Exploring gene expression biomarker candidates for neurobehavioral impairment from total sleep deprivation

Hilary A. Uyhelji, Doris M. Kupfer, Vicky L. White, Melinda L. Jackson, Hans P. A. Van Dongen, Dennis M. Burian

Supplementary Methods, Results, and Figures

Methods

Sample collection and neurobehavioral data

Blood was collected in a PAXgene^{TM} Blood RNA tube. Of 204 possible samples (12 timepoints, 17 subjects), five blood draws were unsuccessful, resulting in 199 RNA extractions **(Additional file 1: Table S1)**.

As previously described [1], subjects were presented with the PVT for 10 min at 2–4 h intervals during scheduled wakefulness. For the present report only the tests at 4 h intervals corresponding to the timing of the blood draws were used (**Fig. 1**). On average, blood draws and PVT data were collected at approximately 8:00 h, 12:00 h, 16:00 h, and 20:00 h. Subjects were asked to immediately press a button upon observing a visual stimulus on a computer screen, which appeared at random 2–10 s intervals. Stimulus-response times >500 ms were recorded as lapses, and data were recorded as the number of lapses per test bout. Significant differences in PVT lapses over time and between conditions were tested with generalized linear models using Poisson distribution and log link with PVT as the response variable, and progressively adding as predictors Treatment (TSD or C), Phase (Baseline, Experimental, Recovery),

Time of Day, and their interaction terms. In all models a random intercept was included for subject; also, a random term for observation was included to correct for overdispersion as in [2]. Models were constructed in R v. 3.2.1 using the glmer function in the lme4 v. 1.1-11 package [3], with the “nlminb” optimizer.

RNA isolation and microarray data collection

Blood samples in PAXgene™ tubes were shipped to CAMI and thawed. Total RNA was isolated with a QIAcube robotic workstation using reagents from the PAXgene™ Blood miRNA Kit and the manufacturer’s instructions. Following isolation, aliquots were used for quantification on a NanoDrop 1000 Spectrophotometer (ThermoFisher), and quality assessment including calculation of the RNA Integrity Number on an Agilent Technologies 2100 Bioanalyzer®. Total RNA samples were reverse transcribed and amplified with the NuGEN Inc. Ovation® Pico WTA System V2 kit. Concentration of the amplified cDNA was determined with the NanoDrop as above. Fragmentation and biotin labeling was performed using the Encore® Biotin module (NuGEN Inc.), followed by hybridization to GeneChip Human Gene 1.0 ST array 169 format chips and scanning on an Affymetrix Scanner, according to the manufacturer’s protocol. Randomization was applied to RNA extractions and microarray preparations.

Differential gene expression: Treatment and PVT effect

Transcript Cluster-level expression values were derived by background correction, quantile normalization, and median polish summarization with the RMA algorithm [4] in the R/oligo package v. 1.32.0 [5]. Data were filtered for low expression,

only retaining a Transcript Cluster if at least 6 samples had $\log_2$ expression >6, corresponding to the median (and mean) expression of the antigenomic Transcript Clusters. Raw Affymetrix expression data, and results of the RMA algorithm post-filtering, are archived at the NCBI GEO online repository accession GSE98582. Array data quality was assessed with hierarchical clustering and tools in R/arrayQualityMetrics v. 3.24.0 [6]. Six arrays from six different TSD subjects were considered outliers: one at noon and one at 20:00 h on day four (Experimental), one at noon and one at 20:00 h on day six (Recovery), and two at 16:00 h on day six. In tests of differential gene expression, outliers were included but de-emphasized using the internal array weighting function of limma [7]. For other analyses without an internal weighting function (WGCNA, Mfuzz, and RIF; see below), a separate RMA expression file was created omitting these six outliers altogether.

Visual examination of the PVT data suggested that three of the 11 TSD subjects were fatigue resistant in terms of neurobehavioral (vigilance) performance (**Fig. S1**). These were subjects s6093, s6125, and s6311 (**Additional file 1: Table S1**). It was hypothesized that the inter-individual variability could impede detection of a gene expression Treatment effect between C and TSD persons, as the fatigue resistant subjects are outliers to the average TSD response, at least in terms of PVT lapses. Thus, to identify genes differentially expressed respective to Treatment, the three fatigue resistant subjects were excluded from the RMA expression file for Treatment effect analysis, and limma analyses (see below) were run to test for a difference between the 6 C and the remaining 8 TSD persons. A comparison was made by running this model of a Treatment effect, with one including the fatigue resistant individuals.

About 90% of genes different between TSD and C subjects at the Experimental phase in models with the fatigue resistant subjects also were significant in models without these subjects at FDR <0.05. However, models without these three subjects detected over three times as many differentially expressed Transcript Clusters. Some of these additional genes such as *Interleukin 1B* (*IL1B*) were known from literature review to respond to sleep levels [8, 9]. Hence for biomarker discovery, models without the fatigue resistant subjects were considered more comprehensive and potentially more accurate indicators of a Treatment response. In future studies if Control subjects are used that have previously been profiled for their fatigue resistance, fatigue susceptibility could be incorporated into the Treatment model as an additional variable rather than eliminating the resistant subjects. Alternatively, a cross-over design could be explored with each subject undergoing two study runs, one with Total Sleep Deprivation and one with the well-rested Control condition.

Tests for differential expression were conducted for biomarker discovery with R/limma v. 3.24.15 using linear models [7]. All models included array weights, incorporation of the inter-subject correlation (using the duplicateCorrelation and block functions in limma), and a term to account for differences attributable to Time of Day. The Treatment list was defined as Transcript Clusters with a significant difference between C and TSD persons at the Experimental or Recovery phase (FDR <0.05), but not at Baseline. Plots of Transcript Clusters were reviewed to ensure mean $\log_2$ expression ±1 SE overlapped between C and TSD subjects at all four Baseline timepoints, when no differences were expected. If plots showed separation in

expression between C and TSD at Baseline, the Transcript Cluster was discarded from the final Treatment list to reduce false positive discoveries.

Similarly, linear models were applied in limma to test for a significant relation of PVT lapses to gene expression. Here, data for all 11 TSD and 6 C subjects were included because PVT lapses inherently provided a means of incorporating inter-individual variability in the response to sleep loss into the statistical models. Besides array weights, the inter-subject correlation, and PVT lapses, the model contained terms for Treatment (TSD or C) and Time of Day. No Transcript Clusters were associated with PVT lapses at a threshold of FDR <0.05. Because exploring biomarkers for neurobehavioral impairment from sleep deprivation was a primary aim, it was important to avoid missing candidate biomarkers in this first screening study. Hence the threshold for significance was relaxed from FDR <0.05 to FDR <0.10. As before, expression was plotted for TSD and C subjects, and Transcript Clusters were eliminated from the PVT list if the mean $\log_2$ expression ±1 SE showed separation at any Baseline timepoint.

Models of the PVT effect were re-run with inclusion of additional factors, but these did not appear to improve results. Use of age in the model led to detection of a single Transcript Cluster at FDR<0.1, which was considered a potential false positive due to separation of mean ±1 SE $\log_2$ expression for C and TSD subjects at Baseline. Use of body mass index or BMI as calculated with the online tool at https://www.nhlbi.nih.gov/health/educational/lose_wt/BMI/bmi-m.htm [10] added seven new Transcript Clusters, but all were eliminated based on separation in expression at Baseline, and most were annotated as ribosomal RNA. While re-running the model accounting for gender added detection of an additional 19 Transcript Clusters (inclusive

of the 7 found in the model with BMI), 13 of them were deemed potential false positives due to separation of expression between C and TSD at baseline, and by that logic the false positive rate was increased by 11% from the original model without BMI, age, or gender. Hence the results reported here (for both PVT and Treatment) reflect the simpler design excluding factors of age, gender, and BMI. Future work is needed to more thoroughly evaluate the impact of such factors on biomarker detection, using a larger and more diverse population.

Fifteen Transcript Clusters significantly related to PVT lapses were not found in the Treatment list (**Additional file 5: Table S4**). To verify that this result was not simply a reflection of the Treatment data set excluding the fatigue resistant individuals while the PVT analysis included them, the Treatment analysis was run again with all 17 subjects. During the Experimental phase, the lowest FDR for the difference between TSD and C subjects in these 15 Transcript Clusters was 0.35, suggesting that the finding was robust to inclusion or exclusion of the fatigue resistant subjects. Sample plots were made of a Transcript Cluster significantly related to PVT lapses and Treatment, as well as a Transcript Cluster significantly related to PVT but not Treatment using R package ggplot2 v. 2.2.1 [11]. Data from outlier arrays (see above) was omitted in creating these plots.

Co-expression and temporal networks

Weighted Gene Co-expression Network Analysis was performed using R/WGCNA v. 1.47 and data from all 17 subjects. This approach has been described in detail by Langfelder and Horvath [12]. Essentially the analysis serves to group genes based on similarity of expression across samples (n=193; 199 successful blood draws

minus 6 outliers). All Transcript Clusters passing the low-expression filter were included, not just those differentially expressed, for a complementary approach to the limma models above. Signed networks were constructed, allowing for positive or inverse relationships based on Pearson correlations among genes. Using internal functions plotting scale free topology, a soft power threshold of 13 was selected, and the minimum number of Transcript Clusters per group was set to 30. Each group of co-expressed Transcript Clusters was termed a module, which the package designated by a color. Pearson correlations were computed separately between each module's eigengene and three variables: Treatment, PVT lapses, and Time of Day. The eigengene was a representative metric of the expression profile of Transcript Clusters in the module [12]. The WGCNA software also allowed identification of the top hub Transcript Cluster for each module, namely, the most highly connected Transcript Cluster within the module. Here, connectivity was based on the correlation of expression among genes [12].

A second temporal clustering approach was performed using R/Mfuzz v. 2.28.0 [13]. This analysis was not designed to test for treatment differences, but rather focuses on portraying time series. Hence the three fatigue resistant individuals and the Controls were omitted, and Mfuzz was used to cluster and plot the mean $\log_2$ expression of the 8 TSD subjects across the 12 timepoints. Expression values for the differentially expressed Treatment effect Transcript Clusters were z-transformed and then clustered across timepoints using fuzzy c-means clustering. An internal function was used to select the fuzzification parameter (m=1.33). A combination of internal functions and plotting was used to decide on the number of clusters, which was set to three. The

same clustering strategy was employed on the Transcript Clusters associated with PVT lapses, based on the average data from the 8 TSD subjects, to show temporal changes in expression of these Transcript Clusters during TSD. The fuzzification parameter was m=1.76, again with creation of three clusters. Plotting was re-run using the same approach, on the Control subjects' data. Fuzzification parameters were the same (m=1.33 for the Treatment list run, and m=1.76 for PVT), and the number of clusters was set to three.

Transcription factor regulators

Regulatory Impact Factor (RIF) scores were assigned to known human transcription factors as a means of ranking potential regulators of the differentially expressed genes in a differential co-expression analysis [14]. In synopsis, transcription factors were separately correlated to all differentially expressed genes in each of two conditions (fatigue resistant subjects were omitted). Then the difference between the correlation of the transcription factor to gene expression in the first condition, and its correlation to gene expression in the second condition, was computed and squared. The result was weighted by the average abundance of the differentially expressed gene across all samples, and its difference in expression between the two conditions. The final value was converted to a z-score and reported for each transcription factor to rank their regulatory potential. Based on the differences in correlation and in expression between the two conditions (here, C and TSD), the RIF z-scores were either positive or negative. Larger absolute values of the z-scores were interpreted as stronger evidence for a regulatory role.

Three input sets of data were necessary for running this analysis: (1) a list of transcription factors, (2) a list of differentially expressed genes (Transcript Clusters), and (3) the expression values of the transcription factors and differentially expressed genes in each condition. The list of transcription factors (1) consisted of all human transcription factors in QIAGEN BIOBASE TRANSFAC v. 2015.4 for which there was data in the present study, after filtering for low expression levels (see above). For the second item two lists of differentially expressed genes were used running the RIF analysis twice. In one run the Treatment effect Transcript Clusters were used (2a), and in the other run the PVT effect Transcript Clusters were employed (2b) as the differentially expressed gene list. Finally, averages were computed for each of the 12 timepoints for C and for TSD individuals, and these 12 values served as the input expression data (3) for each of the two conditions (C, TSD). (It was necessary computationally to have the same number of expression data points for both conditions, making it impossible to use the raw data for the 6 C vs. 8 TSD subjects.) The script published in the supplemental material of Uyhelji et al. [15] was run to compute RIF z-scores, with minor modifications including use of Pearson rather than Spearman correlations.

In addition to the RIF differential co-expression analysis, the BIOBASE F-match tool at http://www.biobase-international.com [16] was utilized to search for regulatory transcription factors based on the promoter sequence of differentially expressed genes. Analysis was based on the BIOBASE TRANSFAC® v. 2015.4 vertebrate non-redundant profile [17], with default settings modified to minSum for optimization of both false positive and false negative errors, and a P-value threshold of 0.05.

The BIOBASE tool was applied to scan for over-represented binding sites, first in Treatment effect Transcript Clusters, and then in PVT effect Transcript Clusters. For the Treatment effect, a background set was randomly selected from Transcript Clusters with raw P-values >0.10 (FDR>0.337) in the limma differential expression analysis between C and TSD subjects at the Experimental phase. For the PVT effect, the background set was randomly selected from Transcript Clusters with raw P-values >0.10 (FDR>0.513) for PVT lapses. Each background list contained the same number of Transcript Clusters as the foreground list (Treatment or PVT effect Transcript Clusters), per the manufacturer's recommendations. Ten background lists were created for each foreground set, and only transcription factors appearing on at least nine iterations of F-match were considered for further analysis.

Functional enrichment and pathway analysis

Affymetrix's online tool NetAffx™ [18] was used to annotate gene lists, with emphasis on the first annotation provided for genes with mixed hybridization targets per Affymetrix's recommendation (pers. comm.). Also, the DAVID v. 6.7 bioinformatics tool [19] was used to characterize functional enrichment. For DAVID, Transcript Clusters in the list of interest served as the foreground input, with all Transcript Clusters in the corresponding RMA expression set passing the low-expression filter as the background. As suggested by [20], such user-defined backgrounds can be important to reduce bias toward tissue-specific expression (e.g., detection of pathways found in blood due to using blood as the RNA source). Analysis focused on DAVID functional clusters with enrichment scores >1.3, corresponding to P <0.05.

The Ingenuity Pathway Analysis® (IPA®, QIAGEN Redwood City, http://www.qiagen.com/ingenuity) Core Analysis tool was used to explore molecular pathways and networks based on previously published interactions among genes. Parameters used included consideration of both direct and indirect relationships, exclusion of endogenous chemicals, inclusion of Causal Network analysis, use of Ingenuity expert and Ingenuity supported third party information from experimentally observed data (vs. predictions), and restriction of species to mammals (human, mouse, rat). The foreground in these IPA® runs consisted of the gene list of interest (e.g., Treatment or PVT effect Transcript Clusters, converted internally by IPA® to genes), and as with DAVID, the background was all microarray genes passing the $\log_2 > 6$ low-expression threshold in at least six samples.

Because PVT lapses are ordinal rather than binary, fold changes for the PVT effect could not be computed directly for use in IPA®. Instead, the $\log_2$ fold change values were used for TSD relative to C subjects (omitting the three fatigue resistant persons) at the Experimental phase, taken from the Treatment effect analysis. Based on the sign of the correlation coefficient of a Pearson correlation between PVT lapses and gene expression, it was confirmed that the direction of expression indicated by the Treatment effect fold change reflected the direction relative to PVT. That is, Transcript Clusters with a positive Pearson correlation between PVT and gene expression (higher expression with more PVT lapses) also showed a positive fold change for the Treatment effect (higher expression in TSD than C subjects). This information was used by IPA® to depict whether a given gene was up-regulated or down-regulated, and whether connected molecules and biological functions were activated or inhibited. Causal

Network pathways of interest were reviewed for connections between differentially expressed genes and upstream regulatory molecules [21]. Fisher's exact tests were used to check for the presence of a greater number of differentially expressed genes related to the Causal Network than would be expected due to chance.

microRNA quantitative PCR

A 300 ng aliquot of total RNA from each blood sample was used for analysis of microRNA (miRNA) expression with 30 TaqMan® assays (Life Technologies, Grand Island NY) according to the recommendations from Fluidigm® for miRNA analysis on the BioMark system. The TaqMan microRNA Reverse Transcription kit (Life Technologies) and KAPA Probe Fast qPCR Kit Master Mix Universal (KAPA Biosystems) were used with custom reverse transcription and preamplification primer pool. Following a 17 cycle preampflication, the product was diluted 1:10 with 10 mM Tris pH 8.0, 0.1 mM EDTA. The qPCR amplification reaction was prepared according to the manufacturer's 96x96 protocol using KAPA Probe Fast qPCR Kit Master Mix Universal (KAPA Biosystems) on Biomark IFC Controller and HD instruments (Fluidigm Corp.).

Initial analysis of the miRNA plate assays was done with settings as recommended (Fluidgm Real-Time PCR Analysis software, ver. 4.1.2): 0.65 quality threshold, Linear (Derivative) baseline correction, and Auto (Global) Ct threshold method. Efficiencies for each gene were determined via duplicate standard curves run on the same plate. The miRNAs *microRNA let-7a-1 (MIRLET7A1)* and *microRNA let-7d (MIRLET7D)* were chosen as Normalizers using GeNorm and Normfinder. Finally, normalized gene expression values were analyzed via linear mixed-effects models with

R package nlme v. 3.1-126, using marginal Type III sum of squares. Models predicted expression of each miRNA as a function of PVT lapses, Treatment (TSD or C), and Time of Day (encoded as factor), including a random intercept for subject. Plots of expression of miRNAs significantly related to PVT were made using R packages ggplot2 v. 2.2.1 [11] and gridExtra v. 2.3 (https://cran.r-project.org/web/packages/gridExtra/index.html).

Results

Comparison of gene lists with published datasets

Four publications with lists of differentially expressed genes responding to total sleep deprivation, as identified from human blood samples, were compared to the Treatment and PVT gene lists in the current study (**Additional file 4: Table S3**). For genes with multiple possible annotations, only the first listed gene symbol in the current study's supplemental Treatment and PVT datasets was used (**Additional file 3: Table S2, Additional file 5: Table S4**), and compared for a perfect match with the gene symbol or identifier reported in literature. In a study by Pellegrino and colleagues [22], a supplemental data file was provided of genes with a difference between both baseline vs. sleep deprivation, as well as during sleep deprivation vs. recovery. None of these genes were found in the Treatment or PVT lists reported here. Likewise, neither of the two genes related to duration of sleep deprivation in a study by Arnadottir et al. [23] were found in the Treatment or PVT lists. However when analysed for a difference among states (phases) of normal sleep, sleep deprivation, and recovery, Arnadottir et al. reported 48 differentially expressed genes in a supplemental file [23], of which one

(*CAMP*) was also found in the Treatment list. There was no overlap with the PVT list. There was only one Treatment gene (*ABCA1*) but no PVT list members that overlapped with the condition of control sleep followed by total sleep deprivation in Möller-Levet et al. 2013 [24]. Finally, of the 1855 genes with rhythmic expression in the control condition in [24], 62 were found in the Treatment list, and another six in the PVT list. Of the six genes overlapping with PVT, three genes were found in both Treatment and PVT analyses (*LITAF, IPMK,* and *AQP9*), and the remainder were unique to the PVT list (*LPCAT2, KCNJ15,* and *MSL1*).

There were likely several reasons for the extremely low congruence between the results in the current study and the genes responding to sleep deprivation in these published datasets, not the least of which would be differences in study design and analytical models. Besides differences in the length of sleep deprivation, microarray platforms varied: Pellegrino et al. used a distinct Affymetrix array from the current study [22], Arnadottir et al. [23] used custom Affymetrix arrays, and Möller-Levet et al. used Agilent arrays [24]. Additionally, some genes possess multiple annotations (due to uncertainty of the best match of a Transcript Cluster or probeset to a gene, or the existence of multiple names for a single gene), which may have led to underestimation in the overlap of gene lists between studies. Moreover, in some cases distinct members of the same gene family were identified. For example, Arnadottir et al. [23] reported *Dedicator of Cytokinesis 3* (*DOCK3)* among their 48 state-responsive genes, whereas the present study identified *Dedicator of Cytokinesis 5 (DOCK5)* in the Treatment list. Future analyses of the candidate Treatment and PVT biomarkers in the present study should include comparison to additional published datasets, but a complete review of

the literature was beyond the scope of this study. Also, it would be interesting to compare genes affected by sleep restriction in these and other studies.

Supplemental Figures

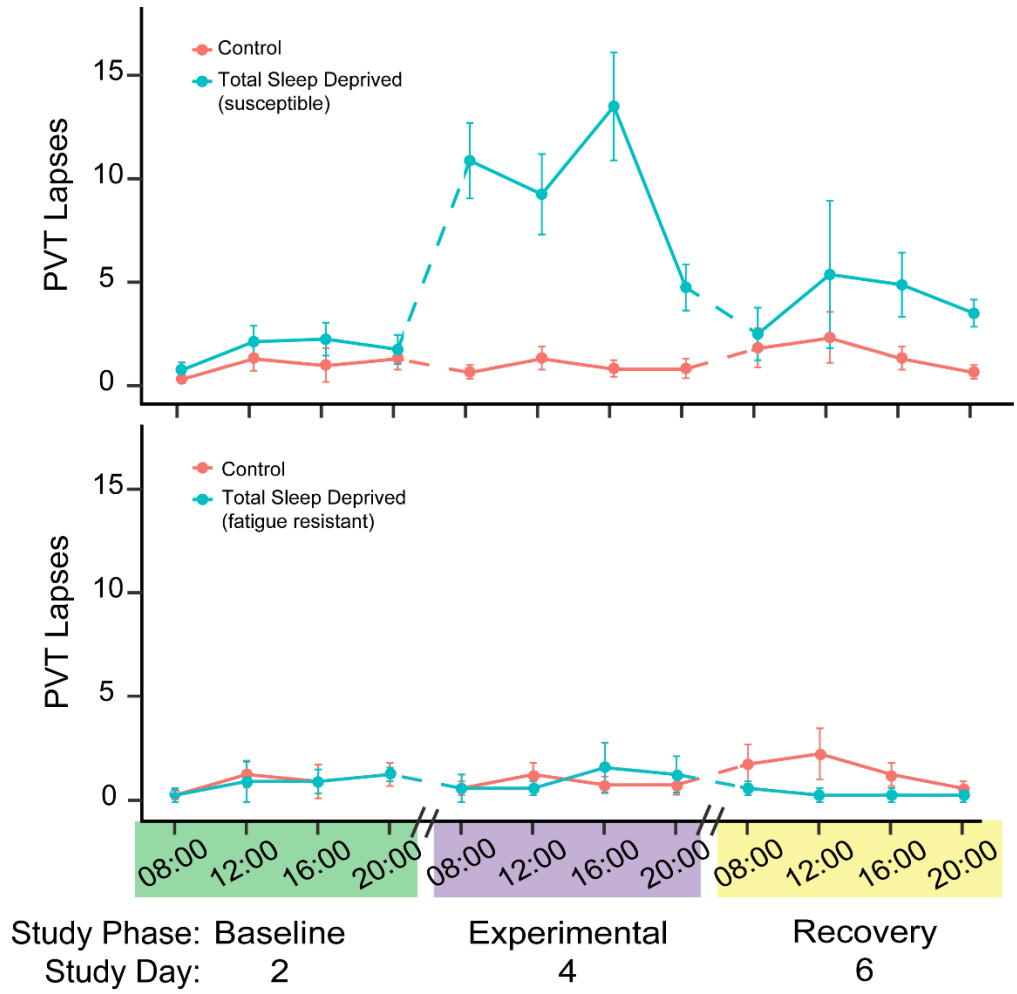


Figure S1

Mean (± 1 SE) Psychomotor Vigilance Test (PVT) lapses for Control (C) vs. susceptible and fatigue resistant Total Sleep Deprivation (TSD) subjects. One chart (top) compares the 6 C individuals to the 8 TSD subjects susceptible to the TSD treatment, whereas the other (bottom) compares the 6 C individuals to the 3 TSD subjects identified as fatigue resistant.

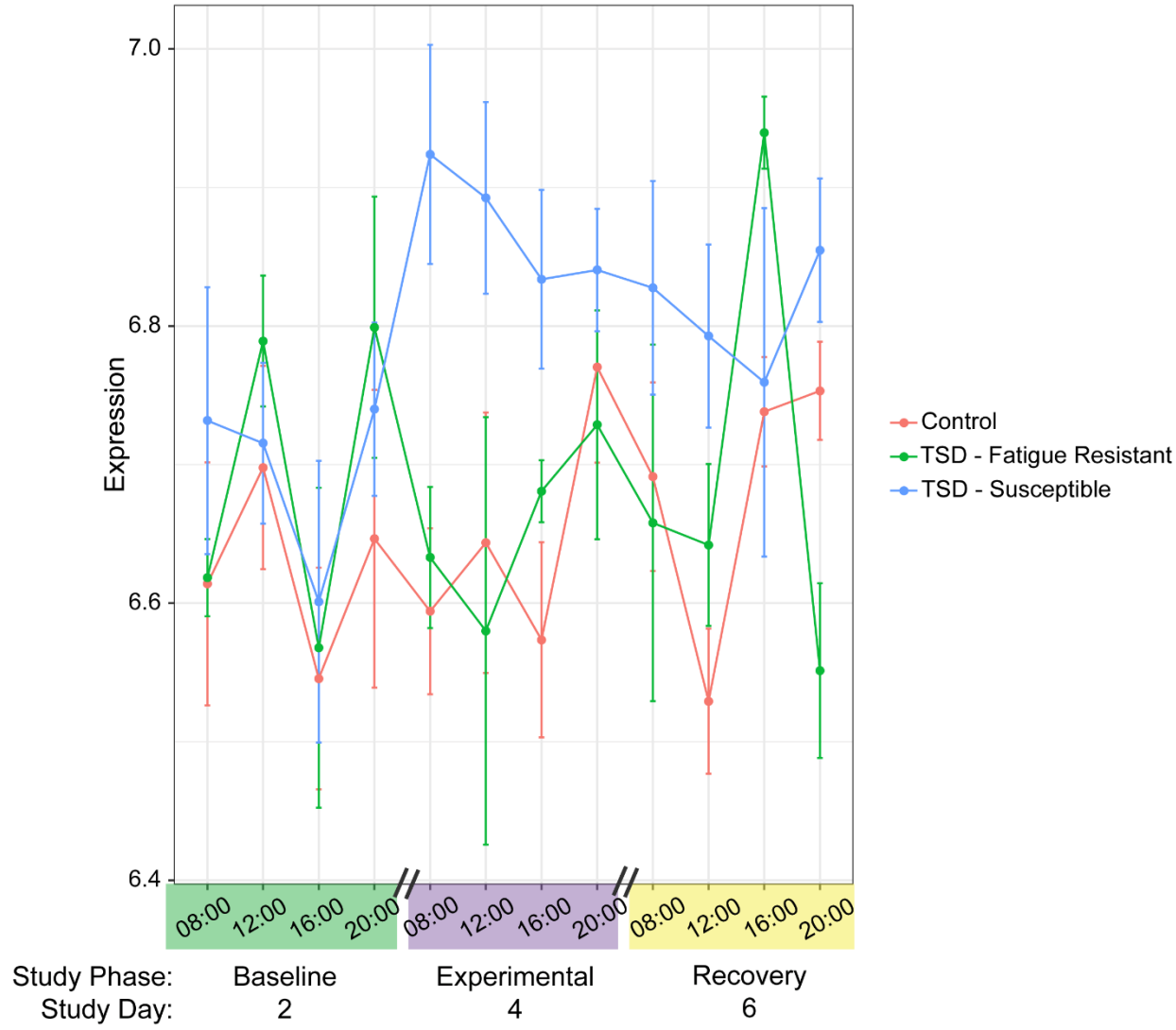


Figure S2

Mean (± 1 SE) $\log_2$ expression of Transcript Cluster 8133209, a *SPDY* gene family member. This Transcript Cluster is shown as a sample Transcript Cluster present in both Treatment and PVT lists.

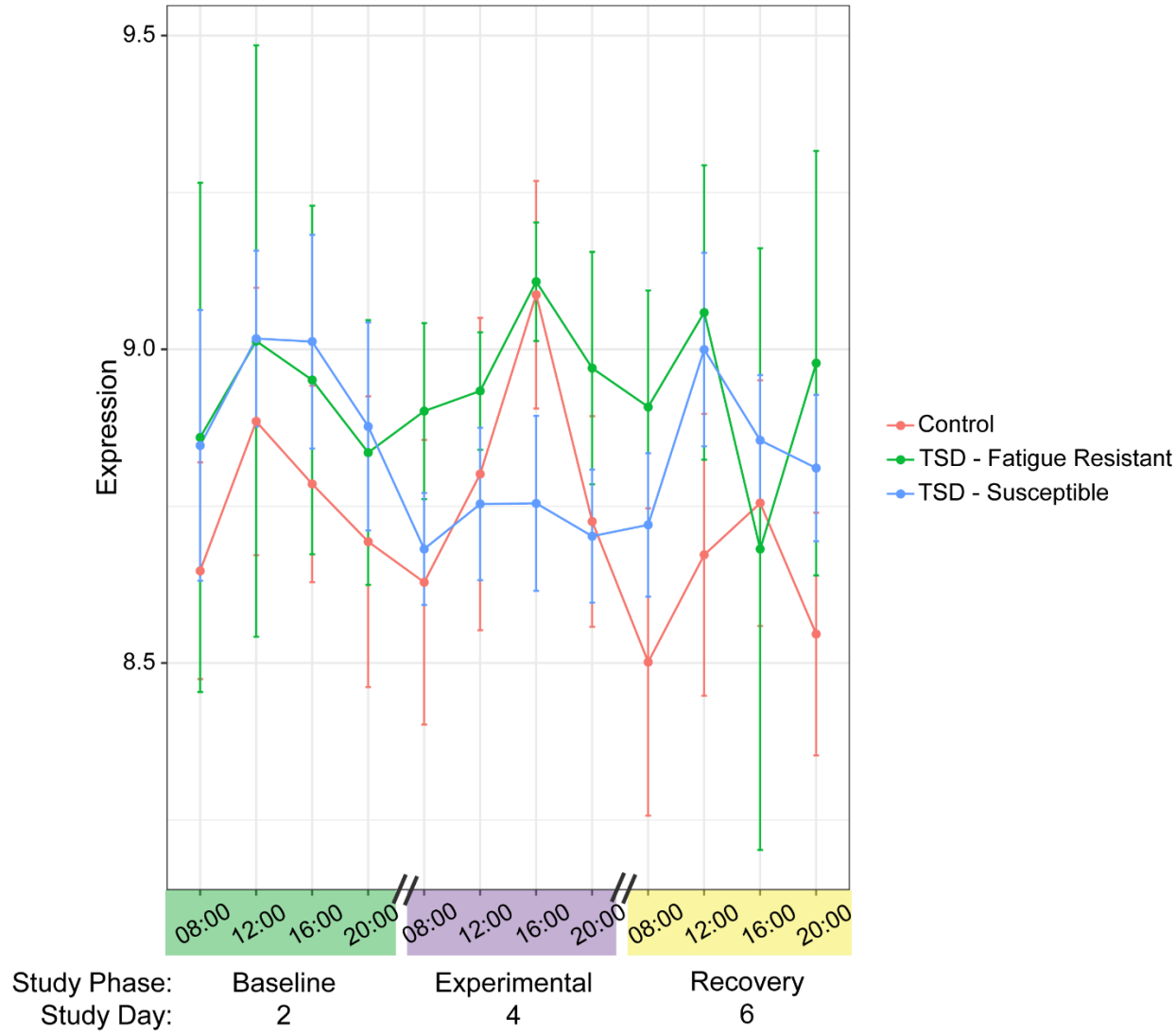


Figure S3

Mean (± 1 SE) $\log_2$ expression of Transcript Cluster 8068583, representing gene *KCNJ15*, as measured on microarrays. This Transcript Cluster is shown as a sample Transcript Cluster present in the PVT list but not the Treatment list.

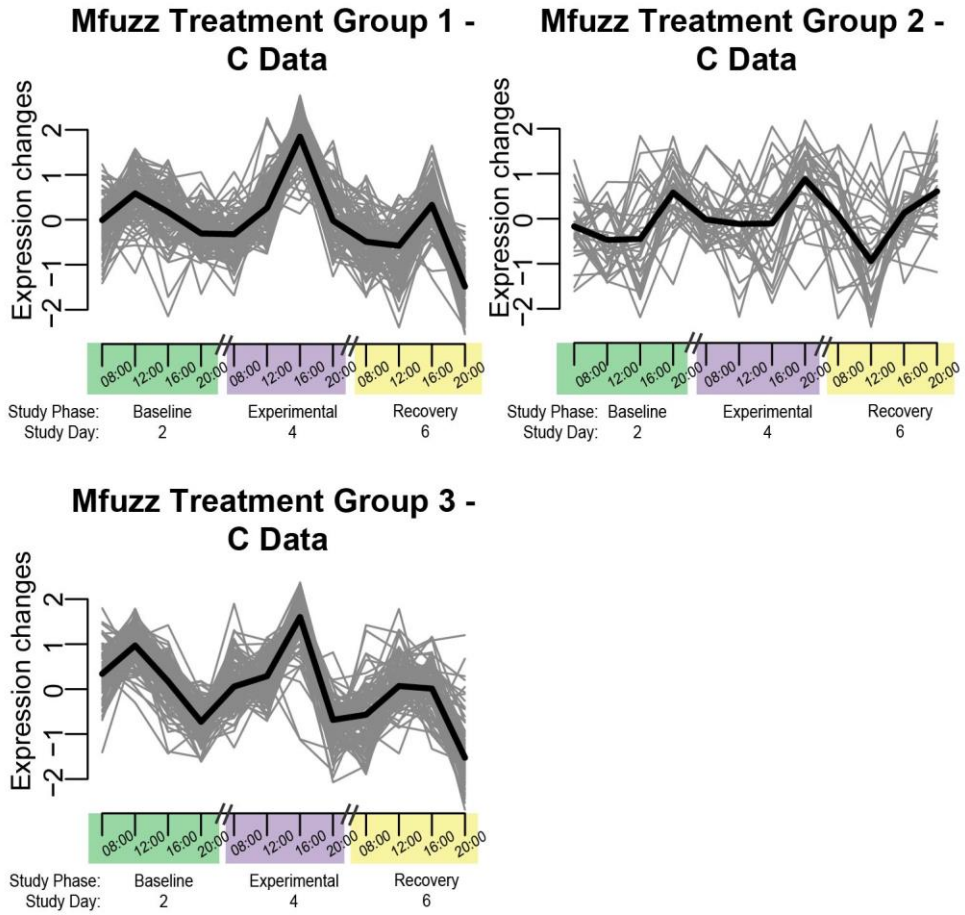


Figure S4

Temporal expression of microarray Transcript Clusters that are significant for the Treatment effect. Gray lines represent individual Transcript Clusters, and the thick black line represents the group center. Values represent normalized expression based on data from Control (C) subjects during three of the seven consecutive study days, consisting of one day each during the Baseline, Experimental, and Recovery phases.

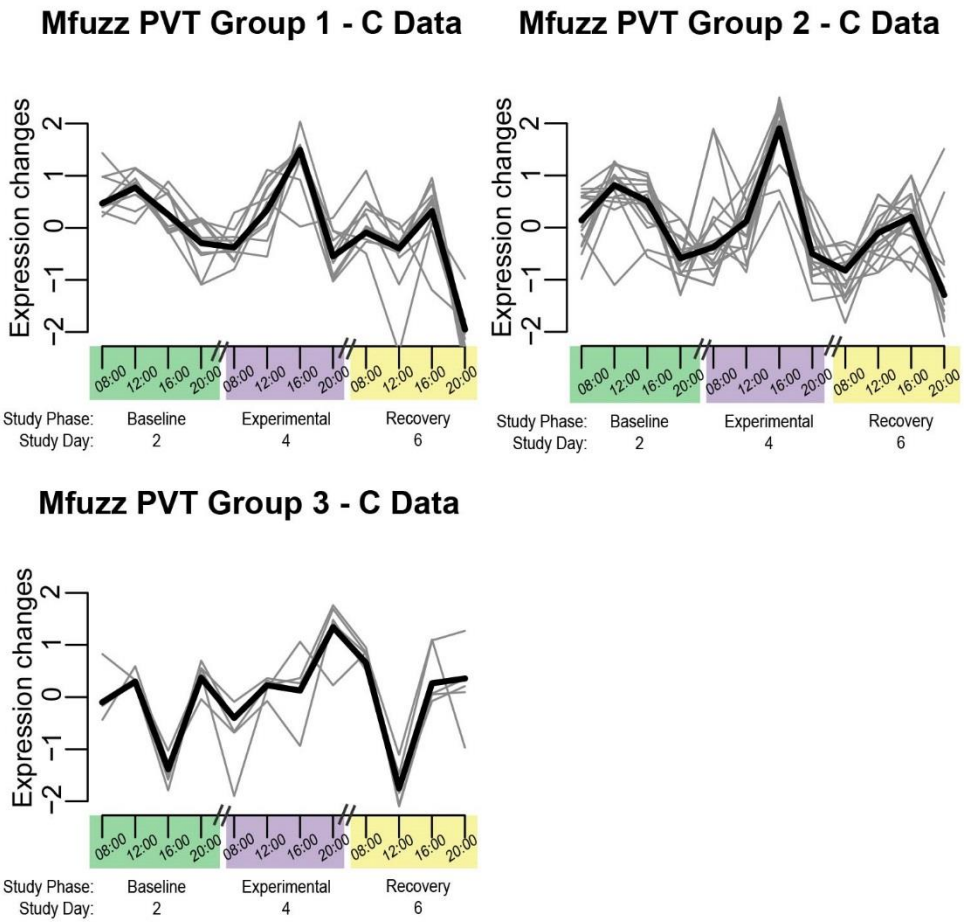
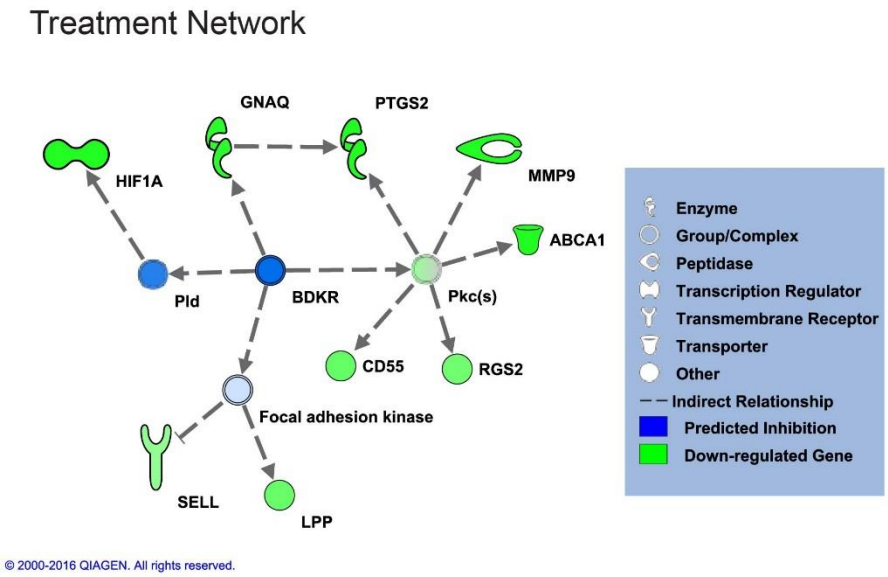


Figure S5

Temporal expression of microarray Transcript Clusters significant for the Psychomotor Vigilance Test effect. Gray lines represent individual Transcript Clusters, and the thick black line represents the group center. Values represent normalized expression based on data from Control (C) subjects during three of the seven consecutive study days, consisting of one day each during the Baseline, Experimental, and Recovery phases.



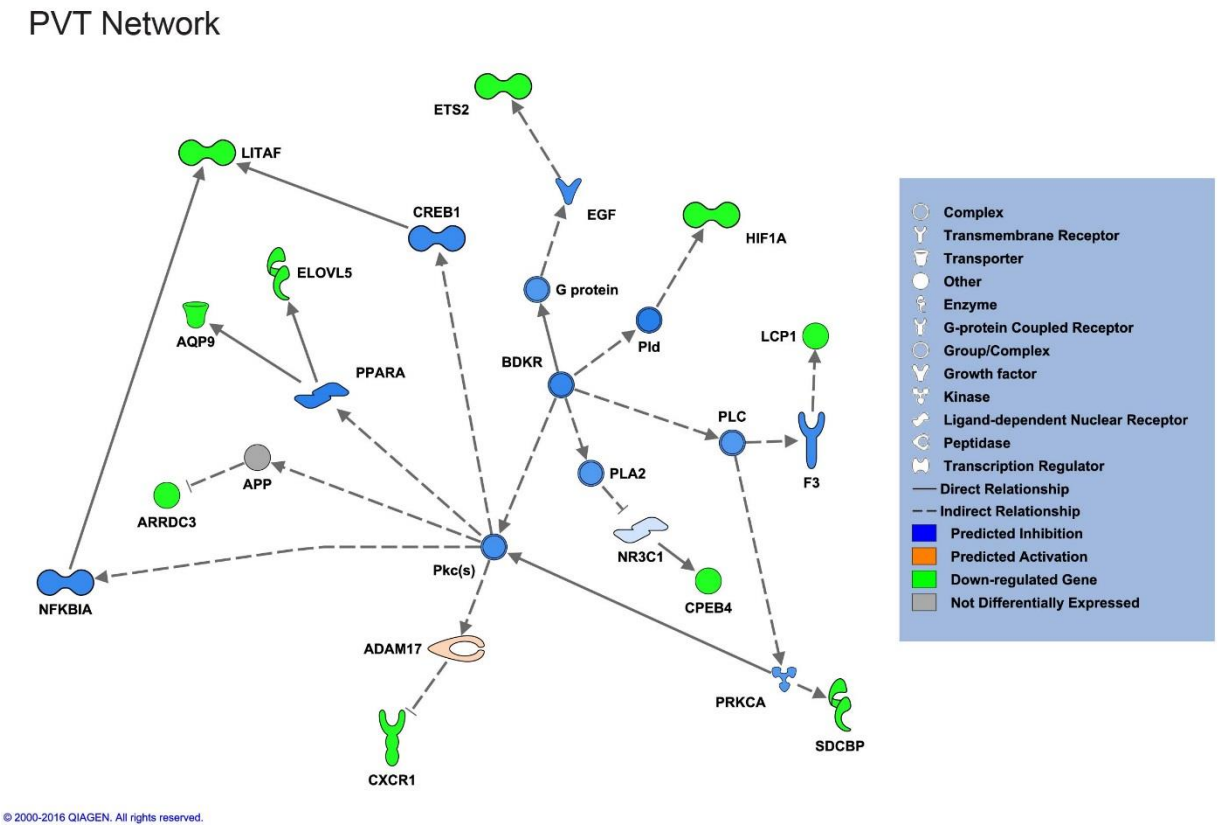
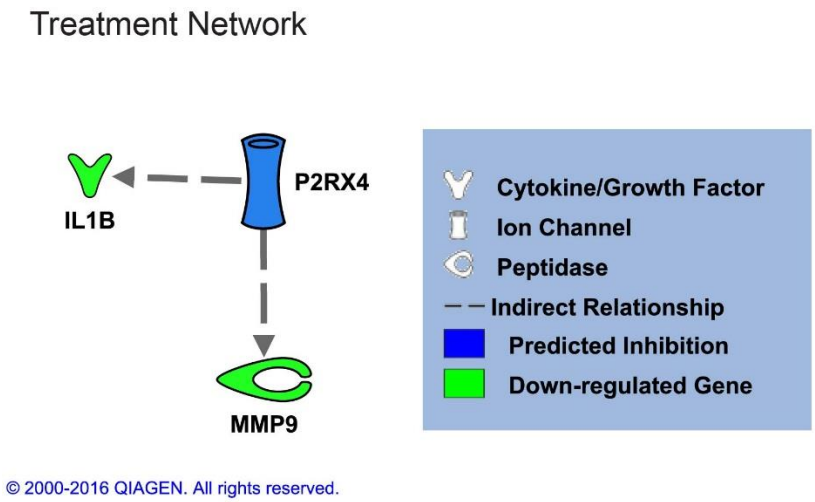


Figure S6

Ingenuity Pathway Analysis® Causal Network pathway B with master regulator *BDKR* for Treatment effect genes (top), and PVT effect genes (bottom). Lighter molecule color

(e.g., *NR3C1*, *ADAM17*) suggests less confidence in prediction of the direction of expression; darker indicates more confidence. In the PVT network, no prediction was made for *APP*, which was tested for differential expression but not significant. In the Treatment network, PKC complex prediction of light green reflects down-regulation of complex member *PRKCB*, but no differential expression for complex genes *PRKCA, PRKCD,* or *PRKCQ*.



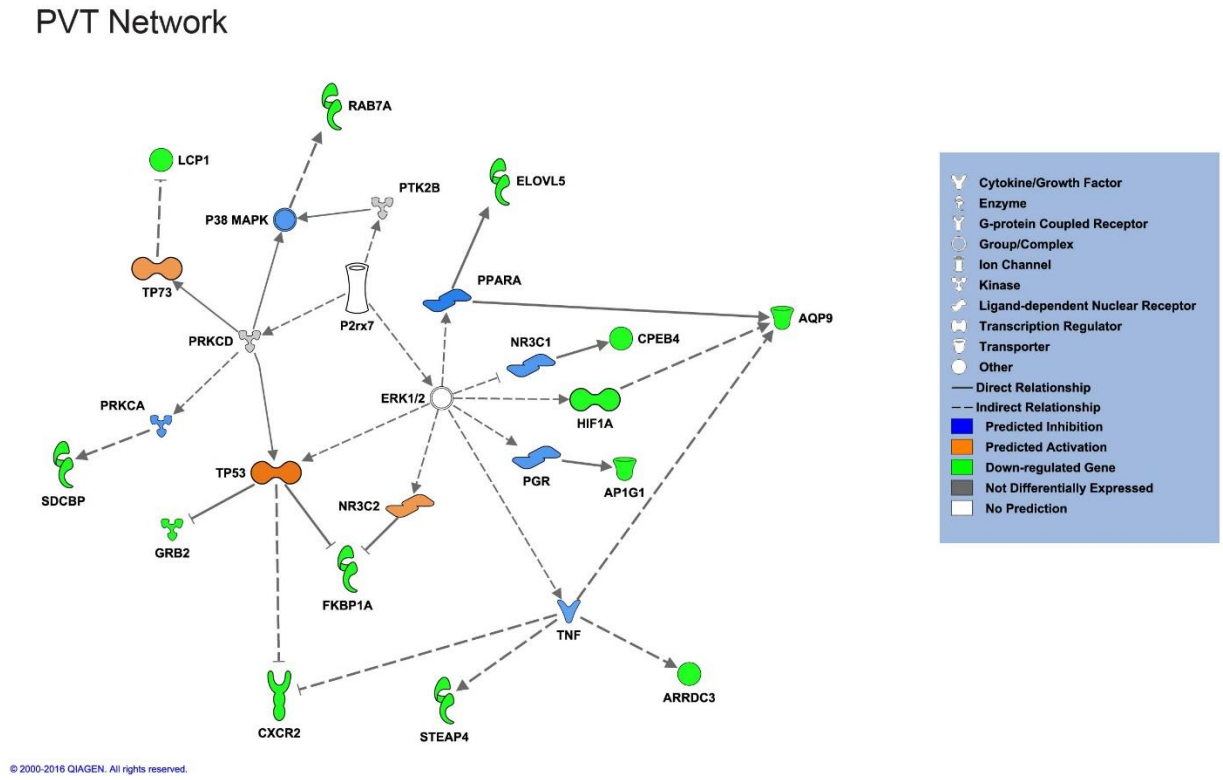


Figure S7

Ingenuity Pathway Analysis® Causal Network pathway P with purinergic receptor master regulator for Treatment (top), and Psychomotor Vigilance Test (PVT, bottom).

Lighter molecule color suggests less confidence in prediction of the direction of expression; darker indicates more confidence. In the PVT network, the software was unable to make predictions of direction of expression for *P2RX7, ERK1/2, PTK2B,* or *PRKCD*. Of these, *PTK2B* and *PRKCD* passed the low-expression thresholds but were not significantly related to PVT lapses.

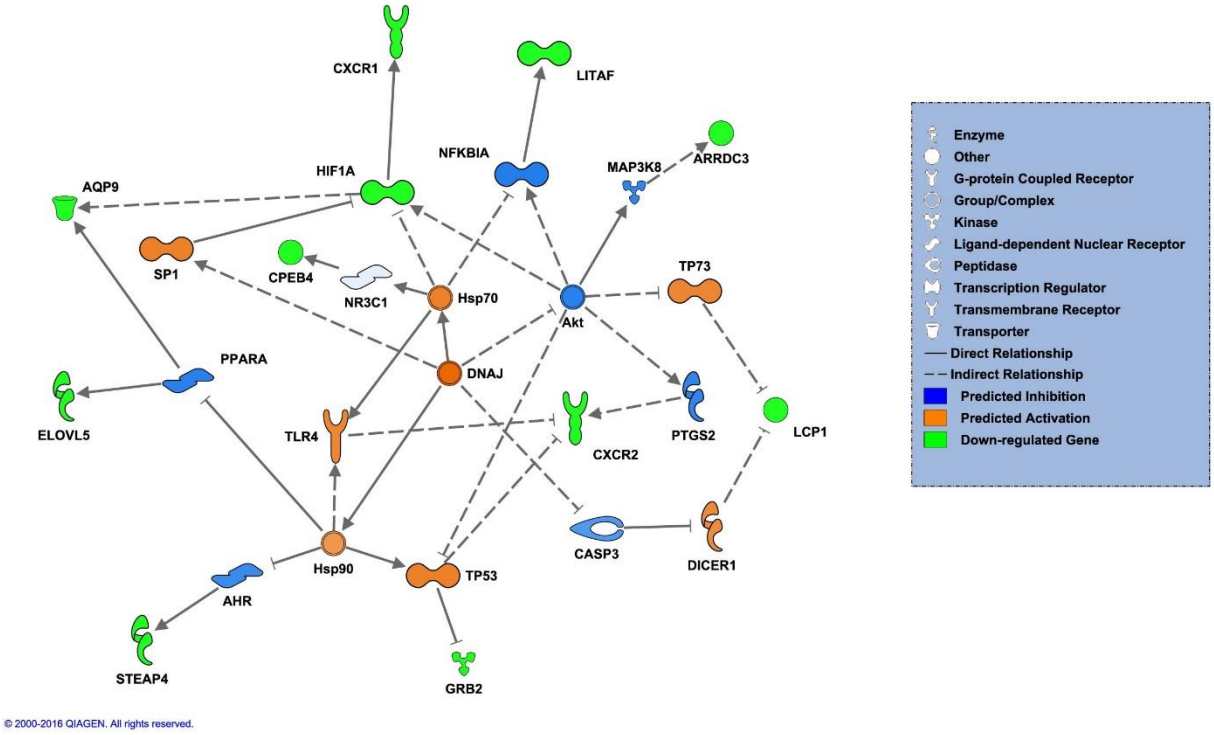


Figure S8

Ingenuity Pathway Analysis® Causal Network pathway D for Psychomotor Vigilance Test (PVT) effect genes with master regulator *DNAJ*. Lighter molecule color (e.g., *NR3C1*) suggests less confidence in prediction of the direction of expression; darker indicates stronger prediction.

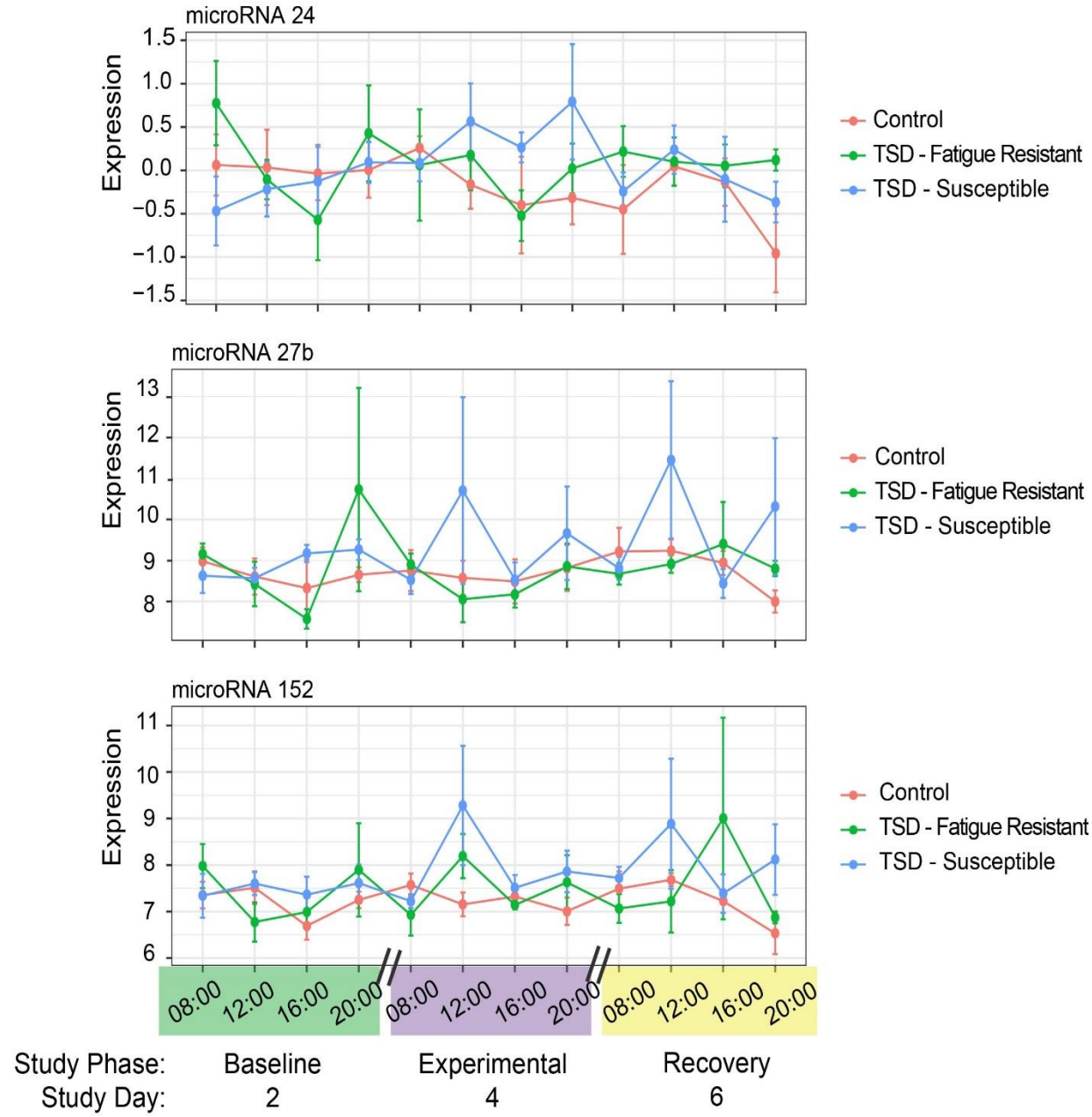


Figure S9

Mean (± 1 SE) expression profile of three miRNAs significantly related to PVT lapses, based on efficiency-corrected, normalized qPCR data.

References

1. Whitney P, Hinson JM, Jackson ML, Van Dongen HPA. Feedback blunting: total sleep deprivation impairs decision making that requires updating based on feedback. Sleep. 2015;38:745-54. doi: 10.5665/sleep.4668.
2. Elston DA, Moss R, Boulinier T, Arrowsmith C, Lambin X. Analysis of aggregation, a worked example: numbers of ticks on red grouse chicks. Parasitol. 2001;122:563-9. doi: doi:10.1017/S0031182001007740.
3. Bates D, Mächler M, Bolker B, Walker S. Fitting linear mixed-effects models using lme4. J. Stat Softw 2015;67:1-48. doi: 10.18637/jss.v067.i01.
4. Irizarry RA, Hobbs B, Collin F, Beazer-Barclay YD, Antonellis KJ, Scherf U, et al. Exploration, normalization, and summaries of high density oligonucleotide array probe level data. Biostatistics. 2003;4:249-64. doi: 10.1093/biostatistics/4.2.249.
5. Carvalho BS, Irizarry RA. A framework for oligonucleotide microarray preprocessing. Bioinformatics. 2010;26:2363-7. doi: 10.1093/bioinformatics/btq431.
6. Kauffmann A, Gentleman R, Huber W. arrayQualityMetrics—a bioconductor package for quality assessment of microarray data. Bioinformatics. 2009;25:415-6. doi: 10.1093/bioinformatics/btn647.
7. Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W, et al. *limma* powers differential expression analyses for RNA-sequencing and microarray studies. Nucleic Acids Res. 2015. doi: 10.1093/nar/gkv007.
8. da Costa Souza A, Ribeiro S. Sleep deprivation and gene expression. In: Meerlo P, Benca RM, Abel T, editors. Sleep, Neuronal Plasticity, and Brain Function. Current Topics in Behavioral Neurosciences. 25. New York: Springer; 2015. p. 65-90.

9. Krueger JM. The role of cytokines in sleep regulation. Curr Pharm Des. 2008;14:3408-16. doi: http://dx.doi.org/10.2174/138161208786549281.

10. National Heart, Lung, and Blood Institute; National Institutes of Health; U.S. Department of Health and Human Services: Calculate your body mass index. https://www.nhlbi.nih.gov/health/educational/lose_wt/BMI/bmi-m.htm. Accessed 18 December 2017.

11. Wickham H. ggplot2. Wiley Interdiscip Rev Comput Stat. 2011;3:180-5. doi: 10.1002/wics.147.

12. Langfelder P, Horvath S. WGCNA: an R package for weighted correlation network analysis. BMC Bioinformatics. 2008;9:1-13. doi: 10.1186/1471-2105-9-559.

13. Kumar L, Futschik ME. Mfuzz: a software package for soft clustering of microarray data. Bioinformation. 2007;2:5-7. http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2139991/.

14. Reverter A, Hudson NJ, Nagaraj SH, Pérez-Enciso M, Dalrymple BP. Regulatory impact factors: unraveling the transcriptional regulation of complex traits from expression data. Bioinformatics. 2010;26:896-904. doi: 10.1093/bioinformatics/btq051.

15. Uyhelji HA, Cheng C, Besansky NJ. Transcriptomic differences between euryhaline and stenohaline malaria vector sibling species in response to salinity stress. Mol Ecol. 2016;25:2210-25. doi: 10.1111/mec.13609.

16. Kel A, Voss N, Jauregui R, Kel-Margoulis O, Wingender E. Beyond microarrays: finding key transcription factors controlling signal transduction pathways. BMC Bioinformatics. 2006;7:S13. doi: 10.1186/1471-2105-7-s2-s13.

17. Matys V, Kel-Margoulis OV, Fricke E, Liebich I, Land S, Barre-Dirrie A, et al. TRANSFAC® and its module TRANSCompel®: transcriptional gene regulation in eukaryotes. Nucleic Acids Res. 2006;34:D108-D10. doi: 10.1093/nar/gkj143.

18. Liu G, Loraine AE, Shigeta R, Cline M, Cheng J, Valmeekam V, et al. NetAffx: Affymetrix probesets and annotations. Nucleic Acids Res. 2003;31:82-6. doi: 10.1093/nar/gkg121.

19. Huang DW, Sherman BT, Lempicki RA. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. Nat Protoc. 2008;4:44-57. doi: http://www.nature.com/nprot/journal/v4/n1/suppinfo/nprot.2008.211_S1.html.

20. Timmons JA, Szkop KJ, Gallagher IJ. Multiple sources of bias confound functional enrichment analysis of global -omics data. Genome Biol. 2015;16:1-3. doi: 10.1186/s13059-015-0761-7.

21. Krämer A, Green J, Pollard J, Jr., Tugendreich S. Causal analysis approaches in Ingenuity Pathway Analysis. Bioinformatics. 2014;30:523-30. doi: 10.1093/bioinformatics/btt703.

22. Pellegrino R, Sunaga DY, Guindalini C, Martins RCS, Mazzotti DR, Wei Z, et al. Whole blood genome-wide gene expression profile in males after prolonged wakefulness and sleep recovery. Physiol Genomics. 2012;44:1003-12. doi: 10.1152/physiolgenomics.00058.2012.

23. Arnardottir ES, Nikonova EV, Shockley KR, Podtelezhnikov AA, Anafi RC, Tanis KQ, et al. Blood-gene expression reveals reduced circadian rhythmicity in individuals resistant to sleep deprivation. Sleep. 2014;37:1589-600. doi: 10.5665/sleep.4064.

24. Möller-Levet CS, Archer SN, Bucca G, Laing EE, Slak A, Kabiljo R, et al. Effects of insufficient sleep on circadian rhythmicity and expression amplitude of the human blood transcriptome. Proc Natl Acad Sci U S A. 2013;110:E1132-E41. doi: 10.1073/pnas.1217154110.