Single nucleotide polymorphism in the neuroplastin locus associates with cortical thickness and intellectual ability in adolescents

Sylvane Desrivières[1,2], Anbarasu Lourdusamy[1,2], Chenyang Tao[3], Roberto Toro[4,5], Tianye Jia[1,2], Eva Loth[1,2], Lourdes Martinez Medina[1,2], Agnieszka Kepa[1,2], Alinda Fernandes[1,2], Barbara Ruggeri[1,2], Fabiana M. Carvalho[1,2] , Graham Cocks[1], Tobias Banaschewski[6,7], Gareth J Barker[1], Arun LW Bokde[8], Christian Büchel[9], Patricia J Conrod[1,10], Herta Flor[11], Andreas Heinz[12], Jürgen Gallinat[12], Hugh Garavan[13,14], Penny Gowland[14], Rüdiger Brühl[15], Claire Lawrence[16], Karl Mann[6], Marie-Laure Paillere Martinot[17], Frauke Nees[11], Mark Lathrop[18], Jean-Baptiste Poline[19], Marcella Rietschel[6], Paul Thompson[20], Mira Fauth-Bühler[21], Michael N Smolka[22,23], Zdenka Pausova[24], Tomáš Paus[16,25,26], Jianfeng Feng[3,27], Gunter Schumann[1,2] and the IMAGEN consortium (www.imagen-europe.com)

[1]Institute of Psychiatry, King's College London, United Kingdom; [2]MRC Social, Genetic and Developmental Psychiatry (SGDP) Centre, London, United Kingdom; [3]Center for Computational Systems Biology, Fudan University; [4]Human Genetics and Cognitive Functions, Institut Pasteur, Paris, France; [5]CNRS URA 2182, Genes, synapses and cognition, Institut Pasteur, Paris, France; [6]Central Institute of Mental Health, Mannheim, Germany; [7]Medical Faculty Mannheim, University of Heidelberg, Germany; [8]Institute of Neuroscience, Trinity College Dublin, Dublin, Ireland; [9]Universitaetsklinikum Hamburg Eppendorf, Hamburg, Germany; [10]Department of Psychiatry, Université de Montreal, CHU Ste Justine Hospital, Canada; [11]Department of Cognitive and Clinical Neuroscience, Central Institute of Mental Health, Medical Faculty Mannheim, Heidelberg University, Mannheim, Germany; [12]Department of

Psychiatry and Psychotherapy, Campus Charité Mitte, Charité – Universitätsmedizin Berlin, Germany; [13]Institute of Neuroscience, Trinity College Dublin, Dublin, Ireland; [14]Departments of Psychiatry and Psychology, University of Vermont, USA; [15]Physikalisch-Technische Bundesanstalt (PTB), Braunschweig und Berlin, Germany; [16]School of Psychology, University of Nottingham, United Kingdom; [17]Institut National de la Santé et de la Recherche Médicale, INSERM CEA Unit 1000 "Imaging & Psychiatry", University Paris Sud, Orsay, and AP-HP Department of Adolescent Psychopathology and Medicine, Maison de Solenn, University Paris Descartes, Paris, France; [18]Centre National de Génotypage, Evry, France; [19]Neurospin, Commissariat à l'Energie Atomique et aux Energies Alternatives, Paris, France; [20]Imaging Genetics Center / Laborarory of Neuro Imaging, UCLA School of Medicine, Los Angeles, CA; [21]Central Institute of Mental Health, Medical Faculty Mannheim / Heidelberg University, Department of Addictive Behaviour and Addiction Medicine, Germany; [22]Department of Psychiatry and Psychotherapy, Technische Universität Dresden, Germany; [23]Neuroimaging Center, Department of Psychology, Technische Universität Dresden, Germany; [24]The Hospital for Sick Children, University of Toronto, Toronto, Canada; [25]Rotman Research Institute, University of Toronto, Toronto, Canada; [26]Montreal Neurological Institute, McGill University, Canada; [27]Department of Computer Science and Centre for Scientific Computing, Warwick University, United Kingdom.

Correspondence: Dr Sylvane Desrivières, MRC-SGDP Centre at the Institute of Psychiatry **(**King's College London)**, 16** De Crespigny Park, Denmark Hill, London SE5 8AF, United Kingdom
Tel: +44(0)20 7848 0528, Fax: +44(0)20 7848 0866, E-mail: sylvane.desrivieres@kcl.ac.uk

Running title: Genetic association study of cortical thickness

Abstract

Despite the recognition that cortical thickness is heritable and correlates with intellectual ability in children and adolescents, the genes contributing to individual differences in these traits remain unknown. We conducted a large-scale association study in 1,583 adolescents to identify genes affecting cortical thickness. Single nucleotide polymorphisms (SNPs; n = 54,837) within genes whose expression changed between stages of growth and differentiation of a human neural stem cell line were selected for association analyses with average cortical thickness. We identified a variant, rs7171755, associating with thinner cortex in the left hemisphere ($P = 1.12 \times 10^{-7}$), particularly in the frontal and temporal lobes. Localized effects of this SNP on cortical thickness differently affected verbal and non-verbal intellectual abilities. The rs7171755 polymorphism acted in *cis* to affect expression in the human brain of the synaptic cell adhesion glycoprotein-encoding gene *NPTN*. We also found that cortical thickness and *NPTN* expression were on average higher in the right hemisphere, suggesting that asymmetric *NPTN* expression may render the left hemisphere more sensitive to the effects of *NPTN* mutations, accounting for the lateralized effect of rs7171755 found in our study. Altogether, our findings support a potential role for regional synaptic dysfunctions in forms of intellectual deficits.

Keywords: cortical thickness, neuroimaging, adolescent, intelligence, lateralization, neural stem cell

Introduction

Genetic factors have a significant contribution in defining brain structure and cognition. In particular, cortical thickness is heritable, with the strongest genetic influences (heritability range 0.50 to 0.90) showing region- and age- specific variations [1] that seem to follow patterns of brain maturation from childhood to early adulthood. Cortical thickness also closely correlates with intellectual ability in normally developing children and adolescents [2, 3]. Yet, little is known about the genetic factors accounting for inter-individual differences in both of these traits.

Advances in neuroimaging studies have enabled the demonstration of spatio-temporal alterations in brain structure and function that occur over a lifetime. This plasticity is particularly important during adolescence, when both hormonal and social environments change dramatically. Whereas white matter increases linearly during this period [4], regional changes in cortical gray matter are non-linear [5]. Localized, region-specific brain gray matter maturation progresses in patterns that appear to follow cognitive and functional maturation [6]. Roughly, areas involved in spatial orientation (parietal lobes) and more advanced functions (frontal lobe) mature around adolescence, after areas of the brain associated with more basic functions (occipital lobe) and before the temporal cortex. Measures of cortical thickness revealed asymmetrical changes in the brain of normally developing children and adolescents. Of notable significance, changes in cortical thickness in the left hemisphere have been found to correlate with performance of children on a test of general verbal intellectual functioning [2]. This plasticity appears to be important in

shaping behaviours and cognitive processes that contribute to normal into adulthood.

Twin studies have demonstrated that brain structure is under significant genetic influence [7], with cortical thickness showing high heritability in children [1, 8] and adults [9, 10]. Differences in heritability are nonetheless notable. Firstly, comparison of estimates of genetic effects in the left and right hemispheres indicate that these values are higher for the left hemisphere, suggesting that the language-dominant left cerebral cortex may be under stronger genetic control than the right cortex [8]. Secondly, age-related differences in the heritability of cortical thickness in children and adolescents have been reported: while regions of primary sensory and motor cortex, which develop earlier, show relatively greater genetic effects in childhood than in adolescence, regions within the frontal cortex, parietal and temporal lobes, associated with complex cognitive processes such as language executive function and social cognition, show relatively greater genetic effects in adolescence [1]. Thus, as suggested by these studies, it would seem necessary to consider both region-specific effects and developmental stage (i.e., age) of individuals when investigating links between genes, cortical thickness and behavior.

At the cellular level, changes in cortical thickness during adolescence are consistent with known cellular maturational alterations, such as the changes in synaptic density [11] and intra-cortical myelination [12] occurring during this developmental stage. Thus, the age-related changes in heritability noted above may be linked to the timing of the expression of given set of genes involved in specific stages of neural

development. It has been postulated that cortical thickness is determined by the number of neurons within radial units of the cortex, and that a diminished ability of the neurogenic progenitors contained in these units to proliferate or to generate neurons will result in a thinner cortex [13, 14]. According to this, genes involved in neural progenitor cell division and/or differentiation are expected to influence cortical thickness. Yet, although cortical thickness is heritable and closely correlates with cognitive ability in children and adolescents [3], the genes influencing these traits remain to be identified.

Therefore, aiming at identifying genes influencing cortical thickness, we decided to focus our analysis on genes relevant for neural progenitor cell proliferation and differentiation. For this purpose, we selected genes whose expression changed between various stages of growth and differentiation of a human neural progenitor cell line. We then conducted association analyses of genetic variations at the selected gene loci and cortical thickness at each hemisphere. Using this approach, we uncovered a gene linking cortical thickness to cognition.

Materials and Methods

Participants

Data analyzed in this study were obtained from 1,583 14-year-old adolescents, participants of the IMAGEN project, for which magnetic resonance images passing quality control procedures were available. Recruitment procedures have been described previously [15] and written informed consent was obtained from all participants and their legal guardians. Individuals completed an extensive battery of neuropsychological, clinical, personality and drug use assessments online and at the testing centers. Participants were excluded if they had contra-indications for MRI (for example, metal implants, claustrophobia). Some individuals were only included in part of the analyses, depending on availability of the genotype, imaging and cognitive data for each participant. The characteristics of this sample are described in **Table 1**.

Cognitive assessment

The Block Design and Matrix Reasoning subtests of the Wechsler Intelligence Scale for Children – Fourth Edition (WISC-IV) [16], (short form of the WISC-IV), were computed to generate a Perceptual Reasoning index and assess non-verbal intelligence (non-verbal IQ). The Similarities and Vocabulary subtests were computed to generate a Verbal Comprehension index measuring verbal concept formation, i.e. the subjects' ability to verbally reason (referred to as verbal IQ). For this, single test scores were converted to more precise age-equivalent scores values. Score values of the relevant subtests were summed to generate indices for

Perceptual Reasoning or Verbal Comprehension. To control for differences in developmental status between participants, pubertal status of the sample was assessed using the Puberty Development Scale [17], which provides an eight-item self-report measure of physical development based on the Tanner stages.

SNP genotyping and Quality control

DNA purification and genotyping were performed by the Centre National de Génotypage in Paris. DNA was purified from whole blood samples (~10ml) preserved in BD Vacutainer EDTA tubes (Becton, Dickinson and Company) using the Gentra Puregene Blood Kit (Qiagen) according to the manufacturer's instructions. A total of 705 and 1382 individuals were genotyped with the Illumina Human610-Quad Beadchip and Illumina Human660-Quad Beadchip, respectively. For each genotyping platform the following quality control was performed separately. SNPs with call rates < 95%, minor allele frequency < 5%, deviation from the Hardy-Weinberg equilibrium ($P \leq 1 \times 10^{-3}$), and non-autosomal SNPs were excluded from the analyses. Individuals with excessive missing genotypes (failure rate > 5%) were also excluded. Population homogeneity was examined with the Structure software, using HapMap populations as reference groups [18]. Individuals with divergent ancestry (from CEU) were excluded. Identity-by-state clustering and multidimentional scaling (MDS) were used to estimate cryptic relatedness for each pair of individuals using the PLINK software [19] and closely related individuals were eliminated from the subsequent analysis. We applied principal components analysis (PCA) to remove remaining outliers [20], defined as individuals located at more than 4 SD of the mean PCA scores on one of the first 20 dimensions. Finally, the integrated genotypes from both Illumina Human610-

Quad BeadChip and Human660-Quad BeadChip were combined and platform-specific SNPs were removed. After the quality control measures, we obtained a total of 466,125 SNPs in 1,834 individuals.

Magnetic Resonance Imaging

Full details of the MRI acquisition protocols and quality checks have been described previously [21]. Brain images were segmented with the FreeSurfer software package and the entire cortex of each individual was inspected for inaccuracies. Individuals with major malformations of the cerebral cortex were excluded from further analysis. Out of 1,909 images, 1,584 passed these quality control checks. In addition to global mean thickness of the left and right cerebral hemispheres, neuroimaging measures included cortical thickness for 33 individual regions per hemisphere. These were combined to produce weighted average thickness (weighted for surface at each region) for the 4 cerebral lobes (i.e., frontal, temporal, parietal and occipital). The effect of MRI site was controlled by adding it as a nuisance covariate in all statistical analyses.

Human neural stem cell culture

The human neural stem cell line SPC-04 was generated from 10-week-old human fetal spinal cord [22] and was cultured mainly as previously described [23]. Briefly, cells were plated on tissue culture flasks that had been freshly coated with laminin (20 μg/ml in DMEM:F12 for 3h 37°C), at a density of 20,000 cells/cm^2 and routinely grown into a reduced minimum media (RMM) formulation consisting of DMEM:F12 with 0.03% human serum albumin, 100 μg/ml human Apo-transferrin, 16.2 μg/ml

putrescine dihydrochloride, 5 μg/ml human insulin, 60 ng/ml progesterone, 2 mM L-glutamine, 40 ng/ml sodium selenite. This RMM was also supplemented with growth factors (10 ng/ml bFGF and 20 ng/ml EGF) and 100 nM 4-hydroxy-tamoxifen (4-OHT). Cell differentiation was triggered when cells reached about 80% confluence by depleting the medium of growth factors and 4-OHT. This was achieved in 2 steps. First, the growth factors- and 4-OHT-depleted medium was supplemented with 10 μM of the γ-secretase inhibitor DAPT and 100 nM all-trans- retinoic acid for 48h. We referred to this stage as "pre-differentiation". Afterwards, differentiation was achieved by maintaining the cells in RMM without any supplements for up to 7 days, with media change every 2 days.

RNA extraction and microarray analyses and SNP selection

RNA was extracted from triplicate SPC04 differentiation experiments using the RNeasy Mini Kit (Qiagen), according to the manufacturer's instructions. Total RNA samples were processed using the TargetAmp™-Nano Labeling Kit (Cambio Ltd.) and hybridized to Illumina HumanHT-12 v4 Expression BeadChips according to the manufacturers' instructions at the Biomedical Genomics microarray core facility of the University of California, San Diego. Raw data were extracted by the Illumina BeadStudio software and further processed in R statistical environment (http://www.r-project.org) using Bioconductor packages. Raw expression data were log2 transformed and normalized by quantile normalization. Differential expression between each differentiated versus undifferentiated conditions was assessed using the linear model for microarray analyses (LIMMA) package. *P* values were adjusted for multiple testing according to the false discovery rate (FDR) procedure of

Benjamini and Hochberg and differentially expressed genes were selected at FDR < 5%. See Supplementary **Table 1** for the list of differentially expressed genes. The functional annotation clustering tool, part of the Database for Annotation, Visualisation and Integrated Discovery (DAVID) [24] was used to determine enrichment of functional groups in genes list generated from the microarray analyses. SNPs (n = 59,643) lying within ±10 kB of each differentially expressed autosomal gene were selected for genetic association studies, of these n = 54,837 passed genetic quality controls and were used in further association analyses.

Genetic associations

Linear regression analyses were performed in PLINK [19] using average cortical thickness of the left or right hemisphere was used as a dependent variable and the additive dosage of each SNP as an independent variable of interest, controlling for covariates of age, sex, puberty and the first four principal components from MDS analysis. Dummy covariates were also used to control for different scanning sites. Genome-wide complex trait analysis (GCTA) [25] was used to estimate the proportion of phenotypic variance in left cortical thickness explained by all genotyped SNPs and SNPs selected from our differential gene expression analyses. The GCTA was fitted using a restricted maximum likelihood method. The Broad Institute's SNAP online plotting tool [26] was used to generate the regional association and recombination rate plots.

The same conditions were used when investigating the association between rs7171755 and IQ except that ethnicity was also included as a nuisance covariate. Given correlations between brain volume (i.e., the sum of all cortical and subcortical

gray and white matter, excluding ventricle and cerebrospinal fluid), cortical thickness, cortical surface area and IQ, left surface area was also included as covariate when using brain volume or IQ as a variable. For the associations of rs7171755 with brain volume, linear regression analyses were performed using site, sex, left surface area and 4 MDS components as covariates. Handedness influenced none of the above associations and was not included as a covariate in our analyses. Mediation analyses between SNP x left (average or frontal) cortical thickness x non-verbal IQ were performed in SPSS using the PROCESS boostrapping procedure [27] with 1000 boostrap samples used to calculate 95% confidence interval (CI) estimates of indirect effects.

Bonferroni corrections adjusting for the total number of tests in each analysis were performed to control for multiple testing. For the genotypes x cortical thickness association analyses with the selected 54,837 SNPs, on the left and right hemispheres, the corresponding significance threshold was $P = 4.56 \times 10^{-7}$.

Meta-analytic association of rs7171755 with brain volumes in ENIGMA

We have used the ENIGMA dataset, the largest meta-analysis of gene x neuroimaging phenotypes, to investigate association of rs7171755 with total brain volume, the brain phenotype most closely related to cortical thickness available in this dataset. Association of rs7171755 with brain volume was performed using the online tool EnigmaVis [28], generating an interactive association plot. Only the healthy subsample (N = 5775) of ENIGMA, for which this brain volume was available, were included in the meta-analysis.

Bootstraping procedure

To provide bias-reduced estimates of the associations reported above, we used a bootstrap re-sampling approach [29] for Linear Regression Models as follows. First, subjects were resampled with replacement from the subjects passing quality controls criteria, here referred to as the bootstrap sample. Second, the coefficient β_{SNP} for the SNP of interest from the bootstrap sample was calculated. We shuffled the SNP column of the bootstrap sample 100,000 times and recalculated the β_{SNP}, generating a NULL distribution of β_{SNP} for the bootstrap sample, denoted as β_{NULL}. Third, the empirical p-value (P_{emp}) of the bootstrap sample was determined as the portion of β_{NULL} greater than β_{SNP}. We repeated this bootstrap procedure 10,000 times to obtain an empirical distribution of the p-values for each variable of interest.

Least Square Kernel Machine (LSKM) Association tests for candidate genes

As genetic association testing based on single SNPs might suffer from low power, we have also used a more sophisticated LSKM procedure that we have recently developed to analyze joint effects of several SNPs with imaging traits [30] to detect possible genetic influences on cortical thickness. In short, this procedure compares the individuals' allele profiles, composes a similarity matrix (Kernel Matrix), and then determines to what extent the similarity matrix explains variations in the phenotype. A summary statistics is used to evaluate the significance under null hypothesis. We considered SNPs within +/- 10Kb of a gene's transcript region as 'belonging to' the corresponding gene. In the current analysis, a gene-wide IBS matrix was used as similarity matrix. After quality control, 2,659 out of the ~3,540 genes differentially expressed in our microarray analyses were retained and subjected to the LSKM

analysis. As for the single SNP association analyses, recruitment site, gender, age, puberty, ethnicity and the four first MDS components were used as covariates in the LSKM analyses.

NPTN expression on mouse brain samples

RNA samples extracted from CD1 mouse brains at embryonic day 10 (E10), E14, E18, and at postnatal (P) stages 1 week, 1 month, or 6 months were obtained from AMS Biotechnology. Whole brain mouse RNAs extracted from pools of 5 and 3 embryos were used for the E10 and E14 stages, respectively. RNAs extracted from the frontal cortex were used for later developmental stages (i.e, E18 to P6months). In this case, triplicate samples from independent brains were analyzed for each stage, except for the P6month stage for which data was derived from a single mouse brain. cDNAs obtained by reverse-transcription using the SuperScript® III First-Strand Synthesis System (Invitrogen) following the manufacturer's instructions, were amplified by PCR with *GAPDH* as an internal control, using the following forward and reverse primers: *GAPDH*-F 5'-TGTTCCTACCCCCAATGTGT-3'; *GAPDH*-R 5'- CCTGCTTCACCACCTTCTTG-3'; *NPTN*-F 5'-GCCTTTCTTGGGAATTCTGGC-3'; *NPTN*-R 5'-AGAGTTGGTTTTCATTGGTCCAG-3'. PCRs were run in triplicate in the Applied Biosystems real-time PCR device (7900HT Fast Real Time PCR system) in 20µl reactions containing 4µl cDNA, 0.5µM of each forward and reverse primers and 1x Power SYBR Green Mix (Applied Biosystems) using the following cycles: 95°C for 15 minutes and 40 cycles at 95°C for 30 seconds and 59°C for 30 seconds. The PCR reaction products were evaluated by a melting curve analysis. Relative quantification of the PCR products was performed using the SDS software (Applied Biosystems)

comparing threshold cycles (Ct). *NPTN* mRNA levels were first normalized to that of *GAPDH* (deltaCt = Ct_{NPTN} - Ct_{GAPDH}) at each developmental stage, and changes in expression relative to E10 were calculated as $2^{-(deltaCt - deltaCtE10)}$. Statistical analysis (one-way ANOVA, followed by Bonferroni-based post-hoc analysis with $\alpha = 0.05$, 2-sided) was performed comparing expression of triplicates at the E18, P1week and P1month stages to that at E10.

***NPTN* expression in human brain samples**

Expression of NPTN in the human brain was investigated using two databases. To study effects of rs7171755 on *NPTN* expression (probe 33624, targeting NM_001161363 and NM_012428), we used the publicly available BrainCloud database (http://BrainCloud.jhmi.edu/), which includes data on gene expression and genotypes from post-mortem dorsolateral prefrontal cortex samples collected from 272 subjects across the lifetime. In this database, transcript expression levels were measured on Illumina Oligoset array of 49152 probes, and genotyping was performed using Illumina Infinium II or HD Gemini 1M Duo BeadChips [31]. The genetic datasets were obtained from dbGaP at http://www.ncbi.nlm.nih.gov/gap through dbGaP accession number phs000417.v1.p1. Submission of the data phs000417.v1.p1 to dbGaP was provided by Drs. Barbara Lipska and Joel Kleinman. Data collection was through a collaborative study sponsored by the NIMH Intramural Research Program. Initial report on this dataset is from Colantuoni *et al.* [31]. For this study, we considered only samples with good RNA quality (RIN ≥ 8). Statistical analyses measuring effects of rs7171755 on the postnatal expression of *NPTN* were

performed on 147 samples (individuals ≥ 0.5-year old) by general linear models controlling for age, ethnicity and RNA quality.

To investigate possible differences in *NPTN* expression between brain hemispheres, we analyzed a database (GEO series GSE25219) containing genome-wide gene expression data from 16 brain regions on both hemispheres, collected from 57 subjects across the lifetime (N=1,340 post-mortem brain samples) [32]. Paired samples t-tests were performed comparing expression on *NPTN* on the right and the left hemisphere for each sample, controlling for the developmental stage and RNA integrity factor.

Results

Selection of genes involved in neural progenitor function

We first selected genes differentially expressed at any stage of proliferation and differentiation of a human neural stem cell line, SPC04. These cells proliferated readily in undifferentiated conditions and acquired a typical neural morphology, with well-developed neurites as early as 3 days after induction of differentiation (**Fig. 1a**). Microarray analyses, comparing gene expression profiles of undifferentiated cells with pre-differentiated cells, or cells that have been induced to differentiate for 3 or 7 days, led to identification of ~3,540 genes that were differentially expressed between these stages, with most of the changes in gene expression occurring 7 days after differentiation (**Fig. 1b**). Gene ontology clustering analyses indicated enrichment of genes down-regulated (n=1,605) at differentiation day 7 for genes involved in cell cycle (enrichment score: 44.57) and DNA metabolic processes (enrichment score: 16.92). Up-regulated genes (n=1,675) were mainly enriched for genes involved in cell adhesion (enrichment score: 7.97), synaptic transmission (enrichment score: 5.83), neuron morphogenesis and differentiation (enrichment score: 5.46) and synapse formation and organization (enrichment score: 4.43). SNPs (n=59,643) located within ± 10kb of differentially expressed autosomal genes were selected for association with cortical thickness.

Large-scale association studies with cortical thickness in adolescents

Given that left-right asymmetry of the brain is a well-known phenomenon [33, 34] that may be triggered by left-right differential gene expression [35, 36], we analyzed each hemisphere separately. Highest associations with left cortical thickness were found

for SNPs on Chromosome 15 (**Fig. 2a**, **Table 2** and **Supplementary Fig. 1)**, with one SNP, rs7171755 (β = -0.01973; P = 1.12 x 10^{-7}), passing the threshold of Bonferroni-corrected significance (the Bonferroni-adjusted significance threshold for association with the selected 54,837 SNPs, on the left and right hemispheres, was P = 4.56 $\times 10^{-7}$). In the right hemisphere, highest associations with cortical thickness were found on Chromosome 11 (**Supplementary Fig. 2**); however, none remained significant after Bonferroni correction for multiple testing. Rs7171755 was associated with right cortical thickness at P = 3.22 x 10^{-4} (β = -0.0134; **Table 2**). Neither handedness nor ethnicity influenced this association. It is worth pointing out that our gene selection procedure resulted in significant gene enrichment: estimation of the variance explained by the SNPs using Genome-wide Complex Trait Analysis (GCTA) [25] indicated that the 59,643 selected SNPs explain 13.3% (SE = 0.093, P = 0.02) of the total variance in left cortical thickness, a five-fold enrichment relative to the 22.2% (SE = 0.195, P = 0.03) variance explained by considering all 506,932 genotyped SNPs simultaneously.

The number of minor alleles at rs7171755 was inversely correlated with average cortical thickness. In the left hemisphere, we observed a decrease of 0.0189 mm (i.e., 0.7% of the average left cortical thickness) per risk allele, explaining 2% of variance. To investigate if effects of rs7171755 on cortical thickness differed across brain regions, we processed the segmented left and right cortical lobes (frontal, temporal, parietal and occipital) into 66 cortical subregions [37] and performed linear regressions, analyzing associations of rs7171755 with cortical thickness within each region. Region-specific effects of rs7171755 on cortical thickness were observed, with most significant overall influences on the cortical thickness in the left temporal

(β = -0.0275; P = 1.23 x 10^{-7}), frontal (β = -0.0212; P = 6.98 x 10^{-7}) and parietal (β = -0.0170; P = 1.684 x 10^{-4}) lobes. In the right hemisphere, associations were significant only for the frontal and temporal lobes (β = -0.0169; P = 8.91 x 10^{-5} and β = -0.0165; P = 1.667 x 10^{-3}, respectively). A further refined neuroanatomical segmentation revealed that these asymmetric associations occurred throughout the left frontal cortex, including the lateral orbito-frontal, the caudal middle-frontal and the superior frontal cortex, the para- and pre-central region and the pars orbitalis. Other significant associations were observed in the left superior and middle temporal cortices and in the left supramarginal region (**Table 3**).

The SNP rs7171755 is located less than 2 KB downstream of the *NPTN* gene and is in high linkage disequilibrium (LD) with other SNPs within the *NPTN* locus. Regional association analysis for SNPs around rs7171755 clearly show that *NPTN* is the candidate gene associated with this signal: the SNPs with the smallest p-values, all in high LD with rs7171755, are located across this gene (**Fig. 3**).

To confirm our finding and to test for a possible significance of joint contribution of multiple SNPs within the *NPTN* locus to left cerebral cortex thickness, we performed gene-wide SNP-sets analyses using the Least Square Kernel Machine (LSKM) approach [30, 38, 39]. The results indicate that, in addition to rs7171755, eight SNPs: rs7176637, rs11854138, rs8028749, rs12185108, rs1564492, rs899981, rs7178269 and rs4075802, in the *NPTN* locus, jointly show significant association with average left cortical thickness (P = 1.264 x 10^{-8}; **Fig. 2b**). As for the single SNP analyses, the most significant associations were observed, in decreasing order, in the left temporal (P = 1.97 x 10^{-10}), left frontal (P = 1.87 x 10^{-8}), left parietal (P = 4.59 x 10^{-}

^{5}) and right frontal (P = 6 x 10^{-4}) lobes. More refined region-specific analyses confirmed the single SNP associations described above (**Supplementary Table 2**)

To demonstrate the stability of the above associations and obtain unbiased estimation of the genetic effects, we used a bootstrapping resampling procedure [29, 40]. Effects of rs7171755 on left cortical thickness were confirmed, with a decrease of 0.0196 mm per risk allele. Effects of this variant were also confirmed for the hemispheric lobes (left frontal lobe: β = -0.022, P_{emp} = 1 x 10^{-6}; left temporal lobe: β = -0.027, P_{emp} = 1 x 10^{-6}; left parietal lobe: β = -0.017, P_{emp} = 7.1 x 10^{-5}; right frontal lobe: β = -0.017, P = 3.3 x 10^{-5} and right temporal lobe: β = -0.017, P = 6.64 x 10^{-4}) and individual ROIs (**Table 3** and **Supplementary Fig. 3**)

In order to have sufficient power to unambiguously reject an observed association, a sample size larger than that of the original study is required [41]. However, a replication sample larger than the IMAGEN sample, with comparable phenotypic characteristics, including assessment of cortical thickness during adolescence is not yet available. We have nonetheless attempted to overcome these limitations by further testing rs7171755 in the ENIGMA dataset, a meta-analysis of gene x neuroimaging phenotypes, where we analyzed its association with brain volume. As cortical thickness measures were not available in the ENIGMA samples, brain volume was the most closely related brain phenotype available [42]. The correlation between brain volume and cortical thickness on the left hemisphere in the IMAGEN sample was high (r(1188) = 0.491; P = 4.29 x 10^{-73}). Upon measuring association of rs7171755 with brain volume, we found in the IMAGEN sample, an association of rs7171755 with significant decrease of brain volume of 3080 mm^3 (β = -3080, P = 0.0457) per risk allele. We have replicated this finding in the subsample of

healthy individuals (*N* = 5775) of ENIGMA, using the EnigmaVis tool [28], confirming the negative effects of the risk allele on brain volume (decrease of 5945.91 mm^3 per risk allele: β = -5945.91, *P* = 0.00327; **Supplementary Fig. 4**). Together, these results further support a role for *NPTN*-related genotypes in influencing brain structure.

Association of rs7171755 with adolescents' intellectual ability

The results presented above, along with previous findings showing relationships between intellectual ability and cortical thickness in healthy subjects, predominantly in frontal and temporal cortical regions [2, 3], suggest that rs7171755 might influence cognitive ability. To test this, and assess gene-brain-behaviour relationships, we estimated Pearson's correlations between indices of intellectual ability and cortical thickness in our sample and found significant positive correlations between average cortical thickness and non-verbal IQ, that were more pronounced in the left hemisphere ($r(1168) = 0.074$; $P = 0.012$ and $r(1168) = 0.06$; $P = 0.041$, in the left and right hemisphere, respectively) (see **Supplementary Table 3**). A positive correlation between left cortical thickness and school performance was also observed ($r(1168) = 0.062$; $P = 0.033$). Correlations with verbal IQ were not significant at an unadjusted $P < 0.05$. In the regions of the left cerebral cortex most significantly affected by rs7171755, i.e., the left temporal and frontal cortices, correlations were also significant for non-verbal IQ ($r(1170) = 0.061$; $P = 0.036$; and $r(1170) = 0.075$; $P = 0.011$, respectively); there was also borderline significance for correlations with verbal IQ (in the temporal lobe only ($r(1170) = 0.059$; $P = 0.044$)).

These results suggested that, by affecting cortical thickness, rs7171755 might influence IQ. Mediation analyses performed to test this hypothesis indicated that the

minor A-allele at rs7171755 associates with lower scores for non-verbal IQ (β = -1.239; *P* = 0.0219). This association was mediated by significant indirect effects (i.e., via left frontal lobe thickness) of this SNP on non-verbal IQ (β = -0.1851; 95% CI[-0.391; -0.046]), while direct effects of the SNP on non-verbal IQ were not significant (95% CI[-2.12; 0.023]). Surprisingly, rs7171755 also associated with verbal IQ (β = -1.5048; *P* = 0.0076), an association that was not mediated by indirect effects on mean or temporal thickness. This suggested that more localized effects of rs7171755 on brain structure might underlie this association. To test this, we investigated correlations between verbal IQ and cortical thickness in language-related ROIs in the left frontal and temporal lobes, where effects of rs7171755 on cortical thickness were strongest (see **Table 3**): the pars orbitalis and the middle temporal and superior temporal regions. Positive correlations between verbal IQ and cortical thickness were found in the pars orbitalis (*r*(1170) = 0.080; *P* = 0.006), while a trend was also found in the middle temporal gyrus (*r*(1170) = 0.055; *P* = 0.060). No correlation was observed with thickness in the superior temporal gyrus. Mediation analyses indicated that indirect effects (β = -0.1486; 95% CI[-0.3347; -0.0376]) of rs7171755 on left pars orbitalis thickness partially contributed to its association with verbal IQ, with other factors accounting for the remaining effects (β = -1.3562; *P* = 0.0165).

Boostrapping analysis revealed similarly negative effect of rs717175 on IQ with a decrease in intelligence by about 1.81 points and 1.41 points per allele for verbal and non-verbal IQ, respectively (β = -1.808, *P* = 0.002 and β = -1.407, *P* = 0.008; for verbal and non-verbal IQ, respectively) (**Supplementary Fig. 3**), accounting

for 0.7% and 0.5% of the total variance in IQ, respectively. Altogether, these analyses indicate that the minor allele at rs7171755, via its effects on cortical thickness, particularly in the left frontal lobe, negatively impacts intellectual abilities.

Effects of rs7171755 on *NPTN* expression

NPTN, a gene selected for our analyses due to its increased expression in differentiating human neural progenitor cells (1.5 fold increase at differentiation day 7 compared with undifferentiated cells; FDR < 0.05), encodes splice isoforms of neuroplastin, a synaptic cell adhesion glycoprotein [43]. This induction of *NPTN* occurs at a time when neurites are well developed, and appears to coincide with induction of genes involved in cell adhesion and synaptic transmission (see above). To confirm this, we investigated patterns of *NPTN* expression in the brain. First, we investigated changes in *NPTN* expression in the mouse brain during stages of embryonic and postnatal development. One-way ANOVA indicated that while levels of *NPTN* mRNA are low in the mouse neocortex during embryonic development, expression of this gene is markedly increased in the first week after birth, reaching maximum levels 1 month after birth ($F(3, 8) = 53.83$; $P = 1.2 \times 10^{-5}$), a time period which corresponds to adolescence in mice (**Fig. 4a**). To confirm relevance of our findings to human brain development, we interrogated the BrainCloud database, which contains genome-wide expression data of the prefrontal cortex of 272 individuals across the lifespan as well as their genotype information [31]. Investigations of changes in *NPTN* expression in the human prefrontal cortex across lifetime, confirmed the expression patterns observed in the mouse brain. While levels of *NPTN* (isoforms NM_001161363 and NM_012428) were low during early fetal development, its

expression increased at later stages of development to reach maximum levels in childhood through early adulthood, after which expression declines (**Fig. 4b**). To gain functional insight into these changing expression patterns, we searched for genes whose expression correlated with that of *NPTN* in the human prefrontal cortex across the lifespan and examined their enrichment for functional gene groups. Expression of *NPTN* positively correlated with a cluster of 721 genes (r>0.6) enriched for genes involved in energy metabolism (n = 36 (7%), P = 1.08 x 10^{-11}), synaptic transmission (n = 27 (5%), P = 7.98 x 10^{-7}) as well as learning and memory (n = 14 (3%), P = 2.3 x 10^{-5}). Lists of correlated genes and their grouping into functional clusters are contained in **Supplementary Tables 4 and 5**, respectively.

We then investigated possible *cis*-effects of rs7171755 on *NPTN* expression, by testing if rs7171755 genotypes correlated with differences in *NPTN* expression. We found that expression of *NPTN* differed by genotypes; individuals homozygotic for the minor A-allele at rs7171755 had lower expression of this gene (P = 0.009; **Fig. 4c**). Remarkably, this difference was most notable from adolescence to early adulthood (late 20s; **Fig. 4b**), suggesting age-dependent effects of rs7171755.

The results presented above point to a lateralized effect of rs7171755, associated with cortical thickness predominantly in the left hemisphere. We analyzed this further, investigating possible asymmetries in cortical thickness and *NPTN* expression. For this purpose, we performed paired samples t-tests that indicated that while cortical thickness correlated well between hemispheres (r(1583) = 0.864) in our sample, the cortex was on average 0.012 mm thicker on the right hemisphere (t(1582) = 9.818, P = 3.977 x 10^{-22}). To test for asymmetric expression of *NPTN* in the human brain, we analyzed a database (GEO series GSE25219) containing

gene expression data from 16 brain regions on both hemispheres (N = 1,340 post-mortem brain samples) [32]. Paired samples t-tests comparing expression on *NPTN* in the right vs. the left hemisphere indicated that RNA levels of this gene where higher in the right hemisphere then in the left ($NPTN_{left}$ - $NPTN_{right}$ = -0.0377, $t(523) = -2.703$, $P = 0.007$). These results illustrate asymmetries in the human brain; with both cortical thickness and *NPTN* expression being more pronounced in the right hemisphere. The observed asymmetry in *NPTN* expression may render the left hemisphere more sensitive to the effects of *NPTN* mutations, accounting for the lateralized effects of rs7171755 found in our study.

Discussion

In this study, we have used a large sample of healthy adolescents to investigate the genetic basis of inter-individual variations in cortical thickness and relevant cognitive phenotypes. We performed transcriptional profiling of human neural progenitor cells for neural gene enrichment to allow targeted SNP selection for association analyses with structural neuroimaging and cognitive phenotypes. Using this combined, hypothesis driven, approach we were able to identify the *NPTN* locus as contributing to individual differences in brain structure and cognition. The SNPs within *NPTN* associate with cortical thickness in the left hemisphere, most significantly in areas associated with higher cognitive functions including regions throughout the left frontal and temporal cortices and the left supramarginal area. The minor allele at rs7171755, which associates with lower cortical thickness at those regions and decreased performance of adolescents on tests of intellectual ability, also associates with lower expression of *NPTN* in the human prefrontal cortex. We have provided additional corroborative evidence from the ENIGMA study, the largest gene x neuroimaging meta-analysis study to date by demonstrating the association of *NPTN* rs7171755 with brain volume, a measure of brain structure related to cortical thickness. We also provide evidence for asymmetries in the human brain and propose that asymmetry in *NPTN* expression may render the left hemisphere more sensitive to the effects of *NPTN* mutations, accounting for the lateralized effects of rs7171755 found in our study.

In keeping with our data, asymmetric genetic influences on brain structure have previously been reported, specifically in the frontal and language-related left temporal cortices, where cortical gray matter distribution displays high heritability [7]. Our study, on a cohort very homogeneous for age (i.e., 14-yr), yielded results consistent with a previous report [3] which described positive correlations, peaking in late childhood/early adolescence, between cortical thickness and levels of intelligence, particularly in the prefrontal cortex. This age homogeneity is a critical characteristic of our sample, given the reported changes in correlations between intelligence and cortical thickness from childhood to early adulthood [3]. Our data also support the notion that cortical thickness differentially impacts verbal and non-verbal abilities. While average thickness, particularly in the prefrontal cortex, influenced non-verbal cognitive abilities, more regionally restricted structural effects may control verbal abilities. In this context, our identification of the pars orbitalis as a region mediating such effects is notable, as this is a part of the Broca language area selectively involved in processing the semantic aspects of sentences [44]. Regionally specific cortical thinning in the pars orbitalis, has been documented in individuals with DiGeorge, velocardiofacial syndrome [45], whose cognitive deficits include language and speech delays [46].

In line the proposed role of *NPTN* in neurite outgrowth, we found that induction of this gene in cultured neural progenitor cells occurs at a time when neurites are well developed, coinciding with induction of genes involved in cell adhesion and synaptic transmission. We also found that *NPTN* is expressed in the brain at periods of intense neuronal activation and synaptic activity, which fits well with the emerging

role of this gene as encoding a cell adhesion protein regulating neuritogenesis and synaptic plasticity [47-49]. Our results also indicate that expression of *NPTN* in the cerebral cortex is highest around adolescence, a period that in humans is accompanied by decrease in gray matter in frontal, parietal and temporal areas [5]. This and the proposed role of *NPTN* in neurite outgrowth and synaptic plasticity suggest that, at the cellular level, synaptic architecture of the cerebral cortex underlie the observed differences in cortical thickness and cognitive abilities.

A role for deregulation of *NPTN* in disorders of the nervous system is also emerging. *NPTN* and other genes involved in neurite outgrowth have recently been identified as direct targets of FOXP2 [50], a transcription factor that when mutated causes a monogenic speech and language disorder in humans [51] and whose reduced dosage impairs synaptic plasticity, motor-skill learning and ultrasonic vocalizations in mice [52, 53] and disrupts vocal learning in songbirds [54]. In agreement with our data, this suggests that like *FOXP2*, *NPTN* may be involved in learning vocal and non-vocal skills. Furthermore, functional polymorphisms in the *NPTN* promoter that may confer susceptibility to schizophrenia have been identified [55]. Analyzing data from the 1000 Genomes Project, we found substantial LD (D' = 1, r^2 = 0.502; data not shown) between rs7171755 and rs3743500, one of these promoter polymorphisms associated with schizophrenia. Taken together, these data highlight a potential role for *NPTN* and, more generally, synaptic dysfunctions in forms of intellectual deficits.

Such aspects of neural development have long been thought to underlie formation of higher order cortical functions. The synaptic architecture of the cortex has been

proposed to define the extent of intellectual capacity: changes in dendritic arborisation and spine structure are commonly observed in brain tissue of patients with various types of intellectual disabilities [56, 57] and mutations are found in many different types of cognitive disorders, including intellectual disability (ID), schizophrenia and autism spectrum disorders (ASDs) which affect synaptic morphology and plasticity [58-60]. The most recent observations using animal models of ID/ASD indicate that the pace of maturation of dentritic spine synapses in early postnatal life is vital for normal intellectual development [61]. It is of interest that those dendritic spines which become larger and functionally stronger (i.e., more stable synapses) too early in development trigger subsequent cognitive deficits [61].

It should be noted that the effect sizes observed in our experiments are small, as might be expected from mutations in human genes that regulate late events in neural differentiation. Such mutations may not cause gross cortical malformations, but rather more subtle cognitive and behavioral defects. Given this and the age-specificity of our observations, a major challenge remains to generate additional studies to replicate our findings. Nonetheless, we have partly overcome these limitations, further testing the relevance of *NPTN* genotypes for inter-individual variations in brain structure in our sample and in the ENIGMA consortium for meta-analysis of large neuroimaging and genetics dataset, and demonstrated the negative association of the rs7171755 risk allele with brain volume, further supporting a role for *NPTN* in influencing brain structure. There still is a need to directly replicate our findings. Even more thrilling is the prospect of applying our approach to the

longitudinal study of normal as well as learning disabled and psychiatric samples to investigate spatio-temporal alterations in the genetic influences reported here.

Acknowledgments

We wish to thank Dr. Gary Hardiman for the microarray hybridizations and Prof. Jack Price for giving us the SPC04 cells.

This work was supported by the European Union-funded FP6 Integrated Project IMAGEN (Reinforcement- related behaviour in normal brain function and psychopathology) (LSHM-CT- 2007-037286), the German Ministry of Education and Research (BMBF grant # 01EV0711 and eMED "Alcoholism"), the FP7 project IMAGEMEND (Development of effective imaging tools for diagnosis, monitoring and management of mental disorders) and the Innovative Medicine Initiative Project EU-AIMS (115300-2), as well as the Medical Research Council Programme Grant "Developmental pathways into adolescent substance abuse" (G93558), the Swedish Research Council (FORMAS) and the United Kingdom National Institute for Health Research (NIHR) Biomedical Research Centre Mental Health.

Conflict of interest

Dr Barker receives honoraria for teaching from General Electric and acts as a consultant for IXICO. Dr Banaschewski served in an advisory or consultancy role for Hexal Pharma, Lilly, Medice, Novartis, PCM scientific, Shire and Viforpharma. He received conference attendance support and conference support or received speaker's fee by Lilly, Janssen McNeil, Medice, Novartis and Shire. He is/has been involved in clinical trials conducted by Lilly, Shire & Viforpharma. The present work is unrelated to the above grants and relationships. The other authors declare no potential conflict of interest.

Tables

Table 1: Characteristics of study participants.

	Number [a]	Mean ± SD
Gender		
Females	847	
Males	735	
Ethnicity		
Both parents Caucasian	1353	
Father or mother not Caucasian	103	
Both parents not Caucasian	74	
Handedness		
Right	1376	
Left	159	
Ambidextrous	9	
rs7171755 genotypes		
GG	488	
AG	624	
AA	225	
Age	1473	14.41 ± 0.72
Puberty (Tannerstage)	1570	4.15 ± 0.95
Verbal IQ	1494	110.78 ± 15.51
Non-verbal IQ	1494	107.03 ± 14.29
Average cortical Thickness		
Left hemisphere	1583	2.68 [b] ± 0.09
Right hemisphere	1583	2.70 [b] ± 0.09

[a] Number of participants for which measurements are available. [b] in millimeters.

Table 2: Top 20 SNPs associating with left cortical thickness

				Cortical thickness			
				Left		Right	
CHR	SNP	POSITION	Minor allele	β[a]	[b]*P*	β[a]	[b]*P*
15	rs7171755	71637633	A	-0.01973	1.12E-07	-0.01343	3.22E-04
15	rs7176637	71637807	A	-0.0178	1.38E-06	-0.01031	0.005336
15	rs16944739	89055629	T	-0.01987	2.27E-06	-0.01542	2.54E-04
15	rs899981	71722382	A	-0.01744	2.71E-06	-0.01128	0.002487
15	rs12185108	71680005	T	0.01593	1.29E-05	0.008699	0.01759
7	rs245974	29262381	C	-0.01586	2.62E-05	-0.01151	0.002328
7	rs4722754	28113977	C	0.02535	2.72E-05	0.01814	0.00275
15	rs922687	71635861	A	0.01495	4.36E-05	0.007401	0.04359
6	rs2318064	124231122	A	-0.01793	7.05E-05	-0.01407	0.001849
18	rs3875089	22699431	C	-0.01967	7.63E-05	-0.01765	3.91E-04
18	rs17614110	22683167	T	-0.02191	8.79E-05	-0.01991	3.71E-04
15	rs1491636	66717647	T	0.01509	9.44E-05	0.01189	0.002133
15	rs12901345	66720414	C	0.01443	1.26E-04	0.01008	0.007592
7	rs4571657	139901359	A	-0.02581	1.36E-04	-0.02143	0.00156
11	rs721607	56754408	C	0.0164	1.45E-04	0.01924	8.20E-06
9	rs1571930	100459653	G	-0.01861	1.57E-04	-0.0192	9.74E-05
11	rs2282624	56758487	C	0.01486	1.67E-04	0.01576	6.58E-05
9	rs2900547	100464426	A	-0.01853	1.70E-04	-0.01897	1.18E-04
6	rs4897561	124233916	C	-0.01762	1.70E-04	-0.01384	0.00319
16	rs1870846	81266758	C	0.02995	1.71E-04	0.03208	5.67E-05

[a]β is the regression coefficient that represents changes in the average hemispheric cortical thickness values due to the additive effect of the minor alleles of the SNPs (i.e., positive means minor allele increases thickness).

[b]*P* is the significance of β (uncorrected for multiple comparisons).

Table 3: Associations of rs7171755 across brain regions

	Linear regression				Bootstrap			
	Left hemisphere		Right hemisphere		Left hemisphere		Right hemisphere	
	β	*P*	β	*P*	β	P_{emp}	β	P_{emp}
Frontal lobe								
Caudal middle-frontal	-0.024	**1.53E-04**	-0.020	1.22E-03	-0.024	**5.8E-05**		
Lateral orbitofrontal	-0.027	**2.47E-05**	-0.016	9.60E-03	-0.027	**1.10E-05**		
Paracentral	-0.025	**2.04E-04**	-0.018	1.17E-02	-0.025	**1.1E-04**		
Pars orbitalis	-0.034	**3.48E-04**	-0.025	8.62E-03	-0.037	**5.4E-05**		
Precentral	-0.024	**2.73E-05**	-0.014	1.58E-02	-0.025	**1.2E-05**		
Superior frontal	-0.022	**1.23E-04**	-0.018	2.66E-03	-0.023	**3.7E-05**		
Parietal lobe								
Supramarginal	-0.026	**2.81E-05**	-0.017	4.91E-03	-0.026	**1.3E-05**		
Temporal lobe								
Bank of the STS	-0.027	9.02E-04	-0.028	**5.88E-04**			-0.024	2.93E-03
Fusiform	-0.024	**6.73E-05**	-0.014	1.46E-02	-0.023	**5.6E-05**		
Inferior temporal	-0.030	**7.98E-05**	-0.012	1.12E-01	-0.029	**7.2E-05**		
Middle temporal	-0.032	**1.36E-05**	-0.015	3.48E-02	-0.032	**3.0E-06**		
Superior temporal	-0.027	**6.95E-05**	-0.021	1.36E-03	-0.027	**4.0E-05**		

Only indicated in this Table are regions for which *P*-values remained significant after Bonferroni correction for multiple testing (highlighted in bold). Regression coefficients (β) and *P*-values obtained using the original linear regression model or the bootstrap approach are indicated.

Figure Legends

Figure 1: Differentiation of SPC04 neural progenitor cells in culture. (**a**) Changes that accompany differentiation are evident when comparing morphology of undifferentiated, proliferating cells (a) or pre-differentiated cells (b) with that of cells that have been induced to differentiate for 3 days (c) or 7 days (d). (**b**) Venn diagram representing number of genes differentially expressed between undifferentiated cells and each of the 3 stages of differentiation and their intersection. Abbreviations: und-, undifferentiated; pre-, pre-differentiated; 3d, differentiated for 3 days and 7d, differentiated for 7 days. Scale bars represent 100 μm.

Figure 2: Genetic associations with cortical thickness on the left hemisphere. (**a**) Manhattan plots of single SNPs associations. SNP markers are plotted according to chromosomal location on the *x*-axis, while the *y*-axis $-\log_{10}$ (*P* values) indicate the significance of the additive effect of the number of minor alleles of each SNP on average cortical thickness for the left hemisphere. A conservative Bonferroni-corrected *P* value threshold (red horizontal line) for significance was set to $P = 8.4 \times 10^{-7}$. (**b**) Manhattan plots of LSKM gene-wide associations. Each dot represents a gene (SNPs set), plotted according to chromosomal location on the *x*-axis. The *NPTN* gene, which is most significantly associated with left cortical thickness, is indicated by the arrow.

Figure 3: SNPs within the *NPTN* locus are associated with cortical thickness in the left hemisphere. Regional association and recombination rate plots for SNPs around

rs7171755, genotyped in our sample. The SNP with the most significant association is denoted with a red diamond. The left y-axis represents $-\log_{10}$ *P* values for association with cortical thickness in the left hemisphere, the right y-axis represents the recombination rate, and the x-axis represents base-pair positions along the chromosome (human genome Build 36).

Figure 4: Developmental stage- and genotype-specific expression of *NPTN* in the cerebral cortex. (**a**) Highest expression of *Nptn* in the adolescent mouse brain. *Nptn* mRNA levels in the whole mouse brain at embryonic day 14 (E14), and in the frontal cortex at E18 or 1 week, 1 month and 6 months postnatally (P) were calculated relative to expression in the brain at E10. Statistical analysis compared expression at E10, E18, P1week and P1month. ***$P < 5 \times 10^{-4}$. (**b**) Changes in expression of *NPTN* across lifespan in the dorsolateral prefrontal cortex of individuals stratified by rs7171755 genotypes. Each subject is colored to indicate its rs7171755 genotype, with the thick dotted curves representing an estimate of the local mean (loess) of *NPTN* expression for each genotype as it varies across age. Only samples with RNA quality RIN ≥ 8 are displayed. (**c**) Statistical analysis of a subset of the data displayed in (b), visualizing differences in *NPTN* expression between rs7171755 genotypes in the postnatal brain (age ≥ 0.5-year). The *y*-axis represents *NPTN* expression after controlling for age, ethnicity and RNA quality (RIN). The *x*-axis represents genotype groups. For (b) and (c), rs7171755 genotypes: Dark blue = AA; light blue = AG; pink = GG.

Supplementary Tables 2 and 3

Supplementary Table 2: multi-loci based association of *NPTN* SNPs with regional cortical thickness

	Left hemisphere	Right hemisphere
	P	*P*
Frontal Lobe	**1.87E-08**	**0.0006**
Caudal anterior cingulate	0.6247	0.6687
Caudal middle-frontal	**0.0004**	0.2065
Frontal pole	0.7484	0.0271
Lateral orbitofrontal	0.0021	0.3724
Medial orbitofrontal	0.1370	0.4943
Paracentral	0.0029	0.1461
Pars opercularis	0.0274	0.2693
Pars orbitalis	**5.07E-08**	0.0035
Pars triangularis	**2.61E-06**	0.3705
Posterior cingulate	0.0192	0.1663
Precentral	**1.44E-06**	0.3993
Rostral anteriorcingulate	0.0160	0.7284
Rostral middle-frontal	0.0017	0.5446
Superior frontal	**8.79E-06**	0.1774
Parietal Lobe	**4.59E-05**	0.0073
Inferior parietal	0.0297	0.2226
Isthmus cingulate	0.1148	0.5894
Postcentral	0.0036	0.2230
Precuneus	0.0202	0.4940
Superior parietal	0.0021	0.2680
Supramarginal	**3.89E-06**	0.0759
Temporal Lobe	**1.97E-10**	0.0040
Bank of the STS	**0.0002**	0.0178
Entorhinal	0.0223	0.1873
Fusiform	**2.48E-06**	0.4460
Inferior temporal	**3.73E-06**	0.3509
Middle temporal	**1.50E-07**	0.0975
Parahippocampal	0.8372	0.5024
Superior temporal	**2.47E-07**	0.1355
Temporal pole	0.0621	0.8214
Transverse temporal	0.0656	0.4262
Occipital Lobe	0.0969	0.6926
Lateral occipital	0.0241	0.6560
Cuneus	0.1662	0.6329
Pericalcarine	0.3281	0.5655
Lingual	0.9132	0.6560

P-values were estimated using the Least Square Kernel Machine (LSKM) approach. Those remaining significant after Bonferroni correction for multiple testing are highlighted in bold.

Supplementary Table 3: Correlations between cortical thickness and indices of intellectual ability

Control Variables			Verbal IQ	Non-verbal IQ	Average school grade last term[a]
Centre, gender, age, stage of pubertal development, ethnicity, 4 MDS components & left cortical surface area	Right Mean Cortical Thickness	Correlation	.015	.060	-.038
		Significance (2-tailed)	.612	.041	.190
		df	1168	1168	1168
	Left Mean Cortical Thickness	Correlation	.029	.074	-.062
		Significance (2-tailed)	.320	.012	.033
		df	1168	1168	1168

[a]Grades were based on percentage of correct answers as follows, with lower grades indicating higher performance: A+ (93-100%) = 1; A (90-92%) = 2; B+ (87-89%) = 3; B (83-86%) = 4; B- (80-82%) = 5; C+ (77-79%) = 6; C (73-76%) = 7; C- (70-72%)= 8

Supplementary Figures Legends

Supplementary Figure 1: Quantile-quantile (Q-Q) plots for the associations shown in **Figure 2b**. *P*-values for all SNPs were plotted against their expected values from a (0, 1) uniform distribution. Genomic inflation factor = 1.0119

Supplementary Fig. 2: Manhattan plots of single SNPs associations with cortical thickness on the right hemisphere. SNP markers are plotted according to chromosomal location on the *x*-axis, while the *y*-axis -$\log_{10}$ (*P* values) indicate the significance of the additive effect of the number of minor alleles of each SNP on average cortical thickness for the right hemisphere.

Supplementary Fig. 3: Bootstrapping resampling for association of rs7171755 with cortical thickness at the indicated ROIs or with verbal and non-verbal IQ. For each phenotype, the upper panel shows the estimation of beta and the lower panel the empirical *P*-value of the estimated beta. All distributions of the estimated *P*-values significantly deviate from the uniform distribution expected from the NULL model.

Supplementary Fig. 4: Meta-analytic association of the *NPTN* SNP rs7171755 with brain volume in normal subjects of the ENIGMA datasets. Regional association and recombination rate plots for SNPs around rs7171755, genotyped in our sample. A vertical line indicates rs7171755, the most significantly associated SNP on the plot. The information box above rs7171755 includes its identifier, chromosome and basepair location, effect allele frequency, meta-analytic p-value, meta-analytic effect size and meta-analytic standard error. The left y-axis represents -$\log_{10}$ *P* values for

association with brain volume, the right y-axis represents the recombination rate, and the x-axis represents base-pair positions along the chromosome (human genome Build 36).

References

1. Lenroot RK, Schmitt JE, Ordaz SJ, Wallace GL, Neale MC, Lerch JP *et al.* Differences in genetic and environmental influences on the human cerebral cortex associated with development during childhood and adolescence. *Hum Brain Mapp* 2009; **30**(1)**:** 163-174.

2. Sowell ER, Thompson PM, Leonard CM, Welcome SE, Kan E, Toga AW. Longitudinal mapping of cortical thickness and brain growth in normal children. *J Neurosci* 2004; **24**(38)**:** 8223-8231.

3. Shaw P, Greenstein D, Lerch J, Clasen L, Lenroot R, Gogtay N *et al.* Intellectual ability and cortical development in children and adolescents. *Nature* 2006; **440**(7084)**:** 676-679.

4. Perrin JS, Herve PY, Leonard G, Perron M, Pike GB, Pitiot A *et al.* Growth of white matter in the adolescent brain: role of testosterone and androgen receptor. *J Neurosci* 2008; **28**(38)**:** 9519-9524.

5. Giedd JN, Blumenthal J, Jeffries NO, Castellanos FX, Liu H, Zijdenbos A *et al.* Brain development during childhood and adolescence: a longitudinal MRI study. *Nat Neurosci* 1999; **2**(10)**:** 861-863.

6. Gogtay N, Giedd JN, Lusk L, Hayashi KM, Greenstein D, Vaituzis AC *et al.*

Dynamic mapping of human cortical development during childhood through early adulthood. *Proc Natl Acad Sci U S A* 2004; **101**(21)**:** 8174-8179.

7. Thompson PM, Cannon TD, Narr KL, van Erp T, Poutanen VP, Huttunen M *et al.* Genetic influences on brain structure. *Nat Neurosci* 2001; **4**(12)**:** 1253-1258.

8. Yoon U, Fahim C, Perusse D, Evans AC. Lateralized genetic and environmental influences on human brain morphology of 8-year-old twins. *Neuroimage* 2010; **53**(3)**:** 1117-1125.

9. Joshi AA, Lepore N, Joshi SH, Lee AD, Barysheva M, Stein JL *et al.* The contribution of genes to cortical thickness and volume. *Neuroreport* 2011; **22**(3)**:** 101-105.

10. Panizzon MS, Fennema-Notestine C, Eyler LT, Jernigan TL, Prom-Wormley E, Neale M *et al.* Distinct genetic influences on cortical surface area and cortical thickness. *Cereb Cortex* 2009; **19**(11)**:** 2728-2735.

11. Huttenlocher PR. Synaptic density in human frontal cortex-developmental changes and effects of aging. *Brain Res* 1979; **163:** 195-205.

12. Yakovlev PI, Lecours AR. The myelogenetic cycles of regional maturation of the brain. In: Minkowski A (ed). *Regional Development of the Brain in Early*

Life. Blackwell Scientific Boston, MA, 1967, pp 3-70.

13. Rakic P. Specification of cerebral cortical areas. *Science* 1988; **241**(4862)**:** 170-176.

14. Pontious A, Kowalczyk T, Englund C, Hevner RF. Role of intermediate progenitor cells in cerebral cortex development. *Dev Neurosci* 2008; **30**(1-3)**:** 24-32.

15. 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; **15**(12)**:** 1128-1139.

16. Wechsler D. The Wechsler intelligence scale for children—fourth edition. Technical and Interpretive Manual. Psychological Association: San Antonio, TX, 2003.

17. Petersen AC, Crockett L, Richards M, Boxer A. A self-report measure of pubertal status: Reliability, validity, and initial norms. *Journal of Youth and Adolescence* 1988; **17**(2)**:** 117-133.

18. Pritchard JK, Stephens M, Donnelly P. Inference of population structure using multilocus genotype data. *Genetics* 2000; **155**(2)**:** 945-959.

19. 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; **81**(3)**:** 559-575.

20. Price AL, Patterson NJ, Plenge RM, Weinblatt ME, Shadick NA, Reich D. Principal components analysis corrects for stratification in genome-wide association studies. *Nat Genet* 2006; **38**(8)**:** 904-909.

21. Schilling C, Kuhn S, Paus T, Romanowski A, Banaschewski T, Barbot A *et al.* Cortical thickness of superior frontal cortex predicts impulsiveness and perceptual reasoning in adolescence. *Mol Psychiatry* 2012.

22. Cocks G, Romanyuk N, Amemori T, Jendelova P, Forostyak O, Jeffries AR *et al.* Conditionally immortalized stem cell lines from human spinal cord retain regional identity and generate functional V2a interneurons and motorneurons. *Stem Cell Res Ther* 2013; **4**(3)**:** 69.

23. Pollock K, Stroemer P, Patel S, Stevanato L, Hope A, Miljan E *et al.* A conditionally immortal clonal stem cell line from human cortical neuroepithelium for the treatment of ischemic stroke. *Exp Neurol* 2006; **199**(1)**:** 143-155.

24. Huang DW, Sherman BT, Lempicki RA. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat Protoc* 2009; **4**(1)**:**

44-57.

25. Yang J, Lee SH, Goddard ME, Visscher PM. GCTA: a tool for genome-wide complex trait analysis. *Am J Hum Genet* 2011; **88**(1)**:** 76-82.

26. Johnson AD, Handsaker RE, Pulit SL, Nizzari MM, O'Donnell CJ, de Bakker PI. SNAP: a web-based tool for identification and annotation of proxy SNPs using HapMap. *Bioinformatics* 2008; **24**(24)**:** 2938-2939.

27. Hayes AF. *Introduction to Mediation, Moderation, and Conditional Process Analysis: A Regression-Based Approach*. 1 edition (17 Jun 2013) edn. The Guilford Press: New York, USA, 2013.

28. Novak NM, Stein JL, Medland SE, Hibar DP, Thompson PM, Toga AW. EnigmaVis: online interactive visualization of genome-wide association studies of the Enhancing NeuroImaging Genetics through Meta-Analysis (ENIGMA) consortium. *Twin Res Hum Genet* 2012; **15**(3)**:** 414-418.

29. Sun L, Bull SB. Reduction of selection bias in genomewide studies by resampling. *Genet Epidemiol* 2005; **28**(4)**:** 352-367.

30. Ge T, Feng J, Hibar DP, Thompson PM, Nichols TE. Increasing power for voxel-wise genome-wide association studies: the random field theory, least square kernel machines and fast permutation procedures. *Neuroimage* 2012; **63**(2)**:**

858-873.

31. Colantuoni C, Lipska BK, Ye T, Hyde TM, Tao R, Leek JT *et al.* Temporal dynamics and genetic control of transcription in the human prefrontal cortex. *Nature* 2011; **478**(7370)**:** 519-523.

32. Kang HJ, Kawasawa YI, Cheng F, Zhu Y, Xu X, Li M *et al.* Spatio-temporal transcriptome of the human brain. *Nature* 2011; **478**(7370)**:** 483-489.

33. Galaburda AM, LeMay M, Kemper TL, Geschwind N. Right-left asymmetrics in the brain. *Science* 1978; **199**(4331)**:** 852-856.

34. Toga AW, Thompson PM. Mapping brain asymmetry. *Nat Rev Neurosci* 2003; **4**(1)**:** 37-48.

35. Sun T, Patoine C, Abu-Khalil A, Visvader J, Sum E, Cherry TJ *et al.* Early asymmetry of gene transcription in embryonic human left and right cerebral cortex. *Science* 2005; **308**(5729)**:** 1794-1798.

36. Chang S, Johnston RJ, Jr., Hobert O. A transcriptional regulatory cascade that controls left/right asymmetry in chemosensory neurons of C. elegans. *Genes & development* 2003; **17**(17)**:** 2123-2137.

37. Hagmann P, Cammoun L, Gigandet X, Meuli R, Honey CJ, Wedeen VJ *et al.*

Mapping the structural core of human cerebral cortex. *PLoS Biol* 2008; **6**(7)**:** e159.

38. Kwee LC, Liu D, Lin X, Ghosh D, Epstein MP. A powerful and flexible multilocus association test for quantitative traits. *Am J Hum Genet* 2008; **82**(2)**:** 386-397.

39. Liu D, Lin X, Ghosh D. Semiparametric regression of multidimensional genetic pathway data: least-squares kernel machines and linear mixed models. *Biometrics* 2007; **63**(4)**:** 1079-1088.

40. Faye LL, Sun L, Dimitromanolakis A, Bull SB. A flexible genome-wide bootstrap method that accounts for ranking and threshold-selection bias in GWAS interpretation and replication study design. *Stat Med* 2011; **30**(15)**:** 1898-1912.

41. Tversky A, Kahneman D. Belief in the law of small numbers. *Psychol Bull* 1971; **75**(2)**:** 105-110.

42. Stein JL, Medland SE, Vasquez AA, Hibar DP, Senstad RE, Winkler AM *et al.* Identification of common variants associated with human hippocampal and intracranial volumes. *Nat Genet* 2012; **44**(5)**:** 552-561.

43. Hill IE, Selkirk CP, Hawkes RB, Beesley PW. Characterization of novel

glycoprotein components of synaptic membranes and postsynaptic densities, gp65 and gp55, with a monoclonal antibody. *Brain Res* 1988; **461**(1)**:** 27-43.

44. Dapretto M, Bookheimer SY. Form and content: dissociating syntax and semantics in sentence comprehension. *Neuron* 1999; **24**(2)**:** 427-432.

45. Bearden CE, van Erp TG, Dutton RA, Tran H, Zimmermann L, Sun D *et al.* Mapping cortical thickness in children with 22q11.2 deletions. *Cereb Cortex* 2007; **17**(8)**:** 1889-1898.

46. Gerdes M, Solot C, Wang PP, Moss E, LaRossa D, Randall P *et al.* Cognitive and behavior profile of preschool children with chromosome 22q11.2 deletion. *Am J Med Genet* 1999; **85**(2)**:** 127-133.

47. Smalla KH, Matthies H, Langnase K, Shabir S, Bockers TM, Wyneken U *et al.* The synaptic glycoprotein neuroplastin is involved in long-term potentiation at hippocampal CA1 synapses. *Proc Natl Acad Sci U S A* 2000; **97**(8)**:** 4327-4332.

48. Empson RM, Buckby LE, Kraus M, Bates KJ, Crompton MR, Gundelfinger ED *et al.* The cell adhesion molecule neuroplastin-65 inhibits hippocampal long-term potentiation via a mitogen-activated protein kinase p38-dependent reduction in surface expression of GluR1-containing glutamate receptors. *J Neurochem* 2006; **99**(3)**:** 850-860.

49. Owczarek S, Soroka V, Kiryushko D, Larsen MH, Yuan Q, Sandi C *et al.* Neuroplastin-65 and a mimetic peptide derived from its homophilic binding site modulate neuritogenesis and neuronal plasticity. *J Neurochem* 2011; **117**(6)**:** 984-994.

50. Vernes SC, Oliver PL, Spiteri E, Lockstone HE, Puliyadi R, Taylor JM *et al.* Foxp2 regulates gene networks implicated in neurite outgrowth in the developing brain. *PLoS Genet* 2011; **7**(7)**:** e1002145.

51. Lai CS, Fisher SE, Hurst JA, Vargha-Khadem F, Monaco AP. A forkhead-domain gene is mutated in a severe speech and language disorder. *Nature* 2001; **413**(6855)**:** 519-523.

52. Groszer M, Keays DA, Deacon RM, de Bono JP, Prasad-Mulcare S, Gaub S *et al.* Impaired synaptic plasticity and motor learning in mice with a point mutation implicated in human speech deficits. *Curr Biol* 2008; **18**(5)**:** 354-362.

53. Shu W, Cho JY, Jiang Y, Zhang M, Weisz D, Elder GA *et al.* Altered ultrasonic vocalization in mice with a disruption in the Foxp2 gene. *Proc Natl Acad Sci U S A* 2005; **102**(27)**:** 9643-9648.

54. Haesler S, Wada K, Nshdejan A, Morrisey EE, Lints T, Jarvis ED *et al.* FoxP2 expression in avian vocal learners and non-learners. *J Neurosci* 2004; **24**(13)**:**

3164-3175.

55. Saito A, Fujikura-Ouchi Y, Kuramasu A, Shimoda K, Akiyama K, Matsuoka H *et al.* Association study of putative promoter polymorphisms in the neuroplastin gene and schizophrenia. *Neurosci Lett* 2007; **411**(3)**:** 168-173.

56. Huttenlocher PR. Dendritic and synaptic pathology in mental retardation. *Pediatr Neurol* 1991; **7**(2)**:** 79-85.

57. Purpura DP. Dendritic spine "dysgenesis" and mental retardation. *Science* 1974; **186**(4169)**:** 1126-1128.

58. Gilman SR, Chang J, Xu B, Bawa TS, Gogos JA, Karayiorgou M *et al.* Diverse types of genetic variation converge on functional gene networks involved in schizophrenia. *Nat Neurosci* 2012.

59. van Bokhoven H. Genetic and epigenetic networks in intellectual disabilities. *Annu Rev Genet* 2011; **45:** 81-104.

60. Gilman SR, Iossifov I, Levy D, Ronemus M, Wigler M, Vitkup D. Rare de novo variants associated with autism implicate a large functional network of genes involved in formation and function of synapses. *Neuron* 2011; **70**(5)**:** 898-907.

61. Clement JP, Aceti M, Creson TK, Ozkan ED, Shi Y, Reish NJ *et al.* Pathogenic

SYNGAP1 Mutations Impair Cognitive Development by Disrupting Maturation of Dendritic Spine Synapses. *Cell* 2012; **151**(4)**:** 709-723.

Figure 1

A

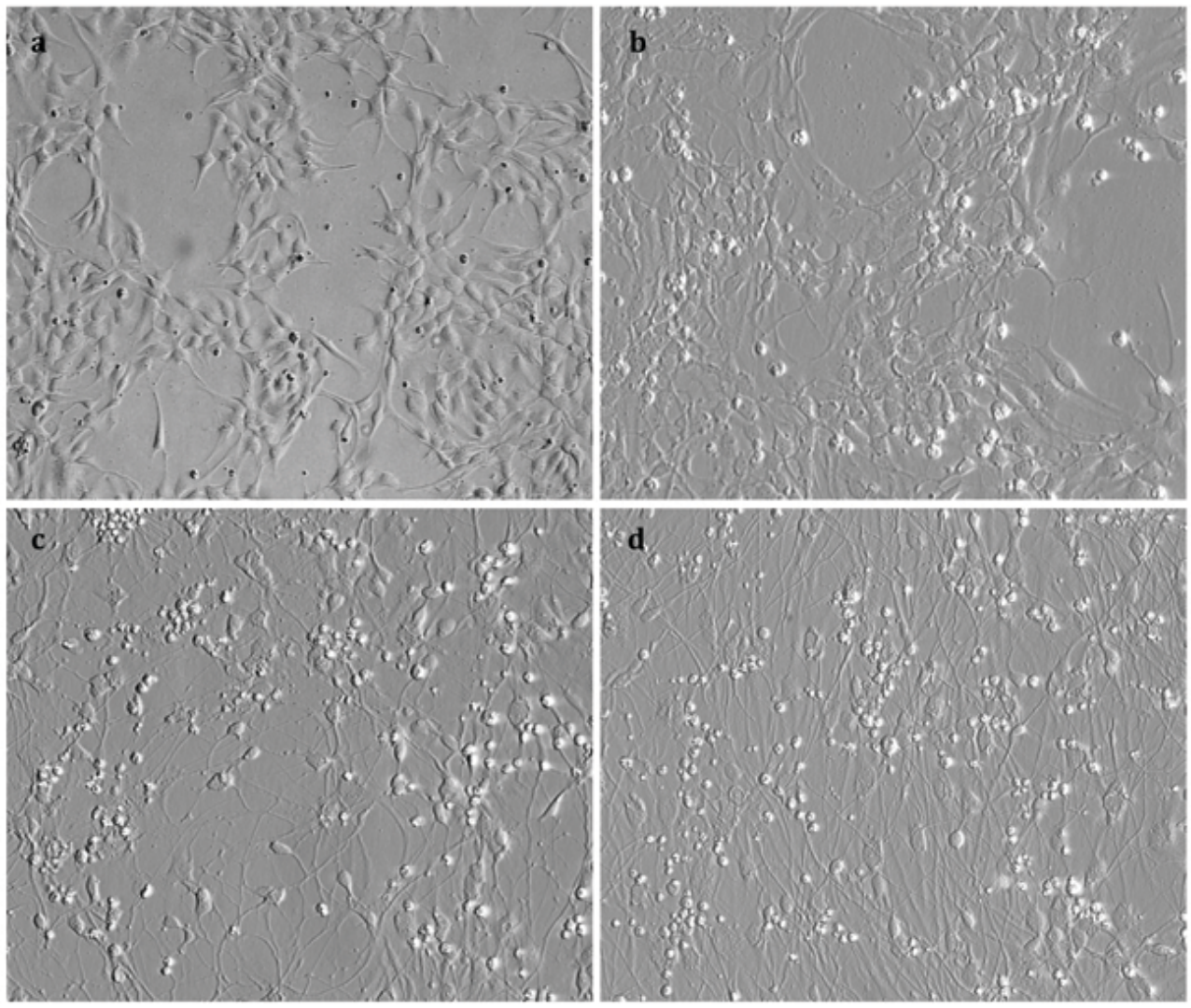


Figure 1

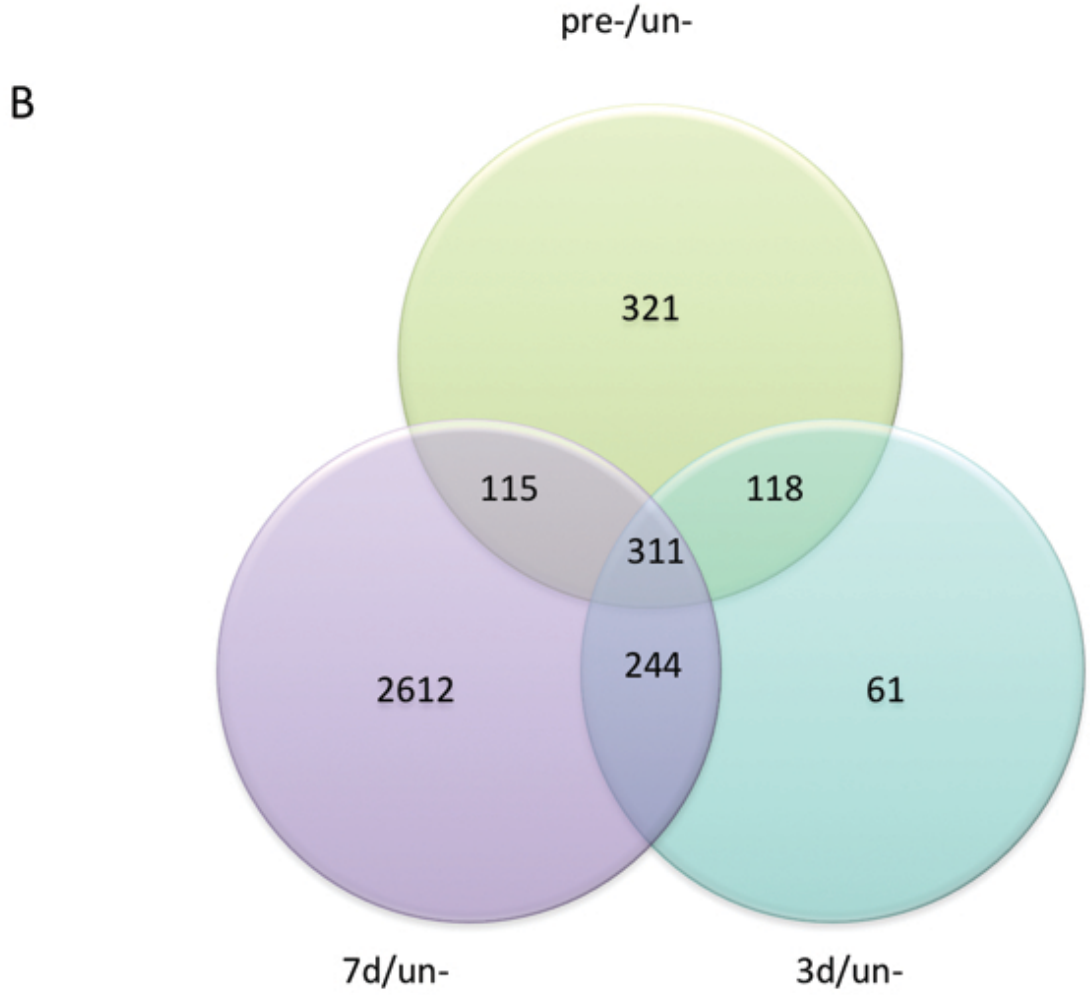


Figure 2A

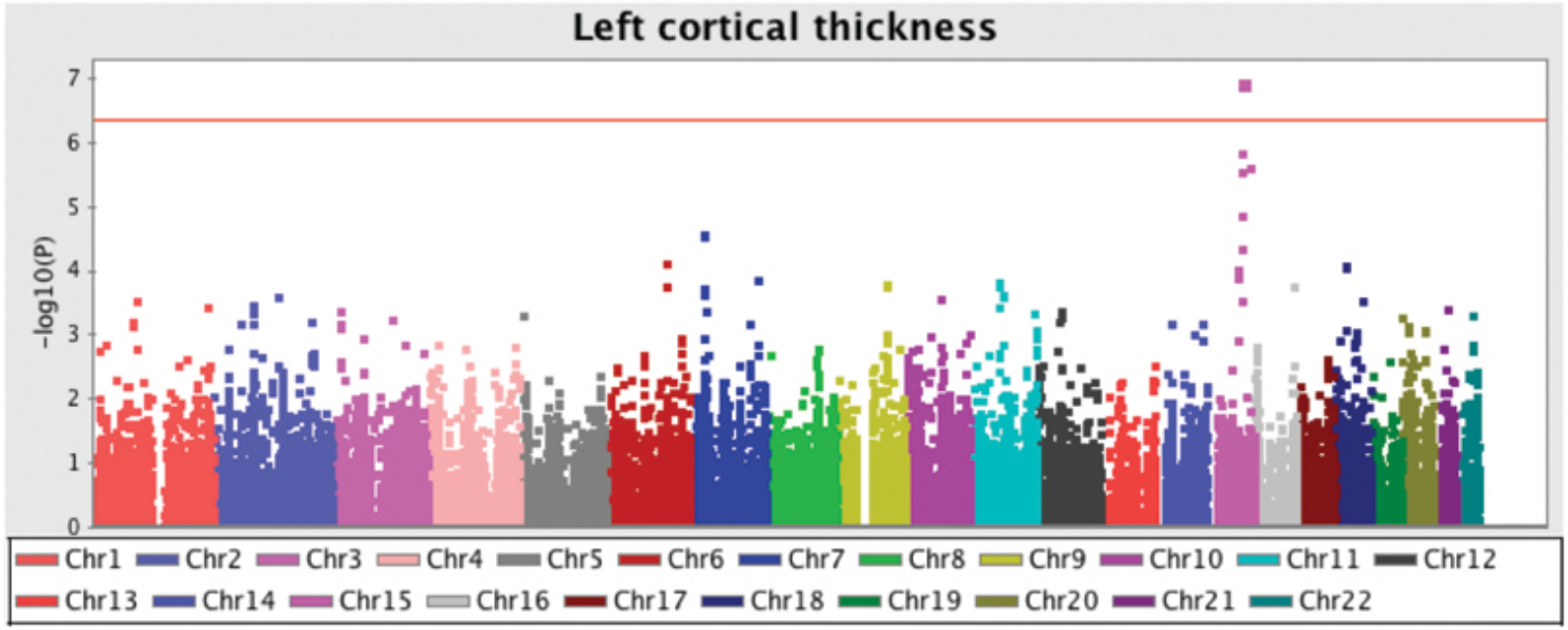

Left cortical thickness
-log10(P)
0
1
2
3
4
5
6
7
Chr1
Chr2
Chr3
Chr4
Chr5
Chr6
Chr7
Chr8
Chr9
Chr10
Chr11
Chr12
Chr13
Chr14
Chr15
Chr16
Chr17
Chr18
Chr19
Chr20
Chr21
Chr22

Figure 2B

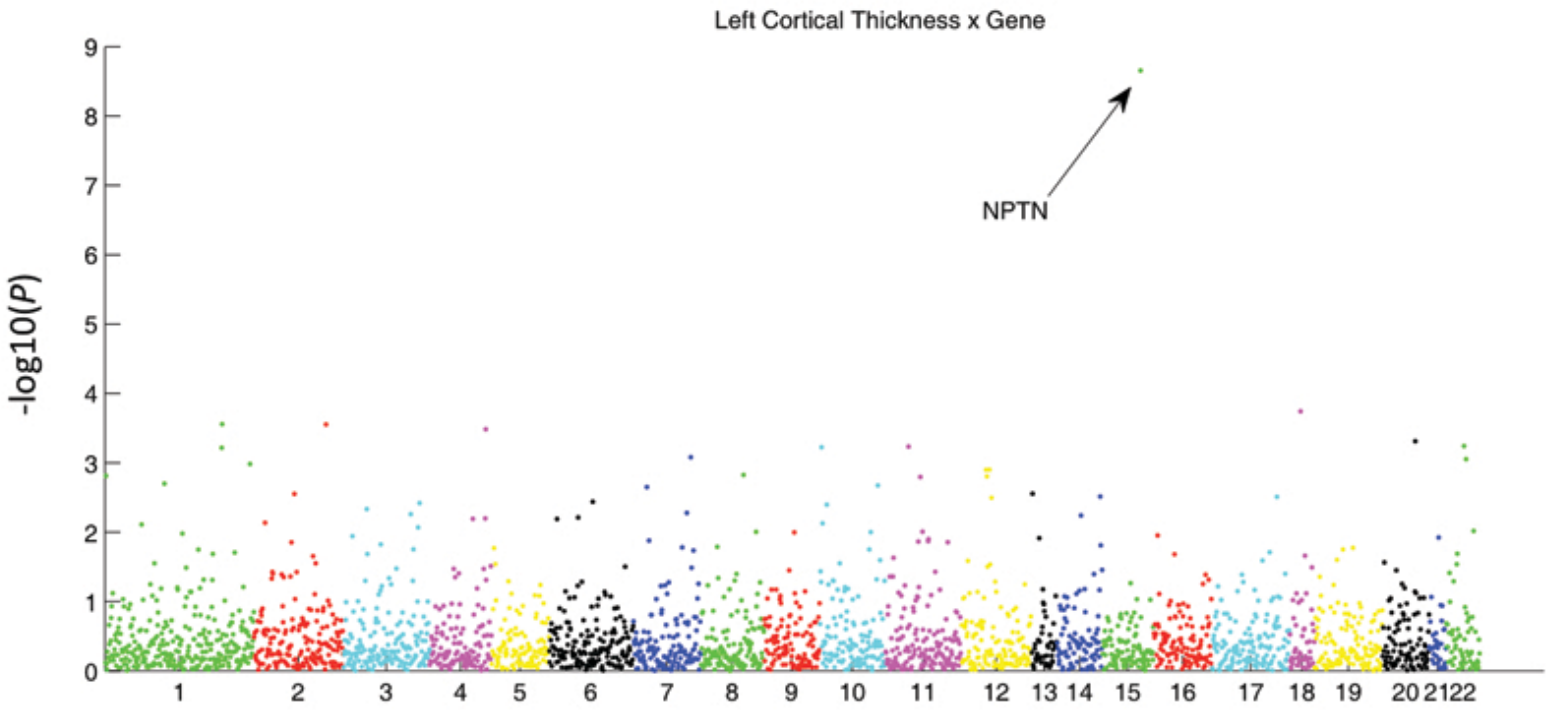

Left Cortical Thickness x Gene
NPTN
-log10(P)
0
1
2
3
4
5
6
7
8
9
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22

Figure 3

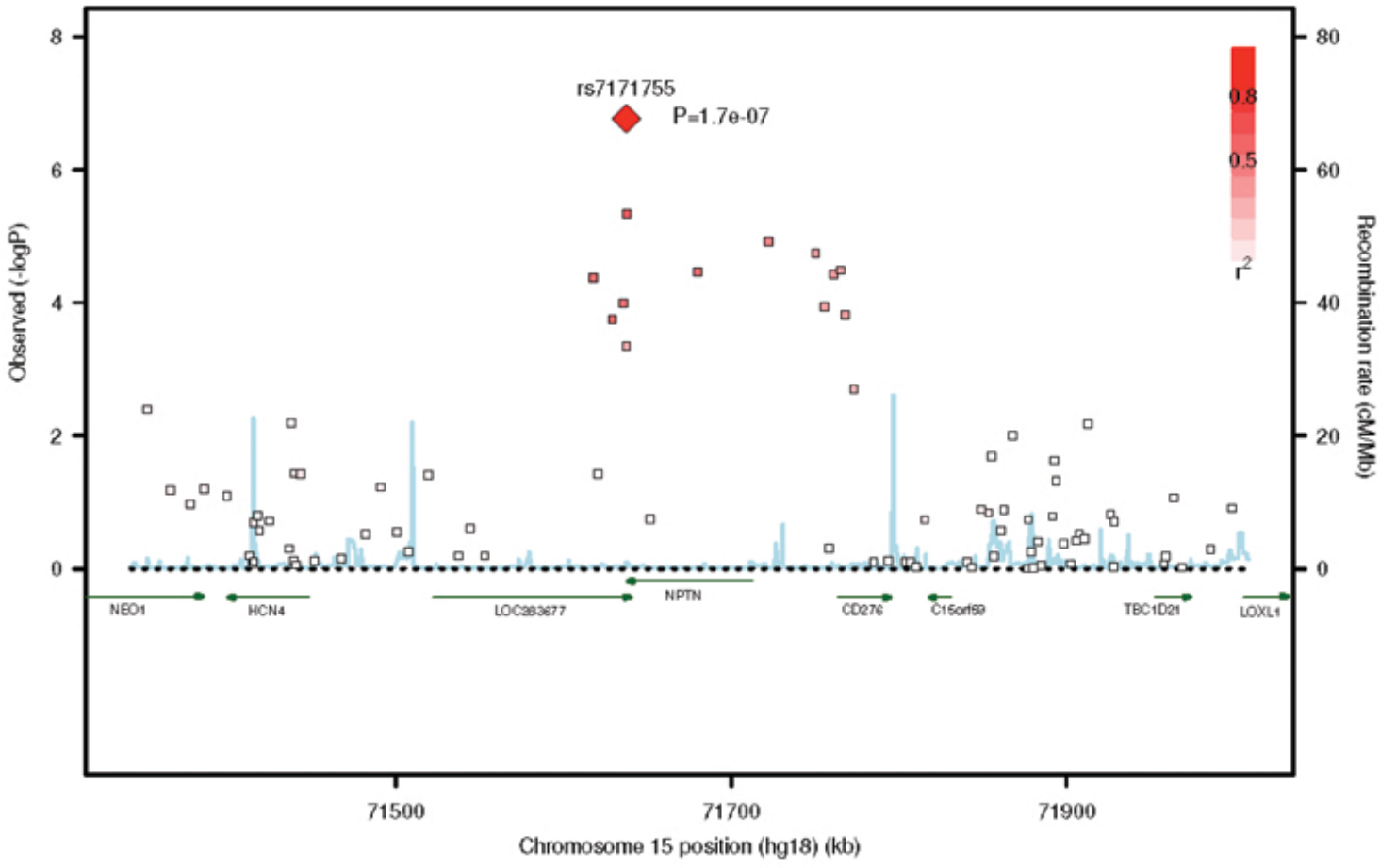

rs7171755
P=1.7e-07
0.8
0.5
r2
Observed (-logP)
Recombination rate (cM/Mb)
8
6
4
2
0
80
60
40
20
0
NEO1
HCN4
LOC283677
NPTN
CD276
C15orf59
TBC1D21
LOXL1
71500
71700
71900
Chromosome 15 position (hg18) (kb)

Figure 4A

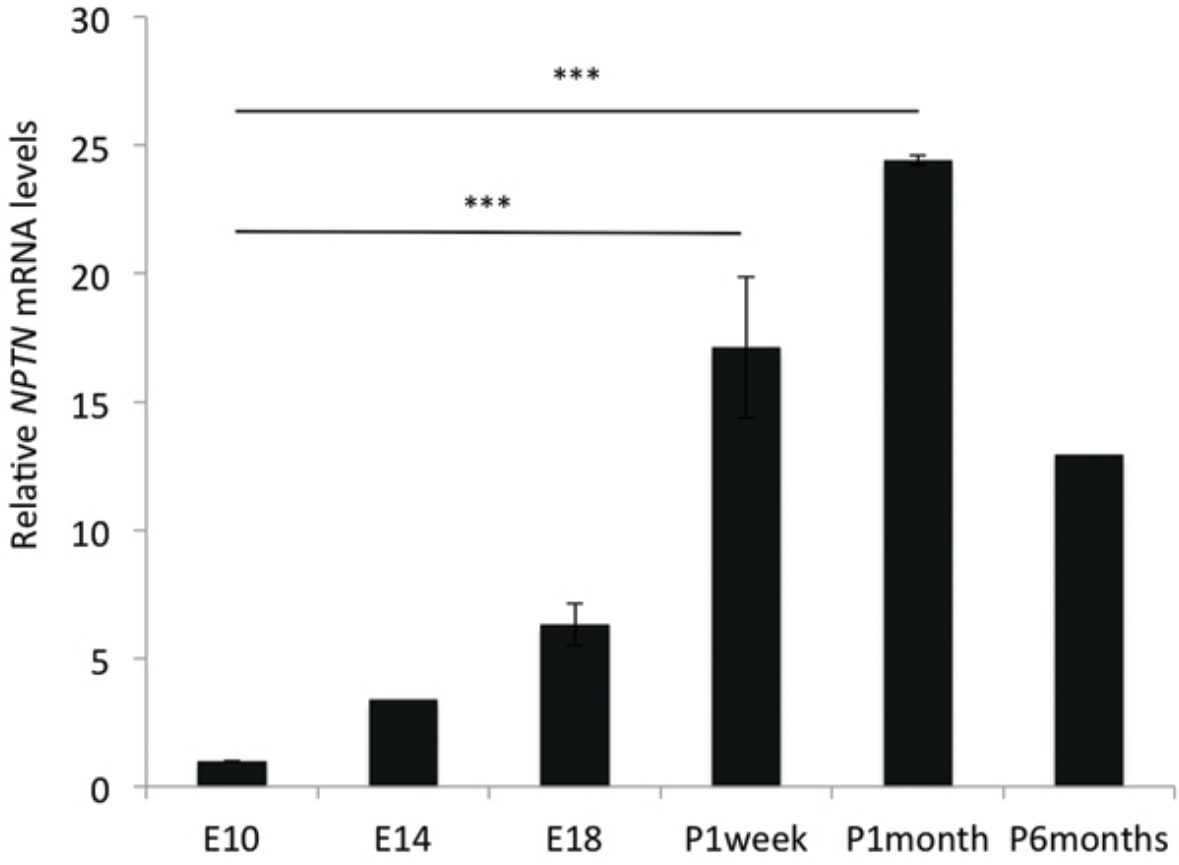

Relative NPTN mRNA levels
30
25
20
15
10
5
0
***
***
E10
E14
E18
P1week
P1month
P6months

Figure 4B

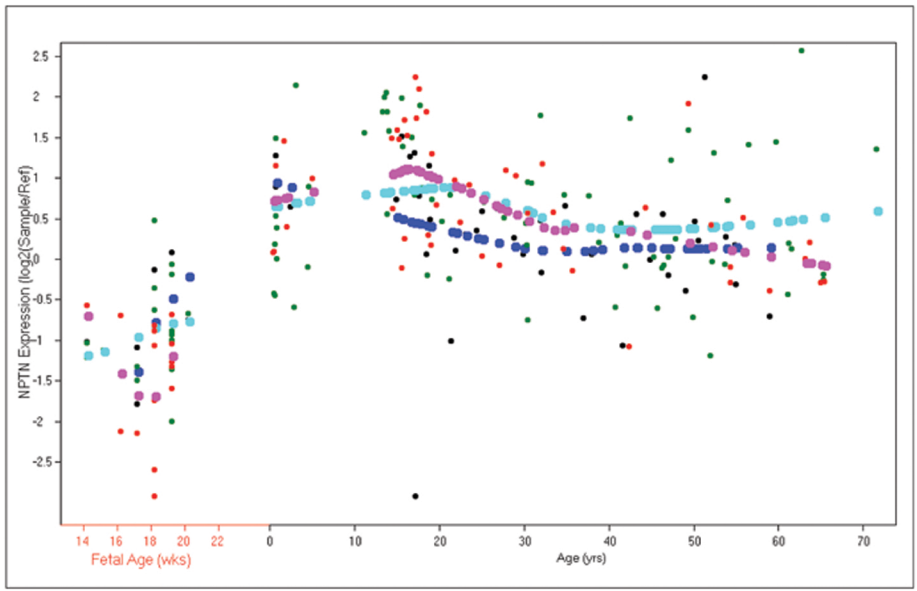
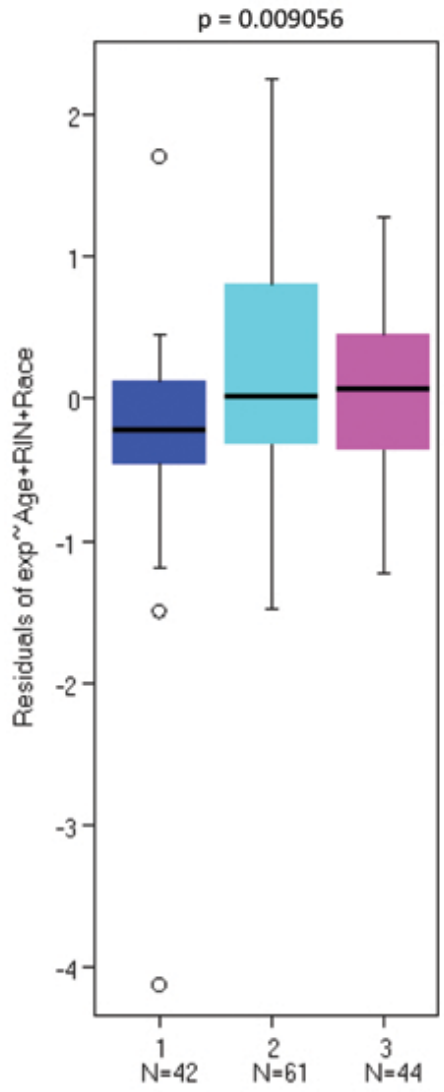

NPTN Expression (log2(Sample/Ref)
2.5
2
1.5
1
0.5
0
-0.5
-1
-1.5
-2
-2.5
14
16
18
20
22
Fetal Age (wks)
0
10
20
30
40
50
60
70
Age (yrs)

Figure 4C


p = 0.009056
Residuals of exp~Age+RIN+Race
2
1
0
-1
-2
-3
-4
1
N=42
2
N=61
3
N=44