Supplementary Table 1. Metabolites identified and quantified.

Metabolites identified and quantified with the LCECA platform are listed.

Supplementary Table 2. Associations between metabolites and clinical phenotypes.

Associations of baseline and changes in metabolite concentrations with clinical measures of response are shown. Of the metabolites measured with the LCECA platform, serotonin was associated with the largest number of clinical response measures. P-values <0.05 are highlighted in yellow and bolded.

Supplementary Table 3. Top SNP p-values and beta estimates for the baseline plasma serotonin GWAS.

SNPs from the baseline serotonin GWAS are listed by lowest p-value. Beta values indicate that the *TSPAN5* SNP minor alleles were associated with higher baseline serotonin concentrations while *ERICH3* SNP minor alleles were associated with lower baseline serotonin concentrations.

Supplementary Table 4. Number of patients with each genotype combination.

The number of patients with each allele combination for the *TSPAN5* SNP (rs11947402) and the *ERICH3* SNPs (rs11210490 and rs11580409).

Supplementary Table 5. *ERICH3* and *TSPAN5* SNP associations with response.

The associations between clinical response for the *TSPAN5* "top" SNP and the *ERICH3* nsSNP (rs11580409) that resulted in proteasome-mediated degradation are shown for three independent SSRI clinical studies (PGRN-AMPS, ISPC, and STAR*D). The *ERICH3* nsSNP (rs11580409) was associated with response at four weeks in the ISPC population and with response at six weeks in the STAR*D population.

Supplementary Table 6. EMSA probesets.

Probesets used to perform the EMSA experiments. Two sets of each probe were prepared: those with or without biotin labels. The red, bolded bp are those that were varied at SNP locations.

Supplementary Table 7. Western blot antibodies.

Antibodies used to perform the Western blot analyses are listed.

Supplementary Table 8. Primers for *ERICH3* site-directed mutagenesis.

Primers used to generate *ERICH3* variant allozyme expression constructs are shown. The red, bolded bp represent the variant SNP bases.

Supplementary Figure 1. GWAS for four and eight week change in serotonin concentrations.

GWAS for change in plasma serotonin concentrations after (**a**) four weeks and (**b**) eight weeks of SSRI therapy. The SNP clusters for the baseline serotonin concentration GWAS on chromosome 4 (*TSPAN5*) and chromosome 1 (*ERICH3*) were also identified in these GWAS for change in plasma serotonin concentrations.

Supplementary Figure 2. GWAS QQ plots.

QQ plots for the GWAS for (**a**) baseline plasma serotonin and for change after (**b**) four or (**c**) eight weeks of SSRI therapy.

Supplementary Figure 3. Average baseline and four and eight week change in plasma serotonin concentrations by genotype.

The variant *TSPAN5* SNP (rs11947402) was associated with higher (**a**) baseline plasma serotonin concentrations and a larger decrease of serotonin concentrations after (**b**) four and (**c**) eight weeks of SSRI treatment. The variant *ERICH3* nsSNP (rs11580409) was associated with lower (**d**) baseline plasma serotonin concentrations and smaller decreases in serotonin concentrations after (**e**) four weeks and (**f**) eight weeks of SSRI treatment when compared to baseline. *** = $p<0.0001$.

Supplementary Figure 4. TSPAN5 expression in human tissues.

TSPAN5 expression level in different human tissues obtained from the Genotype-Tissue Expression (GTEx) database (Data Version 6).[38] Brain regions were among the tissues that demonstrated the highest *TSPAN5* expression.

Supplementary Figure 5. ERICH3 expression in human tissues.

ERICH3 expression level in different human tissues obtained from the Genotype-Tissue Expression (GTEx) database (Data Version 6).[38] Brain region were among the tissues that demonstrated the highest *ERICH3* expression.

Supplementary Figure 6. *TSPAN5* SNP EMSA.

EMSA for three SNPs 5' of *TSPAN5* (rs1918743, rs59961429 and rs56095565) that were predicted by TRANSFAC to disrupt or create transcription factor binding sites. WT = wild type SNP genotype; V = variant SNP genotype.

Supplementary Figure 7. RBPJ- κ knockdown.

RBPJ-κ knockdown in SK-N-BE(2) neuroblastoma cells was associated with increased expression of *SLC6A4*, *TPH1*, *TPH2*, and *DDC* enzymes when compared to baseline. * = $p < 0.05$; ** = $p < 0.01$.

Supplementary Figure 8. Average baseline and change in plasma serotonin concentrations by *TSPAN5* and *ERICH3* SNP genotypes in combination.

Patients homozygous WT for the *TSPAN5* SNP (rs11947402) and homozygous variant for the nonsynonymous (ns) *ERICH3* SNP (rs11580409) that resulted in accelerated ERICH3 proteasome-mediated degradation were associated with lower average **(a)** baseline serotonin concentrations and smaller decreases in serotonin concentrations after **(b)** four and **(c)** eight weeks of SSRI treatment. Moreover, patients who were either homozygous variant or heterozygous for the *TSPAN5* SNP (rs11947402) and homozygous WT for one of the *ERICH3* SNPs (rs11580409 or rs11210490) displayed the highest average **(a)** baseline serotonin concentrations and the largest decreases in serotonin concentration after four or **(c)** eight weeks of SSRI treatment. The values listed in the graphs represent average **(a)** baseline plasma serotonin concentration or average decreases in relative serotonin concentrations after either (b) four weeks or (c) eight weeks of SSRI therapy.

SUPPLEMENTARY TEXT

Study design and samples

Patients who met MDD diagnostic criteria without active suicidality, psychosis or mania and who had a score ≥ 14 on the 17-item Hamilton Depression Rating Scale (HAMD-17) were enrolled in the Mayo Clinic PGRN-AMPS SSRI trial.[36] Informed consent was obtained from all subjects prior to initiating the study. Response to SSRI therapy was assessed by serial determinations of the HAMD-17 and the Quick Inventory of Depressive Symptomology (QIDS)-C16. Treatment outcomes were defined as response (≥ 50% reduction in QIDS-C16 score from baseline), remission (QIDS-C16 score of ≤5) and percent change in QIDS-C16 scores. These were the same response criteria used in the STAR*D study.[9] Blood samples were collected at baseline and after four and eight weeks of SSRI treatment for SNP genotyping, plasma metabolomic analyses and blood drug and drug metabolite concentration measurements.[11, 37] DNA from PGRN-AMPS SSRI trial patients was genotyped at the RIKEN Center for Genomic Medicine (Yokohama, Japan) using Illumina Human610-Quad BeadChips (Illumina, San Diego, CA, USA), as described previously.[11] Of the patients for whom baseline, four and eight week plasma samples and clinical assessment data were available, 306 were chosen in a random order irrespective of clinical outcomes for metabolomic analyses.

Statistical analyses

Of the 306 subjects for whom metabolite concentrations were measured, six non-Caucasian and 10 non-compliant patients (based on blood drug assays) were removed from the analysis. A sample size of 300 patients provides 80% power to detect SNP associations with a quantitative

metabolite level that explains 12.4% of the variation in metabolite level at a genome-wide significant level of 5E-08.

Associations of baseline plasma metabolite concentrations and changes in metabolite concentrations with QIDS-C16 percent change were assessed using Spearman (partial) correlations. Metabolite concentrations and changes in metabolite concentrations after SSRI treatment were tested for association with dichotomous clinical outcomes (response and remission) using likelihood ratio tests based on logistic regression models.

GWAS for plasma serotonin concentrations and change in serotonin concentrations after four and eight weeks used linear regression to evaluate SNP genotype associations with Van der Waerden transformed baseline serotonin concentrations or changes in serotonin concentrations.

GWAS data for response in the PGRN-AMPS[11], STAR*D[24], and ISPC[10] studies were used to determine whether *TSPAN5* or *ERICH3* SNPs were associated with SSRI response.

Cell culture

All cells were cultured at 37°C with 5% CO_2. Human neuroblastoma SK-N-BE(2) (CRL-2271™) and human kidney HEK 293-T/17 (CRL-11268™) cells were purchased from the American Type Culture Collection (ATCC) and were cultured in Dulbecco's modified Eagle's medium (DMEM-F12; Invitrogen, Grand Island, NY USA) supplemented with 10% fetal bovine serum (FBS, Invitrogen). SK-N-SH cells were purchased from ATCC (HTB-11™) and cultured in Eagle's Minimum Essential Medium (EMEM; Invitrogen) supplemented with 10% FBS. Human glioblastoma U-251 MG were purchased from Sigma-Aldrich (09063001) and cultured in Dulbecco's Modified Eagle Medium (DMEM; Thermo Fisher Scientific) supplemented with

10% FBS. Human Neural Progenitor Cells (hNPCs) (ax0016) were purchased from Axol Bioscience (United Kingdom) and differentiated into neurons and grown per their recommendations using their media (ax0047), supplements (ax0045), and coating (ax0041) solutions. The "Human Variation Panel" of lymphoblastoid cell lines (LCLs) was obtained from the Coriell Institute (Camden, NJ, USA) and were cultured in Roswell Park Memorial Institute 1640 Medium (RPMI-1640, Invitrogen) supplemented with 15% FBS. The "Human Variation Panel" consists of 300 LCLs (100 European-American, 100 African-American and 100 Han Chinese-American subjects). We SNP genotyped and performed microarray analysis to determine basal levels of gene expression for all 300 cell lines, as described previously.[41] The Human Variation Panel model system has been utilized repeatedly both to generate pharmacogenomics hypotheses and to test pharmacogenomic hypotheses resulting from clinical GWAS.[40-44]

Electrophoretic mobility shift assays

Electrophoretic mobility shift assays (EMSAs) were performed using the LightShift Chemiluminescent EMSA Kit (PIERCE, Rockford, IL, USA). Nuclear extracts from SK-N-BE(2) cells were prepared with the Nuclear Extract Kit (Active Motif, Carlsbad, CA, USA). Oligonucleotides were designed based on NCBI DNA sequences (GRCh38 assembly; **Supplementary Table 6**) and were synthesized by Integrated DNA Technologies (IDT) (Coralville, IA, USA).

Quantitative real-time PCR (qRT-PCR)

Total RNA was isolated from cells using the RNeasy mini kit (QIAGEN). One hundred ng of total RNA was used to perform qRT-PCR. Gene expression analyses were performed using the ddCt method.[75] Values were compared using Student's *t*-test (two-sided).

Western blots

Total protein was extracted from cells using cell lysis buffer (Cell Signaling, Danvers, MA, USA or Reporter Lysis Buffer Kit, Promega). Protein quantification was performed using the BCA method (PIERCE). Specifically, protein samples were mixed with loading buffer and denatured at 95°C for 5 minutes, followed by separation on a 4-20% gradient SDS–PAGE (Bio-Rad, Hercules, CA, USA), and were transferred to polyvinylidene fluoride membranes that were blocked with 5% nonfat milk for 1 hour, followed by overnight incubation at 4°C with the primary antibody. The membranes were then washed three times with TBST for 10 minutes, followed by incubation for 1 hour at room temperature with appropriate horseradish peroxidase-conjugated secondary antibody (Cell Signaling) at a 1:10,000 dilution. Immunoreactive bands were visualized using chemiluminescence. ImageJ software or ChemiDoc™ Touch Image System was used to quantify protein band densities and compared by the student's *t*-test (two-sided). Data were represented as the mean of three or more independent experiments and the variation from the mean is represented as the standard error of the mean. Antibodies are listed in **Supplementary Table 7**.

Dual luciferase reporter gene assays

DNA sequences 200-300 bp in length were amplified from LCLs homozygous for wild type or variant *TSPAN5* SNP genotypes and cloned into the pGL3-promoter luciferase vector (Promega Corporation, Fitchburg, WI, USA). Plasmid genotypes were verified by sequencing. One microgram of each reporter gene construct and 40 ng of the pRL-TK renilla luciferase vector were cotransfected into SK-N-BE(2) cells. Dual-luciferase assays (Promega) were performed 24 hours after transfection, and two independent transfections were performed in triplicate for each reporter gene construct.

ERICH3 nonsynonymous variant protein levels

Wild type (WT) human *ERICH3* cDNA (OriGene Technologies, Rockville, MD, USA) was cloned into pCMV6-Entry plasmid (OriGene Technologies). Primers (**Supplementary Table 8**) were used to perform site-directed mutagenesis to create plasmids with the nsSNPs identified during the GWAS (rs11580409 or rs11210490) using the QuikChange Lightning kit (Agilent Technologies, Santa Clara, CA, USA). Sequences of each plasmid construct were confirmed by sequencing. Reporter gene constructs were cotransfected with pSV-ß-galactosidase (Promega) into HEK-293T/17 cells. *ERICH3* mRNA expression was validated by qRT-PCR. Total protein was extracted 48 hours after transfection with Reporter Lysis Buffer Kit (Promega) and stored at -80℃ until analysis by western blot. Transfection efficiency was determined by assay of β-galactosidase activity and ERICH3 protein levels were determined by Western blot analysis.

In an attempt to understand the SNP-dependent degradation of ERICH3, HEK293-T/17 cells were transfected with *ERICH3* cDNA constructs and, after 24 hours, were treated with the proteasome inhibitor MG132 (10 μM), the autophagy inhibitor 3-mehtyladenosine (3MA) (10 μM) or DMSO for 24 hours. Cells were lysed, transfection efficiency determined by β-galactosidase activity and ERICH3 protein levels were determined by Western blot analysis. Each of these experiments were performed more than three independent times.

Knockdown and over expression experiments

For TSPAN5 knockdown (KD) studies, SK-N-BE(2) cells were transiently transfected with 200 nM *TSPAN5* shRNA (Dharmacon) using the Lonza nucleofection kit V (Lonza), were transduced with lentiviral particles containing *TSPAN5* siRNA (GenTarget, San Diego, CA, USA) or with RBPJ-κ siRNA. For TSPAN5 over expression (OE) studies, SK-N-BE(2) cells were transiently

transfected with a cDNA clone (Origene) containing the *TSPAN5* gene using Lipofectamine 2000. The cells were cultured for 72 hours after KD or OE before analyzing mRNA or protein. Human Neural Progenitor Cell (hNPC)-derived neurons were used for some of the ERICH3 KD experiments because they highly express ERICH3. hNPC-derived neurons were transduced with lentiviral particles containing *ERICH3* siRNA. ERICH3 was overexpressed in SK-N-BE(2) neuroblastoma cells because they do not express ERICH3. The ERICH3 expression construct was transfected by electroporation using Lonza Solution V kit (Lonza). With the exception of the RBPJ-κ KD studies, each was repeated two or more times. KD and OE efficiency and gene expression were determined by qRT-PCR. All KD and OE experiments were performed in at least two independent experiments.

Cell culture serotonin concentration

Cell culture media was collected after TSPAN5 KD or OE and ERICH3 KD or OE. Cell extracts were analyzed by qRT-PCR to ensure KD and OE. Samples of the media were sent to Bioanalytical Systems, Inc. (BASi) for the assay of serotonin concentrations once per sample using HPLC MS/MS. This method was developed by BASi and used a Restek PFP column for separation and API4000 for detection.