Title:

Analysis of differentially methylated regions in primates and non-primates

provides support for the evolutionary hypothesis of schizophrenia

Niladri Banerjee[a,b], Tatiana Polushina[a,b], Francesco Bettella[c,d],

Vidar M. Steen[a,b], Ole A. Andreassen[c,d], Stephanie Le Hellard[a,b]

a. NORMENT - K.G. Jebsen Center for Psychosis Research, Department of Clinical

Science, University of Bergen, Bergen, Norway

b. Dr. Einar Martens Research Group for Biological Psychiatry, Department of

Medical Genetics, Haukeland University Hospital, Bergen, Norway

c. NORMENT - K.G. Jebsen Center for Psychosis Research, Institute of Clinical

Medicine, University of Oslo, Oslo, Norway

d. NORMENT - K.G. Jebsen Centre, Division of Mental Health and Addiction, Oslo

University Hospital, Oslo, Norway

Correspondence to: Prof. Stéphanie Le Hellard, Department of Clinical Medicine,

Laboratory Building, Haukeland University Hospital, N-5021 Bergen, Norway.

Telephone: +47-55 97 53 37

Fax: +47-55 97 54 79

Email: stephanie.hellard@uib.no

Authors' email addresses

Niladri Banerjee   niladri.banerjee@uib.no

Tatiana Polushina   tatiana.polushina@uib.no

Francesco Bettella   francesco.bettella@medisin.uio.no

Vidar M. Steen   vidar.martin.steen@helse-bergen.no

Ole A. Andreassen   o.a.andreassen@medisin.uio.no

Stéphanie Le Hellard   stephanie.hellard@uib.no

Abstract

Introduction: The persistence of schizophrenia in human populations separated by

geography and time led to the evolutionary hypothesis that proposes schizophrenia as

a by-product of the higher cognitive abilities of modern humans. To explore this

hypothesis, we used here an evolutionary epigenetics approach building on

differentially methylated regions (DMRs) of the genome.

Methods: We implemented a polygenic enrichment testing pipeline using the

summary statistics of genome-wide association studies (GWAS) of schizophrenia and

12 other phenotypes. We investigated the enrichment of association of these traits

across genomic regions with variable methylation between modern humans and great

apes (orangutans, chimpanzees and gorillas; primate DMRs) and between modern

humans and recently extinct hominids (Neanderthals and Denisovans; non-primate

DMRs).

Results: Regions that are hypo-methylated in humans compared to great apes show

enrichment of association with schizophrenia only if the major histocompatibility

complex (MHC) region is included. With the MHC region removed from the analysis,

only a modest enrichment for SNPs of low effect persists. The INRICH pipeline

confirms this finding after rigorous permutation and bootstrapping procedures.

Conclusion: The analyses of regions with differential methylation changes in humans

and great apes do not provide compelling evidence of enrichment of association with

schizophrenia, in contrast to our previous findings on more recent methylation

differences between modern humans, Neanderthals and Denisovans. Our results

further support the evolutionary hypothesis of schizophrenia and indicate that the

origin of some of the genetic susceptibility factors of schizophrenia may lie in recent

human evolution.

Key Words: schizophrenia; evolutionary hypothesis; epigenetics; differentially

methylated regions; primates; Neanderthals.

1. Introduction

Schizophrenia is a psychiatric disorder with a prevalence rate of 2.7-8.3/1,000 persons (Messias et al., 2007) and heritability estimated between 60-90% (Cardno et al., 1999; Lichtenstein et al., 2009; Skre et al., 1993; Sullivan et al., 2003). It occurs at quite similar rates across populations worldwide (Ayuso-Mateos, 2002; Brüne, 2004; WHO, 1973) and written records describing its symptoms exist dating back 5,000 years (Jeste et al., 1985). This consistent persistence of the disease despite reduced fecundity (Brüne, 2004; Nichols, 2009) and increased mortality is a paradox (Bassett et al., 1996; Brown, 1997; Larson and Nyman, 1973), since the reduced fecundity of patients afflicted with schizophrenia does not appear to eliminate the disease from the population (Power et al., 2013) Part of the reason may be due to afflicted individuals reproducing prior to the onset of the disease (Markow, 2012). Another contributing factor could be that schizophrenia risk variants may have provided an advantage to the kin of the affected by conferring superior creative and intellectual abilities upon them (Kyaga et al., 2011; Nichols, 2009). To explain the constant occurrence of the disease, TJ Crow (Crow, 1997, 1995) proposed the so-called evolutionary hypothesis of schizophrenia, which suggests that the disease is a consequence of human evolution: the higher cognitive abilities of modern-day humans, including language, may predispose to psychiatric illnesses such as schizophrenia (Crow, 2008, 2000, 1997).

In the post-genomic era (Lander et al., 2001; Venter et al., 2001), emerging lines of evidence are lending support to this hypothesis. Crespi et al. (Crespi et al., 2007) were amongst the first to show that genes with evidence of recent positive selection in

humans are also implicated more frequently in schizophrenia. More evidence has

been provided by studies based on comparative genomics (Pollard et al., 2006;

Srinivasan et al., 2015; Xu et al., 2015), a field in which genomes of progressively

older species are compared to identify substitutions and mutations that help estimate

divergence between the species. For instance, a group of regions defined by negative

Neanderthal selective sweep (NSS) scores describe the selective evolution of genomic

regions in modern-day humans over Neanderthals (Burbano et al., 2010; Green et al.,

2010). These regions were shown by Srinivasan et al. (2015) to be enriched for

schizophrenia risk markers, in line with the evolutionary hypothesis of schizophrenia.

Other regions known as human accelerated regions (HARs) (Gittelman et al., 2015;

Pollard et al., 2006; Xu et al., 2015), first described by Pollard et al. (2006), show

accelerated evolution in humans compared to primates or mammals. HARs have also

provided some evidence of enrichment of association with schizophrenia (Xu et al.,

2015), but these findings may have been driven by a few genes since they were not

replicated using a polygenic approach (Srinivasan et al., 2017, 2015) .

While several studies have looked at the evolution of the genome (Bird et al., 2007;

Bush and Lahn, 2008; Gittelman et al., 2015; Paaby and Rockman, 2014; Pollard et

al., 2006), there are reports that the epigenome is evolving as well (Gokhman et al.,

2014; Hernando-Herraez et al., 2015, 2013; Mendizabal et al., 2014; Molaro et al.,

2011). This provides new insights into events leading to the speciation and divergence

of modern humans. The epigenome refers to the layer of chemical modifications, such

as methylation and histone modifications, to the genome that regulate gene expression

(Bernstein et al., 2007; Kundaje et al., 2015; Rivera and Ren, 2013). For instance,

Gokhman et al. (2014) compared the methylomes of humans with Neanderthals and

Denisovans. They reported that while 97% of the methylome was comparable

between humans, Neanderthals and Denisovans, some regions showed differential

methylation between the three hominids. Previously (Banerjee et al., 2017), we

analysed the differentially methylated regions (DMRs) identified for Neanderthals,

Denisovans and modern humans by Gokhman et al. (2014), and found evidence that

the regions of the genome with human-specific DMRs harbour relatively more genetic

variants associated with schizophrenia than the rest of the genome, i.e. the DMRs

were enriched for SCZ markers both at the single-nucleotide polymorphism (SNP)

level and at the gene level. These human-specific DMRs thus provide evidence of

enrichment of methylation changes in regions harbouring genetic variants associated

with schizophrenia, at least since the divergence from Neanderthals and Denisovans

(Banerjee et al., 2017).

Here, we sought to determine if evolutionarily older methylation differences can

provide a further timeframe for the origin of schizophrenia risk markers in the human

lineage. We asked whether we can find epigenetic evidence that the origin of

schizophrenia risk markers predates the origins of the Homo genus, i.e. before the

divergence of chimpanzees and humans around 6-8 million years ago (MYA) (Glazko

and Nei, 2003; Langergraber et al., 2012). We tested this hypothesis by analysing

primate DMRs that trace an evolutionary history of at least 13 million years (Glazko

and Nei, 2003; Hasegawa et al., 1985; Rannala and Yang, 2003). We used the same

statistical analyses as described by Lee et al. (2012), Schork et al. (2013), and

Srinivasan et al. (2015) to test for polygenic enrichment of a set of markers from

genome-wide association studies (GWAS). We interrogated regions of the human

genome which are hypo- or hyper-methylated in comparison to the corresponding

ones in chimpanzees, gorillas and orangutans for enrichment of genetic variants

associated with schizophrenia or other human traits.

2. Materials and methods

2.1. GWAS data

Summary statistics for thirteen different phenotypes were obtained from their respective published GWAS studies: schizophrenia (SCZ) (Ripke et al., 2014), bipolar disorder (BPD) (Sklar et al., 2011), attention deficit hyperactivity disorder (ADHD) (Demontis et al., 2017), rheumatoid arthritis (RA) (Stahl et al., 2010), blood lipid markers (high density lipoprotein (HDL), low density lipoprotein (LDL), triglycerides (TG), total cholesterol (TC)) (Teslovich et al., 2010), blood pressure (systolic blood pressure (SBP), diastolic blood pressure (DBP)) (Ehret et al., 2011), body mass index (BMI) (Locke et al., 2015), height (Wood et al., 2014) and intelligence (Sniekers et al., 2017). For studies published with hg18 coordinates (BPD, SBP, DBP, HDL, LDL, TG, TC, RA), conversion to hg19 was performed using the command line version of the liftOver tool from the UCSC Genome Browser (Karolchik et al., 2014) (http://hgdownload.cse.ucsc.edu/downloads.html #utilities_downloads). For BMI and height SNPs, the genomic coordinates were obtained by mapping them to the assembly of 1,000 Genomes Project (1KGP) Phase 1 reference panel SNPs (Durbin et al., 2012).

2.2. Human hypo- and hyper-methylated regions from primate DMRs

These methylated regions were retrieved from the study by Hernando-Herraez et al. (2015), who identified them by comparing the methylation profile of DNA from peripheral blood samples of orangutans, chimpanzees and gorillas to that of humans. Since the DMRs are determined by comparing humans with other primates, we refer to this set collectively as primate DMRs. Both hypo- and hyper-methylated DMRs

from humans were analysed. As these DMRs are identified in the same tissues in all

samples, they are considered to represent species-specific methylation differences, not

tissue-specific methylation differences (Gokhman et al., 2014). Altogether, the human

hypo- and hyper-methylated DMRs can be used to represent an evolutionary course of

history spanning from at least 13 MYA (Glazko and Nei, 2003; Langergraber et al.,

2012), when orangutans diverged from the common ancestors, to 6 MYA, when the

chimpanzees and humans diverged from each other (Glazko and Nei, 2003;

Langergraber et al., 2012). Since our interest was in human-specific enrichment, we

focused the analyses on human hypo- and hyper-methylated DMRs.

2.3. Differentially methylated regions (DMRs) from Neanderthals, Denisovans and

modern humans

As previously described (Gokhman et al., 2014), these methylated regions have been

identified by comparing the methylomes of osteoblasts from modern-day humans

with those from Neanderthals and Denisovans. We refer to them in this paper as non-

primate DMRs. Gokhman et al. (2014) devised a strategy utilizing information in the

form of cytosine (C) to thymine (T) ratios to decipher the ancient methylomes of

Neanderthals and Denisovans. Subsequently, they compared the methylomes of

Neanderthals, Denisovans and modern humans and inferred the species in which the

methylation variation likely took place; this information was used to classify the

DMRs as Neanderthal-specific, Denisovan-specific and human-specific. These DMRs

represent species-specific methylation (Gokhman et al., 2014).

2.4. Neanderthal selective sweep (NSS) data

We obtained NSS marker data from Srinivasan et al. (2015). Negative scores for NSS

markers indicate positive selection in humans. Markers with such scores were used in

the downstream analyses.

2.5. SNP assignment with LDsnpR

The previously published R-based software package LDsnpR (Christoforou et al.,

2012) was utilized for assigning SNPs to the respective DMRs using LD (linkage

disequilibrium)-based binning at $r^2 \geq 0.8$ in R (R Core Team, 2017). LD-based binning

makes it possible to determine whether SNPs from a specific GWAS are in LD with

the DMR of interest. Using LD allows the capture of a greater number of relevant

SNPs in comparison to an approach where only physically overlapping SNPs are

considered. The LD file utilized was in HDF5 format and was constructed from the

European reference population of 1KGP and can be publicly downloaded at:

http://services.cbu.uib.no/software/ldsnpr/Download.

2.6. Enrichment analyses based on stratified quantile-quantile (QQ) plots

QQ plots are an essential method used in GWASs to depict the presence of true

signals. They help to visually observe the spread of data and deviations from the null

distribution. Under the null hypothesis, no difference is expected between the

observed and expected distributions of data. As such, a line of no difference or null

line is obtained that is equidistant from both X and Y axes. However, if the null

hypothesis were to be false, there would be a deviation of the observed data

distribution from the expected data distribution. As described in depth by Schork et al.

(2013), a leftward deflection of the observed distribution from the null line represents

enrichment – the greater the leftward deflection, the stronger the enrichment of true

signals. This method has been used recently not only to show how specific genomic

annotation affects the distribution of disease SNPs with true signals (Schork et al.,

2013), but also to demonstrate that regions of recent evolution are enriched for

schizophrenia markers (Banerjee et al., 2017; Srinivasan et al., 2015). We took the

SNPs that are in LD with the DMR regions and plotted their *p*-value distributions

from various GWASs. The observed *p*-value distributions were then determined to be

enriched or not using conditional Q-Q plots as described by Schork et al. (2013).

Genomic inflation was corrected by $\lambda_{GC.}$

2.7. INRICH-based enrichment analysis

The stratified QQ plots provide a visual depiction of data distributions and enrichment

of true signals within a stratum of data, but they do not quantify this enrichment.

Therefore, we used the INterval EnRICHment (INRICH) analysis tool to statistically

quantify the enrichment observed. This pipeline performs permutation and

bootstrapping procedures to determine with statistical confidence whether LD-

implicated genomic intervals are enriched in specific gene sets (Lee et al., 2012). The

INRICH analysis takes into account several potential biases that can otherwise lead to

false positives, such as variable gene size, SNP density within genes, LD between and

within genes, and overlapping genes in the gene sets. We used the same procedure

reported previously (Banerjee et al., 2017; Xu et al., 2015) with SNPs in the extended

MHC region and SNPs with MAF <0.05 excluded from the analysis. Additional

details can be found in the Supplementary Information.

3. Results

3.1. Co-localisation of human hypo-methylated regions and genetic variants

associated with schizophrenia in the MHC.

We ascertained whether there is any enrichment of human hypo- and hyper-

methylated regions in schizophrenia-associated SNPs. Using previously published

methodology (Christoforou et al., 2012), we mapped schizophrenia markers to human

hypo-methylated regions (hypo-DMRs) and hyper-methylated regions (hyper-DMRs).

Out of a total of ~9.4 million SCZ markers obtained from the GWAS, 10,165 markers

tagged hypo-DMRs and 4,503 tagged hyper-DMRs.

Figure 1A shows the conditional QQ plots for schizophrenia markers (all markers, the

hypo-DMR set and the hyper-DMR set) including those in the MHC region. For

hypo-DMR markers (Supplementary Dataset 1), we observed a significant enrichment

as depicted by the leftward deviation. No enrichment was observed for hyper-DMR

markers. Since the MHC region is a region of extended linkage disequilibrium, which

can bias the enrichment estimates, and since it is the main region of association with

schizophrenia, we also tested the enrichment with the MHC region removed (Figure

1B). Under these conditions there is a trend for enrichment of hypo-DMR markers at

higher *p*-value thresholds, but this enrichment is substantially less than when the

MHC is included (Figure 1A).

3.2. Enrichment of markers is not seen for other human traits

Next, we tested if the human hypo- and hyper-methylated regions are enriched for

other human traits and phenotypes. We tested a total of thirteen different phenotypes,

full details of which can be found in section 2.1. Each GWAS had been performed

with a different number of genotyped SNPs, and this difference could potentially bias

our results. To circumvent this, we created a list of ~2.4 million common SNPs that

were genotyped across all the phenotypes investigated in the present study. Only

SNPs on this list were used for enrichment analysis.

As can be seen in Fig. 2, no enrichment was observed in any of the traits, with the

possible exception of height at higher *p*-value threshold markers. The common list of

markers did not contain the MHC region and as such no enrichment is observed for

schizophrenia either.

3.3. Evidence of enrichment for hypo-methylated regions with SNPs at high p-values

The enrichment plots allowed us to visually ascertain enrichment in the datasets.

However, they did not give any indication of the statistical robustness of the

enrichment. To ascertain if the human hypo- and hyper-methylated regions are

statistically enriched for schizophrenia and height markers, we implemented the

INRICH pipeline, which performs 10,000 permutations and 5,000 bootstrapping

calculations, to determine with statistical confidence the enrichment observed (Lee et

al., 2012).

The INRICH analysis confirmed a significant ($p<0.05$) enrichment of association for

human hypo-DMRs, but not hyper-DMRs, with schizophrenia at SNPs of higher *p*-

value thresholds *($p<10e-3$ to $p<10e-4$)* (Fig. 3). This enrichment was at the gene level,

and complemented the enrichment observed at the SNP level for higher *p*-value

thresholds (Fig. 1B). Importantly, this enrichment persisted upon testing a pruned

schizophrenia dataset (Supplementary Fig. 1). The enrichment was however not

significant at the genome-wide threshold ($p<5x10e-8$) and was much weaker than that

observed for non-primate DMRs (Fig. 3). We also observed a similar trend for height

where there was enrichment at SNPs of higher but not lower p-value thresholds. This

enrichment was similarly less pronounced than for non-primate DMRs

(Supplementary Fig. 2).

4. Discussion

In our study, we investigated if regions of the human genome whose methylation has

evolved since the divergence of modern humans from great apes are enriched for

markers of schizophrenia. We found evidence that there is enrichment for hypo-

methylated DMRs driven by the MHC locus, a known risk region that harbours the

most significant schizophrenia GWAS markers (Ripke et al., 2014). When the MHC

region was excluded from the analysis, there remained a trend towards enrichment of

hypo-DMRs driven by SNPs of higher p-value thresholds. This finding was

complemented by the INRICH analyses that indicated significant enrichment among

SNPs of higher p-value thresholds. When analysing a global SNP list common to

GWAS of several traits, we failed to find evidence of enrichment of any trait with the

possible exception of height at higher SNP p-value thresholds. We tested this further

with the INRICH pipeline, which revealed gene-level enrichment of LD intervals for

height markers below the genome-wide threshold ($p<5\text{x}10\text{e-}8$). Compared to our

previous study, in which we demonstrated enrichment of association with

schizophrenia for non-primate DMRs that were derived by comparing human,

Neanderthal and Denisovan methylomes (Banerjee et al., 2017), the primate DMRs

tested here show far less enrichment. The primate and non-primate DMRs have very

little overlap, which suggests that the methylation changes that took place since the

divergence of modern humans from Neanderthals and Denisovans occurred in

different regions of the genome compared to those that took place since divergence

from great apes.

The central role of the MHC region in the enrichment of human hypo-methylated

regions poses interesting questions. The MHC region is known for its complex LD

architecture, which renders the interpretation of genetic signals very challenging.

Other groups have previously reported that the MHC region is one of the fastest

evolving regions of the human genome (Meyer et al., 2017) and have implicated it in

mate preference (Bernatchez and Landry, 2003; Kromer et al., 2016; Potts and

Wakeland, 1990; Roberts et al., 2008; Winternitz et al., 2017), odour perception

(Roberts et al., 2008; Santos et al., 2005) and immune response (Benacerraf, 1981;

Horton et al., 2004). Recently it was shown that a large proportion of the association

of the region with schizophrenia can be explained by complement C4 haplotypes that

include C4 copy number variation (Sekar et al., 2016). Nevertheless, there remains a

part of the association in this region that is unexplained (Gejman et al., 2011) and will

need further investigation. It is interesting to consider the possibility that the MHC

region and the immune system in general play a central role in evolution at the

epigenomic as well as at the genomic level (Meyer et al., 2017; Potts and Wakeland,

1990; Sommer, 2005; Traherne, 2008). The mechanisms by which hypo-methylation

could influence the aforementioned processes are open to speculation since the MHC

region has more than 200 genes in close physical proximity and LD with one another

(Beck et al., 1999). This makes it hard to interpret the exact biological consequences

of our findings.

Interestingly, the gene-level analysis via INRICH seems to suggest enrichment of

SNPs of higher *p*-value thresholds in primate DMRs for both schizophrenia and

height. This enrichment is far lower than what we found for non-primate DMRs for

both schizophrenia and height (Banerjee et al., 2017) and which persisted for

schizophrenia even with pruned datasets.

The very small overlap between primate and non-primate DMRs might suggest that

the divergence from Neanderthals and Denisovans brought about more significant

methylation changes in regions implicated in the aetiology of schizophrenia and

height than the divergence from great apes. In other words, our results might suggest

that the evolutionary factors that regulate methylation variation acted on different

segments of the genome at different time points. So while the methylation variation

since the divergence from Neanderthals and Denisovans may mark a genome-wide

increase of schizophrenia susceptibility (Banerjee et al., 2017), the methylation

variation from the time period between 13 and 6 MYA appears not to have

significantly increased the risk for schizophrenia (except possibly for some markers in

the MHC region),

Our results are also in line with the findings of Srinivasan et al. (2017), who failed to

find evidence of enrichment of schizophrenia using genomic markers of evolution

dating back to 200 MYA. The same authors also reported enrichment of association

for regions of more recent evolution in modern humans (Srinivasan et al., 2015).

Interestingly, one of the evolutionary proxies used by Srinivasan and colleagues

(2017), namely HARs, also showed enrichment for height, similar to our recent study

(Banerjee et al., 2017). This suggests that regions controlling both genomic and

epigenomic variation in height may also be driven by recent evolution. Finally, our

results agree well with the observation by Srinivasan et al. (2017) of some

involvement of the MHC in an early evolutionary context.

Although our results are in line with several findings in the field, the current methods

have some limitations. Highly polygenic traits such as schizophrenia have a large

number of genetic loci contributing to the aetiology of a disease (Bulik-Sullivan et al.,

2015; Schork et al., 2016). The ability to detect these large numbers of genetic loci is

dependent on the sample size and adequate statistical power (Schork et al., 2016).

Consequently, the polygenic enrichment methods may be limited by the statistical

power of the respective GWAS and trait polygenicity. Furthermore, in the INRICH

analysis that uses LD-clumping of SNPs at $p<10e\text{-}3$ to $p<10e\text{-}8$, higher p-value

thresholds (e.g. $p<10e\text{-}3$) still include SNPs of lower p-values, even though they

become progressively smaller minorities. Thus, although higher p-values increase the

number of LD-clumps tested, we do not expect this to increase the Type I error rate

(Lee et al., 2012).

In conclusion, our results suggest that methylation markers tracing an evolutionary

period dating back to 13 MYA (primate DMRs) are not enriched for schizophrenia

markers, unlike methylation markers from a recent timeframe (non-primate DMRs)

(Banerjee et al., 2017). Taken in consideration with previous studies of genomic

markers of evolution dating back 200 MYA (Srinivasan et al., 2017), our results

support the hypothesis that the origins of schizophrenia lie in more recent

evolutionary events, possibly after the divergence of modern-day humans from

Neanderthals and Denisovans.

Appendix A. Supplementary data

Supplementary Information: Additional Methods, Figures and Tables

Supplementary Dataset 1: Annotation of human hypo-methylated regions with

markers of schizophrenia

References

Ayuso-Mateos, J.L., 2002. Global Burden of schizophrenia in the year 2000-Version 1 Estimates. World Health Organization.

Banerjee, N., Polushina, T., Bettella, F., Giddaluru, S., Steen, V.M., Andreassen, O.A., Le Hellard, S., 2017. Human-specific methylated regions are enriched in schizophrenia. bioRxiv 113175.

Bassett, A.S., Bury, A., Hodgkinson, K.A., Honer, W.G., 1996. Reproductive fitness in familial schizophrenia. Schizophr. Res. 21, 151–60.

Beck, S., Geraghty, D., Inoko, H., Rowen, L., 1999. Complete sequence and gene map of a human major histocompatibility complex. Nature 401, 921–923.

Benacerraf, B., 1981. Role of MHC gene products in immune regulation. Science 212, 1229–1238.

Bernatchez, L., Landry, C., 2003. MHC studies in nonmodel vertebrates: what have we learned about natural selection in 15 years? J. Evol. Biol. 16, 363–77.

Bernstein, B.E., Meissner, A., Lander, E.S., 2007. The mammalian epigenome. Cell 128, 669 – 681.

Bird, C.P., Stranger, B.E., Liu, M., Thomas, D.J., Ingle, C.E., Beazley, C., Miller, W., Hurles, M.E., Dermitzakis, E.T., 2007. Fast-evolving noncoding sequences in the human genome. Genome Biol. 8, R118.

Brown, S., 1997. Excess mortality of schizophrenia. A meta-analysis. Br. J. Psychiatry 171, 502–508.

Brüne, M., 2004. Schizophrenia—an evolutionary enigma? Neurosci. Biobehav. Rev. 28, 41–53.

Bulik-Sullivan, B.K., Loh, P.-R., Finucane, H.K., Ripke, S., Yang, J., Patterson, N., Daly, M.J., Price, A.L., Neale, B.M., Schizophrenia Working Group of the Psychiatric Genomics Consortium, 2015. LD Score regression distinguishes confounding from polygenicity in genome-wide association studies. Nature Genet. 47, 291.

Burbano, H.A., Hodges, E., Green, R.E., Briggs, A.W., Krause, J., Meyer, M., Good, J.M., Maricic, T., Johnson, P.L.F., Xuan, Z., Rooks, M., Bhattacharjee, A., Brizuela, L., Albert, F.W., de la Rasilla, M., Fortea, J., Rosas, A., Lachmann, M., Hannon, G.J., Pääbo, S., 2010. Targeted investigation of the Neandertal genome by array-based sequence capture. Science 328, 723–5.

453 Bush, E.C., Lahn, B.T., 2008. A genome-wide screen for noncoding elements
454   important in primate evolution. BMC Evol. Biol. 8, 17.
455

456 Cardno, A.G., Marshall, E.J., Coid, B., Macdonald, A.M., Ribchester, T.R., Davies,
457   N.J., Venturi, P., Jones, L.A., Lewis, S.W., Sham, P.C., et al., 1999. Heritability
458   estimates for psychotic disorders: the Maudsley twin psychosis series. Arch. Gen.
459   Psychiatry 56, 162–168.
460

461 Christoforou, A., Dondrup, M., Mattingsdal, M., Mattheisen, M., Giddaluru, S.,
462   N□then, M.M., Rietschel, M., Cichon, S., Djurovic, S., Andreassen, O.A., et al.,
463   2012. Linkage-disequilibrium-based binning affects the interpretation of GWASs.
464   Am. J. Hum. Genet .90, 727–733.
465

466 Crespi, B., Summers, K., Dorus, S., 2007. Adaptive evolution of genes underlying
467   schizophrenia. Proc. Roy. Soc. Lond. [Biol] 274, 2801–2810.
468

469 Crow, T.J., 1995. A Darwinian approach to the origins of psychosis. Br. J. Psychiatry
470   167, 12–25.
471

472 Crow, T.J., 1997. Is schizophrenia the price that Homo sapiens pays for language?
473   Schizophr. Res. 28, 127–41.
474

475 Crow, T.J., 2000. Schizophrenia as the price that Homo sapiens pays for language: a
476   resolution of the central paradox in the origin of the species. Brain Research
477   Reviews 31, 118–129.
478

479 Crow, T.J., 2008. The □big bang□ theory of the origin of psychosis and the faculty of
480   language. Schizophr. Res. 102, 31–52.
481

482 Demontis, D., Walters, R.K., Martin, J., Mattheisen, M., Als, T.D., Agerbo, E.,
483   Belliveau, R., Bybjerg-Grauholm, J., Bækved-Hansen, M., Cerrato, F., et al., 2017.
484   Discovery of the first genome-wide significant risk loci for ADHD. bioRxiv
485   145581.
486

487 Durbin, R., Abecasis, G., Bentley, D., Chakravarti, A., Clark, A., Donnelly, P.,
488   Eichler, E., Flicek, P., Gabriel, S., Gibbs, R., et al., 2012. An integrated map of
489   genetic variation from 1,092 human genomes. Nature 497, 56-65.
490

491 Ehret, G.B., Munroe, P.B., Rice, K.M., Bochud, M., Johnson, A.D., Chasman, D.I.,
492   Smith, A.V., Tobin, M.D., Verwoert, G.C., Hwang, S.-J., et al., 2011. Genetic
493   variants in novel pathways influence blood pressure and cardiovascular disease
494   risk. Nature 478, 103.
495

496 Gejman, P.V., Sanders, A.R., Kendler, K.S., 2011. Genetics of schizophrenia: new
497   findings and challenges. Annu. Rev. Genomics Hum. Genet. 12, 121–44.
498

499 Gittelman, R.M., Hun, E., Ay, F., Madeoy, J., Pennacchio, L., Noble, W.S., Hawkins,
500   R.D., Akey, J.M., 2015. Comprehensive identification and analysis of human
501   accelerated regulatory DNA. Genome Res. 25, 1245–1255.
502

503 Glazko, G.V., Nei, M., 2003. Estimation of divergence times for major lineages of
504   primate species. Mol. Biol. Evol. 20, 424–34.
505
506 Gokhman, D., Lavi, E., Prüfer, K., Fraga, M.F., Riancho, J.A., Kelso, J., Pääbo, S.,
507   Meshorer, E., Carmel, L., 2014. Reconstructing the DNA methylation maps of the
508   Neandertal and the Denisovan. Science 344, 523–527.
509
510 Green, R.E., Krause, J., Briggs, A.W., Maricic, T., Stenzel, U., Kircher, M., Patterson,
511   N., Li, H., Zhai, W., Fritz, M.H.-Y., et al., 2010. A draft sequence of the
512   Neandertal genome. Science 328, 710–722.
513
514 Hasegawa, M., Kishino, H., Yano, T., 1985. Dating of the human-ape splitting by a
515   molecular clock of mitochondrial DNA. J. Mol. Evol. 22, 160–74.
516
517 Hernando-Herraez, I., Heyn, H., Fernandez-Callejo, M., Vidal, E., Fernandez-Bellon,
518   H., Prado-Martinez, J., Sharp, A.J., Esteller, M., Marques-Bonet, T., 2015. The
519   interplay between DNA methylation and sequence divergence in recent human
520   evolution. Nucleic Acids Res. 43, 8204–8214.
521
522 Hernando-Herraez, I., Prado-Martinez, J., Garg, P., Fernandez-Callejo, M., Heyn, H.,
523   Hvilsom, C., Navarro, A., Esteller, M., Sharp, A.J., Marques-Bonet, T., 2013.
524   Dynamics of DNA methylation in recent human and great ape evolution. PLoS
525   Genet 9, e1003763.
526
527 Horton, R., Wilming, L., Rand, V., Lovering, R.C., Bruford, E.A., Khodiyar, V.K.,
528   Lush, M.J., Povey, S., Talbot Jr, C.C., Wright, M.W., et al., 2004. Gene map of the
529   extended human MHC. Nat. Rev. Genet. 5, 889-899.
530
531 Jeste, D.V., Del Carmen, R., Lohr, J.B., Wyatt, R.J., 1985. Did schizophrenia exist
532   before the eighteenth century? Compr. Psychiatry 26, 493–503.
533
Karolchik, D., Barber, G.P., Casper, J., Clawson, H., Cline, M.S., Diekhans, M.,
Dreszer, T.R., Fujita, P.A., Guruvadoo, L., Haeussler, M., Harte, R.A., Heitner, S.,
Hinrichs, A.S., Learned, K., Lee, B.T., Li, C.H., Raney, B.J., Rhead, B.,
Rosenbloom, K.R., Sloan, C.A., Speir, M.L., Zweig, A.S., Haussler, D., Kuhn,
R.M., Kent, W.J., 2014. The UCSC Genome Browser database: 2014 update.
Nucleic Acids Res. 42, D764–70.
540
541 Kromer, J., Hummel, T., Pietrowski, D., Giani, A., Sauter, J., Ehninger, G., Schmidt,
542   A., Croy, I., 2016. Influence of HLA on human partnership and sexual satisfaction.
543   Scientific Reports 6, 32550.
544
545 Kundaje, A., Meuleman, W., Ernst, J., Bilenky, M., Yen, A., Heravi-Moussavi, A.,
546   Kheradpour, P., Zhang, Z., Wang, J., Ziller, M.J., et al., 2015. Integrative analysis
547   of 111 reference human epigenomes. Nature 518, 317–330.
548
549 Kyaga, S., Lichtenstein, P., Boman, M., Hultman, C., Långstr□m, N., Landén, M.,
550   2011. Creativity and mental disorder: family study of 300 000 people with severe
551   mental disorder. Brit. J. Psych. 199, 373–379.
552

553 Lander, E.S., Linton, L.M., Birren, B., Nusbaum, C., Zody, M.C., Baldwin, J., Devon,
554   K., Dewar, K., Doyle, M., FitzHugh, W., et al., 2001. Initial sequencing and
555   analysis of the human genome. Nature 409, 860-921.
556
Langergraber, K.E., Prüfer, K., Rowney, C., Boesch, C., Crockford, C., Fawcett, K.,
Inoue, E., Inoue-Muruyama, M., Mitani, J.C., Muller, M.N., Robbins, M.M.,
Schubert, G., Stoinski, T.S., Viola, B., Watts, D., Wittig, R.M., Wrangham, R.W.,
Zuberbühler, K., Pääbo, S., Vigilant, L., 2012. Generation times in wild
chimpanzees and gorillas suggest earlier divergence times in great ape and human
evolution. Proc. Natl. Acad. Sci. U.S.A. 109, 15716–15721.
563
564 Larson, C.A., Nyman, G.E., 1973. Differential fertility in schizophrenia. Acta
565   Psychiatr Scand 49, 272–80.
566
567 Lee, P.H., O'Dushlaine, C., Thomas, B., Purcell, S.M., 2012. INRICH: interval-based
568   enrichment analysis for genome-wide association studies. Bioinformatics 28,
569   1797–9.
570
571 Lichtenstein, P., Yip, B.H., Bj□rk, C., Pawitan, Y., Cannon, T.D., Sullivan, P.F.,
572   Hultman, C.M., 2009. Common genetic determinants of schizophrenia and bipolar
573   disorder in Swedish families: a population-based study. Lancet 373, 234–239.
574
Lindblad-Toh, K., Garber, M., Zuk, O., Lin, M.F., Parker, B.J., Washietl, S.,
Kheradpour, P., Ernst, J., Jordan, G., Mauceli, E., Ward, L.D., Lowe, C.B.,
Holloway, A.K., Clamp, M., Gnerre, S., Alf□ldi, J., Beal, K., Chang, J., Clawson,
H., Cuff, J., Di Palma, F., Fitzgerald, S., Flicek, P., Guttman, M., Hubisz, M.J.,
Jaffe, D.B., Jungreis, I., Kent, W.J., Kostka, D., Lara, M., Martins, A.L.,
Massingham, T., Moltke, I., Raney, B.J., Rasmussen, M.D., Robinson, J., Stark, A.,
Vilella, A.J., Wen, J., Xie, X., Zody, M.C., Baldwin, J., Bloom, T., Whye Chin, C.,
Heiman, D., Nicol, R., Nusbaum, C., Young, S., Wilkinson, J., Worley, K.C.,
Kovar, C.L., Muzny, D.M., Gibbs, R. a., Cree, A., Dihn, H.H., Fowler, G.,
Jhangiani, S., Joshi, V., Lee, S., Lewis, L.R., Nazareth, L.V., Okwuonu, G.,
Santibanez, J., Warren, W.C., Mardis, E.R., Weinstock, G.M., Wilson, R.K.,
Delehaunty, K., Dooling, D., Fronik, C., Fulton, L., Fulton, B., Graves, T., Minx,
P., Sodergren, E., Birney, E., Margulies, E.H., Herrero, J., Green, E.D., Haussler,
D., Siepel, A., Goldman, N., Pollard, K.S., Pedersen, J.S., Lander, E.S., Kellis, M.,
2011. A high-resolution map of human evolutionary constraint using 29 mammals.
Nature 478, 476–482.
591
592 Locke, A.E., Kahali, B., Berndt, S.I., Justice, A.E., Pers, T.H., Day, F.R., Powell, C.,
593   Vedantam, S., Buchkovich, M.L., Yang, J., et al., 2015. Genetic studies of body
594   mass index yield new insights for obesity biology. Nature 518, 197–206.
595
596 Markow, T.A., 2012. Developmental instability: its origins and evolutionary
597   implications: proceedings of the International Conference on Developmental
598   Instability: Its Origins and Evolutionary Implications, Tempe, Arizona, 14-15 June
599   1993. Springer Science and Business Media.
600

Mendizabal, I., Keller, T., Zeng, J., Soojin, V.Y., 2014. Epigenetics and evolution. Integr. Comp. Biol. 54, 31–42.

Messias, E.L., Chen, C.-Y., Eaton, W.W., 2007. Epidemiology of schizophrenia: review of findings and myths. Psychiatr. Clin. North Am. 30, 323–38.

Meyer, D., Aguiar, V.R., Bitarello, B.D., Brandt, D.Y., Nunes, K., 2017. A genomic perspective on HLA evolution. Immunogenetics 70, 1–23.

Molaro, A., Hodges, E., Fang, F., Song, Q., McCombie, W.R., Hannon, G.J., Smith, A.D., 2011. Sperm methylation profiles reveal features of epigenetic inheritance and evolution in primates. Cell 146, 1029–1041.

Nichols, C., 2009. Is there an evolutionary advantage of schizophrenia? Pers. Individ. Diff. 46, 832–838.

Paaby, A.B., Rockman, M.V., 2014. Cryptic genetic variation: evolution's hidden substrate. Nat. Rev. Genet. 15, 247-258

Pollard, K.S., Salama, S.R., King, B., Kern, A.D., Dreszer, T., Katzman, S., Siepel, A., Pedersen, J.S., Bejerano, G., Baertsch, R., et al., 2006. Forces shaping the fastest evolving regions in the human genome. PLoS Genet 2, e168.

Potts, W.K., Wakeland, E.K., 1990. Evolution of diversity at the major histocompatibility complex. Trends Ecol. Evol. (Amst.) 5, 181–7.

Power, R.A., Kyaga, S., Uher, R., MacCabe, J.H., Långstr□m, N., Landen, M., McGuffin, P., Lewis, C.M., Lichtenstein, P., Svensson, A.C., 2013. Fecundity of patients with schizophrenia, autism, bipolar disorder, depression, anorexia nervosa, or substance abuse vs their unaffected siblings. JAMA Psychiatry 70, 22–30.

Rannala, B., Yang, Z., 2003. Bayes estimation of species divergence times and ancestral population sizes using DNA sequences from multiple loci. Genetics 164, 1645–1656.

R Core Team, 2017. R: A Language and Environment for Statistical Computing. Vienna, Austria.

Ripke, S., Neale, B.M., Corvin, A., Walters, J.T., Farh, K.-H., Holmans, P.A., Lee, P., Bulik-Sullivan, B., Collier, D.A., Huang, H., et al., 2014. Biological insights from 108 schizophrenia-associated genetic loci. Nature 511, 421-427.

Rivera, C.M., Ren, B., 2013. Mapping human epigenomes. Cell 155, 39–55.

Roberts, S.C., Gosling, L.M., Carter, V., Petrie, M., 2008. MHC-correlated odour preferences in humans and the use of oral contraceptives. Proc. Biol. Sci. 275, 2715–2722.

Santos, P.S.C., Schinemann, J.A., Gabardo, J., Bicalho, M. da G., 2005. New evidence that the MHC influences odor perception in humans: a study with 58

Southern Brazilian students. Horm. Behav. 47, 384–388.

Schork, A.J., Thompson, W.K., Pham, P., Torkamani, A., Roddey, J.C., Sullivan,
P.F., Kelsoe, J.R., O'Donovan, M.C., Furberg, H., Schork, N.J., et al., 2013. All
SNPs are not created equal: genome-wide association studies reveal a consistent
pattern of enrichment among functionally annotated SNPs. PLoS Genet 9,
e1003449.

Schork, A.J., Wang, Y., Thompson, W.K., Dale, A.M., Andreassen, O.A., 2016. New
statistical approaches exploit the polygenic architecture of schizophrenia—
implications for the underlying neurobiology. Curr. Opin.. Neurobiol. 36, 89–98.

Sekar, A., Bialas, A.R., de Rivera, H., Davis, A., Hammond, T.R., Kamitaki, N.,
Tooley, K., Presumey, J., Baum, M., Van Doren, V., et al., 2016. Schizophrenia
risk from complex variation of complement component 4. Nature 530, 177–183.

Sklar, P., Ripke, S., Scott, L.J., Andreassen, O.A., Cichon, S., Craddock, N.,
Edenberg, H.J., Nurnberger, J.I., Rietschel, M., Blackwood, D., et al., 2011. Large-
scale genome-wide association analysis of bipolar disorder identifies a new
susceptibility locus near ODZ4. Nature Genet. 43, 977–983.

Skre, I., Onstad, S., Torgersen, S., Lygren, S., Kringlen, E., 1993. A twin study of
DSM-III-R anxiety disorders. Acta Psychiatr. Scand. 88, 85–92.

Sniekers, S., Stringer, S., Watanabe, K., Jansen, P.R., Coleman, J.R., Krapohl, E.,
Taskesen, E., Hammerschlag, A.R., Okbay, A., Zabaneh, D., et al., 2017. Genome-
wide association meta-analysis of 78,308 individuals identifies new loci and genes
influencing human intelligence. Nature Genet. 49, 1107-1112.

Sommer, S., 2005. The importance of immune gene variability (MHC) in
evolutionary ecology and conservation. Front. Zool. 2, 16.

Srinivasan, S., Bettella, F., Hassani, S., Wang, Y., Witoelar, A., Schork, A.J.,
Thompson, W.K., Collier, D.A., Desikan, R.S., Melle, I., et al., 2017. Probing the
association between early evolutionary markers and schizophrenia. PloS One 12,
e0169227.

Srinivasan, S., Bettella, F., Mattingsdal, M., Wang, Y., Witoelar, A., Schork, A.J.,
Thompson, W.K., Zuber, V., Winsvold, B.S., Zwart, J.-A., et al., 2015. Genetic
markers of human evolution are enriched in schizophrenia. Biol. Psychiatry 80,
284-292.

Stahl, E.A., Raychaudhuri, S., Remmers, E.F., Xie, G., Eyre, S., Thomson, B.P., Li,
Y., Kurreeman, F.A., Zhernakova, A., Hinks, A., et al., 2010. Genome-wide
association study meta-analysis identifies seven new rheumatoid arthritis risk loci.
Nature Genet. 42, 508–514.

Sullivan, P.F., Kendler, K.S., Neale, M.C., 2003. Schizophrenia as a complex trait:
evidence from a meta-analysis of twin studies. Arch. Gen. Psychiatry 60, 1187–

1192.

Teslovich, T.M., Musunuru, K., Smith, A.V., Edmondson, A.C., Stylianou, I.M.,
Koseki, M., Pirruccello, J.P., Ripatti, S., Chasman, D.I., Willer, C.J., Johansen,
C.T., Fouchier, S.W., Isaacs, A., Peloso, G.M., Barbalic, M., Ricketts, S.L., Bis,
J.C., Aulchenko, Y.S., Thorleifsson, G., Feitosa, M.F., Chambers, J., Orho-
Melander, M., Melander, O., Johnson, T., Li, X., Guo, X., Li, M., Shin Cho, Y., Jin
Go, M., Jin Kim, Y., Lee, J.-Y., Park, T., Kim, K., Sim, X., Twee-Hee Ong, R.,
Croteau-Chonka, D.C., Lange, L.A., Smith, J.D., Song, K., Hua Zhao, J., Yuan, X.,
Luan, J., Lamina, C., Ziegler, A., Zhang, W., Zee, R.Y.L., Wright, A.F., Witteman,
J.C.M., Wilson, J.F., Willemsen, G., Wichmann, H.-E., Whitfield, J.B.,
Waterworth, D.M., Wareham, N.J., Waeber, G., Vollenweider, P., Voight, B.F.,
Vitart, V., Uitterlinden, A.G., Uda, M., Tuomilehto, J., Thompson, J.R., Tanaka,
T., Surakka, I., Stringham, H.M., Spector, T.D., Soranzo, N., Smit, J.H., Sinisalo,
J., Silander, K., Sijbrands, E.J.G., Scuteri, A., Scott, J., Schlessinger, D., Sanna, S.,
Salomaa, V., Saharinen, J., Sabatti, C., Ruokonen, A., Rudan, I., Rose, L.M.,
Roberts, R., Rieder, M., Psaty, B.M., Pramstaller, P.P., Pichler, I., Perola, M.,
Penninx, B.W.J.H., Pedersen, N.L., Pattaro, C., Parker, A.N., Pare, G., Oostra,
B.A., O'Donnell, C.J., Nieminen, M.S., Nickerson, D.A., Montgomery, G.W.,
Meitinger, T., McPherson, R., McCarthy, M.I., McArdle, W., Masson, D., Martin,
N.G., Marroni, F., Mangino, M., Magnusson, P.K.E., Lucas, G., Luben, R., Loos,
R.J.F., Lokki, M.-L., Lettre, G., Langenberg, C., Launer, L.J., Lakatta, E.G.,
Laaksonen, R., Kyvik, K.O., Kronenberg, F., König, I.R., Khaw, K.-T., Kaprio, J.,
Kaplan, L.M., Johansson, A., Jarvelin, M.-R., Janssens, A.C.J.W., Ingelsson, E.,
Igl, W., Kees Hovingh, G., Hottenga, J.-J., Hofman, A., Hicks, A.A.,
Hengstenberg, C., Heid, I.M., Hayward, C., Havulinna, A.S., Hastie, N.D., Harris,
T.B., Haritunians, T., Hall, A.S., Gyllensten, U., Guiducci, C., Groop, L.C.,
Gonzalez, E., Gieger, C., Freimer, N.B., Ferrucci, L., Erdmann, J., Elliott, P.,
Ejebe, K.G., Döring, A., Dominiczak, A.F., Demissie, S., Deloukas, P., de Geus,
E.J.C., de Faire, U., Crawford, G., Collins, F.S., Chen, Y.I., Caulfield, M.J.,
Campbell, H., Burtt, N.P., Bonnycastle, L.L., Boomsma, D.I., Boekholdt, S.M.,
Bergman, R.N., Barroso, I., Bandinelli, S., Ballantyne, C.M., Assimes, T.L.,
Quertermous, T., Altshuler, D., Seielstad, M., Wong, T.Y., Tai, E.-S., Feranil,
A.B., Kuzawa, C.W., Adair, L.S., Taylor, H.A., Borecki, I.B., Gabriel, S.B.,
Wilson, J.G., Holm, H., Thorsteinsdottir, U., Gudnason, V., Krauss, R.M., Mohlke,
K.L., Ordovas, J.M., Munroe, P.B., Kooner, J.S., Tall, A.R., Hegele, R.A.,
Kastelein, J.J.P., Schadt, E.E., Rotter, J.I., Boerwinkle, E., Strachan, D.P., Mooser,
V., Stefansson, K., Reilly, M.P., Samani, N.J., Schunkert, H., Cupples, L.A.,
Sandhu, M.S., Ridker, P.M., Rader, D.J., van Duijn, C.M., Peltonen, L., Abecasis,
G.R., Boehnke, M., Kathiresan, S., 2010. Biological, clinical and population
relevance of 95 loci for blood lipids. Nature 466, 707–713.

Traherne, J., 2008. Human MHC architecture and evolution: implications for disease
association studies. Int. J. Iimmunogenet. 35, 179–192.

Venter, J.C., Adams, M.D., Myers, E.W., Li, P.W., Mural, R.J., Sutton, G.G., Smith,
H.O., Yandell, M., Evans, C.A., Holt, R.A., et al., 2001. The sequence of the
human genome. Science 291, 1304–1351.

WHO, 1973. Report of the international pilot study of schizophrenia.
750
751 Winternitz, J., Abbate, J., Huchard, E., Havl'\i\vcek, J., Garamszegi, L., 2017.
752   Patterns of MHC-dependent mate selection in humans and nonhuman primates: a
753   meta-analysis. Molecular ecology 26, 668–688.
754
755 Wood, A.R., Esko, T., Yang, J., Vedantam, S., Pers, T.H., Gustafsson, S., Chu, A.Y.,
756   Estrada, K., Luan, J., Kutalik, Z., et al., 2014. Defining the role of common
757   variation in the genomic and biological architecture of adult human height. Nature
758   Genet 46, 1173–1186.
759
760 Xu, K., Schadt, E.E., Pollard, K.S., Roussos, P., Dudley, J.T., 2015. Genomic and
761   network patterns of schizophrenia genetic variation in human evolutionary
762   accelerated regions. Mol Biol Evol 32, 1148–1160.
763
764

Figure legends

Fig. 1: Enrichment plots of hypo-DMR and hyper-DMR SNPs in schizophrenia

Quantile-quantile (QQ) plots of GWAS SNPs for Schizophrenia (SCZ) with the

extended MHC region (chr6: 25-35Mb) unmasked (A) and masked (B). Expected -

$\log_{10} p$-values under the null hypothesis are shown on the X-axis. Observed -$\log_{10} p$-

values are on the Y-axis. The values for all GWAS SNPs are plotted in dark green

while the values for SNPs in linkage disequilibrium (LD) with hypo-methylated

DMRs are plotted in blue and SNPs in LD with hyper-methylated DMRs are plotted

in pink. A leftward deflection of the plotted p-values from the line for all GWAS

SNPs indicates enrichment of true signals – the greater the leftward deflection, the

stronger the enrichment. Genomic correction was performed on all SNPs with global

lambda.

Fig. 2: Enrichment plots of hypo-DMR and hyper-DMR SNPs across

multiple traits

Thirteen different GWASs were analysed using a common set of ~2.4 million

SNPs. The p-values for the common set of GWAS SNPs are plotted in dark

green; p-values for SNPs that tag hypo-methylated DMRs are plotted in blue;

and p-values for SNPs that tag hyper-methylated DMRs are plotted in pink.

ADHD, attention deficit hyperactivity disorder; BMI, body mass index; BPD,

bipolar disorder; DBP, diastolic blood pressure; HDL, high density lipoprotein;

LDL, low density lipoprotein; RA, rheumatoid arthritis; SBP, systolic blood

pressure; SCZ, schizophrenia; TC, total cholesterol; TG, triglycerides. The

MHC region was absent from the common set of SNPs.

Fig. 3: INRICH test for enrichment of association of DMR gene sets and NSS

genes with SCZ, MHC masked

A visual heatmap depicting *p*-values from bootstrapping with 5,000 iterations. The

various evolutionary annotations compared are as follows. HypoDMR – human hypo-

methylated DMRs; HyperDMR – human hyper-methylated DMRs. HypoDMR and

HyperDMR were taken from the study by Hernando-Herraez et al. (2013). dmrH –

human-specific DMRs (Gokhman et al, 2014), which are referred to as non-primate

DMRs in this manuscript. NSS - Neanderthal selective sweep. Datasets marked with *

have been previously reported by Banerjee et al. (2017) and are presented here for

comparison only.

Figure 1

Figure 2

Figure 3

Bootstrapped P-values

	Nominal p<= 1e-3	Nominal p<= 1e-4	Nominal p<= 1e-5	Nominal p<= 1e-6	Nominal p<= 1e-7	Nominal p<= 1e-8	
0.0046	0.04059	0.14377	0.05519	0.19456	0.23535	HypoDMR
0.41032	0.4927	0.9764	0.97001	0.59028	0.69606	HyperDMR
0.0046	0.07459	0.03039	0.04439	0.22216	0.06299	dmrH*
0.0024	0.06579	0.10438	0.18896	0.07179	0.71866	NSS*