Analysis of genetic variation in microRNA-mediated regulation and the susceptibility to anxiety disorders

Margarita Muiños Gimeno

DOCTORAL THESIS UPF / 2009

THESIS DIRECTOR

Dr. Yolanda Espinosa Parrilla (Genes and Disease Department)

THESIS CO-DIRECTOR

Dr. Xavier Estivill Pallejà (Genes and Disease Department)

A mi familia

Acknowledgements

Moltíssimes gràcies a tots! Vull que sabeu que amb aquesta tesi no m'enduc només el títol de doctora, sinó que molt més important encara, m'enduc una pila d'amics i un munt de bons records i amor al cor.... NO CANVIEU!

En primer lloc, a en **Xavier Estivill** per donar-me l'oportunitat de realitzar la tesi doctoral al seu laboratori. Per ser un exemple de feina incansable i ben feta i pels seus consells i orientació. Gràcies!

A la **Yolanda Espinosa**... per tantes coses! Per ser una "jefa" atenta, rigurosa, accessible i perspicaç, però sobretot per ser una amiga. Perquè m'has ensenyat moltes coses de ciència, de la vida i de mi mateixa.... De veritat creus que importa si enlloc d'imprimir les tesis en paper net, ho faig amb uns fulls que em sobren amb unes taquetes de cafè de res?... Síiiiii! Sóc una xapusses... Perquè espero que seguim en contacte per allà on siguem!

A **Monica Guidi** por ser una hermana mayor en toda regla. Honesta, noble y algo exagerada... Porque aunque ahora mismo te parezca mentira, tienes el don de saber lo que quieres y porque cuando te desquicias eres divertidísima. Muchas gracias por compartir conmigo tu experiencia y sobretodo por estos últimos años... en los que creo que nos hemos unido mucho.

To **Birgit** for your invaluable help, modest character and gentle nature. Thank you for being so helpful and self-motivated. Finally, I would also like to thank you for your "sun-bathing obsession" that kept us outside in the terrace during our multiple conversations.

A la **Nina** per ser una mena de gurú pel Txema i per mi; per la teva seguretat al parlar que fan que et seguim com gossets, per molt estúpid que sigui el que diguis: "perdoni… truco d'aquí, de l'hotel… I és que voldria fer-li el pont a un cotxe…". Pel teu

esperit de contradicció, la teva discreció i noblesa diurna i els llargs monòlegs nocturns… Per les nostres converses infinites de gastronomia i la vida, perquè em fas riure com ningú i perquè ets la viva imatge de la dita que en el pot petit hi ha la bona confitura. Gràcies Ninota, ets una amiga entre un milió i l'única que sap escriure amb el cul!

Al **Txema** per ser com ets. Per ensenyar-me moltíssimes coses tant al pla professional (on t'admiro profundament) com al personal. Perquè vas ser la primera persona que em vas fer sentir una més del laboratori… te'n recordes? El dia que em vas regalar aquella punta groga defectuosa; la que semblava que anés despentinada…. Perquè quan dorms sues com els bebès (els malpensats que parin; que sóc una dona casada!!!), perquè em vas perdre pel bosc al Canadà però després vas accedir a baixar a camp través perquè no em posés "nirviosa"… Perquè em vas tirar la formamida a sobre i perquè quan et poses nerviós et fregues els dits compulsivament… En fi, perquè com et dic sempre (veu de bleda): "Eeeets taaaaan mooonooo!!".

A l'**Estre Saus,** la "Marquesona" del lab! Per ser polifacètica en la teva labor de científica, secretària, professora d'esplai, psicòloga i actriu. Per les nostres cerques incansables al google (font de saviesa i diversió). Perquè se't fa un rínxol-avisador als cabells quan estàs al "borde de un ataque de nervios" I per les teves contribucions valiosíssimes en els moments difícils: "un moment, amb calma, espereu, pareu… I si passem de buscar el DNI, li fem un fotocòpia i avall…?". En definitiva, per ser una persona amb un cor gegant darrera d'unes… gegants. T'estimo petita! (només t'ho dic a tu perquè ets més cursi que un "repollo").

A la **Nina** i a l'**Ester** els hi vull agrair també les ganes que tenen de voler ser sòcies meves. Ara només falta trobar el negoci que a les *senyoretes* els hi faci gràcia… perquè: "si això no té cap sentit… amb això no tenim ni per menjar… això no té res a veure amb ciència…". Caram! Nenes! No costa tant… fem un karaoke-xino-bingo-piscina i "andando"!

A la **Marta** perquè *"a partir d'aaaaaaraaaaaaaa…."* et trobaré moltíssim a faltar! Perquè m'encantes quan ets nena-dolenta i quan t'enfades perquè et fem regals d'aniversari. Pel teu admirable sentit de la justícia, el teu amor a altres cultures (i anar a veure vídeos de "suRfistes") i per fer-nos una mica de "mama" quan ho necessitem. Per saber que la cursi no sóc jo, sinó la Saus i per la teva mania de les abraçades… (que en el fons saps que m'encanten!). Maaaartaaaaa, per què no véns a posar ordre allà on jo vagi?

A tots aquells ex-p13s que sempre portaré al cor. A **Celia Cerrato**, única persona en el mundo capaz de comerse una patata cruda a rodajitas y mantener su silueta y "glamour" intactos; L'**Ester Ballana** per tenir comentari per tot, encara que no sempre els entenguem, pels atacs de riure que fan patir per la seva vida i per fer-nos de mestra al laboratori; l'**Immota**, per estar sempre pendent de tothom, per ser una font de coneixement tipus google, però de més qualitat i per estar sempre disposada a ajudar.

Al meu prefe, el **Manel**, per aguantar-nos a totes i el que és més difícil fer-nos sentir a totes especials. Marit, preferit, pare biològic, pacient i "manitas-arregla-todo" a jornada completa! Perquè ets un sant que saps gaudir amb tot…

A la **Bruni**… ai Bruneto! Per ser única en la teva espècie. Per tenir una llengua pròpia, per tallar les "*farafernàlies"* de "soca-roc", per fer "*taulex*" enlloc de taules a l'excel i per accedir a portar a les nenes als musicals de "*Brooklyn"* que són millors que els de Broadway, sens dubte i a més queden a prop dels famosíssims cafès de l'"*Strasburg"*. Ai Bruni, que sàpigues que si *te gusto no me acuerdo*!

A l'**Elisa** per fer-nos de doctora i de pallasso picantón alhora i per no tenir *"miedu"* de la vida i viure-ho tot al màxim. A **Susa** por poner la nota castiza en el laboratorio Ayyyy! Te quierooo maaa' que a mi vidaaaaaa! Y por ser una fuente de energía inagotable. Ah! Y por supuesto, por hacernos de profe de esas

coreografías de danza moderna… A **Eva** por enseñarnos a lidiar con las situaciones difíciles, tu sentido del humor y tu par de…

Als post-docs i tècnics del lab: A la "**Grata**" no et pots fer una idea del què m'has ajudat i fet riure alhora; Amb els haplotips, amb els estudis d'associació, amb la falsificació de firmes… moltes gràcies per ser l'única persona que es volia suïcidar amb mi quan va sortir el punyetero Hapmap II! Al **Lluís** i la **Lorena** per l'ajuda amb l'R i els ordinadors… per Déu… quins moments angoixants! Al **Sergi** per haver compartit amb mi durant tants anys la responsabilitat de les màquines de PCR. A la **Kelly** per ser tan organitzada, anar sempre de cara i tenir sempre ganes de passar una bona estona! A la **Sílvia**, la **Bañez** i l'**Eulàlia** per compartir passió pels microRNAs i les ganes de participar. Al **Mario**, la **Mariona** i la **Geòrgia**.

I per suposat tots aquells que han anat arribant més tard però dels que tinc molt bons records: la **Laia** i la seva parsimònia envejable; **Raquel,** sigue siempre así de dulce, eres un encanto; **Alexander** porque es una pena que hayamos coincidido poco, porque tu punto de vista me parece siempre divertidísimo y muy acertado; **Elena**, **Johanna** i **Elisabet**.

Al CEGEN: A la **Carre** per la seva eficiència a l'hora de treballar, per acceptar dormir amb mi algun dia, per la teva paciència i saber fer i perquè quan et poses les pestanyes roses estàs guapíssima. Moltes gràcies Carre per ser una companya de viatges ideal i una font d'alegria inesgotable. Al "**Carlitus**", per ajudar-me tant amb els pyros, perquè ens ho hem passat de conya intentant que aquell programot funcionés i per ser com ets. Molts petons per tu i pel teu gos, el tequieromucho. A la **Puig**, perquè ens coneixem des que jo tenia 5 anys… però fruit de l'empanamenta de les dues ens n'hem assabentat fa uns pocs mesos… per tota l'ajuda de seqüenciació i els teus comentaris encertats. A la **Cecilia** i a la **Magda**, per l'ajuda i la *Illumina*ció®. I també al **Sebas,** a la **Sílvia** i a la **Bayés**.

Al **Bar Mundial** i en especial a la Sra. Malallet per tractar-nos tan malament-genial a les nostres hores d'esbarjo. També els hi

vull agrair a totes les navalles, tallarines i braves que han deixat la vida per nosaltres.

A **Cris** y a **Mar** por recordarme que hiciera los deberes, por ordenarme las carpetas y porque sin su ayuda en el "cole" y la "uni", esta tesis nunca hubiera sido posible. A todos aquellos investigadores y biólogos que habéis compartido carrera u oficio conmigo, me habéis aguantado y me habéis divertido: **Rusi**, **Kinti**, **Ami**, **Pilarina**, **Mossa**, **Iris** y **Carme.**

A mi abuela y al "titi" por cuidarme y estar siempre de mi lado. Al "titi" por ser una fuente de buen humor en la familia y a la abuela por esa autoestima que roza la soberbia y que hace extensa a todos los suyos. A mis abuelos paternos por su cariño y a mi abuela Pilarín por despertar mi curiosidad científica desde muy pequeña: "abuela, si eres africana ¿por qué eres tan blanquita?". "Mira niña, caprichos de la genética; cada uno sale como sale". A mi padrino "el psiquiatra-intelectual" que tras intentar despertar el interés de su sobrina por el arte y la música aceptó la poca sensibilidad de ésta y le familiarizó con temas más relacionados con su profesión, la psiquiatría. A "Thomatito" por ayudarme con la tesis; por tu espíritu crítico, tu paciencia al ver que no seguía tus consejos y tu valentía para sacarnos a los dos adelante. Por ser la persona más buena que conozco, por cuidarme, consolarme y querer "tirar patatas" a los que hacen daño. A mis padres por enseñarme lo que es el espíritu de equipo, por demostrarme que la unión hace la fuerza, por vuestra confianza y por ser unos padres atípicos. A mi madre por saber escuchar venciendo su falta de paciencia, por tu mente organizada y tu carácter dedicado… que no delicado! A mi padre por enseñarme sin esfuerzo cosas tan importantes como: reírse de uno mismo, la pasión por la ciencia y el apreciar el esfuerzo de los otros… A los dos por demostrarme desde pequeña que mi opinión también cuenta y que estáis ahí para lo que necesite… ¡MUCHAS GRACIAS!

Abstract

We have investigated genetic variation in microRNA-mediated regulation as a susceptibility factor for anxiety disorders following two different approaches. We first studied two isoforms of the candidate gene *NTRK3* by re-sequencing its different 3'UTRs in patients with Panic (PD) and Obsessive Compulsive disorders (OCD) as well as controls. Two rare variants that altered microRNA-mediated regulation were identified in PD. Conversely, association of a common SNP with OCD hoarding subtype was found. Moreover, we have also studied a possible involvement of microRNAs in anxiety disorders. Consequently, we have analysed the genomic organisation and genetic variation of miRNA-containing regions to construct a panel of SNPs for association analysis. Case-control studies revealed several associations. However, it is worth remarking the associations of miR-22 and miR-488 with PD; two microRNAs for which functional assays and transcriptome analysis after microRNA overexpression showed significant repression of a subset of genes involved in physiological pathways linked to PD development.

Hem investigat la variació genètica a la regulació mediada per microRNAs com a factors de susceptibilitat pels trastorns d'ansietat seguint dues aproximacions diferents. Primer vam estudiar dues isoformes del gen candidat *NTRK3* mitjançant la reseqüenciació dels seus diferents 3'UTRs a pacients de pànic (TP), a pacients amb trastorn obsessiu compulsiu (TOC) i a controls. Dues variants rares que alteren la regulació mediada per microRNAs foren identificades per TP. D'altra banda, es trobà associació d'un SNP comú amb el subtipus acumulador de TOC. A més, també hem estudiat la possible implicació dels microRNAs als trastorns d'ansietat. Conseqüentment, hem analitzat l'organització genòmica i la variació genètica a regions que contenen microRNAs per construir un panell d'SNPs per fer anàlisis d'associació. Els estudis cas-control van revelar algunes associacions. Tanmateix, val la pena destacar les associacions del miR-22 i el miR-488 amb TP; dos microRNAs pels quals assajos funcionals i anàlisis de transcriptoma després de la seva sobreexpressió han mostrat una repressió significativa d'un grup de gens implicats en vies fisiològiques lligades al desenvolupament del TP.

Though this be madness, yet there is method in 't.

William Shakespeare, "Hamlet", Act 2 scene 2

Contents

Preface

In this introductory chapter, I am going to firstly address the topic of psychiatric disorders, the difficulty of their diagnosis and genetic study and the different approaches that have been used for this purpose. Secondly, I will focus on anxiety disorders, the group of disorders studied in this thesis. In this second section of the introduction, candidate genes for these disorders will be summarized broadly. However, emphasis will be made for one of the candidate genes – the Neurotrophin Tyrosine Kinase Receptor type 3*, NTRK3*- for which we have demonstrated post-transcriptional regulation by microRNAs. Finally, the last section of this introductory chapter consists of an extensive overview of the recently emerged "microRNA world". I will firstly present, the biogenesis and mechanisms of action of these molecules and then I will concentrate on their functions and relation to disease, insisting on their important role in the central nervous system and on their involvement in neurological disorders.

Introduction

The statistics on sanity are that one out of every four Americans is suffering from some form of mental illness. Think of your three best friends. If they're okay, then it's you.

Rita Mae Brown

1. Psychiatric disorders

1.1 Diagnosis and classification

Psychiatric disorders are psychological or behavioural patterns that an individual may present. They are manifested by breakdowns in the adaptive process expressed primarily as abnormalities of thought, feeling, and behaviour producing either distress or impairment of function that is not expected as part of normal development or culture. The recognition and understanding of psychiatric disorders has changed over time and across cultures. Definitions, assessments, and classifications of mental disorders can vary, but guideline criteria listed in the International Statistical Classification of Diseases and Related Health Problems (ICD), Diagnostic and Statistical Manual of Mental Disorders (DSM, for detailed information see Box 1) and other manuals are widely accepted by mental health professionals. In many cases there is no single accepted or consistent cause for mental disorders, although they are often explained through an interaction between the genes and the action of different environmental catalysts, such as stress (figure 1). Psychiatric disorders have been found to be common, with over a third of people in most countries reporting sufficient criteria at some point in their life. Mental health professionals diagnose individuals using different methodologies, often relying on case history and interview.

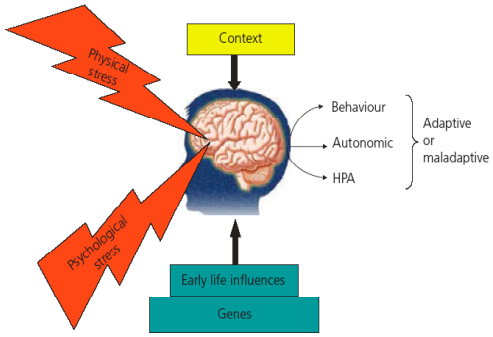


Figure 1: Features involved in the development of psychiatric disorders.

Psychotherapy and psychiatric medication are two major treatment options, as well as supportive interventions and self-help. The latest version from the Diagnostic and Statistical Manual of Mental Disorders Fourth Edition (DSM-IV) (Association, 1994) classifies psychiatric disorders into the following 13 categories: Adjustment Disorders, Anxiety Disorders, Dissociative Disorders, Eating Disorders, Impulse-Control Disorders, Mood Disorders, Sexual Disorders, Sleep Disorders, Psychotic Disorders, Sexual Dysfunctions, Somatoform Disorders, Substance Disorders and Personality Disorders. Anxiety Disorders will be further addressed in the next chapter of this thesis.

BOX 1. Diagnostic and Statistical Manual of Mental Disorders (DSM)

The Diagnostic and Statistical Manual of Mental Disorders (DSM), published by the American Psychiatric Association, provides diagnostic criteria for mental disorders. There have been five revisions since it was first published in 1952, being the the DSM-IV the last major revision (Association, 1994). This manual covers mental health disorders for both children and adults. It also lists known causes of these disorders, statistics in terms of gender, age at onset, and prognosis as well as some research concerning the optimal treatment approaches. Mental Health Professionals use this manual when working with patients in order to better understand their illness and potential treatment and to help 3rd party players (e.g., insurance) understand the needs of the patient. The DSM uses a multiaxial or multidimensional approach to diagnosing. It assesses five dimensions as described below:

Axis I: Clinical Syndromes. This is what we typically think of as diagnosis (e.g., depression, anorexia).

Axis II: Developmental Disorders and Personality Disorders. Developmental disorders include autism and mental retardation, disorders which are typically first evident in childhood. Personality disorders have more long lasting symptoms and encompass the individual's way of interacting with the world.

Axis III: Physical Conditions which play a role in the development, continuance, or exacerbation of Axis I and II Disorders. Physical conditions such as brain injury that can result in symptoms of mental illness are included here.

Axis IV: Severity of Psychosocial Stressors. Events in a person's life, such as death of a loved one, or unemployment, that can have an impact on the disorders listed in Axis I and II. These events are both listed and rated for this axis.

Axis V: Highest Level of Functioning. The clinician rates the person's level of functioning, at the present time and the highest level within the previous year.

1.2 Psychiatric genetics

Emil Kraeplin recognised the hereditary basis for psychiatric disorders at the turn of the nineteenth century. After that, the genetic influence on all major psychiatric disorders was consistently demonstrated by twin and adoption studies (Plomin et al., 1994), confirming work that had been started in the 1930s (Slater, 1936).

In fact, estimated heritabilities for bipolar disorder, schizophrenia and autism (80% to more than 90%) (Bespalova and Buxbaum, 2003; Gupta and State, 2007; Kieseppa et al., 2004; McGuffin et al., 2003; Sullivan et al., 2003) have been shown to be much higher than that of breast cancer (5% to 60%) for instance (Locatelli et al., 2004; Schildkraut et al., 1989), for which several genetic factors are now well established (Plomin, et al., 1994). However, psychiatric disorders, like other complex disorders are complicated by locus heterogeneity, imprecisely specified traits, incomplete penetrance and interaction with non-genetic factors, resulting in a low contribution of each individual risk allele (odds ratio <2) (Burmeister et al., 2008).

The completion of the draft sequence of the entire human genome (Lander et al., 2001) opened up many new resources for the study of genetic determinants of disease; this is, however, the beginning of a whole process. During the last years, large amount of genetic and genomic data have been generated including genomic sequences from multiple organisms and large genome-wide SNP genotyping data provided by projects such as HapMap (for detailed information see Box 2).

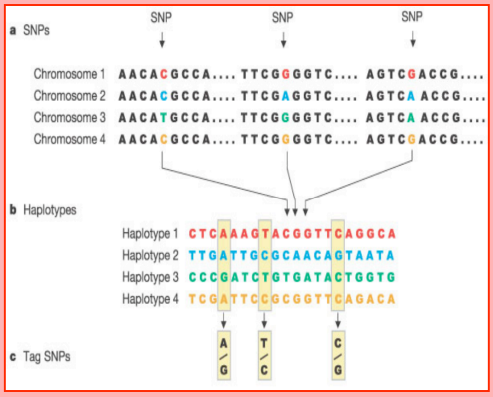
These data have become available and provide the required Linkage Disequilibrium (LD) information for custom design of SNP panels that have maximal power to capture the genetic variation in a specific genomic region of interest (Frazer, et al., 2007; Hinds et al., 2005). Description of the genomic and/or genetic variation of the human genomes is, therefore, the basic tool that geneticists need to conduct their studies.

A condition of any genetic analysis is, however, a valid and accurate phenotype, in order to homogenise the genetic background of the so-called cases (patients). Consequently, psychiatric disorders deal here with an important drawback in comparison to other diseases. As a matter of fact, there is no biochemical or physiological test available for psychiatric disorders, equivalent to, for example, the one that measures glucose level for type 2 diabetes. In contrast, psychiatric diagnosis needs to be made through clinical examination, usually by direct interview, and summarized on the basis of diagnostic criteria. These diagnoses are, as a result, highly time-consuming and albeit being rigorous and highly reproducible from centre to centre or diagnostically useful, might not reflect genetic aetiology and in consequence might not be genetically valid. Problems arise when symptoms in patients are not clear and the boundaries of the disease blur. In addition, in some cases categories overlap; some clear examples of this affirmation are the fact that a panic attack, for instance, can be part of at least two different diagnoses: panic disorder (PD) and agoraphobia. On the other hand, a psychosis (illusions or hallucinations) can be part of at least three different diagnoses, schizophrenia, bipolar disorder and psychotic depression. Only the course of the disease, during the years, will unravel the nature of the

symptoms (Burmeister et al., 2008). Without an accurate specification of the phenotype, mapping the gene/s responsible for disease susceptibility becomes difficult, if not, impossible to perform. In this sense, the success of psychiatric genetics has everything to do with nosology. In many cases, alternative phenotypes such as intermediate traits or endophenotypes, as well as unique families with rare disorders have shown to help in gene identification. Moreover, the argument can be made that for the purposes of identifying susceptibility genes, the practice of condensing phenotypic data into dichotomous or categorical diagnoses results in a loss of information. That is why some genetic factors apparently influencing behaviour or personality traits, which do not fall into specific diagnostic categories, might help in the search for susceptibility to these disorders. Therefore, the use of clinically descriptive, experimental and subjective dimensions of psycopathology is likely to change the current classification scheme in psychiatry. In fact, nowadays many psychiatric genetic studies tend, to narrow criteria: in the case of bipolar disorder for example the phenotype tested would be that the subject had documented evidence of acute mania (required for the diagnosis of bipolar I disorder) a clinical state that is strikingly similar across patients. Most studies, in summary, use a "final best" process that uses data from the interview as well as clinical records and family history. In this context, the identification of genetic risk factors might provide a better understanding of the up to now limited knowledge we have in the aetiology of psychiatric disorders (Burmeister et al., 2008; Smoller and Tsuang, 1998).

BOX 2. The HapMap project

The HapMap (www.hapmap.org) is a catalogue of SNPs that occur in human beings. It describes what these variants are, where they occur in our DNA, and how they are distributed among people within populations and among four populations from different parts of the world. By identifying most of the approximately 10 million SNPs estimated to occur commonly in the human genome, the International HapMap Project is identifying the basis for a large fraction of the genetic diversity in the human species. However, the International HapMap Project is also aiming to identify common haplotypes, as well as "tag" SNPs that uniquely identify these haplotypes. By genotyping tag SNPs researchers are able to identify the collection of haplotypes in a person's DNA. The number of tag SNPs that contain most of the information about the patterns of genetic variation is estimated to be about 300,000 to 600,000, which is far fewer than the expected 10 million common SNPs (Frazer et al., 2007). The DNA samples for the HapMap project have come from a total of 270 people: 30 trios from the Yoruba people of Ibadan, Nigeria; 45 unrelated individuals from the Tokyo area, Japan; 45 unrelated individuals from Beijing, China; 30 trios from U.S. residents with northern and western European ancestry provided by the Centre d'Etude du Polymorphisme Humain (CEPH).



The construction of the HapMap occurs in three steps. (a) SNPs are identified in DNA samples from multiple individuals. (b) Adjacent SNPs that are inherited together are compiled into "haplotypes." (c) "Tag" SNPs within haplotypes are identified that uniquely identify those haplotypes. By genotyping these tag SNPs it is possible to identify which haplotypes are present in each individual. *From www.hapmap.org.*

1.3 Genetic approaches to the study of psychiatric disorders

The most common approaches and methods used for the study of complex disorders that are also employed in psychiatric genetics will be reviewed in this section.

1.3.1 Familial aggregation

The first step in the study of a potentially genetic trait is to see if it tends to aggregate in families more than it would be expected by chance and to examine how that familial tendency is modified by the degree or type of relationship, age, or environmental factors. This is often based on case-control comparisons of family history or on twin or adoption studies. Twin studies aim to separate the effects of genes and environment. The classical twin method involves identification of twin pairs through affected members and comparison of the concordance rates between MZ (monozygotic, identical) and DZ (dizygotic, or fraternal) twins. Assuming that MZ and DZ twins share environmental factors to a similar degree but differ in their genetic similarity, this comparison allows the estimation of the heritability and environmentality of the disease. This is particularly useful in the case of psychiatric disorders where environmental factors, such as learnt behaviours in response to certain situations could play a triggering role on the disease. Adoption studies also aim to separate the effects of genes and environment, although by comparing individuals who share a common environment but have different ancestry or vice-versa (Thomas, 2004).

1.3.2 Linkage analysis

It is possible to identify the approximate chromosomal location of a major gene as a result of the phenomenon of recombination. These analyses are based on the principle that genes on different chromosomes segregate independently, so there can be no linkage between them. The second principle is that the probability of recombination between two loci in the same chromosome increases with the physical distance between them, being 1/2 the limiting value, and the value for two loci located in different chromosomes. Thus, if a genetic marker has a low recombination rate with the disease, it can be inferred that the genetic marker is close to the disease causative gene (Thomas, 2004). The linkage era (1980-2005) failed to identify any single locus that was unequivocally replicated for psychiatric disorders across multiple independent samples. The scientific community suspected that the reason for this lack of replication was the low power of individual studies. In order to address this issue, scientists organised collaborative efforts and meta-analyses. However, no trully significant or replicable results were obtained. Even so, the existing linkage results are now being used in positional candidate gene association studies and pathway analyses. Linkage studies have the disadvantage that they have low power to detect low-risk alleles, like the ones predicted to cause many psychiatric disorders (Burmeister et al., 2008).

1.3.3 Genetic association studies

Risch and Merikangas demonstrated that when genetic variants have small individual effect on risk, association studies are more powerful than linkage studies, even when properly accounting for multiple testing (Risch and Merikangas, 1996). Nonetheless,

these studies deal with the problem that they are only powerful if the risk allele is common (that is to say with a minor allele frequency > 0.05) (Burmeister et al., 2008). Association studies are based on the comparison of the genotypes at the candidate loci between cases and controls, to test whether an allele or a genotype differs significantly between cases and controls.

Genetic association studies can work either with candidate genes or, in the particular case of genome wide association studies (GWAS), with the whole genome. A candidate gene can be either biological (a gene whose function makes it conspicuous of being affected on the studied disease) or positional (a gene that maps within a chromosomal region that has been previously implicated in the disease by linkage studies).

Biological candidate genes

Association can be tested with genetic variants in biological candidate genes in a metabolic pathway that is either known or hypothesized to be connected to the studied disorder, or in genes that encode for protein target of psychoactive drugs (Thomas, 2004). A well-known example of such biological candidate is a functional null allele of aldehyde dehydrogenase 2 (*ALDH2*), an enzyme involved alcohol degrading. Reduced activity of ALDH2 leads to unpleasant reaction when drinking alcohol and therefore, protects from alcoholism (Ehrig et al., 1990). Another example involves a promoter variant in the serotonin transporter (*SLC6A4*), which is a target for many antidepressive drugs and is now considered an established risk for depression (Caspi et al., 2003).

Positional candidate genes

In a positional candidate gene approach what is tested is the possible association of SNPs located in a previously identified linkage peak where a gene maps (Thomas, 2004). Known examples that have shown evidence for association identified by this method are variants in the gamma-aminobutyric acid (GABA) receptor alpha 2 subunit gene (*GABRA2*) as risk factors for alcoholism (Edenberg et al., 2004) and the identification of a novel transcript, D-amino-oxidase activator (*DAOA*) in a gene poor, schizophrenia-linked interval where some SNPs showed association with this disorder. SNPs in the same region were also associated with bipolar disorder (Chumakov et al., 2002; Hattori et al., 2003).

Genome wide association studies (GWAS)

Association studies have been the major focus of psychiatric genetics since 1996, more recently these association studies have become genome wide (Burmeister et al., 2008; Thomas, 2004). GWAS have become possible in the last years by the development of high-throughput techniques in addition to the development of human variation catalogues, like the one provided by HapMap. The first published GWAS in psychiatrics was a study comprising 2000 cases of bipolar disorder compared to 3000 population controls and performed by the Welcome Trust Case-Control Consortium (WTCCC) (Consortium, 2007). Since then, many more GWAS in psychiatry have followed, studying schizophrenia, autism and bipolar disorders among others. However, little or no replication of these GWAS has been obtained. In contrast, each new association study has brought new loci to the surface. This fact can be explained, partly, by the fact that unlike other GWAS involving other diseases, such as

type 2 diabetes, psychiatric disorders GWAS have used smaller sample sizes (38000 samples in the type 2 diabetes study v.s. 2000 cases in WTCCC study for bipolar disorders). It remains to be seen, however, if large sample sizes might shed new light in the near future.

Transmission disequilibrium test

A particular case of association studies is the Transmission Disequilibrium test, which is based on the comparison of the parental alleles transmitted and not transmitted to a proband. Under the null hypothesis, the alternative genotypes are equally likely to have been transmitted to the case, so any deviation from this expected distribution in the case is evidence of association with that gene. This type of study offers the advantage that, using parents, it offers an appealing way of obtaining a matched control group (Thomas, 2004).

1.4 Candidate genes and pathways for psychiatric disorders

Genetic influences on psychiatric disorders, as reviewed before, have been well established. However, localization of genes responsible for these effects has proven to be extremely difficult. This is probably due to three aspects, among others: 1) Problems arising from the difficult diagnosis of Psychiatric disorders; 2) The probable multifactorial and polygenic origin of these disorders; 3) The high level of regulatory control and gene interactions to what human brain and behaviour are exposed. The search for susceptibility genes for psychiatric disorders has classically focused on neurotransmitters and members of neurotransmitter synthesis and degradation pathways. More recently, other groups of molecules have been or are beginning to be considered as good candidate genes for psychiatric disorders. Other candidates include genes involved in neurodevelopment and synaptic plasticity, since their misfunction could lead to developmental alterations of neural circuitries or to impairment of neural adaptation (Gratacos et al., 2007). In addition, regulator elements are also beginning to be studied due to their action as orchestra directors of other molecules intervening in biological processes.

Box 3: Neurotransmitters

Acetylcholine
Acetylcholine (ACh) was the first neurotransmitter identified, in 1921. It is involved in muscle contraction including heartbeat. Recent evidence suggests that it may also be critical for normal attention, memory and sleep.

Amino Acids
Certain amino acids can serve as neurotransmitters in the brain. For instance, while glutamate and aspartate act as excitatory signals, glycine and gamma-aminobutyric acid (GABA) inhibit the firing of neurons. They have been implicated in learning and memory and interestingly, GABA activity is increased by benzodiazepines and by anticonvulsant drugs.

Monoamines
Catecholamines: The most abundant catecholamines are epinephrine (adrenaline), norepinephrine (noradrenaline) and dopamine. Catecholamines control movement and regulate hormonal responses, apart from playing an important role in the arousal and reward systems, which have been implicated in the pathophysiology of addictions, anxiety and depression.
Indolamines: A common example of an indolamine is serotonin, which acts in the regulation of mood, anxiety, sleep/wake cycles and body temperature. Drugs that alter serotonin's action have relieved symptoms of depression and anxiety disorders, including obsessive-compulsive disorder.

Peptides
Three examples are endorphins, Substance P and Cholecystokinin (CCK) that has been suggested to cause nausea, anxiety, and induce a satiating effect.

Gases and single ions
Single ions, such as synaptically released zinc and gaseous molecules such as nitric oxide (NO) and carbon monoxide (CO) can also be considered neurotransmitters.

1.4.1 Genes encoding for specific elements of the neurotransmitter system

Neurotransmitters are involved in neuromodulation, which is the process by which one neuron may use several classes of neurotransmitters to regulate diverse other populations of neurons. This differs from direct synaptic transmission where one neuron directly influences a postsynaptic partner, that is to say, one neuron reaches another one. Neurons expressing certain types of neurotransmitters sometimes form distinct

systems, where activation of the system affects large volumes of the brain, called volume transmission. The effects that these neurotransmitter systems have, made the inclusion of the genes involved in them an obvious first attempt in the study of candidate genes for psychiatric disorders. There are many neurotransmitters and many different ways to classify them, Box 3 summarizes some of the different neurotransmitters and their roles. On the other hand, genes encoding for neurotransmitters are not the only genes involved in neuromodulation; neurotransmitter receptors, precursors, transporters or synthesis/degradation enzymes are also involved in these processes and are thus good candidates as well (Dorland, 1994; Neuroscience, 2005; Rang, 2003).

1.4.2 Genes involved in neurodevelopment and plasticity

This group includes many molecules involved in the development and plasticity of the central nervous system (CNS). However, we will only make especial emphasis on neurotrophic factors and hormones.

Hormones

The brain contains receptors for the thyroid hormone and the six classes of steroid hormones –estrogens, androgens, progestins, glucocorticoids, and vitamin D-. They alter the production of gene products that participate in synaptic neurotransmission as well as the structure of brain cells. Hormones are, therefore, important agents of protection and adaptation. In particular, stress and stress hormones can also alter brain function, including learning. Severe prolonged stress can cause permanent brain damage (Neuroscience, 2005).

Neurotrophic Factors

These, are small proteins that are synthesised in neurons, released locally in the brain and bind to receptors expressed by specific neurons. They are necessary for the development, function and survival of specific groups of neurons. Besides their essential role in regulating the differentiation and proliferation of neuronal precursors, these molecules have also been involved in processes that underlie memory formation and cognitive attributes such as axonal guidance mediation (Teng and Hempstead, 2004). In fact, numerous data argue in favour of the hypothesis that both reduced neurotrophic factor support and the inability of neuronal systems to exhibit appropriate plasticity could contribute to the pathogenesis of a number of psychiatric disorders including depression, eating disorders and anxiety disorders among other (Bremner et al., 2000; Duman, 2002). The action of neurotrophins and their receptors, as well as, their role in psychiatric disorders will be further addressed in future chapters of this thesis.

1.4.3 Regulator elements

Up to now, genetic studies on psychiatric diseases have been focused on functional candidate genes. However, the complexity of the CNS requires not only of a precise function of its formal components but also of an accurate gene regulation of the system. Increasing evidence indicates that genetic variation in regulatory regions could be a major contributor to phenotypic diversity in human populations (Buckland et al., 2005; Knight, 2005; Rockman and Wray, 2002).

Gene regulation is a highly complex process with multiple levels of control at different stages of transcription and translation. A

common molecular basis for much of the control of gene expression is the binding of RNA/protein factors to cis-acting regulatory nucleic acid sequences. The latter can be genomic DNA sequences found in the vicinity of a gene or even within it, or RNA transcript sequences at the level of precursor RNA or mRNA. In contrast, as the protein factors engaged in regulating gene expression are themselves encoded by sometimes distantly located genes, they have to migrate to their site of action, and so, are called trans-acting factors. Some examples for trans-acting elements are transcription factors, chromatin-remodelling complexes, or silencing elements, such as microRNAs (miRNAs) (Strachan and Read, 2004). Taking into account the importance that minimal changes in gene regulation could have in gene dosage, and gene dosage in turn, to genetic susceptibility to disease, regulator elements acting in the brain should be included in every psychiatric candidate gene classification. As this thesis is mainly focused on the role that miRNAs may play in these disorders a more in depth explanation of these small non-coding RNAs will be performed later in this document.

A neurosis is a secret that you don't know you are keeping.

Kenneth Tynan

2 Anxiety disorders: *NTRK3*

2.1 Classification of anxiety disorders

According to the DSM-IV (American Psychiatric Association, 1994) anxiety disorders include a broad category of heterogeneous disorders where the primary feature is abnormal or inappropriate anxiety. The central and unifying features of all anxiety disorders are heightened sense of arousal or fear that is episodic or continuous and may be related to exposure to a specific trigger. There are both psychological and physiological components to anxiety disorders such as worry and fear or increased heart rate and sweating, respectively. Symptoms of anxiety are also part of a normal process called the 'fight or flight' phenomenon. This means that the body is preparing itself to either fight or protect itself or to flee a dangerous situation. These symptoms become problematic when they occur without any recognisable stimulus or when the stimulus does not warrant such a reaction. Population prevalence of these disorders is 29% and the age at onset is variable, some of them starting in the early childhood. Nowadays the DSM-IV classifies anxiety disorders into: Acute Stress Disorder, Agoraphobia (with or without a history of panic disorder), Generalised Anxiety Disorder (GAD), Obsessive-Compulsive Disorder (OCD), Panic Disorder (PD, with or without Agoraphobia), Phobias (including Social Phobia) and Posttraumatic Stress Disorder (PTSD). However, categorical classification of these disorders has not been static during the years and remains still controversial (Association, 1994). Problems arise from the overlapping of phenotypes within the broader context of anxiety disorders, variable expressivity of panic and anxiety or depression, or

phenocopies within a family. On the other hand, since the diagnostic remains purely clinical, genetic studies are arduous to perform. The use of dimensional personality traits, such as shyness, behavioural inhibition and neuroticism, in order to better define an anxiety phenotype instead of clinical diagnoses has been proposed as a solution to this problem (Hettema et al., 2001; Kessler et al., 2005; Smoller and Tsuang, 1998). Some of the most relevant types of anxiety disorders for this thesis are addressed in more detail below.

2.1.1 Agoraphobia (with or without panic disorder)

Agoraphobia, like other phobias, consists of extreme anxiety and fear. Different from other phobias, however, is the generalization what occurs. Agoraphobia is the anxiety about being in places where escape might be difficult or embarrassing or in which help may not be available should a panic attack develop. It can be sub diagnosed as either 'with' or 'without' PD. Typically situations that invoke anxiety are avoided and in extreme cases, the person may never or rarely leave their home. Agoraphobia can develop out of simple phobias or it can be a result of extreme trauma, although it is often a result of numerous panic attacks such as those found in PD. The prognosis for agoraphobia is good, especially if the individual has some insight into the development of the disorder. The treatment may involve anxiety reduction techniques aimed at increasing the control a person feels over his or her anxiety and fears. Other approaches require the patients to work through their anxiety in relation to interpersonal or childhood issues (Barlow, 2004).

2.1.2 Generalised anxiety disorder (GAD)

As its name implies, GAD is evidenced by general feelings of anxiety such as mild heart palpitations, dizziness, and excessive worry. The symptoms are difficult to control by the individual, are not related to a specific event (such as in PTSD) and are not as severe as those found with PD. The categorisation of GAD into anxiety disorders has been controversial over the years (Kendler and Prescott, 2006), and some people nowadays think that it should be re-categorised as a mood disorder (Simon, 2009). Typically GAD develops over a period of time and may not be noticed until it is significant enough to cause problems with functioning. Prognosis is good for the more extreme symptoms, but those associated with underlying fears are more difficult to treat (such as excessive worry). Working through childhood issues can also be helpful (e.g., over-controlling parental styles, sexual abuse, childhood phobias) (Association, 1994). Medication and/or psychotherapy have been found to be helpful, especially therapy aimed at teaching the patient how to gain control over the symptoms.

2.1.3 Simple phobias (including social phobia)

Often a traumatic event is the precursor for a phobia, which may or may not be at the conscious level. Symptoms include either extreme anxiety and fear associated with the object or situation of avoidance. To be diagnosed, the symptoms must be disruptive to everyday functioning. Treatment is often behavioural in nature, with the therapist guiding the patient through exercises resembling the feared object or situation. Exploring underlying issues can also be beneficial. The prognosis, however, is very good if treated effectively (Association, 1994).

2.1.4 Obsessive-compulsive disorder (OCD)

Obsessive-compulsive disorder (OCD) is characterised by obsessions and compulsions. Obsessions are recurrent intrusive and unwanted thoughts that the sufferer cannot dispel. Common themes of the obsessive thoughts include thoughts that the person may cause harm to others or that harm may befall others, or thoughts that the person or others are contaminated. Other common themes are centred on the need for order, symmetry or perfection. The obsessive thoughts are associated with negative feelings, usually anxiety, but other emotions such as disgust, guilt or shame may also be experienced. As a response to these feelings generated by the obsessive thoughts, the person may perform compulsions, and performance of the compulsions temporarily decreases the negative affect. The compulsions are stereotypic, ritualised behaviours that are usually observable but which may include covert mental rituals. Common rituals include repetitive checking, washing or cleaning, or repetitive rearranging and ordering of objects. Examples of covert mental rituals include repetitive counting, praying or thinking magical statements (Gelder MG, 2001). In summary, the obsessions and compulsions are distressing, time-consuming, and often lead to impairment in occupational, scholastic, or social functioning. According to the Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition (DSM-IV) to meet criteria for OCD, the individual must report ineffective efforts to resist, neutralise, or suppress thoughts or behaviours with other thoughts or actions. In addition, the thoughts must be distinct from those associated with other anxiety disorders, and the individual must acknowledge that the thoughts are a product of his or her own mind (Association, 1994).

The disorder affects approximately 1–3% of adults (Kessler et al., 2005) and is ranked by the World Health Organization as among the ten most disabling medical conditions (Grisham et al., 2008). Although there is strong evidence that OCD has a genetic component (Hettema et al., 2001), definitive single-domain or integrative models have not yet been established (Stein, 2002). The broader OCD phenotype has been divided into subgroups that are potentially more etiologically and genetically homogenous: symmetry/order, aggressive/checking, contamination/cleaning and hoarding (Mataix-Cols et al., 2005). A more controversial fifth dimension may be also included, consisting of somatic, sexual, religious obsessions and mental rituals (‘‘pure obsessional’’) (Mataix-Cols, 2006). OCD usually begins before age 25 years and often in childhood or adolescence. Evidence for sex dimorphism of OCD clinical features has been observed. Women with OCD may have more aggressive and contamination obsessions and cleaning rituals, while men tend to report more frequently primary obsessive slowness, sexual, exactness and symmetry obsessions and odd rituals (Grisham et al., 2008).

Medication is often prescribed for individuals with OCD even though psychotherapy can also be helpful in learning ways to feel more in control, cope better with stressors, and explore the underlying issues associated with the obsessive thoughts (Association, 1994). Only serotonin reuptake inhibitor (SRI) antidepressants (and not other antidepressants) are effective in OCD. This fact has led to the development of the serotonergic hypothesis (Barr et al., 1993), which suggests that there is some abnormality (presumably a reduction of function) of the serotonergic system in OCD or that the serotonergic system is

implicated in some way in the pathophysiology of OCD. Nevertheless, serotonin depletion studies do not result in reversal of anti-obsessional drug action or exacerbation of OCD symptoms, as it would be predicted by this hypothesis (Delgado and Moreno, 1998). Therefore, it may be concluded that serotonin plays some part in the disorder, but that perhaps its role may not be primary, but secondary and modulatory. Finally, neuroimaging studies suggest abnormalities of orbito-frontal region and basal ganglia (Saxena et al., 1998).

Prognosis for this disorder has a wide range, depending upon how the individual responds to medication and how deep rooted the underlying issues are (Association, 1994).

2.1.5 Panic disorder (with or without agoraphobia)

The DSM-IV (American Psychiatric Association, 1994) defines PD as the spontaneous, unexpected occurrence of panic attacks followed by persistent concern, worry, and anxiety about having additional panic attacks. Panic attacks are defined as a discrete period of intense fear or discomfort (no more than 30 minutes) that develops abruptly and reaches a peak within 10 minutes, in which at least 4 of 13 symptom criteria are met. Some of these criteria include cardiac palpitations, sweating, feelings of choking, fear of losing control, and fear of dying. Panic attacks often mimic symptoms of physical complaints such as a heart attack or other life-threatening illnesses (Association, 1994) and typically occur spontaneously, with no apparent trigger. However, there is evidence that shows that for the majority of patients mild phobic or hypocondriacally symptoms precede the panic attacks. Therefore, it has been proposed that panic attacks are more unpredictable than unexpected (Gratacos et al., 2007).

PD may manifest with or without accompanying agoraphobia. However, agoraphobia can also occur without PD, and panic attacks can occur in the absence of PD. Comorbidity with depressive and addictive disorders is frequent as much as a lifetime prevalence rate of 50-60%. Life prevalence of PD is 1-2% (Regier et al., 1990) and twice as many women suffer from the disorder if compared to men (Weiller et al., 1998; Weissman et al., 1997). PD has a bimodal age at onset distribution, with highest incidence in late adolescence and a second peak in the mid 30s (Sansone et al., 1998). This disorder is, however, less often observed in the elderly (Gratacos et al., 2007).

PD has moderate estimated heritability rate –proportion of phenotypic variation that is explained by genetic variation- of 44% (van den Heuvel et al., 2000), and results from a metaanalysis performed on genetic epidemiology studies showed that there is a significant association between PD in the probands and PD in the first-degree relatives ($p<0.0001$) (Hettema et al., 2001). Furthermore, risk for PD increases in adult first-degree relatives, when the age at onset of the proband is under twenty age (17-fold to 6-fold), suggesting that age at onset might be useful in differentiating familial subtypes of PD (Goldstein et al., 1997). In addition, PD and agoraphobia with panic attacks were shown to be more than 5 times more frequent in monozygotic (MZ) twins than in dizygotic (DZ) twins (Torgersen, 1983). This was further corroborated by a later study which found a significantly higher concordance among MZ than DZ twins for PD (73% vs. 0%), confirming a role for genetic factors in PD, but not for spontaneous panic attacks (57% vs. 43%) (Perna et al., 1997). Agoraphobia is also thought

to be a more severe variant of PD as suggested by the fact that the risk of PD is increased among the relatives of patients with agoraphobia (8.3%) and the relatives of patients with PD (17.3%). However, while the risk for agoraphobia is also increased among the relatives of patients with agoraphobia (11.6%), it is not among the relatives of patients with PD (1.9%) (Noyes et al., 1986).

Patients suffering from PD are generally treated with Selective serotonin reuptake inhibitors (SSRIs) as first-line agents, followed by tricyclics antidepressants. Benzodiazepines can also achieve long-term control but should be reserved for patients with refractory PD. Cognitive and behavioural psychotherapy can be used alone or in addition to pharmacotherapy. The combination approach yields superior results to either single modality. Cognitive therapy helps patients understand false beliefs/distortions and provides information about panic attacks. Behavioural therapy involves sequentially greater exposure of the patient to anxiety-provoking stimuli; over time, the patient becomes desensitized to the experience. Relaxation techniques also help control patients' levels of anxiety. Respiratory training can help control hyperventilation during panic attacks. Prognosis for this disorder is very good when treatment is pursued. Left untreated, however, symptoms can worsen and agoraphobia can develop. In these cases, the individual has developed such an intense fear that leaving the safety of home feels impossible.

2.2 Neurobiology of anxiety

Biological theories on anxiety disorders suggest abnormalities of the neurobiological pathways associated with modulating normal and pathological fear and/or stress states. In fact, animal models of stress have delineated major components of the stress response (Sullivan et al., 1999). However, nowadays, imaging studies also seek to identify the unique patterns and combinations of activated or deregulated brain regions involved in certain anxiety disorders (Gorman et al., 2000).

A model that focuses on the amygdala and its interconnected structures has been proposed for anxiety disorders. The amygdala and hippocampus are important nuclei of the limbic system that regulate emotions and memory storage, respectively. The amygdala seems to play an essential role in the acquisition of conditioned fear and the expression of innate and learned fear responses (Li et al., 1996). Specifically, the model describes 2 parallel pathways carrying information to the amygdala.

Thalamic-amygdala pathway: This pathway involves projections leading from the sensory thalamus directly to the amygdala. This pathway, thought to be automatic and involuntary, operates via a single synapse, and bypasses the neocortex and therefore, higher information processing. This circuitry is involved in identifying potential threat from the content-impoverished information, which is integrated in the amygdala immediately. From here, it may trigger the fight/flight response. From a survival perspective, it is better to rapidly

respond to any potential threat than take the time to carefully evaluate threat potential.

Cortico-amygdala pathway: This second pathway involves cortical projections, which bring thinking, reasoning, evaluation, and consciousness to this otherwise primitive unit of processing. The second pathway sacrifices speed of transmission in favour of more complex processing of emotional stimuli. Additionally, experiments demonstrated that projections from the cortex to the amygdala may be addressed through the hippocampus, where elaboration of contextual information, interpretation of the environments associated with conditioned responses, and storage of information is performed (Li et al., 1996; Rogan et al., 1997). This information processed in the sensory cortex may go directly to the amygdala or via the sensory thalamus to the amygdala.

Sensory information (coming either from the Thalamic-amygdala pathway or the Cortico-amygdala pathway) enters the lateral amygdala, from which processed information passes to the central nucleus. The central nucleus, then, acts as the central component of the fear neural circuitry and projects to multiple brain systems involved in the physiologic and behavioural responses to fear.

Projections from the amygdala to the hypothalamus take 2 primary routes: 1) the lateral hypothalamus, leading to sympathetic activation; and 2) the paraventricular nucleus, leading to activation of the hypothalamus-pituitary-adrenal axis. Projections to the lateral hypothalamus that activate the sympathetic nervous system induce the release of stress

hormones, such as corticotropin-releasing factor (CRF) (Gorman et al., 2000).

The central nucleus of the amygdala also projects to the periaquaductal grey (PAG), which initiates the body's endogenous opioid response as well as defensive actions (eg, freezing behaviour) (Gorman et al., 2000). Direct stimulation of the PAG in humans with seizure disorder actually leads to a syndrome that looks very much like PD (Graeff et al., 1993).

The amygdala also innervates the locus ceruleus (LC), a brain nucleus that is also believed to be a key element in mediation of fear responses. Norepinephrine (NE) neurons originating in the LC are activated during stress and send projections to cortical regions, the hippocampus, the PAG, the hypothalamus, the thalamus and also the amygdala. There is much evidence that supports that release of CRF, mediated by the amygdala, and the release of NE (mediated by the LC) are subject to cross-regulation. CRF neurons in the paraventricular nucleus of the hypothalamus (PVN) also project to the NE neurons of the brain stem, such as those in the LC, which reciprocally project back to the PVN, linking what are thought to be the two main central constituents of the stress system (figure 2) (Sullivan et al., 1999).

The serotonergic system also appears to be involved in anxiety disorders. Some evidences to this statement are that, for instance, the selective serotonin reuptake inhibitors (SSRIs) are effective in the treatment of some of these disorders (van der Linden et al., 2000). Moreover, increased serotonergic function is associated with dominant status in nonhuman primates

(Raleigh et al., 1991). It is thought that these serotonergic circuits modulate by inhibition the amygdala fear pathways through the rostral raphe nucleus (Tancer, 1993). More concretely, it has been hypothesised that hypofunction of serotoninergic neurons arising from the rostral raphe nucleus may result in a lack of inhibitory effect on the putative panic or obsessive pathways in the brain (Gorman et al., 2000; Neuroscience, 2005).

There are several lines of evidence that show that dopamine is involved in Social Phobia. The density of striatal dopamine reuptake site was found to be markedly decreased in patients suffering from this disorder. It has been also demonstrated that the administration of dopamine blockers results in an increase in social anxiety symptoms (Tiihonen et al., 1997). Other research groups have found that individuals with tic disorders who are treated with dopamine blockers may develop social anxiety symptoms (Schneier et al., 2000). In fact, anxiety disorders, including social phobia, occur in up to 40% of patients with Parkinson's disease. This is a higher rate than that in comparable disorders (Richard et al., 1996).

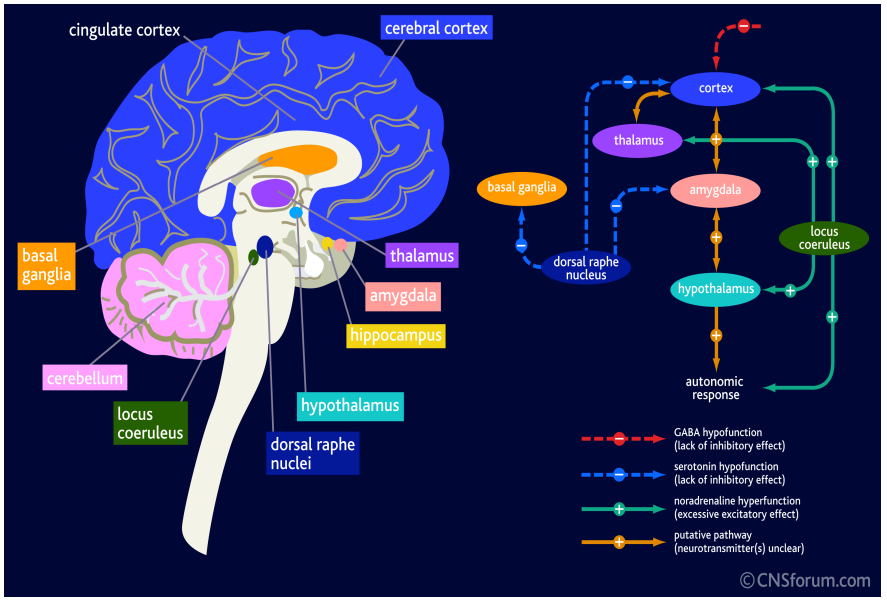


Figure 2: Neurobiology of anxiety. *From CNS forum (www.cnsforum.com).*

2.3 Anxiety disorder: main candidate genes

In anxiety disorders, the linkage era (1980-2005) failed to identify a single locus that was able to replicate across multiple independent samples (Burmeister et al., 2008). The exception came from two linkage studies on obsessive compulsive disorder that identified a peak in the same area of chromosome 9 (Hanna et al., 2002; Willour et al., 2004). However, a third larger study failed to confirm this region (Shugart et al., 2006). The existing linkage results, however, have been joined to biological data in order to select for good candidate genes for anxiety disorders. Usually the biological data used for the selection of candidate genes becomes from animal models, clinical case observations, drug efficacy in their treatment and biological plausibility; being the latter extracted from the several non–mutually exclusive

theories, that have been explained in the last section and that have been proposed to explain the aetiology of anxiety disorders. In this chapter we will present a summary of the most relevant genes that have presented a positive association with anxiety disorders.

Serotonin receptors

As explained before, serotonergic pathways have been involved in the pathogenesis of anxiety disorders, mainly because of the observation that patients with anxiety disorders respond well to serotonergic medications and because the occurrence of panic attacks has been reported after administration of serotonergic agonists (Bell and Nutt, 1998; Maron and Shlik, 2006).

One of the serotonin receptors that have been associated with PD is the 5HT-2A receptor (*HTR2A*), in fact, patients suffering from PD do not respond clinically to its antagonists, which may even exacerbate some of the symptoms (Fehr et al., 2000). Significant association was reported for this gene in a study conducted in a Japanese population. This association was shown to increase in the group of patients with PD with agoraphobia (Inada et al., 2003). More recently, 15 SNPs within *HTR2A* were genotyped in patients suffering from PD and a control sample. An association was found between rs2296972 and the severity of symptoms. This SNP was shown to form part of a haplotype - rs3742278, rs2296972, and rs277029 - which may be associated with higher susceptibility for PD (Unschuld et al., 2007).

Several studies have shown association between PD and another serotonin receptor, the serotonin 1A receptor (*HTR1A*). A

significant association with the -1019C–G polymorphism of *HTR1A* and PD phenotypes (Maron et al., 2005; Rothe et al., 2004) or with panic attacks in other psychiatric disorders has been reported (Huang et al., 2004). However, controversy also exists and other studies have been unable to confirm this association (Hettema et al., 2008; Strug et al., 2008). On the other hand, biologic investigations also provide mixed support for the role of this polymorphism in PD pathophysiology. For instance, patients with symptomatic PD were found to have a significant reduction of 5-HT1A receptors in pre- and various post-synaptic brain regions (Nash et al., 2008; Neumeister et al., 2004), whereas patients in remission after antidepressant treatment demonstrated a degree of normalization in post-synaptic 5-HT1A receptor functioning (Nash et al., 2008). In accordance with this study, another group showed that the -1019C–G variant produces a significant lower activation of the receptor in several brain regions, in PD patients homozygous for the G high-risk allele during processing of anxiety-related facial expressions (Domschke et al., 2006). However, an earlier study postulated that the -1019 G allele results in impaired repression of the 5-HTR1A gene, leading to elevated levels of 5-HTR1A autoreceptor and inhibition of the basal raphe neuronal activity (Lemonde et al., 2003). This effect has been demonstrated in both healthy adults and patients with depression carrying the –1019G allele (Parsey et al., 2006); but not by another neuroimaging study (David et al., 2005).

Finally, the serotonin receptor 2C (*HTR2C*) has been also analysed as a candidate gene for PD and a nominal association with for this gene was found in an Estonian population (Maron et al., 2005).

Serotonin transporters

Interestingly, another gene of the serotonergic pathway that has been positively associated with anxiety disorders is the serotonin transporter gene, *SLC6A4*. The short variant of a 44bp insertion/deletion polymorphism in the promoter region of this gene was reported to reduce its transcriptional activity. Association studies in two independent samples revealed that this polymorphism accounts for 3-4% of total variation and 7-9% percent of inherited variance in anxiety-related personality traits in individuals as well as sibships (Lesch et al., 1996). Nonetheless, no differences in allelic frequencies or in the genotypes of this polymorphism were observed between cases and controls in posterior studies (Deckert et al., 1997; Ishiguro et al., 1997; Matsushita et al., 1997). A posterior meta-analysis on all the association studies for *SLC6A4*, suggested that the association between *SLC6A4* and trait anxiety albeit small was real (Schinka et al., 2004). On the other hand, associations for other members of the serotonin transporter genes, have also been identified. For instance, two SNPs on *SLC6A2* showed to be associated with PD in a group of patients without Agoraphobia (Lee et al., 2005) and five out of twenty studied SNPs on *SLC6A1* showed nominal association with PD (Thoeringer et al., 2009).

Cholecystokinin *(CCK)* and receptors

CCK is the most abundant neuropeptide in the mammalian brain, which if administered intravenously induces panic attacks in 50% of the probands (Bradwejn et al., 1991). A short tandem repeat (STR) in the 5'UTR (5' untranslated region) of the gene has been found to be strongly associated with PD. However, the mechanisms underlying this association remain unknown

(Hansen et al., 2000). Regarding *CCK* receptors, *CCKAR* has been associated with PD twice (Maron et al., 2005; Miyasaka et al., 2004); nevertheless, this association had not been detected two other times (Ise et al., 2003; Kennedy et al., 1999). Similarly, for *CCKBR* mixed evidence for association has been found. Two case control studies found association of a repeat polymorphism in this gene with PD (Hosing et al., 2004; Kennedy et al., 1999), however this association could not be found in a family-based study (Hamilton et al., 2001).

Catechol-O-methyltransferase *(COMT)*

COMT, that is widely expressed in the brain, modulates one of the major degrading pathways of catecholaminergic neurotransmitters. This enzyme is one of the most promising candidate genes for PD: in humans a G to A transition that results in a Valine to Methionine change at the protein level was shown to decrease three to four-fold the activity of this enzyme (Lachman et al., 1996). After that, Woo et al, found a prevalence of 19% of the low activity allele (Metionine-allele) in PD patients in comparison to the prevalence of 2% in the control group ($p=0.005$). In addition the frequency of the Met/Met genotype was also significantly higher in patients with PD and it provided carriers with a worse response to treatment (Woo et al., 2002). More evidence for these associations has been obtained from linkage and transmission disequilibrium test analyses (Hamilton et al., 2002).

In OCD, two studies have found a possible association with the *COMT* gene, which appears to be gender-related, being present only in males (Kessler et al., 2005; Schindler et al., 2000). Other studies have found different results and a meta-analysis

conducted on data from the different studies showed little evidence to support an association between OCD and the *COMT* gene (Azzam and Mathews, 2003).

Corticotropin releasing hormone receptor 1 (*CRHR1*)

Implication of *CRHR1* in the development of PD mainly comes from the observation that mice lacking these receptors in limbic regions show reduced anxiety-related behaviour (Shimizu et al., 2004). Moreover, several SNPs in this gene region showed nominal associations with PD, the most important of which was with rs242937 (Keck et al., 2008).

Monoamine Oxidase A *(MAOA)*

This gene encodes for monoamine oxidase A, an enzyme that degrades catechol- and indolamine neurotransmitters. Monoamine oxidase inhibitors are one of the major classes of drugs prescribed for the treatment of anxiety and mood disorders, even though they are last line treatment due to risk of the drug's interaction with diet or other drugs. A variable tandem repeat (VNTR) was found in the promoter region of this gene, being the longer alleles more active than the shorter. An excess of high activity alleles was identified in female patients of PD ($p<0.001$) (Deckert et al., 1999). In addition, the high activity alleles were afterwards associated with suicidal behaviour in bipolar and major depression patients (Ho et al., 2000; Schulze et al., 2000).

Adenosine A2A receptor *(ADORA2A)*

Adenosine has been long known to be a neuromodulator, with both inhibitory and excitatory functions (Moreau and Huber, 1999). Even though it is not stored or released as a classical

neurotransmitter, it is involved the fine-tuning of other neuromodulators (Sebastiao and Ribeiro, 2000). All four of the known adenosine receptors are expressed in the brain. In particular, *ADORA2A* has been implicated in a number of behaviours including locomotion, anxiety, aggression, reward, nociception sleep, seizures, psychotic-like behaviours, and aging, according to animal work. These observations have pointed to this molecule as a candidate gene for Parkinson's disease, schizophrenia, anxiety disorders, and Alzheimer's disease (Moreau and Huber, 1999). This interest has been intensified in the case of PD where caffeine has been shown to be able to produce panic-like symptoms in humans. Caffeine is an antagonist for the four type of adenosine receptors, however, its action is more potent with the ADORA2A receptor (Boulenger et al., 1984). A mutational screening comprising the coding regions of *ADORA1* and *ADORA2A* reported a significant association of the *ADORA2A* gene with PD. The silent polymorphism identified in exon 2 was reported to have a higher prevalence of the 1083T allele (p=0.01) and 1083T/T genotype (p=0.024) in PD patients (Deckert et al., 1998). Even though this association could not be later replicated in a Japanese population (Yamada et al., 2001) linkage analysis revealed elevated LOD scores for 1083C/T (Hamilton et al., 2004).

Regulator of G protein signalling (RGS)

RGS proteins are regulatory and structural components of G protein-coupled receptor complexes. They accelerate transit through the cycle of GTP binding and hydrolysis to GDP, thereby terminating signal transduction, but paradoxically, also accelerate receptor-stimulated activation. Interestingly, the expression of *RGS2* has been demonstrated to be a quantitative

trait (Yalcin et al., 2004) and Rgs2 knock-out mice have been shown to have a more anxious behaviour than their wild-type counterparts (Oliveira-Dos-Santos et al., 2000).

Genetic variations in two RGS genes -*RGS2* and *RGS7*- were associated with PD in a German sample (Hohoff et al., 2008; Leygraf et al., 2006). On the other hand, the C allele of SNP rs4606 in *RGS2* has been also associated with GAD in an American sample of 607 adults exposed to the 2004 Florida Hurricanes (Koenen et al., 2009), even though it has not been found to be associated with PD in a sample of European Americans (Strug et al., 2008). Remarkably, both associations described by Koenen *et al.* and Leygraf *et al.* were confined to a haplotype within *RGS2* 3'UTR (Koenen et al., 2009; Leygraf et al., 2006).

Brain Derived Neurotrophic Factor *(BDNF)*

Alterations in brain maturational and developmental processes, together with plastic reorganisation in response to experience may contribute to the altered processing and storage of emotional information typical to anxiety disorders. *BDNF,* being a neurotrophin, could be involved in these processes. In fact, *BDNF* has been proposed to participate in serotonin-promoted neuronal plasticity during treatment of mood and anxiety disorders. The evaluation of *BDNF* in genetic association studies has revealed a significant link between a non-conservative Val66Met polymorphism and anxiety, major depression, bipolar disorder and smoking addiction (Hwang et al., 2006; Jiang et al., 2005; Lang et al., 2007; Neves-Pereira et al., 2002; Sklar et al., 2002). More information on the physiological roles of *BDNF,* as

well as of other neurotrophins and their receptors will be provided in section 2.3.

Neurotrophin Tyrosine Kinase Receptor type *3 (NTRK3)*

NTRK3, which will be further addressed in its own section, has also been associated with anxiety disorders. Firstly, a duplication in the genomic region where this receptor is encoded (DUP25) was identified in patients with PD and phobic disorders (Gratacos et al., 2001). Moreover, a SNP in the 5'UTR of this gene was shown to be moderately significantly associated with PD, ($p = 0.02$) suggesting a tendency to heterozygosity in PD subjects (Armengol et al., 2002).

2.4 Neurotrophins and their receptors: *NTRK3*

Altered processing and storage of emotional information like the one seen in PD might be explained by a defective brain maturation and development or deregulated plastic reorganization in response to experience. Due to their participation in the development and function of the nervous system, neurotrophins and their receptors are also considered good candidates for anxiety disorders (Gratacos et al., 2007). Moreover, it has been described that levels of neurotrophic factors in the hippocampus and amygdala correlate with anxiety and fear-related behaviour in mice (Yee et al., 2007; Zhu et al., 2006).

In mammals, the neurotrophin family consists of four well-known secreted proteins: Nerve growth factor (NGF), brain derived neurotrophic factor (BDNF), neurotrophin-3 (NT3) and neurotrophin 4/5 (NT4/5). They play essential roles in the nervous system, the nature of which depends –among other factors- on the type of neurotrophin and the expression pattern of the receptors involved (Reichardt, 2006). Neurotrophins bind to two types of plasma membrane receptors -Tropomyosine-

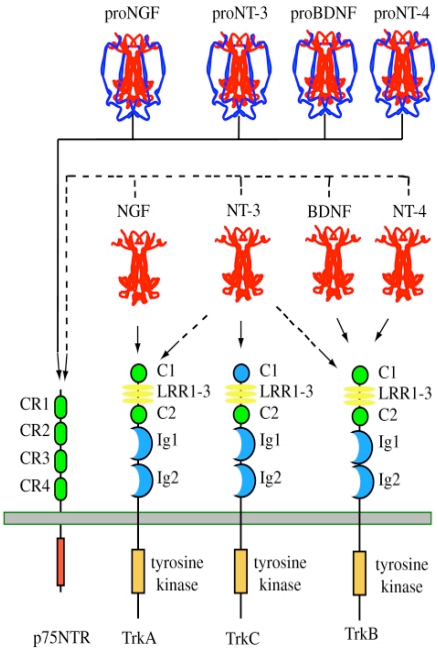


Figure 3. Major interactions of four mammalian neurotrophins. Continuous line: high-affinity binding. Dashed line: low-affinity binding. *Reichardt, 2006.*

related kinase receptors (TRK) and the p75 pan-neurotrophin receptor- to mediate their actions in the central and peripheral nervous system. Trk receptors belong to the family of receptor tyrosine kinases, and three Trk genes have been identified in mammals: *TrkA* (also called *NTRK1*, neurotrophic tyrosine kinase receptor type 1), *TrkB* (also called *NTRK2*, neurotrophic tyrosine kinase receptor type 2) and *TrkC* (also called *NTRK3*, neurotrophic tyrosine kinase receptor type 3). Due to historical reasons Trk is the most commonly used alias; yet *NTRK* is the official gene symbol approved by the HUGO Gene Nomenclature Committee (HGNC) and it will be mostly used for neurotrophin receptor genes and proteins throughout this thesis. While all neurotophins may bind the p75 receptor, each of them has a preference for one of the NTRK receptors. NGF is the preferred ligand for NTRK1, BDNF and NT4/5 are preferred for NTRK2 and NT3 for NTRK3 (Barbacid, 1994). However, cross-interaction between neurotrophins and NTRK receptors may also occur, and for instance, NT3 binds predominantly to NTRK3 even though it can also activate NTRK1 and NTRK2, albeit with lower efficiency (figure 3). When neurotrophins bind to their NTRK receptors, these homodimerise and autophosphorilate in their intracellular tyrosine residues activating different pathways such as PhosphatylInositol 3-kinase (PI3K)/Akt, Mitogen Activated Protein Kinase (MAPK) and Phospholipase C γ (PLC-γ) pathways; that lead to the regulation of proliferation and differentiation of neuronal precursors and survival of neurons in the adult nervous system (Blum and Konnerth, 2005). On the other hand, the interaction between neurotrophins and the p75NTR receptor are critical in the government of death and survival decisions in neural and non-neural cells (figure 3). Furthermore, additional actions for NTRK receptors have been revealed more recently

such as: axon and dendrite growth, axonal guidance mediation, synaptic plasticity and injury protection in neurons, actions that underlie memory formation and cognitive attributes (Teng and Hempstead, 2004).

Neurotrophins are synthesised as pre/proneurotrophin precursors of ~30kDa that are translated from a single coding exon. These precursors can be glycosilated in the N-terminal pro-domain and they dimerise forming ~60kDa species. They are further processed to yield mature C-terminal proteins of ~13.5 KDa that are also secreted as ~28KDa dimers, considered as the active forms. Nonetheless, it is now thought that these forms may not be the only biologically active neurotrophin peptides. After the discovery that pro-NGF is a high affinity ligand for p75NTR, and taking into account the fact that proneurotrophins contain conserved philogenetically cleavage sites, it is now postulated that proneurotrophins may also have specific functions as important as those established for the "mature" neurotrophin forms (Lessmann et al., 2003; Teng and Hempstead, 2004). In fact, it is nowadays thought that while mature forms activate preferentially NTRK receptors to promote survival, proneurotrophins are high affinity ligands that preferentially activate p75NTR receptor to mediate apoptosis. Supporting this idea, it has been shown that a point mutation in the pro-region of BDNF is associated with memory deficits and abnormal hippocampal function in humans (Blum and Konnerth, 2005).

At the genomic level, there is substantial evidence that all neurotrophins arose through successive duplications of a common gene of an ancestral chordate; therefore, they are

similar in sequence and structure (Huang and Reichardt, 2001). They all consist of several short exons, most of them untranslatable, dispersed in the 5' upstream region of the last encoding exon. Albeit all of these similarities, all neurotrophins have a different regional distribution, cellular localization, and synthesis in each developmental stage between one another.

2.4.1 Neurotrophin receptors

Neurotrophins have shown to directly bind and dimerise NTRK receptors. NTRK receptors consist broadly of 3 domains: the extracellular domain, the transmembrane domain and the cytoplasmic domain. While the extracellular part contains two C2-type immunoglobulin-like domain, two cysteine clusters and a leucine-rich motif, and is responsible for the interaction with neurotrophins, the cytoplasmic domain is important for kinase activity. The existence of several splicing isoforms of *NTRK2* and *NTRK3* differing in the cytoplasmic domain has been described. Some of them lack the tyrosine kinase domains, and therefore give rise to non-catalytic isoforms. NTRK receptors contain 10 evolutionary conserved tyrosines in their cytoplasmic domains, of which three are present in the autoregulatory loop of the kinase domain that controls tyrosine kinase activity. Phosphorylation of these residues further activates the receptor. Phosphorylation of the other tyrosine residues promotes signalling by creating docking sites for adapter proteins containing phosphotyrosine-binding (Huang and Reichardt, 2001; Huang and Reichardt, 2003).

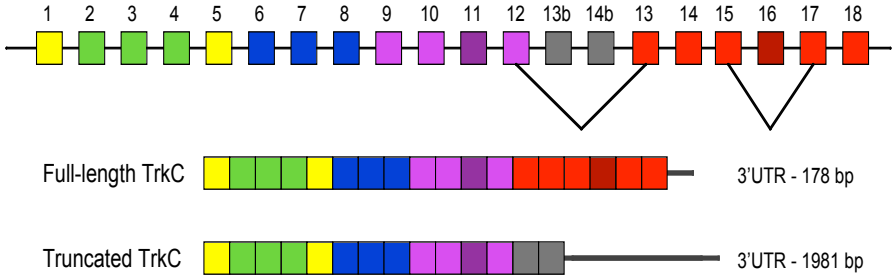


Figure 4: Schematic representation of the human *NTRK3* gene with the alternative splicing patterns. Colors represent the distribution of exons across the receptor functional domains: cystein clusters (yellow), leucine-rich motifs (green), immunoglobulin-like domains (blue), transmembrane (dark purple) and juxtamembrane (light purple) domains, tyrosine kinase domain (red) of the full-length isoform with the alternatively spliced 14aa insert (exon 16) and intracellular domain of the truncated receptor (grey). *Kindly provided by Monica Guidi.*

2.4.2 The neurotrophic tyrosine kinase receptor type 3 gene, *NTRK3*

Human *NTRK3* gene is located in chr15: 86,220,992-86,600,665 (human assembly hg18) and expands ~380kb of genomic DNA. The *NTRK3* gene locus encodes for at least four splice variants receptor isoforms in humans, with two isoforms being more abundantly expressed: an isoform coding for a full length catalytic isoform and another coding for a truncated isoform. The full-length isoform is encoded by a transcript of 19 exons, which give rise to a protein of about 150kDa that contains a fully active kinase domain. The transcript coding for the truncated, non-catalytic form lacks the last 6 exons at the 3' end and therefore lacks the kinase domain. Instead of this 3' region it has a different 3'end and an extra exon and encodes for a protein of about 60 kDa. A third isoform has also been well

characterised: a full-length isoform with an insertion of 14 amino acids in the tyrosine kinase domain, adjacent to the autoregulatory activation loop (figure 4) (Shelton et al., 1995). This isoform is less efficient at ligand-induced autophosphorylation and kinase activity in relation to the canonical full-length form. As it is less abundant than the other two isoforms and shows reduced signalling potential it is thought that its function is mainly regulatory (Tsoulfas et al., 1996). The expression patterns of the full-length and truncated isoforms in the nervous system have been analyzed in mice (Menn et al., 1998), where both transcripts have been detected during embryonic development, starting at day E11.5 and increasingly in later stages. In the embryo, full-length and truncated *NTRK3* transcripts are coexpressed in several regions, like the telencephalon, mesencephalon, rhomboencephalon, spinal chord and in neural crest derivatives in the peripheral nervous system. On the other hand, only the catalytic form is present in the diencephalon. In adults the two isoforms are codistributed in many brain structures (cerebral cortex, hippocampus, cerebellum and olfactory bulb) but the truncated form is totally absent in the thalamus and hypothalamus. The relative expression and subcellular localization of the two isoforms varies throughout neural development, as demonstrated by immunohistochemistry and immunofluorescence experiments. For instance, in differentiating neurons only catalytic receptors are expressed in the growth cone, whereas extending neurites express both variants; in mature neurons the dendritic compartment shows a predominance of truncated receptors but in the axonal compartment both isoforms coexist. This has raised the hypothesis that the catalytic isoform may mediate axon guidance whereas the truncated receptor may play a role

in the maintenance of the differentiated state, especially of the dendritic arborization, being possibly involved in short-term and long-term synaptic plasticity (Menn et al., 2000). Transfection of both isoforms into primary neurons has shown that different ratios of catalytic and non-catalytic receptors are associated with different axonal morphology: the overexpression of the full-length isoform increases the formation of axonal primary processes, while the overexpression of the truncated isoform reduces it and shifts the distribution of branch points to the proximal region of axons (Ichinose and Snider, 2000). In fact, mice overexpressing truncated *NTRK3* die in the first postnatal day and show severe developmental defects in the peripheral nervous system and in the heart (Palko et al., 1999). Moreover, *NTRK3*-deficient mice show as well a high proportion of sensory neuron loss, which are more severe in mice lacking all *NTRK3* isoforms than in those carrying only the kinase-negative mutation, implicating additional important functions for the truncated receptor (Liebl et al., 1997). The high level of sequence conservation of the intracellular domains of truncated receptors across species actually supports the hypothesis of other functions, such as the interaction with adaptor proteins and the activation of specific signalling pathways.

However, while it had been well described that the full-length NTRK3 receptor is involved in neuronal survival signalling, the main function attributed to the kinase-deficient truncated isoform had only been inhibition of the kinase active isoform, by a dominant negative effect or by a ligand sequestering mechanism. More recently a new pathway activated only by the truncated but not by the full-length receptor has been proposed (Esteban et al., 2006). In this newly described pathway,

truncated NTRK3 interacts with scaffold protein tamalin after NT3 binding, which activates a signalling cascade that is involved in membrane trafficking and actin reorganization through Rac1 GTPase and successive activation of adenosine diphosphate-ribosylation factor 6 (Arf6) (figure 5). It is hypothesised that this pathway may act in a similar way to that activated by metabotropic glutamate receptors, which play an important role in neural development and neuronal plasticity through a mechanism involved in membrane ruffling and the formation of cellular protrusions by Arf6 translocation to the membrane. These discoveries give a new role for NTRK3 truncated isoform and offer a possible explanation for the abundant expression of the truncated isoform during development (Esteban et al., 2006).

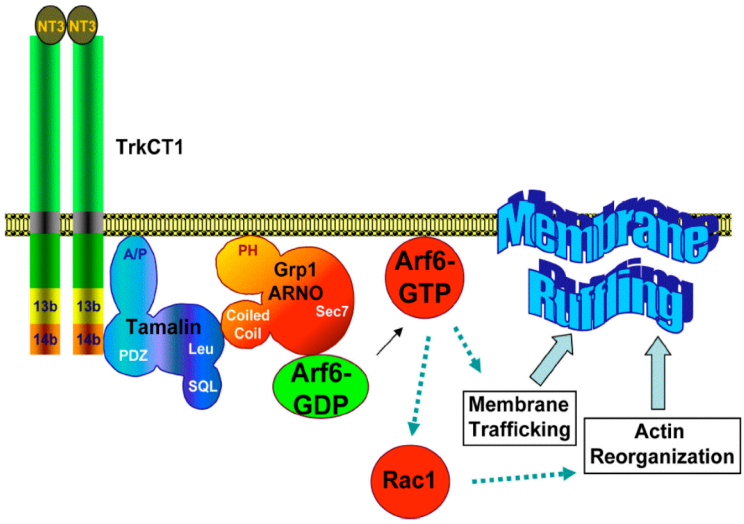


Figure 5: Schematic representation of the newly identified signaling pathway activated by NT3 through the truncated TrkC (NTRK3) receptor. *Esteban et al., 2006.*

On the other hand, evidence from a study that has examined *NTRK3* protein and mRNA levels in human prefrontal cortex post-mortem samples, from individuals ranging from 1 month to 86 years of age, indicates that levels of truncated NTRK3 protein increase during post-natal life, while those of the full-length protein remain relatively low and constant throughout development. In contrast to this increase in the levels of truncated NTRK3 protein, truncated *NTRK3* mRNA levels are quite stable throughout post-natal life, decreasing only in aged people. The fact that, at least for some of the *NTRK3* isoforms, levels of mRNA and protein are discordant during the human lifespan, suggests that post-transcriptional regulation may be contributing to the regulation of these levels throughout human lifespan (Beltaifa et al., 2005). Importance of post-transcriptional gene regulation is increasingly recognised and it has been demonstrated that miRNAs confer a novel layer to this regulation, being implicated in the development of the nervous system (Kosik, 2006). It has been observed that genes that are more prompt to miRNA regulation have on average longer 3'UTRs with a significantly higher density of target sites per kb of 3' UTR sequence. Reciprocally, genes with few target sites tend to have shorter 3'UTRs. This indicates that 3'UTRs have been under selective pressure to acquire, maintain, avoid or eliminate miRNA target sites (Stark et al., 2005). Both *NTRK3* isoforms, which have different 3'UTRs, contain putative target sites for miRNAs. Interestingly, the 3'UTR for the truncated isoform, which is considered to be long (1981 bp), has a higher number of target sites than the short 3'UTR of the full-length isoform. This observation joined to the fact that its mRNA and protein levels do not correlate during human lifespan supports the hypothesis that the two *NTRK3* isoforms may be

differentially regulated by miRNAs and makes the study of *NTRK3* 3' ends extremely important. In this context interesting results have recently emerged from a study focusing on miRNA expression in a neuroblastoma cell line differentiated with retinoic acid. This work demonstrates that miR-9 and miR-125a/b are induced upon retinoic acid treatment and are able to regulate the expression of the truncated isoform of the NTRK3 receptor (Laneve et al., 2007).

If you think you are too small to be effective, you have never been in bed with a mosquito.

Betty Reese

3. Gene regulation: microRNAs

The human genome contains about 25,000 genes. The expression of each of these individual genes needs to be appropriately orchestrated to suit the function and environment of each cell at every moment or condition. Gene expression is controlled transcriptionally and post-transcriptionally (Table 1).

As a first step, cells may regulate transcription by means of structural regulation, for instance by the control of histone packaging. Modifications on histones ''tails'' modulate how tightly the histones are packaged and influence the accessibility to transcription factors of the DNA segments that are associated with them. Another type of regulation is DNA methylation, a common method of gene silencing by which the enzyme methyltransferase typically methylates cytosine nucleotides in CpG dinocleotide sequences (also called "CpG islands" when densely clustered). However, the most common mechanism to control gene expression is to alter the availability, localization, quantity, or activity of transcription factors.

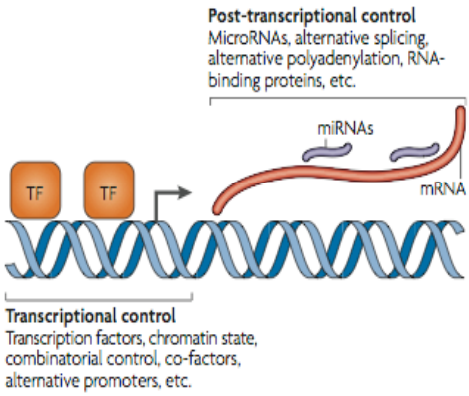


Figure 6: Scheme representing a few examples of transcriptional and post-transcriptional regulation. *Chen and Rajewsky, 2007.*

Gene regulation also occurs on the level of the RNA transcript. In fact, about 30–40% of all human genes are believed to undergo alternative splicing. Alternative splicing is a major mechanism to generate distinct mRNA isoforms from the same gene. Different isoforms are composed from different selections of exons, resulting in variant proteins (Huang et al., 2005). Some other examples of post-transcriptional regulation are length of the polyadenylation tail or presence of sequence elements in the regulatory region of the transcript. These features influence RNA stability and, as a result, are thought to control the amount of corresponding protein in the cell. Genome-wide analyses suggest that up to 50% of all RNAs change their stability significantly in response to cellular signals. As it will be extensively explained later, miRNAs function at this level to alter target RNA stability by means of interference mechanisms (figure 6).

Finally, genes are also extensively regulated at the protein level. Proteins can have active or inactive conformations, or can require chemical modifications or cofactors to be functionally active. This allows rapid and efficient protein control in the cell. In addition, cellular levels of proteins can be downregulated by targeted breakdown, such as the well-studied ubiquitine proteasome pathway mechanism.

In conclusion, genes act together as refined and complex networks controlling gene action or expression on several different levels. In any case, since the discovery of miRNAs, post-transcriptional regulation, previously underestimated, is growing in importance and is thought to play an important role on mammalian development and disease (Sun and Tsao, 2008).

Table 1: Overview of the regulation of gene expression in human cells

EPIGENETICS and CHROMATIN MODIFICATIONS
Allelic exclusion
Long-range control by chromatin structure
Short -range signalling dependent on cell position
TRANSCRIPTIONAL
Binding of tissue-specific transcription factors to cis-acting elements
Binding of hormones/growth factors to response elements (inducible transcription)
Using alternative promoters in a single gene
POST-TRANSCRIPTIONAL
Alternative splicing
Alternative polyadenilation
Tissue-specific RNA editing
Translational control
Regulation by microRNAs

3.1 non-coding RNAs

Contrary to what was thought during many years, it is nowadays accepted that about half of the human genome is transcribed and that most of the generated transcripts (~98%) are actually non-protein coding (Szymanski et al., 2005).

RNAs that do not code for proteins and directly function as RNAs are called non-coding RNAs (ncRNAs). Importance of ncRNAs is supported by the fact that their accumulation tends to increase with organism complexity along the evolutionary scale. Only a limited number of ncRNAs have been classically studied; mainly, ribosomal and transfer RNAs. However, in the last few years additional species of ncRNAs have increasingly been discovered, among which, small ncRNAs attract particular attention because of their role in processes such as RNA silencing and modification (Kawaji and Hayashizaki, 2008). Endogenous small RNAs related

with RNA silencing are diverse and can be categorised into three main classes: miRNAs, small interfering RNAs (siRNAs) and Piwi-interacting RNAs (piRNAs). piRNAs are slightly longer (25-29 nucleotides) than the other two classes and their biogenesis and function is still poorly understood. However, they have been implicated in germline development, silencing of selfish DNA elements, and in maintaining germline DNA integrity (Klattenhoff and Theurkauf, 2008). On the other hand, both miRNAs and siRNAs (~22 nucleotides in length) guide sequence-specific gene silencing at the post-transcriptional level and in contrast to piRNAs, they are produced in a Dicer-dependent mechanism. As the active forms of siRNAs and miRNAs are sometimes biochemically or functionally indistinguishable, they are classified based on their origins: miRNAs are generated from the double-stranded RNA (dsRNA) region of hairpin-shaped precursors while siRNAs are derived from long dsRNAs (Kim, 2005). Furthermore, miRNAs recognise target mRNAs by partial complementarity, whereas siRNAs show perfect complementarity to their target mRNA sequences.

3.2 microRNAs

A little over ten years ago, the Ambros and Ruvkun laboratories found that larval development of the nematode *Caenorhabditis elegans* required of the action of a tiny RNA to inhibit the expression of lin-14 mRNA. This small RNA was found to be lin-4, a 21 nucleotide RNA that recognises complementary sites in the 3'UTR of the lin-14 messenger. By doing so, it downregulates the translation of lin-14 during the transition from the first to the second larval stage of development (Lee et al., 1993). Lin-4 was the first miRNA to be described. Since

then, miRNAs have been found in plants and animal branches of Eukaryotes and have been shown to be encoded by a bewildering array of genes (Carthew and Sontheimer, 2009). In humans, miRNAs are scattered in the genome within all chromosomes, with the exception of the Y chromosome (Ro et al., 2007). The last version of the miRNA database (September 2009, Sanger miRBase, release 14.0) recognises a total of 721 known human miRNAs; yet this number has been growing quickly in the last few years and it is postulated that there are many more miRNAs to be identified. It is estimated that miRNAs will comprise 1%-5% of animal genes (Bartel, 2004), being, in consequence, one of the most abundant classes of regulators in the genome.

3.2.1 Genomic organisation and biogenesis

It was initially thought that miRNAs were encoded mostly in "intergenic regions". This was because the majority of miRNA loci were located in non-coding transcriptional units (TU) or in poorly characterised coding TUs. Nevertheless, more careful annotation of the genomic position of miRNAs indicated that about half of the miRNAs are located inside protein-coding (40%) or non-coding known RNAs (10%). It is thought that these miRNAs can be transcribed as part of the host gene transcription unit (Kim et al., 2009; Kim and Kim, 2007; Rodriguez et al., 2004). On the other hand, more than half of the miRNAs have been found in close proximity to other miRNAs and seem to be transcribed from a single polycistronic TU. It has been proven that, in many cases, clustered miRNAs share a common regulatory region and as a result, these miRNAs may be transcribed in the same developmental stages or biological

processes, suggesting that they are functionally related (figure 7) (Kim and Kim, 2007).

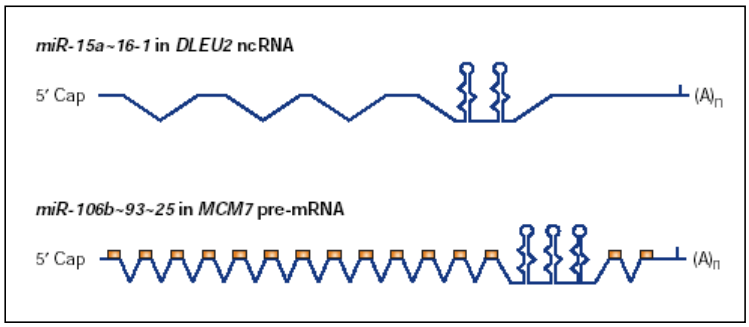


Figure 7: Example of two polycistronic miRNAs. This is an example of 5 miRNAs transcribed in a policystronic way. On the upper part, the policystron is transcribed from a non-coding mRNA. On the lower part of the panel, the policystron is transcribed from a coding RNA. *From Kim et al, 2005.*

Many of the bilaterian animal miRNAs are phylogenetically conserved, fact that indicates that miRNAs have had important roles throughout animal evolution (Ibanez-Ventoso et al., 2008). Nevertheless, animal miRNAs seem to have evolved separately from those in plants, because their sequences, precursor structure and biogenesis mechanisms are different (Chapman and Carrington, 2007; Millar and Waterhouse, 2005). On the other hand, most mammalian miRNAs have paralogues that are thought to result from duplication events. For instance, the human genome has 12 loci for the let-7 family of miRNAs. Paralogue miRNAs usually share the same seed region; which is the region comprising nucleotides 2-7 relative to the 5' end of the mature miRNA sequence. Because it is thought that the seed region is the most critical region for target recognition, paralogue miRNAs are thought to act redundantly. However, due to the fact that 3' sequences of miRNAs also contribute to target

binding and that paralogue miRNAs may be differentially expressed, members of the same seed family might also have distinct roles *in vivo* (Kim et al., 2009).

miRNAs are single stranded RNAs of 19-25 nucleotides in length that are generated from endogenous hairpin-shaped transcripts, fact that, as explained before, differentiates them from siRNAs (Kim, 2005)(figure 8). This hairpin-shaped transcripts are called pri-miRNAs, can be hundreds to thousands nucleotides long and are mainly transcribed by RNA polymerase II (polII) Therefore, they contain a 5'-methyl7G cap structure and are polyadenylated much like mRNAs (Cai et al., 2004). MiRNA gene promoters also share many characteristics with typical targets of polymerase II such as the presence of TATA boxes and conserved motifs that may serve as transcription factor-binding sites (Kim and Nam, 2006). It is noteworthy that miRNA promoter regions are twice as conserved as mRNA promoters in animals, highlighting the importance of miRNAs in the regulation of gene expression (Mahony et al., 2007). Nonetheless, a small fraction of miRNA genes may be instead transcribed by RNA polymerase III, being,

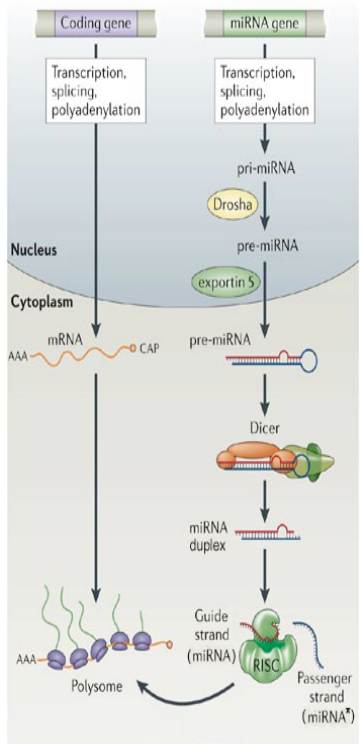


Figure 8: microRNA biogenesis pathway. miRNA genes generally have RNA polymerase II promoters, and their transcripts can undergo 5' capping, polyadenylation and splicing. *Adapted from Kosik, 2006.*

in consequence, regulated independently from canonical mRNA transcription units (Zhou et al., 2007).

Pri-miRNAs are cleaved, in the nucleus, at the stem of the hairpin structure by the nuclear RNaseIII–type protein, Drosha. The cleavage releases a stem-loop structure of 60-80 nucleotides, which possesses a short stem plus a 2nt 3' overhang, called precursor miRNA (pre-miRNA). Drosha, however, requires a cofactor to exert its function, DGCR8 (DiGeorge syndrome critical region gene 8) in humans or Pasha in *D. melanogaster* and *C. elegans.* DGCR8 is a double-stranded RNA-binding protein that together with Drosha forms a large complex (650kDa in humans) called the Microprocessor complex. Recent studies show that pri-miRNA processing of genic miRNAs might be a co-transcriptional process; this, was first based on the finding that Drosha processing of intronic miRNAs precedes the splicing of the host intron. Importantly, the cleavage of the intron by Drosha does not impair splicing, consistently with the "exon-tethering model" that postulates that the exons of PolII transcripts are cotranscriptionally assembled into the spliceosome. Thus, Drosha processing might take place after the transcript is tied to the splicing commitment complex, but before the intron is excised (Kim and Kim, 2007).

The pre-mRNA is subsequently exported to the cytoplasm by a heterodimer consisting of the transport factor Exportin-5 and the GTP-bound form of its cofactor, Ran. Following the export, the excised hairpin is processed by another RNaseIII enzyme called Dicer, which removes the loop region of the hairpin, releasing an imperfect RNA duplex known as the miRNA:miRNA* duplex (~22 nt). Dicer also interacts with RNA-binding cofactors that are not

required for the cleavage reaction itself, but stabilise the miRNA and effect the formation of a multi-protein complex called RISC (RNA-induced silencing complex), which mediates RNA silencing (He and Hannon, 2004; Kim, 2005). During the assembly of the RISC complex, one strand of the miRNA is retained while the other strand is lost. Strand retention is based on the relative thermodynamic stability of the duplex's ends; the strand with lower stability base pairing at nucleotides 2-4 at the 5' end is preferentially loaded onto the RISC and becomes the active miRNA, while the complementary miRNA* strand is in general removed and degraded (figure 8). This rule is not absolute and in some cases both strands may be selected from the opposite arms of the same hairpin precursor to give rise to two different mature miRNA sequences. To distinguish between them, these mature sequences were currently named using the extensions -5p (5' arm) and -3p (3' arm) depending on which arm of the hairpin they derive from - for example miR-17-5p and miR-17-3p - (Griffiths-Jones, 2006). However, this nomenclature has been recently modified and nowadays miRNA* (miR-146a*) is being used to refer to the passenger strand and miRNA (miR-146a) is used for the leading and most abundant strand.

Mature miRNAs loaded onto the RISC are ready to direct their activity as guide molecules in post-transcriptional gene silencing by base pairing with target mRNAs usually located in the 3'UTR of the transcript (Kim, 2005). However, the dogma that miRNA target sites lie preferentially in the 3'UTRs of the mRNA remains controversial and caution should be kept when working under this assumption. In fact, target prediction programs seem to be biased due to technical reasons; as most of them rely on evolutionary conservation, they were designed to focus on

3'UTRs, in order to avoid masking by codon-imposed evolutionary conservation of open reading frames (ORFs). This reduces the amount of DNA sequence to analyse, and therefore, false positives (Brodersen and Voinnet, 2009), but has the limitation that ~30% of human genes lack definitive 3'UTR boundaries (Maziere and Enright, 2007) and that this approach misses possible miRNA target sites within ORFs (Brodersen and Voinnet, 2009).

3.2.2 microRNA-mediated gene signalling

The molecular mechanisms of miRNA action remain intensely debated, however, a general way of action is nowadays widely accepted. The recognition of the target mRNA is mediated by the complementarity between the miRNA incorporated into the RISC complex (miRISC) and sequences located in the target mRNAs that are referred to as miRNA targets. In plants, most miRNAs base-pair to mRNAs with nearly perfect complementarity and induce mRNA degradation by an RNAi-like mechanism — the mRNA is cleaved endonucleolytically in the middle of the miRNA–mRNA duplex. In contrast, generally, metazoan miRNAs target mRNA transcripts through imperfect base-pairing to multiple sites in 3'UTRs, following a set of rules that have been identified by experimental and bioinformatic analyses (figure 9) and that are the basis of the algorithms used by the different target site prediction programs that have been developed (see Box 4).

1. Watson-Crick pairing to the seed region of the miRNA (nucleotides 2 to 8) is thought to be crucial for this targeting and is thought to act as the nucleus of the miRNA–mRNA association. G:U wobbles or mismatches and bulges in the seed

region greatly reduce silencing efficacy. Nevertheless, miRNAs target prediction programs tolerate G:U wobbles within the seed region and nowadays examine compensatory base pairing with the 3' end of miRNA. It has been also seen that an A residue across position 1 of the miRNA, and an A or U across position 9, improve the site efficiency, although they do not need to base pair with miRNA nucleotides.

2. Another rule is that bulges or mismatches must be present in the central region of the miRNA–mRNA duplex (no pairing of nucleotides 9-12), precluding the Argonaute (AGO)-mediated endonucleolytic cleavage of mRNA.

3. The third rule is that, although less important, 3' end pairing also contributes to target recognition. Even though, mismatches and bulges are generally tolerated in this region, good base pairing of nucleotides 13–16 of the miRNA become particularly necessary in the cases where sites have weaker miRNA seed matches.

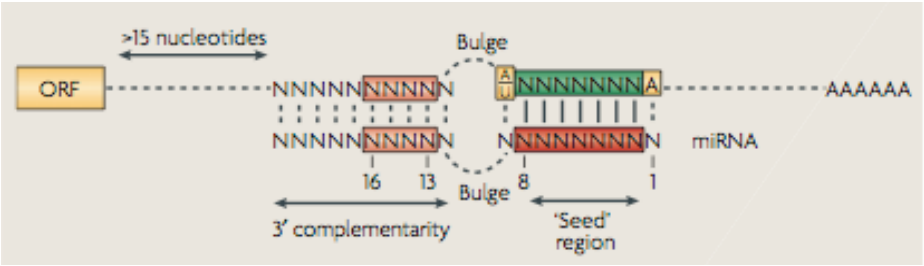


Figure 9: Principles of miRNA-mRNA interactions. The seed region is depicted in dark red, for the miRNA, and in green for the mRNA; positions 1 and 9 are depicted in yellow and nucleotides 13-16 are depicted in dark orange. *Filipowicz, 2008*

As for other factors, intrinsic to the mRNA target, that may improve site efficacy include an AU-rich neighbourhood and, in

the case of long 3'UTRs, a position that is not too far away from the poly(A) tail or the termination codon. Both factors contribute to a less structured and therefore, more accessible 3'UTR region. In addition, combinations of sites can require a specific configuration (for example, separation by a stretch of nucleotides of specific sequence and length) for efficient repression. It has been experimentally described that multiple sites for the same or different miRNAs are a favouring factor in effective repression and if localised near to each other (10–40 nucleotides apart) their action tends to be synergistic, that is, their effect exceeds that expected from the independent contributions of two single sites (Filipowicz et al., 2008).

It has been widely proposed that imperfect miRNA-mRNA hybrids that constitute central bulges (no pairing of nucleotides 9-12) enable translational inhibition or mRNA decay. In contrast, highly complementary targets with central pairing – which are rarely found in animals- result in mRNA cleavage, through what is called slicing process. This is, however, a generalisation and it is thought that other factors might influence this process, such as the type of RISC onto which small RNAs are loaded. The exact composition of the RISC complex is currently unknown, but all biochemical purifications performed so far have revealed the presence of at least one member of the Argonaute (Ago) family of proteins. These proteins are considered to be the core components of the RISC complex and in mammals four members have been described: AGO1 to AGO4. Of them, only AGO2 has endonucleolytic activity and is therefore, the only one that can contribute to target repression through mRNA cleavage (Brodersen and Voinnet, 2009). Research on the process through which miRNAs suppress protein synthesis is still

ongoing, and although controversy is present many agree that these small RNAs may act by different mechanisms, including, mRNA cleavage (previously mentioned), mRNA decay or deadenylation and direct translational repression (figure 10).

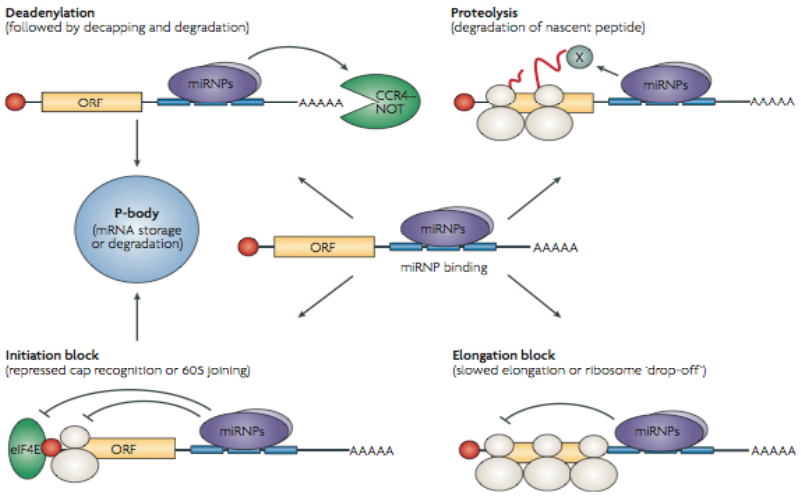


Figure 10: Possible mechanisms of the microRNA-mediated post-transcriptional gene repression in animal cells. *Filipowicz, 2008*

As for what direct translational repression is concerned, miRNAs are thought to act at all the steps of translation: initiation, elongation and termination. However, the precise way of action remains still unclear. At the initiation step, the miRISC has been shown to lead to translational repression by either interfering with translation initiation factors (eIF4E-eIF4G) or by directly binding to the m7Gcap (the 7-methylguanosine linked to the first transcribed nucleoside at the 5'end of eukaryotic mRNAs) of the mRNA, preventing the association of translation initiation factors to the targeted mRNA (Kiriakidou et al., 2007). As it is thought that the AGO proteins have a lower affinity for the m7Gcap than translation initiation factors, this mechanism would explain the requirement of multiple miRISC complexes for robust

repression (Filipowicz et al., 2008). Other mechanisms inhibiting the initiation step have also been reported, such as miRISC targeting of the poly-A tail (Wakiyama et al., 2007) or repression of translation by preventing ribosomal 60S subunit joining (Chendrimada et al., 2007); however complete interpretation of these mechanisms is still complicated. As there are results that are partially incompatible with the initiation model, it is likely that there are other mechanisms by which miRNAs bring about mRNA repression. These other mechanisms include premature termination of translation (Petersen et al., 2006), repression of Internal ribosomal entry sites (IRES)-containing reporters (Lytle et al., 2007) and decelerated translation.

On the other hand, even though initial studies suggested that levels of miRNA-inhibited mRNAs remain mostly unchanged, more recent work suggests that this repression can be under some circumstances associated with mRNA destabilization (Filipowicz et al., 2008). In relation to this, it has been seen that miRISC complexes along with suppressed mRNAs accumulate in the cytoplasm in foci called processing bodies (P-bodies). Once in P-bodies, mRNA degradation can follow two pathways, each of which is initiated by gradual shortening of the mRNA poly(A) tail. The first pathway involves miRISC recruitment of the exosome and subsequent degradation by progressive 3'→5' decay. And the second one, consists in the removal of the cap followed by 5'→3' degradation, that is catalysed by the exonuclease *Xrn* I (Parker and Song, 2004). These mechanisms involving miRNA-mediated mRNA destabilization are also thought to require miRISC recruitment of: GW182 protein, with decay promoting

activity; CCR4-NOT deadenylating complex; and decapping complex proteins (DCP1 and DCP2) (Filipowicz et al., 2008).

Interestingly, mRNAs transported to p-bodies can also be stored for later translational activation. This evidences that, under certain conditions or in specific cell types, miRNA-mediated repression may be reversible as it has been reported in the case of the reversion of CAT1 mRNA regulation by the liver-specific miR-122 in hepatoma cells following certain types of stress such as amino-acid starvation (Bhattacharyya et al., 2006). It has been also suggested that while miRNAs repress translation in proliferating mammalian cells, they may induce translation upregulation of target mRNAs under certain conditions, such as upon cell-cycle arrest (Vasudevan et al., 2007).

Another interesting observation that illustrates how much more complicated miRNA-mediated regulation of mRNAs can be is based on initial studies that show that miRNAs are prone to tissue-specific RNA editing. Editing is a posttranscriptional mechanism, by which some RNA molecules are altered to contain bases not encoded in the genome (specific nucleotides are either deleted, inserted or modified to change one nucleotide into another). Such editing events alter the properties of miRNAs and seem to regulate alternative mRNA:miRNA interactions. This has been, at least, demonstrated by miR-376 targeting a different set of genes after RNA editing in different tissues (Erson and Petty, 2008; Kawahara et al., 2007).

BOX 4. Target prediction Programs

The basic principle of many computational tools is that they are constructed from experimental examples, which means they are based on machine-learning algorithms (Yoon and De Micheli, 2006); however, since the mechanisms behind miRNAs' regulation are not completely revealed, elucidation of miRNA targets remains a major bottleneck and a challenge for the effectiveness of computational methods. In general, prediction criteria include the following parameters in their algorithms: **1**. Complementarity of the so-called *seed* region (however G:U wobbles within the seed region are tolerated (Miranda et al., 2006) and extensive base pairing with the remainder of the miRNA may compensate missing complementarity of the seed (Brennecke et al., 2005)) **2.** The thermodynamics of miRNA-mRNA duplexes **3.** Conservation of target sites among related genomes is the most commonly used property for reducing the search space (Ambros et al., 2003). However, the idea that all miRNAs are deeply conserved has been challenged, and a large number of taxon-specific miRNAs have now been identified (Bentwich, 2005).
The different predictions programs are basically combinations of these criteria, together with other particular requirements. The three programs used for this thesis which are in addition, the most commonly used are summarized in the table below:

Program	Sequence complementarity	Energy filter	Conservation	Other	web	False Positive Rate
TargetScanS	seed matches (6nucleotides) preceded by an adenosine	None	allows prediction for poorly conserved target sites		www.targetscan.org	22-31%
miRanda	algorithm that does not consider the seed region but gives higher weight to matches at the 5' end of the mature miRNA	Hybridization energy threshold (RNA fold)	yes		www.microrna.org; http://microrna.sanger.ac.uk	24-39%
PicTar	seed matches (7nucleotides) or compensatory	Hybridization energy threshold (RNA fold)	yes	able to calculate the likelihood depending on combinatorial targeting by different miRNAs	http://pictar.mdc-berlin.de	~30%

3.2.3 General functions

It is known that a single miRNA can target as many as several hundred genes, but it is also known that one gene can be targeted synergistically by more than one miRNA. Taken together, miRNAs form an interconnected regulatory network that does not simply turn genes on or off, but are thought to ''tune'' the expression level of their target genes (Sun and Tsao, 2008). The importance of miRNAs is made evident by their conservation along evolution and by the multiple processes in which they are implicated, such as: developmental timing, cell differentiation and morphogenesis (Stark et al., 2005). In order to test the global importance of miRNAs, Dicer gene was knocked out in a mouse model, thus, inhibiting miRNA production. This approach showed that Dicer deficient embryos die at an early embryonic age and are depleted of pluripotent stem cells, supporting the crucial role of miRNAs in proper embryogenesis and stem cell development (Bernstein et al., 2003). Following a similar approach, conditional knockouts of Dicer have been generated for different tissues or organs. This has demonstrated roles for Dicer in morphogenesis of several organs, including the lungs, limbs, and muscles, and in T-cell differentiation. There is, yet, one important caveat to interpreting these lines of knockout experiments; the assumption that Dicer does not play other roles in addition to miRNA and siRNA processing (Sun and Tsao, 2008). Investigators, however, have pinpointed specific miRNA functions in invertebrate and vertebrate development and differentiation such as; stem cell proliferation, division and differentiation (e.g. adipocyte, cardiac, neural and hematopoietic lineages) (table 2).

Table 2: Roles of miRNAs in development. *Adapted from Dalmay, 2008.*

Name of miRNA	Function	Target gene
Lin-4	*Caenorhabditis elegans* development	Lin-14, lin-28
Let-7	*C. elegans* development	Lin-14, lin-28, lin-41, lin-42, daf-12
miR-196	Embryo patterning	HOXD8
miR-1	Cardyomyocyte differentiation	HDAC4
		Myostatin
miR-133	Suppression of cardyomyocyte differentiation	Serum response factor
miR-124	Neuronal differentiation	Laminin gamma 1
		Integrin beta 1
		SCP1
		PTBP1
miR-34	Dendritic spine development	Limk1
miR-133b	Maturation of dopaminergic neurons	Pitx3
miR-181	B lymphocyte differentiation	Not known
	Myoblast differentiation	HOX-A11
miR-375	Insulin secretion	Myotrophin
miR-122	Liver development and function	Many genes
miR-143	Adipocyte differentiation	ERK5

During embryogenesis many miRNAs show tissue-specific expression patterns, whereas stem cells express specific sets of miRNAs that conversely become downregulated during differentiation (Suh et al., 2004). Examples of miRNA activity at an early stage are the targeting of *HOX* genes (a family of genes involved in developmental patterning) by miR-196 (Yekta et al., 2004), the promotion of muscle myogenesis in murine cells by miR-26a, which targets a suppressor of skeletal muscle cell differentiation, *Ezh2*, resulting with upregulation of *myoD* and myogenin mRNA expression (Wong and Tellam, 2008), the role of miR-181, miR-223, and miR-142 differentiation of the hematopoietic lineage (Chen et al., 2004), liver development and function by miR-122 and adipocyte differentiation by miR-143 (Table 3) (Dalmay, 2008).

Apart from having functions in development and differentiation, miRNAs have also shown to be important in the normal functioning of the body. For instance, emerging evidence suggests that miRNAs play a key role in the regulation of immunological functions including innate and adaptive immune responses, development and differentiation of immune cells and the prevention of autoimmunity (Pauley and Chan, 2008). Moreover, stress responses can also involve miRNAs, as different types of cellular stress have been shown to alter miRNA levels. For example, hypoxia-responsive transcription factors such as nuclear factor-kappa B and p53 induce miRNA genes (He et al., 2007; Taganov et al., 2006). In general, studies on oxidative stress, cold stress and nutrient deprivation indicate that long-term stress may have an impact on miRNAs and on global gene expression, perhaps leaving tissues more susceptible to pathogenic processes. It is thought that chronic stress can initiate cellular reprogramming through alterations in miRNA expression or activity, leading to sustained changes in gene expression and cellular physiology (Hudder and Novak, 2008). Finally, it is worth mentioning that miRNAs are also involved in cell cycle progression and apoptosis (Carleton et al., 2007), as emphasised by their implication in cancer, in learning and memory (Fiore and Schratt, 2007) or in endocrine regulation of energy homeostasis (insulin secretion and miR-375), among other (Poy et al., 2004).

3.2.4 miRNAs in central nervous system

In the last few years, miRNAs have emerged as important players in gene regulation in the nervous system; in relation to this, various studies on miRNA expression profiles showed a high degree of temporal and spatial specificity for miRNAs and indicated a fundamental implication of miRNAs in neuronal fate decision and CNS development, as well as in later stages of neuronal maturation and synapse development. Further, Dicer deficient zebrafish, which lack all mature miRNAs, showed abnormal morphogenesis including defects in neural development. Notably the various defects in neural development were rescued by the injection of a single family of miRNAs, miR-430, which is ubiquitously expressed during early stages of development (Giraldez et al., 2005). As for results concerning tissue-specific knock-down of Dicer, it has been seen that in mouse Purkinje cells (Schaefer et al., 2007) and in dopaminergic neurons it causes progressive death of neuronal cells (Kim et al., 2007). In addition, as revealed by an expression study, there is a subset of over a hundred brain-expressed or brain enriched miRNAs in mice and humans that have a role in neuronal differentiation and synaptic function. Among them, seven brain-specific (i.e. exclusively detected in brain) miRNAs including miR-9, miR-124a and miR-124b and seven brain-enriched (i.e. whose expression was at least two-fold higher in brain than in other organs) miRNAs including miR-9*, miR-125a, miR-125b and miR-128 were identified. Interestingly these miRNAs also show a high conservation of expression between mouse and human. In general, these results showed that there is a high percentage of brain-specific and brain-enriched miRNAs compared to other organs, as well as an overall prevalence of miRNAs expressed in the brain. In the same study, the authors

also analysed which of the brain-expressed miRNAs were associated with neuronal differentiation, As a result, all of the brain-specific and brain-enriched miRNAs listed above were found to be induced by retinoic acid exposure, indicating that they might contribute to the specification of neuronal identity. In addition, other brain-non-enriched miRNAs were as well upregulated, like let-7a and let-7b among others (Sempere et al., 2004). Other important evidence for the role of miRNAs in establishing and maintaining neuronal cell identity comes from studies carried out in mouse embryonic stem cells. Krichevsky *et al.* demonstrated that neuronal differentiation of embryonic stem cell is accompanied by the upregulation of several miRNAs. Among them miR-9, miR-9*, miR-22, miR-124a and miR-125b are highly expressed in brain and in primary neurons. In particular, overexpression of miR-9 and miR-124a in neural precursors alters the neuron-to-astroglia ratio during neural lineage differentiation, reducing the number of astrocytes, while inhibition of the two miRNAs causes a reduction in the number of neurons (Krichevsky et al., 2006). Finally, profiling experiments performed to define the temporal expression of miRNAs in the mouse brain consistently showed that miR-124, miR-125 and miR-128 accumulate in parallel to neuronal maturation, while miR-23 and miR-29 are expressed in astrocytes and at low levels in embryonic development (Smirnova et al., 2005). MiRNAs have been also shown to play an essential role in synaptic plasticity. The term synaptic plasticity refers to the ability of a synapse to change its efficiency and anatomical organisation in response to repeated stimulation, resulting in increased synaptic strength; likewise, if a synapse has not been used for some time its strength will decrease. The prime example of a miRNA implicated in synaptic plasticity in

mammalian neurons is miR-134. In rats, miR-134 inhibits Lim-domain-containing protein kinase 1 (*Limk1)* translation negatively controlling dendritic spine size. Limk1 is involved in

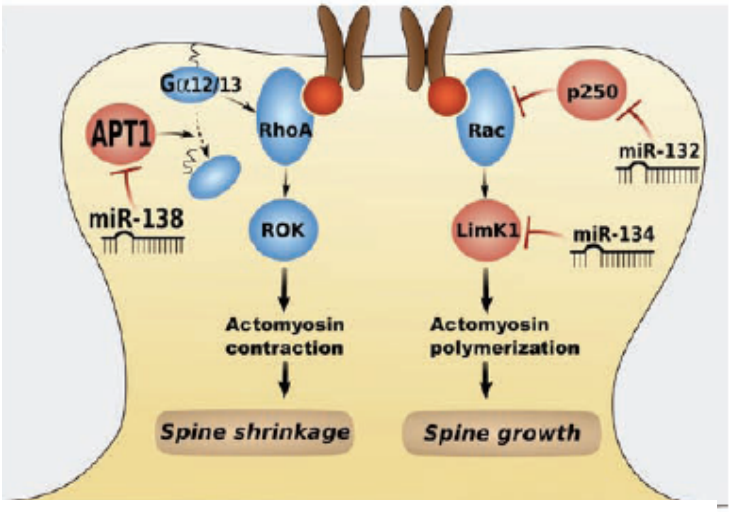


Figure 11: Model for the role of miRNAs in the regulation of dendritic spine morphogenesis. *Siegel,G. 2009.*

actin filament dynamics, a key step in the cytoskeletal modifications of spines associated with plasticity. However, this effect is reversed by the release of BDNF, which ends up with the silencing of limk1 and promotes spine growth (Schratt et al., 2006). On the other hand, miR-138 also seems to have a role on spine structure by regulating the depalmitoylation enzyme APT1. $G\alpha_{13}$, an activator of Rho downstream G-protein-coupled receptors is one of the known APT1 substrates. Therefore, a miR-138-mediated increase in $G\alpha_{13}$ palmitoylation and membrane localization could result in elevated Rho activity, which in turn would trigger spine shrinkage. Electrophysiological measurements showed that miR-138-mediated spine-shrinkage is correlated with decreased postsynaptic function. Interestingly, it has also been demonstrated that in turn, miR-138 expression

and activity are negatively regulated by calcium influx. All in all, it is intriguing how neuronal activity can regulate the expression and activity of a variety of synaptic miRNAs and how these miRNAs can contribute recessively to activity-dependent fine-tuning of signalling pathways that coordinate the structural and functional plasticity of spine synapses (figure 11) (Siegel et al., 2009). Finally, another indication of the involvement of miRNAs in controlling local protein translation and synaptic function comes from a recent study that demonstrated that miR-128 is deregulated in HIV-1 encephalopathy (a manifestation of HIV-1 infection that often results in neuronal damage and dysfunction) and that, in addition, miR-128 inhibits the expression of SNAP25, a pre-synaptic protein that regulates Ca^{++} responsiveness (Eletto et al., 2008).

miRNAs are also rapidly emerging as switches of key regulatory pathways in the cell, such as critical alternative splicing mechanisms, which may contribute to tissue specificity. Importantly, miR-124 has been shown to be involved in the regulation of neuron-specific alternative splicing. One validated target for miR-124 is the polypyrimidine tract-binding protein PTBP1. PTBP1 is an important splicing regulator that represses alternative pre-mRNA splicing in non-neuronal cells. During neuronal differentiation miR-124 reduces PTBP1 levels, which in turn, permits the correct splicing and expression of its neuronal homolog PTBP2 (previously repressed by the higher levels of PTBP1). This switch has substantial consequences on the splicing patterns of genes involved in neuronal functions (Makeyev et al., 2007). On the other hand, miR-133 has been shown to downregulate the PTBP2 during myoblast differentiation, mirroring the example of miR-124 in the nervous system (Boutz

et al., 2007). In fact, the best studied neuronal miRNA is certainly miR-124, which constitutes 25-50% of total miRNAs expressed in the brain (Johnston and Hobert, 2003) and its overexpression in HeLa cells downregulates more than 100 non-neuronal mRNAs, producing a neuron-like expression profile (Lim et al., 2005). Moreover, during chicken spinal cord development miR-124 is needed to preserve neuronal identity. On the other hand, in cortical neurons upon miR-124 knockdown several non-neuronal mRNA transcripts are increased. This is partly mediated by the fact that miR-124 has been validated functionally to target *SCP1* (small C-terminal domain phosphatase-1), an activator of *REST* (RE1 silencing transcription factor). *REST* is a transcriptional repressor of neural genes in non-neuronal tissues, including the miR-124 gene (Conaco et al., 2006). miR-124 contributes to neurogenesis inducement by timely down-regulating *SCP1* (Visvanathan et al., 2007).

3.2.5 miRNAs and disease

In addition to the modulation of physiological functions and due to their important regulatory role in processes of physiopathological relevance, a close relationship between miRNAs and human diseases has been described. Regulatory changes that alter miRNA activity can be caused by both in *cis* and *trans* factors (relative to the locus coding the miRNA), making possible to schematically group known miRNA-related diseases based on the specific process altered. *Cis* deregulation of miRNAs can be caused by chromosomal alterations, epigenetic modifications, polymorphic promoter elements and polymorphisms within the miRNA itself (pri-, pre- and mature

miRNA sequences). *Trans* factors affecting miRNA activity, on the other hand, include functional mutations in the proteins involved in miRNA transcription, processing and targeting, and polymorphisms mutations in miRNA target sites (poly-miRTS). Therefore, we will divide this important section of the introduction into *cis* and *trans* deregulation of miRNA activity that contributes to disease. In both cases, especial emphasis will be made to CNS diseases.

3.2.5.1 *Cis*-deregulation of miRNAs

The implication of miRNAs in cancer has been one of the topics for major research activity in relation to miRNAs in the last years. In fact, early clues linking chromosomal alterations in miRNA loci and disease came from observations in chronic lymphocytic leukemia (CLL), where chromosome band 13q14, commonly lost or altered in CLL patients, was found to harbour miR-15a and miR-16a (Calin et al., 2002). Both miR-15a and miR-16 were later shown to target the 3'UTR of *BCL2*, a well-known anti-apoptotic oncogene (Cimmino et al., 2005). Initial observations that miRNA genes were located on genomic instability and fragile sites by Calin *et al.* (Calin et al., 2004) lead to further analyses that have demonstrated deregulated miRNA expression profiles in various diseases. These different expression profiles have been intensively studied in cancer, where miRNA arrays providing information on deregulated miRNA expression profiles have demonstrated to be specific enough to be effective for tumour classification purposes. Actually, differentially expressed miRNA profiles have been proven to be more reliable than mRNA profiles to define tumour subclasses in some cases (Lu et al., 2005). One of the many miRNAs that was shown to be deregulated in cancers was miR-

21, which is overexpressed in a variety of different tumours including breast cancer. So far, identified targets of miR-21 are *PTEN* and *PDCD4*, genes that are well known to be involved in cell survival and transformation processes (Asangani et al., 2008; Meng et al., 2007). Another example of miRNAs that behave like oncogenes is the polycistronic miRNA cluster miR-17-92 (on 3q31.3). This cluster is thought to enhance cell proliferation and has been found to be overexpressed in several tumour types such as lung (Hayashita et al., 2005), and, more recently, in colorectal cancers (Lanza et al., 2007). On the other hand, miRNAs can also behave as tumour suppressor genes. Examples on such type of miRNAs, that have been found to be downregulated in cancer, have also been documented. Let-7, for instance, targets the widely recognised oncogene -*RAS*-, among others, and has also been found to be downregulated in cancer (Johnson et al., 2005).

With respect to disorders of the CNS an increase in miR-9, miR-125b and miR-128 levels was detected in the hippocampus of Alzheimer's disease (AD) brains (Lukiw, 2007). Later, Wang *et al.* (Wang et al., 2008b) analysed miRNA expression profiles of human brain tissue from non-demented individuals with negligible clinical features of AD, non-demented individuals with incipient clinical features of AD, individuals with mild cognitive impairment with moderate clinical AD, and individuals with full-blown clinical features of AD. They reported significantly decreased miR-107 levels in patients with even the earliest stages of the disease. When analysing further the role of this miRNA, they concluded that miR-107 might be involved in accelerated disease progression through regulation of the beta-site amyloid precursor protein-cleaving enzyme 1 (BACE1).

Recently, Hebert *et al.* (Hebert et al., 2008) demonstrated that miR-29a, miR-29b-1, and miR-9 can also regulate *BACE1* expression *in vitro* and that these miRNAs were also decreased in AD patients resulting in high BACE1 protein levels in patients. On the other hand, reduced expression of miR-133b has been observed in dopaminergic neurons of Parkinson's disease patients and reduced level of this miRNA have been, thus associated with the typical degeneration of this type of neurons in Parkinson disease (Kim et al., 2007). Interestingly, distinct miRNA expression patterns have also been implicated in chronic psychiatric disorders: miR-26b, miR-30b, miR-29b, miR-195, miR-92, miR-24, miR-30e were shown by microarray and quantitative reverse transcriptase-polymerase chain reaction to be decreased in samples from individuals with schizophrenia (Perkins et al., 2007). However, these miRNAs and their targets and how they may be involved in common complex neurological and psychiatric disease states are yet to be examined. Concerning mutations in miRNAs or in their promotor sequences and in relation to schizophrenia, two known SNPs, located in the adjacent +/- 100 bp genomic region of miR-206 and miR-198 were claimed to be associated with schizophrenia. One of them, in miR-206 adjacent genomic region, remained significant after correction for multiple testing. In order to elucidate what biological signalling network might be the one affected by these mRNAs in schizophrenia, the authors performed target predictions for both miRNAs. Target predictions rendered a list of 15 genes that were predicted to be corregulated by both miRNAs. Interestingly, two of the common targets had been previously related to schizophrenia; *CCND2* had been shown to be deregulated in post-mortem schizophrenia brains and *PTPN1* had been positioned under a significant linkage peak (Hansen et

al., 2007). Nevertheless, the authors did not go further on the association and did not show true biological evidence of the implication of these miRNAs in schizophrenia or of these polymorphisms in miRNA expression. Similar studies have implicated mutations in miRNAs or their regulatory regions with different diseases, such as a mutation in miR-27a that has been associated with gastric mucosal atrophy in Japanese men (Arisawa et al., 2007), or several SNPs in pre-and pri-miRNAs that showed borderline association with bladder cancer in whites (Yang et al., 2008). Nevertheless, little biological insight into the participation of these SNPs in the etiogenesis of the disease has been provided. In fact, few examples of mutations on miRNA genes or their regulatory regions have been exhaustively reported so far. Perhaps the best reported study of this type was published last year. In this study, pri-miR-146a and pri-miR-146b were resequenced in 15 patients with papillary thyroid carcinoma (PTC) who had shown a 19-fold increase in the quantity of miR-146 with respect to controls. As a result, a common SNP in pre-miR-146a was reported to be strongly associated with PTC. Remarkably, both homozygous states were shown to be protective whereas heterozygotes had an increased risk of acquiring the disease (OR=1.62). More importantly, the heterozygote state was proven to be enriched in tumour tissues, evidencing that the SNP contributes to the genetic predisposition to PTC, and plays a role in tumorigenesis through somatic mutation. Because localization of this SNP was confined to the passenger strand (miRNA*), the authors failed to offer a good explanation of the role of this variant in PTC formation (Jazdzewski et al., 2008). However, more recently, the authors demonstrated that both miR-146a passenger (miRNA*) and leading strand (miRNA) were expressed, and in the case of

heterozygotes three different forms of this miRNA existed; miR-146a (leading strand), miR-146a-G (passenger strand) and miR-146a-C (passenger strand), each with a distinct set of target genes. Moreover, 7 out of 8 tumours studied, showed a widely different transcriptome (as assessed by microarray) and a 1.5- to 2.6 fold overexpression of polymorphic miR-146a (passenger strand) when compared with the unaffected part of the same gland (Jazdzewski et al., 2009).

3.2.5.2 *Trans*-deregulation of miRNAs

Trans-deregulation of miRNAs can be broadly summarized into structural alterations involving genes that are important in miRNA biogenesis (e.g Fragile X and DiGeorge syndromes), or mutations in miRNAs target mRNA sequences (e.g. Tourette syndrome and pathological aggressiveness). A few examples of human disease shown to be caused by deregulation in the miRNA pathway have been reported, such as cancer or Fragile X Syndrome (Gong et al., 2005). Fragile X syndrome (FX), for instance, is one of the most common forms of mental retardation, and is characterised by abnormalities in the structural development of dendritic spines. It is caused by a CGG repeat in the 5'UTR of the *FMR1* gene, which is located on the long arm of chromosome X. The condition becomes clinically manifest when the repeat expands as the gene is passed from generation to generation, until its transcription is completely shut down in the full-blown syndrome (Penagarikano et al., 2007). The current view is that Fragile X protein (FMRP) associates with endogenous miRNAs and with Ago1- in mammals- to translationally repress a subset of dendritic mRNAs (Jin et al., 2004) and that the disease is caused by the deregulated expression of its mRNA targets, which encode

factors required for synaptic plasticity and development. In relation to this, another study carried out in Drosophila has shown an interesting overlapping between the composition of FMRP-containing neuronal granules and P-bodies, which suggests that these classes of granules might be similar not only in composition but also in function (Hillebrand et al., 2007). Similarly, most cases of DiGeorge syndrome result from a deletion of chromosome 22q11.2 (the DiGeorge syndrome chromosome region, or DGCR); this deletion includes the *DGCR8* gene. DiGeorge syndrome is a developmental disorder characterised by mental retardation, as well as structural and functional palate anomalies, conotruncal cardiac malformations, immunodeficiency, hypocalcemia, and typical facial anomalies. DGCR8, as explained before, is required for the maturation of miRNA primary transcripts. In fact, its knockdown leads to accumulation of pri-miRNAs and reduction of pre-miRNAs and mature miRNAs (Landthaler et al., 2004), indicating again a plausible involvement of miRNAs in the aetiology of the disease. More recently, Melo *et al*. demonstrated that the presence of inactivating mutations in the gene encoding the RNA-binding protein, *TRBP*, in sporadic and hereditary carcinomas. Importantly, the authors show that the reintroduction of TRBP in the deficient cells restores efficient miRNA processing and synthesis and blocks cancer cell growth. TRBP, like DGCR8, is an RNA-binding protein that acts as a catalytic partner of DROSHA (Melo et al., 2009).

Since 2005 -a year after of the start of this project- until now, the number of described mutations in miRNAs target mRNA sequences (poly-miRTS) has been growing exponentially, being the majority of these studies published in the last two years

(Table 3). Abelson *et al.* were the first to associate a sequence variant in a miRNA target site with disease in 2005 (Abelson et al., 2005). They reported a rare sequence variant enhancing a target site for miR-189 in the *SLITRK1* (Slit and TRK-like family member 1) gene in two patients with Tourette's syndrome and in none of the controls tested. An altered interaction between this miRNA and the *SLITRK1* mRNA in the developing brain was suggested to contribute to this neuropsychiatric disorder. However this study has been recently treated with scepticism since a follow-up study indicated that the variant associated with Tourette Syndrome was overrepresented in certain white subgroups such as Ashkenazi Jews, thereby complicating the interpretation of the results in a sample where cases and controls could happen to be not appropriately matched (Keen-Kim et al., 2006). In addition, other replication studies in Italians and Taiwanese did not reproduce the association with Tourette Syndrome, due to the fact that the involved variant was absent among cases (Deng et al., 2006; Fabbrini et al., 2007). After that, the putative importance of poly-miRTS was strengthened by a study from *Clop et al.* (Clop et al., 2006). In a rigorous study, the authors provided *in vivo* evidence that G to A transition in the 3' UTR of *GDF8* generated an illegitimate target site for miR-1 and miR-206, two miRNAs highly expressed in skeletal muscle. This illegitimate target site causes translational inhibition of the myostatin gene and hence contributes to the muscular hypertrophy of Texel sheep (this study is not included in Table 3 as this poly-miRTS is not implicated in human disease). Since then, more studies have followed, linking poly-miRTS with disease. Regarding those affecting the CNS two mutations were identified in the 3'UTR of the receptor expression enhancing protein 1 gene (*REEP1*), after previous

association of this region with hereditary spastic paraplegia (Zuchner et al., 2006). These mutations were predicted to strengthen miR-140 mediated repression. Independently, in a similar approach Beetz *et al.* identified one of these variants and discovered a third variant putatively affecting miR-691 regulation in two other hereditary spastic paraplegia families (Beetz et al., 2008). However, no functional studies were performed in these studies. Last year, Jensen *et al.*, aimed to study the role of miRNAs in human behaviours; as the deletion of serotonin receptor 1B (*HTR1B*) has been shown to cause aggressive behaviour in mice, the authors analysed the gene's 3'UTR. A common polymorphism in the serotonin receptor 1B affecting the target site for miR-96 was shown to be associated with aggressiveness in humans (Jensen et al., 2008). In addition, a cross-species approach to identify genes that regulate anxiety-like behaviour pinpointed a SNP in the 3'UTR of the δ-aminolevulinate dehydratase (*ALAD*) that was part of a haplotype that was shown to be associated with social phobia. Even though this is not a true poly-miRTS study, this association is of particular interest for the topic of this thesis since the authors, at the end of their discussion, comment briefly on the possibility that this SNP generates an illegitimate target site for miR-211 and miR-204 -as predicted by a miRNA target prediction program- (Donner et al., 2008). Finally, a polymorphism in fibroblast growth factor 20 (*FGF20*) conferring a risk to Parkinson's disease was shown to disrupt a miR-433 target site. Unlike other studies, the authors went further to provide human in vivo validation of this target site and a molecular mechanism by which differential miR-433 targeting leads to Parkinson's disease. According to their report, disruption of miR-433 target site leads to increased *FGF20*

translation, which in turn, increases alpha-synuclein expression and ultimately causes Parkinson's disease. Another aspect that makes this study attractive is the fact that it is not confounded by population stratification because of the use of family-based association study design (Wang et al., 2008a).

On this section we have only focused on those poly-miRTS affecting CNS; nevertheless, other poly-miRTS have been implicated in different disorders such as breast cancer (Adams et al., 2007), hypertension (Martin et al., 2007; Sethupathy et al., 2007), methotraxate resistence (Mishra et al., 2007), Childhood asthma (Tan et al., 2007), colorectal cancer (Landi et al., 2008). All of these examples are briefly summarized in Table 3.

Table 3: poly-miRTS implicated in human disease

Associated disease or trait (uncorrected P value)	miRNA	Target gene	Putative risk allele (putative effect on miRNA targeting)	Risk allele frequency (sample)	Functional assay to test for allele-specific effects on miRNA targeting	Sample size	Population (quantitative correction for potential population stratification)	Notes on the associations (Curiosities)
Tourette Syndrome (0.0056)	miR-189	*SLITRK1*	var321-SLITRK1 [A] (+)	~1.5% (study sample)	*In vitro* : reporter gene assay in Neuro2A mouse neuroblastoma cell	139 cases and 2188 controls	>80% white (none)	Not replicated in three independent studies
Hereditary spastic paraplegia 31 (not given) Hereditary spastic paraplegia 31 (not given) Hereditary spastic paraplegia 31 (not given)	miR-140 miR-140 miR-691	*REEP1* *REEP1* *REEP1*	606+50 [A] (+) 606+43 [T] (+) 606+14 [A] (+)	~1% <1% <1%	*In silico* : miRNA target prediction program *In silico* : miRNA target prediction program *In silico* : miRNA target prediction program	90 cases and 365 controls 535 cases and 1000 controls 535 cases and 1000 controls	Of European descent (none) Of European descent (none) Of European descent (none)	Risk allele absent from controls Risk allele absent from controls Risk allele absent from controls
Breast cancer (none)	miR-206	*ESR1*	rs9341070 [C] (-)	99% (Hapmap CEU)	*In vitro* : reporter gene assay in MCF7 breast cancer cell line	-	-	No case-control association test; another variant in this receptor has been shown to affect miR-453, very recently, in an independent study
Hypertension (<0.05)	miR-155	*AGTR1*	rs5186 [C] (-)	~30% (mixed whites)	(i) *In vitro* : reporter gene assay in 293T and CHO cell lines (ii) *In vitro* : Radioligand binding studies	Various previous studies	various including among the positively associated: various European populations, Tibetan and Turk populations (none)	Not replicated in 22 of 40 studies (Two similar studies were published simultaneously. One by Sethupathy *et al* and another one by Martin, MM *et al*)
Methotrexate resistance (none)	miR-24	*DHFR*	rs34764978 [T] (-)	Unknown	*In vitro* : reporter gene assay in DG44 CHO cell line	-	-	No case-control association test
Childhood asthma (0.001)	miR-148a; miR-148b; miR-152	*HLA-G*	rs1063320 [C] (-) and [G] (+)	50% (Hapmap CEU)	*In vitro* : reporter gene assay in JEG3 cell line	180 COAST children	White Americans (none)	Logistic regression for interaction effects with mother's affection status; No replication study
Arson or property damage (0.008)	miR-96	*HTRB1*	rs13212041 [A] (+)	85% (Hapmap CEU)	*In vitro* : reporter gene assay in HeLa cell line	231 A homozygotes and 128 G carriers	White college students (none)	No replication study and not significant after multiple testing correction
Colorectal cancer (0.009)	miR-337; miR-582; miR-200a*; miR-184; miR-212	*CD86*	rs17281995 [C] (+ or -, depending on the miRNA)	12% (Hapmap CEU)	*In silico* : miRNA target prediction program	697 cases and 624 controls	From Czech republic (none)	No replication study
Parkinson's disease (0.0001)	miR-433	*FGF20*	rs12720208 [T] (-)	14% (Hapmap CEU)	(i) *In vitro* : reporter gene assay in Neuro2A mouse neuroblastoma cell (ii) *in vivo* : immunoblot analysis in 3 human brains with different genotypes at the polymiRTS	1089 cases and 1165 controls (from 729 families)	White Americans (Family-based association studies are immune to population stratification effects)	Pedigree disequilibrium test (PDT) and association in the presence of linkage (APL); No replication study
Diarrhea predominant irritable bowel syndrome (0.033)	miR-510	*HTR3E*	rs62625044 [A] (-)	~4% (British cohort)	*In vitro* : reporter gene assay in HEK293 and Colo320 cell line	68 cases	British (none)	Replication study in German cohort: 75 cases and 132 controls (p=0.0046)
Social Phobia (corrected: 0.006)	miR-204; miR-211	*ALAD*	rs8177822 [A] (+)	<1% (study sample)	*In silico* : miRNA target prediction program	58 cases and 133 controls	Finnish (none)	The association was found for the haplotype

Adapted and completed from Sethupathy *et al*, 2008 (Sethupathy and Collins, 2008)

Objectives

The hypothesis of this thesis is that genetic variation, either in miRNAs themselves or in miRNA target sites of candidate genes, could affect miRNA-mediated regulation and could therefore alter the dosage of proteins involved in fundamental pathways of brain function, contributing to the susceptibility to anxiety disorders.

The main objective of this thesis is thus, the description and analysis of the genetic variability involving these regulatory regions that could be underlying the molecular basis of anxiety disorders. The more concrete objectives can be grouped in two different categories depending on the two different approximations used: a single gene approach, based on the study of a particular candidate gene, and a more genomic approach, based on the analysis of the whole collection of miRNAs.

A. Candidate gene approach

1. To identify new allelic variants and study the genetic variation in two different 3'UTR regions which contain putative target sites for miRNAs in *NTRK3*, a candidate gene for anxiety disorders.
2. To validate predicted miRNA target sites in the 3'UTR of *NTRK3* that may be disrupted by the identified allelic variants and evaluate the consequences of these variants by functional experiments using a luciferase reporter-based assay.
3. To perform case-control studies in patients with panic disorder and obsessive-compulsive disorder using the new identified and already known informative allelic variants of *NTRK3.*

B. Analysis of miRNAs at the genomic level

1. To perform an "*in silico*" analysis of the localization and distribution of miRNAs at the genomic level regarding other transcriptional units as well as their aggregation in clusters.
2. To study the genetic variability and SNP coverage of miRNA regions in order to select and design a panel of informative SNPs covering genomic regions containing known miRNAs.
3. To perform case-control studies in Spanish patients with panic disorder and obsessive-compulsive disorder using the designed panel of SNPs and to carry out a replication study for panic disorder using two different European populations.
4. To investigate the possible involvement of miRNAs associated with panic disorder in the Spanish and North European populations in the postranscriptional regulation of candidate genes for this disorder by mean of a luciferase reporter-based system.
6. To analyse the consequences of the overexpression of two neuronal miRNAs associated with panic disorder at the transcriptome level in order to identify altered pathways and molecules of interest.

Results

The results section of this thesis is divided into the two different categories defined on the objectives: the study of *NTRK3,* as a candidate gene approach and the genomic approach based on the analysis of the whole collection of miRNAs.

A. Candidate gene approach

The plan for this first part of the results was to analyse *NTRK3* as a candidate gene for PD and OCD. In particular we studied whether genetic variation in miRNA target sites of *NTRK3* could be a predisposition factor for anxiety disorders in our population. *NTRK3* was selected as a candidate gene for the functional and genetic reasons exposed in the introductory chapter of this thesis. The implication of this gene in other psychiatric disorders such as eating disorders (Mercader et al., 2008) and depression (Nestler et al., 2002) made of *NTRK3* an excellent candidate to be involved in the development of anxiety disorders and framed the starting point of this study for which results are reflected in the accepted paper included in this chapter.

Our aim was not to use a linkage disequilibrium based strategy but a more functional approach and, in addition, to study if different isoforms of *NTRK3* were prone to post-transcriptional regulation by miRNAs. This last objective was carried out in parallel by our group demonstrating, at the RNA and protein levels, that *NTRK3* full length and truncated isoforms are regulated by different sets of miRNAs (these results are part of a paper recently submitted to publication that is presented as an appendix).

In the first article included in this chapter, we investigated genetic variation in these two *NTRK3* isoforms as candidate susceptibility factors for anxiety. To do that, we re-sequenced *NTRK3* 3'UTRs in both isoforms in patients with PD, OCD and controls. As a result, six new allelic variants and an already known SNP were identified within ten putative target sites for miRNAs regulation. We functionally validated the target sites where allelic variants were found using a luciferase reporter-based system. According to this functional assay, only five out of the ten target sites were functional, rendering a total of four out of the seven allelic variants located within functional miRNA target sites. Association analysis of these variants revealed that the C allele of rs28521337, located in the functional target site for miR-485-3p, was significantly associated with the hoarding phenotype of OCD although this variant does not seem to significantly change the affinity and efficiency of the miR-485-3p binding. On the other hand, two new rare allelic variants -ss102661458 and ss102661460- were identified in a chomosome of a patient with panic disorder each. Variant ss102661458 is located in the functional target site for miR-765, and variant ss102661460 in the functional target sites for miR-509 and miR-128, being the latter, a brain-enriched microRNA involved in neuronal differentiation and synaptic processing. This article also shows that these two variants significantly alter the miRNA-mediated regulation of *NTRK3* and result in the recovery of gene expression. Altogether, this study implicates miRNAs as key post-transcriptional regulators of *NTRK3* and for the first time provides a framework for allele-specific miRNA regulation of *NTRK3* in anxiety disorders.

Muiños-Gimeno M, Guidi M, Kagerbauer B, Martín-Santos R, Navinés R, Alonso P, Menchón JM, Gratacòs M, Estivill X, Espinosa-Parrilla Y.
Allele variants in functional MicroRNA target sites of the neurotrophin-3 receptor gene (NTRK3) as susceptibility factors for anxiety disorders.
Hum Mutat. 2009 Jul;30(7):1062-71.

B. Analysis of miRNAs at the genomic level

The plan for this second part of the results was to analyse the possibility that changes in miRNAs function or dosage, due to genetic variation in miRNAs or in their regulatory regions, could contribute as a source of phenotypic diversity to the susceptibility to psychiatric disease, in particular PD or OCD. This second chapter of the results consists of two articles, the first describing the design of a panel of SNPs in miRNA regions and the second including results of association analyses on PD, as well as a series of results of association studies in OCD.

Considering the organisation of miRNAs in transcriptional units, the first article presented in this section shows that thirty-seven percent of the miRNAs were inside known protein-coding genes, which were significantly associated with biological functions regarding neurological, psychological or nutritional disorders. In addition, the characterisation of the SNP coverage in miRNA regions revealed a lower SNP density in these regions in comparison to the average of the genome. Further genotyping of 340 unrelated Spanish individuals and comparison of the minor allele frequencies between the Spanish and HapMap population samples revealed two miRNA regions showing geographical allelic frequency variation among the four HapMap populations and confirmed the applicability of this SNP panel to the study of complex disorders among the Spanish population.

The second article describes the application of the previously described SNP panel to the study of PD. In this paper, case-

control studies in a Spanish population of patients with PD, with or without accompanying agoraphobia, and controls revealed association for SNPs at several miRNA regions -miR-22, miR-138-2, miR-339, miR-488 and miR-491- with different PD phenotypes. Moreover, association with AAO for a SNP tagging miR-148a was also found. Replication studies in a population from Finland and Estonia also showed association of miR-22, miR-138-2, miR-148a, miR-339, and miR-488 with different phenotypes related to anxiety and PD. These associations, however, failed to pass correction for multiple testing. In this second paper we further evaluated if these associated miRNAs were regulating any PD candidate gene. Functional studies showed that miR-22, miR-138-2, miR-148a and miR-488 cause an important repression of several genes that had been previously associated with PD such as *BDNF*, *CCKBR*, *GABRA*6, *HTR2C*, *MAOA*, *POMC* and *RGS2*.

Finally, this second chapter of the results section also includes the results obtained from the application of the miRNA-SNP panel to the study of another anxiety disorder, namely, OCD in a Spanish population. For this part, association analyses yielded a total of 3 miRNA regions associated with OCD, 6 miRNA regions associated with different subphenotypes of this disorder and 4 miRNA regions associated with AAO. Remarkably, six out of the nine miRNAs associated with OCD or OCD subphenotypes are brain expressed. Finally, in order to identify common genetic factors for OCD and PD, we studied the intersections of nominal associations between both disorders and found several shared associations between them.

Muiños-Gimeno M, Montfort M, Bayés M, Estivill X, Espinosa-Parrilla Y.
Design and evaluation of a panel of single-nucleotide polymorphisms in microRNA genomic regions for association studies in human disease.
Eur J Hum Genet. 2010 Feb;18(2):218-26. Epub 2009 Oct 7.

Human miR-22, miR-138-2, miR-148a and miR-488 are associated with panic disorder and regulate several anxiety candidate genes and related pathways

[1,2*]Margarita Muiños-Gimeno, [1,2*]Yolanda Espinosa-Parrilla, [1,2]Monica Guidi, [1]Birgit Kagerbauer, [3]Tessa Sipilä, [4]Kristi Petai, [5]Eduard Maron, [6,7]Rocío Martin-Santos, [6]Ricard Navinés, [1]Mònica Gratacòs, [5]Andres Metspalu, [4]Iiris Hovatta, [1,2]Xavier Estivill

[1]Genes and Disease Program, Centre for Genomic Regulation (CRG) and Public Health and Epidemiology Network Biomedical Research Center (CIBERESP), 08003 Barcelona, Spain.

[2]Experimental and Health Sciences Department, Pompeu Fabra University, Barcelona, Catalonia, Spain

[3] Research Program of Molecular Neurology, Biomedicum, University of Helsinki, Finland

[4] Institute of Molecular and Cell Biology, University of Tartu, Tartu, Estonia.

[5] Department of Psychiatry, University of Tartu, Tartu, Estonia

[6]Neuropsychopharmacology Programme, IMIM-Hospital Mar, Barcelona, Catalonia, Spain

[7]Psychiatry Service, Neurosciences Institute, Hospital Clínic, CIBERSAM, Barcelona, Spain

* These authors contributed equally to this work

Corresponding authors:

Yolanda Espinosa Parrilla, Center for Genomic Regulation (CRG), Genes and Disease Program, Dr. Aiguader 88, 08003 Barcelona, Spain. Tel. +34 93 3160233, Fax: +34 93 3160099, e-mail: yolespinosa@gmail.com

Xavier Estivill, Center for Genomic Regulation (CRG), Genes and Disease Program, Dr. Aiguader 88, 08003 Barcelona, Spain. Tel: +34 93 3160159, Fax: +34 93 3160099, e-mail: xavier.estivill@crg.cat

Abstract

Involvement of microRNAs in the control of neuronal differentiation and synaptic plasticity suggests a role for these molecules in the aetiology of psychiatric disorders. To evaluate microRNAs as candidate genes for panic disorder we have performed association studies using a panel of 768 SNPs at human microRNA regions (326 human miRNAs, MiRBase 7.1). Case-control studies in 203 patients with panic disorder, with or without accompanying agoraphobia, and 341 controls revealed two SNPs to be strongly associated with panic disorder, rs6502892 tagging miR-22 (p=0.00016) and rs11763020 tagging miR-339 (p=0.00007785). Several other SNPs at miR-138-2, miR-488 and miR-491 regions were associated with different panic disorder phenotypes, and rs73531, tagging miR-148a, was found to be associated with age at onset. Replication studies in a sample of 114 patients from Finland and 102 patients from Estonia also showed nominal associations with anxiety and panic disorder for several of these miRNAs. Funtional studies revealed that miR-138-2, miR-148a and miR-488 cause an important repression (30-60%) of panic disorder candidate genes *GABRA6*, *CCKBR* and *POMC*, respectively, and that miR-22 at least regulates four other candidate genes: *BDNF*, *HTR2C*, *MAOA* and *RGS2*. Furthermore, transcriptome analysis of neuroblastoma cells transfected with miR-22 and miR-488, two brain expressed microRNAs, showed altered expression of a subset of predicted target genes for these miRNAs as well as genes that may be affecting in several ways not only the fine functioning of the central nervous system but also other physiological pathways linked to de development of panic disorder. We report for the first time a possible implication of miRNAs in the aetiology of panic disorder.

Introduction

The Statistical Manual of Mental Disorders (DSM-IV[1]) defines panic disorder (PD) as the spontaneous, unexpected occurrence of panic attacks followed by persistent concern, worry, and anxiety about the recurrence of these episodes. Panic attacks often mimic symptoms of physical complaints such as a heart attack or other life-threatening illnesses and typically occur spontaneously, with no apparent trigger. PD is comprised within the category of anxiety disorders that, among others, also include generalized anxiety disorder (GAD), agoraphobia and obsessive-compulsive disorder (OCD). Furthermore, PD can manifest with or without accompanying agoraphobia, which can also occur without PD. Life prevalence of the disorder is 1-2% and twice as many women suffer from PD if compared to men.[2] It has a young adult age at onset (AAO), usually in the late twenty's, and the presence of agoraphobia and AAO under 20 are thought to conform more severe variants of the disease.[3,4]

The search for susceptibility genes for PD has proven to be extremely difficult, in part, due to the multifactorial and polygenic origin of the disorder but also due to the high level of regulatory control and gene interactions to what human brain and behaviour are exposed. However, some candidate genes and positive associations have raised when studying molecules involved in the neurotransmitter synthesis and degradation pathways, such is the case of the serotonin receptors *HTR2A, HTR2C* and *HTR1A,* [5-7] the serotonin transporters *SLC6A4*, *SLC6A1* and *SLC6A2,* [8,9] the Cholecystokinin receptors *CCKAR* and *CCKBR,* [10] or some neurotransmitters degradation enzymes, such as the monoamine oxidase A gene (*MAOA*). [11] In addition, other genes involved in signalling pathways, such as the regulators of G protein signalling *RGS2* and *RGS7* [12,13] or more recently, molecules involved in neuronal growth and differentiation have also been associated with PD. [2]

From a physiological point of view, stress has shown to have a critical role in the development of anxiety disorders, being even able to influence the onset of the disease.[14] Stress may induce anxiety, at least partially, through mechanisms related to neural plasticity: synaptic connections in the brain undergo experience-dependent functional or morphological changes through complex pathways that are not yet fully understood but for which microRNAs (miRNAs) may

have a critical role mediating dendritic patterning and synaptogenesis.[15,16] miRNAs are endogenous small non-coding RNAs that regulate gene expression by means of partial complementarity to miRNA binding sites located in the 3' untranslated regions of the target genes[17] and that have emerged as key regulators of almost every biological process including developmental timing, cell differentiation and morphogenesis[18]. The distribution of miRNAs in the developing and adult nervous system is spatially and/or temporally restricted,[19] indicating that they may contribute to the fine-tuning of neuronal gene expression. Increasing evidence supports that miRNAs are major contributors to phenotypic diversity in human populations[20] and they may thus have a role in the pathophysiology of several disorders. In fact, miRNAs have already been involved in several disorders of the central nervous system such as Alzheimer's disease,[21,22] Parkinson disease,[23] schizophrenia[24,25] aggressive human behaviour[26] and in Tourette Syndrome.[27]

All in all, we propose that the particular miRNA enrichment in the brain and the capability of these molecules to finely modulate many genes simultaneously points at them as good candidates for the study of the genetic susceptibility to many psychiatric disorders. With this aim in mind, we had previously constructed a panel of informative SNPs in miRNAs-containing regions suitable for the study of complex disorders [28] and, in this article, we have performed a case-control study on panic disorder. Association analyses together with functional studies indicate, for the first time, that at least four miRNAs, miR-22, miR-138-2, miR-148a and miR-488, may be involved in the pathophysiology of PD.

Material and methods

Spanish Sample

Between 2001 and 2006, 203 consecutive adult Spanish Caucasian outpatients (mean age 35.89 ± 9.74 years, 151 females; mean age of PD onset 29.44 ± 9.44 years) with PD recruited from the Psychiatry outpatient unit in Hospital del Mar (Barcelona) were studied. The diagnosis of PD was independently assigned by two senior psychiatrists using the Structured Clinical Interview for DSM-IV Disorders - Clinician Version (SCID-CV). Concurrent agoraphobia was present in 135 (66.5%) of the patients. Exclusion criteria were: age under 18, organic brain syndromes, psychoactive abuse disorders (except nicotine abuse), any other DSM-IV Axis I comorbid disorder apart from other anxiety disorders and life prevalence of mood disorder and severe organic or neurological pathology including partial epilepsy. A clinician administered version of the Panic and Agoraphobia scale,[29] the Fear Survey Schedule (F100) of Wolpe[30] and the 21-item HDRS[31] were used to assess the severity of panic and agoraphobia, fears and depressive symptoms, respectively. The control group consisted of 341 healthy blood donors (mean age 39.76 ± 11.92 years; 140 females) recruited from the Blood and Tissue Bank of the Catalan Health Service, all were of Spanish origin (Catalonia, at the north-east of Spain) and gave informed consent. After obtaining written informed consent approved by the hospital ethic committee, genomic DNA of control and patients was extracted from peripheral blood lymphocytes using automatic DNA extraction and Standard protocols. To detect population admixture in our control and PD sample, a structured association method was used to further test each sample set for stratification between cases and controls as previously described. [32]

Finnish Sample

The Finnish sample derives from the Finnish population-based Health 2000 Study carried out in 2000-2001 and has been previously described.[33] Briefly, the 12-month prevalence of DSM-IV mental disorders was estimated from a representative sample *(*n = 6005) of the Finnish general adult (≥ 30 years of age) population with the Composite International Diagnostic Interview (CIDI). From this cohort, 320 individuals with anxiety (mean age 49.62 ± 12.67 years; 204 females) were selected according to DSM-IV anxiety disorder diagnosis but also including

individuals with sub-threshold diagnoses as defined by the CIDI (extended diagnostic group). Among that sample, 106 individuals (mean age 50.48 ± 12.52 years, 65 females) were diagnosed as GAD (including 31 sub-threshold cases) and 114 patients (mean age 46.41 ± 11.27 years; 77 females) as PD. A total of 668 Finnish controls (mean age 49.61 ± 12.62 years; 426 females) were also analyzed. The work was approved by the institutional ethics committee of the National Public Health Institute, and written consent was obtained for all individuals.

Estonian Sample

The Estonian sample consisted of 102 outpatients with PD (mean age 38.7 ± 12.6 years, 87 females; mean age of PD onset 31.0 ± 11.6 years) recruited at the Psychiatry Clinic of the Tartu University Hospital, Estonia. Diagnoses of PD according to DSM-IV criteria were verified with the Mini International Neuropsychiatric Interview (M.I.N.I. 5.0.0) and substantiated by psychiatric history and medical records. Concurrent agoraphobia was present in 60 (59%) of the patients. The PD patients with history of comorbidity with mood disorders or with current social anxiety disorder were included in the study, but no other history or current psychiatric comorbidity was allowed. PD was considered the primary diagnosis in all patients, based on being the main diagnosis at the time of the investigation and/or an earlier onset of PD in the course of illness. The Human Studies Ethics Committee of the University of Tartu approved the study protocol, and all participants provided written informed consent.

Genotyping of the miRNA SNP panel

The selection, design and validation of the 768 SNPs of the panel have been previously described,[28] Genotyping was performed using the GoldenGate assay on an Illumina BeadStation 500G in accordance with the manufacturer's standard recommendations. Allele calling were performed using the BeadStudio program (Illumina Inc, San Diego, CA). All SNPs were examined for standard quality control after genotyping as described before,[28] this evaluation yielded a set of 714 SNPs typed (supplementary table 1) which corresponds to a final cleaned data of 92.97%. Genotypes for the non-excluded SNPs were consistent with Hardy-Weinberg equilibrium (HWE) except for 2 SNPs that were eliminated. Both genotype concordance and correct Mendelian inheritance were verified.

Genotyping of the SNPs for the replication analysis in the Finnish sample

Genotyping for the 19 out of the 768 SNPs of the miRNA SNP panel that were selected for the replication study in the Finnish sample was performed with Sequenom MassARRAY genotyping technology (Sequenom, San Diego, California) with iPLEX chemistry in stage I and iPLEX Gold chemistry in stage II. Assay conditions were as recommended by the manufacturer. Quality control measures, verification of Hardy-Weinberg equilibrium and confirmation of Mendelian inheritance were performed as previously described.[33]

Genotyping of the SNPs for the replication analysis in the Estonian sample

Genotyping in the PD Estonian sample was performed using the Golden Gate assay of VeraCode technology using the BeadXpress Reader System according to the manufacturer's protocol (Illumina, San Diego, CA). Genotyping of 1000 Estonian controls was performed with Illumina Infinium II technology on the HumanCNV-370Duo Beadchip as previously described.[34]

Association Studies

Association analyses were assessed by means of multivariate regression methods under dominant, codominant, recessive, and additive models adjusting for sex and categorizing cases according to clinical status and AAO. We estimated the crude odds ratio (OR) and 95% confidence intervals (95% CI). Correction for multiple testing according to permutation test was applied taking into account 714 independent SNPs and setting the corrected level of significance to 0.0000852. To correct for the use of the four different genetic models, a simulation study was carried out obtaining a factor of correction of 2.5 as the effective number of tests performed.[35] Using these criteria, the corrected level of significance was set up equal to 0.00003408. In the case of the replication study and according to Bonferroni correction the level of significance was settled to 0.0026 for the 19 analyzed SNPs and 0.001 when also considering the use of four genetic models (19 SNPs x 2.5 effective tests = 47.5 comparisons). The association analyses were performed with the SNPassoc R package. [35] Linkage disequilibrium (LD) between polymorphisms was evaluated with the Haploview software version 3.2 (http://www.broad.mit.edu/mpg/haploview/). For the haplotype estimations the haplo.stats R package was used.

PCR and Direct Sequencing

Fragments ranging from 600 bp to 748 bp were amplified by polymerase chain reaction (PCR) in a final volume of 25 µl using standard protocols and the primers listed in the supplementary table 2. The sequencing reaction was performed using the ABI PRISM® BigDye Terminator v3.1 Sequencing Kit (Applied Biosystems, Foster City, CA, USA) and PCR products were purified with ExoSAP-IT (usb, Cleveland, OH, USA) following standard procedures. Sequence purification was performed with Millipore Montage seq96 (Millipore Corporation, Billerica, MA, USA). Sequencing was performed using an ABI PRISM® 3730*xl* DNA Analyzer (Applied Biosystems, Foster City, CA, USA).

Firefly luciferase constructs

The 3'UTRs of the mRNAs corresponding to the twelve analyzed genes (*ADORA 2A, BDNF*, *CCKAR*, *CCKBR*, *CRHR2*, *GABRA6, HTR2C*, *MAOA, NTRK3, POMC, RGS2* and *SCL6A2*) were PCR-amplified from genomic DNA with *PfuTurbo*® DNA polymerase (Stratagene, La Jolla, CA), using primers containing an *Xba*I restriction site at the 5' end (supplementary table 2). PCR fragments ranging 288 to 2969bp were purified, *Xba*I-digested and cloned into an *Xba*I site located downstream of the firefly luciferase reporter gene in the pGL4.13 vector (Promega Corporation, Madison, WI). Constructs were propagated in *E. coli* One Shot® TOP 10 cells (Invitrogen) and the authenticity and orientation of the inserts was confirmed by sequencing.

Cell culture and transfection

HeLa cells and SH-SY5Y neuroblastoma cells were maintained in Dulbecco's Modified Eagle's Medium supplemented with 10% fetal bovine serum (heat inactivated for 45 min at 56°C prior to use in the case of SH-SY5Y cells), 2 mM L-glutamine, 100 units/ml penicillin and 100 µg/ml Streptomycin (GIBCO ™, Invitrogen). Transfection with small RNAs was performed with Lipofectamine 2000 (Invitrogen) and optimized using a fluorescein-labeled double-stranded RNA oligomer. Small RNAs that mimic endogenous mature miRNAs and the related negative controls (miRIDIAN™ miRNA Mimics and miRIDIAN™ miRNA Mimics Negative Controls #2 and #4) were obtained from Dharmacon Inc.

Luciferase activity assay

HeLa cells were seeded at 1.3 x 10^4 cells/well in 96-well plates and cotransfected 24 h later with the Firefly reporter constructs described above or the empty pGL4.13 vector (10-24 ng), the Renilla reporter plasmid pGL4.75 (3 ng) and the appropriate miRNA mimic (10 nM). The activity of Firefly and Renilla luciferases was determined 24 h after transfection using the Dual-Glo™ Luciferase Assay System (Promega). Relative reporter activity was obtained by normalization to the Renilla luciferase activity. In order to correct for vector-dependent unspecific effects, each relative reporter activity was normalized to the empty vector cotransfected with the corresponding miRNA. Results were then compared to the mean of the two negative controls. Each experiment was done in triplicate and at least three independent experiments were performed for each miRNA. For the validation of the significant results two more series of experiments were performed at two different concetrations of miRNA mimic (10 nM and 30 nM). Data are reported as means ± S.E. Statistical significance was determined using Student's *t* test ($p<0.05$).

Whole-genome expression analysis in SH-SY5Y cells using beadchip microarrays

Undifferentiated SH-SY5Y cells were plated at 2 x 10^5 cells per well in 6-well plates and were transfected 24 h later with 100nM miRNA mimic (miR-22, miR-488 or the related negative controls); 48 h after transfection cells were lysed and total RNA was extracted using the using the RNeasy® Mini Kit (QIAGEN). Starting from 200 ng total RNA, biotin-labeled cRNA was synthesized and hybridized on HumanRef-8 BeadChips from Illumina, which target 24,500 well-annotated RefSeq transcripts, according to the manufacturer's instructions. Four independent experiments were performed. Data were analyzed using the Array File Maker (AFM) 4.0 software package.

Computational methods

Two web-based miRNA target prediction methods were used: miRanda (www.microrna.org, 2005 release) and TargetScan (www.targetscan.org, release 5.1) Genomic coordinates are according to the human assembly release of March 2006 (H. sapiens hg18). Sequence analysis was performed using the 4peaks software (http://mekentosj.com/4peaks/) and Multalin

(http://prodes.toulouse.inra.fr/multalin/multalin.html). Primers were designed using the Primer3 software (http://frodo.wi.mit.edu/cgi-bin/primer3/primer3.cgi). The list of the different genes associated with panic disorder was obtained from the association database at the nih website (http://geneticassociationdb.nih.gov/). Pathway analysis from microarray expression data was performed with the Ingenuity Pathway Analysis Software (IPA) version 6.3 (www.ingenuity.com).

RESULTS

Association of miRNA regions with PD in a Spanish population

Case control studies in 203 Spanish patients with PD and in 341 controls were performed for the 712 SNPs in the miRNA SNP panel that passed QC criteria and were shown to be in Hardy-Weinberg equilibrium ($p<0.00008525$). Association analyses were performed under different models using logistic regression and adjusting for sex. Table 1 shows the most significant associations found for SNPs in miRNA regions with PD. The strongest association was found for rs11763020, which remained significant under the recessive model (unadjusted $p= 0.00007785$, OR= 9.45) after correction for the 712 SNPs tested ($p<0.00008525$). However, the significance was lost when correction for the 2.5 effective tests was performed ($p<0.000034$). The calculated odds ratio suggests that the homozygous TT genotype, with a frequency of 1.2% in the controls and 6.4% in the PD sample, could confer a risk in the susceptibility to PD. When looking at other SNPs in the miR-339 region we found that one of the two SNPs tagging the promoter region of the miR-339 host gene (*C7orf50*) was associated with PD but none of the other three SNPs tagging miR-339 were (Table 1). Besides miR-339, two SNPs tagging miR-22 were also found to be associated with PD; rs6502892 under the codominant (unadjusted, $p=0.00016$, OR=0.45) and the recessive (unadjusted, $p=0.00115$, OR=0.34) models and rs8076112 (unadjusted $p=0.00581$, OR=10.59), under the recessive model. It is worth remarking the case of miR-491 for which three out of the four SNPs tagging this miRNA but none of the two SNPs tagging its host gene, *KIAA1797*, were found to be associated with PD (Table 1). These associations, however, failed to pass correction for multiple testing.

Besides PD, 67% of the patients presented with agoraphobia, therefore and in order to take into account this phenotype, association analysis was repeated after stratifying for the agoraphobic phenotype. As shown in table 2, associations for the SNPs tagging miR-339 and miR-22 decreased when considering separately the two groups of patients (either presenting or not with agoraphobia). In contrast, in the case of miR-491, the associations for the three SNPs tagging this miRNA increased for the non-agoraphobic phenotype (table 2). Interestingly, when stratifying by agoraphobia, new associations that pointed towards two other miRNAs, miR-138-2

and miR-488, appeared (table 2): One SNP tagging miR-138-2, rs12921781, that was associated only nominally with the complete PD sample, showed an unadjusted p=0.00075 under the codominant model with PD in the absence of agoraphobia, and similarly, two SNPs tagging miR-488 increased their association but, in this case, with the agoraphobic phenotype: a nominal association for rs2076072 appeared and rs4652220 increased its p-value to unadjusted p=0.00129 and p=0.00253 under the codominant and dominant models, respectively (Table 2).

Finally, since PD presents a particular distribution of age at onset (AAO), with two peaks one in late adolescence and a second peak in the mid 30s, [35] we analyzed the possible association of these miRNA regions with AAO, the association analysis showed that rs73531, tagging the intergenic miRNA miR-148a, was associated with AAO under the recessive model (unadjusted p= 0.001235) with an average AAO of 23 years for the GG homozygotes and 30 years for the AG heterozygotes and AA homozygotes (Table 3).

Haplotype and interaction analyses between the different SNPs showing association in our PD samples were also performed but no increase in the significance level was observed in any case.

Replica of the miRNA associations with PD in North European populations

After the finding of 6 miRNA regions associated with PD in the Spanish population, we wanted to replicate the results in a different population. We constructed a panel of 19 SNPs including all the SNPs analyzed in these 6 miRNA regions and genotyped a Finnish sample of 668 controls and 320 patients with anxiety, including PD (114 patients) and GAD (106 patients) among other anxiety disorders. Association analyses were performed under different models using logistic regression and adjusting for sex. The analyses showed one SNP, rs2076072 in the miR-488 region, to be associated with anxiety (unadjusted p= 0.03805, OR= 0.73, dominant model). When assessing the association with GAD, three SNPs were found to be associated with this phenotype under the dominant model: rs1057558, in the miR-339 region, which showed an unadjusted p= 0.00117 (OR=0.49), rs6502892 in the miR-22 region with an unadjusted p= 0.03720 (OR=1.56) and rs4722551 in the miR-148a region (P= 0.00602, OR=0.53). Finally, another SNP in the miR-148a region, rs735316, was found to be associated with PD under the

recessive model (unadjusted p= 0.00983, OR= 2.05). Further, we performed a second replication study on a different North European sample and genotyped 102 PD patients from Estonia using the panel of 19 selected SNPs. Genotypic data from 1000 Estonian controls coming from a whole genome scan were used for comparisons. Even though the whole genome panel only included 6 out of the 19 SNPs of interest (supplementary table 3), a positive association with PD was identified for rs4790814 tagging miR-22 (P= 0.04140) under the recessive model.

Since these associations did not have enough power to pass multiple test corrections, we performed a joint analysis in the north European populations using the PD samples from Estonia and Finland. All SNPs with exception of four (rs12684948, rs7022666, rs4722551 and rs12920659) showed non-significantly different allele frequencies between both populations (supplementary table 3) and were thus used in the joint case-control study. Even though the significance of the association did not increase, two nominal associations with PD were identified, rs1436424 tagging miR-138-2 (P= 0.01296, codominant model) and rs735316 tagging miR-148a (p= 0.01024, recessive model).

Resequencing of miRNAs in PD patients

In order to search for functional allele variants that could be responsible for the associations found with PD we sequenced the precursor region of miR-22, miR-138-2, miR-339, miR-488 and miR-491 together with their ~1kb upstream and downstream sequences in genomic DNA from 90 patients with PD. Twelve rare allele variants (MAF<0.05) and 8 common variants, 2 of which were new, were found together with two new rare and one common short deletions (Table 4). Even though 6 out of the 8 common allele variants were in LD (D'=1) with the SNPs that we had previously found to be associated with PD, none of them is located in neither the mature nor the pre-miRNA sequences and therefore no obvious disruption of miRNA function could be assigned.

Six candidate genes for anxiety disorders are repressed by miRNAs associated with PD

Next, we wanted to evaluate if there were candidate genes for PD that might be regulated by any of the miRNAs that were associated with PD in the Spanish and North European sample (miR-22, miR-138-2, miR-148a, miR-339 and miR-488). Using the TargetScan and miRanda programs we predicted mRNAs that were targeted by these five miRNAs and crossed this information with a list of 40 genes that had been previously associated with PD, according to the association database from the nih, or that had proven roles in anxiety related pathways. Merging of information provided with a list of twelve candidate genes for anxiety disorders that were also predicted to be regulated by at least one of these five miRNAs (Table 5). To functionally validate these predictions, the 3'UTRs of the twelve candidate miRNA targets were cloned immediately downstream of the firefly luciferase open reading frame in a pGL4.13 plasmid. The five miRNAs and two non-targeting controls were tested on all twelve constructs and the empty plasmid by cotransfection into HeLa cells. A renilla reporter plasmid pGL4.75 was used to control for transfection efficiency. As shown in figure 1 and table 5, a statistically significant reduction of the luciferase activity was observed for seven mRNAs (*BDNF*, *CCKBR, GABRA6*, *HTR2C, MAOA*, *POMC* and *RGS2*) when cotransfected with at least one of the five associated miRNAs tested. Remarkably, miR-22 was the miRNA with the highest number of functional targets with four genes being potentially repressed by this miRNA: the neurotrophic factor *BDNF*, the serotonin receptor *HTR2C*, the monoamine oxidase A, *MAOA*, and the regulator of G protein signalling *RGS2* (figure 1 and table 5). miR-22 was also the miRNA that produced the strongest down-regulation causing a repression of about two thirds of *RGS2* and rendering a 33.7% of total luciferase activity for the *RGS2* construct. In contrast, no candidate gene appeared to be regulated by miR-339 even though this miRNA was predicted to regulate four genes. Finally, miR-138-2, miR-148a and miR-488, which were predicted to regulate four, three and six mRNAs, respectively, were found to strongly regulate at least one gene: miR-148a and miR-488 reduced to half the associated luciferase activity of the *CCKBR and POMC* constructs, respectively, and miR-138-2 repressed GABRA6 in almost one third (table 5, figure 1).

Transcriptome analysis of SH-SY5Y cells transfected with miR-22 and miR-488

To gain insight into the role of some of these miRNAs, we overexpressed some of these miRNAs in neuroblastoma SH-SY5Y cells to see the effects on the transcriptome by whole genome expression microarrays (Illumina's HumanRef-8 v3.0 beadchips). miRNAs, miR-22 and miR-488, were selected because they were the ones predicted to regulate the higher number of candidate genes (6 out of 12 each miRNA) and, in fact, miR-22 was proven to strongly repress at least four target mRNAs. Considering a fold-change (FC) cutoff of 1.2 and a q-value <5, we could identify a total of 1209 deregulated genes after miR-22 overexpression (579 downregulated and 630 upregulated) and 182 deregulated genes after miR-488 overexpression (122 downregulated and 60 upregulated). The top ten upregulated and downregulated genes are listed in Table 6. Interestingly, among them there are several genes that had been previously implicated in the aetiology of anxiety such as the regulators of G protein signalling *RGS2* and *RGS4* as well as the galanin prepropeptide *(GAL)*. The most striking example is *RGS2* (FC = -2.086), a gene that has been associated with PD and that we proved to be strongly repressed by miR-22. Further, microarray results were compared with target predictions in order to check whether some of the deregulated genes could be potential direct targets of these miRNAs. As shown in table 6, according to TargetScan all top ten downregulated genes in the case of miR-488 (except for *RN7SK* which is not considered in TargetScan database) and more than half for miR-22 are predicted to be direct targets, in contrast with only two out of the top-ten upregulated genes by miR-488, indicating an enrichment for predicted target genes in the group of down-regulated genes.

Further, the genes deregulated upon either miR-22 or miR-488 overexpression and the corresponding expression values were uploaded into the Ingenuity Pathway Analysis software. The program was interrogated about biological functions, canonical pathways and molecular networks that could be affected by the deregulation. In the case of miR-22 we found significant associations (-log(p-value) >3) with canonical pathways related to Aminoacyl-tRNA Biosynthesis (p value=2.12E-04), Lysine Degradation (4.05E-03) and Citrate Cycle (4.6E-03). In the case of miR-488 the most significant associations were found for Pentose Phosphate Pathway (2.93E-

02) and Corticotropin Releasing Hormone (CRH) Signalling (1.81E-02, supplementary Figure 1). Interestingly, CRH is downstream from the pro-opiomelanocortin gene *POMC*, which we found to be strongly regulated by miR-488, in an important pathway that is activated in response to stress.

DISCUSSION

Increasing evidence indicates that genetic variation in regulatory regions rather than in protein-coding sequences could be a major contributor to phenotypic diversity in human populations.[36,37] This may be particularly true in the case of psychiatric disorders; human brain is subject to a great amount of gene regulation, where sophisticated circuitries that remain poorly understood are responsible for the different feelings, mood changes or even actions in humans. In fact, changes in regulatory elements leading to small variations in the dosage of proteins involved in neuronal pathways may lead to a disruption in the fine-tuned equilibrium of complex brain functions, and contribute to the development of psychiatric disorders. Recently, miRNAs have emerged as important genomic regulators with a key role in the developing and in the adult nervous system, contributing to the correct calibration of neuronal gene expression. We have analyzed miRNA regions as PD candidate genes by performing association studies in a Spanish population and have found association for miRNAs miR-22, miR-491 and miR-339 with PD. The analyses performed taking into account the agoraphobic phenotype and age at onset also revealed association for miR-138-2, miR-488 and miR-148a. Associations alone did not result to be conclusive due to the fact that they were not strong enough to resist correction for multiple testing. Modest associations, however, are repeatedly identified in most studies on PD, a disorder for which multiple genes of small effect in interaction with each other and/or with non-genetic factors have been proposed to underlie its susceptibility.[38] In fact, results on whole genome association studies point out that susceptibility alleles are likely to be modest in effect size and require large samples for detection.[39] This model tallies with the results obtained from the North European populations, where associations for the replica, albeit being identified for miR-22, miR-339, miR-488, miR-148a and miR-138-2, were weak. Nonetheless, the importance of performing functional studies rather than replicating associations in other cohorts is growing nowadays, as recent association studies of complex disorders do not have enough power and have failed to replicate.

In order to search for possible causal variants that might be in linkage disequilibrium with the associations found, we resequenced the pre-miRNA sequences of the six associated miRNAs

as well as their flanking regions. This analysis identified 9 common and 14 rare allelic variants, nevertheless, none of them was located within the mature or pre-miRNA sequences and no effect in the targeting spectrum of the studied miRNA could be predicted to occur. However, it is important to take into account that miRNA biogenesis still is far to be understood and that many gaps remain in the comprehension of the transcriptional and post-transcriptional mechanisms involved. It has been demonstrated, nonetheless, that efficient processing of pri-miRNAs by Drosha requires the presence of extended single-stranded RNA on both 3'- and 5'-ends of the pre-miRNA hairpin molecule. The length and secondary structure of these ssRNA motifs is of high importance if processing is to take place at all and thus allelic variants in the proximity of a pre-miRNA may alter its expression, transport, or stability. For instance, Drosha will not cleave efficiently in the absence of an additional helical RNA turn beyond the cleavage site.[40] However, even though posttranscriptional regulation for miR-22, miR-138-2 and miR-488 is known to occur[41,42] the mechanisms regarding miRNA processing are still unknown and therefore, it is difficult to raise any conclusion on the functionality of these variants. In any case, effects derived from variants in the proximity of pre-miRNA sequences, if any, would be related to changes in miRNA dosage but not in the targeting spectrum of the studied miRNA. Indeed, variants affecting miRNA expression and processing could explain the neuronal disequilibrium proposed for psychiatric disorders, where correct dosage could be crucial.

Regardless the genetic mechanism involved in these associations, the development of the phenotype could depend upon the expression and activity of these miRNAs. Some of these miRNAs are known to be expressed or have important functions in brain, miR-22, for example, is expressed ubiquitously in diverse tissues including pituitary and midbrain; miR-488 is a brain enriched miRNA that is more abundantly expressed hippocampus and cerebellum and miR-138 is highly enriched in brain, including expression in the frontal cortex, hippocampus and midbrain. Further, miR-138 is localized within dendrites and it is known to negatively regulate the size of dendritic spines in rat hippocampal neurons.[43] As the type of changes that may be underlying the associations identified would not probably alter the target spectrum of the analysed miRNAs but rather could have an effect in miRNA dosage, we decided to identify candidate genes for

panic disorder among those genes predicted to be targeted by the associated miRNAs and functionally validate the predictions through miRNA overexpression in a Hela cell system. We could identify a reduction of the luciferase activity compatible with a putative repression of *RGS2, BDNF, HTR2C*, and *MAOA* by miR-22, *POMC* by miR-488, *GABRA6* by miR-138-2 and *CCKBR* by miR-148a. All of these genes have been implicated in the aetiology of anxiety disorders often in a dosage dependent manner. For instance, serotonergic pathways have been involved in the pathogenesis of anxiety disorders, mainly because of the observation that patients with anxiety disorders respond well to serotonergic medications and because the occurrence of panic attacks has been reported after administration of serotonergic agonists.[44,45] On the other hand, POMC is the precursor molecule of several important components of the hypothalamic-pituitary-adrenal axis, which has been involved in the neurobiology of mood and anxiety disorders.[46] Interestingly, the expression of *RGS2* has been demonstrated to be a quantitative trait,[47] for which association with a haplotype within *RGS2* 3'UTR has been reported.[13,48] Moreover, RGS2 knock-out mice show more anxious behaviour than their wild-type counterparts.[49] Remarkably, expression of *RGS2* was significantly reduced when overexpression of miR-22 was simulated in neuroblastoma cells, more interestingly another gene that was downregulated is *ASCL1*, a gene that has been demonstrated to be essential for the development of central serotonergic neurons and that has been proposed as a candidate for a rare disorder of the chemical control of breathing.[50] Other downregulated genes also have important neuronal functions such as *CHGA* with roles in neuroendocrine secretion[51] or the promotion of dendritic outgrouth by *NPTX2.*[52] Further, it is remarkable the finding of the deregulation of the corticotropin releasing hormone (CRH) signalling pathway associated with the overexpression of miR-488, which is a crucial pathway activated in response to stress and in which the miR-488 deregulated pro-opiomelanocortin gene *POMC* is an important molecule. These results strengthen the idea that alterations in the complex circuitry of gene regulation in which either miR-22 or miR-488 are involved may be affecting in several ways not only the fine functioning of the central nervous system but also other physiological pathways linked to de development of anxiety disorders and particularly PD. All these evidences put together fit with

the miRNA-induced differential dosage hypothesis that we propose here and points towards several miRNAs to be involved in the aetiology of PD.

Acknowledgements

We would like to thank L. Kananen for her help in the genotyping of the Finnish sample and A. Carreras, A. Puig, M. Montfort and C. García for their technical support. We would also like to thank all patients and blood donors participating in this study, as well as clinical collaborators that are not among the authors. This work was supported by the "Instituto Carlos III and Fondo de Investigaciones Sanitarias" [CIBER-CB06/02/0058, CIBER-SAM, FIS/ISCIII:P1052565, ISCIII:GO3/184], the "Fundació la Marató-TV3" [014331], the "Departament d'Universitats Innovació i Empresa, Generalitat de Catalunya" [2005SGR00008] and the European Union Sixth Framework Programme Integrated Project SIROCCO [Grant LSHG-CT-2006-037900]. The Spanish National Genotyping Center (CeGen) is supported by "Genoma España". M.M.G is a recipient of a FIS fellowship [FI05/0006]. Y.E.P was supported by the "Ramón y Cajal" Program [Spanish Ministry of Science and Innovation].

Legends to figures

Figure 1: Results of the luciferase-reporter assay used to test the interaction between the associated miRNAs and candidate genes for PD. This figure only shows the results for those genes that showed a significant reduction of the luciferase expression by at least one of the associated miRNAs. Ratios of the Firefly and Renilla luciferase luminescence are presented after normalization to the empty plasmid pGL4.13 and to the mean of two different mimic controls. Each experiment was done in triplicate and at least three independent experiments were performed for each miRNA tested, at a concentration of 10 nM. Data reported here are the means ± s.e of all experiments performed. Dark-colour bars correspond to those miRNAs found to significantly reduce the luciferase expression ($p<0.05$, Student's t test). For the validation of the significant results two more series of experiments were performed at two different concetrations of miRNA mimic (10 nM and 30 nM).

References

1 Association AP: *Diagnostic and Statistical Manual of Mental Disorders*. Washington, DC., American Psychiatric Press, 1994.

2 Gratacos M, Sahun I, Gallego X, Amador-Arjona A, Estivill X, Dierssen M: Candidate genes for panic disorder: insight from human and mouse genetic studies. *Genes Brain Behav* 2007; **6 Suppl 1**: 2-23.

3 Goldstein RB, Wickramaratne PJ, Horwath E, Weissman MM: Familial aggregation and phenomenology of 'early'-onset (at or before age 20 years) panic disorder. *Arch Gen Psychiatry* 1997; **54**: 271-278.

4 Noyes R, Jr., Crowe RR, Harris EL, Hamra BJ, McChesney CM, Chaudhry DR: Relationship between panic disorder and agoraphobia. A family study. *Arch Gen Psychiatry* 1986; **43**: 227-232.

5 Maron E, Nikopensius T, Koks S *et al*: Association study of 90 candidate gene polymorphisms in panic disorder. *Psychiatr Genet* 2005; **15**: 17-24.

6 Rothe C, Gutknecht L, Freitag C *et al*: Association of a functional 1019C>G 5-HT1A receptor gene polymorphism with panic disorder with agoraphobia. *Int J Neuropsychopharmacol* 2004; **7**: 189-192.

7 Unschuld PG, Ising M, Erhardt A *et al*: Polymorphisms in the serotonin receptor gene HTR2A are associated with quantitative traits in panic disorder. *Am J Med Genet B Neuropsychiatr Genet* 2007; **144B**: 424-429.

8 Strug LJ, Suresh R, Fyer AJ *et al*: Panic disorder is associated with the serotonin transporter gene (SLC6A4) but not the promoter region (5-HTTLPR). *Mol Psychiatry* 2008.

9 Thoeringer CK, Ripke S, Unschuld PG *et al*: The GABA transporter 1 (SLC6A1): a novel candidate gene for anxiety disorders. *J Neural Transm* 2009; **116**: 649-657.

10 Kennedy JL, Bradwejn J, Koszycki D *et al*: Investigation of cholecystokinin system genes in panic disorder. *Mol Psychiatry* 1999; **4**: 284-285.

11 Deckert J, Catalano M, Syagailo YV *et al*: Excess of high activity monoamine oxidase A gene promoter alleles in female patients with panic disorder. *Hum Mol Genet* 1999; **8**: 621-624.

12 Hohoff C, Neumann A, Domschke K *et al*: Association analysis of Rgs7 variants with panic disorder. *J Neural Transm* 2008.

13 Leygraf A, Hohoff C, Freitag C *et al*: Rgs 2 gene polymorphisms as modulators of anxiety in humans? *J Neural Transm* 2006; **113**: 1921-1925.

14 Lupien SJ, McEwen BS, Gunnar MR, Heim C: Effects of stress throughout the lifespan on the brain, behaviour and cognition. *Nat Rev Neurosci* 2009; **10**: 434-445.

15 Schratt GM, Tuebing F, Nigh EA *et al*: A brain-specific microRNA regulates dendritic spine development. *Nature* 2006; **439**: 283-289.

16 Smalheiser NR, Lugli G: microRNA Regulation of Synaptic Plasticity. *Neuromolecular Med* 2009.

17 Bartel DP: MicroRNAs: genomics, biogenesis, mechanism, and function. *Cell* 2004; **116**: 281-297.

18 Stark A, Brennecke J, Bushati N, Russell RB, Cohen SM: Animal MicroRNAs confer robustness to gene expression and have a significant impact on 3'UTR evolution. *Cell* 2005; **123**: 1133-1146.

19 Kosik KS: The neuronal microRNA system. *Nat Rev Neurosci* 2006; **7**: 911-920.

20 Feng Y, Vetro A, Kiss E *et al*: Association of the Neurotrophic Tyrosine Kinase Receptor 3 (NTRK3) Gene and Childhood-Onset Mood Disorders. *Am J Psychiatry* 2008; **165**: 610-616.

21 Lukiw WJ: Micro-RNA speciation in fetal, adult and Alzheimer's disease hippocampus. *Neuroreport* 2007; **18**: 297-300.

through regulation of beta-site amyloid precursor protein-cleaving enzyme 1. *J Neurosci* 2008; **28**: 1213-1223.
23 Kim J, Inoue K, Ishii J *et al*: A MicroRNA feedback circuit in midbrain dopamine neurons. *Science* 2007; **317**: 1220-1224.
24 Perkins DO, Jeffries CD, Jarskog LF *et al*: microRNA expression in the prefrontal cortex of individuals with schizophrenia and schizoaffective disorder. *Genome Biol* 2007; **8**: R27.
25 Hansen T, Olsen L, Lindow M *et al*: Brain expressed microRNAs implicated in schizophrenia etiology. *PLoS ONE* 2007; **2**: e873.
26 Jensen KP, Covault J, Conner TS, Tennen H, Kranzler HR, Furneaux HM: A common polymorphism in serotonin receptor 1B mRNA moderates regulation by miR-96 and associates with aggressive human behaviors. *Mol Psychiatry* 2008.
27 Abelson JF, Kwan KY, O'Roak BJ *et al*: Sequence variants in SLITRK1 are associated with Tourette's syndrome. *Science* 2005; **310**: 317-320.
28 Muinos-Gimeno M, Montfort M, Bayes M, Estivill X, Espinosa-Parrilla Y: Design and evaluation of a panel of single-nucleotide polymorphisms in microRNA genomic regions for association studies in human disease. *Eur J Hum Genet* 2009.
29 Bandelow B: Assessing the efficacy of treatments for panic disorder and agoraphobia. II. The Panic and Agoraphobia Scale. *Int Clin Psychopharmacol* 1995; **10**: 73-81.
30 Arrindell WA, Eisemann M, Richter J *et al*: Phobic anxiety in 11 nations. Part I: Dimensional constancy of the five-factor model. *Behav Res Ther* 2003; **41**: 461-479.
31 Hamilton M: A rating scale for Depression. *Journal of Neurology, Neurosurgery and Psychiatry* 1960; **23**: 56-61.
32 Gratacos M, Soria V, Urretavizcaya M *et al*: A brain-derived neurotrophic factor (BDNF) haplotype is associated with antidepressant treatment outcome in mood disorders. *Pharmacogenomics J* 2008; **8**: 101-112.
33 Donner J, Pirkola S, Silander K *et al*: An association analysis of murine anxiety genes in humans implicates novel candidate genes for anxiety disorders. *Biol Psychiatry* 2008; **64**: 672-680.
34 Nelis M, Esko T, Magi R *et al*: Genetic structure of Europeans: a view from the North-East. *PLoS One* 2009; **4**: e5472.
35 Gonzalez JR, Armengol L, Sole X *et al*: SNPassoc: an R package to perform whole genome association studies. *Bioinformatics* 2007; **23**: 644-645.
36 Sansone RA, Sansone LA, Righter EL: Panic disorder: the ultimate anxiety. *J Womens Health* 1998; **7**: 983-989.
37 Knight JC: Regulatory polymorphisms underlying complex disease traits. *J Mol Med* 2005; **83**: 97-109.
38 Rockman MV, Wray GA: Abundant raw material for cis-regulatory evolution in humans. *Mol Biol Evol* 2002; **19**: 1991-2004.
39 Smoller JW, Tsuang MT: Panic and phobic anxiety: defining phenotypes for genetic studies. *Am J Psychiatry* 1998; **155**: 1152-1162.
40 Sklar P, Smoller JW, Fan J *et al*: Whole-genome association study of bipolar disorder. *Mol Psychiatry* 2008; **13**: 558-569.
41 Zeng Y, Cullen BR: Efficient processing of primary microRNA hairpins by Drosha requires flanking nonstructured RNA sequences. *J Biol Chem* 2005; **280**: 27595-27603.
42 Luciano DJ, Mirsky H, Vendetti NJ, Maas S: RNA editing of a miRNA precursor. *Rna* 2004; **10**: 1174-1177.
43 Obernosterer G, Leuschner PJ, Alenius M, Martinez J: Post-transcriptional regulation of microRNA expression. *Rna* 2006; **12**: 1161-1167.
44 Siegel G, Obernosterer G, Fiore R *et al*: A functional screen implicates microRNA-138-dependent regulation of the depalmitoylation enzyme APT1 in dendritic spine morphogenesis. *Nat Cell Biol* 2009; **11**: 705-716.
45 Bell CJ, Nutt DJ: Serotonin and panic. *Br J Psychiatry* 1998; **172**: 465-471.
46 Maron E, Shlik J: Serotonin function in panic disorder: important, but why? *Neuropsychopharmacology* 2006; **31**: 1-11.

47 Spencer RL, Hutchison KE: Alcohol, aging, and the stress response. *Alcohol Res Health* 1999; **23**: 272-283.
48 Yalcin B, Willis-Owen SA, Fullerton J *et al*: Genetic dissection of a behavioral quantitative trait locus shows that Rgs2 modulates anxiety in mice. *Nat Genet* 2004; **36**: 1197-1202.
49 Koenen KC, Amstadter AB, Ruggiero KJ *et al*: RGS2 and generalized anxiety disorder in an epidemiologic sample of hurricane-exposed adults. *Depress Anxiety* 2009; **26**: 309-315.
50 Oliveira-Dos-Santos AJ, Matsumoto G, Snow BE *et al*: Regulation of T cell activation, anxiety, and male aggression by RGS2. *Proc Natl Acad Sci U S A* 2000; **97**: 12272-12277.
51 Pattyn A, Simplicio N, van Doorninck JH, Goridis C, Guillemot F, Brunet JF: Ascl1/Mash1 is required for the development of central serotonergic neurons. *Nat Neurosci* 2004; **7**: 589-595.
52 Taupenot L, Harper KL, O'Connor DT: The chromogranin-secretogranin family. *N Engl J Med* 2003; **348**: 1134-1149.
53 Tsui CC, Copeland NG, Gilbert DJ, Jenkins NA, Barnes C, Worley PF: Narp, a novel member of the pentraxin family, promotes neurite outgrowth and is dynamically regulated by neuronal activity. *J Neurosci* 1996; **16**: 2463-2478.

Table 1. Adjusted association analysis of SNPs in miRNA regions with PD under four genetic models

MiRNA region	SNP	Codominant [OR (CI) AIC]	Dominant [OR (CI) AIC]	Recessive [OR (CI) AIC]	Log-additive [OR (CI) AIC]
hsa-miR-488	rs2076072	0.16701	0.08235	0.23079	0.05982
hsa-miR-488	rs4652220	**0.01686** [0.57 (0.37-0.87) 623.4]	**0.04236** [0.67 (0.45-0.99) 625.4]	0.26241	0.43176
hsa-miR-488	rs12041859	0.47995	0.46588	0.25489	0.30093
***ASTN1* -miR-488**	rs12021769	0.69211	0.41681	0.94417	0.48054
***ASTN1* -miR-488**	rs6665725	0.52478	-	-	-
miR-339	rs11763020	**0.00041** [9.44 (2.80-31.80) 617.7]	0.24095	**0.000078** [9.45 (2.82-31.68) 615.7]	**0.02025** [1.54 (1.07-2.21) 625.9]
miR-339	rs1057558	0.57276	0.32092	0.93184	0.40477
***C7orf50* -miR-339**	rs10252922	0.10444	**0.03951** [1.57 (1.02-2.43) 625.2]	0.24026	**0.04596** [1.32 (1-1.73) 625.4]
***C7orf50* -miR-339**	rs12531999	0.38366	0.3416	0.22394	0.22689
hsa-miR-148a	rs4719839	0.74194	0.47194	0.96451	0.58149
hsa-miR-148a	rs735316	0.77093	0.6228	0.51229	0.50077
hsa-miR-148a	rs6977848	0.75172	0.50556	0.9037	0.68714
hsa-miR-148a	rs4722551	**0.03557** [3.91 (0.88-17.46) 627.5]	0.24986	**0.04801** [4.26 (0.96-18.86) 628.3]	0.62154
hsa-miR-491	rs4977831	**0.02725** [2.13 (1.23-3.69) 627.0]	**0.00821** [2.08 (1.21-3.58) 625.2]	0.98791	**0.0117** [1.96 (1.16-3.31) 625.8]
hsa-miR-491	rs2039391	**0.01964** [1.75 (1.11-2.76) 624.9]	**0.00769** [1.83 (1.17-2.85) 623.7]	0.14341	**0.00511** [1.78 (1.19-2.66) 623.0]
hsa-miR-491	rs12684948	0.56583	0.51412	0.58804	0.93917
hsa-miR-491	rs7022666	**0.01629** [0.40 (0.21-0.77) 625.5]	**0.04953** [0.68 (0.46-1) 627.9]	**0.0094** [0.46 (0.25-0.84) 625.0]	**0.00688** [0.68 (0.51-1.90) 624.4]
***KIAA1797* -miR-491**	rs10811383	0.71342	0.42445	0.93427	0.45216
***KIAA1797* -miR-491**	rs13290184	0.30773	-	-	-
hsa-miR-138-2	rs1436424	0.19146	0.36311	0.07488	0.10284
hsa-miR-138-2	rs12920659	0.37298	0.88849	0.16175	0.65319
hsa-miR-138-2	rs12921781	0.05806	**0.02817** [1.53 (1.05-2.25) 627.3]	0.84198	0.10408
hsa-miR-22	rs8076112*	**0.02059** [10.84 (1.28-91.74) 626.4]	0.33086	**0.00581** [10.59 (1.26-89.34) 624.6]	0.10975
hsa-miR-22	rs6502892	**0.00016** [0.45 (0.21-0.94) 616.7]	0.11698	**0.00115** [0.34 (0.17-0.68) 621.6]	0.74163
hsa-miR-22	rs4790814	0.10965	0.55064	0.08875	0.53209

ASTN1 , astrotactin 1; C7orf50, chromosome 7 open reading frame 50; *KIAA1797* , KIAA1797 mRNA

CI, 95% confidence interval; OR, odds ratio; SNP, single nucleotide polymorphism

AIC means Akaike information criterion, which attempts to find the minimal model that correctly explains the data

OR, CI and AIC for signficant P values, in bold, are shown

* HWE:0.0166

Table 2. Association analysis of SNPs in miRNA regions with PD with and without agorafobia

Agoraphobic phenotype

MiRNA region	SNP	Codominant [OR]	Dominant [OR]	Recessive [OR]	Log-additive [OR]
hsa-miR-488	rs2076072	**0.04014** [1.88]	**0.01182** [1.91]	0.3576	**0.01261** [1.76]
hsa-miR-488	rs4652220	**0.00129** [0.41]	**0.00253** [0.51]	0.44714	0.09101
hsa-miR-488	rs12041859	0.39669	0.34339	0.22672	0.21337
miR-339	rs11763020	**0.00399** [9.37]	0.30439	**0.0009** [9.30]	0.05093
miR-339	rs1057558	0.77091	0.51826	0.88624	0.60508
hsa-miR-148a	rs4719839	0.49066	0.24054	0.87176	0.31793
hsa-miR-148a	rs735316	0.51694	0.68018	0.2515	0.40511
hsa-miR-148a	rs6977848	0.54676	0.29607	0.98163	0.4545
hsa-miR-148a	rs4722551	0.2093	0.81239	0.09709	0.82024
hsa-miR-491	rs4977831	0.21352	0.19559	0.30478	0.26313
hsa-miR-491	rs2039391	0.25635	0.12978	0.323	0.10235
hsa-miR-491	rs12684948	0.93032	0.70394	0.8884	0.75376
hsa-miR-491	rs7022666	0.09593	0.29546	**0.03364** [0.489]	0.07143
hsa-miR-138-2	rs1436424	0.32102	0.21596	0.22315	0.1318
hsa-miR-138-2	rs12920659	0.34467	0.37196	0.33667	0.55669
hsa-miR-138-2	rs12921781	0.60136	0.31702	0.63806	0.33514
hsa-miR-22	rs8076112*	**0.04374** [9.31]	0.47492	**0.01255** [9.20]	0.18552
hsa-miR-22	rs6502892	**0.00116** [1.76]	0.14862	**0.00536** [0.35]	0.80995
hsa-miR-22	rs4790814	0.06875	0.42227	0.07734	0.59583

Non-agoraphobic phenotype

MiRNA region	SNP	Codominant [OR]	Dominant [OR]	Recessive [OR]	Log-additive [OR]
hsa-miR-488	rs2076072	0.51359	0.76987	0.32994	0.95461
hsa-miR-488	rs4652220	0.39492	0.709	0.17351	0.31381
hsa-miR-488	rs12041859	0.92331	0.93633	0.7223	0.95303
miR-339	rs11763020	**0.01207** [9.32]	0.41604	**0.00297** [9.39]	0.08631
miR-339	rs1057558	0.65309	0.3855	0.94377	0.46442
hsa-miR-148a	rs4719839	0.95988	0.88928	0.84448	0.97777
hsa-miR-148a	rs735316	0.80283	0.74887	0.661	0.98025
hsa-miR-148a	rs6977848	0.93151	0.75612	0.76491	0.707
hsa-miR-148a	rs4722551	**0.00811** [0.38]	**0.03817** [0.48]	0.07826	0.13865
hsa-miR-491	rs4977831	**0.0087** [3.12]	**0.00208** [3.14]	0.44241	**0.00254** [2.84]
hsa-miR-491	rs2039391	**0.01767** [2.41]	**0.00476** [2.47]	0.32282	**0.00521** [2.22]
hsa-miR-491	rs12684948	0.1806	0.37341	0.23047	0.88271
hsa-miR-491	rs7022666	**0.04564** [0.34]	**0.02579** [0.53]	0.07751	**0.01297** [0.59]
hsa-miR-138-2	rs1436424	0.18144	0.94601	0.08113	0.32864
hsa-miR-138-2	rs12920659	0.144	0.1224	0.14107	0.08443
hsa-miR-138-2	rs12921781	**0.00075** [2.78]	**0.00375** [2.31]	0.11492	0.09861
hsa-miR-22	rs8076112*	0.15725	0.45993	0.06034	0.26151
hsa-miR-22	rs6502892	**0.01135** [1.66]	0.38049	**0.01384** [0.27]	0.63616
hsa-miR-22	rs4790814	0.60228	0.85046	0.39027	0.70062

The odds ratio (OR) for signficant P values (in bold) are shown

* HWE:0.0166

Table 3. Association analysis for the miR-148a region with AAO in PD

SNP	p-value [Difference (CI) AIC]
	Codominant
rs4719839	0.27622
rs735316	**0.0033** [-7.7430 (-12.540--29462) 1324]*
rs6977848	0.35714
rs4722551	0.25346
	Dominant
rs4719839	0.11069
rs735316	0.447
rs6977848	0.23438
rs4722551	0.55716
	Recessive
rs4719839	0.47212
rs735316	**0.00073** [-7.9257 (-12.523--3.3284) 1322]
rs6977848	0.246
rs4722551	0.09676
	Log-additive
rs4719839	0.12268
rs735316	**0.03841** [-2.2118 (-4.306--0.1180) 1329]
rs6977848	0.15165
rs4722551	0.28669

*Difference on the AAO and CI for the Homozygous state

Table 4. Sequence variants identified at the five miRNA regions sequenced

miRNA	Change	Distance	MAF in PD	Alleles
miR-488 (chr1)	rs12041859*⌘	192bp	G=0.3118	A/G
	rs10753142*⌘	131bp	T=0.3824	A/T
	chr1:175,265,269-71	65bp	del= 0.0056	-/GAA
	rs56297461*⌘	349bp	T=0.118	C/T
	chr1:175,265,792	588bp	A=0.006	A/G
	chr1:175,266,087	883bp	G=0.006	A/G
	chr1:175,266,110	906bp	C=0.006	C/T
	chr1:175,266,127-28	923bp	del=0.0119	-/T
miR-339 (chr7)	rs2960827	96bp	C=0.0115	A/G
	rs3735686*⌘	42bp	G=0.4091	A/G
	rs1057561*⌘	41bp	T=0.0843	C/T
	chr7:1,029,086	9bp	G=0.0057	A/G
	rs59633269	122bp	T=0.0057	C/T
	chr7:1,029,322	134bp	C=0.0057	C/T
	chr7:1,029,456	268bp	T=0.0057	C/T
miR-491 (chr9)	rs2383148*⌘	293bp	T=0.178	C/T
	chr9:20,706,069	35bp	C=0.0057	C/T
	chr9:20,706,250⌘	63bp	G=0.0625	A/G
	chr9:20,706,311⌘	124bp	T=0.0568	C/T
miR-138-2 (chr16)	rs35474169#⌘	216bp	del=0.5125	-/T
miR-22 (chr17)	chr16:55,450,094	80bp	A=0.0122	A/G
	chr17:1,564,113	82bp	T=0.0059	C/T
	chr17:1,564,125	94bp	G=0.0059	C/G

*, d'=1 with the associated SNP
#, SNPs that are not in HWE in the estonian population
⌘, common allele variatns (MAF>0.05)
The distance between the variant and the pre-miRNA 5' or 3' ends is indicated

Table 5. Target site prediction and luciferase assay for five associated miRNAs

	miR-22	miR-138-2	miR-148a	miR-339	miR-488
ADORA 2A	-	T	-	T	T
BDNF	53.7%^{M}	T	-	T	-
CCKAR	M	-	-	-	M
CCKBR	-	-	42.5%$^{M, T}$	-	-
CRHR2	-	-	-	T	-
GABRA6	-	69.9%^{T}	-	-	-
HTR2C	58.5%^{T}	-	M, T	-	M
MAOA	55.1%$^{M, T}$	-	M, T	-	-
NTRK3	T	M	-	-	M
POMC	-	-	-	-	52.2%^{T}
RGS2	33.7%$^{M, T}$	-	-	-	-
SLC6A2	-	-	-	M	M

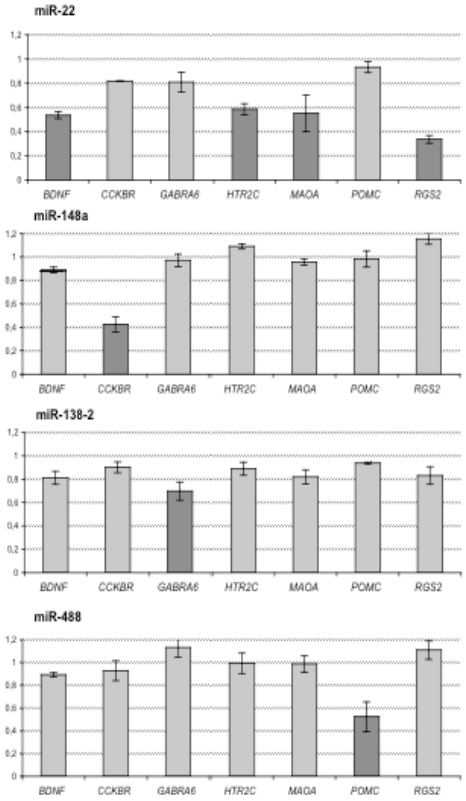
TargetScan (T) and MiRanda (M)
Statistically significant reduction of the luciferase activity (%) is shown in grey

Table 6. Top 10 upregulated and downregulated genes upon overexpression of either miR-22 or miR-488 in SH-SY5Y cells and overlapping TargetScan predictions

miR-22		miR-488	
Upregulated	**Fold Change**	**Upregulated**	**Fold Change**
MOCOS	3.815	*IFRD1*	1.714
VIP5 [#]	3.575	*C6ORF48*	1.601
NQO1	3.223	*TMEM79*	1.515
VGF	3.100	***PLD5****	1.486
ASNS	2.781	***ISLR2****	1.486
THBS2	2.715	*PPFIBP2*	1.461
TRIB3	2.591	*DNALI1*	1.448
FAM129A	2.533	*GAL*	1.426
KLF2	2.516	*DDIT3*	1.413
WARS	2.510	*IRX3*	1.376
Downreregulated	**Fold Change**	**Downreregulated**	**Fold Change**
ASCL1*	-3.404	***SLC35D3****	-2.165
CHGA*	-3.149	***PXDN****	-2.020
DOK4	-2.549	***CTSC****	-1.996
HMGB2	-2.324	***POLR1D****	-1.989
NPTX2*	-2.284	***SDHD****	-1.934
SLC35D3	-2.190	***CANT1****	-1.918
HNRNPA3*	-2.175	***IMPAD1****	-1.870
RFXANK*	-2.136	***FAM57A****	-1.824
NHP2	-2.090	***RGS4****	-1.823
RGS2 *	-2.086	*RN7SK* [#]	-1.696

Deregulated genes predicted as targets by TargetScan (*) are displayed inside grey boxes
[#], Genes that are not recognized by TargetScan

Figure 1


miR-22
1,2
1
0,8
0,6
0,4
0,2
0
BDNF
CCKBR
GABRA6
HTR2C
MAOA
POMC
RGS2
miR-148a
1,2
1
0,8
0,6
0,4
0,2
0
BDNF
CCKBR
GABRA6
HTR2C
MAOA
POMC
RGS2
miR-138-2
1,2
1
0,8
0,6
0,4
0,2
0
BDNF
CCKBR
GABRA6
HTR2C
MAOA
POMC
RGS2
miR-488
1,2
1
0,8
0,6
0,4
0,2
0
BDNF
CCKBR
GABRA6
HTR2C
MAOA
POMC
RGS2

Supplementary Table1. SNPs studied in the miRNA SNP panel

ID	Genomic Position	Genomic Region	Alleles	MA	MAF controls	AF panic
rs4970362	chr1:1084601	cluster1.1	A/G	A	0.333	0.319
rs9660710	chr1:1089205	cluster1.1	A/C	A	0.079	0.056
rs4970420	chr1:1096336	cluster1.1	G/A	A	0.205	0.198
rs11260542	chr1:1099882	cluster1.1	A/G	G	0.068	0.061
rs4281343	chr1:9132545	hsa_mir_34a	G/A	A	0.271	0.287
rs12128240	chr1:9133890	hsa_mir_34a	C/T	T	0.245	0.233
rs2666433	chr1:9135764	hsa_mir_34a	A/G	A	0.125	0.098
rs7520583	chr1:9136959	hsa_mir_34a	A/C	A	0.236	0.250
rs2780950	chr1:40925196	NFYC_cl1.2	C/T	C	0.198	0.229
rs965949	chr1:40926994	NFYC_cl1.2	G/A	G	0.379	0.418
rs3818758	chr1:40987662	cluster1.2	T/A	A	0.425	0.423
rs12729696	chr1:40993358	cluster1.2	C/T	T	0.199	0.225
rs4660453	chr1:40994486	cluster1.2	T/C	T	0.102	0.093
rs17357062	chr1:40994802	cluster1.2	C/T	T	0.040	0.037
rs928508	chr1:40996001	cluster1.2	G/A	G	0.382	0.415
rs7536540	chr1:65297170	miR_101_1	C/G	C	0.224	0.230
rs1011210	chr1:65298764	miR_101_1	T/C	T	0.358	0.376
rs555146	chr1:65299995	miR_101_1	A/C	A	0.090	0.074
rs547812	chr1:65300765	miR_101_1	A/G	A	0.056	0.032
rs17090778	chr1:71304184	miR_186	A/G	G	0.140	0.128
rs11583800	chr1:71309162	miR_186	G/C	C	0.102	0.143
rs9970600	chr1:71310046	miR_186	T/G	G	0.127	0.135
rs11209745	chr1:71310531	miR_186	A/G	G	0.193	0.181
rs12035210	chr1:71318558	ZNF265_miR_186	G/A	A	0.192	0.180
rs7549846	chr1:71320710	ZNF265_miR_186	C/G	G	0.058	0.061
rs17090808	chr1:71323733	ZNF265_miR_186	G/C	C	0.080	0.098
rs2802535	chr1:98280846	miR_137	C/T	C	0.180	0.209
rs2660303	chr1:98285732	miR_137	T/A	A	0.121	0.128
rs3525	chr1:109938369	miR_197	T/A	A	0.153	0.169
rs559479	chr1:109940059	miR_197	T/A	T	0.274	0.283
rs505303	chr1:109942500	miR_197	C/T	T	0.344	0.368
rs10127988	chr1:109946148	miR_197	C/T	T	0.445	0.481
rs16837375	chr1:154653791	miR_9_1	C/G	G	0.270	0.273
rs16837376	chr1:154659079	miR_9_1	A/T	T	0.059	0.074
rs7548323	chr1:154660624	miR_9_1	G/A	A	0.162	0.156
rs10494305	chr1:154666684	C1orf61_miR_9_1	C/T	T	0.087	0.114
rs12138127	chr1:154670305	C1orf61_miR_9_1	C/G	C	0.298	0.325
rs17277008	chr1:170371785	cluster1.3	T/C	C	0.317	0.330
rs2761184	chr1:170373429	cluster1.3	A/C	C	0.062	0.040
rs2819530	chr1:170379577	cluster1.3	T/C	T	0.407	0.437
rs10911111	chr1:170384303	cluster1.3	A/G	G	0.202	0.189
rs2076072	chr1:175260861	miR_488	C/T	T	0.111	0.143
rs4652220	chr1:175263576	miR_488	T/C	C	0.405	0.380
rs12041859	chr1:175264930	miR_488	T/C	C	0.226	0.255
rs12021769	chr1:175398690	ASTN1_miR_488	C/T	T	0.141	0.126
rs6665725	chr1:175402740	ASTN1_miR_488	A/G	G	0.055	0.064
rs12131140	chr1:197092252	cluster1.4	G/A	A	0.102	0.108
rs10753837	chr1:197092369	cluster1.4	C/T	C	0.058	0.066
rs12406238	chr1:197092583	cluster1.4	G/A	A	0.099	0.111
rs10919595	chr1:197093847	cluster1.4	A/G	A	0.378	0.378

rs7548935	chr1:203679584	miR_135b	C/A	C	0.180	0.148
rs12121596	chr1:203681389	miR_135b	G/C	C	0.304	0.336
rs10793730	chr1:203681971	miR_135b	T/C	T	0.081	0.079
rs1999131	chr1:203684757	miR_135b	A/G	G	0.402	0.423
rs1474742	chr1:206043289	cluster1.5	C/A	A	0.330	0.304
rs2745972	chr1:206047186	cluster1.5	T/C	C	0.426	0.405
rs12035599	chr1:207669875	miR_205	T/G	G	0.032	0.045
rs10779506	chr1:207669983	miR_205	G/C	G	0.312	0.415
rs7516946	chr1:207673545	miR_205	G/A	A	0.099	0.130
rs2356929	chr1:207674344	miR_205	A/G	A	0.201	0.257
rs1039206	chr1:218355055	cluster1.6	A/G	G	0.049	0.072
rs17007135	chr1:218355581	cluster1.6	A/C	C	0.139	0.173
rs6756818	chr2:56063383	cluster2.1	T/A	A	0.292	0.285
rs2009613	chr2:56072032	cluster2.1	G/A	A	0.118	0.114
rs3806502	chr2:136004743	R3HDM_miR_128-1	C/T	T	0.207	0.241
rs17331388	chr2:136006035	R3HDM_miR_128-1	C/T	T	0.078	0.093
rs10194951	chr2:136007133	R3HDM_miR_128-1	T/A	A	0.124	0.161
rs3769018	chr2:136135211	miR_128-1	T/C	C	0.127	0.136
rs1446584	chr2:136136431	miR_128-1	A/G	G	0.417	0.455
rs1374330	chr2:136137378	miR_128-1	A/C	C	0.230	0.229
rs2289959	chr2:136140374	miR_128-1	C/T	T	0.264	0.261
rs6736786	chr2:176717160	mir_10b	A/G	G	0.453	0.458
rs1867863	chr2:176723216	mir_10b	G/T	T	0.285	0.312
rs4972505	chr2:176729647	mir_10b	G/A	A	0.343	0.356
rs2279014	chr2:218969420	CTDSP1_miR_26b	C/T	T	0.363	0.375
rs2227251	chr2:218974667	miR_26b	C/A	A	0.313	0.267
rs2227255	chr2:218976025	miR_26b	C/T	C	0.429	0.389
rs921968	chr2:218980538	miR_26b	G/T	G	0.475	0.434
rs371825	chr2:219571900	miR_375	A/G	G	0.190	0.196
rs422433	chr2:219572646	miR_375	A/C	C	0.189	0.193
rs1983298	chr2:219864078	miR_153_1	T/C	C	0.148	0.124
rs4674373	chr2:219864718	miR_153_1	C/G	G	0.195	0.190
rs4485556	chr2:219865662	miR_153_1	A/C	C	0.206	0.229
rs11889536	chr2:219871787	miR_153_1	A/G	G	0.201	0.190
rs7591986	chr2:219880820	PTPRN_miR_153_1	A/C	C	0.158	0.140
rs2090163	chr2:219886679	PTPRN_miR_153_1	T/C	C	0.382	0.365
rs6716265	chr2:241022204	GPC1_miR_149	T/C	C	0.400	0.344
rs3828336	chr2:241039189	miR_149	C/T	T	0.280	0.320
rs4678788	chr3:35658440	ARPP_21_miR_128-2	G/A	G	0.302	0.302
rs7621392	chr3:35756391	miR_128-2	G/A	A	0.072	0.087
rs13094224	chr3:35759585	miR_128-2	A/G	G	0.252	0.272
rs6550367	chr3:35760300	miR_128-2	T/C	C	0.450	0.439
rs6801590	chr3:35762465	miR_128-2	A/C	C	0.282	0.262
rs608516	chr3:37878977	CTDSPL_miR_26a_1	G/A	A	0.000	0.000
rs11719907	chr3:37984704	miR_26a_1	C/T	T	0.139	0.172
rs7372209	chr3:37985712	miR_26a_1	T/C	T	0.304	0.274
rs6789468	chr3:37988497	miR_26a_1	A/G	G	0.186	0.205
rs6441821	chr3:44127825	miR_138_1	A/G	G	0.121	0.122
rs11715269	chr3:44132941	miR_138_1	T/C	C	0.102	0.098
rs3087866	chr3:49029696	cluster3.1	T/C	T	0.217	0.222
rs11706052	chr3:49039114	cluster3.1	A/G	G	0.103	0.111
rs11717018	chr3:52268899	let_7g	A/G	G	0.114	0.117
rs6793317	chr3:52274026	let_7g	T/C	C	0.115	0.120
rs3733063	chr3:52303573	miR_135a_1	G/A	A	0.117	0.120

rs2488	chr3:121595976	miR_198	C/T	T	0.478	0.492
rs1700	chr3:121597327	miR_198	C/T	T	0.143	0.148
rs1259327	chr3:121599612	miR_198	C/T	C	0.332	0.312
rs1147695	chr3:121602136	miR_198	C/T	C	0.115	0.132
rs9631455	chr3:121657140	FSTL1_miR_198	T/C	C	0.350	0.360
rs4680580	chr3:161597698	SMC4L1_cl3.2	C/T	T	0.442	0.430
rs10936201	chr3:161604533	cluster3.2	C/A	A	0.202	0.215
rs7634108	chr3:161607389	cluster3.2	C/A	A	0.441	0.428
rs4679883	chr3:161608297	cluster3.2	G/A	A	0.153	0.116
rs1842306	chr3:189413160	LPP_miR_28	T/C	C	0.261	0.267
rs7623253	chr3:189888174	miR_28	T/C	T	0.476	0.452
rs6791355	chr3:189888490	miR_28	A/G	G	0.461	0.473
rs17670540	chr3:189888911	miR_28	G/A	A	0.289	0.325
rs6770503	chr3:189889673	miR_28	C/G	G	0.364	0.415
rs11719436	chr3:189890561	miR_28	G/A	A	0.077	0.090
rs11729198	chr4:8054827	miR_95	G/A	A	0.339	0.291
rs4696664	chr4:8059164	miR_95	A/G	A	0.429	0.495
rs17180256	chr4:8060659	miR_95	G/A	G	0.327	0.312
rs13125250	chr4:8061012	miR_95	C/T	C	0.223	0.193
rs13132877	chr4:8214460	ABLIM2_miR_95	T/G	G	0.184	0.186
rs564041	chr4:19861365	SLIT2_miR_218_1	G/A	G	0.295	0.295
rs7655084	chr4:19864404	SLIT2_miR_218_1	G/T	G	0.285	0.288
rs7690660	chr4:20136962	miR_218_1	T/G	G	0.175	0.149
rs3775815	chr4:20138768	miR_218_1	G/T	T	0.146	0.119
rs9992559	chr4:20139472	miR_218_1	T/C	C	0.274	0.272
rs2168801	chr4:20140079	miR_218_1	G/A	A	0.409	0.402
rs13136737	chr4:113788584	cluster4.1	G/T	G	0.487	0.497
rs4383675	chr4:113792068	cluster4.1	A/G	A	0.323	0.381
rs6533631	chr4:113797715	cluster4.1	G/A	A	0.136	0.093
rs1129065	chr4:113798058	cluster4.1	C/T	T	0.158	0.114
rs3747725	chr5:54497747	miR_449	A/G	G	0.071	0.096
rs336115	chr5:54504678	miR_449	C/T	C	0.243	0.299
rs2304608	chr5:87998054	miR_9_2	C/A	A	0.175	0.188
rs1501672	chr5:87999517	miR_9_2	A/G	G	0.151	0.151
rs17796757	chr5:148786894	cluster5.1	A/T	T	0.133	0.128
rs353293	chr5:148787419	cluster5.1	C/T	T	0.398	0.397
rs353291	chr5:148790939	cluster5.1	T/C	C	0.403	0.388
rs476741	chr5:148795397	cluster5.1	T/G	G	0.432	0.444
rs17797652	chr5:149085286	PPARGC1B_miR_378	T/G	G	0.107	0.117
rs12516673	chr5:149085798	PPARGC1B_miR_378	A/T	T	0.167	0.205
rs17711165	chr5:149087092	PPARGC1B_miR_378	A/C	C	0.145	0.146
rs1076063	chr5:149092388	miR_378	T/A	T	0.087	0.098
rs9986116	chr5:149093289	miR_378	G/T	T	0.276	0.278
rs10060424	chr5:149094417	miR_378	T/C	C	0.426	0.426
rs6878034	chr5:159847699	miR_146a	T/C	C	0.398	0.394
rs158896	chr5:167915925	miR_103_1	T/C	T	0.304	0.310
rs158895	chr5:167918191	miR_103_1	C/T	C	0.326	0.322
rs11745534	chr5:167919878	miR_103_1	T/G	G	0.150	0.191
rs158900	chr5:167924532	miR_103_1	A/G	G	0.379	0.333
rs149476	chr5:167939278	PANK3_miR_103_1	T/G	T	0.338	0.331
rs11745421	chr5:168123644	miR_218_2	T/C	C	0.108	0.085
rs6860854	chr5:168126754	miR_218_2	C/T	C	0.465	0.396
rs1368355	chr5:168129277	miR_218_2	C/G	C	0.232	0.259
rs6877545	chr5:168129537	miR_218_2	A/G	G	0.093	0.101

rs4242184	chr5:168658664	SLIT3_miR_218_2	G/A	G	0.216	0.266
rs1422028	chr5:168663218	SLIT3_miR_218_2	C/G	G	0.329	0.290
rs10068039	chr5:179370406	miR_340	T/C	T	0.339	0.391
rs3797763	chr5:179372122	miR_340	C/T	C	0.159	0.193
rs11739706	chr5:179374254	miR_340	C/G	G	0.178	0.196
rs11249664	chr5:179376873	miR_340	C/T	C	0.181	0.198
rs4700862	chr5:179433328	RNF130_miR_340	G/A	G	0.369	0.413
rs365339	chr6:33280883	miR_219_1	T/C	T	0.227	0.267
rs421446	chr6:33282761	miR_219_1	A/G	G	0.305	0.336
rs1892275	chr6:52114216	miR-206	T/C	C	0.096	0.066
rs6920648	chr6:52116575	miR-206	A/G	G	0.422	0.391
rs1537670	chr6:52117559	miR-206	A/G	G	0.061	0.042
rs17578851	chr6:52117669	miR-206	C/T	T	0.214	0.204
rs10484882	chr6:52119428	miR-133b	G/A	A	0.093	0.087
rs10948678	chr6:52120245	miR-133b	G/A	A	0.229	0.233
rs10948679	chr6:52120832	miR-133b	C/T	T	0.490	0.449
rs1192037	chr6:72143796	miR_30c_2	A/T	A	0.215	0.222
rs1192038	chr6:72144062	miR_30c_2	C/T	T	0.269	0.266
rs1192039	chr6:72144678	miR_30c_2	A/G	A	0.090	0.098
rs1192040	chr6:72146300	miR_30c_2	T/C	T	0.146	0.164
rs1405848	chr6:72166043	miR_30a	C/T	T	0.135	0.127
rs6921100	chr6:72167425	miR_30a	C/G	C	0.277	0.269
rs2222722	chr6:72170433	miR_30a	G/A	A	0.205	0.204
rs763354	chr6:72172342	miR_30a	C/T	T	0.081	0.066
rs11763020	chr7:1026814	miR_339	C/T	T	0.128	0.172
rs1057558	chr7:1028892	miR_339	T/C	C	0.183	0.162
rs13232101	chr7:1029100	miR_339	G/T	T	0.000	0.000
rs10252922	chr7:1143977	C7orf50_miR_339	G/A	A	0.450	0.5102
rs12531999	chr7:1148127	C7orf50_miR_339	T/A	A	0.146	0.175
rs4719839	chr7:25953098	miR_148a	G/A	G	0.273	0.257
rs735316	chr7:25954710	miR_148a	T/C	C	0.330	0.313
rs6977848	chr7:25956045	miR_148a	T/G	G	0.394	0.386
rs4722551	chr7:25958351	miR_148a	T/C	C	0.133	0.135
rs3801776	chr7:27171807	miR_196b	A/G	A	0.252	0.235
rs17471555	chr7:27173820	miR_196b	A/T	T	0.000	0.000
rs886340	chr7:27174026	miR_196b	G/A	G	0.131	0.096
rs2237336	chr7:27174543	miR_196b	C/T	T	0.439	0.441
rs2074121	chr7:92947019	miR_489	T/A	A	0.494	0.481
rs10488542	chr7:92952300	miR_489	T/G	T	0.472	0.497
rs6965274	chr7:93040548	CALCR_miR_489	T/C	T	0.105	0.120
rs2251307	chr7:93044972	CALCR_miR_489	A/G	G	0.460	0.447
rs13236112	chr7:93045809	CALCR_miR_489	G/A	A	0.091	0.085
rs4928	chr7:99528626	cluster7.1	G/C	G	0.071	0.053
rs13242458	chr7:99533067	cluster7.1	G/A	A	0.118	0.132
rs1534309	chr7:99534306	cluster7.1	C/G	G	0.161	0.146
rs2070215	chr7:99534733	MCM7_cl7.1	T/C	C	0.270	0.302
rs4319008	chr7:99535951	MCM7_cl7.1	G/A	A	0.071	0.053
rs999885	chr7:99539112	MCM7_cl7.1	G/A	G	0.447	0.431
rs4731416	chr7:127631331	miR_129_1	A/G	A	0.077	0.077
rs1451006	chr7:127637891	miR_129_1	G/A	A	0.140	0.161
rs4731419	chr7:127638031	miR_129_1	C/T	C	0.415	0.380
rs6467264	chr7:129194497	cluster7.2	C/T	T	0.378	0.434
rs6965643	chr7:129194800	cluster7.2	A/G	G	0.219	0.220
rs4335057	chr7:129203271	cluster7.2	A/G	A	0.444	0.476

rs2693727	chr7:129204966	cluster7.2	C/T	C	0.081	0.108
rs17391475	chr7:129908597	MEST_miR_335	T/A	T	0.360	0.367
rs12706940	chr7:129916029	MEST_miR_335	C/T	C	0.445	0.481
rs2301335	chr7:129920062	miR_335	G/A	G	0.411	0.460
rs2072573	chr7:129924138	miR_335	G/T	G	0.412	0.458
rs7349991	chr7:130208719	cluster7.3	G/A	A	0.322	0.288
rs207217	chr7:130210863	cluster7.3	C/T	C	0.416	0.378
rs24168	chr7:130211825	cluster7.3	A/G	A	0.459	0.447
rs7781163	chr7:130213364	cluster7.3	T/C	C	0.157	0.184
rs157906	chr7:130213783	cluster7.3	A/G	A	0.316	0.275
rs157907	chr7:130213977	cluster7.3	G/A	G	0.313	0.275
rs17165207	chr7:130216237	cluster7.3	G/C	C	0.096	0.093
rs1424574	chr7:136199573	CHRM2_miR_490	A/G	G	0.158	0.148
rs6957496	chr7:136202377	CHRM2_miR_490	A/G	G	0.104	0.101
rs12535371	chr7:136234615	miR_490	T/C	C	0.363	0.327
rs1364402	chr7:136234903	miR_490	T/C	C	0.124	0.116
rs10488599	chr7:136239634	miR_490	C/T	T	0.143	0.116
rs2350780	chr7:136243509	miR_490	G/A	G	0.306	0.375
rs221295	chr7:157055372	miR_153_2	C/A	A	0.420	0.426
rs221296	chr7:157056251	miR_153_2	G/A	G	0.036	0.019
rs6459797	chr7:157057867	miR_153_2	G/A	A	0.476	0.481
rs7808805	chr7:157061230	miR_153_2	C/G	C	0.459	0.452
rs2335158	chr7:158077841	PTPRN2_miR_153_2	C/T	T	0.215	0.213
rs592420	chr8:9797830	miR_124a_1	A/G	G	0.230	0.206
rs531564	chr8:9798109	miR_124a_1	G/C	G	0.125	0.127
rs13252510	chr8:9803298	miR_124a_1	G/A	A	0.093	0.066
rs4345555	chr8:14746477	miR_383	T/C	C	0.425	0.447
rs4602887	chr8:14763659	miR_383	A/C	A	0.333	0.323
rs1126677	chr8:22161447	miR_320	T/C	T	0.340	0.370
rs2645	chr8:22164292	miR_320	C/T	T	0.419	0.434
rs4317621	chr8:41635738	miR_486	A/G	A	0.497	0.492
rs16890723	chr8:41636221	miR_486	T/C	C	0.024	0.040
rs565491	chr8:41637017	miR_486	G/A	A	0.279	0.229
rs298210	chr8:65452923	miR_124a_2	A/G	A	0.143	0.127
rs298213	chr8:65453820	miR_124a_2	A/C	A	0.150	0.135
rs2289418	chr8:65457385	miR_124a_2	C/T	T	0.063	0.079
rs298219	chr8:65457805	miR_124a_2	G/T	G	0.217	0.215
rs7011656	chr8:135879497	miR-30b	G/A	A	0.283	0.251
rs7460121	chr8:135881598	miR-30b	G/A	A	0.111	0.103
rs10095483	chr8:135881730	miR-30b	A/C	C	0.223	0.213
rs17709119	chr8:135882558	miR-30d	C/T	T	0.149	0.153
rs16905252	chr8:135882875	miR-30d	A/G	G	0.135	0.098
rs11997161	chr8:141807769	miR_151	T/C	T	0.404	0.418
rs4961288	chr8:141810650	miR_151	G/A	A	0.379	0.373
rs170068	chr8:142082654	PTK2_miR_151	G/A	A	0.366	0.362
rs4740801	chr9:4780165	RCL1_miR_101_2	A/G	A	0.416	0.468
rs1260011	chr9:4782044	RCL1_miR_101_2	G/T	T	0.474	0.497
rs423955	chr9:4782339	RCL1_miR_101_2	C/T	C	0.290	0.319
rs10815095	chr9:4838297	miR_101_2	A/G	G	0.176	0.168
rs295254	chr9:4838630	miR_101_2	A/G	G	0.463	0.497
rs159433	chr9:4838875	miR_101_2	G/T	G	0.121	0.120
rs12375841	chr9:4840141	miR_101_2	T/C	C	0.223	0.198
rs460182	chr9:4840613	miR_101_2	A/G	A	0.344	0.320
rs10811383	chr9:20642174	KIAA1797_miR_491	G/A	A	0.125	0.143

rs13290184	chr9:20645564	KIAA1797_miR_491	G/T	T	0.037	0.026
rs4977831	chr9:20703068	miR_491	A/G	A	0.053	0.095
rs2039391	chr9:20704014	miR_491	C/T	C	0.097	0.160
rs12684948	chr9:20706880	miR_491	C/T	T	0.499	0.484
rs7022666	chr9:20709867	miR_491	G/A	A	0.395	0.315
rs9942973	chr9:21497130	mir_31	G/C	C	0.055	0.053
rs1537275	chr9:21497589	mir_31	C/T	C	0.410	0.384
rs10491567	chr9:21500160	mir_31	C/A	A	0.061	0.034
rs2210437	chr9:21504640	mir_31	C/A	C	0.312	0.299
rs540651	chr9:72615220	miR_204	T/C	T	0.003	0.005
rs1415232	chr9:72615588	miR_204	C/G	C	0.201	0.199
rs718447	chr9:72616673	miR_204	G/A	A	0.245	0.243
rs718446	chr9:72616734	miR_204	C/T	T	0.044	0.045
rs3812532	chr9:72673414	TRPM3_miR_204	G/T	T	0.432	0.468
rs2993010	chr9:72673741	TRPM3_miR_204	A/G	A	0.371	0.378
rs7856482	chr9:72675266	TRPM3_miR_204	A/G	G	0.230	0.262
rs10868951	chr9:72925585	TRPM3_miR_204	A/C	C	0.459	0.457
rs4620343	chr9:72926463	TRPM3_miR_204	C/T	C	0.479	0.497
rs1329747	chr9:72927687	TRPM3_miR_204	G/A	A	0.465	0.442
rs296886	chr9:85781846	miR_7_1	A/G	G	0.211	0.210
rs17086750	chr9:85783635	HNRPK_miR_7_1	G/T	T	0.109	0.114
rs296892	chr9:85787626	HNRPK_miR_7_1	A/G	A	0.479	0.458
rs296893	chr9:85788010	HNRPK_miR_7_1	C/T	T	0.262	0.280
rs10739971	chr9:95977501	cluster9.1	G/A	A	0.294	0.331
rs10993081	chr9:95979539	cluster9.1	A/G	G	0.063	0.079
rs8115	chr9:95981398	cluster9.1	A/G	G	0.155	0.160
rs10512231	chr9:95985569	cluster9.1	C/T	T	0.328	0.278
rs396538	chr9:96527525	C9orf3_cl9.2	A/G	G	0.111	0.117
rs4743988	chr9:96886786	cluster9.2	A/G	A	0.236	0.243
rs1045276	chr9:96888505	cluster9.2	C/T	T	0.201	0.201
rs4744443	chr9:96891704	cluster9.2	T/C	C	0.355	0.328
rs513852	chr9:110847973	miR_32	G/C	C	0.493	0.479
rs7045668	chr9:110849139	miR_32	A/G	G	0.142	0.166
rs10512391	chr9:110849351	miR_32	A/G	G	0.241	0.269
rs839011	chr9:110850394	miR_32	C/G	G	0.194	0.241
rs10979694	chr9:110922615	C9orf5_miR_32	C/T	T	0.146	0.184
rs10816772	chr9:110925690	C9orf5_miR_32	T/C	C	0.221	0.243
rs10984879	chr9:122045392	miR_147	T/A	T	0.398	0.405
rs10739558	chr9:122051033	miR_147	C/T	C	0.151	0.151
rs4838200	chr9:126490878	cluster9.3	C/T	C	0.493	0.479
rs6478680	chr9:126498289	cluster9.3	G/A	G	0.499	0.471
rs10987942	chr9:130041513	miR_199b	G/C	C	0.249	0.230
rs1004464	chr9:130043498	miR_199b	A/G	G	0.182	0.176
rs11789170	chr9:130045181	miR_199b	A/C	C	0.209	0.183
rs2240943	chr9:130193738	miR_219_2	A/G	G	0.034	0.016
rs4880117	chr9:138672282	EGFL7_miR_126	C/T	C	0.046	0.069
rs7041558	chr9:138678948	EGFL7_miR_126	A/G	A	0.465	0.463
rs2297537	chr9:138684295	miR_126	C/G	G	0.414	0.394
rs1051828	chr9:138686644	miR_126	A/G	A	0.166	0.198
rs1051851	chr9:138686817	miR_126	G/A	A	0.229	0.204
rs525830	chr10:17925916	miR_511_1 & miR_511_2	A/C	NA	NA	0.367
rs2437258	chr10:17931828	miR_511_1 & miR_511_2	C/T	T	0.209	0.237
rs691461	chr10:17933215	miR_511_1 & miR_511_2	C/T	C	0.494	0.489
rs2607863	chr10:88014423	miR_346	T/C	T	0.065	0.063

rs10887569	chr10:88014533	miR_346	C/T	C	0.357	0.333
rs3814614	chr10:88118261	GRID1_miR_346	A/G	A	0.484	0.481
rs11185773	chr10:91339311	miR_107	T/C	C	0.136	0.124
rs7095066	chr10:91339738	miR_107	A/G	G	0.121	0.122
rs17389917	chr10:91341894	miR_107	T/A	A	0.075	0.079
rs2296616	chr10:91342946	miR_107	G/A	G	0.481	0.465
rs10509578	chr10:91392796	PANK1_miR_107	G/A	A	0.034	0.029
rs17125278	chr10:91392857	PANK1_miR_107	T/C	C	0.255	0.221
rs997456	chr10:91393960	PANK1_miR_107	G/A	A	0.245	0.272
rs1409312	chr10:104181790	miR_146b	C/T	C	0.201	0.196
rs1572530	chr10:104183824	miR_146b	G/A	A	0.208	0.180
rs12355840	chr10:134911103	miR_202	C/T	C	0.198	0.205
rs3124440	chr10:134912329	miR_202	T/A	T	0.139	0.172
rs4838710	chr10:134913018	miR_202	G/T	G	0.308	0.328
rs11101659	chr10:134913643	miR_202	C/T	T	0.077	0.093
rs7927267	chr11:557391	miR_210	A/G	A	0.384	0.399
rs2585	chr11:2107020	miR_483	T/C	T	0.291	0.349
rs734351	chr11:2112789	miR_483	G/A	G	0.369	0.409
rs4244808	chr11:2119686	IGF2_miR_483	T/G	G	0.395	0.365
rs2679059	chr11:43556367	miR_129_2	A/T	A	0.405	0.415
rs731384	chr11:57164958	miR_130a	G/A	A	0.305	0.307
rs1783827	chr11:57166114	miR_130a	G/A	G	0.431	0.413
rs11606204	chr11:64412645	cluster11.1	G/A	A	0.075	0.050
rs7944276	chr11:64413333	cluster11.1	G/A	A	0.114	0.132
rs7949144	chr11:64414975	cluster11.1	G/C	C	0.071	0.056
rs11231898	chr11:64415412	cluster11.1	G/A	A	0.002	0.003
rs7939874	chr11:64417977	cluster11.1	G/A	A	0.301	0.370
rs10793036	chr11:71999827	miR_139	C/T	T	0.484	0.455
rs755933	chr11:72003130	miR_139	G/A	A	0.142	0.140
rs341058	chr11:72063373	PDE2A_miR_139	C/T	C	0.127	0.103
rs11823971	chr11:72066209	PDE2A_miR_139	G/A	A	0.077	0.066
rs616191	chr11:74722446	miR_326	T/C	C	0.466	0.428
rs476364	chr11:74723548	miR_326	C/G	G	0.351	0.423
rs2851459	chr11:74726702	miR_326	G/A	A	0.288	0.294
rs2510894	chr11:74739826	ARRB1_miR_326	T/G	T	0.438	0.495
rs12278108	chr11:74744137	ARRB1_miR_326	T/C	C	0.179	0.140
rs955744	chr11:110879544	cluster11.2	C/T	T	0.410	0.376
rs17662588	chr11:110894604	cluster11.2	G/A	A	0.056	0.037
rs2156580	chr11:110899538	cluster11.2	C/G	G	0.358	0.344
rs664409	chr11:121469779	cluster11.3	G/A	A	0.388	0.375
rs2081443	chr11:121475956	cluster11.3	T/G	G	0.179	0.146
rs17126105	chr11:121482538	cluster11.3	T/G	G	0.053	0.045
rs547008	chr11:121523071	cluster11.3	G/A	A	0.157	0.135
rs562052	chr11:121523760	cluster11.3	T/C	T	0.329	0.312
rs537989	chr11:121525538	cluster11.3	C/A	C	0.166	0.152
rs4936674	chr11:121526493	cluster11.3	C/A	A	0.102	0.101
rs1834306	chr11:121528397	cluster11.3	A/G	A	0.435	0.428
rs638742	chr11:121531148	cluster11.3	C/G	C	0.494	0.487
rs7966756	chr12:6932652	cluster12.1	A/G	A	0.075	0.071
rs759052	chr12:6939881	cluster12.1	T/C	T	0.121	0.128
rs16933011	chr12:6944066	cluster12.1	G/T	T	0.075	0.108
rs2159887	chr12:6946497	cluster12.1	G/C	G	0.176	0.161
rs12304647	chr12:52671214	miR_196a_2	A/C	C	0.252	0.220
rs11614913	chr12:52671866	miR_196a_2	C/T	T	0.369	0.389

rs1109391	chr12:52673644	miR_196a_2	G/T	G	0.409	0.433
rs12231393	chr12:52999040	COPZ1_miR_148b	T/C	C	0.102	0.114
rs4759076	chr12:53016139	miR_148b	T/C	C	0.433	0.479
rs11170877	chr12:53020556	miR_148b	A/G	G	0.099	0.108
rs11172351	chr12:56499752	miR_26a_2	C/T	T	0.253	0.234
rs870392	chr12:56505440	miR_26a_2	C/A	A	0.224	0.183
rs1599750	chr12:56528812	CTDSP2_miR_26a_2	T/G	G	0.223	0.175
rs10877886	chr12:61282328	let_7i	G/A	A	0.358	0.330
rs12818198	chr12:61288340	let_7i	A/G	G	0.268	0.285
rs7962769	chr12:93750359	miR_492	T/G	G	0.181	0.140
rs2289030	chr12:93752417	miR_492	G/C	G	0.065	0.040
rs2289029	chr12:93752762	miR_492	T/C	T	0.031	0.034
rs6538555	chr12:93758409	miR_492	G/A	G	0.253	0.213
rs7966079	chr12:94223371	miR_331	T/C	C	0.484	0.495
rs11107973	chr12:94226516	miR_331	T/C	C	0.385	0.357
rs17023960	chr12:94228714	miR_331	A/C	C	0.050	0.092
rs11109106	chr12:96476013	mir_135a_2	G/T	T	0.174	0.180
rs10860218	chr12:96479233	mir_135a_2	T/C	C	0.158	0.157
rs11109108	chr12:96482239	mir_135a_2	T/C	T	0.290	0.310
rs17334962	chr12:96484192	mir_135a_2	C/T	T	0.072	0.063
rs2476391	chr13:49516505	cluster13.1	T/C	C	0.081	0.082
rs9535416	chr13:49522141	cluster13.1	G/A	G	0.494	0.457
rs2740545	chr13:49524760	cluster13.1	T/C	C	0.204	0.148
rs2177313	chr13:49527409	cluster13.1	T/C	C	0.162	0.176
rs2066575	chr13:49554583	DLEU2_cl13.1	T/A	A	0.202	0.148
rs12100048	chr13:49555363	DLEU2_cl13.1	G/T	T	0.143	0.127
rs790939	chr13:49558882	DLEU2_cl13.1	T/A	T	0.081	0.087
rs6492538	chr13:90791747	cluster13.2	A/C	A	0.176	0.180
rs17642969	chr13:90794757	cluster13.2	T/C	C	0.118	0.085
rs4284505	chr13:90799473	cluster13.2	A/G	A	0.407	0.360
rs9589207	chr13:90801590	cluster13.2	G/A	A	0.000	0.000
rs7318578	chr13:90803470	cluster13.2	C/A	C	0.270	0.274
rs1428	chr13:90804771	cluster13.2	C/A	C	0.380	0.341
rs769040	chr13:90808658	cluster13.2	C/G	C	0.177	0.168
rs8006357	chr14:22923469	miR_208	T/C	C	0.297	0.288
rs178642	chr14:22926736	miR_208	A/G	G	0.472	0.460
rs17091434	chr14:22930548	miR_208	G/T	T	0.087	0.087
rs365990	chr14:22931651	miR_208	A/G	G	0.386	0.346
rs7153499	chr14:22945833	MYH6_miR_208	T/C	C	0.124	0.122
rs12147570	chr14:22951760	MYH6_miR_208	G/T	T	0.142	0.116
rs2284651	chr14:22951984	MYH6_miR_208	T/C	C	0.384	0.365
rs1190956	chr14:99602225	EVL_miR_342	T/C	T	0.121	0.130
rs3783330	chr14:99642707	miR_342	G/T	G	0.350	0.346
rs748354	chr14:99646748	miR_342	T/C	T	0.114	0.094
rs2146026	chr14:99650821	miR_342	A/G	A	0.387	0.384
rs1059264	chr14:99838916	miR_345	C/T	T	0.329	0.325
rs1059263	chr14:99839120	miR_345	C/A	A	0.317	0.362
rs8007427	chr14:99841325	miR_345	C/T	T	0.225	0.222
rs4905945	chr14:99847546	miR_345	T/C	T	0.340	0.354
rs721910	chr14:100399462	cluster14.1	C/A	A	0.420	0.434
rs721909	chr14:100399556	cluster14.1	G/A	A	0.254	0.280
rs8008201	chr14:100399679	cluster14.1	G/A	A	0.178	0.209
rs1950622	chr14:100400287	cluster14.1	G/A	G	0.331	0.298
rs11627690	chr14:100402567	cluster14.1	G/C	C	0.279	0.301

rs8021312	chr14:100403399	cluster14.1	T/C	C	0.211	0.239
rs1077412	chr14:100410387	cluster14.1	A/G	G	0.290	0.285
rs11851174	chr14:100415257	cluster14.1	C/T	T	0.206	0.225
rs12884005	chr14:100417161	cluster14.1	G/A	A	0.000	0.000
rs11623267	chr14:100418337	cluster14.1	C/G	G	0.235	0.245
rs3825569	chr14:100420051	cluster14.1	T/C	T	0.332	0.328
rs3809404	chr14:100425763	cluster14.1	A/G	G	0.441	0.474
rs11852040	chr14:100442494	miR_370	T/G	G	0.143	0.153
rs1956124	chr14:100449349	miR_370	T/G	G	0.331	0.338
rs1956125	chr14:100451317	miR_370	A/G	G	0.391	0.384
rs12435162	chr14:100557661	cluster14.2	T/A	A	0.037	0.029
rs12892719	chr14:100559611	cluster14.2	C/T	T	0.235	0.265
rs12435862	chr14:100560165	cluster14.2	A/G	G	0.399	0.410
rs8018987	chr14:100568527	cluster14.2	G/A	A	0.240	0.267
rs11621499	chr14:100574993	cluster14.2	A/G	G	0.098	0.103
rs7161441	chr14:100582998	cluster14.2	G/T	T	0.223	0.228
rs4525426	chr14:100594909	cluster14.2	A/G	G	0.349	0.352
rs7342570	chr14:100595664	cluster14.2	G/A	A	0.171	0.173
rs8003403	chr14:100599066	cluster14.2	G/A	A	0.199	0.220
rs8015875	chr14:100603041	cluster14.2	A/G	G	0.230	0.228
rs8016185	chr14:100603137	cluster14.2	C/T	T	0.176	0.175
rs8016966	chr14:100603678	cluster14.2	A/G	G	0.096	0.090
rs10151229	chr14:100604060	cluster14.2	A/G	G	0.496	0.428
rs4900482	chr14:100604512	cluster14.2	C/T	T	0.213	0.243
rs941714	chr14:100605882	cluster14.2	A/G	A	0.279	0.324
rs1557012	chr14:103650825	miR_203	C/T	T	0.166	0.148
rs1771007	chr14:103654748	miR_203	T/C	T	0.134	0.116
rs919001	chr15:29144430	miR_211	A/G	A	0.431	0.391
rs8039189	chr15:29146495	miR_211	G/T	T	0.120	0.165
rs3784599	chr15:29146749	miR_211	T/G	G	0.487	0.476
rs3784600	chr15:29146844	miR_211	C/T	T	0.089	0.132
rs8033330	chr15:29183925	TRPM1_miR_211	G/A	A	0.184	0.196
rs17815839	chr15:29185149	TRPM1_miR_211	G/C	C	0.127	0.201
rs4638523	chr15:60723313	TLN2_miR_190	C/G	G	0.087	0.124
rs11852582	chr15:60724122	TLN2_miR_190	C/T	C	0.123	0.128
rs2899683	chr15:60725461	TLN2_miR_190	G/A	G	0.285	0.291
rs16945799	chr15:60900855	miR_190	T/C	C	0.130	0.122
rs12595365	chr15:60905890	miR_190	A/G	A	0.189	0.159
rs12594144	chr15:61948404	miR_422a	C/A	A	0.117	0.090
rs16947523	chr15:61949855	miR_422a	G/A	A	0.067	0.063
rs12324851	chr15:61950548	miR_422a	G/A	G	0.241	0.275
rs12592616	chr15:61953027	miR_422a	C/G	G	0.288	0.262
rs7177356	chr15:77285515	miR_184	A/G	A	0.490	0.463
rs12903401	chr15:77289151	miR_184	G/C	G	0.465	0.495
rs4779008	chr15:77292976	miR_184	T/C	C	0.145	0.130
rs16970715	chr15:77293946	miR_184	T/C	C	0.416	0.402
rs8023916	chr15:86951532	miR_7_2	T/C	C	0.460	0.452
rs11856451	chr15:86958990	miR_7_2	G/C	C	0.350	0.335
rs176644	chr15:87714636	miR_9_3	G/T	T	0.385	0.397
rs176645	chr15:87718874	miR_9_3	C/G	C	0.426	0.407
rs878340	chr16:14302659	cluster16.1	T/C	C	0.395	0.421
rs30235	chr16:14307294	cluster16.1	C/T	C	0.339	0.348
rs30232	chr16:14309463	cluster16.1	G/A	G	0.339	0.328
rs30229	chr16:14311723	cluster16.1	G/T	G	0.292	0.295

rs12924081	chr16:14315676	cluster16.1	C/T	T	0.152	0.140
rs12449163	chr16:15643012	miR_484	T/C	C	0.459	0.434
rs877063	chr16:15645359	miR_484	C/T	T	0.064	0.064
rs1436424	chr16:55452535	miR_138_2	G/T	T	0.487	0.449
rs12920659	chr16:55453374	miR_138_2	C/T	C	0.150	0.130
rs12921781	chr16:55453690	miR_138_2	A/G	G	0.258	0.291
rs11700	chr16:65790185	miR_328	A/G	G	0.075	0.066
rs12051247	chr16:65798783	miR_328	C/T	T	0.077	0.066
rs4985523	chr16:68354693	WWP2_miR_140	T/A	A	0.217	0.217
rs904808	chr16:68515996	WWP2_miR_140	T/C	T	0.192	0.198
rs2270838	chr16:68516724	WWP2_miR_140	G/A	A	0.310	0.304
rs7205289	chr16:68524506	miR_140	C/A	A	0.000	0.000
rs8047818	chr16:68525233	miR_140	G/A	G	0.200	0.210
rs3762177	chr16:68529466	miR_140	A/G	A	0.189	0.214
rs8076112	chr17:1558949	miR_22	A/C	C	0.133	0.175
rs6502892	chr17:1564358	miR_22	T/C	T	0.357	0.339
rs4790814	chr17:1568927	miR_22	G/A	A	0.479	0.471
rs11870150	chr17:1895351	cluster17.1	C/T	T	0.384	0.349
rs8065878	chr17:1898829	cluster17.1	C/T	T	0.297	0.332
rs14309	chr17:6859817	cluster17.2	T/C	T	0.127	0.135
rs312460	chr17:6866854	cluster17.2	A/C	A	0.055	0.063
rs2074222	chr17:7070698	miR_324	A/G	A	0.373	0.365
rs222837	chr17:7073280	miR_324	C/T	C	0.437	0.444
rs901975	chr17:24213900	cluster17.3	A/G	A	0.186	0.189
rs4294864	chr17:25464264	miR_423	G/C	C	0.488	0.471
rs7216082	chr17:26914065	miR_193a	A/G	G	0.118	0.122
rs1551360	chr17:26915568	miR_193a	T/C	T	0.275	0.207
rs121224	chr17:26926113	miR_365_2	G/C	G	0.336	0.348
rs223143	chr17:26927449	miR_365_2	C/T	C	0.252	0.238
rs7209072	chr17:43465468	miR_152	A/G	G	0.104	0.108
rs2240119	chr17:43466171	miR_152	T/C	C	0.110	0.093
rs7210041	chr17:43468821	COPZ2_miR_152	C/T	T	0.099	0.101
rs9907987	chr17:44014227	miR_10a	T/C	T	0.083	0.077
rs2288278	chr17:44016291	miR_10a	G/A	G	0.306	0.312
rs2288280	chr17:44016571	miR_10a	C/T	T	0.082	0.090
rs2303487	chr17:44059535	miR_196a_1	G/T	G	0.391	0.434
rs16942777	chr17:44074689	miR_196a_1	G/A	A	0.041	0.043
rs2680691	chr17:53760857	miR_142	T/C	C	0.245	0.274
rs2632516	chr17:53764088	miR_142	G/C	C	0.475	0.463
rs3744097	chr17:53767652	miR_142	C/T	T	0.037	0.021
rs7342863	chr17:54581792	miR_301	G/A	A	0.269	0.265
rs7502947	chr17:54584498	miR_301	G/A	G	0.348	0.345
rs7217845	chr17:54591682	FAM33A_miR_301	G/A	A	0.265	0.265
rs17681660	chr17:55270206	miR_21	T/G	G	0.112	0.114
rs1292037	chr17:55273690	miR_21	T/C	C	0.179	0.201
rs13137	chr17:55273813	miR_21	A/T	A	0.179	0.201
rs4969364	chr17:76637220	miR_338	G/A	A	0.205	0.196
rs8073904	chr17:76709739	miR_338	G/A	G	0.107	0.154
rs4969253	chr17:76720445	AATK_ miR_338	T/C	T	0.170	0.197
rs4075838	chr17:76720746	AATK_ miR_338	G/T	G	0.130	0.146
rs4969415	chr17:76757647	AATK_ miR_338	C/G	G	0.302	0.307
rs11872553	chr18:17574821	cluster18.1	G/A	A	0.099	0.090
rs12968949	chr18:17658855	cluster18.1	T/C	C	0.193	0.199
rs8089787	chr18:17660599	cluster18.1	T/C	T	0.103	0.090

rs4591246	chr18:17662527	cluster18.1	G/A	A	0.212	0.238
rs3810042	chr18:17666529	cluster18.1	A/G	A	0.316	0.328
rs2155860	chr18:31745681	miR_187	A/G	A	0.055	0.095
rs4940704	chr18:54266487	miR_122a	A/G	A	0.177	0.172
rs9319929	chr18:54266805	miR_122a	A/C	A	0.268	0.251
rs10514884	chr18:54270132	miR_122a	C/T	T	0.251	0.225
rs11663046	chr18:54270609	miR_122a	G/C	C	0.186	0.222
rs12979279	chr19:4715306	C19orf30_miR_7_3	A/G	G	0.196	0.197
rs10408455	chr19:4716524	C19orf30_miR_7_3	T/G	G	0.167	0.184
rs11668223	chr19:4718780	miR_7_3	A/G	G	0.189	0.205
rs3760955	chr19:4723202	miR_7_3	A/C	C	0.365	0.374
rs3760954	chr19:4723343	miR_7_3	T/C	C	0.167	0.181
rs3786719	chr19:10788100	miR_199a_1	G/C	C	0.366	0.349
rs11085748	chr19:10788540	miR_199a_1	T/C	C	0.320	0.290
rs1005039	chr19:10789081	miR_199a_1	C/T	T	0.004	0.003
rs3745453	chr19:13803221	cluster19.1	A/G	G	0.283	0.262
rs1531212	chr19:13812830	cluster19.1	G/A	A	0.217	0.189
rs897790	chr19:13849040	cluster19.2	C/T	T	0.222	0.191
rs10413657	chr19:13851126	cluster19.2	G/A	A	0.220	0.190
rs7252448	chr19:50833893	miR_330	C/T	T	0.252	0.270
rs2286755	chr19:50836435	miR_330	C/G	G	0.176	0.206
rs2286756	chr19:50836865	miR_330	G/T	G	0.459	0.476
rs2280401	chr19:54691821	miR_150	G/A	A	0.187	0.164
rs2077300	chr19:54695065	miR_150	C/T	T	0.153	0.164
rs1868941	chr19:56879795	cluster19.3	A/C	C	0.226	0.196
rs12975333	chr19:56888340	cluster19.3	G/T	T	0.000	0.000
rs8109852	chr19:56889942	cluster19.3	T/G	G	0.198	0.231
rs2167418	chr19:56896754	cluster19.3	G/A	A	0.218	0.238
rs6509807	chr19:58859082	cluster19.4	G/A	G	0.277	0.261
rs7259011	chr19:58865759	cluster19.4	A/G	A	0.040	0.042
rs4803160	chr19:58866849	cluster19.4	G/A	G	0.274	0.259
rs7258128	chr19:58869085	cluster19.4	T/C	C	0.333	0.365
rs10853864	chr19:58872038	cluster19.4	G/A	A	0.159	0.183
rs13345110	chr19:58878506	cluster19.4	T/A	A	0.049	0.059
rs2060230	chr19:58880031	cluster19.4	G/A	G	0.148	0.162
rs987021	chr19:58887052	cluster19.4	G/A	A	0.415	0.425
rs12232836¥	chr19:58895820	cluster19.4	T/C	C	0.114	0.116
rs7254177	chr19:58898500	cluster19.4	T/C	C	0.212	0.203
rs7255628	chr19:58902546	cluster19.4	G/C	C	0.000	0.000
rs2865851	chr19:58907560	cluster19.4	A/G	G	0.319	0.362
rs2865857	chr19:58908718	cluster19.4	A/G	G	0.373	0.421
rs13382089	chr19:58911666	cluster19.4	G/T	T	0.000	0.000
rs2082432	chr19:58913626	cluster19.4	G/A	G	0.466	0.420
rs7248768	chr19:58918275	cluster19.4	T/G	T	0.326	0.275
rs4803168	chr19:58919098	cluster19.4	G/A	G	0.339	0.285
rs7253869	chr19:58923089	cluster19.4	C/T	T	0.050	0.040
rs3859497	chr19:58927419	cluster19.4	A/T	A	0.445	0.388
rs11667544	chr19:58932852	cluster19.4	A/G	G	0.044	0.037
rs7254493	chr19:58935578	cluster19.4	A/G	G	0.132	0.124
rs4806645	chr19:58940187	cluster19.4	C/T	C	0.448	0.439
rs8106202	chr19:58946998	cluster19.4	C/T	T	0.286	0.275
rs3902015	chr19:58948744	cluster19.4	A/G	A	0.397	0.407
rs305941	chr19:58949799	cluster19.4	C/T	T	0.148	0.175
rs683060	chr19:58953481	cluster19.4	G/T	G	0.268	0.259

rs674315	chr19:58959953	cluster19.4	A/C	C	0.258	0.261
rs10427077	chr19:58978708	cluster19.5	G/C	C	0.229	0.249
rs3848577	chr19:58980313	cluster19.5	C/T	C	0.127	0.093
rs3848580	chr19:58980658	cluster19.5	C/G	G	0.491	0.444
rs12983508	chr19:58982532	cluster19.5	C/G	G	0.140	0.133
rs12983273	chr19:58983644	cluster19.5	C/T	T	0.120	0.122
rs3859502	chr19:58984844	cluster19.5	G/C	C	0.128	0.138
rs3859503	chr19:58986212	cluster19.5	A/C	C	0.152	0.141
rs17305227	chr19:58986294	cluster19.5	C/T	T	0.044	0.042
rs6052147	chr20:3816355	PANK2_miR_103_2	A/G	G	0.074	0.088
rs6116087	chr20:3836337	miR_103_2	A/G	A	0.258	0.307
rs6052170	chr20:3844742	miR_103_2	C/T	C	0.263	0.311
rs6084513	chr20:3845014	miR_103_2	C/A	C	0.488	0.497
rs6120777	chr20:33023833	MYH7B_miR_499	G/A	A	0.163	0.217
rs6087657	chr20:33028350	MYH7B_miR_499	G/A	A	0.242	0.238
rs6120778	chr20:33028830	MYH7B_miR_499	T/C	T	0.491	0.457
rs7261167	chr20:33039866	miR_499	G/A	A	0.069	0.063
rs1885120	chr20:33040650	miR_499	C/G	C	0.035	0.032
rs1885114	chr20:33041022	miR_499	A/G	G	0.494	0.455
rs3746444	chr20:33041912	miR_499	A/G	G	0.198	0.241
rs2425009	chr20:33042560	miR_499	T/C	T	0.237	0.237
rs6015380	chr20:56821249	miR_296	G/A	A	0.301	0.259
rs13042402	chr20:56828288	miR_296	G/A	A	0.302	0.274
rs6062278	chr20:60555386	C20orf166_cl20.1	A/G	A	0.462	0.468
rs6122014	chr20:60561960	cluster20.1	G/A	A	0.000	0.000
rs6062267	chr20:60563810	cluster20.1	T/C	T	0.471	0.489
rs6062261	chr20:60565961	cluster20.1	A/G	A	0.315	0.282
rs11906462	chr20:60569397	cluster20.1	C/T	T	0.214	0.238
rs4072471	chr20:60573809	cluster20.1	T/C	T	0.368	0.368
rs11698727	chr20:60575198	cluster20.1	C/T	T	0.224	0.249
rs6062234	chr20:60576985	cluster20.1	C/T	C	0.457	0.484
rs2064611	chr20:61282165	miR_124_a_3	A/G	G	0.294	0.312
rs2064612	chr20:61283377	miR_124_a_3	T/C	T	0.388	0.349
rs17209757	chr21:16363516	c21orf34_cl21.1	A/G	G	0.052	0.059
rs9984017	chr21:16363636	c21orf34_cl21.1	T/C	C	0.071	0.082
rs2823596	chr21:16364888	c21orf34_cl21.1	T/C	C	0.478	0.492
rs2051347	chr21:16485140	c21orf34_cl21.1	C/A	C	0.372	0.367
rs238952	chr21:16485565	c21orf34_cl21.1	T/C	T	0.049	0.058
rs17276131	chr21:16834367	cluster21.1	T/C	C	0.041	0.048
rs11702768	chr21:16852276	cluster21.1	C/T	T	0.127	0.093
rs2823884	chr21:16857333	cluster21.1	T/A	A	0.116	0.104
rs2823888	chr21:16865351	cluster21.1	C/T	T	0.072	0.071
rs17210190	chr21:16874432	cluster21.1	T/C	C	0.063	0.077
rs7279730	chr21:16883415	cluster21.1	A/G	G	0.176	0.178
rs2829803	chr21:25870181	miR_155	G/A	G	0.258	0.238
rs8132093	chr21:25872325	miR_155	A/G	A	0.134	0.130
rs2531716	chr22:18384704	C22orf25_miR_185	C/T	C	0.330	0.246
rs2531700	chr22:18386144	C22orf25_miR_185	T/C	T	0.287	0.288
rs2078749	chr22:18399768	miR_185	A/G	A	0.416	0.481
rs6518597	chr22:18401863	miR_185	C/T	T	0.362	0.434
rs2520733	chr22:18402498	miR_185	C/G	C	0.049	0.045
rs861844	chr22:20336219	miR_130b	G/T	T	0.193	0.180
rs447001	chr22:20340918	miR_130b	C/T	T	0.372	0.434
rs5758487¥	chr22:40556341	SREBF2_miR_33	C/T	T	0.449	0.426

rs9620000	chr22:40626649	miR_33	T/C	C	0.072	0.074
rs17379759	chr22:40629481	miR_33	A/G	G	0.127	0.116
rs11090910	chr22:44886334	cluster22.1	T/C	C	0.301	0.225
rs11090911	chr22:44886442	cluster22.1	C/T	T	0.074	0.045
rs4823529	chr22:44889405	cluster22.1	G/A	G	0.075	0.090
rs7062217	chrX:45485613	clusterX.1	A/G	G	0.066	0.034
rs2745712	chrX:45485908	clusterX.1	G/C	C	0.201	0.156
rs2858060	chrX:45491952	clusterX.1	C/G	G	0.308	0.297
rs974704	chrX:45495475	clusterX.1	G/A	G	0.125	0.122
rs179830	chrX:49663097	clusterX.2	A/G	G	0.245	0.222
rs12387959	chrX:49669031	clusterX.2	G/A	G	0.273	0.275
rs179833	chrX:49669782	clusterX.2	C/A	A	0.423	0.407
rs12559227	chrX:53597922	clusterX.3	T/C	C	0.003	0.005
rs17276588	chrX:53601143	clusterX.3	G/A	A	0.022	0.026
rs5933596	chrX:53603290	clusterX.3	C/T	T	0.000	0.008
rs2036658	chrX:53699057	HUWE1_clX.3	C/G	C	0.444	0.481
rs3848900	chrX:65154736	miR_223	A/G	G	0.139	0.114
rs1044165	chrX:65158451	miR_223	G/A	A	0.135	0.148
rs174145	chrX:73425515	miR_374	A/G	A	0.122	0.106
rs232964	chrX:76053352	clusterX.4	C/T	T	0.168	0.184
rs7049980	chrX:76141062	clusterX.4	A/G	G	0.139	0.118
rs10442398	chrX:85043859	miR_361	G/A	A	0.367	0.282
rs12391037	chrX:85047678	miR_361	A/G	G	0.372	0.277
rs1883795	chrX:85187415	CHM_miR_361	A/G	A	0.459	0.393
rs1883796	chrX:85187633	CHM_miR_361	G/A	G	0.445	0.370
rs475717	chrX:113721638	HTR2C_miR_448	T/G	G	0.134	0.208
rs498207	chrX:113724372	HTR2C_miR_448	A/G	G	0.302	0.383
rs3813928	chrX:113724538	HTR2C_miR_448	G/A	A	0.148	0.172
rs4243981	chrX:113963054	miR_448	T/C	T	0.127	0.201
rs4512568	chrX:113966377	miR_448	C/T	C	0.127	0.201
rs2851730	chrX:122519197	miR_220	T/G	G	0.086	0.074
rs2851732	chrX:122520082	miR_220	G/T	G	0.043	0.069
rs5911659	chrX:122522892	miR_220	G/A	A	0.494	0.476
rs5911663	chrX:122527266	miR_220	C/A	C	0.431	0.439
rs5975446	chrX:133129292	clusterX.5	G/A	A	0.066	0.056
rs5977965	chrX:133137554	clusterX.5	G/C	C	0.062	0.053
rs757308	chrX:133501182	clusterX.6	A/G	A	0.257	0.249
rs17396	chrX:133507278	clusterX.6	T/C	T	0.252	0.253
rs757309	chrX:133508213	clusterX.6	G/A	G	0.392	0.386
rs739817	chrX:133508668	clusterX.6	G/A	A	0.086	0.085
rs657	chrX:137573847	miR_504	T/G	G	0.084	0.071
rs5976185	chrX:137577663	miR_504	A/T	A	0.000	0.000
rs3761591	chrX:137622725	FGF13_miR_504	A/G	G	0.155	0.161
rs2157213	chrX:137622841	FGF13_miR_504	G/C	G	0.438	0.434
rs5931482	chrX:137650576	FGF13_miR_504	G/A	A	0.355	0.383
rs5908669	chrX:138833023	miR_505	G/A	G	0.021	0.016
rs5955049	chrX:138834364	miR_505	C/A	A	0.002	0.000
rs7060854	chrX:146102550	miR_513_1	A/G	G	0.012	0.013
rs5904722	chrX:146104700	miR_513_1	A/G	G	0.271	0.241
rs2018562	chrX:146114853	clusterX.7	T/C	C	0.295	0.241
rs6525827	chrX:146124174	clusterX.7	A/C	C	0.439	0.375
rs2504172	chrX:146147969	miR_509	C/T	T	0.000	0.000
rs5905024	chrX:146153703	miR_509	C/T	T	0.317	0.259
rs1597029	chrX:146158289	clusterX.8	C/T	T	0.164	0.159

rs2442040	chrX:146159067	clusterX.8	G/T	T	0.441	0.402
rs2392720	chrX:146159701	clusterX.8	A/G	G	0.139	0.138
rs2504169	chrX:146165697	clusterX.8	A/C	C	0.426	0.434
rs5905034	chrX:146176797	clusterX.8	A/G	G	0.266	0.241
rs1061420	chrX:150873428	clusterX.9	A/G	G	0.120	0.119
rs928953	chrX:150879418	clusterX.9	T/A	T	0.327	0.370
rs2256882	chrX:150880478	clusterX.9	G/A	A	0.145	0.164
rs2266856	chrX:150892187	GABRE_clX.9	T/C	T	0.294	0.296
rs2266858	chrX:150892631	GABRE_clX.9	T/A	A	0.457	0.410
rs5925077	chrX:150893453	GABRE_clX.9	C/T	T	0.377	0.299
rs5925079	chrX:150896187	GABRE_clX.9	A/G	A	0.262	0.314
rs5970291	chrX:151310797	clusterX.10	G/A	A	0.296	0.266
rs5970292	chrX:151311524	clusterX.10	G/A	G	0.402	0.343
rs5970293	chrX:151313926	clusterX.10	G/C	C	0.311	0.277
rs6627595	chrX:151368659	GABRA3_clX.10	G/C	C	0.404	0.398
rs7391474	chrX:151370364	GABRA3_clX.10	T/G	T	0.229	0.238

The minor alleles (MA) and minor allele frequencies (MAF) in the Spanish (Catalan) control population are indicated
Allele frequencies (AF) for the panic disorder patients are also showed
¥, SNPs that were not consistent with HWE

Supplementary Table 2. Primers used for miRNA sequencing and 3'UTR cloning

Name	Amplified region	Use	Primer sequence
miR-339F	miR-339	Sequencing	5'-TCCCTCCCCAGTTACTTCCT-3'
miR-339R	miR-339	Sequencing	5'-TGTGCGTCTTCCCTAACCTC-3'
miR-138-2F	miR-138-2	Sequencing	5'-CTCAATTTCTTGGTCCTGGTT-3'
miR-138-2R	miR-138-2	Sequencing	5'-CCTCTAACAATGCTGATTCCA-3'
miR-491F	miR-491	Sequencing	5'-TATGTCTGAAGAGTATGAGGGATT-3'
miR-491R	miR-491	Sequencing	5'-CCCAAGGATAAATCAGCCTAC-3'
miR-22F	miR-22	Sequencing	5'-ACTTTTGGACCTCATCTGGAA-3'
miR-22R	miR-22	Sequencing	5'-CACCCTGTCTGAATGCTGAG-3'
miR-488AF	miR-488	Sequencing	5'-GGGTCCTGTGATTCTACCAG-3'
miR-488AR	miR-488	Sequencing	5'-TCTCCCAGATTATCAGTCAGC-3'
miR-488BF	miR-488	Sequencing	5'-AAGCGTTTTCAGTTCTGGTTT-3'
miR-488BR	miR-488	Sequencing	5'-ACCTAGCACATCTGTCTGCTG-3'
ADORA2A-F	*ADORA2A*	3'UTR Cloning	5'-ACACACTCTAGAAGGATGGAGCAGGAGTGTC-3'
ADORA2A-R	*ADORA2A*	3'UTR Cloning	5'-ACACACTCTAGAGGCCTAGCTATGGAGTAGGAA-3'
BDNF-F	*BDNF*	3'UTR Cloning	5'-ACACACTCTAGACATTGACCATTAAAAGGGGAAG-3'
BDNF-R	*BDNF*	3'UTR Cloning	5'-ACACACTCTAGACCCAGAACCCCCAGATTTTA-3'
CCKAR-F	*CCKAR*	3'UTR Cloning	5'-ACACACTCTAGACTCTGTCCAGGTTCTCGTACA-3'
CCKAR-R	*CCKAR*	3'UTR Cloning	5'-ACACACTCTAGAAACTCCATTTGCAGTATGAGC-3'
CCKBR-F	*CCKBR*	3'UTR Cloning	5'-ACACACTCTAGACTGGTCTACTGCTTCATGCAC-3'
CCKBR-R	*CCKBR*	3'UTR Cloning	5'-ACACACTCTAGACCAGAAGAATCCAGAAAGGAA-3'
CRHR2-F	*CRHR2*	3'UTR Cloning	5'-ACACACTCTAGACAGCTTCCACAGCATCAAG3'
CRHR2-R	*CRHR2*	3'UTR Cloning	5'-ACACACTCTAGAACATCTGACTGGCTCTTCTCA-3'
GABRA6-F	*GABRA6*	3'UTR Cloning	5'-ACACACTCTAGATCTCGAATTCTCTTCCCAGTT-3'
GABRA6-R	*GABRA6*	3'UTR Cloning	5'-ACACACTCTAGAACATTGGGGTAGAGGTAACCA-3'
HTR2C-F	*HTR2C*	3'UTR Cloning	5'-ACACACTCTAGAGTGGTTAGCGAAAGGATTAGC-3'
HTR2C-R	*HTR2C*	3'UTR Cloning	5'-ACACACTCTAGATGGGCACGTATGGAAATACT-3'
MAOA-F	*MAOA*	3'UTR Cloning	5'-ACACACTCTAGAGACATCTAGTCTCAGTGCTAGCTT-3'
MAOA-R	*MAOA*	3'UTR Cloning	5'-ACACACTCTAGATTTCTGCATAGAGCCATCAAC-3'
NTRK3-F	*NTRK3*	3'UTR Cloning	5'-ACACACTCTAGAAATAAGCCTTCCCGGACATT–3'
NTRK3-R	*NTRK3*	3'UTR Cloning	5'-ACACACTCTAGATGCAAAATTTCCAAATAAGAGG–3'
POMC-F	*POMC*	3'UTR Cloning	5'-ACACACTCTAGACATCAAGAACGCCTACAAGAA-3'
POMC-R	*POMC*	3'UTR Cloning	5'-ACACACTCTAGAGGAAACCACTGTGCTCCTC-3'
RGS2-F	*RGS2*	3'UTR Cloning	5'-ACACACTCTAGAGGACTTGTGTAAAAAGCCACA-3'
RGS2-R	*RGS2*	3'UTR Cloning	5'-ACACACTCTAGACTTTATCTGCCACATTTCCAG-3'
SCL6A2-F	*SCL6A2*	3'UTR Cloning	5'-ACACACTCTAGAATCAAGTCCTCGCTGTCTTTC-3'
SCL6A2-R	*SCL6A2*	3'UTR Cloning	5'-ACACACTCTAGAAATTTGAGGCAAAGCACCTT-3'

Supplementary Table 3. Allele frequency comparison between Finnish and Estonian PD patients

miRNA	SNP	Allele	Allele frequencies Estonian, Finnish	Chi square	P value
miR_339	rs11763020	T	0.151, 0.102	2.619	0.1056
miR_339	rs1057558#	T	0.762, 0.722	1.035	0.309
C7orf50 -miR-339	rs10252922	A	0.582, 0.558	0.283	0.5949
C7orf50 -miR-339	rs12531999	T	0.788, 0.774	0.132	0.7159
miR_22	rs8076112#	C	0.158, 0.137	0.417	0.5183
miR_22	rs6502892	C	0.740, 0.694	1.264	0.2609
miR_22	rs4790814#	A	0.562, 0.481	3.188	0.0742
miR_491	rs4977831	A	0.072, 0.046	1.422	0.233
miR_491	rs2039391#	A	0.897, 0.897	0.0	0.9981
miR_491	rs12684948*	T	0.628, 0.519	5.999	0.0143
miR_491	rs7022666*	G	0.630, 0.505	8.012	0.0046
miR_148a	rs735316#	A	0.623, 0.565	1.68	0.195
miR_148a	rs4722551*#	T	0.807, 0.722	4.968	0.0258
miR_138_2	rs1436424	T	0.394, 0.393	0.0	0.9887
miR_138_2	rs12920659*	C	0.207, 0.134	4.506	0.0338
miR_138_2	rs12921781	A	0.719, 0.671	1.352	0.2449
miR_488	rs2076072	A	0.216, 0.173	1.43	0.2317
miR_488	rs12041859	C	0.185, 0.175	0.09	0.7644
miR_488	rs10753142	A	0.406, 0.388	0.173	0.6771

*, SNPs presenting significant allele frequency diffrences among both populations
#, SNPs for which 1000 Estonian controls were gentoyped

Association analysis of microRNA regions with obssesive compulsive dsorder in a Spanish population

Case control studies were performed for the 712 SNPs in the panel that passed QC criteria and were shown to be in Hardy-Weinberg equilibrium in our control population, in 210 patients with OCD and in 340 controls. The OCD sample consisted in 210 consecutive Spanish outpatients with OCD (107 males and 103 females, average age 33.6 (18-62)), recruited from the OCD Clinic and Research Unit of Bellvitge Hospital (Barcelona, Catalonia) who full-filled the criteria described in the first article of this thesis. Genotyping and association procedures were performed according to description provided at the Materials and Methods section of the third article included in this thesis.

Association analysis was performed under different models using logistic regression and adjusting for sex. Remarkably, association for rs4970362 (unadjusted p= 0.00005099, OR= 0.15; 95% CI= 0.05-0.49; AIC= 507.1; under the recessive model) tagging a miRNA cluster in chromosome 1 (cluster 1.1: miR-200b, miR-200a and miR-429) remained significant after correction for multiple testing ($p<0.00008525$) (table 1). The calculated odds ratios suggest that the homozygous AA genotype, with a frequency of the genotype of 15.1% in the controls and 2.5% in the OCD sample, could have a protective effect. Unfortunately, significance was lost when correction for the 2.5 effective tests was performed ($p<0.000034$). Another SNP, rs4970420, from the 5 tagging this cluster showed nominal association ($p<0.05$, under the codominant model). Moreover, rs7041558, in the

promoter region of the gene hosting miR-126, also showed a pvalue < 0.0005 (unadjusted p= 0.0004541, OR= 2.63; 95% CI= 1.47-4.70; AIC= 510.7; under the dominant model) (table 1). Interestingly, two out of the three SNPs tagging this miRNA (rs2297537 and rs1051851) also showed nominal association with OCD.

As sexual dimorphism has been suggested in OCD, we stratified our sample according to gender status and repeated the association analysis. Although several nominal associations showed up for both the male and female subgroups, only one association under a 0.0005 threshold was found in the case of female patients for one SNP tagging miR-383 under the dominant model (rs4602887; unadjusted p = 0.00026, OR=0.32; 95% CI=0.17-0.60; AIC=236.7).

Association of miRNA regions with OCD regarding the nature of the obsessions or compulsions

The Y-BOCS Symptom Checklist was employed to ascertain scores on the following symptom dimensions: symmetry/ordering, aggressive/checking, sexual/religious, contamination/cleaning and hoarding and analyses were, then, repeated taking these dimensions into account.

Interestingly, when stratifying the sample according to the presence of contamination or cleaning obsessions (93 patients with contamination obsessions and 117 without) the association of miR-126 increased in the sample without this type of

obsessions. Not only rs7041558 –tagging the promoter region of the gene hosting this miRNA- showed association under the p<0.0005 threshold (unadjusted p = 0.00021, OR= 4.52; 95% CI= 1.76-11.66; AIC= 328.8; under the dominant model), but this time rs2297537, tagging the same miRNA region, also showed association under this threshold (unadjusted p = 0.00018, OR= 0.08; 95% CI= 0.01-0.59; AIC= 328.5; under the recessive model). Again, rs1051851 showed nominal association with this sub-sample. It is worth mentioning that patients with non-contamination/cleaning subphenotype also showed association with a SNP tagging the brain-enriched miR-9-1, rs16837375 (unadjusted p = 0.00265, OR= 2.46; 95% CI= 1.38 - 4.38; AIC= 333.0; under the codominant model).

Similarly, the subgroup formed by those patients with aggressive/checking obsessions (146 patients with aggressive/checking obsessions and 64 without) also showed an increase in the association of miR-126 when compared with the whole OCD sample for SNPs rs7041558 (unadjusted p = 0.00021, OR= 3.61; 95% CI= 1.67 -7.80; AIC= 391.4; under the dominant model) and rs2297537 (unadjusted p = 0.00252, OR= 0.96; 95% CI= 0.59 -1.58; AIC= 405.1; under the recessive model). In addition a moderate association was found for rs2237336 (unadjusted p = 0,00048, OR= 2.13 for the homozygous genotype; 95% CI= 1.19-3.82; AIC= 391.8; under the codominant model) tagging miR-196b on this subgroup.

As for stratification of the OCD sample according to the presence of hoarding obsessions (56 hoarder patients and 154 non-

hoarder patients), the most significant associations were found for two SNPs tagging miR-490, rs12535371 (unadjusted p = 0.00056, OR= 0.19; 95% CI= 0.06 - 0.61; AIC= 404.9; under the recessive model) and rs2350780 (unadjusted p = 0.00032, OR= 2.56; 95% CI= 1.50 - 4.34; AIC= 403.2; under the dominant model)

The OCD sample was also stratified according to the presence of sexual/religious obsessions (50 patients with sexual/religious obsessions and 160 without). This analysis resulted in the association of a SNP tagging miR-449, rs3747725, (unadjusted p = 0.00016, OR= 4.01; 95% CI= 2.04-7.89; AIC= 162.9; under the log-additive model) in those with sexual/religious obsessions. The only other SNP tagging this miRNA region also showed nominal association with this subphenotype. In contrast, those without sexual/religious obsessions showed again association for the region containing miR-126 and for the same SNPs: rs7041558 (unadjusted p = 0.00023, OR= 3.15; 95% CI= 1.61 - 6.19; AIC= 442.9; under the dominant model), rs2297537 (unadjusted p = 0.00073, OR= 0.55; 95% CI= 0.38-0.79; AIC= 445.1; under the log-additive model) and rs1051851, showing nominal association.

Finally, the sample was stratified according to the presence of symmetry/ordering obsessions or compulsions (69 patients with sexual/religious obsessions and 141 without). For those OCD patients with symmetry/ordering obsessions and compulsions association was revealed for all the 7 SNPs tagging miR-128-1 rs1446584 (unadjusted p = 0.00004, OR= 2.60; 95% CI= 1.62 -

4.19; AIC= 260.8; under the log-additive model), rs1374330 (unadjusted p = 0.00012, OR= 2.61; 95% CI= 1.60 - 4.25; AIC= 262.8; under the log-additive model), rs2289959 (unadjusted p = 0.00145, OR= 2.90; 95% CI= 1.46 - 5.74; AIC= 267.5; under the dominant model), rs17331388 (unadjusted p = 0.00209, OR= 3.06; 95% CI= 1.55 - 6.05; AIC= 268; under the codominant model), rs3806502 (unadjusted p = 0.00289, OR= 2.61; 95% CI= 1.37 - 4.98; AIC= 268.0; under the codominant model) and nominal associations were found for rs3769018 and rs10194951. However, it is important to take into account that the first three SNPs described for this region and the ones showing stronger associations (rs1446584, rs1374330 and rs2289959) would be discarded according to significantly different allele frequencies observed between the Catalan and the European HapMap studied populations, and thus, would not be applicable for the study of our population. On the other hand, for those patients without this type of obsessions a strong association was found for rs7944276 tagging a miRNA cluster containing miR-192 and miR-194-2 (unadjusted p = 0.00006, OR=0.16; 95% CI= 0.05 -0.53; AIC= 370.8; under the log-additive model). Nominal association was also found for rs7939874, another SNP tagging this cluster.

Association of miRNA regions with OCD regarding age at onset (AAO)

OCD has a variable AAO beginning before age 25 but being able to start as soon as in childhood or adolescence. For this reason we decided to analyse a possible association of these miRNA

regions with AAO stratifying AAO into two groups, those with AAO under 10 years old and those with AAO under 15, in every case adjusting for sex.

Analysis for those presenting a lower AAO than 15 years old showed association for 3 SNPs of the four tagging miR-153-2: rs221295 (unadjusted p = 0.00022; under the codominant model), rs6459797 (unadjusted p = 0.00074; under the recessive model) and nominal association for rs7808805. This subgroup also showed association for one SNP tagging miR-107, rs2296616, (unadjusted p = 0.00018, OR=2.77; 95% CI= 1.57-4.88; AIC= 142.7; under the log-additive model).

As for the separate analysis of the group of OCD patients presenting an extremely low AAO (under 10 years old), association analysis revealed association of 5 of the 7 SNPs tagging miR-208 and the promoter region of the gene where it is hosted: rs2284651 (unadjusted p = 0.00036, under the log-additive model), rs365990 (unadjusted p = 0.00079; under the log-additive model), rs178642 (unadjusted p = 0.00484; under the dominant model) and nominal association for rs12147570 and rs17091434. Association with this group was also observed for rs2632516 (unadjusted p = 0.00042; under the dominant model) and rs2680691 (unadjusted p = 0.00581; under the codominant model), two of the three SNPs tagging miR-142.

Finally, haplotype and interaction analyses between the different SNPs showing association in our OCD samples were performed but no increase in the significance level was observed.

Table 1. Associated miRNAs with different OCD phenotypes and comparison to PD

miRNA	Chromosome	Association OCD (SNP)	Association OCD subphenotypes (SNP)	CNS expression
miR-200a, miR-200b, miR-429	Chr: 1	YES; (rs4970362, rs4970420)	YES	YES
miR-126	Chr: 9	YES; (rs7041558, rs2297537, rs1051851)	YES; Non-contamination/cleaning (rs7041558, rs2297537, rs1051851) YES; Aggressive/checking (rs7041558, rs2297537) YES; Non-sexual/religious (rs7041558, rs2297537, rs1051851)	YES
miR-383	Chr: 8	YES; females (rs4602887, rs4345555)	YES	YES (brain enriched)
miR-9-1	Chr: 1	YES; (rs16837375, rs10494305)	YES; Non-contamination/cleaning (rs16837375)	YES (brain enriched)
miR-196b	Chr: 7	YES; (rs886340, rs2237336)	YES; Aggressive/checking (rs2237336)	NO
miR-490	Chr: 7	YES; (rs2350780, rs1364402, rs12535371, rs6957496)	YES; Hoarding (rs12535371, rs2350780)	YES (brain enriched)
miR-449	Chr: 5	YES; (rs336115)	YES; sexual/religious (rs3747725, SNP)	NO
miR-128-1	Chr: 2	YES; (rs1446584#, rs1374330#, rs2289959#, rs17331388)	YES; symmetry/ordering (rs1446584#, rs1374330#, rs2289959#, rs17331388, rs3806502, rs3769018, rs10194951)	YES (brain enriched)
miR-192 and miR-194-2	Chr: 11	YES; (rs7939874, rs7944276)	YES; Non-symmetry/ordering (rs7944276, rs7939874)	NO

Associations identified with OCD subphenotypes are only explicited when the association identified for the subphenotypes was stronger than with the complete OCD sample

#, SNPs showing significant allele frequency differences between the Catalan and the European HapMap studied populations

Identification of common genetic factors underlying the susceptibility to panic disorder and OCD

With the aim of identifying common genetic factors underlying OCD and PD, the two anxiety disorders studied in this thesis, we analysed the intersections of those miRNAs that showed nominal associations with both disorders.

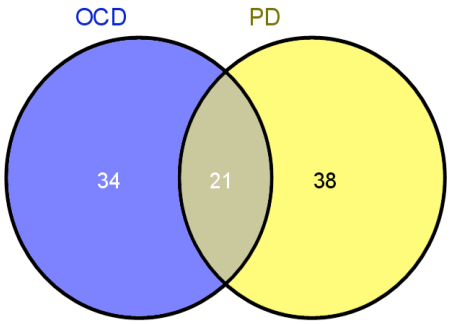


Figure 1: Intersection between the nominal associations identified with OCD and PD.

Remarkably, 21 of the associated miRNA regions are shared between OCD and PD, when analysed without stratifying for subphenotypes. In total, 36%-38% of the associated regions were common to both disorders. It is worthwhile to acknowledge that two of the miRNA regions showing particularly strong associations with PD -miR-22 and miR-491- are also associated with OCD; rs4790814 tagging the miR-22 region (unadjusted p = 0.03177, OR= 0.55; 95% CI= 0.31 - 0.97; AIC= 518.9; under the recessive model) and rs12684948 tagging the miR-491 (unadjusted p = 0.00963, OR= 0.69; 95% CI= 0.51 - 0.92; AIC= 514.4; under the log-additive model).

Finally, the intersection between PD, OCD, OCD Hoarding and OCD non-Hoarding showed that while OCD and OCD non-hoarding share about 40% of the nominally associated regions with PD, the Hoarding sample shared a 30%. In addition, almost half of the associations identified for the OCD Hoarders (44%) are not shared neither with the PD sample, nor with the non-Hoarding sample. These results may suggest that, as proposed in the *NTRK3* study and by others (Miguel et al., 2005; Samuels et al., 2007), OCD hoarding subclinical type has a different pattern of genetic inheritance and may constitute a neurobiologically and etiologically distinct variant of OCD.

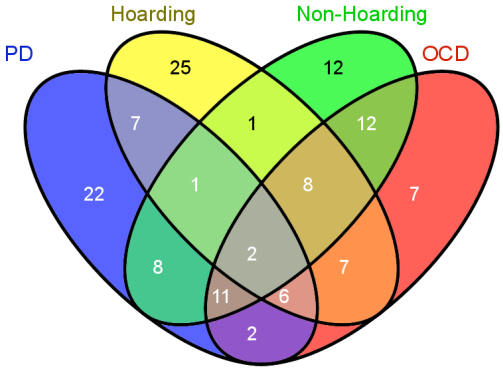


Figure 2: Intersection between the nominal associations identified in PD, OCD, OCD hoarding and OCD non-hoarding.

Discussion

Investigation of how genetic variation on complex gene regulatory networks results in phenotypic alterations may represent a useful approach towards the understanding of human evolution and disease. In this regard, genetic discoveries can contribute to the identification of genes and pathways underlying the susceptibility to psychiatric disorders. However, this has shown to be difficult due to the complexity of both, the genetics and the phenotypes of these disorders. In fact, even though the estimated heritability of psychiatric disorders is high, most genetic risk alleles for these disorders have still not been identified, leading to the conclusion that either major risk alleles are scarce or that these alleles increase risks only marginally, that is to say that their associated OR is low. Genetic heterogeneity for complex disorders is widely accepted and, in addition, it has been suggested that non-standard factors, such as epigenetic or regulatory changes or combinations of these factors could be involved in the aetiology of psychiatric disorders (Burmeister et al., 2008). In fact, there is increasing evidence at a population and experimental level that genetic variation at regulatory regions underlie differences in gene expression (Buckland et al., 2005; Knight, 2005; Rockman and Wray, 2002), which may be critical for the delicate equilibrium required for the proper function of the Central Nervous System (CNS).

Evidence from recent studies on genome-wide transcription demonstrate that more of the genome than the represented by current annotations is transcribed, and much of this transcribed RNA could be non-coding (Cao et al., 2006). In this context, miRNAs have arisen as key regulators of gene expression and cellular machinery and could be considered as one of the most significant discoveries in molecular biology of the last decade,

due to their enormous potential. Recent estimates indicate that miRNAs regulate more than 30% of all protein-coding genes, building complex regulatory networks that control almost every cellular process (Filipowicz et al., 2008). Posttranscriptional regulation by miRNAs may thus represent an important mechanism through which the CNS accomplishes its demands for precise but rapid changes in gene regulation. In fact, both the synthesis and degradation of RNAs are likely to require less time and energy than those of proteins. Consequently, non-coding RNAs and particularly, miRNAs, are suitable candidates for the regulation of a constantly changing microenvironment, like the CNS. In this respect, our study offers an innovative approach for the study of psychiatric disorders in general and particularly, for the study of anxiety disorders. In addition, this is to our knowledge, the first time that the whole collection of miRNAs have been analysed for association purposes.

This thesis is based on the hypothesis that changes in miRNA-mediated regulation, originated from changes in the miRNAs themselves or on their target sites, could alter the dosage of proteins involved in fundamental pathways of brain function, affecting the fine-tuned CNS homeostasis and contributing to the development of anxiety disorders. In the attempt of identifying functional regulatory polymorphisms we divided our strategy into the study of genetic variation in regulatory regions of a candidate gene for anxiety disorders, *NTRK3,* and the genetic study of miRNAs *"per se"*. Results on these studies have indicated that polymorphisms affecting miRNA-mediated regulation may be determinant of a range of human traits related to anxiety disorders. However, as stated before, major risk alleles have not been found. The study of the 3'UTRs of

NTRK3, for instance, led to the identification of several variants in the 3'UTR of the truncated isoform of this gene. Remarkably, one common SNP (rs28521337) was found to be associated with the hoarding sub-clinical type of OCD and two rare variants altered the regulation of *NTRK3* by three different miRNAs (miR-128, miR-509 and miR-765). Nonetheless, contribution to disease of the two rare allelic variants found remains unclear, due to the fact that these variants were only identified in one male patient with PD and agoraphobia, each. While it is widely accepted that rare allelic variants contribute little to heterozygosity, their putative role on disease cannot be ruled out as emphasized by the results on recent whole genome association studies, which have failed to identify major alleles for most of the disorders studied. These studies point out that susceptibility alleles are likely to be modest in effect size requiring large samples for detection (Sklar et al., 2008). Similarly, a rare allelic variant in the 3'UTR of *SLITRK1* affecting the binding of miR-189 was only identified in two patients with Tourette syndrome and in none of the controls tested (Abelson et al., 2005). However, scepticism surrounded the finding since other groups, such as Keen-Kim *et al*. (2006) were unable to replicate the study and identified this rare allelic variant among cases and controls within the same families. In their opinion these results indicate that the variant does not segregate with the disease and that the results from Abelson *et al.* might be confounded by hidden population stratification (Keen-Kim et al., 2006). Confounding might originate on the fact that the screening for new causative allelic variants was carried out only in cases. In order to exclude a possible ascertainment bias due to population stratification, in our study re-sequencing was performed on both, patients and controls, and population

homogeneity was assessed. Nevertheless, that the rare variant did not segregate completely with the disease may be explained, at least partly, by the fact that in Tourette syndrome, as in anxiety disorders, diagnoses may not always be correctly assessed as most cases are mild and may remain undiagnosed. Moreover, genetic and environmental factors are likely to be involved in their aetiology and that derives on the consequence that among family members of an affected person, it is difficult to predict who else may be at risk of developing the condition, in other words, their inheritance patterns are unclear. This fact underlies a heterogenic basis of the disorder, in which cases substantially differ from one another, even though they are considered to suffer from the same disease. Thus, most probably, environmental factors, genetic epistasis and/or accumulation of rare variants might be on the back of the susceptibility to these disorders. Consistently, the case-control studies that we performed with the SNP panel covering miRNA regions in PD or OCD resulted on moderate associations for several miRNA regions and, in addition, increase of risk in most of the cases was in general moderate as indicated by ORs and CIs. These findings are therefore in line with the genetic heterogeneity theory proposed for anxiety disorders.

On the other hand, association for the common SNP identified in the 3'UTR of *NTRK3*, rs28521337, did not resist correction for multiple testing in the OCD sample and was only statistically significant for the hoarding sub-clinical type, suggesting a different pattern of genetic inheritance for this group of patients, which would be in agreement with recent reports that indicate that hoarding subphenotype may constitute a neurobiologically and etiologically distinct variant of OCD (Miguel et al., 2005;

Samuels et al., 2007), being highly heritable as a quantitative trait (Mathews et al., 2007). This suggestion is further supported by the fact that intersection between the nominal associations found for PD, OCD, OCD hoarding and OCD non-hoarding showed that OCD non-hoarders are the group that show a higher proportion of associations that are not shared with other disorders, proposing that hoarding behaviour can indeed constitute a separate category of disorders. Actually, the fact that hoarding can be also associated to schizophrenia, dementia, anorexia, depression and even the so-called Diogenes syndrome (Rosenthal et al., 1999; Seedat and Stein, 2002) reinforces this proposition.

If anxiety disorders do exhibit genetic and or locus heterogeneity where multiple alleles (rare or common) contribute independently to marginal increases of risk; then, increasing sample sizes for their study would be of the utmost importance and the best method of identifying risk variants. Collecting this type of big samples is always arduous scientifically and bureaucratically time-consuming. Indeed, this is particularly true for psychiatric disorders, for which assessment of the phenotype is laborious and no objective experimental diagnostic has still been provided. However, several public and private repositories have been established, such as the National Institute of Mental Health and the Genetic Research Exchange, and other groups collaborate in meta-analyses of their samples. Even though the risk of increasing genetic heterogeneity by pooling data exists, these types of approximations are worth trying. On the other hand, problems arising from the controversial diagnosis of psychiatric disorders could result in some risk variants being associated only with

specific subtypes of a psychiatric disorder, as it was the case of OCD hoarding sub-phenotype. In such case, focusing the analysis on subtypes might allow to improve the power of association studies. In the association studies performed for this thesis, we have tried when possible and appropriate to use information on the different subtypes of PD and OCD. However, owing to correction for multiple phenotype testing restrictions and difficulty in obtaining the data we have not performed studies of the different anxiety endophenotypes (heritable and measurable phenotypes associated with the disease). The use of such traits, when relevant, may prove more powerful than imposing artificial thresholds. Moreover, the situation where variants only affect risk if they co-occur with other genetic or environmental risk factors may also be underlying the susceptibility to anxiety disorders. Although complicating analysis considerably, incorporating known genetic and environmental factors into the model could improve the power substantially (Burmeister et al., 2008). Regarding this attempt, and taking into account the existence of recall bias, the use of large scale longitudinal studies would be worth the effort, especially for anxiety disorders; disorders for which stressful life-events seem to have an impact on their onset (McLaughlin and Hatzenbuehler, 2009) and socio-economical reasons are important, due to their substantial social and health service costs (Gratacos et al., 2007). However, if accumulation of rare variants is contributing to the susceptibility to suffer from anxiety disorders, namely panic disorder, studies based on LD might be underestimating the contribution of the analysed gene. In other words, if heterogeneity was so high that many mutations were unique to each family, traditional methods of comparing allele frequency differences would not be successful,

except for isolated or inbred populations, and new high-throughput sequencing technologies would be needed (Burmeister et al., 2008). In fact, the improvement and price reduction of these technologies are starting to make this type of analysis feasible nowadays. Withal, once rare allelic variants are identified, their contribution to disease is always controversial and may be only assessed by means of functional experiments that demonstrate their possible involvement in disease. In this regard, the more functional approach used for this thesis offers a suitable and modern approximation for the study of anxiety disorders and has proven to be useful in order to identify candidate biological pathways to be involved in their pathophysiology.

Identification of genes targeted by miRNAs is a key step towards understanding the role of miRNAs in gene regulatory networks. As part of the effort to understand interactions between miRNAs and their targets, computational algorithms have been developed based on observed rules for features such as the degree of hybridization between the two RNA molecules. These *in silico* approaches provide important tools for miRNA target detection, and together with experimental validation, help to reveal regulated targets of miRNAs. Our first studies started out in 2005, when the use of target prediction methods was beginning to become popular, and was based on the 326 human miRNAs annotated in miRbase release 7.1. Taking into account the novelty of the miRNAs topic, the low amount of experimentally supported mammalian targets (to use as training sets by the programs' algorithms), and the low amount of time since the development of the first predictions, false positive rate estimations for the most commonly used programs are

satisfactory (about 30% for TargetScan, PicTar and miRanda; as reviewed in the introduction (Bentwich, 2005; Lewis et al., 2003; Sethupathy et al., 2006)). In our hands, five out of the 10 miRNAs predicted to target *NTRK3* truncated isoform, were shown to be functional, according to the luciferase assay, rendering a higher false positive rate (50%). On the other hand, 7 out of the 23 selected predictions using TargetScan and miRanda for those miRNAs found to be associated with PD after analysis with the miRNA SNP panel were validated experimentally. Even though we are conscious that more analyses would be needed to demonstrate it, these results indicate that, performance of the three programs is lower than estimated. However, the rapid growth in the discovery of functional miRNA targets, as well as the constant creation of better algorithms will probably lead to the generation of proficient target prediction programs that will support the scientific community on "à la carte" analysis of miRNA dependent processes. In fact, in the process of writing these lines a few new promising alternatives that permit user defined searches have been implemented and are now available (Maragkakis et al., 2009). In conclusion, many efforts have been dedicated to the computational prediction and identification of miRNA target genes, and so it is reflected in the literature, where many articles on polymiRTS have been published. In contrast, little is known on how polymorphisms might be affecting pre-miRNA transcription, structure and mature miRNA expression and processing. This can be applied both at the functional and prediction level, where few programs are able to analyse the consequences, in terms of stability of the secondary RNA structure, of mutations affecting the pre-miRNA at those regions that will not form part of the mature miRNA itself. We,

ourselves, encountered this problem and after re-sequencing those miRNAs regions that were associated with PD we could only identify variants outside the pre-miRNA sequence and thus, we could not predict the consequences on either RNA stability or mature miRNA expression of the variants identified. The reason for the low amount of information concerning this subject is, obviously, the enormous complexity of the task. After all, despite the importance of ncRNAs is nowadays accepted, it was only a few years ago that they were mostly ignored.

Only few examples of mutations on miRNA genes or their regulatory regions have been exhaustively reported so far. The best documented study of this type involves a common SNP in pre-miR-146a that was reported to be strongly associated with papillary thyroid carcinoma. As this SNP is confined to the passenger strand (miRNA*), little credibility was given to the possibility that this SNP was responsible for the observed 19-fold increase in the quantity of miR-146 in patients. However, more recently, the authors demonstrated that both miR-146a passenger (miRNA*) and leading strand (miRNA) are expressed, and that heterozygotes express three different forms of this miRNA, with a distinct set of target genes each. In addition, tumours from heterozygote patients showed a widely different transcriptome (as assessed by microarray) and a 1.5- to 2.6 fold overexpression of polymorphic miR-146a (passenger strand) when compared with the unaffected part of the same gland (Jazdzewski et al., 2009). This study is one of the first studies that evidences that polymorphisms in miRNA coding regions can lead to disease and contrary to what is generally thought, it proposes that mature miRNAs from the passenger strand may, as well, regulate many genetic processes. This is particularly

interesting since the study uncovers the fact that miRNA processing and action is still to be deeply studied and that, as in every emerging topic of research, there might be false dogmas that should be redirected.

The lack of mutations identified at the mature miRNAs is in agreement with the reported negative selection acting at both miRNAs as well as miRNA target sites. Existence of negative selection in conserved miRNA target sites at 3'UTRs was described a couple of years ago. This work demonstrated a lower SNP density (0.5 SNPs/kb) in conserved miRNA target sites at 3'UTRs than in their flanking control sites (0.73 SNPs/kb) (Chen and Rajewsky, 2006). Likewise, it has been extensively proposed that SNP density is lower in miRNA loci with respect to their flanking regions (Saunders et al., 2007). Accordingly, when we analysed the SNP coverage on miRNAs we found a lower density of SNPs in miRNA regions than in the rest of the genome. In fact, we could only map 24 SNPs (dbSNP 125) within miRNAs sequences; this represents a density of 0.86 SNPs per kb at miRNA regions compared with the observed average of 3.99 SNPs/kb SNP for the whole human genome. However, screening of SNPs from public databases deals with the problem of SNP ascertainment bias, mainly due to underrepresentation of low-frequency variants and the fact that not all the genome has been equally characterised. Remarkably, very recently Quach *et al*. (2009) confirmed this lower SNP density by resequencing 117 miRNAs in four different human reference populations, therefore avoiding ascertainment bias coming from public databases. Their analyses reported a lower SNP density in miRNAs than in other non-coding regions, which were shown to be twice as dense. The study also showed that strong purifying selection affects the

sequence corresponding to the mature miRNA (particularly the first 14 nt, where no mutation is tolerated) as well as the complementary miRNA sequence (miRNA*), stem region and loop (Quach et al., 2009). In summary, these studies indicate that mutations in miRNA hairpins or in miRNA binding sites, such as the previously mentioned SNP occurring in miR-146a, are likely to be deleterious and may have severe phenotypic consequences on human health. Therefore, extensive re-sequencing in patients and controls of 3'UTRs and of miRNAs themselves, like we modestly attempted in our study would be definitely interesting to test the putative role of miRNA-mediated regulation in the susceptibility to anxiety disorders. Unfortunately, only 117 out of the actual 721 miRNAs could be resequenced in the study of Quach *et al.* (2009). Nevertheless, as it happened in our case, the fast increase in the number of new discovered miRNAs is being one of the main handicaps that researchers are facing. This is further complicated by the frequent corrections and modifications miRNAs suffer in their annotation (as far as sequences, nomenclature, etc). Remarkably, since the time that we started the project until now, the number of identified miRNAs has been doubled. The founding member of the miRNA family, lin-4, was identified in C. elegans in 1993 through a genetic screen for defects in the temporal control of post-embryonic development (Lee et al., 1993). Withal, it was not until 2001 that miRNAs were recognised as a large and phylogenetically extensive family of non-coding RNAs representing a new layer of gene regulation (Lagos-Quintana et al., 2001; Lau et al., 2001; Lee and Ambros, 2001). Since then, the number of identified miRNAs has increased vertiginously, for example in humans, their growth has ranged from none, in 2001, to 721 nowadays. Consequently,

publications and knowledge on this class of small RNAs has also increased considerably. As a matter of fact, the number of pubmed publications on miRNAs, in 2001 was of five, while nowadays there are roughly 6000.

The exponential increase on the number of miRNAs has also been a handicap for the design of the SNP panel that we used later for association analyses in disease, since the number of known miRNAs increased in a 55% from the beginning of the design to the moment we performed the association analyses. Apart from the analyses in disease, we also took advantage of the construction of the SNPs panel to study variability in miRNAs regions among different populations. After genotyping a group of 341 Spanish control individuals, allele frequencies between the HapMap European population and our specific North-East Spanish (Catalan) population were compared and pointed out to two genomic regions that show geographic genetic variation among populations. Remarkably, one of these regions is the LCT region (containing hsa-miR-128-1), a region that has already been shown to be under selective pressure (Beja-Pereira et al., 2003; Hollox, 2005). Lactase persistent phenotype is predominant in two different groups: central and northern Europeans (Scandinavian, German, northern French and British) and nomadic and milk-dependent populations from North and Central Africa and Arabia. These two groups are separated by a peri-mediterranean group of people (including our population) that are mainly non-lactase persistent (Kozlov et al., 1998). It is interesting to mention that seven SNPs tagging the hsa-miR-128-1 region were found to be associated with OCD in our population, however, since three of these SNPs showed different allelic frequencies between our and the Hapmap populations, we

should be cautious when facing the analysis of this particular region in order to avoid spurious associations. In addition, one of the rare allelic variants identified in the *NTRK3* study was disrupting a target site for miR-128. It would be intriguing to study if there are differences in the OCD prevalences among European populations and if, as it happens with malaria and sickle cell anemia, what produces a selective advantage on the one hand, like being lactase tolerant, is also a disadvantage and a susceptibility factor for OCD. Analysing the extension of the regions showing geographical differences as well as the ancestral alleles, in order to discern if the real cause for these differences is natural selection, could be a good step beyond.

Another interesting topic is that of miRNA editing. It has been reported that certain pri-miRNAs are subject to A→I editing, which is detected as an A→G change of the cDNA sequence (Blow et al., 2006). Misregulated RNA editing might be responsible for a number of human diseases or pathological processes, as it has been formerly described for the serotonin receptor 2C and schizophrenia, depression and suicide (Kawahara et al., 2008). Specifically, pri-miRNA editing has been predicted to affect about 16% of human pri-miRNAs (Kawahara et al., 2008) and it has been involved in altered miRNA processing and altered miRNA targeting spectrum (Kawahara et al., 2007; Yang et al., 2006). These results suggest that editing of pri-miRNAs may have a large impact on the processing and function of the different miRNAs and may thus be affecting gene expression at a genome scale. Taking these data into account and knowing that editing occurs in a tissue-specific manner (Funahashi et al., 1995) evidences the importance of extensive use of deep sequencing technologies to

analyse miRNA sequences in the different tissues, not only at the DNA but also at the RNA level. Importantly, new technologies are developing and very recently Helicos BioSciences, for instance, have claimed to be able to directly sequence single molecules of RNA, without prior conversion to cDNA; a capability that could have advantages for analysing short RNAs and for providing unbiased sampling of transcriptomes (Ozsolak et al., 2009). Indeed, it would have been extremely interesting to analyse the RNA sequence of those miRNAs that we found to be associated with PD and to study the putative implication of RNA editing in the susceptibility to disease. However, for blatant reasons, comparison of brain RNA between patients and controls is definitely out of the question. In general, invasive studies directly assessing brain tissue are not suited to the study of psychiatric illness. In addition, despite animal studies have attempted to bridge this gap, it is difficult to reach conclusions regarding human psychiatric disorders on the basis of animal observations. In conclusion, even though the importance of performing functional studies rather than replicating in other cohorts has been remarked lately for complex disorders (Burmeister et al., 2008), mapping genetic susceptibility to psychiatric disorders has proven to be experimentally challenging.

At present most genetic association studies on psychiatric disorders do not answer the question as to whether a disease-associated polymorphism is itself functionally important or only acting as a marker for a co-inherited, perhaps yet unidentified, genetic polymorphism. In our studies we tried to overcome this difficulty by performing functional studies based on luciferase assays in HeLa cells. Yet, we are conscious that luciferase

assays, albeit being a convenient and useful tool to rapidly screen many miRNAs, create an environment where the encounters between miRNA and target mRNA are extremely forced, as both are overexpressed, increasing the likelihood of false positives. In addition, HeLa cells, that have cervical cancer origin, might lack additional cofactors required for brain-specific miRNA mediated regulation. In this sense, we tried to analyse the functional consequences of the SNP associated to OCD hoarding subphentoype in a more biologically relevant context; we tried using SH-SY5Y neuroblastoma cell line. However, the high specificity of the expression of the truncated isoform of *NTRK3* made this approach complicated. Another functional approach that we used was transcriptome analysis after overexpression of specific miRNAs in neuroblastoma cells. One interesting application of this analysis is the identification of potential targets: by combining microarray results and target predictions it is in fact possible to identify genes that are repressed by the overexpression and are likely to contain functional target sites. Withal, this approach can be useful to discriminate targets regulated through mRNA cleavage rather than translational repression. Indeed, those targets inhibited through translational repression will not be detected through these microarray experiments. Interestingly, when we performed these experiments with miR-22 and miR-488, in both cases predicted targets that resulted to be deregulated, showed a clear predominance of downregulation over upregulation, supporting the hypothesis that those genes could be targeted by these miRNAs and evidencing that perhaps more genes than initially suggested are dowregulated through RNA degradation in animals. A clear example is that of *RGS2,* for which its expression was significantly reduced when we overexpressed

miR-22 in neuroblastoma cells. In spite of all, this method cannot substitute conventional target validation of each gene, in order to confirm that such repression is directly caused by the miRNAs and not by some secondary effect.

Finally, as far as functional analyses on psychiatric disorders are concerned, it is worth mentioning that data from genetics, cognitive function, and neuroimaging have begun to converge in an auspicious kind of studies. As an example, these studies examine the relation between well-known functional polymorphisms and different brain activation (through comparison of functional magnetic resonance imaging measured blood oxigenation) while the test individual is performing a cognitive task (Dager et al., 2008).

Anxiety disorders have long been believed to have abnormal neural regulatory mechanisms underlying symptom manifestation (Dager et al., 2008). We propose that this deregulation is mediated by defective miRNA action that consequently derives in dosage changes of proteins involved in CNS function. Our results manifest that this proposed deregulation could indeed be true, and that mutations in miRNA target sites, as well as anomalous dosage of miRNAs themselves could be responsible for a disruption in the fine-tuned equilibrium of complex brain functions, contributing to the development of anxiety disorders. To evaluate the possible involvement of miRNAs in disease, we have resequenced the 3'UTRs of *NTRK3*, conducted association and functional studies with the allelic variants found, and demonstrated that at least 2 of them, albeit being rare might be affecting miRNA regulation; we have constructed a panel of SNPs covering miRNA regions

and we have analysed these regions as anxiety disorder candidate genes by performing association studies in OCD and PD. Association with PD was found for 6 miRNAs, and we demonstrated that, at least four of them regulate genes that have been previously involved in anxiety disorders, namely, *NTRK3, RGS2, GABRA6, CCKBR, BDNF, HTR2C and MAOA*. Moreover, results on miR-22 or miR-488 overexpression strengthen the idea that their alteration may be affecting in several ways not only the fine functioning of the CNS, but also other physiological pathways linked to de development of anxiety disorders and particularly PD. All these evidences put together fit with the miRNA-induced differential dosage hypothesis that we propose to be involved in the aetiology of PD. Moreover, the associations identified for PD and OCD are in line with the current view that psychiatric disorders exhibit genetic, allelic and phenotypic heterogeneity. This complexity, together with the fact that psychiatric disorders are also influenced by environmental factors, is challenging and therefore, a joint action of scientists with knowledge in bioinformatics, statistics, neuroscience, genetics, proteomics, gene expression and animal models would be needed. Their results will definitely help to shape the definitions of anxiety disorders and dozens of PDs or OCDs will be described, just as there are now dozens of genetically defined forms of deafness or spinocerebellar ataxia. In conclusion, these genetic findings will probably help to the recognition of biological pathways affected in the aetiology of the different disorders, and may allow the classification of patients on the basis of the pathway involved in their aetiology. This classification, will probably lead to the development of new pathway-oriented therapies, that given the

genetic heterogeneity supposed for these disorders will be more effective than personalised medicine based on individual alleles.

Conclusions

Well, Art is Art, isn't it? Still, on the other hand, water is water. And east is east and west is west and if you take cranberries and stew them like applesauce they taste much more like prunes than rhubarb does. Now you tell me what you know.

Groucho Marx

1. Changes in regulatory elements that lead to minor variations in the dosage of proteins involved in neuronal pathways may contribute to the susceptibility to anxiety disorders.

- Two rare variants in the 3'UTR of the *NTRK3* truncated isoform have been identified in a patient with PD and agoraphobia each. These two variants altered the affinity and efficiency of the miRNA-related down-regulation of miR-765, miR-509 and miR-128 in comparison to the wild-type allele on *NTRK3* truncated isoform. These variants may be involved in the susceptibility to PD.

- Association analysis of miRNA regions in a Spanish population with PD and replication studies in two North European populations revealed association for SNPs tagging miRNAs –miR-22, miR-138-2, miR-148a, miR-339, and miR-488- with PD and other anxiety phenotypes. While no allelic variants could be identified within their pre-miRNA sequences, a putative effect on miRNA dosage of variants identified in the proximity of these sequences cannot be discarded.

- Association for a common SNP (rs28521337) in the miRNA target site for miR-485-3p was found in OCD hoarding subtype. This variant does not seem to affect the regulation of *NTRK3* by this miRNA. Contribution of this variant to disease remains unclear.

- Association analysis of miRNA regions in a Spanish population with OCD revealed nominal associations for SNPs tagging many miRNAs with OCD, various OCD subtypes and AAO. The finding of several specific associations (with different miRNA regions and with *NTRK3* truncated isoform) for the hoarding subclinical type suggests a different pattern of inheritance for this group of patients that could constitute a neurobiologically and etiologically distinct variant of OCD.

- The association analyses with anxiety disorders reported in this thesis point to multiple alleles (rare or common) that contribute independently to marginal increases of risk and, in some cases (rare alleles), with a low weight in the general population.

2. Luciferase-based assays performed in HeLa cells revealed a statistically significant reduction of the luciferase activity, compatible with a putative regulation by miRNAs, for several candidate genes for anxiety disorders.

-The truncated isoform of *NTRK3* is regulated by at least five miRNAs (miR-128, miR-485-3p, miR-509, miR-625 and miR-765). Two of these miRNAs (miR-509 and miR-625) caused a reduction in the luciferase activity of more than 50%.

- Four different anxiety candidate genes, *MAOA, BDNF, HTR2C* and *RGS2* are regulated by miR-22, a miRNA that has been genetically associated with PD, and that causes a conspicuous down-regulation of *RGS2* of around 66%.

- Three other anxiety candidate genes, *POMC*, *GABRA6* and CCKBR, are regulated by miR-488, miR-138-2 and miR-148a, respectively; three miRNAs that were found to be genetically associated with PD.

3. A comprehensive analysis of the genomic organisation of miRNAs regarding other transcriptional units as well as their localization in clusters has been performed. The genetic variability in these regions has also been analysed.

- Analysis of the genomic organisation of 325 human miRNAs (release 7.1, miRBase) showed that 40% of the miRNAs are isolated while the remaining 60% are arranged in 48 miRNA clusters. Moreover, this analysis also demonstrated that 37% of the miRNAs are inside known protein-coding genes.

- Analysis of the biological functions of the host genes containing miRNAs showed enrichment for genes involved in neurological disease, psychological disorders or nutritional disease.

- The SNP density in miRNA regions is lower than in intronic or intergenic regions. This lower SNP density agrees with the reported action of negative selection on miRNA regions.

- Allele frequencies comparison for the miRNA SNP panel between four HapMap populations and our specific North-East Spanish (Catalan) population pointed to two genomic regions that show geographic genetic variation among populations. One

of these regions is the *LCT* region (containing hsa-miR-128-1), which has already been shown to be under selective pressure.

4. Association analyses alone showed to be insufficient to identify variants underlying the susceptibility to anxiety disorders probably due to controversial diagnosis and genetic heterogeneity. In this regard, the inclusion of a more functional approach has proven to be useful in order to identify candidate biological pathways to be involved in the pathophysiology of the studied disorders.

Bibliography

Abelson, J. F., Kwan, K. Y., O'Roak, B. J., Baek, D. Y., Stillman, A. A., Morgan, T. M., Mathews, C. A., Pauls, D. L., Rasin, M. R., Gunel, M., Davis, N. R., Ercan-Sencicek, A. G., Guez, D. H., Spertus, J. A., Leckman, J. F., Dure, L. S. t., Kurlan, R., Singer, H. S., Gilbert, D. L., Farhi, A., Louvi, A., Lifton, R. P., Sestan, N., and State, M. W. (2005). *Science* **310,** 317-20.

Adams, B. D., Furneaux, H., and White, B. A. (2007). *Mol Endocrinol* **21,** 1132-47.

Ambros, V., Bartel, B., Bartel, D. P., Burge, C. B., Carrington, J. C., Chen, X., Dreyfuss, G., Eddy, S. R., Griffiths-Jones, S., Marshall, M., Matzke, M., Ruvkun, G., and Tuschl, T. (2003). *Rna* **9,** 277-9.

Arisawa, T., Tahara, T., Shibata, T., Nagasaka, M., Nakamura, M., Kamiya, Y., Fujita, H., Hasegawa, S., Takagi, T., Wang, F. Y., Hirata, I., and Nakano, H. (2007). *Dig Dis Sci* **52,** 1691-7.

Armengol, L., Gratacos, M., Pujana, M. A., Ribases, M., Martin-Santos, R., and Estivill, X. (2002). *Mol Psychiatry* **7,** 928-30.

Asangani, I. A., Rasheed, S. A., Nikolova, D. A., Leupold, J. H., Colburn, N. H., Post, S., and Allgayer, H. (2008). *Oncogene* **27,** 2128-36.

Association, A. P. (1994). "*Diagnostic and Statistical Manual of Mental Disorders*." American Psychiatric Press, Washington, DC.

Azzam, A., and Mathews, C. A. (2003). *Am J Med Genet B Neuropsychiatr Genet* **123B,** 64-9.

Barbacid, M., Lamballe, F., Pulido, D., and Klein, R. (1991). *Biochim Biophys Acta* **1072,** 115-27.

Barlow, D. (2004). "*Anxiety and its disorders: The nature and treatment of anxiety and panic*." Guilford Press, New York.

Barr, L. C., Goodman, W. K., and Price, L. H. (1993). *Int Clin Psychopharmacol* **8 Suppl 2,** 79-82.

Bartel, D. P. (2004). *Cell* **116,** 281-97.

Beetz, C., Schule, R., Deconinck, T., Tran-Viet, K. N., Zhu, H., Kremer, B. P., Frints, S. G., van Zelst-Stams, W. A., Byrne, P., Otto, S., Nygren, A. O., Baets, J., Smets, K., Ceulemans, B., Dan, B., Nagan, N., Kassubek, J., Klimpe, S., Klopstock, T., Stolze, H., Smeets, H. J., Schrander-Stumpel, C. T., Hutchinson, M., van de Warrenburg, B. P., Braastad, C., Deufel, T., Pericak-Vance, M., Schols, L., de Jonghe, P., and Zuchner, S. (2008). *Brain* **131,** 1078-86.

Beja-Pereira, A., Luikart, G., England, P. R., Bradley, D. G., Jann, O. C., Bertorelle, G., Chamberlain, A. T., Nunes, T. P., Metodiev, S., Ferrand, N., and Erhardt, G. (2003). *Nat Genet* **35,** 311-3.

Bell, C. J., and Nutt, D. J. (1998). *Br J Psychiatry* **172,** 465-71.

Beltaifa, S., Webster, M. J., Ligons, D. L., Fatula, R. J., Herman, M. M., Kleinman, J. E., and Weickert, C. S. (2005). *Eur J Neurosci* **21,** 2433-44.

Bentwich, I. (2005). *FEBS Lett* **579,** 5904-10.

Bernstein, E., Kim, S. Y., Carmell, M. A., Murchison, E. P., Alcorn, H., Li, M. Z., Mills, A. A., Elledge, S. J., Anderson, K. V., and Hannon, G. J. (2003). *Nat Genet* **35,** 215-7.

Bespalova, I. N., and Buxbaum, J. D. (2003). *Ann Med* **35,** 274-81.

Bhattacharyya, S. N., Habermacher, R., Martine, U., Closs, E. I., and Filipowicz, W. (2006). *Cold Spring Harb Symp Quant Biol* **71,** 513-21.

Blow, M. J., Grocock, R. J., van Dongen, S., Enright, A. J., Dicks, E., Futreal, P. A., Wooster, R., and Stratton, M. R. (2006). *Genome Biol* **7,** R27.

Blum, R., and Konnerth, A. (2005). *Physiology (Bethesda)* **20,** 70-8.

Boulenger, J. P., Uhde, T. W., Wolff, E. A., 3rd, and Post, R. M. (1984). *Arch Gen Psychiatry* **41,** 1067-71.

Boutz, P. L., Chawla, G., Stoilov, P., and Black, D. L. (2007). *Genes Dev* **21,** 71-84.

Bradwejn, J., Koszycki, D., and Shriqui, C. (1991). *Arch Gen Psychiatry* **48,** 603-10.

Bremner, J. D., Innis, R. B., White, T., Fujita, M., Silbersweig, D., Goddard, A. W., Staib, L., Stern, E., Cappiello, A., Woods, S., Baldwin, R., and Charney, D. S. (2000). *Biol Psychiatry* **47,** 96-106.

Brennecke, J., Stark, A., Russell, R. B., and Cohen, S. M. (2005). *PLoS Biol* **3,** e85.

Brodersen, P., and Voinnet, O. (2009). *Nat Rev Mol Cell Biol* **10,** 141-8.

Buckland, P. R., Hoogendoorn, B., Coleman, S. L., Guy, C. A., Smith, S. K., and O'Donovan, M. C. (2005). *Hum Mutat* **26,** 214-23.

Burmeister, M., McInnis, M. G., and Zollner, S. (2008). *Nat Rev Genet* **9,** 527-40.

Cai, X., Hagedorn, C. H., and Cullen, B. R. (2004). *Rna* **10,** 1957-66.

Calin, G. A., Dumitru, C. D., Shimizu, M., Bichi, R., Zupo, S., Noch, E., Aldler, H., Rattan, S., Keating, M., Rai, K., Rassenti, L., Kipps, T., Negrini, M., Bullrich, F., and Croce, C. M. (2002). *Proc Natl Acad Sci U S A* **99,** 15524-9.

Calin, G. A., Sevignani, C., Dumitru, C. D., Hyslop, T., Noch, E., Yendamuri, S., Shimizu, M., Rattan, S., Bullrich, F., Negrini, M., and Croce, C. M. (2004). *Proc Natl Acad Sci U S A* **101,** 2999-3004.

Cao, X., Yeo, G., Muotri, A. R., Kuwabara, T., and Gage, F. H. (2006). *Annu Rev Neurosci* **29,** 77-103.

Carleton, M., Cleary, M. A., and Linsley, P. S. (2007). *Cell Cycle* **6,** 2127-32.

Carthew, R. W., and Sontheimer, E. J. (2009). *Cell* **136,** 642-55.

Caspi, A., Sugden, K., Moffitt, T. E., Taylor, A., Craig, I. W., Harrington, H., McClay, J., Mill, J., Martin, J., Braithwaite, A., and Poulton, R. (2003). *Science* **301,** 386-9.

Chapman, E. J., and Carrington, J. C. (2007). *Nat Rev Genet* **8,** 884-96.

Chen, C. Z., Li, L., Lodish, H. F., and Bartel, D. P. (2004). *Science* **303,** 83-6.

Chen, K., and Rajewsky, N. (2006). *Nat Genet* **38,** 1452-6.

Chendrimada, T. P., Finn, K. J., Ji, X., Baillat, D., Gregory, R. I., Liebhaber, S. A., Pasquinelli, A. E., and Shiekhattar, R. (2007). *Nature* **447,** 823-8.

Chumakov, I., Blumenfeld, M., Guerassimenko, O., Cavarec, L., Palicio, M., Abderrahim, H., Bougueleret, L., Barry, C., Tanaka, H., La Rosa, P., Puech, A., Tahri, N., Cohen-Akenine, A., Delabrosse,

S., Lissarrague, S., Picard, F. P., Maurice, K., Essioux, L., Millasseau, P., Grel, P., Debailleul, V., Simon, A. M., Caterina, D., Dufaure, I., Malekzadeh, K., Belova, M., Luan, J. J., Bouillot, M., Sambucy, J. L., Primas, G., Saumier, M., Boubkiri, N., Martin-Saumier, S., Nasroune, M., Peixoto, H., Delaye, A., Pinchot, V., Bastucci, M., Guillou, S., Chevillon, M., Sainz-Fuertes, R., Meguenni, S., Aurich-Costa, J., Cherif, D., Gimalac, A., Van Duijn, C., Gauvreau, D., Ouellette, G., Fortier, I., Raelson, J., Sherbatich, T., Riazanskaia, N., Rogaev, E., Raeymaekers, P., Aerssens, J., Konings, F., Luyten, W., Macciardi, F., Sham, P. C., Straub, R. E., Weinberger, D. R., Cohen, N., and Cohen, D. (2002). *Proc Natl Acad Sci U S A* **99,** 13675-80.

Cimmino, A., Calin, G. A., Fabbri, M., Iorio, M. V., Ferracin, M., Shimizu, M., Wojcik, S. E., Aqeilan, R. I., Zupo, S., Dono, M., Rassenti, L., Alder, H., Volinia, S., Liu, C. G., Kipps, T. J., Negrini, M., and Croce, C. M. (2005). *Proc Natl Acad Sci U S A* **102,** 13944-9.

Clop, A., Marcq, F., Takeda, H., Pirottin, D., Tordoir, X., Bibe, B., Bouix, J., Caiment, F., Elsen, J. M., Eychenne, F., Larzul, C., Laville, E., Meish, F., Milenkovic, D., Tobin, J., Charlier, C., and Georges, M. (2006). *Nat Genet* **38,** 813-8.

Conaco, C., Otto, S., Han, J. J., and Mandel, G. (2006). *Proc Natl Acad Sci U S A* **103,** 2422-7.

Consortium, T. W. T. C. C. (2007). *Nature* **447,** 661-78.

Dager, S. R., Corrigan, N. M., Richards, T. L., and Posse, S. (2008). *Top Magn Reson Imaging* **19,** 81-96.

Dalmay, T. (2008). *J Intern Med* **263,** 366-75.

David, S. P., Murthy, N. V., Rabiner, E. A., Munafo, M. R., Johnstone, E. C., Jacob, R., Walton, R. T., and Grasby, P. M. (2005). *J Neurosci* **25,** 2586-90.

Deckert, J., Catalano, M., Heils, A., Di Bella, D., Friess, F., Politi, E., Franke, P., Nothen, M. M., Maier, W., Bellodi, L., and Lesch, K. P. (1997). *Psychiatr Genet* **7,** 45-7.

Deckert, J., Catalano, M., Syagailo, Y. V., Bosi, M., Okladnova, O., Di Bella, D., Nothen, M. M., Maffei, P., Franke, P., Fritze, J., Maier, W., Propping, P., Beckmann, H., Bellodi, L., and Lesch, K. P. (1999). *Hum Mol Genet* **8,** 621-4.

Deckert, J., Nothen, M. M., Franke, P., Delmo, C., Fritze, J., Knapp, M., Maier, W., Beckmann, H., and Propping, P. (1998). *Mol Psychiatry* **3,** 81-5.

Delgado, P. L., and Moreno, F. A. (1998). *Br J Psychiatry Suppl***,** 21-5.

Deng, H., Le, W. D., Xie, W. J., and Jankovic, J. (2006). *Acta Neurol Scand* **114,** 400-2.

Domschke, K., Braun, M., Ohrmann, P., Suslow, T., Kugel, H., Bauer, J., Hohoff, C., Kersting, A., Engelien, A., Arolt, V., Heindel, W., and Deckert, J. (2006). *Int J Neuropsychopharmacol* **9,** 349-55.

Donner, J., Pirkola, S., Silander, K., Kananen, L., Terwilliger, J. D., Lonnqvist, J., Peltonen, L., and Hovatta, I. (2008). *Biol Psychiatry* **64,** 672-80.

Dorland, W. A. N. (1994). "*Dorland's Illustrated Medical Dictionary*." W.B. Saunders Company, Edinburgh.

Duman, R. S. (2002). *Mol Psychiatry* **7 Suppl 1,** S29-34.

Edenberg, H. J., Dick, D. M., Xuei, X., Tian, H., Almasy, L., Bauer, L. O., Crowe, R. R., Goate, A., Hesselbrock, V., Jones, K., Kwon, J., Li, T. K., Nurnberger, J. I., Jr., O'Connor, S. J., Reich, T., Rice, J., Schuckit, M. A., Porjesz, B., Foroud, T., and Begleiter, H. (2004). *Am J Hum Genet* **74,** 705-14.

Ehrig, T., Bosron, W. F., and Li, T. K. (1990). *Alcohol Alcohol* **25,** 105-16.

Eletto, D., Russo, G., Passiatore, G., Del Valle, L., Giordano, A., Khalili, K., Gualco, E., and Peruzzi, F. (2008). *J Cell Physiol* **216,** 764-770.

Erson, A. E., and Petty, E. M. (2008). *Clin Genet* **74,** 296-306.

Esteban, P. F., Yoon, H. Y., Becker, J., Dorsey, S. G., Caprari, P., Palko, M. E., Coppola, V., Saragovi, H. U., Randazzo, P. A., and Tessarollo, L. (2006). *J Cell Biol* **173,** 291-9.

Fabbrini, G., Pasquini, M., Aurilia, C., Berardelli, I., Breedveld, G., Oostra, B. A., Bonifati, V., and Berardelli, A. (2007). *Mov Disord* **22,** 2229-34.

Fehr, C., Szegedi, A., Anghelescu, I., Klawe, C., Hiemke, C., and Dahmen, N. (2000). *Psychiatr Genet* **10,** 59-65.

Filipowicz, W., Bhattacharyya, S. N., and Sonenberg, N. (2008). *Nat Rev Genet* **9,** 102-14.

Fiore, R., and Schratt, G. (2007). *ScientificWorldJournal* **7,** 167-77.

Frazer, K. A., Ballinger, D. G., Cox, D. R., Hinds, D. A., Stuve, L. L., Gibbs, R. A., Belmont, J. W., Boudreau, A., Hardenbol, P., Leal, S. M., Pasternak, S., Wheeler, D. A., Willis, T. D., Yu, F., Yang, H., Zeng, C., Gao, Y., Hu, H., Hu, W., Li, C., Lin, W., Liu, S., Pan, H., Tang, X., Wang, J., Wang, W., Yu, J., Zhang, B., Zhang, Q., Zhao, H., Zhao, H., Zhou, J., Gabriel, S. B., Barry, R., Blumenstiel, B., Camargo, A., Defelice, M., Faggart, M., Goyette, M., Gupta, S., Moore, J., Nguyen, H., Onofrio, R. C., Parkin, M., Roy, J., Stahl, E., Winchester, E., Ziaugra, L., Altshuler, D., Shen, Y., Yao, Z., Huang, W., Chu, X., He, Y., Jin, L., Liu, Y., Shen, Y., Sun, W., Wang, H., Wang, Y., Wang, Y., Xiong, X., Xu, L., Waye, M. M., Tsui, S. K., Xue, H., Wong, J. T., Galver, L. M., Fan, J. B., Gunderson, K., Murray, S. S., Oliphant, A. R., Chee, M. S., Montpetit, A., Chagnon, F., Ferretti, V., Leboeuf, M., Olivier, J. F., Phillips, M. S., Roumy, S., Sallee, C., Verner, A., Hudson, T. J., Kwok, P. Y., Cai, D., Koboldt, D. C., Miller, R. D., Pawlikowska, L., Taillon-Miller, P., Xiao, M., Tsui, L. C., Mak, W., Song, Y. Q., Tam, P. K., Nakamura, Y., Kawaguchi, T., Kitamoto, T., Morizono, T., Nagashima, A., Ohnishi, Y., et al. (2007). *Nature* **449,** 851-61.

Funahashi, T., Giannoni, F., DePaoli, A. M., Skarosi, S. F., and Davidson, N. O. (1995). *J Lipid Res* **36,** 414-28.

Gelder MG, L.-I. J., Andreasen (2001). "*New Oxford Textbook of Psychiatry*." Oxford University Press, Oxford.

Giraldez, A. J., Cinalli, R. M., Glasner, M. E., Enright, A. J., Thomson, J. M., Baskerville, S., Hammond, S. M., Bartel, D. P., and Schier, A. F. (2005). *Science* **308,** 833-8.

Goldstein, R. B., Wickramaratne, P. J., Horwath, E., and Weissman, M. M. (1997). *Arch Gen Psychiatry* **54,** 271-8.

Gong, H., Liu, C. M., Liu, D. P., and Liang, C. C. (2005). *Med Res Rev* **25,** 361-81.

Gorman, J. M., Kent, J. M., Sullivan, G. M., and Coplan, J. D. (2000). *Am J Psychiatry* **157,** 493-505.

Graeff, F. G., Silveira, M. C., Nogueira, R. L., Audi, E. A., and Oliveira, R. M. (1993). *Behav Brain Res* **58,** 123-31.

Gratacos, M., Nadal, M., Martin-Santos, R., Pujana, M. A., Gago, J., Peral, B., Armengol, L., Ponsa, I., Miro, R., Bulbena, A., and Estivill, X. (2001). *Cell* **106,** 367-79.

Gratacos, M., Sahun, I., Gallego, X., Amador-Arjona, A., Estivill, X., and Dierssen, M. (2007). *Genes Brain Behav* **6 Suppl 1,** 2-23.

Griffiths-Jones, S. (2006). *Methods Mol Biol* **342,** 129-38.

Grisham, J. R., Anderson, T. M., and Sachdev, P. S. (2008). *Eur Arch Psychiatry Clin Neurosci* **258,** 107-16.

Gupta, A. R., and State, M. W. (2007). *Biol Psychiatry* **61,** 429-37.

Hamilton, S. P., Slager, S. L., De Leon, A. B., Heiman, G. A., Klein, D. F., Hodge, S. E., Weissman, M. M., Fyer, A. J., and Knowles, J. A. (2004). *Neuropsychopharmacology* **29,** 558-65.

Hamilton, S. P., Slager, S. L., Heiman, G. A., Deng, Z., Haghighi, F., Klein, D. F., Hodge, S. E., Weissman, M. M., Fyer, A. J., and Knowles, J. A. (2002). *Biol Psychiatry* **51,** 591-601.

Hamilton, S. P., Slager, S. L., Helleby, L., Heiman, G. A., Klein, D. F., Hodge, S. E., Weissman, M. M., Fyer, A. J., and Knowles, J. A. (2001). *Mol Psychiatry* **6,** 59-65.

Hanna, G. L., Veenstra-VanderWeele, J., Cox, N. J., Boehnke, M., Himle, J. A., Curtis, G. C., Leventhal, B. L., and Cook, E. H., Jr. (2002). *Am J Med Genet* **114,** 541-52.

Hansen, T., Olsen, L., Lindow, M., Jakobsen, K. D., Ullum, H., Jonsson, E., Andreassen, O. A., Djurovic, S., Melle, I., Agartz, I., Hall, H., Timm, S., Wang, A. G., and Werge, T. (2007). *PLoS ONE* **2,** e873.

Hansen, T. V., Rehfeld, J. F., and Nielsen, F. C. (2000). *Mol Psychiatry* **5,** 443-7.

Hattori, E., Liu, C., Badner, J. A., Bonner, T. I., Christian, S. L., Maheshwari, M., Detera-Wadleigh, S. D., Gibbs, R. A., and Gershon, E. S. (2003). *Am J Hum Genet* **72,** 1131-40.

Hayashita, Y., Osada, H., Tatematsu, Y., Yamada, H., Yanagisawa, K., Tomida, S., Yatabe, Y., Kawahara, K., Sekido, Y., and Takahashi, T. (2005). *Cancer Res* **65,** 9628-32.

He, L., and Hannon, G. J. (2004). *Nat Rev Genet* **5,** 522-31.

He, L., He, X., Lim, L. P., de Stanchina, E., Xuan, Z., Liang, Y., Xue, W., Zender, L., Magnus, J., Ridzon, D., Jackson, A. L., Linsley, P. S., Chen, C., Lowe, S. W., Cleary, M. A., and Hannon, G. J. (2007). *Nature* **447,** 1130-4.

Hebert, S. S., Horre, K., Nicolai, L., Papadopoulou, A. S., Mandemakers, W., Silahtaroglu, A. N., Kauppinen, S., Delacourte, A., and De Strooper, B. (2008). *Proc Natl Acad Sci U S A* **105,** 6415-20.

Hettema, J. M., An, S. S., van den Oord, E. J., Neale, M. C., Kendler, K. S., and Chen, X. (2008). *Am J Med Genet B Neuropsychiatr Genet* **147B,** 661-6.

Hettema, J. M., Neale, M. C., and Kendler, K. S. (2001). *Am J Psychiatry* **158,** 1568-78.

Hillebrand, J., Barbee, S. A., and Ramaswami, M. (2007). *ScientificWorldJournal* **7,** 178-90.

Hinds, D. A., Stuve, L. L., Nilsen, G. B., Halperin, E., Eskin, E., Ballinger, D. G., Frazer, K. A., and Cox, D. R. (2005). *Science* **307,** 1072-9.

Ho, L. W., Furlong, R. A., Rubinsztein, J. S., Walsh, C., Paykel, E. S., and Rubinsztein, D. C. (2000). *Am J Med Genet* **96,** 36-42.

Hohoff, C., Neumann, A., Domschke, K., Jacob, C., Maier, W., Fritze, J., Bandelow, B., Krakowitzky, P., Rothermundt, M., Arolt, V., and Deckert, J. (2008). *J Neural Transm*.

Hollox, E. (2005). *Eur J Hum Genet* **13,** 267-9.

Hosing, V. G., Schirmacher, A., Kuhlenbaumer, G., Freitag, C., Sand, P., Schlesiger, C., Jacob, C., Fritze, J., Franke, P., Rietschel, M., Garritsen, H., Nothen, M. M., Fimmers, R., Stogbauer, F., and Deckert, J. (2004). *J Neural Transm Suppl***,** 147-56.

Huang, E. J., and Reichardt, L. F. (2001). *Annu Rev Neurosci* **24,** 677-736.

Huang, E. J., and Reichardt, L. F. (2003). *Annu Rev Biochem* **72,** 609-42.

Huang, H. D., Horng, J. T., Lin, F. M., Chang, Y. C., and Huang, C. C. (2005). *Nucleic Acids Res* **33,** D80-5.

Huang, Y. Y., Battistuzzi, C., Oquendo, M. A., Harkavy-Friedman, J., Greenhill, L., Zalsman, G., Brodsky, B., Arango, V., Brent, D. A., and Mann, J. J. (2004). *Int J Neuropsychopharmacol* **7,** 441-51.

Hudder, A., and Novak, R. F. (2008). *Toxicol Sci* **103,** 228-40.

Hwang, J. P., Tsai, S. J., Hong, C. J., Yang, C. H., Lirng, J. F., and Yang, Y. M. (2006). *Neurobiol Aging* **27,** 1834-7.

Ibanez-Ventoso, C., Vora, M., and Driscoll, M. (2008). *PLoS ONE* **3,** e2818.

Ichinose, T., and Snider, W. D. (2000). *J Neurosci Res* **59,** 365-71.

Inada, Y., Yoneda, H., Koh, J., Sakai, J., Himei, A., Kinoshita, Y., Akabame, K., Hiraoka, Y., and Sakai, T. (2003). *Psychiatry Res* **118,** 25-31.

Ise, K., Akiyoshi, J., Horinouchi, Y., Tsutsumi, T., Isogawa, K., and Nagayama, H. (2003). *Am J Med Genet B Neuropsychiatr Genet* **118B,** 29-31.

Ishiguro, H., Arinami, T., Yamada, K., Otsuka, Y., Toru, M., and Shibuya, H. (1997). *Psychiatry Clin Neurosci* **51,** 333-5.

Jazdzewski, K., Liyanarachchi, S., Swierniak, M., Pachucki, J., Ringel, M. D., Jarzab, B., and de la Chapelle, A. (2009). *Proc Natl Acad Sci U S A* **106,** 1502-5.

Jazdzewski, K., Murray, E. L., Franssila, K., Jarzab, B., Schoenberg, D. R., and de la Chapelle, A. (2008). *Proc Natl Acad Sci U S A* **105,** 7269-74.

Jensen, K. P., Covault, J., Conner, T. S., Tennen, H., Kranzler, H. R., and Furneaux, H. M. (2008). *Mol Psychiatry*.

Jiang, X., Xu, K., Hoberman, J., Tian, F., Marko, A. J., Waheed, J. F., Harris, C. R., Marini, A. M., Enoch, M. A., and Lipsky, R. H. (2005). *Neuropsychopharmacology* **30,** 1353-61.

Jin, P., Zarnescu, D. C., Ceman, S., Nakamoto, M., Mowrey, J., Jongens, T. A., Nelson, D. L., Moses, K., and Warren, S. T. (2004). *Nat Neurosci* **7,** 113-7.

Johnson, S. M., Grosshans, H., Shingara, J., Byrom, M., Jarvis, R., Cheng, A., Labourier, E., Reinert, K. L., Brown, D., and Slack, F. J. (2005). *Cell* **120,** 635-47.

Johnston, R. J., and Hobert, O. (2003). *Nature* **426,** 845-9.

Kawahara, Y., Megraw, M., Kreider, E., Iizasa, H., Valente, L., Hatzigeorgiou, A. G., and Nishikura, K. (2008). *Nucleic Acids Res* **36,** 5270-80.

Kawahara, Y., Zinshteyn, B., Chendrimada, T. P., Shiekhattar, R., and Nishikura, K. (2007a). *EMBO Rep* **8,** 763-9.

Kawahara, Y., Zinshteyn, B., Sethupathy, P., Iizasa, H., Hatzigeorgiou, A. G., and Nishikura, K. (2007b). *Science* **315,** 1137-40.

Kawaji, H., and Hayashizaki, Y. (2008). *PLoS Genet* **4,** e22.

Keck, M. E., Kern, N., Erhardt, A., Unschuld, P. G., Ising, M., Salyakina, D., Muller, M. B., Knorr, C. C., Lieb, R., Hohoff, C., Krakowitzky, P., Maier, W., Bandelow, B., Fritze, J., Deckert, J., Holsboer, F., Muller-Myhsok, B., and Binder, E. B. (2008). *Am J Med Genet B Neuropsychiatr Genet* **147B,** 1196-204.

Keen-Kim, D., Mathews, C. A., Reus, V. I., Lowe, T. L., Herrera, L. D., Budman, C. L., Gross-Tsur, V., Pulver, A. E., Bruun, R. D., Erenberg, G., Naarden, A., Sabatti, C., and Freimer, N. B. (2006). *Hum Mol Genet* **15,** 3324-8.

Kendler, K. S., and Prescott, C. (2006). "*Genes, Environment, and Psychopathology*." Guilford Press, New York and London.

Kennedy, J. L., Bradwejn, J., Koszycki, D., King, N., Crowe, R., Vincent, J., and Fourie, O. (1999). *Mol Psychiatry* **4,** 284-5.

Kessler, R. C., Berglund, P., Demler, O., Jin, R., Merikangas, K. R., and Walters, E. E. (2005). *Arch Gen Psychiatry* **62,** 593-602.

Kieseppa, T., Partonen, T., Haukka, J., Kaprio, J., and Lonnqvist, J. (2004). *Am J Psychiatry* **161,** 1814-21.

Kim, J., Inoue, K., Ishii, J., Vanti, W. B., Voronov, S. V., Murchison, E., Hannon, G., and Abeliovich, A. (2007). *Science* **317,** 1220-4.

Kim, V. N. (2005). *Nat Rev Mol Cell Biol* **6,** 376-85.

Kim, V. N., Han, J., and Siomi, M. C. (2009). *Nat Rev Mol Cell Biol* **10,** 126-39.

Kim, V. N., and Nam, J. W. (2006). *Trends Genet* **22,** 165-73.

Kim, Y. K., and Kim, V. N. (2007). *Embo J* **26,** 775-83.

Kiriakidou, M., Tan, G. S., Lamprinaki, S., De Planell-Saguer, M., Nelson, P. T., and Mourelatos, Z. (2007). *Cell* **129,** 1141-51.

Klattenhoff, C., and Theurkauf, W. (2008). *Development* **135,** 3-9.

Knight, J. C. (2005). *J Mol Med* **83,** 97-109.

Koenen, K. C., Amstadter, A. B., Ruggiero, K. J., Acierno, R., Galea, S., Kilpatrick, D. G., and Gelernter, J. (2009). *Depress Anxiety* **26,** 309-15.

Kosik, K. S. (2006). *Nat Rev Neurosci* **7,** 911-20.

Kozlov, A. I., Balanovskaia, E. V., Nurbaev, S. D., and Balanovskii, O. P. (1998). *Genetika* **34,** 551-61.

Krichevsky, A. M., Sonntag, K. C., Isacson, O., and Kosik, K. S. (2006). *Stem Cells* **24,** 857-64.

Lachman, H. M., Papolos, D. F., Saito, T., Yu, Y. M., Szumlanski, C. L., and Weinshilboum, R. M. (1996). *Pharmacogenetics* **6,** 243-50.

Lagos-Quintana, M., Rauhut, R., Lendeckel, W., and Tuschl, T. (2001). *Science* **294,** 853-8.

Lander, E. S., Linton, L. M., Birren, B., Nusbaum, C., Zody, M. C., Baldwin, J., Devon, K., Dewar, K., Doyle, M., FitzHugh, W., Funke, R., Gage, D., Harris, K., Heaford, A., Howland, J., Kann, L., Lehoczky, J., LeVine, R., McEwan, P., McKernan, K., Meldrim, J., Mesirov, J. P., Miranda, C., Morris, W., Naylor, J., Raymond, C., Rosetti, M., Santos, R., Sheridan, A., Sougnez, C., Stange-Thomann, N., Stojanovic, N., Subramanian, A., Wyman, D., Rogers, J., Sulston, J., Ainscough, R., Beck, S., Bentley, D., Burton, J., Clee, C., Carter, N., Coulson, A., Deadman, R., Deloukas, P., Dunham, A., Dunham, I., Durbin, R., French, L., Grafham, D., Gregory, S., Hubbard, T., Humphray, S., Hunt, A., Jones, M., Lloyd, C., McMurray, A., Matthews, L., Mercer, S., Milne, S., Mullikin, J. C., Mungall, A., Plumb, R., Ross, M., Shownkeen, R., Sims, S., Waterston, R. H., Wilson, R. K., Hillier, L. W., McPherson, J. D., Marra, M. A., Mardis, E. R., Fulton, L. A., Chinwalla, A. T., Pepin, K. H., Gish, W. R., Chissoe, S. L., Wendl, M. C., Delehaunty, K. D., Miner, T. L., Delehaunty, A., Kramer, J. B., Cook, L. L., Fulton, R. S., Johnson, D. L., Minx, P. J., Clifton, S. W., Hawkins, T., Branscomb, E., Predki, P., Richardson, P., Wenning, S., Slezak, T., Doggett, N., Cheng, J. F., Olsen, A., Lucas, S., Elkin, C., Uberbacher, E., Frazier, M., et al. (2001). *Nature* **409,** 860-921.

Landi, D., Gemignani, F., Naccarati, A., Pardini, B., Vodicka, P., Vodickova, L., Novotny, J., Forsti, A., Hemminki, K., Canzian, F., and Landi, S. (2008). *Carcinogenesis* **29,** 579-84.

Landthaler, M., Yalcin, A., and Tuschl, T. (2004). *Curr Biol* **14,** 2162-7.

Laneve, P., Di Marcotullio, L., Gioia, U., Fiori, M. E., Ferretti, E., Gulino, A., Bozzoni, I., and Caffarelli, E. (2007). *Proc Natl Acad Sci U S A* **104,** 7957-62.

Lang, U. E., Sander, T., Lohoff, F. W., Hellweg, R., Bajbouj, M., Winterer, G., and Gallinat, J. (2007). *Psychopharmacology (Berl)* **190,** 433-9.

Lanza, G., Ferracin, M., Gafa, R., Veronese, A., Spizzo, R., Pichiorri, F., Liu, C. G., Calin, G. A., Croce, C. M., and Negrini, M. (2007). *Mol Cancer* **6,** 54.

Lau, N. C., Lim, L. P., Weinstein, E. G., and Bartel, D. P. (2001). *Science* **294,** 858-62.

Lee, R. C., and Ambros, V. (2001). *Science* **294,** 862-4.

Lee, R. C., Feinbaum, R. L., and Ambros, V. (1993). *Cell* **75,** 843-54.
Lee, Y. J., Hohoff, C., Domschke, K., Sand, P., Kuhlenbaumer, G., Schirmacher, A., Freitag, C. M., Meyer, J., Stober, G., Franke, P., Nothen, M. M., Fritze, J., Fimmers, R., Garritsen, H. S., Stogbauer, F., and Deckert, J. (2005). *Neurosci Lett* **377,** 40-3.
Lemonde, S., Turecki, G., Bakish, D., Du, L., Hrdina, P. D., Bown, C. D., Sequeira, A., Kushwaha, N., Morris, S. J., Basak, A., Ou, X. M., and Albert, P. R. (2003). *J Neurosci* **23,** 8788-99.
Lesch, K. P., Bengel, D., Heils, A., Sabol, S. Z., Greenberg, B. D., Petri, S., Benjamin, J., Muller, C. R., Hamer, D. H., and Murphy, D. L. (1996). *Science* **274,** 1527-31.
Lessmann, V., Gottmann, K., and Malcangio, M. (2003). *Prog Neurobiol* **69,** 341-74.
Lewis, B. P., Shih, I. H., Jones-Rhoades, M. W., Bartel, D. P., and Burge, C. B. (2003). *Cell* **115,** 787-98.
Leygraf, A., Hohoff, C., Freitag, C., Willis-Owen, S. A., Krakowitzky, P., Fritze, J., Franke, P., Bandelow, B., Fimmers, R., Flint, J., and Deckert, J. (2006). *J Neural Transm* **113,** 1921-5.
Li, X. F., Stutzmann, G. E., and LeDoux, J. E. (1996). *Learn Mem* **3,** 229-42.
Liebl, D. J., Tessarollo, L., Palko, M. E., and Parada, L. F. (1997). *J Neurosci* **17,** 9113-21.
Lim, L. P., Lau, N. C., Garrett-Engele, P., Grimson, A., Schelter, J. M., Castle, J., Bartel, D. P., Linsley, P. S., and Johnson, J. M. (2005). *Nature* **433,** 769-73.
Locatelli, I., Lichtenstein, P., and Yashin, A. I. (2004). *Twin Res* **7,** 182-91.
Lu, J., Getz, G., Miska, E. A., Alvarez-Saavedra, E., Lamb, J., Peck, D., Sweet-Cordero, A., Ebert, B. L., Mak, R. H., Ferrando, A. A., Downing, J. R., Jacks, T., Horvitz, H. R., and Golub, T. R. (2005). *Nature* **435,** 834-8.
Lukiw, W. J. (2007). *Neuroreport* **18,** 297-300.
Lytle, J. R., Yario, T. A., and Steitz, J. A. (2007). *Proc Natl Acad Sci U S A* **104,** 9667-72.
Mahony, S., Corcoran, D. L., Feingold, E., and Benos, P. V. (2007). *Genome Biol* **8,** R84.
Makeyev, E. V., Zhang, J., Carrasco, M. A., and Maniatis, T. (2007). *Mol Cell* **27,** 435-48.
Maragkakis, M., Alexiou, P., Papadopoulos, G. L., Reczko, M., Dalamagas, T., Giannopoulos, G., Goumas, G., Koukis, E., Kourtis, K., Simossis, V. A., Sethupathy, P., Vergoulis, T., Koziris, N., Sellis, T., Tsanakas, P., and Hatzigeorgiou, A. G. (2009). *BMC Bioinformatics* **10,** 295.
Maron, E., Nikopensius, T., Koks, S., Altmae, S., Heinaste, E., Vabrit, K., Tammekivi, V., Hallast, P., Koido, K., Kurg, A., Metspalu, A., Vasar, E., Vasar, V., and Shlik, J. (2005). *Psychiatr Genet* **15,** 17-24.
Maron, E., and Shlik, J. (2006). *Neuropsychopharmacology* **31,** 1-11.
Martin, M. M., Buckenberger, J. A., Jiang, J., Malana, G. E., Nuovo, G. J., Chotani, M., Feldman, D. S., Schmittgen, T. D., and Elton, T. S. (2007). *J Biol Chem* **282,** 24262-9.

Mataix-Cols, D. (2006). *Curr Opin Psychiatry* **19,** 84-9.
Mataix-Cols, D., Rosario-Campos, M. C., and Leckman, J. F. (2005). *Am J Psychiatry* **162,** 228-38.
Mathews, C. A., Nievergelt, C. M., Azzam, A., Garrido, H., Chavira, D. A., Wessel, J., Bagnarello, M., Reus, V. I., and Schork, N. J. (2007). *Am J Med Genet B Neuropsychiatr Genet* **144,** 174-82.
Matsushita, S., Muramatsu, T., Kimura, M., Shirakawa, O., Mita, T., Nakai, T., and Higuchi, S. (1997). *Mol Psychiatry* **2,** 390-2.
Maziere, P., and Enright, A. J. (2007). *Drug Discov Today* **12,** 452-8.
McGuffin, P., Rijsdijk, F., Andrew, M., Sham, P., Katz, R., and Cardno, A. (2003). *Arch Gen Psychiatry* **60,** 497-502.
McLaughlin, K. A., and Hatzenbuehler, M. L. (2009). *J Abnorm Psychol* **118,** 659-69.
Melo, S. A., Ropero, S., Moutinho, C., Aaltonen, L. A., Yamamoto, H., Calin, G. A., Rossi, S., Fernandez, A. F., Carneiro, F., Oliveira, C., Ferreira, B., Liu, C. G., Villanueva, A., Capella, G., Schwartz, S., Jr., Shiekhattar, R., and Esteller, M. (2009). *Nat Genet* **41,** 365-70.
Meng, F., Henson, R., Wehbe-Janek, H., Ghoshal, K., Jacob, S. T., and Patel, T. (2007). *Gastroenterology* **133,** 647-58.
Menn, B., Timsit, S., Calothy, G., and Lamballe, F. (1998). *J Comp Neurol* **401,** 47-64.
Menn, B., Timsit, S., Represa, A., Mateos, S., Calothy, G., and Lamballe, F. (2000). *Eur J Neurosci* **12,** 3211-23.
Mercader, J. M., Saus, E., Aguera, Z., Bayes, M., Boni, C., Carreras, A., Cellini, E., de Cid, R., Dierssen, M., Escaramis, G., Fernandez-Aranda, F., Forcano, L., Gallego, X., Gonzalez, J. R., Gorwood, P., Hebebrand, J., Hinney, A., Nacmias, B., Puig, A., Ribases, M., Ricca, V., Romo, L., Sorbi, S., Versini, A., Gratacos, M., and Estivill, X. (2008). *Hum Mol Genet*.
Miguel, E. C., Leckman, J. F., Rauch, S., do Rosario-Campos, M. C., Hounie, A. G., Mercadante, M. T., Chacon, P., and Pauls, D. L. (2005). *Mol Psychiatry* **10,** 258-75.
Millar, A. A., and Waterhouse, P. M. (2005). *Funct Integr Genomics* **5,** 129-35.
Miranda, K. C., Huynh, T., Tay, Y., Ang, Y. S., Tam, W. L., Thomson, A. M., Lim, B., and Rigoutsos, I. (2006). *Cell* **126,** 1203-17.
Mishra, P. J., Humeniuk, R., Mishra, P. J., Longo-Sorbello, G. S., Banerjee, D., and Bertino, J. R. (2007). *Proc Natl Acad Sci U S A* **104,** 13513-8.
Miyasaka, K., Yoshida, Y., Matsushita, S., Higuchi, S., Shirakawa, O., Shimokata, H., and Funakoshi, A. (2004). *Am J Med Genet B Neuropsychiatr Genet* **127B,** 78-80.
Moreau, J. L., and Huber, G. (1999). *Brain Res Brain Res Rev* **31,** 65-82.
Nakamura, M., Ueno, S., Sano, A., and Tanabe, H. (1999). *Mol Psychiatry* **4,** 155-62.
Nash, J. R., Sargent, P. A., Rabiner, E. A., Hood, S. D., Argyropoulos, S. V., Potokar, J. P., Grasby, P. M., and Nutt, D. J. (2008). *Br J Psychiatry* **193,** 229-34.

Nestler, E. J., Gould, E., Manji, H., Buncan, M., Duman, R. S., Greshenfeld, H. K., Hen, R., Koester, S., Lederhendler, I., Meaney, M., Robbins, T., Winsky, L., and Zalcman, S. (2002). *Biol Psychiatry* **52,** 503-28.

Neumeister, A., Bain, E., Nugent, A. C., Carson, R. E., Bonne, O., Luckenbaugh, D. A., Eckelman, W., Herscovitch, P., Charney, D. S., and Drevets, W. C. (2004). *J Neurosci* **24,** 589-91.

Neuroscience, S. f. (2005). "*Brain Facts A primer on the brain and nervous system*." Society for Neuroscience, Washington DC.

Neves-Pereira, M., Mundo, E., Muglia, P., King, N., Macciardi, F., and Kennedy, J. L. (2002). *Am J Hum Genet* **71,** 651-5.

Noyes, R., Jr., Crowe, R. R., Harris, E. L., Hamra, B. J., McChesney, C. M., and Chaudhry, D. R. (1986). *Arch Gen Psychiatry* **43,** 227-32.

Oliveira-Dos-Santos, A. J., Matsumoto, G., Snow, B. E., Bai, D., Houston, F. P., Whishaw, I. Q., Mariathasan, S., Sasaki, T., Wakeham, A., Ohashi, P. S., Roder, J. C., Barnes, C. A., Siderovski, D. P., and Penninger, J. M. (2000). *Proc Natl Acad Sci U S A* **97,** 12272-7.

Ozsolak, F., Platt, A. R., Jones, D. R., Reifenberger, J. G., Sass, L. E., McInerney, P., Thompson, J. F., Bowers, J., Jarosz, M., and Milos, P. M. (2009). *Nature* **461,** 814-8.

Palko, M. E., Coppola, V., and Tessarollo, L. (1999). *J Neurosci* **19,** 775-82.

Parker, R., and Song, H. (2004). *Nat Struct Mol Biol* **11,** 121-7.

Parsey, R. V., Olvet, D. M., Oquendo, M. A., Huang, Y. Y., Ogden, R. T., and Mann, J. J. (2006). *Neuropsychopharmacology* **31,** 1745-9.

Pauley, K. M., and Chan, E. K. (2008). *Ann N Y Acad Sci* **1143,** 226-39.

Penagarikano, O., Mulle, J. G., and Warren, S. T. (2007). *Annu Rev Genomics Hum Genet* **8,** 109-29.

Perkins, D. O., Jeffries, C. D., Jarskog, L. F., Thomson, J. M., Woods, K., Newman, M. A., Parker, J. S., Jin, J., and Hammond, S. M. (2007). *Genome Biol* **8,** R27.

Perna, G., Caldirola, D., Arancio, C., and Bellodi, L. (1997). *Psychiatry Res* **66,** 69-71.

Petersen, C. P., Bordeleau, M. E., Pelletier, J., and Sharp, P. A. (2006). *Mol Cell* **21,** 533-42.

Plomin, R., Owen, M. J., and McGuffin, P. (1994). *Science* **264,** 1733-9.

Poy, M. N., Eliasson, L., Krutzfeldt, J., Kuwajima, S., Ma, X., Macdonald, P. E., Pfeffer, S., Tuschl, T., Rajewsky, N., Rorsman, P., and Stoffel, M. (2004). *Nature* **432,** 226-30.

Quach, H., Barreiro, L. B., Laval, G., Zidane, N., Patin, E., Kidd, K. K., Kidd, J. R., Bouchier, C., Veuille, M., Antoniewski, C., and Quintana-Murci, L. (2009). *Am J Hum Genet* **84,** 316-27.

Raleigh, M. J., McGuire, M. T., Brammer, G. L., Pollack, D. B., and Yuwiler, A. (1991). *Brain Res* **559,** 181-90.

Rang, H. P. (2003). "*Pharmacology*." Churchill Livingstone, Edinburgh.

Regier, D. A., Narrow, W. E., and Rae, D. S. (1990). *J Psychiatr Res* **24 Suppl 2,** 3-14.

Reichardt, L. F. (2006). *Philos Trans R Soc Lond B Biol Sci* **361,** 1545-64.

Richard, I. H., Schiffer, R. B., and Kurlan, R. (1996). *J Neuropsychiatry Clin Neurosci* **8,** 383-92.
Risch, N., and Merikangas, K. (1996). *Science* **273,** 1516-7.
Ro, S., Park, C., Sanders, K. M., McCarrey, J. R., and Yan, W. (2007). *Dev Biol* **311,** 592-602.
Rockman, M. V., and Wray, G. A. (2002). *Mol Biol Evol* **19,** 1991-2004.
Rodriguez, A., Griffiths-Jones, S., Ashurst, J. L., and Bradley, A. (2004). *Genome Res* **14,** 1902-10.
Rogan, M. T., Staubli, U. V., and LeDoux, J. E. (1997). *Nature* **390,** 604-7.
Rosenthal, M., Stelian, J., Wagner, J., and Berkman, P. (1999). *Isr J Psychiatry Relat Sci* **36,** 29-34.
Rothe, C., Gutknecht, L., Freitag, C., Tauber, R., Mossner, R., Franke, P., Fritze, J., Wagner, G., Peikert, G., Wenda, B., Sand, P., Jacob, C., Rietschel, M., Nothen, M. M., Garritsen, H., Fimmers, R., Deckert, J., and Lesch, K. P. (2004). *Int J Neuropsychopharmacol* **7,** 189-92.
Samuels, J. F., Bienvenu, O. J., 3rd, Pinto, A., Fyer, A. J., McCracken, J. T., Rauch, S. L., Murphy, D. L., Grados, M. A., Greenberg, B. D., Knowles, J. A., Piacentini, J., Cannistraro, P. A., Cullen, B., Riddle, M. A., Rasmussen, S. A., Pauls, D. L., Willour, V. L., Shugart, Y. Y., Liang, K. Y., Hoehn-Saric, R., and Nestadt, G. (2007). *Behav Res Ther* **45,** 673-86.
Sansone, R. A., Sansone, L. A., and Righter, E. L. (1998). *J Womens Health* **7,** 983-9.
Saunders, M. A., Liang, H., and Li, W. H. (2007). *Proc Natl Acad Sci U S A* **104,** 3300-5.
Schaefer, A., O'Carroll, D., Tan, C. L., Hillman, D., Sugimori, M., Llinas, R., and Greengard, P. (2007). *J Exp Med* **204,** 1553-8.
Schildkraut, J. M., Risch, N., and Thompson, W. D. (1989). *Am J Hum Genet* **45,** 521-9.
Schindler, K. M., Richter, M. A., Kennedy, J. L., Pato, M. T., and Pato, C. N. (2000). *Am J Med Genet* **96,** 721-4.
Schinka, J. A., Busch, R. M., and Robichaux-Keene, N. (2004). *Mol Psychiatry* **9,** 197-202.
Schneier, F. R., Liebowitz, M. R., Abi-Dargham, A., Zea-Ponce, Y., Lin, S. H., and Laruelle, M. (2000). *Am J Psychiatry* **157,** 457-9.
Schratt, G. M., Tuebing, F., Nigh, E. A., Kane, C. G., Sabatini, M. E., Kiebler, M., and Greenberg, M. E. (2006). *Nature* **439,** 283-9.
Schulze, T. G., Muller, D. J., Krauss, H., Scherk, H., Ohlraun, S., Syagailo, Y. V., Windemuth, C., Neidt, H., Grassle, M., Papassotiropoulos, A., Heun, R., Nothen, M. M., Maier, W., Lesch, K. P., and Rietschel, M. (2000). *Am J Med Genet* **96,** 801-3.
Sebastiao, A. M., and Ribeiro, J. A. (2000). *Trends Pharmacol Sci* **21,** 341-6.
Seedat, S., and Stein, D. J. (2002). *Psychiatry Clin Neurosci* **56,** 17-23.
Sempere, L. F., Freemantle, S., Pitha-Rowe, I., Moss, E., Dmitrovsky, E., and Ambros, V. (2004). *Genome Biol* **5,** R13.

Sethupathy, P., Borel, C., Gagnebin, M., Grant, G. R., Deutsch, S., Elton, T. S., Hatzigeorgiou, A. G., and Antonarakis, S. E. (2007). *Am J Hum Genet* **81,** 405-13.

Sethupathy, P., Megraw, M., and Hatzigeorgiou, A. G. (2006). *Nat Methods* **3,** 881-6.

Shelton, D. L., Sutherland, J., Gripp, J., Camerato, T., Armanini, M. P., Phillips, H. S., Carroll, K., Spencer, S. D., and Levinson, A. D. (1995). *J Neurosci* **15,** 477-91.

Shimizu, E., Hashimoto, K., Kobayashi, K., Mitsumori, M., Ohgake, S., Koizumi, H., Okamura, N., Koike, K., Kumakiri, C., Nakazato, M., Komatsu, N., and Iyo, M. (2004). *Neurosci Lett* **363,** 81-3.

Shugart, Y. Y., Samuels, J., Willour, V. L., Grados, M. A., Greenberg, B. D., Knowles, J. A., McCracken, J. T., Rauch, S. L., Murphy, D. L., Wang, Y., Pinto, A., Fyer, A. J., Piacentini, J., Pauls, D. L., Cullen, B., Page, J., Rasmussen, S. A., Bienvenu, O. J., Hoehn-Saric, R., Valle, D., Liang, K. Y., Riddle, M. A., and Nestadt, G. (2006). *Mol Psychiatry* **11,** 763-70.

Siegel, G., Obernosterer, G., Fiore, R., Oehmen, M., Bicker, S., Christensen, M., Khudayberdiev, S., Leuschner, P. F., Busch, C. J., Kane, C., Hubel, K., Dekker, F., Hedberg, C., Rengarajan, B., Drepper, C., Waldmann, H., Kauppinen, S., Greenberg, M. E., Draguhn, A., Rehmsmeier, M., Martinez, J., and Schratt, G. M. (2009). *Nat Cell Biol* **11,** 705-16.

Simon, N. M. (2009). *J Clin Psychiatry* **70 Suppl 2,** 10-4.

Sklar, P., Gabriel, S. B., McInnis, M. G., Bennett, P., Lim, Y. M., Tsan, G., Schaffner, S., Kirov, G., Jones, I., Owen, M., Craddock, N., DePaulo, J. R., and Lander, E. S. (2002). *Mol Psychiatry* **7,** 579-93.

Sklar, P., Smoller, J. W., Fan, J., Ferreira, M. A., Perlis, R. H., Chambert, K., Nimgaonkar, V. L., McQueen, M. B., Faraone, S. V., Kirby, A., de Bakker, P. I., Ogdie, M. N., Thase, M. E., Sachs, G. S., Todd-Brown, K., Gabriel, S. B., Sougnez, C., Gates, C., Blumenstiel, B., Defelice, M., Ardlie, K. G., Franklin, J., Muir, W. J., McGhee, K. A., MacIntyre, D. J., McLean, A., VanBeck, M., McQuillin, A., Bass, N. J., Robinson, M., Lawrence, J., Anjorin, A., Curtis, D., Scolnick, E. M., Daly, M. J., Blackwood, D. H., Gurling, H. M., and Purcell, S. M. (2008). *Mol Psychiatry* **13,** 558-69.

Slater, E. (1936). *Proc. R. Soc. Med* **29,** 39-48.

Smirnova, L., Grafe, A., Seiler, A., Schumacher, S., Nitsch, R., and Wulczyn, F. G. (2005). *Eur J Neurosci* **21,** 1469-77.

Smoller, J. W., and Tsuang, M. T. (1998). *Am J Psychiatry* **155,** 1152-62.

Stark, A., Brennecke, J., Bushati, N., Russell, R. B., and Cohen, S. M. (2005). *Cell* **123,** 1133-46.

Stein, D. J. (2002). *Lancet* **360,** 397-405.

Strachan, T., and Read, P. (2004). "*Human Molecular Genetics*." Garland Science, London and New York.

Strug, L. J., Suresh, R., Fyer, A. J., Talati, A., Adams, P. B., Li, W., Hodge, S. E., Gilliam, T. C., and Weissman, M. M. (2008). *Mol Psychiatry*.

Suh, M. R., Lee, Y., Kim, J. Y., Kim, S. K., Moon, S. H., Lee, J. Y., Cha, K. Y., Chung, H. M., Yoon, H. S., Moon, S. Y., Kim, V. N., and Kim, K. S. (2004). *Dev Biol* **270,** 488-98.

Sullivan, G. M., Coplan, J. D., Kent, J. M., and Gorman, J. M. (1999). *Biol Psychiatry* **46,** 1205-18.

Sullivan, P. F., Kendler, K. S., and Neale, M. C. (2003). *Arch Gen Psychiatry* **60,** 1187-92.

Sun, B. K., and Tsao, H. (2008). *J Am Acad Dermatol* **59,** 725-37; quiz 738-40.

Szymanski, M., Barciszewska, M. Z., Erdmann, V. A., and Barciszewski, J. (2005). *Biochim Biophys Acta* **1756,** 65-75.

Taganov, K. D., Boldin, M. P., Chang, K. J., and Baltimore, D. (2006). *Proc Natl Acad Sci U S A* **103,** 12481-6.

Tan, Z., Randall, G., Fan, J., Camoretti-Mercado, B., Brockman-Schneider, R., Pan, L., Solway, J., Gern, J. E., Lemanske, R. F., Nicolae, D., and Ober, C. (2007). *Am J Hum Genet* **81,** 829-34.

Tancer, M. E. (1993). *J Clin Psychiatry* **54 Suppl,** 26-30.

Teng, K. K., and Hempstead, B. L. (2004). *Cell Mol Life Sci* **61,** 35-48.

Thoeringer, C. K., Ripke, S., Unschuld, P. G., Lucae, S., Ising, M., Bettecken, T., Uhr, M., Keck, M. E., Mueller-Myhsok, B., Holsboer, F., Binder, E. B., and Erhardt, A. (2009). *J Neural Transm* **116,** 649-57.

Thomas, D. C. (2004). "*Statistical Methods in Genetic Epidemiology*." Oxford University Press, New York.

Tiihonen, J., Kuikka, J., Bergstrom, K., Lepola, U., Koponen, H., and Leinonen, E. (1997). *Am J Psychiatry* **154,** 239-42.

Torgersen, S. (1983). *Arch Gen Psychiatry* **40,** 1085-9.

Tsoulfas, P., Stephens, R. M., Kaplan, D. R., and Parada, L. F. (1996). *J Biol Chem* **271,** 5691-7.

Unschuld, P. G., Ising, M., Erhardt, A., Lucae, S., Kloiber, S., Kohli, M., Salyakina, D., Welt, T., Kern, N., Lieb, R., Uhr, M., Binder, E. B., Muller-Myhsok, B., Holsboer, F., and Keck, M. E. (2007). *Am J Med Genet B Neuropsychiatr Genet* **144B,** 424-9.

van den Heuvel, O. A., van de Wetering, B. J., Veltman, D. J., and Pauls, D. L. (2000). *J Clin Psychiatry* **61,** 756-66.

van der Linden, G. J., Stein, D. J., and van Balkom, A. J. (2000). *Int Clin Psychopharmacol* **15 Suppl 2,** S15-23.

Vasudevan, S., Tong, Y., and Steitz, J. A. (2007). *Science* **318,** 1931-4.

Visvanathan, J., Lee, S., Lee, B., Lee, J. W., and Lee, S. K. (2007). *Genes Dev* **21,** 744-9.

Wakiyama, M., Takimoto, K., Ohara, O., and Yokoyama, S. (2007). *Genes Dev* **21,** 1857-62.

Wang, G., van der Walt, J. M., Mayhew, G., Li, Y. J., Zuchner, S., Scott, W. K., Martin, E. R., and Vance, J. M. (2008a). *Am J Hum Genet* **82,** 283-9.

Wang, W. X., Rajeev, B. W., Stromberg, A. J., Ren, N., Tang, G., Huang, Q., Rigoutsos, I., and Nelson, P. T. (2008b). *J Neurosci* **28,** 1213-23.

Weiller, E., Bisserbe, J. C., Maier, W., and Lecrubier, Y. (1998). *Br J Psychiatry Suppl***,** 18-23.

Weissman, M. M., Bland, R. C., Canino, G. J., Faravelli, C., Greenwald, S., Hwu, H. G., Joyce, P. R., Karam, E. G., Lee, C. K., Lellouch, J., Lepine, J. P., Newman, S. C., Oakley-Browne, M. A., Rubio-Stipec, M., Wells, J. E., Wickramaratne, P. J., Wittchen, H. U., and Yeh, E. K. (1997). *Arch Gen Psychiatry* **54,** 305-9.

Willour, V. L., Yao Shugart, Y., Samuels, J., Grados, M., Cullen, B., Bienvenu, O. J., 3rd, Wang, Y., Liang, K. Y., Valle, D., Hoehn-Saric, R., Riddle, M., and Nestadt, G. (2004). *Am J Hum Genet* **75,** 508-13.

Wong, C. F., and Tellam, R. L. (2008). *J Biol Chem* **283,** 9836-43.

Woo, J. M., Yoon, K. S., and Yu, B. H. (2002). *Am J Psychiatry* **159,** 1785-7.

Yalcin, B., Willis-Owen, S. A., Fullerton, J., Meesaq, A., Deacon, R. M., Rawlins, J. N., Copley, R. R., Morris, A. P., Flint, J., and Mott, R. (2004). *Nat Genet* **36,** 1197-202.

Yamada, K., Hattori, E., Shimizu, M., Sugaya, A., Shibuya, H., and Yoshikawa, T. (2001). *J Neural Transm* **108,** 837-48.

Yang, H., Dinney, C. P., Ye, Y., Zhu, Y., Grossman, H. B., and Wu, X. (2008). *Cancer Res* **68,** 2530-7.

Yang, W., Chendrimada, T. P., Wang, Q., Higuchi, M., Seeburg, P. H., Shiekhattar, R., and Nishikura, K. (2006). *Nat Struct Mol Biol* **13,** 13-21.

Yee, B. K., Zhu, S. W., Mohammed, A. H., and Feldon, J. (2007). *J Neural Transm* **114,** 431-44.

Yekta, S., Shih, I. H., and Bartel, D. P. (2004). *Science* **304,** 594-6.

Yoon, S., and De Micheli, G. (2006). *Birth Defects Res C Embryo Today* **78,** 118-28.

Zhou, X., Ruan, J., Wang, G., and Zhang, W. (2007). *PLoS Comput Biol* **3,** e37.

Zhu, L., Scelfo, B., Tempia, F., Sacchetti, B., and Strata, P. (2006). *Neuroscience* **140,** 801-10.

Zuchner, S., Wang, G., Tran-Viet, K. N., Nance, M. A., Gaskell, P. C., Vance, J. M., Ashley-Koch, A. E., and Pericak-Vance, M. A. (2006). *Am J Hum Genet* **79,** 365-9.

Abbreviations

Aa	aminoacid
AAO	age at onset
ABCG1	ATP-binding cassette, sub-family G-Drosophila white
Ach	acetylcholine
AD	Alzheimer's disease
ADORA	adenosine receptor
Ago	argonaute
ALDH2	aldehyde dehydrogenase 2
BACE1	beta-site amyloid precursor protein-cleaving enzyme 1
BDNF	brain-derived neurotrophic factor
bp	base pair
CCK	cholecystokinin
CCKAR, CCKBR	cholecystokinin Receptors A and B
CDK	cyclin-dependent Kinase
CEPH	Centre d'Etude du Polymorphisme Humain
CLL	chronic lymphocytic leukemia
CNS	central nervous system
CO	carbon monoxide
COMT	catechol-O-methyltransferase
CRF	corticotropin-releasing factor
CRHR1	corticotropin releasing hormone receptor 1
Da, kDa	Dalton, kiloDalton
DAOA	D-amino-oxidase activator
DNA	deoxyribonucleic acid
DSM	Diagnostic and Statistical Manual of Mental Disorders
dsRNA	double-stranded RNA
DZ	dizygotic
ERK	extracellular signal-regulated kinase
EST	expressed sequence tag
FBS	fetal bovine serum
FL-NTRK3	full-length NTRK3
FMRP	fragile X protein
FX	fragile X
GABA	gamma-aminobutyric acid
GABRA2	gamma-aminobutyric acid receptor alpha 2 subunit gene
GAD	generalised anxiety disorder
GAPDH	glyceraldehyde 3-phosphate dehydrogenase
GBM	glioblastoma
GTP, GDP	Guanosine-5'-triphosphate, Guanosine-5'-diphosphate
GWAS	genome wide association studies
HTR	serotonin receptor
ICD	International Statistical Classification of Diseases and Related Health Problems
IPA	Ingenuity pathway analysis
IRES	internal ribosomal entry sites
LC	locus ceruleus
LD	linkage disequilibrium

LOD	logarithm of the odds
MAOA	monoamine Oxidase A
MAPK	mitogen activated protein kinase
Met	methionine
miRISC	miRNA incorporated into the RISC complex
miRNA	microRNA
mRNA	messenger RNA
MZ	monozygotic
ncRNA	non-coding RNA
NE	norepinephrine
NGF	neurotrophic growth factor
NO	nitric oxide
nt	nucleotide
NT	neurotrophin
NTRK	neurotrophic tyrosine kinase receptor
OCD	obsessive-compulsive disorder
ORF	open reading frame
p75NTR	p75 neurotrophin receptor
PAG	periaquaductal grey
P-body	processing body
PCR	polymerase chain reaction
PD	panic disorder
PI3K	phosphoinositide 3-kinase
piRNA	piwi interacting RNA
PolII	RNA polymerase II
pre-miRNA	precursor miRNA
pri-miRNA	primary miRNA
PTBP1	polypyrimidine tract-binding protein 1
PTC	papillary thyroid carcinoma
PTSD	posttraumatic stress disorder
PVN	paraventricular nucleus of the hypothalamus
RGS	regulator of G protein signaling
RISC	RNA-induced silencing complex
RNA	ribonucleic acid
RNP complex	ribonucleoprotein complex
RT-PCR	reverse transcription PCR
SE	standard error
siRNA	small interfering RNA
SLC6	serotonin transporter
SNP	single nucleotide polymorphism
SSRI	selective serotonin reuptake inhibitor
STR	short tandem repeat
TK	tyrosine kinase
TRBP	RNA-binding protein
TR-NTRK3	truncated NTRK3
TU	transcriptional unit
UTR	untranslated region
Val	valine
VNTR	variable number of tandem repeats
WTCCC	Welcome Trust Case-Control Consortium

Annex

CLASSIFICATION:

Biological Sciences

Cell Biology

TITLE:

Overexpression of miR-128 specifically inhibits the truncated isoform of *NTRK3* and upregulates *BCL2* in SH-SY5Y neuroblastoma cells

Monica Guidi[a,b,c], Margarita Muiños-Gimeno[a,b,c], Birgit Kagerbauer[a,b], Eulàlia Martí[a,b], Xavier Estivill[a,b,c,1], Yolanda Espinosa-Parrilla[a,b]

[a]Center for Genomic Regulation (CRG), Genes and Disease Program, Dr. Aiguader 88, 08003 Barcelona, Spain

[b]Public Health and Epidemiology Network Biomedical Research Center (CIBERESP)

[c]Pompeu Fabra University (UPF), Dr. Aiguader 80, 08003 Barcelona, Spain

[1]Corresponding author

Xavier Estivill, Center for Genomic Regulation (CRG), Genes and Disease Program, Dr. Aiguader 88, 08003 Barcelona, Spain. Tel: +34 93 3160159, Fax: +34 93 3160099, E-mail: Xavier.estivill@crg.cat

Footnote

Author contributions: M.G., Y.E.P. and X.E. designed research; E.M. and M.M.G helped with the design of some experiments and contributed new reagents; M.G. and B.K. performed research; M.G., M.M.G. and Y.E.P. analyzed data; M.G. and Y.E.P wrote the paper; X.E. supervised and revised the manuscript.

The authors declare no conflict of interest.

Abstract

Neurotrophins and their receptors are key molecules in the regulation of neuronal differentiation; they mediate the survival of neurons during development and adulthood and are implicated in synaptic plasticity. The human neurotrophin-3 receptor gene*NTRK3* yields two major isoforms, a full-length kinase-active form and a truncated non-catalytic form, which activates a specific pathway affecting membrane remodeling and cytoskeletal reorganization. The two variants present non-overlapping 3'UTRs, indicating that they might be differentially regulated at the post-transcriptional level. MicroRNAs are small non-coding RNAs with important roles in gene regulation, especially in the nervous system. We provide evidence that the two isoforms of *NTRK3* are targeted by different sets of microRNAs. We identify two microRNAs (miR-24 and miR-151-3p) that repress the full-length isoform and four microRNAs (miR-128, miR-485-3p, miR-765 and miR-768-5p) that repress the truncated isoform. We show that overexpression of miR-128, a brain-enriched microRNA, causes morphological changes in SH-SY5Y neuroblastoma cells, similar to those observed using an siRNA specifically directed against truncated *NTRK3*. Transcriptome analysis of cells transfected with miR-128 reveals altered expression of genes implicated in cytoskeletal organization, indicating that miR-128 may be involved in this process at least in part through the inhibition of *NTRK3*. In addition, a significant increase in cell number is observed upon miR-128 overexpression, with consistent deregulation of genes involved in apoptosis, cell survival and proliferation, such as *BCL2*. Our results show that the regulation of *NTRK3* by microRNAs is isoform-specific and suggest that neurotrophin-mediated processes are strongly linked to microRNA-dependent mechanisms.

Introduction

Neurotrophins are a family of growth factors that play important roles in the nervous system. They exert multiple functions, being crucial in the survival and maintenance of the central and peripheral nervous system as well as in axon and dendrite patterning. More recent evidence has shown that neurotrophins can also act as synaptic modulators with a key role in synaptic plasticity and are consequently involved in cognitive processes, learning and memory formation (1).

In mammals, the neurotrophin family is composed of four members: nerve growth factor (NGF), brain-derived neurotrophic factor (BDNF), neurotrophin-3 (NT3), and neurotrophin-4/5 (NT4/5). Each of them binds with high affinity to a specific neurotrophic tyrosine kinase (NTRK) receptor: NGF is the preferred ligand for NTRK1 (TrkA), BDNF and NT4/5 for NTRK2 (TrkB), and NT3 for NTRK3 (TrkC) (2). Upon neurotrophin-induced stimulation, NTRK receptors can activate the Ras/MAPK pathway, the PI3K (phosphoinositide 3-kinase) pathway, and/or PLC-γ1 (phospholipase C gamma 1)-dependent signaling, respectively promoting cell survival, differentiation and activity-dependent plasticity.

The human *NTRK3* gene is located on chromosome 15q25 and spans ~380 kb of genomic DNA (chr15: 86,220,992-86,600,665, March 2006, hg18). It contains 19 introns and undergoes alternative splicing. In humans, three transcript variants have been well characterized: a full-length catalytic form (NM_002530) containing a tyrosine kinase (TK) domain, a full-length isoform (NM_001012338) with an insertion of 14 amino acids in the TK domain - which is less abundant and shows reduced signaling potential (3) - and a single non-catalytic truncated form (NM_001007156) that completely lacks the TK domain (4). Two major protein isoforms have correspondingly been detected in human brain samples: the full-length, kinase-active 150-kDa receptor (FL-NTRK3) and a truncated non-catalytic isoform (TR-NTRK3) of 50 kDa (5). The main function that has long been attributed to the truncated isoform is the inhibition of catalytic receptors, which is achieved through a dominant-negative or a ligand-sequestering mechanism. However, recent work has identified a new signaling pathway activated by the truncated

isoform that links NT3 to downstream molecules affecting membrane remodeling, cytoskeletal reorganization and cell movement, such as the scaffold protein tamalin, the Rac1 GTPase and the adenosine diphosphate-ribosylation factor Arf6 (6).

In humans the full-length receptor is expressed at low but relatively constant levels throughout development, while the truncated receptor is expressed at moderate levels early in development and increases to reach mature levels by adolescence. In contrast, both full-length and truncated transcripts are uniformly expressed throughout postnatal life and decline in ageing (5). This discrepancy between changes in protein and mRNA levels of NTRK3 suggests that post-transcriptional regulation may play a role in controlling the expression of the two isoforms, in addition to alternative splicing.

microRNAs (miRNAs) are post-transcriptional regulators that have been shown to play a key role in the nervous system. The distribution of miRNAs in the developing and adult nervous system is spatially and/or temporally restricted (7), indicating that they may contribute to the fine-tuning of neuronal gene expression. Mature miRNAs are small RNAs of approximately 22 nucleotides in length that regulate gene expression by means of partial complementarity to miRNA binding sites located in the 3′ untranslated regions (3′UTRs) of target mRNAs (8). In particular, perfect complementarity between nucleotide #2 through #7 or #8 at the 5′ end of the mature miRNA - the so-called *seed* region - and the 3′UTR of genes is considered determinant for successful binding (9). Genes with long 3′UTRs are more prone to miRNA-mediated regulation as compared to genes with short 3′UTRs, which tend to be specifically deprived of miRNA target sites (10). Interestingly, full-length and truncated *NTRK3* transcripts show different 3′UTR regions that do not overlap, supporting the hypothesis that the two isoforms may be differentially regulated by miRNAs. Furthermore, it has recently been shown that miRNAs miR-9 and miR-125a/b are able to specifically regulate the expression of the truncated isoform of the NTRK3 receptor in human neuroblastoma cells (11).

Only one study has previously analyzed the contribution of miRNAs to determining the balance between different variants of the same gene (12). Here, we show that the full-length and truncated isoforms of *NTRK3* are regulated by different sets of miRNAs. In

addition, we demonstrate that, among the miRNAs that inhibit the truncated isoform, the overexpression of miR-128 – a brain-enriched miRNA – in SH-SY5Y cells alters the expression profile of genes involved in cytoskeletal organization and of genes related with apoptosis and cell cycle regulation, including the anti-apoptotic factor BCL2.

Results

Functional screening of miRNA target sites in NTRK3

The full-length and the truncated isoform of *NTRK3* were searched for putative target sites with three widely used miRNA target prediction programs - miRanda, TargetScan and PicTar. Taking into account the 325 human miRNAs annotated in the miRbase database at the time of the analysis (version 7.1, October 2005) we could identify three miRNAs predicted to target the full-length isoform and 29 miRNAs predicted to target the truncated isoform (Table 1).

The entire 3'UTRs of the full-length and truncated isoforms of *NTRK3* (178 and 1981 bp respectively) were cloned immediately downstream of the firefly luciferase open reading frame in the pGL4.13 plasmid, and the resulting constructs were designated pGL4.13-FL and pGL4.13-TR. The renilla reporter plasmid pGL4.75 was used to control for transfection efficiency. A total of 55 miRNAs and two non-targeting controls were tested on both pGL4.13-FL and pGL4.13-TR. Tested miRNAs included the 32 mentioned above; six additional miRNAs that appeared in more recent versions of the miRbase database (miR-485-3p, miR-509, miR-617, miR-625, miR-765 and miR-768-5p) and had putative target sites in the truncated isoform were also included because allelic variants were found in those target sites in patients with anxiety disorders (13). Finally, seventeen miRNAs with no predicted target sites according to the interrogated methods were also analyzed for comparison (Table 1).

The specific luciferase activity of the pGL4.13-FL construct was significantly decreased by three miRNAs (Figure 1), two of which (miR-151-3p and miR-185) were predicted to target the full-length isoform, whereas one (miR-24) was not. The strongest repression was observed with miR-185 (54% reduction), followed by miR-151-3p (20%) and miR-24

(10%). In the case of pGL4.13-TR, luciferase activity was significantly reduced by eight miRNAs (Figure 1), all of which were predicted by at least one program: miR-128, miR-324-5p, miR-330, miR-485-3p, miR-509, miR-625, miR-765 and miR-768-5p. The most conspicuous inhibition was detected with miR-625 (62% reduction), miR-509 (47%) and miR-128 (32%), while the other five miRNAs gave a reduction ranging 13-30%.

Even though miR-24 had no predicted target sites in the full-length isoform - due to the lack of perfect complementarity with the seed region of miR-24 - it was indeed able to repress FL-NTRK3. We therefore checked for miR-24 potential binding sites using RNAhybrid, a program that finds the energetically most favorable hybridizations of a small RNA to a large RNA (14) leaving perfect complementarity constraints in the 5′ part of the miRNA. This analysis indicated a number of potential target sites for miR-24, the top hit being located at nucleotide positions 1726-1758 of the 3′UTR of FL-NTRK3, with a minimum free energy of -28.1 kcal/mol (Figure S1).

miRNA-mediated regulation of endogenous NTRK3 in neuroblastoma cells

SH-SY5Y neuroblastoma cells were used to investigate whether miRNAs causing a significant decrease in luciferase activity were also able to downregulate endogenous NTRK3. Upon exposure to retinoic acid (RA), SH-SY5Y cells reduce their growth rate and differentiate by extending neurites into cells that are biochemically, ultrastructurally and electrophysiologically similar to neurons (15). Furthermore, RA treatment has been shown to induce the appearance of functional FL-NTRK3 receptors in this cell line (16).

We characterized the expression of NTRK3 in SH-SY5Y cells at different time points during RA exposure by western blotting. As previously described for another neuroblastoma cell line (11), we could observe an increase in the levels of the full-length isoform and a decrease of the truncated isoform. Full-length protein levels were ~2.5-fold higher after three days of RA treatment, ~4-fold higher after six days and reached up to ~8-fold after ten days. On the other hand the expression of the truncated isoform was reduced by approximately 50% after three days of exposure, and kept decreasing very slightly through day 10 (Figure 2). Given these expression profiles, untreated cells

were used to study the regulation of the truncated isoform and RA-differentiated cells were used for the full-length isoform.

Endogenous expression for the eight luciferase-validated miRNAs was also analyzed using custom oligonucleotide microarrays in SH-SY5Y cells (Table 2). With the exception of miR-24, the expression of these miRNAs is in general low, making miRNA overexpression the method of choice to analyze the regulation of endogenous NTRK3 in this cell system. Luciferase-validated miRNAs were therefore transfected into either undifferentiated (miR-128, miR-324-5p, miR-330, miR-485-3p, miR-509, miR-625, miR-765 and miR-768-5p) or differentiated SH-SY5Y cells (miR-24, miR-151-3p and miR-185), and protein levels were assessed by western blotting 72 h after transfection (Figure 3A). In agreement with the luciferase assay data, FL-NTRK3 levels were significantly reduced by miR-24 and miR-151-3p; a slight inhibition was also observed with miR-185 but did not reach statistical significance. The degrees of reduction varied compared to those observed with the luciferase constructs: the strongest downregulation was in fact obtained with miR-24 (41%), followed by miR-151-3p (34%). To investigate whether such inhibition of FL-NTRK3 was able to affect the efficiency of NT3-induced signaling, we analyzed ERK1/2 phosphorylation by acutely stimulating transfected cells with NT3, after serum starvation. ERK phosphorylation was used as a readout for the activation of the Ras/MAPK pathway, which is triggered by stimulation of full-length, kinase-active NTRK receptors (17). As expected, we could detect a reduction in ERK phosphorylation with miR-24 and miR-151-3p but not with miR-185 (Figure 3B).

As for TR-NTRK3, a significant downregulation ranging between 20% and 30% was observed with miR-128, miR-485-3p, miR-765 and miR-768-5p, the strongest repression being obtained with miR-128 (32% reduction) and miR-485-3p (30%). The remaining four miRNAs (miR-324-5p, miR-330-3p, miR-509 and miR-625) caused a slight reduction in the expression levels (maximum 15%), but none of them reached statistical significance (Figure 3C).

Finally, full-length and truncated *NTRK3* transcripts were quantified by real-time quantitative RT-PCR after miRNA overexpression, using isoform-specific primers. mRNA

levels were not affected by any of the regulating miRNAs indicating that they do not act by destabilizing *NTRK3* transcripts and suggesting that the observed downregulation is achieved through translational repression.

miR-128 overexpression affects the morphology and number of SH-SY5Y cells

After transfection, cells were examined under a phase-contrast microscope to check for possible alterations induced by miRNA overexpression. While in most cases there were no appreciable differences, considerable changes were observed after transfection with miR-128 (Figure 4A): cells acquired rounded bodies with shorter neurites, the overall cell size looked smaller than control cells and the culture confluence appeared to be higher, suggesting an increase in the total number of cells. Given that miR-128 downregulates TR-NTRK3, it was reasonable to speculate that the repression of this variant could be responsible for at least part of the observed effects. Cells were therefore transfected with an siRNA directed against TR-NTRK3, which targets an isoform-specific sequence located within the 3'UTR region and reduces TR-NTRK3 levels by approximately 25% - a degree of repression comparable to that observed with miR-128. Interestingly, the morphology of cells was similar to that described for miR-128 (Figure 4B), supporting the hypothesis that TR-NTRK3 plays a part in the morphological phenotype.

In order to characterize these changes, transfected cells were counted with a Coulter cell and particle counter. The total number of cells in plates transfected with miR-128 was 27% higher than in control plates (Figure 5A). On the other hand, two-parameter forward/side scatter (FSC/SSC) flow cytometry revealed no shift in the population of cells transfected with miR-128 (Figure 5B), indicating that there is no variation in the actual size or cytoplasmic complexity of cells and that the observed alterations are due to other factors, possibly involving a modification of the adhesive and motility properties of the cell.

miR-128 is a brain-enriched miRNA, whose expression has been shown to correlate and increase with neuronal differentiation (18) (19). The expression of miR-128 was analyzed by real-time quantitative RT-PCR in a set of human tissues (adult brain, colon, heart, kidney, liver, lung, ovary, skeletal muscle, spleen, testis, thymus and placenta) as well

as in SH-SY5Y cells at different stages of RA treatment (Figure 6). The analysis confirmed that miR-128 is strongly expressed in brain, and high levels were also detected in skeletal muscle, followed by thymus and kidney. In SH-SY5Y cells, in agreement with the miRNA microarray experiment described before, miR-128 showed low levels of expression, with high crossing point (Ct) values. However, we could detect an increase in miR-128 expression upon RA treatment, which is consistent with the hypothesis that miR-128 contributes to the repression of TR-NTRK3 during RA-mediated differentiation of SH-SY5Y cells (Figure 6B).

Transcriptome analysis of SH-SY5Y cells transfected with miR-128

In order to gain insights into the role of miR-128, the effects of its overexpression were further analyzed using whole genome expression microarrays (Illumina's HumanRef-8 v3.0 beadchips). Considering a fold-change (FC) cutoff of 1.2 and a q-value <5, we could identify a total of 183 deregulated genes after miR-128 overexpression in SH-SY5Y cells - 116 downregulated and 67 upregulated - the maximum FC being ~1.8 for upregulated genes and -2 for downregulated genes. The top ten upregulated and downregulated genes are listed in Tables S1 and S2, respectively. Interestingly, among them are several genes implicated in apoptosis, cell death/survival and proliferation, with a general tendency for those leading to cell death to be downregulated and for those leading to survival to be upregulated. The most striking example is *BCL2* (FC = 1.69), a well-known antiapoptotic gene that inhibits caspase activity. The upregulation of *BCL2* in miR-128-transfected cells could explain the observed increase in cell number, which is consistent with enhanced apoptosis inhibition. Also interesting is *PAIP2* (FC = -2.03), a translational repressor that inhibits the Vascular Endothelial Growth Factor (*VEGF*), a potent mitogen and survival factor with neuroprotective functions in the brain. The deregulation of other genes like *NGFRAP1*, *PRSS3*, *TRAIF3IP2, CREB3L2* (Tables S1 and S2) and of the proapoptotic factors *TXNIP* (Thioredoxin Interacting Protein, FC = -1.38) and *DAP3* (Death Associated Protein 3, FC = -1.36) follows the same trend. Other strongly deregulated genes are involved in vesicular transport, like *DYNC1I1*, and in cytoskeletal organization like *TMSB10*, *TRAPPC4* and *VBP1* (Table S2). This is in accordance with the

observed morphological phenotype and with the involvement of TR-NTRK3 in membrane remodeling, actin reorganization and cell movement through the activation of Arf6-Rac1 signaling (6).

Finally, the 183 genes deregulated upon miR-128 overexpression and the corresponding expression values were uploaded into the Ingenuity Pathway Analysis software, and the program was interrogated about biological functions, canonical pathways and molecular networks that could be affected. We found associations with biological functions such as cell cycle, cancer, neurological disease and cell death (-log(p-value) >3) and with metabolic canonical pathways (-log(p-value) >3). Enzymes implicated in these pathways show a marked predominance of downregulation over upregulation, indicating that metabolic processes may in general be inhibited or impaired in miR-128-transfected cells. The molecular network showing the highest percentage of deregulated genes (Network 1) is shown in Figure S2.

Overexpression of miR-128 upregulates BCL2

The upregulation of *BCL2* (FC = 1.69, third most upregulated gene and one of the central nodes of Network 1) was validated at the protein level by western blotting, which showed an increase of approximately 1.5-fold (Figure 7A) - similar to that observed for the transcript in the microarray experiment. The possible effect of the upregulation of BCL2 on the activation of apoptotic markers, such as caspase-3 and caspase-9, in cells transfected with miR-128 was also analyzed by western blotting. As shown in Figure 7B, the active forms of both caspases are practically undetectable, indicating that the basal level of apoptosis in SH-SY5Y cells is already low; as a consequence, it was not possible to appreciate any decrease in the activation of the two caspases. However, although better known as an antiapoptotic factor, BCL2 also plays a role in the regulation of the cell cycle, which appears to be at least in part separate from its antiapoptotic function. It has been shown that BCL2 has an antiproliferative effect, driving cells into enhanced G0 arrest. Enhanced G0 could possibly explain the increase in cell number associated with miR-128 and is compatible with the general downregulation of metabolic pathways indicated by the microarray analysis. MTT assays performed on SH-SY5Y cells revealed

no change in total mitochondrial activity after transfection with miR-128 despite the larger number of cells, again in agreement with a general reduction of metabolic activity, which is typical of the quiescent state.

Discussion

In this study we explore how miRNAs take part in the regulation of two isoforms of the neurotrophin-3 receptor *NTRK3,* which present non-overlapping 3'UTRs. We demonstrate that the two isoforms are targeted by different sets of miRNAs, providing new evidence for a role of miRNAs in determining the balance among different splice variants of a gene. In fact, we identify 2 miRNAs regulating the full-length isoform of *NTRK3* (miR-24 and miR-151-3p) and four regulating the truncated isoform (miR-128, miR-485-3p, miR-765 and miR-768-5p).

Our approach followed two steps: a first screening of miRNAs with predicted targets in the 3'UTR of either isoform (performed in HeLa cells using luciferase constructs), and the subsequent validation of the ability of luciferase-positive miRNAs to repress endogenous NTRK3 receptors in SH-SY5Y cells. If we compare the results obtained with the two strategies, 6 out of 11 luciferase-positive miRNAs were validated on endogenous NTRK3 (54%). Given that the constitutive expression of the remaining 5 miRNAs is similarly low in both HeLa and SH-SY5Y cell lines, we could rule out the possibility that high levels of endogenous miRNAs in SH-SY5Y cells could affect the results of the validation, making the overexpression of miRNA mimics less effective. Therefore, such rate of false-positives is probably related with the artificial context in which the 3'UTRs are inserted in the luciferase assay strategy. Interestingly, miR-151-3p and miR-185 have partially overlapping target sequences in the full-length isoform and, in a similar way, miR-128, miR-509 and miR-768-5p target the same region of the 3'UTR of the truncated isoform (Figure S3). It is worth of notice that while both miR-128 and miR-509 cause a strong reduction in luciferase activity (30-50%), only miR-128 seems to repress the corresponding protein isoform in SH-SY5Y cells. Intriguingly, miR-128 is expressed in brain, whereas miR-509 is not present in brain but shows a strong expression in kidney

and testis; a similar situation can be observed in the case of miR-151-3p and miR-185 for FL-NTRK3. This suggests that the binding of these miRNAs to their target sequences might be mutually exclusive and that the regulation of *NTRK3* by miRNAs could be coordinated in a tissue specific fashion.

The second part of this study focused on miR-128, one of the miRNAs repressing the truncated isoform. miR-128 is a brain-enriched miRNA, whose expression has been shown to positively correlate with neuronal differentiation (18) (19). Furthermore, accumulation of miR-128 has been detected in the hippocampus of Alzheimer's disease brains (20). Here we show that the overexpression of miR-128 causes morphological changes in SH-SY5Y cells, which resemble those observed using an siRNA specifically directed against TR-NTRK3. In accordance with the morphological changes, as revealed by microarray analysis, miR-128 alters the expression of genes involved in cytoskeletal organization, a process with which the only pathway so far identified that is activated by the truncated isoform is related. These results suggest an involvement of miR-128 in the organization of the cytoskeleton through the regulation of NTRK3, possibly in concert with other miR-128 targets.

Another consequence of the overexpression of miR-128 in SH-SY5Y cells is an increase in cell number, which is accompanied by the deregulation of genes involved in apoptosis, cell death/survival and proliferation, with a remarkable upregulation of BCL2. *BCL2* codes for an outer mitochondrial membrane protein that blocks cytochrome *c* release from mitochondria and inhibits caspase activity, suppressing apoptosis. However, the analysis of caspase-3 and caspase-9 activation revealed that in normal culture conditions there is virtually no apoptosis, making the hypothesis of an increased inhibition of this process unsuitable for our results. It would be therefore interesting to analyze the effects of miR-128 in the presence of apoptosis-promoting agents.

Besides apoptosis, BCL2 also plays a role in the regulation of the cell cycle, having an antiproliferative effect that drives cells into enhanced G0 arrest. This is thought to promote cell survival, especially in unfavorable environments, since quiescent cells are more resistant to killing than proliferating cells (21) and is compatible with the increase

in cell number and general reduction of metabolic activity observed in our analyses. Further investigation is needed to conclusively confirm this hypothesis and to exhaustively evaluate the contribution of *BCL2* to the observed phenotype, as well as that of other proliferation-regulating genes whose expression is altered by miR-128.

Materials and methods

Firefly luciferase constructs

The 3'UTRs of the full-length isoform and truncated isoforms of *NTRK3* were PCR-amplified from BAC CTD-2508H23 with *PfuTurbo*® DNA polymerase (Stratagene, La Jolla, CA), using primers containing an *Xba*I restriction site at the 5' end: sense 5'–acacactctagagtctgccccaaagaggtgta–3' and antisense 5'–acacactctagaccaaactgccttacagggttt–3' for the full-length isoform and sense 5'-acacactctagaaataagccttcccggacatt–3' and antisense 5'-acacactctagatgcaaaatttccaaataagagg–3' for the truncated isoform. PCR fragments of 334 and 2110 bp respectively were purified, *Xba*I-digested and cloned into an *Xba*I site downstream of the firefly luciferase reporter gene in the pGL4.13 vector (Promega Corporation, Madison, WI) as previously described (13).

Cell culture and transfection

HeLa and SH-SY5Y neuroblastoma cells were maintained in Dulbecco's Modified Eagle's Medium supplemented with 10% fetal bovine serum, 2 mM L-glutamine, 100 units/ml penicillin and 100 µg/ml Streptomycin (GIBCO ™, Invitrogen). Transfection with small RNAs was performed with Lipofectamine 2000 (Invitrogen) and optimized using a fluorescein-labeled double-stranded RNA oligomer. Small RNAs that mimic endogenous mature miRNAs and the related negative controls (miRIDIAN™ miRNA Mimics and miRIDIAN™ miRNA Mimics Negative Controls #2 and #4) were obtained from Dharmacon Inc.

Luciferase activity assay

HeLa cells were seeded at 1.3 x 10^4 cells/well in 96-well plates and cotransfected 24 h later with the Firefly reporter constructs described above or the empty pGL4.13 vector

(10-24 ng), the Renilla reporter plasmid pGL4.75 (3 ng) and the appropriate miRNA mimic (30 nM). The activity of Firefly and Renilla luciferases was determined 24 h after transfection using the Dual-Glo™ Luciferase Assay System (Promega). Relative reporter activity was obtained by normalization to the Renilla luciferase activity. To correct for vector-dependent unspecific effects, each relative reporter activity was normalized to the empty vector cotransfected with the corresponding miRNA. Results were then compared to the mean of the two negative controls. Each experiment was done in triplicate and at least three independent experiments were performed for each miRNA. Data are reported as means ± S.E. Statistical significance was determined using Student's *t* test ($p<0.05$).

Western blotting

For the analysis of NTRK3 expression during RA-induced differentiation, SH-SY5Y cells were treated with 10 µM *all-trans*-RA and harvested at time 0 (untreated), day 3, day 6 and day 10 of RA treatment. For TF-NTRK3, undifferentiated SH-SY5Y cells were plated at 2 x 10^5 cells per well in 6-well plates and were transfected 24 h later with 100nM miRNA mimic; 72 h after transfection cells were lysed and analyzed by western blotting. For FL-NTRK3, SH-SY5Y cells were differentiated with 10 µM *all-trans*-RA, plated in 6-well plates at a concentration of 3 x 10^5 cells per well at day 3 of RA treatment, transfected with 100 nM miRNA mimic at day 4 of RA treatment and analyzed 72 h after transfection (day 7 of RA treatment).

For protein extraction cells were rapidly rinsed with ice-cold PBS and solubilized with RIPA buffer and 1x protease inhibitors (Complete Mini, EDTA-free, Roche). Cells were then scraped off, incubated on ice for 15 min and centrifuged at 12000 rpm for 15 min. Samples were resolved in NuPAGE® 4-12% Bis-Tris polycrylamide gels using the NuPAGE® MES SDS Running Buffer and transferred to nitrocellulose membranes using the iBlot™ Dry Blotting System (Invitrogen). Before blotting, gels were equilibrated in 100 ml equilibration buffer (2x NuPAGE® Transfer Buffer, 10% methanol and 1:1000 NuPAGE® Antioxidant) for 20 min at room temperature.

Immunodetection was performed using the ODYSSEY® infrared imaging system (LI-COR® Biosciences) Membranes were blocked in ODYSSEY® blocking buffer for 1 h at RT and

incubated with the appropriate primary antibody (Ab) for 1 h at RT (anti-TrkC goat polyclonal IgG, Upstate, # 07-226; anti-Bcl-2 (N-19), Santa Cruz, # sc-492; anti-Caspase-3, Millipore, # 06-735; anti-Caspase-9, Cell Signaling, # 9508).

As a loading control, membranes were simultaneously incubated with an anti-GAPDH Ab (anti-GAPDH mouse monoclonal Ab, Chemicon, # MAB374). Blots were subsequently probed with the appropriate fluorophore-labeled secondary antibodies (LI-COR® Biosciences) and scanned on an ODYSSEY® infrared scanner. Fluorescent bands were quantified using the ODYSSEY® software. Each experiment was repeated four times; data are reported as means ± S.E and statistical significance was determined using Student's *t* test.

For the analysis of ERK phosphorylation, differentiated cells were transfected as explained above and 72 h later cells were rinsed twice, incubated for 5 h with serum-free medium and acutely stimulated for 10 min with 30 ng/mL neurotrophin-3 (Alomone Labs). Cells were then lysed in 50 mM Tris-HCl pH 7.4, 50 mM NaF, 10 mM sodium pyrophosphate and 0.005% Triton X-100. After transfer, membranes were incubated with an anti-phopsho-ERK (Anti-MAP kinase, activated mouse monoclonal Ab, Sigma, #M9692) and an anti-pan-ERK Ab (Anti-p44/42 MAP kinase rabbit polyclonal Ab, Cell Signaling, # 9102).

miRNA expression analysis using custom microarrays (Agilent, 11k)

Total RNA was extracted from HeLa and SH-SY5Y cells using the RNeasy® Mini Kit (QIAGEN). miRNA expression was analyzed using custom 11k oligonucleotide microarrays (Agilent Technologies) including probes for 325 known human miRNAs (Sanger miRBase release 7.1). Negative controls included probes for 20 *B. subtilis*-specific sequences and 20 non-human sequences. Two µg of total RNA were labeled with the Hy5™ or Hy3™ fluorescent labels using the miRCURY™ LNA microRNA Labelling kit (Exiqon), following the manufacturer's instructions. Pairs of labeled samples were hybridized to dual-channel microarrays for 40 h at 55°C using Agilent hybridization reagents. Microarray images were quantified using the GenePix 6.0 (Axon) software. Only spots with signal intensities twice above the local background and not saturated

were considered reliable and used for subsequent analysis. Statistical analyses were performed using the MMARGE tool.

Whole-genome expression analysis using beadchip microarrays

Total RNA samples obtained from four independent experiments (SH-SY5Y cells transfected with miR-128 and the related negative controls) were analyzed on HumanRef-8 BeadChips from Illumina, which target 24,500 well-annotated RefSeq transcripts. Starting from 200 ng total RNA, biotin-labeled cRNA was synthesized and hybridized according to the manufacturer's instructions. Data were analyzed using the Array File Maker (AFM) 4.0 software package.

Real-time quantitative RT-PCR

DNase-treated RNA extracted from transfected cells was retrotranscribed with the SuperScript™ III First-Strand Synthesis System for RT-PCR (Invitrogen). RNA levels of the full-length and truncated isoforms of *NTRK3* were analyzed by real-time quantitative RT-PCR using SYBR Green I (Roche), following the manufacturer's instructions. Specific primers were designed for the two *NTRK3* isoforms (FL-NTRK3 forward 5'-cagccatggttccaactctc-3', FL-NTRK3 reverse 5'-ccagcatgacatcgtacacc-3', TR-NTRK3 forward 5'-tccagagtggggaagtgtct-3', TR-NTRK3 reverse 5'-ccatggttaagaggcttgga-3') and for *GAPDH* (forward 5'-ctgacttcaacagcgacacc-3', reverse 5'-ccctgttgctgtagccaaat-3'), which was used as a housekeeping gene. The expression levels of miR-128 in different human tissues and in SH-SY5Y cells were analyzed using TaqMan® MicroRNA Assay, following the manufacturer's instructions.

Cell counting

Cells cultured on 6-well plates and transfected with miR-128, the TR-NTRK3-specific siRNA (sense 5-gagucuaugccuuuggcaatt-3', antisense 5'-uugccaaaggcauagacuctt-3', Gene Link) and the related negative controls were trypsinized and resuspended in 1 mL DMEM; 100 µL of each sample were then diluted in 10 mL of Coulter Isoton II diluent (Beclman Coulter) and counted using a Z2™ Series Coulter Particle Count and Size Analyzer (Beckman Coulter).

Flow cytometry

SH-SY5Y cells transfected with miR-128 and the related negative controls were trypsinized and resuspended in PBS 72 h after transfection. Samples were examined using a BD FACSCanto Flow Cytometer and the BD FACSDiva software (BD Biosciences). A two-parameter forward/side scatter (FSC/SSC) analysis was performed. 5000 events were recorded in each analysis.

Computational methods

Four web-based miRNA target prediction methods were used: miRanda (www.microrna.org, 2005 release; (22)), TargetScan (www.targetscan.org, releases 2.0, 2.1, 3.0 and 3.1; (9)), PicTar (http://pictar.mdc-berlin.de; (23)). and RNA hybrid (http://bibiserv.techfak.uni-bielefeld.de/rnahybrid). Genomic coordinates are according to the human assembly release of March 2006 (H. sapiens hg18). Sequence analysis was performed using the 4peaks software (http://mekentosj.com/4peaks/) and Multalin (http://prodes.toulouse.inra.fr/multalin/multalin.html). Primers were designed using the Primer3 software (http://frodo.wi.mit.edu/cgi-bin/primer3/primer3.cgi). Pathway analysis was performed with the Ingenuity Pathway Analysis Software (IPA) version 6.3 (www.ingenuity.com).

Acknowledgements

We would like to thank S. Porta, M. Bañez and E. Miñones for their support and advice. This work was supported by the "Instituto Carlos III and Fondo de Investigaciones Sanitarias" (CIBER-CB06/02/0058, FIS/ISCIII:P1052565, ISCIII:GO3/184), the "Departament d'Universitats Innovació i Empresa, Generalitat de Catalunya" (2005SGR00008), the Spanish Ministry of Science and Innovation (SAF2008-00357) and the European Union Sixth Framework Programme Integrated Project SIROCCO (Grant LSHG-CT-2006-037900). M. G. was funded by an FPU fellowship. M.M.G is a recipient of a FIS fellowship (FI05/0006). Y.E.P was supported by the "Ramón y Cajal" Program

References

1. Lu B., Pang P. T.,Woo N. H. (2005) The yin and yang of neurotrophin action *Nat Rev Neurosci* 6, 603-614.
2. Barbacid M. (1994) The Trk family of neurotrophin receptors *J Neurobiol* 25, 1386-1403.
3. Tsoulfas P., Stephens R. M., Kaplan D. R.,Parada L. F. (1996) TrkC isoforms with inserts in the kinase domain show impaired signaling responses *J Biol Chem* 271, 5691-5697.
4. Shelton D. L. *et al.* (1995) Human trks: molecular cloning, tissue distribution, and expression of extracellular domain immunoadhesins *J Neurosci* 15, 477-491.
5. Beltaifa S. *et al.* (2005) Discordant changes in cortical TrkC mRNA and protein during the human lifespan *Eur J Neurosci* 21, 2433-2444.
6. Esteban P. F. *et al.* (2006) A kinase-deficient TrkC receptor isoform activates Arf6-Rac1 signaling through the scaffold protein tamalin *J Cell Biol* 173, 291-299.
7. Kosik K. S. (2006) The neuronal microRNA system *Nat Rev Neurosci* 7, 911-920.
8. Bartel D. P. (2004) MicroRNAs: genomics, biogenesis, mechanism, and function *Cell* 116, 281-297.
9. Lewis B. P., Burge C. B.,Bartel D. P. (2005) Conserved seed pairing, often flanked by adenosines, indicates that thousands of human genes are microRNA targets *Cell* 120, 15-20.
10. Stark A., Brennecke J., Bushati N., Russell R. B.,Cohen S. M. (2005) Animal MicroRNAs confer robustness to gene expression and have a significant impact on 3'UTR evolution *Cell* 123, 1133-1146.
11. Laneve P. *et al.* (2007) The interplay between microRNAs and the neurotrophin receptor tropomyosin-related kinase C controls proliferation of human neuroblastoma cells *Proc Natl Acad Sci U S A* 104, 7957-7962.
12. Pietrzykowski A. Z. *et al.* (2008) Posttranscriptional regulation of BK channel splice variant stability by miR-9 underlies neuroadaptation to alcohol *Neuron* 59, 274-287.

13. Muinos-Gimeno M. *et al.* (2009) Allele variants in functional MicroRNA target sites of the neurotrophin-3 receptor gene (NTRK3) as susceptibility factors for anxiety disorders *Hum Mutat* 30**,** 1062-1071.

14. Rehmsmeier M., Steffen P., Hochsmann M.,Giegerich R. (2004) Fast and effective prediction of microRNA/target duplexes *Rna* 10**,** 1507-1517.

15. Kaplan D. R., Matsumoto K., Lucarelli E.,Thiele C. J. (1993) Induction of TrkB by retinoic acid mediates biologic responsiveness to BDNF and differentiation of human neuroblastoma cells. Eukaryotic Signal Transduction Group *Neuron* 11**,** 321-331.

16. Encinas M., Iglesias M., Llecha N.,Comella J. X. (1999) Extracellular-regulated kinases and phosphatidylinositol 3-kinase are involved in brain-derived neurotrophic factor-mediated survival and neuritogenesis of the neuroblastoma cell line SH-SY5Y *J Neurochem* 73**,** 1409-1421.

17. Kaplan D. R.,Miller F. D. (1997) Signal transduction by the neurotrophin receptors *Curr Opin Cell Biol* 9**,** 213-221.

18. Sempere L. F. *et al.* (2004) Expression profiling of mammalian microRNAs uncovers a subset of brain-expressed microRNAs with possible roles in murine and human neuronal differentiation *Genome Biol* 5**,** R13.

19. Smirnova L. *et al.* (2005) Regulation of miRNA expression during neural cell specification *Eur J Neurosci* 21**,** 1469-1477.

20. Lukiw W. J. (2007) Micro-RNA speciation in fetal, adult and Alzheimer's disease hippocampus *Neuroreport* 18**,** 297-300.

21. Zinkel S., Gross A.,Yang E. (2006) BCL2 family in DNA damage and cell cycle control *Cell Death Differ* 13**,** 1351-1359.

22. John B. *et al.* (2004) Human MicroRNA targets *PLoS Biol* 2**,** e363.

23. Krek A. *et al.* (2005) Combinatorial microRNA target predictions *Nat Genet* 37**,** 495-500.

Figure Legends

Figure 1. Luciferase assay results: FL-NTRK3 and TR-NTRK3 regulation by miRNAs. HeLa cells were cotransfected with pGL4.13-FL or pGL4.13-TR and the indicated miRNA mimics (X axis, C=Control) at a concentration of 10 nM. Luciferase activities were measured 24 h after transfection; firefly luciferase activity was normalized to Renilla luciferase activity, and results from at least three independent experiments are presented as means ± SE. Only statistical significant results are shown ($p<0.05$).

Figure 2. Characterization of the expression of NTRK3 in RA-differentiated SH-SY5Y cells. **A.** Immunoblotting of the full-length and truncated isoforms of NTRK3 in SH-SY5Y cells treated with RA for the indicated times. GAPDH was used as a loading control. **B.** Protein levels were quantified by densitometry and are reported as arbitrary units relative to untreated cells.

Figure 3. Western blot analysis of NTRK3 levels in cells transfected with the indicated miRNA mimics. **A.** RA-differentiated cells were transfected with mimics (100 nM) targeting FL-NTRK3; NTRK3 levels were analyzed 72h after transfection and are reported relative to control cells. **B.** Differentiated SH-SY5Y cells were transfected and serum-starved 72 h after transfection for 5 h. After acute stimulation with NT3 (30ng/mL) ERK1/2 phosphorylation was analyzed. Basal phosphorylation in unstimulated cells is also shown (-NT3). **C.** Undifferentiated cells were transfected with mimics targeting TR-NTRK3 and NTRK3 levels were analyzed 72h after transfection. Data from four independent experiments are presented as means ± SE (*$p<0.05$).

Figure 4. Phase-contrast micrographs of undifferentiated SH-SY5Y cells. **A**. Cells were transfected with a non-targeting control (A,B,C) and with miR-128 (D,E,F); morphological changes (rounded bodies, shorter neurites and smaller cell size) are observed in cells transfected with miR-128. **B.** Cells were transfected with a non-silencing control (A,B) and with an siRNA directed against the truncated isoform of NTRK3 (C,D). The morphology of cells transfected with the siRNA resembles that of cells transfected with miR-128. In both cases small RNAs were used at 100 nM concentration, and images were taken 72 h after transfection. Images are shown at increasing magnification.

Figure 5. Characterization of cells transfected with miR-128 and a TR-NTRK3-specific siRNA. **A.** Cell counting: undifferentiated SH-SY5Y cells were transfected with miR-128, the TR-NTRK3-specific siRNA and the corresponding negative controls. 72 h after transfection cells were trypsinized, resuspended in the appropriate buffer and passed through a Coulter cell counter. Results from three independent experiments are reported as means ± SE, relative to the respective controls (*$p<0.05$). **B.** Two-parameter forward/side scatter (FSC/SSC) flow cytometry analysis of SH-SY5Y cells transfected with a non-targeting control and miR-128. No profile shift was observed.

Figure 6. Real-time RT-PCR analysis of miR-128 expression (TaqMan® miRNA assays), using small nucleolar RNA RNU58B as an endogenous control for normalization. **A.** Results for different human tissues are reported as -ΔCts between each miRNA and the endogenous control. ΔCt=0 corresponds to the expression level of RNU58B in each sample. **B.** Results for SH-SY5Y cells during RA-mediated differentiation are reported as -ΔΔCts relative to untreated cells (reference sample: -RA).

Figure 7. Western blot analysis of BCL2 levels and caspase activation in undifferentiated cells transfected with miR-128. **A.** BCL2 levels show a ~1.5-fold increase relative to the negative control (three independent experiments, *$p<0.05$). **B.** Analysis of caspase-3 (green) and caspase-9 (red) activation reveals that neither caspase is present in its activated form (~17 KDa and ~35 KDa respectively).

Figure S1. Minimum free energy hybridization of miR-24 (green) and the 3'UTR of *FL-NTRK3* (red, nucleotide positions 1726-1758 downstream of the stop codon) according to the RNAhybrid software.

Figure S2. Molecular network showing the highest percentage of genes altered by the overexpression of miR-128, according to the Ingenuity Pathway Analysis software.

Figure S3. Relative location of the target sequences of validated miRNAs in the 3'UTRs of FL-NTRK3 and TR-NTRK3. An enlargement of overlapping target sites is shown, where sequence pairing to the 3'UTRs (lines) and miRNA seed regions (underlines) is indicated.

Table 1. Tested miRNAs and their predicted target sites in the two isoforms of NTRK3, according to miRanda (M), TargetScan (TS) and PicTar (P). The number of predicted sites is indicated in brackets. NP = not predicted for either isoform.

miRNA	FL-NTRK3	TR-NTRK3
hsa-let-7b		M(1)
hsa-let-7e	M(1)	
hsa-miR-1	NP	NP
hsa-miR-9		TS(1)
hsa-miR-10a	NP	NP
hsa-miR-15a	NP	NP
hsa-miR-16	NP	NP
hsa-miR-17-3p	NP	NP
hsa-miR-17-5p		TS(2)
hsa-miR-18a		M(1)
hsa-miR-18a*	NP	NP
hsa-miR-20a		TS(2)
hsa-miR-24		TS(1)
hsa-miR-30e-3p	NP	NP
hsa-miR-30e-5p	NP	NP
hsa-miR-93		TS(1)
hsa-miR-103	NP	NP
hsa-miR-106a		TS(2)
hsa-miR-106b		TS(2)
hsa-miR-107	NP	NP
hsa-miR-125a		M(2)
hsa-miR-125b		M(2)
hsa-miR-128		M(1), TS(1)
hsa-miR-133a		M(1)
hsa-miR-141	NP	NP
hsa-miR-149		M(1), TS(1)
hsa-miR-151-3p	M(1)	
hsa-miR-182	NP	NP
hsa-miR-185	M(1), PT(1)	
hsa-miR-188		M(1)
hsa-miR-198		M(1)
hsa-miR-200a	NP	NP
hsa-miR-200b	NP	NP
hsa-miR-204		M(1)
hsa-miR-206	NP	NP
hsa-miR-211		M(1)
hsa-miR-221	NP	NP
hsa-miR-296		M(1)
hsa-miR-324-5p		M(1), TS(1)
hsa-miR-326		M(2)
hsa-miR-330-3p		M(2), TS(2)
hsa-miR-331		M(2)
hsa-miR-340		M(1)
hsa-miR-345		TS(1)
hsa-miR-374	NP	NP
hsa-miR-384		TS(1)
hsa-miR-412		M(1)
hsa-miR-422a		M(1)
hsa-miR-449		M(1)
hsa-miR-485-3p		TS(1)
hsa-miR-509		M(1), TS(2)
hsa-miR-617		TS(1)
hsa-miR-625		TS(2)
hsa-miR-765		TS(1)
hsa-miR-768-5p		TS(1)

Table 2. miRNA expression in HeLa cells, untreated SH-SY5Y cells and differentiated SH-SY5Y cells (6 days RA exposure)

miRNA	HeLa	SH-SY5Y	SH-SY5Y (RA 6d)
miR-9	-	-	-
miR-24	++++	+++++	+++++
miR-124	-	-	-
miR-125a	+	+++	++++
miR-125b	++	+++	++++
miR-128	-	+	+
miR-151-3p	-	-	-
miR-185	-	-	-
miR-324-5p	-	-	-
miR-330	-	+	+
miR-485-3p	-	+	+
miR-509	-	-	-

Expression levels are indicated relative to background signal as a scale from undetectable (<5-fold relative to background) to strongly expressed (>100-fold)

Figure 1

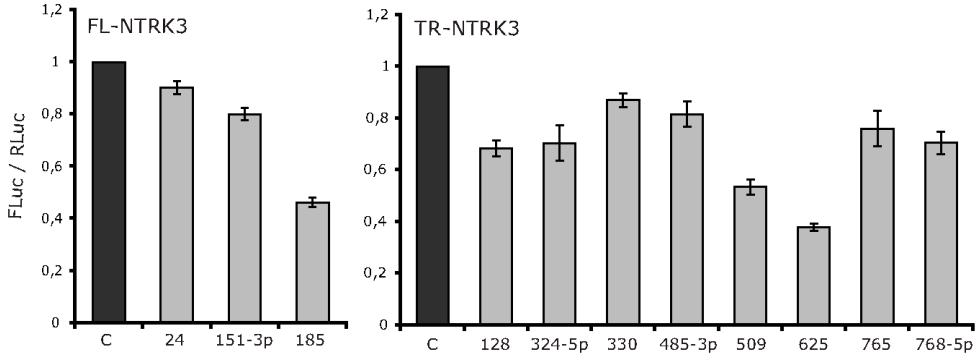


Figure 2

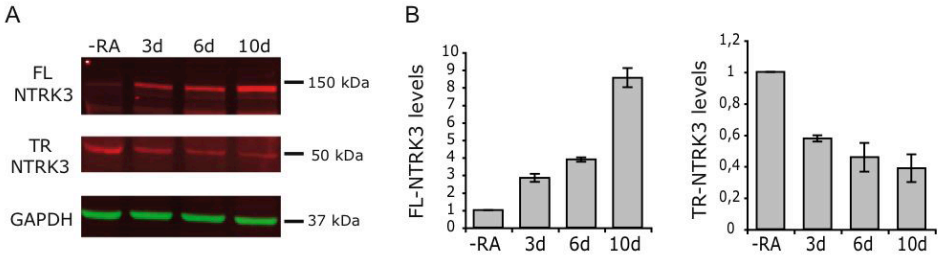


Figure 3

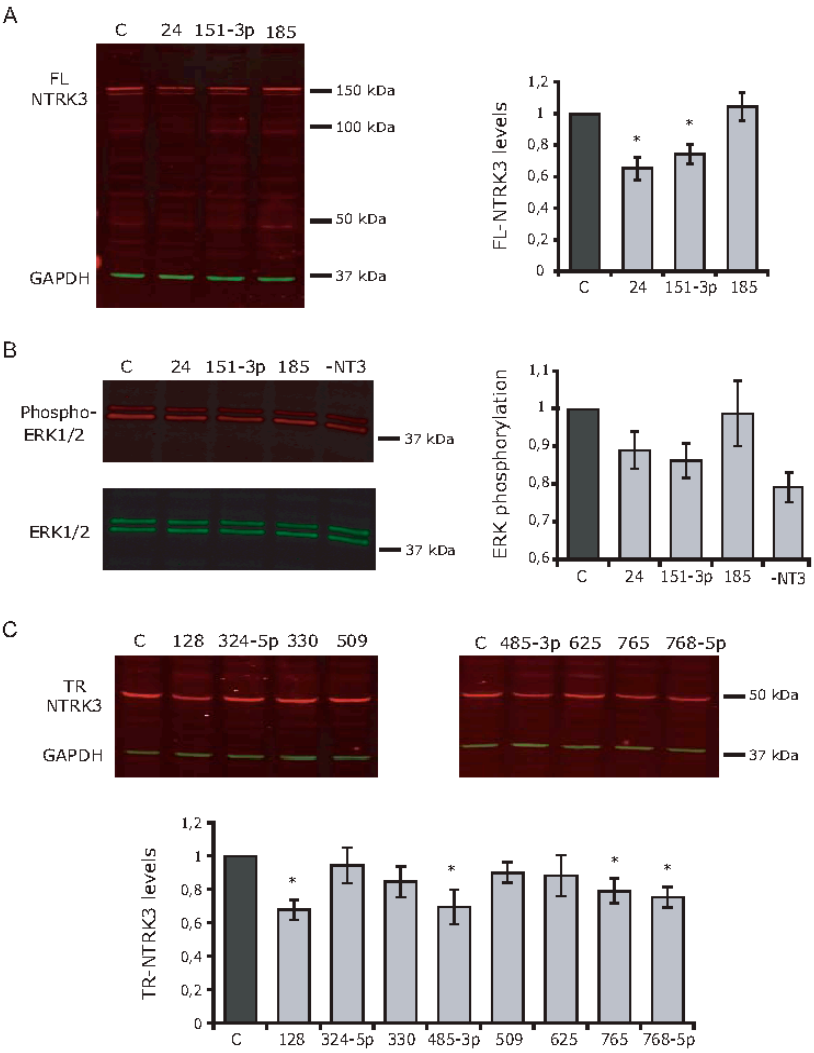


Figure 4

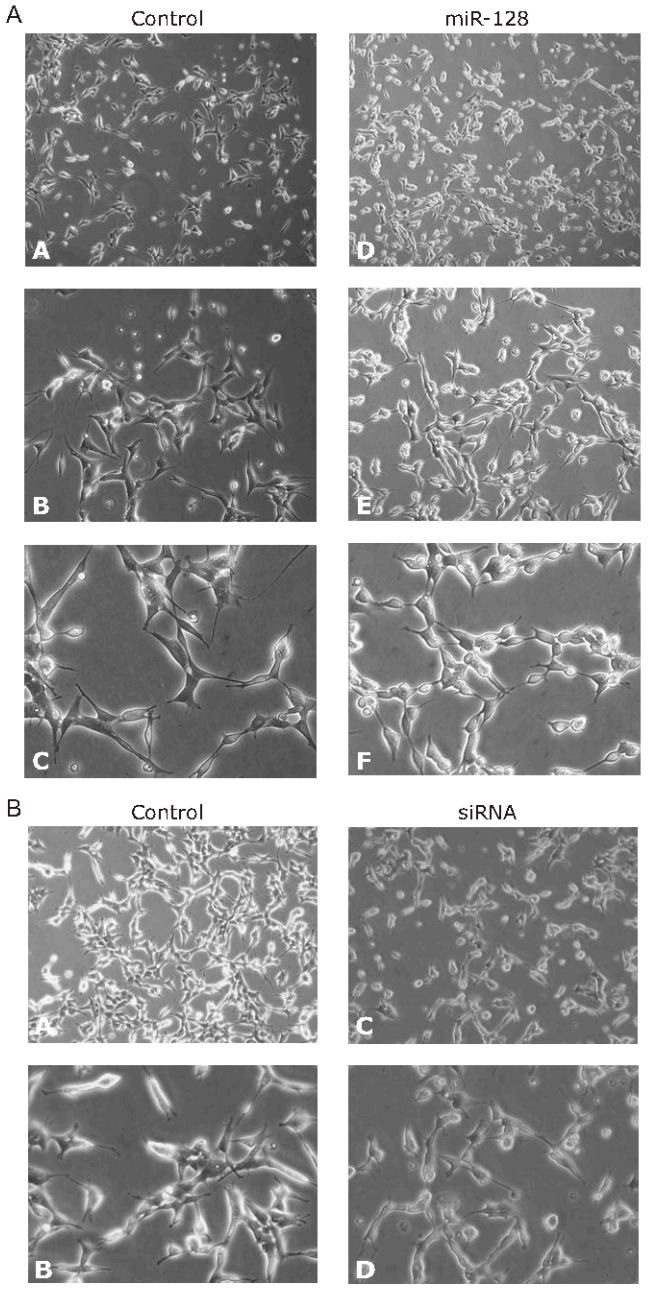


Figure 5

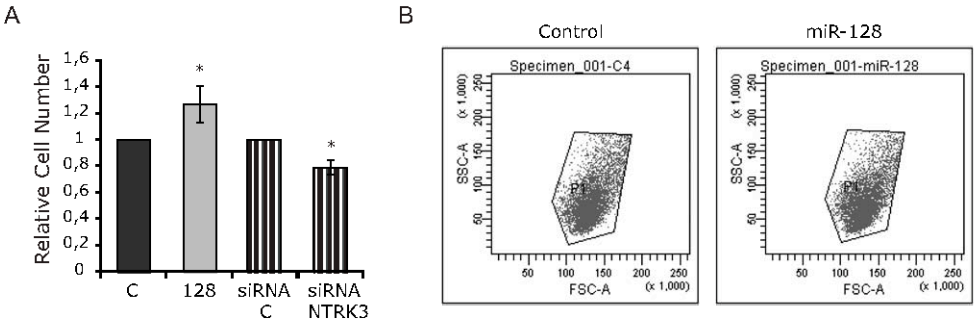


Figure 6

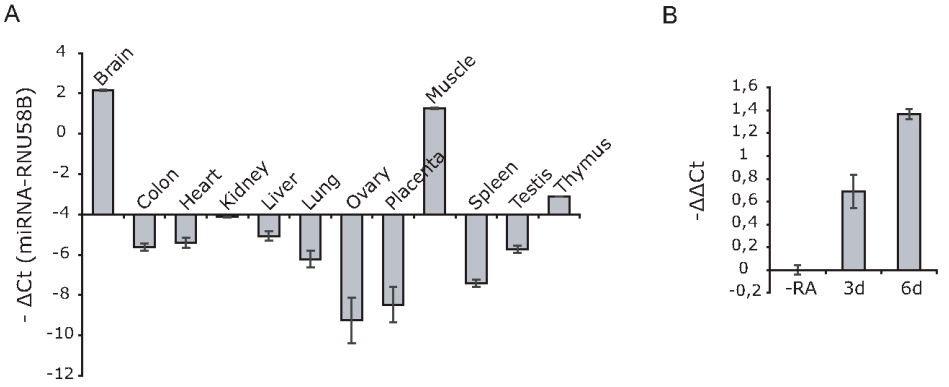


Figure 7