Supplementary Information Guide (pdf)

Supplementary methods

This section describes in detail:

- criteria for selection of candidate genes (belonging to the serotonergic and dopaminergic domains) that were sequenced in this study

- deep sequencing strategy and methodology using an Illumina Genome Analyzer

- RNA and protein extraction protocols

- protocol for detection of RNA nonsense mediated decay

- protocol for western blotting and densitometry

- Finnish follow-up datasets

- a case report of homozygosity for the HTR2B stop codon

- protocol for detection of HTR2B expression in multiple regions of the adult human brain

- investigation of impulsive behavior in Htr2b-/- knockout mice using five measures: novelty-induced locomotion, locomotor reactivity in response to a dopamine D1 receptor agonist, exposure to a novel object, delay discounting and novelty suppressed feeding

Supplementary data

This section describes in detail:

- distribution of HTR2B *20 allele in Finland

- additional sequencing results

- sequencing accuracy

- Cerebrospinal fluid (CSF) monoamine metabolite measurements and results

- CSF testosterone measurement and results

- Average alcohol consumption in the Finnish case/control cohort

- Brown Goodwin Lifetime Aggression Scale score distributions in cases and control and in HTR2B *20 carriers and non-carriers

- Tridimensional Personality Questionnaire (TPQ) scores for violent offenders who were HTR2B *20 carriers and non-carriers

Supplementary figures and tables

Supplementary **Figure 1** is a flow chart illustrating the six components of the study.

Methods and results are supported by Supplementary figures 2-13 and tables 1-12.

Figure 2: Finnish ancestry of HTR2B *20 carriers and noncarriers.
Figure 3: Geographic distribution of HTR2B *20 in Finland
Figure 4: Comparison of ancestry of Finnish cases and controls
Figure 5: Reproducible detection of sequence variants using a test genome (ΦX174)
Figure 6: Error rate by sequencing cycle for a test genome (ΦX174)
Figure 7: Allele counts from pooled sequencing predict HapMap allele frequencies
Figure 8: Genotyping validates allele frequencies determined by pooled sequencing
Figure 9: HTR2B *20 occurs on a single haplotype background
Figure 10: Truncated 5-HT2B is undetectable in *20 heterozygotes (western blots)
Figure 11: High CSF testosterone in HTR2B *20 male carriers
Figure 12: Working memory impairment in HTR2B *20 male carriers
Figure13: HTR2B RNA expression in 13 human brain regions

Table 1: Finnish case-control resequencing and association datasets
Table 2: Region-specific primers for amplification of 108 target regions in fourteen genes
Table 3: Sequence variants detected by exon-centric sequencing of 14 genes
Table 4: Frequencies and secondary single marker associations for 21 additional nsSNPs
Table 5: Predicted functionality of missense/nonsense variants via PolyPhen and SIFT
Table 6: Global test of association for four putatively functional SNPs, and case/control association of HTR2B *20
Table 7: HTR2B *20 co-segregates with ASPD in eight informative families
Table 8: HTR2B *20 was undetectable in other populations worldwide
Table 9: Frequency (0.012) of HTR2B *20 in Finland including case/control, psychosis, and twin datasets (total N = 6554)
Table 10: Distribution of HTR2B *20 in a Finnish psychosis dataset (N = 1885)
Table 11: Distribution of HTR2B *20 in three Finnish twin cohorts
Table 12: Selective impairment of working memory in male HTR2B *20 carriers

Supplementary references are included to support the methodologies used in this study.

SUPPLEMENTARY METHODS

Gene Selection

Dopamine and serotonin releasing neurons are prominent in brain regions that regulate impulse control, and dysregulated activity of the monoamine neurotransmitters has been demonstrated to be involved in impulsivity in neuropharmacological, gene knock-out, and genetic association studies. We performed exon-centric sequencing for fourteen genes involved in dopamine or serotonin function.

Dopamine System. Dopamine regulates cognitive function, attention and responses to reward, all of which are factors in impulsivity. We sequenced genes involved in pre-synaptic, post-synaptic and down-stream (signaling) actions of dopamine (*DRD1, DRD2, SLC6A3* and *PPP1R1B*). There is empirical evidence that variation within these specific genes could influence brain function and moderate risk for impulsive behavior. Decreased levels of D2 receptor in the nucleus accumbens predicts spontaneous impulsivity in rats[13]. The D1 and D2 dopamine receptors modulate DARPP-32 (*PPP1R1B*), a phosphoprotein which plays a role in regulating responses to drugs of abuse and whose cortical and striatal activity are involved in motivation, attention and reward-related learning[40]. *SLC6A3* variation appears to moderate risk for ADHD, which is characterized by both hyperactivity and impulsivity[11].

Serotonin System. The serotonin system has long been implicated in impulsivity and, in particular, impulsive aggression and suicide. We sequenced genes involved in synthesis

and metabolism of serotonin (TPH2 and MAOA) and in pre- and post-synaptic signaling (5-HT receptors *HTR1A, HTR1B, HTR2A, HTR2B, HTR2C, HTR3A, HTR3B, SLC6A4*). The selection of candidates was based on empirical evidence of involvement of these genes (or gene products) in the expression of impulsivity and aggression. Tryptophan hydroxylase 2 (*TPH2*) encodes the enzyme which catalyzes the rate-limiting step for brain serotonin biosynthesis, and a *TPH2* haplotype may predict decreased cerebrospinal fluid levels of the serotonin metabolite 5-hydroxyindoleacetic acid and suicide attempt[41]. *MAOA* knockout mice have higher levels of monoamines and increased aggressive behavior[17], and a functional variable number tandem repeat in the *MAOA* regulatory region (*MAOA-LPR*) appears to moderate the effect of maltreatment on vulnerability to develop antisocial behavior[8,18]. Others have shown that a stop codon variant that produces complete deficiency of *MAOA* activity co-segregates with severe impulsivity[6].

Individual differences in serotonin 1A receptor binding have been associated with life-time aggression[42], and genetic and gene-knockout studies performed in mice implicated the *HTR1B* gene in both alcohol preference and aggression[43]. In one association study in humans, a non-synonymous mutation (Gly861Cys) in the *HTR1B* gene was linked to antisocial alcoholism in two independent populations, including the Finnish violent offenders cohort[44]. Many pharmacological studies with serotonin 2A, 2B and 2C antagonists and agonists have implicated these receptors in impulsive behavior[45]. Two subunits of serotonin type 3 receptors, 3A and 3B, have so far been identified; in one study variation within the *HTR3B* gene was associated with AUD comorbid with ASPD[46]. Stress-modified associations have been reported also for a polymorphism in the serotonin transporter (*HTTLPR*) to suicidality[19,20]. In terms of the relationship of these

serotonin and dopamine related genes to function, it is important to recognize that most of the genetic association results are preliminary, have in several cases been contradicted, and are in need of replication and further exploration at the functional level. The primary impetus to resequence genes involved in dopamine and serotonin neurotransmission in impulsive individuals was evidence from functional neurobiological investigations, including behavioral pharmacology, neurochemistry, and gene knockouts.

Resequencing

Exon-centric targeting of 14 candidate genes: We custom-designed or used Applied Biosystem oligonucleotide primers to amplify 108 target regions that covered exons, flanking regions and ~ 800-1000 bp of the upstream regions of 14 genes, for a total of 82 kb (Supplementary Table 2).

DNA pools: DNA samples were individually quantified in three replicates by RT-PCR, using TaqMan RNase P Detection Reagent kits (FAM) and Roche human DNA standards, and were normalized to 10 ng/µL. Eight pools (12 subjects per pool) were made with equal amounts of DNA from 96 Finnish cases and in parallel fashion eight pools were made from 96 Finnish controls.

DNA Amplification: DNA pools were amplified in 108 separate PCR reactions. Five micrograms of pooled DNA was amplified in a 50 uL reaction volume containing 5 ul PCR buffer, 1 ul of a 10mM dNTP mixture, 2 ul of 50 mM $MgSO_4$, 6 uL primer mix

(containing 5 mM forward primer and 5 mM reverse primer), 0.2 ul Platinum Taq High Fidelity polymerase (Invitrogen), and 30.8 ul distilled water. Polymerase chain reactions were performed at 94°C for 1 min, followed by 35 cycles each at 94°C for 20 sec. Annealing temperatures ranged from 54° to 68° for 20 sec, 68°C for 90 sec, and a final extension step of 72°C for 10 min. Amplifications were verified by presence of a DNA fragment band of the expected size on 1% agarose gels with 5 μl of each amplification product.

DNA Sequencing: Prior to sequencing, amplicon concentrations were normalized using SequalPrep Normalization Plate kits (Invitrogen). All amplicons from the same DNA pool were combined. The DNA was sheared by sonication and purified with QIAquick PCR purification kits (QIAGEN). Genomic DNA Preparation kits and protocol (Illumina) were used to prepare sequencing libraries. Adapters were added to the ends of DNA fragments and fragments ranging in size from 200 to 300 bp were isolated by gel purification, followed by solid phase amplification and analysis on the Illumina Cluster Station and Genome Analyzer. As determined by sequencing of a 5.8 kb ΦX174 genome 24 at 20,000-fold coverage, sequencing detection of sequence variants was reproducible and the error rate was < 1% (Supplementary Fig. 5 and 6). Each subject was therefore re-sequenced at 80-fold coverage (1000-fold for each pool of 12 individuals), a level sufficient to overcome unevenness of target capture and to reliably detect a single heterozygote within the pools of 12. Exon nucleotide coverage was 95%. Exon nucleotide coverage was 95%.

Analysis of sequence data: Sequences were called from image files with the Illumina Genome Analyzer Pipeline and aligned to human reference sequence from NCBI build 36.3 using the Illumina Eland software. The output from the Illumina Genome Analyzer was 36 base-length sequence reads. Each read was uniquely mapped to the human reference genome. Sequence reads with more than 2 mismatches were excluded. Sequence reads with alternative alleles that did not exactly match the reference genome did uniquely map to the corresponding location in the reference sequence. Additional results are described in supplementary data.

HTR2B cDNA Sequencing for nonsense-mediated decay

To detect nonsense-mediated RNA decay we sequenced cDNA from *HTR2B* Q20/Q20 homozygotes and Q20/*20 heterozygotes, measuring the ratio of the Q20 and *20 alleles within heterozygous cDNAs (Fig. 2).

Total RNA was extracted from lymphoblastoid cell lines derived from blood samples of 12 Finnish non stop codon carriers and 14 Finnish stop codon carriers using the TRIzol LS reagent protocol (Invitrogen). cDNA was synthesized by reverse transcription of 2 ug of total RNA with Multiscribe reverse transcriptase (High-Capacity cDNA Reverse Transcription kits protocol, Invitrogen) according to standard protocols.

The sequences of the upstream and downstream oligonucleotides were as follows: 5'-gagtgtttggcatggttaca-3' and 3'-accaggcaggacatagaaca-5'. PCR amplification were performed with 0.2 ul Platinum Taq DNA polymerase high fidelity (Invitrogen), 5ul of

10X high fidelity PCR buffer, 1 ul of a 10mM dNTP mixture, 2 ul of 50 mM $MgSO_4$, 2 ul primer mix (10 uM each), 34.8 ul sterile water, 5 ul cDNA. Polymerase chain reactions were performed at 94°C for 1 min, followed by 40 cycles each at 94°C for 20 sec, 60°C for 30 sec, 68°C for 30 sec, and a final extension step of 72°C for 10 min. Sanger sequencing was performed using the BigDye Terminator Sequencing Mix (Applied Biosystems) and analyzed on the Applied Biosystems 3730 DNA Analyzer.

Quantification of the *HTR2B* transcript was carried out by comparing the intensity peaks of the Q20 and *20 alleles within cDNA derived from each heterozygous individual, with homozygous individuals serving as control to verify the low *20 background signal. The cDNA sequencing revealed different levels of nonsense mediated decay (NMD) ranging from none to nearly complete (Fig. 2). NMD is a complex mechanism that ensures the accuracy of gene expression by rapid degradation of mRNAs containing premature translation termination codons, and regulates the expression of alternative splice products. However, it has become clear during recent years that many physiological mRNAs are also NMD substrates, indicating a role for NMD beyond mRNA quality control. Also, several features of physiological mRNAs can render them NMD sensitive, whereas some mRNAs have evolved stabilizing elements that protect them from NMD[47].

The detection of various levels of un-degraded RNA in the lymphoblastoid cells of heterozygote carriers may indicate different degrees of efficiency of NMD, as has recently emerged[48,49].

Western Blots

To measure the effect of the *HTR2B* stop codon on expression of the receptor, we used western blotting with three antibodies directed against different regions of the receptor. HT2B protein was measured in 12 Finnish Q20/Q20 homozygotes and 14 Finnish Q20/*20 heterozygotes. Total protein was extracted from lymphoblastoid cell lines using the TRIzol LS reagent protocol (Invitrogen). Western blots were prepared using 50 ug of protein per lane on a 10% Bis-Tris gel (Invitrogen). Separated proteins were transferred to nitrocellulose using the iBlot transfer system (NuPage - Invitrogen). Blots were stained with Amido Black to check protein loading and transfer prior to blocking in Tris buffered saline containing 0.1% Tween 20 (TBST) and 5% milk. Blots were probed with antisera raised against either the N-terminal (mouse monoclonal Ab, Novus Biologicals), internal (goat polyclonal Ab, Santa Cruz Biotechnology) or C-terminal (rabbit polyclonal Ab, Santa Cruz Bioechnology) regions of the 5-HT2B protein, and GAPDH Ab (Millipore). Antibody binding was visualized on X-ray film (Kodak XAR) using chemiluminescence (ECL Plus, GE Healthcare). Western blot X-ray films were scanned and densitometry performed using NIH ImageJ. Ratios between the 5-HT2B receptor and the housekeeping protein GAPDH were calculated to normalize 5-HT2B protein quantity. Comparing 26 cell lines of the two genotypes, the mean Q20/Q20 to Q20/*20 5-HT2B protein ratio was 1.93, pval = 0.03, demonstrating that the Q20* stop codon blocks the expression of the 5-HT2B receptor (Fig. 2). No truncated 5-HT2B protein was detected that was capable of cross-reacting with antibodies raised against either the 5-HT2B C-terminus or the N-terminus (Supplementary Fig. 10).

HTR2B Q20* FOLLOW-UP IN FINNISH DATASETS

To confirm the allele frequency of the *HTR2B* *20 variant in the Finnish population (Supplementary Table 9) and test for phenotype/genotype associations we genotyped two family-based Finnish cohorts and three Finnish twin cohorts collected in Finland as a national resource for genetic epidemiological studies. We genotyped the samples for which DNA was available with no genotype or phenotype-based ascertainment bias. The genotyping for *HTR2B* Q20* was performed as previously described.

Bipolar disorder and schizophrenia nuclear families

Nuclear families with at least one member affected with schizophrenia or bipolar disorder were collected in Finland through the National Hospital Discharge Register. All individuals born between 1940 and 1976 who were hospitalized due to an ICD-8 or DSM-III-R diagnosis of bipolar disorder or schizophrenia, schizophreniform disorder, or schizoaffective disorder between 1969 and 1991 [50,51] were identified. Available medical records of the family (siblings and parents) were collected. Probands were contacted through their treating psychiatrist and, if the proband gave permission, other family members were contacted for detailed investigation. The Ministry of Social Affairs and Health and the Ethics Committee of the National Public Health Institute (since January 1, 2009 the National Institute for Health and Welfare) approved the study. All participants to the study signed a written informed consent.

Altogether 32 of these individuals were heterozygous *HTR2B* *20 carriers. Of these, 7 nested within 5 bipolar families and 25 within 13 schizophrenia families, with a frequency of 0.0087 in unrelated individuals. No *20/*20 homozygotes were detected.

1885 individuals were genotyped. Clinical information on the presence of psychosis was available. Diagnosed psychosis was numerically, but not statistically, more prevalent in the *20 carriers (Supplementary Table 10). Early-onset schizophrenia was observed among the *HTR2B* *20 carriers (mean 20.9; median 19 years) compared to Q20/Q20 individuals (mean 24.5; median 22 years), p = 0.035 (Student's two-tailed unequal variance T-test).

The Older Finnish Twin Cohort

This dataset consists of Finnish monozygotic and dizygotic twin pairs born before 1958[52]. The collection of this dataset started in 1975 with follow-up questionnaires in 1981, 1990 and 2000. The subjects of this study were recruited through the Finnish Twin Cohort, compiled from the Central Population Registry of Finland which is a databank of personal information on all Finnish citizens from 1967 onward. In 1975 and 1981 the twins received structured questionnaires assessing their demographic, social, environmental, medical and lifestyle characteristics. Drinking and smoking habits were assessed. In 1999 the twin pairs both alive and age 65 years and over were interviewed with the TELE instrument to assess their cognitive status[53].

In this present study a total of 2388 twins of known zygosity were successfully genotyped for *HTR2B* Q20* (Supplementary Table 11). We analyzed the outcome of items describing externalizing behavior belonging to the Bortner scale[54] and from the 1981 Questionnaire[55]. No difference was detected between 26 *HTR2B* *20 carriers and 1867 non carriers (data not shown) for whom the questionnaire was available.

FinnTwin 16 cohort (FT16)

The FinnTwin 16 study (FT16) is a longitudinal study of twins born between 1975 and 1979. The dataset comprises twin pairs chosen for their extreme discordance or concordance in self-reports on the Rutgers Alcohol Problem Index (RAPI) at age 18.5. In 2001-2003 the twins received an extensive clinical assessment through blood testing, neuropsychological testing, and structured psychiatric interviews (DSM-IIIR)[56,57]. Zygosity was determined for all same-sex twin pairs using multiple genetic markers assayed at the laboratories of the National Public Health Institute in Helsinki.

FinnTwin 12 cohort (FT12)

This study was started in 1994 in 11-12 year old twins to examine genetic and environmental precursors of health related behaviors, with a particular focus on use and abuse of alcohol. Twins were assessed at age 14 and 17.5 years. In 2006 a fourth wave of data collection was initiated on young adults. Information on lifestyle, substance use and

abuse, psychological and physical health, psychiatric conditions, personality, cognitive functions, and work life were collected[56,57].

A total of 701 FT16 individuals and 710 FT12 individuals were successfully genotyped for *HTR2B* Q20*, for a total of 1110.5 (FT12+FT16) genetically independent twins (Supplementary Table 11).

In the FT16 sample 511 Finnish genotyped individuals had neuropsychological test results (252 males and 259 females), while neuropsychological tests results were available for 172 male and 250 female genotyped individuals in the FT12 study.

Neuropsychological assessment

Neuropsychological assessment was conducted on both the combined FT16 and FT12 cohorts for measures of verbal intellectual ability, working memory and executive function:

- Verbal intellectual ability was assessed with the Vocabulary subtest (WAIS-R).

- Working memory was assessed with the Digit Span Forward and Backward subtests of the Wechsler Memory Scale-Revised (WMS-R) and the Letter-Number Sequencing subtest of WAIS-III.

- Executive function was studied with the Trail Making Test (TMT) and the California Stroop Test.

Results

Working memory was assessed with the Digit Span Forward and Backward subtests of the Wechsler Memory Scale-Revised (WMS-R). We analyzed the combined FT16 and FT12 datasets (Supplementary Table 12) for the WMS-R Digit Span Forward, Backward and Total tests, verbal intellectual ability and executive function tests available for both FT16 and FT12 datasets.

A linear regression model was constructed using performance on the working memory test as the dependent variable and sex and genotype as independent variables. Sex was a significant predictor, so the sample was stratified into male and female. Male heterozygotes performed significantly worse on the Digit Span backward and forward tests, and combined score (Supplementary Table 12 and Supplementary Fig. 12). Cognition in heterozygous males is overall normal, however these results may indicate a specific frontal lobe dysfunction modulated by sex.

No associations with performance on the other neuropsychological tests listed above were identified. All statistical analyses were conducted using Stata (version 11, Stata Corp, College Station, Texas, USA). Criterion for statistical significance was set at 0.05.

Bonferroni correction for multiple testing was applied, as presented in Supplementary Table 12.

A case report of homozygosity for the *HTR2B* stop codon

One homozygous *HTR2B* Q20* male belonging to the FT16 cohort was identified in a total of 6554 Finns that have been genotyped. He completed questionnaires at ages 16, 17, 18 and was interviewed in person at age 26. This young adult fulfilled criteria for DSM-III alcohol dependence and had two positive criteria for Antisocial Personality Disorder (Semi-Structured Assessment for the Genetics of Alcoholism - SSAGA interview). He was born at 38 wks gestational age, with a birth weight of 1960 g. Development was reported as normal by the mother. He was cognitively normal (with normal performance on the digit span test), successfully completed high school and military service (which is compulsory for male Finnish citizens), and was gainfully employed. At age 25, he was unmarried but in a stable relationship with no children. He had no major congenital defects or history of major medical illnesses. However, as an adolescent he had a history of weekly drinking with alcohol-related problems which brought him to the attention of the police and social services. At age 16 he endorsed five of nine items related to problems related to alcohol including a tendency to get into fights while under the influence of alcohol (the questionnaire was developed at the University of Helsinki, www.twinstudy.helsinki.fi). His biological mother and father and his fraternal twin brother were both heterozygous for the stop codon. On the basis of questionnaires returned by each parent, neither had pathological behavior, alcohol-related

or otherwise. However, only the brother was directly interviewed and he fulfilled criteria for DSM-III alcohol dependence. Thus, clinical information on this kindred is limited.

QUANTITATIVE PCR FOR *HTR2B* IN HUMAN BRAIN

We evaluated *HTR2B* expression in thirteen human brain regions. Total RNA was available for frontal lobe, temporal lobe, parietal lobe, occipital lobe, olfactory region, cerebellum, diencephalon, hippocampus, thalamus, pituitary gland, pons, medulla oblongata (Biochain) and nucleus accumbens (Clontech). First-strand cDNA was synthesized by reverse transcription of 2 ug of total RNA with Multiscribe reverse transcriptase (High-Capacity cDNA Reverse Transcription protocol, Invitrogen). ABI Taqman gene expression assays directed against the 2-3 and 3-4 exon boundaries of the *HTR2B* gene were used according to protocol (ABI Taqman gene expression assays Hs01118766 and Hs00168362). B-actin expression was used as an internal control (B-Actin Taqman gene expression assay, ABI).

HTR2B expression was detected in all analyzed human brain regions via both qPCR assays. Expression was highest in cerebellum, occipital lobe and frontal lobe and overall the levels of expression determined by the two assays correlated with an r = 0.98 (Supplementary Fig. 13).

$HTR2B$ KNOCKOUT MICE

Htr2b[-/-] knockout mice (50% males and 50% females) were made in a pure 129Sv/PAS background. Wild type (WT) 129/SvPAS mice (8-10 week old), bred in-house, were used as a control group.

Impulsive behavior in *Htr2b*[-/-] knockout mice was investigated using five experimental measures: novelty-induced locomotion, locomotor reactivity in response to a dopamine D1 receptor agonist, exposure to a novel object, delay discounting and novelty-suppressed feeding.

a. Locomotor activity in a novel environment. In order to study their reaction to a novel environment, we compared spontaneous locomotor activity of *Htr2b*[-/-] and WT mice.

Locomotor activity was measured in a circular corridor with four infrared beams placed at every 90° corner (Imetronic, France). Counts were incremented as consecutive interruptions of two adjacent beams (i.e., mice moving through one-quarter of the corridor).

Analysis of locomotor pattern revealed initially elevated activity levels in *Htr2b*[-/-] mice (n = 20) compared to controls (n = 20). However, when locomotion was followed over 24 hrs it was observed that *Htr2b*[-/-] exhibited hyperlocomotion only at the beginning of the day session (0-2h) and the beginning of the night session (4-6h) (Fig. 3, a). After habituation, mice of both genotypes had similar locomotor activity. These data suggest

that *Htr2b*[-/-] mice display an increased locomotor response to environmental novelty, a finding classically associated with impulsivity.

b. Locomotor activity following a D1 agonist. Genetic association studies have implicated several genes involved in dopaminergic neurotransmission in disorders marked by impulsivity, such as ADHD, including the dopamine D1 receptor[58]. Local infusions of D1 receptor agonists results in increased impulsive choice[59,60]. Mice were placed individually in activity boxes for 30 min for 3 days consecutively to habituate them after which D1 agonist-stimulated locomotion was measured.

Compared to wild type mice (n = 8), a single injection of dopamine D1 receptor agonist SKF 81297 induced greater locomotor activation in *Htr2b*[-/-] mice (n = 8) (Fig. 3, b), indicating either more reactive dopaminergic neurotransmission or reduced suppression of locomotion.

c. Exposure to novel object. A novel object (a plastic blue cap from a 15 ml Falcon tube) was introduced in the same corner of the home cage of each mouse. The object was placed in the cage after the animal was momentarily removed. The subject was then returned to the cage by gently setting it down on the opposite corner with respect to the location of the novel object. Data recording started immediately and measured physical and non-physical (sniffing) interactions of the animal with the object for 10 minutes. The latency to approach (measured as first sniffing) and number of contacts, were analyzed by a researcher blind to the genotype. The latency in approaching the novel object was

unmodified but the number of contacts was significantly increased for $Htr2b^{-/-}$ (n = 8) mice compared to wild type (n = 8) (Fig. 3, c).

d. Delay-discounting task. The task measures the reduction of a reward's subjective value as a function of delay. Computer-controlled operant chambers (Imetronic, Pessac, France) were used as experimental apparatus. The chambers, constructed in plexiglas and with a grid floor, were provided with a chamber light, two nose-poking holes, two hole lights, one feeder device and one food magazine. Nose poking into the hole was detected by infra-red photocell and was recorded by a computer which also controlled the food delivery. Each mouse was tested in the same chamber. The grids of the operant chamber were cleaned after each animal was tested.

All animals were habituated to the reinforcer (pellet of sucrose 20 mg, Bioserv, Thiais, France) outside the chamber for three days. The following three days, animals were familiarized with the operant chamber for 25 min, and habituated to the food magazine with the pellet freely dropping every 60 sec. For the following 10 days mice were daily trained for 25 min to respond with a nose poke to a visual stimulus. Before each training session, mice were food-deprived for 8 - 10 hours. At the beginning of each session, the chamber light was switched-on for 3 sec before the hole lights were switched-on. Nose-poking in one of the two holes (termed small and soon – SS - hole) resulted in the delivery of 1 pellet of reinforcer, whereas nose-poking in the other hole (termed large and late – LL - hole) resulted in the delivery of 5 pellets of sucrose. When a nose-poke occurred the corresponding light remained switched on for 1 sec. At the end of the 10

days of training, mice (both wild type and *Htr2b*[-/-]) were tested to ensure, in the absence of delay, their preference for the LL reinforcer (see delay 0 sec, Fig. 3, d).

During the testing phase (7 days), a delay was inserted between nose-poking in the LL hole and the delivery of the five-pellet reward. The chamber light was switched-on during the full length of this delay. Any additional nose poking during this interval was not reinforced. The delay was kept fixed for each of the two daily sessions and progressively increased across subsequent days (0, 2, 4, 6, 8, 10, 12 and 14 sec).

As delays became longer the preference for the large reward appeared to decrease more rapidly for *Htr2b*[-/-] mice (n = 6) (4 sec delay) than for wild type mice (n = 6) (10 sec delay) (Fig. 3, d).

e. Novelty suppressed feeding. This paradigm is a conflict test that measures a consequence of competing motivations: the drive to eat versus fear of venturing into the center of brightly lit arena where the food is located. The testing apparatus consisted of a plastic box, 37 cm x 57 cm x 10 cm, directly illuminated by white light. The floor was covered with 2 cm sawdust. Eighteen hours before the test, food was removed from the cages. At the time of testing, a single pellet of regular food was placed at the center of the box. The animal was placed in a corner of the box, and a stopwatch was immediately started. The latency to start eating (defined as the mouse sitting on its haunches and biting the pellet with use of forepaws) was recorded within a 5 min period[61]. Compared to wild type mice (n = 4), the time spent before eating food was significantly reduced in *Htr2b*[-/-] mice (n = 4), also indicating that they were more impulsive (Fig 3, e).

f. Testosterone plasma levels. Male *Htr2b*[-/-] mice (n = 12) had three-fold higher basal

levels of testosterone compared to control mice (n = 12) (Fig 3, f). These data are echoed

by higher testosterone CSF levels observed in human male *Htr2b* *20 carriers. These may

indicate an interaction between the 5-HT2B receptor and regulation of testosterone. The

mechanism is unknown, but the consequences are potentially important in impulsivity.

Behavioral assays were analyzed by unpaired t test and Fisher test (b, c, e, f), by one-way

analysis of variance (ANOVA) (a) and by repeated measures two-way ANOVAs

(genotype × delay, $F_{(5,50)} = 2.14$, $p = 0.076$; delay, $F_{(5,50)} = 2.64$, $p < 0.05$; genotype,

$F_{(1,10)} = 10.07$, $p < 0.01$) (d). Bonferroni correction was used for post hoc comparisons.

P<0.05 was predetermined as the threshold for statistical significance.

SUPPLEMENTARY DATA

Distribution of *HTR2B* *20 in Finland

The Finnish population was founded by two waves of migration: Eastern Uralic founders arrived 4000 years ago, followed by Indo-European speakers 2000 years later[23]. Population losses and bottlenecks left an imprint of restricted Y-chromosome diversity, providing evidence of the dual origins of Finns[62]. The Finnish population is equally diverse at autosomal loci as other European populations but isolation and genetic drift have molded the gene pool of Finns, which have been extensively analyzed as a population isolate[63]. Within Finland are important sub-isolates. Information on 78 twins heterozygous for *20, and belonging to twin cohorts described in supplementary information, revealed that *20 is distributed throughout Finland. *20 carriers originated from all three regions of Finland[64]: Eastern (2.7%), Western (3.1%) and Middle (5.7%) with an excess in Middle Finland (Pearson chisquare = 8.89, p = 0.012) (Supplementary Fig. 3).

Resequencing

SNP calling: 360 variants passed Genome Analyzer pipeline quality control and the criteria established to identify SNPs (Supplementary Table 3). No base pair in a read had a quality score < 10; a locus, to be identified as a SNP, had to have a quality score >=20. Quality scores were estimated by ELAND using a modified Phred formula, which

calculates four quality scores and selects the most probable base. At least 20 read counts for the alternative allele at a specific locus with minimum alternative allelic frequency of 0.5% were necessary to call a SNP. In the pooled samples it was impossible to distinguish rare insertion/deletion polymorphisms from sequencing errors and here we have analyzed nucleotide substitutions only.

Transitions constituted 60% of the 360 detected SNPs, in accord with previous data[65]. In contrast, for sequence errors, the transition/transversion ratio would be expected to be 1:2. Non-coding SNPs were found every 144 bases, and coding SNPs every 783 bases. Within coding sequence, nonsynonymous and synonymous SNPs were found in a ratio of approximately 1:1, as previously reported[66]. A non-synonymous SNP (nsSNP) was found every 1673 bases and a synonymous SNP every 1472 bases (Supplementary Table 3).

In this study we identified 96% of the abundant (0.05 allele frequency or higher) SNPs recorded in HapMap as residing in the target regions of the 14 genes. Within the 82 kb target region, 172 SNPs are recorded in dbSNP (HapMap) for the CEU population (about 60 unrelated individuals of European origin). 95/172 (56%) of these SNPs were identified. However, the HapMap SNPs that we did not observe were all either monomorphic or low in frequency (<0.04). 73 of 77 HapMap SNPs that we did not identify are recorded in that database as either monomorphic or having an allele frequency <0.05. The pooling design allowed the identification of alleles present in only one copy in the pool, and reasonably accurate estimation of allele frequency by sequencing representation. Major allele frequencies from the HapMap CEU genotyping

and re-sequencing correlated with an $r^2 = 0.88$ (Supplementary Fig. 7). We detected 188 SNPs that were not present in dbSNP and as shown by the individual genotyping of 34 loci (22/26 nsSNPs, 1/1 sSNP and 7/7 intronic SNPs validated), the majority are valid.

Genotyping

The 34 individually genotyped SNPs, including 22 validated nsSNPs (Supplementary Table 4), were genotyped by TaqMan (5' exonuclease) assays predesigned and custom-designed (Applied Biosystems). Primer sequences are available on request. We genotyped 228 Finnish cases and 295 Finnish controls described elsewhere (Finnish sample). Minor allele frequencies estimated by Illumina Genome Analyzer and allele frequencies from genotyping correlated with an $r = 0.94$, demonstrating the accuracy of allele frequency estimation from the sequencing of pooled samples (Supplementary Fig. 8). Genotypes of controls and parents did not deviate from Hardy-Weinberg equilibrium ($P > 0.05$).

The genotyping was performed utilizing ABI 7900 Taqman platform (Applied Biosystems).

Cerebrospinal fluid (CSF) monoamine metabolite measurements

CSF monoamine metabolites were measured in 225 Finnish individuals (112 controls and 113 cases). Lumbar CSF was obtained while they were on the research ward and after

they had been there a minimum of three weeks. They were free of medications, including selective serotonin reuptake inhibitors, alcohol-free and had been maintained on a low monoamine diet. CSF was obtained by lumbar puncture between 8 and 9 a.m. after overnight rest with only water permitted after 8 p.m. Concentrations of 5-HIAA, HVA and MHPG were measured in the first 12 ml of CSF collected, by liquid chromatography and electrochemical detection[67].

CSF monoamine metabolite levels (5-HIAA, MHPG and HVA) were determined for 9 of the *20 cases and 112 controls. *20 cases did not differ significantly from controls. For 5-HIAA, cases had mean = 55.4 pg/ml, SD ± 24.4 and controls 66.9 pg/ml, SD ± 24.5. For, HVA, cases had mean = 139.5 pg/ml, SD ± 46.6 and controls 176.4 pg/ml, SD ± 70.2). For MHPG, cases had mean = 35.9 pg/ml, SD ± 6.7 and controls 41.1 pg/ml, SD ± 10.3.

CSF testosterone measurement

CSF testosterone was measured in 41 controls and 91 cases. Due to the skewness of these data (<0.96) a log transformation was applied. CSF testosterone levels were significantly higher in the nine *20 cases as compared to controls (t-ratio = 2.32, p-val = 0.024), indicating the potential for a gene x endocrine interaction. CSF testosterone measurements were not available for *20 controls. This raises the possibility that an interaction between *20 and testosterone contributes to the severe impulsive behaviors seen in male criminal offenders (Supplementary Fig. 11).

Average alcohol consumption

Average annual alcohol consumption ranged from 0-163 kg/yr (median = 4.6; mean = 15.2; s.e. = 1.07) in subjects without *20 and from 2.8-57.1 kg/yr in subjects with *20 (median = 16.7; mean = 21.2; s.e. = 3.4). Among cases without *20, alcohol consumption ranged from 0.3-163 kg/yr (median = 28.9; mean = 34.8; s.e. = 2.1), and between 4 and 57 kg/yr in cases with *20 cases (median = 29.7; mean = 26.9; s.e. = 3.7). None of these differences were significant.

Brown Goodwin Lifetime Aggression Scale (BGLAS)

In the case/control dataset the subjects carrying *20 had higher mean BGLAS scores ($\chi2$=5.66, pval = 0.017 nonparametric Wilcoxon test). We analyzed the BGLAS scores for cases without the *20 stop codon are compared to BGLAS scores for 13 cases who were heterozygous for *20 and for whom these data were available. No statistical significance was detected but the aggression scores of *20 carriers are located at the high end, even among cases.

Tridimensional Personality Questionnaire (TPQ)

The Tridimensional Personality Questionnaire was available for 424 subjects, including 11 *20 carriers. Comparing cases to controls, differences were observed in three dimensions of personality. Cases had a combination of high novelty seeking (NS) (t ratio

24

= 6.8, p = 0.0001), high harm avoidance (HA) (t-ratio = 8.8, p =0.0001) and low reward dependence (RD) (t ratio=-2.6, p=0.009), consistent with previous analyses in this sample[68]. As compared to controls, *20 cases had high NS (t-ratio = 2.8, p = 0.005), high HA (t-ratio = 1.8, p = 0.06) but were not different in RD (t-ratio = -0.18, p = 0.85).

All analyses were conducted using JMP software v7.0 (SAS Institute, Cary, NC). Criterion for statistical significance was set at 0.05.

SUPPLEMENTARY FIGURES
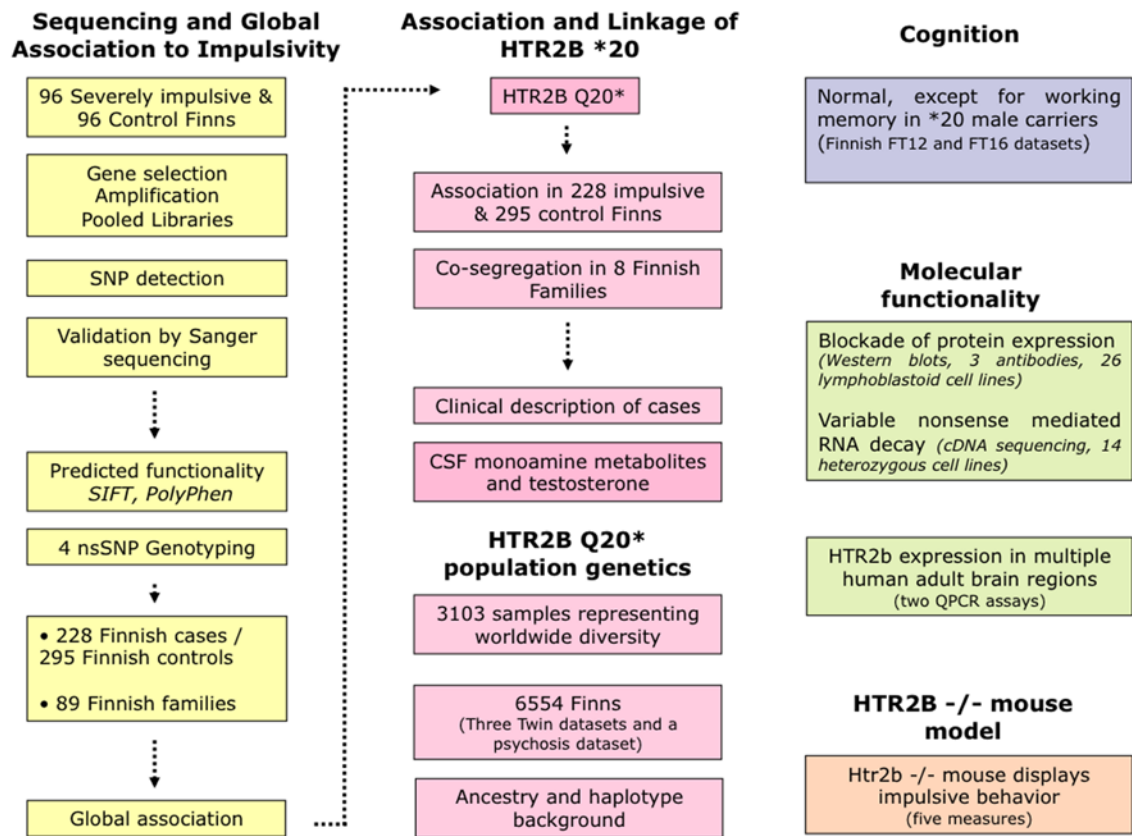
Supplementary Figure 1. Flow chart illustrating the six components of this study:

resequencing and a global test of association of putatively functional variants in severe

probands from a founder population, association and linkage of an *HTR2B* stop codon,

population genetics in Finnish datasets and the Human Genome Diversity Panel (HGDP),

cognitive effects, gene expression and functionality, and studies of the *HTR2B*^{-/-} mouse.
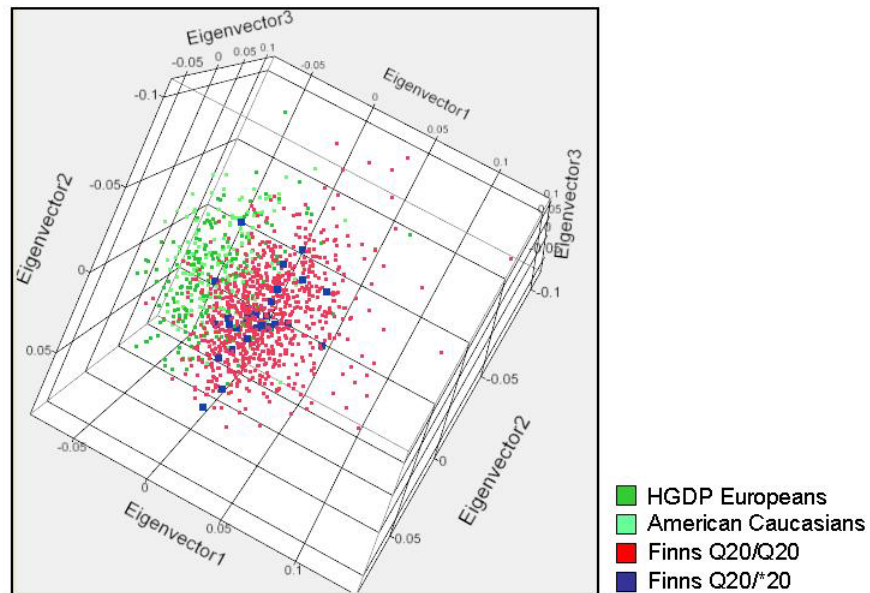
Supplementary Figure 2. Finnish ancestry of *20 carriers and noncarriers.

In ancestry, Q20* carriers (blue dots) more closely resemble Finns than individuals from other European populations. Ancestry was measured using 177 ancestry informative markers in 29 Q20* carriers, 580 other Finns, and 200 individuals representing 10 European populations in HGDP. Principal component analysis was performed with EIGENSTRAT.
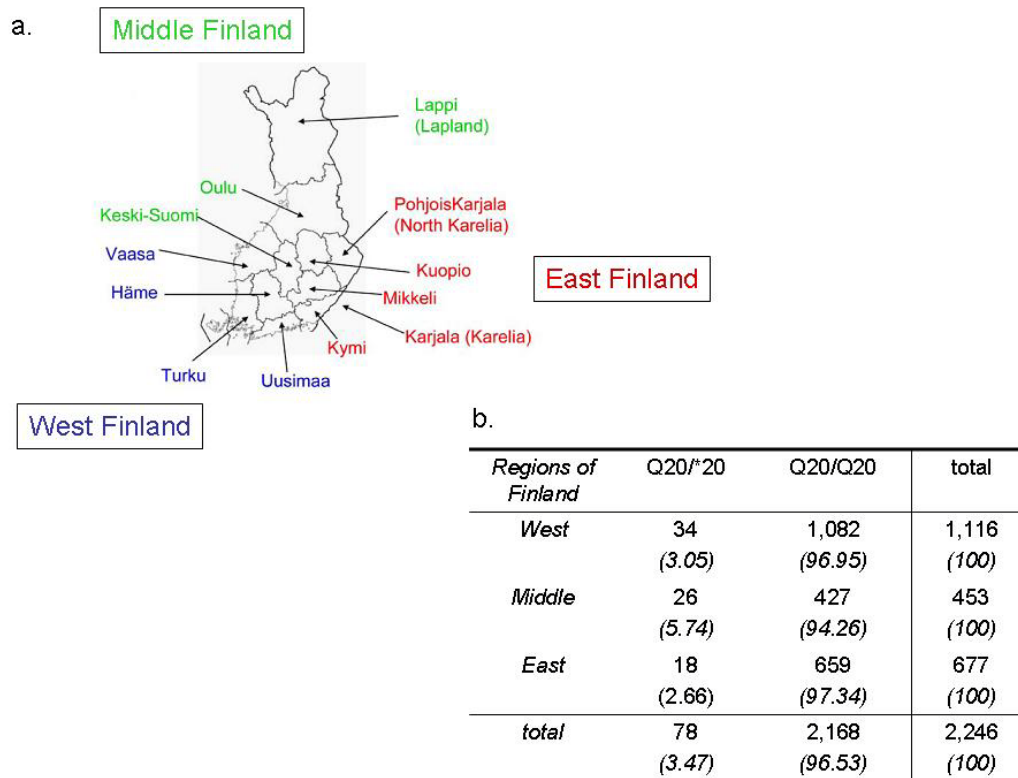
Supplementary Figure 3. Geographic distribution of *HTR2B* *20 in Finland.

The origins of 54 *HTR2B* *20 twins belonging to the Old Finnish Cohort and 24 *20 carriers belonging to the FT12 and FT16 Twin cohorts are depicted.

a) Finland is divided into three historical provinces: East, West and Middle (adapted from [64]). For individuals born 1926-1936, province of origin was defined by birthplace, and for individuals born 1975-1979 parental province of origin was used, including only individuals both of whose parents came from the same region. The latter definition was chosen for younger individuals since it filters out the effects of internal migrations from the late 1950s to 1970s. b) Overall, an excess of *20 carriers appeared to come from middle Finland (Pearson $\chi2 = 8.89$, p = 0.012), but *20 was widely distributed in Finland.
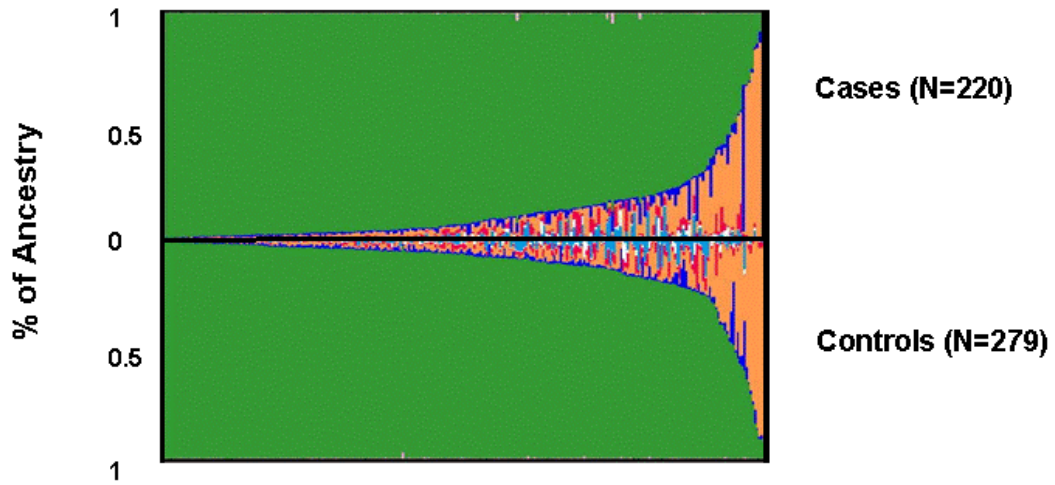
		Case (s.e.)	Control (s.e.)	t p-value
	Average score			
🟪	Africa	0.002 (0.0003)	0.004 (0.003)	0.39
🟩	Europe	0.860 (0.01)	0.868 (0.01)	0.58
🟦	Middle East	0.021 (0.003)	0.017 (0.001)	0.32
🟧	Central Asia	0.077 (0.01)	0.071 (0.008)	0.66
🟥	Far East Asia	0.016 (0.001)	0.015 (0.001)	0.72
⬜	Oceania	0.008 (0.0007)	0.007 (0.0006)	0.33
🟦	North America	0.018 (0.001)	0.018 (0.001)	0.92

Supplementary Figure 4. Comparison of ancestry of Finnish cases and controls.

No difference was detected between cases (ASPD, BPD and IED) and controls in proportions of ancestries. The figure compares the pattern of measured ancestry for seven ancestry factors derived separately for each control (N = 279) and case (N = 220) with reference to the Human Genome Diversity Panel (HGDP) (1051 DNAs representing 51 populations worldwide).
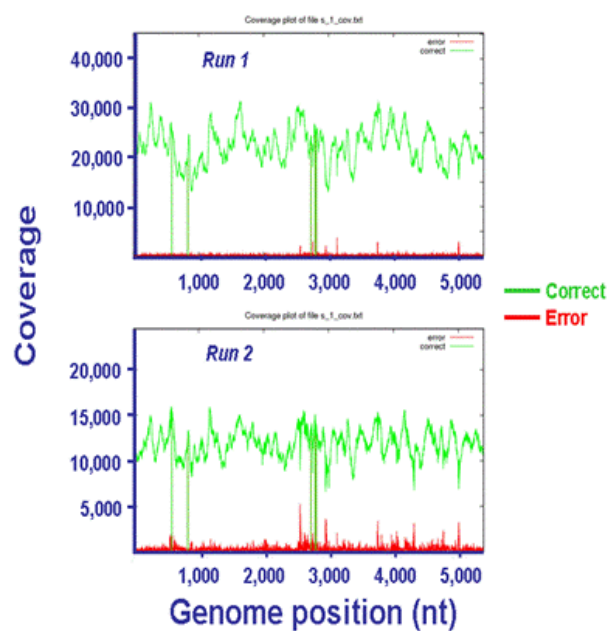
Supplementary Figure 5. **Reproducible detection of sequence variants using a test genome (ΦX174).**

φX174 was resequenced using an Illumina Genome Analyzer at 10,000 X (bottom) or 20,000 X coverage. The 25 bp reads were mapped back to the canonical φX174 sequence (5.8 Kb) using Eland. Error rate was < 1%.
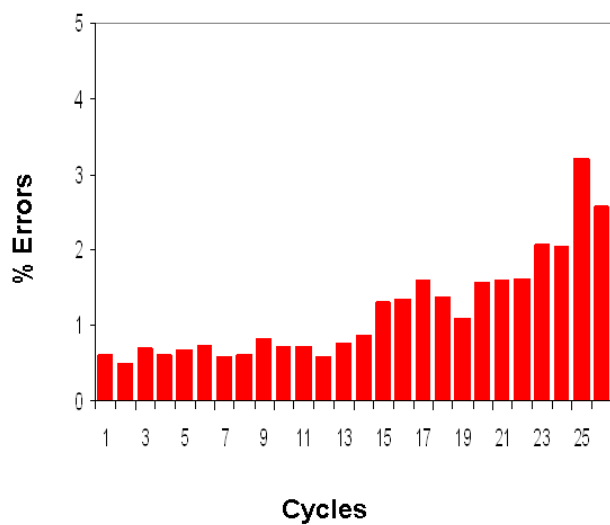
Supplementary Figure 6. Error rate by sequencing cycle for a test genome (ΦX174).

The 5.8 Kb φX174 genome was sequenced at an average coverage of 20,000-fold using a

Solexa Genome Analyzer (Illumina).
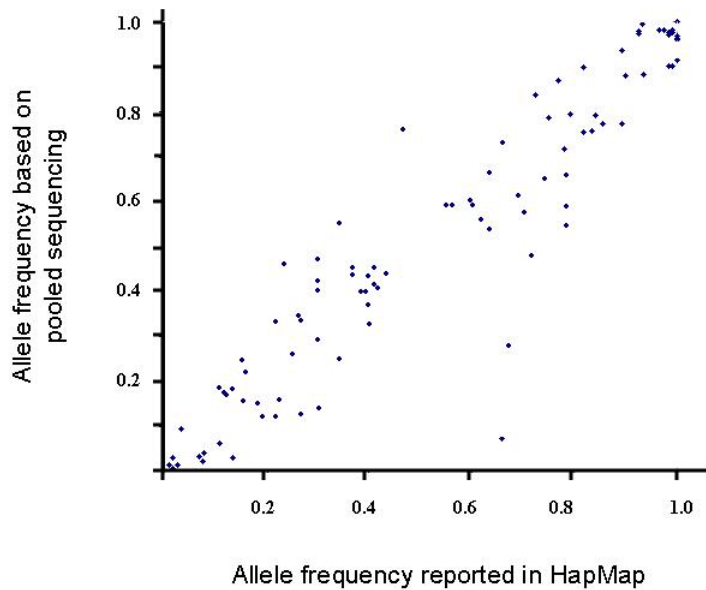
Supplementary Figure 7. Allele counts from pooled sequencing predict HapMap allele frequencies. For 95 loci, allele frequencies estimated by allele representation from resequencing 192 Finnish cases and controls in pools of 12 individuals predicted CEU HapMap allele frequencies in 60 unrelated individuals (r = 0.96).
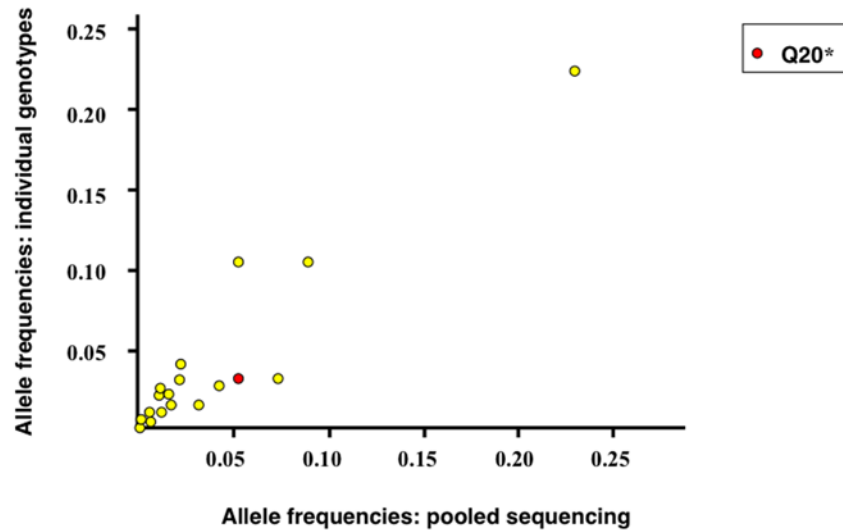
Supplementary Figure 8. Genotyping validates allele frequencies determined by pooled sequencing. For 22 nsSNPs, allele frequencies determined by allele counts from resequencing 192 Finnish cases and controls in pools of 12 individuals predicted allele frequencies determined by individual Taqman assays (r = 0.94).
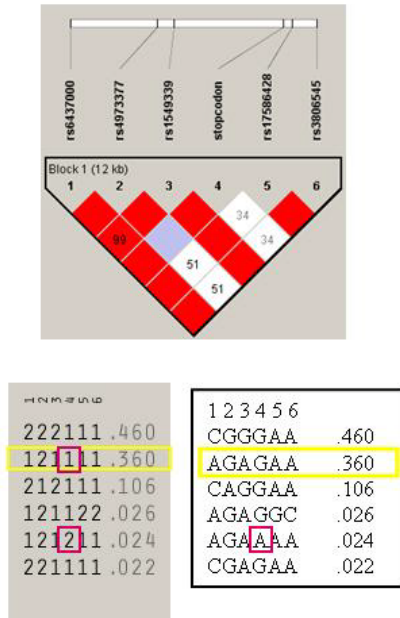
Supplementary Figure 9. *HTR2B* ***20 occurs on a single haplotype background**.

Five tagging SNPs were selected to cover the *HTR2B* gene and genotyped in the Finnish case/control cohort via a 1536 array Goldengate assay (Illumina) as described[39]. The haplotype carrying the Q20* stop codon variant is highlighted.

34
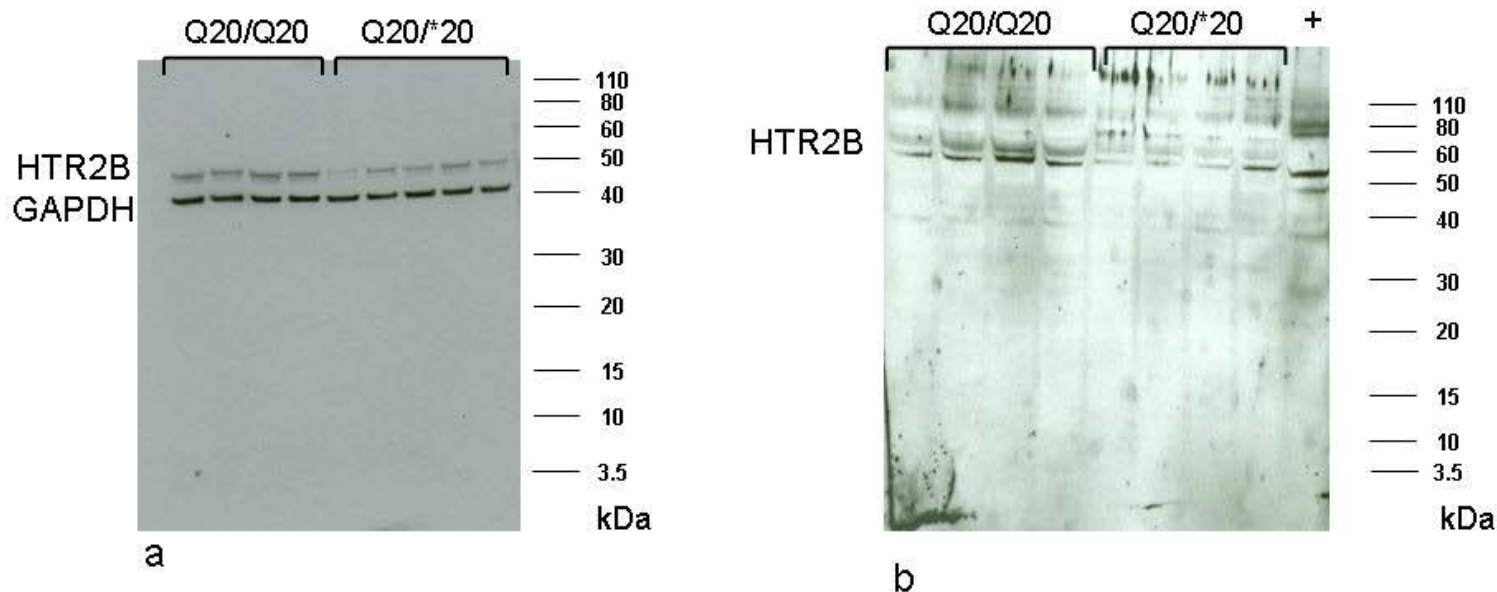
Supplementary Figure 10. Truncated 5-HT2B is undetectable in *20 heterozygotes (western blots).

No truncated 5-HT2B protein was observable in Q20/*20 stop codon carriers using antisera raised against the N-terminal (a) or C-terminal (b) regions of the receptor.
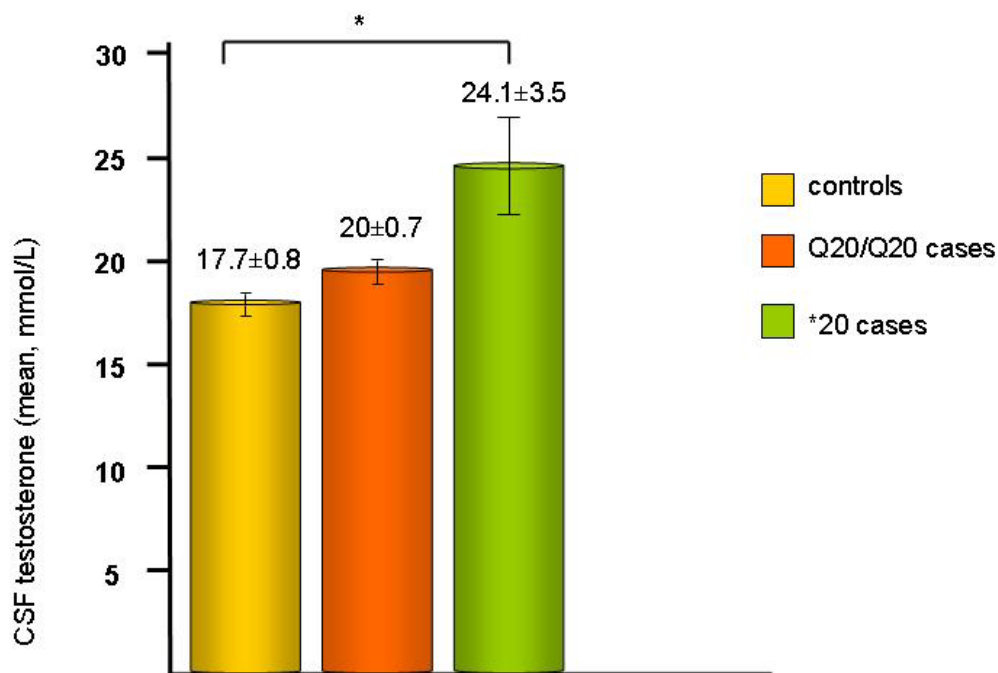
Supplementary Figure 11. High CSF testosterone in *HTR2B* *20 male carriers.

Q20* may influence testosterone levels or interact with testosterone to cause aggressive behavior. CSF testosterone was measured in 41 controls and 91 cases. Due to the skewness of these data (<0.96) log transformation was applied. CSF testosterone levels were significantly higher in the nine Q20* cases as compared to controls (t-ratio = 2.32, p-val = 0.024). Testosterone levels for *20 controls were unavailable.
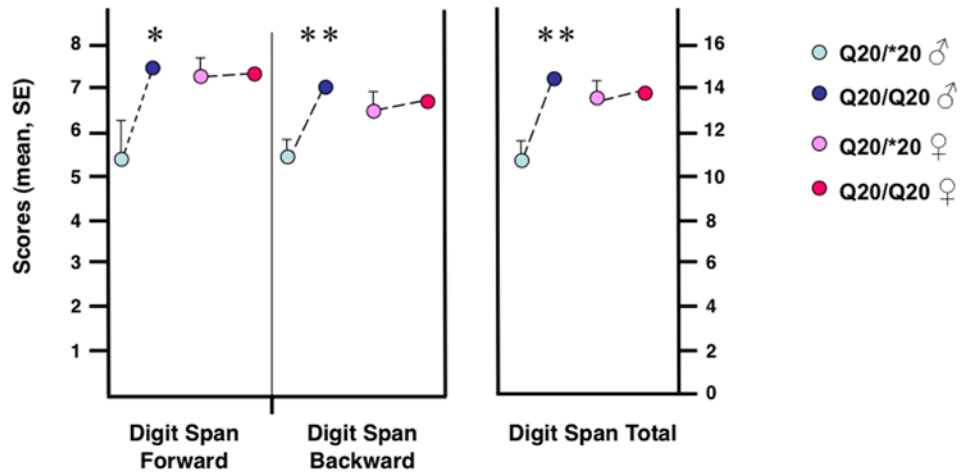
Supplementary Figure 12. Working memory impairment in *HTR2B* *20 male carriers.

Male *HTR2B* Q20/*20 heterozygotes (N = 7) had significantly worse working memory, as indicated by digit span, but were otherwise normal cognitively. Where not visible, SEM's are within the symbols. * p = 0.002; ** p ≤ 0.001.

Brain Region	ΔCt Taqman assay #1	ΔCt Taqman assay #2
cerebellum	1.3	2.5
occipital lobe	1.9	3.3
frontal lobe	2.0	3.3
parietal lobe	3.5	4.5
medulla oblongata	3.5	4.6
temporal lobe	3.7	5.0
pituitary gland	4.1	5.8
nucleus accumbens	6.0	7.8
pons	6.9	7.0
olfactory region	7.5	8.0
diencephalon	7.5	8.7
hippocampus	7.6	8.3
thalamus	7.7	8.2
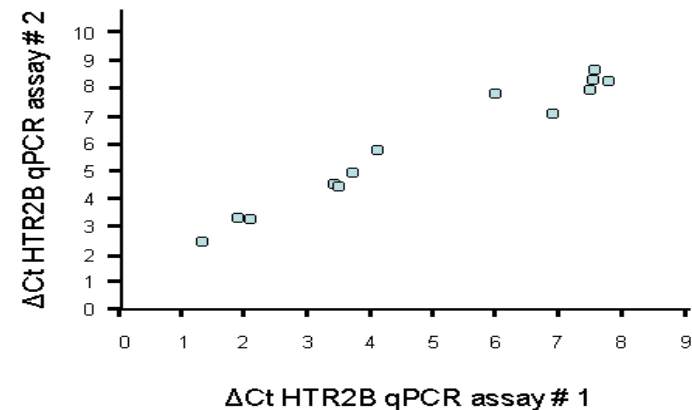
Supplementary Figure 13. *HTR2B* RNA expression in 13 human brain regions.

Rank order of *HTR2B* gene expression in human brain regions (highest expression at top) based on ΔCt values (ΔCt = Ct *HTR2B* – Ct B-Actin) from two *HTR2B* qPCR assays employing different combinations of primers and probes. Higher ΔCt indicates lower *HTR2B* expression (highest expressing regions are cerebellum, occipital cortex and frontal cortex), and each integer difference represents an approximately two-fold difference in RNA level.

SUPPLEMENTARY TABLES

Supplementary Table 1. Finnish case-control resequencing and association datasets.

Finnish case/control dataset	Cases (N = 228 ♂)	Controls (N = 295 ♂)
DSM-IIIR diagnoses	83 ASPD 39 BPD 28 IED 78 ASPD+BPD	186 Axis I and II diagnosis-free 109 other DSM diagnosis – (primarily Major Depression, anxiety disorders, addictions)
Comorbid alcohol use disorder	97%	27%
Age	32.1 (SD ± 9.2)	31.8 (SD ± 10.5)
Brown-Goodwin aggression scale total score	20.0 (SD ± 8.4)	11.3 (SD ± 7.2)
Re-sequencing subset	**Cases (N=96 ♂)**	**Controls (N=96 ♂)**
DSM-IIIR diagnoses	29 ASPD 16 BPD 15 IED 36 ASPD+BPD	No Axis I No Axis II
Comorbid alcohol use disorder	99%	None
Age	33.0 (SD ± 10.1)	32.0 (SD ± 9.5)
Brown-Goodwin aggression scale total score	23.7 (SD ± 4.9),	8.1 (SD ± 4.9)

The study protocol was approved by the Institutional Review Board (IRB) of the National

Institutes of Health and the National Institute of Mental Health (NIMH), by the Office for

Protection from Research Risks (OPRR), by the University of Helsinki Department of Psychiatry IRB, and by the University of Helsinki Central Hospital IRB. All subjects were 17 years or older at the time of the psychiatric interview. Written informed consent was obtained from each participant.

Supplementary Table 2. Region-specific primers for amplification of 108 target regions in fourteen genes.

	Chr	5'/exon	Forward Primer	Reverse Primer	bp
HTR1A	5				
		5'	GCAAAGACGCTGAGCTAGAGGGA	GGTCGGTGACCGCCAAAGAG	908
		1	GGCAACAACACCACATCACCA	CCAGGTCTGCAAGCCGTGAG	1365
HTR1B	6				
		5'	GAGCAGCCGCAACTCCAGC	CCGGTTCCTCCATGGCTCTC	693
		1	TTCAGCTGTGAACCTGGGCG	CCGCGAAAGAAGATTCGACCT	1367
HTR2A	13				
		5'	GAGCTGAAATTCCTGACAGCAGC	GCCAGAACTTGTAGCAGATGAGGTG	590
		1	CTTACAGCCACCACAGTTCAGTTCTTT	TGAAGGCTGGAATATTGGTTGGG	953
		2	TTGTGTTGCCCTTTGCCACC	GCAATACAAATTCATGTTGTCAGCCA	530
		3	GCCAATATTATGTGATGATGCTCTTGC	TCACAGCTGAAATGCTGTCAGAACA	1070
		3b	TCAAAGCAAGATGCCAAGACAACA	GCCGGAAACCTGAGTTCATCC	1015
		3c	CACAAGGTTGCTGGCATTTGCT	CCAAAGGTTTAGGGCAAACACTATTTC	1368
HTR2B	2				
		5', 1	CCAATAACATTCAACATGAGCAAAGGA	CCTTCCCAGTTGCCTTGAGAAGA	1162
		2	AGGGAAGGGCTATGAAGTTCAAGTG	AAACAAGGGACTGTTTACTTGCCGT	901
		3	CCAGGTGTCACAGAGACTCCCA	GCTGGCTTGACCTCTCATGCTG	751
		4a	CCTGCATTGTGGGTCCCAGG	TTTCCTCATGGAATAATCTTGTCCAAA	1179
		4b	TGAGGCTCCGAAGTTCAACCA	AATGGCAATCTTAACTCACCTCCAGA	568
HTR2C	X				
		5'	TGCTGCCTTGACCTTGCATTG	CATATGCAATCGGCAGGTAAGGG	927
		1	CCAGCAGGCTCCAGATGCAC	AACCTCCGCTAGGGCCAGTG	849
		2	TGCCTCAACTGTCTGGGAATGC	TGAACTTCGCTAACCTAATTGTCCCA	660

	3	TGAAGCCATGTCTACTCCAAGAGGG	ACATCGCCAAGCCCTATCCC	521
	4	TGATGATGACAATGATGCTGATGATG	CAACTCCTTCTGTCACACGATTTGC	584
	5	TTATCATGATCACCATCATCATCCCT	GAACCCGATCAAACGCAAATG	479
	6a	TGCCGTGAATAGCTTTCAACAAA	TGCAAATGCCACCTGAGATGAA	1311
	6b	CAGGTGGCATTTGCAGGTGA	CACAGCATAGACTGAATGGCCTGA	1074
	6c	AAGTGGCAAGCTCTTCATTGCAC	GGCAGTCTGTTGCACGTGTCTT	1268
	6d	TCGTAAGTTCTGTGCAGTTTGGTATGA	AGGAATTGGTGATGCTGTTGTTGA	600
HTR3A **11**				
	5', 1	AAACGTTAAGAGTGGCTGTCTCTGAGT	CCCTTCTTTGCAAGCTCCGC	1128
	2	TGCCACTTGCTCTTCCAAGCC	CCATCCTGAGCTGTGGTGCC	554
	3	TGGCCAGGGAGCTCCAGTCT	CCCAATTTGGTTGCTGTCCAA	582
	4	GCGCCACTGACACCAGCTTC	TCACTGGTTTCTTTCCCTGTGGC	521
	5	CTGCCTTGTGTGGCCAGGAG	AGCAGAGGCCCAGGGCATTA	551
	6	AGGTTGAGGTTGGGCCATCG	GGCAAACAGGCTCTGCTCCC	595
	7	TGGAGAAACTCCATGAATGTCCCTT	TGACATGCTCAGCCTCTGGGA	982
	8	CATGTTCAGGTCACCACCCG	TCCCAAAGCTGCTGACAAATGG	1163
HTR3B **11**				
	5', 1	TCATCTTCAATAAGCCTAGAAGCATCC	TGCAATCATAGGAGCCAAGTCACA	944
	2	TCCATTGGGCACTTAGGCCA	AGCGCAACTTCCCTGCCAGT	482
	3, 4	GCTCATAGCGACAGGTGTGCATAAT	CCATGAGATGATTCACCCTGGC	1066
	5, 6	GCTCATCTTTGCCAGGGTGAATC	GCCCGCTGAAGTCCCTGTTC	1099
	7	TTGGCCAAATCTGAGTCTGTTGG	GCCAGGTGTAGTGGTGTGTGCC	578
	8	TGCTGCATTTCAGACTGGAGGC	GAAAGCCAGAAGCAGGCAGC	417
	9	CTTTGCCTCGCCTAATCGGG	GCTTTGGGTGAGTGGACTGGG	951
DRD1 **5**				
	5', 1	ACCTGGGCGAAGCGGGTT	CAGGACTGGTCACCTGGGCA	1279
	1, 2a	GGGCTCTCGAAAGGAAGCCA	ACACTGATCAGGATGAAGGCTGC	1383
	2b	GGCAGTGGCTGAGATTGCTGG	CACTGTTGATTCTTTGCCCTATTTGG	1408
	2c	GAGCCCTCTGCTGCTTTCCA	GCAGGGTTTGAGTTTGGTCCC	1066

DRD2	**11**			
	5', 1	GAAGAACGAGGTTCCTCTGCAGCCC	AAACTTCTGGTCCTGGCCTCGCTCC	1041
	2	TCCCTGTTCCATTTCCTACTCCC	GGGATCCTTGCCGGTTACTCC	1292
	3	CCAGCTGGATGCCCACAAGA	TCCGTCAGAATCACAATCTTTCCC	519
	4	CGTTGTGTCCTGGGCCACTG	TGCCAACTGTCCTCGCCATC	543
	5	GGGAGCAGTCCATCATCACAGC	GCTCCACTGAGGTCCATGCC	600
	6	GGCCCTTAGCAGGTGCTCCA	CCATGCCTTGCAGAACCGAG	520
	7	CCTCAGTGACATCCTTGCCTCC	GCTTGGCCTGTGCCTGAGGA	597
	8a	TGGGAAAGGGACAAATGAGGG	GGGTGCGGGAGGGAGCTAAG	992
	8b	GCTCTAGGGTTGCTGGAGCCTG	GAAGCAGGAGCTGCAACAGGG	933
DAT	**5**			
	5', 1	CCTCGGTGCCTTCTAAGGACCTGGA	ACTGTTGGCGACTTTGGAGACGGC	1118
	2	CCCAGAGGAATGCCCGTGA	GGGAGCTCCGTCTTCACGCA	569
	3	GGCCTCCCTGTGCTGCTCTT	ATGATGCGGCTGGCTTGCTT	485
	4	TGCTGTCTGGCATCAGGGCT	GAGGGCTCGCCCTACAGGAG	591
	5	CGTGCTCCACGGGTTTCAAG	CGTCTACCCAAGCCAACCCG	600
	6	GCATTTGCTGAGCTGGCGTC	TCTCCGGGAGGCTTTCTCCA	558
	7	CCCGGGAGGGTTCACTTCCT	GCGACACACCCACTGGCTTC	505
	8	TTCTGCCCACCCTGAGCCTT	CCATCCTTCAAGAGTGAGGCAGC	509
	9	GGACCCAGGTAATTTGGAGGAGCCC	CAGCCCCGAGAAAGGTCTCCCAAAT	544
	10	GCCATAGAAGCCCTCGCCAA	AGGAGGCAGCTTCAGCCTGG	557
	11	GCCAAGTTCCGCATTCGGTT	CCCTCTGCGGAGCTGTGATG	491
	12	CACGGCTCTTCCTGGTGCAT	TTTCGGTGGGTCTAGAGGGGAG	543
	13	GATGCCCAGCTGCCATGTCT	GGCCACCTCCAGTCTCCTCC	452
	14	GTGGGCGCTGGTGATATGGA	ATGCCCAGGTCTTTGCAGGG	502
	15a	CGGCACCTGTCAAGGGTGTG	ACGCTCCTCTCAGGCCGTTC	1395
	15b	CTGGTATGTCTCACCAGGAAATTCTGT	AGCTTTGCTGTGGGTGGGCT	1439
PPP1R1B	**17**			
	5'	CTCCCGCTCCAGAGACACACACAA	TGCGGTCTCCCCAGAACTTG	808

			Forward Primer	Reverse Primer	Size
		1	CCAGGACCACGAAGGGGTTAAAGCC	CGGGGAAAAGTTGCACGAGTCTCCA	1028
		2	ACAGTCCTGGGCCTAAGCAGC	TCCCTTCCTATCTTCTCCCACCA	586
		3, 4, 5	TGGTGGGAGAAGATAGGAAGGGA	GACAGCGGGCTATGTGTGGC	1322
		6	CTGCCCTGGTCTGACACCCA	CCCATGGAGCGAGAAGACAGG	527
		7, 8a	TCCATGCTTCCATGATGCCC	ACTAGGCGGCATTCCTGGCA	942
		8b	CCCTTAATCACCCTTGCTCCTCC	AGGCCTGGCACTGAGGGAAA	550
MAOA	**X**				
		5'	CACCAGTACCCGCACCAGTA	CTGGGTCGTAGGCACAGGAGT	1150
		1	CTAAACCTAATAACTCTCGC	TACACTACACTCTCGCATGC	468
		2	ATCAGGATTTCTGTATGAGG	ACGCTTATTTGATGGGAGTA	599
		3	GCCTTACCAATAACTTTACC	CCAGGAGACTTTAGAGATTTT	557
		4	CAGCACCAGGGAGATTCTGC	GGTTCCCTTGCTATGAGGTC	533
		5	GTTTTGGCAATGGCATTAGG	GGCTGCATAGTATTCCATGG	642
		6	CAACTCATAATTTCTTAAGCAGGAG	GCTACCAAAAGGAAGGAACT	600
		7, 8	ATCATTGAATTGTCTTGGTC	CCTAAAAGACTGTGATTTCT	851
		9	CCCTTTATGTAAAGGTGATGGAGAT	GGTACAAGTTCAAAATGGTGTCAAT	543
		10	TGCCAAAAACCATGTTAAAC	GAAATCCTACCTCTTAACTACTGTA	547
			TTCCCATTTTTATCTCGAGAAG	GCTTTATACTCTCTTCATAAGACTG	
			TGGGGAAGTAAAGTCTGTTA	ACAGGTCTTGGATTAATTGG	
			CTTTAATGATTGATGATCTGATGGC	CCTGGGATAGTTGCATCTATAATTA	
		15a	TATAGATGCAACTATCCCAG	GTAAGTTGAGAGGTATGTCTAA	
			AACTTCAGGAAAGCAGCAGT	AAGCCATTGAACTTTCTGCA	
		VNTR	TAGAGTCACTTCTCCCCGCCCCTGA	GTGTCCTGATTGGAGGCTGTGACGG	643
SLC6A4	**17**				
		VNTR	TGTTCATCTGAAAGGAGGAGGCCCC	GGATTCTGGTGCCACCTAGACGCCA	554
		1	AGGCGCACACCTGCTCCTTT	ATGCACCTCCTCGCCTCCTC	584
		2	GCTGAGCTGACTTCCCTGCG	CTGCCTTTCAGCAGGCAGGA	544
		3	TGGACTGCCATGTAGCAAATAGGA	GACATAATCTGTCTTCTGGCCTCTCAA	861
		4, 5	GACACTCACATCATCCGGCGA	CAGGCCGTGGAGCACTTGAG	1077

	6, 7	GGGCACCCTCAAAGGAGCAG	GCCACATTTCAGTTGTCCCTACTGTC	1385
	8	TTGAGGCCACTGGATGGGTTT	CGGCCTGGATGAGTGCTCTG	501
	9	TCTAAGGCAGGGCAGGCACA	AGGCATCGAGCTTCTCTCAAATCTT	595
	10, 11	CCCAGGGTTAATGGAGGGCA	GCCAGGGCACTGTGTGAGATG	1236
	12	GCAAAGGCCAGTCCAAGAAC	CATCGGGAGGTCACATCTTG	675
	13	CAGCTCTCAGGAAGGCGGTG	GGAAGTCTTTCGCCAGGGCA	536
	14	TCAGGGAAAGCACCTGGCTG	TTTGCCCTTGCATTTGAGGC	586
	15	GGCCGCAGAAGTCTCCTCTGT	GCAGAGCATGTTGTAGTGATTGTACCC	1152
TPH2	**12**			
	5'	TGGAGTTGGTCAATGAGTAA	GAAACTCATTTCTGGTTGCC	600
	1	CCCTTATGTATTGTTCTCCACCACCC	GCAGACAGGAGGAAAGCAGGG	600
	2	CACGGCAACTTCACCTTAAG	CCTGGCAATAGTCTTCAGAG	533
	3, 4	TGAGGCAGGACAACCTATACACAGAA	CCTCCCTAGTTTCCCAGAGGCA	970
	5	GCCTGAATTGCCACACATCCT	AGCACTTGGCATGTGGCTCA	503
	6	TGTGCTCCACTTTGCCAGGG	GCCACAAAGAGGCCTCAGCA	522
	7	TCTCCTGGAAACCGTCATTTGAGA	CCCACGGACCCAGATGAGGA	531
	8	TGAGATGCAGTGGGATTTCAGACA	TCTGACAGAGTCCTTACGACCCACA	580
	9	TTTGACTCATAGGGCTTCTG	TGACCCATTTTCACAAGTGC	528
	10, 11	CCCTGCACACAGGAGAGTTCCA	CACATGCAAGCACTGGGACG	1533

Supplementary Table 3. **Sequence variants detected by exon-centric sequencing of 14 genes.**

Exon-centric resequencing of 14 genes (81.8 kb) was performed in 96 unrelated Finnish cases and 96 controls, yielding 360 SNPs. Non-coding SNPs: 1/144 bp. Coding SNPs: 1/783 bp, including synonymous SNPs: 1/1472 bp and non-synonymous SNPs: 1/1673 bp.

Gene	Total bp	Coding			Non-coding	
		bp screened	Synonymous SNPs	Missense and Stop	bp screened	SNPs
DRD1	4692	3373	1	2	1319	12
DRD2	5209	2579	2	2	2630	22
HTR1A	2114	1269	1	0	845	0
HTR1B	1979	1173	2	1	806	6
HTR2A	4838	3009	4	3	1829	20
HTR2B	4414	2227	0	4	2187	9
HTR2C	7050	4169	0	1	2881	22
HTR3A	5906	2221	4	0	3685	26
HTR3B	5429	1817	1	3	3612	31
MAOA	10583	4073	3	1	6510	27
PPP1R1B	4681	1841	0	1	2840	13
SLC6A3	9186	3925	4	1	5213	54
SLC6A4	8987	2766	0	2	6221	44
TPH2	6776	2360	3	1	4416	27
TOTAL	81796	36802	25	22	44994	313

Supplementary Table 4. Frequencies and secondary single marker associations for 21 additional nsSNPs.

Subsequent to global analysis with four putatively functional SNPs and single locus analysis of Q20*, 21 additional nsSNPs were secondarily gentoyped in the Finnish case/control dataset. Single marker associations were performed using the Pearson chi-square test. Shown are uncorrected p values. Cases were either ASPD, BPD or IED.

						Cases N=228		Controls N=295			
Gene	id	Ss #	Exon	A1	A2	%A1	%A2	% A1	%A2	χ^2	pval
DRD1	Lys81Arg	159816269	1	A	G	0.92	0.08	0.91	0.09	0.32	0.571
DRD1	Ser259Tyr	159816261	1	G	T	0.98	0.02	0.99	0.01	1.046	0.307
DRD2	Lys327Glu	159816263	6	T	C	0.99	0.01	0.99	0.01	0.262	0.609
DRD2	rs1801028		6	G	C	0.97	0.03	0.98	0.02	1.568	0.210
HTR1B	rs1130060		1	A	C	0.99	0.01	0.98	0.02	1.720	0.190
HTR2A	rs6314		3	G	A	0.97	0.03	0.96	0.04	0.222	0.638
HTR2A	rs6308		3	C	T	0.99	0.01	1.00	0.00	5.252	**0.022**
HTR2A	rs1805055		1	G	T	0.98	0.02	0.97	0.03	1.081	0.298
HTR2B	Arg388Trp	159816262	3	C	T	0.98	0.02	0.99	0.01	1.641	0.200
HTR2B	Phe173Leu	159816264	2	A	G	0.97	0.03	0.97	0.03	0.005	0.945
HTR2B	Gln45Glu	159816265	1	G	C	1.00	0.00	0.998	0.002	0.777	0.378
HTR2C	rs6318		2	G	C	0.88	0.12	0.87	0.13	0.059	0.807
HTR3B	rs17116138		6	G	A	0.91	0.09	0.91	0.09	0.002	0.966
HTR3B	rs1176744		5	A	C	0.70	0.30	0.70	0.30	0.007	0.935
HTR3B	Ser156Arg	159816266	5	A	C	0.998	0.002	0.998	0.002	0.035	0.853
MAOA	Glu188Lys	159816267	6	G	A	0.996	0.004	1.00	0.00	1.280	0.258
PPP1R1B	rs35797948		7	G	A	0.97	0.03	0.97	0.03	0.479	0.489
SLC6A3	Val471Ile	159816268	10	C	T	1.00	0.00	0.998	0.002	0.775	0.379
SLC6A4	rs6352		12	A	C	0.96	0.04	0.95	0.05	0.283	0.595
SLC6A4	rs6355		1	G	C	0.98	0.02	0.99	0.01	0.272	0.602
TPH2	rs17110563		5	C	T	0.998	0.002	1.00	0.00	1.275	0.259

Supplementary Table 5. Predicted functionality of missense/nonsense variants via PolyPhen and SIFT.

From among 27 missense/nonsense variants, variants predicted to be probably damaging or damaging for protein function via PolyPhen and SIFT amino acid substitution prediction methods are shown. Four variants (*DRD1* S259Y*, HTR2B* R388W*, HTR2B* Q20*, and *TPH2* P206S - rs17110563) scored as damaging or intolerant by both methods were used in a global test of proportion of rare functional variants in cases and controls. (Sequencing estimated frequencies for allele 1 (A1) and allele 2 (A2) are shown).

Gene	Position	Cases N=96		Controls N=96		aa>aa	Exon	Ref aa	Mut aa	PolyPhen scores	SIFT
		A1	A2	A1	A2						
DRD1	S259Y	0.98	0.02	1	0	c>a	1	S	Y	2.101	intolerant
DRD2	rs1801028	0.97	0.03	1	0	c>g	6	S	C	1.841	intolerant
HTR2A	rs6314	0.96	0.04	0.98	0.02	c>t	3	H	Y	1.503	tolerant
HTR2B	R388W	0.98	0.02	0.995	0.005	c>t	3	R	W	2.913	intolerant
HTR2B	Q20*	0.97	0.03	0.99	0.01	c>t	1	Q	Stop codon		intolerant
HTR3B	S156R	0.995	0.005	0.994	0.006	a>c	5	S	R	1.76	intolerant
TPH2	rs17110563	0.99	0.01	1	0	c>t	6	P	S	2.096	intolerant

- damaging
- probably damaging
- possibly damaging

Supplementary Table 6. Global test of association for four putatively functional SNPs, and case/control association of *HTR2B* *20.

In a global test for excess of rare functional variants in cases (ASPD, BPD or IED), genotypes of four SNPs classified as functional by both PolyPhen and SIFT (*DRD1* S259Y, *HTR2B* R388W, *HTR2B* Q20*, and *TPH2* rs17110563) were collapsed so that an individual was coded as 1 if a rare allele was present and otherwise as 0. Frequencies of putatively functional variants were globally compared between cases and controls, with the null hypothesis being a lack of difference between cases and controls in the proportion carrying the putatively functional variants.

Case-control association test was also performed for *HTR2B* Q20* alone.

Pearson $\chi 2$ test was used to test the null hypothesis.

Global association test for four putatively functional SNPs in the Finnish case-control cohort.
PolyPhen functionality prediction scores >2, and predicted "intolerant" via SIFT analysis

Cases N=228		Controls N=294		$\chi 2$	p value
1	0	1	0		
30	198	19	275	6.76	**0.009**

Association test for HTR2B Q20*alone.

Cases N=228		Controls N=295		$\chi 2$	p value
Q20	*20	Q20	*20		
439 (.963)	17 (.037)	583 (.988)	7(.012)	7.26	**0.007**

Supplementary Table 7. *HTR2B* *20 co-segregates with ASPD in eight informative families.

Informative meioses in 8 Finnish pedigrees, from among 89 that where comprehensively genotyped.

A = affected (ASPD)

U = unaffected

Affected status was defined as presence of ASPD, BPD, IED, and in these families all cases where ASPD. Meioses supporting linkage of *20 to impulsivity are on the left and meioses with genotype/phenotype configurations contrary to the hypothesis of linkage are on the right side of the table. Superscripts indicate the number of individuals in the same category within the family, so for example in Family one, 2/2 informative meioses are in the linked configuration, with one of the offspring in that family affected (and carrying *20) and one unaffected offspring being a Q20/Q20 homozygote.

Q20* stop codon

Family	linked		unlinked
1	A	U	
2			U
3	A	U^4	U
4	A	U^2	U
5	A		
6		U^2	U
7	A	U	
8	A		A

Supplementary Table 8. *HTR2B* *20 was undetectable in other populations worldwide.

The *HTR2B* *20 stop codon was genotyped in >3100 individuals representative of worldwide ethnic diversity and including DNAs from the Human Genome Diversity Panel (HGDP). Only one additional *20 carrier was observed, a female alcoholic with a Finnish surname.

Populations	ASPD/IED/BPD cases	Controls	Q20*
Caucasians	247	473	1
American Indians	56	334	0
African-Americans	268	661	0
CEPH	0	1064	0
Total	571	2532	1

Supplementary Table 9. Frequency (0.012) of *HTR2B* *20 in Finland including case/control, psychosis, and twin datasets (total N = 6554).

The *HTR2B* *20 variant has an allele frequency of 0.012 in the Finnish population, as represented by the epidemiologically ascertained Older Finnish twin cohort, FinnTwin12 and FinnTwin16 cohorts. This allele frequency is in agreement with the allele frequency of 0.012 determined in controls belonging to the Finnish violent offender case/control and similar to the frequency in Finnish Bipolar and Schizophrenia families (0.009). The frequency of *20 is 0.037 among Finnish violent offenders as compared to 0.012 in controls in that same case/control sample (p= 0.007).

Allele frequencies in the Finnish Twin Cohorts are corrected for genetic relationship, and cases and controls in the violent offender case/control study are unrelated.

Finnish datasets	*HTR2B* *20 Allele freq
Finnish violent offenders - cases	**0.037**
Finnish violent offenders - controls	0.012
Finnish Bipolar and Schizophrenia families	0.009
The Older Finnish Twin Cohort	0.013
FinnTwin12 + FinnTwin16 studies	0.012

Supplementary Table 10. Distribution of *HTR2B* *20 in a Finnish psychosis dataset (N = 1885).

Finnish families with at least one family member affected by bipolar disorder or schizophrenia. Among 1885 genotyped individuals psychosis was numerically, but not statistically, more prevalent in *20 carriers. Early-onset schizophrenia was observed among the *HTR2B* *20 carriers (mean 20.9; median 19 years) compared to Q20/Q20 individuals (mean 24.5; median 22 years), p = 0.035 (Student's two-tailed unequal variance T-test).

Bipolar Disorder + Schizophrenia Families dataset	Q20/Q20	Q20/*20
Psychosis	668	16 (0.012)
Control	1185	16 (0.007)
Total	1853	32

Supplementary Table 11. Distribution of *HTR2B* *20 in three Finnish twin cohorts.

Finnish Older Twin cohort, FinnTwin12 (FT12) and FinnTwin16 (FT16) datasets were genotyped for *HTR2B* Q20*. One *20/*20 homozygous individual was identified in a dizygotic twin pair belonging to the FT16 cohort. Allele frequencies are reported for each cohort after correction for genetic relationships. Each monozygotic twin pair, because of genotypic identity, was counted as one individual. Dizygotic twin pairs were corrected for 1.5 independent genomes, because 50% of their genome is shared.

	Twin pairs								
The Older Finnish Twin Cohort	*1-1*	*1-2*	*2-2*	*2-3*	*1-U*	*2-U*	*Corrected N*	*Corrected N of Q20/*20*	**20 Allele frequency*
MZ	237	0	11	0	0	0	248	11	**0.0131**
DZ	541	19	6	0	743	17	1609	37.7	
FT12 + FT16									
MZ	344	0	7	0	0	0	351	7	**0.0120**
DZ	420	14	4	1	98	3	759.5	19.7	

Twin pairs:

1 = Q20/Q20
2 = Q20/*20
3 = *20/*20
U = unknown

Supplementary Table 12. Selective impairment of working memory in male *HTR2B* *20 carriers.

Relationship of Q20* to working memory in the **FT16** and **FT12** cohorts (424 males and 509 women). Working memory was assessed with the Digit Span Forward and Backward subtests of the Wechsler Memory Scale-Revised (WMS-R). A linear regression model that included sex and genotype was applied to test for difference in working memory between *20 carriers and non-carriers (means and SE are reported). Male and female *20 carriers showed no difference in performance on other cognitive tests listed in Supporting Online Materials. p values (Bonferroni corrected) are shown in italics in parenthesis.

Working memory	♂ N=7 *20	♂ N=417 Q20	♀ N=15 *20	♀ N=494 Q20	♂ P value	♀ P value	♂ & ♀ P value
Digit Span Forward	5.4 (0.7)	7.5 (0.1)	7.2 (0.4)	7.3 (0.1)	**0.002** *(0.008)*	0.92	0.08
Digit Span Backward	5.4 (0.3)	6.9 (0.1)	6.4 (0.5)	6.6 (0.1)	**<0.001** *(0.004)*	0.72	0.06
Digit Span total	10.8 (0.9)	14.5 (0.2)	13.7 (0.8)	13.9 (0.1)	**0.001** *(0.004)*	0.74	0.02

REFERENCES

40. Svenningsson, P. et al. DARPP-32: an integrator of neurotransmission. *Annu. Rev. Pharmacol. Toxicol.* **44**, 269-296 (2004).

41. Zhou, Z. et al. Haplotype-based linkage of tryptophan hydroxylase 2 to suicide attempt, major depression, and cerebrospinal fluid 5-hydroxyindoleacetic acid in 4 populations. *Arch. Gen. Psychiatry.* **62**, 1109-1118 (2005).

42. Parsey, R.V. et al. Effects of sex, age, and aggressive traits in man on brain serotonin 5-HT1A receptor binding potential measured by PET using [C-11]WAY-100635. *Brain Res.* **954,** 173-182 (2002).

43. Crabbe, J.C. et al. Elevated alcohol consumption in null mutant mice lacking 5-HT1B serotonin receptors. *Nat. Genet.* **14**, 98-101 (1996).

44. Lappalainen, J. et al. Linkage of antisocial alcoholism to the serotonin 5-HT1B receptor gene in 2 populations. *Arch. Gen. Psychiatry* **55**, 989-994 (1998).

45. Talpos, J. C., Wilkinson, L.S., Robbins, T.W. A comparison of multiple 5-HT receptors in two tasks measuring impulsivity. *J. Psychopharmacol.* **20**, 47-58 (2006).

46. Ducci, F. et al. HTR3B is associated with alcoholism with antisocial behavior and alpha EEG power--an intermediate phenotype for alcoholism and co-morbid behaviors. *Alcohol.* **43**, 73-84 (2009).

47. Nicholson, P. et al. Nonsense-mediated mRNA decay in human cells: mechanistic insights, functions beyond quality control and the double-life of NMD factors. *Cell. Mol. Life Sci.* **67**, 677-700 (2010).

48. Geiger, S.K. et al. Incomplete nonsense-mediated decay of mutant lamin A/C mRNA provokes dilated cardiomyopathy and ventricular tachycardia. *J. Mol. Med.* **86**, 281-289 (2008).

49. Magyar, I. et al. Quantitative sequence analysis of FBN1 premature termination codons provides evidence for incomplete NMD in leukocytes. *Human Mutation* **30,** 1355-1364 (2009).

50. Palo, O.M. et al. Identification of susceptibility loci at 7q31 and 9p13 for bipolar disorder in an isolated population. *Am. J. Med. Genet. B Neuropsychiatr. Genet.* **153B**, 723-735 (2009).

51. Paunio, T. et al. Genome-wide scan in a nationwide study sample of schizophrenia families in Finland reveals susceptibility loci on chromosomes 2q and 5q. *Hum. Mol. Genet.* **10**, 3037-3048 (2001).

52. Kaprio, J., Koskenvuo, M. Genetic and environmental factors in complex diseases: the older Finnish Twin Cohort. *Twin Research* **5**, 358-365 (2002).

53. Järvenpää, T., Rinne, J.O., Koskenvuo, M., Räihä, I., Kaprio, J. Binge drinking in midlife and dementia risk. *Epidemiology* **16,** 766-771 (2005).

54. Bortner, R.W. A short rating scale as a potential measure of pattern A behavior. *J. Chron. Dis*. **22**, 87-91 (1969).

55. Romanov, K. et al. Self-reported hostility and suicidal acts, accidents, and accidental deaths: a prospective study of 21,443 adults aged 25 to 59. *Psychosom. Med*. **56,** 328-326(1994).

56. Kaprio, J., Pulkkinen, L., Rose, R. J. Genetic and environmental factors in health-related behaviors: studies on Finnish twins and twin families. *Twin Research* **5**, 366-371 (2002).

57. Kaprio, J. Twin studies in Finland 2006. *Twin Research and Human Genetics* **9**, 772-777 (2006).

58. van Gaalen, M.M., van Koten, R., Schoffelmeer, A.N., Vanderschuren, L.J. Critical involvement of dopaminergic neurotransmission in impulsive decision making. *Biol. Psychiatry* **60**, 66-73 (2006).

59. Loos, M. et al. Dopamine receptor D1/D5 gene expression in the medial prefrontal cortex predicts impulsive choice in rats. *Cereb. Cortex* **20,** 1064-1070 (2009).

60. Pezze, M.A., Dalley, J.W., Robbins, T.W. Differential roles of dopamine D1 and D2 receptors in the nucleus accumbens in attentional performance on the five-choice serial reaction time task. *Neuropsychopharmacology* **32**, 273-283 (2007).

61. Santarelli, L. et al. Requirement of hippocampal neurogenesis for the behavioral effects of antidepressants. *Science* **301**, 805-809 (2003).

62. Kittles, R.A. et al. Dual origins of Finns revealed by Y chromosome haplotype variation. *Am. J. Hum. Genet.* **62**, 1171-1179 (1998).

63. Nelis, M. et al.Genetic structure of Europeans: a view from the North-East. *PLoS One* **4**, e5472 (2009).

64. McEvoy, B.P. et al. Geographical structure and differential natural selection among North European populations. *Genome Res*. **19**, 804-814 (2009).

65. Zhao, Z., Boerwinkle, E. Neighboring-nucleotide effects on single nucleotide polymorphisms: a study of 2.6 million polymorphisms across the human genome. *Genome Research* **12**, 1679-1686 (2002).

66. Cargill, M. et al. Characterization of single-nucleotide polymorphisms in coding regions of human genes. *Nat. Genet*. **22**, 231-238 (1999).

67. Scheinin, M., Chang, W.H., Kirk, K.L., Linnoila, M. Simultaneous determination of 3-methoxy-4-hydroxyphenylglycol, 5-hydroxyindoleacetic acid, and homovanillic acid in cerebrospinal fluid with high-performance liquid chromatography using electrochemical detection. *Anal. Biochem*. **131**, 246-253 (1983).

68. Ducci, F. et al. Increased anxiety and other similarities in temperament of alcoholics with and without antisocial personality disorder across three diverse populations. *Alcohol* **41**, 3-12 (2007).