BMJ open

Multi-Gene Interactions and the Prediction of Depression in the Wisconsin Longitudinal Study

Journal:	*BMJ Open*
Manuscript ID:	bmjopen-2012-000944
Article Type:	Research
Date Submitted by the Author:	26-Jan-2012
Complete List of Authors:	Roetker, Nicholas; UW-Madison, Sociology Yonker, James; UW-Madison, Sociology Lee, Chee; UW-Madison, Sociology Chang, Vicky; UW-Madison, Sociology Basson, Jacob; UW-Madison, Medicine Roan, Carol; UW-Madison, Sociology Hauser, Taissa; UW-Madison, Sociology Hauser, Robert; UW-Madison, Sociology Atwood, Craig; University of Wisconsin, Medicine
Primary Subject Heading:	Mental health
Secondary Subject Heading:	Epidemiology, Mental health, Neurology, Genetics and genomics, Public health
Keywords:	EPIDEMIOLOGY, MENTAL HEALTH, Neurogenetics < NEUROLOGY, Anxiety disorders < PSYCHIATRY, Depression & mood disorders < PSYCHIATRY

Multi-Gene Interactions and the Prediction of Depression in the Wisconsin Longitudinal Study

Nicholas S. Roetker[1], James A. Yonker[1], Chee Lee[1], Vicky Chang[1], Jacob Basson[2], Carol L. Roan[1], Taissa S. Hauser[1], Robert M. Hauser[1] and Craig S. Atwood[2,3].

[1]Department of Sociology, University of Wisconsin-Madison, Madison, WI 53706, USA.

[2]Geriatric Research, Education and Clinical Center, Veterans Administration Hospital and Department of Medicine, University of Wisconsin-Madison School of Medicine and Public Health, Madison, WI 53705, USA.

[3]School of Exercise, Biomedical and Health Sciences, Edith Cowan University, Joondalup, 6027 WA, Australia.

Running title: Gene interactions and depression

Address Correspondence and Reprint Requests to:
Craig S. Atwood, Ph.D.
University of Wisconsin-Madison School of Medicine and Public Health
Wm S. Middleton Memorial VA (GRECC 11G)
2500 Overlook Terrace, Madison, WI 53705, USA
Tel. 608 256-1901, Ext. 11664
Fax. 608 280-7291
Email: csa@medicine.wisc.edu

ABSTRACT

Objectives: Single genetic loci offer little predictive power for the identification of depression. This study examined whether an analysis of gene-gene interactions of 84 single nucleotide polymorphisms in genes associated with depression and age-related diseases would identify significant interactions with increased predictive power for depression.

Design: A retrospective cohort study.

Setting: A survey of participants in the Wisconsin Longitudinal Study.

Participants: A total of 4,792 persons (2,459 females and 2,333 males) who provided saliva for genotyping; the group comes from a randomly selected sample of Wisconsin high school graduates from the class of 1957 as well as a randomly selected sibling, almost all of whom are non-Hispanic white.

Primary outcome measure: Depression as determine by the Composite International Diagnostic Interview short-form (CIDI-SF).

Results: Using a classification tree approach (recursive partitioning (RP)) we identified a number of candidate gene-gene interactions associated with depression. The primary SNP splits revealed by RP (*ANKK1* rs1800497 in men and *DRD2* rs224592 in women) were found to be significant as single factors by logistic regression (LR) after controlling for multiple testing (P=0.001 for both). Without considering interaction effects, only 1 of the 5 subsequent RP splits reached nominal significance in logistic regression (*FTO* rs1421085 in women; P-value=0.008). However, after controlling for gene-gene interactions by running logistic regression on RP-specific subsets, every split became significant and grew larger in magnitude (OR [before]→[after]: Men: *GNRH1* novel SNP: [1.43 → 1.57]; Women: *APOC3* rs2854116: [1.28 → 1.56], *ACVR2B* rs3749386: [1.11 → 2.16], *FTO* rs1421085: [1.32 → 1.63], *IL6* rs1800795: [1.12 → 1.85]).

Conclusions: Our results suggest that examining gene-gene interactions improves the identification of genetic associations predictive of depression. Four of the SNPs identified in these interactions were located in two pathways well-known to impact depression: neurotransmitter (*ANKK1* and *DRD2*) and

neuroendocrine (*GNRH1* and *ACVR2B*) signaling. This study demonstrates the utility of RP analysis as an efficient and powerful exploratory analysis technique for uncovering genetic and molecular pathway interactions associated with disease etiology.

INTRODUCTION

Depression is a widespread mental disorder associated with a host of undesirable health, social, and economic outcomes. One in six Americans is diagnosed with depression in his or her lifetime (1). While many environmental factors—such as socioeconomic status, childhood abuse, and major life events—have important ties with depression, so too does gender and many genetic and epigenetic factors, making the disorder heterogeneous in nature (2). Another major risk factor for depression is age, with depression reaching its highest levels in adults over 80 years of age (3).

It has been demonstrated from twin studies that genetic factors typically account for 40–70% of the risk for developing major depressive disorder (MDD), and adoption studies have confirmed the role of genetic risk factors in the development of MDD (see (4) and references therein). Genetic studies, including recent genome-wide association studies (GWAS), have identified genetic alterations in over 50 genes known to be associated with depression (5). However, individually, the genetic alterations found within these genes (primarily single nucleotide polymorphisms (SNPs)) have little predictive value. There is a similar lack of predictive value from GWAS of other major age-related diseases (6).

Given this lack of predictive power among individual genetic alterations for depression together with the complex nature of aging-related diseases, it would seem prudent to examine epistatic effects on this age-related condition. In this respect, we have previously demonstrated that G x G interactions greatly modulate risk for complex age-related diseases (7, 8). Recent studies of depression also have identified epistatic effects. In particular, associations have been identified between *BDNF* Val66Met (brain-derived neurotrophic factor; rs6265) and *5-HTTLPR* (serotonin transporter linked promoter region (9); *GSK3B* rs6782799 (glycogen synthase kinase 3β), *BDNF* rs7124442 and *BDNF* Val66Met (10); *BDNF* Val66Met

and SNPs in *NTRK2* (neurotrophic tyrosine kinase receptor 2; (11)), and *5-HTTLPR* short allele and a chromosome 4 gene (12).

In this study, we have assessed the epistatic effects of known genetic alterations that link to depression and age-related diseases in the Wisconsin Longitudinal Study (WLS). Using recursive partitioning (RP) and logistic regression (LR) we identified associations between dopaminergic genes and depression in men and women, as well as G x G interactions involving neuroendocrine signaling pathways, with increased significance compared with single genetic associations.

METHODS

Study Participants and Surveys

Data were collected from the WLS, a random sample originally comprised of 10 317 men and women who graduated from Wisconsin high schools in 1957. Later in 1977, the WLS began interviewing one randomly selected sibling of each graduate, when possible. The cohort consists almost entirely of non-Hispanic white persons whose average level of educational attainment was 1.5 years of post-high school education at the time of interview in 2004. Ages of participants in the WLS ranged from 35 to 90 years old at this time, with 83% of participants being between 60 and 70 years old. Further characteristics of the WLS cohort may be found in detail elsewhere (13). Health and psychological well-being phenotypic data was taken from mail and phone surveys given in 2004-2005. Our main measure of depression is based on a variation of the Composite International Diagnostic Interview short-form (CIDI-SF). All participants answered a single stem question: “Have you ever had a time in life lasting two weeks or more when nearly every day you felt sad, blue, depressed, or when you lost interest in most things like work, hobbies, or things you usually liked to do for fun?” Only those who answered YES and whose depression was not always caused by alcohol, drugs, medications, or physical illness were asked further depression symptom questions. Symptom questions asked whether the two week period was accompanied with any weight loss, trouble sleeping, feeling tired, feeling bad upon waking, losing interest, trouble concentrating, or thoughts about death. Those answering YES to 3 or more of these symptom questions were classified as having depression (14).

Genotyping

7 101 participants (4 569 graduates & 2 532 siblings) provided saliva samples in Oragene DNA sample collection kits from which DNA was extracted and genotyped for 84 SNPs that were selected based on their association with depression and age-related conditions and diseases. Genotyping was performed by KBioscience (Hoddesdon, UK) with use of a homogeneous Fluorescent Resonance Energy Transfer

technology coupled to competitive allele specific PCR. All SNP genotypes described in our results were in Hardy-Weinberg equilibrium and their frequencies matched those reported in the literature for European samples.

Statistical Analysis

Of those participants that provided DNA and that also completed the survey depression questions (4 792), the following analyses were performed:

Recursive Partitioning (RP). RP is a data mining tool for revealing trends that relate a dependent variable (depression incidence) to various predictor variables (SNPs). Zhang and Bonney have shown how RP can be used in genetic association studies to identify disease genes (15). RP helps control for heterogeneity in the population and confounding factors by allowing for the segregation of the sample population according to any condition. Thus, RP is a useful way to handle complex datasets that might confound regression analysis due to the complexity of the relationship between the independent and dependent variables and due to missing information.

RP classification trees (using R package rpart) were used to identify potential interactions among the 84 SNPs in relation to depression. The trees split the data along branches according to criteria determined by the rpart package algorithm, which is originally based off the work of Breiman's classification and regression trees (CART) algorithm (16). Basically, the CART algorithm first considers all depressed and non-depressed subjects pooled together in a heterogeneous root node. Based on considering every possible "yes-no" binary partition that can be made by each independent variable, the single split which maximizes homogeneity between the two resulting sub-nodes as compared to the root node is made. Each sub-node can then be treated independently as a new root node for all subsequent splits, and the pattern continues until every subject constitutes a terminal node, resulting in a very large and complex tree. A 10-part cross validation procedure seeking to minimize misclassification and complexity

determines optimal pruning. See Therneau and Atkinson (17) for specific details of the rpart package. Tree nodes were re-created in Microsoft Visio to display depression incidence (in %) and total number of participants rather than the default number of controls/cases as presented by rpart.

Logistic Regression (LR). Variables found in association with depression based on RP analysis were considered in single factor LR models, separate by gender, using the specific dichotomous splitting of genotypes as designated by RP trees. Regression models for all seven SNP splits were first run on the full dataset to represent single main factor effects. Then each split was run on the respective subset of data as represented by the preceding RP split criteria. Thus, we attempt to mirror RP splits within a more formal LR framework in order to measure the significance of interactions presented by the trees. Multiple testing of 84 SNPs in RP for both male and females followed by 14 LR models resulted in a modified FDR significance level of 0.008.

RESULTS

Of the 4,792 participants with complete survey information on CIDI-SF depression (2 459 females and 2 333 males), we identified 711 participants (481 females and 230 males) with depression. Given that the independent variable gender (when included as a factor in the full dataset) was the primary split on RP trees; that women are over two times as likely to be diagnosed with depression than men; and since the female etiology of depression has been reported to be associated with unique social, psychological, and biological factors (18), all subsequent analyses were performed by gender.

Recursive Partition Analysis

To examine multi-gene interactions for association with depression we screened our dataset using RP. The two-factor RP tree (*ANKK1*/*GNRH1*) was the optimized pruning for men (Fig. 1), while the five-factor tree (*DRD2*/*APOC3*/*ACVR2B*/*FTO*/*IL6*) was the optimized pruning for women (Fig. 2). The best overall split for men was *ANKK1* rs1800497, where the incidence of depression increased 2.3-fold in those with no C-alleles compared to those with one or two C-alleles. Considering interaction between *ANKK1* and *GNRH1* widened the disparity in incidence, where those with at least one C-allele in both *ANKK1* rs1800497 and the novel SNP in *GNRH1* had a 3-fold lower incidence than those without a C-allele in *ANKK1* rs1800497.

For women, the best overall split was *DRD2* rs2242592, where those with one or two C-alleles had 1.3-fold higher incidence of depression compared to those without any C-alleles. G x G interactions associated with the highest incidence of depression included: *DRD2* rs2242592 T/T + *APOC3* rs45537037 T/T + *ACVR2B* rs3749386 C/C or T/T, accounting for a 1.4-fold increase in depression compared to baseline incidence.

Single Main-Factor Effects

Specific SNP interactions identified by RP were next analyzed by LR (see Table 1, Full Data). The

primary SNP splits in males and females were significant at the modified FDR level. Men with no C-alleles for *ANKK1* rs1800497 had 2.6 times higher odds [P=0.001 (1.5, 4.6)] of depression compared with men with at least 1 C-allele. Women with at least 1 C-allele for *DRD2* rs2242592 had 1.3 times higher odds [P=0.006 (1.1-1.6)] of depression compared with women with no C-alleles. One other split reached nominal significance; women homozygous (C/C or T/T) for *FTO* rs1421085 had 1.32 times higher odds [P=0.008 (1.1-1.6)] for depression than women with a heterozygous genotype. SNP splits of *GNRH1*, *APOC3*, *ACVR2B*, and *IL6* did not significantly associate with depression.

Gene-Gene Interactions Enhance Predictability for Depression

Specific SNP interactions identified by RP were next analyzed by LR as RP-specific subsets (see Table 1, RP-Subsetted Data). All 5 of the secondary and tertiary RP splits were found to be significant at the modified FDR level when considered as subsets. Among only men with at least one C-allele in *ANKK1* rs1800497, those with no C-allele in the novel SNP of *GNRH1* had 1.57 times higher odds [P=0.002 (1.2-2.1)] for depression than men with 1 or 2 C-alleles. For the subset of women in the first right-hand split of Fig. 2, those homozygous for *FTO* rs1421085 had 1.63 times higher odds [P=0.0006 (1.2-2.2)] for depression than women with a heterozygous genotype. For the remaining subset of women in the second right-hand split of Fig. 2, those homozygous for *IL6* rs1800795 had 1.85 times higher odds [P=0.007 (1.2-2.9)] for depression than women with a heterozygous genotype. For the subset of women in the first left-hand split of Fig. 2, those with no C-alleles for *APOC3* rs45537037 had 1.56 times higher odds [P=0.004 (1.2-2.1)] for depression than women with 1 or 2 C-alleles. For the subset of women in the second left-hand split of Fig. 2, those homozygous for *ACVR2B* rs3749386 had 2.16 times higher odds [P=0.001 (1.4-3.4)] for depression than women with a heterozygous genotype.

DISCUSSION

Utilizing RP as a screening tool to find potential multi-gene interactions, followed by verification of multi-gene interactions with LR, our data demonstrate that multi-gene interactions predict depression with a greater certainty than single main factor associations. RP provided us with primary dichotomous genotype splits in men and women (*ANKK1* rs1800497 and *DRD2* rs2242592, respectively) that were both significant in LR models at the modified FDR level (Table 1). Considering the 5 subsequent RP splits in LR over the entire dataset, only 1 reached a nominal level of significance (barely), which was *FTO* rs1421085 in women. However, after running LR on specific subsets of data according to the pattern of RP branches, every split was found to be significant and every odds ratios grew larger (Table 1; P-values [before]→[after]: Male Left: 1.43 → 1.57, Female Left 1: 1.28 → 1.56, Female Left 2: 1.11 → 2.16, Female Right 1: 1.32 → 1.63, Female Right 2: 1.12 → 1.85). Thus, RP provides two unique and important criteria: dichotomous genotype splitting instructions and gene-gene interaction patterns. These criteria go beyond the traditional single factor SNP approach to genetic association studies and allow identification of important multi-gene pathways that more suitably characterize the etiology of complex diseases.

The Utility of Recursive Partitioning, Multi-factor Dimensionality Reduction and Logistic Regression for Identification of Gene-Gene Interactions

With recent advances in genotyping allowing for high-dimensional SNP identification, it is now possible to examine genetic datasets not only for single main factor effects, but also G x G interactions. The requirement for G x G analyses as a better predictor of age-related diseases is obvious from the standpoint that humans are complex biological systems composed of numerous molecular interactions, and from recent studies indicating disease risk is modulated by G x G interactions (7). Notwithstanding this, the development of analytical tools for the identification of G x G interactions has not kept pace with the technological advances in identifying genetic alterations among individuals. In this respect, we have previously used MDR, LR and LD to identify G x G interactions among a small set of SNPs (7). However,

large datasets require a screening tool to identify potential multi-gene interactions. In this study, we have used RP to screen for multi-gene interactions, a data-mining technique that is currently under-utilized in genetic studies. RP serves as an efficient and powerful exploratory analysis technique, especially when looking for interactions in data sets with a large number of independent variables. This screening allows for the identification of G x G interactions (with greater explanatory power), that might otherwise not have been identified, and that can then be confirmed using more traditional statistical techniques. As illustrated in this paper, this data-mining methodology has the advantage of identification of genetic interactions *between* pathways involved in the etiology of depression, in keeping with the etiological heterogeneity of this disorder (see later).

Our study provides proof of principle for the use of RP in higher-dimensional analyses such as GWAS, where a comprehensive list of SNPs may fully explore genetic predisposition to depression and other age-related disease. The WLS is an ideal candidate for future GWAS studies given its large sample size, rich covariate composition and longitudinal nature.

In this genetic study we aimed to identify underlying genetic predispositions to depression and thus have not yet tested environmental/phenotypic data. Future analyses using RP to examine the impact of phenotypic and environmental factors on the development of depression would be anticipated to identify gene-phenotype/environment and multi-phenotype/environment interactions. Indeed, the predictive gains of G x G analyses were stronger for men than women, despite the fact that depression occurs disproportionately in women (~2:1 female-to-male; (19-23)). This suggests that environmental factors may be needed in addition to genetic factors in understanding the etiological pathways for women. Indeed, biological factors such as hormonal changes related to reproductive status (24, 25) may impact environmental factors such as psychosocial experiences (trauma, stress, interpersonal relationships, etc) and general health issues in the development of depression.

Genetic and Biological Correlates of Depression

Numerous studies have identified SNPs that associate with depression. Many of the SNPs associated with depression from other studies were not significantly associated in our study. This is perhaps not surprising, since a single factor is unlikely to provide consistent association especially in a complex condition such as depression, where multiple pathways intersect in regulating the risk of the disease. For example, if a SNP within the serotonin pathway also requires a SNP in the glutamatergic pathway in order for the patient to present with depression, the presence of either SNP in the absence of the other will not be predictive of depression. Moreover, as indicated by Shi and Weinberg, since the human genome contains genetic redundancy, disruption of a single gene may be selectively neutral, but the malfunction of several genes in a pathway might result in expression of a particular phenotype (26).

Both the primary splits in men and women were SNPs linked with *DRD2* (dopamine receptor D2), a gene that has previously been linked with depression and social phobia (27-29). The primary male genotype split rs1800497, technically found in gene *ANKK1*, is historically known as the *DRD2* Taq1A allele because of its known association with decreased dopamine receptor D2 density (in those with T alleles) (30-33). The Taq1A allele has also been previously associated with depressive symptoms in children, where those with the A1 allele (T) were more likely to have depressive symptoms (34). We saw a similar association between A1 and depression in WLS men, where those with two A1 alleles had 2.6 times higher odds for depression compared to those with one or no A1 alleles. The primary split in women (DRD2 rs2242592) has previously been found to be associated with schizophrenia, where the C-allele was associated with higher susceptibility for schizophrenia (35). Interestingly, this same study also found the Taq1A allele to also associate with schizophrenia.

The secondary and tertiary right-hand splits in the female RP tree—*FTO* (fat mass and obesity associated) rs1421085 and *IL6* (interleukin 6) rs1800795—have also been found to relate with mental illness and depression in previous studies (36)(37). There is evidence that activin receptor signaling also is involved in affective disorders, especially when considering interaction with GABAergic pathways (38). Although we did not see an interaction between SNPs in GABA/activin receptor genes and depression,

ACVR2B was associated with depression in women. No previous associations between depression and *APOC3*, *ACVR2B*, or *GNRH1* have been reported.

That these genetic variants are associated with *neuroendocrine* pathways (*GnRH1, ACVR2B*) that are known to regulate *neurotransmitter* release and cognitive behavior (39-40) supports these associations as relevant to the etiology of depression and underlines the benefits of using RP to identify meaningful G x G interactions associated with disease.

Limitations

Given the numerous genetic, phenotypic and environmental influences that are linked to depression, and the small number of SNPs analyzed, it is not surprising that predictability from our models was low (although our predictability was superior to previous studies examining only single main factors). Also, the predictive value of our statistical models was further limited due to user bias in selection of SNPs (from nearly two-million SNPs in the human genome) used in this study. As a result of this, interactions we have found could potentially be moderated by another gene that we have not considered in this study. Nonetheless, we identified significant G x G interactions between known, and newly identified, loci associated with depression. Importantly, 4 of the 7 SNPs identified in these interactions were primarily located in two pathways well-known to impact depression: neurotransmitter and neuroendocrine signaling.

The results from the RP analyses conducted in this study were confirmed by LR, demonstrating the utility of RP as a screening tool for identifying meaningful G x G interactions. Future development of algorithms for RP analysis should not only maximize the distance between branches of the next best split (i.e. rpart), but consider subsequent future split combinations that could potentially result in trees with "better" overall predictability.

Summary

Our data indicate that G x G interaction analyses allows for enhanced predictability of conditions and

diseases of aging. RP is an efficient and powerful exploratory analysis technique for elucidating G x G interactions in large datasets and combined with LR provides an important statistical analysis for the identification of well supported G x G interactions. We predict that such analytical methods will play an increasingly important role in the identification of epistatic effects in future large GWAS. Finally, our studies illustrate how RP analyses can be used to find interacting pathways involved in the etiology of a disease or condition such as depression.

ACKNOWLEDGMENTS

This research uses data from the Wisconsin Longitudinal Study (WLS) of the University of Wisconsin-Madison. Since 1991, the WLS has been supported principally by the National Institute on Aging (AG-9775, AG-21079 and AG-033285), with additional support from the Vilas Estate Trust, the National Science Foundation, the Spencer Foundation, and the Graduate School of the University of Wisconsin-Madison. A public use file of data from the Wisconsin Longitudinal Study is available from the Wisconsin Longitudinal Study, University of Wisconsin-Madison, 1180 Observatory Drive, Madison, Wisconsin 53706 and at http://www.ssc.wisc.edu/wlsresearch/data/. This material is the result of work supported with resources at the William S. Middleton Memorial Veterans Hospital, Madison, WI. The opinions expressed herein are those of the authors. The contents do not represent the views of the Dept. of Veterans Affairs or the United States Government.

Contributorship Statement

CSA, RMH and TSH conceptualized the study. RMH, TSH, CLR, NSR, CL and CSA collected saliva samples and performed genotyping analyses. NSR, JAY, CL, VC and JB performed statistical analyses on the WLS dataset. CSA and RMH directed the statistical analyses. NSR and CSA drafted the manuscript. All authors critically reviewed the manuscript and approved the final version.

Data Sharing

Genetic and environmental data for the WLS is available online at http://www.ssc.wisc.edu/wlsresearch/

Funding

Supported by the National Institute on Aging, grant numbers AG-9775, AG-21079 and AG-033285

Competing Interests

None

FIGURE LEGENDS

Figure 1. Recursive Partitioning Tree of CIDI-SF Depression in Males of the WLS. Upper and lower numbers in nodes represent the proportion of participants with depression and the number of participants in that node, respectively. Blue and purple boxes/circles indicate lower and higher rates of depression relative to the primary node, respectively. Split information indicates gene, SNP, and genotype criteria, respectively. M1 is subset of data referenced in Table 1. Sensitivity: 0.526, Specificity: 0.598, Accuracy: 0.591.

Figure 2. Recursive Partitioning Tree of CIDI-SF Depression in Females of the WLS. Upper and lower numbers in nodes represent the proportion of participants with depression and the number of participants in that node, respectively. Blue and purple boxes/circles indicate lower and higher rates of depression relative to the primary node, respectively. Split information indicates gene, SNP, and genotype criteria, respectively. F1-F4 are subsets referenced in Table 1. Sensitivity: 0.615, Specificity: 0.549, Accuracy: 0.562.

Table 1.
Single-factor logistic regression models based directly off male and female RP tree split criteria (see Figures 1 & 2). Each SNP split was first run on the full dataset to represent single main factor effects ("Full Data") for both males and females. Then the same SNP splits were run on specific subsets of data per RP tree splits (M1, F1-F4; "RP-Subsetted Data").

					Full Data		RP-Subsetted Data		
Gender	RP Split	Gene	SNP	Genotypes	OR (95% CI)	P-value	Subset	OR (95% CI)	P-value
Male	Primary	*ANKK1*	rs1800497	T/T vs. C/C + C/T	2.60 (1.47-4.61)	0.001 *	--	----	----
	Left	*GNRH1*	novel SNP	T/T vs. C/C + T/C	1.43 (1.08-1.88)	0.011	M1	1.57 (1.18-2.09)	0.002 *
Female	Primary	*DRD2*	rs2242592	C/C + T/C vs. T/T	1.33 (1.09-1.62)	0.001 *	--	----	----
	Left 1	*APOC3*	rs2854116	T/T vs. C/C + T/C	1.28 (1.05-1.57)	0.017	F1	1.56 (1.16-2.10)	0.004 *
	Left 2	*ACVR2B*	rs3749386	C/C + T/T vs. T/C	1.11 (0.91-1.36)	0.302	F2	2.16 (1.36-3.42)	0.001 *
	Right 1	*FTO*	rs1421085	C/C + T/T vs. T/C	1.32 (1.07-1.61)	0.008 *	F3	1.63 (1.23-2.17)	0.0006 *
	Right 2	*IL6*	rs1800795	C/C + G/G vs. C/G	1.12 (0.91-1.36)	0.283	F4	1.85 (1.18-2.88)	0.007 *

RP, recursive partitioning; OR, odds ratio; CI, confidence interval
M1: LR analysis was run for only those with genotype *DRD2* rs1800497 C/C or C/T
F1: LR analysis was run for only those with genotype *DRD2* rs2242592 T/T
F2: LR analysis was run for only those with genotypes *DRD2* rs2242592 T/T and *APOC3* rs2854116 T/T
F3: LR analysis was run for only those with genotype *DRD2* rs2242592 C/C or T/C
F4: LR analysis was run for only those with genotypes *DRD2* rs2242592 C/C or T/C and *FTO* rs1421085 T/C

REFERENCES

1. CDC. Anxiety and Depression. Atlanta, GA: Centers for Disease Control and Prevention; 2009 [cited 2010 Oct. 18]; Available from: http://www.cdc.gov/Features/dsBRFSSDepressionAnxiety/.

2. Eley TC, Sugden K, Corsico A, et al. Gene-environment interaction analysis of serotonin system markers with adolescent depression. Mol Psychiatry. 2004;9(10):908-15.

3. Mirowsky J, Ross CE. Age and Depression. Journal of Health and Social Behavior. 1992;33(3):187-205.

4. Zubenko GS, Zubenko WN, Spiker DG, et al. Malignancy of recurrent, early-onset major depression: a family study. Am J Med Genet. 2001 Dec 8;105(8):690-9.

5. Raymer KA, Waters RF, Price CR. Proposed multigenic Composite Inheritance in major depression. Medical Hypotheses. 2005;65(1):158-72.

6. Harold D, Abraham R, Hollingworth P, et al. Genome-wide association study identifies variants at CLU and PICALM associated with Alzheimer's disease. Nat Genet. 2009 Oct;41(10):1088-93.

7. Haasl R, Ahmadi MR, Meethal S, et al. A luteinizing hormone receptor intronic variant is significantly associated with decreased risk of Alzheimer's disease in males carrying an apolipoprotein E epsilon4 allele. BMC Medical Genetics. 2008;9(1):37.

8. Basson J, Wilson A, Haasl R, et al. Multi locus interactions in steroidogenic pathway genes predict Alzheimer's disease. Alzheimer's and Dementia. 2008;4(4 (Suppl. 2)):T579: P3-199.

9. Pezawas L, Meyer-Lindenberg A, Goldman AL, et al. Evidence of biologic epistasis between BDNF and SLC6A4 and implications for depression. Mol Psychiatry. 2008;13(7):709-16.

10. Zhang K, Yang C, Xu Y, et al. Genetic association of the interaction between the BDNF and GSK3B genes and major depressive disorder in a Chinese population. Journal of Neural Transmission. 2010;117(3):393-401.

11. Lin E, Hong CJ, Hwang JP, et al. Gene-gene interactions of the brain-derived neurotrophic-factor and neurotrophic tyrosine kinase receptor 2 genes in geriatric depression. Rejuvenation Res. 2009;12(6):387-93.

12. Neff CD, Abkevich V, Potter J, et al . Evidence for epistasis between SLC6A4 and a chromosome 4 gene as risk factors in major depression. American Journal of Medical Genetics Part B: Neuropsychiatric Genetics. 2010;153B(1):321-2.

13. Sewell WH. As We Age : The Wisconsin Longitudinal Study, 1957-2001: Center for Demography and Ecology University of Wisconsin--Madison; 2001.

14. Nelson CB, Kessler RC, Mroczek D. Scoring the World Health Organization's Composite International Diagnostic Interview Short Form (CIDI-SF; v1.0 NOV98). 1998.

15. Zhang H, Bonney G. Use of classification trees for association studies. Genet Epidemiol. 2000 Dec;19(4):323-32.

16. Breiman L, Friedman J, Olshen R, et al. Classification and Regression Trees. Boca Raton: Chapman and Hall/CRC; 1984.

17. Therneau T, Atkinson E, editors. Technical Report Series No. 61, An Introduction to Recursive Partitioning Using the RPART Routines. Rochester, MN: Department of Health Science Research, Mayo Clinic; 1997.

18. Accortt EE, Freeman MP, Allen JJB. Women and Major Depressive Disorder: Clinical Perspectives on Causal Pathways. Journal of Women's Health. 2008;17(10):1583-90.

19. Kessler RC, McGonagle KA, Swartz M, et al. Sex and depression in the National Comorbidity Survey. I: Lifetime prevalence, chronicity and recurrence. J Affect Disord. 1993 Oct-Nov;29(2-3):85-96.

20. Kessler RC, McGonagle KA, Nelson CB, et al. Sex and depression in the National Comorbidity Survey. II: Cohort effects. J Affect Disord. 1994 Jan;30(1):15-26.

21. Kessler RC, McGonagle KA, Zhao S, et al. Lifetime and 12-month prevalence of DSM-III-R psychiatric disorders in the United States. Results from the National Comorbidity Survey. Arch Gen Psychiatry. 1994 Jan;51(1):8-19.

22. Weissman MM, Bland R, Joyce PR, et al. Sex differences in rates of depression: cross-national perspectives. J Affect Disord. 1993 Oct-Nov;29(2-3):77-84.

23. Weissman MM, Olfson M. Depression in women: implications for health care research. Science. 1995 Aug 11;269(5225):799-801.

24. Pajer K. New strategies in the treatment of depression in women. The Journal of clinical psychiatry. [Review]. 1995;56 Suppl 2:30-7.

25. Paykel ES. Depression in women. Br J Psychiatry Suppl. [Review]. 1991 May(10):22-9.

26. Shi M, Weinberg CR. How Much Are We Missing in SNP-by-SNP Analyses of Genome-wide Association Studies? Epidemiology. 2011 Nov;22(6):845-7.

27. Lawford BR, Young R, Noble EP, et al. The D2 dopamine receptor (DRD2) gene is associated with co-morbid depression, anxiety and social dysfunction in untreated veterans with post-traumatic stress disorder. European Psychiatry. 2006;21(3):180-5.

28. Klimek V, Schenck JE, Han H, et al. Dopaminergic abnormalities in amygdaloid nuclei in major depression: a postmortem study. Biol Psychiatry. 2002 Oct 1;52(7):740-8.

29. Schneier FR, Liebowitz MR, Abi-Dargham A, et al. Low dopamine D(2) receptor binding potential in social phobia. Am J Psychiatry. 2000 Mar;157(3):457-9.

30. Noble EP, Blum K, Ritchie T, et al. Allelic Association of the D2 Dopamine Receptor Gene With Receptor-Binding Characteristics in Alcoholism or Gene ism. Arch Gen Psychiatry. 1991 July 1, 1991;48(7):648-54.

31. Thompson J, Thomas N, Singleton A, et al. D2 dopamine receptor gene (DRD2) Taq1 A polymorphism: reduced dopamine D2 receptor binding in the human striatum associated with the A1 allele. Pharmacogenetics. 1997 Dec;7(6):479-84.

32. Pohjalainen T, Rinne JO, Nagren K, et al. The A1 allele of the human D2 dopamine receptor gene predicts low D2 receptor availability in healthy volunteers. Mol Psychiatry. 1998 May;3(3):256-60.

33. Jonsson EG, Nothen MM, Grunhage F, et al. Polymorphisms in the dopamine D2 receptor gene and their relationships to striatal dopamine receptor density of healthy volunteers. Mol Psychiatry. 1999 May;4(3):290-6.

34. Hayden EP, Klein DN, Dougherty LR, et al. The dopamine D2 receptor gene and depressive and anxious symptoms in childhood: associations and evidence for gene-environment correlation and gene-environment interaction. Psychiatr Genet. 2010 Dec;20(6):304-10.

35. Dubertret C, Bardel C, Ramoz N, et al. A genetic schizophrenia-susceptibility region located between the ANKK1 and DRD2 genes. Progress in Neuro-Psychopharmacology and Biological Psychiatry. 2010;34(3):492-9.

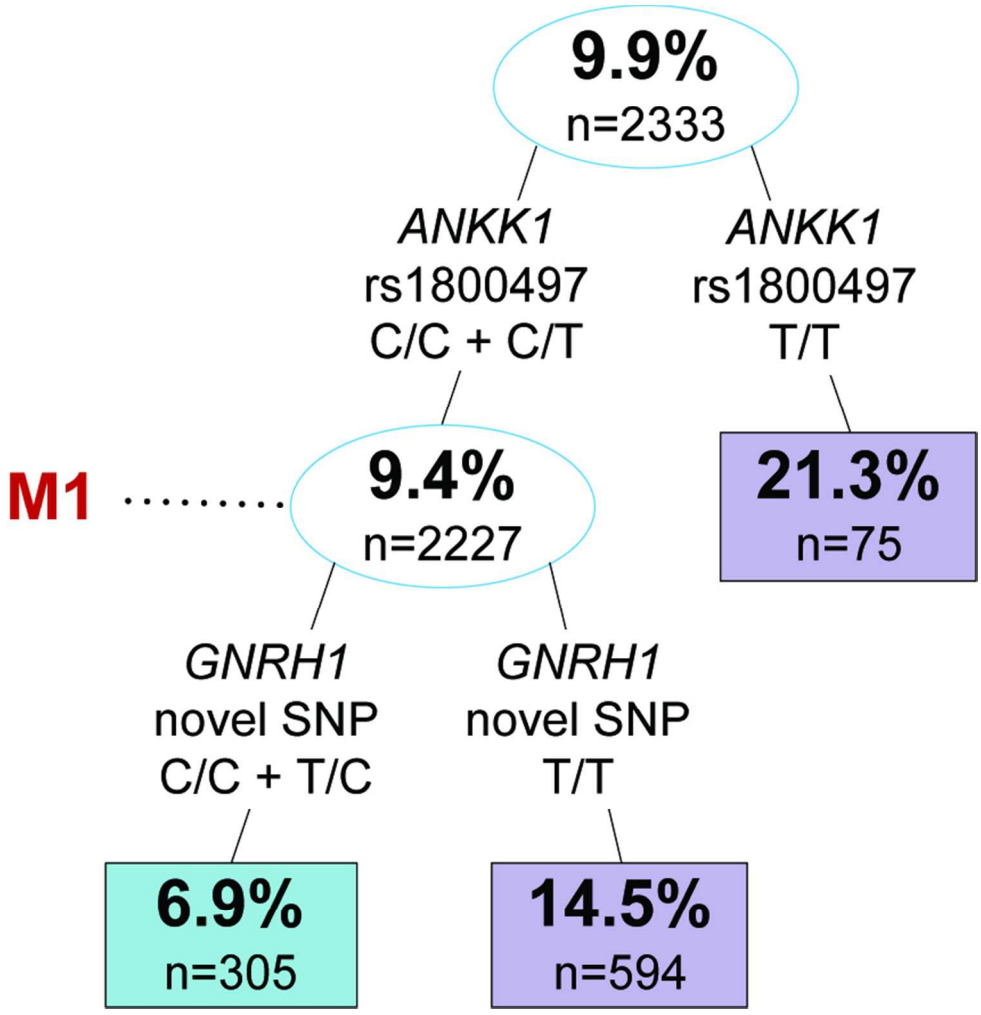
36. Kivimaki M, Jokela M, Hamer M, et al. Examining Overweight and Obesity as Risk Factors for Common Mental Disorders Using Fat Mass and Obesity-Associated (FTO) Genotype-Instrumented Analysis. Am J Epidemiol. 2011 Feb 15;173(4):421-9.

37. Bull SJ, Huezo-Diaz P, Binder EB, et al. Functional polymorphisms in the interleukin-6 and serotonin transporter genes, and depression and fatigue induced by interferon-alpha and ribavirin treatment. Mol Psychiatry. 2009 Dec;14(12):1095-104.

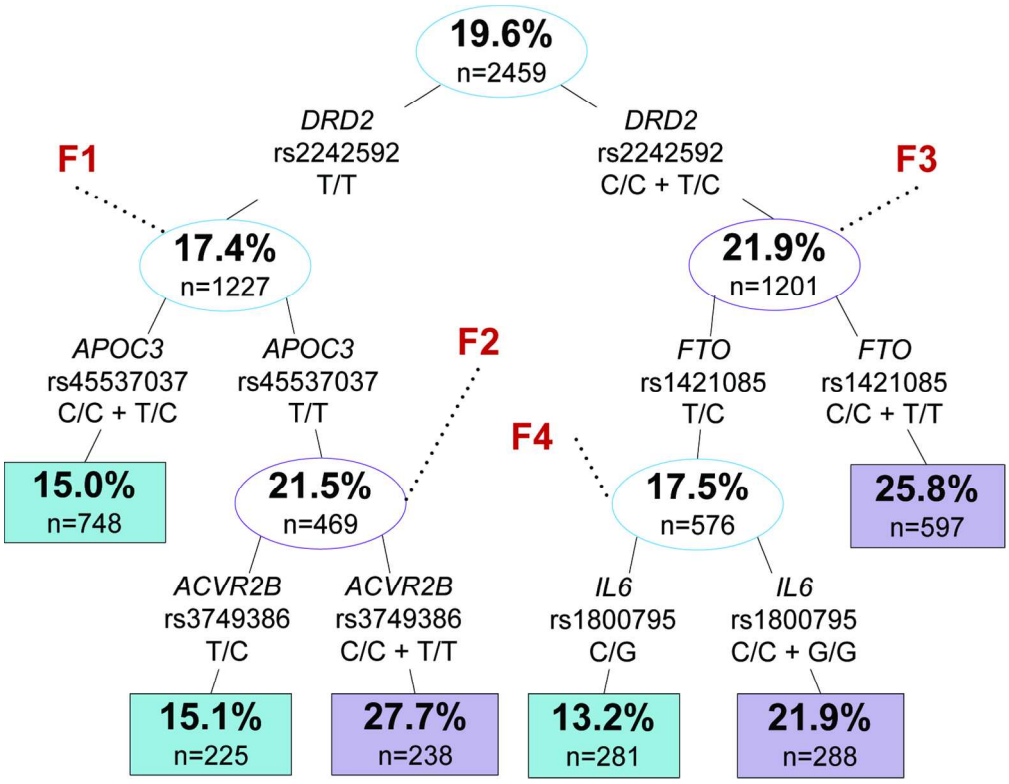
38. Zheng F, Adelsberger H, Muller MR, Fet al. Activin tunes GABAergic neurotransmission and modulates anxiety-like behavior. Mol Psychiatry. 2008;14(3):332-46.

39. Dong Y, Zhang H, Wang X, et al. A Leu184Val polymorphism in PCK1 gene is associated with type 2 diabetes in Eastern Chinese population with BMI < 23 kg/m2. Diabetes Research and Clinical Practice. 2009;83(2):227-32.

40. Atwood CS, Meethal SV, Liu T, et al. Dysregulation of the hypothalamic-pituitary-gonadal axis with menopause and andropause promotes neurodegenerative senescence. J Neuropathol Exp Neurol. 2005;64(2):93-103.



74x74mm (300 x 300 DPI)



132x102mm (300 x 300 DPI)

Supplementary Table 1. Depression-Associated SNP Identified in the WLS

Gene	SNP	Name	Alleles	Chr#/Location	Residue	Associated disease/behavior
ACVR2B	rs3749386	activin receptor IIB	T/C	3/intron 1	--	left-right axis malformations*, (1)
APOC3	rs2854116	apolipoprotein C-III	T/C	11/promoter (-455)	--	nonalcoholic fatty liver disease, insulin resistance[2]
DRD2/ANKK1	rs1800497	dopamine receptor D2/ankyrin repeat and kinase domain containing 1	C/T	11/exon (ANKK1)	Glu713Lys	obesity, drug addiction[3]
DRD2	rs2242592	dopamine receptor D2	T/C	11/3'	--	Schizophrenia[4]
FTO	rs1421085	fat mass and obesity associated	T/C	16/intron 1	--	Obesity[5-7]; mental disorders (in men)[8]
GNRH1	novel SNP	gonadotropin-releasing hormone promoter	T/C	8/promoter	--	Alzheimer's disease[9]
IL6	rs1800795	interleukin 6 (interferon, beta 2)	C/G	7/promoter (-174)	--	Arthritis[10], breast cancer[11]; type II diabetes[12]; depression[13]

*Gene association only

References

1. Kosaki R, Gebbia M, Kosaki K, Lewin M, Bowers P, Towbin JA, et al. Left–right axis malformations associated with mutations in ACVR2B, the gene for human activin receptor type IIB. American Journal of Medical Genetics. 1999;82(1):70-6.

2. Petersen KF, Dufour S, Hariri A, Nelson-Williams C, Foo JN, Zhang X-M, et al. Apolipoprotein C3 Gene Variants in Nonalcoholic Fatty Liver Disease. New England Journal of Medicine. 2010;362(12):1082-9.

3. Blum K, Braverman ER, Wood RC, Gill J, Li C, Chen TJ, et al. Increased prevalence of the Taq I A1 allele of the dopamine receptor gene (DRD2) in obesity with comorbid substance use disorder: a preliminary report. Pharmacogenetics. 1996 Aug;6(4):297-305.

4. Dubertret C, Bardel C, Ramoz N, Martin P-M, Deybach J-C, Adès J, et al. A genetic schizophrenia-susceptibility region located between the ANKK1 and DRD2 genes. Progress in Neuro-Psychopharmacology and Biological Psychiatry. 2010;34(3):492-9.

5. Dina C, Meyre D, Gallina S, Durand E, Korner A, Jacobson P, et al. Variation in FTO contributes to childhood obesity and severe adult obesity. Nat Genet. [10.1038/ng2048]. 2007;39(6):724-6.

6. Tung Y-CL, Yeo GSH. From GWAS to biology: lessons from FTO. Annals of the New York Academy of Sciences. 2011;1220(1):162-71.

7. Chabris C, Hebert B, Benjamin D, Beauchamp J, Cesarini D, van der Loos M, et al. Most reported genetic assocations with general intelligence are probably false positives. Psychological Science. 2011;In press.

8Kivimaki M, Jokela M, Hamer M, Geddes J, Ebmeier K, Kumari M, et al. Examining Overweight and Obesity as Risk Factors for Common Mental Disorders Using Fat Mass and Obesity-Associated (FTO) Genotype-Instrumented Analysis. Am J Epidemiol. 2011 Feb 15;173(4):421-9.

9. Wilson AC. Investigations into gonadotropin-releasing hormone receptor signaling in extrahypothalamic neurons of the brain: Implications for normal cognition and the pathogenesis of Alzheimer's disease: Thesis. University of Wisconsin, Madison, USA; 2009.

10. Fishman D, Faulds G, Jeffery R, Mohamed-Ali V, Yudkin JS, Humphries S, et al. The effect of novel polymorphisms in the interleukin-6 (IL-6) gene on IL-6 transcription and plasma IL-6 levels, and an association with systemic-onset juvenile chronic arthritis. The Journal of Clinical Investigation. 1998;102(7):1369-76.

11. Slattery ML, Curtin K, Sweeney C, Wolff RK, Baumgartner RN, Baumgartner KB, et al. Modifying Effects of IL-6 Polymorphisms on Body Size-Associated Breast Cancer Risk. Obesity. 2008;16(2):339-47.

12. Huth C, Heid IM, Vollmert C, Gieger C, Grallert H, Wolford JK, et al. IL6 Gene Promoter Polymorphisms and Type 2 Diabetes. Diabetes. 2006 October 2006;55(10):2915-21.

13. Bull SJ, Huezo-Diaz P, Binder EB, Cubells JF, Ranjith G, Maddock C, et al. Functional polymorphisms in the interleukin-6 and serotonin transporter genes, and depression and fatigue induced by interferon-alpha and ribavirin treatment. Mol Psychiatry. 2009 Dec;14(12):1095-104.

STROBE Statement—Checklist of items that should be included in reports of ***cohort studies***

	Item No	Recommendation
Title and abstract	1	(*a*) Indicate the study's design with a commonly used term in the title or the abstract
		(*b*) Provide in the abstract an informative and balanced summary of what was done and what was found
Introduction		
Background/rationale	2	Explain the scientific background and rationale for the investigation being reported
Objectives	3	State specific objectives, including any prespecified hypotheses
Methods		
Study design	4	Present key elements of study design early in the paper
Setting	5	Describe the setting, locations, and relevant dates, including periods of recruitment, exposure, follow-up, and data collection
Participants	6	(*a*) Give the eligibility criteria, and the sources and methods of selection of participants. Describe methods of follow-up
		(*b*) For matched studies, give matching criteria and number of exposed and unexposed
Variables	7	Clearly define all outcomes, exposures, predictors, potential confounders, and effect modifiers. Give diagnostic criteria, if applicable
Data sources/ measurement	8*	For each variable of interest, give sources of data and details of methods of assessment (measurement). Describe comparability of assessment methods if there is more than one group
Bias	9	Describe any efforts to address potential sources of bias
Study size	10	Explain how the study size was arrived at
Quantitative variables	11	Explain how quantitative variables were handled in the analyses. If applicable, describe which groupings were chosen and why
Statistical methods	12	(*a*) Describe all statistical methods, including those used to control for confounding
		(*b*) Describe any methods used to examine subgroups and interactions
		(*c*) Explain how missing data were addressed
		(*d*) If applicable, explain how loss to follow-up was addressed
		(*e*) Describe any sensitivity analyses
Results		
Participants	13*	(a) Report numbers of individuals at each stage of study—eg numbers potentially eligible, examined for eligibility, confirmed eligible, included in the study, completing follow-up, and analysed
		(b) Give reasons for non-participation at each stage
		(c) Consider use of a flow diagram
Descriptive data	14*	(a) Give characteristics of study participants (eg demographic, clinical, social) and information on exposures and potential confounders
		(b) Indicate number of participants with missing data for each variable of interest
		(c) Summarise follow-up time (eg, average and total amount)
Outcome data	15*	Report numbers of outcome events or summary measures over time
Main results	16	(*a*) Give unadjusted estimates and, if applicable, confounder-adjusted estimates and their precision (eg, 95% confidence interval). Make clear which confounders were adjusted for and why they were included
		(*b*) Report category boundaries when continuous variables were categorized
		(*c*) If relevant, consider translating estimates of relative risk into absolute risk for a meaningful time period

Other analyses	17	Report other analyses done—eg analyses of subgroups and interactions, and sensitivity analyses
Discussion		
Key results	18	Summarise key results with reference to study objectives
Limitations	19	Discuss limitations of the study, taking into account sources of potential bias or imprecision. Discuss both direction and magnitude of any potential bias
Interpretation	20	Give a cautious overall interpretation of results considering objectives, limitations, multiplicity of analyses, results from similar studies, and other relevant evidence
Generalisability	21	Discuss the generalisability (external validity) of the study results
Other information		
Funding	22	Give the source of funding and the role of the funders for the present study and, if applicable, for the original study on which the present article is based

*Give information separately for exposed and unexposed groups.

Note: An Explanation and Elaboration article discusses each checklist item and gives methodological background and published examples of transparent reporting. The STROBE checklist is best used in conjunction with this article (freely available on the Web sites of PLoS Medicine at http://www.plosmedicine.org/, Annals of Internal Medicine at http://www.annals.org/, and Epidemiology at http://www.epidem.com/). Information on the STROBE Initiative is available at http://www.strobe-statement.org.

BMJ open

Multi-Gene Interactions and the Prediction of Depression in the Wisconsin Longitudinal Study

Journal:	*BMJ Open*
Manuscript ID:	bmjopen-2012-000944.R1
Article Type:	Research
Date Submitted by the Author:	30-Apr-2012
Complete List of Authors:	Roetker, Nicholas; UW-Madison, Sociology Yonker, James; UW-Madison, Sociology Lee, Chee; UW-Madison, Sociology Chang, Vicky; UW-Madison, Sociology Basson, Jacob; UW-Madison, Medicine Roan, Carol; UW-Madison, Sociology Hauser, Taissa; UW-Madison, Sociology Hauser, Robert; UW-Madison, Sociology Atwood, Craig; University of Wisconsin, Medicine
Primary Subject Heading:	Mental health
Secondary Subject Heading:	Epidemiology, Mental health, Neurology, Genetics and genomics, Public health
Keywords:	EPIDEMIOLOGY, MENTAL HEALTH, Neurogenetics < NEUROLOGY, Anxiety disorders < PSYCHIATRY, Depression & mood disorders < PSYCHIATRY

STROBE Statement—Checklist of items that should be included in reports of ***cohort studies***

	Item No	Recommendation
Title and abstract	1	(*a*) Indicate the study's design with a commonly used term in the title or the abstract
		(*b*) Provide in the abstract an informative and balanced summary of what was done and what was found
Introduction		
Background/rationale	2	Explain the scientific background and rationale for the investigation being reported
Objectives	3	State specific objectives, including any prespecified hypotheses
Methods		
Study design	4	Present key elements of study design early in the paper
Setting	5	Describe the setting, locations, and relevant dates, including periods of recruitment, exposure, follow-up, and data collection
Participants	6	(*a*) Give the eligibility criteria, and the sources and methods of selection of participants. Describe methods of follow-up
		(*b*) For matched studies, give matching criteria and number of exposed and unexposed
Variables	7	Clearly define all outcomes, exposures, predictors, potential confounders, and effect modifiers. Give diagnostic criteria, if applicable
Data sources/ measurement	8*	For each variable of interest, give sources of data and details of methods of assessment (measurement). Describe comparability of assessment methods if there is more than one group
Bias	9	Describe any efforts to address potential sources of bias
Study size	10	Explain how the study size was arrived at
Quantitative variables	11	Explain how quantitative variables were handled in the analyses. If applicable, describe which groupings were chosen and why
Statistical methods	12	(*a*) Describe all statistical methods, including those used to control for confounding
		(*b*) Describe any methods used to examine subgroups and interactions
		(*c*) Explain how missing data were addressed
		(*d*) If applicable, explain how loss to follow-up was addressed
		(*e*) Describe any sensitivity analyses
Results		
Participants	13*	(a) Report numbers of individuals at each stage of study—eg numbers potentially eligible, examined for eligibility, confirmed eligible, included in the study, completing follow-up, and analysed
		(b) Give reasons for non-participation at each stage
		(c) Consider use of a flow diagram
Descriptive data	14*	(a) Give characteristics of study participants (eg demographic, clinical, social) and information on exposures and potential confounders
		(b) Indicate number of participants with missing data for each variable of interest
		(c) Summarise follow-up time (eg, average and total amount)
Outcome data	15*	Report numbers of outcome events or summary measures over time
Main results	16	(*a*) Give unadjusted estimates and, if applicable, confounder-adjusted estimates and their precision (eg, 95% confidence interval). Make clear which confounders were adjusted for and why they were included
		(*b*) Report category boundaries when continuous variables were categorized
		(*c*) If relevant, consider translating estimates of relative risk into absolute risk for a meaningful time period

Other analyses	17	Report other analyses done—eg analyses of subgroups and interactions, and sensitivity analyses
Discussion		
Key results	18	Summarise key results with reference to study objectives
Limitations	19	Discuss limitations of the study, taking into account sources of potential bias or imprecision. Discuss both direction and magnitude of any potential bias
Interpretation	20	Give a cautious overall interpretation of results considering objectives, limitations, multiplicity of analyses, results from similar studies, and other relevant evidence
Generalisability	21	Discuss the generalisability (external validity) of the study results
Other information		
Funding	22	Give the source of funding and the role of the funders for the present study and, if applicable, for the original study on which the present article is based

*Give information separately for exposed and unexposed groups.

Note: An Explanation and Elaboration article discusses each checklist item and gives methodological background and published examples of transparent reporting. The STROBE checklist is best used in conjunction with this article (freely available on the Web sites of PLoS Medicine at http://www.plosmedicine.org/, Annals of Internal Medicine at http://www.annals.org/, and Epidemiology at http://www.epidem.com/). Information on the STROBE Initiative is available at http://www.strobe-statement.org.

Multi-Gene Interactions and the Prediction of Depression in the Wisconsin Longitudinal Study

Nicholas S. Roetker[1], James A. Yonker[1], Chee Lee[1], Vicky Chang[1], Jacob Basson[2], Carol L. Roan[1], Taissa S. Hauser[1], Robert M. Hauser[1] and Craig S. Atwood[2,3].

[1]Department of Sociology, University of Wisconsin-Madison, Madison, WI 53706, USA.

[2]Geriatric Research, Education and Clinical Center, Veterans Administration Hospital and Department of Medicine, University of Wisconsin-Madison School of Medicine and Public Health, Madison, WI 53705, USA.

[3]School of Exercise, Biomedical and Health Sciences, Edith Cowan University, Joondalup, 6027 WA, Australia.

Running title: Gene interactions and depression

Address Correspondence and Reprint Requests to:
Craig S. Atwood, Ph.D.
University of Wisconsin-Madison School of Medicine and Public Health
Wm S. Middleton Memorial VA (GRECC 11G)
2500 Overlook Terrace, Madison, WI 53705, USA
Tel. 608 256-1901, Ext. 11664
Fax. 608 280-7291
Email: csa@medicine.wisc.edu

ABSTRACT

Objectives: Single genetic loci offer little predictive power for the identification of depression. This study examined whether an analysis of gene-gene interactions of 78 single nucleotide polymorphisms in genes associated with depression and age-related diseases would identify significant interactions with increased predictive power for depression.

Design: A retrospective cohort study.

Setting: A survey of participants in the Wisconsin Longitudinal Study.

Participants: A total of 4,811 persons (2,464 females and 2,347 males) who provided saliva for genotyping; the group comes from a randomly selected sample of Wisconsin high school graduates from the class of 1957 as well as a randomly selected sibling, almost all of whom are non-Hispanic white.

Primary outcome measure: Depression as determine by the Composite International Diagnostic Interview short-form (CIDI-SF).

Results: Using a classification tree approach (recursive partitioning (RP)) we identified a number of candidate gene-gene interactions associated with depression. The primary SNP splits revealed by RP (*ANKK1* rs1800497 (also known as *DRD2* Taq1A) in men and *DRD2* rs224592 in women) were found to be significant as single factors by logistic regression (LR) after controlling for multiple testing (P=0.001 for both). Without considering interaction effects, only 1 of the 5 subsequent RP splits reached nominal significance in logistic regression (*FTO* rs1421085 in women; P-value=0.008). However, after controlling for gene-gene interactions by running logistic regression on RP-specific subsets, every split became significant and grew larger in magnitude (OR [before]→[after]: Men: *GNRH1* novel SNP: [1.43 → 1.57]; Women: *APOC3* rs2854116: [1.28 → 1.55], *ACVR2B* rs3749386: [1.11 → 2.17], *FTO* rs1421085: [1.32 → 1.65], *IL6* rs1800795: [1.12 → 1.85]).

Conclusions: Our results suggest that examining gene-gene interactions improves the identification of genetic associations predictive of depression. Four of the SNPs identified in these interactions were located in two pathways well-known to impact depression: neurotransmitter (*ANKK1* and *DRD2*) and

Formatted: Font: Italic

neuroendocrine (*GNRH1* and *ACVR2B*) signaling. This study demonstrates the utility of RP analysis as an efficient and powerful exploratory analysis technique for uncovering genetic and molecular pathway interactions associated with disease etiology.

INTRODUCTION

Depression is a widespread mental disorder associated with a host of undesirable health, social, and economic outcomes. One in six Americans is diagnosed with depression in his or her lifetime (1). While many environmental factors—such as socioeconomic status, childhood abuse, and major life events—have important ties with depression, so too does gender and many genetic and epigenetic factors, making the disorder heterogeneous in nature (2). Another major risk factor for depression is age, with depression reaching its highest levels in adults over 80 years of age (3).

It has been demonstrated from twin studies that genetic factors typically account for 40–70% of the risk for developing major depressive disorder (MDD), and adoption studies have confirmed the role of genetic risk factors in the development of MDD (see (4) and references therein). Genetic studies, including recent genome-wide association studies (GWAS), have identified genetic alterations in over 50 genes known to be associated with depression (5). However, individually, the genetic alterations found within these genes (primarily single nucleotide polymorphisms (SNPs)) have little predictive value. There is a similar lack of predictive value from GWAS of other major age-related diseases (6).

Given this lack of predictive power among individual genetic alterations for depression together with the complex nature of aging-related diseases, it would seem prudent to examine epistatic effects on this age-related condition. In this respect, we have previously demonstrated that G x G interactions greatly modulate risk for complex age-related diseases (7, 8). Recent studies of depression also have identified epistatic effects. In particular, associations have been identified between *BDNF* Val66Met (brain-derived neurotrophic factor; rs6265) and *5-HTTLPR* (serotonin transporter linked promoter region (9); *GSK3B* rs6782799 (glycogen synthase kinase 3β), *BDNF* rs7124442 and *BDNF* Val66Met (10); *BDNF* Val66Met and SNPs in *NTRK2* (neurotrophic tyrosine kinase receptor 2; (11)), and *5-HTTLPR* short allele and a chromosome 4 gene (12).

The goals of this study were therefore to 1) explore G x G interactions that might better predict the genetic factors involved in the etiology of depression, and 2) to determine the utility of machine learning

algorithms (recursive partitioning) to identify genetic interactions. Using genotypic data from the ~~In this study, we have assessed the epistatic effects of known genetic alterations that link to depression and age-related diseases in the~~ Wisconsin Longitudinal Study (WLS)~~. Using recursive partitioning (RP) and logistic regression (LR)~~ we identified associations between dopaminergic genes and depression in men and women, as well as G x G interactions involving neuroendocrine signaling pathways, with increased significance compared with single genetic associations.

METHODS

Study Participants and Surveys

Data were collected from the WLS, a random sample originally comprised of 10,317 men and women who graduated from Wisconsin high schools in 1957. Later in 1977, the WLS began interviewing one randomly selected sibling of each graduate, when possible. The cohort ~~consists~~ reflects the ancestral makeup of the late-1950s Wisconsin population in that participants are almost entirely ~~of~~ non-Hispanic white ~~persons~~males and females. ~~whose average level of educational attainment was 1.5 years of post-high school education at the time of interview in 2004. Ages of participants in the WLS ranged from 35 to 90 years old at this time, with 83% of participants being between 60 and 70 years old~~. In general, the sample is broadly representative of older white Americans with at least a high school education (13). Further characteristics of the WLS cohort may be found in detail elsewhere (14). Health and psychological well-being phenotypic data was taken from mail and phone surveys given in 2004-2005. Our main measure of depression is based on a variation of the Composite International Diagnostic Interview short-form (CIDI-SF). All participants answered a single stem question: "Have you ever had a time in life lasting two weeks or more when nearly every day you felt sad, blue, depressed, or when you lost interest in most things like work, hobbies, or things you usually liked to do for fun?" Only those who answered YES and whose depression was not always caused by alcohol, drugs, medications, or physical illness were asked further depression symptom questions. Symptom questions asked whether the two week period was accompanied with a) any weight loss, b) trouble sleeping, c) feeling tired, d) feeling bad upon waking, e) losing interest, f) trouble concentrating, or g) thoughts about death. Those answering YES to 3 or more of these symptom questions were classified as having depression (15). Those answering YES to 2 or fewer symptom questions and all those answering NO to the initial stem question were classified as controls.

Formatted: Font color: Red

Genotyping

7,101 participants (4,569 graduates & 2,532 siblings) provided saliva samples in Oragene DNA sample

collection kits from which DNA was extracted and genotyped for 78 SNPs that were selected based on their association with depression and age-related conditions and diseases (see Supplementary Information 1). Genotyping was performed by KBioscience (Hoddesdon, UK) with use of a homogeneous Fluorescent Resonance Energy Transfer technology coupled to competitive allele specific PCR. All SNP genotypes described in our results were in Hardy-Weinberg equilibrium and their frequencies matched those reported in the literature for European samples.

Statistical Analysis

Analyses were limited to the 4,811 pooled graduates and siblings for whom we have depression and genotype information (Note: individuals with more than 10% missing genotype data were not included). The average age among this sample was just under 65 years in 2004. 80% were married, and the average amount of post-high school educational attainment was 2 years. Median household income in 1993 was $56,700.

Formatted: Indent: Left: 0", First line: 0"

Recursive Partitioning (RP). RP is a data mining tool for revealing trends that relate a dependent variable (~~depression incidence~~depressed vs. non-depressed) to various predictor variables (SNPs). Zhang and Bonney have shown how RP can be used in genetic association studies to identify disease genes (16). RP helps control for heterogeneity in the population and confounding factors by allowing for the segregation of the sample population according to any condition. Thus, RP is a useful way to handle complex datasets that might confound regression analysis due to the complexity of the relationship between the independent and dependent variables and due to missing information.

RP classification trees (using R package rpart) were used to identify potential interactions among the 78 SNPs in relation to depression. The trees split the data along branches according to criteria determined by the rpart package algorithm, which is originally based off the work of Breiman's classification and regression trees (CART) algorithm (17). Basically, the CART algorithm first considers all depressed and

non-depressed subjects pooled together in a heterogeneous root node. Based on considering every possible "yes-no" binary partition that can be made by each independent variable, the single split which maximizes homogeneity between the two resulting sub-nodes as compared to the root node is made. Each sub-node can then be treated independently as a new root node for all subsequent splits, and the pattern continues until every subject constitutes a terminal node, resulting in a very large and complex tree. A 10-part cross validation procedure seeking to minimize misclassification and complexity determines optimal pruning. See Therneau and Atkinson (18) for specific details of the rpart package. Priors were set to 0.5, 0.5. The usesurrogate parameter was set to 0 so that subjects missing the primary split variable do not progress further down the tree, and maxsurrogate was set to 0 to cut computation time in half. The threshold complexity parameter (cp) was set to 0.01. Tree nodes were re-created in Microsoft Visio to display percentage depressed ~~depression incidence (in %) and total number of participants rather than~~and the default number of controls/cases as presented by rpart.

Logistic Regression (LR). Variables found in association with depression based on RP analysis were considered in single factor LR models, separate by gender, using the specific dichotomous splitting of genotypes as designated by RP trees. Regression models for all seven SNP splits were first run on the full dataset to represent single main factor effects. Then each split was run on the respective subset of data as represented by the preceding RP split criteria. Thus, we attempt to mirror RP splits within a more formal LR framework in order to measure the significance of interactions presented by the trees. Multiple testing of 78 SNPs in RP for both male and females followed by 14 LR models resulted in a modified FDR significance level of 0.009.

RESULTS

Of the 4,811 participants ~~with complete survey information on CIDI-SF depression~~ (2,464 females and 2,347 males) under examination in this study, we identified 713 participants (481 females and 232 males) with depression (14.8 %). Given that the independent variable gender (when included as a factor in the full dataset) was the primary split on RP trees; that women are over two times as likely to be diagnosed with depression than men; and since the female etiology of depression has been reported to be associated with unique social, psychological, and biological factors (19), all subsequent analyses were performed by gender.

Recursive Partition Analysis

To examine multi-gene interactions for association with depression we screened our dataset using RP. The two-factor RP tree (*ANKK1*/*GNRH1*) was the optimized pruning for men (Fig. 1), while the five-factor tree (*DRD2*/*APOC3*/*ACVR2B*/*FTO*/*IL6*) was the optimized pruning for women (Fig. 2). For more detailed information on the 7 SNPs found by RP, see Supplementary Information 2. The best overall split for men was *ANKK1* rs1800497 (historically known as the *DRD2* Taq1A allele), where the incidence of depression increased 2.2-fold in those with no C-alleles compared to those with one or two C-alleles. Considering interaction between *ANKK1* and *GNRH1* widened the disparity in incidence, where those with at least one C-allele in both *ANKK1* rs1800497 and the novel SNP in *GNRH1* had a 2.7-fold lower incidence than those without a C-allele in *ANKK1* rs1800497.

For women, the best overall split was *DRD2* rs2242592, where those with one or two C-alleles had 1.3-fold higher incidence of depression compared to those without any C-alleles. G x G interactions associated with the highest incidence of depression included: *DRD2* rs2242592 T/T + *APOC3* rs45537037 T/T + *ACVR2B* rs3749386 C/C or T/T, accounting for a 1.4-fold increase in depression compared to baseline incidence.

Single Main-Factor Effects

Specific SNP interactions identified by RP were next analyzed by LR (see Table 1, Full Data). The primary SNP splits in males and females were significant at the modified FDR level. Men with no C-alleles for *ANKK1* rs1800497 had 2.55 times higher odds [P=0.001 (1.44, 4.51)] of depression compared with men with at least 1 C-allele. Women with at least 1 C-allele for *DRD2* rs2242592 had 1.32 times higher odds [P=0.006 (1.08-1.62)] of depression compared with women with no C-alleles. One other split reached nominal significance; women homozygous (C/C or T/T) for *FTO* rs1421085 had 1.32 times higher odds [P=0.008 (1.08-1.62)] for depression than women with a heterozygous genotype. SNP splits of *GNRH1*, *APOC3*, *ACVR2B*, and *IL6* did not significantly associate with depression.

Gene-Gene Interactions Enhance Predictability for Depression

Specific SNP interactions identified by RP were next analyzed by LR as RP-specific subsets (see Table 1, RP-Subsetted Data). All 5 of the secondary and tertiary RP splits were found to be significant at the modified FDR level when considered as subsets. Among only men with at least one C-allele in *ANKK1* rs1800497, those with no C-allele in the novel SNP of *GNRH1* had 1.57 times higher odds [P=0.002 (1.18-2.08)] for depression than men with 1 or 2 C-alleles. For the subset of women in the first right-hand split of Fig. 2, those homozygous for *FTO* rs1421085 had 1.65 times higher odds [P=0.0005 (1.24-2.18)] for depression than women with a heterozygous genotype. For the remaining subset of women in the second right-hand split of Fig. 2, those homozygous for *IL6* rs1800795 had 1.85 times higher odds [P=0.006 (1.19-2.89)] for depression than women with a heterozygous genotype. For the subset of women in the first left-hand split of Fig. 2, those with no C-alleles for *APOC3* rs45537037 had 1.55 times higher odds [P=0.004 (1.15-2.09)] for depression than women with 1 or 2 C-alleles. For the subset of women in the second left-hand split of Fig. 2, those homozygous for *ACVR2B* rs3749386 had 2.17 times higher odds [P=0.001 (1.37-3.44)] for depression than women with a heterozygous genotype.

DISCUSSION

Utilizing RP as a screening tool to find potential multi-gene interactions, followed by verification of multi-gene interactions with LR, our data demonstrate that multi-gene interactions predict depression with a greater certainty than single main factor associations. RP provided us with primary dichotomous genotype splits in men and women (*ANKK1* rs1800497 and *DRD2* rs2242592, respectively) that were both significant in LR models at the modified FDR level (Table 1). Considering the 5 subsequent RP splits in LR over the entire dataset, only 1 reached a nominal level of significance (barely), which was *FTO* rs1421085 in women. However, after running LR on specific subsets of data according to the pattern of RP branches, every split was found to be significant and every odds ratios grew larger (Table 1; OR [before]→[after]: Male Left: 1.43 → 1.57, Female Left 1: 1.28 → 1.55, Female Left 2: 1.11 → 2.17, Female Right 1: 1.32 → 1.65, Female Right 2: 1.12 → 1.85). Thus, RP provides two unique and important criteria: dichotomous genotype splitting instructions and gene-gene interaction patterns. These criteria go beyond the traditional single factor SNP approach to genetic association studies and allow identification of important multi-gene pathways that more suitably characterize the etiology of complex diseases.

The Utility of Recursive Partitioning, ~~Multi-factor Dimensionality Reduction~~ and Logistic Regression for Identification of Gene-Gene Interactions

With recent advances in genotyping allowing for high-dimensional SNP identification, it is now possible to examine genetic datasets not only for single main factor effects, but also G x G interactions. The requirement for G x G analyses as a better predictor of age-related diseases is obvious from the standpoint that humans are complex biological systems composed of numerous molecular interactions, and from recent studies indicating disease risk is modulated by G x G interactions (7). Notwithstanding this, the development of analytical tools for the identification of G x G interactions has not kept pace with

the technological advances in identifying genetic alterations among individuals. In this respect, we have previously used MDR, LR and LD to identify G x G interactions among a small set of SNPs (7). However, large datasets require a screening tool to identify potential multi-gene interactions. In this study, we have used RP to screen for multi-gene interactions, a data-mining technique that is currently under-utilized in genetic studies. RP serves as an efficient and powerful exploratory analysis technique, especially when looking for interactions in data sets with a large number of independent variables. This screening allows for the identification of G x G interactions (with greater explanatory power), that might otherwise not have been identified, and that can then be confirmed using more traditional statistical techniques. As illustrated in this paper, this data-mining methodology has the advantage of identification of genetic interactions *between* pathways involved in the etiology of depression, in keeping with the etiological heterogeneity of this disorder (see later).

Our study provides proof of principle for the use of RP in higher-dimensional analyses such as GWAS, where a comprehensive list of SNPs may fully explore genetic predisposition to depression and other age-related disease. The WLS is an ideal candidate for future GWAS studies given its large sample size, rich covariate composition and longitudinal nature.

In this genetic study we aimed to identify underlying genetic predispositions to depression and thus have not yet tested environmental/phenotypic data. Future analyses using RP to examine the impact of phenotypic and environmental factors on the development of depression would be anticipated to identify gene-phenotype/environment and multi-phenotype/environment interactions. Indeed, the predictive gains of G x G analyses were stronger for men than women, despite the fact that depression occurs disproportionately in women (~2:1 female-to-male; (20-24)). This suggests that environmental factors may be needed in addition to genetic factors in understanding the etiological pathways for women. Indeed, biological factors such as hormonal changes related to reproductive status (25, 26) may impact environmental factors such as psychosocial experiences (trauma, stress, interpersonal relationships, etc) and general health issues in the development of depression.

Genetic and Biological Correlates of Depression

Numerous studies have identified SNPs that associate with depression. Many of the SNPs associated with depression from other studies were not significantly associated in our study. This is perhaps not surprising, since a single factor is unlikely to provide consistent association especially in a complex condition such as depression, where multiple pathways intersect in regulating the risk of the disease. For example, if a SNP within the serotonin pathway also requires a SNP in the glutamatergic pathway in order for the patient to present with depression, the presence of either SNP in the absence of the other will not be predictive of depression. Moreover, as indicated by Shi and Weinberg, since the human genome contains genetic redundancy, disruption of a single gene may be selectively neutral, but the malfunction of several genes in a pathway might result in expression of a particular phenotype (27).

Both the primary splits in men and women were SNPs linked with *DRD2* (dopamine receptor D2), a gene that has previously been linked with depression and social phobia (28-30). The primary male genotype split rs1800497, technically found in gene *ANKK1*, is historically known as the *DRD2* Taq1A allele because of its known association with decreased dopamine receptor D2 density (in those with T alleles) (31-34). The Taq1A allele has also been previously associated with depressive symptoms in children, where those with the A1 allele (T) were more likely to have depressive symptoms (35). We saw a similar association between A1 and depression in WLS men, where those with two A1 alleles had 2.6 times higher odds for depression compared to those with one or no A1 alleles. The primary split in women (DRD2 rs2242592) has previously been found to be associated with schizophrenia, where the C-allele was associated with higher susceptibility for schizophrenia (36). Interestingly, this same study also found the Taq1A allele to also associate with schizophrenia.

The secondary and tertiary right-hand splits in the female RP tree—*FTO* (fat mass and obesity associated) rs1421085 and *IL6* (interleukin 6) rs1800795—have also been found to relate with mental illness and depression in previous studies (37, 38). There is evidence that activin receptor signaling also

is involved in affective disorders, especially when considering interaction with GABAergic pathways (39). Although we did not see an interaction between SNPs in GABA/activin receptor genes and depression, *ACVR2B* was associated with depression in women. No previous associations between depression and *APOC3*, *ACVR2B*, or *GNRH1* have been reported.

That these genetic variants are associated with *neuroendocrine* pathways (*GnRH1, ACVR2B*) that are known to regulate *neurotransmitter* release and cognitive behavior (39-40) supports these associations as relevant to the etiology of depression and underlines the benefits of using RP to identify meaningful G x G interactions associated with disease.

Formatted: Font color: Red

Limitations

Given the numerous genetic, phenotypic and environmental influences that are linked to depression, and the small number of SNPs analyzed, it is not surprising that predictability from our models was low (although our predictability was superior to previous studies examining only single main factors). Also, the predictive value of our statistical models was further limited due to user bias in selection of SNPs (from nearly two-million SNPs in the human genome) used in this study. As a result of this, interactions we have found could potentially be moderated by another gene that we have not considered in this study. Nonetheless, we identified significant G x G interactions between known, and newly identified, loci associated with depression. Importantly, 4 of the 7 SNPs identified in these interactions were primarily located in two pathways well-known to impact depression: neurotransmitter and neuroendocrine signaling.

The results from the RP analyses conducted in this study were confirmed by LR, demonstrating the utility of RP as a screening tool for identifying meaningful G x G interactions. Future development of algorithms for RP analysis should not only maximize the distance between branches of the next best split (i.e. rpart), but consider subsequent future split combinations that could potentially result in trees with "better" overall predictability.

Summary

Our data indicate that G x G interaction analyses allows for enhanced predictability of conditions and diseases of aging. RP is an efficient and powerful exploratory analysis technique for elucidating G x G interactions in large datasets and combined with LR provides an important statistical analysis for the identification of well supported G x G interactions. We predict that such analytical methods will play an increasingly important role in the identification of epistatic effects in future large GWAS. Finally, our studies illustrate how RP analyses can be used to find interacting pathways involved in the etiology of a disease or condition such as depression.

ACKNOWLEDGMENTS

This research uses data from the Wisconsin Longitudinal Study (WLS) of the University of Wisconsin-Madison. Since 1991, the WLS has been supported principally by the National Institute on Aging (AG-9775, AG-21079 and AG-033285), with additional support from the Vilas Estate Trust, the National Science Foundation, the Spencer Foundation, and the Graduate School of the University of Wisconsin-Madison. A public use file of data from the Wisconsin Longitudinal Study is available from the Wisconsin Longitudinal Study, University of Wisconsin-Madison, 1180 Observatory Drive, Madison, Wisconsin 53706 and at http://www.ssc.wisc.edu/wlsresearch/data/. This material is the result of work supported with resources at the William S. Middleton Memorial Veterans Hospital, Madison, WI. The opinions expressed herein are those of the authors. The contents do not represent the views of the Dept. of Veterans Affairs or the United States Government.

FIGURE LEGENDS

Figure 1. Recursive Partitioning Tree of CIDI-SF Depression in Males of the WLS. Upper and lower numbers in nodes represent the ~~proportion~~ percentage of participants with depression and the number of ~~participants~~ controls/cases in that node, respectively. Blue and purple boxes/circles indicate lower and higher rates of depression relative to the primary node, respectively. Split information indicates gene, SNP, and genotype criteria, respectively. M1 is subset of data referenced in Table 1. Sensitivity: 0.526, Specificity: 0.598, Accuracy: 0.591. Due to missing genotype information, we lose approximately 1.5% of participants per split. *rs1800497 is historically known as the *DRD2* Taq1A allele

Figure 2. Recursive Partitioning Tree of CIDI-SF Depression in Females of the WLS. Upper and lower numbers in nodes represent the ~~proportion~~ percentage of participants with depression and the number of ~~participants~~ controls/cases in that node, respectively. Blue and purple boxes/circles indicate lower and higher rates of depression relative to the primary node, respectively. Split information indicates gene, SNP, and genotype criteria, respectively. F1-F4 are subsets referenced in Table 1. Sensitivity: 0.607, Specificity: 0.563, Accuracy: 0.572. Due to missing genotype information, we lose approximately 1.4% of participants per split.

Table 1.
Single-factor logistic regression models based directly off male and female RP tree split criteria (see Figures 1 & 2). Each SNP split was first run on the full dataset to represent single main factor effects ("Full Data") for both males and females. Then the same SNP splits were run on specific subsets of data per RP tree splits (M1, F1-F4; "RP-Subsetted Data").

					Full Data		RP-Subsetted Data		
Gender	RP Split	Gene	SNP	Genotypes	OR (95% CI)	P-value	Subset	OR (95% CI)	P-value
Male	Primary	*ANKK1**	rs1800497	T/T vs. C/C + C/T	2.55 (1.44-4.51)	0.001 *	--	----	----
	Left	*GNRH1*	novel SNP	T/T vs. C/C + T/C	1.43 (1.09-1.88)	0.011	M1	1.57 (1.18-2.08)	0.002 *
Female	Primary	*DRD2*	rs2242592	C/C + T/C vs. T/T	1.32 (1.08-1.62)	0.006 *	--	----	----
	Left 1	*APOC3*	rs2854116	T/T vs. C/C + T/C	1.28 (1.04-1.57)	0.018	F1	1.55 (1.15-2.09)	0.004 *
	Left 2	*ACVR2B*	rs3749386	C/C + T/T vs. T/C	1.11 (0.91-1.36)	0.302	F2	2.17 (1.37-3.44)	0.001 *
	Right 1	*FTO*	rs1421085	C/C + T/T vs. T/C	1.32 (1.08-1.62)	0.007 *	F3	1.65 (1.24-2.18)	0.0005 *
	Right 2	*IL6*	rs1800795	C/C + G/G vs. C/G	1.12 (0.92-1.37)	0.269	F4	1.85 (1.19-2.89)	0.006 *

RP, recursive partitioning; OR, odds ratio; CI, confidence interval
M1: LR analysis was run for only those with genotype *DRD2* rs1800497 C/C or C/T
F1: LR analysis was run for only those with genotype *DRD2* rs2242592 T/T
F2: LR analysis was run for only those with genotypes *DRD2* rs2242592 T/T and *APOC3* rs2854116 T/T
F3: LR analysis was run for only those with genotype *DRD2* rs2242592 C/C or T/C
F4: LR analysis was run for only those with genotypes *DRD2* rs2242592 C/C or T/C and *FTO* rs1421085 T/C
*rs1800497 is historically known as the *DRD2* Taq1A allele

20

REFERENCES

1. CDC. Anxiety and Depression. Atlanta, GA: Centers for Disease Control and Prevention; 2009 [cited 2010 Oct. 18]; Available from: http://www.cdc.gov/Features/dsBRFSSDepressionAnxiety/.

2. Eley TC, Sugden K, Corsico A, Gregory AM, Sham P, McGuffin P, et al. Gene-environment interaction analysis of serotonin system markers with adolescent depression. Mol Psychiatry. 2004;9(10):908-15.

3. Mirowsky J, Ross CE. Age and Depression. Journal of Health and Social Behavior. 1992;33(3):187-205.

4. Zubenko GS, Zubenko WN, Spiker DG, Giles DE, Kaplan BB. Malignancy of recurrent, early-onset major depression: a family study. Am J Med Genet. 2001 Dec 8;105(8):690-9.

5. Raymer KA, Waters RF, Price CR. Proposed multigenic Composite Inheritance in major depression. Medical Hypotheses. 2005;65(1):158-72.

6. Harold D, Abraham R, Hollingworth P, Sims R, Gerrish A, Hamshere ML, et al. Genome-wide association study identifies variants at CLU and PICALM associated with Alzheimer's disease. Nat Genet. 2009 Oct;41(10):1088-93.

7. Haasl R, Ahmadi MR, Meethal S, Gleason C, Johnson S, Asthana S, et al. A luteinizing hormone receptor intronic variant is significantly associated with decreased risk of Alzheimer's disease in males carrying an apolipoprotein E epsilon4 allele. BMC Medical Genetics. 2008;9(1):37.

8. Basson J, Wilson A, Haasl R, Atwood C. Multi locus interactions in steroidogenic pathway genes predict Alzheimer's disease. Alzheimer's and Dementia. 2008;4(4 (Suppl. 2)):T579: P3-199.

9. Pezawas L, Meyer-Lindenberg A, Goldman AL, Verchinski BA, Chen G, Kolachana BS, et al. Evidence of biologic epistasis between BDNF and SLC6A4 and implications for depression. Mol Psychiatry. 2008;13(7):709-16.

Comment [NR1]: Note that we have added 1 reference (#13).

10. Zhang K, Yang C, Xu Y, Sun N, Yang H, Liu J, et al. Genetic association of the interaction between the BDNF and GSK3B genes and major depressive disorder in a Chinese population. Journal of Neural Transmission. 2010;117(3):393-401.

11. Lin E, Hong CJ, Hwang JP, Liou YJ, Yang CH, Cheng D, et al. Gene-gene interactions of the brain-derived neurotrophic-factor and neurotrophic tyrosine kinase receptor 2 genes in geriatric depression. Rejuvenation Res. 2009;12(6):387-93.

12. Neff CD, Abkevich V, Potter J, Riley R, Shattuck D, Katz DA. Evidence for epistasis between SLC6A4 and a chromosome 4 gene as risk factors in major depression. American Journal of Medical Genetics Part B: Neuropsychiatric Genetics. 2010;153B(1):321-2.

13. Carr D, Khodyakov D. End-of-life health care planning among young-old adults: An assessment of psychosocial influences. J Gerontol B-Psychol. 2007 Mar;62(2):S135-S41.

14. Sewell WH. As We Age : The Wisconsin Longitudinal Study, 1957-2001: Center for Demography and Ecology University of Wisconsin--Madison; 2001.

15. Nelson CB, Kessler RC, Mroczek D. Scoring the World Health Organization's Composite International Diagnostic Interview Short Form (CIDI-SF; v1.0 NOV98). 1998.

16. Zhang H, Bonney G. Use of classification trees for association studies. Genet Epidemiol. 2000 Dec;19(4):323-32.

17. Breiman L, Friedman J, Olshen R, Stone C. Classification and Regression Trees. Boca Raton: Chapman and Hall/CRC; 1984.

18. Therneau T, Atkinson E, editors. Technical Report Series No. 61, An Introduction to Recursive Partitioning Using the RPART Routines. Rochester, MN: Department of Health Science Research, Mayo Clinic; 1997.

19. Accortt EE, Freeman MP, Allen JJB. Women and Major Depressive Disorder: Clinical Perspectives on Causal Pathways. Journal of Women's Health. 2008;17(10):1583-90.

20. Kessler RC, McGonagle KA, Swartz M, Blazer DG, Nelson CB. Sex and depression in the National Comorbidity Survey. I: Lifetime prevalence, chronicity and recurrence. J Affect Disord. 1993 Oct-Nov;29(2-3):85-96.

21. Kessler RC, McGonagle KA, Nelson CB, Hughes M, Swartz M, Blazer DG. Sex and depression in the National Comorbidity Survey. II: Cohort effects. J Affect Disord. 1994 Jan;30(1):15-26.

22. Kessler RC, McGonagle KA, Zhao S, Nelson CB, Hughes M, Eshleman S, et al. Lifetime and 12-month prevalence of DSM-III-R psychiatric disorders in the United States. Results from the National Comorbidity Survey. Arch Gen Psychiatry. 1994 Jan;51(1):8-19.

23. Weissman MM, Bland R, Joyce PR, Newman S, Wells JE, Wittchen HU. Sex differences in rates of depression: cross-national perspectives. J Affect Disord. 1993 Oct-Nov;29(2-3):77-84.

24. Weissman MM, Olfson M. Depression in women: implications for health care research. Science. 1995 Aug 11;269(5225):799-801.

25. Pajer K. New strategies in the treatment of depression in women. The Journal of clinical psychiatry. [Review]. 1995;56 Suppl 2:30-7.

26. Paykel ES. Depression in women. Br J Psychiatry Suppl. [Review]. 1991 May(10):22-9.

27. Shi M, Weinberg CR. How Much Are We Missing in SNP-by-SNP Analyses of Genome-wide Association Studies? Epidemiology. 2011 Nov;22(6):845-7.

28. Lawford BR, Young R, Noble EP, Kann B, Ritchie T. The D2 dopamine receptor (DRD2) gene is associated with co-morbid depression, anxiety and social dysfunction in untreated veterans with post-traumatic stress disorder. European Psychiatry. 2006;21(3):180-5.

29. Klimek V, Schenck JE, Han H, Stockmeier CA, Ordway GA. Dopaminergic abnormalities in amygdaloid nuclei in major depression: a postmortem study. Biol Psychiatry. 2002 Oct 1;52(7):740-8.

30. Schneier FR, Liebowitz MR, Abi-Dargham A, Zea-Ponce Y, Lin SH, Laruelle M. Low dopamine D(2) receptor binding potential in social phobia. Am J Psychiatry. 2000 Mar;157(3):457-9.

31. Noble EP, Blum K, Ritchie T, Montgomery A, Sheridan PJ. Allelic Association of the D2 Dopamine Receptor Gene With Receptor-Binding Characteristics in Alcoholism or Gene ism. Arch Gen Psychiatry. 1991 July 1, 1991;48(7):648-54.

32. Thompson J, Thomas N, Singleton A, Piggott M, Lloyd S, Perry EK, et al. D2 dopamine receptor gene (DRD2) Taq1 A polymorphism: reduced dopamine D2 receptor binding in the human striatum associated with the A1 allele. Pharmacogenetics. 1997 Dec;7(6):479-84.

33. Pohjalainen T, Rinne JO, Nagren K, Lehikoinen P, Anttila K, Syvalahti EK, et al. The A1 allele of the human D2 dopamine receptor gene predicts low D2 receptor availability in healthy volunteers. Mol Psychiatry. 1998 May;3(3):256-60.

34. Jonsson EG, Nothen MM, Grunhage F, Farde L, Nakashima Y, Propping P, et al. Polymorphisms in the dopamine D2 receptor gene and their relationships to striatal dopamine receptor density of healthy volunteers. Mol Psychiatry. 1999 May;4(3):290-6.

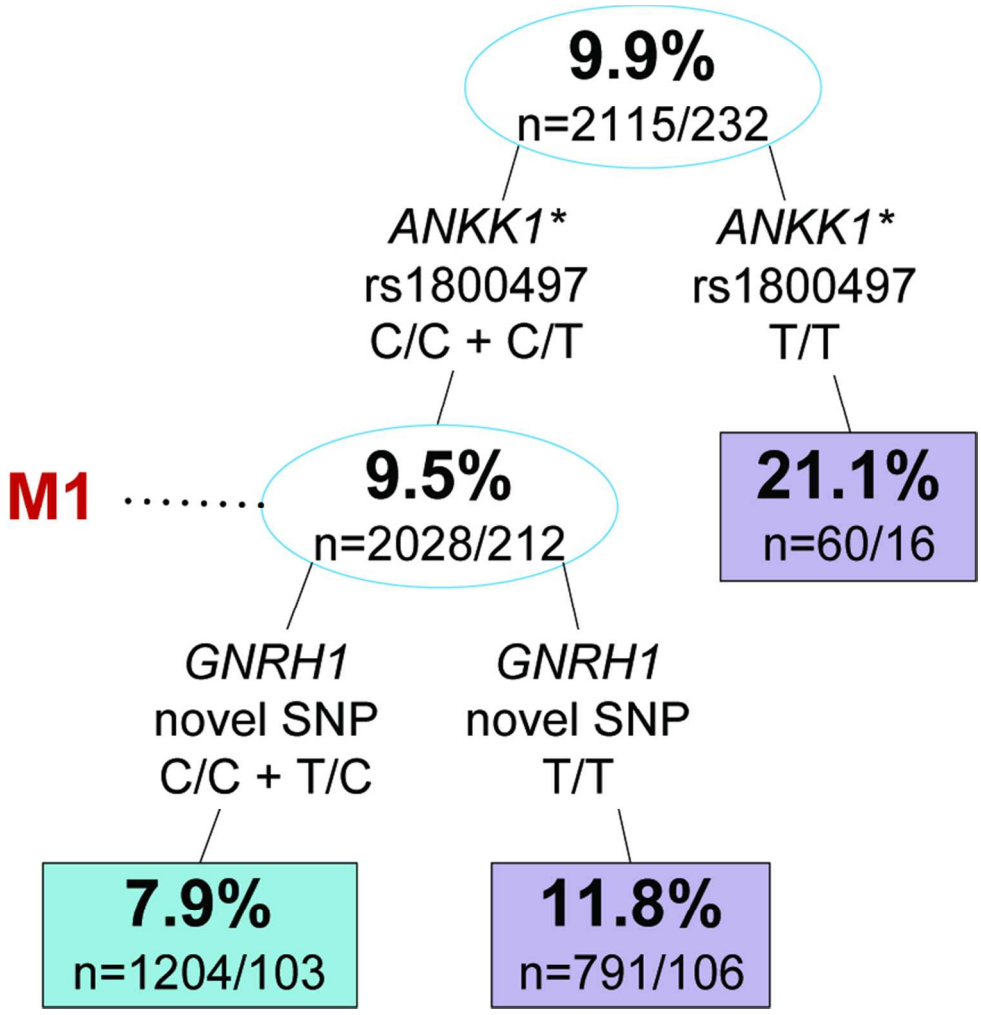
35. Hayden EP, Klein DN, Dougherty LR, Olino TM, Laptook RS, Dyson MW, et al. The dopamine D2 receptor gene and depressive and anxious symptoms in childhood: associations and evidence for gene-environment correlation and gene-environment interaction. Psychiatr Genet. 2010 Dec;20(6):304-10.

36. Dubertret C, Bardel C, Ramoz N, Martin P-M, Deybach J-C, Adès J, et al. A genetic schizophrenia-susceptibility region located between the ANKK1 and DRD2 genes. Progress in Neuro-Psychopharmacology and Biological Psychiatry. 2010;34(3):492-9.

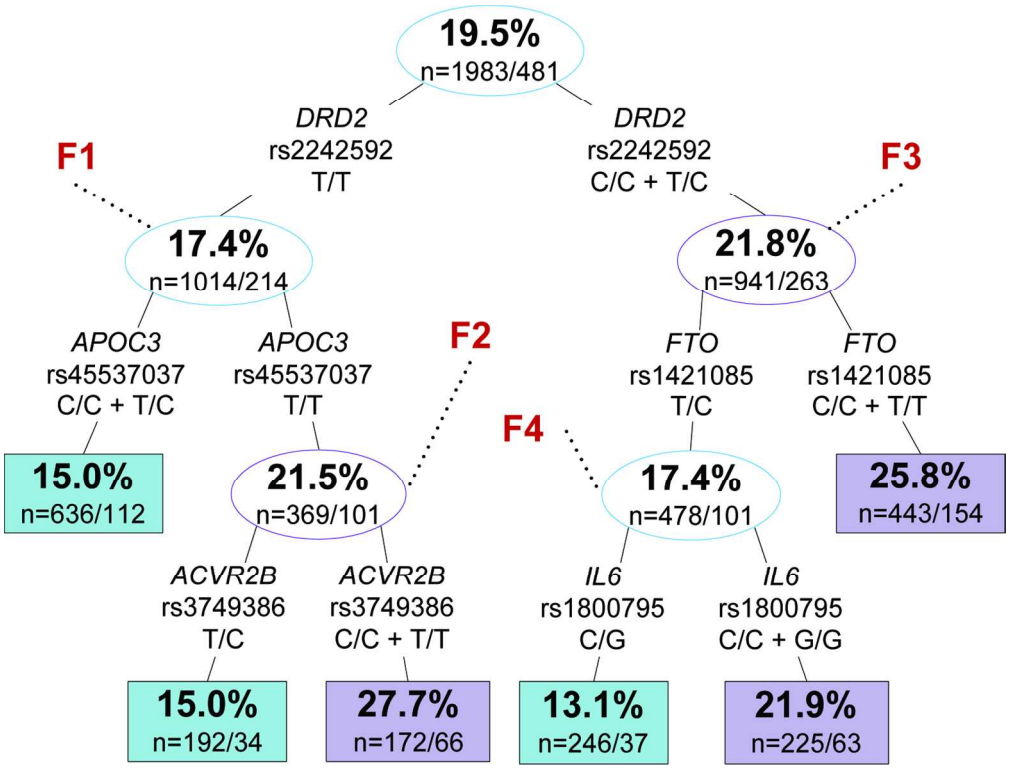
37. Kivimaki M, Jokela M, Hamer M, Geddes J, Ebmeier K, Kumari M, et al. Examining Overweight and Obesity as Risk Factors for Common Mental Disorders Using Fat Mass and Obesity-Associated (FTO) Genotype-Instrumented Analysis. Am J Epidemiol. 2011 Feb 15;173(4):421-9.

38. Bull SJ, Huezo-Diaz P, Binder EB, Cubells JF, Ranjith G, Maddock C, et al. Functional polymorphisms in the interleukin-6 and serotonin transporter genes, and depression and fatigue induced by interferon-alpha and ribavirin treatment. Mol Psychiatry. 2009 Dec;14(12):1095-104.

39. Zheng F, Adelsberger H, Muller MR, Fritschy JM, Werner S, Alzheimer C. Activin tunes GABAergic neurotransmission and modulates anxiety-like behavior. Mol Psychiatry. 2008;14(3):332-46.



74x74mm (300 x 300 DPI)



134x102mm (300 x 300 DPI)

Supplementary Table 1. Single Nucleotide Polymorphisms Assessed in the WLS

Gene	Encodes	SNP	Associated disease/behavior
A2M	alpha-2-macroglobulin	rs669	Alzheimer's disease (1)
ACVR2A	activin receptor IIA	rs1424954	pre-eclampsia (2)
ACVR2B	activin receptor IIB	rs3749386	--
ADIPOQ	adiponectin, C1Q and collagen domain containing	rs1501299	diabetes II (3, 4), obesity (5, 6), breast cancer (7)
ADIPOQ	adiponectin, C1Q and collagen domain containing	rs2241766	diabetes II (3, 4), obesity (8), breast cancer (7)
ACVRL1	activin receptor-like kinase 1	rs2071219	brain arteriovenous malformations (9)
APOC-3	apolipoprotein C-III	rs2854116	nonalcoholic fatty liver disease (10)
ApoE	apolipoprotein E	rs429358	Alzheimer's disease (11, 12)
ApoE	apolipoprotein E	rs7412	Alzheimer's disease (11, 12)
AR	androgen receptor	rs6152	male pattern baldness (13)
BCKDHB	branched chain keto acid dehydrogenase E1, beta polypeptide	rs4502885	premature ovarian failure (14)
BDNF	brain-derived neurotrophic factor	rs6265	depression (15-17), alcohol dependence-related depression (18), bipolar disorder (19), schizophrenia (20), cognition (21), BMI (22)
BDNF	brain-derived neurotrophic factor	rs908867	antidepressant response (23)
BRCA1	breast cancer 1, early onset	rs1799966	breast cancer (24)
BRCA2	breast cancer 2, early onset homolog	rs144848	breast cancer (24)
CH25H	cholesterol 25-hydroxylase	rs3802657	--
CHRM2	cholinergic receptor, muscarinic 2	rs2061174	alcohol dependence, depression (25)
CHRM2	cholinergic receptor, muscarinic 2	rs8191992	cognition (26)
COMT	catechol-O-methyltransferase	rs4680	ADHD (27), substance abuse (28-31), depression (32), antidepressant response (33), bipolar disorder (34), cognition (35)
CTSD	cathepsin D	rs17571	Alzheimer's disease (36)
CYP11A1	cytochrome P450, family 11, subfamily A, polypeptide 1	rs8039957	breast cancer (37)
CYP11B2	cytochrome P450, family 11, subfamily B, polypeptide 2	rs1799998	stroke (38), cardiovascular disease (39)
DAT1	human dopamine transporter	rs11564774	ADHD (40)
DAT1	human dopamine transporter	rs2963238	alcohol-withdrawal seizures (41)
DISC1	disrupted in schizophrenia 1	rs821616	schizophrenia (42), cognitive aging (43)
DRD2	dopamine receptor D2	rs17529477	--
DRD2/ANKK1	dopamine receptor D2/ ankyrin repeat and kinase domain containing 1	rs1800497	obesity, drug addiction (44)
DRD2	dopamine receptor D2	rs2242592	schizophrenia (45)
DRD2	dopamine receptor D2	rs4245147	--
DRD2	dopamine receptor D2	rs6277	schizophrenia (46), PTSD (47)
DRD4	dopamine receptor D4	rs1800955	ADHD (48), heroine addiction (49)
DTNBP1	dystrobrevin-binding protein 1	rs1018381	schizophrenia (50), cognitive ability (51)
DTNBP1	dystrobrevin-binding protein 1	rs760761	schizophrenia (52)
ESR1	estrogen receptor 1	rs7761133	--
ESR1	estrogen receptor 1	rs3853248	--
FADS2	fatty acid desaturase 2	rs1535	breastfeeding & IQ (53)

FADS2	fatty acid desaturase 2	rs174575	breastfeeding & IQ (53)
FMR1	fragile X mental retardation 1	rs1805420	--
FSH	follicle stimulating hormone	rs6169	--
FSHR	follicle stimulating hormone receptor	rs6166	sterility (54), osteoporosis (55)
FST	follistatin	rs12152850	--
FST	follistatin	rs3797297	--
FTO	fat mass and obesity associated	rs1421085	obesity (56-58), mental disorders (59)
GABBR2	γ-aminobutyric acid B receptor 2	rs1435252	nicotine addiction (60)
GABBR2	γ-aminobutyric acid B receptor 2	rs2779562	nicotine addiction (60)
GNRH1	gonadotropin-releasing hormone	novel SNP	Alzheimer's disease (61)
HERC	hect domain and RLD 2	rs12913832	eye color (62, 63)
HFE	hemochromatosis	rs1799945	hemochromatosis(64)
HSD17B1	estradiol 17β-dehydrogenase 1	rs12602084	steroid metabolism (65)
HSD17B1	estradiol 17β-dehydrogenase 1	rs592389	vasomotor symptoms (66), cognition (67)
5-HTR1A	5-hydroxytryptamine (serotonin) receptor 1A	rs878567	mood disorders (68)
5-HTR2A	5-hydroxytryptamine (serotonin) receptor 2A	rs6312	--
5-HTR2A	5-hydroxytryptamine (serotonin) receptor 2A	rs6314	antidepressant response (69), bipolar disorder (70)
5-HTR2A	5-hydroxytryptamine (serotonin) receptor 2A	rs7997012	antidepressant response (71)
5-HTR2C	5-hydroxytryptamine (serotonin) receptor 2C	rs6318	bipolar disorder (72), depression (73)
5-HTT	5-hydroxytryptamine transporter	rs25533	antidepressant response (74)
5-HTT	5-hydroxytryptamine transporter	rs8076005	depressive symptoms (75)
IGF1	insulin-like growth factor 1	rs12313279	--
IL1A	interleukin 1, alpha	rs17561	chronic rhinosinusitis (76), BMI (77)
IL6	interleukin 6	rs1800795	arthritis (78), breast cancer (79), diabetes (80), depression (81)
INHA	inhibin alpha	rs2059693	testicular cancer (82)
INHA	inhibin alpha	rs35118453	--
INHBA	inhibin beta A	rs2237436	--
INHBB	inhibin beta B	rs11902591	--
KIBRA	kidney and brain protein (WWC1)	rs17070145	Alzheimer's disease (83), episodic memory (84)
LEPR	leptin receptor	rs1137100	diabetes II (85), atherosclerosis (86)
LHR	luteinizing hormone receptor	rs4073366	Alzheimer's disease (87)
MAOA	monoamine oxidase A	rs3788862	pain (88)
OXTR	oxytocin receptor	rs2254298	autism (89, 90), social loneliness (91), depressive symptoms & anxiety (92)
PCK1	phosphoenolpyruvate carboxykinase 1	rs707555	diabetes II (93)
PGR	progesterone receptor	rs1042838	ovarian cancer (94), migraine (95), menstruation (96), pregnancy loss (97)
SNAP25	synaptosomal-associated protein 25	rs363050	intelligence (98, 99)
SSADH	succinic semialdehyde dehydrogenase	rs2760118	--
StAR	steroidogenic acute regulatory protein	rs3990403	--
TFAM	transcription factor A, mitochondrial	rs1937	Alzheimer's disease (100)
TFAM	transcription factor A, mitochondrial	rs2306604	Parkinson's disease (101)
TPH1	first tryptophan hydroxylase isoform	rs1799913	heroine addiction (102)

TPH2	first tryptophan hydroxylase isoform	rs11178997	bipolar disorder (103), PTSD symptoms (104)

REFERENCES

1. Zappia M, Manna I, Serra P, Cittadella R, Andreoli V, La Russa A, et al. Increased risk for Alzheimer disease with the interaction of MPO and A2M polymorphisms. Arch Neurol. 2004 Mar;61(3):341-4.
2. Moses EK, Fitzpatrick E, Freed KA, Dyer TD, Forrest S, Elliott K, et al. Objective prioritization of positional candidate genes at a quantitative trait locus for pre-eclampsia on 2q22. Molecular Human Reproduction. 2006 August 2006;12(8):505-12.
3. Hara K, Boutin P, Mori Y, Tobe K, Dina C, Yasuda K, et al. Genetic Variation in the Gene Encoding Adiponectin Is Associated With an Increased Risk of Type 2 Diabetes in the Japanese Population. Diabetes. 2002 February 1, 2002;51(2):536-40.
4. Gu HF, Abulaiti A, Ostenson CG, Humphreys K, Wahlestedt C, Brookes AJ, et al. Single nucleotide polymorphisms in the proximal promoter region of the adiponectin (APM1) gene are associated with type 2 diabetes in Swedish caucasians. Diabetes. 2004 Feb;53 Suppl 1:S31-5.
5. Bouatia-Naji N, Meyre D, Lobbens S, Seron K, Fumeron F, Balkau B, et al. ACDC/adiponectin polymorphisms are associated with severe childhood and adult obesity. Diabetes. 2006 Feb;55(2):545-50.
6. Yang W-S, Yang Y-C, Chen C-L, Wu I-L, Lu J-Y, Lu F-H, et al. Adiponectin SNP276 is associated with obesity, the metabolic syndrome, and diabetes in the elderly. The American Journal of Clinical Nutrition. 2007 August 1, 2007;86(2):509-13.
7. Kaklamani VG, Sadim M, Hsi A, Offit K, Oddoux C, Ostrer H, et al. Variants of the adiponectin and adiponectin receptor 1 genes and breast cancer risk. Cancer Res. 2008 May 1;68(9):3178-84.
8. Beckers S, Peeters AV, De Freitas F, Mertens IL, Verhulst SL, Haentjens D, et al. Association Study and Mutation Analysis of Adiponectin Shows Association of Variants in APM1 with Complex Obesity in Women. Annals of Human Genetics. 2009;73(5):492-501.
9. Young WL, Kwok P-Y, Pawlikowska L, Lawton MT, Kim H, Hysi PG, et al. Arteriovenous Malformation. Journal of Neurosurgery. 2007;106(4):731-2.
10. Petersen KF, Dufour S, Hariri A, Nelson-Williams C, Foo JN, Zhang X-M, et al. Apolipoprotein C3 Gene Variants in Nonalcoholic Fatty Liver Disease. New England Journal of Medicine. 2010;362(12):1082-9.
11. Corder EH, Saunders AM, Strittmatter WJ, Schmechel DE, Gaskell PC, Small GW, et al. Gene dose of apolipoprotein E type 4 allele and the risk of Alzheimer's disease in late onset families. Science. 1993;261(5123):921-3.
12. Saunders AM, Strittmatter WJ, Schmechel D, George-Hyslop PH, Pericak-Vance MA, Joo SH, et al. Association of apolipoprotein E allele epsilon 4 with late-onset familial and sporadic Alzheimer's disease. Neurology. 1993;43(8):1467-72.
13. Ellis JA, Stebbing M, Harrap SB. Polymorphism of the Androgen Receptor Gene is Associated with Male Pattern Baldness. 2001;116(3):452-5.
14. Kang H, Lee SK, Cho S-W, Lee S-H, Kwack K. Branched chain [alpha]-keto acid dehydrogenase, E1-[beta] subunit gene is associated with premature ovarian failure. Fertility and Sterility. 2008;89(3):728-31.

15. Hwang J-P, Tsai S-J, Hong C-J, Yang C-H, Lirng J-F, Yang Y-M. The Val66Met polymorphism of the brain-derived neurotrophic-factor gene is associated with geriatric depression. Neurobiology of Aging. 2006;27(12):1834-7.
16. Czira ME, Wersching H, Baune BT, Berger K. Brain-derived neurotrophic factor gene polymorphisms, neurotransmitter levels, and depressive symptoms in an elderly population. Age (Dordr). 2011 Sep 7.
17. Suchanek R, Owczarek A, Kowalczyk M, Kucia K, Kowalski J. Association between C-281A and val66met functional polymorphisms of BDNF gene and risk of recurrent major depressive disorder in Polish population. J Mol Neurosci. 2011 Mar;43(3):524-30.
18. Su N, Zhang L, Fei F, Hu H, Wang K, Hui H, et al. The brain-derived neurotrophic factor is associated with alcohol dependence-related depression and antidepressant response. Brain Research. 2011;1415(0):119-26.
19. Sears C, Markie D, Olds R, Fitches A. Evidence of associations between bipolar disorder and the brain-derived neurotrophic factor (BDNF) gene. Bipolar Disorders. 2011;13(7-8):630-7.
20. Zakharyan R, Boyajyan A, Arakelyan A, Gevorgyan A, Mrazek F, Petrek M. Functional variants of the genes involved in neurodevelopment and susceptibility to schizophrenia in an Armenian population. Human Immunology. 2011;72(9):746-8.
21. McAllister TW, Tyler AL, Flashman LA, Rhodes CH, McDonald BC, Saykin AJ, et al. Polymorphisms in the brain-derived neurotrophic factor gene influence memory and processing speed one month after brain injury. J Neurotrauma. 2012 Apr 10;29(6):1111-8.
22. Hong K-W, Lim J-E, Go MJ, Shin Cho Y, Ahn Y, Han B-G, et al. Recapitulation of the Association of the Val66Met Polymorphism of BDNF Gene With BMI in Koreans. Obesity. 2011.
23. Gratacos M, Soria V, Urretavizcaya M, Gonzalez JR, Crespo JM, Bayes M, et al. A brain-derived neurotrophic factor (BDNF) haplotype is associated with antidepressant treatment outcome in mood disorders. Pharmacogenomics J. 2007;8(2):101-12.
24. Johnson N, Fletcher O, Palles C, Rudd M, Webb E, Sellick G, et al. Counting potentially functional variants in BRCA1, BRCA2 and ATM predicts breast cancer susceptibility. Human Molecular Genetics. 2007 May 1, 2007;16(9):1051-7.
25. Wang JC, Hinrichs AL, Stock H, Budde J, Allen R, Bertelsen S, et al. Evidence of common and specific genetic effects: association of the muscarinic acetylcholine receptor M2 (CHRM2) gene with alcohol dependence and major depressive syndrome. Human Molecular Genetics. 2004 September 1, 2004;13(17):1903-11.
26. Jones KA, Porjesz B, Almasy L, Bierut L, Goate A, Wang JC, et al. Linkage and linkage disequilibrium of evoked EEG oscillations with CHRM2 receptor gene polymorphisms: implications for human brain dynamics and cognition. International Journal of Psychophysiology. 2004;53(2):75-90.
27. Eisenberg J, Mei-Tal G, Steinberg A, Tartakovsky E, Zohar A, Gritsenko I, et al. Haplotype relative risk study of catechol-O-methyltransferase (COMT) and attention deficit hyperactivity disorder (ADHD): Association of the high-enzyme activity val allele with adhd impulsive-hyperactive phenotype. American Journal of Medical Genetics. 1999;88(5):497-502.
28. Wang T, Franke P, Neidt H, Cichon S, Knapp M, Lichtermann D, et al. Association study of the low-activity allele of catechol-O-methyltransferase and alcoholism using a family-based approach. Mol Psychiatry. 2001 Jan;6(1):109-11.
29. Horowitz R, Kotler M, Shufman E, Aharoni S, Kremer I, Cohen H, et al. Confirmation of an excess of the high enzyme activity COMT val allele in heroin addicts in a family-based haplotype relative risk study. American Journal of Medical Genetics. 2000;96(5):599-603.

30. Vandenbergh DJ, Rodriguez LA, Miller IT, Uhl GR, Lachman HM. High-activity catechol-O-methyltransferase allele is more prevalent in polysubstance abusers. American Journal of Medical Genetics. 1997;74(4):439-42.
31. Li T, Chen C-k, Hu X, Ball D, Lin S-K, Chen W, et al. Association analysis of the DRD4 and COMT genes in methamphetamine abuse. American Journal of Medical Genetics Part B: Neuropsychiatric Genetics. 2004;129B(1):120-4.
32. Åberg E, Fandiño-Losada A, Sjöholm LK, Forsell Y, Lavebratt C. The functional Val158Met polymorphism in catechol-O-methyltransferase (COMT) is associated with depression and motivation in men from a Swedish population-based study. Journal of Affective Disorders. 2011;129(1–3):158-66.
33. Baune BT, Hohoff C, Berger K, Neumann A, Mortensen S, Roehrs T, et al. Association of the COMT val158met Variant with Antidepressant Treatment Response in Major Depression. Neuropsychopharmacology. 2007;33(4):924-32.
34. Lee S-Y, Chen S-L, Chen S-H, Huang S-Y, Tzeng N-S, Chang Y-H, et al. The COMT and DRD3 genes interacted in bipolar I but not bipolar II disorder. World Journal of Biological Psychiatry. 2011;12(5):385-91.
35. Dickinson D, Elvevåg B. Genes, cognition and brain through a COMT lens. Neuroscience. 2009;164(1):72-87.
36. Albayrak Ö, Tirniceriu A, Riemenschneider M, Kurz A, Scherag A, Egensperger R. The Cathepsin D (224C/T) Polymorphism Confers an Increased Risk to Develop Alzheimer's Disease in Men. The Journals of Gerontology Series A: Biological Sciences and Medical Sciences. 2010 March 1, 2010;65A(3):219-24.
37. Yaspan BL, Breyer JP, Cai Q, Dai Q, Elmore JB, Amundson I, et al. Haplotype Analysis of CYP11A1 Identifies Promoter Variants Associated with Breast Cancer Risk. Cancer Research. 2007 June 15, 2007;67(12):5673-82.
38. Tu Y, Cui G, Xu Y, Bao X, Wang X, Wang D. Genetic polymorphism of CYP11B2 gene and stroke in the Han Chinese population and a meta-analysis. Pharmacogenet Genomics. 2011;21(3):115-20.
39. White PC, Hautanen A, Kupari M. Aldosterone synthase (CYP11B2) polymorphisms and cardiovascular function. J Steroid Biochem Mol Biol. 1999;69(1-6):409-12.
40. Brookes K, Xu X, Chen W, Zhou K, Neale B, Lowe N, et al. The analysis of 51 genes in DSM-IV combined type attention deficit hyperactivity disorder: association signals in DRD4, DAT1 and 16 other genes. Mol Psychiatry. 2006;11(10):934-53.
41. Le Strat Y, Ramoz N, Pickering P, Burger V, Boni C, Aubin H-J, et al. The 3' Part of the Dopamine Transporter Gene DAT1/SLC6A3 Is Associated With Withdrawal Seizures in Patients With Alcohol Dependence. Alcoholism: Clinical and Experimental Research. 2008;32(1):27-35.
42. Callicott JH, Straub RE, Pezawas L, Egan MF, Mattay VS, Hariri AR, et al. Variation in DISC1 affects hippocampal structure and function and increases risk for schizophrenia. Proc Natl Acad Sci U S A. 2005;102(24):8627-32.
43. Thomson PA, Harris SE, Starr JM, Whalley LJ, Porteous DJ, Deary IJ. Association between genotype at an exonic SNP in DISC1 and normal cognitive aging. Neuroscience Letters. 2005;389(1):41-5.
44. Blum K, Braverman ER, Wood RC, Gill J, Li C, Chen TJ, et al. Increased prevalence of the Taq I A1 allele of the dopamine receptor gene (DRD2) in obesity with comorbid substance use disorder: a preliminary report. Pharmacogenetics. 1996 Aug;6(4):297-305.
45. Dubertret C, Bardel C, Ramoz N, Martin P-M, Deybach J-C, Adès J, et al. A genetic schizophrenia-susceptibility region located between the ANKK1 and DRD2 genes. Progress in Neuro-Psychopharmacology and Biological Psychiatry. 2010;34(3):492-9.
46. Monakhov M, Golimbet V, Abramova L, Kaleda V, Karpov V. Association study of three polymorphisms in the dopamine D2 receptor gene and schizophrenia in the Russian population. Schizophrenia Research. 2008;100(1–3):302-7.

47. Voisey J, Swagell CD, Hughes IP, Morris CP, van Daal A, Noble EP, et al. The DRD2 gene 957C>T polymorphism is associated with Posttraumatic Stress Disorder in war veterans. Depression and Anxiety. 2009;26(1):28-33.
48. Yang J-W, Jang W-S, Hong SD, Ji YI, Kim DH, Park J, et al. A case-control association study of the polymorphism at the promoter region of the DRD4 gene in Korean boys with attention deficit-hyperactivity disorder: Evidence of association with the - 521 C/T SNP. Progress in Neuro-Psychopharmacology and Biological Psychiatry. 2008;32(1):243-8.
49. Lai JH, Zhu YS, Huo ZH, Sun RF, Yu B, Wang YP, et al. Association study of polymorphisms in the promoter region of DRD4 with schizophrenia, depression, and heroin addiction. Brain Research. 2010;1359:227-32.
50. Funke B, Finn CT, Plocik AM, Lake S, DeRosse P, Kane JM, et al. Association of the DTNBP1 Locus with Schizophrenia in a U.S. Population. The American Journal of Human Genetics. 2004;75(5):891-8.
51. Burdick KE, Lencz T, Funke B, Finn CT, Szeszko PR, Kane JM, et al. Genetic variation in DTNBP1 influences general cognitive ability. Human Molecular Genetics. 2006 15 May 2006;15(10):1563-8.
52. Riley B, Kuo P-H, Maher BS, Fanous AH, Sun J, Wormley B, et al. The dystrobrevin binding protein 1 (DTNBP1) gene is associated with schizophrenia in the Irish Case Control Study of Schizophrenia (ICCSS) sample. Schizophrenia Research. 2009;115(2–3):245-53.
53. Caspi A, Williams B, Kim-Cohen J, Craig IW, Milne BJ, Poulton R, et al. Moderation of breastfeeding effects on the IQ by genetic variation in fatty acid metabolism. Proc Natl Acad Sci U S A. 2007 Nov 20;104(47):18860-5.
54. Binder H, Dittrich R, Hager I, Muller A, Oeser S, Beckmann MW, et al. Association of FSH receptor and CYP19A1 gene variations with sterility and ovarian hyperstimulation syndrome. Reproduction. 2008;135(1):107-16.
55. Rendina D, Gianfrancesco F, De Filippo G, Merlotti D, Esposito T, Mingione A, et al. FSHR gene polymorphisms influence bone mineral density and bone turnover in postmenopausal women. Eur J Endocrinol. 2010;163(1):165-72.
56. Dina C, Meyre D, Gallina S, Durand E, Korner A, Jacobson P, et al. Variation in FTO contributes to childhood obesity and severe adult obesity. Nat Genet. 2007;39(6):724-6.
57. Tung Y-CL, Yeo GSH. From GWAS to biology: lessons from FTO. Annals of the New York Academy of Sciences. 2011;1220(1):162-71.
58. Chabris C, Hebert B, Benjamin D, Beauchamp J, Cesarini D, van der Loos M, et al. Most reported genetic assocations with general intelligence are probably false positives. Psychological Science. 2011;In press.
59. Kivimaki M, Jokela M, Hamer M, Geddes J, Ebmeier K, Kumari M, et al. Examining Overweight and Obesity as Risk Factors for Common Mental Disorders Using Fat Mass and Obesity-Associated (FTO) Genotype-Instrumented Analysis. Am J Epidemiol. 2011 Feb 15;173(4):421-9.
60. Beuten J, Ma JZ, Payne TJ, Dupont RT, Crews KM, Somes G, et al. Single- and Multilocus Allelic Variants within the GABAB Receptor Subunit 2 (GABAB2) Gene Are Significantly Associated with Nicotine Dependence. The American Journal of Human Genetics. 2005;76(5):859-64.
61. Wilson AC. Investigations into gonadotropin-releasing hormone receptor signaling in extrahypothalamic neurons of the brain: Implications for normal cognition and the pathogenesis of Alzheimer's disease: Thesis. University of Wisconsin, Madison, USA; 2009.
62. Eiberg H, Troelsen J, Nielsen M, Mikkelsen A, Mengel-From J, Kjaer K, et al. Blue eye color in humans may be caused by a perfectly associated founder mutation in a regulatory element located within the HERC2 gene inhibiting OCA2 expression. Human Genetics. 2008;123(2):177-87.

63. Sturm RA, Duffy DL, Zhao ZZ, Leite FPN, Stark MS, Hayward Nicholas K, et al. A Single SNP in an Evolutionary Conserved Region within Intron 86 of the HERC2 Gene Determines Human Blue-Brown Eye Color. The American Journal of Human Genetics. 2008;82(2):424-31.
64. Beutler E. The significance of the 187G (H63D) mutation in hemochromatosis. Am J Hum Genet. 1997 Sep;61(3):762-4.
65. Sun T, Oh WK, Jacobus S, Regan M, Pomerantz M, Freedman ML, et al. The Impact of Common Genetic Variations in Genes of the Sex Hormone Metabolic Pathways on Steroid Hormone Levels and Prostate Cancer Aggressiveness. Cancer Prevention Research. 2011 December 1, 2011;4(12):2044-50.
66. Crandall CJ, Crawford SL, Gold EB. Vasomotor Symptom Prevalence Is Associated with Polymorphisms in Sex Steroid-Metabolizing Enzymes and Receptors. The American Journal of Medicine. 2006;119(9, Supplement 1):S52-S60.
67. Kravitz HM, Meyer PM, Seeman TE, Greendale GA, Sowers MR. Cognitive Functioning and Sex Steroid Hormone Gene Polymorphisms in Women at Midlife. The American Journal of Medicine. 2006;119(9, Supplement 1):S94-S102.
68. Brezo J, Bureau A, Merette C, Jomphe V, Barker ED, Vitaro F, et al. Differences and similarities in the serotonergic diathesis for suicide attempts and mood disorders: a 22-year longitudinal gene-environment study. Mol Psychiatry. 2010;15(8):831-43.
69. Wilkie MJV, Smith G, Day RK, Matthews K, Smith D, Blackwood D, et al. Polymorphisms in the SLC6A4 and HTR2A genes influence treatment outcome following antidepressant therapy. Pharmacogenomics J. 2008;9(1):61-70.
70. McAuley EZ, Fullerton JM, Blair IP, Donald JA, Mitchell PB, Schofield PR. Association between the serotonin 2A receptor gene and bipolar affective disorder in an Australian cohort. Psychiatr Genet. 2009 Oct;19(5):244-52.
71. McMahon FJ, Buervenich S, Charney D, Lipsky R, Rush AJ, Wilson AF, et al. Variation in the Gene Encoding the Serotonin 2A Receptor Is Associated with Outcome of Antidepressant Treatment. The American Journal of Human Genetics. 2006;78(5):804-14.
72. Mazza M, Mandelli L, Martinotti G, Di Nicola M, Tavian D, Negri G, et al. Further evidence supporting the association between 5HTR2C gene and bipolar disorder. Psychiatry Res. 2010 Dec 30;180(2-3):151-2.
73. Lerer B, Macciardi F, Segman RH, Adolfsson R, Blackwood D, Blairy S, et al. Variability of 5-HT2C receptor cys23ser polymorphism among European populations and vulnerability to affective disorder. Mol Psychiatry. 2001 Sep;6(5):579-85.
74. Lin E, Chen PS, Chang HH, Gean P-W, Tsai HC, Yang YK, et al. Interaction of serotonin-related genes affects short-term antidepressant response in major depressive disorder. Progress in Neuro-Psychopharmacology and Biological Psychiatry. 2009;33(7):1167-72.
75. Su S, Zhao J, Bremner JD, Miller AH, Tang W, Bouzyk M, et al. Serotonin Transporter Gene, Depressive Symptoms, and Interleukin-6 / CLINICAL PERSPECTIVE. Circulation: Cardiovascular Genetics. 2009 December 1, 2009;2(6):614-20.
76. Mfuna Endam L, Cormier C, Bosse Y, Filali-Mouhim A, Desrosiers M. Association of IL1A, IL1B, and TNF Gene Polymorphisms With Chronic Rhinosinusitis With and Without Nasal Polyposis: A Replication Study. Arch Otolaryngol Head Neck Surg. 2010 February 1, 2010;136(2):187-92.
77. Um J-Y, Rim H-K, Kim S-J, Kim H-L, Hong S-H. Functional Polymorphism of IL-1 Alpha and Its Potential Role in Obesity in Humans and Mice. PLoS ONE. 2011;6(12):e29524.
78. Fishman D, Faulds G, Jeffery R, Mohamed-Ali V, Yudkin JS, Humphries S, et al. The effect of novel polymorphisms in the interleukin-6 (IL-6) gene on IL-6 transcription and plasma IL-6 levels, and an association with systemic-onset juvenile chronic arthritis. The Journal of Clinical Investigation. 1998;102(7):1369-76.

79. Slattery ML, Curtin K, Sweeney C, Wolff RK, Baumgartner RN, Baumgartner KB, et al. Modifying Effects of IL-6 Polymorphisms on Body Size-Associated Breast Cancer Risk. Obesity. 2008;16(2):339-47.
80. Huth C, Heid IM, Vollmert C, Gieger C, Grallert H, Wolford JK, et al. IL6 Gene Promoter Polymorphisms and Type 2 Diabetes. Diabetes. 2006 October 2006;55(10):2915-21.
81. Bull SJ, Huezo-Diaz P, Binder EB, Cubells JF, Ranjith G, Maddock C, et al. Functional polymorphisms in the interleukin-6 and serotonin transporter genes, and depression and fatigue induced by interferon-alpha and ribavirin treatment. Mol Psychiatry. 2009 Dec;14(12):1095-104.
82. Purdue MP, Graubard BI, Chanock SJ, Rubertone MV, Erickson RL, McGlynn KA. Genetic Variation in the Inhibin Pathway and Risk of Testicular Germ Cell Tumors. Cancer Research. 2008 April 15, 2008;68(8):3043-8.
83. Rodríguez-Rodríguez E, Infante J, Llorca J, Mateo I, Sánchez-Quintana C, García-Gorostiaga I, et al. Age-dependent association of KIBRA genetic variation and Alzheimer's disease risk. Neurobiology of Aging. 2009;30(2):322-4.
84. Almeida OP, Schwab SG, Lautenschlager NT, Morar B, Greenop KR, Flicker L, et al. KIBRA genetic polymorphism influences episodic memory in later life, but does not increase the risk of mild cognitive impairment. Journal of Cellular and Molecular Medicine. 2008;12(5a):1672-6.
85. Salopuro T, Pulkkinen L, Lindstrom J, Eriksson JG, Valle TT, Hamalainen H, et al. Genetic variation in leptin receptor gene is associated with type 2 diabetes and body weight: The Finnish Diabetes Prevention Study. Int J Obes Relat Metab Disord. 2005;29(10):1245-51.
86. Saukko M, Kesaniemi YA, Ukkola O. Leptin receptor Lys109Arg and Gln223Arg polymorphisms are associated with early atherosclerosis. Metab Syndr Relat Disord. 2010;8(5):425-30.
87. Haasl R, Ahmadi MR, Meethal S, Gleason C, Johnson S, Asthana S, et al. A luteinizing hormone receptor intronic variant is significantly associated with decreased risk of Alzheimer's disease in males carrying an apolipoprotein E epsilon4 allele. BMC Medical Genetics. 2008;9(1):37.
88. Kim H, Lee H, Rowan J, Brahim J, Dionne R. Genetic polymorphisms in monoamine neurotransmitter systems show only weak association with acute post-surgical pain in humans. Molecular Pain. 2006;2(1):24.
89. Lerer E, Levi S, Salomon S, Darvasi A, Yirmiya N, Ebstein RP. Association between the oxytocin receptor (OXTR) gene and autism: relationship to Vineland Adaptive Behavior Scales and cognition. Mol Psychiatry. 2007;13(10):980-8.
90. Liu X, Kawamura Y, Shimada T, Otowa T, Koishi S, Sugiyama T, et al. Association of the oxytocin receptor (OXTR) gene polymorphisms with autism spectrum disorder (ASD) in the Japanese population. J Hum Genet. 2010;55(3):137-41.
91. Lucht MJ, Barnow S, Sonnenfeld C, Rosenberger A, Grabe HJ, Schroeder W, et al. Associations between the oxytocin receptor gene (OXTR) and affect, loneliness and intelligence in normal subjects. Progress in Neuro-Psychopharmacology and Biological Psychiatry. 2009;33(5):860-6.
92. Thompson RJ, Parker KJ, Hallmayer JF, Waugh CE, Gotlib IH. Oxytocin receptor gene polymorphism (rs2254298) interacts with familial risk for psychopathology to predict symptoms of depression and anxiety in adolescent girls. Psychoneuroendocrinology. 2011;36(1):144-7.
93. Dong Y, Zhang H, Wang X, Feng Q, Chen X, Su Q. A Leu184Val polymorphism in PCK1 gene is associated with type 2 diabetes in Eastern Chinese population with BMI < 23 kg/m2. Diabetes Research and Clinical Practice. 2009;83(2):227-32.
94. Pearce CL, Hirschhorn JN, Wu AH, Burtt NP, Stram DO, Young S, et al. Clarifying the PROGINS Allele Association in Ovarian and Breast Cancer Risk: A Haplotype-Based Analysis. Journal of the National Cancer Institute. 2005 January 5, 2005;97(1):51-9.

95. Joshi G, Pradhan S, Mittal B. Role of the oestrogen receptor (ESR1 PvuII and ESR1 325 C→G) and progesterone receptor (PROGINS) polymorphisms in genetic susceptibility to migraine in a North Indian population. Cephalalgia. 2010 March 1, 2010;30(3):311-20.
96. Taylor KC, Small CM, Epstein MP, Sherman SL, Tang W, Wilson MM, et al. Associations of Progesterone Receptor Polymorphisms with Age at Menarche and Menstrual Cycle Length. Hormone Research in Paediatrics. 2010;74(6):421-7.
97. Su MT, Lee IW, Chen YC, Kuo PL. Association of progesterone receptor polymorphism with idiopathic recurrent pregnancy loss in Taiwanese Han population. J Assist Reprod Genet. 2011 Mar;28(3):239-43.
98. Gosso MF, de Geus EJC, van Belzen MJ, Polderman TJC, Heutink P, Boomsma DI, et al. The SNAP-25 gene is associated with cognitive ability: evidence from a family-based study in two independent Dutch cohorts. Mol Psychiatry. 2006;11(9):878-86.
99. Gosso MF, De Geus EJC, Polderman TJC, Boomsma DI, Heutink P, Posthuma D. Common variants underlying cognitive ability: further evidence for association between the SNAP-25 gene and cognition using a family-based study in two independent Dutch cohorts. Genes, Brain and Behavior. 2008;7(3):355-64.
100. Zhang Q, Yu J-T, Wang P, Chen W, Wu Z-C, Jiang H, et al. Mitochondrial transcription factor A (TFAM) polymorphisms and risk of late-onset Alzheimer's disease in Han Chinese. Brain Research. 2011;1368(0):355-60.
101. Gaweda-Walerych K, Safranow K, Maruszak A, Bialecka M, Klodowska-Duda G, Czyzewski K, et al. Mitochondrial transcription factor A variants and the risk of Parkinson's disease. Neuroscience Letters. 2010;469(1):24-9.
102. Nielsen DA, Barral S, Proudnikov D, Kellogg S, Ho A, Ott J, et al. TPH2 and TPH1: association of variants and interactions with heroin addiction. Behav Genet. 2008 Mar;38(2):133-50.
103. Cichon S, Winge I, Mattheisen M, Georgi A, Karpushova A, Freudenberg J, et al. Brain-specific tryptophan hydroxylase 2 (TPH2): a functional Pro206Ser substitution and variation in the 5'-region are associated with bipolar affective disorder. Human Molecular Genetics. 2008 January 1, 2008;17(1):87-97.
104. Goenjian AK, Bailey JN, Walling DP, Steinberg AM, Schmidt D, Dandekar U, et al. Association of TPH1, TPH2, and 5HTTLPR with PTSD and depressive symptoms. Journal of Affective Disorders. 2012(in Press).

Supplementary Table 2. Depression-Associated SNP Identified in the WLS

Gene	Encodes	SNP	Alleles	Chr#/Location	Residue	Associated disease/behavior
ACVR2B	activin receptor IIB	rs3749386	T/C	3/intron 1	--	left-right axis malformations*(1)
APOC3	apolipoprotein C-III	rs2854116	T/C	11/promoter (-455)	--	nonalcoholic fatty liver disease(2)
DRD2/ANKK1	dopamine receptor D2/ankyrin repeat and kinase domain containing 1	rs1800497	C/T	11/exon (ANKK1)	Glu713Lys	obesity, drug addiction (3)
DRD2	dopamine receptor D2	rs2242592	T/C	11/3'	--	schizophrenia (4)
FTO	fat mass and obesity associated	rs1421085	T/C	16/intron 1	--	obesity (5-7), mental disorders (8)
GNRH1	gonadotropin-releasing hormone	novel SNP	T/C	8/promoter	--	Alzheimer's disease (9)
IL6	interleukin 6	rs1800795	C/G	7/promoter (-174)	--	arthritis (10), breast cancer (11), diabetes (12), depression (13)

*Gene association only

References

1. Kosaki R, Gebbia M, Kosaki K, Lewin M, Bowers P, Towbin JA, et al. Left–right axis malformations associated with mutations in ACVR2B, the gene for human activin receptor type IIB. American Journal of Medical Genetics. 1999;82(1):70-6.
2. Petersen KF, Dufour S, Hariri A, Nelson-Williams C, Foo JN, Zhang X-M, et al. Apolipoprotein C3 Gene Variants in Nonalcoholic Fatty Liver Disease. New England Journal of Medicine. 2010;362(12):1082-9.
3. Blum K, Braverman ER, Wood RC, Gill J, Li C, Chen TJ, et al. Increased prevalence of the Taq I A1 allele of the dopamine receptor gene (DRD2) in obesity with comorbid substance use disorder: a preliminary report. Pharmacogenetics. 1996 Aug;6(4):297-305.
4. Dubertret C, Bardel C, Ramoz N, Martin P-M, Deybach J-C, Adès J, et al. A genetic schizophrenia-susceptibility region located between the ANKK1 and DRD2 genes. Progress in Neuro-Psychopharmacology and Biological Psychiatry. 2010;34(3):492-9.
5. Dina C, Meyre D, Gallina S, Durand E, Korner A, Jacobson P, et al. Variation in FTO contributes to childhood obesity and severe adult obesity. Nat Genet. 2007;39(6):724-6.
6. Tung Y-CL, Yeo GSH. From GWAS to biology: lessons from FTO. Annals of the New York Academy of Sciences. 2011;1220(1):162-71.
7. Chabris C, Hebert B, Benjamin D, Beauchamp J, Cesarini D, van der Loos M, et al. Most reported genetic assocations with general intelligence are probably false positives. Psychological Science. 2011;In press.
8. Kivimaki M, Jokela M, Hamer M, Geddes J, Ebmeier K, Kumari M, et al. Examining Overweight and Obesity as Risk Factors for Common Mental Disorders Using Fat Mass and Obesity-Associated (FTO) Genotype-Instrumented Analysis. Am J Epidemiol. 2011 Feb 15;173(4):421-9.
9. Wilson AC. Investigations into gonadotropin-releasing hormone receptor signaling in extrahypothalamic neurons of the brain: Implications for normal cognition and the pathogenesis of Alzheimer's disease: Thesis. University of Wisconsin, Madison, USA; 2009.
10. Fishman D, Faulds G, Jeffery R, Mohamed-Ali V, Yudkin JS, Humphries S, et al. The effect of novel polymorphisms in the interleukin-6 (IL-6) gene on IL-6 transcription and plasma IL-6 levels, and an association with systemic-onset juvenile chronic arthritis. The Journal of Clinical Investigation. 1998;102(7):1369-76.
11. Slattery ML, Curtin K, Sweeney C, Wolff RK, Baumgartner RN, Baumgartner KB, et al. Modifying Effects of IL-6 Polymorphisms on Body Size-Associated Breast Cancer Risk. Obesity. 2008;16(2):339-47.
12. Huth C, Heid IM, Vollmert C, Gieger C, Grallert H, Wolford JK, et al. IL6 Gene Promoter Polymorphisms and Type 2 Diabetes. Diabetes. 2006 October 2006;55(10):2915-21.
13. Bull SJ, Huezo-Diaz P, Binder EB, Cubells JF, Ranjith G, Maddock C, et al. Functional polymorphisms in the interleukin-6 and serotonin transporter genes, and depression and fatigue induced by interferon-alpha and ribavirin treatment. Mol Psychiatry. 2009 Dec;14(12):1095-104.

BMJ open

Multi-Gene Interactions and the Prediction of Depression in the Wisconsin Longitudinal Study

Journal:	*BMJ Open*
Manuscript ID:	bmjopen-2012-000944.R2
Article Type:	Research
Date Submitted by the Author:	18-May-2012
Complete List of Authors:	Roetker, Nicholas; UW-Madison, Sociology Yonker, James; UW-Madison, Sociology Lee, Chee; UW-Madison, Sociology Chang, Vicky; UW-Madison, Sociology Basson, Jacob; UW-Madison, Medicine Roan, Carol; UW-Madison, Sociology Hauser, Taissa; UW-Madison, Sociology Hauser, Robert; UW-Madison, Sociology Atwood, Craig; University of Wisconsin, Medicine
Primary Subject Heading:	Mental health
Secondary Subject Heading:	Epidemiology, Mental health, Neurology, Genetics and genomics, Public health
Keywords:	EPIDEMIOLOGY, MENTAL HEALTH, Neurogenetics < NEUROLOGY, Anxiety disorders < PSYCHIATRY, Depression & mood disorders < PSYCHIATRY

STROBE Statement—Checklist of items that should be included in reports of ***cohort studies***

	Item No	Recommendation
Title and abstract	1	(*a*) Indicate the study's design with a commonly used term in the title or the abstract
		(*b*) Provide in the abstract an informative and balanced summary of what was done and what was found
Introduction		
Background/rationale	2	Explain the scientific background and rationale for the investigation being reported
Objectives	3	State specific objectives, including any prespecified hypotheses
Methods		
Study design	4	Present key elements of study design early in the paper
Setting	5	Describe the setting, locations, and relevant dates, including periods of recruitment, exposure, follow-up, and data collection
Participants	6	(*a*) Give the eligibility criteria, and the sources and methods of selection of participants. Describe methods of follow-up
		(*b*) For matched studies, give matching criteria and number of exposed and unexposed
Variables	7	Clearly define all outcomes, exposures, predictors, potential confounders, and effect modifiers. Give diagnostic criteria, if applicable
Data sources/ measurement	8*	For each variable of interest, give sources of data and details of methods of assessment (measurement). Describe comparability of assessment methods if there is more than one group
Bias	9	Describe any efforts to address potential sources of bias
Study size	10	Explain how the study size was arrived at
Quantitative variables	11	Explain how quantitative variables were handled in the analyses. If applicable, describe which groupings were chosen and why
Statistical methods	12	(*a*) Describe all statistical methods, including those used to control for confounding
		(*b*) Describe any methods used to examine subgroups and interactions
		(*c*) Explain how missing data were addressed
		(*d*) If applicable, explain how loss to follow-up was addressed
		(*e*) Describe any sensitivity analyses
Results		
Participants	13*	(a) Report numbers of individuals at each stage of study—eg numbers potentially eligible, examined for eligibility, confirmed eligible, included in the study, completing follow-up, and analysed
		(b) Give reasons for non-participation at each stage
		(c) Consider use of a flow diagram
Descriptive data	14*	(a) Give characteristics of study participants (eg demographic, clinical, social) and information on exposures and potential confounders
		(b) Indicate number of participants with missing data for each variable of interest
		(c) Summarise follow-up time (eg, average and total amount)
Outcome data	15*	Report numbers of outcome events or summary measures over time
Main results	16	(*a*) Give unadjusted estimates and, if applicable, confounder-adjusted estimates and their precision (eg, 95% confidence interval). Make clear which confounders were adjusted for and why they were included
		(*b*) Report category boundaries when continuous variables were categorized
		(*c*) If relevant, consider translating estimates of relative risk into absolute risk for a meaningful time period

Other analyses	17	Report other analyses done—eg analyses of subgroups and interactions, and sensitivity analyses
Discussion		
Key results	18	Summarise key results with reference to study objectives
Limitations	19	Discuss limitations of the study, taking into account sources of potential bias or imprecision. Discuss both direction and magnitude of any potential bias
Interpretation	20	Give a cautious overall interpretation of results considering objectives, limitations, multiplicity of analyses, results from similar studies, and other relevant evidence
Generalisability	21	Discuss the generalisability (external validity) of the study results
Other information		
Funding	22	Give the source of funding and the role of the funders for the present study and, if applicable, for the original study on which the present article is based

*Give information separately for exposed and unexposed groups.

Note: An Explanation and Elaboration article discusses each checklist item and gives methodological background and published examples of transparent reporting. The STROBE checklist is best used in conjunction with this article (freely available on the Web sites of PLoS Medicine at http://www.plosmedicine.org/, Annals of Internal Medicine at http://www.annals.org/, and Epidemiology at http://www.epidem.com/). Information on the STROBE Initiative is available at http://www.strobe-statement.org.

Multi-Gene Interactions and the Prediction of Depression in the Wisconsin Longitudinal Study

Nicholas S. Roetker[1], James A. Yonker[1], Chee Lee[1], Vicky Chang[1], Jacob Basson[2], Carol L. Roan[1], Taissa S. Hauser[1], Robert M. Hauser[1] and Craig S. Atwood[2,3].

[1]Department of Sociology, University of Wisconsin-Madison, Madison, WI 53706, USA.

[2]Geriatric Research, Education and Clinical Center, Veterans Administration Hospital and Department of Medicine, University of Wisconsin-Madison School of Medicine and Public Health, Madison, WI 53705, USA.

[3]School of Exercise, Biomedical and Health Sciences, Edith Cowan University, Joondalup, 6027 WA, Australia.

Running title: Gene interactions and depression

Address Correspondence and Reprint Requests to:
Craig S. Atwood, Ph.D.
University of Wisconsin-Madison School of Medicine and Public Health
Wm S. Middleton Memorial VA (GRECC 11G)
2500 Overlook Terrace, Madison, WI 53705, USA
Tel. 608 256-1901, Ext. 11664
Fax. 608 280-7291
Email: csa@medicine.wisc.edu

ABSTRACT

Objectives: Single genetic loci offer little predictive power for the identification of depression. This study examined whether an analysis of gene-gene interactions of 78 single nucleotide polymorphisms in genes associated with depression and age-related diseases would identify significant interactions with increased predictive power for depression.

Design: A retrospective cohort study.

Setting: A survey of participants in the Wisconsin Longitudinal Study.

Participants: A total of 4,811 persons (2,464 females and 2,347 males) who provided saliva for genotyping; the group comes from a randomly selected sample of Wisconsin high school graduates from the class of 1957 as well as a randomly selected sibling, almost all of whom are non-Hispanic white.

Primary outcome measure: Depression as determine by the Composite International Diagnostic Interview short-form (CIDI-SF).

Results: Using a classification tree approach (recursive partitioning (RP)) we identified a number of candidate gene-gene interactions associated with depression. The primary SNP splits revealed by RP (*ANKK1* rs1800497 (also known as *DRD2* Taq1A) in men and *DRD2* rs224592 in women) were found to be significant as single factors by logistic regression (LR) after controlling for multiple testing (P=0.001 for both). Without considering interaction effects, only 1 of the 5 subsequent RP splits reached nominal significance in logistic regression (*FTO* rs1421085 in women; P-value=0.008). However, after controlling for gene-gene interactions by running logistic regression on RP-specific subsets, every split became significant and grew larger in magnitude (OR [before]→[after]: Men: *GNRH1* novel SNP: [1.43 → 1.57]; Women: *APOC3* rs2854116: [1.28 → 1.55], *ACVR2B* rs3749386: [1.11 → 2.17], *FTO* rs1421085: [1.32 → 1.65], *IL6* rs1800795: [1.12 → 1.85]).

Conclusions: Our results suggest that examining gene-gene interactions improves the identification of genetic associations predictive of depression. Four of the SNPs identified in these interactions were located in two pathways well-known to impact depression: neurotransmitter (*ANKK1* and *DRD2*) and

neuroendocrine (*GNRH1* and *ACVR2B*) signaling. This study demonstrates the utility of RP analysis as an efficient and powerful exploratory analysis technique for uncovering genetic and molecular pathway interactions associated with disease etiology.

INTRODUCTION

Depression is a widespread mental disorder associated with a host of undesirable health, social, and economic outcomes. One in six Americans is diagnosed with depression in his or her lifetime (1). While many environmental factors—such as socioeconomic status, childhood abuse, and major life events—have important ties with depression, so too does gender and many genetic and epigenetic factors, making the disorder heterogeneous in nature (2). Another major risk factor for depression is age, with depression reaching its highest levels in adults over 80 years of age (3).

It has been demonstrated from twin studies that genetic factors typically account for 40–70% of the risk for developing major depressive disorder (MDD), and adoption studies have confirmed the role of genetic risk factors in the development of MDD (see (4) and references therein). Genetic studies, including recent genome-wide association studies (GWAS), have identified genetic alterations in over 50 genes known to be associated with depression (5). However, individually, the genetic alterations found within these genes (primarily single nucleotide polymorphisms (SNPs)) have little predictive value. There is a similar lack of predictive value from GWAS of other major age-related diseases (6).

Given this lack of predictive power among individual genetic alterations for depression together with the complex nature of aging-related diseases, it would seem prudent to examine epistatic effects on this age-related condition. In this respect, we have previously demonstrated that G x G interactions greatly modulate risk for complex age-related diseases (7, 8). Recent studies of depression also have identified epistatic effects. In particular, associations have been identified between *BDNF* Val66Met (brain-derived neurotrophic factor; rs6265) and *5-HTTLPR* (serotonin transporter linked promoter region (9); *GSK3B* rs6782799 (glycogen synthase kinase 3β), *BDNF* rs7124442 and *BDNF* Val66Met (10); *BDNF* Val66Met and SNPs in *NTRK2* (neurotrophic tyrosine kinase receptor 2; (11)), and *5-HTTLPR* short allele and a chromosome 4 gene (12). The machine learning tool recursive partitioning has recently been used by Wong (13) to assess complex gene-gene interactions in depression. Wong notes that recursive partitioning is useful in that it quickly explores high dimensional data for non-linear effects that are non-

biased and easily interpretable.

The goals of this study were therefore to 1) explore G x G interactions that might better predict the genetic factors involved in the etiology of depression, and 2) to further demonstrate the utility of machine learning algorithms (recursive partitioning) to identify genetic interactions. Using genotypic data from the Wisconsin Longitudinal Study (WLS) we identified associations between dopaminergic genes and depression in men and women, as well as G x G interactions involving neuroendocrine signaling pathways, with increased significance compared with single genetic associations.

METHODS

Study Participants and Surveys

Data were collected from the WLS, a random sample originally comprised of 10,317 men and women who graduated from Wisconsin high schools in 1957. Later in 1977, the WLS began interviewing one randomly selected sibling of each graduate, when possible. The cohort reflects the ancestral makeup of the late-1950s Wisconsin population in that participants are almost entirely non-Hispanic white males and females. . In general, the sample is broadly representative of older white Americans with at least a high school education (14). Further characteristics of the WLS cohort may be found in detail elsewhere (15). Health and psychological well-being phenotypic data was taken from mail and phone surveys given in 2004-2005. Inclusion criteria for depression included any member of the WLS cohort who was depressed according to the Composite International Diagnostic Interview short-form (CIDI-SF). Individuals who answered YES to the question “Have you ever had a time in life lasting two weeks or more when nearly every day you felt sad, blue, depressed, or when you lost interest in most things like work, hobbies, or things you usually liked to do for fun?” and whose depression was not caused by alcohol, drugs, medications, or physical illness were asked further depression symptom questions. Symptom questions asked whether the two week period was accompanied with a) any weight loss, b) trouble sleeping, c) feeling tired, d) feeling bad upon waking, e) losing interest, f) trouble concentrating, or g) thoughts about death. Those answering YES to 3 or more of these symptom questions were classified as having depression (16). Those answering YES to 2 or fewer symptom questions and all those answering NO to the initial stem question were classified as controls.

Genotyping

7,101 participants (4,569 graduates & 2,532 siblings) provided saliva samples in Oragene DNA sample collection kits from which DNA was extracted and genotyped for 78 SNPs that were selected based on their association with depression and age-related conditions and diseases (see Supplementary

Information 1). Genotyping was performed by KBioscience (Hoddesdon, UK) with use of a homogeneous Fluorescent Resonance Energy Transfer technology coupled to competitive allele specific PCR. All SNP genotypes described in our results were in Hardy-Weinberg equilibrium and their frequencies matched those reported in the literature for European samples.

Statistical Analysis

Analyses were limited to the 4,811 pooled graduates and siblings for whom we have depression and genotype information (Note: individuals with more than 10% missing genotype data were not included). The average age among this sample was just under 65 years in 2004. 80% were married, and the average amount of post-high school educational attainment was 2 years. Median household income in 1993 was $56,700.

Recursive Partitioning (RP). RP is a data mining tool for revealing trends that relate a dependent variable (depressed vs. non-depressed) to various predictor variables (SNPs). Zhang and Bonney have shown how RP can be used in genetic association studies to identify disease genes (17). RP helps control for heterogeneity in the population and confounding factors by allowing for the segregation of the sample population according to any condition. Thus, RP is a useful way to handle complex datasets that might confound regression analysis due to the complexity of the relationship between the independent and dependent variables and due to missing information.

RP classification trees (using R package rpart) were used to identify potential interactions among the 78 SNPs in relation to depression. The trees split the data along branches according to criteria determined by the rpart package algorithm, which is originally based off the work of Breiman's classification and regression trees (CART) algorithm (18). Basically, the CART algorithm first considers all depressed and non-depressed subjects pooled together in a heterogeneous root node. Based on considering every possible "yes-no" binary partition that can be made by each independent variable, the single split which

maximizes homogeneity between the two resulting sub-nodes as compared to the root node is made. Each sub-node can then be treated independently as a new root node for all subsequent splits, and the pattern continues until every subject constitutes a terminal node, resulting in a very large and complex tree. A 10-part cross validation procedure seeking to minimize misclassification and complexity determines optimal pruning. See Therneau and Atkinson (19) for specific details of the rpart package. Priors were set to 0.5, 0.5. The usesurrogate parameter was set to 0 so that subjects missing the primary split variable do not progress further down the tree, and maxsurrogate was set to 0 to cut computation time in half. The threshold complexity parameter (cp) was set to 0.01. Tree nodes were re-created in Microsoft Visio to display percentage depressed and the default number of controls/cases as presented by rpart.

Logistic Regression (LR). Variables found in association with depression based on RP analysis were considered in single factor LR models, separate by gender, using the specific dichotomous splitting of genotypes as designated by RP trees. Regression models for all seven SNP splits were first run on the full dataset to represent single main factor effects. Then each split was run on the respective subset of data as represented by the preceding RP split criteria. Thus, we attempt to mirror RP splits within a more formal LR framework in order to measure the significance of interactions presented by the trees. Multiple testing of 78 SNPs in RP for both male and females followed by 14 LR models resulted in a modified FDR significance level of 0.009.

RESULTS

Of the 4,811 participants (2,464 females and 2,347 males) under examination in this study, we identified 713 participants (481 females and 232 males) with depression (14.8 %). Given that the independent variable gender (when included as a factor in the full dataset) was the primary split on RP trees; that women are over two times as likely to be diagnosed with depression than men; and since the female etiology of depression has been reported to be associated with unique social, psychological, and biological factors (20), all subsequent analyses were performed by gender.

Recursive Partitioning Analysis

To examine multi-gene interactions for association with depression we screened our dataset using RP. The two-factor RP tree (*ANKK1*/*GNRH1*) was the optimized pruning for men (Fig. 1), while the five-factor tree (*DRD2*/*APOC3*/*ACVR2B*/*FTO*/*IL6*) was the optimized pruning for women (Fig. 2). For more detailed information on the 7 SNPs found by RP, see Supplementary Information 2.

The best overall split for men was *ANKK1* rs1800497 (historically known as the *DRD2* Taq1A allele), where the incidence of depression increased 2.2-fold in those with no C-alleles compared to those with one or two C-alleles. Considering interaction between *ANKK1* and *GNRH1* widened the disparity in incidence, where those with at least one C-allele in both *ANKK1* rs1800497 and the novel SNP in *GNRH1* had a 2.7-fold lower incidence than those without a C-allele in *ANKK1* rs1800497.

For women, the best overall split was *DRD2* rs2242592, where those with one or two C-alleles had 1.3-fold higher incidence of depression compared to those without any C-alleles. G x G interactions associated with the highest incidence of depression included: *DRD2* rs2242592 T/T + *APOC3* rs45537037 T/T + *ACVR2B* rs3749386 C/C or T/T, accounting for a 1.4-fold increase in depression compared to baseline incidence.

Single Main-Factor Effects

Specific SNP interactions identified by RP were next analyzed by LR (see Table 1, Full Data). The primary SNP splits in males and females were significant at the modified FDR level. Men with no C-alleles for *ANKK1* rs1800497 had 2.55 times higher odds [P=0.001 (1.44, 4.51)] of depression compared with men with at least 1 C-allele. Women with at least 1 C-allele for *DRD2* rs2242592 had 1.32 times higher odds [P=0.006 (1.08-1.62)] of depression compared with women with no C-alleles. One other split reached nominal significance; women homozygous (C/C or T/T) for *FTO* rs1421085 had 1.32 times higher odds [P=0.008 (1.08-1.62)] for depression than women with a heterozygous genotype. SNP splits of *GNRH1*, *APOC3*, *ACVR2B*, and *IL6* did not significantly associate with depression.

Gene-Gene Interactions Enhance Predictability for Depression

Specific SNP interactions identified by RP were next analyzed by LR as RP-specific subsets (see Table 1, RP-Subsetted Data). All 5 of the secondary and tertiary RP splits were found to be significant at the modified FDR level when considered as subsets. Among only men with at least one C-allele in *ANKK1* rs1800497, those with no C-allele in the novel SNP of *GNRH1* had 1.57 times higher odds [P=0.002 (1.18-2.08)] for depression than men with 1 or 2 C-alleles. For the subset of women in the first right-hand split of Fig. 2, those homozygous for *FTO* rs1421085 had 1.65 times higher odds [P=0.0005 (1.24-2.18)] for depression than women with a heterozygous genotype. For the remaining subset of women in the second right-hand split of Fig. 2, those homozygous for *IL6* rs1800795 had 1.85 times higher odds [P=0.006 (1.19-2.89)] for depression than women with a heterozygous genotype. For the subset of women in the first left-hand split of Fig. 2, those with no C-alleles for *APOC3* rs45537037 had 1.55 times higher odds [P=0.004 (1.15-2.09)] for depression than women with 1 or 2 C-alleles. For the subset of women in the second left-hand split of Fig. 2, those homozygous for *ACVR2B* rs3749386 had 2.17 times higher odds [P=0.001 (1.37-3.44)] for depression than women with a heterozygous genotype.

DISCUSSION

Utilizing RP as a screening tool to find potential multi-gene interactions, followed by verification of multi-gene interactions with LR, our data demonstrate that multi-gene interactions predict depression with a greater certainty than single main factor associations. RP provided us with primary dichotomous genotype splits in men and women (*ANKK1* rs1800497 and *DRD2* rs2242592, respectively) that were both significant in LR models at the modified FDR level (Table 1). Considering the 5 subsequent RP splits in LR over the entire dataset, only 1 reached a nominal level of significance (barely), which was *FTO* rs1421085 in women. However, after running LR on specific subsets of data according to the pattern of RP branches, every split was found to be significant and every odds ratios grew larger (Table 1; OR [before]→[after]: Male Left: 1.43 → 1.57, Female Left 1: 1.28 → 1.55, Female Left 2: 1.11 → 2.17, Female Right 1: 1.32 → 1.65, Female Right 2: 1.12 → 1.85). Thus, RP provides two unique and important criteria: dichotomous genotype splitting instructions and gene-gene interaction patterns. These criteria go beyond the traditional single factor SNP approach to genetic association studies and allow identification of important multi-gene pathways that more suitably characterize the etiology of complex diseases.

The Utility of Recursive Partitioning and Logistic Regression for Identification of Gene-Gene Interactions

With recent advances in genotyping allowing for high-dimensional SNP identification, it is now possible to examine genetic datasets not only for single main factor effects, but also G x G interactions. The requirement for G x G analyses as a better predictor of age-related diseases is obvious from the standpoint that humans are complex biological systems composed of numerous molecular interactions, and from recent studies indicating disease risk is modulated by G x G interactions (7). Notwithstanding this, the development of analytical tools for the identification of G x G interactions has not kept pace with the technological advances in identifying genetic alterations among individuals. In this respect, we have previously used MDR, LR and LD to identify G x G interactions among a small set of SNPs (7). However, large datasets require a screening tool to identify potential multi-gene interactions. In this study, we have

used RP to screen for multi-gene interactions, a data-mining technique that is currently under-utilized in genetic studies. RP serves as an efficient and powerful exploratory analysis technique, especially when looking for interactions in data sets with a large number of independent variables. This screening allows for the identification of G x G interactions (with greater explanatory power), that might otherwise not have been identified, and that can then be confirmed using more traditional statistical techniques. As illustrated in this paper, this data-mining methodology has the advantage of identification of genetic interactions *between* pathways involved in the etiology of depression, in keeping with the etiological heterogeneity of this disorder (see later).

Our study provides proof of principle for the use of RP in higher-dimensional analyses such as GWAS, where a comprehensive list of SNPs may fully explore genetic predisposition to depression and other age-related disease. The WLS is an ideal candidate for future GWAS studies given its large sample size, rich covariate composition and longitudinal nature.

In this genetic study we aimed to identify underlying genetic predispositions to depression and thus have not yet tested environmental/phenotypic data. Future analyses using RP to examine the impact of phenotypic and environmental factors on the development of depression would be anticipated to identify gene-phenotype/environment and multi-phenotype/environment interactions. Indeed, the predictive gains of G x G analyses were stronger for men than women, despite the fact that depression occurs disproportionately in women (~2:1 female-to-male; (21-25)). This suggests that environmental factors may be needed in addition to genetic factors in understanding the etiological pathways for women. Indeed, biological factors such as hormonal changes related to reproductive status (26, 27) may impact environmental factors such as psychosocial experiences (trauma, stress, interpersonal relationships, etc) and general health issues in the development of depression.

Genetic and Biological Correlates of Depression

Numerous studies have identified SNPs that associate with depression. Many of the SNPs associated

with depression from other studies were not significantly associated in our study. This is perhaps not surprising, since a single factor is unlikely to provide consistent association especially in a complex condition such as depression, where multiple pathways intersect in regulating the risk of the disease. For example, if a SNP within the serotonin pathway also requires a SNP in the glutamatergic pathway in order for the patient to present with depression, the presence of either SNP in the absence of the other will not be predictive of depression. Moreover, as indicated by Shi and Weinberg, since the human genome contains genetic redundancy, disruption of a single gene may be selectively neutral, but the malfunction of several genes in a pathway might result in expression of a particular phenotype (28).

Both the primary splits in men and women were SNPs linked with *DRD2* (dopamine receptor D2), a gene that has previously been linked with depression and social phobia (29-31). The primary male genotype split rs1800497, technically found in gene *ANKK1*, is historically known as the *DRD2* Taq1A allele because of its known association with decreased dopamine receptor D2 density (in those with T alleles) (32-35). The Taq1A allele has also been previously associated with depressive symptoms in children, where those with the A1 allele (T) were more likely to have depressive symptoms (36). We saw a similar association between A1 and depression in WLS men, where those with two A1 alleles had 2.6 times higher odds for depression compared to those with one or no A1 alleles. The primary split in women (DRD2 rs2242592) has previously been found to be associated with schizophrenia, where the C-allele was associated with higher susceptibility for schizophrenia (37). Interestingly, this same study also found the Taq1A allele to also associate with schizophrenia.

The secondary and tertiary right-hand splits in the female RP tree—*FTO* (fat mass and obesity associated) rs1421085 and *IL6* (interleukin 6) rs1800795—have also been found to relate with mental illness and depression in previous studies (38, 39). There is evidence that activin receptor signaling also is involved in affective disorders, especially when considering interaction with GABAergic pathways (40). Although we did not see an interaction between SNPs in GABA/activin receptor genes and depression, *ACVR2B* was associated with depression in women. No previous associations between depression and

APOC3, *ACVR2B*, or *GNRH1* have been reported.

That these genetic variants are associated with *neuroendocrine* pathways (*GnRH1, ACVR2B*) that are known to regulate *neurotransmitter* release and cognitive behavior (39-40) supports these associations as relevant to the etiology of depression and underlines the benefits of using RP to identify meaningful G x G interactions associated with disease.

Limitations

Given the numerous genetic, phenotypic and environmental influences that are linked to depression, and the small number of SNPs analyzed, it is not surprising that predictability from our models was low (although our predictability was superior to previous studies examining only single main factors). Also, the predictive value of our statistical models was further limited due to user bias in selection of SNPs (from nearly two-million SNPs in the human genome) used in this study. As a result of this, interactions we have found could potentially be moderated by another gene that we have not considered in this study. Nonetheless, we identified significant G x G interactions between known, and newly identified, loci associated with depression. Importantly, 4 of the 7 SNPs identified in these interactions were primarily located in two pathways well-known to impact depression: neurotransmitter and neuroendocrine signaling.

The results from the RP analyses conducted in this study were confirmed by LR, demonstrating the utility of RP as a screening tool for identifying meaningful G x G interactions. Future development of algorithms for RP analysis should not only maximize the distance between branches of the next best split (i.e. rpart), but consider subsequent future split combinations that could potentially result in trees with "better" overall predictability.

Summary

Our data indicate that G x G interaction analyses allows for enhanced predictability of conditions and diseases of aging. RP is an efficient and powerful exploratory analysis technique for elucidating G x G

interactions in large datasets and combined with LR provides an important statistical analysis for the identification of well supported G x G interactions. We predict that such analytical methods will play an increasingly important role in the identification of epistatic effects in future large GWAS. Finally, our studies illustrate how RP analyses can be used to find interacting pathways involved in the etiology of a disease or condition such as depression.

ACKNOWLEDGMENTS

This research uses data from the Wisconsin Longitudinal Study (WLS) of the University of Wisconsin-Madison. Since 1991, the WLS has been supported principally by the National Institute on Aging (AG-9775, AG-21079 and AG-033285), with additional support from the Vilas Estate Trust, the National Science Foundation, the Spencer Foundation, and the Graduate School of the University of Wisconsin-Madison. A public use file of data from the Wisconsin Longitudinal Study is available from the Wisconsin Longitudinal Study, University of Wisconsin-Madison, 1180 Observatory Drive, Madison, Wisconsin 53706 and at http://www.ssc.wisc.edu/wlsresearch/data/. This material is the result of work supported with resources at the William S. Middleton Memorial Veterans Hospital, Madison, WI. The opinions expressed herein are those of the authors. The contents do not represent the views of the Dept. of Veterans Affairs or the United States Government.

FIGURE LEGENDS

Figure 1. Recursive Partitioning Tree of CIDI-SF Depression in Males of the WLS. Upper and lower numbers in nodes represent the percentage of participants with depression and the number of controls/cases in that node, respectively. Blue and purple boxes/circles indicate lower and higher rates of depression relative to the primary node, respectively. Split information indicates gene, SNP, and genotype criteria, respectively. M1 is subset of data referenced in Table 1. Sensitivity: 0.526, Specificity: 0.598, Accuracy: 0.591. Due to missing genotype information, we lose approximately 1.5% of participants per split. *rs1800497 is historically known as the *DRD2* Taq1A allele

Figure 2. Recursive Partitioning Tree of CIDI-SF Depression in Females of the WLS. Upper and lower numbers in nodes represent the percentage of participants with depression and the number of controls/cases in that node, respectively. Blue and purple boxes/circles indicate lower and higher rates of depression relative to the primary node, respectively. Split information indicates gene, SNP, and genotype criteria, respectively. F1-F4 are subsets referenced in Table 1. Sensitivity: 0.607, Specificity: 0.563, Accuracy: 0.572. Due to missing genotype information, we lose approximately 1.4% of participants per split.

Table 1.
Single-factor logistic regression models based directly off male and female RP tree split criteria (see Figures 1 & 2). Each SNP split was first run on the full dataset to represent single main factor effects ("Full Data") for both males and females. Then the same SNP splits were run on specific subsets of data per RP tree splits (M1, F1-F4; "RP-Subsetted Data").

					Full Data		RP-Subsetted Data		
Gender	RP Split	Gene	SNP	Genotypes	OR (95% CI)	P-value	Subset	OR (95% CI)	P-value
Male	Primary	*ANKK1**	rs1800497	T/T vs. C/C + C/T	2.55 (1.44-4.51)	0.001 *	--	----	----
	Left	*GNRH1*	novel SNP	T/T vs. C/C + T/C	1.43 (1.09-1.88)	0.011	M1	1.57 (1.18-2.08)	0.002 *
Female	Primary	*DRD2*	rs2242592	C/C + T/C vs. T/T	1.32 (1.08-1.62)	0.006 *	--	----	----
	Left 1	*APOC3*	rs2854116	T/T vs. C/C + T/C	1.28 (1.04-1.57)	0.018	F1	1.55 (1.15-2.09)	0.004 *
	Left 2	*ACVR2B*	rs3749386	C/C + T/T vs. T/C	1.11 (0.91-1.36)	0.302	F2	2.17 (1.37-3.44)	0.001 *
	Right 1	*FTO*	rs1421085	C/C + T/T vs. T/C	1.32 (1.08-1.62)	0.007 *	F3	1.65 (1.24-2.18)	0.0005 *
	Right 2	*IL6*	rs1800795	C/C + G/G vs. C/G	1.12 (0.92-1.37)	0.269	F4	1.85 (1.19-2.89)	0.006 *

RP, recursive partitioning; OR, odds ratio; CI, confidence interval
M1: LR analysis was run for only those with genotype *DRD2* rs1800497 C/C or C/T
F1: LR analysis was run for only those with genotype *DRD2* rs2242592 T/T
F2: LR analysis was run for only those with genotypes *DRD2* rs2242592 T/T and *APOC3* rs2854116 T/T
F3: LR analysis was run for only those with genotype *DRD2* rs2242592 C/C or T/C
F4: LR analysis was run for only those with genotypes *DRD2* rs2242592 C/C or T/C and *FTO* rs1421085 T/C
*rs1800497 is historically known as the *DRD2* Taq1A allele

REFERENCES

1. CDC. Anxiety and Depression. Atlanta, GA: Centers for Disease Control and Prevention; 2009 [cited 2010 Oct. 18]; Available from: http://www.cdc.gov/Features/dsBRFSSDepressionAnxiety/.

2. Eley TC, Sugden K, Corsico A, et al. Gene-environment interaction analysis of serotonin system markers with adolescent depression. Mol Psychiatry. 2004;9(10):908-15.

3. Mirowsky J, Ross CE. Age and Depression. Journal of Health and Social Behavior. 1992;33(3):187-205.

4. Zubenko GS, Zubenko WN, Spiker DG, et. al. Malignancy of recurrent, early-onset major depression: a family study. Am J Med Genet. 2001 Dec 8;105(8):690-9.

5. Raymer KA, Waters RF, Price CR. Proposed multigenic Composite Inheritance in major depression. Medical Hypotheses. 2005;65(1):158-72.

6. Harold D, Abraham R, Hollingworth P, et al. Genome-wide association study identifies variants at CLU and PICALM associated with Alzheimer's disease. Nat Genet. 2009 Oct;41(10):1088-93.

7. Haasl R, Ahmadi MR, Meethal S, et al. A luteinizing hormone receptor intronic variant is significantly associated with decreased risk of Alzheimer's disease in males carrying an apolipoprotein E epsilon4 allele. BMC Medical Genetics. 2008;9(1):37.

8. Basson J, Wilson A, Haasl R, et. al. Multi locus interactions in steroidogenic pathway genes predict Alzheimer's disease. Alzheimer's and Dementia. 2008;4(4 (Suppl. 2)):T579: P3-199.

9. Pezawas L, Meyer-Lindenberg A, Goldman AL, et al. Evidence of biologic epistasis between BDNF and SLC6A4 and implications for depression. Mol Psychiatry. 2008;13(7):709-16.

10. Zhang K, Yang C, Xu Y, et al. Genetic association of the interaction between the BDNF and GSK3B genes and major depressive disorder in a Chinese population. Journal of Neural Transmission. 2010;117(3):393-401.

11. Lin E, Hong CJ, Hwang JP, et al. Gene-gene interactions of the brain-derived neurotrophic-factor and neurotrophic tyrosine kinase receptor 2 genes in geriatric depression. Rejuvenation Res. 2009;12(6):387-93.

12. Neff CD, Abkevich V, Potter J, et. al. Evidence for epistasis between SLC6A4 and a chromosome 4 gene as risk factors in major depression. American Journal of Medical Genetics Part B: Neuropsychiatric Genetics. 2010;153B(1):321-2.

13. Wong ML, Dong C, Andreev V, et. al. Prediction of susceptibility to major depression by a model of interactions of multiple functional genetic variants and environmental factors. Mol Psychiatry. 2012 Mar 27.

14. Carr D, Khodyakov D. End-of-life health care planning among young-old adults: An assessment of psychosocial influences. J Gerontol B-Psychol. 2007 Mar;62(2):S135-S41.

15. Sewell WH. As We Age : The Wisconsin Longitudinal Study, 1957-2001: Center for Demography and Ecology University of Wisconsin--Madison; 2001.

16. Nelson CB, Kessler RC, Mroczek D. Scoring the World Health Organization's Composite International Diagnostic Interview Short Form (CIDI-SF; v1.0 NOV98). 1998.

17. Zhang H, Bonney G. Use of classification trees for association studies. Genet Epidemiol. 2000 Dec;19(4):323-32.

18. Breiman L, Friedman J, Olshen R, et. al. Classification and Regression Trees. Boca Raton: Chapman and Hall/CRC; 1984.

19. Therneau T, Atkinson E, editors. Technical Report Series No. 61, An Introduction to Recursive Partitioning Using the RPART Routines. Rochester, MN: Department of Health Science Research, Mayo Clinic; 1997.

20. Accortt EE, Freeman MP, Allen JJB. Women and Major Depressive Disorder: Clinical Perspectives on Causal Pathways. Journal of Women's Health. 2008;17(10):1583-90.

21. Kessler RC, McGonagle KA, Swartz M, et. al. Sex and depression in the National Comorbidity Survey. I: Lifetime prevalence, chronicity and recurrence. J Affect Disord. 1993 Oct-Nov;29(2-3):85-96.

22. Kessler RC, McGonagle KA, Nelson CB, et. al. Sex and depression in the National Comorbidity Survey. II: Cohort effects. J Affect Disord. 1994 Jan;30(1):15-26.

23. Kessler RC, McGonagle KA, Zhao S, et al. Lifetime and 12-month prevalence of DSM-III-R psychiatric disorders in the United States. Results from the National Comorbidity Survey. Arch Gen Psychiatry. 1994 Jan;51(1):8-19.

24. Weissman MM, Bland R, Joyce PR, et. al. Sex differences in rates of depression: cross-national perspectives. J Affect Disord. 1993 Oct-Nov;29(2-3):77-84.

25. Weissman MM, Olfson M. Depression in women: implications for health care research. Science. 1995 Aug 11;269(5225):799-801.

26. Pajer K. New strategies in the treatment of depression in women. The Journal of clinical psychiatry. [Review]. 1995;56 Suppl 2:30-7.

27. Paykel ES. Depression in women. Br J Psychiatry Suppl. [Review]. 1991 May(10):22-9.

28. Shi M, Weinberg CR. How Much Are We Missing in SNP-by-SNP Analyses of Genome-wide Association Studies? Epidemiology. 2011 Nov;22(6):845-7.

29. Lawford BR, Young R, Noble EP, et. al. The D2 dopamine receptor (DRD2) gene is associated with co-morbid depression, anxiety and social dysfunction in untreated veterans with post-traumatic stress disorder. European Psychiatry. 2006;21(3):180-5.

30. Klimek V, Schenck JE, Han H, et. al. Dopaminergic abnormalities in amygdaloid nuclei in major depression: a postmortem study. Biol Psychiatry. 2002 Oct 1;52(7):740-8.

31. Schneier FR, Liebowitz MR, Abi-Dargham A, et. al. Low dopamine D(2) receptor binding potential in social phobia. Am J Psychiatry. 2000 Mar;157(3):457-9.

32. Noble EP, Blum K, Ritchie T et. al. Allelic Association of the D2 Dopamine Receptor Gene With Receptor-Binding Characteristics in Alcoholism or Gene ism. Arch Gen Psychiatry. 1991 July 1, 1991;48(7):648-54.

33. Thompson J, Thomas N, Singleton A, et al. D2 dopamine receptor gene (DRD2) Taq1 A polymorphism: reduced dopamine D2 receptor binding in the human striatum associated with the A1 allele. Pharmacogenetics. 1997 Dec;7(6):479-84.

34. Pohjalainen T, Rinne JO, Nagren K, et al. The A1 allele of the human D2 dopamine receptor gene predicts low D2 receptor availability in healthy volunteers. Mol Psychiatry. 1998 May;3(3):256-60.

35. Jonsson EG, Nothen MM, Grunhage F, et al. Polymorphisms in the dopamine D2 receptor gene and their relationships to striatal dopamine receptor density of healthy volunteers. Mol Psychiatry. 1999 May;4(3):290-6.

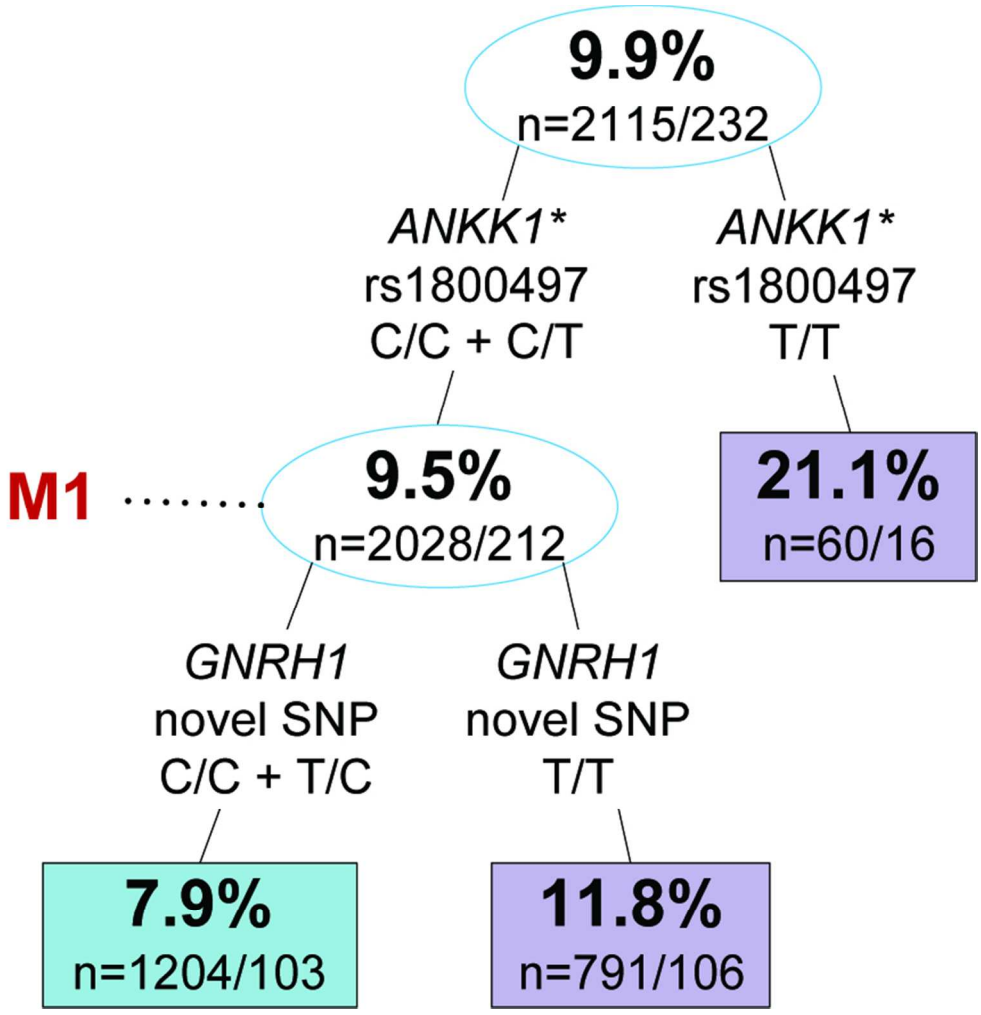
36. Hayden EP, Klein DN, Dougherty LR, et al. The dopamine D2 receptor gene and depressive and anxious symptoms in childhood: associations and evidence for gene-environment correlation and gene-environment interaction. Psychiatr Genet. 2010 Dec;20(6):304-10.

37. Dubertret C, Bardel C, Ramoz N, et al. A genetic schizophrenia-susceptibility region located between the ANKK1 and DRD2 genes. Progress in Neuro-Psychopharmacology and Biological Psychiatry. 2010;34(3):492-9.

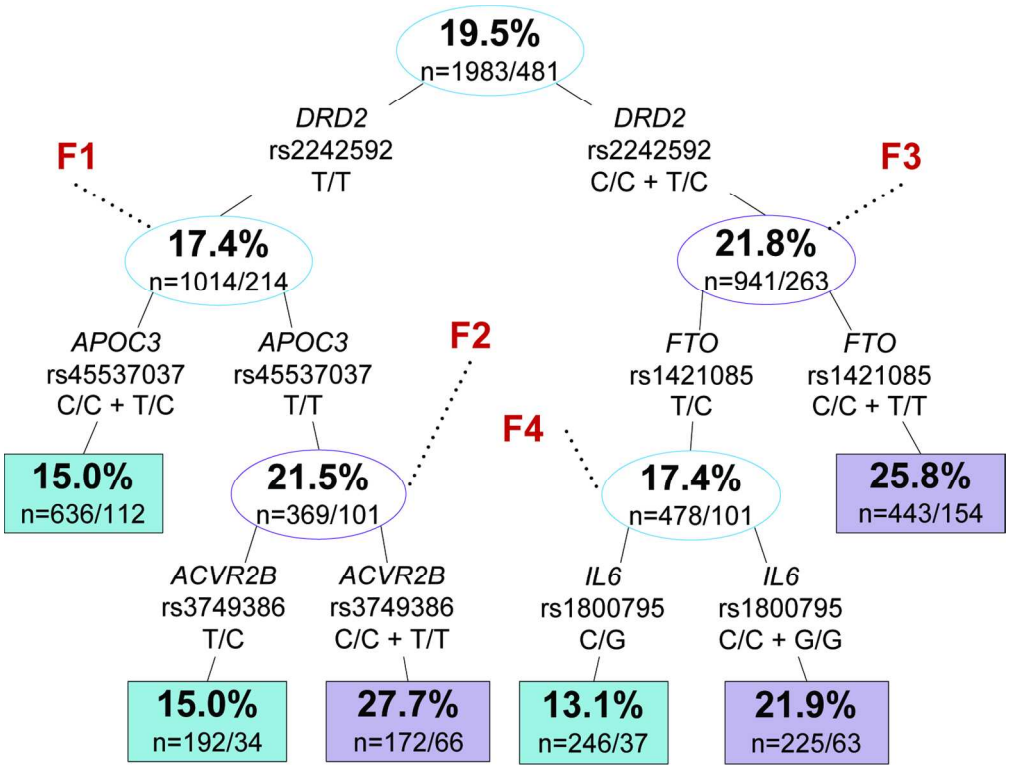
38. Kivimaki M, Jokela M, Hamer M, et al. Examining Overweight and Obesity as Risk Factors for Common Mental Disorders Using Fat Mass and Obesity-Associated (FTO) Genotype-Instrumented Analysis. Am J Epidemiol. 2011 Feb 15;173(4):421-9.

39. Bull SJ, Huezo-Diaz P, Binder EB, et al. Functional polymorphisms in the interleukin-6 and serotonin transporter genes, and depression and fatigue induced by interferon-alpha and ribavirin treatment. Mol Psychiatry. 2009 Dec;14(12):1095-104.

40. Zheng F, Adelsberger H, Muller MR, et. al. Alzheimer C. Activin tunes GABAergic neurotransmission and modulates anxiety-like behavior. Mol Psychiatry. 2008;14(3):332-46.



74x74mm (300 x 300 DPI)



134x102mm (300 x 300 DPI)

Supplementary Table 1. Single Nucleotide Polymorphisms Assessed in the WLS

Gene	Encodes	SNP	Associated disease/behavior
A2M	alpha-2-macroglobulin	rs669	Alzheimer's disease (1)
ACVR2A	activin receptor IIA	rs1424954	pre-eclampsia (2)
ACVR2B	activin receptor IIB	rs3749386	--
ADIPOQ	adiponectin, C1Q and collagen domain containing	rs1501299	diabetes II (3, 4), obesity (5, 6), breast cancer (7)
ADIPOQ	adiponectin, C1Q and collagen domain containing	rs2241766	diabetes II (3, 4), obesity (8), breast cancer (7)
ACVRL1	activin receptor-like kinase 1	rs2071219	brain arteriovenous malformations (9)
APOC-3	apolipoprotein C-III	rs2854116	nonalcoholic fatty liver disease (10)
ApoE	apolipoprotein E	rs429358	Alzheimer's disease (11, 12)
ApoE	apolipoprotein E	rs7412	Alzheimer's disease (11, 12)
AR	androgen receptor	rs6152	male pattern baldness (13)
BCKDHB	branched chain keto acid dehydrogenase E1, beta polypeptide	rs4502885	premature ovarian failure (14)
BDNF	brain-derived neurotrophic factor	rs6265	depression (15-17), alcohol dependence-related depression (18), bipolar disorder (19), schizophrenia (20), cognition (21), BMI (22)
BDNF	brain-derived neurotrophic factor	rs908867	antidepressant response (23)
BRCA1	breast cancer 1, early onset	rs1799966	breast cancer (24)
BRCA2	breast cancer 2, early onset homolog	rs144848	breast cancer (24)
CH25H	cholesterol 25-hydroxylase	rs3802657	--
CHRM2	cholinergic receptor, muscarinic 2	rs2061174	alcohol dependence, depression (25)
CHRM2	cholinergic receptor, muscarinic 2	rs8191992	cognition (26)
COMT	catechol-O-methyltransferase	rs4680	ADHD (27), substance abuse (28-31), depression (32), antidepressant response (33), bipolar disorder (34), cognition (35)
CTSD	cathepsin D	rs17571	Alzheimer's disease (36)
CYP11A1	cytochrome P450, family 11, subfamily A, polypeptide 1	rs8039957	breast cancer (37)
CYP11B2	cytochrome P450, family 11, subfamily B, polypeptide 2	rs1799998	stroke (38), cardiovascular disease (39)
DAT1	human dopamine transporter	rs11564774	ADHD (40)
DAT1	human dopamine transporter	rs2963238	alcohol-withdrawal seizures (41)
DISC1	disrupted in schizophrenia 1	rs821616	schizophrenia (42), cognitive aging (43)
DRD2	dopamine receptor D2	rs17529477	--
DRD2/ANKK1	dopamine receptor D2/ ankyrin repeat and kinase domain containing 1	rs1800497	obesity, drug addiction (44)
DRD2	dopamine receptor D2	rs2242592	schizophrenia (45)
DRD2	dopamine receptor D2	rs4245147	--
DRD2	dopamine receptor D2	rs6277	schizophrenia (46), PTSD (47)
DRD4	dopamine receptor D4	rs1800955	ADHD (48), heroine addiction (49)
DTNBP1	dystrobrevin-binding protein 1	rs1018381	schizophrenia (50), cognitive ability (51)
DTNBP1	dystrobrevin-binding protein 1	rs760761	schizophrenia (52)
ESR1	estrogen receptor 1	rs7761133	--
ESR1	estrogen receptor 1	rs3853248	--
FADS2	fatty acid desaturase 2	rs1535	breastfeeding & IQ (53)

FADS2	fatty acid desaturase 2	rs174575	breastfeeding & IQ (53)
FMR1	fragile X mental retardation 1	rs1805420	--
FSH	follicle stimulating hormone	rs6169	--
FSHR	follicle stimulating hormone receptor	rs6166	sterility (54), osteoporosis (55)
FST	follistatin	rs12152850	--
FST	follistatin	rs3797297	--
FTO	fat mass and obesity associated	rs1421085	obesity (56-58), mental disorders (59)
GABBR2	γ-aminobutyric acid B receptor 2	rs1435252	nicotine addiction (60)
GABBR2	γ-aminobutyric acid B receptor 2	rs2779562	nicotine addiction (60)
GNRH1	gonadotropin-releasing hormone	novel SNP	Alzheimer's disease (61)
HERC	hect domain and RLD 2	rs12913832	eye color (62, 63)
HFE	hemochromatosis	rs1799945	hemochromatosis(64)
HSD17B1	estradiol 17β-dehydrogenase 1	rs12602084	steroid metabolism (65)
HSD17B1	estradiol 17β-dehydrogenase 1	rs592389	vasomotor symptoms (66), cognition (67)
5-HTR1A	5-hydroxytryptamine (serotonin) receptor 1A	rs878567	mood disorders (68)
5-HTR2A	5-hydroxytryptamine (serotonin) receptor 2A	rs6312	--
5-HTR2A	5-hydroxytryptamine (serotonin) receptor 2A	rs6314	antidepressant response (69), bipolar disorder (70)
5-HTR2A	5-hydroxytryptamine (serotonin) receptor 2A	rs7997012	antidepressant response (71)
5-HTR2C	5-hydroxytryptamine (serotonin) receptor 2C	rs6318	bipolar disorder (72), depression (73)
5-HTT	5-hydroxytryptamine transporter	rs25533	antidepressant response (74)
5-HTT	5-hydroxytryptamine transporter	rs8076005	depressive symptoms (75)
IGF1	insulin-like growth factor 1	rs12313279	--
IL1A	interleukin 1, alpha	rs17561	chronic rhinosinusitis (76), BMI (77)
IL6	interleukin 6	rs1800795	arthritis (78), breast cancer (79), diabetes (80), depression (81)
INHA	inhibin alpha	rs2059693	testicular cancer (82)
INHA	inhibin alpha	rs35118453	--
INHBA	inhibin beta A	rs2237436	--
INHBB	inhibin beta B	rs11902591	--
KIBRA	kidney and brain protein (WWC1)	rs17070145	Alzheimer's disease (83), episodic memory (84)
LEPR	leptin receptor	rs1137100	diabetes II (85), atherosclerosis (86)
LHR	luteinizing hormone receptor	rs4073366	Alzheimer's disease (87)
MAOA	monoamine oxidase A	rs3788862	pain (88)
OXTR	oxytocin receptor	rs2254298	autism (89, 90), social loneliness (91), depressive symptoms & anxiety (92)
PCK1	phosphoenolpyruvate carboxykinase 1	rs707555	diabetes II (93)
PGR	progesterone receptor	rs1042838	ovarian cancer (94), migraine (95), menstruation (96), pregnancy loss (97)
SNAP25	synaptosomal-associated protein 25	rs363050	intelligence (98, 99)
SSADH	succinic semialdehyde dehydrogenase	rs2760118	--
StAR	steroidogenic acute regulatory protein	rs3990403	--
TFAM	transcription factor A, mitochondrial	rs1937	Alzheimer's disease (100)
TFAM	transcription factor A, mitochondrial	rs2306604	Parkinson's disease (101)
TPH1	first tryptophan hydroxylase isoform	rs1799913	heroine addiction (102)

TPH2	first tryptophan hydroxylase isoform	rs11178997	bipolar disorder (103), PTSD symptoms (104)

REFERENCES

1. Zappia M, Manna I, Serra P, Cittadella R, Andreoli V, La Russa A, et al. Increased risk for Alzheimer disease with the interaction of MPO and A2M polymorphisms. Arch Neurol. 2004 Mar;61(3):341-4.
2. Moses EK, Fitzpatrick E, Freed KA, Dyer TD, Forrest S, Elliott K, et al. Objective prioritization of positional candidate genes at a quantitative trait locus for pre-eclampsia on 2q22. Molecular Human Reproduction. 2006 August 2006;12(8):505-12.
3. Hara K, Boutin P, Mori Y, Tobe K, Dina C, Yasuda K, et al. Genetic Variation in the Gene Encoding Adiponectin Is Associated With an Increased Risk of Type 2 Diabetes in the Japanese Population. Diabetes. 2002 February 1, 2002;51(2):536-40.
4. Gu HF, Abulaiti A, Ostenson CG, Humphreys K, Wahlestedt C, Brookes AJ, et al. Single nucleotide polymorphisms in the proximal promoter region of the adiponectin (APM1) gene are associated with type 2 diabetes in Swedish caucasians. Diabetes. 2004 Feb;53 Suppl 1:S31-5.
5. Bouatia-Naji N, Meyre D, Lobbens S, Seron K, Fumeron F, Balkau B, et al. ACDC/adiponectin polymorphisms are associated with severe childhood and adult obesity. Diabetes. 2006 Feb;55(2):545-50.
6. Yang W-S, Yang Y-C, Chen C-L, Wu I-L, Lu J-Y, Lu F-H, et al. Adiponectin SNP276 is associated with obesity, the metabolic syndrome, and diabetes in the elderly. The American Journal of Clinical Nutrition. 2007 August 1, 2007;86(2):509-13.
7. Kaklamani VG, Sadim M, Hsi A, Offit K, Oddoux C, Ostrer H, et al. Variants of the adiponectin and adiponectin receptor 1 genes and breast cancer risk. Cancer Res. 2008 May 1;68(9):3178-84.
8. Beckers S, Peeters AV, De Freitas F, Mertens IL, Verhulst SL, Haentjens D, et al. Association Study and Mutation Analysis of Adiponectin Shows Association of Variants in APM1 with Complex Obesity in Women. Annals of Human Genetics. 2009;73(5):492-501.
9. Young WL, Kwok P-Y, Pawlikowska L, Lawton MT, Kim H, Hysi PG, et al. Arteriovenous Malformation. Journal of Neurosurgery. 2007;106(4):731-2.
10. Petersen KF, Dufour S, Hariri A, Nelson-Williams C, Foo JN, Zhang X-M, et al. Apolipoprotein C3 Gene Variants in Nonalcoholic Fatty Liver Disease. New England Journal of Medicine. 2010;362(12):1082-9.
11. Corder EH, Saunders AM, Strittmatter WJ, Schmechel DE, Gaskell PC, Small GW, et al. Gene dose of apolipoprotein E type 4 allele and the risk of Alzheimer's disease in late onset families. Science. 1993;261(5123):921-3.
12. Saunders AM, Strittmatter WJ, Schmechel D, George-Hyslop PH, Pericak-Vance MA, Joo SH, et al. Association of apolipoprotein E allele epsilon 4 with late-onset familial and sporadic Alzheimer's disease. Neurology. 1993;43(8):1467-72.
13. Ellis JA, Stebbing M, Harrap SB. Polymorphism of the Androgen Receptor Gene is Associated with Male Pattern Baldness. 2001;116(3):452-5.
14. Kang H, Lee SK, Cho S-W, Lee S-H, Kwack K. Branched chain [alpha]-keto acid dehydrogenase, E1-[beta] subunit gene is associated with premature ovarian failure. Fertility and Sterility. 2008;89(3):728-31.

15. Hwang J-P, Tsai S-J, Hong C-J, Yang C-H, Lirng J-F, Yang Y-M. The Val66Met polymorphism of the brain-derived neurotrophic-factor gene is associated with geriatric depression. Neurobiology of Aging. 2006;27(12):1834-7.
16. Czira ME, Wersching H, Baune BT, Berger K. Brain-derived neurotrophic factor gene polymorphisms, neurotransmitter levels, and depressive symptoms in an elderly population. Age (Dordr). 2011 Sep 7.
17. Suchanek R, Owczarek A, Kowalczyk M, Kucia K, Kowalski J. Association between C-281A and val66met functional polymorphisms of BDNF gene and risk of recurrent major depressive disorder in Polish population. J Mol Neurosci. 2011 Mar;43(3):524-30.
18. Su N, Zhang L, Fei F, Hu H, Wang K, Hui H, et al. The brain-derived neurotrophic factor is associated with alcohol dependence-related depression and antidepressant response. Brain Research. 2011;1415(0):119-26.
19. Sears C, Markie D, Olds R, Fitches A. Evidence of associations between bipolar disorder and the brain-derived neurotrophic factor (BDNF) gene. Bipolar Disorders. 2011;13(7-8):630-7.
20. Zakharyan R, Boyajyan A, Arakelyan A, Gevorgyan A, Mrazek F, Petrek M. Functional variants of the genes involved in neurodevelopment and susceptibility to schizophrenia in an Armenian population. Human Immunology. 2011;72(9):746-8.
21. McAllister TW, Tyler AL, Flashman LA, Rhodes CH, McDonald BC, Saykin AJ, et al. Polymorphisms in the brain-derived neurotrophic factor gene influence memory and processing speed one month after brain injury. J Neurotrauma. 2012 Apr 10;29(6):1111-8.
22. Hong K-W, Lim J-E, Go MJ, Shin Cho Y, Ahn Y, Han B-G, et al. Recapitulation of the Association of the Val66Met Polymorphism of BDNF Gene With BMI in Koreans. Obesity. 2011.
23. Gratacos M, Soria V, Urretavizcaya M, Gonzalez JR, Crespo JM, Bayes M, et al. A brain-derived neurotrophic factor (BDNF) haplotype is associated with antidepressant treatment outcome in mood disorders. Pharmacogenomics J. 2007;8(2):101-12.
24. Johnson N, Fletcher O, Palles C, Rudd M, Webb E, Sellick G, et al. Counting potentially functional variants in BRCA1, BRCA2 and ATM predicts breast cancer susceptibility. Human Molecular Genetics. 2007 May 1, 2007;16(9):1051-7.
25. Wang JC, Hinrichs AL, Stock H, Budde J, Allen R, Bertelsen S, et al. Evidence of common and specific genetic effects: association of the muscarinic acetylcholine receptor M2 (CHRM2) gene with alcohol dependence and major depressive syndrome. Human Molecular Genetics. 2004 September 1, 2004;13(17):1903-11.
26. Jones KA, Porjesz B, Almasy L, Bierut L, Goate A, Wang JC, et al. Linkage and linkage disequilibrium of evoked EEG oscillations with CHRM2 receptor gene polymorphisms: implications for human brain dynamics and cognition. International Journal of Psychophysiology. 2004;53(2):75-90.
27. Eisenberg J, Mei-Tal G, Steinberg A, Tartakovsky E, Zohar A, Gritsenko I, et al. Haplotype relative risk study of catechol-O-methyltransferase (COMT) and attention deficit hyperactivity disorder (ADHD): Association of the high-enzyme activity val allele with adhd impulsive-hyperactive phenotype. American Journal of Medical Genetics. 1999;88(5):497-502.
28. Wang T, Franke P, Neidt H, Cichon S, Knapp M, Lichtermann D, et al. Association study of the low-activity allele of catechol-O-methyltransferase and alcoholism using a family-based approach. Mol Psychiatry. 2001 Jan;6(1):109-11.
29. Horowitz R, Kotler M, Shufman E, Aharoni S, Kremer I, Cohen H, et al. Confirmation of an excess of the high enzyme activity COMT val allele in heroin addicts in a family-based haplotype relative risk study. American Journal of Medical Genetics. 2000;96(5):599-603.

30. Vandenbergh DJ, Rodriguez LA, Miller IT, Uhl GR, Lachman HM. High-activity catechol-O-methyltransferase allele is more prevalent in polysubstance abusers. American Journal of Medical Genetics. 1997;74(4):439-42.
31. Li T, Chen C-k, Hu X, Ball D, Lin S-K, Chen W, et al. Association analysis of the DRD4 and COMT genes in methamphetamine abuse. American Journal of Medical Genetics Part B: Neuropsychiatric Genetics. 2004;129B(1):120-4.
32. Åberg E, Fandiño-Losada A, Sjöholm LK, Forsell Y, Lavebratt C. The functional Val158Met polymorphism in catechol-O-methyltransferase (COMT) is associated with depression and motivation in men from a Swedish population-based study. Journal of Affective Disorders. 2011;129(1–3):158-66.
33. Baune BT, Hohoff C, Berger K, Neumann A, Mortensen S, Roehrs T, et al. Association of the COMT val158met Variant with Antidepressant Treatment Response in Major Depression. Neuropsychopharmacology. 2007;33(4):924-32.
34. Lee S-Y, Chen S-L, Chen S-H, Huang S-Y, Tzeng N-S, Chang Y-H, et al. The COMT and DRD3 genes interacted in bipolar I but not bipolar II disorder. World Journal of Biological Psychiatry. 2011;12(5):385-91.
35. Dickinson D, Elvevåg B. Genes, cognition and brain through a COMT lens. Neuroscience. 2009;164(1):72-87.
36. Albayrak Ö, Tirniceriu A, Riemenschneider M, Kurz A, Scherag A, Egensperger R. The Cathepsin D (224C/T) Polymorphism Confers an Increased Risk to Develop Alzheimer's Disease in Men. The Journals of Gerontology Series A: Biological Sciences and Medical Sciences. 2010 March 1, 2010;65A(3):219-24.
37. Yaspan BL, Breyer JP, Cai Q, Dai Q, Elmore JB, Amundson I, et al. Haplotype Analysis of CYP11A1 Identifies Promoter Variants Associated with Breast Cancer Risk. Cancer Research. 2007 June 15, 2007;67(12):5673-82.
38. Tu Y, Cui G, Xu Y, Bao X, Wang X, Wang D. Genetic polymorphism of CYP11B2 gene and stroke in the Han Chinese population and a meta-analysis. Pharmacogenet Genomics. 2011;21(3):115-20.
39. White PC, Hautanen A, Kupari M. Aldosterone synthase (CYP11B2) polymorphisms and cardiovascular function. J Steroid Biochem Mol Biol. 1999;69(1-6):409-12.
40. Brookes K, Xu X, Chen W, Zhou K, Neale B, Lowe N, et al. The analysis of 51 genes in DSM-IV combined type attention deficit hyperactivity disorder: association signals in DRD4, DAT1 and 16 other genes. Mol Psychiatry. 2006;11(10):934-53.
41. Le Strat Y, Ramoz N, Pickering P, Burger V, Boni C, Aubin H-J, et al. The 3' Part of the Dopamine Transporter Gene DAT1/SLC6A3 Is Associated With Withdrawal Seizures in Patients With Alcohol Dependence. Alcoholism: Clinical and Experimental Research. 2008;32(1):27-35.
42. Callicott JH, Straub RE, Pezawas L, Egan MF, Mattay VS, Hariri AR, et al. Variation in DISC1 affects hippocampal structure and function and increases risk for schizophrenia. Proc Natl Acad Sci U S A. 2005;102(24):8627-32.
43. Thomson PA, Harris SE, Starr JM, Whalley LJ, Porteous DJ, Deary IJ. Association between genotype at an exonic SNP in DISC1 and normal cognitive aging. Neuroscience Letters. 2005;389(1):41-5.
44. Blum K, Braverman ER, Wood RC, Gill J, Li C, Chen TJ, et al. Increased prevalence of the Taq I A1 allele of the dopamine receptor gene (DRD2) in obesity with comorbid substance use disorder: a preliminary report. Pharmacogenetics. 1996 Aug;6(4):297-305.
45. Dubertret C, Bardel C, Ramoz N, Martin P-M, Deybach J-C, Adès J, et al. A genetic schizophrenia-susceptibility region located between the ANKK1 and DRD2 genes. Progress in Neuro-Psychopharmacology and Biological Psychiatry. 2010;34(3):492-9.
46. Monakhov M, Golimbet V, Abramova L, Kaleda V, Karpov V. Association study of three polymorphisms in the dopamine D2 receptor gene and schizophrenia in the Russian population. Schizophrenia Research. 2008;100(1–3):302-7.

47. Voisey J, Swagell CD, Hughes IP, Morris CP, van Daal A, Noble EP, et al. The DRD2 gene 957C>T polymorphism is associated with Posttraumatic Stress Disorder in war veterans. Depression and Anxiety. 2009;26(1):28-33.
48. Yang J-W, Jang W-S, Hong SD, Ji YI, Kim DH, Park J, et al. A case-control association study of the polymorphism at the promoter region of the DRD4 gene in Korean boys with attention deficit-hyperactivity disorder: Evidence of association with the - 521 C/T SNP. Progress in Neuro-Psychopharmacology and Biological Psychiatry. 2008;32(1):243-8.
49. Lai JH, Zhu YS, Huo ZH, Sun RF, Yu B, Wang YP, et al. Association study of polymorphisms in the promoter region of DRD4 with schizophrenia, depression, and heroin addiction. Brain Research. 2010;1359:227-32.
50. Funke B, Finn CT, Plocik AM, Lake S, DeRosse P, Kane JM, et al. Association of the DTNBP1 Locus with Schizophrenia in a U.S. Population. The American Journal of Human Genetics. 2004;75(5):891-8.
51. Burdick KE, Lencz T, Funke B, Finn CT, Szeszko PR, Kane JM, et al. Genetic variation in DTNBP1 influences general cognitive ability. Human Molecular Genetics. 2006 15 May 2006;15(10):1563-8.
52. Riley B, Kuo P-H, Maher BS, Fanous AH, Sun J, Wormley B, et al. The dystrobrevin binding protein 1 (DTNBP1) gene is associated with schizophrenia in the Irish Case Control Study of Schizophrenia (ICCSS) sample. Schizophrenia Research. 2009;115(2–3):245-53.
53. Caspi A, Williams B, Kim-Cohen J, Craig IW, Milne BJ, Poulton R, et al. Moderation of breastfeeding effects on the IQ by genetic variation in fatty acid metabolism. Proc Natl Acad Sci U S A. 2007 Nov 20;104(47):18860-5.
54. Binder H, Dittrich R, Hager I, Muller A, Oeser S, Beckmann MW, et al. Association of FSH receptor and CYP19A1 gene variations with sterility and ovarian hyperstimulation syndrome. Reproduction. 2008;135(1):107-16.
55. Rendina D, Gianfrancesco F, De Filippo G, Merlotti D, Esposito T, Mingione A, et al. FSHR gene polymorphisms influence bone mineral density and bone turnover in postmenopausal women. Eur J Endocrinol. 2010;163(1):165-72.
56. Dina C, Meyre D, Gallina S, Durand E, Korner A, Jacobson P, et al. Variation in FTO contributes to childhood obesity and severe adult obesity. Nat Genet. 2007;39(6):724-6.
57. Tung Y-CL, Yeo GSH. From GWAS to biology: lessons from FTO. Annals of the New York Academy of Sciences. 2011;1220(1):162-71.
58. Chabris C, Hebert B, Benjamin D, Beauchamp J, Cesarini D, van der Loos M, et al. Most reported genetic assocations with general intelligence are probably false positives. Psychological Science. 2011;In press.
59. Kivimaki M, Jokela M, Hamer M, Geddes J, Ebmeier K, Kumari M, et al. Examining Overweight and Obesity as Risk Factors for Common Mental Disorders Using Fat Mass and Obesity-Associated (FTO) Genotype-Instrumented Analysis. Am J Epidemiol. 2011 Feb 15;173(4):421-9.
60. Beuten J, Ma JZ, Payne TJ, Dupont RT, Crews KM, Somes G, et al. Single- and Multilocus Allelic Variants within the GABAB Receptor Subunit 2 (GABAB2) Gene Are Significantly Associated with Nicotine Dependence. The American Journal of Human Genetics. 2005;76(5):859-64.
61. Wilson AC. Investigations into gonadotropin-releasing hormone receptor signaling in extrahypothalamic neurons of the brain: Implications for normal cognition and the pathogenesis of Alzheimer's disease: Thesis. University of Wisconsin, Madison, USA; 2009.
62. Eiberg H, Troelsen J, Nielsen M, Mikkelsen A, Mengel-From J, Kjaer K, et al. Blue eye color in humans may be caused by a perfectly associated founder mutation in a regulatory element located within the HERC2 gene inhibiting OCA2 expression. Human Genetics. 2008;123(2):177-87.

63. Sturm RA, Duffy DL, Zhao ZZ, Leite FPN, Stark MS, Hayward Nicholas K, et al. A Single SNP in an Evolutionary Conserved Region within Intron 86 of the HERC2 Gene Determines Human Blue-Brown Eye Color. The American Journal of Human Genetics. 2008;82(2):424-31.
64. Beutler E. The significance of the 187G (H63D) mutation in hemochromatosis. Am J Hum Genet. 1997 Sep;61(3):762-4.
65. Sun T, Oh WK, Jacobus S, Regan M, Pomerantz M, Freedman ML, et al. The Impact of Common Genetic Variations in Genes of the Sex Hormone Metabolic Pathways on Steroid Hormone Levels and Prostate Cancer Aggressiveness. Cancer Prevention Research. 2011 December 1, 2011;4(12):2044-50.
66. Crandall CJ, Crawford SL, Gold EB. Vasomotor Symptom Prevalence Is Associated with Polymorphisms in Sex Steroid-Metabolizing Enzymes and Receptors. The American Journal of Medicine. 2006;119(9, Supplement 1):S52-S60.
67. Kravitz HM, Meyer PM, Seeman TE, Greendale GA, Sowers MR. Cognitive Functioning and Sex Steroid Hormone Gene Polymorphisms in Women at Midlife. The American Journal of Medicine. 2006;119(9, Supplement 1):S94-S102.
68. Brezo J, Bureau A, Merette C, Jomphe V, Barker ED, Vitaro F, et al. Differences and similarities in the serotonergic diathesis for suicide attempts and mood disorders: a 22-year longitudinal gene-environment study. Mol Psychiatry. 2010;15(8):831-43.
69. Wilkie MJV, Smith G, Day RK, Matthews K, Smith D, Blackwood D, et al. Polymorphisms in the SLC6A4 and HTR2A genes influence treatment outcome following antidepressant therapy. Pharmacogenomics J. 2008;9(1):61-70.
70. McAuley EZ, Fullerton JM, Blair IP, Donald JA, Mitchell PB, Schofield PR. Association between the serotonin 2A receptor gene and bipolar affective disorder in an Australian cohort. Psychiatr Genet. 2009 Oct;19(5):244-52.
71. McMahon FJ, Buervenich S, Charney D, Lipsky R, Rush AJ, Wilson AF, et al. Variation in the Gene Encoding the Serotonin 2A Receptor Is Associated with Outcome of Antidepressant Treatment. The American Journal of Human Genetics. 2006;78(5):804-14.
72. Mazza M, Mandelli L, Martinotti G, Di Nicola M, Tavian D, Negri G, et al. Further evidence supporting the association between 5HTR2C gene and bipolar disorder. Psychiatry Res. 2010 Dec 30;180(2-3):151-2.
73. Lerer B, Macciardi F, Segman RH, Adolfsson R, Blackwood D, Blairy S, et al. Variability of 5-HT2C receptor cys23ser polymorphism among European populations and vulnerability to affective disorder. Mol Psychiatry. 2001 Sep;6(5):579-85.
74. Lin E, Chen PS, Chang HH, Gean P-W, Tsai HC, Yang YK, et al. Interaction of serotonin-related genes affects short-term antidepressant response in major depressive disorder. Progress in Neuro-Psychopharmacology and Biological Psychiatry. 2009;33(7):1167-72.
75. Su S, Zhao J, Bremner JD, Miller AH, Tang W, Bouzyk M, et al. Serotonin Transporter Gene, Depressive Symptoms, and Interleukin-6 / CLINICAL PERSPECTIVE. Circulation: Cardiovascular Genetics. 2009 December 1, 2009;2(6):614-20.
76. Mfuna Endam L, Cormier C, Bosse Y, Filali-Mouhim A, Desrosiers M. Association of IL1A, IL1B, and TNF Gene Polymorphisms With Chronic Rhinosinusitis With and Without Nasal Polyposis: A Replication Study. Arch Otolaryngol Head Neck Surg. 2010 February 1, 2010;136(2):187-92.
77. Um J-Y, Rim H-K, Kim S-J, Kim H-L, Hong S-H. Functional Polymorphism of IL-1 Alpha and Its Potential Role in Obesity in Humans and Mice. PLoS ONE. 2011;6(12):e29524.
78. Fishman D, Faulds G, Jeffery R, Mohamed-Ali V, Yudkin JS, Humphries S, et al. The effect of novel polymorphisms in the interleukin-6 (IL-6) gene on IL-6 transcription and plasma IL-6 levels, and an association with systemic-onset juvenile chronic arthritis. The Journal of Clinical Investigation. 1998;102(7):1369-76.

79. Slattery ML, Curtin K, Sweeney C, Wolff RK, Baumgartner RN, Baumgartner KB, et al. Modifying Effects of IL-6 Polymorphisms on Body Size-Associated Breast Cancer Risk. Obesity. 2008;16(2):339-47.
80. Huth C, Heid IM, Vollmert C, Gieger C, Grallert H, Wolford JK, et al. IL6 Gene Promoter Polymorphisms and Type 2 Diabetes. Diabetes. 2006 October 2006;55(10):2915-21.
81. Bull SJ, Huezo-Diaz P, Binder EB, Cubells JF, Ranjith G, Maddock C, et al. Functional polymorphisms in the interleukin-6 and serotonin transporter genes, and depression and fatigue induced by interferon-alpha and ribavirin treatment. Mol Psychiatry. 2009 Dec;14(12):1095-104.
82. Purdue MP, Graubard BI, Chanock SJ, Rubertone MV, Erickson RL, McGlynn KA. Genetic Variation in the Inhibin Pathway and Risk of Testicular Germ Cell Tumors. Cancer Research. 2008 April 15, 2008;68(8):3043-8.
83. Rodríguez-Rodríguez E, Infante J, Llorca J, Mateo I, Sánchez-Quintana C, García-Gorostiaga I, et al. Age-dependent association of KIBRA genetic variation and Alzheimer's disease risk. Neurobiology of Aging. 2009;30(2):322-4.
84. Almeida OP, Schwab SG, Lautenschlager NT, Morar B, Greenop KR, Flicker L, et al. KIBRA genetic polymorphism influences episodic memory in later life, but does not increase the risk of mild cognitive impairment. Journal of Cellular and Molecular Medicine. 2008;12(5a):1672-6.
85. Salopuro T, Pulkkinen L, Lindstrom J, Eriksson JG, Valle TT, Hamalainen H, et al. Genetic variation in leptin receptor gene is associated with type 2 diabetes and body weight: The Finnish Diabetes Prevention Study. Int J Obes Relat Metab Disord. 2005;29(10):1245-51.
86. Saukko M, Kesaniemi YA, Ukkola O. Leptin receptor Lys109Arg and Gln223Arg polymorphisms are associated with early atherosclerosis. Metab Syndr Relat Disord. 2010;8(5):425-30.
87. Haasl R, Ahmadi MR, Meethal S, Gleason C, Johnson S, Asthana S, et al. A luteinizing hormone receptor intronic variant is significantly associated with decreased risk of Alzheimer's disease in males carrying an apolipoprotein E epsilon4 allele. BMC Medical Genetics. 2008;9(1):37.
88. Kim H, Lee H, Rowan J, Brahim J, Dionne R. Genetic polymorphisms in monoamine neurotransmitter systems show only weak association with acute post-surgical pain in humans. Molecular Pain. 2006;2(1):24.
89. Lerer E, Levi S, Salomon S, Darvasi A, Yirmiya N, Ebstein RP. Association between the oxytocin receptor (OXTR) gene and autism: relationship to Vineland Adaptive Behavior Scales and cognition. Mol Psychiatry. 2007;13(10):980-8.
90. Liu X, Kawamura Y, Shimada T, Otowa T, Koishi S, Sugiyama T, et al. Association of the oxytocin receptor (OXTR) gene polymorphisms with autism spectrum disorder (ASD) in the Japanese population. J Hum Genet. 2010;55(3):137-41.
91. Lucht MJ, Barnow S, Sonnenfeld C, Rosenberger A, Grabe HJ, Schroeder W, et al. Associations between the oxytocin receptor gene (OXTR) and affect, loneliness and intelligence in normal subjects. Progress in Neuro-Psychopharmacology and Biological Psychiatry. 2009;33(5):860-6.
92. Thompson RJ, Parker KJ, Hallmayer JF, Waugh CE, Gotlib IH. Oxytocin receptor gene polymorphism (rs2254298) interacts with familial risk for psychopathology to predict symptoms of depression and anxiety in adolescent girls. Psychoneuroendocrinology. 2011;36(1):144-7.
93. Dong Y, Zhang H, Wang X, Feng Q, Chen X, Su Q. A Leu184Val polymorphism in PCK1 gene is associated with type 2 diabetes in Eastern Chinese population with BMI < 23 kg/m2. Diabetes Research and Clinical Practice. 2009;83(2):227-32.
94. Pearce CL, Hirschhorn JN, Wu AH, Burtt NP, Stram DO, Young S, et al. Clarifying the PROGINS Allele Association in Ovarian and Breast Cancer Risk: A Haplotype-Based Analysis. Journal of the National Cancer Institute. 2005 January 5, 2005;97(1):51-9.

95. Joshi G, Pradhan S, Mittal B. Role of the oestrogen receptor (ESR1 PvuII and ESR1 325 C→G) and progesterone receptor (PROGINS) polymorphisms in genetic susceptibility to migraine in a North Indian population. Cephalalgia. 2010 March 1, 2010;30(3):311-20.
96. Taylor KC, Small CM, Epstein MP, Sherman SL, Tang W, Wilson MM, et al. Associations of Progesterone Receptor Polymorphisms with Age at Menarche and Menstrual Cycle Length. Hormone Research in Paediatrics. 2010;74(6):421-7.
97. Su MT, Lee IW, Chen YC, Kuo PL. Association of progesterone receptor polymorphism with idiopathic recurrent pregnancy loss in Taiwanese Han population. J Assist Reprod Genet. 2011 Mar;28(3):239-43.
98. Gosso MF, de Geus EJC, van Belzen MJ, Polderman TJC, Heutink P, Boomsma DI, et al. The SNAP-25 gene is associated with cognitive ability: evidence from a family-based study in two independent Dutch cohorts. Mol Psychiatry. 2006;11(9):878-86.
99. Gosso MF, De Geus EJC, Polderman TJC, Boomsma DI, Heutink P, Posthuma D. Common variants underlying cognitive ability: further evidence for association between the SNAP-25 gene and cognition using a family-based study in two independent Dutch cohorts. Genes, Brain and Behavior. 2008;7(3):355-64.
100. Zhang Q, Yu J-T, Wang P, Chen W, Wu Z-C, Jiang H, et al. Mitochondrial transcription factor A (TFAM) polymorphisms and risk of late-onset Alzheimer's disease in Han Chinese. Brain Research. 2011;1368(0):355-60.
101. Gaweda-Walerych K, Safranow K, Maruszak A, Bialecka M, Klodowska-Duda G, Czyzewski K, et al. Mitochondrial transcription factor A variants and the risk of Parkinson's disease. Neuroscience Letters. 2010;469(1):24-9.
102. Nielsen DA, Barral S, Proudnikov D, Kellogg S, Ho A, Ott J, et al. TPH2 and TPH1: association of variants and interactions with heroin addiction. Behav Genet. 2008 Mar;38(2):133-50.
103. Cichon S, Winge I, Mattheisen M, Georgi A, Karpushova A, Freudenberg J, et al. Brain-specific tryptophan hydroxylase 2 (TPH2): a functional Pro206Ser substitution and variation in the 5'-region are associated with bipolar affective disorder. Human Molecular Genetics. 2008 January 1, 2008;17(1):87-97.
104. Goenjian AK, Bailey JN, Walling DP, Steinberg AM, Schmidt D, Dandekar U, et al. Association of TPH1, TPH2, and 5HTTLPR with PTSD and depressive symptoms. Journal of Affective Disorders. 2012(in Press).

Supplementary Table 2. Depression-Associated SNP Identified in the WLS

Gene	Encodes	SNP	Alleles	Chr#/Location	Residue	Associated disease/behavior
ACVR2B	activin receptor IIB	rs3749386	T/C	3/intron 1	--	left-right axis malformations*(1)
APOC3	apolipoprotein C-III	rs2854116	T/C	11/promoter (-455)	--	nonalcoholic fatty liver disease(2)
DRD2/ANKK1	dopamine receptor D2/ankyrin repeat and kinase domain containing 1	rs1800497	C/T	11/exon (ANKK1)	Glu713Lys	obesity, drug addiction (3)
DRD2	dopamine receptor D2	rs2242592	T/C	11/3'	--	schizophrenia (4)
FTO	fat mass and obesity associated	rs1421085	T/C	16/intron 1	--	obesity (5-7), mental disorders (8)
GNRH1	gonadotropin-releasing hormone	novel SNP	T/C	8/promoter	--	Alzheimer's disease (9)
IL6	interleukin 6	rs1800795	C/G	7/promoter (-174)	--	arthritis (10), breast cancer (11), diabetes (12), depression (13)

*Gene association only

References

1. Kosaki R, Gebbia M, Kosaki K, Lewin M, Bowers P, Towbin JA, et al. Left–right axis malformations associated with mutations in ACVR2B, the gene for human activin receptor type IIB. American Journal of Medical Genetics. 1999;82(1):70-6.
2. Petersen KF, Dufour S, Hariri A, Nelson-Williams C, Foo JN, Zhang X-M, et al. Apolipoprotein C3 Gene Variants in Nonalcoholic Fatty Liver Disease. New England Journal of Medicine. 2010;362(12):1082-9.
3. Blum K, Braverman ER, Wood RC, Gill J, Li C, Chen TJ, et al. Increased prevalence of the Taq I A1 allele of the dopamine receptor gene (DRD2) in obesity with comorbid substance use disorder: a preliminary report. Pharmacogenetics. 1996 Aug;6(4):297-305.
4. Dubertret C, Bardel C, Ramoz N, Martin P-M, Deybach J-C, Adès J, et al. A genetic schizophrenia-susceptibility region located between the ANKK1 and DRD2 genes. Progress in Neuro-Psychopharmacology and Biological Psychiatry. 2010;34(3):492-9.
5. Dina C, Meyre D, Gallina S, Durand E, Korner A, Jacobson P, et al. Variation in FTO contributes to childhood obesity and severe adult obesity. Nat Genet. 2007;39(6):724-6.
6. Tung Y-CL, Yeo GSH. From GWAS to biology: lessons from FTO. Annals of the New York Academy of Sciences. 2011;1220(1):162-71.
7. Chabris C, Hebert B, Benjamin D, Beauchamp J, Cesarini D, van der Loos M, et al. Most reported genetic assocations with general intelligence are probably false positives. Psychological Science. 2011;In press.
8. Kivimaki M, Jokela M, Hamer M, Geddes J, Ebmeier K, Kumari M, et al. Examining Overweight and Obesity as Risk Factors for Common Mental Disorders Using Fat Mass and Obesity-Associated (FTO) Genotype-Instrumented Analysis. Am J Epidemiol. 2011 Feb 15;173(4):421-9.
9. Wilson AC. Investigations into gonadotropin-releasing hormone receptor signaling in extrahypothalamic neurons of the brain: Implications for normal cognition and the pathogenesis of Alzheimer's disease: Thesis. University of Wisconsin, Madison, USA; 2009.
10. Fishman D, Faulds G, Jeffery R, Mohamed-Ali V, Yudkin JS, Humphries S, et al. The effect of novel polymorphisms in the interleukin-6 (IL-6) gene on IL-6 transcription and plasma IL-6 levels, and an association with systemic-onset juvenile chronic arthritis. The Journal of Clinical Investigation. 1998;102(7):1369-76.
11. Slattery ML, Curtin K, Sweeney C, Wolff RK, Baumgartner RN, Baumgartner KB, et al. Modifying Effects of IL-6 Polymorphisms on Body Size-Associated Breast Cancer Risk. Obesity. 2008;16(2):339-47.
12. Huth C, Heid IM, Vollmert C, Gieger C, Grallert H, Wolford JK, et al. IL6 Gene Promoter Polymorphisms and Type 2 Diabetes. Diabetes. 2006 October 2006;55(10):2915-21.
13. Bull SJ, Huezo-Diaz P, Binder EB, Cubells JF, Ranjith G, Maddock C, et al. Functional polymorphisms in the interleukin-6 and serotonin transporter genes, and depression and fatigue induced by interferon-alpha and ribavirin treatment. Mol Psychiatry. 2009 Dec;14(12):1095-104.

Multi-Gene Interactions and the Prediction of Depression in the Wisconsin Longitudinal Study

Nicholas S. Roetker[1], James A. Yonker[1], Chee Lee[1], Vicky Chang[1], Jacob Basson[2], Carol L. Roan[1], Taissa S. Hauser[1], Robert M. Hauser[1] and Craig S. Atwood[2,3].

[1]Department of Sociology, University of Wisconsin-Madison, Madison, WI 53706, USA.

[2]Geriatric Research, Education and Clinical Center, Veterans Administration Hospital and Department of Medicine, University of Wisconsin-Madison School of Medicine and Public Health, Madison, WI 53705, USA.

[3]School of Exercise, Biomedical and Health Sciences, Edith Cowan University, Joondalup, 6027 WA, Australia.

Running title: Gene interactions and depression

Address Correspondence and Reprint Requests to:
Craig S. Atwood, Ph.D.
University of Wisconsin-Madison School of Medicine and Public Health
Wm S. Middleton Memorial VA (GRECC 11G)
2500 Overlook Terrace, Madison, WI 53705, USA
Tel. 608 256-1901, Ext. 11664
Fax. 608 280-7291
Email: csa@medicine.wisc.edu

ABSTRACT

Objectives: Single genetic loci offer little predictive power for the identification of depression. This study examined whether an analysis of gene-gene interactions of 78 single nucleotide polymorphisms in genes associated with depression and age-related diseases would identify significant interactions with increased predictive power for depression.

Design: A retrospective cohort study.

Setting: A survey of participants in the Wisconsin Longitudinal Study.

Participants: A total of 4,811 persons (2,464 females and 2,347 males) who provided saliva for genotyping; the group comes from a randomly selected sample of Wisconsin high school graduates from the class of 1957 as well as a randomly selected sibling, almost all of whom are non-Hispanic white.

Primary outcome measure: Depression as determine by the Composite International Diagnostic Interview short-form (CIDI-SF).

Results: Using a classification tree approach (recursive partitioning (RP)) we identified a number of candidate gene-gene interactions associated with depression. The primary SNP splits revealed by RP (*ANKK1* rs1800497 (also known as *DRD2* Taq1A) in men and *DRD2* rs224592 in women) were found to be significant as single factors by logistic regression (LR) after controlling for multiple testing (P=0.001 for both). Without considering interaction effects, only 1 of the 5 subsequent RP splits reached nominal significance in logistic regression (*FTO* rs1421085 in women; P-value=0.008). However, after controlling for gene-gene interactions by running logistic regression on RP-specific subsets, every split became significant and grew larger in magnitude (OR [before]→[after]: Men: *GNRH1* novel SNP: [1.43 → 1.57]; Women: *APOC3* rs2854116: [1.28 → 1.55], *ACVR2B* rs3749386: [1.11 → 2.17], *FTO* rs1421085: [1.32 → 1.65], *IL6* rs1800795: [1.12 → 1.85]).

Conclusions: Our results suggest that examining gene-gene interactions improves the identification of genetic associations predictive of depression. Four of the SNPs identified in these interactions were located in two pathways well-known to impact depression: neurotransmitter (*ANKK1* and *DRD2*) and

Formatted: Font: Italic

neuroendocrine (*GNRH1* and *ACVR2B*) signaling. This study demonstrates the utility of RP analysis as an efficient and powerful exploratory analysis technique for uncovering genetic and molecular pathway interactions associated with disease etiology.

INTRODUCTION

Depression is a widespread mental disorder associated with a host of undesirable health, social, and economic outcomes. One in six Americans is diagnosed with depression in his or her lifetime (1). While many environmental factors—such as socioeconomic status, childhood abuse, and major life events—have important ties with depression, so too does gender and many genetic and epigenetic factors, making the disorder heterogeneous in nature (2). Another major risk factor for depression is age, with depression reaching its highest levels in adults over 80 years of age (3).

It has been demonstrated from twin studies that genetic factors typically account for 40–70% of the risk for developing major depressive disorder (MDD), and adoption studies have confirmed the role of genetic risk factors in the development of MDD (see (4) and references therein). Genetic studies, including recent genome-wide association studies (GWAS), have identified genetic alterations in over 50 genes known to be associated with depression (5). However, individually, the genetic alterations found within these genes (primarily single nucleotide polymorphisms (SNPs)) have little predictive value. There is a similar lack of predictive value from GWAS of other major age-related diseases (6).

Given this lack of predictive power among individual genetic alterations for depression together with the complex nature of aging-related diseases, it would seem prudent to examine epistatic effects on this age-related condition. In this respect, we have previously demonstrated that G x G interactions greatly modulate risk for complex age-related diseases (7, 8). Recent studies of depression also have identified epistatic effects. In particular, associations have been identified between *BDNF* Val66Met (brain-derived neurotrophic factor; rs6265) and *5-HTTLPR* (serotonin transporter linked promoter region (9); *GSK3B* rs6782799 (glycogen synthase kinase 3β), *BDNF* rs7124442 and *BDNF* Val66Met (10); *BDNF* Val66Met and SNPs in *NTRK2* (neurotrophic tyrosine kinase receptor 2; (11)), and *5-HTTLPR* short allele and a chromosome 4 gene (12). The machine learning tool recursive partitioning has recently been used by Wong (13) in order to assess complex gene-gene interactions in depression. Wong notes that recursive partitioning is useful in that it quickly explores high dimensional data for non-linear effects that are non-

biased and easily interpretable.

The goals of this study were therefore to 1) explore G x G interactions that might better predict the genetic factors involved in the etiology of depression, and 2) to further demonstrate the utility of machine learning algorithms (recursive partitioning) to identify genetic interactions. Using genotypic data from the Wisconsin Longitudinal Study (WLS) we identified associations between dopaminergic genes and depression in men and women, as well as G x G interactions involving neuroendocrine signaling pathways, with increased significance compared with single genetic associations.

METHODS

Study Participants and Surveys

Data were collected from the WLS, a random sample originally comprised of 10,317 men and women who graduated from Wisconsin high schools in 1957. Later in 1977, the WLS began interviewing one randomly selected sibling of each graduate, when possible. The cohort reflects the ancestral makeup of the late-1950s Wisconsin population in that participants are almost entirely non-Hispanic white males and females. . In general, the sample is broadly representative of older white Americans with at least a high school education (14). Further characteristics of the WLS cohort may be found in detail elsewhere (15). Health and psychological well-being phenotypic data was taken from mail and phone surveys given in 2004-2005. Inclusion criteria for depression included any member of the WLS cohort who was depressed according to the Composite International Diagnostic Interview short-form (CIDI-SF). Individuals who answered YES to the question "Have you ever had a time in life lasting two weeks or more when nearly every day you felt sad, blue, depressed, or when you lost interest in most things like work, hobbies, or things you usually liked to do for fun?" and whose depression was not caused by alcohol, drugs, medications, or physical illness were asked further depression symptom questions. Symptom questions asked whether the two week period was accompanied with a) any weight loss, b) trouble sleeping, c) feeling tired, d) feeling bad upon waking, e) losing interest, f) trouble concentrating, or g) thoughts about death. Those answering YES to 3 or more of these symptom questions were classified as having depression (16). Those answering YES to 2 or fewer symptom questions and all those answering NO to the initial stem question were classified as controls.

~~Data were collected from the WLS, a random sample originally comprised of 10,317 men and women who graduated from Wisconsin high schools in 1957. Later in 1977, the WLS began interviewing one randomly selected sibling of each graduate, when possible. The cohort consists reflects the ancestral makeup of the late-1950s Wisconsin population in that participants are almost entirely of non-Hispanic white personsmales and females. whose average level of educational attainment was 1.5 years of post-~~

high school education at the time of interview in 2004. Ages of participants in the WLS ranged from 35 to 90 years old at this time, with 83% of participants being between 60 and 70 years old. In general, the sample is broadly representative of older white Americans with at least a high school education (14). Further characteristics of the WLS cohort may be found in detail elsewhere (15). Health and psychological well-being phenotypic data was taken from mail and phone surveys given in 2004-2005. Our main measure of depression is based on a variation of the Composite International Diagnostic Interview short-form (CIDI-SF). All participants answered a single stem question: "Have you ever had a time in life lasting two weeks or more when nearly every day you felt sad, blue, depressed, or when you lost interest in most things like work, hobbies, or things you usually liked to do for fun?" Only those who answered YES and whose depression was not always caused by alcohol, drugs, medications, or physical illness were asked further depression symptom questions. Symptom questions asked whether the two week period was accompanied with a) any weight loss, b) trouble sleeping, c) feeling tired, d) feeling bad upon waking, e) losing interest, f) trouble concentrating, or g) thoughts about death. Those answering YES to 3 or more of these symptom questions were classified as having depression (16). Those answering YES to 2 or fewer symptom questions and all those answering NO to the initial stem question were classified as controls.

Formatted: Font color: Red

Genotyping

7,101 participants (4,569 graduates & 2,532 siblings) provided saliva samples in Oragene DNA sample collection kits from which DNA was extracted and genotyped for 78 SNPs that were selected based on their association with depression and age-related conditions and diseases (see Supplementary Information 1). Genotyping was performed by KBioscience (Hoddesdon, UK) with use of a homogeneous Fluorescent Resonance Energy Transfer technology coupled to competitive allele specific PCR. All SNP genotypes described in our results were in Hardy-Weinberg equilibrium and their frequencies matched those reported in the literature for European samples.

Statistical Analysis

Analyses were limited to the 4,811 pooled graduates and siblings for whom we have depression and genotype information (Note: individuals with more than 10% missing genotype data were not included). The average age among this sample was just under 65 years in 2004. 80% were married, and the average amount of post-high school educational attainment was 2 years. Median household income in 1993 was $56,700.

Recursive Partitioning (RP). RP is a data mining tool for revealing trends that relate a dependent variable (depressed vs. non-depressed) to various predictor variables (SNPs). Zhang and Bonney have shown how RP can be used in genetic association studies to identify disease genes (17). RP helps control for heterogeneity in the population and confounding factors by allowing for the segregation of the sample population according to any condition. Thus, RP is a useful way to handle complex datasets that might confound regression analysis due to the complexity of the relationship between the independent and dependent variables and due to missing information.

RP classification trees (using R package rpart) were used to identify potential interactions among the 78 SNPs in relation to depression. The trees split the data along branches according to criteria determined by the rpart package algorithm, which is originally based off the work of Breiman's classification and regression trees (CART) algorithm (18). Basically, the CART algorithm first considers all depressed and non-depressed subjects pooled together in a heterogeneous root node. Based on considering every possible "yes-no" binary partition that can be made by each independent variable, the single split which maximizes homogeneity between the two resulting sub-nodes as compared to the root node is made. Each sub-node can then be treated independently as a new root node for all subsequent splits, and the pattern continues until every subject constitutes a terminal node, resulting in a very large and complex tree. A 10-part cross validation procedure seeking to minimize misclassification and complexity determines optimal pruning. See Therneau and Atkinson (19) for specific details of the rpart package.

Priors were set to 0.5, 0.5. The usesurrogate parameter was set to 0 so that subjects missing the primary split variable do not progress further down the tree, and maxsurrogate was set to 0 to cut computation time in half. The threshold complexity parameter (cp) was set to 0.01. Tree nodes were re-created in Microsoft Visio to display percentage depressed ~~depression incidence (in %) and total number of participants rather than~~and the default number of controls/cases as presented by rpart.

Logistic Regression (LR). Variables found in association with depression based on RP analysis were considered in single factor LR models, separate by gender, using the specific dichotomous splitting of genotypes as designated by RP trees. Regression models for all seven SNP splits were first run on the full dataset to represent single main factor effects. Then each split was run on the respective subset of data as represented by the preceding RP split criteria. Thus, we attempt to mirror RP splits within a more formal LR framework in order to measure the significance of interactions presented by the trees. Multiple testing of 78 SNPs in RP for both male and females followed by 14 LR models resulted in a modified FDR significance level of 0.009.

RESULTS

Of the 4,811 participants (2,464 females and 2,347 males) under examination in this study, we identified 713 participants (481 females and 232 males) with depression (14.8 %). Given that the independent variable gender (when included as a factor in the full dataset) was the primary split on RP trees; that women are over two times as likely to be diagnosed with depression than men; and since the female etiology of depression has been reported to be associated with unique social, psychological, and biological factors (20), all subsequent analyses were performed by gender.

Recursive Partitioning Analysis

To examine multi-gene interactions for association with depression we screened our dataset using RP. The two-factor RP tree (*ANKK1*/*GNRH1*) was the optimized pruning for men (Fig. 1), while the five-factor tree (*DRD2*/*APOC3*/*ACVR2B*/*FTO*/*IL6*) was the optimized pruning for women (Fig. 2). For more detailed information on the 7 SNPs found by RP, see Supplementary Information 2. Note that subjects are lost in every step down each tree due to missing genotype information. We lose approximately 1.5% of data per split in men and 1.4% of data per split in women.

The best overall split for men was *ANKK1* rs1800497 (historically known as the *DRD2* Taq1A allele), where the incidence of depression increased 2.2-fold in those with no C-alleles compared to those with one or two C-alleles. Considering interaction between *ANKK1* and *GNRH1* widened the disparity in incidence, where those with at least one C-allele in both *ANKK1* rs1800497 and the novel SNP in *GNRH1* had a 2.7-fold lower incidence than those without a C-allele in *ANKK1* rs1800497.

For women, the best overall split was *DRD2* rs2242592, where those with one or two C-alleles had 1.3-fold higher incidence of depression compared to those without any C-alleles. G x G interactions associated with the highest incidence of depression included: *DRD2* rs2242592 T/T + *APOC3* rs45537037 T/T + *ACVR2B* rs3749386 C/C or T/T, accounting for a 1.4-fold increase in depression compared to baseline incidence.

Single Main-Factor Effects

Specific SNP interactions identified by RP were next analyzed by LR (see Table 1, Full Data). The primary SNP splits in males and females were significant at the modified FDR level. Men with no C-alleles for *ANKK1* rs1800497 had 2.55 times higher odds [P=0.001 (1.44, 4.51)] of depression compared with men with at least 1 C-allele. Women with at least 1 C-allele for *DRD2* rs2242592 had 1.32 times higher odds [P=0.006 (1.08-1.62)] of depression compared with women with no C-alleles. One other split reached nominal significance; women homozygous (C/C or T/T) for *FTO* rs1421085 had 1.32 times higher odds [P=0.008 (1.08-1.62)] for depression than women with a heterozygous genotype. SNP splits of *GNRH1*, *APOC3*, *ACVR2B*, and *IL6* did not significantly associate with depression.

Gene-Gene Interactions Enhance Predictability for Depression

Specific SNP interactions identified by RP were next analyzed by LR as RP-specific subsets (see Table 1, RP-Subsetted Data). All 5 of the secondary and tertiary RP splits were found to be significant at the modified FDR level when considered as subsets. Among only men with at least one C-allele in *ANKK1* rs1800497, those with no C-allele in the novel SNP of *GNRH1* had 1.57 times higher odds [P=0.002 (1.18-2.08)] for depression than men with 1 or 2 C-alleles. For the subset of women in the first right-hand split of Fig. 2, those homozygous for *FTO* rs1421085 had 1.65 times higher odds [P=0.0005 (1.24-2.18)] for depression than women with a heterozygous genotype. For the remaining subset of women in the second right-hand split of Fig. 2, those homozygous for *IL6* rs1800795 had 1.85 times higher odds [P=0.006 (1.19-2.89)] for depression than women with a heterozygous genotype. For the subset of women in the first left-hand split of Fig. 2, those with no C-alleles for *APOC3* rs45537037 had 1.55 times higher odds [P=0.004 (1.15-2.09)] for depression than women with 1 or 2 C-alleles. For the subset of women in the second left-hand split of Fig. 2, those homozygous for *ACVR2B* rs3749386 had 2.17 times higher odds [P=0.001 (1.37-3.44)] for depression than women with a heterozygous genotype.

DISCUSSION

Utilizing RP as a screening tool to find potential multi-gene interactions, followed by verification of multi-gene interactions with LR, our data demonstrate that multi-gene interactions predict depression with a greater certainty than single main factor associations. RP provided us with primary dichotomous genotype splits in men and women (*ANKK1* rs1800497 and *DRD2* rs2242592, respectively) that were both significant in LR models at the modified FDR level (Table 1). Considering the 5 subsequent RP splits in LR over the entire dataset, only 1 reached a nominal level of significance (barely), which was *FTO* rs1421085 in women. However, after running LR on specific subsets of data according to the pattern of RP branches, every split was found to be significant and every odds ratios grew larger (Table 1; OR [before]→[after]: Male Left: 1.43 → 1.57, Female Left 1: 1.28 → 1.55, Female Left 2: 1.11 → 2.17, Female Right 1: 1.32 → 1.65, Female Right 2: 1.12 → 1.85). Thus, RP provides two unique and important criteria: dichotomous genotype splitting instructions and gene-gene interaction patterns. These criteria go beyond the traditional single factor SNP approach to genetic association studies and allow identification of important multi-gene pathways that more suitably characterize the etiology of complex diseases.

The Utility of Recursive Partitioningand Logistic Regression for Identification of Gene-Gene Interactions

With recent advances in genotyping allowing for high-dimensional SNP identification, it is now possible to examine genetic datasets not only for single main factor effects, but also G x G interactions. The requirement for G x G analyses as a better predictor of age-related diseases is obvious from the standpoint that humans are complex biological systems composed of numerous molecular interactions, and from recent studies indicating disease risk is modulated by G x G interactions (7). Notwithstanding this, the development of analytical tools for the identification of G x G interactions has not kept pace with the technological advances in identifying genetic alterations among individuals. In this respect, we have

previously used MDR, LR and LD to identify G x G interactions among a small set of SNPs (7). However, large datasets require a screening tool to identify potential multi-gene interactions. In this study, we have used RP to screen for multi-gene interactions, a data-mining technique that is currently under-utilized in genetic studies. RP serves as an efficient and powerful exploratory analysis technique, especially when looking for interactions in data sets with a large number of independent variables. This screening allows for the identification of G x G interactions (with greater explanatory power), that might otherwise not have been identified, and that can then be confirmed using more traditional statistical techniques. As illustrated in this paper, this data-mining methodology has the advantage of identification of genetic interactions *between* pathways involved in the etiology of depression, in keeping with the etiological heterogeneity of this disorder (see later).

Our study provides proof of principle for the use of RP in higher-dimensional analyses such as GWAS, where a comprehensive list of SNPs may fully explore genetic predisposition to depression and other age-related disease. The WLS is an ideal candidate for future GWAS studies given its large sample size, rich covariate composition and longitudinal nature.

In this genetic study we aimed to identify underlying genetic predispositions to depression and thus have not yet tested environmental/phenotypic data. Future analyses using RP to examine the impact of phenotypic and environmental factors on the development of depression would be anticipated to identify gene-phenotype/environment and multi-phenotype/environment interactions. Indeed, the predictive gains of G x G analyses were stronger for men than women, despite the fact that depression occurs disproportionately in women (~2:1 female-to-male; (21-25)). This suggests that environmental factors may be needed in addition to genetic factors in understanding the etiological pathways for women. Indeed, biological factors such as hormonal changes related to reproductive status (26, 27) may impact environmental factors such as psychosocial experiences (trauma, stress, interpersonal relationships, etc) and general health issues in the development of depression.

Genetic and Biological Correlates of Depression

Numerous studies have identified SNPs that associate with depression. Many of the SNPs associated with depression from other studies were not significantly associated in our study. This is perhaps not surprising, since a single factor is unlikely to provide consistent association especially in a complex condition such as depression, where multiple pathways intersect in regulating the risk of the disease. For example, if a SNP within the serotonin pathway also requires a SNP in the glutamatergic pathway in order for the patient to present with depression, the presence of either SNP in the absence of the other will not be predictive of depression. Moreover, as indicated by Shi and Weinberg, since the human genome contains genetic redundancy, disruption of a single gene may be selectively neutral, but the malfunction of several genes in a pathway might result in expression of a particular phenotype (28).

Both the primary splits in men and women were SNPs linked with *DRD2* (dopamine receptor D2), a gene that has previously been linked with depression and social phobia (29-31). The primary male genotype split rs1800497, technically found in gene *ANKK1*, is historically known as the *DRD2* Taq1A allele because of its known association with decreased dopamine receptor D2 density (in those with T alleles) (32-35). The Taq1A allele has also been previously associated with depressive symptoms in children, where those with the A1 allele (T) were more likely to have depressive symptoms (36). We saw a similar association between A1 and depression in WLS men, where those with two A1 alleles had 2.6 times higher odds for depression compared to those with one or no A1 alleles. The primary split in women (DRD2 rs2242592) has previously been found to be associated with schizophrenia, where the C-allele was associated with higher susceptibility for schizophrenia (37). Interestingly, this same study also found the Taq1A allele to also associate with schizophrenia.

The secondary and tertiary right-hand splits in the female RP tree—*FTO* (fat mass and obesity associated) rs1421085 and *IL6* (interleukin 6) rs1800795—have also been found to relate with mental illness and depression in previous studies (38, 39). There is evidence that activin receptor signaling also is involved in affective disorders, especially when considering interaction with GABAergic pathways (40).

Although we did not see an interaction between SNPs in GABA/activin receptor genes and depression, *ACVR2B* was associated with depression in women. No previous associations between depression and *APOC3*, *ACVR2B*, or *GNRH1* have been reported.

That these genetic variants are associated with *neuroendocrine* pathways (*GnRH1, ACVR2B*) that are known to regulate *neurotransmitter* release and cognitive behavior (39-40) supports these associations as relevant to the etiology of depression and underlines the benefits of using RP to identify meaningful G x G interactions associated with disease.

Limitations

Given the numerous genetic, phenotypic and environmental influences that are linked to depression, and the small number of SNPs analyzed, it is not surprising that predictability from our models was low (although our predictability was superior to previous studies examining only single main factors). Also, the predictive value of our statistical models was further limited due to user bias in selection of SNPs (from nearly two-million SNPs in the human genome) used in this study. As a result of this, interactions we have found could potentially be moderated by another gene that we have not considered in this study. Nonetheless, we identified significant G x G interactions between known, and newly identified, loci associated with depression. Importantly, 4 of the 7 SNPs identified in these interactions were primarily located in two pathways well-known to impact depression: neurotransmitter and neuroendocrine signaling.

The results from the RP analyses conducted in this study were confirmed by LR, demonstrating the utility of RP as a screening tool for identifying meaningful G x G interactions. Future development of algorithms for RP analysis should not only maximize the distance between branches of the next best split (i.e. rpart), but consider subsequent future split combinations that could potentially result in trees with "better" overall predictability.

Summary

Our data indicate that G x G interaction analyses allows for enhanced predictability of conditions and diseases of aging. RP is an efficient and powerful exploratory analysis technique for elucidating G x G interactions in large datasets and combined with LR provides an important statistical analysis for the identification of well supported G x G interactions. We predict that such analytical methods will play an increasingly important role in the identification of epistatic effects in future large GWAS. Finally, our studies illustrate how RP analyses can be used to find interacting pathways involved in the etiology of a disease or condition such as depression.

ACKNOWLEDGMENTS

This research uses data from the Wisconsin Longitudinal Study (WLS) of the University of Wisconsin-Madison. Since 1991, the WLS has been supported principally by the National Institute on Aging (AG-9775, AG-21079 and AG-033285), with additional support from the Vilas Estate Trust, the National Science Foundation, the Spencer Foundation, and the Graduate School of the University of Wisconsin-Madison. A public use file of data from the Wisconsin Longitudinal Study is available from the Wisconsin Longitudinal Study, University of Wisconsin-Madison, 1180 Observatory Drive, Madison, Wisconsin 53706 and at http://www.ssc.wisc.edu/wlsresearch/data/. This material is the result of work supported with resources at the William S. Middleton Memorial Veterans Hospital, Madison, WI. The opinions expressed herein are those of the authors. The contents do not represent the views of the Dept. of Veterans Affairs or the United States Government.

FIGURE LEGENDS

Figure 1. Recursive Partitioning Tree of CIDI-SF Depression in Males of the WLS. Upper and lower numbers in nodes represent the percentage of participants with depression and the number of controls/cases in that node, respectively. Blue and purple boxes/circles indicate lower and higher rates of depression relative to the primary node, respectively. Split information indicates gene, SNP, and genotype criteria, respectively. M1 is subset of data referenced in Table 1. Sensitivity: 0.526, Specificity: 0.598, Accuracy: 0.591. Due to missing genotype information, we lose approximately 1.5% of participants per split. *rs1800497 is historically known as the *DRD2* Taq1A allele

Figure 2. Recursive Partitioning Tree of CIDI-SF Depression in Females of the WLS. Upper and lower numbers in nodes represent the percentage of participants with depression and the number of controls/cases in that node, respectively. Blue and purple boxes/circles indicate lower and higher rates of depression relative to the primary node, respectively. Split information indicates gene, SNP, and genotype criteria, respectively. F1-F4 are subsets referenced in Table 1. Sensitivity: 0.607, Specificity: 0.563, Accuracy: 0.572. Due to missing genotype information, we lose approximately 1.4% of participants per split.

Table 1.
Single-factor logistic regression models based directly off male and female RP tree split criteria (see Figures 1 & 2). Each SNP split was first run on the full dataset to represent single main factor effects ("Full Data") for both males and females. Then the same SNP splits were run on specific subsets of data per RP tree splits (M1, F1-F4; "RP-Subsetted Data").

					Full Data		RP-Subsetted Data		
Gender	RP Split	Gene	SNP	Genotypes	OR (95% CI)	P-value	Subset	OR (95% CI)	P-value
Male	Primary	*ANKK1**	rs1800497	T/T vs. C/C + C/T	2.55 (1.44-4.51)	0.001 *	--	----	----
	Left	*GNRH1*	novel SNP	T/T vs. C/C + T/C	1.43 (1.09-1.88)	0.011	M1	1.57 (1.18-2.08)	0.002 *
Female	Primary	*DRD2*	rs2242592	C/C + T/C vs. T/T	1.32 (1.08-1.62)	0.006 *	--	----	----
	Left 1	*APOC3*	rs2854116	T/T vs. C/C + T/C	1.28 (1.04-1.57)	0.018	F1	1.55 (1.15-2.09)	0.004 *
	Left 2	*ACVR2B*	rs3749386	C/C + T/T vs. T/C	1.11 (0.91-1.36)	0.302	F2	2.17 (1.37-3.44)	0.001 *
	Right 1	*FTO*	rs1421085	C/C + T/T vs. T/C	1.32 (1.08-1.62)	0.007 *	F3	1.65 (1.24-2.18)	0.0005 *
	Right 2	*IL6*	rs1800795	C/C + G/G vs. C/G	1.12 (0.92-1.37)	0.269	F4	1.85 (1.19-2.89)	0.006 *

RP, recursive partitioning; OR, odds ratio; CI, confidence interval
M1: LR analysis was run for only those with genotype *DRD2* rs1800497 C/C or C/T
F1: LR analysis was run for only those with genotype *DRD2* rs2242592 T/T
F2: LR analysis was run for only those with genotypes *DRD2* rs2242592 T/T and *APOC3* rs2854116 T/T
F3: LR analysis was run for only those with genotype *DRD2* rs2242592 C/C or T/C
F4: LR analysis was run for only those with genotypes *DRD2* rs2242592 C/C or T/C and *FTO* rs1421085 T/C
*rs1800497 is historically known as the *DRD2* Taq1A allele

REFERENCES

1. CDC. Anxiety and Depression. Atlanta, GA: Centers for Disease Control and Prevention; 2009 [cited 2010 Oct. 18]; Available from: http://www.cdc.gov/Features/dsBRFSSDepressionAnxiety/.

2. Eley TC, Sugden K, Corsico A, Gregory AM, Sham P, McGuffin P, et al. Gene-environment interaction analysis of serotonin system markers with adolescent depression. Mol Psychiatry. 2004;9(10):908-15.

3. Mirowsky J, Ross CE. Age and Depression. Journal of Health and Social Behavior. 1992;33(3):187-205.

4. Zubenko GS, Zubenko WN, Spiker DG, Giles DE, Kaplan BB. Malignancy of recurrent, early-onset major depression: a family study. Am J Med Genet. 2001 Dec 8;105(8):690-9.

5. Raymer KA, Waters RF, Price CR. Proposed multigenic Composite Inheritance in major depression. Medical Hypotheses. 2005;65(1):158-72.

6. Harold D, Abraham R, Hollingworth P, Sims R, Gerrish A, Hamshere ML, et al. Genome-wide association study identifies variants at CLU and PICALM associated with Alzheimer's disease. Nat Genet. 2009 Oct;41(10):1088-93.

7. Haasl R, Ahmadi MR, Meethal S, Gleason C, Johnson S, Asthana S, et al. A luteinizing hormone receptor intronic variant is significantly associated with decreased risk of Alzheimer's disease in males carrying an apolipoprotein E epsilon4 allele. BMC Medical Genetics. 2008;9(1):37.

8. Basson J, Wilson A, Haasl R, Atwood C. Multi locus interactions in steroidogenic pathway genes predict Alzheimer's disease. Alzheimer's and Dementia. 2008;4(4 (Suppl. 2)):T579: P3-199.

9. Pezawas L, Meyer-Lindenberg A, Goldman AL, Verchinski BA, Chen G, Kolachana BS, et al. Evidence of biologic epistasis between BDNF and SLC6A4 and implications for depression. Mol Psychiatry. 2008;13(7):709-16.

10. Zhang K, Yang C, Xu Y, Sun N, Yang H, Liu J, et al. Genetic association of the interaction between the BDNF and GSK3B genes and major depressive disorder in a Chinese population. Journal of Neural Transmission. 2010;117(3):393-401.

11. Lin E, Hong CJ, Hwang JP, Liou YJ, Yang CH, Cheng D, et al. Gene-gene interactions of the brain-derived neurotrophic-factor and neurotrophic tyrosine kinase receptor 2 genes in geriatric depression. Rejuvenation Res. 2009;12(6):387-93.

12. Neff CD, Abkevich V, Potter J, Riley R, Shattuck D, Katz DA. Evidence for epistasis between SLC6A4 and a chromosome 4 gene as risk factors in major depression. American Journal of Medical Genetics Part B: Neuropsychiatric Genetics. 2010;153B(1):321-2.

13. Wong ML, Dong C, Andreev V, Arcos-Burgos M, Licinio J. Prediction of susceptibility to major depression by a model of interactions of multiple functional genetic variants and environmental factors. Mol Psychiatry. 2012 Mar 27.

14. Carr D, Khodyakov D. End-of-life health care planning among young-old adults: An assessment of psychosocial influences. J Gerontol B-Psychol. 2007 Mar;62(2):S135-S41.

15. Sewell WH. As We Age : The Wisconsin Longitudinal Study, 1957-2001: Center for Demography and Ecology University of Wisconsin--Madison; 2001.

16. Nelson CB, Kessler RC, Mroczek D. Scoring the World Health Organization's Composite International Diagnostic Interview Short Form (CIDI-SF; v1.0 NOV98). 1998.

17. Zhang H, Bonney G. Use of classification trees for association studies. Genet Epidemiol. 2000 Dec;19(4):323-32.

18. Breiman L, Friedman J, Olshen R, Stone C. Classification and Regression Trees. Boca Raton: Chapman and Hall/CRC; 1984.

19. Therneau T, Atkinson E, editors. Technical Report Series No. 61, An Introduction to Recursive Partitioning Using the RPART Routines. Rochester, MN: Department of Health Science Research, Mayo Clinic; 1997.

20. Accortt EE, Freeman MP, Allen JJB. Women and Major Depressive Disorder: Clinical Perspectives on Causal Pathways. Journal of Women's Health. 2008;17(10):1583-90.

21. Kessler RC, McGonagle KA, Swartz M, Blazer DG, Nelson CB. Sex and depression in the National Comorbidity Survey. I: Lifetime prevalence, chronicity and recurrence. J Affect Disord. 1993 Oct-Nov;29(2-3):85-96.

22. Kessler RC, McGonagle KA, Nelson CB, Hughes M, Swartz M, Blazer DG. Sex and depression in the National Comorbidity Survey. II: Cohort effects. J Affect Disord. 1994 Jan;30(1):15-26.

23. Kessler RC, McGonagle KA, Zhao S, Nelson CB, Hughes M, Eshleman S, et al. Lifetime and 12-month prevalence of DSM-III-R psychiatric disorders in the United States. Results from the National Comorbidity Survey. Arch Gen Psychiatry. 1994 Jan;51(1):8-19.

24. Weissman MM, Bland R, Joyce PR, Newman S, Wells JE, Wittchen HU. Sex differences in rates of depression: cross-national perspectives. J Affect Disord. 1993 Oct-Nov;29(2-3):77-84.

25. Weissman MM, Olfson M. Depression in women: implications for health care research. Science. 1995 Aug 11;269(5225):799-801.

26. Pajer K. New strategies in the treatment of depression in women. The Journal of clinical psychiatry. [Review]. 1995;56 Suppl 2:30-7.

27. Paykel ES. Depression in women. Br J Psychiatry Suppl. [Review]. 1991 May(10):22-9.

28. Shi M, Weinberg CR. How Much Are We Missing in SNP-by-SNP Analyses of Genome-wide Association Studies? Epidemiology. 2011 Nov;22(6):845-7.

29. Lawford BR, Young R, Noble EP, Kann B, Ritchie T. The D2 dopamine receptor (DRD2) gene is associated with co-morbid depression, anxiety and social dysfunction in untreated veterans with post-traumatic stress disorder. European Psychiatry. 2006;21(3):180-5.

30. Klimek V, Schenck JE, Han H, Stockmeier CA, Ordway GA. Dopaminergic abnormalities in amygdaloid nuclei in major depression: a postmortem study. Biol Psychiatry. 2002 Oct 1;52(7):740-8.

31. Schneier FR, Liebowitz MR, Abi-Dargham A, Zea-Ponce Y, Lin SH, Laruelle M. Low dopamine D(2) receptor binding potential in social phobia. Am J Psychiatry. 2000 Mar;157(3):457-9.

32. Noble EP, Blum K, Ritchie T, Montgomery A, Sheridan PJ. Allelic Association of the D2 Dopamine Receptor Gene With Receptor-Binding Characteristics in Alcoholism or Gene ism. Arch Gen Psychiatry. 1991 July 1, 1991;48(7):648-54.

33. Thompson J, Thomas N, Singleton A, Piggott M, Lloyd S, Perry EK, et al. D2 dopamine receptor gene (DRD2) Taq1 A polymorphism: reduced dopamine D2 receptor binding in the human striatum associated with the A1 allele. Pharmacogenetics. 1997 Dec;7(6):479-84.

34. Pohjalainen T, Rinne JO, Nagren K, Lehikoinen P, Anttila K, Syvalahti EK, et al. The A1 allele of the human D2 dopamine receptor gene predicts low D2 receptor availability in healthy volunteers. Mol Psychiatry. 1998 May;3(3):256-60.

35. Jonsson EG, Nothen MM, Grunhage F, Farde L, Nakashima Y, Propping P, et al. Polymorphisms in the dopamine D2 receptor gene and their relationships to striatal dopamine receptor density of healthy volunteers. Mol Psychiatry. 1999 May;4(3):290-6.

36. Hayden EP, Klein DN, Dougherty LR, Olino TM, Laptook RS, Dyson MW, et al. The dopamine D2 receptor gene and depressive and anxious symptoms in childhood: associations and evidence for gene-environment correlation and gene-environment interaction. Psychiatr Genet. 2010 Dec;20(6):304-10.

37. Dubertret C, Bardel C, Ramoz N, Martin P-M, Deybach J-C, Adès J, et al. A genetic schizophrenia-susceptibility region located between the ANKK1 and DRD2 genes. Progress in Neuro-Psychopharmacology and Biological Psychiatry. 2010;34(3):492-9.

38. Kivimaki M, Jokela M, Hamer M, Geddes J, Ebmeier K, Kumari M, et al. Examining Overweight and Obesity as Risk Factors for Common Mental Disorders Using Fat Mass and Obesity-Associated (FTO) Genotype-Instrumented Analysis. Am J Epidemiol. 2011 Feb 15;173(4):421-9.

39. Bull SJ, Huezo-Diaz P, Binder EB, Cubells JF, Ranjith G, Maddock C, et al. Functional polymorphisms in the interleukin-6 and serotonin transporter genes, and depression and fatigue induced by interferon-alpha and ribavirin treatment. Mol Psychiatry. 2009 Dec;14(12):1095-104.

40. Zheng F, Adelsberger H, Muller MR, Fritschy JM, Werner S, Alzheimer C. Activin tunes GABAergic neurotransmission and modulates anxiety-like behavior. Mol Psychiatry. 2008;14(3):332-46.