Supplementary Information

Supplemental Methods

Bioinformatic modeling of molecular effects of *GSK-3β* variants

All 1703 SNPs within flanking sequences **within *GSK-3β*** and minor allele frequencies (MAF) were extracted from NCBI dbSNP (release 131) and HapMap (phase 3 release 3) databases (www.hapmap.org). Promoter sequences were downloaded from UCSC Genome Browser tool (version hg19). We considered as promoter the region 2000 bp upstream and 500 bp downstream the transcription start site as adopted in a previous work (1). Effects of polymorphisms in the promoter region were predicted by using rVISTA 2.0 (http://rvista.dcode.org). SpliceAid 2 was used to predict position of splicing regulatory sequences. This tool adopts only experimentally assessed target RNA sequences in humans, thus minimizing false positive predictions (2). 5' and 3' splice site detection was performed by NNSPLICE (http://www.fruitfly.org/seq_tools/splice.html). To assess compatibility of aminoacidic substitution, we used rBLOSUM64 (the recently revised BLOSUM matrix), SIFT and PolyPhen, as previously described (1). **We used these different tools because there is no evidence for greater reliability of one algorithm relative to another. Thus, we explored these tools and we report detection of functional roles of genetic variants.** To predict miRNA target site alterations in the gene 3′UTR regions, we used miRBase (http://microrna.sanger.ac.uk/). All predictions were filtered by brain expression data using both transcriptomic and proteomic data. With this aim, ArrayExpress (www.ebi.ac.uk/microarray-as/ae/), Human Transcriptome Map (http://bioinfo.amc.uva.nl/HTMseq/controller), Human Protein Atlas (www.proteinatlas.org/) and Human Protein Reference Database (www.hprd.org/) were used. **In particular, we used as a first option protein expression data as they are the most reliable. If protein expression data were not available, we used mRNA expression data.**

Association of *GSK-3β* SNPs with GSK-3β mRNA expression in post-mortem prefrontal cortex of healthy subjects

The publicly available collection used in the present study included post-mortem human brains taken from autopsy primarily from the Offices of the Chief Medical Examiner of the District of Columbia, and of the Commonwealth of Virginia, Northern District, all with informed consent from the legal next of kin (protocol 90-M-0142 approved by the NIMH/NIH Institutional Review Board). Additional post-mortem fetal, infant, child and adolescent brain tissue samples were provided by the National Institute of Child Health and Human Development Brain and Tissue Bank for Developmental Disorders (http://www.BTBank.org) under contracts NO1-HD-4-3368 andNO1-HD-4-3383. The Institutional Review Board of the University of Maryland at Baltimore and the State of Maryland approved the protocol, and the tissue was donated to the NIMH under the terms of a Material Transfer Agreement. Clinical characterization, diagnoses, and macro- and microscopic neuropathological examinations were performed on all CBDB cases using a standardized paradigm. Details of tissue acquisition, handling, processing, dissection, clinical characterization, diagnoses, neuropathological examinations, RNA extraction and quality control measures were described previously (3-4). The Brain and Tissue Bank cases were handled in a similar fashion (http://medschool.umaryland.edu/BTBank/ProtocolMethods.html). Toxicological analysis was performed on every case. Subjects with evidence of macro- or microscopic neuropathology, drug use, alcohol abuse, or psychiatric illness were excluded.

From each subject in the brain collection, RNA from prefrontal grey matter was analyzed using spotted oligonucleotide microarrays yielding data from 30,176 gene expression probes and allowing us to focus on GSK-3β mRNA expression (accession number: NM__002093; probe position: chr3:121028292-121028361). In particular, total RNA was extracted, amplified and fluorescently labeled. Reference RNA was pooled from all samples and treated identically to sample RNAs. Labeled RNAs were hybridized to two-color custom-spotted arrays from the NHGRI microarray core facility. After normalization (5), $\log_2$ intensity ratios were further adjusted to

reduce the impact of known and unknown sources of systematic noise on gene expression measures using surrogate variable analysis (6).

DNA from cerebellar tissue was studied with Illumina BeadChips producing 625,439 SNP genotypes called using the BeadExpress software for each subject as previously described (4).

Association of *GSK-3β*rs12630592 with GSK-3β expression in PBMCs, prefrontal activity and prefrontal cortical thickness in healthy humans

Subjects and genotyping

Healthy subjects were evaluated with the Structured Clinical Interview for DSM-IV (SCID) (7) to exclude any psychiatric disorder. Further exclusion criteria were history of drug or alcohol abuse, active drug use in the past year, head trauma with loss of consciousness, and any significant medical condition. All subjects were genotyped for *GSK-3β* rs12630592 by Real-Time PCR with TaqMan® SNP Genotyping Assay (Assay ID: C___1849026_20; Applied Biosystems). Genotype groups did not display Hardy-Weinberg equilibrium (minor allele frequency=0.42; p=0.005). This result is consistent with the allele frequency found in another smaller Italian sample (Tuscans, HAP-TSI, http://hapmap.ncbi.nlm.nih.gov/; minor allele frequency=0.41; Hardy-Weinberg p=0.06). However, to exclude the potential for errors, genotyping was double checked and repeated in two independent labs in 192 and 63 subjects respectively. No discordant calls were obtained.

Association of *GSK-3β* rs12630592 with GSK-3β and β-catenin protein levels and GSK-3β phosphorylation in peripheral blood mononuclear cells (PBMCs) of healthy humans

- Data acquisition: Blood samples were collected between 8:30 and 9:30 AM. All individuals were not affected by any metabolic diseases. Furthermore, all subjects were asked to avoid heavy meals, sleepless nights and intense use of caffeine in the 48 hours before the

drawing. Moreover, all individuals included were non-smokers, thus preventing potential effects of nicotine on AKT1 phosphorylation (8) and thus on GSK-3β signaling.

- PBMCs isolation: PBMCs were isolated from blood samples by Ficoll density gradient (ICN Biomedical, Inc.). Cells were rapidly rinsed in ice-cold PBS and solubilized in Triton X-100 lysis buffer (10 mM Tris HCl, pH 7.4, 150 mM NaCl, 1% Triton X-100, 1 mM EDTA, 10% glycerol, 1mM phenylmethylsulphonyl fluoride, 10 µg/ml leupeptin, 10 µg/ml aprotinin, 1 mM sodium orthovanadate, 50 mM sodium fluoride and 10 mM β-glicerophosphate).

- Protein determination: aliquots (2µl) of samples were used for the protein determination by Bio-Rad Protein Assay kit (Bio-Rad, Hercules, CA). Equal amounts of total proteins for each sample were loaded onto 10% polyacrylamide gels. Proteins were separated by SDS-PAGE and transferred overnight to membranes (PVDF) (Amersham Pharmacia Biotech, Uppsala, Sweden). Membranes were immunoblotted overnight using selective antibodies against GSK-3β (1:500, Cell Signaling Technology), P-Ser9-GSK-3β (1:500, Cell Signaling Technology), β-catenin (1:250, Cell Signaling Technology). Selective antibodies against α-tubulin were also used (1:1000, Sigma, St. Louis, MO). Blots were then incubated in horseradish peroxidase-conjugated secondary antibodies and target proteins visualized by ECL detection (Pierce, Rockfors, IL), followed by quantification by Quantity One software (Bio-Rad).

Association of *GSK-3β* rs12630592 with prefrontal physiology during cognition

fMRI tasks

N-back**:** "N-back" refers to how far back in the sequence of stimuli that the subject had to recall. The stimuli consisted of numbers (1–4) shown in random sequence and displayed at the points of a diamond-shaped box (9). There was a non-memory-guided control condition (0-back) that presented the same stimuli but simply required subjects to identify the stimulus currently seen. As memory load increased, the task required the recollection of a stimulus seen one (1-back), two (2-back), or three stimuli (3-Back) beforehand while continuing to encode additionally incoming

stimuli. We used a simple block design in which each block consisted of eight alternating 0- and 1-, 2-back or 3-back (each lasting 30 seconds). Each task combination was obtained in 4 minutes and 8 seconds.

VAC: Each stimulus was composed of arrows of three different sizes pointing either to the right or to the left; small arrows were embedded in medium sized arrows, that were in turn embedded in a large arrow (10-13). Subjects were instructed by a cue word (big, medium or small) displayed above each stimulus to press a button corresponding to the direction of the large, medium or small arrows (right or left). To increase the level of attentional control required, the direction of the arrows were congruent or incongruent across all three sizes. This procedure resulted in the following conditions:

- Low level of attentional control (LOW): all three sizes of arrows were congruent in direction with each other. The cue was the word BIG.
- Intermediate level of attentional control (INT). Two stimuli were used: the big arrow was incongruent in direction to the small and the medium arrows in both stimuli; the cue was BIG in one stimulus, SMALL in the other one.
- High level of attentional control (HIGH). Two stimuli were used: the medium size arrows were incongruent in direction to the big and the small arrows in both stimuli; the cue was SMALL in one stimulus, MEDIUM in the other one.
- a simple bold arrow pointing either to the left or right, was used as a sensorimotor condition.

Subjects were instructed to respond to task stimuli with the right hand using a button box (right button for 'right' response, left button for 'left' response), and to press the response button as fast and accurately as possible. Furthermore, they were asked to move their thumb to a small plastic knob placed between buttons after each response. All subjects were trained on the task prior to the fMRI session. Each stimulus was presented for 800 ms, and the order of the stimuli was randomly distributed across the session. The total number of stimuli was 241: 50 HIGH (25 stimuli of each of the two stimulus types that subtended this level of conflict), 68 INT (34 stimuli of each of the two

stimulus types that subtended this level of conflict), 57 LOW, and 66 simple bold arrows; the total duration of the task was 10 minutes, 8 seconds. A fixation cross-hair was presented during the interstimulus interval, which ranged from 2000 to 6000 ms.

For both tasks, stimuli were presented via a back-projection system. Responses were recorded through a fiber optic response box allowing measurement of accuracy and reaction time.

fMRI data acquisition and preprocessing

fMRI data were acquired with a 3.0 Tesla General Electric MR scanner with gradient-echo-planar-imaging sequences, as follows: N-back: repetition time (TR)/echo time (TE): 2000/30 ms; matrix: 64x64; field of view: 24 cm; 20 interleaved slices; thickness=4mm, gap=1mm; voxel size 3.75x3.75x5mm. scans=120. VAC: TR/TE=2000/30; matrix = 64x64; field of view = 24cm; 26 interleaved slices; thickness=4mm, gap=1mm; voxel size 3.75x3.75x5mm; scans=300).

Images for each subject were realigned to the first volume in the time series to correct for head motion, spatially normalized into a standard stereotactic space (Montreal Neurological Institute, MNI, template) using a 12 parameter affine model and spatially smoothed (10 mm Gaussian filter). After realignment, data sets were also screened for small motion correction (<2mm in translation, <1.5° in rotation).

For each experimental condition of the N-back task(9), a box car model convolved with the haemodynamic response function at each voxel was modeled. To account for potential differences in head movement between groups, residual movement was also modeled as a regressor of no interest. Pre-determined condition effects at each voxel were calculated using a t-statistic, producing a statistical image for the contrast of 3-back versus 0-back, 2-back versus 0-back, 1-back versus 0-back.

fMRI responses during the VAC task (10-13) were modeled using a canonical hemodynamic response function and temporally filtered using a high-pass filter of 128 Hz and a hrf-shape low-

pass filter. Vectors were created for each condition using the timing of correct responses. To account for potential differences in head movement between groups, residual movement was also modeled as a regressor of no interest. A *t* statistic was then used to produce a statistical image for BOLD responses relative to brain processing of stimuli for each level of attentional control (high level - HIGH, intermediate level - INT, and low level - LOW).

Group analyses were constrained by masks obtained by combining group activation maps of all genotype groups ($p<0.05$). Brodmann's areas were assigned to activated clusters using the Talairach Daemon (http://ric.uthscsa.edu/projects/talairachdaemon.html) after converting the MNI coordinates of the local maxima in the activated clusters to Talairach coordinates (http://www.mrc-cbu.cam.ac.uk/Imaging/Common/mnispace.shtml).

Association of *GSK-3β* rs12630592 with prefrontal cortical thickness

3D images were acquired using a T1-weighted SPGR sequence (TE=min full; flip angle= 6°; bandwidth=31.25; matrix size=256x256; field of view=25x25 cm; 124 axial slices; slice thickness=1.3 mm) and processed with Freesurfer (v. 5.0.0) image analysis suite (http://surfer.nmr.mgh.harvard.edu/),which allows to obtain maps of cortical thickness created using spatial intensity gradients across tissue classes and are therefore not simply reliant on absolute signal intensity. Procedures for the measurement of cortical thickness have been validated against histological analysis(14) and manual measurements (15-16). Details of image processing in Freesurfer have been described in prior publications (17-18, 19, 20-23). Briefly, this procedure involves motion correction, removal of non brain tissue using a hybrid watershed/surface deformation procedure, automated Talairach transformation, segmentation of the subcortical white matter and deep gray matter volumetric structures, intensity normalization, tessellation of the gray matter white matter boundary, automated topology correction, and surface deformation following intensity gradients to optimally place the gray/white and gray/CSF borders at the location where the greatest shift in intensity defines the transition to the other tissue class. Once the cortical models are

complete, a number of deformable procedures are performed for further data processing and analysis, including surface inflation, registration to a spherical atlas, which used individual cortical folding patterns to match cortical geometry across subjects, parcellation of the cerebral cortex into units based on gyral and sulcal structure, and creation of a variety of surface-based data. This method uses both intensity and continuity information from the entire three-dimensional MR volume in segmentation and deformation procedures to produce representations of cortical thickness, calculated as the closest distance from the gray/white boundary to the gray/CSF boundary at each vertex on the tessellated surface.

All reconstructed surfaces were visually inspected and any inaccuracy in Talairach-transformed, skull stripped and segmentation were manually corrected and re-inspected.

Supplementary Results

Bioinformatic modeling of molecular effects of *GSK-3β* variants

We analyzed all the 1703 polymorphism lying in the genomic sequence of *GSK-3β* (NCBI dbSNP - release 131, HapMap - phase 3 release 3, UCSC Genome Browser - version hg19) in order to select those that might be potentially functional. These analyses indicated that 7 of these SNPs may alter micro-RNA binding, 36 may affect splicing, and 4 may alter binding of transcription factors. These splicing and promoter allelic variants may alter the quantity of the transcript. Furthermore, 2 SNPs may cause the substitution of aminoacids with others predicted to have medium biochemical compatibility, while 1 SNP may cause a frameshift.

Investigation of race effect on the association between rs12630592 and GSK-3β mRNA expression

Factorial ANOVA with rs12630592 and race as predictors indicated a main effect of rs12630592 genotype (F=5.3; p=0.005), with GG individuals having greater *GSK-3β* expression relative to GT (Fisher' post hoc test p=0.002) and TT subjects (p=0.00001). No main effect of race (p=0.28) or interaction between genotype and race (p=0.27) were found.

Association of *GSK-3β* rs12630592 with prefrontal cortical thickness in healthy humans

Supplemental Table 1: prefrontal clusters showing association of *GSK-3β* rs12630592 with cortical thickness. Only clusters with corrected statistics are indicated.

	Vertex number at the maximum	Talairach coordinates	Prefrontal region	Surface area (mm2)	Number of cluster vertices	Maximum p *value* (log_{10})	Corrected p
GG>TT	152189	x 42 y 38 z 26	Middle frontal gyrus	92.6	148	3.0988	0.0001
	76109	x -39 y 2 z 50	Middle frontal gyrus	154.0	258	3.4427	0.0001
GG>GT	83812	x -21 y 26 z -15	Orbito-frontal gyrus	209.6	385	4.5290	0.001
GT>TT	19927	x 34 y 6 z 32	Middle frontal gyrus	125.0	250	3.0663	0.0001

Supplemental Table 2: prefrontal clusters showing association of *GSK-3β* rs12630592 with cortical thickness outside the prefrontal cortex at the uncorrected threshold of p<0.001. These clusters did not survive statistical correction.

	Vertex number at the maximum	Talairach coordinates	Prefrontal region	Surface area (mm2)	Number of cluster vertices	Maximum p *value* (log_{10})
GG>TT	12961	x 27 y -86 z 9	Lateral occipital gyrus	77.8	107	3.8737
	33517	x 8 y -35 z 9	Posterior Cingulate	19.9	42	3.5025
	148117	x 37 y 2 z -4	Insula	24.3	50	3.4547
	156759	x 12 y -84 z 10	Pericalcarine	12.5	18	3.0820
	39418	x -6 y -35 z 34	Isthmus Cingulate	11.3	24	3.1639
GG>GT	143631	x 15 y -89 z 3	Pericalcarine	60.4	99	3.6834
	161534	x 9 y 38 z 38	Fusiform	47	75	3.7061
	122218	x -21 y 26 z -15	Precuneus	31.1	55	3.5526
GT>TT	21152	x 19 y -73 z 7	Pericalcarine	19.1	35	3.3179

Correlation analysis between prefrontal activity and prefrontal cortical thickness

There was no statistically significant correlation between BOLD responses in the clusters that differentiated the groups at the cognitive tasks with prefrontal grey matter thickness in subjects who underwent both volumetric and fMRIprocedures (N-back n=48; r=0.2; p=0.1; VAC n=74; r=0.09; p=0.4), suggesting that the two effects are not strongly related.

References

1. Piva F, Giulietti M, Nardi B, Bellantuono C, Principato G. An improved in silico selection of phenotype affecting polymorphisms in SLC6A4, HTR1A and HTR2A genes. Hum Psychopharmacol. 2010;25(2):153-61.

2. Piva F, Giulietti M, Burini AB, Principato G. SpliceAid 2: A database of human splicing factors expression data and RNA target motifs. Hum Mutat. 2012;33(1):81-5.

3. Lipska BK, Deep-Soboslay A, Weickert CS, Hyde TM, Martin CE, Herman MM, et al. Critical factors in gene expression in postmortem human brain: Focus on studies in schizophrenia. Biol Psychiatry. 2006;60(6):650-8.

4. 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):519-23.

5. Colantuoni C, Henry G, Zeger S, Pevsner J. SNOMAD (Standardization and NOrmalization of MicroArray Data): web-accessible gene expression data analysis. Bioinformatics. 2002;18(11):1540-1.

6. Leek JT, Storey JD. Capturing heterogeneity in gene expression studies by surrogate variable analysis. PLoS Genet. 2007;3(9):1724-35.

7. First MB, Gibbon M, Spitzer RL, Williams JBW. Guide for the structured clinical interview for DSM-IV axis I disorders-Research version. New York: Biometrics Research; 1996.

8. Nakayama H, Numakawa T, Ikeuchi T. Nicotine-induced phosphorylation of Akt through epidermal growth factor receptor and Src in PC12h cells. J Neurochem. 2002;83(6):1372-9.

9. Bertolino A, Taurisano P, Pisciotta NM, Blasi G, Fazio L, Romano R, et al. Genetically determined measures of striatal D2 signaling predict prefrontal activity during working memory performance. PLoS One. 2010;5(2):e9348.

10. Blasi G, Goldberg TE, Elvevag B, Rasetti R, Bertolino A, Cohen J, et al. Differentiating allocation of resources and conflict detection within attentional control processing. Eur J Neurosci. 2007;25(2):594-602.

11. Blasi G, Mattay VS, Bertolino A, Elvevag B, Callicott JH, Das S, et al. Effect of catechol-O-methyltransferase val158met genotype on attentional control. J Neurosci. 2005;25(20):5038-45.

12. Blasi G, Taurisano P, Papazacharias A, Caforio G, Romano R, Lobianco L, et al. Nonlinear response of the anterior cingulate and prefrontal cortex in schizophrenia as a function of variable attentional control. Cereb Cortex. 2010;20(4):837-45.

13. Blasi G, Napolitano F, Ursini G, Taurisano P, Romano R, Caforio G, et al. DRD2/AKT1 interaction on D2 c-AMP independent signaling, attentional processing, and response to olanzapine treatment in schizophrenia. Proc Natl Acad Sci U S A. 2011;108(3):1158-63.

14. Rosas HD, Liu AK, Hersch S, Glessner M, Ferrante RJ, Salat DH, et al. Regional and progressive thinning of the cortical ribbon in Huntington's disease. Neurology. 2002;58(5):695-701.

15. Kuperberg GR, Broome MR, McGuire PK, David AS, Eddy M, Ozawa F, et al. Regionally localized thinning of the cerebral cortex in schizophrenia. Arch Gen Psychiatry. 2003;60(9):878-88.

16. Salat DH, Buckner RL, Snyder AZ, Greve DN, Desikan RS, Busa E, et al. Thinning of the cerebral cortex in aging. Cereb Cortex. 2004;14(7):721-30.

17. Dale AM, Fischl B, Sereno MI. Cortical surface-based analysis. I. Segmentation and surface reconstruction. Neuroimage. 1999;9(2):179-94.

18. Fischl B, Sereno MI, Dale AM. Cortical surface-based analysis. II: Inflation, flattening, and a surface-based coordinate system. Neuroimage. 1999;9(2):195-207.

19. Fischl B, Salat DH, Busa E, Albert M, Dieterich M, Haselgrove C, et al. Whole brain segmentation: automated labeling of neuroanatomical structures in the human brain. Neuron. 2002;33(3):341-55.

20. Fischl B, van der Kouwe A, Destrieux C, Halgren E, Segonne F, Salat DH, et al. Automatically parcellating the human cerebral cortex. Cereb Cortex. 2004;14(1):11-22.

21. Fischl B, Dale AM. Measuring the thickness of the human cerebral cortex from magnetic resonance images. Proc Natl Acad Sci U S A. 2000;97(20):11050-5.

22. Segonne F, Dale AM, Busa E, Glessner M, Salat D, Hahn HK, et al. A hybrid approach to the skull stripping problem in MRI. Neuroimage. 2004;22(3):1060-75.

23. Jovicich J, Czanner S, Greve D, Haley E, van der Kouwe A, Gollub R, et al. Reliability in multi-site structural MRI studies: effects of gradient non-linearity correction on phantom and human data. Neuroimage. 2006;30(2):436-43.