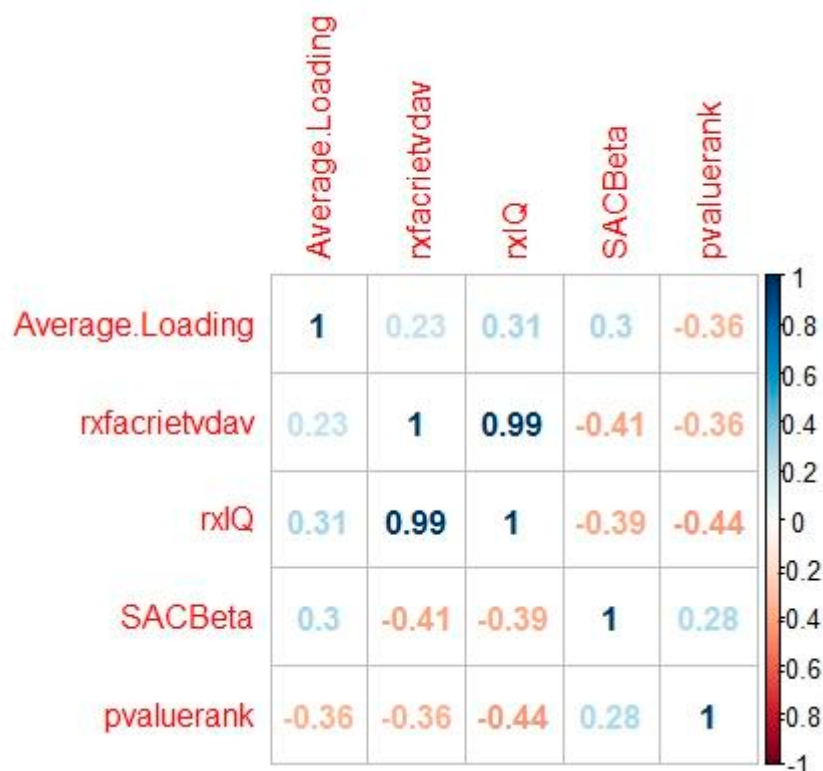
6

("SAC Beta" for short). The values of these indicators are reported in table 4 for each of the 7 SNPs sets, along with their p value rank.

Table 4. Validity of factor scores from 7 subsets of the Okbay et al. (2016) SNPs, ranked by their p values

SNP set	Average Loading	Fac. r x Fac Rietv_Dav	r x IQ	SAC Beta	P value rank
Set 1	0.39	0.608	0.698	0.122	1
Set 2	0.221	0.896	0.715	-0.076	2
Set 3	0.051	-0.847	-0.720	-0.011	3
Set 4	0.152	0.199	0.094	-0.090	4
Set 5	0.199	0.684	0.643	-0.138	5
Set 6	0.046	0.560	0.394	0.096	6
Set 7	0.269	-0.813	-0.782	0.204	7
Mean	**0.190**	**0.184**	**0.149**	**0.015**	

Table 5 reports the intercorrelations between the accuracy measures and p value rank.

Table 5. Intercorrelations between the accuracy measures and p value rank

A novel measure of factor accuracy ("meta-accuracy") was calculated as the mean between the four indicators. In turn, the Spearman-rank correlation between the meta-accuracy vector and p value rank was computed. A negative correlation was found: r= -0.408.

With the aim of validating the meta-accuracy measure, a meta-factor was created by factor analyzing the scores of the 7 factors. The factor loadings ("meta-loadings") were in turn correlated

to the meta-accuracy vector, thus producing a "meta-Jensen coefficient". The correlation between the two meta-vectors was r= 0.969.

Figure 1a. Regression of population IQ on factor extracted from the Okbay et al. (2016) dataset.
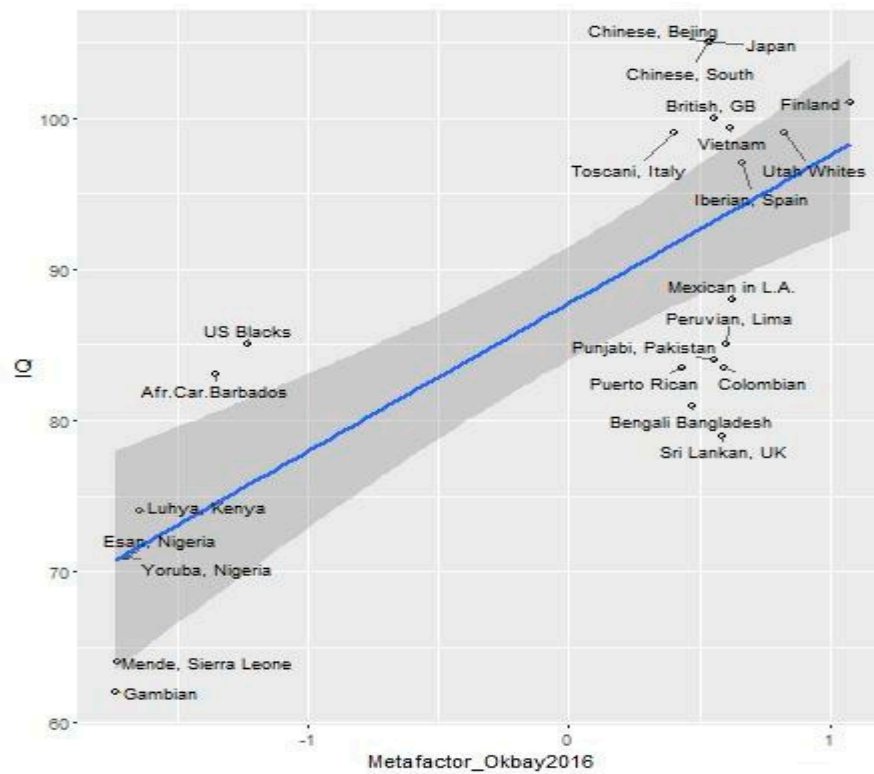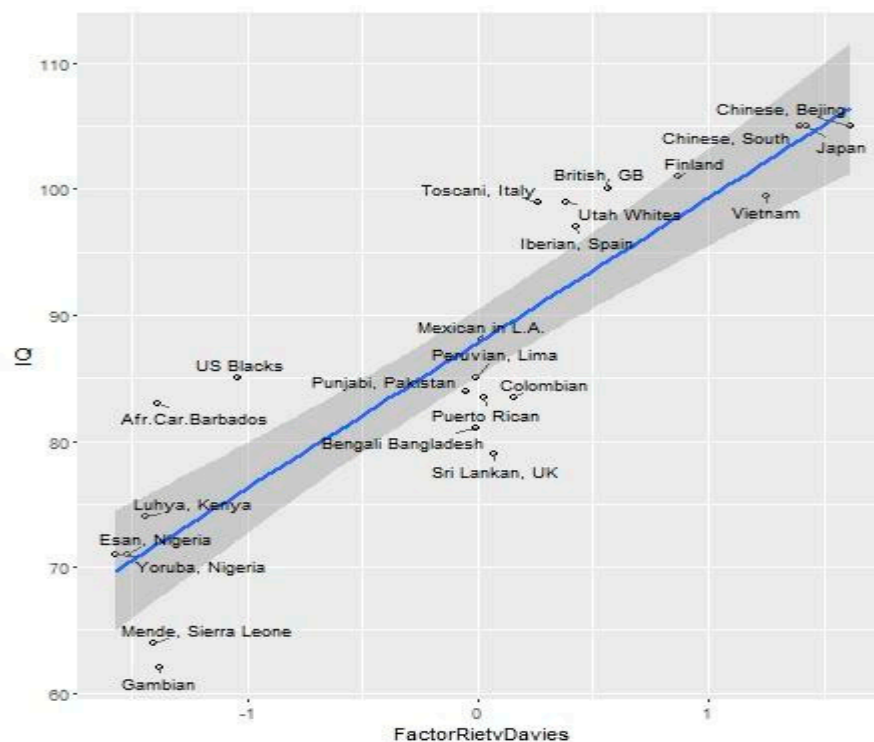

Figure 1b. Regression of population IQ on factor extracted from the Rietveld et al. (2013) & Davies et al. (2016) datasets.

There were substantial intercorrelations between the meta-factor, the Rietveld+Davies factor scores and IQ. SAC-control was applied to the meta-factor. This produced a very weak SAC-free effect (B= 0.097; Fst B= 0.508).

To extract a reliable estimate of polygenic selection, the average of the two factors ("average factor") was computed. The correlation between the average factor and population IQ was r= 0.858.

3.1.5 LD pruning

Cross-GWAS linkage was checked by feeding SNPSNAP with the list of 86 SNPs, with LD thresholds of 500kb and r= 0.5.

In total, 8 SNP pairs were found to be in LD. One SNP was present in two GWAS datasets (rs9320913).

A list of replicated or pseudo-replicated (in LD across studies) SNPs was created, composed of one of the two linked SNPs (one for each pair) and the 8 SNPs in LD across GWAS (table 13). The polygenic scores from the linked SNPs are reported in table 6. The correlation between the two scores is r= 0.919.

Table 6. Pseudo-replicated and replicated SNPs. Sites in LD (r>0.5).

Publication	Index SNP	Publication	Linked SNP
Davies et al., 2016	rs12042107	rs1008078	Okbay et al., 2016
Rietveld et al., 2013	rs11584700	rs11588857	Okbay et al., 2016
Rietveld et al., 2013	rs4851266	rs12987662	Okbay et al., 2016
Davies et al., 2016	rs13086611	rs148734725	Okbay et al., 2016
Davies et al., 2016	rs11130222	rs11712056	Okbay et al., 2016
Davies et al., 2016	rs55686445	rs62263923	Okbay et al., 2016
Davies et al., 2016	rs12553324	rs13294439	Okbay et al., 2016
Davies et al., 2016	rs4799950	rs12969294	Okbay et al., 2016
Rietveld et al., 2013	rs9320913*	rs9320913	Okbay et al., 2016

*Replicated

Frequencies of the replicated hits were also calculated for the 5 super-populations (i.e. races) of 1000 Genomes for both SNP sets. A boxplot is shown in figures 2a and 2b.

Figure 2a. PS of linked/replicated SNPs by race. SNPs from Rietveld et al. (2013) and Davies et al. (2016).
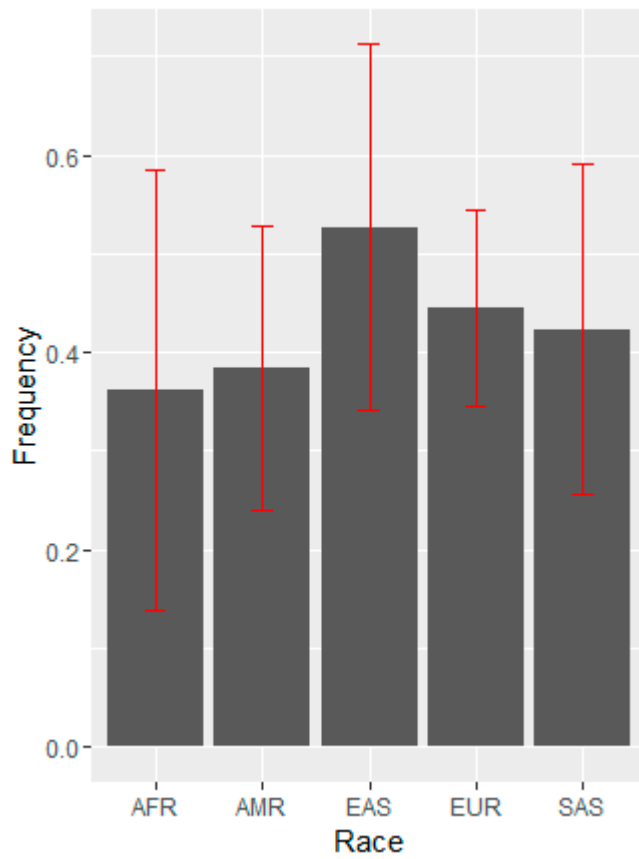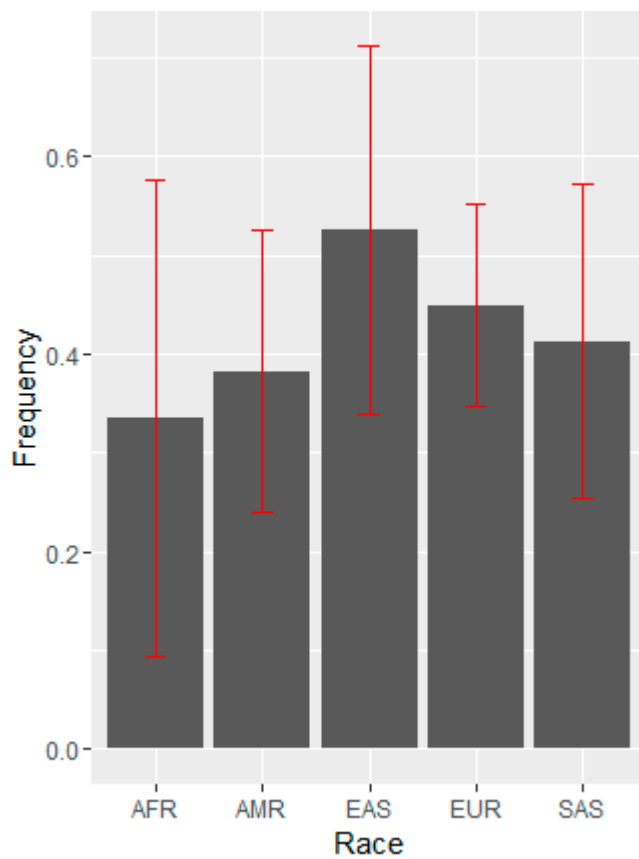
Figure 2b. PS of linked/replicated SNPs by race. SNPs from Okbay et al. (2016).

The replicated SNPs were factor analyzed. The two factors were almost identical (r= 0.998).
Finally, a list of cross-GWAS clumped SNPs was created by keeping only one SNP for each LD pair (e.g. rs12042107 (Davies) - rs1008078 (Okbay). Only the latter (rs1008078) was preserved). Obviously, the replicated SNP (rs9320913) was counted only once.This resulted in a list of "LD-clumped" (86-8-1)= 77 SNPs.
LD-clumped polygenic score (LD clumping across independent hits from three GWAS. Pre-clumping N=86; Post clumping and overlap: N=77).

The LD-clumped PS had the following correlations with the other variables: r x IQ: 0.766; r x FactorRietvDavies: 0.835; r x Metafactor: 0.475.

The population IQ variable had some missing cases so the correlations are reported both with and without IQ (table 7a and 7b, respectively).

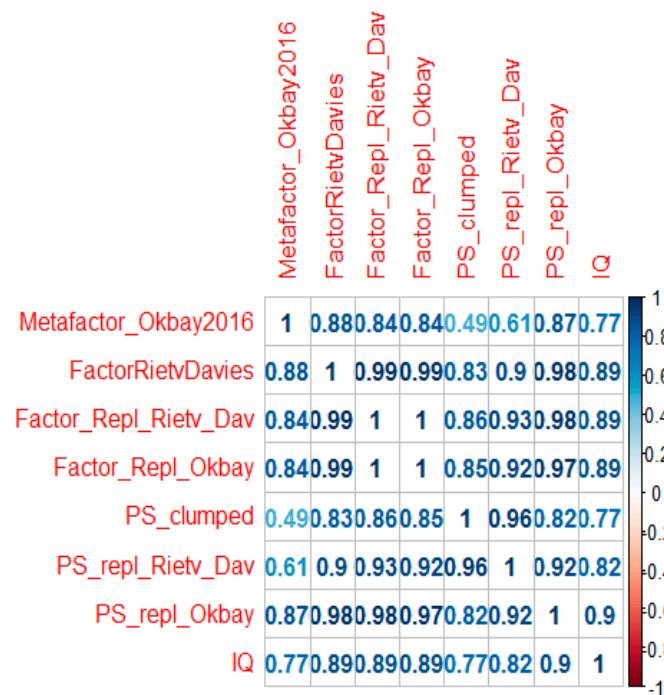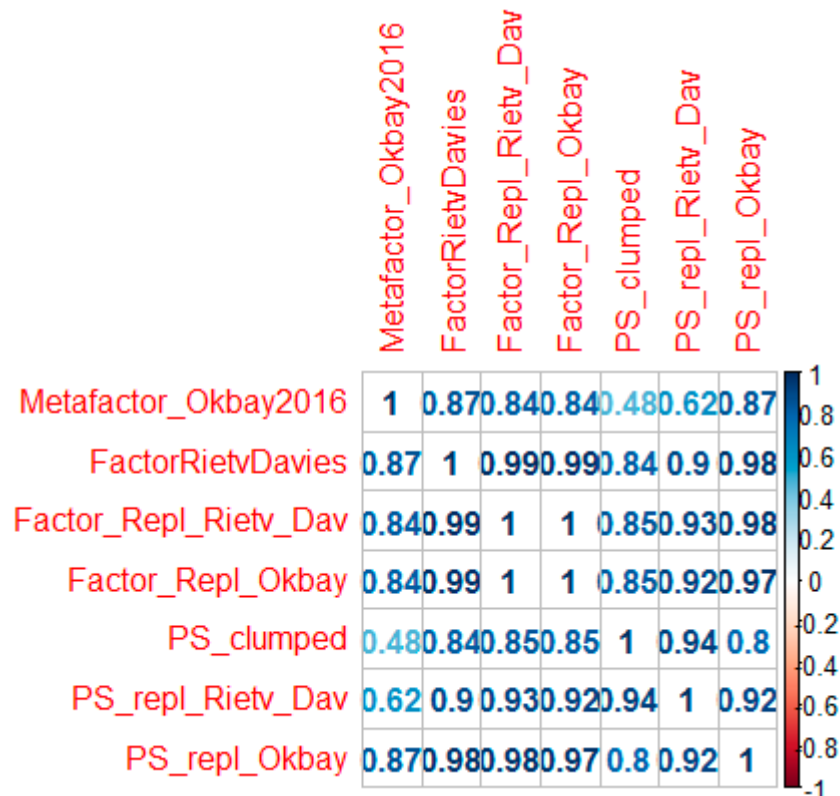**Table 7a. Correlation plot (all polygenic and factor scores). With IQ, N = 26 countries.**

	Metafactor_Okbay2016	FactorRietvDavies	Factor_Repl_Rietv_Dav	Factor_Repl_Okbay	PS_clumped	PS_repl_Rietv_Dav	PS_repl_Okbay	IQ
Metafactor_Okbay2016	1	0.88	0.84	0.84	0.49	0.61	0.87	0.77
FactorRietvDavies	0.88	1	0.99	0.99	0.83	0.9	0.98	0.89
Factor_Repl_Rietv_Dav	0.84	0.99	1	1	0.86	0.93	0.98	0.89
Factor_Repl_Okbay	0.84	0.99	1	1	0.85	0.92	0.97	0.89
PS_clumped	0.49	0.83	0.86	0.85	1	0.96	0.82	0.77
PS_repl_Rietv_Dav	0.61	0.9	0.93	0.92	0.96	1	0.92	0.82
PS_repl_Okbay	0.87	0.98	0.98	0.97	0.82	0.92	1	0.9
IQ	0.77	0.89	0.89	0.89	0.77	0.82	0.9	1

Table 7b.Correlation plot (all polygenic and factor scores). Without IQ, N =23 countries.

Okbay et al. (2016) reported 162 independent SNPs that reached genome-wide significance (P < $5*10^{-8}$) in the pooled-sex EduYears meta-analysis of the discovery and replication samples (N =405,072) and 7 SNPs that replicated across the discovery and replication sample at p <$5*10^{-8}$).
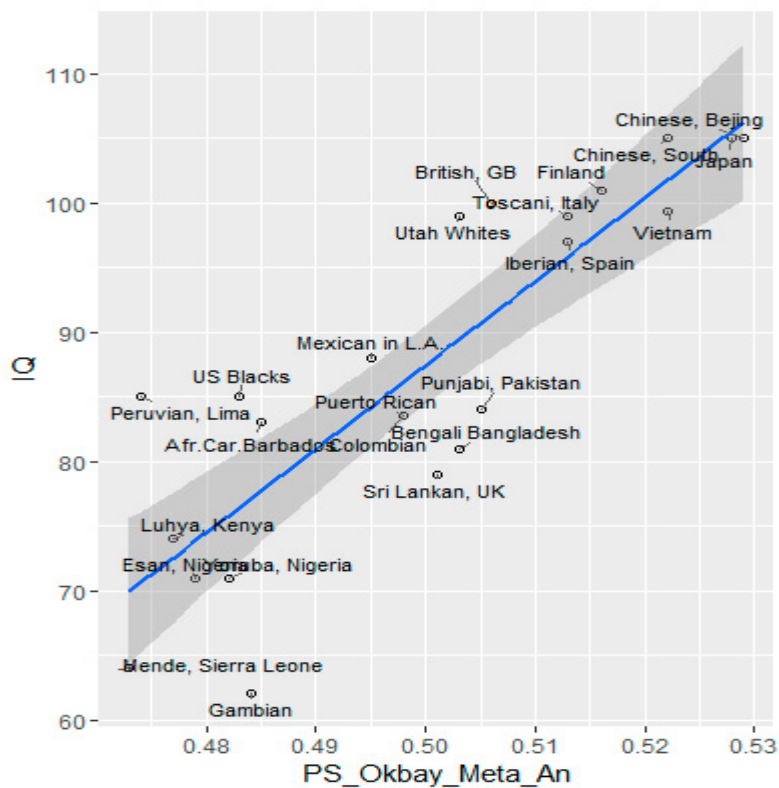
161 SNPs were found in 1000 Genomes. The polygenic score was computed (table 8). Its correlation to population IQ was r= 0.856 (scatterplot figure 3) and PISA 2012 (r=0.851). The polygenic score of the 7 hits that replicated with a genome-wide significant p value had a correlation r= 0.833 with population IQ and PISA 2012 (r= 0.681).

Table 8. Polygenic scores from the meta-analysis. Pooled polygenic score with 161 SNPs (PS_161) and 7 SNPs replicating across samples (PS_7).

Population	PS_161	PS_7
Afr.Car.Barbados	0.485	0.256
US Blacks	0.483	0.286
Bengali Bangladesh	0.503	0.360
Chinese Dai	0.516	0.361
Utah Whites	0.503	0.383
Chinese, Bejing	0.529	0.399
Chinese, South	0.522	0.387
Colombian	0.498	0.356
Esan, Nigeria	0.479	0.263
Finland	0.516	0.449

British, GB	0.506	0.390
Gujarati Indian, Tx	0.506	0.363
Gambian	0.484	0.243
Iberian, Spain	0.513	0.389
Indian Telegu, UK	0.504	0.350
Japan	0.528	0.359
Vietnam	0.522	0.381
Luhya, Kenya	0.477	0.251
Mende, Sierra Leone	0.473	0.263
Mexican in L.A.	0.495	0.329
Peruvian, Lima	0.474	0.298
Punjabi, Pakistan	0.505	0.364
Puerto Rican	0.498	0.339
Sri Lankan, UK	0.501	0.364
Toscani, Italy	0.513	0.389
Yoruba, Nigeria	0.482	0.267

Figure 3. Relationship between P.S. computed from hits by Okbay et al. (2016)'s pooled meta-analysis and population IQ.

Mean scores for 4 populations representing four super-populations/races were calculated and are reported in figure 4.

13

Figure 4. Boxplot of polygenic scores for four populations. BEB= Bengali in Bangladesh; GBR= British in Great Britain; CHS= Chinese South; YRI= Yoruba in Nigeria.

Method of correlated vectors

For the 161 significant SNPs in the meta-analysis, the correlations of p value and GWAS effect size with the IQ correlation (meta-correlation) were r= -0.009 and -0.048, respectively.

Spatial autocorrelation (SAC): clumped and replicated SNPs

Spatial autocorrelation analysis was run on the three scores. The effect size (Standardized beta) is reported in table 9.

Table 9. SAC control for polygenic and factor scores.

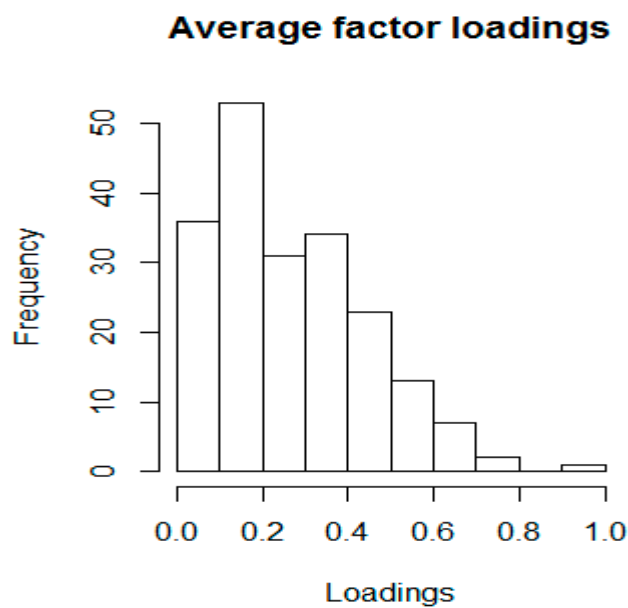| Source	Fst	Factor/PS
PS clumped. B=	0.476	0.250
PS replicated_Rietv_Davies. B=	0.395	0.352
PS replicated_Okbay. B=	-0.062	0.791
PS meta-analysis_Okbay. B=	0.273	0.458
Factor pseudo-replicated. B=	0.002	0.695

3.1.6 Simulation

Factor loadings

14

In order to show that factors formed from educational attainment SNPs are more coherent (i.e. have higher loadings) than factor formed from random SNPs, a simulation was carried out.

100 sets of 10 SNPs matched to the top significant SNPs in Okbay et al. (2016) were obtained from SNPsnap (http://www.broadinstitute.org/mpg/snpsnap/). After removal of problematic SNPs (when frequency was 0 for a population, that population was not counted, creating mismatch between rows, hence these had to be removed). Among those, the first 200 sets (to speed up computation) of 10 random SNPs were chosen for a simulation. Factor analysis was iterated over each set. The average factor loading was 0.268 (SD=0.176).

The distribution was non-normal, with a strong positive skew (figure 5).

Figure 5. Distribution of mean of factor loadings (N=200).

This information was used as a baseline, null model to test against polygenic selection.

Factor scores

The correlations between the factor scores for the 200 sets of 10 random SNPs and population IQ were computed. The average Pearson's r was 0.22 (95% CI= -0.757; 0.823; 99% CI= -0.826; 0.886). Thus, the correlations between the factors (pseudoreplicated hits) and IQ (r=0.89) is significant according to the conventional p value (0.05).

Polygenic scores

The set of 7914 random SNPs was divided into 52 sets of 150 SNPs. N=150 was chosen because it is close to the number of SNPs in the pooled Okbay et al. (2016) sample.

The average correlation between the polygenic scores for 52 sets of 150 SNPs and IQ was 0.467 (95% CI= -0.100; 0.817). The upper limit of the 95 % CI was almost identical to that obtained for the factor scores simulation (0.823). Hence, the correlation between the 152 SNPs GWAS hits polygenic score and population IQ (r= 0.863) is significant according to the conventional p value (0.05).

Meta-accuracy

The 161 SNPs were divided into 32 sets of 5. Factor analysis was carried out on each set. The average factor loadings for each set and the correlation of the factor scores to population IQ are reported in table 20. The average factor loading was 0.225, and the average corr x IQ was 0.112.

The correlations of factor loadings and corr x pop IQ with p value were r= -0.273 and -0.008, respectively. Moreover, the two vectors (factor loadings and corr x pop IQ) were intercorrelated (r= 0.223).

The top 4 significant sets were chosen because their loadings, p value and correlation with population IQ suggested the presence of stronger signal (average corr x population IQ= 0.83; average loading= 0.383). The correlation between population IQ with the average factor score and the polygenic score were r=0.923 and 0.867.

3.1.7 Controlling for baseline population differences in derived allele frequencies

At a theoretical level, an ancestral allele is the allele that was carried by the last common ancestor between humans and other primates whereas an allele is derived when it arose in the human lineage after the split from other primates. In practice, this allele is usually ascertained via comparison with chimpanzees. One limitation of this procedure is that if a mutation arose in chimpanzees after the split from humans, then the ancestral allele is not the chimp allele. Thus, 1000 Genomes infers ancestral alleles via alignment with 6 primate species (Ensembl, 2015).
Allele status could be ascertained for 156 out of 162 trait-increasing alleles in the Okbay et al. (2016) pooled meta-analysis. Of those, 86 and 70 were derived (55.13%) and ancestral, respectively. Since this is not an equal representation, it creates a potential confounding factor. Average derived and ancestral allele frequencies for positive effect alleles (DAF and AAF, respectively) were computed. These indeed confirmed previous findings that non-African populations have higher frequencies of derived alleles (Henn et al., 2015). A new polygenic score was computed by averaging the two scores. This gives equal weights to derived and ancestral alleles, hence eliminating the potential bias due to different baseline levels of derived alleles among populations. The correlation between this "corrected" score and population IQ was r= 0.828. The derived and ancestral polygenic scores had opposite correlations to population IQ (r= 0.848 and - 0.516, respectively).

3.1.8 Do "educational" SNPs population frequencies reflect selection pressures on general intelligence?

Piffer (2015) analysed 9 GWAS hits of educational attainment and IQ. Five of these were related to general cognitive function (rs10457441 C; rs17522122 G; rs10119 G; rs236330 C; rs17518584 T), as reported in two recent studies (Davies et al, 2015; Ibrahim-Verbaas, 2015) . The polygenic score is reported in table 1 (col. 5). Its correlation to population IQ is r= 0.847. Its correlations to the four educational attainment polygenic scores (in the order reported in the table) are: 0.896; 0.822; 0.655;0.856.

3.2. ALFRED

3.2.1 Polygenic scores

The analysis was replicated using ALFRED (lhttps://alfred.med.yale.edu/), which contains more populations than 1000 Genomes (52 after pairwise deletion of missing values) but has smaller samples and much lower coverage. Thus, only a minority of the SNPs available in 1000 Genomes are present in ALFRED. In order to overcome this limitation, two strategies were used: 1) For the SNPs reported in Rietveld et al. (2013) and Davies et al. (2016), the full list of significant hits was searched, comprising also the non-independent hits. Then, the hits in LD (r>0.5) were removed. 2) Okbay's 162 independent hits (pooled meta-analysis) were searched for those in LD (r>0.9) using SNPSnap. In total, 12 SNPs from Rietveld + Davies and 72 SNPs from Okbay et al. (2016) were found in ALFRED.

Polygenic scores were computed for both datasets and are reported in table 10. No valid IQ estimates are available for those populations, so this is not reported.

The correlation between the two polygenic scores was r=0.781 (Figure 6).

Seven racial groups were identified and average polygenic scores computed for them (table 11).

Table 10. Polygenic scores calculated using ALFRED. SNPs published by Okbay et al. (2016) + Rietveld et al. (2013) and Davies et al. (2016). N=72 and 16, respectively.
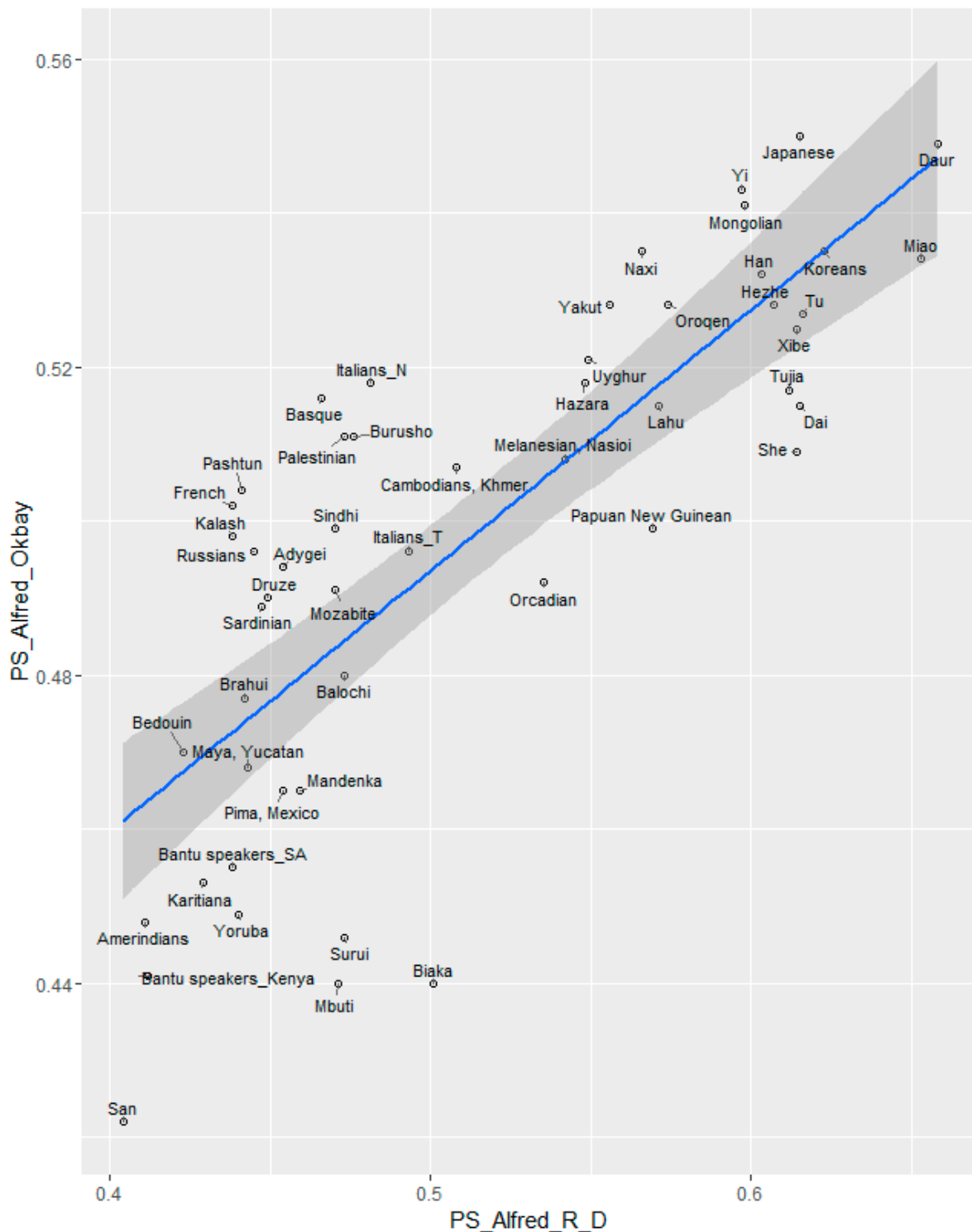
Population	PS_Alfred_Okbay	PS_Alfred_R_D
Bantu speakers_SA	0.455	0.438
Bantu speakers_Kenya	0.441	0.412
Biaka	0.440	0.501
Mandenka	0.465	0.459
Mbuti	0.440	0.471
Mozabite	0.491	0.470
San	0.422	0.404
Yoruba	0.449	0.440
Balochi	0.48	0.473
Bedouin	0.470	0.423
Brahui	0.477	0.442
Burusho	0.511	0.476
Druze	0.490	0.449
Hazara	0.518	0.548
Kalash	0.498	0.438
Mongolian	0.541	0.598
Oroqen	0.528	0.574
Palestinian	0.511	0.473
Pashtun	0.504	0.441
Sindhi	0.499	0.470
Cambodians, Khmer	0.507	0.508
Dai	0.515	0.615
Daur	0.549	0.658
Han	0.532	0.603
Hezhe	0.528	0.607
Japanese	0.550	0.615
Koreans	0.535	0.623
Lahu	0.515	0.571
Miao	0.534	0.653
Naxi	0.535	0.566

She	0.509	0.614
Tu	0.527	0.616
Tujia	0.517	0.612
Uyghur	0.521	0.549
Xibe	0.525	0.614
Yi	0.543	0.597
Adygei	0.494	0.454
Basque	0.516	0.466
French	0.502	0.438
Italians_Tosc	0.496	0.493
Italians_North	0.518	0.481
Orcadian	0.492	0.535
Russians	0.496	0.445
Sardinian	0.489	0.447
Maya, Yucatan	0.468	0.443
Pima, Mexico	0.465	0.454
Melanesian, Nasioi	0.508	0.542
Papuan New Guinean	0.499	0.569
Yakut	0.528	0.556
Amerindians	0.448	0.411
Karitiana	0.453	0.429
Surui	0.446	0.473

Table 11. Polygenic scores for racial groups.

Racial group	Polygenic Score	Ethnic group
African	0.446	Bantu, Biaka,Mandenka,Mbuti,San,Yoruba
Middle Eastern	0.469	Palestinian, Bedouin, Druze
Central Asian	0.484	Burusho, Kalash,Pashtun,Balochi.Brahui,Hazara, Sindhi
East Asian	0.568	Dai,Daur,Han,Hezhe,Japanese,Koreans,Lahu,Miao,Nazi,She,Tu,Tujia
European	0.490	Basque, French, Italian, Orcadian, Russian
Native American	0.449	Maya,Pima,Amerindians,Karitiana,Surui
Papuan	0.529	Papuan New Guinea, Melanesian Nasioi

Figure 6. Scatterplot showing the correlation between the polygenic scores obtained from three GWAS publications. ALFRED populations. r= 0.781

3.2.2 Spatial Autocorrelation and distance from Africa.

Lawson-Handley et al. (2007) calculated geographic distance from Addis Ababa (AA), regarded as the centre of the human expansion, assuming an origin of modern humans in Eastern Africa.
Geographic distances for the ALFRED populations were obtained from an analysis of genetic variation using the HGDP-CEPH data, which showed that 77% of the genetic variance (Fst) could
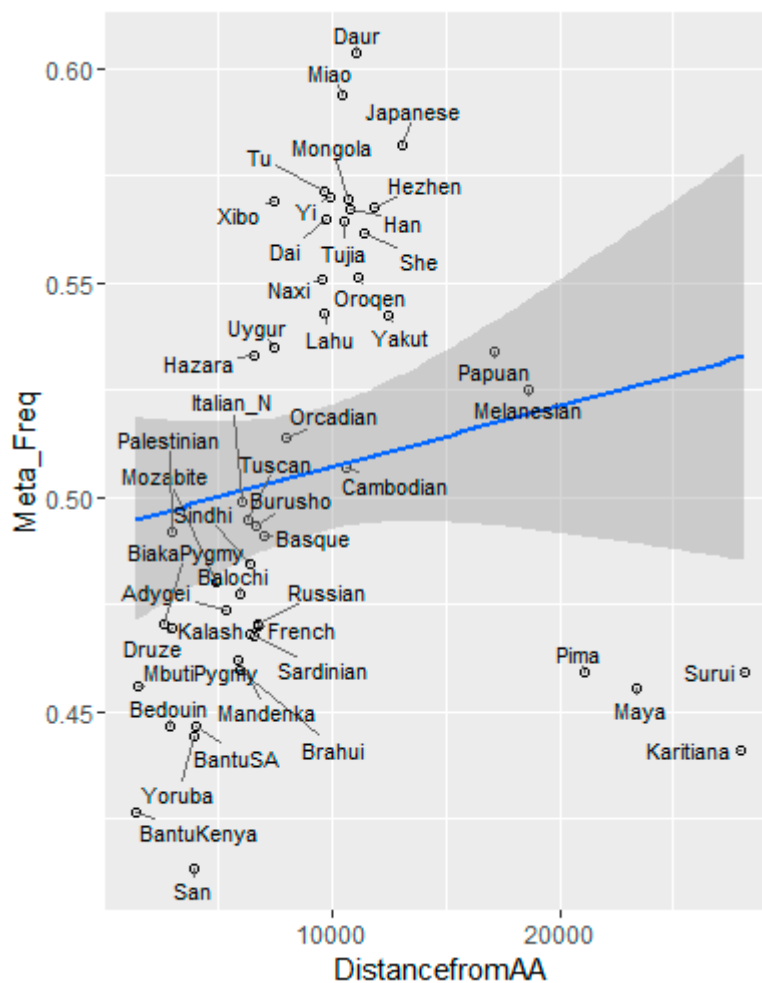
be explained by geographic distance between populations (Lawson Handley, Manica, Goudet and Balloux, 2007). There were 54 and 52 populations in the HGDP and ALFRED datasets, respectively. Of these, 49 were overlapping and were used for calculation of PS distances (N= 1177). Correlations are reported in table 12.

Table 12. Measures of spatial autocorrelation (correlation coefficient between pairwise distances). N= 1176.

	PS Okbay	PS Rietveld + Davies
r x Fst	0.46	0.161
r x Geo	0.377	0.164
r x Distance from AA	0.065	0.207

The two polygenic scores were averaged. The correlation between the average PS and distance from AA is r=0.171, indicating little effect of neutral population history on polygenic scores (Fig. 7).

Figure 7. Scatterplot showing correlation between geographic distance from Addis Ababa and averaged polygenic scores.

4. Discussion

The two largest GWAS of educational attainment, published in 2016, appear to confirm the population-level findings based on the first large GWAS of this kind (Rietveld et al., 2013). Specifically, higher frequencies of alleles with a positive effect on educational attainment are observed among East Asians than in Europeans, and the latter have higher frequencies than Africans. Moreover, the polygenic scores obtained from two independent samples are strongly correlated (table 2) to those obtained from the first GWAS (r= 0.68-0.88).

The frequencies of 4 SNPs associated with general cognitive ability in previous studies (reviewed by Piffer, 2015) are strongly correlated to the educational attainment polygenic scores.

The genetic variants identified by three large genome-wide association studies (GWAS) of educational attainment were used to test a polygenic selection model.

Strong inter-correlations among population-level polygenic scores of alleles found by three independent GWAS to be associated with educational attainment were observed. Moreover, these polygenic scores were substantially correlated to estimates of average population IQ (table 2).

The polygenic score of 161 SNPs that reached genome-wide significance in the meta-analysis by Okbay et al. (2016) of the discovery and replication samples (N =405,072) was highly correlated to population IQ (r=0.854). A similar correlation was produced when using only the 7 SNPs that reached significance in both datasets separately (r= 0.833).

Using the hits by Okbay et al. (2016), the method of correlated vectors revealed the presence of a "Jensen effect" (meta-correlation) of SNP p value on population IQ and factor from the two previous GWAS (r= -.25). That is, frequencies of alleles with lower p value (more GWAS significance and higher likelihood of being true positives) had stronger correlations to population IQ and a factor extracted from an independent dataset. Factor analysis produced similar estimates of polygenic selection strength for educational attainment across the three datasets. The SNPs from the largest GWAS (Okbay et al., 2016) were subset by p value (N= 7) and factor analyzed. Variables indicating reliability (factor loadings), predictive (r x population IQ and SAC-free Beta) and convergent validity (r x factor extracted from previous GWAS) were computed. The correlation with p value rank was in the expected direction in three out of four (except p value rank x SAC Beta) instance and of low magnitude. A composite index was created by calculating the average of the four variables. This index substantially correlated to SNP set's P value rank (r= -0.4). There was also a positive correlation between average factor loading and corr. x IQ, corr. x factor, corr. x SAC-free effect.

This result also highlights a potential problem, that is the high noise present in the data. Some of the SNPs with very small effects on educational attainment may have been subject to such minimal selection that chance rather than selection is the main reason for their present allele frequencies. In other cases, very small effects on education may indicate that their main phenotypic effect is somewhere else. There could be small effects on the risk of psychiatric and other diseases, for example, that were more important than the marginal effects on learning ability or interest in learning.

Factor analysis was carried out on the 7 factors belonging to the 7 sets. The loadings on this "meta-factor" were strongly correlated to the composite index of factor accuracy (r=0.96).

That is, the factors' independent correlations to measures of accuracy were in turn correlated to their loadings on the "meta-factor". The factor extracted from the hits of the first two GWAS (Rietveld et al., 2013; Davies et al,, 2016) survived control for phylogenetic correlation using Fst distances (test of the hypothesis that random factors such as drift or migration confound the results, see Piffer, 2015) quite well (B= 0.861), but among the Okbay et al. 2016 hits, only the top 10 significant ones had some residual predictive validity (B= 0.122). This discrepancy was also reflected in the average factor loadings, which were much higher in the former than in the latter (0.49 vs. 0.19).

21

There were substantial intercorrelations (r= 0.77-0.89) between the meta-factor, the Rietveld+Davies factor scores and IQ (table 7a).

LD clumping was performed on the entire SNPs set from the three GWAS (N=86). This produced 77 LD-clumped SNPs, 1 SNP in common between two publications and 8 linked SNPs. The polygenic score calculated from the LD-clumped SNPs was correlated to population IQ (r=0.77) and to the factor (r= 0.85) and polygenic scores extracted from the 9 replicated/linked SNPs (r= 0.94 and 0.8) (tables 7a and 7b). The two PS obtained from the cross-publication linked/replicated hits, were strongly correlated to population IQ (r=0.82; 0.9). The two factors of the replicated/linked hits were similarly correlated to population IQ (r=0.89).

All four vectors survived control for spatial autocorrelation by partialling out of Fst distances (table 9), reaching Betas= 0.25; 0.35; 0.79; 0.69). The PS obtained from the 161 SNPs of the Okbay et al. (2016) meta-analysis had "SAC-free" B=0.500. Another PS was created using 7 hits that reached genome-wide significance of $p<5 \times 10^{-8}$ in Okbay et al. 2016's study using the UK Biobank replication sample and this had a similar correlation to population IQ (r= 0.83).

The correlation between polygenic, factor scores and population IQ exceeded the confidence interval obtained via simulations with hundreds of random SNPs. Factor analysis also produced more coherent factors than those extracted from random sets of SNPs. The average factor loadings for random SNPs was 0.268, but it was generally higher for the replicated loci or the hits from Rietveld et al. (2013) and Davies et al. (2016). However, the SNPs from Okbay et al. (2016) did not produce loadings higher than chance expectations.

Both controlling for spatial autocorrelation (SAC) via regression analysis and simulations with random SNPs revealed that the high correlation coefficients between polygenic scores and IQ are inflated due to population history mechanisms (drift, migrations, etc.) unrelated to natural selection. That is, both random SNPs and Fst distances had substantial correlations to population IQ and to polygenic scores, although not as strong as the correlation between the latter two. Thus, the reported effects are overestimates and the magnitude of the real effect is probably closer to the "SAC-free" estimate ranging from 0.25-0.79 obtained with polygenic scores (table 9).It should also be pointed out that it is most likely not the SNPs or the polygenes per se that explain the group differences in IQ. These are probably only indicators, picking up signal of differential selection strength, supposedly acting on many other genes, scattered across the genome, whose association with IQ is not known yet. A prediction is thus that as future GWAS will produce more hits, they will fit the pattern found by this study. This study is important because it confirms the prediction of an earlier one (Piffer, 2015) - which used fewer SNPs - that cognition-enhancing alleles identified by future GWAS would show similar population frequencies. The fact that the addition in the present study of over 150 SNPs to the original analysis carried out by Piffer using only 9 SNPs produced very similar polygenic scores is remarkable and suggests that the selection signal is robust and relatively independent of the variance explained by the polygenic score.

However, since the SNPs themselves only contribute to a small % of phenotypic IQ, it is still possible that discovery of additional SNPs could substantially change the PS-group IQ correlations.

Factor analysis carried out on 32 sub-sets (N=5) of the 161 SNPs from Okbay et al. (2016) confirmed a tendency for lower p value (more significant) SNPs to have higher loadings (r= -0.23) and a stronger correlation of their factor scores to population IQ (r= -0.22). Moreover, the two vectors (factor loadings and corr x pop IQ) were intercorrelated (r= 0.27). Polygenic and factor scores computed using the top 20 significant SNPs showed very high correlation to population IQ (r=0.88; 0.9).

Controlling for differing frequencies of derived alleles across populations due to drift or GWAS ascertainment bias and the overrepresentation of derived alleles among alleles with positive effect did not significantly alter the polygenic score nor its relationship to population IQ. A new polygenic score was computed by averaging the two scores. This gives equal weights to derived and ancestral alleles, hence eliminating the potential bias due to different baseline levels of derived alleles among populations. The correlation between this "corrected" score and population IQ was r= 0.851.

However, there is a possibility that giving equal weight to derived and ancestral alleles leads to overcorrection if intelligence-enhancing alleles tend to be derived. This is likely because the increase in intelligence during evolution of the *Homo* lineage has been favoured by mutations that arose after the split from the last common ancestor with chimpanzees.

The major difference was an attenuation of the African-Eurasean differences due to derived alleles being in general at lower frequencies among Africans, and derived alleles being also overrepresented among those with positive GWAS effect. Previous findings that non-African populations have higher frequencies of derived alleles were confirmed within this sample of SNPs (Henn et al., 2015).

The replicated/linked SNPs exhibit a clear clustering of alleles across 5 major racial groups (figure 2), showing a pattern matching cross-racial differences in IQ: East Asians>Europeans>South Asians> Hispanics>Africans.

A replication of the analysis using 52 populations from ALFRED revealed a similar pattern (figure 6).

Surprisingly, distance from the putative ancestral environment in Eastern Africa (Addis Abeba) was not correlated to polygenic score (table 12), suggesting that selection rather than drift due to geographic distance or isolation is responsible for the observed pattern (figure 7). This is reflected also in the lower polygenic scores obtained by Native Americans, despite their genetic similarity to East Asians. If genuine, this finding would suggest that selection pressures on higher IQ for East Asians date after the split from Americans ca. 13Kya and are thus relatively recent. Lower population density in America could account for relaxed selection pressure, possibly mediated by lower social or sexual competition.

Another potential issue raised by one reviewer is that "the extrapolation to non-European populations is still problematic because the accuracy of the polygenic score declines in such populations as a result of differing LD patterns (Scutari et al., 2015)".

In fact differences in LD should simply reduce the frequency differences at the tag SNPs between populations, compared to the real causal SNPs. This is due to a phenomenon called "attenuation". Indeed, correction for attenuation is used "to rid a correlation coefficient from the weakening effect of measurement error (Jensen, 1998). This scenario works in the case that the frequency differences between tag and causal SNPs are due to random error, so that the mean frequency of the cognitive ability alleles is equal to the (genome-wide) background frequency (which for a mathematical reasons, is 50%). If a non-European population's tag SNPs (GWAS hits) are at frequencies lower than 50%, LD decay implies that the causal SNPs are at even lower frequencies because attenuation will push frequencies up towards the baseline frequency of 50%.Conversely, if tag SNPs are at frequencies >0.5, then the frequency of causal SNPs will be underestimated by the tag SNPs. Hence, controlling for LD decay would actually increase frequency differences when non-European populations exhibit frequencies that markedly differ from 50%. For example, the average frequencies of the replicated sites are all markedly lower than 50% (CEU= 38%; Chinese: 39%, Nigerian: 26%), suggesting the causal SNPs are at even lower frequencies, and this effect will be stronger for populations whose frequencies are farther away from the mean of 50%. Future studies should examine which of the SNPs predict educational attainment in non-European populations, because the linkage phase between tag SNP and causal SNP will in some cases be different in different populations. In some cases, the causal SNP may even be monomorphic in some parts of the world, especially if it arose recently. Initially, this means that comparisons between populations should be limited to those SNPs that reliably predict educational attainment or related traits in all populations. In addition, different linkage phase in different populations is not only a nuisance factor that limits the reproducibility and usefulness of reported associations. It is also a great tool for trying to find the causal polymorphism among clusters of SNPs that may have very high linkage disequilibrium in any one population or haplotype.

Another potential problem is that of population stratification in the original GWAS. Population stratification refers to the "allele frequency differences between cases and controls due to systematic ancestry differences" (Liu et al., 2013). Obviously controlling for population stratification is

important because if the putatively causal SNPs were actually ancestry markers, the cross-population differences in allele frequencies would be due to population stratification signal picked by the GWAS, instead of genetic variants which are causally related to the phenotype. Fortunately, the GWAS publications dealt with the issue of population stratification, and identified stratification effects that were small in magnitude, about 8% of the signal due to population stratification and not to polygenic signal (Okbay et al., 2016). Okbay et al. (2016) also employed within-family analysis, which is a more direct approach to isolate polygenic association signal from population stratification but is not powered for individual SNP association testing.

Overall, this study provides strong evidence that there has been recent polygenic and diversifying selection on educational attainment, hence producing different levels of cognitive capacity and other traits related to educational attainment among populations. The fuzzy nature of the results is expected given the highly polygenic nature of the trait, the small effect sizes associated with each SNPs and the likely presence of many false positives. Pleiotropy, random genetic drift, different linkage equilibrium patterns and differential distribution of ancestral and derived alleles are also expected to increase the noise in the data. Given all these factors, it is remarkable that a relatively clear pattern emerged, especially when focusing on the associations produced by the largest sample or the replicated SNPs and loci.

A limitation of this study is the reliance on GWAS hits for a complex phenotype such as educational attainment, which shares the majority of additive genetic variation with general intelligence, but also other personality and health-related traits (Krapohl et al., 2014 and 2015).

It is possible that there are other SNPs affecting cognitive variables unrelated to educational attainment and these may not necessarily be subject to the same selection pressures. Future studies should carry out "deep phenotyping" with SNPs from each chromosomal region that has produced hits in the educational attainment studies, to find out which of them are related to IQ (and its components, such as long-term memory, working memory, and abstract reasoning), and which are related to non-cognitive traits such as ambition, self-control, preference for abstract thinking, or incompetence in practical pursuits. There are reports indicating that school success is as much related to preferences and personality as to intelligence, and that these non-cognitive traits, if properly measured, are as heritable as IQ (Kovas et al., 2015).

Another limitation is the reliance on estimates of population IQ as the phenotypic variable, which are not perfectly accurate, besides reflecting environmental and economic differences between populations.

Appendix

Raw data and code can be downloaded from: osf.io/z5nju

References

Armstrong, E. L., Woodley, M. A., & Lynn, R. (2014). Cognitive abilities amongst the Sàmi population. Intelligence, 46, 35–39. http://dx.doi.org/10.1016/j.intell.2014.03.009.

Berg, J. J., & Coop, G. (2014). A population genetic signal of polygenic adaptation. PLoS Genetics, 10, e1004412.

Danecek, P., Auton, A., Abecasis, G., Albers, C. A., Banks, E., DePristo, M. A., et al. (2011). The variant call format and VCF tools. *Bioinformatics*, 27, 2156–2158

Davies, G., Armstrong, N., Bis, J. C., Bressler, J., Chourake, V., Giddaluru, S., et al. (2015). Genetic contributions to variation in general cognitive function: a meta-analysis of

genome-wide association studies in the CHARGE consortium (N = 53949). Molecular Psychiatry, 20, 183–192. http://dx.doi.org/10.1038/mp.2014.188.

Davies, G., Marioni, R.E., Liewald, D.C., et al. (2016). Genome-wide association study of cognitive functions and educational attainment in UK Biobank (N =112 151). *Molecular Psychiatry*, 1-10. doi:10.1038/mp.2016.45

Ensembl,2015:   http://www.1000genomes.org/faq/where-does-ancestral-allele-information-your-variants-come

Field,Y., Boyle, E.A., Telis, N.,Gao, Z., Gaulton, K.J., Golan, D., Yengo, L, Rocheleau, G.,Froguel, P., McCarthy, P., Pritchard, J.K. Detection of human adaptation during the past 2,000 years. bioRxiv doi: http://dx.doi.org/10.1101/052084

Fisher, Ronald A. (1918). The Correlation Between Relatives on the Supposition of Mendelian Inheritance". Philosophical Transactions of the Royal Society of Edinburgh. 52, 399–433. doi:10.1017/s0080456800012163

Jensen, A. R., and Weng, L.J. (1994). What is a good g? *Intelligence*, 18, 231-258.

Jensen, A.R. (1998). *The* g *Factor: The Science of Mental Ability*. Praeger, Connecticut, USA.

Henn BM, Botigué LR, Peischl S, *et al.*: Distance from sub-Saharan Africa predicts mutational load in diverse human genomes. *Proc Natl Acad Sci U S A.* 2015; pii: 201510805

Holsinger, K. E., & Weir, B. S. (2009). Genetics in geographically structured populations: defining, estimating and interpreting Fst. *Nature Reviews Genetics*, 639–650

Ibrahim-Verbaas, C. A., Bressler, J., Debette, S., Schuur, M., Smith, A. V., Bis, J. C., et al. (2015). GWAS for executive function and processing speed suggests involvement of the CADM2 gene. Molecular Psychiatryhttp://dx.doi.org/10.1038/mp.2015.37.

Kovas, Y., Garon-Carrier, G., Boivin, M., Petrill, S.A. et al. (2015). Why children differ in motivation to learn: Insights from over 13,000 twins from 6 countries. *Personality and Individual Differences* 80: 51-63

Krapohl, E., Rimfeld, K., Shakeshaft, N.G., Trzaskowski, M., McMillan, A., Pingault, J.-B., Asbury, K., Harlaar, N., Kovas, Y., Dale, P.S. & Plomin, R. (2014).The high heritability of educational achievement reflects many genetically influenced traits, not just intelligence. PNAS, 111, 15273–15278, doi: 10.1073/pnas.1408777111

Krapohl, E., Euesden, J., Zabaneh, D., Pingault, J.B., Rimfeld, K., von Stumm, S., Dale, P.S., Breen, G., O'Reilly, P.F., and Plomin, R. (2015). Phenome-wide analysis of genome-wide polygenic scores. *Molecular Psychiatry*, 1-6. doi:10.1038/mp.2015.126

Lawson Handley, L.J., Manica, A., Goudet, J., Balloux, F. (2007). Going the distance: human population genetics in a clinal world. *TRENDS in Genetics*, 23, 432-439.

Leinonen, T., Scott McCairns, R.J., O'Hara, R.B., Merila, J. (2013). Qst-Fst comparisons: evolutionary and ecological insights from genomic heterogeneity. *Nature Reviews*, 14, 179-189. Doi: 10.1038/nrg3395.

Liu, L., Zhang, D., Liu, H., Arendt, C. (2013). Robust methods for population stratification in genome wide association studies. *BMC Bioinformatics*, 14, doi: 10.1186/1471-2105-14-132

OECD, 2012. PISA 2012 Results in Focus. https://www.oecd.org/pisa/keyfindings/pisa-2012-results-overview.pdf

OECD, 2014. 5 Performance of Puerto Rico 15-Year- Old Students in Mathematics, Science, and Reading Literacy in an International Context. http://abrepr.org/sites/default/files/PISA%20Draft.pdf

Okbay, A., Beauchamp, J.P., Fontana, M.A., Lee, J., Pers, T.H., et al. (2016). Genome-wide association study identifies 74 loci associated with educational attainment. *Nature*, doi:10.1038/nature17671

Piffer, D. (2013). Factor Analysis of Population Allele Frequencies as a Simple, Novel Method of Detecting Signals of Recent Polygenic Selection: The Example of Educational Attainment and IQ. *Mankind Quarterly*, 54, 168-200.

Piffer, D. (2015). A review of intelligence GWAS hits: Their relationship to country IQ and the issue of spatial autocorrelation. *Intelligence*, 53, 43-50.

Piffer, D. (2016). Evidence of polygenic selection on human stature inferred from spatial distribution of allele frequencies. F1000Research, 4:15. http://dx.doi.org/10.12688/f1000research.6002.3

Rietveld, C.A., Medland, S.E., Derringer, J., Yang, J., Esko, T., Martin, N.W., et al. (2013). GWAS of 126,559 individuals identifies genetic variants associated with educational attainment. *Science*, 340, 1467-1471. doi: http://doi.org/10.1126/science.1235488

Robinson, M.R., Hemani, G., Medina-Gomez, C., Mezzavilla, M., Esko, T., …. & Visscher, P.M. (2015). Population genetic differentiation of height and body mass index across Europe. *Nature Genetics*, 47, 1357-1362. doi: 10.1038/ng.3401.

Scutari, M., Mackay, I., & Balding, D. J. Using genetic distance to infer the accuracy of genomic prediction. arXiv:1509.00415.

Selzam, S., Krapohl, E., von Stumm, S., O'Reilly, P.F., Rimfeld, K., Kovas, Y., Dale, P.S., Lee, J.J., and Plomin, R. (2016). Predicting educational achievement from DNA. Molecular Psychiatry, doi: 10.1038/mp.2016.107

Weir,B. S.,&Cockerham, C. C. (1984). Estimating F-statistics for the analysis of population structure. Evolution, 38, 1358–1370.