Finding Genes for Schizophrenia

KAROLINA ÅBERG
Dissertation presented at Uppsala University to be publicly examined in Rudbecksalen, Rudbecklaboratoriet, Dag Hammarskjölds väg 20, Uppsala, Thursday, September 22, 2005 at 13.15 for the degree of Doctor of Philosophy. The examination will be conducted in English.

Abstract

Schizophrenia is one of our most common psychiatric diseases. It severely affects all aspects of psychological functions and results in loss of contact with reality. No cure exists and the treatments available today produce only partial relief for disease symptoms. The aim of this work is to better understand the etiology of schizophrenia by identification of candidate genes and gene pathways involved in the development of the disease.

In a preliminarily study, the effects of medication and genetic factors were investigated in a candidate gene, serotonin 2C receptor. This study distinguished pharmacological effects, caused by neuroleptics, and/or genetic effects, caused by unique polymorphisms, from other effects responsible for mRNA expression changes on candidate genes.

The core of the thesis describes a new candidate gene for schizophrenia, the quaking homolog, KH domain RNA binding (mouse) or QKI, located on chromosome 6q26-q27. The identification of QKI is supported by previous linkage studies, current association studies and mRNA expression studies using three different sample sets. The investigated samples included a 12-generation pedigree with 16 distantly related schizophrenic cases and their parents, 176 unrelated nuclear families with at least one affected child in each family and human brain autopsies from 55 schizophrenic cases and from 55 controls. Indirect evidence showing involvement of QKI in myelin regulation of central nervous system is presented. Myelin plays an important role in development of normal brains and disruption of QKI might lead to schizophrenia symptoms.

In a forth sample set, including extended pedigrees originated from a geographically isolated area above the Arctic Circle, in northeast Sweden, two additional schizophrenia susceptibility loci were identified, 2q13 and 5q21. Both these regions have previously been highlighted as potential schizophrenia loci in several other investigations, including a large Finnish study. This suggests common schizophrenia susceptibility loci for Nordic populations.

A pilot investigation including a genome wide haplotype analysis is presented. This statistical strategy could be further developed and applied to the artic Swedish families, including analysis of 900 microsatellites and 10,000 SNPs.

These findings will facilitate the understanding of the schizophrenia etiology and may lead to development of more efficient treatments for patients that suffer from schizophrenia.

Keywords: psychiatric genetics, QKI, large pedigree, haplotype investigation, mRNAexpression, genetic linkage, myelin, HTR2C

Karolina Åberg, Department of Evolution, Genomics and Systematics, Evolutionary Biology, Norbyv. 18D, Uppsala University, SE-75236 Uppsala, Sweden

© Karolina Åberg 2005

ISSN 1651-6214
ISBN 91-554-6308-8
urn:nbn:se:uu:diva-5894 (http://urn.kb.se/resolve?urn=nbn:se:uu:diva-5894)
“Science is organized knowledge. Wisdom is organized life.”
- Immanuel Kant (1724-1804)
List of Papers


IV Åberg, K., Saetre, P. and Jazin, E. Human QKI, a Potential Regulator of mRNA Expression in Myelin Specific Genes *(Submitted)*

V Åberg, K., Lindholm, E., Saetre, P., Wetterberg, L., Pettersson, U. and Jazin, E. Genome Wide Investigation of an Isolated Schizophrenia Population Using a Dense Map of Microsatellites and SNPs in Combination. *(Manuscript)*

Reprints/preprints were made with permission from the publisher
Contents

Introduction ................................................................................................... 11
Schizophrenia – a severe disease that affects the function of the brain ..... 11
  Prevalence differences ................................................................. 11
  Standardized diagnostic methods ........................................... 12
  Medical treatment ...................................................................... 13
  Genetic risk factors ................................................................. 13
Proceedings in schizophrenia research ....................................................... 15
  Neuropathology ...................................................................... 15
  Candidate genes found by linkage and association studies ......... 15
  Candidate genes/pathways found by mRNA investigations ..... 16
Hypotheses in concordance with genetic findings ........................................ 16
  Involvement of neurotransmitter systems in schizophrenia ..... 16
  Involvement of myelin in schizophrenia ..................................... 18
Basic concepts of statistical genetics ........................................................ 19
  Genetic linkage ...................................................................... 19
  Linkage disequilibrium ........................................................... 21
  Linkage analysis versus association analyses ......................... 22
  Haplotype investigations ........................................................ 22
Research Aims .......................................................................................... 26
  Aims for included papers ........................................................ 26
Present Investigations ................................................................................... 27
  Paper I – Serotonin Receptor 2C (HTR2C) and Schizophrenia:
  Examination of Possible Medication and Genetic Influences on
  Expression Levels ...................................................................... 27
  Paper II – Reconstruction of ancestral haplotypes in a 12-generation
  schizophrenia pedigree ......................................................... 28
  Paper III – Human QKI, a New Candidate Gene for Schizophrenia
  Involved in Myelination .......................................................... 30
  Paper IV – Human QKI, a Potential Regulator of mRNA Expression in
  Myelin Specific Genes Involved in Schizophrenia ...................... 32
  Paper V – Genome Wide Investigation of an Isolated Schizophrenia
  Population Using Dense Maps of Microsatellites and SNPs in
  Combination ............................................................................. 33
Future Perspectives and Conclusions ..........................................................35
  Evaluation of QKI as a candidate gene for schizophrenia ....................35
  Novel genome-wide investigations .....................................................35
  General conclusion .............................................................................36

Summary in Swedish ....................................................................................37
  Gener som orsakar schizofreni .............................................................37
  Schizofreni – en allvarlig sjukdom som påverkar hjärnfunktionen .... 37
  Framsteg inom schizofreni forskningen ..............................................37
  Forskningsmål .....................................................................................38
  Artikel I – Serotonin Receptor 2C (HTR2C) och Schizofreni: 
  Undersökning av Möjlig Medicineringspåverkan och Genetisk 
  Påverkan av Uttrycksnivåer .................................................................39
  Artikel II – Rekonstruktion av Nedärvda Haplotyper i ett Tolv 
  Generationers Schizofrenisläktträd ......................................................39
  Artikel III – Humana QKI, en Ny Kandidatgen för Schizofreni 
  Involverad i Myelinering .....................................................................40
  Artikel IV – Humana QKI, en Potentiell Regulator av mRNA 
  Uttryck i Myelin Specifika Gener Involverade i Schizofreni ..............40
  Artikel V – Genom-omfattande Undersökning av en Isolerad 
  Schizofrenipopulation Genom att Använda Mikrosatelliter och 
  SNPs i Kombination ............................................................................40
  Slutsatser ..............................................................................................41

Acknowledgements .......................................................................................42

References .....................................................................................................44
## Abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Full Form</th>
</tr>
</thead>
<tbody>
<tr>
<td>cDNA</td>
<td>complementary DNA</td>
</tr>
<tr>
<td>CNS</td>
<td>central nervous system</td>
</tr>
<tr>
<td>DNA</td>
<td>deoxyribonucleic acid</td>
</tr>
<tr>
<td>Mb</td>
<td>Mega base pair</td>
</tr>
<tr>
<td>mRNA</td>
<td>messenger RNA</td>
</tr>
<tr>
<td>PCR</td>
<td>polymerase chain reaction</td>
</tr>
<tr>
<td>RNA</td>
<td>ribonucleic acid</td>
</tr>
<tr>
<td>RT</td>
<td>reverse transcriptase</td>
</tr>
<tr>
<td>SNP</td>
<td>single nucleotide polymorphism</td>
</tr>
</tbody>
</table>
Introduction

Schizophrenia is one of our most common psychiatric diseases. It severely affects all aspects of psychological functions resulting in loss of contact with reality. No cure exists and the treatments available today produce only partial relief for disease symptoms. To be able to develop more efficient treatments for patients suffering from schizophrenia it is essential to better understand the etiology of the disease.

Schizophrenia – a severe disease that affects the function of the brain

Schizophrenia, which was defined by Eugen Bleuler in 1911, is a common mental disorder with a lifetime prevalence of 1% in the population worldwide (Bear et al. 2001). The symptoms of the disease can vary substantially between different patients. However, all schizophrenic patients exhibit psychotic episodes, characterized by lost contact with reality. Psychotic episodes are characterized by symptoms such as delusions (a false belief based on incorrect inference about external reality), hallucinations (a false sensory perception in the absence of an external stimulus), disorganized speech and bizarre behavior. These symptoms are referred to as positive symptoms. Other common schizophrenia symptoms are social withdrawal, absence of normal emotional feelings and expressions, lack of energy, lack of interest and motivation and absence of purpose of life. These symptoms are referred to as negative symptoms (Gelder et al. 1996). Schizophrenia usually has its onset between the age of 15-25 for males and the age of 15-30 for females but variations from early childhood up to 50 years of age have been observed (Hafner et al. 1998). The disease is normally persistent for the rest of life resulting in life long treatment (Bear et al. 2001).

Prevalence differences

Schizophrenia affects the population worldwide (Bear et al. 2001). Deviations from the average prevalence (1 %) have been reported from Slovenia, western Ireland and northern Sweden (Böök 1953; Lindholm 2001) while
lower prevalence have been detected among the Hutterites in the United States (Eaton and Weil 1955).

Differences in prevalence may have several explanations. Environmental factors are likely to be of significance. Cultural tolerance and social structure may be differently well suited for schizophrenic individuals and their families, which might result in different rates of migration from different populations. Other environmental factors such as bacterial and virus infection, climate and level of education are also hypothesized to be of importance for the development of the disease (Mortensen et al. 1999; Chotai et al. 2000).

Causative or preventing factors (environmental or genetic) could be conserved and accumulated during time in populations founded by a few ancestors and/or in populations that have been isolated for a large number of generations. If such factors were absent from the founders and/or not introduced to the isolate the prevalence rate may be decreased.

Possibly of largest importance for the prevalence in different geographical areas and for prevalence in different studies, is the fact that schizophrenia is difficult to define and diagnose correctly. Not until the 1970s standardized diagnostic methods such as CATEGO (Wing et al. 1974) and the Feighner criteria (Feighner et al. 1972) were introduced.

Standardized diagnostic methods

Psychiatric diseases are diagnosed by evaluation of occurrence and duration of certain symptoms.

Table 1. Diagnostic criteria for schizophrenia according to DSMIV and ICD10, respectively. Adapted from Gelder 1996 (Gelder et al. 1996).

<table>
<thead>
<tr>
<th>DSMIV</th>
<th>ICD10</th>
</tr>
</thead>
<tbody>
<tr>
<td>At least two of the symptoms in &quot;a&quot; must be present for a minimum of one month. Additionally, &quot;b&quot; must be present for six months and criteria &quot;c-e&quot; must be fulfilled.</td>
<td></td>
</tr>
<tr>
<td>a) delusions, hallucinations, bizarre behavior and negative symptoms.</td>
<td></td>
</tr>
<tr>
<td>b) occupational or social dysfunction</td>
<td></td>
</tr>
<tr>
<td>c) schizoaffective or mood disorder exclusion</td>
<td></td>
</tr>
<tr>
<td>d) disturbance must not be due to medication or drug abuse</td>
<td></td>
</tr>
<tr>
<td>e) if a patient has a pervasive development disorder, prominent delusions or hallucinations must be present for one month.</td>
<td></td>
</tr>
<tr>
<td>Minimum of one clear symptoms, or at least two if less clear cut, listed in &quot;a-d&quot; or at least two of the symptoms in &quot;e-i&quot; should be present for one month or more.</td>
<td></td>
</tr>
<tr>
<td>a) thought echo, thought insertion or withdrawal, and thought broadcasting</td>
<td></td>
</tr>
<tr>
<td>b) delusion of control or passivity</td>
<td></td>
</tr>
<tr>
<td>c) hallucinatory voices</td>
<td></td>
</tr>
<tr>
<td>d) persistent delusions</td>
<td></td>
</tr>
<tr>
<td>e) significant and consistent change of personal behavior</td>
<td></td>
</tr>
<tr>
<td>f) negative symptoms that are not due to depression or neuroleptic treatment</td>
<td></td>
</tr>
<tr>
<td>g) persistent hallucinations accompanied by deletions or by persistent over-valued ideas consistently occurring for weeks or months</td>
<td></td>
</tr>
<tr>
<td>h) incoherent or irrelevant speech</td>
<td></td>
</tr>
<tr>
<td>i) catatonic behavior</td>
<td></td>
</tr>
</tbody>
</table>

Today the most frequently applied standardized diagnostic methods are the Diagnostic and Statistical Manual of Mental Disorder, fourth edition
The main difference is the duration time of symptoms before schizophrenia can be diagnosed (Gelder et al. 1996). More details about the diagnostic criteria for DSMIV and ICD10, respectively are presented in table 1.

Medical treatment
Schizophrenia is normally treated with a combination of psychotherapy and social adjustments as well as with drug administration (Wetterberg 1997). The neuroleptics used to treat schizophrenic patients are usually divided in to two main classes, typical neuroleptics and atypical neuroleptics. Typical neuroleptics were introduced to the market in the early 1950s (Kandel 2000). They act by blocking the action of the neurotransmitter dopamine at the receptor level, mainly the D2 receptor. The lower level of dopamine stimulation mainly reduces positive symptoms. However, decreased dopamine level results in increased acetylcholine levels that may cause side effects called extra pyramidal symptoms (EPS) (Kandel 2000). Some frequently used typical neuroleptics are chlorpromazine, promazine, haloperidol, thioridazine, stelazine, trifluoperazine, thiothixene, and sulpiride (Kapur and Remington 2001). In 1988 a new class of neuroleptics, the atypical neuroleptics, became available. These neuroleptics e.g. clozapine and risperidone, have a more heterogeneous binding pattern to the dopamine receptor subtypes than the typical drugs, but the exact binding pattern is not completely known. Atypical neuroleptics have effect on positive symptoms as well as on negative symptoms (Kane et al. 1988; Kandel 2000; Kapur and Remington 2001). The atypical neuroleptics do not cause as severe EPS as typical neuroleptics do (Kane et al. 1988; Kandel 2000; Kapur and Remington 2001). A minority of the individuals diagnosed with schizophrenia responds well to treatment and recover after the first incident. Most patients is only partly helped or do not respond at all to any of the treatments available today (Wetterberg 1997).

Genetic risk factors
The primary cause to the development of schizophrenia is not known, but there is convincing evidence from family, twin and adoption studies, that there is a substantial genetic contribution to the disease (McGuffin et al. 1995; Cannon et al. 1998).
The closer related an individual is to an affected patient, the higher the risk of developing the disease (Figure 1). This figure clearly shows that the risk of developing the disease is dependent on at least one genetic factor. However, the fact that two monozygotic twins have a concordance of only 48% shows that also environmental factors affect the development of the disease (Gottesman 1991). Although schizophrenia is more common in some families/populations than in others, no simple Mendelian inheritance pattern is observed. This suggests that more than one gene is involved in the development of the disease (Riley and McGuffin 2000). It is possible that the causative mutations in these genes are frequent in the population in general and therefore is hard to identify. A likely hypothesis is that several causative mutations in combination with one or several environmental factors together trigger the development of schizophrenia.
Proceedings in schizophrenia research

Neuropathology

Schizophrenia is associated with different macroscopic and histological findings such as ventricular enlargement, reduction of brain volume, cortical thickness, cerebral asymmetry and aberrantly located or clustered neurons (Harrison and Weinberger 2005). However, there is no common pathological marker associated with the majority of schizophrenia cases. The magnitude of observed changes are usually small, the results are inconsistent and some times contradictory (Harrison and Weinberger 2005).

Candidate genes found by linkage and association studies

During the last decades a large number of linkage and association studies have been performed. Results from these studies have suggested susceptibility loci for schizophrenia on most chromosomes, e.g. 1, 2, 5, 6, 7, 8, 9, 10, 13, 15, 18, 22 and X (Riley and McGuffin 2000). The most strongly supported regions include 1q21-q22, 1q42, 5q21-q33, 6p24-p22, 6q16-q26, 8p22-p21, 10p15-p11, 13q22-q34 and 22q11-q12.

Recently, evidence for involvement of several candidate genes for the development of schizophrenia, located in some of these regions, have been reported from different scientists around the world, including regulator of G-protein signaling-4 gene (RGS4) (1q21) (Brzustowicz et al. 2000; Mirnics et al. 2000; Chowdari et al. 2002; Williams et al. 2004), neuregulin 1 gene (NRG1) (8p12) (Stefansson et al. 2002; Stefansson et al. 2003), D-amino-acid oxidase activator (DAOA earlier referred to as G72) (13q22) (Chumakov et al. 2002; Schumacher et al. 2004), and catechol-O-methyltransferase (COMT) (22q11) (Shifman et al. 2002; Glatt et al. 2003).

Chromosome 6 harbors interesting linkage regions on both chromosomal arms. Several groups have reported linkage to 6pter-6p23 (Schwab et al. 1995; Straub et al. 1995; Wang et al. 1995; Lindholm et al. 1999; Schwab et al. 2000). Lately, the dysbindin gene (DTNBP1) located at 6p22.3 was suggested as a potential candidate gene (Straub et al. 2002a). Linkage finding have also been reported on chromosome 6q13-6q26 by several groups including our own (Cao et al. 1997; Kaufmann et al. 1998; Martinez et al. 1999; Levinson et al. 2000; Lindholm et al. 2001; Lerer et al. 2003; Duan et al. 2004). Duan et al have proposed the trace amine receptor 4 gene (TRAR4 or TA4), located at 6q23.2, as a susceptibility gene for schizophrenia (Duan et al. 2004). In an investigation included in this thesis, paper III, we proposed the quaking homolog, KH domain RNA binding (mouse) (QKI) (6q26-q27) as an additional new candidate for schizophrenia located in this region (Aberg et al. 2005).
Candidate genes/pathways found by mRNA investigations

During the last years genome-wide expression studies, simultaneously measuring mRNA expression of thousands of genes, have been performed using microarray technique (Duggan et al. 1999; Lipshutz et al. 1999). These studies mainly show decreased mRNA expression of myelination-related genes and transcripts encoding proteins that regulate pre-synaptic function suggesting oligodendrocyte and synaptic dysfunction involved in the development of schizophrenia (Mirnics et al. 2000; Hakak et al. 2001; Tkachev et al. 2003; Aston et al. 2004; Aston et al. 2005). Several microarray studies have been successfully replicated with real-time RT-PCR (Tkachev et al. 2003; Aston et al. 2004; Aston et al. 2005) and in situ hybridization (Mirnics et al. 2001). In general, the mRNA expression studies investigating synaptic dysfunction and myelination related genes have been less contradictory than other schizophrenia research findings such as linkage and association studies. For example, the MAG gene consistently showed significant down-regulation in schizophrenic cases compared to controls in the majority of investigations. However, with some exceptions (Castensson et al. 2003), most mRNA expression investigations are performed on very small sample sets, including only about ten to fifteen samples in each category (Mirnics et al. 2000; Hakak et al. 2001; Mirnics et al. 2001; Tkachev et al. 2003; Aston et al. 2005).

Hypotheses in concordance with genetic findings

Involvement of neurotransmitter systems in schizophrenia

The neurotransmitter systems interact with each other in several different ways (Laruelle et al. 2003; Meltzer et al. 2003; Coyle 2004). Therefore, it is hard to conclude whether a detected dysfunction is a potential cause to the development of schizophrenia or whether it is a secondary effect of a related dysfunction.

More than forty years ago Carlsson et al pointed out the importance of the dopamine (DA) system (Carlsson and Lindqvist 1963). Based on the correlation between clinical doses of antipsychotic drugs and their potency to block DA D2 receptors, it was proposed that hyperactivity of DA transmission in sub-cortical regions was responsible for the positive symptoms observed in schizophrenia (Carlsson and Lindqvist 1963).

Negative symptoms observed in the disease might also arise from alterations in DA transmission. Modified prefrontal cortical functions in schizophrenic patients have been observed in functional brain imaging investigations. In prefrontal cortex another DA receptor, the DA D1 receptor, is important for normal function (Weinberger 1987; Kandel 2000). These findings
have generated the hypothesis that insufficient D1 receptor dependent DA transmission is responsible for negative symptoms (Laruelle et al. 2003).

Several studies have shown genetic association between schizophrenia and *DA D2 receptor* (Arinami et al. 1994; Dubertret et al. 2001) and *DA D3 receptor* (Crocq et al. 1992; Mant et al. 1994) but conflicting results have also been observed (Kaneshima et al. 1997; Tallerico et al. 1999).

At least one of the candidate genes proposed for schizophrenia is modified by DA signaling. That is *regulator of G-protein signaling 4* (*RGS4*) which is involved in neuronal differentiation (Geurts et al. 2002).

In addition to imbalance in the DA system, other neurotransmitter systems are proposed to be involved in the development of the disease, including serotonin (5-HT), glutamate and gamma-aminobutyric acid (GABA) neurotransmitters (Carlsson 1988).

Serotonin, a regulator of mood and emotional behavior, exerts a tonic control on the DA system by stimulating 5-HTR2C receptors (Meltzer et al. 2003). Serotonin has also the opposite effect on the DA system through the action of 5-HTR2A receptor antagonists that enhance DA release (Di Matteo et al. 2002). These findings suggest that imbalanced serotonin might lead to disruption of the DA system. In addition, it is hypothesized that atypical neuroleptics have a better effect on negative symptoms and cause less extra pyramidal side effects than typical neuroleptics because of their relatively high affinity for 5-HT receptors compared to their affinity to DA D2 receptors (Meltzer et al. 2003). Furthermore, mRNA expression of the 5-HTR2C receptor has been reported as down-regulated in schizophrenic patients (Castensson et al. 2003).

Dysfunction of glutamate transmission has been proposed to involve the N-methyl-D-aspartate (NMDA) receptor. This receptor is involved in the regulation of neuronal migration, neuronal differentiation and elaboration of synaptic spines (Coyle 2004). These are factors that are of potential importance for the development of the disease. In fact, it is suggested that deregulation of DA is a secondary effect of a disrupted NMDA function (Laruelle et al. 2003). Some schizophrenia candidate genes are involved with glutamate transmission. For example, genetic linkage studies have identified the DAOA gene as a potential candidate for schizophrenia (Chumakov et al. 2002; Schumacher et al. 2004), and this gene is involved in the activation of NMDA receptors.

Gamma-aminobutyric acid has been reported as differentially expressed in schizophrenic cases (Coyle 2004). The GABAergic system mediates synaptic inhibition in CNS and it is an interesting candidate pathway involved in the development of the disease.

Finally, two schizophrenia candidate genes named *dysbindin* (*DTNBP1*) and *neuregulin 1* (*NRG1*) are both involved in neurotransmission of NMDA and GABA (Husi et al. 2000; Inoue and Okabe 2003).
Involvement of myelin in schizophrenia

During normal post-adolescent development and maturation of prefrontal and association areas of the brain, a white matter expansion occurs in concert with a grey matter reduction. The white matter increase is due to myelination, a process that continues into middle age (Yakolev and Lecours 1967). If normal myelination is disrupted, normal adult brain function is likely to become impaired. In such a circumstance, the cortical gray matter pruning of neuronal connectivity would not be compensated by the increase connectivity provided by myelination and could result in a brain dysfunction manifested as symptoms of schizophrenia (Maier and Ron 1996; Bartzokis 2002). In support of this hypothesis, several studies have reported decreased white matter volumes associated to schizophrenia, including abnormalities in oligodendroglia and myelin in the disease pathology (Maier and Ron 1996; Sanfilipo et al. 2000; Wible et al. 2001; Bartzokis 2002; Davis et al. 2003). Furthermore, mRNA expression levels of several genes involved in myelination were found to be down-regulated in schizophrenia, including 2',3'-cyclic nucleotidase (CNP), myelin and lymphocyte protein (MAL), gelsolin (GSN), transferrin (TF) neuregulin receptor Her 3 (ErbB3) (Hakak et al. 2001), Myelin-associated glycoprotein (MAG) (Hakak et al. 2001; Tkachev et al. 2003; Aston et al. 2005), Proteolipid Protein (PLP1) and Myelin Basic Protein (MBP) (Tkachev et al. 2003).

The schizophrenia candidate gene, NRG1, binds to ErbB3. The NRG1-ErbB3 complex leads to activation of downstream signaling pathways, which in turn might affect myelination. Another schizophrenia candidate gene, RGS4, also interacts with ErbB3 (Thaminy et al. 2003). The schizophrenia candidate gene QKI regulates, in mice, splicing and mRNA expression of myelination specific genes such as MAG, MBP and PLP (Wu et al. 2002). As described in in this thesis, paper III, the relative mRNA expression of two QKI splice-variants is differently expressed in schizophrenic cases compared with controls (Aberg et al. 2005).

Recently, it was suggested that also neurotransmitters are involved, direct or indirect, in myelination. The GABAergic and glutaminergic systems are involved in neurodevelopment. Both are expressed in oligodendrocyte precursor-cells and are able to depolarize myelinating cells (Butt and Tutton 1992). The NMDA receptors are involved in myelination by regulating the migration that is an essential step prior to myelination (Wang et al. 1996). Dopamine may be involved in the myelination more indirect by regulating RGS4 which interact with to the myelination related ErbB3.
Basic concepts of statistical genetics

Results presented in the papers included in this thesis have mainly been obtained by the below described statistical concepts. The statistical strategies are modified to uniquely fit different types of sample sets and have been implemented in to user-friendly software.

Genetic linkage

Genetic linkage investigation is probably the most commonly used strategy to identify a genetic region that is likely to house a functional mutation of importance for the studied phenotype. The statistical calculations can be conducted in several different ways. However, all calculations have to be based on either a single-point or a multi-point strategy. In addition, the use of a non-parametric or a parametric analysis has to be selected. All genetic linkage investigations are based on detailed information about a certain number of genetic markers.

Genetic linkage is defined as physical association of genes located on the same chromosome, or non-independence among alleles at more than one locus (Liu 1998). In other words, linkage analysis evaluates the probability of recombination between two studied loci in a pedigree. Therefore, the recombination fraction is an essential parameter when calculating genetic linkage.

The recombination fraction ($\theta$) is defined as the number of recombinations that occur in a pedigree divided by the number of meiotic events in the same pedigree. The overall effect of double crossovers always produce 50% recombinants, independently whether it includes two, three or four strands. Therefore, only recombination fraction between 0 and 0.5 is meaningful (Strachan and Read 1999).

To calculate genetic linkage in a population the recombination fraction for each pedigree is needed. However, it is possible to measure genetic linkage by calculating the overall likelihood of the pedigree.

The ratio of the overall likelihood and the likelihood of no linkage gives the odds of linkage. If the loci are linked the observed $\theta < 0.5$ and if the loci are unlinked $\theta = 0.5$. The logarithm of the odds is called the LOD-score ($Z$) (Strachan and Read 1999; Lange 2002).

$$Z(\theta) = \log_{10} \left[ \frac{L(\theta)}{L(0.5)} \right]$$

When this is applied to a two-generation pedigree with unknown genotypes from the grandparents the LOD-score is calculated as followed:

$$Z(\theta) = \log_{10} \left[ \frac{1}{2} \left[ \theta^R (1-\theta)^{n-R} / (1/2)^n \right] + \frac{1}{2} \left[ \theta^{n-R} (1-\theta)^R / (1/2)^n \right] \right]$$
where \( n \) is the total number of children in the pedigree and \( R \) is the number of recombinants.

Positive LOD-score give evidence in favor for linkage and negative LOD give evidence against linkage. As a rule of thumb \( Z = 3.0 \) is the threshold for accepting linkage on an autosome chromosome and \( 2.3 \) is the threshold for significance on the x-chromosome, if \( \alpha = 0.05 \). Linkage can be rejected if \( Z < -2.0 \) (Strachan and Read 1999).

**Genetic markers**

The human genome consists of approximately three billion base pairs (Venter et al. 2001). Today, we do not have a technique that allows for screening of every single base in the genome for each individual, in a time and cost efficient way. Instead we study genetic markers, known polymorphic sites, to estimate how often recombination events occur. The most frequently used genetic markers are microsatellites. Microsatellites are short nucleotide units of 2-5 bases repeated 5-50 times within a locus. Lately, single nucleotide polymorphisms (SNP) have become more common as genetic markers. SNPs are any polymorphic variation at a single nucleotide. As genetic markers the most frequently used SNPs are substitutions. In linkage studies, investigating e.g. schizophrenic individuals and healthy individuals, one also refers to the phenotype as a marker to be included in the statistical analysis.

**Single-point and multi-point linkage analysis**

In single-point linkage analysis (also called two-point linkage analysis) one calculates the recombination events between the phenotype and one genetic marker at a time. In multi-point linkage analysis, one searches for recombination events between many markers simultaneously. Therefore multi-point analysis is usually more powerful than single-point. However, the marker order is crucial in multi-point analysis; small difference in the internal order between markers may have large effect on the analysis (Strachan and Read 1999).

**Parametric and non-parametric linkage analyses**

Non-parametric linkage (NPL) analysis, also known as allele sharing statistics, is very beneficial when the mode of inheritance is unknown, as it does not require specification of a specific inheritance model. If a marker is linked to a disease locus it is expected that few marker alleles descend from the pedigree founders. This statistics is based only on affected individuals, but information from unaffected pedigree members may be used to extract information about inheritance vectors. Parametric linkage analysis performed with a true model is more powerful than non-parametric strategies. However, for most complex diseases the true model is unknown. When working with a parametric strategy it is needed to state an inheritance model by specifying
the level of penetrance as well as the phenocopy rate. For non-parametric strategies as well as for parametric strategies allele frequencies for all markers and for the trait locus need to be specified.

**Linkage disequilibrium**

Linkage disequilibrium (LD) has traditionally been used to further investigate genetic regions found by genetic linkage. Similar to genetic linkage investigations, LD calculations require detailed information about a certain number of genetic markers. Linkage disequilibrium is the statistical association of sequence variants at different positions along chromosomes as they occur in gametes. In other words, two alleles at different loci that appear together in the gametes more frequently than expected by chance are in LD. The significance of LD between two loci is measured by contingency table tests. Typically, Fisher’s exact test is used to determine if two alleles are segregating independently of each other or not. Thus, the null hypothesis of linkage equilibrium for two bi-allelic loci \( (1 = \text{allele one}, 2 = \text{allele two}, A = \text{locus one and } B = \text{locus two}) \) can be written (Nordborg and Tavare 2002):

\[
H_0: p_{A1} p_{B1} = p_{A1B1}
\]

where \( p_{A1} \) is the frequency of the first allele in loci one, \( p_{B1} \) the first allele in locus two, and similar for the second alleles. Thus \( p_{A1B1} \) is the frequency of the observed \( A1B1 \) haplotype.

One way to measure LD is to calculate \( D \).

\[
D = p_{A1B1} - p_{A1} p_{B1}
\]

This measure is strongly dependent on allele frequency and is therefore most often transformed to the more stable LD calculation, \( |D'| \).

\[
|D'| = \frac{|D|}{D_{\text{max}}}
\]

where \( D_{\text{max}} \) is the smaller of \( p_{A1B2} \) and \( p_{A2B1} \) or \( p_{A1B1} \) and \( p_{A2B2} \). If only three of the four possible haplotypes are present in the population, LD is “complete” and \( |D'| = 1 \), whereas a total shuffling among the marker alleles would result in lack of LD and \( |D'| = 0 \). The benefit of calculating \( D' \) when measuring LD is that it is scaled to remove allele frequency effects. Another stable way to measure the MAGnitude of LD is the \( r^2 \) that is the square of the correlation coefficient between the A and B loci.

\[
r^2 = \frac{D^2}{p_{A1} p_{A2} p_{B1} p_{B2}}
\]
The calculated $r^2$ is typically lower than the $D'$ for any chromosomal distance. Both those measures are symmetric, in the sense that it does not matter which locus is the disease locus and which is the marker locus (Nordborg and Tavare 2002; Weiss and Clark 2002).

Linkage analysis versus association analyses

Both linkage analysis and association analysis are used to search for genetic regions that have influence on different phenotypes. Traditionally, the first step in a genetic study has been to calculate genetic linkage based on genotypes from approximately 400 microsatellite markers. Second, the most promising regions that showed genetic linkage are further investigated by performing association analysis on a denser marker map.

With today’s well-equipped genome centers denser marker maps can be used in an initial stage. This allows for linkage analysis as well as association analysis to be conducted simultaneously.

There is a main difference between linkage analysis and association analysis that has to be kept in mind when working with both methods in one population. Linkage analysis highlights regions (loci) that are connected to the disease due to less recombination events between the linkage region and the actual locus that regulates the phenotype, not the actual phenotype. Association analysis, on the other hand, pinpoints specific alleles within a locus that segregate together with the actual phenotype. This results in the fact that linkage and association merge if the investigated population consists of one single large family (Strachan and Read 1999).

Haplotype investigations

Haplotypes are defined as a set of alleles found at loci on a single chromosome (Strachan and Read 1999). If no recombination events have occurred between the alleles in two or more loci during the meiosis these alleles will be inherited as a unit, a non-recombinant haplotype.

When the measure of LD between two loci that appears on the same chromosome is larger than zero, the observed haplotype frequencies are skewed from the haplotype frequencies expected from independent segregation. The expected haplotype frequencies ($h_1, h_2, h_3, h_4$) can be calculated as below:

\[
P(h_1) = p_{A_1} \cdot p_{B_1}
\]
\[
P(h_2) = p_{A_2} \cdot p_{B_1}
\]
\[
P(h_3) = p_{A_1} \cdot p_{B_2}
\]
\[
P(h_4) = p_{A_2} \cdot p_{B_2}
\]
were \( p_{A1} \) and \( p_{A2} \) are the allele frequencies for locus A and \( p_{B1} \) and \( p_{B2} \) are the allele frequencies for locus B.

To be able to correctly construct haplotypes for a genotyped individual it is necessary to know on which of the two chromosomes the allele occurs. That is the phase of the alleles. The phase shows from which parent the allele is inherited. It is preferred to genotype the parents. However, even though the individual of interest and its parents are fully genotyped it might be impossible to decide the phase of the alleles. In this case, siblings or other closely related relatives such as grandparents or aunts and uncles can be genotyped and the information used to determine the phase of the individual of interest. In the case when one or two of the parents are not available for genotyping, it might still be possible to decide the phase of the alleles, if several other closely related relatives are genotyped e.g. larger sib ships.

There are several software that simulate or estimate haplotypes e.g. Simwalk2 (Sobel and Lange 1996), GENHUNTER (Kruglyak et al. 1996) and UNPHASE (Dudbridge 2003). When working with such software, one has to keep in mind that they present the most likely haplotype for each individual in the pedigree, according to the specific algorithm used in each program, which is not necessarily the true inherited haplotype. Both Simwalk2 and GENHUNTER simulate the haplotypes based on several parameters e.g. genotypes, pedigree information, allele frequencies and phase information, while UNPHASE only considers genotypes and allele frequencies when estimating haplotypes.

Traditionally, haplotypes have been used to narrow-down a genetic region that has been suggested by linkage and association studies. For a dominant, monogenic, completely penetrant disease one would expect that the affected individuals share a region that is not shared by healthy individuals (Figure 2A). For an incomplete penetrant disease, the haplotype pattern will be more complex. If the penetrance is 60% one would expect that 40% of the individuals that have inherited the causative haplotype will not show the disease phenotype (Figure 2B).
Figure 2. Haplotypes spanning 13 markers. Shared region inherited from affected parent to child is drawn as gray vertical bars for 14 individuals. Filled squares/circles indicate a monogenic and dominant phenotype of interest. Double bars show location for causative allele. A) Completely penetrant phenotype. B) Incomplete penetrant phenotype. Arrows highlight individuals that are genetic carriers, they have the causative allele but they do not express the phenotype.

For a disease such as schizophrenia, which most likely is polygenic, with incomplete penetrance, and possibly is controlled by several different dominance patterns and/or by imprinting, one would expect to see an extremely complex haplotype pattern (Figure 3). In addition, for heterogeneous diseases, like schizophrenia, it is possible that a certain number of individuals are phenocopies, meaning that they show the phenotype but do not share a common haplotype.
Figure 3. Haplotypes spanning 13 markers. Shared regions, inherited from parents to child, are drawn as gray vertical bars for 14 individuals. Filled squares/circles indicate a digenic, incomplete penetrant phenotype of interest. The phenotype is caused by a dominant allele in this haplotype and a recessive allele (X) in another gene that together show a co-dominant inheritance pattern. Arrows highlight individuals that are genetic carriers.

For polygenic diseases with complex inheritance patterns it is needed to reduce the complexity to be able to perform the above described haplotype comparisons. One way is to include affected individuals only. By doing this, the penetrance parameter is removed. Another way to reduce the complexity is to include individuals from an isolated population and/or from one single large family.
Research Aims

The general aim of this thesis is to better understand the etiology of schizophrenia by identify candidate genes and gene pathways involved in the development of the disease.

Aims for included papers

I To evaluate whether differences in mRNA expression levels of HTR2C between cases and controls are influenced by neuroleptic treatment and/or are associated to investigated polymorphisms in HTR2C promoter region.

II To identify candidate haplotypes inherited identical by descent in 19 related schizophrenic cases and to examine whether these haplotypes are equally common among unrelated unaffected individuals.

III To fine-map a schizophrenia susceptibility locus on chromosome 6q in a large pedigree and to investigate whether similar findings can be observed in 176 unrelated nuclear families, as well as to investigate whether the QKI-gene located in this region is differently expressed in brain autopsies from schizophrenic cases and controls.

IV To investigate whether there is correlation between mRNA expression levels of three myelin-specific genes and whether this correlation is affected by the QKI-gene in human brain autopsies from schizophrenic cases and controls.

V To search for susceptibility locus/loci for schizophrenia in an isolated population by using dense maps of microsatellites and SNPs.
Present Investigations

Paper I – Serotonin Receptor 2C (HTR2C) and Schizophrenia: Examination of Possible Medication and Genetic Influences on Expression Levels

Due to the involvement of serotonin receptor 2C (HTR2C) in regulation of dopamine activity as well as a previously observed decrease of HTR2C mRNA expression in prefrontal cortex in schizophrenic cases, (Castensson et al. 2003) we have further investigated mRNA variations in relation to neuroleptic treatment and promoter polymorphisms in 55 schizophrenic cases and 55 controls without psychiatric diagnoses.

Methods

Polymorphisms in the promoter region were genotyped using dynamic allele-specific hybridization (DASH) (Howell et al. 1999). Sequencing was used to verify the SNPs found by DASH and to search for new polymorphisms within the candidate region. We measured mRNA expression of HTR2C in brain autopsies from prefrontal cortex in the 55 cases and the 55 controls with real-time RT-PCR. We normalized the real-time RT-PCR expression data from HTR2C with respect to reference genes and removed systematic variation due to age, sex, brain-bank and post-mortem interval (Castensson et al. 2003; Sundberg et al. 2005). To test whether mRNA expression was affected by neuroleptic treatment we compared expression levels between three schizophrenic subgroups (atypical neuroleptic treated patients (n = 7), typical neuroleptic treated patients (n = 19), untreated patients (n = 11) and controls. We used a two-way ANOVA to test whether any mRNA differences were associated to the polymorphisms in the promoter region.

Results

When the patients were subdivided according to medication received, no significant difference was observed between the three subgroups (p-value
0.4). On the other hand, all three groups were significantly different from controls. The subgroup that received typical neuroleptics showed a 1.3-fold decrease (p-value 0.05), the untreated patients showed a 1.6-fold decrease (p-value 0.01) and the patients that had received atypical neuroleptics showed a 2.0-fold decrease (p-value 0.002) compared to controls.

There was no evidence for association between any investigated promoter polymorphism and schizophrenia. Neither, there were any evidence for association between expression differences and promoter polymorphisms.

Discussion
In this investigation we subdivided the investigated schizophrenic cases in three subgroups according to neuroleptic medication received. All three groups were significantly down regulated compared to controls. The down-regulation observed in untreated schizophrenic cases suggests that the mRNA expression difference of HTR2C is not an effect of short-term medication but rather an effect of the disease itself. We observed a tendency of a larger down-regulation in patients treated with atypical neuroleptics and a slightly smaller down-regulation in patients treated with typical neuroleptics. However, no significant difference was observed between the three groups.

No association was observed between HTR2C mRNA expression and promoter polymorphisms and/or schizophrenia. Thus, we conclude that the investigated polymorphisms had no clear effect on transcription in neither patients nor controls. However, the possibility still remains that there are other genetic variants that affect expression levels of the receptor.

Paper II – Reconstruction of ancestral haplotypes in a 12-generation schizophrenia pedigree
In this investigation we search for haplotypes that are likely to be inherited identical by decent (IBD) in schizophrenic patients from a large family. We reasoned that IBD haplotypes, shared by multiple cases, in extended pedigrees with high frequency of schizophrenia are likely to harbor schizophrenia susceptibility loci. To evaluate whether the haplotypes that were observed as IBD are associated to schizophrenia we estimated the haplotype frequencies in additional 43 schizophrenia cases and 46 controls.

Methods
We used 371 previously genotyped microsatellite markers (Lindholm et al. 2001) and constructed haplotypes by hand using the minimal number of recombination principle for each pair of adjacent markers in 19 related
schizophrenic cases. If a minimum certainty-score of three was achieved, corresponding to the sum of scores from haplotypes with fully known phase (score 1/each) and/or haplotypes with uncertain phase (score 0.5/each) and/or haplotypes with uncertain alleles (score 0.25/each), from affected individuals located in different branches in the pedigree, the region was selected for additional investigation. An additional marker located between the original two, was genotyped. We estimated haplotype frequencies for 43 affected cases and 46 controls with pmPlus from the EH program (Zhao et al. 2000).

Results

In this investigation we observed 77 candidate haplotypes of two microsatellite markers that had a minimum certainty-score of three. Apart from the previously reported region on chromosome 6q25, eleven genomic regions with a certainty-score of three or more remained when an additional marker was genotyped between the previous two. The regions were located on chromosome 5q21-23, 8p22-21, 9p23-22, 10p15, 11q24-25, 12p13, 13q12-13, 15q22-24, 17q24-25, 21q21 and 22q12-13. These regions are likely to be inherited IBD in affected individuals from a large pedigree. The haplotypes observed in the large family were significantly more frequent in the 43 cases than in the 46 controls in the regions on chromosome 8p, 9p, 13q and 17q.

Discussion

Based on previous investigations (Ewald et al. 1999) and correcting for the number of markers included we believe this study has power to detect ancestral linkage disequilibrium (LD) blocks that are larger than 10 cM shared by more than two affected individuals in the large family. However, smaller LD regions may not be detected. In fact only two observed regions, 5q21-23 and 6q25, were approximately as large as 10 cM suggesting that LD blocks of this size are rare in this large pedigree. The haplotypes found in the large family on chromosome 8p, 9p, 13q and 17q were more frequent in affected individuals than in controls. Two of those regions, 8p and 13q, overlap with the five genomic regions that are located in, or close to, regions previously reported as susceptible loci for schizophrenia. Apart from the susceptibility loci on chromosome 6q this investigation gives support for at least two previously reported susceptibility loci for schizophrenia (8p and 13q) and two new loci (9p and 17q).
Paper III – Human QKI, a New Candidate Gene for Schizophrenia Involved in Myelination

In this study we further investigate the significance of the schizophrenia susceptibility loci on chromosome 6q25 (Lindholm et al. 2001), also mentioned in paper II in this thesis, and found a new candidate gene for schizophrenia. To evaluate the impact of the gene the region was investigated further in nuclear families with at least one affected child. In addition we measured mRNA expression of the candidate gene.

At the time of the start of this investigation, the human genome sequence was released. When comparing the genetic marker map, previously used to locate the 6q loci, with the new physical marker maps from the Celera database (Venter et al. 2001) and the Human Genome Browser (Kent et al. 2002; Karolchik et al. 2003) we realized that the order of the markers in this region was slightly rearranged. We also observed some differences in the marker order between the two new physical marker maps.

Methods

In an initial descriptive stage we performed genotyping of 27 microsatellite markers and 15 SNP in 16 schizophrenic individuals from an extended pedigree, previously used in paper II, and one or two of their parents. All markers were located in a region on chromosome 6q25-q27 spanning 10.7 Mb. We constructed haplotypes with Simwalk2 (Sobel and Lange 1996) for all 16 affected individuals and found a region inherited identical by decent (IBD). We evaluated the significance of the region in 176 unrelated nuclear families from the same geographical area as the large family. This region contained three SNPs. The transmission rate of the haplotype of interest was investigated in affected and healthy siblings using a gamete competition type test, were the association between transmission rate and disease status was assessed by $\chi^2$-statistics.

The mRNA expression levels of the only gene located in this region, quaking homolog, KH domain RNA binding (mouse) (QKI) was investigated with a real-time RT-PCR strategy in brain autopsies from 55 schizophrenic cases and 55 controls. Real-time RT-PCR expression data was normalized with respect to reference genes and systematic variation due to age, sex, brain-bank and post-mortem interval was removed. To test whether mRNA expression levels of QKI and its splice-variants differed between cases and controls and whether neuroleptic treatment affected the mRNA expression levels we used an ANCOVA type model.
Results
In the 16 affected individuals from the large pedigree we found a potential ancestral haplotype of 4.7 Mb. Approximately 0.5 Mb of this region, including three SNPs, was shared by eleven affected individuals. The haplotype found in this region in the large family was more frequently transmitted to affected siblings than to healthy siblings in 176 unrelated nuclear families.

Messenger RNA expression of two QKI splice-variants, in relation to the total QKI mRNA expression, were down regulated in schizophrenic cases (p-value 0.0002 and 0.0164 for QKI-7kb and QKI-7kbB, respectively). The mRNA expression relative to that of reference gene expression was clearly affected by the type of neuroleptics that the patients had received (p-values for QKI-5kb, QKI-6kb, QKI-7kb and QKI-7kbB correspond to 0.0014, 0.0009, <0.0001 and <0.0001, respectively). Patients treated with typical neuroleptics had approximately twice as high mRNA levels as patients treated with atypical neuroleptics as well as patients without neuroleptic treatment, for all four splice-variants.

Discussion
In this investigation we present evidence from three different sample sets, 16 schizophrenic individuals from a large pedigree, 176 unrelated nuclear families and brain autopsies from 55 cases and 55 controls, that propose association between schizophrenia and the QKI-gene located on chromosome 6q26-27. The function of QKI in humans is likely to be similar to its homolog in mouse that is involved in regulation of myelination related genes and neural development (Li et al. 2002; Wu et al. 2002). Both these functions have been suggested as potentially involved in the development of schizophrenia (Maier and Ron 1996; Bartzokis 2002). Lately, several myelin specific genes such as PLP, MAG and MBP have been reported to express down regulated mRNA levels in schizophrenic cases (Hakak et al. 2001; Tkachev et al. 2003; Aston et al. 2005). Studies in mouse suggest that the mouse quaking gene is of importance for splicing and expression of such myelin specific genes (Wu et al. 2002). Taken together we propose QKI as a possible candidate gene for the development of schizophrenia that should be further studied.
Paper IV – Human *QKI*, a Potential Regulator of mRNA Expression in Myelin Specific Genes Involved in Schizophrenia

We formulated three general hypotheses about the possible importance of human *QKI* for regulation of myelination related genes. The hypotheses were based on knowledge on the function of the mouse quaking gene, and recent mRNA expression studies, described in paper III in this thesis, that report down regulated myelination related genes in schizophrenia (Aberg et al. 2005). The three hypotheses can be summarized as follows; i) Due to a common regulator factor the mRNA expression levels of myelin specific genes should be correlated. ii) Human *QKI*-gene is a common regulatory gene that influences mRNA expression levels of myelin specific genes. iii) A specific *QKI* splice variant or the ratio between different *QKI* splice-variants are of importance for determination of mRNA levels of myelin specific genes and/or splice-variants of these genes. Consequently, we also investigate whether the disruption of *QKI* mRNA expression levels associated with schizophrenia and neuroleptic treatment, observed in paper III, can explain variations of mRNA expression differences in myelin specific genes observed in schizophrenic cases.

Methods

The mRNA expression levels of the myelin specific genes as well as the mRNA expression levels of the *QKI*-gene, previously described in paper III, were obtained with real-time RT-PCR in brain autopsies from 55 schizophrenic cases and 55 controls. Prior to all analyses we normalized the real-time RT-PCR expression data with respect to reference genes and removed systematic variation due to age, sex, brain-bank and post-mortem interval.

Correlation between three myelin specific genes, *MAG*, *MBP* and *PLP*, was explored with principal component analysis. To examine whether variation in the total *QKI* level and/or variation in *QKI* splice-variants, not related to the total *QKI* level, can predict the expression variation in myelin specific genes, we used a linear regression approach. Furthermore, to test whether disease status and/or neuroleptic treatment affected mRNA expression of myelin specific genes we used univariate ANOVA.

Results

The expression levels of *MAG*, *PLP* and *MBP* are highly correlated ($r>0.5$). The total expression level of *QKI* could explain 16% to 40% of the expression variation of individual myelin specific genes. In addition, the amount of specific splice-variants also modified the expression of myelin specific
genes. Messenger RNA expression of \textit{MAG} (1.7-fold, p-value 0.002) and \textit{PLP} (1.3-fold, p-value 0.021) were down regulated in schizophrenic cases compared to controls. In paper III, we reported that mRNA expressions of all \textit{QKI} splice-variants were clearly affected by neuroleptic treatment (Aberg et al. 2005). In this investigation we showed that neuroleptics also modify the mRNA expression of myelin specific genes on top of any disease effect. In general, typical neuroleptics tend to increase mRNA levels of myelin specific genes, whereas atypical neuroleptics tend to decrease the expression below the level of untreated patients.

\section*{Discussion}
Indirect evidence suggests that the function of human \textit{QKI} is related to myelination of human CNS by regulating mRNA expression of three major myelin specific genes, \textit{MAG}, \textit{PLP} and \textit{MBP}. We have reported that neuroleptic treatment affects mRNA expression of \textit{QKI} as well as of myelin specific genes. Interestingly, we have noted that if we account for differences in total \textit{QKI} mRNA levels, there is no longer any significant effect of drug treatment of myelin specific genes in schizophrenic patients. This suggests that the effect of neuroleptic treatment, observed in myelin specific genes, most likely is a down stream effect of the variation observed in \textit{QKI}.

\section*{Paper V – Genome Wide Investigation of an Isolated Schizophrenia Population Using Dense Maps of Microsatellites and SNPs in Combination}
In this study we have aimed to reduce the heterogeneity in the sample set by investigating closely related individuals from a geographical isolate. By performing affected-only analysis we can disregard the unknown penetrance level within this population. To receive maximal information from all investigated markers we have exclusively performed multipoint investigations.

\section*{Methods}
In this study we have investigated 35 individuals from a geographically isolated area. We have genotyped approximately 900 microsatellites and 10 000 SNP and performed a four step linkage genome scan. First, we performed non-parametric linkage (NPL) analysis with the microsatellites using Simwalk2 (Sobel and Lange 1996). Thereafter, we calculated NPL for the SNPs for the chromosomal regions that had received NPL-score above one in the initial step. For the regions with NPL-score above two in step one and/or in step two we calculated NPL-score with a combined marker map of microsa-
tellites and SNPs. For the loci that after combining the markers still received NPL-score above two we calculated multipoint LOD-score using Simwalk2.

Results

Microsatellite markers detected twenty-nine loci with NPL-score above one. Twenty-five of these loci were confirmed with SNP markers. Four loci, 1q23, 2q13, 5q21 and 6q16, received NPL-score above two. When combining microsatellites and SNPs in those four loci two regions remained NPL-score above two, the 2q13 locus (NPL-score 2.22, p-value 0.0061) and the 5q21 locus (NPL-score 2.15, p-value 0.0072). Multipoint parametric linkage investigations revealed a maximum LOD-score of 2.2 for chromosome 2q13 and a LOD-score of 1.1 for chromosome 5q21.

Discussion

In this study we use a stepwise statistical strategy, with an increased cutoff value in each step, to perform a genome wide investigation. Non-parametric linkage results received by microsatellites are confirmed by SNPs before the marker map density is increased and additional non-parametric investigations followed by parametric investigations are performed. With this strategy we hope to minimize the number of false positive regions due to incorrect genotyping of individual markers, unique allele frequencies and possibly also regions detected due to linkage disequilibrium (LD) between markers that is present in only one of the two data sets. In addition, this stepwise strategy dramatically decreases the time for computational calculation.

In this investigation we found two potential susceptibility loci for schizophrenia, chromosome 2q13 and 5q21. Both these loci have previously been reported as potential regions to harbor schizophrenia genes. (Straub et al. 1997; Mowry et al. 2000; Paunio et al. 2001; DeLisi et al. 2002; Straub et al. 2002b; Lindholm et al. 2004). Interestingly, these two loci (2q13 and 5q21, respectively) are located within the most significant chromosomal region and right next to the second most significant chromosomal region that, according to a large meta-analysis, are most likely to harbor schizophrenia susceptibility loci common for multiple populations (Lewis et al. 2003).
Future Perspectives and Conclusions

Evaluation of QKI as a candidate gene for schizophrenia

In paper III, we reported evidence for association between schizophrenia and a specific haplotype, located within QKI on chromosome 6q, in two unrelated sample sets from northern Sweden. In addition, we observed mRNA expression differences of the QKI-gene in brain autopsies from schizophrenic cases and controls. To be able to understand the significance of this gene in the development of schizophrenia further research has to be performed.

Sequencing the complete gene, including introns and exons as well as promoter region and 3’ un-translated regions, in a large number of cases and controls would reveal polymorphisms that possibly are associated with the disease. The polymorphisms may be associated either by being direct functional mutations or by being in strong linkage disequilibrium with another functional mutation.

In paper IV, we found that mRNA expression of the QKI-gene to a great extent can predict mRNA expression of other myelin specific genes and our hypothesis is that the QKI-gene is a common regulator of myelin specific genes in humans. To prove or disprove this hypothesis, functional studies of the QKI-gene in human cell lines could be performed. One of the most interesting studies could include the use of small interfering RNA (siRNA) to evaluate the regulation of expression of myelin genes (McManus and Sharp 2002).

Similar to the investigations performed in paper I, it would be of great interest to perform association investigations between polymorphisms in the gene and mRNA expression of QKI.

Novel genome-wide investigations

In paper V, we performed a genome-wide linkage investigation with very dense marker maps of microsatellite markers and/or SNPs. Using non-parametric statistics (NPL-score > 2), we found two chromosomal regions that we further evaluated with parametric statistics. In the future, it would be of interest to calculate parametric statistics for regions that received slightly lower NPL-score. In spite of the fact that we have performed a multipoint
genome-wide investigation based on a high-density marker map, genotyped in related individuals, we have the possibility to further extend this investigation to also include a more novel genome-wide haplotype investigation. No clear statistical strategy is so far developed for this purpose. A possible approach would be to use NPL-statistics to select regions of interest for haplotype construction and measure linkage disequilibrium between the haplotypes in each chromosomal region and calculate their association to schizophrenia. Instead of considering the two most likely haplotypes as true haplotypes for each individual it would be of interest to use the likelihood for all possible haplotypes for each individual. With this strategy it would be possible to detect interaction between multiple loci that are involved in the development of the disease.

**General conclusion**

Many of previously proposed schizophrenia candidate genes are well in line with the established neurotransmitter hypothesis and/or the more novel myelin hypothesis. The investigations presented here have mainly contributed to the schizophrenia research field by identifying a new candidate gene named \emph{QKI}, located on chromosome 6q26-q27. We have shown indirect evidence that \emph{QKI} regulates other myelin specific genes. Myelin plays an important role in the development of normal brains and disruption of this function might lead to symptoms such as schizophrenia.

Furthermore, we have located two additional schizophrenia susceptibility loci, 2q13 and 5q21. Both these loci have previously been highlighted as potential schizophrenia loci in several other investigations, including a Finnish study. This suggests common susceptibility loci for schizophrenia in Nordic populations. These findings support the theory that schizophrenia is a complex disease caused by heterogeneous genetics in combination with environmental involvement.

I believe that this work has contributed with substantial information that will aid the understanding of the schizophrenia etiology. However, identification of all factors involved in the development of schizophrenia, genetic factors as well as environmental factors, is far from completed. Therefore, it is of great importance to continue with advanced and innovative research, that includes collaboration between specialists from different disciplines of natural science, worldwide.
Gener som orsakar schizofreni

Schizofreni – en allvarlig sjukdom som påverkar hjärnfunktionen

Schizofreni är en av våra vanligaste psykiatriska sjukdomar. Sjukdomen drabbar i genomsnitt 1 % av världens befolkning men högre prevalens har uppmätt i flera mindre, förhållandevis isolerade, populationer t.ex. i Slovenia, västra Irland och i norra Sverige. Sjukdomen yttrar sig genom så kallade positiva symtom dvs vanföreställningar, hallucinationer och bisarrt beteende, men också genom negativa symtom så som frånvaro av känsloyttningar, brist på motivation och social tillbakadragenhet. Symtomen kan variera mycket mellan olika patienter vilket leder till att sjukdomen är svår att diagnostisera. Idag tillämpas huvudsakligen två standardiserade diagnostiseringar metoder, DSMIV (Diagnostic and Statistical Manual of Mental Disorder, forth edition) och ICD10 (International Classification of Diseases, tenth edition). De båda diagnostiseringarskriterierna baseras på att särskilda symtom ska ha förekommit under en viss tid. DSMIV kräver att symtom ska ha observerats i minst sex månader medan ICD10 endast kräver en månad.

Ingen av dagens behandlingsmetoder kan fullständigt bota schizofreni. Det finns dock finns två typer av neuroleptika, atypiska och typiska, som delvis kan lindra symtomen. Typiska neuroleptika är främst effektiva mot positiva symtom, medan atypiska läkemedel även har effekt mot negativa symtom.

Framsteg inom schizofreni forskningen

Forskningssstudier har påvisat makroskopiska och histologiska fynd i form av minskad hjärnvolym, förstorade ventriklar, förtjockad cortex, asymmetri av hjärnan samt felaktigt placerade och ihopklumpade neuroner hos vissa schizofrena patienter. Det finns dock inget fynd som förekommer hos majoriteten av alla patienter.

Under senare år har kopplings- och associationsstudier identifierat ett stort antal potentiella schizofrenikänsliga regioner spridda över det humana genomet. De starkaste bevisen för koppling och/eller association finns främst
på kromosom 1q21-q22, 1q42, 5q21-q33, 6p24-p22, 6q16-q26, 8p22-p21, 10p15-p11, 13q22-q34 och på kromosom 22q11-q12. I vissa av dessa regioner har möjliga kandidatgener föreslagits. Dessa gener är framför allt RGS4 (regulator of G-protein signaling-4) på kromosom 1q, DTNBP1 (dysbindin) på kromosom 6p, NRG1 (neuregulin 1) på kromosom 8p, DAOA (D-amino-acid oxidase activator), även kallad G72, på kromosome 13q och COMT (catechol-O-metyltransferase) på kromosom 22q. På kromosom 6q har två potentiella kandidatgener föreslagits TRAR4 (trace amine receptor 4) och QKI (quaking homolog, KH domain RNA binding (mouse)).

Nyutvecklad mikromatristeknik i kombination med realtids-PCR har påvisat förändrat mRNA-uttryck i hjärnautopsier från schizofrena patienter. Förändringarna har främst antytt att oligodendrocytfunktionen och synapsfunktionen är involverade i utvecklandet av schizofreni.

Under många år har schizofreniforskare diskuterat huruvida flera av hjärnans neurotransmitter, så som dopamin, serotonin, glutamat och GABA, är involverade i sjukdomens utveckling. Neurotransmitter är av stor betydelse för att synapserna ska fungera tillfredsställande. Flera av de kandidatgener som har föreslagits genom kopplings- och associationsstudier (t.ex. DAOA, DTNBP1 och NRG1) har en klar inverkan på neurotransmittermekanismen.

Hos en frisk människa sker en viktig del av hjärnutvecklingen under adolescens och postadolescens vilket sammanfaller med den tidsperiod då de flesta schizofrenipatienter får sin diagnos. Under denna utveckling ökas normalt mängden myelin i hjärnan. Minst tre av de kandidatgener som föreslagits genom kopplings- och associationsstudier (NRG1, RGS4 och QKI) är inblandade i myelinbildning. QKI är dessutom en av de gener som har uppvisat förändrad mRNA-nivå i hjärnan på schizofrena patienter. Nyligen har det föreslagits att även neurotransmitterer påverkar bildandet av myelin.

Forskningsmål

För att kunna utveckla nya läkemedel för schizofrenipatienter är det viktigt att förstå den fysiska orsaken till sjukdomens uppkomst. Huvudmålet för denna avhandling är därför att identifiera kandidatgener och genkaskader som är involverade i utvecklandet av schizofreni.

Artikel I  Att undersöka om förändringar av mRNA-uttryck påverkas av neuroleptika och/eller studerade DNA-variationer i promotorregionen av HTR2C (serotonin receptor 2C).

Artikel II  Att identifiera kandidat-haplotyper i 19 besläktade schizofrenapatienter samt att undersöka om dessa haplotyper är lika vanliga bland obesläktade friska kontrollindivider.

Artikel III  Att kartlägga en schizofreniregion på kromosom 6q, i en stor släkt, för att undersöka om samma genetiska mönster kan observeras i 176 orelaterade kärnfamiljer, samt att undersöka huruvida QKI-genen, som finns
i denna region, har förändrat mRNA-uttryck i hjärnautopsier från schizofrena patienter.

Artikel IV Att undersöka om det finns en korrelation mellan mRNA-uttrycket i tre myelinspecifika gener samt att undersöka om denna korrelation är påverkad av QKI-genen i hjärnautopsier från schizofrena patienter och kontrollindivider.

Artikel V Att söka efter nya schizofrenikänsliga regioner i en isolerad population, med hjälp av tätt placerade mikrosatellitter och SNPs.

Artikel I – Serotonin Receptor 2C (HTR2C) och Schizofreni: Undersökning av Möjlig Medicineringspåverkan och Genetisk Päverkan av Uttrycksnivåer

I denna studie undersökte vi huruvida neuroleptika och/eller tre studerade SNPs påverkar mRNA uttrycket i hjärnautopsier från 55 schizofrena patienter i förhållande till 55 friska kontrollindivider.


Artikel II – Rekonstruktion av Nedärvda Haplotyper i ett Tolv Generationers Schizofrenisläktträdd

I denna studie undersökte vi 19 besläktade schizofrena patienter för att finna haplotyper som sannolikt har nedärvts från en gemensam stamfader. Haplotyper som erhållit ett ”säkerhetsvärde” på minst tre undersöktes i 43 sjuka individer och i 46 kontrollindivider. (Säkerhetsvärdet baseras på antalet individer som delar respektive haplotyp och huruvida fasen för varje enskild individs haplotyp är känd.) Förutom den region på kromosom 6q25, som i tidigare studier av detta familjematerial rapporterats som nedärvd från en gemensam stamfader, fann vi ytterligare elva regioner. I fyra av dessa regioner (8p, 9p, 13q och 17q) observerade vi en signifikant skillnad i haployp-frekvenser mellan de 43 sjuka och de 46 friska individerna. Två av dessa regioner (8p och 13q) har i andra studier rapporterats som potentiella kandidatregioner för schizofreni.
Artikel III – Humana \textit{QKI}, en Ny Kandidatgen för Schizofreni Involverad i Myelinering

I denna studie undersökte vi den schizofreniregion på kromosom 6q25, som nämns i artikel II, i 16 besläktade schizofrena individer. Vi undersökte om den vanligaste haplotypen bland de 16 schizofrena även är mer frekvent hos schizofrena patienter än hos deras friska syskon i 176 obesläktade kärnfamiljer. Vi fann en haplotyp som delas av 11 av de 16 besläktade schizofrena. Denna haplotyp är också signifikant mer frekvent bland schizofrena i de 176 familjerna än bland deras friska syskon. I detta område finns endast en gen, \textit{QKI}.

Vi undersökte mRNA-nivåer av \textit{QKI}-genen i hjärnautopsier från 55 schizofrena patienter och från 55 kontrollindivider. Hos schizofrena patienter observerade vi att mRNA-uttrycket av två splitsningsvarianter av \textit{QKI} var signifikant nedreglerad i förhållande till det totala \textit{QKI}-uttrycket. Vi såg också att typisk neuroleptika signifikant ökar uttrycket av alla undersökta \textit{QKI} splitsningsvarianter såväl som av det totala \textit{QKI}-uttrycket. \textit{QKI}-genen är hos mus känt för att vara involverad i myelinering genom att reglera mRNA av viktiga myelinkomponenter.

Artikel IV – Humana \textit{QKI}, en Potentiell Regulator av mRNA Uttryck i Myelin Specifika Gener Involverade i Schizofreni

I denna studie undersökte vi huruvida mRNA-uttryck av tre myelinspecifika gener (\textit{PLP}, \textit{MAG} och \textit{MBP}) är korrelerade samt om deras korrelation och uttryck påverkas av \textit{QKI}-mRNAuttryck. Vi såg att mRNA-uttryck av de myelinspecifika genera är väl korrelerat. Vi såg också att det totala \textit{QKI}-uttrycket kan förklara en stor del av variationen för varje individuell myelin specifik gen. Dessutom kan mängden av vissa splitsningsvarianter ytterligare förklara en del av variationen.

Vi har tidigare visat att neuroleptika påverkar uttrycket av \textit{QKI} (artikel III). I denna studie visar vi att även uttrycket av myelinspecifika gener påverkas av neuroleptika och att detta sannolikt är en följd av förändringarna observerade i \textit{QKI}.

Artikel V – Genom-omfattande Undersökning av en Isolerad Schizofrenipopulation Genom att Använda Mikrosatellitter och SNPs i Kombination

I denna studie sökte vi efter nya kromosomområden som kan vara kopplade till utvecklandet av schizofreni. Vi undersökte 900 mikrosatellitter och 10000 SNPs i 35 individer från ett geografiskt isolerat område och tillämpade en statistisk strategi som är uppdelad i fyra nivåer. För att en kromosomregion
ska undersökas på nästföljande nivå måste särskilda statistiska krav uppfyl-
las. Denna strategi inleds med modellfria kopplingsanalyser och avslutas
med modellbaserade kopplingsanalyser. Såväl de modellfria som de modell-
baserade analyserna är beräknade med flerpunktsstatistik.

I denna studie fann vi två kromosomregioner som är potentiellt kopplade
till uppkomsten av schizofreni, 2q13 och 5q21. Båda dessa regioner ligger i
anslutning till två kromosomområden som av en stor internationell metaana-
lys anses vara mest sannolika att innehålla schizofrenikänslig arvsmassa.

Slutsatser
Sammanfattningsvis så har arbetet bakom denna avhandling främst bidragit
till schizofreniforskningen genom att identifiera en ny kandidatgen, *quaking*
*homolog, KH domain RNA binding (mouse)*, allmänt kallad *QKI*. Denna gen
är placerad på kromosom 6q26-q27. Vi har lagt fram indirekta bevis för att
*QKI* reglerar andra myelingener. Myelin spelar en viktig roll i hjärnans nor-
mala utveckling. Rubbningar av denna funktion kan sannolikt leda till sym-
tom såsom schizofreni.

Genom att även identifiera andra kromosomregioner som potentiellt schi-
zofrenikänsliga områden har vi visat att sjukdomen är högst komplex. Mest
sannolikt är schizofreni en effekt av flera förändringar i hjärnan, till följd av
arvsmassa och/eller miljöpåverkan. För att kunna förstå sjukdomens etiologi
och i förlängningen kunna utveckla behandlingsmetoder för de drabbade
patienterna krävs att lovande forskningsresultat följs upp. Det är därför av
yttersta betydelse att även fortsättningsvis bedriva schizofreniforskning på
högsta nivå såväl nationellt som internationellt.
Acknowledgements

First, I would like to thank all former and present colleagues in the group of Behavioral Genetics at Uppsala University. Especially: my supervisor Elena Jazin for shared joy in good times, for all your encouraging support and the ability "to keep me calm" ;-) when things do not go as easy as I would wish; Eva Lindholm for introducing me to the mystery of schizophrenia; Peter Saetre for never ending enthusiasm when discussing the “schizo-data”, no matter if we are at the running track or at the coffee machine, and for never saying no when I ask you to recode and rerun my data for another “last time”; Anja Castensson for sharing the passion for psychiatric genetics and all the concern it brings to us; Maria Norberg for skilful and patient work with my favorite gene; Julia Lindberg for all the laughs and lively discussions about life of science and science of life; Lina Emilsson for all fun times at work and elsewhere, including the tyrannosaurus incident, and for all “egg on talk” at tuff times. Thank you all for creating such a stimulating atmosphere in our office!

Second, I would like to thank all my colleges at the section for Evolutionary Biology and at the section for Medical Genetics at Uppsala University and elsewhere for creating such a friendly and inspiring atmosphere to work in. In particular I would like to thank; Ulf Pettersson for your valuable opinions and for always believing in my investigations; Lennart Wetterberg for all your support and your encouraging mails with comments about recent research proceedings; Hans Ellegren for accepting me as a PhD student; Siv Strömberg and Lars Pilström for introducing me to the exciting field of genetics and for teaching me all the basic genetic concepts; all co-authors, in previous investigations Shane McCarty, Björn Andersson, Birgit Ekholm and Rolf Adolfsson for successful collaboration; Åsa Johansson and Malin Engelmärk for educational discussions about statistical genetics and other essential things; Inger Jonasson for skilful microsatellite genotyping; Johan Lidén, Marika Rönnholm and David Brodin for help with SNP genotyping; Sonja Färeby for help with genealogy; Håkan Svensson, Micke Brandström and Viktor Persson for computer support; Rose-Marie Löfberg, Carolina Wallström-Pan, Nisse and Henrik for fixing things that no one else know how to do.

Finally, I would like to thank all my friends and family for always being there for me. Especially I would like to express my gratitude to: Pelle P for always being loving, caring and loyal and especially for putting up with me
also when I am hungry; Jenny, Åsa, Björn, Josef, Lotta and Pelle L for good friendship and an infinite number of fun times; Jenny, Anna, Malin and Linda for long, long, long friendship that is not disrupted by thousands of miles or months without an email; Richard M for always answering your phone, three o’clock in the morning is not an exception, to give me your backing and encouragement; the Daddario family Alex, Matt, Catharine, Christina and Richard for always caring about me and for teaching me the English language (I believe it was essential for this book.); my grandparents Saimie and Rune for your heartening support; my grandmother Margaretha for your cheerful words of wisdom; and last but not the least my wonderful family: my mom Chatarina and my dad Peter, my brother Pierre and my sister Rose-Marie for always believing in me, encouraging me in all businesses, enriching me with all your genuineness and showing what really matters in life.

The Beijer Foundation, Torsten and Ragnar Söderberg’s Foundation, the Karolinska Institute, Stockholm and the School of Biology, Uppsala University financially funded this work.
References


Strachan T, Read AP (1999) Human Molecular Genetics 2

Straub RE, MacLean CJ, Ma Y, Webb BT, Myakishev MV, Harris-Kerr C, Wormley B, Sadek H, Kadambi B, O'Neil FA, Walsh D, Kendler KS (2002b) Genome-wide scans of three independent sets of 90 Irish multiplex schizophrenia families and follow-up of selected regions in all families provides evidence for multiple susceptibility genes. Mol Psychiatry 7:542-559


Sundberg R, Castensson A, Jazin E (2005) Statistical modelling in case-control real-time RT-PCR assays, for identification of differentially expressed genes in schizophrenia. Biostatistics


Weinberger DR (1987) Implications of normal brain development for the pathogenesis of schizophrenia. Arch Gen Psychiatry 44:660-669


Wetterberg L (1997) Psykiatri i fickformat. In, Stockholm, pp 54-71


Wing JK, Cooper JE, Sartorius N (1974) Measurement and classification of psychiatric symptoms


Zhao JH, Curtis D, Sham PC (2000) Model-free analysis and permutation tests for allelic associations. Hum Hered 50:133-139
Acta Universitatis Upsaliensis

Digital Comprehensive Summaries of Uppsala Dissertations from the Faculty of Science and Technology 78

Editor: The Dean of the Faculty of Science and Technology

A doctoral dissertation from the Faculty of Science and Technology, Uppsala University, is usually a summary of a number of papers. A few copies of the complete dissertation are kept at major Swedish research libraries, while the summary alone is distributed internationally through the series Digital Comprehensive Summaries of Uppsala Dissertations from the Faculty of Science and Technology. (Prior to January, 2005, the series was published under the title “Comprehensive Summaries of Uppsala Dissertations from the Faculty of Science and Technology”.)

Distribution: publications.uu.se
urn:nbn:se:uu:diva-5894