Genome-wide haplotype association study in imaging genetics using whole-brain sulcal openings of 16,304 UK Biobank subjects.

Slim Karkar[*‡], Claire Dandine-Roulland[†], Jean-François Mangin[*]

Yann Le Guen[*], Cathy Philippe[*], Jean-François Deleuze[†]

Morgane Pierre-Jean[†], Edith Le Floch[†] and Vincent Frouin[*‡]

August 26, 2020

[*]Université Paris-Saclay, CEA, Neurospin, 91191, Gif-sur-Yvette, France.

[†]Université Paris-Saclay, CEA, Centre National de Recherche en Génomique Humaine, 91057, Evry, France

[‡]Corresponding authors : karkar.slim@gmail.com ; vincent.frouin@cea.fr

Running title   Haplotype association study for imaging-genetics

This manuscript is under review : 031-20-EJHGR

Abstract

Neuroimaging-genetics cohorts gather two types of data: brain imaging and genetic data. They allow the discovery of associations between genetic variants and brain imaging features. They are invaluable resources to study the influence of genetics and environment in the brain features variance observed in normal and pathological populations. This study presents a genome wide haplotype analysis for 123 brain sulcus opening value (a measure of sulcal width) across the whole brain that include 16,304 subjects from UK Biobank. Using genetic maps, we defined 119,548 blocks of low recombination rate distributed along the 22 autosomal chromosomes, and analyzed 1,051,316 haplotypes. To test associations between haplotypes and complex traits, we designed three statistical approaches. Two of them use a model that includes all the haplotypes for a sin gle block, while the last approach considers one model by haplotype. All the statistics produced were assessed as rigorously as possible. Thanks to the rich imaging dataset at hand, we used resampling techniques to assess False Positive Rate for each statistical approach in a genome-wide and brain-wide context. The results on real data show that genome-wide haplotype analyses are more sensitive than single- SNP approach and account for local complex Linkage Disequilibrium (LD) structure, which makes genome-wide haplotype analysis an interesting and statistically sound alternative to the single-SNP counterpart.

Keywords Imaging-genetics study ; Haplotype Association Analysis ; UK Biobank

1   Introduction

Numerous population-imaging studies have been built since the beginning of years 2000 based on earlier pioneering studies [1] to support researches mainly in vascular, neurodegenerative diseases or psychiatric syndromes [2] and now include genetics [3, 4, 5]. Our work is based on the UK Biobank resources [6, 7, 8], currently the most emblematic imaging-genetics cohort available as open-data. The UK Biobank cohort brings the unique opportunity to study the genetic and environmental dissection of numerous diseases or complex traits related to brain *via* imaging endophenotypes. These endophenotypes are structural or functional imaging-derived characteristics (or imaging-derived phenotypes - IDP [7] in UK Biobank) and DNA genotyping arrays provide the genetic measures [6]. In aging of normal and pathological brains, we observe the phenomenon of sulcal widening for which several related IDP like sulcal depth, sulcal opening and grey matter thickness in brain were shown by our group to be highly heritable [9]. In the case of grey matter thickness and sulcal opening, our group in [10] has also shown associations with several new markers using Genome-Wide Association Study (GWAS).

In order to yield robust inferences and interpretations from GWAS approaches, one should only consider the hits passing a strict genomic significance threshold [11] to account for the large number of tests. Moreover, in imaging-genetics studies, an additional correction for multiple phenotype testing is mandatory. After these corrections and in the case of a non-synonymous SNP hit, a straightforward interpretation can be carried out to hypothesise an association with the disease. In most cases, GWAS results are harder to interpret because the association is carried by a group of SNPs located in a non-coding region. In some cases, the grouping can be explained by a leading causal SNP which signal is spread amongst the neighbouring SNPs *via* Linkage Desequilibrium (LD). Using imputed SNPs, one can identify putative unmeasured causal variants. Nowadays more and more GWAS are using imputed SNPs to do fine mapping or allow replication study. An alternative to imputation is haplotype analysis that can also capture unmeasured variants.

One might also suspect that the association is actually carried by more than one SNP

in the same region. In the case of multiple causal SNPs, burden-test or collapsing test have shown a high power to detect small effect sizes [12]. A gain in sensitivity can also be obtained by considering the combinations of alleles from several genotyped SNPs in the form of haplotypes [13].

Haplotypes in one individual consist in the combinations of several SNP alleles to form nucleotide sequences. These sequences of alleles can be obtained from phased SNPs that define which alleles are on the same chromosome and inherited together from each parent.

Both phased SNPs and imputed variants are available in the UK Biobank resources, based on large reference panels (1000 Genomes and the Haplotype Reference Consortium (HRC), see [6]). The high quality of the phasing dataset (assessed using trios of parents-child) is an incentive to use this information in haplotype-based analyses. The method described in this paper differs from the commonly used haplotype-based approaches on several points. First it applies at the genome level. Second, the haplotype-based tests are based on phased data which is unusual: out of the 9 haplotype-based methods cited in [14], only one (WHaIT, [15]) uses phase information as input. Moreover our method identifies individual haplotype associations by testing all haplotypes in a given genomic interval, while the WHaIT method only performs a global test, i.e. it does not identify which individual haplotype in the block is associated with the phenotype. Also of note is that the WHaIT method is only designed for categorical traits.

In this work, we propose to push the genome-wide haplotype association approaches to fit the specificity of the IDP obtained in imaging-genetics, more precisely sulcal opening measurements derived from the UK Biobank high quality imaging data. The aim of the paper is three-fold. First, we present this set of quantitative IDP and a normalization of their distributions across the subjects. We show their spatial variability which reveals that IDP are many and varied, and which motivates our search for even more sensitive association methods to compensate for the multiple testing issue. Second, we detail three tests for genome-wide haplotype associations with the traits. This part includes defini-

tion of blocks along the genome based on a genetic map, which is a prerequisite to any haplotype-based test definition. Third, we compare association hits obtained from these three haplotype-based tests and also with the regular genome-wide association test based on single SNPs. Finally, we carried out phenotype permutations to check the validity of our tests on the whole genome.

2 Material and Method

2.1 Samples

The present analyses were conducted under UK Biobank data application number 25251. The UK Biobank is a health research resource including genotypic data of about 500,000 people aged between 45 and 73 years old, that were recruited in the general population across United Kingdom.

The UK Biobank is expected to provide multi-modal MR brain images in 100,000 participants and, by March 2018 it had 20,060 subjects with a T1-weighted MRI [8]. These data were processed locally through BrainVisa/Morphologist pipeline [16] and quality controlled yielding a set of labelled cortical sulci for each of 18,175 subjects (see [10] for details).

We relied on the Quality Control carried out by the UK Biobank consortium on the genotyping data which excluded individuals with high missingness, high heterozygosity, first degree related individuals or sex mismatches [6]. With $658,720$ phased SNPs (Haplotype dataset), $784,256$ genotyped SNPs and $93,095,623$ imputed variants, $16,304$ subjects passed the image processing and the genetic QC (48% of males and 52% of females).

2.2   Data Processing

Imaging data, sulcal opening For each subject, 123 labelled sulci were extracted from T1-weighted images. For each sulcus, a measure of its width - a feature called opening - is computed as the ratio of the volume of the Cerebrospinal Fluid the sulcus contains to the surface of the sulcus (Figure 1).

For each sulcus, after adjusting for age and sex using linear regression, we identified and excluded outliers in the residual distribution using the robust interquartile range (IQR) method [17].

Then, since the distributions of sulcal opening values could exhibit deviation from the normal distribution, we evaluated this deviation and normalized them using a one-parameter Box-Cox transformation (power transformation) [18], see Section Supplementary Material 1
. We selected the optimal power λ using goodness-of-fit for normal distribution (available in MASS R package) for each sulcus. The value of λ quantifies the similarity with a normal distribution. For a sulcus with an original normal distribution, $\lambda = 1$ and for a sulcus with an original log-normal distribution, $\lambda = 0$.

In the following, the Box-Cox transformations of the sulcal openings, $\tilde{Y}$, were considered as phenotypes.

Definition of Haplotypes blocks in UK Biobank When a group of SNPs are in linkage disequilibrium (LD), they are not inherited independently: the presence/absence of one allele at one locus depends on the alleles of neighbouring loci in LD. Using allele frequencies and LD information available in reference panels, phasing softwares like SHAPEIT [19] estimate for each chromosome the sequence of SNP alleles that have been inherited together from each parent.

Using this approach, the UK Biobank consortium has released a curated genome-wide dataset of $658,720$ phased SNPs over the 22 autosomal chromosomes [6]. From these phased data we define haplotypes, which are combinations of neighbouring SNP alleles on a single chromosome. We propose the following two-fold procedure to ensure

the correct haplotype estimation.

First, we determine haplotype blocks that includes adjacent SNPs in high LD or in other words, present a low recombination rate. We used the GRCH37 Rutgers genetic map [20] that includes positions in base pairs (bp) and centiMorgans (cM) based on the 1000 Genomes project. For variants of the UK Biobank arrays that are not present in the genetic map, the position in cM was estimated using linear interpolation based on the local recombination rate (in cM/bp) within the interval defined by the two closest neighbouring variants in the map. For first (resp. last) assessed variants on a chromosome that fall outside the genetic map, recombination rates were estimated by linear extrapolation on the whole chromosome. The complete genetic map allows to define non-overlapping haplotype blocks in which any two consecutive SNP loci are less distant than δ cM. A distance greater than δ cM defines the start of a new block. Second, using the phased SNPs dataset and the blocks described above, we determine the haplotypes.

To insure an homogeneous cover of the genome, we chose δ value equals to 0.001 cM, leading to 119,548 haplotype blocks across the whole genome (see Supplementary Material 4
 and [21])

2.3   Genome-wide haplotype association

2.3.1   Count matrix of each haplotype block

For each haplotype block obtained previously, we defined the reference haplotype h_0 as the most common one, and the *count matrix*, $\mathbf{H} = [\mathbf{h_1} \ldots \mathbf{h_m}] \in \{0, 1, 2\}^{(N \times m)}$, with N the sample size and m the number of alternative haplotypes denoted by $h_1, \ldots, h_m$. Each element $h_{i,j}$ of $\mathbf{H}$ corresponds to the number of copies of the alternative haplotype $h_j, 1 \leq j \leq m$ for the individual $i, 1 \leq i \leq N$. In this way, the count matrix $\mathbf{H}$ codes for an additive haplotype model - the effect of each haplotype block corresponds to the sum of the effects of observed alternative haplotypes. Section Supplementary Material 2

shows an example of matrix $\mathbf{H}$, obtained with 3 phased SNPs and 3 subjects.

2.3.2 Association tests

We defined three association tests between haplotypes h_j included in matrix $\mathbf{H}$, and the phenotypes $\tilde{\mathbf{Y}}$, the sulcal opening values obtained after Box-Cox transformation as described previously. In the following, $\mathbf{X}$ denotes the matrix of covariates containing age, sex and the first ten principal components provided by UK Biobank to account for population stratification. Although all presented methods are using similar statistics, the three tests are addressing three different questions.

Haplotype block model test In this test, referred as "block-test", the following linear model is considered:

$$\tilde{\mathbf{Y}} = \beta_0 + \mathbf{X}\beta + \mathbf{H}\gamma + \epsilon \tag{1}$$

with $\tilde{\mathbf{Y}}$ the phenotype vector, $\mathbf{X}$ the matrix of covariates and $\mathbf{H}$ the matrix of the m alternative haplotypes, included with fixed effects β and γ respectively, and ϵ the error vector. β_0 is the intercept containing the effect of the reference haplotype.

This first test assesses the association between a given block, defined as the set of its haplotypes, and each phenotype i.e. $H_0 : \gamma = \mathbf{0_m}$ versus $H_1 : \gamma_1 \neq 0$ or ... or $\gamma_m \neq 0$

The significance of the association is estimated using a total variance test for nested linear model.

Complete model haplotype test The second test is a complete model haplotype test and is referred as "complete-test". It aims to test the association between one phenotype and each haplotype h_j inside the block, while considering in the model the other haplotypes of the block. For this purpose, we used the same linear model regression as in the block-test (eq. 1) and we considered for each haplotype h_j, $1 \leq j \leq m$, the null hypothesis : $H_0 : \gamma_j = 0$ versus $H_1 : \gamma_j \neq 0$. For this test, we computed a two-sided p-value of the t-statistic.

Single haplotype model test To test the effect of each haplotype versus the others,

the third test referred as "single-test" is based on the following linear model:

$$\tilde{\mathbf{Y}} = \beta_0 + \mathbf{X}\beta + \mathbf{h_j}\gamma_j + \epsilon \qquad \text{for } 0 \leq j \leq m \tag{2}$$

with $\tilde{\mathbf{Y}}$ the phenotype vector, $\mathbf{X}$ the matrix of covariates and $\mathbf{h_j}$ the count vector for the haplotype h_j, included with fixed effects β and γ_j respectively, and ϵ the error vector. β_0 is the intercept containing the mean effect of all haplotypes excluding the considered haplotype j. As previously, we computed a standard two-sided p-value of the t-statistic for the null hypothesis $H_0 : \gamma_j = 0$ versus $H_1 : \gamma_j \neq 0$.

All three tests were implemented in R v.3.4.0 to run on a distributed HPC (code available upon request).

2.3.3 Comparison with single-SNP association used in classical GWAS

To our knowledge, a comprehensive power analysis of haplotype association for quantitative traits on large cohorts (more than 10,000 subjects) has not been done. To gain insight into the power of the haplotype association tests, we propose a comparative study, where we match the results of haplotype-based tests with the results from the single SNP association used in classical GWAS considering both genotyped and imputed SNPs. The test used in classical GWAS using linear model in PLINK [22] is described in Section Supplementary Material 3

.

2.3.4 False Positive Rates in the Genome-wide haplotype association tests

In order to evaluate the validity of the different haplotype association tests scrutinized in this study, we constructed datasets under the null hypothesis using permuted phenotypes. Then, we computed the p-values distributions and the associated False Positive Rate (FPR). The FPR is the percentage of tests that show a significant association in the permuted dataset at the significance threshold with the Bonferroni correction. We fixed

the significance threshold to $\alpha_{N_t} = \alpha/N_t$, with $\alpha = 0.05$, and N_t is the number of hypotheses tested.

Correlations between variables are present in genetic and imaging datasets; between parts of the genome and within the brain respectively. To quantify the impact of these correlations, we computed the FPR under null hypothesis for several scenarios that keep or not the correlation structures among haplotype blocks and among phenotypes.

False Positive Rate while preserving the correlation within each haplotype block In the first permutation scenario, a phenotype is randomly permuted for each block. Therefore, correlations between haplotype blocks across the genome present in the original dataset were removed in this scenario, preserving only the correlation within each haplotype block. The complete-test and the single-test produced several statistics per haplotype block (depending on the number of haplotypes), while the block-test produced a unique statistic per block. We computed $FPR = N_{FP}/N_{t,p}$, where N_{FP} is the observed number of false positives and $N_{t,p}$ is the number of statistics produced by each of the three tests for the permuted phenotype p.

This permutation scenario was replicated on three phenotypes of sulcal opening: two that represent the range of λ values and one for the highest association found in previous study [10], with λ values ≈ 0.2, ≈ 0.8 and ≈ 0.6 respectively. The same permutation order was applied to the three phenotypes.

False Positive Rate while preserving the correlation within the genome In the second permutation scenario, we used all the phenotypes and produced a permutated dataset where the phenotypes were randomly shuffled with the same permutation for all haplotype blocks along the genome. For each of the three tests, we considered all the statistics produced and computed $FPR_p = N_{FP}/N_{t,p}$ where $N_{t,p}$ is the number of statistics produced for a given permutated phenotype p. This analysis produced three FPRs per permuted phenotype that preserves the structure of the correlation within the genome. Reporting FPRs for each phenotype will enable the detection of phenotypes with low quality measurements.

This procedure was replicated ten times. To reduce the computational burden in this case, we considered the residuals $\tilde{\mathbf{Y}}^* = \tilde{\mathbf{Y}} - (\hat{\beta}_0 + \mathbf{X}\hat{\beta})$, that is, the phenotypes $\tilde{\mathbf{Y}}$ adjusted for the covariates $\mathbf{X}$.

False Positive Rate while preserving the correlation within the genome and among the phenotypes Last, we used the above described permutated dataset and, instead of considering each phenotype independently, we pooled the results across the phenotypes. For each permutation and each haplotype test, we considered all the statistics produced. We computed $FPR_T = N_{FP}/N_T$ where N_T is the number of statistics produced by each test *for all permutated phenotypes*. This analysis produced three FPRs per permutation, that preserve both the structure of the correlation within the genome as well as the correlation among phenotypes.

3 Results

In this part, we present results obtained on the original data : the opening measurements of 123 brain sulci and genome-wide haplotypes of $16,304$ subjects. First, we present the effects of normalization of the phenotype distributions, and the characteristics of the haplotypes generated with the original data. Second, we present the significant association hits obtained with our genome-wide haplotype association study and provide the genomic location and length of significant haplotype blocks. Finally, we compare the differences in sensitivity on original data for the three tests proposed above and we propose a comparison study with the single-SNP association test.

A study of the False Positive Rate is also presented, where we study the FPR for the three models under the null hypothesis, using permutated datasets as described in Section 2.3.4.

Imaging data processing The distributions of sulcal opening measurements across the brain exhibit a large range of density shapes from quasi-normal distributions (max. $\lambda = 0.8$) to log-normal distributions (min $\lambda = 0.2$). Fig. 1 displays the λ parameter associated

with the opening distribution for each sulcus across the brain.

We observed a pattern, with larger (in terms of median value) and more normally distributed sulcal opening measurements in parietal and occipital lobes. Conversely, we observed narrower, more log-normally distributed opening values for lower and inferior temporal sulci. This could be related to the differential aging rate of brain structures, such as the widening sulci rates that might differ across the brain [23].

Genomic data processing Across the 22 autosomal chromosomes, we tested over 1 million haplotypes from 119,548 blocks for association with the opening of 123 sulci. Candidate haplotypes had a length, on the genetic map, ranging from 1.4×10^{-8} to 1.7×10^{-2} cM (see Fig. 2, right panel), with a median value of 6×10^{-4} cM, corresponding to consecutive SNP runs of length ranging from 2 to 64. Left panel of Fig. 2 shows the coverage of the blocks along the 22 autosomal chromosomes. Details regarding the distribution can be found in the Section Supplementary Material 4
. Of note is that the variation of the block length across the chromosomes remains small as it can be seen from the curve on the right of Fig 2 (similar distributions).

Haplotype tests exhibit different sensitivity compared to single SNP-based associations We carried out a comparative study of the discovery power for the three haplotype tests. To do so, all significant phenotype-haplotype associations for the three tests were matched with the conventional SNP association results. In Fig. 3 top panel, the hits obtained with the three tests were matched with genotyped SNPs and with the imputed SNPs in the bottom panel of Fig. 3. For each significant phenotype-haplotype association, we used the lowest p-value of single-SNP associations observed in the SNPs of that haplotype block. For the complete-test and the single-test, if more than one haplotype per block was significantly associated, we selected the one with the lowest p-value. When comparing results among the three tests, we found that:

(i) the block-test, that combines all haplotypes present in one block, leads to the lowest p-values for the strongest associations. However it appears globally more conservative

than the single-SNP approach, particularly for associations in a range close to the genomic threshold (10^{-8} to 10^{-10}). The negative intercept of the fitted linear trend (blue lines in Fig. 3) indicates that, in the significant blocks, p-values from this test are on average less significant than their single-SNP counterparts.

(ii) the complete-test appears to be the most sensitive, with p-values on average being 3-fold lower than their single-SNP counterparts.

(iii) the single-test performs similarly to the single-SNP approach, with p-values being slightly closer to those obtained with genotyped SNPs ($R^2 = 0.952$) compared to those obtained with imputed SNPs ($R^2 = 0.944$). Indeed, a large number of haplotypes are differentiated from the other haplotypes by a single SNP in the same block, and thus are expected to carry the same signal.

In Table 1 we used three available SNP datasets: (a; genotyped SNPs) : SNPs called from the UK-Biobank arrays; (b; Imputed SNPs): imputed SNPs with a MAF>0.01 see [6]; (c; Rare SNPs) : imputed SNPs with MAF<0.01. Similarly to Fig. 3, in order to match a given haplotype hit with a single-SNP hit in Table 1, we used the lowest p-value of single-SNP associations observed in the SNPs of that significant haplotype block.

The comparative analysis given in Table 1 revealed the following salient points :

(i) Haplotype association tests have similar or better detection power than imputed or genotyped SNP counterparts. Indeed, in the majority of hits, there was at least one haplotype model that was more significant than the single-SNP test (18 vs. 11). Moreover, for 9 phenotype/genotype significant associations, the haplotype p-values were at least 10-fold lower than their single-SNP counterparts while the opposite was true only once.

(ii) For strong associations involving several haplotypes in the same block, haplotype approaches can lead to p-values that are several order lower, for example, haplotypes association in block chr1:215146807-215154276 with FCMpost Left leads to p-values in the range of 10^{-24} compared to 10^{-19} with single-SNP associations. Additionally, for 6 other phenotypes, p-values obtained using haplotype associations are also 2 or 3 orders of magnitude lower.

(iii) Imputed rare variants do not exhibit any significant associations. Indeed, the smallest p-values associated with rare variants are 6 to 20 orders of magnitude higher than the significant haplotype hits.

False Positive Rates in the Genome-wide haplotype association tests   Table 2 shows an estimation, for each test, of $FPR_T = N_{FP}/N_t$, the FPR for the null distributions that preserve the correlation among phenotypes and within the genome (third scenario described in Section 2.3.4). The Q-Q plots presented in Fig. 4 indicate that in this permutation scenario, we do not observe an inflation of p-values under the null hypothesis. When using Bonferroni correction for significance thresholds (see 2.3.4), none of the three tests seem to show a systematic inflation of FPR (see Table 2). In Section Supplementary Material 5
, we investigated FPR in the two other scenarios of null hypotheses as described in Section 2.3.4 and reported results in Fig 3 (for scenario 1) and Table 1 (for scenario 2). We show that in the second permutation scenario, out of the two phenotypes (SPoCsup_left and SRh_left) showing systematic inflated FPR, none of them were involved in significant hits on real data. In Fig 3, we did not observed any inflation of FPR under the first permutation scenario. In addition, we did not observed any False Positive.

4   Discussion

Box-Cox transformation normalizes phenotypes across brain We used Box-Cox transformation of sulcal opening values to ensure a spatially homogeneous distribution of the opening measurements across the brain. The spatial distribution of the λ parameter across the brain highlights the necessity of such a transformation. This procedure ensures a homogeneous specificity of the analysis across the brain sulci. Indeed, the spatial pattern in the λ values does not reflect the spatial distribution of the FPR. In the analysis on permuted datasets, we observed an overall homogeneous behaviour across the brain for the three haplotype tests (see Section 1).

Choosing δ to define blocks We set δ parameter to the most conservative values among the ones we studied. The objective is to ensure the reliability of the phase along the block. We observed that the distribution of the block genetic length $L[h]$ for this value is the most homogeneous across the different chromosomes. Section Supplementary Material 4

shows that with δ increasing from 0.001 to 0.025, the distribution becomes bi-modal and less homogeneous across the chromosomes. With higher values of δ, $L[h]$ exhibits heterogeneous distributions depending on the chromosome. Moreover, with a higher δ value, a larger number of recombination is allowed within the same haplotype block which may lead to a lower sensitivity.

Sulcal opening patterns are associated with haplotypes In this study we show significant associations of sulcal opening measurements with haplotypes in a genome-wide approach; bringing about genome wide inferences that differ from a local haplotype fine-mapping operation performed after a classical genome wide single-SNP analysis. Along with the previous findings in the upstream region of KCNK2 (chr1:215MB) [10], the haplotype block test revealed 4 other hits on chromosomes 7, 9, 12 and 16, detailed in Table 3. For the following genes, located in these regions, that are predominantly expressed in the brain: LPAR1 (chr9); NUAK1 (chr12), lrcRNA RP11-178L8.8 and FBXO31 (chr16), we provide an overview of expression data, publicly available in Supplementary Material 6

. We also display brain regions associated with each hit in Supplementary Material 7

. Except for chr9:114, all hits reported on Table 1 are associated bilaterally with sulci such as (FCMpost_L, FCMpost_R) for chr1:215, (STsterascant_R, STsterascpost_L) for chr12:106 and (SFinfant_L, SFinfant_R) for chr16:87.

Among the bilateral associations, the hit chr7:134 stands out: its associated brain areas - the insula (left and right) - are not associated with any other genome region. Further studies could be motivated to investigate this genomic region since the insula is a cortical region known to integrates emotional, cognitive, and motivational signals, and that hap-

lotypes of the block chr7:13441632-134416604 are located in a region carrying several genes and markers associated with Autism Spectrum disorders ([24, 25, 26], see details in Table 3).

Multiple testing inside a block For complete-test and single-test, the number of independent tests probably lies between the number of blocks and the number of haplotypes. For example, a block could contain a main, very common haplotype, and several rare haplotypes associated with low-frequency SNPs. In such a case, it is likely that the association p-values of the different haplotypes within the same block will be dependent. The bump in the Q-Q plots in Fig. 4 (middle and right panels) could come from the sample with the main population carrying the main haplotype, and only a few individuals driving an association for the minor haplotypes. In a similar situation where many haplotypes exhibit the same structure in the population, Howard et al. [27] proposed a correction using an estimation of the number of independent haplotypes, which would be less stringent than the Bonferroni correction.

Disentangling multiple neighbouring single SNP hits using haplotypes Out of the three haplotype tests studied, the complete model haplotype test turned out to be the most sensitive overall. However, results show that in the presence of a complex LD background, like in the region upstream of KCNK2, the haplotype modelling underlying the block-test model produces p-values several orders of magnitude lower. This result calls for the development of innovative multivariate approaches based on haplotypes, such as those the authors have started to propose [28].

Limitations of the study

We choose to use a low δ value in order to obtain the most reliable haplotypes. One drawback is that this also allows higher LD between the blocks, meaning a greater correlation between blocks than with higher δ values. Using Bonferroni correction for multiple testing is expected to be conservative in the presence of such dependencies.

In our study of FPR, we purposely focus on realistic null distributions of measurements, that exhibit the true correlation *between all the phenotypes and along the whole*

genome. To fully control the FPR, one would need to run a larger number of runs which is computationally intensive and out of reach using our realistic null distributions. However, with the number of permutations used, we could reasonably detect when a large inflation of FPR occurs.

By comparing with single-SNPs analyses, we did not directly estimate the discovery power of our tests. To our knowledge, for whole-genome haplotype associations, there is no practical way to create a synthetic and realistic ground truth signal for multiple correlated phenotypes.

Conclusion and future works In the context of imaging-genetics, we studied three haplotype tests to find associations between haplotypes and quantitative traits measured on brain MRI. Beyond the process of extracting features from the imaging-genetics data, we normalized the measurement distributions of the sulcal openings using Box-Cox transformations and obtained spatially homogeneous distributions across the brain. We compared three haplotype association tests and achieved the best performance with the haplotype block model test (lowest p-values observed) and the complete model haplotype test (best sensitivity). These two tests use a multivariate model of haplotypes that account for the whole set of haplotypes present in a given block. We show that these tests outperformed the single-SNP based classical approach for haplotype blocks with association hits close to the genomic threshold (10^{-8} to 10^{-10}), specifically in the case of haplotype blocks that exhibit complex LD structures. Based on the results of this work, we do not argue for the systematic use of haplotype modelling over GWAS based on imputed variants. More particularly, in the case of the UK Biobank, all samples come from a homogeneous population with an extensive imputation panel [6]. In this case, the discovery power of haplotypes is often in the same range as that of imputed SNPs. However, in the context of imaging-genetics, we show that using haplotypes showed more reliable results: significant haplotypes were associated with a larger number of phenotypes, increasing the explanatory value of each discovery. This study relies on the definition of blocks using a genetic map and a single value of δ. Future works could take a step further and locally

define block boundaries in order to find the more relevant ones in terms of association [29, 30].

Conflict of interest Authors declare no conflict of interest.

References

[1] Vermeer SE, Koudstaal PJ, Oudkerk M, Hofman A, Breteler MM. Prevalence and risk factors of silent brain infarcts in the population-based Rotterdam Scan Study. Stroke. 2002 Jan;33(1):21–25.

[2] de Wilde A, van Maurik IS, Kunneman M, Bouwman F, Zwan M, Willemse EAJ, et al. Alzheimer's biomarkers in daily practice (ABIDE) project: Rationale and design. "Alzheimer's & Dementia: Diagnosis, Assessment & Disease Monitoring". 2017;6:143 – 151.

[3] Mueller SG, Weiner MW, Thal LJ, Petersen RC, Jack C, Jagust W, et al. The Alzheimer's disease neuroimaging initiative. Neuroimaging Clin N Am. 2005 Nov;15(4):869–877.

[4] Schumann G, Loth E, Banaschewski T, Barbot A, Barker G, Buchel C, et al. The IMAGEN study: reinforcement-related behaviour in normal brain function and psychopathology. Mol Psychiatry. 2010 Dec;15(12):1128–1139.

[5] Loth E, Charman T, Mason L, Tillmann J, Jones EJH, Wooldridge C, et al. The EU-AIMS Longitudinal European Autism Project (LEAP): design and methodologies to identify and validate stratification biomarkers for autism spectrum disorders. Mol Autism. 2017;8:24.

[6] Bycroft C, Freeman C, Petkova D, Band G, Elliott LT, Sharp K, et al. The UK Biobank resource with deep phenotyping and genomic data. Nature. 2018 10;562(7726):203–209.

[7] Elliott LT, Sharp K, Alfaro-Almagro F, Shi S, Miller KL, Douaud G, et al. Genome-wide association studies of brain imaging phenotypes in UK Biobank. Nature. 2018 10;562(7726):210–216.

[8] Alfaro-Almagro F, Jenkinson M, Bangerter NK, Andersson JLR, Griffanti L, Douaud G, et al. Image processing and Quality Control for the first 10,000 brain imaging datasets from UK Biobank. NeuroImage. 2018;166:400 – 424.

[9] Le Guen Y, Leroy F, Auzias G, Riviere D, Grigis A, Mangin JF, et al. The chaotic morphology of the left superior temporal sulcus is genetically constrained. NeuroImage. 2018;174(March):297–307. Publisher: Elsevier Ltd.

[10] Le Guen Y, Philippe C, Riviere D, Lemaitre H, Grigis A, Fischer C, et al. eQTL of KCNK2 regionally influences the brain sulcal widening: evidence from 15,597 UK Biobank participants with neuroimaging data. Brain Structure and Function. 2019 Mar;224(2):847–857.

[11] Spencer CC, Su Z, Donnelly P, Marchini J. Designing genome-wide association studies: sample size, power, imputation, and the choice of genotyping chip. PLoS Genet. 2009 May;5(5):e1000477.

[12] Ionita-Laza I, Lee S, Makarov V, Buxbaum JD, Lin X. Sequence kernel association tests for the combined effect of rare and common variants. Am J Hum Genet. 2013 Jun;92(6):841–853.

[13] Lin WY, Yi N, Zhi D, Zhang K, Gao G, Tiwari HK, et al. Haplotype-based methods for detecting uncommon causal variants with common SNPs. Genet Epidemiol. 2012 Sep;36(6):572–582.

[14] Datta AS, Biswas S. Comparison of haplotype-based statistical tests for disease association with rare and common variants. Briefings in Bioinformatics. 2015 09;17(4):657–671.

[15] Li Y, Byrnes AE, Li M. To identify associations with rare variants, just WHaIT: Weighted haplotype and imputation-based tests. Am J Hum Genet. 2010 Nov;87(5):728–735.

[16] Rivière D, Geffroy D, Denghien I, Souedet N, Cointepas Y. BrainVISA: an extensible software environment for sharing multimodal neuroimaging data and processing tools. NeuroImage. 2009 Jul;47:S163. Publisher: Academic Press.

[17] Mcgill R, Tukey JW, Larsen WA. Variations of Box Plots. The American Statistician. 1978;32(1):12–16.

[18] Box GEP, Cox DR. An analysis of transformations. Journal of the Royal Statistical Society Series B (Methodological. 1964;p. 211–252.

[19] Delaneau O, Marchini J, Zagury JF. A linear complexity phasing method for thousands of genomes. Nat Methods. 2011 Dec;9(2):179–181.

[20] Matise TC, Chen F, Chen W, De La Vega FM, Hansen M, He C, et al. A second-generation combined linkage physical map of the human genome. Genome Res. 2007 Dec;17(12):1783–1786.

[21] Karkar S, Le Guen Y, Philippe C, Dandine-Roulland C, Pierre-Jean M, Mangin JF, et al. A study of feasibility for genome-wide haplotype association of complex traits in imaging genetics. In: 2018 IEEE International Conference on Bioinformatics and Biomedicine (BIBM); 2018. p. 2764–2766.

[22] Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D, et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. Am J Hum Genet. 2007 Sep;81(3):559–575.

[23] Alexander-Bloch AF, Mathias SR, Fox PT, Olvera RL, Goring HHH, Duggirala R, et al. Human Cortical Thickness Organized into Genetically-determined Communities across Spatial Resolutions. Cereb Cortex. 2019 Jan;29(1):106–118.

[24] S Beyer K, M Klauck S, Wiemann S, Poustka A. Construction of a physical map of an autism susceptibility region in 7q32.3-q33. Gene. 2001 08;272:85–91.

[25] International Molecular Genetic Study of Autism Consortium I. A Full Genome Screen for Autism with Evidence for Linkage to a Region on Chromosome 7q. Human Molecular Genetics. 1998 03;7(3):571–578.

[26] Shao Y, Wolpert CM, Raiford KL, Menold MM, Donnelly SL, Ravan SA, et al. Genomic screen and follow-up analysis for autistic disorder. Am J Med Genet. 2002 Jan;114(1):99–105.

[27] Howard DM, Hall LS, Hafferty JD, Zeng Y, Adams MJ, Clarke TKK, et al. Genome-wide haplotype-based association analysis of major depressive disorder in Generation Scotland and UK Biobank. Translational Psychiatry. 2017 Nov;7(11):1263. Publisher: Springer US ISBN: 4139801700.

[28] Karkar S, Gloaguen A, Le Guen Y, Pierre-Jean M, Dandine-Roulland C, Le Floch E, et al. Multivariate Haplotype Analysis Of 96 Sulci Opening For 15,612 UK-Biobank Sujects. In: 2019 IEEE 16th International Symposium on Biomedical Imaging (ISBI 2019); 2019. p. 392–395.

[29] Guinot F, Szafranski M, Ambroise C, Samson F. Learning the optimal scale for GWAS through hierarchical SNP aggregation. BMC Bioinformatics. 2018 Dec;19(1):459. Publisher: BioMed Central.

[30] Balliu B, Houwing-Duistermaat JJ, Bohringer S. Powerful testing via hierarchical linkage disequilibrium in haplotype association studies. Biom J. 2019 May;61(3):747–768.

[31] Courchet J, Lewis TL, Lee S, Courchet V, Liou DY, Aizawa S, et al. Terminal axon branching is regulated by the LKB1-NUAK1 kinase pathway via presynaptic mitochondrial capture. Cell. 2013 Jun;153(7):1510–1525.

[32] Stankoff B, Barron S, Allard J, Barbin G, Noël F, Aigrot MS, et al. Oligodendroglial Expression of Edg-2 Receptor: Developmental Analysis and Pharmacological Responses to Lysophosphatidic Acid. Molecular and Cellular Neuroscience. 2002;20(3):415–428.

Figure legends

Figure 1: In a sample of $16,304$ subjects, sulcal opening measurements have different distributions across the brain, from normal distribution (red, $\lambda = 1$) to log-normal (blue, $\lambda = 0$). (Top) Distributions of uncorrected sulcal opening measurements, for 123 sulci, coloured according to λ values. Inlet panels are histograms of opening values for two sulci with fitted curves superimposed: (left) Anterior occipito-temporal sulcal opening distribution is almost log-normal (blue line); (right) Superior parietal sulcal opening distribution is almost normal (red line). (Bottom) Spatial map of λ values for sulci of the left hemisphere, coloured according to the λ value of their distribution. From left to right: top view, interior view, parietal view and bottom view.

Figure 2: (Left) Genome coverage of haplotype blocks, in percent per Mbp. For each window of 1 Mbp, we computed the total length in bp covered by the 119,912 haplotype blocks (see section 2.2 for details).(Right) Distributions of $L[h]$ (block length in cM), as defined by the genetic distance between bordering markers of each block. Blocks were defined over the 22 autosomal chromosomes using $\delta = 0.001cM$, a parameter that account for the maximal length in cM between two consecutive markers in a same haplotype (see 2.2 and [21] for details). The distribution of block lengths is homogeneous across the 22 autosomal chromosomes.

Figure 3: Comparison for main hits using the three proposed haplotype association tests (Y-axis) and single-SNP association (PLINK, X-Axis) with : (top) genotyped SNPs and (bottom) imputed SNPs. For each significant haplotype-phenotype association (Y-axis), the lowest p-value observed for the SNPs inside the block was reported (X-axis). For points above the grey line at x=y, p-values for haplotype associations test are lower than their single-SNPs counterpart. Blue lines indicates the fit of haplotype block test p-values ($y = 1.13x - 2.67$) with imputed SNPs and ($y = 1.18x - 2.94$) with genotyped SNPs ; Red lines indicates the fit of complete model haplotype test p-values ($y = 0.845x + 2.05$) with imputed SNPs and ($y = 0.908x - 1.5$) with genotyped SNPs; Yellow lines indicates the fit for single haplotype test p-values ($y = 0.877x + 1.16$) with imputed SNPs and ($y = 0.974x + 0.202$) with genotyped SNPs

Figure 4: Aggregation of QQ-plots for 10 runs of permutated datasets (see 2.3.4, third scenario). From left to right: block-test, complete-test and single-test. The shaded area define the hull (minimum and maximum values) of the Q-Q plots for the 10 runs and the average Q-Q plot is given by the coloured, dotted line. The black, dashed line at $y = x$ indicates where the expected distribution lies.

Tables

Hit Region (Mbp)	Start bp	End bp	Length (Kbp)	Phenotype	p-value hapotypes (all tests)	p-value genotyped SNPs	p-value imputed SNPs	p-value rare SNPs
chr1:215	215146807	215154276	7.469	FCMpost_L	**4.96E-24**	2.01E-18	1.16E-19	3.91E-03
				FCMpost_R	**9.33E-22**	4.01E-19	1.19E-20	8.06E-04
				FIP_L	**7.29E-12**	9.29E-10	9.48E-10	7.39E-03
				FIP_R	**2.26E-13**	4.43E-12	4.02E-12	1.68E-02
				FPO_R	**3.82E-14**	2.65E-13	2.96E-13	5.60E-03
				SC_R	**2.65E-13**	2.51E-11	2.56E-11	1.01E-02
				SFint_L	**4.62E-10**	6.37E-09	1.95E-09	7.91E-04
				SFsup_R	**9.98E-11**	9.25E-09	1.32E-08	5.18E-03
				SPaint_L	**5.30E-13**	5.25E-10	4.18E-10	7.22E-03
				SpC_L	**7.52E-15**	8.83E-12	6.56E-12	6.91E-03
				SsP_L	**2.75E-14**	1.79E-11	2.62E-11	4.71E-03
chr7:134	134416326	134416604	0.278	INSULA_L	**8.32E-10**	7.14E-09	5.11E-09	7.94E-01
				INSULA_R	3.36E-10	6.34E-10	**3.33E-10**	2.65E-01
chr9:114	113659382	113665419	6.037	FCMpost_R	2.74E-10	1.58E-10	**1.28E-10**	4.96E-03
				SPaint_L	**2.27E-10**	2.94E-10	2.30E-10	9.50E-03
chr12:106	106476140	106477376	1.236	FCLp_R	**7.52E-13**	1.17E-12	8.29E-13	4.80E-02
				SFint_R	**6.55E-14**	3.71E-13	9.28E-14	1.74E-02
				SPeCinter_L	**1.03E-10**	2.04E-10	1.12E-10	8.85E-02
				STsterascant_R	1.92E-11	**5.08E-12**	1.29E-11	1.76E-02
				STsterascpost_L	**1.44E-10**	1.86E-10	1.91E-10	1.93E-02
chr16:87	87226206	87229190	2.984	SFinf_L	2.03E-13	1.22E-13	**6.38E-14**	4.75E-02
				SFinfant_L	2.51E-10	**1.37E-10**	1.98E-10	2.82E-02
				SFinfant_R	5.61E-17	**9.44E-18**	3.45E-17	2.23E-02
				SFinter_R	1.10E-11	8.08E-12	**3.44E-13**	2.66E-02
				SFmarginal_L	**9.73E-11**	2.17E-10	3.44E-10	1.82E-02
	87245155	87246608	1.453	SFinf_L	1.31E-12	**7.75E-13**	8.66E-13	9.10E-02
	87254180	87257820	3.64	SFinf_L	8.92E-11	5.36E-12	**3.70E-12**	3.49E-02
				SFinfant_L	4.30E-11	1.19E-11	**9.77E-12**	1.28E-02
				SFinfant_R	4.15E-11	1.77E-11	**9.09E-12**	3.30E-02

Table 1: Significant hits found in haplotype blocks. For each block, the lowest p-value over the 3 haplotype tests is compared with the lowest p-value in the block for genotyped SNPs, imputed common SNPs (MAF > 1%) and rare SNPs (MAF < 1%). For each row, the lowest p-value is highlighted in bold. All positions are given for GRCh37/hg19 assembly.

Test	Average number of tests per run	Average discovery threshold $\alpha' = \alpha/N_T$	Average FPR	Number of run with FPR $\geq 1/N_T$	Min FPR	Max FPR
Block	14,673,826	3.4E-09	6.8E-9	1	0	6.8E-8
Complete	107,683,292	4.6E-10	1.8E-9	1	0	9.2E-09
Single	124,951,506	4.0E-10	1.6E-9	1	0	8.0E-09

Table 2: False Positive Rate under null hypothesis for 10 runs of permutations using the third scenario (see section 2.3.4). For each run of permutation, tests were corrected by the Bonferroni procedure $\alpha' = \alpha/N_T$ for N_T tests with risk $\alpha = 0.05$ i.e. the FPR should be lower than $1/N_T$ in 5% of the runs.

Haplotype position	Gene / Region position	Description
chr1: 215146807-215154276	upstream of KCNK2	eQTL of KCNK2 [10]
chr7: 13441632-134416604	chr7: 134346283-134446935	autism susceptibility region 7q32.3-q33 [24]
	chr7: 129202420-138543686	AUTS1 region : D7S530 to D7S684 [25, 26]
chr9: 113659382-113665419	chr9: 113635543-113800981 (intronic)	LPAR1 : Homo sapiens lysophosphatidic acid receptor 1, transcript variant 2, mRNA. Protein, most expressed in Brain - Spinal cord (cervical c-1) (GTEx V7)
chr12: 106476140-106477376	chr12: 106457118-106533811 (intron, exon 3 & 4)	NUAK1 : Homo sapiens NUAK family, SNF1-like kinase, 1, mRNA. Protein, most expressed in Brain - Frontal Cortex (BA9) (GTEx V7), involved in axon branching [31]
chr16: 87226206-87257820	chr16: 87360593-87361190 (102kpb, upstream)	RP11-178L8.8 (AC010531.7): novel transcript, antisense to FBXO31 and C16orf95 readthrough LncRNA, most expressed in Brain - Cerebellum(GTEx V6)
	chr16: 87360593-87422364 (102kpb, downstream)	FBXO31 : F-box protein 31,mRNA. Protein, most expressed in Brain - Cerebellum(GTEx V8)

Table 3: Genes or regions of interest related to haplotypes significantly associated with the opening of at least one sulcus. All positions are given for GRCh37/hg19 assembly.
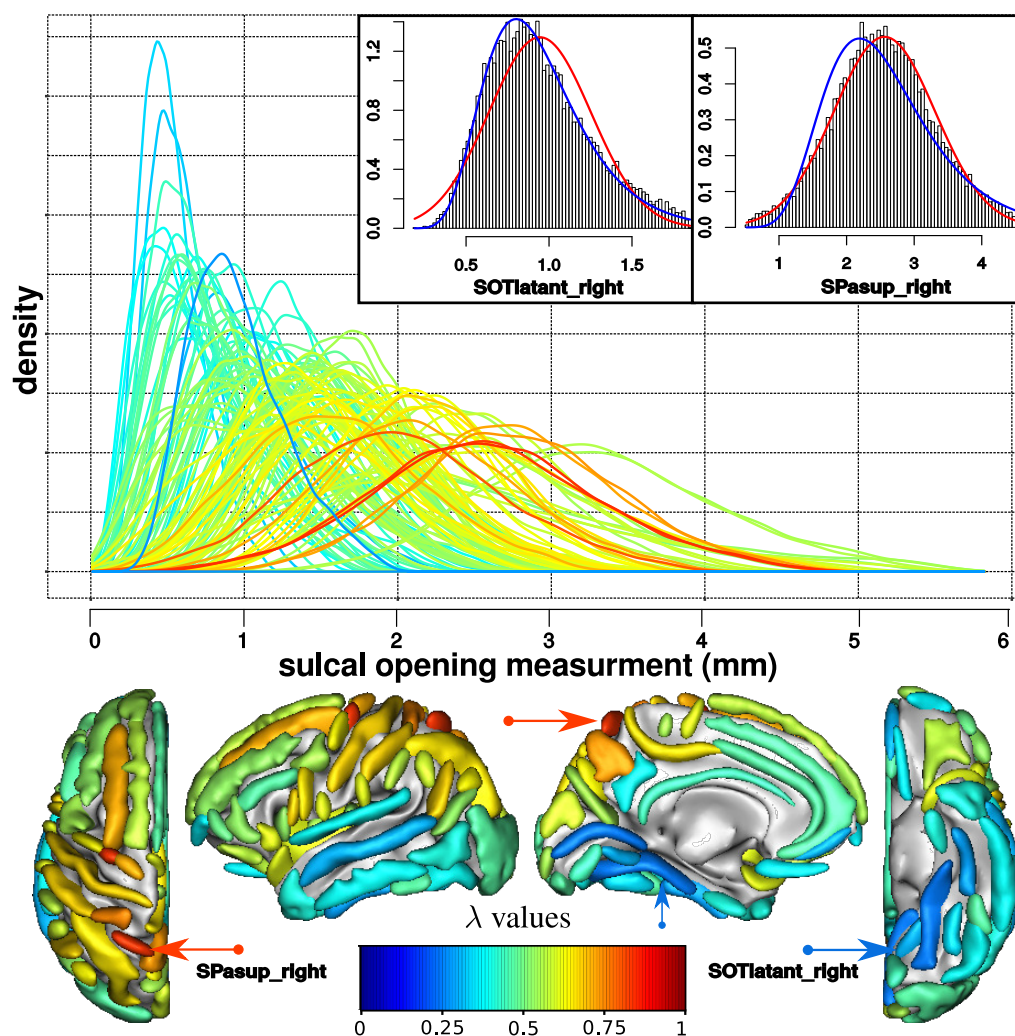
Figures

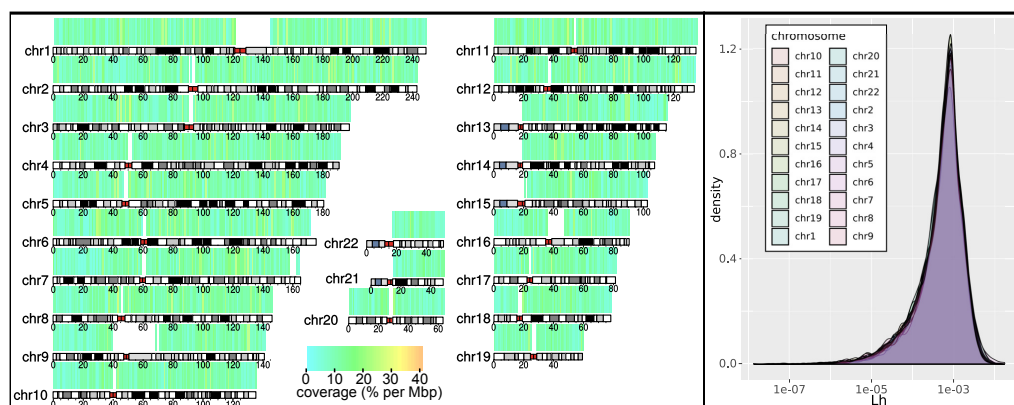Figure 1: Distributions of sulcal opening measurements of $16,304$ subjects

Figure 2: (Left) Genome coverage of haplotype blocks and (Right) Distribution of $L[h]$, the length of haplotype blocks in cM
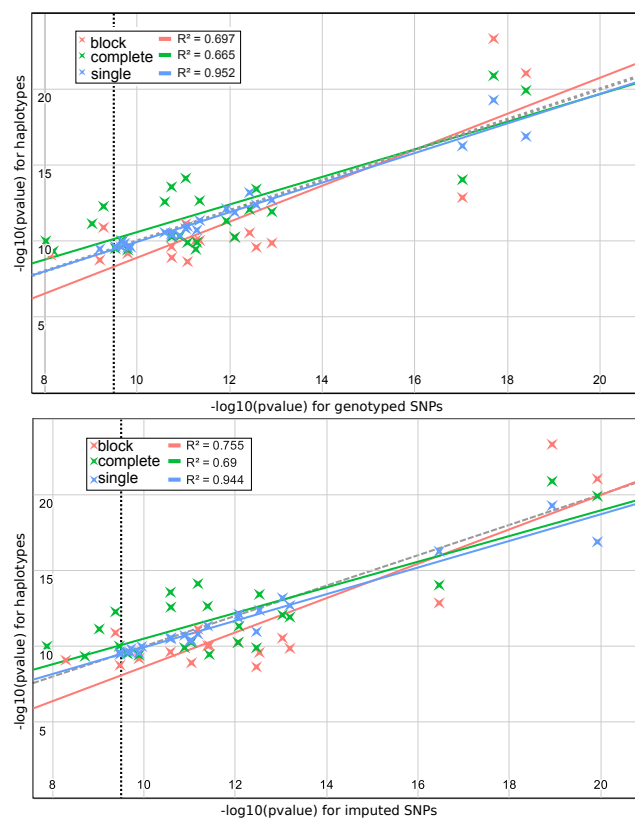
Figure 3: Comparison of p-values between main hits detected with haplotype association tests and their single-SNP approach counterpart.
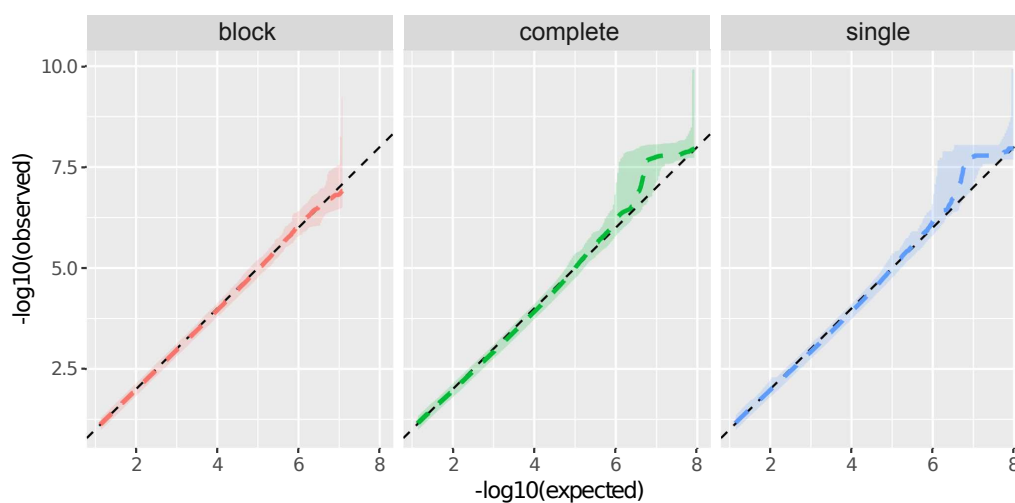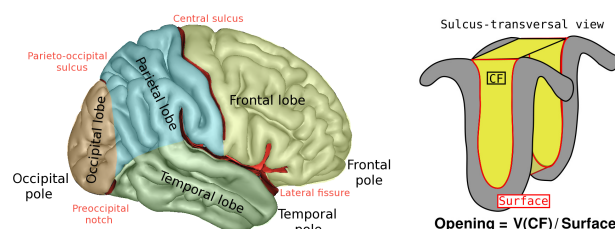
Figure 4: Aggregation of QQ-plots for 10 runs of permutated datasets

Supplementary Material 1   Box-Cox transformation of Sulcal Opening

Supplementary Figure 1: (left) Main brain structures. Sulci are the main furrows. Here only two sulci are labelled. For example, Central Sulcus is separating Frontal and Parietal lobes (Blausen Medical Encyclopedia). (right) Sulcus opening, a measure of the average sulcus width, is computed as the ratio of the volume of Cerebrospinal Fluid to the surface of the sulcus.

For each subject, 123 labelled sulci were extracted from T1-weighted images. For each sulcus, a measure of its width - a feature called opening - is computed as the ratio of the volume of the Cerebrospinal Fluid the sulcus contains to its surface. For each sulcus, after adjusting for age and sex using linear regression, we identified and excluded outliers in the residual distribution using the robust interquartile range (IQR) method [17]. The distributions of sulcal opening values could exhibit deviation from the normal distribution. We evaluated these deviations and normalized them using a one-parameter Box-Cox transformation (power transformation) : Let $\mathbf{Y}$ and $\tilde{\mathbf{Y}}$ be the vectors of the initial and Box-Cox transformed sulcal opening values respectively. The one-parameter Box-Cox transformation is the following:

$$\tilde{\mathbf{Y}} = \begin{cases} \frac{\mathbf{Y}^{\lambda}-1}{\lambda} & if \quad \lambda \neq 0 \\ log(\mathbf{Y}) & if \quad \lambda = 0 \end{cases}$$

Supplementary Material 2   Haplotype count matrix

The following shows an example of Haplotype count matrix $\mathbf{H}$, obtained with 3 phased SNPs and 3 subjects: the matrix $\mathbf{H}$ accounts for the number of copies, per subject, of the two observed alternative haplotypes $h_1 = [101]$ and $h_2 = [010]$. In our example, subject S_1 is homozygous for haplotype h_0 and subjects S_2 and S_3 are heterozygous with haplotypes h_1 and h_2, and haplotypes h_0 and h_2 respectively.

$$
\begin{array}{c}
\begin{array}{ccc} Snp_1 & Snp_2 & Snp_3 \end{array} \\
\begin{array}{c} S_{1,1} \\ S_{1,2} \\ S_{2,1} \\ S_{2,2} \\ S_{3,1} \\ S_{3,2} \end{array}
\left[\begin{array}{ccc}
0 & 0 & 0 \\
0 & 0 & 0 \\
1 & 0 & 1 \\
0 & 1 & 0 \\
0 & 0 & 0 \\
0 & 1 & 0
\end{array}\right] \\
\underbrace{\qquad\qquad\qquad}_{} \\
\text{UK Biobank} \\
\text{haplotype dataset}
\end{array}
\longrightarrow
\begin{array}{c}
\begin{array}{ccc} h_0 & h_1 & h_2 \end{array} \\
\begin{array}{c} S_1 \\ S_2 \\ S_3 \end{array}
\left[\begin{array}{ccc}
2 & 0 & 0 \\
0 & 1 & 1 \\
1 & 0 & 1
\end{array}\right] \\
\underbrace{\qquad}_{\mathbf{H}}
\end{array}
\qquad (3)
$$

Supplementary Material 3   Univariate single-SNP test classically used in GWAS
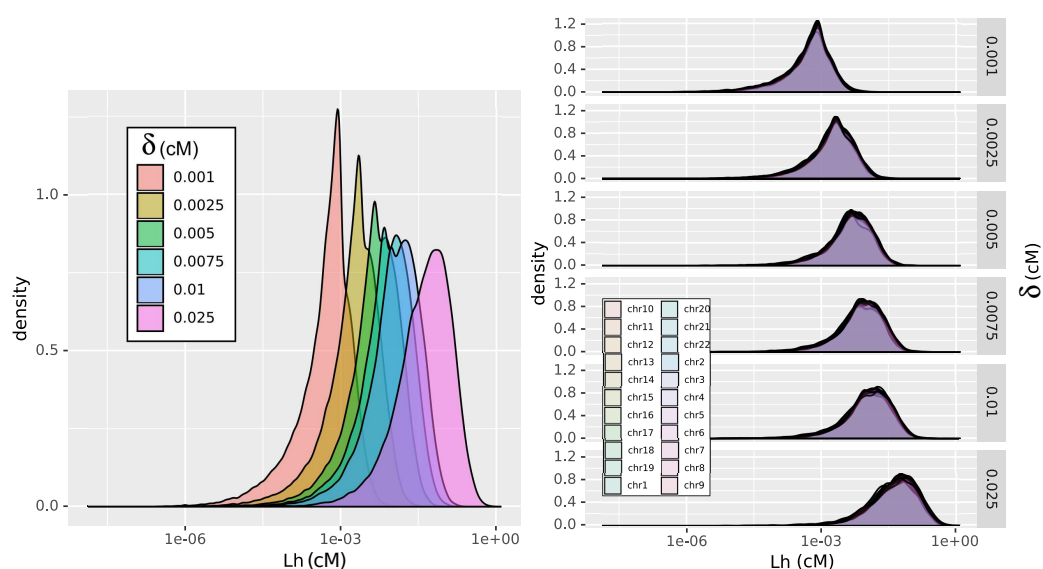
The test for each of the p SNPs in classical GWAS for quantitative traits uses the following linear model :

$$\tilde{\mathbf{Y}} = \mathbf{X}\beta + \mathbf{s_j}\omega_j + \epsilon \qquad \text{for } 1 \leq j \leq p \tag{4}$$

with $\tilde{\mathbf{Y}}$ the phenotype vector, $\mathbf{X}$ the matrix of covariates and $\mathbf{s_j}$ the count vector of minor alleles for the SNP s_j, included with fixed effects β and ω_j respectively, and ϵ the error vector. Using PLINK, we computed a standard two-sided p-value of the t-statistic to test the following null hypothesis :
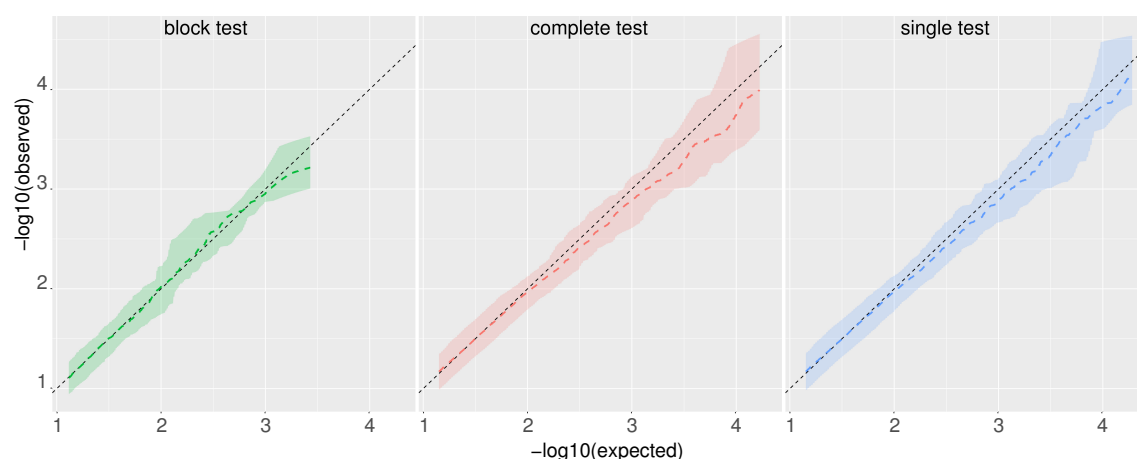
$$H_0 : \omega_j = 0 \qquad \text{vs} \qquad H_1 : \omega_j \neq 0$$

Supplementary Material 4   Length of haplotype blocks along the genetic map

Supplementary Figure 2: Distribution of haplotype block length, in cM for increasing value of parameter δ

Supplementary Material 5   False Positive Rate under null hypothesis

Supplementary Figure 3: Aggregated QQ-plots under the null-hypothesis while preserving the correlation within the haplotype blocks (see section 2.3.4 first scenario, for details) using block-test, complete-test and single-test. We selected three phenotypes : two that represent the range of λ values (see 2.2 for details) and one for the most significant association found with the original dataset. For each test, we computed a QQ-plot for each of the three phenotypes. The shaded area define the hull (minimum and maximum values) of the Q-Q plots for the 3 phenotypes and the average Q-Q plot is given by the coloured, dotted line. The black, dashed line at $y = x$ indicates where the expected distribution lies. For the three tests and the three phenotypes, the procedure did not produce any False Positive using $\alpha = 0.05$ and Bonferroni correction.

Phenotype	Block	Complete	Single
FCalant_ScCal_left	3	5	0
FCalant_ScCal_right	0	0	0
FCLa_left	0	0	0
FCLa_right	0	1	2
FCLp_left	0	0	0
FCLp_right	0	2	0
FCLrant_left	0	0	0
FCLrant_right	0	0	0
FCLrasc_left	0	0	0
FCLrasc_right	0	0	0
FCLrdiag_left	0	0	0
FCLrdiag_right	2	0	0
FCLrretroCtr_left	0	0	0
FCLrretroCtr_right	0	0	0
FCLrscant_left	1	0	1
FCLrscant_right	0	0	0
FCLrscpost_left	0	1	0
FCLrscpost_right	0	0	2
FCMant_left	0	0	0
FCMant_right	2	2	0
FCMpost_left	1	0	5
FCMpost_right	0	0	0
FColl_left	0	1	0
FColl_right	0	0	0
FIP_left	1	0	1
FIP_right	0	0	1
FIPPoCinf_left	0	0	0
FIPPoCinf_right	0	1	0
FIPrint1_left	0	0	0
FIPrint1_right	0	0	0
FIPrint2_left	0	0	0
FIPrint2_right	1	0	0
FPO_left	0	2	0
FPO_right	0	0	0
INSULA_left	0	0	1
INSULA_right	0	1	1
OCCIPITAL_left	0	0	0
OCCIPITAL_right	0	0	0
SC_left	0	1	1
SC_right	0	0	1
SCall_left	0	0	0
SCall_right	0	0	0
SCLPC_left	0	1	0
SCLPC_right	2	1	0
SCsylvian_left	0	0	0
SCsylvian_right	0	0	0
SCu_left	0	0	0
SCu_right	0	0	1
SFinf_left	0	0	0
SFinf_right	0	1	0
SFinfant_left	0	0	0
SFinfant_right	0	0	0
SFint_left	0	0	1
SFint_right	1	1	1
SFinter_left	0	0	0
SFinter_right	0	0	0
SFmarginal_left	1	2	1
SFmarginal_right	0	0	0
SFmedian_left	1	1	1
SFmedian_right	1	0	1
SForbitaire_left	0	0	0
SForbitaire_right	0	0	0
SFpolairetr_left	0	1	0
SFpolairetr_right	0	0	0

Phenotype	Block	Complete	Single
SFsup_left	0	0	0
SFsup_right	1	0	0
SGSM_left	0	0	0
SLiant_left	0	1	0
SLiant_right	0	0	0
SLipost_left	0	0	0
SLipost_right	1	0	0
SOlf_left	0	0	1
SOlf_right	1	0	0
SOp_left	0	0	0
SOp_right	0	0	2
SOr_left	1	0	0
SOr_right	2	0	2
SOTlatant_left	1	1	2
SOTlatant_right	2	0	0
SOTlatint_left	0	0	0
SOTlatint_right	1	2	1
SOTlatmed_left	1	0	0
SOTlatmed_right	0	2	0
SOTlatpost_left	0	1	0
SOTlatpost_right	2	1	0
SPaint_left	1	0	0
SPaint_right	0	0	0
SPasup_left	2	0	0
SPasup_right	1	2	1
SPat_left	2	0	1
SPat_right	1	1	1
SpC_left	0	0	0
SpC_right	1	0	0
SPeCinf_left	2	0	0
SPeCinf_right	0	0	0
SPeCinter_left	0	0	0
SPeCinter_right	0	0	0
SPeCmarginal_left	1	0	0
SPeCmarginal_right	0	0	0
SPeCmedian_left	0	0	0
SPeCmedian_right	0	0	0
SPeCsup_left	0	0	0
SPeCsup_right	0	0	0
SPoCsup_left	1	0	**10**
SPoCsup_right	0	0	3
SRh_left	**8**	**10**	0
SRh_right	3	3	0
SRinf_left	0	0	1
SRinf_right	1	1	1
SsP_left	0	1	0
SsP_right	0	0	1
STiant_left	0	0	0
STiant_right	2	1	1
STipost_left	0	0	0
STipost_right	0	0	1
STpol_left	0	0	0
STpol_right	0	0	0
STs_left	1	0	2
STs_right	2	1	1
STsterascant_left	0	0	0
STsterascant_right	3	0	1
STsterascpost_left	0	2	1
STsterascpost_right	0	1	0

	Block	Complete	Single
Phenotypes where >0 run with (FPR$\geq 1/N_t$)	38	33	35
Phenotypes where >1 run with (FPR$\geq 1/N_t$)	15	11	9
Phenotypes where >2 run with (FPR$\geq 1/N_t$)	4	3	3
Phenotypes where >5 run with (FPR$\geq 1/N_t$)	1	1	1
Frequency of runs with FPR$\geq 1/N_t$	0.05040650	0.04552846	0.04552846

Supplementary Table 1: Study of the False Positive Rate under the null hypothesis for 10 runs of permutation using the second scenario (see section 2.3.4 for details). P-values are corrected using Bonferroni (α/N_t) for N_t hypotheses, and discovery threshold set to $\alpha = 0.05$. The top table shows the number of runs (out of 10) where we observed an estimated FPR $\geq 1/N_t$. The bottom table aggregates these counts per test, and shows that, when we pool the FPR results of all permuted phenotypes, we seem to control the family-wise error rate at 5% under the global null hypothesis for each of the three tests. Two phenotypes raise concern regarding inflation of FPR : SPoCsup_left and SRh_left but were not found significantly associated with any haplotype in the real dataset.

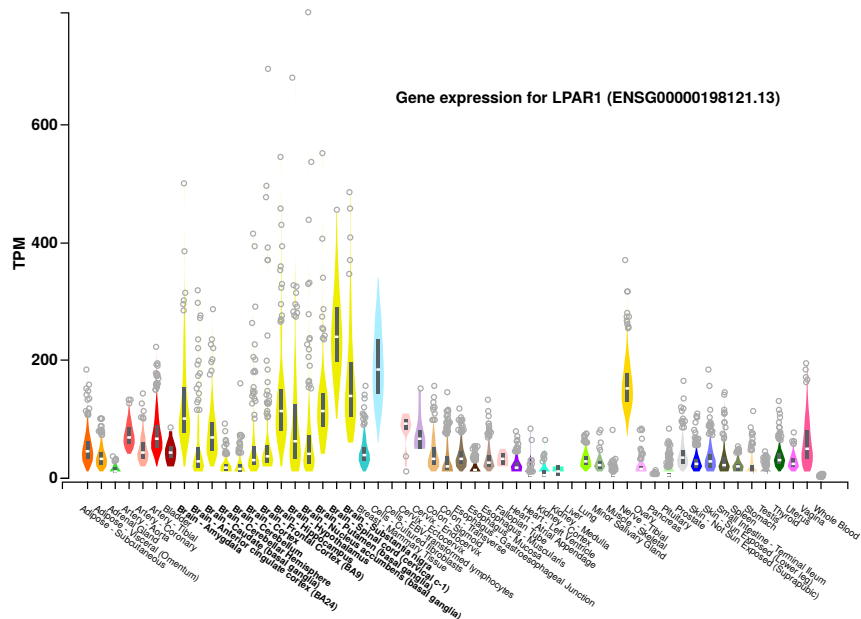# Supplementary Material 6   Tissue expression of gene related to significant hits

The hit located on chr.7q is found only associated to Insula (left and right) openning. The insula is a cortical region known to integrates emotional, cognitive, and motivational signals. Haplotypes of the block chr7:13441632-134416604 are located in a genomic region comprising several genes and markers associated with Autism Spectrum disorders (see Table 3). This hit could carry a signal distinct from the others because it is the only single phenotype (bilateral) - single haplotype block significant association.

Haplotype found on chr9:113 are located within an intronic region of LPAR1 gene, which is encoding for a protein used in cell signalling, notably in inhibition of neuroblastoma cell differentiation. Stankoff *et al.* [32] showed that transcripts are not detected during early stages of oligodendroglial development, but are expressed only in mature oligodendrocytes, shortly before the onset of myelination. Transcript are expressed in different brain tissues in GTEx v8 (see Fig 4)
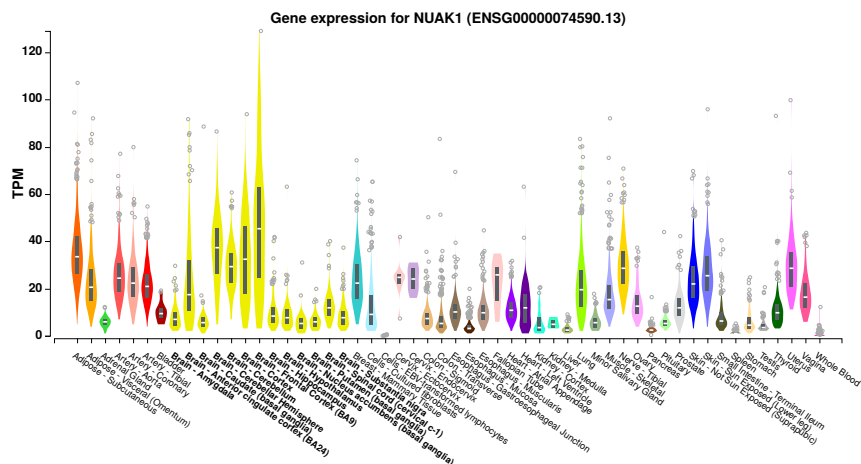
Haplotypes in block on chr.12:106476140-106477376 are located within the gene NUAK1, covering large intronic parts as well as 2 small exons. NUAK1 codes for a protein involved in cell proliferation, for example in arborization of mammalian neurons, regulating axon branching [31]. Particularly, it is expressed in several brain tissues (see Fig. 5).
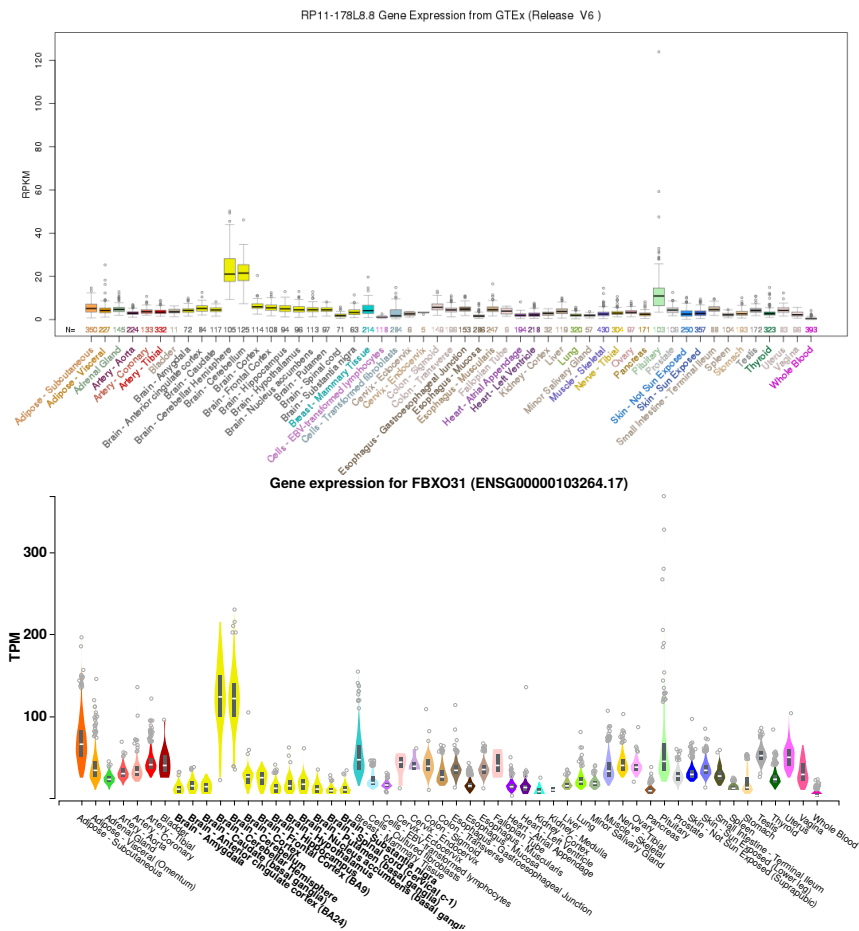
Haplotypes found in the chr16:87Mbp region are associated with several sulci located on both sides of the frontal region, in particular the Inferior Frontal Sulcus (SFinf). Located within gene C16orf95, haplotypes are 100kbp upstream of RP11-178L8.8 (or AC010531.7), a non-coding lncRNA mostly expressed in the brain (GTEx v6). This locus is antisens to FXBO31, a gene predominently expressed in Brain tissues in GTEx v8 (see Fig. 6).

Supplementary Figure 4: Expression of LPAR1 gene in different tissues of the brain (GTEx v8), related to top hit on chromosome 9.
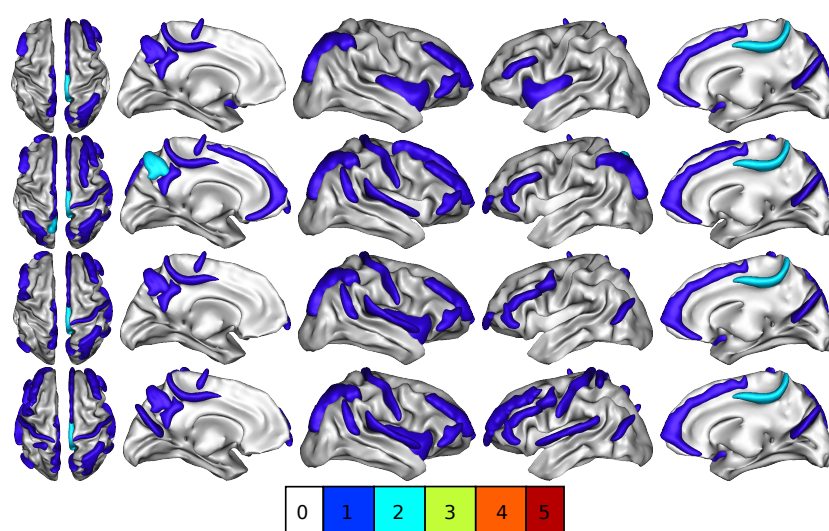

Supplementary Figure 5: Expression of NUAK1gene in different tissues of the brain (GTEx v8), related to top hit for chromosome 12.
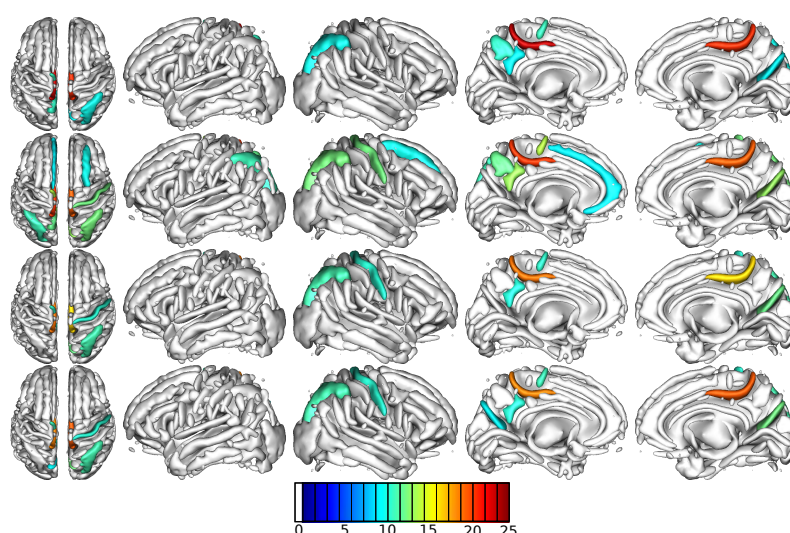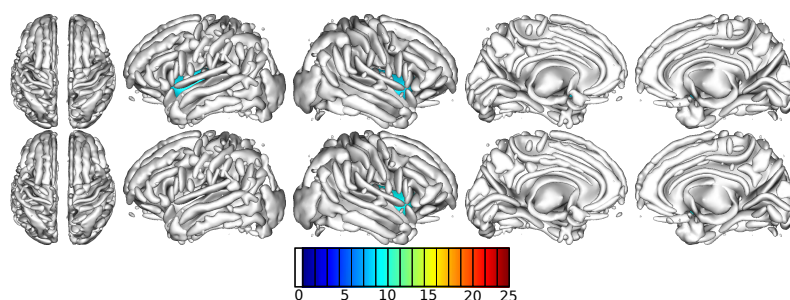
Supplementary Figure 6: (Top) Expression of antisens lrcRNA RP11-178L8.8 in different tissues of the brain (GTEx v6), 102kbp downstream of top hit for chromosome 16 (Bottom) Expression of FBXO31 gene in different tissues of the brain (GTEx v8), 102kbp upstream of top hit for chromosome 16.

Supplementary Material 7 Sulci of which openning is significantly associated with at least one genomic region
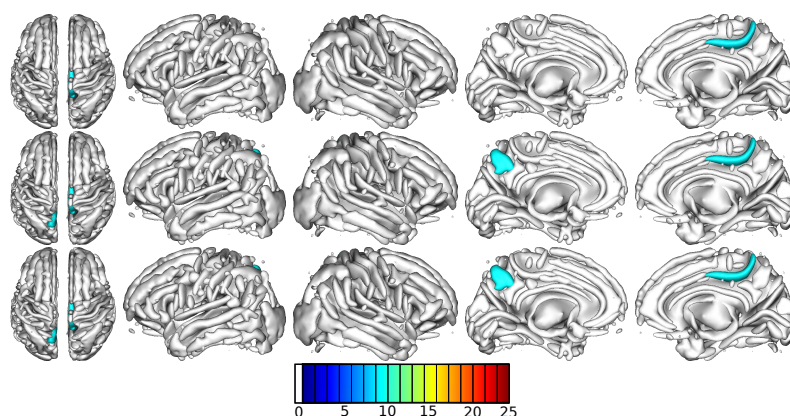

Supplementary Figure 7: Brain sulci coloured according to the number of significant hits. Hits - genomic region of 1 Mbp with at least one significant p-value among the three tests. Sulci colored in dark blue color are associated with one hit, light blue are associated with 2 hits (chr1:215 and chr9:114). In rows, from top to bottom, hits are determined using block-test, complete-test, single-test and imputed SNPs with PLINK..
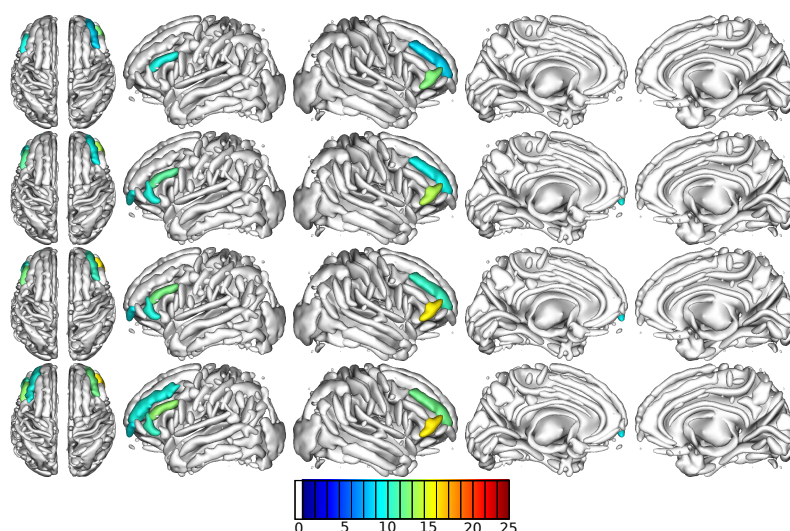

Supplementary Figure 8: Brain sulci coloured according to $-log10(p)$ of significant hits in the chr1:215 region. Hits - genomic region of 1 Mpb with at least one significant p-value among the three tests. In rows, from top to bottom, hits are determined using block-test, complete-test, single-test, imputed SNPs with PLINK.

Supplementary Figure 9: Brain sulci coloured according to $-log10(p)$ of significant hits in the chr7:134 region. Hits - genomic region of 1 Mpb with at least one significant p-value among the three tests. In rows, from top to bottom, hits are determined using block-test and single-test.

Supplementary Figure 10: Brain sulci coloured according to $-log10(p)$ of significant hits in the chr9:114 region. Hits - genomic region of 1 Mpb with at least one significant p-value among the three tests. In rows, from top to bottom, hits are determined using block-test, complete-test, single-test

Supplementary Figure 11: Brain sulci coloured according to $-log10(p)$ of significant hits in the chr16:87 region. Hits - genomic region of 1 Mpb with at least one significant p-value among the three tests. In rows, from top to bottom, hits are determined using block-test, complete-test and single-test.