TECHNISCHE UNIVERSITÄT MÜNCHEN

Lehrstuhl für Entwicklungsgenetik

A Candidate Gene and a Genome-wide Association Study in Depression

I A Candidate Gene Study Reveals Variants in NTRK2 Conferring Risk to Mood Disorder Patients for a Life History of Attempted Suicide

II A Genome-Wide Study Identifies SLC6A15 as a Novel Susceptibility Gene for Major Depression

Martin A. Kohli

Vollständiger Abdruck der von der Fakultät Wissenschaftszentrum Weihenstephan für Ernährung, Landnutzung und Umwelt der Technischen Universität München zur Erlangung des akademischen Grades eines

Doktors der Naturwissenschaften

genehmigten Dissertation.

Vorsitzender: Univ.-Prof. Dr. E. Grill

Prüfer der Dissertation:

1. Univ.-Prof. Dr. W. Wurst
2. Univ.-Prof. Dr. Chr.-C. Schön

Die Dissertation wurde am 04.08.2008 an der Technischen Universität München eingereicht und durch die Fakultät Wissenschaftszentrum Weihenstephan für Ernährung, Landnutzung und Umwelt am 19.03.2009 angenommen.

CONTENTS

ABSTRACT

Genetic association analysis is a powerful tool to unravel genetic risk variants of human disease. The identification of genes involved in disease etiology allows investigations of pathologies on a molecular basis, which might provide a rational strategy for future drug development.

Major depressive disorder (MDD) is a common psychiatric disease with a lifetime prevalence rate of 15-17%. Its heritability has been estimated to range from 34 to 75%. The defining features of MDD are marked and persistent dysphoria plus additional cognitive and physiological symptoms. Moreover, suicidal behavior is most frequent in mood disorder patients, but seems to be inherited independently from psychiatric diagnosis. Previous attempts to identify susceptibility genes for MDD or suicidal behavior revealed several genes, however, only few of these initial reports have been confirmed unambiguously by subsequent studies or in meta-analyzes. Phenotypic diversity and genetic heterogeneity as well as the large environmental contribution inherent to this disorder have been considered to represent major obstacles for the identification of causative variants.

In the past, genetic association analysis was technically constrained to candidate gene approaches, whereas today hypothesis-free genome-wide association studies promise hope for the identification of genes so far unrelated to human disease together with their regulatory regions sometimes far apart from protein coding sequences.

I first conducted a candidate gene association study to further validate a hypothesis derived from clinical, post-mortem gene expression and pharmacological animal studies on the contribution of neurotrophins to the pathogenesis of MDD and suicidal behavior by extensively covering the brain-derived neurotrophic factor (*BDNF) gene* and its high-affinity receptor gene, neurotrophic tyrosine kinase receptor 2 (*NTRK2)* with single nucleotide polymorphic (SNP) markers. Although there was no evidence for association of these two genes with MDD in this study, several independent variants in *NTRK2*

could be related to lifetime attempted suicide (SA) in mood disorder patients of two independent German samples. Intriguingly, two variants showed gene-gene interaction to predict SA and patients carrying all risk genotypes attempted suicide 4.5 times more frequently than those bearing only protective genotypes. The second part of this doctoral thesis presents data from one of the first hypothesis-free genome-wide case-control association studies (GWAS) on MDD. An experiment-wide significant result for a variant on chromosome 12q21.31 was obtained and could be replicated together with other correlated SNPs in another independent German sample (combined N=1270/1388). The closest annotated gene to the region of association is *SLC6A15*, which belongs to the solute carrier 6 gene family and codes for a sodium-dependent branched-chain amino acid transporter. *SLC6A15* is expressed in neurons of the brain with highest expression in the hippocampus. GENEVAR expression array data from lymphoblastoid cell lines (N=210) showed that MDD risk allele carriers exhibit lower *SLC6A15* mRNA expression levels than individuals homozygous for the protective allele. The relevance of *SLC6A15* for MDD could further be supported by a mouse model of chronic social stress, which models relevant features of MDD. *SLC6A15* mRNA levels were reduced in the hippocampus of stress susceptible compared to stress resistant animals. Moreover, nuclear magnet resonance spectroscopy (^{1}H-NMR) revealed lower N-acetylaspartate (NAA) and glutamate/glutamine (Glx) levels in the hippocampus of healthy controls (N=80) carrying the MDD risk alleles compared to protective allele homozygotes. Higher NAA levels constitute a functional correlate for neuronal integrity and decreased NAA levels have been reported in patients with bipolar and unipolar depressive disorder. In the here presented study, these findings could be confirmed in unipolar depressed patients. In addition, data show a risk genotype by diagnosis interaction on hippocampal volumes. Depressed patients carrying the risk genotypes showed smaller hippocampal volumes than controls. This genotype-specific effect was not observed in healthy controls. In conclusion, these results support *SLC6A15* as a novel candidate gene for MDD. This work shows on the one hand the potential of hypothesis-driven genetic association analyses to further validate a well established candidate pathway,

which has previously been defined by findings in post-mortem, clinical studies and functional experiments in animals. On the other hand, this work also describes the discovery of a novel candidate gene by a hypothesis-free genome-wide association approach with a subsequent validation of the found candidate gene by diverse functional approaches.

ZUSAMMENFASSUNG

Assoziationsstudien vermögen genetische Risikovarianten für humane Erkrankungen aufzudecken. Die Identifikation von Genen, die eine Rolle in der Ätiologie einer Krankheit spielen, ermöglicht die Erforschung der molekularen Grundlagen dieser Erkrankung, was einer rational-geleiteten Medikamenten-Entwicklung zu neuen Strategien verhelfen mag.

Unipolare Depression (UD) ist eine häufige psychiatrische Erkrankung mit einer Lebenszeitprävalenz von 15-17%. Der erbliche Anteil für UD wird auf 34-75% geschätzt. Die Merkmale der Erkrankung sind ausgeprägte und persistierende Dysphorie zusammen mit anderen kognitiven Beeinträchtigungen und auch körperlichen Symptomen. Suizidales Verhalten kann während einer depressiven Episode sehr ausgeprägt sein. Dennoch scheint Suizidalität unabhängig von Depression vererbt zu werden.

Bisherige Versuche Suszeptibilitätsgene für UD oder Suizidalität zu finden, waren nur zum Teil erfolgreich und erste Befunde oft schwierig zu reproduzieren, so dass die Datenlage für einzelne Gene meist noch zu inkomplett ist, als daß Meta-Analyzen mehr Aufschluß versprächen. Ursachen dafür scheinen einerseits phänotypische wie genetische Heterogenität und andererseits ein beträchtlicher Beitrag von Umweltfaktoren an der Entstehung der Depression zu sein oder daß die wesentlichen Gene noch nicht mit UD oder Suizidalität in Zusammenhang gebracht werden konnten.

Bis vor kurzem waren in der Assoziationsgenetik nur Kanditatengenansätze möglich. Hypothesenfreie genomweite Studien könnten ganz neue Kandidatengene identifizieren. Darüber hinaus vermögen sie auch regulatorische

Sequenzen weit ab von kodierenden Regionen zu identifizieren, die in Kandidatengenansätzen meist nicht berücksichtigt werden. Im Lichte des noch sehr inkompletten Bildes über die Biologie von Depression und Suizidalität erscheinen Hypothese-freie Ansätze als wegweisend.

In dieser Arbeit wurde zuerst in einem Kandidatengenansatz die durch klinische, postmortem Expressions- und Tier-experimentellen Studien etablierte Hypothese einer Beteiligung von Neurotrophinen in der Ätiologie von UD und suizidalem Verhalten auf suszeptibiliätsgenetischer Ebene überprüft. Das *BDNF*-Gen zusammen mit dem dazugehörigen hoch-affinen Rezeptorgen für *BDNF*, *NTRK2*, wurden dazu mit einer hohen SNP-Markerdichte abgedeckt (*BDNF*: brain-derived neurotrohic factor, *NTRK2*: neurotrophin tyrosine kinase receptor 2, SNP: single nucleotide polymorphism).

Obwohl keine Assoziation zwischen genotypisierten SNPs und UD gefunden wurden, waren mehrere nicht korrelierte *NTRK2*-SNPs in zwei unabhängigen Studien mit Lebenszeit-Suizidversuch (SV) in depressiven Patienten assoziiert. Interessanterweise zeigten zwei SNPs eine Interaktion auf SV. Patienten, die Risikogenotypen aller assoziierten SNPs trugen, begingen 4.5-mal häufiger einen SV als Patienten, die keine Risikogenotypen aufwiesen.

Der zweite Teil dieser Doktorarbeit präsentiert Daten aus einer der ersten genomweiten Assoziationsstudien in UD. Ein SNP auf Chromosom 12q21.31 zeigte Experiment-weite Signifikanz und konnte in einer weiteren Studie zusammen mit anderen korrelierten SNPs repliziert werden (insgesamt 1270 Patienten und 1388 Kontrollen). Das nächstgelegene Gen zu dieser Region ist *SLC6A15*, das für einen Natriumionen-abhängigen, verzweigtkettigen Aminosäurentransporter kodiert. *SLC6A15* ist im Gehirn ausschliesslich neuronal und am stärksten im Hippokampus exprimiert. GENEVAR-Expressionsdaten aus humanen lymphoblastoiden Zelllinien von gesunden Probanden aus verschiedenen Populationen (N=210) zeigten geringere *SLC6A15*-mRNA-Expression in UD-Risikoallelträgern auf als Homozygote für das protektive Allel. Chronischer Stress ist ein Risikofaktor für UD. In diesem Zusammenhang konnte die Relevanz von *SLC6A15* für UD in einem Mausmodell zu chronischem Stress erfolgreich validiert werden. Stress-suszeptible Tiere zeigten

verminderte *SLC6A15*-mRNA-Expression im Hippokampus auf als Stress-resistente Mäuse. Zusätzlich konnte mittels Magnet-Resonanz-Spektroskopie in gesunde Probanden (N=80) nachgewiesen werden, daß Risikoallelträger niedrigere N-Azetylaspartat- (NAA) und Glutamat/Glutamin-Spiegel im Hippokampus aufwiesen. Hohe NAA-Werte sind ein Maß für neuronale Integrität und ein erniedrigter NAA-Spiegel im Hippokampus ist sowohl von bipolar wie unipolar depressiven Patienten aus der wissenschaftlichen Literatur bekannt. In dieser Studie konnte dieser Zusammenhang für unipolar depessive Patienten bestätigt werden. Zudem konnte eine Interaktion zwischen Risikogenotypen und Diagnose auf das Hippokampusvolumen nachgewiesen werden. Nur an UD erkrankte Risikovariantenträger hatten kleinere Hippokampi als Homozygote für das protektive Allel. In gesunden Kontrollen kam dieser Genotyp-abhängige Effekt nicht zum tragen. Insgesamt sprechen diese Befunde dafür, daß *SLC6A15* ein neues Suszeptibilitätsgen für UD sein könnte.

Diese Arbeit zeigt also in jeweils erfolgreichen Beispielen, daß Assoziationsanalyzen es erlauben, Kandidatengen-basierte Hypothesen zu validieren, die sich aus vorausgegangenen Postmortem-, Expressions- oder Tier-experimentellen Studien aufgedrängt haben oder in Hypothese-freien genomweiten Studien Kandidatengene zu identifizieren, die nachfolgend mit funktionalen Methoden wiederum erhärtet werden können.

1 INTRODUCTION

1.1 Human Genetic Association in Complex Disease

1.1.1 Overview

A prerequisite for human disease gene mapping is a reasonably high heritable contribution to the overall risk for the disease of interest. The heritability of a disease can be estimated by twin, adoption and epidemiological family studies (*1*). A second prerequisite is a dense set of genetic marker loci spanning the human genome. Today, the most commonly used markers for gene mapping are single nucleotide polymorphisms (SNPs). SNPs are almost always bi-allelic markers, which are evenly distributed throughout an organism's genome at a very high number. For instance, 6.57 million validated SNPs were annotated in the recent human genome, (NCBI´s dbSNP build 129) (2). Thus, on average, there is about one validated SNP every 500 base pairs throughout the human genome (3).

Disease gene hunting can be done by genetic linkage analysis or genetic association studies. Linkage analysis refers to the ordering of genetic loci on a chromosome and to estimating genetic distances among them, where these distances are determined on the basis of a statistical phenomenon, i.e. the recombination (crossover) frequency occurring between two markers on a gamete. Because recombination events can be recognized only on the basis of haplotypes passed from parents to children, linkage analysis cannot be carried out with unrelated individuals but requires disease status information in relatives (4).

In 1996, Risch and Merikangas demonstrated that linkage analysis has limited power to detect risk alleles for common disease that are relatively frequent in the general population. Especially, when the effect size of a genetic risk variant is smaller than two odds ratios (OR) statistical power becomes critical (OR=2: double disease frequency in risk variant carriers). In fact, the authors

demonstrated that power considerations are in favor of association testing under such conditions (5, 6).

Genetic association testing is based on the linkage disequilibrium (LD) structure in an organism's genome, i.e. the genotype correlation structure between different loci on a chromosome. Given that a typed marker is in high LD with the disease causing variation, it is overrepresented in the patient group (cases) compared to a group of healthy subjects (controls). A significant accumulation of an allele (or genotype, haplotype) of genetic markers in cases compared to controls is referred as to genetic association between this allele and the disease under investigation. Association testing can be performed in familial data and between unrelated cases and controls. The affection status (case or control) is usually displayed in rows and the number of different alleles or genotypes per marker in columns of a so-called contingency table (a 2x2 or 2x3 cell table for allelic or genotypic association of a bi-allelic marker, respectively). A simple way to test for genetic association is performing a chi-square (χ^2) test of independence. If covariates, e.g. age or gender, might modulate the affection status, they can be included by logistic regression analysis. In the case of quantitative traits, analysis of variance with covariates (ANCOVA) might also be performed. Ordered sub-set analysis (OSA) in a case-control design with cases ordered by a quantitative trait, e.g. severity or age-at-onset of disease, may reveal susceptibility genes, which confers disease risk to a subgroup of patients (7).

Complex or multifactorial disease are characterized by the interplay of several genetic and environmental risk factors with low effect sizes, of which some might interact in complex manner (8). Among many others, commonly mentioned complex diseases are schizophrenia, major depression, obesity, diabetes mellitus or coronary heart disease. Risch and Merikangas illustrated that genetic association testing would be the method of choice to detect common variants with small effect sizes in common multifactorial disease (5, 6).

Genetic association can be hypothesis-driven, where candidate genes are assessed, which have already been brought in relation to the pathophysiology of an disease, for instance as a result of pharmocological and behavioral animal

models or blood, biopsy or post-mortem gene expression studies in humans. More recently, we have the possibility to also perform hypothesis-free genome-wide approaches. In 1996, Risch and Merikangas stated that genome-wide association studies (GWAS) are appealing in theory, but out of reach in practice. However, fast progress in high-throughput molecular biotechnology, i. e. genotyping technology enabled the first GWAS already in the first half of this decade challenging human genetic computation simply by the huge bulk of data, but also by means of statistical interpretation. Especially in genome-wide approaches, in which for instance 500'000 SNPs are genotyped, the number of false positive results is remarkable due to the high number of statistical tests performed in one experiment (9, 10).

1.1.2 Major factors influencing genetic association analysis

1.1.2.1 Linkage disequilibrium

Two or more polymorphic loci, e.g. SNPs, are in linkage disequilibrium (LD), when certain alleles of these loci are observed together more often than would be expected by chance in a given population (8). Thus, under random chromosomal segregation, a prerequisite for LD is physical co-occurrence of both SNPs one the same chromosome. The amounts of LD between any two markers is influenced by the classical forces of recombination, natural selection, mutation, genetic drift, ancestral population demographics and mating patterns (8, 11). Association mapping takes advantage of LD in the way that it is sufficient to type a maker, which is in LD with the unknown disease-causing variant.

Most commonly used measures for the strength of LD are pairwise for bi-allelic sites, e.g. SNPs. The most important ones are D' and r^2. Both measures are built on the pairwise disequilibrium coefficient, D, which is the difference between the probability of observing two marker alleles on the same haplotype and observing them independently in the population: $D = f(A_1B_1) - f(A_1) * f(B_1)$, where "A" and "B" refer to two genetic markers with alleles "1" and "2", and "f " is their frequency. Than $D' = D / D_{max}$, and $r^2 = D^2 / (f(A_1) * f(B_1) * f(A_2) * f(B_2))$ (8,

11). Both measures scales from 0 (no LD) to 1 (complete LD). D' is the less stringent measure and becomes 1, when three of the totally four possible haplotypes were observed. Particularly, when the rare allele of one marker occurs exclusively with just one of the two possible alleles of the other marker, D' is 1. This situation is called "complete LD". The measure r^2 represents the statistical correlation between two sites, and takes the value of 1, if only two haplotypes are present. In this case the alleles of both markers show the same frequency. If so, both sites are called to be in "perfect" LD. In this case, there was no historical recombination event between the two SNPs (12).

These two measures introduced above behave very differently, and high values of D' may not be inconsistent with low values of r^2, since r^2 becomes 1.0 only for SNPs with the same minor allele frequences (MAF), whereas D' can show a value of 1.0 for SNPs with unequal MAFs. In particular, there seems to be much more random variation in values of D' at a given recombination distance (11, 13, 14). D' tends toward a minimum value of ~ 0.4 for markers that are effectively independent, while r^2 tends towards 0.0 for independent SNPs (15).

R^2 is arguably the most relevant LD measure for association mapping, because there is a simple inverse relationship between r^2 and the sample size required to detect association between the unkown disease causing variant and the investigated marker. For instance, supposing that a known SNP marker was genotyped in a given number of cases and controls and that this marker is in LD to a certain degree with a so far unkown disease susceptibility locus, the sample size would need to be increased by a factor of $1/ r^2$(marker,risk variant) to achieve the same statistical power to detect the association as would have been reached if it had been possible to directly genotype the unkown risk variant (11, 13, 16, 17).

Patterns of LD are well known for being noisy and unpredictable. Many studies have shown that although the background LD is significantly related to genetic distance, small physical distance do not guarantee high level of LD (18-21). For example, pairs of sites that are tens of kilobases apart might be in complete LD, whereas nearby pairs of sites from the same region might be in weak LD (11, 14).

Genotyping of large numbers of SNPs in different human populations by the International HapMap Project (see also next paragraph) revealed that most of the human genome is contained in LD or haplotype blocks of substantial size (21). Genomic regions with high LD are interrupted by mostly smaller regions of considerably less LD, which often contain hotspots of recombination. It has been estimated that half of the human genome exists in blocks of 22 kb or larger in African and African-American samples and in blocks of 44 kb or larger in European and Asian samples. Within each block, a very small number of common haplotypes (three to five) typically capture about 90% of all chromosomes in each population. Both the boundaries of blocks and the specific haplotypes observed are shared to a remarkable extent across populations. This overall picture of the LD or haplotype block structure of the human genome among different population is supportive of a single "out of Africa" origin (22, 23) for both the European and Asian populations. This data suggest a considerable bottleneck in the ancestry of these two populations, with only a subset of the diversity (of SNPs, haplotypes, and recombinant chromosomes) found in Africa in the two non-African populations (21).

In consequence, the probability of any particular pair of SNPs being in LD should be determined empirically for any specific genomic region in any population (15, 18). The International HapMap Project sufficiently described the genome-wide LD structure in four distinct human populations by the year 2007 (24).

1.1.2.2 The International HapMap project – tagging SNPs became available

The International HapMap Project was launched in 2002 with the aim of providing a sufficiently accurate LD map of different human populations as a public resource to accelerate medical genetic research (24). The recently published Phase II HapMap data characterizes over 3.1 million human SNPs genotyped in a total of 270 individuals from four geographically diverse populations – Yoruba in Ibadan, Nigeria (YRI), Utah residents from northern and western European ancestry from the Centre d'Etude du Polymorphisme Humain (CEPH) collection (CEU), Han Chinese from Beijing, China (CHB) and Japanese from Tokyo, Japan (JPT). The resulting HapMap has a SNP density

of approximately one SNP per kilobase and is estimated to contain approximately 25–35% of all the 9–10 million common SNPs (MAF> 0.05) in the assembled human genome. It has been estimated that Phase II HapMap marker sets capture the overwhelming majority of all common variants at high r^2. For common variants the mean maximum r^2 of any SNP to a typed one is 0.90 in YRI, 0.96 in CEU and 0.95 in CHB&JPT. The prior knowledge of the LD structure in a given population and genomic region coming from HapMap data allows to set up an optimized marker panel with regards to genetic information coverage in an association study by avoiding redundant SNPs. This allows to cut down genotyping costs and labor. A SNP marker being in maximal overall LD with a set of other SNPs can be used as a surrogate marker to cover the information content of all SNPs of this SNP set. Such a SNP is commonly called *tagging SNP (tagSNP)*. Several algorithms and computer programs have been developed to define sets of tagSNPs. For association studies, it was demonstrated that tagSNPs perform substantially better than randomly spaced SNPs which were chosen without considering their inter-dependency (25-27).
Using a simple pairwise tagging approach, the HapMap Consortium showed that 1.09 million SNPs are required to capture all common Phase II SNPs with a $r^2 \geq 0.8$ in YRI, which is more than the 500'000 SNPs required in CEU or CHB+JPT. This is already similar to the estimated number of tagSNPs required to adequately cover the information content of all possible common SNPs in the human genome (14, 17, 21).

1.1.2.3 The effect size

The effect size of a disease susceptibility factor also strongly modulates statistical power of an association study. The more a single genetic risk factor predicts disease status, the easier it is, to detect it. The effect size of a single genetic risk factor can be assessed by the *genotypic relative risk (GRR)*. For bi-allelic polymorphisms like SNPs, the GRR is expressed the following way: Given that the mutant allele (M) and the normal allele (N) with their frequencies m and n in the general population, and the event that an individual became affected by the disease (D), the probabilities of disease conditional on the genotypes of the SNP, can be defined as f_2 = P(D I MM), f_1 = P(D I MN) and f_o =

P(D I NN), whereas the subscript "j" of "*f* " indicates the number of copies of M. Then, the genotypic relative risks are f_1/f_o and f_2/f_o, which measure the relative increase in disease probabilities for heterozygous MN individuals and homozygous MM individuals, respectively, when compared to the probability for homozygous NN individuals (28). The GRR most commonly is not expressed by probability (*f*), but by *odds ratios (OR)*. OR are defined by the odds of exposure to the risk genotype among cases divided by the odds of exposure to the risk genotype among controls, where odds is equal to *f* /(1- *f).* For instance, the GRR of individuals which carry both risk alleles (MM homozygotes) relative to homozygotes without the risk allele expressed in OR is calculated the following: The number of MM homozygotes divided by the number of NN homozygotes in cases (odds in cases) divided by the analogous quotient in controls (odds in controls). Or, as probability: f_2/f_o = OR / (1+OR). The interpretation of the OR is simple. If it is 1, there is no GRR, if it is 2, the GRR is doubled. If the lower 95% confidence intervals of the OR is greater than 1, the GRR becomes statistically significant at a significance level of 5%. If it is smaller than 1, the M allele is protective from disease (29).

The effect size assessed by an association study depends on four parameters: the true relative risk of the disease causing allele, the extend of LD between marker and disease allele, the marker allele frequency and the disease allele frequency (30).

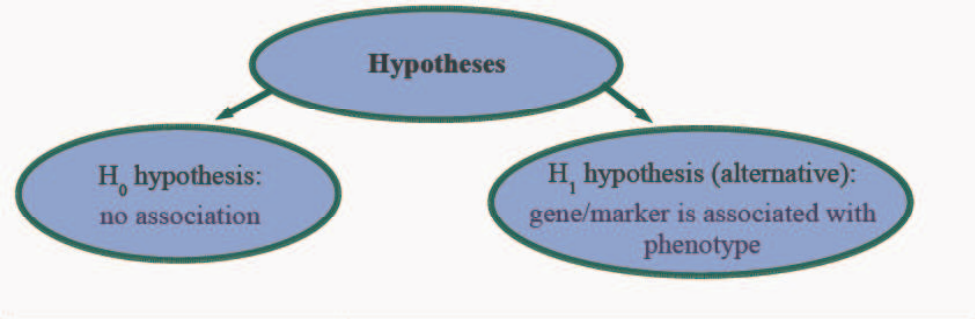
The GRR in complex disease for common SNPs has been shown to be mostly in the range of 1.15–1.5, with the majority in the range of 1.15–1.25. These data are derived from a large meta-analysis of reproducible case-control associations in various complex diseases as diabetes type I and II, hypertension, coronary artery disease, myocardial infarction, Alzheimer's disease, bipolar disorder, schizophrenia, and different types of cancer. These effect sizes are certainly small and need huge sample sizes in order to detect them, especially in a genome-wide study. Of course, there are examples of relative risks being greater than 2, also in complex disease, but most probably they will be the exceptions (31, 32). For instance, in age-related macular degeneration (AMD), which was described to be a complex genetic disorder,

recent genome-wide and candidate-gene studies identified an exonic, non-synonymous SNP in the complement factor H (CFH) gene, for which meta-analysis indicated a multiplicative model with each C allele increasing the odds of AMD by about 2.5-fold (33).

1.1.2.4 Multiple testing

Most researchers use the mentioned test statistics (as described in chapter 1.1.1) to decide whether a genetic variation is associated with a disease or not. Inherent to test statistics is a so-called *multiple testing problem* as soon as more than one hypothesis is tested. This can be apprehended by considering traditional test statistical theory, especially when focusing on its possible errors. The multiple testing problem arises from the type I (or α) error. An α error is made when a true null-hypothesis is rejected. Or, to put it differently, if a variant is declared to be associated with a disease, when in reality it is not. Supposing we have m null-hypotheses H_0, of which m_0 are true. And, A is the number of accepted an R the number of rejected H_0. A and R are observable random variables; U, V, S and T are unobservable random variables and correspond to the four possible cases of a test statistical decision process (fig. 1). The overall type I error rate, the family-wise type I or false positive error rate, can be described as V/m. By scientific tradition the type I error rate is most commonly set to 0.05 in order to call a result significant. This corresponds to a probability for an falsely accepted H_1 lower than 0.05 to be explained by chance. As long as just one hypothesis is tested (m=1) this claim is met. Whenever more than one hypothesis is tested the significance level must be lowered in order to prevent an inflation of the type I error rate. Or, in other words, the more tests are performed, the more false positive results are obtained at a significance level of 0.05.

Empirical data of genome-wide association studies seem to fit well with these theoretical derivations. For instance, in the genome-wide association study, which is a part of this dissertation, I tested 365'676 SNPs and observed 19'229 SNPs with a nominal association p-value below 0.05 (allelic test). This number is close to the randomly expected number of 18'284 (= 365'676 * 0.05) assuming independence among the tests/SNPs.



	Declared non-significant	Declared significant	Total
True H_0	U	V	m_0
Non-true H_0	T	S	$m - m_0$
	A	R	m

m: number of tested hypotheses
m_0: number of true H_0
U,V,T,S: unkown numbers of theoretically possible assignments
A: number of accepted H_0
R: number of rejected H_0

critical are:

V / m = type I (or α) error rate or rate of false positive results
T / m = type II (or β) error rate or rate of false negative results

Figure 1 Schematic illustration on test statistical theory, i.e. the type I and II error. With an increasing number of tested hypotheses m, there is an increasing number of expected false positive results. This necessitates correction for multiple comparisons.

A simple, but robust way to correct for multiple testing is the so-called Bonferroni correction. Each p-value is multiplied by the number of performed independent tests and assessed at a significance level of 0.05. Or, the significance level of 0.05 is divided by the number of performed tests to obtain a significance level adapted to multiple comparisons. Hence, in our example the Bonferroni-corrected genome-wide significance level would be 0.05/365'676, which is equal to 1.37e-07. As was introduced in chapter 1.1.2.1, SNPs on cetain genomic regions are often correlated, thus not independent. Moreover, I applied different genetic models (see *methods*) which are also not independent form each other. Therefore, by using a Bonferroni-correction for multiple testing I would have corrected in an overly conservative way.

There are many different multiple testing procedures which try to take correlation structures among performed tests into account. The methods of choice are those based on re-sampling (permutation) (12). In this work I used the permutation-based minimum p method proposed by Westfall and Young (34, 35). According to this method, all test statistics of interest are calculated in a given dataset of m variables. For example, for each of the 365'676 SNPs a case-control association p-value is calculated and the obtained p-values are ordered from the lowest to the highest one. Then, in the same dataset the affection states (patient or control) are randomly re-assigned to each individual's genotype. By doing so, LD between the SNPs is unaltered, so that the correlation structure of the actual dataset is maintained. After this re-sampling step, again association p-values for all SNPs are calculated and aligned in an ascending way. This permutation step with subsequent re-calculation of the test-statistics is repeated n times, for instance 100'000 times. The p-value corrected for multiple comparisons ($p_{1,cor}$) for the smallest observed nominal p-value ($p_{1,nom}$) of the original dataset is equal to the sum of smallest p-values ($_xp_{1,per}$) from all permutations n which are smaller than $p_{1,nom}$ divided by n, or formally: $p_{1,cor} = {}^n\sum_{x=1} ({}_xp_{1,per} \mid {}_xp_{1,per} < p_{1,nom}) / n$. For the second smallest observed p-value $p_{2,nom}$ the corrected p-value $p_{2,cor}$ is calculated the following: $p_{2,cor} = {}^n\sum_{x=1} ({}_xp_{2,per} \mid {}_xp_{2,per} < p_{2,nom}) / n$, and analogously for $p_{3,cor}$, $p_{4,cor}$, …, $p_{m,cor}$. The advantage of this procedure is that it reflects not only the correlation structure of an empirical dataset, but also its degree of random fluctuation.

1.1.2.5 Disease and marker allele frequencies

Rare markers generally provide less statistical power than frequent ones and require very large sample size to be reliably detected (30). Thus, in association studies, SNPs with MAFs below 5%, and for sure below 1%, provide inadequate power (36-38) (fig.2), unless huge sample sizes are available like those of the Type 1 Diabetes Genetics Consortium which recently presented results of association analyzes in about 30'000 individuals (39).

1.1.2.6 Misclassification Errors

Misclassification errors may potentially inflate the type I and II error rate in genetic studies. On the phenotyping side, a wrong assignment of probands either to cases or controls would lead to a misclassification error. Especially in highly prevalent diseases with a divers and complex catalogue of symptoms, phenotyping might be difficult and prone to misclassification errors. Furthermore, when many genetic risk factors with low penetrance are supposed and large sample sizes are required which prohibit extensive phenotyping, phenotyping errors may be more common and could further decrease the power of genetic association studies. Beside phenotypic misclassification, genotyping errors have been reported to be another potential source of misclassification in genetic studies. For family based data, genotyping errors can increase both type I and II errors (40, 41). For population-based studies, i.e. case-control association, genotyping errors can increase type II errors and thereby decrease power (42, 43). When genotyping does not occur blinded to case-control status in a randomly assigned sample-to-well or -plate design, the risk for inflated type I error rates might be high. Moreover, false positive association has been reported also under randomized experimental settings, probably due to disproportional genotyping errors among cases and controls introduced by chance (44). Additionally, genotyping errors can bias linkage disequilibrium (LD) measurements (45).

Genotyping errors arise from technical failures. But even under optimal experimental conditions genotyping errors occur because different genomic SNP target sequences vary widely in their ease of genotypic discrimination. For instance, unfavorable SNP assays can lead to a systematic misclassification of heterozygotes to one of the two alternative homozygotes. Another possible observation is an additional genotype clusters due to underlying copy number polymorphisms or SNPs in the region of PCR primer binding. Automated genotype calling algorithms might ignore the cluster most distant from predefined genotype cluster coordinates. Such and other genotype calling failure will not necessarily be independent of genotype class, a phenomenon called ‘informative missingness’, leading to bias in allele and genotype

frequencies that are estimated using only called samples. But, even SNPs that show 100% call rates can suffer from a differential bias (44). Nevertheless, close to complete call rates represents a meaningful genotype quality criterion on a statistical level. Although, I did not observe an inflation of the type I error rate, when I used a case wise call rate cut-off per SNP of 50% in a genome-wide analysis as compared to the 98% cut-off that was applied for the presented analysis of this dissertation, we have seen, that on other than Illumina SNP genotyping platforms, stringent call rate cut-offs are required.

Due to the potential for misclassification errors inherent to genotyping, quality controlling becomes indispensable. Besides reducing systematic differential biases by an optimal experimental design, manual checks of genotype clusters can be done for the most interesting association findings. In a GWAS, however, visual inspections of clusters for all SNPs is certainly too laborious and time-consuming. The introduction of a certain number of individuals genotyped in duplicates, paired with re-genotyping of the most associated SNPs with different genotyping methods based on distinct detection principles are crucial elements of quality control in candidate gene association approaches as well as in GWAS.

A screening procedure allowing genotype quality control for a high number of SNPs is Hardy-Weinberg disequilibrium testing as is dealt with in the following paragraph.

1.1.2.7 Deviation from Hardy-Weinberg equilibrium (DHW)

In sufficiently large randomly mating populations, not subject to genetic and population parameters affecting allele frequencies, the genotypes for an individual marker should distribute according to the principle of Hardy-Weinberg equilibrium (HWE). Supposing the frequencies of the two alleles of a bi-allelic marker are p and q (p+q=1) in the general population, then according to the HWE the genotype frequencies are p^2 and q^2 for the two alternative homozygotes and 2pg for heterozygotes ($p^2+2pq+q^2=1$) (46). Naturally occurring factors for deviation from HWE (DHW) are protective alleles, natural selection, population admixture / stratification, inbreeding, deletions or simply

chance (47-50). Nevertheless, DHW testing is usually performed to detect potential genotyping errors (51). Heavy genotyping errors characterized by erroneous genotype clusters, as have been described in the last chapter, can usually be detected by DHW testing. But, DHW testing has generally low power to detect subtle genotyping errors even in large samples. In contrast, power for the detection of pseudo-SNPs (paralogous and ectopic sequence variants) by DHW testing is > 0.8 in as small sample sizes as 50 individuals (52).

It should be noted, that DHW can also be related to true genetic association, because associations are often rather genotypic than allelic, and genotypic differences may result in DHW (12). Depending on the underlying genetic model, DHW may occur in cases and/or controls with a lack or an excess of heterozygous genotypes. The observed deviations from the expected distribution in heterozygotes should be in opposite directions in cases and controls. DHW in cases and/or controls may be present in recessive, dominant and additive models. DHW in patients in contrast is never expected for a multiplicative model, and is less likely if the susceptibility-allele frequencies are extreme small or large (53). An empirical example for a SNP departing from HWE in the control but not the patient group is a non-synonymous SNP in the *P2RX7* gene, for which associations with bipolar and unipolar depression have been reported in at least three independent studies (54-56).

1.1.2.8 Population stratification

Case-control association studies are highly sensitive to population stratification, since they measure allelic or genotypic differences between cases and controls. As mentioned in the paragraph on LD, allele frequencies might vary drastically between different human populations. Assuming a case-control association study is conducted in a geographic region, where multi-ethnic population admixture is predominant, disproportions in ethnic affiliation between cases and controls would directly result in strong association signals which would reflect differences in the ethnic background, but not at all association to disease. In order to avoid population stratification bias in association studies, ethnic homogeneity or a matching of cases and controls according to their ethnicity should be assured.

In GWAS, the method of genomic control allows to test for population stratification, since a sufficient number of markers are interrogated to decide whether significant differences in allele frequencies between cases and controls occur more often than expected under the null-hypothesis of no stratification. The quotient between the median of the observed allelic χ^2 test statistics of all tested SNPs between cases and controls, and the median of the theoretical distribution is a measure for population stratification, most commonly called λ. If λ is equal to 1.0, there is no population stratification. If it is above 1.0, each association test statistic should be corrected by the factor λ (57-59).

An alternative approach to assess population stratification is the EIGENSTRAT method based on principle component analysis. Individuals are structured according to their genome-wide similarity of marker allele distributions and compared to each other. This methods allows the identification of outlying individuals or to group individuals with regards to their ethnic background to enable a unbiased association analysis (60).

Available data on population stratification in Germans, suggest that there is little to no population stratification within individuals from German ethnicity. A minor degree of population substructure from Northern to Southern Germany has been reported, but it appears to be too low to be detectable without prior knowledge of subpopulation affiliation of a proband (61).

1.1.3 Power calculation in association studies

The statistical power of a study is the probability of successfully detecting an effect of a particular size. If β is the probability for a type II error, then power is defined as $1 - \beta$. In association studies power depends on many factors: the prevalence of the disease, the LD between the genetic marker and the disease causing variant, the effect size (genetic relative risk), the frequencies of marker and disease alleles, the required level of statistical significance α (type I error rate), sample size and the accuracy of genotyping and population stratification. There are open-source software-packages available for power calculations in a case-control association design (fig.2) (62-64).

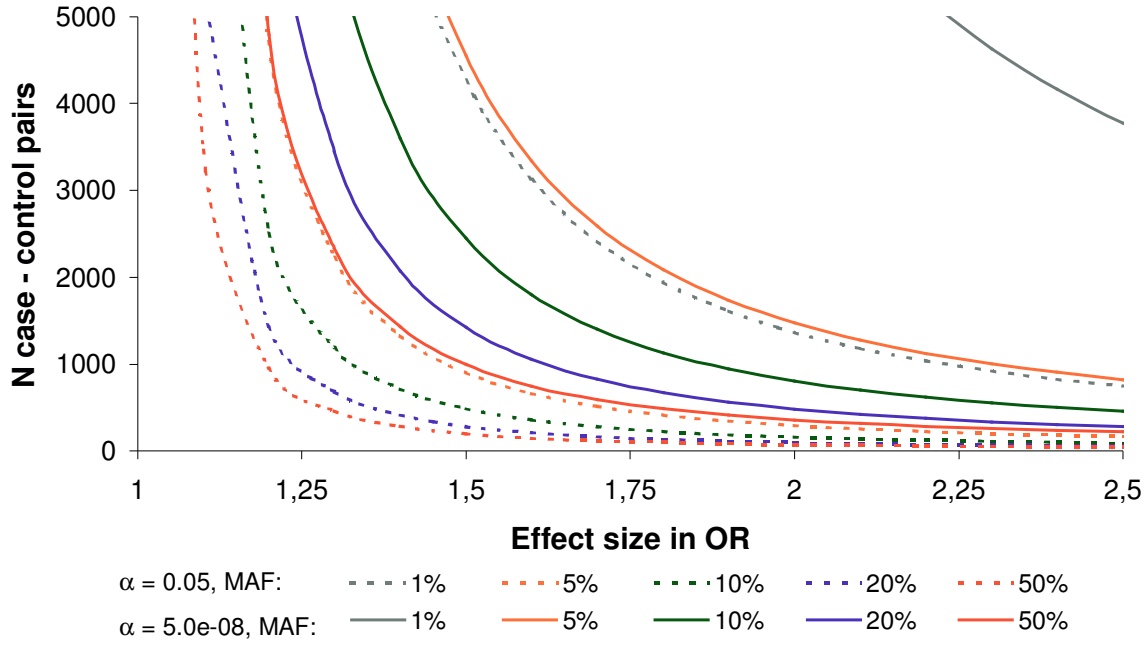


Figure 2: Statistical power of a case-control study in MDD. The effect size of a genetic associtation under a log-linear model is plotted against the number of required matched case-control pairs for a single marker (α = 0.05 , dotted lines) and one million tested markers (α = 5.0e-08, full lines) with low and high minor allele frequncies in the geneneral population (MAF: 1%, 5%, 10%, 20% and 50%) and 80% statistical power (lifetime prevalence of MDD: 16%, used software: Quanto (63)).

1.2 Genetics of Mood Disorders

Major depressive disorder (MDD) is the most common psychiatric disease with an estimated life time prevalence of about 16% (65). The burden of this disorders, in terms of both economic costs and human suffering, is immense (66). In the United States, the annual costs associated with mood disorders have been estimated to exceed $100 billion (67). MDD is associated with substantial disability, medical morbidity, and even mortality (from associated medical illness, suicide, and other sequelae).

In this chapter I summarize the current state of knowledge on the genetic epidemiology and the molecular genetic basis of MDD and suicidal behavior. There is an extensive body of research in this area and therefore only the most robust findings are highlighted (for review: (66, 68-71)). Additionally, this chapter introduces to the more specific background of the here presented studies by giving the reasoning why they were conducted and why the applied experimental approachs were chosen.

1.2.1 Clinical Features and Epidemiology of MDD

MDD is characterized by one or more depressive episodes that are not better accounted for by bereavement, substance use, or a general medical condition (72). A major depressive episode is defined by two weeks or more of persistent depressed mood and/or loss of interest in usual activities that is accompanied by three or more additional symptoms comprising significant changes in appetite/weight, insomnia or hypersomnia, psychomotor agitation or retardation, fatigue or decreased energy, feelings of worthlessness or excessive guilt, impaired concentration or indecisiveness, and recurrent thoughts of death or suicide. Recognized risk factors for MDD include family history of mood disorder, childhood adversity and parental loss, and stressful life events in adolescence and adulthood (73, 74).

The estimated lifetime prevalence of MDD from the National Comorbidity Survey Replication study is 16.2%, and the rate is 1.7-fold higher in women than in men (65). The most common lifetime comorbidities associated with depression are anxiety disorders, substance use disorders, and impulse-control

disorders (65). The burden of disease associated with MDD is enormous. At the beginning of this century, MDD was identified as the leading cause of years lived with disability worldwide and the fourth leading cause of disease burden (75). The direct and indirect costs of MDD exceed $80 billion annually, in the US (76). In addition to the disability associated with MDD, the risk of mortality due to suicide is substantial, with lifetime risk estimates as high as 15-20% (77).

1.2.2 Genetic Epidemiology of MDD

A substantial body of evidence has established that MDD is a familial phenotype (78). In case-control family studies, the risk of depression in relatives of depressed probands has been significantly higher than the risk in relatives of unaffected control subjects, with relative risks ranging from approximately two- to nine-fold (79-83). In a meta-analysis of high-quality family studies, Sullivan et al. found that the prevalence of MDD was threefold higher in the relatives of affected probands, compared to the relatives of unaffected control subjects (84). Twin studies of MDD have provided consistent evidence that genes account for a substantial proportion of the familial aggregation of MDD. In almost all of the more recent twin studies, concordance rates for monozygotic (MZ) twins have exceeded those for dizygotic (DZ) twins (84). For studies published since 1985, the MZ concordance rates have typically been in the range of 30-50%, and DZ concordance rates have typically ranged from 12-40%, with somewhat higher rates seen in female twin-pairs, compared with male twin-pairs (84). Estimates of heritability in these studies ranged from 17-78%. Combining data from these studies, Sullivan et al. estimated the summary heritability at 37% (95% CI: 33-42%), with a larger share of the variance explained by individual-specific environment (63%; 95% CI: 58-67%). The absence of a significant effect of shared family environment suggests that the familial aggregation of MDD is mostly or entirely due to genetic influences. These estimates are consistent with those of the largest and most recent twin study comprising more than 15'000 Swedish twin-pairs, in which the heritability of MDD was estimated at 42% for women and 29% for men (85).

1.2.3 Molecular Genetic Studies in MDD

Genetic linkage studies of MDD have implicated several chromosome regions as harboring susceptibility genes, including 1p (86-89), 2q (90), 12q (86, 91), and 15q (86, 92), although these findings await confirmation. A recent large linkage study of families with recurrent early-onset MDD provided strong support for the 15q region (93, 94). Overall, however, compared to schizophrenia and bipolar disorder, no genetic locus has so far achieved genome-wide significance in linkage studies on MDD (95-102).

Of the many candidate genes examined, the most widely studied has been *SLC6A4* on chromosome 17, which is the gene encoding the serotonin transporter, the therapeutic target of selective serotonin reuptake inhibitor antidepressants. In particular, a common polymorphism in the promoter of the serotonin transporter gene (the serotonin transporter length polymorphic region, 5-HTTLPR) has been the focus of association studies in mood disorders (68). Two common alleles exist and are distinguishable by the insertion (long allele) or deletion (short allele) of a 44-base pair sequence (103, 104). The short allele has been associated with reduced *SLC6A4* function (103, 105). Several meta-analysis of association studies of this polymorphism in MDD have been reported, with mixed results (106-109). More consistent results have been obtained in studies that have incorporated analyzes of gene-by-environment interaction. In a landmark study, Caspi et al. examined the 5-HTTLPR polymorphism in a longitudinal birth cohort of 847 individuals from New-Zealand. These authors found that the short allele was associated with risk for MDD, depressive symptoms, and suicidality among individuals who experienced stressful life events but not in absence of these events (110). At least 11 subsequent studies have examined 5-HTTLPR-environment interactions and depression (111-121), with most supporting an effect on depressive episodes or depressive symptoms, although the two largest studies had negative findings (112, 118). There is also evidence that environmental influences can counteract genetic and environmental risk factors for depression. Kaufman et al. reported that among children with both genetic (5-HTTLPR short alleles) and environmental vulnerabilities (childhood maltreatment), positive

environmental factors (i.e., social support) exerted a protective effect on risk of depressive symptoms (114, 122).

Numerous other genes have been associated with risk of depression in multiple studies, e.g. *BDNF* (123), *ACE* (124), *P2RX7* (54), *TPH2* (125), *PDE9A*, *PDE11A* (126), *DISC1* (127), and *GRIK3* (128). A very recent meta-analysis of studies of 18 genes, found statistically significant evidence of association for six genes: *APOE*, *DRD4*, *GNB3*, *MTHFR*, *SLC6A3* and *SLC6A4* (129), but most others initial results could not be unambiguously confirmed.

In summary, so far neither linkage nor candidate studies revealed susceptibility genes for MDD of unambiguous evidence and none unbiased genome-wide association studies on MDD has been reported, yet. In order to unravel the genetics of MDD, genome-wide case-control, gene-environment interaction and multi-gene model studies might be more promising.

1.2.4 Epidemiology of suicidal behavior

Suicide represents a major public health problem worldwide. There are an estimated 10-20 million attempted suicides each year and one million completed suicides. Since the year 2000, suicide is among the three leading causes of death among those aged 15-44 years (130). The EUROSAVE (European Review of Suicide and Violence Epidemiology) project reported suicide rates in the range of 2.8 to 22.8 per 100'000 inhabitants varying among different European countries for the year 1998 (131). Lifetime prevalence of attempted suicide (SA) in the general population was found to be more stable with average rates between 3-5 % across different sites all over the world (132, 133). Though suicide exerts its fatal effect on numerous lives, there is still a lack of knowledge concerning its underlying cause and pathological mechanism.

In general, patients suffering from psychiatric disorders are at high risk to commit suicide. Especially, mood disorder patients show the highest relative risk for lifetime SA (OR = 7.8-29.9), followed by patients attributed to other diagnostic categories of psychiatric disorders (OR=2.1-6.5) (133). This is also reflected by reported estimates for lifetime prevalence of completed suicide of

8.6% for suicidal inpatients, 4.0% for mood disorder inpatients and 2.2% for mood disorder outpatients, compared to 0.5% for the general population (134).

1.2.5 Genetic bases of suicidal behavior

Suicide and suicidal behavior appear to be part of the same clinical phenotype. Both, suicide and SA are more common in the relatives of probands with either completed or attempted suicide (for review: (135). A higher rate of SA has been found in the relatives of suicide completers (136, 137). Conversely, a high rate of completed suicide has been found in the relatives of patients who attempted suicide (138, 139). Even though patients with mood disorders are at the highest risk to commit suicide, the majority of patients do not even attempt suicide, which suggests that the genetic predisposition for suicidal behavior might be independent of the one for mood disorders. Indeed, twin studies show an increased risk of SA if a co-twin has attempted suicide, even after controlling for psychiatric, demographic, and other salient risk factors for suicide, e.g. childhood physical abuse (140, 141). Furthermore, several family studies of suicide and of suicidal behavior demonstrate transmission of suicidal behavior even after controlling for the transmission of psychiatric disorder and for proband diagnosis (136, 137, 142). The greater concordance for suicidal behavior of mono-zygote versus di-zygote twins (143), together with a high correlation between adoptee suicide and suicide in biological relatives with who the adoptees have had no contact (144), further support a genetic transmission of risk factors for suicidal behavior. In summary, these studies provide evidence for genetic transmission of suicidal behavior being independent from the one of psychiatric diagnosis.

1.2.6 Molecular Genetic Studies of suicidal behavior

The supposed genetic risk for suicidal behavior by epidemiological studies has been specified by human genetic association analysis, especially for genes involved in the monoaminergic system. Recent large meta-analyzes support evidence for association of two functional promoter polymorphisms with suicidal behavior: the A218C/A779C polymorphism in the tryptophan hydroxylase (TPH)

gene and the insertion-deletion (5-HTTLPR) polymorphism in the serotonin transporter gene (107, 145, 146) as well as the functional Val58Met polymorphism in the catechol-O-methyltransferase (COMT) gene (147). The involvment of the serotonergic system as has been suggested by association genetics might partially be supported by post-mortem studies comparing serotonin (5-HT) levels, serotonin transporter (5-HTT) activity and serotonin receptor (5-HT1A/2A) density in different brain regions of suicide victims versus controls. Overall, many studies found reduced 5-HT levels, less 5-HTT activity and an increased serotonin receptor density in certain brain regions of suicidal subjects. An increased receptor density might reflect an adoption to lower 5-HT levels in suicide victims. Of note, there are also many negative studies, which did not find any of these alterations (148).

Beside the monoaminergic system the stress response regulating system, the hypothalamus-pituitary-adrenal (HPA) axis, and the neurotrophic system has been related to suicidal behavior (69, 149-151).

1.2.7 The neurotrophin hypothesis of MDD and suicidal behavior

Neurotrophins regulate neuronal development, survival and function in the peripheral and central nervous system (152). The family of neurotrophic factors is composed of nerve growth factor (NGF), brain-derived neurotrophic factor (*BDNF*), neurotrophin-3 (NTF3) and –4/5 (NTF4/5), and their respective high-affinity receptors – receptor tyrosin kinases A - trkA (NTRK1), B - trkB (*NTRK2*) and C - trkC (NTRK3) as well as the low-affinity nerve growth factor receptor (NGFR, $p75^{NTR}$) (153).

BDNF, for instance, is widely distributed in the brain with high expression levels in the hippocampus and other limbic structures. *BDNF* has been shown to increase synaptic strength, survival, and maintenance of neuronal plasticity in the adult brain through activation of its high-affinity receptor *NTRK2* (154).

Substantial work has raised the possibility that among the many long-term targets of antidepressant treatments may be regulation of neurotrophins (155). An increase of *BDNF* expression or the induction of its high-affinity receptor *NTRK2* in the rodent brain has been shown to occur in response to different

types of treatment with antidepressant efficacy, regardless of whether this implies administration of antidepressant drugs (156, 157) treatment with electroconvulsive seizures (156) or repetitive transcranial stimulation (rTMS) (158). Furthermore, antidepressant treatment might also increase *BDNF* levels in the human brain (159). The relevance of *BDNF*/*NTRK2* induction by antidepressant treatments is supported by animal studies reporting that infusion of *BDNF* either into the midbrain region (160) or bilaterally into the hippocampus (161) as well as overexpression of the full-length *NTRK2* in neurons (162) results in similar behavioural effects as those observed after antidepressant treatment. In addition, infusion of inhibitors of *NTRK2* block the antidepressant-like effects of *BDNF* or antidepressant drug treatment. Using an inducible knockout system, Monteggia et al. could show that the loss of forebrain *BDNF* attenuates the actions of the antidepressant desipramine in the forced-swim-test, further supporting the involvement of neurotrophins in molecular mechanisms of antidepressant treatment (163).

A reduced neurotrophic support in the brain of mood disorder patients is consistent with studies showing anatomical chances, including atrophy and cell death in the limbic system and altered volumes of the hippocampus and prefrontal cortex (PFC) (164-167).

Changes in the neurotrophic system appear to be accentuated in suicide. Plasma *BDNF* levels in depressed patients who had recently attempted suicide were reduced compared to non-suicidal depressed patients and healthy controls (168). And, serum *BDNF* levels were lower in suicide attempters without an established psychiatric diagnosis (169). Strong and most consistent evidence for the involvement of neurotrophic factors especially in suicidal behavior might come from post-mortem brain studies comparing suicidal subject with non-psychiatric controls. Diwivedi et al. first reported reduced *BDNF* and *NTRK2* mRNA and protein levels in the prefrontal cortex (PFC) and the hippocampus of suicide victims relative to controls (170). Karege et al. replicated the findings for *BDNF* and further found NTF3 also to be reduced in the hippocampus, but not in the PFC (171). In both studies the results were independent of psychiatric diagnosis. Karege et al. showed that reduced

neurotrophin protein levels were only found in antidepressant untreated suicidal patients. Furthermore, Dwivedi et al. (2005) reported reduced mRNA and protein levels of NGF, NTF3 and NTF4/5 in the hippocampus of suicide victims compared to controls. In the PFC, however, mRNA and protein levels of NTF4/5 and only protein levels of NGF were reduced. In concordance with Karege et al. (2005) NTF3 levels were not reduced in the PFC (172).

Beside the serotonic system, the stress system (HPA axis) revealed compelling evidence to be implicated in a variety of psychiatric disease including MDD (173-177). With regard to suicidal behavior the HPA axis seems to play a pivotal role too. Nemeroff et al. found reduced corticotropin-releasing hormone (CRH) binding sites in the frontal cortex of suicidal patients compared to controls (178). CRH is a key regulator of the HPA axis in the CNS. However, a later study failed to replicate this initial finding (179). Higher pro-opiomelanocortin (POMC) and glucocorticoid receptor (GR) mRNA levels were detected in the anterior pituitary in suicide victims relative to controls, both key regulators of the HPA axis (180). Furthermore, there is strong evidence for a more severe dysregulation of the HPA axis in suicidal versus non-suicidal psychiatric patients, mainly showed by the dexamethasone suppression test (DST). Seven of nine prospective studies on depressive patients found a higher proportion of DST non-suppressors among suicide completers with a 14-fold risk compared to DST suppressors (71, 181). Another study on the effect of SA on the HPA axis assessed by the combined Dexamethasone/CRH test revealed an intermediate HPA system deflection in depressive patients after a suicide attempt relative to depressed non-suicidal patients and healthy controls (151).

Interrelationships between the neurotrophic-, the serotonergic- and HPA system have also been established in the pathophysiology of depression (182, 183). *BDNF* is a major downstream target of 5-HT signaling. Animal as well as clinical studies revealed elevated *BDNF* and *NTRK2* levels when 5-HT signaling was influenced by chronic administration of selective serotonin reuptake inhibitors (SSRI) (156, 159, 184). Vice versa, injection of *BDNF* in the mouse brain increases 5-HT innervations (185). It has also been shown that stress lowers *BDNF* levels (186), and inversely, *BDNF* administered to stressed animals

resulted in an antidepressant-like behavioral effect (160, 161) and that antidepressants increase neurotrophic signaling (187). Moreover, glucocorticoids (stress hormones) prevents *BDNF*-mediated maturation of synaptic function in developing hippocampal neurons (188). For review of the neurotrophic model for stress-related mood disorders see also Duman and Monteggia (189).

1.2.8 Genetic association studies on neurotrophins in mood disorders

In the current literature, there are several human genetic association studies of the *BDNF* gene and mood disorders; at least ten in bipolar disorder (BP) (123, 190-198) and six in MDD (123, 192, 197, 199-201). These studies focused on a frequent SNP in the most 3' coding exon of the human *BDNF* gene, which results in a valine (val) to methionine (met) amino acid substitution at codon 66 (val66met, dbSNP: rs6265) of the precursor *BDNF* peptide. Although the substitution does not affect *BDNF* protein function per se, it impacts the regulated secretion of the mature peptide by dramatically altering intracellular trafficking and packaging of the precursor *BDNF* peptide (202, 203). In healthy subjects, the met allele has been associated with abnormal hippocampal activation assayed with functional magnetic resonance imaging (fMRI) (202, 204), diminished levels of hippocampal N-acetyl-aspartate (NAA) assayed with MRI spectroscopy, a putative marker of neuronal integrity and synaptic abundance (202, 205) and reduced hippocampal volume (167, 206-208). This functional characteristic of the val66met polymorphism supports the notion that it might confer susceptibility to psychiatric disease. Indeed, the val66met polymorphism has been associated with BP in three independent studies comprising subjects of European origin (190, 196, 198). However, association studies in Belgian, Chinese, and Japanese samples failed to reconcile this finding with the val66met polymorphism and BP (192, 195, 197). A reason for this discrepancy might be the small sample size of these studies. But recent association studies with much larger sample sizes again obtained conflicting

results. Out of four independent studies just the one from Lohoff et al. reached statistical significance in an European sample (123, 191, 193, 194).

With respect to MDD, one study in Mexican-Americans found positive association with the *BDNF* val66met polymorphism (200) and four other studies failed to show significant single-marker case-control association with this or any other *BDNF* polymorphism (192, 197, 199, 201). Nevertheless, Schumacher et al. showed a significant assocation of a three-marker haplotype (rs988748-$(GT)_n$-rs6265) which withstood correction for multiple testing in the original German sample and replication in another independent German sample (123) and Kaufman et. al. reported a significant three-way interaction between the *BDNF* val66met and the serotonin transporter length (5-HTTLPR) polymorphisms and maltreatment history in children affected by MDD (122).

There are less studies published assessing association of neurotrophic polymorphisms on attempted suicide. Kunugi et al., found a non-synonymous SNP (S205L, rs2072446) in the neurotrophin low-affinity receptor gene $p75^{NTR}$ to be associated with SA in a Japanese sample (209), which could, however, not be replicated in an European-American sample (210).

1.3 Study Design of this Dissertation

1.3.1 The hypothesis-driven candidate gene association approach

Since there is strong evidence for *BDNF* and its high-affinity receptor *NTRK2* to be involved in the pathophysiology of MDD as well as suicidal behavior as has been illustrated in previous paragraphs of this dissertation and there are no published case-control data on MDD and SA with a high genetic marker coverage of *BDNF* and *NTRK2* so far, I tagged (27) both genes with a total of 83 SNPs from the 5' to the 3' end of their longest mRNA isoforms annotated by RefSeq (211) ± 20 kb of flanking sequences. The markers were selected to obtain a genomic information content coverage close to 100% according to the inter-marker (SNP) correlation (LD) structure based on the Human HapMap Project Phase II data for the Central European (CEU) population (24).

Genetic association with MDD and SA was tested in a discovery sample of 405 German mood disorder patients and 366 non-psychiatric controls from the Munich Antidepressant Response Signature (MARS) study (www.mpipsykl.mpg.de) (212). In the patient group 113 individuals were positive for at least one lifetime SA and 281 were negative for SA. Significant associations were replicated in an independent German case-control sample of the Max Planck Institute of Psychiatry (MPIP) in Munich consisting of 920 recurrent unipolar depressive patients, of which 152 had positive history of SA and 1024 healthy controls.

1.3.2 The hypothesis-free genome-wide association approach

Several attempts to identify susceptibility genes for MDD by linkage and candidate gene approaches have been undertaken and several genes have been reported to be associated with the disease, however, less or none of these initial reports has been confirmed unambiguously by subsequent studies or in meta-analyzes (68, 129) (see also chapter 1.2.3). Phenotypic diversity and genetic heterogeneity as well as the large environmental contribution inherent to this disorder have been considered to represent major obstacles for the identification of causative variants. In contrast to bipolar disorder, genome-wide association testing has not been reported so far for MDD. To close this gap, I performed a genome-wide association study (GWAS) in unipolar depressive patients (N=353) and screened healthy controls (N=366), followed by an attempt for replication of the most promising result in an independent larger German case-control sample of recurrent unipolar depression (920/1024).

In order to validate the newly identified susceptiblility locus on its biological relevance for an involvement in the physiopathology of MDD, I searched for genotypic influence of associated SNPs on functional variables. I tested genotype-dependency of the associated SNPs on lymphocyte expression profiles based on data from HapMap individuals provided by GENEVAR. Such a gene expression - genotype correlation analysis might result in a functional identification of the underlying susceptibility gene located in or nearby the region of association.

Decreased hippocampal integrity is considered to be related to MDD susceptibility, because depressed patients showed reduced hippocampal volumes assayed by magnetic resonance imaging (MRI) (213). Therefore, I started a collaboration with Philipp Sämann of the MRI group at the MPIP. In vivo high resolution structural imaging (MRI) and ^{1}H-NMR spectroscopy (^{1}H-NMR) was performed as independent functional assays to investigate potential associations of the identified variants with biological risk markers for MDD in a subset of patients and controls of the replication sample.

Since the newly identified susceptibility locus of MDD could be confirmed by the obtained imaging genomics results and was related to gene expressional chances mostly on one specific nearby gene, I went for another independent validation of the so far achieved assignment of the associated region to a novel candidate gene for MDD. Chronic stress represents a risk factor for MDD by increasing exposed individuals risk of about 2-3 OR (214). In collaboration with Mathias Schmidt of the molecular stressphysiology group at the MPIP who established a well-validated mouse model of chronic social stress (215, 216), a possible role of the novel candidate gene in stress vulnerability was assessed by interrogating its hippocampal expression in stress vulnerable compared to stress resistant animals.

2 MATERIALS AND METHODS

2.1 Human Genetic Association Studies

2.1.1 Recruitment and sample characterization

2.1.1.1 Discovery sample

Patients

405 mood disorder patients (178 males, 227 females) were recruited for the Munich Antidepressant Response Signature of the Max Planck Institute of Psychiatry (MPIP) in Munich, Germany (212) (www.mpipsykl.mpg.de). The mean age was 49.0 ± 14.4 years (males: 48.0 ± 13.3, females: 49.8 ± 15.3 years). Patients were included in the study within 1-3 days of admission to the hospital of the MPIP, and diagnosis was ascertained by trained psychiatrists according to the Diagnostic and Statistical Manual of Mental Disorders (DSM) IV criteria. Patients fulfilling the criteria for at least a moderate depressive episode on the 21-item Hamilton Depression Rating Scale (HAM-D ≥ 14) entered the analysis. Patients with depressive disorders due to a medical or neurological condition were excluded. Patients suffered from a first depressive episode (32.8 %), from recurrent depressive disorder (55.6 %) or from bipolar disorder (11.6%). Ethnicity was recorded using a self-report sheet for nationality, first language and ethnicity of the subject and of all 4 grandparents. All included patients were of European descent and 83.2% were of German origin defined as having all four grand-parents of German nationality. Another 10.1% of all patients had at least one grand-parent with German nationality. Life history of suicide attempts (SA) information was available for 394 mood disorder patients (97.3%). There was no difference between patients with and without SA in relation to age, ethnicity, nationality, main psychiatric diagnosis, number of previous depressive episodes and 1st degree family history of unipolar depression. SA was more frequent in female patients ($p=2.4x10^{-3}$). And, attempters had an earlier age at onset of depressive disorder ($p=7.9x10^{-5}$) and

a longer illness duration (defined as age minus age at onset of depression) than non-attempters ($p=2.4x10^{-6}$).

The genome-wide association study (GWAS) in MDD was performed in the 353 unipolar depressive patients (155 males, 198 females). Their mean age was 49.5 ± 14.3 (males: 48.4 ± 13.4, females: 50.4 ± 15.0 years. Unipolar patients suffered from a first depressive episode (36.8 %) or from recurrent depressive disorder (63.2 %). All included patients were of European descent and 88.7% were of German origin.

Controls

366 healthy subjects were recruited at the MPIP. They were selected randomly from a Munich-based community sample and underwent a screening-procedure as follows: Subjects who declared to be of European descent, had never sought psychological or psychiatric help, who did not suffer from severe somatic diseases (e.g. any cancer currently under therapy, Parkinson's disease, multiple sclerosis, Huntington's chorea, amyotrophic lateral sclerosis, cirrhosis of the liver, Cushing's disease, Addison's disease, dyspnea with need of oxygen supply, renal failure necessitating dialysis) according to a telephone-interview were invited to participate in the study. At the time of their visit, persons older than 65 years were tested with the Mini Mental State-Examination test, and subjects with values ≥ 26 were excluded from further procedures (N=10). The remainder of the persons underwent face-to-face computer-assisted interviews based on the Munich version of the Composite International Diagnostic Interview (M-CIDI) (217-219). The M-CIDI is an updated version of the World Health Organization's CIDI version 1.2 (WHO-CIDI; World Health Organization, 1992), which incorporates questions to cover DSM-IV (American Psychiatric Association, 1994) and ICD-10 (World Health Organization, 1990) diagnostic criteria. Exclusion criteria were lifetime history of one of the following psychiatric diseases: mood disorders, schizophrenia, anxiety disorders, alcohol dependence, drug abuse, obsessive/compulsive disorders, post-dramatic stress disorder, dissociative disorders, somatoform disorders, and eating disorders. The overall exclusion rate of all contacted probands was 49.7 %. These subjects thus represent a group of individuals from the general population who

are expected to never have suffered from an axis I psychiatric disorder. Age, gender, ethnicity and nationality were not different from the patient sample.

2.1.1.2 Replication sample

Patients

921 patients (302 males, 619 females) suffering from recurrent major depression were recruited at the MPIP and psychiatric hospitals in Augsburg and Ingolstadt (all in Bavaria, Germany). The mean age was 51.0 ± 13.8 years (males: 49.8 ± 13.6, females: 51.6 ± 13.8 years). Patients were diagnosed by WHO-certified raters according to the Diagnostic and Statistical Manual of Mental Disorders (DSM-IV) using the WHO Schedule for Clinical Assessment in Neuropsychiatry (SCAN, version 2.1). Patients of European descent over 18 years of age with at least two moderately severe depressive episodes were included. Exclusion criteria were the presence of manic episodes, mood incongruent psychotic symptoms, the presence of a lifetime diagnosis of intravenous drug abuse and depressive symptoms only secondary to alcohol or substance abuse or to medical illness or medication. Ethnicity was recorded using a self-report sheet for perceived nationality, first language and ethnicity of the subject and all four grandparents. Information on the lifetime prevalence of suicide attempts was available for 744 patients (80.8%). There was no difference between patients with and without SA in relation to gender, ethnicity, nationality, depressive diagnosis, illness duration and 1st degree family history of MDD. The attempters were younger than non-attempters ($p=7.3x10^{-3}$), had an earlier age at onset of depression ($p=1.1x10^{-4}$) and a higher number of previous depressive episodes ($p=0.03$).

Controls

1029 controls matched to the patient sample for age, gender and ethnicity were randomly selected from a Munich-based community sample and screened for the presence of anxiety and affective disorders using the Composite International Diagnostic-Screener (219). Only individuals negative for the above-named disorders were included and thus represent a group of healthy

individuals with regard to depression and anxiety. The overall exclusion rate of all contacted probands was 18.2 %.

For demographic and illness related data see table 1 and 2.

All studies have been approved by the ethics committee of the Ludwig-Maximilians-University (LMU) in Munich and written informed consent was obtained from all subjects.

Table 1: Demographic data of the discovery and the replication sample. There was no difference in age, gender and ethnicity between cases and controls per sample. For depressive patients illness related variables are shown as well. There were significant differences for many variables between both independent case-control samples.

Depressive patients and healthy controls of the whole study (N=2721)							
Characteristics	**Discovery sample (N=771)**			**Replication sample (N=1950)**			**Betw. samples**
	Controls (N=366)	**Patients (N=405)**	**P[1]**	**Controls (1029)**	**Patients (921)**	**P[1]**	**P[2]**
Women (%)	205 (56.0)	227 (56.0)	1.0	693 (67.3)	619 (67.2)	0.96	$<10^{-4}$
Age (SD)	48.6 (13.4)	49.0 (14.4)	0.67	50.7 (13.9)	51.0 (13.8)	0.68	$<10^{-3}$
Of European descent, %	100	100	1.0	100	100	1.0	1.00
German origin (%)	299 (81.7)	337 (83.2)	0.64	845 (82.1)	779 (84.6)	0.16	0.45
Unipolar depression (MDD, %)	-	353 (87.2)	-	-	921 (100)	-	$<10^{-4}$
Recurrent unipolar depression (%)	-	225 (55.6)	-	-	921 (100)	-	$<10^{-4}$
Bipolar depression (%)	-	47 (11.6)	-	-	0 (0)	-	$<10^{-4}$
Dysthymia (%)	-	5 (1.2)	-	-	0 (0)	-	-
Life history of suicide attempt (%)	-	113 (28.7)	-	-	152 (20.4)	-	$<10^{-2}$
Illness-related variables for unipolar depressive patient (MDD) group:							
Age at onset of depression (SD)	-	37.2 (15.7)	-	-	36.0 (13.9)	-	0.145
Previous episodes (SD)	-	2.8 (5.0)	-	-	5.0 (4.2)	-	$<10^{-4}$
Illness duration[3] in years (SD)	-	11.8 (12.4)	-	-	15.0 (12.1)	-	$<10^{-4}$
1st degree family history of MDD (%)	-	148 (37.1)	-	121 (12.8)	396 (47.8)	$<10^{-4}$	$<10^{-3}$

[1] P-value of a Chi-square test for dichomtome variables and of an ANOVA for quantitative variables with 1 degree of freedom.

[2] P-value of a Chi-square test for dichomtome variables and of an ANOVA for quantitative variables with 3 degrees of freedom.

[3] Illness duration: age minus age at onset of depression. SD: standard deviation. - : no data available.

Table 2: Demographics, depressive diagnosis and illness-related variables are displayed for mood disorder patients with available information about life history of attempted suicide (SA). The patient discovery and replication samples were split according to patients positive and negative for a life history of attempted suicide (SA). In the discovery sample the proportion of women with attempted suicide was significantly higher, but attempters were not younger. Vice versa, younger age, but not female sex was over-represented in attempters in the replication sample. There was no difference in the proportion of patients from German origin in both samples. Suicide attempters had a lower mean age-at-onset of depressive disorder consistently in both samples. Moreover, suicide attempters hat a higher number of previous episodes and longer illness duration of depression becoming significant in the combined discovery and replication sample.

Mood disorder patients with and without life history of suicide attempts (SA) of the whole study (N=1138)							
Characteristics	**Discovery patients with SA information (N=394)**			**Replication patients with SA information (N=744)**			**Betw. samples**
	Without SA (N=281)	**With SA (N=113)**	**P[1]**	**Without SA (N=592)**	**With SA (N=152)**	**P[1]**	**P[2]**
Women (%)	143 (50.9)	77 (68.1)	$<10^{-2}$	692 (66.2)	106 (69.7)	0.44	$<10^{-4}$
Age (SD)	49.1 (14.3)	48.6 (14.5)	0.74	51.5 (13.5)	48.2 (13.7)	$<10^{-2}$	$<10^{-2}$
Of European descent, %	100%	100%	1.0	100%	100%	1.0	1.00
German origin (%)	235 (83.6)	92 (81.4)	0.66	514 (86.8)	125 (82.2)	0.76	$<10^{-2}$
Unipolar depression (%)	247 (87.2)	78 (85.8)	0.74	592 (100)	152 (100)	1.0	$<10^{-4}$
Recurrent unipolar depression (%)	155 (63.2)	66 (68.0)	0.41	592 (100)	152 (100)	1.0	$<10^{-4}$
Bipolar depression (%)	33 (11.7)	13 (11.7)	0.95	0 (0)	0 (0)	1.0	$<10^{-4}$
Age at onset of depression (SD)	39.1 (15.5)	32.3 (14.5)	$<10^{-4}$	37.3 (14.2)	32.4 (12.7)	$<10^{-3}$	$<10^{-4}$
Previous episodes (SD)	2.7 (5.3)	3.3 (4.1)	0.35	4.8 (3.9)	5.8 (5.1)	0.028	$<10^{-4}$
Illness duration[3] in years (SD)	10.0 (10.4)	16.4 (15.4)	$<10^{-4}$	14.1 (11.8)	15.9 (12.5)	0.095	$<10^{-4}$
1st degree family history of MDD (%)	105 (37.9)	41 (36.9)	0.86	251 (46.8)	62 (48.4)	0.74	0.027

[1] P-value of a Chi-square test for dichomtome variables and of an ANOVA for quantitative variables with 1 degree of freedom.
[2] P-value of a Chi-square test for dichomtome variables and of an ANOVA for quantitative variables with 3 degrees of freedom.
[3] Illness duration: age minus age at onset of depression. SD: standard deviation. - : no data available.

2.1.1.3 Epidemiological sample

For research groups within the German National Genome Research Network (NGFN) genotypes from 550k Genotyping BeadChips (Illumina Inc., San Diego, USA) were available of 493 participants from PopGen (Population Genetic Cohort), an on-going cross-sectional epidemiological survey of a population from Northern Germany(61).

2.1.2 DNA preparation

On enrollment in the study, up to 40 ml of EDTA blood were drawn from each patient and DNA was extracted from fresh blood using standard DNA extraction procedures – the Puregene whole blood DNA-extraction kit (Gentra Systems Inc; MN) (220).

2.1.3 SNP selection and genotyping

For the hypothesis-driven candidate gene appraoch in MDD and SA a high SNP marker coverage was applied for the full-length mRNA isoform of *NTRK2* (RefSeq: NM_006180, 355.04 kb) on chromosome 9q22 and for the longest *BDNF* isoform (NM_170731, 66.86 kb) on 11p14.1 comprising 20 kb of flanking sequences on both, the 5'- and 3'-ends of the genes according to Human HapMap Project Phase II data for the Central European (CEU) population (24). A tagging SNP approach was performed using the tagger software (27) integrated in the HapMap project home page (http://www.hapmap.org). Pairwise r^2 was set to ≥ 0.8 with a MAF cut-off of ≥ 0.1 for *NTRK2* and ≥ 0.01 for *BDNF*, respectively. The suggested number of tagging SNPs for *NTRK2* was 69 and 18 for *BDNF* (totally 87 tagging SNPs). Genotype data of a genome-wide case-control experiment in the discovery sample were available. The genome-wide experiment was performed on the Sentrix Human-1 100k and 300k Genotyping BeadChips (Illumina Inc., San Diego, USA) according to the manufacturer's standard protocols. The genotypes of all 100 SNPs located in the above mentioned loci were extracted from the genome-wide dataset. Prior to any association testing, tagging SNPs were drawn of this dataset and assigned to the r^2-bin structure of HapMap data. Our genome-wide dataset covered 46 HapMap bins of both, the *BDNF* and *NTRK2* loci. The missing 41 tagging SNPs were genotyped on a MALDI-TOF mass-spectrometer (MassArray® system) employing the AssayDesigner software (Sequenom Inc., San Diego, USA) for primer selection, multiplexing and assay design, and the homogeneous mass-extension (hMe) process for producing primer extension products. MALDI-TOF SNP genotyping was performed at the facilities of the Helmholtz Zentrum Munich, Germany. The 12 SNPs, which were assessed in the replication

sample, were also genotyped on the MALDI-TOF platform. All primer sequences used are available upon request.

For the GWAS in MDD genome-wide SNP genotyping for the discovery sample was performed on Sentrix Human-1 100k and 300k Genotyping BeadChips (Illumina Inc., San Diego, USA) according to the manufacturer's standard protocols. On the Illumina Human-1 100k Genotyping BeadChip about 109,000 exon-centric SNPs can be interrogated. Nearly 25,000 of the loci are located in transcripts and more than 73,000 loci are within 10 kb of coding sequences. The Illumina Human-1 300k Genotyping BeadChip comprises about 317,000 SNPs, which have been selected from the entire genome, mainly based on HapMap tagging SNPs and to a minor part with an exon-centric focus. Genome-wide genotyping was performed at the Center for Applied GenoTyping (CAGT), a core facility of the MPIP. The average call rate achieved was better than 99%, with samples below 98% being either retyped or excluded from the study. The reproducibility for samples (N=3) genotyped twice was 99.999% or better.

To exclude the possibility of genotyping errors at least for the most relevant associations, two SNPs (rs1545843 and rs1031681) were exemplary re-genotyped in 339 patients and 365 controls of the discovery sample with the independent hybridization probe method on a Light Cycler 480 PCR Analysis System (F. Hoffmann-La Roche Ltd, Basel, Switzerland) according to the manufacturer's standard protocols. The genotype concordance rate with the Illumina chip genotypes was 99.93% (1 discordant out of 1408).

The epidemiological sample has been genotyped on 550k Genotyping BeadChips at the Illumina facilities (Illumina Inc., San Diego, USA).

SNPs of the GWAS selected for replication were genotyped on a MALDI-TOF mass-spectrometer (MassArray® system, Sequenom Inc., San Diego, USA) at the facilities of the Helmholtz Zentrum Munich, Germany. The eight selected SNPs were part of three highly multiplexed iPlexes together with SNPs of completely different projects. They were genotyped on six 384-well plates. The mean call rate over all plates for the 7 SNPs fulfilling quality control criteria was 97.8 ± 2.4 % and these SNPs were in HWE ($p<0.05$) on each single plate and over all plates. All primer sequences used are available upon request. To

assess the reliabilty of MALDI-TOF genotyping rs1545843 was exemplarily genotyped twice in a total of 1581 cases and controls of the replication sample and a genotype concordance rate of 99.94 % (1 discordant out of 1581) was obtained.

2.1.4 Statistics

2.1.4.1 Power calculation

For power calculations the Genetic Power Calculator (62) and Quanto (63) were used. (http://pngu.mgh.harvard.edu/~purcell/gpc and http://hydra.usc.edu/gxe). Given a prevalence of unipolar depression of 16% (65), a marker in LD with the causative variation (D'=1) under a dominant or recessive genetic model and 80% statistical power by testing the tagging SNP marker panel for *BDNF* and *NTRK2* at a significance level $\alpha = 6.02 \times 10^{-4}$ (= 0.05 / 83 SNPs), the candidate gene approach was able to detect an effect size of ≥ 1.8 OR in the discovery sample. Under an log-additive model the study was able to detect effect sizes of ≥ 1.6 OR.

The prevalence of life history of SA among mood disorder patients was estimated to be 23% from the combined study data, which is consistent with reported prevalence (133). Under the above mentioned settings for a dominant or recessive model, the study achieved 80% statistical power to detect effect sizes of ≥ 2.2 OR on SA among mood disorder patients in the discovery sample.

The GWAS on MDD had 80% statistical power to detect an effect size of ≥ 2.6 and ≥ 2.0 OR under a dominant or recessive and a log-additive model, respectively, at a significance level $\alpha = 1.4 \times 10^{-7}$ (= 0.05 / 365,676 SNPs) under the same settings as applied above.

2.1.4.2 Genomic controls

Genomic controls (59) for the case-control MDD phenotype were calculated with R-2.5.0 (http://cran.r-project.org) on a genome-wide level in the discovery sample by including genotypes of all 365'676 quality controlled SNPs. In

addition, population stratification was tested with EIGENSTRAT implemented in EIGENSOFT which is based on principle component analysis (60). (http://genepath.med.harvard.edu/~reich/EIGENSTRAT.htm).

2.1.4.3 Association testing

Case-control association and HWE testing as well as MAF and OR calculation were conducted by applying the WG-Permer software (http://www.mpipsykl.mpg.de/wg-permer/). For post-hoc analysis R-2.5.0 (http://cran.r-project.org) and SPSS for Windows (Releases 12, SPSS Inc., Chicago, Illinois 60606, USA) were used.

SNPs deviating from HWE at a significance level of $\alpha = 5.7 \times 10^{-4}$ (= 0.05 / 87 SNPs) in combined cases and controls, with MAFs below 1% and case-wise callrates per SNP below 98% were exluded from association testing in the discovery and replication sample of the candidate gene approach. In the GWAS in MDD, SNPs deviating from HWE at a significance level of 10^{-5} with a MAF below 5% and a case-wise call rate per SNP below 98% or 95% in the discovery and replication sample, respectively, were excluded from association testing.

SNPs were tested for association with major depression (MDD: unipolar depressive episode and unipolar recurrent depression) in a case-control design using Chi square test statistics under allelic and both alternative recessive-dominant modes of inheritance. Genetic associations with life history of SA among mood disorder patients were tested by ANOVA with residual variables which were generated by logistic regression analysis correcting for age and gender. All association results were verified with logistic regression analysis including gender and age as covariates. P-values from logistic regression analysis with covariates did not differ from those of the Chi square or ANOVA test statistics for all reported associations of this work. All nominal p-values given in this dissertation were further validated by calculating empirical p-values under an appropriate numer of permutations ($N_{permutation} \geq$ [1 / nominal p-value]*10). Each here reported nominal p-value did not deviate from its empirical p-value.

The level of significance was set to 0.05 and p-values were corrected for multiple comparisons by the permutation-based minimum p method proposed by Westfall and Young (WY) (34, 35) under 10^5 permutations over all performed tests (number of tested genetic models and SNPs) per phenotype. In addition, the permutation-based Fisher product method (FPM) (221) was applied to test whether there is genetic association over all tested markers of a candidate gene with the phenotype of interest. FPM was exerted under 10^5 permutations over all tested SNPs and genetic models for both phenotypes (MDD, SA) and genes (*BDNF*, *NTRK2*). The resulting 4 p-values of this method were Bonferroni-corrected to adjusted for the number of performed FPM tests ($\alpha = 0.05 / 4 = 0.0125$). Both procedures (WY, FPM) were also applied in the replication sample in an analogous way, but with the SNP-wise best genetic model and with the same direction of the effect as was observed in the discovery sample, thus allowing one-sided testing.

In the GWAS on MDD the WY method over all SNPs and all three tested genetic models was used to correct for multiple testing in the discovery as well as in the replication sample. Since the direction of the association as observed in the discovery sample was preset for each tested genetic model in the replication sample, one-sided p-values were calculated for the replication analysis.

Haplotype association testing was performed with Haploview 4.0 (222) for the SNPs on 12q21.31 associated with MDD (http://www.broad.mit.edu/mpg/haploview).

Multi-locus interaction testing on SA among mood disorder patients in the combined sample was performed using step-wise logistic regression analysis conducted in R-2.5.0 (http://cran.r-project.org). In order to maintain statistical power just the three most significant SNPs of the joint sample analysis under their best fitting genetic models, their pair-wise interaction terms, age and gender were included to the model.

Sample demographic statistics and post-hoc tests on age, gender and German origin, life events, recurrence of MDD, age at onset of depression, number of previous depressive episodes, first degree family history of MDD and life time

attempted suicide status were performed by logistic regression analysis and ANCOVA. Post-hoc tests with additional covariates in genetic association testing on SA among mood disorder patients (age at onset of depression and the number of previous episodes of depression) or comparisons of mood disorder suicide attempters with healthy controls were performed by logistic regression analysis and ANCOVA.

2.1.4.4 Linkage disequlibrium

Linkage disequilibrium (LD) pattern and haplotype block delineation were determined by performing Haploview 4.0 (222) (http://www.broad.mit.edu/mpg/haploview). Blocks were defined using the confidence interval method described by Gabriel et al. (21). Pairwise LD measures (r^2 and D') were calculated in the 366 healthy controls of the discovery sample and compared to the CEPH population (Utah residents with ancestry from northern and western Europe, CEU, N=60) (24) (http://www.hapmap.org).

2.2 SNP-genotype functional correlation analyzes

2.2.1 Genotype-dependent human mRNA levels

Publicly available genome-wide expression data from GENEVAR - GENe Expression VARiation - a large resource-generating scientific project of the Wellcome Trust Sanger Institute (http://www.sanger.ac.uk/humgen/genevar/) were explored (223, 224). These gene expression data were generated from Epstein-Barr Virus (EBV) transformed lymphoblastoid cell lines from all 270 HapMap individuals by applying Illumina genome-wide expression arrays (wg GEX human 6). Each individual had at least 4 replicate hybridizations. The available data were quantile normalized across replicates and medial normalized across individuals. Expression and genotype data of the 210 unrelated HapMap individuals of the following populations were extracted: Yoruba in Ibadan, Nigeria (YRI, N=60), Japanese in Tokyo, Japan (JPT, N=45), Han Chinese in Beijing, China (CHB, N=45) and the CEU (N=60)

(http://www.hapmap.org) (24, 225). All RefSeq annotated genes (211) located within 1.5 megabase on both sites of the experiment-wide significant SNP of the GWAS (rs1545843, total sequence of 3 Mb) were considered. The five following genes intersect with the defined genomic region (hybridisation probes in brackets): *TMTC2* (GI_22749210-S), *SLC6A15* (GI_33354280-A, GI_21361692 -I, GI_33354280-I), *TSPAN19* (GI_37541880-S), *LRRIQ1* (hmm2373-S) and *ALX1* (GI_5901917-S]) (226). A residual expression variable for each probe was built by regression analysis to correct for ethnicity. Out of the eight associated SNPs being in LD with each other, two tagging SNPs were determined by applying the pairwise tagging algorithm integrated in Haploview 4.0 (r^2>0.75) in the combinded discovery and replication sample. Proposed tagging SNPs were the experiment-wide significant SNP of the GWAS (rs1545843) and rs1031681. The allelic and both alternative recessive-dominant genetic models per SNP for each of the probes (N=7) were tested by performing ANOVA under 10^5 permutations. P-values were corrected for multiple comparisons by the WY method (34, 35) under 10^5 permutations over the performed tests. Subsequently, this analysis was repeated by including data of all available probes for expressed sequence tags (EST) from GenBank (NCBI) in the same genomic window (N=+6: Hs.365699-S, Hs.506230-S, GI_41149683-S, Hs.208111-S, GI_41149726-S, hmm21473-S, tab.3). Data of four ESTs were excluded from the analysis, because their probes did not map completely or uniquley to any target EST sequence of the current GenBank database (GI_37541937-S, hmm21470-S, GI_37541941-S, hmm21472-S). Target sequences of all probes included in this analyses are all devoid of known common variations denominated by dbSNP build 129 (NCBI).

Table 3: Transcripts, Human expression probes, RefSeq gene and EST annotation. All probes of the GENEVAR dataset (Illumina GEX human 6 gene expression chip) located within 1.5 megabase on both sites of the experiment-wide significant SNP (rs1545843) of the GWAS in MDD were analysed (total genomic sequence of 3 Mb on 12q21.31, FL: full-length transcript, S: shorter, truncated transcript).

Transcript	Probe	RefSeq Gene	UniGene	GenBank
SLC6A15 FL	GI_33354280-I	NM_182767	Hs.44424	GI:60115819
SLC6A15 S	GI_21361692-I	NM_018057	Hs.44424	GI:60223401
TMTC2	GI_22749210-S	NM_152588	Hs.577775	GI:22749210
ALX1	GI_5901917-S	NM_006982	Hs.41683	GI:154813200
TSPAN19	GI_37541880-S	NM_001100917	Hs.156962	GI:155369268
Hs.557975	hmm21473-S	no; unspliced single hit EST	Hs.557975	GI:24993973
Hs.677048	Hs.365699-S	no; unspliced single hit EST	Hs.677048	GI:10432863
Hs.339071	GI_41149726-S	no; unspliced 10-hit EST	Hs.339071	GI:13746404
GI:875314	GI_41149683-S	no; unspliced single hit EST	-	GI:875314
Hs.208111	Hs.208111-S	no; spliced 5-hit EST	Hs.208111	GI:13744001
SLC6A15, FL/S	GI_33354280-A	NM_182767, NM_018057	Hs.44424	GI:60115819,GI:60223401
LRRIQ1 FL/S	hmm2373-S	NM_001079910, NM_032165	Hs.402200	GI:24038146,GI:14017818
Hs.675906	Hs.506230-S	no; unspliced 4-hit EST	Hs.675906	GI:34528241

2.2.2 Imaging genomics

2.2.2.1 Nuclear magnetic resonance spectroscopy (^{1}H-NMR)

^{1}H-NMR-spectra of the left hippocampus (fig. 13A) were available from a total of 204 subjects of the replication sample, comprising 109 patients and 96 control subjects. The subjects' affective condition at the time of scanning was documented using the Beck Depression Inventory (BDI) (227). For 18 subjects, no genotypes were available. Three control cases with outlier cerebrospinal fluid (CSF) fraction values within the spectral voxel were radiologically reviewed and excluded due to obvious signs of neurodegeneration.

After strict spectral quality control (see below), groups entering statistical analysis of genotype associations with ^{1}H-NMR metabolites comprised 161 subjects (80 patients [mean age 50.2 (SD 13.8), 62.5% female]), 81 control subjects [mean age 42.5 (SD 13.9), 50.6% female]), for creatine (Cr), myo-inositol (ml), combined N-acetylaspartate and N-acetylaspartylglutamate (NAA),

choline and phosphocholine containing compounds (Cho), and unresolved glutamate/glutamine (Glx) (tab.4).

Table 4: Demographic and clinical information on the MRI/^{1}H-NMR sample.

^{1}H-NMR / Regional volumetry sample chracteristics	Patients	Controls
Number of subjects	80	81
Mean age in years (standard deviation)	50.2 (13.8)	42.5 (13.9)
Females, %	62.5	50.6
Patients with Beck Depression Index > 14, %	40.0	NA
Patients with antidepressant therapy at MRI / ^{1}H-NMR, %	73.8	NA
Rs1031681 genotypes (AA / AG / GG)	18 / 43 / 19	13 / 36 / 32
Rs1545843 genotypes (AA / AG / GG)	22 / 39 / 19	17 / 40 / 24

^{1}H-NMR protocol

Single voxel spectra sized 30 × 17 × 12 mm^3 (6.12 mL) were acquired at 1.5 Tesla using a standard quadrature head coil and a PRESS acquisition (probe-p, General Electrics Medical Systems) (TR = 2000 ms, TE = 35 ms). Voxels were positioned on a high resolution T1-weighted image (spoiled gradient echo sequence, TR 10.3 s, TE 3.4 ms, 124 sagittal slices, voxel size 0.8975 × 0.8975 × [1.2-1.4] mm^3, flip angle 90°) with maximum inclusion of the head and tail of the left hippocampus and minimal inclusion of CSF. The number of averages was standardized to 128 resulting in a total acquisition time of 6 minutes. Shimming was performed by automated first order gradient shim. Metabolite concentrations were calculated using a time-domain fitting algorithm (228). Metabolites Cr, ml, NAA, Cho, Glu and Glx were analyzed and expressed normalized to internal water in institutional units (IU). The tissue composition within the spectrum voxel was determined using template supported segmentation of the high resolution T1-weighted anatomical image into grey matter (GM), white matter (WM) and CSF in native space (Statistical Parametric Mapping software package, version SPM2, http://www.fil.ion.ucl.ac.uk/

pmsoftware/spm2). Volumes of all compartments were extracted from the spectral voxel using in-house IDL software and transformed to fractional values between 0 and 1. Spectral quality and accuracy were analysed qualitatively by visual inspection of residuals after fitting, and quanitatively by using the linewidth (full width at half maximum, fwhm) of the entire spectrum and metabolite-specific fit accuracy values. Long-term stability of the scanner system was assured by analysis of phantom spectra performed two times per week during the study interval (fig.3a). Spectra were required to have a fwhm ≤ 0.115 ppm and fit accuracies ≤ 20% (Cramer Lao lower bound) for NAA, Cho, mI, Cr and Glx, as recommended (228). For glutamate and glutamine which are difficult to separate at a field strength of 1.5 Tesla sample sizes decreased to 125 and 27 subjects, respectively. Thus, Glx effects were explored with regard to the contribution of glutamate at a lenient fit accuracy threshold of ≤ 25% for glutamate that retained a sample size of 147 subjects. No difference of the spectral linewidth between patients and controls, or between genotypes within the control group was detected (T-tests for independent samples, two-sided, all p-values > 0.05).

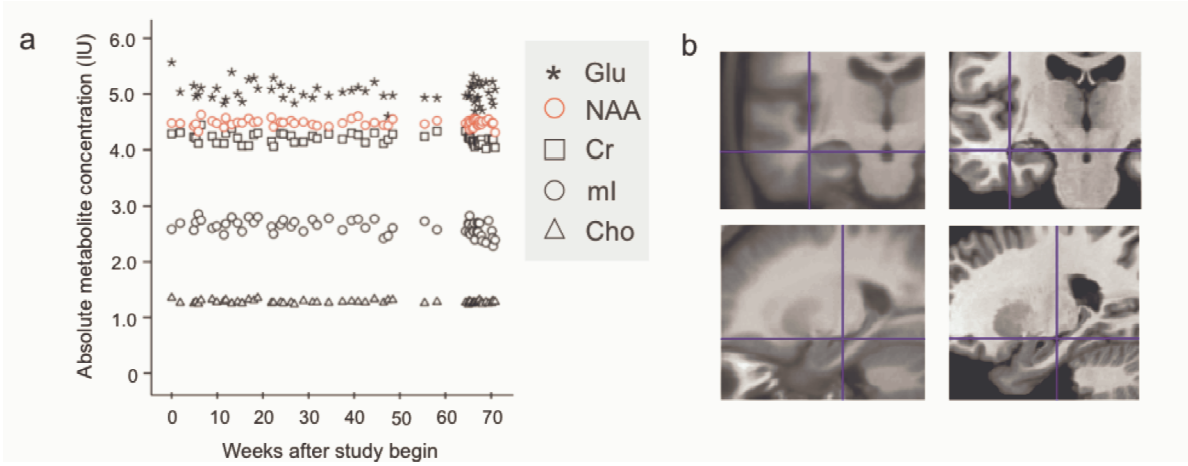


Figure 3: Stability of ^{1}H-NMR spectroscopy across study interval and exemplary depiction of coregistration accuracy for regional volumetry. (a) Phantom measurements of standard metabolites over the study period are shown, based on post-processing identical to the clinical sample, with no indication of systematic drifts. Particularly, no changes were seen for NAA (highlighted in red, r=-0.18, p=0.218). Glu: glutamate; NAA: acetylaspartate; Cr: creatine; mI: myo-inositol; Cho: choline; IU: institutional units. **(b)** Left panel shows average of 30 randomly selected T1 whole head images of the study sample (15 patients, 15 controls) after application of linear/non-linear transformations previously estimated by the combined normalisation/segmentation algorithm in SPM5, and resliced using trilinear interpolation. Right panel shows the two identical cursor positions as projected on a representative single subject template brain ('Colin27 average') in Montreal Neurological Institute (MNI) space at 1×1×1 mm^3 resolution.

Statistical analysis

As proof-of-concept analysis, one-factorial multivariate analysis of covariance (MANCOVA, Wilks' lambda statistics) was performed between patients and controls, and between patients currently non-depressed (Beck Depression Index [BDI] ≤ 14, N=42) and depressed (BDI > 14, N=32) (229) to investigate group effects on NAA, Cho, mI, Cr and Glx, covarying for age and gender effects (230) and the fraction of CSF within the spectrum voxel. Equally, one-factorial MANCOVA was employed in the control and patient group to investigate genotypic effects of the tagging SNPs rs1545843 and rs1031681 on metabolites with covariates as described. Genetic models were the allelic model (A vs. G) and the model comparing carriers vs. non-carriers of the risk allele (AA+AG vs. GG).

2.2.2.2 Regional volumetry

Image preprocessing

T1-weighted high resolution images were subjected to inhomogeneity correction, spatial normalisation and segmentation into grey matter (GM), white matter (WM) and CSF using the unified segmentation algorithm (231) as implemented in SPM5 (http://www.fil.ion.ucl.ac.uk/ spm/software/spm5) and prior probability maps of the standard SPM5 distribution in MNI152 space (resolution $2 \times 2 \times 2$ mm^3) as based on the International Consortium for Brain Mapping maps (ICBM, http:// www.loni.ucla.edu/ICBM). Further settings were: ICBM space European brains template for affine regularization; [2, 2, 2, 4] Gaussian distributions to represent the intensity distribution for GM, WM, CSF and other; warping regularization strength: 1; DCT bases cutoff: 25 mm; bias regularization factor: 0.0001; Gaussian smoothness of bias for non-uniformity correction (fwhm): 60 mm; sampling distance: 3 mm. A hidden Markov random field was applied to account for dependency of neighboring voxels (http://www.fil.ion.ucl.ac.uk/spm/ext/#VBMtools), and Jacobian modulation was performed to preserve local volumes. Normalized, modulated maps of GM, WM and CSF were written out at a resolution of $1 \times 1 \times 1$ mm^3. Global volumes of GM, WM and CSF were calculated as total of unsmoothed, modulated tissue

volume maps after masking out non-brain and non-CSF voxels. A respective mask which separated external CSF spaces from the soft tissues was generated from a database of 498 subjects pre-processed in the identical manner. Figure 3b exemplifies hippocampal coregistration accuracy between MNI space and study images.

Automated regional volumetry

Based on histologically validated, published cytoarchitectonic probability maps of hippocampal subregions (232) the following binary maximum probability maps were derived (resolution 1 x 1 x 1 mm^3) using the Anatomy Toolbox (233): (1–2) left and right hippocampus (HIP; including cornu ammonis [CA1–3], ref. to as CA; fascia dentata together with CA4, ref. to as dentate gyrus [DG], and subiculum [SUB]), (3–4) left and right CA, (5–6) left and right SUB, and (7–8) left and right DG (fig.14D). As control region (9) the combined left and right precentral gyrus from the AAL atlas system was defined (http://www.fil.ion.ucl.ac.uk/spm/ext/#WFU_PickAtlas). The sum of all GM voxels and all WM voxel values within the regional masks was calculated using in-house software programmed in IDL (http://www.creaso.com) and is referred to as regional GM and or WM volume, respectively. The sum of regional GM and WM volumes is referred to as regional brain volume.

Statistical analysis

To test for an association between genotypes of the two tagging SNPs (rs1031681 and rs1545843) and left or right total HIP volumes, analyses of covariance across patients and controls with independent factors *genotype* (model selection as described for the ^{1}H-NMR analysis) and *group (case-control)*, covarying for age, gender and total brain volume were performed. MANCOVA of three subregions were applied for the left and right side to define the subregion and compartment driving the effects seen in the total hippocampus. Analyses were repeated for GM and WM with conservative Bonferroni adjusted significance level for univariate comparisons to 0.05 / 72 = 0.00069 (3 compartments [brain, GM, WM], 6 subregions [3 left and 3 right subregions], 2 SNPs, 2 genetic models).

2.3 A mouse model on chronic stress

2.3.1.1 Animal housing

120 male CD1 outbred mice were used for all experiments. Animals were 28 days old at the day of arrival and were kept on day cycle consisting of a 12-hours-light and a 12-hours-dark period. Food and water was provided *ad libitum*. The experiments were carried out in accordance with European Communities Council Directive 86/609/EEC. All efforts were made to minimize animal suffering during the experiments. The protocols were approved by the committee for the Care and Use of Laboratory Animals of the Government of Upper Bavaria, Germany.

2.3.1.2 Chronic stress paradigm

The chronic social stress procedure was performed as described previously (215, 216). In this paradigm mice are exposed to a highly unstable social and hierarchical situation during the adolescence and young adult period. Briefly, after a habituation period of five days following arrival, the group composition in each cage was changed twice per week for seven weeks, so that each time four mice from different cages were put together in a new, clean cage. The rotation schedule was randomized to minimize the likelihood of a repeated encounter of the same mice throughout the experiment. Control mice remained with the same cage mates. After the seven weeks stress procedure all animals were single housed for 5 weeks. Blood corticosterone levels were measured prior to the stress period (T_0) and after one (T_1) and five weeks (T_2) of stress-free single housing (fig. 14b). Adrenal weight was determined after five weeks of rest (T_2). An extension of this chronic stress animal model is the subdivision of the formally stressed animals into stress resistant and stress susceptible animals after one week of rest (T_1). Stress resistant animals shows strong stress-related disturbances immediately after the stress period of 7 weeks, but recovered to the level of unstressed control animals already after one week under stress-free conditions, whereas stress susceptible animals show persistent elevated corticosterone levels even after five weeks of rest (fig.15). For expression

profiling the 6 most stress resistant and the 6 most stress susceptible mice were chosen (fig. 16).

2.3.1.3 Behavioral analysis

Anxiety-related behavior was assessed using the novelty-induced suppression of feeding test as described previously (234) with some adaptations. Three days before testing the animals received a piece of almond in their home cage for two consecutive days. The consumption time of the almond in the home cage environment was less than 30 seconds for all mice on the second day, with no differences between the groups. Testing was then performed in an open field arena as described below (light intensity 50 lux), with a small piece of almond put in the center. The animals were familiar with the almond, but were unfamiliar with the open field arena. The latency until the initiation of food intake was recorded. The total test time was 30 minutes. Shorter consumption times are interpreted as lower anxiety levels.

2.3.1.4 Tissue dissection and expression profiling

Frozen brains were sectioned at the level of the dorsal hippocampus and the subregions CA1 and dentate gyrus were laser-microdissected (fig. 4). Extracted RNA was quality checked, subjected to two rounds of linear amplification and hybridized to genome-wide Illumina Sentrix BeadChip microarrays. The data of the whole experiment have been deposited in NCBI's Gene Expression Omnibus (GEO, http://www.ncbi.nlm.nih.gov/geo/) and are accessible through GEO Series accession number GSE112211.

Tissue dissection was performed according to Datson et al. with some minor modifications (235). Brain slices of the hippocampal region were cut at 20 µm and thaw-mounted on membrane-coated slides (P.A.L.M. Microlaser Technologies, Bernried, Germany). Laser dissection of CA1 and dentate gyrus material was performed using a laser capture microscope (P.A.L.M. Microlaser Technologies, Bernried, Germany). Extracted samples were immediately dissolved in 100µl TRIZOL (Invitrogen Life Technologies, Carlsbad, USA) and RNA was extracted according the the manufacturer's instructions.

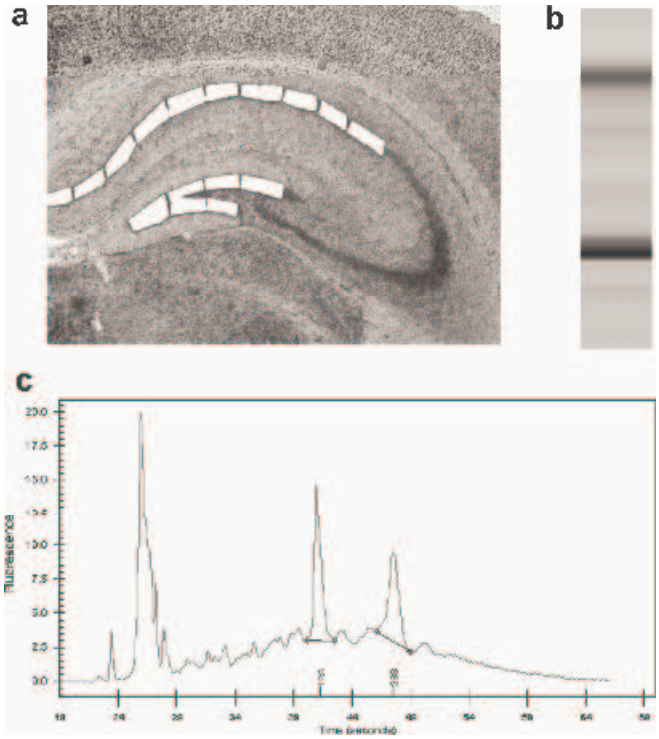


Figure 4: Mouse mRNA sampling by laser-assisted microdissection.
(a) Sample picture of laser-captured CA1 and dentate gyrus slice counter-stained with cressyl violet. RNA quality was determined on an Agilent 2100 Bio-analyzer using the RNA 6000 Pico Lab-Chip kit (Agilent Technologies, USA).
(b) and **(c)** Example gel and electro-pherogram pictures for RNA quality assessment.

RNA quality and quantity was checked by analyzing 1µl of RNA on the Agilent 2100 Bioanalyzer using the RNA 6000 Pico LabChip Kit (Agilent Technologies, Palo Alto, CA, USA) according to the manufacturer's instructions. High quality RNA samples were then amplified twice using the Amino Allyl MessageAmp RNA Amplification Kit (Ambion Inc. Austin, Texas, USA).

Pooled amplified RNA samples were then hybridized on Illumina mouse BeadChips (N=4 per group) and detected in the Illumina BeadArray Reader (Illumina, Inc., San Diego, CA). Gene expression was analyzed using the Illumina BeadStudio software (version 1.5.1.3). A differential score, which takes into account background noise and sample variability (236), of 50 and a fold regulation of 1.5 were set as cut-off criteria.

2.3.1.5 Gene expression analysis in stress vulnerable versus stress resistant mice

The same procedure to select genes adjacent to the region of association for their validation in the described mouse paradigm was applied as chosen for the human expression analysis. Expression differencies were checked for *TMTC2* (NM_025775.1; scl066807.1_5-S), *SLC6A15* (NM_175328.1; scl0003791.1), *TSPAN19* (no mouse homolog), *LRRIQ1* (XM_137221.4; scl074978.1_15-S) and *ALX1* (NM_009423.2; scl022032.1).

Differentially expressed genes were validated by *in-situ* hybridization as described previously (215). The antisense cRNA hybridisation probe of *SLC6A15* was 487 base pairs long (left primer: TGCCGTGAGCTTTGTTTATG; right primer: CAGTGTTGGGGAACCACTTT covering exons 11 to 13 of the gene. The slides were apposed to Kodak Biomax MR films (Eastman Kodak Co., Rochester, NY) and developed. Autoradiographs were digitized, and relative expression was determined by computer-assisted optical densitometry (Scion Image, Scion Corporation).

Statistical analysis

The commercially available program SPSS version 12 was used for statistical analysis. Group comparisons were performed using the two-tailed paired t-test to determine statistical significance (*P<0.05; **P<0.01; ***P<0.001). Data are presented as mean ± s.e.m.

3 RESULTS

3.1 Results of the candidate gene association study on MDD and a life history of attempted suicide

Genetic association analyses was performed in 394 patients with depressive disorder of which 113 had a history of suicide attempts (SA) and healthy matched controls (N=366) with SNPs located in two neurotrophic candidate genes for MDD and suicidal behavior encoding brain-derived neurotrophic factor (*BDNF*) and its high-affinity receptor, receptor tyrosine kinase 2 (*NTRK2*). The reasoning behind this gene selection has been outlined in chapter 1.2.7.

3.1.1 Quality control and SNP marker coverage

The *NTRK2* and *BDNF* loci were tagged with a total of 87 tagging SNPs for most complete genetic coverage in European populations. Data of four single r^2 bin SNPs of *NTRK2* were excluded from association testing, because of call rates smaller than 90% in cases and controls (rs11140831, rs10868231, rs17087959 and rs4504715). Two other *NTRK2* tagging SNPs (rs4406490, rs614886) showed overall call rates of 93.5% and 97.0%, respectively, which is below the claimed quality cut-off of 98% of this study. Nevertheless, both SNPs were included into the association analysis, because both were in HWE and visual examination of their genotype clusters did not suppose any type of genotyping misclassification error. Moreover, the obtained genotype distributions of these SNPs were in corcordance to the ones in the HapMap population of European descent (CEU). The mean (± SD) SNP-wise call rate of all 83 tagging SNPs included in association testing was 99.7 ± 0.82 % and none of these SNPs displayed deviation from HWE after Bonferroni-correction for multiple testing. Rs1659412 was the only SNP showing nominal deviation from HWE (nominal p=0.003, corrected p=0.23; suppl. tab. 1 and 2). Overall, the 18 quality controlled *BDNF* tagging SNPs provided a 100% SNP marker coverage of the *BDNF* gene with 20 kb of flanking sequences according to HapMap

Phase I and II data ($r^2 \geq 0.8$, MAF ≥ 0.01). The 65 quality controlled *NTRK2* tagging SNPs reached a marker coverage of 94.2% for the much larger *NTRK2* gene ($r^2 \geq 0.8$, MAF ≥ 0.1).

3.1.2 Genetic associations in the discovery sample

3.1.2.1 Case-control association with MDD

No case-control association with MDD in *NTRK2*, nor in *BDNF* was detected which withstood correction for multiple comparisons. The smallest p-value (nominal $p=9.3x10^{-3}$, corrected p=0.38) was obtained for rs9969765, a SNP located in the 3' untranslated region of the *NTRK2*-T1 mRNA isoform (NM_001018065 and NM_001018066). There were 7 SNPs in *NTRK2* and 4 SNPs in *BDNF*, which showed nominal p-values below 0.05 (suppl. tab. 1 and 2). For instance, the often-reported functional val66met SNP (rs6265) located in *BDNF* showed nominal significant association with MDD (p=0.035). Neither *NTRK2*, nor *BDNF* was associated with MDD over all tested SNPs per locus assessed by the Fisher product method (FPM).

3.1.2.2 Associations with SA among patients

There was one SNP in the promoter region of *NTRK2*, which was associated with SA among mood disorder patients after correction for multiple comparison over all tested SNPs and genetic models (rs11140714: nominal $p=2.6x10^{-4}$, corrected p=0.043). Overall, 12 SNPs located in *NTRK2* revealed to be associated with SA on a nominal significance level of 0.05 (fig. 12, suppl. tab. 2), whereas in *BDNF* there was none (suppl. tab. 1). The FPM suggested association of *NTRK2* with SA among mood disorder patients: this analysis resulted in a p-value of $4.1x10^{-3}$ over all SNPs tested in the *NTRK2* locus and the three tested genetic models. The p-values for each tested single genetic model (allelic and both alternative recessive-dominant) were in the same range. When correction for all four FPM tests performed in this study (2 phenotypes x 2 genes) was applied, *NTRK2* was still significantly associated with SA among mood disorder patients (Bonferroni-corrected FPM p-value: 0.016). In contrast

to *NTRK2*, *BDNF* showed not even nominally significant association with SA by the FPM ($p>0.05$).

3.1.3 Replication of associations with SA

Since association between *NTRK2* and SA among mood disorder patients was obtained in the discovery sample, the 12 nominally significant SNPs of *NTRK2* associated with SA were selected to be genotyped in an independent German replication sample consisting of 739 patients with unipolar depression (152 with SA) and 1024 matched controls. Again, the FPM on SA was performed with the best genetic model per replicated SNP according to association testing results in the discovery sample. The FPM p-value was significant in the replication sample ($p=3.3x10^{-3}$) indicating real association of the replicated SNPs in *NTRK2* with SA among recurrent depressive patients. This result is carried by 5 out of the 12 tested SNPs (rs10868235, rs1147198, rs1867283, rs1187286 and rs11140800) which show nominal significant association under the same model as in the discovery sample (tab.5). This number of nominally replicated associations is about 16.7 times higher than would have been expected by chance. However, the experiment-wide associated SNP of the discovery sample, rs11140714, was insignificant in the replication sample (suppl. tab. 3), which might be due to differences between the two samples (tab. 2).

3.1.4 Interdependency of the associated SNPs and their location in NTRK2

The linkage disequilibrium (LD) structure among all 65 tested SNPs of the *NTRK2* locus is shown in figure 5C. In particular, LD among the five *NTRK2* SNPs associated with SA among mood disorder patients in both independent samples is shown in figure 5D. By definition, tagging SNPs are not in high LD with each other. Rs10868235 and rs11140800 correlated most among the associated SNPs with an r^2 of 0.63 ($D'=0.94$). On the other extreme, rs1147198 revealed to be independent of the other four associated SNPs. Its pairwise r^2-value with each other SNP was 0.0 ($D'_{max} = 0.14$ with rs1867283).

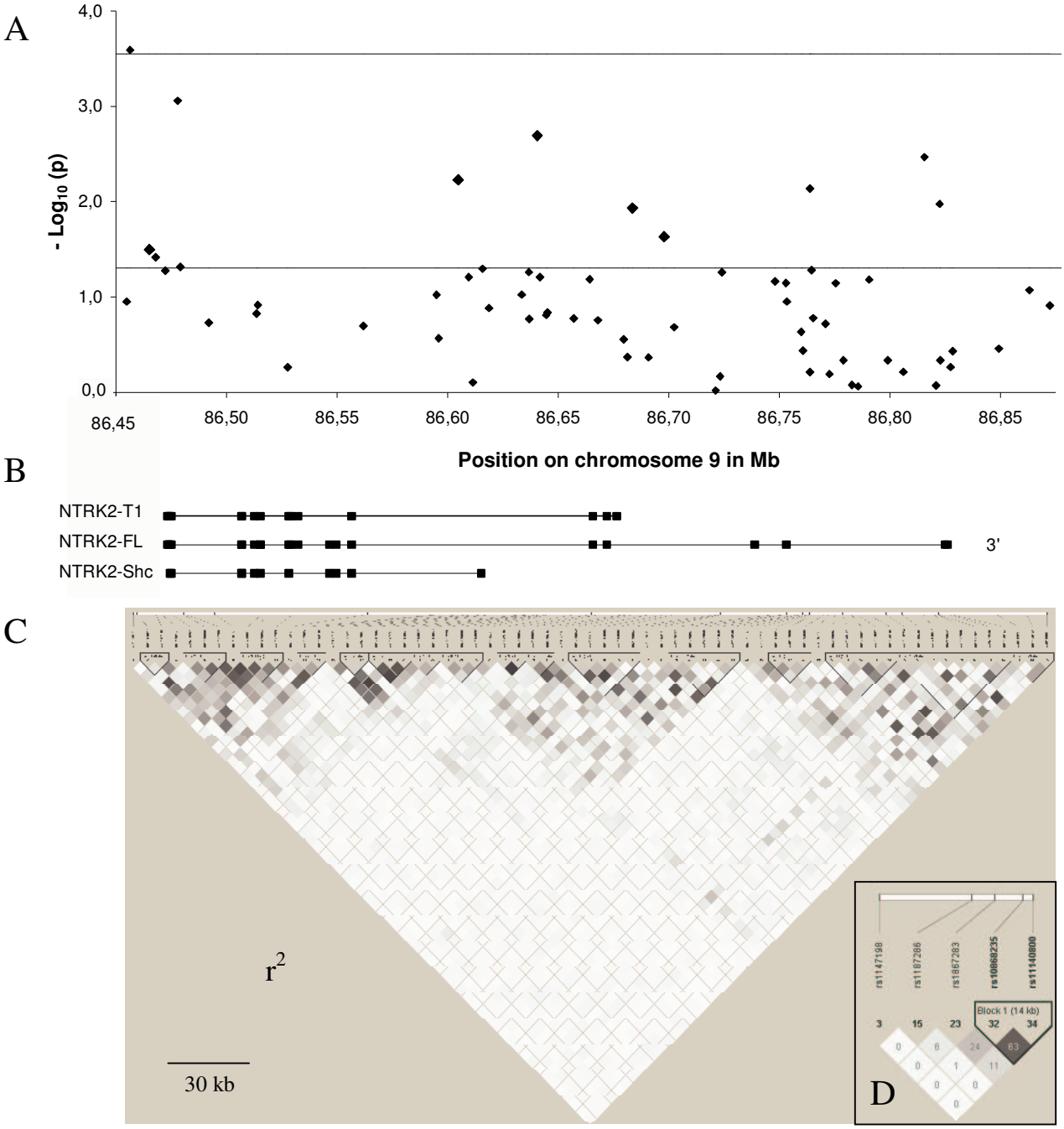


Figure 5: Essential association results of NTRK2 with SA among mood disorder patients are summarized. **(A)** The negative common logarithm of the best genetic model p-value of the 65 tagging SNPs of the discovery sample were plotted against the SNP's chromosome position throughout the *NTRK2* locus on chromosome 9q22. The lower horizontal line indicates the nominal significance level of 0.05 and the upper line the significance level after correction for multiple comparisons over all performed tests (SNPs and genetic models) of the association study on SA in the discovery sample. The five bigger dotes mark the SNPs that could be replicated in the independent German replication sample. **(B)** The most prominent isoforms of *NTRK2* are symbolized and aligned to fig. 5A from their 5'- to 3'-end. The boxes and lines depict the exon-intron structure of *NTRK2*. **(C)** The LD structure over all genotyped tagging SNPs in controls of the discovery sample is shown in pairwise r^2 measurs. The genetic information content of *NTRK2* spanning over 355 kb, decays into 16 haplotype blocks based on D' (21) **(D)** LD between the five SNPs associated with SA among mood disorder patients is displayed (r^2 *100).

The low correlation between the associated SNPs suggests that they map at least four independent casual risk variants for SA in the *NTRK2* gene.
Rs1147198 is located 8 kb 5' of the transcription start site into the putative promoter region of *NTRK2*. The other four associated SNPs are located in intron 14 and 16 of the full-length *NTRK2* transcript (fig. 5A and 5B).

3.1.5 Association with SA in the combined sample

3.1.5.1 Single SNP associations

In the combined discovery and replication sample rs10868235 became significantly associated with SA among mood disorder patients over all 83 tested tagging SNPs with a nominal p-value of 5.7×10^{-4} under a recessive mode of inheritance (Bonferroni-corrected p=0.048). In the separate samples the same recessive model gave lowest nominal p-values. The second to fifth lowest p-values in the combined sample were obtained for the other four SNPs (rs1147198, rs1867283, rs1187286 and rs11140800) which were nominally associated in both separate samples under the same genetic models. In the combined sample, their p-values were lower than in the separate samples (tab.5).
When mood disorder suicide attempters were tested versus healthy controls for the five SNPs mentioned above, again significant association was observed, albeit the effect sizes were smaller than when compared with mood disorder non-suicide attempters (tab.5). These five SNPs did not show case-control association with MDD in the combined sample, indicating a genetic effect on attempted suicide rather than on depressive disorder.
Age at onset of mood disorder was 5.6 years lower in suicide attempters compared with non-suicide attempters in the combined sample ($p=5.1 \times 10^{-8}$). This finding was also significant in the separated samples (tab. 2) and is consistent with reported results from the STAR*D sample (237), in which also more previous depressive episodes are reported among patients with a life history of SA. The same direction of this effect was observed in the samples from the MPIP with approaching significance in the combined sample (p=0.061).

Table 5: Associated SNPs with SA among mood disorder patients. Results are presented for the study-wide most significantly associated SNPs with live history of attempted suicide (SA) among mood disorder patients. All five SNPs are located in *NTRK2*. Genotype distributions, HWE and association p-values of the best genetic and allelic model and odds ratios are given, tested in both independent and the combined sample. In the last column, p-values are shown from post-hoc testing between combined sample mood disorder suicide attempters and healthy controls devoid of SA and a lifetime diagnosis of mood disorders.

SNP	Discovery sample patients		Replication sample patients		Combined sample	
	- SA [1]	+ SA [2]	- SA	+ SA	- SA vs. +SA	Co[3] vs. +SA
rs10868235						
N [all genotypes] (100%)	281	113	589	150	1133	1648
N [CC] (%)	51 (18.1)	33 (29.2)	125 (21.2)	45 (30.0)	254	386
N [CT] (%)	143 (50.9)	52 (46.0)	311 (52.8)	74 (49.3)	580	826
N [TT] (%)	87 (31.0)	28 (24.8)	153 (26.0)	31 (20.7)	299	436
HWE(p)	0.56	0.41	0.17	0.95	0.25	0.89
Allelic p-value	**0.020**		**0.011**		**1.3×10^{-3}**	**4.4×10^{-3}**
Allelic OR (low 95% C.I.)	**1.41 (1.04)**		**1.33 (1.07)**		**1.34 (1.10)**	**1.28 (1.12)**
T-carrier[4] p-value	**0.012**		**8.6x10-3**		**5.7×10^{-4}**	**4.3×10^{-3}**
T-carrier OR (low 95% C.I.)	**1.86 (1.12)**		**1.59 (1.14)**		**1.66 (1.22)**	**1.47 (1.20)**
rs1867283						
N [all genotypes] (100%)	281	113	583	149	1126	1638
N [GG] (%)	63 (22.4)	38 (33.6)	132 (22.6)	45 (30.2)	278	435
N [AG] (%)	141 (50.2)	59 (52.2)	304 (52.1)	71 (47.7)	575	834
N [AA] (%)	77 (27.4)	16 (14.2)	147 (25.2)	33 (22.1)	273	369
HWE(p)	0.92	0.34	0.29	0.62	0.51	0.42
Allelic p-value	**2.0×10^{-3}**		**0.041**		**1.2×10^{-3}**	**0.014**
Allelic OR (low 95% C.I.)	**1.64 (1.20)**		**1.24 (1.00)**		**1.39 (1.14)**	**1.24 (1.06)**
A-carrier[4] p-value	**0.034**		**0.013**		**3.5×10^{-3}**	**0.023**
A-carrier OR (low 95% C.I.)	**1.75 (1.08)**		**1.48 (1.06)**		**1.59 (1.17)**	**1.35 (1.03)**
rs1147198						
N [all genotypes] (100%)	281	113	581	149	1124	1639
N [CC] (%)	10 (3.6)	10 (8.8)	26 (4.5)	14 (9.4)	60	104
N [AC] (%)	101 (35.9)	34 (30.1)	208 (35.8)	44 (29.5)	387	546
N [AA] (%)	170 (60.5)	69 (61.1)	347 (59.7)	91 (61.1)	677	989
HWE(p)	0.29	0.08	0.47	**0.02**	0.62	**0.02**
Allelic p-value	0.500		0.284		0.392	0.287
Allelic OR (low 95% C.I.)	1.14 (0.79)		1.11 (0.86)		1.12 (0.89)	1.07 (0.93)
A-carrier[4] p-value	**0.032**		**0.016**		**2.1×10^{-3}**	**0.025**
A-carrier OR (low 95% C.I.)	**2.63 (1.06)**		**2.21 (1.25)**		**2.31 (1.35)**	**1.63 (1.05)**
rs11140800						
N [all genotypes] (100%)	280	113	581	149	1123	1638
N [CC] (%)	39 (13.9)	22 (19.5)	79 (13.6)	34 (22.8)	174	290
N [AC] (%)	140 (50.0)	63 (55.8)	304 (52.3)	65 (43.6)	572	772
N [AA] (%)	101 (36.1)	28 (24.8)	198 (34.1)	50 (33.6)	377	576
HWE(p)	0.39	0.21	**0.03**	0.15	0.52	0.89
Allelic p-value	**0.023**		**0.039**		**5.9×10^{-3}**	**0.015**
Allelic OR (low 95% C.I.)	**1.41 (1.03)**		**1.22 (1.00)**		**1.30 (1.06)**	**1.25 (1.06)**
A-carrier[4] p-value	0.168		**1.6×10^{-3}**		**1.5×10^{-3}**	**0.045**
A-carrier OR (low 95% C.I.)	1.49 (0.84)		**1.88 (1.29)**		**1.71 (1.20)**	**1.33 (1.01)**
rs1187286						
N [all genotypes] (100%)	281	113	589	151	1134	1652
N [CC] (%)	12 (4.3)	13 (11.5)	27 (4.6)	10 (6.6)	62	102
N [AC] (%)	98 (34.9)	38 (33.6)	191 (32.4)	58 (38.4)	385	607
N [AA] (%)	171 (60.9)	62 (54.9)	371 (63.0)	83 (55.0)	687	943
HWE(p)	0.66	0.07	0.70	0.98	0.87	0.87
Allelic p-value	**0.039**		**0.035**		**6.5×10^{-3}**	0.085
Allelic OR (low 95% C.I.)	**1.42 (1.00)**		**1.33 (1.04)**		**1.38 (1.10)**	1.16 (0.97)
A-carrier[4] p-value	**5.9×10^{-3}**		0.167		**0.012**	**0.033**
A-carrier OR (low 95% C.I.)	**2.91 (1.29)**		1.48 (0.79)		**2.03 (1.19)**	**1.58 (1.05)**

[1] -SA: non-suicide attempters; [2] +SA: suicide attempters; [3] Co: healthy controls; [4] A-carrier: BB vs. AA+AB

Thus, a post hoc analysis of the described genetic associations with SA among mood disorder patients was performed, which controlled for the effects of age at onset of depression and the number of previous depressive episodes, in addition to age and gender. In these models, all 5 SNPs remained significant predictors for SA, with gender and age at onset of depression as significant covariates ($p < 0.05$).

3.1.5.2 Multilocus model and SNP-SNP interaction

Since four of the five SNPs associated with SA among mood disorder patients are in limited or no LD (rs10868235, rs1147198, rs1867283, rs1187286, fig. 5D), they might reflect more than one causal variant of the *NTRK2* gene conferring risk for SA. Therefore, multi-locus analysis was performed in the combined sample allowing two-way and three-way interactions between these SNPs, with age and gender as covariates. Only the three associated SNPs (rs10868235, rs1147198 and rs1867283) with the lowest p-values in the combined sample were included in the interaction analysis to conserve statistical power. The model showed a multiplicative effect of the combined SNPs on SA and became highly significant with a p-value of 4.7×10^{-7}. Main effects for each single SNP, age and gender and the interaction between rs10868235 and rs1147198 were significant predictors ($p<0.05$). The proportion of explained variance of this model was 3.2%, thus small. Figure 6 illustrates the interaction between genotypes of rs10868235 and rs1147198.

The effect size of the association with SA was 3.3 OR (95% C.I. 1.7-6.1, $p=1.5 \times 10^{-4}$) when patients carrying risk genotypes in both interacting loci (CC or CT + CC) were compared with patients without risk genotypes (TT + AA or AC). In analogy, the effect size became even larger when risk genotype carriers in all three markers (rs10868235, rs1147198, rs1867283 - CC or CT + CC + AG or GG) were compared to those without any risk genotype in any marker (TT + AA or AC + AA): 4.5 OR (95% C.I. 2.1-9.8, $p=5.1 \times 10^{-5}$).

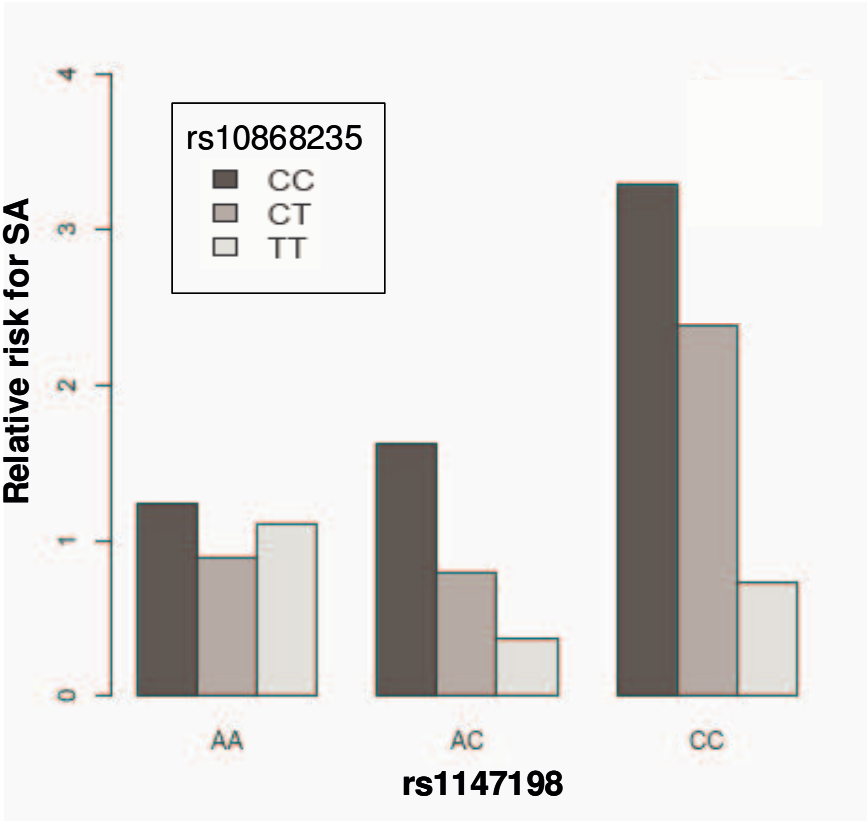


Fig. 6: SNP-SNP interaction effect on life history of SA among mood disorder patients. The SNPs rs10868235 and rs1147198 both located in the *NTRK2* locus showed a main effect on SA among mood disorder patients in both independent samples and an significant interaction on SA in the combinded patient sample. The relative risk for SA is visualized for the genotype combinations of the two independent markers (fig. 5D).

3.2 Results of the Genome-wide Association Study in MDD

3.2.1 Quality control and population stratification in the discovery sample

407'406 SNPs were genotyped in a sample of 353 cases and 366 controls using Illumina Sentrix Human-1 100k and 300k BeadChips. SNPs with a case-wise call rate of less than 98% (15'999 SNPs) and/or displaying deviation from Hardy-Weinberg equilibrium (HWE) at a significance level of below 10^{-5} (12'126 SNPs) were excluded from further analyzes. SNPs with a minor allele frequency (MAF) below 5% (15'401 SNPs) were also ignored due to insufficient power (see methods). This resulted in 365'676 SNPs entering the case-control analysis. Of these SNPs, 16'169 (4.4%) showed a nominally significant deviation from HWE ($p < 0.05$), which is close to the expected number of false positive findings under the null hypothesis of no HWE deviations. The average case-wise call rate for the remaining SNPs was 99.3%.

To test for the possible effects of population stratification, the method of genomic controls was applied. P-values for the allelic Chi square test for all SNPs passing quality control were calculated and the distribution of these p-values was compared with the theoretically expected Chi square distribution (fig. 7). Thereby, the quotient termed lambda (λ) was determined, which is a measure for stratification. This quotient was 1.023 in the discovery sample, which implies that no large effects of population stratification were present. By running EIGENSTRAT, a method based on principal component analysis, no outliers were detected which is another indicator that population stratification was not likely to be a confounder in the discovery sample of the GWAS. Population stratification in the replication sample has been shown to be low (54).

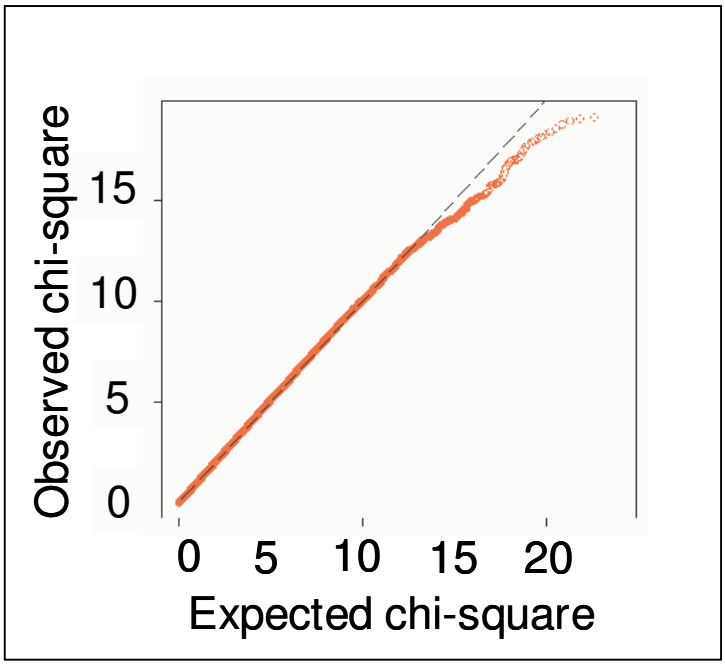


Figure 7: Genomic controls in the discovery sample.
The quantile-quantile plot of the expected against the observed Chi square test statistics for allelic tests of all quality controlled SNPs of the GWAS in MDD is depicted. The quotient between the median of the test statistics and the median of the theoretical distribution is a measure for population stratification. This quotient was 1.023, which implies that no large effects of stratification were present.

3.2.2 Association results from the GWAS in MDD

The common SNP rs1545843 on chromosome 12q21.31 showed experiment-wide significance in a recessive mode of inheritance (AA vs. AG+GG) after permutation-based correction for multiple comparisons over all tested SNPs and genetic models as proposed by Westfall and Young (34, 35) (fig. 8 and tab. 6, N=353/366, nominal $p=7.4x10^{-8}$, corrected p=0.041, OR=2.84). The statistical power to detect such an effect was 93%. Under an allelic model (A vs. G) rs1545843 was nominally significant ($p=9.2x10^{-5}$, OR=1.52). The effect sizes for this association were unexpectedly large when relating it to effect sizes that have previously been reported for variants associated with psychiatric disorders (238, 239).

Seven additional SNPs which are in LD with rs1545843 were also nominally associated with MD ($p=3.2x10^{-4}$ - $3.4x10^{-5}$, tab. 6, fig. 11B and suppl. tab. 5). All these associated SNPs are located in a region spanning about 450 kb on 12q21.31

The 20 best hits on other chromosomal regions are summarized in table 7.

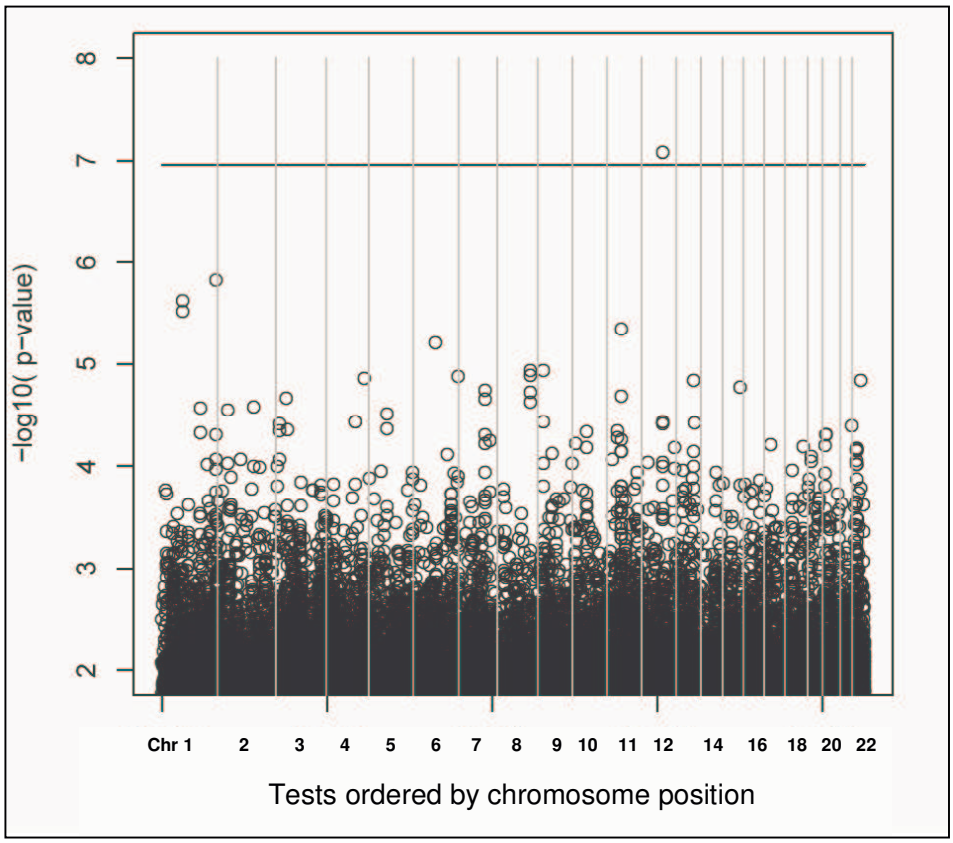


Figure 8: Genome-wide case-control association results in MDD. Rs1545843 on chr12q21.31 revealed to be the only marker with experiment-wide significance. SNPs are aligned according to the number of autosomal chromosomes and their chromosome position (x-axis) and plotted against the negative common logarithm (-log10) of the nominal p-values of each tested SNP's best fitting genetic model (y-axis). The horizontal line represents the experiment-wide significance level after correction for all tested SNPs and genetic models. Tests with -log10(p) < 2 were omitted from the plot.

3.2.3 LD structure among the associated SNPs on 12q21.31

The pairwise r^2-values of the eight associated SNPs ranged from 0.40 to >0.99 in the controls of the discovery sample, a genotype distribution that was in concordance with data from the HapMap population of European descent (CEU) (24) (fig. 9). Thus, all eight SNPs might map the same underlying causative variation associated with MDD.

Table 6: SNPs associated with MDD on 12q21.31. Association results for the experiment-wide significant SNP (rs1545843) in the discovery sample of the GWAS and one representative SNP (rs1031681) of the seven other nominally associated SNPs which are in LD with rs1545843. Results of the discovery, replication as well as of the combined sample are shown (results of all eight associated SNPs are provided in suppl. tab. 5).

SNP	Discovery sample				Replication sample				Combined sample		
Minor / major allele, MAF = RAF, HWE, r^2 in controls, Δ chromosome position in kb	N	P nom.	P corr.	OR	N	P nom.	P corr.	OR	N	P nom.	OR (95% C.I.)
	Co Ca All	Rec. Dom. Allelic	Rec. Dom. Allelic	Rec. Dom. Allelic	Co Ca All	Rec. Dom. Allelic	Rec. Dom. Allelic	Rec. Dom. Allelic	Co Ca All	Rec. Dom. Allelic	Rec. Dom. Allelic
rs1545843	366	7.4e-08	**0.041**	2.84	1022	0.018	0.22	1.27	1388	5.2e-06	1.54 (1.28-1.86)
	353	0.18	1.0	0.80	917	0.013	0.16	0.79	1270	8.0e-03	0.79 (0.67-0.94)
	719	9.2e-05	1.0	1.52	1939	3.6e-03	**0.049**	1.19	2658	1.4e-05	1.27 (1.14-1.42)
Min./maj. allele	A/G				A/G				A/G		
MAF (co,ca,all)	0.41	0.51	0.46		0.45	0.49	0.47		0.44	0.50	0.47
P(HWE, co,ca,all)	3.5e-04	0.14	0.48		0.92	0.98	0.89		0.10	0.40	0.70
pair wise r^2 (co)	0.67				0.67				0.67		
Δ chr. pos (kb)	356										
rs1031681	366	2.8e-04	1.0	2.17	1016	0.014	0.17	1.31	1382	1.8e-04	1.47 (1.20-1.81)
	353	0.102	1.0	0.77	915	6.5e-03	0.089	0.78	1268	2.6e-03	0.78 (0.66-0.92)
	719	2.8e-03	1.0	1.38	1931	1.8e-03	**0.025**	1.21	2650	5.1e-05	1.25 (1.12-1.40)
Min./maj. allele	A/G				A/G				A/G		
MAF (co,ca,all)	0.37	0.45	0.41		0.40	0.45	0.42		0.39	0.45	0.42
P(HWE, co,ca,all)	7.0e-03	1.00	0.27		0.90	0.89	0.93		0.14	0.91	0.32

MAF: Minor allele frequency; RAF: MDD risk allele frequency; HWE: Hardy-Weinberg equilibrium; Co: controls; Ca: cases; Genetic models: Rec.: Recessive (AA vs. AG+GG), Dom.: Dominant (GG vs. AA+AG), Allelic: Additive allele dosage (A vs. G); P-values: nom.: nominal, corr.: multiple comparison corrected.

3.2.4 Association results in the replication and combined sample

The experiment-wide significant SNP (rs1545843) of the discovery sample was genotyped together with the seven correlated SNPs in an independent German case-control sample (N=920/1024) on a Sequenom MALDI-TOF mass-spectrometry platform. Data of one SNP (rs7967594) were excluded because of quality control insufficiencies. The remaining 7 SNPs had a mean case-wise call rate of 97.8% (SD=2.4%) and all SNPs were in HWE ($p<0.05$) over all and on each experimental plate.

Six SNPs were significantly associated with MDD after correction for multiple comparisons (tab. 6 and suppl. tab. 5). Rs1545843 showed the highest significance in an allelic mode of inheritance (nominal $p=3.6x10^{-3}$, corrected $p=0.049$; OR=1.19).

Table 7: The 20 most significant SNPs of the GWAS in MDD. Rs1545843 was the only marker which showed experiment-wide significant association under a recessive-dominant mode of inheritance in the discovery sample (N=353/366). The markers are ordered according to their best model association p-values. Chromosomal region, position and the marker's relative location in a RefSeq annotated gene or to the 3' or 5' end of the transcriptional unit of the closest gene are given in kilobases (kb). HWE p-values and MAFs are shown for the complete sample enclosing both controls and cases.

SNP	Chromosome	Position	RefSeq gene (next gene)	Relative to gene (next gene)	HWE p	MAF	Model	nominal p	corrected p	OR
rs1545843	12q21.31	83088198	(SLC6A15)	670 kb 3'	0.43	0.46	AA vs. AG+GG	$8.3x10^{-8}$	0.038	2.84
rs1342872	1q43	240046844	(WDR64)	15 kb 3'	1.00	0.49	GG vs. AG+AA	$1.5x10^{-6}$	0.510	2.34
rs2125372	1p22.3	85361532	WDR63	intronic	0.50	0.30	GG vs. AG+AA	$2.4x10^{-6}$	0.672	3.56
rs491005	11q12.1	58200376	(GLYAT)	32 kb 3'	0.89	0.49	AA vs. AG+GG	$4.5x10^{-6}$	0.880	2.21
rs163981	6q14.1	80663255	(ELOVL4)	18 kb 3'	0.21	0.36	AA vs. AG+GG	$6.1x10^{-6}$	0.950	2.79
rs1494340	9p22.3	14723895	(FREM1)	3.3 kb 3'	0.38	0.45	CC vs. CT+TT	$1.2x10^{-5}$	0.996	2.07
rs4871503	8q24.13	125651279	MTSS1	intronic	0.63	0.18	C vs. A	$1.2x10^{-5}$	0.996	1.83
rs11755586	6q27	166906756	RPS6KA2	intronic	0.34	0.34	GG vs. AG+AA	$1.3x10^{-5}$	0.999	2.83
rs4544692	4q34.1	174997120	(MORF4)	223 kb 3'	0.85	0.27	GG vs. AG+AA	$1.4x10^{-5}$	0.999	4.12
rs138999	22q13.2	41945184	SCUBE1	intronic	0.52	0.47	TT vs. CT+CC	$1.4x10^{-5}$	0.999	2.07
rs7983386	13q31.3	89899943	(GPC5)	949 kb 5'	0.77	0.46	CC vs. CT+TT	$1.5x10^{-5}$	0.999	2.26
rs4777959	15q26.1	90663219	(ST8SIA2)	75 kb 5'	0.08	0.45	TT vs. CT+CC	$1.7x10^{-5}$	1.000	2.36
rs12673082	7q.31.1	107798556	NRCAM	intronic	0.3	0.19	T vs. C	$1.8x10^{-5}$	1.000	1.79
rs11129469	3p23	31875069	OSBPL10	intronic	0.66	0.44	TT vs. CT+CC	$2.2x10^{-5}$	1.000	2.31
rs10190546	2q23.3	152642720	CACNB4	intronic	0.85	0.25	G vs. A	$2.6x10^{-5}$	1.000	1.68
rs1327136	1q25.3	182206707	GLT25D2	intronic	0.18	0.42	CC vs. CT+TT	$2.7x10^{-5}$	1.000	2.38
rs12466450	2p22.2	36474456	CRIM1	intronic	0.42	0.19	TT vs. CT+CC	$2.8x10^{-5}$	1.000	11.80
rs7713917	5q14.1	78865004	(HOMER1)	20 kb 5'	0.88	0.37	AA vs. AG+GG	$3.1x10^{-5}$	1.000	1.90
rs7690766	4q28.2	131993657	(C4orf33)	1.74 Mb 3'	0.82	0.20	AA vs. AG+GG	$3.7x10^{-5}$	1.000	1.91
rs12685826	9p22.3	14691615	(ZDHHC21)	8 kb 5'	0.92	0.23	G vs. A	$3.7x10^{-5}$	1.000	1.68

The recessive model, which was experiment-wide significant in the discovery sample, showed a nominally significant association (nominal p=0.018, corrected p=0.22, OR=1.27, statistical power: 71%).

In the combined sample, rs1545843 showed the lowest p-values under the recessive model (AA vs. AG+GG) (tab. 6, N=1270/1388, nominal $p=5.2x10^{-6}$, OR=1.52). The seven correlated SNPs showed intermediate p-values ranging from $3.3x10^{-4}$ to $5.2x10^{-6}$ under the recessive model in the combined sample. The p-values for the allelic model were in average 1.6 orders of magnitude lower in the combined sample than in both separate samples over all seven SNPs (mean N=2558, mean $p=9.2x10^{-5}$, mean OR=1.26, 95% C.I. 1.13-1.41, see suppl. tab. 5 for the results of all SNPs).

While 6 tested SNPs independently replicated, the odds ratios of the associations in the second sample were significantly lower than in the discovery

sample and closer to the range expected for psychiatric disorders. These observations warranted a closer inspection of the two samples.

Analysis of haplotypes did not yield associations stronger than those for individual genotypes (data not shown).

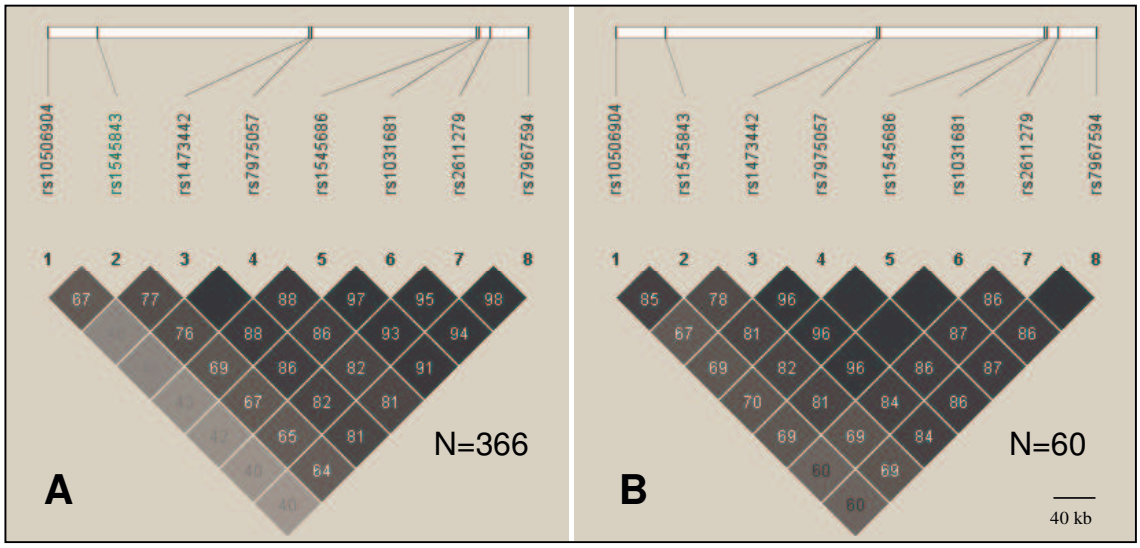


Figure 9: The LD structure of the 8 SNPs associated with MDD on 12q21.31 is presented in **(A)** controls of the discovery sample (N=366) and **(B)** in the HapMap Central European population (CEU, N=60). Pairwise r^2 values multiplied with 100 are shown for each SNP pair. The experiment-wide significant SNP (rs1545843, SNP 2) of the GWAS is in LD with the other seven nominally associated SNPs. LD data are concordant between both samples.

3.2.5 Nominal DHW in controls of the discovery sample

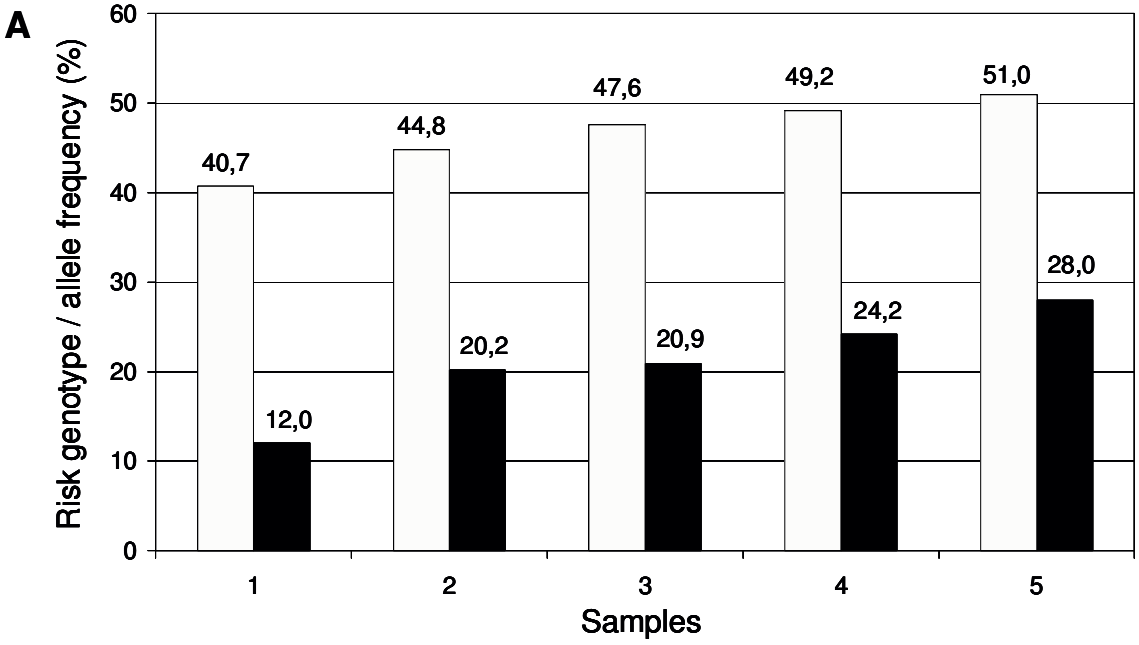
The SNPs associated with MDD all showed nominal DHW in the controls of the discovery sample (tab. 6 and suppl. tab. 5). The observed deviations were insignificant after correction for multiple testing and might thus be due to chance. However, DHW may also reflect genotyping errors (51, 240), which seems to be unlikely for eight linked SNPs. Moreover, since they show a comparable LD structure as the one obtained of the HapMap (24) individuals from the Central European (CEU) population (fig. 9b). However, to exclude the possibility of genotyping errors, two SNPs (rs1545843 and rs1031681) were re-genotyped in 339 patients and 365 controls of the discovery sample with an independent method based on double-stranded DNA melting curves. The obtained concordance rate with the Illumina genotypes was 99.93% (1

discordant genotype out of 1408). All these SNPs showed heterozygote excess in the discovery controls. An excess of heterozygotes might indicate population admixture. This second source for DHW was excluded, since neither the genomic control method (λ = 1.023, fig. 7) nor EIGENSTRAT analysis gave any indication for population stratification. Finally, DHW might also reflect selection introduced by sample screening protocols (53). Actually, the controls of the discovery sample underwent a most stringent screening procedure, with history of any axis I psychiatric disorders as exclusion criterion (see next paragraph). This hypothesis is also supported by the fact that the SNPs were in HWE in the less strictly screened replication controls and in a population reference sample of unscreened individuals from Northern Germany (N=493, fig.10, sample 3). In fact, differences in screening procedures would alter the number of misclassified controls in these samples, which is highly relevant in MDD with a lifetime prevalence as high as 16% (65) and might be related to the differences in the observed effect sizes in the discovery and replication sample.

3.2.6 The number of risk-allele carriers is correlated with the expected risk for MDD

The controls of the discovery sample underwent the most stringent screening procedure of the 3 control samples investigated in this study. Using the face to face Composite International Diagnostic Interview (CIDI) (217, 218), subjects positive for one or more life-time diagnosis of any axis I psychiatric disorder were excluded from the study, including mood disorders, schizophrenia, anxiety disorders, alcohol dependence, drug abuse, obsessive/compulsive disorders, post-traumatic stess disorder, dissociative disorders, somatoform disorders, and eating disorders. This screening procedure resulted in an overall exclusion rate of 49.7%.

The controls of the replication sample were only screened for the absence of anxiety and mood disorders using the Composite International Diagnostic Screener (CID-S) (219). This is a written questionnaire including only the CIDI core items. The CID-S is considered to be less stringent than the CIDI.



B

In-patient portion:

X	(4) Replication Patients		(5) Discovery Patients	
	OR [95 % C.I]		OR [95 % C.I]	
	A vs. G	AA vs. AG+GG	A vs. G	AA vs. AG+GG
(1) Discov. contr.	1.41 [1.22-1.63]	2.34 [1.74-3.13]	1.52 [1.27-1.80]	2.84 [2.05-3.95]
(2) Repl. contr.	1.19 [1.07-1.33]	1.27 [1.06-1.52]	1.28 [1.11-1.48]	1.54 [1.22-1.94]
(3) Pop. ref.	1.07 [0.94-1.21]	1.21 [0.97-1.51]	1.15 [0.98-1.35]	1.47 [1.13-1.92]

Screening stringency

Figure 10: Comparison of associated SNPs to a population reference sample. (A) The MDD risk A-allele and risk AA-genotype frequencies (in %) between most stringently screened controls (discovery sample, 1), less stringently screened controls (replication sample, 2) and the unscreened population reference sample (PopGen, 3) and patients from the replication (4) and the discovery sample (5) exemplified by rs1545843, the SNP which reached experiment-wide significant association with MDD in both case-control samples. **(B)** Odds ratios (OR) and the 95% confidence intervals (C.I.) between different sample comparisons are illustrated. Inpatient proportions in the patient samples are indicated by pie charts (sample 4: 60.3% inpatients, sample 5: 100% inpatients).

For instance, 18.2 % (± 2.3) of CID-S negatives (no item endorsed) were assessment positives for a lifetime DSM-IV psychiatric diagnosis in a subsequent CIDI (219). The exclusion rate in the replication controls was 19.2% which is in concordance with the combined population prevalence of anxiety and mood disorders (241).

Finally, the population reference sample from Northern Germany is likely to contain individuals with a history of MDD or at risk for MDD at the population based prevalence rate of 16% (65).

Moreover, the discovery sample patients might have been more severely affected by MDD than the patients from the replication sample, since discovery patients were all inpatients of a tertiary clinic, whereas up to 40% of replication patients were outpatients from doctor's practices.

When the five case and control samples were ordered according to an increasing number of potential individuals at risk for MDD, a steady increase in risk allele and risk genotype frequencies was observed (fig. 10A). When all controls were tested against all patients (N=3149) the recessive model p-value for rs1545843 was 7.0×10^{-6}. However, comparison of any of the two patient samples against the discovery controls revealed similarly high ORs (2.4 to 2.8; recessive model) while comparison with the replication controls showed lower but still significant OR for both case samples (1.2 and 1.4). Comparison with the population reference still confirmed an overrepresentation of the risk alleles and genotypes in cases, but only the recessive model comparing discovery patients with population controls was nominally significant (fig. 10B). The differences in ORs between our samples thus seem to be mainly explained by the screening stringency of controls.

3.2.7 Rs1545843 associations with demographic and illness-related variables

No genotype-dependent differencies of rs1545843 were observed regarding age, gender and German origin tested in both case and control samples separately. Furthermore, in both case samples no associations were observed for recurrence of MDD, age at onset of MDD, number of previous depressive episodes, first degree family history of MDD and life time attempted suicide status (data not shown).

3.2.8 Genomic context of the associated region on 12q21.31

The eight associated SNPs span a gene desert region of about 450 kb in size mapping neither to any annotated gene nor to predicted human mRNAs with the exception of some small human spliced expressed sequence tags (EST, fig. 11A). Among unspliced ESTs in the associated region, there are at least seven different single hit or multi-hit ESTs which have been detected in brain tissue (H10492, H11641, H18024, AV746466, CD518115, CD519334 and BF967336) (226).

The next gene annotated by RefSeq (211, 226) ends 287 kb further distal to the 450 kb region of association on 12q21.31. This gene belongs to the solute carrier 6 family and codes a sodium-dependent branched-chain amino acid transporter (*SLC6A15*, NM_182767) with highest gene expression observed in neurons of the human brain (242, 243). The nearest gene on the proximal side, transmembrane and tetratricopeptide repeat containing 2 gene (*TMTC2*, NM_152588) ends 989 kb apart from the region of association. It is expressed in a variety of human tissues but its function is largely unknown. According to HapMap (226, 244, 245) and Perlegen (225) genotyping data several recombinational hotspots were predicted between the associated region and the flanking genes (fig. 11A), making it unlikely that the underlying functional variation might directly hit a classical promoter region or the open reading frame of a known gene.

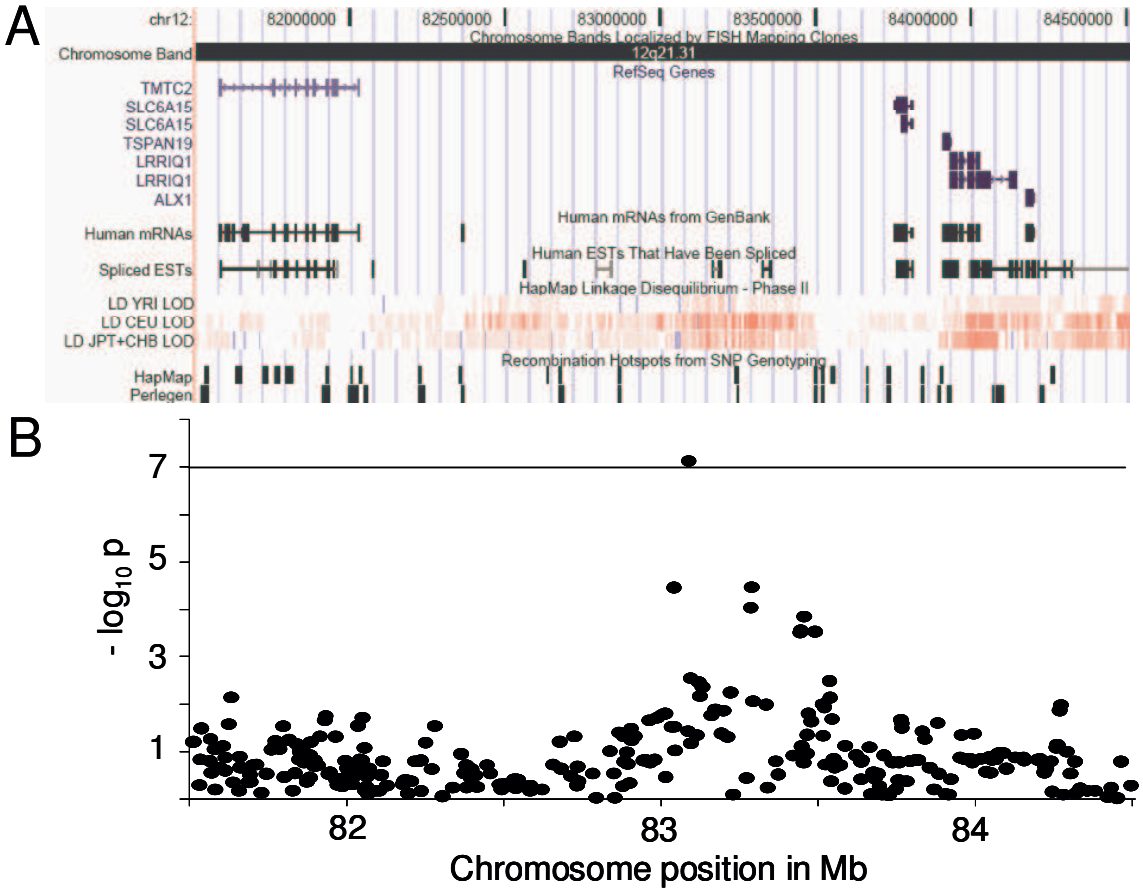


Figure 11: Genomic context of the associated region on 12q21.31. (A) shows relevant features of the genomic architecture of a 3 Mb region comprising rs1545843 according to the UCSC Genome Browser: RefSeq annotated genes (blue), human mRNAs and expressed sequence tags from Genebank (black), HapMap Linkage Disequilibrium (red: high LD, white: low LD) and recombination hotspots from SNP genotyping data provided by HapMap and Perlegen (black). The associated region did not map to any known gene (compare with 2B). The flanking next genes to the region of association are *SLC6A15* (+287 kb), a solute carrier family 6 gene that codes for a sodium-dependent branched amino acid transporter with the highest gene expression observed in neurons of the brain and *TMTC2* (-989 kb), the transmembrane and tetratricopeptide repeat containing 2 gene which has a provisional annotation status and a rather unknown function. **(B)** The negative common logarithm ($-\log_{10}$) of the best model p-values (y-axis) of all tested SNPs in the shown region from genome-wide case-control association testing in the discovery sample were plotted against the SNP's chromosome positions (x-axis). The horizontal line across this figure indicates the genome-wide significance level of the experiment. The dot above this line represents the $-\log_{10}$ p-value of rs1545843. Seven other SNPs reached high nominal significance.

3.3 Genotype-related functional correlates with SNPs from the GWAS

3.3.1 Risk allele carriers display lower SLC6A15 mRNA levels

To adress the functional role of the identified locus, I analysed genome-wide mRNA expression data from lymphoblastoid cell lines of unrelated HapMap individuals (223, 224). A previous genome-wide study had shown that the median distance between SNPs and genes whose mRNA expression is significantly regulated by them is approximately 30 kb, with distances up to a maximum of 1 Mb (246). Therefore, all five RefSeq annotated genes within 1.5 Mb proximal and distal of rs1545843 on 12q21.31 (fig.11A*, TMTC2*, *SLC6A15*, *TSPAN19*, *LRRIQ1*, *ALX1*) were assessed. Expression levels of all seven available probes (3 for *SLC6A15*) were related to genotypes of two SNPs which tag the total of the eight SNPs associated with MDD, rs1545843 and rs1031681 (tab.6). The allelic and both alternative recessive-dominant genetic models per SNP and probe were tested and corrected for ethnicity and multiple comparisons. Only the association of rs1031681 with the expression of the full-length mRNA isoform of *SLC6A15* reached experiment-wide signficance (fig.12A and B). The risk-allele carrier group (AG+AA) showed lower mRNA expression levels of *SLC6A15* than the non-risk genotype (GG) carrier group (nominal $p=1.5x10^{-3}$, corrected p=0.041, N=204). Rs1031681 showed also nominally significant associations with the short mRNA isoform of *SLC6A15* as well as with *TMTC2* (fig.12A). The second tested tagging SNP (rs1545843) failed to reach significant association in this analysis. However, when the analysis was restricted to the subset of individuals of European descent both SNPs showed nominally significant association with full-length *SLC6A15* mRNA expression (rs1031681: N=57, p=0.019; rs1545843: N=53, p=0.044).

A

rs1031681	AA+AG vs. GG		A vs. G	
mRNA	P nom.	P corr.	P nom.	P corr.
SLC6A15 FL	**0.0015**	**0.041**	0.0023	0.086
SLC6A15 S	0.105	0.904	0.018	0.349
TMTC2	0.043	0.356	0.031	0.353
ALX1	0.062	0.621	0.053	0.608
TSPAN19	0.826	1	0.300	0.993
SLC6A15 FL/S	0.414	1	0.715	1
LRRIQ1	0.476	1	0.461	1

B

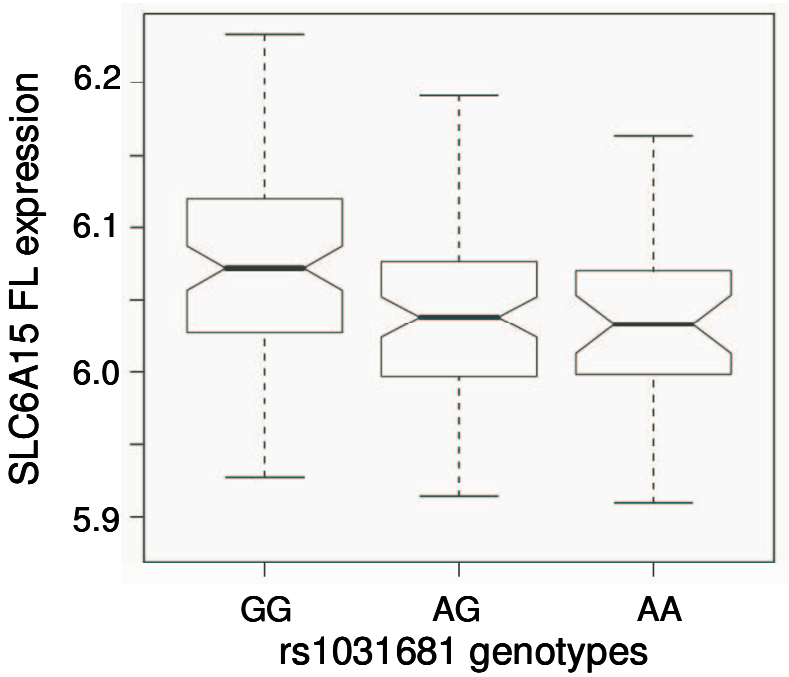


Figure 12: *SLC6A15* mRNA expression per rs1031681 genotype group measured in EBV-transformed lymphoblastoid cell lines of unrelated HapMap subjects from different human populations (GENEVAR and HapMap data). **(A)** MDD risk allele carrier (AA+AG) revealed to have reduced full-length (FL) *SLC6A15* mRNA expression levels compared to non-risk allele carriers (GG) signified by non-overlapping notches between the named groups in part B. Several other SNPs nominally associated with MDD which are in high LD (r^2>0.9) with rs1031681 showed comparable effects (data not shown). None of the other genes (*TMTC2*, *ALX1*) also flanking the region of association with MDD showed experiment-wide significant geno-type-specific alteration in expression levels. The nominal p-values (P nom.) are corrected for ethnicity and the corrected p-values (P corr.) are further adjusted to multiple comparisons. *SLC6A15* S: Short mRNA isoform of *SLC6A15*. N=204 for all tested probes **(B)** On the x-axis the 3 genotype groups of rs1031681 are plotted against normalized *SLC6A15* mRNA levels on the y-axis (group means: solid horizontal lines). The upper and lower ends of the dotted lines per group represent the maximal and minimal expression levels, respectively. The boxes upper and lower edges indicate the upper and lower quartiles.

Since the functional variant might regulate a so far not annotated gene, this analysis was repeated by including all available probes (N=+6), also those for expressed sequence tags (EST) from GenBank in the same genomic window. None of these ESTs showed nominally significant genotype-dependent expression (tab.8, for EST and RefSeq gene annotation see tab.3).

Table 8: SNP with gene expression association analysis with all probes of a 3 Mb genomic window including the region of association with MDD on 12q21.31.

Transcript	rs1031681			rs1545843		
	Best model [4]	[5] P nom.	P emp.	Best model [1]	P nom.	P emp.
SLC6A15 FL [1]	AA+AG vs. GG	0.0015	0.0015	A vs. G	0.469	0.487
SLC6A15 S [2]	A vs. G	0.018	0.027	AA vs. AG+GG	0.807	0.811
TMTC2	A vs. G	0.031	0.044	AA+AG vs. GG	0.114	0.111
ALX1	A vs. G	0.053	0.069	AA+AG vs. GG	0.109	0.109
TSPAN19	AA vs. AG+GG	0.101	0.106	AA+AG vs. GG	0.645	0.650
Hs.557975	AA+AG vs. GG	0.104	0.107	AA+AG vs. GG	0.34	0.337
Hs.677048	AA+AG vs. GG	0.263	0.253	A vs. G	0.077	0.091
Hs.339071	AA vs. AG+GG	0.308	0.305	AA vs. AG+GG	0.314	0.317
GI:875314	AA vs. AG+GG	0.331	0.323	AA vs. AG+GG	0.248	0.248
Hs.208111	AA vs. AG+GG	0.371	0.377	A vs. G	0.349	0.374
SLC6A15, FL/S [3]	AA+AG vs. GG	0.414	0.414	AA+AG vs. GG	0.333	0.334
LRRIQ1 FL/S	A vs. G	0.461	0.490	A vs. G	0.271	0.290
Hs.675906	AA+AG vs. GG	0.562	0.563	A vs. G	0.291	0.313

[1] FL: full-length mRNA isoform, [2] S: short isoform, 3 FL/S: both isoforms, [4] Best genetic model
[5] P-values: nom: nominal (ethnicity adjusted), emp: empirical (permutation based)

In conclusion, MDD risk allele carriers showed lower *SLC6A15* mRNA expression levels in EBV-transformed lymphoblastoid cell lines from different human ethnicities. The associated SNP showed less if any association with expression levels of other RefSeq genes or ESTs annotated in this region.
Because *SLC6A15* expression is highest in the hippocampal region (247, 248), the following experiments addressing the role of the associated SNPs in biomarkers for MDD and animal models of chronic stress focused on this brain area.

3.3.2 Healthy risk allele carriers display lower hippocampal N-acetylaspartate and glutamine/glutamate levels

To further validate the functional revalence of the SNPs associated with MDD in the GWAS, I started a colaboration with Philipp Säeman from the research group of Michael Czisch at the MPIP's magnetic resonance imaging (MRI) facility.

In a subsample of patients and controls of the replications sample, brain imaging data were available. Demographic and clinical characteristics of the MRI/^{1}H-NMR sample are given in table 4 of the material and method part of this dissertation.

Left hippocampal ^{1}H-NMR spectroscopy detected higher values for choline and phosphocholine containing compounds (Cho) in patients (N=80) compared with controls (N=81, $F_{1,156}$=5.530, p=0.020). Lower levels of N-acetylaspartate (NAA, $F_{1,75}$=9.170, p=0.003) and unresolved glutamate/glutamine (Glx, $F_{1,75}$=5.127, p=0.026) were found in patients fulfilling criteria for MDD (N=32) compared with patients not fulfilling criteria of MDD at the time of ^{1}H-NMR (N=48) (fig.13B), converging with the reported sensitivity of ^{1}H-NMR spectroscopy to neurochemical alterations in mood disorders (249).

Because mood state and medication influence hippocampal neurochemistry, genotype effects on these measures were assessed in healthy, unmedicated control subjects. The same tagging SNPs were analysed (rs1545843 and rs1031681) as for association testing wih gene expression. Multivariate analysis detected an allelic effect of rs1031681 on hippocampal metabolites (Wilks' lambda: 0.683, $F_{2,75}$=2.976, p=0.002) with univariate comparisons pointing towards NAA ($F_{2,75}$=6.143, p=0.003) and creatine (Cr) ($F_{2,75}$=4.941, p=0.010), and also towards Glx with approaching significance ($F_{2,75}$=2.839, p=0.067). Equally, the dominant model was significant at the multivariate level (Wilks' lambda: 0.761, $F_{1,76}$=4.529, p=0.001), demonstrating significantly lower NAA, Glx and Cr in risk allele carriers (AA+AG) compared with non-risk allele carriers (GG) (NAA: $F_{1,76}$=5.575, p=0.021; Glx: $F_{1,76}$=5.752, p=0.019; Cr: $F_{1,76}$=4.009, p=0.049) (fig.13C). Analysis of glutamate revealed lower mean levels in risk allele carriers (conservative fit accuracy threshold ≤ 20% [N=61]: $F_{1,56}$=1.754,

p=0.190, lenient fit accuracy threshold ≤ 25% [N=73]: $F_{1,68}$=4.460, p=0.038). Multivariate analysis was also significant for rs1545843 with the univariate analysis pointing selectively towards NAA (dominant model: $F_{1,76}$=5.333, p=0.024).

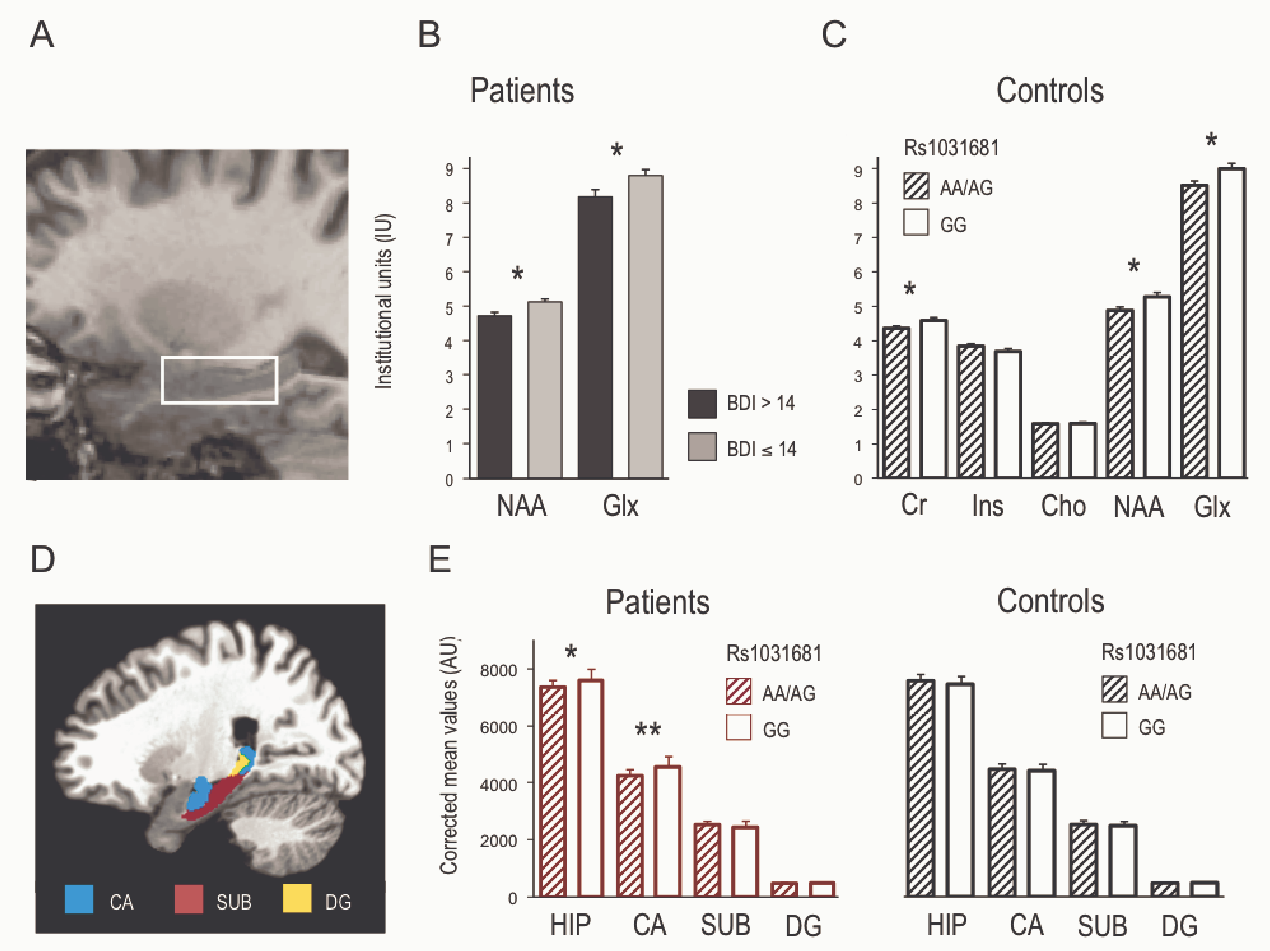


Fig. 13: Genotype dependent differences of hippocampal neurochemistry and macroscopical structure. (A) Typical position of single voxel [30×17×12 mm3] for acquisition of ^{1}H-NMR from the left hippocampus as defined in a sagittal T1-weighted anatomical image. **(B)** Significantly lower NAA and Glx levels were found in the subgroup of patients with high depression scores at the time of ^{1}H-NMR. **(C)** Healthy controls carrying the risk allele of the rs1031681 polymorphism exhibited lower Cr, NAA and Glx levels. Mean values corrected for age, gender and cerebrospinal fluid content of the voxel and one standard error of mean (S.E.M.) are shown. **(D)** Sagittal view of left hippocampal subregions as derived from cytoarchitectonic probability maps of cornu ammonis (CA), subiculum (SUB) and dentate gyrus including fascia dentata and CA4 (DG), all constituting the hippocampal formation (HIP). A projection of the binary maps of the left hemisphere onto a single subject template (C. Holmes brain) in Montreal Neurological Institute (MNI) space is shown. **(E)** Genotype-dependent differences of the total HIP volume, strongest in the CA subregion, were detected in the patient group, but not the control group. P-values of univariate analyses of covariance within the patient and control group are given. Mean values corrected for age, gender and total brain volume, and five S.E.M. are depicted. *p<0.05. **p=0.001 within patients; respective group × genotype interaction effect robust to correction for multiple tests (nominal p=0.0006, corrected p=0.043).

3.3.3 Genetic associations with volumetric measures of the hippocampal complex

In a subsample of patients and controls of the replication sample, brain imaging data were available. Sample characteristics are given in table 4. Again, the two tagging SNPs (rs1545843 and rs1031681) were tested. For rs1031681, a case-control × genotype interaction effect was found for the left total hippocampal (HIP) volume ($F_{3,154}$=9.103, p=0.0030), with approaching significance also for the right HIP volume ($F_{3,154}$=3.549, p=0.061). Subregional analysis revealed a case-control × genotype interaction for the left cornu ammonis (CA) ($F_{3,154}$=12.349, p=0.0006; Bonferroni-corrected p=0.043) and genotype main effects for the left CA (lower volumes in risk allele carriers; $F_{3,154}$=5.343, p=0.0221) and the left subiculum (higher volumes in risk allele carriers; $F_{3,154}$=6.062, p=0.0149) (tab.9). Analysis of the separate groups demonstrated that interaction effects were driven by stronger genotype effects in patients compared with controls (fig.13D and E). No effects were observed for the dentate gyrus of the hippocampus and the control region (precentral gyrus). For SNP rs1545843, interaction effects with left total HIP volumes were observed with approaching significance (dominant model, left total HIP volume, case-control × genotype interaction effect, $F_{3,154}$=3.128, p=0.079).

Table 9: Regional volumetry results for rs1031681. Results of a joint analysis of the regional brain (sum of GM and WM compartments) volumes across patients and controls for rs1031681 (AA+AG vs. GG). Separate analysis of patients and controls (data not shown) located case-control × genotype interaction effect into the patient group with spurious genotype effects in the control group. Additional analysis of GM and WM compartments demonstrated that the case-control × genotype interaction effect on total hippocampal volumes was mainly driven by WM (left: $F_{3,154}$=7.187, p=0.008; right: $F_{3,154}$=3.489, p=0.064), whereas the subregional left CA effect was driven about equally by GM ($F_{3,154}$=5.005, p=0.027) and WM ($F_{3,154}$=7.131, p=0.008).
[1] Each multivariate test was performed on the left and right subregions (left: p<0.05 for genotype and group × genotype effect; right: n. s.). F and p values of univariate tests are given.
* p<0.05
** Exact p-value 0.00058, significant at Bonferroni corrected threshold of 0.05 / 72 = 0.00069.
Abbreviations: CA, cornu ammonis; SUB: subiculum; DG: dentate gyrus.

Region	Case-control		Genotype		Case-control x Genotype	
	F	P	F	p	F	p
Left total hippocampus	<0.001	0.987	0.469	0.494	9.130	**0.003***
Right total hippocampus	0.232	0.631	0.106	0.745	3.549	0.061
Left hippocampal subregions[1]						
CA subregion	0.017	0.896	5.343	**0.022***	12.349	**<0.001****
SUB subregion	0.047	0.828	6.062	**0.015***	0.065	0.798
DG subregion	0.002	0.963	0.190	0.664	0.109	0.742
Right hippocampal subregions[1]						
CA subregion	0.368	0.535	1.821	0.179	2.645	0.106
SUB subregion	0.001	0.970	2.042	0.155	1.405	0.238
DG subregion	0.255	0.614	0.237	0.627	0.383	0.537

3.4 Evidence for a role of SLC6A15 in stress vulnerability

3.4.1 Reduced *SLC6A15* hippocampal mRNA levels in stress vulnerable mice

Chronic stress represents a main risk factor for a variety of diseases, including anxiety and depression (73, 74, 250, 251). Therefore, I started a collaboration with Mathias Schmidt of the MPIP´s molecular stressphysiology group to exploit his recently developed and extensively validated mouse paradigm on chronic social stress (215, 216) to study *SCL6A15* expression in stress susceptible versus stress resistant mice . According to this animal model, CD1 outbred mice are exposed to a unpredictable social stressor (unknown mates) to which they can not adapt for seven weeks during their adolescence and young adulthood (fig. 14a). In this experiment 120 mice were subjected to this chronic social stress paradigm. The experimental time course is given in figure 14b.

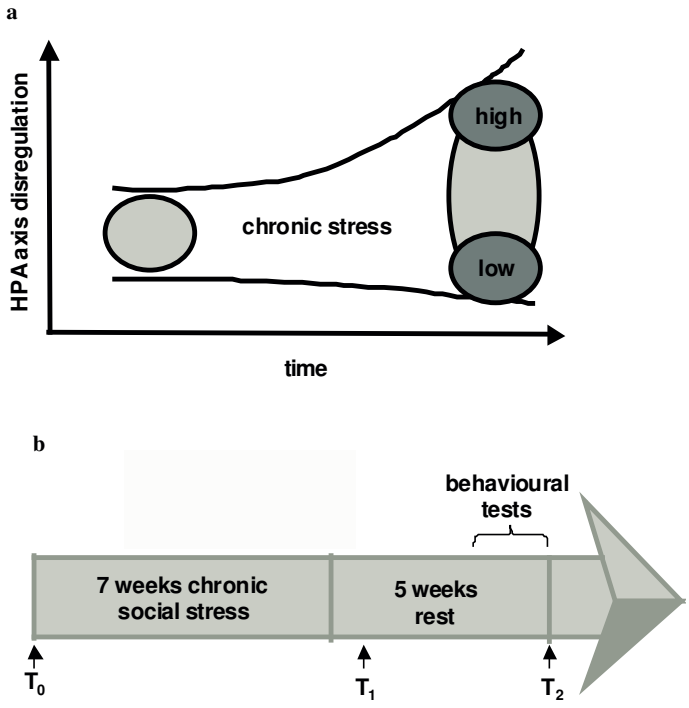


Figure 12: A mouse model of chronic social stress.
(a) Schematic hypothesis of stress susceptibility. If a large cohort of outbred animals (e.g. CD1 mice) is subjected to chronic stress, a subgroup of those animals will be lastingly affected by the stress exposure (high-responders; stress susceptible phenotype), while others will recover quickly (low responders; stress resistant phenotype).
(b) Experimental time course.
120 animals are subjected to 7 weeks of chronic social stress starting at postnatal day 28 (T_0). After one week of recovery, a blood sample is taken for corticosterone determination (T_1). At this time animals are divided into the two groups of stress resistance and stress susceptibility according to stress-related phenotypes (peripheral corticosterone secretion).
Behavioral tests took place between 4 and 5 weeks after the end of stress exposure (T_2).

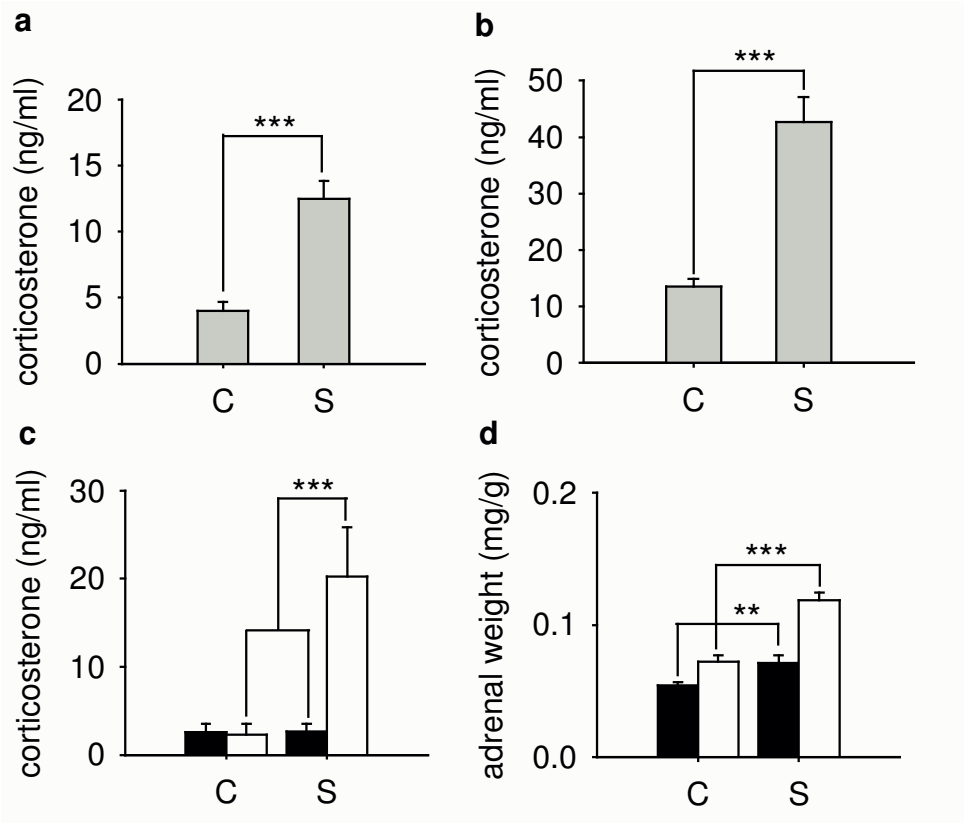


Figure 15: Hormonal and physiological changes following 7 weeks of chronic social stress exposure. (a) After one week of rest under single housed conditions, animals (S, N=120) that were exposed to the chronic stress procedure still show a significantly increased basal secretion of corticosterone compared to unstressed control animals (C, N=40). **(b)** As the groups consist of a large number of animals a big inter-group variation was observed: The top 20%of control and stress exposed animals in terms of basal corticosterone secretion displayed significantly different corticosterone levels (condition: $F[1,44] = 35.872$, $p<0.001$). In contrast, there was no significant difference between the low 20% of each population (data not shown). **(c)** To investigate the stability of this effect basal corticosterone levels of these animals 5 weeks after the end of the chronic stress procedure were assessed. At this late time point still significantly higher basal corticosterone levels were observed in animals that belonged to the top 20% group (white bars) one week after stress exposure compared to the low 20% animals (black bars) from the stress exposed group and controls (condition: $F[1,32] = 5.217$, $p<0.05$; extremes: $F[1,32] = 4.788$, $p<0.05$; interaction: $F[1,32] = 5.052$, $p<0.05$). **(d)** The persistent effects in basal corticosterone secretion are mirrored by the adrenal weights of the animals after 5 weeks of rest (condition: $F[1,30] = 31.483$, $p<0.001$; extremes: $F[1,30] = 33.215$, $p<0.001$; interaction: $F[1,30] = 6.947$, $p<0.05$). **P<0.01; ***P<0.001.

As was previously reported (215, 216), stressed mice again showed severe physiological and endocrine changes related to chronic stress compared to control animals including increased adrenal weight, reduced thymus weight and elevated morning blood corticosterone levels after one week of rest under stress-free conditions (fig. 15a). As the groups consist of a large number of

animals a large inter-group variation could be observed which allowed to group the animals in terms of basal corticosterone secretion. Stress susceptible animals were definded as the subgroup of stressed animals with the top 20% of corticosterone secretion displaying significantly different mean corticosterone levels compared to the top 20% sub-group of control animals (fig. 15b). The low 20% sub-group of stressed animals were defined as stress resistant, since in contrast, there was no significant difference between the low 20% sub-groups of stess exposed and control animals (data not shown).

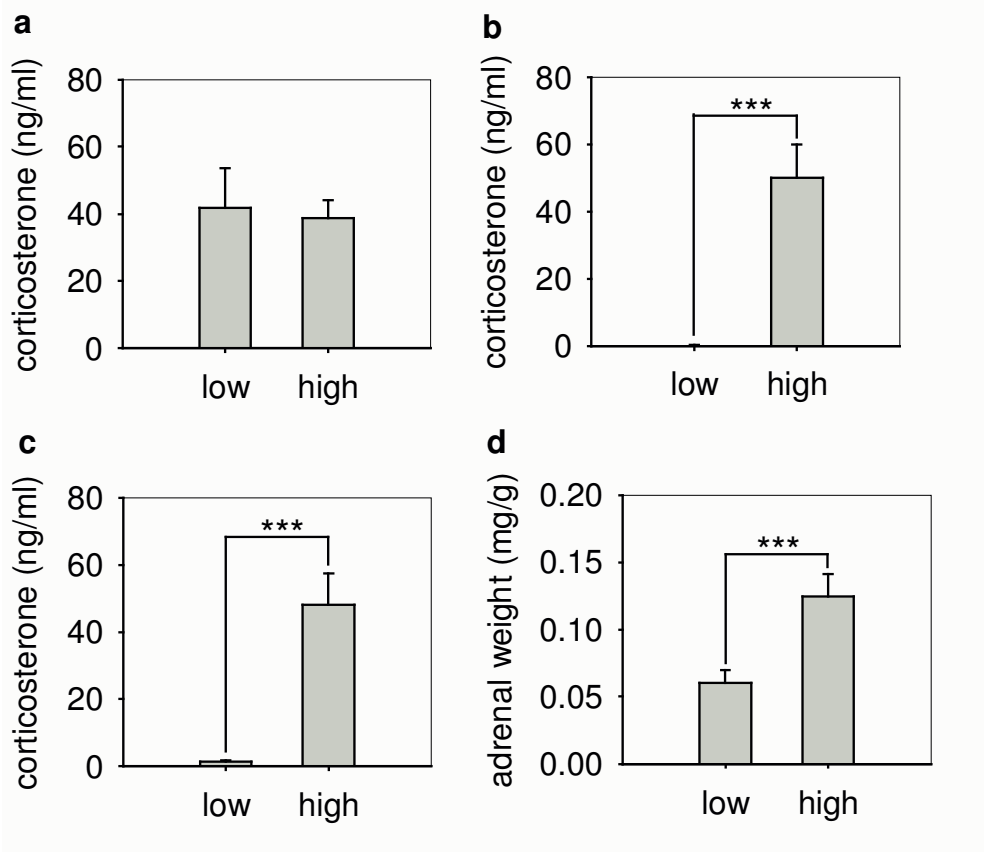


Figure. 16: Neuroendocrine characteristics of animals selected for microarray analysis. (a) At the age of 28 days animals displayed no difference in basal corticosterone secretion (low: 6 most stress resistant, high: 6 most stress susceptible animals). **(b)** After one week of rest following the chronic stress paradigm animals with a very low basal corticosterone secretion (resistant phenotype) and animals that still displayed very high basal corticosterone levels (susceptible phenotype) were selected. **(c)** The resistant and susceptible phenotype in these animals was stable over time, as the same difference in basal morning corticosterone secretion was observed 5 weeks following the end of the stress paradigm. **(d)** Animals with low or high basal corticosterone secretion also differed significantly in relative adrenal weight. ***P<0.001.

To investigate the stability of this effect, basal corticosterone levels of these animals 5 weeks after the end of the chronic stress procedure were assessed. At this late time point still significantly higher basal corticosterone levels were observed in animals that belonged to the top 20% group (susceptible group) one week after stress exposure compared to the low 20% animals (resistant group) from the stress exposed group and compared to controls (fig. 15c). Moreover, the persistent effects in basal corticosterone secretion are mirrored by the adrenal weights of the animals after 5 weeks of rest (fig. 15d).

The 6 most stress-affected and the 6 most resistant individuals were selected from the group extremes in the formerly chronically stressed animals. The neuroendocrine characteristics of these animals are given in figure 16. Pooled mRNA samples of laser-assisted micro dissections from the CA1 region of the hippocampus from both experimental groups (fig. 4) were subjected to genome-wide Illumina Sentrix BeadChip microarrays.

When expression data for the probes specific for the genes *TMTC2*, *SLC6A15*, *TSPAN19*, *LRRIQ1*, and *ALX1* were analyzed, *SLC6A15* mRNA levels were 1.9-fold reduced in the CA1 hippocampal region in stress susceptible versus stress resistant mice. Expression levels of the other genes did not exceed background noise in the CA1, and are thus likely not expressed at higher levels in this brain region (tab. 10). The human *TSPAN19* could not be examinded in this analysis, because there is no mouse homologue.

Table 10: Gene expression microarray data from a mouse model of chronic social stress.
SLC6A15 mRNA levels were significantly reduced in stress resistant (R) compared to stress susceptible (S) mice in the CA1 region of the hippocampus (high differential score and high signals). *TMTC2*, *LRRIQ1* and *ALX1* mRNA levels were below background (low score and low signals) in the same brain region. There is no mouse homologue of *TSPAN19* in current genome database.

RefSeq gene	Accession #	Signal R	Signal S	Detection R	Detection S	Differential Score	Fold regulation
TMTC2	NM_025775.1	259	352	0.945	0.9735	**2.9**	1.36
SLC6A15	NM_175328.1	4660	2406	0.999	0.9988	**-102.2**	**-1.94**
TSPAN19	no mouse homolog	NA	NA	NA	NA	**NA**	NA
LRRIQ1	XM_137221.4	-122,3	-100,7	0,0772	0,1484	**0.6**	1.21
ALX1	NM_009423.2	87	113	0.854	0.8733	**0.7**	1.30

The reduction of *SLC6A15* expression could be validated by individual *in-situ* hybridizations (fig. 17a and b). Moreover, a significant reduction in *SLC6A15* expression was detected in the dentate gyrus of stress susceptible animals. Reduced *SLC6A15* mRNA expression was also measured in the visual cortex of stress susceptible mice with approaching significance (fig. 17c and d).

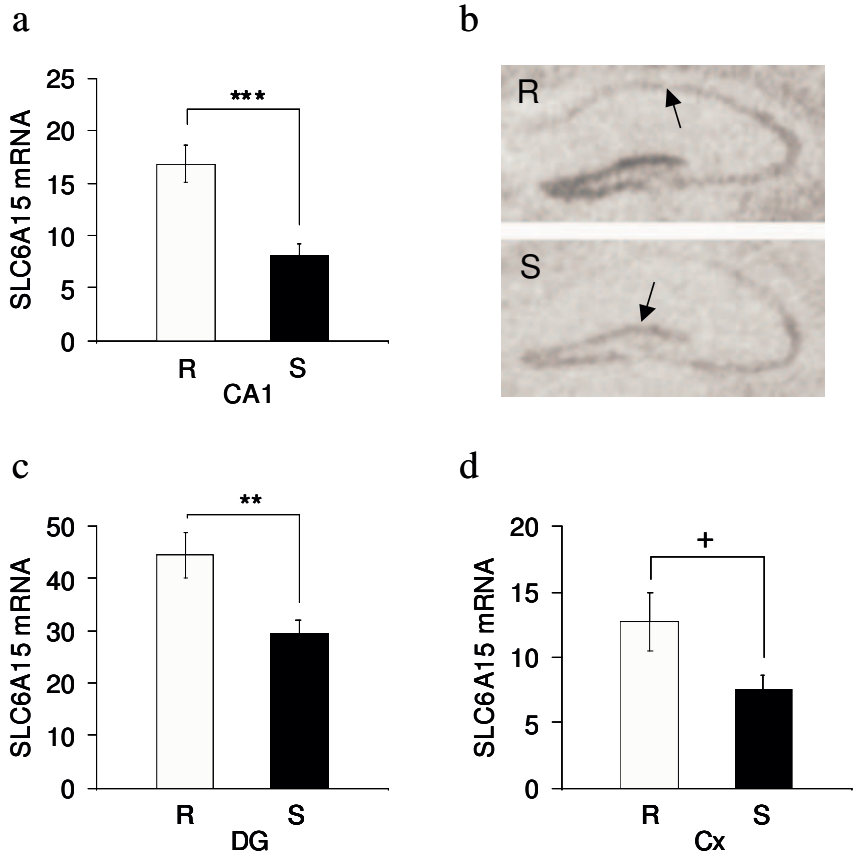


Fig. 17: *SLC6A15* mRNA levels in the hippocampus of stress resistant and susceptible mice.
(a) The significant reduction in *SLC6A15* mRNA levels in the CA1 hippocampal region between stress resistant (R) and susceptible (S) mice detected by microarray analysis could be confirmed by in-situ hybridization (N=9/9, -2.1-fold reduction). **(b)** Two representative radiograms of hippocampal slices from one animal per group are shown. The arrow in the upper picture points toward the CA1 region, the arrow in the lower picture towards the DG **(c,d)** *SLC6A15* mRNA was also significantly reduced in the dentate gyrus (DG, -1.5-fold) and by trend reduced in the visual cortex (Cx, -1.7-fold). $^{+}$ P<0.06; **P<0.01; ***P<0.001.

3.5 SLC6A15 in post-mortem brain studies of bipolar patients

A search of the NCBI's Gene Expression Omnibus (GEO) database (http:// www.ncbi.nlm.nih.gov/geo/) revealed no post-mortem brain expression profile studies for unipolar, but for bipolar depressed patients and controls. Genome-wide gene expression data were available of two different studies (252). *SLC6A15* and *ALX1* mRNA expression data were available, but none of the other genes adjacent to the region of association with MDD. *SLC6A15* (fig.18, N=30/31, p=0.041), but not *ALX1* mRNA expression was reduced in bipolar depressed patients compared to healthy controls in the dorsolateral prefrontal cortex. In the second dataset of different bipolar patients and controls *SLC6A15* mRNA expression levels were lowered in the orbitofrontal cortex (N=10/11, p>0.05).

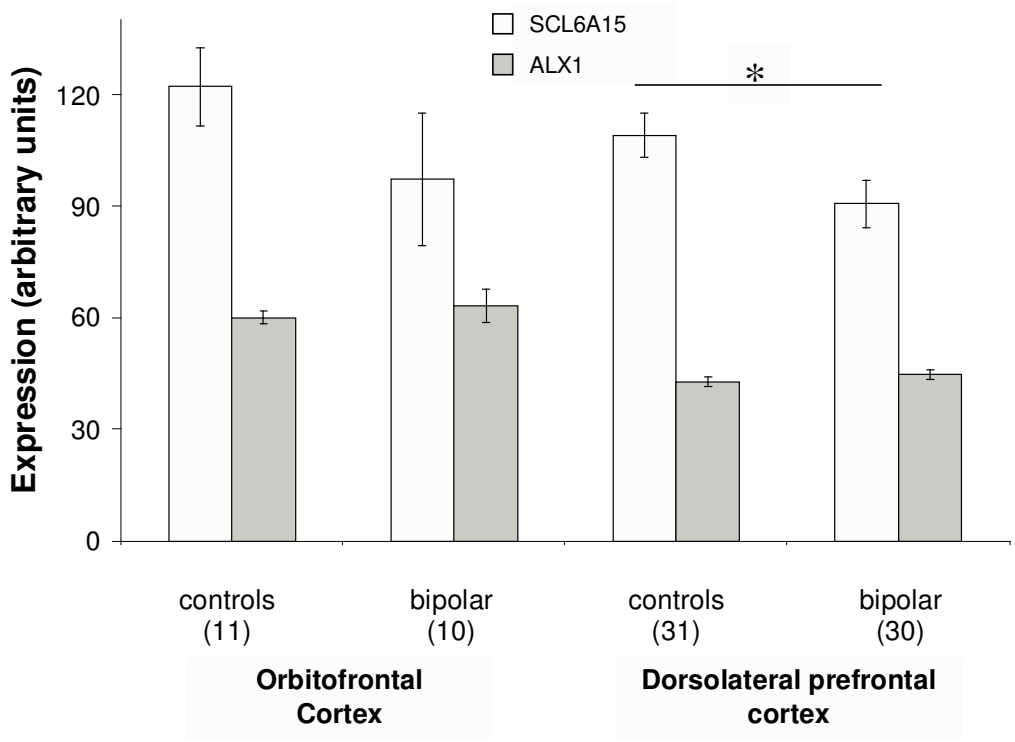


Fig. 18: Reduced post-mortem *SLC6A15* mRNA levels in the cortex of bipolar patients.
NCBI's Gene Expression Omnibus data (GEO series accession numbers: GDS2190 and GDS2191) from post-mortem genome-wide gene expression studies in the cortex of bipolar patients and controls show reduced *SLC6A15* mRNA expression in the dorsolateral prefrontal cortex of bipolar patients (*: N=30/31, p=0.041). This is further evidence for *SLC6A15* being involved in the pathophysiology of mood disorders.

4 DISCUSSION

4.1 A Candidate Gene Study Reveals Variants in *NTRK2* Conferring Risk to Mood Disorder Patients for a Life History of Attempted Suicide

4.1.1 Case-control association with MDD

Besides testing for association with SA, I also tested association with major depression with the complete marker panel for *BDNF* and *NTRK2* in the discovery sample. The analysis comprised 353 unipolar depressive patients and 366 non-psychiatric controls. Neither *BDNF*, nor *NTRK2* were associated over all tested markers and no single SNP association withstood correction for multiple testing. The presented study in MDD had statistical power ≥ 80% for the detection of effect sizes ≥ 1.6 OR under an log-additive genetic model, therefore lacking power for effect sizes of smaller magnitude as has often been reported for complex disease (1.1-1.4 OR) (32, 238).

However, the functional val66met BNDF polymorphism (rs6265) showed nominal significant association with MDD, but not with SA in the here presented study (suppl. tab. 1). Associations with this SNP and unipolar or bipolar disorder as well as gene-by-environment interactions have been reported by several groups, although conflicting results were obtained at least among the case-control studies (122, 123, 190-201) (for a more detailed discussion of these previously reported associations see chapter 1.2.8).

4.1.2 Phenotypic specificity of SNPs in NTRK2 associated with SA

Five tagging SNPs (rs10868235, rs1147198, rs1867283, rs1187286 and rs11140800) located in *NTRK2* were associated with SA within patients with mood disorders in two independent German samples. This association appears to be carried by several independent risk loci within this gene and carriers of the risk genotypes in the three strongest markers had a 4.5 fold higher risk for

suicide attempt than mood disorder patients carrying the non-risk genotypes. In contrast, no polymorphisms within *BDNF* were found to be associated with SA, although an highly complete SNP marker coverage of this locus was applied.
Overall, the presented genetic association study thus supports the large body of evidence implicating the neurotrophic system in the pathophysiology of suicidal behavior.
The fact that these associations were stronger in the comparison of patients with and without SA than in the comparison of suicide attempters with healthy controls and the fact that no significant case-control association with MDD was observed with any of these polymorphisms, suggest that these effects are likely specific to SA and not to the underlying psychiatric diagnosis. The discovery patient sample consisted of bipolar patients (11.6%), unipolar patients with first depressive episode (32.8%) and patients with recurrent unipolar depression (55.6%), whereas the replication patient sample comprised of recurrent unipolar depressed patients only (tab. 1 and 2). The associations with SA showed very similar effect sizes in these two diagnostically different samples and were unaffected by introducing type of mood disorder as covariate (data not shown), which further points toward a specificity of these genetic associations to SA.

4.1.3 MDD-related clinical parameters in SA

Moreover, this study replicates a previous study showing an earlier age at onset and a higher number of previous depressive episodes in suicide attempters compared to the rest of depressive patients (237). Co-varying for these variables did, however, not change the strength of the genetic association, suggesting that they did not confound the analysis. In conclusion, the presented results are in concordance with the hypothesis that psychiatric disease, social and other environmental risk factors are independent or interactive factors for genetic susceptibility specific to suicidal behavior (chapter 1.2.4-5). Furthermore, results from a linkage analysis in recurrent early-onset MDD families revealed several highly significant linkage peaks on a genome-wide level, when multipoint LOD scores without SA as covariate (base model) were subtracted by

the ones including SA as covariate (253), which supports genetic susceptibility specific for suicidal behavior.

4.1.4 Power considerartions and putative confounders

The two independent samples consist of a total of 265 mood disorder patients with SA, 873 mood disorder patients without a life history of SA and 1395 controls free of mood disorders and SA (tab. 1 and 2). The presented study had sufficient statistical power to detect the five described associations with SA among mood disorder patients in the combined sample with a power ranging from 68-84% depending on the minor allele frequency and effect size of the single markers. Nevertheless, more independent attempts of replication are needed to validate these results.

Population stratification is a concern in genetic association studies, since an unrecognized bias in the ancestral structure between the two compared groups, might lead to spurious association results with the phenotype under investigation. Therefore, the study design included recording of the ethnicity of parents, and all four grandparents of each proband. Affiliation to an ethnicity other than of European descent of one or more relatives led to the exclusion of a proband to statistical analysis in this study. The proportion of mood disorder patients with SA having all four grand-parents of German nationality over both studies was 89.6%, which was not different from the proportion in mood disorder patients without SA (90.2%, see also tab. 2). When the analysis for the five associated SNPs was repeated after exclusion of all probands with one or more grandparents with another nationality than German no substantial chance was detected in the strength of association with SA among mood disorder patients. Yet, published data suggest that there is almost no easily detectable population stratification within individuals with German ethnicity. A minor degree of population substructure from North to South of Germany has been reported, but it appears to be too low to be detectable with methods which do not include prior information of subpopulation membership (61). Both samples explored in this study were recruited in Bavaria, thus exclusively in Southern Germany.

Therefore, it is not likely that the here reported results are substantially biased by population stratification.

4.1.5 Multi-SNP-locus interaction analysis

Multi-locus interaction analysis in the combined patient sample revealed that the three most significant SNPs (rs10868235, rs1147198, rs1867283) contribute to the risk for SA in a multiplicative fashion. The model showed a p-value of $4.7x10^{-7}$ with a significant interaction between two of these SNPs (rs10868235, rs1147198). When patients carrying risk genotypes in all three markers were compared to those without any risk genotype in any marker the effect size was 4.5 OR (95% C.I. 2.1-9.8). Four of the five SNPs found to be associated with SA are independent from each other according to their LD structure in healthy controls (fig. 5D). Thus, suggesting that a combination of several independent risk alleles within the *NTRK2* locus is associated with SA in mood disorder patients. The presence of independent disease risk alleles in a single locus is well known from Mendelian diseases, e.g. the autosomal recessive inherited cystic fibrosis (CF), where currently more than 1500 mutations in the cystic fibrosis transmembrane conductance regulator (CFTR) gene have been related to CF (254). Most of these mutations are rare, but the most prominent one occurs in about two third of all affected individuals (255). Another example, but for complex disease, might be represented by two rare functional non-synonymous variants and one common regulatory variant in the tryptophan hydroxylase 2 (TPH2) gene which were reported to be associated with depressive disorder (256, 257).

The presented study achieved an almost complete tagging SNP coverage according to HapMap phase I and II data over both genes under investigation. This approach allowed to identify several independent common markers associated with SA among mood disorder patients in *NTRK2*. However, variants with a minor allele frequency smaller than 10% in the *NTRK2*- and 1% in the *BDNF* locus, respectively were not genotyped in this study due to a lack of statistical power. The reason behind the differently applied MAF threshold in the mapping approach of these two genes is that *NTRK2* is a large receptor gene

spanning over 360 kb in size and contains several regions of decayed LD (fig.5C), whereas the *BDNF* locus is about six times smaller and shows a continuous structure of relatively high LD (data not shown). Therefore, in order to cover the *NTRK2* region to the same extend as the *BDNF* locus many more tagging SNPs would have been required, which exceeded the scale of the study resources. Moreover, variants with a MAF smaller than 10% were most probably underpowered in the discovery sample of this study with respect to the rather small effect sizes generally expected in complex phenotypes (32, 238). As a consequence, the presence of additional even rarer risk loci for SA especially in the *NTRK2* locus cannot be excluded. These need to be investigated in larger trials, possibly by using re-sequencing approaches in order to discover also so far unkown rare variants.

4.1.6 Hypothesized functional relevance of the associated SNPs

As stated in the last paragraph, the performed multiplicative genetic model indicates at least four functionally independent risk loci for SA in *NTRK2*. Rs1147198 for example maps 8 kb 5' of the transcription start site of *NTRK2*. It is in high LD (HapMap, CEU) with SNPs located in the putative promoter region of *NTRK2* as predicted by the Genomatix® software (258), several of which hit a classical CpG-island (226, 259) of the *NTRK2* promoter. SNPs altering single CpG configurations have been associated with overall changes in the methylation level of the CpG island. For a CpG-dinucleotide abolishing SNP located in the promoter CpG-island of the *LRP1B* gene, for example, the presence of the C allele has been correlated with a hyper-methylation of all 16 adjacent CpGs in the CpG-island of the *LRP1B* promoter (260). Another study reported a similar finding for the val66met polymorphism (rs6265) located in the terminal exon of *BDNF*. The G allele constitutes a CpG-dinucleotide and its presence was not only correlated with CpG methylation at rs6265, but also with methylation of nearby CpGs (261). One functional consequences tagged by rs1147198 could thus be an altered methylation pattern of the *NTRK2* promoter's CpG-island.

The remaining four SNPs associated with SA in this study are all intronic and map to intron 14 or 16 of the full-length *NTRK2* gene, within a region where multiple alternative splicing events occurs leading to the truncated *NTRK2* isoforms (fig. 5A and 5B). The truncated *NTRK2* isoforms can act as dominant negative inhibitors of *NTRK2* tyrosine kinase activity if both isoforms are expressed in *cis* on the same cell (262, 263). The same isoforms can also function as scavenger receptors in *trans* by sequestering *BDNF* and removing it from the extracellular space via internalization (153, 264). While the associated SNPs are located in or in LD with functionally relevant regions of the *NTRK2* locus, further fine-mapping and re-sequencing efforts will be necessary to identify the the functional variants.

In summary, this study shows association of several independent common variants of *NTRK2* with attempted suicide in mood disorder patients in two independent German samples. This supports the large body of evidence that alterations in neurotrophin signaling might be involved in the pathophysiology of suicidal behavior.

4.2 A Genome-Wide Study Identifies SLC6A15 as a Novel Susceptibility Gene for MDD

SNP on chr12q21.31 show association with MDD in two independent German samples comprising over 2,600 subjects with an OR of 1.54 for the best fitting genetic model in the combined sample.

The observed OR in the GWAS sample was 2.8, while that of the replication study was only 1.3. This divergence most likely arises from different screening procedures of the control samples, as decreasing risk allele frequencies in parallel with an increasing stringency of exclusion criteria in the selection of the control samples were observed. On the other hand, risk allele frequencies did not differ between the two case samples (tab. 6, fig.10). The overall exclusion rate in the recruitment of replication controls was 20% and therefore well within

expectations when applying a screening procedure with mood disorders and anxiety as exclusion criteria. The GWAS controls, however, represent a subset of the population devoid of any current or past psychiatric disorder. The lower frequency of risk allele carriers in this group might therefore point towards a role of the associated SNPs reaching beyond MDD, such as an underlying vulnerability for the negative effects of chronic stress which represents a major risk factor for a variety of psychiatric disorders, including anxiety and depression. Furthermore DHWE was observed in the extensively screened controls for the associated polymorphisms, a fact that may point towards a truly associated region (53).

Thus, these results underline the importance of using carefully screened controls for association studies in highly prevalent disorders. In fact, the association p-value of $8.6x10^{-3}$ that would have been observed when comparing the discovery cases to the population-based controls would not have passed selection criteria for subsequent replication attempts within a GWAS. In summary, the applied study design is likely the reason for the observed relatively high OR of 2.8 in our GWAS sample unlike the reported results of GWA studies in bipolar disorder with ORs between 1.15 and 1.4 with less strictly or unscreened controls (238, 239, 265). This would also explain why smaller ORs were observed in our replication sample.

4.2.1 Functional relevance of the associated locus

The associated polymorphisms are located in a gene desert on chr12q21.31 with the next known gene annotated at a distance of about 300 kb. This gene codes for a sodium-dependent branched-chain amino acid transporter (*SLC6A15*) with the highest expression in the human brain (www. ncbi.nlm.nih.gov/UniGene). The solute carrier family 6 (SLC6) gene family also includes the monoamine and GABA transporters (266). According to *in situ* hybridizations, *SLC6A15* is most strongly expressed in the hippocampal region and expression is restricted to neurons (247, 248, 267). The CNS distribution pattern of this transporter may point to a role in neuronal metabolism and the provision of substrates for neurotransmitter synthesis (242). Knockout mice

lacking functional *SLC6A15* protein show no obvious gross phenotypic differences compared to wildtype animals and are similiarly viable and fertile. Experimental data from these transgenic animals indicate a moderate contribution of *SLC6A15* to total proline and leucine transport into cortical synaptosomes of about 15% (268). Proline, the amino acid with the highest affinity for *SLC6A15*, and leucine may act as precursors for glutamate synthesis (242) and this transporter could thus be involved in the regulation of glutamate transmission (269).

Recently, the prokaryotic leucine transporter homologue ($LeuT_{aa}$) of *SLC6A15* has been crystallized from Aquifex aeolicus and shown to bind tricyclic antidepressant drugs that can directly block leucine transport by closing the molecular gate for the substrate in a noncompetitive manner (270). Due to the high degree of conservation of the antidepressant binding sites, these drugs may also bind to the human transporter which would further support its role in the pathophysiology of major depression and make this molecule an attractive drug target.

The next gene proximal of the associated region, *TMTC2*, is expressed in the brain but has not yet been characterized in more detail. Neither of the two genes have so far been implicated in the causality of mood disorders, but regions on chromosome 12q22-q24 have shown evidence for linkage in three independent studies (86, 91, 271). However, these linkage peaks do not overlap with the region associated in this study.

The LD of the associated region on chr12q21.31 does not extend into regions of the adjacent genes associated with classical promoter sites or amino acid coding sequences. However, long-range regulatory effects are possible (272, 273). To address this issue, I took advantage of the publicly available gene expression data from the HapMap project, produced in lymphoblastoid cell lines. Only the expression of the full-length transcript of *SCL6A15* but not of other genes in the region were significantly associated with the risk variants after correction for multiple testing. In fact, risk allele carrier status was associated with less *SLC6A15* transcript. While this is a strong indication of the regulatory relevance of the region associated with MDD for *SLC6A15*

expression, there is still the possibility that these variants might also influence the expression of six unspliced ESTs detected in brain tissue which have been mapped to the region of association and were not probed by the used Illumina chip (fig.11A and chapter 3.2.8). Additional non-annotated transcripts, as described in the ENCODE pilot project in regions of the genome previously thought to be transcriptionally silent (274) might also be functionally relevant for this association. In the publicly available cortex mRNA expression data set of Myers *et al.* (246), all RefSeq annotated genes discussed in this study had missing data rates over 50%, so that informative analyses in this dataset were not possible.

Even though other functional effects of the associated variants cannot be excluded, their regulation of *SLC6A15* expression might be relevant for the pathophysiology of MDD, since in a well-established mouse model of chronic social stress (215, 216), *SLC6A15* but not *TMTC2*, *TSPAN19* (no mouse homolog), *LRRIQ1* and *ALX1* expression was downregulated in stress vulnerable compared to stress resistant mice. Moreover, post-mortem brain expression data in the NCBI's Gene Expression Omnibus (GEO) database (252) shows that mRNA expression of *SLC6A15* but not *ALX1* was reduced in the dorsolateral prefrontal cortex of patients suffering from bipolar disorder compared to healthy controls (fig.18), thus providing further independent evidence that *SLC6A15* might be involved in the pathophysiology of mood disorders.

4.2.2 MRI and ^{1}H-NMR studies reveal association of risk variants with measures of hippocampal integrity

Additional evidence for a functional relevance of the detected genetic variants in MDD is provided by the imaging genomics results presented in this dissertation. Healthy risk allele carriers exhibited lower hippocampal NAA compared to non-risk allele carriers. NAA is a sensitive marker of neuronal viability with additional implications in cell signalling. It has been found decreased in first episode affective psychosis (275), bipolar disorder (276), post-traumatic stress disorder (277) and was also decreased in currently depressed unipolar patients in this

study. In animal models, hippocampal NAA can be decreased by chronic stress paradigms (278, 279). Thus, a genetic predisposition towards lower hippocampal NAA, similar to a condition induced by chronic stress experiments, may impair an individual's resilience to stress.

In addition, lower total left hippocampal volumes, particularly of the cornu ammonis, in rs1031681 risk allele carriers of the patient- but not the control group was found in the here presented study. Hippocampal morphology is subject to stronger environmental influences compared with other brain regions (280), and interactions between recurrent depression and specific genetic predispositions as indicated by the here presented results may advance hippocampal atrophy (281). Volumetric effects were strongest in the cornu ammonis (CA) subregion, which is in accordance with results from the presented mouse model, where down-regulation of *SLC6A15* expression in response to stress was most prominent in the CA1 region of the hippocampus. Since altered expression of *SLC6A15* in CA1 was associated with susceptibility to long term negative effects of chronic social stress, these results support a role for hippocampal *SLC6A15* function in stress sensitivity and thus the pathophysiology of MDD.

The observed genotype-dependent differences of glutamate/glutamine in unaffected control subjects also fit well with a proposed role of *SLC6A15* as transporter of neurotransmitter precursors in the glutamate system and the notion that NAA may function as reservoir for glutamate (282). Decreased hippocampal NAA/NAAG and glutamate/glutamine, related to genetic factors, may limit excitatory signalling capacity with secondary effects on stress response regulation and other hippocampal functions (283). For all presented neuroimaging results risk allele carrier status was associated with patterns of hippocampal metabolism and morphology previously observed in mood disorders, providing additional support for biological validity of the newly discovered genetic susceptiblity markers.

4.2.3 *SLC6A15* - a novel candidate gene for MDD

Epidemiological studies on MDD report a 2-3 fold risk increase for individuals exposed to chronic stress (214) and twin studies clearly point to an increased susceptibility for MDD as a result of a combination of environmental and genetic risk factors (73, 74). The above presented human genetics, expression, spectroscopy, imaging and mouse model results all support the notion that lower *SLC6A15* expression, especially in the hippocampus, could increase an individual's stress susceptibility by altering excitatory neurotransmission in this brain region. This hypothesis is supported by the fact that *SLC6A15* homozygous knockout mice show a trend towards increased acute stress-induced anxiety-related phenotypes without baseline differences in the same behavioral paradigms when compared to wildtype animals (268). Individuals carrying the identified risk variants might be more vulnerable to the stress-induced downregulation of this gene thus passing the threshold for dysfunctions in disease-relevant brain circuits more easily.

All data from clinical and preclinical studies presented here converge to support *SLC6A15* as a susceptibility gene for MDD. Because this molecule appears amenable to drug targeting, this results may incite the development of a novel class of antidepressant drugs.

4.3 Overall conclusion

This dissertation presents results from both, a candidate gene- as well as a genome-wide association study and shows that both approaches can be important contributors in the understanding of the biology of mood disorders.

The hypothesis-driven genetic association study in major depression and attempted suicide represents an example of a high-density SNP marker association screening in two neurotrophic candidate genes with a subsequent replication of several independent SNPs associated with lifetime attempted suicide among depressed patients. It shows the potential of hypothesis-driven association analysis to further validate a well-established candidate pathway, whose importance for suicidal behavior has previously been defined by findings

in post-mortem and clinical studies as well as functional experiments in animals. High-throughput genotyping technology together with sufficient knowledge about the complexity of the human genome with regards to genetic variants as was provided by the International HapMap Project enabled large scale hypothesis-driven association analyses in the absence of known functional variants and became an important tool to unravel genetic susceptibility for complex human disease. The example highlighted in this dissertation suggests that multiple independent genetic variants in a disease-relevant gene contribute to the pathophysiology of a disorder as previously described for monogenetic diseases. Mapping with almost complete coverage of the genomic information content as used here, is crucial for the identification of most common contributing alleles. This study shows that multi-variant models in one candidate gene lead to a better differentiation between probands at high- and low-risk for disease than single SNP associations. Therefore, multi-locus analysis seems to be promising in order to better understand complex genetic traits. In the described analysis interaction among only the three most associated variants within one locus were explored, but not interactions among different loci due to a lack of statistical power. For polygenic traits gene x gene interactions are likely the rule rather than the exception. However, up to now these interactions have rarely been examined due to large sample requirements. With case and control samples now often in the 10,000 provided by multi-center consortiums, these analyses should begin to be sufficiently powered.

However, there are also limitiations of genetic association compared to more functional approaches beside its simple correlative nature. Missing evidence from genetic association does not mean that a candidate gene is not substantially involved in the etiology of a disease by reason that the loci might be devoid of functional variants with relevant impact, because of its high importance for the homeostasis of vital processes. Moreover, genetic association findings in one human population can not be generalized to other populations, since the underlying genetic variants can be population specific. Even if different populations share an important functional disease causing variant and genetic association with disease could be established in one

population with markers common in all populations, the same markers might fail to show association in other populations due to a population-specific underlying linkage disequilibrium structure in the region of association. In this respect, the here described associations await validation in other than European populations.

On the other hand, this work also highlights the discovery of a novel candidate gene by a hypothesis-free genome-wide association approach with a subsequent validation of the found candidate gene by diverse functional studies. The conclusions drawn in this study are based on converging results from a experiment-wide significant case-control association result with subsequent positive replication in a larger independent sample, gene expression profiling, genomic imaging and a well-validated animal model mimicking the human disease state. Converging results of diverse and complementary experimental and technical approaches might highlight future discoveries in genetics of complex traits. In genetics of mood disorders where most genetic associations have been difficult to reproduce converging data from different approaches together with consideration of gene-by-environment interactions might be promising for the more reproducible establishment of novel candidate genes.

Another important conclusion from the here reported genome-wide association study results might be especially relevant for the genetics of common psychiatric disease. So far, researchers did not obtained genome-wide significance in case-control studies in MDD nor in bipolar disorder, even though the later studies have already reached much larger sample sizes and bipolar disorder has a much higher estimated heritability than MDD (BP:80-90% versus MDD: 33-42%) (284). In the GWA studies in bipolar disorder (238, 239, 265), controls were not extensively screened for the absence of any axis I psychiatric disorder as practiced in the GWAS described in this dissertation. Thus, the number of misclassified controls was most certainly larger in those studies. Moreover, lifetime prevalence of bipolar disorder is estimated to be about 1.0-2.4% (285), whereas for MDD estimates are much higher at 16.2% (95% C.I. 15.1-17.3%) (65) indicating a much higher necessity of thoroughly screened

controls. In addition, susceptibility genes for psychiatric disorders may be genes conferring susceptibility to environmental stressor in general and the final phenotypic presentation, such as MDD or anxiety disorders for example, could be dependent on other environmental or genetic factors, thus underlining the importance of equally complete phenotypic characterization of cases and controls.

5 ABBREVIATIONS

AMD	Age-related Macula Degeneration
BDI	Beck Depression Inventory
BDNF	Brain-Derived Neurotrophic Factor
BP	Bipolar Disorder
CA	Cornu ammonis
CAGT	Center of Applied GenoTyping
CEU	Utah residents from northern and western European ancestry from the Centre d'Etude du Polymorphisme Humain (CEPH) collection
CHB	Han Chinese from Beijing, China
CI	Confidence Interval
CIDI	Composite International Diagnostic Interview
CID-S	Composite International Diagnostic Screener
CSF	CerebroSpinal Fluid
DG	Dentate Gyrus
DHW	Deviation from Hardy-Weinberg Equlibrium
DSM IV	Diagnostic and Statistical Manual of Mental Disorders IV
DST	Dexamethasone Suppression Test
DZ	DiZygotic
EBV	Epstein-Barr Virus
EST	Expressed Sequence Tag
FPM	Fisher Product Method
GENEVAR	GENe Expression VARiation
GEO	Gene Expression Omnibus
Glx	Glutamate/Glutamine
GM	Grey Matter
GRR	Genotypic Relative Risk
GWAS	Genome-Wide Association Study
HapMap	International HapMap Project

HIP	Hippocampus
HPA	Hypothalamic-Pituitary-Adrenal axis
HWE	Hardy Weinberg Equilibrium
JPT	Japanese from Tokyo, Japan
Kb	Kilobase
LD	Linkage Disequilibrium
MAF	Minor Allele Frequency
MALDI-TOF	Matrix Assisted Laser Desorption/Ionisation – Time Of Flight
MARS	Munich Antidepressant Response Signature
Mb	Megabase
MDD	Major Depressive Disorder
MPIP	Max Planck Institute of Psychiatry
MRI	Magnetic Resonance Imaging
MZ	MonoZygotic
^{1}H-NMR	Nuclear Magnetic Resonance Spectroscopy
NAA	N-Acetyl-Aspartate
NCBI	National Center for Biotechnology Information
NTRK2	Neurotrophine Receptor Tyrosine Kinase 2
OR	Odds Ratio
OSA	Ordered Sub-set Analysis
PFC	PreFrontal Cortex
PopGen	Population Genetic Cohort
RAF	Risk Allele Frequency
SA	Suicide Attempt
SD	Standard Deviation
SLC6A15	Solute Carrier 6 family member gene A15
SNP	Single Nucleotide Polymorphism
TagSNP	Tagging SNP
WM	White matter
WY	Permutation-based minimum p method for multiple comparisons proposed by Westfall and Young
YRI	Yoruba in Ibadan, Nigeria

6 DIRECT CONTRIBUTORS

Martin A. Kohli (MAK), M.Sc, the author of this dissertation designed and conducted all aspects of the presented studies other than patient and control recruitment under supervision of Elisabeth B. Binder (EBB), Ph.D., MD.

Susanne Lucae, Ph.D., MD., was responsible for ascertainment, study design and management of the GWAS. MAK performed the genetic statistics of the GWAS and the genotyping of the replication samples, whereas whole-genome genotyping was carried out at the Center for Applied GenoTyping Munich, CAGT-Munich, a facility of the Max Planck Institute of Psychiatry (MPIP) in Munich, run by Thomas Bettecken, Ph.D. MAK performed all the bioinformatic work to interpret the results and the human expression analysis of the GENEVAR data. The GWAS was further supervised by Prof. Bertram Müller-Myhsok, MD, and EBB.

Mathias Schmidt, Ph.D., and Marianne Müller, MD, provided the microarray data of the mouse model and validated the findings described in this work.

Philipp Sämann, MD, David Höhn, MD and Michael Czisch, MD, conducted the magnet resonance imaging study. MAK ran all genetic association tests with the imaging and spectroscopy data.

7 ACKNOWLEDGEMENTS

I would like to thank my supervisor Elisabeth B. Binder for supporting me with her profound knowledge of psychiatric genetics and inspiring me with her great drive and fascination for realizing scientific goals.

My special thanks to Daria Salyakina for her daily help in scientific work, especially regarding theoretical and practical aspects of genetic statistics.

My thanks to Susanne Lucae for genetic and clinical study management, her support in scientific writing and psychiatric phenotyping and many other advice.

Many thanks also to Stephan Ripke for providing me with the Whole-Genome Permer software and Prof. Bertram Müller-Myhsok for his deep understanding in statistical genetics. Thanks to Benno Pütz, Darina Röske, Michael Specht, Hildegard Pfister and Marcus Ising for your support in statistics and data processing.

This project would not have been possible without the work of many clinicians, psychologists and researchers who were responsible for study ascertainment. Many thanks to them all: Susanne Lucae, Stephan Kloiber, Andreas Menke, Sonja Horstmann, Angelika Erhardt, Paul Gerson Unschuld, Derek Spieler, Angela Heck, Tanja Brückl, Petra Zimmermann, Hildegard Pfister and Markus Ising.

I like to thank the laboratory heads Manfred Uhr and Thomas Bettecken for their support and the technician team for their great service. Tons of thanks to Alina Tontsch, Susann Sauer, Sabine Damast, Maik Ködel and Angelika Sangl.

Thanks to the direct contributors for their contribution to the functional validation of the results of the GWAS in MDD, Mathias Schmidt and Marianne Müller who provided me with their mouse model data, and Philipp Sämann, David Höhn and Michael Czisch who performed the MRI study.

At last, but not least, special thanks to Prof. Florian Holsboer for my position at the Max Planck Institute of Psychiatry (MPIP) and for providing an outstanding multi disciplinary infrastructure at the MPIP.

8 REFERENCES

1. Falconer, DS Inheritance Of Liability To Certain Diseases Estimated From Incidence Among Relatives. Annals Of Human Genetics, 1965; *29*(1)*:* 51.
2. dbSNP build 127, NCBI. (http://www.ncbi.nlm.nih.gov/projects/SNP/snp_summary.cgi), 2007.
3. Collins, FS, Lander, ES, Rogers, J, and Waterston, RH Finishing the euchromatic sequence of the human genome. Nature, 2004; *431*(7011)*:* 931-945.
4. Ott, J Analysis of Human Genetic Linkage. The Johns Hopkins University Press. Thrid edition 1999, p. 1 and 15.
5. Risch, N and Merikangas, K The future of genetic studies of complex human diseases. Science, 1996; *273*(5281)*:* 1516-1517.
6. Risch, NJ Searching for genetic determinants in the new millennium. Nature, 2000; *405*(6788)*:* 847-856.
7. Macgregor, S, Craddock, N, and Holmans, PA Use of phenotypic covariates in association analysis by sequential addition of cases. European Journal Of Human Genetics, 2006; *14*(5)*:* 529-534.
8. Zondervan, KT and Cardon, LR The complex interplay among factors that influence allelic association. Nature Reviews Genetics, 2004; *5*(2)*:* 89-U14.
9. Dudbridge, F and Gusnanto, A Estimation of significance thresholds for genomewide association scans. Genetic Epidemiology, 2008; *32*(3)*:* 227-234.
10. Pe'er, I, Yelensk, R, Altshuler, D, and Daly, MJ Estimation of the multiple testing burden for genomewide association studies of nearly all common variants. Genetic Epidemiology, 2008; *32*(4)*:* 381-385.
11. Wall, JD and Pritchard, JK Haplotype blocks and linkage disequilibrium in the human genome. Nature Reviews Genetics, 2003; *4*(8)*:* 587-597.
12. Salyakina, D Candidate gene association testing in the dissection of genetic causes for depressive disorders and the response to antidepressant treatment. Technical University Munich. Dissertation, 2007; *http://mediatum2.ub.tum.de/doc/614023/614023.pdf* (
13. Pritchard, JK and Przeworski, M Linkage disequilibrium in humans: Models and data. American Journal Of Human Genetics, 2001; *69*(1)*:* 1-14.
14. Patil, N, Berno, AJ, Hinds, DA*, et al.* Blocks of limited haplotype diversity revealed by high-resolution scanning of human chromosome 21. Science, 2001; *294*(5547)*:* 1719-1723.
15. Wang, N, Akey, JM, Zhang, K, Chakraborty, R, and Jin, L Distribution of recombination crossovers and the origin of haplotype blocks: The interplay of population history, recombination, and mutation. American Journal Of Human Genetics, 2002; *71*(5)*:* 1227-1234.

16. Weiss, KM and Clark, AG Linkage disequilibrium and the mapping of complex human traits. Trends In Genetics, 2002; *18*(1)*:* 19-24.
17. Carlson, CS, Eberle, MA, Rieder, MJ, Smith, JD, Kruglyak, L, and Nickerson, DA Additional SNPs and linkage-disequilibrium analyses are necessary for whole-genome association studies in humans. Nature Genetics, 2003; *33*(4)*:* 518-521.
18. Stephens, JC, Schneider, JA, Tanguay, DA*, et al.* Haplotype variation and linkage disequilibrium in 313 human genes. Science, 2001; *293*(5529)*:* 489-493.
19. Service, SK, Ophoff, RA, and Freimer, NB The genome-wide distribution of background linkage disequilibrium in a population isolate. Human Molecular Genetics, 2001; *10*(5)*:* 545-551.
20. Abecasis, GR, Noguchi, E, Heinzmann, A*, et al.* Extent and distribution of linkage disequilibrium in three genomic regions. American Journal Of Human Genetics, 2001; *68*(1)*:* 191-197.
21. Gabriel, SB, Schaffner, SF, Nguyen, H*, et al.* The structure of haplotype blocks in the human genome. Science, 2002; *296*(5576)*:* 2225-2229.
22. Cann, RL, Brown, WM, and Wilson, AC Polymorphic Sites And The Mechanism Of Evolution In Human Mitochondrial-Dna. Genetics, 1984; *106*(3)*:* 479-499.
23. Stringer, CB and Andrews, P Genetic And Fossil Evidence For The Origin Of Modern Humans. Science, 1988; *239*(4845)*:* 1263-1268.
24. Frazer, KA, Ballinger, DG, Cox, DR*, et al.* A second generation human haplotype map of over 3.1 million SNPs. Nature, 2007; *449*(7164)*:* 851-U853.
25. Goldstein, DB, Ahmadi, KR, Weale, ME, and Wood, NW Genome scans and candidate gene approaches in the study of common diseases and variable drug responses. Trends In Genetics, 2003; *19*(11)*:* 615-622.
26. Carlson, CS, Eberle, MA, Rieder, MJ, Yi, Q, Kruglyak, L, and Nickerson, DA Selecting a maximally informative set of single-nucleotide polymorphisms for association analyses using linkage disequilibrium. American Journal Of Human Genetics, 2004; *74*(1)*:* 106-120.
27. de Bakker, PIW, Yelensky, R, Pe'er, I, Gabriel, SB, Daly, MJ, and Altshuler, D Efficiency and power in genetic association studies. Nature Genetics, 2005; *37*(11)*:* 1217-1223.
28. Schaid, DJ and Sommer, SS Genotype Relative Risks - Methods For Design And Analysis Of Candidate-Gene Association Studies. American Journal Of Human Genetics, 1993; *53*(5)*:* 1114-1126.
29. Bland, JM and Altman, DG Statistics notes - The odds ratio. British Medical Journal, 2000; *320*(7247)*:* 1468-1468.
30. Risch, N and Teng, J The relative power of family-based and case-control designs for linkage disequilibrium studies of complex human diseases - I. DNA pooling. Genome Research, 1998; *8*(12)*:* 1273-1288.
31. Bourgain, C, Genin, E, Cox, N, and Clerget-Darpoux, F Are genome-wide association studies all that we need to dissect the genetic component of complex human diseases? European Journal Of Human Genetics, 2007; *15*(3)*:* 260-263.

32. Lohmueller, KE, Pearce, CL, Pike, M, Lander, ES, and Hirschhorn, JN Meta-analysis of genetic association studies supports a contribution of common variants to susceptibility to common disease. Nature Genetics, 2003; *33*(2)*:* 177-182.
33. Thakkinstian, A, Han, P, McEvoy, M*, et al.* Systematic review and meta-analysis of the association between complementary factor HY402H polymorphisms and age-related macular degeneration. Human Molecular Genetics, 2006; *15*(18)*:* 2784-2790.
34. Westfall, PH and Young, SS Resampling-Based Multiple Testing. John Wiley & Sons, Inc., New York., 1993.
35. Westfall, PH, Zaykin, DV, and Young, SS Multiple tests for genetics effects in association studies: biostatistical methods. Humana Press, Totawa, NJ., 2001.
36. Goncalo, R, Abecasis, GR, Cookson, WOC, and Cardon, LR The power to detect linkage disequilibrium with quantitative traits in selected samples. American Journal Of Human Genetics, 2001; *68*(6)*:* 1463-1474.
37. Müller-Myhsok, B and Abel, L Genetic analysis of complex diseases. Science, 1997; *275*(5304)*:* 1328-1329.
38. Tu, IP and Whittemore, AS Power of association and linkage tests when the disease alleles are unobserved. American Journal Of Human Genetics, 1999; *64*(2)*:* 641-649.
39. Concannon, P, Onengut-Gumuscu, S, Todd, JA*, et al.* A human type 1 diabetes susceptibility locus maps to chromosome 21q22.3. Diabetes, 2008; *57*(10)*:* 2858-2861.
40. Gordon, D, Matise, TC, Heath, SC, and Ott, J Power loss for multiallelic transmission/disequilibrium test when errors introduced: GAW11 simulated data. Genetic Epidemiology, 1999; *17*(587-592.
41. Gordon, DL, SM. Heath, SC. Ott, J. An analytic solution to single nucleotide polymorphism error-detection rates in nuclear families: implications for study design. Pacific Symposium on Biocomputing, 2000; 663-674.
42. Gordon, DO, J. Assessment and management of single nucleotide polymorphism genotype errors in genetic association analysis. Pacific Symposium on Biocomputing, 2001; 18-29.
43. Gordon, D, Finch, SJ, Nothnagel, M, and Ott, J Power and sample size calculations for case-control genetic association tests when errors are present: Application to single nucleotide polymorphisms. Human Heredity, 2002; *54*(1)*:* 22-33.
44. Clayton, DG, Walker, NM, Smyth, DJ*, et al.* Population structure, differential bias and genomic control in a large-scale, case-control association study. Nature Genetics, 2005; *37*(11)*:* 1243-1246.
45. Akey, JM, Zhang, K, Xiong, MM, Doris, P, and Jin, L The effect that genotyping errors have on the robustness of common linkage-disequilibrium measures. American Journal Of Human Genetics, 2001; *68*(6)*:* 1447-1456.
46. Hardy, GH Mendelian proportions in a mixed population. Science, 1908; *28*(1)*:* 49-50.

47. Cockerha, CC Group Inbreeding And Coancestry. Genetics, 1967; *56*(1)*:* 89.
48. Park, YC Variance Of Gene-Frequencies Due To Random Drift In A Selection Experiment. Genetics, 1975; *80*(3)*:* S62-S63.
49. Deng, HW, Chen, WM, and Recker, RR Population admixture: Detection by Hardy-Weinberg test and its quantitative effects on linkage-disequilibrium methods for localizing genes underlying complex traits. Genetics, 2001; *157*(2)*:* 885-897.
50. Cockerha, CC Variance Of Gene Frequencies. Evolution, 1969; *23*(1)*:* 72.
51. Hosking, L, Lumsden, S, Lewis, K*, et al.* Detection of genotyping errors by Hardy-Weinberg equilibrium testing. European Journal Of Human Genetics, 2004; *12*(5)*:* 395-399.
52. Leal, SM Detection of genotyping errors and pseudo-SNPs via deviations from Hardy-Weinberg equilibrium. Genetic Epidemiology, 2005; *29*(3)*:* 204-214.
53. Wittke-Thompson, JK, Pluzhnikov, A, and Cox, NJ Rational inferences about departures from Hardy-Weinberg equilibrium. American Journal Of Human Genetics, 2005; *76*(6)*:* 967-986.
54. Lucae, S, Salyakina, D, Barden, N*, et al.* P2RX7, a gene coding for a purinergic ligand-gated ion channel, is associated with major depressive disorder. Human Molecular Genetics, 2006; *15*(16)*:* 2438-2445.
55. Barden, N, Harvey, M, Gagne, B*, et al.* Analysis of single nucleotide polymorphisms in genes in the chromosome 12Q24.31 region points to P2RX7 as a susceptibility gene to bipolar affective disorder. American Journal Of Medical Genetics Part B-Neuropsychiatric Genetics, 2006; *141B*(4)*:* 374-382.
56. Rahman, OA, Nagy, G, Mate, A, Somogyi, A, Sasvari-Szekely, M, and Nemoda, Z Association of P2RX7 polymorphism with depression in diabetic patients with poor glycemic control. European Neuropsychopharmacology, 2008; *18*(63-63.
57. Devlin, B, Bacanu, SA, and Roeder, K Genomic control to the extreme. Nature Genetics, 2004; *36*(11)*:* 1129-1130.
58. Bacanu, SA, Devlin, B, and Roeder, K The power of genomic control. American Journal Of Human Genetics, 2000; *66*(6)*:* 1933-1944.
59. Freedman, ML, Reich, D, Penney, KL*, et al.* Assessing the impact of population stratification on genetic association studies. Nature Genetics, 2004; *36*(4)*:* 388-393.
60. Price, AL, Patterson, NJ, Plenge, RM, Weinblatt, ME, Shadick, NA, and Reich, D Principal components analysis corrects for stratification in genome-wide association studies. Nature Genetics, 2006; *38*(8)*:* 904-909.
61. Steffens, M, Lamina, C, Illig, T*, et al.* SNP-based analysis of genetic substructure in the German population. Human Heredity, 2006; *62*(1)*:* 20-29.
62. Purcell, S, Cherny, SS, and Sham, PC Genetic Power Calculator: design of linkage and association genetic mapping studies of complex traits. Bioinformatics, 2003; *19*(1)*:* 149-150.

63. Gauderman, WJ Sample size requirements for matched case-control studies of gene-environment interaction. Statistics In Medicine, 2002; *21*(1)*:* 35-50.
64. Gauderman, WJ Sample size requirements for association studies of gene-gene interaction. American Journal Of Epidemiology, 2002; *155*(5)*:* 478-484.
65. Kessler, RC, Berglund, P, Demler, O*, et al.* The epidemiology of major depressive disorder - Results from the National Comorbidity Survey Replication (NCS-R). Jama-Journal Of The American Medical Association, 2003; *289*(23)*:* 3095-3105.
66. Smoller, JW, Sheidley, BS, and Tsuang, MT Psychiatric genetics: applications in clinical practice. American Psychiatric Publishing, Inc. Washington, DC, London, England, 2008; *First Edition*(131-176.
67. Uhl, GR and Grow, RW The burden of complex genetics in brain disorders. Archives Of General Psychiatry, 2004; *61*(3)*:* 223-229.
68. Levinson, DF The genetics of depression: A review. Biological Psychiatry, 2006; *60*(2)*:* 84-92.
69. Mann, JJ Neurobiology of suicidal behaviour. Nature Reviews Neuroscience, 2003; *4*(10)*:* 819-828.
70. Mann, JJ The neurobiology of suicide. Nature Medicine, 1998; *4*(1)*:* 25-30.
71. Mann, JJ and Currier, D A Review of Prospective Studies of Biologic Predictors of Suicidal Behavior in Mood Disorders. Archives of Suicide Research, 2007; *11*(1)*:* 3 - 16.
72. Diagnostic and Statistical Manual of Mental Disorders. 4th Edition, Text Revision. Washington, DC, American Psychiatric Association, 2000.
73. Kendler, KS, Gardner, CO, and Prescott, CA Toward a comprehensive developmental model for major depression in men. American Journal Of Psychiatry, 2006; *163*(1)*:* 115-124.
74. Kendler, KS, Gardner, CO, and Prescott, CA Toward a comprehensive developmental model for major depression in women. American Journal Of Psychiatry, 2002; *159*(7)*:* 1133-1145.
75. Ustun, TB, Ayuso-Mateos, JL, Chatterji, S, Mathers, C, and Murray, CJL Global burden of depressive disorders in the year 2000. British Journal Of Psychiatry, 2004; *184*(386-392.
76. Greenberg, PE, Kessler, RC, Birnbaum, HG*, et al.* The economic burden of depression in the United States: How did it change between 1990 and 2000? Journal Of Clinical Psychiatry, 2003; *64*(12)*:* 1465-1475.
77. Ebmeier, KP, Donaghey, C, and Steele, JD Recent developments and current controversies in depression. Lancet, 2006; *367*(9505)*:* 153-167.
78. Smoller, JW and Perlis, RH Family and genetic studies of depression, in Handbook of Chronic Depression., 2004; 97-138.
79. Bland, RC, Newman, SC, and Orn, H Recurrent And Nonrecurrent Depression - A Family Study. Archives Of General Psychiatry, 1986; *43*(11)*:* 1085-1089.
80. Farmer, A, Harris, T, Redman, K, Sadler, S, Mahmood, A, and McGuffin, P Cardiff Depression Study - A sib-pair study of life events and familiality in major depression. British Journal Of Psychiatry, 2000; *176*(150-155.

81. Gershon, ES, Hamovit, J, Guroff, JJ*, et al.* A Family Study Of Schizoaffective, Bipolar-I, Bipolar-Ii, Unipolar, And Normal Control Probands. Archives Of General Psychiatry, 1982; *39*(10)*:* 1157-1167.
82. Kutcher, S and Marton, P Affective-Disorders In 1st-Degree Relatives Of Adolescent Onset Bipolars, Unipolars, And Normal Controls. Journal Of The American Academy Of Child And Adolescent Psychiatry, 1991; *30*(1)*:* 75-78.
83. Warner, V, Mufson, L, and Weissman, MM Offspring At High And Low-Risk For Depression And Anxiety - Mechanisms Of Psychiatric-Disorder. Journal Of The American Academy Of Child And Adolescent Psychiatry, 1995; *34*(6)*:* 786-797.
84. Sullivan, PF, Neale, MC, and Kendler, KS Genetic epidemiology of major depression: Review and meta-analysis. American Journal Of Psychiatry, 2000; *157*(10)*:* 1552-1562.
85. Kendler, KS, Gatz, M, Gardner, CO, and Pedersen, NL A Swedish national twin study of lifetime major depression. American Journal Of Psychiatry, 2006; *163*(1)*:* 109-114.
86. McGuffin, P, Knight, J, Breen, G*, et al.* Whole genome linkage scan of recurrent depressive disorder from the depression network study. Human Molecular Genetics, 2005; *14*(22)*:* 3337-3345.
87. Zubenko, GS, Hughes, HB, Maher, BS, Stiffler, JS, Zubenko, WN, and Marazita, ML Genetic linkage of region containing the CREB1 gene to depressive disorders in women from families with recurrent, early-onset, major depression. American Journal Of Medical Genetics, 2002; *114*(8)*:* 980-987.
88. Zubenko, GS, Hughes, HB, Stiffler, JS, Zubenko, WN, and Kaplan, BB Genome survey for susceptibility loci for recurrent, early-onset major depression: Results at 10cM resolution. American Journal Of Medical Genetics, 2002; *114*(4)*:* 413-422.
89. Zubenko, GS, Hughes, HB, Stiffler, JS, Zubenko, WN, and Kaplan, BB D2S2944 identifies a likely susceptibility locus for recurrent, early-onset, major depression in women. Molecular Psychiatry, 2002; *7*(5)*:* 460-467.
90. Zubenko, GS, Maher, B, Hughes, HB*, et al.* Genome-wide linkage survey for genetic loci that influence the development of depressive disorders in families with recurrent, early-onset, major depression. American Journal Of Medical Genetics Part B-Neuropsychiatric Genetics, 2003; *123B*(1)*:* 1-18.
91. Abkevich, V, Camp, NJ, Hensel, CH*, et al.* Predisposition locus for major depression at chromosome 12q22-12q23.2. American Journal Of Human Genetics, 2003; *73*(6)*:* 1271-1281.
92. Holmans, P, Zubenko, GS, Crowe, RR*, et al.* Genomewide significant linkage to recurrent, early-onset major depressive disorder on chromosome 15q. American Journal Of Human Genetics, 2004; *74*(6)*:* 1154-1167.
93. Holmans, P, Weissman, MM, Zubenko, GS*, et al.* Genetics of recurrent early-onset major depression (GenRED): Final genome scan report. American Journal Of Psychiatry, 2007; *164*(2)*:* 248-258.

94. Levinson, DF, Evgrafov, OV, Knowles, JA*, et al.* Genetics of recurrent early-onset major depression (GenRED): Significant linkage on chromosome 15q25-q26 after fine mapping with single nucleotide polymorphism markers. American Journal Of Psychiatry, 2007; *164*(2)*:* 259-264.
95. Almasy, L, Gur, RC, Haack, K*, et al.* A genome screen for quantitative trait loci influencing schizophrenia and neurocognitive phenotypes. American Journal Of Psychiatry, 2008; *165*(9)*:* 1185-1192.
96. Irmansyah, Schwab, SG, Heriani*, et al.* Genome-wide scan in 124 Indonesian sib-pair families with schizophrenia reveals genome-wide significant linkage to a locus on chromosome 3p26-21. American Journal of Medical Genetics Part B: Neuropsychiatric Genetics, 2008; *147B*(7)*:* 1245-1252.
97. Bulayeva, KB, Glatt, SJ, Bulayev, OA, Pavlova, TA, and Tsuang, MT Genome-wide linkage scan of schizophrenia: A cross-isolate study. Genomics, 2007; *89*(2)*:* 167-177.
98. Venken, T, Alaerts, M, Souery, D*, et al.* Chromosome 10q harbors a susceptibility locus for bipolar disorder in Ashkenazi Jewish families. Molecular Psychiatry, 2008; *13*(4)*:* 442-450.
99. Escamilla, MA, Ontiveros, A, Nicolini, H*, et al.* A genome-wide scan for schizophrenia and psychosis susceptibility loci in families of Mexican and Central American ancestry. American Journal Of Medical Genetics Part B-Neuropsychiatric Genetics, 2007; *144B*(2)*:* 193-199.
100. Herzberg, I, Jasinska, A, Garcia, J*, et al.* Convergent linkage evidence from two Latin-American population isolates supports the presence of a susceptibility locus for bipolar disorder in 5q31-34. Human Molecular Genetics, 2006; *15*(21)*:* 3146-3153.
101. Jones, I, Hamshere, M, Nangle, JM*, et al.* Bipolar affective puerperal psychosis: Genome-wide significant evidence for linkage to chromosome 16. American Journal Of Psychiatry, 2007; *164*(7)*:* 1099-1104.
102. Cassidy, F, Zhao, C, Badger, J*, et al.* Genome-wide scan of bipolar disorder and investigation of population stratification effects on linkage: Support for susceptibility loci at 4q21, 7q36, 9p21, 12q24, 14q24, and 16p13. American Journal of Medical Genetics Part B: Neuropsychiatric Genetics, 2007; *144B*(6)*:* 791-801.
103. Heils, A, Teufel, A, Petri, S*, et al.* Allelic variation of human serotonin transporter gene expression. Journal Of Neurochemistry, 1996; *66*(6)*:* 2621-2624.
104. Lesch, KP, Bengel, D, Heils, A*, et al.* Association of anxiety-related traits with a polymorphism in the serotonin transporter gene regulatory region. Science, 1996; *274*(5292)*:* 1527-1531.
105. Heinz, A, Jones, DW, Mazzanti, C*, et al.* A relationship between serotonin transporter genotype and in vivo protein expression and alcohol neurotoxicity. Biological Psychiatry, 2000; *47*(7)*:* 643-649.
106. Anguelova, M, Benkelfat, C, and Turecki, G A systematic review of association studies investigating genes coding for serotonin receptors and the serotonin transporter: I. Affective disorders. Molecular Psychiatry, 2003; *8*(6)*:* 574-591.

107. Anguelova, M, Benkelfat, C, and Turecki, G A systematic review of association studies investigating genes coding for serotonin receptors and the serotonin transporter: II. Suicidal behavior. Molecular Psychiatry, 2003; *8*(7)*:* 646-653.
108. Lotrich, FE and Pollock, BG Meta-analysis of serotonin transporter polymorphisms and affective disorders. Psychiatric Genetics, 2004; *14*(3)*:* 121-129.
109. Furlong, RA, Ho, L, Walsh, C*, et al.* Analysis and meta-analysis of two serotonin transporter gene polymorphisms in bipolar and unipolar affective disorders. American Journal Of Medical Genetics, 1998; *81*(1)*:* 58-63.
110. Caspi, A, Sugden, K, Moffitt, TE*, et al.* Influence of life stress on depression: Moderation by a polymorphism in the 5-HTT gene. Science, 2003; *301*(5631)*:* 386-389.
111. Eley, TC, Sugden, K, Corsico, A*, et al.* Gene-environment interaction analysis of serotonin system markers with adolescent depression. Molecular Psychiatry, 2004; *9*(10)*:* 908-915.
112. Gillespie, NA, Whitfield, JB, Williams, B, Heath, AC, and Martin, NG The relationship between stressful life events, the serotonin transporter (5-HTTLPR) genotype and major depression. Psychological Medicine, 2005; *35*(1)*:* 101-111.
113. Grabe, HJ, Lange, M, Wolff, B*, et al.* Mental and physical distress is modulated by a polymorphism in the 5-HT transporter gene interacting with social stressors and chronic disease burden. Molecular Psychiatry, 2005; *10*(2)*:* 220-224.
114. Kaufman, J, Yang, BZ, Douglas-Palumberi, H*, et al.* Social supports and serotonin transporter gene moderate depression in maltreated children. Proceedings Of The National Academy Of Sciences Of The United States Of America, 2004; *101*(49)*:* 17316-17321.
115. Kendler, KS, Kuhn, JW, Vittum, J, Prescott, CA, and Riley, B The interaction of stressful life events and a serotonin transporter polymorphism in the prediction of episodes of major depression - A replication. Archives Of General Psychiatry, 2005; *62*(5)*:* 529-535.
116. Scheid, JM, Holzman, CB, Jones, N*, et al.* Depressive symptoms in mid-pregnancy, lifetime stressors and the 5-HTTLPR genotype. Genes Brain And Behavior, 2007; *6*(5)*:* 453-464.
117. Sjoberg, RL, Nilsson, KW, Nordquist, N*, et al.* Development of depression: sex and the interaction between environment and a promoter polymorphism of the serotonin transporter gene. International Journal Of Neuropsychopharmacology, 2006; *9*(4)*:* 443-449.
118. Surtees, PG, Wainwright, NWJ, Willis-Owen, SAG, Luben, R, Day, NE, and Flint, J Social adversity, the serotonin transporter (5-HTTLPR) polymorphism and major depressive disorder. Biological Psychiatry, 2006; *59*(3)*:* 224-229.
119. Taylor, SE, Way, BM, Welch, WT, Hilmert, CJ, Lehman, BJ, and Eisenberger, NI Early family environment, current adversity, the serotonin transporter promoter polymorphism, and depressive symptomatology. Biological Psychiatry, 2006; *60*(7)*:* 671-676.

120. Wilhelm, K, Mitchell, PB, Niven, H*, et al.* Life events, first depression onset and the serotonin transporter gene. British Journal Of Psychiatry, 2006; *188*(210-215.
121. Zalsman, G, Huang, YY, Oquendo, MA*, et al.* Association of a triallelic serotonin transporter gene promoter region (5-HTTLPR) polymorphism with stressful life events and severity of depression. American Journal Of Psychiatry, 2006; *163*(9)*:* 1588-1593.
122. Kaufman, J, Yang, BZ, Douglas-Palumberi, H*, et al.* Brain-derived neurotrophic factor-5-HTTLPR gene interactions and environmental modifiers of depression in children. Biological Psychiatry, 2006; *59*(8)*:* 673-680.
123. Schumacher, J, Abou Jamra, R, Becker, T*, et al.* Evidence for a relationship between genetic variants at the brain-derived neurotrophic factor (*BDNF*) locus and major depression. Biological Psychiatry, 2005; *58*(4)*:* 307-314.
124. Baghai, TC, Binder, EB, Schule, C*, et al.* Polymorphisms in the angiotensin-converting enzyme gene are associated with unipolar depression, ACE activity and hypercortisolism. Molecular Psychiatry, 2006; *11*(11)*:* 1003-1015.
125. Zill, P, Baghai, TC, Schule, C*, et al.* SNP and haplotype analysis of a novel tryptophan hydroxylase isoform (TPH2) gene provide evidence for association with major depression. American Journal Of Medical Genetics Part B-Neuropsychiatric Genetics, 2004; *130B*(1)*:* 32-32.
126. Wong, ML, Whelan, F, Deloukas, P*, et al.* Phosphodiesterase genes are associated with susceptibility to major depression and antidepressant treatment response. Proceedings Of The National Academy Of Sciences Of The United States Of America, 2006; *103*(41)*:* 15124-15129.
127. Hashimoto, R, Numakawa, T, Ohnishi, T*, et al.* Impact of the DISC1 Ser704Cys polymorphism on risk for major depression, brain morphology and ERK signaling. Human Molecular Genetics, 2006; *15*(20)*:* 3024-3033.
128. Schiffer, HH and Heinemann, SF Association of the human kainate receptor GluR7 gene (GRIK3) with recurrent major depressive disorder. American Journal Of Medical Genetics Part B-Neuropsychiatric Genetics, 2007; *144B*(1)*:* 20-26.
129. López-León, S, Janssens, ACJW, González-Zuloeta Ladd, AM*, et al.* Meta-analyses of genetic studies on major depressive disorder. Molecular Psychiatry, 2008; *13*(772-785.
130. WHO Suicide prevention (SUPRE) http://www.who.int/mental_health/prevention/suicide/suicideprevent/en/index.html. 2007.
131. Chishti, P, Stone, DH, Corcoran, P, Williamson, E, and Petridou, E Suicide mortality in the European Union. European Journal Of Public Health, 2003; *13*(2)*:* 108-114.
132. Weissman, MM, Bland, RC, Canino, GJ*, et al.* Prevalence of suicide ideation and suicide attempts in nine countries. Psychological Medicine, 1999; *29*(1)*:* 9-17.

133. Kessler, RC, Borges, G, and Walters, EE Prevalence of and risk factors for lifetime suicide attempts in the national comorbidity survey. Archives Of General Psychiatry, 1999; *56*(7)*:* 617-626.
134. Bostwick, JM and Pankratz, VS Affective disorders and suicide risk: A reexamination. American Journal Of Psychiatry, 2000; *157*(12)*:* 1925-1932.
135. Brent, DA and Mann, JJ Family genetic studies, suicide, and suicidal behavior. American Journal Of Medical Genetics Part C-Seminars In Medical Genetics, 2005; *133C*(1)*:* 13-24.
136. Brent, DA, Bridge, J, Johnson, BA, and Connolly, J Suicidal behavior runs in families - A controlled family study of adolescent suicide victims. Archives Of General Psychiatry, 1996; *53*(12)*:* 1145-1152.
137. Cheng, ATA, Chen, THH, Chen, CC, and Jenkins, R Psychosocial and psychiatric risk factors for suicide - Case-control psychological autopsy study. British Journal Of Psychiatry, 2000; *177*(360-365.
138. Roy, A Relation of family history of suicide to suicide attempts in alcoholics. American Journal Of Psychiatry, 2000; *157*(12)*:* 2050-2051.
139. Roy, A Genetic and biologic risk factors for suicide in depressive disorders. Psychiatric Quarterly (New York), 1993; *64*(4)*:* 345-358.
140. Fu, Q, Heath, AC, Bucholz, KK*, et al.* A twin study of genetic and environmental influences on suicidality in men. Psychological Medicine, 2002; *32*(1)*:* 11-24.
141. Glowinski, AL, Bucholz, KK, Nelson, EC*, et al.* Suicide attempts in an adolescent female twin sample. Journal Of The American Academy Of Child And Adolescent Psychiatry, 2001; *40*(11)*:* 1300-1307.
142. Powell, J, Geddes, J, Deeks, J, Goldacre, M, and Hawton, K Suicide in psychiatric hospital in-patients - Risk factors and their predictive power. British Journal Of Psychiatry, 2000; *176*(266-272.
143. Statham, DJ, Heath, AC, Madden, PAF*, et al.* Suicidal behaviour: an epidemiological and genetic study. Psychological Medicine, 1998; *28*(4)*:* 839-855.
144. Wender, PH, Kety, SS, Rosenthal, D, Schulsinger, F, Ortmann, J, and Lunde, I Psychiatric-Disorders In The Biological And Adoptive Families Of Adopted Individuals With Affective-Disorders. Archives Of General Psychiatry, 1986; *43*(10)*:* 923-929.
145. Li, DW and He, L Further clarification of the contribution of the tryptophan hydroxylase (TPH) gene to suicidal behavior using systematic allelic and genotypic meta-analyses. Human Genetics, 2006; *119*(3)*:* 233-240.
146. Li, D and He, L Meta-analysis supports association between serotonin transporter (5-HTT) and suicidal behavior. Molecular Psychiatry, 2007; *12*(1)*:* 47-54.
147. Kia-Keating, BM, Glatt, SJ, and Tsuang, MT Meta-analyses suggest association between COMT, but not HTR1B, alleles, and suicidal behavior. American Journal Of Medical Genetics Part B-Neuropsychiatric Genetics, 2007; *144B*(8)*:* 1048-1053.
148. Kamali, M, Oquendo, MA, and Mann, JJ Understanding the neurobiology of suicidal behavior. Depression And Anxiety, 2001; *14*(3)*:* 164-176.

149. Westrin, A Stress system alterations and mood disorders in suicidal patients. A review. Biomedicine & Pharmacotherapy, 2000; *54*(3)*:* 142-145.
150. Strohle, A and Holsboer, F Stress responsive neurohormones in depression and anxiety. Pharmacopsychiatry, 2003; *36*(S207-S214.
151. Pfennig, A, Kunzel, HE, Kern, N*, et al.* Hypothalamus-pituitary-adrenal system regulation and suicidal behavior in depression. Biological Psychiatry, 2005; *57*(4)*:* 336-342.
152. Poo, MM Neurotrophins as synaptic modulators. Nature Reviews Neuroscience, 2001; *2*(1)*:* 24-32.
153. Alderson, RF, Curtis, R, Alterman, AL, Lindsay, RM, and DiStefano, PS Truncated TrkB mediates the endocytosis and release of *BDNF* and neurotrophin-4/5 by rat astrocytes and Schwann cells in vitro. Brain Research, 2000; *871*(2)*:* 210-222.
154. Soppet, D, Escandon, E, Maragos, J*, et al.* The Neurotrophic Factors Brain-Derived Neurotrophic Factor And Neurotrophin-3 Are Ligands For The Trkb Tyrosine Kinase Receptor. Cell, 1991; *65*(5)*:* 895-903.
155. Castren, E Neurotrophic effects of antidepressant drugs. Current Opinion In Pharmacology, 2004; *4*(1)*:* 58-64.
156. Nibuya, M, Morinobu, S, and Duman, RS Regulation Of Bdnf And Trkb Messenger-Rna In Rat-Brain By Chronic Electroconvulsive Seizure And Antidepressant Drug Treatments. Journal Of Neuroscience, 1995; *15*(11)*:* 7539-7547.
157. Russo-Neustadt, AA, Alejandre, H, Garcia, C, Ivy, AS, and Chen, MJ Hippocampal brain-derived neurotrophic factor expression following treatment with reboxetine, citalopram, and physical exercise. Neuropsychopharmacology, 2004; *29*(12)*:* 2189-2199.
158. Muller, MB, Toschi, N, Kresse, AE, Post, A, and Keck, ME Long-term repetitive transcranial magnetic stimulation increases the expression of brain-derived neurotrophic factor and cholecystokinin mRNA, but not neuropeptide tyrosine mRNA in specific areas of rat brain. Neuropsychopharmacology, 2000; *23*(2)*:* 205-215.
159. Chen, B, Dowlatshahi, D, MacQueen, GM, Wang, JF, and Young, LT Increased hippocampal *BDNF* immunoreactivity in subjects treated with antidepressant medication. Biological Psychiatry, 2001; *50*(4)*:* 260-265.
160. Siuciak, JA, Lewis, DR, Wiegand, SJ, and Lindsay, RM Antidepressant-like effect of brain-derived neurotrophic factor (*BDNF*). Pharmacology Biochemistry And Behavior, 1997; *56*(1)*:* 131-137.
161. Shirayama, Y, Chen, ACH, Nakagawa, S, Russell, DS, and Duman, RS Brain-derived neurotrophic factor produces antidepressant effects in behavioral models of depression. Journal Of Neuroscience, 2002; *22*(8)*:* 3251-3261.
162. Koponen, E, Lakso, M, and Castren, E Overexpression of the full-length neurotrophin receptor trkB regulates the expression of plasticity-related genes in mouse brain. Molecular Brain Research, 2004; *130*(1-2)*:* 81-94.
163. Monteggia, LM, Barrot, M, Powell, CM*, et al.* Essential role of brain-derived neurotrophic factor in adult hippocampal function. Proceedings

Of The National Academy Of Sciences Of The United States Of America, 2004; *101*(29)*:* 10827-10832.
164. Drevets, WC, Price, JL, Simpson, JR*, et al.* Subgenual prefrontal cortex abnormalities in mood disorders. Nature, 1997; *386*(6627)*:* 824-827.
165. Soares, JC and Mann, JJ The anatomy of mood disorders - Review of structural neuroimaging studies. Biological Psychiatry, 1997; *41*(1)*:* 86-106.
166. Bremner, JD, Narayan, M, Anderson, ER, Staib, LH, Miller, HL, and Charney, DS Hippocampal volume reduction in major depression. American Journal Of Psychiatry, 2000; *157*(1)*:* 115-117.
167. Frodl, T, Schuele, C, Schmitt, G*, et al.* Association of the brain-derived neurotrophic factor Val66Met polymorphism with reduced hippocampal volumes in major depression. Archives Of General Psychiatry, 2007; *64*(4)*:* 410-416.
168. Kim, YK, Lee, HP, Won, SD*, et al.* Low plasma *BDNF* is associated with suicidal behavior in major depression. Progress In Neuro-Psychopharmacology & Biological Psychiatry, 2007; *31*(1)*:* 78-85.
169. Deveci, A, Aydemir, O, Taskin, O, Taneli, F, and Esen-Danaci, A Serum *BDNF* Levels in Suicide Attempters Related to Psychosocial Stressors: A Comparative Study with Depression. Neuropsychobiology, 2007; *56*(93-97.
170. Dwivedi, Y, Rizavi, HS, Conley, RR, Roberts, RC, Tamminga, CA, and Pandey, GN Altered gene expression of brain-derived neurotrophic factor and receptor tyrosine kinase B in postmortem brain of suicide subjects. Archives Of General Psychiatry, 2003; *60*(8)*:* 804-815.
171. Karege, F, Vaudan, G, Schwald, M, Perroud, N, and La Harpe, R Neurotrophin levels in postmortem brains of suicide victims and the effects of antemortem diagnosis and psychotropic drugs. Molecular Brain Research, 2005; *136*(1-2)*:* 29-37.
172. Dwivedi, Y, Mondal, AC, Rizavi, HS, and Conley, RR Suicide brain is associated with decreased expression of neurotrophins. Biological Psychiatry, 2005; *58*(4)*:* 315-324.
173. Holsboer, F The rationale for corticotropin-releasing hormone receptor (CRH-R) antagonists to treat depression and anxiety. Journal Of Psychiatric Research, 1999; *33*(3)*:* 181-214.
174. Holsboer, F Psychiatric Implications Of Altered Limbic-Hypothalamic-Pituitary-Adrenocortical Activity. European Archives Of Psychiatry And Clinical Neuroscience, 1989; *238*(5-6)*:* 302-322.
175. Demitrack, MA and Crofford, LJ Evidence for and pathophysiologic implications of hypothalamic-pituitary-adrenal axis dysregulation in fibromyalgia and chronic fatigue syndrome. Neuroimmunomodulation, 1998; *840*(684-697.
176. Yehuda, R, Giller, EL, Southwick, SM, Lowy, MT, and Mason, JW Hypothalamic Pituitary-Adrenal Dysfunction In Posttraumatic-Stress-Disorder. Biological Psychiatry, 1991; *30*(10)*:* 1031-1048.
177. Arborelius, L, Owens, MJ, Plotsky, PM, and Nemeroff, CB The role of corticotropin-releasing factor in depression and anxiety disorders. Journal Of Endocrinology, 1999; *160*(1)*:* 1-12.

178. Nemeroff, CB, Owens, MJ, Bissette, G, Andorn, AC, and Stanley, M Reduced Corticotropin Releasing-Factor Binding-Sites In The Frontal-Cortex Of Suicide Victims. Archives Of General Psychiatry, 1988; *45*(6)*:* 577-579.
179. Hucks, D, Lowther, S, Crompton, MR, Katona, CLE, and Horton, RW Corticotropin-releasing factor binding sites in cortex of depressed suicides. Psychopharmacology, 1997; *134*(2)*:* 174-178.
180. Lopez, JF, Palkovits, M, Arato, M, Mansour, A, Akil, H, and Watson, SJ Localization And Quantification Of Proopiomelanocortin Messenger-Rna And Glucocorticoid Receptor Messenger-Rna In Pituitaries Of Suicide Victims. Neuroendocrinology, 1992; *56*(4)*:* 491-501.
181. Coryell, W and Schlesser, M The dexamethasone suppression test and suicide prediction. American Journal Of Psychiatry, 2001; *158*(5)*:* 748-753.
182. Porter, RJ, Gallagher, P, Watson, S, and Young, AH Corticosteroid-serotonin interactions in depression: a review of the human evidence. Psychopharmacology, 2004; *173*(1-2)*:* 1-17.
183. Lanfumey, L, Mongeau, R, Cohen-Salmon, C, and Hamon, M Corticosteroid-serotonin interactions in the neurobiological mechanisms of stress-related disorders. Neuroscience And Biobehavioral Reviews, 2008; *32*(6)*:* 1174-1184.
184. Castren, E Neurotrophins as mediators of drug effects on mood, addiction, and neuroprotection. Molecular Neurobiology, 2004; *29*(3)*:* 289-301.
185. Mamounas, LA, Blue, ME, Siuciak, JA, and Altar, CA Brain-derived neurotrophic factor promotes the survival and sprouting of serotonergic axons in rat brain. Journal Of Neuroscience, 1995; *15*(12)*:* 7929-7939.
186. Smith, MA, Makino, S, Kvetnansky, R, and Post, RM Stress And Glucocorticoids Affect The Expression Of Brain-Derived Neurotrophic Factor And Neurotrophin-3 Messenger-Rnas In The Hippocampus. Journal Of Neuroscience, 1995; *15*(3)*:* 1768-1777.
187. Duman, RS, Malberg, J, and Thome, J Neural plasticity to stress and antidepressant treatment. Biological Psychiatry, 1999; *46*(9)*:* 1181-1191.
188. Kumamaru, E, Numakawa, T, Adachi, N*, et al.* Glucocorticoid prevents brain-derived neurotrophic factor-mediated maturation of synaptic function in developing hippocampal neurons through reduction in the activity of mitogen-activated protein kinase. Molecular Endocrinology, 2008; *22*(3)*:* 546-558.
189. Duman, RS and Monteggia, LM A neurotrophic model for stress-related mood disorders. Biological Psychiatry, 2006; *59*(12)*:* 1116-1127.
190. Geller, B, Badner, JA, Tillman, R, Christian, SL, Bolhofner, K, and Cook, EH Linkage disequilibrium of the brain-derived neurotrophic factor Val66Met polymorphism in children with a prepubertal and early adolescent bipolar disorder phenotype. American Journal Of Psychiatry, 2004; *161*(9)*:* 1698-1700.
191. Green, EK, Raybould, R, MacGregor, S*, et al.* Genetic variation of brain-derived neurotrophic factor (*BDNF*) in bipolar disorder - Case-control

study of over 3000 individuals from the UK. British Journal Of Psychiatry, 2006; *188*(21-25.
192. Hong, CJ, Huo, SJ, Yen, FC, Tung, CL, Pan, GM, and Tsai, SJ Association study of a brain-derived neurotrophic-factor genetic polymorphism and mood disorders, age of onset and suicidal behavior. Neuropsychobiology, 2003; *48*(4)*:* 186-189.
193. Kunugi, H, Iijima, Y, Tatsumi, M*, et al.* No association between the Val66Met polymorphism of the brain-derived neurotrophic factor gene and bipolar disorder in a Japanese population: a Multicenter study. Biological Psychiatry, 2004; *56*(5)*:* 376-378.
194. Lohoff, FW, Sander, T, Ferraro, TN, Dahl, JP, Gallinat, J, and Berrettini, WH Confirmation of association between the Val66Met polymorphism in the brain-derived neurotrophic factor (*BDNF*) gene and bipolar I disorder. American Journal Of Medical Genetics Part B-Neuropsychiatric Genetics, 2005; *139B*(1)*:* 51-53.
195. Nakata, K, Ujike, H, Sakai, A*, et al.* Association study of the brain-derived neurotrophic factor (*BDNF*) gene with bipolar disorder. Neuroscience Letters, 2003; *337*(1)*:* 17-20.
196. Neves-Pereira, M, Mundo, E, Muglia, P, King, N, Macciardi, F, and Kennedy, JL The brain-derived neurotrophic factor gene confers susceptibility to bipolar disorder: Evidence from a family-based association study. American Journal Of Human Genetics, 2002; *71*(3)*:* 651-655.
197. Oswald, P, Del-Favero, J, Massat, I*, et al.* No implication of brain-derived neurotrophic factor (*BDNF*) gene in unipolar affective disorder: Evidence from Belgian first and replication patient-control studies. European Neuropsychopharmacology, 2005; *15*(5)*:* 491-495.
198. Sklar, P, Gabriel, SB, McInnis, MG*, et al.* Family-based association study of 76 candidate genes in bipolar disorder: *BDNF* is a potential risk locus. Molecular Psychiatry, 2002; *7*(6)*:* 579-593.
199. Tsai, SJ, Cheng, CY, Yu, YWY, Chen, TJ, and Hong, CJ Association study of a brain-derived neurotrophic-factor genetic polymorphism and major depressive disorders, symptomatology, and antidepressant response. American Journal Of Medical Genetics Part B-Neuropsychiatric Genetics, 2003; *123B*(1)*:* 19-22.
200. Ribeiro, L, Busnello, JV, Cantor, RM*, et al.* The brain-derived neurotrophic factor rs6265 (Val66Met) polymorphism and depression in Mexican-Americans. Neuroreport, 2007; *18*(12)*:* 1291-1293.
201. Surtees, PG, Wainwright, NWJ, Willis-Owen, SAG*, et al.* No association between the *BDNF* Val66Met polymorphism and mood status in a non-clinical community sample of 7389 older adults. Journal Of Psychiatric Research, 2007; *41*(5)*:* 404-409.
202. Egan, MF, Kojima, M, Callicott, JH*, et al.* The *BDNF* val66met polymorphism affects activity-dependent secretion of *BDNF* and human memory and hippocampal function. Cell, 2003; *112*(2)*:* 257-269.
203. Chen, ZY, Patel, PD, Sant, G*, et al.* Variant brain-derived neurotrophic factor (*BDNF*) (Met66) alters the intracellular trafficking and activity-

dependent secretion of wild-type *BDNF* in neurosecretory cells and cortical neurons. Journal Of Neuroscience, 2004; *24*(18)*:* 4401-4411.
204. Hariri, AR, Goldberg, TE, Mattay, VS*, et al.* Brain-derived neurotrophic factor val(66)met polymorphism affects human memory-related hippocampal activity and predicts memory performance. Journal Of Neuroscience, 2003; *23*(17)*:* 6690-6694.
205. Stern, AJ, Marenco, S, van der Veen, JW*, et al.* Impact of the *BDNF* va166met polymorphism on levels of hippocampal N-acetyl aspartate: A magnetic resonance spectroscopic imaging study. Biological Psychiatry, 2007; *61*(8)*:* 158.
206. Bueller, JA, Aftab, M, Sen, S, Gomez-Hassan, D, Burmeister, M, and Zubieta, JK *BDNF* val(66) met allele is associated with reduced hippocampal volume in healthy subjects. Biological Psychiatry, 2006; *59*(9)*:* 812-815.
207. Pezawas, L, Verchinski, BA, Mattay, VS*, et al.* The brain-derived neurotrophic factor val66met polymorphism and variation in human cortical morphology. Journal Of Neuroscience, 2004; *24*(45)*:* 10099-10102.
208. Szeszko, PR, Lipsky, R, Mentschel, C*, et al.* Brain-derived neurotrophic factor val66met polymorphism and volume of the hippocampal formation. Molecular Psychiatry, 2005; *10*(7)*:* 631-636.
209. Kunugi, H, Hashimoto, R, Yoshida, M, Tatsumi, M, and Kamijima, K A missense polymorphism (S205L) of the low-affinity neurotrophin receptor p75(NTR) gene is associated with depressive disorder and attempted suicide. American Journal Of Medical Genetics Part B-Neuropsychiatric Genetics, 2004; *129B*(1)*:* 44-46.
210. McGregor, S, Strauss, J, Bulgin, N*, et al.* p75(NTR) Gene and suicide attempts in young adults with a history of childhood-onset mood disorder. American Journal Of Medical Genetics Part B-Neuropsychiatric Genetics, 2007; *144B*(5)*:* 696-700.
211. Pruitt, KD, Tatusova, T, and Maglott, DR NCBI Reference Sequence (RefSeq): a curated non-redundant sequence database of genomes, transcripts and proteins. Nucleic Acids Research, 2005; *33*(501-504.
212. Hennings, JM, Owashi, T, Binder, EB*, et al.* Clinical characteristics and treatment outcome in a representative sample of depressed inpatients - Findings from the Munich Antidepressant Response Signature (MARS) project. Journal of Psychiatric Research, 2008; *Epub ahead of print*(
213. Frodl, T, Meisenzahl, EM, Zetzsche, T*, et al.* Hippocampal changes in patients with a first episode of major depression. American Journal Of Psychiatry, 2002; *159*(7)*:* 1112-1118.
214. Wang, JL Work stress as a risk factor for major depressive episode(s). Psychological Medicine, 2005; *35*(6)*:* 865-871.
215. Schmidt, MV, Sterlemann, V, Ganea, K*, et al.* Persistent neuroendocrine and behavioral effects of a novel, etiologically relevant mouse paradigm for chronic social stress during adolescence. Psychoneuroendocrinology, 2007; *32*(5)*:* 417-429.
216. Sterlemann, V, Ganea, K, Liebl, C*, et al.* Long-term behavioral and neuroendocrine alterations following chronic social stress in mice:

Implications for stress-related disorders. Hormones And Behavior, 2008; *53*(386-394.
217. Robins, LN, Wing, J, Wittchen, HU*, et al.* The Composite International Diagnostic Interview - An Epidemiologic Instrument Suitable For Use In Conjunction With Different Diagnostic Systems And In Different Cultures. Archives Of General Psychiatry, 1988; *45*(12)*:* 1069-1077.
218. Semler, G, Wittchen, HU, Joschke, K*, et al.* Test-Retest Reliability Of A Standardized Psychiatric Interview (Dis/Cidi). European Archives Of Psychiatry And Clinical Neuroscience, 1987; *236*(4)*:* 214-222.
219. Wittchen and UH Screening for mental disorders: Performance of the Composite International Diagnostic-Screener (CID-S). International Journal of Methods in Psychiatric Research, 1999; 8)*:* 59-70.
220. Binder, EB, Salyakina, D, Lichtner, P*, et al.* Polymorphisms in FKBP5 are associated with increased recurrence of depressive episodes and rapid response to antidepressant treatment. Nature Genetics, 2004; *36*(12)*:* 1319-1325.
221. Fisher, RA Statistical Methods for research workers. London: Oliver and Boyd., 1932.
222. Barrett, JC, Fry, B, Maller, J, and Daly, MJ Haploview: analysis and visualization of LD and haplotype maps. Bioinformatics, 2005; *21*(2)*:* 263-265.
223. Stranger, BE, Nica, AC, Forrest, MS*, et al.* Population genomics of human gene expression. Nature Genetics, 2007; *39*(1217-1224.
224. Stranger, BE, Forrest, MS, Clark, AG*, et al.* Genome-wide associations of gene expression variation in humans. Plos Genetics, 2005; *1*(6)*:* 695-704.
225. Hinds, DA, Stuve, LL, Nilsen, GB*, et al.* Whole-genome patterns of common DNA variation in three human populations. Science, 2005; *307*(5712)*:* 1072-1079.
226. UCSC Human Genome Browser version 'March 2006' (hg18). See: www.genome.ucsc.edu/cgi-bin/hgGateway.
227. Beck, AT, Erbaugh, J, Ward, CH, Mock, J, and Mendelsohn, M An Inventory For Measuring Depression. Archives Of General Psychiatry, 1961; *4*(6)*:* 561.
228. Provencher, SW Estimation Of Metabolite Concentrations From Localized In-Vivo Proton Nmr-Spectra. Magnetic Resonance In Medicine, 1993; *30*(6)*:* 672-679.
229. Viinamaki, H, Tanskanen, A, Honkalampi, K*, et al.* Is the Beck Depression Inventory suitable for screening major depression in different phases of the disease? Nordic Journal Of Psychiatry, 2004; *58*(1)*:* 49-53.
230. Grachev, ID and Apkarian, AV Chemical heterogeneity of the living human brain: a proton MR spectroscopy study on the effects of sex, age, and brain region. Neuroimage, 2000; *11*(5 Pt 1)*:* 554-563.
231. Ashburner, J and Friston, KJ Unified segmentation. Neuroimage, 2005; *26*(3)*:* 839-851.
232. Amunts, K, Kedo, O, Kindler, M*, et al.* Cytoarchitectonic mapping of the human amygdala, hippocampal region and entorhinal cortex: intersubject

variability and probability maps. Anatomy And Embryology, 2005; *210*(5-6)*:* 343-352.
233. Eickhoff, SB, Stephan, KE, Mohlberg, H*, et al.* A new SPM toolbox for combining probabilistic cytoarchitectonic maps and functional imaging data. Neuroimage, 2005; *25*(4)*:* 1325-1335.
234. Merali, Z, Levac, C, and Anisman, H Validation of a simple, ethologically relevant paradigm for assessing anxiety in mice. Biological Psychiatry, 2003; *54*(5)*:* 552-565.
235. Datson, NA, Meijer, L, Steenbergen, PJ*, et al.* Expression profiling in laser-microdissected hippocampal subregions in rat brain reveals large subregion-specific differences in expression. European Journal Of Neuroscience, 2004; *20*(10)*:* 2541-2554.
236. Chudin, E, Kruglyak, S, Baker, SC, Oeser, S, Barker, D, and McDaniel, TK A model of technical variation of microarray signals. Journal Of Computational Biology, 2006; *13*(4)*:* 996-1003.
237. Claassen, CA, Trivedi, MH, Rush, AJ*, et al.* Clinical differences among depressed patients with and without a history of suicide attempts: findings from the STAR*D trial. Journal of affective disorders, 2007; *97*(1-3)*:* 77-84.
238. WTCCC Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. Nature, 2007; *447*(7145)*:* 661-U667.
239. Baum, AE, Akula, N, Cabanero, M*, et al.* A genome-wide association study implicates diacylglycerol kinase eta (DGKH) and several other genes in the etiology of bipolar disorder. Molecular Psychiatry, 2008; *13*(197-207.
240. Leal, SM The detection of genotyping errors and pseudo-SNPs via deviations from Hardy Weinberg equilibrium. Annals Of Human Genetics, 2005; *69*(769-769.
241. Blazer, DG, Kessler, RC, McGonagle, KA, and Swartz, MS The prevalence and distribution of major depression in a national community sample: the National Comorbidity Survey. American Journal of Psychiatry, 1994; *151*(7)*:* 979-986.
242. Broer, A, Tietze, N, Kowalczuk, S*, et al.* The orphan transporter v7-3 (slc6a15) is a Na+-dependent neutral amino acid transporter (B(0)AT2). Biochemical Journal, 2006; *393*(421-430.
243. UniGene NCBI, EST Profile Viewer. http://www.ncbi.nlm.nih.gov/sites/entrez?db=unigene.
244. Myers, S, Bottolo, L, Freeman, C, McVean, G, and Donnelly, P A fine-scale map of recombination rates and hotspots across the human genome. Science, 2005; *310*(5746)*:* 321-324.
245. McVean, GAT, Myers, SR, Hunt, S, Deloukas, P, Bentley, DR, and Donnelly, P The fine-scale structure of recombination rate variation in the human genome. Science, 2004; *304*(5670)*:* 581-584.
246. Myers, AJ, Gibbs, JR, Awebster, J*, et al.* A survey of genetic human cortical gene expression. Nature Genetics, 2007; *39*(1494-1499.
247. Lein, ES, Hawrylycz, MJ, Ao, N*, et al.* Genome-wide atlas of gene expression in the adult mouse brain. Nature, 2007; *445*(7124)*:* 168-176.

248. Allen Brain Atlas [Internet]. Seattle (WA): Allen Institute for Brain Science. © 2008*Available from:* http://www.brain-map.org.(
249. Capizzano, AA, Jorge, RE, Acion, LC, and Robinson, RG In vivo proton magnetic resonance spectroscopy in patients with mood disorders: A technically oriented review. Journal Of Magnetic Resonance Imaging, 2007; *26*(1378-1389.
250. Shekhar, A, Truitt, W, Rainnie, D, and Sajdyk, T Role of stress, corticotrophin releasing factor (CRF) and amygdala plasticity in chronic anxiety. Stress-The International Journal On The Biology Of Stress, 2005; *8*(4)*:* 209-219.
251. Bale, TL Stress sensitivity and the development of affective disorders. Hormones And Behavior, 2006; *50*(4)*:* 529-533.
252. Ryan, MM, Lockstone, HE, Huffaker, SJ, Wayland, MT, Webster, MJ, and Bahn, S Gene expression analysis of bipolar disorder reveals downregulation of the ubiquitin cycle and alterations in synaptic genes. Molecular Psychiatry, 2006; *11*(10)*:* 965-978.
253. Zubenko, GS, Maher, BS, Hughes, HB, Zubenko, WN, Stiffler, JS, and Marazita, ML Genome-wide linkage survey for genetic loci that affect the risk of suicide attempts in families with recurrent, early-onset, major depression. American Journal Of Medical Genetics Part B-Neuropsychiatric Genetics, 2004; *129B*(1)*:* 47-54.
254. Cystic Fibrosis Mutation Database. In: http://www.genet.sickkids.on.ca/cftr/. 2008.
255. Quinton Paul, M Too much salt, too little soda: cystic fibrosis. Shengli Xuebao, 2007; *59*(4)*:* 397-415.
256. Cichon, S, Winge, I, Mattheisen, M*, 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; *17*(87-97.
257. Zhang, XD, Gainetdinov, RR, Beaulieu, JM*, et al.* Loss-of-function mutation in tryptophan hydroxylase-2 identified in unipolar major depression. Neuron, 2005; *45*(1)*:* 11-16.
258. Cartharius, K, Frech, K, Grote, K*, et al.* MatInspector and beyond: promoter analysis based on transcription factor binding sites. Bioinformatics, 2005; *21*(13)*:* 2933-2942.
259. Gardinergarden, M and Frommer, M Cpg Islands In Vertebrate Genomes. Journal Of Molecular Biology, 1987; *196*(2)*:* 261-282.
260. Taylor, KH, Kramer, RS, Davis, JW*, et al.* Ultradeep bisulfite sequencing analysis of DNA methylation patterns in multiple gene promoters by 454 sequencing. Cancer Research, 2007; *67*(18)*:* 8511-8518.
261. Mil, J, Tang, T, Kaminsky, Z*, et al.* Epigenomic profiling reveals DNA-Methylation changes associated with major psychosis. American Journal Of Human Genetics, 2008; *82*(3)*:* 696-711.
262. Ninkina, N, Adu, J, Fischer, A, Pinon, LGP, Buchman, VL, and Davies, AM Expression and function of TrkB variants in developing sensory neurons. Embo Journal, 1996; *15*(23)*:* 6385-6393.
263. Eide, FF, Vining, ER, Eide, BL, Zang, KL, Wang, XY, and Reichardt, LF Naturally occurring truncated trkB receptors have dominant inhibitory

effects on brain-derived neurotrophic factor signaling. Journal Of Neuroscience, 1996; *16*(10)*:* 3123-3129.
264. Biffo, S, Offenhauser, N, Carter, BD, and Barde, YA Selective Binding And Internalization By Truncated Receptors Restrict The Availability Of Bdnf During Development. Development, 1995; *121*(8)*:* 2461-2470.
265. Sklar, P, Smoller, JW, Fan, J*, et al.* Whole-genome association study of bipolar disorder. Molecular Psychiatry, 2008; *13*(6)*:* 558-569.
266. Broer, S The SLC6 orphans are forming a family of amino acid transporters. Neurochemistry International, 2006; *48*(6-7)*:* 559-567.
267. Masson, J, Pohl, M, Aidouni, Z, Giros, B, Hamon, M, and ElMestikawy, S The two orphan Na+/Cl--dependent transporters Rxt1 and V-7-3-2 have an overlapping expression pattern in the rat central nervous system. Receptors & Channels, 1996; *4*(4)*:* 227-242.
268. Dronova, J, Liu, QR, Hall, FS, Krieger, RM, and Uhl, GR Deletion of v7-3 (*SLC6A15*) transporter allows assessment of its roles in synaptosomal proline uptake, leucine uptake and behaviors. Brain Research, 2007; *1183*(10-20.
269. Tapiero, H, Mathe, G, Couvreur, P, and Tew, KD Dossier: Free amino acids in human health and pathologies - II. Glutamine and glutamate. Biomedicine & Pharmacotherapy, 2002; *56*(9)*:* 446-457.
270. Zhou, Z, Zhen, J, Karpowich, NK*, et al.* LeuT-desipramine structure reveals how antidepressants block neurotransmitter reuptake. Science, 2007; *317*(5843)*:* 1390-1393.
271. Morissette, J, Villeneuve, A, Bordeleau, L*, et al.* Genome-wide search for linkage of bipolar affective disorders in a very large pedigree derived from a homogeneous population in Quebec points to a locus of major effect on chromosome 12q23-q24. American Journal Of Medical Genetics, 1999; *88*(5)*:* 567-587.
272. Kleinjan, DA, Seawright, A, Mella, S*, et al.* Long-range downstream enhancers are essential for Pax6 expression. Developmental Biology, 2006; *299*(2)*:* 563-581.
273. van Heyningen, V, McBride, D, and Kleinjan, DA Long-range gene regulation in development and disease. Journal Of Medical Genetics, 2006; *43*(S15-S15.
274. Birney, E, Stamatoyannopoulos, JA, Dutta, A*, et al.* Identification and analysis of functional elements in 1% of the human genome by the ENCODE pilot project. Nature, 2007; *447*(7146)*:* 799-816.
275. Blasi, G, Bertolino, A, Brudaglio, F*, et al.* Hippocampal neurochemical pathology in patients at first episode of affective psychosis: a proton magnetic resonance spectroscopic imaging study. Psychiatry Research-Neuroimaging, 2004; *131*(2)*:* 95-105.
276. Deicken, RF, Pegues, MP, Anzalone, S, Feiwell, R, and Soher, B Lower concentration of hippocampal N-acetylaspartate in familial bipolar I disorder. American Journal Of Psychiatry, 2003; *160*(5)*:* 873-882.
277. Schuff, N, Neylan, TC, Fox-Bosetti, S*, et al.* Abnormal N-acetylaspartate in hippocampus and anterior, cingulate in posttraumatic stress disorder. Psychiatry Research-Neuroimaging, 2008; *162*(2)*:* 147-157.

278. Czeh, B, Michaelis, T, Watanabe, T*, et al.* Stress-induced changes in cerebral metabolites, hippocampal volume, and cell proliferation are prevented by antidepressant treatment with tianeptine. Proceedings Of The National Academy Of Sciences Of The United States Of America, 2001; *98*(22)*:* 12796-12801.
279. Li, CX, Wang, Y, Gao, H*, et al.* Cerebral Metabolic Changes in a Depression-like Rat Model of Chronic Forced Swimming Studied by Ex vivo High Resolution (1)H Magnetic Resonance Spectroscopy. Neurochemical Research, 2008 (in press).
280. Glahn, DC, Thompson, PM, and Blangero, J Neuroimaging endophenotypes: Strategies for finding genes influencing brain structure and function. Human Brain Mapping, 2007; *28*(6)*:* 488-501.
281. Videbech, P and Ravnkilde, B Hippocampal volume and depression: A meta-analysis of MRI studies. American Journal Of Psychiatry, 2004; *161*(11)*:* 1957-1966.
282. Benarroch, EE N-acetylaspartate and N-acetylaspartylglutamate: neurobiology and clinical significance. Neurology, 2008; *70*(16)*:* 1353-1357.
283. Herman, JP, Figueiredo, H, Mueller, NK*, et al.* Central mechanisms of stress integration: hierarchical circuitry controlling hypothalamo-pituitary-adrenocortical responsiveness. Frontiers In Neuroendocrinology, 2003; *24*(3)*:* 151-180.
284. Craddock, N and Forty, L Genetics of affective (mood) disorders. European Journal Of Human Genetics, 2006; *14*(6)*:* 660-668.
285. Merikangas, KR, Akiskal, HS, Angst, J*, et al.* Lifetime and 12-month prevalence of bipolar spectrum disorder in the national comorbidity survey replication. Archives Of General Psychiatry, 2007; *64*(5)*:* 543-552.

9 SUPPLEMENTARY TABLES

Suppl. table. 1: Quality control and association results of BDNF locus mapping in the discovery sample. 18 tagging SNPs fully covering the *BDNF* locus according to HapMap Phase I and II data fulfilled quality control criteria: MAF>1%, no nominal deviation from HWE and with case-wise call rates per SNP higher than 99%. The second last column (CC-MDD-logp) shows the negative decadean logarithm of the case-control association p-values in MDD association testing. The last column (SA-logp) gives the corresponding association results for attempted suicide (SA) among mood disorder patient. Position (hg18) represents the chromosome position of each marker according to UCSC (hg18) / dbSNP Built 128.

SNP	Position hg18	Min./Maj. Allele	MAF	HWE p	N	Callrate %	CC-MDD -logp	SA -logp
rs10835211	27657941	A/G	0,22	0,67	771	100	0,40	0,28
rs11030094	27616351	A/G	0,43	0,71	770	99,9	0,61	0,58
rs11030109	27653527	A/G	0,02	1,00	771	100	0,34	0,70
rs1157659	27714198	C/T	0,50	0,72	771	100	1,18	0,34
rs11602246	27617502	G/C	0,10	0,23	770	99,9	0,32	0,36
rs12273363	27701435	C/T	0,17	0,44	769	99,7	0,14	0,26
rs1491850	27706301	C/T	0,44	0,21	771	100	0,91	0,45
rs1491851	27709339	T/C	0,44	0,83	771	100	1,76	0,11
rs1552736	27719853	A/G	0,39	0,06	769	99,7	0,98	0,51
rs2049046	27680351	A/T	0,45	0,61	771	100	1,15	0,40
rs2049048	27707162	T/C	0,16	0,78	771	100	0,57	0,20
rs4923468	27682351	A/C	0,01	1,00	771	100	0,15	0,08
rs6265	27636492	A/G	0,22	0,25	771	100	1,46	0,62
rs7130131	27619402	C/T	0,06	0,20	770	99,9	0,17	0,24
rs727155	27707025	T/C	0,04	0,63	771	100	0,20	0,00
rs7482257	27711874	T/C	0,33	0,10	770	99,9	1,77	0,26
rs908867	27702340	A/G	0,09	0,67	771	100	0,67	0,04
rs925946	27623778	T/G	0,27	0,47	771	100	1,35	0,32
Genotypes of 18 BDNF tagging SNPs fulfilling quality control criteria		**minimum:** **maximum:** **mean:**	**0,01** **0,50** **0,24**	**0,06** **1,00** **0,53**	**769** **771** **771**	**99,7** **100** **99,9**	**0,14** **1,77** **0,78**	**0,00** **0,70** **0,32**

Suppl. table. 2: Quality control and association results of the NTRK2 locus mapping in the discovery sample.

SNP	Position hg18	Min./Maj. Allele	MAF	HWE p	N	Callrate %	CC-MDD -logp	SA -logp
rs1047896	86615793	C/T	0,19	0,29	763	99,0	0,19	1,29
rs10512159	86798819	A/G	0,17	0,37	771	100	0,35	0,33
rs1078947	86753072	A/G	0,15	1,00	771	100	0,26	1,15
rs10868223	86454946	T/C	0,12	0,13	771	100	0,32	0,95
rs10868235	86683575	C/T	0,48	0,94	771	100	1,28	1,93
rs10868241	86782848	A/G	0,29	0,19	770	99,9	1,37	0,08
rs11140714	86456520	G/A	0,22	0,68	770	99,9	1,78	3,59
rs11140771	86633182	A/G	0,21	0,83	771	100	0,43	1,02
rs11140776	86636747	T/G	0,47	1,00	764	99,1	0,52	1,26
rs11140778	86636978	T/A	0,21	0,58	771	100	0,29	0,77
rs11140783	86644637	T/C	0,10	0,56	771	100	0,76	0,81
rs11140793	86681300	C/A	0,16	0,41	771	100	1,28	0,37
rs11140800	86697957	C/A	0,42	0,82	768	99,6	0,87	1,63
rs11140803	86702489	A/G	0,18	0,90	771	100	0,53	0,68
rs11140810	86764628	T/G	0,45	0,19	770	99,9	0,89	1,28
rs1147198	86465168	C/A	0,23	0,09	771	100	1,26	1,49
rs1187274	86609609	G/C	0,30	0,73	771	100	0,49	1,21
rs1187286	86604848	C/A	0,23	0,31	771	100	0,20	2,23
rs1187329	86479100	A/G	0,45	0,38	771	100	0,36	1,31
rs1187363	86527740	A/G	0,20	0,17	771	100	0,05	0,26
rs12338909	86770879	T/G	0,09	0,83	768	99,6	0,67	0,72
rs1387924	86822813	A/C	0,18	0,12	771	100	0,80	0,33
rs1387926	86822483	A/G	0,13	0,52	770	99,9	0,63	1,97
rs1439050	86478013	T/G	0,33	0,94	771	100	1,24	3,06
rs1443440	86595988	G/A	0,28	0,59	764	99,1	0,67	0,56
rs1443445	86723978	C/T	0,21	0,19	771	100	0,72	1,26
rs1490402	86863073	G/A	0,14	0,76	765	99,2	0,36	1,07
rs1490403	86828326	A/T	0,29	0,79	771	100	1,55	0,43
rs1619120	86492016	T/C	0,41	0,50	769	99,7	0,92	0,73
rs1627784	86618686	G/A	0,27	0,36	759	98,4	0,10	0,88
rs1659412	86467972	C/T	0,11	0,003	757	98,2	0,81	1,42
rs1662695	86595032	C/T	0,12	0,74	771	100	0,60	1,02
rs17418241	86815668	T/C	0,09	0,64	769	99,7	0,16	2,47
rs1778931	86513595	A/G	0,47	0,72	771	100	0,78	0,83
rs1778933	86514231	C/T	0,35	0,64	768	99,6	0,74	0,91
rs1822420	86657096	T/C	0,14	0,31	769	99,7	1,18	0,77
rs1867283	86640586	A/G	0,48	0,89	771	100	0,52	2,69
rs1948308	86806077	C/T	0,45	0,38	771	100	0,72	0,21
rs2165893	86721244	G/A	0,06	0,53	771	100	0,23	0,02
rs2277192	86763797	G/A	0,16	0,79	771	100	0,54	0,21
rs2277193	86763829	C/T	0,30	0,67	768	99,6	1,17	2,13
rs2289656	86753382	T/C	0,18	1,00	770	99,9	0,74	0,95
rs2378672	86820854	G/A	0,08	0,64	771	100	0,23	0,07
rs2586566	86760732	G/A	0,23	0,41	770	99,9	0,45	0,43
rs2808707	86748114	T/G	0,41	0,46	771	100	0,56	1,16
rs3739570	86827398	T/C	0,09	0,53	771	100	0,71	0,26
rs3739804	86611451	G/A	0,08	0,79	771	100	0,24	0,10
rs3758317	86472435	T/C	0,18	0,81	771	100	2,03	1,28
rs3824519	86759824	A/G	0,07	1,00	771	100	0,14	0,63
rs3860945	86775444	G/A	0,13	0,44	771	100	0,64	1,14
rs4361832	86785554	A/G	0,18	0,12	771	100	0,80	0,06
rs4406490	86790617	C/A	0,46	0,82	721	93,5	1,42	1,18
rs4486281	86645167	G/A	0,30	0,34	770	99,9	0,22	0,84
rs614886	86872295	T/G	0,31	0,23	748	97,0	0,56	0,91
rs6559836	86723209	G/A	0,19	0,35	771	100	0,99	0,17
rs6559840	86765320	T/C	0,27	0,41	764	99,1	0,14	0,78
rs681329	86849219	C/T	0,13	0,63	771	100	0,55	0,52
rs7023589	86561917	G/A	0,34	0,42	771	100	0,32	0,70
rs7026417	86778985	C/T	0,10	0,07	771	100	0,15	0,34
rs7048015	86667955	C/A	0,19	0,34	771	100	0,73	0,75
rs7855888	86641878	C/T	0,26	0,17	771	100	1,02	1,21
rs7859023	86664240	G/A	0,16	0,51	770	99,9	1,63	1,19
rs7875184	86690918	T/C	0,15	0,89	771	100	0,50	0,36
rs984430	86772647	A/G	0,18	0,23	771	100	0,78	0,19
rs9969765	86679605	G/C	0,34	0,34	768	99,6	2,03	0,55
Genotypes of 65 NTRK2 tagging SNPs fulfilling quality control criteria		**minimum:**	**0,06**	**0,003**	**721**	**93,5**	**0,05**	**0,02**
		maximum:	**0,48**	**1,00**	**771**	**100**	**2,03**	**3,59**
		mean:	**0,24**	**0,53**	**768**	**99,7**	**0,71**	**0,97**

65 tagging SNPs, highly covering the *NTRK2* locus according to HapMap data, fulfilled quality control criteria: MAF > 10%, no nominal deviation from HWE and with case-wise call rates per SNP higher than 98%. The second last column (CC – MDD - logp) shows the negative decadean logarithm of the case -control association testing in MDD. The last column (SA-log p) gives the association results for attempted suicide (SA) among mood disorder patient. Chromosome positions of each marker according to UCSC (hg18) / dbSNP Built 128 are given.

Suppl. table. 3: Quality control and association results of the replication study.

All 12 SNPs which were selected for replication fulfilled quality control criteria: MAF>5%, no deviation from HWE after correction for multiple testing (p>0.2), and with case-wise call rates per SNP higher than 98%. The best genetic model with the same direction of the effect as was observed in the discovery sample was tested in the replication. The second last column (CC-MDD-logp) shows the negative decadean logarithm of the case-control association test with unipolar depression (MDD). The last column (SA-logp) shows association results for attempted suicide (SA) among mood disorder patient. Position (hg18) represents the chromosome position of each marker according to UCSC (hg18) / dbSNP Built 128.

SNP	Position hg18	Min./Maj. Allele	MAF	HWE p	N	Callrate %	CC-MDD -logp	SA -logp
rs10868235	86683575	C/T	0,48	0,08	1934	99,6	1,24	2,07
rs11140714	86456520	G/A	0,21	0,17	1915	98,7	1,45	0,05
rs11140800	86697957	C/A	0,40	0,37	1917	98,8	0,49	1,41
rs1147198	86465168	C/A	0,22	0,84	1916	98,7	0,44	1,80
rs1187286	86604848	C/A	0,23	0,66	1938	99,8	0,64	0,78
rs1187329	86479100	A/G	0,43	0,55	1933	99,6	0,45	0,05
rs1387926	86822483	A/G	0,14	0,50	1934	99,6	0,88	0,74
rs1439050	86478013	T/G	0,32	0,47	1937	99,8	1,03	0,07
rs1659412	86467972	C/T	0,10	0,90	1938	99,8	1,13	0,05
rs17418241	86815668	T/C	0,10	0,36	1918	98,8	1,00	1,23
rs1867283	86640586	A/G	0,50	0,44	1917	98,8	0,76	1,39
rs2277193	86763829	C/T	0,31	0,02	1913	98,6	0,97	0,79
Quality control of the 12 NTRK2 SNPs genotyped in the replication sample		**minimum:**	**0,10**	**0,02**	**1913**	**98,6**	**0,44**	**0,05**
		maximum:	**0,50**	**0,90**	**1938**	**99,8**	**1,45**	**2,07**
		mean:	**0,29**	**0,45**	**1926**	**99,2**	**0,87**	**0,87**

Suppl. tab. 4: Genotypes and HWE data for the associated SNPs on 12q21.31

SNP	Discovery sample (GWAS)		Replication sample	
	Controls	Cases	Controls	Cases
rs10506904	N(GG) = 52 (*63.13*) N(AG) = *200* (177.75) N(AA) = 114 (*125.13*) f(G) = 0.42 +/- 0.017 HetEx: 0.125 HWE(p) = **0.018**	N(GG) = *94* (92.30) N(AG) = 173 (*176.41*) N(AA) = *86* (84.30) f(G) = 0.51 +/- 0.019 HetEx: -0.019 HWE(p) = 0.749	N(GG) = 200 (*207.45*) N(AG) = *505* (490.09) N(AA) = 282 (*289.45*) f(G) = 0.46 +/- 0.011 HetEx: 0.030 HWE(p) = 0.369	N(GG) = *222* (216.05) N(AG) = 433 (*444.91*) N(AA) = *235* (229.05) f(G) = 0.49 +/- 0.012 HetEx: -0.027 HWE(p) = 0.422
rs1545843	N(AA) = 44 (*60.66*) N(AG) = *210* (176.68) N(GG) = 112 (*128.66*) f(A) = 0.41 +/- 0.016 HetEx: 0.189 HWE(p) = **3.5e-04**	N(AA) = *99* (91.78) N(AG) = 162 (*176.43*) N(GG) = *92* (84.78) f(A) = 0.51 +/- 0.020 HetEx: -0.082 HWE(p) = 0.136	N(AA) = *206* (205.25) N(AG) = 504 (*505.50*) N(GG) = *312* (311.25) f(A) = 0.45 +/- 0.011 HetEx: -0.003 HWE(p) = 0.950	N(AA) = *222* (221.81) N(AG) = 458 (*458.38*) N(GG) = *237* (236.81) f(A) = 0.49 +/- 0.012 HetEx: -0.0008 HWE(p) = 1.0
rs1473442	N(CC) = 36 (*49.92*) N(AC) = *195* (167.16) N(AA) = 126 (*139.92*) f(A) = 0.37 +/- 0.017 HetEx: 0.167 HWE(p) = **2.2e-03**	N(CC) = *72* (69.73) N(AC) = 168 (*172.54*) N(AA) = *109* (106.73) f(A) = 0.45 +/- 0.019 HetEx: -0.026 HWE(p) = 0.665	N(CC) = 156 (*156.27*) N(AC) = *481* (480.45) N(AA) = 369 (*369.27*) f(A) = 0.39 +/- 0.011 HetEx: 0.001 HWE(p) = 1.0	N(CC) = 173 (*177.74*) N(AC) = *461* (451.51) N(AA) = 282 (*286.74*) f(A) = 0.44 +/- 0.011 HetEx: 0.021 HWE(p) = 0.547
rs7975057	N(AA) = 36 (*49.06*) N(AG) = *196* (169.88) N(GG) = 134 (*147.06*) f(A) = 0.37 +/- 0.016 HetEx: 0.154 HWE(p) = **3.5e-03**	N(AA) = *74* (70.27) N(AG) = 167 (*174.45*) N(GG) = *112* (108.27) f(A) = 0.45 +/- 0.019 HetEx: -0.043 HWE(p) = 0.451	N(AA) = *160* (154.36) N(AG) = 465 (*476.2***7**) N(GG) = *373* (367.36) f(A) = 0.39 +/- 0.011 HetEx: -0.024 HWE(p) = 0.466	N(AA) = 173 (*175.07*) N(AG) = *447* (442.86) N(GG) = 278 (*280.07*) f(A) = 0.44 +/- 0.012 HetEx: 0.009 HWE(p) = 0.839
rs1545686	N(AA) = 38 (*48.60*) N(AG) = *190* (168.81) N(GG) = 136 (*146.60*) f(A) = 0.37 +/- 0.017 HetEx: 0.126 HWE(p) = **0.018**	N(AA) = *71* (69.83) N(AG) = 172 (*174.35*) N(GG) = *110* (108.83) f(A) = 0.44 +/- 0.019 HetEx: -0.013 HWE(p) = 0.829	N(AA) = 157 (*158.68*) N(AG) = *467* (463.64) N(GG) = 337 (*338.68*) f(A) = 0.41 +/- 0.011 HetEx: 0.007 HWE(p) = 0.841	N(AA) = 172 (*178.81*) N(AG) = *438* (424.38) N(GG) = 245 (*251.81*) f(A) = 0.46 +/- 0.012 HetEx: 0.032 HWE(p) = 0.371
rs1031681	N(AA) = 38 (*50.54*) N(AG) = *196* (170.93) N(GG) = 132 (*144.54*) f(A) = 0.37 +/- 0.016 HetEx: 0.147 HWE(p) = **7.0e-03**	N(AA) = 71 (*71.17*) N(AG) = *175* (174.66) N(GG) = 107 (*107.17*) f(A) = 0.45 +/- 0.019 HetEx: 0.002 HWE(p) = 1.0	N(AA) = 161 (*162.24*) N(AG) = *490* (487.52) N(GG) = 365 (*366.24*) f(A) = 0.40 +/- 0.011 HetEx: 0.005 HWE(p) = 0.896	N(AA) = 181 (*182.37*) N(AG) = *455* (452.25) N(GG) = 279 (*280.37*) f(A) = 0.45 +/- 0.012 HetEx: 0.006 HWE(p) = 0.894
rs2611279	N(AA) = 35 (*48.69*) N(AC) = *197* (169.61) N(CC) = 134 (*147.69*) f(A) = 0.36 +/- 0.016 HetEx: 0.161 HWE(p) = **2.3e-03**	N(AA) = *69* (67.62) N(AC) = 171 (*173.76*) N(CC) = *113* (111.62) f(A) = 0.44 +/- 0.019 HetEx: -0.016 HWE(p) = 0.747	N(AA) = *155* (154.69) N(AC) = 486 (*486.62*) N(CC) = *383* (382.69) f(A) = 0.39 +/- 0.011 HetEx: -0.003 HWE(p) = 1.0	N(AA) = *175* (174.48) N(AC) = 450 (*451.04*) N(CC) = *292* (291.48**)** f(A) = 0.44 +/- 0.012 HetEx: -0.002 HWE(p) = 0.947
rs7967594	N(AA) = 35 (*47.97*) N(AC) = *195* (169.06) N(CC) = 136 (*148.97*) f(G) = 0.36 +/- 0.016 HetEx: 0.153 HWE(p) = **4.5e-03**	N(AA) = *67* (66.75) N(AC) = 173 (*173.50*) N(CC) = *113* (112.75) f(G) = 0.43 +/- 0.019 HetEx: -0.003 HWE(p) = 1.0	NA	NA

N(XX): observed number of genotype XX (expected number according to the HWE), f(X): frequency of allele X, HetEx: Heterozygote excess according to the HWE, bold: bigger number of observed and expected genotypes, NA: not analysed.

Suppl. tab. 5: Results of all SNPs on 12q21.31 associated with MDD

SNPs on 12q21.31				Discovery sample						Replication sample						Combined sample				
SNP	Δ Chr. Pos. (kb)	r^2	Minor/ Major allele	N	MAF = RAF	HWE	Association with MDD			N	MAF = RAF	HWE	Association with MDD			N	MAF = RAF	HWE	Association with MDD	
						P nom.	P nom.	P corr.	OR			P nom.	P nom.	P corr.	OR			P nom.	P nom.	OR (95% C.I.)
				Co Ca All	Co Ca All	Co Ca All	Rec. Dom. Allelic	Rec. Dom. Allelic	Rec. Dom. Allelic	Co Ca All	Co Ca All	Co Ca All	Rec. Dom. Allelic	Rec. Dom. Allelic	Rec. Dom. Allelic	Co Ca All	Co Ca All	Co Ca All	Rec. Dom. Allelic	Rec. Dom. Allelic
rs10506904	46.7	0.67	G/A	366 353 719	0.42 0.51 0.47	0.02 0.75 0.43	3.5e-05 0.042 2.6e-04	1.0 1.0 1.0	2.19 0.71 1.47	987 890 1877	0.46 0.49 0.47	0.37 0.42 0.96	8.9e-03 0.159 0.019	0.12 0.88 0.23	1.31 0.90 1.15	1353 1243 2596	0.45 0.50 0.47	0.05 0.39 0.50	2.9e-05 0.050 2.2e-04	1.49 (1.24-1.80) 0.84 (0.71-1.00) 1.23 (1.10-1.37)
rs1545843	197.7	1.00	A/G	366 353 719	0.41 0.51 0.46	3.5e-04 0.14 0.48	7.4e-08 0.18 9.2e-05	**0.041** 1.0 1.0	2.84 0.80 1.52	1022 917 1939	0.45 0.49 0.47	0.92 0.98 0.89	0.018 0.013 3.6e-03	0.22 0.16 **0.049**	1.27 0.79 1.19	1388 1270 2658	0.44 0.50 0.47	0.10 0.40 0.70	5.2e-06 8.0e-03 1.4e-05	1.54 (1.28-1.86) 0.79 (0.67-0.94) 1.27 (1.14-1.42)
rs1473442	2.5	0.77	C/A	357 349 706	0.37 0.47 0.41	2.2e-03 0.66 0.29	9.9e-05 0.25 5.3e-03	1.0 1.0 1.0	2.32 0.83 1.35	1006 916 1922	0.39 0.44 0.42	1.00 0.55 0.74	0.029 3.7e-03 2.0e-03	0.32 0.051 **0.028**	1.27 0.77 1.21	1363 1265 2628	0.39 0.44 0.41	0.12 0.82 0.24	2.8e-04 3.4e-03 8.5e-05	1.46 (1.19-1.80) 0.78 (0.66-0.92) 1.25 (1.12-1.39)
rs7975057	154.5	0.76	A/G	366 353 719	0.37 0.47 0.40	3.5e-03 0.45 0.50	3.4e-05 0.17 2.0e-03	1.0 1.0 1.0	2.43 0.80 1.39	998 898 1896	0.39 0.44 0.42	0.47 0.84 0.67	0.037 1.9e-03 1.5e-03	0.39 **0.027** **0.020**	1.25 0.75 1.22	1364 1251 2615	0.39 0.44 0.41	0.42 0.86 0.81	2.5e-04 1.3e-03 3.0e-05	1.47 (1.19-1.80) 0.77 (0.65-0.90) 1.26 (1.13-1.41)
rs1545686	1.6	0.69	A/G	364 353 717	0.37 0.44 0.40	0.02 0.83 0.55	3.1e-04 0.08 2.2e-03	1.0 1.0 1.0	2.16 0.76 1.39	961 855 1816	0.41 0.46 0.43	0.84 0.37 0.50	0.021 2.0e-03 1.1e-03	0.25 **0.028** **0.015**	1.29 0.74 1.23	1325 1208 2533	0.40 0.45 0.42	0.19 0.56 0.22	3.3e-04 7.2e-04 2.5e-05	1.46 (1.19-1.80) 0.75 (0.63-0.89) 1.27 (1.14-1.42)
rs1031681	11.1	0.67	A/G	366 353 719	0.37 0.45 0.41	7.0e-03 1.00 0.27	2.8e-04 0.102 2.8e-03	1.0 1.0 1.0	2.17 0.77 1.38	1016 915 1931	0.40 0.45 0.42	0.90 0.89 0.93	0.014 6.5e-03 1.8e-03	0.17 0.088 **0.025**	1.31 0.78 1.21	1382 1268 2650	0.39 0.45 0.42	0.14 0.91 0.32	1.8e-04 2.6e-03 5.1e-05	1.47 (1.20-1.81) 0.78 (0.66-0.92) 1.25 (1.12-1.40)
rs2611279	34.8	0.65	A/C	366 353 719	0.36 0.44 0.40	2.3e-03 0.83 0.20	1.4e-04 0.194 4.8e-03	1.0 1.0 1.0	2.30 0.82 1.36	1024 917 1941	0.39 0.44 0.41	1.00 0.95 0.85	0.012 5.8e-03 1.5e-03	0.15 0.078 **0.020**	1.32 0.78 1.22	1390 1270 2660	0.38 0.44 0.41	0.14 0.82 0.47	1.1e-04 4.1e-03 8.5e-05	1.50 (1.22-1.85) 0.79 (0.67-0.93) 1.25 (1.12-1.40)
rs7967594	NA	0.64	A/C	366 353 719	0.36 0.44 0.40	4.5e-03 1.00 0.18	3.2e-04 0.14 4.7e-03	1.0 1.0 1.0	2.21 0.79 1.36	NA						NA				

Δ Chr. Pos. (kb): Distance form one SNP to the next in kilobases; r^2: pair wise r-square relative to rs1545843 in discovery controls; MAF: Minor allele frequency; RAF: MDD risk allele freque equilibrium; Co: controls; Ca: cases; Genetic models: Rec.: Recessive (AA vs. AG+GG), Dom.: Dominant (AA+AG vs. GG), Allelic: Additive allele dosage (A vs. G); P-values: nom.: nomina corrected; NA: not analysed

10 APPENDIX

10.1 Curriculum vitae

Name:	Martin A. Kohli
Adresse:	Kraepelinstrasse 14, D-80804 München, Deutschland
Geburtsjahr, -ort:	1975, Muri bei Bern, Schweiz

Ausbildungsgang

1991 – 1996	Matura (Abitur) am Real-Gymnasium Bern-Kirchenfeld, Schweiz, mit einem Schwerpunkt auf Mathematik und Naturwissenschaften.
1996 – 1998	Grundstudium der Biologie an der Universität Bern (Bachelor degree).
1998 – 2000	Zivildiensteinsatz als Pfleger am Hôspital de Lavaux in der Nähe von Lausanne (Hôpital de Lavaux, Chemin Colombaires 31, Cully, VD, CH-1096, Suisse).
2000 – 2002	Hauptstudium der Biologie an der Universität Bern. Schwerpunkte: Zoologie, Biochemie, Molekulare Zellbiologie und Genetik. Thema der Diplomarbeit: "Molekulare Mechanismen des Stop-Codon-vermittelten Abbaus von Boten-RNA (nonsense-mediated mRNA decay, NMD) in menschlichen Zellkulturen. Hauptmethodiken: Zellkultivierung, Zelltransfektion, RT-PCR, Gel-Electrophorese, RNase Protection Assays, Real-time quantitiative PCR (TaqMan-Proben, ABI system). Projektveranwortlicher: Assistenzprofessor Dr. Oliver Mühlemann am Institut für Zellbiologie (IZB). (Baltzerstrasse 4, CH-3012 Bern (www.izb.unibe. ch/index.php).
Februar, 2002	Erhalt eines Diploms in Biologie der Universität Bern (Master of Science degree).
2002 – 2003	Projektorientierte Position am Institut für Wirtschaft und Verwaltung, an der Fachhochschule Bern. Statistische Auswertung und redaktionelle Arbeiten an eGoverment/eDemocracy-Projekten. Verantwortlicher: Prof. Dieter Spahni. Dreimonatiger Zivildiensteinsatz bei „Swiss Peace", einem praxisorientierten Friedensforschungsinstitut mit Sitz in Bern (www.swisspeace.ch/).
Sep. 2003 – heute	Stipendium der Max-Planck-Gesellschaft für eine Doktorarbeit am Max-Planck-Institut für Psychiatrie in München. Doktorand in der Forschungsgruppe von Prof. Dr. Dr. Florian Holsboer und Dr. Dr.

Elisabeth Binder. (Kraepelinstrasse 2-10, D-80804 München, www.mpipsykl.mpg.de). Akademischer Betreuer: Prof. Dr. Wolfgang Wurst an der Technischen Universität München (Helmholtz Zentrum München, Institut für Entwicklungsgenetik, Ingolstädter Landstraße 1, D-85764 Neuherberg, http://www.helmholtz-muenchen.de/en/idg/).

2003 – Juni '04 Kopplungsungleichgewichts-Studien an ENU-mutierten Mauslinien mit Depressions- und Angst-bezogenen Verhaltensabnormalitäten.

2004 – 2008 Kandidatengen- und genom-weite humane Assoziationsanalyzen in unipolarer Depression, Suizidversuch, Therapie-Ansprechen auf Antidepressiva, Adipositas und Multipler Sklerose.

Aufgabenbereiche: Kandidatengen- und Markerselektion, Zusammenstellen von TaggingSNP-basierten Marker-Sets aus HapMap-Projekt-Daten für Gen-Screening- und ‚fine-mapping'-Experimente (LD-mapping). Experimentelle Planung, Qualitätsprüfung und Auswertung in Genotypisierungs- und Re-Sequenzierungsprojekten: ‚MALDI-TOF high-throughput multiplex genotyping (Sequenom), Illumina Goldengate customized und 100k / 300k whole-genome Beadchip assays' und Genotypisierungs- und Expressionsmessungen auf einem Roche Light Cycler. Re-Sequenzierung auf einem ABI 3730 DNA Analyzer.

Statistische Genetik Anwendungen als 'open-source user' von Kandidatengen- und genomweiten Assoziationsanalysen in einem Fall-Kontroll-Design, 'ordered-subset analysis' ('sequential addition of cases'), Power-Berechnungen, permutationsbasierte Korrekturen für multiples Testen, diverse bioinformatische Ansätze zur Generierung funktionaler Hypothesen zu Assoziationsbefunden.

Aug. 2008 Einreichung dieser Dissertation an der Fakultät „Wissenschaftszentrum Weihenstephan für Ernährung, Landnutzung und Umwelt" der Technischen Universität München. (Arcisstrasse 21, D-80333 München).

10.2 List of Publications

Binder, E. B. Salyakina, D. Lichtner, P. Wochnik, G. M. Ising, M. Putz, B. Papiol, S. Seaman, S. Lucae, S. **Kohli, M. A.** Nickel, T. Kunzel, H. E. Fuchs, B. Majer, M. Pfennig, A. Kern, N. Brunner, J. Modell, S. Baghai, T. Deiml, T. Zill, P. Bondy, B. Rupprecht, R. Messer, T. Kohnlein, O. Dabitz, H. Bruckl, T. Muller, N. Pfister, H. Lieb, R. Mueller, J. C. Lohmussaar, E. Strom, T. M. Bettecken, T. Meitinger, T. Uhr, M. Rein, T. Holsboer, F. Muller-Myhsok, B.
Polymorphisms in FKBP5 are associated with increased recurrence of depressive episodes and rapid response to antidepressant treatment. Nature Genetics (2004), 36 (12):1319-1325

Unschuld PG, Ising M, Erhardt A, Lucae S, Kloiber S**, Kohli M**, Salyakina D, Welt T, Kern N, Lieb R, Uhr M, Binder EB, Müller-Myhsok B, Holsboer F, Keck ME.
Polymorphisms in the Serotonin Receptor Gene HTR2A Are Associated With Quantitative Traits in Panic Disorder. Am J Med Genet Part B (2007), 144B:424–429.

Uhr M, Tontsch A, Namendorf C, Ripke S, Lucae S, Ising M, Dose T, Ebinger M, Rosenhagen M, **Kohli MA**, Salyakina D, Bettecken T, Specht M., Pütz B, Binder EB, Müller-Myhsok B, Holsboer F.
Polymorphisms in the drug transporter gene ABCB1 predict antidepressant treatment response in depression. Neuron (2008), 57: 203–209.

Unschuld PG, Ising M, Erhardt A, Lucae S, **Kohli M**, Kloiber S, Salyakina D, Thoeringer CK, Kern N, Lieb R, Uhr M, Binder EB, Müller-Myhsok B, Holsboer F, Keck ME.
Polymorphisms in the galanin gene are associated with symptom-severity in female patients suffering from panic disorder. Journal of Affective Disorders (2008), 105: 177–184.

Weber F, Fontaine B, Cournu-Rebeix I, Kroner A, Knop M, Lutz S, Müller-Sarnowski F, Uhr M, Bettecken T, **Kohli M**, Ripke S, Ising M, Rieckmann P, Brassat D, Semana G, Babron M-C, S Mrejen S, Gout C, Lyon-Caen O, Yaouanq J, Edan G, Clanet M, Holsboer F, Clerget-Darpoux F and Müller-Myhsok B.
IL2RA and IL7RA genes confer susceptibility for multiple sclerosis in two independent European populations. Genes and Immunity (2008), 9: 259–263.

Kloiber S, Ripke S, **Kohli MA**, Reppermund S, Salyakina D, Saemann P, Bettecken T, Horstmann S, Dose T, Unschuld PG, Zihl J, Muller-Myhsok B, Holsboer F, Lucae S.
Polymorphisms in Leptin are associated with resistance to antidepressant treatment and cognitive impairment in major depression. Accepted for publication by Pharmacogenomics on 3rd of Oct. 2008.

Ising M.*, Lucae S.*, Binder EB., Bettecken T., Uhr M., Ripke S., **Kohli MA.**, Hennings JM., Kloiber S., Menke A., Bondy B., Rupprecht R., Domschke K., Baune BT., Arolt V., Rush AJ., Florian Holsboer F. & Müller-Myhsok B.
A genome-wide association study points to multiple loci predicting treatment outcome in depression. Archives of General Psychiatry (2009), Epub ahead of print.

Kloiber S, **Kohli MA**, Brueckl T, Ripke S, Ising M, Uhr M, Menke A, Unschuld PG, Horstmann S, Salyakina D, Müller-Myhsok B, Binder EB, Holsboer F, Lucae S.
Variations in Tryptophan Hydroxylase 2 linked to Decreased Serotonergic Activity are associated with elevated Risk for Metabolic Syndrome in Depression. Molecular Psychiatry (2009), 1-12, Epub ahead of print.

Erhardt A, Czibere L, Roeske D, Lucae S, Unschuld PG, Ripke R, Specht M, **Kohli MA**, Kloiber S, Weber P, Deussing J, Ising M, Heck A, Ellgas A, Pfister H, Lieb R, Pütz B, Uhr M, Hohoff C, Maier W, Bandelow B, Domschke K, Krakowitzky P, Jacob C, Deckert J, Landgraf R, Bettecken T, Keck ME, Müller-Myhsok B, Holsboer F, Binder EB.

Genomewide association study in patients with panic disorder identifies transmembrane protein 132D (TMEM132D) as susceptibility gene for anxiety-related phenotypes. To be submitted to Nature Neutroscience.

Kohli MA*, Lucae S*, Schmidt M, Müller MB, Saemann P, Hoehn D, Czisch M, Salyakina D, Ripke S, Heck A, Ising M, Roeske D, Menke A, Kloiber S, Specht M, Uhr M, Bettecken T, Holsboer F, Binder EB^{+}, Müller-Myhsok B^{+}.
Variants regulating the expression of the neuronal amino acid transporter gene *SLC6A15* are associated with Major Depression. To be submitted to Nature Neuroscience.

Kohli MA, Salyakina D, Pfennig A, Lucae S, Heck A, Ising M, Horstmann S, Kloiber S, Menke A, Hennings J, Müller-Myhsok B, Holsboer F, B. Binder EB.
Genetic variants in the *NTRK2* gene are associated with life history of suicide attempts in patients of mood disorders. In review at Archives of General Psychiatry.

10.3 Lectures and Talks

July 2005	"SNPs in the Neurotrophin Tyrosine Kinase II (*NTRK2*) gene are associated with attempted suicide". Hold at the summer symposium of the Max Planck Institute for Psychiatry (MPIP).
February 2007	"Candidate-gene and whole-genome human genetic association studies in psychiatry". Hold at the Cell Biology Seminar of the Institute for Cell Biology (ICB), University of Bern.
May 2008	"A genome-wide association study in major depression reveals association of SNPs on chromosome 12q21.31". Hold at the "Freshest Data Session" of the 63rd Annual Scientific Convention & Program of the Society of Biological Psychiatry (SOBP) in Washington.
July 2008	"A genome-wide association study reveals evidence for *SLC6A15* – a neuronal amino acid transporter gene – as a susceptibility gene for major depression". Hold at the summer symposium of the Max Planck Institute for Psychiatry (MPIP).

10.4 Major conferences and workshops

Feb. 16 – 19, 2004	Workshop on "Practical Analysis of Gene Expression Data, Multiple testing problem" at the Ludwig-Maximilians University of Munich (LMU).
June 12 – 15, 2004	"European Human Genetics Conference 2004" of the European Society of Human Genetics (ESHG) in Munich.
Oct. 26 – 30, 2004	"54th Annual Meeting of the American Society of Human Genetics" in Toronto.
Dec. 13 – 14, 2004	"BCB-Workshop on Gene Annotation Analysis and Alternative Splicing" at the Berlin Center for Genome Based Bioinformatics (BCB).
Mar. 30 – 31, 2006	"NeuroNet Spring-Meeting 2006 - Recent Progress in the Genetic Analysis of Common CNS Disorders" in Bonn.
Oct. 16 – 17, 2006	"Epigenetics Workshop" at the Institute of Neuropathology at the University of Bonn (supervisor: Dr. Andreas Waha).
Oct. 28 – 31, 2006	"XIV World Congress on Psychiatric Genetics 2006" of the International Society of Psychiatric Genetics (ISPG) in Cagliari.
May 1 – 3, 2008	"63rd Annual Scientific Convention & Program" of the Society of Biological Psychiatry (SOBP) in Washington.
July 13 – 17, 2008	"XXVI CINP Congress" of the Collegium Internationale Psychopharmacologicum in Munich.