



Strong negative relationship between population-level general intelligence and ADHD genetic factors inferred from allele frequencies

DAVIDE PIFFER, EMIL KIRKEGAARD

ABSTRACT

ADHD and general intelligence are negatively correlated (within populations) and this correlation is driven by common genetic variants shared between the two phenotypes. This paper analyzes the population frequency patterns of alleles associated with ADHD and intelligence in two samples of 26 and 50 populations (1000 genomes and ALFRED). Factor analysis of allele frequencies was used to estimate the strength of natural selection on the two traits. The two factors, indicating selection for general intelligence and ADHD, show strong negative correlations in both 1000 Genomes and ALFRED samples ($r = -0.93$ and -0.90 , respectively). Alleles with lower p-values would be less likely to be false positives, so the more significant ADHD GWAS hits are expected to be more strongly negatively correlated with the general intelligence SNP and the ADHD SNP factors, which were also found ($r = -0.26$ and 0.37 , respectively).

The ADHD factor predicted national IQs also after accounting for a measure of population structure (F_{st}).

Results are interpreted in a framework based on evolutionary convergent selection pressure for higher general intelligence and lower ADHD.

READ REVIEWS

WRITE A REVIEW

CORRESPONDENCE:
pifferdavide@gmail.com

DATE RECEIVED:
June 10, 2015

KEYWORDS:
intelligence, adhd, polygenic selection

© Piffer et al. This article is distributed under the terms of the [Creative Commons Attribution 4.0 International License](#), which permits unrestricted use, distribution, and redistribution in any medium, provided that the original author and source are

INTRODUCTION

Attention deficit hyperactivity disorder (ADHD) is a psychiatric disorder characterized by having trouble concentrating and impulsive behavior. Prevalence depends on specific criteria used (e.g. self-report vs. clinical diagnosis) but is around 5-7% (Willcutt, 2012). Genetically informative designs find high heritabilities of ADHD. One review of 20 twin studies found a mean heritability of 76% (Faraone et al, 2005). ADHD is related to many other psychiatric problems such as bipolar disorder and drug dependency (Kessler et al, 2006). More importantly, ADHD is negatively related to general intelligence (GI; Kuntsi et al, 2003) and learning ability (Mayes et al, 2000). The correlation with GI is about $-.3$. This may seem small, but in practice it means that the ADHD group has an IQ about 9 points below the non-ADHD group (see Frazier 2004 for a meta-analysis of ADHD x IQ correlations). Moreover, Kuntsi et al. (2003) found that the genetic overlap between ADHD and GI was 100%; the same genes have an influence on both traits.

Martin et al. (2014) found that a genetic composite risk score, based on a case-control genome-wide association study (GWAS) for clinical ADHD (Stergiakouli et al., 2012), was independently associated with lower IQ. However, the association was rather weak ($\beta = -.05$). On the other hand, the discovery sample was quite small for a GWAS (700 with ADHD, 5100 controls), so the association will

credited.



likely be larger in a better powered GWAS.

Molecular genetic studies of height, general intelligence and educational attainment

Over the last few years, researchers have started moving away from the study of genetic evolution using a single-gene, Mendelian approach towards models that examine many genes together (polygenic). The more genes are involved in a given phenotype, the more the signal of natural selection will be “diluted” across different genomic regions (because each gene accounts for a tiny effect) making it difficult to detect it using approaches focused on a single gene (Pritchard et al., 2010; Piffer, 2014; Davies et al, 2011; Groen-Blokhus et al, 2014). A first attempt at empirically identifying polygenic selection was made by Turchin et al. (2012) on two populations (Northern and Southern Europeans), providing evidence for higher frequency of height-increasing alleles (obtained from GWAS studies) among Northern Europeans. A drawback of that study was the reliance on populations from a single continent and that crude pairwise comparisons (e.g. French vs. Italian) were used without correlating frequency differences to average population height. Moreover, the strength of selection was not determined.

Rietveld et al. (2013)’s meta-analysis found ten SNPs that increased educational attainment, comprising three with nominal genome-wide significance and seven with suggestive significance. A recent study has replicated the positive effect of these top three SNPs, rs9320913, rs11584700 and rs4851266, on mathematics and reading performance in an independent sample of school children (Ward et al., 2014).

Two different approaches to identify selection based on the correlation of allele frequencies across different populations have been recently developed by Piffer (2013) and Berg & Coop (2014).

Piffer (2013) obtained two samples comprising 14 and 50 populations (1000 Genomes and ALFRED databases, respectively) and applied principal components analysis to the frequencies of the ten alleles reported in Rietveld et al. (2013). The alleles loaded highly and in the expected direction (positively) on a single factor accounting for most of the variance. The factor scores were correlated to indexes of country educational achievement (PISA) and IQ, producing high correlations (r ’s around 0.9). This factor was interpreted as indicating the strength of polygenic selection. This was the first time that genetic frequencies had been used from a cross-racial sample and an estimate of selection strength was provided, thus correlating it with measured average phenotypic scores.

The genetic correlation between ADHD and GI within a population makes it plausible that the same pattern will be found across populations (Jensen, 1998, “default hypothesis”). Based on the above, we formed the following hypotheses:

1. If the ADHD SNPs are true positives and have been selected for or against, then there should be a general ADHD SNP selection factor across populations.
2. There should be a negative correlation between the ADHD SNP general factor and measures of national GI, as well as with the previously reported GI SNP selection factor.
3. There should be a negative correlation between the p-values of ADHD SNPs and their relationships with the GI/educational attainment SNP genetic factor and the ADHD SNP general factor because SNPs with higher p-values are more likely to be false positives.

METHODS

IQ/educational attainment-increasing alleles: Piffer’s factor, extracted from 4 SNPs affecting general intelligence, was used (Piffer, 2015a). The top three SNPs (not in LD), were obtained from the most recent GWAS of cognitive function (Davies et al., 2015). This study focused on fluid intelligence.

ADHD risk alleles were obtained from the largest GWAS to this date, including 5,621 clinical patients and 13,589 controls (Grohen-Blokuis et al., 2014).

National/Ethnic IQs were obtained from Lynn and Vanhanen (2012). The IQ for Tuscany was obtained from Piffer & Lynn, 2014, as the IQ of Central Italy. There were 3 missing cases (Chinese Dai, Gujarati Indian, Indian Telegu). See suppl. material (“gadhdfactors” spreadsheet).

As previously done, we conducted the analysis on two independent databases of population genetic data: 1000 Genomes (<http://browser.1000genomes.org/index.html>), comprising 26 populations, and ALFRED(<http://alfred.med.yale.edu/>), comprising 50 populations after list-wise deletion of missing cases.

RESULTS (STUDY 1 AND 2)

Since the analyses were generally identical in both databases, we go through the first in detail and mostly summarize the results for the second.

STUDY 1: 1000 Genomes (phase 3)

Factor analysis of GI-increasing alleles including the latest three GWAS hits (Davies et al., 2015).

Factor analysis was on the three new hits and the four SNPs (Piffer, 2015a) minus one SNP (rs9320913) which was in strong LD with one of the three new hits (rs10457441).

Table 1. Factor loadings of 3 “new” SNPs

Chr. #	Gen. region	SNP	Factor loading
6	98678841	rs10457441.C	0.83
14	32372633	rs17522122.G	-0.76
19	50098513	rs10119.G	0.39

Table 2. Factor loadings of 4 SNPs minus overlapping one.

Chr. #	Gen. region	SNP	Factor loading
1	204576983	rs11584700.G	0.83
2	100818479	rs4851266.T	0.91
1	94059554	rs236330.C	0.83

The two factors were strongly correlated with each other ($r=0.79$, $N=26$, $p<0.0001$).

A factor analysis was then carried out on the 6 SNPs together. 5 of 6 had strongly positive loadings, but one SNP (rs17522122) loaded in the wrong direction (-0.55).

Table 3. Factor loadings of 6 GI SNPs.

--	--	--	--

Chr. #	Gen. region	SNP	Factor loading
6	98678841	rs10457441.C	0.69
14	32372633	rs17522122.G	-0.55
19	50098513	rs10119.G	0.76
1	204576983	rs11584700.G	0.82
2	100818479	rs4851266.T	0.92
1	94059554	rs236330.C	0.85

The correlation between the 6 SNPs GI factor and national IQs was $r=0.87$ ($N=23$, $p<0.0001$). Factor scores are reported in Table 4.

Table 4. Factor scores for two GI SNP sets and total set of 6 SNPs.

	3 "new" snps	3 old snps (4 minus 1 in LD)	6 SNPs	IQ	
ACB	-1.20		-1.22	-1.28	83
ASW	-1.10		-1.23	-1.24	85
ESN	-1.21		-1.43	-1.47	71
GWD	-1.03		-1.38	-1.41	62
LWK	-1.71		-1.47	-1.77	74
MSL	-0.38		-1.29	-1.16	64
YRI	-1.00		-1.46	-1.47	71
CLM	-0.08		-0.25	-0.26	83.5
MXL	-0.30		0.05	0.06	88
PEL	-0.86		0.08	0.12	85
PUR	0.73		-0.17	0.04	83.5
CDX	0.94		1.06	1.14	N/A
CHB	0.18		1.31	1.22	105

CHS	0.27	1.24	1.22	105
JPT	0.01	1.22	1.07	105
KHV	0.63	1.45	1.40	99.4
CEU	1.12	0.52	0.51	99
FIN	1.27	0.47	0.56	101
GBR	0.88	0.60	0.55	100
IBS	1.39	0.33	0.46	97
TSI	1.41	0.29	0.52	99
BEB	0.24	-0.09	0.08	81
GIH	0.25	0.56	0.51	N/A
ITU	-0.20	0.31	0.29	N/A
PJL	-0.18	0.34	0.21	84
STU	-0.07	0.14	0.08	79

Factor analysis of ADHD risk alleles (1000 Genomes)

Analysis including linked loci

Since the probability that an SNP is a true positive depends on its p-value, only alleles with a p-value 5×10^{-5} were included. This is already a very liberal threshold, because the conventional threshold to be considered significant is 5×10^{-8} after correction for multiple testing issues (Clarke et al., 2011). However, the SNPs with the lowest p-value in this study were only $< 2 \times 10^{-6}</math>, so we chose as a-priori convention to include all SNPs with a p-value of an order of magnitude lower. This resulted in a set of 42 SNPs. For allele frequencies, see the supplementary material. The set of SNPs (N=42) included many hits within the same genomic region (500Kb). To avoid redundancy, linked loci were not counted more than once and, in order to reduce noise, if two or more SNPs were in linkage, only the one with the best p-value was included. Since having a case-to-variable ratio of at least 2:1 is recommended for factor analysis (Zhao, 2009) and there were a total of 17 SNPs, we created two sets of 9 and 8 SNPs, ordered according to their p-value (i.e. the first set comprising the half with the lower p-values).$

A factor analysis using minimum residuals was carried out and factor scores saved with the Thurstone method. Other extraction methods were not used as they produced nearly identical solutions. The same factor analytic procedure was used throughout this paper.

Most of the alleles' loadings (11/17) were in the right direction (positive), $p=0.166$. These are reported in Tables 5 and 6.

The two factors (obtained from the two separate and unlinked SNP sets) were highly correlated with each other ($r=0.96$) and with the genetic population GI SNP factor (from the 4 alleles) ($r=-0.90$ and -0.89 , respectively).

Table 5. First set of ADHD risk SNPs.

Chr. #	Gen. region	SNP	Factor loading
18	64443239	rs17232800.C	-0.76
5	78146288	rs6453417.T	0.93
19	8675892	rs301407.T	0.45
11	120025839	rs4245040.G	0.74
2	232026998	rs16828074.G	0.89
14	96126480	rs7160641.T	0.95
4	97191397	rs7674977.G	-0.59
20	6624355	rs6038589.C	0.97
8	24338666	rs7012077.	-0.26

Table 6. Second set of ADHD risk SNPs.

Chr. #	Gen. region	SNP	Factor loading
9	133903199	rs10512416.A	0.7
1	203260558	rs2802837.C	1
7	85478161	rs17160554.C	0.52
6	40761488	rs9471400.C	0.67
3	62876746	rs4635724.G	-0.9
17	85123530	rs17712565.G	-0.73
12	33865802	rs12582885.T	0.33
22	36781026	rs8142185.C	-0.21

Factor scores for GI and ADHD alleles are reported in Table 7.

Table 7. Factor scores: GI-enhancing and ADHD risk alleles in 1000 Genomes.

Population	ADHD SNP second order factor	ADHD unlinked first set, SNP factor	ADHD unlinked second set, SNP factor	ADHD Composite	GI 4 SNP factor	GI 6 SNP factor
Afr.Car.Barbados	1.53	1.17	1.43	1.30	-1.26	-1.28
US Blacks	1.21	1.12	1.05	1.09	-1.21	-1.24
Esan Nigeria	1.93	1.69	1.65	1.67	-1.45	-1.47
Gambian	1.47	1.48	1.53	1.50	-1.45	-1.41
Luhya Kenya	1.50	1.47	1.63	1.55	-1.54	-1.77
Mende Sierra Leo	1.71	1.64	1.65	1.65	-1.24	-1.16
Yoruba	1.66	1.66	1.70	1.68	-1.46	-1.47
Colombian	-0.32	-0.78	-0.46	-0.62	-0.12	-0.26
Mexican LA	-0.42	-0.58	-0.57	-0.58	0.02	0.06
Peruvian	-0.46	-0.69	-1.07	-0.88	-0.30	0.12
Puerto Rican	-0.32	-0.41	-0.17	-0.29	0.01	0.04
Chinese Dai	-0.82	-1.05	-0.96	-1.00	1.18	1.14
Han Chin. Beijing	-0.65	-1.06	-0.80	-0.93	1.40	1.22
Han Chin. South	-0.83	-0.80	-0.97	-0.89	1.30	1.22
Japanese	-0.25	-1.32	-0.90	-1.11	1.23	1.07
Vietnam	-0.77	-0.91	-0.83	-0.87	1.60	1.40
UtahWhites	-0.57	-0.75	-0.77	-0.76	0.76	0.51
Finns	-0.63	-0.30	-0.98	-0.64	0.71	0.56
British	-0.66	-0.46	-0.58	-0.52	0.85	0.55
Spanish	-0.55	-0.75	-0.68	-0.71	0.60	0.46
Tuscan (Italy)	-0.66	-0.70	-0.44	-0.57	0.57	0.52
Bengali Bangladesh	-0.68	-0.06	-0.22	-0.14	-0.26	0.08

Gujarati Ind. Tx	-0.51	0.22	-0.24	-0.01	0.47	0.51
Indian Telegu UK	-0.71	0.18	-0.06	0.06	0.02	0.29
Punjabi Pakistan	-0.54	-0.11	0.23	0.06	0.19	0.21
Sri LankanUK	-0.67	0.11	-0.18	-0.03	-0.61	0.08

Given the high correlation between the factors extracted from the two sets of unlinked alleles, a composite factor was obtained as the mean of the two vectors, to make the analysis more parsimonious and less erroneous.

Figures 1(a,b) shows a scatterplot with the factor scores for the ADHD (composite) and GI SNP factors. The correlations between the two variables were $r = -0.91$ and -0.93 for the 4 and 6 SNPs factors, respectively.

Figure 1a. Regression of ADHD SNP factor on the 4 SNPs GI SNP factor.

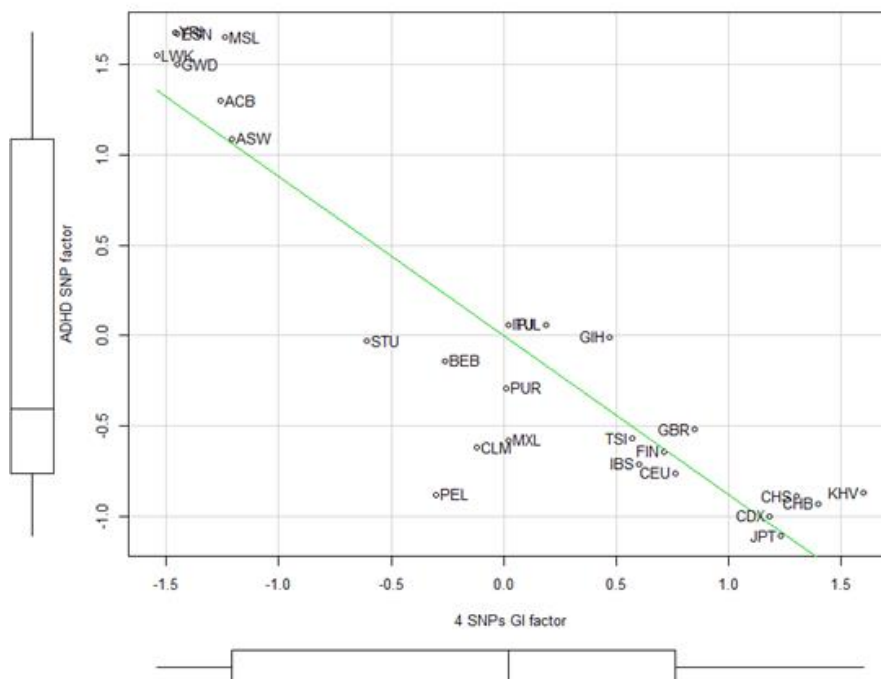


Figure 1b. Regression of ADHD SNP factor on the 6 SNPs GI SNP factor.



Visual inspection of the plot revealed that the African groups were outliers. After removal of the African groups, the correlations between the ADHD SNP composite factor and the 4 or 6 SNPs GI factors were still negative and significant (respectively, $r = -0.64$ and $r = 0.61$, $N = 19$, $p < 0.05$). This indicates that the relationship also persists across the other continents.

Method of correlated vectors

The method of correlated vectors consists of correlating the factor loadings of indicator variables with

the correlations between each indicator variable and the criteria variable (see Kirkegaard, 2014). The method was employed to test the hypothesis that SNPs with higher p-values would be more likely to be false positives, hence if there has been natural selection, less selection signal would be detected. The composite ADHD factor (Table 10, 5th col.) was a criterion variable and the unlinked SNPs were correlated with it. Spearman rank-order correlation was slightly negative (in line with predictions) but nonsignificant: $r = -0.26$ ($N = 17$, $p = 0.31$). Using all the 42 SNPs, the Spearman rank-order correlation was $r = 0.37$ and significant ($p = .015$). This relationship is plotted in Figure 2.

Figure 2. Regression of ADHD risk alleles' correlation with ADHD composite genetic factor on alleles' p-values.



Correlation of population genetic factors with aggregate phenotypic measures

A systematic review and meta-analysis concluded that “the large variability of ADHD/HD prevalence rates worldwide resulted mainly from methodological differences across studies” (Polanczyk et al., 2007). Since prevalence rates do not reflect an underlying phenotype, correlating them with genetic scores would not be useful. However, since there is a genetic correlation between GI and ADHD, and natural selection has stronger effects upon higher-order constructs with pervasive effects on life outcomes and survival, such as GI (Jensen, 1998; Gottfredson, 1997; Gordon, 1997; Gottfredson, 2004) we expect to find a correlation between population-level ADHD genetic scores and aggregate measures of cognitive capacity.

The correlation between country/ethnic phenotypic IQ and the ADHD genetic factor was highly negative ($r = -0.91$, $N = 23$, $p < .001$).

STUDY 2 (ALFRED dataset)

The 4 SNPs influencing general intelligence used by Piffer (2015a) were searched on ALFRED. When a SNP was not found in ALFRED, the SNP in close linkage disequilibrium ($r > 0.8$) was used. Linkage calculator (<http://www.broadinstitute.org/mpg/snap/>) based on 1000 Genomes phase 1, CEU data. Corresponding SNPs are in brackets in the spreadsheet file.

None of the 3 new SNPs was found in ALFRED with this method, apart from the one loading in the wrong direction (rs17522122) in the 1000 Genomes analysis. Thus, analysis of ALFRED was limited to the 4 SNPs.

Table 8 shows the factor loadings of the 4 GI-related SNPs in ALFRED.

Table 8. Factor loadings of 4 GI-related SNPs in ALFRED.

SNP	Factor loading
rs9320913 A	0.43
rs11584700 G	0.70
rs4851266 T	0.84
rs236330 C	0.72

Factor analysis of ADHD risk alleles

In total, 13 (out of the 42) SNPs were found in ALFRED (including those in close LD). Results of factor analysis are reported in Table 9.

Table 9. Factor loadings of 13 ADHD SNPs in ALFRED.

SNP	Factor Loading
rs507533.G	-0.22
rs6453417.T	0.89
rs7160641.T	0.81
rs4936536.G	-0.31
rs10026084.A	0.63
rs4673145.A	-0.24
rs2206922.T	0.72
rs1982863.T	-0.48
rs2841633.T	0.72
rs12668989.T	0.32
rs4635724.G	-0.79
rs4580847.G	0
rs3026685.T	-0.40

The factor was negatively correlated to the GI SNP factor ($r=-0.81$, $N=50$).

Since the method of correlated vectors showed an abundance of noise in the SNPs with lower significance, another factor analysis was carried out including only the SNPs with the p-value in the top half of the total 13 ($N=6$). Results are reported in Table 10.

Table 10. Top 6 ADHD SNPs loadings.

SNP	Factor Loading
rs507533.G	-0.28
rs6453417.T	0.90

rs7160641.T	0.75
rs4936536.G	-0.30
rs10026084.A	0.65
rs4673145.A	-0.28

In accord with the hypothesis of stronger selection signal among the genuine hits, the correlation between the extracted factor and the GI SNP factor slightly improved over the all-inclusive factor analysis ($r = -0.9$ vs -0.81). Factor scores are reported in Table 11.

Visual inspection of the regression plot suggests that the distribution of scores is not driven by outliers as in the 1000 Genomes database but is fairly uniform across continental clusters and populations (Figure 3).

Figure 3. Relationship between 4 GI SNP factor and ADHD (top 6) SNP risk factor, by ethnic group.



Table 11. Factor scores for ADHD risk and GI-related alleles in ALFRED.

Continent	Population	ADHD factor	ADHD top 6	4 GI factor
Africa	Bantu	1.88	1.66	-2.13
Africa	San	2.52	2.58	-1.85
Africa	Biaka	1.91	1.83	-1.48
Africa	Mbuti	1.90	1.76	-1.68
Africa	Yoruba	1.86	1.35	-2.07
Africa	Mandenka	1.66	1.14	-1.70
Africa	Mozabite	0.49	0.51	-0.67
Middle East	Bedouin	0.15	-0.12	-0.43
Middle East	Druze	0.01	0.31	0.10
Middle East	Palestinian	0.47	0.36	-0.14
Europe	Adygey	-0.01	0.08	0.37
Europe	Basque	-0.06	0.16	-0.27
Europe	French	0.02	0.08	0.06

Europe	Italians	-0.17	-0.14	0.23
Europe	Orcadian	0.06	0.00	0.63
Europe	Russian	0.32	0.43	-0.32
Europe	Sardinian	0.20	0.64	-0.36
C Asia	Burusho	-0.05	-0.24	0.06
C Asia	Kalash	0.36	0.50	0.33
C Asia	Pashtun	0.49	0.55	-0.36
C Asia	Balochi	0.37	0.30	0.29
C Asia	Brahui	0.26	0.26	-0.17
C Asia	Hazara	-0.42	-0.15	0.64
C Asia	Sindhi	0.80	0.79	-0.35
E Asia	Mongolia	-0.90	-0.81	1.27
E Asia	Dai	-0.29	-0.30	0.65
E Asia	Daur	-1.21	-1.21	1.15
E Asia	Han	-0.48	-0.24	0.87
E Asia	Hezhe	-1.29	-1.80	1.18
E Asia	Japanese	-1.01	-0.96	0.59
E Asia	Koreans	-0.89	-0.80	0.93
E Asia	Lahu	-0.67	-0.81	0.71
E Asia	Miao	-0.73	-1.19	0.62
E Asia	Naxi	-1.07	-0.92	0.05
E Asia	Oroquen	-0.68	-0.68	0.75
E Asia	She	-1.07	-1.27	0.62
E Asia	Tu	-0.50	-0.14	0.61
E Asia	Tujia	-0.64	-0.61	1.42
E Asia	Uyghur	-0.38	-0.33	0.66
E Asia	Xibe	-1.04	-1.13	1.04
E Asia	Yi	-0.76	-0.45	0.98

SE Asia	Cambodians	-0.84	-0.24	0.13
Oceania	Papuan N.G	1.12	1.06	-1.16
Melanesian	Melanesian	1.74	1.77	0.26
Siberia	Yakut	-0.84	-0.95	0.54
America	Pima,Mx	-0.21	-0.21	-0.25
America	Maya,Yuc	-0.79	-0.71	-0.70
America	Amerindians	-0.29	-0.18	-1.33
America	Karitiana	-0.70	-0.88	0.07
America	Surui	-0.60	-0.67	-0.40

Controlling for migration and drift

Piffer (2015c) developed a method to assess the impact of population structure on estimates of selection inferred from allele frequency patterns.

The correlations between IQ distances and Fst, ADHD factor distances were, respectively: $r= 0.588$, 0.63 .

Accordingly, Fst distances published by Piffer (2015c) were used in a multiple linear regression with ADHD factor distances to predict IQ distances between populations. After list-wise deletion there were 253 caes (NA=72). Standardized Betas were 0.168 and 0.483 for Fst and ADHD factor, respectively.

4 and 6 SNPs g factor distances were also employed as dependent variables to assess the predictive power of ADHD distances net of population structure. Fst and ADHD had similar Betas (0.409,0.443), (0.5,0.429) for the 4 and 6 SNPs factors, respectively.

DISCUSSION

Factor analysis of allele frequencies obtained from GWAS hits by independent studies on different but partly overlapping phenotypes (ADHD and GI) shows that they follow an inverse spatial distribution.

Robustness (reliability) of the findings was provided by the internal consistency of the measures: factor analysis of two sets of unliked ADHD risk alleles (using 1000 Genomes) located on different chromosomes yielded two very similar factors ($r= 0.96$). Moreover, most of the factor loadings (11/17) were positive (in the right direction), ($p= 0.17$). Rietveld's and Davies' three top hits were related to two distinct phenotypes (educational attainment and fluid intelligence, respectively), yet the two sets of alleles produced similar frequency patterns ($r=0.79$). Moreover, the effect on educational attainment of a genomic region on chromosome 6 was replicated by Davies for fluid intelligence.

Validity was confirmed by the strong negative correlation of the factor of ADHD SNPs with a factor extracted from an independent set of 6 GI SNPs produced by several studies, with frequencies obtained from the ALFRED and 1000 Genomes datasets, comprising 50 and 26 populations respectively. GWAS hits with a lower p-value had a stronger relationship with the ADHD SNP factor,

confirming the hypothesis that a stronger signal of selection would be found in alleles that are likely to be true positives. Moreover, the factor was negatively related to average phenotypic country/ethnic GI.

The 6 GI-related SNPs were not in linkage disequilibrium (LD) with the 17 unlinked ADHD alleles, so the correlation between the two factors is not confounded by LD. However, two of them (GI-related rs11584700 and ADHD risk rs2802837) were in nearby genomic regions (around 1.2 Mbp apart), on chromosome 1.

Another noteworthy finding is that groups which are the most distant genetically (Africans and Oceanians) appear to have the most similar genetic scores for cognitive traits, whereas genetically similar groups (East Asians and Native Americans) differ by almost a standard deviation (Tables 15 and 16). This fact is hard to square with an explanation of the results in terms of genetic drift and instead favours a model of differential selection pressures faced by isolated populations living in different environments.

We controlled for the effect of population structure using the method outlined by Piffer(2015c). This showed that although the ADHD factor was a stronger predictor of IQ distances between populations than F_{st} , as indicated by F_{st} , it lost part of its predictive power (Beta= 0.483). Moreover, it did not have more predictive power of GI factor distances than F_{st} . These results together are weaker than what found by Piffer for the 4 SNPs g factor(2015c), which retained all its predictive power on IQ distances when regressed with F_{st} .

We see at least two possible explanations for the general results: 1) The SNPs found by the ADHD GWAS are in reality GI-increasing alleles and have no specific effect on ADHD. 2) ADHD and GI have undergone opposite selection pressure, that is ancestral environments that selected for higher GI also selected against behavioral predispositions to ADHD, such as inattention and impulsivity. In such a hypothetical scenario, harsher environments could have placed more demands on focused attention and problem solving skills that resulted in survival and fitness differentials among carriers of different alleles.

Obviously, a third possible explanation is that the ADHD hits were not under selection and represent population structure just as any other random set of SNPs. The not entirely satisfying results of regressing these with F_{st} distances does not allow us to rule out this possibility.

Another limitation of this study is that ADHD prevalence rates are too affected by different diagnostic criteria to provide a universal measure that can be used for a cross-cultural comparison.

Future research

Future studies should replicate the present analysis when new GI and ADHD SNPs are reported.

REFERENCES

- Benyamin, B., Pourcain, B.St., Davis, O.S., Davies, G., Hansell, N.K., Brion, M.-J.A. et al. (2013). Childhood intelligence is heritable, highly polygenic and associated with FBNP1L. *Molecular Psychiatry*, 19: 253-258.
- Berg, J.J. & Coop, G. (2014). A Population Genetic Signal of Polygenic Adaptation. *PLOS Genetics*, Doi: 10.1371/journal.pgen.1004412
- Clarke G.M., Anderson .CA., Pettersson F.H., Cardon L.R., Morris A.P., Zondervan K.T. (2011). Basic statistical analysis in genetic case-control studies. *Nat Protoc* 6: 121–33. doi:10.1038/nprot.2010.182
- Davies, G., Tenesa, A., Payton, A., Yang, J., Harris, S.E., Liewald, D., Xiayi, K., Le Hellard, S. et al. (2011). Genome-wide association studies establish that human intelligence is highly heritable and

polygenic. *Molecular Psychiatry*, 16: 996-1005.

Davies, G., Armstrong, N.,... Deary, I.J. (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*, 1-10.[doi:10.1038/mp.2014.188](https://doi.org/10.1038/mp.2014.188)

Faraone, S. V., Perlis, R. H., Doyle, A. E., Smoller, J. W., Goralnick, J. J., Holmgren, M. A., & Sklar, P. (2005). Molecular genetics of attention-deficit/hyperactivity disorder. *Biological psychiatry*, 57(11), 1313-1323.

Frazier, T. W., Demaree, H. A., & Youngstrom, E. A. (2004). Meta-analysis of intellectual and neuropsychological test performance in attention-deficit/hyperactivity disorder. *Neuropsychology*, 18(3), 543.

Gordon, R. A. (1997). Everyday life as an intelligence test: Effects of intelligence and intelligence context. *Intelligence*, 24(1), 203-320.

Gottfredson, L. S. (2004). Life, death, and intelligence. *Journal of Cognitive Education and Psychology*, 4(1), 23-46.

Gottfredson, L. S. (1997). Why g matters: The complexity of everyday life. *Intelligence*, 24(1), 79-132.

Groen-Blokhuis, M. M., Middeldorp, C. M., Kan, K. J., Abdellaoui, A., van Beijsterveldt, C. E., Ehli, E. A., ... & Psychiatric Genomics Consortium ADHD Working Group. (2014). Attention-Deficit/Hyperactivity Disorder Polygenic Risk Scores Predict Attention Problems in a Population-Based Sample of Children. *Journal of the American Academy of Child & Adolescent Psychiatry*, 53(10), 1123-1129.

Jensen, A. R. (1998). *The g factor: The science of mental ability*. Westport, CT: Praeger.

Kessler, R., Adler, L., Barkley, R., Biederman, J., Conners, C., Demler, O., ... & Zaslavsky, A. (2006). The prevalence and correlates of adult ADHD in the United States: results from the National Comorbidity Survey Replication. *American Journal of Psychiatry*, 163(4), 716-723.

Kirkegaard, E. O. W. (2014). Crime, income, educational attainment and employment among immigrant groups in Norway and Finland. *Open Differential Psychology*.

Kuntsi, J., Eley, T. C., Taylor, A., Hughes, C., Asherson, P., Caspi, A., & Moffitt, T. E. (2004). Co-occurrence of ADHD and low IQ has genetic origins. *American Journal of Medical Genetics Part B: Neuropsychiatric Genetics*, 124(1), 41-47.

Martin, J., Hamshere, M.L., Stergiakouli, E., O'Donovan, M.C., Thapar, A. (2014). Neurocognitive abilities in the general population and composite genetic risk scores for attention-deficit hyperactivity disorder. *Journal of Child Psychology and Psychiatry*, [doi:10.1111/jcpp.12336](https://doi.org/10.1111/jcpp.12336)

Mayes, S. D., Calhoun, S. L., & Crowell, E. W. (2000). Learning disabilities and ADHD overlapping spectrum disorders. *Journal of learning disabilities*, 33(5), 417-424.

Lynn, R., & Vanhanen, T. (2012). *Intelligence: A unifying construct for the social sciences*. London: Ulster Institute for Social Research.

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: 168200.

Piffer, D. (2014). Simple statistical tools to detect signals of recent polygenic selection. *IBC* 2014;6:1, 1-6 • DOI: 10.4051 / ibc.2014.6.1.0001

Piffer, D. & Lynn, R. (2014). New evidence for differences in fluid intelligence between North and South Italy and against school resources as an explanation for the north-south IQ differential. *Intelligence*, 46:

246-249.

Piffer, D. (2015a). Estimating the genotypic intelligence of populations and assessing the impact of socioeconomic factors and migrations., *The Winnower* 2:e142299.93508 (2015). DOI:10.15200/winn.142299.93508

Piffer, D. (2015b). Evidence of polygenic selection on human stature inferred from spatial distribution of allele frequencies [v1; ref status: awaiting peer review, <http://f1000r.es/4ye>] *F1000Research* 2015, 4:15 (doi:10.12688/f1000research.6002.1)

Piffer, D. (2015c). Intelligence GWAS hits: Selection signal or population structure? A test of the null hypothesis, *The Winnower* 2:e142593.30710 (2015). DOI:10.15200/winn.142593.30710

Polanczyk, G., Silva de Lima, M., Horta, B.L., Biederman, J., Rohde, L.A. (2007). The Worldwide Prevalence of ADHD: A Systematic Review and Metaregression Analysis. *American Journal of Psychiatry*, 164: 942-948.

Rietveld, C. A., Medland, S. E., Derringer, J., Yang, J., Esko, T., Martin, N. W., Westra, H. J., Shakhbazov, K., Abdellaoui, A., Agrawal, A., et al. (2013). GWAS of 126,559 individuals identifies genetic variants associated with educational attainment. *Science* 340, 1467-1471.

Stergiakouli, E., Hamshere, M., Holmans, P., Langley, K., Zaharieva, I., Hawi, Z., ... & Thapar, A. (2012). Investigating the contribution of common genetic variants to the risk and pathogenesis of ADHD. *American Journal of Psychiatry*, 169, 186–194.

Turchin, M. C., Chiang, C. W., Palmer, C. D., Sankararaman, S., Reich, D., Genetic Investigation of, A. T. C., and Hirschhorn, J. N. (2012). Evidence of widespread selection on standing variation in Europe at height-associated SNPs. *Nat Genet* 44, 1015-1019.

Ward, M.E., McMahon, G., St Pourcain, B., Evans, D.M., Rietveld, C.A., et al. (2014) Genetic Variation Associated with Differential Educational Attainment in Adults Has Anticipated Associations with School Performance in Children. *PLoS ONE* 9(7): e100248. doi:10.1371/journal.pone.0100248

Willcutt, E. G. (2012). The prevalence of DSM-IV attention-deficit/hyperactivity disorder: a meta-analytic review. *Neurotherapeutics*, 9(3), 490-499.

Zhao, N. (2009). The minimum sample size in factor analysis. URL <https://www.encorewiki.org/display/~nzhao/The+Minimum+Sample+Size+in+Factor+Analysis> accessed 19th Dec. 2014.